TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

Unterschrift der Begutachter

....

....

D I S S E R T A T I O N

Underlying metric structures of some evolution equations with focus on the Wasserstein distance

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der Naturwissenschaften unter der Leitung von

Univ.-Prof. Dr. rer. nat. Ansgar Jüngel
E101
Institut für Analysis und Scientific Computing

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformatik

von

Mag. rer. nat. Philipp Fuchs
Matrikelnummer: 021730
Grundsteingasse 5/15, 1160 Wien

Wien, am 15.09.2015

Kurzfassung

In dieser Arbeit werden Evolutionsgleichungen von verschiedenen metrischen Blickpunkten aus untersucht. Zum einen wird die Herleitung quantenmechanischer Modelle aus Lagrangeschen Prinzipien auf dem Raum der Wahrscheinlichkeitsmaße untersucht, zum anderen die Lösung von Gradientenflussgleichungen mit Hilfe numerischer Verfahren höherer Ordnung.

Im ersten Teil beschäftigen wir uns mit der oben angesprochenen Herleitung quantenmechanischer Modelle. Wir nutzen dazu die Darstellung verschiedener quantenmechanischer Prozesse als fluidmechanische Gleichungen. Von diesen können wir zeigen, dass sie als Euler-Lagrange-Gleichungen gewisser Lagrangefunktionen, auf dem Raum der Wahrscheinlichkeitsmaße, aufgefasst werden können. Hierbei spielt die Wassersteinmetrik eine wesentliche Rolle. Wir zeigen wie diese Lagrangefunktionen auf dem Raum der Wahrscheinlichkeitsmaße von bekannten Lagrangefunktionen für Punktteilchen abgeleitet werden können. Wir illustrieren die Herleitung der Modelle aus Lagrangeschen Prinzipien anhand der elektromagnetischen Schrödingergleichung. Weiters untersuchen wir die Hamiltonsche Struktur der elektromagnetischen Schrödingergleichung, wie sie in unserer Herangehensweise zu Tage tritt. Wir vergleichen diese mit einer bekannten Darstellung der elektromagnetischen Schrödingergleichung als Hamiltonschen Fluss. Wir erweitern die Menge der Modelle, welche durch unseren Zugang dargestellt werden können, indem wir Lagrangefunktionen mit dissipativen Potentialen betrachten. Davon leiten wir exemplarisch die sogenannte Quanten-Navier-Stokes-Gleichung ab.

Im zweiten Teil werden sogenannte One-Leg Schemata für Gradientenflüsse auf Hilberträumen untersucht. Wir sammeln einige in der Literatur bekannte Resultate, geben aber alternative Beweise an. Einerseits zeigen wir, dass Lösungen für die die One-Leg Schemata definierenden Gleichungen eindeutig gegeben sind, und zwar für Gradientenflüsse, deren rechte Seite durch sogenannte λ-konvexe Entropien gegeben sind. Zum anderen geben wir einen Beweis für die Konvergenzordnung der untersuchten Verfahren an, welcher nach unserer Ansicht besser für den Fall unendlichdimensionaler Hilberträume geeignet ist. An die Konvergenz- und Existenzanalyse schließen wir eine Untersuchung der strukturerhaltenden Eigenschaften der diskreten Lösung an. Genauer untersuchen wir, ob die diskreten Lösungen die, aus physikalischer Sicht, wünschenswerte Eigenschaft tragen, die Entropie zu dissipieren. Dies ist im Allgemeinen jedoch nicht der

Fall,wie wir zeigen können. Allerdings stellt sich heraus, dass eine andere Größe, die sogenannte G-Norm, von der diskreten Lösung dissipiert wird, falls die Entropie ein konvexes Funktional ist. Wir schließen die Untersuchung der Gradientenflüsse auf Hilberträumen mit numerischen Experimenten zur Porösen-Medien-Gleichung in einer Raumdimension, welche durch einen Gradientenfluss bezüglich der H^{-1}-Norm gegeben ist. Wir wählen den Exponenten in der Porösen-Medien-Gleichung so, dass sie einem Modell aus der Halbleitersimulation entspricht.

Im dritten Teil beschäftigen wir uns wieder mit Gradientenflüssen, dieses Mal allerdings auf dem Raum der Wahrscheinlichkeitsmaße, ausgestattet mit der Wassersteinmetrik. In der analytischen Untersuchung von Wassersteingradientenflüssen spielt eine von dem BDF-1 Verfahren abgeleitete Zeitdiskretisierung eine wesentliche Rolle. Ausgehend von dieser stellen wir Diskretisierungen höherer Ordnung in der Zeit vor. Wir nutzen dann ein an das BDF-2 Verfahren angelehnte Methode, um eine nichtlineare Diffusionsgleichung, welche eine Formulierung als Wassersteingradientenfluss zulässt, numerisch zu lösen. Um ein vollständig diskretisiertes Problem zu erhalten verwenden wir einen Ansatz mit finiten Elementen zweiter Ordnung in einer Raumdimension. Unsere numerischen Untersuchungen zeigen Konvergenz zweiter Ordnung, sowohl bezüglich der Raumdiskretisierung, als auch der Zeitdiskretisierung. Im Rahmen unserer numerischen Experimente untersuchen wir auch die numerische Abklingrate verschiedener Entropiefunktionale und deren Abhängigkeit von verschiedenen numerischen Parametern.

Contents

Chapter 1

Introduction

1.1 Abstract

This thesis is concerned with the investigation of evolution equations from different "metric" points of view. We derive quantum mechanical models from Lagrangian principles on the so called Wasserstein space. Furthermore we study the numerical treatment of gradient flows. Gradient flows are defined by equations of the form $u_t = \nabla^* \phi(u)$, where $*$ indicates the dependence of the gradient on the underlying metric structure. Higher order fully discrete numerical schemes in one space dimension are studied for gradient flows on Hilbert spaces, as well as fully discrete higher order schemes in one space dimension for gradient flows on the Wasserstein space.

1.2 Elements of optimal transport the Wasserstein metric and Otto's calculus

In this section we present some preliminaries on the Wasserstein metric, the Wasserstein space and Otto's calculus which will be needed later. The Wasserstein metric is a metric defined on the space of probability $\mathscr{P}(X)$ measures over some space X. It arises as the minimal cost of an optimal transportation problem. In optimal transport theory, as the name indicates, one is interested in transporting mass in an optimal way. A mathematical description of this problem is given by the Kantorovich problem. Let two probability spaces (X, μ) and (Y, ν) and a cost function $c : X \times Y \to \mathbb{R}$ be given. We call a measure π on the product space $X \times Y$ a transport plan, and the value $\pi(A, B)$ is the amount of mass transported from A to B. The measure should be constructed so that all the mass that lies in A is transported somewhere in Y, and conversely that all the mass that is transported from X to B is the amount needed at B. More precisely, π is subject to the following conditions

$$\pi[A \times Y] = \mu(A), \qquad \pi[X \times B] = \nu(B)$$

for all measurable $A \subseteq X$ and $B \subseteq Y$. The cost of transporting μ to ν according to the plan π is then given by

$$\int_{X \times Y} c(x,y) d\pi(x,y).$$

We denote the set of all transport plans by $\Pi(\mu, \nu)$. The minimal cost is

$$T_c(\mu,\nu) := \min \left\{ \int_{X \times Y} c(x,y) d\pi(x,y) : \pi \in \Pi(\mu,\nu) \right\}$$

If (X, d) is a metric space and $Y = X$ we can take $c(x, y) = d^p(x, y)$ and define the p-th Wasserstein distance between μ and ν by:

$$W_p^p(\mu,\nu) = \min \left\{ \int_{X \times Y} d^p(x,y) d\pi(x,y), \quad \pi \in \Pi(\mu,\nu) \right\}$$

It can be shown that there exists a $\pi \in \Pi(\mu, \nu)$ realizing the minimal transportation cost, see [62]. In this thesis we only consider the case $p = 2$ and call the W_2 the Wasserstein metric for short. Determining the transport plan π minimizing the cost answers our question of transporting mass in an optimal way. However, the plan π just tells us how much mass has to be transported from A to B and not which "path" each unit of mass has to take. Under suitable regularity assumptions such a dynamic picture is given via the so called Benamou-Brenier formula, see [6, 62]. We only sketch the idea. Suppose you are given a velocity field $v(t, x)$ on $\mathbb{R}^n$ and a mass density $\mu(t, x)$ flowing along this field (assuming absolute continuity of the measure μ we by abuse of notation identify the measure in the following with its density). At each time we define the total kinetic energy by

$$E(t) = \int_{\mathbb{R}^n} \mu(t,x)|v(t,x)|^2 dx$$

and the action by

$$A[\mu, v] = \int_0^1 \left(\int_{\mathbb{R}^n} \mu(t,x)|v(t,x)|^2 dx \right) dt.$$

Consider two mass densities μ_0, μ_1 with compact support. Then it turns out that $W_2^2(\mu_0, \mu_1)$ is obtained by minimizing the action:

$$W_2^2(\mu_0,\mu_1) = \inf \left\{ A[\mu, v] : (\mu, v) \in V(\mu_0,\mu_1) \right\},$$

where $V(\mu_0, \mu_1)$ is the set of reasonable (μ, v), for details see [61, Chapter 5].

In other words, the Wasserstein distance is the minimal action of all velocity fields v_t pushing μ_0 to μ_1. The relation between the action A and the Wasserstein metric already hints towards the Riemannian structure of $\mathscr{P}(\mathbb{R}^n)$. The action depends on the norm of the velocity vector that is calculated with respect to the norm on the tangent space. Here we identify the tangent space $T\mathscr{P}(\mathbb{R}^n)$ with the set of all velocity vectors $\partial_t\mu$. To furnish the space of probability measures with a geometric structure that yields the Wasserstein metric as an intrinsic metric (i.e. geodesic distance) we define the norm of a tangent vector by

$$\|\partial_t\mu\|_\mu^2 = \inf_{v\in L^2(\mathrm{d}\mu,\mathbb{R}^n)} \left\{ \int_{\mathbb{R}^n} \mu|v|^2 dx :\ \partial_t\mu + \mathrm{div}(\mu v) = 0 \right\}.$$

Consequently, the norm of the velocity vector is given via the total kinetic energy, minimized over all velocity fields for that the continuity equation $\partial_t\mu_t + \mathrm{div}(\mu v) = 0$ holds true. Since there are many solutions to the continuity equation, it is necessary to minimize over all admissible velocity fields. Heuristically this minimization is motivated as follows: The flow μ_t describes the flow of the particle density and hence does not give full information about the flow of each particle. If we want to associate to each flow of measures a velocity field, we have to choose a velocity field which contains the "measure relevant" information. Imagine a radially symmetric measure on the unit circle and let v_r be the velocity field of particles rotating around the origin. From a measure point of view nothing happens, therefore v^r produces the same flow as the zero vector field $v^0 \equiv 0$. It seems reasonable to think of v^0 as the correct velocity field to the constant flow of measures $\mu_t = \mu_0$.
It can be shown (again under regularity assumptions on μ and v), that the infimum in the definition of the norm above is attained by a unique velocity field of gradient type ∇u. Hence we end up with the the definition

$$\|\partial_t\mu\|_\mu = \int \mu|\nabla u|^2 \ \mathrm{d}x; \quad \partial_t\mu + \mathrm{div}(\mu\nabla u) = 0.$$

Using this definition of a norm we get the desired identity

$$W_2^2(\mu_0,\mu_1) = \inf\left\{ \int_0^1 \|\partial_t\mu\|_\mu^2 dt; \quad \mu(0) = \mu_0,\ \mu(1) = \mu_1 \right\},$$

which identifies the Wasserstein metric as an intrinsic metric. Looking at this result from a less rigorous point of view one could interpret the norm as follows. We want to give a meaning to velocity and kinetic energy, respectively of a flow of measures. The "mathematical" velocity $\partial_t\mu$ is itself a measure and does not contain the desired physical information. However, we can associate a vector field to each flow via the continuity equation. This vector field contains information of particles flowing around in such a way as the particles density changes (and has the correct physical dimension of a velocity). Ultimately, we use the kinetic

energy of this velocity field to define the kinetic energy of $\partial_t \mu$.
The above derivation of the Riemannian structure of the space of measures is purely formal and taken from [61, Chapters 1,5,8]. The idea of furnishing the space of probability measures with such a geometric structure goes back to Felix Otto and is called Otto's calculus. We already pointed out that to make these ideas precise, one has to make regularity assumptions. A rigorous treatment of the Riemannian structure of the space of measures under such regularity assumptions is given in [43].

1.3 Summary of Chapter 1

Chapter 1 is dedicated to the formal derivation of quantum fluid models from Lagrangian principles on the Wasserstein space. It is well known that the Schrödinger equation, via the so called Madelung transform, admits a fluid mechanical description. It was shown in [55] by von Renesse that the Schrödinger equation in its Madelung representation is a lift of Newton's second law using Otto's calculus (as sketched above). Von Renesse used the formulas of Otto's calculus developed by Lott in [43] to derive a formulation of the Schrödinger equation as Newton's second law on the space of probability measures. Inspired by this result we pursue another direction and derive general quantum fluid models from Lagrangian principles. Recall that in Lagrangian mechanics the equation of motion is obtained from the principle of least action. The action is the functional

$$\int_0^T L(x(t), \dot{x}(t))dt.$$

where $L : TM \to \mathbb{R}$ is the Lagrangian, and TM is the tangent bundle over some manifold $M \subseteq \mathbb{R}^d$ (called the configuration space). We set $M = \mathbb{R}^d$. The criticality of a curve $\gamma : [0,T] \to \mathbb{R}^d$ is (formally) equivalent to the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}(\gamma, \dot{\gamma}) - \frac{\partial L}{\partial x}(\gamma, \dot{\gamma}) = 0, \quad t \in (0,T).$$

We define a lift of Lagrangians $L : T\mathbb{R}^d \to \mathbb{R}$ to Lagrangians $\mathscr{L} : T\mathscr{P}(\mathbb{R}^d) \to \mathbb{R}$ by:

$$\mathscr{L}(\mu, \eta) := \inf \left\{ \int L(x, v(x))\mu(dx) : v \in C^\infty(\mathbb{R}^d; \mathbb{R}^d), \eta + \operatorname{div}(\mu v) = 0 \right\},$$

Compare the choice of the Lagrangian as the infimum over all "reasonable" velocity fields with the definition of $\|\mu_t\|_\mu^2$ *as sketched above.*

We show that this definition induces the Wasserstein metric as an intrinsic distance on the space of measures.

We show (Theorem 2.3.1) that the lifted Lagrangian for a quantum particle in an electromagnetic field is calculated to be

$$\mathscr{L}(\mu,\eta) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 - \frac{1}{2}|A|^2 - \Phi(x) - \frac{\hbar^2}{8}|\nabla \log \mu|^2 \right) \mu(dx)$$

and that the Euler-Lagrange equations are given by the continuity equation for the particle density μ and the Hamilton-Jacobi equation for the velocity potential S

$$\begin{aligned} &\partial_t \mu + \operatorname{div}\big(\mu(\nabla S - A)\big) = 0, \\ &\partial_t S + \frac{1}{2}|\nabla S - A|^2 + \Phi(x) - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} = 0 \quad \text{in } \mathbb{R}^d,\ t > 0, \end{aligned}$$

where, with a slight abuse of notation, $\hbar$ is the *scaled* Planck constant. We introduce the wave function $\Psi = \sqrt{\mu}\exp(iS/\hbar)$ using the so-called Madelung transform, for smooth solutions (μ, S) with positive density (or mass distribution) μ. Then Ψ solves the Schrödinger equation with vector potential A and electric potential $\Phi(x)$, i.e.

$$i\hbar\partial_t\Psi = \frac{1}{2}\left(\frac{\hbar}{i}\nabla - A\right)^2 \Psi + \Phi(x)\Psi \quad \text{in } \mathbb{R}^d,\ t > 0.$$

We present a systematic analysis of the Madelung transform as a symplectic map between Hamiltonian systems, preserving the electromagnetic Schrödinger Hamiltonian (see Theorem 2.3.6). The term $\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}$ in the Hamilton-Jacobi equation above is called Bohm potential. The fluid mechanical representation of the Schrödinger equation is closely related to the de Broglie-Bohm theory. One of the main attempts of the de Broglie-Bohm theory is the formulation of quantum mechanics avoiding the problem of quantum mechanical measurement. The analysis of the interplay of optimal transport and the de Broglie-Bohm theory is beyond the scope of this thesis and subject to further research. However, for the reader interested in more details on the Bohmian representation of quantum mechanics we recommend [23, 63]. To derive a wider class of quantum fluid models we augment the Lagrangians by dissipation potentials and show that the lifted Euler-Lagrange equations with linear friction lead to the quantum Navier-Stokes equations. After identifying vector fields modulo rotational components, these equations read as (see Theorem 2.4.1)

$$\partial_t \mu + \operatorname{div}(\mu v) = 0, \tag{1.1}$$

$$\partial_t(\mu v) + \operatorname{div}(\mu v \otimes v) + \nabla p(\mu) + \mu\nabla\Phi(x) - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) = \alpha \operatorname{div}(\mu D(v)), \tag{1.2}$$

where the velocity is given by $v = \nabla S$, $v \otimes v$ is a matrix with components $v_j v_k$, $p(\mu)$ is the pressure, and $D(v) = \frac{1}{2}(\nabla v + \nabla v^\top) = \nabla v$ is the symmetric velocity gradient. Some words on the model here: This system was first derived by Brull and Méhats [9] from the Wigner-BGK equation (named after Bhatnagar, Gross, and Krook) using a Chapman-Enskog expansion. An alternative derivation from the Wigner-Fokker-Planck model by just applying a moment method was proposed in [38]. For systems including the energy equation, we refer to [38, 39]. Our approach yields a third way to derive the quantum Navier-Stokes equations. An advantage of our method is that we can propose more general friction terms, leading to a variety of nonlinear viscosities (see Remark 3). The selection of quantum mechanically correct dissipation terms remains a research topic for the future (see [3] for a Lindblad equation approach).

Surprisingly, system (1.1)-(1.2) allows for *two* different energies, as observed in [36]. Indeed, a formal computation shows that the Hamiltonian

$$\mathscr{H}_{\mathrm{Q}} = \int_{\mathbb{R}^d} \left(\frac{1}{2}|v|^2 + U(\mu) + \Phi(x) + \frac{\hbar^2}{8}|\nabla \log \mu|^2 \right) \mu dx$$

is a Lyapunov functional along the solutions to (1.1)-(1.2), see Proposition 3. Here, the internal energy U relates to the pressure p by $p'(s) = sU''(s)$, $s > 0$. Furthermore, the energy

$$\mathscr{H}_{\mathrm{Q}}^* = \int_{\mathbb{R}^d} \left(\frac{1}{2}|v + v_{\mathrm{os}}|^2 + U(\mu) + \Phi(x) + \left(\frac{\hbar^2}{8} - \frac{\alpha^2}{2} \right) |\nabla \log \mu|^2 \right) \mu dx,$$

where $v_{\mathrm{os}} = \alpha \nabla \log \mu$ is the osmotic velocity, is another Lyapunov functional. We will explain this fact by a variant of the Noether theorem. Indeed, time invariance of the system leads to dissipation of the Hamiltonian $\mathscr{H}_{\mathrm{Q}}$ (since we have friction, the energy is not a constant of motion). Interestingly, a special transformation of the variables (t, μ) leads to a Noether current which equals $\mathscr{H}_{\mathrm{Q}}^*$ (see Theorem 2.4.2). Thus, the existence of the second energy functional is a consequence of a "Noether symmetry", showing that the quantum Navier-Stokes equations exhibit a certain geometric structure.

1.4 Summary of Chapter 2

Chapter 2 is dedicated to the study of numerical schemes for the solution of gradient flows on Hilbert spaces, i.e. equations of the form:

$$u_t = \nabla^* \phi(u),$$

where $*$ denotes that the gradient taken with respect to the inner product structure of the space. We collect some results regarding existence, convergence, and dissipation of the scheme. Numerous evolution equations in science and engineering turn out to admit a gradient flow formulation on a Hilbert space. Linear equations like the heat equation, Fisher's equation for population dynamics or the Allen-Cahn equation modeling phase separation in iron alloy, are known to constitute gradient flows with respect to the L^2-norm. Gradient flows with respect to the L^2-inner product also play a role in computer vision in active contour problems. As examples of nonlinear equations (which admit a gradient flow structure on a Hilbert space) we mention the Cahn-Hilliard equation, modeling phase separation, which constitutes a gradient flow with respect to the H^{-1}-norm. Furthermore we mention the porous medium equation, which we investigate numerically at the end of Chapter 2. The first part of the Chapter part is dedicated to the analysis of so-called one-leg schemes for gradient flows on Hilbert spaces for λ-convex entropies ϕ. One-leg schemes are a concept introduced by Germund Dahlquist in order to study general multistep schemes. One-leg schemes are a well known and well studied concept. For results on one-leg scheme we recommend [34]. However, we give an alternative proof on the convergence of the scheme which seem more suitable for our analysis. One-leg schemes are derived from linear multistep schemes and related via the following definition. Given a general linear multistep scheme

$$\sum_{i=0}^{k} \alpha_i x_{m+i} = \tau \sum_{i=0}^{k} \beta_i \nabla \phi(x_{m+i})$$

the corresponding one-leg scheme is defined by:

$$\sum_{i=0}^{k} \alpha_i x_{m+i} = \tau \nabla \phi \left(\sum_{i=0}^{k} \beta_i x_{m+i} \right).$$

In Theorem 3.1.9 we show the existence of solutions to the scheme for λ-convex ϕ. To prove existence we use a formulation of the scheme as a minimization problem. Therefore we call the schemes minimizing movement schemes (MMS) as well. An other reason for this labeling is that in the theory of gradient flows on metric spaces the schemes used to show existence of flows are called minimizing movement schemes as well. In this sense (within this thesis) it seems consistent to call the schemes with this familiar structure (on Hilbert or on Wasserstein space) by the same name. In Theorem 3.1.17 we give a proof of the convergence of the schemes and their respective order. After that, we tackle the question of structure preserving properties of the scheme. By definition the solution of the flow dissipates entropy. A desirable property of the scheme is that the discrete solution dissipates the discrete entropy as well. Unfortunately this will not hold true for general schemes. Moreover, we show that no other Lyapunov functional, depending only on the state of the system at time t_k, will be dissipated by the discrete solution. We illustrate the violation of dissipation of the discrete solution by a counterexample. Although there cannot exist a quantity depending on a single value of the discrete solution we observe that the so called G-norm, a quantity depending on the discrete solution at several time steps is dissipated, at least for convex functionals. The G-norm was introduced by Dahlquist for the error analysis and is employed in our proof of the convergence of the scheme.

