Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

On evolutionary Γ-convergence for gradient systems

In memory of Eduard, Waldemar, and Elli Mielke

Alexander Mielke [1,2]

submitted: January 6, 2014

[1] Weierstraß-Institut
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: alexander.mielke@wias-berlin.de

[2] Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof
Germany

No. 1915
Berlin 2014

2010 *Mathematics Subject Classification.* 35B27 35K55 47H20 49S05 49J40 49J45.

Key words and phrases. Variational evolution, energy functional, dissipation potential, dissipation distance, gradient flows, Gamma convergence, Mosco convergence, well-prepared initial conditions, rate-independent systems, abstract chain-rule, energy-dissipation balance, integrated evolutionary variational estimate, energetic solutions.

The research was partially supported by the DFG within the SFB 910 (subproject A5) and by the ERC under AdG 267802 AnaMultiScale.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In these notes we discuss general approaches for rigorously deriving limits of generalized gradient flows. Our point of view is that a generalized gradient system is defined in terms of two functionals, namely the energy functional $\mathcal{E}_\varepsilon$ and the dissipation potential $\mathcal{R}_\varepsilon$ or the associated dissipation distance. We assume that the functionals depend on a small parameter and that the associated gradient systems have solutions u_ε. We investigate the question under which conditions the limits u of (subsequences of) the solutions u_ε are solutions of the gradient system generated by the Γ-limits $\mathcal{E}_0$ and $\mathcal{R}_0$. Here the choice of the right topology will be crucial as well as additional structural conditions.

We cover classical gradient systems, where $\mathcal{R}_\varepsilon$ is quadratic, and rate-independent systems as well as the passage from classical gradient to rate-independent systems. Various examples, such as periodic homogenization, are used to illustrate the abstract concepts and results.

1 Introduction

This work is devoted to the study of evolutionary systems of the form $\dot{u} = \boldsymbol{V}_\varepsilon(u)$, $u(0) = u_\varepsilon^0$ depending on a small parameter ε, which may arise as a microscopic length scale. The aim is to find effective equations such that any suitable limit u of solutions u_ε satisfies the effective model $\dot{u} = \boldsymbol{V}_0(u)$. We are interested where the dependence on the parameter is singular, i.e. the dependence $\varepsilon \to \boldsymbol{V}_\varepsilon(v)$ is not continuous, but we still hope that the solutions u_ε have a limit. Thus, our results on evolutionary Γ-convergence can be seen as *singular limits* in the sense of [FeN09]. As in the case of static Γ-convergence, the aim is to derive effective limit equations that still capture the main effects of the original model with $0 < \varepsilon \ll 1$, but are simple in the sense that they do not longer include the small scales, and hence are easier to analyze or to solve numerically.

We emphasize that the justification of such multiscale limits corresponds to showing that it is possible to interchange the time evolution $u_\varepsilon(0) \rightsquigarrow u_\varepsilon(t) = \boldsymbol{S}_\varepsilon(t, u_\varepsilon(0))$ with the limit in $\varepsilon \to 0$. More precisely, we have to show that the diagram in Figure 1.1 commutes, i.e. we have to prove that

$$\lim_{\varepsilon\to 0} \mathbb{M}_\varepsilon \circ \boldsymbol{S}_\varepsilon(t, \cdot) = \boldsymbol{S}_0(t, \lim_{\varepsilon\to 0} \mathbb{M}_\varepsilon(\cdot)), \tag{1.1}$$

where $\mathbb{M}_\varepsilon$ denotes a suitable upscaling operator, which in most of our cases is simply given by id. In the latter case we say that the evolution equations $\dot{u} = \boldsymbol{V}_\varepsilon(u)$ *semigroup converges* to $\dot{u} = \boldsymbol{V}_0(u)$. In principle the justification of this limiting process can be done directly on the (partial) differential equation $\dot{u}_\varepsilon = \boldsymbol{V}_\varepsilon(u_\varepsilon)$.

However, in this work we will concentrate on special evolutionary systems, namely *generalized gradient system* $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$. This triple consist of a state space $\boldsymbol{X}$ containing the states u, which is usually a Banach space $\boldsymbol{X}$, a closed subset, or even a metric space. The driving functional $\mathcal{E} : \boldsymbol{X} \to \mathbb{R} := \mathbb{R} \cup \{\infty\}$ is usually a (free) energy or the negative on an entropy. The function $\mathcal{R} : \boldsymbol{X} \times \boldsymbol{X} \to [0, \infty]$ is called dissipation potential, which means that $\mathcal{R}(u, \cdot)$ is lower semicontinuous (lsc), proper, and convex and that $\mathcal{R}(u, 0) = 0$. Here $\partial_{\dot{u}}\mathcal{R}(u, \dot{u})$ gives the dissipative force, while $-\mathrm{D}\mathcal{E}(u)$ is the potential restoring force.

	microscopic system $\dot u_\varepsilon = \boldsymbol{V}_\varepsilon(u_\varepsilon)$	$\varepsilon \to 0$	macroscopic system $\dot u = \boldsymbol{V}_0(u)$
initial state	u_ε^0	$\xrightarrow[\mathbb{M}_\varepsilon]{\text{upscaling}}$	u^0
time evolution	$\downarrow$		$\downarrow$
time $t > 0$	$u_\varepsilon(t) = \boldsymbol{S}_\varepsilon(t, u_\varepsilon^0)$	$\xrightarrow[\mathbb{M}_\varepsilon]{\text{upscaling}}$	$u(t) = \boldsymbol{S}_0(t, u^0)$

Figure 1.1: Illustration of upscaling and semigroup convergence, if $\mathbb{M}_\varepsilon = \mathrm{id}$

Using the dual dissipation potential $\mathcal{R}^*(u, \xi)$ (cf. Section 2.3) we can write the gradient flow induced by $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$ in three equivalent forms

$$
\begin{aligned}
&0 = \partial_{\dot u}\mathcal{R}(u, \dot u) + \mathrm{D}\mathcal{E}(u) && \text{force balance, Biot's equation}\\
&\dot u = \partial_\xi \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) && \text{rate equation, Onsager system}\\
&\mathcal{R}(u, \dot u) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) = -\langle \mathrm{D}\mathcal{E}(u), \dot u\rangle && \text{power balance}
\end{aligned}
$$

If $\mathcal{R}(u, \cdot)$ is quadratic, i.e. $\mathcal{R}(u, v) = \frac{1}{2}\langle \mathbb{G}(u)v, v\rangle$ and $\mathcal{R}^*(u, \xi) = \frac{1}{2}\langle \xi, \mathbb{K}(u)\xi\rangle$ with $\mathbb{K} = \mathbb{G}^{-1}$, we call $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$ a (classical) gradient system, and the evolution equations read

$$
\dot u = \boldsymbol{V}(u) = -\mathbb{K}(u)\mathrm{D}\mathcal{E}(u) \quad \Longleftrightarrow \quad 0 = \mathbb{G}(u)\dot u + \mathrm{D}\mathcal{E}(u). \tag{1.2}
$$

The focus of this work is to derive sufficient conditions for semigroup convergence for the special class of evolutionary systems that are induced by generalized gradient systems. The emphasis is on methods that fully exploit the gradient structure. In particular, it is desirable to derive sufficient conditions on the convergence of the pair $(\mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ to effective functionals $(\mathcal{E}_0, \mathcal{R}_0)$ such that $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ semigroup converges to $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$ in the sense of (1.1) with $\mathbb{M}_\varepsilon = \mathrm{id}$.

We also emphasize that there is a rich literature on semigroup convergence for general evolutionary systems $\dot u + \boldsymbol{A}_\varepsilon(u) = 0$ that often use that $\boldsymbol{A}_\varepsilon$ is defined via a maximal monotone operator on a *fixed* Banach space $\boldsymbol{X}$ (see e.g. [Bré73, Thm. 3.18], [Att84, Thm. 3.74], and the discussion in [Ste08]), which includes classical gradient systems $\mathbb{G}\dot u = -\mathrm{D}\mathcal{E}_\varepsilon(u)$ if the dissipation potential $\mathcal{R}_\varepsilon(u, v) = \frac{1}{2}\langle \mathbb{G}_\varepsilon(u)v, v\rangle$ is independent of u and ε. Note that writing the system in the form $\dot u + \boldsymbol{A}_\varepsilon(u) = 0$ with $\boldsymbol{A}_\varepsilon(u) = \mathbb{K}_\varepsilon(u)\mathrm{D}\mathcal{E}_\varepsilon(u)$ the monotonicity for all $\varepsilon \in [0, 1]$ is lost if working with a fixed Hilbert-space norm.

For gradient systems it is also important to keep track of the energies $\mathcal{E}_\varepsilon(u_\varepsilon(t))$ along the solutions and to control their convergence in the limit $\varepsilon \to 0$. In analogy to the notion of static Γ-convergence (cf. [Dal93, Bra02] or Section 2.2) we then follow the naming suggested in [SaS04] and speak of *evolutionary* Γ*-convergence for gradient system*.

More precisely, we introduce two notions of evolutionary Γ-convergence depending of whether we need control of the energy at the initial time or not. We define the statements

$$
\mathfrak{S}(t) = \text{“}\big(\, u_\varepsilon(t) \to u(t)\,\big)\text{”} \qquad \text{and} \qquad \mathfrak{E}(t) = \text{“}\big(\, \mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u(t))\,\big)\text{”}.
$$

Then, semigroup convergence simply means “$\mathfrak{S}(0) \implies \forall\, t > 0 :\ \mathfrak{S}(t)$”. For gradient

systems we will use the following two definitions for evolutionary Γ-convergence:

$$\begin{aligned}&\textit{E-convergence}, \quad \text{written } (\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon) \xrightarrow{\mathsf{E}} (\boldsymbol{X},\mathcal{E}_0,\mathcal{R}_0):\\ &\qquad \mathfrak{S}(0) \quad\Longrightarrow\quad \forall\, t>0:\ \mathfrak{S}(t) \text{ and } \mathfrak{E}(t);\end{aligned}$$

$$\begin{aligned}&\textit{well-prepared E-convergence}, \quad \text{written } (\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon) \xrightarrow{\mathsf{pE}} (\boldsymbol{X},\mathcal{E}_0,\mathcal{R}_0):\\ &\qquad \mathfrak{S}(0) \text{ and } \mathfrak{E}(0) \quad\Longrightarrow\quad \forall\, t>0:\ \mathfrak{S}(t) \text{ and } \mathfrak{E}(t).\end{aligned}$$

Here the letter "E" stands for both 'E'volutionary as well as 'E'nergy convergence. In the second definition the assumptions $\mathfrak{S}(0)$ and $\mathfrak{E}(0)$ at $t=0$ are called *well-'P'reparedness of the initial conditions*. Clearly, well-prepared E-convergence is implied by the stronger *E*-convergence, since the latter has a weaker assumptions at $t=0$, whereas the conclusions for $t>0$ are the same. We also use the notations $\xrightarrow{\mathsf{E}}$ and $\xrightarrow{\mathsf{pE}}$ if the convergences in $\mathfrak{S}(0)$ and $\mathfrak{S}(t)$ is replaced by the weak convergence in $\boldsymbol{X}$.

In Section 2 we give preparatory work including the basic modeling ideas of gradient systems (Section 2.1), basic facts about and a few examples for weak and strong Γ-convergence, Mosco and continuous convergence including the relations with the Legendre transform (Sections 2.2 and 2.3), and finally some examples of gradient systems that are used to illustrate the abstract theory in this work or that highlight the development of the general field (Section 2.5 to 2.8).

The main Sections 3 to 5 are devoted to the question, which types of convergences $\mathcal{E}_\varepsilon \rightsquigarrow \mathcal{E}_0$ and $\mathcal{R}_\varepsilon \rightsquigarrow \mathcal{R}_0$ are sufficient to guarantee that the gradient system $(\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$ evolutionary Γ-converges to the limit system $(\boldsymbol{X},\mathcal{E}_0,\mathcal{R}_0)$. In Section 3 we will discuss convergence results based on the *energy-dissipation balance* (EDB), also called *De Giorgi's* $(\mathcal{R},\mathcal{R}^*)$ *formulation*:

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon(t),\dot u_\varepsilon(t)) + \mathcal{R}_\varepsilon^*\big(u_\varepsilon(t),-\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t))\big)\,\mathrm{d}t \;=\; \mathcal{E}_\varepsilon(u_\varepsilon(0)). \tag{1.3}$$

This formulation was the starting point of the fundamental paper [SaS04], see also [Ser11], where the crucial conditions for well-prepared E-convergence are given by suitable liminf estimates for $\mathcal{R}_\varepsilon$ and $\mathcal{R}_\varepsilon^*(\cdot,-\mathrm{D}\mathcal{E}_\varepsilon(\cdot))$, respectively. We formulate a corresponding result on evolutionary Γ-convergence in Theorem 3.6 and present the more restrictive Theorem 3.3, which is based on Mosco convergence and was developed in [Ste08, Thm. 7.2]:

$$\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0 \ \text{ and } \ \Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0 \quad\Longrightarrow\quad (\boldsymbol{X},\mathcal{E}_\varepsilon,\Psi_\varepsilon) \xrightarrow{\mathsf{pE}} (\boldsymbol{X},\mathcal{E}_0,\Psi_0).$$

We continue to use the notation Ψ_ε for the special case of dissipation potentials $\mathcal{R}_\varepsilon$ that do not depend on the state variable u, i.e. $\mathcal{R}_\varepsilon(u,v)=\Psi_\varepsilon(v)$. Such $\mathcal{R}_\varepsilon$ are also called translation invariant.

The result in [Ste08] is based on a version of the Brézis-Ekeland principle that strongly uses that the translation invariance and the convexity of the energies $\mathcal{E}(t,\cdot)$. We present the approach developed in [MRS13b] which is based on the EDB (1.3) and works for general dissipation potentials $\mathcal{R}_\varepsilon$.

These abstract convergence results are applied to various examples in Section 3.5, in particular, to periodic homogenization of parabolic equations, to Tartar's ODE example

$\dot{u}(t,x) = -a(x/\varepsilon)u(t,x)$, and to the passage from a wiggly-energy system with small viscous dissipation to a rate-independent hysteresis model.

In Section 4 we show that under certain conditions the differential form of gradient systems is equivalent to an *evolutionary variational inequality* (EVI) that is formulated solely in terms of u, $\mathcal{E}$, and $\mathcal{R}$, and does not contain any derivatives, i.e. $\dot{u}$ and $\partial\mathcal{E}$ do not occur. In the case of convex energies $\mathcal{E}_\varepsilon$ and a translation invariant, quadratic dissipation potential $\mathcal{R}_\varepsilon(u,v) = \Psi_\varepsilon(v) := \frac{1}{2}\langle \mathbb{G}_\varepsilon v, v\rangle$ it takes the simple form of the Integrated Evolutionary Variational Estimate $(\mathrm{IEVE})_{\lambda=0}$:

$$\forall\, 0 \le s < t \le T,\ w \in \boldsymbol{X}: \quad \Psi_\varepsilon(u(t)-w) - \Psi_\varepsilon(u(s)-w) \le (t-s)\big(\mathcal{E}_\varepsilon(w)-\mathcal{E}_\varepsilon(u(t))\big). \tag{1.4}$$

This formulation is a variant of Bénilan's integral solutions (cf. in [Bén72]) for the Banach-space setting $\mathcal{R}(u,v) = \frac{1}{2}\|v\|^2_{\boldsymbol{X}}$, while the setting of metric spaces $(\mathcal{Q},\mathcal{D})$ is discussed extensively in [AGS05, Sav11], see Section 4.3 for the general form of the IEVE for geodesically λ-convex gradient systems on a geodesic space $(\mathcal{Q},\mathcal{D})$. Since only the functionals (but not their derivatives) appear, the IEVE formulation is particularly useful for passing to the limit $\varepsilon \to 0$. However, it again relies on a uniform convexity condition and is restricted to the case of classical, but possibly metric gradient systems. A simplified variant of the general result on E-convergence from [Sav11] is given in Theorem 4.5:

$$\mathcal{D}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{D}_0 \text{ and } \mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0 \quad \Longrightarrow \quad (\mathcal{Q},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon) \xrightarrow{\mathsf{E}} (\mathcal{Q},\mathcal{E}_0,\mathcal{D}_0),$$

where $\xrightarrow{\mathsf{C}}$ means continuous convergence, see (2.7). We emphasize that this E-convergence does not need well-preparedness of the initial conditions.

The final Section 5 is devoted to rate-independent systems $(\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$, which are a special case of generalized gradient flows where the dissipation potential is positively homogenous of degree 1 in the rate, i.e. $\forall\,\gamma>0,\ u,v\in\boldsymbol{X}: \quad \mathcal{R}_\varepsilon(u,\gamma v) = \gamma\mathcal{R}_\varepsilon(u,v)$. As in the case of classical gradient systems, the dissipation potential generates a dissipation distance $\mathcal{D}: \boldsymbol{X}\times\boldsymbol{X}\to[0,\infty]$ leading to an energetic rate-independent system $(\boldsymbol{X},\mathcal{E},\mathcal{D})$. In fact, one does not need a Banach-space structure and can work on general topological spaces $\mathcal{Q}$ with a suitable metric $\mathcal{D}$. A function $u:[0,T]\to\mathcal{Q}$ is called *energetic solution* for $(\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon)$ if (S) and (E) hold:

$$\text{(S)} \quad \mathcal{E}_\varepsilon(t,u(t)) \le \mathcal{E}_\varepsilon(t,w) + \mathcal{D}_\varepsilon(u(t),w) \text{ for all } t\in[0,T] \text{ and } w\in\mathcal{Q}, \tag{1.5}$$

$$\text{(E)} \quad \mathcal{E}_\varepsilon(T,u(T)) + \mathrm{Diss}_{\mathcal{D}_\varepsilon}(u,[0,T]) = \mathcal{E}_\varepsilon(0,u(0)) + \int_0^T \partial_s\mathcal{E}_\varepsilon(s,u(s))\,\mathrm{d}s \tag{1.6}$$

with $\mathrm{Diss}_{\mathcal{D}_\varepsilon}(u,[0,T]) = \sup\sum_{j=1}^N \mathcal{D}_\varepsilon(u(t_{j-1}),u(t_j))$, where the supremum is taken over all $N\in\mathbb{N}$ and all partitions $0=t_0<t_1<\cdots<t_{N-1}<t_N=T$. Again, we have a derivative-free formulation such that evolutionary Γ-convergence can be easily connected to Γ-convergence of the energies $\mathcal{E}_\varepsilon$ and of the dissipation distances $\mathcal{D}_\varepsilon$. We present some simplified versions of the results in [MRS08] and [LiM11]. Theorem 5.4 reads

$$\mathcal{D}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{D}_0 \text{ and } \mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0 \quad \Longrightarrow \quad (\mathcal{Q},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon) \xrightarrow{\mathsf{pE}} (\mathcal{Q},\mathcal{E}_0,\mathcal{D}_0).$$

If the energy is quadratic, i.e. $\mathcal{E}_\varepsilon(t,u) = \frac{1}{2}\langle A_\varepsilon u,u\rangle - \langle \ell_\varepsilon(t),u\rangle$, and $\mathcal{D}_\varepsilon(u_1,u_2) = \Psi_\varepsilon(u_2-u_1)$ on a Hilbert space $\boldsymbol{H}$, we can weaken the convergence for Ψ_ε slightly (cf. Theorem 5.7):

$$\Psi_\varepsilon \xrightarrow{\mathsf{C}} \Psi_0,\ \Psi_\varepsilon \xrightarrow{\Gamma} \Psi_0, \text{ and } \mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0 \quad \Longrightarrow \quad (\boldsymbol{H},\mathcal{E}_\varepsilon,\Psi_\varepsilon) \xrightarrow{\mathsf{pE}} (\boldsymbol{H},\mathcal{E}_0,\Psi_0).$$

As we see from the above results, it is often necessary to have the rather strong Mosco convergence to establish the desired E-convergence of the whole gradient flow. At first glance, such a result looks quite restrictive; however, we made the experience that a proper understanding of the properties (like microstructure or sharp interfaces) of the solutions of $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ is needed even in more direct approaches. Using that knowledge it is often possible to find a suitable blowup or unfolding of the microstructure that turns the usual Γ-convergence into Mosco convergence. A typical example is that of replacing homogenization by two-scale homogenization, as was done in [MiT07] for rate-independent elastoplasticity.

Throughout this survey, we try to give the main themes and to highlight the principal approaches. Hence, we are not always able to give the full details or the optimal results. Moreover, we may not always list all the implicit assumptions, so in any case of doubt the original papers should be consulted.

2 Gradient systems and Γ-convergence

We first define our notion of gradient systems in the sense of modeling. Then, we recall the main definitions for (static) Γ-convergence and discuss a few properties needed especially for the evolutionary context, such as the relation to Legendre transforms. The definitions of evolutionary Γ-convergence are discussed in Section 2.4, and finally we give some examples that will be used later to highlight different aspects of the theory.

2.1 Gradient systems from the modeling point of view

To begin with, we start from the case that the state space is a smooth (finite-dimensional) manifold $\mathcal{X}$ such that at each state $u \in \mathcal{X}$ the tangent space $\mathrm{T}_u\mathcal{X}$ and its dual space $\mathrm{T}_u^*\mathcal{X} = (\mathrm{T}_u\mathcal{X})^*$, the so-called co-tangent space, are well defined. For the modeling it is important to distinguish the states u, the rates or velocities $v = \dot{u} \in \mathrm{T}_u\mathcal{X}$, and the (thermodynamically conjugate) forces $\xi \in \mathrm{T}_u^*\mathcal{X}$. We denote by $\langle \xi, v \rangle$ the dual pairing between $\mathrm{T}_u^*\mathcal{X}$ and $\mathrm{T}_u\mathcal{X}$. By $\mathrm{T}\mathcal{X}$ we denote the tangent bundle $\cup_{u\in\mathcal{X}}(u, \mathrm{T}_u\mathcal{X})$.

A gradient system is a triple $(\mathcal{X}, \mathcal{E}, \mathcal{R})$ with a state space $\mathcal{X}$ as above, a smooth energy functional $\mathcal{E} : [0,T] \times \mathcal{X} \to \mathbb{R}$, and a dissipation potential $\mathcal{R} : \mathrm{T}\mathcal{X} \to [0,\infty]$, i.e. $\mathcal{R}$ depends on the state u and the rate $\dot{u}$. A function $\mathcal{R}$ is called dissipation potential, if for all $u \in \mathcal{X}$, the function $\mathcal{R}(u,\cdot) : \mathrm{T}_u\mathcal{X} \to [0,\infty]$ is convex, lower semicontinuous, and satisfies $\mathcal{R}(u,0) = 0$.

We consider $\xi = \mathrm{D}_u\mathcal{E}(t,u) \in \mathrm{T}_u\mathcal{X}^*$ as (the negative of) the potential restoring force generated by the energy $\mathcal{E}$, while $\partial_t\mathcal{E}(t,u) \in \mathbb{R}$ will be called the power of the external forces. The partial derivative $\eta = \mathrm{D}_{\dot{u}}\mathcal{R}(u,\dot{u}) \in \mathrm{T}_u^*\mathcal{X}$ is the dissipative force induced by the changes of u. The induced evolution of the gradient system $(\mathcal{X}, \mathcal{E}, \mathcal{R})$ is now given in terms of the force balance also called *Biot's equation* (see e.g. [Bio55, Eqn. (2.14)])

$$0 = \mathrm{D}_{\dot{u}}\mathcal{R}(u(t), \dot{u}(t)) + \mathrm{D}_u\mathcal{E}(t, u(t)) \quad \in \mathrm{T}_u^*\boldsymbol{X}. \tag{2.1}$$

In the simplest case the dissipation potential is a quadratic form $\mathcal{R}(u,v) = \frac{1}{2}\langle \mathbb{G}(u)v, v\rangle$ giving a linear constitutive law $\eta = \mathbb{G}(u)\dot{u}$, which is usually called viscous dissipation. We call these systems *classical gradient systems* and also write $(\mathcal{X}, \mathcal{E}, \mathbb{G})$ or $(\mathcal{X}, \mathcal{E}, \mathbb{K})$.

If $\mathcal{R}(u,\cdot)$ is nonquadratic we call $(\mathcal{X},\mathcal{E},\mathcal{R})$ a *generalized gradient system*. See [MPR13] for a natural occurrence of generalized gradient systems obtained via a large-deviation principle and Section 5 for rate-independent systems, which are characterized by the fact that $\mathcal{R}(u,\cdot)$ is positively homogeneous of degree 1, see (5.1).

In the classical case we can invert the operator $\mathbb{G}(u)$ and solve for $\dot u$ to obtain a rate equation, which we also call *Onsager equation* (cf. [Ons31, Eqn. (1.11)])

$$\dot u(t) = -\mathbb{K}(u(t))\mathrm{D}_u\mathcal{E}(t,u(t)) \ \in \mathrm{T}_u\boldsymbol{X}, \quad \text{where } \mathbb{K}(u)=\mathbb{G}(u)^{-1}. \tag{2.2}$$

We call $\mathbb{K}(u)$ the Onsager operator to honor Onsager's seminal contributions to the thermodynamics of dissipative processes, cf. [Ons31, OnM53]. For systems that are always close to local thermodynamic equilibrium, he derived that the relation $\dot u = f(\xi)$ between the thermodynamic driving force ξ and the rate must be linear, i.e. i.e. $\dot u = \mathbb{K}\xi$ with a symmetric and positive semidefinite $\mathbb{K}$, where "$\mathbb{K}=\mathbb{K}^\top$" are his famous reciprocal relations. In his setting $\xi = \mathrm{D}\mathcal{S}(u)$ with $\mathcal{S}$ denoting the entropy.

In a more general setting, we use the dual dissipation potential $\mathcal{R}^*(u,\xi)$, which reads $\mathcal{R}^*(u,\xi)=\frac12\langle\xi,\mathbb{K}(u)\xi\rangle$ in the quadratic case and is defined in (2.11) for the general case, to write the rate equation in the form

$$\dot u(t) = \mathrm{D}_\xi\mathcal{R}^*\big(u(t),-\mathrm{D}_u\mathcal{E}(t,u(t))\big) \ \in \mathrm{T}_u\boldsymbol{X}. \tag{2.3}$$

In principle, equations (2.1) and (2.3) are equivalent, but depending on the context one may prefer one over the other. For the gradient structures for time-continuous discrete Markov chains (see [Maa11]) or the reaction-diffusion systems (see [Mie11b, GlM13, Mie13]) it is crucial that $\mathcal{R}^*$ (or the Onsager operator $\mathbb{K}(u)$) is a sum of different dissipative effects. Thus, $\mathbb{K}(u)$ is explicit, whereas $\mathbb{G}(u)$ or $\mathcal{R}(u,\cdot)$ are only implicitly defined.

An important feature in modeling is energy conservation, when taking the dissipation properly into account. For this it is important to realize that the dissipation potential is usually different from the dissipation functions $\Phi(u,\dot u)$ or $\Phi_*(u,\eta)$, where

$$\Phi(u,\dot u)=\langle \mathrm{D}_{\dot u}\mathcal{R}(u,\dot u),\dot u\rangle \quad\text{and}\quad \Phi_*(u,\xi)=\langle\eta,\mathrm{D}_\xi\mathcal{R}^*(u,\xi)\rangle.$$

For quadratic functions we have $\Phi=2\mathcal{R}$ and $\Phi_*=2\mathcal{R}^*$, but in general there is no such identity. Along solutions of (2.1) or (2.3) the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t,u(t)) - \partial_t\mathcal{E}(t,u(t)) = \langle \mathrm{D}_u\mathcal{E}(t,u(t)),\dot u(t)\rangle = \Phi(u(t),\dot u(t)) \text{ and} \tag{2.4a}$$

$$\Phi(u(t),\dot u(t)) = \Phi_*\big(u(t),-\mathrm{D}_u\mathcal{E}(t,u(t))\big) = \mathcal{R}(u(t),\dot u(t))+\mathcal{R}^*\big(u(t),-\mathrm{D}_u\mathcal{E}(t,u(t))\big). \tag{2.4b}$$

For the latter relation we refer to the Legendre-Fenchel equivalence discussed in Proposition 2.7.

Integrating (2.4a), with Φ replaced by $\mathcal{R}+\mathcal{R}^*$ as in (2.4b), over $t\in[0,T]$ we find the *energy-dissipation balance* (EDB), namely

$$\begin{aligned}\mathcal{E}(T,u(T)) + \int_0^T \mathcal{R}(u(t),\dot u(t))+\mathcal{R}^*\big(u(t),-\mathrm{D}_u\mathcal{E}(t,u(t))\big)\,\mathrm{d}t \\ = \mathcal{E}(0,u(0)) + \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t \quad \text{for all } T>0.\end{aligned} \tag{2.5}$$

At first sight it seems that the EDB is a simple consequence of (2.1) or (2.3). However, writing the total dissipation in the form $\mathcal{R}+\mathcal{R}^*$ (rather than using e.g. $\frac{2}{9}\Phi+\frac{7}{9}\Phi_*$) it can be shown that the EDB is equivalent to (2.1) and (2.3), see Theorem 3.2.

We emphasize here that each gradient system $(\mathcal{X},\mathcal{E},\mathcal{R})$ has a well-defined associated evolution equation $\dot u(t)=\boldsymbol{V}_{\mathcal{E},\mathcal{R}}(t,u(t)):=\mathrm{D}_\xi\mathcal{R}^*\big(u(t),-\mathrm{D}_u\mathcal{E}(t,u(t))\big)$. However, for one given evolution equation $\dot u(t)=\boldsymbol{V}_{\mathcal{E},\mathcal{R}}(t,u(t))$ there may exist many different gradient structures. So, the gradient structure is an additional information which contains additional physical information on the model under investigation.

Example 2.1 *Here we show that ODEs may have several genuinely different gradient structures. For the scalar linear ODE* $\dot u=-u=:\boldsymbol{V}(u)$ *we take any smooth, convex* $\phi:\mathbb{R}\to\mathbb{R}$ *with* $\phi(0)=\phi'(0)=0$ *and* $\phi''(0)>0$. *Then,* $\mathcal{E}(u)=\phi(u)$ *and* $\mathbb{K}(u)=\frac{u}{\phi'(u)}>0$ *defines a gradient structure* $(\mathbb{R},\mathcal{E},\mathbb{K})$ *for* $\dot u=-u$.

Similar construction also works for systems of ODEs. For example, consider

$$\dot u_1=u_2-u_1,\quad \dot u_2=-3u_2,$$

choose a uniformly convex ϕ *and set* $\mathcal{E}(u_1,u_2)=\phi(u_1)+\phi(u_2)+\phi(-u_1-u_2)$ *and*

$$\mathcal{R}^*(u_1,u_2,\xi_1,\xi_2)=\tfrac{u_2-u_1}{\phi'(u_2)-\phi'(u_1)}\big(\xi_2-\xi_1\big)^2+\tfrac{2u_1+u_2}{\phi'(u_1)-\phi'(-u_1-u_2)}\,\xi_1^2+\tfrac{u_1+2u_2}{\phi'(u_2)-\phi'(-u_1-u_2)}\,\xi_2^2.$$

See [Maa11] for similar gradient structures for time-continuous Markov chains.

