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

Preprint ISSN 2198-5855

Non-equilibrium steady states as saddle points and EDP-convergence for slow-fast gradient systems

Alexander Mielke

submitted: March 14, 2023

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
and
Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin
Germany
E-Mail: alexander.mielke@wias-berlin.de

No. 2998
Berlin 2023

2020 *Mathematics Subject Classification.* 35G25, 35K90, 47J35, 70K70, 82C05, 92E20.

Key words and phrases. Slow-fast gradient system, dual and primal dissipation potentials, EDP-convergence, chemical reaction system, reaction-diffusion equation, constrained saddle points, effective kinetic relation.

Research partially funded by Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center SFB 1114 "*Scaling Cascades in Complex Systems*" (Project no. 235221301), Subproject C05 "Effective models for materials and interfaces with multiple scales".

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

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

Non-equilibrium steady states as saddle points and EDP-convergence for slow-fast gradient systems

Alexander Mielke

Abstract

The theory of slow-fast gradient systems leads in a natural way to non-equilibrium steady states, because on the slow time scale the fast subsystem stays in steady states that are driven by the interaction with the slow system. Using the theory of convergence of gradient systems in the sense of the energy-dissipation principle shows that there is a natural characterization of these non-equilibrium steady states as saddle points of a Lagrangian where the slow variables are fixed. We give applications to slow-fast reaction-diffusion systems based on the so-called cosh-type gradient structure for reactions. It is shown that two binary reaction give rise to a ternary reaction with a state-dependent reaction coefficient. Moreover, we show that a reaction-diffusion equation with a thin membrane-like layer convergences to a transmission condition, where the formerly quadratic dissipation potential for diffusion convergences to a cosh-type dissipation potential for the transmission in the membrane limit.

1 Introduction

A gradient system (GS) is a triple $(M, \mathcal{E}, \mathcal{R})$ where M is the state space, which is either a smooth manifold or a convex subset of a Banach space X such that the tangent spaces $\mathrm{T}_u M$ and cotangent spaces $\mathrm{T}^*_u M$ are well defined for $u \in M$. For notational simplicity we restrict to the case that M is equal to a reflexive Banach space X such that $\mathrm{T}_u M = X$ and $\mathrm{T}^*_u M = X^*$.

The energy functional $\mathcal{E} : X \to \mathbb{R}_\infty =: \,]-\infty, \infty]$ is assumed to be differentiable in a suitable subset $\mathrm{dom}(\mathrm{D}\mathcal{E})$. The function $\mathcal{R} : \mathrm{T}M = X{\times}X \to [0, \infty]$ denotes the dissipation potential, which means that for all $u \in X$, the function $\mathcal{R}(u, \cdot) : \mathrm{T}_u X \to [0, \infty]$ is lower semi-continuous, convex and satisfies $\mathcal{R}(u, 0) = 0$. By $\mathcal{R}^* : \mathrm{T}^*M = X{\times}X^* \to [0, \infty]$ we denote the dual dissipation potential which is defined via

$$\mathcal{R}^*(u, \xi) := \sup \big\{ \langle \xi, v\rangle - \mathcal{R}(u, v) \,\big|\, v \in X \big\}.$$

The gradient-flow equation associated with the GS $(X, \mathcal{E}, \mathcal{R})$ can be written in two equivalent forms, namely

$$\text{(I)}\;\; 0 \in \partial_v \mathcal{R}(u, \dot u) + \mathrm{D}\mathcal{E}(u) \quad \Longleftrightarrow \quad \text{(II)}\;\; \dot u \in \partial_\xi \mathcal{R}^*\big(u, -\mathrm{D}\mathcal{E}(u)\big),$$

where $\partial_v \mathcal{R}$ and $\partial_\xi \mathcal{R}^*$ denote the subdifferentials of the convex functions $\mathcal{R}(u, \cdot)$ and $\mathcal{R}^*(u, \cdot)$, respectively, see [Mie16] and the references therein.

Under suitable technical assumptions (such as the validity of a suitable abstract chain rule) there is a third equivalent formulation according to the *energy-dissipation principle* (EDP). If $u : [0, T] \to X$ satisfies $\int_0^T \big[\mathcal{R}(u, \dot u) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))\big]\,\mathrm{d}t < \infty$, then (I) and (II) hold a.e. in $[0, T]$ if and only if the *energy-dissipation inequality* (EDI) holds, namely

$$\mathcal{E}(u(T)) + \int_0^T \Big(\mathcal{R}(u, \dot u) + \mathcal{R}^*\big(u, -\mathrm{D}\mathcal{E}(u)\big)\Big)\,\mathrm{d}t \le \mathcal{E}(u(0)). \tag{1.1}$$

DOI 10.20347/WIAS.PREPRINT.2998

If we now have a family $(X,\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)_{\varepsilon>0}$ of GS with a small parameter $\varepsilon>0$, we say that this family converges in the sense of the EDP to the limit $(X,\mathcal{E}_{\mathrm{eff}},\mathcal{R}_{\mathrm{eff}})$ if we have the following Γ-convergences

$$\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_{\mathrm{eff}} \text{ in } X \quad \text{and} \quad \mathfrak{D}_\varepsilon \xrightarrow{\Gamma} \mathfrak{D}_0 \text{ in } \mathrm{L}^2([0,T];X), \tag{1.2}$$

where the dissipation functionals $\mathfrak{D}_\varepsilon : \mathrm{L}^2([0,T];X \to [0,\infty]$ are defined as follows:

$$\mathfrak{D}_\varepsilon\big(u(\cdot)\big) := \int_0^T \Big(\mathcal{R}_\varepsilon(u,\dot u)+\mathcal{R}_\varepsilon^*\big(u,-\mathrm{D}\mathcal{E}_\varepsilon(u)\big)\Big)\,\mathrm{d}t \qquad \text{for } \varepsilon>0,$$

$$\mathfrak{D}_0\big(u(\cdot)\big) := \int_0^T \Big(\mathcal{R}_{\mathrm{eff}}(u,\dot u)+\mathcal{R}_{\mathrm{eff}}^*\big(u,-\mathrm{D}\mathcal{E}_{\mathrm{eff}}(u)\big)\Big)\,\mathrm{d}t \qquad \text{for } \varepsilon=0.$$

We refer to [LM*17] for the first discussion of this concept, to [MMP21] for refinements, and to [DFM19, Fre19, MPS21, FrL21, PeS22] for various applications of this approach.

We emphasize two important properties of EDP-convergence: The first simply states that if u_ε are solutions to $(X,\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$ and we have convergence $u_\varepsilon(t)\to u(t)$ in a suitable way and we have well-prepared initial conditions, i.e. $\mathcal{E}_\varepsilon(u_\varepsilon(0))\to\mathcal{E}_{\mathrm{eff}}(u(0))$, then u is a solution of the effective gradient system $(X,\mathcal{E}_{\mathrm{eff}},\mathcal{R}_{\mathrm{eff}})$. The second property states that $\mathcal{R}_{\mathrm{eff}}$ can be different a potentially existing Γ-limit $\mathcal{R}_0$, i.e. $\mathcal{R}_\varepsilon \xrightarrow{\Gamma} \mathcal{R}_0$. The point is that $\mathfrak{D}_\varepsilon$ involves a nonlinear construction for the pair $(\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$, which allows to transfer microscopic information of the energy (encoded in $\mathcal{E}_\varepsilon-\mathcal{E}_{\mathrm{eff}}$) into the dissipation $\mathcal{R}_{\mathrm{eff}}$. We will see this below in Section 5.3 where $\mathcal{R}_\varepsilon^*(u,\cdot)$ is quadratic and has a quadratic limit $\mathcal{R}_0^*$ but $\mathcal{R}_{\mathrm{eff}}^*$ contains a cosh-type membrane part for the transmission through the membrane.

The slow-fast GS under consider are assumed to be of the following form

$$X=X_{\mathrm{slow}}\times X_{\mathrm{fast}},\quad \mathcal{E}_\varepsilon(U,w)=E(U)+\varepsilon\, e(w),\quad \mathcal{R}_\varepsilon^*(U,w;\Xi,\xi)=\overline{\mathcal{R}}^*\big(U,w;\Xi,\tfrac{1}{\varepsilon}\,\xi\big),$$

which provides only one class of GS where slow-fast effects can be studied (see [MiS20, MPS21] for other scalings). The associated gradient-flow equation reads

$$\begin{pmatrix}\dot U\\ \varepsilon\,\dot w\end{pmatrix} = \partial_{\Xi,\zeta}\overline{\mathcal{R}}^*\big(U,w;-\mathrm{D}E(U),-\mathrm{D}e(w)\big),$$

which shows nicely the slow-fast structure, because ε only appears once, namely in front of the time derivative $\dot w$ of the fast variable $w\in X_{\mathrm{fast}}$.

To pass to the EDP limit we observe that $\mathfrak{D}_\varepsilon$ takes the simple form

$$\mathfrak{D}_\varepsilon\big(U,w\big) := \int_0^T \Big(\overline{\mathcal{R}}(U,w;\dot U,\varepsilon\,\dot w)+\overline{\mathcal{R}}^*\big(U,w;-\mathrm{D}E(U),-\mathrm{D}e(w)\big)\Big)\,\mathrm{d}t$$

and it is tempting to drop the term $\varepsilon\dot w$ and minimize the integrand for each $t\in[0,T]$ with respect to the w. However, we will see that this approach is not correct because we have to find the correct *non-equilibrium steady states* which create a nontrivial flux as a limit of $\varepsilon\dot w$.

We follow the approach in [LM*17] and estimate $\overline{\mathcal{R}}$ from below via

$$\overline{\mathcal{R}}(U,w;\dot U,\varepsilon\,\dot w) \geq \Big\langle \begin{pmatrix}\Xi\\ \zeta\end{pmatrix}, \begin{pmatrix}\dot U\\ \varepsilon\,\dot w\end{pmatrix}\Big\rangle - \overline{\mathcal{R}}^*(U,w;\Xi,\zeta),$$

where $(\Xi,\zeta):[0,T]\to X_{\mathrm{slow}}^*\times X_{\mathrm{fast}}^*$ are smooth test functions. Thus, we have

$$\mathfrak{D}_\varepsilon(U,w) \geq \int_0^T \Big(\Big\langle \begin{pmatrix}\Xi\\ \zeta\end{pmatrix}, \begin{pmatrix}\dot U\\ \varepsilon\,\dot w\end{pmatrix}\Big\rangle - \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta)\Big)\,\mathrm{d}t$$

with $\mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta)=\overline{\mathcal{R}}^*(U,w;\Xi,\zeta)-\overline{\mathcal{R}}^*\big(U,w;-\mathrm{D}E(U),-\mathrm{D}e(w)\big)$.

It turns our that we can now pass to the limit $\varepsilon \to 0$ by omitting the term $\varepsilon\,\dot w$. Then, we can then maximize with respect to ζ and minimize with respect to w for each $t \in [0,T]$. Hence, in terms of the Lagrangian $\mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}$ we are lead to the following sup-inf problem:

$$\mathscr{L}_{\rm red}(U,\Xi) := \sup_{w\in X_{\rm fast}} \inf_{\zeta\in X^*_{\rm fast}} \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta). \tag{1.3}$$

We say that the reduced Lagrangian $\mathscr{L}_{\rm red}$ has a *duality structure* if there exists $\mathcal{R}_{\rm eff} : X_{\rm slow}{\times}X_{\rm slow} \to [0,\infty]$ such that it can be written as

$$\mathscr{L}_{\rm eff}(U,\Xi) = \mathcal{R}_{\rm eff}(U,\Xi) - \mathcal{R}^*_{\rm eff}\big(U,-{\rm D}E(U)\big). \tag{1.4}$$

Using the EDP backwards, we see that the effective GS $(X_{\rm slow}, E, \mathcal{R}_{\rm eff})$ with the gradient-flow equation

$$\dot U = \partial_\Xi \mathcal{R}^*_{\rm eff}\big(U,-{\rm D}\mathcal{E}(U)\big)$$

indeed describes the limiting dynamics.

Thus, the main point in applying this theory successfully is to show the existence of the duality structure $(E,\mathcal{R}_{\rm eff})$ for the reduced Lagrangian $\mathscr{L}_{\rm red}$. And it is here were the theory of NESS comes into play. The definition of NESS in the above context means that we fix $\overline U \in X_{\rm slow}$ and want to find the NESS $\overline w \in X_{\rm fast}$ such that

$$\begin{pmatrix} 0\\ 0\end{pmatrix} = \begin{pmatrix} {\rm D}_\Xi \overline{\mathcal{R}}^*\big(U,w;-{\rm D}E(U),-{\rm D}e(w)\big) \\ {\rm D}_\zeta \overline{\mathcal{R}}^*\big(U,w;-{\rm D}E(U),-{\rm D}e(w)\big)\end{pmatrix} + \begin{pmatrix} \overline V\\ 0\end{pmatrix}, \quad U=\overline U \in X_{\rm slow}, \quad \overline V \in X_{\rm slow}. \tag{1.5}$$

We refer to (2.10) for the general case involving port mappings $P : X \to Y$ and $P^\circ : X^* \to Y^*$, which in (1.5) take the simple form $Y = X_{\rm slow}$, $P(U,w)=U$, and $P^\circ(\Xi,\zeta)=\Xi$. We observe that fixing $U=\overline U$ artificially generates a flux $\overline V$ which is generated by the NESS $\overline w$ associated with $\overline U$.

The first major link between the theory of NESS and the above saddle-point reduction for Lagrangians is the fact that NESS $\overline w$ solving (1.5) give rise to a *global null-saddle* $(w,\zeta) = \big(\overline w, -{\rm D}e(\overline w)\big)$ for $\mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}\big(\overline U,\cdot;-{\rm D}E(\overline U),\cdot\big)$, i.e.

$$\begin{aligned} \forall\,(w,\zeta)\in X_{\rm fast}{\times}X^*_{\rm fast}:\quad & \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}\big(\overline U,w;-{\rm D}E(\overline U),-{\rm D}e(\overline w)\big)\\ &\le \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}\big(\overline U,\overline w;-{\rm D}E(\overline U),-{\rm D}e(\overline w)\big) = 0 \le \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}\big(\overline U,\overline w;-{\rm D}E(\overline U),\zeta\big). \end{aligned}$$

Proposition 2.7 provides conditions under which null-saddles automatically have the form $\big(\overline w,-{\rm D}e(\overline w)\big)$, where $\overline w$ is NESS solving (1.5). There seem to exist a number of different variational characterizations (also called extremum principles) of NESS, but to the best of the author's knowledge the saddle-point formulation given here is new. We refer to [ASGB95, StW98, DD*12], [DeM84, Cha. V], and [Tsc00, Ch. 30, pp. 213-215]. In particular, [StW98] has the appealing title "*Maximum of the Local Entropy Production Becomes Minimal in Stationary Processes*".

The second important link arises from the fact that the existence of null-saddles implies $\mathscr{L}_{\rm red}\big(\overline U,-{\rm D}\mathcal{E}(\overline U)\big) = 0$ with $\mathscr{L}_{\rm red}$ from (1.3). However, Proposition 2.14 shows that this condition (for all $\overline U \in X_{\rm slow}$) is exactly the crucial condition for the existence of a duality structure in the sense of (1.4). Theorem 2.15 provides the main result giving the explicit construction of $\mathcal{R}_{\rm red}$, which involves a nontrivial series of duality arguments. In particular, the convexity of $\zeta \mapsto \mathscr{L}_{\rm red}(w,\zeta)$ needs a special argument.

Section 3 gives a more detailed account of the reduction of slow-fast gradient systems as discussed above. On particular, Section 3.2 it also treats the case where the slow component $U \in X_{\rm slow}$ and the fast component $w \in X_{\text{fast}}$ only interact by a constraint $P_{\text{slow}}U = P_{\text{fast}}w$, where $P_{\text{slow}} : X_{\text{slow}} \to Y$ and $P_{\text{fast}} : X_{\text{fast}} \to Y$ are the port mappings. In that case the effective dual dissipation potential is the sum

$$\mathcal{R}^*_{\text{eff}}(U,\Xi) = \mathcal{R}^*_{\text{slow}}(U,\Xi) + \mathsf{R}^*_Y\big(P_{\text{slow}}U, P^\circ_{\text{slow}}\Xi\big),$$

i.e. R_Y encodes all the information on the NESS in X_{fast}.

Section 4 provides two ODE examples, the first being that of a general quadratic dissipation potential and quadratic energies E and e. Everything can be explicitly calculated such that this case is helpful to obtain guidance when the abstract theory may be overwhelming. The second example treat a reaction-rate equation for four species A, B, C, and D undergoing two binary reaction pairs $A+B \rightleftharpoons D$ and $A+D \rightleftharpoons C$. Starting with constant reaction coefficients $\kappa_{1,2}$ for the two reaction and assuming that the vector of equilibrium densities is $(a_*, b_*, c_*, d_\varepsilon)$ with $d_\varepsilon = \varepsilon w_\varepsilon$ the transformation $d(t) = \varepsilon w(t)$ provides exactly a slow-fast GS as above, where the energies E and e are relative Boltzmann entropies and $\overline{\mathcal{R}}^*$ is of cosh-type. Applying the above reduction method via NESS we find an effective GS of cosh-type for the density vector (a,b,c) that corresponds to the ternary reaction pair $2A{+}B \rightleftharpoons C$. The interesting point is that, in contrast to the result in [MPS21], the cosh-type gradient structure is preserved, but now the effective reaction coefficient depends o the density a.

Section 5 revisits the results obtained in [LM*17] but now from a more general perspective. Moreover, the results are generalized by allowing for a reaction term which models sorption into and desorption from the background. The model starts from a one-dimensional diffusion on an interval, where the diffusion coefficient in the central membrane region $]{-\varepsilon},\varepsilon[$ is scaled by ε. Using Otto's gradient structure (see [Ott96, Ott98, JKO98, Ott01]) we start again from relative Boltzmann entropies E and e and from quadratic dual dissipation potentials $\overline{\mathcal{R}}^*(U,w;\,\cdot\,,\,\cdot\,)$. In the limit $\varepsilon \to 0$ the membrane part collapses to an interface generating transmission conditions. Our methods shows that R_Y is of cosh-type, which shows that it has inherited properties from the Boltzmann entropy e. Indeed, the Boltzmann function $\lambda_{\mathrm{B}}(z) = z\log z - z + 1$ with $\lambda'_{\mathrm{B}}(z) = \log z$ generates by the saddle-point problem the cosh-type function $\mathsf{C}^*(\zeta) = 4\cosh(\zeta/2) - 4$. Theorem 5.1 contains a much shorter derivation of $\mathcal{R}_{\text{eff}}$ than in [LM*17, PeS22], and Theorem 5.2 generalizes the result to the case including a reaction term that scales like $1/\varepsilon$ in the membrane region $]{-\varepsilon},\varepsilon[$.

Finally, Appendix A provides classical result on global saddle points as discussed in [EkT74]. For the readers convenience, we include a full proof for the existence of saddle points for convex-concave Lagrangians.

2 Constrained saddle points

We first collect some basic facts about unconstrained saddle points, then introduce the notion of constrained saddle points using the port mappings $P : X \to Y$ and $P^\circ : X^* \to Y^*$ and show that under a suitable additional condition that these constrained saddle points are indeed NESS solving (1.5) if they are null-saddles. Section 2.3 shows a further characterization if NESS as null-minimizers. Section 2.4 provides the main result concerning the duality structure for reduced Lagrangians $\mathscr{L}_{\text{red}}$ if the associated NESS are null-saddles.

2.1 Classical saddle points

For a gradient system $(X, \mathcal{E}, \mathcal{R})$ we consider the Lagrange functional

$$\mathscr{L}_{\mathcal{E},\mathcal{R}}(u,\xi) = \mathcal{R}^*(u,\xi) - \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)), \tag{2.1}$$

which is defined on $X \times X^*$. It will be the source of a series of results concerning NESS. We will simply write $\mathscr{L}$ in place of $\mathscr{L}_{\mathcal{E},\mathcal{R}}$ if the relevant GS $(X, \mathcal{E}, \mathcal{R})$ is clear.

Remark 2.1 (Slope dissipation term) *In the definition of $\mathscr{L}_{\mathcal{E},\mathcal{R}}$, we use the formula $\mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))$ to denote the so-called* slope dissipation*, which should properly be defined by its weak lower semicontinuous hull, namely*

$$\mathcal{S}(u) := \inf \Big\{ \liminf_{n\to\infty} \mathcal{R}^*(u_n, -\mathrm{D}\mathcal{E}(u_n)) \ \Big| \ u_n \rightharpoonup u, \ u_n \in \mathrm{dom}(\mathrm{D}\mathcal{E}) \Big\}. \tag{2.2}$$

For example the linear diffusion equation $\dot u = \Delta u$ with no-flux boundary conditions is the gradient-flow equation associated with the Otto gradient system $(\mathrm{P}(\Omega), \mathcal{E}_{\mathrm{Bz}}, \mathcal{R}^*_{\text{Otto}})$ with $\mathcal{E}_{\mathrm{Bz}}(u) = \int_\Omega \lambda_{\mathrm{B}}(u)\,\mathrm{d}x$, and $\mathcal{R}^*_{\text{Otto}}(u,\xi) = \int_\Omega \frac12 |\nabla\xi|^2 u\,\mathrm{d}x$. We obtain the Fisher information $\mathcal{S}(u) = \int_\Omega 2|\nabla\sqrt{u}|^2\,\mathrm{d}x$, which is well defined even when $u=0$ in a set of positive measure, whereas $u \in \mathrm{dom}(\mathrm{D}\mathcal{E})$ needs $u>0$ a.e.

