Technische Universität München
Fakultät für Mathematik

Lehrstuhl für Höhere Mathematik und Analytische Mechanik

Systems of Evolution Equations with Gradient Flow Structure

Jonathan Simon Zinsl

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. rer. nat. Johannes Müller
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. rer. nat. Daniel Matthes
	2.	Univ.-Prof. Dr. rer. nat. Alexander Mielke, Humboldt-Universität zu Berlin
	3.	Univ.-Prof. Dr. rer. nat. Martin Brokate

Die Dissertation wurde am 08.10.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 18.01.2016 angenommen.

Acknowledgement

First of all, I would like to express my particular gratitude to my advisor Prof. Dr. Daniel Matthes without whom this thesis would not have been possible. I am very grateful for all the hours of constructive discussion, listening and reading and — above all — for the guidance through the maze of all the research projects and for giving me the freedom to choose the topic to my interest. Thank you for your unrestricted support!

Next, my special gratitude also goes to Prof. Dr. Alexander Mielke for his kind hospitality at the Weierstraß-Institut in Berlin and his valuable feedback on my research. I would also like to thank my collaborators Dr. Karoline Disser and Dr. Matthias Liero from the Weierstraß-Institut for all the time we spent on our project and in particular for the friendly atmosphere during my visits in Berlin.

Of course, I would also like to say thank you to Prof. Dr. Jürgen Scheurle and all my colleagues from the reseach unit M8 for their outstanding collegiality and helpfulness. I shall definitely miss the cookies after lunch! A special thanks goes to my "academic brother" and office mate Horst Osberger who calmly beared all my spleens during the last three years and always was ready to help.

The following people at the faculty deserve my gratitude for their helpfulness in academic (mostly teaching) matters: Dr. Carl-Friedrich Kreiner, Dr. Stephan Schmitz and Dr. Hans-Peter Kruse. With your advice and help, the teaching part of my work became considerably easier.

Especially, I would like to thank my family for always supporting my choice of being a mathematician. Finally, thank you so much, Marion, for your love and just always being there.

Zusammenfassung

In dieser Arbeit werden mathematische Analysen verschiedener bestimmter Systeme von gewöhnlichen und evolutionären partiellen Differentialgleichungen durchgeführt, die sich auf eine zu Grunde liegende formale Gradientenflussstruktur stützen.

Zunächst werden Varianten des Keller-Segel-Modells für Chemotaxis untersucht und die Existenz schwacher Lösungen durch Approximation mittels des *minimizing movement scheme* bewiesen. In besonderen Fällen kann die exponentiell schnelle Konvergenz jener schwacher Lösungen zum eindeutigen Gleichgewichtszustand des Systems gezeigt werden. Eine ähnliche Methode findet Anwendung bei der Untersuchung des Langzeitverhaltens schwacher Lösungen des Poisson-Nernst-Planck-Systems, eines Modells für den Transport geladener Teilchen.

Weiter wird ein Multikomponentensystem aus nichtlokalen Interaktionsgleichungen analysiert. Wir erhalten eine hinreichende Bedingung für die Konvexität des zu Grunde liegenden Interaktionsenergiefunktionals entlang von Geodäten bezüglich einer Wasserstein-artigen Metrik im Raum der endlichen Vektormaße. Im nicht-konvexen Fall werden qualitative Eigenschaften der Lösung wie z.B. die gleichmäßige Beschränktheit des Trägers oder die Stabilität stationärer Zustände, hergeleitet.

Ein weiterer Teil der Arbeit beschäftigt sich mit mehrkomponentigen, degenerierten Diffusionssystemen mit nichtlinearer, matrixwertiger Mobilität. Wir zeigen die variationelle Struktur solcher Systeme auf, indem wir eine neue Metrik zwischen vektorwertigen, messbaren Funktionen definieren, bezüglich derer das System eine formale Gradientenstruktur besitzt. Wir untersuchen die topologischen Eigenschaften dieser Metrik und die geodätische Konvexität von Funktionalen. Mit Hilfe des *minimizing movement scheme* beweisen wir die Existenz schwacher Lösungen für bestimmte Klassen von Systemen zweiter sowie vierter Ordnung.

Des weiteren betrachten wir ein System aus gewöhnlichen Differentialgleichungen, das ein Netzwerk aus schnellen und langsamen chemischen Reaktionen modelliert, und untersuchen das Konvergenzverhalten im Limes unendlich großer Geschwindigkeiten der schnellen Reaktionen. Mittels der verallgemeinerten Gradientenstruktur des Systems können wir eine dimensionsreduzierende evolutionäre Konvergenz des Systems gegen ein Limes-Gradientensystem und dessen Evolutionsgleichungen herleiten.

Abstract

In this thesis, we perform mathematical analyses for several specific systems of ordinary and evolutionary partial differential equations using an underlying formal gradient flow structure.

We first investigate variants of the Keller-Segel model for chemotaxis and prove the existence of weak solutions by approximation via the minimizing movement scheme. In a specialized setting, those weak solutions are shown to converge at exponential rate to the unique equilibrium of the system. A similar strategy is used to study the long-time behaviour of weak solutions to the Poisson-Nernst-Planck system modelling ion transport.

Second, we investigate a multi-species system of nonlocal interaction equations and find a sufficient criterion for convexity along geodesics of the underlying interaction energy functional with respect to a metric of Wasserstein type on the space of vector-valued finite measures. Moreover, we obtain — in the non-convex scenario — results on the qualitative behaviour of the solution such as confinement of the support or stability of steady states.

In the third part, we consider multi-species systems of degenerate diffusion equations with a nonlinear mobility matrix function and prove that it arises as formal gradient flow with respect to a novel transportation distance between vector-valued measurable functions. We investigate the topological properties of this new distance and study the geodesic convexity of functionals. Using the minimizing movement scheme, the existence of weak solutions to specific classes of second- and fourth-order systems of this kind is shown.

Furthermore, we consider a system of ordinary differential equations modelling a network of slow and fast chemical reactions and investigate the limit behaviour as the fast reaction rates tend to infinity. Using the generalized gradient structure of the system of equations, we obtain a notion of dimension-reducing evolutionary convergence to a limit gradient system, thereby deriving its governing equations.

Contents

Part I

Introduction and Overview

CHAPTER I.1

Introduction and main results

During the last decades, both entropy as well as variational methods have been proved to be extremely useful for the analysis of nonlinear evolution equations. They can not only be used for proving the existence of (generalized) solutions, but the study of Lyapunov functionals also sometimes makes it possible to extract specific information on the qualitative behaviour of solutions, e.g. the speed of convergence to equilibrium. One class of evolution equations exhibiting a variational structure consists of those associated to *gradient systems* or *gradient flows*. Heuristically, a gradient flow follows the direction of steepest descent in the landscape of an *energy* (also called *entropy*) functional, with respect to the curved structure of the underlying (often non-flat) metric space.

Since the seminal paper by Jordan, Kinderlehrer and Otto [105] on the variational structure of the Fokker-Planck equation, various evolutionary partial differential equations have been interpreted as gradient flows on the space of probability measures $\mathscr{P}(\mathbb{R}^d)$, endowed with the so-called L^2-Wasserstein distance (which is a central object in optimal transportation theory [179]). Apart from the application to nonlinear diffusion equations [57, 154, 1, 49, 68] such as the porous medium equation, aggregation equations [54, 55, 47] and equations of fourth order [86, 88, 132, 125] have been considered, also in a spatially discrete framework [130, 78, 89, 79].

The interpretation as a gradient flow on the space of probability measures can be of use for the analysis of nonlinear evolution equations in several ways. For instance, Jordan, Kinderlehrer and Otto [105] proved the existence of *nonnegative* weak solutions to the Fokker-Planck equation using the so-called *minimizing movement scheme* (for details, see Chapter I.2 below). As a byproduct of the construction of a continuous flow in the space $\mathscr{P}(\mathbb{R}^d)$, the nonnegativity of solutions is automatically obtained, thus leading to a sensible solution from the modelling point of view. Therefore, the minimizing movement scheme might allow one to construct physically or biologically sensible solutions also to higher-order equations or to coupled systems when comparison principles are not at hand in general. The application of this variational method for proving existence requires relatively little on the corresponding energy functional, since boundedness from below, lower semicontinuity with respect to a suitable topology and a suitable version of coercivity usually suffices to prove the existence of minimizers of the associated *Yosida penalization* (as needed for the scheme) with the classical direct method from the calculus of variations. An existence analysis via this approach has been made e.g. by Gianazza, Savaré and Toscani for the quantum drift-diffusion equation [88] and has later been generalized by Matthes, McCann and Savaré to a more general class of fourth-order equations [132] comprising e.g. the thin film equation.

Apart from the space of probability measures $\mathscr{P}(\mathbb{R}^d)$ equipped with the L^2-Wasserstein distance, one can also consider gradient flows on abstract metric spaces. A thorough investigation in this more general framework has been made by Ambrosio, Gigli and Savaré in their monograph [4]. On grounds of their theory and the results by Dolbeault, Nazaret and Savaré [74] and Lisini and Marigonda [124] on the construction of generalized Wasserstein distances, the existence of weak solutions to a class of nonlinear fourth-order equations (e.g., the Cahn-Hilliard and nonlinear versions of the thin film equation) has been proved by Lisini, Matthes and Savaré [125] using the formal gradient flow structure of the equation with respect to these generalized Wasserstein distances. More recently, also systems of evolution equations came into consideration in the context of gradient flows. There, much attention was devoted to the Keller-Segel system for chemotaxis and its various conceivable variants. For the parabolic-elliptic Keller-Segel model, which can be reduced to a single scalar nonlocal equation possessing a formal gradient flow structure on the space $\mathscr{P}(\mathbb{R}^d)$, an existence analysis with the minimizing movement scheme has been performed by Blanchet, Calvez and Carrillo [23]. The variational structure of the parabolic-parabolic Keller-Segel system (which is a genuine, coupled system of two equations) is different: it can formally be written as a gradient flow with respect to a compound metric on the hybrid product space $\mathscr{P}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. It was proved by Blanchet and Laurençot [28] in higher spatial dimensions for Keller-Segel-type systems comprising nonlinear diffusion with a critical diffusion exponent and by Blanchet *et al.* [25] on the plane in the (most delicate) case of linear diffusion that an approximation via the minimizing movement scheme leads to the existence of nonnegative weak solutions. In comparison with the vast literature on the existence of solutions to systems of Keller-Segel type, the interpretation as a gradient system permits to consider broader parameter ranges as well as more general initial data than for the use of non-variational methods, usually arriving at a weaker — but still biologically reasonable — notion of solution. In a similar hybrid product space, the existence of solutions to a system modelling the Janossy effect in liquid crystals has been shown by Kinderlehrer and Kowalczyk [111] with essentially the same method. Obviously, it is not limited to hybrid product spaces: in the product $\mathscr{P}(\mathbb{R}^d) \times \mathscr{P}(\mathbb{R}^d)$ of two Wasserstein spaces, one can e.g. consider the Poisson-Nernst-Planck system [112] modelling ion transport or a thin film approximation of the Muskat problem [118] which both allow for the construction of weak solutions via the minimizing movement scheme, employing the formal gradient structure with respect to a certain energy functional.

In the framework of abstract metric spaces, Ambrosio, Gigli and Savaré [4] are mostly concerned with gradient flows of functionals which are λ-convex along geodesics in the respective metric space for some $\lambda \in \mathbb{R}$. Convexity along geodesics is a suitable notion of convexity in non-Euclidean spaces, taking the curvature of the underlying space into account. In the special framework of the space of probability measures, McCann [133] studied internal (as e.g. *Boltzmann's entropy*), potential and interaction energy functionals to obtain criteria for the λ-convexity along (even generalized — see Chapter I.2 below) geodesics in $\mathscr{P}(\mathbb{R}^d)$ with the approach of *displacement interpolation*. Using the λ-convexity along geodesics, Ambrosio, Gigli and Savaré [4] construct a discrete approximative gradient flow via the minimizing movement scheme which converges in the limit of vanishing step size to a continuous gradient flow. They also obtain optimal error estimates for the approximative discrete flow in comparison with the sought-for continuous gradient flow. The notion of λ-convexity along geodesics leads us to the second main benefit of the interpretation as a gradient system: λ-convexity of the energy functional implies the *λ-contractivity* of the flow; thus, solutions are unique and converge at an exponential rate to the unique equilibrium of the system coinciding with the unique minimizer of the underlying energy functional, if the modulus of convexity λ is strictly positive. In his preceding seminal article on the porous medium equation, Otto [154] was able to quantify the self-similar convergence of solutions to the porous medium equation against the so-called *Barenblatt(-Pattle) profile*. In a nutshell, the transformation of the equation to self-similar variables leads to a system which has a gradient structure in the Wasserstein space $\mathscr{P}(\mathbb{R}^d)$ with respect to a λ-convex functional with positive λ. The derivation of the corresponding contractivity estimate for the flow of the transformed equation made it possible to obtain an explicit exponential rate of convergence to the Barenblatt profile which has not been known before. This seminal result initiated many studies of the long-time behaviour of solutions to certain evolution equations. For instance, Otto's result has been generalized first to a more general class of nonlinear diffusion equations by Carrillo, di

Francesco and Toscani [49] and later by Agueh [1] to porous medium equations with a p-Laplacian. For more general second-order evolution equations that are gradient flows in the Wasserstein space $\mathscr{P}(\mathbb{R}^d)$ of a geodesically λ-convex functional consisting of an internal, a potential and an interaction energy part, the respective contraction estimates have been shown by Carrillo, McCann and Villani [55] on grounds of Otto's method. For purely nonlocal interaction equations with a non-smooth interaction potential, the qualitative behaviour of the gradient flow semigroup has been studied by Carrillo *et al.* in [47]

Energy-dissipation methods involving the Wasserstein gradient structure were also applied to investigate the long-time behaviour of higher-order equations. In the aforementioned article by Gianazza, Savaré and Toscani [88] on the quantum drift-diffusion equation, which can be seen as a Wasserstein gradient flow of the so-called *Fisher information* functional, the authors demonstrate the exponential convergence to equilibrium for certain confinement potentials. There, it is made use of a link of the fourth-order quantum drift diffusion equation to the second-order Fokker-Planck equation: the Fisher information functional coincides with the squared Wasserstein subdifferential (as later established in [4]) of the geodesically convex functional inducing the Fokker-Planck equation as Wasserstein gradient flow. Later, by Matthes, McCann and Savaré [132], their result was generalized to a wider class of fourth-order equations. Moreover, the so-called *flow interchange* technique (which has previously been applied in concrete examples, see e.g. [88]) was firstly formalized in a more abstract setting. More recently, Blanchet, Carlen and Carrillo [24] used the gradient structure of the parabolic-elliptic Keller-Segel system with critical mass to study the basins of attraction of steady states. The long-time asymptotics in the parabolic-parabolic case, however, have not been investigated by variational methods yet.

Aim of the thesis. The goal of this thesis is to analyse specific coupled systems with a formal gradient structure, for example as an extension of the theory for certain classes of scalar equations to the case of genuine systems. We shall use methods from the theory of gradient flows to prove the existence of solutions to the problems at hand as well as to investigate their qualitative behaviour, e.g. in the long-time limit. Moreover, the thesis aims at finding new properties of these multi-species systems not known for the corresponding scalar equations and at detecting the possible limitations due to the inherent more complex character of systems.

In the following, we give a brief and descriptive overview on the main results of this thesis.

A Keller-Segel-type models. In Part II, we consider several variants of the *Keller-Segel model* for chemotaxis of the following form:

$$\partial_t u(t,x) = K_u \Delta u^m(t,x) + \operatorname{div}\left(u(t,x)\mathrm{D}\left[W(x) + \chi\phi(v(t,x))\right]\right) \qquad \text{(bacterial density)},$$
$$\partial_t v(t,x) = K_v \Delta v(t,x) - \kappa v(t,x) - \chi u(t,x)\phi'(v(t,x)) \qquad \text{(signal strength)},$$

where $t > 0$ and $x \in \mathbb{R}^d$, which formally possesses a gradient structure on the (hybrid) product space $\mathscr{P}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ with respect to the energy functional

$$\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^d} \left(\frac{K_u u^m}{m-1} + uW + \frac{K_v}{2}|\mathrm{D}v|^2 + \frac{\kappa}{2}v^2 + \chi u\phi(v)\right) \mathrm{d}x & \text{if } m > 1, \\ \int_{\mathbb{R}^d} \left(K_u u\log u + uW + \frac{K_v}{2}|\mathrm{D}v|^2 + \frac{\kappa}{2}v^2 + \chi u\phi(v)\right) \mathrm{d}x & \text{if } m = 1. \end{cases} \tag{I.1.1}$$

In this variant of the Keller-Segel model, the central objects are the exponent $m \geq 1$ of nonlinear diffusion and the (possibly nonlinear) chemotactic sensitivity function ϕ.

We obtain the following main results:

- For $d \geq 3$, linear sensitivity ϕ and supercritical diffusion exponent $m > 2 - \frac{2}{d}$, weak solutions can be constructed by approximation via the minimizing movement scheme in the space $\mathscr{P}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$.
- For $d = 3$, $m = 2$ and decreasing and convex sensitivity ϕ, a similar existence result is obtained. Moreover, if the equations above are weakly coupled ($\chi > 0$ is small), we demonstrate the exponential convergence to the unique stationary state of the system of the weak solution constructed beforehand by the study of a suitable Lyapunov functional. Results of this kind for parabolic-parabolic Keller-Segel-type systems are — even in the regime of small coupling — new in the literature.
- In one spatial dimension $d = 1$, analogous results as above can be found even allowing for linear diffusion $m = 1$.

B **Poisson-Nernst-Planck-type systems.** The remainder of Part II is concerned with a variant of the *Poisson-Nernst-Planck system* modelling the transport of charged particles (e.g. ions) in an electrically neutral medium:

$$\partial_t u(t,x) = \operatorname{div}(u(t,x)\mathrm{D}[2u(t,x) + U(x) + \varepsilon\psi(t,x)]) \qquad \text{(positively charged particles)},$$
$$\partial_t v(t,x) = \operatorname{div}(v(t,x)\mathrm{D}[2v(t,x) + V(x) - \varepsilon\psi(t,x)]) \qquad \text{(negatively charged particles)},$$

coupled by *Poisson's equation*

$$-\Delta\psi = u - v \qquad \text{(electric potential)},$$

considered in three spatial dimensions. Above, the functions $U, V \in C^2(\mathbb{R}^3)$ are assumed to be quadratically growing λ_0-convex *confinement potentials*, with $\lambda_0 > 0$, and $\varepsilon > 0$ is a parameter reflecting the relative permittivity of the surrounding medium. This two-species model has (similarly to the Keller-Segel model) a formal gradient flow structure, now, in contrast to before, in the non-hybrid product space $\mathscr{P}(\mathbb{R}^d) \times \mathscr{P}(\mathbb{R}^d)$. The free energy functional $\mathcal{E}$ here is

$$\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^3} (u^2 + v^2 + uU + vV + \frac{\varepsilon}{2}|\mathrm{D}\psi|^2)\,\mathrm{d}x & \text{if } (u,v) \in L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3), \\ +\infty & \text{otherwise.} \end{cases}$$

Our main results are twofold and can be obtained via similar strategies as for the Keller-Segel system:

- We show that the system above possesses a unique steady state.
- On grounds of an existence result from [112], we find the exponential convergence to this unique equilibrium in the case of small coupling $0 < \varepsilon \ll 1$.

C **Systems of nonlocal interaction equations.** Part III is devoted to an n-component system of *nonlocal interaction equations* of the form

$$\begin{aligned} \partial_t \mu_1 &= \operatorname{div}[m_1\mu_1\nabla(W_{11} * \mu_1 + W_{12} * \mu_2 + \ldots + W_{1n} * \mu_n)], \\ \partial_t \mu_2 &= \operatorname{div}[m_2\mu_2\nabla(W_{21} * \mu_1 + W_{22} * \mu_2 + \ldots + W_{2n} * \mu_n)], \\ &\vdots \\ \partial_t \mu_n &= \operatorname{div}[m_n\mu_n\nabla(W_{n1} * \mu_1 + W_{n2} * \mu_2 + \ldots + W_{nn} * \mu_n)]. \end{aligned}$$

This system arises as an abstract force balance for particle densities $(\mu_1, \ldots, \mu_n)$ subject to nonlocal attractive or repulsive forces generated by the densities of all species via the (by assumption regular) potentials W_{ij}. It is — as a completely natural generalization of the scalar interaction equation — governed by the *interaction potential* $W : \mathbb{R}^d \to \mathbb{R}^{n \times n}$ and possesses a formal gradient structure on a suitable subspace $\mathscr{P}$ of the space of finite n-vector-valued measures on $\mathbb{R}^d$ endowed with an n-product distance of Wasserstein type compounds, w.r.t. the energy

$$\mathcal{E}(\mu) := \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} W_{ij}(x-y)\,\mathrm{d}\mu_i(x)\,\mathrm{d}\mu_j(y).$$

Having certain symmetry, convexity and growth properties of W at hand, our main results are summarized as follows:

- We derive, by adapting McCann's method [133], a sufficient condition for convexity along generalized geodesics on $\mathscr{P}$ of the interaction energy functional $\mathcal{E}$ — including McCann's criterion for convexity [133] — and conclude the existence and uniqueness of solutions to the system of nonlocal equations above using the abstract theory from [4]. This criterion is novel in the literature; and it does not emerge from an obvious generalization of the scalar theory.
- In one spatial dimension, we derive a criterion under which the support of the solution stays confined to a compact interval even if $\mathcal{E}$ is not uniformly convex along geodesics.
- We investigate, also in the non-convex case, the long-time behaviour of solutions to find that the ω-limit set of the dynamical system associated to the partial differential equation above contains only steady states.
- We study the linear and nonlinear stability of stationary states (generalizing a method used by Fellner and Raoul [82]). By analysing the spectrum of the linearization, we exclude the linear asymptotic stability for non-discrete steady states. On the other hand, we derive a sufficient criterion for the local nonlinear asymptotic stability of discrete stationary states.

D **Degenerate diffusion systems with nonlinear mobility.** In Part IV, we present a new variational structure for the following system of $n \in \mathbb{N}$ degenerate diffusion equations in one spatial dimension:

$$\partial_t\mu(t,x) = \partial_x\left[\mathbf{M}(\mu(t,x))\partial_x\mathcal{E}'(\mu(t,x))\right],$$

where we assume that the sought-for n-vector-valued function $\mu : (0,\infty)\times\mathbb{R} \to S$ possesses values in a prescribed convex and compact set $S \subset \mathbb{R}^n$. A central position in the above system is occupied by the *mobility matrix function* $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ which is assumed to be (in a suitable sense for matrices) positive and concave and to take into account degeneracy effects on the boundary of the value space S. This system has already been studied in the scalar case $n = 1$: It is known to have a gradient flow structure with respect to a modified Wasserstein distance (defined by Dolbeault *et al.* [74] and Lisini and Marigonda [124]) as a generalization of the dynamical formulation for the quadratic Wasserstein distance found by Benamou and Brenier [11]. In this thesis, we extend this theory to the case of genuine systems $n > 1$:

- We give a rigorous construction of a (pseudo-)distance $\mathbf{W}_{\mathbf{M}}$ on the space $\mathscr{M}(\mathbb{R};S)$ of n-vector-valued measurable functions with values in S which formally reads

$$\mathbf{W}_{\mathbf{M}}(\mu_0,\mu_1)^2 = \inf\Bigg\{\int_0^1\int_{\mathbb{R}} w_t^{\mathrm{T}}\mathbf{M}(\mu_t)^{-1}w_t\,\mathrm{d}x\,\mathrm{d}t :$$
$$\partial_t\mu_t = -\partial_x w_t \text{ in the sense of distributions on } [0,1]\times\mathbb{R},\ \mu|_{t=0} = \mu_0, \mu|_{t=1} = \mu_1\Bigg\},$$

  for $\mu_0,\mu_1 \in \mathscr{M}(\mathbb{R};S)$. We study the topological properties of this distance and discuss criteria that guarantee the finiteness of $\mathbf{W}_{\mathbf{M}}$.
- As an application of the *Eulerian calculus* by Daneri and Savaré [64] in the abstract form of Liero and Mielke [123], we find a sufficient condition for λ-convexity on geodesics in the space $\mathscr{M}(\mathbb{R};S)$ of *internal energy functionals* $\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu)\,\mathrm{d}x$, comprising the generalized McCann

condition found by Carrillo *et al.* [53] in the scalar case $n = 1$. As a byproduct, the multi-component diffusion equation $\partial_t \mu = \partial_{xx} \mu$ is seen to correspond to a 0-contractive gradient flow w.r.t. this new distance $\mathbf{W_M}$ assuming that the mobility is *induced* by a energy functional $\mathcal{H}(\mu) = \int_{\mathbb{R}} h(\mu)\,\mathrm{d}x$ in the following sense:

$$\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \quad \text{for all } z \text{ in the interior of } S.$$

However, as already evident in the scalar case [53], λ-convexity on geodesics is a very rare property. Nevertheless, we present some nontrivial examples which arise as perturbations of the heat entropy $\mathcal{H}$ above.

- Even in the non-convex case, we can — under certain assumptions on $\mathbf{M}$ and $\mathcal{E}$ — construct weak solutions to the system above again by using the variational minimizing movement scheme. Specifically, we require the mobility matrix to be decoupled in such a way that the distance $\mathbf{W_M}$ is a product distance of n factors each corresponding to a distance for scalar densities (as studied in [74, 124]). Nonetheless, we are able to consider second- as well as fourth-order systems with our strategy.

E Networks of slow and fast chemical reactions. In the last main part (Part V) of the thesis, we move to a more general framework of gradient systems and specifically consider a system of $I \in \mathbb{N}$ ordinary differential equations modelling a network of chemical reactions with mass-action kinetics:

$$\dot{c}_\varepsilon(t) = -\sum_{r=1}^{R} k_\varepsilon^r \big(c_\varepsilon(t)^{\alpha^r} - c_\varepsilon(t)^{\beta^r}\big)(\alpha^r - \beta^r),$$

where $\alpha^r, \beta^r \in \mathbb{N}_0^I$ are the stoichiometric vectors of the system of chemical reactions. The main feature of this system is reflected by the dependency of the reaction rates k_ε^r on the (small) parameter $\varepsilon > 0$. The R reactions can be divided into two classes: the first R_s reactions are assumed to be *slow* in the sense that k_ε^r is of order 1 as $\varepsilon \searrow 0$ and the subsequent R_f reactions are *fast*; $k_\varepsilon^r \in O\left(\frac{1}{\varepsilon}\right)$.

Those systems of chemical reactions with a mass-action law are *generalized gradient systems* in the sense of Mielke [134]: with the energy functional

$$\mathcal{E}_\varepsilon(c) := \sum_{i=1}^{I} (c_i \log(c_i) - c_i + 1) \qquad \text{for } c \in [0,\infty)^I,$$

there exists a *dissipation potential* Ψ_ε such that, along the solution $c_\varepsilon : [0,\infty) \to \mathbb{R}^I$ to the system of ordinary differential equations above, the following energy dissipation balance holds:

$$\mathcal{E}_\varepsilon(c_\varepsilon(T)) + \int_0^T \left(\Psi_\varepsilon(c_\varepsilon; \dot{c}_\varepsilon) + \Psi_\varepsilon^*(c_\varepsilon; -\mathrm{D}\mathcal{E}_\varepsilon(c_\varepsilon))\right) \mathrm{d}t = \mathcal{E}_\varepsilon(c_\varepsilon(0)).$$

Our main interest lies in the behaviour in the limit $\varepsilon \searrow 0$:

- Using the gradient structure in combination with the analysis of suitable Lyapunov functions, we obtain the convergence of solutions to the system above in the limit $\varepsilon \searrow 0$ to solutions to a limit system of lower dimension. Our result includes a related theorem by Bothe [30].
- The limit system is found in a more natural way than in [30] by analysing the Γ-limit behaviour of $\mathcal{E}_\varepsilon$ and Ψ_ε as $\varepsilon \searrow 0$. We recover the notion of *evolutionary Γ-convergence* — developed originally by Sandier and Serfaty [160] — in the sense of Mielke [138] of the family of gradient systems associated to the evolution equation.

In advance of the detailed exposition of our work, we begin with a short summary on the theory of gradient flows and gradient systems for the sake of motivation and preparation.

CHAPTER I.2

A synopsis on the theory of gradient systems

In this chapter, we give a (non-exhaustive) overview on the theory of gradient flows and generalized gradient systems as a preparation for the novel studies in this thesis.

In finitely many dimensions, the dynamics of gradient flows are almost trivial. Indeed, for sufficiently smooth $F : \mathbb{R}^d \to \mathbb{R}$, solutions to the ordinary differential equation

$$\dot{u}(t) = -\nabla F(u(t)) \tag{I.2.1}$$

associated to the gradient flow of F on $\mathbb{R}^d$ either converge to a critical point of F or diverge as $t \to \infty$. In particular, periodic solutions cannot exist — which is also true in the general case. The flow of equation (I.2.1) can moreover be characterized by a variational principle, the so-called *evolution variational estimate* (with parameter $\kappa \in \mathbb{R}$)

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|u(t) - v|^2 + \frac{\kappa}{2}|u(t) - v|^2 \le F(v) - F(u(t)) \qquad \text{for all } v \in \mathbb{R}^d \text{ and } t > 0, \tag{I.2.2}$$

since the following is true: $u : (0, \infty) \to \mathbb{R}^d$ satisfies (I.2.2) if and only if u is a solution to (I.2.1) and F is κ-convex, i.e. $z \mapsto F(z) - \frac{\kappa}{2}|z|^2$ is convex. Furthermore, the following *energy dissipation balance* holds along the gradient flow:

$$F(u(T)) + \frac{1}{2}\int_0^T |\dot{u}(t)|^2 \,\mathrm{d}t + \frac{1}{2}\int_0^T |\nabla F(u(t))|^2 \,\mathrm{d}t = F(u(0)) \quad \text{for all } T > 0, \tag{I.2.3}$$

having the chain rule $\frac{\mathrm{d}}{\mathrm{d}t}F(u(t)) = \langle \nabla F(u(t)), \dot{u}(t)\rangle$ in mind.

In non-Euclidean settings — e.g. in a metric space with non-zero curvature — the relation (I.2.1) can possibly not be given any sense. However, one can generalize (I.2.2)&(I.2.3) in a suitable way to the metric setting (see Section I.2.1 below).

I.2.1. Gradient flows in abstract metric spaces

We begin with a summary on relevant definitions and facts about gradient flows in complete metric spaces $(\mathbf{X}, \mathbf{d})$ (consult the monograph by Ambrosio, Gigli and Savaré [4] for more details). Let a functional $\mathcal{A} : \mathbf{X} \to (-\infty, +\infty] =: \mathbb{R}_\infty$ be given. The *proper domain* $\mathrm{Dom}(\mathcal{A})$ is defined as the set $\mathrm{Dom}(\mathcal{A}) := \{w \in \mathbf{X} : \mathcal{A}(w) < \infty\}$ where $\mathcal{A}$ is finite on. We shall always assume that $\mathcal{A}$ is *proper*, i.e. $\mathrm{Dom}(\mathcal{A}) \neq \varnothing$. A functional $\mathcal{A}$ is called *(sequentially) lower semicontinuous* if $w_k \to w$ in $(\mathbf{X}, \mathbf{d})$ as $k \to \infty$ implies $\mathcal{A}(w) \le \liminf_{k\to\infty} \mathcal{A}(w_k)$. Given $\tau > 0$ and $\widetilde{w} \in \mathbf{X}$, we define the *Yosida penalization* for $\mathcal{A}$ as the functional

$$\mathcal{A}_\tau(\cdot|\widetilde{w}) : \mathbf{X} \to \mathbb{R}_\infty, \qquad \mathcal{A}_\tau(w|\widetilde{w}) := \frac{1}{2\tau}\mathbf{d}(w, \widetilde{w})^2 + \mathcal{A}(w).$$

This functional plays a pivotal role in the construction of *gradient flows* by approximation with *minimizing movements*, see below. There, minimizers of $\mathcal{A}_\tau(\cdot|\widetilde{w})$ are involved, motivating the following definition: A functional $\mathcal{A}$ is called *coercive* if there exist $\tau_* > 0$ and $w_* \in \mathbf{X}$ such that $\mathcal{A}_{\tau_*}(\cdot|w_*)$ is bounded from below on $\mathbf{X}$. Obviously, if $\mathcal{A}$ is bounded from below itself, it also is coercive.

In this purely metric framework, a notion of convexity is defined via the interpolation along *curves* in $\mathbf{X}$.

Definition I.1 (Convexity along (geodesic) curves [133], [4, §2.4]).

(a) *A functional* $\mathcal{A}$ *is called* λ-convex *along a curve* $\gamma : [0,1] \to \mathbf{X}$ *for some* $\lambda \in \mathbb{R}$ *if for all* $s \in [0,1]$*, the following holds:*

$$\mathcal{A}(\gamma_s) \le (1-s)\mathcal{A}(\gamma_0) + s\mathcal{A}(\gamma_1) - \frac{\lambda}{2}s(1-s)\mathbf{d}^2(\gamma_0,\gamma_1).$$

(b) *A curve* $\gamma : [0,1] \to \mathbf{X}$ *is called* constant-speed geodesic *if for all* $s,t \in [0,1]$*:*

$$\mathbf{d}(\gamma_s,\gamma_t) = \mathbf{d}(\gamma_0,\gamma_1)|s-t|.$$

(c) *A functional* $\mathcal{A}$ *is called* geodesically λ-convex *for some* $\lambda \in \mathbb{R}$ *if for each* $w_0, w_1 \in \mathrm{Dom}(\mathcal{A})$*, there exists a constant-speed geodesic* γ *with* $\gamma_0 = w_0$ *and* $\gamma_1 = w_1$ *such that* $\mathcal{A}$ *is* λ*-convex along* γ*.*

For geodesically λ-convex functionals, the following notion of *gradient flow* has been developed in [4]:

Definition I.2 (κ-flow). *Let* $\mathcal{A} : \mathbf{X} \to \mathbb{R}_\infty$ *be a lower semicontinuous functional on the metric space* $(\mathbf{X},\mathbf{d})$ *and let* $\kappa \in \mathbb{R}$*. A continuous semigroup* $\mathsf{S}^{\mathcal{A}}$ *on* $(\mathbf{X},\mathbf{d})$ *is called* κ-flow *(or* κ-contractive (gradient) flow*) of* $\mathcal{A}$ *for some* $\kappa \in \mathbb{R}$*, if the following holds:*

(a) $\mathsf{S}^{\mathcal{A}}_s(u) \in \mathrm{Dom}(\mathcal{A})$ *for all* $s > 0$ *and all* $u \in \mathbf{X}$*.*

(b) *The map* $s \mapsto \mathcal{A}\left(\mathsf{S}^{\mathcal{A}}_s(u)\right)$ *is nonincreasing in* $s \ge 0$ *for each fixed* $u \in \mathbf{X}$*.*

(c) *The* evolution variational estimate *(with parameter* κ*)*

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}s}\mathbf{d}^2(\mathsf{S}^{\mathcal{A}}_s(w),\widetilde{w}) + \frac{\kappa}{2}\mathbf{d}^2(\mathsf{S}^{\mathcal{A}}_s(w),\widetilde{w}) + \mathcal{A}(\mathsf{S}^{\mathcal{A}}_s(w)) \le \mathcal{A}(\widetilde{w}), \tag{I.2.4}$$

holds for arbitrary $w, \widetilde{w} \in \mathrm{Dom}(\mathcal{A})$*, and for all* $s \ge 0$*, where* $\frac{\mathrm{d}^+}{\mathrm{d}s}$ *denotes the right-sided derivative with respect to* s*.*

Notice that κ-flows are κ-contractive [64]: For all $w, \widetilde{w} \in \mathcal{A}$ and all $s \ge 0$, it holds that

$$\mathbf{d}(\mathsf{S}^{\mathcal{A}}_s(w),\mathsf{S}^{\mathcal{A}}_s(\widetilde{w})) \le e^{-\kappa s}\mathbf{d}(w,\widetilde{w}). \tag{I.2.5}$$

Especially, κ-flows are unique. The relation between geodesic λ-convexity and the existence of a κ-flow for a lower semicontinuous functional $\mathcal{A}$ can be characterized as follows:

Theorem I.3 (Convexity and gradient flows [4, Thm. 4.0.4], [64]). *Let* $\mathcal{A} : \mathbf{X} \to \mathbb{R}_\infty$ *be lower semicontinuous and coercive. The following statements hold for each* $\kappa \in \mathbb{R}$*:*

(a) *If* $\mathcal{A}$ *is bounded from below and there exists a* κ*-flow* $\mathsf{S}^{\mathcal{A}}$ *of* $\mathcal{A}$*, then* $\mathcal{A}$ *is geodesically* κ*-convex.*

(b) *Assume that the following condition* (C) *is satisfied:*

For each $\widetilde{w}, w_0, w_1 \in \mathrm{Dom}(\mathcal{A})$*, there exists a curve* γ *connecting* w_0 *and* w_1 *such that the functional*

$$w \mapsto \mathcal{A}_\tau(w|\widetilde{w}) \text{ is } \left(\frac{1}{\tau}+\kappa\right)\text{-convex along } \gamma \text{ for every } \tau \in \left(0,\frac{1}{\kappa_-}\right), \text{ with } \kappa_- := \max(0,-\kappa). \tag{C}$$

Then, if $\mathcal{A}$ *is geodesically* κ*-convex, then there exists a* κ*-flow of* $\mathcal{A}$*. Moreover, for each* $w_0 \in \mathrm{Dom}(\mathcal{A})$*, the curve* $w(s) = \mathsf{S}^{\mathcal{A}}_s(w_0)$ *for* $s \ge 0$ *is a* curve of maximal slope*:* w *is locally* L^2*-absolutely continuous with respect to the distance* $\mathbf{d}$ *(write* $w \in AC^2_{\mathrm{loc}}([0,\infty);(\mathbf{X},\mathbf{d}))$*), and the* energy dissipation balance *holds for all* $T > 0$*:*

$$\mathcal{A}(w_0) = \mathcal{A}(w(T)) + \frac{1}{2}\int_0^T |w'|^2(s)\,\mathrm{d}s + \frac{1}{2}\int_0^T |\partial\mathcal{A}|^2(w(s))\,\mathrm{d}s, \tag{I.2.6}$$

where

$$|w'|(s) := \lim_{t\to s} \frac{\mathbf{d}(w(s), w(t))}{|s-t|} \tag{I.2.7}$$

is the metric derivative *along w, and*

$$|\partial\mathcal{A}|(v) := \limsup_{\widetilde{v}\to v} \frac{\max\left(0, \mathcal{A}(v) - \mathcal{A}(\widetilde{v})\right)}{\mathbf{d}(v,\widetilde{v})} \tag{I.2.8}$$

is the local slope *of $\mathcal{A}$ at $v \in \mathrm{Dom}(\mathcal{A})$.*

The special case $\kappa > 0$ deserves to be explained in more detail. First, there exists exactly one minimizer $w_{\min}$ of $\mathcal{A}$, for which the following holds at each $w \in \mathrm{Dom}(\mathcal{A})$ [4, Lemma 2.4.13]:

$$\frac{\kappa}{2}\mathbf{d}^2(w, w_{\min}) \le \mathcal{A}(w) - \mathcal{A}(w_{\min}) \le \frac{1}{2\kappa}|\partial\mathcal{A}|^2(w). \tag{I.2.9}$$

This estimate is of particular interest since it e.g. provides a notion of uniform coercivity of the functional $\mathcal{A}$. Moreover, using the κ-contractivity of the flow $\mathsf{S}^{\mathcal{A}}$, one immediately deduces the convergence to $w_{\min}$ at exponential rate κ:

$$\mathbf{d}(\mathsf{S}^{\mathcal{A}}_s(w), w_{\min}) \le e^{-\kappa s}\mathbf{d}(w, w_{\min}).$$

Hence, the dynamics of the gradient flow $\mathsf{S}^{\mathcal{A}}$ can be characterized completely in the case of strictly positive modulus of convexity κ. The case $\kappa \le 0$ is more involved.

The additional convexity condition (C) does not only involve the behaviour of the functional $\mathcal{A}$, but is also a structural condition on the metric space $(\mathbf{X}, \mathbf{d})$. For example, if $(\mathbf{X}, \mathbf{d})$ is a Hilbert space (with the distance induced by the inner product) and $\mathcal{A}$ is κ-convex along geodesics (which are straight line segments in flat spaces), condition (C) is fulfilled. Indeed, for all $\widetilde{w}, w_0, w_1 \in \mathbf{X}$ and all $s \in [0,1]$, one has

$$\begin{aligned}
\mathcal{A}((1-s)w_0 + sw_1|\widetilde{w}) &\le \frac{1}{2\tau}(1-s)\|w_0-\widetilde{w}\|^2 + \frac{1}{2\tau}s\|w_1-\widetilde{w}\|^2 - \frac{1}{2\tau}s(1-s)\|w_0-w_1\|^2 \\
&\quad + (1-s)\mathcal{A}(w_0) + s\mathcal{A}(w_1) - \frac{\kappa}{2}s(1-s)\|w_0-w_1\|^2 \\
&= (1-s)\mathcal{A}_\tau(w_0|\widetilde{w}) + s\mathcal{A}_\tau(w_1|\widetilde{w}) - \frac{1}{2}\left(\frac{1}{\tau}+\kappa\right)s(1-s)\|w_0-w_1\|^2,
\end{aligned}$$

even for all $\tau > 0$. In contrast, on the space of probability measures $\mathscr{P}(\mathbb{R}^d)$ endowed with the L^2-Wasserstein distance (see Section I.2.2 below), (C) is *not* satisfied. There, a stronger version of convexity than that in Definition I.1 is needed to conclude the existence of gradient flows.

Example I.4 (Dirichlet energy). *Consider the Hilbert space $\mathbf{X} := L^2(\mathbb{R}^d)$ and let $\kappa \ge 0$. The functional $\mathcal{A} : \mathbf{X} \to \mathbb{R}_\infty$ defined by*

$$\mathcal{A}(w) = \begin{cases} \frac{1}{2}\|\mathrm{D}w\|^2_{L^2} + \frac{1}{2}\kappa\|w\|^2_{L^2} & \text{if } w \in W^{1,2}(\mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases} \tag{I.2.10}$$

is bounded from below, lower semicontinuous and κ-convex (along geodesics) in $\mathbf{X}$. Its associated κ-flow $\mathsf{S}^{\mathcal{A}}$ (given by Theorem I.3(b)) is a solution to the diffusion equation with decay:

$$\partial_s\mathsf{S}^{\mathcal{A}}_s(w) = \Delta\mathsf{S}^{\mathcal{A}}_s(w) - \kappa\mathsf{S}^{\mathcal{A}}_s(w).$$

For $\kappa = 0$, $\mathcal{A}$ is called Dirichlet energy.

The cornerstone of the proof for Theorem I.3(b) it the so-called *Moreau-Yosida approximation* of $\mathcal{A}$ by

$$\widetilde{w} \mapsto \inf_{w\in\mathbf{X}} \mathcal{A}_\tau(w|\widetilde{w}) \qquad \text{on } \mathbf{X},$$

which allows one to construct a discrete approximation of the sought-for gradient flow by the *minimizing movement scheme* (dating back to de Giorgi [65]): Given a step size $\tau > 0$ and an initial datum $w_0 \in \mathbf{X}$, define a sequence $(w_\tau^n)_{n\in\mathbb{N}}$ recursively by

$$w_\tau^0 := w^0, \qquad w_\tau^n \in \operatorname*{argmin}_{w\in\mathbf{X}} \mathcal{A}_\tau(w|w_\tau^{n-1}) \quad (n \in \mathbb{N}). \tag{I.2.11}$$

The *discrete solution* $w_\tau : [0,\infty) \to \mathbf{X}$ can be constructed as the piecewise constant interpolation along $(w_\tau^n)_{n\in\mathbb{N}}$, that is

$$w_\tau(0) := w_0, \quad w_\tau(t) := w_\tau^n \text{ for } t \in ((n-1)\tau, n\tau] \text{ and } n \in \mathbb{N}. \tag{I.2.12}$$

The minimizing movement scheme is well-posed also for more general functionals which are not κ-convex for any $\kappa \in \mathbb{R}$. Moreover, useful estimates on the behaviour of the associated discrete solution might be derived using the following Theorem.

Theorem I.5 (Flow interchange lemma [132, Thm. 3.2]). *Let $\mathcal{B}$ be a proper, lower semicontinuous and geodesically λ-convex functional on $(\mathbf{X}, \mathbf{d})$ and assume that there exists a λ-flow $\mathsf{S}^{\mathcal{B}}$. Let furthermore $\mathcal{A}$ be another proper, lower semicontinuous functional on $(\mathbf{X}, \mathbf{d})$ such that* $\mathrm{Dom}(\mathcal{A}) \subset \mathrm{Dom}(\mathcal{B})$. *Assume that, for arbitrary $\tau > 0$ and $\widetilde{w} \in \mathbf{X}$, the functional $\mathcal{A}_\tau(\cdot|\widetilde{w})$ possesses a minimizer w on $\mathbf{X}$.*

Then, the following holds:

$$\mathcal{B}(w) + \tau \mathrm{D}^{\mathcal{B}}\mathcal{A}(w) + \frac{\lambda}{2}\mathbf{d}^2(w,\widetilde{w}) \le \mathcal{B}(\widetilde{w}).$$

There, $\mathrm{D}^{\mathcal{B}}\mathcal{A}(w)$ *denotes the* dissipation *of the functional $\mathcal{A}$ along the λ-flow $\mathsf{S}^{\mathcal{B}}$ of the functional $\mathcal{B}$, i.e.*

$$\mathrm{D}^{\mathcal{B}}\mathcal{A}(w) := \limsup_{h\searrow 0} \frac{\mathcal{A}(w) - \mathcal{A}(\mathsf{S}_h^{\mathcal{B}}(w))}{h}.$$

Often, the behaviour of the discrete solution in the continuous-time limit of vanishing step size $\tau \searrow 0$ shall be investigated. The following theorem provides an extension of the classical Aubin-Lions compactness lemma to this metric framework.

Theorem I.6 (Extension of the Aubin-Lions lemma [159, Thm. 2]). *Let $\mathbf{Y}$ be a Banach space and let $\mathcal{A} : \mathbf{Y} \to [0,\infty]$ be lower semicontinuous and have relatively compact sublevels in $\mathbf{Y}$. Let furthermore $\mathbf{W} : \mathbf{Y}\times\mathbf{Y} \to [0,\infty]$ be lower semicontinuous and such that $\mathbf{W}(u,\widetilde{u}) = 0$ for $u,\widetilde{u} \in$* $\mathrm{Dom}(\mathcal{A})$ *implies $u = \widetilde{u}$.*

If for a sequence $(U_k)_{k\in\mathbb{N}}$ of measurable functions $U_k : (0,T) \to \mathbf{Y}$, one has

$$\sup_{k\in\mathbb{N}} \int_0^T \mathcal{A}(U_k(t))\,\mathrm{d}t < \infty \qquad \textit{and} \tag{I.2.13}$$

$$\lim_{h\searrow 0} \sup_{k\in\mathbb{N}} \int_0^{T-h} \mathbf{W}(U_k(t+h), U_k(t))\,\mathrm{d}t = 0, \tag{I.2.14}$$

then there exists a subsequence that converges in measure w.r.t. $t \in (0,T)$ to a limit $U : (0,T) \to \mathbf{Y}$.

I.2.2. Gradient flows in spaces of measures

This section is mostly devoted to the theory of gradient flows in the space of probability measures on $\mathbb{R}^d$. We begin with some measure theory (see [4, Ch. 5] for more details).

By $\mathscr{P}(\mathbb{R}^d)$, we denote the space of probability measures on $\mathbb{R}^d$. The subspace $\mathscr{P}_2(\mathbb{R}^d)$ (also denoted by $\mathscr{P}_2$ for brevity) is meant to be the space of those measures $\mu \in \mathscr{P}(\mathbb{R}^d)$ with finite second moment

$$\mathbf{m}_2(\mu) = \int_{\mathbb{R}^d} |x|^2\,\mathrm{d}\mu(x).$$

A sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}(\mathbb{R}^d)$ is said to converge *narrowly* (or *weakly$*$*) to some limit probability measure $\mu \in \mathscr{P}(\mathbb{R}^d)$ if for all continuous and bounded maps $\phi : \mathbb{R}^d \to \mathbb{R}$, one has

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \phi(x)\,\mathrm{d}\mu_n(x) = \int_{\mathbb{R}^d} \phi(x)\,\mathrm{d}\mu(x). \tag{I.2.15}$$

Actually, on the space $\mathbb{R}^d$, it suffices to verify property (I.2.15) for $\phi \in C_c^\infty(\mathbb{R}^d)$, showing that narrow and distributional convergence are equivalent in this case. An important condition for relative compactness with respect to the narrow topology is given by Prokhorov's theorem [4, Thm. 5.1.3]. We state an equivalent condition here.

Theorem I.7 (Prokhorov [4, Rem. 5.1.5]). *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{P}(\mathbb{R}^d)$. If there exists a function $\psi : \mathbb{R}^d \to [0,\infty]$ with compact sublevels and*

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} \psi(x)\,\mathrm{d}\mu_n(x) < \infty,$$

then there exists a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ and a limit $\mu \in \mathscr{P}(\mathbb{R}^d)$ such that μ_{n_k} converges to μ narrowly as $k \to \infty$.

Example I.8 (Bounded second moments). *Clearly, Theorem I.7 admits the choice $\psi(x) = |x|^2$. Hence, if for a sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}_2$ the sequence of second moments $(\mathfrak{m}_2(\mu_n))_{n\in\mathbb{N}}$ is bounded, $(\mu_n)_{n\in\mathbb{N}}$ is relatively compact in the narrow topology.*

In many applications, one also likes to consider *unbounded* functions ϕ in (I.2.15). The following theorem provides a criterion to do so:

Theorem I.9 (Unbounded integrands [4, Lemma 5.1.7]). *Let a narrowly convergent sequence $\mu_n \to \mu$ in $\mathscr{P}(\mathbb{R}^d)$, a continuous function $f : \mathbb{R}^d \to \mathbb{R}$ and a lower semicontinuous function $g : \mathbb{R}^d \to \mathbb{R}_\infty$ be given. The following statements hold:*

(a) If $|f|$ is uniformly integrable *with respect to the set $\{\mu_n : n \in \mathbb{N}\}$, that is*

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{\{x:|f(x)|\geq R\}} |f(x)|\,\mathrm{d}\mu_n(x) = 0,$$

then

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu_n(x) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}\mu(x). \tag{I.2.16}$$

(b) If $g_- := \max(0,-g)$ is uniformly integrable w.r.t. $\{\mu_n : n \in \mathbb{N}\}$, then

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} g(x)\,\mathrm{d}\mu_n(x) \geq \int_{\mathbb{R}^d} g(x)\,\mathrm{d}\mu(x).$$

(c) The property (I.2.16) *holds for every continuous f with* (at most) quadratic growth, *i.e. $|f(x)| \leq A|x|^2 + B$ for suitable $A, B \in \mathbb{R}$, if and only if the map $\psi(x) = |x|^2$ is uniformly integrable w.r.t. $\{\mu_n : n \in \mathbb{N}\}$.*

Remark I.10 (Semicontinuity and uniform integrability of second moments). *If $\mu_n \to \mu$ narrowly and the sequence of second moments converges as $\mathfrak{m}_2(\mu_n) \to \mathfrak{m}_2(\mu)$, then one concludes with the help of Vitali's theorem that $\psi(x) = |x|^2$ is uniformly integrable w.r.t. $\{\mu_n : n \in \mathbb{N}\}$. Hence, Theorem I.9(c) is applicable. By Theorem I.9(a), we have that $\mathfrak{m}_2$ is lower semicontinuous with respect to narrow convergence.*

In this thesis, we sometimes consider absolutely continuous measures (with respect to the Lebesgue measure on $\mathbb{R}^d$) only. For the sake of presentation, we will — by a slight abuse of notation — often identify an absolutely continuous measure with its corresponding Lebesgue density.

I.2.2.1. The Wasserstein distance

The space $\mathscr{P}_2$ of probability measures on $\mathbb{R}^d$ with finite second moment $\mathfrak{m}_2$ has a metric structure with is related to optimal transportation (see [179] for more details on this topic): it can be endowed with the so-called L^2-*Wasserstein distance* $\mathbf{W}_2$ defined as

$$\mathbf{W}_2(\mu_0, \mu_1) := \left[\inf\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\,\mathrm{d}\gamma(x,y):\ \gamma\in\Gamma(\mu_0,\mu_1)\right\}\right]^{1/2}, \tag{I.2.17}$$

where $\Gamma(\mu_0, \mu_1)$ is the set of all *transport plans* or *couplings*, e.g. $\gamma \in \Gamma(\mu_0, \mu_1)$ is a probability measure on the product space $\mathbb{R}^d \times \mathbb{R}^d$ with μ_0 and μ_1 as first and second marginal, respectively. Note that the infimum above is always attained [179, Ch. 2]. We summarize some elementary properties of relevance:

Proposition I.11 (Properties of $\mathbf{W}_2$). *The following statements hold:*

(a) The metric space $(\mathscr{P}_2, \mathbf{W}_2)$ is complete.
(b) $\mathbf{W}_2$ is lower semicontinuous with respect to narrow convergence in both arguments.
(c) A sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}_2$ converges to $\mu \in \mathscr{P}_2$ w.r.t. $\mathbf{W}_2$ if and only if $\mu_n \to \mu$ narrowly and the sequence of second moments converges, $\mathfrak{m}_2(\mu_n) \to \mathfrak{m}_2(\mu)$.
(d) One has $\mathbf{W}_2(\mu, \delta_0)^2 = \mathfrak{m}_2(\mu)$, where δ_0 is the Dirac measure concentrated at the origin $0 \in \mathbb{R}^d$.

Often, one deals with absolutely continuous measures. If $\mu_0 \in \mathscr{P}_2$ is absolutely continuous, the quadratic Wasserstein distance to an arbitrary measure $\mu_1 \in \mathscr{P}_2$ is given by

$$\mathbf{W}_2(\mu_0, \mu_1) = \left[\inf\left\{\int_{\mathbb{R}^d} |t(x)-x|^2\mu_0(x)\,\mathrm{d}x:\ t:\mathbb{R}^d\to\mathbb{R}^d \text{ is measurable and } t_\#\mu_0=\mu_1\right\}\right]^{1/2}.$$

Above, $t_\#\mu_0$ is the *push-forward* or *image measure* of μ_0 via the Borel measurable map t. As before, the infimum is (uniquely) attained by an *optimal transport map* which can be expressed as the (weak) gradient of a convex function [179, Thm. 2.12].

In analogy, one can also define L^p-Wasserstein (pseudo-)distances $\mathbf{W}_p$ on $\mathscr{P}(\mathbb{R}^d)$ for values of $p \in [1, \infty]$ other than $p = 2$:

$$\begin{aligned}\mathbf{W}_p(\mu_0, \mu_1) &= \left[\inf\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p\,\mathrm{d}\gamma(x,y):\ \gamma\in\Gamma(\mu_0,\mu_1)\right\}\right]^{1/p} \quad \text{for } p\in[1,\infty),\\ \mathbf{W}_\infty(\mu_0,\mu_1) &= \inf\left\{\|x-y\|_{L^\infty(\mathbb{R}^d\times\mathbb{R}^d;\,\mathrm{d}\gamma)}:\ \gamma\in\Gamma(\mu_0,\mu_1)\right\},\end{aligned}$$

with the convention that the distance can attain the value $+\infty$ when considering arbitrary elements of $\mathscr{P}(\mathbb{R}^d)$ disregarding finiteness of moments.

Observe that the Wasserstein distances $\mathbf{W}_p$ can also be considered on the space of finite Borel measures with fixed mass $m \in (0, \infty)$ yielding the same metric, up to the p-dependent scaling factor $m^{1/p}$.

In one spatial dimension $d = 1$, the Wasserstein distances can be expressed via the so-called *inverse distribution functions* [179, Ch. 7]: Denoting by $F_\mu : \mathbb{R} \to [0, 1]$ the cumulative distribution function, i.e. $F_\mu(x) = \mu((-\infty, x])$ for $x \in \mathbb{R}$, the *inverse distribution function* $u_\mu : [0, 1] \to \overline{\mathbb{R}}$ is defined as

$$u_\mu(z) = \inf\{x\in\mathbb{R}:\ F_\mu(x) > z\} \quad \text{at each } z\in[0,1].$$

Observe that cumulative distribution functions and inverse distribution functions are increasing and *càdlàg*. The L^p-Wasserstein distance between two measures $\mu_0, \mu_1 \in \mathscr{P}(\mathbb{R})$ is equal to the L^p-distance of their corresponding inverse distribution functions:

$$\mathbf{W}_p(\mu_0, \mu_1) = \|u_{\mu_0} - u_{\mu_1}\|_{L^p([0,1])}.$$

We conclude our introduction to the Wasserstein distance with the following *dynamical characterization* of the L^2-Wasserstein distance on $\mathscr{P}_2$ found by Benamou and Brenier [11]:

$$\mathbf{W}_2(\mu_0,\mu_1)^2 = \inf\left\{\int_0^1\int_{\mathbb{R}^d}|v_t|^2\,\mathrm{d}\mu_t\,\mathrm{d}t : \right. \tag{I.2.18}$$
$$\left.\partial_t\mu_t = -\mathrm{div}(\mu_t v_t) \text{ in the sense of distributions on } [0,1]\times\mathbb{R}^d,\ \mu|_{t=0}=\mu_0, \mu|_{t=1}=\mu_1\right\}.$$

There, μ_0 and μ_1 are linked by the dynamical transport induced by the curve (μ_t, v_t) which is a solution to the *continuity equation* $\partial_t\mu_t = -\mathrm{div}(\mu_t v_t)$. The formulation (I.2.18) gave rise to several generalizations of the Wasserstein distances, e.g. introducing nonlinear mobility functions (see Section I.2.2.3 below).

I.2.2.2. Geodesic convexity and gradient flows

This paragraph is concerned with geodesic λ-convexity of functionals defined on $(\mathscr{P}_2, \mathbf{W}_2)$ and their associated λ-contractive gradient flows. As mentioned above, λ-convexity along geodesics in $\mathscr{P}_2$ does *not* yield the existence of a λ-flow immediately via Theorem I.3 since the additional convexity condition (C) is *not* satisfied in $(\mathscr{P}_2, \mathbf{W}_2)$ in multiple spatial dimensions $d \geq 2$ [4, Ch. 9]. However, this problem can be circumvented using a stronger version of geodesic convexity.

We first introduce the following notation on *projections*: Given a set of coordinates $y^1, \dots, y^K \in \mathbb{R}^d$, $K \in \mathbb{N}$ and an ordered subset $\mathcal{K} \subset \{1,\dots,K\}$, the projection $\pi^{\mathcal{K}}$ is defined as

$$\pi^{\mathcal{K}} : \mathbb{R}^{Kd} \to \mathbb{R}^{|\mathcal{K}|d}, \quad (y^1, \dots y^k) \mapsto (y_k)_{k\in\mathcal{K}}.$$

Definition I.12 (Convexity along generalized geodesics). *Given $\lambda \in \mathbb{R}$, we say that $\mathcal{A} : \mathscr{P}_2 \to \mathbb{R}_\infty$ is λ-*convex along generalized geodesics *in $(\mathscr{P}_2, \mathbf{W}_2)$, if for any triple $\mu^1, \mu^2, \mu^3 \in \mathscr{P}_2$, there exists a Borel measure $\boldsymbol{\mu}$ on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$ such that:*

- *For all $k \in \{1,2,3\}$: $\mu^k = {\pi^k}_\#\boldsymbol{\mu}$.*
- *For $k \in \{2,3\}$, the measure ${\pi^{(1,k)}}_\#\boldsymbol{\mu}$ is optimal in $\Gamma(\mu^1,\mu^k)$, i.e. it realizes the minimum in*

$$\inf\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d}|x^1-x^k|^2\,\mathrm{d}\gamma(x^1,x^k)\middle|\gamma\in\Gamma(\mu^1,\mu^k)\right\}.$$

- *Defining for $s \in [0,1]$ the* generalized geodesic *μ_s connecting μ^2 and μ^3 (with* base point *μ^1) by*

$$\mu_s := \left[(1-s)\pi^2 + s\pi^3\right]_\#\boldsymbol{\mu},$$

  one has for all $s \in [0,1]$:

$$\mathcal{A}(\mu_s) \leq (1-s)\mathcal{A}(\mu^2) + s\mathcal{A}(\mu^3) - \frac{\lambda}{2}s(1-s)\int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d}|x^3-x^2|^2\,\mathrm{d}\boldsymbol{\mu}(x^1,x^2,x^3).$$

The notion of λ-convexity along generalized geodesics is obviously stronger than λ-convexity along geodesics as introduced in Definition I.1 above: just choose $\mu^1 = \mu^3$ — the generalized geodesic μ_s connecting μ^2 and μ^3 then coincides with the actual geodesic curve connecting μ^2 and μ^3. The main benefit we get from this stronger notion is

Proposition I.13 (Existence of gradient flows [4, Lemma 9.2.7]). *Let $\mathcal{A} : \mathscr{P}_2 \to \mathbb{R}_\infty$ be lower semicontinuous, coercive and κ-convex along generalized geodesics for some $\kappa \in \mathbb{R}$. Then, condition* (C) *from Theorem I.3(b) is satisfied and the conclusion from Theorem I.3(b) holds.*

The main classes of λ-convex functionals along generalized geodesics on $(\mathscr{P}_2, \mathbf{W}_2)$ are given by McCann's criteria [133] which we summarize in the following (see also [4, Ch. 9] or [179, Ch. 5] for more details):

Theorem I.14 (Criteria for geodesic convexity on $(\mathscr{P}_2, \mathbf{W}_2)$). *The following statements are true:*

(a) (Internal energy functionals) *Let* $h \in C^0([0,\infty))$ *with* $h(0) = 0$ *be given such that* $\liminf_{z\searrow 0} z^{-\alpha} h(z) > -\infty$ *for some* $\alpha > \frac{d}{d+2}$ *and the map* $r \mapsto r^d h(r^{-d})$ *is convex and nonincreasing on* $(0,\infty)$*. Then, the functional* $\mathcal{A}$ *on* $\mathscr{P}_2$ *defined by*

$$\mathcal{A}(\mu) := \begin{cases} \int_{\mathbb{R}^d} h(w(x))\,\mathrm{d}x & \text{if } \mu = w \cdot \mathcal{L}^d \text{ is absolutely continuous,} \\ +\infty & \text{otherwise,} \end{cases}$$

is 0*-convex along generalized geodesics in* $(\mathscr{P}_2, \mathbf{W}_2)$*.*

(b) (Potential energy functionals) *Let a function* $V \in C^0(\mathbb{R}^d)$ *be given and assume that* V *is* λ*-convex for some* $\lambda \in \mathbb{R}$*. Then, the functional* $\mathcal{A}$ *on* $\mathscr{P}_2$ *defined by*

$$\mathcal{A}(\mu) := \begin{cases} \int_{\mathbb{R}^d} V\,\mathrm{d}\mu & \text{if } V \in L^1(\mathbb{R}^d; \mathrm{d}\mu), \\ +\infty & \text{otherwise,} \end{cases}$$

is λ*-convex along generalized geodesics in* $(\mathscr{P}_2, \mathbf{W}_2)$*.*

(c) (Interaction energy functionals) *Let a function* $W \in C^0(\mathbb{R}^d)$ *be given and assume that* W *is* λ*-convex for some* $\lambda \in \mathbb{R}$*. Then, the functional* $\mathcal{A}$ *on* $\mathscr{P}_2$ *defined by*

$$\mathcal{A}(\mu) := \begin{cases} \frac{1}{2}\int_{\mathbb{R}^d} W * \mu\,\mathrm{d}\mu & \text{if } W * \mu \in L^1(\mathbb{R}^d; \mathrm{d}\mu), \\ +\infty & \text{otherwise,} \end{cases}$$

is $\min(0,\lambda)$*-convex along generalized geodesics in* $(\mathscr{P}_2, \mathbf{W}_2)$*.*

For fixed, but arbitrary $E \in \mathbb{R}^d$*, the functional* $\mathcal{A}^E$ *on* $\mathscr{P}_2$ *defined as*

$$\mathcal{A}^E(\mu) := \begin{cases} \mathcal{A}(\mu) & \text{if } \int_{\mathbb{R}^d} x\,\mathrm{d}\mu(x) = E, \\ +\infty & \text{otherwise,} \end{cases}$$

i.e. as the restriction of $\mathcal{A}$ *to the subspace of those measures in* $\mathscr{P}_2$ *having mean* E*, is* λ*-convex along generalized geodesics in* $(\mathscr{P}_2, \mathbf{W}_2)$*.*

In our forthcoming analysis of evolution equations, we consider the following specific choices for internal energy functionals meeting the assumptions from Theorem I.14(a):

Example I.15 (Boltzmann and Rényi entropy).

(a) *Choosing* $h(z) = z \log z$ *in Theorem I.14(a) above leads to* Boltzmann's entropy

$$\mathcal{A}(u) = \int_{\mathbb{R}^d} u \log u\,\mathrm{d}x,$$

which induces the heat *or* diffusion equation $\partial_t u = \Delta u$ *as its gradient flow [105].*

(b) *Given* $m > 1$*, choosing* $h(z) = \frac{1}{m-1} z^m$ *in Theorem I.14(a) above leads to the* Rényi entropy

$$\mathcal{A}(u) = \frac{1}{m-1}\int_{\mathbb{R}^d} u^m\,\mathrm{d}x,$$

which induces the porous medium equation $\partial_t u = \Delta u^m$ *as its gradient flow [154].*

Observe that since h *is of superlinear growth as* $z \to \infty$ *in both cases,* $\mathcal{A}$ *is lower semicontinuous on* $(\mathscr{P}_2, \mathbf{W}_2)$ *[3].*

Formally, the corresponding evolution equation for a functional $\mathcal{A}$ on $(\mathscr{P}_2, \mathbf{W}_2)$ can be written as

$$\partial_t \mathsf{S}_t^{\mathcal{A}}(w) = \mathrm{div}\left(\mathsf{S}_t^{\mathcal{A}}(w)\mathrm{D}\left(\frac{\delta\mathcal{A}}{\delta w}(\mathsf{S}_t^{\mathcal{A}}(w))\right)\right),$$

where $\frac{\delta \mathcal{A}}{\delta w}$ stands for the usual first variation of the functional $\mathcal{A}$ on $L^2(\mathbb{R}^d)$. It is this formal gradient structure which serves as an indicator for the well-posedness of evolution systems, even if $\mathcal{A}$ is not convex along generalized geodesics.

I.2.2.3. Transportation distances induced by nonlinear mobilities

In this section, we sketch a possible generalization of the Benamou-Brenier formula (I.2.18) for the L^2-Wasserstein distance. Specifically, as Dolbeault, Nazaret and Savaré [74] and Lisini and Marigonda [124] have demonstrated, one may replace the linear *mobility function* $\mathbf{m}(z) = z$ occurring in (I.2.18) by a nonlinear function $\mathbf{m} : S \to \mathbb{R}$ on a closed interval $S \subset \mathbb{R}$ to define the (pseudo-)distance

$$\mathbf{W_m}(\mu_0, \mu_1)^2 = \inf \Bigg\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|w_t|^2}{\mathbf{m}(\mu_t)} \,\mathrm{d}x \,\mathrm{d}t : \\ \partial_t \mu_t = -\mathrm{div}(w_t) \text{ in the sense of distributions on } [0,1] \times \mathbb{R}^d,\ \mu|_{t=0} = \mu_0, \mu|_{t=1} = \mu_1 \Bigg\}$$

on the space $\mathscr{M}(\mathbb{R}^d; S)$ of locally integrable functions μ with values in S. The main assumptions on the *nonlinear mobility* $\mathbf{m} \in C^2(\mathrm{int}(S))$ are

- *Positivity*: $\mathbf{m}(z) > 0$ for all $z \in \mathrm{int}(S)$;
- *Concavity*: $\mathbf{m}''(z) \leq 0$ for all $z \in \mathrm{int}(S)$;
- *Degeneracy*: $\mathbf{m}(z) = 0$ on ∂S.

Formally, $\mathbf{W_m}$ induces a variational structure for evolution equations of the form

$$\partial_t \mu = \mathrm{div} \left(\mathbf{m}(\mu) \mathrm{D} \left(\frac{\delta \mathcal{A}}{\delta w}(\mu) \right) \right); \tag{I.2.19}$$

as their solutions $\mu(t, \cdot)$ may be viewed as curves of steepest descent in the energy landscape of $\mathcal{A}$ with respect to a formal Riemannian structure on $\mathscr{M}(\mathbb{R}^d; S)$. In the article by Lisini, Matthes and Savaré [125], this formal variational structure is the starting point for an existence proof of solutions to a class of fourth-order equations of this type generalizing the Cahn-Hilliard and thin film equations to the case of nonlinear mobility.

The distance $\mathbf{W_m}$ has similar topological properties as the L^2-Wasserstein distance, e.g.,

- it is lower semicontinuous with respect to weak$*$-convergence of the absolutely continuous signed Radon measures $\mu_n \cdot \mathcal{L}^d$ associated to the sequence of densities $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{M}(\mathbb{R}^d; S)$;
- it is convex in both arguments w.r.t. the usual linear interpolation;
- $\mathbf{W_m}$-bounded sets are weak$*$-relatively compact.

In contrast to this rather straightforward generalization, the issue of geodesic convexity w.r.t. those generalized Wasserstein distances is more delicate for nonlinear mobilities $\mathbf{m}$. A generalized McCann condition for internal energy functionals of the form $\mathcal{A}(\mu) = \int_{\mathbb{R}^d} f(\mu) \,\mathrm{d}x$ has been found by Carrillo, Lisini, Savaré and Slepčev [53]:

THEOREM I.16 (Generalized McCann condition [53]). *Let $f : S \to \mathbb{R}$ be smooth and define a functional $\mathcal{A} : \mathscr{M}(\mathbb{R}^d; S) \to \mathbb{R}_\infty$ via $\mathcal{A}(\mu) = \int_{\mathbb{R}^d} f(\mu) \,\mathrm{d}x$. Assume that $\mathcal{A}$ is proper, bounded from below, lower semicontinuous with respect to the distance $\mathbf{W_m}$ and that the following* d-dimensional generalized McCann condition for the mobility $\mathbf{m}$ *holds:*

$$f''(z)\mathbf{m}(z)^2 \geq \left(1 - \frac{1}{d}\right) H(z) \geq 0 \quad \textit{for all } z \in \mathrm{int}(S), \quad \textit{where } H'(z) = f''(z)\mathbf{m}(z)\mathbf{m}'(z). \tag{I.2.20}$$

Then, $\mathcal{A}$ is 0-convex along geodesics in $(\mathscr{M}(\mathbb{R}^d; S), \mathbf{W_m})$ and generates a 0-flow $\mathsf{S}^{\mathcal{A}}$.

EXAMPLE I.17 (Heat equation). *Choosing f such that $f''(z) = \mathbf{m}(z)^{-1}$ for all $z \in \operatorname{int}(S)$ yields* (I.2.20) *since $H(r) = \mathbf{m}(r) \geq 0$. Hence, the functional $\mathcal{A}_{\text{heat}}(\mu) = \int_{\mathbb{R}^d} f(\mu)\,\mathrm{d}x$ is 0-convex along geodesics in $(\mathscr{M}(\mathbb{R}^d;S),\mathbf{W}_{\mathbf{m}})$ and its associated 0-flow is a solution to the heat equation.*

In contrast to that, potential (and similarly also interaction) energy functionals of the form $\int_{\mathbb{R}^d} \mu V\,\mathrm{d}x$ are — even for $V \in C_c^\infty(\mathbb{R}^d)$ — never λ-convex along geodesics, unless $\mathbf{m}$ is linear [53]. However, in combination with the energy inducing the heat flow $\mathcal{A}_{\text{heat}}$ from Example I.17, one can prove that, given $\nu > 0$, the functional $\mathcal{A}_\nu$ defined as

$$\mathcal{A}_\nu(\mu) := \nu\mathcal{A}_{\text{heat}}(\mu) + \int_{\mathbb{R}^d} \mu V\,\mathrm{d}x$$

generates a λ-flow in $(\mathscr{M}(\mathbb{R}^d;S),\mathbf{W}_{\mathbf{m}})$ with $\lambda = -\frac{C}{\nu}$ for a constant $C > 0$ depending on the confinement potential V and on the spatial dimension d [125].

I.2.3. Generalized gradient systems

Gradient structures as introduced above can also be seen in the purely Riemannian framework (see, e.g., [134]). Consider the abstract evolution equation

$$\dot{u} = -\mathbf{F}(u), \tag{I.2.21}$$

where u is an element of the so-called *state space* $\mathbf{X}$ with dual $\mathbf{X}^*$. Equation (I.2.21) is said to have *gradient structure* if there exists a driving entropy functional $\mathcal{E} : \mathbf{X} \to \mathbb{R}_\infty$ and, for each fixed $u \in \mathbf{X}$, a symmetric, positive semidefinite operator $\mathbf{G}(u) : \mathsf{T}_u\mathbf{X} \to \mathsf{T}_u^*\mathbf{X}$ mapping the tangent space of $\mathbf{X}$ at u to its cotangent space, such that

$$\dot{u} = -\mathbf{F}(u) \quad \Leftrightarrow \quad -\mathbf{G}(u)\dot{u} = \mathrm{D}\mathcal{E}(u) \quad \Leftrightarrow \quad \dot{u} = -\nabla_{\mathbf{G}}\mathcal{E}(u). \tag{I.2.22}$$

The (formal) "gradient" operator $\nabla_{\mathbf{G}}$ is calculated by means of the so-called *Onsager operator* $\mathbf{K}(u) = \mathbf{G}(u)^{-1}$ (in reference to *Onsager's principle* [152]) yielding the rate equation

$$\dot{u} = -\mathbf{K}(u)\mathrm{D}\mathcal{E}(u).$$

The triple $(\mathbf{X},\mathcal{E},\mathbf{G})$ (or, equivalently, $(\mathbf{X},\mathcal{E},\mathbf{K})$) is called *(classical) gradient system*. Note that if a suitable version of the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u) = \langle \mathrm{D}\mathcal{E}(u), \dot{u}\rangle$$

is available, the rate equation (I.2.22) is equivalent to the *energy dissipation balance*

$$\mathcal{E}(u(T)) + \frac{1}{2}\int_0^T \langle \mathbf{G}(u)\dot{u}, \dot{u}\rangle\,\mathrm{d}t + \frac{1}{2}\int_0^T \langle \mathrm{D}\mathcal{E}(u), \mathbf{K}(u)\mathrm{D}\mathcal{E}(u)\rangle\,\mathrm{d}t = \mathcal{E}(u(0)), \tag{I.2.23}$$

at each $T > 0$. Actually, this formulation is formally equivalent to the metric energy dissipation balance (I.2.6) from Section I.2.1.

EXAMPLE I.18 (Gradient flows in $(\mathscr{P}_2(\mathbb{R}^d),\mathbf{W}_2)$ and $L^2(\mathbb{R}^d)$). *Consider the case $\mathbf{X} = L^2(\mathbb{R}^d)$. A seemingly trivial example for an Onsager operator is given by*

$$\mathbf{K}(u) = \mathrm{id} \quad \textit{for all } u \in \mathbf{X}.$$

The associated evolution equation is that of a gradient flow with respect to the metric induced by the $L^2(\mathbb{R}^d)$ norm:

$$\dot{u} = -\mathrm{D}_u\mathcal{E}(u).$$

Defining on $\mathbf{X} = \mathscr{P}_2(\mathbb{R}^d)$ the mapping

$$\mathbf{K}(u)\xi = -\mathrm{div}(u\mathrm{D}_x\xi) \quad \textit{for } \xi \in \mathsf{T}_u^*\mathbf{X}$$

leads to a formal gradient flow equation w.r.t. the L^2-Wasserstein distance (see Section I.2.2.2),

$$\dot{u} = \operatorname{div}(u\mathrm{D}_x\mathrm{D}_u\mathcal{E}(u)).$$

In view of Examples I.4&I.15, a specific evolution equation (there, the heat equation) can possess several truely different gradient structures.

The operator $\mathbf{G}$ can be viewed as a metric tensor inducing the (pseudo-)distance $\mathbf{d_G} : \mathbf{X} \times \mathbf{X} \to [0, \infty]$, defined by the following analog of the Benamou-Brenier formula (I.2.18) for the quadratic Wasserstein distance [123]:

$$\mathbf{d_G}(u_0, u_1) = \left[\inf \left\{ \int_0^1 \langle \mathbf{G}(\gamma_s)\dot{\gamma}_s, \dot{\gamma}_s \rangle \, \mathrm{d}s : \gamma \in C^1([0,1]; \mathbf{X}),\ \|\dot{\gamma}_s\| = 1\ \forall s \in [0,1],\ \gamma_0 = u_0,\ \gamma_1 = u_1 \right\} \right]^{1/2}.$$

A sufficient criterion for λ-convexity along geodesics in the space $(\mathbf{X}, \mathbf{d_G})$ has been established by Liero and Mielke in [123] (see also [135]):

Theorem I.19 (Abstract convexity condition). *Assume that the rate equation* (I.2.22) *corresponding to the gradient system* $(\mathbf{X}, \mathcal{E}, \mathbf{G})$ *possesses a local-in-time continuous semiflow* S *and assume that* $\mathcal{E}$ *is proper, lower semicontinuous w.r.t.* $\mathbf{d_G}$ *and bounded from below. If for given* $\lambda \in \mathbb{R}$*, all* $u \in \mathbf{X}$ *and all* $\xi \in \mathsf{T}^*_u\mathbf{X}$*, one has*

$$\langle \xi, \mathrm{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi \rangle - \frac{1}{2} \langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle \geq \lambda \langle \xi, \mathbf{K}(\mu)\xi \rangle ,$$

then S *satisfies the evolution variational estimate* (I.2.4) *for* $\mathcal{E}$ *and hence defines a* λ*-flow on* $(\mathbf{X}, \mathbf{d_G})$*. Further,* $\mathcal{E}$ *is* λ*-convex along geodesics in* $(\mathbf{X}, \mathbf{d_G})$*.*

The proof of this theorem in [123] is based on the *Eulerian calculus* developed by Otto and Westdickenberg [155] and Daneri and Savaré [64]. Natural applications of Theorem I.19 involve metrics of transportation type such as the L^2-Wasserstein distance $\mathbf{W}_2$ or their generalizations with nonlinear mobility $\mathbf{W_m}$ from Section I.2.2.3.

As a further generalization, one may consider an energy dissipation identity of the form (I.2.23) where the dissipation density does not necessarily consist of quadratic terms anymore:

$$\mathcal{E}(u(T)) + \int_0^T \{\Psi(u; \dot{u}) + \Psi^*(u; -\mathrm{D}\mathcal{E}(u))\} \, \mathrm{d}t = \mathcal{E}(u(0)). \qquad \textbf{(EDB)}$$

Above, $\Psi : \mathsf{T}\mathbf{X} \to [0, \infty]$ and $\Psi^* : \mathsf{T}^*\mathbf{X} \to [0, \infty]$ are called *primal* and *dual dissipation potentials*, and are defined as Legendre duals, i.e.,

$$\Psi(z; v) := \sup_{\xi \in \mathsf{T}^*_z\mathbf{X}} \{\langle \xi, v \rangle - \Psi^*(z; \xi)\} \quad \text{for } v \in \mathsf{T}_z\mathbf{X},$$

assuming that Ψ^* is convex and lower semicontinuous and that $\Psi(z; 0) = 0$ holds. For more details on convex analysis, we refer to the monograph by Ekeland and Témam [77].

There are various models which can be interpreted as (generalized) gradient systems, see, for instance, [134, 135, 136, 91, 140, 139], or — as an overview — the article [138]. Note that one may also consider non-autonomous problems by allowing for an explicit dependency of $\mathcal{E}$ on the time t. For simplicity, we consider the autonomous case here. According to [138, Thm. 3.2] (again assuming that the chain rule holds), (**EDB**) is equivalent to the *upper energy dissipation estimate*

$$\mathcal{E}(u(T)) + \int_0^T \left\{\Psi(u; \dot{u}) + \Psi^*(u; -\mathrm{D}\mathcal{E}(u))\right\} \mathrm{d}t \leq \mathcal{E}(u(0)). \qquad \textbf{(UEDE)}$$

By means of the so-called *Young-Fenchel estimate*

$$\Psi(z;v) + \Psi^*(z;\xi) \geq \langle \xi, v \rangle \,, \tag{I.2.24}$$

which holds for all $z \in \mathbf{X}$, $\xi \in \mathsf{T}_z^*\mathbf{X}$ and $v \in \mathsf{T}_z\mathbf{X}$ by definition of the Legendre transform, one arrives at the following equivalent formulations for (**EDB**):

$$\begin{aligned} \Psi(u(t);\dot u(t)) + \Psi^*(u(t);-\mathrm{D}\mathcal{E}(u(t))) = -\langle \mathrm{D}\mathcal{E}(u(t)), \dot u(t)\rangle \qquad &\text{(power balance);}\\ 0 \in \partial_v \Psi(u(t);\dot u(t)) + \mathrm{D}\mathcal{E}(u(t)) \qquad &\text{(force balance);}\\ \dot u(t) \in \partial_\xi \Psi^*(u(t);-\mathrm{D}\mathcal{E}(u(t))) \qquad &\text{(rate equation).} \end{aligned}$$

There, ∂ denotes the *convex subdifferential*: recall that the convex subdifferential of a proper, convex, lower semicontinuous function $f : \mathbf{Y} \to \mathbb{R}_\infty$ at $y \in \mathbf{Y}$ is given by

$$\partial f(y) = \{\eta \in \mathbf{Y}^* \mid f(y') - f(y) \geq \langle \eta, y' - y \rangle \;\; \forall y' \in \mathbf{Y}\}.$$

We call the triple $(\mathbf{X}, \mathcal{E}, \Psi)$ (or, equivalently, $(\mathbf{X}, \mathcal{E}, \Psi^*)$) *generalized gradient system*.

In this thesis, we are primarily interested in convergence of families of generalized gradient systems $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)_{\varepsilon>0}$ in the (possibly singular) limit $\varepsilon \searrow 0$. For this purpose, a notion of evolutionary convergence for generalized gradient systems was defined by Mielke in [138], originally developed by Sandier and Serfaty in the seminal paper [160] where their notion of variational convergence was applied to the *Ginzburg-Landau* functional.

We first recall a notion of variational convergence for static functionals (for a detailed exposition, consult the monographs by Attouch [8] or Braides [33]):

Definition I.20 (Gamma- and Mosco-convergence). *Let $(\Phi_\varepsilon)_{\varepsilon>0}$ be a family of functionals $\Phi_\varepsilon : \mathbf{Y} \to \mathbb{R}_\infty$, where $\mathbf{Y}$ is a Banach space, and let $\Phi : \mathbf{Y} \to \mathbb{R}_\infty$.*

(a) *$(\Phi_\varepsilon)_{\varepsilon>0}$ is said to* Γ-converge *to Φ with respect to the strong [respectively, weak] convergence in $\mathbf{Y}$ (write $\Phi_\varepsilon \xrightarrow{\Gamma} \Phi$ [$\Phi_\varepsilon \overset{\Gamma}{\rightharpoonup} \Phi$]) as $\varepsilon \searrow 0$, if the following conditions are satisfied:*
   - *(i)* (Liminf estimate) *If $u_\varepsilon \to u$ [$u_\varepsilon \rightharpoonup u$], then $\liminf_{\varepsilon\searrow 0} \Phi_\varepsilon(u_\varepsilon) \geq \Phi(u)$.*
   - *(ii)* (Recovery sequences) *For all $u \in \mathbf{Y}$, there exists a sequence $(\widehat{u}_\varepsilon)_{\varepsilon>0}$ with $\widehat{u}_\varepsilon \to u$ [$\widehat{u}_\varepsilon \rightharpoonup u$] and $\limsup_{\varepsilon\searrow 0} \Phi_\varepsilon(\widehat{u}_\varepsilon) \leq \Phi(u)$.*

(b) *$(\Phi_\varepsilon)_{\varepsilon>0}$ is said to* converge in the sense of Mosco *to Φ as $\varepsilon \searrow 0$ (write $\Phi_\varepsilon \xrightarrow{\mathsf{M}} \Phi$), if both $\Phi_\varepsilon \xrightarrow{\Gamma} \Phi$ and $\Phi_\varepsilon \overset{\Gamma}{\rightharpoonup} \Phi$.*

Clearly, if $\mathbf{Y}$ is finite-dimensional, Gamma- and Mosco-convergence are equivalent. We will make use of the following notion of *evolutionary Γ-convergence*:

Definition I.21 (Evolutionary Γ-convergence [138]). *Let a family of gradient systems $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)_{\varepsilon>0}$ be given. We say that $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)_{\varepsilon>0}$* (strongly) E-converges *to a limit gradient system $(\mathbf{X}, \mathcal{E}, \Psi)$ as $\varepsilon \searrow 0$ and write $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \xrightarrow{\mathsf{E}} (\mathbf{X}, \mathcal{E}, \Psi)$, if, given a sequence of solutions $u_\varepsilon : [0,T] \to \mathbf{X}$ to $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon)$ with $u_\varepsilon(0) \to u^0$ as $\varepsilon \searrow 0$, there exists a limit solution $u : [0,T] \to \mathbf{X}$ to $(\mathbf{X}, \mathcal{E}, \Psi)$ with $u(0) = u^0$ and a subsequence $\varepsilon_k \searrow 0$ $(k \to \infty)$ such that for all $t \in (0,T]$:*

$$u_{\varepsilon_k}(t) \to u(t) \quad \textit{and} \quad \mathcal{E}_{\varepsilon_k}(u_{\varepsilon_k}(t)) \to \mathcal{E}(u(t)) \quad \textit{as } k \to \infty.$$

Notice that in order to obtain evolutionary Γ-convergence, it does in general *not* suffice that $\mathcal{E}_\varepsilon \xrightarrow{\mathsf{M}} \mathcal{E}$ and $\Psi_\varepsilon \xrightarrow{\mathsf{M}} \Psi$. For $(\mathbf{X}, \mathcal{E}_\varepsilon, \Psi_\varepsilon) \xrightarrow{\mathsf{E}} (\mathbf{X}, \mathcal{E}, \Psi)$, certain additional compatibility conditions between the energy and dissipation functionals have to be fulfilled (depending on the specific problem at hand) [138].

Part II

Two-species systems modelling cell motion and ion transport

CHAPTER II.1

Introduction to Part II

This part of the thesis is based on the author's articles [185, 187, 188] and the joint work [189] with Daniel Matthes.

II.1.1. Keller-Segel models

It is mostly devoted to the mathematical analysis of the following variant of the Keller-Segel model for chemotaxis: for $x \in \mathbb{R}^d$ and $t \in [0, \infty)$, we consider

$$\begin{aligned} \partial_t u(t,x) &= K_u \Delta u^m(t,x) + \operatorname{div}\left(u(t,x)\mathrm{D}\left[W(x) + \chi\phi(v(t,x))\right]\right), \\ \partial_t v(t,x) &= K_v \Delta v(t,x) - \kappa v(t,x) - \alpha u(t,x)\phi'(v(t,x)). \end{aligned} \tag{II.1.1}$$

Above, K_u, K_v and κ are nonnegative parameters, $m \geq 1$ is the diffusion exponent. The map $W : \mathbb{R}^d \to \mathbb{R}$ is a confining potential with additional properties to be specified below. We always assume that the nonlinearity $\phi \in C^2(\mathbb{R})$ is convex and strictly decreasing. Formally, if $\chi \in \mathbb{R}$ and $\alpha \in \mathbb{R}$ have the same sign, system (II.1.1) possesses a gradient flow structure with respect to the compound Wasserstein-L^2 distance

$$\mathbf{d}((u,v),(\widetilde{u},\widetilde{v})) := \sqrt{\mathbf{W}_2(u,\widetilde{u})^2 + \frac{\chi}{\alpha}\|v - \widetilde{v}\|_{L^2}^2}, \tag{II.1.2}$$

on the metric space $\mathbf{X} = \mathscr{P}_2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. The associated energy functional $\mathcal{E} : \mathbf{X} \to \mathbb{R}_\infty$ reads

$$\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^d} \left(\frac{K_u u^m}{m-1} + uW + \frac{K_v\chi}{2\alpha}|\mathrm{D}v|^2 + \frac{\chi\kappa}{2\alpha}v^2 + \chi u\phi(v)\right) \mathrm{d}x & \text{if } m > 1, \\ \int_{\mathbb{R}^d} \left(K_u u\log u + uW + \frac{K_v\chi}{2\alpha}|\mathrm{D}v|^2 + \frac{\chi\kappa}{2\alpha}v^2 + \chi u\phi(v)\right) \mathrm{d}x & \text{if } m = 1, \end{cases} \tag{II.1.3}$$

with the convention that $\mathcal{E}(u,v) := +\infty$ if one of the integrals on the r.h.s. in (II.1.3) is not well-defined, e.g. if u is not absolutely continuous. Without loss of generality, we set $K_v = 1$ and $\chi = \alpha \neq 0$ in the following.

II.1.1.1. Modelling background

Systems of the form (II.1.1) arise as spatial models for population dynamics, in particular for microbial growth and movement. The first model of this kind — with the linear sensitivity function $\phi(w) = -w$ — has been set up by Keller and Segel [110] as description of slime mould aggregation. There, the individuals of a population respond to gradients of chemical substances (*chemotaxis*). Chemotactic processes occur in many (and highly different) biological systems; for the biological details, we refer to the book by Eisenbach [76]. For example, many bacteria like *Escherichia coli* possess *flagella* driven by small motors which respond to gradients of signalling molecules in the environment. Chemotaxis also plays an important role in embryonal development, e.g. in the development of blood vessels (*angiogenesis*), which is also a crucial step in tumour growth. Starting from the classical Keller-Segel model, many different model extensions are conceivable. A broad range of those is summarized in the review articles by Hillen and Painter [96] and Horstmann [100]. Details on the modelling aspects can be found e.g. in the books by Murray [148] and Perthame [156].

In the model (II.1.1) under consideration here, u is the time-dependent spatial density of the cells, and v is the time-dependent concentration of the signalling substance which either attracts or repels the cells: depending on the sign of χ, the microorganisms move towards regions with higher concentration of the signalling substance and produce it ($\chi > 0$) or to regions with lower concentration and degrade it ($\chi < 0$). Two special aspects are included in this particular model: nonlinear diffusion, i.e., the use of a non-constant, u-dependent mobility coefficent for the diffusive motion of the bacteria, and signal-dependent chemotactic sensitivity, i.e. the use of the — in general nonlinear — response $\phi(v)$. Biological populations might often be described more accurately by diffusion of porous medium type than by Fick's laws since a population-dependent, increasing diffusion coefficient prevents overcrowding effects (see e.g. [96]). Our second extension of the classical model is motivated by the fact that the conversion of an external signal into a reaction of the considered microorganism (*signal transduction*) often occurs by binding and dissociation of molecules to certain receptors. The movement of the cell is then caused rather by gradients in the number of receptors occupied by signalling molecules than by concentration gradients of signalling molecules themselves. For growing concentrations, the number of bound receptors may exhibit a saturation, such that the gradient vanishes. In [96, 163, 117], this was included into the model by the *chemotactic sensitivity function*

$$\phi'(w) = -\frac{1}{(1+w)^2} \qquad \text{for } w \geq 0.$$

In our model, three paradigmatic examples arise for $w \geq 0$:

$$\begin{aligned} \phi(w) &= -w && \text{(classical Keller-Segel model)},\\ \phi(w) &= -\log(1+w) && \text{(weak saturation effect)},\\ \phi(w) &= \frac{1}{1+w} && \text{(strong saturation effect)}. \end{aligned}$$

For the dynamics of the signalling substance, we assume linear diffusion according to Fick's laws and degradation with a constant, exponential rate κ. The term $-\chi\phi'(v)$ models the production or degradation of signalling substance by the microorganisms; here it is taken into account that the cells might be the less active in producing additional substance the higher its local concentration already is (in agreement with the models presented in [100, Sect. 6]). Finally, an external background potential W is included in order to generate a spatial confinement of the bacterial population.

II.1.1.2. Main results

More specifically, we are concerned with three versions of the model (II.1.1) in Chapters II.2 and II.3. In Chapter II.2, we first study a system in at least three spatial dimensions with supercritical diffusion exponent m and linear sensitivity ϕ (see also [185]):

Assumption II.1 (Model with linear sensitivity). *We require $d \geq 3$, $m > 2 - \frac{2}{d}$ and $\phi(w) = -w$. Moreover, $W \in C^2(\mathbb{R}^d)$ shall be bounded from below and grow at most quadratically, that is $W(x) \leq A|x|^2 + B$ for all $x \in \mathbb{R}^d$ and suitable $A, B \geq 0$; and we assume that $\Delta W \in L^\infty(\mathbb{R}^d)$. Finally, we set $K_u = 1$ and $\kappa = 0$ for the sake of readability.*

In Section II.2.1, we prove the existence of weak solutions to (II.1.1) in the setting of Assumption II.1. Our main result is the following

Theorem II.2 (Existence of weak solutions to (II.1.1)). *Consider* (II.1.1) *together with the initial condition*

$$(u(0,\cdot), v(0,\cdot)) = (u_0, v_0) \qquad \text{on } \mathbb{R}^d, \tag{II.1.4}$$

and assume that Assumption II.1 holds. Let $(u_0, v_0) \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d)$*. Define for each* $\tau > 0$ *a function* (u_τ, v_τ) *via the scheme* (I.2.11) *and* (I.2.12)*. Then, there is a sequence* $(\tau_k)_{k\in\mathbb{N}}$ *with* $\tau_k \searrow 0$ *such that* (u_{τ_k}, v_{τ_k}) *converges to a weak solution* $(u,v) : [0,\infty) \times \mathbb{R}^d \to [0,\infty] \times \overline{\mathbb{R}}$ *to* (II.1.1) *in the following sense: For each* $T > 0$*:*

$$u_{\tau_k}(t,\cdot) \to u(t,\cdot) \text{ narrowly in } \mathscr{P}(\mathbb{R}^d)\text{, pointwise with respect to } t \in [0,T],$$
$$v_{\tau_k}(t,\cdot) \to v(t,\cdot) \text{ in } L^2(\mathbb{R}^d)\text{, uniformly with respect to } t \in [0,T].$$

Furthermore,

$$\partial_t u = \Delta u^m + \operatorname{div}(u\mathrm{D}[W - \chi v]) \quad \text{in the sense of distributions on } (0,\infty) \times \mathbb{R}^d,$$
$$\partial_t v = \Delta v + \chi u \qquad \text{almost everywhere in } (0,\infty) \times \mathbb{R}^d,$$

and (II.1.4) *is attained for almost all* $x \in \mathbb{R}^d$*.*

Moreover, for all $T > 0$*, one has*

$$\begin{aligned} u &\in C^{1/2}([0,T]; \mathscr{P}_2(\mathbb{R}^d)) \cap L^\infty([0,T]; L^m(\mathbb{R}^d)) \cap L^2([0,T]; L^2(\mathbb{R}^d)),\\ v &\in C^{1/2}([0,T]; L^2(\mathbb{R}^d)) \cap L^\infty([0,T]; W^{1,2}(\mathbb{R}^d)) \cap L^2([0,T]; W^{2,2}(\mathbb{R}^d)) \cap W^{1,2}([0,T]; L^2(\mathbb{R}^d)),\\ u^{m/2} &\in L^2([0,T]; W^{1,2}(\mathbb{R}^d)). \end{aligned}$$

In the special case $\chi > 0$*, additionally* $v(t,\cdot) \geq 0$ *holds a.e. on* $\mathbb{R}^d$ *for all* $t > 0$*, given a nonnegative initial datum, viz.* $v_0 \geq 0$ *a.e. on* $\mathbb{R}^d$*.*

The remaining part of Chapter II.2 is then devoted to the analysis of a more specific system with nonlinear sensitivity ϕ (see also [189]). There, the central point is the analysis of the long-time behaviour of solutions. The existence proof for solutions will be omitted in this case since the method of Section II.2.1 applies *mutatis mutandis* and an analogous statement to Theorem II.2 holds (see [189]). More in detail, we require:

Assumption II.3 (Model with nonlinear sensitivity). *We assume that* $d = 3$*,* $m = 2$*,* $\kappa > 0$ *and* $\chi > 0$*. The nonlinearity* $\phi \in C^2(\mathbb{R})$ *is convex and decreasing, with*

$$0 < -\phi'(w) \leq \overline{\phi'} < \infty \quad \text{and} \quad 0 \leq \phi''(w) \leq \overline{\phi''} < \infty \qquad \text{for all } w \in \mathbb{R},$$

and some constants $\overline{\phi'} > 0$ *and* $\overline{\phi''} \geq 0$*. Furthermore, the confinement potential* $W \in C^2(\mathbb{R}^3)$ *shall be* λ_0*-uniformly convex (that is,* $\mathrm{D}^2 W(x) \geq \lambda_0 \mathbb{1}$ *in the sense of symmetric matrices, at each* $x \in \mathbb{R}^3$*) for some* $\lambda_0 > 0$ *and its partial derivatives of second order shall be uniformly bounded. For convenience, we set* $K_u = \frac{1}{2}$*.*

Assuming a small coupling strength $\chi > 0$ (which will be denoted by ε in the following to emphasize the smallness), we obtain the following result on exponential convergence to equilibrium:

Theorem II.4 (Exponential convergence of solutions to (II.1.1)). *Consider* (II.1.1) *together with an initial datum* (u_0, v_0) *and assume that Assumption II.3 holds.*

There are constants $\overline{\varepsilon} > 0$ *and* $\overline{L} > 0$ *such that for all* $\delta, \delta' > 0$*, there exists* $C_{\delta,\delta'} > 0$ *such that the following is true for every* $\chi = \varepsilon \in (0, \overline{\varepsilon})$ *and initial conditions* $(u_0, v_0) \in (\mathscr{P}_2(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)) \times W^{1,2}(\mathbb{R}^3)$ *with* $v_0 \in L^{6/5}(\mathbb{R}^3)$*:*

A weak solution (u,v) *to* (II.1.1) *obtained as a limit of the scheme* (I.2.11) *(see Theorem II.22 below) converges to the nonnegative unique stationary solution* $(u_\infty, v_\infty) \in (\mathscr{P}_2(\mathbb{R}^3) \cap W^{1,2}(\mathbb{R}^3)) \times (C^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3))$ *of* (II.1.1) *exponentially fast with rate* $\Lambda_\varepsilon := \min(\lambda_0, \kappa) - \overline{L}\varepsilon > 0$ *in the following sense:*

$$\begin{aligned} &\mathbf{W}_2(u(t,\cdot), u_\infty) + \|u(t,\cdot) - u_\infty\|_{L^2} + \|v(t,\cdot) - v_\infty\|_{W^{1,2}} \\ &\qquad \leq C_{\delta,\delta'}(1 + \|v_0\|_{L^{6/5}})^{1+\delta'} \left(\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + 1\right)^{1+\delta} e^{-\Lambda_\varepsilon t} \qquad \text{for all } t \geq 0. \end{aligned} \tag{II.1.5}$$

In Chapter II.3, we perform a similar analysis for a version of (II.1.1) in space dimension $d = 1$ (see also [187]):

Assumption II.5 (One-dimensional model: requirements for existence). *Let $d = 1$, $m = 1$, $K_u = 1$ and $\phi(w) = -w$ and assume that the confinement potential $W \in C^2(\mathbb{R})$ is bounded from below and has a globally bounded second derivative.*

These assumptions guarantee the existence of weak solutions, again by the method presented in Chapter II.2:

Theorem II.6 (Existence of weak solutions to (II.1.1) in one dimension). *Consider* (II.1.1) *under Assumption II.5, together with an initial condition $(u_0, v_0) \in \mathscr{P}_2(\mathbb{R}) \times W^{1,2}(\mathbb{R})$ with $\int_{\mathbb{R}} u_0 \log u_0 \,\mathrm{d}x < \infty$. Define, for each $\tau > 0$, a discrete solution by means of* (I.2.11)&(I.2.12). *Then, there exists a vanishing sequence $\tau_k \searrow 0$ $(k \to \infty)$ such that (u_{τ_k}, v_{τ_k}) converges to a weak solution (u, v) to* (II.1.1) *in the sense that the differential equation for u holds in the sense of distributions, whereas the equation for v holds almost everywhere in $[0, \infty) \times \mathbb{R}$. Specifically, one has for all $T > 0$:*

$$
\begin{aligned}
&u_{\tau_k} \rightharpoonup u \text{ narrowly in } \mathscr{P}(\mathbb{R}), \text{ pointwise with respect to } t \in [0, T],\\
&v_{\tau_k} \to v \text{ in } L^2(\mathbb{R}), \text{ uniformly with respect to } t \in [0, T],\\
&u \in C^{1/2}([0, T]; (\mathscr{P}_2(\mathbb{R}), \mathbf{W}_2)) \cap L^1([0, T]; L^\infty(\mathbb{R})) \cap L^2([0, T]; L^2(\mathbb{R})),\\
&\sqrt{u} \in L^2([0, T]; W^{1,2}(\mathbb{R})), \quad u \log u \in L^\infty([0, T]; L^1(\mathbb{R})),\\
&v \in C^0([0, T] \times \mathbb{R}) \cap W^{1,2}([0, T]; L^2(\mathbb{R})) \cap L^\infty([0, T]; W^{1,2}(\mathbb{R})) \cap L^2([0, T]; W^{2,2}(\mathbb{R})).
\end{aligned}
$$

Again, for the analysis of the long-time behaviour, we need the stronger

Assumption II.7 (One-dimensional model: requirements for convergence). *In addition to Assumption II.5, assume $\chi > 0$, $\kappa > 0$ and uniform convexity of W, i.e. $W_{xx}(x) \geq \lambda_0$ for all $x \in \mathbb{R}$, with some $\lambda_0 > 0$.*

In analogy to Theorem II.4, our result reads:

Theorem II.8 (Long-time behaviour of (II.1.1) in one dimension). *Consider* (II.1.1) *under Assumption II.7, together with an initial condition $(u_0, v_0) \in \mathscr{P}_2(\mathbb{R}) \times W^{1,2}(\mathbb{R})$ with $\int_{\mathbb{R}} u_0 \log u_0 \,\mathrm{d}x < \infty$. There exist $\bar{\varepsilon} > 0$, $C > 0$ and $L > 0$ such that for all $\chi = \varepsilon \in (0, \bar{\varepsilon})$, the following statements hold:*

(a) The system (II.1.1) *possesses a unique stationary state $(u_\infty, v_\infty) \in (\mathscr{P}_2 \cap L^\infty)(\mathbb{R}) \times W^{2,2}(\mathbb{R})$ satisfying*

$$
\begin{aligned}
u_\infty &= U_\varepsilon \exp(-W + \varepsilon v_\infty), \quad \text{with } U_\varepsilon > 0 \text{ such that } \|u_\infty\|_{L^1} = 1,\\
\partial_{xx} v_\infty &= \kappa v_\infty - \varepsilon u_\infty.
\end{aligned}
$$

(b) One has $\Lambda_\varepsilon := \min(\kappa, \lambda_0) - \varepsilon L > 0$ and for all $t \geq 0$, the weak solution (u, v) to (II.1.1) *from Theorem II.6 admits the estimate*

$$
\begin{aligned}
&\|u(t, \cdot) - u_\infty\|_{L^1} + \mathbf{W}_2(u(t, \cdot), u_\infty) + \sup_{x \in \mathbb{R}} |v(t, \cdot) - v_\infty| + \|v(t, \cdot) - v_\infty\|_{W^{1,2}}\\
&\qquad \leq C(\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty))^{1/2} e^{-\Lambda_\varepsilon t},
\end{aligned}
\tag{II.1.6}
$$

i.e. $(u(t, \cdot), v(t, \cdot))$ converges exponentially fast with rate Λ_ε to the equilibrium (u_∞, v_∞) as $t \to \infty$.

