Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

On microscopic origins of generalized gradient structures

Matthias Liero[1], Alexander Mielke[1,2], Mark A. Peletier[3], D.R. Michiel Renger[1]

submitted: July 23, 2015

[1] Weierstrass Institute
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: matthias.liero@wias-berlin.de
        alexander.mielke@wias-berlin.de
        d.r.michiel.renger@wias-berlin.de

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

[3] Department of Mathematics and Computer Science and
Institute for Complex Molecular Systems (ICMS)
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven
the Netherlands
E-Mail: m.a.peletier@tue.nl

Abstract

Classical gradient systems have a linear relation between rates and driving forces. In generalized gradient systems we allow for arbitrary relations derived from general non-quadratic dissipation potentials. This paper describes two natural origins for these structures.

A first microscopic origin of generalized gradient structures is given by the theory of large-deviation principles. While Markovian diffusion processes lead to classical gradient structures, Poissonian jump processes give rise to cosh-type dissipation potentials.

A second origin arises via a new form of convergence, that we call EDP-convergence. Even when starting with classical gradient systems, where the dissipation potential is a quadratic functional of the rate, we may obtain a generalized gradient system in the evolutionary Γ-limit. As examples we treat (i) the limit of a diffusion equation having a thin layer of low diffusivity, which leads to a membrane model, and (ii) the limit of diffusion over a high barrier, which gives a reaction-diffusion system.

1 Introduction

We consider evolution equations $\dot{u} = V(t, u)$ that are generated by gradient systems (GS). By a GS we understand a triple $(\mathbf{X}, \mathcal{E}, \mathcal{R})$, where the state space $\mathbf{X}$ is a weakly closed convex subset of a Banach space containing the states $u(t)$. The functional $\mathcal{E} : \mathbf{X} \to \mathbb{R} \cup \{\infty\}$ is called energy, but in applications it may be a free energy, a relative entropy, or the negative of the entropy. Finally $\mathcal{R}$ is the dissipation potential depending on the state u and the rate $\dot{u}$ such that $\mathrm{D}_{\dot{u}} \mathcal{R}(u, \dot{u}) \in \mathbf{X}^*$ denotes the dissipation force. The induced evolution equation is the force balance

$$0 = \mathrm{D}_{\dot{u}} \mathcal{R}(u(t), \dot{u}(t)) + \mathrm{D}_u \mathcal{E}(t, u(t)), \tag{1.1}$$

where the symbol D denotes the (partial) Gateaux derivative or the convex subdifferential. Quite often, we will use the dual dissipation potential $\mathcal{R}^*$ that is defined by the Legendre-Fenchel transform of $\mathcal{R}(u, \cdot)$. Then, the evolution equation can be rewritten as

$$\dot{u}(t) = \mathrm{D}_\xi \mathcal{R}^* \big(u(t), -\mathrm{D}\mathcal{E}(u(t)) \big), \tag{1.2}$$

see Section 2.1 for the details. Since $\mathcal{R}$ and $\mathcal{R}^*$ are in one-to-one correspondence, we will sometimes denote $(\mathbf{X}, \mathcal{E}, \mathcal{R})$ also by $(\mathbf{X}, \mathcal{E}, \mathcal{R}^*)$, in particular if $\mathcal{R}^*$ is given explicitly.

A third equivalent formulation of the gradient flow is given via the *energy-dissipation principle (EDP)*, also called De Giorgi's $(\mathcal{R}, \mathcal{R}^*)$ principle, cf. [DMT80, AGS05]. This states that, under suitable technical assumptions, a curve $u : [0, T] \to \mathbf{X}$ is a solution of (1.1) or (1.2) if and only if it satisfies the energy-dissipation estimate

$$\mathcal{E}(u(T)) + \mathscr{D}(u) \leq \mathcal{E}(u(0)) \text{ with } \mathscr{D}(u) := \int_0^T \mathcal{R}(u(t), \dot{u}(t)) + \mathcal{R}^* \big(u(t), -\mathrm{D}\mathcal{E}(u(t)) \big) \, \mathrm{d}t. \tag{1.3}$$

We call $\mathscr{D}$ the *De Giorgi dissipation functional*.

A GS is called *classical* if the dissipation potential $\mathcal{R}(u, \cdot)$ is quadratic, i.e. $\mathcal{R}(u, \dot{u}) = \frac{1}{2} \langle \mathbb{G}(u) \dot{u}, \dot{u} \rangle$ for a linear, symmetric, and positive definite operator $\mathbb{G}(u) : \mathbf{X} \to \mathbf{X}^*$. If we want to emphasize that a GS is not classical, we call it a *generalized GS*. The aim of

this work is to show that generalized GS arise in two natural ways. First, it is shown in [MPR14, MP*15] that they appear via large-deviation principles from a microscopic N-particle system for $N \to \infty$, see Section 2.4 for a brief summary of the main result. Second, generalized GS occur as suitable multiscale limits of classical GS.

Obviously, every GS generates exactly one gradient-flow evolution equation by (1.1) or (1.2), but a given evolution equation $\dot{u} = V(t, u)$ may be generated by many GS. If there exists at least one such GS, we say that the evolution equation has a gradient structure, if we do not want to specify the particular GS. As an elementary example we treat the scalar ODE

$$\dot{p} = 1 - 2p \text{ with } p(t) \in [0, 1] = \text{Prob}(\{1, 2\}),$$

which we interpret as the Kolmogorov forward equation for a Markov process $(X_t)_{t \geq 0}$ with $X_t \in \{1, 2\}$. Obviously, this ODE is generated by the GS $([0, 1], \mathcal{E}_2, \mathcal{R}_2)$ with

$$\mathcal{E}_2(p) = a \left(p - 1/2\right)^2 \quad \text{and} \quad \mathcal{R}_2(p, \dot{p}) = \frac{a}{2} \dot{p}^2$$

for any $a > 0$. Of course, GS that simply differ by a scaling constant such as $a > 0$ are not considered as different. Motivated by a Markovian large-deviation principle, a truly different GS is obtained for $a > 0$ by

$$\mathcal{E}_{\text{Mv}}(p) = a \left(p \log p + (1-p) \log(1-p)\right) \text{ and } \mathcal{R}_{\text{Mv}}^*(p, \xi) = a \sqrt{p(1-p)} \, \mathscr{C}^*(\xi/a),$$

where the function $\mathscr{C}$ and its Legendre dual $\mathscr{C}^*$ are given by

$$\mathscr{C}(v) = 2v \operatorname{arsinh}(v/2) - 2\sqrt{4+v^2} + 4 \quad \text{and} \quad \mathscr{C}^*(\xi) = 4\left(\cosh(\xi/2) - 1\right). \qquad (1.4)$$

The functions $\mathscr{C}$ and $\mathscr{C}^*$ will play a fundamental role, so we give some elementary relations:

$$\mathscr{C}(v) = \tfrac{1}{2}v^2 + O(v^4), \quad \mathscr{C}^*(\xi) = \tfrac{1}{2}\xi^2 + O(\xi^4), \quad \mathscr{C}'(v) = 2\operatorname{arsinh}(v/2),$$

$$\sqrt{pq} \, \mathscr{C}^*(\log p - \log q) = 2\left(\sqrt{p} - \sqrt{q}\right)^2, \qquad \sqrt{pq} \, (\mathscr{C}^*)'(\log p - \log q) = p - q.$$

Indeed, using the last relation and $\text{D}\mathcal{E}_{\text{Mv}}(p) = a\left(\log p - \log(1-p)\right)$ we easily find

$$\dot{p} = \text{D}_\xi \mathcal{R}_{\text{Mv}}\left(p, -\text{D}\mathcal{E}_{\text{Mv}}(p)\right) = 1 - 2p.$$

Moreover, using $(\mathscr{C}^*)'(\xi) = 2\sinh(\xi/2)$ we see that the evolution takes the exponential form

$$\dot{p} = -2\sqrt{p(1-p)} \sinh\left(\frac{1}{2}\text{D}\mathcal{E}_{\text{Mv}}(p)\right).$$

This form is derived and extensively studied in [BoP14], it occurs in mechanics [RRG00, Eqn. (5)] and in chemistry, see the discussion at the end of Section 2.3.2.

The usage of generalized GS is common in the modeling of materials, e.g. for plasticity, ferromagnetism, etc., where the nonsmoothness and nonlinearity of the constitutive law $\dot{u} \mapsto \text{D}_{\dot{u}}\mathcal{R}(u, \dot{u})$ for the dissipative forces is essential, see Section 2.3.1 and the survey [Mie15b]. The mathematical usage of generalized GS in smooth models such as reaction-diffusion equations and systems is rather new. One of the remarkable origins of gradient structures arises from the interpretation of a macroscopic system as a Kolmogorov forward equation

$$\dot{\rho} = \mathbb{Q}^*\rho, \quad \text{where } \rho(t) \in \text{Prob}(S), \qquad (1.5)$$

2

for a Markov process $(X(t))_{t \geq 0}$ on the set S with generator $\mathbb{Q}$. Considering N independent particles $X_j(t)$, $j = 1, \ldots, N$ one can define the empirical process $\rho^N(t) = \frac{1}{N} \sum_{j=1}^N \delta_{X_j(t)} \in$ Prob(S). For $N \to \infty$ the process ρ^N converges to a solution ρ of (1.5). Moreover, according to the program in [AD*11, AD*13] ρ^N satisfies a large-deviation principle that gives rise to a rate functional $\mathscr{I}(\rho(\cdot)) = \int_0^T \mathcal{L}(\rho(t), \dot{\rho}(t)) \, \mathrm{d}t$, where $\mathcal{L}$ can be characterized explicitly by $\mathbb{Q}$. The main observation in [MPR14] is that $\mathcal{L}$ defines a GS (Prob(S), $\mathcal{E}$, $\mathcal{R}$) via the explicit representation

$$\mathcal{L}(\rho, \dot{\rho}) = \mathcal{R}(\rho, \dot{\rho}) + \mathcal{R}^*(\rho, -\mathrm{D}\mathcal{E}(\rho)) + \mathrm{D}\mathcal{E}(\rho)[\dot{\rho}], \tag{1.6}$$

whenever $\mathbb{Q}$ has a unique steady state $\pi \in$ Prob(S) and $\mathbb{Q}$ satisfies the detailed-balance condition with respect to π (i.e. the Markov process is reversible). We refer to Section 2.4 for details, where we also highlight that the arising gradient systems are classical only in the case of diffusion processes. In case of jumps, one obtains generalized GS involving the function $\mathscr{C}$. In particular, for $\dot{p} = 1 - 2p$ one finds ([0,1], $\mathcal{E}_{\mathrm{Mv}}, \mathcal{R}_{\mathrm{Mv}}$) with $a = \frac{1}{2}$.

We consider the above stochastic approach as a first microscopic origin of GS. The second origin involves the concept of evolutionary Γ-convergence for GS, see the surveys [Ser11, Mie15a] for the general ideas. Here we concentrate on convergence results based on the EDP, cf. (1.3), which is an ideal tool for doing a limit passage for solutions $u_\varepsilon : [0, T] \to \mathbf{X}$ for a family $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ of GS depending on a small parameter ε. The aim is then to derive a limiting GS $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ such that a limit u of the solutions u_ε is indeed a solution for the limiting GS. Our Definition 3.2 introduces the concept of EDP-convergence: A family of GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ converges to the GS $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ in the EDP sense, if the following holds:

$$\left. \begin{array}{c} u_\varepsilon : [0, T] \to \mathbf{X} \text{ is a} \\ \text{solution of } (\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon), \\ u_\varepsilon(0) \rightharpoonup u^0, \text{ and} \\ \mathcal{E}_\varepsilon(0, u_\varepsilon(0)) \to \mathcal{E}_0(0, u^0) < \infty \end{array} \right\} \implies \left\{ \begin{array}{c} \exists\, u \text{ sol. of } (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0) \text{ with } u(0) = u^0 \\ \text{and a subsequence } \varepsilon_k \to 0 : \\ \forall\, t \in\,]0, T]: u_{\varepsilon_k}(t) \rightharpoonup u(t) \text{ and} \\ \mathcal{E}_{\varepsilon_k}(u_{\varepsilon_k}(t)) \to \mathcal{E}_0(u(t)); \end{array} \right. \tag{1.7a}$$

$$\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0 \text{ in } \mathbf{X}; \tag{1.7b}$$

$$\left. \begin{array}{c} \widetilde{u}_\varepsilon(\cdot) \xrightarrow{*} \widetilde{u}(\cdot) \text{ in } \mathrm{L}^\infty([0,T]; \mathbf{X}) \text{ and} \\ \sup_{\varepsilon \in\,]0,1],\, t \in [0,T]} \mathcal{E}_\varepsilon(\widetilde{u}_\varepsilon(t)) \leq C < \infty \end{array} \right\} \implies \liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(\widetilde{u}_\varepsilon) \geq \mathscr{D}_0(\widetilde{u}). \tag{1.7c}$$

When asking only for condition (1.7a) we speak of pE-convergence, see Definition 3.1. Note that (1.7c) enforces a liminf estimate of De Giorgi's dissipation functionals $\mathscr{D}_\varepsilon$ along general functions $\widetilde{u}_\varepsilon$, not only along the solutions of the GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$. Having this liminf estimate, it is easy to pass to the limit in the ε-dependent energy-dissipation estimate (1.3), since the initial energy on the right-hand side is assumed to converge according to (1.7a). Then, applying the EDP for the limiting GS $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ we see that u is a solution.

In fact, many approaches to evolutionary Γ-convergence establish EDP-convergence, but do not explicitly state condition (1.7c) as a main result. E.g. the Sandier-Serfaty approach [SaS04, Ser11], where the terms $\int_0^T \mathcal{R}_\varepsilon \, \mathrm{d}t$ and $\int_0^T \mathcal{R}_\varepsilon^* \, \mathrm{d}t$ are treated separately, provides EDP-convergence. Our approach is more general than the latter, since we only ask that the sum $\int_0^T \mathcal{R}_\varepsilon \, \mathrm{d}t + \int_0^T \mathcal{R}_\varepsilon^* \, \mathrm{d}t$ behaves well, but not necessarily the individual terms. This has two effects: (i) we can allow for general functions $\widetilde{u}_\varepsilon$, and (ii) it can lead to exchanges between the two terms in the limit $\varepsilon \to 0$. Point (i) is important to explore $\mathscr{D}_0$ outside of the set of solutions and thus providing the full information

3

about the GS $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$, while the set of solutions of the limit equation $\dot{u} = V_0(t, u) := \mathrm{D}_{\dot{u}} \mathcal{R}_0(u, -\mathrm{D}_u \mathcal{E}_0(t, u))$ only contains information on V_0. Point (ii) is relevant for another important message of this paper. The EDP-limit of classical GS can be a generalized GS. This phenomenon is considered as another microscopic origin of generalized GS.

Here we provide three different examples for point (ii), the first of which is an ODE example in Section 3.3.3, while Sections 4 and 5 contain more elaborate examples treating the membrane limit of a thin-layer and the limit of diffusion to reaction, respectively.

For the membrane limit we consider a diffusion equation with a thin layer with very small diffusivity. In [Lie12, Lie13] pE-convergence to the membrane limit was established; however EDP-convergence was not studied. We start with the diffusion equation which is the gradient flow for the classical GS $(\mathrm{Prob}(\Omega), \mathcal{E}, \mathcal{R}_\varepsilon^*)$ with $\mathcal{E}(u) = \int_\Omega u \log(2u) \, \mathrm{d}x$ and $\mathcal{R}_\varepsilon^*(u, \xi) = \frac{1}{2} \int_\Omega a_\varepsilon(x)(\partial_x \xi)^2 u \, \mathrm{d}x$. Using suitable scalings for the diffusion coefficient $a_\varepsilon(x)$ Theorem 4.1 provides EDP-convergence to the generalized GS $(\mathrm{Prob}(\Omega), \mathcal{E}, \mathcal{R}_0^*)$ with

$$
\mathcal{R}_0^*(u, \xi) = \int_{-1}^0 \frac{a}{2} (\partial_x \xi)^2 u \, \mathrm{d}x + a_* \sqrt{u(0^-)u(0^+)} \, \mathscr{C}^*(\xi(0^+) - \xi(0^-)) + \int_0^1 \frac{a}{2}(\partial_x \xi)^2 u \, \mathrm{d}x,
$$

where $u(0^-)$ and $u(0^+)$ denote the limit of $u(x)$ at $x = 0$ from the left and from the right, respectively. Thus, $\mathcal{R}_0^*$ involves $\mathscr{C}^*$ and is therefore non-quadratic.

Section 5 follows [PSV10, PSV12, AM*12] by considering the limit from pure diffusion in physical space and along a reaction-path variable $y \in \Upsilon \subset \mathbb{R}$ to a limit of a reaction-diffusion system on Ω. The Fokker-Planck equation reads

$$
\dot{u} = m_\Omega \Delta_x u + \tau_\varepsilon \partial_y \Big(\partial_y u + \frac{1}{\varepsilon} u \, \partial_y V(y) \Big),
$$

where V is a potential with two global minima y_0 and y_1 and one global maximum in-between. This equation is generated by the classical GS $(\mathrm{Prob}(\Omega \times \Upsilon), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon^*)$ where $\mathcal{E}_\varepsilon$ is the relative entropy and $\mathcal{R}_\varepsilon^*$ is the quadratic Wasserstein dissipation potential, see (5.2). Theorem 5.2 establishes EDP-convergence to a generalized GS $(\mathrm{Prob}(\Omega \times \{y_0, y_1\}), \mathbf{E}, \mathbf{R}^*)$, where $\mathbf{R}^*$ again involves the non-quadratic function $\mathscr{C}^*$.

We conclude our introduction by a general and surprising observation. The three main models in this work (i.e. the ODE, the membrane, and the reaction-to-diffusion model in Sections 3.3.2, 4, and 5, respectively) can be seen as Kolmogorov forward equations for naturally associated Markov processes. Thus, the large-deviation theory of Section 2.4 is applicable and provides entropic GS $(\mathrm{Prob}(S), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ for the associated Kolmogorov forward equations $\dot{\rho}_\varepsilon = \mathbb{Q}_\varepsilon^* \rho_\varepsilon$ for each $\varepsilon \in \,]0, 1[$ as well as for $\varepsilon = 0$. The limit for $\varepsilon = 0$ can be also defined in terms of the classical convergence for Markov processes asking $\rho_\varepsilon(t) = \mathrm{e}^{t\mathbb{Q}_\varepsilon^*} \rho(0) \stackrel{*}{\rightharpoonup} \rho(t) = \mathrm{e}^{t\mathbb{Q}_0^*} \rho(t)$. Ignoring the (linear) Markovian structure, we can also consider EDP-convergence of the induced entropic GS. In all our three examples we find the surprising result that the EDP-limit is exactly the entropic GS of the limiting Markov process. This means that applying the described large-deviation principle and taking the limit $\varepsilon \to 0$ (either on the level of Markov semigroups or as EDP-convergence for GS) commute, see Figure 1.1. This result appears naturally, if we use representation (1.6) of the rate function $\mathscr{I}$ giving

$$
\mathscr{I}(\rho) = \int_0^T \mathcal{R}(\rho, \dot{\rho}) + \mathcal{R}^*(\rho, -\mathrm{D}\mathcal{E}(\rho)) + \mathrm{D}\mathcal{E}(\rho)[\dot{\rho}] \, \mathrm{d}t = \mathscr{D}(\rho) + \mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)).
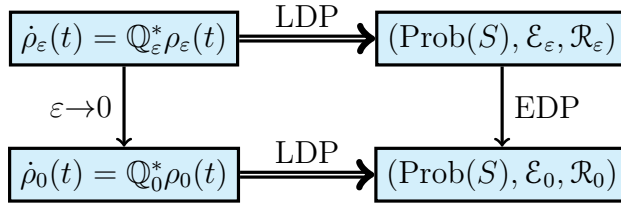$$

Figure 1.1: For reversible, time-continuous Markov processes the large-deviation principle (LPD) of Section 2.4 provides a (generalized) gradient structure. This mapping commutes with taking the limit $\varepsilon \to 0$ and EDP-convergence, respectively.

Hence, the above large-deviation principle exactly encodes the energy-dissipation principle, and EDP-convergence for the induced entropic GS can be interpreted as Γ-convergence of the rate functionals.

The question how general this observation about the interchangeability of the suitable large-deviation principles and the EDP-convergence is, seems to be challenging, but goes beyond the scope of this work. We mention that in [BoP14] similar relations between large-deviation principles and evolutionary Γ-convergence are studied.

As a final general remark, we emphasize that this paper focuses on the modeling aspects of the emergence of generalized GS. Thus, we do not give the full analytical details in terms of estimates and convergences in the proper functional spaces, but rather highlight the structures and manipulations needed to understand the corresponding limit procedures.