Subsequently, we will still write $\mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))$ to emphasize the structure of the problem, but whenever analysis is done, we replace this term by $\mathcal{S}$.

Obviously, for all $u \in X$ the functions $\mathscr{L}(u,\cdot) : X^* \to \mathbb{R}$ are convex, and in some cases we have concavity of $\mathscr{L}(\cdot,\xi)$ for all $\xi \in X^*$. In the case of quadratic energy $\mathcal{E}(u) = \frac12\langle \mathbb{A}u, u\rangle - \langle \ell, u\rangle$ and a quadratic dual dissipation potential $\mathcal{R}^*(u,\xi) = \frac12\langle \xi, \mathbb{K}\xi\rangle$ we obtain the simple quadratic Lagrange functional

$$\mathscr{L}_{\mathrm{quadr}}(u,\xi) = \frac12 \langle \xi, \mathbb{K}\xi\rangle - \frac12 \big\langle \mathbb{A}u{-}\ell, \mathbb{K}(\mathbb{A}u{-}\ell)\big\rangle. \tag{2.3}$$

which has the above-mentioned concave-convex property on $X \times X^*$.

Definition 2.2 (Global saddle points) *Given two Banach spaces X and Y and a functional $\mathscr{L} : X \times Y \to \mathbb{R}$, we call a point $(\overline{x}, \overline{y}) \in X \times Y$ a* (global) saddle point *for $\mathscr{L}$ if*

$$\forall\, x \in X,\ y \in Y: \quad \mathscr{L}(x, \overline{y}) \le \mathscr{L}(\overline{x}, \overline{y}) \le \mathscr{L}(\overline{x}, y).$$

Thus, we are in the situation of classical saddle-point theory, see [EkT74] and Appendix A that collects the most important facts. In particular, we will use the the fact that the infimum over $\xi \in X$ and the supremum over $u \in X$ can be interchanged if a saddle point exists, see Lemma A.1:

$$\text{(a)} \qquad \mathrm{SI}_{\mathscr{L}} := \sup_{u\in X} \inf_{\xi \in X^*} \mathscr{L}(u,\xi) \ \le\ \inf_{\xi\in X^*} \sup_{u\in X} \mathscr{L}(u,\xi) := \mathrm{IS}_{\mathscr{L}} \tag{2.4a}$$

$$\text{(b)} \qquad \text{saddle point } (\overline{u}, \overline{\xi}) \text{ exists} \implies \mathrm{SI}_{\mathscr{L}} = \mathrm{IS}_{\mathscr{L}} = \mathscr{L}(\overline{u}, \overline{\xi}). \tag{2.4b}$$

For $\mathscr{L}_{\mathrm{quadr}}$ in (2.3) with invertible $\mathbb{A}$ we see that $(\overline{u}, \overline{\xi})$ is a saddle point if and only if $\overline{\xi} = 0$ (use $\mathbb{K} > 0$) and $\mathrm{D}\mathcal{E}(\overline{u}) = \mathbb{A}\overline{u}{-}\ell = 0$, viz. $\overline{u} = \mathbb{A}^{-1}\ell$. We then have $\mathscr{L}(\overline{u}, 0) = 0$. If $\mathbb{A}$ is not invertible, we have multiple saddle points, namely all $\overline{u}$ minimizing $u \mapsto \frac12\big\langle \mathbb{A}u{-}\ell, \mathbb{K}(\mathbb{A}u{-}\ell)\big\rangle$. Then, one has $\mathscr{L}(\overline{u}, 0) = -\min\big\{ \frac12\langle \mathbb{A}u{-}\ell, \mathbb{K}(\mathbb{A}u{-}\ell)\rangle \ \big|\ u \in X \big\}$.

As a second example we consider

$$X = \mathbb{R}^{i_*}, \quad \mathcal{E}(u) = \frac{1}{2}\langle \mathbb{A}u, u\rangle, \quad \text{and } \mathcal{R}(v) = \sum_{i=1}^{i_*} \big(\sigma|v_i| + \frac{\nu}{2}|v_i|^2\big) = \sigma|v|_1 + \frac{\nu}{2}|v|_2^2. \tag{2.5}$$

Now we have $\mathcal{R}^*(\xi) = \sum_{i_1}^{i_*} \frac{1}{2\nu}\big(\max\{|\xi_i|-\sigma, 0\}\big)^2$, which means $\mathcal{R}^*(u,\xi) = 0$ for $|\xi|_\infty \leq \sigma$. Hence, we have many saddle points $(\overline{u}, \overline{\xi})$, namely all pairs with $|\overline{\xi}|_\infty \leq \sigma$ and $|\mathbb{A}\overline{u}|_\infty \leq \sigma$. Again all saddle points satisfy $\mathscr{L}(\overline{u}, \overline{\xi}) = 0$.

A general characterization is the following. We also complement the result with a discussion of critical points $(\widetilde{u}, \widetilde{\xi})$ of $\mathscr{L}$ is they happen to be of the form $\widetilde{\xi} = -\mathrm{D}\mathcal{E}(\widetilde{u})$.

Theorem 2.3 (Unconstrained saddle points) *Consider a GS $(X, \mathcal{E}, \mathcal{R})$ and set $\mathscr{L} = \mathscr{L}_{\mathcal{E},\mathcal{R}}$ as in (2.1).*

(a) A pair $(\overline{u}, \overline{\xi})$ is a (global) saddle point of $\mathscr{L}$ if and only if

$$\mathcal{R}^*(\overline{u}, \overline{\xi}) = 0 \quad \textit{and} \quad \mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})) = \min_{u\in X} \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)).$$

(b) If there exists $u_\mathrm{eq} \in X$ with $\mathrm{D}\mathcal{E}(u_\mathrm{eq}) = 0$, then all saddle points satisfy $\mathcal{R}^*\big(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})\big) = 0$, and hence $\mathscr{L}(\overline{u}, \overline{\xi}) = 0$.

(c) If in addition to the condition in (b), the dual dissipation potentials $\mathcal{R}^*(u, \cdot) : X^* \to \mathbb{R}$ are strictly convex, then all saddle points $(\overline{u}, \overline{\xi})$ satisfy $\overline{\xi} = 0$ and $\mathrm{D}\mathcal{E}(\overline{u}) = 0$.

Proof. *Part (a).* Minimizing $\mathscr{L}$ with respect to $\xi \in X^*$ and using use $0 = \mathcal{R}^*(u, 0) \leq \mathcal{R}^*(u, \xi)$ yields

$$\mathrm{SI}_\mathscr{L} = \sup_{u\in X} \big(-\mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))\big) =: \overline{S} \leq 0.$$

Moreover, choosing $\xi = 0$ we obtain an upper bound for $\mathrm{IS}_\mathscr{L}$, namely $\mathrm{IS}_\mathscr{L} \leq \overline{S}$. Thus, with (2.4a) we conclude $\mathrm{SI}_\mathscr{L} = \mathrm{IS}_\mathscr{L} = \overline{S}$.

Hence, we conclude that a saddle point $(\overline{u}, ol\xi)$ must satisfy $\mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})) = -\overline{S}$ and $\mathcal{R}^*(\overline{u}, \overline{\xi}) = 0$, which is the desired result (a).

Part (b). We obtain $\overline{S} = 0$ and the result follows.

Part (c). This is an immediate consequence of the implication $\mathcal{R}^*(u, \xi) = 0 \;\Rightarrow\; \xi = 0$ and of Part (b). ■

The following result will not be used in the sequel, but it gives a first insight why the saddle-point theory for $\mathscr{L}_{\mathcal{E},\mathcal{R}}$ is useful. The point is that there is a certain redundancy in the Euler-Lagrange equation for critical points $(\widetilde{u}, \widetilde{\xi})$ of $\mathscr{L}_{\mathcal{E},\mathcal{R}}$, when the critical point satisfies $\widetilde{\xi} = -\mathrm{D}\mathcal{E}(\widetilde{u})$.

Lemma 2.4 (Euler-Lagrange equations if $\widetilde{\xi} = -\mathrm{D}\mathcal{E}(\widetilde{u})$) *The $(\widetilde{u}, \widetilde{\xi}) = (\widetilde{u}, -\mathrm{D}\mathcal{E}(\widetilde{u})) \in X\times X^*$ is a critical point of $\mathscr{L} = \mathscr{L}_{\mathcal{E},\mathcal{R}}$ if and only if $\partial_\xi \mathcal{R}^*(\widetilde{u}, -\mathrm{D}\mathcal{E}(\widetilde{u})) = 0 \in X$, i.e. $\widetilde{u}$ is a steady state for the gradient system $(X, \mathcal{E}, \mathcal{R})$.*

Proof. We have $\mathrm{D}_\xi \mathscr{L}(u,\xi)[\widetilde{\xi}] = \mathrm{D}_\xi \mathcal{R}^*(u,\xi)[\widetilde{\xi}]$ and

$$\mathrm{D}_u \mathscr{L}(u,\xi)[\widetilde{u}] = \mathrm{D}_u \mathcal{R}^*(u,\xi)[\widetilde{u}] - \mathrm{D}_u \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))[\widetilde{u}] + \mathrm{D}_\xi \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u))\big[\mathrm{D}^2\mathcal{E}(u)[\widetilde{u}]\big].$$

Inserting $(u,\xi)=(\widetilde{u}-\mathrm{D}\mathcal{E}(\widetilde{u})$ we see a cancellation and the two equations for a critical point reduce to

$$0=\mathrm{D}_\xi\mathscr{L}(u_*,\xi_*)[\widetilde{\xi}]=\mathrm{D}_\xi\mathcal{R}^*(u_*,\xi_*)[\widetilde{\xi}],\quad 0=\mathrm{D}_u\mathscr{L}(u_*,\xi_*)[\widetilde{u}]=\mathrm{D}_\xi\mathcal{R}^*(u_*,\xi_*)\big[\mathrm{D}^2\mathcal{E}(u_*)[\widetilde{u}]\big].$$

Thus, we see that it is necessary and sufficient to satisfy $\partial_\xi\mathcal{R}^*(u_*,\xi_*)=0$. ■

Remark 2.5 (NESS in perturbed gradient systems) *If a the gradient-flow equation of a GS $(X,\mathcal{E},\mathcal{E})$ is perturbed by a general vector field V in the form*

$$\dot{u}=V(u)+\partial_\xi\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)), \tag{2.6}$$

then steady states can still be obtained as stationary points of a Lagrangian $\widetilde{\mathscr{L}}$, namely

$$\widetilde{\mathscr{L}}(u,\xi)=\mathcal{R}^*(u,\xi)-\langle\xi,V(u)\rangle-\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))-\langle\mathrm{D}\mathcal{E}(u),V(u)\rangle. \tag{2.7}$$

Assume that u_* is a steady state for (2.6)*, namely*

$$0=V(u_*)+\partial_\xi\mathcal{R}^*(u_*,-\mathrm{D}\mathcal{E}(u_*)), \tag{2.8}$$

then $(u,\xi)=(u_*,-\mathrm{D}\mathcal{E}(u_*))$ is a stationary point for $\widetilde{\mathscr{L}}$ and obviously the critical value is 0, i.e. $\widetilde{\mathscr{L}}(u_*,-\mathrm{D}\mathcal{E}(u_*))=0$.

To see the stationarity we observe $\mathrm{D}_\xi\widetilde{\mathscr{L}}(u,\xi)=\mathrm{D}_\xi\mathcal{R}^*(u,\xi)-V(q)$, and (2.8) *yields $\mathrm{D}_\xi\widetilde{\mathscr{L}}(u_*,-\mathrm{D}\mathcal{E}(u_*))=0$ as desired. For the derivative with respect to u we have*

$$\begin{aligned}\mathrm{D}_u\widetilde{\mathscr{L}}(u,\xi)[w]&=\mathrm{D}_u\mathcal{R}^*(u,\xi)[w]-\langle\xi,\mathrm{D}V(u)[w]\rangle-\mathrm{D}_u\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))[w]\\&\quad-\mathrm{D}_\xi\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))[-\mathrm{D}^2\mathcal{E}(u)[w,\cdot]\rangle-\mathrm{D}^2\mathcal{E}(u)[w,V(u)]-\mathrm{D}\mathcal{E}(u)\big[\mathrm{D}V(u)[w]\big].\end{aligned}$$

Inserting $\xi=-\mathrm{D}\mathcal{E}(q)$ the first term cancels the third, and the second term cancels the last. Moreover, the forth and the fifth terms cancel if we additionally use (2.8)*. Hence, $\mathrm{D}_u\widetilde{\mathscr{L}}(u_*,-\mathrm{D}\mathcal{E}(u_*))=0$, and $(u_*,-\mathrm{D}\mathcal{E}(u_*))$ is indeed a stationary point for $\widetilde{\mathscr{L}}$.*

2.2 Constrained saddle points

Hence, we now study the constrained case, where the port mappings $P:X\to Y$ and $P^\circ:X^*\to Y^*$ are used to drive the GS $(X,\mathcal{E},\mathcal{R})$. We start by introducing a constrained saddle-point problem and then relate the existence of constrained saddle points to the existence of NESS.

Problem 2.6 (Constrained saddle-point problem (CSPP)) *Given the GS $(X,\mathcal{E},\mathcal{R})$ with Lagrangian $\mathscr{L}_{\mathcal{E},\mathcal{R}}$ and the port mapping $P^\circ:X^*\to Y^*$, the* constrained saddle-point problem *for $\eta\in Y^*$ consists in finding a saddle point $(\overline{u}_\eta,\overline{\xi}_\eta)\in X{\times}X^*$ for*

$$\begin{aligned}&\forall\ u\in X \text{ with } P^\circ\mathrm{D}\mathcal{E}(u)=\eta\ \ \forall\ \xi\in X^* \text{ with } P^\circ\xi=-\eta:\\&\qquad\qquad\mathscr{L}_{\mathcal{E},\mathcal{R}}(u,\overline{\xi}_\eta)\leq\mathscr{L}_{\mathcal{E},\mathcal{R}}(\overline{u}_\eta,\overline{\xi}_\eta)\leq\mathscr{L}_{\mathcal{E},\mathcal{R}}(\overline{u}_\eta,\xi).\end{aligned} \tag{2.9}$$

The saddle point $(\overline{u}_\eta,\overline{\xi}_\eta)\in X{\times}X^*$ is called a null-saddle *if $\mathscr{L}_{\mathcal{E},\mathcal{R}}(\overline{u}_\eta,\overline{\xi}_\eta)=0$.*

If for some $\eta\neq0$ we find a saddle-point $(u_\eta,-\mathrm{D}\mathcal{E}(u_\eta))$, then u_η is called a Non-Equilibrium Steady State (NESS) *corresponding to the constraint $\eta\in Y^*$.*

In light of our theory, it will be important to conclude that a constrained saddle point is actually a NESS. Under natural assumption this can be concluded for null saddle points.

Proposition 2.7 (Null-saddles and NESS) *If a constrained saddle point* $(\overline{u}_\eta, \overline{\xi}_\eta) \in X{\times}X^*$ *is a NESS, then it is a null-saddle.*

If $\mathcal{R}^*(\overline{u}_\eta, \cdot) : X^*$ *is strictly convex and* $(\overline{u}_\eta, \overline{\xi}_\eta) \in X{\times}X^*$ *is a null-saddle, then it is a NESS, i.e.* $\overline{\xi} = -\mathrm{D}\mathcal{E}(\overline{u})$.

Proof. The first statement follows directly from $\mathscr{L}(u, -\mathrm{D}\mathcal{E}(u)) = 0$ for all $u \in X$.

For the opposite implication we start from a general null-saddle $(\overline{u}, \overline{\xi})$. From $0 = \mathscr{L}(\overline{u}, \overline{\xi}) \leq \mathscr{L}(\overline{u}, \xi)$ for all ξ with $P^\circ\xi = \eta$ we see that $\xi = \overline{\xi}$ and $\xi = -\mathrm{D}\mathcal{E}(\overline{u})$ are global minimizers. By strict convexity the minimizer is unique, which proves the assertion. ■

We recall the example in (2.5) where $\mathcal{R}^*$ is not strictly convex, because of $\mathcal{R}^*(\xi) = 0$ for $|\xi|_\infty \leq \sigma$. The saddle points $(\overline{u}, \overline{\xi})$ are characterized by $|\mathbb{A}\overline{u}|_\infty \leq \sigma$ and $|\overline{\xi}|_\infty \leq \sigma$ and all of them are null-saddles. However, only the ones satisfying additionally $\overline{\xi} = -\mathbb{A}\overline{u}$ are NESS. This shows that the result does not hold without a further condition like our strict convexity.

The next result shows that a NESS obtained from as a constrained saddle point satisfy the desired Euler-Lagrange equation (1.5), where port mapping P° features twice, namely first as constraint on the state and secondly to insert the Lagrange multiplier $v \in Y$, which denotes the necessary fluxes to support the NESS $\overline{u}$ induced by the constraint $P^\circ\mathrm{D}\mathcal{E}(u) = -\eta$.

Proposition 2.8 (Euler-Lagrange equations for NESS) *If the saddle point in the CSPP* (2.6) *has the form* $(\overline{u}, \overline{\xi}) = (\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u}))$*, then the corresponding Euler-Lagrange equations for NESS reads*

$$0 = \mathrm{D}_\xi\mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})) - P^{\circ *}\overline{v}, \quad P^\circ\mathrm{D}\mathcal{E}(\overline{u}) = -\eta \in Y^*, \quad \overline{v} \in Y. \tag{2.10}$$

Proof. In (2.9) we may consider variations $\widehat{\xi}$ and $\widehat{u}$ with $P^\circ\widehat{\xi} = 0$ and $P^\circ\mathrm{D}^2\mathcal{E}(u)[\widehat{u}] = 0$. Thus, we obtain

$$0 = \mathrm{D}_\xi\mathscr{L}(u,\xi)[\widehat{\xi}] = \mathrm{D}_\xi\mathcal{R}^*(u,\xi)[\widehat{\xi}] = \langle\widehat{\xi}, \mathrm{D}_\xi\mathcal{R}^*(u,\xi)\rangle_X,$$
$$0 = \mathrm{D}_u\mathscr{L}(u,\xi)[\widehat{u}] = \mathrm{D}_u\mathcal{R}^*(u,\xi)[\widehat{u}] - \mathrm{D}_u\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))[\widehat{u}] + \mathrm{D}_\xi\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\big[\mathrm{D}^2\mathcal{E}(u)[\widehat{u}]\big].$$

Inserting $\overline{\xi} = -\mathrm{D}\mathcal{E}(\overline{u})$ we obtain a cancellation in the second line leading to

$$0 = \langle\widehat{\xi}, \mathrm{D}_\xi\mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u}))\rangle_X \quad \text{and} \quad 0 = \langle\mathrm{D}^2\mathcal{E}(\overline{u})[\widehat{u}], \mathrm{D}_\xi\mathcal{R}^*(\overline{u}, \overline{\xi})\rangle_X.$$

However, by the choice of admissible variations, we see that the second relation follows from the first. Hence we have $\mathrm{D}_\xi\mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})) \in \big(\mathrm{ker}(P^\circ)\big)^\perp$.

To conclude, we simply use Fredholm's alternative (theorem):

$$\big(\mathrm{ker}(P^\circ)\big)^\perp := \big\{\, x \in X \,\big|\, P^\circ\xi = 0 \ \Rightarrow\ \langle\xi, x\rangle_X = 0 \,\big\} \ = \ \mathrm{ran}(P^{\circ *}) := \big\{\, P^{\circ *}y \,\big|\, y \in Y \,\big\}.$$

With this we have $\mathrm{D}_\xi\mathcal{R}^*(\overline{u}, -\mathrm{D}\mathcal{E}(\overline{u})) \in \big\{\, P^{\circ *}y \,\big|\, y \in Y \,\big\}$, which gives $\overline{y} \in Y$ such that (2.10) holds. ■

Under the assumption that for all $\eta \in Y^*$ there exists a unique NESS $\overline{u}_\eta$ of (2.10) with Lagrange parameter $v = \overline{v}_\eta$, we can define the *port relation*

$$\mathfrak{P} : Y^* \to Y;\ \eta \mapsto \overline{v}_\eta.$$

It is this port relation which will play a crucial role in the sequel. As a first result we observe, that in the case that $\mathcal{R}$ is independent of the state, the port relation can be obtained easily from $\mathcal{R}$, it is independent of the energy $\mathcal{E}$, and it is given as the differential of an effective potential $\mathsf{R}_{\mathfrak{P}}$. We refer to Section 4.1 to a simple and explicit case.