After analysis of the scheme we use the BDF-2 scheme (which belongs to the class of one-leg schemes) for numerical analysis of the porous-medium equation. We study the special case

$$\begin{cases} u_t = \Delta u^{5/3}, & x \in [0,1];\ t \in (0,T), \\ u(t,0) = u(t,1) = 0, & t \in [0,T], \\ u(0,x) = u_0, & x \in [0,1]. \end{cases}$$

This equation arises in semiconductor modeling, see [35]. We fully discretize the equation by a linear FEM ansatz in space. We show numerically first order convergence of the scheme in space and the predicted second order convergence in time. We show that the scheme, as expected, dissipates the G-norm. Moreover, the entropy in our example is dissipated, although as mentioned above, in general we may not hope for dissipation. We compare the decay of the G-norm and the entropy. Finally, for the sake of completeness we give a numerical example where the entropy is not dissipated by the scheme.

1.5 Summary of Chapter 3

In Chapter 3 we turn our attention to gradient flows on the Wasserstein space. The investigation of Wasserstein gradient flows has become a active field of research in recent years, since it turns out that a wide class of evolution equations constitute Wasserstein gradient flows. The theory on Wasserstein gradient flows is covered by [2]. Central to the investigation of Wasserstein gradient flows is the minimizing movement scheme.

Minimizing movement schemes for evolution equations with an underlying gradient-flow structure were first suggested by De Giorgi [21] in an abstract framework. Jordan, Kinderlehrer, and Otto [40] have shown that the solution to the linear Fokker-Planck equation can be obtained by minimizing the logarithmic entropy in the Wasserstein space. Since then, many nonlinear evolution equations have been shown to constitute Wasserstein gradient flows, for instance the porous-medium equation [51], the Keller-Segel model [12], equations for interacting gases [18], and a nonlinear fourth-order equation for quantum fluids [31].

The aim of Chapter 3 is to study fully discrete higher order variants of minimizing movement schemes in one spatial dimension on the Wasserstein space for the nonlinear diffusion equation

$$\partial_t u = \alpha^{-1}\Delta(u^\alpha) \quad \text{in } \mathscr{T}^d,\ t > 0, \quad u(0) = u^0, \tag{1.3}$$

with negative exponent $\alpha < 0$, where $\mathscr{T}^d$ is the d-dimensional torus. A short computation reveals that equation (4.1) can be written as the gradient flow of the entropy $S[u] = (\alpha(\alpha-1))^{-1}\int_{\mathscr{T}^d} u^\alpha dx$ with respect to the Wasserstein distance. For the (formal) derivation of the Wasserstein gradient see Appendix A We calculate

$$\nabla^W S[u] = -\operatorname{div}(u(\nabla DS[u])) = -\operatorname{div}\left(u\frac{1}{\alpha(\alpha-1)}(\alpha(\alpha-1)u^{\alpha-2}\nabla u)\right),$$

which yields after a straight forward calculation

$$-\nabla^W S[u] = \alpha^{-1}\Delta(u^\alpha).$$

The minimizing movement scheme can be interpreted as an implicit Euler semi-discretization with respect to the Wasserstein gradient-flow structure. Due to the high computational cost, there are not many results on numerical approximations of evolution equations using this scheme. In one space dimension, the optimal transport metric becomes flat when re-parametrized by means of inverse cumulative functions, which simplifies the numerical solution; see e.g. [1, 12, 44, 50]. For multi-dimensional situations, one approach is based on the Eulerian representation of the discrete solution on a fixed grid. The resulting problem can be solved by using interior point methods [16], finite elements [15], or finite volumes [17]. Another approach employs the Lagrangian representation, which is well adapted to optimal transport. Examples are moving meshes [14], linear finite elements for a fourth-order equation [22], and entropic smoothing using the Kullback-Leibler divergence [53]. The connection between Lagrangian schemes and the gradient-flow structure was investigated in [41]. In this thesis, we will use the Lagrangian viewpoint.

The minimizing movement scheme of De Giorgi is of first order in time only since it is based on the implicit Euler method. Concerning higher-order schemes, we are only aware of the paper [64]. There, second-order gradient-flow schemes were suggested for the Euler equations, with finite differences in space and the two-step BDF (Backward Differentiation Formula) method or diagonally implicit Runge-Kutta (DIRK) schemes in time.

In this thesis, we propose a fully discrete second-order minimizing movement scheme using quadratic finite elements in space and the two-step BDF method in time. We consider periodic point-symmetric solutions. The finite-dimensional minimization problem, constrained by the mass conservation, is solved by the method of Lagrange multipliers which leads to a sequential quadratic programming problem.

By construction, our numerical scheme is of second order both in time and space, it conserves the mass and dissipates the G-norm of the Lagrangian weight vector $\mathbf{g}^k$ at time step k,

$$\|(\mathbf{g}^{k+1},\mathbf{g}^k)\|_G^2 = \frac{5}{2}(\mathbf{g}^{k+1})^\top M_w \mathbf{g}^{k+1} - 2(\mathbf{g}^{k+1})^\top M_w \mathbf{g}^k + \frac{1}{2}(\mathbf{g}^k)^\top M_w \mathbf{g}^k, \tag{1.4}$$

where the matrix M_w is defined in the approximation of the Wasserstein metric on the space of the quadratic ansatz functions. We refer to Section 4.2 for details. It turns out that numerically, the relative G-norm decays exponentially fast to zero. Although we cannot expect for the multistep scheme that the discrete entropy decays exponentially fast, this holds true for the numerical experiments performed in this thesis. Furthermore, also the discrete variance of the original variable u and the Lagrangian variable decay exponentially fast.

The dependence of the numerical decay rates on the time step size τ and the (spatial) grid number N is somehow surprising. The numerical tests indicate that the decay rate of the G-norm is increasing with respect to τ and N, i.e., the numerical values are worse for coarser meshes. Moreover, the decay rates of the discrete entropy and the variance are decreasing in N, i.e., the discrete decay rates are better than the corresponding value of the continuous equation. This result is in accordance with the findings of [45] for a finite-volume approximation of a one-dimensional linear Fokker-Planck equation.

Chapter 2

Lagrangian mechanics on Wasserstein space

In this chapter we derive quantum fluid models from Lagrangian principles on the Wasserstein space.

The chapter is organized as follows. The basic setup of Lagrangian mechanics on the set of probability measures is introduced in Section 2.1. The following sections are concerned with three applications of the Lagrangian method. For the particle motion in a potential field, we recover the usual flow equations, showing that our approach includes the classical case (Section 2.2). The Euler-Lagrange equation for a charged particle in an electromagnetic field is computed in Section 2.3, and the symplectic structure of the flow equations is analyzed. Section 2.4 is devoted to the derivation of the quantum Navier-Stokes equations and the relation between energy functionals and the Noether theorem.

2.1 Basic setup

In this section, we extend the classical Lagrangian mechanics to a configuration space consisting of probability measures. A similar approach is contained in the work of Lafferty [42]. We recall the definition of the phase space, introduce the Lagrangians considered in this thesis, and formulate the (dissipative) Euler-Lagrange equations.

2.1.1 Phase space

Let $\mathscr{P}(\mathbb{R}^d)$ $(d \geq 1)$ be the set of probability measures on $\mathbb{R}^d$. Obviously, the space $\mathbb{R}^d$ is embedded in $\mathscr{P}(\mathbb{R}^d)$ via the Dirac masses $x \mapsto \delta_x$. A physical interpretation of $\mu \in \mathbb{R}^d$ is that μ represents a (possibly diffuse) distribution of mass with fixed total amount. The following arguments may be made rigorous on the set $\mathscr{P}^\infty(\mathbb{R}^d)$ of absolutely continuous probability measures with smooth positive density and finite exponential moments, as pointed out by Lott [43]. However, similarly to the previous works [43, 51, 52, 55], we shall not try to find the maximal subset of $\mathscr{P}(\mathbb{R}^d)$ on which our formulas remain valid, and therefore, we assume that the measures $\mu \in \mathscr{P}(\mathbb{R}^d)$ are sufficiently smooth for the formulas to hold. In the following, we often identify the measure $\mu \in \mathscr{P}(\mathbb{R}^d)$ with its density $d\mu/dx \simeq \mu$ and we write $\mathscr{P}$ instead of $\mathscr{P}(\mathbb{R}^d)$.

Given $\mu \in \mathscr{P}$ we introduce the tangent space of $\mathscr{P}$ at μ by

$$T_\mu \mathscr{P} = \{\eta \in \mathscr{S}'(\mathbb{R}^d) : \exists v \in C^\infty(\mathbb{R}^d;\mathbb{R}^d),\ \eta + \operatorname{div}(\mu v) = 0\},$$

where $\mathscr{S}'(\mathbb{R}^d)$ is the dual of the Schwartz space, which is the collection of infinitesimal variations of μ by smooth flows. The tangent bundle

$$T\mathscr{P} = \bigcup_{\mu \in \mathscr{P}} T_\mu \mathscr{P}$$

serves as the physical phase space for our Lagrangian mechanics of mass distributions. We remark that the motion of a single particle with velocity u is included in our formalism by means of the representation $\eta = -\operatorname{div}(\delta_x v)$, where v is any vector field on $\mathbb{R}^d$ satisfying $v(x) = u$. We also notice that in Hamiltonian mechanics, the phase space is defined by the pairs of generalized coordinates in $T\mathscr{P}$ and generalized momenta in the dual space $T^*\mathscr{P}$. We refer to Section 2.3.2 for details.

We note that the case of concentration of measures on singular sets may be more delicate. Therefore, we restrict ourselves to the case of absolutely continuous probability measures with smooth positive density and finite exponential moments. The concept of geometric tangent cones [2, Section 12.4] may help to reformulate some of the ideas presented here.

2.1.2 Lagrangians

A function $\mathscr{L} : T\mathscr{P} \to \mathbb{R}$ is called a Lagrangian. Below, we shall mostly be concerned with Lagrangians $\mathscr{L}$, which are obtained as lifts from classic Lagrange functions $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, defined by

$$\mathscr{L}(\mu, \eta) = \inf \left\{ \int_{\mathbb{R}^d} L(x, v(x))\mu(dx) : v \in C^\infty(\mathbb{R}^d;\mathbb{R}^d),\ \eta + \operatorname{div}(\mu v) = 0 \right\}, \quad (2.1)$$

where $\mu \in \mathscr{P}$ and $\eta \in T_\mu\mathscr{P}$. The infimum is necessary since the map $v \mapsto -\operatorname{div}(\mu v) \in T_\mu\mathscr{P}$ is generally not injective (*compare the choice of a minimizing velocity v for the lifted Lagrangian $\mathscr{L}$ with the choice of a minimizing velocity v for the definition of the norm of $\|\partial_t \mu\|$ in 1.2*). We prefer the notation $\mathscr{L}(\mu, \eta)$ instead of the simpler (and geometrically more consistent) notation $\mathscr{L}(\eta)$ in order to emphasize the importance of the referring base point for η in $T_\mu\mathscr{P}$. Notice that the classical case is embedded in this situation since

$$\mathscr{L}(\delta_x, -\operatorname{div}(\delta_x v)) = \int_{\mathbb{R}^d} L(x, v(x))\delta(dx) = L(x, v(x)).$$

We present some examples studied in this thesis.

Single-particle dynamics.

We lift the kinetic energy $L(q,\dot{q}) = \frac{1}{2}|\dot{q}|^2$ to $\mathscr{P}(\mathbb{R}^d)$ and show that the lifted Lagrangian $\mathscr{L}$ is by definition consistent with Otto's geometric structure on $\mathscr{P}(\mathbb{R}^d)$ as derived in the Introduction 1.2. A standard duality argument shows that the infimum in (2.1) is attained. Indeed, we compute formally, for $\mu \in \mathscr{P}$ and $\eta \in T_\mu\mathscr{P}$:

$$\begin{aligned}\mathscr{L}(\mu,\eta) &= \inf_v \sup_\chi \int_{\mathbb{R}^d} \left(\frac{1}{2}|v|^2 + \frac{\eta\chi}{\mu} - v\cdot\nabla\chi\right)\mu(dx)\\ &= \sup_\chi \inf_v \int_{\mathbb{R}^d} \left(\frac{1}{2}|v-\nabla\chi|^2 - \frac{1}{2}|\nabla\chi|^2 + \frac{\eta\chi}{\mu}\right)\mu(dx).\end{aligned}$$

The infimum is realized at $v = \nabla\chi$:

$$\mathscr{L}^* = \mathscr{L}(\mu,\eta) = \sup_\chi \int_{\mathbb{R}^d}\left(\frac{\eta\chi}{\mu} - \frac{1}{2}|\nabla\chi|^2\right)\mu(dx).$$

Defining $S = \operatorname{argsup}\mathscr{L}^*$ and inserting $v = \nabla S$, $\chi = S$ into $\mathscr{L}$, we find that

$$\mathscr{L}(\mu,\eta) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla S|^2 \mu(dx).$$

We recall that $S : \mathbb{R}^d \to \mathbb{R}$ is the (up to constants) unique solution to $-\operatorname{div}(\mu\nabla S) = \eta$ in $\mathbb{R}^d$. The function S is called the velocity potential of the variation η with respect to the state μ. We introduce the notation

$$\Delta_\mu S = \operatorname{div}(\mu\nabla S) \quad \text{in } \mathbb{R}^d. \tag{2.2}$$

The minimizer defines a quadratic form on the tangent space $T_\mu\mathscr{P}$:

$$\|\eta\|^2_{T_\mu\mathscr{P}} = \int_{\mathbb{R}^d} |\nabla S(x)|^2 \mu(dx).$$

This is Otto's Riemannian tensor on $T\mathscr{P}$ and shows the consistency of the lifted $\mathscr{L}$ with Otto's Calculus.

Charged particles in an electromagnetic field.

The Lagrange function $L(q,\dot{q}) = \frac{1}{2}|\dot{q}|^2 + \dot{q}\cdot A - \Phi(x)$ models the motion of a charged particle in an electromagnetic field, where $A : \mathbb{R}^d \to \mathbb{R}^d$ is the magnetic vector potential [58, Section 12.6] and $\Phi : \mathbb{R}^d \to \mathbb{R}$ is the electric potential. By a similar computation as in the previous example, for $\mu \in \mathscr{P}$ and $\eta \in T_\mu\mathscr{P}$,

$$\mathscr{L}(\mu,\eta) = \sup_\chi \inf_v \int_{\mathbb{R}^d}\left(\frac{1}{2}|v + (A-\nabla\chi)|^2 - \frac{1}{2}|A - \nabla\chi|^2 + \frac{\eta\chi}{\mu} - \Phi(x)\right)\mu(dx).$$

Then, taking $v^* = \nabla\chi - A$ to realize the infimum and $S = \operatorname{argsup} \mathscr{L}$, $\chi = S$, it holds that

$$\begin{aligned}\mathscr{L}(\mu, \eta) &= \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S - A|^2 + (\nabla S - A) \cdot A - \Phi(x) \right) \mu(dx) \\ &= \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 - \frac{1}{2}|A|^2 - \Phi(x) \right) \mu(dx), \end{aligned} \tag{2.3}$$

and $S : \mathbb{R}^d \to \mathbb{R}$ is the (up to constants) unique solution to

$$\eta = -\operatorname{div}(\mu v^*) = -\operatorname{div}(\mu(\nabla S - A)) = -\operatorname{div}(\mu\nabla S) + \operatorname{div}(\mu A) \quad \text{in } \mathbb{R}^d.$$

With the notation (2.2), we have $S = -\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))$ in $\mathbb{R}^d$.

Charged quantum particles.

Subtracting from the kinetic energy of the previous example the Fisher information $I(\mu)$, defined by

$$I(\mu) = \int_{\mathbb{R}^d} |\nabla \log \mu|^2 \mu(dx), \tag{2.4}$$

the lifted Lagrangian

$$\mathscr{L}(\mu, \eta) = \frac{1}{2}\|\eta\|^2_{T_\mu \mathscr{P}} - V(\mu) - \frac{\hbar^2}{8} I(\mu) \tag{2.5}$$

was considered by Lafferty [42] and von Renesse [55] to formulate the Schrödinger equation by means of the Madelung equations. We remark that Feng and Nguyen [26] employed $-I(\mu)$ instead of $I(\mu)$ to derive compressible Euler-type equations from minimizers of an action functional defined on probability measure-valued paths. One may augment $\mathscr{L}$ also by the internal energy term

$$-\int_{\mathbb{R}^d} U(\mu)\mu(dx), \tag{2.6}$$

where $U : \mathbb{R} \to \mathbb{R}$ is the (smooth) internal energy potential.

2.1.3 Smooth curves in $\mathscr{P}$

Let $\mu : [0, T] \to \mathscr{P}$ be a smooth curve, i.e., its time derivative $\dot\mu_t := \partial_t \mu(t)$ exists in the distributional sense and $\dot\mu_t \in T_{\mu_t}\mathscr{P}$ for all $t \in [0, T]$. For instance, $\dot\mu_t \in \mathscr{S}'(\mathbb{R}^d)$ may be defined for each $t \in [0, T]$ by

$$\partial_t \langle \mu_t, \xi \rangle = \langle \dot\mu_t, \xi \rangle \quad \text{for all } \xi \in \mathscr{S}(\mathbb{R}^d),$$

where $\langle \cdot, \cdot \rangle$ is the dual product between $\mathscr{S}'(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$.

Let $\mu : [0,T] \to \mathscr{P}$ be a smooth curve. If $\dot\mu_t \in \mathscr{S}'(\mathbb{R}^d)$ is regular and $\mu_t \in \mathscr{P}^\infty(\mathbb{R}^d)$ (see Section 2.1.1 for the definition of $\mathscr{P}^\infty(\mathbb{R}^d)$), standard elliptic theory provides the existence of (up to an additive constant) unique smooth solution $S_t : \mathbb{R}^d \to \mathbb{R}$ to the problem

$$-\operatorname{div}(\mu_t \nabla S_t) = \dot\mu_t \quad \text{in } \mathbb{R}^d.$$

In particular, the curve $\dot\mu : (0,T) \to T\mathscr{P}$, $t \mapsto \eta_t := \dot\mu_t = -\operatorname{div}(\mu_t\nabla S_t)$ is well defined and, by definition of the tangent space, $\eta_t \in T_{\mu_t}\mathscr{P}$. Again, the single-particle motion $c : [0,T] \to \mathbb{R}^d$ is included by taking $\gamma_t = \delta_{c(t)}$ and $\eta_t = -\operatorname{div}(v_t\delta_{c(t)}) \in \mathscr{S}'(\mathbb{R}^d)$, where v_t is some vector field such that $v_t(x) = \dot c(t)$ for $x \in \mathbb{R}^d$.

2.1.4 Action functional and critical points

Given a Lagrangian $\mathscr{L}$ on $\mathscr{P}$ (see Section 2.1.2), we define the action functional on smooth curves $\gamma : [0,T] \to \mathscr{P}$ by

$$\mathscr{A}(\gamma) = \int_0^T \mathscr{L}(\gamma_t, \dot\gamma_t)dt.$$

A critical point of $\mathscr{A}$ is a curve γ which satisfies

$$\frac{d}{ds}\mathscr{A}(\gamma^s)\Big|_{s=0} = 0$$

for all smooth variations $\gamma : [-\varepsilon,\varepsilon] \times [0,T] \to \mathscr{P}$, $(s,t) \mapsto \gamma_t^s$, satisfying $\gamma_t^0 = \gamma_t$ for $t \in [0,T]$. Hence, assuming differentiability of $\mathscr{L}$, a curve is a critical point if and only if it satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\eta}(\gamma,\dot\gamma) - \frac{\partial\mathscr{L}}{\partial\mu}(\gamma,\dot\gamma) = 0, \tag{2.7}$$

where $\partial\mathscr{L}/\partial\eta$ and $\partial\mathscr{L}/\partial\mu$ are the variational derivatives of $\mathscr{L}$ with respect to η and μ, respectively (see Section 2.3.1). A Lagrangian system on $\mathscr{P}$ with friction is modeled by means of a dissipative potential $\mathscr{D} : T\mathscr{P} \to \mathbb{R}$:

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\eta}(\gamma,\dot\gamma) - \frac{\partial\mathscr{L}}{\partial\mu}(\gamma,\dot\gamma) + \frac{\partial\mathscr{D}}{\partial\eta}(\gamma,\dot\gamma) = 0. \tag{2.8}$$

Von Renesse identified in [55] the flow (2.7), with $\mathscr{L}$ given by (2.5), with the Schrödinger equation in its Madelung representation. We extend this concept in the following sections for more general Lagrangians.

2.2 Example 1: Particle motion in a potential field

We show that the formalism of Section 2.1 includes as a special case the motion of a single particle in a potential $\Phi(x)$. Indeed, choosing the Lagrangian as the lift of the classical Lagrangian $L(q, \dot{q}) = \frac{1}{2}|\dot{q}|^2 - \Phi(x)$, the arguments in Section 2.1.2 yield, for vector fields $v \in C^\infty(\mathbb{R}^d; \mathbb{R}^d)$,

$$\mathscr{L}(\delta_x, -\operatorname{div}(\delta_x v)) = L(x, v(x)) = \frac{1}{2}|v(x)|^2 - \Phi(x).$$

Elementary computations show that curves $\gamma_t = \delta_{x_t}$ with $\ddot{x}_t = -\nabla\Phi(x_t)$ are critical flows for the corresponding lifted action functional $\mathscr{A}$, i.e., γ_t is a critical point for $\mathscr{A}$ (see Section 2.1.4).

Clearly, the case of a collection of point masses moving in a joint potential is more interesting. When the particle system is coalescing (corresponding to inelastic particle collisions), the system may eventually collapse to single Dirac measures moving along a classical particle trajectory. This situation is described by the above Lagrangian. An example is the chemotactic movement of cells modeled by a Keller-Segel system, which may exhibit finite-time blow-up. After blow-up, collapsed parts seems to consist of evolving Dirac measures.