2 Classical and generalized gradient systems

We now convert the formal ideas from the introduction into rigorous mathematical statements. We call a triple $(\mathbf{X}, \mathcal{E}, \mathcal{R})$ a gradient system (GS), if $\mathbf{X}$ is a Banach space, $\mathcal{E} : [0,T] \times \mathbf{X} \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ is a functional (such as the free energy, the negentropy, etc.), and $\mathcal{R} : \mathbf{X} \times \mathbf{X} \to [0,\infty]$ is a dissipation potential, which means that for all $q \in \mathbf{X}$ the functional $\mathcal{R}(u, \cdot) : \mathbf{X} \to \mathbb{R}_\infty$ is lower semicontinuous, nonnegative, convex, and satisfies $\mathcal{R}(u, 0) = 0$. In this section, we allow for the case that the energy functional depends on the time variable $t \in [0,T]$ to show that the abstract principle is valid in this general case. However, for notational convenience we will restrict to the autonomous case (i.e. $\partial_t \mathcal{E}(t, u) \equiv 0$) in all other parts.

We speak of a *classical GS*, if $\mathcal{R}(u, \cdot)$ is quadratic, i.e. there exists a symmetric and positive definite operator $\mathbb{G}$ such that $\mathcal{R}(u, v) = \frac{1}{2}\langle \mathbb{G}(u)v, v \rangle$. However, plasticity requires non-quadratic dissipation potentials, e.g. of the form $\mathcal{R}(\dot{\pi}) = \sigma_{\text{yield}}\|\dot{\pi}\|_{L^1} + \frac{1}{2}\mu_{\text{visc}}\|\dot{\pi}\|_{L^2}^2$, see [Mie03, MiR15]. In particular, the rate-independent case is based on $\mathcal{R}(u, \lambda v) = \lambda \mathcal{R}(u, v)$ for all $\lambda > 0$, which is incompatible with a quadratic form. If $\mathcal{R}(u, \cdot)$ is non-quadratic, we call $(\mathbf{X}, \mathcal{E}, \mathcal{R})$ a *generalized GS*.

2.1 Variational principles for gradient systems

The following proposition from convex analysis shows that there are several completely equivalent formulations of the generalized force balance (1.1). The equivalences of the points (ii) to (iv) below are also called Fenchel equivalences, cf. [Fen49]. The essential

tool is the Legendre-Fenchel transform $\Psi^* : \mathbf{X}^* \to \mathbb{R}_\infty$ of a convex function $\Psi : \mathbf{X} \to \mathbb{R}_\infty$ defined via

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

In a reflexive Banach space we have $(\Psi^*)^* = \Psi$.

Proposition 2.1 (Equivalent formulations) *Let $\mathbf{X}$ be a reflexive Banach space and $\Psi : \mathbf{X} \to \mathbb{R}_\infty$ be proper, convex, and lower semicontinuous. Then, for every $\xi \in \mathbf{X}^*$ and every $v \in \mathbf{X}$ the following five statements are equivalent:*

$$(i) \quad v \in \operatorname*{Arg\,min}_{w \in \mathbf{X}} \big(\Psi(w) - \langle \xi, w \rangle\big); \qquad (ii)\ \xi \in \partial\Psi(v);$$

$$(iii) \quad \Psi(v) + \Psi^*(\xi) = \langle \xi, v \rangle;$$

$$(iv) \quad v \in \partial\Psi^*(\xi); \qquad\qquad (v)\ \ \xi \in \operatorname*{Arg\,min}_{\eta \in \mathbf{X}^*} \big(\Psi^*(\eta) - \langle \eta, v \rangle\big).$$

Note that the definition of Ψ^* immediately implies the Young-Fenchel inequality $\Psi(w) + \Psi^*(\eta) \geq \langle \eta, w \rangle$ for all w and η. Thus, (iii) expresses an optimality as well.

Defining the dual dissipation potential $\mathcal{R}^*$ via $\mathcal{R}^*(u, \cdot) := (\mathcal{R}(u, \cdot))^*$ we can apply these equivalences to reformulate (1.1) in the following ways:

(I) Rayleigh principle [Ray71]

$$(RP) \quad \dot{u} \in \operatorname*{Arg\,min}_{v \in \mathbf{X}} \Big(\mathcal{R}(u, v) - \langle \mathrm{D}\mathcal{E}(t, u), v \rangle\Big);$$

(II) Force balance in $\mathbf{X}^*$ Rayleigh-Biot equation [Ray71, Bio55]

$$(FB) \quad 0 \in \partial_{\dot{u}} \mathcal{R}(u, \dot{u}) + \mathrm{D}\mathcal{E}(t, u) \ \in \mathbf{X}^*;$$

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

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

(IV) Rate equation in $\mathbf{X}$ Onsager equation [Ons31]

$$(RE) \quad \dot{u} \in \partial_\xi \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(t, u)) \ \in \mathbf{X};$$

(V) Maximum dissipation principle   cf. e.g. [HaF08]

$$(MDP) \quad \mathrm{D}\mathcal{E}(t, u) \in \operatorname*{Arg\,max}_{\xi \in \mathbf{X}^*} \Big(\langle \xi, \dot{u} \rangle - \mathcal{R}^*(u, \xi)\Big).$$

In fact, [Ray71, Eqn. (26)] also includes the kinetic energy $\mathcal{T}$, which we omit in our approximation, namely $\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{D}_{\dot{u}}\mathcal{T}(u, \dot{u})\big) + \mathrm{D}_{\dot{u}}\mathcal{R}(u, \dot{u}) + \mathrm{D}_q\mathcal{E}(t, u) = 0$.

Before returning to the general situation, we highlight the three different cases (II)–(IV) for the classical viscous dissipation, 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 $\mathbb{K}(u) = \mathbb{G}(u)^{-1}$. Then, we have

$$(FB) \quad \mathbb{G}(u)\dot{u} = -\mathrm{D}\mathcal{E}(u) \qquad (RE) \quad \dot{u} = -\mathbb{K}(u)\mathrm{D}\mathcal{E}(u) =: -\nabla_\mathbb{G}\mathcal{E}(u)$$

$$(PB) \quad \frac{1}{2}\langle \mathbb{G}(u)\dot{u}, \dot{u} \rangle + \frac{1}{2}\langle \mathrm{D}\mathcal{E}(u), \mathbb{K}(u)\mathrm{D}\mathcal{E}(u) \rangle = -\langle \mathrm{D}\mathcal{E}(u), \dot{u} \rangle,$$

where (RE) can be seen as a *"gradient-flow equation"*, as $\nabla_\mathbb{G}$ is the gradient operator.

2.2 The energy-dissipation principle

The above formulations can already be understood in a variational sense, since the evolution is expressed by extremizing a functional or by variations or derivatives of the two functionals $\mathcal{E}$ and $\mathcal{R}$. However, for mathematical purposes it is desirable to have formulations in terms of a minimization problem for the whole solution trajectories $u : [0, T] \to \mathbf{X}$. One such principle can be derived on the basis of the power balance (PB) by integration in time and using the chain rule and finally employing the Young-Fenchel inequality $\Psi(w) + \Psi^*(\eta) \geq \langle \eta, w \rangle$, cf. [DMT80] or the survey [Mie15a]. This leads to the celebrated energy-dissipation principle, also called De Giorgi's $(\mathcal{R}, \mathcal{R}^*)$ principle, see [AGS05] for some historical remarks.

Theorem 2.2 (De Giorgi's energy-dissipation principle) *Under suitable technical conditions on $(\mathbf{X}, \mathcal{E}, \mathcal{R})$ a function $u : [0, T] \to \mathbf{X}$ satisfies (I)–(V) from above for almost all $t \in [0, T]$ if and only if the* energy-dissipation balance (EDB) *holds:*

$$\text{(EDB)} \quad \begin{cases} \mathcal{E}(T, u(T)) + \displaystyle\int_0^T \mathcal{R}(u, \dot{u}) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(t, u)) \, \mathrm{d}t \\ \qquad\qquad = \ \mathcal{E}(0, u(0)) + \displaystyle\int_0^T \partial_t \mathcal{E}(t, u(t)) \, \mathrm{d}t. \end{cases}$$

Under additional technical conditions it is sufficient to have only the upper estimate where "=" is replaced by "≤". In this case, we speak of the energy-dissipation estimate (EDE).

2.3 Examples of generalized gradient structures

Here we give some examples of generalized gradient structures. First, we discuss dissipative material models like plasticity or shape-memory materials that form a huge class of generalized GS. Second, we treat nonlinear reaction-diffusion systems (RDS), which will be closer to the main theme of this paper. The third class of examples concerns reversible Markov processes, where the Kolmogorov forward equation has a gradient structure with the relative entropy as energy functional. This latter class is so important that it is treated in the subsequent Subsection 2.4.

2.3.1 Dissipative material models

The state of a body $\Omega \subset \mathbb{R}^d$, composed of so-called dissipative materials (also called standard generalized materials), is given in terms of the elastic displacement $\mathbf{u} : \Omega \to \mathbb{R}^d$ and an additional internal variable $z : \Omega \to \mathbb{R}^k$. The latter may describe plastic deformations, damage, phase-field variables, magnetization, or other internal states of the material. The total stored energy $\mathcal{E}$ depends on $\mathbf{u}$, z, and usually also on a process time $t \in [0, T]$, i.e. $\mathcal{E}(t, \mathbf{u}, z)$. As introduced in the theory of standard generalized materials in [HaN75] the dissipative forces are given in terms of a (primal) dissipation potential $\mathcal{R}$ that also may include viscoelastic terms:

$$\mathcal{R}(\dot{\mathbf{u}}, \dot{z}) = \mathcal{R}_{\mathrm{diss}}(\dot{z}) + \mathcal{R}_{\mathrm{visc}}(e(\dot{\mathbf{u}})),$$

where $e(\dot{\mathbf{u}}) = \frac{1}{2}(\nabla \dot{\mathbf{u}} + (\nabla \dot{\mathbf{u}})^\top)$. As before, the corresponding force balance equations (FB) are

$$0 = \mathrm{D}_{\dot{\mathbf{u}}} \mathcal{R}(\dot{\mathbf{u}}, \dot{z}) + \mathrm{D}_{\mathbf{u}} \mathcal{E}(t, \mathbf{u}, z), \quad 0 \in \partial_{\dot{z}} \mathcal{R}(\dot{\mathbf{u}}, \dot{z}) + \mathrm{D}_z \mathcal{E}(t, \mathbf{u}, z)$$

with $\partial_{\dot{z}}\mathcal{R}(\dot{\mathbf{u}}, \dot{z})$ denoting the set-valued convex subdifferential.

While the viscoelastic potential $\mathcal{R}_{\mathrm{visc}}$ is assumed to be quadratic in many applications, the potential $\mathcal{R}_{\mathrm{diss}}$ for the internal variables z is often supposed to be non-quadratic. E.g. in viscoplasticity with yields stress σ_{yield} one takes the form

$$\mathcal{R}_{\mathrm{diss}}(\dot{z}) = \int_{\Omega} \left(\sigma_{\mathrm{yield}} |\dot{z}|_1 + \nu |\dot{z}|^{1+\delta} \right) \mathrm{d}x,$$

where $\delta > 0$ is usually taken small, e.g. in $\delta = 0.012$ in [ZR*06]. The weak growth of order $1+\delta$ is sometimes even replaced by a growth $O(|\dot{z}| \log |\dot{z}|)$ as given by our function $\mathscr{C}$ (see e.g. [RRG00, Eqn. (5)] and [BoP14]).

For later reference, we mention the very simple scalar hysteresis model of a so-called play operator. It is given by the generalized GS $(\mathbb{R}, \mathcal{E}_{\mathrm{play}}, \mathcal{R}_{\mathrm{play}})$ with

$$\mathcal{E}_{\mathrm{play}}(t, z) = \frac{1}{2}z^2 - \ell(t)z \quad \text{and} \quad \mathcal{R}_{\mathrm{play}}(\dot{z}) = r|\dot{z}| \text{ with } r > 0. \tag{2.1}$$

It serves as a limit for evolutionary Γ-convergence in Example 3.3 as well as a large-deviation limit in [BoP14].

2.3.2 Nonlinear reaction-diffusion systems

We consider concentrations $\mathbf{c}(t) : \Omega \to [0, \infty[^I$ of chemical species $C_1, \dots C_I$ that can react according to R reactions of mass action type given by a stoichiometric relation

$$\alpha_i^r C_1 + \dots + \alpha_I^r C_I \underset{k_r^{\mathrm{b}}}{\overset{k_r^{\mathrm{f}}}{\rightleftharpoons}} \beta_1^r C_1 + \dots + \beta_I^r C_I,$$
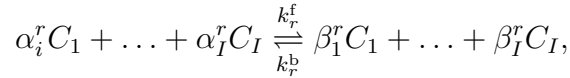
where $r = 1, \dots, R$ is the index of the reaction, k_r^{f} and k_r^{b} are the forward and backward reaction coefficients, and the stoichiometric coefficients α_i^r and β_i^r are nonnegative integers. The reaction-diffusion system (RDS) for the concentrations $\mathbf{c} = (c_1, \dots, c_I)$ takes the form

$$\dot{\mathbf{c}} = \mathbb{D}\Delta\mathbf{c} - \mathbf{R}(\mathbf{c}) \quad \text{with } \mathbf{R}(\mathbf{c}) := \sum_{r=1}^{R} \left(k_r^{\mathrm{f}} \mathbf{c}^{\boldsymbol{\alpha}^r} - k_r^{\mathrm{b}} \mathbf{c}^{\boldsymbol{\beta}^r} \right) \left(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r \right) \tag{2.2}$$

and $\mathbb{D} = \mathrm{diag}(\delta_i)_{i=1,\dots,I}$, where $\delta_i > 0$. With the stoichiometric vectors $\boldsymbol{\alpha}^r = (\alpha_i^r)_i$ and $\boldsymbol{\beta}^r = (\beta_i^r)_i \in \mathbb{N}_0^I$ we define the monomials in the form $\mathbf{c}^{\boldsymbol{\alpha}} := \prod_{i=1}^{I} c_i^{\alpha_i}$.

It was shown in [Mie11, Mie13b] that (2.2) has a (classical) gradient structure under the additional assumption of the detailed-balance condition, which means that

$$\exists \mathbf{w} = (w_i)_i : \quad w_i > 0 \quad \text{and} \quad k_r^{\mathrm{f}} \mathbf{w}^{\boldsymbol{\alpha}^r} = k_r^{\mathrm{b}} \mathbf{w}^{\boldsymbol{\beta}^r} \text{ for } r = 1, \dots, R. \tag{2.3}$$

Using the Boltzmann function $\lambda_{\mathrm{B}}(z) = z \log z - z + 1$ we define the relative entropy

$$\mathcal{E}(\mathbf{c}) = \int_{\Omega} \sum_{i=1}^{I} \lambda_{\mathrm{B}}(c_i(x)/w_i) w_i \, \mathrm{d}x,$$

which gives rise to the vector of thermodynamic driving forces (also called chemical potentials) $\boldsymbol{\mu} = \mathrm{D}\mathcal{E}(\mathbf{c})$ with $\mu_i = \log(c_i/w_i)$. Because of the logarithm laws they satisfy

the relation $\boldsymbol{\alpha}^r \cdot D\mathcal{E}(\mathbf{c}) = \sum_{i=1}^{I} \alpha_i^r \log(c_i/w_i) = \log(\mathbf{c}^{\boldsymbol{\alpha}^r}) - \log(\mathbf{w}^{\boldsymbol{\alpha}^r})$. Thus, using the detailed-balance conditions we obtain the relation

$$\left(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\right) \cdot \left(-D\mathcal{E}(\mathbf{c})\right) = \log(k_r^{\mathrm{b}}\mathbf{c}^{\boldsymbol{\beta}^r}) - \log(k_r^{\mathrm{f}}\mathbf{c}^{\boldsymbol{\alpha}^r}). \tag{2.4}$$

To construct the dual dissipation potentials we may choose any scalar, strictly convex dual dissipation functional $\psi : \mathbb{R} \to \mathbb{R}$ with $\psi(0) = \psi'(0) = 0$ and $\psi''(0) > 0$ and let

$$\mathcal{R}^*(\mathbf{c}, \boldsymbol{\mu}) := \int_\Omega \left[\frac{1}{2} \sum_{i=1}^{I} \delta_i c_i |\nabla \mu_i|^2 \;+\; \sum_{r=1}^{R} H_\psi^r(\mathbf{c}) \psi\big((\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r)\cdot\boldsymbol{\mu}\big) \right] \mathrm{d}x$$

$$\text{with } H_\psi^r(\mathbf{c}) := \frac{k_r^{\mathrm{b}}\mathbf{c}^{\boldsymbol{\beta}^r} \;-\; k_r^{\mathrm{f}}\mathbf{c}^{\boldsymbol{\alpha}^r}}{\psi'\big(\log(k_r^{\mathrm{b}}\mathbf{c}^{\boldsymbol{\beta}^r}) - \log(k_r^{\mathrm{f}}\mathbf{c}^{\boldsymbol{\alpha}^r})\big)}.$$

Using $D_{\boldsymbol{\mu}}\mathcal{R}(\mathbf{c}, \boldsymbol{\mu}) = -\big(\mathrm{div}(\delta_i c_i \nabla \mu_i)\big)_i + \sum_{r=1}^{R} H_\psi^r(\mathbf{c})\psi'\big((\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r)\cdot\boldsymbol{\mu}\big)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big)$ and (2.4), we easily see that the nonlinear RDS (2.2) satisfying the detailed-balance condition is generated by the pair $\mathcal{E}$ and $\mathcal{R}$, i.e. $\dot{\mathbf{c}} = \mathbb{D}\Delta\mathbf{c} - \mathbf{R}(\mathbf{c}) = D_{\boldsymbol{\mu}}\mathcal{R}\big(\mathbf{c}, -D\mathcal{E}(\mathbf{c})\big)$. Thus, we have found a family of generalized gradient structures of the nonlinear RDS (2.2).

The case of quadratic $\mathcal{R}(\mathbf{c}; \cdot)$, i.e., $\psi(\eta) = \eta^2/2$ was introduced in [Mie11], leading to the logarithmic means $H_{\mathrm{quadr}}^r(\mathbf{c}) = \Lambda(k_r^{\mathrm{f}}\mathbf{c}^{\boldsymbol{\alpha}^r}, k_r^{\mathrm{b}}\mathbf{c}^{\boldsymbol{\beta}^r})$ with $\Lambda(a, b) = (a-b)/(\log a - \log b)$, see also [Maa11, ErM12, MaM15a]. However, it was already criticized in the 1930s that the linear relation $\dot{\mathbf{c}} = -\mathbb{K}(\mathbf{c})\boldsymbol{\mu}$ (i.e. $\mathcal{R}^*(\mathbf{c}, \boldsymbol{\mu}) = \frac{1}{2}\langle \boldsymbol{\mu}, \mathbb{K}(\mathbf{c})\boldsymbol{\mu}\rangle$ is quadratic) arising from Onsager's principle is not suitable for chemical reactions if one wants to model systems that are not very close to thermal equilibrium. As a solution Marcelin and de Donder introduced exponential dependencies between $\boldsymbol{\mu}$ and $\dot{\mathbf{c}}$, see [Fei72, Def. 3.3] or [GK*00, Eqn. (11)]. In [Grm10] Remark iii on p. 77 gives some historical comments and Eqn. (69) explicitly features an exponential dissipation potential Ξ involving the function $\big(\mathrm{e}^{\xi/2} + \mathrm{e}^{-\xi/2} - 2\big)$. Since the choice $\psi(\xi) = \mathscr{C}^*(\xi)$ is central for our paper, we give $\mathcal{R}^*$ explicitly for this case, viz.

$$\mathcal{R}^*(\mathbf{c}, \boldsymbol{\mu}) = \int_\Omega \left[\sum_{i=1}^{I} \frac{\delta_i c_i}{2} |\nabla c_i|^2 + \sum_{r=1}^{R} \sqrt{k_r^{\mathrm{f}}\mathbf{c}^{\boldsymbol{\alpha}^r} k_r^{\mathrm{b}}\mathbf{c}^{\boldsymbol{\beta}^r}} \, \mathscr{C}^*\big((\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r) \cdot \boldsymbol{\mu}\big) \right] \mathrm{d}x. \tag{2.5}$$

We will see that exactly the same structure, up to a trivial scaling factor $1/2$, arises via the large-deviation principle described next, see also [MP*15].

2.4 Markov processes, large deviations, and GS

Here we give a rough sketch of the theory in [MPR14] about gradient structures for the Kolmogorov forward equation $\dot{\rho} = \mathbb{Q}^*\rho$ of Markov processes satisfying a detailed-balance condition, which are also called reversible Markov processes, for short. The idea that large-deviation principles generate gradient structures goes back to [OnM53] (see Eqn. (4-21) therein for a quadratic version of the energy-dissipation principle derived by large-deviations, called Boltzmann's principle). The mathematical theory was developed only recently, see [AD*11, AD*13, MPR14].