Of course, in general thermodynamical modeling we are interested in partial differential equations, where the above theory has to be generalized suitably. In particular, the underlying space $\mathcal{X}$ may not be a smooth manifold but is usually a subset of a Banach space $\boldsymbol{X}$. Moreover, the functionals $\mathcal{E}$ and $\mathcal{R}$ will no longer be smooth but may attain the value $+\infty$. Hence, the derivatives $\mathrm{D}_u\mathcal{E}(t,u)$, $\mathrm{D}_v\mathcal{R}(u,v)$, and $\mathrm{D}_\xi\mathcal{R}^*(u,\xi)$ must be generalized in a suitable way. We refer to Section 3.1 for a mathematical precise setup. The main point to be remembered is, that the state u, the rates $v=\dot u$, and the forces ξ lie in different spaces. Only for the rates and the forces we have linear spaces and a duality pairing $\langle\xi,\dot u\rangle$, which denotes a power.

Working with partial differential equations, we may proceed formally without specifying a function space $\boldsymbol{X}$, which is the typical approach in thermomechanical modeling, see e.g. [Mie11b, Mie13, MiT14]. Functional derivatives can be interpreted as variational derivatives by assuming that all functions are sufficiently smooth. As an example consider the PDE

$$\dot u=\Delta u \text{ in } \Omega\subset\mathbb{R}^d,\quad \nabla u\cdot\nu=0 \text{ on } \partial\Omega, \tag{2.6}$$

where Ω is a smooth bounded domain in $\mathbb{R}^d$. There are two simple quadratic gradient structures, namely $(\mathrm{L}^2(\Omega),\mathcal{E}_0,\mathbb{K}_0)$ and $(\mathrm{H}^1(\Omega),\mathcal{E}_1,\mathbb{K}_1)$ with

$$\mathcal{E}_0(u)=\int_\Omega\frac{u^2}{2}\,\mathrm{d}x,\quad \mathbb{K}_0\xi=-\Delta\xi,\qquad \mathcal{E}_1(u)=\int_\Omega\frac{1}{2}|\nabla u|^2\,\mathrm{d}x,\quad \mathbb{K}_1\xi=\xi.$$

However, if we consider (2.6) a heat equation for the absolute temperature u of a gas, we choose the entropy $\mathcal{S}$ for $\dot u=+\mathbb{K}_{\mathrm{heat}}(u)\mathrm{D}\mathcal{S}(u)$ with

$$\mathcal{S}(u)=\int_\Omega\log(u(x))\,\mathrm{d}x\quad\text{and}\quad \mathbb{K}_{\mathrm{heat}}(u)\xi=-\operatorname{div}\big(u^2\nabla\xi\big),$$

see [Mie11b, Mie13]. If (2.6) describes the diffusion of a density u, then the Wasserstein gradient structure for the relative entropy $\mathcal{H}$, introduced in [JKO98, Ott01], is the right choice:

$$\mathcal{H}(u) = \int_\Omega u \log u \,\mathrm{d}x \quad \text{and} \quad \mathbb{K}_{\mathrm{Wass}}(u)\xi = -\operatorname{div}(u\nabla\xi).$$

Of course, the different choices of the gradient structures may influence the theory of evolutionary Γ-convergence. In Section 3.5.2 we discuss the simple model $\dot{u}(t,x) = -a(\frac{1}{\varepsilon}x)u(t,x)$ and show that different gradient structures lead to different evolutionary Γ-limits, see Corollary 3.8.

2.2 Γ-convergence for (static) functionals

We consider a reflexive Banach space $\boldsymbol{X}$ and functionals $\mathcal{J}_\varepsilon : \boldsymbol{X} \to \mathbb{R}_\infty$. Strong and weak convergence in $\boldsymbol{X}$ will be denoted by $u_k \to u$ and $v_k \rightharpoonup v$, respectively. We first introduce more classical notions of convergence of functionals, namely the pointwise convergence $\mathcal{J}_\varepsilon \xrightarrow{\mathrm{pw}} \mathcal{J}_0$ and the strong or weak continuous convergence defined via

$$\mathcal{J}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{J}_0, \quad \text{if } u_\varepsilon \to u \;\Rightarrow\; \mathcal{J}_\varepsilon(u_\varepsilon) \to \mathcal{J}_0(u); \tag{2.7a}$$

$$\mathcal{J}_\varepsilon \xrightarrow{\mathsf{C}_\rightharpoonup} \mathcal{J}_0, \quad \text{if } u_\varepsilon \rightharpoonup u \;\Rightarrow\; \mathcal{J}_\varepsilon(u_\varepsilon) \to \mathcal{J}_0(u). \tag{2.7b}$$

In the context of minimization of functionals, the concept of Γ-*convergence* is more natural, see Theorem 2.4. This convergence was originally called variational convergence or epi-graph convergence (cf. [DeF75, Att84]), but nowadays the term Γ-convergence is more common and we refer to [Dal93, Bra02, Bra06, Bra13] for further details.

Definition 2.2 (Γ and Mosco convergence) *Let $\boldsymbol{X}$ be a reflexive Banach space.*
We say that $\mathcal{J}_\varepsilon$ weakly Γ-converges to $\mathcal{J}_0$ and write $\mathcal{J}_\varepsilon \xrightarrow{\Gamma_\rightharpoonup} \mathcal{J}_0$, if (G1w) and (G2w) hold:

(G1w) $u_\varepsilon \rightharpoonup u \implies \mathcal{J}_0(u) \le \liminf_{\varepsilon\to 0} \mathcal{J}_\varepsilon(u_\varepsilon)$ *(liminf estimate)*

(G2w) $\forall\, \widehat{u}\ \exists\, (\widehat{u}_\varepsilon)_\varepsilon$: $\widehat{u}_\varepsilon \rightharpoonup \widehat{u}$ *and* $\mathcal{J}_0(\widehat{u}) = \lim_{\varepsilon\to 0} \mathcal{J}_\varepsilon(\widehat{u}_\varepsilon)$ *(recovery seq. exist)*

We say that $\mathcal{J}_\varepsilon$ strongly Γ-converges to $\mathcal{J}_0$ and write $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$, if (G1s) and (G2s) hold:

(G1s) $u_\varepsilon \to u \implies \mathcal{J}_0(u) \le \liminf_{\varepsilon\to 0} \mathcal{J}_\varepsilon(u_\varepsilon)$ *(liminf estimate)*

(G2s) $\forall\, \widehat{u}\ \exists\, (\widehat{u}_\varepsilon)_\varepsilon$: $\widehat{u}_\varepsilon \to \widehat{u}$ *and* $\mathcal{J}_0(\widehat{u}) = \lim_{\varepsilon\to 0} \mathcal{J}_\varepsilon(\widehat{u}_\varepsilon)$ *(recovery seq. exist)*

We say that $\mathcal{J}_\varepsilon$ Mosco converges to $\mathcal{J}_0$ and write $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{J}_0$, if $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ and $\mathcal{J}_\varepsilon \xrightarrow{\Gamma_\rightharpoonup} \mathcal{J}_0$.

Since all strongly converging sequences are also weakly converging we have the implications (G1w) $\Rightarrow$ (G1s) and (G2s) $\Rightarrow$ (G2w). Hence, for Mosco convergence one needs to check only (G1w) and (G2s). We will see in Lemma 2.6 that there are simple quadratic functionals for which weak and strong Γ-limits exist, but they are different.

The conditions (G2w) and (G2s) on the existence of recovery sequences is also called "limsup estimate". The recovery sequences are crucial since they capture the correct microscopic behavior that is needed to recover the correct (namely the lowest possible) macroscopic energy $\mathcal{J}_0(\widehat{u})$.

Clearly, continuous convergence is much stronger than Γ-convergence. We have the following relations.

Lemma 2.3 *Assume that $\boldsymbol{X}$ is a reflexive Banach space and $\mathcal{J}_\varepsilon$, $\mathcal{K}_\varepsilon : \boldsymbol{X} \to \mathbb{R}_\infty$.*

(a) $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ *and* $\mathcal{K}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{K}_0 \implies \mathcal{J}_\varepsilon+\mathcal{K}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0+\mathcal{K}_0$

(b) $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ *and* $\mathcal{K}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{K}_0 \implies \mathcal{J}_\varepsilon+\mathcal{K}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0+\mathcal{K}_0$

(c) $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{J}_0 \implies \mathcal{J}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{J}_0$

(d) If $(\mathcal{J}_\varepsilon)_{\varepsilon\in[0,1]}$ is strongly equicontinuous, then $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ implies $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{J}_0$.

The origin for the definition of Γ-convergence, which is clearer in the original name "variational convergence", is the following convergence of minimizers, see [Dal93, Bra02].

Theorem 2.4 (Convergence of minimizers) *Assume $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ in $\boldsymbol{X}$ and that* $\inf \mathcal{J}_0 =: \alpha_0 > -\infty$. *Moreover, assume that there exists a closed bounded set $B \subset \boldsymbol{X}$ such that $\mathcal{J}_\varepsilon(u) \le \alpha_0+1$ implies $u \in B$, then for every sequence $(u_{\varepsilon_k})_{k\in\mathbb{N}}$ with $\varepsilon_k \to 0$ we have:*

$$\text{if } u_\varepsilon \rightharpoonup \widetilde{u} \text{ and } \mathcal{J}_{\varepsilon_k}(u_{\varepsilon_k}) \to \alpha_0, \quad \text{then } \mathcal{J}_0(\widetilde{u}) = \alpha_0.$$

In particular, if u_ε is a minimizer of $\mathcal{J}_\varepsilon$ (i.e. $\mathcal{J}_\varepsilon(u_\varepsilon) = \inf_{\boldsymbol{X}} \mathcal{J}_\varepsilon$), then any accumulation point of $(u_\varepsilon)_{\varepsilon\in]0,1[}$ is a minimizer of $\mathcal{J}_0$.

The following useful result seems to be folklore, but is not easy to find.

Proposition 2.5 (Γ- versus Mosco convergence) *Assume that $\boldsymbol{X}$ and $\boldsymbol{Z}$ are reflexive Banach spaces such that $\boldsymbol{Z}$ is compactly embedded in $\boldsymbol{X}$, written $\boldsymbol{Z} \Subset \boldsymbol{X}$. Moreover, assume that the functionals $\mathcal{J}_\varepsilon$ are equicoercive in $\boldsymbol{Z}$, i.e.*

$$\forall\, J > 0\ \exists\, R > 0\ \forall\, \varepsilon > 0,\ u \in \boldsymbol{X}: \quad \mathcal{J}_\varepsilon(u) \le J \ \Rightarrow\ \|u\|_{\boldsymbol{Z}} \le R: \tag{2.8}$$

Then, $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{J}_0$ in $\boldsymbol{X}$ is equivalent to $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ in $\boldsymbol{Z}$.

Proof: The equicoercivity is meant in the way that all $\mathcal{J}_\varepsilon$ take the value $+\infty$ on $\boldsymbol{X} \setminus \boldsymbol{Z}$.

"⇒" We start from $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{J}_0$ in $\boldsymbol{X}$. If $u_\varepsilon \rightharpoonup u$ in $\boldsymbol{Z}$, then this also holds in $\boldsymbol{X}$. Hence, the liminf estimate follows. To construct a recovery sequence $\widehat{u}_\varepsilon \rightharpoonup \widehat{u}$ in $\boldsymbol{Z}$ for arbitrary $\widehat{u} \in \boldsymbol{Z}$, we first assume $\mathcal{J}_0(\widehat{u}) < \infty$. We choose the recovery sequence $\widehat{u}_\varepsilon$ guaranteed by $\mathcal{J}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{J}_0$ in $\boldsymbol{X}$, i.e. we know $\widehat{u}_\varepsilon \to \widehat{u}$ in $\boldsymbol{X}$. The equicoercivity (2.8) and $\mathcal{J}_\varepsilon(\widehat{u}_\varepsilon) \to \mathcal{J}_0(\widehat{u}) < \infty$ imply $\|\widehat{u}_\varepsilon\|_{\boldsymbol{Z}} \le R$. Hence, $\widehat{u}_\varepsilon \rightharpoonup \widehat{u}$ in $\boldsymbol{Z}$ by reflexivity of $\boldsymbol{Z}$. If $\mathcal{J}_0(\widehat{u}) = \infty$, we choose $\widehat{u}_\varepsilon = \widehat{u}$ giving $\widehat{u}_\varepsilon \to \widehat{u}$ in $\boldsymbol{Z}$. Hence, the liminf estimate yields $\infty = \mathcal{J}_0(\widehat{u}) \le \liminf_{\varepsilon\to 0} \mathcal{J}_\varepsilon(\widehat{u})$, which shows that we have a recovery sequence in $\boldsymbol{Z}$.

"⇐" Given $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ in $\boldsymbol{Z}$, we take any sequence $u_\varepsilon \rightharpoonup u$ in $\boldsymbol{X}$. If $\liminf_{\varepsilon\to 0} \|u_\varepsilon\|_{\boldsymbol{Z}} = \infty$, then the equicoercivity implies $\mathcal{J}_\varepsilon(u_\varepsilon) \to \infty$ and the liminf estimate holds. If for some subsequence $\|u_{\varepsilon_k}\|_{\boldsymbol{Z}} \le C$, then $u_{\varepsilon_k} \rightharpoonup u$ in $\boldsymbol{Z}$, and the liminf estimate follows from that of $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ in $\boldsymbol{Z}$. For the construction of recovery sequences, we can choose $\widehat{u}_\varepsilon = \widehat{u}$ if $\widehat{u} \in \boldsymbol{X} \setminus \boldsymbol{Z}$. If $\widehat{u} \in \boldsymbol{Z}$ we choose a recovery sequence $\widehat{u}_\varepsilon \rightharpoonup \widehat{u}$ in $\boldsymbol{Z}$. By the compact embedding we have $\widehat{u}_\varepsilon \to \widehat{u}$ in $\boldsymbol{X}$ and the proof is finished. ∎

The following lemma presents a simple quadratic example in which the weak and the strong Γ-limits exist but they are different. We define

$$\mathcal{F}_\varepsilon(w) = \int_\Omega \frac{1}{2} w(x) \cdot \mathbb{A}\big(\frac{1}{\varepsilon}x\big) w(x)\,\mathrm{d}x \quad \text{for } w \in \boldsymbol{X} = \mathrm{L}^2(\Omega;\mathbb{R}^m),$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $\mathbb{A} \in \mathrm{L}^\infty(\mathbb{R}^d; \mathbb{R}^{m\times m}_{\mathrm{sym}})$ is 1-periodic, i.e. $\mathbb{A}(y+n) = \mathbb{A}(y)$ for all $y \in \mathbb{R}^d$ and all $n \in \mathbb{Z}^d$. Moreover, we assume that $\mathbb{A}$ is uniformly positive definite, i.e. $\underline{a}|w|^2 \le w \cdot \mathbb{A}(y)w \le \overline{a}|w|^2$ for $\overline{a} > \underline{a} > 0$.

Lemma 2.6 *Define the arithmetic and harmonic mean of $\mathbb{A}$ via*

$$\mathbb{A}_{arith} := \int_{[0,1]^d} \mathbb{A}(y)\,\mathrm{d}y \quad \text{and} \quad \mathbb{A}_{harm} := \Big(\int_{[0,1]^d} \mathbb{A}(y)^{-1}\,\mathrm{d}y \Big)^{-1}$$

and the two functionals

$$\mathcal{F}_{arith}(w) = \int_\Omega \frac{1}{2} w(x) \cdot \mathbb{A}_{arith} w(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{F}_{harm}(w) = \int_\Omega \frac{1}{2} w(x) \cdot \mathbb{A}_{harm} w(x)\,\mathrm{d}x.$$

In $\boldsymbol{X} = \mathrm{L}^2(\Omega;\mathbb{R}^m)$ we have $\mathcal{F}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{F}_{harm}$ and $\mathcal{F}_\varepsilon \overset{\mathsf{C}}{\to} \mathcal{F}_{arith}$, which implies $\mathcal{F}_\varepsilon \overset{\Gamma}{\to} \mathcal{F}_{arith}$.

Proof: We first prove $\mathcal{F}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{F}_{\mathrm{harm}}$.

For the liminf estimate assume $w_\varepsilon \rightharpoonup w$ in $\mathrm{L}^2(\Omega)$. Writing $\mathbb{A}_\varepsilon(x) = \mathbb{A}(\frac{1}{\varepsilon}x)$ we have

$$\mathcal{F}_\varepsilon(w_\varepsilon) = \frac{1}{2}\int_\Omega w_\varepsilon \cdot \mathbb{A}_\varepsilon w_\varepsilon\,\mathrm{d}x = \tag{2.9}$$

$$= \frac{1}{2}\int_\Omega \underbrace{(w_\varepsilon - \mathbb{A}_\varepsilon^{-1}\mathbb{A}_{\mathrm{harm}}w)\cdot\mathbb{A}_\varepsilon(w_\varepsilon - \mathbb{A}_\varepsilon^{-1}\mathbb{A}_{\mathrm{harm}}w)}_{\ge 0} + 2\underbrace{w_\varepsilon}_{\to w}\cdot\mathbb{A}_{\mathrm{harm}}w - \mathbb{A}_{\mathrm{harm}}w\cdot\underbrace{\mathbb{A}_\varepsilon^{-1}}_{\overset{*}{\rightharpoonup}\mathbb{A}_{\mathrm{harm}}^{-1}}\mathbb{A}_{\mathrm{harm}}w\,\mathrm{d}x.$$

Dropping the nonnegative term and taking the limit $\varepsilon \to 0$ give the desired lower estimate $\liminf_\varepsilon \mathcal{F}_\varepsilon(w_\varepsilon) \ge \frac{1}{2}\int_\Omega 0 + 2w\cdot\mathbb{A}_{\mathrm{harm}}w - w\cdot\mathbb{A}_{\mathrm{harm}}w\,\mathrm{d}x = \mathcal{F}_0(w)$.

For the limsup-estimate we use the same reformulation of $\mathcal{F}_\varepsilon$ as in (2.9). For a given $\widehat{w}$ we choose $\widehat{w}_\varepsilon = \mathbb{A}_\varepsilon^{-1}\mathbb{A}_{\mathrm{harm}}\widehat{w}$. Since by construction the first term in the integral is 0 we find $\mathcal{F}_\varepsilon(\widehat{w}_\varepsilon) = \frac{1}{2}\int_\Omega 0 + 2\mathbb{A}_\varepsilon^{-1}\mathbb{A}_{\mathrm{harm}}\widehat{w}\cdot\mathbb{A}_{\mathrm{harm}}\widehat{w} - \mathbb{A}_{\mathrm{harm}}\widehat{w}\cdot\mathbb{A}_\varepsilon^{-1}\mathbb{A}_{\mathrm{harm}}\widehat{w}\,\mathrm{d}x \to \mathcal{F}_{\mathrm{harm}}(\widehat{w})$.

For the proof of strong continuous convergence take any $w_\varepsilon \to w$ in $\mathrm{L}^2(\Omega)$ and write

$$\mathcal{F}_\varepsilon(w_\varepsilon) = \frac{1}{2}\int_\Omega w\cdot\underbrace{\mathbb{A}_\varepsilon w}_{\to\mathbb{A}_{\mathrm{arith}}w} - 2w\cdot\mathbb{A}_\varepsilon\underbrace{(w-w_\varepsilon)}_{\to 0} + \underbrace{(w-w_\varepsilon)}_{\to 0}\cdot\mathbb{A}_\varepsilon(w-w_\varepsilon)\,\mathrm{d}x \;\to\; \mathcal{F}_{\mathrm{arith}}(w). \tag{2.10}$$

This proves the strong continuous and hence the strong Γ-convergence. ∎

2.3 Prerequisites from convex analysis

For each $u \in \boldsymbol{X}$, the dissipation potentials $\mathcal{R}_\varepsilon(u,\cdot) : \boldsymbol{X} \to [0,\infty]$ are always convex, and lower semicontinuous (lsc). So we can apply the Legendre-Fenchel theory for convex functionals $\Psi : \boldsymbol{X} \to \mathbb{R}_\infty$, where we always assume that $\boldsymbol{X}$ is on a reflexive Banach space. The Legendre-Fenchel transform $\Psi^* = \mathrm{Leg}(\Psi) : \boldsymbol{X}^* \to \mathbb{R}_\infty$ is defined via

$$\Psi^*(\xi) := \sup\{\, \langle \xi, v\rangle - \Psi(v) \mid v \in \boldsymbol{X}\,\}, \tag{2.11}$$

where $\langle\cdot,\cdot\rangle$ is the natural dual pairing of $\boldsymbol{X}^*$ and $\boldsymbol{X}$, see [Roc70, EkT76]. Clearly, Ψ^* is again convex, lsc, and satisfies $\Psi^*(0)=0$. In particular, the dual dissipation potential $\mathcal{R}_\varepsilon^*$ is defined via $\mathcal{R}_\varepsilon^*(u,\cdot)=\mathrm{Leg}\big(\mathcal{R}_\varepsilon(u,\cdot)\big)$. Elementary examples are

$$\Psi(v)=\frac{1}{2}\langle\mathbb{G}v,v\rangle \iff \Psi^*(\xi)=\frac{1}{2}\langle\xi,\mathbb{G}^{-1}\xi\rangle$$
$$\Psi(v)=\frac{1}{p}\|v\|_{\boldsymbol{X}}^p \iff \Psi^*(\xi)=\frac{1}{p^*}\|\xi\|_{\boldsymbol{X}^*}^{p^*},\quad \text{where } 1<p<\infty \text{ and } p^*=\frac{p}{p-1}.$$

The fundamental properties of the Legendre-Fenchel transform are the duality relation $\mathrm{Leg}\big(\mathrm{Leg}\Psi\big)=\Psi$ or $\Psi^{**}=\Psi$ and the

$$\text{Young-Fenchel estimate:}\qquad \forall\, v\in\boldsymbol{X}\ \forall\,\xi\in\boldsymbol{X}^*:\ \Psi(v)+\Psi^*(\xi)\geq\langle\xi,v\rangle. \tag{2.12}$$

To discuss the case of equality in this estimate we need the *subdifferential of the convex functional* Ψ, which is defined via

$$\partial\Psi(v):=\{\,\eta\in\boldsymbol{X}^*\,|\,\forall\, w\in\boldsymbol{X}:\ \Psi(w)\geq\Psi(v)+\langle\eta,w{-}v\rangle\,\}\subset\boldsymbol{X}^*.$$

Note that $\partial\Psi$ is set-valued and we will write $\partial\Psi:\boldsymbol{X}\rightrightarrows\boldsymbol{X}^*$. However, if Ψ is differentiable with Gateaux derivative $\mathrm{D}\Psi(v)$, then $\partial\Psi(v)=\{\mathrm{D}\Psi(v)\}$.

The Fenchel equivalence characterizes equality in the Young-Fenchel estimate (2.12).

Proposition 2.7 (Fenchel equivalence [Fen49]) *Let $\boldsymbol{X}$ be a reflexive Banach space and $\Psi:\boldsymbol{X}\to\mathbb{R}_\infty$ be proper, convex, and lsc. Then, we have*

$$\textit{(i)}\ \ \xi\in\partial\Psi(v)\quad\Longleftrightarrow\quad \textit{(ii)}\ \ v\in\partial\Psi^*(\xi)\quad\Longleftrightarrow\quad \textit{(iii)}\ \ \Psi(v)+\Psi^*(\xi)=\langle\xi,v\rangle.$$

For a proof we refer to [EkT76]. We emphasize that the relation (i) is a relation in dual space $\boldsymbol{X}^*$, (ii) is a relation in $\boldsymbol{X}$, and (iii) is a relation in $\mathbb{R}$.

A further fundamental property of the Legendre transform is related to its continuity with respect to weak or strong Γ-convergence. This result will be important for studying E-convergence in Section 3.

Theorem 2.8 ([Att84, pp. 271]) *Let $\boldsymbol{X}$ be a reflexive Banach space and assume that all $\mathcal{F}_\varepsilon:\boldsymbol{X}\to[0,\infty]$ are convex, lsc, and satisfy $\mathcal{F}_\varepsilon(0)=0$. Then,*

$$\mathcal{F}_\varepsilon\ \overset{\Gamma}{\rightharpoonup}\ \mathcal{F}\qquad\Longleftrightarrow\qquad \mathcal{F}_\varepsilon^*\ \overset{\Gamma}{\to}\ \mathcal{F}^*\,.$$

The duality and the switch between the weak and strong convergence appears natural, because the definition of the Legendre transform involves the duality product $\langle\xi,v\rangle$. A well-known result from linear functional analysis states that $\langle\xi_\varepsilon,v_\varepsilon\rangle\to\langle\xi,v\rangle$ if either $v_\varepsilon\rightharpoonup v$ and $\xi_\varepsilon\to\xi$ or vice versa $v_\varepsilon\to v$ and $\xi_\varepsilon\rightharpoonup\xi$.

Under the assumptions of the above theorem the definition of Mosco convergence gives the equivalences

$$\mathcal{F}_\varepsilon\overset{\mathsf{M}}{\to}\mathcal{F}\quad\Longleftrightarrow\quad \mathcal{F}_\varepsilon^*\overset{\mathsf{M}}{\to}\mathcal{F}^*\quad\Longleftrightarrow\quad\Big(\ \mathcal{F}_\varepsilon\overset{\Gamma}{\rightharpoonup}\mathcal{F}\ \text{and}\ \mathcal{F}_\varepsilon^*\overset{\Gamma}{\rightharpoonup}\mathcal{F}^*\ \Big). \tag{2.13}$$

Lemma 2.6 provides an interesting example for the application of Theorem 2.8. In fact, we have $\mathcal{F}_\varepsilon^*(\xi)=\frac{1}{2}\int_\Omega\xi\cdot\mathbb{A}_\varepsilon^{-1}\xi\,\mathrm{d}x$. Thus, the strong convergence for $\mathcal{F}_\varepsilon^*$ leads to an effective matrix $\mathrm{arith}(\mathbb{A}^{-1})=\mathrm{harm}(\mathbb{A})^{-1}$.

Another important tool of convex analysis is the weak-strong closedness of the graphs of the subdifferentials $\partial\mathcal{E}_\varepsilon:\boldsymbol{X}\rightrightarrows\boldsymbol{X}^*$ (which is in fact equivalent to Mosco convergence).

Proposition 2.9 (Strong weak-closedness, [Att84, Thm. 3.66]) *Assume that all* $\mathcal{E}_\varepsilon : \boldsymbol{X} \to [0,\infty]$ *are lsc and convex and that* $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$. *Then, we have*

$$u_\varepsilon \to u,\ \mathcal{E}_\varepsilon(u_\varepsilon) \to e_0 \in \mathbb{R},\ \partial\mathcal{E}_\varepsilon(u_\varepsilon) \ni \xi_\varepsilon \rightharpoonup \xi \quad \Longrightarrow \quad e_0 = \mathcal{E}_0(u) \text{ and } \xi \in \partial\mathcal{E}_0(u). \tag{2.14}$$

The same conclusion holds under the assumptions $u_\varepsilon \rightharpoonup u$ *and* $\partial\mathcal{E}_\varepsilon(u_\varepsilon) \ni \xi_\varepsilon \to \xi$.

Proof: The convexity of $\mathcal{E}_\varepsilon$ gives $\mathcal{E}_\varepsilon(w) \ge \mathcal{E}_\varepsilon(u_\varepsilon) + \langle \xi_\varepsilon, w-u_\varepsilon\rangle$ of all $w \in \boldsymbol{X}$. For fixed $\widehat{u}$ the Mosco convergence provides a recovery sequence $\widehat{u}_\varepsilon$ with $\widehat{u}_\varepsilon \to \widehat{u}$, $\mathcal{E}_\varepsilon(\widehat{u}_\varepsilon) \to \mathcal{E}_0(\widehat{u})$. Hence, inserting $w = \widehat{u}_\varepsilon$ yields $\mathcal{E}_\varepsilon(\widehat{u}_\varepsilon) \ge \mathcal{E}_\varepsilon(u_\varepsilon) + \langle \xi_\varepsilon, \widehat{u}_\varepsilon - u_\varepsilon\rangle$.

Taking the limit $\varepsilon \to 0$, all three terms converge (where we use weak-strong continuity of the duality product $\langle\cdot,\cdot\rangle$), and we obtain the relation $\mathcal{E}_0(\widehat{u}) \ge e_0 + \langle \xi, \widehat{u}-u\rangle$ and choosing $\widehat{u} = u$ yields $\mathcal{E}_0(u) \ge e_0$. However, the liminf estimate for $u_\varepsilon \rightharpoonup u$ gives $e_0 \ge \mathcal{E}_0(u)$. Thus, $e_0 = \mathcal{E}_0(u)$ and we conclude $\xi \in \partial\mathcal{E}_0(u)$ as desired. ∎

2.4 Definitions for Γ-convergence of evolutionary systems

According to Theorem 2.4 the definition of (static) Γ-convergence implies the convergence of minimizers. If we interpret minimizers as "solutions of a static variational problem", we may state that variational convergence implies that "the solutions of variational problems converge". We now define several versions of evolutionary convergence in a similar way.

We start with *semiflow convergence* as was defined in (1.1). Assume that for all $\varepsilon \in [0,1]$ the evolutionary systems $\dot{u}_\varepsilon = \boldsymbol{V}_\varepsilon(u_\varepsilon)$ have for each $u_\varepsilon^0 \in \boldsymbol{X}$ at least one solution $u_\varepsilon : [0,T] \to \boldsymbol{X}$ with $u_\varepsilon(0) = u_\varepsilon^0$. We say that $(\boldsymbol{X}, \boldsymbol{V}_\varepsilon)$ *strongly (or weakly) semiflow-converge* to $(\boldsymbol{X}, \boldsymbol{V}_0)$, if $u_\varepsilon(0) \to u^0$ (or $u_\varepsilon(0) \rightharpoonup u^0$) implies that there exist a sequence $\varepsilon_k \to 0$ and a solution $u : [0,T] \to \boldsymbol{X}$ of $\dot{u} = \boldsymbol{V}_0(u)$ with $u(0) = u^0$ such that $u_{\varepsilon_k}(t) \to u(t)$ (or $u_{\varepsilon_k}(t) \rightharpoonup u(t)$) for all $t \in]0,T]$.

For gradient systems $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ we will use a special types of semiflow convergence that includes the additional condition on energy convergence, which is the reason why we include the symbol Γ in the name. Throughout this work we call $u_\varepsilon : [0,T] \to \boldsymbol{X}$ a solution for the gradient system $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ if it satisfies the gradient evolution $0 \in \partial_{\dot{u}}\mathcal{R}_\varepsilon(u,\dot{u}) + \mathrm{D}\mathcal{E}_\varepsilon(t,u(t))$ in a suitable sense.

Definition 2.10 (Evolutionary Γ-convergence) *For* $\varepsilon \in [0,1]$ *consider gradient systems* $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$.