2.3 Example 2: The electromagnetic Schrödinger equation

We consider the motion of a charged quantum particle in an electromagnetic field with magnetic vector potential A. According to Section 2.1.2, the Lagrangian reads as

$$\mathscr{L}_{\mathrm{M}}(\mu, \eta) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 - \frac{1}{2}|A|^2 - \Phi(x) - \frac{\hbar^2}{8}|\nabla \log \mu|^2 \right) \mu(dx), \tag{2.9}$$

where $\mu \in \mathscr{P}$, $\eta \in T_\mu\mathscr{P}$, and $S = -\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))$. The corresponding action functional becomes

$$\mathscr{A}_{\mathrm{M}}(\gamma) = \int_0^T \mathscr{L}_{\mathrm{M}}(\gamma_t, \dot{\gamma}_t) dt, \tag{2.10}$$

where $\gamma : [0, T] \to \mathscr{P}$ is a smooth curve.

2.3.1 Electromagnetic Madelung equations

We show that the critical points for $\mathscr{A}_{\mathrm{M}}$ solve Madelung-type and quantum hydrodynamic equations.

Theorem 2.3.1 (Electromagnetic Madelung equations)**.** *A smooth curve* $\mu : [0,T] \to \mathscr{P}$ *is a critical point for* $\mathscr{A}_{\mathrm{M}}$*, i.e.*

$$\frac{d}{dt}\frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial \eta} - \frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial \mu} = 0, \tag{2.11}$$

if and only if the flow of the generalized momenta $S_t : \mathbb{R}^d \to \mathbb{R}$*,* $t \in [0,T]$*, of*

$$\partial_t \mu + \operatorname{div}\big(\mu(\nabla S - A)\big) = 0 \quad \text{in } \mathbb{R}^d \tag{2.12}$$

solves the Hamilton-Jacobi equation

$$\partial_t S + \frac{1}{2}|\nabla S - A|^2 + \Phi(x) - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} = 0 \quad \text{in } \mathbb{R}^d. \tag{2.13}$$

For the proof of the above theorem, we need an auxiliary result. Let denote

$$\mathscr{M} = \Big\{\xi \text{ smooth signed measure on } \mathbb{R}^d : \langle \xi, 1\rangle = 0,$$
$$\int_{\mathbb{R}^d} e^{\alpha|x|}|\xi|(dx) < \infty \text{ for all } \alpha > 0\Big\}$$

the set of smooth signed measures with zero mean and finite exponential absolute moments. Here, $\langle\cdot,\cdot\rangle$ denotes the dual product between the space of finitely additive measures on $\mathbb{R}^d$ and the space $L^\infty(\mathbb{R}^d)$. Then, for $\mu \in \mathscr{P}$ and $S \in \mathscr{S}(\mathbb{R}^d)$, the differential operator $\Delta_\mu(S) = \operatorname{div}(\mu\nabla S)$ is well defined. Furthermore, we write

$$\delta_* F(\mu,\xi) = \frac{d}{d\varepsilon}F(\mu+\varepsilon\xi)\Big|_{\varepsilon=0}, \quad \mu \in \mathscr{P},\ \xi \in \mathscr{M},$$

for the first variation of F at μ in the direction of ξ. If $\delta_* F(\mu,\xi) = \int_{\mathbb{R}^d} G\xi dx$, we set $G = \partial F/\partial\mu$, the variational derivative of F with respect to μ.

Lemma 2.3.2. *For smooth measures* $\mu \in \mathscr{P}$*, the operator-valued functions* $\mu \mapsto \Delta_\mu$ *and* $\mu \mapsto \Delta_\mu^{-1}$ *are differentiable in the direction of* $\xi \in \mathscr{M}$*, and their first variations are given by*

$$\delta_*\Delta_{(\mu,\xi)} = \Delta_\xi, \quad \delta_*\Delta^{-1}_{(\mu,\xi)} = -\Delta_\mu^{-1}\Delta_\xi\Delta_\mu^{-1}.$$

Proof. The first claim follows from

$$\delta_*\Delta_{(\mu,\xi)}(S) = \frac{d}{d\varepsilon}\operatorname{div}\big((\mu+\varepsilon\xi)\nabla S\big)\Big|_{\varepsilon=0} = \operatorname{div}(\xi\nabla S) = \Delta_\xi S.$$

To prove the second claim, we notice that $\Delta_{(\mu,\xi)}\Delta^{-1}_{(\mu,\xi)}(S) = S$ implies, by the Leibniz rule, that

$$0 = \delta_*\big(\Delta_{(\mu,\xi)}\Delta^{-1}_{(\mu,\xi)}\big)(S) = \delta_*\Delta_{(\mu,\xi)}(\Delta_\mu^{-1}S) + \Delta_\mu\delta_*\Delta^{-1}_{(\mu,\xi)}(S).$$

By the first claim, this can be written as

$$0 = \Delta_\xi\Delta_\mu^{-1}S + \Delta_\mu\delta_*\Delta^{-1}_{(\mu,\xi)}(S),$$

and multiplication by Δ_μ^{-1} from the left shows the result. □

Proof of Theorem 2.3.1. The theorem is proved by calculating the derivatives in the Euler-Lagrange equation (2.11). To this aim, we set $\mathscr{L} = \mathscr{T} - \mathscr{V}$, where

$$\mathscr{T}(\mu,\eta) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla S|^2 \mu(dx) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla \Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))|^2 \mu(dx),$$
$$\mathscr{V}(\mu) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|A|^2 + \Phi(x) + \frac{\hbar^2}{8}|\nabla \log \mu|^2\right)\mu(dx)$$

are the "kinetic energy" and "potential energy" terms. First, we find that, for fixed $\mu \in \mathscr{P}$ and for any $\xi \in \mathscr{M}$,

$$\begin{aligned}\delta_* \mathscr{T}(\eta,\xi) &= \frac{1}{2}\frac{d}{d\varepsilon}\int_{\mathbb{R}^d} |\nabla \Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A) + \varepsilon\xi)|^2 \mu(dx)\Big|_{\varepsilon=0} \\ &= \int_{\mathbb{R}^d} \nabla \Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A)) \cdot \nabla \Delta_\mu^{-1}\xi \mu(dx) \\ &= -\int_{\mathbb{R}^d} \nabla S \cdot \nabla \Delta_\mu^{-1}\xi\mu(dx) = -\int_{\mathbb{R}^d} \mu \nabla S \cdot \nabla \Delta_\mu^{-1}\xi dx.\end{aligned}$$

Then, by integrating by parts and using the definition of Δ_μ,

$$\delta_* \mathscr{T}(\eta,\xi) = \int_{\mathbb{R}^d} \Delta_\mu^{-1} \operatorname{div}(\mu\nabla S)\xi dx = \int_{\mathbb{R}^d} \Delta_\mu^{-1}(\Delta_\mu S)\xi dx = \int_{\mathbb{R}^d} S\xi dx,$$

showing that $\partial\mathscr{T}/\partial\eta = S$. The expression $\mathscr{V}$ does not depend on η, and hence, $\partial\mathscr{V}/\partial\eta = 0$. Thus,

$$\frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial \eta} = S. \tag{2.14}$$

Next, we compute $\partial\mathscr{T}/\partial\mu$. To illuminate the dependency of $\mathscr{T}$ on μ and especially on Δ_μ^{-1} we observe that $\mathscr{T}$ can be reformulated as

$$\begin{aligned}\mathscr{T}(\mu,\eta) &= \frac{1}{2}\int_{\mathbb{R}^d} \mu\nabla S \cdot \nabla S dx = -\frac{1}{2}\int_{\mathbb{R}^d} \operatorname{div}(\mu\nabla S) S dx \\ &= -\frac{1}{2}\int_{\mathbb{R}^d} S\Delta_\mu S dx = -\frac{1}{2}\int_{\mathbb{R}^d} (\eta - \operatorname{div}(\mu A))\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))dx,\end{aligned}$$

using $S = -\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))$. Hence, the first variation reads as

$$\delta_* \mathscr{T}(\mu,\xi) = -\frac{1}{2}\frac{d}{d\varepsilon}\int_{\mathbb{R}^d} \big(\eta - \operatorname{div}((\mu+\varepsilon\xi)A)\big)\Delta_{\mu+\varepsilon\xi}^{-1}\big(\eta - \operatorname{div}((\mu+\varepsilon\xi)A)\big)dx\Big|_{\varepsilon=0}.$$

We employ the product rule and Lemma 2.3.2 to compute Δ_μ^{-1}:

$$\begin{aligned}\delta_* \mathscr{T}(\mu,\xi) = &\int_{\mathbb{R}^d} \operatorname{div}(\xi A)\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))dx \\ &+ \frac{1}{2}\int_{\mathbb{R}^d} (\eta - \operatorname{div}(\mu A))\Delta_\mu^{-1}\Delta_\xi\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))dx.\end{aligned}$$

The first term becomes, after an integration by parts,

$$\int_{\mathbb{R}^d} \operatorname{div}(\xi A)\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))dx = -\int_{\mathbb{R}^d} \operatorname{div}(\xi A) S dx = \int_{\mathbb{R}^d} (A\cdot \nabla S)\xi dx.$$

For the second term, we find that, by the definition of Δ_ξ,

$$\begin{aligned}
&\frac{1}{2}\int_{\mathbb{R}^d} (\eta - \operatorname{div}(\mu A))\Delta_\mu^{-1}\Delta_\xi\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))dx \\
&\qquad = \frac{1}{2}\int_{\mathbb{R}^d} \Delta_\mu^{-1}(\eta - \operatorname{div}\mu A)\operatorname{div}\left(\xi\nabla\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))\right)dx \\
&\qquad = -\frac{1}{2}\int_{\mathbb{R}^d} \nabla\Delta_\mu^{-1}(\eta - \operatorname{div}\mu A)\cdot\left(\xi\nabla\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A))\right)dx \\
&\qquad = -\frac{1}{2}\int_{\mathbb{R}^d} |\nabla S|^2\xi dx.
\end{aligned}$$

We conclude that

$$\delta_*\mathscr{T}(\mu,\xi) = \int_{\mathbb{R}^d}\left((A\cdot\nabla S) - \frac{1}{2}|\nabla S|^2\right)\xi dx$$

and therefore, the variational derivative equals

$$\frac{\partial\mathscr{T}}{\partial\mu} = A\cdot\nabla S - \frac{1}{2}|\nabla S|^2. \tag{2.15}$$

It remains to calculate $\partial\mathscr{V}/\partial\mu$. The first two terms in the integral of $\mathscr{V}$ depend only linearly on μ which shows that

$$\frac{\partial}{\partial\mu}\int_{\mathbb{R}^d}\left(\frac{1}{2}|A|^2 + \Phi(x)\right)\mu dx = \frac{1}{2}|A|^2 + \Phi(x).$$

The first variation of the Fisher information becomes

$$\begin{aligned}
&\delta_*\left(\int_{\mathbb{R}^d}|\nabla\log\mu|^2\mu dx\right)(\mu,\xi) = \frac{d}{d\varepsilon}\int_{\mathbb{R}^d}|\nabla\log(\mu+\varepsilon\xi)|^2(\mu+\varepsilon\xi)dx\Big|_{\varepsilon=0} \\
&\quad = \frac{d}{d\varepsilon}\int_{\mathbb{R}^d}|\nabla\log\mu|^2(\mu+\varepsilon\xi)dx\Big|_{\varepsilon=0} + \frac{d}{d\varepsilon}\int_{\mathbb{R}^d}|\nabla\log(\mu+\varepsilon\xi)|^2\mu dx\Big|_{\varepsilon=0} \\
&\quad = \int_{\mathbb{R}^d}|\nabla\log\mu|^2\xi dx + 2\frac{d}{d\varepsilon}\int_{\mathbb{R}^d}\nabla\log(\mu+\varepsilon\xi)\cdot\nabla(\log\mu)\mu dx\Big|_{\varepsilon=0} \\
&\quad = \int_{\mathbb{R}^d}\frac{|\nabla\mu|^2}{\mu^2}\xi dx + 2\int_{\mathbb{R}^d}\nabla\frac{\xi}{\mu}\cdot\nabla\mu dx \\
&\quad = \int_{\mathbb{R}^d}\left(\frac{|\nabla\mu|^2}{\mu} - 2\Delta\mu\right)\frac{\xi}{\mu}dx = -4\int_{\mathbb{R}^d}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\xi dx.
\end{aligned}$$

We infer that

$$\frac{\partial}{\partial \mu}\frac{\hbar^2}{8}\int_{\mathbb{R}^d} |\nabla \log \mu|^2 \mu dx = -\frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}.$$

Summarizing, we conclude that

$$\frac{\partial \mathscr{V}}{\partial \mu} = \frac{1}{2}|A|^2 + \Phi(x) - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} \tag{2.16}$$

and for the Lagrangian

$$\begin{aligned}\frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial \mu} &= A\cdot\nabla S - \frac{1}{2}|\nabla S|^2 - \frac{1}{2}|A|^2 - \Phi(x) + \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} \\ &= -\frac{1}{2}|\nabla S - A|^2 - \Phi(x) + \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}},\end{aligned}$$

which finishes the proof. □

We call (2.12)-(2.13) the electromagnetic Madelung equations. The expression $(\hbar^2/2)\ \times\Delta\sqrt{\mu}/\sqrt{\mu}$ is referred to as the Bohm potential. It is the quantum correction to the (electromagnetic) hydrodynamic equations. Via the Madelung transformation $\Psi = \sqrt{\mu}\exp(iS/\hbar)$, smooth solutions (μ, S) to (2.12)-(2.13) with initial data $\mu(\cdot,0) = \mu_0$, $S(\cdot,0) = S_0$ in $\mathbb{R}^d$ yield solutions to the electromagnetic Schrödinger equation

$$i\hbar\partial_t\Psi = \frac{1}{2}\left(\frac{\hbar}{i}\nabla - A\right)^2\Psi + \Phi(x)\Psi,\ t>0, \quad \Psi(\cdot,0) = \sqrt{\mu_0}\exp(iS_0/\hbar) \text{ in } \mathbb{R}^d. \tag{2.17}$$

Remark 1. Taking the gradient of (2.13), multiplying the resulting equation by μ and employing (2.12) similarly as in the proof of Theorem 14.1 in [35], we find the quantum hydrodynamic equations

$$\begin{aligned}&\partial_t\mu + \mathrm{div}(\mu v) = 0,\\ &\partial_t(\mu v) + \mathrm{div}(\mu v\otimes v) - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) + \mu\nabla\Phi(x) = 0, \quad t>0,\\ &\mu(\cdot,0) = \mu_0, \quad (\mu v)(\cdot,0) = \mu_0(\nabla S_0 - A) \quad \text{in } \mathbb{R}^d,\end{aligned}$$

where $v = \nabla S - A$ and $v\otimes v$ denotes the matrix with components $v_j v_k$. Here, we have used the fact that A does not depend on time. Thus the dynamics of a charged particle in an electromagnetic field is formally the same as that of a charged particle in an electric field, with different initial conditions and a different velocity function v. □

Remark 2. Including the internal energy (2.6) into the Lagrangian (2.3), without electromagnetic field,

$$\mathscr{L}(\mu,\eta) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 - U(\mu) - \Phi(x) - \frac{\hbar^2}{8}|\nabla \log \mu|^2 \right) \mu(dx), \quad S = \nabla \Delta_\mu^{-1}\eta,$$

we can derive the nonlinear Schrödinger equation. Indeed, curves of the corresponding action functional are critical if and only if (μ, S) solves

$$\begin{aligned} &\partial_t \mu + \operatorname{div}(\mu \nabla S) = 0, \\ &\partial_t S + \frac{1}{2}|\nabla S|^2 + \Phi(x) + U'(\mu) - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} = 0. \end{aligned}$$

Taking the gradient, multiplying by μ, and setting $\Psi = \sqrt{\mu}\exp(iS/\hbar)$, we arrive at the nonlinear Schrödinger equation

$$i\hbar\partial_t \Psi = -\frac{\hbar^2}{2}\Delta\Psi + f(|\Psi|^2)\Psi + \Phi(x)\Psi,$$

where f is defined by $f(s) = s^{-1/2}U'(s)$ $(s > 0)$. □

2.3.2 Almost symplectic equivalence of measure and wave function dynamics

We have mentioned in Section 2.3.1 that solutions (μ, S) to (2.12)-(2.13) yield solutions to the electromagnetic Schrödinger equation (2.17) via the Madelung transform $(\mu, S) \mapsto \Psi = \sqrt{\mu}\exp(iS/\hbar)$. Similarly to the treatment of the standard Schrödinger case in [55], we shall now give a systematic analysis of this transformation as a symplectic map between two Hamiltonian systems, which turn out to be almost equivalent, as specified in Theorem 2.3.6 below.

Hamiltonian formulation of the electromagnetic Madelung flow.

The first step is to identify the Hamiltonian description of the Lagrangian flow (2.12) – (2.13) by means of the Legendre transform on $T\mathscr{P}$ induced by the lifted Lagrangian (2.9). Since in the current situation, $\mathscr{L}_{\mathrm{M}}$ is no longer quadratic in $\eta \in T_\mu\mathscr{P}$, its induced Legendre transform is not a simple Riesz isomorphism on the Hilbert space $(T_\mu\mathscr{P}, \|\cdot\|_{T_\mu\mathscr{P}})$. As a consequence, the distinct roles played by tangent space $T\mathscr{P}$ of generalized coordinates and its dual space $T^*\mathscr{P}$ of generalized momenta become apparent.

We recall that the cotangent bundle $T^*\mathscr{P}$ consists of all pairs (μ, F), where $\mu \in \mathscr{P}$ and $F : T_\mu\mathscr{P} \to \mathbb{R}$ is linear. From the definition of the tangent space $T_\mu\mathscr{P}$ follows that any distribution η in $T\mathscr{P}$ annihilates the constant functions. Therefore, in our situation, $T^*\mathscr{P}$ can be defined by

$$T^*\mathscr{P} = \big\{(\mu, f) : \mu \in \mathscr{P},\ f \in \mathscr{S}_0(\mathbb{R}^d)\big\},$$

where

$$\mathscr{S}_0 = \{f = \phi + c : \phi \in \mathscr{S}, c \in \mathbb{R}\}/\sim$$

is the space of equivalence classes of shifted Schwartz functions, with $f \sim g$ if and only if $f - g = \text{const}$. Note that the density of the measure μ is supposed to be positive.

In analogy to the classical approach, one defines the Hamiltonian $\mathscr{H}_{\mathrm{M}} : T^*\mathscr{P} \to \mathbb{R}$ associated to the Lagrangian $\mathscr{L}_{\mathrm{M}} : T\mathscr{P} \to \mathbb{R}$ as its Legendre transform, i.e.

$$\mathscr{H}_{\mathrm{M}}(\mu, f) = \sup_{\eta \in T_\mu \mathscr{P}} \big(\langle \eta, f\rangle - \mathscr{L}_{\mathrm{M}}(\mu, \eta)\big),$$

where $(\mu, f) \in \mathscr{P} \times \mathscr{S}_0(\mathbb{R}^d)$ and $\langle \cdot, \cdot \rangle$ denotes the dual bracket in $\mathscr{S}'(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$. Thanks to the strict convexity of $\mathscr{L}_{\mathrm{M}}$, the supremum is attained at $\eta^* \in T_\mu\mathscr{P}$ which is the unique solution to $f = (\partial \mathscr{L}_{\mathrm{M}}/\partial\eta)(\mu, \eta^*)$, and hence,

$$\mathscr{H}_{\mathrm{M}}(\mu, f) = \langle \eta^*, f\rangle - \mathscr{L}_{\mathrm{M}}(\mu, \eta^*).$$

Now, the variational derivative $\partial\mathscr{L}_{\mathrm{M}}/\partial\eta$ has been computed in Section 2.3.1, see formula (2.14). Therefore, $f = (\partial\mathscr{L}_{\mathrm{M}}/\partial\eta)(\mu, \eta^*) = S^*$, where $S^* = -\Delta_\mu^{-1}(\eta^* - \operatorname{div}(\mu A))$, and S^* is unique as a solution in $\mathscr{S}_0(\mathbb{R}^d)$. As a result, we have identified the change of coordinates

$$T\mathscr{P} \to T^*\mathscr{P}, \quad (\mu, \eta) \mapsto (\mu, S), \quad S = -\Delta_\mu^{-1}(\eta - \operatorname{div}(\mu A)),$$

as the Legendre transform from the physical phase space of variations $T\mathscr{P}$ to the space of generalized momenta $T^*\mathscr{P}$.

Inserting the identification $\eta^* = -\Delta_\mu S^* + \operatorname{div}(\mu A)$ into the Hamiltonian gives an explicit expression for $\mathscr{H}_{\mathrm{M}}$:

$$\begin{aligned}\mathscr{H}_{\mathrm{M}}(\mu, S^*) &= \langle -\Delta_\mu S^* + \operatorname{div}(\mu A), S^*\rangle - \mathscr{L}_{\mathrm{M}}(\eta^*, \mu) \\ &= -\int_{\mathbb{R}^d} \operatorname{div}\big(\mu(\nabla S^* - A)\big) S^* dx - \mathscr{L}_{\mathrm{M}}(\eta^*, \mu).\end{aligned}$$

Integrating by parts in the first integral and using the definition of $\mathscr{L}_{\mathrm{M}}$ gives

$$\begin{aligned}\mathscr{H}_{\mathrm{M}}(\mu, S^*) &= \int_{\mathbb{R}^d} \mu|\nabla S^*|^2 dx - \int_{\mathbb{R}^d} \mu A \cdot \nabla S^* dx \\ &\quad - \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S^*|^2 - \frac{1}{2}|A|^2 - \Phi(x) - \frac{\hbar^2}{8}|\nabla \log\mu|^2\right)\mu dx \\ &= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla S^* - A|^2 \mu dx + \int_{\mathbb{R}^d} \Phi(x)\mu dx + \frac{\hbar^2}{8}\int_{\mathbb{R}^d} |\nabla \log\mu|^2 \mu dx.\end{aligned}$$

We see that the Hamiltonian is, as expected, the sum of the magnetic, potential, and quantum energies, respectively. Indeed, the classical electromagnetic Hamiltonian is $H_{\mathrm{M}} = \frac{1}{2}|p - A|^2 + \Phi(x)$, where p is the momentum. In the lifted version, the momentum becomes ∇S, and therefore, $\mathscr{H}_{\mathrm{M,mag}} = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla S - A|^2 \mu(dx)$, which is the above expression.

As a second ingredient for a Hamiltonian description of the associated flow of generalized momenta on $T^*\mathscr{P}$, we introduce a symplectic form on $T^*\mathscr{P}$, similarly as in [55] on the physical phase space $T\mathscr{P}$. We recall that a symplectic form ω on a vector space is a skew-symmetric, non-degenerate, bilinear form, i.e. $\omega(v,w) = -\omega(w,v)$ for all u, v and $\omega(v,w) = 0$ for all w implies that $v = 0$.