In Section 2.4.1 we first describe a time-dependent large-deviation principle for general Markov processes providing a formula for the rate function $\mathscr{I}(\rho(\cdot)) = \int_0^T \mathcal{L}(\rho(t), \dot{\rho}(t)) \, \mathrm{d}t$ and then present the result of [MPR14], which shows that for reversible Markov processes the functional $\mathcal{L}$ is induced by an EDP for a GS $(\mathrm{Prob}(S), \mathcal{E}, \mathcal{R})$. In Sections 2.4.2 to 2.4.4 we then discuss a few applications of the abstract result in Theorem 2.3.

2.4.1 Gradient structures obtained via large deviations

We consider a compact metric space S and denote by $\mathrm{Prob}(S)$ the subset of probability Radon measures on S equipped with the narrow convergence $\overset{*}{\rightharpoonup}$ defined by duality with continuous, bounded functions. The Kolmogorov forward equation $\dot{\rho} = \mathbb{Q}^*\rho$ describes the evolution of the probability laws $\rho(t)$ of a Markov process $(X_t)_{t\geq 0}$, if the law of X_0 is given by $\rho(0) \in \mathrm{Prob}(S)$. The Markov generator is given as $\mathbb{Q}$ acting on functions on S, while its dual $\mathbb{Q}^*$ acts on measures such that $\int_S (\mathbb{Q}f)\,\mathrm{d}\rho = \int_S f\,\mathrm{d}(\mathbb{Q}^*\rho)$.

Considering N independent realizations $(X_t^{(i)})_{t\geq 0}$, $i = 1,\ldots,N$ of the underlying Markov process, the measure-valued empirical process $\rho^N(t) := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{(i)}} \in \mathrm{Prob}(S)$ can be defined. Using the law of large numbers the limit $N \to \infty$ gives $\rho^N(t) \overset{*}{\rightharpoonup} \rho(t)$, which solves the Kolmogorov forward equation $\dot{\rho} = \mathbb{Q}^*\rho$, see e.g. [Ren13, Thm. 2.3.1]. Under suitable assumptions, see [MPR14], it is shown in [FeK06] that the empirical process ρ^N satisfies a large-deviation principle with a rate function $\mathscr{I}(\rho(\cdot))$, i.e.

$$\mathbb{P}\big(\rho^N(\cdot) \approx \widehat{\rho}(\cdot)\big) \;\simeq\; \mathrm{e}^{-N\mathscr{I}(\widehat{\rho}(\cdot))},$$

see the above references for the proper definition of "$\simeq$". The main result is that $\mathscr{I}$ has the form $\mathscr{I}(\rho(\cdot)) = \int_0^T \mathcal{L}(\rho(t), \dot{\rho}(t))\,\mathrm{d}t$, where $\mathcal{L}$ is the Legendre transform of the functional $\mathcal{H}(\rho, \cdot)$ (i.e. $\mathcal{L}(\rho, v) = \sup \int_\Omega \xi\mathrm{d}v - \mathcal{H}(\rho, \xi)$) given via the explicit formula

$$\mathcal{H}(\rho, \xi) := \int_S \mathrm{e}^{-\xi(s)}(\mathbb{Q}\,\mathrm{e}^\xi)(s)\,\mathrm{d}\rho(s).$$

We emphasize the simplicity of this formula and the (separate) linearity in ρ and in $\mathbb{Q}$.

A main observation in [MPR14] is that the deterministic case, which is given by the relation

$$\mathscr{I}(\rho(\cdot)) = \int_0^T \mathcal{L}(\rho(t), \dot{\rho}(t))\,\mathrm{d}t = 0,$$

can be interpreted as an energy-dissipation principle if and only if the Markov process is reversible, which is the same as asking for the detailed balance condition (cf. (2.3)) for the linear Kolmogorov forward equation $\dot{\rho} = \mathbb{Q}^*\rho$. Hence, we now further assume that there exists a stationary measure $\pi \in \mathrm{Prob}(S)$ which has, without loss of generality, the full set S as its support. We say that $\mathbb{Q}$ *satisfies the detailed balance condition with respect to π,* if

$$\int_S f\,(\mathbb{Q}\,g)\,\mathrm{d}\pi = \int_S g\,(\mathbb{Q}\,f)\,\mathrm{d}\pi \tag{2.6}$$

for all f and g in the domain if $\mathbb{Q}$. Choosing $g \equiv 1$, we find that $\mathbb{Q}^*\pi = 0$, i.e. the detailed balance condition implies the stationarity.

This version of the detailed-balance condition for Markov processes coincides with the detailed-balance condition for chemical reactions in (2.3). Indeed, if the ODE case $\dot{\mathbf{c}} = -\mathbf{R}(\mathbf{c})$ is linear, i.e. $\dot{\mathbf{c}} = \mathbb{A}\mathbf{c}$ with $\mathbb{A}\mathbf{w} = 0$, then (2.3) means $\mathbb{A}_{ij}w_j = \mathbb{A}_{ji}w_i$. Setting $S = \{1,\ldots,I\}$ and $\mathbb{Q} = \mathbb{A}^*$ gives (2.6).

In the sequel we will use the Radon-Nikodym derivative of ρ with respect to π denoted by $f = \frac{\mathrm{d}\rho}{\mathrm{d}\pi} \in \mathrm{L}^1_{\geq 0}(S, \pi)$ and defined via $\int_B 1\,\mathrm{d}\rho = \int_B f\,\mathrm{d}\pi$ for all Borel sets $B \subset S$.

Theorem 2.3 ([MPR14, Sec. 3]) *If the Markov process $(X_t)_{t\geq 0}$ on S is reversible, i.e. a stationary measure $\pi \in \mathrm{Prob}(S)$ satisfying the detailed-balance condition (2.6) exists*

for the Kolmogorov forward equation $\dot{\rho} = \mathbb{Q}^*\rho$, then the large-deviation rate functional $\int_0^T \mathcal{L}(\rho, \dot{\rho}) \mathrm{d}t$ has the form of an energy-dissipation principle, namely

$$\mathcal{L}(\rho, v) = \mathcal{R}(\rho, v) + \mathcal{R}^*(\rho, -\mathrm{D}\mathcal{E}(\rho)) + \int_S \mathrm{D}\mathcal{E}(\rho) \mathrm{d}v,$$

where the gradient structure $(\mathrm{Prob}(S), \mathcal{E}, \mathcal{R}^*)$ is given by $\mathcal{E}(\rho) = \frac{1}{2} \int_S \lambda_B\left(\frac{\mathrm{d}\rho}{\mathrm{d}\pi}\right) \mathrm{d}\pi$ and the dual dissipation potential

$$\mathcal{R}^*(\rho, \xi) = \int_S \left((\sqrt{f}\, \mathrm{e}^{-\xi})(\mathbb{Q}(\sqrt{f}\, \mathrm{e}^{\xi})) - \sqrt{f}(\mathbb{Q}\sqrt{f}) \right) \mathrm{d}\pi, \quad \text{where } f = \frac{\mathrm{d}\rho}{\mathrm{d}\pi}. \qquad (2.7)$$

The cited reference contains not only a full proof, but also specifies under what assumptions this implication is in fact an equivalence, i.e. the existence of a gradient structure implies the existence of a steady state satisfying the detailed-balance condition.

Since the arguments and proofs in [MPR14] are quite involved, we highlight here the main structures and formal calculations to see that $(\mathrm{Prob}(S), \mathcal{E}, \mathcal{R}^*)$ is a GS and that it generates the Kolmogorov equation $\dot{\rho} = \mathbb{Q}^*\rho$.

We first observe that $\mathcal{R}^*$ is defined in terms of $\mathcal{H}$ via

$$\mathcal{R}^*(\rho, \xi) = \mathcal{H}\left(\rho, \xi + \frac{1}{2} \log f\right) - \mathcal{H}\left(\rho, \frac{1}{2} \log f\right), \quad \text{where } \frac{1}{2} \log f = \mathrm{D}\mathcal{E}(\rho) = \frac{1}{2} \log \frac{\mathrm{d}\rho}{\mathrm{d}\pi}.$$

Obviously, $\mathcal{R}^*(\rho, \cdot)$ is convex if and only if $\mathcal{H}(\rho, \cdot)$ is convex. The latter is independent of the detailed-balance condition and can be established as follows. Consider the Markov semigroup $P_t = \mathrm{e}^{t\mathbb{Q}}$ for $t \geq 0$. For fixed $\rho \in \mathrm{Prob}(S)$ and $t \geq 0$ define

$$\mathcal{A}_t(\xi) := \int_S \mathrm{e}^{-\xi} P_t(\mathrm{e}^{\xi}) \mathrm{d}\rho = \int_{S \times S} \mathrm{e}^{-\xi(x) + \xi(y)} p_t(x, \mathrm{d}y) \rho(\mathrm{d}x),$$

where p_t with $p_t(x, \cdot) \in \mathrm{Prob}(S)$ denotes the time-dependent Markov kernel. From the convexity of $\xi \mapsto \mathrm{e}^{-\xi(x) + \xi(y)}$ and the nonnegativity of p_t and ρ, we conclude that $\xi \mapsto \mathcal{A}_t(\xi)$ is convex. Using $\frac{1}{t}(P_t \eta - \eta) \to \mathbb{Q}\eta$ and $\mathcal{A}_0(\xi) \equiv 1$, we see that

$$\mathcal{H}(\rho, \xi) = \lim_{t \to 0} \frac{1}{t}(\mathcal{A}_t(\xi) - \mathcal{A}_0(\xi)) = \lim_{t \to 0} \frac{1}{t}(\mathcal{A}_t(\xi) - 1)$$

is also convex in ξ.

By definition we have $\mathcal{R}^*(\rho, 0) = 0$, and the detailed-balance condition implies the time reversibility $\mathcal{R}^*(\rho, -\xi) = \mathcal{R}^*(\rho, \xi)$. Since this implies $\mathrm{D}\mathcal{R}^*(\rho, 0) = 0$, convexity gives the positivity $\mathcal{R}^*(\rho, \xi) \geq 0$. Thus, $\mathcal{R}^*$ is indeed a dual dissipation potential.

To derive the induced gradient-flow evolution for the GS $(\mathrm{Prob}(S), \mathcal{E}, \mathcal{R}^*)$ we observe

$$\mathrm{D}_\xi \mathcal{R}^*(\rho, -\mathrm{D}\mathcal{E}(\rho))[\eta] = \mathrm{D}_\xi \mathcal{H}(\rho, 0)[\eta] = \int_S \left(\mathrm{e}^0(-\eta)\mathbb{Q}(\mathrm{e}^0) + \mathrm{e}^0 \mathbb{Q}(\mathrm{e}^0 \eta) \right) \mathrm{d}\rho = \int_S \eta \, \mathrm{d}(\mathbb{Q}^*\rho),$$

where we used $\mathbb{Q}1 \equiv 0$. This provides $\dot{\rho} = \mathrm{D}_\xi \mathcal{R}^*(\rho, -\mathrm{D}\mathcal{E}(\rho)) = \mathbb{Q}^*\rho$, which is the expected Kolmogorov forward equation.

2.4.2 A finite-state Markov process

We consider the finite state space $S = \{1, \ldots, I\}$ such that

$$\mathrm{Prob}(S) = \big\{\, \rho = \mathbf{c} = (c_1, \ldots, c_I) \in [0,1]^I \ \big|\ \textstyle\sum_{i=1}^{I} c_i = 1 \,\big\}.$$

The Kolmogorov forward equation is the ODE

$$\dot{\mathbf{c}} = \mathbb{A}\mathbf{c} \quad \text{with } \mathbb{A} \in \mathbb{R}^{I \times I}.$$

Note that the Markov generator is given by $\mathbb{Q}_{\mathrm{finite}} = \mathbb{A}^\top$, and the conditions for a Markov generator are

$$\mathbb{A}_{ij} \geq 0 \text{ for all } i \neq j \quad \text{and} \quad \forall\, i = 1, \ldots, I : \ 0 = \sum_{j=1}^{I} \mathbb{A}_{ji}.$$

We further assume that there is a unique positive steady state $\pi = \mathbf{w} \in \mathrm{Prob}(S)$ such that the detailed-balance condition holds, namely $\mathbb{A}_{ij} w_j = \mathbb{A}_{ji} w_i$.

Thus, the induced energy functional is $\mathcal{E}_{\mathrm{finite}}(\mathbf{c}) = \frac{1}{2} \sum_{i=1}^{I} c_i \lambda_{\mathrm{B}}(c_i/w_i)$. To calculate the dissipation potential we use that $\mathbb{Q}$ can be split

$$\mathbb{Q}_{\mathrm{finite}} = \sum_{i=1}^{I-1} \sum_{j=i+1}^{I} \mathbb{Q}^{i \leftrightarrow j} \text{ with } \mathbb{Q}^{i \leftrightarrow j} := m_{ij}\Big(\frac{1}{w_j}\mathbf{e}^j \otimes \mathbf{e}^i + \frac{1}{w_i}\mathbf{e}^i \otimes \mathbf{e}^j\Big),$$

where $m_{ij} := \mathbb{A}_{ij} w_j = \mathbb{A}_{ji} w_i$ and $\mathbf{e}^k$ denotes the kth unit vector in $\mathbb{R}^I$. Using the linearity in $\mathbb{Q}_{\mathrm{finite}}$ of the formula (2.7) for $\mathcal{R}_{\mathrm{finite}}$ we can first calculate

$$\mathcal{R}^*_{i \leftrightarrow j}(\mathbf{c}, \boldsymbol{\xi}) = \sum_{k,l=1}^{I} \left[\Big(\frac{c_k}{w_k}\Big)^{1/2} \mathrm{e}^{-\xi_k} \mathbb{Q}^{i \leftrightarrow j}_{kl} \Big(\frac{c_l}{w_l}\Big)^{1/2} \mathrm{e}^{\xi_l} - \Big(\frac{c_k}{w_k}\Big)^{1/2} \mathbb{Q}^{i \leftrightarrow j}_{kl} \Big(\frac{c_l}{w_l}\Big)^{1/2} \right] w_k$$

$$= m_{ij} \Big(\frac{c_i c_j}{w_i w_j}\Big)^{1/2} \big(\mathrm{e}^{\xi_i - \xi_j} + \mathrm{e}^{\xi_j - \xi_i} - 2\big).$$

Summing these terms and using the function $\mathscr{C}^*$ we find

$$\mathcal{R}^*_{\mathrm{finite}}(\mathbf{c}, \boldsymbol{\xi}) = \frac{1}{2} \sum_{i=1}^{I-1} \sum_{j=i+1}^{I} m_{ij} \Big(\frac{c_i c_j}{w_i w_j}\Big)^{1/2} \mathscr{C}^*\big(2(\xi_i - \xi_j)\big)$$

and conclude by Theorem 2.3 that the equation $\dot{\mathbf{c}} = \mathbb{Q}_{\mathrm{finite}}^\top \mathbf{c}$ is induced by the GS $(\mathrm{Prob}(S), \mathcal{E}_{\mathrm{finite}}, \mathcal{R}_{\mathrm{finite}})$.

2.4.3 Linear reaction-diffusion systems

We now return to RDS as discussed in Section 2.3.2, but now consider only linear reactions where all stoichiometric vectors $\boldsymbol{\alpha}^r$ and $\boldsymbol{\beta}^r$ are given by unit vectors $\mathbf{e}^i$ and $\mathbf{e}^j$, respectively. This means that the reaction is a simple exchange reaction $C_i \rightleftharpoons C_j$. The linear RDS on a bounded smooth domain $\Omega \subset \mathbb{R}^d$ takes the form

$$\dot{\mathbf{c}} = \mathbb{D}\Delta\mathbf{c} + \mathbb{A}\mathbf{c}, \quad \text{where } \mathbb{D} = \mathrm{diag}(\delta_j)_{j=1,\ldots,I} \text{ with } \delta_j \geq 0 \tag{2.8}$$

complemented by no-flux boundary conditions. The matrix $\mathbb{A}$ is as before. Now, $c_j(t, \cdot) \in L^1(\Omega)$ is the nonnegative concentration of the chemical species C_i.

This system can be understood as the Kolmogorov forward equation on the state space $S = \Omega \times \{1, \ldots, I\}$, where the random variable $Y_t = (X_t, i(t))$ undergoes a Brownian motion in Ω with diffusion constant δ_j as long as $i(t) = j$. At discrete times the particle can change its type within $\{1, \ldots, I\}$, according to the jump process induced by the generator $\mathbb{Q}_{\text{finite}} = \mathbb{A}^\top$, and then continue a Brownian motion with the new diffusion constant. The full generator is

$$(\mathbb{Q}f)(x, i) = \delta_i \Delta f(x, i) + \sum_{k=1}^{I} \mathbb{A}_{ki} f(x, k), \quad \nabla f(x, i) \cdot \nu = 0 \text{ on } \partial\Omega.$$

We now assume that the linear reaction system satisfies the detailed-balance condition, i.e. we assume that there is an equilibrium state $\mathbf{w}$ with $w_i > 0$ and $\mathbb{A}_{ij} w_j = \mathbb{A}_{ji} w_i$ for all i and j. Then, the steady state $\pi \in \text{Prob}(S)$ is given by the product of the d-dimensional Lebesgue measure on Ω and $\mathbf{w}$, up to a suitable normalization factor:

$$\pi = \frac{1}{Z} \, \mathrm{d}x \otimes w \text{ where } Z = \sum_{i=1}^{I} w_i \, \text{vol}(\Omega).$$

By normalizing $\mathbf{w}$ suitably, we may assume $Z = 1$ subsequently.

Using the Neumann boundary conditions, it is easy to check that the generator $\mathbb{Q}$ satisfies the detailed-balance condition (2.6) with respect to π.

Hence, we can apply Theorem 2.3 which provides the large-deviation GS for (2.8). Note that $\rho \in \text{Prob}(S)$ is absolutely continuous with respect to π if and only if $\rho = (c_i \, \mathrm{d}x)_{i=1,\ldots,I}$ with $\mathbf{c} = (c_i)_{i=1,\ldots,I} \in L^1_{\geq 0}(\Omega)^I$. Moreover, $\frac{\mathrm{d}\rho}{\mathrm{d}\pi} = (c_i/w_i)_{i=1,\ldots,I}$ shows that the probability density $\frac{\mathrm{d}\rho}{\mathrm{d}\pi}$ equals the vector of relative concentrations c_i/w_i.

The driving functional $\mathcal{E}$ is the relative entropy up to a factor $1/2$, viz.

$$\mathcal{E}(\rho) = \frac{1}{2} \int_S \lambda_{\mathrm{B}}\left(\frac{\mathrm{d}\rho}{\mathrm{d}\pi}\right) \mathrm{d}\pi = \frac{1}{2} \int_\Omega \left(\sum_{i=1}^{I} \lambda_{\mathrm{B}}(c_i(x)/w_i) w_i \right) \mathrm{d}x. \tag{2.9}$$

For calculating the dissipation potential $\mathcal{R}$ we can take advantage of the linearity in $\mathbb{Q}$. In fact, $\mathbb{Q}$ can be split into I diffusion processes and the reaction part, namely $\mathbb{Q} = \sum_{i=1}^{I} \mathbb{Q}_{\text{diff}}^{(i)} + \mathbb{Q}_{\text{finite}}$ with $\mathbb{Q}_{\text{finite}}$ as above. The corresponding functionals $\mathcal{H}_j$ for diffusion processes $\mathbb{Q}_{\text{diff}}^{(j)}$ take the form

$$\mathcal{H}_j(\rho, \xi) = \delta_j \int_\Omega \left(|\nabla \xi(x, j)|^2 + \Delta \xi(x, j) \right) \mathrm{d}\rho(x, j).$$

The dual dissipation potential $\mathcal{R}^*$ is obtained by replacing $\xi(x, j)$ by $\xi(x, j) + \frac{1}{2} \log\left(\frac{c_j(x)}{w_j}\right)$ and $\mathrm{d}\rho(x, j) = c_j(x) \, \mathrm{d}x$, where we also use

$$\int_\Omega \nabla c_j(x) \cdot \nabla \xi(x, j) + c_j(x) \Delta \xi(x, j) \, \mathrm{d}x = \int_{\partial\Omega} c_j(x) \nabla \xi(x, j) \cdot \nu \, \mathrm{d}a = 0.$$

Subtracting the term at $\xi = 0$, writing $\boldsymbol{\xi} = (\xi_j)_j$ with $\xi_j(x) = \xi(x, j)$, and using the result for $\mathbb{Q}_{\text{finite}}$ from above, we arrive at the formula

$$\mathcal{R}^*(\mathbf{c}, \boldsymbol{\xi}) = \int_\Omega \sum_{j=1}^{I} \delta_j |\nabla \xi_j(x)|^2 c_j(x) \, dx$$

$$+ \int_\Omega \sum_{i=1}^{I-1} \sum_{k=2}^{I} \frac{1}{2} \sqrt{\mathbb{A}_{ki} c_i(x) \mathbb{A}_{ik} c_k(x)} \, \mathscr{C}^* \big(2(\xi_k(x) - \xi_i(x)) \big) \, dx.$$

This is the same dual dissipation potential as given in (2.5), except for the factors $\frac{1}{2}$ and 2 outside and inside of $\mathscr{C}^*$. However, these scaling factors arise since in the large-deviation result in Theorem 2.3 a factor $\frac{1}{2}$ appears in the definition of $\mathcal{E}(\rho)$ (see (2.9)), which is one-half of the usual relative entropy.