(A) We say $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ E-converges to $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$ *and write* $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \xrightarrow{\mathsf{E}} (\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$, *if*

$$\left.\begin{array}{c} u_\varepsilon : [0,T] \to \boldsymbol{X} \\ \text{is sol. of } (\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \\ \text{and } u_\varepsilon(0) \to u^0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{c} \exists\, u \text{ sol. of } (\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0) \text{ with } u(0) = u^0 \\ \text{and a subsequence } \varepsilon_k \to 0 : \\ \forall\, t \in]0,T] :\ u_{\varepsilon_k}(t) \to u(t) \text{ and} \\ \mathcal{E}_\varepsilon(u_{\varepsilon_k}(t)) \to \mathcal{E}_0(u(t)). \end{array}\right. \tag{2.15}$$

(B) We say that $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ E-converges with well-prepared initial conditions to $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$,

and shortly write $(\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon) \overset{\mathsf{pE}}{\longrightarrow} (\boldsymbol{X},\mathcal{E}_0,\mathcal{R}_0)$*, if*

$$\left.\begin{array}{c} u_\varepsilon : [0,T] \to \boldsymbol{X} \\ \text{is sol. of } (\boldsymbol{X},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon), \\ u_\varepsilon(0) \to u^0, \text{ and} \\ \mathcal{E}_\varepsilon(u_\varepsilon(0)) \to \mathcal{E}_0(u^0) < \infty \end{array}\right\} \implies \left\{\begin{array}{c} \exists\, u \text{ sol. of } (\boldsymbol{X},\mathcal{E}_0,\mathcal{R}_0) \text{ with } u(0)=u^0 \\ \text{and a subsequence } \varepsilon_k \to 0 : \\ \forall\, t \in]0,T] : \ u_{\varepsilon_k}(t) \to u(t) \text{ and} \\ \mathcal{E}_\varepsilon(u_{\varepsilon_k}(t)) \to \mathcal{E}_0(u(t)). \end{array}\right. \tag{2.16}$$

(C) If all the strong convergences in (A) or (B) are replaced by weak convergences, then we write $\overset{\mathsf{E}}{\rightharpoonup}$ *or* $\overset{\mathsf{pE}}{\rightharpoonup}$*, respectively. Any of these four types of convergence is called* evolutionary Γ-convergence.

As in [SaS04] we use the symbol Γ in the name "evolutionary Γ-convergence" to indicate the relation to Γ-convergence for the energy. In fact, we typically expect that $\mathcal{E}_\varepsilon \overset{\Gamma}{\to} \mathcal{E}_0$ is a necessary condition for evolutionary Γ-convergence. Then, the statement that the energies $\mathcal{E}_\varepsilon(u_\varepsilon(t))$ converge to $\mathcal{E}_0(u(t))$ means that $u_\varepsilon(t)$ is a recovery sequence for $u(t)$. If one asks this condition at the initial time $t = 0$, one speaks of *well-preparedness of the initial conditions*, i.e. the initial conditions must be a recovery sequence as well. This explains our name "well-prepared E-convergence".

Note that E-convergence implies well-prepared E-convergence, since in the later case the assumptions are stronger while the conclusions are the same. Often the energy convergence improves the convergence $u_\varepsilon(t) \rightharpoonup u_0(t)$ into strong convergence or weak convergence in a better space.

The aim of these notes is to provide conditions for suitable convergences for $(\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$ towards a limiting pair $(\mathcal{E}_0,\mathcal{R}_0)$ on a suitable space $\boldsymbol{X}$ such that evolutionary Γ-convergence is guaranteed. We emphasize that separate convergence of the two functionals will not be sufficient. It is important to realize that there has to be some *compatibility* between the convergences $\mathcal{E}_\varepsilon \rightsquigarrow \mathcal{E}_0$ and $\mathcal{R}_\varepsilon \rightsquigarrow \mathcal{R}_0$. This will be the topic of Sections 3 and 4, while the rest of the present section introduces a few examples that will be used later for applying the abstract theory.

2.5 An ODE example

In the first example we discuss a linear ODE. *Since the solutions and the* Γ*-limits can be explicitly constructed we can check the validity of evolutionary* Γ*-convergence without any abstract theory. In particular, we will show that in general it is not enough to have separate Mosco convergences* $\mathcal{E}_\varepsilon \overset{\mathsf{M}}{\to} \mathcal{E}_0$ *and* $\Psi_\varepsilon \overset{\mathsf{M}}{\to} \Psi_0$ *to conclude evolutionary* Γ*-convergence.*

For $\boldsymbol{X} = \mathbb{R}^2$ we consider the gradient system $(\mathbb{R}^2,\mathcal{E}_\varepsilon,\Psi_\varepsilon)$ with

$$\begin{aligned} \mathcal{E}_\varepsilon(u) &= \frac{1}{2}u\cdot\mathbb{A}_\varepsilon u = \frac{1}{2}u_1^2 + \frac{1}{2\varepsilon^2}(u_2-\varepsilon u_1)^2, \\ \Psi_\varepsilon(v) &= \frac{1}{2}\dot u\cdot\mathbb{G}_\varepsilon\dot u = \frac{1}{2}\dot u_1^2 + \frac{1}{2\varepsilon^\beta}\dot u_2^2, \end{aligned} \qquad \text{where } \mathbb{A}_\varepsilon = \begin{pmatrix} 2 & -\frac{1}{\varepsilon} \\ -\frac{1}{\varepsilon} & \frac{1}{\varepsilon^2} \end{pmatrix},\ \mathbb{G}_\varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\varepsilon^\beta} \end{pmatrix}.$$

The solution of the gradient evolution $\mathbb{G}_\varepsilon\dot u_\varepsilon = -\mathbb{A}_\varepsilon u_\varepsilon$ with general initial conditions $u_\varepsilon^0 = (a_\varepsilon,b_\varepsilon)^\top$ can be calculated explicitly for all $\varepsilon > 0$ and all $\beta > 0$.

Before discussing the question of the types of evolutionary convergence, we study the convergence properties of the functionals $\mathcal{E}_\varepsilon$ and Ψ_ε. Because the weak and the strong topology coincide on $\mathbb{R}^2$, Γ-convergence equals Mosco convergence. In fact, we have

$$\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0 : u \mapsto \begin{cases} \frac{1}{2}u_1^2 & \text{for } u_2 = 0, \\ \infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0 : v \mapsto \begin{cases} \frac{1}{2}v_1^2 & \text{for } v_2 = 0, \\ \infty & \text{otherwise.} \end{cases}$$

We also mention that $\mathcal{E}_\varepsilon \xrightarrow{\text{pw}} \mathcal{E}_{\text{pw}} = 2\mathcal{E}_0 \gneqq \mathcal{E}_0$, but neither $\mathcal{E}_\varepsilon$ nor Ψ_ε converge continuously. Finally we consider well-preparedness of initial conditions, namely $(a_\varepsilon, b_\varepsilon) \to (a, b)$ and $\mathcal{E}_\varepsilon(a_\varepsilon, b_\varepsilon) \to \mathcal{E}_0(a, b) < \infty$, which implies

$$b = 0 \quad \text{and} \quad b_\varepsilon = \varepsilon a + o(\varepsilon) \text{ for } \varepsilon \to 0. \tag{2.17}$$

The general solutions can be calculated explicitly. With $\delta_\varepsilon = \varepsilon^{\beta-2}/2$ the matrix $\mathbb{G}_\varepsilon^{-1}\mathbb{A}_\varepsilon$ has the eigenvalues $\mu_1 = 1 + \delta_\varepsilon - \big(1+\delta_\varepsilon^2\big)^{1/2}$ and $\mu_2 = 1 + \delta_\varepsilon + \big(1+\delta_\varepsilon^2\big)^{1/2}$, and we obtain

$$u_\varepsilon(t) = \tfrac{\varepsilon(\mu_2-2)a+b}{\varepsilon(\mu_2-\mu_1)}\, \mathrm{e}^{-\mu_1 t} \begin{pmatrix} 1 \\ \varepsilon(2-\mu_1) \end{pmatrix} + \tfrac{b+\varepsilon(\mu_1-2)a}{\varepsilon(\mu_2-\mu_1)}\, \mathrm{e}^{-\mu_2 t} \begin{pmatrix} -1 \\ \varepsilon(\mu_2-2) \end{pmatrix}.$$

In the case $\beta \in]0, 2[$ we have $\delta_\varepsilon \to \infty$, $\mu_1 \to 1$, $\mu_2 \to \infty$, and hence for $t > 0$ we have $\mathrm{e}^{-\mu_2 t} \to 0$ faster than $O(\varepsilon^k)$ for all $k \in \mathbb{N}$. This fact leads to the convergences

$$u_\varepsilon(t) \to a\, \mathrm{e}^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \frac{a^2}{2}\, \mathrm{e}^{-2t}.$$

Thus, we conclude $(\mathbb{R}^2, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \xrightarrow{\mathsf{E}} (\mathbb{R}^2, \mathcal{E}_0, \Psi_0)$ for $\beta \in]0, 2[$.

For $\beta > 2$ we have $\delta_\varepsilon \to 0$, $\mu_1 \to 0$, and $\mu_2 \to 2$. For $b_\varepsilon \to b \neq 0$ the solutions are unbounded and $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \infty$ for $t \in [0, T_b[$ for some $T(b) > 0$. So we consider the case $b_\varepsilon/\varepsilon \to \widehat{b}$ for some $\widehat{b} \in \mathbb{R}$ and find the limits

$$u_\varepsilon(t) \to \Big(\frac{\widehat{b}}{2} + \big(a - \frac{\widehat{b}}{2}\big)\, \mathrm{e}^{-2t}\Big) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \frac{\widehat{b}^2}{4} + \big(a - \frac{\widehat{b}}{2}\big)^2 \mathrm{e}^{-4t}.$$

Note that well-prepared initial conditions require $\widehat{b} = a$, while the case $\widehat{b} = 0$ implies semigroup convergence to the solutions of the gradient system $(\mathbb{R}^2, \mathcal{E}_{\text{pw}}, \Psi_0)$, but we do not have evolutionary Γ-convergence.

For the crossover case $\beta = 2$ the assumption $b_\varepsilon/\varepsilon \to \widehat{b}$ leads to

$$u_\varepsilon(t) \to \begin{pmatrix} w(t) \\ 0 \end{pmatrix} \quad \text{with } w(t) = \frac{1}{2\sqrt{5}} \Big(\big(\sqrt{5}-1\big)a + 2\widehat{b}\big) \mathrm{e}^{-\overline{\mu}_1 t} + \big((\sqrt{5}+1)a - 2\widehat{b}\big) \mathrm{e}^{-\overline{\mu}_2 t} \Big),$$

where $\overline{\mu}_{1,2} = (3 \pm \sqrt{5})/2$. Assuming $\widehat{b} = 0$ we obtain the linear evolution $w(t) = B(t)w(0)$, where $\dot{B}(t)/B(t)$ is not constant. Hence, the limit evolution cannot be described by an autonomous equation $\dot{w} = \boldsymbol{V}(w)$, so we do not have semigroup convergence.

Remark 2.11 (Scaling changes evolutionary convergence) *We note that a rescaling of the variables may change the convergence behavior dramatically. E.g. using* $z = (z_1, z_2) = (u_1, u_2/\varepsilon)$, *we obtain*

$$\widetilde{\mathcal{E}}_\varepsilon(z) = \mathcal{F}(z) := \frac{1}{2}z_1^2 + \frac{1}{2}(z_2 - z_1)^2 \quad \textit{and} \quad \widetilde{\Psi}_\varepsilon(\dot{z}) = \frac{1}{2}\dot{z}_1^2 + \frac{\varepsilon^{2-\beta}}{2}\dot{z}_2^2,$$

where $\widetilde{\mathcal{E}}_\varepsilon$ is independent of ε and hence continuously converging to $\mathcal{F}$. For $\beta > 2$ we obtain $(\mathbb{R}^2, \mathcal{F}, \Psi_\varepsilon) \xrightarrow{\mathsf{E}} (\mathbb{R}^2, \mathcal{F}, \Psi_0)$.

The alternative scaling $w = (u_1, \varepsilon^{-\beta/2} u_2)$ makes the dissipation distance ε-independent:

$$\overline{\mathcal{E}}_\varepsilon(w) = \frac{1}{2} w_1^2 + \frac{1}{2}\big(w_1 - \varepsilon^{\beta/2-1} w_2\big)^2 \quad \text{and} \quad \overline{\Psi}(\dot{w}) = \frac{1}{2}|\dot{w}|^2.$$

For $\beta \geq 2$ we have the continuous convergence $\overline{\mathcal{E}}_\varepsilon \xrightarrow{\mathsf{C}} \overline{\mathcal{E}}_0 : w \mapsto w_1^2$, while for $\beta \in]0,2[$ we have $\overline{\mathcal{E}}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$. In all cases we have $(\mathbb{R}^2, \overline{\mathcal{E}}_\varepsilon, \overline{\Psi}) \xrightarrow{\mathsf{E}} (\mathbb{R}^2, \overline{\mathcal{E}}_0, \overline{\Psi})$ cf. Theorem 3.3.

An important conclusion of this remark is that linear transformations, here scalings, may change the convergence behavior dramatically.

2.6 Homogenization of a 1D parabolic equation

Our main guiding example for the different levels of the theory will be the homogenization of a simple parabolic equation, since for this case it is possible to do all the calculation by hand. We emphasize that the theory works equally well for any space dimension, of course then the calculations are no longer explicit.

On the domain $\Omega =]0, \ell[$ for $t > 0$ we consider the parabolic equation

$$c(\tfrac{x}{\varepsilon})\dot{u}(t,x) = \big(a(\tfrac{x}{\varepsilon})u_x(t,x)\big)_x - b(\tfrac{x}{\varepsilon})u(t,x), \qquad u_x(t,0) = 0 = u_x(t,\ell). \tag{2.18}$$

Here the coefficient functions $a, b, c \in \mathrm{L}^\infty(\mathbb{R})$ are 1-periodic and are bounded from below by a positive constant $c_0 > 0$. Choosing $\boldsymbol{X} = \mathrm{L}^2(\Omega)$ this equation is induced by the gradient system $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ with

$$\mathcal{E}_\varepsilon(u) = \tfrac{1}{2}\textstyle\int_\Omega a(\frac{x}{\varepsilon})u_x(x)^2 + b(\frac{x}{\varepsilon})u(x)^2 \,\mathrm{d}x \quad \text{and} \quad \Psi_\varepsilon(v) = \tfrac{1}{2}\textstyle\int_\Omega c(\frac{x}{\varepsilon})v(x)^2 \,\mathrm{d}x, \tag{2.19}$$

if the derivatives of Ψ_ε and $\mathcal{E}_\varepsilon$ are calculated in the duality pairing of $\boldsymbol{X} = \mathrm{L}^2(\Omega)$.

We are interested in the question of evolutionary Γ-convergence and expect that in the limit $\varepsilon \to 0$ we find effective functionals

$$\mathcal{E}_{\text{eff}}(u) = \frac{1}{2}\int_\Omega a_{\text{eff}}\, u_x(x)^2 + b_{\text{eff}}\, u^2 \,\mathrm{d}x \quad \text{and} \quad \Psi_{\text{eff}}(v) = \frac{1}{2}\int_\Omega c_{\text{eff}}\, v^2 \,\mathrm{d}x,$$

where the main task remains to determine the effective coefficients a_{eff}, b_{eff}, and c_{eff}. Moreover, when taking the Γ-limits of $\mathcal{E}_\varepsilon$ and Ψ_ε in (2.19) there is major issue in choosing a suitable topology.

The natural function space for the dissipation functionals Ψ_ε is $\boldsymbol{X} = \mathrm{L}^2(\Omega)$, which we will call the *dissipation space*. According to Lemma 2.6 we have $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_{\text{harm}}$ and $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_{\text{arith}}$ in $\mathrm{L}^2(\Omega)$, where $c_{\text{eff}} = c_{\text{harm}}$ and $c_{\text{eff}} = c_{\text{arith}}$, respectively.

Similarly, we can study the energy functionals in their natural space $\boldsymbol{Z} = \mathrm{H}^1(\Omega)$, i.e. the *energy space*. Using the compact embedding of $\mathrm{H}^1(\Omega) \Subset \mathrm{L}^2(\Omega)$, we find

$$\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_{\text{harm}} : u \mapsto \frac{1}{2}\int_\Omega a_{\text{harm}} u_x^2 + b_{\text{arith}} u^2 \,\mathrm{d}x \quad \text{in } \boldsymbol{Z} = \mathrm{H}^1(\Omega),$$
$$\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_{\text{arith}} : u \mapsto \frac{1}{2}\int_\Omega a_{\text{arith}}\, u_x^2 + b_{\text{arith}} u^2 \,\mathrm{d}x \quad \text{in } \boldsymbol{Z} = \mathrm{H}^1(\Omega).$$

For later use we prepare the following result which has an obvious extension to general elliptic homogenization problems, see [Dal93, Ch. 24+25] and [Bra06, Sect. 5].

Topology	a_{eff}	b_{eff}	c_{eff}
L^2 weak	a_{harm}	b_{arith}	c_{harm}
L^2 strong	a_{harm}	b_{arith}	c_{arith}
H^1 weak	a_{harm}	b_{arith}	c_{arith}
H^1 strong	a_{arith}	b_{arith}	c_{arith}

Table 1: Possible effective coefficients in 1D homogenization

Proposition 2.12 *Consider the functionals $\mathcal{E}_\varepsilon : \mathrm{L}^2(\Omega) \to [0,\infty]$ as in (2.19), which are set to ∞ for $u \in \mathrm{L}^2(\Omega) \setminus \mathrm{H}^1(\Omega)$. Then, $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_{harm}$ in $\mathrm{L}^2(\Omega)$.*

Proof: For the liminf estimate in the weak L^2 topology consider $u_\varepsilon \rightharpoonup u$ in $\mathrm{L}^2(\Omega)$ with $\alpha := \liminf \mathcal{E}_\varepsilon(u_\varepsilon)$. For $\alpha = \infty$ there is nothing to be shown. For $\alpha < \infty$ the uniform coercivity of $\mathcal{E}_\varepsilon$ in $\mathrm{H}^1(\Omega)$ implies $u_\varepsilon \rightharpoonup u$ in H^1. Now using $u_\varepsilon \to u$ and $\nabla u_\varepsilon \rightharpoonup \nabla u$ in $\mathrm{L}^2(\Omega)$ together with Lemma 2.6 give $\mathcal{E}_{\mathrm{harm}}(u) \le \liminf \mathcal{E}_\varepsilon(u_\varepsilon)$ as desired.

For the limsup estimate in the strong L^2 topology we construct a recovery sequence that converges weakly in $\mathrm{H}^1(\Omega)$ and hence strongly in $\mathrm{L}^2(\Omega)$. For given $\widehat{u}$ we simply define $\widehat{u}_\varepsilon$ such that $\nabla \widehat{u}_\varepsilon(x) = a_{\mathrm{harm}} \nabla \widehat{u}(x)/a(\frac{1}{\varepsilon}x)$ and $\widehat{u}_\varepsilon(0) = \widehat{u}(0)$. ∎

We leave the question concerning the evolutionary Γ-convergence with the proper effective coefficients open until Sections 3.2 and 4.2. At this moment we observe that we have several choices for the topology that lead to different Γ-limits, see Table 1. *This fact is our motivation to derive mathematical theories giving us guidance for choosing the correct topology via suitable compatibility conditions of the two convergences.*

2.7 Tartar's model equation

In [Tar89, Tar90] a more general version of following example is considered:

$$\dot{u}_\varepsilon(t,x) = -a_\varepsilon(x) u_\varepsilon(t,x) + f(t,x), \qquad t > 0, \quad x \in \Omega, \qquad u_\varepsilon(0,x) = u_\varepsilon^0(x). \tag{2.20}$$

It is shown there that for $u_\varepsilon^0 = 0$ and $(a_\varepsilon)_{\varepsilon \in (0,1)}$ bounded in $\mathrm{L}^\infty(\Omega)$ one can pass to the nonlocal limit equation $\dot{u}_0(t,x) = -a_0(x) u_0(t,x) + f(t,x) + \int_0^t K^0(x, t-s) u_0(s,x) \, \mathrm{d}s$, where the memory kernel K^0 is determined from the sequence a_ε alone (via its Young measure).

This model serves three purposes: it shows that (i) simple Γ-convergence in the weak or the strong topology is not sufficient for evolutionary convergence, (ii) that there can be quite different gradient structures, and (iii) that we may obtain different effective equations when doing pE-convergence for different gradient structures, see Corollary 3.8.

(i) The simplest gradient structure is the one in $\mathrm{L}^2(\Omega)$, namely $(\mathrm{L}^2(\Omega), \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ with

$$\mathcal{E}_\varepsilon(u) = \int_\Omega \frac{a_\varepsilon}{2} u^2 \, \mathrm{d}x \quad \text{and} \quad \Psi_\varepsilon(v) = \int_\Omega \frac{1}{2} v^2 \, \mathrm{d}x,$$

Assuming that $a_\varepsilon(x) = a(\frac{1}{\varepsilon}x)$ with a nontrivial 1-periodic $a \in \mathrm{L}^\infty(\mathbb{R})$, we have $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{E}} \mathcal{E}_{\mathrm{arith}}$ and $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{E}} \mathcal{E}_{\mathrm{harm}}$ but no Mosco convergence, see Lemma 2.6. Since the explicit solution is given by $u_\varepsilon(t,x) = \mathrm{e}^{-a_\varepsilon(x)t} u_\varepsilon(0,x)$ we see that evolutionary Γ-convergence does not hold (not even semigroup convergence).

In [Bra13, Ex. 7.2.5] this gradient structure for Tartar's model was studied via time-incremental minimization, i.e. for a fixed time step τ one solves iteratively

$$u_k^\tau = \arg\min\{\, \mathcal{E}_\varepsilon(u) + \tfrac{1}{\tau}\Psi_\varepsilon(u-u_{k-1}^\tau) \mid u \in \mathrm{L}^2(\Omega)\,\}.$$

Taking first the limit $\varepsilon \to 0$ for fixed τ and then the limit $\tau \to 0$, the convergence to the effective equation $\dot u = -a_{\mathrm{harm}} u$.

(ii) Under the assumption $u^0(x) > 0$, we can introduce further gradient structures by considering u as the density of a measure on Ω. Choosing $\boldsymbol{X} := \boldsymbol{M}_+(\Omega)$, the set of nonnegative measures on Ω, we define the functionals

$$\widetilde{\mathcal{E}}_\varepsilon(u) = \int_\Omega b_\varepsilon(x)\,\mathrm{d}u(x) \quad \text{and} \quad \widetilde{\mathcal{R}}_\varepsilon^*(u,\xi) = \frac{1}{2}\int_\Omega \frac{a_\varepsilon}{b_\varepsilon}\xi^2\,\mathrm{d}u(x),$$

where now a_ε and b_ε are assumed to be positive continuous functions on $\overline{\Omega}$. The dual dissipation potential $\widetilde{\mathcal{R}}_\varepsilon^*$ defines the state-dependent Onsager operator

$$\widetilde{\mathbb{K}}_\varepsilon(u)\xi := \mathrm{D}_\xi \mathcal{R}_\varepsilon^*(u,\xi) = \frac{a_\varepsilon}{b_\varepsilon}\,u\,\xi.$$

Using $\mathrm{D}\widetilde{\mathcal{E}}_\varepsilon(u) = b_\varepsilon$ we easily find the linear gradient flow $\dot u = -\widetilde{\mathbb{K}}_\varepsilon(u)\mathrm{D}\widetilde{\mathcal{E}}_\varepsilon(u) = -a_\varepsilon u$.

(iii) In Section 3.5.2 we will discuss pE-convergence and highlight that different choices of b_ε, which do not change the gradient flow equation, lead to different Γ-limits. In particular, we will emphasize the role of the well-preparedness $\widetilde{\mathcal{E}}_\varepsilon(u_\varepsilon(0)) \to \widetilde{\mathcal{E}}_0(u(0))$.

2.8 Further examples of gradient systems

Here we discuss some additional gradient systems, which highlight that generally both, the energy and the dissipation potential, may depend on the small parameter ε.

2.8.1 Interfaces as limits of thin bulk layers

This example shows that E-convergence can be used for dimension reduction, in particular for deriving new interface models. We consider a bulk material in $\Omega =]-1,1[$ where a layer of width ε around 0 is filled with a material of having quite different material coefficients, that scale suitably with ε. The parabolic equation is given by

$$\gamma_\varepsilon(x)\dot u = \big(A_\varepsilon(x)u_x\big)_x - \partial_u F(x,u) \text{ in } \Omega, \quad u_x(t,-1) = 0 = u_x(t,1).$$

For the coefficients we assume the scalings

$$(\gamma_\varepsilon(x), A_\varepsilon(x)) = \begin{cases} (c\,,\alpha) & \text{for } |x| > \varepsilon/2, \\ (\rho/\varepsilon\,,\beta\varepsilon) & \text{for } |x| < \varepsilon/2. \end{cases}$$

For $\mathcal{E}_\varepsilon(u) = \int_\Omega \frac{A_\varepsilon}{2}u_x^2 + F_\varepsilon(x,u)\,\mathrm{d}x$ and $\Psi_\varepsilon(v) = \int_\Omega \frac{\gamma_\varepsilon}{2}v^2\,\mathrm{d}x$ we find the Γ-limits

$$\mathcal{E}_0(u) = \int_{\Omega_*} \frac{\alpha}{2}u_x^2 + F(x,u)\,\mathrm{d}x + \frac{\beta}{2}\big(u(0^-)-u(0^+)\big)^2 \text{ and } \Psi_0(v) = \int_{\Omega_*}\frac{c}{2}v^2\,\mathrm{d}x + \frac{\rho}{2}v(0)^2,$$

where $\Omega_* =]-1,0[\,\cup\,]0,1[$. Hence, the limit functionals consist of two integral terms as well as a point contribution at the interface $x = 0$. These extra terms at $x = 0$ determine the effective properties of the limiting sharp interface. We refer to [Lie12a, Lie12b, GlM13] and Section 3.5.1 for more details in this direction.

2.8.2 Ginzburg-Landau vortices

This gradient system is included for historical reasons: the fundamental paper [SaS04] is the first that develops a method for using the gradient structure for establishing evolutionary Γ*-convergence*, see Section 3.3 for the abstract approach and [Bra13, Ch. 10] for a short survey. The gradient system $(\mathrm{L}^2(\Omega), \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ is given by Ginzburg-Landau functional for the order parameter $\psi : \Omega \to \mathbb{C}$ in the form

$$\mathcal{E}_\varepsilon(\psi) = \int_\Omega \frac{1}{2}|\nabla\psi|^2 + \frac{1}{\varepsilon^2}(1-|\psi|^2)^2 \,\mathrm{d}x - n\pi\,|\log\varepsilon|$$

and $\Psi_\varepsilon(v) = \frac{1}{2|\log\varepsilon|}\|v\|_2^2$. Well-prepared solutions $\psi_\varepsilon(t)$ are then well-approximated by simple vortices with positions $x_i(t) \in \mathbb{R}^2$, $i = 1, \dots, n$. The associated evolutionary Γ-limit induces an ODE for the evolution of the vortex positions $x_i(t)$.

2.8.3 Slow and fast chemical reactions

This example describes a gradient system where the energy $\mathcal{E}$ *is independent of* ε*, but* $\mathcal{R}_\varepsilon$ *is state-dependent and strongly dependent on* ε*.* In [Mie11b] it is shown that general reaction systems for concentrations $c_1, ..., c_I > 0$ with R reactions of mass-action type can be written as a gradient system on $\mathcal{Q} =]0, \infty[^I$ if the reactions satisfy the detailed-balance condition for some concentration vector $\boldsymbol{w} = (w_1, \dots, w_I) \in \mathcal{Q}$. We then have

$$\dot{\boldsymbol{c}} = -\sum_{r=1}^R \kappa_r \Big(\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}} - \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}} \Big) \Big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r \Big) = -\mathbb{K}(\boldsymbol{c})\mathrm{D}\mathcal{E}(\boldsymbol{c}),$$

where the vectors $\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r \in \mathbb{N}_0^I$ contain the stoichiometric coefficients for the rth reaction and $\boldsymbol{c}^\gamma$ denotes the monomial $\prod_1^I c_i^{\gamma_i}$. The gradient structure $(\mathcal{Q}, \mathcal{E}, \mathcal{R}_\varepsilon)$ is given by

$$\mathcal{E}(\boldsymbol{c}) = \sum_{i=1}^I \big(c_i \log(c_i/w_i) - c_i + w_i \big) \text{ and } \mathcal{R}_\varepsilon^*(\boldsymbol{c}, \boldsymbol{\xi}) = \frac{1}{2}\sum_{r=1}^R \kappa_r \Lambda\big(\tfrac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}}, \tfrac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}\big) \Big(\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \cdot \boldsymbol{\xi} \Big)^2,$$

where $\Lambda(a, b) = (a-b)/(\log a - \log b) > 0$. This defines $\mathbb{K}_\varepsilon$ via $\mathcal{R}_\varepsilon^*(\boldsymbol{c}, \boldsymbol{\xi}) = \frac{1}{2}\langle \mathbb{K}_\varepsilon(\boldsymbol{c})\boldsymbol{\xi}, \boldsymbol{\xi}\rangle$.

Often the reaction coefficients κ_r have quite different magnitudes, i.e. some reactions are very fast while others are slow. Assuming $k_r = \rho_r/\varepsilon$ for $r = 1, \dots, R_{\mathrm{fast}}$ and $\kappa_r = O(1)$ for $r > R_{\mathrm{fast}}$ leads to a decomposition $\mathcal{R}_\varepsilon = \frac{1}{\varepsilon}\mathcal{R}_{\mathrm{fast}} + \mathcal{R}_{\mathrm{slow}}$ and a Mosco limit $\mathcal{R}_0^*(\boldsymbol{c}, \boldsymbol{\xi}) = \mathcal{R}_{\mathrm{slow}}(\boldsymbol{c}, \boldsymbol{\xi})$ if $\mathcal{R}_{\mathrm{fast}}^*(\boldsymbol{c}, \boldsymbol{\xi}) = 0$ and $+\infty$ otherwise. In particular, this is a case where the limit gradient system is only defined on a nonlinear submanifold of $\mathcal{Q}$, namely the one defined by $\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}} = \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}$ for $r = 1, \dots, R_{\mathrm{fast}}$.

3 pE-convergence via the energy-dissipation balance

In this section we discuss formulations that are based on the differential formulations involving derivatives such as $\dot{u}(t)$, $\mathrm{D}\mathcal{E}$, $\mathrm{D}_{\dot{u}}\mathcal{R}(u, \dot{u})$, or $\mathrm{D}_\xi\mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))$. In Section 4 we will see that under certain structural assumptions such relations can be replaced by evolutionary variational inequalities only involving $u(t)$, $\mathcal{E}(w)$, and a dissipation distance $\mathcal{D}(u(t), w)$, i.e. derivatives are not required.