Proposition 2.9 (Port relation for state-independent dissipation) *If $\mathcal{R} : X \to [0, \infty]$ is a state-independent dissipation potential, then the port relation $\mathfrak{P}$ is given by*

$$v = \mathfrak{P}(\eta) = \partial \mathsf{R}_Y^*(v) \quad \textit{with}\ \ \mathsf{R}_Y(v) = \mathcal{R}(P^{\circ *}v).$$

Equivalently, R_Y^* is characterized via $\mathsf{R}_Y^*(\eta) := \inf_{\xi:\, P^\circ \xi = \eta} \mathcal{R}^*(\xi)$.

Proof. By Fenchel's equivalence we have $\xi \in \partial\Psi(v) \iff v \in \partial\Psi^*(\xi)$. Hence, the NESS equation (2.10) can be rewritten as

$$\mathrm{D}\mathcal{R}\big(P^{\circ *}v\big) = -\mathrm{D}\mathcal{E}(\overline{u}), \quad P^\circ \mathrm{D}\mathcal{E}(\overline{u}) = -\eta, \quad v \in Y,$$

where we used that $\mathcal{R}^*$, and hence also $\mathcal{R}$, are independent of u. Hence, we have the relation $\eta = P^\circ \mathrm{D}\mathcal{R}\big(P^{\circ *}v\big) = \mathrm{D}_v \mathsf{R}_Y(v)$, which is independent of u. Applying Fenchel's equivalence once again we obtain the assertion $v = \mathrm{D}_\eta \mathsf{R}_Y^*(\eta)$.

The second characterization of R_Y^* follows by an application of Lemma 2.12. ■

Looking into the proof of the above theorem, we can see that the inverse port relation $\mathfrak{P}^{-1} : Y \to Y^*$ in the state-dependent case has the more general form

$$\eta = P^\circ \mathrm{D}\mathcal{R}\big(\overline{u}(v); v\big),$$

where $\overline{u}(v)$ is the NESS associated with the flux $v \in Y$. It is surprising that also in such cases one can show that $\mathfrak{P}^{-1}(v) = \mathrm{D}\mathcal{R}_{\mathrm{red}}(v)$ for a reduced dissipation potential $\mathcal{R}_{\mathrm{red}}$ that is now depending on $\mathcal{R}$ and $\mathcal{E}$, see Section 3.

We now provide a general existence result for constrained saddle points and for NESS. For this we use the following major assumptions on $\mathscr{L} = \mathscr{L}_{\mathcal{E},\mathcal{R}}$ and $B : X^* \to Z$.

$$\left.\begin{aligned} \forall\, u \in X,\ \xi \in X^* :\ \ \mathscr{L}(u,\cdot) : X^* \to \mathbb{R}\ \text{ and }\ -\mathscr{L}(\cdot,\xi) : X \to \mathbb{R}\ \text{ are} \\ \text{lower semi-continuous, strictly convex, and coercive;} \end{aligned}\right\} \tag{2.11a}$$

$$\forall\, \eta \in Y^* : \quad \big\{\, u \in X \,\big|\, P^\circ \mathrm{D}\mathcal{E}(u) = \eta \,\big\} \text{ is closed and convex.} \tag{2.11b}$$

Theorem 2.10 (Existence of constrained saddle points) *Assume that $(X, \mathcal{E}, \mathcal{R})$ and $P^\circ : X^* \to Y^*$ satisfy* (2.11)*. Then, for each $\eta \in Y^*$ there exists a unique constrained saddle point $(\overline{u}_\eta, \overline{\xi}_\eta)$ for $\mathscr{L}$ (in the sense of* (2.9)*).*

If additionally the mapping $X \ni u \mapsto \mathrm{D}\mathcal{E}(u) \in X^*$ is surjective, then these saddles points are NESS satisfying $\overline{\xi}_\eta = -\mathrm{D}\mathcal{E}(\overline{u}_\eta)$ and (2.10)*.*

Proof. The existence follows by applying Proposition A.2 with $\boldsymbol{U} = \big\{\, \xi \in X^* \,\big|\, P^\circ\xi = -\eta \,\big\}$ and $\widetilde{\boldsymbol{V}} = \big\{\, u \in X \,\big|\, P^\circ \mathrm{D}\mathcal{E}(u) = \eta \,\big\}$, where we extend $\mathscr{L}$ by $-\infty$ outside of $\widetilde{\boldsymbol{V}}$ if it is not a linear space. Thus, we find a unique constrained saddle point $(\overline{u}_\eta, \overline{\xi}_\eta)$ with $P^\circ \mathrm{D}\mathcal{E}(\overline{u}_\eta) = \eta$ and $P^\circ \overline{\xi}_\eta = -\eta$.

Using Proposition 2.7 it is sufficient to show that $(\overline{u}_\eta, \overline{\xi}_\eta)$ is a null-saddle. Because we already have a saddle point, it is sufficient to show $\mathrm{SI}_{\mathscr{L}} \leq 0 \leq \mathrm{IS}_{\mathscr{L}}$.

For the lower estimate we simply use $\inf_{\xi\in\boldsymbol{U}} \mathscr{L}(u,\xi) \leq \mathscr{L}(u, -\mathrm{D}\mathcal{E}(u)) = 0$. Taking the supremum over $u \in \widetilde{\boldsymbol{V}}$ we find $SI_{\mathscr{L}} \leq 0$.

For the upper estimate we start from a general $\xi \in \boldsymbol{U}$ such that the surjectivity of $\mathrm{D}\mathcal{E}$ provides a $u_\xi \in \boldsymbol{V}$ with $\xi = -\mathrm{D}\mathcal{E}(u_\xi)$. With this we have $\sup_{u\in\boldsymbol{V}} \mathscr{L}(u,\xi) \geq \mathscr{L}(u_\xi, \xi) = 0$. Now taking the infimum over $\xi \in \boldsymbol{U}$ yields $\mathrm{IS}_{\mathscr{L}} \geq 0$ as desired. ■

2.3 NESS as minimizers

The main observation of the last section is that the equation (2.10) does not have a simple variational structure. Its characterization via the above saddle-point theory provides some kind of variational structure, but needs a doubling of variables. Moreover, in nonlinear problems (non-quadratic $\mathscr{L}$) the saddle-point theory for solving infinite-dimensional problem like PDEs is technically very demanding.

The naive way of treating the CSPP (2.9) would be to minimize first with respect to ξ providing $\xi = \Xi_B(z,u)$ and such that it remains to study the minimization problem

$$u \;\mapsto\; \mathcal{R}(u, -\mathrm{D}\mathcal{E}(u)) - \mathcal{R}^*(u, \Xi_B(z,u)) \quad \text{subject to } B\mathrm{D}\mathcal{E}(u) = z.$$

This approach is doable but has the disadvantage that it is difficult to keep enough control on the mapping $u \mapsto \Xi_B(z,u)$ to tackle the final minimization problem.

The following result shows that the saddle point can be turned into a minimization problem by applying a suitable Legendre transformation with respect to the constrained variable ξ, but keeping a dual parameter $\Lambda \in Z^*$. Thus, the minimization formulation stays explicit in terms of the constituents of the GS $(X, \mathcal{E}, \mathcal{R})$. Moreover, it is more directly related to the Euler-Lagrange equations (2.10).

Proposition 2.11 (NESS as minimizers) *For all $\eta \in Y^*$ any global minimizer $(\overline{u}, \overline{y}) \in X \times Y$ of the constrained minimization problem*

$$\begin{array}{rl} \textit{minimize} & \mathcal{R}(u, P^{\circ *}y) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) + \langle \eta, y\rangle_Y \\ \textit{over} & (u,y) \in X\times Y \quad \textit{subject to} \;\; P^\circ \mathrm{D}\mathcal{E}(u) = \eta \end{array} \tag{2.12}$$

gives rise to a constrained saddle points $(\overline{u}, \overline{\xi}) \in X \times X^*$ for (2.9) *where we can choose any $\overline{\xi} = \mathrm{Argmin}\big\{\, \mathcal{R}^*(\overline{u}, \xi) \,\big|\, P^\circ\xi = -\eta \,\big\}$. Vice versa, if $(\overline{u}, \overline{\xi})$ is a constrained saddle point for* (2.9)*, then $(\overline{u}, \overline{y})$ with $\overline{y} \in \mathrm{Argmax}\big\{\, \langle \eta, y\rangle - \mathcal{R}(\overline{u}, P^{\circ *}y) \,\big|\, y \in Y \,\big\}$ is a global minimizer for* (2.12)*.*

Moreover, if $(\overline{u}, \overline{y})$ is a null-minimizer, then $(\overline{y}, \overline{\xi})$ is a null-saddle, and under the additional assumption of strict convexity of $\mathcal{R}^*(\overline{u}, \cdot)$ it defines a NESS solving (2.10)*.*

Proof. We define the auxiliary dissipation potentials $\Psi_u : Y \to \mathbb{R}_\infty;\; y \mapsto \mathcal{R}(u, P^{\circ *}y)$ and can now apply Lemma 2.12 below. This gives

$$\inf_{\xi: P^\circ\xi = -\eta} \mathcal{R}^*(u,\xi) = \Psi_u^*(-z) = \sup_{y\in Y} \big(-\langle \eta, y\rangle_Y - \mathcal{R}(u, P^{\circ *}y) \big). \tag{2.13}$$

DOI 10.20347/WIAS.PREPRINT.2998

With this we obtain the following chain of identities:

$$\begin{aligned}\sup_{\substack{u\in X\\ P^\circ \mathrm{D}\mathcal{E}(u)=\eta}} \inf_{\substack{\xi\in X^*\\ P^\circ\xi=-\eta}} \mathscr{L}_{\mathcal{E},\mathcal{R}}(u,\xi) &= \sup_{\substack{u\in X\\ P^\circ \mathrm{D}\mathcal{E}(u)=\eta}} \Big(\Big[\inf_{\substack{\xi\in X^*\\ P^\circ\xi=-\eta}} \mathcal{R}^*(u,\xi)\Big] - \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\Big)\\ &\overset{(2.13)}{=} \sup_{\substack{u\in X\\ P^\circ \mathrm{D}\mathcal{E}(u)=\eta}} \Big(\Big[\sup_{y\in Y}\big(-\langle \eta,y\rangle_Y - \mathcal{R}(u,P^{\circ*}y)\big)\Big] - \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\Big)\\ &= -\inf_{\substack{u: P^\circ \mathrm{D}\mathcal{E}(u)=\eta\\ y\in Y}} \Big(\mathcal{R}(u,P^{\circ*}y) + \langle \eta,y\rangle_Y + \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\Big).\end{aligned}$$

This shows that the minimization problem (2.12) is equivalent to the CSSP (2.9) if we choose $\xi=\overline{\xi}\in X^*$ in (2.13) optimally, i.e. $\overline{\xi}=\mathrm{Argmin}\big\{\,\mathcal{R}^*(\overline{u},\xi)\,\big|\,P^\circ\xi=-\eta\,\big\}$.

Moreover, the values are the same up to a minus sign. Hence, null-minimizers $(\overline{u},\overline{y})\in X{\times}Y$ correspond to null-saddles $(\overline{u},\overline{\xi})\in X{\times}X^*$, and the remaining statement follows from Proposition 2.7. ■

In the above proof the relation in (2.13) relies on the following result.

Lemma 2.12 *For a lower semi-continuous and convex* $\Psi: X\to\mathbb{R}_\infty$ *and linear bounded operator* $B: X^*\to Z$ *we have*

$$\inf_{\xi\in X^*:\, B\xi=z} \Psi^*(\xi) = \sup_{\Lambda\in Z^*}\Big(\langle\Lambda,z\rangle_Z - \Psi(B^*\Lambda)\Big).$$

Proof. Consider a dissipation potential $\Psi: X\to[0,\infty]$ and a bounded linear mapping $A: Y\to X$ and define the dissipation potential $\widetilde{\Psi}: Y\to[0,\infty];\ y\mapsto\Psi(Ay)$. In [MaM20, Prop. 6.1] the identity $\big(\widetilde{\Psi}\big)^*(\eta)=\inf\big\{\,\Psi^*(\xi)\,\big|\,A^*\xi=\eta\,\big\}$ is established. Applying this with $Y=Z^*$ and $A=B^*: Z^*\to X$ the assertion follows. ■

2.4 Constrained Lagrangians, duality structure, and NESS

When doing reduction or Γ-limits of Lagrangians, we may end up with a general function $\mathcal{K}: Y{\times}Y^*\to\mathbb{R}$ and may then ask the question whether this function can be written as a Lagrangian $\mathscr{L}_{\mathsf{E},\mathsf{R}}$.

Definition 2.13 (Duality structure) *We say that a function* $\mathcal{K}: Y{\times}Y^*\to\mathbb{R}$ *has the* duality structure (E,R)*, if* $(Y,\mathsf{E},\mathsf{R})$ *is a gradient system and*

$$\mathcal{K}=\mathscr{L}_{\mathsf{E},\mathsf{R}},\quad \textit{namely}\ \forall\,(y,\eta)\in Y{\times}Y^*:\ \mathcal{K}(y,\eta)=\mathsf{R}^*(y,\eta)-\mathsf{R}^*(y,-\mathrm{D}\mathsf{E}(y)).$$

We observe that for a given $\mathcal{K}$ the dissipation functional R and its dual R^* are uniquely determined by $\mathsf{R}^*(y,\eta)=\mathcal{K}(y,\eta)-\mathcal{K}(y,0)$. Hence, we have the following necessary and sufficient conditions of a duality structure.

Proposition 2.14 (Conditions for duality structure) *Given an energy* $\mathsf{E}: Y\to\mathbb{R}$*, the function* $\mathcal{K}: Y{\times}Y^*\to\mathbb{R}$ *has a duality structure* (E,R) *if and only if*

$$\forall\,(y,\eta)\in Y{\times}Y^*:\quad \mathcal{K}(y,\eta)\geq\mathcal{K}(y,0), \tag{2.14a}$$

$$\forall\,y\in Y:\quad \mathcal{K}(y,\cdot): Y^*\to\mathbb{R}\ \textit{ is convex}, \tag{2.14b}$$

$$\forall\,y\in Y:\quad \mathcal{K}(y,-\mathrm{D}\mathsf{E}(y))=0. \tag{2.14c}$$

Then, R *is given by* $\mathsf{R}^*(y,\eta)=\mathcal{K}(y,\eta)-\mathcal{K}(y,0)$.

Proof. It is obvious that $\mathcal{K}$ satisfies (2.14) if it has the duality structure (E,R).

To show the opposite, we observe that $\mathsf{R}^*_{\mathcal{K}} : (y,\eta) \mapsto \mathcal{K}(y,\eta) - \mathcal{K}(y,0)$ is a dual dissipation potential because of (2.14a) and (2.14b). Inserting the formula for $\mathsf{R}^*_{\mathcal{K}}$ into the condition $0 = \mathcal{K}(y,\eta) - \mathsf{R}^*(y,\eta) + \mathsf{R}^*(y,-\mathrm{D}\mathsf{E}(y))$ for the duality structure, we obtain

$$\begin{aligned}0 &= \mathcal{K}(y,\eta) - \mathsf{R}^*_{\mathcal{K}}(y,\eta) + \mathsf{R}^*_{\mathcal{K}}(y,-\mathrm{D}\mathsf{E}(y))\\ &= \mathcal{K}(y,\eta) - \big(\mathcal{K}(y,\eta)-\mathcal{K}(y,0)\big) + \big(\mathcal{K}(y,-\mathrm{D}\mathsf{E}(y)) - \mathcal{K}(y,0)\big) = \mathcal{K}(y,-\mathrm{D}\mathsf{E}(y)).\end{aligned}$$

Hence, (2.14c) guarantees that this $(\mathsf{E},\mathsf{R}_{\mathcal{K}})$ is the desired duality structure. ■

We return to our constrained saddle point problems by generalizing it in a crucial way. For this we use the second port function $P : X \to Y$ which allows us to impose direct conditions $Pu = y$ on the state variable, whereas $P^\circ\mathrm{D}\mathcal{E}(u) = \eta$ does this indirectly. Nevertheless, we always assume there is an energy $\mathsf{E} : Y \to \mathbb{R}$, such that

$$Pu = y \quad \Longrightarrow \quad P^\circ\mathrm{D}\mathcal{E}(u) = \mathrm{D}\mathsf{E}(y). \tag{2.15}$$

An important point for understanding of the induced kinetic relation generated by the gradient system $(X,\mathcal{E},\mathcal{R})$ with port (P,P°) is to study the reduced Lagrangian $\mathscr{L}_{\mathrm{red}} : Y{\times}Y^* \to \mathbb{R}$ defined via

$$\mathscr{L}_{\mathrm{red}}(y,\eta) := \sup_{\substack{u\in X\\ Pu=y}} \ \inf_{\substack{\xi\in X^*\\ P^\circ\xi=\eta}} \mathscr{L}_{\mathcal{E},\mathcal{R}}(u,\xi). \tag{2.16}$$

In contrast to the previous analysis, we are now using to independent constraints $y \in Y$ and $\eta \in Y^*$, whereas in Section 2.2 we always assumed the compatibility $\eta = -\mathrm{D}\mathsf{E}(y)$, cf. (2.15). However, assuming there are null-saddles under these constraints means that $\mathscr{L}_{\mathrm{red}}(y,-\mathrm{D}\mathsf{E}(y)) = 0$ holds, i.e. the necessary (2.14c) holds. Hence, the major part of the following proof goes into showing that $\eta \mapsto \mathscr{L}_{\mathrm{red}}(y,\eta)$ is convex and attains its minimum value at $\eta = 0$. The convexity in η is nontrivial, because convexity is preserved by taking suprema (over u with $Pu = y$) but not by taking infima (over ξ with $P^\circ\xi = \eta$).

Theorem 2.15 (Duality structure for $\mathscr{L}_{\mathrm{red}}$) *Consider a gradient system $(X,\mathcal{E},\mathcal{R})$ with port mappings $P : X \to Y$ and $P^\circ : X^* \to Y^*$ and a compatible energy E as in* (2.15)*. Assume that for all $y \in Y$ the CSSP* (2.9) *with $\eta = -\mathrm{D}\mathsf{E}(y) = -P^\circ\mathrm{D}\mathcal{E}(u)$ has a null-saddle. Then, the reduced Lagrangian $\mathscr{L}_{\mathrm{red}}$ defined in* (2.16) *has the duality structure (R,E) with $\mathsf{R}^*(y,\eta) = \mathscr{L}_{\mathrm{red}}(y,\eta) - \mathscr{L}_{\mathrm{red}}(y,0)$, namely*

$$\mathscr{L}_{\mathrm{red}}(y,\eta) = \mathsf{R}^*(y,\eta) - \mathsf{R}^*\big(y,-\mathrm{D}\mathsf{E}(y)\big). \tag{2.17}$$

Proof. The proof relies on the following auxiliary functionals:

$$\begin{aligned}
&\mathcal{M}^* : X{\times}Y^* \to \mathbb{R}_\infty, && \mathcal{M}^*(u,\eta) := \inf_{\substack{\xi\in X^*\\ P^\circ\xi=\eta}} \Big(\mathcal{R}^*(u,\xi) - \mathcal{R}^*\big(u,-\mathrm{D}\mathcal{E}(u)\big)\Big),\\
&\mathcal{M} : X{\times}Y \to \mathbb{R}_\infty, && \mathcal{M}(u,v) := \sup_{\eta\in Y^*} \Big(\langle\eta,v\rangle - \mathcal{M}^*(u,\eta)\Big)\\
&\mathcal{N} : Y{\times}Y \to \mathbb{R}_\infty, && \mathcal{N}(y,v) := \inf_{\substack{u\in X\\ Pu=y}} \mathcal{M}(u,v).
\end{aligned}$$

Clearly, we have $\mathscr{L}_{\mathrm{red}}(y,\eta) = \sup\big\{\, \mathcal{M}^*(u,\eta) \,\big|\, Pu = y \,\big\}$.

Step (a): $\mathcal{M}(u,v) = \mathcal{R}^*\big(u,-\mathrm{D}\mathcal{E}(u)\big) + \mathcal{R}(u,P^{\circ*}v)$.

To show this, we simply use the definitions and obtain

$$\begin{aligned}
\mathcal{M}(u,v) &= \sup_{\eta\in Y^*}\big(\langle\eta,v\rangle - \mathcal{M}^*(u,\eta)\big)\\
&= \sup_{\eta\in Y^*}\Big(\langle\eta,v\rangle - \inf_{\xi:\,P^\circ\xi=\eta}\big(\mathcal{R}^*(u,\xi){-}\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u))\big)\Big)\\
&= \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) \;+\; \sup_{\substack{\eta\in Y^*\\ \xi:\,P^\circ\xi=\eta}}\big(\langle\eta,v\rangle - \mathcal{R}^*(u,\xi)\big)\\
&= \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) + \sup_{\xi\in X^*}\big(\langle P^\circ\xi,v\rangle - \mathcal{R}^*(u,\xi)\big) \;=\; \mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) + \mathcal{R}(u,P^{\circ*}v).
\end{aligned}$$

Thus, the desired result of part (a) is established.

