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

Preprint ISSN 0946 – 8633

Gradient structures and geodesic convexity for reaction-diffusion systems

Matthias Liero[1], Alexander Mielke[1,2]

submitted: April 4, 2012

[1] Weierstraß-Institut
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: Matthias.Liero@wias-berlin.de
Alexander.Mielke@wias-berlin.de

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

No. 1701
Berlin 2012

2010 *Mathematics Subject Classification.* 35K57, 53C21, 53C23, 60J60, 82B35.

Key words and phrases. Geodesic convexity, gradient structures, gradient flow, Onsager operator, reaction-diffusion system, Wasserstein metric, relative entropy.

The research of M.L. was supported by DFG via the MATHEON project D22. The research of A.M. was partially sponsored by ERC-2010-AdG no.267802 *AnaMultiScale*.

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

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We consider systems of reaction-diffusion equations as gradient systems with respect to an entropy functional and a dissipation metric given in terms of a so-called Onsager operator, which is a sum of a diffusion part of Wasserstein type and a reaction part. We provide methods for establishing geodesic λ-convexity of the entropy functional by purely differential methods, thus circumventing arguments from mass transportation. Finally, several examples, including a drift-diffusion system, provide a survey on the applicability of the theory.

1 Introduction

In several papers by Otto (see [JKO98, Ott98, Ott01]) it was shown that certain diffusion problems can be interpreted as gradient flows with respect to the free energy or relative entropy and the Wasserstein metric. In [Mie11c] it was shown that general reaction-diffusion systems, with reactions satisfying the detailed balance condition, can be written as a gradient system with respect to the relative entropy. The associated dissipation metric $\mathcal{G}$ is most easily modeled by considering its inverse $\mathcal{K} = \mathcal{G}^{-1}$, called Onsager operator, as the sum of a diffusion part and a reaction part. The diffusion part is a vector-valued version of the Wasserstein metric used for the scalar Fokker-Planck equation in [JKO98, Ott01], namely

$$\mathcal{K}_{\mathsf{diff}}(\boldsymbol{u})\boldsymbol{\xi} = -\operatorname{div}\big(\mathbb{M}(x,\boldsymbol{u})\nabla\boldsymbol{\xi}\big),$$

where $\boldsymbol{u} = (u_1, ..., u_I) \in [0,\infty[^I$ is the vector of the densities of the species and $\boldsymbol{\xi} = (\xi_1, ..., \xi_I) \in \mathbb{R}^I$ is the associated thermodynamical driving force, also called vector of chemical potentials. Here $\mathbb{M}(x,\boldsymbol{u}) \in \mathrm{Lin}(\mathbb{R}^{I\times d};\mathbb{R}^{I\times d})$ is a general density-dependent mobility tensor, which is symmetric and positive semidefinite. Using a symmetric and positive semidefinite matrix $\mathbb{K}_{\mathrm{react}}(x,\boldsymbol{u}) \in \mathbb{R}^{I\times I}$ we obtain the full Onsager operator

$$\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = -\operatorname{div}\big(\mathbb{M}(x,\boldsymbol{u})\nabla\boldsymbol{\xi}\big) + \mathbb{K}(x,\boldsymbol{u})_{\mathrm{react}}\boldsymbol{\xi},$$

and together with the entropy functional $\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(x,\boldsymbol{u}(x))\,\mathrm{d}x$ giving $\boldsymbol{\xi} = \mathrm{D}\mathcal{E}(\boldsymbol{u}) = \mathrm{D}_{\boldsymbol{u}}E(x,\boldsymbol{u}(x))$ we find the gradient system

$$\dot{\boldsymbol{u}} = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u}) = \operatorname{div}\big(\mathbb{M}(x,\boldsymbol{u})\nabla\mathrm{D}_{\boldsymbol{u}}E(x,\boldsymbol{u})\big) - \mathbb{K}(x,\boldsymbol{u})_{\mathrm{react}}\mathrm{D}_{\boldsymbol{u}}E(x,\boldsymbol{u}),$$

which leads to a large class of reaction-diffusion systems.

The focus of this work is to provide conditions on the system such that the driving functional $\mathcal{E} : X \to \mathbb{R}$ is geodesically λ-convex with respect to the metric $\mathcal{G} = \mathcal{K}^{-1}$. This means that $s \mapsto \mathcal{E}(\gamma(s))$ is λ-convex for all geodesic curves $\gamma : [s_\mathrm{a}, s_\mathrm{b}] \to X$, i.e.,

$$\mathcal{E}(\gamma(s_\theta)) \leq (1-\theta)\mathcal{E}(\gamma(s_0)) + \theta\mathcal{E}(\gamma(s_1)) - \lambda\frac{\theta(1-\theta)}{2}(s_1-s_0)^2 \tag{1.1}$$

for all $\theta \in [0,1]$ and $s_0, s_1 \in [s_\mathrm{a}, s_\mathrm{b}]$, where $s_\theta = (1-\theta)s_0 + \theta s_1$. The study of geodesic λ-convexity for scalar drift-diffusion equations given by

$$\mathcal{E}(u) = \int_\Omega E(u) + uV(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{K}(u)\xi = -\operatorname{div}(\mu(u)\nabla\xi), \tag{1.2}$$

where $u \mapsto E(u)$ is convex and $u \mapsto \mu(u)$ is concave, was initiated in [McC97] and is studied extensively since then, see e.g. [OtW05, AGS05, DaS10, CL*10]. An essential tool in this theory is the characterization of the geodesic curves in terms of mass transportation and the optimal transport problem of Monge-Kantorovich type. Presently, such a method is not available for systems of equations or for scalar equations with reaction terms, which destroy the conservation of mass. Instead, this work relies on a differential characterization of geodesic λ-convexity developed in [DaS08].

Thus, after an introduction of gradient structures for reaction-diffusion systems in Section 2 we provide an abstract version of the theory developed in [DaS08]. We mainly address the abstract framework and present the estimates to obtain concrete convexity properties, while the functional analytic aspects as well as the full framework in terms of complete metric spaces are postponed to subsequent work. Moreover, we assume that our evolutionary system

$$\dot{\boldsymbol{u}} = -\mathcal{F}(\boldsymbol{u}) := -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u}) \tag{1.3}$$

generates a suitable smooth local semiflow on a scale of Banach spaces $Z \subset Y \subset H$ with dense embeddings, see Section 3 for the details. The main characterization involves the quadratic form

$$\boldsymbol{\xi} \mapsto \mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) := \langle \boldsymbol{\xi}, \mathrm{D}\mathcal{F}(\boldsymbol{u})\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}\rangle - \frac{1}{2}\langle \boldsymbol{\xi}, \mathrm{D}\mathcal{K}(\boldsymbol{u})[\mathcal{F}(\boldsymbol{u})]\boldsymbol{\xi}\rangle,$$

which can be seen as the form induced by the metric Hessian of $\mathcal{E}$. The main result is that $\mathcal{E}$ is geodesically λ-convex if the estimate

$$\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) \geq \lambda \langle \boldsymbol{\xi}, \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}\rangle \tag{1.4}$$

holds for all suitable $\boldsymbol{u}$ and $\boldsymbol{\xi}$. Our proof is a straightforward generalization of the approach in [DaS08] that is based on the evolutionary variational inequality (EVI)$_\lambda$ given by

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t} d_\mathcal{K}(u(t), w)^2 + \frac{\lambda}{2} d_\mathcal{K}(u(t), w)^2 + \mathcal{E}(u(t)) \leq \mathcal{E}(w), \quad \forall\, w \in X,\ t > 0, \tag{1.5}$$

where $\frac{\mathrm{d}^+}{\mathrm{d}t} f(t) = \limsup_{\tau \searrow 0} \frac{1}{\tau}(f(t+\tau) - f(t))$ and $d_\mathcal{K}$ is the distance induced by $\mathcal{G} = \mathcal{K}^{-1}$. The idea is to show that (1.3) and (1.4) imply (1.5), and finally deduce (1.1).

Condition (1.4) is closely related to the Bakry-Émery conditions [BaÉ85, Bak94] and provides a strengthened version of the classical entropy-dissipation estimate. In fact, defining $\mathcal{D}(\boldsymbol{u}) = \langle \mathrm{D}\mathcal{E}(\boldsymbol{u}), \mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u})\rangle$ and $\mathcal{R}(\boldsymbol{u}) = 2\mathcal{M}(\boldsymbol{u}, \mathrm{D}\mathcal{E}(\boldsymbol{u}))$ the solutions $\boldsymbol{u}$ of (1.3) satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\boldsymbol{u}(t)) = -\mathcal{D}(\boldsymbol{u}(t)) \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{D}(\boldsymbol{u}(t)) = -\mathcal{R}(\boldsymbol{u}(t)).$$

By (1.4) there exists $\alpha \geq \lambda$ such that $\mathcal{R}(\boldsymbol{u}) - 2\alpha\mathcal{D}(\boldsymbol{u}) = \mathcal{P}(\boldsymbol{u}) \geq 0$ for all $\boldsymbol{u}$. Assuming $\alpha > 0$, in [AM*01] the decay estimates

$$\mathcal{D}(\boldsymbol{u}(t)) \leq \mathrm{e}^{-2\alpha t}\mathcal{D}(\boldsymbol{u}(0)) \quad \text{and} \quad \mathcal{E}(\boldsymbol{u}(t)) - \mathcal{E}(\boldsymbol{u}(\infty)) + \int_t^\infty \mathcal{P}(\boldsymbol{u}(s))\,\mathrm{d}s = \frac{1}{2\alpha}\mathcal{D}(\boldsymbol{u}(t))$$

are used to derive convergence for $t \to \infty$. We discuss further useful properties of the geodesic λ-convexity in Section 3.3, also if $\lambda < 0$.

The main part of this work surveys possible applications of the abstract theory, see Section 4. We emphasize that geodesic convexity is a strong structural property of a gradient system that is rather difficult to achieve, in particular with respect to distances that are associated with the Wasserstein metric. Our examples show that there are at least some nontrivial reaction-diffusion equations or systems that satisfy this beautiful property. First we discuss simple reaction kinetics satisfying the detailed balance conditions, i.e. ODE systems in the form

$$\dot{\boldsymbol{u}} = -F(\boldsymbol{u}) = -\mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u}), \quad \text{where } E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i).$$

This includes the case of general reversible Markov chains $\dot{\boldsymbol{u}} = Q\boldsymbol{u}$, where $Q \in \mathbb{R}^{I\times I}$ is a stochastic generator, see also [Maa11, Mie11b, ErM11].

In the following sections we treat partial differential equations or systems where estimate (1.4) heavily relies on a well-chosen sequence of integrations by parts, where the occurring boundary integrals needs to be taken care of. We use the fact that for convex domains Ω and functions $\xi \in \mathrm{H}^2(\Omega)$ with $\nabla\xi \cdot \nu = 0$ on $\partial\Omega$, we have $\nabla\big(|\nabla\xi|^2\big) \cdot \nu \leq 0$ on $\partial\Omega$, Proposition 4.2. In Section 4.2 we give a lower bound for the geodesic convexity of $\mathcal{E}(u) = \int_\Omega u \log u \, \mathrm{d}x$ with respect to the inhomogeneous Wasserstein metric induced by $\mathcal{K}(u)\xi = -\operatorname{div}(\mu(x)u\nabla\xi)$, where $0 < \mu_0 \leq \mu \in \mathrm{W}^{2,\infty}(\Omega)$, thus generalizing results in [Lis09]. Theorem 4.3 provides a new result of geodesic convexity for $\mathcal{E}$ and $\mathcal{K}$ from (1.2), where the concave mobility $u \mapsto \mu(u)$ is allowed to be decreasing, i.e. $\mu'(u) < 0$, thus complementing results in [CL*10].

Sections 4.4 and 4.5 discuss problems with reactions, namely

$$\dot{u} = \Delta u - f(u) \quad \text{and} \quad \begin{pmatrix} \dot{u}_1 \\ \dot{u}_2 \end{pmatrix} = \begin{pmatrix} \delta\Delta u_1 \\ \delta\Delta u_2 \end{pmatrix} + k\begin{pmatrix} u_2 - u_1 \\ u_1 - u_2 \end{pmatrix}.$$

The first case with $f(u) = k(1-u)$ gives geodesic λ-convexity with $\lambda = \frac{1}{2}\min\{k, k^2\}$, while the second case gives geodesic 0-convexity. In Section 4.6 a one-dimensional drift-diffusion system with charged species is considered, where the nonlinear coupling occurs via the electrostatic potential. The final example discusses cross-diffusion of Stefan-Maxwell type for $\boldsymbol{u} = (u_1, ..., u_I)$ under the size-exclusion condition $u_1 + \cdots + u_I \equiv 1$.

There are further interesting applications of gradient flows where methods based on geodesic convexity can be employed, even though the system under investigation may not be geodesically λ-convex, see e.g. the fourth order problems studied in [MMS09, GST09, CL*10]. Possible applications to viscoelasticity are discussed in [MOS12]. In [FiG10] a diffusion equation with Dirichlet boundary conditions, which leads to absorption, is investigated.

2 Gradient structures for reaction-diffusion systems

In this section we give some general background on gradient and Onsager systems. All our arguments are formal and will be made precise in the following sections.

A gradient system is a triple $(X, \mathcal{E}, \mathcal{G})$ where X is the state space containing the states $\boldsymbol{u} \in X$. For simplicity we assume that X is a reflexive Banach space with dual X^*. The driving functional

$\mathcal{E} : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ is assumed to be differentiable (in a suitable way) such that the potential restoring force is given by $-\mathrm{D}\mathcal{E}(\boldsymbol{u}) \in X^*$. The third ingredient is a metric tensor $\mathcal{G}$, i.e. $\mathcal{G}(\boldsymbol{u}) : X \to X^*$ is linear, symmetric and positive (semi-)definite.

The gradient flow associated with $(X, \mathcal{E}, \mathcal{G})$ is the (abstract) force balance

$$\mathcal{G}(\boldsymbol{u})\dot{\boldsymbol{u}} = -\mathrm{D}\mathcal{E}(\boldsymbol{u}) \qquad \Longleftrightarrow \qquad \dot{\boldsymbol{u}} = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u}) := -\nabla_{\mathcal{G}}\mathcal{E}(\boldsymbol{u}), \tag{2.1}$$

where we recall that the "gradient" $\nabla_{\mathcal{G}}\mathcal{E}$ of the functional $\mathcal{E}$ is an element of X (in contrast to the differential $\mathrm{D}\mathcal{E}(\boldsymbol{u}) \in X^*$) and is calculated in terms of $\mathcal{K}(\boldsymbol{u}) = \mathcal{G}(\boldsymbol{u})^{-1}$. We call this equation an abstract force balance, since $\mathcal{G}(\boldsymbol{u})\dot{\boldsymbol{u}}$ can be seen as a viscous force arising from the motion of $\boldsymbol{u}$. We call the linear, symmetric and positive semidefinite operator $\mathcal{K}(\boldsymbol{u}) : X^* \to X$ the Onsager operator and the corresponding triple $(X, \mathcal{E}, \mathcal{K})$ Onsager system.

Since we are mainly interested in reaction-diffusion systems we consider densities $\boldsymbol{u} : \Omega \to]0, \infty[^I$ of diffusive species $X_1, \ldots, X_I$. The driving functional of the evolution $\mathcal{E}$ is of the form

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(x, \boldsymbol{u}(x))\,\mathrm{d}x,$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain and E is a sufficiently smooth energy density. It was shown in [Mie11c] that for a wide class of reaction-diffusion systems gradient/Onsager structures can be specified. The central point is that in the Onsager form we have an additive splitting of the Onsager operator into a diffusive and a reaction part, namely $\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = \mathcal{K}_{\text{diff}}(\boldsymbol{u})\boldsymbol{\xi} + \mathcal{K}_{\text{react}}(\boldsymbol{u})\boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is the thermodynamically conjugated force being dual to the rate $\dot{\boldsymbol{u}}$. We define the diffusion part $\mathcal{K}_{\text{diff}}$, following the Wasserstein approach to diffusion introduced by Otto in [JKO98, Ott01], and the reaction part $\mathcal{K}_{\text{react}}$ as follows:

$$\mathcal{K}_{\text{diff}}(\boldsymbol{u})\boldsymbol{\xi} = -\operatorname{div}\big(\mathbb{M}(\boldsymbol{u})\nabla\boldsymbol{\xi}\big), \quad \text{and} \quad \mathcal{K}_{\text{react}}(\boldsymbol{u})\boldsymbol{\xi} = \mathbb{K}(\boldsymbol{u})\boldsymbol{\xi}. \tag{2.2}$$

Here $\mathbb{M}(x, \boldsymbol{u}) : \mathbb{R}^{I\times d} \to \mathbb{R}^{I\times d}$ and $\mathbb{K}(x, \boldsymbol{u}) \in \mathbb{R}^{I\times I}$ are symmetric, positive semidefinite tensors of order four and two, respectively. The evolution is described by

$$\dot{\boldsymbol{u}} = -\operatorname{div}\Big(\mathbb{M}(x, \boldsymbol{u})\nabla\big(\partial_{\boldsymbol{u}}E(x, \boldsymbol{u})\big)\Big) + \mathbb{K}(x, \boldsymbol{u})\partial_{\boldsymbol{u}}E(x, \boldsymbol{u}), \tag{2.3}$$

subjected to the no-flux boundary condition $\mathbb{M}(x, u)\nabla\big(\partial_{\boldsymbol{u}}E(x, \boldsymbol{u})\big) \cdot \nu(x) = 0$ for $x \in \partial\Omega$.

The symmetry of the tensor $\mathcal{K}(\boldsymbol{u})$ allows us to define the dual dissipation potential

$$\Psi^*(\boldsymbol{u}; \boldsymbol{\xi}) = \frac{1}{2}\langle\boldsymbol{\xi}, \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \frac{1}{2}\int_\Omega \nabla\boldsymbol{\xi}\cdot\mathbb{M}(\boldsymbol{u})\nabla\boldsymbol{\xi} + \boldsymbol{\xi}\cdot\mathbb{K}(\boldsymbol{u})\boldsymbol{\xi}\,\mathrm{d}x.$$

We call Ψ^* the dual dissipation potential since it is the Legendre transform of the dissipation potential $\Psi : (\boldsymbol{u}, \dot{\boldsymbol{u}}) \mapsto \frac{1}{2}\langle\mathcal{G}(\boldsymbol{u})\dot{\boldsymbol{u}}, \dot{\boldsymbol{u}}\rangle$, i.e., we have

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

In the following subsections we will specify the structure of the functional $\mathcal{E}$ and the Onsager operator $\mathcal{K}$ and present some illustrative examples.