2.4.4 Large deviations for a membrane model

We consider a diffusion equation in the interval $\Omega = \,]{-1}, 1[$, where at $x = 0$ there is a membrane giving rise to a transmission condition. The Kolmogorov forward equation takes the form (where $\dot{} = \partial_t$ and $' = \partial_x$)

$$\dot{\rho} = a_\pm \rho'' \text{ for } \pm x \in \,]0, 1[, \qquad 0 = a_\pm \rho'(\pm 1),$$
$$a_+ \rho'(0^+) = b\big(\rho(0^+) - \rho(0^-)\big) = a_- \rho'(0^-).$$

The last relation means first that the mass flowing out of $]{-1}, 0[$ has to equal the flow into $]0, 1[$, and second that this flow is proportional to the difference of the densities.

The invariant measure is $\pi = \frac{1}{2} \, dx$, and the Markov generator $\mathbb{Q}_{\text{memb}}$ takes the form

$$(\mathbb{Q}_{\text{memb}} f)(x) = a_\pm f''(x) \text{ for } \pm x \in \,]0, 1[, \qquad 0 = a_\pm f'(\pm 1),$$
$$a_+ f'(0^+) = b\big(f(0^+) - f(0^-)\big) = a_- f'(0^-).$$

The functional $\mathcal{H}_{\text{memb}}$ takes the form

$$\mathcal{H}_{\text{memb}}(\rho, \zeta) = \int_{]{-1},0[} a_- \big(\zeta'' + (\zeta')^2\big) \rho \, dx + \int_{]0,1[} a_+ \big(\zeta'' + (\zeta')^2\big) \rho \, dx$$
$$\text{with } \zeta'(\pm 1) = 0 \text{ and } a_+ \zeta'(0^+) e^{\zeta(0^+)} = b\big(e^{\zeta(0^+)} - e^{\zeta(0^-)}\big) = a_- \zeta'(0^-) e^{\zeta(0^-)}.$$

Inserting $\zeta = \xi + \frac{1}{2} \log(2\rho)$ and doing an integration by parts using the nonlinear boundary conditions one obtains the dual dissipation potential

$$\mathcal{R}_{\text{memb}}(\rho, \xi) = \int_{]{-1},0[} a_- (\xi')^2 \rho \, dx + \sqrt{\rho(0^-)\rho(0^+)} \, \mathscr{C}^* \big(2(\xi(0^+) - \xi(0^-))\big) + \int_{]0,1[} a_+ (\xi')^2 \rho \, dx,$$

which again features the non-quadratic dissipation function $\mathscr{C}^*$.

3 Evolutionary Γ-convergence

Following the notions in the survey [Mie15a] we consider families of GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)_{\varepsilon \in \,]0,1[}$ and ask the question whether the solutions u_ε for these systems have a limit u for $\varepsilon \to 0$

and whether u is again a solution to a GS $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$. Ideally, one might hope that it is sufficient for $\mathcal{E}_\varepsilon$ and $\mathcal{R}_\varepsilon$ to converge in a suitable topology to $\mathcal{E}_0$ and $\mathcal{R}_0$, respectively. Such results indeed exist and can be found in the surveys [Ser11, Mie15a]. However, the aim of this work is to highlight the fact that starting with classical (i.e. quadratic) dissipation potentials $\mathcal{R}_\varepsilon$ we may end up with a limiting dissipation $\mathcal{R}_0$ that is non-quadratic. Thus, limits of classical GS may be generalized GS. First such examples were given in [Mie12, MiT12] in the context of plasticity.

3.1 pE-convergence of gradient systems

We first recall the general definition of pE-convergence, which is a short name for evolutionary Γ-convergence with well-prepared initial conditions. Hence, the letter "E" stands for both, 'E'volutionary convergence and 'E'nergy convergence, while the letter "p" stands for well'P'reparedness of the initial conditions, i.e., $\mathcal{E}_\varepsilon(0, u_\varepsilon(0)) \to \mathcal{E}_0(0, u^0) < \infty$.

Definition 3.1 (pE-convergence of $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$) *We say that the generalized gradient systems $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ pE-converge to $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$, and write $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathrm{pE}}{\to} (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$, if*

$$\left.\begin{array}{c} u_\varepsilon : [0, T] \to \mathbf{X} \\ \text{is sol. of } (\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon), \\ u_\varepsilon(0) \rightharpoonup u^0, \text{ and} \\ \mathcal{E}_\varepsilon(0, u_\varepsilon(0)) \to \mathcal{E}_0(0, u^0) < \infty \end{array}\right\} \implies \left\{\begin{array}{c} \exists\, u \text{ sol. of } (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0) \text{ with } u(0) = u^0 \\ \text{and a subsequence } \varepsilon_k \to 0 : \\ \forall\, t \in \,]0, T]: u_{\varepsilon_k}(t) \rightharpoonup u(t) \text{ and} \\ \mathcal{E}_{\varepsilon_k}(u_{\varepsilon_k}(t)) \to \mathcal{E}_0(u(t)). \end{array}\right. \tag{3.1}$$

Here $u_\varepsilon \rightharpoonup u$ means the weak convergence in the Banach space $\mathbf{X}$. We emphasize that the notion of pE-convergence asks for convergence of both, the solutions and the energies, but not of the dissipation potentials. However, using the EDP and the convergence of the energies, we easily obtain convergence of the integrated dissipations, namely

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

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

A first systematic study of evolutionary Γ-convergence relying on gradient structures was initiated in Sandier-Serfaty [SaS04], see also [Ser11, Mie15a]. In this approach one derives *sufficient conditions* for pE-convergence based on a limiting passage in the EDB

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon, \dot{u}_\varepsilon) + \mathcal{R}^*(u_\varepsilon, -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))\,\mathrm{d}t = \mathcal{E}_\varepsilon(u_\varepsilon(0)). \tag{3.2}$$

We observe that on the right-hand side we have the initial energy, which converges to the desired limit because of the well-prepared initial conditions. Thus, to obtain an (EDE) for the limiting process it suffices to show three liminf estimates for the terms on the left-hand side, namely

$$\liminf_{\varepsilon \to 0} \mathcal{E}_\varepsilon(u_\varepsilon(T)) \geq \mathcal{E}_0(u(T)); \tag{3.3a}$$

$$\liminf_{\varepsilon \to 0} \int_0^T \mathcal{R}_\varepsilon(u_\varepsilon, \dot{u}_\varepsilon)\,\mathrm{d}t \geq \int_0^T \mathcal{R}_0(u, \dot{u})\,\mathrm{d}t; \tag{3.3b}$$

$$\liminf_{\varepsilon \to 0} \int_0^T \mathcal{R}_\varepsilon^*\big(u_\varepsilon, -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon)\big)\,\mathrm{d}t \geq \int_0^T \mathcal{R}_0^*\big(u, -\mathrm{D}\mathcal{E}_0(u)\big)\,\mathrm{d}t. \tag{3.3c}$$

Of course, it is sufficient that these convergences hold only along (a subsequence of) the solutions u_ε of (EDP). In the following subsection, we will generalize this approach by keeping the terms $\mathcal{R}_\varepsilon$ and $\mathcal{R}_\varepsilon^*$ together.

3.2 EDP-convergence for gradient systems

Here we define a new notion of evolutionary Γ-convergence for GS that, on the one hand, is more restrictive but, one the other hand, gives a more precise information on the limiting dissipation potential $\mathcal{R}_0$. We use the fact that we do not need to have the two convergences (3.3b) and (3.3c) separately. Indeed, it is sufficient that only the integral over the sum of the two terms in (3.2) converges. This approach relaxes the sufficient conditions for pE-convergence substantially, since in the limit $\varepsilon \to 0$ the different parts of the dissipation may be distributed differently.

In the quadratic case we always have equidistribution $\mathcal{R}_\varepsilon(u_\varepsilon, \dot{u}_\varepsilon) = \mathcal{R}_\varepsilon^*(u_\varepsilon, -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))$ for solutions u_ε. So, if (3.3b) and (3.3c) hold, we will still have equidistribution in the limit, but only along the limit solutions, while the limit functionals need not be quadratic. In [Mie12] it is shown that the limit of a classical GS can be a rate-independent system, where we always have $\mathcal{R}_0^*(u, -\mathrm{D}\mathcal{E}_0(u)) = 0$, so there is not even equidistribution along solutions.

We also add a strengthening condition to obtain our notion of "EDP convergence" by asking for the convergence not only along solutions, but rather along a suitable general class of functions $u : [0, T] \to \mathbf{X}$. For this, we associate with the GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)_{\varepsilon \in [0,1]}$ De Giorgi's dissipation functionals $\mathscr{D}_\varepsilon$, which are defined as

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

Definition 3.2 (EDP convergence) *We say that the family* $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)_{\varepsilon > 0}$ *converges in the EDP sense to* $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$, *and shortly write* $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathrm{EDP}}{\rightharpoonup} (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ *if*

$$(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathrm{pE}}{\rightharpoonup} (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0), \tag{3.4a}$$

$$\mathcal{E}_\varepsilon \overset{\Gamma}{\rightharpoonup} \mathcal{E}_0 \ in \ \mathbf{X}, \ and \tag{3.4b}$$

$$\left.\begin{array}{l} \widetilde{u}_\varepsilon(\cdot) \overset{*}{\rightharpoonup} \widetilde{u}(\cdot) \ in \ \mathrm{L}^\infty([0,T]; \mathbf{X}) \ and \\ \sup_{\varepsilon \in]0,1], \, t \in [0,T]} \mathcal{E}_\varepsilon(\widetilde{u}_\varepsilon(t)) \le C < \infty \end{array}\right\} \implies \liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(\widetilde{u}_\varepsilon) \ge \mathscr{D}_0(\widetilde{u}). \tag{3.4c}$$

We emphasize that in condition (3.4c), the functions $\widetilde{u}_\varepsilon$ are arbitrary and need not be solutions of the GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$. From this definition we see that the convergence conditions (3.3) obviously imply EDP convergence. However, we will study cases, where (3.3) does not hold, but we still have EDP convergence.

We emphasize that EDP-convergence is to be expected whenever one uses the EDP principle for establishing pE-convergence. Indeed, from general arguments one presumes that De Giorgi's dissipation functional $\mathscr{D}_\varepsilon$ has (after extraction of a subsequence) a Γ-limit $\mathscr{D}_0$ in the form

$$\mathscr{D}_0(u) = \int_0^T \mathcal{M}_0(u(t), \dot{u}(t)) \,\mathrm{d}t.$$

From the lower semicontinuity of Γ-limits, one expects that $\mathcal{M}(u, \cdot)$ is convex. Hence, one can define $\mathcal{R}_\mathcal{M}$ via
$$\mathcal{R}_\mathcal{M}(u, v) := \mathcal{M}_0(u, v) - \mathcal{M}_0(u, 0)$$
and hope that it is a dissipation potential. For this, one needs to show (i) the positivity $\mathcal{R}_\mathcal{M}(u, v) \geq 0$ and (ii) $\mathcal{M}_0(u, 0) \geq \mathcal{R}_\mathcal{M}^*(u, -\mathrm{D}\mathcal{E}_0(u))$. Often, the positivity (i) follows simply from the evenness $\mathcal{M}_0(u, -v) = \mathcal{M}_0(u, v)$ and convexity. Moreover, if it is possible to show $\mathcal{M}_0(u, v) \geq -\langle \mathrm{D}\mathcal{E}_0(u), v \rangle$ (this holds for $\varepsilon > 0$ in the form $\mathcal{R}_\varepsilon(u, v) + \mathcal{R}_\varepsilon^*(u, -\mathrm{D}\mathcal{E}_\varepsilon(u)) \geq -\langle \mathrm{D}\mathcal{E}_\varepsilon(u), v \rangle$), then we find (ii) via the estimate

$$\mathcal{R}_\mathcal{M}^*(u, -\mathrm{D}\mathcal{E}_0(u)) = \sup_{v \in \mathbf{X}} \left(\langle -\mathrm{D}\mathcal{E}_0(u), v \rangle - \mathcal{M}_0(u, v) + \mathcal{M}_0(u, 0) \right) \leq \mathcal{M}_0(u, 0).$$

Thus, we arrive at the desired EDE $\mathcal{E}_0(u(T)) + \int_0^T \mathcal{R}_\mathcal{M} + \mathcal{R}_\mathcal{M}^* \, \mathrm{d}t \leq \mathcal{E}_0(u(0))$ and EDP-convergence to $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_\mathcal{M})$ is established.

It would be interesting to study more generally the relations between pE and EDP-convergence. Obviously, showing the liminf estimate for $\mathscr{D}_\varepsilon$ is the major step in establishing pE-convergence. Hence, it seems redundant to ask for the pE-convergence explicitly, yet it is not obvious under what additional condition (e.g. the validity of a suitable chain rule) we really can deduce the pE-convergence from the liminf estimate for $\mathscr{D}_\varepsilon$.

We end this section with two examples concerning EDP-convergence. Example 3.3 shows that the model discussed in [Mie12] satisfies pE-convergence but not EDP-convergence. Example 3.4 emphasizes the fact that pE and EDP-convergence are not properties of an evolution equation $\dot{u} = V_\varepsilon(u)$ but of a GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$. Indeed, for a given equation one may have different gradient structures leading to different limits in the EDP sense, which in turn generate different limit evolutions.

Example 3.3 (pE-convergence without EDP-convergence) *We consider the wiggly-energy model introduced in [ACJ96]. It is given via the time-dependent GS $(\mathbb{R}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ with*
$$\mathcal{E}_\varepsilon(t, u) = \frac{1}{2}u^2 - \ell(t)u + r\varepsilon \sin(u/\varepsilon) \quad \text{and} \quad \mathcal{R}_\varepsilon(\dot{u}) = \frac{\varepsilon}{2}\dot{u}^2.$$

For sufficiently smooth loading curves $\ell : [0, T] \to \mathbb{R}$ it was shown in [Mie12, Thm. 3.2] that the GS $(\mathbb{R}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ pE-converge to the generalized GS $(\mathbb{R}, \mathcal{E}_{\mathrm{play}}, \mathcal{R}_{\mathrm{play}})$ defined in (2.1). Obviously, we have the uniform convergence $\mathcal{E}_\varepsilon \to \mathcal{E}_{\mathrm{play}}$, while $\mathscr{D}_\varepsilon$ converges to a limit $\mathscr{D}_0$ that cannot be written in terms of $\mathcal{R}_{\mathrm{play}} + \mathcal{R}_{\mathrm{play}}^*$, see [Mie12, Prop. 3.1]

Example 3.4 (Different limit equations) *Here, we provide an example of an evolution equation $\dot{u} = V_\varepsilon(u)$ with two different gradient structures. Both gradient structures have an evolutionary Γ-limit in the EDP sense, and the surprising fact is that the generated limit evolutions are different. Thus, EDP-convergence and pE-convergence are not properties of the family of evolution equations $\dot{u} = V_\varepsilon(u)$, but of the chosen gradient structures.*

Consider $\Omega = [0, 1]$ and let $\mathbf{X} = \mathbf{M}_{\geq 0}(\Omega)$, the set of nonnegative finite Radon measures. Moreover, consider a continuous periodic function $\mathbb{A} : \mathbb{R} \to]0, \infty[$ such that
$$0 < a_{\min} := \min_{y \in \mathbb{R}} \mathbb{A}(y) < \max_{y \in \mathbb{R}} \mathbb{A}(y) =: a_{\max} < \infty.$$

With $a_\varepsilon(x) = \mathbb{A}(x/\varepsilon)$ we define the simple PDE
$$\dot{u}(t, x) = -a_\varepsilon(x)u(t, x), \quad u(0, x) = u_0(x) > 0 \quad \text{for } x \in \Omega.$$

We introduce two different gradient structures $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ and $(\mathbf{X}, \widetilde{\mathcal{E}}_\varepsilon, \widetilde{\mathcal{R}}_\varepsilon)$ via

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

It is shown in [Mie15a, Cor. 3.8] that these GS converge in the EDP sense to the limit systems $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ and $(\mathbf{X}, \widetilde{\mathcal{E}}_0, \widetilde{\mathcal{R}}_0)$, respectively, where

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

In particular, the limit evolution for the first is $\dot{u} = -a_{\min} u$, while it is $\dot{u} = -a_{\max} u$ for the second. This is not a contradiction, but has its origin in the well-preparedness condition for the initial data. No sequence can be well-prepared for both systems, i.e. if $u_\varepsilon \overset{*}{\rightharpoonup} u$ and $\mathcal{E}_\varepsilon(u_\varepsilon) \to \mathcal{E}_0(u)$, then we have $\widehat{\mathcal{E}}_0(u) \lneqq \liminf_{\varepsilon \to 0} \widehat{\mathcal{E}}_\varepsilon(u_\varepsilon)$, and vice versa.

3.3 EDP-convergence for an ODE example

We discuss a very simple example of a discrete Markov process with state space $S = \{1, 2, 3\}$. The jump rates are such that in the limit $\varepsilon \to 0$ the particles never stay in the state 2. Thus, the limiting Markov process has the state space $\{1, 3\}$ only, see Figure 3.1. We will start with three different GS, namely (i) the quadratic one, where both $\mathcal{E}_\varepsilon$ and $\mathcal{R}_\varepsilon$ are quadratic, (ii) the entropic one with classical $\mathcal{R}_\varepsilon^*$, and (iii) the entropic one with the dual dissipation potential defined in terms of $\mathscr{C}^*$. The interesting point is that in the cases (i) and (iii) the limiting GS obtained via EDP-convergence will still be in the same modeling class. However, in case (ii) we will lose the classical GS and obtain a generalized GS that cannot be described via $\mathscr{C}^*$.

We consider the Kolmogorov forward equation (here an ODE) of a Markov process on the state space $S = \{1, 2, 3\}$ given by

$$\dot{u} = (2+\varepsilon) \begin{pmatrix} -1 & 1/\varepsilon & 0 \\ 1 & -2/\varepsilon & 1 \\ 0 & 1/\varepsilon & -1 \end{pmatrix} u, \quad \mathbf{X} := \mathrm{Prob}(\{1, 2, 3\}). \tag{3.5}$$

The unique equilibrium $w^\varepsilon = \frac{1}{2+\varepsilon}(1, \varepsilon, 1)^\top$ satisfies the detailed-balance condition (2.6).