Step (b): Defining $\overline{\mathsf{R}}(y,v) = \mathcal{N}(y,v) - \mathcal{N}(y,0)$ and $\mathsf{R}(y,\cdot) = \big(\overline{\mathsf{R}}(y,\cdot)\big)^{**}$ we have to show

$$\mathcal{N}(y,v) = \overline{\mathsf{R}}(y,v) + \mathsf{R}^*\big(y,-\mathrm{DE}(y)\big). \tag{2.18}$$

By the definition of $\overline{\mathsf{R}}$, it it is sufficient to derive the identity $\mathcal{N}(y,0) = \mathsf{R}^*\big(y,-\mathrm{DE}(y)\big)$. Since R is the convexification of $\overline{\mathsf{R}}$ with respect to v for fixed y, we can use the identity $\mathsf{R}^*(y,\cdot) = \big(\overline{\mathsf{R}}(y,\cdot)\big)^*$. Using the abbreviation $\widehat{\eta} := -\mathrm{DE}(y) \in Y^*$ we find

$$\begin{aligned}
&\mathsf{R}^*\big(y,\widehat{\eta}\big) - \mathcal{N}(y,0) = \sup_{v\in Y}\Big(\langle\widehat{\eta},v\rangle - \underbrace{\big(\mathcal{N}(y,v){-}\mathcal{N}(y,0)\big)}_{=\overline{\mathsf{R}}(y,v)}\Big) \;-\mathcal{N}(y,0)\\
&= \sup_{v\in Y}\big(\langle\widehat{\eta},v\rangle - \mathcal{N}(y,v)\big) = \sup_{v\in Y}\Big(\langle\widehat{\eta},v\rangle - \inf_{u:\,Pu=y}\mathcal{M}(u,v)\Big)\\
&\overset{(a)}{=} \sup_{v\in Y}\Big(\langle\widehat{\eta},v\rangle - \inf_{u:\,Pu=y}\big(\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) + \mathcal{R}(u,P^{\circ*}v)\big)\Big)\\
&= \sup_{u:\,Pu=y}\Big(-\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) + \sup_{v\in Y}\big(\langle\widehat{\eta},v\rangle - \mathcal{R}(u,P^{\circ*}v)\big)\Big)\\
&\overset{\text{Lem. 2.12}}{=} \sup_{u:\,Pu=y}\Big(-\mathcal{R}^*(u,-\mathrm{D}\mathcal{E}(u)) + \inf_{\xi:\,P^\circ\xi=\widehat{\eta}}\mathcal{R}^*(u,\xi)\Big) = \sup_{Pu=y}\inf_{P^\circ\xi=\widehat{\eta}}\mathscr{L}_{\mathcal{E},\mathcal{R}}(u,\xi) = 0.
\end{aligned}$$

For the last identity we used $\widehat{\eta} = -\mathrm{DE}(y)$ and that we arrived exactly at the desired CSSP for $\mathscr{L}$ with compatible constraints, which has a null-saddle by assumption. Hence, (2.18) is established.

Step (c): We establish (2.17) by simple manipulations:

$$\begin{aligned}
\mathscr{L}_{\mathrm{red}}(y,\eta) &= \sup_{u:\,Pu=y}\mathcal{M}^*(u,\eta) = \sup_{u:\,Pu=y}\Big(\sup_{v\in Y}\big(\langle\eta,v\rangle - \mathcal{M}(u,v)\big)\Big)\\
&= \sup_{v\in Y}\big(\langle\eta,v\rangle - \inf_{u:\,Pu=y}\mathcal{M}(u,v)\big) = \sup_{v\in Y}\big(\langle\eta,v\rangle - \mathcal{N}(y,v)\big)\\
&\overset{(b)}{=} \sup_{y\in Y}\Big(\langle\eta,y\rangle - \overline{\mathsf{R}}(y,v) - \mathsf{R}^*\big(y,-\mathrm{DE}(y)\big)\Big) \;=\; \mathsf{R}^*(y,\eta) - \mathsf{R}^*\big(y,-\mathrm{DE}(y)\big),
\end{aligned}$$

which is the desired result.

Step (d): It remains to show that R^* is a dissipation potential. From $\mathsf{R}^*(y,\cdot) = \big(\overline{\mathsf{R}}(y,\cdot)\big)^*$ we see that $\mathsf{R}^*(y,\cdot)$ is lower semi-continuous and convex.

The formula for $\mathcal{M}$ in (a) shows $\mathcal{M}(u,v) \geq \mathcal{M}(u,0)$. Hence, taking the infimum over u satisfying $Pu = y$, we have $\mathcal{N}(y,v) \geq \mathcal{N}(y,0)$, i.e. $\overline{\mathsf{R}}(y,v) \geq 0$. By definition of $\overline{\mathsf{R}}$ we also have $\overline{\mathsf{R}}(y,0) = 0$, which allows us to conclude $\mathsf{R}^*(y,0) = 0$ and $\mathsf{R}^*(y,\eta) \geq 0$. Hence, $\mathsf{R}^*(y,\cdot) : Y^* \to [0,\infty]$ is a dual dissipation potential. ∎

3 EDP-convergence for slow-fast GSs via NESS

We consider a family of gradient systems $(X, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ where $\varepsilon > 0$ is the small parameter modeling the ratio between fast and slow relaxation times.

We consider two cases distinguished cases: in the first the state space can be decomposed in the form $u = (U, w) \in X_{\rm slow}{\times}X_{\rm fast} = X$ and in the second we have

$$X = \big\{\, u = (U, w) \in X_{\rm slow}{\times}X_{\rm fast} \,\big|\, \mathbb{P}(U, w) := P_{\rm slow}U - P_{\rm fast}w = 0 \,\big\}$$

where $P_{\rm slow} : X_{\rm slow} \to Y$ and $P_{\rm fast} : X_{\rm fast} \to Y$ are suitable port mappings. Here we consider $U \in X_{\rm slow}$ as the slow macroscopic part of the state variables, while $w \in X_{\rm fast}$ is the fast microscopic part, that one wants to eliminate in the limit $\varepsilon \to 0$.

In both setting we assume that the scaling in ε is very particular, but nevertheless we are able to treat a number of prototypical cases. In particular, we assume $\mathcal{E}_\varepsilon(U, w) = E(U) + \varepsilon\, e(w)$.

3.1 Case 1: product space $X = X_{\rm fast}{\times}X_{\rm slow}$

The precise assumptions on the scaling with $\varepsilon > 0$ are the following:

$$\mathcal{E}_\varepsilon(U, w) = E(U) + \varepsilon\, e(w) \qquad \text{additive split of energy,} \tag{3.1a}$$
$$\mathcal{R}^*_\varepsilon(U, w; \Xi, \mu) = \overline{\mathcal{R}}^*\big(U, w; \Xi, \tfrac{1}{\varepsilon}\mu\big) \qquad \text{fast relaxation of } w, \tag{3.1b}$$

where $\overline{\mathcal{R}}^* : X{\times}X^* \to [0, \infty]$ is a general dual dissipation potential.

The associated gradient-flow equation takes a simple form, because the appearance of ε is chosen in a particular way.

$$\dot U = {\rm D}_\Xi \overline{\mathcal{R}}^*\big(U, w; -{\rm D}E(U), -{\rm D}e(w)\big), \tag{3.2a}$$
$$\varepsilon \dot w = {\rm D}_\mu \overline{\mathcal{R}}^*\big(U, w; -{\rm D}E(U), -{\rm D}e(w)\big). \tag{3.2b}$$

Thus, on the formal level, we can drop the term $\varepsilon\dot w$, because w relaxes into a NESS on the time scale ε which is much faster than the evolution of U which happens on time scales of order 1. The microscopic variable w moves into the NESS $w = \widehat w(U)$ satisfying

$$0 = {\rm D}_\mu \overline{\mathcal{R}}^*\big(U, w; -{\rm D}E(U), -{\rm D}e(w)\big). \tag{3.3}$$

Note that $\mu \mapsto \overline{\mathcal{R}}(U, w; \Xi, \mu)$ is not a dual dissipation potential, but after doing a linear correction we see that $\Psi^*(U, w; \Xi, \cdot) : X_{\rm fast} \to [0, \infty]$ defined via

$$\Psi^*(U, w; \Xi, \mu) := \overline{\mathcal{R}}^*(U, w; \Xi, \mu) - \overline{\mathcal{R}}^*(U, w; \Xi, 0) - \big\langle \mu, {\rm D}_\mu \overline{\mathcal{R}}^*(U, w; \Xi, \mu)\big\rangle$$

is a dual dissipation potential. Rewriting (3.3) in terms of Ψ^* we obtain

$$0 = {\rm D}_\mu \Psi^*\big(U, w; -{\rm D}E(U), -{\rm D}e(w)\big) + \overline{\mathcal{R}}^*((U, w; -{\rm D}E(U), 0),$$

which is indeed an equation for a NESS in the sense on (2.10).

Inserting the limiting relation $w = \widehat w(U)$ into the first equation of (3.2a) we obtain the reduced macroscopic problem

$$\dot U = {\rm D}_\Xi \overline{\mathcal{R}}^*\big(U, \widehat w(U); -{\rm D}E(U), -{\rm D}e(\widehat w(U))\big). \tag{3.4}$$

The disadvantage of the above approach is that we lose control over the gradient structures. As we have started with the GSs $(X, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$, it is natural to ask whether the effective equation (3.4) has a natural gradient structure inherited from E, e, and $\overline{\mathcal{R}}$.

This question can be answered by the notion of EDP-convergence, which provides a tool to stay on the level of gradient systems. We follow here the approach developed in [LM*17] which forms the basis of the further developments of EDP-convergence in [DFM19, MMP21]. The abbreviation "EDP" stand for the *energy-dissipation principle* (cf. [Mie16, Thm. 3.3.1]) that shows that under suitable technical assumptions a curve $u_\varepsilon = (U_\varepsilon, w_\varepsilon) : [0,T] \to X$ is a solution of the gradient-flow equation (3.2) if and only if it satisfies the energy-dissipation inequality

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

The idea in [LM*17, MaM20] is to replace the primal dissipation $\mathcal{R}_\varepsilon(u, \dot u)$ by the lower bound $\langle \xi, \dot u\rangle - \mathcal{R}^*_\varepsilon(u, \xi)$ for an arbitrary test function $\xi : [0,T] \to X^*$. Then, the limit $\varepsilon \to 0$ is performed and finally one maximizes with respect to ξ to recover the limiting energy-dissipation balance again.

Thus, for a general smooth function $\xi : [0,T] \to X^*$ we have

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \Big(\big\langle \xi, \dot u_\varepsilon \big\rangle - \mathcal{R}^*_\varepsilon\big(u_\varepsilon; \xi\big) + \mathcal{R}^*_\varepsilon\big(u_\varepsilon; -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon)\big) \Big) \, \mathrm{d}t \le \mathcal{E}_\varepsilon(u_\varepsilon(0)).$$

Using the explicit ε-dependence of $\mathcal{E}_\varepsilon$ and $\mathcal{R}^*_\varepsilon$ imposed in (3.1) and choosing $\xi = (\Xi, \varepsilon\zeta)$ we arrive at

$$\mathcal{E}_\varepsilon(u_\varepsilon(T)) + \int_0^T \Big(\big\langle (\Xi, \varepsilon\zeta), \dot u_\varepsilon \big\rangle - \overline{\mathcal{R}}^*\big(u_\varepsilon; \Xi, \zeta\big) + \overline{\mathcal{R}}^*\big(u_\varepsilon; -\mathrm{D}E(U_\varepsilon), -\mathrm{D}e(w_\varepsilon)\big) \Big) \, \mathrm{d}t \le \mathcal{E}_\varepsilon(u_\varepsilon(0)).$$

Now passing to the limit $\varepsilon \to 0$ the term $\langle \varepsilon\zeta, \dot w_\varepsilon\rangle$ and the terms $\varepsilon e(w_\varepsilon(t))$ vanish. Assuming $(U_\varepsilon, w_\varepsilon) \to (U, w)$ we are left with the inequality

$$E(U(T)) + \int_0^T \Big(\big\langle \Xi, \dot U \big\rangle - \mathscr{L}_{\overline{\mathcal{E}}, \overline{\mathcal{R}}}(U, w; \Xi, \zeta) \Big) \, \mathrm{d}t \le E(U(0)) \ \text{ for all } (\Xi, \zeta) \in \mathrm{L}^\infty([0,T]; X^*),$$

where $\overline{\mathcal{E}}(U, w) = E(U) + e(w)$ and hence

$$\mathscr{L}_{\overline{\mathcal{E}}, \overline{\mathcal{R}}}(U, w; \Xi, \zeta) = \overline{\mathcal{R}}^*\big(U, w; \Xi, \zeta\big) - \overline{\mathcal{R}}^*\big(U, w; -\mathrm{D}E(U), -\mathrm{D}e(w)\big).$$

Since w appear in the integral only via $w(t)$, but not with a derivative $\dot w(t)$ we can eliminate $w(t)$ by a pointwise infimum. Similar, we can eliminate ζ by a pointwise supremum. Hence, defining $\mathscr{L}_{\mathrm{red}} : X_{\mathrm{slow}} \times X^*_{\mathrm{slow}} \to \mathbb{R}$ via

$$\mathscr{L}_{\mathrm{red}}(U, \Xi) := \sup_{w \in X_{\mathrm{fast}}} \inf_{\zeta \in X^*_{\mathrm{fast}}} \mathscr{L}_{\overline{\mathcal{E}}, \overline{\mathcal{R}}}(U, w; \Xi, \zeta). \tag{3.5}$$

we obtain the inequality

$$E(U(T)) + \int_0^T \Big(\big\langle \Xi, \dot U \big\rangle - \mathscr{L}_{\mathrm{red}}(U, \Xi) \Big) \, \mathrm{d}t \le E(U(0)). \tag{3.6}$$

Now it remains to show that $\mathscr{L}_{\mathrm{red}}$ has a duality structure $(E, \mathcal{R}_\ell)$ in the sense of Definition 2.13, i.e. it has the form

$$\mathscr{L}_{\mathrm{red}}(U, \Xi) = \mathcal{R}^*_{\mathrm{eff}}(U; \Xi) - \mathcal{R}^*_{\mathrm{eff}}(U; -\mathrm{D}E(U)), \quad \text{i.e. } \mathscr{L}_{\mathrm{red}} = \mathscr{L}_{E, \mathcal{R}_{\mathrm{eff}}} \tag{3.7}$$

for a suitable effective dissipation potential $\mathcal{R}_{\text{eff}}$.

If this is the case, we can insert this into (3.6) and reverse the Legendre transform to obtain the energy-dissipation inequality

$$E(U(T)) + \int_0^T \Big(\mathcal{R}_{\text{eff}}(U;\dot U) + \mathcal{R}^*_{\text{eff}}\big(U;-\mathrm{D}E(U)\big)\Big)\,\mathrm{d}t \leq E(U(0)). \tag{3.8}$$

Applying the energy-dissipation principle once again, we see that U is a solution of the gradient-flow equation

$$\dot U = \mathrm{D}\mathcal{R}^*_{\text{eff}}(U,-\mathrm{D}E(U))$$

for the reduced gradient system $(X_{\text{slow}}, E, \mathcal{R}_{\text{eff}})$. Clearly, this equation must equal (3.4), but now we have a much cleaner structure.

To achieve this goal it remains to establish the duality structure (3.7). The following result is the analogue of Theorem 2.15.

Theorem 3.1 ($\mathscr{L}_{\text{eff}}$ has duality structure) *For a GS $(X_{\text{slow}}{\times}X_{\text{fast}}, \overline{\mathcal{E}}, \overline{\mathcal{R}})$ with $\overline{\mathcal{E}} = E{\otimes}e$ define $\mathscr{L}_{\text{red}} : X_{\text{slow}}{\times}X^*_{\text{slow}} \to \mathbb{R}$ as in* (3.5)*. If for all $U \in X_{\text{slow}}$ we have that*

$$\mathscr{L}_{\text{eff}}(U,-\mathrm{D}E(U)) := \sup_{w\in X_{\text{fast}}} \inf_{X^*_{\text{fast}}} \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;-\mathrm{D}E(U),\zeta)$$

is a null-saddle (i.e. $\mathscr{L}_{\text{red}}\big(U,-\mathrm{D}E(U)\big) = 0$), then $\mathscr{L}_{\text{red}}$ has the duality structure $(E,\mathcal{R}_{\text{eff}})$ where $\mathcal{R}_{\text{eff}}$ is given via $\mathcal{R}^*_{\text{eff}}(U,\Xi) = \mathscr{L}_{\text{red}}(U,\Xi) - \mathscr{L}_{\text{red}}(U,0)$.

Proof. The result follows directly from Theorem 2.15 if we use the port mappings

$$P(U,w) = U \in X_{\text{slow}} \quad \text{and} \quad P^\circ(\Xi,\zeta) = \Xi \in X^*_{\text{slow}}.$$

Note that $\overline{\mathcal{E}} = E{\otimes}e$ satisfies $\mathrm{D}\overline{\mathcal{E}}(U,w) = \big(\mathrm{D}E(U),\mathrm{D}e(w)\big)$, hence, $E : X_{\text{slow}} \to \mathbb{R}$ is a compatible energy in the sense of (2.15). ■

3.2 Case 2: factored product space $X = (X_{\text{fast}}{\times}X_{\text{slow}})\big/_{\ker \mathbb{P}}$

In some cases it is not easy to decompose the state space X into a product $X_{\text{slow}}{\times}X_{\text{fast}}$, but it is possible to decompose the state with some overlay or joint traces on an interface, namely

$$X = \big\{\, u = (U,w) \in X_{\text{slow}}{\times}X_{\text{fast}} \;\big|\; \mathbb{P}(U,w) := P_{\text{slow}}U - P_{\text{fast}}w = 0 \,\big\}$$

where $P_{\text{slow}} : X_{\text{slow}} \to Y$, $P_{\text{fast}} : X_{\text{fast}} \to Y$, $P^\circ_{\text{slow}} : X^*_{\text{slow}} \to Y^*$, and $P^\circ_{\text{fast}} : X^*_{\text{fast}} \to Y^*$ are suitable port mappings. Below we will show that the chosen ansatz applies diffusion problems, where P_{slow} and P_{fast} are used to define traces from two different sides of an interface, see (5.3) in Section 5.1.

The precise assumptions are the following:

$$\mathcal{E}_\varepsilon(U,w) = E(U) + \varepsilon e(w) \qquad \text{additive split of energy,} \tag{3.9a}$$

$$\mathcal{R}^*_\varepsilon(U,w;\Xi,\xi) = \widetilde{\mathcal{R}}^*(U,w;\Xi,\tfrac{1}{\varepsilon}\xi\big) \qquad \text{fast relaxation of } w, \tag{3.9b}$$

$$\begin{aligned}\widetilde{\mathcal{R}}^*(U,w;\Xi,\zeta) = {}& \mathcal{R}^*_{\text{slow}}(U;\Xi) + \mathcal{R}^*_{\text{fast}}(w;\zeta)\\ &+ \boldsymbol{\delta}_{\{0\}}\big(P^\circ_{\text{fast}}\zeta - P^\circ_{\text{slow}}\Xi\big) \qquad \text{interaction through } Y^*.\end{aligned} \tag{3.9c}$$

In principle, we could allow the more general case $\mathcal{R}^*_{\rm fast}(U,w;\zeta)$ in place of $\mathcal{R}^*_{\rm fast}(w;\zeta)$, but refrain from doing so, because the restricted version better highlights the fact that the U and w can only interact through the ports via Y.

Here $\boldsymbol{\delta}_{\{0\}} : Y \to [0,\infty]$ is the convex function with $\boldsymbol{\delta}_{\{0\}}(0) = 0$ and ∞ otherwise. This function implements the constraint $P^*_{\rm fast}\zeta = P^*_{\rm slow}\Xi$ giving the interaction condition $P^*_{\rm fast}{\rm D}e(w) = P^*_{\rm slow}{\rm D}E(U)$. The subdifferential of $\boldsymbol{\delta}_{\{0\}}$ at $\eta = 0$ is given by $\partial\boldsymbol{\delta}_{\{0\}}(0) = Y$, i.e. the hard constraint can transmit the fluxes $(-P^{\circ*}_{\rm slow}v, P^{\circ*}_{\rm fast}v)$ for arbitrary $v \in Y$.