2.1 Chemical reaction kinetics of mass-action type

Pure chemical reaction systems are ODE systems $\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u})$, where often the right-hand side is written in terms of polynomials associated to the reaction kinetics. It was observed in [Mie11c] that under the assumption of detailed balance (also called reversibility) such systems have a gradient structure with the relative entropy

$$E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i)$$

as the driving functional, where $w_i > 0$ denotes fixed reference densities. We assume that there are R reactions of mass-action type (cf. e.g. [DeM84, GiM04, KjB08]) between the species $X_1, \ldots, X_I$ denoted by

$$\alpha_1^r X_1 + \cdots + \alpha_I^r X_I \underset{k_r^{\mathsf{bw}}}{\overset{k_r^{\mathsf{fw}}}{\rightleftharpoons}} \beta_1^r X_1 + \cdots + \beta_I^r X_I \qquad r = 1, \ldots, R,$$

where k_r^{bw} and k_r^{fw} are the backward and forward reaction rates, and the vectors $\boldsymbol{\alpha}^r,\ \boldsymbol{\beta}^r \in \mathbb{N}_0^I$ contain the stoichiometric coefficients of the rth reaction.

The associated reaction system for the densities (in a spatially homogeneous system, where diffusion can be neglected) reads

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) := -\sum_{r=1}^{R} \Big(k_r^{\mathsf{fw}}(\boldsymbol{u})\, \boldsymbol{u}^{\boldsymbol{\alpha}^r} - k_r^{\mathsf{bw}}(\boldsymbol{u})\, \boldsymbol{u}^{\boldsymbol{\beta}^r}\Big)\Big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\Big), \tag{2.4}$$

where we use the monomial notation $\boldsymbol{u}^{\boldsymbol{\alpha}} = u_1^{\alpha_1} \cdots u_I^{\alpha_I}$.

The main assumption to obtain a gradient structure is that of *detailed balance*, which means that there exists a reference density vector $\boldsymbol{w}$ such that all R reactions are balanced individually, namely

$$\exists\, \boldsymbol{w} \in\,]0,\infty[^I\ \ \forall\, r = 1, \ldots, R\, \forall\, \boldsymbol{u} \in\,]0,\infty[^I :\ k_r^{\mathsf{fw}}(\boldsymbol{u})\boldsymbol{w}^{\boldsymbol{\alpha}^r} = k_r^{\mathsf{bw}}(\boldsymbol{u})\boldsymbol{w}^{\boldsymbol{\beta}^r} =: k_r^*(\boldsymbol{u}). \tag{2.5}$$

Here we have used the freedom to allow for reaction coefficients depending on the densities.

As in [Mie11c] we now define the Onsager matrix

$$\mathbb{K}(\boldsymbol{u}) = \sum_{r=1}^{R} k_r^*(\boldsymbol{u}) \Lambda\big(\tfrac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}}, \tfrac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}\big)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \otimes \big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \text{ with } \Lambda(a,b) = \frac{a-b}{\log a - \log b} \tag{2.6}$$

and find that the reaction system (2.4) takes the form

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) = -\mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u}). \tag{2.7}$$

This follows easily by using the definition of Λ and the rules for logarithms, namely

$$\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \cdot \big(\,\mathbf{log}\,\boldsymbol{u} - \mathbf{log}\,\boldsymbol{w}\big) = \log\big(\boldsymbol{u}^{\boldsymbol{\alpha}^r}/\boldsymbol{w}^{\boldsymbol{\alpha}^r}\big) - \log\big(\boldsymbol{u}^{\boldsymbol{\beta}^r}/\boldsymbol{w}^{\boldsymbol{\beta}^r}\big).$$

2.2 Coupling diffusion and reaction

Now we consider coupled reaction-diffusion systems. The driving functional for the evolution is the total relative entropy $\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(\boldsymbol{u})\, \mathrm{d}x$. The Onsager operator is given by the sum $\mathcal{K}(\boldsymbol{u}) =$

$\mathcal{K}_{\text{diff}}(\boldsymbol{u}) + \mathcal{K}_{\text{react}}(\boldsymbol{u})$ with $\mathcal{K}_{\text{diff}}$ and $\mathcal{K}_{\text{react}}$ as in (2.2). Hence, with $\mathbb{K}$ given in (2.6) the coupled system reads

$$\dot{\boldsymbol{u}} = \operatorname{div}\Big(\mathbb{M}(\boldsymbol{u})\nabla\big(\log \boldsymbol{u} - \log \boldsymbol{w}\big)\Big) + \mathbb{K}(\boldsymbol{u})\big(\log \boldsymbol{u} - \log \boldsymbol{w}\big) = \operatorname{div}\big(\widetilde{\mathbb{M}}(\boldsymbol{u})\nabla \boldsymbol{u}\big) + \boldsymbol{R}(\boldsymbol{u}),$$

where $\mathbb{M}(\boldsymbol{u}) = \widetilde{\mathbb{M}}(\boldsymbol{u})\operatorname{diag}(\boldsymbol{u})$.

As an example for a reaction-diffusion system we consider the quaternary system studied in [DF*07, DeF08], namely the evolution of a mixture of diffusive species X_1, X_2, X_3 and X_4 in a bounded domain Ω undergoing a reversible reaction of the type

$$X_1 + X_2 \underset{k^{\mathsf{bw}}}{\overset{k^{\mathsf{fw}}}{\rightleftharpoons}} X_3 + X_4. \tag{2.8}$$

For the density vector $\boldsymbol{u} = (u_1, u_2, u_3, u_4)$ we introduce the free energy functional

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega \sum_{i=1}^4 u_i \log(u_i/w_i)\,\mathrm{d}x.$$

For simplicity we assume that $k^{\mathsf{fw}} = k^{\mathsf{bw}} = 1$ and can take $w_i = 1$. We have the stoichiometric vectors $\boldsymbol{\alpha} = (1,1,0,0)$, $\boldsymbol{\beta} = (0,0,1,1)$ and thus

$$\mathbb{K}(u_1, u_2, u_3, u_4) = \Lambda\Big(u_1u_2\,,\, u_3u_4\Big)\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix}.$$

The tensor $\mathbb{M}(\boldsymbol{u}) = \operatorname{diag}(\delta_1 u_1, \dots, \delta_4 u_4)$ and the corresponding Onsager operator $\mathcal{K}_{\text{diff}}(\boldsymbol{u})$ leads to the reaction-diffusion system

$$\dot{\boldsymbol{u}} = \operatorname{div}\big(\mathbb{D}\nabla\boldsymbol{u}\big) - (u_1u_2 - u_3u_4)(\boldsymbol{\alpha} - \boldsymbol{\beta}), \quad \text{where } \mathbb{D} = \operatorname{diag}(\delta_1, \dots, \delta_4).$$

In fact, many reaction-diffusion systems studied in the literature (including semiconductor models involving an elliptic equation for the electrostatic potential), see e.g. [GlH05, DeF06, DeF07, Gli09, BoP11], have the structure developed above. Except for the recent work [Mie11c, GlM12, Mie12] the gradient structure was not displayed and exploited explicitly, only the Liapunov property of the free energy $\mathcal{E}$ was exploited for deriving a priori estimates.

2.3 Non-isothermal coupled systems

We now extend the system from the previous section and consider the non-isothermal case when the temperature of the system is not constant but an independent field coupled to the densities $\boldsymbol{u}$. For such systems we have two functionals, namely the total energy, which is preserved during the evolution of the system, and the total entropy, which acts as the driving force. Instead of using the temperature $\theta : \Omega \to \mathbb{R}$ as additional variable it has certain advantages to use the internal energy $e : \Omega \to \mathbb{R}$ as free variable (see [Mie12, Mie11a] for details). Thus, the functionals are

$$\mathcal{E}(\boldsymbol{u}, e) = \int_\Omega e(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{S}(\boldsymbol{u}, e) = \int_\Omega S(x, \boldsymbol{u}(x), e(x))\,\mathrm{d}x.$$

Now the Gibbs relation leads to the definition of the temperature as $\theta = \Theta(\boldsymbol{u}, e) := 1/\partial_e S(\boldsymbol{u}, e)$, where the relation $\partial_e S(\boldsymbol{u}, e) > 0$ is imposed.

The major advantage of the formulation in terms of $(\boldsymbol{u}, e)$ is that energy conservation is a linear constraint. Moreover, following [AGH02] it is reasonable to assume that S is a concave function in $(\boldsymbol{u}, e)$. The equations can be written in $(\boldsymbol{u}, e)$ using the Onsager operator

$$\mathcal{K}(\boldsymbol{u}, e)\begin{pmatrix}\boldsymbol{\xi} \\ \varepsilon\end{pmatrix} = -\operatorname{div}\left(\mathbb{M}(\boldsymbol{u}, e)\nabla\begin{pmatrix}\boldsymbol{\xi} \\ \varepsilon\end{pmatrix}\right) + \mathbb{K}(\boldsymbol{u}, e)\boldsymbol{\xi},$$

where the fourth order tensor $\mathbb{M}(\boldsymbol{u}, e)$ has the block structure

$$\mathbb{M}(\boldsymbol{u}, e) = \begin{pmatrix}\mathbb{M}_{\boldsymbol{u}\boldsymbol{u}}(\boldsymbol{u}, e) & \mathbb{M}_{\boldsymbol{u}e}(\boldsymbol{u}, e) \\ \mathbb{M}_{\boldsymbol{u}e}(\boldsymbol{u}, e) & \mathbb{M}_{ee}(\boldsymbol{u}, e)\end{pmatrix},$$

where $\mathbb{M}_{\boldsymbol{u}\boldsymbol{u}}(\boldsymbol{u}, e), \mathbb{M}_{\boldsymbol{u}e}(\boldsymbol{u}, e)$ and $\mathbb{M}_{ee}(\boldsymbol{u}, e)$ are symmetric and positive semidefinite.

The evolution equations for $(\boldsymbol{u}, e)$ take the form

$$\begin{aligned}\dot{\boldsymbol{u}} &= -\operatorname{div}\Big(\mathbb{M}_{\boldsymbol{u}\boldsymbol{u}}(\boldsymbol{u}, e)\nabla\big(\partial_{\boldsymbol{u}} S(\boldsymbol{u}, e)\big) + \mathbb{M}_{\boldsymbol{u}e}(\boldsymbol{u}, e)\nabla\big(\partial_e S(\boldsymbol{u}, e)\big)\Big) + \mathbb{K}(\boldsymbol{u}, e)\partial_{\boldsymbol{u}} S(\boldsymbol{u}, e), \\ \dot{e} &= -\operatorname{div}\Big(\mathbb{M}^*_{\boldsymbol{u}e}(\boldsymbol{u}, e)\nabla\big(\partial_{\boldsymbol{u}} S(\boldsymbol{u}, e)\big) + \mathbb{M}_{ee}(\boldsymbol{u}, e)\nabla\big(\partial_e S(\boldsymbol{u}, e)\big)\Big).\end{aligned}$$

This form has the major advantage that we can read of “parabolicity” in the sense of Petrovsky (cf. [LSU68, Sect. VII.8]) for the full coupled system by assuming that $\mathbb{M}$ is positive definite and that $\mathrm{D}^2 S$ is negative definite. Hence, local existence results can be obtained from [Ama93].

Moreover, we are able to postulate suitable strongly coupled models by assuming that S has the form

$$S(\boldsymbol{u}, e) = s(e) - \boldsymbol{u} \cdot \big(\mathbf{log}\,\boldsymbol{u} - \mathbf{log}\,\boldsymbol{w}(e)\big), \tag{2.9}$$

where $\boldsymbol{w}(e)$ is the vector of reference densities in the detailed balance condition (2.5), which may now depend on the internal energy (i.e. on the temperature).

2.4 Drift-reaction-diffusion equations

We close this section by considering a drift-diffusion system coming from the theory of semiconductor devices. More precisely, we treat the simplest semiconductor model, namely the van Roosbroeck system. Here, we additionally need to take into account that the electric charge of the species generates an electric potential, whose electric field creates drift forces proportional to the charges of the species. We recite here briefly the results of [Mie11c, Sect. 4] and refer to latter for the full discussion. Moreover, we refer to [GIM12] for drift-diffusion systems exhibiting bulk-interface interaction.

The system's state is described by the electron and hole densities $n : \Omega \to\,]0, \infty[$ and $p : \Omega \to\,]0, \infty[$, respectively. The charged species generate an electrostatic potential $\phi_{n,p}$ being the unique solution of the linear potential equation

$$-\operatorname{div}(\varepsilon\nabla\phi) = \delta + q_n n + q_p p \text{ in } \Omega, \qquad \phi = \phi_{\mathrm{Dir}} \text{ on } \Gamma_{\mathrm{Dir}} \subset \partial\Omega, \tag{2.10a}$$

where $\delta : \Omega \to \mathbb{R}$ is a given doping profile and $q_n = -1$ and $q_p = 1$ are the charge numbers with opposite sign. The evolution of the densities n, p is governed by diffusion, drift according to the electric field $\nabla\phi_{n,p}$, and recombination according to simple creation-annihilation reactions for electron-hole pairs, namely

$$X_n + X_p \rightleftharpoons \emptyset, \quad \text{i.e. } \boldsymbol{\alpha} = \begin{pmatrix}1\\1\end{pmatrix} \text{ and } \boldsymbol{\beta} = \begin{pmatrix}0\\0\end{pmatrix}.$$

With mobilities $\mu_n(n,p), \mu_p(n,p) > 0$ and reaction rate $\kappa(n,p) > 0$ the drift-diffusion system reads

$$\begin{aligned}\dot n &= \operatorname{div}\big(\mu_n(n,p)(\nabla n + q_n n \nabla\phi_{n,p})\big) - \kappa(n,p)(np{-}1),\\ \dot p &= \operatorname{div}\big(\mu_p(n,p)(\nabla p + q_p p \nabla\phi_{n,p})\big) - \kappa(n,p)(np{-}1).\end{aligned} \tag{2.10b}$$

For establishing a gradient structure we define the functional $\mathcal{E}$ as the sum of electrostatic and free energy:

$$\mathcal{E}(n,p) = \int_\Omega \frac{1}{2}|\nabla\phi_{n,p}|^2 + n(\log n - 1) + p(\log p - 1)\,\mathrm{d}x.$$

The thermodynamic conjugated forces, also called quasi-Fermi potentials, read $\mathrm{D}_n\mathcal{E}(n,p) = \log n + q_n\phi_{n,p}$ and $\mathrm{D}_p\mathcal{E}(n,p) = \log p + q_p\phi_{n,p}$. Here we used that $\phi_{n,p}$ solves (2.10a) and depends linearly on n and p. The Onsager operator $\mathcal{K}(n,p)$ takes the form

$$\mathcal{K}(n,p)\begin{pmatrix}\xi_n\\ \xi_p\end{pmatrix} = \begin{pmatrix}-\operatorname{div}(\mu_n n\nabla\xi_n)\\ -\operatorname{div}\mu_p(p\nabla\xi_p)\end{pmatrix} + \kappa(n,p)\Lambda(np,1)\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}\xi_n\\ \xi_p\end{pmatrix}.$$

Thus again we have two Wasserstein terms for the electrochemical potentials coupled with a reaction term. We immediately find that for $q_n = -q_p$ (opposite charges of electron and holes) it holds that $\binom{q_n}{q_p} \in \operatorname{Ker}\mathcal{K}(n,p)$. This means, that the total charge $\mathcal{Q}(n,p) = \int_\Omega \delta{+}q_n n{+}q_p p\,\mathrm{d}x$ is a conserved quantity, i.e., $\frac{\mathrm{d}\mathcal{Q}(n,p)}{\mathrm{d}t} = 0$. Moreover, using that

$$-\mathcal{K}(n,p)\mathrm{D}\mathcal{E}(n,p) = \begin{pmatrix}\operatorname{div}\big(\mu_n n\nabla(\log n + q_n\phi_{n,p})\big) - \kappa\Lambda(np,1)\log(np)\\ \operatorname{div}\big(\mu_p p\nabla(\log p + q_p\phi_{n,p})\big) - \kappa\Lambda(np,1)\log(np)\end{pmatrix}$$

we see that $\binom{\dot n}{\dot p} = -\mathcal{K}(n,p)\mathrm{D}\mathcal{E}(n,p)$ is the desired Onsager structure of the van Roosbroeck system (2.10).

A similar gradient system with only one species was considered in [AmS08], namely

$$\dot u = \operatorname{div}\big(u\nabla\Phi_u\big),\ -\Delta\Phi_u + \Phi_u = u \text{ in } \Omega, \quad \nabla u\cdot\nu = 0,\ \Phi = 1 \text{ on } \partial\Omega.$$

It is a gradient system for the energy $\mathcal{E}(u) = \int_\Omega u + \frac{1}{2}|\nabla\Phi_u|^2 + \frac{1}{2}|\Phi_u{-}1|^2\,\mathrm{d}x$ and the Wasserstein operator $\mathcal{K}(u)\xi = -\operatorname{div}(u\nabla\xi)$.

2.5 On the metric induced by reaction and diffusion

As we have seen above it is most natural to model reaction-diffusion systems in terms of the Onsager operator. Hence, we will formulate the convexity conditions in terms of $\mathcal{E}$, $\mathcal{K}$, and the vector field $\mathcal{F}$. However, from the mathematical point of view the metric $\mathcal{G} = \mathcal{K}^{-1}$ and the induced distance $d_\mathcal{K}$ are important as well. Following the famous Benamou-Brenier formulation [BeB00] we can characterize our $\mathcal{G}$ in a similar fashion

$$\begin{aligned}\langle\mathcal{G}(\boldsymbol{u})\boldsymbol{v},\boldsymbol{v}\rangle = \inf\Big\{ \int_\Omega \boldsymbol{\Xi}{:}\mathbb{M}(\boldsymbol{u})\boldsymbol{\Xi} + \boldsymbol{\xi}\cdot\mathbb{K}(\boldsymbol{u})\boldsymbol{\xi}\,\mathrm{d}x \;\Big|\; &\boldsymbol{\Xi}\in\mathrm{L}^2(\Omega;\mathbb{R}^{I\times d}),\\ &\boldsymbol{\xi}\in\mathrm{L}^2(\Omega;\mathbb{R}^I),\ \boldsymbol{\xi} + \operatorname{div}\boldsymbol{\Xi} = \boldsymbol{v}\Big\}.\end{aligned} \tag{2.11}$$

In particular, concavity of the tensors $\mathbb{M}$ and $\mathbb{K}$ (i.e. for all $\boldsymbol{\xi}$ the mapping $\boldsymbol{u} \mapsto \boldsymbol{\xi}\cdot\mathbb{K}(\boldsymbol{u})\boldsymbol{\xi}$ is concave) we find that $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \langle \mathcal{G}(\boldsymbol{u})\boldsymbol{v}, \boldsymbol{v}\rangle$ is convex, which can be used to establish the existence of geodesic curves.