II.1.1.3. Related studies

Systems of Keller-Segel type have been a very popular object of investigation during the last decades. The rapidly growing mathematical literature about the Keller-Segel model and its manifold variants is devoted primarily to the dichotomy *global existence* versus *finite-time blow-up* of (weak, possibly measure-valued) solutions, but the long-time behaviour of global solutions — and their self-similarity — has been intensively investigated as well.

Relatively much attention has been drawn to the parabolic-elliptic (Patlak-)Keller-Segel model and its variants, where the dynamics for the signalling substance are assumed to be in equilibrium and which thus can be reduced to a single nonlocal scalar equation for the bacterial density (see, for instance, [104, 114, 127, 37, 128, 40, 73, 41])

Global existence and blow-up in the classical parabolic-parabolic Keller-Segel model, which is (II.1.1) with $\phi(w) = -w$, $W \equiv 0$ and linear diffusion $m = 1$, has been thoroughly studied by Calvez and Corrias [39] in space dimension $d = 2$, and by Corrias and Perthame [62] in higher space dimensions $d > 2$, see also [18, 115, 144, 150, 165, 170, 69, 182, 40]. Recently, in [45], uniqueness and long-time behaviour of solutions to the parabolic-parabolic Keller-Segel system was studied by means of a perturbation of the parabolic-elliptic framework.

Variants with nonlinear diffusion and drift have been studied for instance by Sugiyama [168, 169] and by Ishida and Yokota [102, 103]. The results from [168] already indicate that in the model (II.1.1) under consideration, blow-up *never* occurs, in accordance with Theorems II.2 and II.6.

The one-dimensional model on bounded spatial domains has been explicitly investigated by Osaki and Yagi [153] and Hillen and Potapov [97] leading to similar results as proved in Chapter II.3.

The fully parabolic model (II.1.1) with a genuinely nonlinear response function ϕ has not been rigorously analysed so far, with the following exception: in her thesis [157], Post proves existence and uniqueness of solutions to a similar system with linear diffusion and vanishing confinement on a bounded domain by nonvariational methods and obtains convergence to the (spatially homogeneous) stationary solution from compactness arguments. Variants of the classical parabolic-parabolic or parabolic-elliptic Keller-Segel models with a nonlinear chemotactic sensitivity coefficient have also been studied e.g. in [149, 181].

Despite the fact that entropy methods are one of the key tools for the analysis of Keller-Segel-type systems, the use of genuine variational methods is relatively recent in that context.

For the parabolic-elliptic Keller-Segel model, the variational framework was established by Blanchet, Calvez and Carrillo [23], who represented the evolution as a gradient flow of an appropriate potential with respect to the Wasserstein distance and constructed a numerical scheme on these grounds. Later, the gradient flow structure has been used for a detailed analysis of the basin of attraction in the critical mass case by Blanchet, Carlen and Carrillo [48] (see also e.g. [26, 38, 126]).

The parabolic-parabolic Keller-Segel model was harder to fit into the framework, since the two equations are formal gradient flows with respect to *different* metrics: $\mathbf{W}_2$ and L^2. The first rigorous analytical result on grounds of this structure was given by Blanchet and Laurençot in [28] (see also [143]), where they constructed weak solutions for the system with critical exponents of nonlinear diffusion. In the recent work [25] by Blanchet *et al.*, a similar strategy was used to re-prove the result in [39] about the global existence of weak solutions to the classical Keller-Segel system in two spatial dimensions.

II.1.2. Poisson-Nernst-Planck models

The last chapter of Part II is based on the author's article [188] and concerned with the long-time behaviour of a certain Possion-Nernst-Planck-type system, namely

$$\begin{aligned} \partial_t u(t,x) &= \operatorname{div}(u(t,x)\mathrm{D}[2u(t,x) + U(x) + \varepsilon\psi(t,x)]), \\ \partial_t v(t,x) &= \operatorname{div}(v(t,x)\mathrm{D}[2v(t,x) + V(x) - \varepsilon\psi(t,x)]), \end{aligned} \tag{II.1.7}$$

coupled by *Poisson's equation*

$$-\Delta\psi = u - v. \tag{II.1.8}$$

We consider system (II.1.7) on the whole space $\mathbb{R}^3$, so

$$\psi = \mathbf{G} * (u - v),$$

with *Newton's potential*

$$\mathbf{G}(x) = \frac{1}{4\pi|x|} \quad \text{for } x \neq 0.$$

We furthermore assume that $\varepsilon > 0$ is a fixed parameter and the confinement potentials $U, V \in C^2(\mathbb{R}^3)$ are subject to the conditions of W in Assumption II.3 (uniform convexity and bounded second derivative). Similarly to the Keller-Segel-type systems considered in Chapters II.2 and II.3, system (II.1.7) possesses a formal gradient flow structure: here, the underlying metric space is a product of two Wasserstein spaces, that is

$$\mathbf{X} = \mathscr{P}_2(\mathbb{R}^3) \times \mathscr{P}_2(\mathbb{R}^3) \text{ endowed with the distance } \mathbf{d}((u,v),(\widetilde{u},\widetilde{v})) := \sqrt{\mathbf{W}_2(u,\widetilde{u})^2 + \mathbf{W}_2(v,\widetilde{v})^2},$$

and the free energy $\mathcal{E} : \mathbf{X} \to \mathbb{R}_\infty$ is given by

$$\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^3} (u^2 + v^2 + uU + vV + \frac{\varepsilon}{2}|\mathrm{D}\psi|^2)\,\mathrm{d}x & \text{if } (u,v) \in L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3), \\ +\infty & \textit{otherwise}. \end{cases}$$

System (II.1.7) may arise as a model for the dynamics of a system consisting of positively and negatively charged particles (e.g. ions) inside some electrically neutral surrounding medium (e.g. air, water). For further details on the mathematical modelling of those phenomena, we refer to the monographs [131, 106]. Here, both species are confined by means of external potentials U and V and are assumed to diffuse nonlinearly with a diffusive mobility depending linearly on the concentrations u and v, respectively. We assume the Poisson coupling via equation (II.1.8) to be suitably weak ($\varepsilon \ll 1$), i.e. the drift induced by the electromagnetic force to be small. The quantity $\varepsilon^{-1} \gg 1$ corresponds to a large *relative permittivity (dielectric constant)* of the surrounding medium.

In Chapter II.4, we employ a similar strategy as for the Keller-Segel-type systems (Chapters II.2&II.3) to study the long-time behaviour of solutions. First, we characterize the set of equilibria:

Theorem II.9 (Existence and uniqueness of stationary states). *For every $\varepsilon > 0$, there exists a unique minimizer $(u_\infty, v_\infty) \in (W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3))$ of $\mathcal{E}$ on $\mathbf{X}$. (u_∞, v_∞) is a stationary solution to* (II.1.7) *and satisfies*

$$u_\infty = \frac{1}{2}[C_u - U - \varepsilon\psi_\infty]_+, \tag{II.1.9}$$

$$v_\infty = \frac{1}{2}[C_v - V + \varepsilon\psi_\infty]_+, \tag{II.1.10}$$

$$\psi_\infty := \mathbf{G} * (u_\infty - v_\infty),$$

where $C_u, C_v \in \mathbb{R}$ are such that $\|u_\infty\|_{L^1} = 1 = \|v_\infty\|_{L^1}$; $[\cdot]_+$ denoting the positive part. For every $\alpha \in (0,1)$, $u_\infty, v_\infty \in C^{0,\alpha}(\mathbb{R}^3)$ with compact support and $\psi \in L^\infty(\mathbb{R}^3) \cap C^{2,\alpha}(\mathbb{R}^3)$.

Second, we prove for sufficiently small coupling strength $\varepsilon > 0$ the exponential convergence to (u_∞, v_∞):

Theorem II.10 (Exponential convergence to equilibrium). *There are constants $\overline{\varepsilon} > 0$ and $\overline{L} > 0$ such that for all $\delta > 0$, there exists $C_\delta > 0$ such that the following is true for every $\varepsilon \in (0, \overline{\varepsilon})$ and arbitrary initial conditions $(u_0, v_0) \in \mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$: A weak solution (u,v) to* (II.1.7)*, obtained as a limit of the scheme* (I.2.11) *(see Theorem II.46 below) converges to (u_∞, v_∞) exponentially fast with rate $\Lambda_\varepsilon := \lambda_0 - \overline{L}\varepsilon > 0$ in the following* *sense:*

$$\begin{aligned} &\mathbf{W}_2(u(t,\cdot),u_\infty) + \mathbf{W}_2(v(t,\cdot),v_\infty) + \|u(t,\cdot) - u_\infty\|_{L^2} + \|v(t,\cdot) - v_\infty\|_{L^2} \\ &\qquad \le C_\delta \left(\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + 1\right)^{1+\delta} e^{-\Lambda_\varepsilon t} \qquad \textit{for all } t \ge 0. \end{aligned} \tag{II.1.11}$$

A similar system has been considered by Biler, Dolbeault and Markowich [20]. There, a *time-dependent* coupling $\varepsilon(t)$ was introduced, with the crucial assumption that $\varepsilon(t) \to 0$ as $t \to \infty$, i.e. asymptotical damping of the electrostatic potential. Under relatively general requirements on spatial dimension, external potential and diffusive nonlinearity, convergence to equilibrium as $t \to \infty$ is proved for sufficiently regular solutions. Here, we do *not* require asymptotical damping of the Poisson coupling, that is, the system at hand still constitutes a *coupled* system even in the large-time limit $t \to \infty$. To the best of our knowledge, our rigorous result on exponential convergence of *weak* solutions is novel in the case of genuinely nonlinear diffusion on multiple space dimensions, even in the small coupling regime $\varepsilon \ll 1$. Partial results have been obtained in one spatial dimension [70] or for space-dependent diffusion [10] only.

In contrast to that, the case of linear diffusion has already been treated almost exhaustively. In the articles [5, 19, 6] preceding [20], it was shown that the rate of exponential convergence to equilibrium of the system without coupling, for uniformly convex potentials, is (almost) retained for coupled systems. There, the strategy of proof is mainly based on applications of *generalized Sobolev inequalities* the derivation of which require the use of a *Holley-Stroock*-type perturbation lemma [98]. Seemingly, such a strategy might not be applicable in the setting of nonlinear diffusion. On the other hand, systems of the form above possess (at least formally) a *gradient flow* structure (w.r.t. e.g. the L^2-Wasserstein distance) which also is of use for the analysis of the system — and, in contrast, does *not* at all require linear diffusion.

CHAPTER II.2

Systems of Keller-Segel type with porous medium diffusion

II.2.1. Existence of weak solutions

This section is devoted to the proof of Theorem II.2 (cf. [185, 189]). Our strategy will be as follows: At first, we prove some properties of the energy functional $\mathcal{E}$ which usually are at the basis of variational methods: boundedness from below, coercivity and lower semicontinuity in an appropriate sense. With these properties at hand, we are able to construct a discrete solution via the minimizing movement scheme. Together with classical estimates directly derived from the minimizing movement scheme, we set up a discrete weak formulation satisfied by the piecewise constant (in time) discrete solution. Due to the nonlinearity of the problem, additional compactness estimates are necessary for the passage to the limit of vanishing step size. Here, those will be derived by energy-dissipation methods: perturbation of the subsequent minimizers from the minimizing movement scheme along the heat flow yields a higher order of spatial regularity of the discrete solution. To conclude the proof, we pass to the continuous-time limit in a strong sense and verify the (time-continuous) weak formulation of (II.1.1) for the limit curve.

II.2.1.1. Properties of the energy functional $\mathcal{E}$

Recall our definition for the driving energy $\mathcal{E}$: For every $(u,v) \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d)$, we set

$$\mathcal{E}(u,v) := \int_{\mathbb{R}^d} \left(\frac{u^m}{m-1} + uW + \frac{1}{2}\,|\mathrm{D}v|^2 - \chi uv \right) \mathrm{d}x,$$

and $\mathcal{E}(u,v) = +\infty$ otherwise. We first prove several elementary properties of $\mathcal{E}$:

Proposition II.11 (Domain of $\mathcal{E}$ and boundedness from below). *If $(u,v) \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d)$, then $\mathcal{E}(u,v) < \infty$. Moreover,*

$$\mathcal{E}(u,v) \geq C_m \|u\|_{L^m}^m + \inf W - C_k \|u\|_{L^m}^{2(1-\theta)}, \tag{II.2.1}$$

$$\|\mathrm{D}v\|_{L^2}^2 \leq C_0 \left(\mathcal{E}(u,v) + |\inf W| + \|u\|_{L^m}^{2(1-\theta)} \right), \tag{II.2.2}$$

where C_m, C_k and C_0 are positive constants and $\theta \in (0,1)$ is such that $2(1-\theta) < m$. In particular, $\mathcal{E}$ is bounded from below.

Proof. Since $W(x) \leq A|x|^2 + B$ for some $A, B \geq 0$ and the second moment $\mathbf{m}_2(u)$ is finite, one obviously has

$$\int_{\mathbb{R}^d} uW \,\mathrm{d}x \leq A\,\mathbf{m}_2(u) + B < \infty.$$

For the coupling term, we observe that $m > 2 - \frac{2}{d} \geq \frac{2d}{d+2}$ and by the Hölder, Gagliardo-Nirenberg-Sobolev and the L^p-interpolation (cf. [80, Thm. B.2.h]) inequalities that

$$\|uv\|_{L^1} \leq \|u\|_{L^{\frac{2d}{d+2}}} \|v\|_{L^{\frac{2d}{d-2}}} \leq C \|\mathrm{D}v\|_{L^2} \|u\|_{L^1}^{\theta} \|u\|_{L^m}^{1-\theta} < \infty, \tag{II.2.3}$$

where $\theta := \frac{\frac{d+2}{2d} - \frac{1}{m}}{1-\frac{1}{m}} \in (0,1)$. Hence, $\mathcal{E}$ is well-defined and finite. The estimate (II.2.1) follows by a similar estimate using Young's inequality in addition to (II.2.3):

$$\mathcal{E}(u,v) \geq \frac{\|u\|_{L^m}^m}{m-1} + \inf W + \frac{1}{2}\|\mathrm{D}v\|_{L^2}^2 - |\chi|\|uv\|_{L^1} \geq C_m\|u\|_{L^m}^m + \inf W - C_k\|u\|_{L^m}^{2(1-\theta)}.$$

From that, boundedness from below of $\mathcal{E}$ follows because $m > 2 - \frac{2}{d}$ implies $m - 2(1-\theta) > 0$ and the map $[0,\infty) \ni y \mapsto C_l y^l - C_k y^k$ is bounded from below for $C_l > 0$ and $l > k > 0$. We obtain (II.2.2) by using the Hölder, Gagliardo-Nirenberg-Sobolev and Young inequalities:

$$\begin{aligned}\|\mathrm{D}v\|_{L^2}^2 &\leq 2(\mathcal{E}(u,v) + |\inf W| + |\chi|\|uv\|_{L^1}) \\ &\leq 2(\mathcal{E}(u,v) + |\inf W| + \frac{(C|\chi|)^2}{2}\|u\|_{L^m}^{2(1-\theta)} + \frac{1}{2}\|\mathrm{D}v\|_{L^2}^2).\end{aligned}$$

Obviously, the second estimate (II.2.2) follows. □

In preparation for the application of variational principles, we show the following lower semicontinuity property of $\mathcal{E}$:

Proposition II.12 (Weak lower semicontinuity of $\mathcal{E}$). *Let a sequence $(u_n, v_n)_{n\in\mathbb{N}}$ in $\mathscr{P}_2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ with uniformly bounded second moments $\mathfrak{m}_2(u_n)$ be given. Assume furthermore that $(u_n)_{n\in\mathbb{N}}$ converges to $u \in \mathscr{P}_2(\mathbb{R}^d)$ narrowly in the space of probability measures $\mathscr{P}(\mathbb{R}^d)$ and that $(v_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\mathbb{R}^d)$ to $v \in L^2(\mathbb{R}^d)$ as $n \to \infty$. Then,*

$$\mathcal{E}(u,v) \leq \liminf_{n\to\infty} \mathcal{E}(u_n, v_n).$$

Proof. If $\liminf_{n\to\infty} \mathcal{E}(u_n,v_n) = +\infty$, there is nothing to prove. For $\liminf_{n\to\infty} \mathcal{E}(u_n,v_n) < +\infty$, we restrict $(u_n,v_n)_{n\in\mathbb{N}}$ without loss of generality to the subsequence (non-relabelled) converging to $\liminf_{n\to\infty} \mathcal{E}(u_n,v_n)$, so that $\mathcal{E}(u_n,v_n)$ is uniformly bounded by a positive constant C. Let $C_1 > 0$ with $\mathfrak{m}_2(u_n) \leq C_1$ for all $n \in \mathbb{N}$. Elementary properties of weak convergence ensure the existence of a positive constant C_2 such that $\|v_n\|_{L^2} \leq C_2$. To simplify the notation, define $g : [0,\infty) \to \mathbb{R}$,

$$g(y) := C_m y^m - C_k y^{2(1-\theta)}, \tag{II.2.4}$$

which is bounded from below thanks to $m > 2(1-\theta)$. We now claim that $(u_n)_{n\in\mathbb{N}}$ is bounded in $L^m(\mathbb{R}^d)$. Suppose not, then there exists a subsequence $(u_{n_l})_{l\in\mathbb{N}}$ such that $\|u_{n_l}\|_{L^m} \to \infty$ as $l \to \infty$. It is immediate that $g(\|u_{n_l}\|_{L^m}) \to \infty$, which contradicts $g(\|u_n\|_{L^m}) \leq C - \inf W$ obtained from the uniform bound on $\mathcal{E}$. So there exists $C_3 > 0$ such that $\|u_n\|_{L^m} < C_3$ for all $n \in \mathbb{N}$. By the Banach-Alaoglu theorem [75, Theorem V.4.7.], we are able to extract a subsequence such that u_n converges weakly to u in $L^m(\mathbb{R}^d)$.

Using (II.2.2), we obtain the uniform estimate

$$\|\mathrm{D}v_n\|_{L^2} \leq C_0 \left(C + |\inf W| + C_3^{2(1-\theta)}\right)^{1/2} =: C_4.$$

By the above estimate and the bound on $\|v_n\|_{L^2}$, $(v_n)_{n\in\mathbb{N}}$ is bounded in $W^{1,2}(\mathbb{R}^d)$. Again, by the Banach-Alaoglu theorem, there exists a subsequence of $(v_n)_{n\in\mathbb{N}}$ such that v_n converges weakly to v in $W^{1,2}(\mathbb{R}^d)$ as $n \to \infty$. Due to lower semicontinuity of norms and (at most) quadratic growth of W (see Theorem I.9), it only remains to consider the coupling term in $\mathcal{E}$.

We prove $\lim_{n\to\infty} \int_{\mathbb{R}^d} (u_n v_n - uv)\,\mathrm{d}x = 0$ by a truncation argument as in [28]. Let $\beta_R \in C_c^\infty(\mathbb{R}^d)$ for $R > 0$ with

$$\begin{aligned}0 \leq \beta_R &\leq 1, \\ \beta_R(x) = 1 &\qquad \forall x \in \mathbb{B}_R(0), \\ \beta_R(x) = 0 &\qquad \forall x \notin \mathbb{B}_{2R}(0),\end{aligned}$$

and choose $p \in (1,m) \cap (\frac{2d}{d+2}, 2) \neq \emptyset$. Since u_n is uniformly bounded in $L^1(\mathbb{R}^d)$ and converges weakly to u in $L^m(\mathbb{R}^d)$, u_n converges weakly to u in $L^p(\mathbb{R}^d)$, possibly by extracting a subsequence. Note that $\frac{p}{p-1} \in (2, \frac{2d}{d-2})$.

For fixed $R > 0$, the Rellich-Kondrachov compactness theorem [80, Thm. 5.7.1] yields for all $q \in [1, \frac{2d}{d-2})$ the strong convergence of $\beta_R v_n$ to $\beta_R v$ in $L^q(\mathbb{R}^d)$ on a subsequence, in particular for $q = \frac{p}{p-1}$. It follows that

$$\left|\int_{\mathbb{R}^d} \beta_R v(u_n - u)\,\mathrm{d}x\right| + \left|\int_{\mathbb{R}^d} \beta_R u_n(v_n - v)\,\mathrm{d}x\right| \overset{n\to\infty}{\longrightarrow} 0.$$

An analogous application of the Hölder and Gagliardo-Nirenberg-Sobolev inequalities as in (II.2.3) leads to the estimate

$$\begin{aligned}\left|\int_{\mathbb{R}^d} (1-\beta_R) u_n v_n\,\mathrm{d}x\right| &\leq C\|\mathrm{D}v_n\|_{L^2}\|u_n\|_{L^m}^{1-\theta}\left(\int_{\mathbb{R}^d\setminus\mathbb{B}_R(0)} u_n\,\mathrm{d}x\right)^{\theta} \\ &\leq C\|\mathrm{D}v_n\|_{L^2}\|u_n\|_{L^m}^{1-\theta}\left(\int_{\mathbb{R}^d\setminus\mathbb{B}_R(0)} \frac{|x|^2}{R^2} u_n(x)\,\mathrm{d}x\right)^{\theta} \\ &\leq R^{-2\theta} C C_4 C_3^{1-\theta} C_1^{\theta},\end{aligned}$$

in combination with the uniform bounds on $\mathfrak{m}_2(u_n)$, $\|u_n\|_{L^m}$ and $\|\mathrm{D}v_n\|_{L^2}$. By the same argument and weak lower semicontinuity of norms,

$$\left|\int_{\mathbb{R}^d} (1-\beta_R) uv\,\mathrm{d}x\right| \leq R^{-2\theta} C C_4 C_3^{1-\theta} C_1^{\theta}.$$

We apply the triangular inequality:

$$\begin{aligned}\left|\int_{\mathbb{R}^d} (u_n v_n - uv)\,\mathrm{d}x\right| \leq &\left|\int_{\mathbb{R}^d} \beta_R v(u_n - u)\,\mathrm{d}x\right| + \left|\int_{\mathbb{R}^d} \beta_R u_n(v_n - v)\,\mathrm{d}x\right| \\ &+ \left|\int_{\mathbb{R}^d} (1-\beta_R) u_n v_n\,\mathrm{d}x\right| + \left|\int_{\mathbb{R}^d} (1-\beta_R) uv\,\mathrm{d}x\right|.\end{aligned}$$

This leads us to

$$0 \leq \limsup_{n\to\infty} \left|\int_{\mathbb{R}^d} (u_n v_n - uv)\,\mathrm{d}x\right| \leq 2R^{-2\theta} C C_4 C_3^{1-\theta} C_1^{\theta},$$

for all $R > 0$. Passing to the limit $R \to \infty$ yields the claim. □

Remark II.13 (Non-convexity of $\mathcal{E}$). *Unfortunately, for $\chi \neq 0$, the functional $\mathcal{E}$ is not λ-convex along geodesics (for any $\lambda \in \mathbb{R}$) with respect to the distance $\mathbf{d}$ because of the coupling term as the following formal calculation indicates.*

Denote by $\mathcal{E}_c$ the coupling functional

$$\mathcal{E}_c(u,v) := \int_{\mathbb{R}^d} \chi uv\,\mathrm{d}x$$

and let $\lambda \leq 0$. We show that there exist $u \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d)$, $v, w \in W^{1,2}(\mathbb{R}^d)$ and a Borel-measurable function $t : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} \mathcal{E}_c([(1-s)\,\mathrm{id} + st]_{\#}u, (1-s)v + sw) - \lambda\,\mathbf{d}^2((u,v),(t_{\#}u, w)) < 0.$$

Using the density transformation theorem, one obtains

$$\begin{aligned}&\mathcal{E}_c([(1-s)\,\mathrm{id} + st]_{\#}u, (1-s)v + sw) \\ &= \int_{\mathbb{R}^d} \chi u(y)\left[(1-s)v((1-s)y + st(y)) + sw((1-s)y + st(y))\right]\mathrm{d}y.\end{aligned}$$

A straightforward calculation yields

$$\begin{aligned}
&\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0} \mathcal{E}_c([(1-s)\,\mathrm{id} + st]_\# u, (1-s)v + sw) - \lambda\, \mathbf{d}^2((u,v),(t_\# u, w)) \\
&\leq \int_{\mathbb{R}^d} u(y)(t(y)-y)\cdot(\chi \mathrm{D}^2 v(y) - \lambda \mathbb{1})(t(y)-y)\,\mathrm{d}y \\
&\quad + \int_{\mathbb{R}^d} 2\chi u(y)(\mathrm{D}(w-v))(y)\cdot(t(y)-y)\,\mathrm{d}y - \int_{\mathbb{R}^d} \lambda (v(y)-w(y))^2\,\mathrm{d}y.
\end{aligned}$$

Choose $t(y) := 2y$ *and* $u := \frac{\mathbf{1}_{\mathbb{B}_1(0)}}{|\mathbb{B}_1(0)|}$. *Furthermore, let* $v, w \in C_c^\infty(\mathbb{R}^d)$ *such that* $v = w$ *on* $\mathbb{R}^d$ *and* $v(y) = \frac{\lambda-1}{2\chi}|y|^2$ *for* $y \in \mathbb{B}_1(0)$. *Then, the desired result follows:*

$$\begin{aligned}
&\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0} \mathcal{E}_c([(1-s)\,\mathrm{id} + st]_\# u, (1-s)v + sw) - \lambda\, \mathbf{d}^2((u,v),(t_\# u, w)) \\
&\leq \int_{\mathbb{B}_1(0)} \frac{1}{|\mathbb{B}_1(0)|}\left(\chi\frac{\lambda-1}{\chi} - \lambda\right)|y|^2\,\mathrm{d}y = -\frac{\sigma(\partial\mathbb{B}_1(0))}{|\mathbb{B}_1(0)|(d+2)} < 0,
\end{aligned}$$

where σ *denotes the (hyper-)surface measure in* $\mathbb{R}^d$.

II.2.1.2. Time discretization

Recall the discretization scheme from (I.2.11). We introduce the step size $\tau > 0$ and define the associated Yosida penalization $\mathcal{E}_\tau$ of the energy by

$$\mathcal{E}_\tau(u, v \,|\, \widetilde{u}, \widetilde{v}) := \frac{1}{2\tau}\mathbf{d}^2((u,v),(\widetilde{u},\widetilde{v})) + \mathcal{E}(u,v) \tag{II.2.5}$$

for all $(u,v),(\widetilde{u},\widetilde{v}) \in \mathbf{X}$. Set $(u_\tau^0, v_\tau^0) := (u_0, v_0)$ and define the sequence $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$ inductively by choosing

$$(u_\tau^n, v_\tau^n) \in \underset{(u,v)\in\mathbf{X}}{\operatorname{argmin}}\ \mathcal{E}_\tau(u, v \,|\, u_\tau^{n-1}, v_\tau^{n-1}).$$

Proposition II.14 (Well-posedness of minimizing movement scheme). *For every* $(\widetilde{u},\widetilde{v}) \in \mathbf{X}$, *there exists at least one minimizer* $(u,v) \in \mathbf{X}$ *of* $\mathcal{E}_\tau(\cdot \,|\, \widetilde{u},\widetilde{v})$ *that satisfies* $u \in L^m(\mathbb{R}^d)$ *and* $v \in W^{1,2}(\mathbb{R}^d)$.

Proof. The proof is an application of the direct method from the calculus of variations to the functional $\mathcal{E}_\tau(\cdot \,|\, \widetilde{u},\widetilde{v})$.

First, observe thanks to Proposition II.11 that on any given sublevel S of $\mathcal{E}_\tau(\cdot \,|\, \widetilde{u},\widetilde{v})$, both $\mathbf{W}_2(u,\widetilde{u})$ and $\|v-\widetilde{v}\|_{L^2}$ are uniformly bounded. The first bound implies that also the second moment $\mathfrak{m}_2(u)$ is uniformly bounded, and thus the u-components in S belong to a subset of $\mathscr{P}_2(\mathbb{R}^d)$ that is relatively compact in the narrow topology by Prokhorov's theorem (see Theorem I.7). The other bound implies via Alaoglu's theorem that the v-components belong to a weakly relatively compact subset of $L^2(\mathbb{R}^d)$. Hence, the existence of a minimizer follows by lower semicontinuity of $\mathcal{E}$, cf. Proposition II.12. The additional regularity is a consequence of the fact that the proper domain of $\mathcal{E}$ is a subset of $L^m(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d)$. □

Given the sequence $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$, define the *discrete solution* $(u_\tau, v_\tau) : [0,\infty) \to \mathbf{X}$ as in (I.2.12) by piecewise constant interpolation:

$$(u_\tau, v_\tau)(t) := (u_\tau^n, v_\tau^n) \text{ for } t \in ((n-1)\tau, n\tau] \text{ and } n \geq 1.$$

We start by recalling a collection of estimates on (u_τ, v_τ) that follow immediately from the construction by minimizing movements.

Proposition II.15 (Classical estimates). *The following holds for* $T > 0$:

$$\mathcal{E}(u_\tau^n, v_\tau^n) \leq \mathcal{E}(u_\tau^{n-1}, v_\tau^{n-1}) \leq \mathcal{E}(u_0, v_0) < \infty \qquad \forall n \in \mathbb{N}, \tag{II.2.6}$$

$$\sum_{n=1}^{\infty} \mathbf{d}^2((u_\tau^n, v_\tau^n), (u_\tau^{n-1}, v_\tau^{n-1})) \leq 2\tau(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}), \tag{II.2.7}$$

$$\mathbf{d}((u_\tau(s), v_\tau(s)), (u_\tau(t), v_\tau(t))) \leq [2(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}) \max(\tau, |t-s|)]^{1/2} \quad \textit{for all } s, t \in [0, T], \tag{II.2.8}$$

the infimum $\inf \mathcal{E}$ *of* $\mathcal{E}$ *on* $\mathbf{X}$ *being finite.*

Proof. We include the proof for the sake of completeness. It can also be found e.g. in [4, Ch. 3]. By the minimizing property, one has

$$\mathcal{E}(u_\tau^n, v_\tau^n) + \frac{1}{2\tau}\mathbf{d}^2((u_\tau^n, v_\tau^n), (u_\tau^{n-1}, v_\tau^{n-1})) \leq \mathcal{E}(u_\tau^{n-1}, v_\tau^{n-1}).$$

Clearly, (II.2.6) and after summation over $n \geq 1$, (II.2.7) follows. For all $0 \leq s \leq t \leq T$, there exist $m \leq n$ such that

$$\mathbf{d}((u_\tau(s), v_\tau(s)), (u_\tau(t), v_\tau(t))) \leq \sum_{k=m}^{n-1} \mathbf{d}((u_\tau^k, v_\tau^k), (u_\tau^{k+1}, v_\tau^{k+1})).$$

Using Hölder's inequality yields the desired result:

$$\begin{aligned}\sum_{k=m}^{n-1} \mathbf{d}((u_\tau^k, v_\tau^k), (u_\tau^{k+1}, v_\tau^{k+1})) &\leq \left(\sum_{k=m}^{n-1} \tau\right)^{1/2} \left(\sum_{k=m}^{n-1} \frac{\mathbf{d}^2((u_\tau^k, v_\tau^k), (u_\tau^{k+1}, v_\tau^{k+1}))}{\tau}\right)^{1/2} \\ &\leq (2\tau(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})(n-m))^{1/2} \\ &\leq [2(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}) \max(\tau, |t-s|)]^{1/2}.\end{aligned}$$

□

II.2.1.2.1. *Additional regularity*

Due to the nonlinearity of the PDE system (II.1.1), further compactness arguments are needed to enable passage to the continuous-time limit $\tau \searrow 0$. Therefore, regularity estimates on the discrete solution are proved. Here, we follow the method by Blanchet and Laurençot in [28, 22].

Proposition II.16 (Further regularity of the minimizers). *Let* $(\widetilde{u}, \widetilde{v}) \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d) \times W^{1,2}(\mathbb{R}^d)$ *and let* (u, v) *be a minimizer of* $\mathcal{E}_\tau(\cdot \mid \widetilde{u}, \widetilde{v})$*. Then,* $u \in L^2(\mathbb{R}^d)$*,* $u^{m/2} \in W^{1,2}(\mathbb{R}^d)$ *and* $v \in W^{2,2}(\mathbb{R}^d)$ *as:*

$$\begin{aligned}&\|\mathrm{D}u^{m/2}\|_{L^2}^2 + \|\Delta v\|_{L^2}^2 + \|u\|_{L^2}^2 \\ &\leq C_0 \left(\frac{1}{\tau}\left(\mathcal{H}(\widetilde{u}) - \mathcal{H}(u) + \mathcal{F}(\widetilde{v}) - \mathcal{F}(v)\right) + \|u\|_{L^m}^{(1-A)\frac{m}{m-1}} + \|\Delta W\|_{L^\infty}\right),\end{aligned} \tag{II.2.9}$$

where $A := \frac{m-\left(2-\frac{2}{d}\right)}{m-1}$*,* $C_0 > 0$ *and denoting* $\mathcal{H}(u) := \int_{\mathbb{R}^d} u \log u \, \mathrm{d}x$ *and* $\mathcal{F}(v) := \frac{1}{2}\|\mathrm{D}v\|_{L^2}^2$*.*

Proof. The idea of proof used here is based on the flow interchange lemma (Theorem I.5) and is to calculate the dissipation of $\mathcal{E}$ along the gradient flow of an auxiliary functional, namely the diffusion flow. Therefore, we recall from Example I.15 that *Boltzmann's entropy* $\mathcal{H}$ is 0-convex along generalized geodesics in $\mathscr{P}_2(\mathbb{R}^d)$ and its gradient flow $\mathsf{S}^{\mathcal{H}}$ is the diffusion flow satisfying

$$\partial_s \mathsf{S}_s^{\mathcal{H}}(u) = \Delta \mathsf{S}_s^{\mathcal{H}}(u).$$

Moreover, with the evolution variational estimate (I.2.4), we deduce by integration over time using that $\mathcal{H}$ is a Lyapunov functional along $\mathsf{S}^{\mathcal{H}}$:

$$\frac{1}{2}\left(\mathbf{W}_2^2(\mathsf{S}_s^{\mathcal{H}}(u), \widetilde{u}) - \mathbf{W}_2^2(u, \widetilde{u})\right) \leq \int_0^s (\mathcal{H}(\widetilde{u}) - \mathcal{H}(\mathsf{S}_\sigma^{\mathcal{H}}(u)))\, \mathrm{d}\sigma \leq s[\mathcal{H}(\widetilde{u}) - \mathcal{H}(\mathsf{S}_s^{\mathcal{H}}(u))]. \tag{II.2.10}$$

Analogous to that, the *Dirichlet energy* $\mathcal{F}$ is geodesically 0-convex on $L^2(\mathbb{R}^d)$ and its gradient flow $\mathsf{S}^{\mathcal{F}}$ is also associated to the diffusion equation (see Example I.4).

The application of the evolution variational estimate (I.2.4) then shows

$$\frac{1}{2}\left(\|\mathsf{S}^{\mathcal{F}}_s(v)-\widetilde{v}\|_{L^2}^2-\|v-\widetilde{v}\|_{L^2}^2\right)\le\int_0^s\left(\mathcal{F}(\widetilde{v})-\mathcal{F}(\mathsf{S}^{\mathcal{F}}_\sigma(v))\right)\mathrm{d}\sigma\le s[\mathcal{F}(\widetilde{v})-\mathcal{F}(\mathsf{S}^{\mathcal{F}}_s(v))]. \tag{II.2.11}$$

Well-known results of parabolic theory ensure that $(\mathsf{S}^{\mathcal{H}}_s(u),\mathsf{S}^{\mathcal{F}}_s(v))\in\mathbf{X}\cap(L^m(\mathbb{R}^d)\times W^{1,2}(\mathbb{R}^d))$ if $(u,v)\in\mathbf{X}\cap(L^m(\mathbb{R}^d)\times W^{1,2}(\mathbb{R}^d))$. For the sake of clarity, we introduce the notation $(\mathcal{U}_s,\mathcal{V}_s):=(\mathsf{S}^{\mathcal{H}}_s(u),\mathsf{S}^{\mathcal{F}}_s(v))$ and calculate for $s>0$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U}_s,\mathcal{V}_s)&=\int_{\mathbb{R}^d}\left(\left[\frac{m}{m-1}\mathcal{U}_s^{m-1}+W-\chi\mathcal{V}_s\right]\Delta\mathcal{U}_s+\left[-\Delta\mathcal{V}_s-\chi\mathcal{U}_s\right]\Delta\mathcal{V}_s\right)\mathrm{d}x\\&=\int_{\mathbb{R}^d}\left(-\frac{4}{m}\left|\mathrm{D}\mathcal{U}_s^{m/2}\right|^2+\mathcal{U}_s\Delta W-(\Delta\mathcal{V}_s+\chi\mathcal{U}_s)^2+\chi^2\mathcal{U}_s^2\right)\mathrm{d}x,\end{aligned} \tag{II.2.12}$$

where the last step follows by integration by parts and elementary calculations. We first provide an auxiliary estimate on the L^2 norm of u. By the Hölder and L^p-interpolation inequality, we get

$$\chi^2\|\mathcal{U}_s\|_{L^2}^2\le\chi^2\|\mathcal{U}_s\|_{L^{\frac{dm}{(m-1)d+2}}}\|\mathcal{U}_s\|_{L^{\frac{dm}{d-2}}}\le\chi^2\|\mathcal{U}_s\|_{L^1}^A\|\mathcal{U}_s\|_{L^m}^{(1-A)}\|\mathcal{U}_s^{m/2}\|_{L^{\frac{2d}{d-2}}}^{2/m},$$

where A is defined as above; note that $\frac{dm}{(m-1)d+2}\in(1,m)$. Using the Gagliardo-Nirenberg-Sobolev inequality together with Young's inequality, one obtains (recalling $\|\mathcal{U}_s\|_{L^1}=1$)

$$\chi^2\|\mathcal{U}_s\|_{L^2}^2\le\frac{2}{m}\|\mathrm{D}\mathcal{U}_s^{m/2}\|_{L^2}^2+K_0\|\mathcal{U}_s\|_{L^m}^{(1-A)\frac{m}{m-1}}, \tag{II.2.13}$$

with a suitable constant $K_0>0$. Exploiting the monotonicity of the L^m norm along $\mathsf{S}^{\mathcal{H}}$ and using $\Delta W\in L^\infty(\mathbb{R}^d)$, we infer from (II.2.12)&(II.2.13) that

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U}_s,\mathcal{V}_s)\le-\frac{2}{m}\left\|\mathrm{D}\mathcal{U}_s^{m/2}\right\|_{L^2}^2-\|\Delta\mathcal{V}_s+\chi\mathcal{U}_s\|_{L^2}^2+K_0\|u\|_{L^m}^{(1-A)\frac{m}{m-1}}+\|\Delta W\|_{L^\infty}. \tag{II.2.14}$$

As a concluding step, the minimizing property

$$0\le\mathcal{E}_\tau(\mathcal{U}_s,\mathcal{V}_s\,|\,\widetilde{u},\widetilde{v})-\mathcal{E}_\tau(u,v\,|\,\widetilde{u},\widetilde{v})$$

yields together with (II.2.10)&(II.2.11) and (II.2.14):

$$\begin{aligned}&\frac{1}{s}\int_0^s\left(\left\|\mathrm{D}\mathcal{U}_s^{m/2}\right\|_{L^2}^2+\|\Delta\mathcal{V}_s+\chi\mathcal{U}_s\|_{L^2}^2\right)\mathrm{d}\sigma\\&\le K_1\left(\|u\|_{L^m}^{(1-A)\frac{m}{m-1}}+\|\Delta W\|_{L^\infty}+\frac{1}{\tau}(\mathcal{H}(\widetilde{u})-\mathcal{H}(\mathcal{U}_s)+\mathcal{F}(\widetilde{v})-\mathcal{F}(\mathcal{V}_s))\right),\end{aligned}$$

where $K_1>0$ is a constant. Passage to the $\liminf$ as $s\searrow 0$ yields

$$\|\mathrm{D}u^{m/2}\|_{L^2}^2+\|\Delta v+\chi u\|_{L^2}^2\le K_1\left(\frac{1}{\tau}\left(\mathcal{H}(\widetilde{u})-\mathcal{H}(u)+\mathcal{F}(\widetilde{v})-\mathcal{F}(v)\right)+\|u\|_{L^m}^{(1-A)\frac{m}{m-1}}+\|\Delta W\|_{L^\infty}\right),$$

using lower semicontinuity of norms and continuity of the entropies $\mathcal{H}$ and $\mathcal{F}$ along their respective gradient flows. The final estimate (II.2.9) then follows from (II.2.13) applied for $s=0$, in combination with the triangular and Young inequalities. □

II.2.1.2.2. *Discrete weak formulation*

In this paragraph, an approximate weak formulation satisfied by the discrete solution (u_τ,v_τ) will be derived with the now classical *JKO method* [105]; see also e.g. [179, Sect. 8.4]. The main idea is as follows: To calculate the first variation of the functional $\mathcal{E}_\tau(\cdot\,|\,u_\tau^{n-1},v_\tau^{n-1})$, let $\eta,\gamma\in C_c^\infty(\mathbb{R}^d)$ and $\xi:=\mathrm{D}\eta$. Define by X the smooth flow associated with ξ, i.e. the flow of the ODE $\frac{\mathrm{d}}{\mathrm{d}s}X^s(x)=\xi(X^s(x))$ with initial condition $X^0(x) = x$ for $x \in \mathbb{R}^d$ and $s \geq 0$. There exists an optimal transport map $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ with $\varphi_\# u_\tau^{n-1} = u_\tau^n$, see for instance [179, Thm. 2.32]. A perturbation of the minimizer (u_τ^n, v_τ^n) of $\mathcal{E}_\tau(\cdot \mid u_\tau^{n-1}, v_\tau^{n-1})$ along X and the direction γ, respectively, yields the estimate

$$\begin{aligned} 0 &\leq \mathcal{E}_\tau(X^s{}_\# u_\tau^n, v_\tau^n + s\gamma \mid u_\tau^{n-1}, v_\tau^{n-1}) - \mathcal{E}_\tau(u_\tau^n, v_\tau^n \mid u_\tau^{n-1}, v_\tau^{n-1}) \\ &= \frac{1}{2\tau}\left(\mathbf{W}_2^2(X^s{}_\# u_\tau^n, u_\tau^{n-1}) - \mathbf{W}_2^2(u_\tau^n, u_\tau^{n-1})\right) + \frac{1}{2\tau}\left(\|v_\tau^n + s\gamma - v_\tau^{n-1}\|_{L^2}^2 - \|v_\tau^n - v_\tau^{n-1}\|_{L^2}^2\right) \\ &\quad + \frac{1}{m-1}\left(\|X^s{}_\# u_\tau^n\|_{L^m}^m - \|u_\tau^n\|_{L^m}^m\right) + \int_{\mathbb{R}^d} W\left[(X^s{}_\# u_\tau^n) - u_\tau^n\right]) \,\mathrm{d}x \\ &\quad + \frac{1}{2}\left(\|\mathrm{D}(v_\tau^n + s\gamma)\|_{L^2}^2 - \|\mathrm{D}v_\tau^n\|_{L^2}^2\right) - \chi \int_{\mathbb{R}^d} \left((X^s{}_\# u_\tau^n)(v_\tau^n + s\gamma) - u_\tau^n v_\tau^n\right) \,\mathrm{d}x. \end{aligned} \tag{II.2.15}$$

We consider the parts on the r.h.s. of (II.2.15) separately. By the definition of $\mathbf{W}_2$,

$$\begin{aligned} &\frac{1}{s}\left(\mathbf{W}_2^2(X^s{}_\# u_\tau^n, u_\tau^{n-1}) - \mathbf{W}_2^2(u_\tau^n, u_\tau^{n-1})\right) \leq \frac{1}{s}\int_{\mathbb{R}^d}\left(|(X^s \circ \varphi)(x) - x|^2 - |\varphi(x) - x|^2\right) u_\tau^{n-1}(x) \,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \frac{(X^s \circ \varphi)(x) - \varphi(x)}{s} \cdot \left((X^s \circ \varphi)(x) + \varphi(x) - 2x\right) u_\tau^{n-1}(x) \,\mathrm{d}x \\ &\xrightarrow{s \searrow 0} \int_{\mathbb{R}^d} 2\xi(\varphi(x)) \cdot (\varphi(x) - x) u_\tau^{n-1}(x) \,\mathrm{d}x. \end{aligned}$$

Straightforward, one obtains:

$$\frac{1}{s}\left(\|v_\tau^n + s\gamma - v_\tau^{n-1}\|_{L^2}^2 - \|v_\tau^n - v_\tau^{n-1}\|_{L^2}^2\right) \xrightarrow{s \searrow 0} \int_{\mathbb{R}^d} 2(v_\tau^n - v_\tau^{n-1})\gamma \,\mathrm{d}x.$$

Let for $y \in \mathbb{R}^d$ and $s \geq 0$ be $Z_s(y) := \det \mathrm{D}X^s(y) > 0$ the volume distortion due to X. Note that $Z_0(y) = 1$ and $\frac{\mathrm{d}}{\mathrm{d}s}|_{s=0} Z_s(y) = \mathrm{div}\xi(y)$ (see [132, eq. (2.30)] for a proof). By the definition of the push-forward and the density transformation theorem, we get:

$$\frac{1}{s}\left(\|X^s{}_\# u_\tau^n\|_{L^m}^m - \|u_\tau^n\|_{L^m}^m\right) = \frac{1}{s}\int_{\mathbb{R}^d} \frac{u_\tau^n(x)^m}{Z_s(x)^{m-1}}\left(1 - Z_s(x)^{m-1}\right) \mathrm{d}x \xrightarrow{s \searrow 0} -(m-1)\int_{\mathbb{R}^d} u_\tau^n(x)^m \mathrm{div}\xi(x) \,\mathrm{d}x.$$

By straightforward calculation,

$$\begin{aligned} &\frac{1}{s}\int_{\mathbb{R}^d} W(x)\left((X^s{}_\# u_\tau^n)(x) - u_\tau^n(x)\right) \mathrm{d}x = \frac{1}{s}\int_{\mathbb{R}^d} \left(W(X^s(x)) - W(x)\right) u_\tau^n(x) \,\mathrm{d}x \xrightarrow{s \searrow 0} \int_{\mathbb{R}^d} \mathrm{D}W \cdot \xi u_\tau^n \,\mathrm{d}x, \\ &\frac{1}{s}\left(\|\mathrm{D}(v_\tau^n + s\gamma)\|_{L^2}^2 - \|\mathrm{D}v_\tau^n\|_{L^2}^2\right) \xrightarrow{s \searrow 0} 2\int_{\mathbb{R}^d} \mathrm{D}\gamma \cdot \mathrm{D}v_\tau^n \,\mathrm{d}x. \end{aligned}$$

Moreover, taking into account the L^2 regularity of u obtained from Proposition II.16, we get

$$\begin{aligned} &\frac{1}{s}\int_{\mathbb{R}^d} \left((X_\#^s u_\tau^n)(v_\tau^n + s\gamma) - u_\tau^n v_\tau^n\right) \mathrm{d}x = \frac{1}{s}\int_{\mathbb{R}^d} u_\tau^n(x)\left(v_\tau^n(X^s(x)) + s\gamma(X^s(x)) - v_\tau^n(x)\right) \mathrm{d}x \\ &\xrightarrow{s \searrow 0} \int_{\mathbb{R}^d} \left(\gamma + \mathrm{D}v_\tau^n \cdot \xi\right) u_\tau^n \,\mathrm{d}x. \end{aligned}$$

We therefore have

$$\begin{aligned} 0 &\leq \frac{1}{\tau}\int_{\mathbb{R}^d} \xi(\varphi(x)) \cdot (\varphi(x) - x) u_\tau^{n-1}(x) \,\mathrm{d}x + \frac{1}{\tau}\int_{\mathbb{R}^d} (v_\tau^n - v_\tau^{n-1})\gamma \,\mathrm{d}x - \int_{\mathbb{R}^d} u_\tau^n(x)^m \mathrm{div}\xi(x) \,\mathrm{d}x \\ &\quad + \int_{\mathbb{R}^d} \mathrm{D}W \cdot \xi u_\tau^n \,\mathrm{d}x + \int_{\mathbb{R}^d} \mathrm{D}\gamma \cdot \mathrm{D}v_\tau^n \,\mathrm{d}x - \chi\int_{\mathbb{R}^d} \left(\gamma + \mathrm{D}v_\tau^n \cdot \xi\right) u_\tau^n \,\mathrm{d}x. \end{aligned} \tag{II.2.16}$$

With the Taylor expansion $\eta(x) = \eta(\varphi(x)) + (x - \varphi(x)) \cdot \mathrm{D}\eta(\varphi(x)) + \frac{1}{2}(x - \varphi(x)) \cdot \mathrm{D}^2\eta(x_*)(x - \varphi(x))$ and the definition of the push-forward $\varphi_\# u_\tau^{n-1} = u_\tau^n$, the first term in (II.2.16) can be rewritten as

$$\frac{1}{\tau}\int_{\mathbb{R}^d} \xi(\varphi(x)) \cdot (\varphi(x) - x) u_\tau^{n-1}(x) \,\mathrm{d}x \leq \frac{1}{\tau}\int_{\mathbb{R}^d} \eta(u_\tau^n - u_\tau^{n-1}) \,\mathrm{d}x + \frac{1}{2\tau}\|\eta\|_{C^2}\mathbf{W}_2^2(u_\tau^n, u_\tau^{n-1}).$$

As the same calculations can be carried out with $-\xi$ and $-\gamma$ in place of ξ and γ, we conclude from (II.2.16):

$$\begin{aligned}
&-\frac{1}{2\tau}\|\xi\|_{C^1}\mathbf{W}_2^2(u_\tau^n,u_\tau^{n-1})\\
&\leq \frac{1}{\tau}\int_{\mathbb{R}^d}\xi(\varphi(x))\cdot(\varphi(x)-x)u_\tau^{n-1}(x)\,\mathrm{d}x+\frac{1}{\tau}\int_{\mathbb{R}^d}(v_\tau^n-v_\tau^{n-1})\gamma\,\mathrm{d}x-\int_{\mathbb{R}^d}u_\tau^n(x)^m\mathrm{div}\xi(x)\,\mathrm{d}x\\
&\quad+\int_{\mathbb{R}^d}\mathrm{D}W\cdot\xi u_\tau^n\,\mathrm{d}x+\int_{\mathbb{R}^d}\mathrm{D}\gamma\cdot\mathrm{D}v_\tau^n\,\mathrm{d}x-\chi\int_{\mathbb{R}^d}(\gamma+\mathrm{D}v_\tau^n\cdot\xi)\,u_\tau^n\,\mathrm{d}x\\
&\quad\leq\frac{1}{2\tau}\|\xi\|_{C^1}\mathbf{W}_2^2(u_\tau^n,u_\tau^{n-1}).
\end{aligned}\tag{II.2.17}$$

Consider now a nonnegative time-dependent test function $\psi\in C_c^\infty((0,\infty))\cap C([0,\infty))$. We multiply the chain of estimates (II.2.17) with $\tau\psi(n\tau)$ and sum over $n\geq 1$ to obtain

$$\begin{aligned}
&-\tau\|\psi\|_{C^0}\|\eta\|_{C^2}(\mathcal{E}(u_0,v_0)-\inf\mathcal{E})\\
&\leq\tau\sum_{n\geq 0}\int_{\mathbb{R}^d}(\eta(x)u_\tau^n(x)+\gamma(x)v_\tau^n(x))\,\frac{\psi(n\tau)-\psi((n+1)\tau)}{\tau}\,\mathrm{d}x\\
&\quad+\tau\sum_{n\geq 0}\int_{\mathbb{R}^d}\psi(n\tau)\left[-(u_\tau^n)^m\Delta\eta+u_\tau^n\mathrm{D}W\cdot\mathrm{D}\eta+\mathrm{D}\gamma\cdot\mathrm{D}v_\tau^n-\chi\,(\gamma+\mathrm{D}v_\tau^n\cdot\mathrm{D}\eta)\,u_\tau^n\right](x)\,\mathrm{d}x\\
&\quad\leq\tau\|\psi\|_{C^0}\|\eta\|_{C^2}(\mathcal{E}(u_0,v_0)-\inf\mathcal{E}),
\end{aligned}\tag{II.2.18}$$

where we have used (II.2.7) and

$$\begin{aligned}
\sum_{n\geq 1}\psi(n\tau)(u_\tau^n(x)-u_\tau^{n-1}(x))&=\sum_{n\geq 0}u_\tau^n(x)(\psi(n\tau)-\psi((n+1)\tau)),\\
\sum_{n\geq 1}\psi(n\tau)(v_\tau^n(x)-v_\tau^{n-1}(x))&=\sum_{n\geq 0}v_\tau^n(x)(\psi(n\tau)-\psi((n+1)\tau)).
\end{aligned}$$

For sign-changing test functions $\psi\in C_c^\infty((0,\infty))\cap C([0,\infty))$, we decompose ψ in its positive and negative part ψ_+ and ψ_-, viz. $\psi=\psi_+-\psi_-$, and subtract the respective estimates (II.2.18) for ψ_+ and ψ_-, respectively, to also arrive at (II.2.18) for arbitrary ψ. Introducing the notation

$$\psi_\tau(s):=\psi\left(\left\lfloor\frac{s}{\tau}\right\rfloor\tau\right)\qquad\text{for } s\geq 0,$$

we express (II.2.18) in terms of the discrete solution (u_τ,v_τ) to deduce:

Lemma II.17 (Discrete weak formulation). *For all $n\in\mathbb{N}$ and all test functions $\eta,\gamma\in C_c^\infty(\mathbb{R}^d)$ and $\psi\in C_c^\infty((0,\infty))\cap C([0,\infty))$, the following* discrete weak formulation *holds:*

$$\begin{aligned}
&\bigg|\int_0^\infty\int_{\mathbb{R}^d}(\eta(x)u_\tau(t,x)+\gamma(x)v_\tau(t,x))\,\frac{\psi_\tau(t)-\psi_\tau(t+\tau)}{\tau}\,\mathrm{d}x\,\mathrm{d}t\\
&\quad+\int_0^\infty\int_{\mathbb{R}^d}\psi_\tau(t)\bigg[-u_\tau(t,x)^m\Delta\eta(x)+u_\tau(t,x)\mathrm{D}W(x)\cdot\mathrm{D}\eta(x)+\mathrm{D}\gamma(x)\cdot\mathrm{D}v_\tau(t,x)\\
&\qquad\qquad-\chi\,(\gamma(x)+\mathrm{D}v_\tau(t,x)\cdot\mathrm{D}\eta(x))\,u_\tau(t,x)\bigg]\,\mathrm{d}x\,\mathrm{d}t\bigg|\\
&\quad\leq\tau\|\psi\|_{C^0}\|\eta\|_{C^2}(\mathcal{E}(u_0,v_0)-\inf\mathcal{E}).
\end{aligned}\tag{II.2.19}$$

In the special case $\chi>0$, one has $v_\tau^n\geq 0$ a.e. on $\mathbb{R}^d$ for each $n\in\mathbb{N}$, given that $v_0\geq 0$ a.e. on $\mathbb{R}^d$.

Proof. It remains to prove the additional assertion on nonnegativity of the v component, if $\chi>0$ and $v_0\geq 0$ a.e. on $\mathbb{R}^d$. We proceed by induction and assume that $v_\tau^{n-1}\geq 0$ a.e. on $\mathbb{R}^d$ for some fixed, but arbitrary $n \in \mathbb{N}$. Putting $\xi \equiv 0$ in (II.2.17), it immediately follows that v_τ^n is a solution to

$$\frac{v_\tau^n - v_\tau^{n-1}}{\tau} - \Delta v_\tau^n - \chi u_\tau^n = 0 \quad \text{in the sense of distributions on } \mathbb{R}^d.$$

Since thanks to the additional regularity from Proposition II.16, the left-hand side above is an element of $L^2(\mathbb{R}^d)$, we conclude that

$$-\Delta v_\tau^n + \frac{1}{\tau} v_\tau^n = \frac{1}{\tau} v_\tau^{n-1} + \chi u_\tau^n \qquad \text{a.e. on } \mathbb{R}^d.$$

Consequently,

$$v_\tau^n = \mathbf{G}_{\frac{1}{\tau}} * \left(\frac{1}{\tau} v_\tau^{n-1} + \chi u_\tau^n \right),$$

where the nonnegative kernel $\mathbf{G}_{\frac{1}{\tau}}$ is the so-called *Yukawa potential* [121, Thm. 6.23]

$$\mathbf{G}_{\frac{1}{\tau}}(x) = \int_0^\infty (4\pi s)^{-d/2} \exp\left(-\frac{|x|^2}{4s} - \frac{1}{\tau} s \right) \mathrm{d}s. \tag{II.2.20}$$

Since both kernel and right-hand side of the elliptic equation above are nonnegative, the claim follows. □

II.2.1.3. Compactness estimates and passage to continuous time

The proof of Theorem II.2 will be finished by passing to the continuous-time limit $\tau \to 0$. Therefore we need, in addition to Proposition II.15, several estimates on the discrete solution. As a preparation, we prove a classical estimate on Boltzmann's entropy $\mathcal{H}$ which is included here for the sake of completeness.

Lemma II.18 (Estimate on Boltzmann's entropy). *For all $u \in (\mathscr{P}_2 \cap L^m)(\mathbb{R}^d)$, one has for some constant $C > 0$:*

$$|\mathcal{H}(u)| \le C \left(\|u\|_{L^m}^m + (\mathbf{m}_2(u) + 1)^{\frac{d}{d+1}} \right) < \infty. \tag{II.2.21}$$

Proof. Define $\alpha := \frac{d+1}{d} > 1$ and note that (see e.g. [105])

$$r \log r \le C_0 r^m \qquad \text{for all } r \ge 1, \tag{II.2.22}$$

$$r \log \frac{1}{r} \le C_0 r^{1/\alpha} \qquad \text{for all } r \in (0,1], \tag{II.2.23}$$

for some constant $C_0 > 0$. By splitting up the Boltzmann entropy, one sees

$$|\mathcal{H}(u)| \le \int_{\{u \ge 1\}} u \log u \,\mathrm{d}x + \int_{\{u<1\}} u \log \frac{1}{u} \,\mathrm{d}x. \tag{II.2.24}$$

From (II.2.22), we conclude that

$$\int_{\{u \ge 1\}} u \log u \,\mathrm{d}x \le C_0 \|u\|_{L^m}^m. \tag{II.2.25}$$

Second, with (II.2.23) and Hölder's inequality, one gets

$$\int_{\{u<1\}} u \log \frac{1}{u} \,\mathrm{d}x \le C_0 \left(\int_{\mathbb{R}^d} u(x)(|x|^2+1) \,\mathrm{d}x \right)^{1/\alpha} \left(\int_{\mathbb{R}^d} \left(\frac{1}{|x|^2+1} \right)^{\frac{1}{\alpha-1}} \mathrm{d}x \right)^{\frac{\alpha-1}{\alpha}}, \tag{II.2.26}$$

where the last integral is finite since $d - 1 - \frac{2}{\alpha-1} = -1 - d < -1$. Combining (II.2.24), (II.2.25) and (II.2.26) yields the assertion. □

II.2.1.3.1. *Compactness estimates*

PROPOSITION II.19 (Additional *a priori* estimates). *Let (u_τ, v_τ) be the discrete solution obtained by the minimizing movement scheme* (I.2.11)&(I.2.12) *for $\tau > 0$. Then the following holds for $T > 0$:*

$$\mathfrak{m}_2(u_\tau^n) \leq 2\mathfrak{m}_2(u_0) + 4T\mathcal{E}(u_0, v_0) < \infty \qquad \forall n \leq \left\lfloor \frac{T}{\tau} \right\rfloor, \tag{II.2.27}$$

$$\|u_\tau^n\|_{L^m} \leq C_3 < \infty \qquad \forall n \geq 0, \tag{II.2.28}$$

$$\|v_\tau^n\|_{W^{1,2}} \leq C_5 < \infty \qquad \forall n \geq 0, \tag{II.2.29}$$

$$\int_0^T \|u_\tau(t)^{m/2}\|_{W^{1,2}}^2 \,\mathrm{d}t \leq C_6 < \infty, \tag{II.2.30}$$

$$\int_0^T \|u_\tau(t)\|_{L^2}^2 \,\mathrm{d}t \leq C_7 < \infty, \tag{II.2.31}$$

$$\int_0^T \|v_\tau(t)\|_{W^{2,2}}^2 \,\mathrm{d}t \leq C_8 < \infty. \tag{II.2.32}$$

PROOF. By the triangular and Hölder inequalities,

$$\begin{aligned}\sqrt{\mathfrak{m}_2(u_\tau^n)} - \sqrt{\mathfrak{m}_2(u_0)} &\leq \sum_{k=1}^n \mathbf{W}_2(u_\tau^k, u_\tau^{k-1}) \\ &\leq \left(\sum_{k=1}^n 1\right)^{1/2} \left(\sum_{k=1}^n \mathbf{W}_2^2(u_\tau^k, u_\tau^{k-1})\right)^{1/2} \leq (2\tau n(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}))^{1/2}.\end{aligned}$$

Using Young's inequality and $n\tau \leq T$ yields (II.2.27):

$$\mathfrak{m}_2(u_\tau^n) \leq 2\mathfrak{m}_2(u_0) + 4n\tau\mathcal{E}(u_0, v_0) \leq 2\mathfrak{m}_2(u_0) + 4T(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}).$$

By an analogous argument, one obtains

$$\|v_\tau^n\|_{L^2}^2 \leq 2\|v_0\|_{L^2}^2 + 4T(\mathcal{E}(u_0, v_0) - \inf \mathcal{E}).$$

From the energy estimate (II.2.6) and the second part of Proposition II.11, we can deduce similarly to the proof of Proposition II.14 that there exists $C_3 > 0$ independent of τ such that (II.2.28) holds. Furthermore, using in addition the third part of Proposition II.11, one obtains

$$\|\mathrm{D}v_\tau^n\|_{L^2}^2 \leq C_0 \left(\mathcal{E}(u_0, v_0) + |\inf W| + C_3^{2(1-\theta)}\right),$$

and consequently (II.2.29):

$$\|v_\tau^n\|_{W^{1,2}} = \left(\|v_\tau^n\|_{L^2}^2 + \|\mathrm{D}v_\tau^n\|_{L^2}^2\right)^{1/2} \leq C_5 < \infty.$$

The proofs of (II.2.30)–(II.2.32) involve the application of Proposition II.16 and are similar. We give the proof of (II.2.30) as an example. The following estimate holds due to Proposition II.16 for $n = \lfloor \frac{T}{\tau} \rfloor + 1$:

$$\begin{aligned}&\int_0^T \|\mathrm{D}u_\tau(t)^{m/2}\|_{L^2}^2 \,\mathrm{d}t \leq \sum_{k=1}^n \tau \|\mathrm{D}(u_\tau^k)^{m/2}\|_{L^2}^2 \\ &\leq \sum_{k=1}^n C_0\tau \left[\frac{1}{\tau}\left(\mathcal{H}(u_\tau^{k-1}) - \mathcal{H}(u_\tau^k) + \mathcal{F}(v_\tau^{k-1}) - \mathcal{F}(v_\tau^k)\right) + \|u_\tau^k\|_{L^m}^{(1-A)\frac{m}{m-1}} + \|\Delta W\|_{L^\infty}\right] \\ &\leq \sum_{k=1}^n \left[C_0\left(\mathcal{H}(u_\tau^{k-1}) - \mathcal{H}(u_\tau^k) + \mathcal{F}(v_\tau^{k-1}) - \mathcal{F}(v_\tau^k)\right) + \tau C_0 C_3^{(1-A)\frac{m}{m-1}} + \tau C_0 \|\Delta W\|_{L^\infty}\right],\end{aligned}$$

where we have used (II.2.28) in the last step. Subsequently, we apply the estimates (II.2.21) and (II.2.27)–(II.2.29) to deduce a τ-uniform bound on $\int_0^T \|\mathrm{D}u_\tau(t)^{m/2}\|_{L^2}^2\,\mathrm{d}t$.
With the identity $\|u_\tau(t)^{m/2}\|_{L^2}^2 = \|u_\tau(t)\|_{L^m}^m$ and (II.2.28), (II.2.30) follows. □

II.2.1.3.2. *Passage to the continuous-time limit*

We can now prove a first convergence result:

Proposition II.20 (Continuous-time limit of discrete solutions). *Let $(\tau_k)_{k\in\mathbb{N}}$ be a vanishing sequence of step sizes, i.e. $\tau_k \searrow 0$ as $k\to\infty$, and let $(u_{\tau_k}, v_{\tau_k})_{k\geq 0}$ be the corresponding sequence of discrete solutions obtained by the minimizing movement scheme.*

Then for each $T>0$, there exist a subsequence (non-relabelled) and limit curves $u\in C^{1/2}([0,T];\mathscr{P}_2(\mathbb{R}^d))$ and $v\in C^{1/2}([0,T];L^2(\mathbb{R}^d))$ such that the following holds for $k\to\infty$:
(a) For fixed $t\in[0,T]$, $u_{\tau_k}\to u$ narrowly in $\mathscr{P}(\mathbb{R}^d)$,
(b) $v_{\tau_k}\to v$ uniformly with respect to $t\in[0,T]$ in $L^2(\mathbb{R}^d)$,
(c) $(u_{\tau_k})^{m/2}\rightharpoonup u^{m/2}$ weakly in $L^2([0,T];W^{1,2}(\mathbb{R}^d))$,
(d) $u_{\tau_k}\rightharpoonup u$ weakly in $L^2([0,T];L^2(\mathbb{R}^d))$,
(e) $v_{\tau_k}\rightharpoonup v$ weakly in $L^2([0,T];W^{2,2}(\mathbb{R}^d))$,
(f) $u_{\tau_k}\to u$ strongly in $L^m([0,T];L^m(\Omega))$ for all bounded domains $\Omega\subset\mathbb{R}^d$,
(g) $v_{\tau_k}\to v$ strongly in $L^2([0,T];W^{1,2}(\Omega))$ for all bounded domains $\Omega\subset\mathbb{R}^d$.
If $\chi>0$ and $v_0\geq 0$ a.e. on $\mathbb{R}^d$, we have $v(t,\cdot)\geq 0$ a.e. on $\mathbb{R}^d$, at each $t\geq 0$.

Proof. The claims (a)&(b) are consequences of the *a priori* estimate (II.2.27) and Prokhorov's theorem as well as (II.2.29) and the Banach-Alaoglu theorem, respectively. The claimed $\frac{1}{2}$-Hölder continuity of the limit curve (u,v) is obtained via estimate (II.2.8) and a refined version of the Arzelà-Ascoli theorem [4, Thm. 3.3.1]. The claims (c)–(e) are obtained from the Banach-Alaoglu theorem and the estimates (II.2.30)–(II.2.32), respectively. In the special case $\chi>0$, nonnegativity of $v(t,\cdot)$ is inherited from the initial datum since the $(v_\tau^n)_{n\in\mathbb{N}}$ are nonnegative thanks to Lemma II.17 and the set $L^2_+(\mathbb{R}^d):=\{\rho\in L^2(\mathbb{R}^d):\ \rho\geq 0 \text{ a.e. on } \mathbb{R}^d\}$ is a convex and closed subset of $L^2(\mathbb{R}^d)$.

To prove (f), we let $\Omega\subset\mathbb{R}^d$ be an arbitrary bounded domain and seek to apply Theorem I.6. Define the Banach space $\mathbf{Y}:=L^m(\Omega)$, the l.s.c. functional $\mathcal{A}(u):=\|u^{m/2}\|^2_{W^{1,2}(\Omega)}$ and the distance $\mathbf{W}$ on $\mathbf{Y}$ via

$$\mathbf{W}(u,\widetilde{u}):=\inf\Big\{\mathbf{W}_2(\rho,\widetilde{\rho}):\ \rho,\widetilde{\rho}\in\mathscr{P}_2(\mathbb{R}^d),\,\mathbf{m}_2(\rho),\mathbf{m}_2(\widetilde{\rho})\leq 2\mathbf{m}_2(u_0)+4T\mathcal{E}(u_0,v_0),\,\rho|_\Omega=u,\,\widetilde{\rho}|_\Omega=\widetilde{u}\Big\}.$$

Thanks to the topological properties of the L^2-Wasserstein distance $\mathbf{W}_2$, it is immediate that, if the admissible set in the infimum above is nonempty, the infimum above is attained. Furthermore, $\mathbf{W}$ satisfies the requirements of Theorem I.6.

We show that for every $b\in\mathbb{R}$ the set $\{u\in\mathbf{Y}:\ \mathcal{A}(u)\leq b\}$ is relatively compact in $\mathbf{Y}$. By the Rellich-Kondrachov compactness theorem, $W^{1,2}(\Omega)$ is compactly embedded in $L^2(\Omega)$. So, for every sequence $(u_l)_{l\in\mathbb{N}}$ in $\{u\in\mathbf{Y}:\ \mathcal{A}(u)\leq b\}$, there exists a (non-relabelled) subsequence and $\sigma\in L^2(\Omega)$ such that $\|u_l^{m/2}-\sigma\|_{L^2(\Omega)}\to 0$ as $l\to\infty$. On a further subsequence, one has convergence almost everywhere of u_l to $\sigma^{2/m}$. With the identity $\|u_l\|^m_{L^m(\Omega)}=\|u_l^{m/2}\|^2_{L^2(\Omega)}\overset{l\to\infty}{\longrightarrow}\|\sigma\|^2_{L^2(\Omega)}=\|\sigma^{2/m}\|^m_{L^m(\Omega)}$, we conclude that $\|u_l-\sigma^{2/m}\|_{L^m(\Omega)}\to 0$ as $l\to\infty$, i.e. the relative compactness of $\{u\in\mathbf{Y}:\ \mathcal{A}(u)\leq b\}$ in $\mathbf{Y}$.

Consider the sequence $(U_k)_{k\in\mathbb{N}}$ defined by $U_k:=(u_{\tau_k}|_{[0,\infty)\times\Omega})_{k\in\mathbb{N}}$. The first assumption (I.2.13) of Theorem I.6 is an immediate consequence of (II.2.30) as

$$\sup_{k\geq 0}\int_0^T\mathcal{A}(u_{\tau_k}(t))\,\mathrm{d}t=\sup_{k\geq 0}\int_0^T\|u_{\tau_k}(t)^{m/2}\|^2_{W^{1,2}(\Omega)}\,\mathrm{d}t\leq C_6<\infty.$$

In order to show the second assumption (I.2.14) of Theorem I.6, we distinguish two cases. Let without loss of generality $0 < h < 1$ and notice that by construction of $\mathbf{W}$ and estimate (II.2.27):

$$\mathbf{W}(U_k(t+h), U_k(t)) \leq \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t)).$$

Case 1: $0 < h \leq \tau_k$.
Then we obtain by using Hölder's inequality and the distance estimate (II.2.7):

$$\int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t))\,\mathrm{d}t = \sum_{n=1}^{\left\lfloor \frac{T}{\tau_k} \right\rfloor} h\mathbf{W}_2(u_{\tau_k}^n, u_{\tau_k}^{n+1}) \leq \left(\sum_{n=1}^{\left\lfloor \frac{T}{\tau_k} \right\rfloor} \mathbf{W}_2^2(u_{\tau_k}^n, u_{\tau_k}^{n+1}) \right)^{1/2} \left(\sum_{n=1}^{\left\lfloor \frac{T}{\tau_k} \right\rfloor} h^2 \right)^{1/2}$$

$$\leq \sqrt{2(T+1)(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})h} \leq \sqrt{2h(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})}(T+1).$$

Case 2: $\tau_k < h < 1$.
Here, we apply the Hölder-type estimate (II.2.8):

$$\int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t))\,\mathrm{d}t \leq \sqrt{2h(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})}(T-h) \leq \sqrt{2h(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})}(T+1).$$

Obviously, by combination of the two cases, (I.2.14) follows:

$$0 \leq \sup_{k \geq 0} \int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t))\,\mathrm{d}t \leq \sqrt{2h(\mathcal{E}(u_0, v_0) - \inf \mathcal{E})}(T+1) \xrightarrow{h \searrow 0} 0.$$

Hence, the application of Theorem I.6 yields the existence of a subsequence on which $u_{\tau_k}(t, \cdot) \to u(t, \cdot)$ strongly in $L^m(\Omega)$, in measure w.r.t. $t \in (0, T)$ as $k \to \infty$. By the estimate on the L^m norm (II.2.28) in combination with the Radon-Riesz and and the dominated convergence theorems, we conclude that $u_{\tau_k} \to u$ strongly in $L^m([0, T] \times \Omega)$.

The proof of (g) is similar: We apply Theorem I.6 with $\mathbf{Y} := W^{1,2}(\Omega)$, $\mathcal{A}(v) := \|v\|^2_{W^{2,2}(\Omega)}$ and $\mathbf{W}(v, \widetilde{v}) := \|v - \widetilde{v}\|_{L^2(\Omega)}$. Analogous arguments show that the assumptions (I.2.13)&(I.2.14) are satisfied for $(U_k)_{k \in \mathbb{N}}$ defined by $U_k := (v_{\tau_k}|_{[0,\infty) \times \Omega})_{k \in \mathbb{N}}$. We obtain the existence of a subsequence on which $v_{\tau_k}(t, \cdot) \to v(t, \cdot)$ strongly in $W^{1,2(\Omega)}$, in measure w.r.t. $t \in (0, T)$. By the uniform estimate (II.2.29), the dominated convergence and Radon-Riesz theorems, one has strong convergence of v_{τ_k} to v in $L^2([0, T]; W^{1,2}(\Omega))$. By a diagonal argument, setting $\Omega := \mathbb{B}_R(0)$ and letting $R \nearrow \infty$, we deduce that (f)&(g) are true simultaneously for every bounded domain, extracting a further subsequence. Moreover, we may assume that (u_{τ_k}, v_{τ_k}) converges to (u, v) pointwise almost everywhere in $[0, T] \times \mathbb{R}^d$. □

To complete the proof of Theorem II.2, it remains to verify that (u, v) is a solution to (II.1.1) in the sense of distributions. To this end, we show that the discrete weak formulation (II.2.19) converges to the continuous-time weak formulation of system (II.1.1), using the convergence and integrability properties of the test functions η, γ and ψ and those of u_{τ_k} and v_{τ_k} from Proposition II.20. In particular, one has uniform convergence of ψ_{τ_k} to ψ and $\frac{1}{\tau_k}(\psi_{\tau_k}(\cdot + \tau_k) - \psi_{\tau_k})$ to $\partial_t \psi$ on $(0, \infty)$. Since the integrand in (II.2.19) is of compact support, the strong convergence results in Proposition II.20(f)&(g) are applicable. Hence, it follows that

$$0 = \lim_{k \to \infty} \left(\int_0^\infty \int_{\mathbb{R}^d} \left(\eta(x) u_{\tau_k}(t, x) + \gamma(x) v_{\tau_k}(t, x) \right) \frac{\psi_{\tau_k}(t) - \psi_{\tau_k}(t + \tau_k)}{\tau_k} \,\mathrm{d}x\,\mathrm{d}t \right.$$
$$+ \int_0^\infty \int_{\mathbb{R}^d} \psi_{\tau_k}(t) \Big[- u_{\tau_k}(t, x)^m \Delta\eta(x) + u_{\tau_k}(t, x) \mathrm{D}W(x) \cdot \mathrm{D}\eta(x) + \mathrm{D}\gamma(x) \cdot \mathrm{D}v_{\tau_k}(t, x)$$
$$\left. - \chi \left(\gamma(x) + \mathrm{D}v_{\tau_k}(t, x) \cdot \mathrm{D}\eta(x) \right) u_{\tau_k}(t, x) \Big] \,\mathrm{d}x\,\mathrm{d}t \right)$$

$$
\begin{aligned}
&= \int_0^\infty \int_{\mathbb{R}^d} (-\partial_t \psi(t)) \left(\eta(x)u(t,x) + \gamma(x)v(t,x)\right) \,\mathrm{d}x\,\mathrm{d}t \\
&+ \int_0^\infty \int_{\mathbb{R}^d} \psi(t) \Big[- u(t,x)^m \Delta\eta(x) + u(t,x)\mathrm{D}W(x)\cdot \mathrm{D}\eta(x) + \mathrm{D}\gamma(x)\cdot \mathrm{D}v(t,x) \\
&\qquad\qquad - \chi\left(\gamma(x) + \mathrm{D}v(t,x)\cdot \mathrm{D}\eta(x)\right) u(t,x) \Big] \,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}
$$

which finishes the proof of Theorem II.2. □

II.2.2. Convergence to equilibrium

This section is concerned with the proof of Theorem II.4, as a revised form of the joint article [189] with Daniel Matthes. In the following, we always require Assumption II.3.

II.2.2.1. Preliminaries

We first summarize some preliminary results which are of particular importance for the forthcoming analysis of the long-time behaviour. They can be derived with the methods from Section II.2.1, see [189] for more details.

On the metric space $\mathbf{X} = \mathscr{P}_2(\mathbb{R}^3) \times L^2(\mathbb{R}^3)$, the energy $\mathcal{E} : \mathbf{X} \to \mathbb{R}_\infty$ reads as

$$
\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^3} \left(\frac{1}{2}u^2 + uW + \frac{1}{2}|\mathrm{D}v|^2 + \frac{\kappa}{2}v^2 + \varepsilon u\phi(v)\right) \mathrm{d}x & \text{if } (u,v) \in L^2(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3), \\ +\infty & \text{otherwise}. \end{cases}
$$

Proposition II.21 (Properties of the entropy functional $\mathcal{E}$). *The functional $\mathcal{E}$ defined above has the following properties:*

(a) There exist C_0, $C_1 > 0$ such that

$$
\mathcal{E}(u,v) \geq C_0 \left[\|u\|_{L^2}^2 + \mathfrak{m}_2(u) + \|v\|_{W^{1,2}}^2 - C_1 \right]. \tag{II.2.33}
$$

In particular, $\mathcal{E}$ is bounded from below.

(b) $\mathcal{E}$ is weakly lower semicontinuous in the following sense: For every sequence $(u_n, v_n)_{n\in\mathbb{N}}$ in $\mathbf{X}$, where $(u_n)_{n\in\mathbb{N}}$ converges narrowly to some $u \in \mathscr{P}_2(\mathbb{R}^3)$ and where $(v_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\mathbb{R}^3)$ to some $v \in L^2(\mathbb{R}^3)$, one has

$$
\mathcal{E}(u,v) \leq \liminf_{n\to\infty} \mathcal{E}(u_n, v_n).
$$

(c) For sufficiently small $\varepsilon > 0$, $\mathcal{E}$ is λ'-geodesically convex for some $\lambda' > 0$ with respect to the distance induced by the norm $\|(\widetilde{u},\widetilde{v})\|_{L^2\times L^2} := \sqrt{\|\widetilde{u}\|_{L^2}^2 + \|\widetilde{v}\|_{L^2}^2}$.

Proof. For part (a), we observe that due to λ_0-convexity of W, one has

$$
W(x) - W(x_{\min}) \geq \frac{\lambda_0}{4}|x|^2 - \frac{\lambda_0}{2}|x_{\min}|^2,
$$

where $x_{\min} \in \mathbb{R}^3$ is the unique minimizer of W. Moreover, with convexity of ϕ, we deduce

$$
\int_{\mathbb{R}^3} u\phi(v)\,\mathrm{d}x \geq \phi(0) + \phi'(0)\|uv\|_{L^1} \geq \phi(0) + C\phi'(0)\|\mathrm{D}v\|_{L^2}\|u\|_{L^2}^{1/3},
$$

using that $\|u\|_{L^1(\mathbb{R}^3)} = 1$ and the following chain of estimates:

$$
\|uv\|_{L^1} \leq \|u\|_{L^{6/5}}\|v\|_{L^6} \leq C\|\mathrm{D}v\|_{L^2}\|u\|_{L^1}^{2/3}\|u\|_{L^2}^{1/3}. \tag{II.2.34}
$$

All in all, we arrive at

$$
\mathcal{E}(u,v) \geq \frac{1}{2}\|u\|_{L^2}^2 + \frac{\lambda_0}{4}\mathfrak{m}_2(u) - \frac{\lambda_0}{2}|x_{\min}|^2 + W(x_{\min})
$$

$$+\frac{1}{2}\|\mathrm{D}v\|_{L^2}^2+\frac{\kappa}{2}\|v\|_{L^2}^2-\varepsilon|\phi(0)|-\varepsilon C|\phi'(0)|\|\mathrm{D}v\|_{L^2}\|u\|_{L^2}^{1/3}.$$

From this, the desired estimate follows by means of Young's inequality.

Part (b) follows similarly to Proposition II.12. Note that the additional assumption of boundedness of $(\mathfrak{m}_2(u_n))_{n\in\mathbb{N}}$ is not needed here — this is a consequence of the stronger assumption on W (see Assumption II.3) and the resulting coercivity estimate (II.2.33).

To prove (c), consider a geodesic $(u_s,v_s)_{s\in[0,1]}$ with respect to the flat metric induced by $\|\cdot\|_{L^2\times L^2}$, that is $u_s=(1-s)u_0+su_1$ and $v_s=(1-s)v_0+sv_1$ for given $u_0,u_1\in(\mathscr{P}_2\cap L^2)(\mathbb{R}^3)$ and $v_0,v_1\in W^{1,2}(\mathbb{R}^3)$. It then follows that

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathcal{E}(u_s,v_s)&=\int_{\mathbb{R}^3}\big((u_1-u_0)^2+|\mathrm{D}(v_1-v_0)|^2+\kappa(v_1-v_0)^2\\&\qquad+2\varepsilon\phi'(v_s)(u_1-u_0)(v_1-v_0)+\varepsilon u_s\phi''(v_s)(v_1-v_0)^2\big)\,\mathrm{d}x\\&\geq\int_{\mathbb{R}^3}\begin{pmatrix}u_1-u_0\\v_1-v_0\end{pmatrix}^{\mathrm{T}}A_s\begin{pmatrix}u_1-u_0\\v_1-v_0\end{pmatrix}\mathrm{d}x\quad\text{with}\quad A_s:=\begin{pmatrix}1&\varepsilon\phi'(v_s)\\\varepsilon\phi'(v_s)&\kappa\end{pmatrix},\end{aligned}$$

where we have used that ϕ is convex. Thus, $\mathcal{E}$ is λ'-convex with respect to the flat distance above if $A_s\geq\lambda'\mathbb{1}$ for all $s\in[0,1]$. Recalling that $0<-\phi'(v_s)\leq\overline{\phi'}$ by Assumption II.3, it follows from elementary linear algebra that $\varepsilon^2\overline{\phi'}^2<\kappa$ is sufficient to find a suitable $\lambda'>0$ with $A_s\geq\lambda'\mathbb{1}$. □

By essentially the same argumentation as in the previous section, one constructs a time-discrete solution via the minimizing movement scheme (I.2.11) and passes to the continuous-time limit:

Theorem II.22 (Minimizing movement and existence of weak solutions [189]). *For every $\tau>0$ and every $(\widetilde{u},\widetilde{v})\in\mathbf{X}$, there exists at least one minimizer $(u,v)\in\mathbf{X}$ of $\mathcal{E}_\tau(\cdot\,|\,\widetilde{u},\widetilde{v})$. One additionally has $u\in W^{1,2}(\mathbb{R}^3)$ and $v\in W^{2,2}(\mathbb{R}^3)$. If moreover $\widetilde{v}\geq 0$ a.e. on $\mathbb{R}^3$, then also $v\geq 0$ a.e. on $\mathbb{R}^3$.*

For given initial data $(u_0,v_0)\in\mathbf{X}\cap(L^2(\mathbb{R}^3)\times W^{1,2}(\mathbb{R}^3))$, there exists a null sequence $(\tau_k)_{k\in\mathbb{N}}$ such that the corresponding sequence $(u_{\tau_k},v_{\tau_k})_{k\in\mathbb{N}}$ of discrete solutions converges to a solution (u,v) to (II.1.1) *in the sense of distributions on $(0,\infty)\times\mathbb{R}^3$, attaining the initial condition $(u(0,\cdot),v(0,\cdot))=(u_0,v_0)$. Specifically, one has for each $T>0$:*

$$\begin{aligned}&u_{\tau_k}(t,\cdot)\to u(t,\cdot)\text{ narrowly in }\mathscr{P}(\mathbb{R}^3)\text{, pointwise with respect to }t\in[0,T],\\&v_{\tau_k}(t,\cdot)\to v(t,\cdot)\text{ in }L^2(\mathbb{R}^3)\text{, uniformly with respect to }t\in[0,T],\\&(u,v)\in C^{1/2}([0,T];(\mathbf{X},\mathbf{d})).\end{aligned}$$

If moreover $v_0\geq 0$ a.e. on $\mathbb{R}^3$, then also $v(t,\cdot)\geq 0$ a.e. on $\mathbb{R}^3$ for each $t\geq 0$.

II.2.2.2. The stationary solution

We provide the characterization of a stationary state of system (II.1.1) and prove some relevant properties.

II.2.2.2.1. *Existence and uniqueness*

At first, we show existence and uniqueness.

Proposition II.23. *For each sufficiently small $\varepsilon>0$, there exists a unique minimizer $(u_\infty,v_\infty)\in\mathbf{X}$ of $\mathcal{E}$, for which the following holds:*
(u_∞,v_∞) is an element of $W^{1,2}(\mathbb{R}^3)\times W^{2,2}(\mathbb{R}^3)$, a stationary solution to (II.1.1) *and a solution to the Euler-Lagrange system*

$$\Delta v_\infty-\kappa v_\infty=\varepsilon u_\infty\phi'(v_\infty),\tag{II.2.35}$$

$$u_\infty=[U_\varepsilon-W-\varepsilon\phi(v_\infty)]_+,\tag{II.2.36}$$

where $U_\varepsilon \in \mathbb{R}$ is chosen such that $\|u_\infty\|_{L^1} = 1$, and $[\cdot]_+$ denotes the positive part.
Moreover, $u_\infty \in C^{0,1}(\mathbb{R}^3)$ with compact support, $v_\infty \in C^0(\mathbb{R}^3)$ is strictly positive and there exists $V > 0$ independent of $\varepsilon > 0$ such that $\|v_\infty\|_{L^\infty(\mathbb{R}^3)} \leq V$.

Proof. We prove that $\mathcal{E}$ possesses a unique minimizer (u_∞, v_∞). Let therefore be given a minimizing sequence $(u_n, v_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \mathcal{E}(u_n, v_n) = \inf \mathcal{E} > -\infty$. As the sequence $(\mathcal{E}(u_n, v_n))_{n\in\mathbb{N}}$ is bounded, we can, using the coercivity estimate (II.2.33) from Proposition II.21(a), extract a (non-relabelled) subsequence, on which $(u_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\mathbb{R}^3)$ and narrowly in $\mathscr{P}(\mathbb{R}^3)$ to some $u_\infty \in \mathscr{P}_2(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, and $(v_n)_{n\in\mathbb{N}}$ converges weakly in $W^{1,2}(\mathbb{R}^3)$ to some $v_\infty \in W^{1,2}(\mathbb{R}^3)$, as $n \to \infty$. By weak lower semicontinuity, see Proposition II.21(b), (u_∞, v_∞) is indeed a minimizer of $\mathcal{E}$ on $\mathbf{X}$.

Since $(u_\infty, v_\infty) \in \operatorname{argmin} \mathcal{E}_\tau(\cdot \mid u_\infty, v_\infty)$ for arbitrary $\tau > 0$, Proposition II.22 yields that (u_∞, v_∞) is an element of $W^{1,2}(\mathbb{R}^3) \times W^{2,2}(\mathbb{R}^3)$, and because of the continuous embedding of $W^{2,2}(\mathbb{R}^3)$ into $C^0(\mathbb{R}^3)$, it follows that $\|v_\infty\|_{L^\infty} \leq V$ for some $V > 0$.

Uniqueness of the minimizer is, by [179, Thm. 5.32], a consequence of λ'-geodesic convexity of $\mathcal{E}$ with respect to the distance induced by $\|\cdot\|_{L^2\times L^2}$ for some $\lambda' > 0$ as proved in Proposition II.21(c).

We show that there is a set of Euler-Lagrange equations characterizing (u_∞, v_∞). The following variational inequality holds thanks to the minimizing property of (u_∞, v_∞):

$$\begin{aligned} 0 &\leq \left.\frac{\mathrm{d}^+}{\mathrm{d}s}\right|_{s=0} \mathcal{E}(u_\infty + s\widetilde{u}, v_\infty + s\widetilde{v}) \\ &= \int_{\mathbb{R}^3} (u_\infty + W + \varepsilon\phi(v_\infty))\widetilde{u}\,\mathrm{d}x + \int_{\mathbb{R}^3} (-\Delta v_\infty + \kappa v_\infty + \varepsilon u_\infty \phi'(v_\infty))\widetilde{v}\,\mathrm{d}x, \end{aligned} \tag{II.2.37}$$

for arbitrary maps $\widetilde{u}$, $\widetilde{v}$ such that $u_\infty + \widetilde{u} \geq 0$ on $\mathbb{R}^3$ and $\int_{\mathbb{R}^3} \widetilde{u}\,\mathrm{d}x = 0$.

First, we consider the second component and thus set $\widetilde{u} = 0$ in (II.2.37). As there are no constraints on v_∞, it is allowed to replace $\widetilde{v}$ by $-\widetilde{v}$ in (II.2.37), yielding equality and hence, (II.2.35).

Second, we consider the first component and set $\widetilde{v} = 0$ in (II.2.37). For arbitrary ψ such that $\int_{\mathbb{R}^3} \psi\,\mathrm{d}x \leq 1$ and $\psi + u_\infty \geq 0$ on $\mathbb{R}^3$, we put

$$\widetilde{u}_\psi := \frac{1}{2}\psi - \frac{1}{2}u_\infty \int_{\mathbb{R}^3} \psi\,\mathrm{d}x,$$

and observe that $u_\infty + u_\psi \geq 0$ on $\mathbb{R}^3$ and $\int_{\mathbb{R}^3} \widetilde{u}_\psi\,\mathrm{d}x = 0$, since u_∞ has mass equal to 1. By straightforward calculation, we obtain

$$0 \leq \int_{\mathbb{R}^3} (u_\infty + W + \varepsilon\phi(v_\infty) - U_\varepsilon)\psi\,\mathrm{d}x, \tag{II.2.38}$$

for all ψ as above and the constant

$$U_\varepsilon := \int_{\mathbb{R}^3} (u_\infty^2 + W u_\infty + \varepsilon u_\infty \phi(v_\infty))\,\mathrm{d}x \in \mathbb{R}.$$

Fix $x \in \mathbb{R}^3$. If $u_\infty(x) > 0$, choosing ψ supported on a small neighbourhood of x and replacing by $-\psi$ in (II.2.38) eventually yields

$$u_\infty(x) = U_\varepsilon - W(x) - \varepsilon\phi(v_\infty(x)).$$

If $u_\infty(x) = 0$, we obtain

$$U_\varepsilon - W(x) - \varepsilon\phi(v_\infty(x)) \leq 0.$$

Hence, for all $x \in \mathbb{R}^3$,

$$u_\infty(x) = [U_\varepsilon - W(x) - \varepsilon\phi(v_\infty(x))]_+.$$

The claimed Lipschitz continuity of u_∞ is easy to see from this representation. To prove strict positivity of v_∞, we make use of the Yukawa potential (see (II.2.20)) once more: (II.2.35) yields

$$v_\infty = -\varepsilon \mathbf{G}_\kappa * (u_\infty \phi'(v_\infty)).$$

Since there exist sets with positive volume on which u_∞ is strictly positive, it follows that $v_\infty(x) > 0$ for all $x \in \mathbb{R}^3$ (recall that $\mathbf{G}_\kappa > 0$ and $\phi' < 0$). □

II.2.2.2.2. *Structure of the Yukawa potential*

As a preparation to prove some crucial regularity estimates on the stationary solution (u_∞, v_∞), several properties of solutions to the elliptic partial differential equation $-\Delta h + \kappa h = f$ are needed.

Therefore, we again introduce for $\kappa > 0$ the *Yukawa potential* (also called *screened Coulomb* or *Bessel potential*) $\mathbf{G}_\kappa$ which has the following explicit representation on $\mathbb{R}^3$:

$$\mathbf{G}_\kappa(x) := \frac{1}{4\pi|x|}\exp(-\sqrt{\kappa}|x|) \quad \text{for all } x \in \mathbb{R}^3 \setminus \{0\}. \tag{II.2.39}$$

Additionally, we define for $\sigma > 0$ the kernel $\mathbf{Y}_\sigma$ by

$$\mathbf{Y}_\sigma := \frac{1}{\sigma}\mathbf{G}_{\frac{1}{\sigma}}.$$

In subsequent parts of this work, we will need the *iterates* $\mathbf{Y}_\sigma^k$ for $k \in \mathbb{N}$ defined inductively by

$$\mathbf{Y}_\sigma^1 := \mathbf{Y}_\sigma, \quad \mathbf{Y}_\sigma^{k+1} := \mathbf{Y}_\sigma * \mathbf{Y}_\sigma^k.$$

The relevant properties of $\mathbf{G}_\kappa$ and $\mathbf{Y}_\sigma$ are summarized in Lemma II.24 below.

Lemma II.24 (Yukawa potential). *The following statements hold for all $\kappa > 0$, $\sigma > 0$ and $k \in \mathbb{N}$:*

(a) *$\mathbf{G}_\kappa$ and $\mathbf{Y}_\sigma$ are the fundamental solutions to $-\Delta h + \kappa h = f$ and $-\sigma\Delta h + h = f$ on $\mathbb{R}^3$, respectively.*

(b) *Let $p > 1$. If $f \in L^p(\mathbb{R}^3)$, then $\mathbf{G}_\kappa * f \in W^{2,p}(\mathbb{R}^3)$ and*

$$\kappa\|\mathbf{G}_\kappa * f\|_{L^p} + \sqrt{\kappa}\|\mathrm{D}(\mathbf{G}_\kappa * f)\|_{L^p} + \|\mathrm{D}^2(\mathbf{G}_\kappa * f)\|_{L^p} \le C_p\|f\|_{L^p}, \tag{II.2.40}$$

for some p-dependent constant $C_p > 0$. (Note that this fact is not obvious as $\mathrm{D}^2(\mathbf{G}_\kappa) \notin L^1$.)

(c) *For all $x \in \mathbb{R}^3 \setminus \{0\}$,*

$$\mathbf{Y}_\sigma(x) = \int_0^\infty \mathbf{H}_{\sigma t}(x)e^{-t}\,\mathrm{d}t,$$

where $\mathbf{H}_t$ is the heat kernel *on $\mathbb{R}^3$ at time $t > 0$, i.e.*

$$\mathbf{H}_t(\xi) = t^{-3/2}\mathbf{H}_1(t^{-1/2}\xi), \quad \text{with} \quad \mathbf{H}_1(\zeta) = (4\pi)^{-3/2}\exp\Big(-\frac{1}{4}|\zeta|^2\Big).$$

Additionally, one has

$$\mathbf{Y}_\sigma^k = \int_0^\infty \mathbf{H}_{\sigma r}\frac{r^{k-1}e^{-r}}{\Gamma(k)}\,\mathrm{d}r. \tag{II.2.41}$$

Moreover, $\mathbf{Y}_\sigma^k \in W^{1,q}(\mathbb{R}^3)$ for each $q \in [1, \frac{5}{4})$, and there are universal constants Y_q such that

$$\|\mathrm{D}\mathbf{Y}_\sigma^k\|_{L^q(\mathbb{R}^3)} \le Y_q(\sigma k)^{-Q} \quad \text{where} \quad Q := 2 - \frac{3}{2q} \in \Big[\frac{1}{2}, \frac{4}{5}\Big). \tag{II.2.42}$$

Proof. (a) The proof of the first assertion can be found in [121, Thm. 6.23]. From that, the second one follows by elementary calculations.

(b) According to [167, Ch. V, §3.3, Thm. 3], one has for $p > 1$:

$$\|\mathbf{G}_1 * f\|_{W^{2,p}} \le C_p\|f\|_{L^p}. \tag{II.2.43}$$

To prove assertion (b), we use a rescaling of the equation $-\Delta h + \kappa h = f$ by $\widetilde{x} := \sqrt{\kappa}x$. Consequently, $h(\widetilde{x}) = (\mathbf{G}_\kappa * f)\left(\frac{\widetilde{x}}{\sqrt{\kappa}}\right)$ is a solution to $-\Delta_{\widetilde{x}} h + h = \frac{f}{\kappa}$, i.e. $h(\widetilde{x}) = \left(\mathbf{G}_1 * \frac{f}{\kappa}\right)(\widetilde{x})$. By the transformation theorem, we obtain

$$\left(\int_{\mathbb{R}^3} \left|\frac{f(\widetilde{x})}{\kappa}\right|^p \,\mathrm{d}\widetilde{x}\right)^{\frac{1}{p}} = \kappa^{\frac{3}{2}-1}\|f\|_{L^p},$$

$$\left(\int_{\mathbb{R}^3} \left|\left(\mathbf{G}_1 * \frac{f}{\kappa}\right)(\widetilde{x})\right|^p \,\mathrm{d}\widetilde{x}\right)^{\frac{1}{p}} = \kappa^{\frac{3}{2}}\|\mathbf{G}_\kappa * f\|_{L^p},$$

$$\left(\int_{\mathbb{R}^3} \left|\mathrm{D}_{\widetilde{x}}\left(\mathbf{G}_1 * \frac{f}{\kappa}\right)(\widetilde{x})\right|^p \,\mathrm{d}\widetilde{x}\right)^{\frac{1}{p}} = \kappa^{\frac{3}{2}-\frac{1}{2}}\|\mathrm{D}_x(\mathbf{G}_\kappa * f)\|_{L^p},$$

$$\left(\int_{\mathbb{R}^3} \left|\mathrm{D}^2_{\widetilde{x}}\left(\mathbf{G}_1 * \frac{f}{\kappa}\right)(\widetilde{x})\right|^p \,\mathrm{d}\widetilde{x}\right)^{\frac{1}{p}} = \kappa^{\frac{3}{2}-1}\|\mathrm{D}^2_x(\mathbf{G}_\kappa * f)\|_{L^p},$$

which yields (II.2.40) after insertion in (II.2.43) and simplification.

(c) The first statement is a straightforward consequence of the integral-type representation of $\mathbf{G}_\kappa$ in (II.2.20). To prove the first claim of the second statement, we proceed by induction. For $k = 1$, equation (II.2.41) is just the definition of $\mathbf{Y}_\sigma$. Now assume that (II.2.41) holds for some $k \in \mathbb{N}$. Using the semigroup property $\mathbf{H}_{t_1+t_2} = \mathbf{H}_{t_1} * \mathbf{H}_{t_2}$ of the heat kernel, we find that

$$\begin{aligned}\mathbf{Y}_\sigma^{k+1} &= \int_0^\infty \int_0^\infty \mathbf{H}_{\sigma r_1} * \mathbf{H}_{\sigma r_2} e^{-r_1} r_2^{k-1} e^{-r_2} \frac{\mathrm{d}r_1\,\mathrm{d}r_2}{\Gamma(k)} \\ &= \int_0^\infty \int_0^\infty \mathbf{H}_{\sigma(r_1+r_2)} e^{-(r_1+r_2)} r_2^{k-1} \frac{\mathrm{d}r_1\,\mathrm{d}r_2}{\Gamma(k)}.\end{aligned}$$

Now perform a change of variables

$$r := r_1 + r_2,\ s := r_2,$$

which is of determinant 1 and leads to

$$\mathbf{Y}_\sigma^{k+1} = \int_0^\infty \mathbf{H}_{\sigma r} e^{-r}\left(\int_0^r s^{k-1}\,\mathrm{d}s\right)\frac{\mathrm{d}r}{\Gamma(k)} = \int_0^\infty \mathbf{H}_{\sigma r}\frac{e^{-r} r^k\,\mathrm{d}r}{k\Gamma(k)},$$

which is (II.2.41) with $k+1$ in place of k, using that $k\Gamma(k) = \Gamma(k+1)$.