The limit dynamics is easily obtained by setting $u_2 = \varepsilon r$ giving

$$\dot{u}_1 = (2+\varepsilon)(r - u_1), \quad \varepsilon \dot{r} = (2+\varepsilon)(u_1 - 2r + u_3), \quad \dot{u}_3 = (2+\varepsilon)(r - u_3).$$

Thus, in the limit $\varepsilon \to 0$ we find $r = (u_1 + u_3)/2$ and

$$\dot{u}_1 = u_3 - u_1, \quad 0 = u_1 - 2r + u_3, \quad \dot{u}_3 = u_1 - u_3.$$

More precisely, if the initial condition satisfies $u^\varepsilon(0) \to (p_0, 0, 1-p_0)$, then for all $t > 0$ we have

$$u^\varepsilon(t) \to \big(p(t), 0, 1-p(t)\big)^\top \quad \text{where } \dot{p}(t) = 1 - 2p(t), \ p(0) = p_0.$$

We will study the limit $\varepsilon \to 0$ in several gradient structures. For general strictly con-
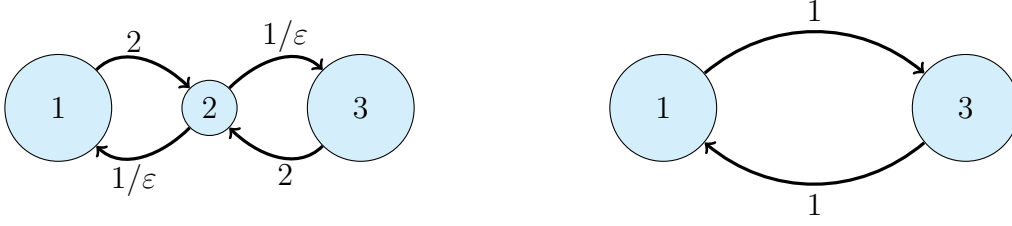
Figure 3.1: Left: Three-state Markov process with high rate of leaving state 2. Right: The limit for $\varepsilon \to 0$ gives a two-state Markov process.

vex and superlinear functions ϕ and ψ we consider the GS $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ with $\mathcal{E}_\varepsilon(u) = \sum_{i=1}^3 w_i^\varepsilon \phi(u_i/w_i^\varepsilon)$ (recall $w^\varepsilon = \frac{1}{2+\varepsilon}(1, \varepsilon, 1)^\top$) and

$$\mathcal{R}_\varepsilon^*(u, \xi) = \sum_{j=1}^2 a_j^\varepsilon(u)\, \psi^*\big(\xi_{j+1} - \xi_j\big), \quad \text{where } a_j^\varepsilon(u) := \frac{u_{j+1}/w_{j+1}^\varepsilon - u_j/w_j^\varepsilon}{(\psi^*)'\big(\phi'(\frac{u_{j+1}}{w_{j+1}^\varepsilon}) - \phi'(\frac{u_j}{w_j^\varepsilon})\big)}.$$

Using the fact that $v = \dot{u}$ satisfies $v_2 = -v_1 - v_3$ we obtain the primal dissipation potential $\mathcal{R}_\varepsilon$ in the form

$$\mathcal{R}_\varepsilon(u, v) = a_1^\varepsilon(u)\psi\big(\frac{v_1}{a_1^\varepsilon(u)}\big) + a_2^\varepsilon(u)\psi\big(\frac{v_3}{a_2^\varepsilon(u)}\big). \tag{3.6}$$

Indeed, in general $\mathcal{R}_\varepsilon$ is an inf-convolution, since $\mathcal{R}_\varepsilon^*$ is a sum over two terms. However, here we can eliminate v_2 and argue as follows:

$$\mathcal{R}_\varepsilon(u, v) = \sup\{\, \xi \cdot v - \mathcal{R}_\varepsilon^*(u, \xi) \mid \xi_1 + \xi_2 + \xi_3 = 0 \,\}$$
$$= \sup\{\, (\xi_1 - \xi_2)v_1 + (\xi_3 - \xi_2)v_3 - \mathcal{R}_\varepsilon^*(u, \xi) \mid \xi_1 + \xi_2 + \xi_3 = 0 \,\},$$

which gives the desired result, since $\mathcal{R}_\varepsilon^*$ only depends on $\xi_1 - \xi_2$ and $\xi_2 - \xi_3$.

We now state the result on EDP-convergence of $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ to $(\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$. Since the limiting GS can be described on the much smaller set $\mathrm{Prob}(\{1, 3\})$, which we identify with $\mathbf{Y} := [0, 1]$, we can formulate the limit GS in terms of the reduced GS $([0, 1], \mathbf{E}, \mathbf{R})$.

Theorem 3.5 *We have* $(\mathbf{X}, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon) \overset{\mathrm{EDP}}{\to} (\mathbf{X}, \mathcal{E}_0, \mathcal{R}_0)$ *with*

$$\mathcal{E}_0(u) = \begin{cases} \mathbf{E}(p) & \text{for } u = (p, 0, 1-p)^\top, \\ \infty & \text{otherwise}; \end{cases} \quad \text{with } \mathbf{E}(p) = \frac{1}{2}\phi(2p) + \frac{1}{2}\phi(2 - 2p)$$

$$\mathcal{R}_0(u, \dot{u}) = \begin{cases} \mathbf{R}(p, \dot{p}) & \text{for } (u, \dot{u}) = (p, 0, 1-p, \dot{p}, 0, -\dot{p})^\top, \\ \infty & \text{otherwise}, \end{cases}$$

where $\mathbf{R}$ *is given via its Legendre dual*

$$\mathbf{R}^*(p, \eta) := \sigma(p) + \sup_{z > 0}\left\{ \inf_{\tau \in \mathbb{R}}\Big(\widehat{a}(p, z)\psi^*(\eta - \tau) + \widehat{a}(1-p, z)\psi^*(\tau)\Big) - \Sigma(p, z) \right\},$$

$$\text{where} \quad \Sigma(p, z) := \widehat{a}(p, z)\psi^*\big(\phi'(2p) - \phi'(z)\big) + \widehat{a}(1-p, z)\psi^*\big(\phi'(2 - 2p) - \phi'(z)\big),$$

$$\widehat{a}(p, r) := (2p - r)/\{(\psi^*)'\big(\phi'(2p) - \phi'(r)\big)\}, \text{ and}$$

$$\sigma(p) := \inf\{\, \Sigma(p, z) \mid z > 0 \,\}.$$

19

In particular, this implies that the limiting ODE $\dot{p} = 1 - 2p$ is induced by the reduced generalized GS $([0,1], \mathbf{E}, \mathbf{R})$, i.e. $\dot{p} = 1 - 2p = D_\eta \mathbf{R}(p, -D\mathbf{E}(p))$. The above theorem follows directly from the next proposition and the general theory described in Section 3.2. For the energy-dissipation principle we consider De Giorgi's dissipation functional

$$\mathscr{D}_\varepsilon(u) = \int_0^T M_\varepsilon(u(t), \dot{u}(t)) \, dt \quad \text{with } M_\varepsilon(u,v) = \mathcal{R}_\varepsilon(u,v) + \mathcal{R}_\varepsilon^*(u, -D\mathcal{E}_\varepsilon(u)).$$

Proposition 3.6 *We have the Γ-limits* $\mathcal{E}_\varepsilon \overset{\Gamma}{\to} \mathcal{E}_0$, $M_\varepsilon \overset{\Gamma}{\to} M_0$, *and* $\mathscr{D}_\varepsilon \overset{\Gamma}{\to} \mathscr{D}_0$ *with*

$$M_0(u,v) = \begin{cases} \mathbf{R}(p,\nu) + \mathbf{R}^*\big(p, \phi'(2p) - \phi'(2-2p)\big) & \text{for } (u,v) = (p, 0, 1-p, \nu, 0, -\nu)^\top, \\ \infty & \text{otherwise}; \end{cases}$$

and $\mathscr{D}_0(u) = \int_0^T M_0(u(t), \dot{u}(t)) \, dt$, *where* $\mathbf{R}$ *is given as in Theorem 3.5.*

Proof: The convergence $\mathcal{E}_\varepsilon \overset{\Gamma}{\to} \mathcal{E}_0$ follows easily from the explicit form $\mathcal{E}_\varepsilon(u) = \sum_1^2 w_i^\varepsilon \phi(u_i/w_i^\varepsilon)$, the convergence $w_2^\varepsilon \to 0$, and the superlinearity of ϕ.

To simplify the Γ-limit of $\mathscr{D}_\varepsilon$ we introduce the scaling $u_2 = \varepsilon r$ and use that $\dot{u}_2 = \varepsilon \dot{r}$ does not explicitly appear in M_ε, see $\mathcal{R}_\varepsilon$ in (3.6). However, the relation $u_1 + u_2 + u_3 \equiv 1$ now takes the ε-dependent form $u_1 + \varepsilon r + u_3 \equiv 1$. Moreover, defining $\widetilde{M}_\varepsilon(u_1, r, u_3, \dot{u}_1, \dot{u}_3) = M_\varepsilon(u_1, \varepsilon r, u_3, \dot{u})$ shows that $\widetilde{M}_\varepsilon$ is continuous in $\varepsilon \in [0,1]$ and all the arguments.

Hence, when considering a sequence of functions $u^\varepsilon : [0,T] \to \mathbf{X}$ with $u^\varepsilon \overset{*}{\rightharpoonup} u_0$ and $\mathcal{E}_\varepsilon(u^\varepsilon(t)) \leq C < \infty$ we find $u_0(t) = (p(t), 0, 1-p(t))^\top$ and

$$\liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(u^\varepsilon) \geq \int_0^T m(p(t), \dot{p}(t)) \, dt \text{ with}$$

$$m(p,v) := \inf_{z>0} \left\{ \widehat{a}(p,z)\psi\Big(\frac{v}{\widehat{a}(p,z)}\Big) + \widehat{a}(1-p,z)\psi\Big(\frac{v}{\widehat{a}(1-p,z)}\Big) + \Sigma(p,z) \right\}.$$

Note that we have no control over the variable $r(t) = u_2^\varepsilon(t)/\varepsilon$ in the limit $\varepsilon \to 0$. So, we simply minimize over all possible values $z \in [0, \infty[$, which certainly provides a good lower bound. Moreover, recovery sequences for the convergence $\mathscr{D}_\varepsilon \overset{\Gamma}{\to} \mathscr{D}_0$ can be obtained in the form $u^\varepsilon(t) = (p(t), 0, 1-p(t)) + \varepsilon\zeta(t)(-p(t), 1, p(t)-1)$, where $\zeta(t)$ is the minimizer in the definition of $m(p(t), \dot{p}(t))$.

Obviously, the definition of σ gives the relation $\sigma(p) = m(p,0)$. Thus, we have derived the reduced dissipation potential $\mathbf{R}$ in the form $\mathbf{R}(p,v) = m(p,v) - \sigma(p)$. Doing a Legendre transform with respect to v we obtain the form of $\mathbf{R}^*$ given in Theorem 3.5, since the sum turns into an inf-convolution. ∎

Next we consider three different choices for ϕ and ψ.

3.3.1 Quadratic energy and dissipation

First, we consider the case

$$\phi(r) = \frac{1}{2}r^2 \quad \text{and} \quad \psi(v) = \frac{1}{2}v^2,$$

which gives $\mathbf{E}(p) = p^2 + (1-p)^2$ and $\widehat{a}(p,r) \equiv 1$ and simplifies all expressions considerably:

$$\Sigma(p,z) = \frac{1}{2}(2p-z)^2 + \frac{1}{2}(2-2p-z)^2, \quad \sigma(p) = (1-2p)^2, \quad \mathbf{R}(p,\eta) = \frac{1}{4}\eta^2.$$

Here, the crucial point in the definition of $\mathbf{R}^*$ is that the inf-convolution involving ψ^* and τ does not involve any dependence on z, so that the term $\sigma(p)$ exactly cancels $\sup_{z>0} -\Sigma(p, z)$, which is generally not the case. Thus, the limiting GS is $([0, 1], \mathbf{E}, \mathbf{R})$ where $\mathbf{R}(p, \nu) = \nu^2$ is quadratic, and $([0, 1], \mathbf{E}, \mathbf{R})$ is again a classical GS.

3.3.2 Entropic energy and $\mathscr{C}$-type dissipation

Next, we consider the case of the Boltzmann entropy and the dissipation defined in terms of $\psi = \mathscr{C}$, which coincides with Section 2.4 except for the trivial scaling factor 2:

$$\phi(r) = \lambda_{\mathrm{B}}(r) = r \log r - r + 1 \quad \text{and} \quad \psi^*(\xi) = \mathscr{C}^*(\xi) = 4\big(\cosh(\xi/2) - 1\big).$$

This gives the reduced energy functional $\mathbf{E}(p) = \frac{1}{2}\lambda_{\mathrm{B}}(2p) + \frac{1}{2}\lambda_{\mathrm{B}}(2-2p)$ and $\widehat{a}(p, z) = \sqrt{2pz}$. In the latter expression and in the definition of Σ we profit from the interaction of $\lambda_{\mathrm{B}}'(r) = \log r$ and the exponential form of $\mathscr{C}^*$, viz.

$$\Sigma(p, z) = 2(\sqrt{2p} - \sqrt{z})^2 + 2(\sqrt{2-2p} - \sqrt{z})^2 = 4 - 4b_p\sqrt{z} + 4z \text{ with } b_p = \sqrt{2p} + \sqrt{2-2p}.$$

Minimizing in $z > 0$ we arrive at

$$\sigma(p) = 4 - b_p^2 = 2 - 4\sqrt{p(1-p)} = 2(\sqrt{p} - \sqrt{1-p})^2.$$

For calculating $\mathbf{R}^*$ we first observe, for $a, b > 0$, the formula

$$\inf_{\tau \in \mathbb{R}} \left(a\mathscr{C}^*(\tau) + b\mathscr{C}^*(\xi - \tau)\right) = 4\sqrt{(a+b)^2 + \frac{ab}{2}\mathscr{C}^*(\xi)} - 4(a+b),$$

which follows by writing the left-hand side via $\mathscr{C}^*(\xi - \tau) = 2e^\xi/x + 2e^{-\xi}x - 4$, where $x = e^\tau$, and minimizing in $x > 0$. With $a = \widehat{a}(p, z)$ and $b = \widehat{a}(1-p, z)$ we find

$$\mathbf{R}^*(p, \eta) = \sigma(p) + \sup_{z>0} \left(4\sqrt{z}\sqrt{b_p^2 + \sqrt{p(1-p)}\mathscr{C}^*(\eta)} - 4b_p\sqrt{z} - \Sigma(p, z)\right)$$

$$= \sigma(p) + \sup_{z>0} \left(4\sqrt{z}\sqrt{b_p^2 + \sqrt{p(1-p)}\mathscr{C}^*(\eta)} - 4 - 4z\right)$$

$$= \sigma(p) - 4 + b_p^2 + \sqrt{p(1-p)}\,\mathscr{C}^*(\eta) = \sqrt{p(1-p)}\,\mathscr{C}^*(\eta).$$

We emphasize that in this minimization with respect to z it is crucial to keep the terms involving the dual dissipation potential $\mathscr{C}^*(\eta)$ and the term Σ together.

We observe that the resulting gradient structure is again the structure, which is obtained from the large-deviation principle of Section 2.4. This confirms the statement that gradient structures obtained from the large-deviation theory are very stable against taking further limits in the sense of EDP-convergence, see Figure 1.1

3.3.3 Entropic energy and quadratic dissipation

In [Maa11, ErM12, Mie13a, MaM15a, MaM15b] the relative entropy was used for the energy functional $\mathcal{E}$ and a quadratic dissipation leading to a classical gradient system:

$$\phi(r) = \lambda_{\mathrm{B}}(r) = r \log r - r + 1 \quad \text{and} \quad \psi^*(\xi) = \frac{1}{2}\xi^2.$$

We obtain the same limit energy $\mathbf{E}(p) = \frac{1}{2}\lambda_{\mathrm{B}}(2p) + \frac{1}{2}\lambda_{\mathrm{B}}(2-2p)$ as in the previous case, but the functions $a_j^\varepsilon(u)$ are quite different as they involve the logarithmic mean $\Lambda(r,s) = \frac{r-s}{\log r - \log s}$. Indeed we have $\widehat{a}(p,z) = \Lambda(2p, z) = \frac{2p-z}{\log(2p) - \log z}$. We further obtain the functions

$$\Sigma(p,z) = \frac{1}{2}\Big((2p-z)\big(\log(2p) - \log z\big) + (2-2p-z)\big(\log(2-2p) - \log z\big)\Big)$$

and have no explicit formula for $\sigma(p) = \inf_{z>0} \Sigma(p,z)$. In the definition of $\mathbf{R}^*$ we can do the inf-convolution explicitly, since ψ^* is quadratic, so we find the formula

$$\mathbf{R}^*(p,\eta) = \sigma(p) + \sup_{z>0}\Big(\frac{\widehat{a}(p,z)\widehat{a}(1-p,z)}{2(\widehat{a}(p,z)+\widehat{a}(1-p,z))}\, \eta^2 \; - \Sigma(p,z)\Big).$$

We claim that the growth of $\mathbf{R}^*(p,\eta)$ is no longer quadratic, but exponential. For this we insert $z = \mathrm{e}^{b\eta}$ for $\eta \gg 1$ for some $b \in \,]0, 1/2[$ into the supremum to obtain a lower bound. From $\Sigma(p,z) \approx z\log z$ and $\widehat{a}(p,z) \approx z/\log z$ for $z \to \infty$ we find the asymptotic lower bound

$$\mathbf{R}^*(p,\eta) \gtrapprox \Big(\frac{1}{4b} - b\Big)\eta\mathrm{e}^{b\eta}.$$

Hence, we see that the growth is at least as $\mathrm{e}^{b|\eta|}$ for all $b \in \,]0, 1/2[$. Moreover, we expect that the function $\mathbf{R}^*(p,\eta)$ does not have a product structure $b(p)\Psi(\eta)$ any more.

Thus, we see that the classical gradient structure for the relative entropy is not stable under EDP, in general. Nevertheless, in [GiM13, DiL14, MaM15a] evolutionary Γ-limits between discrete Markov processes and continuous Fokker-Planck equation are studied, where the classical gradient structure survives.

4 The membrane as a thin-layer limit

In our first major application of the EDP-convergence as a microscopic origin of generalized GS, we follow [Lie12, Lie13] and consider a one-dimensional diffusion equation with a thin layer of very small diffusivity. Assuming that the diffusion coefficient and the width of the layer scale in the proper way, we will arrive at a membrane model in the limit. While the limit passage of the linear diffusion problem to the linear transmission problem at the membrane can be done directly (or with the quadratic gradient structure, see [Lie13]), we prefer to do the somewhat more elaborate EDP-limit using the GS with the relative entropy as energy functional and the classical dissipation potential of Wasserstein type. This case was already studied in [Lie12, Sec. 3.2] in a more special setting and without explicitly calculating $\mathcal{R}_0^*$.

We start from the equation

$$\dot{u} = \Big(a_\varepsilon(x)\big(u' + uV_\varepsilon'(x)\big)'\Big) \text{ in } \Omega := \,]-1, 1[, \quad \partial_x u(t, \pm 1) + u(t, \pm 1)V_\varepsilon'(\pm 1) = 0, \quad (4.1)$$

where $\dot{} = \partial_t$ and $' = \partial_x$. By our choice of the boundary conditions, the total mass $\int_\Omega u(t,x)\,\mathrm{d}x = 1$ is conserved, thus we can interpret the equation as the Fokker-Planck equation of a Markov process. Defining the equilibrium density

$$w_\varepsilon(x) = \frac{1}{Z_\varepsilon}\mathrm{e}^{-V_\varepsilon(x)} \quad \text{with } Z_\varepsilon = \int_{-1}^{1} \mathrm{e}^{-V_\varepsilon(x)}\,\mathrm{d}x,$$

we have the GS $(\mathrm{Prob}(\Omega), \mathcal{E}, \mathcal{R}_\varepsilon^*)$ with

$$\mathcal{E}_\varepsilon(u) = \int_\Omega \lambda_\mathrm{B}\big(u(x)/w_\varepsilon(x)\big) w_\varepsilon(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{R}_\varepsilon^*(u, \xi) = \frac{1}{2}\int_\Omega a_\varepsilon(x) u(x) \xi'(x)^2\,\mathrm{d}x.$$

The nontrivial behavior happens in the thin layer given by the small interval $[0, \varepsilon]$. In particular, we allow a_ε and V_ε to depend non-trivially on x: We assume that there are functions $a_*, a_+, V_*, V_+ \in \mathrm{C}^1([0, 1])$ and $a_-, V_- \in \mathrm{C}^1([-1, 0])$ such that $a_*(x), a_+(x), a_-(-x) \geq \underline{a} > 0$ for all $x \in [0, 1]$, $V_-(0) = V_*(0)$, $V_*(1) = V_+(0)$, and

$$a_\varepsilon(x) = \left\{ \begin{array}{ll} a_+(x) & \text{for } x > \varepsilon, \\ \varepsilon a_*(x/\varepsilon) & \text{for } x \in [0, \varepsilon], \\ a_-(x) & \text{for } x < 0, \end{array} \right. \quad \text{and} \quad V_\varepsilon(x) = \left\{ \begin{array}{ll} V_+(x+\varepsilon) & \text{for } x > \varepsilon, \\ V_*(x/\varepsilon) & \text{for } x \in [0, \varepsilon], \\ V_-(x) & \text{for } x < 0. \end{array} \right. \quad (4.2)$$

Here V_ε is constructed to be continuous on $\overline{\Omega} = [-1, 1]$, while a_ε has jumps for $x \in \{0, \varepsilon\}$.

The pE-convergence result established in [Lie12, Sec. 3] states that the limiting system is given as a membrane problem, where the thin layer is replaced by a transmission condition. The interesting point is that the EDP-convergence reveals that the limiting GS is no longer classical but involves $\mathscr{C}$ for the jump of the driving forces at the membrane.

For passing to the limit we note that the function w_ε converges pointwise to the limit

$$w_0(x) = \left\{ \begin{array}{ll} \frac{1}{Z_0}\mathrm{e}^{-V_+(x)} & \text{for } x > 0, \\ \frac{1}{Z_0}\mathrm{e}^{-V_-(x)} & \text{for } x < 0, \end{array} \right. \quad \text{with } Z_0 = \int_{-1}^0 \mathrm{e}^{-V_-(x)}\,\mathrm{d}x + \int_0^1 \mathrm{e}^{-V_+(x)}\,\mathrm{d}x, \quad (4.3)$$

which may be discontinuous at $x = 0$, but has well-defined limits $w_0(0^-)$ and $w_0(0^+)$ from the left and from the right, respectively. This limit is totally independent of the potential V_* inside the layer. The influence on the layer potential V_* and the layer diffusion profile a_* will only survive in one coefficient A_*.