3 Geodesically λ-convex gradient systems

3.1 Abstract setup

In this section we provide an abstract formulation such that the theory of [DaS08] can be applied to general systems $(X, \mathcal{E}, \mathcal{K})$, in particular to systems of partial differential equations, where $\mathcal{K}$ is allowed to be a partial differential operator as well. The main point of [DaS08] is that it is sufficient to establish the geodesic λ-convexity of $\mathcal{E}$ on a dense set, where all the calculations on functions can be done rigorously. Then the abstract theory allows us to extend the geodesic λ-convexity of $\mathcal{E}$ to the closure of the domain of $\mathcal{E}$.

We consider a set $\mathcal{X}$ which is a closed subset of a Banach space X, e.g. vectors of Radon measures. For the smooth solutions and their velocities we need smaller spaces

$$Z \subset Y \subset X$$

with dense and continuous embeddings. For $u \in Y$ the norm induced by the metric $\mathcal{G}(u)$ will be equivalent to that of the Hilbert space H, for which we assume

$$Y \subset H \text{ with dense and continuous embedding.}$$

We assume that open and connected sets $\mathcal{Z} \subset Z$ and $\mathcal{Y} \subset Y$ exist such that

$$\mathcal{Z} \subset Z \cap \mathcal{X}, \quad \mathcal{Z} \subset \mathcal{Y} \subset Y \cap \mathcal{X}, \text{ and } \mathcal{Z} \text{ is dense in } \mathcal{X}.$$

We consider a gradient systems $(\mathcal{Z}, \mathcal{E}, \mathcal{K})$ satisfying

$$\mathcal{E} \in \mathrm{C}^2(\mathcal{Z};\mathbb{R}), \quad \mathcal{K} \in \mathrm{C}^1(\mathcal{Y};\mathrm{Lin}(H^*;H)), \quad \mathcal{G} = \mathcal{K}^{-1} \in \mathrm{C}^1(\mathcal{Y};\mathrm{Lin}(H;H^*)). \tag{3.1}$$

Thus, the evolution reads

$$\dot{u} = -\mathcal{F}(u) = -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u),$$

where, having in mind PDEs, we assume the smoothness of the vector field $\mathcal{F}$

$$\mathcal{F} \in \mathrm{C}^1(\mathcal{Z};Y) \text{ and } \mathrm{D}\mathcal{F} \in \mathrm{C}^0(\mathcal{Z};\mathrm{Lin}(Z;Y)) \cap \mathrm{C}^0(\mathcal{Z};\mathrm{Lin}(Y;H)), \tag{3.2}$$

which is what one would obtain composing the smoothness of $\mathcal{K}$ and $\mathcal{E}$.

3.2 Geodesic curves and geodesic λ-convexity

The metric tensor $\mathcal{G} = \mathcal{K}^{-1}$ generates a distance $d_\mathcal{K} : X \times X \to [0,\infty]$ in the usual way: For $u_0, u_1 \in X$ we define the set of connecting curves via

$$\boldsymbol{C}(u_0,u_1) = \{\, \gamma \in \mathrm{C}^1([0,1];X) \mid \gamma(0) = u_0,\ \gamma(1) = u_1 \,\}.$$

This allows us to define the distance $d_{\mathcal{K}}$ as follows

$$d_{\mathcal{K}}(u_0,u_1)^2 = \inf\{\, J_{\mathcal{K}}(\gamma) \mid \gamma \in \boldsymbol{C}(u_0,u_1)\,\} \text{ with } J_{\mathcal{K}}(\gamma) = \int_0^1 \boldsymbol{\Gamma}(\gamma(s),\gamma'(s))\,\mathrm{d}s. \qquad (3.3)$$

Here γ' denotes the derivative with respect to the arclength parameter s, and

$$\boldsymbol{\Gamma}(u,v) = \langle \mathcal{G}(u)v, v\rangle \quad \text{if } (u,v)\in \mathcal{Y}\times H \quad \text{and } +\infty \text{ else.}$$

It is easy to see that $d_{\mathcal{K}}$ is symmetric and satisfies the triangle inequality. We assume positivity, i.e.,

$$\forall\, u,w \in \mathcal{Z}: \quad u \neq w \implies d_{\mathcal{K}}(u,w) > 0. \qquad (3.4)$$

Thus, we may consider also the metric gradient system $(X,\mathcal{E},d_{\mathcal{K}})$ in the sense of [AGS05]. We refer to this work or to [CL*10] for distances $d_{\mathcal{K}}$ in more general cases. As in any metric space $(X,d_{\mathcal{K}})$, a geodesic curve connecting u_0 and u_1 is a curve $\gamma\in\boldsymbol{C}(u_0,u_1)$ satisfying

$$\forall\, r,s\in[0,1]: \quad d_{\mathcal{K}}(\gamma(r),\gamma(s)) = |r{-}s|\, d_{\mathcal{K}}(u_0,u_1). \qquad (3.5)$$

Remark 3.1 *If $\mathcal{Y}$ is a convex subset of $Y\subset X$ and $\mathcal{Y}\ni u\mapsto \langle \eta,\mathcal{K}(u)\eta\rangle$ is concave for all η, then $(u,v)\mapsto\langle\mathcal{G}(u)v,v\rangle$ is (jointly) convex on $\mathcal{Y}\times H$. As a consequence the functional $J_{\mathcal{K}}$ in* (3.3) *and hence $d_{\mathcal{K}}^2:\mathcal{Y}\times\mathcal{Y}\to[0,\infty[$ is convex as well.*

Remark 3.2 *Only in very few cases $d_{\mathcal{K}}$ can be calculated explicitly, all relying on the Wasserstein distance d_{Wass}, see [AGS05, Vil09]. For constants $\mu>0$ and $\kappa\geq 0$ consider $\mathcal{K}_{\mu,\kappa}(u) = -\operatorname{div}(\mu u\nabla\xi)+\kappa u\xi$, which is affine in u. For $\kappa=0$ we have, on the set $X=\{\,u\in\mathrm{M}(\Omega)\mid u\geq 0\,\}$ of nonnegative Radon measures, the distance*

$$d_{\mathcal{K}_{\mu,0}}(u_1,u_2) = \sqrt{\mathrm{vol}(u_1)/\mu}\; d_{\mathrm{Wass}}\big(\tfrac{1}{\mathrm{vol}(u_1)}u_1, \tfrac{1}{\mathrm{vol}(u_1)}u_2\big) \quad \textit{if } \mathrm{vol}(u_1)=\mathrm{vol}(u_2) \textit{ and } +\infty \textit{ else.}$$

For κ the Onsager operator $\mathcal{K}_{\mu,0}$ is mass preserving, hence X decomposes into the components $X_\alpha = \{\,u\in X \mid \mathrm{vol}(u)=\alpha\,\}$. For $\mu=0$ there is no spatial interaction, and we find the explicit formula $d_{\mathcal{K}_{0,\kappa}}(u_1,u_2) = \sqrt{4/\kappa}\,\|\sqrt{u_1}{-}\sqrt{u_2}\|_{\mathrm{L}^2(\Omega)}$. For $\mu,\kappa>0$ there are geodesic curves between all points of X, and we conjecture the formula

$$d_{\mathcal{K}_{\mu,\kappa}}(u_1,u_2)^2 = \sup\Big\{ \int_\Omega \eta(0,x)u_1(\mathrm{d}x) - \int_\Omega \eta(1,x)u_2(\mathrm{d}x) \Big| \dot\eta + \tfrac{\mu}{2}|\nabla\eta|^2 \leq \tfrac{\kappa}{2}\eta^2 \Big\}.$$

This and other characterizations of reaction-diffusion distances will be investigated in subsequent work.

For a given $\lambda\in\mathbb{R}$, a functional $\mathcal{E}$ is called *geodesically λ-convex* with respect to the metric $d_{\mathcal{K}}$ if for all geodesics $\gamma:[s_\mathrm{a},s_\mathrm{b}]\to X$ the map $s\to\mathcal{E}(\gamma(s))$ is λ-convex, i.e.

$$\mathcal{E}(\gamma(s_\theta)) \leq (1{-}\theta)\mathcal{E}(\gamma(s_0)) + \theta\mathcal{E}(\gamma(s_1)) - \lambda\frac{\theta(1{-}\theta)}{2}(s_0{-}s_1)^2 \qquad (3.6)$$

for all $\theta\in[0,1]$ and $s_0,s_1\in[s_\mathrm{a},s_\mathrm{b}]$, where $s_\theta = (1{-}\theta)s_0+\theta s_1$.

3.3 Properties of geodesically λ-convex gradient flows

In this section we collect some useful properties of geodesically λ-convex systems. We refer to [AGS05] for the full discussion. First, we have a Lipschitz continuous dependence of the solutions u_j, $j = 1, 2$, on the initial data (also see [DaS08]), namely

$$\text{for all } t \geq 0: \qquad d_\mathcal{K}(u_1(t), u_2(t)) \leq \mathrm{e}^{-\lambda t} d_\mathcal{K}(u_1(0), u_2(0)). \tag{3.7}$$

In particular, for $\lambda \geq 0$ we have a contraction semigroup. If $\lambda > 0$ we obtain exponential decay towards the unique equilibrium state u_*, which minimizes $\mathcal{E}$, i.e.

$$d_\mathcal{K}(u(t), u_*) \leq \mathrm{e}^{-\lambda t} d_\mathcal{K}(u(0), u_*).$$

Second, the time-continuous solutions $u : [0, \infty[\to X$ can be well approximated by interpolants obtained by incremental minimizations: Fixing a time step $\tau > 0$ we define iteratively

$$u^\tau_{k+1} \in \operatorname*{Arg\,min}_{w \in X} \Big(\mathcal{E}(w) + \tfrac{1}{2\tau} d_\mathcal{K}(u^\tau_k, w)^2 \Big).$$

For geodesically λ-convex functionals $\mathcal{E}$ the minimizers are unique for any $\tau \in]0, \tau_0[$ if $1/\tau_0 + \lambda \geq 0$. Moreover, if u is the time-continuous solution with $u(0) = u_0$ and if $\overline{u}^\tau$ is the left-continuous piecewise constant interpolant of $(u^\tau_k)_{k \in \mathbb{N}}$, then

$$d_\mathcal{K}(u(t), \overline{u}^\tau(t)) \leq C(u_0) \sqrt{\tau}\, \mathrm{e}^{-\lambda_\tau t} \qquad \text{for } t \geq 0,$$

see [AGS05, Thms. 4.0.9+10], where $\lambda_\tau = \lambda$ for $\lambda < 0$ and $\lambda_\tau = \frac{1}{\lambda} \log(1+\lambda\tau)$ for $\lambda > 0$.

Finally, it was shown in [DaS08, Prop. 3.1] that for geodesically λ-convex functionals the solutions of the (differential) gradient flow (2.1) satisfy a purely metric formulation in terms of the evolutionary variational inequality (EVI$_\lambda$)

$$\frac{1}{2} \frac{\mathrm{d}^+}{\mathrm{d}t} d^2_\mathcal{K}(u(t), w) + \frac{\lambda}{2} d^2_\mathcal{K}(u(t), w) + \mathcal{E}(u(t)) \leq \mathcal{E}(w), \quad \forall\, w \in X,\ t > 0,$$

where for a function $f : [0, \infty[\to \mathbb{R}$ we set $\frac{\mathrm{d}^+}{\mathrm{d}t} f(t) = \limsup_{h \to 0^+} \frac{1}{h}\big(f(t+h) - f(t)\big)$. The above differential form is equivalent to the integrated form of (EVI$_\lambda$) given by

$$\frac{\mathrm{e}^{\lambda\tau}}{2} d_\mathcal{K}(u(t+\tau), w)^2 - \frac{1}{2} d_\mathcal{K}(u(t), w)^2 \leq \int_0^\tau \mathrm{e}^{\lambda r}\, \mathrm{d}r\ \big(\mathcal{E}(w) - \mathcal{E}(u(t+\tau))\big)$$
$$\text{for all } t, \tau \geq 0 \text{ and } w \in X.$$

In particular, the solutions of (EVI$_\lambda$) satisfy the uniform regularization bound

$$\mathcal{E}(u(t)) \leq \mathcal{E}(w) + \frac{\lambda}{2(\mathrm{e}^{\lambda t} - 1)} d_\mathcal{K}(u(0), w) \text{ for all } w \in X \text{ and } t > 0.$$

Moreover, the solutions are uniformly continuous in time:

$$d_\mathcal{K}(u(t+\tau), u(t))^2 \leq 2\, \frac{\mathrm{e}^{-\lambda\tau} - 1}{\lambda} \Big(\mathcal{E}(u(t)) - \inf_{w \in X} \mathcal{E}(w) \Big).$$

3.4 Completion of smooth gradient flows

In addition to (3.1) and (3.2) we now assume that $(\mathcal{Z},\mathcal{E},\mathcal{K})$ generates a global semiflow in the form $u(t)=\mathcal{S}_t(u(0))$ with a semigroup $\mathcal{S}:[0,\infty[\,\times\,\mathcal{Z}\to\mathcal{Z}$, i.e.

$$\begin{aligned}&\mathcal{S}_t\circ\mathcal{S}_r=\mathcal{S}_{t+r} \text{ for } r,t\geq 0;\\ &\mathcal{S}_t(u)\to u \text{ in } Z \text{ and } \tfrac{1}{t}(\mathcal{S}_t(u){-}u)\to -\mathcal{F}(u) \text{ in } Y \quad \text{for } t\to 0^+.\end{aligned}$$

The assumptions on the semiflow $\mathcal{S}$ are

$$\mathcal{S}\in \mathrm{C}^0([0,\infty[\,{\times}\mathcal{Z};Z)\,\cap\,\mathrm{C}^1([0,\infty[\,{\times}\mathcal{Z};Y)\,\cap\,\mathrm{C}^2([0,\infty[\,{\times}\mathcal{Z};H). \tag{3.8}$$

In particular, this implies that $\mathrm{D}\mathcal{S}$ and $\mathcal{F}(u)=-\partial_t\mathcal{S}_t(u)|_{t=0}$ satisfy

$$(t,u)\mapsto \mathrm{D}\mathcal{S}_t(u)\in \mathrm{C}^0([0,\infty[\,{\times}\mathcal{Z};\mathrm{Lin}(Z;Y))\,\cap\,\mathrm{C}^1([0,\infty[\,{\times}Z;\mathrm{Lin}(Z;H)). \tag{3.9}$$

We define the functionals $\mathcal{A}:\mathcal{Y}\times H\to\mathbb{R}$ and $\mathcal{B}:\mathcal{Z}\times Y\to\mathbb{R}$ via

$$\mathcal{A}(u,v)=\langle\mathcal{G}(u)v,v\rangle,\quad \mathcal{B}(u,v)=\langle\mathcal{G}(u)v,\mathrm{D}\mathcal{F}(u)v\rangle+\frac{1}{2}\langle\mathrm{D}\mathcal{G}(u)[\mathcal{F}(u)]v,v\rangle$$

and obtain the following formulas.

Proposition 3.3 *(i) For $u\in\mathrm{C}^1([t_0,t_1];\mathcal{Y})$ and $v\in\mathrm{C}^1([t_0,t_1];H)$ we have*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(u(t),v(t))=2\langle\mathcal{G}(u)v,\dot v\rangle+\langle\mathrm{D}\mathcal{G}(u)[\dot u]v,v\rangle. \tag{3.10}$$

(ii) For all $u\in\mathcal{Z}$, $v\in Z$, and $t\geq 0$ we have

$$\tfrac{1}{2}\,\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(\mathcal{S}_t(u),\mathrm{D}\mathcal{S}_t(u)v)+\mathcal{B}(\mathcal{S}_t(u),\mathrm{D}\mathcal{S}_t(u)v)=0. \tag{3.11}$$

Proof: Part (i) follows simply by the assumed smoothness of $\mathcal{G}$ and the chain rule for the Fréchet derivative in Banach spaces. Part (ii) is an application of part (i) by using $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_t(u)=-\mathcal{F}(\mathcal{S}_t(u))$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{D}\mathcal{S}_t(u)=-\mathrm{D}\mathcal{F}(\mathcal{S}_t(u))\mathrm{D}\mathcal{S}_t(u)$. ∎

The central idea of [DaS08] is the transport of curves $\gamma_t\in\boldsymbol{C}(u_0,\mathcal{S}_t(u_1))$ defined via

$$\gamma_t(s)=\mathcal{S}_{st}(\gamma(s))\quad \text{for } \gamma\in\boldsymbol{C}(u_0,u_1)\cap\mathrm{C}^2([0,1];\mathcal{Z}).$$

The main tool is the following relation (3.12) for the functions

$$A(s,t):=\mathcal{A}(\gamma_t(s),\gamma_t'(s)),\quad B(s,t):=\mathcal{B}(\gamma_t(s),\gamma_t'(s)),\ \text{ and } E(s,t):=\mathcal{E}(\gamma_t(s)),$$

where $\gamma_t'(s)=\partial_s(\gamma_t(s))\in Y$.