For (II.2.42), first observe that $r \mapsto r^{k-1}e^{-r}/\Gamma(k)$ defines a probability density on $(0,\infty)$. We can thus apply Jensen's inequality to obtain

$$\|\mathrm{D}\mathbf{Y}_\sigma^k\|_{L^q}^q \le \int_0^\infty \|\mathrm{D}\mathbf{H}_{\sigma r}\|_{L^q}^q \frac{r^{k-1}e^{-r}\,\mathrm{d}r}{\Gamma(k)}. \tag{II.2.44}$$

The L^q-norm of $\mathrm{D}\mathbf{H}_{\sigma r}$ is easily evaluated using its definition,

$$\begin{aligned}\|\mathrm{D}\mathbf{H}_{\sigma r}\|_{L^q} &= (\sigma r)^{-3/2}\left(\int_{\mathbb{R}^3} |\mathrm{D}_\zeta \mathbf{H}_1\big((\sigma r)^{-1/2}\zeta\big)|^q\,\mathrm{d}\zeta\right)^{1/q} \\ &= (\sigma r)^{-3/2}\left(\int_{\mathbb{R}^3} |(\sigma r)^{-1/2}\mathrm{D}_\zeta \mathbf{H}_1(\zeta)|^q\,(\sigma r)^{3/2}\,\mathrm{d}\zeta\right)^{1/q} = (\sigma r)^{-Q}\|\mathrm{D}\mathbf{H}_1\|_{L^q}.\end{aligned}$$

By definition of the Γ function, we thus obtain from (II.2.44) that (notice that $Qq = 2q - \frac{3}{2} \in (0,1)$)

$$\|\mathrm{D}\mathbf{Y}_\sigma^k\|_{L^q} \le \|\mathrm{D}\mathbf{H}_1\|_{L^q}\left(\frac{\Gamma(k-Qq)}{\Gamma(k)}\sigma^{-Qq}\right)^{1/q}.$$

For further estimation, observe that the sequence $(a_k)_{k\in\mathbb{N}}$ with $a_k = k^{Qq}\frac{\Gamma(k-Qq)}{\Gamma(k)}$ is monotonically decreasing (to zero). Indeed,

$$\frac{a_{k+1}}{a_k} = \frac{(k+1)k^{Qq}\,(k-Qq)\Gamma(k-Qq)\Gamma(k)}{k^{Qq}\,k\Gamma(k)\Gamma(k-Qq)} = \Big(1+\frac{1}{k}\Big)^{Qq}\Big(1-\frac{Qq}{k}\Big)$$

is always less than 1 since $\xi \mapsto (1+\xi)^{-Qq}$ is convex:

$$\left(1-\frac{1}{k}\right)^{-Qq} \geq 1 + \frac{1}{k}(-Qq).$$

Therefore, $a_k \leq a_1$ for all $k \in \mathbb{N}$, and so (II.2.42) follows with $Y_q := (\Gamma(1-Qq))^{1/q}\|\mathrm{D}\mathbf{H}_1\|_{L^q}$.

□

In addition to the properties above, we also need Hölder estimates for the kernel $\mathbf{G}_\kappa$ providing a kind of *elliptic regularity* which is known for Poisson's kernel $\mathbf{G}_0$ (see [90, 121]). As a preparation, we calculate the derivatives of $\mathbf{G}_\kappa$ in $\mathbb{R}^3\backslash\{0\}$. For all $i,j,k \in \{1,2,3\}$, one has

$$\begin{aligned}
\partial_i\mathbf{G}_\kappa(x) &= -\frac{1}{4\pi}\frac{\exp(-\sqrt{\kappa}|x|)}{|x|^3}(\sqrt{\kappa}|x|+1)x_i,\\
\partial_i\partial_j\mathbf{G}_\kappa(x) &= -\frac{1}{4\pi}\exp(-\sqrt{\kappa}|x|)\left[\left(\frac{\kappa}{|x|^3}+\frac{3\sqrt{\kappa}}{|x|^4}+\frac{3}{|x|^5}\right)x_ix_j - \left(\frac{\sqrt{\kappa}}{|x|^2}+\frac{1}{|x|^3}\right)\delta_{ij}\right],\\
\partial_i\partial_j\partial_k\mathbf{G}_\kappa(x) &= -\frac{1}{4\pi}\exp(-\sqrt{\kappa}|x|)\frac{-\sqrt{\kappa}x_k}{|x|}\left[\left(\frac{\kappa}{|x|^3}+\frac{3\sqrt{\kappa}}{|x|^4}+\frac{3}{|x|^5}\right)x_ix_j - \left(\frac{\sqrt{\kappa}}{|x|^2}+\frac{1}{|x|^3}\right)\delta_{ij}\right]\\
&\quad - \frac{1}{4\pi}\exp(-\sqrt{\kappa}|x|)\left[\left(-\frac{3\kappa}{|x|^4}-\frac{12\sqrt{\kappa}}{|x|^5}-\frac{15}{|x|^6}\right)\frac{x_ix_jx_k}{|x|}\right.\\
&\quad \left. +\delta_{ij}\left(\frac{2\sqrt{\kappa}}{|x|^3}+\frac{3}{|x|^4}\right)\frac{x_k}{|x|} + \left(\frac{\kappa}{|x|^3}+\frac{3\sqrt{\kappa}}{|x|^4}+\frac{3}{|x|^5}\right)(\delta_{ik}x_j+\delta_{jk}x_i)\right],
\end{aligned}$$

where δ_{ij} denotes Kronecker's delta.

We prove the following

Lemma II.25 (Hölder estimate for second derivative). *Let $f \in C^{0,\alpha}(\mathbb{R}^3)$ for some $\alpha \in (0,1)$ and assume that it is of compact support. Then, there exists $C > 0$ such that for all $i,j \in \{1,2,3\}$ the following estimate holds:*

$$\left[\partial_i\partial_j(\mathbf{G}_\kappa * f)\right]_{C^{0,\alpha}} \leq C[f]_{C^{0,\alpha}}.$$

Here,

$$[g]_{C^{0,\alpha}(\mathbb{R}^3)} := \sup_{x,y\in\mathbb{R}^3,\, x\neq y}\frac{|g(x)-g(y)|}{|x-y|}$$

denotes the Hölder seminorm *of $g:\ \mathbb{R}^3 \to \mathbb{R}$.*

Proof. This result is an extension of the respective result for Poisson's equation (corresponding to $\kappa = 0$) proved by Lieb and Loss [121, Thm. 10.3]. Their method of proof is adapted here. In the following, $C, \widetilde{C}$ denote generic nonnegative constants.

The following holds for arbitrary test functions $\psi \in C_c^\infty(\mathbb{R}^3)$:

$$-\int_{\mathbb{R}^3}(\partial_j\psi)(x)(\partial_i(\mathbf{G}_\kappa * f))(x)\,\mathrm{d}x = \int_{\mathbb{R}^3} f(y)\int_{\mathbb{R}^3}(\partial_j\psi)(x)\partial_{x_i}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\;\mathrm{d}y,$$

which can be rewritten by the dominated convergence theorem and integration by parts as

$$\int_{\mathbb{R}^3} f(y)\int_{\mathbb{R}^3}(\partial_j\psi)(x)\partial_{x_i}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\;\mathrm{d}y$$

$$
\begin{aligned}
&= \lim_{\delta\to 0}\int_{\mathbb{R}^3} f(y)\int_{\mathbb{R}^3\setminus\mathbb{B}_\delta(y)}(\partial_j\psi)(x)\partial_{x_i}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\ \mathrm{d}y\\
&= \lim_{\delta\to 0}\int_{\mathbb{R}^3} f(y)\left[-\int_{\partial\mathbb{B}_\delta(y)}\psi(x)\partial_{x_i}\mathbf{G}_\kappa(x-y)\mathsf{e}_j\cdot\nu_{y,\delta}(x)\,\mathrm{d}\sigma(x)\right.\\
&\qquad\left.-\int_{\mathbb{R}^3\setminus\mathbb{B}_\delta(y)}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\right]\mathrm{d}y,
\end{aligned}
$$

where e_j is the j^{th} canonical unit vector in $\mathbb{R}^d$ and $\nu_{y,\delta}(x)=\frac{x-y}{\delta}$ is the unit outward normal vector in x on the sphere $\partial\mathbb{B}_\delta(y)$. The first part can be simplified explicitly by the transformation $z:=\frac{x-y}{\delta}$:

$$
\begin{aligned}
&-\int_{\partial\mathbb{B}_\delta(y)}\psi(x)\partial_{x_i}\mathbf{G}_\kappa(x-y)\mathsf{e}_j\cdot\nu_{y,\delta}(x)\,\mathrm{d}\sigma(x)\\
&\quad=\frac{1}{4\pi}\int_{\partial\mathbb{B}_1(0)}\psi(\delta z+y)\exp(-\sqrt{\kappa}\delta)(\sqrt{\kappa}\delta+1)z_iz_j\,\mathrm{d}\sigma(z),
\end{aligned}
$$

which converges as $\delta\to 0$ to $\psi(y)\frac{\delta_{ij}}{3}$. For the second part, we split the domain of integration $\mathbb{R}^3\setminus\mathbb{B}_\delta(y)$ into two parts:

$$
\begin{aligned}
&\int_{\mathbb{R}^3\setminus\mathbb{B}_\delta(y)}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\\
&=\int_{\mathbb{R}^3\setminus\mathbb{B}_1(y)}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x+\int_{\{1\ge|x-y|\ge\delta\}}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x.
\end{aligned}
$$

We use integration by parts to insert convenient additional terms:

$$
\begin{aligned}
&\int_{\mathbb{R}^3\setminus\mathbb{B}_1(y)}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x+\int_{\{1\ge|x-y|\ge\delta\}}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x=\\
&=\int_{\mathbb{R}^3\setminus\mathbb{B}_1(y)}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x+\int_{\{1\ge|x-y|\ge\delta\}}\psi(x)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\\
&\quad-\int_{\{1\ge|x-y|\ge\delta\}}\psi(y)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}x\\
&\quad+\int_{\partial\mathbb{B}_1(y)}\psi(y)\partial_{x_j}\mathbf{G}_\kappa(x-y)\mathsf{e}_i\cdot\nu_{y,1}(x)\,\mathrm{d}\sigma(x)-\int_{\partial\mathbb{B}_\delta(y)}\psi(y)\partial_{x_j}\mathbf{G}_\kappa(x-y)\mathsf{e}_i\cdot\nu_{y,\delta}(x)\,\mathrm{d}\sigma(x).
\end{aligned}
$$

Now we calculate again explicitly and obtain in the limit $\delta\searrow 0$:

$$
\begin{aligned}
&\int_{\partial\mathbb{B}_1(y)}\psi(y)\partial_{x_j}\mathbf{G}_\kappa(x-y)\mathsf{e}_i\cdot\nu_{y,1}(x)\,\mathrm{d}\sigma(x)-\int_{\partial\mathbb{B}_\delta(y)}\psi(y)\partial_{x_j}\mathbf{G}_\kappa(x-y)\mathsf{e}_i\cdot\nu_{y,\delta}(x)\,\mathrm{d}\sigma(x)\\
&\xrightarrow{\delta\searrow 0}-\frac{1}{3}\delta_{ij}[\exp(-\sqrt{\kappa})(\sqrt{\kappa}+1)-1]\psi(y).
\end{aligned}
$$

In summary, one gets

$$
\begin{aligned}
&-\int_{\mathbb{R}^3}(\partial_j\psi)(x)(\partial_i(\mathbf{G}_\kappa*f))(x)\,\mathrm{d}x\\
&=\int_{\mathbb{R}^3}\psi(x)\left[\frac{1}{3}\delta_{ij}f(x)\exp(-\sqrt{\kappa})(\sqrt{\kappa}+1)+\int_{\mathbb{R}^3\setminus\mathbb{B}_1(x)}f(y)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y\right.\\
&\quad\left.+\lim_{\delta\to 0}\int_{\{1\ge|x-y|\ge\delta\}}(f(x)-f(y))\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y\right]\mathrm{d}x.
\end{aligned}
$$

From α-Hölder continuity of f, we conclude that, independently of δ,

$$
\mathbf{1}_{\{1\ge|x-y|\ge\delta\}}(y)\left|[f(x)-f(y)]\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\right|\le C|x-y|^{\alpha-3},
$$

which is integrable as $\alpha-3+2>-1$.

So, using again the dominated convergence theorem, we have

$$(\partial_i\partial_j v)(x) = \frac{1}{3}\delta_{ij}\exp(-\sqrt{\kappa})(\sqrt{\kappa}+1) + \int_{\mathbb{R}^3\setminus\mathbb{B}_1(x)} f(y)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y + \int_{\mathbb{B}_1(x)}[f(x)-f(y)]\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y \tag{II.2.45}$$

in the sense of distributions. In view of [121, Thm. 6.10], it remains to prove the (Hölder) continuity of the right-hand side in (II.2.45). Obviously, the first term in (II.2.45) is Hölder continuous. For the second term in (II.2.45), we obtain for all $x, z \in \mathbb{R}^3$, $x \neq z$:

$$\begin{aligned}&\left|\int_{\mathbb{R}^3\setminus\mathbb{B}_1(x)} f(y)\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y - \int_{\mathbb{R}^3\setminus\mathbb{B}_1(z)} f(y)\partial_{z_i}\partial_{z_j}\mathbf{G}_\kappa(z-y)\,\mathrm{d}y\right|\\ &= \left|\int_{\mathbb{B}_1(0)}[f(z-a)-f(x-a)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right|,\end{aligned}$$

by the transformation $a := x-y$ in the first and $a := z-y$ in the second integral. From α-Hölder continuity of f, we get the estimate

$$\left|\int_{\mathbb{B}_1(0)}[f(z-a)-f(x-a)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right| \leq C|x-z|^\alpha \int_{\mathbb{R}^3\setminus\mathbb{B}_1(x)} |\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a|,$$

where the integral on the r.h.s. is finite because $\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)$ behaves as $r^{-1}\exp(-r)$ for $r\to\infty$, which is integrable.

The same transformation yields for the third term in (II.2.45):

$$\begin{aligned}&\left|\int_{\mathbb{B}_1(x)}[f(x)-f(y)]\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)\,\mathrm{d}y - \int_{\mathbb{B}_1(z)}[f(z)-f(y)]\partial_{z_i}\partial_{z_j}\mathbf{G}_\kappa(z-y)\,\mathrm{d}y\right|\\ &= \left|\int_{\mathbb{B}_1(0)}[f(z)-f(z-a)-f(x)+f(x-a)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right|.\end{aligned}$$

We now proceed as in [121] and write $\mathbb{B}_1(0) = A\cup B$ with

$$\begin{aligned}A &:= \{a : 0 \leq |a| < 4|x-z|\},\\ B &:= \{a : 4|x-z| < |a| < 1\},\end{aligned}$$

where $B=\emptyset$ for $|x-z|\geq \frac{1}{4}$, and calculate, using that $|\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)|\leq C|a|^{-3}$:

$$\left|\int_A[f(z)-f(z-a)-f(x)+f(x-a)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right| \leq \int_A 2C|a|^{\alpha-3}\,\mathrm{d}a = \widetilde{C}|x-z|^\alpha.$$

It remains to consider the case $|x-z| < \frac{1}{4}$. One has, with the unit normal vector field ν on ∂B, that

$$\left|\int_B[f(z)-f(x)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right| = \left|\int_{\partial B}[f(z)-f(x)]\partial_{a_j}\mathbf{G}_\kappa(a)\mathrm{e}_i\cdot\nu(a)\,\mathrm{d}\sigma(a)\right|,$$

and by similar arguments as above,

$$\begin{aligned}&\left|\int_{\partial B}[f(z)-f(x)]\partial_{a_j}\mathbf{G}_\kappa(a)\mathrm{e}_i\cdot\nu(a)\,\mathrm{d}\sigma(a)\right|\\ &= \frac{1}{3}\delta_{ij}|f(z)-f(x)||\exp(-\sqrt{\kappa})(\sqrt{\kappa}+1)-\exp(-4\sqrt{\kappa}|x-z|)(4\sqrt{\kappa}|x-z|+1)|.\end{aligned}$$

Note that the real-valued map $[0,\infty)\ni r\mapsto \exp(-\sqrt{\kappa}r)(\sqrt{\kappa}r+1)$ is monotonically decreasing. This yields

$$\left|\int_B[f(z)-f(x)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right| \leq C|z-x|^\alpha.$$

By the transformations $b := x - a - z$ for the first and $b := -a$ for the second term, we get

$$\begin{aligned}&\left|\int_B [f(x-a) - f(z-a)]\partial_{a_i}\partial_{a_j}\mathbf{G}_\kappa(a)\,\mathrm{d}a\right| \\ &= \left|\int_B f(z+b)\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b)\,\mathrm{d}b - \int_D f(b+z)\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b-x+z)\,\mathrm{d}b\right|,\end{aligned} \tag{II.2.46}$$

with $D := \{b : 4|x-z| < |b-x+z| < 1\}$.

Note that

$$\int_B \partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b)\,\mathrm{d}b = \int_D \partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b-x+z)\,\mathrm{d}b.$$

This enables us to rewrite (II.2.46) as follows:

$$\begin{aligned}&\left|\int_B f(z+b)\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b)\,\mathrm{d}b - \int_D f(b+z)\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b-x+z)\,\mathrm{d}b\right| \\ &= \left|\int_B [f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b)\,\mathrm{d}b - \int_D [f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b-x+z)\,\mathrm{d}b\right|.\end{aligned} \tag{II.2.47}$$

We consider (II.2.47) separately on the sets $B \cap D$, $B\backslash D$ and $D\backslash B$.

Note that, by the triangular inequality, $B \cap D \subset \{b : 3|x-z| < |b| < 1 + |x-z|\}$ and by Taylor's theorem

$$(\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa)(b) - (\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa)(b-x+z) = \sum_{k=1}^{3}(\partial_k\partial_i\partial_j\mathbf{G}_\kappa)(b^*)(x_k - z_k),$$

for some $b^* = b - \beta(x-z)$ with $\beta \in (0,1)$. Therefore, one has by the triangular inequality that $|b^*| \geq |b| - \beta|x-z| \geq (1 - \frac{\beta}{3})|b| \geq \frac{2}{3}|b|$ on $B \cap D$ and consequently

$$|(\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa)(b) - (\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa)(b-x+z)| \leq C|b^*|^{-4}|x-z| \leq \widetilde{C}|b|^{-4}|x-z|.$$

This allows us to estimate

$$\begin{aligned}&\left|\int_{B\cap D}[f(z+b)-f(z)]\left[\partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b) - \partial_{b_i}\partial_{b_j}\mathbf{G}_\kappa(b-x+z)\right]\,\mathrm{d}b\right| \\ &\leq C|x-z|\int_{3|x-z|}^{1+|x-z|} r^{-4+\alpha+2}\,\mathrm{d}r \\ &\leq \frac{C|x-z|}{1-\alpha}\left[(3|x-z|)^{\alpha-1} - (1+|x-z|)^{\alpha-1}\right] \leq \frac{\widetilde{C}}{1-\alpha}|x-z|^\alpha.\end{aligned}$$

For the remaining terms, we split up as in [121]:

$$\begin{aligned}B\backslash D &\subset E \cup G, \\ D\backslash B &\subset E' \cup G',\end{aligned}$$

where

$$\begin{aligned}E &:= \{b : 4|x-z| < |b| \leq 5|x-z|\}, \\ G &:= \{b : 1 - |x-z| \leq |b| < 1\}, \\ E' &:= \{b : 4|x-z| < |b-x+z| \leq 5|x-z|\}, \\ G' &:= \{b : 1 - |x-z| \leq |b-x+z| < 1\}.\end{aligned}$$

Consider at first the real-valued map $[0, \frac{1}{4}] \ni s \mapsto (1-s)^\beta$ for arbitrary $\beta > 0$. Obviously, it is continuously differentiable and therefore α-Hölder continuous because its domain of definition is compact. Hence, the following holds for all $0 \leq s \leq \frac{1}{4}$:

$$1-(1-s)^{\beta}=(1-0)^{\beta}-(1-s)^{\beta}\leq Cs^{\alpha}. \tag{II.2.48}$$

Now, we estimate the integral on $B\backslash D$, where we use again the estimate $|\partial_{a_i}\partial_{a_j}\mathbf{G}_{\kappa}(a)|\leq C|a|^{-3}$:

$$\begin{aligned}
&\left|\int_{B\backslash D}[f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_{\kappa}(b)\,\mathrm{d}b\right|\leq C\left(\int_{4|x-z|}^{5|x-z|}r^{\alpha-3+2}\,\mathrm{d}r+\int_{1-|x-z|}^{1}r^{\alpha-3+2}\,\mathrm{d}r\right)\\
&=\frac{C}{\alpha}\left[(5|x-z|)^{\alpha}-(4|x-z|)^{\alpha}+1-(1-|x-z|)^{\alpha}\right]\leq\frac{C}{\alpha}(5^{\alpha}+\widetilde{C})|x-z|^{\alpha},
\end{aligned}$$

where we have used (II.2.48) for $\beta:=\alpha$ in the last step.

For the remaining integral on $D\backslash B$, we consider the domains E' and G' separately and note at first that, using the triangular inequality, $E'\subset\{0<|b|\leq 6|x-z|\}$. Subsequently, this yields that $|b-x+z|^{-3}<(4|x-z|)^{-3}\leq C|b|^{-3}$ on E'. Hence, by the estimate $|\partial_{a_i}\partial_{a_j}\mathbf{G}_{\kappa}(a)|\leq C|a|^{-3}$, the following holds:

$$\int_{E'}\left|[f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_{\kappa}(b-x+z)\right|\,\mathrm{d}b\leq C\int_0^{6|x-z|}r^{\alpha-3+2}\,\mathrm{d}r=\widetilde{C}|x-z|^{\alpha}.$$

On G', one has $|b-x+z|\geq 1-|x-z|>\frac{3}{4}$. Consequently, it holds that

$$\begin{aligned}
&\int_{G'}\left|[f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_{\kappa}(b-x+z)\right|\,\mathrm{d}b\leq C\left(\frac{3}{4}\right)^{-3}\int_{1-|x-z|}^{1}r^{\alpha+2}\,\mathrm{d}r\\
&=\widetilde{C}\left[1-(1-|x-z|)^{3+\alpha}\right]\leq\widetilde{C}|x-z|^{\alpha},
\end{aligned}$$

where we have used (II.2.48) for $\beta:=3+\alpha$ in the last step. Together,

$$\left|\int_{D\backslash B}[f(z+b)-f(z)]\partial_{b_i}\partial_{b_j}\mathbf{G}_{\kappa}(b-x+z)\,\mathrm{d}b\right|\leq C|x-z|^{\alpha},$$

and the assertion is proved. □

II.2.2.2.3. *Properties of the steady state*

Now, we are in position to prove several estimates on the stationary solution (u_∞,v_∞):

Proposition II.26 (Estimates on the stationary solution). *The following uniform estimates hold for all $x\in\mathbb{R}^3$:*

(a) $u_\infty(x)\leq U_0-\varepsilon V\phi'(0)$, where $U_0\in\mathbb{R}$ is chosen in such a way that $\int_{\mathbb{R}^3}[U_0-W]_+\,\mathrm{d}x=1$ and $V>0$ is the constant from Proposition II.23.

(b) $|\mathrm{D}v_\infty(x)|\leq C\varepsilon$ for some constant $C>0$.

(c) $-C'\varepsilon\mathbb{1}\leq \mathrm{D}^2v_\infty(x)\leq C'\varepsilon\mathbb{1}$ in the sense of symmetric matrices, for some constant $C'>0$.

Proof. (a) We first prove that $U_\varepsilon\leq U_0+\varepsilon\phi(0)$, which in turn follows if

$$\int_{\mathbb{R}^3}[U_0+\varepsilon\phi(0)-W-\varepsilon\phi(v_\infty)]_+\,\mathrm{d}x\geq 1.$$

One has

$$\begin{aligned}
\int_{\mathbb{R}^3}[U_0+\varepsilon\phi(0)-W-\varepsilon\phi(v_\infty)]_+\,\mathrm{d}x&=\int_{\{U_0-W\geq 0\}}[U_0-W+\varepsilon(\phi(0)-\phi(v_\infty))]\,\mathrm{d}x\\
&+\int_{\{0>U_0-W\geq\varepsilon(\phi(v_\infty)-\phi(0))\}}[U_0-W+\varepsilon(\phi(0)-\phi(v_\infty))]\,\mathrm{d}x.
\end{aligned} \tag{II.2.49}$$

From $\phi(0)-\phi(v_\infty)\geq 0$ (recall that v_∞ is strictly positive and ϕ is decreasing) and the definition of U_0, we deduce that the first term on the r.h.s. of (II.2.49) is larger or equal to 1. The second term on the r.h.s. of (II.2.49) is nonnegative because the integrand is nonnegative on the domain of integration.

Now, if $u_\infty(x) > 0$ for some $x \in \mathbb{R}^3$, we also have thanks to convexity of ϕ:

$$u_\infty(x) \le U_\varepsilon - W(x) - \varepsilon\phi(0) - \varepsilon v_\infty(x)\phi'(0) \le U_0 + \varepsilon\phi(0) - \varepsilon\phi(0) - \varepsilon V\phi'(0),$$

from which the desired estimate follows.

(b) Define

$$f_v : \mathbb{R}^3 \to \mathbb{R}, \quad f_v(x) := \varepsilon[U_\varepsilon - W(x) - \varepsilon\phi(v(x))]_+\phi'(v(x)).$$

Then, $f_v \in L^\infty(\mathbb{R}^3)$ with compact support $\operatorname{supp}(f_v) \subset \mathbb{B}_R(0)$ where $R > 0$ can be chosen independently of $\varepsilon \in (0,1)$. Moreover, by Lemma II.24(a), (u_∞, v_∞) is the solution to the integral equation

$$v = -(\mathbf{G}_\kappa * f_v),$$

with the Yukawa potential $\mathbf{G}_\kappa$ defined in (II.2.39). Since $W^{2,4}(\mathbb{R}^3)$ is continuously embedded in $C^1(\mathbb{R}^3)$ [183, Appendix, sec. (45) et seq.] and $f_v \in L^4(\mathbb{R}^3)$, we deduce from Lemma II.24(b) that

$$\|v\|_{C^1} \le \widetilde{C}\|f_v\|_{L^4},$$

for some constant $\widetilde{C} > 0$. Hence, we obtain (b) by using (a):

$$\|\mathrm{D}v_\infty\|_{L^\infty} \le \widetilde{C}\|f_{v_\infty}\|_{L^4} \le \widetilde{C}\varepsilon(U_0 - \varepsilon V\phi'(0))|\phi'(0)||\mathbb{B}_R(0)|^{1/4} =: C\varepsilon.$$

(c) First, consider $x \in \mathbb{R}^3 \backslash \mathbb{B}_{R+1}(0)$, where $R > 0$ is such that $\operatorname{supp}(f_{v_\infty}) \subset \mathbb{B}_R(0)$. Smoothness of $\mathbf{G}_\kappa$ on $\mathbb{R}^3 \backslash \{0\}$ yields for all $i, j \in \{1,2,3\}$:

$$\begin{aligned} |\partial_i\partial_j v_\infty(x)| &= \left| \int_{\mathbb{B}_R(0)} (\partial_{x_i}\partial_{x_j}\mathbf{G}_\kappa(x-y)) f_{v_\infty}(y)\,\mathrm{d}y \right| \\ &= \left| \int_{\mathbb{B}_R(x)} (\partial_i\partial_j\mathbf{G}_\kappa(z)) f_{v_\infty}(x-z)\,\mathrm{d}z \right|, \end{aligned}$$

where the last equality follows by the transformation $z := x - y$. Since $\mathbb{B}_R(x) \subset \mathbb{R}^3 \setminus \mathbb{B}_1(0)$ and

$$|\partial_i\partial_j\mathbf{G}_\kappa(z)| \le \frac{C(\kappa)\exp(-\sqrt{\kappa}|z|)}{4\pi|z|} \le \frac{C'(\kappa)}{4\pi} \quad \text{for } |z| \ge 1,$$

we obtain the estimate

$$|\partial_i\partial_j v_\infty(x)| \le \frac{1}{3}R^3 C'(\kappa)\|f_{v_\infty}\|_{L^\infty} < \infty.$$

Consider now the case $|x| \le R+1$ and set $y := (R+2)\mathrm{e}_1 \ne x$. By the triangular inequality, we have for $\alpha \in (0,1)$ that

$$|\partial_i\partial_j v_\infty(x)| \le |\partial_i\partial_j v_\infty(y)| + \frac{|\partial_i\partial_j v_\infty(x) - \partial_i\partial_j v_\infty(y)|}{|x-y|^\alpha}|x-y|^\alpha.$$

By the arguments above, f_{v_∞} is α-Hölder continuous for some $\alpha \in (0,1)$ since u_∞ is Lipschitz continuous and of compact support. By Lemma II.25, we know that there exists $C > 0$ such that

$$[\partial_i\partial_j v_\infty]_{C^{0,\alpha}} \le C[f_{v_\infty}]_{C^{0,\alpha}}.$$

Hence, since $|x - y| \le 2R + 3$, one has

$$|\partial_i\partial_j v_\infty(x)| \le |\partial_i\partial_j v_\infty(y)| + C(2R+3)^\alpha[f_{v_\infty}]_{C^{0,\alpha}}.$$

Combining both cases yields

$$\begin{aligned} |\partial_i\partial_j v_\infty(x)| &\le |\partial_i\partial_j v_\infty((R+2)\mathrm{e}_1)| + C(2R+3)^\alpha[f_{v_\infty}]_{C^{0,\alpha}} \\ &\le C_0\|f_{v_\infty}\|_{L^\infty} + C_1\|f_{v_\infty}\|_{W^{1,\infty}}, \end{aligned}$$

for some $C_0, C_1 > 0$ and *all* $x \in \mathbb{R}^3$. Using (a) and (b), it is straightforward to conclude that there exists $C_2 > 0$ with

$$(\|\mathrm{D}f_{v_\infty}\|_{L^\infty} + \|f_{v_\infty}\|_{L^\infty}) \le C_2\varepsilon.$$

All in all, we proved the existence of $C_3 > 0$ such that for all $x \in \mathbb{R}^3$ and all $i, j \in \{1, 2, 3\}$:

$$|\partial_i\partial_j v_\infty(x)| \le C_3\varepsilon.$$

Obviously, this estimate yields the assertion (for a different constant $C' > 0$).

□

The strategy of proof of Theorem II.4 now is as follows: We first decompose the energy $\mathcal{E}$ into a decoupled, uniformly geodesically convex part $\mathcal{L}$ and into a remaining part $\mathcal{L}_*$, which has no useful convexity properties, but can nevertheless be controlled. The convex part $\mathcal{L}$ serves as an auxiliary functional for the flow interchange lemma (Theorem I.5). On a smooth and very formal level, the decomposition $\mathcal{E} = \mathcal{L} + \varepsilon\mathcal{L}_*$ (see formula (II.2.52) below) leads us to

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L} = \nabla\mathcal{L}\cdot\nabla\mathcal{E} = \|\nabla\mathcal{L}\|^2 + \varepsilon\nabla\mathcal{L}\cdot\nabla\mathcal{L}_* \ge \left(1-\frac{\varepsilon}{2}\right)\|\nabla\mathcal{L}\|^2 - \frac{\varepsilon}{2}\|\nabla\mathcal{L}_*\|^2 \ge 2\lambda_\varepsilon\left(1-\frac{\varepsilon}{2}\right)\mathcal{L} - \frac{\varepsilon}{2}\|\nabla\mathcal{L}_*\|^2,$$

using Young's inequality and the λ_ε-convexity along geodesics of $\mathcal{L}$, see Proposition II.28 below. Thus, we are almost in the situation to apply Gronwall's lemma to $\mathcal{L}$, apart from the last term involving the formal gradient of the nonconvex part $\mathcal{L}_*$ which has to be controlled in a suitable manner depending on the problem at hand (cf. also Chapters II.3 and II.4).

We seek to eventually arrive at an exponential estimate for $\mathcal{L}$ along the discrete solution curves constructed via the minimizing movement scheme. The dissipation of the driving energy $\mathcal{E}$ along the gradient flow associated to the auxiliary energy $\mathcal{L}$ first gives us an estimate guaranteeing boundedness of $\mathcal{L}$ for large times. There, one of the key observations is that the gradient of the v-component along the discrete solution is uniformly bounded in a certain function space. Subsequently, the boundedness of $\mathcal{L}$ allows us to prove a revised dissipation estimate which then yields the desired exponential estimate for sufficiently large times. The proof of Theorem II.4 is completed by passage to the continuous-time limit.

Since our claim only concerns the solutions (u, v) to (II.1.1) that are constructed by the minimizing movement scheme, we assume in the following that we are given a family of time-discrete approximations $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$ that converge to the weak solution (u, v) as indicated in Theorem II.22 as $\tau \searrow 0$. Therefore, we may assume without loss of generality that $\tau > 0$ is sufficiently small.

Throughout this section, we shall use the abbreviation $[a]_\tau := \frac{1}{\tau}\log(1 + a\tau)$, where $a > 0$. Note that, for every $\tau > 0$ and a family of indices $m_\tau \in \mathbb{N}$ given such that $m_\tau\tau \ge T$ with a fixed $T \ge 0$, one has

$$(1 + a\tau)^{-m_\tau} \le e^{-[a]_\tau T} \downarrow e^{-aT} \quad \text{as } \tau \downarrow 0. \tag{II.2.50}$$

In order to keep track of the dependencies of certain quantities on ε, we are going to define several positive numbers ε_j such that the estimates in a certain proof are uniform with respect to $\varepsilon \in (0, \varepsilon_j)$. When we want to emphasize that a quantity is independent of $\varepsilon \in (0, \varepsilon_j)$ — and also of τ and the initial condition (u_0, v_0) — we call it a *system constant*. System constants are expressible in terms of λ_0, κ, ϕ and truely universal constants. For brevity, we write $\mathcal{E}_\infty := \mathcal{E}(u_\infty, v_\infty)$.

II.2.2.3. Decomposition of the energy

The key element in the proof of Theorem II.4 is a decomposition of the energy functional. Introduce the *perturbed potential* W_ε by

$$W_\varepsilon(x) := W(x) + \varepsilon\phi(v_\infty(x)). \tag{II.2.51}$$

Recall that (u_∞, v_∞) is the minimizer of $\mathcal{E}$ on $\mathbf{X}$, and define

$$\mathcal{L}_u(u) := \int_{\mathbb{R}^3} \Big(\frac{1}{2}(u^2 - u_\infty^2) + W_\varepsilon(u - u_\infty)\Big)\,\mathrm{d}x,$$

$$\mathcal{L}_v(v) := \int_{\mathbb{R}^3} \frac{1}{2}\big(|\mathrm{D}(v - v_\infty)|^2 + \kappa(v - v_\infty)^2\big)\,\mathrm{d}x,$$

$$\mathcal{L}_*(u,v) := \int_{\mathbb{R}^3} \big(u[\phi(v) - \phi(v_\infty)] - u_\infty\phi'(v_\infty)[v - v_\infty]\big)\,\mathrm{d}x.$$

Finally, let $\mathcal{L}(u,v) := \mathcal{L}_u(u) + \mathcal{L}_v(v)$ denote the *auxiliary entropy*.

Lemma II.27. *The following decomposition holds:*

$$\mathcal{E}(u,v) - \mathcal{E}_\infty = \mathcal{L}(u,v) + \varepsilon\mathcal{L}_*(u,v). \tag{II.2.52}$$

Proof. By the properties of ϕ and the fact that u_∞ has compact support, $\mathcal{L}_*$ is well-defined on all of $\mathbf{X}$, while $\mathcal{L}_u$ and $\mathcal{L}_v$ are finite precisely on $(\mathscr{P}_2 \cap L^2)(\mathbb{R}^3)$ and $W^{1,2}(\mathbb{R}^3)$, respectively. Thus, both sides in (II.2.52) are finite on the same subset of $\mathbf{X}$. Now, for every such pair (u,v), we have on the one hand that

$$\mathcal{L}_u(u) = \int_{\mathbb{R}^3} \Big(\frac{1}{2}u^2 + uW + \varepsilon u\phi(v_\infty)\Big)\,\mathrm{d}x - \int_{\mathbb{R}^3} \Big(\frac{1}{2}u_\infty^2 + u_\infty W + \varepsilon u_\infty\phi(v_\infty)\Big)\,\mathrm{d}x, \tag{II.2.53}$$

and on the other hand that

$$\mathcal{L}_v(v) = \int_{\mathbb{R}^3} \Big(\frac{1}{2}|\mathrm{D}v|^2 + \frac{\kappa}{2}v^2\Big)\,\mathrm{d}x + \int_{\mathbb{R}^3} \Big(\frac{1}{2}|\mathrm{D}v_\infty|^2 + \frac{\kappa}{2}v_\infty^2\Big)\,\mathrm{d}x - \int_{\mathbb{R}^3} (\mathrm{D}v\cdot\mathrm{D}v_\infty + \kappa v v_\infty)\,\mathrm{d}x.$$

Integration by parts in the last integral yields, recalling the defining equation (II.2.35) for v_∞, that

$$-\int_{\mathbb{R}^3} (\mathrm{D}v\cdot\mathrm{D}v_\infty + \kappa v v_\infty)\,\mathrm{d}x = \int_{\mathbb{R}^3} (\Delta v_\infty - \kappa v_\infty)v\,\mathrm{d}x = \varepsilon\int_{\mathbb{R}^3} u_\infty\phi'(v_\infty)v\,\mathrm{d}x.$$

Similarly, integration by parts in the middle integral leads to

$$\int_{\mathbb{R}^3} \Big(\frac{1}{2}|\mathrm{D}v_\infty|^2 + \frac{\kappa}{2}v_\infty^2\Big)\,\mathrm{d}x = -\frac{\varepsilon}{2}\int_{\mathbb{R}^3} u_\infty\phi'(v_\infty)v_\infty\,\mathrm{d}x.$$

And so,

$$\mathcal{L}_v(v) = \int_{\mathbb{R}^3} \Big(\frac{1}{2}|\mathrm{D}v|^2 + \frac{\kappa}{2}v^2\Big)\,\mathrm{d}x - \int_{\mathbb{R}^3} \Big(\frac{1}{2}|\mathrm{D}v_\infty|^2 + \frac{\kappa}{2}v_\infty^2\Big)\,\mathrm{d}x + \varepsilon\int_{\mathbb{R}^3} u_\infty\phi'(v_\infty)(v - v_\infty)\,\mathrm{d}x. \tag{II.2.54}$$

Combining (II.2.53) and (II.2.54) with the definition of $\mathcal{L}_*$ yields (II.2.52). □

We summarize some useful properties of the auxiliary entropy $\mathcal{L}$ in the following.

Proposition II.28 (Properties of $\mathcal{L}$). *There are constants $K, L > 0$ and some $\varepsilon_0 > 0$ such that the following is true for every $\varepsilon \in (0, \varepsilon_0)$:*

(a) $W_\varepsilon \in C^2(\mathbb{R}^3)$ is λ_ε-convex with $\lambda_\varepsilon := \lambda_0 - L\varepsilon > 0$.

(b) $\mathcal{L}_u$ is λ_ε-convex along generalized geodesics in $(\mathscr{P}_2(\mathbb{R}^3), \mathbf{W}_2)$, and for every $u \in (\mathscr{P}_2 \cap W^{1,2})(\mathbb{R}^3)$, one has

$$\frac{1}{2}\|u - u_\infty\|_{L^2}^2 \le \mathcal{L}_u(u) \le \frac{1}{2\lambda_\varepsilon}\int_{\mathbb{R}^3} u|\mathrm{D}(u + W_\varepsilon)|^2\,\mathrm{d}x. \tag{II.2.55}$$

(c) $\mathcal{L}_v$ is κ-convex in $L^2(\mathbb{R}^3)$, and for every $v \in W^{2,2}(\mathbb{R}^3)$, one has

$$\frac{\kappa}{2}\|v - v_\infty\|_{L^2}^2 \le \mathcal{L}_v(v) \le \frac{1}{2\kappa}\int_{\mathbb{R}^3} \big(\Delta(v - v_\infty) - \kappa(v - v_\infty)\big)^2\,\mathrm{d}x. \tag{II.2.56}$$

(d) For every $(u,v) \in \mathbf{X}$,

$$\mathcal{L}(u,v) \le (1 + K\varepsilon)\big(\mathcal{E}(u,v) - \mathcal{E}_\infty\big). \tag{II.2.57}$$

Proof. (a) Since $W_\varepsilon = W + \varepsilon\phi(v_\infty)$, the chain rule yields

$$\mathrm{D}^2W_\varepsilon = \mathrm{D}^2W + \varepsilon\phi''(v_\infty)\mathrm{D}v_\infty \otimes \mathrm{D}v_\infty + \varepsilon\phi'(v_\infty)\mathrm{D}^2v_\infty.$$

Using our assumptions on ϕ and by Proposition II.26, there are some $L > 0$ and some ε_0 such that

$$\phi''(v_\infty)\mathrm{D}v_\infty \otimes \mathrm{D}v_\infty + \phi'(v_\infty)\mathrm{D}^2 v_\infty \geq -L\mathbb{1}$$

holds uniformly with respect to $\varepsilon \in (0, \varepsilon_0)$. And thus also $\mathrm{D}^2 W_\varepsilon \geq \lambda_\varepsilon \mathbb{1}$, with the indicated definition of λ_ε.

(b) Since W_ε is λ_ε-convex, also $\mathcal{L}_u$ is geodesically λ_ε-convex with respect to $\mathbf{W}_2$ because it is the sum of a geodesically 0-convex functional and a geodesically λ_ε-convex functional, see Theorem I.14.

The Wasserstein subdifferential of $\mathcal{L}_u$ has been calculated in [4, Lemma 10.4.1]. Together with (I.2.9), this shows the second inequality in (II.2.55). Concerning the first inequality, observe that

$$\mathcal{L}_u(u) = \frac{1}{2}\int_{\mathbb{R}^3} (u - u_\infty)^2 \,\mathrm{d}x + \int_{\mathbb{R}^3} (W_\varepsilon + u_\infty)(u - u_\infty)\,\mathrm{d}x.$$

It thus suffices to prove nonnegativity of the second integral term for all $u \in \mathscr{P}_2(\mathbb{R}^3)$. First, as u and u_∞ have equal mass, and by the definition of u_∞,

$$0 = \int_{\mathbb{R}^3} (u_\infty - u)\,\mathrm{d}x = \int_{\{U_\varepsilon - W_\varepsilon > 0\}} u_\infty \,\mathrm{d}x - \int_{\mathbb{R}^3} u\,\mathrm{d}x,$$

and consequently

$$\int_{\{U_\varepsilon - W_\varepsilon > 0\}} (u - u_\infty)\,\mathrm{d}x = -\int_{\{U_\varepsilon - W_\varepsilon \leq 0\}} u\,\mathrm{d}x. \tag{II.2.58}$$

Also, by definition of u_∞,

$$\int_{\mathbb{R}^3} (W_\varepsilon + u_\infty)(u - u_\infty)\,\mathrm{d}x = \int_{\{U_\varepsilon - W_\varepsilon > 0\}} U_\varepsilon (u - u_\infty)\,\mathrm{d}x + \int_{\{U_\varepsilon - W_\varepsilon \leq 0\}} W_\varepsilon u\,\mathrm{d}x.$$

Combining this with (II.2.58) yields

$$\int_{\{U_\varepsilon - W_\varepsilon > 0\}} U_\varepsilon (u - u_\infty)\,\mathrm{d}x + \int_{\{U_\varepsilon - W_\varepsilon \leq 0\}} W_\varepsilon u\,\mathrm{d}x = \int_{\{U_\varepsilon - W_\varepsilon \leq 0\}} (W_\varepsilon - U_\varepsilon) u\,\mathrm{d}x \geq 0,$$

as the integrand is nonnegative on the domain of integration.

(c) This is an immediate consequence of (I.2.9) for the L^2 subdifferential of $\mathcal{L}_v$.

(d) Since ϕ is convex, we have

$$\phi(v) - \phi(v_\infty) - \phi'(v_\infty)[v - v_\infty] \geq 0,$$

and so we can estimate $\mathcal{L}_*$ from below as follows:

$$\begin{aligned}
\mathcal{L}_*(u, v) &= \int_{\mathbb{R}^3} (u - u_\infty)[\phi(v) - \phi(v_\infty)]\,\mathrm{d}x + \int_{\mathbb{R}^3} \big(\phi(v) - \phi(v_\infty) - \phi'(v_\infty)[v - v_\infty]\big)\,\mathrm{d}x \\
&\geq -\frac{1}{2}\int_{\mathbb{R}^3} (u - u_\infty)^2\,\mathrm{d}x - \frac{\phi'(0)^2}{2}\int_{\mathbb{R}^3} (v - v_\infty)^2\,\mathrm{d}x \\
&\geq -\mathcal{L}_u(u) - \frac{\phi'(0)^2}{\kappa}\mathcal{L}_v(v),
\end{aligned}$$

using the properties (b) and (c) above. By (II.2.52), we conclude

$$(1 - K'\varepsilon)\mathcal{L}(u, v) = \mathcal{E}(u, v) - \mathcal{E}_\infty \quad \text{with} \quad K' := \max\Big(1, \frac{\phi'(0)^2}{\kappa}\Big),$$

which clearly implies (II.2.57) for all $\varepsilon \in (0, \varepsilon_0)$, possibly after diminishing ε_0.

□

II.2.2.4. Dissipation

We can now formulate the main *a priori* estimate for the time-discrete solution.

Proposition II.29. *Given $(\widetilde{u},\widetilde{v}) \in \mathbf{X}$ with $\mathcal{E}(\widetilde{u},\widetilde{v}) < \infty$, let $(u,v) \in \mathbf{X}$ be a minimizer of the functional $\mathcal{E}_\tau(\,\cdot\,|\widetilde{u},\widetilde{v})$ introduced in* (II.2.5). *Then*

$$\mathcal{L}_u(u) + \tau\mathcal{D}_u(u,v) \leq \mathcal{L}_u(\widetilde{u}) \quad \text{and} \quad \mathcal{L}_v(v) + \tau\mathcal{D}_v(u,v) \leq \mathcal{L}_v(\widetilde{v}), \tag{II.2.59}$$

where the dissipation terms are given by

$$\mathcal{D}_u(u,v) = \Big(1-\frac{\varepsilon}{2}\Big)\int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}^3} u\big|\mathrm{D}\big(\phi(v)-\phi(v_\infty)\big)\big|^2\,\mathrm{d}x, \tag{II.2.60}$$

$$\mathcal{D}_v(u,v) = \Big(1-\frac{\varepsilon}{2}\Big)\int_{\mathbb{R}^3} \big(\Delta(v-v_\infty)-\kappa(v-v_\infty)\big)^2\,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}^3} \big(u\phi'(v)-u_\infty\phi'(v_\infty)\big)^2\,\mathrm{d}x. \tag{II.2.61}$$

Proof. Naturally, these estimates are derived with the flow interchange lemma (Theorem I.5). For given $\nu > 0$, introduce the regularized functional $\mathcal{L}_u^\nu = \mathcal{L}_u + \nu\mathcal{H}$, where

$$\mathcal{H}(u) = \int_{\mathbb{R}^3} u\log u\,\mathrm{d}x.$$

Recall from Lemma II.18 that $\mathcal{H}$ is finite on $(\mathscr{P}_2\cap L^2)(\mathbb{R}^3)$. Moreover, $\mathcal{L}_u^\nu$ is λ_ε-convex along generalized geodesics in $\mathbf{W}_2$ by Theorem I.14. We claim that the λ_ε-flow associated to $\mathcal{L}_u^\nu$ satisfies the evolution equation

$$\partial_s\mathcal{U} = \nu\Delta\mathcal{U} + \frac{1}{2}\Delta\mathcal{U}^2 + \operatorname{div}(\mathcal{U}\mathrm{D}W_\varepsilon). \tag{II.2.62}$$

Since $\nu > 0$, this equation is strictly parabolic. Therefore, for every initial condition $\mathcal{U}_0 \in (\mathscr{P}_2\cap L^2)(\mathbb{R}^3)$, there exists a smooth and positive solution $\mathcal{U} : (0,\infty)\times\mathbb{R}^3 \to \mathbb{R}$ such that $\mathcal{U}(s,\cdot) \to \mathcal{U}_0$ both in $\mathbf{W}_2$ and in $L^2(\mathbb{R}^3)$ as $s \searrow 0$. By [4, Theorem 11.2.8], the solution operator to (II.2.62) can be identified with the λ_ε-flow of $\mathcal{L}_u^\nu$. Now, let $\mathcal{U}$ be the smooth solution to (II.2.62) with initial condition $\mathcal{U}_0 = u$. By smoothness of $\mathcal{U}$, the equation (II.2.62) is satisfied in the classical sense at every time $s > 0$, and the following integration by parts is justified:

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},v) &= -\int_{\mathbb{R}^3}\big[\mathcal{U}+W_\varepsilon+\varepsilon(\phi(v)-\phi(v_\infty))\big]\operatorname{div}\big[\mathcal{U}\mathrm{D}(\mathcal{U}+W_\varepsilon)+\nu\mathrm{D}\mathcal{U}\big]\,\mathrm{d}x \\
&= \int_{\mathbb{R}^3}\mathcal{U}|\mathrm{D}(\mathcal{U}+W_\varepsilon)|^2\,\mathrm{d}x + \varepsilon\int_{\mathbb{R}^3}\mathcal{U}\mathrm{D}(\phi(v)-\phi(v_\infty))\cdot\mathrm{D}(\mathcal{U}+W_\varepsilon)\,\mathrm{d}x \\
&\quad + \nu\int_{\mathbb{R}^3}\mathrm{D}\big[\mathcal{U}+W+\varepsilon\phi(v)\big]\cdot\mathrm{D}\mathcal{U}\,\mathrm{d}x.
\end{aligned}$$

For the last integral, one has by integration by parts, the chain rule and Young's inequality:

$$\begin{aligned}
\int_{\mathbb{R}^3}\mathrm{D}\big[\mathcal{U}+W+\varepsilon\phi(v)\big]\cdot\mathrm{D}\mathcal{U}\,\mathrm{d}x &= \|\mathrm{D}\mathcal{U}\|_{L^2}^2 + \int_{\mathbb{R}^3}\big(-\mathcal{U}\Delta W+\varepsilon\phi'(v)\mathrm{D}\mathcal{U}\cdot\mathrm{D}v\big)\,\mathrm{d}x \\
&\geq -\|\Delta W\|_{L^\infty} - \frac{1}{2}\varepsilon^2\overline{\phi'}^2\|v\|_{W^{1,2}}^2.
\end{aligned}$$

Rewriting the middle integral with Young's inequality, we arrive at

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},v) \geq \Big(1-\frac{\varepsilon}{2}\Big)\int_{\mathbb{R}^3}\mathcal{U}|\mathrm{D}(\mathcal{U}+W_\varepsilon)|^2\,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}^3}\mathcal{U}\big|\mathrm{D}\big(\phi(v)-\phi(v_\infty)\big)\big|^2\,\mathrm{d}x - \nu C\big(1+\|v\|_{W^{1,2}}^2\big),$$

for some $C > 0$. We pass to the limit $s \searrow 0$. Recall that $\mathcal{U}$ converges (strongly) to its initial datum $\mathcal{U}_0 = u$ in $L^2(\mathbb{R}^3)$, and observe that the expressions on the right-hand side are lower semicontinuous with respect to that convergence. In fact, this is clear except perhaps for the first integral, which however can be rewritten, using integration by parts, in the form

$$\int_{\mathbb{R}^3} \mathcal{U}|\mathrm{D}(\mathcal{U}+W_\varepsilon)|^2\,\mathrm{d}x = \frac{4}{9}\int_{\mathbb{R}^3} |\mathrm{D}\mathcal{U}^{3/2}|^2\,\mathrm{d}x - \int_{\mathbb{R}^3} \mathcal{U}^2\Delta W_\varepsilon\,\mathrm{d}x + \int_{\mathbb{R}^3} \mathcal{U}|\nabla W_\varepsilon|^2\,\mathrm{d}x,$$

in which the lower semicontinuity is obvious since $\Delta W_\varepsilon \in L^\infty(\mathbb{R}^3)$. Applying now Theorem I.5, we arrive at

$$\mathcal{L}_u^\nu(u) + (1-\varepsilon)\int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}^3} u\big|\mathrm{D}\big(\phi(v)-\phi(v_\infty)\big)\big|^2\,\mathrm{d}x \le \mathcal{L}_u^\nu(\widetilde{u}) + \nu C\big(1+\|v\|_{W^{1,2}}^2\big).$$

Finally, passage to the limit $\nu \searrow 0$ yields the dissipation (II.2.60).

The dissipation (II.2.61) is easier to obtain. It is immediate that the κ-flow in $L^2(\mathbb{R}^3)$ of $\mathcal{L}_v$ satisfies the linear parabolic equation

$$\partial_s\mathcal{V} = \Delta(\mathcal{V}-v_\infty) - \kappa(\mathcal{V}-v_\infty). \tag{II.2.63}$$

Solutions $\mathcal{V}$ to (II.2.63) exist for arbitrary initial conditions $\mathcal{V}_0 \in L^2(\mathbb{R}^3)$, and they have at least the spatial regularity of v_∞. Hence, with $\mathcal{V}_0 := v$, we have, also recalling the defining equation (II.2.35) for v_∞,

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(u,\mathcal{V}) = \int_{\mathbb{R}^3}\big[\Delta(\mathcal{V}-v_\infty) - \kappa(\mathcal{V}-v_\infty) - \varepsilon(u\phi'(\mathcal{V}) - u_\infty\phi'(v_\infty))\big]\big[\Delta(\mathcal{V}-v_\infty) - \kappa(\mathcal{V}-v_\infty)\big]\,\mathrm{d}x.$$

Another application of Young's inequality yields

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(u,\mathcal{V}) \ge \Big(1-\frac{\varepsilon}{2}\Big)\int_{\mathbb{R}^3}\big[\Delta(\mathcal{V}-v_\infty) - \kappa(\mathcal{V}-v_\infty)\big]^2\,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}^3}(u\phi'(\mathcal{V}) - u_\infty\phi'(v_\infty))^2\,\mathrm{d}x.$$

We pass to the limit $s \searrow 0$, so that $\mathcal{V}$ converges to v in $L^2(\mathbb{R}^3)$. The first integral is obviously lower semicontinuous. Concerning the second integral, note that the integrand converges pointwise a.e. on $\mathbb{R}^3$ on a subsequence, and that it is pointwise a.e. bounded by the integrable function $2\overline{\phi'}^2(u^2+u_\infty^2)$. Hence, we can pass to the limit using the dominated convergence theorem. Now another application of Theorem I.5 yields the desired result. □

Below, we will need two further estimates for the dissipation terms from (II.2.60)&(II.2.61).

Lemma II.30 (Estimate in $L^3(\mathbb{R}^3)$). *There is a constant $\theta > 0$ such that for every $\varepsilon \in (0,\varepsilon_0)$ and every $u \in (\mathscr{P}_2 \cap W^{1,2})(\mathbb{R}^3)$, the following holds:*

$$\|u\|_{L^3}^4 \le \theta\Big(1 + \int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x\Big). \tag{II.2.64}$$

Proof. Integrating by parts, it is easily seen that

$$\frac{4}{9}\int_{\mathbb{R}^3}\big|\mathrm{D}u^{3/2}\big|^2\,\mathrm{d}x + \int_{\mathbb{R}^3} u|\mathrm{D}W_\varepsilon|^2\,\mathrm{d}x = \int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x + \int_{\mathbb{R}^3} u^2\Delta W_\varepsilon\,\mathrm{d}x.$$

By Proposition II.26 on the regularity of u_∞ and v_∞, there exists a constant C such that

$$\Delta W_\varepsilon = \Delta W + \varepsilon\phi'(v_\infty)\Delta v_\infty + \varepsilon\phi''(v_\infty)|\mathrm{D}v_\infty|^2 \le C \quad \text{on } \mathbb{R}^3$$

for all sufficiently small ε. Moreover,

$$\frac{1}{2}\int_{\mathbb{R}^3} u^2\,\mathrm{d}x \le \int_{\mathbb{R}^3} u_\infty^2\,\mathrm{d}x + \frac{1}{\lambda_\varepsilon}\int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x$$

by (II.2.55). Invoking again Proposition II.26, it follows that there exists an ε-uniform constant C' such that

$$\|\mathrm{D}u^{3/2}\|_{L^2}^2 \le C'\Big(1 + \int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x\Big)$$

holds for all $u \in \mathscr{P}_2(\mathbb{R}^3)$. On the other hand, Hölder's and Sobolev's inequalities provide

$$\|u\|_{L^3} \le \|u^{3/2}\|_{L^6}^{1/2}\|u\|_{L^1}^{1/4} \le C''\|\mathrm{D}u^{3/2}\|_{L^2}^{1/2},$$

where we have used that u is of unit mass. Together, this yields (II.2.64). □

Lemma II.31 (Estimate in $W^{2,2}(\mathbb{R}^3)$). *For every $v \in W^{2,2}(\mathbb{R}^3)$,*

$$\min(1,2\kappa,\kappa^2)\|v-v_\infty\|_{W^{2,2}}^2 \le \int_{\mathbb{R}^3}\big(\Delta(v-v_\infty)-\kappa(v-v_\infty)\big)^2\,\mathrm{d}x. \tag{II.2.65}$$

Proof. Set $\widehat{v} := v - v_\infty$ for brevity. Integration by parts yields

$$\begin{aligned}\int_{\mathbb{R}^3}(\Delta\widehat{v}-\kappa\widehat{v})^2\,\mathrm{d}x &= \int_{\mathbb{R}^3}(\Delta\widehat{v})^2\,\mathrm{d}x - 2\kappa\int_{\mathbb{R}^3}\widehat{v}\Delta\widehat{v}\,\mathrm{d}x + \kappa^2\int_{\mathbb{R}^3}\widehat{v}^2\,\mathrm{d}x\\ &= \int_{\mathbb{R}^3}\|\mathrm{D}^2\widehat{v}\|^2\,\mathrm{d}x + 2\kappa\int_{\mathbb{R}^3}|\mathrm{D}\widehat{v}|^2\,\mathrm{d}x + \kappa^2\int_{\mathbb{R}^3}\widehat{v}^2\,\mathrm{d}x,\end{aligned}$$

which clearly implies (II.2.65). □

II.2.2.5. Control of the v component

For our estimates below, we need some preliminaries concerning solutions to the time-discrete heat equation. Here, we use the iterates $\mathbf{Y}_\sigma^k$ defined in (II.2.41) to write a semi-explicit representation of the components v_τ^n for a particular choice of σ.

Lemma II.32 (Semi-explicit representation). *For every $n \in \mathbb{N}$,*

$$v_\tau^n = (1+\kappa\tau)^{-n}\mathbf{Y}_\sigma^n * v_0 + \tau\sum_{m=1}^{n}(1+\kappa\tau)^{-m}\mathbf{Y}_\sigma^m * f_\tau^{n+1-m}, \tag{II.2.66}$$

where we have set

$$f_\tau^k := -\varepsilon u_\tau^k\phi'(v_\tau^k), \quad \sigma := \frac{\tau}{1+\kappa\tau}.$$

Proof. We proceed by induction on n. By the flow interchange lemma (Theorem I.5), using the auxiliary functional $(u,v) \mapsto \int_{\mathbb{R}^3}\gamma v\,\mathrm{d}x$ for an arbitrary test function $\gamma \in C_c^\infty(\mathbb{R}^3)$, one sees as in the proof of (II.2.61) that v_τ^n is the — unique in $L^2(\mathbb{R}^3)$ — distributional solution to

$$v_\tau^n - \sigma\Delta v_\tau^n = (1+\kappa\tau)^{-1}v_\tau^{n-1} + \tau(1+\kappa\tau)^{-1}f_\tau^n.$$

Hence it can be written as

$$v_\tau^n = (1+\kappa\tau)^{-1}\mathbf{Y}_\sigma * v_\tau^{n-1} + \tau(1+\kappa\tau)^{-1}\mathbf{Y}_\sigma * f_\tau^n.$$

For $n=1$, this is (II.2.66) because $v_\tau^0 = v_0$. Now, if $n>1$, and (II.2.66) holds with $n-1$ in place of n, then

$$v_\tau^n = (1+\kappa\tau)^{-n}\mathbf{Y}_\sigma * (\mathbf{Y}_\sigma^{n-1} * v_0) + \tau\sum_{m=1}^{n-1}(1+\kappa\tau)^{-(m+1)}\mathbf{Y}_\sigma * (\mathbf{Y}_\sigma^m * f_\tau^{n-m}) + \tau(1+\kappa\tau)^{-1}\mathbf{Y}_\sigma * f_\tau^n.$$

Using that $\mathbf{Y}_\sigma * (\mathbf{Y}_\sigma^k * f) = \mathbf{Y}_\sigma^{k+1} * f$, we obtain (II.2.66). □

We are now able to prove the main result of this section.

Proposition II.33 (Control of the gradient). *Provided that $v_0 \in L^{6/5}(\mathbb{R}^3)$, then $\mathrm{D}v_\tau^n \in L^{6/5}(\mathbb{R}^3)$ for every $n \in \mathbb{N}$, and the following estimate holds:*

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le a\|v_0\|_{L^{6/5}}e^{-[\kappa]_\tau n\tau}(n\tau)^{-1/2} + \varepsilon M_1, \tag{II.2.67}$$

with the system constants

$$a := (1+\kappa)Y_1, \quad \textit{and} \quad M_1 := \overline{\phi'}Y_{6/5}(1+\kappa)^{3/4}\int_0^\infty(1+\kappa)^{-s}s^{-3/4}\,\mathrm{d}s. \tag{II.2.68}$$

Proof. From the representation formula (II.2.66) it follows that

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le (1+\kappa\tau)^{-n}\|\mathrm{D}\mathbf{Y}_\sigma^n\|_{L^1}\|v_0\|_{L^{6/5}} + \tau\sum_{m=1}^{n}(1+\kappa\tau)^{-m}\|\mathrm{D}\mathbf{Y}_\sigma^m\|_{L^{6/5}}\|f_\tau^{n+1-m}\|_{L^1}.$$

Now apply estimate (II.2.42), once with $q := 1$ and $Q := 1/2$ to the first term, and once with $q := 6/5$ and $Q := 3/4$ to the second term on the right-hand side. Further, since u_τ^n is of unit mass, one has

$$\|f_\tau^k\|_{L^1(\mathbb{R}^3)} = \varepsilon\|u_\tau^k\phi'(v_\tau^k)\|_{L^1(\mathbb{R}^3)} \le \varepsilon\overline{\phi'}.$$

This yields

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le \Upsilon_1\|v_0\|_{L^{6/5}}(1+\kappa\tau)^{-n}(\sigma n)^{-1/2} + \varepsilon\overline{\phi'}\Upsilon_{6/5}\,\tau\sum_{m=1}^{n}(1+\kappa\tau)^{-m}(\sigma m)^{-3/4}. \tag{II.2.69}$$

The sum in (II.2.69) is bounded uniformly in n and τ because

$$\tau\sum_{m=1}^{\infty}(1+\kappa\tau)^{-m}(\sigma m)^{-3/4} \le (1+\kappa\tau)^{3/4}\int_0^\infty e^{-[\kappa]_\tau t}t^{-3/4}\,\mathrm{d}t.$$

Without loss of generality, we assume that $\tau \le 1$. By the monotone convergence $e^{-[\kappa]_\tau t} \downarrow e^{-\kappa t}$ as $\tau \downarrow 0$, we can estimate the sum in (II.2.69) as

$$\tau\sum_{m=1}^{\infty}(1+\kappa\tau)^{-m}(\sigma m)^{-3/4} \le (1+\kappa)^{3/4}\int_0^\infty (1+\kappa)^{-t}t^{-3/4}\,\mathrm{d}t,$$

and the r.h.s. is finite. Thus (II.2.69) implies (II.2.67), with the given constants. □

In view of (II.2.50), we can draw the following conclusion from (II.2.67), with $\varepsilon_1 := \min(\varepsilon_0, 1)$, where $\varepsilon_0 > 0$ was implicitly characterized in Proposition II.28.

Proposition II.34 (Asymptotic boundedness of the gradient). *Assume that $v_0 \in L^{6/5}(\mathbb{R}^3)$, and define for fixed, but arbitrary $\delta' > 0$*

$$T_1 := \max\Big(1, \frac{1+\delta'}{\kappa}\log\frac{a\|v_0\|_{L^{6/5}}}{M_1}\Big), \tag{II.2.70}$$

with the system constants a and M_1 from (II.2.68)*. Then, there exists $\overline{\tau} > 0$ such that for every $\varepsilon \in (0, \varepsilon_1)$, for every $\tau \in (0, \overline{\tau}]$, and for every $n \in \mathbb{N}$ such that $n\tau \ge T_1$, one has*

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le 2M_1. \tag{II.2.71}$$

Proof. For $\varepsilon \le 1$ and $n\tau \ge T_1 \ge 1$, we obtain from (II.2.67) that

$$\begin{aligned}\|\mathrm{D}v_\tau^n\|_{L^{6/5}} &\le a\|v_0\|_{L^{6/5}}\exp\left(-\frac{1}{\tau}\log(1+\kappa\tau)T_1\right)T_1^{-1/2} + M_1\\ &\le a\|v_0\|_{L^{6/5}}\exp\left(-\frac{1}{\tau}\log(1+\kappa\tau)T_1\right) + M_1.\end{aligned} \tag{II.2.72}$$

We distinguish cases and first consider $a\|v_0\|_{L^{6/5}} \le M_1\exp\left(\frac{\kappa}{1+\delta'}\right)$, so that $T_1 = 1$. The convergence $\frac{1}{\kappa\tau}\log(1+\kappa\tau) \nearrow 1$ as $\tau \searrow 0$ yields the existence of $\overline{\tau} > 0$ such that $\frac{1}{\tau}\log(1+\kappa\tau) \ge \frac{\kappa}{1+\delta'}$ for all $\tau \in (0, \overline{\tau}]$. Hence, using (II.2.72), we deduce

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le M_1 + M_1\exp\left(\frac{\kappa}{1+\delta'} - \frac{1}{\tau}\log(1+\kappa\tau)\right) \le 2M_1,$$

for all $\tau \in (0, \overline{\tau}]$. For the converse case $a\|v_0\|_{L^{6/5}} > M_1\exp\left(\frac{\kappa}{1+\delta'}\right) > M_1$, we directly insert the definition of T_1 into (II.2.72), again for $\tau \in (0, \overline{\tau}]$:

$$\|\mathrm{D}v_\tau^n\|_{L^{6/5}} \le M_1 + a\|v_0\|_{L^{6/5}} \exp\left(-\frac{1}{\tau}\log(1+\kappa\tau)\frac{1+\delta'}{\kappa}\log\frac{a\|v_0\|_{L^{6/5}}}{M_1}\right)$$
$$\le M_1 + a\|v_0\|_{L^{6/5}} \exp\left(-\frac{\kappa}{1+\delta'}\frac{1+\delta'}{\kappa}\log\frac{a\|v_0\|_{L^{6/5}}}{M_1}\right) = 2M_1.$$

□

II.2.2.6. Bounds on the auxiliary entropy

We are now in position to prove the main estimate leading towards boundedness and exponential decay of the auxiliary entropy $\mathcal{L}$ along the discrete solution.

LEMMA II.35. *There are system constants L', M' and an $\varepsilon_2 \in (0, \varepsilon_1)$ such that for every $\varepsilon \in (0, \varepsilon_2)$, for every $\tau \in (0, \overline{\tau}]$, and for every n with $n\tau > T_1$, we have that*

$$(1 + 2\Lambda'_\varepsilon\tau)\mathcal{L}(u_\tau^n, v_\tau^n) \le \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}) + \tau\varepsilon M' \tag{II.2.73}$$

with $\Lambda'_\varepsilon := \min(\kappa, \lambda_0) - L'\varepsilon$.

PROOF. For clarity, we simply write u and v in place of u_τ^n and v_τ^n, respectively, and we introduce $\widehat{v} := v - v_\infty$. Since $n\tau > T_1$ by hypothesis, Proposition II.34 implies that

$$\|\mathrm{D}\widehat{v}\|_{L^{6/5}} \le \|\mathrm{D}v\|_{L^{6/5}} + \|\mathrm{D}v_\infty\|_{L^{6/5}} \le 2M_1 + \sup_{0<\varepsilon<\varepsilon_1} \|\mathrm{D}v_\infty\|_{L^{6/5}} =: Z < \infty.$$

Now, since

$$\big|\mathrm{D}\big(\phi(v) - \phi(v_\infty)\big)\big|^2 \le 2\phi'(v)^2|\mathrm{D}\widehat{v}|^2 + 2(\phi'(v) - \phi'(v_\infty))^2|\mathrm{D}v_\infty|^2 \le \alpha|\mathrm{D}\widehat{v}|^2 + \beta\widehat{v}^2,$$

with the system constants

$$\alpha := 2\overline{\phi'}^2, \quad \beta := 2\overline{\phi''}^2 \sup_{0<\varepsilon<\varepsilon_1} \|\mathrm{D}v_\infty\|_{L^\infty}^2, \tag{II.2.74}$$

we conclude that

$$\begin{aligned}
&\int_{\mathbb{R}^3} u\big|\mathrm{D}\big(\phi(v) - \phi(v_\infty)\big)\big|^2\,\mathrm{d}x \le \alpha\int_{\mathbb{R}^3} u|\mathrm{D}\widehat{v}|^2\,\mathrm{d}x + \beta\int_{\mathbb{R}^3} u\widehat{v}^2\,\mathrm{d}x \\
&\le \alpha\|u\|_{L^3}\|\mathrm{D}\widehat{v}\|_{L^3}^2 + \beta\|u\|_{L^1}\|\widehat{v}\|_{L^\infty}^2 \\
&\le \|u\|_{L^3}^4 + \alpha^{4/3}\|\mathrm{D}\widehat{v}\|_{L^3}^{8/3} + \beta\|\widehat{v}\|_{L^\infty}^2 \\
&\le \|u\|_{L^3}^4 + \alpha^{4/3}\big(S_1\|\widehat{v}\|_{W^{2,2}}^{3/4}\|\mathrm{D}\widehat{v}\|_{L^{6/5}}^{1/4}\big)^{8/3} + \beta S_2\|\widehat{v}\|_{W^{2,2}}^2 \\
&\le \theta\left(1 + \int_{\mathbb{R}^3} u|\mathrm{D}(u + W_\varepsilon)|^2\,\mathrm{d}x\right) + \frac{\alpha^{4/3}S_1^{8/3}Z^{2/3} + \beta S_2}{\min(1, 2\kappa, \kappa^2)}\int_{\mathbb{R}^3}\big(\Delta\widehat{v} - \kappa\widehat{v}\big)^2\,\mathrm{d}x,
\end{aligned} \tag{II.2.75}$$

where θ is the constant from (II.2.64), and S_1, S_2 are Sobolev constants. Next, observe that

$$(u\phi'(v) - u_\infty\phi'(v_\infty))^2 \le 2(u - u_\infty)^2\phi'(v)^2 + 2u_\infty^2(\phi'(v) - \phi'(v_\infty))^2 \le \alpha(u - u_\infty)^2 + \beta\|u_\infty\|_{L^\infty}^2\widehat{v}^2,$$

with the same constants as in (II.2.74). Therefore, using (II.2.55), (II.2.56) and Proposition II.26(a),

$$\int_{\mathbb{R}^3}(u\phi'(v) - u_\infty\phi'(v_\infty))^2\,\mathrm{d}x \le \alpha\|u - u_\infty\|_{L^2}^2 + \beta(U_0 - \varepsilon V\phi'(0))^2\|\widehat{v}\|_{L^2}^2$$
$$\le 2\alpha\mathcal{L}_u(u) + \frac{2\beta}{\kappa}(U_0 - \varepsilon V\phi'(0))^2\mathcal{L}_v(v).$$

Altogether, we have shown that there is a system constant M' such that (recall the dissipation terms $\mathcal{D}_u(u,v)$ and $\mathcal{D}_v(u,v)$ from (II.2.60)&(II.2.61))

$$\begin{aligned}\mathcal{D}_u(u,v)+\mathcal{D}_v(u,v) \geq (1-M'\varepsilon)\int_{\mathbb{R}^3} u|\mathrm{D}(u+W_\varepsilon)|^2\,\mathrm{d}x + (1-M'\varepsilon)\int_{\mathbb{R}^3}\big(\Delta\widehat{v}-\kappa\widehat{v}\big)^2\,\mathrm{d}x \\ -M'\varepsilon\mathcal{L}_u(u)-M'\varepsilon\mathcal{L}_v(v)-M'\varepsilon\end{aligned} \tag{II.2.76}$$

for all $\varepsilon\in(0,\varepsilon_1)$. Provided that $M'\varepsilon<1$, we can apply (II.2.55) and (II.2.56) to estimate further:

$$\mathcal{D}_u(u,v)+\mathcal{D}_v(u,v)\geq\big(2\lambda_\varepsilon(1-M'\varepsilon)-M'\varepsilon\big)\mathcal{L}_u(u)+\big(2\kappa(1-M'\varepsilon)-M'\varepsilon\big)\mathcal{L}_v(v)-M'\varepsilon.$$

Finally, we can choose $\varepsilon_2\in(0,\varepsilon_1)$ so small that the coefficients of $\mathcal{L}_u$ and $\mathcal{L}_v$ above are nonnegative for every $\varepsilon\in(0,\varepsilon_2)$, and thus we arrive at the final estimate

$$\mathcal{D}_u(u,v)+\mathcal{D}_v(u,v)\geq 2(\min(\kappa,\lambda_\varepsilon)-L'\varepsilon)\mathcal{L}(u,v)-\varepsilon M',$$

with a suitable choice of L'. Now estimate (II.2.59) implies (II.2.73) with Λ'_ε given as above. □

Diminishing ε_2 such that the constant $1+K\varepsilon_2$ in (II.2.57) is less or equal to two, we derive the following explicit estimate:

Proposition II.36. *Assume that $v_0\in L^{6/5}(\mathbb{R}^3)$, and let T_1 be defined as in* (II.2.70). *Then, for every $\varepsilon\in(0,\varepsilon_2)$, for every $\tau\in(0,\overline{\tau}]$, and for every $n\geq\overline{n}$, with $\overline{n}:=\left\lceil\frac{T_1}{\tau}\right\rceil$, the following estimate holds:*

$$\mathcal{L}(u_\tau^n,v_\tau^n)\leq 2\big(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty\big)(1+2\Lambda'_\varepsilon\tau)^{-(n-\overline{n})}+\frac{M'\varepsilon}{2\Lambda'_\varepsilon}\big(1-(1+2\Lambda'_\varepsilon\tau)^{-(n-\overline{n})}\big). \tag{II.2.77}$$

Proof. We proceed by induction on $n\geq\overline{n}$. For $n=\overline{n}$, (II.2.77) is a consequence of (II.2.57) and the energy estimate $\mathcal{E}(u_\tau^n,v_\tau^n)\leq\mathcal{E}(u_0,v_0)$. Now assume (II.2.77) for some $n\geq\overline{n}$, and apply the iterative estimate (II.2.73):

$$\begin{aligned}\mathcal{L}(u_\tau^{n+1},v_\tau^{n+1}) &\leq (1-2\Lambda'_\varepsilon\tau)^{-1}\mathcal{L}(u_\tau^n,v_\tau^n)+(1+2\Lambda'_\varepsilon\tau)^{-1}\tau M'\varepsilon \\ &\leq 2\big(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty\big)(1+2\Lambda'_\varepsilon\tau)^{-((n+1)-\overline{n})} \\ &\quad+\frac{M'\varepsilon}{2\Lambda'_\varepsilon}\big((1+2\Lambda'_\varepsilon\tau)^{-1}-(1+2\Lambda'_\varepsilon\tau)^{-((n+1)-\overline{n})}\big)+(1+2\Lambda'_\varepsilon\tau)^{-1}\tau M'\varepsilon.\end{aligned}$$

Elementary calculations show that the last expression above is equal to the right-hand side of (II.2.77) with $n+1$ in place of n. □

As before, we deduce a uniform estimate for large times:

Proposition II.37 (Asymptotic boundedness of $\mathcal{L}$). *With the assumptions from Proposition II.34, define for fixed, but arbitrary $\delta>0$ the quantities*

$$M_2:=\frac{M'}{2(\min(\kappa,\lambda_0)-L'\varepsilon_2)},\qquad T_2:=T_1+\max\left(0,\frac{1+2\delta}{2\Lambda'_\varepsilon}\log\frac{2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau})}{M_2}\right).$$

Then, there exists $\overline{\tau}'\in(0,\overline{\tau}]$ such that for every $\varepsilon\in(0,\varepsilon_2)$, $\tau\in(0,\overline{\tau}']$ and all $n\in\mathbb{N}$ with $n\tau\geq T_2$, one has

$$\mathcal{L}(u_\tau^n,v_\tau^n)\leq 2M_2. \tag{II.2.78}$$

Proof. We proceed similarly to the proof of Proposition II.34. First, using (II.2.77) and $\overline{n}\tau\leq T_1+\tau$, we get for $n\tau\geq T_2$ that

$$\begin{aligned}\mathcal{L}(u_\tau^n,v_\tau^n) &\leq M_2+2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)\exp\left(-(T_2-T_1)\frac{1}{\tau}\log(1+2\Lambda'_\varepsilon\tau)\right)(1+2\Lambda'_\varepsilon\tau) \\ &\leq M_2+2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau})\exp\left(-(T_2-T_1)\frac{1}{\tau}\log(1+2\Lambda'_\varepsilon\tau)\right),\end{aligned} \tag{II.2.79}$$

where we used in the last step that $\Lambda'_\varepsilon \leq \lambda_0$. Consider the case $2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau}) \leq M_2$, where $T_2 = T_1$. Trivially, (II.2.78) then follows from (II.2.79).

For the remaining case, we use one more time that $\frac{1}{s}\log(1+s) \nearrow 1$ as $s \searrow 0$: there exists $\overline{s} > 0$ such that $\frac{1}{s}\log(1+s) \geq \frac{1}{1+2\delta}$ for all $s \in (0,\overline{s}]$. Defining $\overline{\tau}' := \min(\overline{\tau}, \frac{s}{2\lambda_0})$ yields $\frac{1}{\tau}\log(1+2\Lambda'_\varepsilon\tau) \geq \frac{1}{1+2\delta}$ for all $\tau \in (0,\overline{\tau}')$.

We conclude with (II.2.79):

$$\begin{aligned}
&\mathcal{L}(u_\tau^n, v_\tau^n)\\
&\quad\leq M_2 + 2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau})\exp\left(-(1+2\delta)\log\left[\frac{2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau})}{M_2}\right]\frac{1}{1+2\delta}\right)\\
&\quad= 2M_2.
\end{aligned}$$

□

We have thus proved that, for $t \geq T_2$, the auxiliary entropy $\mathcal{L}$ is bounded by a system constant. Next, we prove that $\mathcal{L}$ is not only bounded, but actually convergent to zero exponentially fast.

II.2.2.7. Exponential decay for large times

Lemma II.38. *There is a constant $L'' > L'$ and some $\varepsilon_3 \in (0,\varepsilon_2)$ such that for every $\varepsilon \in (0,\varepsilon_3)$, for every sufficiently small $\tau > 0$, and for every n such that $n\tau > T_2$, we have*

$$(1+2\Lambda''_\varepsilon\tau)\mathcal{L}(u_\tau^n, v_\tau^n) \leq \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}), \tag{II.2.80}$$

with $\Lambda''_\varepsilon := \min(\lambda_0,\kappa) - L''\varepsilon$.

Proof. We proceed like in the proof of Lemma II.35, with the following modifications. By Proposition II.37, we know that

$$\mathcal{L}_u(u_\tau^n) \leq \mathcal{L}(u_\tau^n, v_\tau^n) \leq 2M_2.$$

Using the first inequality in (II.2.55), we can estimate the L^2-norm of u_τ^n by a system constant:

$$\|u\|_{L^2} \leq \|u_\infty\|_{L^2} + \|u-u_\infty\|_{L^2} \leq \sup_{0<\varepsilon<\varepsilon_2}\|u_\infty\|_{L^2} + 2\sqrt{M_2} =: Z.$$

This allows us to replace the chain of estimates (II.2.75) by a simpler one:

$$\int_{\mathbb{R}^3} u\big|\mathrm{D}\big(\phi(v)-\phi(v_\infty)\big)\big|^2\,\mathrm{d}x \leq \|u\|_{L^2}\big(\alpha\|\mathrm{D}\widehat{v}\|_{L^4}^2 + \beta\|\widehat{v}\|_{L^4}^2\big),$$

with the constants from (II.2.74). Using the Sobolev inequalities

$$\|\mathrm{D}\widehat{v}\|_{L^4} \leq S\|\widehat{v}\|_{W^{2,2}} \quad\text{and}\quad \|\widehat{v}\|_{L^4} \leq S\|\widehat{v}\|_{W^{1,2}}$$

in combination with (II.2.65) and (II.2.56), respectively, we arrive at

$$\int_{\mathbb{R}^3} u\big|\mathrm{D}\big(\phi(v)-\phi(v_\infty)\big)\big|^2\,\mathrm{d}x \leq \frac{\alpha Z S^2}{\min(1,2\kappa,\kappa^2)}\int_{\mathbb{R}^3}(\Delta\widehat{v}-\kappa\widehat{v})^2\,\mathrm{d}x + \frac{2\beta Z S^2}{\min(1,\kappa)}\mathcal{L}_v(v).$$

This eventually leads to the dissipation estimate (II.2.76) again, with a different constant M', but *without the constant term $-\varepsilon M'$*. By means of (II.2.59), this implies (II.2.80) for appropriate choices of L'' and ε_3. □

By iteration of (II.2.80), starting from (II.2.78), one immediately obtains

Corollary II.39. *For all sufficiently small τ and every n such that $n\tau \geq T_2$, we have*

$$\mathcal{L}(u_\tau^n, v_\tau^n) \leq 2M_2 e^{-2[\Lambda''_\varepsilon]_\tau(n\tau - T_2)}. \tag{II.2.81}$$

II.2.2.8. Passage to continuous time

To complete the proof of Theorem II.4, we consider the limit $\tau \searrow 0$ of the estimates obtained above. This means that we consider a vanishing sequence $(\tau_k)_{k\in\mathbb{N}}$ such that the corresponding sequence of discrete solutions $(u_{\tau_k}, v_{\tau_k})_{k\in\mathbb{N}}$ converges in the sense of Theorem II.22 to a weak solution (u, v) to (II.1.1). The lower semicontinuity properties of $\mathcal{L}$ allow us to conclude that

$$\mathcal{L}(t) := \mathcal{L}\big(u(t,\cdot), v(t,\cdot)\big) \leq \liminf_{k\to\infty} \mathcal{L}\big(u_{\tau_k}(t,\cdot), v_{\tau_k}(t,\cdot)\big) \quad \text{for every } t \geq 0.$$

Recalling (II.2.50), we conclude from (II.2.81) that

$$\mathcal{L}(t) \leq 2M_2 e^{-2\Lambda''_\varepsilon(t-T_2)} \quad \text{for all } t \geq T_2. \tag{II.2.82}$$

Moreover, from (II.2.57) and the energy estimate, we obtain

$$\mathcal{L}(t) \leq 2\big(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty\big) \quad \text{for all } t \geq 0. \tag{II.2.83}$$

We prove that

$$\mathcal{L}(t) \leq C(1 + \|v_0\|_{L^{6/5}})^{2(1+\delta')}(1 + \mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)^{2(1+\delta)} e^{-2\Lambda''_\varepsilon t} \quad \text{for all } t \geq 0. \tag{II.2.84}$$

From this, claim (II.1.5) in Theorem II.4 follows with $\Lambda_\varepsilon := \Lambda''_\varepsilon$ and $C_{\delta,\delta'} := \sqrt{C}$. We distinguish cases and first consider $2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)(1 + 2\lambda_0\overline{\tau}) \leq M_2$. There, $T_2 = T_1$, so combining (II.2.82)&(II.2.83) yields for all $t \geq 0$:

$$\mathcal{L}(t) \leq \max(2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty), 2M_2) \exp(2\Lambda''_\varepsilon T_1) \exp(-2\Lambda''_\varepsilon t).$$

In the other case, we get

$$\mathcal{L}(t) \leq \max(2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty), 2M_2) \exp(2\Lambda''_\varepsilon T_2) \exp(-2\Lambda''_\varepsilon t),$$

from which the following estimate can be obtained using the definition of T_2 and $\Lambda''_\varepsilon \leq \Lambda'_\varepsilon$:

$$\begin{aligned}\mathcal{L}(t) &\leq \max(2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty), 2M_2) \exp(2\Lambda''_\varepsilon T_1) \left(\frac{2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)(1 + 2\lambda_0\overline{\tau})}{M_2}\right)^{\frac{\Lambda''_\varepsilon}{\Lambda'_\varepsilon}(1+2\delta)} \exp(-2\Lambda''_\varepsilon t)\\ &\leq \max(2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty), 2M_2) \exp(2\Lambda''_\varepsilon T_1) \left(\frac{2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)(1 + 2\lambda_0\overline{\tau})}{M_2}\right)^{1+2\delta} \exp(-2\Lambda''_\varepsilon t),\end{aligned}$$

since $\frac{2(\mathcal{E}(u_0,v_0)-\mathcal{E}_\infty)(1+2\lambda_0\overline{\tau})}{M_2} > 1$ in this case. From the combination of both cases, we infer

$$\mathcal{L}(t) \leq 2\max(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty, M_2) \exp(2\Lambda''_\varepsilon T_1) \max\left(1, \frac{2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)(1 + 2\lambda_0\overline{\tau})}{M_2}\right)^{1+2\delta} \exp(-2\Lambda''_\varepsilon t).$$

The definition of T_1 and the fact that $\Lambda''_\varepsilon \leq \kappa$ yield

$$\exp(2\Lambda''_\varepsilon T_1) \leq \exp(2\kappa T_1) \leq \max\left(\exp(2\kappa), \left(\frac{a\|v_0\|_{L^{6/5}}}{M_1}\right)^{2(1+\delta')}\right) \leq C_{\delta'}(1 + \|v_0\|_{L^{6/5}})^{2(1+\delta')}.$$

Likewise,

$$2\max(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty, M_2) \max\left(1, \frac{2(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty)(1 + 2\lambda_0\overline{\tau})}{M_2}\right)^{1+2\delta} \leq C_\delta(\mathcal{E}(u_0, v_0) - \mathcal{E}_\infty + 1)^{2(1+\delta)}.$$

Putting everything together, (II.2.84) follows and the proof of Theorem II.4 is complete. □

CHAPTER II.3

The classical Keller-Segel model in one spatial dimension

In this chapter, we use the methods applied in Chapter II.2 to investigate system (II.1.1) in one spatial dimension. Our main result is that the one-dimensionality allows one to consider arbitrary diffusion exponents $m \geq 1$. We shall treat here the most involved case of linear diffusion ($m = 1$) corresponding to the classical Keller-Segel model. An abridged version of the results in this chapter is contained in the note [187]. Since the overall strategy of proof in this chapter is the same as in the previous one, we skip most of the technical details and only highlight the important differences.

II.3.1. Existence of weak solutions

This section is concerned with a motivation for the proof of Theorem II.6. The crucial step here is to verify that the discrete solution (u_τ, v_τ) is well-defined and regular enough to allow for passage to the continuous-time limit $\tau \searrow 0$ in a strong sense. Once obtained, we can proceed as in Chapter II.2 establishing an approximate weak formulation which turns into the weak formulation of the time-continuous equation as $\tau \searrow 0$. One major obstacle to overcome is the unboundedness from below of the energy functional $\mathcal{E}$ as defined in (II.1.3).

We first prove the following

Proposition II.40 (Minimizing movement scheme). *For each $\tau \in (0,1)$ and $(\widetilde{u}, \widetilde{v}) \in \mathbf{X}$, the functional*

$$\mathcal{E}_\tau(\cdot|\widetilde{u}, \widetilde{v}) := \frac{1}{2\tau}\mathbf{d}^2(\cdot, (\widetilde{u}, \widetilde{v})) + \mathcal{E}$$

possesses a minimizer $(u, v) \in \mathscr{P}_2(\mathbb{R}) \times W^{1,2}(\mathbb{R})$ with $\int_{\mathbb{R}} u \log u \, \mathrm{d}x < \infty$. Moreover, there exist constants $K_0, K_1, K_2 > 0$ such that if in addition $\widetilde{v} \in W^{1,2}(\mathbb{R})$ and $\int_{\mathbb{R}} \widetilde{u} \log \widetilde{u} \, \mathrm{d}x < \infty$, then

$$\begin{aligned} &\tau\|(\sqrt{u})_x\|_{L^2}^2 + \tau\|v_{xx}\|_{L^2}^2 \\ &\quad\leq K_0 \int_{\mathbb{R}} (u \log u - \widetilde{u} \log \widetilde{u}) \, \mathrm{d}x + K_1(\|v\|_{W^{1,2}}^2 - \|\widetilde{v}\|_{W^{1,2}}^2) + K_2\tau(\|v\|_{W^{1,2}}^2 + 1); \end{aligned} \tag{II.3.1}$$

hence, $v \in W^{2,2}(\mathbb{R})$, $\sqrt{u} \in W^{1,2}(\mathbb{R})$ and $u \in L^\infty(\mathbb{R})$.

Proof. First, in one spatial dimension, there exists $C_0 > 0$ such that $\|v\|_{L^\infty} \leq \|v\|_{C^{0,\frac{1}{2}}} \leq C_0\|v\|_{W^{1,2}}$. Moreover, for some $C_1 > 0$, one has

$$\int_{\mathbb{R}} u \log u \, \mathrm{d}x \geq -C_1(\mathbf{m}_2(u) + 1)^{1/2}.$$

From this, we easily see that for all $(u, v) \in \mathscr{P}_2(\mathbb{R}) \times W^{1,2}(\mathbb{R})$ with $\int_{\mathbb{R}} u \log u \, \mathrm{d}x < \infty$, we have

$$\int_{\mathbb{R}} u \log u \, \mathrm{d}x + \underline{W} + \frac{1}{2}\|v_x\|_{L^2}^2 - |\chi| C_0 \|v\|_{W^{1,2}} \leq \mathcal{E}(u, v) < \infty,$$

where $\underline{W} \in \mathbb{R}$ is a lower bound for W. Using the triangle inequality for $\mathbf{d}$ and Young's inequality, we deduce coercivity of $\mathcal{E}_\tau(\cdot|\widetilde{u}, \widetilde{v})$:

$$\mathcal{E}_\tau(u, v|\widetilde{u}, \widetilde{v}) \geq \frac{1}{4}\|v\|_{W^{1,2}}^2 + \frac{1}{4}\mathbf{m}_2(u) - C. \tag{II.3.2}$$

Thus, by the Banach-Alaoglu, Arzelà-Ascoli and Prokhorov theorems, a minimizing sequence $(u_n, v_n)_{n\in\mathbb{N}}$ for $\mathcal{E}_\tau(\cdot|\widetilde{u},\widetilde{v})$ converges — at least on a subsequence — to some limit $(u,v) \in \mathscr{P}_2(\mathbb{R}) \times W^{1,2}(\mathbb{R})$ with $\int_\mathbb{R} u \log u \,\mathrm{d}x < \infty$: $v_n \rightharpoonup v$ in $W^{1,2}(\mathbb{R})$, $v_n \to v$ locally uniformly in $\mathbb{R}$ and $u_n \rightharpoonup u$ narrowly in $\mathscr{P}(\mathbb{R})$. With respect to these convergences $\mathcal{E}_\tau(\cdot|\widetilde{u},\widetilde{v})$ is lower semicontinuous, which is clear except for the term $\int_\mathbb{R} u_n v_n \,\mathrm{d}x$. We employ a truncation argument similar as in the proof of Proposition II.12 to prove the lower semicontinuity of this remaining term, and consequently obtain the minimizing property for (u,v). It remains to prove the additional regularity estimate (II.3.1). We investigate the dissipation of $\mathcal{E}$ along the (auxiliary) 0-flow $(\mathcal{U},\mathcal{V})_{s\geq 0}$ w.r.t. $\mathbf{d}$ generated by the geodesically 0-convex functional

$$\mathcal{A}(u,v) := \int_\mathbb{R} \left[u\log u + \frac{1}{2}v_x^2 + \frac{\kappa}{2}v^2\right] \mathrm{d}x$$

on $\mathbf{X}$. Elementary calculations yield, since we have $\mathcal{U}_s = \mathcal{U}_{xx}$, $\mathcal{V}_s = \mathcal{V}_{xx} - \kappa\mathcal{V}$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V}) \leq \int_\mathbb{R}\left[-4(\sqrt{\mathcal{U}})_x^2 + \|W_{xx}\|_{L^\infty} - \frac{1}{2}(\mathcal{V}_{xx}-\kappa\mathcal{V})^2 + \frac{5}{2}\chi^2\mathcal{U}^2 + \frac{\kappa^2}{2}\mathcal{V}^2\right]\mathrm{d}x.$$

Using the one-dimensional Sobolev inequality

$$\|\eta\|_{L^4} \leq C\|\eta\|_{W^{1,2}}^{1/4}\|\eta\|_{L^2}^{3/4}, \tag{II.3.3}$$

we eventually arrive at

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V}) \leq -2\|(\sqrt{\mathcal{U}})_x\|_{L^2}^2 - \frac{1}{2}\|\mathcal{V}_{xx}-\kappa\mathcal{V}\|_{L^2}^2 + \frac{\kappa^2}{2}\|v\|_{L^2}^2 + C_2. \tag{II.3.4}$$

Finally, we use the flow interchange lemma (Theorem I.5) to obtain $\mathcal{A}(u,v) + \tau\mathrm{D}^\mathcal{A}\mathcal{E}(u,v) \leq \mathcal{A}(\widetilde{u},\widetilde{v})$, which yields (II.3.1) in combination with (II.3.4) and lower semicontinuity as $s \searrow 0$. □

In this specific setting, $\mathcal{E}$ is unbounded from below. Hence, the classical estimates from Proposition II.15 do not provide any information at first sight. However, thanks to the coercivity of $\mathcal{E}_\tau$ from (II.3.2) the following estimate — considering a *finite* time horizon — can be deduced:

Proposition II.41 (Total square distance estimate). *Let $(u_0,v_0) \in \mathbf{X}$ with $\int_\mathbb{R} u_0\log u_0\,\mathrm{d}x < \infty$ and $v_0 \in W^{1,2}(\mathbb{R})$ and a time horizon $T>0$ be given. There exists a constant $C>0$ depending on u_0, v_0 and T such that for all $\tau \in (0,\frac{1}{8})$, the following holds with $N := \left\lfloor \frac{T}{\tau}\right\rfloor$ and the sequence of minimizing movements $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$:*

$$\sum_{n=1}^N \mathbf{d}^2((u_\tau^n,v_\tau^n),(u_\tau^{n-1},v_\tau^{n-1})) \leq C\tau. \tag{II.3.5}$$

Proof. Our method of proof is inspired from [4, Sect. 3.2]. Using the minimizing property of (u_τ^n, v_τ^n), one has

$$\sum_{n=1}^N \mathbf{d}^2((u_\tau^n,v_\tau^n),(u_\tau^{n-1},v_\tau^{n-1})) \leq 2\tau(\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\tau^N,v_\tau^N)). \tag{II.3.6}$$

Additionally, for $\tau_* := \frac{3}{4}$, the following is true:

$$\mathcal{E}(u_\tau^N,v_\tau^N) \geq \inf_{(u,v)\in\mathbf{X}} \mathcal{E}_{\tau_*}(u,v\,|\,u_0,v_0) - \frac{1}{2\tau_*}\mathbf{d}^2((u_\tau^N,v_\tau^N),(u_0,v_0)), \tag{II.3.7}$$

and the infimum is finite thanks to Proposition II.40. It thus remains to provide an appropriate upper bound for $\mathbf{d}^2((u_\tau^N,v_\tau^N),(u_0,v_0))$. Writing as a telescopic sum, we obtain

$$\begin{aligned}&\frac{1}{2}\mathbf{d}^2((u_\tau^N,v_\tau^N),(u_0,v_0))\\ &\quad= \sum_{n=1}^N\left[\mathbf{d}^2((u_\tau^n,v_\tau^n),(u_0,v_0)) - \frac{1}{2}\left(\mathbf{d}^2((u_\tau^n,v_\tau^n),(u_0,v_0)) + \mathbf{d}^2((u_\tau^{n-1},v_\tau^{n-1}),(u_0,v_0))\right)\right]\end{aligned}$$

$$\le \sum_{n=1}^{N} \mathbf{d}((u_\tau^n, v_\tau^n),(u_0,v_0)) \left[\mathbf{d}((u_\tau^n, v_\tau^n),(u_0,v_0)) - \mathbf{d}((u_\tau^{n-1}, v_\tau^{n-1}),(u_0,v_0))\right]$$
$$\le \sum_{n=1}^{N} \mathbf{d}((u_\tau^n, v_\tau^n),(u_0,v_0))\mathbf{d}((u_\tau^{n-1}, v_\tau^{n-1}),(u_\tau^n, v_\tau^n)).$$

We use Young's inequality and the estimates (II.3.6)&(II.3.7):

$$\frac{1}{2}\mathbf{d}^2((u_\tau^N, v_\tau^N),(u_0,v_0)) \le \frac{\tau_*}{2}\sum_{n=1}^{N}\frac{1}{2\tau}\mathbf{d}^2((u_\tau^{n-1}, v_\tau^{n-1}),(u_\tau^n, v_\tau^n)) + \frac{1}{\tau_*}\sum_{n=1}^{N}\tau\mathbf{d}^2((u_\tau^n, v_\tau^n),(u_0,v_0))$$
$$\le \frac{\tau_*}{2}\mathcal{E}(u_0,v_0) - \frac{\tau_*}{2}\inf_{(u,v)\in\mathbf{X}}\mathcal{E}_{\tau_*}(u,v\,|\,u_0,v_0) + \frac{1}{4}\mathbf{d}^2(u_\tau^N, v_\tau^N),(u_0,v_0)) + \frac{1}{\tau_*}\sum_{n=1}^{N}\tau\mathbf{d}^2((u_\tau^n, v_\tau^n),(u_0,v_0)).$$

Rearranging and recalling that $\tau_* = \frac{3}{4}$, we obtain with the binomial formula that

$$\mathbf{d}^2((u_\tau^N, v_\tau^N),(u_0,v_0)) \le \frac{3}{2}\mathcal{E}(u_0,v_0) - \frac{3}{2}\inf_{(u,v)\in\mathbf{X}}\mathcal{E}_{\tau_*}(u,v\,|\,u_0,v_0) + \frac{16}{3}\sum_{n=1}^{N}\tau\mathbf{d}^2((u_\tau^n, v_\tau^n),(u_0,v_0)).$$

Since $\tau < \frac{1}{8} < \frac{3}{16}$, we are in position to apply a discrete Gronwall-type lemma [4, Lemma 3.2.4] which yields

$$\mathbf{d}^2((u_\tau^N, v_\tau^N),(u_0,v_0)) \le \widetilde{C}\exp\left(\frac{16}{3-16\tau}(N-1)\tau\right) \le \widetilde{C}\exp\left(16T\right),$$

for a constant $\widetilde{C} > 0$ depending on u_0, v_0. Thus, the asserted estimate (II.3.5) follows from (II.3.6). □

We can now proceed as in Chapter II.2 and arrive at a weak solution (u,v) to (II.1.1) with the properties

$$v \in L^\infty([0,T];L^2(\mathbb{R})),\ v_x \in L^\infty([0,T];L^2(\mathbb{R})),\ v_t \in L^2([0,T];L^2(\mathbb{R})),$$

for each $T > 0$. We immediately deduce that $v \in L^\infty([0,T]\times\mathbb{R})$. We now show that v is continuous in both arguments. In fact, for all bounded intervals $I \subset \mathbb{R}$, v belongs to the *anisotropic* Sobolev space $W^{1,\mathbf{P}}([0,T]\times I)$ with $\mathbf{P} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, the spectral radius of which is less than 1. Since in this case $W^{1,\mathbf{P}}([0,T]\times I) \Subset C^0([0,T]\times\bar{I})$, the claim follows (for details on anisotropic spaces, see e.g. [17, 116]).

II.3.2. Convergence to equilibrium

This section is devoted to a sketch of the proof for Theorem II.8. Again, the overall strategy is as in Chapter II.2, but the analysis simplifies on a broader range here thanks to the spatial one-dimensionality.

It is easily shown (see Section II.2.2.1) that the additional assumption of λ_0-convexity of the confinement W yields boundedness from below of the energy $\mathcal{E}$. We obtain $(u_\infty, v_\infty) \in (\mathscr{P}_2 \cap L^\infty)(\mathbb{R}) \times W^{2,2}(\mathbb{R})$ as a minimizer of $\mathcal{E}$ similarly to the proof of Proposition II.23. Again, uniqueness is proved by showing strict convexity of $\mathcal{E}$ as a functional on $L^2(\mathbb{R}) \times L^2(\mathbb{R})$, which requires a small coupling strength $\varepsilon > 0$. Note that $u_\infty \in L^\infty(\mathbb{R})$ is crucial here.