3.1 Four equivalent formulations via Legendre transform

We now convert the formal modeling ideas in Section 2.1 into exact mathematical statements by working in a reflexive Banach space $\boldsymbol{X}$. We consider generalized gradient systems $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$ with time-dependent energy functionals $\mathcal{E} : [0,T]\times\boldsymbol{X} \to \mathbb{R}_\infty$. Then, by Fenchel's equivalence in Proposition 2.7 the following three formulations are equivalent, but have different physical interpretations. They are formulated in terms of the dissipation potential $\mathcal{R}$, the dual dissipation potential $\mathcal{R}^*$, or in terms of an extremum principle involving both:

Force balance in $\boldsymbol{X}^*$ Biot's equation [Bio55]

$$\text{(FB)}\quad 0 \in \partial_{\dot u}\mathcal{R}(u(t),\dot u(t)) + \mathrm{D}\mathcal{E}(t,u(t)) \ \in \boldsymbol{X}^* \text{ for a.a. } t\in[0,T].$$

Rate equation in $\boldsymbol{X}$ Onsager's equation [Ons31]

$$\text{(RE)}\quad \dot u(t) \in \partial_{\xi}\mathcal{R}^*(u(t),-\mathrm{D}\mathcal{E}(t,u(t))) \ \in \boldsymbol{X} \text{ for a.a. } t\in[0,T].$$

Power balance in $\mathbb{R}$ De Giorgi's $(\mathcal{R},\mathcal{R}^*)$ formulation [DMT80]

$$\text{(PB)}\quad \mathcal{R}(u(t),\dot u(t)) + \mathcal{R}^*(u(t),-\mathrm{D}\mathcal{E}(t,u(t))) = -\langle \mathrm{D}\mathcal{E}(t,u(t)),\dot u(t)\rangle.$$

Before returning to the general situation, we highlight the three different cases for the classical viscous dissipation, i.e. $\mathcal{R}(u,v) = \frac12\langle \mathbb{G}v, v\rangle$ and $\mathcal{R}^*(u,\xi) = \frac12\langle \xi, \mathbb{K}\xi\rangle$ with $\mathbb{K} = \mathbb{G}^{-1}$. Then, we have

$$\begin{aligned}&\text{(FB)}\quad \mathbb{G}\dot u = -\mathrm{D}\mathcal{E}(u) \qquad\qquad \text{(RE) } \dot u = -\mathbb{K}\mathrm{D}\mathcal{E}(u) = -\nabla_{\mathbb{G}}\mathcal{E}(u)\\ &\text{(PB)}\quad \frac12\langle \mathbb{G}\dot u,\dot u\rangle + \frac12\langle \mathrm{D}\mathcal{E}(u), \mathbb{K}\mathrm{D}\mathcal{E}(u)\rangle = -\langle \mathrm{D}\mathcal{E}(u),\dot u\rangle,\end{aligned}$$

where (RE) can be seen as a "gradient evolution", because $\nabla_{\mathbb{G}}$ is the gradient operator.

The Young-Fenchel estimate (2.12) states that $\geq$ in (PB) always holds. Hence, in limit passages it will be sufficient to control the opposite estimate. Moreover, it is advantageous to use the integrated form. For this we employ a version of the chain rule. Indeed, it will be sufficient to have a chain-rule estimate in terms of a suitable notion of a set-valued subdifferential $\widetilde\partial\mathcal{E}$ for the functional $\mathcal{E}$, see (3.2) for the Fréchet subdifferential or [RoS06] for a general theory. As before $\widetilde\partial\mathcal{E}(t,u)$ denotes a differential for the function $\mathcal{E}(t,\cdot) : \boldsymbol{X} \to \mathbb{R}_\infty$ for fixed t, while $\partial_t\mathcal{E}(t,u)$ denotes the partial derivative $\lim_{h\to0}\frac1h\big(\mathcal{E}(t+h,u)-\mathcal{E}(t,u)\big)$, which is always assumed to exist.

Definition 3.1 (Abstract chain rule) *We say that the triple $(\boldsymbol{X},\mathcal{E},\widetilde\partial\mathcal{E})$ satisfies the* chain rule *if for all $p\geq1$ the following holds. If $u\in \mathrm{W}^{1,p}([0,T];\boldsymbol{X})$ and $\xi\in \mathrm{L}^{p^*}([0,T];\boldsymbol{X}^*)$ with $\xi(t)\in\widetilde\partial\mathcal{E}(t,u(t))$ a.e. in $[0,T]$, then $t\mapsto\mathcal{E}(t,u(t))$ is absolutely continuous and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t,u(t)) = \langle \xi(t),\dot u(t)\rangle + \partial_t\mathcal{E}(t,u(t)) \text{ a.e. in } [0,T]. \tag{3.1}$$

We refer to [RoS06, MRS13b] for a general treatments of such abstract chain rule. In particular, the chain rule holds for λ-convex functionals, i.e. $u\mapsto\mathcal{E}(t,u)-\frac{\lambda}{2}\|u\|_{\boldsymbol{X}}^2$ if $\widetilde\partial\mathcal{E}$ denotes the Fréchet subdifferential

$$\widetilde\partial\mathcal{E}(u) := \{\,\xi\in\boldsymbol{X}^* \mid \mathcal{E}(u{+}w) \geq \mathcal{E}(u) + \langle\xi,w\rangle + o(\|w\|) \text{ for } w\to0\,\}. \tag{3.2}$$

Theorem 3.2 (Upper Energy-Dissipation Estimate) *Assume that* $(\boldsymbol{X}, \mathcal{E}, \widetilde{\partial}\mathcal{E})$ *satisfies the chain rule* (3.1). *Further assume that* $u \in \mathrm{W}^{1,p}([0,T];\boldsymbol{X})$ *and* $\xi \in \mathrm{L}^{p^*}([0,T];\boldsymbol{X}^*)$ *with* $\xi(t) \in \widetilde{\partial}\mathcal{E}(t,u(t))$ *a.e. in* $[0,T]$. *Then, u solves (FB), (RE), or (PB) if and only if the* Upper Energy-Dissipation Estimate *(UEDE) holds:*

$$\textit{UEDE:}\quad \mathcal{E}(T,u(T)) + \int_0^T \mathcal{R}(u(t),\dot{u}(t)) + \mathcal{R}^*(u(t),-\xi(t))\,\mathrm{d}t \le \mathcal{E}(0,u(0)) + \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t.$$

Then, the UEDE is equivalent to the energy-dissipation balance *(EDB)*

$$\textit{EDB:}\quad \mathcal{E}(T,u(T)) + \int_0^T \big(\mathcal{R}(u(t),\dot{u}(t)) + \mathcal{R}^*(u(t),-\xi(t))\,\mathrm{d}t = \mathcal{E}(0,u(0)) + \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t.$$

Proof: By the Fenchel equivalence (FB), (RE), and (PB) are equivalent, where now $\mathrm{D}\mathcal{E}(t,u(t))$ is replaced by $\xi(t)$. Integrating (PB) and using the chain rule (3.1) provide the EDB and hence the UEDE. For the opposite implication we start from the Young-Fenchel estimate (2.12) and obtain the following chain of estimates:

$$\begin{aligned}
&\textstyle\int_0^T -\langle \xi(t),\dot{u}(t)\rangle\,\mathrm{d}t \overset{\text{YF}}{\le} \int_0^T \mathcal{R}(u,\dot{u}) + \mathcal{R}^*(u,-\xi)\,\mathrm{d}t\\
&\overset{\text{UEDE}}{\le} \mathcal{E}(0,u(0)) + \textstyle\int_0^T \partial_t\mathcal{E}(t,u)\,\mathrm{d}t - \mathcal{E}(T,u(T)) \;= \int_0^T \partial_t\mathcal{E}(t,u) - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t,u)\,\mathrm{d}t\\
&\overset{\text{CR}}{=} \textstyle\int_0^T -\langle \xi(t),\dot{u}(t)\rangle\,\mathrm{d}t.
\end{aligned}$$

Thus, we conclude that all the above estimates must be equalities. Moreover, using the Young-Fenchel estimate again with the first estimate $\overset{\text{YF}}{\le}$ being an equality, we see that (PB) has to hold a.e. in $[0,T]$. For the last conclusion we simply use that $g(t) \le h(t)$ and $\int_0^T g\,\mathrm{d}t = \int_0^T h\,\mathrm{d}t$ imply that $g(t) = h(t)$ a.e. Hence, the proof is complete. ∎

We remark that the UEDE relates the final energy $\mathcal{E}(T,u(T))$ plus the dissipated energy $\int_0^T \mathcal{R}+\mathcal{R}^*\,\mathrm{d}t$ to the initial energy $\mathcal{E}(0,u(0))$ plus the external work $\int_0^T \partial_s\mathcal{E}(s,u)\,\mathrm{d}s$, which arises via time-dependence of the system. The main importance of the UEDE is that the final and the dissipated energies only need to have a good upper bound. Hence, in passing to a Γ-limit it will be sufficient to have good liminf estimates for these terms, while the right-hand side can be controlled by the well-preparedness of the initial conditions and proper assumptions on the work of the external forces. In fact, we will mostly assume $\mathcal{E}(t,u) = \mathcal{F}(u) - \langle \ell(t),u\rangle$, so it will be easy to control the linear term $-\int_0^T \langle \dot{\ell},u\rangle\,\mathrm{d}t = \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t$.

3.2 pE-convergence obtained from Mosco convergence

We now provide a first general result on pE-convergence for an ε-dependent family of gradient systems $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ that is based on the UEDE

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon,\dot{u}_\varepsilon) + \mathcal{R}_\varepsilon^*(u_\varepsilon,-\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))\,\mathrm{d}t \le \mathcal{E}_\varepsilon(u_\varepsilon(0)).$$

The proof presented here is a simplified version of the Mosco-convergence theory developed in [MRS13b]. In fact, the results there (cf. Thm. 4.8) are more general, while we present a

version which displays the general structure more clearly by restricting to the case that $\mathcal{E}_\varepsilon$ is independent of time and that $\mathcal{R}_\varepsilon$ is translation invariant, i.e. $\mathcal{R}_\varepsilon(u,v) = \Psi_\varepsilon(v)$. Thus, the restricted results corresponds to [Ste08, Thm. 7.2] for variational doubly nonlinear equations with convex energies $\mathcal{E}(t,\cdot)$, which can be reformulated via a Brézis-Ekeland principle, i.e. u solves $0 \in \partial\Psi(\dot u) + \mathrm{D}\mathcal{F}(u) - \ell(t)$ if and only if there exists $\xi(t) \in \partial\mathcal{E}(t,u(t)$ such that (u,ξ) minimizes $\mathcal{I}$ with $I(u,\xi) = 0$, where $\mathcal{I}$ is given by

$$\begin{aligned}\mathcal{I}(u,\xi) = \left(\int_0^T \Psi(\dot u) + \Psi^*(-\xi) - \partial_t \mathcal{E}(t,u)\,\mathrm{d}t + \mathcal{E}(T,u(t)) - \mathcal{E}(0,u_0) \right)^+ \\ + \int_0^T \mathcal{E}(t,u) + \mathcal{E}^*(t,\xi) - \langle \xi, u\rangle\,\mathrm{d}t + \|u(0)-u_0\|_X^2,\end{aligned}$$

where $(a)^+ = \max a, 0$ and $\mathcal{E}^*(,\cdot)$ is the Fenchel-Legendre transform of the convex function $\mathcal{E}(t,\cdot)$. A more restrictive result for Hilbert spaces $\boldsymbol{X}$ with the constant dissipation potential $\Psi_\varepsilon(v) = \frac12\|v\|_{\boldsymbol{X}}^2$ has already been obtained in [Att84, Thm. 3.74].

Theorem 3.3 (Mosco convergence implies pE-convergence) *Let $\boldsymbol{X}$ be a reflexive Banach space and let $c, C, \lambda_{\mathrm{c}} > 0$, $p > 1$ be such that $\mathcal{E}_\varepsilon(\cdot) + \lambda_{\mathrm{c}}\|\cdot\|_{\boldsymbol{X}}^2$ is convex and that the dissipation potentials satisfy the coercivity assumption*

$$\forall\, v \in \boldsymbol{X},\ \xi \in \boldsymbol{X}^* : \quad \Psi_\varepsilon(v) \geq c\|v\|_{\boldsymbol{X}}^p - C,\ \Psi_\varepsilon^*(\xi) \geq c\|\xi\|_{\boldsymbol{X}^*}^{p^*} - C. \tag{3.3}$$

Moreover, assume that $\boldsymbol{Z}$ is compactly embedded into $\boldsymbol{X}$ such that the energies are uniformly $\boldsymbol{Z}$-coercive, cf. (2.8)*. Then, we have*

$$\big(\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0 \text{ and } \Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0 \text{ in } \boldsymbol{X}\big) \quad \Longrightarrow \quad (\boldsymbol{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \xrightarrow{\mathsf{pE}} (\boldsymbol{X}, \mathcal{E}_0, \Psi_0).$$

In this result the convergence of the two functionals is rather strong, namely Mosco convergence. At first glance, there seems to be no compatibility condition between the two convergences. However, the necessary compatibility reduces to the fact that the Mosco convergences $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ and $\Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0$ occur in the same topology, namely in the dynamic function space $\boldsymbol{X}$. Here the uniform coercivity of Ψ_ε and Ψ_ε^* is crucial.

The proof relies on two important arguments, namely a lower semicontinuity result of Ioffe type (cf. [Iof77]) and the strong-weak closedness of the graphs of $\widetilde{\partial}\mathcal{E}_\varepsilon$.

Proposition 3.4 (Lower semicontinuity, e.g. [MRS13b, Thm. A.2]) *Assume that all $\mathcal{J}_\varepsilon : \boldsymbol{X} \to \mathbb{R}_\infty$ are proper, lsc, and convex and that $\mathcal{J}_\varepsilon \xrightarrow{\Gamma} \mathcal{J}_0$ in $\boldsymbol{X}$, then*

$$w_\varepsilon \rightharpoonup w \in \mathrm{L}^p([0,T],\boldsymbol{X}) \quad \Longrightarrow \quad \textstyle\int_0^T \mathcal{J}_0(w(t))\,\mathrm{d}t \leq \liminf_{\varepsilon\to 0} \int_0^T \mathcal{J}_\varepsilon(w_\varepsilon(t))\,\mathrm{d}t.$$

Definition 3.5 (Strong-weak closedness) *We say that the triples $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \widetilde{\partial}\mathcal{E}_\varepsilon)_{\varepsilon\in[0,1]}$ satisfies the* strong-weak closedness of the graph of $\widetilde{\partial}\mathcal{E}$*, if the following holds:*

$$\textit{If } \left\{ \begin{array}{c} u_\varepsilon \to u \text{ in } \boldsymbol{X},\ \mathcal{E}_\varepsilon(u_\varepsilon) \to e_0 \in \mathbb{R}, \\ \xi_\varepsilon \in \widetilde{\partial}\mathcal{E}_\varepsilon(u_\varepsilon),\ \xi_\varepsilon \rightharpoonup \xi \text{ in } \boldsymbol{X}^*, \end{array} \right\} \textit{ then } \xi \in \widetilde{\partial}\mathcal{E}_0(u) \text{ and } \mathcal{E}_0(u) = e_0. \tag{3.4}$$

Proposition 2.9 shows that strong-weak closedness holds if all $\mathcal{E}_\varepsilon$ are lsc and convex and the convex subdifferential is used. The assumption of convexity of $\mathcal{E}_\varepsilon(\cdot) + \lambda_c\|\cdot\|_{\boldsymbol{X}}^2$, which is also called uniform λ_c convexity, provides the same result for the Fréchet subdifferential defined in (3.2), see [RoS06, MRS13b].

Sketch of proof of Theorem 3.3: We only give the main arguments of the proof and refer to [MRS13b, Thm. 4.8] for the details. We start from a family of solutions $u_\varepsilon : [0,T] \to \boldsymbol{X}$ of $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ and want to show that an accumulation point u is a solution of the limit system $(\boldsymbol{X}, \mathcal{E}_0, \Psi_0)$.

Step 1: Using the well-preparedness of the initial condition we have $\mathcal{E}_\varepsilon(u_\varepsilon(0)) \le C$. Since the solutions u_ε satisfy the EDB, we obtain the uniform bounds $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \le C$, $\int_0^T \Psi_\varepsilon(\dot{u}_\varepsilon)\,\mathrm{d}t \le C$, and $\int_0^T \Psi_\varepsilon^*(-\xi_\varepsilon)\,\mathrm{d}t \le C$, where $\xi_\varepsilon(t) \in \widetilde{\partial}\mathcal{E}(u_\varepsilon(t))$ a.e. The coercivity of $\mathcal{E}_\varepsilon$, Ψ_ε, and Ψ_ε^* implies the uniform a priori bounds:

$$\|u_\varepsilon\|_{\mathrm{L}^\infty([0,T];\boldsymbol{Z})} + \|u_\varepsilon\|_{\mathrm{W}^{1,p}([0,T];\boldsymbol{X})} + \|\xi_\varepsilon\|_{\mathrm{L}^p([0,T];\boldsymbol{X}^*)} \le C.$$

Step 2: Using the reflexivity of $\boldsymbol{X}$ and the compact embedding of $\boldsymbol{Z}$ into $\boldsymbol{X}$, we can extract a convergent subsequence (not relabeled) giving

$$u_\varepsilon \rightharpoonup u \text{ in } \mathrm{W}^{1,p}([0,T];\boldsymbol{X}), \quad \xi_\varepsilon \rightharpoonup \xi \text{ in } \mathrm{L}^p([0,T];\boldsymbol{X}^*), \quad \forall\, t \in [0,T] :\ u_\varepsilon(t) \rightharpoonup u(t) \text{ in } \boldsymbol{Z}.$$

Step 3: Following (2.13), $\Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0$ provides the weak Γ-convergences $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_0$ and $\Psi_\varepsilon^* \xrightarrow{\Gamma} \Psi_0^*$. Hence we can apply Proposition 3.4 to obtain lower semicontinuity of the dissipation:

$$\int_0^T \Psi_0(\dot{u}) + \Psi_0^*(\xi)\,\mathrm{d}t \le \liminf_{\varepsilon\to 0} \int_0^T \Psi_\varepsilon(\dot{u}_\varepsilon) + \Psi_\varepsilon^*(\xi_\varepsilon)\,\mathrm{d}t.$$

Step 4: Next we use the strong-weak closedness to conclude that the two limit functions u and ξ are related. Using a Banach-space valued Young measure ν generated by a suitable subsequence of (ξ_ε), the strong-weak closedness implies that for a.a. $t \in [0,T]$ the support of $\nu(t)$ is concentrated in $\widetilde{\partial}\mathcal{E}_0(u(t))$ a.e. in $[0,T]$. Assuming additionally that $\widetilde{\partial}\mathcal{E}_0$ is single-valued (see the proof of [MRS13b, Thm. 4.4] for the general case) we conclude $\xi_\varepsilon(t) \to \xi(t)$ and hence $\xi(t) \in \widetilde{\partial}\mathcal{E}_0(u(t))$ a.e.

Step 5: We now pass to the limit $\varepsilon \to 0$ in the EDB. Note that the right-hand side converges because of the well-preparedness of the initial data. From $u_\varepsilon(T) \to u(T)$ in $\boldsymbol{X}$ and $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ we find $\mathcal{E}_0(u(T)) \le \liminf_{\varepsilon\to 0} \mathcal{E}_\varepsilon(u_\varepsilon(T))$. Now using the weak convergence of u_ε and ξ_ε we can apply the lsc result of Proposition 3.4 to conclude the UEDE

$$\mathcal{E}_0(u(T)) + \int_0^T \Psi_0(\dot{u}) + \Psi_0^*(\xi)\,\mathrm{d}t \ \le\ \mathcal{E}_0(u(0)) \quad \text{with } \xi(t) \in \widetilde{\partial}\mathcal{E}_0(u(t)),$$

where we used the well-preparedness of the initial data to obtain $\mathcal{E}_0(u(0))$ on the left-hand side. Since the uniform λ-convexity is preserved under Mosco convergence, the limit system $(\boldsymbol{X}, \mathcal{E}_0, \Psi_0)$ satisfies the chain rule, and Theorem 3.2 is applicable. Thus, u is a solution of $(\boldsymbol{X}, \mathcal{E}_0, \Psi_0)$.

Step 6: It remains to show the energy convergence $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u(t))$ for all t. For this we can repeat the derivation of the above UEDE on the interval $[0,t]$ where the energy term and the dissipation term satisfy a liminf estimate separately. However, using the chain rule estimate on this interval, we know that the UEDE is in fact an equality and we conclude that in both liminf estimates we must have a limit:

$$\mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u(t)) \quad \text{and} \quad \int_0^t \Psi_\varepsilon(\dot{u}_\varepsilon(t)) + \Psi_\varepsilon^*(\xi_\varepsilon(t))\,\mathrm{d}t \to \int_0^t \Psi_0(\dot{u}(t)) + \Psi_0^*(\xi(t))\,\mathrm{d}t$$

for all t. Thus, pE-convergence is established. ∎

We now discuss the applicability of the above result to our major examples. Note that the dissipation potentials Ψ_ε determine the choice of the dynamic space $\boldsymbol{X}$, while the energy space $\boldsymbol{Z}$ will be determined by $\mathcal{E}_\varepsilon$.

ODE model on $\boldsymbol{X} = \mathbb{R}^2$ from Section 2.5: Here $\boldsymbol{X} = \mathbb{R}^2$ is the only possible choice, and because in finite dimensions weak and strong convergence coincide we have $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ and $\Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0$. However, $\Psi_\varepsilon(v) = \frac{1}{2}(v_1^2 + v_2^2/\varepsilon^\beta)$ gives the dual dissipation potential $\Psi_\varepsilon^*(\xi) = \frac{1}{2}(\xi_1^2 + \varepsilon^\beta \xi_2^2)$. Thus, we have uniform coercivity of Ψ_ε and Ψ_ε^* only for the case $\beta = 0$, where Theorem 3.3 is applicable.

Moreover, the uniform-coercivity assumption guides us to a proper choice for the rescaling. Using $w = (u_1, \varepsilon^{-\beta/2})$ we obtain the gradient system $(\mathbb{R}^2, \overline{\mathcal{E}}_\varepsilon, \overline{\Psi})$ with

$$\overline{\mathcal{E}}_\varepsilon(w) = \frac{1}{2} w_1^2 + \frac{1}{2}\big(w_1 - \varepsilon^{\beta/2-1} w_2\big)^2 \quad \text{and} \quad \overline{\Psi}(\dot{w}) = \frac{1}{2}|\dot{w}|^2.$$

Obviously, we have uniform coercivity of $\overline{\Psi}$ and $\overline{\Psi}^*$, while uniform coercivity of $\overline{\mathcal{E}}_\varepsilon$ only holds for $\beta \le 2$. Thus, Theorem 3.3 is only applicable to $(\mathbb{R}^2, \overline{\mathcal{E}}_\varepsilon, \overline{\Psi})$ for $\beta \ge 2$. However, the same strategy of proof shows that pE-convergence also holds for $\beta \in [0, 1[$.

1D homogenization from Section 2.6: We have $\Omega =]0, \ell[$, $a_\varepsilon(x) = a(x/\varepsilon)$, $b_\varepsilon(x) = b(x/\varepsilon)$, and $c_\varepsilon(x) = c(x/\varepsilon)$ with upper and lower positive bounds. Since the dissipation potentials are given via $\Psi_\varepsilon(v) = \frac{1}{2}\int_\Omega c_\varepsilon v^2 \,\mathrm{d}x$ and $\Psi_\varepsilon^*(\xi) = \frac{1}{2}\int_\Omega 1/c_\varepsilon\, \xi^2 \,\mathrm{d}x$, we have uniform coercivity of the dissipation in $\boldsymbol{X} = \mathrm{L}^2(\Omega)$. The energies $\mathcal{E}_\varepsilon(u) = \frac{1}{2}\int_0^\ell a_\varepsilon u_x^2 + b_\varepsilon u^2 \,\mathrm{d}x$ are convex and uniformly coercive in $\boldsymbol{Z} = \mathrm{H}^1(\Omega)$, which is compactly embedded in $\boldsymbol{X}$.

By Proposition 2.12 we know $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_{\mathrm{harm}}$ in $\boldsymbol{X}$. Lemma 2.6 gives $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_{\mathrm{harm}}$ and $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_{\mathrm{arith}}$. Thus, we have Mosco convergence if and only if c_ε is constant. In that case, Theorem 3.3 applies, and we conclude pE-convergence, i.e. the solution of $c\dot{u} = (a_\varepsilon u_x)_x - b_\varepsilon u$ converge to the solutions of the effective equation $c\dot{u} = (a_{\mathrm{harm}} u_x)_x - b_{\mathrm{arith}} u$.

The case of an oscillating c_ε will be covered in Section 4.2.

Tartar's model from Section 2.7: We emphasize that even in the case of constant $\Psi_\varepsilon = \Psi_0$ it is not sufficient to have $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ or $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ in $\boldsymbol{X}$. This is seen by considering Tartar's example of Section 2.7 on the Hilbert space $\mathrm{L}^2(\Omega)$ with $\Psi_\varepsilon = \Psi_0 : v \mapsto \frac{1}{2}\|v\|_2^2$ and the energy functionals $\mathcal{E}_\varepsilon(u) = \int_\Omega \frac{a_\varepsilon}{2} u^2 \,\mathrm{d}x$. Then, according to Lemma 2.6 the weak and the strong Γ-limit exist and are different. However, none of them characterizes the limit dynamics.

3.3 The Sandier-Serfaty approach to pE-convergence

The approach of Sandier and Serfaty (cf. [SaS04, Ser11]) is more general than the previous result using Mosco convergence, because it directly imposes assumptions on the liminf estimates for the two dissipation terms in the EDB. For this, we now even allow that the

dissipation potentials depend on u (in a mild way), but restrict, for simplicity, to the case that the subdifferential is at most single-valued, i.e. $\partial\mathcal{E}_\varepsilon(u) = \{\mathrm{D}\mathcal{E}_\varepsilon(u)\}$ or $\emptyset$.

$$\begin{aligned}&\text{If } u_\varepsilon \to u \text{ in } \boldsymbol{X}, \quad v_\varepsilon \rightharpoonup v \text{ in } \boldsymbol{X}^*, \quad \text{and } \mathcal{E}_\varepsilon(u_\varepsilon) \le C, \text{ then}\\ &\quad \mathcal{R}_0(u,v) \le \liminf_{\varepsilon\to 0} \mathcal{R}_\varepsilon(u_\varepsilon, v_\varepsilon) \text{ and} \qquad (3.5a)\\ &\quad \mathcal{R}_0^*(u, -\mathrm{D}\mathcal{E}_0(u)) \le \liminf_{\varepsilon\to 0} \mathcal{R}_\varepsilon^*(u_\varepsilon, -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon)). \qquad (3.5b)\end{aligned}$$

Clearly, the strong-weak closedness of the graph of $(\partial\mathcal{E}_\varepsilon)_{\varepsilon\in[0,1]}$ and the Mosco convergence $\Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi_0$ imposed in Theorem 3.3 imply these two assumptions. In particular, condition (3.5b) implicitly imposes a closedness of the graph of the differential $\mathrm{D}\mathcal{E}_\varepsilon$.

Theorem 3.6 (Sandier-Serfaty [SaS04, Ser11]) *Let $\boldsymbol{X}$ and $\boldsymbol{Z}$ be reflexive Banach spaces with compact embedding $\boldsymbol{Z} \Subset \boldsymbol{X}$. Assume that $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ satisfy the chain rule* (3.1) *and that there are $c, C > 0$ and $p > 1$ such that*

$$\forall\, u, v \in \boldsymbol{X}\ \forall\, \varepsilon \in [0,1]: \quad \mathcal{R}_\varepsilon(u,v) \ge c\|v\|_{\boldsymbol{X}}^p - C, \quad \mathcal{E}_\varepsilon(u) \ge c\|u\|_{\boldsymbol{Z}} - C.$$

Then, $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ in $\boldsymbol{X}$ and (3.5) *imply $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \xrightarrow{\mathsf{pE}} (\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$.*

Proof: We follow the steps in the proof of Theorem 3.3. Consider a family u_ε of solutions with well-prepared initial conditions. The EDB and the well-preparedness of the initial conditions provide the a priori bounds $\|u_\varepsilon\|_{\mathrm{L}^\infty([0,T];\boldsymbol{Z})} + \|u_\varepsilon\|_{\mathrm{W}^{1,p}([0,T];\boldsymbol{X})} \le C$. The compact embedding $\boldsymbol{Z} \Subset \boldsymbol{X}$ allows us to extract a subsequence (not relabeled) such that

$$u_\varepsilon \rightharpoonup u \text{ in } \mathrm{W}^{1,p}([0,T];\boldsymbol{X}) \quad \text{ and } \ \forall\, t \in [0,T]: u_\varepsilon(t) \rightharpoonup u(t) \text{ in } \boldsymbol{Z}.$$

Combining a generalization of Ioffe's lsc result in Proposition 3.4, assumption (3.5a), the strong convergence $u_\varepsilon(t) \to u(t)$ in $\boldsymbol{X}$, and the weak convergence $\dot u_\varepsilon \rightharpoonup u$ in $\mathrm{L}^p([0,T];\boldsymbol{X})$ yields

$$\textstyle\int_0^T \mathcal{R}_0(u(t), \dot u(t))\,\mathrm{d}t \le \liminf_{\varepsilon\to 0} \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon(t), \dot u_\varepsilon(t))\,\mathrm{d}t.$$

Similarly, assumption (3.5b) and Fatou's lemma give

$$\textstyle\int_0^T \mathcal{R}_0^*(u(t), -\mathrm{D}\mathcal{E}_0(u(t)))\,\mathrm{d}t \le \liminf_{\varepsilon\to 0} \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon(t), -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t)))\,\mathrm{d}t.$$

Using $u_\varepsilon(T) \to u(T)$, $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ in $\boldsymbol{X}$, and the well-preparedness of the initial data yield

$$\mathcal{E}_0(u(T)) + \textstyle\int_0^T \mathcal{R}_0(u(t), \dot u(t)) + \mathcal{R}_0^*(u(t), -\mathrm{D}\mathcal{E}_0(u(t)))\,\mathrm{d}t \ \le\ \mathcal{E}_0(u(0)).$$

Since the gradient system $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$ satisfies the chain-rule estimate (3.1), the proof can be completed as for Theorem 3.3. ∎

For applications of this result in a situation where the Mosco convergence of Theorem 3.3 does not hold we refer to [SaS04, Lie12a, Mie14]. In [BB*12] this method is used to show that the solutions of $\dot u = \big(W'(u) - \varepsilon^2 u_{xx}\big)_{xx}$ converge to the solutions of the Stefan problem, which is the H^{-1}-gradient flow for the functional $\mathcal{E}_0(u) = \int_\Omega W^{**}(u)\,\mathrm{d}x$, where W^{**} is the lsc convex envelope of W.