We first observe that the gradient-flow equation takes a simple form, because the appearance of ε is chosen in a particular way.

$$\begin{pmatrix}\dot U\\ \varepsilon\dot w\end{pmatrix} \in \partial\overline{\mathcal{R}}^*\big(U,w;-{\rm D}E(U),-{\rm D}e(w)\big) \iff$$
$$\begin{pmatrix}\dot U\\ \varepsilon\dot w\end{pmatrix} = \begin{pmatrix}{\rm D}_\Xi\mathcal{R}^*_{\rm slow}\big(U,-{\rm D}E(U)\big)\\ {\rm D}_\zeta\mathcal{R}^*_{\rm fast}\big(w,-{\rm D}e(w)\big)\end{pmatrix} + \begin{pmatrix}P^{\circ*}_{\rm slow}v\\ -P^{\circ*}_{\rm fast}v\end{pmatrix} \quad \text{with} \begin{cases} P_{\rm slow}U = P_{\rm fast}w \\ \text{and } v\in Y.\end{cases}$$

Thus, on the formal level, we can drop the term $\varepsilon\dot w$, because w relaxes into a NESS on the time scale ε which is much faster than the evolution of U which happens on time scales of order 1. The microscopic variable w moves along the family of NESS $w = \widehat{w}(U,v)$ generated by the flux $v \in Y$ from

$$0 = {\rm D}_\zeta\mathcal{R}^*_{\rm fast}\big(U,w;-{\rm D}e(w)\big) + P^{\circ*}_{\rm fast}v.$$

As in the previous subsection, we can now involve the energy-dissipation principle to show EDP-convergence, where now $\overline{\mathcal{R}}^*$ is replaced by $\widetilde{\mathcal{R}}^*$ containing the constraint $P^\circ_{\rm slow}\Xi = P^\circ_{\rm fast}\zeta$. We again arrive at the reduced energy inequality (3.6), where now $\mathscr{L}_{\rm slow}$ takes a special form because of the additive splitting of $\widetilde{\mathcal{R}}^*$ in (3.9c):

$$\begin{aligned}\mathscr{L}_{\rm eff}(U,\Xi) &= \sup_{w\in X_{\rm fast}}\ \inf_{\zeta\in X^*_{\rm fast}}\ \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta) = \mathscr{L}_{E,\mathcal{R}_{\rm slow}}(U,\Xi) + \mathscr{L}_{\rm red}(U,\Xi)\\ \text{with } \mathscr{L}_{\rm red}(U,\Xi) &:= \sup_{\substack{w\in X_{\rm fast}\\ P_{\rm fast}w = P_{\rm slow}U}}\ \inf_{\substack{\zeta\in X^*_{\rm fast}\\ P^\circ_{\rm fast}\zeta = P^\circ_{\rm slow}\Xi}}\ \mathscr{L}_{e,\mathcal{R}_{\rm fast}}(w,\zeta)\end{aligned} \tag{3.10}$$

Thus, we see that $\mathscr{L}_{\rm red}$ is exactly obtained as in Section 2.4. Hence, we know that $\mathscr{L}_{\rm red}$ has a duality structure if for all $\eta \in Y$ CSSP (2.9) for $\mathscr{L}_{e,\mathcal{R}_{\rm fast}}$ with constraint $P^\circ{\rm D}e(w) = \eta$ has a null-saddle. In that case we have the duality structure $(\mathsf{E}_Y, \mathsf{R}_Y)$ such that

$$\mathscr{L}_{\rm eff}(U,\Xi) = \mathscr{L}_{E,\mathcal{R}_{\rm eff}}(U,\Xi) \quad \text{with } \mathcal{R}_{\rm eff}(U,\Xi) = \mathsf{R}^*_Y\big(P_{\rm slow}U, P^\circ_{\rm slow}\Xi\big).$$

We see that $\mathscr{L}_{\rm red}$ depends on (U,Ξ) only through the port values $\big(P_{\rm slow}U, P^\circ_{\rm slow}\Xi\big) \in Y\times Y^*$. Returning to $\mathscr{L}_{\rm eff} = \mathscr{L}_{E,\mathcal{R}_{\rm slow}} + \mathscr{L}_{\rm red}$ we obtain

$$\mathscr{L}_{\rm eff} = \mathscr{L}_{E,\mathcal{R}_{\rm eff}} \quad \text{with } \mathcal{R}^*_{\rm eff}(U,\Xi) = \mathcal{R}_{\rm slow}(U,\Xi) + \mathsf{R}_Y\big(P_{\rm slow}U, P^\circ_{\rm slow}\Xi\big).$$

Moreover, we see that $(X_{\rm slow}, E, \mathcal{R}_{\rm eff})$ is the EDP limit of $(X, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ and the effective gradient-flow equation reads

$$\dot U = {\rm D}_\Xi\mathcal{R}^*_{\rm eff}\big(U,-{\rm D}E(U)\big) = {\rm D}_\Xi\mathcal{R}^*_{\rm slow}\big(U,-{\rm D}E(U)\big) + P^{\circ*}_{\rm slow}{\rm D}_\eta\mathsf{R}_Y\big(P_{\rm slow}U, P^\circ_{\rm slow}\Xi\big),$$

which clearly shows that the non-equilibrium flux is given by

$$P^{\circ*}_{\rm slow}v \quad \text{with } v = {\rm D}_\eta\mathsf{R}_Y\big(P_{\rm slow}U, P^\circ_{\rm slow}\Xi\big) \in Y.$$

4 EDP-convergence for two ODE examples

We first treat the linear case as given in (2.3) and with a suitable scaling in $\varepsilon > 0$. Secondly, we consider a nonlinear reaction systems with four species and two binary reactions $A+B \rightleftharpoons D$ and $A+D \rightleftharpoons C$ and show that the limiting system gives the single ternary reaction $2A+B \rightleftharpoons C$.

4.1 Simple quadratic energy and dissipation

On the Hilbert space $X = X_\text{slow}\times X_\text{fast}$ with $u=(U,w)$ we consider the family $(X,\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$ of GSs given by $\mathcal{E}_\varepsilon(U,w) = E(U)+\varepsilon e(w)$ with

$$E(U) = \frac12 \langle \mathbb{A}_\text{s} U - \mu_\text{s}, U\rangle_{X_\text{slow}} \text{ and } e(w) = \frac12 \langle \mathbb{A}_\text{f} w - \mu_\text{f}, w\rangle_{X_\text{fast}}$$

and

$$\mathcal{R}_\varepsilon(\Xi,\xi) = \frac12 \Big\langle \begin{pmatrix}\Xi\\ \frac1\varepsilon \xi\end{pmatrix}, \begin{pmatrix}\mathbb{K}_\text{ss} & \mathbb{K}_\text{sf}\\ \mathbb{K}_\text{fs} & \mathbb{K}_\text{ff}\end{pmatrix}\begin{pmatrix}\Xi\\ \frac1\varepsilon \xi\end{pmatrix}\Big\rangle = \overline{\mathcal{R}}^*\big(\Xi, \frac1\varepsilon \xi\big).$$

Hence, we have the situation treated in Section 3.1.

The linear gradient-flow equations and their limit for $\varepsilon\to 0$ take the form

$$\begin{pmatrix}\dot U_\varepsilon\\ \varepsilon \dot w_\varepsilon\end{pmatrix} = -\begin{pmatrix}\mathbb{K}_\text{ss} & \mathbb{K}_\text{sf}\\ \mathbb{K}_\text{fs} & \mathbb{K}_\text{ff}\end{pmatrix}\begin{pmatrix}\mathbb{A}_\text{s}U_\varepsilon - \mu_\text{s}\\ \mathbb{A}_\text{f} w_\varepsilon - \mu_\text{f}\end{pmatrix} \text{ and } \begin{pmatrix}\dot U\\ 0\end{pmatrix} = -\begin{pmatrix}\mathbb{K}_\text{ss} & \mathbb{K}_\text{sf}\\ \mathbb{K}_\text{fs} & \mathbb{K}_\text{ff}\end{pmatrix}\begin{pmatrix}\mathbb{A}_\text{s}U - \mu_\text{s}\\ \mathbb{A}_\text{f} w - \mu_\text{f}\end{pmatrix}.$$

With the port mappings $P(U,w) = U \in Y := X_\text{slow}$ and $P^\circ(\Xi,\zeta) \to \Xi \in X_\text{fast}$ we obtain the determining equation (2.10) for the NESS

$$\begin{pmatrix}0\\0\end{pmatrix} = -\begin{pmatrix}\mathbb{K}_\text{ss} & \mathbb{K}_\text{sf}\\ \mathbb{K}_\text{fs} & \mathbb{K}_\text{ff}\end{pmatrix}\begin{pmatrix}\mathrm{D}E(U)\\ \mathrm{D}e(w)\end{pmatrix} + \begin{pmatrix}V\\0\end{pmatrix}, \quad \mathrm{D}E(U) = -\Xi \in X_\text{slow}, \quad V\in X_\text{slow}.$$

As Ξ is given, and the upper equation is always true for a suitable V, we find the NESS

$$\mathbb{A}_\text{f} w - \mu_\text{f} = \mathrm{D}e(w) = \mathbb{K}_\text{ff}^{-1}\mathbb{K}_\text{fs}\Xi.$$

The resulting port mapping $\mathfrak{P}: X^*_\text{slow} \to X_\text{slow};\ \Xi \mapsto V$ takes the explicit form

$$V = \mathfrak{P}\Xi = \mathbb{K}_\text{eff}\Xi \quad \text{with } \mathbb{K}_\text{eff} = \mathbb{K}_\text{ss} - \mathbb{K}_\text{sf}\mathbb{K}_\text{ff}^{-1}\mathbb{K}_\text{fs}.$$

In particular, $\mathfrak{P}$ is independent of the energy $\mathcal{E}$, as predicted by Proposition 2.9.

We also want to show that $\mathfrak{P} = \mathrm{D}\mathcal{R}^*_\text{eff}$ can be obtained by the saddle-point reduction of the Lagrangian

$$\mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta) = \overline{\mathcal{R}}^*(\Xi,\zeta) - \overline{\mathcal{R}}^*\big(\mu_\text{s}-\mathbb{A}_\text{s}U, \mu_\text{f}-\mathbb{A}_\text{f}w\big).$$

Assuming that $\mathbb{K}>0$ and $\mathbb{A}_\text{s}>0$, a simple calculation gives

$$\begin{aligned}\mathscr{L}_\text{red}(U,\Xi) &= \sup_{w\in X_\text{fast}} \inf_{\zeta\in X^*_\text{fast}} \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta) = \inf_{\zeta\in X^*_\text{fast}} \overline{\mathcal{R}}^*(\Xi,\zeta) - \sup_{w\in X_\text{fast}} \overline{\mathcal{R}}^*\big(\mu_\text{s}-\mathbb{A}_\text{s}U, \mu_\text{f}-\mathbb{A}_\text{f}w\big)\\ &= \frac12\langle \Xi, \mathbb{K}_\text{eff}\Xi\rangle - \frac12 \langle \mu_\text{s}-\mathbb{A}_\text{s}U, \mathbb{K}_\text{eff}(\mu_\text{s}-\mathbb{A}_\text{s}U)\rangle = \mathscr{L}_{E,\mathcal{R}_\text{eff}}(U,\Xi)\end{aligned}$$

with $\mathcal{R}_\text{eff}(\Xi) = \frac12\langle\Xi,\mathbb{K}_\text{eff}\Xi\rangle$.

4.2 Two binary reaction generate one ternary reaction

We consider four chemical species A, B, C, and D with associated concentrations $a, b, c, d \in [0, \infty[$. The undergo the two binary reversible reaction pairs $A + B \rightleftharpoons D$ and $A + D \rightleftharpoons C$ according to the mass action law. We assume that species D is very unstable and either react fast with an A to create C or decay fast into A and B. In particular, the equilibrium concentrations for D will be $d_\varepsilon := \varepsilon w_*$, while the equilibrium densities a_*, b_*, c_* are positive and independent of ε.

The associated reaction rate equation is the ODE system

$$\begin{pmatrix} \dot a \\ \dot b \\ \dot c \\ \dot d \end{pmatrix} = \kappa_1 \Big(\frac{d}{d_\varepsilon} - \frac{ab}{a_* b_*} \Big) \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \kappa_2 \Big(\frac{c}{c_*} - \frac{ad}{a_* d_\varepsilon} \Big) \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \tag{4.1}$$

where κ_1 and κ_2 are positive reaction coefficients that may depend on a, b, c, d, but make them constant for simplicity.

As above one may replace d by εw and such that the right-hand side becomes independent of ε. Dropping the term εw on the left-hand side leads to the algebraic-differential system

$$\begin{pmatrix} \dot a \\ \dot b \\ \dot c \\ 0 \end{pmatrix} = \kappa_1 \Big(\frac{w}{w_*} - \frac{ab}{a_* b_*} \Big) \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \kappa_2 \Big(\frac{c}{c_*} - \frac{aw}{a_* w_*} \Big) \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \tag{4.2}$$

Solving the last equation for w and inserting the result into the first three equations leads to the reduced ODE

$$\begin{pmatrix} \dot a \\ \dot b \\ \dot c \end{pmatrix} = \kappa_{\text{eff}}(a) \Big(\frac{c}{c_*} - \frac{a^2 b}{a_*^2 b_*} \Big) \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \quad \text{with } \kappa_{\text{eff}}(a) := \frac{\kappa_1 \kappa_2 a_*}{\kappa_1 a_* + \kappa_2 a} \tag{4.3}$$

which is the reaction-rate equation for the ternary reaction $2A + B \rightleftharpoons C$ with an effective reaction coefficient $\kappa_{\text{eff}}(a) \in]0, \kappa_2[$.

The original system has the entropic cosh-gradient structure as derived in [MP*17] and further studied in [MiS20, MPS21]. In our specific case, the reaction-rate equation (4.1) is the gradient-flow equation for the GS $(\mathbb{R}^4, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ given by (where $u = (a, b, c, d)$)

$$\mathcal{E}_\varepsilon(u) = \lambda_{\text{B}}(a/a_*) a_* + \lambda_{\text{B}}(b/b_*) b_* + \lambda_{\text{B}}(c/c_*) c_* + \lambda_{\text{B}}(d/d_\varepsilon) d_\varepsilon \ \text{ and}$$
$$\mathcal{R}^*_\varepsilon(u;\xi) = \kappa_1 \Big(\frac{abd}{a_* b_* d_\varepsilon} \Big)^{1/2} \mathsf{C}^*\big(\xi_1{+}\xi_2{-}\xi_4\big) + \kappa_2 \Big(\frac{acd}{a_* c_* d_\varepsilon} \Big)^{1/2} \mathsf{C}^*\big(\xi_1{-}\xi_3{+}\xi_4\big),$$

where $\lambda_{\text{B}}(z) = z \log z - z + 1$ is the Boltzmann function and $\mathsf{C}^*(\zeta) = 4\cosh(\zeta/2) - 4$.

Doing our standard scaling for the slow and fast variables gives

$$u = (U, \varepsilon w), \quad U = (a, b, c) \in X_{\text{slow}}, \quad \mathcal{E}_\varepsilon(u) = E(U) + \varepsilon e(w) \text{ with } e(w) = \lambda_{\text{B}}(w/w_*) w_*.$$

Moreover, with $\Xi = (\xi_1, \xi_2, \xi_3) \in X^*_{\text{slow}}$ we have $\widetilde{\mathcal{R}}_\varepsilon(U, \varepsilon w; \Xi, \mu) = \overline{\mathcal{R}}\big(U, w; \Xi, \frac{1}{\varepsilon}\mu\big)$ with

$$\overline{\mathcal{R}}^*\big(U, w; \Xi, \zeta\big) = \kappa_1 \Big(\frac{abw}{a_* b_* w_*} \Big)^{1/2} \mathsf{C}^*\big(\xi_1{+}\xi_2{-}\zeta\big) + \kappa_2 \Big(\frac{acw}{a_* c_* w_*} \Big)^{1/2} \mathsf{C}^*\big(\xi_1{-}\xi_3{+}\zeta\big),$$

Thus, we can apply the theory of Section 3.1 and define $\mathscr{L}_{\rm red}$ as in (3.5), namely

$$\mathscr{L}_{\rm red}(U,\Xi) := \sup_{w>0} \inf_{\zeta\in\mathbb{R}} \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta).$$

The sup-inf can be calculated explicitly as is explained in [LM*17, Sec. 3.3.2]. Indeed using the formula

$$\inf_{\zeta\in\mathbb{R}} \big(g\mathsf{C}^*(\zeta)+h\mathsf{C}^*(\rho-\zeta)\big) = 4W(g,h,\rho)-4(g+h) \text{ with } W(g,h,\rho) = \Big((g+h)^2+\frac{gh}{2}\mathsf{C}^*(\rho)\Big)^{1/2},$$

where $\rho = 2\xi_1+\xi_2-\xi_3$, $g = \kappa_1\big(\frac{abw}{a_*b_*w_*}\big)^{1/2}$, and $h = \kappa_2\big(\frac{acw}{a_*c_*w_*}\big)^{1/2}$, a lengthy calculation yields

$$\mathscr{L}_{\rm red}(U,\Xi) := \sup_{w>0} \Big(4W(g,h,\rho) - 2\kappa_1\big(\frac{ab}{a_*b_*}+\frac{w}{w_*}\big) - 2\kappa_2\big(\frac{c}{c_*}+\frac{aw}{a_*w_*}\big)\Big)$$

Noting that g and h are proportional to $\sqrt{w}$, we see that also $W(g,h,\rho)$ is exactly proportional to $\sqrt{w}$. Hence, the maximum with respect to w can be determined and another lengthy calculation gives the explicit expression

$$\mathscr{L}_{\rm red}(U,\Xi) = \kappa_{\rm eff}(a)\big(\frac{a^2\,b\,c}{a_*^2 b_* c_*}\big)^{1/2}\,\mathsf{C}^*\big(2\xi_1+\xi_2-\xi_3\big) - \kappa_{\rm eff}(a)\,2\,\Big(\big(\frac{a^2\,b}{a_*^2 b_*}\big)^{1/2} - \big(\frac{c}{c_*}\big)^{1/2}\Big)^2$$

with $\kappa_{\rm eff}(a)$ from (4.3). Now, it can easily be checked that we have the duality structure $\mathscr{L}_{\rm red}(U,\Xi) = \mathcal{R}^*_{\rm eff}(U,\Xi) - \mathcal{R}^*_{\rm eff}(U,-\mathrm{D}E(U))$.

It seems that the above theory can be generalized to an arbitrary number of species with a density vector $\boldsymbol{c} = (c_1,...,c_{i_*}) \in \mathbb{R}^{i_*}$ and an arbitrary number r_* of reactions following the mass-action law, as long as we have the detailed-balance condition, i.e. there exists a positive steady state $\boldsymbol{c}^*_\varepsilon = (c^*_1,...,c_{j_*},\varepsilon w^*_{j+1},...,\varepsilon w^*_{i_*})$. If this is so, then the interesting question is how the reaction coefficients of the limiting system depend on the reaction coefficients of the original system. Note that even in our simple case, we can start with constant coefficients κ_1 and κ_2 but then find $\kappa_{\rm eff}(a)$ which depends on the state.

In particular, we want to highlight that the effective system has again the expected entropic cosh-gradient structure for the ternary reaction $2A + B \rightleftharpoons C$. We emphasize that this is not automatic, because in [MPS21, Sec. 4.3] an example of a reaction-rate equation is studied where the EDP-limit of the entropy cosh-gradient structure leads to an effective GS $(\mathbb{R}^4,\boldsymbol{E},\boldsymbol{R})$ where $\boldsymbol{E}$ is no longer a Boltzmann entropy and the reaction does no longer follow the mass-action law.

5 Linear diffusion through a membrane

The example in this section is well studied from the context of PDEs and singular limits. We are looking at a diffusion system of i_* mass densities $\boldsymbol{\rho} = (\rho_1,...,\rho_{i_*})$ that diffuse along an interval on the real line, where in the small interval $]{-\varepsilon},\varepsilon[$ representing a membrane the mobility is also of order ε, whereas it is of order 1 outside the membrane.

5.1 The PDE model and its limiting equation

We consider the intervals $\Omega_\varepsilon =]{-1}{-}\varepsilon, 1{+}\varepsilon[$ and define the piecewise affine maps ψ_ε and ϕ_ε between Ω_ε and $\Omega := \Omega_1 =]{-2}, 2[$:

$$\psi_\varepsilon(x) = \begin{cases} x{+}\varepsilon{-}1 & \text{for } x \geq 1, \\ \varepsilon\, x & \text{for } |x| \leq 1, \\ x{-}\varepsilon{+}1 & \text{for } y \leq -1; \end{cases} \quad \text{and} \quad \phi_\varepsilon(y) = \begin{cases} y{-}\varepsilon{+}1 & \text{for } y \geq \varepsilon, \\ y/\varepsilon & \text{for } |y| \leq \varepsilon, \\ y{+}\varepsilon{-}1 & \text{for } y \leq -\varepsilon. \end{cases} \tag{5.1}$$

The original diffusion problem is defined on Ω_ε and we assume that the mobility is given in the form

$$K_\varepsilon(y) = \frac{1}{\phi'_\varepsilon(y)}\, \overline{K}\big(\phi_\varepsilon(y)\big) \quad \text{with } \overline{K} \in \mathrm{PC}^0\big([-2,-1]\overline{\cup}[-1,1]\overline{\cup}[1,2]; \mathbb{R}^{i_*\times i_*}_{\mathrm{sym}}\big), \tag{5.2}$$

which means that $\overline{K}$ is piecewise continuous and has continuous extensions on the three closed intervals $[-2,-1]$, $[-1,1]$, and $[1,2]$, such that the one-sided limits

$$K_\pm := \overline{K}\big(\pm(1{+}0)\big) = \lim_{\delta\to 0^+} \overline{K}\big(\pm(1{+}\delta)\big) \quad \text{and} \quad k_\pm := \overline{K}\big(\pm(1{-}0)\big) = \lim_{\delta\to 0^+} \overline{K}\big(\pm(1{-}\delta)\big)$$

but may be different. Moreover, we assume that $\overline{K}$ is positive definite, i.e. there exists $\kappa > 0$ such that $a\cdot\overline{K}(x)a \geq \kappa|a|^2$ for all $x \in [-2,2]$ and $a \in \mathbb{R}^{i_*}$. Hence, $y \mapsto K_\varepsilon(y)$ is discontinuous at $y = \pm\varepsilon$, because it jumps by a factor of ε.