Using the properties of (u_∞, v_∞), we observe that the energy can be decomposed as follows into a convex part $\mathcal{L}$ (see Proposition II.42 below) and a non-convex, but controllable part $\varepsilon\mathcal{L}_*$:

$$\mathcal{E}(u,v) - \mathcal{E}(u_\infty,v_\infty) = \mathcal{L}(u,v) + \varepsilon\mathcal{L}_*(u,v), \tag{II.3.8}$$

where $\mathcal{L}(u,v) = \mathcal{L}_u(u) + \mathcal{L}_v(v)$,

$$\mathcal{L}_u(u) := \int_{\mathbb{R}} [u\log u - u_\infty\log u_\infty + W_\varepsilon(u-u_\infty)]\,\mathrm{d}x,\ \text{with}\ W_\varepsilon := W - \varepsilon v_\infty,$$

$$\mathcal{L}_v(v) := \frac{1}{2}\|(v-v_\infty)_x\|_{L^2}^2 + \frac{\kappa}{2}\|v-v_\infty\|_{L^2}^2, \quad \mathcal{L}_*(u,v) := -\int_{\mathbb{R}}(u-u_\infty)(v-v_\infty)\,\mathrm{d}x.$$

PROPOSITION II.42 (Properties of $\mathcal{L}$). *Let ε be sufficiently small. Then, the following statements hold:*

(a) There exists $M_1 > 0$ such that the perturbed potential *W_ε is λ_ε-convex, where $\lambda_\varepsilon := \lambda_0 - M_1\varepsilon > 0$.*

(b) The functional $\mathcal{L}_u$ is λ_ε-convex along generalized geodesics in $(\mathscr{P}_2(\mathbb{R}), \mathbf{W}_2)$ and

$$\frac{\lambda_\varepsilon}{2}\mathbf{W}_2^2(u,u_\infty) \le \mathcal{L}_u(u) \le \frac{1}{2\lambda_\varepsilon}\int_{\mathbb{R}} u((\log u + W_\varepsilon)_x)^2\,\mathrm{d}x.$$

(c) The functional $\mathcal{L}_v$ is geodesically κ-convex on $L^2(\mathbb{R})$ and

$$\frac{\kappa}{2}\|v-v_\infty\|_{L^2}^2 \le \mathcal{L}_v(v) \le \frac{1}{2\kappa}\|(v-v_\infty)_{xx} - \kappa(v-v_\infty)\|_{L^2}^2.$$

(d) There exists $M_2 > 0$ such that $\mathcal{L}(u,v) \le (1+M_2\varepsilon)(\mathcal{E}(u,v) - \mathcal{E}(u_\infty,v_\infty))$.

In one spatial dimension, the proof of part (a) simplifies dramatically compared to Section II.2.2, since

$$(W_\varepsilon)_{xx} = W_{xx} - \varepsilon(v_\infty)_{xx} = W_{xx} - \varepsilon(\kappa v_\infty - \varepsilon u_\infty) \ge \lambda_0 - \varepsilon\kappa\|v_\infty\|_{L^\infty} \ge \lambda_0 - \varepsilon\widetilde{C}(\mathcal{E}(u_\infty,v_\infty)+1),$$

for some constant $\widetilde{C} > 0$. The proof of part (d) is mainly a consequence of the (generalized) *Csiszár-Kullback inequality*

$$\|u-u_\infty\|_{L^1}^2 \le C\mathcal{L}_u(u),$$

since $\mathcal{L}_u$ coincides with the *relative entropy*

$$\widetilde{\mathcal{L}}_u(u;u_\infty) := \int_{\mathbb{R}} \left(h(u) - h(u_\infty) - h'(u_\infty)(u-u_\infty)\right)\,\mathrm{d}x,$$

for the admissible choice $h(z) := z\log z$ (cf. [51]).

The following estimate — valid in one spatial dimension only — will be needed for the analysis below.

PROPOSITION II.43 (Estimate in $L^2(\mathbb{R})$). *There exists a constant $C' > 0$ such that for all $u \in \mathscr{P}_2(\mathbb{R}) \cap W^{1,2}(\mathbb{R})$, one has*

$$\|u-u_\infty\|_{L^2}^2 \le C'\int_{\mathbb{R}} u((\log u + W_\varepsilon)_x)^2\,\mathrm{d}x. \tag{II.3.9}$$

PROOF. Define for brevity

$$\mathcal{F}(u) := \int_{\mathbb{R}} u((\log u + W_\varepsilon)_x)^2\,\mathrm{d}x.$$

Thanks to λ_ε-convexity of $\mathcal{L}_u$ and (I.2.9), we have

$$\mathcal{L}_u(u) \le \frac{1}{2\lambda_\varepsilon}\mathcal{F}(u), \tag{II.3.10}$$

and u_∞ satisfies $\log u_\infty + W_\varepsilon = \text{const.}$ on $\mathbb{R}$. With Taylor's theorem and the fact that u and u_∞ have the same mass, we conclude that

$$\mathcal{L}_u(u) = \frac{1}{2}\int_{\mathbb{R}}(u-u_\infty)^2\frac{1}{u_*}\,\mathrm{d}x, \tag{II.3.11}$$

where $u_*(x)$ is between $u(x)$ and $u_\infty(x)$.

Using integration by parts, one obtains that

$$\begin{aligned}\int_{\mathbb{R}} u(\log u)_x^2 \,\mathrm{d}x &= \mathcal{F}(u) - \int_{\mathbb{R}} u\left[2(\log u)_x(W_\varepsilon)_x + (W_\varepsilon)_x^2\right] \,\mathrm{d}x \\ &\leq \mathcal{F}(u) - 2\int_{\mathbb{R}} u_x(W_\varepsilon)_x \,\mathrm{d}x = \mathcal{F}(u) + 2\int_{\mathbb{R}} u(W_\varepsilon)_{xx} \,\mathrm{d}x \leq \mathcal{F}(u) + 2\|(W_\varepsilon)_{xx}\|_{L^\infty}.\end{aligned} \tag{II.3.12}$$

With the Sobolev inequality (II.3.3), we get

$$\|u-u_\infty\|_{L^2}^2 \leq 2\|\sqrt{u}\|_{L^4}^4 + 2\|\sqrt{u_\infty}\|_{L^4}^4 \leq 2C(\|(\sqrt{u})_x\|_{L^2} + \|(\sqrt{u_\infty})_x\|_{L^2}),$$

which can be estimated with (II.3.12) as

$$\|u-u_\infty\|_{L^2}^2 \leq C\sqrt{\mathcal{F}(u)} + C_0, \tag{II.3.13}$$

for a fixed constant $C_0 > 0$. We define $C_1 := \max(1, C_0)$ and distiguish cases.
Case 1: $\mathcal{F}(u) \geq C_1$.
Then, estimate (II.3.13) immediately yields the claim:

$$\|u-u_\infty\|_{L^2}^2 \leq (C+1)\mathcal{F}(u).$$

Case 2: $\mathcal{F}(u) \leq C_1$.
In one spatial dimension, we are allowed to estimate as follows:

$$\begin{aligned}\|u\|_{L^\infty} \leq \widetilde{C}\|u\|_{W^{1,1}} &= \widetilde{C}\left(1 + \int_{\mathbb{R}} |u_x| u^{-1/2}u^{1/2} \,\mathrm{d}x\right) \\ &\leq \widetilde{C}\left[1 + \left(\int_{\mathbb{R}} \frac{u_x^2}{u} \,\mathrm{d}x\right)^{1/2}\right].\end{aligned}$$

With the help of (II.3.12), recalling that $\mathcal{F}(u) \leq C_1$, we conclude that

$$\|u\|_{L^\infty} \leq \widetilde{C}\left[1 + (\mathcal{F}(u) + 2\|(W_\varepsilon)_{xx}\|_{L^\infty})^{1/2}\right] \leq C_2, \tag{II.3.14}$$

for some fixed constant $C_2 > 0$. Since $\mathcal{F}(u_\infty) = 0 \leq C_1$, one has $\|u_\infty\|_{L^\infty} \leq C_2$ as well. The claim now follows by combining (II.3.10), (II.3.11) and (II.3.14):

$$\begin{aligned}\frac{1}{\lambda_\varepsilon}\mathcal{F}(u) &\geq \int_{\{u<u_\infty\}} \frac{1}{u_*}(u-u_\infty)^2 \,\mathrm{d}x + \int_{\{u>u_\infty\}} \frac{1}{u_*}(u-u_\infty)^2 \,\mathrm{d}x \\ &\geq \int_{\{u<u_\infty\}} \frac{1}{u_\infty}(u-u_\infty)^2 \,\mathrm{d}x + \int_{\{u>u_\infty>0\}} \frac{1}{u}(u-u_\infty)^2 \,\mathrm{d}x \geq \frac{1}{C_2}\|u-u_\infty\|_{L^2}^2.\end{aligned}$$

□

As a conclusion to this section, we now prove the central estimate leading to Theorem II.8:

Proposition II.44 (Exponential estimate for $\mathcal{L}$). *Let $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$ be a family of time-discrete approximations obtained by* (I.2.11) *which converges to a weak solution (u,v) as $\tau \searrow 0$ in the sense stated in Theorem II.6. Then, there exist $\bar{\varepsilon} > 0$ and $L > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$ and $n \in \mathbb{N}$, one has*

$$\mathcal{L}(u_\tau^n, v_\tau^n) \leq (1 + M_2\varepsilon)(\mathcal{E}(u^0, v^0) - \mathcal{E}(u_\infty, v_\infty))(1 + 2\Lambda_\varepsilon\tau)^{-n}, \tag{II.3.15}$$

with $\Lambda_\varepsilon := \min(\lambda_0, \kappa) - L\varepsilon > 0$.

Once proved, this result yields exponential convergence of $\mathcal{L}(u(t,\cdot), v(t,\cdot))$ to zero for $t \to \infty$ after passage to the continuous-time limit $\tau \searrow 0$. From this, Theorem II.8 clearly follows.
Proof. We investigate the dissipation of $\mathcal{E}$ along the (auxiliary) $\min(\lambda_\varepsilon, \kappa)$-flow $(\mathcal{U}, \mathcal{V})_{s\geq 0}$ of the geodesically $\min(\lambda_\varepsilon, \kappa)$-convex functional $\mathcal{L}$ on $\mathbf{X}$, which is associated to the evolution system

$$\mathcal{U}_s = (\mathcal{U}_x + \mathcal{U}(W_\varepsilon)_x)_x, \qquad \mathcal{V}_s = (\mathcal{V} - v_\infty)_{xx} - \kappa(\mathcal{V} - v_\infty),$$

together with the initial conditions $(\mathcal{U}(0,\cdot),\mathcal{V}(0,\cdot)) = (u_\tau^n, v_\tau^n)$. First, by elementary calculations, we obtain using the decomposition (II.3.8):

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V}) \leq \left(\frac{\varepsilon}{2}-1\right)\int_{\mathbb{R}} \mathcal{U}((\log\mathcal{U}+W^\varepsilon)_x)^2\,\mathrm{d}x + \frac{\varepsilon}{2}\int_{\mathbb{R}}\mathcal{U}(\mathcal{V}-v_\infty)_x^2\,\mathrm{d}x$$
$$+\frac{\varepsilon}{2}\|\mathcal{U}-u_\infty\|_{L^2}^2 + \left(\frac{\varepsilon}{2}-1\right)\|(\mathcal{V}-v_\infty)_{xx}-\kappa(\mathcal{V}-v_\infty)\|_{L^2}^2.$$

The third term can be controlled by the first one using (II.3.9), whereas the second term is controllable by the fourth term using the inequality $\|\eta\mu_x^2\|_{L^1} \leq C\|\eta\|_{L^1}\|\mu\|_{W^{2,2}}^2$ which is valid in one spatial dimension. Taking into account the properties of $\mathcal{L}$ from Proposition II.42, we end up with

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V}) \geq 2(1-\varepsilon M)\min(\lambda_\varepsilon,\kappa)\mathcal{L}(\mathcal{U},\mathcal{V}), \tag{II.3.16}$$

for some constant $M > 0$ if ε is sufficiently small. The application of the flow interchange lemma (Theorem I.5) eventually yields with (II.3.16):

$$[1+2\tau(1-\varepsilon M)\min(\lambda_\varepsilon,\kappa)]\mathcal{L}(u_\tau^n, v_\tau^n) \leq \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}).$$

By iteration of this estimate and Proposition II.42(d), the desired estimate (II.3.15) follows. □

CHAPTER II.4

A system of Poisson-Nernst-Planck type with quadratic diffusion

The results in this chapter are contained in the article [188]. The methods of proof to deduce the exponential convergence to equilibrium are strongly related to those of Section II.2.2 (see also [189]) and require the existence of a weak solution constructed via the minimizing movement scheme. Using similar methods as those in Chapter II.2, Kinderlehrer, Monsaingeon and Xu proved the existence of such weak solutions in [112]. We recall their main results in the following section.

II.4.1. Preliminaries

In this section, we restate the results proved by Kinderlehrer *et al.* in [112] adapted to our specific setting. First, recall that the state space here is $\mathbf{X} = \mathscr{P}_2(\mathbb{R}^3) \times \mathscr{P}_2(\mathbb{R}^3)$ endowed with the canonical product distance

$$\mathbf{d}((u,v),(\widetilde{u},\widetilde{v})) := \sqrt{\mathbf{W}_2(u,\widetilde{u})^2 + \mathbf{W}_2(v,\widetilde{v})^2},$$

and the free energy is defined as

$$\mathcal{E}(u,v) := \begin{cases} \int_{\mathbb{R}^3} (u^2 + v^2 + uU + vV + \frac{\varepsilon}{2}|\mathrm{D}\psi|^2)\,\mathrm{d}x & \text{if } (u,v) \in L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3), \\ +\infty & \textit{otherwise}. \end{cases}$$

Proposition II.45 (Minimizing movement scheme [112, Prop. 3.3]). *Let $\tau > 0$ and an initial datum $(u_0, v_0) \in \mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$ be given. Then, the sequence $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$ defined by the minimizing movement scheme* (I.2.11) *is well-defined with $(u_\tau^n, v_\tau^n) \in \mathbf{X} \cap (W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3))$ for all $n \in \mathbb{N}$. By definition, the sequence $(\mathcal{E}(u_\tau^n, v_\tau^n))_{n\in\mathbb{N}}$ is nonincreasing.*

The following main result of [112] about the existence of nonnegative solutions to (II.1.7) is at the basis of our subsequent analysis:

Theorem II.46 (Existence of solutions [112, Thm. 2]). *Let $\varepsilon > 0$ and U, V as mentioned above. Define, for initial conditions $(u_0, v_0) \in \mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$ and each $\tau > 0$ a discrete solution (u_τ, v_τ) by* (I.2.11)&(I.2.12). *Then, there exists a sequence $\tau_k \searrow 0$ and a map $(u,v) : [0,\infty) \times \mathbb{R}^3 \to [0,\infty]^2$ such that for each $t > 0$, $u_{\tau_k}(t,\cdot) \rightharpoonup u(t,\cdot)$ and $v_{\tau_k}(t,\cdot) \rightharpoonup v(t,\cdot)$, both narrowly in $\mathscr{P}_2(\mathbb{R}^3)$ as $k \to \infty$. Moreover, (u,v) is a solution to* (II.1.7) *in the sense of distributions, it attains the initial condition and one has for each $T > 0$:*

$$u, v \in C^{1/2}([0,T]; (\mathscr{P}_2(\mathbb{R}^3), \mathbf{W}_2)) \cap L^\infty([0,T]; L^2(\mathbb{R}^3)) \cap L^2([0,T]; W^{1,2}(\mathbb{R}^3)),$$
$$\mathcal{E}(u(T,\cdot), v(T,\cdot)) \le \mathcal{E}(u_0, v_0).$$

II.4.2. The equilibrium state

In this section, we prove Theorem II.9.

Proof. Existence: Trivially, $\mathcal{E}$ is bounded from below. Hence, there exists a minimizing sequence $(u_k, v_k)_{k\in\mathbb{N}}$ in $\mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$ with $\lim_{k\to\infty} \mathcal{E}(u_k, v_k) = \inf_{(u,v)\in\mathbf{X}} \mathcal{E}(u,v)$. Thus, we have for some $C > 0$ that $\|u_k\|_{L^2} \le C$, $\|v_k\|_{L^2} \le C$ for all $k \in \mathbb{N}$. Moreover, using the λ_0-convexity of U and

V on $\mathbb{R}^3$, one obtains $\sup_{k\in\mathbb{N}} \mathfrak{m}_2(u_k) < \infty$ and $\sup_{k\in\mathbb{N}} \mathfrak{m}_2(v_k) < \infty$ with the help of the elementary estimates $U(x) - U(x^U_{\min}) \geq \frac{\lambda_0}{4}|x|^2 - \frac{\lambda_0}{2}|x^U_{\min}|^2$ and $V(x) - V(x^V_{\min}) \geq \frac{\lambda_0}{4}|x|^2 - \frac{\lambda_0}{2}|x^V_{\min}|^2$ (with the unique minimizers $x^U_{\min}, x^V_{\min}$ of U and V on $\mathbb{R}^3$, respectively). We infer with the Prokhorov and Banach-Alaoglu theorems that there exists a subsequence (non-relabelled) and a limit $(u_\infty, v_\infty) \in \mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$ such that $u_k \rightharpoonup u_\infty$ and $v_k \rightharpoonup v_\infty$ both narrowly as probability measures and weakly in $L^2(\mathbb{R}^3)$, as $k \to \infty$. With respect to these convergences, $\mathcal{E}$ is lower semicontinuous. In fact, this is obvious for the quadratic and linear terms in $\mathcal{E}$ since U and V grow quadratically. For the last term containing the Dirichlet energy $\frac{1}{2}\|\mathrm{D}\psi\|^2_{L^2}$, we refer to [112, Prop. 6.1] for a result on lower semicontinuity w.r.t. weak $L^1(\mathbb{R}^3)$ convergence. Hence, it follows that (u_∞, v_∞) is indeed a minimizer of $\mathcal{E}$ on $\mathbf{X}$ and hence also a steady state of (II.1.7).

Uniqueness: We claim that $\mathcal{E}$ is uniformly convex with respect to the flat distance induced by the product norm $\|\cdot\|_{L^2\times L^2}$, which implies the uniqueness of minimizers.

For all $(u,v), (u',v') \in \mathbf{X} \cap (L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$ and all $s \in [0,1]$, we have, thanks to

$$\int_{\mathbb{R}^3} |\mathrm{D}(\mathbf{G} * w)|^2 \,\mathrm{d}x = \int_{\mathbb{R}^3} (\mathbf{G} * w) w \,\mathrm{d}x = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} w(x)\mathbf{G}(x-y)w(y)\,\mathrm{d}x\,\mathrm{d}y \quad \forall w \in \mathbf{X} \cap L^2(\mathbb{R}^3) \tag{II.4.1}$$

that

$$\begin{aligned} &\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0} \mathcal{E}(u + s(u'-u), v + s(v'-v)) \\ &\quad = \int_{\mathbb{R}^3} \Big[2(u'-u)^2 + 2(v'-v)^2 + \varepsilon((u'-u)-(v'-v))\mathbf{G} * ((u'-u)-(v'-v))\Big]\,\mathrm{d}x \\ &\quad \geq 2\|(u'-u, v'-v)\|^2_{L^2\times L^2}, \end{aligned}$$

so $\mathcal{E}$ is 2-convex w.r.t. the distance induced by $\|\cdot\|_{L^2\times L^2}$.

Euler-Lagrange equations: Since (u_∞, v_∞) is the minimizer of $\mathcal{E}$, the following variational inequality holds:

$$\begin{aligned} 0 &\leq \frac{\mathrm{d}^+}{\mathrm{d}s}\bigg|_{s=0} \mathcal{E}(u_\infty + s\widetilde{u}, v_\infty + s\widetilde{v}) \\ &= \int_{\mathbb{R}^3} [2u_\infty + U + \varepsilon\mathbf{G} * (u_\infty - v_\infty)]\,\widetilde{u}\,\mathrm{d}x + \int_{\mathbb{R}^3} [2v_\infty - V - \varepsilon\mathbf{G} * (u_\infty - v_\infty)]\,\widetilde{v}\,\mathrm{d}x, \end{aligned} \tag{II.4.2}$$

for all $\widetilde{u}, \widetilde{v}$ such that both $u_\infty + \widetilde{u} \geq 0$ and $v_\infty + \widetilde{v} \geq 0$ on $\mathbb{R}^3$, and $\int_{\mathbb{R}^3} \widetilde{u}\,\mathrm{d}x = 0 = \int_{\mathbb{R}^3} \widetilde{v}\,\mathrm{d}x$. In order to prove (II.1.9), we set $\widetilde{v} = 0$. Let $\phi : \mathbb{R}^3 \to \mathbb{R}$ be such that $\int_{\mathbb{R}^3} \phi\,\mathrm{d}x \leq 1$ and $\phi + u_\infty \geq 0$ on $\mathbb{R}^3$. The choice

$$\widetilde{u}_\phi := \frac{1}{2}\phi - \frac{1}{2}u_\infty \int_{\mathbb{R}^3} \phi\,\mathrm{d}x$$

is admissible for $\widetilde{u}$ in (II.4.2), hence (recall the notation $\psi_\infty := \mathbf{G} * (u_\infty - v_\infty)$)

$$0 \leq \int_{\mathbb{R}^3} (2u_\infty + U + \varepsilon\psi_\infty - C_u)\phi\,\mathrm{d}x, \tag{II.4.3}$$

with

$$C_u := \int_{\mathbb{R}^3} (2u_\infty^2 + Uu_\infty + \varepsilon u_\infty \psi_\infty)\,\mathrm{d}x \in \mathbb{R}.$$

If $u_\infty(x) > 0$ for some $x \in \mathbb{R}^3$, we are able to choose ϕ supported on a small neighbourhood of x and to replace by $-\phi$ in (II.4.3) and obtain

$$2u_\infty(x) + U(x) + \varepsilon\psi_\infty(x) = C_u.$$

If $u_\infty(x) = 0$ for some x, one has $U(x) - \varepsilon\psi_\infty(x) - C_u \geq 0$, and hence (II.1.9) is true in both cases. The equation for v_∞ (II.1.10) can be derived in analogy.

Properties: First, since (u_∞, v_∞) are admissible as starting condition (u^0_τ, v^0_τ) (for arbitrary $\tau > 0$) in the scheme (I.2.11), we obtain thanks to the minimizing property and Proposition II.45 that $(u_\infty, v_\infty) \in W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3)$. We now show that $\psi_\infty \in L^\infty(\mathbb{R}^3)$. To this end, let $x \in \mathbb{R}^3$ and observe at first that

$$\int_{\mathbb{B}_1(x)} \frac{|u_\infty(y) - v_\infty(y)|}{|x-y|}\,\mathrm{d}y \le \|u_\infty - v_\infty\|_{L^2} \left(\int_{\mathbb{B}_1(x)} \frac{1}{|x-y|^2}\,\mathrm{d}y\right)^{1/2} = 2\sqrt{\pi}\|u_\infty - v_\infty\|_{L^2},$$

independent of x, by Hölder's inequality and the transformation theorem. Furthermore, since $|x-y| \ge 1$ if $y \notin \mathbb{B}_1(x)$ and $\|u_\infty\|_{L^1} = 1 = \|v_\infty\|_{L^1}$, we get

$$\int_{\mathbb{R}^3 \setminus \mathbb{B}_1(x)} \frac{|u_\infty(y) - v_\infty(y)|}{|x-y|}\,\mathrm{d}y \le \|u_\infty - v_\infty\|_{L^1} \sup_{y \notin \mathbb{B}_1(x)} |x-y|^{-1} \le 2.$$

Putting both parts together, we see that $\sup_{x \in \mathbb{R}^3} |\psi_\infty(x)| < \infty$. In view of (II.1.9)&(II.1.10), $\psi_\infty \in L^\infty(\mathbb{R}^3)$ implies that u_∞ and v_∞ have compact support since U and V grow quadratically as $|x| \to \infty$. By classical results on solutions to Poisson's equation [121, Thm. 10.2], we then infer that $\psi_\infty \in C^{0,\alpha}(\mathbb{R}^3)$ for all $\alpha \in (0,1)$, since by the Gagliardo-Nirenberg-Sobolev inequality, one has $(u_\infty, v_\infty) \in L^6(\mathbb{R}^3) \times L^6(\mathbb{R}^3)$. Hence, using (II.1.9)&(II.1.10) again, we conclude that u_∞ and v_∞ also are Hölder continuous. By elliptic regularity for Poisson's kernel [121, Thm. 10.3], it follows that $\psi_\infty \in C^{2,\alpha}(\mathbb{R}^3)$. □

II.4.3. Auxiliary entropy and dissipation

In this section, we define a suitable geodesically convex auxiliary entropy $\mathcal{L}$ and derive the dissipation of the driving energy $\mathcal{E}$ along the gradient flow $\mathsf{S}^{\mathcal{L}}$ of $\mathcal{L}$.

Let $\mathcal{L} : \mathbf{X} \to \mathbb{R}_\infty$ be defined via

$$\mathcal{L}(u,v) := \begin{cases} \int_{\mathbb{R}^3} \left[u^2 - u_\infty^2 + v^2 - v_\infty^2 + (u-u_\infty)U + (v-v_\infty)V + \varepsilon(u-u_\infty)\psi_\infty - \varepsilon(v-v_\infty)\psi_\infty\right]\,\mathrm{d}x \\ \qquad\qquad \text{if } (u,v) \in L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3), \\ +\infty \qquad\qquad \text{otherwise.} \end{cases}$$

Obviously, $\mathcal{L}$ is proper and lower semicontinuous on $(\mathbf{X}, \mathbf{d})$.

Proposition II.47 (Properties of $\mathcal{L}$). *There exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$, the following statements hold:*

(a) There exists $L > 0$ such that $\mathcal{L}$ is geodesically λ_ε-convex w.r.t. $\mathbf{d}$, where $\lambda_\varepsilon := \lambda_0 - L\varepsilon > 0$.

(b) The following holds for all $(u,v) \in \mathbf{X} \cap (W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3))$:

$$\begin{aligned} \|u-u_\infty\|_{L^2}^2 + \|v-v_\infty\|_{L^2}^2 &\le \mathcal{L}(u,v) \\ &\le \frac{1}{2\lambda_\varepsilon} \int_{\mathbb{R}^3} \left[u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2 + v|\mathrm{D}(2v + V - \varepsilon\psi_\infty)|^2\right]\,\mathrm{d}x. \end{aligned} \tag{II.4.4}$$

(c) There exists a constant $K > 0$ independent of ε such that for all $(u,v) \in \mathbf{X}$:

$$\mathcal{L}(u,v) \le \mathcal{E}(u,v) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon. \tag{II.4.5}$$

Proof. (a) In view of Theorem I.14, as $\mathcal{L}$ is decoupled in its arguments u and v, it suffices to prove that there exists $C > 0$ such that $\|\mathrm{D}^2\psi_\infty\|_{L^\infty} \le C$ for all sufficiently small $\varepsilon > 0$. Let $R > 0$ such that $\operatorname{supp} u_\infty \cup \operatorname{supp} v_\infty \subset \mathbb{B}_R(0)$. Since $\psi_\infty \in C^2(\mathbb{R}^3)$ thanks to Theorem II.9, we have $\sup\left\{|\partial_{x_i}\partial_{x_j}\psi_\infty(x)| : x \in \overline{\mathbb{B}_{R+1}(0)}\right\} < \infty$ for each pair $(i,j) \in \{1,2,3\}^2$. Consider now $x \notin \overline{\mathbb{B}_{R+1}(0)}$. One obtains for $z \neq 0$ that

$$\partial_{z_i}\partial_{z_j}\mathbf{G}(z) = \frac{1}{4\pi|z|^3}\left(\frac{3z_iz_j}{|z|^2} - \delta_{ij}\right),$$

where δ_{ij} denotes Kronecker's delta. So, using a linear transformation,

$$|\partial_{x_i}\partial_{x_j}\psi_\infty(x)| = \left|\int_{\mathbb{B}_R(x)} \partial_{z_i}\partial_{z_j}\mathbf{G}(z)(u_\infty(x-z)-v_\infty(x-z))\,\mathrm{d}z\right| \le \frac{1}{3}R^3\|u_\infty - v_\infty\|_{L^\infty},$$

since for all $z \in \mathbb{B}_R(x)$, one has $|z| > 1$ by definition of x. Hence, the desired uniform estimate on $\mathrm{D}^2\psi_\infty$ is proved.

(b) The upper estimate is a straightforward consequence of λ_ε-convexity of $\mathcal{L}$ and the structure of its Wasserstein subdifferential w.r.t. u and v, respectively (see e.g. [4, Lemma 10.4.1]), in combination with (I.2.9). For the lower estimate, we observe that

$$\mathcal{L}(u,v) = \int_{\mathbb{R}^3} \Big[(u-u_\infty)^2 + (v-v_\infty)^2 + (u-u_\infty)(2u_\infty + U + \varepsilon\psi_\infty) + (v-v_\infty)(2v_\infty + V - \varepsilon\psi_\infty)\Big]\,\mathrm{d}x.$$

We prove that $\int_{\mathbb{R}^3}(u-u_\infty)(2u_\infty+U+\varepsilon\psi_\infty)\,\mathrm{d}x \ge 0$. Since the term above involving the v component can be treated in the same way, the claim then follows. Using (II.1.9), we obtain

$$\int_{\mathbb{R}^3}(u-u_\infty)(2u_\infty+U+\varepsilon\psi_\infty)\,\mathrm{d}x = \int_{\{C_u-U-\varepsilon\psi_\infty>0\}}(u-u_\infty)C_u\,\mathrm{d}x + \int_{\{C_u-U-\varepsilon\psi_\infty\le 0\}} u(U+\varepsilon\psi_\infty)\,\mathrm{d}x$$
$$= C_u\int_{\mathbb{R}^3}(u-u_\infty)\,\mathrm{d}x + \int_{\{C_u-U-\varepsilon\psi_\infty\le 0\}} u(U+\varepsilon\psi_\infty - C_u)\,\mathrm{d}x \ge 0,$$

since u and u_∞ have equal mass (hence the first term is equal to zero) and the integrand of the second integral is nonnegative on the domain of integration.

(c) One has for all $(u,v) \in \mathbf{X} \cap (L^2(\mathbb{R}^3)\times L^2(\mathbb{R}^3))$:

$$\frac{1}{\varepsilon}(\mathcal{L}(u,v) - \mathcal{E}(u,v) + \mathcal{E}(u_\infty,v_\infty)) = \int_{\mathbb{R}^3}\left[(u-u_\infty)\psi_\infty - (v-v_\infty)\psi_\infty - \frac{1}{2}|\mathrm{D}\psi|^2 + \frac{1}{2}|\mathrm{D}\psi_\infty|^2\right]\,\mathrm{d}x$$
$$\le \int_{\mathbb{R}^3}\psi_\infty(u - v - \frac{1}{2}u_\infty + \frac{1}{2}v_\infty)\,\mathrm{d}x \le 3\|\psi_\infty\|_{L^\infty} \le K,$$

thanks to (II.4.1) and Theorem II.9.

□

According to Example I.15, the λ_ε-contractive flow $\mathsf{S}^{\mathcal{L}} =: (\mathcal{U},\mathcal{V})$ is characterized by

$$\partial_s\mathcal{U} = \mathrm{div}\left[\mathcal{U}\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)\right], \qquad \partial_s\mathcal{V} = \mathrm{div}\left[\mathcal{V}\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)\right]. \tag{II.4.6}$$

Now, we derive the central *a priori* estimate on the discrete solution:

Proposition II.48 (Dissipation of $\mathcal{E}$ along $\mathsf{S}^{\mathcal{L}}$). *Let $\tau > 0$ and let $(u_\tau^n, v_\tau^n)_{n\in\mathbb{N}}$ be the sequence defined via the minimizing movement scheme* (I.2.11). *Then, for all $n \in \mathbb{N}$:*

$$\mathcal{L}(u_\tau^n, v_\tau^n) + \tau\mathcal{D}(u_\tau^n, v_\tau^n) \le \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}), \tag{II.4.7}$$

the dissipation *being given by*

$$\mathcal{D}(u,v) := \left(1-\frac{\varepsilon}{2}\right)\int_{\mathbb{R}^3}\left(u|\mathrm{D}(2u+U+\varepsilon\psi_\infty)|^2 + v|\mathrm{D}(2v+V-\varepsilon\psi_\infty)|^2\right)\mathrm{d}x$$
$$-\frac{\varepsilon}{2}\int_{\mathbb{R}^3}(u+v)|\mathrm{D}(\psi-\psi_\infty)|^2\,\mathrm{d}x. \tag{II.4.8}$$

Proof. To justify the calculations below, we regularize the flow given by (II.4.6). Define, for $\nu > 0$ and $(u,v) \in \mathbf{X}\cap(L^2(\mathbb{R}^3)\times L^2(\mathbb{R}^3))$ the regularized functional

$$\mathcal{L}_\nu(u,v) := \mathcal{L}(u,v) + \nu\mathcal{H}(u) + \nu\mathcal{H}(v), \quad \text{with } \textit{Boltzmann's entropy } \mathcal{H}(w) := \int_{\mathbb{R}^3} w\log w\,\mathrm{d}x,$$

which is finite on $\mathscr{P}_2(\mathbb{R}^3)\cap L^2(\mathbb{R}^3)$ (cf. Lemma II.18). Furthermore, by Example I.15, $\mathcal{H}$ is 0-convex along generalized geodesics in $\mathscr{P}_2(\mathbb{R}^3)$, so $\mathcal{L}_\nu$ is geodesically λ_ε-convex w.r.t. $\mathbf{d}$ and the associated evolution equation to its λ_ε-flow $(\mathcal{U},\mathcal{V})$ is the strictly parabolic, decoupled system

$$\partial_s\mathcal{U} = \nu\Delta\mathcal{U} + \operatorname{div}\left[\mathcal{U}\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)\right], \qquad \partial_s\mathcal{V} = \nu\Delta\mathcal{V} + \operatorname{div}\left[\mathcal{V}\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)\right]. \tag{II.4.9}$$

Let $(u,v) \in \mathbf{X}\cap(W^{1,2}(\mathbb{R}^3)\times W^{1,2}(\mathbb{R}^3))$. At least for small $s>0$, system (II.4.9) has a smooth and nonnegative solution $(\mathcal{U},\mathcal{V})$ such that $(\mathcal{U}(s,\cdot),\mathcal{V}(s,\cdot)) \to (u,v)$ both strongly in $L^2(\mathbb{R}^3)\times L^2(\mathbb{R}^3)$ and $\mathbf{d}$, as well as weakly in $W^{1,2}(\mathbb{R}^3)\times W^{1,2}(\mathbb{R}^3)$, for $s\searrow 0$. Moreover, this flow operator can be identified with the λ_ε-flow associated to $\mathcal{L}_\nu$ (see e.g. [4, Thm. 11.2.8]). Then, writing $\Psi := \mathbf{G}*(\mathcal{U}-\mathcal{V})$ for brevity:

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V}) = &-\int_{\mathbb{R}^3}[2\mathcal{U}+U+\varepsilon\Psi]\operatorname{div}\left[\nu\mathrm{D}\mathcal{U}+\mathcal{U}\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)\right]\,\mathrm{d}x\\
&-\int_{\mathbb{R}^3}[2\mathcal{V}+V-\varepsilon\Psi]\operatorname{div}\left[\nu\mathrm{D}\mathcal{V}+\mathcal{V}\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)\right]\,\mathrm{d}x.
\end{aligned}$$

We first focus on the viscosity terms and obtain, using that $(\mathcal{U},\mathcal{V})\in\mathbf{X}$:

$$\begin{aligned}
&-\int_{\mathbb{R}^3}\left([2\mathcal{U}+U+\varepsilon\Psi]\Delta\mathcal{U}+[2\mathcal{V}+V-\varepsilon\Psi]\Delta\mathcal{V}\right)\mathrm{d}x\\
&=\int_{\mathbb{R}^3}\left(2|\mathrm{D}\mathcal{U}|^2+2|\mathrm{D}\mathcal{V}|^2-\mathcal{U}\Delta U-\mathcal{V}\Delta V-\varepsilon(\mathcal{U}-\mathcal{V})\Delta\Psi\right)\mathrm{d}x\\
&=2\|\mathrm{D}\mathcal{U}\|_{L^2}^2+2\|\mathrm{D}\mathcal{V}\|_{L^2}^2-\int_{\mathbb{R}^3}(\mathcal{U}\Delta U+\mathcal{V}\Delta V)\,\mathrm{d}x+\varepsilon\|\mathcal{U}-\mathcal{V}\|_{L^2}^2 \geq -\|\Delta U\|_{L^\infty}-\|\Delta V\|_{L^\infty}.
\end{aligned}$$

The remaining terms can be rewritten as

$$\begin{aligned}
&-\int_{\mathbb{R}^3}[2\mathcal{U}+U+\varepsilon\Psi]\operatorname{div}\left[\mathcal{U}\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)\right]\,\mathrm{d}x-\int_{\mathbb{R}^3}[2\mathcal{V}+V-\varepsilon\Psi]\operatorname{div}\left[\mathcal{V}\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)\right]\,\mathrm{d}x\\
&=\int_{\mathbb{R}^3}\mathcal{U}|\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)|^2\,\mathrm{d}x+\int_{\mathbb{R}^3}\mathcal{V}|\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)|^2\,\mathrm{d}x\\
&\quad+\varepsilon\int_{\mathbb{R}^3}\mathcal{U}\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)\cdot\mathrm{D}(\Psi-\psi_\infty)\,\mathrm{d}x-\varepsilon\int_{\mathbb{R}^3}\mathcal{V}\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)\cdot\mathrm{D}(\Psi-\psi_\infty)\,\mathrm{d}x\\
&\geq\left(1-\frac{\varepsilon}{2}\right)\int_{\mathbb{R}^3}\left(\mathcal{U}|\mathrm{D}(2\mathcal{U}+U+\varepsilon\psi_\infty)|^2+\mathcal{V}|\mathrm{D}(2\mathcal{V}+V-\varepsilon\psi_\infty)|^2\right)\mathrm{d}x-\frac{\varepsilon}{2}\int_{\mathbb{R}^3}(\mathcal{U}+\mathcal{V})|\mathrm{D}(\Psi-\psi_\infty)|^2\,\mathrm{d}x,
\end{aligned}$$

using Young's inequality in the final step. All in all, we arrive at

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathcal{U},\mathcal{V})\geq\mathcal{D}(\mathcal{U},\mathcal{V})-\nu\left(\|\Delta U\|_{L^\infty}+\|\Delta V\|_{L^\infty}\right).$$

Observing that the terms appearing in $\mathcal{D}$ are lower semicontinuous w.r.t. the convergence of $(\mathcal{U},\mathcal{V})\to(u,v)$ above, we obtain after passage to the limits $s\searrow 0$ and $\nu\searrow 0$ that $\mathrm{D}^{\mathcal{L}}\mathcal{E}(u,v)\geq\mathcal{D}(u,v)$. The application of the flow interchange lemma (Theorem I.5) completes the proof of (II.4.7). □

The remaining task is to establish appropriate bounds on the dissipation $\mathcal{D}(u_\tau^n,v_\tau^n)$ in terms of $\mathcal{L}(u_\tau^n,v_\tau^n)$ in order to apply a discrete Gronwall lemma and to conclude exponential convergence. Note that, in view of (I.2.9), it will be enough to control the second part of $\mathcal{D}(u_\tau^n,v_\tau^n)$.

II.4.4. Convergence to equilibrium

In this section, we complete the proof of Theorem II.10. Our strategy is as follows: First, we derive a uniform bound (independent of ε and the initial condition) on the auxiliary entropy $\mathcal{L}$ for sufficiently large times. This brings us into position to prove a refined estimate on the dissipation $\mathcal{D}$ strong enough to infer exponential convergence of $\mathcal{L}$ to zero. In the following, for $\tau>0$, we denote by $(u_\tau^n,v_\tau^n)_{n\in\mathbb{N}}$ a sequence given by the minimizing movement scheme (I.2.11).

II.4.4.1. Boundedness of auxiliary entropy

We first need an additional estimate for the dissipation terms in (II.4.8) (compare with Lemma II.30):

Lemma II.49. *There exists a constant $\theta > 0$ such that the following holds for all $\varepsilon \in (0, \varepsilon_0)$ and all $(u, v) \in \mathbf{X} \cap (W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3))$:*

$$\begin{aligned} \|u\|_{L^3}^4 &\le \theta \left(1 + \int_{\mathbb{R}^3} u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2 \,\mathrm{d}x\right), \\ \|v\|_{L^3}^4 &\le \theta \left(1 + \int_{\mathbb{R}^3} v|\mathrm{D}(2v + V - \varepsilon\psi_\infty)|^2 \,\mathrm{d}x\right), \end{aligned} \tag{II.4.10}$$

with the convention that the respective right-hand side is set equal to $+\infty$ if $u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2$ or $v|\mathrm{D}(2v + V - \varepsilon\psi_\infty)|^2$ is not integrable.

Proof. We shall prove the statement for u; the other one can be shown analogously. We assume that the r.h.s. is finite. Expanding the square and integrating by parts, one has

$$\int_{\mathbb{R}^3} u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2 \,\mathrm{d}x = \int_{\mathbb{R}^3} \left(\frac{16}{9}|\mathrm{D}u^{3/2}|^2 - 2u^2\Delta(U + \varepsilon\psi_\infty) + u|\mathrm{D}(U + \varepsilon\psi_\infty)|^2\right) \mathrm{d}x.$$

Since ΔU and $\Delta\psi_\infty = v_\infty - u_\infty$ are essentially bounded, we obtain

$$\frac{16}{9}\|\mathrm{D}u^{3/2}\|_{L^2}^2 \le \int_{\mathbb{R}^3} u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2 \,\mathrm{d}x + C\|u\|_{L^2}^2,$$

for some constant $C > 0$. By the triangle and Young inequalities, $\|u\|_{L^2}^2 \le 2\|u_\infty\|_{L^2}^2 + 2\|u - u_\infty\|_{L^2}^2$. For small $\varepsilon > 0$, we can use (II.4.4) and arrive at

$$\frac{16}{9}\|\mathrm{D}u^{3/2}\|_{L^2}^2 \le \int_{\mathbb{R}^3} \left(1 + \frac{C}{\lambda_\varepsilon}\right) u|\mathrm{D}(2u + U + \varepsilon\psi_\infty)|^2 \,\mathrm{d}x + 2C\|u_\infty\|_{L^2}^2.$$

On the other hand, with the L^p-interpolation and Gagliardo-Nirenberg-Sobolev inequalities, we have (recall $\|u\|_{L^1} = 1$):

$$\|u\|_{L^3} \le \|u\|_{L^9}^{3/4}\|u\|_{L^1}^{1/4} = \|u^{3/2}\|_{L^6}^{1/2} \le C'\|\mathrm{D}u^{3/2}\|_{L^2}^{1/2}.$$

Raising to the fourth power, we end up with (II.4.10). □

We now derive a uniform bound on $\mathcal{L}$ for large times.

Proposition II.50 (Boundedness of $\mathcal{L}$). *(a) There exist $\varepsilon_1 \in (0, \varepsilon_0)$, $L' > 0$ and $M > 0$ such that for all $\varepsilon \in (0, \varepsilon_1)$, all $\tau > 0$ and all $n \in \mathbb{N}$:*

$$(1 + 2\lambda_\varepsilon'\tau)\mathcal{L}(u_\tau^n, v_\tau^n) \le \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}) + \tau\varepsilon M, \tag{II.4.11}$$

where $\lambda_\varepsilon' := \lambda_0 - L'\varepsilon > 0$.

(b) Define, with the quantities from (a) and fixed, but arbitrary $\delta > 0$:

$$M' := \frac{M\varepsilon_1}{2(\lambda_0 - L'\varepsilon_1)} > 0 \quad \textit{and} \quad T_0 := \max\left(0, \frac{1 + 2\delta}{2\lambda_\varepsilon'} \log\frac{\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1}{M'}\right) \ge 0,$$

where $K > 0$ is the constant from (II.4.5)*. Then, there exists $\overline{\tau} > 0$ such that for all $\varepsilon \in (0, \varepsilon_1)$, $\tau \in (0, \overline{\tau}]$ and $n \in \mathbb{N}$ with $n\tau \ge T_0$, one has*

$$\mathcal{L}(u_\tau^n, v_\tau^n) \le 2M'. \tag{II.4.12}$$

Proof. (a) We first estimate the last term appearing in $\mathcal{D}(u, v)$ from (II.4.8). By Hölder's inequality, for $(u, v) \in \mathbf{X} \cap (W^{1,2}(\mathbb{R}^3) \times W^{1,2}(\mathbb{R}^3))$:

$$\int_{\mathbb{R}^3} (u + v)|\mathrm{D}\psi|^2 \,\mathrm{d}x \le (\|u\|_{L^{3/2}} + \|v\|_{L^{3/2}})\|\mathrm{D}\psi\|_{L^6}^2. \tag{II.4.13}$$

The term involving the gradient of ψ can be treated with the Hardy-Littlewood-Sobolev inequality (see e.g. [121, Thm. 4.3], [112, Lemma 3.1]) which is applicable for Poisson's kernel $\mathbf{G}$:

$$\|D\psi\|_{L^6}^2 \leq C\|u-v\|_{L^2}^2 \leq 2C\|u\|_{L^2}^2 + 2C\|v\|_{L^2}^2, \tag{II.4.14}$$

for some constant $C > 0$. Combining (II.4.13)&(II.4.14) and using $\|u\|_{L^1} = 1 = \|v\|_{L^1}$ again, the L^p-interpolation inequality yields for some $\beta, \beta' \in (0,1)$:

$$\begin{aligned}
&\int_{\mathbb{R}^3}(u+v)|D(\psi-\psi_\infty)|^2\,dx \leq 2\int_{\mathbb{R}^3}(u+v)|D\psi|^2\,dx + 2\int_{\mathbb{R}^3}(u+v)|D\psi_\infty|^2\,dx\\
&\leq 4C\big(\|u\|_{L^3}^{\beta}\|u\|_{L^3}^{2\beta'} + \|u\|_{L^3}^{\beta}\|v\|_{L^3}^{2\beta'} + \|v\|_{L^3}^{\beta}\|u\|_{L^3}^{2\beta'} + \|v\|_{L^3}^{\beta}\|v\|_{L^3}^{2\beta'}\big) + 2\int_{\mathbb{R}^3}(u+v)|D\psi_\infty|^2\,dx\\
&\leq C'(\|u\|_{L^3}^4 + \|v\|_{L^3}^4 + 1),
\end{aligned}$$

for some $C' > 0$, by Young's inequality and thanks to finiteness of $\|D\psi_\infty\|_{L^\infty}$. Now, we apply (II.4.10) and obtain

$$\mathcal{D}(u,v) \geq \Big(1 - \frac{\varepsilon}{2}(1+C'')\Big)\int_{\mathbb{R}^3}\big(u|D(2u+U+\varepsilon\psi_\infty)|^2 + v|D(2v+V-\varepsilon\psi_\infty)|^2\big)\,dx - \varepsilon M,$$

for suitable $C'' > 0$ and $M > 0$. For $\varepsilon < \frac{2}{1+C''}$, we further conclude by (II.4.4) that

$$\mathcal{D}(u,v) \geq 2\lambda_\varepsilon\Big(1-\frac{\varepsilon}{2}(1+C'')\Big)\mathcal{L}(u,v) - \varepsilon M.$$

Insertion into (II.4.7) yields (a).

(b) We first prove the following explicit estimate for all $\tau > 0$ and $n \in \mathbb{N}\cup\{0\}$ by induction on n:

$$\mathcal{L}(u_\tau^n, v_\tau^n) \leq (\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)(1+2\lambda'_\varepsilon\tau)^{-n} + \frac{M\varepsilon}{2\lambda'_\varepsilon}(1-(1+2\lambda'_\varepsilon\tau)^{-n}). \tag{II.4.15}$$

Indeed, the claim holds for $n=0$ thanks to (II.4.5). If it holds for an arbitrary $n \in \mathbb{N}\cup\{0\}$, we obtain with (II.4.11):

$$\begin{aligned}
&\mathcal{L}(u_\tau^{n+1}, v_\tau^{n+1}) \leq (1+2\lambda'_\varepsilon\tau)^{-1}\mathcal{L}(u_\tau^n,v_\tau^n) + (1+2\lambda'_\varepsilon\tau)^{-1}\tau\varepsilon M\\
&\leq (1+2\lambda'_\varepsilon\tau)^{-(n+1)}(\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)\\
&\quad + \frac{M\varepsilon}{2\lambda'_\varepsilon}(1+2\lambda'_\varepsilon\tau)^{-1}(1-(1+2\lambda'_\varepsilon\tau)^{-n}) + (1+2\lambda'_\varepsilon\tau)^{-1}\tau\varepsilon M\\
&= (\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)(1+2\lambda'_\varepsilon\tau)^{-(n+1)} + \frac{M\varepsilon}{2\lambda'_\varepsilon}(1-(1+2\lambda'_\varepsilon\tau)^{-(n+1)}).
\end{aligned}$$

Let now $\tau > 0$ and $n \in \mathbb{N}$ with $n\tau \geq T_0$. Thanks to (II.4.15), for each $\delta > 0$,

$$\begin{aligned}
\mathcal{L}(u_\tau^n,v_\tau^n) &\leq \frac{M\varepsilon}{2\lambda'_\varepsilon} + (\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)\exp\Big(-\frac{n\tau}{\tau}\log(1+2\lambda'_\varepsilon\tau)\Big)\\
&\leq (\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)\exp\left(-\frac{T_0}{\tau}\log(1+2\lambda'_\varepsilon\tau)\right) + M'.
\end{aligned}$$

Obviously, we obtain (II.4.12) in the case $\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1 \leq M'$. Consider the converse case. Since $\lim_{s\to 0}\frac{\log(1+s)}{s} = 1$, there exists $\bar{s} > 0$ such that $\frac{\log(1+s)}{s} \geq \frac{1}{1+2\delta}$ for all $s \in (0,\bar{s}]$. Henceforth, defining $\overline{\tau} := \frac{\bar{s}}{2\lambda_0}$ yields $\frac{\log(1+2\lambda'_\varepsilon\tau)}{2\lambda'_\varepsilon\tau} \geq \frac{1}{1+2\delta}$ for all $\tau \in (0,\overline{\tau}]$, and we arrive at the desired estimate by definition of T_0:

$$\mathcal{L}(u_\tau^n,v_\tau^n) \leq M' + (\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1)\exp\left(-\log\frac{\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + K\varepsilon_1}{M'}\right) = 2M'.$$

□

II.4.4.2. Exponential convergence to equilibrium

We are now able to prove — for sufficiently large times — a refined version of Proposition II.50(a):

Proposition II.51 (Exponential estimate for $\mathcal{L}$). *There exist constants $\overline{\varepsilon} \in (0, \varepsilon_1)$ and $\overline{L} > 0$ such that for arbitrary $\delta > 0$, there exists $\overline{\tau} > 0$ such that for all $\varepsilon \in (0, \overline{\varepsilon})$, $\tau \in (0, \overline{\tau}]$ and $n \in \mathbb{N}$ with $n\tau \geq T_0$, we have*

$$(1 + 2\Lambda_\varepsilon \tau)\mathcal{L}(u_\tau^n, v_\tau^n) \leq \mathcal{L}(u_\tau^{n-1}, v_\tau^{n-1}), \tag{II.4.16}$$

with $\Lambda_\varepsilon := \lambda_0 - \overline{L}\varepsilon > 0$ and T_0 as in Proposition II.50(b).

Proof. We write (u, v) instead of (u_τ^n, v_τ^n) for the sake of clarity and consider the last term in $\mathcal{D}(u, v)$ once more. Using as in the proof of Proposition II.50(a) the Hölder, Hardy-Littlewood-Sobolev and L^p-interpolation inequalities (cf. (II.4.13)&(II.4.14)), we get for some $C, C' > 0$ and $\beta \in (0, 1)$:

$$\begin{aligned}
&\int_{\mathbb{R}^3} (u + v)|\mathrm{D}(\psi - \psi_\infty)|^2 \,\mathrm{d}x = \int_{\mathbb{R}^3} ((u - u_\infty) + (v - v_\infty) + (u_\infty + v_\infty))|\mathrm{D}(\psi - \psi_\infty)|^2 \,\mathrm{d}x \\
&\leq C\|(u - u_\infty) - (v - v_\infty)\|_{L^2}^2 \left(\|u - u_\infty\|_{L^2}^\beta \|u - u_\infty\|_{L^1}^{1-\beta} + \|v - v_\infty\|_{L^2}^\beta \|v - v_\infty\|_{L^1}^{1-\beta} + \|u_\infty + v_\infty\|_{L^{3/2}}\right) \\
&\leq C \cdot 2\mathcal{L}(u, v) \cdot C'(1 + \mathcal{L}(u, v)) \leq 2CC'(1 + 2M')\mathcal{L}(u, v),
\end{aligned}$$

with Young's inequality, (II.4.4) and (II.4.12). Now, (II.4.16) follows thanks to (II.4.7), for sufficiently small $\varepsilon > 0$. □

Finally, we prove Theorem II.10.

Proof. Consider a vanishing sequence $(\tau_k)_{k\in\mathbb{N}}$ such that the corresponding sequence of discrete solutions $(u_{\tau_k}, v_{\tau_k})_{k\in\mathbb{N}}$ converges to a weak solution to (II.1.7), in the sense of Theorem II.46. Lower semicontinuity yields $\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq \liminf_{k\to\infty} \mathcal{L}(u_{\tau_k}(t, \cdot), v_{\tau_k}(t, \cdot))$ for all $t \geq 0$. By (II.4.5) and the monotonicity of $\mathcal{E}$ from Proposition II.45, one obtains after passage to $k \to \infty$ that

$$\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq \mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1 \qquad \forall t \geq 0. \tag{II.4.17}$$

Iterating the estimate (II.4.16), assuming without loss of generality that $k \in \mathbb{N}$ is sufficiently large, we get in the limit $k \to \infty$ that

$$\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq 2M' \exp(-2\Lambda_\varepsilon(t - T_0)) \qquad \forall t \geq T_0. \tag{II.4.18}$$

Actually, (II.4.17)&(II.4.18) imply that $\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq A \exp(-2\Lambda_\varepsilon t)$ for all $t \geq 0$, with some constant $A > 0$, the particular structure of which remaining to be identified.
Consider the case $\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1 \leq M'$. Then $T_0 = 0$, so (II.4.18) holds for all $t \geq 0$.
In the other case, (II.4.17)&(II.4.18) yield

$$\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq \max(\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1, 2M') \exp(2\Lambda_\varepsilon T_0) \exp(-2\Lambda_\varepsilon t) \qquad \forall t \geq 0.$$

We insert the definition of T_0 and use that $\Lambda_\varepsilon \leq \lambda_\varepsilon'$ to find

$$\begin{aligned}
\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq{}& \max\left((\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1), 2M'\right) \\
&\cdot \left(\frac{\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1}{M'}\right)^{1+2\delta} \exp(-2\Lambda_\varepsilon t).
\end{aligned}$$

Combining both cases yields

$$\begin{aligned}
\mathcal{L}(u(t, \cdot), v(t, \cdot)) \leq{}& \max(\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1, 2M') \\
&\cdot \max\left(1, \frac{\mathcal{E}(u_0, v_0) - \mathcal{E}(u_\infty, v_\infty) + K\varepsilon_1}{M'}\right)^{1+2\delta} \exp(-2\Lambda_\varepsilon t).
\end{aligned}$$

Thus, we can find $\widetilde{C}_\delta > 0$ such that

$$\mathcal{L}(u(t,\cdot),v(t,\cdot)) \leq \widetilde{C}_\delta(\mathcal{E}(u_0,v_0) - \mathcal{E}(u_\infty,v_\infty) + 1)^{2(1+\delta)} \exp(-2\Lambda_\varepsilon t) \qquad \forall t \geq 0,$$

and the desired exponential estimate (II.1.11) follows by means of (II.4.5) and (I.2.9). $\square$

Part III

Multi-species systems of nonlocal interaction equations

CHAPTER III.1

Introduction to Part III

In this part of the thesis (as an extended version of the article [186]), we analyse the following system of $n \in \mathbb{N}$ nonlocal interaction evolution equations

$$\begin{aligned} \partial_t \mu_1 &= \operatorname{div}[m_1 \mu_1 \nabla (W_{11} * \mu_1 + W_{12} * \mu_2 + \ldots + W_{1n} * \mu_n)], \\ \partial_t \mu_2 &= \operatorname{div}[m_2 \mu_2 \nabla (W_{21} * \mu_1 + W_{22} * \mu_2 + \ldots + W_{2n} * \mu_n)], \\ &\vdots \\ \partial_t \mu_n &= \operatorname{div}[m_n \mu_n \nabla (W_{n1} * \mu_1 + W_{n2} * \mu_2 + \ldots + W_{nn} * \mu_n)]. \end{aligned} \tag{III.1.1}$$

The sought-for n-vector-valued solution $\mu(t) = (\mu_1(t), \ldots, \mu_n(t))$ describes the distribution or concentration of n different populations or agents on $\mathbb{R}^d$ at time $t \geq 0$, $d \in \mathbb{N}$ denoting the spatial dimension. Apart from the constant *mobility magnitudes* $m_1, \ldots, m_n > 0$, system (III.1.1) is mainly governed by the matrix-valued *interaction potential* $W : \mathbb{R}^d \to \mathbb{R}^{n \times n}$ satisfying the following requirements:

(W1) $W(z)$ is a symmetric matrix for each $z \in \mathbb{R}^d$.
(W2) $W_{ij} \in C^1(\mathbb{R}^d; \mathbb{R})$ for all $i, j \in \{1, \ldots, n\}$.
(W3) $W(z) = W(-z)$ for all $z \in \mathbb{R}^d$.
(W4) There exists a matrix $\overline{W} \in \mathbb{R}^{n \times n}$ such that for each $i, j \in \{1, \ldots, n\}$ and all $z \in \mathbb{R}^d$:

$$|W_{ij}(z)| \leq \overline{W}_{ij}(1 + |z|^2).$$

(W5) There exists a symmetric matrix $\kappa \in \mathbb{R}^{n \times n}$ such that, for each $i, j \in \{1, \ldots, n\}$, W_{ij} is κ_{ij}-(semi-)convex, i.e. the map $z \mapsto W_{ij}(z) - \frac{1}{2}\kappa_{ij}|z|^2$ is convex.

System (III.1.1) possesses a formal gradient flow structure: On the subspace $\mathscr{P}$ of those n-vector Borel measures on $\mathbb{R}^d$ with fixed *total masses* $\mu_j(\mathbb{R}^d) = p_j > 0$, fixed *(joint, weighted) center of mass*

$$\sum_{j=1}^n \frac{1}{m_j} \int_{\mathbb{R}^d} x \, d\mu_j(x) = E \in \mathbb{R}^d,$$

and finite second moments $\mathbf{m}_2(\mu_j) := \int_{\mathbb{R}^d} |x|^2 \, d\mu_j(x)$, the *multi-component interaction energy* functional

$$\mathcal{E}(\mu) := \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W_{ij}(x-y) \, d\mu_i(x) \, d\mu_j(y) \tag{III.1.2}$$

induces (III.1.1) as its gradient flow w.r.t. the following compound metric of Wasserstein-type distances for each of the components of the vector measures $\mu^0, \mu^1 \in \mathscr{P}$:

$$\mathbf{W}_{\mathscr{P}}(\mu^0, \mu^1) = \left[\sum_{j=1}^n \frac{1}{m_j} \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \, d\gamma_j(x,y) \,\middle|\, \gamma_j \in \Gamma(\mu_j^0, \mu_j^1) \right\} \right]^{1/2}, \tag{III.1.3}$$

where $\Gamma(\mu_j^0, \mu_j^1)$ is the subset of finite Borel measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ_j^0 and μ_j^1.

It easily follows from the properties of the usual Wasserstein distance for probability measures with finite second moment that $\mathbf{W}_{\mathscr{P}}$ defines a distance on the (geodesic) space $\mathscr{P}$ (see for instance [179, 4] for more details on optimal transport and gradient flows). Note that (W1)–(W5) imply (at least formally) that $\mathscr{P}$ is a positively invariant set along the flow of the evolution (III.1.1). In this part of the thesis, we give a rigorous proof for these formal arguments.

III.1.1. Main results

We obtain in the case of genuine *irreducible* systems (see Definition III.6 below) a novel sufficient condition on the model parameters such that the interaction energy functional $\mathcal{E}$ becomes λ-convex along *generalized* geodesics on $\mathscr{P}$ with respect to the distance $\mathbf{W}_{\mathscr{P}}$ for some $\lambda \in \mathbb{R}$, see Definition III.1 below. Note that λ-convexity along generalized geodesics implies λ-convexity along geodesics in the usual sense, that is, for every pair $\mu^0, \mu^1 \in \mathscr{P}$, there exists a constant-speed geodesic curve $(\mu^s)_{s\in[0,1]}$ in $\mathscr{P}$ connecting μ^0 and μ^1 for which

$$\mathcal{E}(\mu^s) \leq (1-s)\mathcal{E}(\mu^0) + s\mathcal{E}(\mu^1) - \frac{1}{2}s(1-s)\lambda \mathbf{W}_{\mathscr{P}}(\mu^0,\mu^1)^2, \qquad \forall\, s \in [0,1]. \tag{III.1.4}$$

For convexity along *generalized* geodesics, an inequality of the form (III.1.4) is required for a wider class of curves joining μ^0 and μ^1.

We call the energy *uniformly* geodesically convex if it is λ-convex along generalized geodesics, for some $\lambda > 0$. Define, for each $i \in \{1,\dots,n\}$, the numbers $\eta_i := \min_{j\neq i} \kappa_{ij} m_j$. We prove in Section III.2.1 the λ-convexity along generalized geodesics of $\mathcal{E}$ for all

$$\lambda \leq \min_{i\in\{1,\dots,n\}} \left[p_i \min(0, m_i\kappa_{ii} - \eta_i) + \frac{1}{2}\sum_{j=1}^n p_j \left(\eta_j + \eta_i \frac{m_i}{m_j} \right) \right]. \tag{III.1.5}$$

Even if some of the W_{ij} are *not* uniformly convex (i.e. $\kappa_{ij} \leq 0$), we might still obtain a convexity modulus $\lambda > 0$, if attraction dominates repulsion as required in (III.1.5). Using this new condition, we are in position to invoke the theory on (λ-contractive) gradient flows in metric spaces by Ambrosio, Gigli and Savaré ([4]; see Section I.2.1 in the introduction) to obtain (cf. Section III.2.2) the existence of a *gradient flow solution* $\mu \in AC^2_{\mathrm{loc}}([0,\infty);(\mathscr{P},\mathbf{W}_{\mathscr{P}}))$ to (III.1.1) with initial datum $\mu^0 \in \mathscr{P}$. Moreover, uniqueness of solutions follows from the *contraction estimate* for all $t > 0$: $\mathbf{W}_{\mathscr{P}}(\mu(t),\nu(t)) \leq e^{-\lambda t}\mathbf{W}_{\mathscr{P}}(\mu^0,\nu^0)$. If the modulus of geodesic convexity is strictly positive, the measure

$$\mu^\infty := (p_1,\dots,p_n)^{\mathrm{T}}\delta_{x^\infty}, \qquad \text{with} \qquad x^\infty := E\left[\sum_{j=1}^n \frac{p_j}{m_j}\right]^{-1} \in \mathbb{R}^d,$$

is the unique minimizer — the *ground state* — of $\mathcal{E}$ and the unique stationary state of (III.1.1) on $\mathscr{P}$. It is globally asymptotically stable since gradient flow solutions $\mu(t)$ converge exponentially fast in $(\mathscr{P},\mathbf{W}_{\mathscr{P}})$ with rate λ to μ^∞. In contrast, if $\mathcal{E}$ is geodesically λ-convex with only $\lambda \leq 0$, the dynamics of system (III.1.1) are more involved.

There, we restrict to one spatial dimension ($d = 1$) and rewrite the system in terms of *inverse distribution functions*: Given the (scaled) cumulative distribution functions

$$F_i(t,x) = \int_{-\infty}^x \frac{1}{p_i}\,\mathrm{d}\mu_i(t,y) \quad \in [0,1], \tag{III.1.6}$$

let u_i be their corresponding pseudo-inverse, i.e.

$$u_i(t,z) = \inf\{x \in \mathbb{R} : F_i(t,x) > z\} \quad (\text{for } z \in [0,1)). \tag{III.1.7}$$

Then, system (III.1.1) transforms into (cf. Section III.3)

$$\partial_t u_i(t,z) = m_i \sum_{j=1}^n p_j \int_0^1 W_{ij}'(u_j(t,\xi) - u_i(t,z))\,\mathrm{d}\xi \qquad (i = 1,\dots,n). \tag{III.1.8}$$

In terms of system (III.1.8), if $\mu \in AC^2_{\mathrm{loc}}([0,\infty);(\mathscr{P},\mathbf{W}_{\mathscr{P}}))$, one has $u \in AC^2_{\mathrm{loc}}([0,\infty);L^2([0,1];\mathbb{R}^n))$ and for all $t \geq 0$ and all $i \in \{1,\ldots,n\}$, $u_i(t,\cdot)$ is a non-decreasing *càdlàg* function on $(0,1)$. Chapter III.3 is concerned with the analysis of the qualitative behaviour of the solution μ to (III.1.1) by means of investigation of the corresponding solution u to (III.1.8). Our main result is a *confinement* property of the solution: For admissible interaction potentials satisfying (W1)–(W5) only, we prove (cf. Proposition III.14) that

$$\operatorname{supp}\mu_i(t) \subset [-K(T,\mu^0),K(T,\mu^0)] \quad \forall\, t \in [0,T] \quad \forall\, i \in \{1,\ldots,n\}, \tag{III.1.9}$$

for some finite constant $K > 0$ depending on the (compactly supported) initial datum μ^0 and the (finite) time horizon $T > 0$. Due to repulsion effects in this general setting, $K(T,\mu^0) \to \infty$ may occur as $T \to \infty$. We propose a *confining* condition on the interaction potential W (cf. Definition III.15) such that the property above extends to $T = \infty$ (cf. Theorem III.17): In a nutshell, we require W to behave — outside a compact set — like a potential inducing a uniformly geodesically convex energy functional $\mathcal{E}$, in the sense of our criterion (III.1.5). Note that we do *not* require all W_{ij} to be uniformly convex far away from the origin.

Thus, in many cases, mass cannot escape to infinity. In contrast, is it possible to have *concentration* in finite time, i.e., can it occur that absolutely continuous solutions collapse to measures with nonvanishing singular part in finite time? The answer is negative for Lipschitz continuous W'_{ij} and absolutely continuous initial data with continuous and bounded Lebesgue density (cf. Proposition III.18).

Section III.3.3 is devoted to the study of the long-time behaviour of the solution to (III.1.1). We first prove (cf. Theorem III.19) that if the solution is *a priori* confined to a compact set, the ω-limit set of the system only contains steady states of (III.1.1). More specifically, assume that (III.1.9) is true for some $K > 0$ and $T = \infty$ and assume that all W'_{ij} are Lipschitz continuous on the interval $[-2K,2K]$. Then,

$$\lim_{t\to\infty}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\mu(t))\right) = 0. \tag{III.1.10}$$

Moreover, for each sequence $t_k \to \infty$, there exists a subsequence and a steady state $\overline{\mu} \in \mathscr{P}$ of (III.1.1) such that on the subsequence

$$\lim_{k\to\infty}\mathbf{W}_1(\mu_i(t_k),\overline{\mu}_i) = 0 \quad \forall\, i \in \{1,\ldots,n\}.$$

There, $\mathbf{W}_1$ denotes the L^1-Wasserstein distance between finite measures. However, this large-time limit $\overline{\mu}$ is not unique since it depends both on the chosen sequence $(t_k)_{k\in\mathbb{N}}$ and the extracted subsequence.

Even if the interaction potential does neither yield uniform geodesic convexity of the energy nor is confining, we may observe a *δ-separation phenomenon*: If the initial datum has compact support and the model parameters admit $\sum_{j=1}^n \kappa_{ij}p_j > 0$ for all i, the diameter of the support of the solution shrinks exponentially fast over time (cf. Proposition III.21). Still, the solution does in general not converge to a fixed steady state. However, in the uniformly geodesically convex regime ($\lambda > 0$ in (III.1.5)), we obtain convergence even w.r.t. the stronger topology of the L^∞-Wasserstein distance $\mathbf{W}_\infty$,

$$\lim_{t\to\infty}\mathbf{W}_\infty(\mu_i(t),\mu_i^\infty) = 0 \quad \forall\, i \in \{1,\ldots,n\},$$

for initial data with compact support. In contrast to convergence w.r.t. $\mathbf{W}_{\mathscr{P}}$ (cf. Corollary III.12), we do not obtain a specific rate of convergence.

The last chapter of this part (Chapter III.4) is concerned with the existence and stability of steady states for (III.1.1) in one spatial dimension $d = 1$ for sufficiently regular interaction potentials W. First, we show that if W is analytic, only *discrete* steady states $\overline{\mu}$ of the form

$$\overline{\mu}_i = \sum_{k=1}^{N_i} p_i^k \delta_{x_i^k} \qquad (i = 1 \dots, n)$$

can exist. Concerning the local nonlinear stability of those discrete steady states, we give a sufficient condition in Section III.4.2, see Theorem III.29. In a nutshell, if one requires the stability in certain subspaces of the phase space in the *linearized* system at $\overline{\mu}$, local *nonlinear* stability follows. As a preparation, we investigate the linearized system in Section III.4.1. There, we also disprove the linear asymptotic stability of non-discrete steady states by studying the spectrum of the associated linear operator (Theorem III.25).

III.1.2. Modelling background and relation to the literature

System (III.1.1) is a natural generalization of the scalar nonlocal evolution equation

$$\partial_t \mu = \operatorname{div}[m\mu \nabla(W * \mu)], \tag{III.1.11}$$

to multiple components. For the corresponding interaction energy functional

$$\mathcal{E}(\mu) = \frac{1}{2} \int_{\mathbb{R}^d} (W * \mu) \, \mathrm{d}\mu, \tag{III.1.12}$$

McCann provided in his seminal paper [133] a criterion for geodesic λ-convexity with respect to the L^2-Wasserstein distance (see Theorem I.14). In a nutshell, if W is κ-convex in the Euclidean sense on $\mathbb{R}^d$ for some $\kappa \in \mathbb{R}$, then $\mathcal{E}$ is geodesically $\min(0, \kappa)$-convex on the space of probability measures endowed with the L^2-Wasserstein distance. On the subspace of those measures having fixed center of mass, $\mathcal{E}$ is geodesically κ-convex (i.e. the uniform convexity $\kappa > 0$ is retained in the metric framework). It was proved by Ambrosio, Gigli and Savaré in [4] that geodesic convexity essentially leads to existence and uniqueness of weak solutions for the associated gradient flow evolution equation and to contractivity of the associated flow map — cf. Section I.2.1. An immediate consequence of λ-geodesic convexity of functionals — for strictly positive $\lambda \in \mathbb{R}$ — is existence and uniqueness of minimizers (for recent results without using convexity, see e.g. [46, 42]).

Model equations of the form (III.1.11) have arised in the study of population dynamics in many cases (e.g. [24, 27, 34, 50, 60, 109, 113, 129, 145, 174, 175]) often derived as the infinite-particle limit of a individual-based model (e.g. [29, 94, 146]):

- In the parabolic-elliptic (Patlak-)Keller-Segel model for chemotaxis in two spatial dimensions, the interaction potential is given by (the negative of) the Newtonian potential, i.e. $W(z) = \frac{1}{2\pi}\log(|z|)$, which is singular at $z = 0$ and attractive.
- Typical mathematical models of swarming processes include so-called *attractive-repulsive* potentials of the form $W(z) = -C_a e^{-\frac{|z|}{l_a}} + C_r e^{-\frac{|z|}{l_r}}$, a special case of which is the attractive *Morse potential* $W(z) = -e^{-|z|}$. Also, Gaussian-type attractive-repulsive potentials $W(z) = -C_a e^{-\frac{|z|^2}{l_a}} + C_r e^{-\frac{|z|^2}{l_r}}$ are conceivable.

Nonlocal interaction potentials also appear in several models of physical applications such as models for granular media [12, 54, 55, 120, 176, 58], opinion formation [177] or interactions between particles (e.g. in crystals [172] or fluids [184]) with a broad range of reasonable interaction potentials. One can e.g. consider

- convex and C^1-regular potentials, e.g. $W(z) = |z|^q$ for $q > 1$;
- non-convex, but regular potentials such as the *double-well potential* $W(z) = |z|^4 - |z|^2$;
- non-convex and singular potentials, e.g. the *Lennard-Jones* potential.

In the case of a radially symmetric potential $W(z) = w(|z|)$, the effect of the interaction potential is reflected by the sign of w': If w' is positive, the individuals of the population *attract* each other, whereas in the case of negative w' the dynamics are *repulsive*. The force generated by the potential W points towards or away from the origin for positive or negative w', respectively. Radially symmetric potentials describe interactions only depending on the distance of the particles. With the sum of convolutions appearing in the flux on the r.h.s. of system (III.1.1), we take into account that every species generates a — probably long-range — force on every other species.

Naturally, aggregation processes modelled by nonlocal interaction potentials are often combined with diffusive processes yielding (nonlinear) drift-diffusion equations as mathematical models. The question of global existence of solutions to equations of these forms has been addressed in various publications. Using the theory of gradient flows, global existence of measure-valued solutions was proved in [47, 52], also for non-smooth potentials, in generalization of [54, 55]. Methods from optimal transportation theory were useful for proving uniqueness, see e.g. [56, 63]. Well-posedness in the measure-valued sense was also studied in [43], and for a similar system as (III.1.1) in [66] for two species (see below).

A second field of study is the analysis of the qualitative behaviour of solutions to equations like (III.1.11), such as the speed of propagation, finite- and infinite-time blow-up of solutions and possible attractors, also with focus on self-similarity of solutions. It is not surprising that (III.1.11) exhibits blow-ups if the potential is sufficiently attractive. The aforementioned properties were investigated e.g. in [9, 13, 14, 15, 16, 21, 48, 119, 171]. One specific object of study is equation (III.1.11) considered in one spatial dimension. For instance, in [158, 82, 83] by Raoul and Fellner, rewriting in terms of inverse distribution functions allowed for the characterization of the long-time behaviour and the set of possible steady states of (III.1.11). One-dimensional models with nonlinear diffusion have been studied e.g. by Burger and Di Francesco in [35].

Genuine systems of the specific form (III.1.1) have been investigated only in the case of two species as a physical model for two-component mixtures [178], fluids [184] or particle interactions [87]. From a more mathematical point of view, they were analysed by Di Francesco and Fagioli first in [66], where the results from [47] were generalized to the case of two components using gradient flow methods. In their recent work [67], existence and stability of steady states were studied for the two-species system, leading to similar results as presented in Chapter III.4 for an arbitrary number of species. In [66], also the case of non-symmetric interaction was studied. Here, we only focus on the case of regular and symmetric interaction potentials, see assumptions (W1)–(W5), but allow for an arbitrary number of species. Besides, our condition for *uniform* geodesic convexity for genuine systems of the form (III.1.1) is novel.

CHAPTER III.2

Geodesic convexity and existence of gradient flow solutions

In this chapter, we derive a sufficient condition for λ-convexity along generalized geodesics of the interaction energy $\mathcal{E}$ (cf. formula (III.1.2)) and conclude existence and uniqueness of solutions to (III.1.1). Throughout this part, the assumptions (W1)–(W5) above shall be fulfilled.

We begin with our definition of convexity along generalized geodesics, which is a straightforward generalization of the respective definition in the scalar case (see Definition I.12) to our vector-valued setting:

Definition III.1 (λ-convexity along generalized geodesics). *Given $\lambda \in \mathbb{R}$, we say that a functional $\mathcal{A} : \mathscr{P} \to \mathbb{R}_\infty$ is* λ-convex along generalized geodesics *on $(\mathscr{P}, \mathbf{W}_{\mathscr{P}})$, if for any triple $\mu^1, \mu^2, \mu^3 \in \mathscr{P}$, there exists a n-vector-valued Borel measure $\boldsymbol{\mu}$ on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$ such that:*

- *For all $j \in \{1, \ldots, n\}$ and all $k \in \{1, 2, 3\}$: $\mu_j^k = \pi^k{}_{\#}\boldsymbol{\mu}_j$.*
- *For $k \in \{2, 3\}$ and all $j \in \{1, \ldots, n\}$, the measure $\pi^{(1,k)}{}_{\#}\boldsymbol{\mu}_j$ is optimal in $\Gamma(\mu_j^1, \mu_j^k)$, i.e. it realizes the minimum in*

$$\inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x^1 - x^k|^2 \, \mathrm{d}\gamma_j(x^1, x^k) \middle| \gamma_j \in \Gamma(\mu_j^1, \mu_j^k) \right\}.$$

- *Defining for $s \in [0, 1]$ and all $j \in \{1, \ldots, n\}$ the* generalized geodesic *μ_s connecting μ^2 and μ^3 (with* base point *μ^1) by*

$$\mu_{s,j} := \Big[(1-s)\pi^2 + s\pi^3 \Big]_{\#} \boldsymbol{\mu}_j,$$

one has for all $s \in [0, 1]$:

$$\mathcal{A}(\mu_s) \le (1-s)\mathcal{A}(\mu^2) + s\mathcal{A}(\mu^3) - \frac{\lambda}{2} s(1-s) \sum_{j=1}^{n} \frac{1}{m_j} \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} |x^3 - x^2|^2 \, \mathrm{d}\boldsymbol{\mu}(x^1, x^2, x^3). \tag{III.2.1}$$

The following sufficient criterion is useful in verifying convexity along generalized geodesics as it allows to consider *absolutely continuous* measures and transport *maps*.

Theorem III.2 (Sufficient criterion for convexity along generalized geodesics [4, Prop. 9.2.10]). *Let $\mathcal{A} : \mathscr{P} \to \mathbb{R}_\infty$ be lower semicontinuous and such that for all $\mu \in \mathscr{P}$, there exists a sequence $(\mu^k)_{k \in \mathbb{N}}$ on the subspace $\mathscr{P}^{\mathrm{ac}}$ of absolutely continuous measures in $\mathscr{P}$ with $\lim_{k \to \infty} \mathbf{W}_{\mathscr{P}}(\mu^k, \mu) = 0$ and $\lim_{k \to \infty} \mathcal{A}(\mu^k) = \mathcal{A}(\mu)$.*

Assume moreover that for each $\mu \in \mathscr{P}^{\mathrm{ac}}$ and $t_1, \ldots, t_n, \widetilde{t}_1, \ldots, \widetilde{t}_n : \mathbb{R}^d \to \mathbb{R}^d$ such that $t_j - \widetilde{t}_j \in L^2(\mathbb{R}; \mathrm{d}\mu_j)$ for all $j \in \{1, \ldots, n\}$, the following estimate holds along the interpolating curve $(\mu^s)_{s \in [0,1]}$ defined as $\mu_j^s := \big[(1-s)\widetilde{t}_j + st_j\big]_{\#}\mu_j$ for all $s \in [0, 1]$ and $j \in \{1, \ldots, n\}$:

$$\mathcal{A}(\mu^s) \le (1-s)\mathcal{A}(\mu^0) + s\mathcal{A}(\mu^1) - \frac{\lambda}{2} s(1-s) \sum_{j=1}^{n} \frac{1}{m_j} \int_{\mathbb{R}^d} |t_j(x) - \widetilde{t}_j(x)|^2 \, \mathrm{d}\mu_j(x) \quad \textit{for all } s \in [0, 1]. \tag{III.2.2}$$

Then, $\mathcal{A}$ is λ-convex along generalized geodesics on $(\mathscr{P}, \mathbf{W}_{\mathscr{P}})$.

III.2.1. Geodesic convexity of the multi-component interaction energy

We first prove some basic properties of the interaction energy $\mathcal{E}$.

Lemma III.3 (Proper domain and lower semicontinuity). *The following statements hold:*
(a) For all $\mu \in \mathscr{P}$, one has $|\mathcal{E}(\mu)| < \infty$.
(b) $\mathcal{E}$ is continuous on the metric space $(\mathscr{P}, \mathbf{W}_{\mathscr{P}})$.
(c) Let $\rho \in C_c^\infty(\mathbb{R}^d)$ be defined via

$$\rho(x) := Z \exp\left(\frac{1}{|x|^2 - 1}\right) \mathbf{1}_{\mathbb{B}_1(0)}(x),$$

where $Z > 0$ is such that $\|\rho\|_{L^1} = 1$, and put $\rho_\varepsilon(x) := \varepsilon^{-d}\rho\left(\frac{x}{\varepsilon}\right)$ for $\varepsilon > 0$. Then, for each $\mu \in \mathscr{P}$, the sequence $(\mu^k)_{k\in\mathbb{N}}$ with $\mu_j^k := \rho_{\frac{1}{k}} * \mu_j$ for $k \in \mathbb{N}$, $j \in \{1, \dots, n\}$, belongs to $\mathscr{P}^{\mathrm{ac}}$ and $\lim\limits_{k\to\infty} \mathbf{W}_{\mathscr{P}}(\mu^k, \mu) = 0$.

Proof. The key observation for parts (a) and (b) is the sub-quadratic growth of $\mathcal{E}$ as a consequence of condition (W4): there exists a constant $C > 0$ such that for all $\mu \in \mathscr{P}$, one has

$$|\mathcal{E}(\mu)| \le C \sum_{j=1}^n \int_{\mathbb{R}^d} (1 + |x|^2)\,\mathrm{d}\mu_j(x). \tag{III.2.3}$$

Clearly, (a) follows. If $(\mu^k)_{k\in\mathbb{N}}$ is a sequence converging to μ in $(\mathscr{P}, \mathbf{W}_{\mathscr{P}})$, in particular their second moments converge componentwise. Hence, the integrand on the r.h.s. in (III.2.3) is uniformly integrable which yields (b) as in [4, Lemma 5.1.7], using the continuity of W. For part (c), we observe for $j \in \{1, \dots, n\}$ that μ_j^k is an absolutely continuous measure on $\mathbb{R}^d$ with Lebesgue density

$$x \mapsto \int_{\mathbb{R}^d} \rho_{\frac{1}{k}}(x - y)\,\mathrm{d}\mu_j(y).$$

Clearly, $\mu_j^k(\mathbb{R}^d) = p_j$. Moreover, the center of mass E is unchanged by convolution with $\rho_{\frac{1}{k}}$ since for all $j \in \{1, \dots, n\}$ and all $k \in \mathbb{N}$:

$$\int_{\mathbb{R}^d} x\,\mathrm{d}\mu_j^k(x) = \int_{\mathbb{R}^d} x\,\mathrm{d}\mu_j(x).$$

Indeed, by transformation and Fubini's theorem,

$$\int_{\mathbb{R}^d} x\,\mathrm{d}\mu_j^k(x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (z + y)\rho_{\frac{1}{k}}(z)\,\mathrm{d}\mu_j(y)\,\mathrm{d}z = \int_{\mathbb{R}^d} y\,\mathrm{d}\mu_j(y) + p_j \int_{\mathbb{R}^d} z\rho_{\frac{1}{k}}(z)\,\mathrm{d}z.$$

The last integral above vanishes since ρ is an even function. Along the same lines, one proves convergence of the second moments:

$$\begin{aligned}\mathfrak{m}_2(\mu_j^k) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |z + y|^2 \rho_{\frac{1}{k}}(z)\,\mathrm{d}\mu_j(y)\,\mathrm{d}z \\ &= \mathfrak{m}_2(\mu_j) + p_j \int_{\mathbb{R}^d} |z|^2 \rho_{\frac{1}{k}}(z)\,\mathrm{d}z + 2\left(\int_{\mathbb{R}^d} z\rho_{\frac{1}{k}}(z)\,\mathrm{d}z\right)^{\mathrm{T}} \left(\int_{\mathbb{R}^d} y\,\mathrm{d}\mu_j(y)\right).\end{aligned}$$

Since the last term vanishes again, we see

$$\mathfrak{m}_2(\mu_j^k) = \mathfrak{m}_2(\mu_j) + \frac{1}{k^2}\int_{\mathbb{R}^d} |x|^2 \rho(x)\,\mathrm{d}x \xrightarrow{k\to\infty} \mathfrak{m}_2(\mu_j).$$

It remains to prove narrow convergence of μ_j^k to μ_j. Fix $f : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded. Using Fubini's theorem again, we get, since ρ is even,

$$\int_{\mathbb{R}^d} f\,\mathrm{d}\mu_j^k - \int_{\mathbb{R}^d} f\,\mathrm{d}\mu_j = \int_{\mathbb{R}^d} (\rho_{\frac{1}{k}} * f - f)\,\mathrm{d}\mu_j.$$

Since f is continuous, $\rho_{\frac{1}{k}} * f$ converges to f pointwise on $\mathbb{R}^d$ (see, for instance, [80, App. C]). Clearly, $\rho_{\frac{1}{k}} * f$ is k-uniformly bounded. The dominated convergence theorem now yields

$$\int_{\mathbb{R}^d} (\rho_{\frac{1}{k}} * f - f) \, \mathrm{d}\mu_j \overset{k\to\infty}{\longrightarrow} 0,$$

proving the claim. □

Lemma III.4 (Growth control on the gradient). *There exists a matrix $\overline{C} \in \mathbb{R}^{n\times n}$ such that for all $z \in \mathbb{R}^d$ and all $i, j \in \{1, \dots, n\}$:*

$$|\nabla W_{ij}(z)| \le \overline{C}_{ij}(|z| + 1). \tag{III.2.4}$$

Proof. We give a short proof for the sake of completeness. From (W2) and (W5), it easily follows for all $x, y \in \mathbb{R}^d$ that

$$W_{ij}(y) - W_{ij}(x) - \frac{\kappa_{ij}}{2}|y - x|^2 \ge \nabla W_{ij}(x)^{\mathrm{T}}(y - x).$$

Putting

$$\alpha := \begin{cases} |4\overline{W}_{ij} - \kappa_{ij}|^{-1} & \text{if } 4\overline{W}_{ij} > \kappa_{ij}, \\ 1 & \text{otherwise,} \end{cases} \quad \text{and} \quad y := x + \alpha\nabla W_{ij}(x),$$

we get, using (W4) and Young's inequality:

$$\alpha|\nabla W_{ij}(x)|^2 \le \overline{W}_{ij}(2 + 3|x|^2) + \frac{1}{2}\alpha^2(4\overline{W}_{ij} - \kappa_{ij})|\nabla W(x)|^2.$$

Consequently, in both cases, we have

$$|\nabla W_{ij}(x)| \le \left(2\alpha^{-1}\overline{W}_{ij}(2 + 3|x|^2)\right)^{1/2}$$

which implies an estimate of the form (III.2.4) via the elementary estimate $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ holding for $a, b \ge 0$. □

Remark III.5 (Invariants). *Along the flow of system* (III.1.1)*, the set $\mathscr{P}$ is positively invariant. We give a formal indication of this fact: Let an initial datum $\mu^0 \in \mathscr{P}$ be given. Since* (III.1.1) *is in divergence form, we immediately obtain the conservation of mass:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \mathrm{d}\mu_i(t, x) = 0.$$

Furthermore, by formal integration by parts, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^n \mathfrak{m}_2(\mu_i(t)) = -\sum_{i=1}^n 2m_i \int_{\mathbb{R}^d} x^{\mathrm{T}} \left(\sum_{j=1}^n \nabla W_{ij} * \mu_j(t)\right)(x) \, \mathrm{d}\mu_i(t, x),$$

from which it is possible to derive using the Young and Jensen inequalities and Lemma III.4 the estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^n \mathfrak{m}_2(\mu_i(t)) \le A\sum_{i=1}^n \mathfrak{m}_2(\mu_i(t)) + B,$$

for suitable $A, B \in \mathbb{R}$. Gronwall's lemma now yields finiteness of second moments at a fixed time $t \ge 0$. Finally,

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^n \frac{1}{m_i}\int_{\mathbb{R}^d} x \, \mathrm{d}\mu_i(t, x) = -\sum_{i=1}^n\sum_{j=1}^n \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla W_{ij}(x - y) \, \mathrm{d}\mu_j(t, y) \, \mathrm{d}\mu_i(t, x).$$

Using assumptions (W1) *and* (W3) *in combination with Fubini's theorem, we observe that the r.h.s. above is in fact equal to 0.*

Definition III.6 (Irreducible systems). *We call a system of the form* (III.1.1) irreducible, *if the graph* $G = (V_G, E_G)$ *with nodes* $V_G = \{1, \dots, n\}$ *and edges* $E_G = \{(i,j) \in V_G \times V_G : \nabla W_{ij} \not\equiv 0 \text{ on } \mathbb{R}^d\}$ *is connected. That is, irreducible systems cannot be split up into independent subsystems.*

The main result of this section is concerned with the geodesic convexity of the interaction energy $\mathcal{E}$:

Theorem III.7 (Criterion for geodesic convexity). *Let* $n > 1$ *and let* (III.1.1) *be irreducible. Define for* $i \in \{1, \dots, n\}$ *the quantity* $\eta_i := \min_{j \neq i} \kappa_{ij} m_j \in \mathbb{R}$. *Then,* $\mathcal{E}$ *is* λ*-convex along generalized geodesics on* $\mathscr{P}$ *w.r.t.* $\mathbf{W}_{\mathscr{P}}$ *for all* $\lambda \leq \lambda_0$ *with*

$$\lambda_0 := \min_{i \in \{1,\dots,n\}} \left[p_i \min(0, m_i \kappa_{ii} - \eta_i) + \frac{1}{2} \sum_{j=1}^n p_j \left(\eta_j + \eta_i \frac{m_i}{m_j} \right) \right]. \tag{III.2.5}$$

Proof. Thanks to the properties from Lemma III.3, we are allowed to use Theorem III.2. Let therefore $\mu \in \mathscr{P}^{\mathrm{ac}}$ and let $t_1, \dots, t_n, \widetilde{t}_1, \dots, \widetilde{t}_n : \mathbb{R}^d \to \mathbb{R}^d$ such that $t_j - \widetilde{t}_j \in L^2(\mathbb{R}; \mathrm{d}\mu_j)$ for all $j \in \{1, \dots, n\}$. With the notation from Theorem III.2, we have, using condition (W5):

$$\begin{aligned}
\mathcal{E}(\mu^s) &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W_{ij}(\widetilde{t}_j(x) - \widetilde{t}_i(y) + s[t_j(x) - t_i(y) - (\widetilde{t}_j(x) - \widetilde{t}_i(y))])\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y) \\
&\leq (1-s)\mathcal{E}(\mu^0) + s\mathcal{E}(\mu^1) - \frac{1}{2}s(1-s) \cdot \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_{ij} |t_j(x) - t_i(y) - (\widetilde{t}_j(x) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y).
\end{aligned}$$

In view of (III.2.1), we have to verify that

$$\begin{aligned}
&\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_{ij} |t_j(x) - t_i(y) - (\widetilde{t}_j(x) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y) \\
&\geq \lambda_0 \sum_{i=1}^n \frac{1}{m_i} \int_{\mathbb{R}^d} |t_i(x) - \widetilde{t}_i(x)|^2\, \mathrm{d}\mu_i(x).
\end{aligned} \tag{III.2.6}$$

We first split up the l.h.s. of (III.2.6) into its diagonal and off-diagonal part and perform an estimate on the latter introducing the numbers $\eta_i = \min_{j \neq i} \kappa_{ij} m_j$:

$$\begin{aligned}
&\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_{ij} |t_j(x) - t_i(y) - (\widetilde{t}_j(x) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y) \\
&\qquad \geq \frac{1}{2} \sum_i \sum_{j \neq i} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\eta_i}{m_j} |(t_j(x) - \widetilde{t}_j(x)) - (t_i(y) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y) \\
&\qquad\quad + \frac{1}{2} \sum_i \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_{ii} |(t_i(x) - \widetilde{t}_i(x)) - (t_i(y) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_i(x)\, \mathrm{d}\mu_i(y).
\end{aligned}$$

Expanding the squares yields

$$\begin{aligned}
&\frac{1}{2} \sum_i \sum_{j \neq i} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\eta_i}{m_j} |(t_j(x) - \widetilde{t}_j(x)) - (t_i(y) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_j(x)\, \mathrm{d}\mu_i(y) \\
&\quad + \frac{1}{2} \sum_i \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \kappa_{ii} |(t_i(x) - \widetilde{t}_i(x)) - (t_i(y) - \widetilde{t}_i(y))|^2\, \mathrm{d}\mu_i(x)\, \mathrm{d}\mu_i(y) \\
&= \frac{1}{2} \sum_i \sum_{j \neq i} \left(\int_{\mathbb{R}^d} \frac{p_i \eta_i}{m_j} |t_j(x) - \widetilde{t}_j(x)|^2\, \mathrm{d}\mu_j(x) + \int_{\mathbb{R}^d} \frac{p_j \eta_i}{m_j} |t_i(x) - \widetilde{t}_i(x)|^2\, \mathrm{d}\mu_i(x) \right) \\
&\quad - \sum_i \left(\sum_{j \neq i} \int_{\mathbb{R}^d} \frac{1}{m_j} (t_j(x) - \widetilde{t}_j(x))\, \mathrm{d}\mu_j(x) \right)^{\mathrm{T}} \left(\int_{\mathbb{R}^d} \eta_i (t_i(x) - \widetilde{t}_i(x))\, \mathrm{d}\mu_i(x) \right)
\end{aligned} \tag{III.2.7}$$

$$+\sum_i \kappa_{ii}\left(\int_{\mathbb{R}^d} p_i|t_i(x)-\widetilde{t}_i(x)|^2\,\mathrm{d}\mu_i(x)\right)-\sum_i \kappa_{ii}\left|\int_{\mathbb{R}^d}(t_i(x)-\widetilde{t}_i(x))\,\mathrm{d}\mu_i(x)\right|^2.$$

Now, the special structure of $\mathscr{P}$ comes into play: since the weighted center of mass E is fixed on $\mathscr{P}$, one has

$$E=\sum_{j=1}^n \frac{1}{m_j}\int_{\mathbb{R}^d} x\,\mathrm{d}(t_{j\#}\mu_j)=\sum_{j=1}^n \frac{1}{m_j}\int_{\mathbb{R}^d} x\,\mathrm{d}(\widetilde{t}_{j\#}\mu_j),$$

and consequently

$$\sum_{j\neq i}\int_{\mathbb{R}^d}\frac{1}{m_j}(t_j(x)-\widetilde{t}_j(x))\,\mathrm{d}\mu_j(x)=-\int_{\mathbb{R}^d}\frac{1}{m_i}(t_i(x)-\widetilde{t}_i(x))\,\mathrm{d}\mu_i(x).$$

We exploit this fact in order to simplify the second term on the r.h.s. of formula (III.2.7) above:

$$\begin{aligned}\text{r.h.s.}=\sum_i\Bigg\{&\left|\int_{\mathbb{R}^d}(t_i(x)-\widetilde{t}_i(x))\,\mathrm{d}\left(\frac{1}{p_i}\mu_i\right)(x)\right|^2 p_i^2\left(\frac{\eta_i}{m_i}-\kappa_{ii}\right)\\&+\int_{\mathbb{R}^d}|t_i(x)-\widetilde{t}_i(x)|^2\,\mathrm{d}\left(\frac{1}{p_i}\mu_i\right)(x)p_i^2\left(\kappa_{ii}-\frac{\eta_i}{m_i}\right)\\&+\frac{1}{m_i}\int_{\mathbb{R}^d}|t_i(x)-\widetilde{t}_i(x)|^2\,\mathrm{d}\mu_i(x)\cdot\frac{1}{2}\sum_j p_j\left(\eta_j+\eta_i\frac{m_i}{m_j}\right)\Bigg\}=:\sum_i S_i.\end{aligned}$$

We analyse each S_i separately.
If $\frac{\eta_i}{m_i}-\kappa_{ii}\geq 0$, the first term in S_i is nonnegative, so

$$S_i\geq\frac{1}{m_i}\int_{\mathbb{R}^d}|t_i(x)-\widetilde{t}_i(x)|^2\,\mathrm{d}\mu_i(x)\cdot\left[p_i(m_i\kappa_{ii}-\eta_i)+\frac{1}{2}\sum_j p_j\left(\eta_j+\eta_i\frac{m_i}{m_j}\right)\right].$$

If $\frac{\eta_i}{m_i}-\kappa_{ii}<0$, the sum of the first two terms in S_i is nonnegative thanks to Jensen's inequality. Hence,

$$S_i\geq\frac{1}{m_i}\int_{\mathbb{R}^d}|t_i(x)-\widetilde{t}_i(x)|^2\,\mathrm{d}\mu_i(x)\cdot\frac{1}{2}\sum_j p_j\left(\eta_j+\eta_i\frac{m_i}{m_j}\right).$$

Defining λ_0 as in (III.2.5) clearly leads to (III.2.6), completing the proof. □

Remark III.8 (Non-irreducible systems). *If system* (III.1.1) *is* not *irreducible, there exists an I-integer partition ($I\in\mathbb{N}$) of $n\in\mathbb{N}$ into $n_1+n_2+\ldots+n_I=n$ such that* (III.1.1) *decomposes into I independent irreducible subsystems having the same structure as* (III.1.1)*, but with n replaced by $n_1,\ldots,n_I$, respectively. The modulus of geodesic convexity of the interaction energy $\mathcal{E}$ can now be computed as the minimum of the respective convexity moduli of each subsystem: if $n_k>1$ for some $k\in\{1,\ldots,I\}$, formula* (III.2.5) *applies; if $n_k=1$, McCann's criterion [133] applies (and yields convexity modulus $m\kappa p$ for the respective m,κ,p of the k^{th} subsystem in our framework).*

Remark III.9 (Necessary condition for $\lambda_0>0$). *If $\lambda_0>0$ in* (III.2.5)*, then for all $i\in\{1,\ldots,n\}$:*

$$\sum_{j=1}^n\kappa_{ij}p_j>0.$$

This condition is not *sufficient (cf. Example III.10 below).*

Proof. Fix $i\in\{1,\ldots,n\}$. The following holds:

$$m_i\sum_j\kappa_{ij}p_j=p_im_i\kappa_{ii}+\sum_{j\neq i}m_i\kappa_{ij}p_j=p_i(m_i\kappa_{ii}-\eta_i)+p_i\eta_i+\frac{1}{2}\sum_{j\neq i}p_j\kappa_{ji}m_i+\frac{1}{2}\sum_{j\neq i}p_j\kappa_{ij}m_j\frac{m_i}{m_j},$$

where we used the symmetry of κ. Now, we estimate using the definition of η_i, η_j:

$$m_i \sum_j \kappa_{ij} p_j \geq p_i \min(0, m_i\kappa_{ii} - \eta_i) + \frac{1}{2} p_i \eta_i + \frac{1}{2}\sum_{j\neq i} p_j \eta_j + \frac{1}{2} p_i \eta_i \frac{m_i}{m_i} + \frac{1}{2}\sum_{j\neq i} p_j \eta_i \frac{m_i}{m_j}$$
$$= p_i \min(0, m_i\kappa_{ii} - \eta_i) + \frac{1}{2}\sum_j p_j \left(\eta_j + \eta_i \frac{m_i}{m_j}\right) \geq \lambda_0 > 0.$$

□

We conclude this section with several examples for our convexity condition (III.2.5).

Example III.10. *Set, for simplicity, $m_j = 1$ and $p_j = 1$ for all $j \in \{1, \dots, n\}$. Then,* (III.2.5) *simplifies to*

$$\lambda_0 = \min_{i\in\{1,\dots,n\}} \left[\min(0, \kappa_{ii} - \eta_i) + \frac{1}{2}\sum_{j=1}^n \eta_j + \frac{n}{2}\eta_i\right] \qquad (n > 1),$$

with $\eta_i = \min_{j\neq i} \kappa_{ij}$. In the even more specific setting of two species ($n = 2$), one has $\eta_1 = \kappa_{12} = \eta_2$ and

$$\lambda_0 = \min\{\kappa_{11}, \kappa_{12}, \kappa_{22}\} + \kappa_{12} \qquad (n = 2).$$

Given the matrix κ from (W5)*, we obtain the following moduli for geodesic convexity λ_0 , respectively:*

$$\kappa = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \Longrightarrow \lambda_0 = 2, \qquad \kappa = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} \Longrightarrow \lambda_0 = 1,$$
$$\kappa = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \Longrightarrow \lambda_0 = 0, \qquad \kappa = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \Longrightarrow \lambda_0 = -2,$$
$$\kappa = \begin{pmatrix} -1 & -1 & a \\ -1 & -1 & b \\ a & b & -1 \end{pmatrix} \Longrightarrow \lambda_0 = \frac{b-5}{2} \quad (\text{for } a \geq b \geq -1).$$

III.2.2. Existence and uniqueness of gradient flow solutions

With the results of Lemma III.3 and Theorem III.7 at hand, the following statement follows thanks to [4, Ch. 11] (see Section I.2.2):

Theorem III.11 (Existence and uniqueness). *Consider* (III.1.1) *together with an initial datum $\mu^0 \in \mathscr{P}$. Then, there exists a* gradient flow *solution $\mu \in AC^2_{\mathrm{loc}}([0,\infty); (\mathscr{P}, \mathbf{W}_{\mathscr{P}}))$ to this initial-value problem: System* (III.1.1) *holds in the sense of distributions and one has $\mu(0) = \mu^0$. Moreover, with λ_0 from* (III.2.5)*, the* evolution variational estimate *holds for almost every $t > 0$ and all $\nu \in \mathscr{P}$:*

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathbf{W}_{\mathscr{P}}(\mu(t), \nu)^2 + \frac{\lambda_0}{2}\mathbf{W}_{\mathscr{P}}(\mu(t), \nu)^2 \leq \mathcal{E}(\nu) - \mathcal{E}(\mu(t)).$$

Given another initial datum $\nu^0 \in \mathscr{P}$ and the respective gradient flow solution $\nu \in AC^2_{\mathrm{loc}}([0,\infty); (\mathscr{P}, \mathbf{W}_{\mathscr{P}}))$, the following contraction estimate *holds for all $t \geq 0$:*

$$\mathbf{W}_{\mathscr{P}}(\mu(t), \nu(t)) \leq e^{-\lambda_0 t}\mathbf{W}_{\mathscr{P}}(\mu^0, \nu^0), \tag{III.2.8}$$

which implies in particular the uniqueness of solutions.