Proposition 3.4 *For every curve $\gamma\in\boldsymbol{C}(w,u)$ we have*

$$\frac{1}{2}\,\frac{\partial}{\partial t}A(s,t)+\frac{\partial}{\partial s}E(s,t)+sB(s,t)=0. \tag{3.12}$$

Proof: We first observe that the mapping $\Gamma : (s,t) \mapsto \gamma_t(s)$ satisfies

$$\Gamma \in \mathrm{C}^0([0,1]\times [0,\infty[\,;\mathcal{Z}) \cap \mathrm{C}^1([0,1]\times [0,\infty[\,;Y) \cap \mathrm{C}^2([0,1]\times [0,\infty[\,;H).$$

In particular, using the definition of the semiflow $\mathcal{S}_t$ we have the relations

$$\partial_t\gamma_t(s) = -s\mathcal{F}(\gamma_t(s)) \text{ and } \partial_t(\gamma_t'(s)) = \partial_s\partial_t\gamma_t(s) = -\mathcal{F}(\gamma_t(s)) - s\mathrm{D}\mathcal{F}(\gamma_t(s))\gamma_t'(s).$$

Note that we will not need an expression for $\gamma_t'(s)$. Applying Proposition 3.3(i) and the above formulas for $\partial_t\gamma_t(s)$ and $\partial_t(\gamma_t'(s))$ we find

$$\begin{aligned}\tfrac{1}{2}\tfrac{\partial}{\partial t}A(s,t) &= -\langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), \mathcal{F}(\gamma_t(s))\rangle - \langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), s\,\mathrm{D}\mathcal{F}(\gamma_t(s))\gamma_t'(s)\rangle \\ &\quad - \tfrac{1}{2}\langle \mathrm{D}\mathcal{G}(\gamma_t(s))[s\mathcal{F}(\gamma_t(s))]\gamma_t'(s), \gamma_t'(s)\rangle \\ &= -\langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), \mathcal{K}(\gamma_t(s))\mathrm{D}\mathcal{E}(\gamma_t(s))\rangle - s\mathcal{B}(\gamma_t(s), \gamma_t'(s)) \\ &= -\langle \mathrm{D}\mathcal{E}(\gamma_t(s)), \gamma_t'(s)\rangle - sB(s,t) \;=\; -\tfrac{\partial}{\partial s}E(s,t) - sB(s,t),\end{aligned}$$

which is the desired result. ∎

One of the main achievements of [DaS08] was to show that the identity (3.12) can be used to derive the evolutionary variational inequality (EVI$_\lambda$), namely

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}d_\mathcal{K}(\mathcal{S}_t(u), w)^2 + \frac{\lambda}{2}d_\mathcal{K}(\mathcal{S}_t(u), w)^2 + \mathcal{E}(\mathcal{S}_t(u)) \leq \mathcal{E}(w) \text{ for all } t \geq 0 \text{ and } w \in \mathcal{Z}. \tag{3.13}$$

It is especially interesting that this result holds without any completeness of the space $\mathcal{Z}$. The crucial assumption needed is that $B(s,t)$ can be estimated in terms of $A(s,t)$, namely in the form $B(s,t) \geq \lambda A(s,t)$ along the curves γ_t. The following result is an abstract version of the ideas in [DaS08].

Theorem 3.5 *Assume that* $(\mathcal{Z}, \mathcal{E}, \mathcal{K})$ *generates the semigroup* $\mathcal{S}$ *and the above conditions* (3.1)—(3.8) *hold. If additionally*

$$\begin{aligned}&\forall\, u \in \mathcal{Z}\;\forall\, v \in Y: \quad \mathcal{B}(u,v) \geq \lambda\mathcal{A}(u,v), \\ &\textit{i.e. } \langle \mathcal{G}(u)v, \mathrm{D}\mathcal{F}(u)v\rangle + \frac{1}{2}\langle \mathrm{D}\mathcal{G}(u)[\mathcal{F}(u)]v, v\rangle \geq \lambda\langle \mathcal{G}(u)v, v\rangle,\end{aligned} \tag{3.14}$$

then, the semigroup $\mathcal{S}$ *satisfies (EVI$_\lambda$) given in* (3.13).

Proof: We apply the theory in [DaS08, Sect. 5], where the underlying metric space (X, d) is not assumed to be complete. Hence we are able to choose $(\mathcal{Z}, d_\mathcal{K})$. Following the analysis in the proof of [DaS08, Theorem 5.1] (see (5.14) and (5.17) there) we obtain the final estimate

$$\tfrac{\mathrm{e}^{\lambda t}\mathsf{s}(\lambda t)}{2}d_\mathcal{K}(\mathcal{S}_t(u), w)^2 - \tfrac{1}{2}d_\mathcal{K}(u,w)^2 + \mathsf{E}_{2\lambda}(t)\mathcal{E}(\mathcal{S}_t(u)) - t\mathcal{E}(w) = O(t^2).$$

Dividing by t, using $\mathsf{s}(t) = t/\sinh(t) = 1 + O(t^2)$ and $\mathsf{E}_\lambda(t) = \int_0^t \mathrm{e}^{\lambda r}\,\mathrm{d}r = t + O(t^2)$, and taking the limit $t \to 0^+$ we obtain

$$\tfrac{1}{2}\tfrac{\mathrm{d}^+}{\mathrm{d}t}d_\mathcal{K}(\mathcal{S}_t(u), w)^2\big|_{t=0} + \tfrac{\lambda}{2}d_\mathcal{K}(u,w)^2 + \mathcal{E}(u) \leq \mathcal{E}(w) \text{ for } w \in \mathcal{Z}.$$

Since $u \in \mathcal{Z}$ was arbitrary, it can be replaced by $\widetilde{u} = \mathcal{S}_t(u)$, and the desired (EVI$_\lambda$) in (3.13) follows. ∎

Since in applications the metric $\mathcal{G}$ is often not given explicitly (see examples in Section 4), it is desirable to express the fundamental estimate (3.14) in terms of the Onsager operator $\mathcal{K} = \mathcal{G}^{-1}$.

Proposition 3.6 *Assume that*

$$\begin{aligned}&\forall\, u \in \mathcal{Z}\, \forall\, \eta \in \mathcal{G}(u)Y: \quad \mathcal{M}(u,\eta) \geq \lambda\langle \eta, \mathcal{K}(u)\eta\rangle,\\ &\textit{where } \mathcal{M}(u,\eta) := \langle \eta, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\eta\rangle - \tfrac{1}{2}\langle \eta, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\eta\rangle,\end{aligned} \tag{3.15}$$

then estimate (3.14) *holds.*

Proof: The proof is immediate since for a given $v \in Y$ we can use $\eta = \mathcal{G}(u)v$ in (3.15). After using the formula for the derivative of the inverse, namely $\mathrm{D}\mathcal{G}(u)[w] = -\mathcal{G}(u)\mathrm{D}\mathcal{K}(u)[w]\mathcal{G}(u)$ we find (3.14). ∎

The conditions in Proposition 3.6 are similar to the conditions of Bakry and Émery [BaÉ85, Bak94]. We now return to the metric evolution in the larger space $\mathcal{X}$. For this, we assume that $d_\mathcal{K}$ on $\mathcal{Z}$ can be extended to a metric on $\mathcal{X}$ such that

$$(\mathcal{X}, d_\mathcal{K}) \text{ is a complete metric space.} \tag{3.16}$$

Moreover, $\mathcal{E} : \mathcal{Z} \to \mathbb{R}$ is assumed to have a lower semicontinuous extension $\overline{\mathcal{E}} : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ (with respect to the metric topology). Finally, $\mathcal{Z}$ is assumed to be dense, viz.

$$\forall\, u \in \mathcal{X} \text{ with } \overline{\mathcal{E}}(u) < \infty\; \exists\, u_n \in \mathcal{Z}: \quad d_\mathcal{K}(u_n, u) \to 0 \text{ and } \mathcal{E}(u_n) \to \overline{\mathcal{E}}(u). \tag{3.17}$$

Using the Lipschitz continuity (3.7), there is a unique continuous extension $\overline{\mathcal{S}} : [0,\infty[\times \mathcal{X} \to \mathcal{X}$. Then, [DaS08, Thm. 3.3] provides the following result.

Theorem 3.7 *If* (3.16)*,* (3.17) *and the assumptions of Theorem 3.5 hold, then the semiflow $\overline{\mathcal{S}}$ associated with the gradient system $(\mathcal{X}, \overline{\mathcal{E}}, d_\mathcal{K})$ satisfies EVI$_\lambda$* (3.13) *and the Lip-schitz continuity* (3.7) *with $(\mathcal{E}, \mathcal{S})$ replaced by $(\overline{\mathcal{E}}, \overline{\mathcal{S}})$. Moreover, $\overline{\mathcal{E}}$ is geodesically λ-convex on $\mathcal{X}$, i.e. for every arc-length parameterized geodesic curve $\gamma \in \mathrm{C}^0([0,1];\mathcal{X})$ we have*

$$\overline{\mathcal{E}}(\gamma(s)) \leq (1{-}s)\,\overline{\mathcal{E}}(\gamma(0)) + s\,\overline{\mathcal{E}}(\gamma(1)) - \frac{\lambda}{2}s(1{-}s)\,\mathrm{d}_\mathcal{K}(\gamma(0),\gamma(1))^2 \textit{ for } s \in [0,1]. \tag{3.18}$$

4 Examples

This section surveys possible applications of the abstract methods developed above to scalar equations as well as reaction-diffusion systems. In particular, we show geodesic λ-convexity in a smooth setting by establishing the estimate $\mathcal{M}(u,\xi) \geq \lambda\langle \xi, \mathcal{K}(u)\xi\rangle$. In particular, we generalize the known results for scalar drift-diffusion equations (with conserved mass) to systems with reaction terms (non-conserved masses). The discussion of the corresponding metric spaces $(X, d_\mathcal{K})$ is postponed to future research.

4.1 Pure reaction systems and Markov chains

In [Mie11c] an entropy gradient structure was established for general reaction systems of mass-action type that satisfy the detailed balance condition. We consider a vector $\boldsymbol{u} \in \,]0,\infty[^n$ of densities and R polynomial reactions

$$\dot{\boldsymbol{u}} = -\sum_{r=1}^{R} k^r(\boldsymbol{u})\Big(\frac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\alpha^r}} - \frac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}\Big)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big), \qquad \text{where } \boldsymbol{u}^{\boldsymbol{\alpha}^r} = \Pi_{i=1}^n u_i^{\alpha_i^r}. \tag{4.1}$$

Here $\boldsymbol{w} \in \left]0,\infty\right[^n$ is the reference density, which is obviously a steady state and satisfies the detailed balance condition. Moreover, $k^r(\boldsymbol{u}) \geq 0$ is the reaction coefficient (normalized with respect to $\boldsymbol{w}$), and the vectors $\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r \in \left]0,\infty\right[^n$ are the stoichiometric vectors for the forward and backward reactions. Usually the entries are assumed to be nonnegative integers, but this is not necessary here. The gradient system $(\left]0,\infty\right[^n, E, \mathbb{K})$ with

$$E(\boldsymbol{u}) = \sum_{i=1}^{n} u_i \log(u_i/w_i) \text{ and } \mathbb{K}(\boldsymbol{u}) = \sum_{r=1}^{R} k^r(\boldsymbol{u})\Lambda\big(\tfrac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}}, \tfrac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}\big)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \otimes \big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big)$$

gives (4.1). We find $\mathcal{M}(\boldsymbol{u},\boldsymbol{\xi}) = \boldsymbol{\xi}\cdot\boldsymbol{M}(\boldsymbol{u})\boldsymbol{\xi}$, where $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{R}^{I\times I}$ is defined via

$$\boldsymbol{M}(\boldsymbol{u}) = \tfrac{1}{2}\Big(\mathbb{K}(\boldsymbol{u})\mathrm{D}\mathcal{F}(\boldsymbol{u})^{\mathsf{T}} + \mathrm{D}\mathcal{F}(\boldsymbol{u})\mathbb{K}(\boldsymbol{u}) - \mathrm{D}\mathbb{K}(\boldsymbol{u})[\mathcal{F}(\boldsymbol{u})]\Big),$$

see also [Mie11b]. Note that the vector field $\mathcal{F}(\boldsymbol{u}) = \mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u})$ is nonlinear and that the matrices $\mathbb{K}(\boldsymbol{u})$ and $\mathbb{M}(\boldsymbol{u})$ have no homogeneity or concavity properties, in general.

We want to study a few simple cases and discuss the possibility of geodesic λ-convexity. For $R = 1$ we drop the reaction number r and write $\boldsymbol{\gamma} = \boldsymbol{\alpha} - \boldsymbol{\beta}$ and $\boldsymbol{\varrho} = (u_i/w_i)_i$. Then,

$$\begin{aligned}
\mathcal{F}(\boldsymbol{u}) &= \phi(\boldsymbol{u})\boldsymbol{\gamma} && \text{with } \phi(\boldsymbol{u}) = k(\boldsymbol{u})(\boldsymbol{\varrho}^{\boldsymbol{\alpha}} - \boldsymbol{\varrho}^{\boldsymbol{\beta}}),\\
\mathbb{K}(\boldsymbol{u}) &= \kappa(\boldsymbol{u})\boldsymbol{\gamma}\otimes\boldsymbol{\gamma} && \text{with } \kappa(\boldsymbol{u}) = k(\boldsymbol{u})\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}}),\\
\boldsymbol{M}(\boldsymbol{u}) &= m(\boldsymbol{u})\boldsymbol{\gamma}\otimes\boldsymbol{\gamma} && \text{with } m(\boldsymbol{u}) = \kappa(\boldsymbol{u})\mathrm{D}\phi(\boldsymbol{u})\cdot\boldsymbol{\gamma} - \tfrac{1}{2}\phi(\boldsymbol{u})\mathrm{D}\kappa(\boldsymbol{u})\cdot\boldsymbol{\gamma}.
\end{aligned}$$

The general case seems too difficult to be analyzed, hence we reduce to the case $k(\boldsymbol{u}) \equiv 1$. Introducing the matrix $V = \mathrm{diag}(1/u_i)_i$ we have $\mathrm{D}_{\boldsymbol{u}}(\boldsymbol{u}^{\boldsymbol{\alpha}})[\boldsymbol{\gamma}] = \boldsymbol{u}^{\boldsymbol{\alpha}}\boldsymbol{\alpha}\cdot V\boldsymbol{\gamma}$, and after some elementary calculations involving the properties of the function Λ (see [Mie11b]) we find

$$m(\boldsymbol{u}) = \tfrac{1}{2}\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})\big(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}\boldsymbol{\alpha} - \boldsymbol{\varrho}^{\boldsymbol{\beta}}\boldsymbol{\beta} + \Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\boldsymbol{\alpha}-\boldsymbol{\beta})\big)\cdot V(\boldsymbol{\alpha}-\boldsymbol{\beta}).$$

For geodesic λ-convexity we have to show $m(\boldsymbol{u}) \geq \lambda\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})$ which after dividing by $\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})$ leads to the formula

$$\lambda = \tfrac{1}{2}\inf\{\, \textstyle\sum_{i=1}^{n} \frac{(\alpha_i-\beta_i)}{w_i\varrho_i}\big[\boldsymbol{\varrho}^{\boldsymbol{\alpha}}\alpha_i - \boldsymbol{\varrho}^{\boldsymbol{\beta}}\beta_i + \Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\alpha_i-\beta_i)\big] \mid \boldsymbol{\varrho} \in \left]0,\infty\right[^n \,\}.$$

In the special case where $\alpha_i\beta_i = 0$ for all i we find the simpler form

$$\lambda = \tfrac{1}{2}\inf\{\, \textstyle\sum_{i=1}^{n} \frac{1}{w_i\varrho_i}\big(\alpha_i^2\boldsymbol{\varrho}^{\boldsymbol{\alpha}} + \beta_i^2\boldsymbol{\varrho}^{\boldsymbol{\beta}} + \Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\alpha_i^2+\beta_i^2)\big) \mid \boldsymbol{\varrho} \in \left]0,\infty\right[^n \,\} \geq 0.$$

This formula applies to example (2.8) where $\boldsymbol{\alpha} = (1,1,0,0)^{\mathsf{T}}$ and $\boldsymbol{\beta} = (0,0,1,1)^{\mathsf{T}}$. Because of $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \geq 2$ the infimum is $\lambda = 0$ (by choosing $\boldsymbol{\varrho} = \varepsilon(1,1,1,1)$ and $\varepsilon \to 0$).

Example 4.1 *The annihilation-creation reaction modeling the recombination and generation of electron-hole pairs in semiconductors, cf. [Gli08, GlG09] and Section 2.4, reads*

$$\dot{\boldsymbol{u}} = -\big(u_1u_2 - 1\big)\,(1,1)^{\mathsf{T}}, \quad \textit{where} \quad \boldsymbol{\alpha} = (1,1)^{\mathsf{T}} \textit{ and } \boldsymbol{\beta} = (0,0)^{\mathsf{T}}. \tag{4.2}$$

The formula yields $\lambda = \frac{1}{2}\inf\{\, \big(\frac{1}{u_1}+\frac{1}{u_2}\big)\big(u_1u_2+\Lambda(u_1u_2,1)\big) \mid u_1,u_2 > 0 \,\} = \cosh(1) > 0.$

Discrete Markov chains can be seen as special reaction systems where only exchange reactions $X_i \leftrightharpoons X_j$ occur. The reaction system takes the form

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) = Q\boldsymbol{u}, \quad \text{where } Q_{ij} \geq 0 \text{ for } i \neq j \quad \text{and } \sum_{i=1}^{I} Q_{ij} = 0. \tag{4.3}$$

We assume that there is a unique steady state $\boldsymbol{w}$ with $w_i > 0$ for all i (also called irreducibility). A much stronger assumption is the condition of detailed balance, which reads $Q_{ij}w_j = Q_{ji}w_i$ for $i, j = 1, ..., I$. According to [Mie11b, Maa11], (4.3) is induced by the gradient system $(X_{\mathsf{Mkv}}, E_{\mathsf{Mkv}}, \mathbb{K}_{\mathsf{Mkv}})$, where $X_{\mathsf{Mkv}} = \{\, \boldsymbol{u} \in [0,1]^I \mid \sum_{i=1}^{I} u_i = 1 \,\}$, $\mathcal{E}_{\mathsf{Mkv}}(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i)$, and

$$\mathbb{K}_{\mathsf{Mkv}}(\boldsymbol{u}) = \sum_{1\leq i<j\leq I} Q_{ij}w_j\, \Lambda\big(\tfrac{u_i}{w_i}, \tfrac{u_j}{w_j}\big)\, \big(e_i{-}e_j\big) \otimes \big(e_i{-}e_j\big) \;\in \mathbb{R}^{I\times I},$$

where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^I$ are the unit vectors. Moreover, it is shown in [Mie11b] that for all Markov chains there is a $\lambda \in \mathbb{R}$ such that $(\mathcal{E}_{\mathsf{Mkv}}, \mathbb{K}_{\mathsf{Mkv}})$ is geodesically λ-convex. For special classes, like tridiagonal Q, explicit estimates for λ are obtained.