We define a second positive definite function $\overline{A} \in \mathrm{PC}^0\big([-2,-1]\overline{\cup}[-1,1]\overline{\cup}[1,2]; \mathbb{R}^{i_*\times i_*}_{\mathrm{sym}}\big)$ which determines the energy functional

$$\widetilde{\mathcal{E}}_\varepsilon(\boldsymbol{\rho}) := \int_{\Omega_\varepsilon} \frac{1}{2}\boldsymbol{\rho}(y)\cdot\overline{A}(\phi_\varepsilon(y))\boldsymbol{\rho}(y)\,\mathrm{d}y \ \text{ on the space } X_\varepsilon = \mathrm{L}^2(\Omega_\varepsilon;\mathbb{R}^{i_*}).$$

Moreover, we define the dual dissipation potential $\mathcal{R}_\varepsilon$ via

$$\widetilde{\mathcal{R}}_\varepsilon(\boldsymbol{\mu}) = \int_\Omega \frac{1}{2}\partial_y\boldsymbol{\mu}(y)\cdot K_\varepsilon(y)\partial_y\boldsymbol{\mu}(y)\,\mathrm{d}y.$$

The gradient-flow equation for the GS $(X_\varepsilon, \widetilde{\mathcal{E}}_\varepsilon, \widetilde{\mathcal{R}}_\varepsilon)$ is the linear parabolic system

$$\dot{\boldsymbol{\rho}} = \partial_y\Big(K_\varepsilon(y)\,\partial_y\big(A_\varepsilon(y)\boldsymbol{\rho}(t,y)\big)\Big) \ \text{ for } t>0,\ y\in\Omega_\varepsilon, \quad \partial_y\big(A_\varepsilon(y)\boldsymbol{\phi}(t,y)\big)\big|_{y=\pm(1+\varepsilon)} = 0.$$

Note that $M(t) = \int_{\Omega_\varepsilon}\rho(t,y)\,\mathrm{d}y$ is independent of t because of the divergence form and the no-flux boundary conditions.

To study the limit $\varepsilon\to 0$ it is advantageous to transform the PDE to the fixed interval Ω via $\phi_\varepsilon(\Omega_\varepsilon) = \Omega$. For $x\in\Omega$ we set

$$\boldsymbol{u}(t,x) = \frac{1}{\psi'_\varepsilon(x)}\,\boldsymbol{\rho}(t,\psi_\varepsilon(x)) \quad \text{and} \quad \mathcal{E}_\varepsilon(\boldsymbol{u}) = \widetilde{\mathcal{E}}_\varepsilon\big(\tfrac{1}{\phi'_\varepsilon}\boldsymbol{u}\circ\phi_\varepsilon\big) = \int_\Omega \frac{1}{2}\boldsymbol{u}\cdot\overline{A}\boldsymbol{u}\,\psi'_\varepsilon\,\mathrm{d}x.$$

The transformed dissipation potential takes the form

$$\mathcal{R}_\varepsilon(\boldsymbol{\xi}) = \int_\Omega \frac{1}{2}\partial_x\big(\frac{1}{\psi'_\varepsilon}\boldsymbol{\xi}\big)\cdot\overline{K}\,\partial_x\big(\frac{1}{\psi'_\varepsilon}\boldsymbol{\xi}\big)\,\mathrm{d}x,$$

where we used the scaling $\phi'_\varepsilon(y)K_\varepsilon(y) = \overline{K}(x)$ to cancel the powers of ψ'_ε.

The transformed linear diffusion equation reads

$$\psi_\varepsilon'(x)\,\dot{\boldsymbol{u}}(t,x) = \partial_x\Big(\overline{K}(x)\,\partial_x\big(\overline{A}(x)\boldsymbol{u}(t,x)\big)\Big), \quad \partial_x\big(\overline{A}(x)\boldsymbol{u}(t,x)\big)\big|_{x=\pm 2} = \mathbf{0}.$$

Of course, in the above development we have anticipated the scalings in such a way that in the last equation ε only occurs once, namely in the prefactor ψ_ε', namely $\psi_\varepsilon'(x) = \varepsilon$ for $|x| < 1$ and $\psi_\varepsilon'(x) = 1$ for $1 < |x| < 2$. Thus, we are exactly in the situation of a slow-fast gradient system as studied in Section 3.

We make the splitting and the corresponding port mappings explicit. We are in "Case 2" where the product space $X = X_{\rm slow}\times X_{\rm fast}$ needs a factorization along the boundary of the membrane, now placed at $x = \pm 1$. We set

$$\Omega_{\rm fast} = [-1,1], \quad \Omega_{\rm slow} = \,]{-2},-1]\cup[1,2[, \quad X_{\rm fast} = {\rm L}^2(\Omega_{\rm fast};\mathbb{R}^{i_*}), \quad X_{\rm slow} = {\rm L}^2(\Omega_{\rm slow};\mathbb{R}^{i_*}).$$

We introduce the variable $U = u|_{\Omega_{\rm slow}} \in X_{\rm slow}$ and $w = u|_{\Omega_{\rm fast}} \in X_{\rm fast}$. With this we find the transformed energy

$$\mathcal{E}_\varepsilon(U,w) = E(U) + \varepsilon\, e(w) \quad \text{with } E(U) = \int_{\Omega_{\rm slow}} \frac12 U\cdot \overline{A}U\,{\rm d}x \text{ and } e(w) = \int_{\Omega_{\rm fast}} \frac12 w\cdot\overline{A}w\,{\rm d}x.$$

If we similarly write $\boldsymbol{\xi} = (\Xi,\zeta)$ with $\Xi = \boldsymbol{\xi}|_{\Omega_{\rm slow}} \in X^*_{\rm slow}$ and $\zeta = \frac1\varepsilon\boldsymbol{\xi}|_{\Omega_{\rm fast}} \in X^*_{\rm fast}$ we obtain

$$\begin{aligned}&\overline{\mathcal{R}}^*(\Xi,\zeta) = \mathcal{R}^*_{\rm slow}(X) + \mathcal{R}^*_{\rm fast}(\zeta) + \boldsymbol{\delta}_{\{0\}}\big(P^\circ_{\rm slow}\Xi - P^\circ_{\rm fast}\zeta\big)\\ &\text{where } \mathcal{R}^*_{\rm slow}(\Xi) = \int_{\Omega_{\rm slow}}\frac12\partial_x\Xi\cdot\overline{K}\partial_x\Xi\,{\rm d}x \text{ and } \mathcal{R}^*_{\rm fast}(\zeta) = \int_{\Omega_{\rm fast}}\frac12\partial_x\zeta\cdot\overline{K}\partial_x\zeta\,{\rm d}x.\end{aligned} \tag{5.3}$$

Here the compatibility condition $P^\circ_{\rm slow}\Xi = P^\circ_{\rm fast}\zeta$ are crucial. We define $Y = \mathbb{R}^{i_*}\times\mathbb{R}^{i_*}$ and the port mappings

$$P_{\rm slow} : X_{\rm slow} \to Y; U \mapsto \big(U(-1^-),U(1^+)\big) \text{ and } P_{\rm fast} : X_{\rm fast} \to Y; w \mapsto \big(w(-1^+),w(1^-)\big),$$

and similarly $P^\circ_{\rm slow} : X^*_{\rm slow} \to Y^*$ and $P^\circ_{\rm fast}$. Here $f(x^+)$ and $f(x^-)$ denote the limit from the right and from the left, respectively.

The limiting equation for $\varepsilon = 0$ takes the form

$$\begin{aligned}\dot U &= \partial_x\big(\overline{K}\,\partial_x(\overline{A}U)\big) \text{ for } x\in\,]1,2[, && \partial_x(\overline{A}U)|_{x=2},\\ 0 &= \partial_x\big(\overline{K}\,\partial_x(\overline{A}w)\big) \text{ for } x\in\,]{-1},1[, && U(1^+) = w(1^-),\ \partial_x(\overline{A}U)|_{x=1^+} = \partial_x(\overline{A}w)|_{x=1^-}\\ \dot U &= \partial_x\big(\overline{K}\,\partial_x(\overline{A}U)\big) \text{ for } x\in\,]{-2},-1[, && U(-1^-) = w(-1^+),\ \partial_x(\overline{A}U)|_{x=-2},\\ & && \partial_x(\overline{A}U)|_{x=-1^-} = \partial_x(\overline{A}w)|_{x=-1^+}.\end{aligned}$$

The static equation on $\Omega_{\rm fast} = [-1,1]$ can be solved explicitly and we obtain the corresponding *transmission conditions*

$$\overline{K}\partial_x(\overline{A}U)\big|_{x=\pm1} = \mathbb{H}_K\big(A(1)U(1) - A(-1)U(-1)\big), \text{ where } \mathbb{H}_K = \Big(\int_{-1}^1 \overline{K}(x)^{-1}\,{\rm d}x\Big)^{-1}.$$

5.2 Lagrangian EDP-convergence for the quadratic case

To understand the origin of the transmission conditions, we want to use our Lagrangian EDP-convergence as described in Section 3.2. Thus, we want to construct

$$\mathscr{L}_{\mathrm{red}}(U,\Xi) = \sup\inf \mathscr{L}_{\overline{\mathcal{E}},\overline{\mathcal{R}}}(U,w;\Xi,\zeta) = \mathscr{L}_{E,\mathcal{R}_{\mathrm{slow}}}(U,\Xi) + \mathscr{L}_{\mathrm{red}}(U,\Xi)$$

We are in the case where $\overline{\mathcal{R}}$ is independent of the state, such that $\mathcal{R}_{\mathrm{red}}$ has the form

$$\mathcal{R}^*_{\mathrm{red}}(\Xi) = \mathsf{R}^*_Y(P^\circ\Xi) \quad \text{with } \mathsf{R}^*_Y(\eta) := \inf_{\Xi:\ P^\circ\Xi=\eta} \mathcal{R}^*_{\mathrm{fast}}(\Xi).$$

These relations are not the port relations $\mathfrak{P} : Y^* \to Y$, as these relations must be independent of the energy $\overline{\mathcal{E}} = E \otimes e$, which is given in terms of $\overline{A}$. A direct calculation shows that $\mathfrak{P}$ is given in terms of

$$\mathcal{R}^*_{\mathrm{red}}(\Xi) = \mathsf{R}^*_Y(P^\circ\Xi) \quad \text{with } \mathsf{R}_Y(\eta(-1),\eta(1)) = \frac{1}{2}\big(\eta(1){-}\eta(-1)\big)\cdot \mathbb{H}_K\big(\eta(1){-}\eta(-1)\big)$$

which shows $\mathfrak{P}(\eta(1),\eta(-1)) = \big(\mathbb{H}_K(\eta(1){-}\eta(-1))\,,\, \mathbb{H}_K(\eta(-1){-}\eta(1))\big)$. Indeed, $\mathcal{R}_{\mathrm{red}}$ can easily be obtained by minimizing $\mathcal{R}_{\mathrm{fast}}(\zeta)$ over the constraints $P^\circ_{\mathrm{fast}}\zeta = P^\circ\Xi$.

5.3 EDP-convergence in the Otto gradient structure

We reconsider the above linear equation, but now we strict to the one-dimensional case $i_* = 1$, viz. $u(t,x) \in [0,\infty[\in \mathbb{R}^1$. The linear equation can then be interpreted as a Fokker-Planck equation. Our aim is now to redo the EDP-limit $\varepsilon \to 0$ as in the previous subsection, but now for the so-called Otto gradient structure, also called gradient-flow in the Wasserstein space. The the gradient system is the triple $(\mathrm{Prob}(\Omega_\varepsilon), \mathcal{E}^{\mathrm{B}}_\varepsilon, \mathcal{R}^{\mathrm{Otto}}_\varepsilon)$, where the function space is

$$\mathrm{Prob}(\Omega_\varepsilon) := \Big\{\, u \in \mathrm{L}^1(\Omega_\varepsilon) \;\Big|\; u \geq 0,\ \int_{\Omega_\varepsilon} u \,\mathrm{d}y = 1 \,\Big\},$$

the energy is Boltzmann's relative entropy

$$\mathcal{E}^{\mathrm{B}}_\varepsilon(u) = \int_{\Omega_\varepsilon} \lambda_{\mathrm{B}}\big(A_\varepsilon(y)u(y)\big)\frac{\mathrm{d}y}{A_\varepsilon(y)},$$

and the dual dissipation functional reads

$$\mathcal{R}^{\mathrm{Otto}}_\varepsilon(u,\xi) = \int_{\Omega_\varepsilon} \frac{K_\varepsilon(y)}{2}|\partial_y\xi(y)|^2 u(y)\,\mathrm{d}y = \frac{1}{2}\big\langle \mathbb{K}^{\mathrm{Otto}}_\varepsilon(u)\xi,\xi\big\rangle,$$

which is quadratic in ξ and dependent on the state $u \in \mathrm{Prob}(\Omega_\varepsilon)$. Here $\mathbb{K}^{\mathrm{Otto}}_\varepsilon(u)$ can be understood as the self-adjoint nonnegative differential operator

$$\mathbb{K}^{\mathrm{Otto}}_\varepsilon(u)\,\xi = -\partial_y\big(K_\varepsilon u \partial_y\xi\big) \quad \text{with } K_\varepsilon u\partial_y\xi\big|_{y=\pm(1+\varepsilon)} = 0.$$

The associated gradient-flow equation is the Fokker-Planck equation

$$\dot u = \partial_y\big(K_\varepsilon u \partial_y(A_\varepsilon u)\big) = \partial_y\big(K_\varepsilon(\partial_y u + uV'_\varepsilon)\big),$$

if we define the driving potential V_ε by $V_\varepsilon(y) = \log A_\varepsilon(y)$. We refer to [Ott96, JKO97, Ott98, JKO98, Ott01] for the first work treating the Fokker-Planck equation as an gradient-flow equation with respect to this gradient structure.

We now want to do the EDP-limit in this gradient structure, where the new feature is the dependence on the state in $\mathcal{R}^*_\varepsilon$. As a result the limit gradient structure will be quite different. First it will depend in properties of $\overline{A}$ which shows that $\mathcal{R}^*_{\mathrm{slow}}$ cannot be calculated from $\mathcal{R}^*_\varepsilon$ alone. Secondly, we will see that $\mathcal{R}^*_{\mathrm{eff}} = \mathcal{R}^*_{\mathrm{slow}} + \mathcal{R}^*_{\mathrm{red}}$ will no longer be quadratic in ξ, namely $\mathcal{R}^*_{\mathrm{red}}$, which is obtained from the NESS problem of the rescaled membrane, will have a cosh-type behavior given through C^*.

We will not give the analysis in detail, as the result is well-established see [LM*17, Sec. 4], [Fre19, Sec. 4], [PeS22, FrM21]. However, we will give the main formal steps to put the results into the perspective of Section 3.2.

We first transform the problem as in Section with ψ_ε and ψ_ε from (5.1). With the notion from the previous subsection we have $\overline{\mathcal{E}}(U,w) = E(U) + e(w)$ with

$$E(U) = \int_{\Omega_{\mathrm{slow}}} \lambda_{\mathrm{B}}\big(\overline{A}(x)u(x)\big)\frac{1}{\overline{A}(x)}\,\mathrm{d}x \quad \text{and} \quad e(w) = \int_{\Omega_{\mathrm{fast}}} \lambda_{\mathrm{B}}\big(\overline{A}(x)w(x)\big)\frac{1}{\overline{A}(x)}\,\mathrm{d}x$$

and the rescaled dual dissipation potential $\overline{\mathcal{R}}^*(U,w;\Xi,\zeta) = \mathcal{R}^*_{\mathrm{slow}}(U,\Xi) + \mathcal{R}^*_{\mathrm{fast}}(w,\zeta) + \boldsymbol{\delta}_{\{0\}}\big(P^\circ_{\mathrm{slow}}\Xi - P^\circ_{\mathrm{fast}}\zeta\big)$ with

$$\mathcal{R}^*_{\mathrm{slow}}(U,\Xi) = \int_{\Omega_{\mathrm{slow}}} \frac{\overline{K}(Y)}{2}|\partial_x\Xi(x)|^2 U(x)\,\mathrm{d}x \text{ and } \mathcal{R}^*_{\mathrm{fast}}(w,\zeta) = \int_{\Omega_{\mathrm{fast}}} \frac{\overline{K}(Y)}{2}|\partial_x\zeta(x)|^2 w(x)\,\mathrm{d}x.$$

The reduced dissipation potential $\mathcal{R}_{\mathrm{red}}$ is now obtained by the saddle-point reduction, namely

$$\mathscr{L}_{E,\mathcal{R}_{\mathrm{red}}}(U,\Xi) = \sup_{w:\,P_{\mathrm{fast}}w = P_{\mathrm{slow}}U}\ \inf_{\zeta:\,P^\circ_{\mathrm{fast}}\zeta = P^\circ_{\mathrm{slow}}\Xi} \mathscr{L}_{e,\mathcal{R}_{\mathrm{fast}}}(w,\zeta),$$

where $\mathscr{L}_{e,\mathcal{R}_{\mathrm{fast}}}(w,\zeta) : X_{\mathrm{fast}} \times X^*_{\mathrm{fast}} \to \mathbb{R}$ takes the explicit form (using $\mathrm{D}e(w) = \log(\overline{A}w)$)

$$\begin{aligned}\mathscr{L}_{e,\mathcal{R}_{\mathrm{fast}}}(w,\zeta) &= \mathcal{R}^*_{\mathrm{fast}}(w,\zeta) - \mathcal{R}^*_{\mathrm{fast}}\big(w,-\mathrm{D}e(w)\big)\\ &= \int_{-1}^{1} \frac{1}{2}\Big(\overline{K}\,w\,|\partial_x\zeta|^2 - \frac{\overline{K}\,w}{(\overline{A}\,w)^2}\,\big|\partial_x(\overline{A}\,w)\big|^2\Big)\,\mathrm{d}x.\end{aligned}$$

It is surprising that the sup-inf of $\mathscr{L}_{e,\mathcal{R}_{\mathrm{fast}}}$ under given boundary conditions can be evaluated explicitly, see [LM*17, App. A] and also [PeS22, Sec. 1.3]. Here we provide a new and much shorter way of obtaining the desired result.

Theorem 5.1 (Membrane reduction) *Let $\overline{K}, \overline{A} \in \mathrm{L}^\infty([-1,1])$ be given and bounded from below by a positive constant. Then*

$$\mathcal{J}(w_-,w_+;\zeta_-,\zeta_+) := \sup_{\substack{w(1)=w_+\\ w(-1)=w_-}}\ \inf_{\substack{\zeta(1)=\zeta_+\\ \zeta(-1)=\zeta_-}} \mathscr{L}_{e,\mathcal{R}_{\mathrm{fast}}}(w,\zeta)$$

has the explicit form

$$\mathcal{J}(w_-,w_+;\zeta_-,\zeta_+) = K_{\mathrm{eff}}\sqrt{a_-w_-\,a_+w_+}\,\mathsf{C}^*\big(\zeta_+ - \zeta_-\big) - K_{\mathrm{eff}}\,2\big(\sqrt{a_+w_+} - \sqrt{a_-w_-}\big)^2 \quad (5.4)$$

where $K_{\mathrm{eff}} = \Big(\int_{-1}^{1} \overline{A}(x)/\overline{K}(x)\,\mathrm{d}x\Big)^{-1}$ and $a_\pm = \overline{A}(\pm(1-0))$.

Proof. Clearly, $\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}$ is strictly concave-convex and thus has at most one saddle point which is also the only critical point. Hence solving $\mathrm{D}\mathscr{L}_{e,\mathcal{R}_{\text{fast}}} = 0$ gives the solution.

However, it is advantageous to do a transformation first. We set

$$w = v/\overline{A}, \quad \zeta = \log(v/\eta^2), \quad \text{and } \mathcal{I}(v,\eta) = \mathscr{L}_{e,\mathcal{R}_{\text{fast}}}\big(v/\overline{A}, \log(v/\eta^2)\big).$$

An elementary calculation shows that $\mathcal{I}$ has a much simpler form, namely

$$\mathcal{I}(v,\eta) = -2\int_{-1}^{1} \overline{\kappa}\,\eta' \,\big(\frac{v}{\eta}\big)' \,\mathrm{d}x, \quad \text{where } \overline{\kappa}(x) = \overline{K}(x)/\overline{A}(x).$$

It will be particularly useful, that $\mathcal{I}$ is linear in v.