Lemma 2.3.3 (Symplectic form on $T^*\mathscr{P}$). *Each pair $(\phi,\psi) \in \mathscr{S}_0(\mathbb{R}^d) \times \mathscr{S}_0(\mathbb{R}^d)$ induces a vector field $V_{\phi,\psi} : T^*\mathscr{P} \to TT^*\mathscr{P}$ via*

$$V_{\phi,\psi}(\mu, f) = (-\operatorname{div}(\mu\nabla\psi), \phi) \in T_{(\mu,f)}T^*\mathscr{P}, \quad (\mu, f) \in T^*\mathscr{P}.$$

Furthermore, $T^*\mathscr{P}$ is endowed with a unique symplectic form ω, defined on the above vector fields by

$$\omega(V_{\phi_1,\psi_1}, V_{\phi_2,\psi_2}) = \int_{\mathbb{R}^d} (\nabla\phi_1 \cdot \nabla\psi_2 - \nabla\phi_2 \cdot \nabla\psi_1)\mu(dx), \tag{2.18}$$

where $(\phi_j, \psi_j) \in \mathscr{S}_0(\mathbb{R}^d) \times \mathscr{S}_0(\mathbb{R}^d)$, $j = 1, 2$.

Proof. Expression (2.18) clearly defines a skew-symmetric bilinear form. Furthermore, an elementary calculation shows that ω is non-degenerate. Uniqueness follows from the fact that for given $(\mu, f) \in T^*\mathscr{P}$, the set of tangent vectors $\{V_{\phi,\psi}(\mu, f) : \phi, \psi \in \mathscr{D}_0(\mathbb{R}^d)\}$ is total in $T_{(\mu,f)}T^*\mathscr{P}$. □

Recall that a Hamiltonian flow on a manifold M with symplectic form ω is induced by an energy function $\varphi : M \to \mathbb{R}$ via the integral curves of the corresponding Hamiltonian vector field X_φ on M. The latter is uniquely defined by the requirement that in any $p \in M$, it holds that

$$\omega(X_\varphi, Z) = d\varphi(Z) \quad \text{for all } Z \in T_pM.$$

The form (2.18) for $M = T^*\mathscr{P}$ allows us to study Hamiltonian flows for various energy functions φ on $T^*\mathscr{P}$. For $\varphi = \mathscr{H}_{\mathrm{M}}$, we arrive at the following statement, which is the analogue of Proposition 3.4 in [55] (also see Corollary 3.5 in that paper).

Theorem 2.3.4 (Critical points and Hamiltonian flow). *A smooth curve of measures $\gamma : [0,T] \to \mathscr{P}$, $t \mapsto \gamma_t$, is a critical point of the action functional $\mathscr{A}_{\mathrm{M}}$, defined in (2.10), if and only if the corresponding curve $(\gamma_t, S_t) \in T^*\mathscr{P}$ in the space of generalized momenta, where $S_t = \Delta_{\gamma_t}^{-1}(\dot\gamma_t - \operatorname{div}(\gamma_t A))$, is a Hamiltonian flow on $(T^*\mathscr{P}, \omega)$ associated to the Hamiltonian $\mathscr{H}_{\mathrm{M}}$.*

Proof. It suffices to compute the corresponding Hamiltonian vector field $X_{\mathscr{H}_{\mathrm{M}}}$ on $M := T^*\mathscr{P}$. To this aim, fix $p = (\mu, f) \in T^*M$ and choose $Z = V_{\phi,\psi}(\mu, f) \in TT^*\mathscr{P}$ as in Definition 2.3.3. Then

$$
\begin{aligned}
d\mathscr{H}_{\mathrm{M}}(V_{\phi,\psi}(\mu,f)) &= \frac{d}{d\varepsilon}\mathscr{H}_{\mathrm{M}}((\mu - \varepsilon \operatorname{div}(\mu\nabla\psi), f + \varepsilon\phi))\Big|_{\varepsilon=0} \\
&= \int_{\mathbb{R}^d} ((\nabla f - A)\mu)\cdot\nabla\phi dx - \frac{1}{2}\int_{\mathbb{R}^d} |\nabla f - A|^2 \operatorname{div}(\mu\nabla\psi)dx \\
&\quad - \int_{\mathbb{R}^d} \Phi \operatorname{div}(\nabla\psi\mu)dx + \frac{\hbar^2}{2}\int_{\mathbb{R}^d}\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right)\operatorname{div}(\nabla\psi\mu)dx. \\
&= \int_{\mathbb{R}^d} (\nabla f - A)\cdot\nabla\phi\mu dx + \frac{1}{2}\int_{\mathbb{R}^d} \nabla|\nabla f - A|^2\cdot\nabla\psi\mu dx \\
&\quad + \int_{\mathbb{R}^d} \nabla\Phi\cdot\nabla\psi\mu dx - \frac{\hbar^2}{2}\int_{\mathbb{R}^d}\nabla\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\cdot\nabla\psi\mu dx.
\end{aligned}
$$

To reveal the structure of the Hamiltonian flow we reformulate the first integral as follows

$$
\begin{aligned}
&\int_{\mathbb{R}^d}(\nabla f - A)\cdot\nabla\phi\mu dx = -\int_{\mathbb{R}^d}\operatorname{div}((\nabla f - A)\mu)\phi dx \\
&= -\int_{\mathbb{R}^d}\Delta_\mu\Delta_\mu^{-1}(\operatorname{div}((\nabla f - A)\mu))\phi dx = \int_{\mathbb{R}^d}\nabla(\Delta_\mu^{-1}\operatorname{div}((\nabla f - A)\mu))\nabla\phi\mu dx.
\end{aligned}
$$

This yields

$$
\begin{aligned}
d\mathscr{H}_{\mathrm{M}}(V_{\phi,\psi}(\mu,f)) &= \int_{\mathbb{R}^d}\nabla(\Delta_\mu^{-1}\operatorname{div}((\nabla f - A)\mu))\nabla\phi\mu dx \\
&\quad + \frac{1}{2}\int_{\mathbb{R}^d}\nabla|\nabla f - A|^2\cdot\nabla\psi\mu dx \\
&\quad + \int_{\mathbb{R}^d}\nabla\Phi\cdot\nabla\psi\mu dx - \frac{\hbar^2}{2}\int_{\mathbb{R}^d}\nabla\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\cdot\nabla\psi\mu dx.
\end{aligned}
$$

Comparing with (2.18), we find that

$$
X_{\mathscr{H}_{\mathrm{M}}}(\mu, f) = \left(-\Delta_\mu^{-1}\operatorname{div}((\nabla f - A)\mu), \frac{1}{2}|\nabla f - A|^2 + \Phi - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right).
$$

Hence, a smooth curve $t \mapsto (\mu_t, S_t) \in T^*\mathscr{P}$ is an integral curve for $X_{\mathscr{H}_{\mathrm{M}}}$ if and only if the corresponding flow of variations $t \mapsto \dot{\mu}_t \in T\mathscr{P}$ solves (2.12)-(2.13). □

Hamiltonian structure of the electromagnetic Schrödinger flow.

Let us recall the basic fact that the electromagnetic Schrödinger equation has a Hamiltonian structure, too. Indeed, denoting by $\mathscr{C} = C^\infty(\mathbb{R}^d;\mathbb{C})$ the linear space of smooth complex-valued functions on $\mathbb{R}^d$ and identifying as usual the tangent space over an element $\Psi \in \mathscr{C}$ with the space $\mathscr{C}$, the tangent bundle $T\mathscr{C}$ is naturally equipped with the symplectic form

$$\omega_{\mathrm{C}}(F, G) = -2 \int_{\mathbb{R}^d} \Im(F \cdot \overline{G})(x) dx,$$

where $\Im(z)$ is the imaginary part of $z \in \mathbb{C}$ and $\overline{z}$ is its complex conjugate. This way $(\mathscr{C}, \hbar\omega_{\mathrm{C}})$ becomes a symplectic space. On $\mathscr{C}$ we define the energy function $\mathscr{H}_{\mathrm{C}} : \mathscr{C} \mapsto \mathbb{R}$ by

$$\mathscr{H}_{\mathrm{C}}(\Psi) = \frac{1}{2} \int_{\mathbb{R}^d} \left| \left(\frac{\hbar}{i} \nabla - A \right) \Psi \right|^2 dx + \int_{\mathbb{R}^d} \Phi(x) \, |\Psi|^2(x) \, dx,$$

which is the electromagnetic Schrödinger Hamiltonian.

Proposition 1. *A smooth flow of wave functions $t \mapsto \Psi_t \in \mathscr{C}$ solves the electromagnetic Schrödinger equation* (2.17) *if and only if it is a Hamiltonian flow induced from the energy function $\mathscr{H}_{\mathrm{C}}$ on the symplectic space $(\mathscr{C}, \hbar\omega_{\mathrm{C}})$.*

Proof. We only sketch the proof of this classical but mostly forgotten fact. For Ψ, $\zeta \in \mathscr{C}$, we find by a straightforward computation that

$$\begin{aligned}
\frac{d}{d\epsilon} \mathscr{H}_{\mathrm{C}}(\Psi + \epsilon\zeta)\Big|_{\varepsilon=0} &= \Re \int_{\mathbb{R}^d} \left(\left(\frac{\hbar}{i} \nabla - A \right)^2 + 2\Phi \right) \Psi \cdot \overline{\zeta} dx \\
&= \Im \int_{\mathbb{R}^d} i \left(\left(\frac{\hbar}{i} \nabla - A \right)^2 + 2\Phi \right) \Psi \cdot \overline{\zeta} dx \\
&= \omega_{\mathrm{C}} \left(-\frac{i}{\hbar} \left(\frac{1}{2} \left(\frac{\hbar}{i} \nabla - A \right)^2 + \Phi \right) \Psi, \zeta \right).
\end{aligned}$$

This shows that the Hamiltonian vector field $X_{\mathscr{H}_{\mathrm{C}}}$ associated to $\mathscr{H}_{\mathrm{C}}$ on $(\mathscr{C}, \omega_{\mathrm{C}})$ is

$$X_{\mathscr{H}_{\mathrm{C}}}(\Psi) = -\frac{i}{\hbar} \left(\frac{1}{2} \left(\frac{\hbar}{i} \nabla - A \right)^2 + \Phi \right) \Psi.$$

Hence, solutions to the electromagnetic Schrödinger equation (2.17) are precisely the integral curves of the Hamiltonian vector field $X_{\mathscr{H}_{\mathrm{C}}}$. □

Madelung transform: Precise definition and symplectic properties.

Let $\mathscr{C}_* = \{\Psi \in \mathscr{C} : \int_{\mathbb{R}^d} |\Psi|^2 dx = 1, \ \Psi(x) \neq 0 \text{ for all } x \in \mathbb{R}^d\}$ be the set of smooth nowhere vanishing normalized wave functions. Each $\Psi \in \mathscr{C}_*$ admits a decomposition $\Psi = |\Psi| \exp(iS/\hbar)$, where the smooth function $S : \mathbb{R}^d \to \mathbb{R}$

is uniquely defined up to an additive constant of the form $2\pi\hbar k$, $k \in \mathbb{N}$. In particular, the Madelung transform is well defined

$$\sigma : \mathscr{C}_* \to T^*\mathscr{P}, \quad \sigma(\Psi) = (|\Psi(x)|^2 dx, S) \in \mathscr{P} \times \mathscr{S}_0(\mathbb{R}^d). \tag{2.19}$$

Recall that by the definition of $\mathscr{S}_0(\mathbb{R}^d)$ as the space of equivalence classes of shifted Schwartz functions, the map σ is not injective. However, we may apply the abstract notion of a symplectic submersion (see [55]) which is a generalization of a symplectic isomorphism where the injectivity assumption is dropped.

Definition 2.3.5 (Symplectic submersion on manifolds)**.** Let (M, ω_M), (N, ω_N) be symplectic manifolds equipped with the symplectic forms ω_M, ω_N, respectively, and let $s : M \to N$ be a smooth map. Then s is called a *symplectic submersion* if its differential $s_* : TM \to TN$ is surjective and satisfies $\omega_N(s_*X, s_*Y) = \omega_M(X, Y)$ for all X, $Y \in TM$.

Similarly to the isomorphism case one may easily see that Hamiltonian flows are stable under sympletic submersions. This is stated in the following proposition, cf. [55, Prop. 4.2].

Proposition 2 (Submersions between Hamiltonian flows)**.** *Let M, N be symplectic manifolds equipped with the symplectic forms ω_M, ω_N, respectively, and let $s : M \to N$ be a symplectic submersion. If the Hamiltonians $F \in C^\infty(M)$ and $G \in C^\infty(N)$ are related by $F = G \circ s$, the submersion s maps Hamiltonian flows associated to F on (M, ω_M) to Hamiltonian flows associated to G on (N, ω_N).*

We are now ready to state the main result of this section which asserts that the Madelung transform is a symplectic submersion from $\mathscr{C}_*$ to $T^*\mathscr{P}$.

Theorem 2.3.6 (Madelung transform as symplectic submersion)**.** *The Madelung transform $\sigma : \mathscr{C}_* \to T^*\mathscr{P}$, defined in* (2.19)*, is a symplectic submersion from $(\mathscr{C}_*, \hbar\omega_{\mathrm{C}})$ to $(T^*\mathscr{P}, \omega)$, preserving the electromagnetic Schrödinger Hamiltonian,*

$$\mathscr{H}_{\mathrm{C}} = \mathscr{H}_{\mathrm{M}} \circ \sigma.$$

Proof. Since the proof is very similar to the proof of Theorem 4.3 in [55], we give only a sketch. First, we restrict the phase $S/\hbar$ in $|\Psi| \exp(iS/\hbar)$ to the interval $[0, 2\pi\hbar)$ by defining an appropriate bijection. We can prove that the differential s_* is surjective. A calculation shows that $\omega_{T^*\mathscr{P}}(s_* V_{\phi_1,\psi_1}, s_* V_{\phi_2,\psi_2}) = \hbar\omega_{\mathrm{C}}(V_{\phi_1,\psi_1}, V_{\phi_2,\psi_2})$ for all vector fields V_{ϕ_1,ψ_1}, V_{ϕ_2,ψ_2}. Thus, s is a symplectic submersion. The remaining part $\mathscr{H}_{\mathrm{C}} = \mathscr{H}_{\mathrm{M}} \circ \sigma$ is a computation; see [55, Section 4] for details. □

In light of Proposition 2 and Theorem 2.3.6, the electromagnetic Schrödinger equation (2.17) for wave functions can be interpreted as the lift of the physically intuitive Lagrangian flow on probability measures (or mass distributions) (2.13) to the larger space of complex wave functions. The lifted Hamiltonian system is the familiar electromagnetic Schrödinger equation for wave functions and has the advantage that it is linear. However, a disadvantage is that a new and unphysical degree of freedom, incorporated in the constant phase shift for wave functions and describing the same physical state, is introduced.

2.4 Example 3: Quantum Navier-Stokes equations

In this section, we consider the quantum Lagrangian

$$\mathscr{L}_{\mathrm{Q}}(\mu,\eta) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 - U(\mu) - \Phi(x) - \frac{\hbar^2}{8}|\nabla \log \mu|^2 \right) \mu(dx), \tag{2.20}$$

where $\mu \in \mathscr{P}$, $\eta \in T_\mu\mathscr{P}$, $S = -\Delta_\mu^{-1}\eta$, and $U(\mu)$ denotes the internal energy which is assumed to be a smooth function. Here, we are interested in the Lagrangian flow with dissipation

$$\mathscr{D}(\mu,\eta) = \frac{\alpha}{2}\int_{\mathbb{R}^d} |\nabla v|^2 \mu(dx),$$

where $\alpha \geq 0$, and $v = \nabla S$ is the unique potential velocity field inducing the variation η of the state μ.

2.4.1 Quantum Navier-Stokes equations

We show that the dissipative Lagrangian flow on $\mathscr{P}$ can be related to the Navier-Stokes equations including the Bohm potential and a density-depending viscosity. Our result reads as follows.

Theorem 2.4.1 (Quantum Navier-Stokes equations)**.** *A smooth curve $\mu : [0,T] \to \mathscr{P}$ satisfies*

$$\frac{d}{dt}\frac{\partial \mathscr{L}_{\mathrm{Q}}}{\partial \eta}(\mu,\dot{\mu}) - \frac{\partial \mathscr{L}_{\mathrm{Q}}}{\partial \mu}(\mu,\dot{\mu}) + \frac{\partial \mathscr{D}}{\partial \eta}(\mu,\dot{\mu}) = 0 \tag{2.21}$$

if and only if the mass flux $t \mapsto \mu_t v_t$ with $v = -\nabla\Delta_\mu^{-1}\eta$ solves the quantum Navier-Stokes equation

$$\partial_t(\mu v) + \operatorname{div}(\mu v \otimes v) + \nabla p(\mu) + \mu\nabla\Phi(x) - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) = \alpha\mu\nabla\Delta_\mu^{-1}(\nabla^2 : (\mu\nabla v)). \tag{2.22}$$

Here, $v \otimes v$ is a tensor with components $v_j v_k$; the pressure function $p(\mu)$ is defined through $p'(s) = sU''(s)$ for $s \geq 0$; and the product ":" signifies summation over both indices. Identifying vector fields modulo rotational components, we can write this equation as

$$\partial_t(\mu v) + \operatorname{div}(\mu v \otimes v) + \nabla p(\mu) + \mu\nabla\Phi(x) - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) \equiv \alpha \operatorname{div}(\mu D(v)), \tag{2.23}$$

where $A \equiv B$ *if and only if* $\operatorname{div}(A - B) = 0$*, and* $D(v) = \frac{1}{2}(\nabla v + \nabla v^\top) = \nabla v$ *is the symmetric velocity gradient.*

The system of quantum Navier-Stokes equations is given by (2.22) and the continuity equation

$$\partial_t \mu + \operatorname{div}(\mu v) = 0. \tag{2.24}$$

In this model, the viscous stress tensor is $\mathbb{S} = \nu D(v)$, where the viscosity $\nu = \alpha\mu$ depends on the particle density μ. For variants of the stress tensor, see Remark 3.

Proof. We write $\mathscr{L}_{\mathrm{Q}} = \mathscr{T} - \mathscr{V}$, where

$$\mathscr{T}(\mu, \eta) = \|\eta\|^2_{T_\mu \mathscr{P}} = \int_{\mathbb{R}^d} |\nabla \Delta_\mu^{-1} \eta|^2 \mu(dx)$$

corresponds to the "kinetic energy" and

$$\mathscr{V}(\mu, \eta) = \int_{\mathbb{R}^d} \left(\Phi(x) + U(\mu) + \frac{\hbar^2}{8} |\nabla \log \mu|^2 \right) \mu(dx) \tag{2.25}$$

corresponds to the "potential energy". By the proof of Theorem 2.3.1 (see (2.16) with $A = 0$), we have

$$\frac{\partial \mathscr{V}}{\partial \mu} = \Phi(x) + U'(\mu) - \frac{\hbar^2}{2} \frac{\Delta \sqrt{\mu}}{\sqrt{\mu}}. \tag{2.26}$$

Since $\mathscr{V}$ does not depend on η, it follows that $\partial \mathscr{V}/\partial \eta = 0$. Furthermore, by (2.14) and (2.15) (with $A = 0$),

$$\frac{\mathscr{L}_{\mathrm{Q}}}{\partial \eta} = \frac{\partial \mathscr{T}}{\partial \eta} = S, \quad \frac{\partial \mathscr{T}}{\partial \mu} = -\frac{1}{2} |\nabla S|^2. \tag{2.27}$$

It remains to compute $\partial \mathscr{D}/\partial \eta$. To this end, let $\xi \in \mathscr{M}$ and set $\zeta = \Delta_\mu^{-1} \xi$. Since $v = \nabla S = -\nabla \Delta_\mu^{-1} \eta$, we infer that

$$\begin{aligned}
\delta_* \mathscr{D}(\eta, \xi) &= \frac{\alpha}{2} \frac{d}{d\varepsilon} \int_{\mathbb{R}^d} |\nabla^2 \Delta_\mu^{-1} (\eta + \varepsilon \xi)|^2 \mu(dx) \Big|_{\varepsilon=0} \\
&= \frac{\alpha}{2} \frac{d}{d\varepsilon} \int_{\mathbb{R}^d} |\nabla^2 (\Delta_\mu^{-1} \eta + \varepsilon \Delta_\mu^{-1} \xi)|^2 \mu(dx) \Big|_{\varepsilon=0} \\
&= \frac{\alpha}{2} \frac{d}{d\varepsilon} \int_{\mathbb{R}^d} \nabla^2 (-S + \varepsilon \zeta) : \nabla^2 (-S + \varepsilon \zeta) \mu(dx) \Big|_{\varepsilon=0} \\
&= -\alpha \int_{\mathbb{R}^d} \nabla^2 S : \nabla^2 \zeta \mu dx = -\alpha \int_{\mathbb{R}^d} \Delta_\mu^{-1} (\nabla^2 : (\mu \nabla^2 S)) \xi dx.
\end{aligned}$$

This implies that

$$\frac{\partial \mathscr{D}}{\partial \eta} = -\alpha \Delta_\mu^{-1} (\nabla^2 : (\mu \nabla^2 S)). \tag{2.28}$$

Inserting this expression as well as (2.26) and (2.27) into (2.21) gives

$$\partial_t S + \frac{1}{2}|\nabla S|^2 + \Phi(x) + U'(\mu) - \frac{\hbar^2}{2}\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}} = \alpha\Delta_\mu^{-1}\nabla^2 : (\mu\nabla^2 S).$$

We take the gradient, multiply this equation by μ, and replace $\nabla S = v$:

$$\mu\partial_t v + \frac{1}{2}\mu\nabla|v|^2 + \mu\nabla\Phi(x) + \mu U''(\mu)\nabla\mu - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) = \alpha\mu\nabla\Delta_\mu^{-1}\nabla^2 : (\mu\nabla v).$$