4.2 Scalar diffusion equation

We consider a bounded, convex domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, with smooth boundary. In Ω we are given the scalar diffusion equation

$$\dot{u} = \operatorname{div}(a(u)\nabla u) \quad \text{in } \Omega, \qquad \nabla u \cdot \nu = 0 \quad \text{on } \partial\Omega. \tag{4.4}$$

This equation is the gradient flow of the energy $\mathcal{E}$ with respect to the Onsager operator $\mathcal{K}$ given via

$$\mathcal{E}(u) = \int_\Omega E(u)\,\mathrm{d}x \qquad \text{and} \qquad \mathcal{K}(u)\xi = -\operatorname{div}(\mu(u)\nabla\xi),$$

where E and μ are such that $\mu(u)E''(u) = a(u)$ holds. In particular, we assume that $E, \mu \in \mathrm{C}^2(]0,\infty[)$ and the sign conditions

$$\mu(u) \geq 0, \quad \mu''(u) \leq 0, \quad E''(u) > 0 \quad \text{for all } u > 0.$$

We impose that solutions $u : \Omega \to \mathbb{R}$ of (4.4) are sufficiently smooth for given smooth initial conditions such that the assumptions of the last section for the semiflow $\mathcal{S}_t : u(0) \mapsto u(t)$ are satisfied.

In the following we slightly deviate from the setting in Section 3.4 in that we consider $\mathcal{Y}$ and $\mathcal{Z}$ to be open and connected subsets of affine spaces $u_*{+}Y$ and $u_*{+}Z$ where the shift is given by $u_* = 1/|\Omega|$. This modification allows us to extend our theory to the space of probability measures $\rho \in \mathcal{P}(\Omega)$. More precisely, let $X = \mathcal{M}(\Omega)$ the space of Radon measures ρ (using that Ω is bounded all moments $\int_\Omega |x|^p \,\mathrm{d}\rho(x)$ are finite) and $\mathcal{X} = \mathcal{P}(\Omega)$ denotes the subset of probability measures such that

$$\mathcal{X} = \{\, \rho \in X \mid \rho(\Omega) = 1 \text{ and } \rho \geq 0 \,\}.$$

The results of Section 3.4 can be easily adapted to this case.

The quadratic form associated with the operator $\mathcal{K}$ defines in a natural way the spaces

$$H^* = \mathrm{H}^1_{\mathsf{av}}(\Omega) = \{\, \xi \in \mathrm{H}^1(\Omega) \mid \textstyle\int_\Omega \xi \,\mathrm{d}x = 0 \,\}, \quad H = \mathrm{H}^{-1}_{0,\mathsf{av}} = (\mathrm{H}^1_{\mathsf{av}}(\Omega))^*.$$

Moreover, we choose $s \geq 4$ such that $s > 2 + d/2$ and define the spaces

$$
\begin{aligned}
&Y = \{\, v \in \mathrm{H}^{s-2}(\Omega) \mid \textstyle\int_\Omega v \,\mathrm{d}x = 0 \text{ and } \nabla v \cdot \nu = 0 \text{ on } \partial\Omega \,\}, \quad \mathcal{Y} = \{\, u \in u_*{+}Y \mid \inf u > 0 \,\},\\
&Z = \mathrm{H}^s(\Omega) \cap Y, \qquad \mathcal{Z} = \{\, u \in u_*{+}Z \cap \mathcal{Y} \mid \nabla\big(\operatorname{div}(a(u)\nabla u)\big) \cdot \nu = 0 \text{ on } \partial\Omega \,\}.
\end{aligned}
$$

The boundary condition in the definition of the set $\mathcal{Z}$ is necessary to ensure that the semiflow satisfies $\mathcal{S} \in \mathrm{C}^1([0,\infty[\times\mathcal{Z};Y)$. In particular, for a solution $t \mapsto u(t) \in \mathcal{Z}$ holds $\dot u(t) \in Y$. Obviously we have $Z \subset Y \subset X$ and $\mathcal{Y} \subset u_* + Y$ and $\mathcal{Z} \subset u_* + Z$ with dense embeddings.

Our analysis is similar to that in [DaS08, Sect. 4] with the main difference that we have to take care of the boundary conditions when doing integrations by part. There are two crucial observations for the case with boundaries. Firstly, the curvature of the boundary of convex bodies provides a sign for the normal derivative $\nabla(|\nabla\xi|^2) \cdot \nu \leq 0$, whenever $\nabla\xi \cdot \nu = 0$ holds, see Proposition 4.2. Secondly, the testfunctions $\xi \in \mathcal{G}(u)Y$ will satisfy *two boundary conditions*, namely

$$
-\operatorname{div}(\mu(u)\nabla\xi) = v \in Y \quad \Longrightarrow \quad \big(\, \nabla\xi \cdot \nu = 0 \;\text{ and }\; \nabla\big(\operatorname{div}(\mu(u)\nabla\xi)\big) \cdot \nu = 0 \,\big).
$$

In order to show that the assumptions of Proposition 3.6 hold we have to compute the quadratic form $\mathcal{M}(u,\xi) = \langle \xi, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\xi\rangle - \frac12 \langle \xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi\rangle$, with

$$
\mathcal{F}(u) = -\operatorname{div}(a(u)\nabla u) \quad \text{and} \quad \mathrm{D}\mathcal{F}(u)[v] = -\operatorname{div}\big(a'(u)v\nabla u + a(u)\nabla v\big).
$$

For $\xi \in \mathcal{G}(u)Y$ we use the abbreviation $v = \mathcal{K}(u)\xi \in Y$ and obtain

$$
\begin{aligned}
\mathcal{M}_0(u,\xi) &= -\int_\Omega \xi\big(\operatorname{div}(a'(u)v\nabla u + a(u)\nabla v)\big)\,\mathrm{d}x - \frac12 \int_\Omega \mu'(u)\big(-\operatorname{div}(a(u)\nabla u)\big)|\nabla\xi|^2\,\mathrm{d}x\\
&= \int_\Omega \nabla\xi \cdot \big(a'(u)v\nabla u + a(u)\nabla v\big)\,\mathrm{d}x - \int_\Omega a(u)\nabla u \cdot \nabla\big(\mu'(u)\tfrac12|\nabla\xi|^2\big)\,\mathrm{d}x,
\end{aligned}
$$

where in both cases the boundary terms vanish, namely using $(a'(u)v\nabla u{+}a(u)\nabla v) \cdot \nu = 0$ and $a(u)\nabla u \cdot \nu = 0$ from $v \in Y$ and $u \in \mathcal{Z}$. Applying integration by parts one more time yields

$$
\begin{aligned}
\mathcal{M}_0(u,\xi) &= \int_\Omega a(u)\Delta\xi \;\operatorname{div}(\mu(u)\nabla\xi)\,\mathrm{d}x - \int_\Omega a(u)\nabla u \cdot \nabla\big(\mu'(u)\tfrac12|\nabla\xi|^2\big)\,\mathrm{d}x\\
&= \int_\Omega \nabla H(u) \cdot \big((\Delta\xi)\nabla\xi - \nabla(\tfrac12|\nabla\xi|^2)\big) + a(u)\mu(u)(\Delta\xi)^2 - \frac{a(u)\mu''(u)}{2}|\nabla u|^2|\nabla\xi|^2\,\mathrm{d}x,
\end{aligned}
$$

where we have set $H(u) = \int_0^u a(y)\mu'(y)\,\mathrm{d}y$ and used that $\nabla\xi \cdot \nu = 0$. Finally, integrating by parts one last time leads to

$$
\begin{aligned}
\mathcal{M}_0(u,\xi) = &\int_\Omega H(u)|\mathrm{D}^2\xi|^2 + \big(a(u)\mu(u){-}H(u)\big)(\Delta\xi)^2 - \frac{a(u)\mu''(u)}{2}|\nabla u|^2|\nabla\xi|^2\,\mathrm{d}x\\
&- \int_{\partial\Omega} H(u)\nabla(\tfrac12|\nabla\xi|^2) \cdot \nu\,\mathrm{da}.
\end{aligned}
\tag{4.5}
$$

Here we used Bochner's formula $\operatorname{div}((\Delta\xi)\nabla\xi) - \Delta(\frac12|\nabla\xi|^2) = (\Delta\xi)^2 - |\mathrm{D}^2\xi|^2$. The boundary integral is nonpositive using the assumption $H(u) \geq 0$ and Proposition 4.2 below.

Thus, we have shown that $\mathcal{M}(u,\xi) \geq 0$ holds if we assume that $u \mapsto \mu(u)$ is concave and $a\mu \geq \frac{d-1}{d}H \geq 0$. Here the latter condition is due to the elementary estimate

$$
\forall \mathbb{E} \in \mathbb{R}^{d\times d}: \qquad \alpha|\mathbb{E}|^2 - \beta(\operatorname{tr}\mathbb{E})^2 \geq 0 \quad \Leftrightarrow \quad \alpha \geq \max\{0, d\beta\}.
$$

Now, Proposition 3.6 states that $(\mathcal{E}, \mathcal{K})$ is geodesically 0-convex. Since the present result will be a special case of the result in the next subsection, we refer to Theorem 4.3 for the precise statement.

Thus, we have generalized [DaS08, Thm. 4.2] from manifolds without boundary to the case of convex domains in $\mathbb{R}^d$ with smooth boundaries. The condition of convexity is quite natural in the context of optimal transport, since only convex domains are still complete metric length-spaces with respect to the Euclidean distance.

We close this subsection with the result on the signs of $\nabla\big(|\nabla\xi|^2\big) \cdot \nu$ on the boundary. We refer to [Gri85, Ch. 3] and [GST09, Lem. 5.2] for previous proofs, but still give an independent proof of a more general result needed in Section 4.7. It involves the second fundamental form $\mathbb{I}$ of the boundary, i.e. for two tangent vectors $\tau_1, \tau_2 \in \mathrm{T}_x\partial\Omega$ we have $\mathbb{I}(\tau_1, \tau_2) = \tau_1 \cdot \mathrm{D}\nu(x)\tau_2 = \mathbb{I}(\tau_2, \tau_1)$, where ν is the outer normal vector.

Proposition 4.2 *Assume that $\Omega \subset \mathbb{R}^d$ is a domain with C^2 boundary. Then, for functions $\xi_1, \xi_2 \in \mathrm{H}^3(\Omega)$ with $\nabla\xi_1 \cdot \nu = \nabla\xi_2 \cdot \nu = 0$ on $\partial\Omega$ we have the identity*

$$\nabla\big(\nabla\xi_1 \cdot \nabla\xi_2\big) \cdot \nu = -2\mathbb{I}(\nabla_{\|}\xi_1, \nabla_{\|}\xi_2), \tag{4.6}$$

where $\nabla_{\|}\xi$ denotes the tangential part of the gradient $\nabla_{\|}\xi = \nabla\xi - (\nabla\xi\cdot\nu)\nu$. In particular, if Ω is convex and $\xi_2 = \xi_1$, then $\nabla\big(|\nabla\xi_1|^2\big) \cdot \nu \leq 0$ on $\partial\Omega$.

Proof: Without loss of generality we assume that ξ_j is smooth. We denote by $\overline{\nu} \in \mathrm{C}^1(\overline{\Omega})$ a smooth extension of the outer unit normal ν into Ω. For $x \in \Omega$ we compute

$$\begin{aligned} &\nabla\Big(\nabla\xi_1 \cdot \nabla\xi_2\Big) \cdot \overline{\nu}(x) = \nabla\xi_2 \cdot \mathrm{D}^2\xi_1\,\overline{\nu} + \nabla\xi_1 \cdot \mathrm{D}^2\xi_2\,\overline{\nu} \\ &= \nabla\xi_2 \cdot \Big(\nabla\big(\nabla\xi_1 \cdot \overline{\nu}\big) - \mathrm{D}\overline{\nu}\nabla\xi_1\Big) + \nabla\xi_1 \cdot \Big(\nabla\big(\nabla\xi_2 \cdot \overline{\nu}\big) - \mathrm{D}\overline{\nu}\nabla\xi_2\Big). \end{aligned} \tag{4.7}$$

On the boundary the product $\nabla\xi_j \cdot \overline{\nu}$ vanishes identically, such that $\nabla_{\|}(\nabla\xi_j \cdot \overline{\nu}) = 0$ on $\partial\Omega$. Hence, there are scalar functions $\gamma_j : \partial\Omega \to \mathbb{R}$ such that $\nabla\big(\nabla\xi_j \cdot \overline{\nu}\big) = \gamma\nu$ on $\partial\Omega$. Inserting this into (4.7) and using $\nabla\xi_j \cdot \nu = 0$ we have established (4.6).

For a convex body, the second fundamental form is positive semidefinite. Hence, formula (4.6) gives the desired result for $\xi_1 = \xi_2$. ∎

We end this subsection by mentioning that the theory can also be applied to smooth inhomogeneous systems, e.g. where the mobility depends on the spatial variable $x \in \Omega$:

$$\dot{u} = -\mathcal{F}(u) = \operatorname{div}(\mathbb{M}(x)\nabla u), \quad \mathcal{K}(u)\xi = -\operatorname{div}(u\,\mathbb{M}(x)\nabla\xi), \quad \mathcal{E}(u) = \int_\Omega u\log u\,\mathrm{d}x,$$

where $\mathbb{M} \in \mathrm{W}^{2,\infty}(\overline{\Omega}; \mathbb{R}^{d\times d}_{\mathsf{spd}})$, and there exists $\alpha_0 > 0$ with $\boldsymbol{a}\cdot\mathbb{M}(x)\boldsymbol{a} \geq \alpha_0|\boldsymbol{a}|^2$. The appropriate boundary conditions are now $(\mathbb{M}(x)\nabla u(x)) \cdot \nu(x) = 0 = (\mathbb{M}(x)\nabla\xi(x)) \cdot \nu(x)$ for $x \in \partial\Omega$. Doing the appropriate integrations by part we obtain the formula

$$\begin{aligned} &\mathcal{M}(u,\xi) = \langle \mathrm{D}\mathcal{F}(u)\xi, \mathcal{K}(u)\xi\rangle - \tfrac{1}{2}\langle\xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi\rangle \\ &= \int_\Omega \operatorname{div}(\mathbb{M}\nabla\xi)\operatorname{div}(u\mathbb{M}\nabla\xi) + \tfrac{1}{2}\operatorname{div}(\mathbb{M}\nabla u)\nabla\xi \cdot \mathbb{M}\nabla\xi\,\mathrm{d}x \\ &= \int_\Omega u\Big(\nabla\xi\cdot\mathbb{B}\nabla\xi + \nabla\xi\cdot\boldsymbol{B}{:}\mathrm{D}^2\xi + |\mathbb{M}\mathrm{D}^2\xi|^2\Big)\,\mathrm{d}x - \int_{\partial\Omega} u\mathbb{M}\nabla\big(\tfrac{1}{2}\nabla\xi\cdot\mathbb{M}\nabla\xi\big) \cdot \nu\,\mathrm{d}a, \end{aligned}$$

where all terms involving third derivatives of ξ cancel, and the tensors $\mathbb{B}$ and $\boldsymbol{B}$ are given via $\mathbb{M}$, $\mathrm{D}\mathbb{M}$, and $\mathrm{D}^2\mathbb{M}$. Proposition 4.2 can be generalized for spatially dependent mobilities leading to three additional terms due to the spatial derivatives of $\mathbb{M}$:

$$\begin{aligned}\mathbb{M}\nabla\big(\nabla\xi_1\cdot\mathbb{M}\nabla\xi_2\big)\cdot\nu = -\mathbb{I}(\mathbb{M}\nabla\xi_1,\mathbb{M}\nabla\xi_2) + \nabla\xi_1\cdot\mathrm{D}\mathbb{M}[\mathbb{M}\nu]\nabla\xi_2\\ - \nabla\xi_2\cdot\mathrm{D}\mathbb{M}[\mathbb{M}\nabla\xi_1]\nu - \nabla\xi_1\cdot\mathrm{D}\mathbb{M}[\mathbb{M}\nabla\xi_2]\nu.\end{aligned}$$

If the sum of these terms is negative, using $\alpha_0 > 0$ and $\mathbb{M}\in\mathrm{W}^{2,\infty}(\Omega)$ (giving $\boldsymbol{B},\mathbb{B}\in\mathrm{L}^\infty(\Omega)$) and pointwise minimization over $\mathrm{D}^2\xi(x)\in\mathbb{R}^{d\times d}_{\mathsf{sym}}$ provides a $\lambda_\mathbb{M}\in\mathbb{R}$ such that $\mathcal{M}(u,\xi)\geq\lambda_\mathbb{M}\langle\xi,\mathcal{K}(u)\xi\rangle = \lambda_\mathbb{M}\int_\Omega u\,\nabla\xi\cdot\mathbb{M}\nabla\xi\,\mathrm{d}x$.

For an isotropic mobility matrix $\mathbb{M}(x)=\mu(x)I$ satisfying the boundary relation $\mu\mathbb{I}(\tau,\tau)\geq\nabla\mu\cdot\nu|\tau|^2$ for all $x\in\partial\Omega$ and $\tau\in\mathrm{T}_x\partial\Omega$ we obtain the simplified estimate

$$\begin{aligned}\mathcal{M}(u,\xi) = \int_\Omega u\Big(&\nabla\xi\cdot\big((\tfrac12\mu\Delta\mu+\tfrac12|\nabla\mu|^2)I-\mu\mathrm{D}^2\mu\big)\nabla\xi\\ &+2\mu\nabla\xi\cdot\mathrm{D}^2\xi\nabla\mu-\mu\Delta\xi\nabla\mu\cdot\nabla\xi+\mu^2|\mathrm{D}^2\xi|^2\Big)\,\mathrm{d}x\\ &+\int_{\partial\Omega}u\big(\mu^2\mathbb{I}(\nabla\xi,\nabla\xi)-\mu(\nabla\mu\cdot\nu)|\nabla\xi|^2\big)\,\mathrm{d}x\\ \geq\int_\Omega u\Big(&\nabla\xi\cdot\big(\tfrac{\mu}{2}\Delta\mu I-\mu\mathrm{D}^2\mu\big)\nabla\xi-\tfrac{d-2}{4}\big(\nabla\xi\cdot\nabla\mu\big)^2\Big)\,\mathrm{d}x\ \geq\lambda_\mathbb{M}\langle\xi,\mathcal{K}(u)\xi\rangle,\end{aligned}\tag{4.8}$$

where $\lambda_\mathbb{M}=\inf\{\,\tfrac12\Delta\mu(x)-\sigma_{\mathsf{max}}\big(\mathrm{D}^2\mu(x)+\tfrac{d-2}{4\mu(x)}\nabla\mu(x)\otimes\nabla\mu(x)\big)\mid x\in\Omega\,\}$,

where again minimization with respect to $\mathrm{D}^2\xi$ as used in the first estimate. Here $\sigma_{\mathsf{max}}(H)\in\mathbb{R}$ denotes the largest eigenvalue of a symmetric matrix $H\in\mathbb{R}^{d\times d}$. In space dimensions $d=1$ and 2 we obtain

$$d=1:\quad\lambda_\mathbb{M}=\inf\{\,-\mu''(x)/2+(\mu'(x))^2/(4\mu(x))\mid x\in\Omega\,\},\tag{4.9a}$$

$$d=2:\quad\lambda_\mathbb{M}=\inf\{\,\tfrac12\big(\sigma_{\mathsf{min}}(\mathrm{D}^2\mu(x))-\sigma_{\mathsf{max}}(\mathrm{D}^2(\mu(x))\big)\mid x\in\Omega\,\}.\tag{4.9b}$$

Our result can be compared to the estimates obtained in [Lis09, Thm. 1.5] with complete different methods. The results there are formulated using the Wasserstein distance W_I, while our results are formulated in terms of $d_\mathcal{K}$ which is called W_G there, see [Lis09, Eqn. (1.67)]. Thus, our rate $\lambda_\mathbb{M}$ may differ from the contractivity rate α, which takes the form $\alpha=\inf\{\,-\sigma_{\mathsf{max}}\big(\mathrm{D}^2\mu(x)+d\nabla\sqrt{\mu}(x)\otimes\nabla\sqrt{\mu}(x)\big)\mid x\in\Omega\,\}$ in our smooth setting.