If (w_*, ζ_*) is a critical point for $\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}$, then the transformed point (v_*, ζ_*) is a critical point for $\mathcal{I}$, and vice versa. Moreover, the boundary values between the two pairs can be calculated easily. Hence, we have to determine the critical points of $\mathcal{I}$ and observe that

$$\mathrm{D}_v\mathcal{I}(v,\eta) = -\frac{2}{\eta}\,\big(\overline{\kappa}\,\eta'\big)'.$$

As the prefactor $2/\eta$ is irrelevant, we see that η_* is uniquely determined by its boundary values η_- and η_+. In particular, we know that $\overline{\kappa}\eta_*'$ must be constant, namely

$$\overline{\kappa}(x)\eta_*'(x) = K_{\text{eff}}\big(\eta_+ - \eta_-\big) \quad \text{for all } x \in [-1,1].$$

Because of $\mathrm{D}_v\mathcal{I}(v,\eta_*) = 0$, this is enough to evaluate $\mathcal{I}(v,\eta_*)$ explicitly by only knowing the boundary values v_- and v_* of v_*:

$$\mathcal{I}(u_*,\eta_*) = -2\int_{-1}^{1} \underbrace{\overline{\kappa}\,\eta'}_{=\text{const.}} \,\big(\frac{v}{\eta}\big)' \,\mathrm{d}x = 2\,K_{\text{eff}}\big(\eta_+ - \eta_-\big)\,\Big(\frac{v_+}{\eta_+} - \frac{v_-}{\eta_-}\Big).$$

Inserting the boundary conditions $v_\pm = a_\pm w_\pm$ and $\eta_\pm = \big(A_\pm w_\pm\big)^{1/2}\,\mathrm{e}^{-\zeta_\pm/2}$ gives

$$\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}(w_*,\zeta_*) = -2K_{\text{eff}}\Big(a_+w_+ - \sqrt{a_+w_+a_-w_-}\big(\mathrm{e}^{(\zeta_+-\zeta_-)/2} + \mathrm{e}^{(\zeta_--\zeta_+)/2}\big) + a_-w_-\Big),$$

which yields the desired formula (5.4). ■

Using the port conditions $P_{\text{fast}}w = P_{\text{slow}}U$ and $P^\circ_{\text{fast}}\zeta = P^\circ_{\text{slow}}\Xi$, the above result leads to the desired duality structure

$$\mathcal{J}(U_-,U_+;\Xi_-,\Xi_+) = \mathsf{R}^*_Y(U_-,U_+;\Xi_-,\Xi_+) - \mathsf{R}^*_y\big(U_-,U_+;\log(A_-U_-),\log(A_+U_+)\big),$$

where $A_\pm = \overline{A}(\pm(1{+}0))$.

In summary, we obtain the effective gradient system $(X_{\text{slow}}, E, \mathcal{R}_{\text{eff}})$ with

$$\mathcal{R}^*_{\text{eff}}(U,\Xi) = \mathcal{R}^*_{\text{slow}}(U,\Xi) + K_{\text{eff}}\sqrt{A_-U(-1)\,A_+U(1)}\,\mathsf{C}^*\big(\Xi(1){-}\Xi(-1)\big).$$

We clearly see that the effective contribution of the membrane is of cosh-type, and in particular it is not quadratic. Moreover, R_y depends on $\overline{A}$ which is information that stems from $\mathcal{E}_\varepsilon$, which was not present in $\mathcal{R}^*_\varepsilon$. Of course, also the cosh-type function C^* is inherited from $\mathcal{E}_\varepsilon$, namely from the Boltzmann function λ_{B}. Observe that $\mu = \lambda_{\text{B}}'(r) = \log r$ has the inversion $r = \mathrm{e}^\mu$. Using this for the forward and for the backward fluxes it is no longer surprising to obtain C^*.

5.4 Linear reaction-diffusion equation

Before going into the one-dimensional equation with membrane scaling, we note that the general structure of reaction-diffusion systems with detailed balance condition has the following gradient structure. On $X = \mathrm{L}^1(\Omega;\mathbb{R}^{i_*})$ we consider

$$\mathcal{E}(\boldsymbol{c}) = \mathcal{H}(\boldsymbol{c}|\boldsymbol{c}_*) = \int_\Omega \lambda_\mathrm{B}(c_i/c_i^*)c_i^* \,\mathrm{d}x \quad \text{and}$$
$$\mathcal{R}^*(\boldsymbol{c};\boldsymbol{\xi}) = \int_\Omega \Big(\sum_{i=1}^{i_*} \frac{K_i c_i}{2}|\nabla\xi_i|^2 + \sum_{r=1}^{r_*} \mu_r \big(\boldsymbol{c}^{\boldsymbol{\alpha}^r}\boldsymbol{c}^{\boldsymbol{\beta}^r}\big)^{1/2}\mathsf{C}^*\big(\boldsymbol{\xi}\cdot(\boldsymbol{\alpha}^r-\boldsymbol{\beta}^r)\big)\Big)\,\mathrm{d}x,$$

where $K_i \geq 0$ is the diffusion constants of species X_i, while $\mu_r > 0$ is the reaction coefficient of the rth reaction having stoichiometric vectors $\boldsymbol{\alpha}^r$, $\boldsymbol{\beta}^r \in \mathbb{N}_0^{i_*}$. The associated gradient-flow equation is the following system of i_* equations:

$$\dot c_i = \mathrm{div}\,\Big(K_i\big(\nabla c_i - \frac{c_i}{c_i^*}\nabla c_i^*\big)\Big) - \sum_{r=1}^{r_*}\mu_r\Big(\big(\frac{\boldsymbol{c}_*^{\boldsymbol{\beta}^r}}{\boldsymbol{c}_*^{\boldsymbol{\alpha}^r}}\big)^{1/2}\boldsymbol{c}^{\boldsymbol{\alpha}^r} - \big(\frac{\boldsymbol{c}_*^{\alpha^r}}{\boldsymbol{c}_*^{\boldsymbol{\beta}^r}}\big)^{1/2}\boldsymbol{c}^{\boldsymbol{\beta}^r}\Big)\big(\alpha_i^r - \beta_i^r\big).$$

In the same spirit as in the previous section we study again a linear PDE, but now it has diffusion and reaction with the background, i.e. $A \leftrightharpoons \emptyset$. Again we assume that the material parameters K_ε for diffusion and B_ε for reaction scale like ε in a the membrane region $]{-\varepsilon},\varepsilon[$. With $\Omega_\varepsilon =]{-1}{-\varepsilon}, 1{+}\varepsilon[$, the gradient system is given given via $X = \mathrm{L}^1(\Omega_\varepsilon)$,

$$\mathcal{E}_\varepsilon(u) = \mathcal{H}(u|1/A_\varepsilon) = \int_{\Omega_\varepsilon} \lambda_\mathrm{B}(A_\varepsilon u)\frac{1}{A_\varepsilon}\,\mathrm{d}y \quad \text{and}$$
$$\mathcal{R}^*(u,\xi) = \int_{\Omega_\varepsilon}\Big(\frac{K_\varepsilon u}{2}|\xi'|^2 + B_\varepsilon\sqrt{u}\,\mathsf{C}^*(\xi)\Big)\,\mathrm{d}y.$$

Using $\phi_\varepsilon : \Omega_\varepsilon \to \Omega := [-2,2]$ and $\psi_\varepsilon = \phi_\varepsilon^{-1} : \Omega \to \Omega_\varepsilon$ from (5.1) we assume that A_ε, B_ε, and K_ε are given in the form

$$A_\varepsilon(y) = \overline{A}(\psi_\varepsilon(y)), \quad B_\varepsilon(y) = \phi_\varepsilon'(y)\,\overline{B}(\phi_\varepsilon(y)), \quad K_\varepsilon(y) = \frac{1}{\phi_\varepsilon(y)}\,\overline{K}(\phi_\varepsilon(y)) \tag{5.5}$$

for given functions $\overline{A},\ \overline{B},\ \overline{K} \in \mathrm{PC}^0\big([-2,-1]\overline{\cup}[-1,1]\overline{\cup}[1,2]\big)$. To make our theory work we assume that $\overline{A}$ and $\overline{K}$ have a positive lower bound on Ω, whereas for $\overline{B}$ it is sufficient to have $\overline{B}(x) \geq 0$.

Transforming the system to the domain Ω as in the previous subsection, we obtain a slow-fast gradient system $(X_\mathrm{slow}{\times}X_\mathrm{fast},\mathcal{E}_\varepsilon,\mathcal{R}_\varepsilon)$ given by

$$X_\mathrm{slow} = \mathrm{L}^1(\Omega_\mathrm{slow}),\ \ X_\mathrm{fast} = \mathrm{L}^1(\Omega_\mathrm{fast}),\ \ \Omega_\mathrm{slow}[-2,-1]\cup[1,2],\ \ \Omega_\mathrm{fast} = [-1,1],$$
$$\mathcal{E}_\varepsilon(U,w) = E(u) + \varepsilon e(w),\ \ E(U) = \int_{\Omega_\mathrm{slow}} \lambda_\mathrm{B}(\overline{A}U)\frac{\mathrm{d}x}{\overline{A}},\ \ e(w) = \int_{\Omega_\mathrm{fast}}\lambda_\mathrm{B}(\overline{A}w)\frac{\mathrm{d}x}{\overline{A}},$$
$$\mathcal{R}_\varepsilon^*(U,w;\Xi,\xi) = \overline{\mathcal{R}}^*\big(U,w;\Xi,\frac{1}{\varepsilon}\xi\big) \ \text{ with}$$
$$\overline{\mathcal{R}}^*(U,w;\Xi,\zeta) = \overline{\mathcal{R}}^*_\mathrm{slow}(U,\Xi) + \overline{\mathcal{R}}^*_\mathrm{fast}(w,\zeta) + \boldsymbol{\delta}_{\{\mathbf{0}\}}(P^0_\mathrm{slow}\Xi{-}P^\circ_\mathrm{fast}\zeta),$$
$$\overline{\mathcal{R}}^*_\mathrm{slow}(U,\Xi) = \int_{\Omega_\mathrm{slow}}\Big(\frac{\overline{K}}{2}|\Xi'|^2U + \overline{B}\,\sqrt{U}\,\mathsf{C}^*(\Xi)\Big)\,\mathrm{d}x, \ \text{and}$$
$$\overline{\mathcal{R}}^*_\mathrm{fast}(w,\zeta) = \int_{\Omega_\mathrm{fast}}\Big(\frac{\overline{K}}{2}|\zeta'|^2 w + \overline{B}\,\sqrt{w}\,\mathsf{C}^*(\zeta)\Big)\,\mathrm{d}x.$$

As in the previous subsection we obtain the effective gradient structure $(X_{\rm slow}, E, \mathcal{R}_{\rm eff})$ by solving the sup-inf problem for the Lagrangian $\mathscr{L}_{\overline{E},\mathbb{R}}$ in the form $\mathcal{R}^*_{\rm eff} = \mathcal{R}^*_{\rm slow} + \mathcal{R}^*_{\rm red}$ with $\mathcal{R}^*_{\rm red} = \mathsf{R}^*_Y(P_{\rm slow}\cdot, P^\circ_{\rm slow})$, where we obtain an explicit formula for R_Y. To formulate the following result we introduce the two auxiliary functions $H_+,\ H_- : [-1,1] \to \mathbb{R}$ via

$$\big(\frac{\overline{K}}{\overline{A}}\, H'_\pm\big)' = \frac{\overline{B}}{\overline{A}^{1/2}}\, H_\pm \ \text{ in }]-1,1[, \quad H_\pm(\pm1) = 1, \quad H_\pm(\mp1) = 0. \tag{5.6}$$

Simple ODE arguments show $H_\pm(x) \in [0,1]$, $H'_-(x) < 0$, and $H'_+(x) > 0$ for all $x \in [-1,1]$.

Theorem 5.2 (Membrane with reaction and diffusion) *For the fast gradient system $(X_{\rm fast}, e, \mathcal{R}^*_{\rm fast})$ the reduced Lagrangian $\mathscr{L}_{\rm red}$ has the duality structure (e, R^*_Y) with*

$$\begin{aligned}\mathsf{R}^*_Y(w_-, w_+; \zeta_-, \zeta_+) = {}& M_{\rm eff} \sqrt{\overline{A}(-1)w_-\overline{A}(1)w_+}\, \mathsf{C}^*(\zeta_+ - \zeta_-) \\ & + M_- \sqrt{\overline{A}(-1)w_-}\, \mathsf{C}^*(\zeta_-) + M_+ \sqrt{\overline{A}(1)w_+}\, \mathsf{C}^*(\zeta_+),\end{aligned} \tag{5.7}$$

where $M_{\rm eff} = \overline{K}(1)|H'_-(1)|/\overline{A}(1) = \overline{K}(-1)H'_+(-1)/\overline{A}(-1)$ and $M_\pm = \int_{-1}^1 \overline{B}\, H_\pm/\overline{A}^{1/2}\,{\rm d}x.$

In the case of constant coefficients we have

$$M_{\rm eff} = \frac{\overline{B}}{\overline{A}} \frac{\sigma \cosh(2\sigma)}{\sinh(2\sigma)} \quad \textit{and} \quad M_+ = M_- = \frac{\overline{B}}{\overline{A}^{1/2}} \frac{\sinh(\sigma)}{\sigma\cosh(\sigma)}.$$

Proof. As in Theorem 5.1 we do a transformation to characterize the unique saddle point (w_*, ζ_*). With $w = v/\overline{A}$ and $\zeta = \log(\overline{A}w/\eta^2)$, the Lagrangian $\mathscr{L}_{e,\mathcal{R}_{\rm fast}}$ gives

$$\mathcal{I}(v,\eta) := \mathscr{L}_{e,\mathcal{R}_{\rm fast}}\big(v/\overline{A}, \log(v/\eta^2)\big) = \int_{-1}^1 \Big(-2\overline{\kappa}\eta'\big(\frac{v}{\eta}\big)' + 2\overline{\beta}\frac{1-\eta}{\eta}(v-\eta)\Big)\,{\rm d}y, \tag{5.8}$$

where $\overline{\kappa} = \overline{K}/\overline{A}$ and $\overline{\beta} = \overline{B}/\overline{A}^{1/2}$. Here we used the specific interaction of C^* and $\log = \lambda'_{\rm B}$, namely $\mathsf{C}^*(\log\alpha) = \big(\alpha^{1/4} - \alpha^{-1/4}\big)^2$. Of course, the construction is such that $\eta \equiv 1$ leads to $\mathcal{I}(v,1) = 0$.

The surprising and helpful fact is that $\mathcal{I}$ is affine in v which allows us to evaluate $\mathscr{L}_{e,\mathcal{R}_{\rm fast}}(w_*,\zeta_*) = \mathcal{I}(v_*,\eta_*)$ at the unique critical point. In particular, we have

$$0 = {\rm D}_v\mathcal{I}(v,\eta) = \frac{2}{\eta}\Big(\big(\overline{\kappa}\eta'\big)' - \overline{\beta}\,\eta + \overline{\beta}\Big),$$

such that the critical point (v_*,η_*) satisfies the linear ODE $-(\overline{\kappa}\eta')' + \overline{\beta}\eta = \overline{\beta}$. Hence,

$$\begin{aligned}\mathscr{L}_{e,\mathcal{R}_{\rm fast}}(w_*,\zeta_*) = \mathcal{I}(v_*,\eta_*) &= \int_{-1}^1 \Big(-2\overline{\kappa}\eta'_*\big(\frac{v_*}{\eta_*}\big)' + 2\overline{\beta}\frac{1-\eta_*}{\eta}(v_*-\eta_*)\Big)\,{\rm d}y \\ &= \Big[-2\overline{\kappa}\eta'_*\frac{v_*}{\eta_*}\Big]_{x=-1}^1 + \int_{-1}^1 \Big(\frac{2}{\eta_*}\big(\underbrace{(\overline{\kappa}\eta'_*)' + \overline{\beta}(1-\eta_*)}_{=0}\big) + 2\,\underbrace{\overline{\beta}(\eta_*-1)}_{=(\overline{\kappa}\eta'_*)'}\Big)\,{\rm d}x \\ &= 2\overline{\kappa}_-\eta'_*(-1)\big(\frac{v_-}{\eta_-} - 1\big) + 2\overline{\kappa}_+\eta'_*(1)\big(1 - \frac{v_+}{\eta_+}\big),\end{aligned}$$

where $\overline{\kappa}_\pm = \overline{\kappa}(\pm1)$, $v_\pm = v_*(\pm1)$, and $\eta_\pm = \eta_*(\pm1)$.

Using the auxiliary functions $H_\pm$ we have $\eta_* = 1+(\eta_-{-}1)H_- + (\eta_+{-}1)H_+$ which gives $\eta_*'(\pm 1) = (\eta_-{-}1)H_-'(\pm 1) + (\eta_+{-}1)H_+'(\pm 1)$. Abbreviating $b_\pm := \sqrt{v_\pm}$ and $E_\pm := \mathrm{e}^{\zeta_\pm/2}$ and using $\eta_\pm = \sqrt{v_\pm}\mathrm{e}^{-\zeta_\pm/2} = b_\pm E_\pm^{-1}$ we obtain

$$\begin{aligned}\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}(w_*,\zeta_*) = 2\overline{\kappa}_+ H_+'(1)\big(b_+(E_+{+}E_+^{-1}){-}b_-^2{-}1\big) - 2\overline{\kappa}_- H_-'(-1)\big(b_-(E_-{+}E_-^{-1}){-}b_-^2{-}1\big)\\ + 2\overline{\kappa}_- H_+'(-1)(b_+E_+^{-1}{-}1)(b_-E_-{-}1) - 2\overline{\kappa}_+ H_-'(1)(b_-E_-^{-1}{-}1)(b_+E_+{-}1).\end{aligned}$$

To simplify this expression, we use that the Wronski determinant $\overline{\kappa}H_+'H_- - \overline{\kappa}H_-'H_+$ is constant on $[-1,1]$, and we call this constant $M_{\text{eff}} > 0$. Using the boundary conditions of $H_\pm$ we have $M_{\text{eff}} = \overline{\kappa}_- H_+'(-1) = -\overline{\kappa}_+ H_-'(1)$. Moreover, integrating the ODE (5.6) yields

$$\pm\overline{\kappa}_\pm H_\pm'(\pm 1) = \pm\overline{\kappa}_\mp H_\pm'(\mp 1) + \int_{-1}^{1} \overline{\kappa} H_\pm \,\mathrm{d}x = M_{\text{eff}} + M_\pm .$$

With this we arrive at

$$\begin{aligned}\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}(w_*,\zeta_*) = {}& 2M_-\big(b_-(E_-{+}E_-^{-1}{-}2) - (b_-{-}1)^2\big)\\ &+ 2M_+\big(b_+(E_+{+}E_+^{-1}{-}2) - (b_+{-}1)^2\big)\\ &+ 2M_{\text{eff}}\big(b_+b_-(E_+E_-^{-1}{+}E_+^{-1}E_-{-}2) - (b_+{-}b_-)^2\big).\end{aligned}$$

Inserting $E_\pm = \mathrm{e}^{-\zeta_\pm/2}$ and $b_\pm = \sqrt{v_\pm} = \sqrt{a_\pm w_\pm}$ yields $\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}(w_*,\zeta_*) =$

$$\begin{aligned}\mathscr{L}_{e,\mathcal{R}_{\text{fast}}}(w_*,\zeta_*) &= M_-\big(\sqrt{a_-w_-}\,\mathsf{C}^*(\zeta_-) - 2(\sqrt{a_-w_-}{-}1)^2\big)\\ &\quad + M_+\big(\sqrt{a_+w_+}\,\mathsf{C}^*(\zeta_+) - 2(\sqrt{a_+w_+}{-}1)^2\big)\\ &\quad + M_{\text{eff}}\,\big(\sqrt{a_+w_+a_-w_-}\,\mathsf{C}^*(\zeta_+{-}\zeta_-) - 2(\sqrt{a_+w_+}{-}\sqrt{a_-w_-})^2\big)\\ &= \mathsf{R}_Y^*(w_-,w_+;\zeta_-,\zeta_+) - \mathsf{R}_Y^*\big(w_-,w_+;\log(a_-w_-),\log(a_+w_+)\big),\end{aligned}$$

which is the desired general formula (5.7).

The special formula for constant coefficients follows by setting $\sigma^2 = \overline{A}^{1/2}\overline{B}/\overline{K}$ and observing $H_\pm(x) = \sinh(\sigma \pm \sigma x)/\sinh(2\sigma)$. ■

A Classical existence theory for saddle points

We recollect the basic result from saddle point theory as contained in [EkT74, Cha. VI] (La dualité par les minimax).

We consider a Lagrangian functional $\mathfrak{L} : \boldsymbol{U}\times\boldsymbol{V} \to \overline{\mathbb{R}} = [-\infty,\infty]$, where we now want to minimize with respect to $x \in \boldsymbol{U}$ and maximize with respect to $y \in \boldsymbol{V}$. This means that for applying the theory below to the Lagrangians used above we have to set $\boldsymbol{U} = X$, $\boldsymbol{V} = X^*$, and $\mathfrak{L}(u,\xi) = -\mathscr{L}(u,\xi)$. Now, a point (x_*,y_*) is called a *saddle point* of $\mathfrak{L}$ if

$$\forall\, x \in \boldsymbol{U},\ y \in \boldsymbol{V}: \quad \mathfrak{L}(x_*,y) \leq \mathfrak{L}(x_*,y_*) \leq \mathfrak{L}(x,y_*).$$

Thus, we minimize with respect to $x \in \boldsymbol{U}$, and we maximize with respect to y.