Then, employing the continuity equation $v\partial_t\mu + v\,\mathrm{div}(\mu v) = 0$ and rearranging terms, we obtain

$$\partial_t(\mu v) + \mathrm{div}(\mu v\otimes v) + \mu\nabla\Phi(x) + \nabla p(\mu) - \frac{\hbar^2}{2}\mu\nabla\left(\frac{\Delta\sqrt{\mu}}{\sqrt{\mu}}\right) = \alpha\mu\nabla\Delta_\mu^{-1}\nabla^2 : (\mu D(v)),$$

which equals (2.22). The final step is the projection on the space of curl-free fields by taking the divergence which leads to (2.23). Indeed, observing that

$$\mathrm{div}\left(\mu\nabla\Delta_\mu^{-1}\nabla^2 : (\mu D(v))\right) = \Delta_\mu\Delta_\mu^{-1}(\nabla^2 : (\mu D(v)) = \mathrm{div}\left(\,\mathrm{div}(\mu D(v))\right),$$

we conclude the proof. □

Remark 3. The Lagrangian approach allows us to choose other dissipation terms. We consider two simple examples:

$$\mathscr{D}_1(\mu,\eta) = \frac{\alpha}{p}\int_{\mathbb{R}^d}|\nabla v|^p\mu(dx), \quad p \geq 2,$$
$$\mathscr{D}_2(\mu,\eta) = \frac{1}{2}\int_{\mathbb{R}^d} g(\mu)\big(\nu_1|\nabla v|^2 + \nu_2(\mathrm{div}\,v)^2\mathbb{I}\big)\mu(dx),$$

where $g : \mathbb{R} \to [0,\infty)$ is some function and ν_1, $\nu_2 > 0$. The variational derivatives are computed similarly as in the proof of Theorem 2.4.1. The results are as follows:

$$\frac{\partial\mathscr{D}_1}{\partial\eta} = -\alpha\Delta_\mu^{-1}\nabla^2 : (\mu|D(v)|^{p-2}D(v)),$$
$$\frac{\partial\mathscr{D}_2}{\partial\eta} = -\Delta_\mu^{-1}\nabla^2 : \big(\mu g(\mu)(\nu_1 D(v) + \nu_2(\mathrm{div}\,v)\mathbb{I})\big).$$

The viscous term in the quantum Navier-Stokes equations is obtained after taking the gradient, multiplying by μ, and projecting it on the space of curl-free vectors:

$$\mathrm{div}\left(\mu\nabla\frac{\partial\mathscr{D}_1}{\partial\eta}\right) = -\alpha\,\mathrm{div}\left(\mu\nabla\Delta_\mu^{-1}\nabla^2 : (\mu|D(v)|^{p-2}D(v))\right)$$
$$= -\alpha\,\mathrm{div}\left(\,\mathrm{div}(\mu|D(v)|^{p-2}D(v))\right),$$

and similarly for the second expression. The viscous stress tensors become

$$\mathbb{S}_1 = \alpha\mu|D(v)|^{p-2}D(v), \quad \mathbb{S}_2 = \mu g(\mu)(\nu_1 D(v) + \nu_2(\operatorname{div} v)\mathbb{I}).$$

The viscosity $\nu_1 = \alpha\mu|D(v)|^{p-2}$ depends not only on the particle density but also on the velocity gradient. When we choose $g(\mu) = 1/\mu$, the viscosities are constant, which corresponds to the case of Newtonian fluids (see, e.g., [25, Formula (1.16)]). □

2.4.2 Energy-dissipation identities and Noether currents

According to Section 2.3.2, the Hamiltonian $\mathscr{H}_{\mathrm{Q}} : T^*\mathscr{P} \to \mathbb{R}$ associated to the Lagrangian $\mathscr{L}_{\mathrm{Q}} : T\mathscr{P} \to \mathbb{R}$, defined in (2.20), is given by

$$\mathscr{H}_{\mathrm{Q}}(\mu, S) = \langle \eta, S\rangle - \mathscr{L}_{\mathrm{Q}}(\mu, \eta),$$

where $S = (\partial\mathscr{L}_{\mathrm{Q}}/\partial\eta)(\mu, \eta) = -\Delta_\mu^{-1}\eta$. Inserting $\eta = -\Delta_\mu S$ and the definition (2.20) of $\mathscr{L}_{\mathrm{Q}}$ into this expression, we find that

$$\begin{aligned}\mathscr{H}_{\mathrm{Q}}(\mu, S) &= \int_{\mathbb{R}^d} |\nabla S|^2 \mu dx - \mathscr{L}_{\mathrm{Q}}(\mu, \eta) \\ &= \int_{\mathbb{R}^d} \left(\frac{1}{2}|\nabla S|^2 + U(\mu) + \Phi(x) + \frac{\hbar^2}{8}|\nabla \log \mu|^2\right) \mu dx, \end{aligned} \tag{2.29}$$

which is the sum of the kinetic, internal, potential, and quantum energies. In this section, we derive energy-dissipation identities for smooth solutions to the quantum Navier-Stokes equations (2.22) and (2.24).

Proposition 3 (Energy-dissipation identity)**.** *Let* (μ, v) *be a smooth solution to* (2.22) *and* (2.24)*. Then*

$$\frac{d\mathscr{H}_{\mathrm{Q}}}{dt} + \alpha \int_{\mathbb{R}^d} \mu |\nabla v|^2 dx = 0. \tag{2.30}$$

Proof. Multiplying (2.22) by v and (2.24) by

$$-\frac{1}{2}|v|^2 + U'(\mu) + \Phi(x) + (\hbar^2/2)(\Delta\sqrt{\mu}/\sqrt{\mu})$$

and adding the resulting equations, a straightforward computation yields

$$\begin{aligned}&\frac{d}{dt}\int_{\mathbb{R}^d} \left(\frac{1}{2}|v|^2 + U(\mu) + \mu\Phi(x) + \frac{\hbar^2}{8}\mu|\nabla \log \mu|^2\right) dx \\ &\qquad = \alpha \int_{\mathbb{R}^d} \mu v \cdot \nabla\Delta_\mu^{-1}(\nabla^2 : (\mu v)) dx.\end{aligned}$$

The left-hand side equals $d\mathscr{H}_{\mathrm{Q}}/dt$. The right-hand side can be rewritten, using $v = \nabla S$ and integration by parts, as

$$\begin{aligned}
-\alpha \int_{\mathbb{R}^d} \operatorname{div}(\mu \nabla S) \Delta_\mu^{-1}(\nabla^2 : (\mu \nabla^2 S)) dx &= -\alpha \int_{\mathbb{R}^d} \Delta_\mu S \Delta_\mu^{-1}(\nabla^2 : (\mu \nabla^2 S)) dx \\
&= -\alpha \int_{\mathbb{R}^d} \Delta_\mu^{-1} \Delta_\mu S (\nabla^2 : (\mu \nabla^2 S)) dx \\
&= -\alpha \int_{\mathbb{R}^d} \nabla^2 S : (\mu \nabla^2 S) dx = -\alpha \int_{\mathbb{R}^d} \mu |\nabla v|^2 dx,
\end{aligned}$$

proving the claim. □

Remark 4. Proposition 3 is the counterpart of the energy dissipation law for classical damped Lagrangian systems in $\mathbb{R}^n$ in which case the analogue of (2.21) reads as

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) - \frac{\partial L}{\partial q}(q, \dot{q}) + \frac{\partial D}{\partial \dot{q}}(q, \dot{q}) = 0. \tag{2.31}$$

Writing the dynamics in Hamiltonian coordinates $t \mapsto (q(t), p(t))$ via the Legendre transform, i.e. $p = p(q, \dot{q}) = (\partial L / \partial \dot{q})(q, \dot{q})$, for the Hamiltonian we obtain

$$H(q, p(q, \dot{q})) = \langle \dot{q}, \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \rangle - L(q, \dot{q}),$$

which yields, after differentiation with respect to t and inserting (2.31),

$$\frac{dH}{dt}(q(t), p(t)) = -\langle \dot{q}, \frac{\partial D}{\partial \dot{q}}(q, \dot{q}) \rangle.$$

In our case, by the same computation and using (2.28), it follows that

$$\frac{d\mathscr{H}_{\mathrm{Q}}}{dt} = -\langle \eta, \frac{\partial \mathscr{D}}{\partial \eta} \rangle = -\alpha \langle \Delta_\mu S, \Delta_\mu^{-1}(\nabla^2 : (\mu \nabla^2 S)) \rangle = -\alpha \int_{\mathbb{R}^d} |\nabla^2 S|^2 d\mu,$$

which equals (2.30). □

It has been shown in [36] that the *projected* system (2.23)-(2.24) possesses a second energy functional,

$$\mathscr{H}_{\mathrm{Q}}^*(\mu, S) = \int_{\mathbb{R}^d} \left(\frac{1}{2}|w|^2 + U(\mu) + \Phi(x) + \left(\frac{\hbar^2}{8} - \frac{\alpha^2}{2} \right) |\nabla \log \mu|^2 \right) \mu dx, \tag{2.32}$$

where $w = v + v_{\mathrm{os}}$ and $v_{\mathrm{os}} = \alpha \nabla \log \mu$ is the osmotic velocity first introduced by Nelson [49, Formula (26)]. More precisely, let (μ, v) with $v = \nabla S = -\nabla \Delta_\mu^{-1} \eta$ be a smooth solution to

$$\partial_t \mu + \operatorname{div}(nv) = 0, \quad x \in \mathbb{R}^d, \ t > 0, \tag{2.33}$$

$$\partial_t(\mu v) + \operatorname{div}(\mu v \otimes v) + \nabla p(\mu) + \mu \nabla \Phi(x) - \frac{\hbar^2}{2} \mu \nabla \left(\frac{\Delta \sqrt{\mu}}{\sqrt{\mu}} \right) = \alpha \operatorname{div}(\mu D(v)). \tag{2.34}$$

Then a formal computation [36] shows that

$$\frac{d\mathscr{H}_{\mathrm{Q}}^*}{dt} + \alpha \int_{\mathbb{R}^d} \left(\mu |\nabla w|^2 + U''(\mu)|\nabla \mu|^2 + \left(\frac{\hbar^2}{8} - \frac{\alpha^2}{2} \right) \mu |\nabla^2 \log \mu|^2 \right) dx = 0,$$

which provides additional estimates for the solutions if $\hbar^2/4 > \alpha^2$. We wish to understand why system (2.33)-(2.34) possesses *two* dissipative laws.

A first partial answer was given in [38]. There it was shown that the osmotic velocity emerges from gauge field theory by introducing the local gauge transformation $\psi \mapsto \phi = \exp(-i\alpha \log \mu)\psi$, where ψ is a given quantum state. This transformation leaves the particle density invariant but it changes the mass flux $nv = -\Im(\overline{\psi} \nabla \psi)$ according to

$$nw = -\Im(\overline{\phi} \nabla \phi) = -\Im(\overline{\psi} \nabla \psi - i\alpha\mu \nabla \log \mu) = \mu(v + \alpha \nabla \log \mu).$$

Our goal is to show that the new velocity w can be interpreted as a special transformation of (t, μ) and that the Hamiltonian $\mathscr{H}_{\mathrm{Q}}^*$ can be interpreted as the Noether current associated to this transformation.

To this end, we recall some basic facts from classical Noether theory (see, e.g., [10, Chapter 9]). Let a Lagrangian $L(t, q, \dot{q})$ be given. We introduce the transformations $T(t, q; s)$ and $Q(t, q; s)$, where $s > 0$ is a parameter, such that $t = T(t, q; 0)$ and $q = Q(t, q; 0)$. Setting

$$\delta t = \frac{\partial T}{\partial s}(t, q; 0), \quad \delta q = \frac{\partial Q}{\partial s}(t, q; 0),$$

Taylor's expansion gives $T(t, q) = t + s\delta t + O(s^2)$ and $Q(t, q) = q + s\delta q + O(s^2)$ as $\gamma \to 0$. For infinitesimal small $s > 0$, we can formulate the transformation as $t \mapsto t + \delta t$ and $q \mapsto q + \delta q$. Now, the *Noether current* is defined as

$$J = \delta t \left(\frac{\partial L}{\partial \dot{q}} \dot{q} - L \right) - \delta q \frac{\partial L}{\partial \dot{q}}.$$

If the Lagrangian density $L(t, q, \dot{q})$ is invariant under the above transformation, Noether's theorem states that the Noether current is constant along any extremal of the action integral over L.

On the space of probability measures, we define the lifted Noether current as

$$\mathscr{J}(\mu, \eta) = \delta t \Big\langle \frac{\partial \mathscr{L}}{\partial \eta}(\mu, \eta), \eta \Big\rangle - \delta t \mathscr{L}(\mu, \eta) - \Big\langle \frac{\partial \mathscr{L}}{\partial \eta}(\mu, \eta), \delta \mu \Big\rangle, \quad (\mu, \eta) \in T\mathscr{P},$$

where $\langle \cdot, \cdot \rangle$ denotes the dual product in suitable spaces. We prove the following result.

Theorem 2.4.2 (Noether currents)**.** *Let the Lagrangian* $\mathscr{L}_{\mathrm{Q}}$ *be given by* (2.20)*. Then*

- $\delta t = 1$, $\delta\mu = 0$: $\mathscr{J} = \mathscr{H}_{\mathrm{Q}}$, *defined in* (2.29);
- $\delta t = 1$, $\delta\mu = \alpha\Delta\mu$: $\mathscr{J} = \mathscr{H}_{\mathrm{Q}}^*$, *defined in* (2.32).

Proof. The theorem follows by inserting the transformations into the definition of the Noether current. We recall from (2.27) that $\partial\mathscr{L}_{\mathrm{Q}}/\partial\eta = S$, where $S = -\Delta_\mu^{-1}\eta$. Then, if $\delta t = 1$, $\delta\mu = 0$, we find that

$$\mathscr{J} = \int_{\mathbb{R}^d} S\eta dx - \mathscr{L}_{\mathrm{Q}} = \int_{\mathbb{R}^d} \mu|\nabla S|^2 dx - \mathscr{L}_{\mathrm{Q}} = \mathscr{H}_{\mathrm{Q}}.$$

Next, if $\delta t = 1$, $\delta\mu = \alpha\Delta\mu$, we compute

$$\begin{aligned}
\mathscr{J} &= \int_{\mathbb{R}^d} (S\eta - \alpha\Delta\mu S)dx - \mathscr{L}_{\mathrm{Q}} \\
&= \int_{\mathbb{R}^d} \left(\frac{1}{2}\mu|\nabla S|^2 + U(\mu) + \psi(x) + \frac{\hbar^2}{8}\mu|\nabla\log\mu|^2 + \alpha\nabla\mu\cdot\nabla S\right) dx \\
&= \int_{\mathbb{R}^d} \left(\frac{1}{2}\mu|\nabla(S+\alpha\log\mu)|^2 + \left(\frac{\hbar^2}{8} - \frac{\alpha^2}{2}\right)\mu|\nabla\log\mu|^2 + U(\mu) + \psi(x)\right) dx \\
&= \mathscr{H}_{\mathrm{Q}}^*,
\end{aligned}$$

completing the proof. □

Notice that Noether's theorem, which yields energy conservation, can be applied only if $\alpha = 0$, otherwise we have dissipation of energy. For a classical Noether theory including dissipative terms, we refer to [20, 57] or the more recent works [29, 30]. The extension of this theory to our context is an open question.

Chapter 3

Gradient flows on Hilbert spaces

In this chapter we collect results on so called one-leg schemes for gradient flows and investigate entropy dissipating properties of the discrete solution.

The Chapter is organized as follows. First in section 3.1.1 we give prerequisites on λ-convex functions. Then in 3.1.3 we define the one-leg schemes and show that a unique solution to the scheme exists. In section 3.1.4 we answer the question of the convergence order of the schemes. In 3.2 we investigate the structure preserving properties of the scheme. In 3.3 we investigate the porous medium equation numerically.

Remark 5. As mentioned in the introduction the existence and convergence of one-leg schemes is known (see e.g. [34]), however we give an alternative proof, which in our opinion is more suitable in the case of a general Hilbert space. The existence analysis is based on the proof of [24, Chapter 9.6, Theorem 1], however, we extent the result from convex to λ-convex functionals.

3.1 Existence and convergence

3.1.1 Prerequisites

We collect the basic definitions and useful properties of λ-convex functions in this section. In the following we consider H to be a Hilbert space and $\phi : H \to (-\infty, \infty]$ to be:

- proper, i.e.
  ϕ is not identically equal to ∞
- lower semicontinuous, i.e.
  $x_k \to x$ in $H \Rightarrow \phi(x) \leq \liminf_{k\to\infty} \phi(x_k)$, and
- λ-convex, i.e.
  $\phi(x + t(y - x)) \leq \phi(x) + t(\phi(y) - \phi(x)) - \frac{\lambda}{2} t(1 - t)\|x - y\|^2$
  for $x, y \in H$, $t \in [0, 1]$, and $\lambda \in \mathbb{R}$.

Remark 6. Assume $\phi : \mathbb{R} \to \mathbb{R}$ to be C^2. Then the λ in the definition of the λ-convexity is nothing else but a lower bound for the second derivative, i.e. $\phi''(x) \geq \lambda$. In this sense the inequality above is a second order "Taylor estimate". Note furthermore that $\lambda \in \mathbb{R}$ is arbitrary, i.e. can also be negative.

Above we allowed $\phi(x) = \infty$, which should symbolize states which are not accessible. We define the domain of ϕ as the set of accessible points.

Definition 3.1.1. The domain of $\phi : H \to (-\infty, \infty]$ is given by

$$D(\phi) := \{x \in H : \phi(x) < \infty\}.$$

For the concept of a gradient flow, we need the concept of subdifferential for λ-convex functionals.

Definition 3.1.2. For a λ-convex and proper functional $\phi : H \to (-\infty, \infty]$ we define the subdifferential of ϕ at $x \in H$ by:

$$\partial\phi(x) := \{v \in H : \phi(x) + \langle v, y - x\rangle + \frac{\lambda}{2}\|x - y\|^2 \leq \phi(y); \forall y \in H\}.$$

We define the domain of $\partial\phi$ to be:

$$D(\partial\phi) := \{x \in H : \partial\phi(x) \neq \emptyset\}.$$

Note that if ϕ is C^1 the set $\partial\phi(x)$ is single valued and coincides with the usual gradient. Before we come to the investigation of the gradient flow we show the following useful lemma.

Lemma 3.1.3. *Let $\phi : H \to (-\infty, \infty]$ be proper, lower semicontinuous and λ-convex. Then*

- *$D(\partial\phi) \subset D(\phi)$.*
- *Let $x, y \in D(\partial\phi)$. If $v \in \partial\phi(x)$, and $w \in \partial\phi(y)$, then*

$$\langle v - w, x - y\rangle \geq \lambda\|x - y\|^2.$$

- *For $\lambda \geq 0$*

$$\phi(x) = \min_{y\in H} \phi(y) \Leftrightarrow 0 \in \partial\phi(x).$$

Proof. 1. Let $x \in D(\partial\phi), v \in \partial\phi(x)$. The definition of the subdifferential for λ-convex functionals yields:

$$\phi(y) \geq \phi(x) + \langle v, y - x\rangle + \frac{\lambda}{2}\|x - y\|^2.$$

For all $y \in H$. Since ϕ is proper there exists $\tilde{y}$, so that $c = \phi(\tilde{y}) < \infty$. Therefore, $c \geq \phi(x) + \langle v, \tilde{y} - x \rangle + \frac{\lambda}{2}\|x - \tilde{y}\|^2$. From here it follows that $\phi(x) \neq \infty$ and by definition $x \in D(\phi)$.

2. To show the second result assume $v \in \partial\phi(x), w \in \partial\phi(y)$ for $x, y \in D(\partial\phi)$. Again by the definition of the subdifferential for λ-convex functionals we get

$$\phi(y) \geq \phi(x) + \langle v, y - x \rangle + \frac{\lambda}{2}\|x - y\|^2$$
$$\phi(x) \geq \phi(y) + \langle w, x - y \rangle + \frac{\lambda}{2}\|x - y\|^2$$

Adding up the two inequalities we get:

$$\phi(y) + \phi(x) \geq \phi(y) + \langle v, y - x \rangle + \frac{\lambda}{2}\|x - y\|^2 + \phi(x) + \langle w, x - y \rangle + \frac{\lambda}{2}\|x - y\|^2.$$

We collect the $\frac{\lambda}{2}\|x - y\|^2$ terms to get

$$\phi(y) + \phi(x) \geq \phi(y) + \phi(x) + \langle v, y - x \rangle + \langle w, x - y \rangle + \lambda\|x - y\|^2.$$

We now subtract $\phi(y) + \phi(x)$ to arrive at $0 \geq \langle v, y - x \rangle + \langle w, x - y \rangle + \lambda\|x - y\|^2$. Hence $0 \geq \langle w - v, x - y \rangle + \lambda\|x - y\|^2$, which is equivalent to the desired inequality $\langle v - w, x - y \rangle \geq \lambda\|x - y\|^2$.

3. ($\Leftarrow$) Let $0 \in \partial\phi(x)$, by Definition 3.1.2 this is equivalent to

$$\phi(y) \geq \phi(x) + \langle 0, y - x \rangle + \frac{\lambda}{2}\|x - y\|^2$$

for all $y \in H$. As we assumed $\lambda \geq 0$ it follows that $\phi(y) \geq \phi(x)$ for all $y \in H$, hence $\phi(x)$ is minimal.

($\Rightarrow$) Let $\phi(x) = \min_{y \in H} \phi(y)$. Assume to the contrary $0 \notin \partial\phi(x)$, i.e. there exists $y \in H$, so that $\phi(x) + \frac{\lambda}{2}\|x - y\|^2 > \phi(y)$. Hence there exists a $\epsilon > 0$, so that

$$\frac{\lambda}{2}\|x - y\|^2 - \epsilon > \phi(y) - \phi(x). \tag{3.1}$$

Recall now the definition of λ-convexity

$$\phi(x + t(y - x)) \leq \phi(x) + t(\phi(y) - \phi(x)) - \frac{\lambda}{2}t(1 - t)\|x - y\|^2.$$

Using inequality (3.1) and expanding the last term we arrive at:

$$\phi(x + t(y - x)) < \phi(x) + t\left(\frac{\lambda}{2}\|x - y\|^2 - \epsilon\right) - \frac{\lambda t}{2}\|x - y\|^2 + \frac{\lambda t^2}{2}\|x - y\|^2.$$

Collecting terms depending on t (and ignoring the negative term $-\frac{\lambda t}{2}\|x - y\|^2$) we arrive at:

$$\phi(x + t(y - x)) < \phi(x) + \underbrace{t\left(\frac{\lambda t}{2}\|x - y\|^2 - \epsilon\right)}_{c(t)}.$$

This inequality should hold true for all $t \in [0,1]$ by definition. But if we choose $\tilde{t} < \frac{2\epsilon}{\lambda \|x-y\|^2}$, $c(\tilde{t})$ becomes negative and hence $\phi(x + \tilde{t}(y-x)) < \phi(x)$, which contradicts the minimality of $\phi(x)$.