4.3 A scalar drift-diffusion equation with concave mobility

We now generalize the diffusion equation of the previous section by adding drift terms induced by a given potential V. Moreover, we allow the density u to be restricted to a bounded interval, i.e. we assume that there is a bound $U\in\,]0,\infty]$ such that

$$u(t,x)\in\,]0,U[\quad\text{a.e.}$$

Such restriction occur in systems with exclusion principles. We refer to [GaG05, BD*10] and Section 4.7. Our work relates to [CL*10] and [LMS12, Prop. 4.6], where the entropy and the potential energy are studied concerning their geodesic λ-convexity. We make the result of the latter work more precise. We have the total energy and the Onsager operator

$$\mathcal{E}(u) = \int_\Omega E(u(x)) + u(x)V(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{K}(u)\xi = -\operatorname{div}\big(\mu(u)\nabla\xi\big).$$

The drift-diffusion equation takes the form

$$\dot u = \operatorname{div}(a(u)\nabla u + \mu(u)\nabla V\big) \quad \text{in } \Omega, \quad (a(u)\nabla u + \mu(u)\nabla V)\cdot\nu = 0 \quad \text{on } \partial\Omega, \tag{4.10}$$

where $a(u) = \mu(u)E''(u)$. We again impose the sign conditions

$$\mu(u) > 0, \quad \mu''(u) \le 0, \quad E''(u) > 0 \quad \text{for all } u \in \,]0, U[. \tag{4.11}$$

In the case $U < \infty$ we explicitly allow for the case $\mu'(u) < 0$ which occurs in the commonly used mobility $\mu(u) = u - u^2$ on $]0, 1[$. We will see that the non-monotonicity of μ gives rise to new conditions. We emphasize that the following result does not need the condition $\nabla V \cdot \nu = 0$ on $\partial\Omega$ employed in [LMS12, Prop. 4.6].

Theorem 4.3 *Assume that Ω is a convex bounded domain in $\mathbb{R}^d$ with smooth boundary. In addition to* (4.11) *define $H(u) = \int_0^u \mu(y)\mu'(y)E''(y)\,\mathrm{d}y$ and assume*

$$H(u) \ge 0, \quad \mu(u)^2E''(u) \ge \tfrac{d-1}{d}H(u) \quad \textit{for all } u \in \,]0, U[. \tag{4.12}$$

If the potential $V : \Omega \to \mathbb{R}$ satisfies $V \in \mathrm{W}^{2,\infty}(\Omega)$, then $(\mathcal{E}, \mathcal{K})$ are geodesically λ-convex for $\lambda = \lambda_2^V - \lambda_1^V$, where

$$\begin{aligned}
\lambda_1^V &= \frac{9}{8}\|\nabla V\|_{\mathrm{L}^\infty}^2 \sup\{\, -\mu''(u)/E''(u) \mid u \in \,]0, U[\,\} \ \ge 0,\\
\lambda_2^V &= \inf\{\, \mu'(u)\,\boldsymbol{a}\cdot \mathrm{D}^2V(x)\boldsymbol{a} \mid u \in \,]0, U[, x \in \Omega, \boldsymbol{a} \in \mathbb{R}^d \textit{ with } |\boldsymbol{a}| = 1\,\}.
\end{aligned}$$

Before giving the proof of this result note that the case of a linear mobility (i.e. $\mu(u) = u$) for the Wasserstein distance gives the standard result as $\lambda_1^V = 0$. Moreover, λ_2^V simply characterizes the λ-convexity of V on the Euclidean space Ω. Note that in the case $\mu'(u) < 0$ we need λ-concavity of V.

Proof: We proceed exactly as in the previous subsection. We only have the new terms associated with V. Since $\mathcal{K}$ is independent of V and the vector field $\mathcal{F}$ depends linearly on V, the new terms are also linear in V. Together with $\mathcal{M}_0$ from (4.5) we have

$$\begin{aligned}
\mathcal{M}_V(u,\xi) &= \mathcal{M}_0(u,\xi)\\
&\quad + \int_\Omega \mu\mu'\nabla\xi\cdot\mathrm{D}^2V\nabla\xi + \frac{\mu\mu''}{2}\big(2\nabla u\cdot\nabla\xi\;\nabla V\cdot\nabla\xi - |\nabla\xi|^2\nabla V\cdot\nabla u\big)\,\mathrm{d}x.
\end{aligned} \tag{4.13}$$

To reach this result, we emphasize that the integrations by parts have to be done of the full vector field $\mathcal{F}$ such that $w = \mathcal{K}(u)\xi$ in $\int_\Omega \xi\mathrm{D}\mathcal{F}(u)[w]\,\mathrm{d}x$ satisfies the additional boundary condition $[w(a'(u)\nabla u + \mu'(u)\nabla V) + a(u)\nabla w]\cdot\nu = 0$ obtained by differentiating the boundary condition in (4.10).

While the first term in (4.13) can be immediately estimated from below by $\lambda_2^V\mu|\nabla\xi|^2$, the other terms do not have a sign. That is why in [CL*10] it was expected that the potential energy $\int_\Omega uV\,\mathrm{d}x$ is not geodesically convex. However, to estimate the geodesic convexity of $\mathcal{E}$ we can use the nonnegative term $-\mu''\frac{a}{2}|\nabla u|^2|\nabla\xi|^2$ occurring in $\mathcal{M}_0$ and not needed otherwise to show positivity of $\mathcal{M}_0$. Abbreviating $\boldsymbol{U} = \nabla u$ and $\boldsymbol{X} = \nabla\xi$, we have to estimate the following terms from below:

$$\begin{aligned} &-\mu''\tfrac{a}{2}|\boldsymbol{U}|^2|\boldsymbol{X}|^2 + \tfrac{\mu\mu''}{2}\big(2\boldsymbol{U}\cdot\boldsymbol{X}\,\nabla V\cdot\boldsymbol{X} - |\boldsymbol{X}|^2\nabla V\cdot\boldsymbol{U}\big) \geq (-\mu'')|\boldsymbol{X}|^2\big(\tfrac{a}{2}|\boldsymbol{U}|^2 - \tfrac{3}{2}\mu|\boldsymbol{U}||\nabla V|\big) \\ &\geq -(-\mu'')|\boldsymbol{X}|^2\tfrac{9\mu^2}{8a}|\nabla V|^2 = -\tfrac{(-\mu'')}{E''}|\nabla V|^2\mu|\boldsymbol{X}|^2 \geq -\lambda_1^V\mu|\boldsymbol{X}|^2. \end{aligned}$$

Thus, the result is established. ■

We conclude by making the conditions more explicit in the case of $\mu(u) = u - u^2$ on $]0,1[$ and $E''(u) = 1/m(u)$, i.e. $E(u) = u\log u + (1-u)\log(1-u)$. We obtain $\lambda_1^V = 9\|\nabla V\|_\infty^2/16$ and $\lambda_2^V = \|r_{\text{spec}}(\mathrm{D}^2V(\cdot))\|_\infty$, where r_{spec} denotes the spectral radius.

4.4 A scalar nonlinear reaction-diffusion equation

In a convex, bounded, and smooth domain Ω we consider the reaction-diffusion equation

$$\dot u = \Delta u - f(u) \quad \text{in } \Omega, \qquad \nabla u\cdot\nu = 0 \quad \text{on } \partial\Omega.$$

We assume that it is the gradient flow of the free energy $\mathcal{E}$ and the Onsager operator $\mathcal{K}$ defined via

$$\mathcal{E}(u) = \int_\Omega u(\log u - 1)\,\mathrm{d}x \qquad \text{and} \qquad \mathcal{K}(u)\xi = -\operatorname{div}(u\nabla\xi) + \kappa(u)\xi. \tag{4.14}$$

Hence, we assume the relation $f(u) = \kappa(u)\log u$. The reaction coefficient κ satisfies

$$\begin{aligned} &\kappa \in \mathrm{C}^0([0,\infty[) \cap \mathrm{C}^2(]0,\infty[), \\ &\kappa(0) = 0, \quad \kappa(u), \kappa'(u) > 0 \text{ and } \kappa''(u) \leq 0 \quad \text{for all } u > 0. \end{aligned} \tag{4.15}$$

The concavity of κ implies that of $u \mapsto \langle\xi, \mathcal{K}(u)\eta\rangle$, which is the prerequisite of the convexity of $d_{\mathcal{K}}^2$, see Remark 3.1.

Similar to the previous examples we introduce the spaces

$$\begin{aligned} &H^* = \mathrm{H}^1(\Omega), && H = \mathrm{H}_0^{-1}(\Omega), \\ &Y = \{\, u \in \mathrm{H}^{s-2}(\Omega) \mid \nabla u\cdot\nu = 0 \text{ on } \partial\Omega \,\}, && \mathcal{Y} = \{\, u \in Y \mid \inf u > 0 \,\}, \\ &Z = \mathrm{H}^s(\Omega) \cap Y, && \mathcal{Z} = \{\, u \in Z \cap \mathcal{Y} \mid \nabla(\Delta u - f(u))\cdot\nu = 0 \,\} \end{aligned}$$

and calculate $\mathcal{M}(u,\xi)$. With $\mathrm{D}\mathcal{F}(u)[v] = -\Delta v - f'(u)v$ and $v = \mathcal{K}(u)\xi$ we obtain

$$\mathcal{M}(u,\xi) = \int_\Omega \xi\big(-\Delta v + f'(u)v\big)\,\mathrm{d}x - \frac{1}{2}\int_\Omega \big(-\Delta u + f(u)\big)\big(|\nabla\xi|^2 + \kappa'(u)\xi^2\big)\,\mathrm{d}x = I_1 + I_2.$$

Integrating twice the first term in I_1 (using $\nabla\xi\cdot\nu = \nabla v\cdot\nu = 0$ on $\partial\Omega$) and inserting the definition of $v = \mathcal{K}(u)\xi$ we find

$$\begin{aligned} I_1 &= \int_\Omega \big(-\Delta\xi + f'(u)\xi\big)\big(-\operatorname{div}(u\nabla\xi) + \kappa(u)\xi\big)\,\mathrm{d}x \\ &= \int_\Omega -u\nabla\Delta\xi\cdot\nabla\xi + u\nabla\big(f'(u)\xi\big)\cdot\nabla\xi + \nabla\xi\cdot\big(\kappa(u)\xi\big) + f'(u)\kappa(u)\xi^2\,\mathrm{d}x \\ &= \int_\Omega -u\nabla\Delta\xi\cdot\nabla\xi + (uf' + \kappa)|\nabla\xi|^2 + (uf'' + \kappa')\xi\nabla\xi\cdot\nabla u + f'\kappa\xi^2\,\mathrm{d}x. \end{aligned}$$

Similarly, we integrate by parts the first term in I_2 (using $\nabla u \cdot \nu = 0$) and obtain

$$2I_2 = \int_\Omega -\nabla u \cdot \nabla\big(|\nabla\xi|^2{+}\kappa'\xi^2\big) - f|\nabla\xi|^2 - 2f\kappa'\xi\,\mathrm{d}x$$
$$= -\int_{\partial\Omega} u\nabla\big(|\nabla\xi|^2\big){\cdot}\nu\,\mathrm{da} + \int_\Omega u\Delta\big(|\nabla\xi|^2\big){-}f|\nabla\xi|^2{-}2\kappa'\xi\nabla\xi{\cdot}\nabla u{-}\big(\kappa''|\nabla u|^2{+}f\kappa'\big)\xi^2\,\mathrm{d}x.$$

Again using Proposition 4.2 we can estimate the boundary integral and obtain

$$\mathcal{M}(u,\xi) \geq \int_\Omega u|\mathrm{D}^2\xi|^2 + M_1(u)|\nabla\xi|^2 + m_2(u)\xi\,\nabla\xi{\cdot}\nabla u + \big(M_3(u){-}\kappa''|\nabla u|^2/2\big)\xi^2\,\mathrm{d}x$$

where $M_1(u) = uf'(u){+}\kappa(u){-}f(u)/2,\ m_2(u) = uf''(u),\ M_3(u) = f'(u)\kappa(u){-}f(u)\kappa'(u)/2.$

Using assumption (4.15) the last term, which involves $|\nabla u|^2\xi^2$ is nonnegative and can be dropped. We define $M_2(u) = f(0) + uf'(u) - f(u)$ such that $M_2'(u) = m_2(u)$ and $M_2(0) = 0$. The term involving m_2 can be integrated by parts (using $\nabla\xi\cdot\nu = 0$) via

$$\int_\Omega M_2'(u)\nabla u\cdot\xi\nabla\xi\,\mathrm{d}x = -\int_\Omega M_2(u)\Delta\big(\tfrac12\xi^2\big)\,\mathrm{d}x = -\int_\Omega M_2(u)\big(|\nabla\xi|^2{+}\xi\Delta\xi\big)\,\mathrm{d}x.$$

The pointwise estimate $-M_2(u)\xi\Delta\xi \geq -\frac{u}{d}(\Delta\xi)^2 - \frac{dM_2(u)2}{4u}\xi^2$ yields the lower estimate

$$\mathcal{M}(u,\xi) \geq \int_\Omega u\big(|\mathrm{D}^2\xi|^2{-}\tfrac1d(\Delta\xi)^2\big) + \big(M_1(u){-}M_2(u)\big)|\nabla\xi|^2 + \Big(M_3(u){-}\tfrac{dM_2(u)^2}{4u}\Big)\xi^2\,\mathrm{d}x.$$

Thus, we have established the following result.

Theorem 4.4 *Let Ω, κ, and M_j be given as above. Define the values*

$$\lambda_1^* = \inf\{\, \tfrac{M_1(u)-M_2(u)}{u} \mid u>0\,\} \quad \textit{and} \quad \lambda_2^* = \inf\{\, \tfrac{4uM_3(u)-dM_2(u)^2}{4u\kappa(u)} \mid u>0\,\},$$

and set $\lambda^* = \min\{\lambda_1^*, \lambda_2^*\}$. If $\lambda^* > -\infty$, then $(\mathcal{E},\mathcal{K})$ defined in (4.14) *is geodesically λ^*-convex.*

Proof: To conclude the proof we have to establish

$$\mathcal{M}(u,\xi) \geq \lambda^*\langle\xi,\mathcal{K}(u)\xi\rangle = \lambda^*\int_\Omega u|\nabla\xi|^2 + \kappa(u)\xi^2\,\mathrm{d}x$$

for all $u \in \mathcal{Z}$ and $\xi \in \mathcal{G}(u)Y$. Since the first term in the above lower estimate for $\mathcal{M}$ is nonnegative, it suffices to show $M_1(u){-}M_2(u) \geq \lambda^* u$ and $M_3(u) - dM_2(u)^2/(4u) \geq \lambda^*\kappa(u)$ for all $u \geq 0$. Since these estimates are exactly the definitions of λ_j^*, the desired result is established. ■

The following result provides sufficient conditions on the function κ, satisfying (4.15), that lead to a geodesically λ-convex gradient system. It is posed in terms of the ansatz $\kappa(u) = k(u)\Lambda(1,u)$ and shows that k can be chosen to be constant near $u = 0$ given the linear reaction term $f(u) = k(0)(u-1)$ there. For large u one may choose $k(u) = c(\log u)^p$ for $c>0$ and $p\in[0,1]$ leading to the nonlinear reaction term $f(u) = c(u-1)(\log u)^{p-1}$.

Proposition 4.5 *Consider a function κ satisfying* (4.15) *and let $k(u) = \kappa(u)/\Lambda(1,u) > 0$. If there exist $0 < u_0 < 1 < u_1 < \infty$ and positive constants k_j, $j = 0, \ldots, 3$ such that k satisfies the conditions*

$$k \in \mathrm{C}^0([0,\infty[) \text{ with } k(0) = k_0; \tag{4.16a}$$

$$\liminf_{u\to\infty} k(u) \geq k_1; \tag{4.16b}$$

$$k \in \mathrm{C}^1([u_1,\infty[) \text{ and } k \in \mathrm{C}^{1,\alpha}([0,u_0]) \text{ for some } \alpha \in \,]1/2,1]\,; \tag{4.16c}$$

$$k(u){+}uk'(u) \geq k_2 \text{ and } |k(u){+}u^2k'(u)|^2 \leq k_3u^2k(u)/\log u \text{ for } u \geq u_1, \tag{4.16d}$$

then in Theorem 4.4 we have $\lambda^* > -\infty$. The case $k \equiv k_0$ gives $\lambda^* = \frac{1}{2}\min\{k_0, k_0^2\} > 0$.