Corollary III.12 (The uniformly convex case). *If* (III.2.5) *yields $\lambda_0 > 0$, the measure*

$$\mu^\infty := (p_1, \dots, p_n)^{\mathrm{T}}\delta_{x^\infty}, \qquad \text{with} \qquad x^\infty := E\left[\sum_{j=1}^n \frac{p_j}{m_j}\right]^{-1} \in \mathbb{R}^d,$$

is the unique minimizer of $\mathcal{E}$ and the unique stationary state of (III.1.1) *on $\mathscr{P}$. It is globally asymptotically stable: the solution from Theorem III.11 converges exponentially fast in $(\mathscr{P}, \mathbf{W}_{\mathscr{P}})$ at rate λ_0 to μ^∞.*

As for scalar equations of the form (III.1.11), system (III.1.1) can be viewed as a continuum limit of a multi-particle system. To this end, we introduce the concept of *particle solutions* as a conclusion to this chapter.

Remark III.13 (Particle solutions). *Assume that the initial datum is discrete, i.e. each component μ_i^0 is a finite linear combination of Dirac measures:*

$$\mu_i^0 = \sum_{k=1}^{N_i} p_i^k \delta_{x_i^{0,k}} \quad (i = 1, \dots, n).$$

There, the $N_i \in \mathbb{N}$ particles *of species i have* mass $p_i^k > 0$ *and are at initial* position $x_i^{0,k} \in \mathbb{R}^d$*, for $k = 1, \dots, N_i$, respectively. Let $N := \sum_{i=1}^n N_i$ and let a family $\mathrm{x} = (x_i^k)$ ($k = 1, \dots, N_i$; $i = 1, \dots, n$) of L^2-absolutely continuous curves $x_i^k : [0, \infty) \to \mathbb{R}^d$ be given, such that the following initial-value problem for a system of N ordinary differential equations on $\mathbb{R}^d$ is globally solved:*

$$\frac{\mathrm{d}}{\mathrm{d}t} x_i^k(t) = -m_i \sum_{j=1}^n \sum_{l=1}^{N_j} p_j^l \nabla W_{ij}(x_i^k(t) - x_j^l(t)), \qquad x_i^k(0) = x_i^{0,k} \qquad (k = 1, \dots, N_i;\ i = 1, \dots, n). \tag{III.2.9}$$

Then it is easy to verify that the particle solution

$$\mu_i(t) = \sum_{k=1}^{N_i} p_i^k \delta_{x_i^k(t)} \quad (i = 1, \dots, n) \tag{III.2.10}$$

is the unique gradient flow solution to system (III.1.1) *with initial datum μ^0 given above. However, it is a non-trivial question if such* x *exist, since* (W1)–(W5) *do not imply global Lipschitz continuity of the r.h.s. in* (III.2.9). *Nevertheless,* (III.2.9) *admits L^2-absolutely continuous solutions since this system possesses an underlying (discrete) gradient flow structure: Define the finite-dimensional space*

$$\mathscr{P}_\mathrm{d} := \left\{ \mathrm{x} \in \prod_{i=1}^n \prod_{k=1}^{N_i} \mathbb{R}^d \cong \mathbb{R}^{Nd} :\ p_i = \sum_{k=1}^{N_i} p_i^k\ (i = 1, \dots, n);\ E = \sum_{i=1}^n \frac{1}{m_i} \sum_{k=1}^{N_i} p_i^k x_i^k \right\},$$

endowed with the (weighted Euclidean) distance

$$\mathbf{d}(\mathrm{x}, \mathrm{y}) = \left[\sum_{i=1}^n \frac{1}{m_i} \sum_{k=1}^{N_i} p_i^k |x_i^k - y_i^k|^2 \right]^{1/2},$$

and define the discrete interaction energy $\mathcal{E}_\mathrm{d}$ *on $\mathscr{P}_\mathrm{d}$ as*

$$\mathcal{E}_\mathrm{d}(\mathrm{x}) := \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} p_i^k p_j^l W_{ij}(x_i^k - x_j^l).$$

Applying the same method of proof as for Theorem III.7 mutatis mutandis *for the discrete framework, one can show that $\mathcal{E}_\mathrm{d}$ is λ_0-geodesically convex on $(\mathscr{P}_\mathrm{d}, \mathbf{d})$ with* the same *modulus of convexity λ_0 as in the continuous case* (III.2.5). *We can again invoke [4] to obtain the existence and uniqueness of a solution curve* $\mathrm{x} \in AC^2_{\mathrm{loc}}([0, \infty); (\mathscr{P}_\mathrm{d}, \mathbf{d}))$ *to the particle system* (III.2.9). *Conversely, thanks to the uniqueness of solutions to both* (III.1.1) *and* (III.2.9)*, a gradient flow solution μ to* (III.1.1) *of the form* (III.2.10) *can be represented by a solution* x *to* (III.2.9).

CHAPTER III.3

Confinement and qualitative behaviour of solutions

In this chapter, we analyse the qualitative behaviour of the solution from Theorem III.11 in the general scenario, i.e. the criterion for geodesic convexity may only yield $\lambda_0 \leq 0$. In this case, the contraction estimate (III.2.8) does not allow for conclusions on the long-time behaviour of the solution.

From now on, consider (III.1.1) in one spatial dimension $d = 1$; and let μ be the solution to (III.1.1) with initial datum $\mu^0 \in \mathscr{P}$, as given in Theorem III.11. First, we rewrite system (III.1.1) in terms of the inverse distribution functions $u = (u_1, \ldots, u_n)$; recall their definition from (III.1.6)&(III.1.7).

For all $z \in (0,1)$, one has $z = F_i(t, u_i(t,z))$. Differentiation w.r.t. t yields

$$\begin{aligned}0 &= \partial_t F_i(t, u_i(t,z)) + \partial_x F_i(t, u_i(t,z))\partial_t u_i(t,z) \\ &= \int_{-\infty}^{u_i(t,z)} \frac{1}{p_i}\partial_y \left(\sum_{j=1}^n m_i \mu_i(t,y)(W'_{ij} * \mu_j)(t,y)\right) \mathrm{d}y + \frac{1}{p_i}\mu_i(t,u_i(t,z))\partial_t u_i(t,z) \\ &= \frac{m_i}{p_i}\sum_{j=1}^n \mu_i(t,u_i(t,z)) \int_{\mathbb{R}} W'_{ij}(u_i(t,z) - y)\,\mathrm{d}\mu_j(t,y) + \frac{1}{p_i}\mu_i(t,u_i(t,z))\partial_t u_i(t,z).\end{aligned}$$

Rearranging yields with the help of (W3) and the transformation $\xi := F_j(t,y)$:

$$\partial_t u_i(t,z) = m_i \sum_{j=1}^n p_j \int_0^1 W'_{ij}(u_j(t,\xi) - u_i(t,z))\,\mathrm{d}\xi \qquad (i = 1, \ldots, n). \tag{III.3.1}$$

It is a consequence of Theorem III.11 that given a gradient flow solution μ to (III.1.1), the corresponding curve of pseudo-inverse distribution functions $u \in AC^2_{\mathrm{loc}}([0,\infty); L^2([0,1];\mathbb{R}^n))$ solves (III.3.1). Furthermore, since $\mu(t) \in \mathscr{P}$ for all $t \geq 0$, $u_i(t,\cdot)$ is a non-decreasing *càdlàg* function on $(0,1)$. Conservation of the weighted center of mass E over time is reflected in terms of u by the identity

$$E = \sum_{j=1}^n \frac{p_j}{m_j}\int_0^1 u_j(t,z)\,\mathrm{d}z \qquad \forall t \geq 0. \tag{III.3.2}$$

The concept of inverse distribution functions substantially simplifies the analysis of solutions to (III.1.1) since there does not appear any spatial derivative on the right-hand side of (III.3.1) anymore. However, this approach can be employed in one spatial dimension $d = 1$ only.

III.3.1. The purely quadratic case

This paragraph is devoted to another specific example for system (III.1.1), namely the case where all entries in W are purely quadratic functions, i.e. $W_{ij}(z) = \frac{1}{2}\kappa_{ij}z^2$. There, it is possible to solve system (III.3.1) analytically: we obtain by elementary calculations — involving the usage of (III.3.2) — that

$$\partial_t u_i(t,z) = -m_i\left(\sum_j \kappa_{ij}p_j\right)u_i(t,z) + m_i^2\kappa_{ii}E + \sum_{j\neq i} m_i\left(\kappa_{ij} - \kappa_{ii}\frac{m_i}{m_j}\right)\int_0^1 p_j u_j(t,\xi)\,\mathrm{d}\xi. \tag{III.3.3}$$

Define $v_i(t) := \int_0^1 u_i(t,z)\,\mathrm{d}z$ and integrate (III.3.3) w.r.t. $z \in (0,1)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} v_i(t) = -m_i \left(\sum_j \kappa_{ij} p_j \right) v_i(t) + m_i^2 \kappa_{ii} E + \sum_{j \neq i} m_i p_j \left(\kappa_{ij} - \kappa_{ii} \frac{m_i}{m_j} \right) v_j(t). \tag{III.3.4}$$

Consequently, with the definitions

$$A_{ii} := -m_i \sum_{j=1}^{n} \kappa_{ij} p_j, \quad A_{ij} := m_i p_j \left(\kappa_{ij} - \kappa_{ii} \frac{m_i}{m_j} \right) \quad (i \neq j), \quad b_i := m_i^2 \kappa_{ii},$$

for $i, j \in \{1, \ldots, n\}$, the following holds for $v = (v_1, \ldots, v_n)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} v(t) = Av(t) + Eb, \qquad \partial_t (u(t,z) - v(t)) = \operatorname{diag}(A)(u(t,z) - v(t)). \tag{III.3.5}$$

There, $\operatorname{diag}(A)$ is meant to be the diagonal matrix with the same diagonal as A. The linear systems in (III.3.5) can easily be solved; we eventually obtain:

$$\begin{aligned} u(t,z) = \exp(\operatorname{diag}(A)t) \left(u(0,z) - \int_0^1 u(0,\xi)\,\mathrm{d}\xi \right) + \exp(At) \int_0^1 u(0,\xi)\,\mathrm{d}\xi \\ + \left(\sum_{j=1}^{n} \int_0^1 \frac{p_j}{m_j} u_j(0,\xi)\,\mathrm{d}\xi \right) \int_0^t \exp(A(t-s)) b\,\mathrm{d}s \quad \text{for all } z \in (0,1). \end{aligned}$$

We expect exponential convergence of u to the spatially constant equilibrium $(x^\infty, \ldots, x^\infty) \in \mathbb{R}^n$ as $t \to \infty$ if both A and $\operatorname{diag}(A)$ possess eigenvalues with negative real parts only. Clearly, our result on geodesic convexity (cf. Section III.2) shows that if $\lambda_0 > 0$ in (III.2.5), these conditions are bound to hold. The necessary condition from Remark III.9 implies that $\operatorname{diag}(A)$ is negative definite.

The specific case of two species ($n = 2$) deserves a closer look. Thanks to the invariant (III.3.2), the two equations for u_1 and u_2 in (III.3.1) can be separated completely:

$$\begin{aligned} \partial_t u_1(t,z) &= -m_1(\kappa_{11} p_1 + \kappa_{12} p_2) u_1(t,z) + p_1(m_1 \kappa_{11} - m_2 \kappa_{12}) v_1(t) + m_1 m_2 \kappa_{12} E, \\ \partial_t u_2(t,z) &= -m_2(\kappa_{12} p_1 + \kappa_{22} p_2) u_2(t,z) + p_2(m_2 \kappa_{22} - m_1 \kappa_{12}) v_2(t) + m_1 m_2 \kappa_{12} E, \\ \frac{\mathrm{d}}{\mathrm{d}t} v_i(t) &= -v_i(t)(m_1 p_2 + m_2 p_1)\kappa_{12} + m_1 m_2 \kappa_{12} E \quad (\text{for both } i \in \{1,2\}). \end{aligned}$$

Comparison with our criterion for geodesic convexity (III.2.5) shows that the solution to the system above — for generic initial data — is unbounded in time if $\lambda_0 < 0$ in (III.2.5). For example, if $m_1 = 1 = m_2$ and $p_1 = 1 = p_2$, one easily sees that $\kappa_{12} > 0$, $\kappa_{11} + \kappa_{12} \geq 0$ and $\kappa_{22} + \kappa_{12} \geq 0$ are necessary for a bounded solution. These conditions are equivalent to $\lambda_0 \geq 0$ in the case of an irreducible system of two components.

III.3.2. Speed of propagation and confinement

In this section, we investigate the rate of propagation of the solution to (III.1.1) in space over time, given an initial datum with compact support. We first obtain — for arbitrary potentials satisfying (W1)–(W5) — boundedness of the support of $\mu(t)$ for fixed time $t > 0$, and second — under more restrictive requirements on the potential W — *t-uniform* boundedness of $\operatorname{supp} \mu(t)$.

Proposition III.14 (Finite speed of propagation). *Let an initial datum μ^0 with compact support and $T > 0$ be given. Then, there exists a constant $K = K(T, \mu^0) > 0$ such that for all $t \in [0, T]$,*

$$\operatorname{supp} \mu(t) \subset [-K, K].$$

Proof. For $t \in [0,T]$ and $i \in \{1,\dots,n\}$, denote

$$u_i(t,1^-) := \lim_{\varepsilon\searrow 0} u_i(t,1-\varepsilon) \in \mathbb{R}\cup\{+\infty\} \quad \text{and} \quad u_i(t,0^+) := \lim_{\varepsilon\searrow 0} u_i(t,\varepsilon) \in \mathbb{R}\cup\{-\infty\}.$$

The assertion will follow from finiteness of those limits. Let $\varepsilon > 0$. Then,

$$\partial_t(u_i(t,1-\varepsilon)^2) \le 2|u_i(t,1-\varepsilon)| \sum_{j=1}^n m_i p_j \int_0^1 |W'_{ij}(u_j(t,\xi) - u_i(t,1-\varepsilon))| \,\mathrm{d}\xi.$$

Lemma III.4, Hölder's and Young's inequality eventually lead to

$$\begin{aligned}
&2|u_i(t,1-\varepsilon)| \sum_{j=1}^n m_i p_j \int_0^1 |W'_{ij}(u_j(t,\xi) - u_i(t,1-\varepsilon))| \,\mathrm{d}\xi \\
&\le 2|u_i(t,1-\varepsilon)| \sum_{j=1}^n \overline{C}_{ij} m_i p_j \left(\int_0^1 |u_j(t,\xi)| \,\mathrm{d}\xi + |u_i(t,1-\varepsilon)| + 1\right) \qquad \text{(III.3.6)}\\
&\le 2\left[\sum_{j=1}^n \overline{C}_{ij} m_i p_j + 1\right] u_i(t,1-\varepsilon)^2 + \left(\sum_{j=1}^n \overline{C}_{ij} m_i p_j\right)^2 + 2\max_j \left(\overline{C}_{ij}^2 m_i^2 m_j p_j\right) \sum_{j=1}^n \int_0^1 \frac{p_j}{m_j} u_j(t,\xi)^2 \,\mathrm{d}\xi.
\end{aligned}$$

With the transformation $\xi := F_j(t,x)$, we observe that the sum in the last term on the right-hand side of (III.3.6) can be expressed in terms of the second moments $\mathfrak{m}_2(\mu_j(t))$ and of $\mathbf{W}_{\mathscr{P}}(\mu(t),\delta_0 \mathsf{e})$, where $\mathsf{e} = (1,1,\dots,1)^{\mathrm{T}} \in \mathbb{R}^n$:

$$\sum_{j=1}^n \int_0^1 \frac{p_j}{m_j} u_j(t,\xi)^2 \,\mathrm{d}\xi = \sum_{j=1}^n \frac{1}{m_j} \mathfrak{m}_2(\mu_j(t)) = \mathbf{W}^2_{\mathscr{P}}(\mu(t),\delta_0\mathsf{e}).$$

Since $\mu \in AC^2([0,T];(\mathscr{P},\mathbf{W}_{\mathscr{P}}))$, there exists $\varphi \in L^2([0,T])$ such that

$$\mathbf{W}_{\mathscr{P}}(\mu(t),\mu^0) \le \int_0^t \varphi(s)\,\mathrm{d}s \qquad \forall\, t \in [0,T].$$

We obtain

$$\begin{aligned}
\sum_{j=1}^n \int_0^1 \frac{p_j}{m_j} u_j(t,\xi)^2 \,\mathrm{d}\xi &\le 2\mathbf{W}^2_{\mathscr{P}}(\mu(t),\mu^0) + 2\mathbf{W}^2_{\mathscr{P}}(\mu^0,\delta_0\mathsf{e}) \le 2\left(\int_0^t \varphi(s)\,\mathrm{d}s\right)^2 + 2\mathbf{W}^2_{\mathscr{P}}(\mu^0,\delta_0\mathsf{e}) \\
&\le 2T\|\varphi\|^2_{L^2([0,T])} + 2\sum_{j=1}^n \frac{1}{m_j}\mathfrak{m}_2(\mu_j^0),
\end{aligned}$$

which is a constant depending on T and μ^0. Inserting into (III.3.6), we observe

$$\partial_t(u_i(t,1-\varepsilon)^2) \le A u_i(t,1-\varepsilon)^2 + B(T,\mu^0),$$

for suitable constants $A, B > 0$. We apply Gronwall's lemma, let $\varepsilon \searrow 0$ and use that — since μ^0 has compact support by assumption — the limit $u_i(0,1^-)$ exists in $\mathbb{R}$:

$$u_i(t,1^-)^2 \le \left[u_i(0,1^-)^2 + \frac{B}{A}\right] \exp(AT) \quad \forall\, t \in [0,T].$$

Thus, $u_i(t,1^-)$ is a finite value, at each $t \in [0,T]$. Along the same lines, it can be shown that

$$u_i(t,0^+)^2 \le \left[u_i(0,0^+)^2 + \frac{B}{A}\right] \exp(AT) \quad \forall\, t \in [0,T],$$

hence $u_i(t,0^+)$ is finite as well. Since by definition $\operatorname{supp}\mu_i(t) \subset [u_i(t,0^+), u_i(t,1^-)]$, the assertion is proved. □

The statement of Proposition III.14 shows that at fixed $t > 0$, the limits

$$u_i(t, 1^-) := \lim_{\varepsilon \searrow 0} u_i(t, 1-\varepsilon) \quad \text{and} \quad u_i(t, 0^+) := \lim_{\varepsilon \searrow 0} u_i(t, \varepsilon)$$

exist (in $\mathbb{R}$), if they exist at $t = 0$. In order to prove uniform confinement of the solution, we show t-uniform boundedness of those limits. We first introduce a requirement on the potential by the following

Definition III.15 (Confining potentials). *We call an interaction potential W satisfying* (W1)–(W5) confining *if there exists $R > 0$ such that:*

(i) System (III.1.1) *is* irreducible at large distance from the origin, *that is, the graph $G' = (V_{G'}, E_{G'})$ with nodes $V_{G'} = \{1, \dots, n\}$ and edges $E_{G'} = \{(i,j) \in V_{G'} \times V_{G'} : W'_{ij} \not\equiv 0 \text{ on } (R, \infty)\}$ is connected.*

(ii) There exists a matrix $C \in \mathbb{R}^{n\times n}$ such that for each $i, j \in \{1, \dots, n\}$, the map W_{ij} is C_{ij}-(semi-)convex on the interval (R, ∞) and the following holds:
If $n = 1$, then $C > 0$. If $n > 1$, with $\widetilde{\eta}_i := \min_{j\neq i} C_{ij} p_j$ for all $i \in \{1, \dots, n\}$,

$$\widetilde{\lambda}_0 := \min_{i\in\{1,\dots,n\}} \left[p_i \min(0, m_i C_{ii} - \widetilde{\eta}_i) + \frac{1}{2} \sum_{j=1}^{n} p_j \left(\widetilde{\eta}_j + \widetilde{\eta}_i \frac{m_i}{m_j} \right) \right] > 0. \tag{III.3.7}$$

Remark III.16 (Geodesic convexity and confinement). *In the scalar case $n = 1$, uniform geodesic convexity of the interaction energy $\mathcal{E}$ is* equivalent *to κ-convexity of W with $\kappa > 0$ [133]. So, the potential is confining. Also for genuine systems, if $\lambda_0 > 0$ in* (III.2.5), *the definition $C := \kappa$ yields $\widetilde{\lambda}_0 = \lambda_0 > 0$. Hence, our criterion for uniform geodesic convexity of $\mathcal{E}$ necessarily implies that W is a confining potential. Naturally, if the system is not irreducible at large distance from the origin, the independent irreducible subsystems should be considered separately.*

Theorem III.17 (Confinement). *Assume that W is confining and let μ^0 have compact support. Then, there exists a constant $K = K(\mu^0) > 0$ independent of t such that for all $t \geq 0$:*

$$\operatorname{supp} \mu(t) \subset [-K, K]. \tag{III.3.8}$$

Proof. We prove the assertion in the case of genuine systems $n > 1$.
Step 1: L^2 estimate.
Let $\varepsilon > 0$ be sufficiently small such that replacing C_{ij} by $C^\varepsilon_{ij} := C_{ij} - \varepsilon$ in (III.3.7) still yields a number $\widetilde{\lambda}^\varepsilon_0 > 0$, possibly with $\widetilde{\lambda}^\varepsilon_0 < \widetilde{\lambda}_0$. From the C_{ij}-convexity of W_{ij} on (R, ∞) and with the help of Young's inequality, we get for all $z > R$:

$$W_{ij}(z) \geq W_{ij}(R) + W'_{ij}(R)(z-R) + \frac{1}{2} C_{ij}(z-R)^2 \geq \frac{1}{2} C^\varepsilon_{ij} z^2 - D_{ij},$$

for appropriate constants $D_{ij} > 0$. Thanks to (W2)&(W3), enlarging the constants, there exists $D > 0$ such that for all $i, j \in \{1, \dots, n\}$ and *for all $z \in \mathbb{R}$*:

$$W_{ij}(z) \geq \frac{1}{2} C^\varepsilon_{ij} z^2 - D. \tag{III.3.9}$$

We now use the monotonicity of the energy $\mathcal{E}$ along the (gradient flow) solution to obtain with (III.3.9) for all $t \geq 0$:

$$2\mathcal{E}(\mu^0) \geq 2\mathcal{E}(\mu(t)) \geq \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_{\mathbb{R}} \int_{\mathbb{R}} C^\varepsilon_{ij} (x-y)^2 \, \mathrm{d}\mu_j(x) \, \mathrm{d}\mu_i(y) - D \left(\sum_{j=1}^{n} p_j \right)^2.$$

The first term on the r.h.s. has precisely the same structure as the l.h.s. in (III.2.6) for $t_i = t_j \equiv 0$ and $\widetilde{t}_i = \widetilde{t}_j = \mathrm{id}$. Arguing exactly as in the proof of Theorem III.7, we obtain

$$2\mathcal{E}(\mu^0) \geq \widetilde{\lambda}_0^{\varepsilon} \sum_{j=1}^{n} \frac{1}{m_j} \mathbf{m}_2(\mu_j(t)) - D \left(\sum_{j=1}^{n} p_j \right)^2 .$$

All in all, we have proved uniform boundedness of the second moments: there exists $C_2 > 0$ such that for all $t \geq 0$ and all $i \in \{1, \dots, n\}$, one has $\mathbf{m}_2(\mu_i(t)) \leq C_2$.

Step 2: L^∞ estimate.

We first prove an upper bound. For each $t \geq 0$, we consider those indices $i \in \{1, \dots, n\}$, where

$$u_i(t, 1^-) \geq \max_{j \in \{1, \dots, n\}} u_j(t, 1^-) - R.$$

That is, for all $\xi \in [0, 1)$ and all $j \in \{1, \dots, n\}$: $u_i(t, 1^-) \geq u_j(t, \xi) - R$. We thus have, for each $j \in \{1, \dots, n\}$, a partition of $[0, 1)$ into two sets A_1^j and A_2^j, where

$$A_1^j := \{\xi \in [0,1) : u_j(t,\xi) - u_i(t,1^-) < -R\}, \qquad A_2^j := \{\xi \in [0,1) : |u_j(t,\xi) - u_i(t,1^-)| \leq R\}.$$

Since W'_{ij} is continuous thanks to (W2), it is bounded on the interval $[-R, R]$. The C_{ij}-convexity of W on $(-\infty, -R)$ yields

$$W'_{ij}(z) - C_{ij} z \leq W'_{ij}(-R) - C_{ij}(-R) \qquad \forall z < -R,$$

which can be rewritten as follows using (W3):

$$W'_{ij}(z) \leq C_{ij} z + C_{ij} R - W'_{ij}(R) \qquad \forall z < -R.$$

Hence, we obtain

$$\partial_t u_i(t, 1^-) \leq \sum_{j=1}^{n} \int_{A_1^j} m_i C_{ij} p_j (u_j(t,\xi) - u_i(t,1^-)) \, \mathrm{d}\xi + C_0,$$

for some constant $C_0 > 0$. Then, with the help of Hölder's and Young's inequality,

$$\begin{aligned}
&\sum_{j=1}^{n} \int_{A_1^j} m_i C_{ij} p_j (u_j(t,\xi) - u_i(t,1^-)) \, \mathrm{d}\xi \\
&= m_i \sum_{j=1}^{n} C_{ij} p_j \left[\int_0^1 (u_j(t,\xi) - u_i(t,1^-)) \, \mathrm{d}\xi - \int_{A_2^j} (u_j(t,\xi) - u_i(t,1^-)) \, \mathrm{d}\xi \right] \\
&\leq -m_i \sum_{j=1}^{n} C_{ij} p_j u_i(t,1^-) + C' \sum_{j=1}^{n} \int_0^1 p_j u_j(t,\xi)^2 \, \mathrm{d}\xi + C_1 = -m_i \sum_{j=1}^{n} C_{ij} p_j u_i(t,1^-) + C' \sum_{j=1}^{n} \mathbf{m}_2(\mu_j(t)) + C_1,
\end{aligned}$$

for some constants $C', C_1 > 0$. We now employ step 1 and observe that, as in Remark III.9, we have $\sum_{j=1}^{n} C_{ij} p_j \geq \widetilde{\lambda}_0 > 0$:

$$\partial_t u_i(t, 1^-) \leq -m_i \widetilde{\lambda}_0 u_i(t, 1^-) + C'',$$

for $C'' > 0$. Gronwall's lemma yields — thanks to $u_i(0, 1^-) < \infty$ — the existence of a constant $K > 0$ such that $\max_{j \in \{1,\dots,n\}} u_j(t, 1^-) \leq K$ for all $t \geq 0$.

In analogy, we now consider those $i \in \{1, \dots, n\}$ such that

$$u_i(t, 0^+) \leq \min_{j \in \{1, \dots, n\}} u_j(t, 0^+) + R,$$

yielding for each $j \in \{1,\dots,n\}$ a partition $[0,1) = B_1^j \cup B_2^j$ with

$$B_1^j := \{\xi \in [0,1) : u_j(t,\xi) - u_i(t,0^+) > R\}, \qquad B_2^j := \{\xi \in [0,1) : |u_j(t,\xi) - u_i(t,0^+)| \le R\}.$$

Similarly to step 1, using the symmetry property (W3), we get

$$-\partial_t u_i(t,0^+) \le -m_i \sum_{j=1}^n C_{ij} p_j(-u_i(t,0^+)) + C' \sum_{j=1}^n \mathfrak{m}_2(\mu_j(t)) + C_1 \le -m_i \widetilde{\lambda}_0(-u_i(t,0^+)) + C'',$$

allowing us to proceed as before.
Putting the bounds together finishes the proof: $\sup\limits_{t\ge 0} \max\limits_{j\in\{1,\dots,n\}} \|u_j(t,\cdot)\|_{L^\infty([0,1])} \le K$. □

We thus know, given a confining potential, that the solution lives on a fixed compact interval. It is now a natural question to ask if, for absolutely continuous initial data, partial or total collapse of the support can occur in *finite* time. This question is addressed in the following

Proposition III.18 (Exclusion of finite-time blow-up). *Let $i \in \{1,\dots,n\}$ be fixed, but arbitrary. Assume that for all $j \in \{1,\dots,n\}$ the maps W_{ij}' are Lipschitz continuous. Suppose moreover that* $\operatorname{supp} \mu_i^0$ *is a (possibly unbounded) interval and μ_i^0 is absolutely continuous w.r.t. the Lebesgue measure. Assume that its Lebesgue density is continuous on the interior of* $\operatorname{supp} \mu_i^0$ *and globally bounded. Then, $\mu_i(t)$ is absolutely continuous for all $t \ge 0$.*

Proof. Our method of proof is an adaptation of the proof of [35, Thm. 2.9] to the situation at hand. We show that for all $t \ge 0$, there exists $\gamma(t) > 0$ such that for all $z \in (0,1)$ and all $h > 0$ with $z + h < 1$:

$$\frac{1}{h}(u_i(t,z+h) - u_i(t,z)) \ge \gamma(t) > 0. \tag{III.3.10}$$

That is, $u_i(t,\cdot)$ is *strictly* increasing at each $t \ge 0$. The assumptions on the initial datum above ensure that (III.3.10) is true at $t = 0$ with some $\gamma(0) > 0$. If (III.3.10) holds at a given t_0, the cumulative distribution function $F_i(t_0,\cdot)$ is Lipschitz continuous, which implies absolute continuity of $\mu_i(t_0)$.

From (III.3.1), we get

$$\partial_t(u_i(t,z+h) - u_i(t,z)) = m_i \sum_{j=1}^n p_j \int_0^1 \left[W_{ij}'(u_j(t,\xi) - u_i(t,z+h)) - W_{ij}'(u_j(t,\xi) - u_i(t,z)) \right] \mathrm{d}\xi.$$

Denote by $L_{ij} > 0$ the Lipschitz constant of W_{ij}'. From the monotonicity $u_i(t,z+h) - u_i(t,z) \ge 0$, it follows that

$$\partial_t(u_i(t,z+h) - u_i(t,z)) \ge -m_i \sum_{j=1}^n L_{ij} p_j (u_i(t,z+h) - u_i(t,z)).$$

We subsequently obtain for $\widetilde{C}_i := m_i \sum\limits_{j=1}^n L_{ij} p_j$ that $\partial_t[(u_i(t,z+h) - u_i(t,z))e^{\widetilde{C}_i t}] \ge 0$, and hence

$$\frac{1}{h}(u_i(t,z+h) - u_i(t,z)) \ge \frac{1}{h} e^{-\widetilde{C}_i t}(u_i(0,z+h) - u_i(0,z)) \ge e^{-\widetilde{C}_i t}\gamma(0) > 0.$$

Letting $\gamma(t) := e^{-\widetilde{C}_i t}\gamma(0)$, (III.3.10) follows. □

Naturally, the above result does not extend to $t \to \infty$ since e.g. in the uniformly geodesically convex case, the solution collapses to a Dirac measure in the limit $t \to \infty$.

III.3.3. Long-time behaviour

We now analyse the long-time behaviour of the solution to (III.1.1) in the non-uniformly convex case.

Theorem III.19 (Long-time behaviour). *Assume that the solution μ to* (III.1.1) *is* uniformly confined, *i.e. there exists $K > 0$ such that* $\operatorname{supp} \mu_i(t) \subset [-K,K]$ *holds for all $t \ge 0$ and all $i \in \{1,\dots,n\}$ as in* (III.3.8).

Moreover, suppose that the maps W_{ij}' are Lipschitz continuous on the interval $[-2K, 2K]$ for all $i, j \in \{1, \dots, n\}$. Set, for $t \geq 0$, $\mathcal{E}^t := \mathcal{E}(\mu(t))$. The following holds:

(a) There exists $\mathcal{E}^\infty \in \mathbb{R}$ such that

$$\lim_{t\to\infty} \mathcal{E}^t = \mathcal{E}^\infty, \tag{III.3.11}$$

and

$$\lim_{t\to\infty} \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^t\right) = 0. \tag{III.3.12}$$

(b) For each sequence $(t_k)_{k\in\mathbb{N}}$ in $(0,\infty)$ with $t_k \to \infty$ as $k \to \infty$, there exists a subsequence $(t_{k_l})_{l\in\mathbb{N}}$ and a steady state $\overline{\mu} \in \mathscr{P}$ of (III.1.1) *such that for all $i \in \{1, \dots, n\}$:*

$$\lim_{l\to\infty} \mathbf{W}_1(\mu_i(t_{k_l}), \overline{\mu}_i) = 0. \tag{III.3.13}$$

Thus, the ω-limit set of the dynamical system associated to (III.1.1) *can only contain steady states of* (III.1.1).

Proof. We proceed similarly to the proof of [158, Prop. 1] and observe that along the solution μ, the dissipation of $\mathcal{E}$ reads

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^t &= -\sum_{i=1}^n m_i \int_{\mathbb{R}} \left(\sum_{j=1}^n W_{ij}' * \mu_j(t)\right)^2 \mathrm{d}\mu_i(t) \\ &= -\sum_{i=1}^n m_i p_i \int_0^1 \left(\sum_{j=1}^n p_j \int_0^1 W_{ij}'(u_i(t,z) - u_j(t,\xi))\,\mathrm{d}\xi\right)^2 \mathrm{d}z,\end{aligned} \tag{III.3.14}$$

which is nonpositive. By (III.3.8), all u_i are bounded in time and space by the constant K. Since W_{ij}' is Lipschitz continuous, it is differentiable almost everywhere on $[-2K, 2K]$. So, another differentiation of the dissipation w.r.t. t shows

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{E}^t = -2\sum_{i=1}^n m_i p_i \int_0^1 &\left(\sum_{j=1}^n \int_0^1 p_j W_{ij}'(u_i(t,z) - u_j(t,\zeta))\,\mathrm{d}\zeta\right) \\ &\cdot \left(\sum_{j=1}^n p_j \int_0^1 W_{ij}''(u_i(t,z) - u_j(t,\xi)) \right. \\ &\quad \left. \cdot \left[\sum_{k=1}^n m_i p_k \int_0^1 W_{ik}'(u_k(t,\zeta) - u_i(t,z))\,\mathrm{d}\zeta - \sum_{k=1}^n m_j p_k \int_0^1 W_{jk}'(u_k(t,\zeta) - u_j(t,\xi))\,\mathrm{d}\zeta\right] \mathrm{d}\xi\right) \mathrm{d}z.\end{aligned}$$

By elementary estimates, using in particular that $|W_{ij}''(z)| \leq L_{ij}$ a.e. on $[-2K, 2K]$ by Lipschitz continuity, we find

$$\sup_{t\geq 0} \left|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{E}^t\right| \leq C_2, \tag{III.3.15}$$

for some $C_2 > 0$. Furthermore, it is easy to conclude from (W2) and (III.3.8) that

$$\inf_{t\geq 0} \mathcal{E}^t \geq -C_0, \tag{III.3.16}$$

for another constant $C_0 > 0$. Putting (III.3.14) and (III.3.16) together yields the existence of $\mathcal{E}^\infty \in \mathbb{R}$ such that (III.3.11) holds. We now use (III.3.15) to prove (III.3.12): Define for $t > \frac{2}{C_2}\sqrt{\mathcal{E}^0 - \mathcal{E}^\infty} > 0$ the quantity $\tau(t) := \frac{1}{C_2}\sqrt{\mathcal{E}^{t/2} - \mathcal{E}^\infty} > 0$. Since $\mathcal{E}^t$ is nonincreasing, we also have $\tau(t) < \frac{t}{2}$. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^t = \frac{1}{\tau(t)}(\mathcal{E}^t - \mathcal{E}^\infty) - \frac{1}{\tau(t)}(\mathcal{E}^{t-\tau(t)} - \mathcal{E}^\infty) + \frac{1}{\tau(t)}\int_{t-\tau(t)}^t \int_s^t \frac{\mathrm{d}^2}{\mathrm{d}\sigma^2}\mathcal{E}^\sigma\,\mathrm{d}\sigma\,\mathrm{d}s,$$

from which with (III.3.15) and (III.3.11) the desired result (III.3.12) follows:

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^t\right| &\leq \frac{1}{\tau(t)}(\mathcal{E}^{t/2}-\mathcal{E}^\infty)+\frac{1}{\tau(t)}(\mathcal{E}^{t-t/2}-\mathcal{E}^\infty)+\int_{t-\tau(t)}^t\int_s^t C_2\,\mathrm{d}\sigma\,\mathrm{d}s\\ &\leq \frac{2}{\tau(t)}(\mathcal{E}^{t/2}-\mathcal{E}^\infty)+\frac{1}{\tau(t)}C_2\tau(t)^2=(2C_2+1)\sqrt{\mathcal{E}^{t/2}-\mathcal{E}^\infty}\overset{t\to\infty}{\longrightarrow}0.\end{aligned}$$

For part (b), let a sequence of time points $t_k \to \infty$ be given. The family $u_i(t_k,\cdot)$ ($k \in \mathbb{N}$; $i \in \{1,\dots,n\}$) of nondecreasing functions is uniformly bounded in $L^\infty([0,1])$ (by the constant K) — hence bounded in the space $BV([0,1])$. Thus, there exist a subsequence $(t_{k_l})_{l\in\mathbb{N}}$ and nondecreasing maps $\overline{u}_1,\dots,\overline{u}_n$ with $\|\overline{u}_i\|_{L^\infty([0,1])} \leq K$ for all $i \in \{1,\dots,n\}$, such that $u_i(t_{k_l},\cdot)$ converges to $\overline{u}_i$ in $L^1([0,1])$ and almost everywhere on $[0,1]$, as $l \to \infty$ (for details, see e.g. [3]). The corresponding measure $\overline{\mu}$ belongs to $\mathscr{P}$ thanks to the dominated convergence theorem. It remains to show that $\overline{\mu}$ is a steady state of system (III.1.1). Define the number

$$\omega := -\sum_{i=1}^n m_i p_i \int_0^1 \left(\sum_{j=1}^n p_j \int_0^1 W_{ij}'(\overline{u}_i(z)-\overline{u}_j(\xi))\,\mathrm{d}\xi\right)^2 \mathrm{d}z.$$

Elementary calculations — involving Lemma III.4 and the elementary identity $a^2-b^2=(a-b)(a+b)$ — show

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^{t_{k_l}}-\omega\right| \leq C_0\sum_{i=1}^n\sum_{j=1}^n\int_0^1\int_0^1\left|W_{ij}'(u_i(t_{k_l},z)-u_j(t_{k_l},\xi))-W_{ij}'(\overline{u}_i(z)-\overline{u}_j(\xi))\right|\mathrm{d}\xi\,\mathrm{d}z,$$

for a suitable constant $C_0>0$. The Lipschitz continuity of the W_{ij}' on $[-2K,2K]$, the triangle inequality and (III.3.13) then imply

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^{t_{k_l}}-\omega\right| \leq C_0\sum_{i=1}^n\sum_{j=1}^n L_{ij}\big(\|u_i(t_{k_l},\cdot)-\overline{u}_i\|_{L^1([0,1])}+\|u_j(t_{k_l},\cdot)-\overline{u}_j\|_{L^1([0,1])}\big)\overset{l\to\infty}{\longrightarrow}0.$$

Hence, because of (III.3.12), $\omega=0$. Specifically, this means that for each $i\in\{1,\dots,n\}$ and almost every $x\in\operatorname{supp}\overline{\mu}_i$, the following holds:

$$\sum_{j=1}^n\int_{\mathbb{R}}W_{ij}'(x-y)\,\mathrm{d}\overline{\mu}_j(y)=0.$$

So, $\overline{\mu}$ is a t-independent solution to (III.1.1) and the proof is complete. □

Remark III.20. *The result of Theorem III.19 does neither yield the uniqueness of steady states of* (III.1.1) *nor convergence of the entire curve μ to some specific object as $t\to\infty$. Due to the non-strict convexity of $\mathcal{E}$, the ω-limit set might contain more than one element. Still, as it is also the case for a finite-dimensional gradient system, every ω-limit point is an equilibrium. Our result may be viewed as an extension of the well-known* LaSalle invariance principle *from the theory of ordinary differential equations: The existence of the Lyapunov functional $\mathcal{E}$ and certain compactness properties (here, the confinement of the solution) guarantee convergence of the solution orbit to the set with vanishing energy dissipation.*

If there only exists the trivial steady state μ^∞ from Corollary III.12 in the set of those elements of $\mathscr{P}$ with support contained in $[-K,K]$, then Theorem III.19 implies

$$\lim_{t\to\infty}\mathbf{W}_1(\mu_i(t),\mu_i^\infty)=0\qquad\forall i\in\{1,\dots,n\},$$

without obtaining any specific rate of convergence.

If the potential is not confining, convergence may not occur. However, we might observe a *δ-separation* phenomenon: the support of each component μ_i collapses to a single (but not necessarily steady) point as $t \to \infty$.

Proposition III.21 (δ-separation). *Let $i \in \{1, \dots, n\}$ be fixed, but arbitrary. Assume that the support of μ_i^0 is compact and that $S_i := \sum_{j=1}^{n} \kappa_{ij} p_j > 0$ holds. Then,*

$$\operatorname{diam}\operatorname{supp}\mu_i(t) \le e^{-m_i S_i t} \operatorname{diam}\operatorname{supp}\mu_i^0.$$

That is, the support of μ_i contracts at exponential speed.

Proof. Recall that the diameter of the support here is given by $\operatorname{diam}\operatorname{supp}\mu_i(t) = u_i(t, 1^-) - u_i(t, 0^+)$. We have

$$\partial_t(u_i(t,1^-) - u_i(t,0^+)) = \sum_{j=1}^{n} m_i p_j \int_0^1 \left[W'_{ij}(u_j(t,\xi) - u_i(t,1^-)) - W'_{ij}(u_j(t,\xi) - u_i(t,0^+)) \right] \mathrm{d}\xi$$

$$\le \sum_{j=1}^{n} m_i p_j \kappa_{ij} \int_0^1 \left[(u_j(t,\xi) - u_i(t,1^-)) - (u_j(t,\xi) - u_i(t,0^+)) \right] \mathrm{d}\xi = -m_i S_i (u_i(t,1^-) - u_i(t,0^+)),$$

the second-to-last step being a consequence of κ_{ij}-convexity (W5). Applying Gronwall's lemma completes the proof. □

In the regime where Proposition III.21 is applicable for all $i \in \{1, \dots, n\}$, system (III.1.1) behaves asymptotically as $t \to \infty$ like the particle system (III.2.9) in the case of only one (heavy) particle for each component ($N_i = 1$ for all i). Obviously, by Remark III.9, the condition $S_i > 0$ above is met in the scenario with uniformly geodesically convex energy. This enables us to improve the convergence result from Section III.2.2 in one spatial dimension for compactly supported initial data:

Proposition III.22 (The uniformly convex case in one spatial dimension). *Assume that the criterion for geodesic convexity* (III.2.5) *yields $\lambda_0 > 0$ and suppose that μ^0 has compact support. Then, for each $i \in \{1, \dots, n\}$,*

$$\lim_{t\to\infty} \mathbf{W}_\infty(\mu_i(t), \mu_i^\infty) = 0.$$

In view of Corollary III.12, we obtain convergence w.r.t. the stronger topology of the L^∞-Wasserstein distance, but lose the exponential rate of convergence.

Proof. Fix $i \in \{1, \dots, n\}$. Since $\lambda_0 > 0$ and $\operatorname{supp}\mu_i^0$ is compact, we know from Corollary III.12, Theorem III.17 and Proposition III.21 that $\|u_i(t,\cdot) - x^\infty\|_{L^2([0,1])} \to 0$ as $t \to \infty$, $\|u_i(t,\cdot)\|_{L^\infty([0,1])} \le K$ for all $t \ge 0$ and $u_i(t,1^-) - u_i(t,0^+) \to 0$ as $t \to \infty$. Obviously, if $\lim_{t\to\infty} u_i(t,1^-) = x^\infty$ holds, the desired result follows immediately from

$$\|u_i(t,\cdot) - x^\infty\|_{L^\infty([0,1])} = \max(|u_i(t,1^-) - x^\infty|, |u_i(t,0^+) - x^\infty|),$$

since then also $\lim_{t\to\infty} u_i(t,0^+) = x^\infty$ holds. So, assume that $u_i(t,1^-)$ does *not* converge to x^∞ as $t \to \infty$. Then, there exists $\varepsilon > 0$ and a sequence $t_k \to \infty$ such that

$$|u_i(t_k, 1^-) - x^\infty| \ge \varepsilon \qquad \forall k \in \mathbb{N}. \tag{III.3.17}$$

Thanks to the observations above, there exist a subsequence $(t_{k_l})_{l\in\mathbb{N}}$ and $\omega \in \mathbb{R}$ such that

$$\lim_{l\to\infty}(u_i(t_{k_l}, 1^-) - x^\infty) = \omega, \text{ and } \lim_{l\to\infty} u_i(t_{k_l}, z) = x^\infty \text{ for a.e. } z \in (0,1).$$

Immediately, it follows that $\lim_{l\to\infty} u_i(t_{k_l}, 0^+) = x^\infty + \omega$ and consequently $\omega = 0$ by monotonicity of $\lim_{l\to\infty} u_i(t_{k_l}, \cdot)$. But $\omega = 0$ is a contradiction to (III.3.17). □

CHAPTER III.4

Steady states and stability

In this chapter, we study the (non-)linear stability of stationary states of system (III.1.1) in one spatial dimension $d = 1$. Specifically, we identify the class of *discrete distributions* as candidates for stable steady states. This is a reasonable focus since non-discrete steady states might not exist (see Proposition III.23) or are not linearly asymptotically stable (see Theorem III.25) if the potential W is sufficiently regular.

One can narrow down the set of possible steady states to the class of discrete distributions if W is analytic:

Proposition III.23 (Analyticity implies singularity of steady states). *Let $i \in \{1, \dots, n\}$ be fixed, but arbitrary, and suppose that the following assumptions are satisfied:*

(i) For all $j \in \{1, \dots, n\}$, the map W_{ij} is analytic.

(ii) There exist $R > 0$ and $S \in \{-1, 1\}$ such that for almost every $z > R$:

$$\operatorname{sgn}\left(\sum_{j=1}^{n} W'_{ij}(z)\right) = S, \quad \textit{and} \quad \operatorname{sgn}\left(W'_{ij}(z)\right) \in \{0, S\} \quad \textit{for all } j \in \{1, \dots, n\}.$$

Let $\overline{\mu} \in \mathscr{P}$ be a compactly supported steady state of (III.1.1). *Then, the i^{th} component $\overline{\mu}_i$ is a finite linear combination of Dirac measures.*

Remark III.24. *Assumption (ii) in Proposition III.23 is e.g. satisfied if W is confining, cf. Definition III.15, and the C_{ij} from there satisfy $\sum_{j=1}^{n} C_{ij} > 0$ and $C_{ij} \geq 0$ for all j.*

Proof. For all $x \in \operatorname{supp} \overline{\mu}_i$:

$$\sum_{j=1}^{n} \int_{\mathbb{R}} W'_{ij}(x - y) \, \mathrm{d}\overline{\mu}_j(y) = 0. \tag{III.4.1}$$

We proceed similarly to [158, Prop. 2] and assume that $\overline{\mu}_i$ is not a finite linear combination of Dirac measures. Then, $\operatorname{supp} \overline{\mu}_i$ possesses an accumulation point since it is compact and contains an infinite set of points. Since W_{ij} is analytic for all $j \in \{1, \dots, n\}$, the left-hand side in (III.4.1) is analytic in x. By the principle of permanence for analytic functions, (III.4.1) then holds on the *whole space* $\mathbb{R}$. Let $C > 0$ with $\operatorname{supp} \overline{\mu}_k \subset [-C, C]$ for all k and fix some $\widetilde{x} > C + R$. By the linear transformation $t(y) := \widetilde{x} - y$, we obtain from (III.4.1):

$$0 = \int_{[\widetilde{x}-C, \widetilde{x}+C]} \sum_{j=1}^{n} W'_{ij}(z) \, \mathrm{d}(t_{\#}\overline{\mu}_j)(z).$$

Since $\operatorname{supp}(t_{\#}\overline{\mu}_k) \subset [\widetilde{x} - C, \widetilde{x} + C] \subset (R, \infty)$ for all k by construction, and due to assumption (ii), this is a contradiction. Hence, the assertion is proved. □

III.4.1. Linear stability

In most cases, Proposition III.23 is not applicable since analyticity of W is a very strong restriction. If W is less regular, non-discrete steady states might exist — but their linear asymptotic stability (in a suitable space) can be excluded:

THEOREM III.25 (Non-discrete steady states are not linearly asymptotically stable). *Let $W \in C^2(\mathbb{R};\mathbb{R}^{n\times n})$ and let $\overline{\mu}$ be a steady state of* (III.1.1) *for which there exists an index $i \in \{1,\dots,n\}$ such that* $\operatorname{supp}\overline{\mu}_i$ *possesses an accumulation point. Assume that $K \in (0,\infty]$ is such that* $\operatorname{supp}\overline{\mu}_j \subset \overline{(-K,K)}$ *for all $j \in \{1,\dots,n\}$ and that the entries in W'' are α-Hölder continuous and bounded for some $\alpha \in [0,1)$ on the interval $\overline{(-2K,2K)}$. Then, $\overline{\mu}$ is not linearly asymptotically stable in the following sense:*

The linear operator $A : L^1_0([0,1];\mathbb{R}^n) \to L^1_0([0,1];\mathbb{R}^n)$, defined on the Banach space

$$L^1_0([0,1];\mathbb{R}^n) := \left\{ v \in L^1([0,1];\mathbb{R}^n) :\ 0 = \sum_{j=1}^n \int_0^1 \frac{p_j}{m_j} v_j(z)\,\mathrm{d}z \right\} \subset L^1([0,1];\mathbb{R}^n),$$

via

$$(Av)_i(z) := \sum_{j=1}^n m_i p_j \int_0^1 W''_{ij}(\overline{u}_j(\xi) - \overline{u}_i(z))(v_j(\xi) - v_i(z))\,\mathrm{d}\xi \quad \text{for } z \in (0,1), \tag{III.4.2}$$

and all $i \in \{1,\dots,n\}$, where $\overline{u} = (\overline{u}_1,\dots,\overline{u}_n)$ is the associated vector of inverse distribution functions for $\overline{\mu}$, is continuous and 0 *belongs to the spectrum of A.*

REMARK III.26 (Formal linearization). *We give a motivation for the particular choice of A above. A perturbation $u_i = \overline{u}_i + sv_i$ for some $v : [0,1] \to \mathbb{R}^n$ and small $|s|$ is* admissible *if u_i is increasing and càdlàg and condition* (III.3.2) *on the center of mass holds, which transforms into*

$$\sum_{j=1}^n \int_0^1 \frac{p_j}{m_j} v_j(z)\,\mathrm{d}z = 0.$$

Moreover, the linearization of equation (III.3.1) *at $\overline{u}$ (evaluated at $z \in (0,1)$) is given by*

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \sum_{j=1}^n m_i p_j \int_0^1 W'_{ij}(\overline{u}_j(\xi) + sv_j(\xi) - \overline{u}_i(z) - sv_i(z))\,\mathrm{d}\xi \\ &= \sum_{j=1}^n m_i p_j \int_0^1 W''_{ij}(\overline{u}_j(\xi) - \overline{u}_i(z))(v_j(\xi) - v_i(z))\,\mathrm{d}\xi = (Av)_i(z),\end{aligned}$$

for each $i \in \{1,\dots,n\}$. In the following proof of Theorem III.25, it is convenient to use the topology induced by the norm in $L^1([0,1];\mathbb{R}^n)$.

PROOF. It is straightforward to prove that $L^1_0([0,1];\mathbb{R}^n)$ is a closed subspace of the Banach space $L^1([0,1];\mathbb{R}^n)$, thus also complete. We first show that A is well-defined, i.e. $Av \in L^1_0([0,1];\mathbb{R}^n)$ for each $v \in L^1_0([0,1];\mathbb{R}^n)$. We calculate:

$$\sum_{i=1}^n \int_0^1 \frac{p_i}{m_i}(Av)_i(z)\,\mathrm{d}z = \sum_{i=1}^n \sum_{j=1}^n p_i p_j \int_0^1 \int_0^1 W''_{ij}(\overline{u}_j(\xi) - \overline{u}_i(z))(v_j(\xi) - v_i(z))\,\mathrm{d}\xi\,\mathrm{d}z.$$

Using the symmetry of W'' and the fact that the W''_{ij} are even maps, we obtain by interchanging the names for i and j:

$$\sum_{i=1}^n \sum_{j=1}^n p_i p_j \int_0^1 \int_0^1 W''_{ij}(\overline{u}_j(\xi) - \overline{u}_i(z))(v_j(\xi) - v_i(z))\,\mathrm{d}\xi\,\mathrm{d}z$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \int_0^1 \int_0^1 W_{ij}''(\overline{u}_j(z) - \overline{u}_i(\xi))(v_i(\xi) - v_j(z))\, \mathrm{d}\xi\, \mathrm{d}z.$$

Interchanging the variable names of ξ and z, one has with Fubini's theorem that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \int_0^1 \int_0^1 W_{ij}''(\overline{u}_j(z) - \overline{u}_i(\xi))(v_i(\xi) - v_j(z))\, \mathrm{d}\xi\, \mathrm{d}z$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \int_0^1 \int_0^1 W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))(v_i(z) - v_j(\xi))\, \mathrm{d}\xi\, \mathrm{d}z = -\sum_{i=1}^{n} \int_0^1 \frac{p_i}{m_i}(Av)_i(z)\, \mathrm{d}z.$$

Hence,

$$\sum_{i=1}^{n} \int_0^1 \frac{p_i}{m_i}(Av)_i(z)\, \mathrm{d}z = 0,$$

and so $Av \in L^1_0([0,1];\mathbb{R}^n)$. It remains to verify the continuity of A in $L^1([0,1];\mathbb{R}^n)$. For each $v \in L^1_0([0,1];\mathbb{R}^n)$, one gets:

$$\|Av\|_{L^1} \le \sum_{i=1}^{n} \sum_{j=1}^{n} m_i p_j \int_0^1 \int_0^1 |W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))||v_j(\xi) - v_i(z)|\, \mathrm{d}\xi\, \mathrm{d}z.$$

Since $|\overline{u}_j(\xi) - \overline{u}_i(z)| \le 2K$ by definition of K, we are able to use the boundedness of W_{ij}'' and the triangle inequality to estimate further:

$$\|Av\|_{L^1} \le C \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\|v_j\|_{L^1} + \|v_i\|_{L^1}\right) = 2nC\|v\|_{L^1},$$

for some constant $C > 0$, showing the (Lipschitz) continuity of A.

For the proof of $0 \in \operatorname{spec}(A)$, we employ a similar strategy as in [158] and construct a family of *approximate eigenvectors* of A for 0, that is, $(v^\varepsilon)_{\varepsilon>0}$ in $L^1_0([0,1];\mathbb{R}^n)$ with

$$\inf_{\varepsilon>0} \frac{\|Av^\varepsilon\|_{L^1}}{\|v^\varepsilon\|_{L^1}} = 0. \tag{III.4.3}$$

Let $\overline{x} \in \overline{(-K,K)}$ be the accumulation point of $\operatorname{supp} \overline{\mu}_i$. There exists a monotone sequence $(x_k)_{k\in\mathbb{N}}$ in $\operatorname{supp} \overline{\mu}_i$ such that either $x_k \downarrow \overline{x}$ or $x_k \uparrow \overline{x}$ as $k \to \infty$. By definition of the support, one therefore either has

$$\int_{\overline{x}}^{\overline{x}+\varepsilon} \mathrm{d}\overline{\mu}_i > 0 \quad \text{or} \quad \int_{\overline{x}-\varepsilon}^{\overline{x}} \mathrm{d}\overline{\mu}_i > 0, \quad \text{for all } \varepsilon > 0.$$

We define, for each $\varepsilon > 0$, an interval $Z_\varepsilon := [z_0^\varepsilon, z_1^\varepsilon) \subset [0,1)$ as follows: In the first case above, set

$$z_0^\varepsilon := \inf\{z \in [0,1] :\ \overline{u}_i(z) > \overline{x}\},$$
$$z_1^\varepsilon := \sup\{z \in [0,1] :\ \overline{u}_i(z) < \overline{x} + \varepsilon\};$$

and in the second case, set

$$z_0^\varepsilon := \inf\{z \in [0,1] :\ \overline{u}_i(z) > \overline{x} - \varepsilon\},$$
$$z_1^\varepsilon := \sup\{z \in [0,1] :\ \overline{u}_i(z) < \overline{x}\}.$$

Note that in both cases, one has $|\overline{u}_i(z) - \overline{x}| \le \varepsilon$ for all $z \in Z_\varepsilon$. We now introduce a suitable approximation of $\overline{u}_i$: define u_i^ε by

$$u_i^\varepsilon(z) := \begin{cases} \overline{u}_i(z) & \text{if } z \notin Z_\varepsilon, \\ \frac{1}{z_1^\varepsilon - z_0^\varepsilon} \int_{Z_\varepsilon} \overline{u}_i(\zeta)\, \mathrm{d}\zeta & \text{if } z \in Z_\varepsilon, \end{cases}$$

replacing the values of $\overline{u}_i$ by its mean value on the interval Z_ε. Notice that u_i^ε is increasing and càdlàg. The associated displacement v_i^ε is given by

$$v_i^\varepsilon(z) := u_i^\varepsilon(z) - \overline{u}_i(z) = \begin{cases} 0 & \text{if } z \notin Z_\varepsilon, \\ \frac{1}{z_1^\varepsilon - z_0^\varepsilon} \int_{Z_\varepsilon} \overline{u}_i(\zeta)\,\mathrm{d}\zeta - \overline{u}_i(z) & \text{if } z \in Z_\varepsilon. \end{cases}$$

One easily sees that $\int_0^1 v_i^\varepsilon(z)\,\mathrm{d}z = 0$. For all other components $j \in \{1, \dots n\} \setminus \{i\}$, we put $v_j^\varepsilon \equiv 0$. Clearly, $v^\varepsilon := (v_1^\varepsilon, \dots, v_n^\varepsilon) \in L_0^1([0,1];\mathbb{R}^n)$. To prove the assertion (III.4.3), we deduce that there exists a constant $C > 0$ such that for all $\varepsilon > 0$, one has $\|Av^\varepsilon\|_{L^1} \le C\varepsilon^\alpha \|v^\varepsilon\|_{L^1}$.

Beforehand, we note that

$$\sum_{j=1}^n \int_0^1 p_j W_{ij}''(\overline{x} - \overline{u}_j(\xi))\,\mathrm{d}\xi = 0. \tag{III.4.4}$$

Indeed, since $\overline{\mu}$ is a steady state, one has for all $x \in \operatorname{supp} \overline{\mu}_i$ that

$$\sum_{j=1}^n \int_0^1 p_j W_{ij}'(x - \overline{u}_j(\xi))\,\mathrm{d}\xi = 0,$$

especially for $x = x_k$ and $x = \overline{x}$. By differentiability, (III.4.4) follows.

We can now estimate Av^ε in $L^1([0,1];\mathbb{R}^n)$:

$$\begin{aligned} \|Av^\varepsilon\|_{L^1} &= \sum_{k=1}^n \int_0^1 \left| \sum_{j=1}^n m_k p_j \int_0^1 W_{kj}''(\overline{u}_j(\xi) - \overline{u}_k(z))(v_j^\varepsilon(\xi) - v_k^\varepsilon(z))\,\mathrm{d}\xi \right| \mathrm{d}z \\ &\le \sum_{k=1}^n \int_0^1 \left| m_k p_i \int_0^1 W_{ki}''(\overline{u}_i(\xi) - \overline{u}_k(z)) v_i^\varepsilon(\xi)\,\mathrm{d}\xi \right| \mathrm{d}z + \int_0^1 \left| \sum_{j=1}^n m_i p_j \int_0^1 W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))\,\mathrm{d}\xi \right| |v_i^\varepsilon(z)|\,\mathrm{d}z, \end{aligned}$$

using the triangle inequality and $v_j^\varepsilon \equiv 0$ for $j \ne i$ in the last step. We consider both parts of the r.h.s. above separately. First, thanks to $\int_0^1 v_i^\varepsilon(\xi)\,\mathrm{d}\xi = 0$, one has

$$\begin{aligned} &\sum_{k=1}^n \int_0^1 \left| m_k p_i \int_0^1 W_{ki}''(\overline{u}_i(\xi) - \overline{u}_k(z)) v_i^\varepsilon(\xi)\,\mathrm{d}\xi \right| \mathrm{d}z \\ &= \sum_{k=1}^n \int_0^1 \left| m_k p_i \int_0^1 \left[W_{ki}''(\overline{u}_i(\xi) - \overline{u}_k(z)) - W_{ki}''(\overline{x} - \overline{u}_k(z)) \right] v_i^\varepsilon(\xi)\,\mathrm{d}\xi \right| \mathrm{d}z. \end{aligned}$$

Denoting by $H_{ij} \ge 0$ the Hölder constant of W_{ij}'' on $\overline{(-2K, 2K)}$, we estimate:

$$\begin{aligned} &\sum_{k=1}^n \int_0^1 \left| m_k p_i \int_0^1 \left[W_{ki}''(\overline{u}_i(\xi) - \overline{u}_k(z)) - W_{ki}''(\overline{x} - \overline{u}_k(z)) \right] v_i^\varepsilon(\xi)\,\mathrm{d}\xi \right| \mathrm{d}z \\ &\le \sum_{k=1}^n m_k p_i H_{ki} \int_0^1 \int_0^1 |\overline{x} - \overline{u}_i(\xi)|^\alpha |v_i^\varepsilon(\xi)|\,\mathrm{d}\xi\,\mathrm{d}z. \end{aligned}$$

Using that $|\overline{x} - \overline{u}_i(\xi)| \le \varepsilon$ for $\xi \in Z_\varepsilon$, one arrives at:

$$\sum_{k=1}^n m_k p_i H_{ki} \int_0^1 \int_0^1 |\overline{x} - \overline{u}_i(\xi)|^\alpha |v_i^\varepsilon(\xi)|\,\mathrm{d}\xi\,\mathrm{d}z \le p_i \sum_{k=1}^n m_k H_{ki} \varepsilon^\alpha \|v_i^\varepsilon\|_{L^1}.$$

The remaining term to be estimated can be rewritten using (III.4.4):

$$\int_0^1 \left| \sum_{j=1}^n m_i p_j \int_0^1 W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))\,\mathrm{d}\xi \right| |v_i^\varepsilon(z)|\,\mathrm{d}z$$

$$= m_i \int_0^1 \left| \sum_{j=1}^n p_j \int_0^1 \left[W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z)) - W_{ij}''(\overline{u}_j(\xi) - \overline{x}) \right] \, \mathrm{d}\xi \right| |v_i^\varepsilon(z)| \, \mathrm{d}z.$$

By the same arguments as above, we get:

$$m_i \int_0^1 \left| \sum_{j=1}^n p_j \int_0^1 \left[W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z)) - W_{ij}''(\overline{u}_j(\xi) - \overline{x}) \right] \, \mathrm{d}\xi \right| |v_i^\varepsilon(z)| \, \mathrm{d}z \leq m_i \sum_{j=1}^n H_{ij} p_j \varepsilon^\alpha \|v_i^\varepsilon\|_{L^1}.$$

All in all, we have proved that

$$\|Av^\varepsilon\|_{L^1} \leq \varepsilon^\alpha \|v_i^\varepsilon\|_{L^1} \sum_{j=1}^n H_{ij}(m_i p_j + m_j p_i),$$

and the claim follows since $\|v_i^\varepsilon\|_{L^1} = \|v^\varepsilon\|_{L^1}$. □

Thus, under reasonable conditions, non-discrete steady states are not asymptotically stable in the linear sense of Theorem III.25.

III.4.2. Local nonlinear stability

In this section, we investigate the complementary case to Section III.4.1. Specifically, we derive sufficient conditions under which discrete steady states are locally asymptotically stable in the *nonlinear* sense. Before we actually state our main result, we sketch the strategy.

In a nutshell, a strictly stable spectrum of the *linearization* A (see formula (III.4.2)) considered on a specific *subclass* of perturbations $u = \overline{u} + v$ (with $v \in L_0^1([0,1];\mathbb{R}^n)$ as in the previous section) implies the local *nonlinear* stability (in the topology induced by the $\mathbf{W}_\infty$ distance) in the *whole* state space. Those specific perturbations of the steady state $\overline{\mu}$ are divided into two classes: First, so-called *shifts*, where the positions of the Dirac measures in $\overline{\mu}$ are altered; second, so-called *reallocations*, where some Dirac measures are "fattened", keeping the average value conserved. A similar phenomenon has been observed for the scalar case $n = 1$ in [82, 83]. We shall assume $n > 1$ here.

Throughout this section, we consider a *discrete* steady state $\overline{\mu}$ of the form

$$\overline{\mu}_i = \sum_{k=1}^{N_i} p_i^k \delta_{x_i^k} \qquad (i = 1 \ldots, n),$$

with $N_i \in \mathbb{N}$ (without restriction ordered) *positions* $x_i^1 < x_i^2 < \ldots < x_i^{N_i}$ and *masses* $p_i^1, \ldots, p_i^{N_i}$ satisfying $\sum_{k=1}^{N_i} p_i^k = p_i$, for each $i \in \{1, \ldots, n\}$. Recall that the components are related via the center-of-mass condition

$$E = \sum_{i=1}^n \sum_{k=1}^{N_i} \frac{p_i^k x_i^k}{m_i}.$$

The inverse distribution function $\overline{u}_i : [0,1) \to \mathbb{R}$ corresponding to $\overline{\mu}_i$ is the step function

$$\overline{u}_i(z) = \sum_{k=1}^{N_i} x_i^k \mathbf{1}_{Z_i^k}(z), \quad \text{with} \quad Z_i^k := \left[\sum_{\ell=1}^{k-1} \frac{p_i^\ell}{p_i}, \sum_{\ell=1}^{k} \frac{p_i^\ell}{p_i} \right).$$

Notice that the length of Z_i^k is given by $|Z_i^k| = \frac{p_i^k}{p_i}$.

III.4.2.1. Shifts

Definition III.27 (Shift). *An element $v \in L^1_0([0,1];\mathbb{R}^n)$ is called* shift *if, for each $i \in \{1,\dots,n\}$, v_i is of the form*

$$v_i(z) = \sum_{k=1}^{N_i} v_i^k \mathbf{1}_{Z_i^k}(z),$$

with $v_i^1,\dots,v_i^{N_i} \in \mathbb{R}$ satisfying

$$\sum_{i=1}^{n}\sum_{k=1}^{N_i} \frac{p_i^k v_i^k}{m_i} = 0.$$

The space of all shifts is denoted by $\mathscr{S}$.

The space of shifts $\mathscr{S}$ has some useful properties:

- it is an N-dimensional subspace of $L^1_0([0,1];\mathbb{R}^n)$, with $N := \sum_{i=1}^n N_i$;
- it is positively invariant under the linearization operator A defined in (III.4.2), i.e. $A\mathscr{S} \subset \mathscr{S}$.

Indeed, if $v \in \mathscr{S}$, it is obvious that $(Av)_i$ is constant on each interval Z_i^k, for all i. Hence, $Av \in \mathscr{S}$ as well.

We now study the spectrum of the restricted operator $A|_{\mathscr{S}}$. Recall that, for $v \in \mathscr{S}$, we have

$$(A|_{\mathscr{S}} v)_i(z) = \sum_{j=1}^{n} m_i \sum_{k=1}^{N_j} p_j^k W''_{ij}(x_j^k - x_i^\ell)(v_j^k - v_i^\ell) \quad \text{for } z \in Z_i^\ell \quad (\ell \in \{1,\dots,N_i\}).$$

Introducing the new set of variables $w_i^\ell := \frac{p_i^\ell}{p_i} v_i^\ell$ for $\ell = 1,\dots,N_i$, $i = 1,\dots,n$, the relation above can be rewitten as

$$\frac{p_i^\ell}{p_i}(A|_{\mathscr{S}} v)_i(z) = -\left[\sum_{j=1}^{n}\sum_{k=1}^{N_j} m_i p_j^k W''_{ij}(x_j^k - x_i^\ell)\right] w_i^\ell + \sum_{j=1}^{n}\sum_{k=1}^{N_j}\left[m_i p_j W''_{ij}(x_j^k - x_i^\ell)\frac{p_i^\ell}{p_i}\right] w_j^k,$$

for $z \in Z_i^\ell$ ($\ell = 1,\dots,N_i$, $i = 1,\dots,n$).

With the new indices $r = \sum_{a=1}^{i-1} N_a + \ell$ and $s = \sum_{a=1}^{j-1} N_a + k$, running from 1 to N, we obtain the following matrix-vector form of the right-hand side above:

$$\left(-\left[\sum_{j=1}^{n}\sum_{k=1}^{N_j} m_i p_j^k W''_{ij}(x_j^k - x_i^\ell)\right] w_i^\ell + \sum_{j=1}^{n}\sum_{k=1}^{N_j}\left[m_i p_j W''_{ij}(x_j^k - x_i^\ell)\frac{p_i^\ell}{p_i}\right] w_j^k\right)_{\ell=1,\dots,N_i,\ i=1,\dots,n} = -Mw,$$

where $M = D - R$, with a diagonal matrix $D \in \mathbb{R}^{N\times N}$ defined by

$$D_{rr} = m_i \sum_{j=1}^{n}\sum_{k=1}^{N_j} W''_{ij}(x_j^k - x_i^\ell) p_j^k \qquad (r = 1,\dots,N),$$

and $R \in \mathbb{R}^{N\times N}$ with entries

$$R_{rs} = m_i p_j W''_{ij}(x_j^k - x_i^\ell)\frac{p_i^\ell}{p_i} \qquad (r,s = 1,\dots,N).$$

Due to the center-of-mass condition

$$0 = \sum_{i=1}^{n}\sum_{\ell=1}^{N_i} \frac{p_i}{m_i} w_i^\ell, \tag{III.4.5}$$

one can eliminate one variable w_i^ℓ and eventually ends up with the linear map $\mathbb{R}^{N-1} \ni \widetilde{w} \mapsto -M|_U\widetilde{w}$, for the restriction $M|_U \in \mathbb{R}^{(N-1)\times(N-1)}$ of $M \in \mathbb{R}^{N\times N}$ to the $(N-1)$-dimensional subspace $U := \{w \in \mathbb{R}^n : \text{(III.4.5) holds}\}$. We arrive at the following sufficient condition for linear asymptotic stability of $\overline{\mu}$ with respect to shifts (i.e., constraining the spectrum of $A|_{\mathscr{S}}$ to the left half-plane in $\mathbb{C}$):

$$(\mathrm{SS})_\delta \qquad \exists \delta > 0 : \ \mathrm{spec}(M|_U) \subset \{\zeta \in \mathbb{C} : \ \mathrm{Re}(\zeta) \geq \delta\} =: H_\delta^+.$$

III.4.2.2. Reallocations

Definition III.28 (Reallocation). *An element $v \in L_0^1([0,1];\mathbb{R}^n)$ is called* reallocation *if for all $k \in \{1,\dots,N_i\}$ and all $i \in \{1,\dots,n\}$, one has*

$$\int_{Z_i^k} v_i(z)\,\mathrm{d}z = 0.$$

The space of all reallocations is denoted by $\mathscr{R}$.

The following properties hold for the space $\mathscr{R}$:

- it is an (infinite-dimensional) closed subspace of $L_0^1([0,1];\mathbb{R}^n)$;
- it is positively invariant under the linearization operator A defined in (III.4.2), i.e. $A\mathscr{R} \subset \mathscr{R}$.

Notice that for $v \in \mathscr{R}$, one has

$$(Av)_i(z) = -\sum_{j=1}^n m_i \sum_{k=1}^{N_j} W_{ij}''(x_j^k - \overline{u}_i(z))p_j^k v_i(z) + \sum_{j=1}^n m_i p_j \sum_{k=1}^{N_j} W_{ij}''(x_j^k - \overline{u}_i(z)) \int_{Z_j^k} v_j(\xi)\,\mathrm{d}\xi,$$

and the second term on the right-hand side vanishes by definition of $\mathscr{R}$. Hence, $Av \in \mathscr{R}$ because of

$$\int_{Z_i^\ell} (Av)_i(z)\,\mathrm{d}z = -\sum_{j=1}^n m_i \sum_{k=1}^{N_j} W_{ij}''(x_j^k - x_i^\ell)p_j^k \int_{Z_i^\ell} v_i(z)\,\mathrm{d}z = 0.$$

Concerning the spectrum of the restricted operator $A|_{\mathscr{R}}$, we observe that

$$(A|_{\mathscr{R}}v)_i(z) = -\left[\sum_{j=1}^n m_i \sum_{k=1}^{N_j} W_{ij}''(x_j^k - x_i^\ell)p_j^k\right] v_i(z) \quad \text{for } z \in Z_i^\ell.$$

We thus obtain the following sufficient condition on the linear asymptotic stability of $\overline{\mu}$ with respect to reallocations (i.e., constraining the spectrum of $A|_{\mathscr{R}}$ to the left half-plane in $\mathbb{C}$):

$$(\mathrm{SR})_\delta \qquad \exists \delta > 0 : \ \forall i \in \{1,\dots,n\}\ \forall \ell \in \{1,\dots,N_i\} : \ \sum_{j=1}^n m_i \sum_{k=1}^{N_j} W_{ij}''(x_j^k - x_i^\ell)p_j^k \geq \delta.$$

Note that $(\mathrm{SR})_\delta$ coincides with $\mathrm{spec}(D) \in H_\delta^+$, where D is the diagonal matrix from Section III.4.2.1.

III.4.2.3. A result on nonlinear stability

We now are in position to formulate our main theorem about the local nonlinear stability of discrete steady states:

Theorem III.29 (Local nonlinear asymptotic stability of discrete steady states). *Let a discrete steady state $\overline{\mu} \in \mathscr{P}$ of* (III.1.1) *of the form*

$$\overline{\mu}_i = \sum_{k=1}^{N_i} p_i^k \delta_{x_i^k} \qquad (i = 1\dots,n)$$

be given. Assume that $W \in C^2(\mathbb{R};\mathbb{R}^{n\times n})$ and that all W''_{ij} are α-Hölder continuous for some $\alpha \in (0,1]$. Moreover, assume that the conditions $(\mathrm{SS})_\delta$ and $(\mathrm{SR})_\delta$ are satisfied for some $\delta > 0$. Then, $\overline{\mu}$ is locally nonlinearly asymptotically stable in the following sense:

For each $\widetilde{\delta} \in (0,\delta)$, there exist $\varepsilon > 0$ and $C > 0$ such that for all $\mu^0 \in \mathscr{P}$ with $\mathbf{W}_\infty(\mu_i^0, \overline{\mu}_i) \le \varepsilon$ for each $i \in \{1,\dots,n\}$, one has

$$\mathbf{W}_\infty(\mu_i(t), \overline{\mu}_i) \le C\exp(-\widetilde{\delta}t) \qquad \text{for all } t \ge 0 \text{ and each } i \in \{1,\dots,n\}. \tag{III.4.6}$$

In advance of the proof, we need to prove

Lemma III.30 (A version of Gronwall's lemma). *Assume that $f : [0,\infty) \to [0,\infty)$ is continuous and satisfies the following estimate for some $C > 0$, $a > 0$, $k \in \mathbb{N}_0$, $\alpha \in (0,1]$ and all $t \ge 0$:*

$$f(t) \le Cf(0)(1+t^k)e^{-at} + C\int_0^t (1+(t-s)^k)e^{-a(t-s)} f(s)^{1+\alpha}\,\mathrm{d}s. \tag{III.4.7}$$

Then, for each $\eta \in (0,1)$, there exist $\delta > 0$ and $B > 0$ such that, provided $f(0) \le \delta$, one has

$$f(t) \le Be^{-(1-\eta)at} \qquad \forall t \ge 0. \tag{III.4.8}$$

Proof. Let $\eta \in (0,1)$ fixed, but arbitrary. Rewriting (III.4.7) yields

$$f(t) \le f(0)h(t)e^{-(1-\eta)at} + C\int_0^t h(t-s)e^{-(1-\eta)a(t-s)} f(s)^{1+\alpha}\,\mathrm{d}s,$$

where for $\sigma \ge 0$:

$$h(\sigma) := C(1+\sigma^k)e^{-\eta a\sigma} \le \begin{cases} C\left(1+\left(\frac{k}{\eta a e}\right)^k\right) & \text{if } k \in \mathbb{N}, \\ 2C & \text{if } k = 0 \end{cases}.$$

Consequently,

$$f(t) \le C_1 f(0)e^{-(1-\eta)at} + C_1\int_0^t e^{-(1-\eta)a(t-s)} f(s)^{\alpha+1}\,\mathrm{d}s, \tag{III.4.9}$$

for some constant $C_1 > 0$. Introduce the continuous function $g : [0,\infty) \to [0,\infty)$ by $g(t) := f(t)e^{(1-\eta)at}$. Then, (III.4.9) transforms into (note that $g(0) = f(0)$):

$$g(t) \le C_1 f(0) + C_1\int_0^t g(s)^{\alpha+1} e^{\alpha a(1-\eta)s}\,\mathrm{d}s. \tag{III.4.10}$$

Define

$$B := \left(\frac{\alpha a(1-\eta)}{2C_1}\right) \quad \text{and} \quad \delta := B\min\left(1, \frac{1}{2C_1}\right).$$

It remains to show that the following implication holds:

$$f(0) \le \delta \qquad \Longrightarrow \qquad g(t) \le B \quad \forall t \ge 0.$$

Notice that by construction, one has $g(0) < B$. Assume that the above implication is false. Then, by continuity, there exists $t_0 > 0$ such that $g(t_0) = B$ and $g(s) < B$ for all $s \in [0,t_0)$. Then, (III.4.10) implies that

$$g(t_0) < C_1\delta + C_1 B^{\alpha+1}\int_0^{t_0} e^{-s\alpha a(1-\eta)}\,\mathrm{d}s \le C_1\delta + \frac{C_1 B^{\alpha+1}}{(1-\eta)\alpha a} = C_1\delta + \frac{B}{2} \le B,$$

which is a contradiction. □

Now, we prove Theorem III.29.

Proof. We define the displacement $v(t,z) = u(t,z) - \overline{u}(z)$ and derive the claimed estimate (III.4.6) for v in the space $L^\infty([0,1];\mathbb{R}^n)$. We first obtain — for each $i \in \{1,\dots,n\}$ — by insertion into the differential equation (III.3.1) for u:

$$\partial_t v_i(t,z) = \sum_{j=1}^{n} m_i p_j \int_0^1 W_{ij}'(\overline{u}_j(\xi) + v_j(t,\xi) - \overline{u}_i(z) - v_i(t,z))\,\mathrm{d}\xi.$$

Recalling that $\overline{u}$ is a steady state and using Taylor's theorem, one gets

$$\begin{aligned}\partial_t v_i(t,z) &= \sum_{j=1}^{n} m_i p_j \int_0^1 \Big[W_{ij}'(\overline{u}_j(\xi) + v_j(t,\xi) - \overline{u}_i(z) - v_i(t,z)) - W_{ij}'(\overline{u}_j(\xi) - \overline{u}_i(z))\Big]\,\mathrm{d}\xi \\ &= \sum_{j=1}^{n} m_i p_j \int_0^1 W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z) + \vartheta_{ij}(t,\xi,z))(v_j(t,\xi) - v_i(t,z))\,\mathrm{d}\xi,\end{aligned}$$

for some intermediate value $\vartheta_{ij}(t,\xi,z)$ between 0 and $v_j(t,\xi) - v_i(t,z)$. Addition and subtraction of a suitable linear term yields

$$\begin{aligned}\partial_t v_i(t,z) = &\sum_{j=1}^{n} m_i p_j \int_0^1 W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))(v_j(t,\xi) - v_i(t,z))\,\mathrm{d}\xi \\ &+ \sum_{j=1}^{n} m_i p_j \int_0^1 \Big[W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z) + \vartheta_{ij}(t,\xi,z)) - W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))\Big]\,(v_j(t,\xi) - v_i(t,z))\,\mathrm{d}\xi.\end{aligned}$$

Denote, for later reference,

$$\mathcal{N}_i(t,z,v(t,z)) := \sum_{j=1}^{n} m_i p_j \int_0^1 \Big[W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z) + \vartheta_{ij}(t,\xi,z)) - W_{ij}''(\overline{u}_j(\xi) - \overline{u}_i(z))\Big]\,(v_j(t,\xi) - v_i(t,z))\,\mathrm{d}\xi.$$

The map $\mathcal{N}_i$ is of the order $O(\|v(t,\cdot)\|_{L^\infty}^{1+\alpha})$ for fixed $t > 0$ since due to $|\vartheta_{ij}(t,\xi,z)| \le |v_j(t,\xi) - v_i(t,z)|$ and α-Hölder continuity of W_{ij}'' (with Hölder constant $H_{ij} > 0$), we have

$$|\mathcal{N}_i(t,z,v(t,z))| \le \sum_{j=1}^{n} m_i p_j H_{ij} \int_0^1 |v_j(t,\xi) - v_i(t,z)|^{1+\alpha}\,\mathrm{d}\xi \le \sum_{j=1}^{n} m_i p_j H_{ij} 2^{\alpha} \left(\|v_j(t,\cdot)\|_{L^\infty}^{\alpha+1} + \|v_i(t,\cdot)\|_{L^\infty}^{\alpha+1}\right).$$

Define for $\ell = 1,\dots,N_i$, $i = 1,\dots,n$:

$$v_i^\ell(t,z) := v_i(t,z)\mathbf{1}_{Z_i^l}(z) \quad \text{and} \quad w_i^\ell(t) := \int_{Z_i^\ell} v_i^\ell(t,z)\,\mathrm{d}z,$$

recalling that the Z_i^ℓ are those intervals where $\overline{u}_i$ is constant. Using the notation and definitions from Section III.4.2.1, we obtain

$$\partial_t v_i^\ell(t,z) = -D_{rr} v_i^\ell(t,z) + m_i \sum_{j=1}^{n} \sum_{k=1}^{N_j} p_j W_{ij}''(x_j^k - x_i^\ell) w_j^k + \mathcal{N}_i(t,z,v(t,z)) \quad \text{for } z \in Z_i^\ell, \tag{III.4.11}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} w_i^\ell(t) = -D_{rr} w_i^\ell + \sum_{s=1}^{N} R_{rs} w_j^k(t) + \int_{Z_i^\ell} \mathcal{N}_i(t,z,v(t,z))\,\mathrm{d}z. \tag{III.4.12}$$

The system (III.4.12) for w can be rewritten in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} w_r(t) = -\sum_{s=1}^{N} M_{rs} w_s(t) + \int_{Z_i^\ell} \mathcal{N}_i(t,z,v(t,z))\,\mathrm{d}z \qquad (r = 1,\dots,N).$$

As in Section III.4.2.1, using that the weighted center of mass is conserved (cf. (III.3.2)), one can eliminate one of the variables w_r and consider the system on the subspace U introduced above:

$$\partial_t \widetilde{w}(t) = -M|_U \widetilde{w}(t) + \widetilde{\mathcal{N}}(t, v(t,\cdot)), \tag{III.4.13}$$

for $\widetilde{w}(t) \in \mathbb{R}^{N-1}$ and a vector-valued nonlinearity $\widetilde{\mathcal{N}}(t, v(t,\cdot)) \in \mathbb{R}^{N-1}$. Writing $M' := M|_U$ for brevity, the solution to (III.4.13) reads

$$\widetilde{w}(t) = \exp(-M't)\widetilde{w}(0) + \int_0^t \exp(-M'(t-s))\widetilde{\mathcal{N}}(s, v(s,\cdot))\,\mathrm{d}s. \tag{III.4.14}$$

In order to simplify the matrix exponential in (III.4.14), we decompose M' into its Jordan normal form over the field $\mathbb{C}$: there exist an invertible matrix $T \in \mathbb{C}^{(N-1)\times(N-1)}$, a nilpotent matrix $B \in \mathbb{C}^{(N-1)\times(N-1)}$ (i.e., $B^{N-1} = 0$) and a diagonal matrix $V \in \mathbb{C}^{(N-1)\times(N-1)}$ with $\operatorname{spec}(V) = \operatorname{spec}(M')$, such that $M' = T(V+B)T^{-1}$ and $VB = BV$ hold. One easily obtains that

$$\exp(-M't) = T\exp(-Vt)\exp(-Bt)T^{-1} = \sum_{b=0}^{N-2}(-1)^b t^b T\exp(-Vt)B^b T^{-1}.$$

Since $\operatorname{spec}(M') \subset H_\delta^+$ thanks to condition $(\mathrm{SS})_\delta$, one consequently has (with the maximum norm $\|\cdot\|_\infty$ for vectors and matrices in $\mathbb{C}^{N-1}$ and $\mathbb{C}^{(N-1)\times(N-1)}$, respectively):

$$\|\exp(-M'\sigma)\|_\infty \le \|T\|_\infty e^{-\delta\sigma}\sum_{b=0}^{N-2}\sigma^b\|B\|_\infty^b\|T^{-1}\|_\infty \le C(1+\sigma^{N-2})e^{-\delta\sigma},$$

for some constant $C > 0$ and all $\sigma \ge 0$. Thus, we obtain the following upper bound for $\widetilde{w}(t)$ from (III.4.14):

$$\|\widetilde{w}(t)\|_\infty \le C\|\widetilde{w}(0)\|_\infty(1+t^{N-2})e^{-\delta t} + \int_0^t C(1+(t-s)^{N-2})e^{-\delta(t-s)}\|\widetilde{\mathcal{N}}(s, v(s,\cdot))\|_\infty\,\mathrm{d}s.$$

We use this information to perform an estimate for the right-hand side of equation (III.4.11). Note that due to condition $(\mathrm{SR})_\delta$, one has $D_{rr} \ge \delta$. For another constant $C' > 0$ and abbreviating $\|\mathcal{N}(t, v(t,\cdot))\|_\infty := \|\mathcal{N}_i(t,\cdot,v(t,\cdot))\|_{L^\infty([0,1];\mathbb{R}^n)}$, we get:

$$\begin{aligned}\partial_t v_i^\ell(t,z) &\le -\delta v_i^\ell(t,z) + C'\|\widetilde{w}(t)\|_\infty + \mathcal{N}_i(t,z,v(t,z))\\ &\le -\delta v_i^\ell(t,z) + C'C\|\widetilde{w}(0)\|_\infty(1+t^{N_2})e^{-\delta t} + C'C\int_0^t(1+(t-s)^{N-2})e^{-\delta(t-s)}\|\widetilde{\mathcal{N}}(s,v(s,\cdot))\|_\infty\,\mathrm{d}s\\ &\qquad + \|\mathcal{N}(t,v(t,\cdot))\|_\infty.\end{aligned}$$

We apply Gronwall's lemma (in its classical form) to this differential estimate to obtain:

$$\begin{aligned}v_i^\ell(t,z) \le{}& v_i^\ell(0,z)e^{-\delta t} + \int_0^t C'C\|\widetilde{w}(0)\|_\infty e^{-\delta(t-s)}(1+s^{N-2})e^{-\delta s}\,\mathrm{d}s\\ &+ C'C\int_0^t e^{-\delta(t-s)}\int_0^s(1+(s-\sigma)^{N-2})e^{-\delta(s-\sigma)}\|\widetilde{\mathcal{N}}(\sigma,v(\sigma,\cdot))\|_\infty\,\mathrm{d}\sigma\,\mathrm{d}s\\ &+ \int_0^t e^{-\delta(t-s)}\|\mathcal{N}(s,v(s,\cdot))\|_\infty\,\mathrm{d}s.\end{aligned}$$

With the notation $\|v(t,\cdot)\|_\infty := \max_{r=1,\ldots,N}\|v_i^\ell(t,\cdot)\|_{L^\infty([0,1])}$, we obviously have for some constant $C_0 > 0$ that

$$\begin{aligned}\|\widetilde{w}(0)\|_\infty &\le C_0\|v(0,\cdot)\|_\infty,\\ \|\mathcal{N}(s,v(s,\cdot))\|_\infty &\le C_0\|v(s,\cdot)\|_\infty^{1+\alpha},\\ \|\widetilde{\mathcal{N}}(s,v(s,\cdot))\|_\infty &\le C_0\|v(s,\cdot)\|_\infty^{1+\alpha},\end{aligned}$$

and consequently, for another constant $\widetilde{C} > 0$:

$$
\begin{aligned}
\|v(t,\cdot)\|_\infty \le \|v(0,\cdot)\|_\infty e^{-\delta t} &+ \widetilde{C}\|v(0,\cdot)\|_\infty \int_0^t e^{-\delta(t-s)}(1+s^{N-2})e^{-\delta s}\,\mathrm{d}s \\
&+ \widetilde{C}\int_0^t\int_0^s e^{-\delta(t-s)}e^{-\delta(s-\sigma)}(1+(s-\sigma)^{N-2})\|v(\sigma,\cdot)\|_\infty^{1+\alpha}\,\mathrm{d}\sigma\,\mathrm{d}s \\
&+ \widetilde{C}\int_0^t e^{-\delta(t-s)}\|v(s,\cdot)\|_\infty^{1+\alpha}\,\mathrm{d}s.
\end{aligned}
$$

The double integral on the right-hand side above can be rewritten using Fubini's theorem:

$$
\begin{aligned}
&\int_0^t\int_0^s e^{-\delta(t-s)}e^{-\delta(s-\sigma)}(1+(s-\sigma)^{N-2})\|v(\sigma,\cdot)\|_\infty^{1+\alpha}\,\mathrm{d}\sigma\,\mathrm{d}s \\
&= \int_0^t\int_\sigma^t e^{-\delta(t-s)}e^{-\delta(s-\sigma)}(1+(s-\sigma)^{N-2})\|v(\sigma,\cdot)\|_\infty^{1+\alpha}\,\mathrm{d}s\,\mathrm{d}\sigma \\
&= \int_0^t \|v(\sigma,\cdot)\|_\infty^{1+\alpha}e^{-\delta(t-\sigma)}\int_\sigma^t(1+(s-\sigma)^{N-2})\,\mathrm{d}s\,\mathrm{d}\sigma \\
&= \int_0^t \|v(\sigma,\cdot)\|_\infty^{1+\alpha}e^{-\delta(t-\sigma)}\left[(t-\sigma)+\frac{1}{N-1}(t-\sigma)^{N-1}\right]\,\mathrm{d}\sigma.
\end{aligned}
$$

All in all, we can find a suitably large constant $C>0$ such that for all $t\ge 0$:

$$
\|v(t,\cdot)\|_\infty \le C\|v(0,\cdot)\|_\infty(1+t^{N-1})e^{-\delta t} + C\int_0^t(1+(t-s)^{N-1})e^{-\delta(t-s)}\|v(s,\cdot)\|_\infty^{\alpha+1}\,\mathrm{d}s. \tag{III.4.15}
$$

Applying Lemma III.30 to estimate (III.4.15), one obtains that for each $\widetilde{\delta}\in(0,\delta)$, there exist $\varepsilon>0$ and $C>0$ such that $\|v(0,\cdot)\|_{L^\infty([0,1];\mathbb{R}^n)}\le\varepsilon$ implies

$$
\|v(t,\cdot)\|_{L^\infty([0,1];\mathbb{R}^n)} \le Ce^{-\widetilde{\delta}t} \qquad \forall t\ge 0.
$$

This concludes the proof of the asserted estimate (III.4.6). □

Part IV

Degenerate diffusion systems with nonlinear mobility

CHAPTER IV.1

Introduction to Part IV

Parts of the results presented in this part of the thesis have been published in a joint article with Daniel Matthes [190]. We are concerned with the variational structure of the following system of coupled nonlinear evolution equations in one spatial dimension:

$$\partial_t \mu(t,x) = \partial_x \left[\mathbf{M}(\mu(t,x)) \partial_x \mathcal{E}'(\mu(t,x)) \right] \quad \text{at } t > 0 \text{ and } x \in \mathbb{R}, \tag{IV.1.1}$$

for the n components $\mu_1, \dots, \mu_n$ of the sought-for function $\mu : [0,\infty) \times \mathbb{R} \to S$.

Systems of that kind arise e.g. in reaction-diffusion models for chemical agents as well as for semiconductor dynamics [134, 123, 140, 91], or for population dynamics [59, 108, 107, 36] with or without cross-diffusion.

We assume that μ attains values in a convex compact set $S \subset \mathbb{R}^n$ with nonempty interior $\mathrm{int}(S)$. Above, $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ is the *mobility matrix*, and $\mathcal{E}'$ is the first variation of the *driving entropy* or *free energy* functional $\mathcal{E}$ which is defined on $\mathscr{M}(\mathbb{R}; S)$, the space of measurable functions on $\mathbb{R}$ with values in S.

Formally, (IV.1.1) possesses a gradient flow structure: solutions $\mu(t,\cdot)$ are curves of steepest descent in the potential landscape of $\mathcal{E}$, with respect to the Riemannian structure induced on the "manifold" $\mathscr{M}(\mathbb{R}; S)$ by weighted H^{-1}-norms $\|\cdot\|_\mu$ on "tangent vectors" $\dot{\mu}$:

$$\|\dot{\mu}\|_\mu^2 = \int_{\mathbb{R}} \partial_x \Psi^{\mathrm{T}} \mathbf{M}(\mu) \partial_x \Psi \, \mathrm{d}x, \tag{IV.1.2}$$

where the auxiliary function $\Psi : \mathbb{R} \to \mathbb{R}^n$ solves the elliptic problem

$$\dot{\mu} + \partial_x \big(\mathbf{M}(\mu) \partial_x \Psi \big) = 0.$$

A cornerstone in the theory of optimal transportation is the Benamou-Brenier dynamical interpretation of the L^2-Wasserstein distance [11], see formula (I.2.18) from Section I.2.2. Dolbeault *et al.* [74] have used that interpretation to *define* a new class of transportation metrics $\mathbf{W_m}$, corresponding to *nonlinear* mobilities $\mathbf{m}$, recall Section I.2.2.3.

We extend the approach of [74, 124] to densities $\mu : \mathbb{R} \to S$ with values in a convex and compact set $S \subset \mathbb{R}^n$, and a mobility matrix $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ in place of $\mathbf{m}$. Our hypotheses on $\mathbf{M}$ are:

(M0) $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ is continuous, and is smooth on $\mathrm{int}(S)$.
(M1) $\mathbf{M}(z)$ is symmetric and positive definite for each $z \in \mathrm{int}(S)$.
(M2) $\mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta]$ is negative semidefinite for each $z \in \mathrm{int}(S)$ and $\zeta \in \mathbb{R}^n$.
(M3) $\mathbf{M}(z)\nu = 0$ if $z \in \partial S$ and ν is a normal vector to ∂S at z.

Conditions (M1)&(M2) are direct generalizations of positivity and concavity of the mobility function $\mathbf{m}$, and (M0) is a technical hypothesis. Condition (M3) is a natural requirement that is satisfied in all of our examples, but is not substantial for the proofs. Its intepretation is that the values of solutions to (IV.1.1) are confined to S.

Finally, we say that $\mathbf{M}$ is *induced by a function* $h \in C^2(\mathrm{int}(S))$, if

$$\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \quad \text{for all } z \in \mathrm{int}(S). \tag{IV.1.3}$$

This hypothesis allows one to formulate the multi-component heat equation $\partial_t \mu = \partial_{xx}\mu$ in the form (IV.1.1); with the functional $\mathcal{E}(\mu) = \int_{\mathbb{R}} h(\mu(x))\,\mathrm{d}x$.

Under conditions (M0)–(M3), we prove in Chapter IV.2 that the Benamou-Brenier formula with the norms from (IV.1.2) defines a (pseudo-)metric $\mathbf{W_M}$ on the space $\mathscr{M}(\mathbb{R}; S)$.

Specifically, the function $\mathbf{W_M} : \mathscr{M}(\mathbb{R}; S) \times \mathscr{M}(\mathbb{R}; S) \to [0, \infty]$ is defined by

$$\mathbf{W_M}(\mu_0, \mu_1) := \left[\inf \left\{ \int_0^1 \int_{\mathbb{R}} w^{\mathrm{T}} (\mathbf{M}(\mu))^{-1} w \,\mathrm{d}x\,\mathrm{d}t : \ (\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1) \right\} \right]^{1/2}, \tag{IV.1.4}$$

where $\mathscr{C}_1(\mu_0 \to \mu_1)$ denotes the set of all curves $(\mu, w) = (\mu_t, w_t)_{t\in[0,1]}$ satisfying the *continuity equation*

$$\partial_t \mu + \partial_x w = 0, \tag{IV.1.5}$$

having μ_0 and μ_1 as initial and terminal values, respectively.

By a careful transfer of the proofs in [74, 124] to the multi-component setting, we obtain that $\mathbf{W_M}$ inherits the essential topological properties known for the $\mathbf{W_m}$ distances, like

- existence of constant-speed geodesics connecting densities of finite distance;
- lower semicontinuity with respect to weak$*$ convergence;
- weak$*$-relative compactness of bounded sets.

We further discuss under which criteria $\mathbf{W_M}$ is finite.

In practice, the conditions (M0)–(M3) turn out to be quite restrictive, and their validity is fragile under perturbations. A seemingly trivial family of examples is given by *fully decoupled* mobility matrices, i.e.

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & & \\ & \mathbf{m}_2(z_2) & & \\ & & \ddots & \\ & & & \mathbf{m}_n(z_n) \end{pmatrix}, \tag{IV.1.6}$$

with n nonnegative and concave (scalar) mobility functions $\mathbf{m}_k : [a_k, b_k] \to \mathbb{R}$. Since the components do not interact with each other through $\mathbf{M}$, one has that

$$\mathbf{W_M}(\mu, \widetilde{\mu})^2 = \mathbf{W}_{\mathbf{m}_1}(\mu_1, \widetilde{\mu}_1)^2 + \cdots + \mathbf{W}_{\mathbf{m}_n}(\mu_n, \widetilde{\mu}_n)^2,$$

i.e., $\mathbf{W_M}$ is simply the canonical product of the metrics $\mathbf{W}_{\mathbf{m}_k}$ for each of the components. Surprisingly, it turns out that fully decoupled matrices are ungeneric for property (M2) in the sense that any sufficiently general, arbitrarily small perturbation of the components of $\mathbf{M}$ makes (M2) invalid. We shall show how certain fully decoupled matrices can be "stabilized" with a suitably chosen special perturbation such that the perturbed mobility matrix retains (M2).

Apart from the rigorous definition of $\mathbf{W_M}$, we analyse the property of geodesic λ-convexity with respect to that distance (see Chapter IV.3). Generally spoken, geodesic convexity w.r.t. transportation metrics is a very rare property [133]. Up to now, the only known λ-convex functionals $\mathcal{E}$ for the metrics $\mathbf{W_m}$ with nonlinear mobilities $\mathbf{m}$ in space dimension $d = 1$ are the *internal energies*

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu(x))\,\mathrm{d}x, \tag{IV.1.7}$$

provided that f satisfies the generalized McCann condition [53] (see I.16) — which reduces to the usual convexity of f in $d = 1$ — and the *regularized potential energies*

$$\mathcal{V}(\mu) = \int_{\mathbb{R}} \left[\alpha h(\mu(x)) + \rho(x)\mu(x)\right] \,\mathrm{d}x, \tag{IV.1.8}$$

where $h : [0, \infty) \to \mathbb{R}$ is such that $\mathbf{m}h'' \equiv 1$, $\alpha > 0$, and $\rho : \mathbb{R} \to \mathbb{R}$ is a smooth function of compact support [125]. The respective gradient flow evolutions are given by

$$\partial_t \mu = \partial_{xx} P(\mu), \qquad \text{and} \qquad \partial_t \mu = \alpha \partial_{xx} \mu + \partial_x \big(\mathbf{m}(\mu) \partial_x \rho\big),$$

respectively, where $P'(z) = \mathbf{m}(z) f''(z)$.

Both types of functionals (IV.1.7) and (IV.1.8) possess canonical generalizations to densities with multiple components. In (IV.1.7), simply replace f by a smooth function $f : S \to \mathbb{R}$. To make sense of (IV.1.8), assume that $\mathbf{M}$ is induced by $h : S \to \mathbb{R}$, see (IV.1.3), and use a potential $\rho : \mathbb{R} \to \mathbb{R}^n$ with n components. We derive sufficient criteria for the geodesic λ-convexity of these functionals with respect to the new metric $\mathbf{W_M}$. For that, we use the formalism from Section I.2.3 developed by Liero and Mielke [134, 123], which is based on the *Eulerian calculus* for transportation distances, see [64, 155].

Our own generalization of McCann's condition for $\mathcal{E}$ of the form (IV.1.7) is given in Proposition IV.20, see formula (IV.3.9). Our examples for pairs of a (nondiagonal) mobility matrix $\mathbf{M}$ and a function f that satisfy this condition are limited to perturbations of the heat equation. In contrast, if $\mathbf{M}$ is a fully decoupled mobility, $\mathcal{E}$ has to be decoupled in order to satisfy our generalized McCann condition.

Our condition assuring geodesic λ-convexity for functionals of type (IV.1.8) is given in (IV.3.12). Even for smooth ρ of compact support, it imposes an apparently very strong restriction on the function h in $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$.

In Chapter IV.4, we discuss the primary application of the new metric $\mathbf{W_M}$, namely the construction of weak solutions to a class of drift-diffusion equations of the form (IV.1.1) by means of de Giorgi's *minimizing movement scheme* (I.2.11).