□

3.1.2 The schemes

We turn to the investigation of the numerical schemes. To define the gradient flow on a Hilbert, assume for a moment that H is finite dimensional and that the entropy ϕ is C^1. In this case a gradient flow with initial point x_0 is a solution to

$$\begin{cases} \dot{x}(t) = -\nabla\phi(x(t)), \\ x(0) = x_0. \end{cases}$$

In analogy we define a gradient flow for λ-convex functionals on a possibly infinite Hilbert space as follows

Definition 3.1.4. We call $x : (0, +\infty) \to H$ a gradient flow for $\phi : H \to \mathbb{R}$ starting at $x_0 \in H$ if it is locally absolutely continuous and

$$\begin{cases} \dot{x}(t) = -\partial\phi(x(t)), \\ \lim_{t \downarrow 0} x(t) = x_0. \end{cases}$$

The analytical treatment of gradient flows on Hilbert spaces is covered by [8]. What we are going to do next is to introduce the one-leg schemes already mentioned in the introduction. Recall that a general k-step multistep method is of the form

$$\sum_{i=0}^{k} \alpha_i x_{n+i} = -\tau \left(\sum_{i=0}^{k} \beta_i \partial\phi(x_{n+i}) \right), \tag{3.2}$$

where τ denotes the time step size. As usual given an initial point x_0 the $k-1$ initial points for the k-step scheme will be generated by lower step schemes. The coefficients α_i, β_i define the generating polynomials:

$$\rho(\zeta) := \sum_{i=0}^{k} \alpha_i \zeta^i; \qquad \sigma(\zeta) := \sum_{i=0}^{k} \beta_i \zeta^i.$$

The corresponding one-leg scheme is given via

Definition 3.1.5. Let a k-step multistep scheme be given, such that the generating polynomials

$$\rho(\zeta) := \sum_{i=0}^{k} \alpha_i \zeta^i; \qquad \sigma(\zeta) := \sum_{i=0}^{k} \beta_i \zeta^i.$$

are relatively prime. Moreover we assume the normalization $\beta := \sum_{i=0}^{k} \beta_i = 1$. We then call the scheme

$$\sum_{i=0}^{k} \alpha_i x_{n+i} \in -\tau \partial \phi \left(\sum_{i=0}^{k} \beta_i x_{n+i} \right) \tag{3.3}$$

the associated one-leg scheme ore one leg method.

3.1.3 Existence and uniqness

To show the existence of a unique solution to (3.3) we rewrite this equation in a more suitable way. To illustrate this we consider first the BDF-1 approximation of a gradient flow, i.e. x_{n+1} is defined as a solution to $x - x_n \in -\tau \partial \phi(x)$. A general one-leg scheme (3.3) can be rewritten in a BDF-1 like way where x_n is replaced by a vector η known at time t_n and x is replaced by $\tilde{x}$ via a simple change of variables. Hence the solution to (3.3) is equivalent to a solution to

$$\tilde{x} + \eta \in -\tilde{\tau} \partial(\tilde{x}). \tag{3.4}$$

The reason for this reformulation is that a solution to the scheme in the form (3.4) is shown more easily. We now show how to derive (3.4). Let us start with a general one-leg scheme, i.e. x_{n+k} is given as the solution to

$$\alpha_k x + \sum_{i=0}^{k-1} \alpha_i x_{n+i} \in -\tau \partial \phi \left(\beta_k x + \sum_{i=0}^{k-1} \beta_i x_{n+i} \right).$$

Now we perform a change of variables $x \mapsto \tilde{x} = \beta_k x + \sum_{i=0}^{k-1} \beta_i x_{n+i}$ to arrive at

$$\alpha_k \underbrace{\frac{1}{\beta_k} \left(\tilde{x} - \sum_{i=0}^{k-1} \beta_i x_{n+i} \right)}_{x} + \sum_{i=0}^{k-1} \alpha_i x_{n+i} \in -\tau \partial \phi \left(\tilde{x} \right).$$

We separate the new variable $\tilde{x}$ from the variables $x_{n+k-1} \ldots, x_n$ known at time t_n. This yields

$$\frac{\alpha_k}{\beta_k} \tilde{x} + \frac{1}{\beta_k} \left(-\sum_{i=0}^{k-1} \beta_i x_{n+i} \right) + \sum_{i=0}^{k-1} \alpha_i x_{n+i} \in -\tau \partial \phi \left(\tilde{x} \right).$$

Multiplying by $\frac{\alpha_k}{\beta_k}$ yields

$$\underbrace{\tilde{x} + \alpha_k \left(-\sum_{i=0}^{k-1} \beta_i x_{n+i} \right) + \frac{\beta_k}{\alpha_k} \left(\sum_{i=0}^{k-1} \alpha_i x_{n+i} \right)}_{\eta} \in - \underbrace{\frac{\beta_k}{\alpha_k} \tau}_{\tilde{\tau}} \partial\phi\left(\tilde{x}\right),$$

the desired form:

$$\tilde{x} + \eta \in -\tilde{\tau}\phi(\tilde{x}). \tag{3.5}$$

In theorem 3.1.9 below we show that a unique solution $\tilde{x}$ to (3.5) exists. Then the solution x to (3.3) is given by:

$$x := \frac{1}{\beta_k} \left(\tilde{x} - \sum_{i=0}^{k-1} \beta_i x_{n+i} \right).$$

Before we can show the existence of a unique solution to (3.5) we give 3 lemmas which we need in the proof.

Lemma 3.1.6. *Assume that $\phi : H \to (-\infty, +\infty]$ is convex and lower semicontinuous in the strong topology. Then ϕ is lower semicontinuous in the weak topology.*

Proof. A proof is given in [7, Corollary 3.9] □

Lemma 3.1.7. *Assume that $\phi : H \to (-\infty, +\infty]$ is convex and lower semicontinuous. Then there exists a $C \in \mathbb{R}_+$ so that:*

$$\phi(x) \geq -C - C\|x\|.$$

Proof. Assume to the contrary that there exists a sequence $x_k \in H$, so that

$$\phi(x_k) \leq -k - k\|x_k\|.$$

If the sequence x_k is bounded, we subtract a weakly convergent subsequence: $x_k \rightharpoonup x$. But then by Lemma 3.1.6 it follows that $\phi(x) = -\infty$. Therefore assume $\|x_k\| \to \infty$. Let now $x_0 \in H$ be so that $\phi(x_0) < \infty$, and define

$$\tilde{x}_k := \frac{x_k}{\|x_k\|} + \left(1 + \frac{1}{\|x_k\|}\right) x_0,$$

then by convexity of ϕ

$$\phi(\tilde{x}_k) \leq \frac{\phi(x_k)}{\|x_k\|} + \left(1 + \frac{1}{\|x_k\|}\right) \phi(x_0) \leq -k + \phi(x_0).$$

Now $\tilde{x}_k$ is bounded. Subtracting a convergent subsequence: $\tilde{x}_k \rightharpoonup \tilde{x}$ we again derive the contradiction: $\phi(\tilde{x}) = -\infty$. □

Lemma 3.1.8. *Assume that* $\phi : H \to (-\infty, +\infty]$ *is* λ*-convex and lower semicontinuous. Then* $\tilde{\phi}(x) : H \to (-\infty, +\infty]$ *defined by:*

$$\tilde{\phi}(x) := \phi(x) - \frac{\lambda}{2}\|x\|^2$$

is convex and lower semicontinuous.

Proof. We first show the identity

$$\begin{aligned}
\|x + t(y-x)\|^2 &= \|x\|^2 + 2t\langle y-x, x\rangle + t^2\|y-x\|^2 \\
&= \|x\|^2 + 2t\langle y, x\rangle - 2t\|x\|^2 + t^2\|y-x\|^2 \\
&= (1-t)\|x\|^2 + 2t\langle y, x\rangle - t\|x\|^2 + t^2\|y-x\|^2 \\
&+ t\|y\|^2 - t\|y\|^2 \\
&= (1-t)\|x\|^2 + t\|y\|^2 - t(\|x\|^2 - 2\langle y, x\rangle + \|y\|^2) \\
&+ t^2\|y-x\|^2 \\
&= \|x\|^2 + t(\|y\|^2 - \|x\|^2) - t\|y-x\|^2 + t^2\|y-x\|^2 \\
&= \|x\|^2 + t(\|y\|^2 - \|x\|^2) - t(1-t)\|y-x\|^2
\end{aligned}$$

This means, that $\frac{\lambda}{2}\|x\|^2$ is "exactly" λ-convex:

$$\frac{\lambda}{2}\|x + t(y-x)\|^2 = \frac{\lambda}{2}\left(\|x\|^2 + t(\|y\|^2 - \|x\|^2) - t(1-t)\|y-x\|^2\right)$$

Now it is easy to see that:

$$\begin{aligned}
\tilde{\phi}(x + t(y-x)) &= \phi(x + t(y-x)) \\
&- \frac{\lambda}{2}\left(\|x\|^2 + t(\|y\|^2 - \|x\|^2) - t(1-t)\|y-x\|^2\right) \\
&\leq \phi(x) + t(\phi(y) - \phi(x)) - \frac{\lambda}{2}t(1-t)\|y-x\|^2 \\
&- \frac{\lambda}{2}\left(\|x\|^2 + t(\|y\|^2 - \|x\|^2)\right) + \frac{\lambda}{2}t(1-t)\|y-x\|^2 \\
&= \phi(x) + t(\phi(y) - \phi(x)) - \frac{\lambda}{2}\left(\|x\|^2 + t(\|y\|^2 - \|x\|^2)\right) \\
&= \tilde{\phi}(x) + t(\tilde{\phi}(y) - \tilde{\phi}(x)).
\end{aligned}$$

Hence $\tilde{\phi}$ is convex. The lower semicontinuous of $\tilde{\phi}$ follows by the lower semicontinuous of ϕ and the continuity of $\frac{1}{2}\|x\|^2$. □

Now we are prepared to show the desired existence to (3.5).

Theorem 3.1.9. *Assume a functional* $\phi : H \to (-\infty, \infty]$ *which is proper, lower semicontinuous and* λ*-convex. Then a unique solution to*

$$x + \eta \in -\tilde{\tau}\partial\phi(x)$$

exists if $\tilde{\tau}\lambda > -1$*, with* $\tilde{\tau} = \frac{\beta_k}{\alpha_k}\tau$*,* $\tau > 0$*, and* $\eta \in H$*.*

Proof. The proof works as follows. Lemma 3.1.6, 3.1.7 and 3.1.8 shows that the functional $\psi(x) := \frac{1}{2}\|x+\eta\|^2 + \tilde{\tau}\phi(x)$ attains a minimum x_M. We then show that x_M is a solution to (3.5). Uniqueness follows by λ-convexity. Lemma 3.1.7 and 3.1.8 yield the estimate

$$\phi(x) \geq -C - C\|x\| + \frac{\lambda}{2}\|x\|^2. \tag{3.6}$$

For $\psi(x) := \frac{1}{2}\|x+\eta\|^2 + \tilde{\tau}\phi(x)$ inequality (3.6) yields $\psi(x) \geq \frac{1}{2}\|x+\eta\|^2 + \frac{\tilde{\tau}\lambda}{2}\|x\|^2 - \tilde{\tau}C\|x\| - \tilde{\tau}C$. Therefore the behavior of $\psi(x)$ for $\|x\| \to \infty$ is dominated by the quadratic term $\frac{1+\tilde{\tau}\lambda}{2}\|x\|^2$. Hence for $\tilde{\tau}\lambda > -1$

$$\lim_{\|x\|\to\infty} \psi(x) = \infty. \tag{3.7}$$

Now choose a minimizing sequence x_n for $\psi(x)$, i.e. $\psi(x_n) \to \inf_{x\in H}\psi(x) = m$. Now (3.7) (together with $\phi(x) \neq -\infty$) implies that m is a finite number. Therefore the minimizing sequence (x_n) is bounded and we can extract a weakly convergent subsequence $x_{n_k} \rightharpoonup x_M$. Lemma 3.1.6 shows that ψ is lower semicontinuous with respect to the weak topology and therefore attains its minimum at x_M. We now show that x_M is the desired solution. With the above assumptions on τ it is easy to show that $\psi(x)$ is $\tilde{\lambda}$-convex with $\tilde{\lambda} = 1 + \tilde{\tau}\lambda > 0$. By point 3 in Lemma 3.1.3 we conclude that $0 \in \partial\psi(x_M)$. By the definition of the subdifferential this is equivalent to

$$\psi(x_M) + \langle 0, y - x_M\rangle + \frac{1+\tilde{\tau}\lambda}{2}\|y - x_M\|^2 \leq \psi(y),$$

for all $y \in H$. We use the definition of ψ to get

$$\frac{1}{2}\|x_M + \eta\|^2 + \tilde{\tau}\phi(x_M) + \langle 0, y - x_M\rangle + \frac{1+\tilde{\tau}\lambda}{2}\|y - x_M\|^2 \leq \frac{1}{2}\|y+\eta\|^2 + \tilde{\tau}\phi(y).$$

We collect terms on the left hand side

$$\frac{1}{2}\left(\|x_M + \eta\|^2 - \|y+\eta\|^2\right) + \tilde{\tau}\phi(x_M) + \langle 0, y - x_M\rangle + \frac{1+\tilde{\tau}\lambda}{2}\|y - x_M\|^2 \leq \tilde{\tau}\phi(y)$$

and rewrite the left hand side as follows

$$\underbrace{\frac{1}{2}\left(\|x_M + \eta\|^2 - \|y+\eta\|^2 + \|y - x_M\|^2\right)}_{=:A} + \tilde{\tau}\phi(x_M) + \frac{\tilde{\tau}\lambda}{2}\|y - x_M\|^2 \leq \tilde{\tau}\phi(y).$$

A straight forward expansion of A yields $A = -\langle x_M + \eta, y - x_M\rangle$ Hence

$$\tilde{\tau}\phi(x_M) - \langle x_M + \eta, y - x_M\rangle + \frac{\tilde{\tau}\lambda}{2}\|y - x_M\|^2 \leq \tilde{\tau}\phi(y).$$

Again by definition of the subdifferential $-\eta - x_M \in \tilde{\tau}\phi(x_M)$ or equivalently $x_M + \eta \in -\tilde{\tau}\phi(x_M)$. It remains to show uniqueness. Let us assume $x + \tilde{\tau}v = -\eta$, and $y + \tilde{\tau}w = -\eta$ for $v \in \partial\phi(x), w \in \partial\phi(y)$. By Lemma 1 it follows that

$$\lambda\|x-y\|^2 \leq \langle x-y, v-w\rangle = \langle x-y, \frac{y-x}{\tilde{\tau}}\rangle = -\frac{1}{\tilde{\tau}}\|x-y\|^2,$$

which is a contradiction to our assumption $\tilde{\tau}\lambda > -1$.

□

Definition 3.1.10. Let $\phi : H \to (-\infty, \infty]$, be proper lower semicontinuous and λ-convex, and $k-1$ points $x_{n+k-1}, ..., x_n \in H$ be given. We define

$$\tilde{x} := \underset{x\in H}{\operatorname{argmin}} \left\{ \frac{\|x+\eta\|^2}{2} + \tilde{\tau}\phi(x) \right\}, \tag{3.8}$$

and

$$x_{n+k} := \frac{\tilde{x} - \sum_{i=0}^{k-1}\beta_i x_{n+i}}{\beta_k},$$

where $\eta := \sum_{i=0}^{k-1}\left(\frac{\alpha_i\beta_k}{\alpha_k} - \beta_i\right)x_{n+i}$ and $\tilde{\tau} := \tau\frac{\beta_k}{\alpha_k}$ s.t. $\tilde{\tau}\lambda > -1$. Here the α_i and β_i are the coefficients of the generating polynomials as defined in Definition 3.1.5. We call the procedure (3.8) minimizing movement scheme or minimizing movement method.

3.1.4 Convergence

We show the convergence of discrete solutions obtained by minimizing movement schemes as given by definition 3.8 to the exact solution (see definition 3.1.4). Consider the sequence x_j generated by a minimizing movement scheme. To investigate the convergence of the scheme we introduce the concept of G-stability. As the solution at time step $n+k$ depends upon $k-1$ steps in the past we will consequently investigate a quantity depending on $X_n := (x_{n+k-1}, ..., x_n)$. This was the idea of Dahlquist who defined the G-norm as a norm for X_n as follows.

Definition 3.1.11. Let $x_{n+k-1}, ..., x_n \in H$, and $X_n := (x_{n+k-1}, ..., x_n)$. The G-norm of X_n is defined by

$$\|X_n\|_G^2 := \sum_{i=1}^{k}\sum_{j=1}^{k} g_{ij}\langle x_{n+i-1}, x_{n+j-1}\rangle,$$

where $G = (g_{ij})_{i,j=1,...,k}$ is supposed to be real, symmetric and positive definite.

We call a one-leg scheme G-stable, if the errors are contractive in the G-norm.

Definition 3.1.12. The method (3.3) is called G-stable, if there exists a real, symmetric and positive definite matrix G, so that for two numerical solutions x_n and y_n to a gradient flow with convex entropy (i.e. $\lambda = 0$) we have

$$\|X_{n+1} - Y_{n+1}\|_G \leq \|X_n - Y_n\|_G,$$

for all $\tau > 0$.

We mention that the BDF-2 method belongs to the class of G-stable one-leg methods. We give some auxiliary results which will be employed in the proof of convergence of G-stable schemes.

Lemma 3.1.13. *For a G-stable one leg method the following equality holds true:*

$$\langle \sum_{i_0}^{k} \alpha_i e_{n+i}, \sum_{i=0}^{k} \beta_i e_{n+i} \rangle = \|E_{n+1}\|_G^2 - \|E_n\|_G^2 + \| \sum_{i=0}^{k} a_i e_{n+i}\|^2,$$

for $a_i \in \mathbb{R}$, *where* $e_{n+i} = x(t_{n+i}) - x_{n+i}$, $\|E_{n+1}\|_G = \|X_{n+1} - X(t_n)\|_G$ *and* $x(t)$ *is the exact solution.*

Proof. A proof can be found in [5] □

Lemma 3.1.14. *Let* $X_n = (x_{n+k-1}, ...x_n)$ *and a G-norm be given by*

$$\|X_n\|_G^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} g_{ij} \langle x_{n+i-1}, x_{n+j-1} \rangle = \sum_{i=1}^{k} g_{ii} \|x_{n+i-1}\|^2 + \sum_{i \neq j} g_{ij} \langle x_{n+i-1}, x_{n+j-1} \rangle.$$

Then there exists a positive constant C, *so that*

$$\| \sum_{i=0}^{k} a_i x_{n+i}\|^2 \leq C(\|X_{n+1}\|_G^2 + \|X_n\|_G^2).$$

where $a_i \in \mathbb{R}$.

Proof. We want to show that:

$$C(\|X_{n+1}\|_G^2 + \|X_n\|_G^2) - \| \sum_{i=0}^{k} a_i x_{n+i}\|^2 \geq 0 \tag{3.9}$$

for some $C \in \mathbb{R}_+$. We first show that $\| \sum_{i=0}^{k} a_i x_{n+i}\|^2 \leq \tilde{C} \sum_{i=0}^{k} \|x_{n+i}\|^2$ by a straightforward computation.

$$
\begin{aligned}
\|\sum_{i=0}^{k} a_i x_{n+i}\|^2 &\le \left(\sum_{i=0}^{k}\|a_i x_{n+i}\|\right)^2 = \sum_{i=0}^{k}|a_i|^2\|x_{n+i}\|^2 + \sum_{i\neq j}|a_i||a_j| \underbrace{\|x_{n+i}\|\|x_{n+j}\|}_{\le\|x_{n+i}\|^2+\|x_{n+j}\|^2} \\
&\le \max_{i\le k}|a_i|^2\left(\sum_{i=0}^{k}\|x_{n+i}\|^2 + \sum_{i\neq j}(\|x_{n+i}\|^2+\|x_{n+j}\|^2)\right) \\
&\le \max_{i\le k}|a_i|^2\left(\sum_{i=0}^{k}\|x_{n+i}\|^2 + 2(k-1)\sum_{i=0}^{k}\|x_{n+i}\|^2\right) \\
&\le \underbrace{(\max_{i\le k}|a_i|^2)(2k-1)}_{=:\tilde{C}}\sum_{i=0}^{k}\|x_{n+i}\|^2.
\end{aligned}
$$

So (3.9) holds true if $C(\|X_{n+1}\|_G^2+\|X_n\|_G^2)-\tilde{C}\sum_{i=0}^{k}\|x_{n+i}\|^2 \ge 0$. The first term is the sum of G-norms. The second can be written as the sum of G-norms with matrix CI where I is the identity. We use this and absorb the constants C and $\tilde{C}$ in a new matrix and then argue by the eigenvalues of this new matrix. We estimate the left hand side by

$$
\begin{aligned}
&C(\|X_{n+1}\|_G^2+\|X_n\|_G^2)-\tilde{C}\sum_{i=0}^{k}\|x_{n+i}\|^2 \\
&= C(\|X_{n+1}\|_G^2+\|X_n\|_G^2)-\tilde{C}\sum_{i=1}^{k}\|x_{n+i}\|^2-\tilde{C}\|x_n\|^2 \\
&\ge C\|X_{n+1}\|_G^2-\tilde{C}\sum_{i=1}^{k}\|x_{n+i}\|^2+C\|X_n\|_G^2-\tilde{C}\sum_{i=0}^{k-1}\|x_{n+i}\|^2 \\
&= \|X_{k+1}\|_{CG}^2-\|X_{k+1}\|_{\tilde{C}I}^2+\|X_k\|_{CG}^2-\|X_k\|_{\tilde{C}I}^2 \\
&= \|X_{k+1}\|_{CG-\tilde{C}I}^2+\|X_k\|_{CG-\tilde{C}I}^2.
\end{aligned}
$$

The matrix $CG-\tilde{C}I$ is symmetric and real. We only have to show that $CG-\tilde{C}I$ is positive definite for some C. Let λ_{min} be the minimal eigenvalue of G. Then $C\lambda_{min}$ is the minimal eigenvalue of CG. As $\tilde{C}I$ is just a multiple of the identity the minimal eigenvalue of $CG-\tilde{C}I$ is given by $\tilde{\lambda}_{min} = C\lambda_{min}-\tilde{C}$. Hence the matrix $CG-\tilde{C}I$ is positive definite if $\tilde{\lambda}_{min} > 0$, i.e. if we choose C, so that:

$$
C \ge \frac{\tilde{C}}{\lambda_{\min}} = \frac{(\max|a_i|^2)(2k-1)}{\lambda_{\min}}.
$$

□

Remark 7. As $C(\|X_{n+1}\|_G^2+\|X_n\|_G^2) \le C(\|X_{n+1}\|_G+\|X_n\|_G)^2$ also the inequality $\|\sum_{i=0}^{k} a_i x_{n+i}\| \le \sqrt{C}(\|X_{n+1}\|_G+\|X_n\|_G)$ holds true.