3.4 More general results on pE-convergence using the EDB

In fact, the energy-dissipation formulation via $(\mathcal{R}, \mathcal{R}^*)$ is even more flexible. We do not even need an independent bound for each of the terms $\int_0^T \mathcal{R}_\varepsilon \,\mathrm{d}t$ and $\int_0^T \mathcal{R}_\varepsilon^* \,\mathrm{d}t$. Moreover, in passing to the liminf for the total dissipation $\int_0^T \mathcal{R}_\varepsilon + \mathcal{R}_\varepsilon^* \,\mathrm{d}t$ we may even give up the special dual form $\mathcal{R} + \mathcal{R}^*$ of the integrand. This idea, which was applied successfully in [AM*12, Mie12, DiL13], can be summarized as follows.

We define the functional $\mathcal{J}_\varepsilon : \mathrm{W}^{1,1}([0,T];\boldsymbol{X}) \to [0,\infty]$ via

$$\mathcal{J}_\varepsilon(u) := \int_0^T \mathcal{R}_\varepsilon(u,\dot u) + \mathcal{R}_\varepsilon^*(u, -\mathrm{D}\mathcal{E}_\varepsilon(u)) \,\mathrm{d}t,$$

and we have to find a sufficiently good lower bound for the Γ-liminf, namely

$$\text{(i)} \quad u_\varepsilon(\cdot) \rightharpoonup u(\cdot) \text{ in } \mathrm{W}^{1,1}([0,T];\boldsymbol{X}) \quad \Longrightarrow \quad \int_0^T \mathcal{M}_0(u(t),\dot u(t)) \,\mathrm{d}t \le \liminf_{\varepsilon\to 0} \mathcal{J}_\varepsilon(u_\varepsilon),$$

where the integrand $\mathcal{M}_0$ does not need to be of the form $\mathcal{R}_0 + \mathcal{R}_0^*$. Hence, finding the best (i.e. largest) $\mathcal{M}_0$ is nothing else than finding the (static) Γ-limit of the functionals $\mathcal{J}_\varepsilon$. It suffices to find $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$ and $\mathcal{M}_0$ such that

(ii) $\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0$;
(iii) $\mathcal{M}_0(u,v) \ge -\langle \mathrm{D}\mathcal{E}_0(u), v\rangle$;
(iv) chain rule (3.1) holds for$(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$;
(v) $\mathcal{M}_0(u,v) = -\langle \mathrm{D}\mathcal{E}_0(u), v\rangle \implies \mathcal{R}_0(u,v) + \mathcal{R}_0^*(u, -\mathrm{D}\mathcal{E}_0(u)) = -\langle \mathrm{D}\mathcal{E}_0(u), v\rangle$.

As before, we can start from the EDB $\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \mathcal{J}_\varepsilon(u_\varepsilon) = \mathcal{E}_\varepsilon(u_\varepsilon(0))$. Using the well-preparedness of the initial datum, (i), and (ii) we pass to the limit and obtain the UEDE

$$\mathcal{E}_0(u(T)) + \int_0^T \mathcal{M}_0(u(t),\dot u(t)) \,\mathrm{d}t \le \mathcal{E}_0(u(0)).$$

Now using the (iii) and the chain rule (iv) we find

$$\mathcal{E}_0(u(0)) \overset{\text{(iv)}}{=} \mathcal{E}_0(u(T)) - \int_0^T \langle \mathrm{D}\mathcal{E}(u(t)), \dot u(t)\rangle \,\mathrm{d}t \overset{\text{(iii)}}{\le} \mathcal{E}_0(u(T)) + \int_0^T \mathcal{M}_0(u(t),\dot u(t)) \,\mathrm{d}t \le \mathcal{E}_0(u(0)).$$

Thus, we conclude that we must have equality in (iii) for almost all $t \in [0,T]$, such that we can use (v) to conclude that u is a solution for $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{R}_0)$.

Section 3.5.3 summarizes the results of [Mie12], which show that the above strategy can even be applied to justify the passage from small viscous dissipation (i.e. $\mathcal{R}_\varepsilon(u,\cdot)$ is quadratic) to a limit problem with large rate-independent dissipation (i.e. $\mathcal{R}_0(u,\cdot)$ is positively homogeneous of degree 1, see Section 5.1), like dry friction.

3.5 Applications for pE-convergence based on the EDB

We treat three different applications of the abstract approaches described above, which rely on calculating the Γ-limit $\mathcal{J}_0$ of the functionals $\mathcal{J}_\varepsilon(u) = \int_0^T \mathcal{R}(u,\dot u) + \mathcal{R}(u, -\mathrm{D}\mathcal{E}(u)) \,\mathrm{d}t$. The first two are based on the Sandier-Serfaty approach of Section 3.3, while the third uses the generalization of Section 3.4.

3.5.1 Dimension reduction

The theory of evolutionary Γ-convergence was applied to cases of dimension reduction in [Lie12b, Lie12a]. There, the passage from a thin bulk layer to hypersurface was considered. This limit passage is relevant for the modeling of delamination via thin damage layers (cf. [MRT12]), the modeling of electronically active interfaces in thin-film photovoltaics [GlM13, Gli12], or for bio-membranes.

We treat the following simplified one-dimensional diffusion model with $\Omega =]-1,1[$:

$$u_t = (A_\varepsilon(x)u_x)_x \quad \text{for } t>0,\ x\in\Omega, \qquad u_x(t,\pm1)=0, \tag{3.6}$$

where the diffusion coefficient A_ε equals 1 for $|x|>\varepsilon/2$ and $\alpha\varepsilon$ for $|x|\le\varepsilon/2$. While the system can be written as a L^2-gradient flow with quadratic energy (see Section 2.8.1), here we use the logarithmic entropy that is physically more relevant for diffusion, see [JKO98, Ott01, Mie11b]. Thus, we consider the gradient system $(\mathcal{Q},\mathcal{E},\mathcal{R}_\varepsilon)$ with state space $\mathcal{Q}=\{\,u\in \mathrm{L}^1(\Omega)\mid u>0,\ \int_\Omega u(x)\,\mathrm{d}x=1\,\}$, ε-independent energy functional $\mathcal{E}(u)=\int_\Omega u\log u-u\,\mathrm{d}x$, and the dual dissipation potential $\mathcal{R}^*_\varepsilon(u,\xi)=\frac12\int_\Omega A_\varepsilon u\xi_x^2\,\mathrm{d}x$. Using

$$\mathbb{K}_\varepsilon(u)\xi=\mathrm{D}_\xi\mathcal{R}^*_\varepsilon(u,\xi)=-\big(uA_\varepsilon\xi_x\big)_x \quad\text{and}\quad \mathrm{D}\mathcal{E}(u)=\log u,$$

we see that $\dot u=-\mathbb{K}_\varepsilon(u)\mathrm{D}\mathcal{E}(u)=\big(uA_\varepsilon(\log u)_x\big)_x=(A_\varepsilon u_x)_x$ is the diffusion equation (3.6).

The EDB takes the form $\mathcal{E}(u_\varepsilon(T))+\mathcal{J}_\varepsilon(u_\varepsilon)=\mathcal{E}(u_\varepsilon(0))$ with

$$\mathcal{J}_\varepsilon(u)=\int_0^T\mathcal{R}_\varepsilon(u,\dot u)+\mathcal{R}^*_\varepsilon(u,-\mathrm{D}\mathcal{E}_\varepsilon(u))\,\mathrm{d}t=\int_0^T\int_\Omega\Big(\tfrac{1}{2uA_\varepsilon}\big(\textstyle\int_{-1}^x\dot u\,\mathrm{d}y\big)^2+\dfrac{A_\varepsilon}{2u}u_x^2\Big)\,\mathrm{d}x\,\mathrm{d}t.$$

We note that $\mathcal{J}_\varepsilon$ is a convex functional over the space-time domain $\Omega\times[0,T]$ depending on ε via the coefficient A_ε. The calculation of the Γ-limit of $\mathcal{J}_\varepsilon$ can be done with classical methods, see [Lie12b, Sec. 3.2], and for $u_\varepsilon\rightharpoonup u$ one finds

$$\begin{aligned}
&\mathcal{J}_0(u):=\int_0^T\mathcal{R}_0(u,\dot u)+M(u)\,\mathrm{d}t\le\liminf_{\varepsilon\to0}\mathcal{J}_\varepsilon(u_\varepsilon),\ \text{where}\\
&\mathcal{R}_0(u,v)=\int_{-1}^0\frac{1}{2u}\big(\textstyle\int_{-1}^x v\,\mathrm{d}y\big)^2\mathrm{d}x+\dfrac{(\int_0^1 v\,\mathrm{d}y)^2}{2\alpha\Lambda(u(0^-),u(0^+))}+\int_0^1\frac{1}{2u}\big(\textstyle\int_x^1 v\,\mathrm{d}y\big)^2\mathrm{d}x\ \text{ and}\\
&M(u)=\int_{-1}^0\frac{u_x^2}{2u}\,\mathrm{d}x+\frac{\alpha(u(0^-)-u(0^+))^2}{2\Lambda(u(0^-),u(0^+))}+\int_0^1\frac{u_x^2}{2u}\,\mathrm{d}x\quad\text{with }\Lambda(a,b)=\frac{a-b}{\log a-\log b}>0.
\end{aligned}$$

Since $\mathcal{E}$ does not depend on ε, we have $\mathcal{E}_0=\mathcal{E}$, and one can check $\mathcal{R}_0(u,v)+M(u)\ge-\langle\mathrm{D}\mathcal{E}(u),v\rangle=-\int_\Omega(\log u)v\,\mathrm{d}x$, which is condition (iii) for $\mathcal{M}_0$ in Section 3.4. Since v was arbitrary we have $\mathcal{R}^*_=(u,-\mathrm{D}\mathcal{E}(u))\le M(u)$ and conclude

$$\begin{aligned}
&\mathcal{E}(u(T))+\int_0^T\mathcal{R}_0(u,\dot u)+\mathcal{R}^*_0(u,-\mathrm{D}\mathcal{E}(u))\,\mathrm{d}t\ \le\ \mathcal{E}(u(T))+\mathcal{J}_0(u)\\
&\overset{\mathcal{J}_\varepsilon\xrightarrow{\Gamma}\mathcal{J}_0}{\underset{u_\varepsilon(\cdot)\rightharpoonup u(\cdot)}{\le}}\liminf_{\varepsilon\to0}\mathcal{E}(u_\varepsilon(T))+\mathcal{J}_\varepsilon(u_\varepsilon)\ \overset{\text{EDB}_\varepsilon}{=}\ \lim_{\varepsilon\to0}\mathcal{E}_\varepsilon(u_\varepsilon(0))\ \overset{\text{wellpr.}}{=}\ \mathcal{E}(u(0)).
\end{aligned}$$

Invoking the chain rule again, we find that u is a solution of the gradient system $(\mathcal{Q}, \mathcal{E}, \mathcal{R}_0)$, cf. Theorem 3.2. The associated PDE is given by classical diffusion in the two bulk parts $]-1,0[$ and $]0,1[$ and a transmission condition at the interface $x=0$:

$$\begin{aligned}
&\dot u = u_{xx} && \text{for } x \in]0,1[,\\
&0 = u_x(0^+) + \alpha(u(0^+) - u(0^-)) && \text{for } x = 0^+,\\
&0 = u_x(0^-) + \alpha(u(0^-) - u(0^+)) && \text{for } x = 0^-,\\
&\dot u = u_{xx} && \text{for } x \in]-1,0[.
\end{aligned}$$

We refer to [Lie12b, Lie12a, GlM13] for more details concerning gradient structures for reaction-diffusion systems with nontrivial interface conditions.

3.5.2 pE-convergence for Tartar's model

This example highlights the fact that a gradient structure is an additional information for a given evolution equation. In particular, choosing different gradient structures we may obtain different limiting evolution equations via evolutionary Γ*-convergence.*

We now consider a bounded open domain $\Omega \subset \mathbb{R}^d$ and the damped parametrized ODE

$$\dot u(t,x) = -a_\varepsilon(x) u(t,x), \quad u(0,x) = u^0(x) \ge 0. \tag{3.7}$$

Here we will interprete $u(t,\cdot) \in \mathrm{L}^1_+(\Omega)$ as a density of a measure and assume that the absorption coefficient a_ε has the form $a_\varepsilon(x) = \mathbb{A}(x, \frac{1}{\varepsilon}x)$, where $\mathbb{A}$ is continuous on $\overline{\Omega} \times \mathbb{R}^d$ and 1-periodic in the second argument, i.e. $\mathbb{A}(x, y+k) = \mathbb{A}(x,y)$ for all $k \in \mathbb{Z}^d$.

As indicated in Section 2.7 we use the gradient structure $(\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_\varepsilon, \mathcal{R}^{\mathrm{H}}_\varepsilon)$ with

$$\mathcal{E}_\varepsilon(x) = \int_{\partial\Omega} b_\varepsilon(x)\,\mathrm{d}u(x) \quad \text{and} \quad \mathcal{R}^{\mathrm{H}*}_\varepsilon(u,\xi) = \frac{1}{2}\int_{\overline{\Omega}} \frac{a_\varepsilon(x)}{b_\varepsilon(x)} \xi(x)^2\,\mathrm{d}u(x),$$

where b_ε is again assumed in the form $b_\varepsilon(x) = \mathbb{B}(x, \frac{1}{\varepsilon}x)$. In particular, the energy functional is linear, but the dual dissipation potential depends on u, so we are *not* using the Hilbert-space structure of $\mathrm{L}^2(\Omega)$ as in [Tar89, Tar90]. As in [MOS13, LMS14] we call $\mathcal{R}^{\mathrm{H}*}_\varepsilon$ the dual Hellinger-dissipation potential, because it gives rise to a weighted Hellinger distance between nonnegative measures.

We equip $\boldsymbol{M}_+(\overline{\Omega})$ with the weak* convergence and find the Γ-limit of $\mathcal{E}_\varepsilon$ via

$$\mathcal{E}_\varepsilon \overset{\Gamma*}{\rightharpoonup} \mathcal{E}_0 : \; u \mapsto \int_{\overline{\Omega}} b_0(x)\,\mathrm{d}u(x) \text{ with } b_0(x) := \min\{\, \mathbb{B}(x,y) \mid y \in \mathbb{R}^d \,\},$$

where $u_\varepsilon \overset{*}{\rightharpoonup} u$ denotes weak* convergence in $\boldsymbol{M}_+$, i.e. $\int \phi \mathrm{d}u_\varepsilon \to \int \phi \mathrm{d}u$ for all $\phi \in \mathrm{C}(\overline{\Omega})$. For the proof of the liminf estimate simply note $b_\varepsilon(x) \ge b_0(x)$, which implies $\mathcal{E}_\varepsilon(u) \le \mathcal{E}_0(u)$ for all u. Since $\mathcal{E}_0$ is weakly* continuous the liminf estimate follows.

For a given $\widehat u$ we have to find a recovery sequence $\widehat u_\varepsilon$ such that $\widehat u_\varepsilon \overset{*}{\rightharpoonup} \widehat u$ and $\mathcal{E}_\varepsilon(\widehat u_\varepsilon) \to \mathcal{E}_0(u)$. We do the construction as follows. For positive ε decompose Ω into finitely many sets of the form $Q_{k,\varepsilon} = \varepsilon(k + [0,1[^d) \cap \Omega$ with $k \in K_\varepsilon := \{\, k \in \mathbb{Z}^d \mid Q_{k,\varepsilon} \neq \emptyset \,\}$. Denote by $y_{k,\varepsilon}$ the point in $Q_{k,\varepsilon}$ which minimizes b_ε. Then, define the functions

$$\widehat u_\varepsilon = \textstyle\sum_{k\in K_\varepsilon} \int_{Q_{k,\varepsilon}} \mathrm{d}\widehat u(y)\; \delta_{y_{k,\varepsilon}} \quad \text{ and obtain } \widehat u_\varepsilon \overset{*}{\rightharpoonup} \widehat u.$$

Moreover, using $b_\varepsilon(y_{k,\varepsilon}) \le b_0(y_{k,\varepsilon})$ we have $\mathcal{E}_\varepsilon(\widehat{u}_\varepsilon) \le \mathcal{E}_0(\widehat{u}_\varepsilon)$ and conclude the limsup estimate.

Finally we prove the pE-convergence using the EDB. For $\varepsilon > 0$ the solutions satisfy

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \int_{\overline{\Omega}} \Big(\frac{b_\varepsilon}{2a_\varepsilon} \frac{\dot{u}_\varepsilon^2}{u_\varepsilon} + \frac{1}{2} a_\varepsilon b_\varepsilon u_\varepsilon \Big) \mathrm{d}x \, \mathrm{d}t = \mathcal{E}_\varepsilon(u_\varepsilon(0)).$$

To pass to the limit $\varepsilon \to 0$ we define the two functions

$$g_0(x) := \min\{ \mathbb{B}(x,y)/\mathbb{A}(x,y) \mid y \in \mathbb{R}^d \} \quad \text{and} \quad h_0(x) := \min\{ \mathbb{B}(x,y)\mathbb{A}(x,y) \mid y \in \mathbb{R}^d \}. \tag{3.8}$$

Using the well-preparedness of the initial condition on the right-hand side and the liminf estimates any pointwise weak* limit $u(t)$ of $u_\varepsilon(t)$ satisfies

$$\mathcal{E}_0(u(T)) + \int_0^T \mathcal{M}_0(u(t), \dot{u}(t)) \, \mathrm{d}t \le \mathcal{E}_0(u(0)) \quad \text{with } \mathcal{M}_0(u,v) = \int_{\overline{\Omega}} \frac{g_0}{2} \frac{v^2}{u} + \frac{h_0}{2} u \, \mathrm{d}x.$$

To show this, we first estimate $b_\varepsilon/a_\varepsilon \ge g_0$ and $a_\varepsilon b_\varepsilon \ge h_0$ and then use that the integrand is jointly convex in $(u, \dot{u})$ such that weak lsc follows.

To complete the proof of evolutionary Γ-convergence we have to prove $\mathcal{M}_0(u,v) \ge -\langle \mathrm{D}\mathcal{E}_0(u), v \rangle$ (cf. (iii) in Section 3.4), which means

$$\frac{g_0 v^2}{2u} + \frac{h_0 u}{2} \ge -b_0 v \quad \text{for all } u > 0 \text{ and } v \in \mathbb{R}.$$

Since this condition is equivalent to $g_0 h_0 \ge b_0^2$ and since the chain rule holds, we obtain the following result.

Proposition 3.7 *Consider the gradient systems $(\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon^{\mathrm{H}})$ for $\varepsilon > 0$ and the functions b_0, g_0, and h_0 as given in* (3.8)*. Define*

$$\mathcal{E}_0(u) := \int_{\overline{\Omega}} b_0(x) \, \mathrm{d}u(x) \quad \textit{and} \quad \mathcal{R}_0^*(u,\xi) := \int_{\overline{\Omega}} \frac{h_0(x)}{2b_0(x)^2} \, \xi(x)^2 \, \mathrm{d}u(x)$$

and assume $g_0 h_0 \ge b_0^2$, then $(\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathsf{pE}}{\rightarrow}{}^* (\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_0, \mathcal{R}_0)$ and the limiting equation reads $\dot{u} = -\frac{k_0}{b_0} u = -\frac{b_0}{g_0} u$.

We emphasize that this result is in contrast, but not in contradiction, to the results in [Tar89, Tar90]. There, the solutions $u_\varepsilon(t,\cdot)$ converge weakly in $\mathrm{L}^2(\Omega)$. However, this does not fit to our well-preparedness asking $u_\varepsilon(t) \overset{*}{\rightharpoonup} u(t)$ in $\boldsymbol{M}_+(\overline{\Omega})$ and $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u(t))$, which requires a concentration at the minimizing points of $\mathbb{B}(x,\cdot)$.

More importantly, this provides an example where for the same equation $\dot{u} = -a_\varepsilon u$ we have two different gradient structures that produce two different evolutionary Γ-limits having different effective equations.

Corollary 3.8 (Different gradient structures lead to different limit equations) *Equation* (3.7) *has the two gradient structures $(\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_\varepsilon^{(j)}, \mathcal{R}_\varepsilon^{(j)})$, $j = 1, 2$, with*

$$\mathcal{E}_\varepsilon^{(1)}(u) = \int_{\overline{\Omega}} a_\varepsilon \, \mathrm{d}u, \ \mathcal{R}_\varepsilon^{(1)\,*}(u,\xi) = \int_{\overline{\Omega}} \tfrac{1}{2}\xi^2 \, \mathrm{d}u, \quad \mathcal{E}_\varepsilon^{(2)}(u) = \int_{\overline{\Omega}} \tfrac{1}{a_\varepsilon} \, \mathrm{d}u, \ \mathcal{R}_\varepsilon^{(2)\,*}(u,\xi) = \int_{\overline{\Omega}} \tfrac{a_\varepsilon^2}{2}\xi^2 \, \mathrm{d}u.$$

Setting $a_{min}(x) := \min\{\, \mathbb{A}(x,y) \mid y \in \mathbb{R}^d \,\}$ *and* $a_{max}(x) := \max\{\, \mathbb{A}(x,y) \mid y \in \mathbb{R}^d \,\}$ *and*

$$\mathcal{E}_0^{(1)}(u) = \int_{\overline{\Omega}} a_{min}\,\mathrm{d}u,\ \mathcal{R}_0^{(1)\,*}(u,\xi) = \int_{\overline{\Omega}} \tfrac{1}{2}\xi^2\,\mathrm{d}u,\ \ \mathcal{E}_0^{(2)}(u) = \int_{\overline{\Omega}} \tfrac{1}{a_{max}}\,\mathrm{d}u,\ \mathcal{R}_0^{(2)\,*}(u,\xi) = \int_{\overline{\Omega}} \tfrac{a_{max}^2}{2}\xi^2\,\mathrm{d}u,$$

we obtain the evolutionary Γ*-convergences*

$$(\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_\varepsilon^{(j)}, \mathcal{R}_\varepsilon^{(j)}) \ \xrightarrow{\mathsf{pE}}{}^{*}\ (\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_0^{(j)}, \mathcal{R}_0^{(j)}),$$

which give rise to the two different effective limit equations

$$\begin{aligned} (\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_0^{(1)}, \mathcal{R}_0^{(1)}): &\quad \dot{u}(t,x) = -a_{min}(x)u(t,x) \\ (\boldsymbol{M}_+(\overline{\Omega}), \mathcal{E}_0^{(2)}, \mathcal{R}_0^{(2)}): &\quad \dot{u}(t,x) = -a_{max}(x)u(t,x). \end{aligned} \tag{3.9}$$

Proof: We simply apply Proposition 3.7. For $j = 1$ we choose $b_\varepsilon = a_\varepsilon$ (or $\mathbb{B} = \mathbb{A}$), then $b_0(x) = a_{\min}(x) := \min\{\, \mathbb{A}(x,y) \mid y \in \mathbb{R}^d \,\}$, $g_0 = 1$, and $h_0 = a_{\min}^2$, such that condition $g_0 h_0 \geq b_0^2$ holds. Similarly, for $j = 2$ we choose $b_\varepsilon = 1/a_\varepsilon$, then $b_0 = 1/a_{\max}$, $g_0(x) = 1/a_{\max}^2$, and $h_0 = 1$. ∎

Again, there is no contradiction between the two different limit equations. However, the choice of the energy functionals induces different conditions for the convergence of well-prepared initial conditions. Since this well-preparedness is preserved during the full evolution, we obtain truely different effective evolution equations.

The two different limit equations in (3.9) *for Tartar's model* (3.7) *show clearly, that the choice of the gradient structure for given equation is an additional modeling information that cannot be extracted from the equation alone.*

3.5.3 From viscous to rate-independent friction

In this section we discuss evolutionary Γ*-convergence, where we start form classical gradient systems* $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ *with quadratic* Ψ_ε *and obtain a rate-independent generalized gradient system* $(\boldsymbol{X}, \mathcal{E}_0, \Psi_0)$ *with* $\Psi_0(v) = \rho|v|$*, cf. Section 5.1. In particular, we emphasize that in this case* $\Psi_\varepsilon \xrightarrow{\mathsf{C}} 0$ *while* $\Psi_0 \neq 0$. To analyze the evolutionary Γ-limit via the EDB we use that the functional $\mathcal{J}_\varepsilon(u) = \int_0^T \Psi_\varepsilon(\dot{u}) + \Psi_\varepsilon^*(-\mathrm{D}\mathcal{E}_\varepsilon(u))\,\mathrm{d}t$ may lead to a more general $\mathcal{J}_0$, as was discussed in Section 3.4.

First investigations concerning the derivation of dry friction from viscous models can be found in [Pra28], see Figure 3.1. The work was re-initiated in [Jam96] and further investigated by [PuT05, MiT12] to understand hysteresis in materials. Here we summarize the analytical approach in [Mie12] based on the EDB. The main interest is that we have the E-convergence $(\mathbb{R}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \xrightarrow{\mathsf{pE}} (\mathbb{R}, \mathcal{E}_0, \Psi_0)$ where $\Psi_\varepsilon(v) = \frac{\varepsilon^\alpha}{2}v^2$ is a small viscous (i.e. quadratic) dissipation potential while the limit friction $\Psi_0(v) = \rho|v|$ is homogeneous of degree 1, giving Coulomb's rate-independent friction law $\xi = \rho\mathrm{Sign}(v)$. Moreover, the friction coefficient $\rho > 0$ will depend on the size of the wiggles in the wiggly energy landscape $\mathcal{E}_\varepsilon(t,\cdot)$.

To be more precise consider the time-dependent gradient system $(\mathbb{R}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ with

$$\mathcal{E}_\varepsilon(t,u) = \mathcal{E}_0(t,u) + W_\varepsilon(u) \quad \text{with } \mathcal{E}_0(t,u) := \frac{1}{2}u^2 - \ell(t)u \text{ and } W_\varepsilon(u) := \varepsilon\rho\cos(u/\varepsilon).$$

Here the macroscopic part is independent of ε and will be the Γ-limit of $\mathcal{E}_\varepsilon$. For $t \in \mathbb{R}$ we have continuous and Mosco convergence: $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{E}_0$ and $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$. However, the derivative

$$\xi_\varepsilon = \mathrm{D}_u\mathcal{E}_\varepsilon(t,u) = \mathrm{D}_u\mathcal{E}_0(t,u) + \mathrm{D}W_\varepsilon(u) = u - \ell(t) - \rho\sin(u/\varepsilon)$$

has fluctuations of fixed size $\rho > 0$ and period ε. Moreover, for each $t \in \mathbb{R}$ there are many steady states, namely approximately ρ/ε, see Figure 3.1. This explains the name *wiggly energy landscape*.

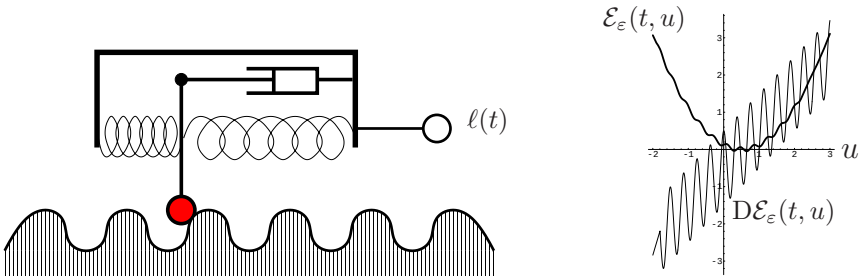


Figure 3.1: Left: Prandtl's Gedankenmodell [Pra28]: a microscopically wiggly energy landscape gives rise to macroscopic dry friction. Right: wiggly energy and its derivative.

The gradient system induces the ODE

$$\varepsilon^\alpha \dot{u} = -\mathrm{D}_u\mathcal{E}_\varepsilon(t,u) = -\big(u - \ell(t)\big) + \rho\sin(u/\varepsilon). \tag{3.10}$$

A plot of a numerical simulation is depicted in Figure 3.2, which shows that the solutions u_ε converge for $\varepsilon \to 0$ to a limit solution u that does not solve the naïve limit equation $0 = -\mathrm{D}_u\mathcal{E}_0(t,u(t))$, which might be guessed from the limits $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{C}} \mathcal{E}_0$ and $\Psi_\varepsilon \xrightarrow{\mathsf{C}} 0$.

The limit passage can be achieved by using the EDB in the form

$$\mathcal{E}_\varepsilon(T,u_\varepsilon(T)) + \int_0^T \frac{\varepsilon^\alpha}{2}\dot{u}_\varepsilon^2 + \frac{1}{2\varepsilon^\alpha}|\mathrm{D}_u\mathcal{E}_\varepsilon(t,u_\varepsilon(t))|^2\,\mathrm{d}t = \mathcal{E}_\varepsilon(0,u_\varepsilon(0)) - \int_0^T \dot{\ell}(t)u_\varepsilon(t)\,\mathrm{d}t,$$

where the last term denotes the work of the external forces. The main difficulty in the proof of evolutionary Γ-convergence is to find a liminf estimate for the dissipation integral.

Proposition 3.9 ([Mie12, Prop. 3.1]) *If $u_\varepsilon(t) \to u(t)$ and $|u_\varepsilon(t) - u_\varepsilon(s)| \le C(\varepsilon^\delta + |t-s|)$ for some $\delta > 0$. Then,*

$$\liminf_{\varepsilon\to 0} \int_0^T \tfrac{\varepsilon^\alpha}{2}\dot{u}_\varepsilon^2 + \tfrac{1}{2\varepsilon^\alpha}|\mathrm{D}_u\mathcal{E}_\varepsilon(t,u_\varepsilon(t))|^2\,\mathrm{d}t \ge \int_0^T \mathcal{M}_0(u(t),\dot{u}(t),t)\,\mathrm{d}t$$

with $\mathcal{M}_0(u,v,t) = |v|K(\ell(t)-u) + \chi_{[-\rho,\rho]}(\ell(t)-u)$ and $K(\xi) = \frac{1}{2\pi}\int_0^{2\pi}|\xi+\rho\cos y|\,\mathrm{d}y$, where $\chi_{[-\rho,\rho]}(\xi) = 0$ for $|\xi| \le \rho$ and ∞ otherwise.