We emphasize that the global existence of solutions to (IV.1.9) is a nontrivial result of independent interest. It does not follow immediately from classical parabolic theory: Indeed, since $\mathbf{M}\nabla_z^2 f$ typically lacks positivity (meaning that $v^{\mathrm{T}}\mathbf{M}\nabla_z^2 f v \geq 0$), the differential operator in (IV.1.9) is not elliptic in the strong sense. The theory for parabolic equations with normally elliptic operators, see e.g. Amann [2], provides existence of solutions only locally in time for sufficiently regular initial data; for extension of those to global solutions, additional estimates would be needed which guarantee that the values of the solution μ stay away from the boundary of the admissible set S.

Specifically, we consider the initial value problem

$$\partial_t \mu = \partial_x (\mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu + \mathbf{M}(\mu) \partial_x \eta), \qquad \mu(0) = \mu^0, \tag{IV.1.9}$$

where the mobility matrix $\mathbf{M}$ is fully decoupled as in (IV.1.6), $S \subset \mathbb{R}^n$ is an n-cuboid, $\eta \in C_c^\infty(\mathbb{R})$ is a smooth potential, and $f : S \to \mathbb{R}$ is uniformly convex, i.e., $\nabla_z^2 f(z) \geq C_f \mathbb{1}$ with $C_f > 0$, but does *not* need to be the sum of functions of the components of μ (thus giving rise to a coupling of the species via $\mathcal{E}$). Thus, the diffusion matrix $\mathbf{M}\nabla_z^2 f$ will not be symmetric nor positive definite in general. Also, the corresponding energy functional

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f\big(\mu(x)\big)\, \mathrm{d}x + \int_{\mathbb{R}} \mu(x)^{\mathrm{T}} \eta(x)\, \mathrm{d}x$$

will not be geodesically λ-convex.

Still, the minimizing movement scheme is well-posed. We prove that in the limit of vanishing time step size, it produces a limit curve that is a weak solution to (IV.1.9), see Theorem IV.30. The overall strategy of proof will be similar to that developed in Part II. In comparison to the classical results, we obtain weaker solutions of lower regularity, but we can allow for more general initial data.

The same methods apply *mutatis mutandis* to solutions to a class of fourth-order systems which formally are $\mathbf{W_M}$-gradient flows of the free energy

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\partial_x \mu, \mu) \,\mathrm{d}x,$$

with a suitable map $f : \mathbb{R}^n \times S \to \mathbb{R}$. The associated evolution equation reads as

$$\partial_t \mu = \partial_x \left(\mathbf{M}(\mu) \left[\nabla^2_{zz} f(\mu_x, \mu)\mu_x - \partial_x(\nabla^2_{pp} f(\mu_x, \mu)\mu_{xx}) - \partial_x(\nabla^2_{pz} f(\mu_x, \mu))\mu_x \right] \right). \tag{IV.1.10}$$

Whereas the first part of the right-hand side in equation (IV.1.10) is a term of second order as in equation (IV.1.9), the two remaining terms are of higher order. More specifically, the middle term resembles a multi-component version of the *Cahn-Hilliard* equation (as considered in, e.g., [125]). The last term takes into account the possible coupling of the two arguments of the energy density f. However, we show that under certain convexity and growth properties of f (similar to those required for the system of second order), this issue is not decisive for the question of global existence of weak solutions (see Theorem IV.40).

We conclude this introductory section with some preliminary remarks.

We use ∇_z for the gradient, ∇_z^2 for the Hessian and D in combination with square brackets for directional derivatives with respect to z. For instance, if $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ and $\mu : \mathbb{R} \to S$, then we write the chain rule as

$$\partial_x \mathbf{M}(\mu) = \mathrm{D}\mathbf{M}(\mu)[\partial_x \mu],$$

where the $n \times n$-matrices $\mathrm{D}\mathbf{M}(z)[\zeta]$ and $\mathrm{D}^2\mathbf{M}(z)[\zeta, \widetilde{\zeta}]$ are defined via

$$\mathrm{D}\mathbf{M}(z)[\zeta]_{ij} := \nabla_z \mathbf{M}_{ij}(z)^{\mathrm{T}} \zeta,$$
$$\mathrm{D}^2\mathbf{M}(z)[\zeta, \widetilde{\zeta}]_{ij} := \widetilde{\zeta}^{\mathrm{T}} \nabla_z^2 \mathbf{M}_{ij}(z) \zeta,$$

for all $\zeta, \widetilde{\zeta} \in \mathbb{R}^n$ and all $i, j \in \{1, \ldots, n\}$. Note that even for symmetric matrices $\mathbf{M}$, the third order-tensor $\mathrm{D}\mathbf{M}$ and the fourth-order tensor $\mathrm{D}^2\mathbf{M}$ are not totally symmetric in general, although $\mathrm{D}\mathbf{M}(z)[\zeta]$ and $\mathrm{D}^2\mathbf{M}(z)[\zeta, \widetilde{\zeta}]$ are symmetric $n \times n$ matrices, for every choice of $\zeta, \widetilde{\zeta} \in \mathbb{R}^n$. Given a multilinear operator or its tensor representative, the norm $\|\cdot\|$ denotes the operator norm. For a nonnegative measurable function $\widetilde{\mu} : \mathbb{R} \to \mathbb{R}^n$, the functional

$$\mathfrak{m}_2(\widetilde{\mu}) := \int_{\mathbb{R}} x^2 \mathsf{e}^{\mathrm{T}} \widetilde{\mu}(x) \,\mathrm{d}x \in [0, \infty]$$

is called the *second moment* of $\widetilde{\mu}$, where $\mathsf{e} := (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$. Given an arbitrary set $D \subset \mathbb{R}^m$ $(m \in \mathbb{N})$ and a closed set $A \subset \mathbb{R}^n$, $\mathscr{M}(D; A)$ denotes the space of all measurable functions $\widetilde{\mu} : D \to A$. We call a sequence of measurable functions $(\widetilde{\mu}_k)_{k\in\mathbb{N}}$ in $\mathscr{M}(D; A)$ weakly$*$-convergent to a limit $\widetilde{\mu} \in \mathscr{M}(D; A)$, if for all $\rho \in C_c^0(D; \mathbb{R}^n)$, one has

$$\lim_{k\to\infty} \int_D \widetilde{\mu}_k^{\mathrm{T}} \rho \,\mathrm{d}x = \int_D \widetilde{\mu}^{\mathrm{T}} \rho \,\mathrm{d}x.$$

Inequalities between vectors, multi-dimensional intervals (also referred to as n-cuboids) $[q^0, q^1] := \prod_{j=1}^n [q_j^0, q_j^1]$ for $q^0, q^1 \in \mathbb{R}^n$, $q^0 \leq q^1$, as well as integration of vector-valued functions are understood component-wise.

CHAPTER IV.2

Distances generated by mobility matrices

This chapter is devoted to the study of transport distances between vector-valued densities on $\mathbb{R}$. Throughout this chapter, let some convex and compact set $S \subset \mathbb{R}^n$ with nonempty interior be fixed, and recall that $\mathscr{M}(\mathbb{R}; S)$ is the space of measurable functions on $\mathbb{R}$ with values in S; moreover, assume that $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ is a mobility matrix that satisfies (M0)–(M3).

We begin by giving a rigorous interpretation of the objects occurring in the definition of $\mathbf{W}_\mathbf{M}$ above.

IV.2.1. Action density

Proposition IV.1 (Properties of the density function). *The* action density $\widetilde{\phi} : \operatorname{int}(S) \times \mathbb{R}^n \to [0, \infty)$, *defined by*

$$\widetilde{\phi}(z, p) := p^{\mathrm{T}} (\mathbf{M}(z))^{-1} p \tag{IV.2.1}$$

has the following properties:
(a) $\widetilde{\phi}$ *is continuous and (jointly) convex.*
(b) $\widetilde{\phi}$ *is nondegenerate:* $\widetilde{\phi}(z, p) > 0$ *for all* $z \in \operatorname{int}(S)$ *and* $p \neq 0$.
(c) $\widetilde{\phi}$ *is 2-homogeneous in its second argument.*

Proof. Since $\mathbf{M}$ is subject to (M0)–(M2), only convexity is not obvious. For the second directional derivative of ϕ at (z, p) in directions (ζ, π) for $\zeta \in \mathbb{R}^n$, $\pi \in \mathbb{R}^n$, we obtain

$$\mathrm{D}^2_{(z,p)} \widetilde{\phi}(z, p)[(\zeta, \pi), (\zeta, \pi)] = \pi^{\mathrm{T}} A \pi + p^{\mathrm{T}} B \pi + p^{\mathrm{T}} C p, \tag{IV.2.2}$$

with $A, B, C \in \mathbb{R}^{n\times n}$ defined as

$$\begin{aligned}
A &:= 2\mathbf{M}(z)^{-1}, \\
B &:= -4\mathbf{M}(z)^{-1} \mathrm{D}\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1}, \\
C &:= 2\mathbf{M}(z)^{-1} \mathrm{D}\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1} \mathrm{D}\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1} - \mathbf{M}(z)^{-1} \mathrm{D}^2 \mathbf{M}(z)[\zeta, \zeta] \mathbf{M}(z)^{-1}.
\end{aligned}$$

We prove that the expression in (IV.2.2) is nonnegative for all admissible choices of (z, p) and (ζ, π). Since, by condition (M1), A is symmetric positive definite, there exists a symmetric positive definite square root $A^{1/2} \in \mathbb{R}^{n\times n}$ such that $A^{1/2} A^{1/2} = A$. Further, B is symmetric. By elementary calculations, we obtain

$$\begin{aligned}
\mathrm{D}^2_{(z,p)} \widetilde{\phi}(z, p)[(\zeta, \pi), (\zeta, \pi)] &= \left| A^{1/2} \pi + \frac{1}{2} A^{-1/2} B p \right|^2 + \frac{1}{4} p^{\mathrm{T}} (4C - B A^{-1} B) p \\
&= \left| A^{1/2} \pi + \frac{1}{2} A^{-1/2} B p \right|^2 - p^{\mathrm{T}} \mathbf{M}(z)^{-1} \mathrm{D}^2 \mathbf{M}(z)[\zeta, \zeta] \mathbf{M}(z)^{-1} p,
\end{aligned}$$

which is nonnegative due to condition (M2). □

Below, we need the action density to be defined up to the boundary. To this end, we replace $\widetilde{\phi}$ by its *lower semicontinuous envelope* $\phi : S \times \mathbb{R}^n \to [0, \infty]$, defined by

$$\phi(\widetilde{z}, \widetilde{p}) := \liminf_{(z,p) \to (\widetilde{z}, \widetilde{p})} \widetilde{\phi}(z, p). \tag{IV.2.3}$$

Thanks to continuity of $\widetilde{\phi}$, we have $\phi \equiv \widetilde{\phi}$ on $\operatorname{int}(S) \times \mathbb{R}^n$.

Example IV.2. *Let* $\mathbf{M}(z) = (\nabla_z^2 h_\varepsilon(z))^{-1}$ *with* h_ε *given by (see also* (IV.2.7) *below)*

$$h_\varepsilon(z) = z_1 \log z_1 + (1-z_1)\log(1-z_1) + z_2 \log z_2 + (1-z_2)\log(1-z_2) + \varepsilon z_1 z_2 (1-z_1)(1-z_2).$$

Then, for $z = (z_1, z_2) \in [0,1]^2$ *and every* $p = (p_1, p_2) \in \mathbb{R}^2$*, we have that*

$$\phi(z,p) = \begin{cases} p^{\mathrm{T}} \nabla_z^2 h_\varepsilon(z) p & \text{if } z \in (0,1)^2, \\ \frac{p_2^2}{z_2(1-z_2)} & \text{if } z_1 \in \{0,1\},\ z_2 \in (0,1),\ p = (0,p_2), \\ \frac{p_1^2}{z_1(1-z_1)} & \text{if } z_2 \in \{0,1\},\ z_1 \in (0,1),\ p = (p_1,0), \\ 0 & \text{if } z \in \{(0,0),(1,0),(1,1),(0,1)\} \text{ and } p = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The key step in the derivation is to observe that if z *tends to a boundary point* $\tilde{z} \in \partial S$ *that is not a corner, then precisely one of the two eigenvalues of* $\nabla_z^2 h(z)$ *converges to zero, and the eigenvector for the non-vanishing eigenvalue is asymptotically parallel to* ∂S *at* $\tilde{z}$*.*

For $\mu \in \mathscr{M}(\mathbb{R}; S)$ and $w \in \mathscr{M}(\mathbb{R}; \mathbb{R}^n)$, we define the *action functional*

$$\mathbf{\Phi}(\mu, w) := \int_{\mathbb{R}} \phi\big(\mu(x), w(x)\big)\, \mathrm{d}x. \tag{IV.2.4}$$

Proposition IV.1 allows us to apply Theorem 2.1 in [74] to obtain:

Proposition IV.3 (Lower semicontinuity of the action functional). *If* $(\mu_k)_{k\in\mathbb{N}}$ *and* $(w_k)_{k\in\mathbb{N}}$ *are weakly∗-convergent sequences to* $\mu \in \mathscr{M}(\mathbb{R}; S)$ *and* $w \in \mathscr{M}(\mathbb{R}; \mathbb{R}^n)$*, respectively, then*

$$\liminf_{k\to\infty} \mathbf{\Phi}(\mu_k, w_k) \geq \mathbf{\Phi}(\mu, w).$$

IV.2.2. Examples

This section is concerned with specific examples of mobility matrices $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ that satisfy conditions (M0)–(M3) stated in the introduction. We will occasionally also consider the following stronger version of (M2):

(M2') The matrix $\mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta] \in \mathbb{R}^{n\times n}$ is negative definite for all $z \in \operatorname{int}(S)$ and $\zeta \in \mathbb{R}^n \backslash \{0\}$.

All of our examples are of the form (IV.1.3), where $\mathbf{M}$ is induced by a convex function h.

IV.2.2.1. Fully decoupled mobilities

Consider concave functions $\mathbf{m}_1, \ldots, \mathbf{m}_n$ with $\mathbf{m}_j : [S_j^\ell, S_j^r] \to \mathbb{R}$, $S_j^\ell < S_j^r$, such that $\mathbf{m}_j(s) > 0$ for $s \in (S_j^\ell, S_j^r)$ and $\mathbf{m}_j(S_j^\ell) = \mathbf{m}_j(S_j^r) = 0$, for each $j \in \{1, \ldots, n\}$. Define a mobility matrix $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ on the n-cuboid $S := [S^\ell, S^r]$ by

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & \\ & \ddots & \\ & & \mathbf{m}_n(z_n) \end{pmatrix}. \tag{IV.2.5}$$

Clearly, $\mathbf{M}$ is of the form (IV.1.3) when

$$h(z) = h_1(z_1) + \cdots + h_n(z_n),$$

and each $h_j : (S_j^\ell, S_j^r) \to \mathbb{R}$ is a second primitive of the respective $\frac{1}{\mathbf{m}_j}$, i.e., $\mathbf{m}_j(s)h_j''(s) = 1$. It is immediately verified that $\mathbf{M}$ satisfies (M0)–(M3). Concerning property (M2), we remark that

$$\beta^{\mathrm{T}}\mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta]\beta = \sum_{j=1}^{n} \mathbf{m}_j''(z_j)\,(\zeta_j\beta_j)^2,$$

hence the sharper condition (M2') is *not* satisfied, even if all $\mathbf{m}_j$ are *strictly* concave functions. This is the reason why the concavity (M2) is lost under generic perturbations of $\mathbf{M}$. In the next example below, we discuss a very special "perturbation" of a particular matrix of type (IV.2.5), for which (M2') is valid.

For obvious reasons, we call mobility matrices $\mathbf{M}$ of the form (IV.2.5) *fully decoupled*: the different species do not influence the mobilities of each other. It is clear that each fully decoupled matrix $\mathbf{M}$ induces a metric on $\mathscr{M}(\mathbb{R};S)$, simply applying the theory from [74, 124] to each component separately.

IV.2.2.2. Perturbations of a fully decoupled mobility

Let us now specialize the previous example by choosing $n = 2$ components, $S = [0,1]^2$ and the map $h_0 : (0,1)^2 \to \mathbb{R}$ with

$$h_0(z) = z_1\log z_1 + (1-z_1)\log(1-z_1) + z_2\log z_2 + (1-z_2)\log(1-z_2). \tag{IV.2.6}$$

From (IV.1.3), we obtain the fully decoupled mobility matrix

$$\mathbf{M}_0(z) = \big(\nabla_z^2 h_0(z)\big)^{-1} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}, \quad \text{with} \quad d_j = \mathbf{m}(z_j),$$

where $\mathbf{m}(s) = s(1-s)$. By the discussion above, (M0)–(M3) are satisfied, but (M2') is not. It is easily seen that for a general (smooth, compactly supported) function $g : (0,1)^2 \to \mathbb{R}$, the matrix $\widetilde{\mathbf{M}}_\varepsilon = (\nabla_z^2(h_0+\varepsilon g))^{-1}$ does *not* satisfy (M2) anymore, no matter how small $\varepsilon > 0$ is.

Let us introduce a very special perturbation h_ε of h_0:

$$h_\varepsilon(z) = h_0(z) + \varepsilon z_1 z_2(1-z_1)(1-z_2) = h_0(z) + \varepsilon d_1 d_2. \tag{IV.2.7}$$

We are going to show that $\mathbf{M}_\varepsilon(z) = (\mathbf{H}_\varepsilon(z))^{-1}$, with

$$\mathbf{H}_\varepsilon(z) := \nabla_z^2 h_\varepsilon(z) = \begin{pmatrix} \frac{1}{d_1} & 0 \\ 0 & \frac{1}{d_2} \end{pmatrix} + \varepsilon \begin{pmatrix} -2d_2 & d_1'd_2' \\ d_1'd_2' & -2d_1 \end{pmatrix}, \quad \text{with} \quad d_1' = 1-2z_1,\ d_2' = 1-2z_2,$$

satisfies (M0)–(M3), and in addition also (M2'), for all sufficiently small $\varepsilon > 0$. Thus, this special perturbation makes the mobility matrix robust with respect to further (smaller) generic perturbations.

First, note that $\mathbf{M}_\varepsilon(z)$ is well-defined at $z \in (0,1)^2$ if

$$\det \mathbf{H}_\varepsilon(z) = \frac{1}{d_1 d_2} - 4\varepsilon + \varepsilon^2\big(4d_1d_2 - (d_1'd_2')^2\big) \tag{IV.2.8}$$

is positive. This is true simultaneously at all $z \in (0,1)^2$ if $\varepsilon > 0$ is sufficiently small. It is further easily seen that $\mathbf{M}_\varepsilon$ extends continuously to the boundary of S by setting $\mathbf{M}_\varepsilon(z) = \mathbf{M}_0(z)$ for $z \in \partial S$; just observe that

$$\mathbf{M}_\varepsilon(z) = \frac{1}{1-\varepsilon d_1 d_2\big[4-\varepsilon\big(4d_1d_2-(d_1'd_2')^2\big)\big]}\left[\mathbf{M}_0(z) - \varepsilon d_1 d_2 \begin{pmatrix} 2d_1 & d_1'd_2' \\ d_1'd_2' & 2d_2 \end{pmatrix}\right],$$

and that $d_1 d_2 \searrow 0$ as $z \to \partial S$. This implies (M0) and (M3) for $\mathbf{M}_\varepsilon$.

Next, since the entries of $\mathbf{M}_\varepsilon$ vary continuously with ε, and since $\det\mathbf{M}_\varepsilon(z) = (\det\mathbf{H}_\varepsilon(z))^{-1}$ never vanishes for any $z \in (0,1)^2$ and any sufficiently small $\varepsilon > 0$, it follows that $\mathbf{M}_\varepsilon$ inherits the positive definiteness of $\mathbf{M}_0$. Thus, also (M1) is verified.

The proof of condition (M2') is more involved. To begin with, observe that $\mathbf{M}_\varepsilon(z) = (\mathbf{H}_\varepsilon(z))^{-1}$ implies

$$D^2\mathbf{M}_\varepsilon(z)[\zeta,\zeta] = -\mathbf{H}_\varepsilon(z)^{-1}D^2\mathbf{H}_\varepsilon(z)[\zeta,\zeta]\mathbf{H}_\varepsilon(z)^{-1} + 2\mathbf{H}_\varepsilon(z)^{-1}D\mathbf{H}_\varepsilon(z)[\zeta]\mathbf{H}_\varepsilon(z)^{-1}D\mathbf{H}_\varepsilon(z)[\zeta]\mathbf{H}_\varepsilon(z)^{-1}.$$

Thus, for proving (M2'), it suffices to show that for all $z \in (0,1)^2$ and all $\beta, \zeta \in \mathbb{R}^n \setminus \{0\}$,

$$P := (d_1d_2)^3\beta^{\mathrm{T}}\Big(\det\mathbf{H}_\varepsilon(z)\,D^2\mathbf{H}_\varepsilon(z)[\zeta,\zeta] - 2D\mathbf{H}_\varepsilon(z)[\zeta]\mathbf{H}_\varepsilon(z)^{+}D\mathbf{H}_\varepsilon(z)[\zeta]\Big)\beta > 0,$$

where $\det\mathbf{H}_\varepsilon$ is given in (IV.2.8), and

$$D\mathbf{H}_\varepsilon(z)[\zeta] = \begin{pmatrix} -\frac{d_1'}{d_1^2}\zeta_1 & 0 \\ 0 & -\frac{d_2'}{d_2^2}\zeta_2 \end{pmatrix} - 2\varepsilon\begin{pmatrix} d_2'\zeta_2 & d_1'\zeta_2 + d_2'\zeta_1 \\ d_1'\zeta_2 + d_2'\zeta_1 & d_1'\zeta_1 \end{pmatrix},$$

$$D^2\mathbf{H}_\varepsilon(z)[\zeta,\zeta] = 2\begin{pmatrix} \frac{1-3d_1}{d_1^3}\zeta_1^2 & 0 \\ 0 & \frac{1-3d_2}{d_2^3}\zeta_2^2 \end{pmatrix} + 4\varepsilon\begin{pmatrix} \zeta_2^2 & 2\zeta_1\zeta_2 \\ 2\zeta_1\zeta_2 & \zeta_1^2 \end{pmatrix},$$

$$\mathbf{H}_\varepsilon(z)^{+} = \det\mathbf{H}_\varepsilon(z)\,\mathbf{H}_\varepsilon(z)^{-1} = \begin{pmatrix} \frac{1}{d_2} & 0 \\ 0 & \frac{1}{d_1} \end{pmatrix} + \varepsilon\begin{pmatrix} -2d_1 & -d_1'd_2' \\ -d_1'd_2' & -2d_2 \end{pmatrix}.$$

A tedious but straightforward calculation leads to the following explicit representation of P, with the abbreviations $\widetilde{\zeta}_1 := d_2\zeta_1$, $\widetilde{\zeta}_2 := d_1\zeta_2$:

$$\begin{aligned} P =& 2\big[\widetilde{\zeta}_1^2 + 2\varepsilon(d_2\widetilde{\zeta}_2)^2\big]\beta_1^2 + 2\big[\widetilde{\zeta}_2^2 + 2\varepsilon(d_1\widetilde{\zeta}_1)^2\big]\beta_2^2 + \varepsilon\big[\widehat{f}_1\widetilde{\zeta}_1^2 + \varepsilon f_2(d_2\widetilde{\zeta}_2)^2 + \widehat{f}_3\widetilde{\zeta}^1(d_2\widetilde{\zeta}_2)\big]\beta_1^2 \\ &+ \varepsilon\big[\check{f}_1\widetilde{\zeta}_2^2 + \varepsilon f_2(d_1\widetilde{\zeta}_1)^2 + \check{f}_3\widetilde{\zeta}_2(d_1\widetilde{\zeta}_1)\big]\beta_2^2 + \varepsilon\big[\widehat{f}_4\widetilde{\zeta}_1(d_1\widetilde{\zeta}_1) + \check{f}_4\widetilde{\zeta}_2(d_2\widetilde{\zeta}_2) + 2f_5\widetilde{\zeta}_1\widetilde{\zeta}_2\big]\beta_1\beta_2, \end{aligned}$$

where the functions f_i, $\widehat{f}_i$ and $\check{f}_i$ are bounded uniformly with respect to $z \in (0,1)^2$ and (small) $\varepsilon > 0$:

$$\begin{aligned} \widehat{f}_1 &:= 2d_2(8d_1 - 3) + \varepsilon(d_2(-64d_1 + 132d_1^2 + 8) - 39d_1^2 - 2 + 18d_1) - 8\varepsilon^2 d_1^3 d_2^2(1 - 4d_2), \\ \check{f}_1 &:= 2d_1(8d_2 - 3) + \varepsilon(d_1(-64d_2 + 132d_2^2 + 8) - 39d_2^2 - 2 + 18d_2) - 8\varepsilon^2 d_2^3 d_1^2(1 - 4d_1), \\ f_2 &:= 4d_1d_2 - d_1 - d_2 - \varepsilon d_1d_2(28d_1d_2 - 10d_1 - 10d_2 + 3), \\ \widehat{f}_3 &:= 4d_1'd_2'\big(1 + \varepsilon d_2(1 - 2\varepsilon(1 - 2d_2)d_1^2)\big), \\ \check{f}_3 &:= 4d_1'd_2'\Big(1 + \varepsilon d_1(1 - 2\varepsilon(1 - 2d_1)d_2^2)\big), \\ \widehat{f}_4 &:= -2d_1'd_2'\big(1 + 4\varepsilon(d_1 - 1 + 2\varepsilon d_2 d_1^2(1 - 2d_2))\big), \\ \check{f}_4 &:= -2d_1'd_2'\big(1 + 4\varepsilon(d_2 - 1 + 2\varepsilon d_1 d_2^2(1 - 2d_1))\big), \\ f_5 &:= 1 + 40d_1d_2 - 6d_1 - 6d_2 - 32\varepsilon d_1^2d_2^2 - 4\varepsilon^2 d_1^2 d_2^2(5 + 56d_1d_2 - 18d_1 - 18d_2). \end{aligned}$$

From elementary calculations — applying the Cauchy-Schwarz and Young inequalities and collecting terms — we conclude that

$$P \geq \big[\widetilde{\zeta}_1^2 + 2\varepsilon(d_2\widetilde{\zeta}_2)^2\big]\beta_1^2 + \big[\widetilde{\zeta}_2^2 + 2\varepsilon(d_1\widetilde{\zeta}_1)^2\big]\beta_2^2$$

for arbitrary $z \in (0,1)^2$, $\beta, \zeta \in \mathbb{R}^n$, and all sufficiently small $\varepsilon > 0$. This implies positivity of P for $\beta \neq 0$ and $\zeta \neq 0$, and therefore proves (M2').

IV.2.2.3. Volume filling mobility

The following example describes the interaction of species that influence the mobilities of each other by competing for limited volume. This example is related but not identical to the one considered in [123], where in addition a microlocal conservation of mass was assumed.

Define the n^{th} *standard simplex*

$$S := \left\{ z \in [0,1]^n : 1 - \sum_{j=1}^n z_j \geq 0 \right\}$$

as value space and the map $h : \operatorname{int}(S) \to \mathbb{R}$ by

$$h(z) := \sum_{j=1}^n z_j \log z_j + \left(1 - \sum_{j=1}^n z_j\right) \log \left(1 - \sum_{j=1}^n z_j\right).$$

The second-order partial derivatives of h amount to

$$\frac{\partial^2 h}{\partial z_i \partial z_j}(z) = \frac{1}{z_i}\delta_{ij} + \left(1 - \sum_{\ell=1}^n z_\ell\right)^{-1},$$

where δ_{ij} denotes Kronecker's delta. By elementary calculations, we obtain the explicit form of the inverse matrix:

$$\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} - zz^{\mathrm{T}}.$$

Property (M0) obviously holds. To verify (M1), let $\gamma \in \mathbb{R}^n$ be given and observe that

$$\gamma^{\mathrm{T}} zz^{\mathrm{T}} \gamma = \sum_{i=1}^n \sum_{j=1}^n z_i z_j \gamma_i \gamma_j \leq \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n z_i z_j (\gamma_i^2 + \gamma_j^2) = \left(\sum_{j=1}^n \gamma_j^2 z_j\right) \left(\sum_{\ell=1}^n z_\ell\right).$$

Therefore,

$$\gamma^{\mathrm{T}} \mathbf{M}(z) \gamma = \sum_{j=1}^n \gamma_j^2 z_j - \gamma^{\mathrm{T}} zz^{\mathrm{T}} \gamma \geq \sum_{j=1}^k \gamma_j^2 z_j \left(1 - \sum_{\ell=1}^n z_\ell\right),$$

which is positive for all $z \in \operatorname{int}(S)$ and $\gamma \neq 0$. Condition (M2) is immediately obtained from

$$\mathrm{D}^2 \mathbf{M}(z)[\zeta, \zeta] = -2\zeta\zeta^{\mathrm{T}},$$

which is negative semidefinite, for arbitrary $z \in \operatorname{int}(S)$ and $\zeta \in \mathbb{R}^n$. Note that $\mathrm{D}^2\mathbf{M}(z)[\zeta, \zeta]$ has rank one, hence the stronger condition (M2') is not satisfied. Finally, let $z \in \partial S$, and let ν be a normal vector to ∂S at z. We distinguish two cases. In the first, z lies on one of the coordinate hyperplanes. Then $\nu_j \neq 0$ only if $z_j = 0$, for $j = 1, \dots, n$, and so clearly $\mathbf{M}(z)\nu = 0$. In the second case, we have $z_1 + \cdots + z_n = 1$. Hence the normal vector is (a multiple of) $\mathsf{e} = (1, \dots, 1)^{\mathrm{T}}$, and therefore

$$\mathbf{M}(z)\mathsf{e} = z - z(z^{\mathrm{T}}\mathsf{e}) = \left(1 - \sum_{\ell=1}^n z_\ell\right) z = 0.$$

This proves (M3).

IV.2.2.4. Radially symmetric mobility

On the n-dimensional closed unit ball $S := \overline{\mathbb{B}_1(0)}$, define $h : S \to \mathbb{R}$ by

$$h(z) = \log\left(1 + \sqrt{1 - |z|^2}\right) - \sqrt{1 - |z|^2}.$$

One easily verifies that

$$\nabla_z^2 h(z) = \frac{1}{1+\sqrt{1-|z|^2}}\mathbb{1} + \frac{1}{\left(1+\sqrt{1-|z|^2}\right)\sqrt{1-|z|^2}}\frac{zz^{\mathrm{T}}}{|z|^2},$$

which obviously is positive definite for $z \in \mathbb{B}_1(0)$. Now define $\mathbf{M}$ by (IV.1.3), i.e.,

$$\begin{aligned}\mathbf{M}(z) &= \left(\nabla_z^2 h(z)\right)^{-1} \\ &= \left(1+\sqrt{1-|z|^2}\right)\mathbb{1} + \left[\left(1+\sqrt{1-|z|^2}\right)\sqrt{1-|z|^2} - \left(1+\sqrt{1-|z|^2}\right)\right]\frac{zz^{\mathrm{T}}}{|z|^2} \\ &= \left(1+\sqrt{1-|z|^2}\right)\mathbb{1} - zz^{\mathrm{T}}.\end{aligned}$$

Conditions (M0) and (M1) obviously hold. Next, for arbitrary $\zeta \in \mathbb{R}^n$ and $z \in \operatorname{int}(S)$, we have that

$$\begin{aligned}\mathrm{D}\mathbf{M}(z)[\zeta] &= -(1-|z|^2)^{-1/2}(z^{\mathrm{T}}\zeta)\mathbb{1} - z\zeta^{\mathrm{T}} - \zeta z^{\mathrm{T}} \\ \mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta] &= -(1-|z|^2)^{-3/2}(z^{\mathrm{T}}\zeta)^2\mathbb{1} - 2\zeta\zeta^{\mathrm{T}} - (1-|z|^2)^{-1/2}|\zeta|^2\mathbb{1}.\end{aligned}$$

The last matrix is obviously negative definite for each $z \in \mathbb{B}_1(0)$ and $\zeta \neq 0$, which shows (M2'). Finally, to verify (M3), let $z \in S$ with $|z| = 1$ be given, and observe that z itself is a normal vector to ∂S at z. One has

$$\mathbf{M}(z)z = \left(1+\sqrt{1-|z|^2}\right)z - |z|^2 z = z - z = 0.$$

IV.2.3. Solutions to the continuity equation

Next, we investigate the structure of solutions to the (multi-component) *continuity equation* (IV.1.5).

Since the components of μ and w are decoupled in (IV.1.5), most of the results below follow from a "component-wise application" of the corresponding results in [74, 124].

Definition IV.4 (The class $\mathscr{C}_T$). *Given $T > 0$, define $\mathscr{C}_T$ as the set of all curves $(\mu, w) = (\mu_t, w_t)_{t\in[0,T]}$ with the following properties:*

(a) $(\mu_t)_{t\in[0,T]}$ is a weakly$*$-continuous curve in $\mathscr{M}(\mathbb{R}; S)$;
(b) $(w_t)_{t\in[0,T]}$ is a Borel-measurable family in $\mathscr{M}(\mathbb{R}; \mathbb{R}^n)$;
(c) For each $R > 0$ and $j \in \{1, \dots, n\}$,

$$\int_0^T \int_{-R}^R |w_j|_t \,\mathrm{d}x\,\mathrm{d}t < \infty;$$

(d) (μ, w) is a distributional solution to (IV.1.5) *on $[0, T] \times \mathbb{R}$.*

Furthermore, we denote by $\mathscr{C}_T(\widehat{\mu} \to \check{\mu})$ the subset of those $(\mu, w) \in \mathscr{C}_T$ with $\mu|_{t=0} = \widehat{\mu}$ and $\mu|_{t=T} = \check{\mu}$.

The continuity property (a) above imposes no restriction on the curve $(\mu_t, w_t)_{t\in[0,T]}$. Indeed, by componentwise application of Lemma 4.1 from [74], one deduces that every $(\mu_t, w_t)_{t\in[0,T]}$ satisfying (b)–(d) possesses a uniquely determined weakly$*$-continuous representative.

Lemma IV.5 (Time rescaling). *Let $\sigma : [0, T'] \to [0, T]$ be almost everywhere equal to a diffeomorphism. Then (μ, w) is a distributional solution of* (IV.1.5) *on $[0, T] \times \mathbb{R}$ if and only if $(\widehat{\mu}, \widehat{w}) := (\mu \circ \sigma, \sigma' \cdot w \circ \sigma)$ is a distributional solution of* (IV.1.5) *on $[0, T'] \times \mathbb{R}$.*

Proof. See [4, Lemma 8.1.3]. □

LEMMA IV.6 (Glueing lemma). *Let $(\widehat{\mu}, \widehat{w}) \in \mathscr{C}_{T_1}(\mu_0 \to \mu_1)$, $(\widehat{\widehat{\mu}}, \widehat{\widehat{w}}) \in \mathscr{C}_{T_2}(\mu_1 \to \mu_2)$. Then the concatenation $(\mu, w) = (\mu_t, w_t)_{t\in[0,T]}$ with $T = T_1 + T_2$, defined by*

$$(\mu_t, w_t) := \begin{cases} (\widehat{\mu}_t, \widehat{w}_t) & \text{for } t \in [0, T_1], \\ (\widehat{\widehat{\mu}}_{t-T_1}, \widehat{\widehat{w}}_{t-T_1}) & \text{for } t \in (T_1, T_1 + T_2], \end{cases}$$

is an element of $\mathscr{C}_T(\mu_0 \to \mu_2)$.

PROOF. This is a direct consequence of Lemma IV.5, see for instance [74]. □

DEFINITION IV.7. *The* energy $\mathbf{E}_T$ *of a curve $(\mu, w) = (\mu_t, w_t)_{t\in[0,T]} \in \mathscr{C}_T$ is defined by*

$$\mathbf{E}_T(\mu, w) := \int_0^T \mathbf{\Phi}(\mu_t, w_t)\,\mathrm{d}t.$$

PROPOSITION IV.8 (Compactness in $\mathscr{C}_T$, part I). *Let $(\mu_k, w_k)_{k\in\mathbb{N}}$ be a sequence in $\mathscr{C}_T$ such that for each fixed $R > 0$, the family*

$$\left\{ t \mapsto \int_{-R}^{R} |w_{k,j}|_t\,\mathrm{d}x : \, k \in \mathbb{N}, \, j \in \{1, \dots, n\} \right\}$$

of maps from $(0, T)$ to $\mathbb{R}$ is k-uniformly integrable. Then, there exists a subsequence (non-relabelled) and a limit $(\mu, w) \in \mathscr{C}_T$ such that for $k \to \infty$:

$$\begin{aligned}
&(\mu_k)_t \stackrel{*}{\rightharpoonup} \mu_t \text{ weakly}* \text{ in } \mathscr{M}(\mathbb{R}; S) \text{ for every } t \in [0, T], \\
&w_k \stackrel{*}{\rightharpoonup} w \text{ weakly}* \text{ in } \mathscr{M}((0, T) \times \mathbb{R}; \mathbb{R}^n), \\
&\mathbf{E}_T(\mu, w) \le \liminf_{k\to\infty} \mathbf{E}_T(\mu_k, w_k).
\end{aligned}$$

PROOF. Apply Lemma 4.5 in [74] componentwise. □

PROPOSITION IV.9 (Compactness in $\mathscr{C}_T$, part II). *Let $(\mu_k, w_k)_{k\in\mathbb{N}}$ be a sequence in $\mathscr{C}_T$ of uniformly bounded energy,*

$$\sup_{k\in\mathbb{N}} \mathbf{E}_T(\mu_k, w_k) < \infty.$$

Then the hypotheses of Proposition IV.8 are fulfilled.

PROOF. To begin with, observe that thanks to continuity of $\mathbf{M}$ by (M0), there exists a constant $C_{\mathbf{M}} > 0$ such that $\|\mathbf{M}(z)\| \le C_{\mathbf{M}}$ for all $z \in S$. Hence $\phi(z, p) \ge C_{\mathbf{M}}^{-1}|p|^2$, for all $(z, p) \in S \times \mathbb{R}^n$. For given $R > 0$, we have that:

$$\begin{aligned}
&\sum_{j=1}^{n} \int_0^T \left[\int_{-R}^{R} |w_{k,j}|_t\,\mathrm{d}x \right]^2 \mathrm{d}t \\
&\le \sum_{j=1}^{n} \int_0^T 2R \int_{\mathbb{R}} |w_{k,j}|_t^2\,\mathrm{d}x\,\mathrm{d}t \le 2RC_{\mathbf{M}} \int_0^T \int_{\mathbb{R}} \phi(\mu_k, w_k)\,\mathrm{d}x\,\mathrm{d}t \le 2RC_{\mathbf{M}} \sup_{k\in\mathbb{N}} \mathbf{E}_T(\mu_k, w_k) < \infty.
\end{aligned}$$

This proves that the family $\int_{-R}^{R} |(w_k)_t|\,\mathrm{d}x$ is k-uniformly bounded in $L^2((0, T); \mathbb{R}^n)$. □

IV.2.4. Distance functional and its topological properties

In this section, we prove that the functional $\mathbf{W}_{\mathbf{M}}$ defined as

$$\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) := \left[\inf\{\mathbf{E}_1(\mu, w) : \, (\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)\}\right]^{1/2} \tag{IV.2.9}$$

is a (pseudo-)metric on $\mathscr{M}(\mathbb{R}; S)$ and investigate topological properties of $\mathbf{W}_{\mathbf{M}}$.

We begin by showing that the infimum in (IV.2.9) is either equal to infinity, or is in fact a minimum. In the latter case, any curve in $\mathscr{C}_1(\mu_0 \to \mu_1)$ that attains the minimal value can be considered as a constant speed minimal geodesic joining μ_0 and μ_1.

Proposition IV.10 (Minimizers and equivalent characterization). *The following statements hold:*

(a) If the infimum W occurring in $\mathbf{W_M}$ is finite, then it is attained by a curve $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$, for which one has

$$\mathbf{\Phi}(\mu_t, w_t) = W \quad \textit{for a.e. } t \in (0,1).$$

Consequently,

$$\mathbf{W_M}(\mu_s, \mu_t) = |t-s|\mathbf{W_M}(\mu_0, \mu_1) \quad \textit{for all } s,t \in [0,1].$$

(b) There are two equivalent characterizations of $\mathbf{W_M}$: for all $T > 0$,

$$\mathbf{W_M}(\mu_0, \mu_1) = [\inf\{T\mathbf{E}_T(\mu, w) : (\mu, w) \in \mathscr{C}_T(\mu_0 \to \mu_1)\}]^{1/2} \tag{IV.2.10}$$

$$= \inf\left\{\int_0^T [\mathbf{\Phi}(\mu_t, w_t)]^{1/2}\,\mathrm{d}t : (\mu, w) \in \mathscr{C}_T(\mu_0 \to \mu_1)\right\}. \tag{IV.2.11}$$

Proof. The proof of (IV.2.11) is essentially the same as in [74, 4], using the rescaling lemma (Lemma IV.5). The other characterization (IV.2.10) can also be obtained by this lemma using a linear rescaling of time.

For the proof of statement (a), assume that $\mathbf{W_M}(\mu_0, \mu_1) = W^{1/2} < \infty$ for $W \geq 0$. Then, there exists a sequence $(\mu_k, w_k)_{k\in\mathbb{N}}$ in $\mathscr{C}_1(\mu_0 \to \mu_1)$ with $\sup_{k\in\mathbb{N}} \mathbf{E}_1(\mu_k, w_k) < \infty$. The application of the Propositions IV.9 and IV.8 yields a limit curve $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$ that is a minimizer of $\mathbf{E}_1$ on $\mathscr{C}_1(\mu_0 \to \mu_1)$ due to weak∗-lower semicontinuity. With (IV.2.11), one deduces

$$W^{1/2} = \int_0^1 \mathbf{\Phi}(\mu_t, w_t)^{1/2}\,\mathrm{d}t,$$

and consequently, since $(0,1) \ni t \mapsto \mathbf{\Phi}^{1/2}(\mu_t, w_t)$ and $(0,1) \ni t \mapsto 1$ yield equality in Hölder's inequality, $\mathbf{\Phi}(\mu_t, w_t) = W$ for almost every $t \in (0,1)$. □

We are now in position to prove that $\mathbf{W_M}$ is a distance.

Proposition IV.11 ($\mathbf{W_M}$ is a pseudometric). $\mathbf{W_M}$ *is a (possibly ∞-valued) metric on the space $\mathscr{M}(\mathbb{R}; S)$.*

Proof. *Symmetry.* This is immediate from the 2-homogeneity of ϕ and the rescaling lemma (Lemma IV.5).
Definiteness. $\mathbf{W_M}(\mu_0, \mu_1) = 0$ if and only if $\mathbf{E}_1(\mu, w) = 0$ for some $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$. From positive definiteness of $\mathbf{M}$, this is the case if and only if $w \equiv 0$ for some $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$, hence if and only if $\mu_0 = \mu_1$.
Triangle inequality. Let $\mu_0, \mu_1, \mu_2 \in \mathscr{M}(\mathbb{R}; S)$. If $\mathbf{W_M}(\mu_0, \mu_1)$ or $\mathbf{W_M}(\mu_1, \mu_2)$ is equal to $+\infty$, there is nothing to prove. If both are finite, we can use the second equivalent characterization of $\mathbf{W_M}$ (IV.2.11) and the glueing lemma (Lemma IV.6) to construct $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$ for which one has that $\mathbf{W_M}(\mu_0, \mu_1) + \mathbf{W_M}(\mu_1, \mu_2) = \int_0^1 \mathbf{\Phi}(\mu_t, w_t)^{1/2}\,\mathrm{d}t$. Again, invoking (IV.2.11), we obtain the triangle inequality. □

The following topological results are a consequence of the compactness results of Section IV.2.3, in particular of Proposition IV.9.

Proposition IV.12 (Topological properties). *The following statements hold:*

(a) $\mathbf{W_M}$ is lower semicontinuous in both components with respect to weak∗ convergence.

(b) *Let $\mu_0 \in \mathscr{M}(\mathbb{R}; S)$ be fixed, but arbitrary and let $K \subset \mathscr{M}(\mathbb{R}; S)$. If there exists $C \in \mathbb{R}$ such that $\mathbf{W_M}(\mu_0, \mu) \leq C$ for all $\mu \in K$, then K is relatively compact in the weak$*$ topology.*
(c) *Let $\mu_0 \in \mathscr{M}(\mathbb{R}; S)$ be fixed, but arbitrary and define $\mathbf{X}[\mu_0] := \{\mu \in \mathscr{M}(\mathbb{R}; S) : \mathbf{W_M}(\mu_0, \mu) < \infty\}$. Then, the metric space $(\mathbf{X}[\mu_0], \mathbf{W_M})$ is complete.*
(d) *$\mathbf{W}^2_\mathbf{M}$ is convex with respect to the linear structure of $\mathscr{M}(\mathbb{R}; S)$: if μ_0, μ_1, $\widetilde{\mu}_0$, $\widetilde{\mu}_1 \in \mathscr{M}(\mathbb{R}; S)$ and $\tau \in [0, 1]$, then*

$$\mathbf{W}^2_\mathbf{M}((1-\tau)\mu_0 + \tau\widetilde{\mu}_0, (1-\tau)\mu_1 + \tau\widetilde{\mu}_1) \leq (1-\tau)\mathbf{W}^2_\mathbf{M}(\mu_0, \mu_1) + \tau\mathbf{W}^2_\mathbf{M}(\widetilde{\mu}_0, \widetilde{\mu}_1).$$

(e) *Let $\Gamma \in C^\infty(\mathbb{R})$ be nonnegative, with support in $[-1, 1]$ and $\|\Gamma\|_{L^1} = 1$, and define $\Gamma_\varepsilon(x) := \frac{1}{\varepsilon}\Gamma\left(\frac{x}{\varepsilon}\right)$ for $\varepsilon > 0$. For all $\mu_0, \mu_1 \in \mathscr{M}(\mathbb{R}; S)$, the following holds:*

$$\mathbf{W_M}(\mu_0 * \Gamma_\varepsilon, \mu_1 * \Gamma_\varepsilon) \leq \mathbf{W_M}(\mu_0, \mu_1),$$
$$\lim_{\varepsilon\to 0} \mathbf{W_M}(\mu_0 * \Gamma_\varepsilon, \mu_1 * \Gamma_\varepsilon) = \mathbf{W_M}(\mu_0, \mu_1).$$

Proof. (a) Let $(\mu_{0,k}, \mu_{1,k})_{k\in\mathbb{N}}$ be weakly$*$ convergent to (μ_0, μ_1) as $k \to \infty$. Without loss of generality, there exists $Z \geq 0$ such that $\sup_{k\in\mathbb{N}} \mathbf{W_M}(\mu_{0,k}, \mu_{1,k}) \leq Z$. From Proposition IV.10(a), we obtain a sequence $(\mu_k, w_k)_{k\in\mathbb{N}}$ with $(\mu_k, w_k) \in \mathscr{C}_1(\mu_{0,k} \to \mu_{1,k})$ such that $\mathbf{W}^2_\mathbf{M}(\mu_{0,k}, \mu_{1,k}) = \mathbf{\Phi}((\mu_k)_t, (w_k)_t) \leq Z^2$ for almost every $t \in [0, 1]$ and all $k \in \mathbb{N}$. Hence, the requirement of Proposition IV.9 is fulfilled. The application of this proposition together with Proposition IV.8 now yields a limit curve $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$ and

$$\liminf_{k\to\infty} \mathbf{W}^2_\mathbf{M}(\mu_{0,k}, \mu_{1,k}) = \liminf_{k\to\infty} \mathbf{E}_1(\mu_k, w_k) \geq \mathbf{E}_1(\mu, w) \geq \mathbf{W}^2_\mathbf{M}(\mu_0, \mu_1).$$

(b) If there exists $C \in \mathbb{R}$ such that $\mathbf{W_M}(\mu_0, \mu) \leq C$ for all $\mu \in K$, we can find by Proposition IV.10(a) for each $k \in \mathbb{N}$ a curve $((\mu_k)_t, (w_k)_t)_{t\in[0,1]}$ in $\mathscr{C}_1(\mu_0 \to \mu_k)$ such that $\mathbf{\Phi}((\mu_k)_t, (w_k)_t) \leq C^2$ for a.e. $t \in [0, 1]$ and all $k \in \mathbb{N}$. The requirement of Proposition IV.9 is again fulfilled. Its application yields in particular that $(\mu_k)_t \overset{*}{\rightharpoonup} \mu_t$ (on a subsequence) for all $t \in [0, 1]$ and some $(\mu_t)_{t\in[0,1]}$.
(c) This proof is analogous to the proof of [74, Thm. 5.7] using (a) and (b) of this proposition.
(d) This is a consequence of convexity of the action density ϕ.
(e) This statement can be obtained as in [74, Thm. 5.15].

□

Recall that Proposition IV.10 implies that any pair (μ_0, μ_1) of densities at finite distance — that is, $\mathbf{W_M}(\mu_0, \mu_1) < \infty$ — can be connected by a constant speed minimal geodesic $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$. We shall not enter into the regularity theory for such geodesic curves, which is delicate already in the scalar case [44]. However, it is easy to show that geodesics can be approximated by smooth solutions to the continuity equation (IV.1.5). This will be useful for application of the *Eulerian calculus* in Section IV.3.1 later. To make the statement about smooth approximation precise, we denote the one-dimensional *heat kernel* by $\mathbf{G}_{(\cdot)} : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, it is given by

$$\mathbf{G}_s(y) := \frac{1}{\sqrt{4\pi s}} \exp\left(-\frac{y^2}{4s}\right) \quad \text{for } y \in \mathbb{R} \text{ and } s > 0, \tag{IV.2.12}$$

and it satisfies $\partial_s \mathbf{G}_s = \partial_{xx}\mathbf{G}_s$. Below, convolution $*$ with $\mathbf{G}_s$ is again understood component-wise.

Proposition IV.13 (Smooth approximation of geodesics). *Assume that the mobility $\mathbf{M}$ is induced by a function $h : S \to \mathbb{R}$. Let $\mu_0, \mu_1 \in \mathscr{M}(\mathbb{R}; S)$ be at finite distance $\mathbf{W_M}(\mu_0, \mu_1) < \infty$, and assume that*

$$\lim_{\delta\searrow 0} \delta \int_\mathbb{R} [h(\mu_i) - h(\mathbf{G}_\delta * \mu_i)] \, \mathrm{d}x = 0 \quad \textit{for } i \in \{0, 1\}. \tag{IV.2.13}$$

For a given minimizer $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$ of $\mathbf{E}_1$ and any $\delta > 0$, define the smooth approximation $(\mu^\delta, w^\delta) \in \mathscr{C}_1(\mu_0^\delta \to \mu_1^\delta)$ by $\mu_s^\delta = \mathbf{G}_\delta * \mu_s$ and $w_s^\delta = \mathbf{G}_\delta * w_s$ for each $s \in [0,1]$. Then, for all $\delta > 0$,

$$\mathbf{W_M}(\mu_0^\delta, \mu_1^\delta)^2 \le \mathbf{E}_1(\mu^\delta, w^\delta) \le \mathbf{W_M}(\mu_0, \mu_1)^2, \tag{IV.2.14}$$

and as $\delta \searrow 0$,

$$\mathbf{W_M}(\mu_i^\delta, \mu_i) \searrow 0 \quad \textit{for } i \in \{0,1\}, \textit{ and} \quad \mathbf{W_M}(\mu_0^\delta, \mu_1^\delta) \to \mathbf{W_M}(\mu_0, \mu_1). \tag{IV.2.15}$$

Proof. Our arguments are similar to the proof of [4, Lemma 8.1.9]. The first inequality in (IV.2.14) is obvious from the definition of $\mathbf{W_M}$. To prove the second inequality, recall that the action density function ϕ is convex. Hence, by Jensen's inequality, we have for every $s \in [0,1]$ the pointwise estimate

$$\phi(\mu_s^\delta, w_s^\delta) = \phi(\mathbf{G}_\delta * \mu_s, \mathbf{G}_\delta * w_s) \le \mathbf{G}_\delta * \phi(\mu_s, w_s),$$

and therefore, using that $\mathbf{G}_\delta$ has unit mass,

$$\begin{aligned} \mathbf{E}_1(\mu^\delta, w^\delta) = \int_0^1 \mathbf{\Phi}(\mu_s^\delta, w_s^\delta)\,\mathrm{d}s &\le \int_0^1 \left[\int_{\mathbb{R}} \mathbf{G}_\delta * \phi(\mu_s, w_s)\,\mathrm{d}x\right] \mathrm{d}s \\ &= \int_0^1 \left[\mathbf{\Phi}(\mu_s, w_s) \int_{\mathbb{R}} \mathbf{G}_\delta(z)\,\mathrm{d}z\right] \mathrm{d}s = \mathbf{W_M}(\mu_0, \mu_1)^2. \end{aligned}$$

Next, we prove the first limit in (IV.2.15), for $i = 0$, by estimating $\mathbf{W_M}(\mu_0^\delta, \mu_0)$ in terms of the energy along a particular curve $(\widetilde{\mu}, \widetilde{w}) \in \mathscr{C}_1(\mu_0 \to \mu_0^\delta)$: define for $t \in [0,1]$

$$\widetilde{\mu}_t := \mathbf{G}_{\delta t} * \mu_0, \quad \widetilde{w}_t := -\delta \partial_x(\mathbf{G}_{\delta t} * \mu_0).$$

Notice that $(\widetilde{\mu}, \widetilde{w})$ is indeed a smooth solution to the continuity equation, thanks to the properties of the heat kernel $\mathbf{G}$. For the energy of this particular curve, we obtain with the identities $\partial_x(\nabla_z h(\mathbf{G}_{\delta t} * \mu_0)) = \nabla_z^2 h(\mathbf{G}_{\delta t} * \mu_0)\partial_x(\mathbf{G}_{\delta t} * \mu_0)$ and $(\mathbf{M}(z))^{-1} = \nabla_z^2 h(z)$ that

$$\begin{aligned} \mathbf{E}_1(\widetilde{\mu}, \widetilde{w}) &= \delta^2 \int_{[0,1]\times\mathbb{R}} \partial_x(\mathbf{G}_{\delta t} * \mu_0)^{\mathrm{T}} \big(\mathbf{M}(\mathbf{G}_{\delta t} * \mu_0)\big)^{-1} \partial_x(\mathbf{G}_{\delta t} * \mu_0)\,\mathrm{d}(t,x) \\ &= -\delta^2 \int_{[0,1]\times\mathbb{R}} \nabla_z h(\mathbf{G}_{\delta t} * \mu_0)^{\mathrm{T}} \partial_{xx}(\mathbf{G}_{\delta t} * \mu_0)\,\mathrm{d}(t,x) \\ &= -\delta \int_{[0,1]\times\mathbb{R}} \partial_t h(\mathbf{G}_{\delta t} * \mu_0)\,\mathrm{d}(t,x) = \delta \int_{\mathbb{R}} \left[h(\mu_0) - h(\mathbf{G}_\delta * \mu_0)\right]\,\mathrm{d}x \to 0 \end{aligned}$$

for $\delta \searrow 0$ because of (IV.2.13). The limit for $i = 1$ is obtained in the same way. The remaining limit in (IV.2.15) follows from the above in combination with the triangle inequality: indeed,

$$0 \le \mathbf{W_M}(\mu_0, \mu_1) - \mathbf{W_M}(\mu_0^\delta, \mu_1^\delta) \le \mathbf{W_M}(\mu_0, \mu_0^\delta) + \mathbf{W_M}(\mu_1^\delta, \mu_1)$$

yields $\mathbf{W_M}(\mu_0^\delta, \mu_1^\delta) \to \mathbf{W_M}(\mu_0, \mu_1)$ in the limit $\delta \searrow 0$. □

Remark IV.14 (Compactly supported velocity). *Combining the convolution by $\mathbf{G}_\delta$ with a smooth cut-off, one can define approximations (μ^δ, w^δ) of a given geodesic (μ, w) in such a way that for fixed $\delta > 0$, the velocity fields w_s^δ have compact support in $\mathbb{R}$, uniformly in $s \in [0,1]$.*

Under specialized conditions, an estimate of $\mathbf{W_M}$ in terms of the second moment $\mathfrak{m}_2$ is possible:

Proposition IV.15 (Distance and second moment). *Consider a value space $S \subset \mathbb{R}^n$ of the following form: there exists $S^\ell \in \partial S$ such that $z - S^\ell \ge 0$ (component-wise) for all $z \in S$. Assume that the mobility $\mathbf{M}$ satisfies, in addition to (M0)–(M3), the following Lipschitz-type condition w.r.t. z:*

$$\mathsf{e}^{\mathrm{T}} \mathbf{M}(z) \mathsf{e} \le L \mathsf{e}^{\mathrm{T}}(z - S^\ell), \quad \textit{for all } z \in S, \tag{IV.2.16}$$

some constant $L > 0$ and the vector $\mathsf{e} := (1, 1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^n$. Then, for all $\mu_0, \mu_1 \in \mathscr{M}(\mathbb{R}; S)$, one has

$$\mathfrak{m}_2(\mu_0 - S^\ell) \le e^L \left(\mathfrak{m}_2(\mu_1 - S^\ell) + \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2 \right).$$

Proof. Since the assertion is trivial otherwise, assume that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) < \infty$ and $\mathfrak{m}_2(\mu_1 - S^\ell) < \infty$. Given $R > 0$, let $\theta_R \in C_c^\infty(\mathbb{R})$ with $\theta_R = \mathrm{id}$ on $[-R, R]$, $\theta_R = 0$ on $\mathbb{R} \setminus [-3R, 3R]$ and $|\theta_R'(x)| \le 1$ for all $x \in \mathbb{R}$. Observe that θ_R^2 increases to $x \mapsto x^2$ as $R \nearrow \infty$. Let $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$ be such that $\mathbf{\Phi}(\mu_t, w_t) = \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2$ for almost all $t \in [0, 1]$, by Proposition IV.10. Let $s \in [0, 1]$ be arbitrary. We first obtain that

$$\int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_s - S^\ell) \, \mathrm{d}x - \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_0 - S^\ell) \, \mathrm{d}x = - \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} \partial_x w_t \, \mathrm{d}x \, \mathrm{d}t.$$

Using condition (M1), which yields the existence of a unique symmetric, positive definite square root $\mathbf{M}(z)^{1/2}$ of $\mathbf{M}(z)$, we get

$$\begin{aligned} - \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} \partial_x w_t \, \mathrm{d}x \, \mathrm{d}t &= \int_0^s \int_{\mathbb{R}} 2\theta_R \theta_R' \mathsf{e}^{\mathrm{T}} \mathbf{M}(\mu_t)^{1/2} \mathbf{M}(\mu_t)^{-1/2} w_t \, \mathrm{d}x \, \mathrm{d}t \\ &\le \int_0^s \int_{\mathbb{R}} (\theta_R \theta_R')^2 \mathsf{e}^{\mathrm{T}} \mathbf{M}(\mu_t) \mathsf{e} \, \mathrm{d}x \, \mathrm{d}t + \mathbf{E}_1(\mu, w), \end{aligned}$$

the last step being a consequence of the Cauchy-Schwarz and Young inequalities. Using the Lipschitz-type condition (IV.2.16) and the bound on θ_R', we end up with

$$\int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_s - S^\ell) \, \mathrm{d}x - \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_0 - S^\ell) \, \mathrm{d}x \le L \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_t - S^\ell) \, \mathrm{d}x \, \mathrm{d}t + \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2.$$

Hence, by Gronwall's lemma,

$$\int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_s - S^\ell) \, \mathrm{d}x \le e^{Ls} \left(\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) + \int_{\mathbb{R}} \theta_R^2 \mathsf{e}^{\mathrm{T}} (\mu_0 - S^\ell) \, \mathrm{d}x \right),$$

from which the assertion follows by monotone convergence $R \nearrow \infty$ for $s = 1$. □

IV.2.5. Densities at finite distance

In this section, we derive sufficient conditions under which $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ is finite. Throughout this section, the value space shall be a n-cuboid $S = [S^\ell, S^r]$.

Proposition IV.16 (Bounds on $\mathbf{W}_{\mathbf{M}}$ in terms of $\mathbf{W}_2$). *Let a mobility $\mathbf{M}$ be given and assume that there exists a fully decoupled mobility $\mathbf{M}_0$ as in* (IV.2.5), *where the scalar mobilities $\mathbf{m}_j$ are* uniformly concave, *viz. $\mathbf{m}_j'' \le -\delta$ for some $\delta > 0$, and that the following condition holds:*

There exists $K > 0$ such that $A_K(z) := K\mathbf{M}_0(z)^{-1} - \mathbf{M}(z)^{-1} \in \mathbb{R}^{n \times n}$ is positive definite. (IV.2.17)

Let $\mu_0, \mu_1 \in \mathscr{M}(\mathbb{R}; S)$ with

$$\int_{\mathbb{R}} (\mu_0 - S^\ell) \, \mathrm{d}x = m = \int_{\mathbb{R}} (\mu_1 - S^\ell) \, \mathrm{d}x$$

for some $m \in [0, \infty)^n$ and $\mathfrak{m}_2(\mu_0 - S^\ell), \mathfrak{m}_2(\mu_1 - S^\ell) < \infty$. Then the following statements hold:

(a) $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ is finite; in particular, one has

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \le C[\mathfrak{m}_2(\mu_0 - S^\ell) + \mathfrak{m}_2(\mu_1 - S^\ell)] \tag{IV.2.18}$$

with a constant $C > 0$ depending on m.

(b) If, moreover, for almost every $x \in \mathbb{R}$, one has $\mu_0(x)$, $\mu_1(x) \le \widetilde{S}^r$ for $\widetilde{S}^r \in \mathrm{int}(S)$, then

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \le \widetilde{C} \sum_{j=1}^n \mathbf{W}_2^2(\mu_{0,j} - S_j^\ell, \mu_{1,j} - S_j^\ell), \tag{IV.2.19}$$

with a constant $\widetilde{C} > 0$ depending on m and $\widetilde{S}^r$.

Proof. For every $(\mu, w) \in \mathscr{C}_1(\mu_0 \to \mu_1)$, one has due to condition (IV.2.17) that

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq \mathbf{E}_1(\mu, w) \leq K \sum_{j=1}^{n} \int_0^1 \int_{\mathbb{R}} \frac{w_j^2}{\mathbf{m}_j(\mu_j)} \, \mathrm{d}x \, \mathrm{d}t.$$

Moreover, since the $\mathbf{m}_j$ are uniformly concave, we have

$$\mathbf{m}_j(\mu_j) \geq \frac{\delta}{4}(\mu_j - S_j^\ell)(S_j^r - \mu_j) =: \widetilde{\mathbf{m}}_j(\mu_j),$$

and hence

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq K \sum_{j=1}^{n} \int_0^1 \int_{\mathbb{R}} \frac{w_j^2}{\widetilde{\mathbf{m}}_j(\mu_j)} \, \mathrm{d}x \, \mathrm{d}t.$$

This estimate allows us to consider each component separately, by the same procedure as in the proof of [124, Thm. 3]. □

In the framework of perturbations of fully decoupled mobilities (cf. Section IV.2.2.2) for $n = 2$ components, we are able to give a sufficient criterion such that condition (IV.2.17) is true.

Proposition IV.17 (Estimate on $\mathbf{M}^{-1}$ for two components). *Assume that, for small $\varepsilon > 0$, the mobility $\mathbf{M}$ is of the form*

$$\mathbf{M}(z) = \mathbf{M}_0(z) + \varepsilon \mathbf{M}_\varepsilon(z), \quad \text{where} \quad \mathbf{M}_0(z) := \begin{pmatrix} \mathbf{m}_1(z_1) & 0 \\ 0 & \mathbf{m}_2(z_2) \end{pmatrix},$$

with a fully decoupled mobility $\mathbf{M}_0$. Assume that, in addition to (M0)–(M2), the following conditions are satisfied for some $C > 0$:

(M3'a) $\dfrac{|\mathbf{M}_{\varepsilon,11}(z)|}{\mathbf{m}_1(z_1)} < C,$

(M3'b) $\dfrac{|\mathbf{M}_{\varepsilon,22}(z)|}{\mathbf{m}_2(z_2)} < C,$

(M3'c) $\dfrac{\mathbf{m}_1(z_1)\mathbf{m}_2(z_2)}{\det \mathbf{M}(z)} < C.$

Then, condition (IV.2.17) in Proposition IV.16 holds.

Proof. We use the tr-det criterion on $A_K(z)$ and have that (omitting the argument for the sake of clarity)

$$\operatorname{tr}(A_K) > 0 \iff K > \frac{\mathbf{m}_1 \mathbf{m}_2}{\det \mathbf{M}} \left(1 + \varepsilon \frac{\mathbf{M}_{\varepsilon,11} + \mathbf{M}_{\varepsilon,22}}{\mathbf{m}_1 + \mathbf{m}_2} \right) > 0, \tag{IV.2.20}$$

$$\det(A_K) > 0 \iff K^2 - K \frac{2\mathbf{m}_1\mathbf{m}_2 + \varepsilon \mathbf{M}_{\varepsilon,11}\mathbf{m}_2 + \varepsilon \mathbf{M}_{\varepsilon,22}\mathbf{m}_1}{\det \mathbf{M}} + \frac{\mathbf{m}_1\mathbf{m}_2}{\det \mathbf{M}} > 0. \tag{IV.2.21}$$

Using the assumptions on $\mathbf{M}$, one easily verifies that (IV.2.21) holds if

$$K > \frac{2\mathbf{m}_1\mathbf{m}_2 + \varepsilon \mathbf{M}_{\varepsilon,11}\mathbf{m}_2 + \varepsilon \mathbf{M}_{\varepsilon,22}\mathbf{m}_1}{\det \mathbf{M}} > 0. \tag{IV.2.22}$$

The middle terms in (IV.2.20) and (IV.2.22) are strictly bounded from above by $2C(1 + \varepsilon C)$, where C is the constant in (M3'a)–(M3'c). Hence, choosing $K := 2C(1 + \varepsilon C)$ yields the assertion. □

CHAPTER IV.3

Gradient flows and geodesic convexity

In this chapter, we formally establish criteria for geodesic λ-convexity of entropy functionals $\mathcal{E}$ appearing in (IV.1.1) with respect to the distance $\mathbf{W_M}$. In advance of the main results, we introduce our method of proof by referring to abstract results in the literature adapted to the situation at hand.

IV.3.1. Preliminaries

We first briefly recall the abstract setting developed in [123, 134] from Section I.2.3, which is a variant of the famous "Otto calculus". The goal is to give the metric space $(\mathscr{M}(\mathbb{R};S), \mathbf{W_M})$ a partial Riemannian structure.

A function $\mu \in \mathscr{M}(\mathbb{R};S)$ is called *regular* if μ is smooth and attains values in $\mathrm{int}(S)$ only. Regular functions are dense in $\mathscr{M}(\mathbb{R};S)$ in the following sense: Let $\mu \in \mathscr{M}(\mathbb{R};S)$ be arbitrary, with the only restriction that its values $\mu(x)$ do not lie in a convex subset of the boundary ∂S for a.e. $x \in \mathbb{R}$. Then μ can be approximated by a sequence of regular functions with respect to $\mathbf{W_M}$. This is achieved by standard smoothing techniques; see Proposition IV.13 above.

At a regular μ_0, regular tangent vectors to $\mathscr{M}(\mathbb{R};S)$ are defined by functions $v : \mathbb{R} \to \mathbb{R}^n$ that can be written in the form $v = -\partial_x w$ for some $w \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$: in accordance with the continuity equation (IV.1.5), the associated infinitesimal curve $\mu_{(\cdot)} : (-\varepsilon, \varepsilon) \to \mathscr{M}(\mathbb{R};S)$ is given by $\mu_s = \mu_0 + sv$. Regular tangent vectors are dense in the tangent space $\mathsf{T}_{\mu_0}\mathscr{M}(\mathbb{R};S)$, see Remark IV.14. Regular cotangent vectors are equivalence classes — with respect to additive constants — of functions $\xi \in C^\infty(\mathbb{R};\mathbb{R}^n)$ such that $\partial_x \xi$ has compact support; these lie dense in the corresponding cotangent space $\mathsf{T}^*_{\mu_0}\mathscr{M}(\mathbb{R};\mathbb{R}^n)$. The pairing between v and ξ is given by the scalar product in $L^2(\mathbb{R};\mathbb{R}^n)$:

$$\langle \xi, v \rangle = \int_{\mathbb{R}} \xi(x)^{\mathrm{T}} v(x)\, \mathrm{d}x.$$

The metric structure of $(\mathscr{M}(\mathbb{R};S), \mathbf{W_M})$ distinguishes an injective map $\mathbf{K}$ — the *Onsager operator* — from cotangent to tangent vectors at regular points μ:

$$\mathbf{K}(\mu) : \mathsf{T}^*_\mu\mathscr{M}(\mathbb{R};S) \to \mathsf{T}_\mu\mathscr{M}(\mathbb{R};S), \quad \mathbf{K}(\mu)\xi = -\partial_x(\mathbf{M}(\mu)\partial_x\xi). \tag{IV.3.1}$$

With these notions, we write (IV.1.1) as an *abstract evolution equation*,

$$\partial_t \mu = -\mathbf{F}(\mu), \tag{IV.3.2}$$

with the nonlinear operator $\mathbf{F} : \mathscr{M}(\mathbb{R};S) \to \mathsf{T}\mathscr{M}(\mathbb{R};S)$ given by

$$\mathbf{F}(\mu) := -\partial_x(\mathbf{M}(\mu)\partial_x\mathcal{E}'(\mu)) = \mathbf{K}(\mu)\mathcal{E}'(\mu). \tag{IV.3.3}$$

In the framework of [123, 134], the verification of geodesic λ-convexity of $\mathcal{E}$ with respect to the distance $\mathbf{W_M}$ is based on the *Eulerian calculus* that has originally been developed in [155]; see also [64]. Theorem IV.18 below summarizes the main result of that theory adapted to the framework of this part of the thesis.

We remark that certain hypotheses are implicitly imposed in order to justify the calculations that lead to that result. The main one is that there is a dense subset $\mathscr{M}_0 \subset \mathscr{M}(\mathbb{R};S)$ of regular functions such that (IV.3.2) possesses a smooth classical solution for each initial condition from $\mathscr{M}_0$, and the associated flow maps $\mathsf{S}^t : \mathscr{M}_0 \to \mathscr{M}(\mathbb{R};S)$ are continuous in the topology of $(\mathscr{M}(\mathbb{R};S), \mathbf{W_M})$, for each time $t \geq 0$. It is then one of the consequences of Theorem IV.18 that $\mathsf{S}^{(\cdot)}$ actually extends in a unique way to a continuous flow on all of $\mathscr{M}(\mathbb{R};S)$. Further, one needs to assume that the underlying entropy functional $\mathcal{E} : \mathscr{M}(\mathbb{R};S) \to \mathbb{R}_\infty$ is proper, lower semicontinuous and bounded from below.

The abstract criterion for λ-convexity is the following (compare with Theorem I.19).

Theorem IV.18 (Condition for convexity [123, Thm. 3.6]). *Let $\lambda \in \mathbb{R}$ and let $\mathcal{E}$, $\mathbf{F}$ and $\mathbf{K}$ be defined as in* (IV.3.1)*&*(IV.3.3)*. If*

$$\langle \xi, \mathrm{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi \rangle - \frac{1}{2} \langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle \geq \lambda \langle \xi, \mathbf{K}(\mu)\xi \rangle \tag{IV.3.4}$$

holds for all regular $\mu \in \mathscr{M}(\mathbb{R};S)$ and $\xi \in C^\infty(\mathbb{R};\mathbb{R}^n)$ with $\partial_x\xi$ of compact support, then $\mathsf{S}^{(\cdot)}$ satisfies the evolution variational estimate (I.2.4) *for $\mathcal{E}$ and hence defines a λ-flow on $(\mathscr{M}(\mathbb{R};S), \mathbf{W_M})$. Further, $\mathcal{E}$ is geodesically λ-convex w.r.t. $\mathbf{W_M}$.*

IV.3.2. The multi-component heat equation

In this section, we apply the theory of Section IV.3.1 to the case of the *multi-component heat equation*,

$$\partial_t \mu = \partial_{xx}\mu, \tag{IV.3.5}$$

which is (IV.3.2) for $\mathbf{F}(\mu) = -\partial_{xx}\mu$. In this case, the flow maps $\mathsf{S}^t : \mathscr{M}(\mathbb{R};S) \to \mathscr{M}(\mathbb{R};S)$ are explicitly known:

$$\mathsf{S}^t(\mu^0) = \mathbf{G}_t * \mu^0,$$

with the heat kernel $\mathbf{G}$ from (IV.2.12), for each $t > 0$ and arbitrary initial data $\mu^0 \in \mathscr{M}(\mathbb{R};S)$. Moreover, if μ^0 is a smooth function with values in $\mathrm{int}(S)$ only, then it follows by classical results that the map $(t,x) \mapsto \big(\mathsf{S}^t(\mu^0)\big)(x)$ is also smooth on $[0,\infty) \times \mathbb{R}$, and attains values in $\mathrm{int}(S)$ only. We are thus in the framework described above and conclude the following with the help of Theorem IV.18.

Proposition IV.19 (The heat flow as a gradient flow). *Assume that $\mathbf{M} : S \to \mathbb{R}^{n\times n}$ satisfies* (M0)–(M3)*, and that $\mathbf{M}$ is induced by h as in* (IV.1.3)*, i.e. $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$ at every $z \in \mathrm{int}(S)$, for a continuous function $h : S \to \mathbb{R}$ which is smooth on $\mathrm{int}(S)$. Suppose that for each $\mu_0, \mu_1 \in \mathscr{M}(\mathbb{R};S)$ with $\mathbf{W_M}(\mu_0,\mu_1) < \infty$, condition* (IV.2.13) *is satisfied. Then the flow map $\mathsf{S}^{(\cdot)}$ for* (IV.3.5) *defined above is a 0-flow on $\mathscr{M}(\mathbb{R};S)$, and it is the gradient flow of the functional $\mathcal{H}(\mu) := \int_\mathbb{R} h(\mu)\,\mathrm{d}x$, which is geodesically 0-convex w.r.t. $\mathbf{W_M}$.*

Proof. To begin with, observe that with $\mathcal{H}$ defined as above,

$$\mathbf{M}(\mu)\partial_x\mathcal{H}'(\mu) = \mathbf{M}(\mu)\nabla_z^2 h(\mu)\partial_x\mu = \mathbb{1}\partial_x\mu,$$

which means that (IV.3.2) simplifies to (IV.3.5). We verify (IV.3.4) for $\lambda = 0$: for a given smooth map $w : \mathbb{R} \to \mathbb{R}^n$, the relevant derivative expressions amount to

$$\begin{aligned} \mathrm{D}\mathbf{F}(\mu)[w] &= -\partial_{xx}w, \\ \mathrm{D}\mathbf{K}(\mu)[w]\xi &= -\partial_x(\mathrm{D}\mathbf{M}(\mu)[w]\partial_x\xi). \end{aligned} \tag{IV.3.6}$$

We substitute this into the left-hand side of (IV.3.4) and integrate by parts to obtain

$$\begin{aligned} \langle \xi, \mathrm{D}\mathbf{F}(\mu)[\mathbf{K}(\mu)\xi] \rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle &= \langle \xi, \partial_{xxx}(\mathbf{M}(\mu)\partial_x\xi) \rangle - \frac{1}{2}\langle \xi, \partial_x(\mathrm{D}\mathbf{M}(\mu)[\partial_{xx}\mu]\partial_x\xi) \rangle \\ &= -\frac{1}{2}\left\langle \partial_x\xi, \mathrm{D}^2\mathbf{M}(\mu)[\partial_x\mu, \partial_x\mu]\partial_x\xi \right\rangle + \langle \partial_{xx}\xi, \mathbf{M}(\mu)\partial_{xx}\xi \rangle, \end{aligned}$$

which is nonnegative because of (M1) and (M2). □

IV.3.3. Internal energy functionals

We now study geodesic convexity of more general functionals of the form

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu(x))\,\mathrm{d}x, \tag{IV.3.7}$$

with a smooth function $f : \mathrm{int}(S) \to \mathbb{R}$. For brevity, we call these functionals *internal energies*, regardless of their actual interpretation in physics or other sciences. Our main result is Proposition IV.20 below, which is a further generalization of the generalized McCann condition established by Carrillo *et al.* [53] for *scalar* nonlinear mobilities ($n = 1$), see Theorem I.16.

IV.3.3.1. A generalized McCann condition

The main result of this section is the following sufficient criterion for 0-contractivity of the flow generated by the evolution equation

$$\partial_t \mu = \partial_x\big(\mathbf{L}(\mu)\partial_x\mu\big), \quad \text{with} \quad \mathbf{L}(z) = \mathbf{M}(z)\nabla_z^2 f(z), \tag{IV.3.8}$$

which is (IV.1.1) for $\mathcal{E}$ from (IV.3.7), i.e., the formal gradient flow of $\mathcal{E}$ in $\mathbf{W_M}$.

Proposition IV.20 (Multi-component McCann condition). *Given a mobility matrix* $\mathbf{M}$ *that satisfies* (M0)–(M2) *and a functional* $\mathcal{E}$ *of the form* (IV.3.7), *assume that for all* $z \in \mathrm{int}(S)$ *and all* $v, \zeta, \beta \in \mathbb{R}^n$ *(omitting the argument* z *from* $\mathbf{M} = \mathbf{M}(z)$ *and from* $\mathbf{L} = \mathbf{M}(z)\nabla_z^2 f(z)$*):*

$$\begin{aligned} 0 \leq &-\frac{1}{2} v^{\mathrm{T}}\,\mathrm{D}^2\mathbf{M}[\zeta, \mathbf{L}\zeta]\,v + \beta^{\mathrm{T}}\,\mathbf{L}\mathbf{M}\,\beta \\ &+ \beta^{\mathrm{T}}\big(\mathbf{L}\mathrm{D}\mathbf{M}[\zeta] - \mathrm{D}\mathbf{M}[\mathbf{L}\zeta]\big)v + v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}[\zeta]\big(\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big) - v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}\big[\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big]\,\zeta. \end{aligned} \tag{IV.3.9}$$

Then, under the assumption of sufficient regularity of the associated flow generated by (IV.3.8), *the functional* $\mathcal{E}$ *is geodesically* 0*-convex w.r.t. the distance* $\mathbf{W_M}$.

Proof. This is another application of Theorem IV.18. Let therefore $\mu \in \mathscr{M}(\mathbb{R}; S)$ be regular and $\xi, w \in C^\infty(\mathbb{R}; \mathbb{R}^n)$, $\partial_x\xi$ with compact support. Observe that

$$\begin{aligned} \mathbf{F}(\mu) &= -\partial_x(\mathbf{L}(\mu)\mu_x), \\ \mathrm{D}\mathbf{F}(\mu)[w] &= -\partial_x(\mathrm{D}\mathbf{L}(\mu)[w]\mu_x) - \partial_x(\mathbf{L}(\mu)w_x), \end{aligned}$$

and in addition, (IV.3.6) holds. Hence, integrating by parts, we obtain

$$\begin{aligned} &\langle \xi, \mathrm{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi\rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi\rangle \\ &= -\langle \xi_x, \mathrm{D}\mathbf{L}(\mu)[\partial_x(\mathbf{M}(\mu)\xi_x)]\mu_x\rangle + \langle \xi_{xx}, \mathbf{L}(\mu)\partial_x(\mathbf{M}(\mu)\xi_x)\rangle + \langle \xi_{xx}, \mathrm{D}\mathbf{L}(\mu)[\mu_x]\partial_x(\mathbf{M}(\mu)\xi_x)\rangle \\ &\quad - \frac{1}{2}\langle \xi_{xx}, \mathrm{D}\mathbf{M}(\mu)[\mathbf{L}(\mu)\mu_x]\xi_x\rangle - \frac{1}{2}\langle \xi_x, \mathrm{D}\mathbf{M}(\mu)[\mathbf{L}(\mu)\mu_x]\xi_{xx}\rangle - \frac{1}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{L}(\mu)\mu_x]\xi_x\right\rangle \\ &= -\frac{1}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{L}(\mu)\mu_x]\xi_x\right\rangle - \langle \xi_{xx}, \mathrm{D}\mathbf{M}(\mu)[\mathbf{L}(\mu)\mu_x]\xi_x\rangle \\ &\quad - \langle \xi_x, \mathrm{D}\mathbf{L}(\mu)[\mathrm{D}\mathbf{M}(\mu)[\mu_x]\xi_x + \mathbf{M}(\mu)\xi_{xx}]\mu_x\rangle + \langle \xi_{xx}, \mathbf{L}(\mu)(\mathrm{D}\mathbf{M}(\mu)[\mu_x]\xi_x + \mathbf{M}(\mu)\xi_{xx})\rangle \\ &\quad + \langle \xi_{xx}, \mathrm{D}\mathbf{L}(\mu)[\mu_x](\mathrm{D}\mathbf{M}(\mu)[\mu_x]\xi_x + \mathbf{M}(\mu)\xi_{xx})\rangle\,. \end{aligned}$$

Condition (IV.3.9) now implies pointwise nonnegativity (substitute $v := \xi_x(x)$, $\beta := \xi_{xx}(x)$, $\zeta := \mu_x(x)$ for $x \in \mathbb{R}$) and consequently (IV.3.4) for $\lambda = 0$. □

Remark IV.21 (Diagonal mobility). *In the case of a fully decoupled mobility matrix*

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & 0 \\ 0 & \mathbf{m}_2(z_2) \end{pmatrix}$$

for $n = 2$ components, where in general $\frac{1}{2}\mathbf{m}_j''\mathbf{m}_j + (\mathbf{m}_j')^2 \neq 0$, the generalized McCann condition (IV.3.9) *is equivalent to*

$$\partial_{11}f(z) \geq 0, \quad \partial_{22}f(z) \geq 0, \quad \partial_{12}f(z) = 0,$$

since (IV.3.9) *reads in this case*

$$\begin{aligned} 0 \geq &\left[\frac{1}{2}v_1^2\zeta_1^2\mathbf{m}_1''\mathbf{m}_1 - \beta_1^2\mathbf{m}_1^2\right]\partial_{11}f + \left[\frac{1}{2}v_2^2\zeta_2^2\mathbf{m}_2''\mathbf{m}_2 - \beta_2^2\mathbf{m}_2^2\right]\partial_{22}f \\ &+ \left[\frac{1}{2}v_1^2\zeta_1\zeta_2\mathbf{m}_1''\mathbf{m}_1 + \frac{1}{2}v_2^2\zeta_1\zeta_2\mathbf{m}_2''\mathbf{m}_2 - 2\beta_1\beta_2\mathbf{m}_1\mathbf{m}_2 - 2v_1v_2\zeta_1\zeta_2\mathbf{m}_1'\mathbf{m}_2'\right]\partial_{12}f \\ &+ \Big[[v_1^2(\mathbf{m}_1')^2 + v_2^2(\mathbf{m}_2')^2]\zeta_1\zeta_2 + 2v_1\beta_1\zeta_2\mathbf{m}_1'\mathbf{m}_1 + 2v_2\beta_2\zeta_1\mathbf{m}_2'\mathbf{m}_2 - 2v_2\beta_1\zeta_2\mathbf{m}_2'\mathbf{m}_1 - 2v_1\beta_2\zeta_1\mathbf{m}_1'\mathbf{m}_2\Big]\partial_{12}f. \end{aligned}$$

Imposing e.g. $\beta = 0$, $v_1 = 1$, $v_2 = 0$ and $\zeta_1 = 1$, one obtains

$$0 \geq \frac{1}{2}\mathbf{m}_1''\mathbf{m}_1\partial_{11}f + \left[\frac{1}{2}\mathbf{m}_1''\mathbf{m}_1 + (\mathbf{m}_1')^2\right]\zeta_2\partial_{12}f,$$

from which necessarily $\partial_{12}f(\mu) = 0$ follows. Hence, the only possible choice is $f(z) := \psi_1(z_1) + \psi_2(z_2)$ with convex functions ψ_1, ψ_2. We solely recover the generalized McCann condition for $n = 1$ (cf. [53]) for each of the two components separately if $\mathbf{M}$ is fully decoupled.

IV.3.3.2. Perturbation results and examples

This paragraph is devoted to examples satisfying condition (IV.3.9) of Proposition IV.20. In particular, we investigate suitable perturbations of the energies having the heat flow as gradient flow, cf. Proposition IV.19. We first start with a more general result involving perturbations of compact support in $\operatorname{int}(S)$ and continue with a specific example where the support of the perturbation extends to all of S.

Proposition IV.22 (Perturbations of compact support). *Let a mobility $\mathbf{M}$ satisfy the conditions* (M0)–(M3) *and the stronger condition* (M2′) *and be induced by h as in* (IV.1.3)*. For $\alpha, \widetilde{\varepsilon} > 0$ and $g \in C_c^\infty(\operatorname{int}(S))$, define $f(z) := \alpha h(z) + \widetilde{\varepsilon}g(z)$ and $\mathcal{E}$ according to* (IV.3.7)*. Then, for $\widetilde{\varepsilon} > 0$ sufficiently small, the generalized McCann condition* (IV.3.9) *is satisfied.*

Proof. If $z \notin \operatorname{supp} g$, the conditions (M1) and (M2′) directly yield the claim. Furthermore, there exists a constant $\delta_g > 0$ such that for all $z \in \operatorname{supp} g$, one has

$$\begin{aligned} \beta^{\mathrm{T}}\mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta]\beta &\leq -\delta_g|\beta|^2|\zeta|^2, \\ -\gamma^{\mathrm{T}}\mathbf{M}(z)\gamma &\leq -\delta_g|\gamma|^2, \end{aligned}$$

for all $\beta, \gamma, \zeta \in \mathbb{R}^n$. Hence, by continuity, we obtain for the r.h.s. in (IV.3.9), recalling

$$\mathbf{L}(z) = \mathbf{M}(z)\nabla_z^2 f(z) = \alpha\mathbb{1} + \widetilde{\varepsilon}\nabla_z^2 g(z):$$

$$\begin{aligned} &-\frac{1}{2}v^{\mathrm{T}}\,\mathrm{D}^2\mathbf{M}[\zeta, \mathbf{L}\zeta]\,v + \beta^{\mathrm{T}}\,\mathbf{L}\mathbf{M}\,\beta \\ &\qquad + \beta^{\mathrm{T}}\big(\mathbf{L}\mathrm{D}\mathbf{M}[\zeta] - \mathrm{D}\mathbf{M}[\mathbf{L}\zeta]\big)v + v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}[\zeta]\big(\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big) - v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}\big[\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big]\,\zeta \\ &\geq \frac{\alpha}{2}\delta_g|\zeta|^2|v|^2 + \alpha\delta_g|\beta|^2 - \widetilde{\varepsilon}C_{g,\mathbf{M}}(|\zeta|^2|v|^2 + |\beta|^2 + |\zeta||v||\beta|), \end{aligned}$$

with a constant $C_{g,\mathbf{M}} > 0$. Using Young's inequality, one immediately deduces that the r.h.s. is nonnegative and thus (IV.3.9) is satisfied, provided that $\widetilde{\varepsilon} \leq \frac{\alpha\delta_g}{3C_{g,\mathbf{M}}}$. □

We conclude this section with a specific example such that the support of the perturbation g extends to all of S.

Example IV.23 (Non-compactly supported perturbations). *Let* $\mathbf{M}$ *be induced by* h_ε *from* (IV.2.6)&(IV.2.7)*:*

$$\begin{aligned} h_\varepsilon(z) &:= z_1\log(z_1) + (1-z_1)\log(1-z_1) + z_2\log(z_2) + (1-z_2)\log(1-z_2) + \varepsilon d_1 d_2, \\ d_j &:= z_j(1-z_j), \end{aligned}$$

and $\varepsilon > 0$ *chosen so small that the conditions* (M0)–(M3) *and the stronger condition* (M2′) *are satisfied. Define furthermore* $\widetilde{g} : [0, \frac{1}{4}]^2 \to \mathbb{R}$ *by*

$$\widetilde{g}(m_1, m_2) := \exp\left(-\frac{1}{m_1} - \frac{1}{m_2}\right)$$

for all $0 < m_1, m_2 \le \frac{1}{4}$*, and* $\widetilde{g}(m_1, 0) = 0 = \widetilde{g}(0, m_2)$*. Consider now for* $\widetilde{\varepsilon} > 0$ *the map* $f(z) := h(z) + \widetilde{\varepsilon}\widetilde{g}(d_1, d_2)$ *and the functional* $\mathcal{E}$ *according to* (IV.3.7)*. Then, for* $\widetilde{\varepsilon} > 0$ *sufficiently small, the generalized McCann condition* (IV.3.9) *is satisfied.*

Our idea of proof relies on the structure of $\mathbf{M}$ in this particular case (cf. Section IV.2.2.2): There exists a positive rational function $r_1 : \left(0, \frac{1}{4}\right]^2 \to (0,\infty)$ with $\lim_{\widetilde{m}\searrow 0} r_1(m_1, \widetilde{m}) = 0 = \lim_{\widetilde{m}\searrow 0} r_1(\widetilde{m}, m_2)$ for all $(m_1, m_2) \in \left(0, \frac{1}{4}\right]^2$, such that

$$\frac{1}{2}\beta^{\mathrm{T}} \mathrm{D}^2\mathbf{M}(z)[\zeta,\zeta]\beta - \gamma^{\mathrm{T}}\mathbf{M}(z)\gamma \le -r_1(d_1,d_2)(|\zeta|^2|\beta|^2 + |\gamma|^2).$$

Furthermore, there exists another rational function $r_2 : \left(0, \frac{1}{4}\right]^2 \to [0,\infty)$ such that the following estimate on the r.h.s. in condition (IV.3.9) is possible:

$$\begin{aligned} &-\frac{1}{2}v^{\mathrm{T}}\,\mathrm{D}^2\mathbf{M}[\zeta, \mathbf{L}\zeta]\,v + \beta^{\mathrm{T}}\,\mathbf{LM}\,\beta \\ &\quad + \beta^{\mathrm{T}}\big(\mathbf{L}\mathrm{D}\mathbf{M}[\zeta] - \mathrm{D}\mathbf{M}[\mathbf{L}\zeta]\big)v + v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}[\zeta]\big(\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big) - v^{\mathrm{T}}\,\mathrm{D}\mathbf{L}\big[\mathrm{D}\mathbf{M}[\zeta]v + \mathbf{M}\beta\big]\,\zeta \\ &\qquad \ge (r_1(d_1,d_2) - \widetilde{g}(d_1,d_2)r_2(d_1,d_2))(|\zeta|^2|\beta|^2 + |\gamma|^2). \end{aligned}$$

Since for all $(m_1, m_2) \in \left(0, \frac{1}{4}\right]^2$, one has

$$\lim_{\widetilde{m}\searrow 0} \widetilde{g}(\widetilde{m}, m_2)\frac{r_2(\widetilde{m}, m_2)}{r_1(\widetilde{m}, m_2)} = 0 = \lim_{\widetilde{m}\searrow 0} \widetilde{g}(m_1, \widetilde{m})\frac{r_2(m_1, \widetilde{m})}{r_1(m_1, \widetilde{m})},$$

we find $\widetilde{\varepsilon}_0 > 0$ sufficiently small such that (IV.3.9) holds for all $0 < \widetilde{\varepsilon} \le \widetilde{\varepsilon}_0$.

IV.3.4. The potential energy

In this section, we study the geodesic λ-convexity of the regularized *potential energy* functional

$$\mathcal{V}(\mu) = \int_{\mathbb{R}} \left[\alpha h(\mu) + \rho(x)^{\mathrm{T}}\mu\right]\,\mathrm{d}x, \tag{IV.3.10}$$

which has a density depending explicitly on the spatial variable x via the potential ρ. Here, $\mathbf{W}_{\mathbf{M}}$ and h are as in Proposition IV.19 and $\alpha > 0$ as well as $\rho \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$ are fixed. The flow associated to $\mathcal{V}$ is generated by the following (regularized) nonlinear *transport equation*:

$$\partial_t\mu = \alpha\partial_{xx}\mu + \partial_x(\mathbf{M}(\mu)\partial_x\rho). \tag{IV.3.11}$$

IV.3.4.1. Convexity

A sufficient condition on convexity of those energies is the following:

Proposition IV.24 (Convexity for the regularized potential energy functional). *Let* $\mathcal{V}$ *be of the form* (IV.3.10) *with* h*,* α *and* ρ *as mentioned above, let* $\mathbf{M} = (\nabla_z^2 h)^{-1}$ *be as in Proposition IV.19 and* $\lambda \in \mathbb{R}$ *be fixed. If for all*

$z \in \mathrm{int}(S)$ and all $v, \zeta \in \mathbb{R}^n$, $q^1, q^2 \in \overline{\mathbb{B}_R(0)}$ with $R := \|\rho\|_{C^2}$, the condition

$$0 \leq -\frac{\alpha}{2} v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta,\zeta] v - \lambda v^{\mathrm{T}} \mathbf{M} v - \frac{1}{2} v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta, \mathbf{M} q^1] v + v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta, \mathbf{M} v] q^1 + v^{\mathrm{T}} \mathrm{D}\mathbf{M}[\mathbf{M} q^2] v \tag{IV.3.12}$$

is satisfied, then $\mathcal{V}$ is geodesically λ-convex w.r.t. the distance $\mathbf{W}_{\mathbf{M}}$ under the assumption of sufficient regularity of the flow generated by (IV.3.11).

Proof. The method of proof is similar to that of Proposition IV.20. Here, one gets

$$\mathbf{F}(\mu) = -\alpha\partial_{xx}\mu - \partial_x(\mathbf{M}(\mu)\rho_x),$$
$$\mathrm{D}\mathbf{F}(\mu)[w] = -\alpha\partial_{xx}w - \partial_x(\mathrm{D}\mathbf{M}(\mu)[w]\rho_x).$$

Consequently, performing essentially the same calculations as in the proofs of the Propositions IV.19 and IV.20,

$$\begin{aligned}
&-\frac{1}{2}\langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi\rangle + \langle \xi, \mathrm{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi\rangle - \lambda\langle \xi, \mathbf{K}(\mu)\xi\rangle \\
&= -\frac{\alpha}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x,\mu_x]\xi_x\right\rangle - \frac{1}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\rho_x]\xi_x\right\rangle \\
&\quad + \alpha\langle \xi_{xx}, \mathbf{M}(\mu)\xi_{xx}\rangle + \left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\xi_x]\rho_x\right\rangle + \langle \xi_x, \mathrm{D}\mathbf{M}(\mu)[\mathbf{M}(\mu)\xi_x]\rho_{xx}\rangle \\
&\quad - \lambda\langle \xi_x, \mathbf{M}(\mu)\xi_x\rangle .
\end{aligned}$$

We use the fact that for all $\gamma, q^1, v \in \mathbb{R}^n$ and all $z \in \mathrm{int}(S)$, one has due to symmetry of the third-order tensor $\mathrm{D}^3 h$:

$$\gamma^{\mathrm{T}}\mathrm{D}\mathbf{M}(z)[\mathbf{M}(\mu)q^1]v = -\mathrm{D}^3h(z)[\mathbf{M}(z)\gamma, \mathbf{M}(z)q^1, \mathbf{M}(z)v] = \gamma^{\mathrm{T}}\mathrm{D}\mathbf{M}(z)[\mathbf{M}(z)v]q^1.$$

Hence, we obtain

$$\begin{aligned}
&\langle \xi, \mathrm{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi\rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi\rangle - \lambda\langle \xi, \mathbf{K}(\mu)\xi\rangle \\
&= -\frac{\alpha}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x,\mu_x]\xi_x\right\rangle + \alpha\langle \xi_{xx}, \mathbf{M}(\mu)\xi_{xx}\rangle - \lambda\langle \xi_x, \mathbf{M}(\mu)\xi_x\rangle \\
&\quad - \frac{1}{2}\left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\rho_x]\xi_x\right\rangle + \left\langle \xi_x, \mathrm{D}^2\mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\xi_x]\rho_x\right\rangle + \langle \xi_x, \mathrm{D}\mathbf{M}(\mu)[\mathbf{M}(\mu)\rho_{xx}]\xi_x\rangle ,
\end{aligned}$$

which is nonnegative due to condition (IV.3.12) and (M1) (substitute $v := \xi_x(x)$, $\zeta := \mu_x(x)$, $q^1 := \rho_x(x)$, $q^2 := \rho_{xx}(x)$ for $x \in \mathbb{R}$) and hence implies (IV.3.4). □

IV.3.4.2. The case of fully decoupled mobility

In this paragraph, we consider the case of a fully decoupled mobility (cf. Section IV.2.2.1)

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & \\ & \ddots & \\ & & \mathbf{m}_n(z_n) \end{pmatrix},$$

on the n-cuboid $S = [S^\ell, S^r]$. We shall assume that the scalar mobilities $\mathbf{m}_j$ are such that

- $\mathbf{m}_j \in C^2([S_j^\ell, S_j^r])$;
- $\mathbf{m}_j(s) > 0$ for $s \in (S_j^\ell, S_j^r)$ and $\mathbf{m}_j(S_j^\ell) = \mathbf{m}_j(S_j^r) = 0$;
- $\mathbf{m}_j''(s) \leq 0$ for $s \in [S_j^\ell, S_j^r]$.

Recall that $\mathbf{M}$ is of the special form (IV.1.3) $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$, where

$$h(z) = \sum_{j=1}^{n} h_j(z_j),$$

h_j being a second primitive of $\frac{1}{\mathbf{m}_j}$.

Proposition IV.25 (λ-convexity of the potential energy). *For a fully decoupled mobility $\mathbf{M}$ as mentioned above, fix $\alpha > 0$ and $\rho \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$ and consider the regularized potential energy functional $\mathcal{V}$ defined in* (IV.3.10).

(a) Let $\overline{z} \in S$ and $\mu^0 \in \mathscr{M}(\mathbb{R};S)$ such that $\mu^0 - \overline{z} \in H^1(\mathbb{R};\mathbb{R}^n)$ and such that μ^0 attains values in int(S) *only. Then, the initial-value problem for* (IV.3.11)

$$\partial_t \mu = \alpha \partial_{xx}\mu + \partial_x(\mathbf{M}(\mu)\partial_x \rho), \qquad \mu(0,\cdot) = \mu^0, \tag{IV.3.13}$$

possesses a unique local-in-time classical solution $\mu\colon [0,T] \to \mathscr{M}(\mathbb{R};S)$ with $\mu - \overline{z} \in C^0([0,T];H^1(\mathbb{R};\mathbb{R}^n))$, where $T = T(\mu^0,\rho) > 0$.

(b) There exists $C = C(\rho) > 0$ such that condition (IV.3.12) *in Proposition IV.24 is satisfied for all $\lambda \le -C(\frac{1}{\alpha}+1)$.*

Hence, Proposition IV.24 is applicable and yields geodesic λ-convexity of the potential energy $\mathcal{V}$.

Proof. (a) Since in the case at hand, the system (IV.3.11) is decoupled, it suffices to prove the assertion in the scalar case $n = 1$, where the mobility $\mathbf{m}$ is a scalar function satisfying the properties of Section IV.3.4.2. Suppose that $\mu^0 \in \mathscr{M}(\mathbb{R};S)$ attains values in int(S) only, with $S = [S^\ell, S^r] \subset \mathbb{R}$ being an interval, and $\mu^0 - \overline{z} \in H^1(\mathbb{R})$ for some $\overline{z} \in S$. Using the transformation $u := \mu - \overline{z}$ and writing $\theta := \rho_x$, we may instead consider the equation

$$\partial_t u = \partial_{xx} u + \partial_x(\mathbf{m}(u+\overline{z})\theta), \tag{IV.3.14}$$

together with the initial condition $u^0 := \mu^0 - \overline{z} \in H^1(\mathbb{R})$ with values in $(S^\ell - \overline{z}, S^r - \overline{z}) \ni 0$.

Inspired from [95, Ch. 3], we write (IV.3.14) as an abstract semilinear evolution equation on $H^1(\mathbb{R})$:

$$\dot{u}(t) = -Au(t) + F(u(t)), \tag{IV.3.15}$$

with $A := -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$, and

$$F(u) := \mathbf{m}'(u+\overline{z})u_x\theta + \mathbf{m}(u+\overline{z})\theta_x.$$

We first prove some properties of the nonlinearity F.

Lemma IV.26 (Properties of F). *(a) F maps bounded subsets of $H^1(\mathbb{R})$ onto bounded subsets of $L^2(\mathbb{R})$, because for all $u \in H^1(\mathbb{R})$, one has*

$$\|F(u)\|_{L^2} \le C_0\|u\|_{H^1} + C_1, \tag{IV.3.16}$$

for some $C_0, C_1 > 0$.

(b) F is locally Lipschitz continuous in the following sense: if $u,v \in H^1(\mathbb{R})$ with $\|u-u^0\|_{H^1} < \delta$ and $\|v-u^0\|_{H^1} < \delta$ for some $\delta > 0$, then

$$\|F(u)-F(v)\|_{L^2} \le C_2\|u-v\|_{H^1}, \tag{IV.3.17}$$

for some $C_2 = C_2(\delta,u^0) > 0$.

Proof. (a) By the triangle inequality, we have

$$\|F(u)\|_{L^2} \le \|\mathbf{m}\|_{C^1}\|\theta\|_{C^0}\|u_x\|_{L^2} + \|[\mathbf{m}(u+\overline{z}) - \mathbf{m}(\overline{z})]\theta_x\|_{L^2} + \|\mathbf{m}(\overline{z})\theta_x\|_{L^2}$$

$$\leq 2\|\mathbf{m}\|_{C^1}\|\theta\|_{C^1}\|u\|_{H^1} + \|\mathbf{m}(\bar{z})\theta_x\|_{L^2},$$

from which the desired estimate follows since θ has compact support.