Lemma 3.1.15. *Assume* $a, b, \hat{C}, \tilde{C}_1, \tilde{C}_2, \lambda \in [0, +\infty)$, *so that*

$$a^2 \leq (1 + \hat{C}\tau\lambda)b^2 + \tilde{C}_1\tau^{(p+1)}(a + b) + \tilde{C}_2\tau^{2(p+1)}.$$

Then there exists $M \in [0, +\infty)$, *so that* $a \leq (1 + \hat{C}\tau\lambda)b + M\tau^{p+1}$.

Proof. To collect now all terms depending on a on the left and all terms depending on b on the right we complete the squares:

$$\left(a - \frac{\tilde{C}_1\tau^{p+1}}{2}\right)^2 - \left(\frac{\tilde{C}_1\tau^{p+1}}{2}\right)^2 \leq \left(\left(\sqrt{1 + \hat{C}\tau\lambda}\right)b + \frac{\tilde{C}_1\tau^{p+1}}{2\sqrt{1 + \hat{C}\tau\lambda}}\right)^2 - \left(\frac{\tilde{C}_1\tau^{p+1}}{2\sqrt{1 + \hat{C}\tau\lambda}}\right)^2 + \tilde{C}_2\tau^{2(p+1)}.$$

We drop the negative term on the right and bring the second term on the left to the right:

$$\left(a - \frac{\tilde{C}_1\tau^{p+1}}{2}\right)^2 \leq \left(\left(\sqrt{1 + \hat{C}\tau\lambda}\right)b + \frac{\tilde{C}_1\tau^{p+1}}{2\sqrt{1 + \hat{C}\tau\lambda}}\right)^2 + \left(\frac{\tilde{C}_1\tau^{p+1}}{2}\right)^2 + \tilde{C}_2\tau^{2(p+1)}$$

As all the terms on the r.h.s are non-negative we use the elementary inequality $\sqrt{x^2 + y^2} \leq x + y$ for $x, y > 0$ to arrive at

$$a - \frac{\tilde{C}_1\tau^{p+1}}{2} \leq \left(\sqrt{1 + \hat{C}\tau\lambda}\right)b + \frac{\tilde{C}_1\tau^{p+1}}{2\sqrt{1 + \hat{C}\tau\lambda}} + \frac{\tilde{C}_1\tau^{p+1}}{2} + \sqrt{\tilde{C}_2}\tau^{(p+1)}.$$

Hence $a \leq \left(\sqrt{1 + \hat{C}h\lambda}\right)b + \frac{3\overline{C}}{2}\tau^{p+1} + \sqrt{\tilde{C}_2}\tau^{p+1}$. Define $M := \max\{\frac{3\tilde{C}_1}{2}, \sqrt{\tilde{C}_2}\}$ so that

$$a \leq \left(\sqrt{1 + \hat{C}\tau\lambda}\right)b + M\tau^{p+1} \leq \left(1 + \hat{C}\tau\lambda\right)b + M\tau^{p+1},$$

where the second inequality holds because $\tau\lambda \geq 0$. □

Before we now can tackle the convergence result, there is one last definition we need in the proof.

Definition 3.1.16. Let $x : [0, T] \to H$. We define the differentiation error δ_D and the integration error δ_I by:

$$\delta_D := \sum_{i=0}^{k} \alpha_i x(t + i\tau) - \tau\dot{x}(t + \beta\tau); \qquad \delta_I := \sum_{i=0}^{k} \beta_i x(t + i\tau) - x(t + \beta\tau);$$

where $\beta = \sum_{i=0}^{k} i\beta_i$.

We are now prepared to show the following convergence result.

Theorem 3.1.17. *Let $x(t)$ be the exact solution to a gradient flow with λ-convex entropy ϕ and x_n the solution generated by a minimizing movement scheme. Define $e_n := x(t_n) - x_n$. If then the scheme is G-stable and the error δ_D is of order p and δ_I is of order p-1. Then*

$$\|e_n\| < C \max_{i=0,\dots,k} \|e_k\| + M\tau^p.$$

For some $C > 0$ and $\tau\tilde{\lambda} < 1$ where $\tilde{\lambda} = K \max\{0, \lambda\}$ and K some constant depending on the G-matrix of the scheme.

Proof. Let $x(t)$ be the exact solution to a gradient flow. Then $x(t)$ is a solution to:

$$\sum_{i=0}^{k} \alpha_i x(t + i\tau) - \delta_D \in -\tau\partial\phi \left(\sum_{i=0}^{k} \beta_i x(t + i\tau) - \delta_I \right).$$

Now the procedure would be to compare this with the approximation (3.3). The order of the error then depends on δ_D and δ_I. However, we now turn to another perturbed problem of the form:

$$\sum_{i=0}^{k} \alpha_i \hat{x}(t + i\tau) - \hat{\delta}_D \in -\tau\partial\phi \left(\sum_{i=0}^{k} \beta_i \hat{x}(t + i\tau) - \hat{\delta}_I \right). \tag{3.10}$$

Where we define $\hat{x}(t) := x(t) - \delta_I(t)$, and

$$\hat{\delta}_D(t) := \delta_D - \sum_{i=0}^{k} \alpha_i \delta_D(t + i\tau); \qquad \hat{\delta}_I(t) := \delta_I - \sum_{i=0}^{k} \beta_i \delta_I(t + i\tau).$$

We do this because the order of $\hat{\delta}_I$ is higher than the order of δ_I whereas the order of $\hat{\delta}_D$ is the same as the order of δ_D:

$$\mathrm{ord}(\hat{\delta}_I) = \mathrm{ord}(\delta_I) + 1, \qquad \mathrm{ord}(\hat{\delta}_D) = \mathrm{ord}(\delta_D). \tag{3.11}$$

Gaining one order in $\hat{\delta}_I$ allows us to assume δ_I to be only of order $p - 1$. It is an easy calculation to show that the error of the approximation to the perturbed problem $\|\hat{e}_n\| = \|x_n - \hat{x}(t_n)\|$ is of the same order as the error as the original error $\|e_n\| = \|x_n - x(t_n)\|$, which concludes the proof. The idea of investigating problem (3.10) goes back to Hundstorfer & Steininger (see [33]).

$$\begin{aligned}\hat{\delta}_I(t) &= \delta_I(t) - \sum_{i=0}^{k} \beta_i \delta_I(t + i\tau) \\ &= \delta_I(t) - \sum_{i=0}^{k} \beta_i \left(\sum_{j=0}^{p} \frac{1}{j!} \delta_I^{(j)}(t) \tau^j i^j + O(\tau^{p+1}) \right) \\ &= \delta_I(t) - \sum_{j=0}^{p} \left(\sum_{i=0}^{k} \beta_i \right) \frac{1}{j!} \delta_I^{(j)}(t) \tau^j i^j + O(\tau^{p+1}) \\ &= \delta_I(t) - \delta_I(t) \left(\sum_{i=0}^{k} \beta_i \right) - \sum_{j=1}^{p} \left(\sum_{i=0}^{k} \beta_i \right) \frac{1}{j!} \delta_I^{(j)} \tau^j i^j + O(\tau^{p+1}).\end{aligned}$$

Recall that in the definition of the one-leg scheme we assumed the normalization $\sum_{i=0}^{k} \beta_i = 1$. Hence

$$\hat{\delta}_I(t) = \delta_I(t) - \delta_I(t) - \sum_{j=0}^{p} \left(\sum_{i=0}^{k} \beta_i \right) \frac{1}{j!} \delta_I^{(j)} \tau^j i^j + O(\tau^{p+1}).$$

For every j we we set $C_j := \sum_{i=0}^{k} \frac{\beta_i i^j}{j!}$ and arrive at $\hat{\delta}_I(t) = -\sum_{j=1}^{p} C_j \delta_I^{(j)} \tau^j + O(\tau^{p+1})$. The $\delta_I^{(j)}$ are given by: $\delta_I^{(j)}(t) = \sum_{i=0}^{k} \beta_i x^{(j)}(t + i\tau) - x^{(j)}(t + \beta\tau)$. As this term is of order $p - 1$ for every function smooth enough, we see, that $\sum_{j=1}^{p} C_j \delta_I^{(j)} \tau^j + O(\tau^{p+1})$ is of order p and hence $\hat{\delta}_I$ is of order p. A similar calculation shows, that $\hat{\delta}_D$ is of order p for $\delta_D = O(\tau^{p+1})$:

$$\begin{aligned}\hat{\delta}_D &= \delta_D - \sum_{i=0}^{k} \alpha_i \delta_I(t + i\tau) \\ &= \delta_D - \sum_{i=0}^{k} \alpha_i \left(\sum_{j=0}^{p} \frac{1}{j!} \delta_I^{(j)} \tau^j i^j + O(\tau^{p+1}) \right) \\ &= \delta_D - \underbrace{\left(\sum_{i=0}^{k} \alpha_i \right)}_{=0} \delta_I(t) - \left(\sum_{j=1}^{p} \sum_{i=0}^{k} \alpha_i \frac{1}{j!} \delta_I^{(j)} \tau^j i^j + O(\tau^{p+1}) \right) \\ &= \delta_D - \left(\sum_{j=1}^{p} C_j \delta_I^{(j)} \tau^j + O(\tau^{p+1}) \right).\end{aligned}$$

As we assumed δ_D to be of order p and the second term in the last line again is of order p, we see, that $\hat{\delta}_D$ is of order p. Now we estimate the error between $\hat{x}(t)$

and the discrete solution x_n generated by the minimizing movement scheme.

$$\sum_{i=0}^{k} \alpha_i \underbrace{(x_{n+i} - \hat{x}(t+i\tau))}_{=:\hat{e}_{n+i}} - \hat{\delta}_D \in -\tau\partial\phi\left(\sum_{i=0}^{k} \beta_i x(t_n + i\tau) - \hat{\delta}_I\right) - \tau\partial\phi\left(\sum_{i=0}^{k} \beta_i x_{n+i}\right).$$

Next we take advantage of the monotonicity inequality (point 2 in Lemma 3.1.3). Therefore we consider the inner product:

$$\langle \sum_{i=0}^{k} \alpha_i \hat{e}_{n+i} - \hat{\delta}_D, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} - \hat{\delta}_I \rangle. \tag{3.12}$$

By (3.10) we know that $\sum_{i=0}^{k} \alpha_i \hat{e}_{n+i} - \hat{\delta}_D = -\tau(v-w)$ for some $v \in \partial\phi(\sum_{i=0}^{k} \beta_i \hat{x}(t_n + ih) - \delta_I)$ and $w \in \partial\phi(\sum_{i=0}^{k} \beta_i x_{n+i})$. Now (3.12) becomes

$$-\tau\langle v - w, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} - \hat{\delta}_I \rangle.$$

Set $x = \sum_{i=0}^{k} \beta_i \hat{x}(t_n + ih) - \delta_I$ and $y = \sum_{i=0}^{k} \beta_i x_{n+i}$ to write (3.12) as $-\tau\langle v-w, x-y\rangle$ with $v \in \partial\phi(x)$ and $w \in \partial\phi(y)$. By the monotonicity inequality (point 2 in Lemma 3.1.3) we get $-\tau\langle v-w, x-y\rangle \leq -\tau\lambda\|x-y\|^2$. Using the relations for v, w, x, y we get:

$$\langle \sum_{i=0}^{k} \alpha_i \hat{e}_{n+i} - \hat{\delta}_D, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} - \hat{\delta}_I \rangle \leq -\tau\lambda\| \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} - \hat{\delta}_I \|^2.$$

For further calculations it will be convenient to expand the left hand side. This yields

$$\langle \sum_{i=0}^{k} \alpha_i \hat{e}_{n+i}, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} \rangle - \langle \sum_{i=0}^{k} \alpha_i \hat{e}_{n+i}, \hat{\delta}_I \rangle - \langle \hat{\delta}_D, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} \rangle + \langle \hat{\delta}_D, \hat{\delta}_I \rangle$$
$$\leq -\tau\lambda\| \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} - \hat{\delta}_I \|^2.$$

We want to get rid of the first term on the left and substitute it by a G-norm dependent term. To this end we use Lemma 3.1.13 to get

$$\langle \sum_{i=0}^{k} \alpha_i \hat{e}_{n+i}, \sum_{i=0}^{k} \beta_i \hat{e}_{n+i} \rangle = \|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 + \| \sum_{i=0}^{k} a_i \hat{e}_{n+i} \|^2,$$

and hence $\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 \leq \langle \sum_{i=0}^k \alpha_i \hat{e}_{n+i}, \sum_{i=0}^k \beta_i \hat{e}_{n+i}\rangle$. Therefore

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 - \|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\|\|\hat{\delta}_I\| - \|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|\|\hat{\delta}_D\| - \|\hat{\delta}_D\|\|\hat{\delta}_I\|$$
$$\leq -\tau\lambda \left(\|\sum_{i=0}^k \beta_i \hat{e}_{n+i} - \hat{\delta}_I\|^2 \right),$$

where for the line in the middle we employed the Cauchy-Schwarz inequality. The "bad" case would be, that $\lambda < 0$. We therefore use from here on the inequality:

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 - \|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\|\|\hat{\delta}_I\| - \|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|\|\hat{\delta}_D\| - \|\hat{\delta}_D\|\|\hat{\delta}_I\|$$
$$\leq -\tau\overline{\lambda} \left(\|\sum_{i=0}^k \beta_i \hat{e}_{n+i} - \hat{\delta}_I\|^2 \right),$$

where we define $\overline{\lambda} := \max\{0, -\lambda\}$. Using the triangle inequality on the right-hand side we arrive at

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 - \|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\|\|\hat{\delta}_I\| - \|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|\|\hat{\delta}_D\| - \|\hat{\delta}_D\|\|\hat{\delta}_I\|$$
$$\leq \tau\overline{\lambda} \left(\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\| + \|\hat{\delta}_I\| \right)^2.$$

We collect the errors on the left hand side which yields

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 \leq \|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\|\|\hat{\delta}_I\| + \|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|\|\hat{\delta}_D\| + \|\hat{\delta}_D\|\|\hat{\delta}_I\|$$
$$+ \tau\overline{\lambda} \left(\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|^2 + \|\hat{\delta}_I\|^2 + 2\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|\|\hat{\delta}_I\| \right).$$

So far we derived an inequality between errors in the G-norm $\|E_n\|_G$ and errors in $\|e_n\|$. We use Lemma 3.1.14 and remark 7 above to show that the terms $\|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\|$ and $\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|$ can be estimated by $\|\hat{E}_n\|_G + \|\hat{E}_{n+1}\|_G$, as well as $\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|^2$ can be estimated by $\|\hat{E}_n\|_G^2 + \|\hat{E}_{n+1}\|_G^2$, and hence deduce an inequality in terms of the G-norm, only. More precisely:

$$\|\sum_{i=0}^k \alpha_i \hat{e}_{n+i}\| \leq C_1(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G); \quad \|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\| \leq C_2(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G);$$

$$\|\sum_{i=0}^k \beta_i \hat{e}_{n+i}\|^2 \leq K(\|\hat{E}_{n+1}\|_G^2 + \|\hat{E}_n\|_G^2).$$

Using this we get

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 \leq K\tau\overline{\lambda}(\|\hat{E}_{n+1}\|_G^2 + \|\hat{E}_n\|_G^2) \tag{3.13}$$
$$+ \tau\overline{\lambda}2C_2(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G)\|\hat{\delta}_I\| \tag{3.14}$$
$$+ \tau\overline{\lambda}\|\hat{\delta}_I\|^2 \tag{3.15}$$
$$+ C_1(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G)\|\hat{\delta}_I\| \tag{3.16}$$
$$+ C_2(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G)\|\hat{\delta}_D\| \tag{3.17}$$
$$+ \|\hat{\delta}_I\|\|\hat{\delta}_I\|. \tag{3.18}$$

The lines (3.16) and (3.17) are of order p, the line (3.14) is of order $p+1$. Hence for $\tau < 1$ we can estimate the 3 lines by a term of the form $C_a\tau^{p+1}(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G)$. The line (3.15) is of order $2(p+1)+1$ and (3.18) is of order $2(p+1)$. Hence for $\tau < 1$ we estimate them by a term $C_b\tau^{2(p+1)}$. We define $\tilde{\lambda} := K\overline{\lambda}$ and we derive the inequality

$$\|\hat{E}_{n+1}\|_G^2 - \|\hat{E}_n\|_G^2 \leq \tau\tilde{\lambda}\left(\|\hat{E}_{n+1}\|_G^2 + \|\hat{E}_n\|_G^2\right) + \overline{C}_1\tau^{p+1}(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G) + \overline{C}_2\tau^{2(p+1)}.$$

Next we collect the $\|E_{n+1}\|_G^2$ on the left and the $\|E_n\|_G^2$ on the right and divide by $(1-\tau\tilde{\lambda})$ to arrive at

$$\|\hat{E}_{n+1}\|_G^2 \leq (1 + \frac{2\tau\tilde{\lambda}}{1-\tau\tilde{\lambda}})\|\hat{E}_n\|_G^2 + \frac{\overline{C}_1}{1-\tau\tilde{\lambda}}\tau^{p+1}(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G) + \frac{\overline{C}_2}{1-\tau\tilde{\lambda}}\tau^{2(p+1)}$$

where we used $\frac{1+\tau\tilde{\lambda}}{1-\tau\tilde{\lambda}} = 1 + \frac{2\tau\tilde{\lambda}}{1-\tau\tilde{\lambda}}$. Assume now $\tau\tilde{\lambda} < 1$ to get

$$\|\hat{E}_{n+1}\|_G^2 \leq (1 + \hat{C}\tau\tilde{\lambda})\|\hat{E}_n\|_G^2 + \tilde{C}_1\tau^{p+1}(\|\hat{E}_{n+1}\|_G + \|\hat{E}_n\|_G) + \tilde{C}_2\tau^{2(p+1)}.$$

This is an inequality of the form in Lemma 3.1.15. Therefore it follows that $\|\hat{E}_{n+1}\|_G \leq \left(1 + \hat{C}\tau\tilde{\lambda}\right)\|\hat{E}_n\|_G + \tilde{M}\tau^{p+1}$. We now apply a discrete version of Gronwall's lemma to get $\|\hat{E}_{n+1}\|_G \leq e^{(\hat{C}\tilde{\lambda}T)}\|\hat{E}_0\|_G + \tilde{M}\tau^p$. It is easy to see (by equivalence of norms), that $\|\hat{E}_0\|_G < c_o \max_{i=0,\dots,k}\|\hat{e}_i\|$, as well as $\|\hat{e}_{n+1}\| < c_{n+1}\|\hat{E}_{n+1}\|_G$. We arrive at $\|\hat{e}_{n+1}\| \leq \hat{C}\max_{i=0,\dots,k}\|\hat{e}_i\| + \tilde{M}\tau^p$. for some $\hat{C} > 0$. So far we estimated the error between the discrete solution x_n and the exact solution of the perturbed problem $\hat{x}(t)$. It remains to estimate $\|e_n\|$ by $\|\hat{e}_n\|$. To this end recall the definition $\hat{x}(t) := x(t) - \delta_I(t)$. We get

$$\begin{aligned}\|e_n\| &= \|x_n - x(t_n)\| = \|x_n - \hat{x}(t_n) - \delta_I\| < \|x_n - \hat{x}(t_n)\| + \|\delta_I\| \\ &< C \max_{i=0,\dots,k} \|\hat{e}_k\| + \hat{M}\tau^p.\end{aligned}$$

It remains to compare $\|e_k\|$ and $\|\hat{e}_k\|$

$$\begin{aligned}\|\hat{e}_k\| &= \|x_k - \hat{x}(t_k)\| = \|x_k - \hat{x}(t_k) + \delta_I - \delta_I\| < \|x_k - x(t_k)\| + \|\delta_I\| \\ &< \|e_k\| + \|\delta_I\|.\end{aligned}$$

Finally we infer the desired result $\|e_n\| < C \max_{i=0,\dots,k} \|e_k\| + M\tau^p$, for some $C > 0$ and some $M > 0$. □

Remark 8. As we defined $\overline{\lambda} := \max\{0, -\lambda\}$, we can set $C = 1$ for convex (i.e. $\lambda \geq 0$) entropies.

Remark 9. We made several assumption on τ. For the solution of (3.3) we assumed $\tau \frac{\beta_k}{\alpha_k}\lambda > -1$. In the proof of convergence of the scheme we needed $K\tau \max\{0, -\lambda\} < 1$, as well as $\tau < 1$, and $\tau\tilde{\lambda} < 1$. We choose

$$\tau < \min\left\{\frac{1}{|\frac{\beta_k}{\alpha_k}\lambda|}, \frac{1}{|K\lambda|}, 1, \frac{1}{\tilde{\lambda}}\right\}.$$

Hence these 4 conditions do not yield any contradiction.

3.2 Entropy properties

By definition, the solution $x(t)$ to the equation

$$\begin{cases}\dot{x}(t) \in -\partial\phi(x(t)) \\ \lim\limits_{t\downarrow 0} x(t) = x_0.\end{cases}$$

dissipates the entropy ϕ, i.e. $\phi(x_{t_1}) < \phi(x_{t_2})$, for $t_1 < t_2$. Unfortunately such a property will not hold true for the discrete solution obtained by a one-leg scheme in general. We want to illustrate this fact by the following counterexample. We show that even on $\mathbb{R}$ the dissipation property will not hold true. Set

$$\phi(x) := \begin{cases}-bx, & x \leq 0 \\ ax, & x > 0\end{cases}$$

3.2.1 Violation of entropy dissipation

We now want to generate a discrete solution to the problem:

$$\begin{cases} \dot{x}(t) \in & -\partial\phi(x(t)) \\ x(0) = & x_0 \end{cases}$$

by the BDF 2 scheme. The G-stability of the scheme can be shown easily (see e.g. [34]). The coefficients or the BDF-2 scheme are given by

- $\alpha_0 = \frac{1}{2}, \alpha_1 = -2, \alpha_2 = \frac{3}{2}$,
- $\beta_0 = \beta_1 = 0, \beta_2 = 1$.