Theorem 4.1 (Membrane limit) $(\mathrm{Prob}(\Omega), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon^*) \overset{\mathrm{EDP}}{\longrightarrow} (\mathrm{Prob}(\Omega), \mathcal{E}_0, \mathcal{R}_0^*)$, *where*

$$\mathcal{E}_0(u) = \int_\Omega \lambda_\mathrm{B}(u/w_0) w_0\,\mathrm{d}x \quad \text{and}$$

$$\mathcal{R}_0^*(u, \xi) = \int_{]-1,0[} \frac{a_-}{2}(\xi')^2 u\,\mathrm{d}x + \int_{]0,1[} \frac{a_+}{2}(\xi')^2 u\,\mathrm{d}x$$

$$+ A_*\sqrt{\frac{u(0^-)u(0^+)}{w_0(0^-)w_0(0^+)}}\,\mathscr{C}^*\big(\xi(0^+)-\xi(0^-)\big) \quad \text{where } A_* = \left(\int_0^1 \frac{Z_0 \mathrm{e}^{V_*(y)}}{a_*(y)}\,\mathrm{d}y\right)^{-1}.$$

Before we go into the details of the proof, some comments are in order. First, we emphasize that the constant Z_0 in the definition of the coupling coefficient A_* is not related to V_*, but only depends on $V_\pm$, see (4.3). Hence, for a large barrier V_* the transmission coefficient A_* becomes indeed small.

Second, the limiting equation is a PDE in the subdomains $\Omega_- =]-1, 0[$ and $\Omega_+ =]0, 1[$ coupled by a transmission condition. It can be obtained easily by considering test functions $\widehat{\xi} \in \mathrm{H}^1(\Omega_-) \times \mathrm{H}^1(\Omega_+)$ in the weak form $\int_\Omega \dot{u}\widehat{\xi}\,\mathrm{d}x = \mathrm{D}_\xi \mathcal{R}_0^*\big(u, -\mathrm{D}\mathcal{E}_0(u)\big)[\widehat{\xi}]$. Using the fact that $\widehat{\xi}$ may have a jump at $x = 0$, the transmission conditions arise via the

boundary terms when integrating by parts. We arrive at

$$\dot{u} = \Big(aw_0\big(u/w_0)\big)'\Big)' \quad \text{for } x \in \Omega_- \cup \Omega_+.$$

$$0 = a_+(0)w_0(0^+)\big(u/w_0)\big)'(0^+) - A_*\Big(\tfrac{u(0^+)}{w_0(0^+)} - \tfrac{u(0^-)}{w_0(0^-)}\Big),$$

$$0 = a_-(0)w_0(0^-)\big(u/w_0)\big)'(0^-) - A_*\Big(\tfrac{u(0^+)}{w_0(0^+)} - \tfrac{u(0^-)}{w_0(0^-)}\Big),$$

$$0 = a_\pm(x)w_0(x)\big(u/w_0)\big)'(x) \quad \text{at } x \in \{-1,1\}.$$

We refer to [GlM13] for a similar derivation of more general nonlinear transmission conditions and active interface conditions using gradient structures.

Finally, we remark that the primal dissipation potential can be written using the integration operator $I[\dot{u}](x) := \int_{-1}^{x} \dot{u}(y)\,\mathrm{d}y = -\int_{x}^{1} \dot{u}(y)\,\mathrm{d}y$, where the last relation follows from $\int_{-1}^{1} \dot{u}\,\mathrm{d}y = 0$, which in turn is due to $u(t) \in \mathrm{Prob}(\Omega)$. Noting that the functions ξ may have a jump at $x = 0$, one has the identity

$$\int_{-1}^{1} \xi\dot{u}\,\mathrm{d}x = -\int_{-1}^{0} I[\dot{u}]\xi'\,\mathrm{d}x - I[\dot{u}](0)\big(\xi(0^+)-\xi(0^-)\big) - \int_{0}^{1} I[\dot{u}]\xi'\,\mathrm{d}x.$$

Here $I[\dot{u}](0)$ is the flux through the membrane, which is thermodynamically conjugate to the jump $\xi(0^+)-\xi(0^-)$ in the driving forces. With this and $I[\dot{u}](-1) = 0 = I[\dot{u}](1)$ the evaluation of the Legendre transform for $\mathcal{R}_0^*$ yields the primal dissipation potential

$$\mathcal{R}_0(u,\dot{u}) = \int_{-1}^{0} \frac{I[\dot{u}]^2}{2a_-u}\,\mathrm{d}x + \int_{0}^{1} \frac{I[\dot{u}]^2}{2a_+u}\,\mathrm{d}x + A_*\sqrt{\tfrac{u(0^-)u(0^+)}{w_0(0^-)w_0(0^+)}}\,\mathscr{C}\Big(\frac{I[\dot{u}](0)}{A_*\sqrt{\tfrac{u(0^-)u(0^+)}{w_0(0^-)w_0(0^+)}}}\Big). \quad (4.4)$$

Proof of Theorem 4.1: We first observe that $\mathcal{E}_\varepsilon \xrightarrow{\Gamma_*} \mathcal{E}$ using [AGS05, Lem. 9.4.2]. Moreover, pE-convergence was established in [Lie12, Sec. 3.2]. It remains to establish the liminf estimate (3.4c), where De Giorgi's dissipation functional $\mathscr{D}_\varepsilon$ takes the explicit form

$$\mathscr{D}_\varepsilon(u) := \int_{0}^{T} \int_{\Omega} \frac{1}{2a_\varepsilon u} I[\dot{u}]^2 + \frac{a_\varepsilon u}{2}\Big(\big(\log(\tfrac{u}{w_\varepsilon})\big)'\Big)^2 \mathrm{d}x\,\mathrm{d}t.$$

Step 1. Blow up = transformation from $\mathscr{D}_\varepsilon$ to $\widehat{\mathscr{D}}_\varepsilon$: To study the Γ-limit $\mathscr{D}_0$ of $\mathscr{D}_\varepsilon$ we blow up the thin layer such that its transformed thickness becomes of order one. For this we use $Y_\varepsilon : [-1,1] \to [-1,2]$ and its inverse $X_\varepsilon = Y_\varepsilon^{-1}$:

$$Y_\varepsilon(x) = \begin{cases} x & \text{for } x \leq 0, \\ \frac{1+\varepsilon}{\varepsilon}x & \text{for } x \in [0,\varepsilon], \\ x+1 & \text{for } x \geq \varepsilon; \end{cases} \quad \text{and} \quad X_\varepsilon(y) = \begin{cases} y & \text{for } y \leq 0, \\ \frac{\varepsilon}{1+\varepsilon}y & \text{for } y \in [0,1+\varepsilon], \\ y-1 & \text{for } y \geq 1+\varepsilon. \end{cases}$$

For $u : [0,T] \times \Omega \to \mathbb{R}$ and $y \in \widehat{\Omega} := \,]-1,2[$ we define the functions

$$U_\varepsilon(t,y) = u(t,X_\varepsilon(y)), \quad W_\varepsilon(y) = w_\varepsilon(X_\varepsilon(y)), \quad A_\varepsilon(y) = \frac{a_\varepsilon(X_\varepsilon(y))}{X_\varepsilon'(y)},$$

and the functionals $\widehat{\mathscr{D}}_\varepsilon$ via $\mathscr{D}_\varepsilon(u) = \widehat{\mathscr{D}}_\varepsilon(U_\varepsilon)$ and find

$$\widehat{\mathscr{D}}_\varepsilon(U) := \int_{0}^{T} \int_{\widehat{\Omega}} \frac{1}{2A_\varepsilon U} \widehat{I}_\varepsilon[\dot{U}]^2 + \frac{A_\varepsilon U}{2}\Big(\big(\log(U/W_\varepsilon)\big)'\Big)^2 \mathrm{d}y\,\mathrm{d}t,$$

where $\widehat{I}_\varepsilon[\dot{U}](Y) = \int_{-1}^y \dot{U}(\eta) X'_\varepsilon(\eta)\, \mathrm{d}\eta = -\int_y^2 \dot{U}(\eta) X'_\varepsilon(\eta)\, \mathrm{d}\eta$.

Following the arguments in [Lie12, Sec. 3.2] it is not difficult to establish the Γ-convergence of $\widehat{\mathscr{D}}_\varepsilon$ to $\widehat{\mathscr{D}}_0$, where the latter is given in the form

$$\widehat{\mathscr{D}}_0(U) := \int_0^T \int_{\widehat{\Omega}} \frac{1}{2\widehat{A}U} \widehat{I}_0[\dot{U}]^2 + \frac{\widehat{A}U}{2}\Big(\big(\log(U/\widehat{W})\big)'\Big)^2 \mathrm{d}y\, \mathrm{d}t$$

$$\text{with} \quad \widehat{I}_0[\dot{U}](y) := \int_{-1}^y \dot{U}(\eta) \widehat{M}(\eta)\, \mathrm{d}\eta = -\int_y^2 \dot{U}(\eta) \widehat{M}(\eta)\, \mathrm{d}\eta,$$

$$\text{where} \ (\widehat{A}(y), \widehat{W}(y), \widehat{M}(y)) := \begin{cases} (a_-(y), \mathrm{e}^{-V_-(y)}/Z_0, 1) & \text{for } y < 0, \\ (a_*(y), \mathrm{e}^{-V_*(y)}/Z_0, 0) & \text{for } y \in [0,1], \\ (a_+(y), \mathrm{e}^{-V_+(y)}/Z_0, 1) & \text{for } y > 1. \end{cases}$$

Step 2. Minimization over the rescaled layer: The main structure in this limit model is that $\widehat{\mathscr{D}}_0$ does not depend on $\dot{U}(t,\cdot)|_{]0,1[}$, since $\widehat{M}(y) = 0$ for $y \in [0,1]$. Moreover, on this interval $\widehat{I}_0[\dot{U}(t,\cdot)]$ is constant, namely $\mu_u(t) := \int_{-1}^0 \dot{U}(\eta)\, \mathrm{d}\eta = \widehat{I}_0[\dot{U}(t)](y)$ for $y \in [0,1]$. Thus, given the value $\mu_u(t)$ we can obtain the optimal profile of $U(t)|_{[0,1]}$ from the boundary values $U(t,0)$ and $U(t,1)$ and minimizing the functional $\widehat{\mathscr{G}}(\mu_u(t); \cdot)$ given via

$$\widehat{\mathscr{G}}(\alpha, U) := \int_0^1 \frac{\alpha^2}{2\widehat{A}U} + \frac{\widehat{A}U}{2}\Big(\big(\log(U/\widehat{W})\big)'\Big)^2 \mathrm{d}y. \tag{4.5}$$

Now, Proposition A.2 in Appendix A.1 provides the explicit formula

$$\widehat{G}(\alpha, u_0, u_1) := \min\{\, \widehat{\mathscr{G}}(\alpha, U) \mid U > 0,\ U(0) = u_0,\ U(1) = u_1 \,\}$$

$$= A_* \sqrt{\tfrac{u_0 u_1}{w_- w_+}}\, \mathscr{C}\Big(\tfrac{1}{A_*}\sqrt{\tfrac{w_- w_+}{u_0 u_1}}\, \alpha\Big) + A_* \sqrt{\tfrac{u_0 u_1}{w_- w_+}}\, \mathscr{C}^*\Big(\log\big(\tfrac{u_0 w_+}{u_1 w_-}\big)\Big)$$

with $w_- = \widehat{W}(0) = w_0(0^-)$, $w_+ = \widehat{W}(1) = w_0(0^+)$, and $A_* = \big(\int_0^1 1/(\widehat{A}(y)\widehat{W}(y))\, \mathrm{d}y\big)^{-1}$. Inserting the definitions of $\widehat{A}$ and $\widehat{W}$ gives exactly the formula for A_* in the theorem.

Thus, we have constructed a simpler functional $\overline{\mathscr{D}}_0$, which is given by

$$\overline{\mathscr{D}}_0(U) := \int_0^T \Bigg[\int_{]-1,0[\cup]1,2[} \Big(\frac{1}{2\widehat{A}U} \widehat{I}_0[\dot{U}]^2 + \frac{\widehat{A}U}{2}\Big(\big(\log(U/\widehat{W})\big)'\Big)^2\Big) \mathrm{d}y$$

$$+ \widehat{G}\Big(\widehat{I}_0[\dot{U}](0), U(t,0), U(t,1)\Big) \Bigg] \mathrm{d}t,$$

satisfies the lower bound $\widehat{\mathscr{D}}_0(U) \geq \overline{\mathscr{D}}_0(U)$ for all U, and has the important property that it does not depend on $U|_{[0,T]\times]0,1[}$.

Step 3. Relation between $\overline{\mathscr{D}}_0$ and $\mathscr{D}_0$: Using the special form of $\mathcal{R}_0$ and $\mathcal{R}_0^*$ stated in (4.4) and the theorem, we define the limiting dissipation functional $\mathscr{D}_0(u) = \int_0^T \mathcal{R}_0(u, \dot{u}) + \mathcal{R}_0^*(u, -\mathrm{D}\mathcal{E}_0(u))\, \mathrm{d}t$. By construction and the special form of $\widehat{G}$ the functional $\overline{\mathscr{D}}_0$ is closely related to $\mathscr{D}_0$ in the following way. For any function $U : [0,T]\times[-1,2] \to \mathbb{R}$, we may define $u : [0,T]\times[-1,1]$ by $u(t,x) = U(t, Y_0(x))$, where $Y_0(x) = x$ for $x \leq 0$ and $Y_0(x) = x+1$ otherwise. Then, we have $\mathscr{D}_0(u) = \overline{\mathscr{D}}_0(U) \leq \widehat{\mathscr{D}}_0(U)$.

Moreover, for any function u one can construct an optimal U as follows. We split u at $x = 0$, the right part is shifted by 1 to the right, and the minimizer of $U(t,\cdot) \in \mathrm{H}^1([0,1])$ of $\widehat{\mathscr{G}}(I[\dot{u}(t)](0), u(t,0^-), u(t,0^+))$ is inserted into the gap. Then, $\mathscr{D}_0(u) = \widehat{\mathscr{D}}_0(U)$.

Step 4. The liminf estimate (3.4c): To establish the fundamental liminf estimate we consider, w.l.o.g., sequences u_ε satisfying $1/R \leq u_\varepsilon \leq R$ for some large $R > 1$. In particular, by minimum and maximum principles these bounds can be expected for solutions of (4.1). Defining $U_\varepsilon(t, y) = u_\varepsilon(t, X_\varepsilon(y))$ we again have $U_\varepsilon(t, y) \in [1/R, R]$. Thus, we find a subsequence $\varepsilon_k \to 0$ such that

$$u_\varepsilon \rightharpoonup u_0 \text{ in } L^2([0, T] \times \Omega) \quad \text{and} \quad U_\varepsilon \rightharpoonup U_0 \text{ in } L^2(0, T; L^2(\widehat{\Omega})).$$

Moreover, we have $u_0(t, x) = U_0(t, Y_0(x))$. Now, using $\mathscr{D}_\varepsilon(u_\varepsilon) = \widehat{\mathscr{D}}_\varepsilon(U_\varepsilon)$ we arrive at the desired liminf estimate

$$\liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(u_\varepsilon) = \liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(U_\varepsilon) \geq \widehat{\mathscr{D}}_0(U_0) \geq \overline{\mathscr{D}}_0(U_0) = \mathscr{D}_0(u_0). \tag{4.6}$$

This concludes the proof of Theorem 4.1. ∎

We conclude this section by observing that the EDP-limit of the thin-layer diffusion system given by the classical GS $(\text{Prob}(\Omega), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon^*)$ is a the generalized GS for the membrane problem. For $\varepsilon > 0$ and for $\varepsilon = 0$ the gradient structures are exactly the ones obtained from the large-deviation principle, see Section 2.4.4. Hence, we again found an instance where the diagram in Figure 1.1 commutes, that means that applying the large-deviation principle can be interchanged with taking the EDP-limit $\varepsilon \to 0$.

5 From diffusion to reaction

In our second major application of EDP-convergence as a microscopic origin of generalized GS, we continue the work in [PSV10, PSV12, AM*12] which show that linear reactions can be obtained as limits of diffusion for a suitably scaled energy barrier. In [PSV10, PSV12] the method relies on a quadratic energy functional and a classical gradient structure. In [AM*12] the pE-convergence for the entropic GS is shown, but only diffusion along the reaction path is allowed. In fact, the result therein gives EDP-convergence, if one takes the addition in [MPR14, Prop. 4,4] into account.

Here we generalize the latter work by also allowing diffusion in a physical space Ω, such that the resulting limit equation will be a (linear) reaction-diffusion system. Our physical domain $\Omega \subset \mathbb{R}^d$ is bounded and has a Lipschitz boundary. For the reaction path we choose $\Upsilon = [0, 7] \subset \mathbb{R}$ and define the cylinder $Q = \Omega \times \Upsilon$. (Indeed, Υ could by any bounded or unbounded interval.)

For densities $u \in L^1(Q)$ the integral $\int_D \int_{y_0}^{y_1} u \, \mathrm{d}y \, \mathrm{d}x$ denotes the number of particles per unit volume that are in the subdomain $D \subset \Omega$ and have a reaction state $y \in [y_0, y_1] \subset \Upsilon$. The evolution of the density u is driven by diffusion in the x-direction with diffusion constant $m_\Omega > 0$ and a much faster diffusion in the y-direction with diffusion constant $\tau_\varepsilon \gg 1$ to allow the particles to overcome a huge potential barrier given by $V_\varepsilon(y) = \frac{1}{\varepsilon} V(y)$, see Figure 5.1.

For simplicity we assume that the total mass $\int_Q u \, \mathrm{d}x \, \mathrm{d}y$ as well as the volume $|\Omega|$ of the physical domain equal 1. Hence, we can again consider the model as a Markov process with continuous paths $t \mapsto (X_t, Y_t) \in \Omega \times \Upsilon = Q$, whose distribution laws can be described by densities $u(t) \in \text{Prob}(Q)$. The Kolmogorov forward equation reads

$$\dot{u} = m_\Omega \Delta_x u + \tau_\varepsilon \partial_y \big(\partial_y u + u \, \partial_y V_\varepsilon \big), \quad (\nabla_x u, \partial_y u + u \, \partial_y V_\varepsilon) \cdot \nu = 0 \text{ on } \partial Q. \tag{5.1}$$
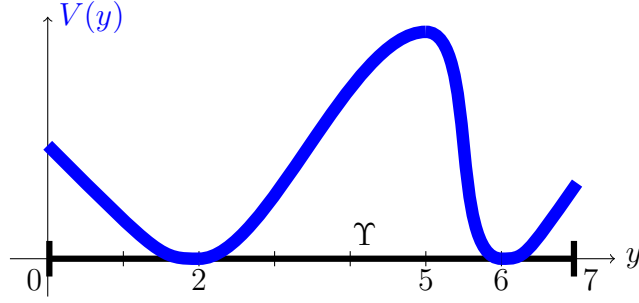
Figure 5.1: The potential V along the reaction path $\Upsilon = [0, 7]$.

Clearly, the unique steady state $\widetilde{w}_\varepsilon$ is independent of x and takes the form

$$\widetilde{w}_\varepsilon(x, y) = w_\varepsilon(y) := \frac{1}{z_\varepsilon} \exp\left(-\frac{1}{\varepsilon} V(y)\right) \text{ with } z_\varepsilon := \int_\Upsilon \exp\left(-\frac{1}{\varepsilon} V(y)\right) dy.$$

Equation (5.1) is a Fokker-Planck equation and, hence, has the Wasserstein gradient structure introduced in [JKO98] with

$$\mathcal{E}_\varepsilon(u) = \iint_Q \lambda_{\mathrm{B}}\left(\frac{u}{w_\varepsilon}\right) w_\varepsilon \, dy \, dx \text{ and } \mathcal{R}_\varepsilon^*(u, \xi) = \iint_Q \left(\frac{m_\Omega}{2} |\nabla_x \xi|^2 + \frac{\tau_\varepsilon}{2} (\partial_y \xi)^2\right) u \, dy \, dx. \quad (5.2)$$

For studying the limit $\varepsilon \to 0$ we now assume that $V \in \mathrm{C}^2(\Upsilon)$ has exactly two non-degenerate minimizers as pure states, where $V = 0$ w.l.o.g, and one global maximum as barrier, namely

$$V(2) = V(6) = 0, \ V(y) > 0 \text{ on } \Upsilon \setminus \{2, 6\}, \ V''(2) > 0, \ V''(6) > 0; \quad (5.3)$$
$$V(5) > V(\overline{y}) \text{ on } \Upsilon \setminus \{5\}, \ V''(5) < 0, \quad (5.4)$$

see Figure 5.1. (Again, any two points in Υ could be taken as the pure states, and any point in between as barrier.) As a consequence w_ε concentrates in the points $y = 2$ and $y = 6$ in the limit, viz.

$$w_\varepsilon \overset{*}{\rightharpoonup} w_0 = \alpha_0 \delta_2 + \alpha_1 \delta_6 \in \mathrm{Prob}(\Upsilon), \quad \alpha_0 = \frac{\sqrt{V''(6)}}{\sqrt{V''(2)} + \sqrt{V''(6)}}, \ \alpha_1 = \frac{\sqrt{V''(2)}}{\sqrt{V''(2)} + \sqrt{V''(6)}}. \quad (5.5)$$

Here the convergence means $\int_\Upsilon \phi(y) w_\varepsilon(y) \, dy \to \alpha_0 \phi(2) + \alpha_1 \phi(6)$ for all $\phi \in \mathrm{C}^0(\Upsilon)$.