Proof: We denote by $\eta_j(u)$ the functions in the infima defining λ_j^* in Theorem 4.4. Since both functions are continuous on $]0,\infty[$ it suffices to estimate η_j near $u = 0$ and $u = \infty$.

ad η_1: By (4.16a) we have $M_1(0) - M_2(0) = -f(0)/2 = k_0/2 > 0$ and conclude $\eta_1(u) \geq 0$ for sufficiently small u. For $u \geq 2$ we have

$$M_1(u) - M_2(u) = \kappa(u) + f(u)/2 - f(0) \geq \tfrac{\kappa(u)}{2}\log u = \tfrac{u-1}{2}k(u) \geq uk(u)/4.$$

Using (4.16b) we obtain $\eta_1(u) \geq k_1/4$ for all sufficiently large u.

ad η_2: For $u \leq 1$ we have $f(u) \leq 0$; using $\kappa' \geq 0$ we conclude $M_3(u) \geq f'(u)\kappa(u)$. Moreover, from $f(u) = (u{-}1)k(u)$ and (4.16c) we conclude $f \in \mathrm{C}^{1,\alpha}([0,u_0])$. Hence, $M_2(u) = \int_0^u f'(u){-}f'(\nu)\,\mathrm{d}\nu$ satisfies $|M_2(u)| \leq Cu^{1+\alpha}$. Together we find

$$\eta_2(u) \geq f'(u) - \tfrac{d}{4}C^2u^{2\alpha-1}/\kappa(u) \geq f'(0) - Cu^\alpha - \tfrac{d}{4}C^2\tfrac{u^{2\alpha-1}|\log u|}{(1-u)k(u)} \geq \lambda_2^- \text{ on }]0,u_0]\,.$$

For large u we use the asymptotics for $u \to \infty$ given via

$$M_2(u) \approx k(u) + u^2k'(u) \text{ and } M_3(u) \approx \tfrac{uk(u)}{2\log u}\big(k(u){+}uk'(u)\big)..$$

Using (4.16d) we find $\eta_2(u) \geq k_2/4 - dk_3/2$ which gives the desired result.

For the last statement note that $M_1(u) = k_0(u{+}1)/2 + k_0\kappa(u) \geq k_0u/2$, $M_2 \equiv 0$, and $M_3(u) = k_0^2(\kappa(u){-}(u{-}1)\kappa'(u)/2) \geq k_0^2\kappa(u)/2$. The latter estimate follows from the explicit relation $(u{-}1)\kappa'(u) = (1{-}\kappa(u)/u)\kappa(u)$, cf. [Mie11b, (A.3)]. ■

4.5 A linear reaction-diffusion system

For $\boldsymbol{u} = (u_1,u_2)$ we consider the system of coupled linear equations

$$\begin{pmatrix}\dot u_1\\ \dot u_2\end{pmatrix} = \begin{pmatrix}\delta_1\Delta u_1\\ \delta_2\Delta u_2\end{pmatrix} + k\begin{pmatrix}u_2{-}u_1\\ u_1{-}u_2\end{pmatrix} \quad \text{in } \Omega, \qquad \nabla u_1\cdot\nu = \nabla u_2\cdot\nu = 0 \quad \text{on } \partial\Omega, \tag{4.17}$$

which is the gradient flow for the energy $\mathcal{E}$ and the Onsager operator $\mathcal{K}$ given via

$$\begin{aligned}\mathcal{E}(\boldsymbol{u}) &= \int_\Omega u_1\log(u_1{-}1) + u_2\log(u_2{-}1)\,\mathrm{d}x \qquad \text{and}\\ \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} &= \begin{pmatrix}-\operatorname{div}(u_1\delta_1\nabla\xi_1)\\ -\operatorname{div}(u_2\delta_2\nabla\xi_2)\end{pmatrix} + k\Lambda\,(u_1,u_2)\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}.\end{aligned} \tag{4.18}$$

Observe that the total mass $\mathcal{Q}(u_1,u_2) = \int_\Omega u_1{+}u_2 \,\mathrm{d}x$ is conserved along solutions of (4.17), i.e., $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}(u_1,u_2) = 0$. We fix a constant state $\boldsymbol{u}^* = (u_1^*, u_2^*) \in {]0,\infty[}^2$, choose the Sobolev index s as before, and define the spaces

$$
\begin{aligned}
&H^* = \{\, \boldsymbol{\xi} \in \mathrm{H}^1(\Omega){\times}\mathrm{H}^1(\Omega) \mid \textstyle\int_\Omega \xi_1{+}\xi_2 \,\mathrm{d}x = 0 \,\},\\
&Y = \{\, \boldsymbol{v} \in \mathrm{H}^{s-2}(\Omega){\times}\mathrm{H}^{s-2}(\Omega) \mid \textstyle\int_\Omega v_1{+}v_2 \,\mathrm{d}x = 0 \text{ and } \nabla v_1{\cdot}\nu{=}\nabla v_2{\cdot}\nu{=}0 \,\},\\
&\mathcal{Y} = \{\, \boldsymbol{u} \in \boldsymbol{u}^*{+}Y \mid \inf u_i > 0,\ i=1,2 \,\},\\
&Z = (H^s(\Omega){\times}\mathrm{H}^s(\Omega)) \cap Y, \quad \mathcal{Z} = \{\, \boldsymbol{u} \in (\boldsymbol{u}^*{+}Z) \cap \mathcal{Y} \mid \nabla(\Delta u_i)\cdot\nu = 0,\ i=1,2 \,\}.
\end{aligned}
$$

Since $\boldsymbol{u} \mapsto \mathcal{F}(\boldsymbol{u}) = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u})$ is linear we compute

$$
\mathrm{D}\mathcal{F}(\boldsymbol{u})[\boldsymbol{v}] = \big({-}\delta_1\Delta v_1 + k(v_1{-}v_2)\,,\ {-}\delta_2\Delta v_2 + k(v_2{-}v_1)\big)^{\mathsf{T}}.
$$

With the shorthand $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}$ we obtain $\mathcal{M}(\boldsymbol{u},\boldsymbol{\xi}) = I_1 + I_2$ with

$$
\begin{aligned}
I_1 &= \int_\Omega \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\cdot\begin{pmatrix}-\delta_1\Delta v_1 + k(v_1{-}v_2)\\ -\delta_2\Delta v_2 + k(v_2{-}v_1)\end{pmatrix}\mathrm{d}x \quad \text{and}\\
I_2 &= -\frac12\int_\Omega \big(\delta_1|\nabla\xi_1|^2{+}k\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\big)\big({-}\delta_1\Delta u_1{+}k(u_1{-}u_2)\big)\,\mathrm{d}x\\
&\quad -\frac12\int_\Omega \big(\delta_2|\nabla\xi_2|^2{+}k\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\big)\big({-}\delta_2\Delta u_2{+}k(u_2{-}u_1)\big)\,\mathrm{d}x
\end{aligned}
$$

Integrating the first term in I_1 by parts twice, using the boundary conditions $\nabla v_i\cdot\nu = \nabla\xi_i\cdot\nu = 0$, and finally substituting $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}$ gives

$$
\begin{aligned}
I_1 &= \int_\Omega \begin{pmatrix}\delta_1\Delta\xi_1{-}k(\xi_1{-}\xi_2)\\ \delta_2\Delta\xi_2{-}k(\xi_2{-}\xi_1)\end{pmatrix}\cdot\begin{pmatrix}\mathrm{div}(u_1\delta_1\nabla\xi_1){-}k\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)\\ \mathrm{div}(u_2\delta_2\nabla\xi_2){-}k\Lambda(\boldsymbol{u})(\xi_2{-}\xi_1)\end{pmatrix}\mathrm{d}x\\
&= \int_\Omega \Big[-\delta_1^2 u_1\nabla\Delta\xi_1{\cdot}\nabla\xi_1 - \delta_2^2 u_2\nabla\Delta\xi_2{\cdot}\nabla\xi_2 + k\nabla(\xi_1{-}\xi_2){\cdot}(\delta_1 u_1\nabla\xi_1{-}\delta_2 u_2\nabla\xi_2)\\
&\qquad - k\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)(\delta_1\Delta\xi_1{-}\delta_2\Delta\xi_2) + 2k^2\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\Big]\,\mathrm{d}x.
\end{aligned}
$$

Similarly, we integrate the second term and obtain

$$
\begin{aligned}
2I_2 &= \int_\Omega \Big[-\delta_1^2\nabla u_1\cdot\nabla\big(|\nabla\xi_1|^2\big) + \Big(k\delta_1\Delta u_1{-}k^2(u_1{-}u_2)\Big)\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\\
&\qquad - \delta_2^2\nabla u_2\cdot\nabla\big(|\nabla\xi_2|^2\big) + \big(k\delta_2\Delta u_2{-}k^2(u_2{-}u_1)\big)\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\Big]\,\mathrm{d}x\\
&= \int_\Omega \Big[\delta_1^2 u_1\Delta\big(|\nabla\xi_1|^2\big) + \Big(k\delta_1\Delta u_1{-}k^2(u_1{-}u_2)\Big)\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\\
&\qquad + \delta_2^2 u_2\Delta\big(|\nabla\xi_2|^2\big) + \big(k\delta_2\Delta u_2{-}k^2(u_2{-}u_1)\big)\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\Big]\,\mathrm{d}x\\
&\quad - \int_{\partial\Omega} \delta_1 u_1\nabla\big(|\nabla\xi_1|^2\big) + \delta_2 u_2\nabla\big(|\nabla\xi_2|^2\big)\,\mathrm{d}a.
\end{aligned}
$$

Thus, using again Bochner's formula and Proposition 4.2 we arrive at

$$
I_1 + I_2 \geq \int_\Omega \delta_1 u_1|\mathrm{D}^2\xi_1|^2 + \delta_2^2 u_2|\mathrm{D}^2\xi_2|^2 + k^2 m(\boldsymbol{u})(\xi_1{-}\xi_2)^2 + kG(\delta_1,\delta_2,\boldsymbol{u},\boldsymbol{\xi})\,\mathrm{d}x
$$
$$
\text{with } m(\boldsymbol{u}) = 2\Lambda(u_1,u_2){-}\tfrac12\big(\partial_{u_1}\Lambda(u_1,u_2){-}\partial_{u_2}\Lambda(u_1,u_2)\big)(u_1{-}u_2).
$$

It was shown in [Mie11b, Example 3.5] that $m(\boldsymbol{u}) \geq 2\Lambda(\boldsymbol{u}) \geq 0$ holds.

The main task is to control the mixed terms with prefactor $k\delta_j$ that are collected in the function G. Unfortunately we can estimate these terms only in the case of equal mobilities $\delta_j = \delta > 0$. For $G(\boldsymbol{u}, \boldsymbol{\xi}) = \frac{2}{\delta} G(\delta, \delta, \boldsymbol{u}, \boldsymbol{\xi})$ some rearrangements yield the identity

$$G(\boldsymbol{u}, \boldsymbol{\xi}) = (\xi_1 - \xi_2)^2 \big(\Delta\Lambda(\boldsymbol{u}) - \mathbb{L}(\boldsymbol{u}) \big) - 2\Lambda(\boldsymbol{u})(\xi_1 - \xi_2)\Delta(\xi_1 - \xi_2) + (u_1 + u_2)|\nabla(\xi_1 - \xi_2)|^2$$

where $\mathbb{L}(\boldsymbol{u}) = \partial^2_{u_1}\Lambda(\boldsymbol{u})|\nabla u_1|^2 + 2\partial_{u_1}\partial_{u_2}\Lambda(\boldsymbol{u})\nabla u_1 \cdot \nabla u_2 + \partial^2_{u_2}\Lambda(\boldsymbol{u})|\nabla u_2|^2$.

Since Λ is a concave function, we have $\mathbb{L}(\boldsymbol{u}) \leq 0$. To estimate $\int_\Omega G \, \mathrm{d}x$ we integrate by parts the very first term twice (using $\nabla\boldsymbol{\xi} \cdot \nu = 0$ and $\nabla\Lambda(\boldsymbol{u}) \cdot \nu = 0$) and find

$$\int_\Omega G(\boldsymbol{u}, \boldsymbol{\xi}) \, \mathrm{d}x = \int_\Omega \big(2\Lambda(\boldsymbol{u}) + u_1 + u_2 \big) |\nabla(\xi_1 - \xi_2)|^2 - (\xi_1 - \xi_2)^2 \mathbb{L}(\boldsymbol{u}) \, \mathrm{d}x \geq 0.$$

Hence, we have established the following result.

Theorem 4.6 *If Ω is smooth and convex and $\delta_1 = \delta_2 > 0$, then the gradient system* (4.17) *generated by $(\mathcal{E}, \mathcal{K})$ from* (4.18) *is geodesically 0-convex.*

4.6 Drift-diffusion system in 1D

We consider the one-dimensional version of the drift-reaction-diffusion system (2.10) for electron and holes in a semiconductor, see Section 2.4. We further simplify the system be neglecting the reaction terms $(np - 1)$.

To highlight the general structure we treat a system with I nonnegative densities $u_i \in \mathrm{L}^1(\Omega)$ with $\Omega = \,]0, 1[$, where the species have the charge vector $\boldsymbol{q} = (q_i)_{i=1,\dots,I} \in \mathbb{Z}^I$. The system takes the form

$$0 = (\varepsilon\phi'_{\boldsymbol{u}})' + \boldsymbol{q} \cdot \boldsymbol{u}, \qquad \text{in } \Omega, \qquad \phi_{\boldsymbol{u}}(0) = 0 = \phi'_{\boldsymbol{u}}(1); \tag{4.19a}$$

$$\dot{u}_i = \big[\mu_i \big(u'_i + u_i V'_i + q_i u_i \phi'_{\boldsymbol{u}} \big) \big]' \qquad \text{in } \Omega; \tag{4.19b}$$

$$0 = \mu_i \big(u'_i + u_i V'_i + q_i u_i \phi'_{\boldsymbol{u}} \big) \qquad \text{for } x \in \{0, 1\}, \tag{4.19c}$$

where $'$ is the partial derivative with respect to x. The potentials $\boldsymbol{V} = (V_1, \dots, V_I)$ are smooth functions and contain possible doping terms. The system is the gradient flow for

$$\mathcal{E}(\boldsymbol{u}) = \int_0^1 \sum_{i=1}^I u_i(\log u_i + V_i) + \frac{\varepsilon}{2}|\phi'_{\boldsymbol{u}}|^2 \, \mathrm{d}x \quad \text{and} \quad \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = -\big(\mu_i u_i \xi'_i \big)'_{i=1,\dots,I}. \tag{4.20}$$

Since we have no reaction between the species and no-flux boundary conditions the individual masses $\int_0^1 u_i \, \mathrm{d}x$ are conserved. The electrostatic potential $\phi_{\boldsymbol{u}}$ is a linear function of $\boldsymbol{q} \cdot \boldsymbol{u}$, viz. $\phi_{\boldsymbol{u}} = L\boldsymbol{q} \cdot \boldsymbol{u}$. In the one-dimensional case we have an explicit solution formula:

$$\big(\, \phi = Lg, \ g = \gamma', \ \gamma(1) = 0 \, \big) \implies \phi' = -\gamma/\varepsilon. \tag{4.21}$$

The function spaces can be introduced as in the above examples. We only give the calculation of the operator $\mathcal{M}$, where now the quadratic nature of $\mathcal{F}$ due to the terms $u_i\phi'_{\boldsymbol{u}}$ has to be observed. Using the two boundary conditions for $\boldsymbol{\xi} = \mathcal{G}(\boldsymbol{u})\boldsymbol{v}$ we find

$$(\mathrm{D}\mathcal{F}(\boldsymbol{u})^*\boldsymbol{\xi})_i = -\mu_i \xi''_i + \mu_i \big(V'_i + q_i \phi'_{\boldsymbol{u}} \big) \xi'_i - q_i L g, \ \text{where } g = \textstyle\sum_{j=1}^I \mu_j q_j \big(u_j \xi'_j \big)' = -\boldsymbol{q} \cdot \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}.$$

Now the quadratic form $\mathcal{M}(\boldsymbol{u},\boldsymbol{\xi}) = I_1 + I_2$ can be calculated as usual:

$$I_1 = \sum_{i=1}^{I} \int_0^1 (\mathrm{D}\mathcal{F}(\boldsymbol{u})^*\boldsymbol{\xi})_i(\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi})_i \,\mathrm{d}x = \int_0^1 \sum_1^I \mu_i^2\Big(-u_i\xi_i'''\xi_i' + u_i\xi_i'\big((V_i'+q_i\phi_{\boldsymbol{u}}')\xi_i'\big)'\Big) + gLg\,\mathrm{d}x,$$
$$I_2 = -\tfrac{1}{2}\sum_{i=1}^{I}\int_0^1 \mu_i^2(\xi_2')^2\mathcal{F}(\boldsymbol{u})_i\,\mathrm{d}x = \sum_{i=1}^{I}\mu_i^2\int_0^1 u_i\big(\xi_i'''\xi_i'+(\xi_i'')^2\big) - \xi_i'\xi_i''u_i(V_i'+q_i\phi_{\boldsymbol{u}}')\,\mathrm{d}x,$$

where used the boundary conditions $\xi_i' = 0$ on $\partial\Omega$. Combining the two terms and using some cancellation we arrive at

$$\mathcal{M}(\boldsymbol{u},\boldsymbol{\xi}) = \int_0^1 \sum_{i=1}^{I} \mu_i^2 u_i\big((\xi_i'')^2 + V_i''(\xi_i')^2\big) + h_{\boldsymbol{\xi}}\phi_{\boldsymbol{u}}'' + gLg\,\mathrm{d}x \ \text{ with } h_{\boldsymbol{\xi}} = \sum_{i=1}^{I}\mu_i^2 q_i u_i(\xi_i')^2.$$

The first two terms can be estimated in the standard way. For the interaction via $\phi_{\boldsymbol{u}}$ and L we note that g is such that formula (4.21) can be applied. When assuming additionally that $\varepsilon \equiv \varepsilon_0$ the third and fourth term can be rewritten as

$$Q_{\boldsymbol{u}}(\boldsymbol{\xi}) = h_{\boldsymbol{\xi}}\phi_{\boldsymbol{u}}'' + gLg = \frac{1}{\varepsilon_0}\Big(-h_{\boldsymbol{\xi}}\,\boldsymbol{q}\cdot\boldsymbol{u} + \big(\textstyle\sum_{i=1}^{I}\mu_j q_j u_j \xi_j'\big)^2\Big)^2.$$

There are two cases in which this quadratic form can be estimated from below. First, in the case $I = 1$ we obviously have $Q_{\boldsymbol{u}} \equiv 0$. For $I = 2$ the expression simplifies to

$$Q_{\boldsymbol{u}}(\boldsymbol{\xi}) = -q_1q_2u_1u_2\big(\mu_1\xi_1' - \mu_2\xi_2'\big)^2.$$

Thus, we find $Q_{\boldsymbol{u}} \geq 0$ if $q_1q_2 \leq 0$, this means that the particles are oppositely charged. Of course, we could add further uncharged particles (i.e. $q_j = 0$), but this is useless as they do not interact with the other particles. We summarize our findings as follows.