(b) With u and v as required, one has

$$\begin{aligned}\|F(u) - F(v)\|_{L^2} &\leq \|\theta\|_{C^1}\|\mathbf{m}(u+\bar{z}) - \mathbf{m}(v+\bar{z})\|_{L^2} + \|\theta\mathbf{m}'(u+\bar{z})[u_x - v_x]\|_{L^2} \\ &\quad + \|\theta u_x^0[\mathbf{m}'(u+\bar{z}) - \mathbf{m}'(v+\bar{z})]\|_{L^2} + \|\theta(v_x - u_x^0)[\mathbf{m}'(u+\bar{z}) - \mathbf{m}'(v+\bar{z})]\|_{L^2} \\ &\leq \|\theta\|_{C^1}\|\mathbf{m}\|_{C^2}\left[\|u-v\|_{L^2} + \|u_x - v_x\|_{L^2} + (\|u^0\|_{H^1} + \|v - u^0\|_{H^1})\|u-v\|_{L^\infty}\right].\end{aligned}$$

Since $H^1(\mathbb{R})$ is continuously embedded into $C^0(\mathbb{R})$ and $\|v - u^0\|_{H^1} < \delta$, the desired estimate follows.

□

Let now $\delta > 0$ fixed, but arbitrary and define

$$K_\delta := \{u \in C^0([0,T]; H^1(\mathbb{R})) \mid \|u(t) - u^0\|_{H^1} \leq \delta \quad \forall t \in [0,T]\},$$

where $T > 0$ is to be determined later. K_δ is a closed subset of the Banach space $C^0([0,T]; H^1(\mathbb{R}))$. Define a mapping B on K_δ by

$$B(u)(t) := e^{-At}u^0 + \int_0^t e^{-A(t-s)}F(u(s))\,\mathrm{d}s \qquad \text{for } t \in [0,T].$$

We prove the following statement:

Lemma IV.27 (B is a contraction). *There exists $T = T(\delta, u^0) > 0$ sufficiently small such that B maps K_δ into itself and is a strict contraction.*

Proof. We first prove that $\|B(u)(t) - u^0\|_{H^1} \leq \delta$ for all $t \in [0,T]$, where T is sufficiently small. For all $s \in (0,t)$, one has

$$e^{-A(t-s)}F(u(s)) = \mathbf{G}_{t-s} * F(u(s)),$$

where $\mathbf{G}$ is the one-dimensional heat kernel from (IV.2.12). Note that for all $\sigma > 0$, we have

$$\|\mathbf{G}_\sigma\|_{L^1} = A_0, \qquad \|\partial_y \mathbf{G}_\sigma\|_{L^1} = A_1\sigma^{-1/2}, \tag{IV.3.18}$$

for some constants $A_0, A_1 > 0$. Elementary kernel estimates yield

$$\|\mathbf{G}_t * u^0 - u^0\|_{H^1} \leq \frac{\delta}{2}, \tag{IV.3.19}$$

for all $t \in [0,T]$, provided that T is sufficiently small. For the other part, we use Young's inequality for convolutions to obtain

$$\left\|\int_0^t e^{-A(t-s)}F(u(s))\,\mathrm{d}s\right\|_{H^1} \leq \int_0^t \left[\|\mathbf{G}_{t-s}\|_{L^1} + \|\partial_y \mathbf{G}_{t-s}\|_{L^1}\right]\|F(u(s))\|_{L^2}\,\mathrm{d}s$$

Using (IV.3.18) and (IV.3.16), together with the fact that $\|u(s)\|_{H^1} \leq \|u^0\|_{H^1} + \delta$ (since $u \in K_\delta$), yields

$$\left\|\int_0^t e^{-A(t-s)}F(u(s))\,\mathrm{d}s\right\|_{H^1} \leq (tA_0 + 2\sqrt{t}A_1)(C_0\|u^0\|_{H^1} + C_0\delta + C_1) \leq \frac{\delta}{2}, \tag{IV.3.20}$$

for all $t \in [0,T]$, provided that T is sufficiently small. Putting (IV.3.19)&(IV.3.20) together yields the claim. Along the same lines, it can be shown that $B(u) \in C^0([0,T]; H^1(\mathbb{R}))$; hence $B(u) \in K_\delta$.

For Lipschitz continuity, we proceed exactly as before using (IV.3.17) instead:

$$\|B(u)(t) - B(v)(t)\|_{H^1} \leq C_2\int_0^t (A_0 + (t-s)^{-1/2}A_1)\|u(t-s) - v(t-s)\|_{H^1}\,\mathrm{d}s$$

$$\leq C_2(A_0T + 2\sqrt{T}A_1)\|u - v\|_{C^0([0,T];H^1)},$$

for all $t \in [0,T]$. Hence, if T is sufficiently small, one has

$$\|B(u) - B(v)\|_{C^0([0,T];H^1)} \leq L\|u - v\|_{C^0([0,T];H^1)},$$

for some $0 \leq L < 1$. □

Now, by Banach's fixed point theorem, B possesses exactly one fixed point u^* in K_δ which is, by means of [95, Lemma 3.3.2], the desired unique smooth solution to (IV.3.15) on $[0,T]$. It remains to prove that $u^*(t,x) \in \operatorname{int}(S)$ for all $x \in \mathbb{R}$ and $t \in [0,T']$, for some sufficiently small $T' > 0$.

Case 1: $\bar{z} \in \operatorname{int}(S)$. Thanks to $u^0 \in H^1(\mathbb{R})$, there exists $\delta_0 > 0$ such that

$$\operatorname{dist}(u^0(x), \partial S) > \delta_0 \quad \forall x \in \mathbb{R}.$$

Since $u^* \in C^0([0,T];H^1(\mathbb{R})) \subset C^0([0,T];C^0(\mathbb{R}))$, there exists $T' \in (0,T]$ such that for all $t \in [0,T']$, one has $\|u^*(t,\cdot) - u^0\|_{C^0} < \frac{\delta_0}{2}$. Hence, we obtain

$$\operatorname{dist}(u^*(t,x), \partial S) > \frac{\delta_0}{2} \quad \forall t \in [0,T'],\ x \in \mathbb{R},$$

which proves the claim.

Case 2: $\bar{z} = S^\ell$. First, as in case 1, there exists $T_1' \in (0,T]$ such that $u^*(t,x) < S^r - S^\ell$ for all $x \in \mathbb{R}$ and all $t \in [0,T_1']$. It remains to prove the lower bound $u^*(t,x) > 0$. Let therefore $R > 0$ such that $\operatorname{supp}(\theta) \subset [-R,R]$. Since u^0 is strictly positive and continuous, there exists $\delta > 0$ such that $u^0(x) > \delta$ for all $x \in [-(R+1), R+1]$. Hence, we can find $T_2' \in (0,T_1']$ such that $u^*(t,x) > \frac{\delta}{2}$ for all $t \in [0,T_2']$ and all $x \in [-(R+1), R+1]$. Moreover, thanks to the smoothness of u^*, there exists $C_0 > 0$ such that $|F(u^*(s))(y)| < C_0$ for all $s \in [0,T_2']$ and all $y \in [-R,R]$.

It remains to consider the case $|x| > R+1$, $t \in [0,T_2']$, where we explicitly analyse u^* by means of its fixed-point property $B(u^*) = u^*$, i.e.

$$u^*(t,x) = \int_{\mathbb{R}} \mathbf{G}_t(x-y)u^0(y)\,\mathrm{d}y + \int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y)F(u^*(s))(y)\,\mathrm{d}y\,\mathrm{d}s. \tag{IV.3.21}$$

For the second part on the right-hand side of formula (IV.3.21), we immediately obtain the estimate

$$\left|\int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y)F(u^*(s))(y)\,\mathrm{d}y\,\mathrm{d}s\right| \leq C_0 \int_{-R}^R \int_0^t \mathbf{G}_s(x-y)\,\mathrm{d}s\,\mathrm{d}y,$$

where we recall that $|x-y| > 1$. Since for fixed $v > 0$, the map

$$g_v : (0,\infty) \to \mathbb{R}, \qquad g_v(s) := \frac{1}{\sqrt{4\pi s}} \exp\left(-\frac{v^2}{4s}\right)$$

is strictly increasing for $s < \frac{v^2}{2}$, we obtain

$$\left|\int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y)F(u^*(s))(y)\,\mathrm{d}y\,\mathrm{d}s\right| \leq C_0 \int_{-R}^R t\mathbf{G}_t(x-y)\,\mathrm{d}y,$$

if $t < \frac{1}{2}$. Hence, for all $t < \min\left(T_2', \frac{1}{2}, \frac{\delta}{2C_0}\right) =: T'$ and all $|x| > R+1$, formula (IV.3.21) yields

$$u^*(t,x) > \int_{-R}^R \left(\frac{\delta}{2} - C_0 t\right) \mathbf{G}_t(x-y)\,\mathrm{d}y,$$

the right-hand side being nonnegative.

Case 3: $\bar{z} = S^r$. Here, argue in analogy to case 2.

(b) We proceed similarly to [125] and observe that for all $z \in \operatorname{int}(S)$ and all $v, \zeta \in \mathbb{R}^n$, $q^1, q^2 \in \overline{\mathbb{B}_R(0)}$, $R := \|\rho\|_{C^2}$, one has

$$\begin{aligned}
&-\frac{\alpha}{2} v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta, \zeta] v - \frac{1}{2} v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta, \mathbf{M} q^1] v + v^{\mathrm{T}} \mathrm{D}^2 \mathbf{M}[\zeta, \mathbf{M} v] q^1 + v^{\mathrm{T}} \mathrm{D} \mathbf{M}[\mathbf{M} q^2] v \\
&= \sum_{j=1}^n \left[-\frac{\alpha}{2} \mathbf{m}_j''(z_j) \zeta_j^2 + \frac{1}{2} \mathbf{m}_j''(z_j) \mathbf{m}_j(z_j) q_j^1 \zeta_j + \mathbf{m}_j'(z_j) \mathbf{m}_j(z_j) q_j^2 \right] v_j^2 \\
&\geq \sum_{j=1}^n \left[-\frac{1}{8\alpha} |\mathbf{m}_j''(z_j)| \mathbf{m}_j(z_j)^2 (q_j^1)^2 + \mathbf{m}_j'(z_j) \mathbf{m}_j(z_j) q_j^2 \right] v_j^2,
\end{aligned}$$

the last step being a consequence of Young's inequality. Using the bounds on $\mathbf{m}_j$, q^1 and q^2, we obtain

$$\sum_{j=1}^n \left[-\frac{1}{8\alpha} |\mathbf{m}_j''(z_j)| \mathbf{m}_j(z_j) (q_j^1)^2 + \mathbf{m}_j'(z_j) q_j^2 \right] \mathbf{m}_j(z_j) v_j^2 \geq -\sum_{j=1}^n \|\mathbf{m}_j\|_{C^2} R \left[\frac{\|\mathbf{m}_j\|_{C^2} R}{8\alpha} + 1 \right] \mathbf{m}_j(z_j) v_j^2.$$

Obviously, for all $\lambda \leq -\max_j \|\mathbf{m}_j\|_{C^2} R \left[\frac{\|\mathbf{m}_j\|_{C^2} R}{8\alpha} + 1 \right]$, (IV.3.12) holds.

□

CHAPTER IV.4

Existence of weak solutions

In this chapter, we prove the existence of weak solutions for a class of initial-value problems of the form (IV.1.1). More specifically, we consider the case of a fully decoupled mobility $\mathbf{M}$ but allow for coupling inside the energy $\mathcal{E}$. Note that, by Remark IV.21, the functional $\mathcal{E}$ will in general *not* be geodesically λ-convex.

IV.4.1. Setting and preliminaries

We again consider as value space an n-cuboid $S = [S^\ell, S^r] \subset \mathbb{R}^n$ and let $h : S \to \mathbb{R}$, $h(z) = \sum_{j=1}^n h_j(z_j)$, where for all $j = 1, \dots, n$:

(H0) h_j is $\widetilde{\alpha}$-Hölder continuous on $[S^\ell_j, S^r_j]$ for some $\widetilde{\alpha} \in (\frac{1}{3}, 1]$ and smooth on (S^ℓ_j, S^r_j);
(H1) h_j is strictly convex;
(H2) $\lim_{s \searrow S^\ell_j} h''_j(s) = +\infty = \lim_{s \nearrow S^r_j} h''_j(s)$;
(H3) $\frac{1}{h''_j}$ is concave and can be extended at the boundary $\{S^\ell_j, S^r_j\}$ of S to a function in $C^2([S^\ell_j, S^r_j])$.

Obviously, the induced fully decoupled mobility $\mathbf{M}$ as in Section IV.3.4.2 satisfies the requirements of that section, in particular also (M0)–(M3), if h satisfies (H0)–(H3).

Furthermore, let $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and $f : S \to \mathbb{R}$ such that

(F) f is smooth and uniformly convex, i.e. $\nabla_z^2 f(z) \geq C_f \mathbb{1}$ for all $z \in S$ and some $C_f > 0$.

We introduce a *reference state* $\overline{z} \in S$, i.e. a constant level relatively to which certain quantities (e.g. the mass of an element in $\mathscr{M}(\mathbb{R}; S)$) will be measured. We distinguish two qualitatively different cases:

(A) Reference state $\overline{z} = S^\ell$.
(B) Reference state $\overline{z} \in \operatorname{int}(S)$.

The respective case will be indicated with (A) and/or (B) in definitions and statements. Note that in case (A), the function $\mu - \overline{z}$ is nonnegative for each $\mu \in \mathscr{M}(\mathbb{R}; S)$. We begin with a detailed exposition of the relevant energy functionals.

Definition IV.28 (Heat and driving entropy). *Let $\overline{z}, f, h, \eta$ be as mentioned above. Define the* heat entropy *functional $\mathcal{H} : \mathscr{M}(\mathbb{R}; S) \to \mathbb{R}_\infty$ by*

$$\mathcal{H}(\mu) = \int_{\mathbb{R}} h_{\overline{z}}(\mu) \, \mathrm{d}x,$$

where

(A) $h_{\overline{z}} := h(z) - h(\overline{z})$;
(B) $h_{\overline{z}}(z) := h(z) - h(\overline{z}) - (z - \overline{z})^{\mathrm{T}} \nabla_z h(\overline{z})$.

The driving entropy *functional $\mathcal{E} : \mathscr{M}(\mathbb{R}; S) \to \mathbb{R}_\infty$ is defined by*

$$\mathcal{E}(\mu) = \begin{cases} \int_{\mathbb{R}} [f(\mu) - f(\overline{z}) - (\mu - \overline{z})^{\mathrm{T}} \nabla_z f(\overline{z}) + \mu^{\mathrm{T}} \eta] \, \mathrm{d}x & \text{if } \mu \in \mathbf{X}_{\overline{z}}, \\ +\infty & \text{otherwise,} \end{cases}$$

where

(A) $\mathbf{X}_{\overline{z}} := \{\mu \in \mathscr{M}(\mathbb{R};S) : \|\mu - \overline{z}\|_{L^1} = m := \|\mu^0 - \overline{z}\|_{L^1} \in (0,\infty),\ \mathfrak{m}_2(\mu - \overline{z}) < \infty\}$, *where* $\mu^0 \in \mathscr{M}(\mathbb{R};S)$ *is the initial condition;*
(B) $\mathbf{X}_{\overline{z}} := \{\mu \in \mathscr{M}(\mathbb{R};S) : \|\mu - \overline{z}\|_{L^2} < \infty\}$.

Note that in both cases, $h_{\overline{z}}(\overline{z}) = 0$ and $h_{\overline{z}}$ is strictly convex with $\nabla_z^2 h_{\overline{z}} = \nabla_z^2 h$. In case (B), $h_{\overline{z}}$ is nonnegative.

Example IV.29. *(a) The paradigmatic example for h satisfying* (H0)–(H3) *is given by*

$$h_j(s) = \begin{cases} (s - S_j^\ell)\log(s - S_j^\ell) + (S_j^r - s)\log\left(S_j^r - s\right) - (S_j^r - S_j^\ell)\log(S_j^r - S_j^\ell) & \text{if } s \in (S_j^\ell, S_j^r), \\ 0 & \text{if } s \in \{S_j^\ell, S_j^r\}, \end{cases}$$

yielding

$$\mathbf{m}_j(s) = \frac{1}{S_j^r - S_j^\ell}(s - S_j^\ell)(S_j^r - s).$$

(b) An admissible choice for f is

$$f(z) = \frac{1}{2}z^{\mathrm{T}}Qz + \varepsilon r(z),$$

where $Q \in \mathbb{R}^{n\times n}$ *is symmetric positive definite,* $r : S \to \mathbb{R}$ *is smooth and* $\varepsilon \geq 0$ *is such that* $Q + \varepsilon\nabla_z^2 r(z)$ *is positive definite for all* $z \in S$.

We summarize the main results of this section in the following

Theorem IV.30 (Existence of weak solutions to (IV.1.9)). *Consider the initial-value problem for the system of degenerate diffusion equations with drift*

$$\partial_t\mu = \partial_x(\mathbf{M}(\mu)\nabla_z^2 f(\mu)\partial_x\mu + \mathbf{M}(\mu)\partial_x\eta) \quad \text{for } t > 0 \text{ and } x \in \mathbb{R}, \tag{IV.4.1}$$

$$\mu(0,x) = \mu^0(x) \quad \text{for } x \in \mathbb{R}, \tag{IV.4.2}$$

where the mobility $\mathbf{M}$ *is fully decoupled on the value space* $S = [S^\ell, S^r] \subset \mathbb{R}^n$ *and of the form* $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \in \mathbb{R}^{n\times n}$ *with* $h : S \to \mathbb{R}$ *satisfying* (H0)–(H3). *Assume that* $f : S \to \mathbb{R}$ *satisfies* (F) *and* $\eta \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$.
Suppose that $\mu^0 \in \mathscr{M}(\mathbb{R};S)$ *and either*
(A) $\mu^0 - \overline{z} \in L^1(\mathbb{R};\mathbb{R}^n)$ *and* $\mathfrak{m}_2(\mu^0 - \overline{z}) < \infty$ *for* $\overline{z} := S^\ell$; *or*
(B) $\mu^0 - \overline{z} \in L^2(\mathbb{R};\mathbb{R}^n)$ *for some* $\overline{z} \in \operatorname{int}(S)$.
Then, there exists a function $\mu : [0,\infty) \times \mathbb{R} \to S$ *with*

$$\mu \in C^{1/2}([0,T];(\mathscr{M}(\mathbb{R};S),\mathbf{W}_{\mathbf{M}})),$$
$$\mu - \overline{z} \in L^\infty([0,T];L^2(\mathbb{R};\mathbb{R}^n)) \cap L^2([0,T];H^1(\mathbb{R};\mathbb{R}^n)),$$

for all $T > 0$, *satisfying* (IV.4.1) *in the sense of distributions and attaining the initial condition* (IV.4.2). *Additionally, in case* (A), *the following holds for all* $t \in [0,T]$:

$$\|\mu(t) - \overline{z}\|_{L^1} = \|\mu^0 - \overline{z}\|_{L^1}, \quad \text{and} \quad \mathfrak{m}_2(\mu(t) - \overline{z}) < \infty.$$

As usual, we first prove several elementary properties of $\mathcal{H}$ and $\mathcal{E}$:

Proposition IV.31 (Properties of heat and driving entropy (A)&(B)). *The following statements hold:*
(a) $\mathcal{H}$ *is finite on* $\mathbf{X}_{\overline{z}}$.
(b) For all $\mu_0, \mu_1 \in \mathbf{X}_{\overline{z}}$ *with* $\mathbf{W}_{\mathbf{M}}(\mu_0,\mu_1) < \infty$, *condition* (IV.2.13) *holds for* $h_{\overline{z}}$ *in place of h.*
(c) The Lipschitz-type condition (IV.2.16) *holds.*

(d) There exist constants $\underline{C}, \overline{C} > 0$ *such that for all* $\mu \in \mathbf{X}_{\overline{z}}$*, the following holds:*

$$\underline{C}(\|\mu - \overline{z}\|_{L^2}^2 - 1) \le \mathcal{E}(\mu) \le \overline{C}(\|\mu - \overline{z}\|_{L^2}^2 + 1).$$

In particular, $\mathcal{E}$ *is finite on* $\mathbf{X}_{\overline{z}}$.

(e) If $\mu_k - \overline{z} \rightharpoonup \mu - \overline{z}$ *weakly in* $L^2(\mathbb{R};\mathbb{R}^n)$*, then*

$$\mathcal{E}(\mu) \le \liminf_{k\to\infty} \mathcal{E}(\mu_k).$$

Proof. (a) We distinguish both cases.

(A) Due to $\widetilde{\alpha}$-Hölder continuity of h, there exists $C > 0$ such that for all $z \in S$:

$$|h_{\overline{z}}(z)| \le C \sum_{j=1}^n |z_j - \overline{z}_j|^{\widetilde{\alpha}}.$$

By Hölder's inequality, we then deduce for $\mu \in \mathbf{X}_{\overline{z}}$:

$$|\mathcal{H}(\mu)| \le C \sum_{j=1}^n \int_{\mathbb{R}} |\mu_j - \overline{z}_j|^{\widetilde{\alpha}}\,\mathrm{d}x \le C \sum_{j=1}^n \left(\int_{\mathbb{R}} (\mu_j - \overline{z}_j)(x^2+1)\,\mathrm{d}x \right)^{\widetilde{\alpha}} \left(\int_{\mathbb{R}} (x^2+1)^{\frac{\widetilde{\alpha}}{\widetilde{\alpha}-1}}\,\mathrm{d}x \right)^{1-\widetilde{\alpha}},$$

which is finite thanks to $\widetilde{\alpha} > \frac{1}{3}$ and the definition of $\mathbf{X}_{\overline{z}}$.

(B) Obviously, since h is smooth in a neighbourhood of $\overline{z}$ and bounded on the whole of S, there exists $C > 0$ such that for all $z \in S$:

$$h_{\overline{z}}(z) \le C|z - \overline{z}|^2,$$

which proves the claim.

(b) (A) Thanks to the properties of the heat kernel, $\mathbf{G}_\delta * \mu \in \mathbf{X}_{\overline{z}}$ if $\mu \in \mathbf{X}_{\overline{z}}$ since mass is conserved and the second moment grows linearly in time along the heat flow. Hence, by part (a), $\mathcal{H}(\mu_i) - \mathcal{H}(\mathbf{G}_\delta * \mu_i)$ is δ-bounded which yields the claim.

(B) $\mathcal{H}(\mu_i) - \mathcal{H}(\mathbf{G}_\delta * \mu_i) \le \mathcal{H}(\mu_i) < \infty$ by nonnegativity and part (a).

(c) This is obvious thanks to smoothness and concavity of the $\mathbf{m}_j$; take $L = \max_j \mathbf{m}_j'(S_j^\ell)$.

(d) This follows from assumption (F) on f and Taylor's theorem, $\eta \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$ and the fact that $\mathscr{M}(\mathbb{R};S) \subset L^\infty(\mathbb{R};\mathbb{R}^n)$. Note that in both cases, $\mathbf{X}_{\overline{z}} \subset L^2(\mathbb{R};\mathbb{R}^n)$ holds.

(e) This is clear thanks to convexity and nonnegativity of f.

□

IV.4.2. The time-discrete approximative solution

We construct a time-discrete solution by means of the minimizing movement scheme and introduce the *Yosida penalized* energy $\mathcal{E}$, i.e.

$$\mathcal{E}_\tau : \mathscr{M}(\mathbb{R};S) \times \mathscr{M}(\mathbb{R};S) \to \mathbb{R}_\infty, \quad \mathcal{E}_\tau(\mu \,|\, \widetilde{\mu}) := \frac{1}{2\tau} \mathbf{W_M}(\mu,\widetilde{\mu})^2 + \mathcal{E}(\mu),$$

where $\tau \in (0,\overline{\tau}]$ is a given step size; and $\overline{\tau} > 0$.

Proposition IV.32 (Minimizing movement scheme (A)&(B)). *Let* $\tau > 0$ *and* $\widetilde{\mu} \in \mathbf{X}_{\overline{z}}$*. Then, there exists a minimizer* $\mu^* \in \mathbf{X}_{\overline{z}}$ *of the functional* $\mathcal{E}_\tau(\cdot \,|\, \widetilde{\mu})$ *on* $\mathscr{M}(\mathbb{R};S)$*. Moreover, one has*

$$\tau \|\partial_x \mu^*\|_{L^2}^2 \le \frac{2}{C_f}[\mathcal{H}(\widetilde{\mu}) - \mathcal{H}(\mu^*)] + C\tau, \tag{IV.4.3}$$

where $C = C(f,\eta) > 0$*. In particular,* $\mu^* - \overline{z} \in H^1(\mathbb{R};\mathbb{R}^n)$.

Proof. By Proposition IV.31(d), $\mathcal{E}$ is bounded from below. Hence, $\mathcal{E}_\tau(\cdot\,|\,\widetilde{\mu})$ is proper and bounded from below. An infimizing sequence $(\mu_k)_{k\in\mathbb{N}}$ in $\mathbf{X}_{\overline{z}}$, viz.

$$\lim_{k\to\infty}\mathcal{E}_\tau(\mu_k\,|\,\widetilde{\mu}) = \inf\mathcal{E}_\tau(\cdot\,|\,\widetilde{\mu}),$$

thus satisfies $\|\mu_k - \overline{z}\|_{L^2} \le C$ (thanks to (F) in case (B); for case (A), this is trivial because of the uniform L^1 and L^∞ bounds on $\mu_k - \overline{z}$) and $\mathbf{W_M}(\mu_k, \widetilde{\mu}) \le C$ for some constant $C > 0$. Using Proposition IV.12(b) and Alaoglu's theorem yield the existence of a (non-relabelled) subsequence and a limit $\mu^* \in \mathbf{X}_{\overline{z}}$ such that $\mu_k - \overline{z} \rightharpoonup \mu^* - \overline{z}$ weakly in $L^2(\mathbb{R};\mathbb{R}^n)$ and $\mu_k \overset{*}{\rightharpoonup} \mu^*$ weakly$*$ in $\mathscr{M}(\mathbb{R};S)$, as $k\to\infty$. Note that in case (A), finiteness of $\mathfrak{m}_2(\mu^* - \overline{z})$ is a consequence of the uniform bound

$$\mathfrak{m}_2(\mu_k - \overline{z}) \le e^L(\mathfrak{m}_2(\widetilde{\mu} - \overline{z}) + C^2) < \infty,$$

using Proposition IV.15. The lower semicontinuity properties from the Propositions IV.12(a) and IV.31(e) show that μ^* is indeed a minimizer of $\mathcal{E}_\tau(\cdot\,|\,\widetilde{\mu})$.

In order to obtain (IV.4.3), recall that the heat entropy $\mathcal{H}$ is geodesically 0-convex w.r.t. $\mathbf{W_M}$, thanks to Proposition IV.19. Application of the flow interchange lemma (Theorem I.5) yields

$$\tau \mathrm{D}^{\mathcal{H}}\mathcal{E}(\mu^*) \le \mathcal{H}(\widetilde{\mu}) - \mathcal{H}(\mu^*). \tag{IV.4.4}$$

For the dissipation, we obtain (write $\mu_s := \mathsf{S}_s^{\mathcal{H}}(\mu^*)$ for brevity) for small $s > 0$:

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mu_s) &= -\int_{\mathbb{R}}(\nabla_z f(\mu_s) - \nabla_z f(\overline{z}) + \eta)^{\mathrm{T}}\partial_{xx}\mu_s\,\mathrm{d}x = \int_{\mathbb{R}}(\partial_x\mu_s^{\mathrm{T}}\nabla_z^2 f(\mu_s)\partial_x\mu_s + \partial_x\eta^{\mathrm{T}}\partial_x\mu_s)\,\mathrm{d}x\\ &\ge \int_{\mathbb{R}}\left[C_f|\partial_x\mu_s|^2 - \frac{1}{2C_f}|\partial_x\eta|^2 - \frac{C_f}{2}|\partial_x\mu_s|^2\right]\mathrm{d}x = \frac{C_f}{2}\|\partial_x\mu_s\|_{L^2}^2 - \widetilde{C},\end{aligned}$$

where we used (F), the Cauchy-Schwarz and the Young inequality. Note that since $\eta \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$, $\widetilde{C} = \widetilde{C}(f,\eta)$ is finite. Passing to $s \searrow 0$ yields thanks to lower semicontinuity of the right-hand side

$$\mathrm{D}^{\mathcal{H}}\mathcal{E}(\mu^*) \ge \frac{C_f}{2}\|\partial_x\mu^*\|_{L^2}^2 - \widetilde{C},$$

from which (IV.4.3) follows by insertion into (IV.4.4). □

The scheme (I.2.11) is well-posed and produces a sequence $(\mu_\tau^k)_{k\in\mathbb{N}}$ for each initial datum $\mu_\tau^0 = \mu^0 \in \mathbf{X}_{\overline{z}}$. We define the *time-discrete solution* $\mu_\tau : [0,\infty) \to \mathbf{X}_{\overline{z}}$ by piecewise constant interpolation as in (I.2.12). The following statements are an immediate consequence of the minimizing movement scheme:

Proposition IV.33 (Classical estimates (A)&(B)). *The following statements hold:*

(a) For all $k \in \mathbb{N}$, one has $\mathcal{E}(\mu_\tau^k) \le \mathcal{E}(\mu^0) < \infty$.

(b) $\displaystyle\sum_{k=1}^{\infty}\mathbf{W}_\mathbf{M}^2(\mu_\tau^k, \mu_\tau^{k-1}) \le 2\tau(\mathcal{E}(\mu^0) - \inf\mathcal{E})$.

(c) For all $T > 0$ and all $s,t \in [0,T]$, one has

$$\mathbf{W_M}(\mu_\tau(s), \mu_\tau(t)) \le \Big[2(\mathcal{E}(\mu^0) - \inf\mathcal{E})\max(\tau, |t-s|)\Big]^{1/2}.$$

Proof. This is classical, see for instance [4, Ch. 3] (or Proposition II.15 from Part II). □

For clarity, we introduce the following notation for a given function $\varphi : [0,\infty) \to \mathbb{R}$: for each $\tau > 0$ and $s \ge 0$, let

$$\varphi_\tau(s) := \varphi\left(\left\lfloor\frac{s}{\tau}\right\rfloor\tau\right), \text{ where } \lfloor r\rfloor := \max\{n \in \mathbb{N}_0 : n \le r\}.$$

Lemma IV.34 (Discrete weak formulation (A)&(B)). *Let $\alpha > 0$, fix test functions $\rho \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$, $\psi \in C_c^\infty((0,\infty)) \cap C^0([0,\infty))$ and set $\lambda = \lambda(\alpha) = -C\left(\frac{1}{\alpha}+1\right)$ with C from Proposition IV.25(b). Then,* *the discrete solution μ_τ obtained from the scheme* (I.2.11) *satisfies the following* discrete weak formulation*:*

$$\begin{aligned}&\left|\int_0^\infty\int_{\mathbb{R}}\left[\rho^{\mathrm{T}}\mu_\tau\frac{\psi_\tau(t)-\psi_\tau(t+\tau)}{\tau}+\psi_\tau(t)[\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau)\nabla_z^2 f(\mu_\tau)\partial_x\mu_\tau+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau)\partial_x\eta]\right]\,\mathrm{d}x\,\mathrm{d}t\right|\\&\leq\left|\alpha\int_0^\infty\int_{\mathbb{R}}\left[h_{\overline{z}}(\mu_\tau)\frac{|\psi|_\tau(t)-|\psi|_\tau(t+\tau)}{\tau}+|\psi|_\tau(t)[\partial_x\mu_\tau^{\mathrm{T}}\nabla_z^2 f(\mu_\tau)\partial_x\mu_\tau+\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau]\right]\,\mathrm{d}x\,\mathrm{d}t\right.\\&\qquad\left.+2\lambda\tau\|\psi\|_{C^0}[\mathcal{E}(\mu^0)-\inf\mathcal{E}]\right|.\end{aligned}\tag{IV.4.5}$$

Proof. Recall that for this choice of λ, the regularized potential energy $\mathcal{V}$ defined in (IV.3.10) is geodesically λ-convex w.r.t. $\mathbf{W}_\mathbf{M}$ (cf. Proposition IV.25). Hence, we are in position to apply the flow interchange lemma (Theorem I.5) to obtain for all $k\in\mathbb{N}$:

$$\mathcal{V}(\mu_\tau^k)+\tau\mathrm{D}^{\mathcal{V}}\mathcal{E}(\mu_\tau^k)+\frac{\lambda}{2}\mathbf{W}_\mathbf{M}^2(\mu_\tau^k,\mu_\tau^{k-1})\leq\mathcal{V}(\mu_\tau^{k-1}).\tag{IV.4.6}$$

For the dissipation, one has (write $\mu_s:=\mathsf{S}_s^{\mathcal{V}}(\mu_\tau^k)$ for brevity) for small $s>0$ that

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mu_s)&=-\int_{\mathbb{R}}[\nabla_z f(\mu_s)-\nabla_z f(\overline{z})+\eta]^{\mathrm{T}}[\alpha\partial_{xx}\mu_s+\partial_x(\mathbf{M}(\mu_s)\partial_x\rho)]\,\mathrm{d}x\\&=\alpha\int_{\mathbb{R}}\left[\partial_x\mu_s^{\mathrm{T}}\nabla_z^2 f(\mu_s)\partial_x\mu_s+\partial_x\eta^{\mathrm{T}}\partial_x\mu_s+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_s)\nabla_z^2 f(\mu_s)\partial_x\mu_s+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_s)\partial_x\eta\right]\,\mathrm{d}x,\end{aligned}$$

and consequently, passing to $s\searrow 0$:

$$\mathrm{D}^{\mathcal{V}}\mathcal{E}(\mu_\tau^k)\geq\alpha\int_{\mathbb{R}}\left[\partial_x\mu_\tau^{k\,\mathrm{T}}\nabla_z^2 f(\mu_\tau^k)\partial_x\mu_\tau^k+\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau^k+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau^k)\nabla_z^2 f(\mu_\tau^k)\partial_x\mu_\tau^k+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau^k)\partial_x\eta\right]\,\mathrm{d}x.$$

Inserting this into (IV.4.6) and repeating this calculation with $-\rho$ in place of ρ yields

$$\begin{aligned}&\alpha\int_{\mathbb{R}}\left[h_{\overline{z}}(\mu_\tau^k)-h_{\overline{z}}(\mu_\tau^{k-1})+\tau\partial_x\mu_\tau^{k\,\mathrm{T}}\nabla_z^2 f(\mu_\tau^k)\partial_x\mu_\tau^k+\tau\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau^k\right]\,\mathrm{d}x+\frac{\lambda}{2}\mathbf{W}_\mathbf{M}^2(\mu_\tau^k,\mu_\tau^{k-1})\\&\leq\int_{\mathbb{R}}\left[\rho^{\mathrm{T}}[\mu_\tau^k-\mu_\tau^{k-1}]+\tau\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau^k)\nabla_z^2 f(\mu_\tau^k)\partial_x\mu_\tau^k+\tau\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau^k)\partial_x\eta\right]\,\mathrm{d}x\\&\leq-\alpha\int_{\mathbb{R}}\left[h_{\overline{z}}(\mu_\tau^k)-h_{\overline{z}}(\mu_\tau^{k-1})+\tau\partial_x\mu_\tau^{k\,\mathrm{T}}\nabla_z^2 f(\mu_\tau^k)\partial_x\mu_\tau^k+\tau\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau^k\right]\,\mathrm{d}x-\frac{\lambda}{2}\mathbf{W}_\mathbf{M}^2(\mu_\tau^k,\mu_\tau^{k-1}).\end{aligned}\tag{IV.4.7}$$

Let $\psi\in C^0([0,\infty))$ be *nonnegative* and have compact support in $(0,\infty)$. We multiply the chain of inequalities (IV.4.7) with $\psi((k-1)\tau)$ and take the sum over all $k\in\mathbb{N}$, recalling Proposition IV.33(b) and observing

$$\sum_{k\in\mathbb{N}}\psi((k-1)\tau)[g(\mu_\tau^k)-g(\mu_\tau^{k-1})]=\sum_{k\in\mathbb{N}}g(\mu_\tau^k)[\psi((k-1)\tau)-\psi(k\tau)],$$

for an arbitrary map $g:\mathbb{R}^n\to\mathbb{R}^d$, $d\in\mathbb{N}$. The resulting chain of inequalities can be expressed in terms of the discrete solution μ_τ as follows:

$$\begin{aligned}&\alpha\int_0^\infty\int_{\mathbb{R}}\left[h_{\overline{z}}(\mu_\tau)\frac{\psi_\tau(t)-\psi_\tau(t+\tau)}{\tau}+\psi_\tau(t)[\partial_x\mu_\tau^{\mathrm{T}}\nabla_z^2 f(\mu_\tau)\partial_x\mu_\tau+\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau]\right]\,\mathrm{d}x\,\mathrm{d}t\\&\qquad+\lambda\tau\|\psi\|_{C^0}[\mathcal{E}(\mu^0)-\inf\mathcal{E}]\\&\leq\int_0^\infty\int_{\mathbb{R}}\left[\rho^{\mathrm{T}}\mu_\tau\frac{\psi_\tau(t)-\psi_\tau(t+\tau)}{\tau}+\psi_\tau(t)[\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau)\nabla_z^2 f(\mu_\tau)\partial_x\mu_\tau+\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_\tau)\partial_x\eta]\right]\,\mathrm{d}x\,\mathrm{d}t\\&\leq-\alpha\int_0^\infty\int_{\mathbb{R}}\left[h_{\overline{z}}(\mu_\tau)\frac{\psi_\tau(t)-\psi_\tau(t+\tau)}{\tau}+\psi_\tau(t)[\partial_x\mu_\tau^{\mathrm{T}}\nabla_z^2 f(\mu_\tau)\partial_x\mu_\tau+\partial_x\eta^{\mathrm{T}}\partial_x\mu_\tau]\right]\,\mathrm{d}x\,\mathrm{d}t\\&\qquad-\lambda\tau\|\psi\|_{C^0}[\mathcal{E}(\mu^0)-\inf\mathcal{E}].\end{aligned}\tag{IV.4.8}$$

For general $\psi \in C_c^\infty((0,\infty)) \cap C^0([0,\infty))$, decompose ψ into its positive and negative part and subtract the respective inequalities (IV.4.8) to obtain (IV.4.5). □

IV.4.3. Passage to the continuous-time limit

Proposition IV.35 (*A priori* estimates (A)). *For given $T > 0$, there exist constants $C_i = C_i(T) > 0$ such that for all $\tau \in (0, \overline{\tau}]$, the following holds:*
(a) $\mathbf{W_M}(\mu_\tau(t), \mu^0) \leq C_1$ *for all* $t \in [0, T]$.
(b) $\|\mu_\tau - \overline{z}\|_{L^\infty([0,T];L^2)} \leq C_2$.
(c) $\mathfrak{m}_2(\mu_\tau(t) - \overline{z}) \leq C_3$ *for all* $t \in [0, T]$.
(d) $\|\mu_\tau - \overline{z}\|_{L^2([0,T];H^1)} \leq C_4$.

Proof. (a) Using Proposition IV.33(c) yields

$$\mathbf{W_M}(\mu_\tau(t), \mu^0) \leq [2(\mathcal{E}(\mu^0) - \inf \mathcal{E}) \max(\tau, t)]^{1/2} \leq C_1,$$

for $0 \leq t \leq T$ and $0 < \tau \leq \overline{\tau}$.
(b) This is obvious thanks to the uniform bounds on $\mu_\tau - \overline{z}$ in $L^1(\mathbb{R}; \mathbb{R}^n)$ and $L^\infty(\mathbb{R}; \mathbb{R}^n)$, respectively.
(c) By part (a) and Proposition IV.15, one has

$$\mathfrak{m}_2(\mu_\tau(t) - \overline{z}) \leq e^L(\mathfrak{m}_2(\mu^0 - \overline{z}) + C_1^2) \quad \text{for all } t \in [0, T].$$

(d) In view of (b), it remains to prove that $\|\partial_x \mu_\tau\|_{L^2([0,T];L^2)}$ is τ-uniformly bounded. Define $K := \left\lfloor \frac{T}{\tau} \right\rfloor + 1 \leq \frac{T+\overline{\tau}}{\tau}$ to obtain

$$\int_0^T \|\partial_x \mu_\tau(t)\|_{L^2}^2 \, \mathrm{d}t \leq \sum_{k=1}^K \tau \|\partial_x \mu_\tau^k\|_{L^2}^2 \leq \sum_{k=1}^K \left[\frac{2}{C_f} (\mathcal{H}(\mu_\tau^{k-1}) - \mathcal{H}(\mu_\tau^k)) + C\tau \right], \tag{IV.4.9}$$

where we used (IV.4.3) in the last step. In the proof of Proposition IV.31(a), we have seen that there exist constants $\widetilde{C}_0$, $\widetilde{C}_1 > 0$ such that for all $\mu \in \mathbf{X}_{\overline{z}}$

$$|\mathcal{H}(\mu)| \leq \widetilde{C}_0 + \widetilde{C}_1(m + \mathfrak{m}_2(\mu - \overline{z})).$$

Using (c) with $T + \overline{\tau}$ in place of T, we eventually end up with

$$\int_0^T \|\partial_x \mu_\tau(t)\|_{L^2}^2 \, \mathrm{d}t \leq C(T + \overline{\tau}) + \frac{2}{C_f} \left(\mathcal{H}(\mu^0) + \widetilde{C}_0 + \widetilde{C}_1(m + C_3) \right).$$

□

Proposition IV.36 (*A priori* estimates (B)). *For given $T > 0$, there exist constants $C_i = C_i(T) > 0$ such that for all $\tau \in (0, \overline{\tau}]$, the following holds:*
(a) $\mathbf{W_M}(\mu_\tau(t), \mu^0) \leq C_1$ *for all* $t \in [0, T]$.
(b) $\|\mu_\tau - \overline{z}\|_{L^\infty([0,T];L^2)} \leq C_2$.
(d) $\|\mu_\tau - \overline{z}\|_{L^2([0,T];H^1)} \leq C_4$.

Notice that an analog to Proposition IV.35(c) is missing in case (B), since $\mu_\tau - \overline{z}$ does not have a sign.
Proof. Part (a) is the same as for Proposition IV.35. For part (b), thanks to Proposition IV.33(a), for all $t > 0$, one has $\mathcal{E}(\mu_\tau(t)) = \mathcal{E}(\mu_\tau^k) \leq \mathcal{E}(\mu^0)$, with $k = \left\lfloor \frac{t}{\tau} \right\rfloor + 1$. Using Proposition IV.31(d) yields

$$\|\mu_\tau(t) - \overline{z}\|_{L^2} \leq C_2 \quad \text{for all } t > 0.$$

For (d), we again proceed as before to arrive at (IV.4.9). From there, the claim obviously follows by nonnegativity of $\mathcal{H}$. □

We now are in position to pass to the limit $\tau \searrow 0$, thereby completing the proof of Theorem IV.30.

Proposition IV.37 (Continuous-time limit (A)&(B)). *Let $T > 0$ be given, $(\tau_k)_{k\in\mathbb{N}}$ be a vanishing sequence of step sizes, i.e. $\tau_k \searrow 0$ as $k \to \infty$, and $(\mu_{\tau_k})_{k\in\mathbb{N}}$ be the corresponding sequence of discrete solutions obtained by the minimizing movement scheme. Then, there exists a (non-relabelled) subsequence and a limit curve $\mu : [0, T] \to \mathbf{X}_{\overline{z}}$ such that as $k \to \infty$:*
(a) For fixed $t \in [0, T]$, $\mu_{\tau_k}(t) \overset{*}{\rightharpoonup} \mu(t)$ weakly∗ in $\mathscr{M}(\mathbb{R}; S)$,
(b) $\mu_{\tau_k} - \overline{z} \rightharpoonup \mu - \overline{z}$ weakly in $L^2([0, T]; H^1(\mathbb{R}; \mathbb{R}^n))$,
(c) $\mu_{\tau_k} - \overline{z} \to \mu - \overline{z}$ strongly in $L^2([0, T]; L^2_{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^n))$,
with the properties

$$\mu \in C^{1/2}([0, T]; (\mathscr{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})), \tag{IV.4.10}$$

$$\mu - \overline{z} \in L^\infty([0, T]; L^2(\mathbb{R}; \mathbb{R}^n)) \cap L^2([0, T]; H^1(\mathbb{R}; \mathbb{R}^n)). \tag{IV.4.11}$$

Moreover, the limit μ is a weak solution to (IV.1.1) *in the following sense: for all test functions $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and $\psi \in C_c^\infty((0, \infty)) \cap C^0([0, \infty))$, one has*

$$\int_0^\infty \int_\mathbb{R} \Big[-\partial_t \psi \rho^\mathrm{T} \mu + \psi [\partial_x \rho^\mathrm{T} \mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu + \partial_x \rho^\mathrm{T} \mathbf{M}(\mu) \partial_x \eta] \Big] \,\mathrm{d}x\,\mathrm{d}t = 0. \tag{IV.4.12}$$

Remark IV.38 (Distributional formulation). *If μ satisfies the weak formulation* (IV.4.12) *above, then it is also a solution to* (IV.1.1) *in the (pure) sense of distributions, since the space of all test functions $\varphi \in C_c^\infty((0, \infty) \times \mathbb{R}; \mathbb{R}^n)$ of the form $\varphi(t, x) = \psi(t)\rho(x)$ with $\psi \in C_c^\infty((0, \infty))$ and $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ lies dense in $C_c^\infty((0, \infty) \times \mathbb{R}; \mathbb{R}^n)$ (see for instance [85, Thm. 4.3.1]).*

Proof. We divide the proof into several steps.
Step 1: Weak convergence and limit properties.
Using the *a priori* estimates in Proposition IV.35/IV.36(a)&(b) together with Proposition IV.12 and Alaoglu's theorem, we deduce the weak convergences (a)&(b) and also the properties of the limit. Note that in case (A), finiteness of $\mathfrak{m}_2(\mu(t) - \overline{z})$ is a consequence of the uniform estimate from Proposition IV.35(c). In both cases, 1/2-Hölder continuity w.r.t. $\mathbf{W}_\mathbf{M}$ can be obtained thanks to Proposition IV.33(c) via a refined version of the Arzelà-Ascoli theorem (cf. [4, Thm. 3.3.1]).
Step 2: Strong convergence.
In order to prove the strong convergence (c), we fix a bounded interval $I \subset \mathbb{R}$ and apply Theorem I.6 with the admissible choices

$$\mathbf{Y} := \{u \in \mathscr{M}(I; S) : u - \overline{z} \in L^2(I; \mathbb{R}^n)\}, \text{ endowed with } \|u\|_\mathbf{Y} := \|u - \overline{z}\|_{L^2(I)},$$

which is isometric to a closed subset of $L^2(I; \mathbb{R}^n)$,

$$\mathcal{A}(u) := \begin{cases} \|u - \overline{z}\|^2_{H^1(I)} & \text{if } u - \overline{z} \in H^1(I; \mathbb{R}^n), \\ +\infty & \text{otherwise,} \end{cases}$$

which has relatively compact sublevels in $\mathbf{Y}$ due to the Rellich-Kondrachov compactness theorem, and the pseudo-distance $\mathbf{W}$ defined by

$$\mathbf{W}(u, \widetilde{u}) := \inf \Big\{ \mathbf{W}_\mathbf{M}(w, \widetilde{w}) : w, \widetilde{w} \in \mathscr{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M}(w, \mu_0), \mathbf{W}_\mathbf{M}(\widetilde{w}, \mu_0) \le C_1, w|_I = u, \widetilde{w}|_I = \widetilde{u} \Big\},$$

where C_1 is the constant from Proposition IV.35/IV.36(a). Thanks to the topological properties of $\mathbf{W}_\mathbf{M}$ (cf. Proposition IV.12), one easily sees that finiteness of the infimum above yields the existence of a minimizer, and that the requirements of Theorem I.6 are fulfilled. We verify the hypotheses (I.2.13)&(I.2.14) for the sequence $(U_k)_{k\in\mathbb{N}}$ defined by $U_k := \mu_{\tau_k}|_{[0,\infty)\times I}$: (I.2.13) is immediate because of the *a priori* estimate from

Proposition IV.35/IV.36(d). For (I.2.14), we first notice by construction of $\mathbf{W}$ that

$$\mathbf{W}(U_k(t+h), U_k(t)) \le \mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t)).$$

We claim that

$$\sup_{k\in\mathbb{N}} \int_0^{T-h} \mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t))\,\mathrm{d}t \le \max\left(1, \sqrt{T+\overline{\tau}}\right)\sqrt{2(\mathcal{E}(\mu^0) - \inf\mathcal{E})(T+\overline{\tau})h}, \tag{IV.4.13}$$

for all $h \in (0, \overline{\tau})$, from which (I.2.14) follows. Indeed, for fixed $k \in \mathbb{N}$ and $h \in (0, \tau_k]$, one has

$$\begin{aligned}\int_0^{T-h} \mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t))\,\mathrm{d}t &= \sum_{i=1}^{\left\lfloor \frac{T}{\tau_k}\right\rfloor} h\mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}^i, \mu_{\tau_k}^{i+1}) \le \sqrt{2(\mathcal{E}(\mu^0) - \inf\mathcal{E})}\sqrt{h^2\left\lfloor \frac{T}{\tau_k}\right\rfloor} \\ &\le \sqrt{2(\mathcal{E}(\mu^0) - \inf\mathcal{E})(T+\overline{\tau})h},\end{aligned}$$

thanks to Hölder's inequality and Proposition IV.33(b). On the other hand, for $h \in (\tau_k, \overline{\tau}]$, we directly get from Proposition IV.33(c):

$$\int_0^{T-h} \mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t))\,\mathrm{d}t \le (T-h)\sqrt{2(\mathcal{E}(\mu^0) - \inf\mathcal{E})h} \le (T+\overline{\tau})\sqrt{2(\mathcal{E}(\mu^0) - \inf\mathcal{E})h}.$$

Hence, (IV.4.13) holds and the application of Theorem I.6 yields the existence of a (non-relabelled) subsequence which converges to (the spatial restriction to I of) μ in measure w.r.t. $t \in (0, T)$. By the uniform estimate in Proposition IV.35/IV.36(b) and the dominated convergence theorem, we conclude that

$$\mu_{\tau_k} - \overline{z} \to \mu - \overline{z} \text{ strongly in } L^2([0,T] \times I; \mathbb{R}^n),$$

proving claim (c) for a prescribed interval I. By a diagonal argument, setting $I_R := [-R, R]$ and letting $R \nearrow \infty$, we deduce that (c) is true simultaneously for every bounded interval I, extracting a further subsequence. Moreover, we may assume that μ_{τ_k} converges to μ almost everywhere in $[0, T] \times \mathbb{R}$.
Step 3: Weak formulation.
Let $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and $\psi \in C_c^\infty((0,\infty)) \cap C^0([0,\infty))$ be given and set $\alpha_k := \sqrt{\tau_k}$ for $k \in \mathbb{N}$. By Lemma IV.34, μ_{τ_k} satisfies the discrete weak formulation (IV.4.5) for each k, putting $\lambda_k = \lambda(\alpha_k)$ according to Lemma IV.34. Note that with this choice of α_k, one has $\lim_{k\to\infty} \lambda_k\tau_k = 0$.

We first prove that

$$\int_0^\infty \int_{\mathbb{R}} \left[h_{\overline{z}}(\mu_{\tau_k})\frac{|\psi|_{\tau_k}(t) - |\psi|_{\tau_k}(t+\tau_k)}{\tau_k} + |\psi|_{\tau_k}(t)[\partial_x\mu_{\tau_k}^{\mathrm{T}}\nabla_z^2 f(\mu_{\tau_k})\partial_x\mu_{\tau_k} + \partial_x\eta^{\mathrm{T}}\partial_x\mu_{\tau_k}]\right]\mathrm{d}x\,\mathrm{d}t \tag{IV.4.14}$$

is bounded w.r.t. $k \in \mathbb{N}$. For the first part, since $\psi \in C_c^\infty((0,\infty))$, there exists $T' > 0$ such that

$$\left|\int_0^\infty \int_{\mathbb{R}} h_{\overline{z}}(\mu_{\tau_k})\frac{|\psi|_{\tau_k}(t) - |\psi|_{\tau_k}(t+\tau_k)}{\tau_k}\,\mathrm{d}x\,\mathrm{d}t\right| \le C\int_0^{T'}\int_{\mathbb{R}} |h_{\overline{z}}(\mu_{\tau_k})|\,\mathrm{d}x\,\mathrm{d}t.$$

In case (A), we obtain

$$\int_0^{T'}\int_{\mathbb{R}} |h_{\overline{z}}(\mu_{\tau_k})|\,\mathrm{d}x\,\mathrm{d}t \le \int_0^{T'}\left[\widetilde{C}_0 + \widetilde{C}_1(m + \mathbf{m}_2(\mu_{\tau_k}(t) - \overline{z}))\right]\mathrm{d}t,$$

which is bounded thanks to Proposition IV.35(c). In case (B), we have

$$\int_0^{T'}\int_{\mathbb{R}} |h_{\overline{z}}(\mu_{\tau_k})|\,\mathrm{d}x\,\mathrm{d}t \le \widetilde{C}\int_0^{T'} \|\mu_{\tau_k}(t) - \overline{z}\|_{L^2}^2\,\mathrm{d}t,$$

so Proposition IV.36(b) yields boundedness. For the second part in (IV.4.14), we use the inclusion $\mathscr{M}(\mathbb{R}; S) \subset L^\infty(\mathbb{R}; \mathbb{R}^n)$ and the Cauchy-Schwarz inequality to obtain

$$\left|\int_0^\infty \int_{\mathbb{R}} |\psi|_{\tau_k}(t)[\partial_x \mu_{\tau_k}^{\mathrm{T}} \nabla_z^2 f(\mu_{\tau_k})\partial_x \mu_{\tau_k} + \partial_x \eta^{\mathrm{T}} \partial_x \mu_{\tau_k}]\,\mathrm{d}x\,\mathrm{d}t\right| \le \overline{C} \int_0^{T'} \int_{\mathbb{R}} \left[\overline{C}'|\partial_x \mu_{\tau_k}|^2 + \frac{1}{2}|\partial_x \eta|^2\right]\,\mathrm{d}x\,\mathrm{d}t.$$

Thanks to $\eta \in C_c^\infty(\mathbb{R};\mathbb{R}^n)$ and Proposition IV.35/IV.36(d), this is bounded.

From the dominated convergence theorem, since μ_{τ_k} converges to μ pointwise a.e., it follows — using $\mathscr{M}(\mathbb{R};S) \subset L^\infty(\mathbb{R};\mathbb{R}^n)$ again — that

$$\begin{aligned}\lim_{k\to\infty} &\left(\int_0^\infty \int_{\mathbb{R}} \left[\rho^{\mathrm{T}} \mu_{\tau_k} \frac{\psi_{\tau_k}(t) - \psi_{\tau_k}(t+\tau_k)}{\tau_k} + \psi_{\tau_k}(t)\partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu_{\tau_k})\partial_x \eta\right]\,\mathrm{d}x\,\mathrm{d}t\right) \\ &= \int_0^\infty \int_{\mathbb{R}} \left[-\partial_t \psi \rho^{\mathrm{T}} \mu + \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\partial_x \eta\right]\,\mathrm{d}x\,\mathrm{d}t.\end{aligned} \tag{IV.4.15}$$

We now prove

$$\lim_{k\to\infty}\left(\int_0^\infty \int_{\mathbb{R}} \psi_{\tau_k}(t)\partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu_{\tau_k})\nabla_z^2 f(\mu_{\tau_k})\partial_x \mu_{\tau_k}\,\mathrm{d}x\,\mathrm{d}t\right) = \int_0^\infty \int_{\mathbb{R}} \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)\partial_x \mu\,\mathrm{d}x\,\mathrm{d}t. \tag{IV.4.16}$$

First, we show that

$$\lim_{k\to\infty}\int_0^\infty \int_{\mathbb{R}} \left(\psi_{\tau_k}(t)\partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu_{\tau_k})\nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)\right)\partial_x \mu_{\tau_k}\,\mathrm{d}x\,\mathrm{d}t = 0. \tag{IV.4.17}$$

Using Hölder's inequality, the fact that ψ and ρ have compact support and Proposition IV.35/IV.36(d) reduces the problem to verifying

$$\lim_{k\to\infty}\int_0^\infty \int_{\mathbb{R}} \left|\psi_{\tau_k}(t)\partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu_{\tau_k})\nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)\right|^2\,\mathrm{d}x\,\mathrm{d}t = 0.$$

We can proceed using the dominated convergence theorem since the integrand converges pointwise a.e. to zero and the following pointwise estimate holds:

$$\left|\psi_{\tau_k}(t)\partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu_{\tau_k})\nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)\right|^2 \le C\mathbf{1}_{\operatorname{supp}\psi}\mathbf{1}_{\operatorname{supp}\rho}.$$

The r.h.s. obviously is integrable on $(0,\infty)\times\mathbb{R}$. Second,

$$\lim_{k\to\infty}\int_0^\infty \int_{\mathbb{R}} \psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)(\partial_x \mu_{\tau_k} - \partial_x \mu)\,\mathrm{d}x\,\mathrm{d}t = 0, \tag{IV.4.18}$$

since $\psi \partial_x \rho^{\mathrm{T}} \mathbf{M}(\mu)\nabla_z^2 f(\mu)$ is bounded and has compact support in $[0,T']\times\mathbb{R}$ for some $T' > 0$ and hence is an element of $L^2([0,T'];L^2(\mathbb{R};\mathbb{R}^n))$: this yields the claim together with the weak convergence of $\partial_x \mu_{\tau_k} \rightharpoonup \partial_x \mu$ in $L^2([0,T'];L^2(\mathbb{R};\mathbb{R}^n))$ using part (b) of this proposition. We have thus proved (IV.4.16). Putting (IV.4.14)–(IV.4.16) together yields (IV.4.12). □

IV.4.4. Extension to systems of fourth order

In this section, we briefly sketch the possible extension of the methods used in this chapter to the case of fourth-order systems of the form (IV.1.10). Most of the arguments coincide with those of the preceding sections, so we omit the technical details where appropriate.

As in Section IV.4.1, we introduce a constant reference state $\overline{z} \in S = [S^\ell, S^r]$, which either is the corner S^ℓ (case (A)) or an element of the interior (case (B)). The density f of the free energy $\mathcal{E}$ is assumed to be a smooth function $f : \mathbb{R}^n \times S \to \mathbb{R}$ subject to the following convexity and growth property:

(F′) There exist $\overline{C}_f > 0$ and $\underline{C}_f > 0$ such that for all $p \in \mathbb{R}^n$, $z \in S$ and $(\pi,\zeta) \in \mathbb{R}^n \times \mathbb{R}^n$, the following holds:

$$\underline{C}_f(|\pi|^2 + |\zeta|^2) \le \mathrm{D}^2_{(p,z)} f(p,z)[(\pi,\zeta),(\pi,\zeta)] \le \overline{C}_f(|\pi|^2 + |\zeta|^2).$$

Compared to assumption (F) from Section IV.4.1, further restrictions on the (joint) Hessian $\nabla^2_{(p,z)} f$ have to be imposed here because a sensible definition of f does not *a priori* confine the p argument —

corresponding to the spatial derivative of μ when evaluating $\mathcal{E}(\mu)$ — to a compact set. The driving entropy $\mathcal{E}$ and the auxiliary space $\mathbf{X}_{\overline{z}}$ are now defined as follows:

Definition IV.39 (Gradient-dependent driving entropy). *Let $\overline{z}$ and f be as mentioned above. The* driving entropy *functional $\mathcal{E} : \mathscr{M}(\mathbb{R}; S) \to \mathbb{R}_\infty$ is defined by*

$$\mathcal{E}(\mu) = \begin{cases} \int_{\mathbb{R}} [f(\mu_x, \mu) - f(0, \overline{z}) - \mu_x^{\mathrm{T}} \nabla_p f(0, \overline{z}) - (\mu - \overline{z})^{\mathrm{T}} \nabla_z f(0, \overline{z})] \, \mathrm{d}x & \text{if } \mu \in \mathbf{X}_{\overline{z}}, \\ +\infty & \text{otherwise,} \end{cases}$$

where

(A) $\mathbf{X}_{\overline{z}} := \{\mu \in \mathscr{M}(\mathbb{R}; S) : \|\mu - \overline{z}\|_{L^1} = m := \|\mu^0 - \overline{z}\|_{L^1} \in (0, \infty), \mathbf{m}_2(\mu - \overline{z}) < \infty, \|\mu_x\|_{L^2} < \infty\}$, *where $\mu^0 \in \mathscr{M}(\mathbb{R}; S)$ is the initial condition;*

(B) $\mathbf{X}_{\overline{z}} := \{\mu \in \mathscr{M}(\mathbb{R}; S) : \|\mu - \overline{z}\|_{H^1} < \infty\}$.

Easily, one shows the following coercivity and growth estimate on $\mathcal{E}$ for each $\mu \in \mathbf{X}_{\overline{z}}$:

$$\underline{C} \|\mu - \overline{z}\|_{H^1}^2 \leq \mathcal{E}(\mu) \leq \overline{C} \|\mu - \overline{z}\|_{H^1}^2,$$

for some constants $\underline{C}, \overline{C} > 0$. Note that since we consider the whole space $\mathbb{R}$ as spatial domain and therefore Poincaré's inequality is not available, both $\mu - \overline{z}$ and μ_x have to be controlled in $L^2(\mathbb{R}; \mathbb{R}^n)$. This is a further reason for the rather restrictive property (F') (which e.g. excludes functions f which are independent from z).

Our main result on the existence of solutions to system (IV.1.10) is of similar form as Theorem IV.30:

Theorem IV.40 (Existence of weak solutions to (IV.1.10)). *Consider the initial-value problem for the system of degenerate diffusion equations of fourth order on $(0, \infty) \times \mathbb{R}$,*

$$\partial_t \mu = \partial_x \big(\mathbf{M}(\mu) \Big[\nabla_{zz}^2 f(\mu_x, \mu) \mu_x - \partial_x (\nabla_{pp}^2 f(\mu_x, \mu) \mu_{xx}) - \partial_x (\nabla_{pz}^2 f(\mu_x, \mu)) \mu_x \Big] \big), \tag{IV.4.19}$$

$$\mu(0, \cdot) = \mu^0, \tag{IV.4.20}$$

where the mobility $\mathbf{M}$ is fully decoupled on the value space $S = [S^\ell, S^r] \subset \mathbb{R}^n$ and of the form $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \in \mathbb{R}^{n \times n}$ with $h : S \to \mathbb{R}$ satisfying (H0)–(H3). *Assume that $f : \mathbb{R}^n \times S \to \mathbb{R}$ satisfies* (F').

Suppose that $\mu^0 \in \mathscr{M}(\mathbb{R}; S)$ and either

(A) $\mu^0 - \overline{z} \in L^1(\mathbb{R}; \mathbb{R}^n)$, $\mathbf{m}_2(\mu^0 - \overline{z}) < \infty$ *and* $\partial_x \mu^0 \in L^2(\mathbb{R}; \mathbb{R}^n)$ *for* $\overline{z} := S^\ell$; *or*

(B) $\mu^0 - \overline{z} \in H^1(\mathbb{R}; \mathbb{R}^n)$ *for some* $\overline{z} \in \operatorname{int}(S)$.

Then, there exists a function $\mu : [0, \infty) \times \mathbb{R} \to S$ with

$$\mu \in C^{1/2}([0, T]; (\mathscr{M}(\mathbb{R}; S), \mathbf{W_M})),$$
$$\mu - \overline{z} \in L^\infty([0, T]; H^1(\mathbb{R}; \mathbb{R}^n)) \cap L^2([0, T]; H^2(\mathbb{R}; \mathbb{R}^n)),$$

for all $T > 0$ satisfying (IV.4.19) *in the sense of distributions and attaining the initial condition* (IV.4.20). *Additionally, in case* (A), *the following holds for all $t \in [0, T]$:*

$$\|\mu(t) - \overline{z}\|_{L^1} = \|\mu^0 - \overline{z}\|_{L^1}, \quad \text{and} \quad \mathbf{m}_2(\mu(t) - \overline{z}) < \infty.$$

As above, the weak solution is constructed by approximation via the minimizing movement scheme:

Proposition IV.41 (Minimizing movement scheme (A)&(B)). *Let $\tau > 0$ and $\widetilde{\mu} \in \mathbf{X}_{\overline{z}}$. Then, there exists a minimizer $\mu^* \in \mathbf{X}_{\overline{z}}$ of the functional $\mathcal{E}_\tau(\cdot \mid \widetilde{\mu})$ on $\mathscr{M}(\mathbb{R}; S)$. Moreover, one has*

$$\tau \|\partial_{xx} \mu^*\|_{L^2}^2 \leq \frac{1}{\underline{C}_f} [\mathcal{H}(\widetilde{\mu}) - \mathcal{H}(\mu^*)], \tag{IV.4.21}$$

where the heat entropy $\mathcal{H}$ is defined as in Definition IV.28. In particular, $\mu^* - \overline{z} \in H^2(\mathbb{R}; \mathbb{R}^n)$.

Proof. We sketch the proof of the additional regularity estimate (IV.4.21). As in the proof of Proposition IV.32, we use the flow interchange lemma (Theorem I.5) with the heat entropy $\mathcal{H}$ as auxiliary entropy. Writing $\mu_s := \mathsf{S}_s^{\mathcal{H}}(\mu^*)$, we obtain for the dissipation of $\mathcal{E}$ along $\mathsf{S}^{\mathcal{H}}$ with integration by parts:

$$
\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mu_s)\\
&=\int_{\mathbb{R}}\Big(\partial_x\mu_s^{\mathrm{T}}\nabla_{zz}^2 f(\partial_x\mu_s,\mu_s)\partial_x\mu_s-\partial_x\mu_s^{\mathrm{T}}\partial_x[\nabla_{pp}^2 f(\partial_x\mu_s,\mu_s)\partial_{xx}\mu_s]-\partial_x\mu_s^{\mathrm{T}}\partial_x[\nabla_{pz}^2 f(\partial_x\mu_s,\mu_s)]\partial_x\mu_s\Big)\,\mathrm{d}x\\
&=\int_{\mathbb{R}}\mathrm{D}^2_{(p,z)}f(\partial_x\mu_s,\mu_s)[(\partial_{xx}\mu_s,\partial_x\mu_s),(\partial_{xx}\mu_s,\partial_x\mu_s)]\,\mathrm{d}x\geq \underline{C}_f\|\partial_{xx}\mu_s\|_{L^2}^2,
\end{aligned}
$$

where we used (F') in the last step. Estimate (IV.4.21) is now a straightforward consequence of the flow interchange lemma (Theorem I.5), letting $s\searrow 0$. □

With the results from Proposition IV.41 at hand, one is able to define the discrete solution μ_τ as usual as the piecewise constant interpolation of the successive minimizers $(\mu_\tau^k)_{k\in\mathbb{N}}$ from the minimizing movement scheme and to set up a discrete weak formulation with the same method as in Section IV.4.2. Similarly to Section IV.4.3, the following *a priori* estimates for μ_τ hold: for each fixed $T>0$ and $\overline{\tau}>0$, there exists a constant $C>0$, such that

$$
\|\mu_\tau-\overline{z}\|_{L^\infty([0,T];H^1)}\leq C\quad\text{and}\quad\|\mu_\tau-\overline{z}\|_{L^2([0,T];H^2)}\leq C\quad\text{for all }\tau\in(0,\overline{\tau}]. \tag{IV.4.22}
$$

It remains to pass to the limit $\tau\searrow 0$ in a suitable sense:

Proposition IV.42 (Continuous-time limit (A)&(B)). *Let $T>0$ be given, $(\tau_k)_{k\in\mathbb{N}}$ be a vanishing sequence of step sizes, i.e. $\tau_k\searrow 0$ as $k\to\infty$, and $(\mu_{\tau_k})_{k\in\mathbb{N}}$ be the corresponding sequence of discrete solutions obtained by the minimizing movement scheme. Then, there exists a (non-relabelled) subsequence and a limit curve $\mu:[0,T]\to\mathbf{X}_{\overline{z}}$ such that as $k\to\infty$:*
(a) For fixed $t\in[0,T]$, $\mu_{\tau_k}(t)\overset{*}{\rightharpoonup}\mu(t)$ weakly$*$ in $\mathscr{M}(\mathbb{R};S)$,
(b) $\mu_{\tau_k}-\overline{z}\rightharpoonup\mu-\overline{z}$ weakly in $L^2([0,T];H^2(\mathbb{R};\mathbb{R}^n))$,
(c) $\mu_{\tau_k}-\overline{z}\to\mu-\overline{z}$ strongly in $L^2([0,T];H^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R}^n))$,
with the properties

$$
\mu\in C^{1/2}([0,T];(\mathscr{M}(\mathbb{R};S),\mathbf{W_M})), \tag{IV.4.23}
$$

$$
\mu-\overline{z}\in L^\infty([0,T];H^1(\mathbb{R};\mathbb{R}^n))\cap L^2([0,T];H^2(\mathbb{R};\mathbb{R}^n)). \tag{IV.4.24}
$$

Moreover, the limit μ is a weak solution to (IV.1.10) *in the following sense: for all $\rho\in C_c^\infty(\mathbb{R};\mathbb{R}^n)$ and all $\psi\in C_c^\infty((0,\infty))\cap C^0([0,\infty))$, one has*

$$
\int_0^\infty\int_{\mathbb{R}}\Big(-\partial_t\psi\rho^{\mathrm{T}}\mu+\psi\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu)\Big[\nabla_{zz}^2 f(\mu_x,\mu)\mu_x-\partial_x(\nabla_{pp}^2 f(\mu_x,\mu)\mu_{xx})-\partial_x(\nabla_{pz}^2 f(\mu_x,\mu))\mu_x\Big]\Big)\,\mathrm{d}x\,\mathrm{d}t=0.
$$

Proof. The convergence properties above can be obtained by the same method as in Section IV.4.3. Concerning the weak formulation, we prove the following: with $q_k:=\psi_{\tau_k}\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu_{\tau_k})$ and $q:=\psi\partial_x\rho^{\mathrm{T}}\mathbf{M}(\mu)$, one has

$$
\begin{aligned}
&\lim_{k\to\infty}\int_0^\infty\int_{\mathbb{R}}\Big(q_k^{\mathrm{T}}\Big[\nabla_{zz}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_x\mu_{\tau_k}-\partial_x(\nabla_{pp}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_{xx}\mu_{\tau_k})-\partial_x(\nabla_{pz}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_x\mu_{\tau_k}\Big]\Big)\,\mathrm{d}x\,\mathrm{d}t\\
&=\int_0^\infty\int_{\mathbb{R}}\Big(q^{\mathrm{T}}\Big[\nabla_{zz}^2 f(\partial_x\mu,\mu)\partial_x\mu-\partial_x(\nabla_{pp}^2 f(\partial_x\mu,\mu)\partial_{xx}\mu)-\partial_x(\nabla_{pz}^2 f(\partial_x\mu,\mu))\partial_x\mu\Big]\Big)\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{IV.4.25}
$$

For the proof of (IV.4.25), we integrate by parts and rearrange terms to observe

$$
\begin{aligned}
\int_0^\infty\int_{\mathbb{R}}\Big[&q_k^{\mathrm{T}}\Big[\nabla_{zz}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_x\mu_{\tau_k}-\partial_x(\nabla_{pp}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_{xx}\mu_{\tau_k})-\partial_x(\nabla_{pz}^2 f(\partial_x\mu_{\tau_k},\mu_{\tau_k})\partial_x\mu_{\tau_k}\Big]\\
&-q^{\mathrm{T}}\Big[\nabla_{zz}^2 f(\partial_x\mu,\mu)\partial_x\mu-\partial_x(\nabla_{pp}^2 f(\partial_x\mu,\mu)\partial_{xx}\mu)-\partial_x(\nabla_{pz}^2 f(\partial_x\mu,\mu))\partial_x\mu\Big]\Big]\,\mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^\infty \int_{\mathbb{R}} \Big[q_k^{\mathrm{T}} \nabla_{zz}^2 f(\partial_x \mu_{\tau_k}, \mu_{\tau_k}) \partial_x \mu_{\tau_k} - q^{\mathrm{T}} \nabla_{zz}^2 f(\partial_x \mu, \mu) \partial_x \mu \\
&\qquad + \partial_x q_k^{\mathrm{T}} \nabla_{pp}^2 f(\partial_x \mu_{\tau_k}, \mu_{\tau_k}) \partial_{xx} \mu_{\tau_k} - \partial_x q^{\mathrm{T}} \nabla_{pp}^2 f(\partial_x \mu, \mu) \partial_{xx} \mu \\
&\qquad + \partial_x q_k^{\mathrm{T}} \nabla_{pz}^2 f(\partial_x \mu_{\tau_k}, \mu_{\tau_k}) \partial_x \mu_{\tau_k} + q_k^{\mathrm{T}} \nabla_{pz}^2 f(\partial_x \mu_{\tau_k}, \mu_{\tau_k}) \partial_{xx} \mu_{\tau_k} \\
&\qquad - \partial_x q^{\mathrm{T}} \nabla_{pz}^2 f(\partial_x \mu, \mu) \partial_x \mu - q^{\mathrm{T}} \nabla_{pz}^2 f(\partial_x \mu, \mu) \partial_{xx} \mu \Big] \,\mathrm{d}x \,\mathrm{d}t \\
&= \int_0^\infty \int_{\mathbb{R}} \Big[\mathrm{D}^2_{(p,z)} f(\partial_x \mu_{\tau_k}, \mu_{\tau_k})[(q_k, \partial_x q_k), (\partial_x \mu_{\tau_k}, \partial_{xx} \mu_{\tau_k})] - \mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), (\partial_x \mu, \partial_{xx} \mu)] \Big] \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned}
$$

Inserting suitable terms, the claim (IV.4.25) can be shown by verifying

$$
\begin{aligned}
&\lim_{k\to\infty} \int_0^\infty \int_{\mathbb{R}} \Big(\mathrm{D}^2_{(p,z)} f(\partial_x \mu_{\tau_k}, \mu_{\tau_k})[(q_k, \partial_x q_k), \cdot] \\
&\qquad\qquad - \mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), \cdot] \Big) \left[(\partial_x \mu_{\tau_k}, \partial_{xx} \mu_{\tau_k}) \right] \,\mathrm{d}x \,\mathrm{d}t = 0, \qquad \text{(IV.4.26)}
\end{aligned}
$$

$$
\lim_{k\to\infty} \int_0^\infty \int_{\mathbb{R}} \mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), (\partial_x \mu_{\tau_k}, \partial_{xx} \mu_{\tau_k}) - (\partial_x \mu, \partial_{xx} \mu)] \,\mathrm{d}x \,\mathrm{d}t = 0. \qquad \text{(IV.4.27)}
$$

Beforehand, we prove that the operator norm $\|\mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), \cdot]\|$ is an element of $L^2([0,\infty) \times \mathbb{R})$. Indeed, using (F'), we have that

$$
\int_0^\infty \int_{\mathbb{R}} \|\mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), \cdot]\|^2 \,\mathrm{d}x \,\mathrm{d}t \le \int_0^\infty \int_{\mathbb{R}} \overline{C}_f^2 (|q|^2 + |\partial_x q|^2) \,\mathrm{d}x \,\mathrm{d}t.
$$

Using the definition of q and Young's inequality, one arrives at

$$
\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} \overline{C}_f^2 (|q|^2 + |\partial_x q|^2) \,\mathrm{d}x \,\mathrm{d}t \\
&\le \int_0^\infty \int_{\mathbb{R}} \overline{C}_f^2 \|\psi\|_{C^0} \|\rho\|_{C^2} \mathbf{1}_{\operatorname{supp}\psi} \mathbf{1}_{\operatorname{supp}\rho} \left[3 \sup_{z\in S} \|\mathbf{M}(z)\|^2 + 2 \sup_{z\in S} \|\mathrm{D}\mathbf{M}(z)\|^2 |\partial_x \mu|^2 \right] \,\mathrm{d}x \,\mathrm{d}t,
\end{aligned}
$$

which is finite. The same arguments show that for all $k \in \mathbb{N}$, one has

$$
\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} \|\mathrm{D}^2_{(p,z)} f(\partial_x \mu_{\tau_k}, \mu_{\tau_k})[(q_k, \partial_x q_k), \cdot]\|^2 \,\mathrm{d}x \,\mathrm{d}t \\
&\le \int_0^\infty \int_{\mathbb{R}} \overline{C}_f^2 \|\psi\|_{C^0} \|\rho\|_{C^2} \mathbf{1}_{\operatorname{supp}\psi} \mathbf{1}_{\operatorname{supp}\rho} \left[3 \sup_{z\in S} \|\mathbf{M}(z)\|^2 + 2 \sup_{z\in S} \|\mathrm{D}\mathbf{M}(z)\|^2 |\partial_x \mu_{\tau_k}|^2 \right] \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned}
$$

Thanks to the strong convergence of $\partial_x \mu_{\tau_k} \to \partial_x \mu$ in $L^2([0,T]; L^2_{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^n)$ and Vitali's convergence theorem, the integrand on the right-hand side of the last expression above is uniformly integrable. Without loss of generality, we may assume that $\mu_{\tau_k} \to \mu$ and $\partial_x \mu_{\tau_k} \to \partial_x \mu$ pointwise almost everywhere on $[0,T] \times \mathbb{R}$ (extracting further subsequences, if necessary). By continuity,

$$
\|\mathrm{D}^2_{(p,z)} f(\partial_x \mu_{\tau_k}, \mu_{\tau_k})[(q_k, \partial_x q_k), \cdot] - \mathrm{D}^2_{(p,z)} f(\partial_x \mu, \mu)[(q, \partial_x q), \cdot]\| \to 0
$$

pointwise almost everywhere on $[0,T] \times \mathbb{R}$ as well. Using Vitali's convergence theorem again, uniform integrability yields that the above convergence also holds in $L^2([0,T]; L^2(\mathbb{R}; \mathbb{R}^n))$. Thanks to the estimates in (IV.4.22), one has $\|(\partial_x \mu_{\tau_k}, \partial_{xx} \mu_{\tau_k})\|_{L^2([0,T]; L^2(\mathbb{R}; \mathbb{R}^n \times \mathbb{R}^n))} \le \|\mu_{\tau_k}\|_{L^2([0,T]; H^2(\mathbb{R}; \mathbb{R}^n))} \le C$ for all $k \in \mathbb{N}$; hence Hölder's inequality allows us to conclude (IV.4.26). The second claim (IV.4.27) is a straightforward consequence of the weak convergence $\partial_x \mu_{\tau_k} \rightharpoonup \partial_x \mu$ and $\partial_{xx} \mu_{\tau_k} \rightharpoonup \partial_{xx} \mu$ in $L^2([0,T]; L^2(\mathbb{R}; \mathbb{R}^n))$. □

Part V

Generalized gradient systems modelling chemical reactions

CHAPTER V.1

Introduction to Part V

The results presented in this part of the thesis are joint work with Karoline Disser and Matthias Liero [72].

We investigate the evolutionary limit of systems modelling slow and fast chemical reactions of mass-action type. More specifically, we are concerned with the evolutionary Γ-convergence (see Definition I.21) as $\varepsilon \searrow 0$ of the family of generalized gradient systems $(\mathbf{X}, \mathcal{E}, \Psi_\varepsilon)$ associated to the system of ordinary differential equations

$$\dot{c}_\varepsilon = -\sum_{r=1}^{R} \left(k^r_{\mathrm{fw}}(c_\varepsilon)c_\varepsilon^{\alpha^r} - k^r_{\mathrm{bw}}(c_\varepsilon)c_\varepsilon^{\beta^r}\right)(\alpha^r - \beta^r) =: -R(c_\varepsilon), \tag{V.1.1}$$

on the state space $\mathbf{X} = [0,\infty)^I$, which is a closed subset of the Banach space $\mathbb{R}^I$. The vector $c_\varepsilon(t) \in \mathbf{X}$ denotes the concentrations of the $I \in \mathbb{N}$ species or chemical agents at time $t \geq 0$ undergoing $R \leq I$ reactions obeying the mass-action law, where $k^r_{\mathrm{fw}} \geq 0$ and $k^r_{\mathrm{bw}} \geq 0$ are the forward and backward reaction rates and $\alpha^r, \beta^r \in \mathbb{N}_0^I$ are the vectors of the stoichiometric coefficients for the r^{th} reaction. Here, we use the monomial notation $z^v = \prod_{i=1}^{I} z_i^{v_i}$ for $z \in \mathbf{X}$ and $v \in \mathbb{N}_0^I$.

It was shown in [134] (see also [136]) that systems of the form (V.1.1) fulfilling a *detailed-balance condition* have a gradient structure. The condition of detailed balance for the reaction system means that there exists a positive steady state $c_{\mathrm{eq}} \in (0,\infty)^I =: \mathring{\mathbf{X}}$ such that

$$k^r_{\mathrm{fw}}(z)c_{\mathrm{eq}}^{\alpha^r} = k^r_{\mathrm{bw}}(z)c_{\mathrm{eq}}^{\beta^r} \qquad \text{for all } r \in \{1,\dots,R\} \text{ and } z \in \mathbf{X}. \tag{V.1.2}$$

Obviously, this implies the steady-state condition $R(z) = 0$ in such a way that each term in the sum in (V.1.1) is equal to zero. Each reaction is in equilibrium. For reaction rates k^r_{fw} and k^r_{bw} which do not depend on the state z, condition (V.1.2) can be recast as a linear system for the unknowns $\log c_{\mathrm{eq},i}$ (for $i = 1,\dots,I$) — the so-called *chemical potentials* — which is then called the system of *Wegscheider conditions*, see e.g. [162]. If the number R of reactions is smaller than the number I of species, these conditions can usually be satisfied easily.

Using the steady state c_{eq}, we define the *driving entropy* $\mathcal{E} : \mathbf{X} \to \mathbb{R}$ and the Onsager matrix $\mathbb{K}(z) \in \mathbb{R}^{I\times I}_{\mathrm{spsd}}$ on the state space $\mathbf{X} = [0,\infty)^I$ by

$$\mathcal{E}(z) := \sum_{i=1}^{I} c_{\mathrm{eq},i} E\big(\tfrac{z_i}{c_{\mathrm{eq},i}}\big), \quad \text{with } E(s) = \begin{cases} s(\log s - 1) + 1 & \text{if } s > 0, \\ 1 & \text{if } s = 0, \end{cases}$$

$$\mathbb{K}(z) := \sum_{r=1}^{R} k^r_{\mathrm{fw}}(z)c_{\mathrm{eq}}^{\alpha^r} \Lambda\left(\frac{z^{\alpha^r}}{c_{\mathrm{eq}}^{\alpha^r}}, \frac{z^{\beta^r}}{c_{\mathrm{eq}}^{\beta^r}}\right)(\alpha^r - \beta^r) \otimes (\alpha^r - \beta^r),$$

with

$$\Lambda(a,b) := \begin{cases} 0 & \text{if } a = 0 \text{ or } b = 0, \\ \frac{a-b}{\log a - \log b} & \text{if } a,b > 0,\ a \neq b, \\ a & \text{if } a = b > 0 \end{cases}$$

denoting the (nonnegative) *logarithmic mean* and $\gamma \otimes \gamma \in \mathbb{R}^{I\times I}$ denoting the tensor product, that is, $(\gamma \otimes \gamma)_{ij} = \gamma_i \gamma_j$.

Using the detailed-balance condition (V.1.2) and the calculation rules for the logarithm, we immediately verify

$$\dot{c}_\varepsilon = -R(c_\varepsilon) \quad \Longleftrightarrow \quad \dot{c}_\varepsilon = -\mathbb{K}(c_\varepsilon)\mathrm{D}\mathcal{E}(c_\varepsilon). \tag{V.1.3}$$

The *stoichiometric subspace* $\mathsf{S} \subset \mathbb{R}^I$ is spanned by the vectors $\gamma^r := \alpha^r - \beta^r \in \mathbb{Z}^I$, viz. $\mathsf{S} = \mathrm{span}\{\gamma^1, \ldots, \gamma^R\}$. The orthogonal complement $\mathsf{S}^\perp = \ker \mathbb{K}$ gives the invariants of the evolution: given $\zeta \in \mathsf{S}^\perp$, the quantity $\zeta^\mathrm{T} c_\varepsilon$ is conserved during the evolution.

The reactions are divided into two classes: slow reactions ($r = 1, \ldots, R_\mathrm{s}$) with rates of order $O(1)$ and fast reactions ($r = R_\mathrm{s}+1, \ldots, R_\mathrm{f}$) with rates of order $O\left(\frac{1}{\varepsilon}\right)$, as $\varepsilon \searrow 0$. Specifically, we require:

(R1) The vectors $\gamma^1, \ldots, \gamma^R$ are linearly independent.

(R2) The forward and backward reaction rates k^r_fw and k^r_bw are equal and do not depend on the state z, but depend on a small model parameter $\varepsilon > 0$ in the following way:

$$\begin{aligned}
k^r_\varepsilon &:= k^r_\mathrm{fw} = k^r_\mathrm{bw} > 0 && \text{for all } r = 1, \ldots, R;\\
k^r_\varepsilon &\in O(1) \text{ as } \varepsilon \searrow 0, \text{ with } k^r := \lim_{\varepsilon\searrow 0} k^r_\varepsilon > 0 && \text{for all } r = 1, \ldots, R_\mathrm{s};\\
k^{R_\mathrm{s}+s}_\varepsilon &\in O\left(\frac{1}{\varepsilon}\right) \text{ as } \varepsilon \searrow 0, \text{ with } \lambda^s := \lim_{\varepsilon\searrow 0} \varepsilon k^{R_\mathrm{s}+s}_\varepsilon > 0 && \text{for all } s = 1, \ldots, R_\mathrm{f}.
\end{aligned}$$

Parameters corresponding to fast reactions (i.e. $k^r_\varepsilon, \alpha^r, \beta^r, \gamma^r$) will be denoted with a tilde as follows: for $s \in \{1, \ldots, R_\mathrm{f}\}$, let e.g.

$$\widetilde{k}^s_\varepsilon := k^{R_\mathrm{s}+s}_\varepsilon.$$

Introduce for $r \in \{1, \ldots, R\}$ the auxiliary functions $\mathcal{C}^r : \mathbf{X} \to \mathbb{R}$, $\mathcal{C}^r(z) := z^{\beta^r} - z^{\alpha^r}$ and the associated algebraic constraint set $\mathsf{C}_\mathrm{f} \subset \mathbf{X}$ via

$$\mathsf{C}_\mathrm{f} = \left\{ z \in \mathbf{X} : \ \mathcal{C}^r(z) = 0 \text{ for all } r \in \{R_\mathrm{s}+1, \ldots, R\} \right\}.$$

Then, system (V.1.1) can be rewritten in the following form:

$$\dot{c}_\varepsilon(t) = \sum_{r=1}^{R_\mathrm{s}} k^r_\varepsilon \mathcal{C}^r(c_\varepsilon(t))\gamma^r + \sum_{s=1}^{R_\mathrm{f}} \widetilde{k}^s_\varepsilon \mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon(t))\widetilde{\gamma}^s, \tag{V.1.4}$$

Additionally, we impose the initial condition

$$c_\varepsilon(0) = c^0_\varepsilon \in \mathring{\mathbf{X}}, \tag{V.1.5}$$

and assume:

(I) The family $(c^0_\varepsilon)_{\varepsilon>0}$ of initial conditions in $\mathring{\mathbf{X}}$ is convergent with limit in $\mathring{\mathbf{X}}$: $\lim_{\varepsilon\searrow 0} c^0_\varepsilon = c^0 \in \mathring{\mathbf{X}}$.

In this framework, we may choose $c_{\mathrm{eq},i} = 1$ for all $i \in I$, hence, the entropy (which does not depend explicitly on the small parameter ε here) reads

$$\mathcal{E}(z) = \sum_{i=1}^{I} E(z_i).$$

According to Section I.2.3, one can write down an energy dissipation balance of the form (**EDB**), where the dual dissipation potential Ψ^*_ε is quadratic (actually, the gradient system is a classical one, cf. (I.2.23)): $\Psi^*_\varepsilon(z; \xi) = \frac{1}{2}\xi^\mathrm{T}\mathbb{K}(z)\xi$. Our main interest is concerned with the evolutionary Γ-limit of the gradient system $(\mathbf{X}, \mathcal{E}, \Psi_\varepsilon)$ in the sense given in Definition I.21. Our main result, which is proved in Chapter V.2, is summarized in the following.

Theorem V.1 (Evolutionary convergence). *Assume that* (R1)&(R2) *and* (I) *hold and let* $\overline{x} \in \mathring{\mathbf{X}}$ *be given by* $\overline{x} := \lim_{\tau\to\infty} x(\tau)$, *where* $x : [0,\infty) \to \mathring{\mathbf{X}}$ *is the solution to*

$$\dot{x}(\tau) = \sum_{s=1}^{R_\mathrm{f}} \lambda^s \mathcal{C}^{R_\mathrm{s}+s}(x(\tau))\widetilde{\gamma}^s \qquad (\tau > 0),$$
$$x(0) = c^0 \in \mathring{\mathbf{X}},$$

see Theorem V.12 below. The following statements hold:

(a) As $\varepsilon \searrow 0$, *the family of solutions* $(c_\varepsilon)_{\varepsilon>0}$ *to* (V.1.4)&(V.1.5) *converges locally uniformly in* $(0,\infty)$ *to the solution* $c : (0,\infty) \to \mathring{\mathbf{X}}$ *of the Cauchy problem*

$$\dot{c}(t) = \left[\mathbb{1} - \mathbf{G}(\mathbf{G}^\mathrm{T}\mathbf{I}(c(t))\mathbf{G})^{-1}\mathbf{G}^\mathrm{T}\mathbf{I}(c(t))\right]\left(\sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c(t))\gamma^r\right) \quad \text{for } t > 0, \tag{V.1.6}$$
$$\lim_{t\searrow 0} c(t) = \overline{x}.$$

Above, $\mathbf{G} \in \mathbb{R}^{I\times R_\mathrm{f}}$ *has columns* $\widetilde{\gamma}^1, \ldots, \widetilde{\gamma}^{R_\mathrm{f}}$ *and* $\mathbf{I}(z) \in \mathbb{R}^{I\times I}$ *is the diagonal matrix with diagonal entries* $z_1^{-1}, \ldots, z_I^{-1}$, *for* $z \in \mathring{\mathbf{X}}$.

(b) Denoting the Γ*-limits (as* $\varepsilon \searrow 0$*) of* Ψ^*_ε *and* Ψ_ε *by* Ψ^* *and* Ψ, *the following energy dissipation balance holds for all* $t \geq 0$*:*

$$\mathcal{E}(c^0) \geq \mathcal{E}(\overline{x}) = \mathcal{E}(c(t)) + \int_0^t [\Psi(c(s);\dot{c}(s)) + \Psi^*(c(s);-\mathrm{D}\mathcal{E}(c(s)))]\,\mathrm{d}s. \tag{V.1.7}$$

The limit curve c is a solution to the differential-algebraic system

$$\dot{c}(t) = \sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c(t))\gamma^r + \sum_{s=1}^{R_\mathrm{f}} \omega^s(t)\widetilde{\gamma}^s \qquad (t > 0),$$
$$0 = \mathcal{C}^{R_\mathrm{s}+s}(c(t)) \quad \forall t \geq 0, \quad \forall s \in \{1, \ldots, R_\mathrm{f}\},$$

where $\omega^1, \ldots, \omega^{R_\mathrm{f}} : (0,\infty) \to \mathbb{R}$ *are continuous functions.*

(c) On all intervals of the form $[t_0, t_1] \subset (0,\infty)$, *one has evolutionary* Γ*-convergence of gradient systems:* $(\mathbf{X}, \mathcal{E}, \Psi_\varepsilon) \xrightarrow{\mathsf{E}} (\mathbf{X}, \mathcal{E}, \Psi)$ *as* $\varepsilon \searrow 0$.

Remark V.2 (The case $t_0 = 0$). *At time* $t_0 = 0$, *an additional amount of energy is dissipated in the limit, namely the quantity* $\mathcal{E}(c^0) - \mathcal{E}(\overline{x})$. *This phenomenon can be seen as an additional external force which brings the solution instantaneously into the fast-reaction equilibrium, only acting at initial time.*

Remark V.3 (Change of variables [72]). *Instead of looking at the solution curves* c_ε *to* (V.1.4) *directly, one can also perform a change of dependent variables via the transformation* $g_\varepsilon = \mathbf{\Gamma}^{-1}c_\varepsilon$, *where the* $I \times I$*-matrix* $\mathbf{\Gamma} = \left(\gamma^1 \ \cdots \ \gamma^R \ \gamma^{R+1} \ \cdots \ \gamma^I\right)$ *is invertible, e.g. with* $\left\{\gamma^{R+1}, \ldots, \gamma^I\right\}$ *being a basis of* $\mathsf{S}^\perp$. *Consequently, one has*

$$\dot{g}_{\varepsilon,i}(t) = \begin{cases} k^i_\varepsilon \mathcal{C}^i(\mathbf{\Gamma}g_\varepsilon(t)) & \text{if } i \in \{1, \ldots, R\}, \\ 0 & \text{if } i \in \{R+1, \ldots, I\}. \end{cases} \tag{V.1.8}$$

The transformed gradient structure $(\widetilde{\mathbf{X}}, \widetilde{\mathcal{E}}, \widetilde{\Psi}^*_\varepsilon)$ *then reads*

$$\widetilde{\mathbf{X}} := \mathbf{\Gamma}^{-1}\mathbf{X}, \quad \widetilde{\mathcal{E}}(g) = \mathcal{E}(\mathbf{\Gamma}g), \quad \widetilde{\Psi}^*_\varepsilon(g;\xi) = \frac{1}{2}\xi^\mathrm{T}\widetilde{\mathbb{K}}(g)\xi,$$

with the transformed Onsager matrix $\widetilde{\mathbb{K}}(g) \in \mathbb{R}^{I\times I}$ *(for* $g \in \widetilde{\mathbf{X}}$*) defined as the* diagonal *matrix with diagonal* *entries*

$$\widetilde{\mathbb{K}}(g)_{ii} = \begin{cases} k_\varepsilon^i \Lambda((\mathbf{\Gamma} g)^{\alpha^i}, (\mathbf{\Gamma} g)^{\beta^i}) & \text{if } i \in \{1, \dots, R\}, \\ 0 & \text{if } i \in \{R+1, \dots, I\}. \end{cases}$$

Indeed, using (V.1.8) *and the chain rule* $\mathrm{D}_g \widetilde{\mathcal{E}}(g) = \mathbf{\Gamma}^{\mathrm{T}} \mathrm{D}_z \mathcal{E}(\mathbf{\Gamma} g)$*, one easily verifies that*

$$\dot{g}_\varepsilon(t) = -\mathbb{K}(g_\varepsilon(t)) \mathrm{D}_g \widetilde{\mathcal{E}}(g_\varepsilon(t).$$

Despite this remarkably simple structure of the transformed Onsager matrix $\widetilde{\mathbb{K}}(g)$*, an analysis of system* (V.1.4) *via its gradient structure might be performed more conveniently on the original c variables.*

The notion of evolutionary Γ-convergence can be used to analyse the behaviour of evolution systems of the form $\dot{u}_\varepsilon(t) = -\mathbf{F}_\varepsilon(u_\varepsilon(t))$ — usually comprising nonlinear partial differential equations — in a possibly *singular* limit $\varepsilon \searrow 0$ (for specific applications see, e.g., [160, 141, 166, 122, 7, 142, 71, 137]). In contrast to this relatively new theory, the instantaneous limit of systems of chemical reactions with mass-action law (with or without diffusion) has been studied quite extensively with classical methods from the theory of ordinary and partial differential equations not relying on a possibly present gradient structure. The purpose of this study is to demonstrate a connection of those two concepts. Specifically, our Theorem V.1 includes a related result by Bothe [30]. In constrast to [30], we aim to mainly use the gradient structure of the system. In the future, diffusion processes could be added to our system. In this direction, only results for very specific reaction systems are available, see [32, 31, 61].

For the analysis of systems of the form (V.1.1), one particular tool is provided by so-called *quasi steady state assumptions*, where one assumes that certain components of the system are in equilibrium [161, 164]. A mathematically rigorous theory is based on the theorem by Tikhonov [173] and Fenichel [84] providing criteria on the reaction terms for a dimension-reducing convergence in the singular limit (see, for example, [151, 92, 93]). With our method, however, also cases not covered by the Tikhonov-Fenichel theory can be treated.

Our strategy of proof is as follows: In Section V.2.1, we perform an exhaustive analysis of the dual and primal dissipation potentials Ψ_ε^* and Ψ_ε and derive their limits in the sense of Mosco (see Definition I.20) as a preparation for passage to the limit $\varepsilon \searrow 0$ in the energy dissipation balance (**EDB**). Afterwards, in Section V.2.2, we study the solution curves c_ε to (V.1.4)&(V.1.5) to obtain suitable ε-uniform *a priori* estimates. A cornerstone is the ε-uniform positivity of c_ε for small times (cf. Proposition V.15) which yields suitable L^∞-estimates on the *derivatives* $\dot{c}_\varepsilon$ by the investigation of the evolution over time of a quantity related to the dual dissipation potential Ψ_ε^* (see Lemma V.16 and Proposition V.17). We consequently are in position to pass to the limit $\varepsilon \searrow 0$ (see Section V.2.3). First, we obtain convergence of the family of curves $(c_\varepsilon)_{\varepsilon>0}$ on a small time interval and find the limit energy dissipation balance (V.1.7). Subsequently, we derive the differential equation (V.1.6) corresponding to (V.1.7). Finally, we show that this local-in-time solution can be extended globally in time, completing the proof of Theorem V.1.

CHAPTER V.2

Instantaneous limit of systems with slow and fast reactions

In this chapter, Theorem V.1 is proved. Our main tool for obtaining the necessary *a priori* estimates is the energy dissipation balance (**EDB**) (see Section I.2.3) associated to the gradient system $(\mathbf{X}, \mathcal{E}, \Psi_\varepsilon)$.

V.2.1. Dissipation potentials

This section is devoted to the study of the dual and primal dissipation potentials and their limits in the sense of Mosco as $\varepsilon \searrow 0$.

To simplify notation, we introduce for a given vector $z \in \mathbf{X}$ the index sets

$$J_z := \left\{ r \in \{1, \dots, R_\mathrm{s}\} \mid \Lambda(z^{\alpha^r}, z^{\beta^r}) \neq 0 \right\},$$
$$\widetilde{J}_z := \left\{ s \in \{1, \dots, R_\mathrm{f}\} \mid \Lambda(z^{\widetilde{\alpha}^s}, z^{\widetilde{\beta}^s}) \neq 0 \right\}.$$

Definition V.4 (Dual dissipation potentials). *For $\varepsilon > 0$, the* dual dissipation potential $\Psi^*_\varepsilon : \mathsf{T}^*\mathbf{X} \to [0, \infty]$ *will be denoted as follows:*

$$\begin{aligned}
\Psi^*_\varepsilon(z;\xi) &= \Psi^*_{\mathrm{s},\varepsilon}(z;\xi) + \Psi^*_{\mathrm{f},\varepsilon}(z;\xi), \qquad \textit{with} \\
\Psi^*_{\mathrm{s},\varepsilon}(z;\xi) &:= \sum_{r \in J_z} \frac{1}{2} k^r_\varepsilon \Lambda(z^{\alpha^r}, z^{\beta^r}) [\xi^\mathrm{T} \gamma^r]^2, \\
\Psi^*_{\mathrm{f},\varepsilon}(z;\xi) &:= \sum_{s \in \widetilde{J}_z} \frac{1}{2} \widetilde{k}^s_\varepsilon \Lambda(z^{\widetilde{\alpha}^s}, z^{\widetilde{\beta}^s}) [\xi^\mathrm{T} \widetilde{\gamma}^s]^2.
\end{aligned} \tag{V.2.1}$$

In this finite-dimensional framework, we can identify both the tangent and cotangent bundles $\mathsf{T}\mathbf{X}$ *and* $\mathsf{T}^*\mathbf{X}$ *with* $\mathbf{X} \times \mathbb{R}^I$.