The important point is now to choose the diffusion constant τ_ε sufficiently large such that the transitions between $y = 2$ and $y = 6$ can occur on times of order 1. According to Kramer's rule (see e.g. [AM*12]), this is achieved by choosing $m_\Upsilon > 0$ and setting

$$\tau_\varepsilon := m_\Upsilon \int_\Upsilon \frac{1}{w_\varepsilon(y)} \, dy, \text{ with gives } \frac{\tau_\varepsilon}{\varepsilon} \exp\left(-V(5)/\varepsilon\right) \to m_\Upsilon \frac{2\pi \left(\sqrt{V''(2)} + \sqrt{V''(6)}\right)}{\sqrt{-V''(5)} \sqrt{V''(2)V''(6)}} > 0.$$

From the concentration of w_ε in the points $\{2, 6\}$ we obtain that $\widetilde{w}_\varepsilon \in \mathrm{Prob}(Q)$ concentrates in the sets $\Omega \times \{2\}$ and $\Omega \times \{6\}$, namely

$$\widetilde{w}_\varepsilon \overset{*}{\rightharpoonup} \widetilde{w}_0 := \chi_\Omega \otimes w_0 \text{ in } \mathrm{Prob}(Q).$$

Recalling that $\mathcal{E}_\varepsilon$ is the relative entropy with respect to w_ε, the Γ-convergence $\mathcal{E}_\varepsilon \xrightarrow{\Gamma}{}^* \mathcal{E}_0$ appears natural. To be more precise concerning densities and measures, we define

$$\mathcal{E}_0(\mu) := \begin{cases} \mathbf{E}((c_0, c_1)) & \text{for } \mu = c_0 \, \mathrm{d}x \otimes \delta_2 + c_1 \, \mathrm{d}x \otimes \delta_6, \\ \infty & \text{otherwise,} \end{cases}$$

$$\text{where } \mathbf{E}((c_0, c_1)) := \int_\Omega \left(\lambda_\mathrm{B}\big(\tfrac{c_0(x)}{\alpha_0}\big)\alpha_0 + \lambda_\mathrm{B}\big(\tfrac{c_1(x)}{\alpha_1}\big)\alpha_1 \right) \mathrm{d}x. \tag{5.6}$$

Proposition 5.1 *We have $\mathcal{E}_\varepsilon \xrightarrow{\Gamma}{}^* \mathcal{E}_0$ in the weak* topology of $\mathrm{Prob}(Q)$.*

Proof: The liminf estimate is established in [AGS05, Lem. 9.4.3].

To construct recovery sequences, we may restrict to the case $\mathcal{E}_0(\mu) < \infty$, since otherwise the liminf estimate provides the result. Hence, we may assume $\mu = c_0 \, \mathrm{d}x \otimes \delta_2 + c_1 \, \mathrm{d}x \otimes \delta_6$ and, using a nonnegative, continuous cut-off function $\chi : y \mapsto \max\{1 - |y|, 0\}$, we can define the measures

$$\mu_\varepsilon = u_\varepsilon(x, y) \, \mathrm{d}x \otimes \mathrm{d}y \text{ with } u_\varepsilon(x, y) = c_0(x) w_\varepsilon(y) \beta_{0,\varepsilon} \chi(y-2) + c_1(x) w_\varepsilon(y) \beta_{1,\varepsilon} \chi(y-6),$$

where the normalization constants $\beta_{j,\varepsilon}$ are given by

$$\beta_{0,\varepsilon} \int_\Upsilon w_\varepsilon(y) \chi(y-2) \, \mathrm{d}y = 1 = \beta_{1,\varepsilon} \int_\Upsilon w_\varepsilon(y) \chi(y-6) \, \mathrm{d}y,$$

which implies $\beta_{j,\varepsilon} \to 1/\alpha_j > 0$ for $\varepsilon \to 0$. Then, we easily find $\mathcal{E}_\varepsilon(u_\varepsilon) \to \mathcal{E}_0(\mu)$. ∎

Thus, the limit evolution will be described by the densities c_0 and c_1 on Ω for the particles being in the pure states $y = 2$ and $y = 6$, respectively. In particular, in the limit $\varepsilon \to 0$ the time that the particles spend along the reaction path away from these points, i.e. in $\Upsilon \setminus \{2, 6\}$, is 0.

One difficulty in deriving the liminf estimate for De Giorgi's dissipation functional

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

is that $\mathcal{R}_\varepsilon$ is only implicitly defined via the Legendre transform of $\mathcal{R}_\varepsilon^*$. Moreover, we are not able to employ the classical Wasserstein gradient flow theory in [AGS05] using the Benamou-Brenier formulation, because of the different roles of the diffusion in x with mobility m_Ω and the diffusion in y with mobility $\tau_\varepsilon \to \infty$. The first step to establish the following result follows the idea in [MaM15a], where one obtains a lower estimate by replacing $\mathcal{R}_\varepsilon(u, \dot{u})$ by the smaller term $\langle \xi_\varepsilon, \dot{u}_\varepsilon \rangle - \mathcal{R}_\varepsilon^*(u_\varepsilon, \xi_\varepsilon)$ and by choosing a suitable recovery sequence $\xi_\varepsilon \to \xi_0$ for the limit passage $\varepsilon \to 0$. Finally, one takes the supremum over all ξ_0 to recover $\mathcal{R}_0$ as dual of $\mathcal{R}_0^*$. The second step involves a suitable transformation of the reaction variable $z = Z_\varepsilon(y)$ (first introduced in [AM*12]) which allows us to control the relative densities $v_\varepsilon := u_\varepsilon / w_\varepsilon$ and the dual potentials ξ_ε along the reaction path Υ.

In the following result we will again describe the limit GS $(\mathrm{Prob}(Q), \mathcal{E}_0, \mathcal{R}_0)$ by a reduced GS $(\mathrm{Prob}(\Omega \times \{0, 1\}), \mathbf{E}, \mathbf{R})$, since in the limit every $\mu \in \mathrm{Prob}(Q)$ with finite relative entropy satisfies $\mu = c_0 \, \mathrm{d}x \otimes \delta_2(y) + c_1 \, \mathrm{d}x \otimes \delta_6(y)$ with $(c_0 \, \mathrm{d}x, c_1 \, \mathrm{d}x) \in \mathrm{Prob}(\Omega \times \{0, 1\})$, see (5.6).

Theorem 5.2 (From diffusion to reaction-diffusion) *The family of gradient systems $(\mathrm{Prob}(Q), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ defined via (5.2) converges in the EDP sense to the gradient system*

$(\mathrm{Prob}(Q), \mathcal{E}_0, \mathcal{R}_0)$, *where $\mathcal{E}_0$ is given in* (5.6) *via* $\mathbf{E}$ *and accordingly $\mathcal{R}_0$ is given via* $\mathbf{R}$, *which is defined in terms of the dual dissipation potential*

$$\mathbf{R}^*(\mathbf{c}, \boldsymbol{\eta}) := \int_\Omega \frac{m_\Omega}{2} \big(c_0 |\nabla_x \eta_0|^2 + c_1 |\nabla_x \eta_1|^2 \big) + m_\Upsilon \sqrt{\tfrac{c_0 c_1}{\alpha_0 \alpha_1}} \mathscr{C}^*(\eta_1 - \eta_0) \, \mathrm{d}x.$$

The above result means that the limiting GS is a generalized gradient system defined for $\mathbf{c} = (c_0, c_1) \in \mathrm{Prob}(\Omega \times \{0, 1\})$, where the limiting system is the coupled system of linear PDEs given in the form

$$\dot{c}_0 = m_\Omega \Delta c_0 - m_\Upsilon \big(c_0/\alpha_0 - c_1/\alpha_1 \big), \quad \dot{c}_1 = m_\Omega \Delta c_1 + m_\Upsilon \big(c_0/\alpha_0 - c_1/\alpha_1 \big),$$

with Neumann boundary conditions $\nabla c_j \cdot \nu = 0$. We emphasize that the original GS $(\mathrm{Prob}(Q), \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ is the classical GS for the Fokker-Planck equation, while the EDP limit provides the generalized gradient structure discussed in Section 2.4.3. We observe that for $\varepsilon > 0$ as well as for $\varepsilon = 0$ we have the GS that is induced by the large-deviation principle discussed in Section 2.4. Thus, we have found another instance of the interchangeability of the large-deviation principle and the EDP-limit, as displayed in Figure 1.1.

Sketch of proof of Theorem 5.2: Since the Γ-convergence $\mathcal{E}_\varepsilon \xrightarrow{\Gamma_*} \mathcal{E}_0$ was already established in Proposition 5.1, it remains to show the liminf estimate for the dissipation functional $\mathscr{D}_\varepsilon$. More precisely, assume $u_\varepsilon(t) \xrightarrow{*} \mu(t) = c_0(t) \, \mathrm{d}x \otimes \delta_2(y) + c_1(t) \, \mathrm{d}x \otimes \delta_6(y)$ in $\mathrm{Prob}(Q)$ for all $t \in [0, T]$ such that $\sup_{t \in [0,T]} \mathcal{E}_\varepsilon(u_\varepsilon) < \infty$; then, we have to show

$$\liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(u_\varepsilon) \geq \mathbf{D}(\mathbf{c}) := \int_0^T \mathbf{R}(\mathbf{c}, \dot{\mathbf{c}}) + \mathbf{R}^*(\mathbf{c}, -\mathrm{D}\mathbf{E}(\mathbf{c})) \, \mathrm{d}t. \tag{5.7}$$

Step 1. Dualization of $\mathcal{R}_\varepsilon$: The first major idea follows [MaM15a] and exploits the definition of $\mathcal{R}_\varepsilon$ as Legendre transform of $\mathcal{R}_\varepsilon^*$. Introducing the functional

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

we easily see that $\mathscr{D}_\varepsilon(u)$ can be reconstructed via $\sup_\xi \mathscr{B}_\varepsilon(u, \xi)$. Using the definitions of $\mathcal{E}_\varepsilon$ and $\mathcal{R}_\varepsilon^*$ we have the explicit form

$$\mathscr{B}_\varepsilon(u, \xi) = \int_0^T \int_Q \Big[\xi \dot{u} - \frac{m_\Omega}{2} |\nabla_x \xi|^2 u - \frac{\tau_\varepsilon}{2} (\partial_y \xi)^2 u$$
$$+ \frac{m_\Omega}{2} \frac{|\nabla_x u|^2}{u} + \frac{\tau_\varepsilon}{2} \big(\partial_y \big(\log(u/w_\varepsilon) \big) \big)^2 u \Big] \, \mathrm{d}y \, \mathrm{d}x \, \mathrm{d}t.$$

Step 2. Rescaling the reaction-path variable. The second major idea follows [AM*12, Sec. 2.1], where no x-direction was present. We define the diffeomorphism $Z_\varepsilon : \Upsilon \to \mathcal{Z} := [0, 1]$ and its inverse $Y_\varepsilon = Z_\varepsilon^{-1} : \mathcal{Z} \to \Upsilon$ via

$$z = Z_\varepsilon(y) := \frac{m_\Upsilon}{\tau_\varepsilon} \int_{\overline{y}=0}^y \frac{1}{w_\varepsilon(\overline{y})} \, \mathrm{d}\overline{y} \quad \text{and} \quad Y_\varepsilon'(z) = \frac{\tau_\varepsilon}{m_\Upsilon} w_\varepsilon(Y_\varepsilon(z)).$$

The transformed equilibrium density $\widehat{w}_\varepsilon$ on $\mathcal{Z}$ is

$$\widehat{w}_\varepsilon(z) := w_\varepsilon(Y_\varepsilon(z)) Y_\varepsilon'(z) \quad \text{and satisfies} \quad \widehat{w}_\varepsilon \xrightarrow{*} \widehat{w}_0 := \alpha_0 \delta_0 + \alpha_1 \delta_1. \tag{5.8}$$

Indeed, for the latter statement we first use that for all $g \in \mathrm{C}^0(\mathcal{Z})$ we have the identity $\int_{\mathcal{Z}} g(z)\widehat{w}_\varepsilon(z)\,\mathrm{d}z = \int_\Upsilon g(Z_\varepsilon(y))w_\varepsilon(y)\,\mathrm{d}y$. Recalling that V has a unique global maximum at $y = 5$, the function Z_ε converges uniformly on compact subsets of $\Upsilon \setminus \{5\}$ to the step function $Z_0(y) = 0$ for $y < 5$ and $Z_0(y) = 1$ for $y > 5$. With this and (5.5) we conclude $\int_\Upsilon g(Z_\varepsilon(y))w_\varepsilon(y)\,\mathrm{d}y \to \alpha_0 g(Z_0(2)) + \alpha_1 g(Z_0(6))$ which is the desired result (5.8).

To estimate $\mathscr{B}_\varepsilon$ in the limit $\varepsilon \to 0$ we use now the independent variable $z = Z_\varepsilon(y)$ and the dependent variables

$$v(t, x, z) = \frac{u(t, x, Y_\varepsilon(z))}{w_\varepsilon(Y_\varepsilon(z))} \quad \text{and} \quad \zeta(t, x, z) = \xi(t, x, Y_\varepsilon(z)).$$

Introducing the domain $\widehat{Q} = \Omega \times \mathcal{Z}$ we find $\mathscr{B}_\varepsilon(u, \xi) = \widehat{\mathscr{B}}_\varepsilon(v, \zeta)$ with

$$\widehat{\mathscr{B}}_\varepsilon(v, \zeta) = \int_0^T\!\!\int_{\widehat{Q}} \zeta \dot{v}\widehat{w}_\varepsilon - \frac{m_\Omega}{2}|\nabla_x \zeta|^2 v\widehat{w}_\varepsilon - \frac{m_\Upsilon}{2}(\partial_z\zeta)^2 v + \frac{m_\Omega}{2}\frac{|\nabla_x v|^2}{v}\widehat{w}_\varepsilon + \frac{m_\Upsilon}{2}\frac{(\partial_z v)^2}{v}\ \mathrm{d}z\,\mathrm{d}x\,\mathrm{d}t.$$

The transformation of $\mathscr{B}_\varepsilon$ to $\widehat{\mathscr{B}}_\varepsilon$ follows easily by using the relations

$$u\,\mathrm{d}y = v\widehat{w}_\varepsilon\,\mathrm{d}z, \quad \partial_y\xi = \frac{\partial_z\zeta}{Y_\varepsilon'(z)}, \quad \text{and} \quad \frac{\tau_\varepsilon}{(Y_\varepsilon'(z))^2} = \frac{m_\Upsilon}{\widehat{w}_\varepsilon(z)}.$$

Step 3. The Γ-limit for $\widehat{\mathscr{B}}_\varepsilon(\cdot, \zeta)$: The importance of the new form $\widehat{\mathscr{B}}_\varepsilon$ is that the dependence on ε only occurs in the weighting measure $\widehat{w}_\varepsilon$. Since $\widehat{w}_\varepsilon$ concentrates in the points $z = 0$ and 1, the three terms that are multiplied by the weight $\widehat{w}_\varepsilon$ will converge to simple integrals over $[0, T] \times \Omega$ for the densities c_0 and c_1 respectively. In contrast there are two terms not involving $\widehat{w}_\varepsilon$, but these terms only involve derivatives in the z-direction. In particular, they control the smoothness of ζ and v in z-direction, namely $\sqrt{v} \in \mathrm{L}^2([0, T] \times \Omega; \mathrm{H}^1(\mathcal{Z}))$ such that $v\widehat{w}_\varepsilon$ indeed has a well-defined limit. With this and $\widehat{w}_\varepsilon \overset{*}{\rightharpoonup} \widehat{w}_0 = \alpha_0\delta_1 + \alpha_1\delta_1$ (cf. (5.8)), it is possible to show that for fixed and sufficiently smooth ζ we have $\widehat{\mathscr{B}}_\varepsilon(\cdot, \zeta) \overset{\Gamma}{\to} \widehat{\mathscr{B}}_0(\cdot, \zeta)$ with

$$\widehat{\mathscr{B}}_0(v, \zeta) = \int_0^T\!\!\int_\Omega \left[\int_{\mathcal{Z}} \frac{m_\Upsilon}{2}\left(\frac{(\partial_z v)^2}{v} - (\partial_z\zeta)^2 v\right)\mathrm{d}z \right.$$
$$\left. + \sum_{j=0}^1 \alpha_j\left(\zeta_j\dot{v}_j - \frac{m_\Omega}{2}v_j|\nabla_x\zeta_j|^2 + \frac{m_\Omega}{2}\frac{|\nabla_x v_j|^2}{v_j}\right) \right] \mathrm{d}x\,\mathrm{d}t,$$

where $v_j(t, x) = v(t, x, j)$ and $\zeta_j(t, x) = \zeta(t, x, j)$ for $j = 0$ and $j = 1$.

Step 4. Minimization over the reduction path profile. Note that in the definition of $\widehat{\mathscr{B}}_0$, the values of the functions v and ζ for $z \in \,]0, 1[$ only occur in the first integrand (with factor m_Υ). Hence, one can eliminate the integral by taking the supremum in ζ and the infimum in v for given boundary values at $z = 0$ and $z = 1$. The relevant functional reads

$$\mathcal{N}(v, \zeta) = \int_0^1 \left(\frac{v'(z)^2}{2v(z)} - \frac{1}{2}\zeta'(z)^2 v(z)\right)\mathrm{d}z,$$

and Proposition A.3 provides the following explicit inf-sup formula

$$\inf\left\{ \sup\left\{ \mathcal{N}(v, \zeta) \,\middle|\, \zeta(0) = \zeta_0,\ \zeta(1) = \zeta_1 \right\} \,\middle|\, v(0) = v_0,\ v(1) = v_1,\ v > 0 \right\}$$
$$= \sqrt{v_0 v_1}\,\mathscr{C}^*(\log v_1 - \log v_0) - \sqrt{v_0 v_1}\,\mathscr{C}^*(\zeta_1 - \zeta_0) =: N(\zeta_1 - \zeta_0, v_0, v_1).$$

Thus, we can reduce $\widehat{\mathscr{B}}_0$ to a functional $\mathbf{B}$ on $\mathbf{v} = (v_0, v_1)$ and $\boldsymbol{\zeta} = (\zeta_0, \zeta_1)$, namely

$$\mathbf{B}(\mathbf{v}, \boldsymbol{\zeta}) := \int_0^T \!\!\! \int_\Omega \left[m_\Upsilon N(\zeta_1 - \zeta_0, v_0, v_1) \right.$$
$$\left. + \sum_{j=0}^1 \alpha_j \left(\zeta_j \dot{v}_j - \frac{m_\Omega}{2} v_j |\nabla_x \zeta_j|^2 + \frac{m_\Omega}{2} \frac{|\nabla_x v_j|^2}{v_j} \right) \right] \mathrm{d}x\,\mathrm{d}t.$$

The inf-sup definition of N provides the following relation between $\widehat{\mathscr{B}}_0$ and $\mathbf{B}$:

$$\forall\, v \text{ with } v|_{[0,T]\times\Omega\times\{0,1\}} = \mathbf{v}\ \exists\, \zeta \text{ with } \zeta|_{[0,T]\times\Omega\times\{0,1\}} = \boldsymbol{\zeta}: \quad \widehat{\mathscr{B}}_0(v,\zeta) \geq \mathbf{B}(\mathbf{v}, \boldsymbol{\zeta}). \quad (5.9)$$

Step 5. Identification of the limits: It now remains to relate the limit functions $\mathbf{v}$ to the weak limit of the sequence u_ε. For this, we consider a sequence u_ε as in (3.4c), i.e. $u_\varepsilon \overset{*}{\rightharpoonup} u$ and $\mathcal{E}_\varepsilon(u_\varepsilon(t)) \leq C < \infty$. By the definition of v we have

$$\int_0^T \!\!\! \int_Q u_\varepsilon(t,x,y)\phi(t,x,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \!\!\! \int_Q v_\varepsilon(t,x,Z_\varepsilon(y))w_\varepsilon(y)\phi(t,x,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t. \quad (5.10)$$