Theorem 4.7 *Consider the gradient system $(\mathcal{E},\mathcal{K})$ defined via* (4.19) *and* (4.20) *with constant ε. Assume either $I = 1$ or $I = 2$ and $q_1q_2 < 0$. If the potentials V_i are λ_i-convex, i.e. $V_i'' \geq \lambda_i$ on Ω, then, $(\mathcal{E},\mathcal{K})$ is geodesically λ^*-convex with $\lambda^* = \min\{\,\mu_i\lambda_i \mid i = 1,\dots,I\,\}$.*

4.7 A multi-particle system with cross-diffusion

In several applications one is interested in reaction-diffusion systems with I species, where the microscopic sites are occupied exactly by one species. We refer to [Gri04, BD*10]. On the macroscopic level this means that the density vector $\boldsymbol{u} = (u_1,\dots,u_I)$ satisfies the pointwise restriction

$$\boldsymbol{u}(x)\cdot\boldsymbol{e} = \sum_{i=1}^{I} u_i(x) = 1 \quad \text{a.e. in } \Omega. \tag{4.22}$$

Moreover, the mobility tensor obeys the Stefan-Maxwell law (see e.g. [Gri04])

$$\mathbb{M}(\boldsymbol{u}) = \mathrm{diag}(\boldsymbol{u}) - \boldsymbol{u}\otimes\boldsymbol{u} = \begin{pmatrix} u_1 - u_1^2 & -u_1u_2 & \cdots & -u_1u_I \\ -u_1u_2 & u_2 - u_2^2 & \cdots & -u_2u_I \\ \vdots & & \ddots & \vdots \\ -u_1u_I & -u_2u_I & \cdots & u_I - u_I^2 \end{pmatrix} \tag{4.23}$$

Using (4.22) we easily see that $\mathbb{M}$ is positive semidefinite, namely

$$\boldsymbol{a}\cdot\mathbb{M}(\boldsymbol{a})\boldsymbol{\xi} = \sum_{i=1}^I u_i a_i^2 - (\boldsymbol{u}\cdot\boldsymbol{a})^2 = \sum_{i=1}^I u_i\big(a_i - \boldsymbol{u}\cdot\boldsymbol{a}\big)^2 \;\geq 0. \tag{4.24}$$

Thus, we consider the energy functional

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(\boldsymbol{u}) + \boldsymbol{u}\cdot\boldsymbol{V}\,\mathrm{d}x, \quad \text{where } E(\boldsymbol{u}) = \sum_{i=1}^I u_i(\log u_i - 1) \tag{4.25a}$$

and $\boldsymbol{V} = (V_1,\ldots,V_I)$ is a vector of potentials with $\boldsymbol{V}\cdot\boldsymbol{e}\equiv 0$. Thus, $\boldsymbol{V}$ determines the equilibrium state $\boldsymbol{w}$ via $w_i = \mathrm{e}^{-V_i}$. Moreover, the Onsager operator acts now on the vector-valued dual variables $\boldsymbol{\xi}\in H^* = \{\,\boldsymbol{\xi}\in \mathrm{H}^1_{\mathrm{av}}(\Omega)^I \mid \boldsymbol{\xi}\cdot\boldsymbol{e}\equiv 0\,\}$ and takes the form

$$\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = \Big(-\operatorname{div}\big(u_i(\nabla\xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\big)\Big)_{i=1,\ldots,I} \quad \text{where } \boldsymbol{\Xi}_{\boldsymbol{u}} = \sum_{j=1}^I u_j\nabla\xi_j. \tag{4.25b}$$

Taking into account the constraint (4.22) when calculating the differentials we find the nonlinear evolutionary systems

$$\dot{\boldsymbol{u}} = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u}) = \Delta\boldsymbol{u} + \Big(\operatorname{div}\big(u_i(\nabla V_i - \boldsymbol{G}_{\boldsymbol{u}})\big)\Big)_{i=1,\ldots,I} \quad \text{where } \boldsymbol{G}_{\boldsymbol{u}} = \sum_{j=1}^I u_j\nabla V_j.$$

Here the diffusion term is linear since $\mathbb{M}(\boldsymbol{u})$ is exactly the inverse of $\mathrm{D}^2E(\boldsymbol{u})$ (taking the constraint into account). We see that the special choice of $\mathbb{M}$ with negative off-diagonal terms simplifies the diffusion terms, while the drift terms from the potential become more involved. This approach was also used in [Gri04, GaG05], while in [BD*10] the off-diagonal terms are not used.

In particular, the mass of each component is preserved during the flow, namely

$$\int_\Omega \boldsymbol{u}(t,x)\,\mathrm{d}x = \int_\Omega \boldsymbol{u}(0,x)\,\mathrm{d}x = \boldsymbol{m}\in\,]0,\infty[^I \quad \text{with } \boldsymbol{m}\cdot\boldsymbol{e} = \mathrm{vol}(\Omega).$$

In the case $I=2$ the system reduces to a scalar equation for $u\in[0,1]$ via $\boldsymbol{u}=(u,1-u)$ of the form

$$\dot{u} = \Delta u + \operatorname{div}\big((u-u^2)\nabla V\big) \quad \text{where } 2V = V_1 = -V_2,$$

which is covered by the analysis treated in Section 4.3.

We now restrict to the case $\boldsymbol{V}\equiv 0$ and leave the general case for future research. Our aim is to show that the pure (uncoupled) diffusion is geodesically 0-convex. This statement is nontrivial since the metric $d_\mathcal{K}$ induced by the mobility tensor $\mathbb{M}$ couples the densities in a nontrivial way. However, since $\mathbb{M}(\boldsymbol{u})$ can be estimated from above by $\mathbb{M}_\mathrm{W}(\boldsymbol{u}) = \mathrm{diag}(\boldsymbol{u})\in\mathbb{R}^{I\times I}$ we see that $d_\mathcal{K}$ can be estimated from above by the componentwise Wasserstein distance, i.e.

$$d_\mathcal{K}(\boldsymbol{u}^1,\boldsymbol{u}^2)^2 \leq d_\mathrm{W}(\boldsymbol{u}^1,\boldsymbol{u}^2)^2 = \sum_{i=1}^I d_\mathrm{W}(u_i^1,u_i^2)^2.$$

Theorem 4.8 *Consider the gradient system $(\mathcal{E},\mathcal{K})$ defined in* (4.25) *with $\boldsymbol{V}\equiv 0$. Then, $\mathcal{E}$ is geodesically 0-convex with respect to $d_\mathcal{K}$.*

To estimate the quadratic form $\mathcal{M}$ we assume as usual that Ω is a convex domain with smooth boundary and define the spaces Z, Y, and H as before in the Sobolev space H^s, H^{s-2} and H^1, respectively. Moreover, for the functions $\boldsymbol{u}$, $\boldsymbol{\xi}$, and $\boldsymbol{v} = (v_1, \ldots, v_I) = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} \in Y$ we have the following boundary conditions:

$$\text{(a) } \nabla u_i \cdot \nu = 0, \quad \text{(b) } \nabla \xi_i \cdot \nu = 0, \quad \text{(c) } \nabla v_i \cdot \nu = 0. \tag{4.26}$$

Using $\mathcal{F}(\boldsymbol{u}) = -\Delta \boldsymbol{u}$ and $\boldsymbol{\xi} \in \mathcal{G}(\boldsymbol{u})Y$ giving $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} \in Y$, we have

$$\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) = \sum_{i=1}^{I} \int_\Omega \xi_i(-\Delta v_i) - \frac{1}{2}(-\Delta u_i)\big(|\nabla \xi_i|^2 - 2\nabla \xi_i \cdot \boldsymbol{\Xi}_{\boldsymbol{u}}\big)\,\mathrm{d}x.$$

Using (c) and (b) we can integrate by parts the first term twice. The second term will be integrated once using (a). After inserting the definition of $\boldsymbol{v}$ we arrive at

$$\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) = \sum_{i=1}^{I} \int_\Omega \Delta \xi_i \ \mathrm{div}\,\big(u_i(\nabla \xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\big) - \nabla u_i \cdot \nabla\big(\tfrac{1}{2}|\nabla \xi_i|^2\big) + \nabla u_i \cdot \nabla(\xi_i \cdot \boldsymbol{\Xi}_{\boldsymbol{u}})\big)\,\mathrm{d}x..$$

The first term will now be integrated by part once again by using (b), which also implies $\boldsymbol{\Xi}_{\boldsymbol{u}} \cdot \nu = 0$. Integrating the second term will generate a boundary integral that will be nonnegative by Proposition 4.2:

$$\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) = \int_\Omega \sum_{i=1}^{I} u_i\big(-\nabla \Delta \xi_i \cdot \nabla \xi_i + \Delta\big(\tfrac{1}{2}|\nabla \xi_i|^2\big)\big) + \mu(\boldsymbol{u}, \boldsymbol{\xi})\,\mathrm{d}x + \beta^1_{\partial\Omega}, \text{ where}$$

$$\mu(\boldsymbol{u}, \boldsymbol{\xi}) = \sum_{i,j=1}^{I} \sum_{\alpha,\beta=1}^{d} \big(u_i u_j \xi_{i\alpha\alpha\beta}\xi_{j\beta} + u_{i\beta}u_{j\beta}\xi_{i\alpha}\xi_{j\alpha} + u_{i\beta}u_j\xi_{i\alpha\beta}\xi_{j\alpha} + u_{i\beta}u_j\xi_{i\alpha}\xi_{j\alpha\beta}\big)$$

$$\text{and } \beta^1_{\partial\Omega} = \int_\Omega \sum_{i=1}^{I} u_i \mathbb{I}(\nabla_\| \xi_i, \nabla_\| \xi_i)\,\mathrm{d}a \geq 0.$$

Here the indices α and β denote partial derivatives with respect to x_α.

The first term in $\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi})$ is positive by Bochner's identity. To estimate μ we interchange the summation indices i and j in the fourth term to find that the last two terms can be combined into $(u_i u_j)_\beta \xi_{i\alpha\beta}\xi_{j\alpha}$. Thus, integration by parts, employing Proposition 4.2, and exploiting the cancellation of the terms involving $\xi_{i\alpha\alpha\beta}$ gives

$$\textstyle\int_\Omega \mu(\boldsymbol{u}, \boldsymbol{\xi})\,\mathrm{d}x = \int_\Omega |\nabla \boldsymbol{u}^T \nabla \boldsymbol{\xi}|^2 - \big|\sum_{i=1}^{I} u_i \mathrm{D}^2 \xi_i\big|^2\,\mathrm{d}x + \beta^2_{\partial\Omega},$$

where $\beta^2_{\partial\Omega} = \int_{\partial\Omega} -\mathbb{I}(\boldsymbol{\Xi}_{\boldsymbol{u}}, \boldsymbol{\Xi}_{\boldsymbol{u}})\,\mathrm{d}x$. Here we used the boundary conditions (b), which give $\boldsymbol{\Xi}_{\boldsymbol{u}} \cdot \nu = 0$ and hence $\sum_{i=1}^{I} u_i \nabla_\| \xi_i = \boldsymbol{\Xi}_{\boldsymbol{u}}$ on $\partial\Omega$. The first term in the above integral is nonnegative, while the other two terms are nonpositive. However, they are dominated by the corresponding positive terms obtained earlier, e.g. $\beta^1_{\partial\Omega} + \beta^2_{\partial\Omega} \geq 0$. Using the same rearrangement as in (4.24) we find the final expression

$$\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) = \int_\Omega \sum_{i=1}^{I} u_i |\mathrm{D}^2 \xi_i - \boldsymbol{H}|^2 + |\nabla \boldsymbol{u}^T \nabla \boldsymbol{\xi}|^2\,\mathrm{d}x + \int_{\partial\Omega} \sum_{i=1}^{I} u_i \mathbb{I}(\nabla_\| \xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}}, \nabla_\| \xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\,\mathrm{d}a,$$

where $\boldsymbol{H} = \sum_{i=1}^{I} u_i \mathrm{D}^2 \xi_i$. Thus, we have established the desired result $\mathcal{M}(\boldsymbol{u}, \boldsymbol{\xi}) \geq 0$.

Acknowledgments. The authors are grateful to Giuseppe Savaré for stimulating discussions and several helpful remarks.

References

[AGH02] G. ALBINUS, H. GAJEWSKI, and R. HÜNLICH. Thermodynamic design of energy models of semiconductor devices. *Nonlinearity*, 15(2), 367–383, 2002.

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

[AM*01] A. ARNOLD, P. MARKOWICH, G. TOSCANI, and A. UNTERREITER. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26(1-2), 43–100, 2001.

[Ama93] H. AMANN. Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In H.-J. Schmeisser and et al., editors, *Function spaces, differential operators and nonlinear analysis*, Teubner-Texte Math. 133, pages 9–126. Teubner, Stuttgart, 1993.

[AmS08] L. AMBROSIO and S. SERFATY. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11), 1495–1539, 2008.

[BaÉ85] D. BAKRY and M. ÉMERY. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[Bak94] D. BAKRY. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[BD*10] M. BURGER, M. DI FRANSCESCO, J.-F. PIETSCHMANN, and B. SCHLAKE. Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Analysis*, 42(6), 2842–2871, 2010.

[BeB00] J.-D. BENAMOU and Y. BRENIER. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3), 375–393, 2000.

[BoP11] D. BOTHE and M. PIERRE. The instantaneous limit for reaction-diffusion systems with a fast irreversible reaction. *Discr. Cont. Dynam. Systems Ser. S*, 8(1), 49–59, 2011.

[CL*10] J. A. CARRILLO, S. LISINI, G. SAVARÉ, and D. SLEPČEV. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4), 1273–1309, 2010.

[DaS08] S. DANERI and G. SAVARÉ. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Analysis*, 40, 1104–1122, 2008.

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

[DeF06] L. DESVILLETTES and K. FELLNER. Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations. *J. Math. Anal. Appl.*, 319(1), 157–176, 2006.

[DeF07] L. DESVILLETTES and K. FELLNER. Entropy methods for reaction-diffusion systems. In *Discrete Contin. Dyn. Syst. (suppl). Dynamical Systems and Differential Equations. Proceedings of the 6th AIMS International Conference*, pages 304–312, 2007.

[DeF08] L. DESVILLETTES and K. FELLNER. Entropy methods for reaction-diffusion equations with degenerate diffusion arising in reversibly chemistry. In *Proceedings of EQUADIFF 2007*, page To appear, 2008.

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

[DF*07] L. DESVILLETTES, K. FELLNER, M. PIERRE, and J. VOVELLE. Global existence for quadratic systems of reaction-diffusion. *Adv. Nonlinear Stud.*, 7(3), 491–511, 2007.

[ErM11] M. ERBAR and J. MAAS. Ricci curvature of finite Markov chains via convexity of the entropy. *arXiv: 1111.2687*, 2011.

[FiG10] A. FIGALLI and N. GIGLI. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *J. Math. Pures Appl.*, 94(2), 107–130, 2010.

[GaG05] H. GAJEWSKI and K. GÄRTNER. On a nonlocal model of image segmentation. *Z. angew. Math. Mech. (ZAMM)*, 56(4), 572–591, 2005.

[GiM04] V. GIOVANGIGLI and M. MASSOT. Entropic structure of multicomponent reactive flows with partial equilibrium reduced chemistry. *Math. Methods Appl. Sci. (MMAS)*, 27(7), 739–768, 2004.

[GlG09] A. GLITZKY and K. GÄRTNER. Energy estimates for continuous and discretized electro-reaction-diffusion systems. *Nonlinear Anal.*, 70(2), 788–805, 2009.

[GlH05] A. GLITZKY and R. HÜNLICH. Global existence result for pair diffusion models. *SIAM J. Math. Analysis*, 36(4), 1200–1225 (electronic), 2005.

[Gli08] A. GLITZKY. Exponential decay of the free energy for discretized electro-reaction-diffusion systems. *Nonlinearity*, 21(9), 1989–2009, 2008.

[Gli09] A. GLITZKY. Energy estimates for electro-reaction-diffusion systems with partly fast kinetics. *Discr. Cont. Dynam. Systems Ser. A*, 25(1), 159–174, 2009.

[GlM12] A. GLITZKY and A. MIELKE. A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. *Z. angew. Math. Phys. (ZAMP)*, 2012. To appear. WIAS preprint 1603 (doi 10.1007/s00033-0012-207-y).

[Gri85] P. GRISVARD. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[Gri04] J. A. GRIEPENTROG. On the unique solvability of a nonlocal phase separation problem for multicomponent systems. In *Nonlocal elliptic and parabolic problems*, volume 66 of *Banach Center Publ.*, pages 153–164. Polish Acad. Sci., Warsaw, 2004.

[GST09] U. GIANAZZA, G. SAVARÉ, and G. TOSCANI. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Rational Mech. Anal.*, 194(1), 133–220, 2009.

[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.

[KjB08] S. KJELSTRUP and D. BEDEAUX. *Non-equilibrium thermodynamics of heterogeneous systems*, volume 16 of *Series on Advances in Statistical Mechanics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[Lis09] S. LISINI. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3), 712–740, 2009.

[LMS12] S. LISINI, D. MATTHES, and G. SAVARÉ. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *Preprint*, 2012. arXiv:1201.2367v1.

[LSU68] O. A. LADYŽENSKAJA, V. A. SOLONNIKOV, and N. N. URAL'CEVA. *Linear and Quasilinear Equations of Parabolic Type*. Transl. Math. Monographs. Amer. Math. Soc., Providence, R.I., 1968.

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

[McC97] R. J. MCCANN. A convexity principle for interacting gases. *Adv. Math.*, 128, 153–179, 1997.

[Mie11a] A. MIELKE. Formulation of thermoelastic dissipative material behavior using GENERIC. *Contin. Mech. Thermodyn.*, 23(3), 233–256, 2011.

[Mie11b] A. MIELKE. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Part. Diff. Eqns.*, 2011. Submitted. WIAS preprint 1650.

[Mie11c] A. MIELKE. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24, 1329–1346, 2011.

[Mie12] A. MIELKE. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discr. Cont. Dynam. Systems Ser. S*, 2012. To appear. WIAS preprint 1661 (Nov. 2011).

[MMS09] D. MATTHES, R. J. MCCANN, and G. SAVARÉ. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12), 1352–1397, 2009.

[MOS12] A. MIELKE, C. ORTNER, and Y. SENGÜL. Quasistativ nonlinear viscoelasticity as a curve of maximal slope. *In preparation*, 2012.

[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 Differential Equations*, 26, 101–174, 2001.

[OtW05] F. OTTO and M. WESTDICKENBERG. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Analysis*, 37, 1227–1255, 2005.

[Vil09] C. VILLANI. *Optimal transport. Old and new.* Berlin: Springer, 2009.