Since $K(\xi) = |\xi|$ for $|\xi| \ge \rho$ and $K(\xi) \gneqq |\xi|$ for $|\xi| < \rho$ we find the desired relation (iii) in Section 3.4, viz.

$$\mathcal{M}_0(u,v,t) \ge |v|\,|\ell(t)-u| \ge -v\,\mathrm{D}\mathcal{E}_0(t,u).$$

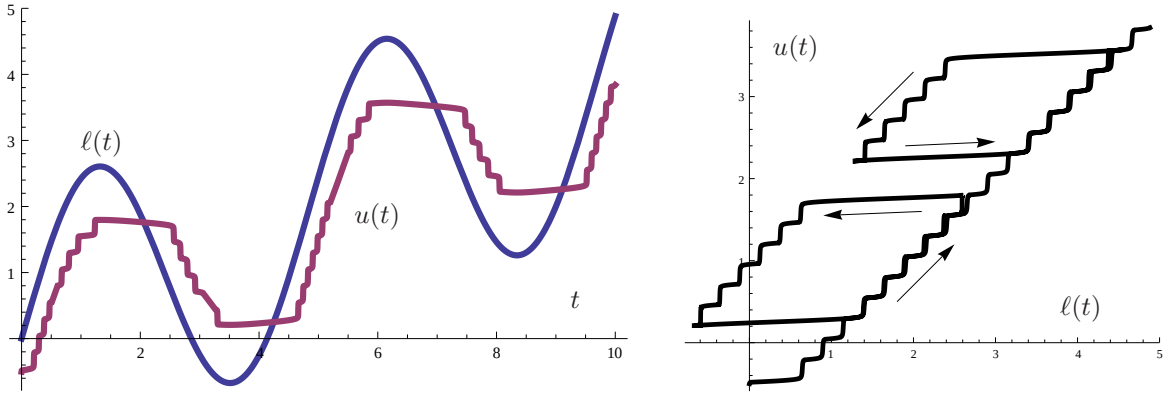


Figure 3.2: Simulation of (3.10) for $\ell(t) = 2.1\sin(1.3t) + 0.4\,t$, $\rho = \alpha = 1$, and $\varepsilon = 0.04$.

Moreover, by the chain rule (here in $\mathbb{R}^1$) the limit functions will realize equality which holds if and only if $|\mathrm{D}\mathcal{E}_0(t,u)| \le \rho$ and $\rho|v| + v\,\mathrm{D}\mathcal{E}_0(t,u) = 0$. Thus, using Lemma 5.1 the limit u of solutions u_ε satisfies the differential inclusion

$$0 \in \partial\Psi_0(\dot u) + \mathrm{D}\mathcal{E}_0(t,u) \quad \text{with } \Psi_0(v) = \rho|v|.$$

Thus, we have established the following evolutionary Γ-convergence result that connects quadratic Ψ_ε to a one-homogeneous Ψ_0 in the limit.

Theorem 3.10 ([PuT05], [Mie12, Th. 3.2]) *For the above gradient structure for* (3.10) *we have* $(\mathbb{R}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \overset{\mathsf{pE}}{\longrightarrow} (\mathbb{R}, \mathcal{E}_0, \Psi_0)$. *Moreover, along well-prepared solutions we have* $\int_0^T 2\Psi_\varepsilon(\dot u_\varepsilon(t))\,\mathrm{d}t \to \int_0^T \Psi_0(\dot u(t))\,\mathrm{d}t$.

Concerning the last convergence of the dissipation, we emphasize that the convergence only occurs after integration. The integrand $2\Psi_\varepsilon(\dot u_\varepsilon(t))$ develops many sharp peaks at distance proportional to ε that have a width proportional to $\varepsilon^{\alpha+1}$ and a height of order $\varepsilon^{-\alpha}$, each of which corresponds to a fast viscous jump from one wiggle to the next.

4 E-convergence using evolutionary variational estimates

Here we derive evolutionary Γ-convergence results based on the *integrated evolutionary variational estimate* (IEVE), which was initiated in [AGS05, Ch. 4.0] and is further developed [DaS10, Sav11]. There, the name *evolutionary variational inequality* and the abbreviation EVI is used. However, to distinguish this notion from the evolutionary (quasi-) variational inequalities (see (5.3) and e.g. [Kre99, BKS04]) we stick with our name IEVE.

The main advantage of the reformulation of a gradient system $\dot u = \mathrm{D}_\xi\mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))$ in terms of an IEVE is that the latter is derivative free (i.e. no occurrence of $\dot u$, $\mathrm{D}\mathcal{E}$, nor $\partial_\xi\mathcal{R}^*$), so we can use Γ-convergence for the functionals $\mathcal{E}_\varepsilon$ and $\mathcal{R}_\varepsilon$ more directly. However, the theory using IEVE is restricted to the case of quadratic dissipation potentials, which can be replaced by the associated dissipation distance $\mathcal{D}$, and an additional strong convexity condition, called geodesic λ-convexity, is needed as well, see Section 4.3.

We first motivate the equivalence of the IEVE and the EDB by looking at an abstract linear gradient flow. This will allow us to treat the one-dimensional homogenization problem with oscillatory coefficient $c_\varepsilon(x) = c(x/\varepsilon)$ for $\dot{u}$, see Section 4.2. Next, we discuss the general metric approach in Section 4.3, where the dissipation distance is a general geodesic distance. In Theorem 4.5 we then provide a general result showing E-convergence.

4.1 The simplest case: quadratic convex functionals

We consider the simplest case that the functionals $\mathcal{E}$ and Ψ are quadratic, i.e. a gradient systems $(\boldsymbol{H}, \mathcal{E}, \Psi)$ where $\boldsymbol{H}$ is a Hilbert space, $\mathcal{E}(u) = \frac{1}{2}\langle \mathbb{L}u, u\rangle \geq 0$, and $\Psi(v) = \frac{1}{2}\langle \mathbb{G}v, v\rangle \geq 0$ with $\mathbb{G}^* = \mathbb{G}$ and $\mathbb{L}^* = \mathbb{L}$. The associated gradient flow is the linear equation $\mathbb{G}\dot{u} = -\mathbb{L}u$.

The following formulation in terms of the IEVE is a special case of the integral solutions developed in [Bén72], which was developed for a general Banach space $\boldsymbol{X}$ and equations $\dot{u} + \mathcal{A}(u) \ni f(t)$, where $\mathcal{A}$ is a possibly multi-valued accretive operator (see [Rou05, Sect. 9.6] for more details on integral solutions).

Proposition 4.1 (IEVE in quadratic case) *For the above gradient system $(\boldsymbol{H}, \mathcal{E}, \Psi)$ the following formulations are equivalent: (FB) $\Leftrightarrow$ (RE) $\Leftrightarrow$ (EDB) $\Leftrightarrow$ (IEVE) where the* Integrated Evolutionary Variational Estimate *formulation is given by*

$$\textit{IEVE:} \qquad \begin{cases} \forall\, 0 \leq s < t \leq T\ \forall\, w \in \boldsymbol{H}: \\ \Psi(u(t){-}w) - \Psi(u(s){-}w) \leq (t{-}s)\big(\mathcal{E}(w){-}\mathcal{E}(u(t))\big). \end{cases} \tag{4.1}$$

The addition "Integrated" is used to distinguish this formulation from the differential evolutionary variational estimate (EVE) given by $\frac{\mathrm{d}}{\mathrm{d}t}\Psi(u(t){-}w) \leq \mathcal{E}(w){-}\mathcal{E}(u(t))$.

Proof: The equivalence between (FB), (RE), and (EDB) is already established in Section 3.1, so it remains to show (i) $\Leftrightarrow$ (IEVE).

"$\Rightarrow$" The quadratic form of Ψ and $\mathcal{E}$ and the linear form $\mathbb{G}\dot{u} = -\mathbb{L}u$ of (i) yield

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\Psi(u{-}w) &= \langle \mathbb{G}\dot{u}, u-w\rangle \overset{\text{(i)}}{=} -\langle \mathbb{L}u, u{-}w\rangle \\ &= \tfrac{1}{2}\langle \mathbb{L}w, w\rangle - \tfrac{1}{2}\langle \mathbb{L}u, u\rangle - \tfrac{1}{2}\langle \mathbb{L}(u{-}w), u{-}w\rangle \leq \mathcal{E}(w) - \mathcal{E}(u) - 0. \end{aligned}$$

Integration over time and the monotonicity of $\mathcal{E}$ along solutions give

$$\Psi(u(t){-}w) - \Psi(u(s){-}w) = \textstyle\int_s^t \mathcal{E}(w){-}\mathcal{E}(u(\tau))\,\mathrm{d}\tau \leq (t{-}s)\big(\mathcal{E}(w){-}\mathcal{E}(u(t))\big),$$

which is the desired IEVE.

"$\Leftarrow$" Rearrangement of quadratic expressions gives

$$\text{IEVE} \ \Leftrightarrow\ \ \forall\, w \in \boldsymbol{H}:\ \frac{1}{2}\langle \mathbb{G}(u(t){-}u(s)), u(t){+}u(s){-}2w\rangle \leq \frac{t{-}s}{2}\langle \mathbb{L}(u(t){+}w), w{-}u(t)\rangle.$$

Choosing $s = t - h$, dividing by h, and letting $h \to 0_+$, we find $\langle \mathbb{G}\dot{u}(t), u(t){-}w\rangle \leq \frac{1}{2}\langle \mathbb{L}(u(t){+}w), w{-}u(t)\rangle$ for all $w \in \boldsymbol{H}$. Setting $w = u(t){-}\delta\widehat{v}$, dividing by δ and letting $\delta \to 0_+$ gives $\langle \mathbb{G}\dot{u}(t), \widehat{v}\rangle \leq -\langle \mathbb{L}u(t), \widehat{v}\rangle$ for all $\widehat{v} \in \boldsymbol{H}$. Replacing $\widehat{v}$ by $-\widehat{v}$, we even have equality and conclude that (i) holds. ∎

This formulation leads to the following simple but new E-convergence result. For this, we introduce the domains $\mathrm{dom}(\mathcal{E}_\varepsilon) := \{\, u \in \boldsymbol{H} \mid \mathcal{E}_\varepsilon(u) < \infty \,\}$ and note that the solutions do not need $u_\varepsilon(0) \in \mathrm{dom}(\mathcal{E}_\varepsilon)$, instead $u_\varepsilon(0) \in \overline{\mathrm{dom}(\mathcal{E}_\varepsilon)}$ is sufficient.

Theorem 4.2 (E-convergence for quadratic case) *Assume that $\mathcal{E}_\varepsilon$ and Ψ_ε are quadratic functionals of a Hilbert space $\boldsymbol{H}$. Further assume that there is a Hilbert space $\boldsymbol{Z} \Subset \boldsymbol{H}$ (compact embedding) such that we have the coercivities*

$$\exists\, c_1, c_0 > 0\ \forall\, \varepsilon \in [0,1],\ u \in \boldsymbol{H}: \quad \mathcal{E}_\varepsilon(u) \geq c_0 \|u\|_{\boldsymbol{Z}}^2,\ c_0\|u\|_{\boldsymbol{H}}^2 \leq \Psi_\varepsilon(u) \leq c_1 \|u\|_{\boldsymbol{H}}^2. \quad (4.2)$$

If $\mathcal{E}_\varepsilon \overset{\mathsf{M}}{\longrightarrow} \mathcal{E}_0$ in $\boldsymbol{H}$ and $\Psi_\varepsilon \overset{\mathsf{C}}{\to} \Psi_0$ in $\boldsymbol{H}$, then $(\boldsymbol{H}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \overset{\mathsf{E}}{\to} (\boldsymbol{H}, \mathcal{E}_0, \Psi_0)$ (i.e. without well-prepared initial conditions), namely if $u_\varepsilon(0) \in \overline{\mathrm{dom}(\mathcal{E}_\varepsilon)}$ for all $\varepsilon \in [0,1]$, then

$$u_\varepsilon(0) \to u_0(0) \text{ in } \boldsymbol{H} \quad \Longrightarrow \quad \forall\, t > 0:\ u_\varepsilon(t) \rightharpoonup u_0(t) \text{ in } \boldsymbol{Z} \text{ and } \mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u_0(t)),$$

where $u_\varepsilon : [0,\infty[\to \boldsymbol{H}$ are solutions for $(\boldsymbol{H}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$.

The following proof is a simplified version of the more general proof for Theorem 4.5 in the metric setting for uniformly geodesically λ-convex gradient systems, which is given in [DaS10, Sav11]. Comparing this result with Theorem 3.3, we see that the Mosco convergence of the energies $\mathcal{E}_\varepsilon$ is the same, but there the Mosco convergence $\Psi_\varepsilon \overset{\mathsf{M}}{\longrightarrow} \Psi_0$ is much stronger than the continuous convergence $\Psi_\varepsilon \overset{\mathsf{C}}{\to} \Psi_0$ needed here.

Proof: The solutions u_ε of $\mathbb{G}_\varepsilon \dot{u} = -\mathbb{L}_\varepsilon u$ with $u_\varepsilon(0) \in \boldsymbol{H}$ satisfy

$$\text{IEVE}_\varepsilon: \qquad \Psi_\varepsilon(u_\varepsilon(t)-w) - \Psi_\varepsilon(u_\varepsilon(s)-w) \leq (t-s)\big(\mathcal{E}_\varepsilon(w) - \mathcal{E}_\varepsilon(u_\varepsilon(t))\big)$$

for $0 \leq s < t \leq T$ and $w \in \boldsymbol{H}$. Hence we find the following a priori estimates:

$$\Psi_\varepsilon(u_\varepsilon(t)) + t\mathcal{E}_\varepsilon(u_\varepsilon(t)) \leq \Psi_\varepsilon(u_\varepsilon(0)) \leq C_0, \quad (4.3a)$$

$$\mathcal{E}_\varepsilon(u_\varepsilon(t)) + \int_s^t \Psi_\varepsilon(u_\varepsilon) + \Psi_\varepsilon^*(\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))\,\mathrm{d}\tau = \mathcal{E}_\varepsilon(u_\varepsilon(s)) \leq C_0/s. \quad (4.3b)$$

$$\Psi_\varepsilon^*(\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t))) \geq \Psi_\varepsilon^*(\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(s))) \text{ for } 0 < t < s \leq T. \quad (4.3c)$$

Here (4.3a) follows from the IEVE by setting $s = 0$ and $w = 0$ using $\mathcal{E}_\varepsilon(0) = 0$, while (4.3b) is a consequence of (EDB) and (4.3a). The monotonicity (4.3c) follows from $\frac{\mathrm{d}}{\mathrm{d}t}\Psi_\varepsilon^*(\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon)) = -\langle \mathbb{L}_\varepsilon \dot{u}_\varepsilon, \dot{u}_\varepsilon\rangle \leq 0$. Choosing $s = t/2$ in (4.3b) and using (4.3c) we obtain $\Psi_\varepsilon^*(\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t)) \leq 4C_0/t^2$.

Hence, for each $t > 0$ we have a uniform bound for u_ε in the spaces $\mathrm{H}^1([t,T];\boldsymbol{H}) \subset \mathrm{C}^{1/2}([t,T];\boldsymbol{H})$ and $\mathrm{L}^\infty([t,T];\boldsymbol{Z})$. By Arzela-Ascoli's theorem we find a subsequence (not relabeled) such that $u_\varepsilon(t) \rightharpoonup U(t)$ in $\boldsymbol{Z}$ for all $t > 0$. We also set $U(0) = u_0(0)$ and prove now that $U : [0,T] \to \boldsymbol{H}$ is a solution of $\text{IEVE}_{\varepsilon=0}$ as follows.

For a given $\widehat{w}$ in $\boldsymbol{H}$ we choose a recovery sequence $\widehat{w}_\varepsilon \to \widehat{w}$ in $\boldsymbol{H}$ with $\mathcal{E}_\varepsilon(\widehat{w}_\varepsilon) \to \mathcal{E}_0(\widehat{w})$. Inserting $\widehat{w}_\varepsilon$ into IEVE_ε we can pass to the limit using $\mathcal{E}_\varepsilon \overset{\Gamma}{\to} \mathcal{E}_0$ and $\Psi_\varepsilon \overset{\mathsf{C}}{\to} \Psi_0$ and find

$$\text{IEVE}_0: \qquad \Psi_0(U(t)-\widehat{w}) - \Psi_0(U(s)-\widehat{w}) \leq (t-s)\big(\mathcal{E}_0(\widehat{w}) - \mathcal{E}_0(U(t))\big).$$

Thus, U solves IEVE_0. To conclude that it is indeed the solution u_0 of $\mathbb{G}_0\dot{u} = -\mathbb{L}_0 u$ with $u(0) = u_0(0)$ we still need to show that it is continuous at $t = 0$, i.e. $U(t) \to U(0) = u_0(0)$. Taking $s = 0$ in IEVE_0 and $\widehat{w} \in \mathrm{dom}(\mathcal{E}_0)$ and using $\mathcal{E}_0(U(t)) \geq 0$ we find

$$\limsup_{t\searrow 0} \|U(t)-\widehat{w}\|_{\boldsymbol{H}}^2 \leq \frac{1}{c_0} \limsup_{t\searrow 0} \Psi_0(U(t)-\widehat{w}) \leq \frac{1}{c_0}\Psi_0(U(0)-\widehat{w}) \leq \frac{c_1}{c_0}\|U(0)-\widehat{w}\|_{\boldsymbol{H}}^2.$$

Choosing $\widehat{w} = w_k \to U(0)$ in $\boldsymbol{H}$, where we used $U(0) \in \overline{\mathrm{dom}(\mathcal{E}_0)}$, we conclude $U(t) \to U(0)$ as desired. Hence, $U(t) = u_0(t)$ because of the uniqueness of solutions.

Finally, we show energy convergence along the solutions. From (4.3) we have derived the ε-independent bound $\|\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t))\|_{\boldsymbol{H}^*} \le C_*/t^2$, which implies

$$\mathcal{E}_\varepsilon(u) \ge \mathcal{E}_\varepsilon(u_\varepsilon(t)) - \tfrac{C_*}{t^2}\|u-u_\varepsilon(t)\|_{\boldsymbol{H}}^2 \quad \text{for all } u \in \boldsymbol{H}.$$

Fixing $t > 0$ and inserting a recovery sequence $u = \widehat{u}_\varepsilon \to u_0(t)$ in $\boldsymbol{H}$ with $\mathcal{E}_\varepsilon(\widehat{u}_\varepsilon) \to \mathcal{E}_0(u_0(t))$ into the above estimate yield $\mathcal{E}_0(u_0(t)) \ge \limsup_{\varepsilon\to 0} \mathcal{E}_\varepsilon(u_\varepsilon(t))$, because of $\|u_\varepsilon(t)-\widehat{u}_\varepsilon\|_{\boldsymbol{H}} \to 0$. Since the opposite estimate holds by $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$, we obtain $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u_0(t))$, and the proof is complete. ∎

4.2 Linear parabolic homogenization via E-convergence

Here we show that the linear parabolic homogenization problem can be handled using the IEVE. We recall the linear parabolic equation

$$c_\varepsilon \dot{u} = \mathrm{div}(a_\varepsilon \nabla u) - b_\varepsilon u \quad \text{in } \Omega, \qquad a_\varepsilon \nabla u \cdot \nu = 0 \text{ on } \partial\Omega, \tag{4.4}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain. The coefficients are $b_\varepsilon(x) = b(x/\varepsilon)$ and $c_\varepsilon(x) = c(x/\varepsilon) > 0$ and $a_\varepsilon(x) = a(x/\varepsilon) \in \mathbb{R}^{d\times d}_{\mathrm{spd}}$, where a, b, and c are 1-periodic.

The Allen-Cahn-type gradient structure is given by $\boldsymbol{H} = \mathrm{L}^2(\Omega)$ and the functionals

$$\Psi_\varepsilon(v) = \int_\Omega \frac{c_\varepsilon(x)}{2}\, v(x)^2 \,\mathrm{d}x \quad \text{and} \quad \mathcal{E}_\varepsilon(u) = \int_\Omega \frac{1}{2}\nabla u(x)\cdot a_\varepsilon(x)\nabla u(x) + \frac{b_\varepsilon(x)}{2}\, u(x)^2 \,\mathrm{d}x.$$

Now the assumptions of Theorem 4.2 are satisfied with $\boldsymbol{Z} = \mathrm{H}^1(\Omega)$. In particular, we have $\Psi_\varepsilon \xrightarrow{\mathsf{C}} \Psi_0$ and $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ in $\boldsymbol{H}$ with

$$\mathcal{E}_0(u) = \int_\Omega \frac{1}{2}\nabla u \cdot a_{\mathrm{eff}} \nabla u + \frac{b_{\mathrm{arith}}}{2}\, u^2 \,\mathrm{d}x \text{ and } \Psi_0(v) = \int_\Omega \frac{c_{\mathrm{arith}}}{2}\, v^2 \,\mathrm{d}x,$$

where in the 1D case we have $a_{\mathrm{ell}} = a_{\mathrm{harm}}$, see Lemma 2.6. For homogenization with $d \ge 2$ we refer to [Dal93, Bra06], where $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ is shown in $\mathrm{H}^1(\Omega)$, which gives $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ in $\mathrm{L}^2(\Omega)$ via Proposition 2.5.

Theorem 4.2 guarantees that solutions u_ε of (4.4) converge to solutions u_0 of the effective equation

$$c_{\mathrm{arith}} \dot{u} = \mathrm{div}\left(a_{\mathrm{eff}} \nabla u\right) - b_{\mathrm{arith}} u \text{ in } \Omega, \quad a_{\mathrm{eff}} \nabla u \cdot \nu = 0 \text{ on } \partial\Omega, \tag{4.5}$$

in the following way:

$$u_\varepsilon(0) \xrightarrow{\mathrm{L}^2} u_0(0) \quad \Longrightarrow \quad \forall\, t > 0 :\ u_\varepsilon(t) \xrightarrow{\mathrm{H}^1} u_0(t) \ \text{ and } \ \mathcal{E}_\varepsilon(u_\varepsilon(t)) \to \mathcal{E}_0(u_0(t)). \tag{4.6}$$

Again, we emphasize that neither well-preparedness nor finite energy are needed. It can be shown that the convergence for $t > 0$ implies $a_\varepsilon \nabla u_\varepsilon(t) \to a_{\mathrm{eff}} \nabla u(t)$ in $\mathrm{L}^2(\Omega;\mathbb{R}^d)$.

A second way of formulating (4.4) as a gradient system relates to the gradient structure for reaction-diffusion systems developed in [Mie11b] as a generalization of the Wasserstein

gradient structure for diffusion equations of [JKO98]. For this, we introduce the density $\rho = c_\varepsilon u$ and obtain an equation in the Onsager form (cf. Section 3.1)

$$\dot\rho = \operatorname{div}\Big(a_\varepsilon \nabla\big(\frac{\rho}{c_\varepsilon}\big)\Big) - \frac{b_\varepsilon}{c_\varepsilon}\,\rho = -\mathbb{K}_\varepsilon \mathrm{D}\mathcal{F}_\varepsilon(\rho),$$

where we introduced the Onsager operator $\mathbb{K}_\varepsilon$ and the free energy $\mathcal{F}_\varepsilon$ via

$$\mathbb{K}_\varepsilon \xi := -\operatorname{div}\big(a_\varepsilon \nabla \xi\big) + b_\varepsilon \xi \quad \text{and} \quad \mathcal{F}_\varepsilon(\rho) = \int_\Omega \frac{1}{2c_\varepsilon}\,\rho^2 \,\mathrm{d}x.$$

Here $\mathbb{K}_\varepsilon$ can be seen as the operator generating the dual dissipation potential $\mathcal{R}^*_\varepsilon$, viz.

$$\mathcal{R}^*_\varepsilon(\xi) = \frac{1}{2}\langle \xi, \mathbb{K}_\varepsilon \xi\rangle := \int_\Omega \frac{1}{2}\nabla\xi \cdot a_\varepsilon \nabla\xi + \frac{b_\varepsilon}{2}\,\xi^2 \,\mathrm{d}x.$$

Because of this form, we see that the relevant Hilbert space $\widetilde{\boldsymbol{H}}$ is defined such that $\widetilde{\boldsymbol{H}}^* = \mathrm{H}^1(\Omega)$, i.e. $\widetilde{\boldsymbol{H}}$ is a closed subspace of $\mathrm{H}^{-1}(\Omega)$. Again, Theorem 4.2 is applicable with $\widetilde{\boldsymbol{Z}} = \mathrm{L}^2(\Omega)$, which is the space for which the functionals $\mathcal{F}_\varepsilon$ are equi-coercive. Moreover, we have the convergences $\mathcal{F}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{F}_0$ in $\mathrm{L}^2(\Omega)$ and $\mathcal{R}^*_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{R}^*_0$ in $\mathrm{H}^1(\Omega)$, where

$$\mathcal{F}_0(\rho) = \int_\Omega \frac{1}{2c_{\mathrm{arith}}}\,\rho^2 \,\mathrm{d}x \quad \text{and} \quad \mathcal{R}^*_0(\xi) = \int_\Omega \frac{1}{2}\nabla\xi \cdot A_{\mathrm{eff}} \nabla\xi + \frac{b_{\mathrm{arith}}}{2}\,\xi^2 \,\mathrm{d}x.$$

By Proposition 2.5 we conclude $\mathcal{F}_\varepsilon \overset{\mathsf{M}}{\longrightarrow} \mathcal{F}_0$ in $\widetilde{\boldsymbol{H}}$, and by Theorem 2.8 and Lemma 2.3(c) we have $\mathcal{R}_\varepsilon \overset{\mathsf{C}}{\to} \mathcal{R}_0$ in $\widetilde{\boldsymbol{H}}$.

Again, Theorem 4.2 is applicable and we obtain $(\widetilde{\boldsymbol{H}}, \mathcal{F}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathsf{E}}{\to} (\widetilde{\boldsymbol{H}}, \mathcal{F}_0, \mathcal{R}_0)$. In particular, we conclude that the solutions ρ_ε converge to ρ_0 in the following sense:

$$\rho_\varepsilon(0) \overset{\mathrm{H}^{-1}}{\to} \rho_0(0) \implies \forall\, t>0:\ \rho_\varepsilon(t) \overset{\mathrm{L}^2}{\rightharpoonup} \rho_0(t) \text{ and } \mathcal{F}_\varepsilon(\rho_\varepsilon(t)) \to \mathcal{F}_\varepsilon(\rho_0(t)). \tag{4.7}$$

The latter convergences imply $u_\varepsilon(t) = \rho_\varepsilon(t)/c_\varepsilon \to u_0(t) = \rho_0(t)/c_{\mathrm{arith}}$, where we used the original definition of ρ_ε.

Hence, we see that the second gradient structure yields a stronger convergence result, because the necessary assumptions on the initial conditions are much weaker. In the original variables u_ε the convergence in (4.7) means that we need $u_\varepsilon(0)/c_\varepsilon \to u_0(0)/c_{\mathrm{arith}}$ in $\widetilde{\boldsymbol{H}} \approx \mathrm{H}^{-1}(\Omega)$ to obtain the weak convergence $u_\varepsilon(t)/c_\varepsilon \to u_0(t)/c_{\mathrm{arith}}$ as well as $u_\varepsilon(t) \to u_0(t)$ in $\mathrm{L}^2(\Omega)$. Hence, we can apply the convergence (4.6) afterwards.

4.3 Metric gradient systems and IEVE$_\lambda$

We present here a few basic facts on the general theory of evolutionary variational estimates on metric spaces. We refer to [AGS05, DaS10, Sav11] for the general theory. In the next section we then show that there is a natural approach to establish E-convergence via the IEVE.

A metric gradient system is given by a triple $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$ where $(\mathcal{Q}, \mathcal{D})$ is a complete, geodesic space, that is a complete metric space admitting constant-speed geodesic curves $\gamma : [0,1] \to \mathcal{Q}$ between every two points $u_0,\, u_1 \in \mathcal{Q}$, i.e.

$$\gamma(0) = u_0,\ \gamma(1) = u_1, \text{ and } \mathcal{D}(\gamma(s),\gamma(t)) = |s-t|\mathcal{D}(u_0,u_1) \text{ for } s,t \in [0,1].$$

The notion of geodesic spaces is motivated by classical gradient systems $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$ with a quadratic dissipation potential $\mathcal{R}(u,v) = \frac{1}{2}\langle \mathbb{G}(u)v, v\rangle$. Under suitable conditions we can define the associated *dissipation distance* via

$$\mathcal{D}_{\mathcal{R}}(u_0,u_1)^2 = \inf\{ \int_0^1 2\mathcal{R}(\gamma(s),\dot\gamma(s))\,\mathrm{d}s \mid \gamma(0)=u_0,\ \gamma(1)=u_1,\ \gamma\in \mathrm{W}^{1,1}([0,1];\boldsymbol{X})\,\},$$

and the minimizers are constant-speed geodesic curves.

For general metric gradient systems the notion of gradient flows was introduced in [DMT80], see [AGS05] for historical remarks, using the metric velocity $|u'|_{\mathcal{D}}(t)$ and the metric slope $|\partial\mathcal{E}|_{\mathcal{D}}(u)$. These notions can be seen as generalizations of the terms $\big(2\mathcal{R}(u,\dot u)\big)^{1/2}$ and $\big(2\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\big)^{1/2}$, respectively, and are defined via

$$|u'|_{\mathcal{D}}(t) = \lim_{h\to 0} \frac{\mathcal{D}(u(t+h),u(t))}{h} \quad \text{and} \quad |\partial\mathcal{E}|_{\mathcal{D}}(u) = \limsup_{w\to u} \frac{\max\{\mathcal{E}(w)-\mathcal{E}(u),0\}}{\mathcal{D}(w,u)}.$$

Both definitions do not use any linear structure on the metric space $(\mathcal{Q},\mathcal{D})$. We only take difference quotients in the real numbers. In [DMT80] metric gradient flows were introduced by generalizing the energy-dissipation balance (EDB) (1.3) (cf. Section 3.1):

$$\mathcal{E}(u(T)) + \int_0^T \frac{1}{2}|u'|_{\mathcal{D}}(t)^2 + \frac{1}{2}|\partial\mathcal{E}|_{\mathcal{D}}(u(t))^2\,\mathrm{d}t = \mathcal{E}(u(0)). \tag{4.8}$$

As in Theorem (3.2) the EDB can be replaced by an Upper Energy-Dissipation Estimate (UEDE) if we have a suitable lower chain-rule estimate, namely

$$|u'|_{\mathcal{D}}(\cdot),\ |\partial\mathcal{E}|_{\mathcal{D}}(u(\cdot)) \in \mathrm{L}^2([0,T]) \implies \left\{ \begin{array}{c} \mathcal{E}(u(\cdot)) \in \mathrm{W}^{1,1}([0,T]) \text{ and} \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) \ge -|u'|_{\mathcal{D}}(t)\,|\partial\mathcal{E}|_{\mathcal{D}}(u(t)) \text{ a.e.} \end{array} \right. \tag{4.9}$$

Solutions of the UEDE are called *curves of maximal slope*.

The general class of metric gradient systems is especially adapted to construct solutions by taking limits $\tau\to 0$ for the time-discretized (backward-Euler) minimization problems

$$u^\tau_{k+1} \in \operatorname{Arg\,min}\{\, \mathcal{E}(u) + \tfrac{1}{2\tau}\mathcal{D}(u^\tau_k,u)^2 \mid u\in\mathcal{Q}\,\} \text{ for } k=1,...,N, \tag{4.10}$$

where $\tau = T/N > 0$ is the time step. We refer to [AGS05, DaS10, MRS13b] for general surveys.