The aim is to find a saddle point from general principles. For this one looks at $\sup_{y\in\boldsymbol{V}} \inf_{x\in\boldsymbol{U}} \mathfrak{L}(x,y)$ and $\inf_{x\in\boldsymbol{U}} \sup_{y\in\boldsymbol{V}} \mathfrak{L}(x,y)$. We obviously always have a one-sided estimate, and the major question in constructing saddle points is when we have equality.

Lemma A.1 (Simple facts on saddles points)

$$\text{(a)}\qquad \mathrm{SI}_{\mathfrak{L}} := \sup_{y\in \boldsymbol{V}} \inf_{x\in \boldsymbol{U}} \mathfrak{L}(x,y) \;\leq\; \inf_{x\in \boldsymbol{U}} \sup_{y\in \boldsymbol{V}} \mathfrak{L}(x,y) := \mathrm{IS}_{\mathfrak{L}} \tag{A.1}$$

$$\text{(b)}\qquad \text{saddle point } (x_*,y_*) \text{ exists} \;\Longrightarrow\; \mathrm{SI}_{\mathfrak{L}} = \mathrm{IS}_{\mathfrak{L}}. \tag{A.2}$$

In the latter case, we have $\mathfrak{L}(x_*,y_*) = \mathrm{SI}_{\mathfrak{L}} = \mathrm{IS}_{\mathfrak{L}}$.

Proof. To show (a), we start from $\mathfrak{L}(x,y) \leq \sup_{\overline{y}} \mathfrak{L}(x,\overline{y})$. Taking the infimum over x we obtain $\inf_x \mathfrak{L}(x,y) \leq \mathrm{IS}_{\mathfrak{L}}$. Now taking the supremum over y in the left-hand side leads to the desired estimate $\mathrm{SI}_{\mathfrak{L}} \leq \mathrm{IS}_{\mathfrak{L}}$.

To show (b) simply note that the saddle-point property implies

$$\inf_{x\in \boldsymbol{U}} \mathfrak{L}(x,y_*) = \mathfrak{L}(x_*,y_*) = \sup_{y\in \boldsymbol{V}} \mathfrak{L}(x_*,y).$$

Thus, we find $\mathrm{SI}_{\mathfrak{L}} \geq \mathfrak{L}(x_*,y_*) \geq \mathrm{IS}_{\mathfrak{L}}$. With (a) this implies the desired equality. ∎

The quantity $\delta_{\mathfrak{L}} = \mathrm{IS}_{\mathfrak{L}} - \mathrm{SI}_{\mathfrak{L}} \geq 0$ is called the *duality gap*. The function $\mathfrak{L}(x,y) = \tanh(x{-}y)$ on $\mathbb{R}{\times}\mathbb{R}$ shows that $\delta_{\mathfrak{L}}$ can be positive. Indeed, $\mathrm{SI}_{\tanh} = -1$ and $\mathrm{IS}_{\tanh} = +1$ such that $\delta_{\tanh} = 2$.

The opposite implication in (A.2) is not valid. To see this consider $\boldsymbol{U} = \boldsymbol{V} = \mathbb{R}$ and $\mathfrak{L}(x,y) = \mathrm{e}^x - \mathrm{e}^{-y}$. Clearly, $\inf_x \mathfrak{L}(x,y) = -\mathrm{e}^{-y}$ and hence, $\mathrm{SI}_{\mathfrak{L}} = 0$ and similarly $\mathrm{IS}_{\mathfrak{L}} = 0$. However, no saddle-point exists. Even in cases where no saddle-point exists it is an interesting question under what conditions the duality gap is 0, see e.g. [EkT74, Ch. III, Prop. 2.3].

If two saddle points (x_j,y_j) with $j = 1,2$ exist, we have

$$\mathfrak{L}(x_1,y_2) \leq \mathfrak{L}(x_1,y_1) \leq \mathfrak{L}(x_2,y_1) \leq \mathfrak{L}(x_2,y_2) \leq \mathfrak{L}(x_1,y_2),$$

which means that all four points have the same value. If each $\mathfrak{L}(\cdot,y_j)$ is convex and each $\mathfrak{L}(x_j,\cdot)$ concave, then we conclude $\mathfrak{L}(x,y) = \mathfrak{L}(x_1,y_1)$ for all $x = (1{-}s)x_1 + sx_2$ and $y = (1{-}r)y_1 + ry_2$ with arbitrary $r,s \in [0,1]$.

A standard existence result for saddle points can be found in [EkT74, Ch. VI, Prop. 2.1]. We provide a variant that is adjusted to our purposes.

Proposition A.2 (Existence of saddle points) *Consider reflexive Banach spaces* $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *and assume that the following conditions hold:*

$$\forall\, y \in \boldsymbol{V}: \quad x \mapsto \mathfrak{L}(x,y) \text{ is convex and lsc}, \tag{A.3a}$$

$$\forall\, x \in \boldsymbol{U}: \quad y \mapsto -\mathfrak{L}(x,y) \text{ is convex and lsc}, \tag{A.3b}$$

$$\exists\, y_0 \in \boldsymbol{V}: \quad \mathfrak{L}(\,\cdot\,,y_0) \text{ is coercive}, \tag{A.3c}$$

$$\exists\, x_0 \in \boldsymbol{U}: \quad -\mathfrak{L}(x_0,\,\cdot\,) \text{ is coercive}. \tag{A.3d}$$

Then, a saddle point (x_*,y_*) *exists and*

$$\mathfrak{L}(x_*,y_*) = \min_{x\in \boldsymbol{U}} \sup_{y\in \boldsymbol{V}} \mathfrak{L}(x,y) = \max_{y\in \boldsymbol{V}} \inf_{x\in \boldsymbol{U}} \mathfrak{L}(x,y).$$

If moreover, in (A.3a) *and* (A.3b) *we have strict convexity, then the saddle point is unique.*

Proof. *Step 1: Saddle points on balls using strict convexity.* We additionally impose that

$$\forall\, y \in \boldsymbol{V}: \quad \mathfrak{L}(\,\cdot\,, y) : \boldsymbol{U} \to \overline{\mathbb{R}} \text{ is strictly convex.} \tag{A.4}$$

For $R \geq R_0 := \max\{\|x_0\|_{\boldsymbol{U}}, \|y_0\|_{\boldsymbol{V}}\}$ we consider the closed and convex balls $\boldsymbol{U}_R = \{\, x \in \boldsymbol{U} \mid \|x\|_{\boldsymbol{U}} \leq R\,\}$ and similarly $\boldsymbol{V}_R$.

For all R we obtain a saddle point (x_R, y_R) as follows. For all $y \in \boldsymbol{V}_R$ the direct method of the calculus of variations provides a minimizer $x = \widehat{x}_R(y) \in \boldsymbol{U}_R$ for $\mathfrak{L}(\cdot, y)|_{\boldsymbol{U}_R}$, i.e. $\mathfrak{L}(\widehat{x}_R(y), y) = \min_{x\in\boldsymbol{U}_R} \mathfrak{L}(x, y) =: \lambda_R(y)$. By the strict convexity in (A.4) $\widehat{x}_R(y)$ is uniquely determined.

We first observe that $-\lambda_R : \boldsymbol{V}_R \to \overline{\mathbb{R}}$ is convex and lsc, as it is the supremum of the convex and lsc functions $-\mathfrak{L}(x, \cdot)$. Moreover, by (A.3d) the function $-\lambda_R$ is bounded from below by the proper, lsc, convex function $-\mathfrak{L}(x_0, \cdot)$. Hence, λ_R attains its maximum in a point $y^R \in \boldsymbol{V}_R$.

Our aim is now to show that $(\widehat{x}_R(y^R), y^R)$ is a saddle point of $\mathfrak{L}$ on $\boldsymbol{U}_R{\times}\boldsymbol{V}_R$. For this we choose arbitrary $y \in \boldsymbol{V}_R$ and $\theta \in [0, 1]$ and set $x_\theta(y) := \widehat{x}_R((1{-}\theta)y^R{+}\theta y)$ and obtain

$$\begin{aligned}\lambda_R(y^R) &\geq \lambda_R\big((1{-}\theta)y^R{+}\theta y\big) = \mathfrak{L}\big(x_\theta(y), (1{-}\theta)y^R{+}\theta y\big)\\ &\overset{-\mathfrak{L}(x_\theta(y),\cdot)\text{ cvx}}{\geq} (1{-}\theta)\mathfrak{L}(x_\theta, y^R) + \theta\mathfrak{L}(x_\theta(y), y) \geq (1{-}\theta)\lambda_R(y^R) + \theta\mathfrak{L}(x_\theta(y), y).\end{aligned}$$

In particular, for $\theta \in \,]0, 1]$ and all $y \in \boldsymbol{V}_R$ we conclude

$$\lambda_R(y^R) = \mathfrak{L}(\widehat{x}_R(y^R)) \geq \mathfrak{L}(x_\theta(y), y) \quad \text{for all } y \in \boldsymbol{V}_R. \tag{A.5}$$

Choosing $\theta = 1/k$ for $k \in \mathbb{N}$, we obtain $x_k := x_{1/k}(y) \in \boldsymbol{V}_R$ and may select a weakly convergent subsequence (not relabeled) with $x_k \rightharpoonup x^R$. We claim that $x^R = \widehat{x}_R(y^R)$ and hence is independent of y. Indeed, for our fixed $y \in \boldsymbol{V}_R$ and arbitrary $\widetilde{x} \in \boldsymbol{U}_R$ we have

$$\begin{aligned}\mathfrak{L}(x^R, y^R) &\overset{\text{(A.3a),lsc}}{\leq} \liminf_{k\to\infty} \mathfrak{L}(x_k, y^R)\\ &\overset{\text{(A.3b),cvx}}{\leq} \limsup_{k\to\infty} \frac{1}{1{-}\frac{1}{k}}\Big(\mathfrak{L}(x_k, (1{-}\tfrac{1}{k})y^R{+}\tfrac{1}{k}y) - \frac{1}{k}\mathfrak{L}(x_k, y)\Big)\\ &\overset{\text{def. }\lambda_R,\,\widehat{x}_R}{\leq} \limsup_{k\to\infty} \Big(\frac{k}{k{-}1}\mathfrak{L}(\widetilde{x}, (1{-}\tfrac{1}{k})y^R{+}\tfrac{1}{k}y) - \frac{1}{k{-}1}\lambda_R(y)\Big)\\ &\overset{\lambda_R(y)<\infty}{\leq} \limsup_{k\to\infty} \mathfrak{L}(\widetilde{x}, (1{-}\tfrac{1}{k})y^R{+}\tfrac{1}{k}y) \overset{\text{(A.3b),lsc}}{\leq} \mathfrak{L}(\widetilde{x}, y^R),\end{aligned}$$

where we used $(1{-}\frac{1}{k})y^R{+}\frac{1}{k}y \to y^R$ in the last step. Since $\widetilde{x} \in \boldsymbol{U}_R$ was arbitrary we obtain $\lambda_R(y^R) \leq \mathfrak{L}(x^R, y^R) \leq \min_{\widetilde{x}\in\boldsymbol{U}_R} \mathfrak{L}(\widetilde{x}, y^R) = \lambda_R(y^R)$. Hence, $\boldsymbol{U}^R$ is a minimizer of $\mathfrak{L}(\cdot, y^R)$ and hence coincides with $\widehat{x}_R(y^R)$ because of the strict convexity (A.4).

Because of the uniqueness of the limit we conclude that for all $y \in \boldsymbol{V}_R$ we have $x_\theta(y) \rightharpoonup x^R = \widehat{x}_R(y^R)$ for $\theta \to 0^+$. Thus, taking the limit $\theta \to 0^+$ in (A.5) and exploiting the lsc from (A.3a) we obtain

$$\forall\, y \in \boldsymbol{V}_R\, \forall\, \widetilde{x} \in \boldsymbol{U}_R: \quad \mathfrak{L}(x^R, y) \leq \lambda_R(y^R) = \mathfrak{L}(x^R, y^R) \leq \mathfrak{L}(\widetilde{x}, y^R).$$

This shows that (x^R, y^R) is a saddle point for $\mathfrak{L}$ restricted to $\boldsymbol{U}_R{\times}\boldsymbol{V}_R$.

Step 2: Saddle points on balls without strict convexity. If we only have convexity we consider

$$\mathfrak{L}_\varepsilon(x, y) = \mathfrak{L}(x, y) + \varepsilon\|x\|^2 \quad \text{with } \varepsilon > 0,$$

where we can choose a strictly convex norm $\|\cdot\|$ on the reflexive space $\boldsymbol{U}$. By Step 1 we obtain a saddle point $(x_\varepsilon^R, y_\varepsilon^R) \in \boldsymbol{U}_R \times \boldsymbol{V}_R$. Hence, we have

$$\forall\, y \in \boldsymbol{V}_R\, \forall\, x \in \boldsymbol{U}_R : \quad \mathfrak{L}(x_\varepsilon^R, y) + \varepsilon\|x_\varepsilon^R\|^2 \leq \mathfrak{L}(x_\varepsilon^R, y_\varepsilon^R) + \varepsilon\|x_\varepsilon^R\|^2 \leq \mathfrak{L}(x, y_\varepsilon^R) + \varepsilon\|x\|^2. \quad \text{(A.6)}$$

We may choose a subsequence (not relabeled) with $(x_\varepsilon^R, y_\varepsilon^R) \rightharpoonup (\overline{x}^R, \overline{y}^R)$ in $\boldsymbol{U} \times \boldsymbol{V}$. Dropping the middle term in (A.6) we can pass to the limit using the lsc in (A.3a) and (A.3b) and arrive at

$$\forall\, y \in \boldsymbol{V}_R\, \forall\, x \in \boldsymbol{U}_R : \quad \mathfrak{L}(\overline{x}^R, y) \leq \mathfrak{L}(x, \overline{y}^R).$$

Thus, $(\overline{x}^R, \overline{y}^R)$ is indeed a saddle point for $\mathfrak{L}$ restricted to restricted to $\boldsymbol{U}_R \times \boldsymbol{V}_R$.

Step 3: Unbounded case. We now consider the limit $R \to \infty$. Using the coercivities (A.3c) and (A.3d). For $R \geq R_0$ the saddle points (x^R, y^R) from Step 2 satisfy

$$\mathfrak{L}(x_R, y_0) \leq \mathfrak{L}(x^R, y^R) \leq \mathfrak{L}(x_0, y^R). \quad \text{(A.7)}$$

Since $\mathfrak{L}(\cdot, y_0)$ and $-\mathfrak{L}(x_0, \cdot)$ are lsc and coercive (cf. (A.3c) and (A.3d)), they are bounded from below:

$$\exists\, M > 0\, \forall\, x \in \boldsymbol{U}\, \forall\, y \in \boldsymbol{V} : \quad \mathfrak{L}(x, y_0) \geq -M \text{ and } \mathfrak{L}(x_0, y) \leq M.$$

Combining this with (A.7), we have

$$\forall\, R \geq R_0 : \quad \text{(i) } \mathfrak{L}(x^R, y_0) \geq -M \text{ and (ii) } \mathfrak{L}(x_0, y^R) \leq M.$$

With (A.7) we obtain $|\mathfrak{L}(x^R, y^R)| \leq M$. Using the coercivity (A.3c) and (ii) we find $\|y^R\| \leq C_{\boldsymbol{V}}$, and similarly (A.3d) and (i) give $\|x^R\| \leq C_{\boldsymbol{U}}$. Thus, using the reflexivity of $\boldsymbol{U}$ and $\boldsymbol{V}$ we find a subsequence (x^R, y^R) (not relabeled) such that

$$\mathfrak{L}(x^R, y^R) \to \lambda_*, \quad x^R \rightharpoonup x_* \text{ in } \boldsymbol{U}, \quad y^R \rightharpoonup y_* \text{ in } \boldsymbol{V}.$$

For arbitrary $x \in \boldsymbol{U}$ we choose $R > \max\{R_0, \|x\|\}$ and obtain $\mathfrak{L}(x, y^R) \geq \mathfrak{L}(x^R, y^R)$. Taking the limit $R \to \infty$ (along the subsequence) and using the lsc of $-\mathfrak{L}(x, \cdot)$ we arrive at

$$\mathfrak{L}(x, y_*) \geq \limsup_{R\to\infty} \mathfrak{L}(x, y^R) \geq \limsup_{R\to\infty} \mathfrak{L}(x^R, y^R) = \lambda_*,$$

where $x \in \boldsymbol{U}$ was arbitrary. Similarly, we obtain $\mathfrak{L}(x_*, y) \leq \lambda_*$ which gives the desired saddle-point property for $(x_*, y_*) \in \boldsymbol{U} \times \boldsymbol{V}$:

$$\forall\, x \in \boldsymbol{U},\ y \in \boldsymbol{V} : \quad \lambda(x_*, y) \leq \lambda_* = \lambda(x_*, y_*) \leq \mathfrak{L}(x, y_*).$$

Step 4: Uniqueness under strict convexity. This was shown already in Step 1.

This completes the proof of Proposition A.2. ■

References

[ASGB95] A. A. SANTOS, V. GARZÓ, and J. J. BREY. On the validity of a variational principle for far-from-equilibrium steady states. *Europhys. Lett.*, 29(9), 693–698, 1995.

[DD*12] A. DE MARTINO, D. DE MARTINO, R. MULET, and G. UGUZZONI. Reaction networks as systems for resource allocation: A variational principle for their non-equilibrium steady states. *PLoS ONE*, 7(7), e39849/1–9, 2012.

[DeM84] S. DE GROOT and P. MAZUR. *Non-Equilibrium Thermodynamics*. Dover Publ., New York, 1984.

[DFM19] P. DONDL, T. FRENZEL, and A. MIELKE. A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM Control Optim. Calc. Var.*, 25, 68/1–45, 2019.

[EkT74] I. EKELAND and R. TEMAM. *Analyse Convexe et Problèmes Variationnels*. Dunod, 1974.

[Fre19] T. FRENZEL. *On the derivation of effective gradient systems via EDP-convergence*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2019. Defense date 26.6.2019, publication date 10.6.2020, epub HU Berlin.

[FrL21] T. FRENZEL and M. LIERO. Effective diffusion in thin structures via generalized gradient systems and EDP-convergence. *Discr. Cont. Dynam. Systems Ser. S*, 14(1), 395–425, 2021.

[FrM21] T. FRENZEL and A. MIELKE. Deriving the kinetic flux relation for nonlinear diffusion through a membrane using edp-convergence. *In preparation*, 2021.

[JKO97] R. JORDAN, D. KINDERLEHRER, and F. OTTO. Free energy and the fokker-planck equation. *Physica D*, 107(2-4), 265–271, 1997.

[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.

[LM*17] M. LIERO, A. MIELKE, M. A. PELETIER, and D. R. M. RENGER. On microscopic origins of generalized gradient structures. *Discr. Cont. Dynam. Systems Ser. S*, 10(1), 1–35, 2017.

[MaM20] J. MAAS and A. MIELKE. Modeling of chemical reaction systems with detailed balance using gradient structures. *J. Stat. Physics*, 181, 2257–2303, 2020.

[Mie16] A. MIELKE. On evolutionary Γ-convergence for gradient systems (Ch. 3). In A. Muntean, J. Rademacher, and A. Zagaris, editors, *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, Lecture Notes in Applied Math. Mechanics Vol. 3, pages 187–249. Springer, 2016. Proc. of Summer School in Twente University, June 2012.

[MiS20] A. MIELKE and A. STEPHAN. Coarse graining via EDP-convergence for linear fast-slow reaction systems. *Math. Models Meth. Appl. Sci. (M^3AS)*, 30(9), 1765–1807, 2020. (In the published version, Lemma 3.4 is wrong. See arXiv:1911.06234v2 for the correction.).

[MMP21] A. MIELKE, A. MONTEFUSCO, and M. A. PELETIER. Exploring families of energy-dissipation landscapes via tilting — three types of EDP convergence. *Contin. Mech. Thermodyn.*, 33, 611–637, 2021.

[MP*17] A. MIELKE, R. I. A. PATTERSON, M. A. PELETIER, and D. R. M. RENGER. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. *SIAM J. Appl. Math.*, 77(4), 1562–1585, 2017.

[MPS21] A. MIELKE, M. A. PELETIER, and A. STEPHAN. EDP-convergence for nonlinear fast-slow reaction systems with detailed balance. *Nonlinearity*, 34(8), 5762–5798, 2021.

[Ott96] F. OTTO. Double degenerate diffusion equations as steepest descent. Preprint no. 480, SFB 256, University of Bonn, 1996.

[Ott98] F. OTTO. Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Rational Mech. Anal.*, 141(1), 63–103, 1998.

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

[PeS22] M. A. PELETIER and A. SCHLICHTING. Cosh gradient systems and tilting. *Preprint*, 2022. arXiv:2203.05435.

[StW98] H. STRUCHTRUP and W. WEISS. Maximum of the local entropy production becomes minimal in stationary processes. *Phys. Rev. Lett.*, 80(23), 5048–5051, 1998.

[Tsc00] N. W. TSCHOEGL. *Fundamentals of Equilibrium and Steady-State Thermodynamics*. Elsevier, 2000.