Without loss of generality we assume $\infty > C \geq \mathscr{D}_\varepsilon(u_\varepsilon) = \sup_\xi \mathscr{B}_\varepsilon(u_\varepsilon, \xi) = \sup_\zeta \widehat{\mathscr{B}}(v_\varepsilon, \zeta)$, which gives the bound

$$\|\sqrt{v_\varepsilon}\|_{\mathrm{L}^2([0,T]\times\Omega;\mathrm{H}^1(\mathcal{Z}))} \leq C. \quad (5.11)$$

This implies Hölder continuity of $v(t,x,\cdot): \mathcal{Z} \to \mathbb{R}$. Moreover, Z_ε converges uniformly to 0 and 1 near $y = 2$ and $y = 6$. Hence, we can pass to the limit in (5.10) and obtain

$$\int_0^T \!\!\! \int_\Omega c_1(t,x)\phi(t,x,2) + c_2(t,x)\phi(t,x,6)\,\mathrm{d}x\,\mathrm{d}t = \lim_{\varepsilon\to 0} \int_0^T \!\!\! \int_Q u_\varepsilon(t,x,y)\phi(t,x,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t$$
$$= \lim_{\varepsilon\to 0} \int_0^T \!\!\! \int_Q v_\varepsilon(t,x,Z_\varepsilon(y)w_\varepsilon(y)\phi(t,x,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \!\!\! \int_\Omega \sum_0^1 v_j(t,x)\alpha_j\phi(t,x,2+4j)\,\mathrm{d}x\,\mathrm{d}t,$$

which means $\mathbf{c} = (c_0, c_1) = (\alpha_0 v_0, \alpha_1 v_1)$. Using the explicit form of $\mathbf{R}^*$ and $\mathbf{E}$ implies

$$\int_0^T \langle \boldsymbol{\zeta}, \dot{\mathbf{c}} \rangle - \mathbf{R}^*(\mathbf{c}, \boldsymbol{\zeta}) + \mathbf{R}^*(\mathbf{c}, -\mathrm{D}\mathbf{E}(\mathbf{c}))\,\mathrm{d}t = \mathbf{B}\big((c_1/\alpha_0, c_2/\alpha_1), \boldsymbol{\zeta}\big). \quad (5.12)$$

Step 6. The liminf estimate: With these preparations we can now complete the liminf estimate. By the construction of v_ε and $\xi_\varepsilon(t,x,y) = \zeta(t,x,Z_\varepsilon(y))$ we obtain the relations

$$\mathscr{D}_\varepsilon(u_\varepsilon) \geq \mathscr{B}_\varepsilon(u_\varepsilon, \xi_\varepsilon) = \widehat{\mathscr{B}}_\varepsilon(v_\varepsilon, \zeta),$$

where ζ is now fixed. For the sequence u_ε as given in Step 5, we can further assume that $v_\varepsilon \rightharpoonup v$ in $\mathrm{L}^2([0,T]\times\Omega;\mathrm{C}^0(\mathcal{Z}))$, using (5.11). According to Step 3 the liminf for $\varepsilon \to 0$ yields

$$\liminf_{\varepsilon\to 0} \mathscr{D}_\varepsilon(u_\varepsilon) \geq \liminf_{\varepsilon\to 0} \widehat{\mathscr{B}}_\varepsilon(v_\varepsilon, \zeta) \geq \widehat{\mathscr{B}}_0(v, \zeta) \geq \mathbf{B}\big((c_0/\alpha_0, c_1/\alpha_1), \boldsymbol{\zeta}\big),$$

where for the last estimate we have to choose ζ according to (5.9) to fit the limit v and $\boldsymbol{\zeta} = \zeta|_{[0,T]\times\Omega\times\{0,1\}}$. Nevertheless, the functions $\boldsymbol{\zeta} = (\zeta_0, \zeta_1)$ are still free. Using the

characterization (5.12) and taking the supremum over all $\boldsymbol{\zeta}$ gives the desired lower bound:

$$\liminf_{\varepsilon \to 0} \mathscr{D}_\varepsilon(u_\varepsilon) \ \geq \ \sup_\zeta \widehat{\mathscr{B}}_0(v, \zeta) \ \geq \ \sup_\zeta \int_0^T \big(\langle \boldsymbol{\zeta}, \dot{\mathbf{c}} \rangle - \mathbf{R}^*(\mathbf{c}, \boldsymbol{\zeta}) + \mathbf{R}^*(\mathbf{c}, -\mathrm{D}\mathbf{E}(\mathbf{c})) \big)\, \mathrm{d}t$$

$$= \int_0^T \big(\mathbf{R}(\mathbf{c}, \dot{\mathbf{c}}) + \mathbf{R}^*(\mathbf{c}, -\mathrm{D}\mathbf{E}(\mathbf{c})) \big)\, \mathrm{d}t \ = \ \mathbf{D}(\mathbf{c}).$$

Thus the desired estimate (5.7) is established, which finishes the proof of Theorem 5.2. ∎

A   Evaluation of some functionals

Here we give explicit calculations for the functional $\mathcal{G}$ occurring in the membrane limit and the functional $\mathcal{N}$ occurring in the limit of diffusion to reaction. It is surprising that both functional are closely related, see (A.5).

A.1   Derivation of the potential $G(\alpha, u_0, u_1)$

We first give the result of the standard case of constant coefficients $\widehat{A}$ and $\widehat{W}$, which was already derived in [MPR14, Prop. 4.4] under the restriction $u_0 + u_1 = 1$. For the functional

$$\mathcal{G}(\alpha, u) := \int_0^1 \frac{\alpha^2 + u'(x)^2}{2u(x)}\, \mathrm{d}x,$$

we define the value function

$$G(\alpha, u_0, u_1) := \min \Big\{\, \mathcal{G}(\alpha, u) \ \Big|\ u \in \mathrm{H}^1(0,1),\ u(0) = u_0,\ u(1) = u_1,\ u > 0 \,\Big\}, \qquad \text{(A.1)}$$

and give a full proof of the derivation of the explicit formula.

Proposition A.1 *For all $\alpha \in \mathbb{R}$ and u_0, u_1 we have*

$$G(\alpha, u_0, u_1) = \sqrt{u_0 u_1}\, \mathscr{C}\Big(\frac{\alpha}{\sqrt{u_0 u_1}}\Big) + \sqrt{u_0 u_1}\, \mathscr{C}^*(\log u_1 - \log u_0), \qquad \text{(A.2)}$$

where the last term simplifies to $G(0, u_0, u_1) = 2(\sqrt{u_0} - \sqrt{u_1})^2$. Moreover, the unique minimizer is given by

$$u(x) = (1-x)u_0 + xu_1 + b(x^2 - x) \quad \text{with } b = u_0 + u_1 - \sqrt{\alpha^2 + 4u_0 u_1}. \qquad \text{(A.3)}$$

Proof: Since the integrand is convex, there is a unique minimizer u. Denoting the integrand by $f(u, u')$ the Euler-Lagrange equations $-\big(\partial_{u'} f(u, u')\big)' + f(u, u') = 0$ are $uu'' - (u')^2 + \alpha^2 = 0$. By Noether's theorem we also have the first integral $u' \partial_{u'} f(u, u') - f(u, u') = \big((u')^2 - \alpha^2\big)/(2u) = \gamma/2 = $const. From $(u')^2 = \alpha^2 + \gamma u$ it is now easy to see that all solutions of the Euler-Lagrange equations are parabolas. Using the boundary conditions we find u in (A.3), where $\gamma = 4b$.

To evaluate the integral we restrict to the case $u'(x) > 0$ on $[0,1]$, which means $2u_0 < \sqrt{\alpha^2 + 4u_0 u_1} < 2u_1$. In the other cases, one can do the calculation on all monotone parts in a similar fashion and add the result. We use (A.3) and $(u')^2 = \alpha^2 + \gamma u$ to obtain

$$\int_0^1 \frac{\alpha^2 + u'(x)^2}{2u(x)}\, \mathrm{d}x = \frac{\gamma}{2} + \alpha^2 \int_0^1 \frac{\mathrm{d}x}{u(x)} = 2b + \alpha^2 \int_{u_0}^{u_1} \frac{\mathrm{d}u}{u\sqrt{\alpha^2 + \gamma u}},$$

$$= 2b - 2\alpha\, \mathrm{artanh}\, \sqrt{1 + \tfrac{\gamma}{\alpha^2} u_1} + 2\alpha\, \mathrm{artanh}\, \sqrt{1 + \tfrac{\gamma}{\alpha^2} u_0}.$$

To proceed we first observe $\alpha^2+\gamma u_j = \left(\sqrt{\alpha^2+4u_0u_1} - 2u_j\right)^2$, which gives $\sqrt{\alpha^2+\gamma u_0} = \sqrt{\alpha^2+4u_0u_1} - 2u_0$ and $\sqrt{\alpha^2+\gamma u_1} = 2u_1 - \sqrt{\alpha^2+4u_0u_1}$. Now employing the addition rule $\operatorname{artanh}(x) + \operatorname{artanh}(y) = \operatorname{arsinh}\left((x+y)/\sqrt{(1-x^2)(1-y^2)}\right)$ and $\gamma = 4b$ gives (A.2). ∎

In Section 4 we need a more general version with non-constant functions $\widehat{A}$ and $\widehat{W}$:

$$\widehat{\mathcal{G}}(\alpha, U) := \int_0^1 \frac{\alpha^2}{2\widehat{A}U} + \frac{\widehat{A}U}{2}\left(\left(\log(U/\widehat{W})\right)'\right)^2 \mathrm{d}y. \tag{A.4}$$

We will show that the influence of the coefficient functions $\widehat{A}$ and $\widehat{W}$ can be calculated from Proposition A.1 by a suitable rescaling of the layer variable in the form $x = X(y)$.

Proposition A.2 *We have the following formula:*

$$\widehat{G}(\alpha, u_0, u_1) := \min\{\,\widehat{\mathcal{G}}(\alpha, U) \mid U > 0,\ U(0) = u_0,\ U(1) = u_1\,\}$$
$$= A_* \sqrt{\frac{u_0u_1}{w_0w_1}}\, \mathscr{C}\left(\frac{1}{A_*}\sqrt{\frac{w_0w_1}{u_0u_1}}\,\alpha\right) + A_* \sqrt{\frac{u_0u_1}{w_0w_1}}\, \mathscr{C}^*\left(\log\left(\frac{u_0w_1}{u_1w_0}\right)\right),$$

where $w_0 = \widehat{W}(0)$, $w_1 = \widehat{W}(1)$, and $A_* = \operatorname{Harm}(\widehat{A}\widehat{W}) = \left(\int_0^1 1/(\widehat{A}(y)\widehat{W}(y))\,\mathrm{d}y\right)^{-1}$.

Proof: We define the new independent variable z and a new function $v(z)$ via

$$z = Z(y) := A_* \int_0^y \frac{\mathrm{d}\eta}{\widehat{A}(\eta)\widehat{W}(\eta)} \quad \text{and} \quad v(Z(y)) = \frac{U(y)}{\widehat{W}(y)},$$

where $A_* = \left(\int_0^1 1/(\widehat{A}(\eta)\widehat{W}(\eta))\,\mathrm{d}\eta\right)^{-1}$. By definition we have $Z(0) = 0$ and $Z(1) = 1$, and the inverse Y of Z maps $[0,1]$ into itself again. Hence, using $Z'(y) = A_*/(\widehat{A}(y)\widehat{W}(y))$ the functional $\widehat{\mathcal{G}}$ from (A.4) is transformed into $\mathcal{G}$ via $\widehat{\mathcal{G}}(\alpha, U) = A_*\,\mathcal{G}(\alpha/A_*, v)$. The result of Proposition A.2 now follows from Proposition A.1 via

$$\widehat{G}(\alpha, u_0, u_1) = \min\{\,\widehat{\mathcal{G}}(\alpha, U) \mid U(0) = u_0,\ U(1) = u_1\,\}$$
$$= \min\{\,A_*\mathcal{G}(\alpha/A_*, v) \mid v(0) = U(0)/\widehat{W}(0),\ v(1) = U(1)/\widehat{W}(1)\,\}.$$

Thus, the asserted formula is established. ∎

A.2 Derivation of the potential N

We consider the functional

$$\mathcal{N}(v, \zeta) = \int_0^1 \left(\frac{v'(z)^2}{2v(z)} - \frac{1}{2}\zeta'(z)^2 v(z)\right) \mathrm{d}z$$

for functions $v > 0$. Hence, $\mathcal{N}$ is convex in v and concave in ζ. We are interested in the inf-sup for given boundary values, namely

$$N(\delta, v_0, v_1) := \inf\left\{\,\sup\left\{\,\mathcal{N}(v, \zeta) \mid \zeta(1)-\zeta(0) = \delta\,\right\}\,\Big|\, v(0) = v_0,\ v(1) = v_1,\ v > 0\,\right\}.$$

The following result provides an explicit formula in terms of the dual dissipation potential $\mathscr{C}^*$. It is based on the following surprising relation between $\mathcal{N}$ and $\mathcal{G}$ from (A.1):

$$\mathcal{M}(\delta, v) := \max\left\{\,\mathcal{N}(v, \zeta) \mid \zeta(1)-\zeta(0) = \delta\,\right\} \overset{!!}{=} \min\left\{\,\mathcal{G}(\alpha, v) - \alpha\delta \mid \alpha \in \mathbb{R}\,\right\}. \tag{A.5}$$

The equality $\overset{!!}{=}$ can be checked by elementary calculations, since in both cases we find

$$\mathcal{M}(\delta, v) = \int_0^1 \frac{v'(z)^2}{2v(z)} \, \mathrm{d}z - \frac{\delta^2}{2} \mathrm{Harm}(v), \quad \text{where } \mathrm{Harm}(v) = \left(\int_0^1 \frac{\mathrm{d}z}{v(z)} \right)^{-1}.$$

Using the strong link (A.5) between $\mathcal{N}$ and $\mathcal{G}$ we show that N can be calculated from G.

Proposition A.3 *We have the relation*

$$N(\delta, v_0, v_1) = \sqrt{v_0 v_1} \, \mathscr{C}^*(\log v_1 - \log v_0) - \sqrt{v_0 v_1} \, \mathscr{C}^*(\delta).$$

Proof: Using (A.5) we want to show that N is related to the Legendre transform $G^*(\delta, v_0, v_1) := \sup_{\alpha \in \mathbb{R}} \delta\alpha - G(\alpha, v_0, v_1)$ of G from (A.1). For this we keep $\delta \in \mathbb{R}$ fixed.

The functional $(\alpha, v) \mapsto \mathcal{G}(\alpha, v) - \delta\alpha$ is jointly convex, such that it can be minimized in any desired order of α and v. Letting $V := \{\, v \mid v > 0,\ v(0) = v_0,\ v(1) = v_1 \,\}$ we have

$$\begin{aligned} N(\delta, v_0, v_1) &= \inf_{v \in V} \mathcal{M}(\delta, v) = \inf_{v \in V} \left(\inf_{\alpha \in \mathbb{R}} \mathcal{G}(\alpha, v) - \delta\alpha \right) \\ &= \inf_{\alpha \in \mathbb{R}} \left(\inf_{v \in V} \mathcal{G}(\alpha, v) - \delta\alpha \right) = \inf_{\alpha \in \mathbb{R}} G(\delta, v_0, v_1) - \delta\alpha = -G^*(\delta, u_0, u_1). \end{aligned}$$

Thus, evaluating G^* with G from (A.2) explicitly gives the desired result. ∎

References

[ACJ96] R. ABEYARATNE, C.-H. CHU, and R. JAMES. Kinetics of materials with wiggly energies: theory and application to the evolution of twinning microstructures in a Cu-Al-Ni shape memory alloy. *Phil. Mag. A*, 73, 457–497, 1996.

[AD*11] S. ADAMS, N. DIRR, M. A. PELETIER, and J. ZIMMER. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Comm. Math. Phys.*, 307(3), 791–815, 2011.

[AD*13] S. ADAMS, N. DIRR, M. PELETIER, and J. ZIMMER. Large deviations and gradient flows. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 371(2005), 20120341, 17, 2013.

[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.

[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.

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

[BoP14] G. A. BONASCHI and M. A. PELETIER. Quadratic and rate-independent limits for a large-deviations functional. *Preprint*, 2014. arXiv:1409.4350.

[DiL14] K. DISSER and M. LIERO. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Networks Heterg. Materials*, 2014. To appear. WIAS preprint 1899.

[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.

[ErM12] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Rational Mech. Anal.*, 206(3), 997–1038, 2012.

[Fei72] M. Feinberg. On chemical kinetics of a certain class. *Arch. Rational Mech. Anal.*, 46, 1–41, 1972.

[FeK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.

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

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

[GK*00] A. N. Gorban, I. V. Karlin, V. B. Zmievskii, and S. V. Dymova. Reduced description in the reaction kinetics. *Physica A*, 275, 361–379, 2000.

[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.

[Grm10] M. Grmela. Multiscale equilibrium and nonequilibrium thermodynamics in chemical engineering. *Adv. Chem. Eng.*, 39, 75–128, 2010.

[HaF08] K. Hackl and F. D. Fischer. On the relation between the principle of maximum dissipation and inelastic evolution given by dissipation potentials. *Proc. R. Soc. A*, 464, 117–132, 2008.

[HaN75] B. Halphen and Q. S. Nguyen. Sur les matériaux standards généralisés. *J. Mécanique*, 14, 39–63, 1975.

[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.

[Lie12] M. Liero. *Variational Methods for Evolution*. PhD thesis, Institut für Mathematik, Humboldt-Universität zu Berlin, 2012.

[Lie13] M. Liero. Passing from bulk to bulk-surface evolution in the Allen-Cahn equation. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 20(3), 919–942, 2013.

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

[MaM15a] J. Maas and A. Mielke. On gradient structures for chemical reactions with detailed balance: I. modeling and large-volume limit. *In preparation*, 2015.

[MaM15b] J. Maas and A. Mielke. On gradient structures for chemical reactions with detailed balance: II. dissipation distances and geodesic convexity. *In preparation*, 2015.

[Mie03] A. Mielke. Energetic formulation of multiplicative elasto–plasticity using dissipation distances. *Contin. Mech. Thermodyn.*, 15, 351–382, 2003.

[Mie11] 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.

[Mie13a] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Part. Diff. Eqns.*, 48(1), 1–31, 2013.

[Mie13b] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discr. Cont. Dynam. Systems Ser. S*, 6(2), 479–499, 2013.

[Mie15a] A. Mielke. On evolutionary Γ-convergence for gradient systems. In A. Muntean, J. Rademacher, and A. Zagaris, editors, *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, 55 pp. Springer, 2015. To appear. WIAS Preprint 1915.

[Mie15b] A. Mielke. Variational approaches and methods for dissipative material models with multiple scales. In K. Hackl and S. Conti, editors, *Analysis and Computation of Microstructure in Finite Plasticity*, volume 78 of *Lect. Notes Appl. Comp. Mechanics*, pp. 125–155. Springer, 2015.

[MiR15] A. Mielke and T. Roubíček. *Rate-Independent Systems: Theory and Application.* Springer. In print, 2015.

[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.

[MP*15] A. Mielke, R. I. A. Patterson, M. A. Peletier, and D. R. M. Renger. Non-equilibrium thermodynamical principles for nonlinear chemical reactions and systems with coagulation and fragmentation. *In preparation*, 2015.

[MPR14] A. Mielke, M. A. Peletier, and D. R. M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Analysis*, 41(4), 1293–1327, 2014.

[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).

[PSV10] M. A. Peletier, G. Savarè, and M. Veneroni. From diffusion to reaction via Γ-convergence. *SIAM J. Math. Analysis*, 42(4), 1805–1825, 2010.

[PSV12] M. A. Peletier, G. Savaré, and M. Veneroni. Chemical reactions as Γ-limit of diffusion [revised reprint of [PSV10]. *SIAM Rev.*, 54(2), 327–352, 2012.

[Ray71] L. Rayleigh (Hon. J. W. Strutt). Some general theorems relating to vibrations. *Proc. London Math. Soc.*, s1-4, 357–368, 1871.

[Ren13] D. R. M. Renger. *Microscopic interpretation of Wasserstein gradient flows.* PhD thesis, Technische Universiteit Eindhoven, 2013.

[RRG00] F. Roters, D. Raabe, and G. Gottstein. Work hardening in heterogeneous alloys—a microstructural approach based on three internal variables. *Acta Materialia*, 48, 4181–4189, 2000.

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

[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.

[ZR*06] N. Zaafarani, D. Raabe, R. N. Singh, F. Roters, and S. Zaefferer. Three-dimensional investigation of the texture and microstructure below a nanoindent in a Cu single crystal using 3D EBSD and crystal plasticity finite element simulations. *Acta Materialia 54 (2006) 1863–1876*, 54, 1863–1876, 2006.