Remark 4.3 (Generalized metric gradient systems) *The above metric gradient flow corresponds to classical gradient flows in the sense that the dissipation potential $\frac{1}{2}|u'|_{\mathcal{D}}(t)^2$ is quadratic in the velocity. Considering a general convex, lsc, and monotone function $\psi : [0,\infty[\to [0,\infty]$ we obtain a generalized metric gradient system $(\mathcal{Q},\mathcal{E},\mathcal{D},\psi)$. The associated generalized metric gradient flow is defined by the convex dual $\psi^*(\beta) = \sup\{\,\nu\beta - \psi(\nu) \mid \nu\ge 0\,\}$ and the energy-dissipation balance*

$$\mathcal{E}(u(T)) + \int_0^T \psi\big(|u'|_{\mathcal{D}}(t)\big) + \psi^*\big(|\partial\mathcal{E}|_{\mathcal{D}}(u(t))\big)\,\mathrm{d}t = \mathcal{E}(u(0)).$$

The p-gradient systems are given by $\psi(\nu) = \nu^p/p$ and are studied in [AGS05]. The case of rate-independent systems is given by $\psi(\nu) = c\,\nu$ and will be studied in more detail in Section 5. Rate-independent systems with small viscosity are given by the function $\psi(\nu) = \sigma_{yield}\nu + \delta\nu^2/2$. The limit $\delta \searrow 0$ is treated in [MRS09, MRS13a].

The important point is that convexity of functionals can be generalized to geodesic convexity on geodesic spaces $(\mathcal{Q},\mathcal{D})$ as follows. A functional $\mathcal{E}:\mathcal{Q}\to\mathbb{R}_\infty$ is called *geodesically λ-convex* on $(\mathcal{Q},\mathcal{D})$ if for all u_0, $u_1\in\mathrm{dom}(\mathcal{E}):=\{\,u\in\mathcal{Q}\,|\,\mathcal{E}(u)<\infty\,\}$, there exists a constant-speed geodesic $\gamma:[0,1]\to\mathcal{Q}$ with

$$\gamma(0)=u_0,\ \gamma(1)=u_1,\ s\mapsto\mathcal{E}(\gamma(s))+\tfrac{\lambda s(1-s)}{2}\mathcal{D}(\gamma(0),\gamma(1))^2 \text{ is convex on } [0,1]. \quad (4.11)$$

The case $\lambda=0$ is simply called geodesic convexity, $\lambda>0$ improves the convexity, and $\lambda<0$ relaxes the convexity. A trivial but useful and important special case is given by the Hilbert space setting with $\mathcal{Q}=\boldsymbol{H}$ and $\Psi(v)=\frac12\langle\mathbb{G}v,v\rangle$. Then, we have $\mathcal{D}(u_0,u_1)=(2\Psi(u_1-u_0))^{1/2}$ and the constant-speed geodesics are $\gamma(s)=(1-s)u_0+s\,u_1$. Moreover, for smooth energies $\mathcal{E}$, geodesic λ-convexity is equivalent to $\mathrm{D}^2\mathcal{E}\ge\lambda\mathbb{G}$.

For general geodesically λ-convex gradient systems $(\mathcal{Q},\mathcal{E},\mathcal{D})$ there is a rather complete existence and uniqueness theory, see [AGS05, DaS10], which we summarize here: For each $u_0\in\overline{\mathrm{dom}(\mathcal{E})}$ there exists a unique solution $u:[0,\infty[\,\to\mathcal{Q}$ of (4.8) with $u(0)=u_0$. Moreover the solutions depend Lipschitz continuously on the initial data, i.e. any two solutions u_1 and u_2 satisfy

$$\mathcal{D}(u_1(t),u_2(t))\le \mathrm{e}^{-\lambda(t-s)}\mathcal{D}(u_1(s),u_2(s)) \ \text{ for } 0\le s<t.$$

Moreover, under slightly stronger assumptions it can be shown that the time-incremental solutions of (4.10) converge strongly, namely

$$\mathcal{D}(u(k\tau),u_k^\tau)\le C(u_0)\sqrt{\tau}\,\mathrm{e}^{-\lambda_\tau k\tau} \ \text{ for } k=1,...,N=T/\tau,$$

where $\lambda_\tau=\frac1\tau\log(1-\lambda\tau)$. See [MOS13] for an application in one-dimensional viscoelasticity with a true metric (not translation invariant).

The main tool for the analysis of geodesically λ-convex metric gradient systems is the reformulation in terms of the integrated evolutionary variational estimate IEVE_λ, where now Proposition 4.1 is generalized by the appearance of λ, which was chosen 0 in Section 4.1, and by replacing $\Psi(u-w)$ with $\frac12\mathcal{D}(u,w)^2$. We have the following equivalence:

Theorem 4.4 ([DaS10, Thm. 2.11]) *Assume that $(\mathcal{Q},\mathcal{E},\mathcal{D})$ is a geodesically λ-convex gradient system as introduced above. Then, every solution $u:[0,\infty[\,\to\mathcal{Q}$ of the metric EDB (4.8) is a solution of the Integrated Evolutionary Variational Estimate (IEVE)$_\lambda$:*

$$\begin{aligned}&\forall w\in\mathcal{Q}\ \forall s\ge0,\ \forall\, t>s:\\ &\frac{\mathrm{e}^{\lambda(t-s)}}{2}\mathcal{D}(u(t),w)^2-\frac12\mathcal{D}(u(s),w)^2\le \mathsf{m}_\lambda(t-s)\big(\mathcal{E}(w)-\mathcal{E}(u(t))\big),\end{aligned} \quad (4.12)$$

where $\mathsf{m}_\lambda(r)=\int_0^r\mathrm{e}^{\lambda\rho}\,\mathrm{d}\rho$. Moreover, every solution u of (IEVE)$_\lambda$ with $\mathcal{E}(u(0))<\infty$ is a solution of (4.8).

Again, we see that the IEVE is a formulation without any derivative, i.e. we do neither need the metric velocity $|u'|_{\mathcal{D}}$ nor the slope $|\partial\mathcal{E}|_{\mathcal{D}}$. Thus, it is natural to study evolutionary Γ-limits $\varepsilon\to0$ for families $(\mathcal{Q},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon)$ of metric gradient systems. The importance here is that we need to be able to choose $\lambda\in\mathbb{R}$ independent of ε, i.e. we need uniform geodesic convexity.

The crucial a priori estimate, which generalizes (4.3), is

$$\begin{aligned}&\frac{\mathrm{e}^{\lambda t}}{2}\mathcal{D}(u(t),w)^2+\mathsf{m}_\lambda(t)\mathcal{E}(u(t))+\frac{\mathsf{m}_\lambda(t)^2}{2}|\partial\mathcal{E}(u(t))|_{\mathcal{D}}^2\\&\le\frac{1}{2}\mathcal{D}(u(0),w)^2+\mathsf{m}_\lambda(t)\mathcal{E}(w)\text{ for all }t>0\text{ and }w\in\mathcal{Q},\end{aligned}\tag{4.13}$$

see [DaS10, Thm. 2.6].

4.4 E-convergence for metric gradient systems

The following result on evolutionary Γ-convergence is a slight variant of [DaS10, Thm. 2.17] or [Sav11, Thm. 6.2], since we allow $\mathcal{D}_\varepsilon$ to depend on ε, while it was assumed to be constant there. We refer to [Sav11, Thm. 7.4, Cor. 8.6] for more general results including abstract Gromov-Hausdorff convergence of metric spaces.

We first list the precise assumptions on the gradient system $(\mathcal{Q},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon)_{\varepsilon\in[0,1]}$. Note that we assume that all the metrics $\mathcal{D}_\varepsilon$ on the space $\mathcal{Q}$ are equivalent, cf. (4.14a). Hence, we can write $u_n\xrightarrow{\mathcal{D}}u$ for $\mathcal{D}_{\varepsilon_n}(u_n,u)\to0$ for $n\to\infty$, where $\varepsilon_n\in[0,1]$ can be arbitrary. However, for the geodesic convexity of $\mathcal{E}_\varepsilon$ it is crucial to consider the dissipation distance $\mathcal{D}_\varepsilon$ with the same ε, see (4.14e):

$$\exists\, c_1,c_0>0\ \forall\,\varepsilon\in[0,1],\ u_1,u_2\in\mathcal{Q}:\ c_0\mathcal{D}_0(u_1,u_2)\le\mathcal{D}_\varepsilon(u_1,u_2)\le c_1\mathcal{D}_0(u_1,u_2);\tag{4.14a}$$

$$\mathcal{D}_\varepsilon\xrightarrow{\mathsf{C}}\mathcal{D}_0,\quad\text{i.e. }u_\varepsilon^{(j)}\xrightarrow{\mathcal{D}}u^{(j)}\implies\mathcal{D}_\varepsilon(u_\varepsilon^{(1)},u_\varepsilon^{(2)})\to\mathcal{D}_0(u^{(1)},u^{(2)});\tag{4.14b}$$

$$\forall\,E>0\ \exists\,\mathcal{C}_E\text{ compact in }(\mathcal{Q},\mathcal{D}_0)\ \forall\,\varepsilon\in[0,1]:\{\,u\in\mathcal{Q}\mid\mathcal{E}_\varepsilon(u)\le E\,\}\subset\mathcal{C}_E;\tag{4.14c}$$

$$\mathcal{E}_\varepsilon\xrightarrow{\Gamma}\mathcal{E}_0\text{ in }(\mathcal{Q},\mathcal{D}_0);\tag{4.14d}$$

$$\exists\,\lambda\in\mathbb{R}\ \forall\,\varepsilon\in[0,1]:\ \mathcal{E}_\varepsilon\text{ is geodesically }\lambda\text{-convex in }(\mathcal{Q},\mathcal{D}_\varepsilon).\tag{4.14e}$$

Thus, the crucial assumptions are the uniform compactness (4.14c), the uniform geodesic λ-convexity (4.14e) as well as the convergences $\mathcal{E}_\varepsilon\xrightarrow{\Gamma}\mathcal{E}_0$ and $\mathcal{D}_\varepsilon\xrightarrow{\mathsf{C}}\mathcal{D}_0$ in $(\mathcal{Q},\mathcal{D}_0)$.

Theorem 4.5 (E-convergence for IEVE) *If the assumptions* (4.14) *hold, then* $(\mathcal{Q},\mathcal{E}_\varepsilon,\mathcal{D}_\varepsilon)\xrightarrow{\mathsf{E}}(\mathcal{Q},\mathcal{E}_0,\mathcal{D}_0)$ *(no well-preparedness of init. cond.). More precisely,*

$$\begin{aligned}&\textit{if }\ u_\varepsilon(0)\in\overline{\mathrm{dom}(\mathcal{E}_\varepsilon)}\ \textit{for}\ \varepsilon\in[0,1]\ \ \textit{and}\ \ u_\varepsilon(0)\xrightarrow{\mathcal{D}}u_0(0),\\&\textit{then}\ \ \forall\,t>0:\ u_\varepsilon(t)\to u_0(t)\ \textit{and}\ \mathcal{E}_\varepsilon(u_\varepsilon(t))\to\mathcal{E}_0(u_0(t)).\end{aligned}$$

We highlight that the whole sequence u_ε converges to u_0, because we have uniqueness of the solutions. Even though well-preparedness of the initial conditions is not needed, we obtain energy convergence for all $t>0$.

Sketch of proof: For the full proof we refer to the above references. We just highlight the main steps, which are exactly the same as in the proof of Theorem 4.2.

Based on the a priori estimate (4.13) evaluated for $t\in[t_0,T]$ with arbitrary $0<t_0<T$ and the uniform compactness (4.14c) we can extract a subsequence (not relabeled) with $u_\varepsilon(t)\xrightarrow{\mathcal{D}}U(t)$. Choosing a recovery sequence $\widehat{w}_\varepsilon\xrightarrow{\mathcal{D}}\widehat{w}$ with $\mathcal{E}_\varepsilon(\widehat{w}_\varepsilon)\to\mathcal{E}_0(\widehat{w})$ we can pass to the limit $\varepsilon\to0$ in $(\mathrm{IEVE})_\lambda^\varepsilon$ and find that U is a solution of the limit equation $(\mathrm{IEVE})_\lambda^0$.

Continuity at $t = 0$ and energy convergence then follow as for Theorem 4.2. ∎

For an application of this theory we refer to [GiM13], where the heat equation on the torus $\mathbb{T}_d = \mathbb{R}^d/\mathbb{Z}^d$ is considered as generated by the the Kantorovich-Wasserstein gradient system $(\mathrm{Prob}(\mathbb{T}_d), \mathcal{E}_0, \mathcal{D}_0)$ with the entropy functional $\mathcal{E}_0(\rho) = \int_{\mathbb{T}_d} \rho \log \rho \,\mathrm{d}x$ and $\mathcal{D}_0 = \mathcal{D}_{\mathrm{Wass}}$. It is obtained as evolutionary Γ-limit of discrete Markov chains in the periodic d-dimensional lattice $(\mathbb{Z}/N\mathbb{Z})^d$ for $N \to \infty$.

An application to the justification of amplitude equations is given in [Mie14]. In fact, this application is a Hilbert-space case, but nevertheless using the approach via the IEVE improves the results considerably. The original equation is the ε-dependent, fourth-order parabolic Swift-Hohenberg equation

$$\dot{u} = -\frac{1}{\varepsilon^2}\big(1+\varepsilon^2\partial_x^2\big)^2 u + \mu u - u^3 \text{ on } \mathbb{S} := \mathbb{R}/2\pi\mathbb{Z}, \tag{4.15}$$

which is the L^2-gradient flow of the functional $\mathcal{F}_\varepsilon^{\mathrm{SH}}(u) = \int_{\mathbb{S}} \frac{1}{2\varepsilon^2}(u+\varepsilon^2 u_{xx})^2 - \frac{\mu}{2}u^2 + \frac{1}{4}u^4 \,\mathrm{d}x$. Because of the linear operator the typical solutions (e.g. well-prepared solutions) will spatially oscillate on the scale ε and are approximately of the form $u(t,x) \approx \mathrm{Re}\,\big(A(t,x)\mathrm{e}^{\mathrm{i}x/\varepsilon}\big)$. Using Theorem 4.5 it is possible to justify that the Ginzburg-Landau equation

$$\dot{A} = 4A_{xx} + \mu A - \frac{3}{4}|A|^2 A$$

is the evolutionary Γ-limit. Using a bijection $\mathbb{M}_\varepsilon$, which satisfies $u = \mathrm{Re}\,\big((\mathbb{M}_\varepsilon u)\mathrm{e}^{\mathrm{i}x/\varepsilon}\big)$, it is shown in [Mie14, Thm. 2.3] that the solutions u_ε of the Swift-Hohenberg equation converge to solutions A of Ginzburg-Landau equation:

$$\mathbb{M}_\varepsilon u_\varepsilon(0) \overset{\mathrm{L}^2(\mathbb{S})}{\to} A(0) \quad \Longrightarrow \quad \forall\, t > 0 :\ \mathbb{M}_\varepsilon u_\varepsilon(t) \overset{\mathrm{H}^1(\mathbb{S})}{\rightharpoonup} A(t) \text{ and } \mathcal{F}_\varepsilon^{\mathrm{SH}}(u_\varepsilon(t)) \to \mathcal{F}^{\mathrm{GL}}(A(t)).$$

The last application of the IEVE shows that even in the case of a Hilbert space with a quadratic and translation-invariant dissipation potential the metric concepts are extremely helpful and give a new look to semilinear parabolic equations, in particular concerning question of evolutionary Γ*-convergence.*

5 Rate-independent systems (RIS)

We again consider generalized gradient systems $(\boldsymbol{X}, \mathcal{E}, \mathcal{R})$ on a Banach space $\boldsymbol{X}$, but now we focus to the special case that the dissipation potential $\mathcal{R}$ is positively homogeneous of degree 1, i.e.

$$\forall\, u,\, v \in \boldsymbol{X}\ \forall\, \gamma > 0 : \quad \mathcal{R}(u, \gamma v) = \gamma^1 \mathcal{R}(u, v). \tag{5.1}$$

In particular, $\mathcal{R}(u, \cdot)$ cannot be smooth, but using convexity we still have a subdifferential $\partial_v \mathcal{R}(u, v)$ that is set-valued now. In particular, the relation between the rate $v = \dot{u}$ and the dissipative forces $\xi \in \partial\mathcal{R}(u, \dot{u})$ is positively homogeneous of degree 0, i.e. $\partial_v \mathcal{R}(u, \gamma v) = \gamma^0 \partial_v \mathcal{R}(u, v)$, which explains the name "rate independence" of the dissipative constitutive law. Rate-independent systems are ideal to describe hysteretic effects and occur in the case of Coulomb friction, in plasticity, magnetism, or phase transformations.

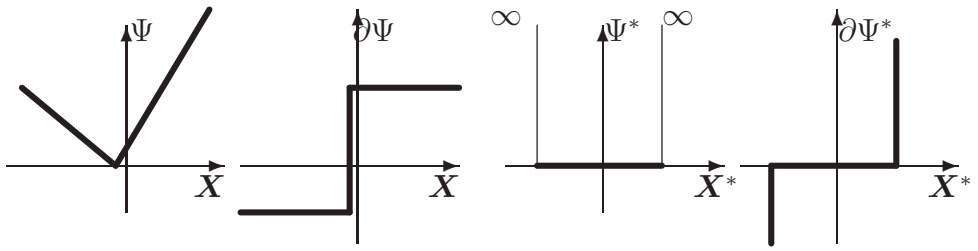


Figure 5.1: Primal and dual dissipation potential for RIS.

5.1 Subdifferential formulation of RIS

The evolution equation in Biot's form (cf. (2.1)) is the subdifferential inclusion

$$0 \in \partial_{\dot u}\mathcal{R}(u(t),\dot u(t)) + \mathrm{D}_u\mathcal{E}(t,u(t)) \in \boldsymbol{X}^*, \tag{5.2}$$

where we now included a time-dependence into the energy functional $\mathcal{E}$, as in Section 2.1. This is essential for rate-independent systems because they do not have any internal time scale any more. So they do not move without an external time-dependent stimulus.

As before, Biot's equation (5.2) is only one of the three equivalent formulations of the problem. However, in the rate-independent case the 1-homogeneity of the dissipation induces a very particular structure, see also Figure 5.1.

Lemma 5.1 *Assume that $\Psi : \boldsymbol{X} \to [0,\infty]$ is convex, lsc, and 1-homogeneous, and set $K := \partial\Psi(0)$. Then,*

$$\xi \in \partial\Psi(v) \iff \left\{ \begin{array}{c} \xi \in K \\ \langle \xi, v\rangle = \Psi(v) \end{array} \right. \qquad \text{and} \quad \Psi^*(\xi) = \left\{ \begin{array}{ll} 0 & \text{if } \xi \in K, \\ \infty & \text{else.} \end{array} \right.$$

To understand the rate equation (often called flow rule) $\dot u \in \partial_\eta\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(t,u))$, we introduce the closed and convex set $K(u) := \partial_{\dot u}\mathcal{R}(u,0) \subset \boldsymbol{X}^*$ and its outer normal cone

$$\mathrm{N}_{K(u)}(\eta) := \{\, v \in \boldsymbol{X} \mid \forall\, \xi \in K(u) :\ \langle \xi - \eta, v\rangle \ge 0 \,\},$$

then we have $\partial_\eta\mathcal{R}^*(u,-\eta) = -\mathrm{N}_{K(u)}(\eta)$ (cf. [Roc70]) and find the rate equation

$$\dot u(t) \in \partial_\eta\mathcal{R}^*\big(u(t), -\mathrm{D}\mathcal{E}(t,u(t))\big), \quad \text{or} \quad -\dot u(t) \in \mathrm{N}_{K(u(t))}(\mathrm{D}\mathcal{E}(t,u(t))), \tag{5.3}$$

which is also called an evolutionary quasi-variational inequality in [Kre99, BKS04, Mie05].

For treating more general systems the energy-dissipation balance (2.5) is more relevant. However, for RIS the dual dissipation potential $\mathcal{R}^*$ does not give a contribution to the balance, because it only attains the values 0 or $+\infty$. The latter value leads to a constraint. Thus, the EDB takes the form

$$\forall\, t \in [0,T] : -\mathrm{D}\mathcal{E}(t,u(t)) \in K(u(t)) = \partial_{\dot u}\mathcal{R}(u(t),0), \tag{5.4a}$$

$$\mathcal{E}(T,u(T)) + \int_0^T \mathcal{R}(u,\dot u)\,\mathrm{d}t = \mathcal{E}(0,u(0)) + \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t. \tag{5.4b}$$

Here (5.4a) is a totally static *local stability* condition saying that the potential restoring force is balanced by one of the possible dissipative friction forces, while (5.4b) is a reduced energy balance, where the last integral is the work of the external forces.

5.2 Energetic solutions of RIS

Since the dissipation $\int_0^T \mathcal{R}(u(t), \dot{u}(t))\,\mathrm{d}t$ controls the BV-norm of u with respect to the time only, in general cases the solutions u will not be absolutely continuous such that $\dot{u}$ is not properly defined because of jumps with $u(t^-) \neq u(t^+)$. The following definition of *energetic solutions* is adapted to this situation, because it does not need any continuity or differentiability with respect to time. We use a general state space $\mathcal{Q}$, which can be a general topological space without a differential structure, and a general dissipation distance $\mathcal{D} : \mathcal{Q}\times\mathcal{Q} \to [0,\infty]$ which is assumed to satisfy the triangle inequality $\mathcal{D}(u_1,u_3) \le \mathcal{D}(u_1,u_2)+\mathcal{D}(u_2,u_3)$ (but not necessarily the symmetry and positivity of usual distances, see [Mie05, Ch. 5] or [Mie11a]). We call a triple $(\mathcal{Q},\mathcal{E},\mathcal{D})$ an *energetic RIS*. Starting from a rate-independent dissipation potential $\mathcal{R}$, it is possible to find the associated dissipation distance $\mathcal{D}(u_0,u_1)$ by minimizing $\int_0^1 \mathcal{R}(u(s),\dot{u}(s))\,\mathrm{d}s$ along all curves connecting u_0 and u_1, which is similar but not equal to the definition of $\mathcal{D}$ in Section 4.3.

Definition 5.2 (Energetic solutions for RIS) *A function $u : [0,T] \to \mathcal{Q}$ is called an* energetic solution *for the RIS $(\mathcal{Q},\mathcal{E},\mathcal{D})$ if **stability (S)** and **energy balance (E)** hold:*

$$\begin{aligned} &\textbf{(S)} \quad \mathcal{E}(t,u(t)) \le \mathcal{E}(t,w) + \mathcal{D}(u(t),w) \text{ for all } t\in[0,T] \text{ and } w\in\mathcal{Q},\\ &\textbf{(E)} \quad \mathcal{E}(T,u(T)) + \mathrm{Diss}_{\mathcal{D}}(u,[0,T]) = \mathcal{E}(0,u(0)) + \int_0^T \partial_t\mathcal{E}(t,u(t))\,\mathrm{d}t. \end{aligned} \tag{5.5}$$

In many cases this definition is equivalent to the notion of *quasistatic (irreversible) evolutions* studied in [FrL03, DFT05, Neg14].

The dissipated energy can be expressed solely by the dissipation distance

$$\mathrm{Diss}_{\mathcal{D}}(u,[0,T]) := \sup\Big\{ \textstyle\sum_{j=1}^N \mathcal{D}(u(t_{j-1}),u(t_j)) \;\Big|\; \text{all partitions of } [0,T] \Big\}$$

and coincides with $\int_0^T \mathcal{R}(u(t),\dot{u}(t))\,\mathrm{d}t$ in the smooth case.

The major advantage of the concept of energetic solutions is that it is a derivative-free formulation, which only features the values of the functionals $\mathcal{E}$ and $\mathcal{D}$ and the function values $u(t)$. Thus, it shares the same properties as the IEVE discussed in Section 4. In particular, existence results and evolutionary Γ-convergence can be attacked by tools from the calculus of variations. In fact, energetic solutions are the limits of the incremental minimization problems with time step $\tau = T/N \to 0$ (cf. [Mie05, Thm. 5.2]):

$$u_k^\tau \in \mathrm{Arg\,Min}\{\, \mathcal{E}(k\tau,u) + \mathcal{D}(u_{k-1}^\tau,u) \mid u\in\mathcal{Q}\,\}. \tag{5.6}$$

In contrast to the backward-Euler algorithm (4.10) for the metric gradient flow, now there is no explicit dependence on the time step τ (reflecting rate independence) and the dissipation distance has the power 1.

In fact, the reformulation of the rate formulations for RIS and the energetic solutions are equivalent in good cases, but are different in general, cf. [MiR14, Ch. 3].

Proposition 5.3 (Equivalence of formulations for convex energies) *Consider a RIS $(\boldsymbol{X},\mathcal{E},\mathcal{D})$, where $\boldsymbol{X}$ is a Banach space and assume that the energies $\mathcal{E}(t,\cdot) : \boldsymbol{X} \to \mathbb{R}_\infty$ are convex for all $t\in[0,T]$. Further assume that the dissipation distance $\mathcal{D}$ has the form $\mathcal{D}(u_0,u_1) = \Psi(u_1-u_0)$ for a lsc, convex and 1-homogeneous function $\Psi : \boldsymbol{X} \to [0,\infty]$. Then, $u \in \mathrm{W}^{1,1}([0,T];\boldsymbol{X})$ is an energetic solution if and only if u solves* (5.4).

5.3 pE-convergence for energetic solutions

We now consider a family $(\mathcal{Q}, \mathcal{E}_\varepsilon, \mathcal{D}_\varepsilon)$ of RIS, where for the simplicity of the presentation we restrict the discussion to the case that $\mathcal{Q}$ is a reflexive Banach space $\boldsymbol{X}$ and the energies $\mathcal{E}_\varepsilon$ have the form

$$\mathcal{E}_\varepsilon(t,u) = \mathcal{F}_\varepsilon(u) - \langle \ell_\varepsilon(t), u\rangle. \tag{5.7a}$$

We still keep general dissipation distances $\mathcal{D}_\varepsilon$ and do not assume convexity of $\mathcal{F}_\varepsilon$. Thus, all the results can be generalized to the general topological case, see [MRS08, MiR14]. Our precise assumptions are as follows:

$$\exists\, c, C > 0\ \forall\, \varepsilon \geq 0,\ u \in \boldsymbol{X}:\ \mathcal{F}_\varepsilon(u) \geq c\|u\|^2 - C; \tag{5.7b}$$

$$\forall\, \varepsilon \geq 0: \quad \mathcal{F}_\varepsilon : \boldsymbol{X} \to \mathbb{R}_\infty \text{ is weakly lsc}; \tag{5.7c}$$

$$\exists\, C > 0\ \forall\, \varepsilon \geq 0:\ \|\ell_\varepsilon\|_{\mathrm{C}^1([0,T])} \leq C; \tag{5.7d}$$

$$\forall\, t \in [0,T]: \qquad \dot{\ell}_\varepsilon(t) \to \dot{\ell}_0(t) \text{ in } \boldsymbol{X}^*; \tag{5.7e}$$

$$\forall\, \varepsilon \geq 0\ \forall\, u_j \in \boldsymbol{X}:\ \begin{cases} \mathcal{D}_\varepsilon(u_1,u_3) \leq \mathcal{D}_\varepsilon(u_1,u_2) + \mathcal{D}_\varepsilon(u_2,u_3), \\ \mathcal{D}_\varepsilon(u_1,u_2) = 0 \implies u_1 = u_2. \end{cases} \tag{5.7f}$$

In general, these conditions are not strong enough to show existence of solutions. Indeed, even if we assume existence of solutions for $\varepsilon > 0$, we cannot expect to conclude existence of solutions for $\varepsilon = 0$ by a limit argument. We need additional properties, e.g. the weak continuity of $\mathcal{D}_0$.

However, we can already address another nice property of the energetic formulation, namely the general validity of an appropriate generalization of the chain rule (3.1), which allows us to turn an upper energy estimate into an energy-dissipation balance as in Theorem 3.2. If $(\boldsymbol{X}, \mathcal{E}_0, \mathcal{D}_0)$ satisfies (5.7) and $u : [0,T] \to \boldsymbol{X}$ satisfies the global stability (S) in (5.5), then we have the lower energy estimate (see [Mie05, Prop. 5.7])

$$\mathcal{E}_0(t,u(t)) + \mathrm{Diss}_{\mathcal{D}_0}(u,[s,t]) \geq \mathcal{E}_0(s,u(s)) + \int_s^t \partial_r \mathcal{E}(r,u(r))\,\mathrm{d}r, \tag{5.8}$$

We present two results for evolutionary Γ-convergence. The first assumes that the dissipation distances $\mathcal{D}_\varepsilon$ weakly continuously converge to $\mathcal{D}_0$.

Theorem 5.4 (pE-convergence for RIS [MRS08]) *Assume that the RIS $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{D}_\varepsilon)$ satisfies (5.7) and $\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0$ and $\mathcal{D}_\varepsilon \overset{\mathsf{C}}{\rightharpoonup} \mathcal{D}_0$ in $\boldsymbol{X}$, then $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \mathcal{D}_\varepsilon) \overset{\mathsf{pE}}{\rightharpoonup} (\boldsymbol{X}, \mathcal{E}_0, \mathcal{D}_0)$, i.e. for energetic solutions $u_\varepsilon : [0,T] \to \boldsymbol{X}$ the following holds:*

$$\begin{aligned} &\text{if } \ u_\varepsilon(0) \rightharpoonup u_0(0) \text{ and } \mathcal{E}_\varepsilon(0,u_\varepsilon(0)) \to \mathcal{E}_0(0,u_0(0)) < \infty, \\ &\text{then } \ u_\varepsilon(t) \rightharpoonup u_0(t) \text{ and } \mathcal{E}_\varepsilon(t,u_\varepsilon(t)) \to \mathcal{E}_0(t,u_0(t)) \ \text{ for } 0 < t \leq T. \end{aligned} \tag{5.9}$$

Proof: From (5.7b) and (5.7d) we find $C_1, \Lambda > 0$ such that the power control $|\partial\mathcal{E}_\varepsilon(t,u)| \leq \Lambda\big(\mathcal{E}_\varepsilon(t,u)+C_1\big)$ holds. Inserting this estimate into the energy balance (E) we obtain the a priori bound

$$\mathcal{E}_\varepsilon(t,u_\varepsilon(t)) + \mathrm{Diss}_{\mathcal{D}_\varepsilon}(u_\varepsilon,[0,t]) \leq \mathrm{e}^{\Lambda t}\big(\mathcal{E}_\varepsilon(0,u_\varepsilon(0)) + C_1\big) - C_1 \leq E_*,$$

where we used the well-preparedness $\mathcal{E}_\varepsilon(0,u_\varepsilon(0)) \to \mathcal{E}_0(0,u_0(0)) < \infty$. Using (5.7b) once again we find $\|u_\varepsilon(t)\| \le C_2$ for all t and $\varepsilon > 0$, and Helly's selection principle (cf. [Mie05, Thm. 5.1]) allows us to extract a (not relabeled) subsequence such that $u_\varepsilon(t) \rightharpoonup u_0(t)$ for all t.