We give an explicit characterization of the primal dissipation potential using

Definition V.5 (Infimal convolution). *Given $\Phi_1, \Phi_2 : \mathbf{Y} \to \mathbb{R}_\infty$ on a Banach space $\mathbf{Y}$, their* infimal convolution $\Phi_1 \overset{\inf}{\Delta} \Phi_2$ *is given by*

$$\Phi_1 \overset{\inf}{\Delta} \Phi_2(y) = \inf\{\Phi_1(y_1) + \Phi_2(y_2) | y = y_1 + y_2\}.$$

Recall the definition of the Legendre-Fenchel transform $\mathcal{L}$: For a map $\Phi : \mathbf{Y} \to \overline{\mathbb{R}}$ defined on a reflexive Banach space $\mathbf{Y}$, we set

$$\mathcal{L}[\Phi] : \mathbf{Y}^* \to \overline{\mathbb{R}}, \quad \mathcal{L}[\Phi](\zeta) := \sup_{y \in \mathbf{Y}} (\langle \zeta, y \rangle - \Phi(y)).$$

Lemma V.6 (Legendre transform and summation [8, Prop. 3.4]). *If Φ_1 and Φ_2 are proper, convex and lower semicontinuous and* $\mathrm{Dom}(\mathcal{L}[\Phi_1]) - \mathrm{Dom}(\mathcal{L}[\Phi_2])$ *contains a neighbourhood of the origin, then*

$$\mathcal{L}[\mathcal{L}[\Phi_1] + \mathcal{L}[\Phi_2]] = \Phi_1 \overset{\inf}{\Delta} \Phi_2. \tag{V.2.2}$$

Proposition V.7 (Primal dissipation potential). *Let $z \in \mathbf{X}$ be given. Then, the following holds for all $v \in \mathbb{R}^I$:*

$$\Psi_{\mathrm{s},\varepsilon}(z;v) := \mathcal{L}[\Psi^*_{\mathrm{s},\varepsilon}(z;\cdot)](v) = \begin{cases} \frac{1}{2}\sum_{r\in J_z} v_r^2 & \text{if there exist } v_r \in \mathbb{R}\ (r\in J_z) \text{ such that } v = \sum_{r\in J_z} v_r \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\gamma^r, \\ +\infty & \text{otherwise.}\end{cases} \tag{V.2.3}$$

$$\Psi_{\mathrm{f},\varepsilon}(z;v) := \mathcal{L}[\Psi^*_{\mathrm{f},\varepsilon}(z;\cdot)](v) = \begin{cases} \frac{1}{2}\sum_{s\in \widetilde{J}_z} \widetilde{v}_s^2 & \text{if there exist } \widetilde{v}_s \in \mathbb{R}\ (s\in \widetilde{J}_z) \text{ such that } v = \sum_{s\in \widetilde{J}_z} \widetilde{v}_s \sqrt{\Lambda(z^{\widetilde{\alpha}^s}, z^{\widetilde{\beta}^s}) \widetilde{k}^s_\varepsilon}\widetilde{\gamma}^s, \\ +\infty & \text{otherwise.}\end{cases} \tag{V.2.4}$$

$$\Psi_{\varepsilon}(z;v) = \mathcal{L}[\Psi^*_{\varepsilon}(z;\cdot)](v) = \begin{cases} \frac{1}{2}\sum_{r\in J_z} v_r^2 + \frac{1}{2}\sum_{s\in \widetilde{J}_z} \widetilde{v}_s^2 & \text{if there exist } v_r, \widetilde{v}_s \in \mathbb{R}\ (r\in J_z,\ s\in \widetilde{J}_z) \text{ such that} \\ & v = \sum_{r\in J_z} v_r \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\gamma^r + \sum_{s\in \widetilde{J}_z} \widetilde{v}_s \sqrt{\Lambda(z^{\widetilde{\alpha}^s}, z^{\widetilde{\beta}^s}) \widetilde{k}^s_\varepsilon}\widetilde{\gamma}^s, \\ +\infty & \text{otherwise.}\end{cases} \tag{V.2.5}$$

Proof. We first prove (V.2.3). Obviously, $\mathcal{L}[\Psi_{\mathrm{s},\varepsilon}(z;\cdot)](v) = +\infty$ if

$$v \notin \operatorname{span}\{\gamma^r : r \in J_z\} = \operatorname{span}\{\sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\gamma^r : r \in J_z\}.$$

Let therefore $v = \sum_{r\in J_z} v_r \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\gamma^r$ with suitable $v_r \in \mathbb{R}$, $r \in J_z$. Then, for all $\xi \in \mathbb{R}^I$:

$$\begin{aligned} \xi^{\mathrm{T}} v - \Psi^*_{\mathrm{s},\varepsilon}(z,\xi) &= \sum_{r\in J_z}\left[v_r \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\xi^{\mathrm{T}}\gamma^r - \frac{1}{2} k^r_\varepsilon \Lambda(z^{\alpha^r}, z^{\beta^r})(\xi^{\mathrm{T}}\gamma^r)^2\right] \\ &= \frac{1}{2}\sum_{r\in J_z}\left[v_r^2 - \left(v_r - \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\xi^{\mathrm{T}}\gamma^r\right)^2\right]. \end{aligned}$$

Since the γ^r are linearly independent, the claim follows by observing that there exists a solution $\xi \in \mathbb{R}^I$ to the linear system $v_r - \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r}) k^r_\varepsilon}\xi^{\mathrm{T}}\gamma^r = 0$ $(r \in J_z)$.
The proof of (V.2.4) goes along the same lines; combining (V.2.3)&(V.2.4) with formula (V.2.2) then yields (V.2.5). □

We now derive the Mosco limits of the dissipation potentials Ψ^*_ε and Ψ_ε as $\varepsilon \searrow 0$.

Definition V.8 (Limit dissipation potentials). *The* limit dual *and* primal dissipation protentials *are defined as* $\Psi^* : \mathsf{T}^*\mathbf{X} \to [0,\infty]$ *and* $\Psi : \mathsf{T}\mathbf{X} \to [0,\infty]$, *via*

$$\Psi^*(z;\xi) := \sum_{r\in J_z} \frac{1}{2} k^r \Lambda(z^{\alpha^r}, z^{\beta^r})[\xi^{\mathrm{T}}\gamma^r]^2 + \chi^*(z;\xi), \tag{V.2.6}$$

with

$$\chi^*(z;\xi) := \begin{cases} 0 & \text{if } \xi^{\mathrm{T}}\widetilde{\gamma}_s = 0 \text{ for all } s \in \widetilde{J}_z, \\ +\infty & \text{otherwise,}\end{cases} \tag{V.2.7}$$

and

$$\Psi(z;v) := \begin{cases} \frac{1}{2}\sum_{r\in J_z} v_r^2 & \text{if there exist } v_r \in \mathbb{R}\,(r\in J_z) \text{ and } \widetilde{v}\in \operatorname{span}\{\widetilde{\gamma}^s : s\in \widetilde{J}_z\} \\ & \text{such that } v = \sum_{r\in J_z} v_r \sqrt{\Lambda(z^{\alpha^r}, z^{\beta^r})k^r}\gamma^r + \widetilde{v}, \\ +\infty & \text{otherwise.} \end{cases} \tag{V.2.8}$$

For the sake of presentation, we introduce for a given family $(z_\varepsilon)_{\varepsilon>0}$ in $\mathbf{X}$ the following auxiliary proper, convex and lower semicontinuous functionals on $\mathbb{R}^I$:

$$\Phi_\varepsilon^*(\xi) := \Psi_\varepsilon^*(z_\varepsilon;\xi),$$
$$\Phi_\varepsilon(v) := \Psi_\varepsilon(z_\varepsilon;v).$$

Lemma V.9 (Mosco limit of auxiliary functionals). *Let $z_\varepsilon \to z$ as $\varepsilon \searrow 0$ for some $z\in\mathbf{X}$. Then, the families $(\Phi_\varepsilon^*)_{\varepsilon>0}$ and $(\Phi_\varepsilon)_{\varepsilon>0}$ converge in the sense of Mosco as $\varepsilon\searrow 0$ to the limit functionals $\Psi^*(z;\cdot) : \mathbb{R}^I \to [0,\infty]$ and $\Psi(z;\cdot) : \mathbb{R}^I \to [0,\infty]$, respectively.*

Proof. We first prove the Mosco-convergence of $(\Phi_\varepsilon^*)_{\varepsilon>0}$ to $\Psi^*(z;\cdot)$.

(i) *Liminf estimate.* Consider a convergent sequence $\xi_\varepsilon \to \xi$ in $\mathbb{R}^I$.
<u>Case 1:</u> $\xi^{\mathrm{T}}\widetilde{\gamma}^s = 0$ for all $s\in\widetilde{J}_z$.
From the pointwise estimate

$$\Phi_\varepsilon^*(\xi_\varepsilon) \ge \Psi_{\mathrm{s},\varepsilon}^*(z_\varepsilon;\xi_\varepsilon),$$

it is easily seen that

$$\liminf_{\varepsilon\searrow 0}\Phi_\varepsilon^*(\xi_\varepsilon) \ge \lim_{\varepsilon\searrow 0}\Psi_{\mathrm{s},\varepsilon}^*(z_\varepsilon;\xi_\varepsilon) = \sum_{r=1}^{R_{\mathrm{s}}}\frac{1}{2}k^r\Lambda(z^{\alpha^r},z^{\beta^r})[\xi^{\mathrm{T}}\gamma^r]^2.$$

<u>Case 2:</u> There exists $s_*\in\widetilde{J}_z$ such that $\xi^{\mathrm{T}}\widetilde{\gamma}^{s_*}\neq 0$.
Then the quantity

$$\frac{1}{2}\widetilde{k}_\varepsilon^{s_*}\Lambda(z_\varepsilon^{\widetilde{\alpha}^{s_*}}, z_\varepsilon^{\widetilde{\beta}^{s_*}})(\xi_\varepsilon{}^{\mathrm{T}}\widetilde{\gamma}^{s_*})^2$$

diverges to $+\infty$ as $\varepsilon\searrow 0$, thus showing that

$$\liminf_{\varepsilon\searrow 0}\Phi_\varepsilon^*(\xi_\varepsilon) = +\infty.$$

(ii) *Recovery sequences.* Let $\xi\in\mathbb{R}^I$ be given. One easily verifies that the constant sequence $\widehat{\xi}_\varepsilon := \xi$ is a recovery sequence.

The Mosco-convergence of $(\Phi_\varepsilon)_{\varepsilon>0}$ to the limit $\mathcal{L}[\Psi^*(z;\cdot)]$ is a consequence of [8, Thm. 3.7]. We first determine the Legrendre transform of $\xi\mapsto\chi^*(z;\xi)$:

$$\mathcal{L}[\chi^*(z;\cdot)](v) = \sup\{\xi^{\mathrm{T}}v : \xi\in\mathbb{R}^I,\ \xi^{\mathrm{T}}\widetilde{\gamma}^s = 0\ \forall s\in\widetilde{J}_z\}$$
$$= \begin{cases} 0 & \text{if } v\in\operatorname{span}\{\widetilde{\gamma}^s : s\in\widetilde{J}_z\}, \\ +\infty & \text{otherwise.}\end{cases}$$

A similar calculation as in the proof of Proposition V.7 shows, with the help of (V.2.2), that $\mathcal{L}[\Psi^*(z;\cdot)] = \Psi(z;\cdot)$ holds. □

Proposition V.10 (Mosco limit of dual and primal dissipation potentials). *The families of dual and primal dissipation potentials $(\Psi_\varepsilon^*)_{\varepsilon>0}$ and $(\Psi_\varepsilon)_{\varepsilon>0}$ converge in the sense of Mosco as $\varepsilon\searrow 0$ to their respective limits Ψ^* and Ψ.*

PROOF. This is immediately obtained from Lemma V.9. □

V.2.2. A priori estimates

For fixed $\varepsilon > 0$, the dynamics of (V.1.4) are well-known (cf. e.g. [81, 99, 180, 30]). The following two results are essentially contained in [30, Thm. 1] and are at the basis of our subsequent analysis.

THEOREM V.11 (Dynamics of (V.1.4)). *Let $\varepsilon > 0$ be fixed, but arbitrary. The following statements hold:*

(a) For each initial condition $c_\varepsilon^0 \in \mathring{\mathbf{X}}$, there exists a unique steady state $\overline{c}_\varepsilon \in (c_\varepsilon^0 + \mathsf{S}) \cap \mathring{\mathbf{X}}$ of the dynamical system induced by (V.1.4).

(b) The solution $c_\varepsilon \in C^\infty([0,\infty);\mathbb{R}^I)$ to the initial-value problem (V.1.4)&(V.1.5) *exists globally in time, it is unique and it satisfies $c_\varepsilon(t) \in \mathring{\mathbf{X}}$ for all $t \geq 0$.*

(c) With $\overline{c}_\varepsilon$ from (a), the map

$$\mathcal{E}(\cdot;\overline{c}_\varepsilon) : \mathring{\mathbf{X}} \to \mathbb{R}, \quad \mathcal{E}(z;\overline{c}_\varepsilon) := \sum_{i=1}^{I} \overline{c}_{\varepsilon,i} E\left(\frac{z_i}{\overline{c}_{\varepsilon,i}}\right) = \mathcal{E}(z) + \sum_{i=1}^{I} \left[\overline{c}_{\varepsilon,i} - z_i \log(\overline{c}_{\varepsilon,i}) - 1\right], \tag{V.2.9}$$

is a Lyapunov function for (V.1.4). *Moreover, one has*

$$\lim_{t\to\infty} c_\varepsilon(t) = \overline{c}_\varepsilon \quad \text{and} \quad \lim_{t\to\infty} \mathcal{E}(c_\varepsilon(t);\overline{c}_\varepsilon) = 0.$$

For later reference, we state an analogous theorem for the system governed by fast reactions only:

THEOREM V.12 (System of fast reactions). *Consider the initial-value problem*

$$\begin{aligned} \dot{x}(\tau) &= \sum_{s=1}^{R_\mathrm{f}} \lambda^s \mathcal{C}^{R_\mathrm{s}+s}(x(\tau))\widetilde{\gamma}^s \qquad (\tau > 0), \\ x(0) &= x^0 \in \mathring{\mathbf{X}}. \end{aligned} \tag{V.2.10}$$

The following statements hold:

(a) For every $x^0 \in \mathring{\mathbf{X}}$, the dynamical system induced by (V.2.10) *possesses a unique steady state $\overline{x} \in (x^0 + \mathrm{span}\{\widetilde{\gamma}^1, \ldots, \widetilde{\gamma}^{R_\mathrm{f}}\}) \cap \mathring{\mathbf{X}}$.*

(b) The solution $x \in C^\infty([0,\infty);\mathbb{R}^I)$ to (V.2.10) *exists globally in time, it is unique and it satisfies $x(\tau) \in \mathring{\mathbf{X}}$ for all $\tau \geq 0$.*

(c) With $\overline{x}$ from (a), the map $\mathcal{E}(\cdot,\overline{x})$ defined in analogy to (V.2.9) *from Theorem V.11 is a Lyapunov function for the dynamical system induced by* (V.2.10) *and one has*

$$\lim_{\tau\to\infty} x(\tau) = \overline{x} \quad \text{and} \quad \lim_{\tau\to\infty} \mathcal{E}(x(\tau);\overline{x}) = 0.$$

Note that the differential equation in (V.2.10) can be obtained from (V.1.4) by the linear rescaling of time $\tau := t\varepsilon^{-1}$ in the limit $\varepsilon \searrow 0$. The new time scale τ is referred to as *fast time scale*. Its introduction allows for a closer investigation of the region $t \approx 0$. More details on asymptotic analysis in the context of reaction kinetics can be found e.g. in [147, Ch. 6].

We define for $z \in \mathring{\mathbf{X}}$ and $\zeta \in \mathbb{R}^I$ the *total dissipation functional*

$$\mathcal{M}_\varepsilon(z;\zeta) = \Psi_\varepsilon(z;\zeta) + \Psi_\varepsilon^*(z;-\mathrm{D}\mathcal{E}(z)).$$

With this, we can rewrite the evolution in (V.1.4) in the form as in **(EDB)**:

$$\mathcal{E}(c_\varepsilon(T)) + \int_0^T \mathcal{M}_\varepsilon(c_\varepsilon(t);\dot{c}_\varepsilon(t))\,\mathrm{d}t = \mathcal{E}(c_\varepsilon(0)), \tag{V.2.11}$$

at each $T > 0$.

Lemma V.13. *For all $t > 0$, one has*

$$\mathcal{M}_\varepsilon(c_\varepsilon(t);\dot{c}_\varepsilon(t)) = 2\Psi_\varepsilon(c_\varepsilon(t);\dot{c}_\varepsilon(t)). \tag{V.2.12}$$

Proof. For brevity, we omit the argument t in the following. By applying the rules for the logarithm, we easily see that

$$\Psi^*_\varepsilon(c_\varepsilon;-\mathrm{D}\mathcal{E}(c_\varepsilon)) = \sum_{r\in J_{c_\varepsilon}} \frac{1}{2}k^r_\varepsilon \frac{\mathcal{C}^r(c_\varepsilon)^2}{\Lambda(c^{\alpha^r},c^{\beta^r})} + \sum_{s\in\widetilde{J}_{c_\varepsilon}} \frac{1}{2}\widetilde{k}^s_\varepsilon \frac{\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)^2}{\Lambda(c^{\widetilde{\alpha}^s},c^{\widetilde{\beta}^s})}.$$

Thanks to (V.1.4), we observe that

$$\dot{c}_\varepsilon = \sum_{r\in J_{c_\varepsilon}} v_r\sqrt{\Lambda(c_\varepsilon^{\alpha^r},c_\varepsilon^{\beta^r})k^r_\varepsilon}\gamma^r + \sum_{s\in\widetilde{J}_{c_\varepsilon}} \widetilde{v}_s\sqrt{\Lambda(c_\varepsilon^{\widetilde{\alpha}^s},c_\varepsilon^{\widetilde{\beta}^s})\widetilde{k}^s_\varepsilon}\widetilde{\gamma}^s,$$

with

$$v_r = \sqrt{\frac{k^r_\varepsilon}{\Lambda(c_\varepsilon^{\alpha^r},c_\varepsilon^{\beta^r})}}\mathcal{C}^r(c_\varepsilon) \quad \text{for } r\in J_{c_\varepsilon},$$

$$\widetilde{v}_s = \sqrt{\frac{\widetilde{k}^s_\varepsilon}{\Lambda(c_\varepsilon^{\widetilde{\alpha}^s},c_\varepsilon^{\widetilde{\beta}^s})}}\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon) \quad \text{for } s\in\widetilde{J}_{c_\varepsilon}.$$

Hence, we have $\Psi_\varepsilon(c_\varepsilon,\dot{c}_\varepsilon) = \Psi^*_\varepsilon(c_\varepsilon;-\mathrm{D}\mathcal{E}(c_\varepsilon))$. □

We will now derive suitable *a priori* estimates on $(c_\varepsilon)_{\varepsilon>0}$ by means of (V.1.4)&(V.2.11):

Proposition V.14 (*A priori* estimates). *Let c_ε be the solution to* (V.1.4)&(V.1.5)*. Then*

$$\mathbf{C} := \sup_{\varepsilon>0}\|c_\varepsilon\|_{L^\infty([0,\infty);\mathbb{R}^I)} < \infty, \tag{V.2.13}$$

thus the family of solutions $(c_\varepsilon)_{\varepsilon>0}$ to (V.1.4)&(V.1.5) *is uniformly bounded by a constant independent of ε. Moreover, for all $s\in\{1,\dots,R_\mathrm{f}\}$:*

$$\sup_{\varepsilon>0}\sqrt{\widetilde{k}^s_\varepsilon}\|\mathcal{C}^{R_\mathrm{s}+s}\circ c_\varepsilon\|_{L^2([0,\infty))} < \infty. \tag{V.2.14}$$

Hence, $\mathcal{C}^{R_\mathrm{s}+s}\circ c_\varepsilon \to 0$ in $L^2([0,\infty))$ as $\varepsilon\searrow 0$ for each $s\in\{1,\dots,R_\mathrm{f}\}$.

Proof. For the proof of (V.2.13), we combine (V.2.11) with the elementary estimate

$$E(s) \geq (\sqrt{s}-1)^2 \quad \text{for } s\geq 0$$

to obtain for arbitrary $t\geq 0$:

$$+\infty > \sup_{\varepsilon>0}\mathcal{E}(c^0_\varepsilon) \geq \sup_{\varepsilon>0}\mathcal{E}(c_\varepsilon(t)) \geq \sup_{\varepsilon>0}\sum_{i=1}^I\left(\sqrt{c_{\varepsilon,i}(t)}-1\right)^2,$$

from which (V.2.13) obviously follows.

The second estimate (V.2.14) can also be obtained from (V.2.11), using that the uniform bound (V.2.13) on $c_\varepsilon(t)$ implies $\Lambda(c_\varepsilon(t)^{\alpha^s},c_\varepsilon(t)^{\beta^s}) \leq C$ for some constant C independent of t, s and ε:

$$+\infty > \sup_{\varepsilon>0}\mathcal{E}(c^0_\varepsilon) \geq \sup_{\varepsilon>0}\int_0^T \Psi_\varepsilon(c_\varepsilon(t),\dot{c}_\varepsilon(t))\,\mathrm{d}t \geq C\sup_{\varepsilon>0}\int_0^T \sum_{s\in\widetilde{J}_{c_\varepsilon(t)}} \widetilde{k}^s_\varepsilon\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon(t))^2\,\mathrm{d}t. \tag{V.2.15}$$

Using that $c_\varepsilon(t) \in \mathring{\mathbf{X}}$ for all $t \geq 0$ thanks to Theorem V.11, we infer that $\widetilde{J}_{c_\varepsilon(t)} = \{1, \dots, R_\mathrm{f}\}$ for all t and therefore

$$\int_0^T \sum_{s \in \widetilde{J}_{c_\varepsilon(t)}} \widetilde{k}_\varepsilon^s \mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon(t))^2 \, \mathrm{d}t = \int_0^T \sum_{s=1}^{R_\mathrm{f}} \widetilde{k}_\varepsilon^s \mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon(t))^2 \, \mathrm{d}t.$$

Letting $T \to \infty$ in (V.2.15) yields (V.2.14). Since $\widetilde{k}_\varepsilon^s \to +\infty$ as $\varepsilon \searrow 0$, one has $\mathcal{C}^{R_\mathrm{s}+s} \circ c_\varepsilon \to 0$ in $L^2([0,\infty))$ as $\varepsilon \searrow 0$ for each $s \in \{1, \dots, R_\mathrm{f}\}$. □

We now prove ε-uniform positivity of the solution c_ε on a (possibly small) time interval. This lower bound will allow us to prove a suitable uniform estimate on the derivatives of c_ε.

Proposition V.15 (Uniform positivity). *There exist $\varepsilon_0 > 0$, $T_+ > 0$ and $\delta > 0$ such that*

$$c_{\varepsilon,i}(t) \geq \delta \quad \text{for all } i \in \{1, \dots, I\}, \text{ all } t \in [0, T_+] \text{ and all } \varepsilon \leq \varepsilon_0. \tag{V.2.16}$$

Proof. Our method of proof is an adaptation of the respective proof in [30] (see also [180]) to the situation at hand. First, we introduce a new family of dependent variables $(x_\varepsilon)_{\varepsilon>0}$ by the linear rescaling of time $\tau := t\varepsilon^{-1}$, i.e. $x_\varepsilon(\tau) := c_\varepsilon(\varepsilon\tau)$. It is easily shown that x_ε satisfies

$$\dot{x}_\varepsilon(\tau) = \sum_{r=1}^{R_\mathrm{s}} \varepsilon k_\varepsilon^r \mathcal{C}^r(x_\varepsilon(\tau))\gamma^r + \sum_{s=1}^{R_\mathrm{f}} \varepsilon \widetilde{k}_\varepsilon^s \mathcal{C}^{R_\mathrm{s}+s}(x_\varepsilon(\tau))\widetilde{\gamma}^s \tag{V.2.17}$$

for $\tau > 0$, together with $x_\varepsilon(0) = c_\varepsilon^0$. Using (R2) and (I), we infer that $x_\varepsilon \to x$ in $C([0,T]; \mathbb{R}^I)$ as $\varepsilon \searrow 0$ (for fixed, but arbitrary $T > 0$), where x solves

$$\dot{x}(\tau) = \sum_{s=1}^{R_\mathrm{f}} \lambda^s \mathcal{C}^{R_\mathrm{s}+s}(x(\tau))\widetilde{\gamma}^s, \qquad x(0) = c^0.$$

From Theorem V.12, we know that there exists $\overline{x} \in \mathring{\mathbf{X}}$ such that $x(\tau) \to \overline{x}$ and $\mathcal{E}(x(\tau); \overline{x}) \downarrow 0$ as $\tau \uparrow \infty$. By continuity, there exist $\varepsilon_0 > 0$, $\tau_+ > 0$ and $\eta > 0$ such that

$$\begin{aligned} x_{\varepsilon,i}(\tau) &\geq \eta \quad \text{for all } \varepsilon \leq \varepsilon_0, \text{ all } \tau \in [0, \tau_+] \text{ and all } i \in \{1, \dots, I\}; \\ \mathcal{E}(x_\varepsilon(\tau_+); \overline{x}) &\leq \frac{1}{2} \min_i \overline{x}_i \quad \text{for all } \varepsilon \leq \varepsilon_0. \end{aligned}$$

Rewriting in terms of c_ε means

$$\begin{aligned} c_{\varepsilon,i}(t) &\geq \eta \quad \text{for all } \varepsilon \leq \varepsilon_0, \text{ all } t \in [0, \varepsilon\tau_+] \text{ and all } i \in \{1, \dots, I\}; \\ \mathcal{E}(c_\varepsilon(\varepsilon\tau_+); \overline{x}) &\leq \frac{1}{2} \min_i \overline{x}_i \quad \text{for all } \varepsilon \leq \varepsilon_0. \end{aligned}$$

Since $\mathcal{E}(c_\varepsilon(\cdot))$ is nonincreasing, we also have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(c_\varepsilon(t); \overline{x}) \leq -\sum_{i=1}^{I} \log(\overline{x}_i)\dot{c}_{\varepsilon,i} = \sum_{i=1}^{I} \sum_{r=1}^{R_\mathrm{s}} k_\varepsilon^r \mathcal{C}^r(c_\varepsilon(t)) \log(\overline{x}_i)\gamma_i^r + \sum_{s=1}^{R_\mathrm{f}} \widetilde{k}_\varepsilon^s \mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon(t)) \sum_{i=1}^{I} \log(\overline{x}_i)\widetilde{\gamma}_i^s.$$

Using (V.2.13), we see that the first double sum above is bounded by a constant $K_0 > 0$ independent of ε and t. The second part is equal to zero since for every $s \in \{1, \dots, R_\mathrm{f}\}$, one has

$$\sum_{i=1}^{I} \log(\overline{x}_i)\widetilde{\gamma}_i^s = \log(\overline{x}^{\widetilde{\gamma}^s}) = 0, \tag{V.2.18}$$

because $\overline{x} \in \mathring{\mathbf{X}}$ is a steady state of system (V.2.10), so $\mathcal{C}^{R_\mathrm{s}+s}(\overline{x}) = 0$ and hence $\overline{x}^{\widetilde{\gamma}^s} = 1$ holds for each $s \in \{1, \dots, R_\mathrm{f}\}$. All in all, we have proved that $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(c_\varepsilon(t); \overline{x}) \leq K_0$ for all $t \geq 0$ and all $\varepsilon > 0$. Consequently, this implies for $T_+ := \frac{1}{2}\left(\frac{1}{4K_0} \min_i \overline{x}_i + \varepsilon_0\tau_+\right)$ — after possibly diminishing ε_0 further —

that $\mathcal{E}(c_\varepsilon(t);\overline{x}) \leq \frac{3}{4}\min_i \overline{x}_i$ for all $t \in [\varepsilon\tau_+, T_+]$ and all $\varepsilon \leq \varepsilon_0$. Inserting the definition of $\mathcal{E}(c_\varepsilon(t);\overline{x})$, we obtain for every $i \in \{1,\ldots,I\}$ that

$$\frac{c_{\varepsilon,i}(t)}{\overline{x}_i}\log\left(\frac{c_{\varepsilon,i}(t)}{\overline{x}_i}\right) - \frac{c_{\varepsilon,i}(t)}{\overline{x}_i} + \frac{1}{4} \leq 0 \quad \forall t \in [\varepsilon\tau_+, T_+] \quad \forall \varepsilon \leq \varepsilon_0.$$

Elementary calculations then show $\frac{c_{\varepsilon,i}(t)}{\overline{x}_i} \geq \zeta$ for the constant $\zeta = \min\{\rho > 0 : \ E(\rho) = \frac{3}{4}\} > 0$, so $c_{\varepsilon,i}(t) \geq \zeta \min_l \overline{x}_l$ for all $t \in [\varepsilon\tau_+, T_+]$ and all $\varepsilon \leq \varepsilon_0$. Putting $\delta := \min(\eta, \zeta\min_l \overline{x}_l)$ yields the claim. □

We want to establish an ε-uniform bound on $\dot{c}_\varepsilon$ under the assumption of uniform positivity as in (V.2.16). To this end, we perform a *Bakry-Emery*-type argument for the "fast" part of the dual dissipation potential $\Psi^*_{\mathrm{f},\varepsilon}$ along solutions to (V.1.4). In view of Bothe's proof [30], it seems to be convenient — at least for $R_{\mathrm{f}} > 1$ — not to use $\Psi^*_{\mathrm{f},\varepsilon}$ directly in order to avoid technicalities.

Lemma V.16 (Dissipation estimate for auxiliary map). *Assume that there exist $\varepsilon_0' > 0$, $T_+' > 0$ and $\delta' > 0$ such that*

$$c_{\varepsilon,i}(t) \geq \delta' \quad \textit{for all } i \in \{1,\ldots,I\},\ \textit{all } t \in [0,T_+']\ \textit{and all } \varepsilon \leq \varepsilon_0'. \tag{V.2.19}$$

Define the map

$$\mathcal{D} : [0,\infty) \to [0,\infty),\ \mathcal{D}(t) := \sum_{s=1}^{R_{\mathrm{f}}} \varepsilon\widetilde{k}^s_\varepsilon \mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)^2 c_\varepsilon(t)^{-\widetilde{\alpha}^s}.$$

Then, there exist $C_0, C_1, C_2 > 0$ such that for all $t \in [0, T_+']$ and all $\varepsilon \leq \varepsilon_0$:

$$\dot{\mathcal{D}}(t) \leq -\frac{C_0}{\varepsilon}\mathcal{D}(t) + \frac{C_1}{\varepsilon}\mathcal{D}(t)^{3/2} + C_2\mathcal{D}(t)^{1/2}. \tag{V.2.20}$$

Proof. We omit the argument t for the sake of brevity. As a preparation, observe that for $z \in \mathring{\mathbf{X}}$, $i \in \{1,\ldots,I\}$ and $s \in \{1,\ldots,R_{\mathrm{f}}\}$, one has

$$\begin{aligned}
\partial_{z_i} z^{\widetilde{\alpha}^s} &= \frac{\widetilde{\alpha}^s_i}{z_i} z^{\widetilde{\alpha}^s}, \\
\partial_{z_i}\mathcal{C}^{R_{\mathrm{s}}+s}(z) &= \frac{1}{z_i}(\widetilde{\beta}^s_i z^{\widetilde{\beta}^s} - \widetilde{\alpha}^s_i z^{\widetilde{\alpha}^s}) = \frac{1}{z_i}(\widetilde{\beta}^s_i \mathcal{C}^{R_{\mathrm{s}}+s}(z) - \widetilde{\gamma}^s_i z^{\widetilde{\alpha}^s}).
\end{aligned} \tag{V.2.21}$$

Differentiation of $\mathcal{D}$ w.r.t. t eventually yields

$$\dot{\mathcal{D}} = \sum_{i=1}^{I}\left\{\sum_{s=1}^{R_{\mathrm{f}}}\frac{\varepsilon\widetilde{k}^s_\varepsilon}{c_{\varepsilon,i}}\left[(2\widetilde{\beta}^s_i - \widetilde{\alpha}^s_i)c_\varepsilon^{-\widetilde{\alpha}^s}\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)^2 - 2\widetilde{\gamma}^s_i\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)\right]\left(\sum_{r=1}^{R_{\mathrm{s}}} k^r_\varepsilon \mathcal{C}^r(c_\varepsilon)\gamma^r_i + \sum_{s'=1}^{R_{\mathrm{f}}}\widetilde{k}^{s'}_\varepsilon\mathcal{C}^{R_{\mathrm{s}}+s'}(c_\varepsilon)\widetilde{\gamma}^{s'}_i\right)\right\}.$$

We estimate terms by means of (V.2.13) and (V.2.19). First, it is easy to see that

$$\sum_{i=1}^{I}\sum_{s=1}^{R_{\mathrm{f}}}\sum_{r=1}^{R_{\mathrm{s}}}\frac{\varepsilon\widetilde{k}^s_\varepsilon}{c_{\varepsilon,i}}(2\widetilde{\beta}^s_i - \widetilde{\alpha}^s_i)c_\varepsilon^{-\widetilde{\alpha}^s}\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)^2 k^r_\varepsilon\mathcal{C}^r(c_\varepsilon)\gamma^r_i \leq C_{2,0}\mathcal{D} \leq C_{2,1}\mathcal{D}^{1/2}.$$

Second, by insertion of appropriate terms, we have

$$\begin{aligned}
&-\sum_{i=1}^{I}\sum_{s=1}^{R_{\mathrm{f}}}\sum_{r=1}^{R_{\mathrm{s}}} 2\frac{\varepsilon\widetilde{k}^s_\varepsilon}{c_{\varepsilon,i}}\widetilde{\gamma}^s_i\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)k^r_\varepsilon\mathcal{C}^r(c_\varepsilon)\gamma^r_i \leq \sum_{i=1}^{I}\sum_{s=1}^{R_{\mathrm{f}}}\sum_{r=1}^{R_{\mathrm{s}}} 2|\gamma^r_i\mathcal{C}^r(c_\varepsilon)|k^r_\varepsilon\sqrt{\varepsilon\widetilde{k}^s_\varepsilon\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)^2 c_\varepsilon^{-\widetilde{\alpha}^s}}\sqrt{\varepsilon\widetilde{k}^s_\varepsilon c_\varepsilon^{\widetilde{\alpha}^s}} \\
&\leq C_{2,2}\left(\sum_{s=1}^{R_{\mathrm{f}}}\sqrt{\varepsilon\widetilde{k}^s_\varepsilon\mathcal{C}^{R_{\mathrm{s}}+s}(c_\varepsilon)^2 c_\varepsilon^{-\widetilde{\alpha}^s}}^{\,2}\right)^{1/2} = C_{2,2}\mathcal{D}^{1/2},
\end{aligned}$$

the last estimate being a consequence of Hölder's inequality.

Similarly, we obtain

$$\sum_{i=1}^{I}\sum_{s=1}^{R_\mathrm{f}}\sum_{s'=1}^{R_\mathrm{f}}\frac{\varepsilon\widetilde{k}_\varepsilon^s}{c_{\varepsilon,i}}(2\widetilde{\beta}_i^s-\widetilde{\alpha}_i^s)c_\varepsilon^{-\widetilde{\alpha}^s}\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)^2\widetilde{k}_\varepsilon^{s'}\mathcal{C}^{R_\mathrm{s}+s'}(c_\varepsilon)\widetilde{\gamma}_i^{s'}$$
$$\leq\frac{C_{1,0}}{\varepsilon}\sum_{s=1}^{R_\mathrm{f}}\varepsilon\widetilde{k}_\varepsilon^s c_\varepsilon^{-\widetilde{\alpha}^s}\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)^2\sum_{s'=1}^{R_\mathrm{f}}\varepsilon\widetilde{k}_\varepsilon^{s'}|\mathcal{C}^{R_\mathrm{s}+s'}(c_\varepsilon)|\leq\frac{C_{1,1}}{\varepsilon}\mathcal{D}^{3/2},$$

again by Hölder's inequality. For the last term, we introduce for $s\in\{1,\dots,R_\mathrm{f}\}$ the vectors $w_\varepsilon^s\in\mathbb{R}^I$ via $w_{\varepsilon,i}:=c_{\varepsilon,i}^{-1/2}\widetilde{\gamma}_i^s$ and observe

$$-2\sum_{i=1}^{I}\sum_{s=1}^{R_\mathrm{f}}\sum_{s'=1}^{R_\mathrm{f}}\frac{\varepsilon\widetilde{k}_\varepsilon^s}{c_{\varepsilon,i}}\widetilde{\gamma}_i^s\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)\widetilde{k}_\varepsilon^{s'}\mathcal{C}^{R_\mathrm{s}+s'}(c_\varepsilon)\widetilde{\gamma}_i^{s'}=-\frac{2}{\varepsilon}\left|\sum_{s=1}^{R_\mathrm{f}}\varepsilon\widetilde{k}_\varepsilon^s\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)w_\varepsilon^s\right|^2.$$

Since the vectors $\widetilde{\gamma}^1,\dots,\widetilde{\gamma}^{R_\mathrm{f}}$ are linearly independent by assumption (R1), and thanks to (V.2.19), there exists $\rho>0$ *independent of* ε *and* t such that

$$\left|\sum_{s=1}^{R_\mathrm{f}}\varepsilon\widetilde{k}_\varepsilon^s\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)w_\varepsilon^s\right|^2\geq\rho\sum_{s=1}^{R_\mathrm{f}}\left[\varepsilon\widetilde{k}_\varepsilon^s\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)\right]^2.$$

By assumption (R2) and (V.2.19), we then obtain

$$-\frac{2}{\varepsilon}\left|\sum_{s=1}^{R_\mathrm{f}}\varepsilon\widetilde{k}_\varepsilon^s\mathcal{C}^{R_\mathrm{s}+s}(c_\varepsilon)w_\varepsilon^s\right|^2\leq-\frac{C_0}{\varepsilon}\mathcal{D},$$

which completes the proof. □

We now are in position to prove the central estimate on $(c_\varepsilon)_{\varepsilon>0}$:

Proposition V.17 (Estimate on $\dot{c}_\varepsilon$). *Assume that estimate* (V.2.19) *from Lemma V.16 is fulfilled. The following statements hold:*

$$\sup_{\varepsilon\in(0,\varepsilon_0']}\|\dot{c}_\varepsilon\|_{L^1([0,T_+'];\mathbb{R}^I)}<\infty,\tag{V.2.22}$$
$$\|\dot{c}_\varepsilon\|_{L^\infty([T_0,T_+'];\mathbb{R}^I)}\leq C_{T_0}\quad\forall\varepsilon\leq\varepsilon_0',\tag{V.2.23}$$

for fixed, but arbitrary $T_0\in(0,T_+')$.

Proof. For all $i\in\{1,\dots,I\}$ and all $t\in[0,T_+']$, the differential equation (V.1.4), estimate (V.2.19) and Hölder's inequality yield

$$|\dot{c}_{\varepsilon,i}(t)|\leq K_0+\frac{K_1}{\varepsilon}\mathcal{D}(t)^{1/2},$$

with suitable ε-uniform constants $K_0,K_1>0$. Since $\mathcal{D}(t)=0$ for some $t\in[0,T_+']$ necessarily implies $\dot{\mathcal{D}}(t)=0$, we infer from Lemma V.16 that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{D}^{1/2}(t)=\frac{1}{2}\mathcal{D}(t)^{-1/2}\dot{\mathcal{D}}(t)\leq-\frac{C_0}{2\varepsilon}\mathcal{D}(t)^{1/2}+\frac{C_1}{2\varepsilon}\mathcal{D}(t)+\frac{C_2}{2}=\left[-\frac{C_0}{2\varepsilon}+\frac{C_1}{2\varepsilon}\mathcal{D}^{1/2}(t)\right]\mathcal{D}^{1/2}(t)+\frac{C_2}{2},$$

for almost every $t\in[0,T_+']$. Hence, one obtains with Gronwall's lemma:

$$\begin{aligned}\mathcal{D}(t)^{1/2}\leq{}&\mathcal{D}(0)^{1/2}\exp\left(\frac{C_1}{2\varepsilon}\int_0^t\mathcal{D}(s)^{1/2}\,\mathrm{d}s-\frac{C_0}{2\varepsilon}t\right)\\&+\frac{C_2}{2}\int_0^t\exp\left(\frac{C_1}{2\varepsilon}\int_s^t\mathcal{D}(\sigma)^{1/2}\,\mathrm{d}\sigma-\frac{C_0}{2\varepsilon}(t-s)\right)\mathrm{d}s.\end{aligned}\tag{V.2.24}$$

Recall that, thanks to (V.2.14) and (V.2.19), we have for all $t \in [0, T'_+]$:

$$\int_0^t \mathcal{D}(s)\,\mathrm{d}s \leq \widetilde{C}\varepsilon \|\mathcal{C}^{R_\mathrm{s}+s} \circ c_\varepsilon\|^2_{L^2([0,\infty))} \leq \varepsilon \widetilde{C}' \sup_{\varepsilon>0} \mathcal{E}(c^0_\varepsilon). \tag{V.2.25}$$

Young's inequality and (V.2.25) yield the existence of a constant $C > 0$ depending on T'_+ such that

$$\frac{C_1}{2\varepsilon}\int_s^t \mathcal{D}(\sigma)^{1/2}\,\mathrm{d}\sigma - \frac{C_0}{2\varepsilon}(t-s) \leq -\frac{C_0}{4\varepsilon}(t-s) + C,$$

for all $0 \leq s \leq t \leq T'_+$. Inserting this into (V.2.24), we obtain for all $t \in [0, T'_+]$

$$\begin{aligned} \mathcal{D}(t)^{1/2} &\leq \mathcal{D}(0)^{1/2}\exp(Ct)\exp\left(-\frac{C_0}{4\varepsilon}t\right) + \frac{C_2}{2}\int_0^t \exp(C(t-s))\exp\left(-\frac{C_0}{4\varepsilon}(t-s)\right)\mathrm{d}s \\ &\leq \mathcal{D}(0)^{1/2}\exp(CT'_+)\exp\left(-\frac{C_0}{4\varepsilon}t\right) + \frac{2C_2}{C_0}\exp(CT'_+)\varepsilon. \end{aligned}$$

Hence, for some constants $K_2, K_3, K_4 > 0$:

$$|\dot{c}_{\varepsilon,i}(t)| \leq K_2 + \frac{K_3}{\varepsilon}\exp\left(-\frac{K_4}{\varepsilon}t\right).$$

Clearly, this implies the desired ε-uniform bounds on the L^1 and L^∞ norm, respectively. □

Note that the result of Proposition V.17 does not extend to the case $T_0 = 0$. As a matter of fact, (V.1.4) exhibits an initial layer as $\varepsilon \searrow 0$ resulting in an ε-unbounded initial slope $|\dot{c}_\varepsilon(0)|$. Despite that, we are able to prove local uniform convergence of the curves c_ε in $(0, T_+]$ and find the one-sided limit of the limit curve at $t = 0$ afterwards.

V.2.3. Convergence to the evolutionary limit

The *a priori* estimates from the previous section now allow us to pass to the limit $\varepsilon \searrow 0$. First, we shall do so for the family of curves $(c_\varepsilon)_{\varepsilon>0}$ on the small time interval given by (V.2.16). Second, we consider the evolutionary limit as $\varepsilon \searrow 0$ of the gradient systems $(\mathbf{X}, \mathcal{E}, \Psi_\varepsilon)$ introduced above to obtain an energy dissipation balance and the corresponding rate equation in the limit. We then extend this local-in-time solution to a global-in-time solution and pass to the limit as $\varepsilon \searrow 0$ globally in time.

Proposition V.18 (Convergence of solutions for small times). *Let a sequence $\varepsilon_k \searrow 0$ be given. If* (R1)&(R2) *and* (I) *hold, there exist a subsequence (non-relabelled) and a curve $c^\sharp : [0, T_+] \to \mathbf{X}$ such that the sequence $(c_{\varepsilon_k})_{k\in\mathbb{N}}$ of solutions to* (V.1.4)*&*(V.1.5) *converges to $c^\sharp$ as $k \to \infty$ in the following sense:*

$$\begin{aligned} c_{\varepsilon_k}(t) &\to c^\sharp(t) \quad \text{for all } t \in [0, T_+], \\ c_{\varepsilon_k} &\to c^\sharp \quad \text{locally uniformly on } (0, T_+], \\ \dot{c}_{\varepsilon_k} &\rightharpoonup \dot{c}^\sharp \quad \text{weakly in } L^2([T_0, T_+]; \mathbb{R}^I) \quad \text{for each } T_0 > 0. \end{aligned}$$

The limit curve $c^\sharp$ is equal to an L^1-absolutely continuous function everywhere except for the point $t = 0$ and satisfies $c^\sharp(t) \in \mathring{\mathbf{X}} \cap \mathsf{C}_\mathrm{f}$ for all $t \in (0, T_+]$ and $\|c^\sharp\|_{L^\infty([0,T_+];\mathbb{R}^I)} \leq \mathbf{C}$ for all $t \in [0, T_+]$. Moreover,

$$\lim_{t\searrow 0} c^\sharp(t) = \overline{x},$$

with $\overline{x} \in \mathring{\mathbf{X}} \cap \mathsf{C}_\mathrm{f}$ from Theorem V.12(a), applied for $x^0 := c^0$.

Proof. From the uniform estimates (V.2.13) and (V.2.16), we infer with the help of Proposition V.17 that $\sup_{\varepsilon\in(0,\varepsilon_0]} \|c_\varepsilon\|_{C^1([T_0,T_+];\mathbb{R}^I)}$ is finite for fixed, but arbitrary $T_0 > 0$. The compact embedding of $C^1([T_0, T_+]; \mathbb{R}^I)$ into $C^{0,\frac{1}{2}}([T_0, T_+]; \mathbb{R}^I)$ and a diagonal argument yield local uniform convergence in $(0, T_+]$ of $(c_{\varepsilon_k})_{k\in\mathbb{N}}$ to some $c^\sharp \in C((0, T_+]; \mathbb{R}^I)$ and thus also pointwise convergence on $[0, T_+]$ defining $c^\sharp(0) := c^0$ (recall assumption (I)) on a suitable (non-relabelled) subsequence. Furthermore, one has weak convergence of $\dot{c}_{\varepsilon_k} \rightharpoonup \dot{c}^\sharp$ in $L^2([T_0, T_+]; \mathbb{R}^I)$. Consequently, on the finite intervals $[T_0, T_+]$, $\dot{c}_{\varepsilon_k} \rightharpoonup \dot{c}^\sharp$ also in $L^1([T_0, T_+]; \mathbb{R}^I)$. Weak lower semicontinuity of the L^1 norm then yields

$$\|\dot{c}^\sharp\|_{L^1([T_0,T_+];\mathbb{R}^I)} \leq \liminf_{k\to\infty} \|\dot{c}_\varepsilon\|_{L^1([T_0,T_+];\mathbb{R}^I)} \leq \sup_{\varepsilon\in(0,\varepsilon_0']} \|\dot{c}_\varepsilon\|_{L^1([0,T_+'];\mathbb{R}^I)} < \infty,$$

thanks to estimate (V.2.22) from Proposition V.17. By monotone convergence as $T_0 \searrow 0$, we conclude that $\dot{c}^\sharp \in L^1([0, T_+]; \mathbb{R}^I)$; hence, $c^\sharp$ is a.e. equal to an L^1-absolutely continuous function and has the additional properties mentioned above. In particular, the limit $c^\sharp(0^+) := \lim_{t\searrow 0} c^\sharp(t)$ exists. For all vectors $\zeta \in (\operatorname{span}\{\widetilde{\gamma}^1, \ldots, \widetilde{\gamma}^{R_\mathrm{f}}\})^\perp$ and all $t \geq 0$, one has

$$|\zeta^\mathrm{T}(c_\varepsilon(t) - c_\varepsilon^0)| = \left|\int_0^t \zeta^\mathrm{T}\dot{c}_\varepsilon(s)\,\mathrm{d}s\right| = \left|\int_0^t \sum_{r=1}^{R_\mathrm{s}} k_\varepsilon^r \mathcal{C}^r(c_\varepsilon(s))\zeta^\mathrm{T}\gamma^r\,\mathrm{d}s\right| \leq \int_0^t C\,\mathrm{d}s = Ct,$$

for some $C > 0$ thanks to (V.2.13), and hence after passage to $\varepsilon \searrow 0$ and $t \searrow 0$ that $\zeta^\mathrm{T}c^\sharp(0^+) = \zeta^\mathrm{T}c^0$. Moreover, $c^\sharp(0^+) \in \mathring{\mathbf{X}} \cap \mathsf{C}_\mathrm{f}$ by uniform positivity (V.2.16) and Proposition V.14. By Theorem V.12(a), it follows that $c^\sharp(0^+) = \overline{x}$ thanks to the uniqueness of steady states of the differential equation (V.2.10) in $(c^0 + \operatorname{span}\{\widetilde{\gamma}^1, \ldots, \widetilde{\gamma}^{R_\mathrm{f}}\}) \cap \mathring{\mathbf{X}}$. □

We now are in position to pass to the limit $\varepsilon \searrow 0$ in the energy dissipation balance **(EDB)** for $t \in [0, T_+]$.

Proposition V.19 (Convergence of the evolution for small times). *Assume that* (R1)&(R2) *and* (I) *hold. Then, the following holds for $c^\sharp$ from Proposition V.18:*

(a) For all $t \in [0, T_+]$, one has

$$\mathcal{E}(c^0) \geq \mathcal{E}(\overline{x}) = \mathcal{E}(c^\sharp(t)) + \int_0^t [\Psi(c^\sharp(s); \dot{c}^\sharp(s)) + \Psi^*(c^\sharp(s); -\mathrm{D}\mathcal{E}(c^\sharp(s)))]\,\mathrm{d}s. \tag{V.2.26}$$

(b) There exists a continuous map $\omega = (\omega^1, \ldots, \omega^{R_\mathrm{f}}) : (0, T_+] \to \mathbb{R}^{R_\mathrm{f}}$, such that for all $t \in (0, T_+]$:

$$\dot{c}^\sharp(t) = \sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c^\sharp(t))\gamma^r + \sum_{s=1}^{R_\mathrm{f}} \omega^s(t)\widetilde{\gamma}^s. \tag{V.2.27}$$

(c) Define the matrices $\mathbf{G} \in \mathbb{R}^{I\times R_\mathrm{f}}$ and $\mathbf{I}(z) \in \mathbb{R}^{I\times I}$ for $z \in \mathring{\mathbf{X}}$ via

$$\mathbf{G}\mathsf{e}_s := \widetilde{\gamma}^s, \qquad \mathbf{I}(z)\mathsf{e}_i := z_i^{-1}\mathsf{e}_i,$$

where e_l denotes the l^{th} canonical unit vector in $\mathbb{R}^d$ for the appropriate $d \in \mathbb{N}$, respectively. Then, for all $t \in (0, T_+]$:

$$\dot{c}^\sharp(t) = \left[\mathbb{1} - \mathbf{G}(\mathbf{G}^\mathrm{T}\mathbf{I}(c^\sharp(t))\mathbf{G})^{-1}\mathbf{G}^\mathrm{T}\mathbf{I}(c^\sharp(t))\right]\left(\sum_{r=1}^{R_\mathrm{s}} k^r C^r(c^\sharp(t))\gamma^r\right). \tag{V.2.28}$$

Proof. (a) Let $T_0 > 0$ and $t \in [T_0, T_+]$. From the energy dissipation balance **(EDB)** for (V.1.4), we have for $\varepsilon \in (0, \varepsilon_0]$ that

$$\mathcal{E}(c_\varepsilon^0) \geq \mathcal{E}(c_\varepsilon(T_0)) = \mathcal{E}(c_\varepsilon(t)) + \int_{T_0}^t [\Psi_\varepsilon(c_\varepsilon(s); \dot{c}_\varepsilon(s)) + \Psi_\varepsilon^*(c_\varepsilon(s); -\mathrm{D}\mathcal{E}(c_\varepsilon(s)))]\,\mathrm{d}s.$$

Recall that $\dot{c}_\varepsilon \rightharpoonup \dot{c}^\sharp$ weakly in $L^1([T_0, T_+]; \mathbb{R}^I)$ and $c_\varepsilon \to c^\sharp$ uniformly on $[T_0, T_+]$ as $\varepsilon \searrow 0$ thanks to Proposition V.18. In particular, since $c_\varepsilon(t) \in \mathring{\mathbf{X}}$ and $c^\sharp(t) \in \mathring{\mathbf{X}}$ for all $t \in [0, T_+]$, also $\mathrm{D}\mathcal{E}(c_\varepsilon) \to \mathrm{D}\mathcal{E}(c^\sharp)$ uniformly on $[T_0, T_+]$. The convergence in the sense of Mosco of the (auxiliary) dissipation potentials (cf. Lemma V.9, Proposition V.10) and Ioffe's lower semicontinuity theorem [101, Thm. 1] (see also

[138, Prop. 3.4]) then yield as $\varepsilon \searrow 0$:

$$\mathcal{E}(c^0) \geq \mathcal{E}(c^\sharp(T_0)) \geq \mathcal{E}(c^\sharp(t)) + \int_{T_0}^t \Big[\Psi(c^\sharp(s);\dot{c}^\sharp(s)) + \Psi^*(c^\sharp(s);-\mathrm{D}\mathcal{E}(c^\sharp(s)))\Big]\,\mathrm{d}s.$$

Monotone convergence as $T_0 \searrow 0$ implies

$$\mathcal{E}(c^0) \geq \mathcal{E}(\overline{x}) \geq \mathcal{E}(c^\sharp(t)) + \int_0^t [\Psi(c^\sharp(s);\dot{c}^\sharp(s)) + \Psi^*(c^\sharp(s);-\mathrm{D}\mathcal{E}(c^\sharp(s)))]\,\mathrm{d}s.$$

Since $c^\sharp(t) \in \mathring{\mathbf{X}}$ for all $t \in [0, T_+]$, the chain rule $\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(c^\sharp(s)) = \mathrm{D}\mathcal{E}(c^\sharp(s))^{\mathrm{T}}\dot{c}^\sharp(s)$ holds. The Young-Fenchel estimate (I.2.24) $\Psi(c^\sharp(s);\dot{c}^\sharp(s)) + \Psi^*(c^\sharp(s);-\mathrm{D}\mathcal{E}(c^\sharp(s))) \geq -\mathrm{D}\mathcal{E}(c^\sharp(s))^{\mathrm{T}}\dot{c}^\sharp(s)$ yields equality in the (second) estimate above and hence (V.2.26).

(b) Having the chain rule at hand, (V.2.26) is equivalent to its corresponding rate equation

$$\dot{c}^\sharp(t) \in \partial_\xi \Psi^*(c^\sharp(t);-\mathrm{D}\mathcal{E}(c^\sharp(t))) \quad \text{for almost all } t \in (0, T_+], \tag{V.2.29}$$

together with the Cauchy condition $\lim_{t\searrow 0} c^\sharp(t) = \overline{x}$.

Since both parts in $\Psi^*(c^\sharp(t);\cdot)$ are finite at $\xi = 0$ and the first part is continuous at $\xi = 0$ (recall (V.2.6)), the convex subdifferential of their sum can be computed as the sum of the subdifferentials of each part separately (see e.g. [77, Sect. I.5.3]). We first determine the subdifferential of $\xi \mapsto \chi^*(z,\xi)$ for fixed, but arbitrary $z \in \mathbf{X}$ and find

$$\partial_\xi \chi^*(z;\xi) = \begin{cases} \operatorname{span}\{\widetilde{\gamma}^s : s \in \widetilde{J}_z\} & \text{if } \xi^{\mathrm{T}}\widetilde{\gamma}^s = 0 \text{ for all } s \in \widetilde{J}_z, \\ \emptyset & \text{otherwise.} \end{cases}$$

We eventually obtain

$$\partial_\xi \Psi^*(z;\xi) = \begin{cases} \sum_{r\in J_z} k^r \Lambda(z^{\alpha^r}, z^{\beta^r})\xi^{\mathrm{T}}\gamma^r\,\gamma^r + \operatorname{span}\{\widetilde{\gamma}^s : s \in \widetilde{J}_z\} & \text{if } \xi^{\mathrm{T}}\widetilde{\gamma}^s = 0 \text{ for all } s \in \widetilde{J}_z, \\ \emptyset & \text{otherwise.} \end{cases}$$

From (V.2.29), we know that $\partial_\xi \Psi^*(c^\sharp(t);-\mathrm{D}\mathcal{E}(c^\sharp(t))) \neq \emptyset$ for almost every $t \in (0, T_+]$. Together with $c^\sharp(t) \in \mathring{\mathbf{X}}$, this yields the existence of a continuous map $\omega = (\omega^1, \ldots, \omega^{R_\mathrm{f}}) : (0, T_+] \to \mathbb{R}^{R_\mathrm{f}}$, such that for almost all $t \in (0, T_+]$, (V.2.27) holds.

(c) We omit the argument t for brevity. It remains to verify that

$$\omega = -(\mathbf{G}^{\mathrm{T}}\mathbf{I}(c^\sharp)\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{I}(c^\sharp)\left(\sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c^\sharp)\gamma^r\right). \tag{V.2.30}$$

To this end, we introduce two more matrices: $\mathbf{J}(z) \in \mathbb{R}^{I\times R_\mathrm{f}}$ and $\mathbf{A}(z) \in \mathbb{R}^{R_\mathrm{f}\times R_\mathrm{f}}$ for $z \in \mathring{\mathbf{X}}$, defined as

$$\begin{aligned} \mathbf{J}(z)\mathsf{e}_s &:= \mathrm{D}_z\mathcal{C}^{R_\mathrm{s}+s}(z), \\ \mathbf{A}(z)\mathsf{e}_s &:= z^{\widetilde{\alpha}^s}\mathsf{e}_s, \end{aligned}$$

for $s \in \{1, \ldots, R_\mathrm{f}\}$. Recalling (V.2.21) and $c^\sharp \in \mathsf{C}_\mathrm{f}$, we observe that

$$\mathbf{J}(c^\sharp) = -\mathbf{I}(c^\sharp)\mathbf{G}\mathbf{A}(c^\sharp).$$

By differentiation of $\mathcal{C}^{R_\mathrm{s}+s}(c^\sharp) = 0$ w.r.t. t, one has $\mathbf{J}(c^\sharp)^{\mathrm{T}}\dot{c}^\sharp = 0$ and hence, using (V.2.27):

$$0 = \sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c^\sharp)\mathbf{J}(c^\sharp)^{\mathrm{T}}\gamma^r + \mathbf{J}(c^\sharp)^{\mathrm{T}}\mathbf{G}\omega.$$

The matrix $\mathbf{J}(c^\sharp)^{\mathrm{T}}\mathbf{G} = -\mathbf{A}(c^\sharp)\mathbf{G}^{\mathrm{T}}\mathbf{I}(c^\sharp)\mathbf{G}$ is regular since it is a product of the two regular matrices $-\mathbf{A}(c^\sharp)$ (which is diagonal with strictly positive diagonal entries; recall that $c^\sharp \in \mathring{\mathbf{X}}$) and $\mathbf{G}^{\mathrm{T}}\mathbf{I}(c^\sharp)\mathbf{G}$

(because $\mathbf{G}$ is of full rank and $\mathbf{I}(c^\sharp)$ is also diagonal with positive diagonal entries). So, we are able to express ω in terms of $c^\sharp$ as

$$\omega = -(\mathbf{J}(c^\sharp)^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{J}(c^\sharp)^{\mathrm{T}}\left(\sum_{r=1}^{R_{\mathrm{s}}} k^r \mathcal{C}^r(c^\sharp)\gamma^r\right),$$

which yields (V.2.30) by elementary calculations.

□

We now prove that the limit $c^\sharp : [0, T_+] \to \mathring{\mathbf{X}}$ can be extended in time by means of the differential equation (V.2.28).

Lemma V.20 (Extension of $c^\sharp$). *Consider the initial-value problem*

$$\begin{aligned} &\dot{c}(t) = \left[\mathbb{1} - \mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{I}(c(t))\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{I}(c(t))\right]\left(\sum_{r=1}^{R_{\mathrm{s}}} k^r\mathcal{C}^r(c(t))\gamma^r\right) \quad \textit{for } t > 0, \\ &\lim_{t\searrow 0} c(t) = \overline{x}. \end{aligned} \tag{V.2.31}$$

There exists a unique solution $c : [0, \overline{T}) \to \mathring{\mathbf{X}}$ to (V.2.31)*, where $\overline{T} \in (T_+, \infty]$ denotes the maximal time of existence.*

Proof. The unique local-in-time solution is given by $c^\sharp : [0, T_+] \to \mathbb{R}^I$ above. Since the right-hand side of the differential equation is locally Lipschitz continuous in $\mathring{\mathbf{X}}$, $c^\sharp$ can be uniquely extended to the desired solution c until strict positivity or finiteness of c fails. The bounds $c_i^\sharp(T_+) \in [\delta, \mathbf{C}]$ imply that $\overline{T}$ is *strictly* larger than T_+. □

We now aim both for *global-in-time* existence of the solution c to (V.2.31) (i.e. $\overline{T} = \infty$) and for local uniform convergence of $(c_\varepsilon)_{\varepsilon>0}$ to c as $\varepsilon \searrow 0$. We first prove

Proposition V.21 (Local uniform convergence on interval of existence). *As $\varepsilon \searrow 0$, $c_\varepsilon \to c$ locally uniformly in $(0, \overline{T})$.*

Proof. Assume that the assertion is false, i.e. there is $T_- < \overline{T}$ such that

$$T_- = \sup\{T \geq T_+ : c_\varepsilon \to c \text{ locally uniformly on } (0, T]\}. \tag{V.2.32}$$

We now show that there exists $T_+' > T_-$ such that (V.2.19) holds, by a similar argument as in the proof of Proposition V.15. This implies a contradiction since we can perform — with the help of Proposition V.17 — the same proof as for the Propositions V.18 and V.19 to conclude local uniform convergence $c_\varepsilon \to c$ on $(0, T_+']$, since the extended solution c to (V.2.31) is unique.

We have, as in the proof of Proposition V.15, since $c(t) \in \mathsf{C}_{\mathrm{f}}$ for all $t \in (0, \overline{T})$ that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(c_\varepsilon(t); c(t')) \leq \widetilde{K}_0,$$

for all $t' \in \left(0, \frac{1}{2}(\overline{T} + T_-)\right)$, with a constant $\widetilde{K}_0 > 0$ independent of ε and t. Define

$$\widetilde{t} := \min\left\{\frac{1}{8\widetilde{K}_0}\min_{s\in[0,T_-]}\min_i c_i(s), \frac{1}{2}(\overline{T} - T_-), \frac{1}{2}T_-\right\} > 0.$$

By the convergence $c_\varepsilon(T_- - \widetilde{t}) \to c(T_- - \widetilde{t})$ as $\varepsilon \searrow 0$ and the continuity of $\mathcal{E}(\cdot; c(T_- - \widetilde{t}))$, one has

$$\mathcal{E}(c_\varepsilon(T_- - \widetilde{t}); c(T_- - \widetilde{t})) \leq \frac{1}{2}\min_{s\in[0,T_-]}\min_i c_i(s),$$

for all $\varepsilon \leq \varepsilon_0'$, with a suitably small $\varepsilon_0' > 0$. We conclude that

$$\mathcal{E}(c_\varepsilon(t); c(T_- - \widetilde{t})) \leq \frac{3}{4}\min_{s\in[0,T_-]}\min_i c_i(s) \qquad \forall t \in [T_- - \widetilde{t}, T_- + \widetilde{t}].$$

From here, we can proceed as in the proof of Proposition V.15 to obtain (V.2.19) (and consequently the assertion) for $T'_+ := T_- + \widetilde{t} > T_-$. □

Proposition V.22 (Global existence). *The solution c to* (V.2.31) *exists globally in time, i.e. $\overline{T} = \infty$, and is subject to $\|c\|_{L^\infty([0,\infty);\mathbb{R}^I)} \le \mathbf{C}$.*

Proof. Assume that $\overline{T} < \infty$. Due to Proposition V.21 and the uniform bound (V.2.13), $\|c\|_{L^\infty([0,\overline{T});\mathbb{R}^I)} \le \mathbf{C}$. Hence, blow-up at time $\overline{T}$ cannot occur. So, strict positivity of c fails at time $\overline{T}$. Let $J := \{i \in \{1,\dots,I\} : \lim_{t\nearrow\overline{T}} c_i(t) = 0\}$. Define the curve of chemical potentials $\mu \in C([0,\overline{T});\mathbb{R}^I)$ via $\mu_i(t) := \log(c_i(t))$ for $i \in \{1,\dots,I\}$ and observe that $\mu(t) \in (\operatorname{span}\{\widetilde{\gamma}^1,\dots,\widetilde{\gamma}^{R_\mathrm{f}}\})^\perp$ since $c(t) \in \mathsf{C}_\mathrm{f}$, for all $t \in (0,\overline{T})$. Clearly, $\lim_{t\nearrow\overline{T}} \mu_i(t) = -\infty$ for $i \in J$.

Let a sequence $(t_n)_{n\in\mathbb{N}}$ in $(0,\overline{T})$ with $\lim_{n\to\infty} t_n = \overline{T}$ be given. The bounded sequence $\left(\frac{\mu(t_n)}{|\mu(t_n)|}\right)_{n\in\mathbb{N}}$ in $(\operatorname{span}\{\widetilde{\gamma}^1,\dots,\widetilde{\gamma}^{R_\mathrm{f}}\})^\perp$ possesses (thanks to the Bolzano-Weierstraß theorem) an accumulation point $\nu \in (\operatorname{span}\{\widetilde{\gamma}^1,\dots,\widetilde{\gamma}^{R_\mathrm{f}}\})^\perp$ with $|\nu| = 1$, $\nu_i = 0$ for all $i \notin J$ and $\nu_i \le 0$ for all $i \in J$. In particular, there exists $j \in J$ such that $\nu_j < 0$. For $t \in (0,\overline{T})$, one then first obtains

$$-\frac{\mathrm{d}}{\mathrm{d}t}\nu^\mathrm{T} c(t) = -\sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(c(t))\nu^\mathrm{T}\gamma^r.$$

Define, for $i = 1,\dots,I$, maps $f_i : \mathbf{X} \to \mathbb{R}$ by $f_i(z) = \sum_{r=1}^{R_\mathrm{s}} k^r \mathcal{C}^r(z)\gamma_i^r$. One easily verifies that $(f_1,\dots,f_I)$ is *quasi-positive* [30], that is, $z_i = 0$ for some $z \in \mathbf{X}$ and some $i \in \{1,\dots,I\}$ implies $f_i(z) \ge 0$. Thus, denoting the joint Lipschitz constant of all f_i on $[0,\mathbf{C}]^I$ by $L > 0$, one gets for each $i \in \{1,\dots,I\}$ and $z \in [0,\mathbf{C}]^I$:

$$-f_i(z) \le -f_i(z) + f_i(z - z_i\mathrm{e}_i) \le |f_i(z - z_i\mathrm{e}_i) - f_i(z)| \le L z_i.$$

Consequently (recall that $\nu_i \le 0$ for all i):

$$-\frac{\mathrm{d}}{\mathrm{d}t}\nu^\mathrm{T} c(t) \ge L\nu^\mathrm{T} c(t).$$

Gronwall's lemma implies $-c(t)^\mathrm{T}\nu \ge -\overline{x}^\mathrm{T}\nu\exp(-Lt) \ge -\overline{x}^\mathrm{T}\nu\exp(-L\overline{T})$, which is a *strictly* positive constant. Hence, for all $n \in \mathbb{N}$:

$$0 < -\overline{x}^\mathrm{T}\nu\exp(-L\overline{T}) \le -c(t_n)^\mathrm{T}\nu = -\sum_{i=1}^{I}\nu_i\exp(\mu_i(t_n)) = \sum_{i\in J}(-\nu_i)\exp(\mu_i(t_n)) \xrightarrow{n\to\infty} 0,$$

which is a contradiction. Therefore, the claim $\overline{T} = \infty$ follows. □

Since c exists for all times, we immediately obtain with Proposition V.21 that $c_\varepsilon \to c$ as $\varepsilon \searrow 0$ uniformly on compact subsets of $(0,\infty)$. We can also re-prove Proposition V.19 on intervals $[0,T]$ for arbitrary time horizon $T > 0$ to complete the proof of Theorem V.1.

Part VI

Conclusion and outlook

We give some concluding remarks and sketch several open problems to be studied in the future.

Part II. In Part II, we have shown the applicability of variational methods to a variety of two-species models. By approximation with minimizing movements, we have, on the one hand, been able to prove the well-posedness of evolution equations with a formal gradient flow structure. On the other hand, for systems with a suitably weak coupling, we developed a new method to establish the exponential convergence to equilibrium of the before-constructed weak solutions.

Our strategy of proof to obtain the existence of weak solutions to our variant of the Keller-Segel model can successfully be applied in many different parameter situations allowing for even broader ranges as described in Part II. The same is true for the Poisson-Nernst-Planck system, see [112] for an almost exhaustive existence analysis. However, at the moment, it is unclear if our method for proving convergence to equilibrium can be transferred to more general situations. Specifically, for the analysis of the Keller-Segel model, the choice $m = 2$ for the exponent of nonlinear diffusion seems to be crucial. Apart from the possible existence of more than one minimizer of the energy functional, the right energy-dissipation estimates are currently not at hand. For example, one might require an analog to (II.2.55) from Proposition II.28, that is

$$\|u - u_\infty\|_{L^2}^2 \leq C \int_{\mathbb{R}^d} (u^m - u_\infty^m)\, \mathrm{d}x,$$

to have available a complementary estimate to (II.2.56) from Proposition II.28, which estimates $\|v - v_\infty\|_{L^2}^2$ from above and is unchanged since the L^2-variational structure of the v-component is conserved. Unfortunately, since u_∞ will in general not be uniformly positive, the estimate above does not hold for $m \neq 2$.

In contrast, our analysis of the long-time behaviour of the Poisson-Nernst-Planck model may allow for a generalization to non-quadratic diffusion, since two Wasserstein spaces $\mathscr{P}_2$ are more compatible than $\mathscr{P}_2$ and L^2 when it comes to a change in the diffusion exponent. For example, it could be a reasonable strategy to have conjugate diffusion exponents, i.e. $m_u > 1$ and $m_v > 1$ such that $1 = \frac{1}{m_u} + \frac{1}{m_v}$. This variation of the model might also be conceivable from a modelling point of view: one can take into account effects caused by the different "size" of the particles (e.g. sodium and chloride ions) on the rate of diffusion through the environment.

One may ask if there is a general principle for two-component systems behind our machinery to analyse the long-time asymptotics. The answer is probably affirmative, at least under relatively specialized assumptions on the coupling between the species and on the structural compatibility of both the underlying metric spaces and the proper domain of the energy functional.

Part III. We have found in Part III a novel condition for λ-convexity along generalized geodesics of a multi-component version of the interaction energy in a cartesian product of Wasserstein-type spaces. This condition provides a better insight into the behaviour of systems of interaction equations, showing new effects which do not arise in the scalar case. In one spatial dimension, we performed a thorough analysis of solutions to the associated multi-component interaction equation. We successfully generalized results for the scalar interaction equation (from, e.g., [35, 82, 158]) to the case of genuine systems.

One apparent idea for generalization of the results obtained is to add diffusion to the system. At least formally, the weighted center of mass is still invariant: for example, for the pure (scalar) diffusion equation

$$\partial_t \mu = \operatorname{div}(\mu \mathrm{D} f'(\mu)),$$

seen as the evolution equation associated to the gradient flow of $\mathcal{E}(\mu) = \int_{\mathbb{R}^d} f(\mu)\,\mathrm{d}x$ (for smooth $f : \mathbb{R} \to \mathbb{R}$ subject to McCann's criteria) w.r.t. the L^2-Wasserstein distance, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} x\mu(t,x)\,\mathrm{d}x = -\int_{\mathbb{R}^d} f''(\mu)\mu \mathrm{D}\mu\,\mathrm{d}x = -\int_{\mathbb{R}^d} \mathrm{D}\left(\int_0^{\mu(t,x)} f''(z)z\,\mathrm{d}z\right)\,\mathrm{d}x = 0,$$

under the assumption that $x \mapsto \int_0^{\mu(t,x)} f''(z)z\,\mathrm{d}z$ is integrable on $\mathbb{R}^d$ [35]. Consequently, an energy functional defined as a sum of a diffusion and an interaction part will still be λ_0-convex along generalized geodesics where $\lambda_0 \in \mathbb{R}$ is the modulus of convexity for the interaction part of the functional. Hence, existence and uniqueness of solutions to the resulting diffusion-interaction system can be deduced using the theory from [4]. Concerning the one-dimensional case $d = 1$ and $\lambda_0 < 0$, the methods from [35] might also be generalized to systems with diffusion in order to investigate the qualitative behaviour of solutions.

Easily, one sees that adding a general drift term to the system destroys the invariance of the weighted center of mass. Including such a term might cause a loss of uniform convexity along generalized geodesics ($\lambda > 0$), but still, the resulting energy functional could be λ-convex along generalized geodesics for some $\lambda \leq 0$ which is sufficent for proving existence and uniqueness of solutions [4]. Hence, further studies should include an analysis of the detailed qualitative behaviour of the gradient flow associated to

$$\mathcal{E}(\mu) = \sum_{i=1}^n \int_{\mathbb{R}^d} f_i(\mu_i)\,\mathrm{d}x + \sum_{i=1}^n \int_{\mathbb{R}^d} V_i(x)\,\mathrm{d}\mu_i(x) + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} W_{ij}(x-y)\,\mathrm{d}\mu_i(x)\,\mathrm{d}\mu_j(y).$$

In particular, it would be interesting to characterize the set of steady states and their stability which might lead to a relatively complex analysis (as already in the scalar case [35, 82, 83]).

Part IV. In Part IV, we have successfully endowed a certain class of degenerate diffusion systems with a variational structure, generalizing the theory for transportation distances with nonlinear mobility (cf. [74, 124]) to the vector-valued case. We have derived criteria for the geodesic λ-convexity of functionals with respect to our novel metric $\mathbf{W_M}$, rigorously showing that the heat flow is a $\mathbf{W_M}$-gradient flow in certain situations. By approximation with minimizing movements, we demonstrated the existence of weak solutions to a broad class of second- and fourth-order systems, even if the underlying energy functional is not geodesically convex.

Observe that in principle, one can also consider multiple space dimensions ($d > 1$) with the same assumptions on the mobility matrix function $\mathbf{M}$: A pseudo-distance $\mathbf{W_M}$ on the space $\mathscr{M}(\mathbb{R}^n; S)$ can be defined via

$$\mathbf{W_M}(\mu_0,\mu_1)^2 = \inf\left\{\int_0^1\int_{\mathbb{R}^d} \operatorname{tr}\left(W_t^{\mathrm{T}}\mathbf{M}(\mu_t)^{-1}W_t\right)\,\mathrm{d}x\,\mathrm{d}t :\ \partial_t\mu_t = -\operatorname{div}(W_t),\ \mu|_{t=0} = \mu_0,\ \mu|_{t=1} = \mu_1\right\},$$

where the flux W now is a $n \times d$-matrix-valued curve. Consequently, one can also derive a sufficient condition for geodesic convexity. Due to its presumable complexity, however, it might be of limited use in practice.

Concerning our existence analysis via the minimizing movement scheme, it might be possible to extend the results also to certain classes of gradient-dependent energy functionals where the integrand $f : \mathbb{R}^n \times S \to \mathbb{R}$ is not uniformly convex. A promising ansatz is to replace f by

$$|\partial_x g(z)|^2 + |g(z)|^2,$$

for a suitable bijective function $g : \mathbb{R}^n \to \mathbb{R}^n$. Indeed, one obtains via the direct method of the calculus of variations that the associated Yosida penalization

$$\mathcal{E}_\tau(\mu \,|\, \widetilde{\mu}) = \frac{1}{2\tau}\mathbf{W}_{\mathbf{M}}(\mu, \widetilde{\mu})^2 + \|g(\mu)\|_{H^1}^2$$

admits minimizers.

Part V. In Part V, we gave an alternative proof for a result in [30] on the behaviour of systems of ordinary differential equations modelling chemical reactions when some of the reaction rates tend to infinity. Our main improvement is of methodical nature: using the generalized gradient structure of the system allowed us — without any heuristics on advance — to rigorously establish the evolutionary convergence to a limit (gradient) system.

An appealing generalization of the presented results lies in the inclusion of diffusion to the system leading to a coupled system of reaction-diffusion equations of the form

$$\partial_t c_\varepsilon = \operatorname{div}(A(x)\mathrm{D}c_\varepsilon) - R_\varepsilon(c_\varepsilon),$$

together with no-flux boundary conditions $A(\overline{x})\mathrm{D}c_\varepsilon(t,\overline{x})\nu(\overline{x}) = 0$ on the boundary $\partial\Omega$ of a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary. Above, $A \in \mathrm{C}^1(\overline{\Omega}; \mathbb{R}^{I\times I})$ is a diagonal matrix comprising the (possibly space-dependent) diffusion coefficients and R_ε are reaction terms of mass-action type (comprising slow and fast reactions) as in Part V. Following the procedure in [134], one can assign a generalized gradient structure to the system above and write down the corresponding energy-dissipation balance. Note that, apart from the diffusion part of the system, the energy $\mathcal{E}$ and the dual dissipation potential Ψ_ε^* will turn out to be the (w.r.t. $x \in \Omega$) integrated quantities from Part V. The question of evolutionary convergence as $\varepsilon \searrow 0$ of those gradient systems is open; the necessary *a priori* estimates are at the moment not available. It is a relatively straightforward consequence of the energy-dissipation balance for $\varepsilon > 0$ in combination with well-prepared initial conditions $\lim_{\varepsilon\searrow 0}\mathcal{E}(c_\varepsilon^0) = \mathcal{E}(c^0)$ that

$$\sup_{\varepsilon>0}\left(\|c_\varepsilon\|_{L^\infty([0,T];L^1)} + \max_{i\in\{1,\dots,I\}}\|\sqrt{c_{\varepsilon,i}}\|_{L^2([0,T];H^1)} + \|c_\varepsilon\|_{L^1([0,T];W^{1,1})}\right) < \infty,$$

assuming that c_ε is a nonnegative function. In order to deduce convergence, the Aubin-Lions compactness lemma might be the appropriate tool — however, the uniform estimate above does not guarantee its applicability.

Bibliography

[1] M. Agueh. Rates of decay to equilibria for p-Laplacian type equations. *Nonlinear Anal.*, 68(7):1909–1927, 2008.

[2] H. Amann. Dynamic theory of quasilinear parabolic systems. III. Global existence. *Math. Z.*, 202(2):219–250, 1989.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[4] 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, second edition, 2008.

[5] A. Arnold, P. Markowich, and G. Toscani. On large time asymptotics for drift-diffusion-Poisson systems. *Transport Theory Statist. Phys.*, 29(3-5):571–581, 2000.

[6] 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.

[7] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: from diffusion to reaction. *Calc. Var. Partial Differential Equations*, 44(3-4):419–454, 2012.

[8] H. Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, 1984.

[9] D. Balagué, J. A. Carrillo, and Y. Yao. Confinement for repulsive-attractive kernels. *Discrete Contin. Dyn. Syst. Ser. B*, 19(5):1227–1248, 2014.

[10] N. Ben Abdallah, F. Méhats, and N. Vauchelet. A note on the long time behavior for the drift-diffusion-Poisson system. *C. R. Math. Acad. Sci. Paris*, 339(10):683–688, 2004.

[11] 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.

[12] D. Benedetto, E. Caglioti, and M. Pulvirenti. A kinetic equation for granular media. *RAIRO Modél. Math. Anal. Numér.*, 31(5):615–641, 1997.

[13] A. L. Bertozzi and J. Brandman. Finite-time blow-up of L^∞-weak solutions of an aggregation equation. *Commun. Math. Sci.*, 8(1):45–65, 03 2010.

[14] A. L. Bertozzi, J. A. Carrillo, and T. Laurent. Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity*, 22(3):683–710, 2009.

[15] A. L. Bertozzi and T. Laurent. Finite-time blow-up of solutions of an aggregation equation in $\mathbf{R}^n$. *Comm. Math. Phys.*, 274(3):717–735, 2007.

[16] A. L. Bertozzi, T. Laurent, and J. Rosado. L^p theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64(1):45–83, 2011.

[17] O. V. Besov, V. P. Il′in, and S. M. Nikol′skiĭ. *Integral representations of functions and imbedding theorems. Vol. I.* V. H. Winston & Sons, Washington, D.C.; Halsted Press, New York-Toronto, Ont.-London, 1978.

[18] P. Biler, L. Corrias, and J. Dolbeault. Large mass self-similar solutions of the parabolic-parabolic Keller-Segel model of chemotaxis. *J. Math. Biol.*, 63(1):1–32, 2011.

[19] P. Biler and J. Dolbeault. Long time behavior of solutions of Nernst-Planck and Debye-Hückel drift-diffusion systems. *Ann. Henri Poincaré*, 1(3):461–472, 2000.

[20] P. Biler, J. Dolbeault, and P. A. Markowich. Large time asymptotics of nonlinear drift-diffusion systems with Poisson coupling. *Transport Theory Statist. Phys.*, 30(4-6):521–536, 2001.

[21] P. Biler, G. Karch, and P. Laurençot. Blowup of solutions to a diffusive aggregation model. *Nonlinearity*, 22(7):1559–1568, 2009.

[22] A. Blanchet. A gradient flow approach to the Keller-Segel systems. *RIMS Kokyuroku's lecture notes*, 2014.

[23] A. Blanchet, V. Calvez, and J. A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.

[24] A. Blanchet, E. A. Carlen, and J. A. Carrillo. Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. *J. Funct. Anal.*, 262(5):2142–2230, 2012.

[25] A. Blanchet, J. A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot, and S. Lisini. A hybrid variational principle for the Keller-Segel system in $\mathbb{R}^2$. *ESAIM Math. Model. Numer. Anal.*, 49(6):1553–1576, 2015.

[26] A. Blanchet, J. A. Carrillo, and P. Laurençot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calc. Var. Partial Differential Equations*, 35(2):133–168, 2009.

[27] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations*, 2006(44):1–33, 2006.

[28] A. Blanchet and P. Laurençot. The parabolic-parabolic Keller-Segel system with critical diffusion as a gradient flow in $\mathbb{R}^d$, $d \geq 3$. *Comm. Partial Differential Equations*, 38(4):658–686, 2013.

[29] M. Bodnar and J. J. L. Velázquez. An integro-differential equation arising as a limit of individual cell-based models. *J. Differential Equations*, 222(2):341–380, 2006.

[30] D. Bothe. Instantaneous limits of reversible chemical reactions in presence of macroscopic convection. *J. Differential Equations*, 193(1):27–48, 2003.

[31] D. Bothe and M. Pierre. The instantaneous limit for reaction-diffusion systems with a fast irreversible reaction. *Discrete Contin. Dyn. Syst. Ser. S*, 5(1):49–59, 2012.

[32] D. Bothe, M. Pierre, and G. Rolland. Cross-diffusion limit for a reaction-diffusion system with fast reversible reaction. *Comm. Partial Differential Equations*, 37(11):1940–1966, 2012.

[33] A. Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[34] M. Burger, V. Capasso, and D. Morale. On an aggregation model with long and short range interactions. *Nonlinear Anal. Real World Appl.*, 8(3):939–958, 2007.

[35] M. Burger and M. Di Francesco. Large time behavior of nonlocal aggregation models with nonlinear diffusion. *Netw. Heterog. Media*, 3(4):749–785, 2008.

[36] M. Burger, M. Di Francesco, P. A. Markowich, and M.-T. Wolfram. Mean field games with nonlinear mobilities in pedestrian dynamics. *Discrete Contin. Dyn. Syst. Ser. B*, 19(5):1311–1333, 2014.

[37] V. Calvez and J. A. Carrillo. Volume effects in the Keller-Segel model: energy estimates preventing blow-up. *J. Math. Pures Appl. (9)*, 86(2):155–175, 2006.

[38] V. Calvez and J. A. Carrillo. Refined asymptotics for the subcritical Keller-Segel system and related functional inequalities. *Proc. Amer. Math. Soc.*, 140(10):3515–3530, 2012.

[39] V. Calvez and L. Corrias. The parabolic-parabolic Keller-Segel model in $\mathbb{R}^2$. *Commun. Math. Sci.*, 6(2):417–447, 2008.

[40] V. Calvez, L. Corrias, and M. A. Ebde. Blow-up, concentration phenomenon and global existence for the Keller-Segel model in high dimension. *Comm. Partial Differential Equations*, 37(4):561–584, 2012.

[41] J. F. Campos and J. Dolbeault. Asymptotic estimates for the parabolic-elliptic Keller-Segel model in the plane. *Comm. Partial Differential Equations*, 39(5):806–841, 2014.

[42] J. A. Cañizo, J. A. Carrillo, and F. S. Patacchini. Existence of compactly supported global minimisers for the interaction energy. *Arch. Ration. Mech. Anal.*, 217(3):1197–1217, 2015.

[43] J. A. Cañizo, J. A. Carrillo, and J. Rosado. A well-posedness theory in measures for some kinetic models of collective motion. *Math. Models Methods Appl. Sci.*, 21(3):515–539, 2011.

[44] P. Cardaliaguet, G. Carlier, and B. Nazaret. Geodesics for a class of distances in the space of probability measures. *Calc. Var. Partial Differential Equations*, 48(3-4):395–420, 2013.

[45] K. Carrapatoso and S. Mischler. Uniqueness and long time asymptotic for the parabolic-parabolic Keller-Segel equation, 2014. Preprint. arXiv:1406.6006.

[46] J. A. Carrillo, D. Castorina, and B. Volzone. Ground states for diffusion dominated free energies with logarithmic interaction. *SIAM J. Math. Anal.*, 47(1):1–25, 2015.

[47] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[48] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Confinement in nonlocal interaction equations. *Nonlinear Anal.*, 75(2):550–558, 2012.

[49] J. A. Carrillo, M. Di Francesco, and G. Toscani. Intermediate asymptotics beyond homogeneity and self-similarity: long time behavior for $u_t = \Delta\phi(u)$. *Arch. Ration. Mech. Anal.*, 180(1):127–149, 2006.

[50] J. A. Carrillo, M. R. D'Orsogna, and V. Panferov. Double milling in self-propelled swarms from kinetic theory. *Kinet. Relat. Models*, 2(2):363–378, 2009.

[51] J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.*, 133:1–82, 2001.

[52] J. A. Carrillo, S. Lisini, and E. Mainini. Gradient flows for non-smooth interaction potentials. *Nonlinear Anal.*, 100:122–147, 2014.

[53] 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.

[54] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[55] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[56] J. A. Carrillo and J. Rosado. Uniqueness of bounded solutions to aggregation equations by optimal transport methods. In *European Congress of Mathematics*, pages 3–16. Eur. Math. Soc., Zürich, 2010.

[57] J. A. Carrillo and G. Toscani. Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity. *Indiana Univ. Math. J.*, 49(1):113–142, 2000.

[58] J. A. Carrillo and G. Toscani. Wasserstein metric and large-time asymptotics of nonlinear diffusion equations. In *New trends in mathematical physics*, pages 234–244. World Sci. Publ., Hackensack, NJ, 2004.

[59] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differential Equations*, 224(1):39–59, 2006.

[60] R. M. Colombo and M. Lécureux-Mercier. Nonlocal crowd dynamics models for several populations. *Acta Math. Sci. Ser. B Engl. Ed.*, 32(1):177–196, 2012.

[61] F. Conforto and L. Desvillettes. Rigorous passage to the limit in a system of reaction-diffusion equations towards a system including cross diffusions. *Commun. Math. Sci.*, 12(3):457–472, 2014.

[62] L. Corrias and B. Perthame. Asymptotic decay for the solutions of the parabolic-parabolic Keller-Segel chemotaxis system in critical spaces. *Math. Comput. Modelling*, 47(7-8):755–764, 2008.

[63] G. Crippa and M. Lécureux-Mercier. Existence and uniqueness of measure solutions for a system of continuity equations with non-local flow. *NoDEA Nonlinear Differential Equations Appl.*, 20(3):523–537, 2013.

[64] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.

[65] E. De Giorgi. New problems on minimizing movements. *Notes Appl. Math.*, 2(4):81–89, 1983.

[66] M. Di Francesco and S. Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity*, 26(10):2777–2808, 2013.

[67] M. Di Francesco and S. Fagioli. A nonlocal swarm model for predators-prey interactions. *Math. Models Methods Appl. Sci.*, 26(2):319–355, 2016.

[68] M. Di Francesco and D. Matthes. Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. Partial Differential Equations*, 50(1-2):199–230, 2014.

[69] M. Di Francesco and J. Rosado. Fully parabolic Keller–Segel model for chemotaxis with prevention of overcrowding. *Nonlinearity*, 21(11):2715, 2008.

[70] M. Di Francesco and M. Wunsch. Large time behavior in Wasserstein spaces and relative entropy for bipolar drift-diffusion-Poisson models. *Monatsh. Math.*, 154(1):39–50, 2008.

[71] K. Disser and M. Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Netw. Heterog. Media*, 10(2):233–253, 2015.

[72] K. Disser, M. Liero, and J. Zinsl. On the evolutionary Gamma-convergence of gradient systems modeling slow and fast chemical reactions, 2016. WIAS Preprint No. 2227.
Available at http://www.wias-berlin.de/preprint/2227/wias_preprints_2227.pdf.

[73] J. Dolbeault and J. Campos. A functional framework for the Keller-Segel system: logarithmic Hardy-Littlewood-Sobolev and related spectral gap inequalities. *C. R. Math. Acad. Sci. Paris*, 350(21-22):949–954, 2012.

[74] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

[75] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory.

[76] M. Eisenbach and J. Lengeler. *Chemotaxis*. Imperial College Press, London, 2004.

[77] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1999.

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

[79] M. Erbar and J. Maas. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.*, 34(4):1355–1374, 2014.

[80] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, second edition, 2010.

[81] M. Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Rational Mech. Anal.*, 132(4):311–370, 1995.

[82] K. Fellner and G. Raoul. Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.*, 20(12):2267–2291, 2010.

[83] K. Fellner and G. Raoul. Stability of stationary states of non-local equations with singular interaction potentials. *Math. Comput. Modelling*, 53(7-8):1436–1450, 2011.

[84] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979.

[85] F. Friedlander. *Introduction of the theory of distributions. With additional material by M. Joshi.* Cambridge University Press, Cambridge, second edition, 1998.

[86] L. Giacomelli and F. Otto. Variational formulation for the lubrication approximation of the Hele-Shaw flow. *Calc. Var. Partial Differential Equations*, 13(3):377–403, 2001.

[87] G. Giacomin, J. L. Lebowitz, and R. Marra. Macroscopic evolution of particle systems with short- and long-range interactions. *Nonlinearity*, 13(6):2143–2162, 2000.

[88] U. Gianazza, G. Savaré, and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Ration. Mech. Anal.*, 194(1):133–220, 2009.

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

[90] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001.

[91] A. Glitzky and A. Mielke. A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. *Z. Angew. Math. Phys.*, 64(1):29–52, 2013.

[92] A. Goeke, C. Schilli, S. Walcher, and E. Zerz. Computing quasi-steady state reductions. *J. Math. Chem.*, 50(6):1495–1513, 2012.

[93] A. Goeke and S. Walcher. Quasi-steady state: searching for and utilizing small parameters. In *Recent trends in dynamical systems*, volume 35 of *Springer Proc. Math. Stat.*, pages 153–178. Springer, Basel, 2013.

[94] F. Golse. The mean-field limit for the dynamics of large particle systems. In *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. IX, 1–47. Univ. Nantes, Nantes, 2003.

[95] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1981.

[96] T. Hillen and K. J. Painter. A user's guide to PDE models for chemotaxis. *J. Math. Biol.*, 58(1-2):183–217, 2009.

[97] T. Hillen and A. Potapov. The one-dimensional chemotaxis model: global existence and asymptotic profile. *Math. Methods Appl. Sci.*, 27(15):1783–1801, 2004.

[98] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.*, 46(5-6):1159–1194, 1987.

[99] F. Horn and R. Jackson. General mass action kinetics. *Arch. Rational Mech. Anal.*, 47:81–116, 1972.

[100] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.*, 105(3):103–165, 2003.

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

[102] S. Ishida and T. Yokota. Global existence of weak solutions to quasilinear degenerate Keller-Segel systems of parabolic-parabolic type. *J. Differential Equations*, 252(2):1421–1440, 2012.

[103] S. Ishida and T. Yokota. Global existence of weak solutions to quasilinear degenerate Keller-Segel systems of parabolic-parabolic type with small data. *J. Differential Equations*, 252(3):2469–2491, 2012.

[104] W. Jäger and S. Luckhaus. On explosions of solutions to a system of partial differential equations modelling chemotaxis. *Trans. Amer. Math. Soc.*, 329(2):819–824, 1992.

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

[106] A. Jüngel. *Quasi-hydrodynamic semiconductor equations*. Progress in Nonlinear Differential Equations and their Applications, 41. Birkhäuser Verlag, Basel, 2001.

[107] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.

[108] A. Jüngel and I. V. Stelzer. Entropy structure of a cross-diffusion tumor-growth model. *Math. Models Methods Appl. Sci.*, 22(7):1250009, 26, 2012.

[109] K. Kang, B. Perthame, A. Stevens, and J. J. L. Velázquez. An integro-differential equation model for alignment and orientational aggregation. *J. Differential Equations*, 246(4):1387–1421, 2009.

[110] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.

[111] D. Kinderlehrer and M. Kowalczyk. The Janossy effect and hybrid variational principles. *Discrete Contin. Dyn. Syst. Ser. B*, 11(1):153–176, 2009.

[112] D. Kinderlehrer, L. Monsaingeon, and X. Xu. A Wasserstein gradient flow approach to Poisson-Nernst-Planck equations. *ESAIM Control Optim. Calc. Var.*, 2015. doi:10.1051/cocv/2015043.

[113] T. Kolokolnikov, Y. Huang, and M. Pavlovski. Singular patterns for an aggregation model with a confining potential. *Phys. D*, 260:65–76, 2013.

[114] R. Kowalczyk. Preventing blow-up in a chemotaxis model. *J. Math. Anal. Appl.*, 305(2):566–588, 2005.

[115] H. Kozono and Y. Sugiyama. Global strong solution to the semi-linear Keller-Segel system of parabolic-parabolic type with small data in scale invariant spaces. *J. Differential Equations*, 247(1):1–32, 2009.

[116] P. Krejčí and L. Panizzi. Regularity and uniqueness in quasilinear parabolic systems. *Appl. Math.*, 56(4):341–370, 2011.

[117] I. Lapidus and R. Schiller. Model for the chemotactic response of a bacterial population. *Biophysical Journal*, 16(7):779–789, 1976.

[118] P. Laurençot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the Muskat problem. *Calc. Var. Partial Differential Equations*, 47(1-2):319–341, 2013.

[119] T. Laurent. Local and global existence for an aggregation equation. *Comm. Partial Differential Equations*, 32(10-12):1941–1964, 2007.

[120] H. Li and G. Toscani. Long-time asymptotics of kinetic models of granular flows. *Arch. Ration. Mech. Anal.*, 172(3):407–428, 2004.

[121] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, second edition, 2001.

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

[123] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. R. Soc. A*, 371(2005):20120346/1–20120346/28, 2013.

[124] S. Lisini and A. Marigonda. On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscripta Math.*, 133(1-2):197–224, 2010.

[125] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.

[126] J. López-Gómez, T. Nagai, and T. Yamada. The basin of attraction of the steady-states for a chemotaxis model in $\mathbb{R}^2$ with critical mass. *Arch. Ration. Mech. Anal.*, 207(1):159–184, 2013.

[127] S. Luckhaus and Y. Sugiyama. Large time behavior of solutions in super-critical cases to degenerate Keller-Segel systems. *M2AN Math. Model. Numer. Anal.*, 40(3):597–621, 2006.

[128] S. Luckhaus and Y. Sugiyama. Asymptotic profile with the optimal convergence rate for a parabolic equation of chemotaxis in super-critical cases. *Indiana Univ. Math. J.*, 56(3):1279–1297, 2007.

[129] S. Luckhaus, Y. Sugiyama, and J. J. L. Velázquez. Measure valued solutions of the 2D Keller-Segel system. *Arch. Ration. Mech. Anal.*, 206(1):31–80, 2012.

[130] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[131] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser. *Semiconductor equations*. Springer-Verlag, Wien, 1990.

[132] 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.

[133] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

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

[135] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations*, 48(1-2):1–31, 2013.

[136] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discrete Contin. Dyn. Syst. Ser. S*, 6(2):479–499, 2013.

[137] A. Mielke. Deriving amplitude equations via evolutionary Γ-convergence. *Discrete Contin. Dyn. Syst.*, 35(6):2679–2700, 2015.

[138] A. Mielke. On evolutionary Γ-convergence for gradient systems. In *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, volume 3 of *Lecture Notes in Applied Mathematics and Mechanics*, pages 187–249. Springer International Publishing, Cham, 2016.

[139] A. Mielke, R. Rossi, and G. Savaré. BV solutions and viscosity approximations of rate-independent systems. *ESAIM Control Optim. Calc. Var.*, 18(1):36–80, 2012.

[140] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Partial Differential Equations*, 46(1-2):253–310, 2013.

[141] A. Mielke, T. Roubíček, and U. Stefanelli. Γ-limits and relaxations for rate-independent evolutionary problems. *Calc. Var. Partial Differential Equations*, 31(3):387–416, 2008.

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

[143] Y. Mimura. The variational formulation of the fully parabolic Keller-Segel system with degenerate diffusion, 2012. Preprint.

[144] N. Mizoguchi. Global existence for the Cauchy problem of the parabolic-parabolic Keller-Segel system on the plane. *Calc. Var. Partial Differential Equations*, 48(3-4):491–505, 2013.

[145] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *J. Math. Biol.*, 38(6):534–570, 1999.

[146] D. Morale, V. Capasso, and K. Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.*, 50(1):49–66, 2005.

[147] J. D. Murray. *Mathematical biology. I: An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2002.

[148] J. D. Murray. *Mathematical biology. II: Spatial models and biomedical applications*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2003.

[149] T. Nagai and T. Senba. Global existence and blow-up of radial solutions to a parabolic-elliptic system of chemotaxis. *Adv. Math. Sci. Appl.*, 8(1):145–156, 1998.

[150] T. Nagai, R. Syukuinn, and M. Umesako. Decay properties and asymptotic profiles of bounded solutions to a parabolic system of chemotaxis in $\mathbf{R}^n$. *Funkcial. Ekvac.*, 46(3):383–407, 2003.

[151] L. Noethen and S. Walcher. Tikhonov's theorem and quasi-steady state. *Discrete Contin. Dyn. Syst. Ser. B*, 16(3):945–961, 2011.

[152] L. Onsager. Reciprocal relations in irreversible processes. II. *Physical Review*, 38(12):2265, 1931.

[153] K. Osaki and A. Yagi. Finite dimensional attractor for one-dimensional Keller-Segel equations. *Funkcial. Ekvac.*, 44(3):441–469, 2001.

[154] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[155] F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255, 2005.

[156] B. Perthame. *Transport equations in biology*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2007.

[157] K. Post. *A non-linear parabolic system modeling chemotaxis with sensitivity functions*. PhD thesis, Humboldt-Universität zu Berlin, Institut für Mathematik, 1999.

[158] G. Raoul. Nonlocal interaction equations: stationary states and stability analysis. *Differential Integral Equations*, 25(5-6):417–440, 2012.

[159] R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.

[160] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[161] M. Schauer and R. Heinrich. Quasi-steady-state approximation in the mathematical modeling of biochemical reaction networks. *Math. Biosci.*, 65(2):155–170, 1983.

[162] S. Schuster and R. Schuster. A generalization of Wegscheider's condition. Implications for properties of steady states and for quasi-steady-state approximation. *J. Math. Chem.*, 3(1):25–42, 1989.

[163] L. Segel. A theoretical study of receptor mechanisms in bacterial chemotaxis. *SIAM J. Appl. Math.*, 32(3):653–665, 1977.

[164] L. A. Segel and M. Slemrod. The quasi-steady-state assumption: a case study in perturbation. *SIAM Rev.*, 31(3):446–477, 1989.

[165] T. Senba and T. Suzuki. A quasi-linear parabolic system of chemotaxis. *Abstr. Appl. Anal.*, 2006:1–21, 2006.

[166] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.

[167] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, 1970.

[168] Y. Sugiyama. Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems. *Differential Integral Equations*, 19(8):841–876, 2006.

[169] Y. Sugiyama. Time global existence and asymptotic behavior of solutions to degenerate quasi-linear parabolic systems of chemotaxis. *Differential Integral Equations*, 20(2):133–180, 2007.

[170] Y. Sugiyama and H. Kunii. Global existence and decay properties for a degenerate Keller-Segel model with a power factor in drift term. *J. Differential Equations*, 227(1):333–364, 2006.

[171] H. Sun, D. Uminsky, and A. L. Bertozzi. Stability and clustering of self-similar solutions of aggregation equations. *J. Math. Phys.*, 53(11):115610, 18, 2012.

[172] F. Theil. A proof of crystallization in two dimensions. *Comm. in Math. Phys.*, 262(1):209–236, 2006.

[173] A. N. Tihonov. Systems of differential equations containing small parameters in the derivatives. *Mat. Sbornik N. S.*, 31(73):575–586, 1952.

[174] C. M. Topaz and A. L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.*, 65(1):152–174, 2004.

[175] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.

[176] G. Toscani. One-dimensional kinetic models of granular flows. *M2AN Math. Model. Numer. Anal.*, 34(6):1277–1291, 2000.

[177] G. Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006.

[178] A. E. van Giessen and B. Widom. Path dependence of surface–tension scaling in binary mixtures. *Fluid Phase Equilibria*, 164(1):1–12, 1999.

[179] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2003.

[180] A. I. Vol′pert. Differential equations on graphs. *Mat. Sb. (N.S.)*, 88(130):578–588, 1972.

[181] M. Winkler. Absence of collapse in a parabolic chemotaxis system with signal-dependent sensitivity. *Math. Nachr.*, 283(11):1664–1673, 2010.

[182] T. Yamada. Improvement of convergence rates for a parabolic system of chemotaxis in the whole space. *Math. Methods Appl. Sci.*, 34(17):2103–2124, 2011.

[183] E. Zeidler. *Nonlinear functional analysis and its applications. II/B: Nonlinear monotone operators*. Springer-Verlag, New York, 1990.

[184] J. Zhang and D. Y. Kwok. A mean-field free energy lattice Boltzmann model for multicomponent fluids. *The European Physical Journal Special Topics*, 171(1):45–53, 2009.

[185] J. Zinsl. Existence of solutions for a nonlinear system of parabolic equations with gradient flow structure. *Monatsh. Math.*, 174(4):653–679, 2014.

[186] J. Zinsl. Geodesically convex energies and confinement of solutions for a multi-component system of nonlocal interaction equations, 2014. Preprint. arXiv:1412.3266.

[187] J. Zinsl. A note on the variational analysis of the parabolic–parabolic Keller–Segel system in one spatial dimension. *C. R. Math. Acad. Sci. Paris*, 353(9):849–854, 2015.

[188] J. Zinsl. Exponential convergence to equilibrium in a Poisson-Nernst-Planck-type system with nonlinear diffusion. *Discrete and Continuous Dynamical Systems*, 36(5):2915–2930, 2016.

[189] J. Zinsl and D. Matthes. Exponential convergence to equilibrium in a coupled gradient flow system modeling chemotaxis. *Anal. PDE*, 8(2):425–466, 2015.

[190] J. Zinsl and D. Matthes. Transport distances and geodesic convexity for systems of degenerate diffusion equations. *Calc. Var. Partial Differential Equations*, 54(4):3397–3438, 2015.