Next, we show that all $u_0(t)$ satisfy the stability condition (S). We know that (S) is true for $\varepsilon > 0$, i.e.

$$\mathcal{E}_\varepsilon(t,u_\varepsilon(t)) \le \mathcal{E}_\varepsilon(t,w_\varepsilon)+\mathcal{D}_\varepsilon(u_\varepsilon(t),w_\varepsilon) \text{ for all } t\in[0,T] \text{ and } w_\varepsilon \in \boldsymbol{X}. \tag{5.10}$$

For a given test state w we choose a recovery sequence $w_\varepsilon \rightharpoonup w$ with $\mathcal{E}_\varepsilon(t,w_\varepsilon) \to \mathcal{E}_0(t,w)$ and pass to the limit, viz.

$$\mathcal{E}_0(t,u_0(t)) \overset{\Gamma}{\le} \liminf_{\varepsilon\to 0} \mathcal{E}_\varepsilon(t,u_\varepsilon(t)) \le \mathcal{E}_0(t,w) + \mathcal{D}_0(u_0(t),w),$$

where we used $\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0$ and $\mathcal{D}_\varepsilon \overset{\mathsf{C}}{\rightharpoonup} \mathcal{D}_0$. Thus, stability of $u_0(t)$ is established.

In the energy balance $(\mathrm{E})_\varepsilon$ we can pass to the limit using (5.7d) and (5.7e) to obtain

$$\int_0^T \partial_t\mathcal{E}_\varepsilon(t,u_\varepsilon(t))\,\mathrm{d}t = -\int_0^T \langle \dot\ell_\varepsilon, u_\varepsilon\rangle\,\mathrm{d}t \to -\int_0^T \langle \dot\ell_0, u_0\rangle\,\mathrm{d}t = \int_0^T \partial_t\mathcal{E}_0(t,u_0(t))\,\mathrm{d}t.$$

Since the initial energies converge, the right-hand sides in $(\mathrm{E})_\varepsilon$ converge to that of $(\mathrm{E})_0$.

Using $\mathcal{D}_\varepsilon \overset{\mathsf{C}}{\rightharpoonup} \mathcal{D}_0$ we estimate the total dissipation on the interval $[0,t]$ via

$$\lim_{\varepsilon\to 0} \sum_{j=1}^N \mathcal{D}_\varepsilon(u_\varepsilon(t_{j-1}),u_\varepsilon(t_j)) = \sum_{j=1}^N \mathcal{D}_0(u_0(t_{j-1}),u_0(t_j)) \ge \mathrm{Diss}_{\mathcal{D}_0}(u_0,[0,T]) - \rho,$$

where $\rho > 0$ can be made arbitrary small by choosing a suitable partition $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = t$. Hence, $\liminf_{\varepsilon\to 0} \mathrm{Diss}_{\mathcal{D}_\varepsilon}(u_\varepsilon,[0,t]) =: D_0(t) \ge d_0(t) := \mathrm{Diss}_{\mathcal{D}_0}(u_0,[0,t])$. Because of $\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0$ for all $t\in[0,T]$ we have

$$e_0(t) := \mathcal{E}_0(t,u_0(t)) \le E_0(t) := \liminf_{\varepsilon\to 0} \mathcal{E}_\varepsilon(t,u_\varepsilon(t)) \le E_1(t) := \limsup_{\varepsilon\to 0} \mathcal{E}_\varepsilon(t,u_\varepsilon(t)).$$

Using the upper energy estimate on $[0,t]$ yields

$$e_0(t) + d_0(t) \le E_1(t) + D_0(t) \le e_0(0) + \int_0^t \partial_s\mathcal{E}_0(s,u_0(s))\,\mathrm{d}s \le e_0(t) + d_0(t),$$

where the last estimate follows from the lower energy estimate (5.8). Hence, we conclude $e_0(t) = E_0(t) = E_1(t)$ which is the desired energy convergence in (5.9). ∎

We apply this result to a rate-independent homogenization problem with nonquadratic energies. We first show that the result fails, even if $\mathcal{D}_\varepsilon = \mathcal{D}_0$, e.g. if $\mathcal{D}_0$ is not weakly continuous.

Example 5.5 (Counterexample) *We consider*

$$\boldsymbol{X} = \mathrm{L}^2(\Omega), \quad \mathcal{E}_\varepsilon(t,u) = \textstyle\int_\Omega \frac12 a(\frac1\varepsilon x)u(x)^2 - \ell(t)u(x)\,\mathrm{d}x, \quad \text{and } \Psi_\varepsilon(v) = \textstyle\int_\Omega c(\frac1\varepsilon x)|v(x)|\,\mathrm{d}x.$$

Clearly, $\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0 : u \mapsto \int_\Omega \frac{a_{harm}}{2} u^2 - \ell u \,\mathrm{d}x$ *while* $\Psi_\varepsilon \overset{\Gamma}{\to} \Psi_{arith} := c_{arith}\|\cdot\|_{\mathrm{L}^1}$ *and* $\Psi_\varepsilon \overset{\Gamma}{\rightharpoonup} \Psi_{min} = c_{min}\|\cdot\|_{\mathrm{L}^1}$*, if the set* $\{\, y \in [0,1]^d \mid c(y) = c_{min} \,\}$ *has positive* $\mathcal{L}^d$ *measure.*

For the loading we assume $\ell(0) = 0$ *and* $\dot\ell(t) > 0$ *for* $t > 0$*. Starting from the well-prepared initial condition* $u_\varepsilon(0,x) = 0$ *we find the unique solution*

$$u_\varepsilon(t,x) = \max\{0, \ell(t) - c(\tfrac{1}{\varepsilon}x)\}/a(\tfrac{1}{\varepsilon}x).$$

Hence, for $c \equiv$ *const we obtain weak evolutionary* Γ*-convergence to* $(\mathrm{L}^2(\Omega), \mathcal{F}_0, \Psi_0)$ *with* $\Psi_0 = \Psi_{arith} = \Psi_{min}$*, while for* $c(\cdot)$ *nonconstant the weak limit* $U(t)$ *of* $u_\varepsilon(t)$ *cannot be described by a RIS system* $(\mathrm{L}^2(\Omega), \mathcal{F}_0, \Psi_0)$ *for any* Ψ_0*.*

Example 5.6 (Homogenization of RIS) *We consider* $\boldsymbol{X} = \mathrm{H}^1_0(\Omega)$ *for a bounded Lipschitz domain* $\Omega \subset \mathbb{R}^d$*, and energy functional* $\mathcal{F}_\varepsilon(u) = \int_\Omega \frac12 \nabla u \cdot A(\frac{1}{\varepsilon}x)\nabla u + F(\frac{1}{\varepsilon}x, u)\,\mathrm{d}x$*, and a dissipation distance* $\mathcal{D}_\varepsilon(u_1,u_2) = \Psi_\varepsilon(u_2 - u_1)$ *with* $\Psi_\varepsilon(v) = \int_\Omega c(\frac{1}{\varepsilon}x)|v(x)|\,\mathrm{d}x$*, where* $F(y,\cdot) \ge 0$ *is continuous and convex and* A*,* c*, and* $F(\cdot,u)$ *are 1-periodic in* $y = \frac{1}{\varepsilon}x$*.*

By Proposition 5.3, the energetic solutions are solutions of the differential inclusion

$$0 \in c(\tfrac{1}{\varepsilon}x)\mathrm{Sign}(\dot u) - \mathrm{div}\left(A(\tfrac{1}{\varepsilon}x)\nabla u\right) + \partial_u F(\tfrac{1}{\varepsilon}x, u) - \ell_\varepsilon(t).$$

Weak convergence in $\mathrm{H}^1(\Omega)$ *implies strong convergence in* $\mathrm{L}^2(\Omega)$*, so in* $\boldsymbol{X}$ *we have*

$$\Psi_\varepsilon \overset{\mathsf{C}}{\rightharpoonup} \Psi_0 : v \mapsto \int_\Omega c_{arith}|v|\,\mathrm{d}x \quad \text{and} \quad \mathcal{F}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{F}_0 : u \mapsto \int_\Omega \frac12 \nabla u \cdot A_{eff}\nabla u + F_{arith}(u)\,\mathrm{d}x.$$

Theorem 5.4 is applicable giving $(\boldsymbol{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \overset{\mathsf{pE}}{\rightharpoonup} (\boldsymbol{X}, \mathcal{E}_0, \Psi_0)$*, where the limit equation reads* $0 \in c_{arith}\,\mathrm{Sign}(\dot u) - \mathrm{div}\left(A_{eff}\nabla u\right) + \partial_u F_{arith}(u) - \ell_0(t)$*.*

Our second result on evolutionary Γ-convergence does not need the weak continuity of $\mathcal{D}_0$, but requires the Hilbert-space setting with

$$\mathcal{F}_\varepsilon(u) = \frac12\langle A_\varepsilon u, u\rangle \quad \text{and} \quad \mathcal{D}_\varepsilon(u_1,u_2) = \Psi_\varepsilon(u_2 - u_1) \text{ with } \Psi_\varepsilon(\gamma v) = \gamma^1 \Psi_\varepsilon(v). \tag{5.11}$$

This case is important in classical models of elastoplasticity, and for fixed ε it can be reduced to the analysis of monotone operators, see [Alb98]. For evolutionary Γ-convergence with Ψ_ε truely dependent on ε the notion of energetic solutions is more flexible. The following convergence result was developed for two-scale homogenization and dimension reduction in elastoplastic models in [MiT07] and [LiM11], respectively.

The homogenization of RIS without compactness does not work directly (see Example 5.5), however the method of periodic unfolding developed for two-scale homogenization turns weak Γ-convergence of the functionals into Mosco-convergence in the two-scale setting. Hence, the homogenization results for elastoplasticity in [MiT07, GiM11, Han11] can be derived using the following abstract evolutionary Γ-convergence result.

Theorem 5.7 (pE-convergence for quadratic RIS) *Assume that* $\boldsymbol{H}$ *is a Hilbert space and that the RIS* $(\boldsymbol{H}, \mathcal{E}_\varepsilon, \mathcal{D}_\varepsilon)$*,* $\varepsilon \in [0,1]$*, satisfy* (5.7) *and* (5.11)*. If additionally*

$$\mathcal{E}_\varepsilon \overset{\mathsf{M}}{\to} \mathcal{E}_0, \qquad \Psi_\varepsilon \overset{\mathsf{C}}{\to} \Psi_0, \quad \textit{and} \quad \Psi_\varepsilon \overset{\Gamma}{\rightharpoonup} \Psi_0 \quad \textit{in } \boldsymbol{H},$$

then $(\boldsymbol{H}, \mathcal{E}_\varepsilon, \mathcal{D}_\varepsilon) \overset{\mathsf{pE}}{\rightharpoonup} (\boldsymbol{H}, \mathcal{E}_0, \mathcal{D}_0)$ *as in* (5.9)*.*

Proof: We can proceed as in the proof of Theorem 5.4 for all parts except for the stability of the weak limits $u_0(t)$, since now weak convergence is not enough to pass to the limit in $\Psi_\varepsilon(w_\varepsilon - u_\varepsilon(t))$, but for the upper energy estimate we need $\Psi_\varepsilon \xrightarrow{\Gamma} \Psi_0$.

We prove the desired stability of $u_0(t)$ under the additional assumption $\mathcal{F}_\varepsilon(u) \le C_2\|u\|^2$. For the general case we refer to [LiM11] and [MiR14, Ch. 3.5.4]. Then, for a test state v we can choose the recovery sequence $v_\varepsilon = A_\varepsilon^{-1}A_0 v$. Indeed, since v_ε minimizes the functional $u \mapsto \mathcal{F}_\varepsilon(u) - \langle A_0 w, u\rangle$ and $\mathcal{F}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{F}_0$ we conclude $v_\varepsilon \to v$ and $\mathcal{F}_\varepsilon(v_\varepsilon) \to \mathcal{F}_0(v)$ by a variant of Theorem 2.4.

Using the stability of $u_\varepsilon(t)$ we have, for all w_ε,

$$0 \le \mathcal{E}_\varepsilon(t, w_\varepsilon) - \mathcal{E}_\varepsilon(t, u_\varepsilon) + \Psi_\varepsilon(w_\varepsilon - u_\varepsilon(t)). \tag{5.12}$$

For a given limit test state w we choose $w_\varepsilon = u_\varepsilon + A_\varepsilon^{-1}A_0(w-u_0)$, which guarantees

$$w_\varepsilon \rightharpoonup w \quad \text{and} \quad w_\varepsilon - u_\varepsilon \to w - u_0, \text{ giving } \Psi_\varepsilon(w_\varepsilon - u_\varepsilon) \to \Psi_0(w-u_0), \tag{5.13}$$

because of $\Psi_\varepsilon \xrightarrow{\mathsf{C}} \Psi_0$. Moreover, using the quadratic structure of $\mathcal{E}_\varepsilon$ we have

$$\begin{aligned}
\mathcal{E}_\varepsilon(t, w_\varepsilon) - \mathcal{E}_\varepsilon(t, u_\varepsilon) &= \frac{1}{2}\langle A_\varepsilon(w_\varepsilon - u_\varepsilon), w_\varepsilon + u_\varepsilon\rangle - \langle \ell_\varepsilon, w_\varepsilon - u_\varepsilon\rangle \\
&= \frac{1}{2}\langle A_0(w-u_0), w_\varepsilon + u_\varepsilon\rangle - \langle \ell_\varepsilon, w_\varepsilon - u_\varepsilon\rangle \\
&\to \frac{1}{2}\langle A_0(w-u_0), w_0 + u_0\rangle - \langle \ell_\varepsilon, w_0 - u_0\rangle = \mathcal{E}_0(t, w) - \mathcal{E}_0(t, u_0).
\end{aligned}$$

Using this convergence and (5.13) in (5.12) yields the desired stability of $u_0(t)$. ∎

The following simple example from [MRS08] shows that Mosco convergence of $\mathcal{E}_\varepsilon$ and Ψ_ε is not sufficient for evolutionary Γ-convergence even in finite dimensions.

Example 5.8 *For $\varepsilon > 0$ consider the RIS $(\boldsymbol{H}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ with*

$$\boldsymbol{H} = \mathbb{R}^2, \quad \mathcal{E}_\varepsilon(t, q) = \frac{1}{2}q_1^2 + \frac{1}{2}(q_1 - \frac{q_2}{\varepsilon})^2 - tq_1, \quad \mathcal{R}_\varepsilon(v) = |v_1| + \tfrac{1}{\varepsilon^2}|v_2|.$$

In $\mathbb{R}^2$ all Γ-limits are Mosco limits, i.e. we have $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}_0$ and $\mathcal{R}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{R}_0$ with

$$\mathcal{E}_0(t, q) = \begin{cases} \frac{1}{2}q_1^2 - tq_1 & \text{for } q_2 = 0, \\ \infty & \text{for } q_2 \neq 0; \end{cases} \quad \text{and} \quad \mathcal{R}_0(v) = \begin{cases} |v_1| & \text{for } v_2 = 0, \\ \infty & \text{for } v_2 \neq 0. \end{cases}$$

For the unique solutions q_ε with $q_\varepsilon(0) = 0$ we find

$$q_0(t) = \begin{pmatrix} \max\{t-1,\, 0\} \\ 0 \end{pmatrix} \neq \lim_{\varepsilon \to 0} q_\varepsilon(t) = \begin{pmatrix} \max\{0,\, t/2-1\} \\ 0 \end{pmatrix},$$

which contradicts evolutionary Γ-convergence.

Remark 5.9 (Mutual recovery sequences) *In both convergence results the important step in the proof is the stability of the limits $u_0(t)$. The same problem already appears in the existence theory via the time-incremental minimization* (5.6)*. Thus, in*

[MRS08, MiR14] the notion of mutual recovery sequences $(w_\varepsilon)_{\varepsilon>0}$ *is introduced for a given stable sequence* u_ε *and a test state* w *by asking for the condition*

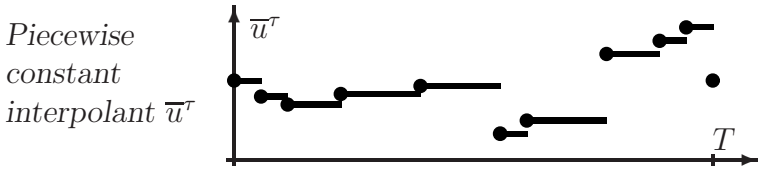
$$\limsup_{\varepsilon\to 0} \big(\mathcal{E}_\varepsilon(t,w_\varepsilon) - \mathcal{E}_\varepsilon(t,u_\varepsilon) + \mathcal{D}_\varepsilon(w_\varepsilon,u_\varepsilon)\big) \ \le\ \mathcal{E}_0(t,w) - \mathcal{E}_0(t,u_0) + \mathcal{D}_0(w,u_0),$$

see the occurrence of w_ε *in* (5.10) *and* (5.12) *in the proofs of Theorems 5.4 and 5.7, respectively. An important nontrivial construction in that direction was the so-called* jump transfer lemma *for rate-independent models for brittle fracture, see [FrL03, DFT05].*

The above abstract results have a variety of applications. In [MiS13] it is shown that linearized elastoplasticity can be obtained by pE-convergence from finite-strain elastoplasticity in the limit of small loadings and small yield stress. In [Gia05] brittle fracture is obtained as evolutionary Γ-limit of the Ambrosio-Tortorelli approximation, while [GiP06a] discusses homogenization for fracture evolution.

Remark 5.10 (Numerical approximation of RIS) *Both of the above convergence results can be used to establish convergence of numerical schemes for a given RIS* $(\boldsymbol{X},\mathcal{E},\mathcal{D})$*, see [MiR09] for various versions. The main idea is to consider the incremental minimization problem* (5.6) *restricted to finite-dimensional subspaces* $\boldsymbol{X}_m$ *of the underlying Banach space* $\boldsymbol{X}$*, to keep the distance* $\mathcal{D}$ *or* Ψ *independent of* m*, while defining* $\mathcal{E}_m(t,u) = \mathcal{E}(t,u)$ *for* $u\in\boldsymbol{X}_m$ *and* $\mathcal{E}_m(t,u)=\infty$ *for* $u\in\boldsymbol{X}\setminus\boldsymbol{X}_m$*.*

We say that the subspaces $\boldsymbol{X}_m$ *approximate* $\boldsymbol{X}$*, if for each* $u\in\boldsymbol{X}$ *there exists* $u_m\in\boldsymbol{X}_m$ *such that* $u_m\to u$ *for* $m\to\infty$*. Under mild conditions on* $\mathcal{E}(t,\cdot)$ *one can show that* $\mathcal{E}_m \overset{\mathsf{M}}{\longrightarrow} \mathcal{E}$ *in* $\boldsymbol{X}$*, and then the above theorems are applicable.*



In fact, one can strengthen the result for the piecewise constant approximants $\overline{u}^{\tau,m}$ *by proving that the joint limit* $\tau\to 0$ *and* $m\to\infty$ *always has a subsequence converging to a limit* u_0 *which is an energetic solution for* $(\boldsymbol{X},\mathcal{E},\mathcal{D})$*, cf. [GiP06b, MRS08, MiR09].*

In [Bra13] the general interplay between incremental minimization with time step $\tau>0$ *and* Γ*-convergence of the functionals* $\mathcal{E}_\varepsilon$ *and* Ψ_ε *for* $\varepsilon\to 0$ *is studied.*

A completely different approach to pE-convergence for RIS is developed in [Neg14] which involves the notion of parametrized solutions for RIS, which is a variant of the *balanced-viscosity solutions* defined in [MRS09, MRS13a]. The convergence result in [Neg14] is based on an adaptation of the theory in Section 3.3 using the energy-dissipation balance.

Acknowledgments. The research was partially supported by the DFG within the SFB 910 (subproject A5) and by the ERC under AdG 267802 AnaMultiScale. These notes are based on lectures given at Twente University in June 2012 and at the Centro De Giorgi (Pisa) in November 2013. The author gives special thanks to Karoline Disser, Matthias Liero, Michiel Renger, Sina Reichelt, and Marita Thomas for reading and considerably improving earlier versions of this work. Moreover, he is grateful for stimulating discussions with Dorothee Knees, Mark Peletier, Riccarda Rossi, Tomáš Roubíček, Giuseppe Savaré, Ulisse Stefanelli, Lev Truskinovsky, and Augusto Visintin.

References

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[Alb98] H.-D. Alber. *Materials with Memory*, volume 1682 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.

[AM*12] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: from diffusion to reaction. *Calc. Var. Part. Diff. Eqns.*, 44, 419–454, 2012.

[Att84] H. Attouch. *Variational Convergence of Functions and Operators*. Pitman Advanced Publishing Program. Pitman, 1984.

[BB*12] G. Bellettini, L. Bertini, M. Mariani, and M. Novaga. Convergence of the one-dimensional Cahn-Hilliard equation. *SIAM J. Math. Anal.*, 44(5), 3458–3480, 2012.

[Bén72] P. Bénilan. Solutions intégrales d'équations d'évolution dans un espace de Banach. *C. R. Acad. Sci. Paris Sér. A-B*, 274, A47–A50, 1972.

[Bio55] M. A. Biot. Variational principles in irreversible thermodynamics with applications to viscoelasticity. *Phys. Review*, 97(6), 1463–1469, 1955.

[BKS04] M. Brokate, P. Krejčí, and H. Schnabel. On uniqueness in evolution quasivariational inequalities. *J. Convex Analysis*, 11, 111–130, 2004.

[Bra02] A. Braides. Γ-*Convergence for Beginners*. Oxford University Press, 2002.

[Bra06] A. Braides. A handbook of Γ-convergence. In M. Chipot and P. Quittner, editors, *Handbook of Differential Equations. Stationary Partial Differential Equations. Volume 3*. Elsevier, 2006.

[Bra13] A. Braides. *Local minimization, Variational Evolution and Gamma-convergence*. Lect. Notes Math. Vol. 2094. Springer, 2013.

[Bré73] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973.

[Dal93] G. Dal Maso. *An Introduction to* Γ-*Convergence*. Birkhäuser Boston Inc., Boston, MA, 1993.

[DaS10] S. Daneri and G. Savaré. Lecture notes on gradient flows and optimal transport. arXiv:1009.3737v1, 2010.

[DeF75] E. De Giorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 58(6), 842–850, 1975.

[DFT05] G. Dal Maso, G. Francfort, and R. Toader. Quasistatic crack growth in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 176, 165–225, 2005.

[DiL13] K. Disser and M. Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *WIAS preprint 1899*, 2013.

[DMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3), 180–187, 1980.

[EkT76] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North Holland, 1976.

[Fen49] W. Fenchel. On conjugate convex functions. *Canadian J. Math.*, 1, 73–77, 1949.

[FeN09] E. Feireisl and A. Novotný. *Singular Limits in Thermodynamics of Viscous Fluids*. Birkhäuser, 2009.

[FrL03] G. A. Francfort and C. Larsen. Existence and convergence for quasi-static evolution of brittle fracture. *Comm. Pure Applied Math.*, 56, 1495–1500, 2003.

[Gia05] A. Giacomini. Ambrosio-Tortorelli approximation of quasistatic evolution of brittle fractures. *Calc. Var. Part. Diff. Eqns.*, 22, 129–172, 2005.

[GiM11] A. Giacomini and A. Musesti. Two-scale homogenization for a model in strain gradient plasticity. *ESAIM Control Optim. Calc. Var.*, 17(4), 1035–1065, 2011. Published online. DOI 10.1051/cocv/2010036.

[GiM13] N. Gigli and J. Maas. Gromov-Hausdorff convergence of discrete transportation metrics. *SIAM J. Math. Analysis*, 45(2), 879–899, 2013.

[GiP06a] A. Giacomini and M. Ponsiglione. A Γ-convergence approach to stability of unilateral minimality properties in fracture mechanics and applications. *Arch. Ration. Mech. Anal.*, 180(3), 399–447, 2006.

[GiP06b] A. Giacomini and M. Ponsiglione. Discontinuous finite element approximation of quasistatic crack growth in nonlinear elasticity. *Math. Models Methods Appl. Sci.*, 16, 77–118, 2006.

[Gli12] A. Glitzky. An electronic model for solar cells including active interfaces and energy resolved defect densities. *SIAM J. Math. Analysis*, 44, 3874–3900, 2012.

[GlM13] A. Glitzky and A. Mielke. A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. *Z. angew. Math. Phys. (ZAMP)*, 64, 29–52, 2013.

[Han11] H. Hanke. Homogenization in gradient plasticity. *Math. Models Meth. Appl. Sci. (M^3AS)*, 21(8), 1651–1684, 2011.

[Iof77] A. D. Ioffe. On lower semicontinuity of integral functionals. I. *SIAM J. Control Optimization*, 15(4), 521–538, 1977.

[Jam96] R. D. James. Hysteresis in phase transformations. In *ICIAM 95 (Hamburg, 1995)*, volume 87 of *Math. Res.*, pages 135–154. Akademie Verlag, Berlin, 1996.

[JKO98] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Analysis*, 29(1), 1–17, 1998.

[Kre99] P. Krejčí. Evolution variational inequalities and multidimensional hysteresis operators. In *Nonlinear differential equations (Chvalatice, 1998)*, volume 404 of *Chapman & Hall/CRC Res. Notes Math.*, pages 47–110. Chapman & Hall/CRC, Boca Raton, FL, 1999.

[Lie12a] M. Liero. Passing from bulk to bulk-surface evolution in the Allen-Cahn equation. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 2012. To appear. WIAS preprint 1676.

[Lie12b] M. Liero. *Variational Methods for Evolution*. PhD thesis, Institut für Mathematik, Humboldt-Universität zu Berlin, 2012.

[LiM11] M. Liero and A. Mielke. An evolutionary elastoplastic plate model derived via Γ-convergence. *Math. Models Meth. Appl. Sci. (M^3AS)*, 21(9), 1961–1986, 2011.

[LMS14] M. Liero, A. Mielke, and G. Savaré. Optimal transport in competition with reaction – the Hellinger–Kantorovich distance. *In preparation*, 2014.

[Maa11] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261, 2250–2292, 2011.

[Mie05] A. Mielke. Evolution in rate-independent systems (Ch. 6). In C. Dafermos and E. Feireisl, editors, *Handbook of Differential Equations, Evolutionary Equations, vol. 2*, pages 461–559. Elsevier B.V., Amsterdam, 2005.

[Mie11a] A. Mielke. Differential, energetic, and metric formulations for rate-independent processes. In L. Ambrosio and G. Savaré, editors, *Nonlinear PDE's and Applications*, pages 87–170. Springer, 2011. (C.I.M.E. Summer School, Cetraro, Italy 2008, Lect. Notes Math. Vol. 2028).

[Mie11b] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24, 1329–1346, 2011.

[Mie12] A. Mielke. Emergence of rate-independent dissipation from viscous systems with wiggly energies. *Contin. Mech. Thermodyn.*, 24(4), 591–606, 2012.

[Mie13] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discr. Cont. Dynam. Systems Ser. S*, 6(2), 479–499, 2013.

[Mie14] A. Mielke. Deriving amplitude equations via evolutionary Γ-convergence. *Discr. Cont. Dynam. Systems Ser. A*, 2014. Submitted. (WIAS Preprint 1914).

[MiR09] A. Mielke and T. Roubíček. Numerical approaches to rate-independent processes and applications in inelasticity. *Math. Model. Numer. Anal. (M2AN)*, 43, 399–428, 2009.

[MiR14] A. Mielke and T. Roubíček. *Rate-Independent Systems: Theory and Application*. In preparation, 2014.

[MiS13] A. Mielke and U. Stefanelli. Linearized plasticity is the evolutionary Γ-limit of finite plasticity. *J. Europ. Math. Soc.*, 15(3), 923–948, 2013.

[MiT07] A. Mielke and A. M. Timofte. Two-scale homogenization for evolutionary variational inequalities via the energetic formulation. *SIAM J. Math. Analysis*, 39(2), 642–668, 2007.

[MiT12] A. Mielke and L. Truskinovsky. From discrete visco-elasticity to continuum rate-independent plasticity: rigorous results. *Arch. Rational Mech. Anal.*, 203(2), 577–619, 2012.

[MiT14] A. Mielke and M. Thomas. GENERIC — a powerful tool for thermomechanical modeling. *In preparation*, 2014.

[MOS13] A. Mielke, C. Ortner, and Y. Sengül. An approach to nonlinear viscoelasticity via metric gradient flows. *SIAM J. Math. Analysis*, 2013. To appear. WIAS preprint 1816 (July 2013).

[MPR13] A. Mielke, M. Peletier, and M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *WIAS preprint 1868*, 2013. Submitted.

[MRS08] A. Mielke, T. Roubíček, and U. Stefanelli. Γ-limits and relaxations for rate-independent evolutionary problems. *Calc. Var. Part. Diff. Eqns.*, 31, 387–416, 2008.

[MRS09] A. Mielke, R. Rossi, and G. Savaré. Modeling solutions with jumps for rate-independent systems on metric spaces. *Discr. Cont. Dynam. Systems Ser. A*, 25(2), 585–615, 2009.

[MRS13a] A. Mielke, R. Rossi, and G. Savaré. Balanced-viscosity (BV) solutions to infinite-dimensional rate-independent systems. *WIAS preprint 1845*, 2013. Submitted. http://arxiv.org/abs/1309.6291.

[MRS13b] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Part. Diff. Eqns.*, 46(1-2), 253–310, 2013.

[MRT12] A. Mielke, T. Roubíček, and M. Thomas. From damage to delamination in nonlinearly elastic materials at small strains. *J. Elasticity*, 109, 235–273, 2012.

[Neg14] M. Negri. Quasi-static rate-independent evolutions: characterization, existence, approximation and application to fracture mechanics. *ESAIM Control Optim. Calc. Var.*, 2014. To appear.

[OnM53] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91(6), 1505–1512, 1953.

[Ons31] L. Onsager. Reciprocal relations in irreversible processes, I+II. *Physical Review*, 37, 405–426, 1931. (part II, 38:2265-2279).

[Ott01] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26, 101–174, 2001.

[Pra28] L. Prandtl. Gedankenmodel zur kinetischen Theorie der festen Körper. *Z. angew. Math. Mech. (ZAMM)*, 8, 85–106, 1928.

[PuT05] G. Puglisi and L. Truskinovsky. Thermodynamics of rate-independent plasticity. *J. Mech. Phys. Solids*, 53, 655–679, 2005.

[Roc70] R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[RoS06] R. Rossi and G. Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM Control Optim. Calc. Var.*, 12, 564–614, 2006.

[Rou05] T. Roubíček. *Nonlinear Partial Differential Equations with Applications*. Birkhäuser Verlag, Basel, 2005.

[SaS04] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, LVII, 1627–1672, 2004.

[Sav11] G. Savaré. Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds. In preparation, 2011.

[Ser11] S. Serfaty. Gamma-convergence of gradient flows on Hilbert spaces and metric spaces and applications. *Discr. Cont. Dynam. Systems Ser. A*, 31(4), 1427–1451, 2011.

[Ste08] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.*, 47(3), 1615–1642, 2008.

[Tar89] L. Tartar. Nonlocal effects induced by homogenization. In *Partial Differential Equations and the Calculus of Variations, vol. 2*, pages 925–938. Birkhäuser, 1989.

[Tar90] L. Tartar. Memory effects and homogenization. *Arch. Rational Mech. Anal.*, 111(2), 121–133, 1990.