By definition 3.1.10 we have to solve first the minimization problem

$$\tilde{x} := \operatorname*{argmin}_{x \in H} \left\{ \frac{\|x + \eta\|^2}{2} + \tilde{\tau}\phi(x) \right\},$$

and then find x_{n+k} by

$$x_{n+k} := \frac{\tilde{x} - \sum_{i=0}^{k-1} \beta_i x_{n+i}}{\beta_k}.$$

Using the definition of β_i it follows that $x_{n+2} = \tilde{x}$. Hence the problem is to find

$$x_{n+2} := \operatorname*{argmin}_{x \in H} \underbrace{\left\{ \frac{\|x + \eta\|^2}{2} + \tau\frac{2}{3}\phi(x) \right\}}_{\psi(x)}. \tag{3.19}$$

Using the definition of the α_i we get $\eta := \frac{1}{3}x_n - \frac{4}{3}x_{n+1}$. Assume for the moment that x_n, x_{n+1} are chosen in such a way that so that $x_{n+2} = 0$. Now the solution should not move away from 0 as this is the state of minimal entropy. Think of the BDF-2 approximation in the form (3.19) as a BDF-1 approximation with initial condition $-\eta$ for the entropy $\frac{2}{3}\phi$. If the entropy is too flat the flow starting at $-\eta$ will be too slow to reach 0 within one time step. Hence $x_{n+3} \neq 0$ and consequently $\phi(x_{n+3}) > 0$. We now show that this is possible for any τ. Therefore let τ be given. Let

$$x_0 \in ((n+1)\tau a + \frac{1}{3}\tau a, (n+2)\tau a)$$

where $n > 2$. We find x_1 by a BDF-1 approximation, i.e.

$$x_1 = \operatorname*{argmin}_{x \in H} \left\{ \frac{\|x - x_0\|^2}{2} + \tau\phi(x) \right\}.$$

We first show that $x_1 \in \mathbb{R}_+$. Assume to the contrary $x_1 \leq 0$, then

$$\psi(x) > \frac{\|x_1 - x_0\|^2}{2} > \frac{\|x_0\|^2}{2} = \frac{(\tau a m)^2}{2},$$

for $m \in (n+1+\frac{1}{3}, n+2)$. As $n > 2$, it follows that $\frac{m}{2} \geq 1$, hence

$$\frac{(\tau am)^2}{2} = (\tau a)\frac{m}{2}\tau am \geq (\tau a)\tau am = \tau a x_0 = \tau\phi(x_0) = \psi(x_0).$$

Hence if $x_1 \in \mathbb{R}\backslash\mathbb{R}_+$ it can not be the minimum and therefore $x_1 \in \mathbb{R}_+$. On $\mathbb{R}_+$ the entropy ϕ is smooth and we find the minimum by calculating the first derivative and setting it zero. This yields $x_1 = x_0 - \tau a$. We calculate x_2. To this end we first calculate $\eta = \frac{1}{3}x_0 - \frac{4}{3}x_1 = -x_0 + (1+\frac{1}{3})\tau a$. We have to solve the problem:

$$x_2 = \underset{x \in H}{\operatorname{argmin}} \left\{ \frac{\|x-(-\eta)\|^2}{2} + \tau\phi(x) \right\}.$$

As $-\eta \in (n+\frac{1}{3}, n+1)$, by the same arguments as above we find that $x_2 \in \mathbb{R}_+$ and $x_2 = x_1 - \tau a = x_0 - 2\tau a$. Repeating the arguments we find that for $2 < i \leq n$: $x_i = x_0 - i\tau a$. For x_{n+1} we find

$$x_{n+1} = x_0 - (n+1)\tau a \in ((n+1)\tau a + \frac{1}{3}\tau a, (n+2)\tau a) - (n+1)\tau a = (\frac{1}{3}\tau a, \tau a).$$

The η for the calculation of x_{n+2} is given by

$$\begin{aligned} \eta &= \frac{1}{3}x_n - \frac{4}{3}x_{n+1} \\ &= \frac{1}{3}(x_0 - n\tau a) - \frac{1}{3}(x_0 - (n+1)\tau a) - (x_0 - (n+1)\tau a) \\ &= \frac{1}{3}\tau a - x_0 + (n+1)\tau a \\ &\in -((n+2)\tau a, -\frac{1}{3}\tau a - (n+1)\tau a) + \frac{1}{3}\tau a + (n+1)\tau a \\ &= (-\frac{2}{3}\tau a, 0). \end{aligned}$$

We now show that η is such that $x_{n+2} = 0$. As ϕ is not differentiable at zero we have to solve the minimization problem. Assume $x \in (-\infty, 0)$ then

$$\psi(x) = \frac{x^2 + \overbrace{2x\eta}^{>0} + \eta^2}{2} + \tau\frac{2}{3}\overbrace{bx}^{>0} > \frac{\eta^2}{2} = \psi(0).$$

On the other hand let $x \in (0, \infty)$ then

$$\psi(x) = \frac{x^2 + 2x\eta + \eta^2}{2} + \tau\frac{2}{3}ax = \frac{x^2+\eta^2}{2} + \underbrace{x\eta}_{>-\tau\frac{2}{3}ax} + \tau\frac{2}{3}ax > \frac{\eta^2}{2} = \psi(0).$$

So $x_{n+2} = 0$. We show that $x_{n+3} < 0$. The new η is given by $\frac{1}{3}x_{n+1}$. We have $x_{n+1} \in (\frac{1}{3}\tau a, \tau a)$. Let now b be so that $\tau\frac{2}{3}b < \frac{\eta}{2}$. For $\psi(x)$ we get

$$\psi(x) = \frac{x^2+\eta^2}{2} + x\eta - \tau\frac{2}{3}bx = \frac{x^2+\eta^2}{2} + x\underbrace{\left(\eta - \tau\frac{2}{3}b\right)}_{>\frac{\eta}{2}}.$$

If $x < 0$ we get

$$\psi(x) < \frac{\eta^2}{2} + \underbrace{\frac{x^2+x\eta}{2}}_{A}.$$

If $x \in (-\frac{\eta}{2}, 0)$ A is negative, hence

$$\psi(x) < \frac{\eta^2}{2} = \psi(0); \qquad \forall x \in (-\frac{\eta}{2}, 0).$$

Hence $x_{n+3} \neq 0$ (in fact $x_{n+3} = -\frac{\eta}{2}$) and the scheme does not dissipate the energy as

$$\phi(x_{n+3}) = -bx_{n+3} > 0 = \phi(x_{n+2}).$$

Remark 10. We derive a relation between a and b which allows for the violation of energy dissipation. Above we set $x_{n+1} \in \left(\frac{1}{3}\tau a, \tau a\right)$. We observed that η in the critical step is given by $\eta = \frac{1}{3}x_{n+1}$, i.e. $\eta \in \left(\frac{1}{9}\tau a, \frac{1}{3}\tau a\right)$. We observed that the entropy is not dissipated if $\tau\frac{2}{3}b < \frac{\eta}{2}$. This holds true iff $\tau\frac{2}{3}b < \frac{1}{18}\tau a$, i.e. $b < \frac{a}{12}$ revealing that the condition is independent of τ.

3.2.2 The G-norm: A discrete entropy

We now want to find some quantity that is dissipated by the discrete solution. The example above shows that even if x_n is already the point of minimal energy x_M, it is possible that in the next step we get some other point x_{n+1}. If we would take some other quantity as an entropy which is evaluated at one point only, then for this quantity let's say $\varphi(x)$ it should hold true that x_M is also the point where $\varphi(x)$ attains its unique minimum. But then this quantity can as well not be an entropy for the discrete solution (as $\varphi(x_{n+1}) > \varphi(x_n)$). Therefore we need a function of several x_i. For example a function depending on the k points of the scheme. It turns out that the G-norm, which we used to show convergence of the scheme, is dissipated by the discrete solution. At least for convex functionals (i.e. $\lambda \geq 0$), as the following calculation shows. Assume we are given a convex function $\phi(x)$, and a gradient flow with respect to ϕ, i.e. $\dot{x}(t) \in -\partial\phi(x(t))$. By the third point in Lemma 3.1.3 we get

$$\langle \tilde{v} - \tilde{w}, x - y\rangle \geq \lambda\|x-y\|^2 \geq 0$$

for $\tilde{v} \in \partial\phi(x), \tilde{w} \in \partial\phi(y)$. Hence for $v \in -\partial\phi(x), w \in -\partial\phi(y)$ we get:

$$\langle v - w, x - y\rangle \leq 0 \tag{3.20}$$

Assume without loss of generality $\phi(0) = \min_{x\in X}\{\phi(x)\}$ and let

$$\sum_{i=0}^{k} \alpha_i x_{n+i} \in -h\partial\phi\left(\sum_{i=0}^{k} \beta_i x_{n+i}\right).$$

Then by (3.20) $\langle v - w, \sum_{i=0}^{k} \beta_i x_{n+i} - 0\rangle \leq 0$ where $v \in -\partial\phi(\sum_{i=0}^{k} \beta_i x_{n+i})$ and $w \in -\partial\phi(0)$. We know, as $\phi(0)$ is the minimum, that: $0 \in \partial\phi(x)$. Therefore set $v := \sum_{i=0}^{k} \alpha_i x_{n+i}$ and $w := 0$ Using this in inequality (3.20) we arrive at

$$\langle\sum_{i=0}^{k} \alpha_i x_{n+i}, \sum_{i=0}^{k} \beta_i x_{n+i}\rangle \leq 0. \tag{3.21}$$

From Lemma 3.1.13 we know that

$$\langle\sum_{i=0}^{k} \alpha_i x_{n+i}, \sum_{i=0}^{k} \beta_i x_{n+i}\rangle = \|X_{n+1}\|_G^2 - \|X_n\|_G^2 + \|\sum_{i=0}^{k} a_i x_{n+i}\|. \tag{3.22}$$

From (3.21) and (3.22) we infer the inequality $\|X_{n+1}\|_G^2 \leq \|X_n\|_G^2$. We have shown the following lemma.

Lemma 3.2.1. *Let ϕ be convex and let x_n be the discrete solution to*

$$\begin{cases} \dot{x}(t) \in -\partial\phi(x(t)), \\ \lim_{t\downarrow 0} x(t) = x_0, \end{cases}$$

generated by a G-stable one-leg scheme. Then the G-norm

$$\|X_n\|_G = \left(\sum_{i=1}^{k}\sum_{j=1}^{k} g_{ij}\langle x_{n+i-1}, x_{n+j-1}\rangle\right)^{\frac{1}{2}}$$

of the discrete solution decays monotonically, i.e. $\|X_{n+1}\|_G \leq \|X_n\|_G$.

3.3 Numerics

3.3.1 The H^{-1}-norm

So far we investigated minimizing movement schemes for gradient flows. What we are interested in is to apply these schemes to partial differential equations. To this end we have to find a gradient flow formulation of the equation in question. It would be convenient if the problem could be interpreted as a gradient flow on the most simple Hilbert space $L^2(\Omega)$. However, it turns out that the more interesting cases (i.e. non-linear partial differential equation) correspond to gradient flows on spaces equipped with a more complicated norm e.g. H^{-1}-norm. It is well known that the porous medium equation

$$\begin{cases} u_t = \Delta u^m, & x \in \Omega;\ t \in (0,T), \\ u(t,x) = 0, & x \in \partial\Omega;\ t \in [0,T], \\ u(0,x) = u_0, & x \in \Omega. \end{cases}$$

constitutes a gradient flow for the entropy

$$\phi(u) = \frac{1}{m+1} \int_\Omega u^{m+1}(\omega) d\omega,$$

with respect to the H^{-1}-norm. We sketch the derivation of the gradient flow formulation. Recall that the gradient of a functional is defined via the inner product. To this end we first sketch the derivation of the H^{-1} inner product. Let the inner product on $H_0^1(\Omega)$ be defined via

$$\langle f, g \rangle := \int_\Omega \nabla f(\omega) \nabla g(\omega) d\omega.$$

Let now $H^{-1}(\Omega)$ be the dual space to $H_0^1(\Omega)$. By Riesz's representation theorem we know that for $u \in H^{-1}(\Omega)$ there exists a $f \in H_0^1(\Omega)$, so that

$$u(g) = \langle u, g \rangle = \int_\Omega \nabla f(\omega) \nabla g(\omega) d\omega, \qquad \forall g \in H_0^1(\Omega).$$

In this sense we identify u and $-\Delta f$. We now define the H^{-1} inner product by:

$$\langle u, v \rangle_{H^{-1}} = \int_\Omega \nabla(\Delta^{-1} u(\omega)) \nabla(\Delta^{-1} v(\omega)) d\omega.$$

Recall that the gradient of a functional with respect to a given inner product is the co-vector $\nabla\phi$ so that

$$D\phi(u)(v) = \langle \nabla\phi(u), v \rangle, \qquad \forall v \in H.$$

Let $u, v \in H^{-1}(\Omega)$. An easy calculation shows that the variation of $\phi(u) = \frac{1}{m+1} \int_\Omega u^{m+1}(\omega) d\omega$ with respect to v is given by:

$$D\phi(u)(v) = \int_\Omega u^m(\omega) v(\omega) d\omega.$$

We now show that $\nabla\phi(u)$ with respect to $\langle\cdot,\cdot\rangle_{H^{-1}}$ is given by $-\Delta u^m$.

$$\begin{aligned}D\phi(u)(v) &= \int_\Omega u^m(\omega)v(\omega)d\omega = \int_\Omega u^m(\omega)\Delta\Delta^{-1}v(\omega)d\omega \\ &= -\int_\Omega \nabla u^m(\omega)\nabla\Delta^{-1}v(\omega)d\omega = -\int_\Omega \nabla\Delta^{-1}\underbrace{\Delta u^m(\omega)}_{=:\nabla\phi}\nabla\Delta^{-1}v(\omega)d\omega \\ &= \langle\nabla\phi(u), v\rangle_{H^{-1}}.\end{aligned}$$

Hence $-\Delta u^m$ is the gradient of $\phi(u)$ and we can write the porous medium equation as a gradient flow, i.e.:

$$u_t = -\nabla\phi(u) = \Delta u^m.$$

3.3.2 Discretization

Space discretization

For our numerical experiments we choose $\Omega := [0,1]$. We use a uniform gird Ω_N with:

$$\Omega_N := \left\{\omega_i := \frac{i}{N};\ \text{for } i = 1, ..., N\right\}.$$

We define the set of ansatz functions by $\mathcal{A}_N := \{\varphi_i : 1 \leq i \leq N-1\}$ where:

$$\begin{cases}\varphi_i(\omega) := \frac{\omega-\omega_{i-1}}{\omega_i-\omega_{i-1}}; & \omega\in[\omega_{i-1},\omega_i] \\ \varphi_i(\omega) := \frac{\omega_{i+1}-\omega}{\omega_{i+1}-\omega_i}; & \omega\in[\omega_i,\omega_{i+1}] \\ \varphi_i(\omega) := 0; & \text{else}\end{cases}$$

For the sake of convenience we define the shape functions:

$$\begin{cases}\varphi_i^r(\omega) := \varphi_i(\omega)|_{[\omega_i,\omega_{i+1}]} \\ \varphi_i^l(\omega) := \varphi_i(\omega)|_{[\omega_{i-i},\omega_i]}\end{cases}$$

Recall that by Definition 3.1.10 solving the gradient flow via a minimizing movement scheme is to find first

$$\tilde{u} := \underset{u\in H}{\operatorname{argmin}} \underbrace{\left\{\frac{\|u+\eta\|^2}{2\tau} + \frac{\beta_k}{\alpha_k}\phi(u)\right\}}_{\psi(u)}$$

and then to set

$$u_{n+k} = \frac{\tilde{u} - \sum_{i=0}^{k-1}\beta_i u_{n+i}}{\beta_k}.$$

To this end we calculate the discrete version of $\psi(u)$.

Discrete H^{-1}-norm

We split the functional ψ in its summands and first discretize the norm part. The discretized version $(u+\eta)_d$ of $u+\eta$ is given by:

$$(u+\eta)_d =: u_a = \sum_{i=1}^{N-1} a_i \varphi_i(\omega),$$

where $a_i = u(\omega_i) + \eta(\omega_i)$. Above we have shown that the H^{-1}-norm is given via:

$$\|u_a\|_{H_0^{-1}} = \int \nabla f \nabla f dx.$$

Where $u_a = -\Delta f$ with homogeneous Dirichlet boundary conditions. Hence to calculate the norm, we have to solve the elliptic problem

$$\begin{cases} u_a = -\Delta f \\ u_a(0) = u_a(1) = 0. \end{cases}$$

We assume f as well to be a linear combination of our ansatz functions, i.e. $f = \sum_{i=1}^{N-1} f_i \varphi_i(\omega)$. The problem in its weak form then reads to be

$$\int_0^1 \left(\sum_{i=1}^{N-1} a_i \varphi_i(\omega) \right) \varphi_j(\omega) d\omega = \int_0^1 \nabla \left(\sum_{i=1}^{N-1} f_i \varphi_i(\omega) \right) \nabla \varphi_j(\omega) d\omega, \qquad (3.23)$$

for every $\varphi_j, j \in \{1, ..., N-1\}$. Pick an arbitrary φ_j. As $\text{supp}(\varphi_i) = [\omega_{i-1}, \omega_{i+1}]$ and $\text{supp}(\varphi_j) = [\omega_{j-1}, \omega_{j+1}]$ we get:

$$\begin{aligned} \varphi_i \varphi_j &= 0, && \text{for } i \notin \{j-1, j, j+1\}, \\ \varphi_i \varphi_1 &= 0, && \text{for } i \notin \{1, 2\}, \\ \nabla\varphi_i \nabla\varphi_j &= 0, && \text{for } i \notin \{j-1, j, j+1\}, \\ \nabla\varphi_i \nabla\varphi_{N-1} &= 0, && \text{for } i \notin \{N-2, N-1\}. \end{aligned}$$

Hence the left hand side of (3.23) for φ_j reads to be:

$$a_{j-1} \underbrace{\int_{\omega_{j-1}}^{\omega_j} \varphi_{j-1}(\omega) \varphi_j(\omega) d\omega}_{A_{j-1,j}} + a_j \underbrace{\int_{\omega_{j-1}}^{\omega_{j+1}} \varphi_j(\omega) \varphi_j(\omega) d\omega}_{A_{j,j}} + a_{j+1} \underbrace{\int_{\omega_j}^{\omega_{j+1}} \varphi_{j+1}(\omega) \varphi_j(\omega) d\omega}_{A_{j+1,j}}.$$

Whereas the right-hand side becomes:

$$f_{j-1}\underbrace{\int_{\omega_{j-1}}^{\omega_j} \nabla\varphi_{j-1}(\omega)\nabla\varphi_j(\omega)d\omega}_{B_{j-1,j}} + f_j \underbrace{\int_{\omega_{i-j}}^{\omega_{j+1}} \nabla\varphi_j(\omega)\nabla\varphi_j(\omega)d\omega}_{B_{j,j}}$$
$$+ f_{j+1}\underbrace{\int_{\omega_i}^{\omega_{i+1}} \nabla\varphi_{j+1}(\omega)\nabla\varphi_j(\omega)d\omega}_{B_{j+1,j}}.$$

In matrix form the discrete problem then is given by:

$$B^{-1}Aa = f. \tag{3.24}$$

Where the coefficients of A and B are easy to be calculated. See Appendix B. By (3.24) we get the weights f_i. To finally get the norm, we have to calculate:

$$\begin{aligned}\int_0^1 \|\nabla \sum_{i=1}^{N-1} f_i\varphi_i(\omega)\|^2 d\omega &= \int_0^1 \|\sum_{i=1}^{N-1} f_i \nabla\varphi_i(\omega)\|^2 d\omega \\ &= \int_0^1 \sum_{j=1}^{N-1}\sum_{i=1}^{N-1} f_i f_j \nabla\varphi_i(\omega)\nabla\varphi_j(\omega)d\omega \\ &= \sum_{j=1}^{N-1}\sum_{i=1}^{N-1} f_i f_j \underbrace{\int_0^1 \nabla\varphi_i(\omega)\nabla\varphi_j(\omega)d\omega}_{C_{i,j}}.\end{aligned}$$

This is a quadratic form with respect to the matrix C. Comparing the $C_{i,j}$ with the $B_{i,j}$ we find that both coincide ($C = B$). So the norm of f is given by:

$$\|f\|_{H_0^1} = f^T B f$$

As by (3.24) $f = B^{-1}Aa$ we end up with

$$\begin{aligned}\|u_a\|_{H^{-1}} &= (B^{-1}Aa)^T BB^{-1}Aa = a^T A^T (B^{-1})^T BB^{-1}Aa = a^T A^T (B^{-1})^T Aa \\ &= a^T M a,\end{aligned}$$

a simple quadratic form in the weights a.

Discrete entropy

We now derive the discrete entropy. We call the discrete form of u:

$$u_b(\omega) := \sum_{i=1}^{N-1} u_i\varphi_i(\omega,)$$

where $u_i := u(\omega_i)$. We calculate the entropy in its discrete form for a general F:

$$
\begin{aligned}
\int_0^1 F(u_b)d\omega &= \int_0^1 F\left(\sum_{i=1}^{N-1} u_i\varphi_i(\omega)\right)d\omega \\
&= \sum_{j=0}^{N-1}\left(\int_{\omega_j}^{\omega_{j+1}} F\left(\sum_{i=1}^{N-1} u_i\varphi_i(\omega)\right)d\omega\right) \\
&= \int_{\omega_0}^{\omega_1} F(u_1\varphi_1(\omega))d\omega \\
&+ \sum_{j=1}^{N-2}\left(\int_{\omega_j}^{\omega_{j+1}} F\left(u_j\varphi_j(\omega) + u_{j+1}\varphi_{j+1}(\omega)\right)d\omega\right) \\
&+ \int_{\omega_{N-1}}^{\omega_N} F(u_{N-1}\varphi_{N-1}(\omega))d\omega \\
&= \underbrace{\int_{\omega_0}^{\omega_1} F\left(u_1\frac{\omega-\omega_0}{\omega_1-\omega_0}\right)d\omega}_{\Phi_1} \\
&+ \sum_{j=1}^{N-2}\left(\underbrace{\int_{\omega_j}^{\omega_{j+1}} F\left(u_j\frac{\omega_{j+1}-\omega}{\omega_{j+1}-\omega_j} + u_{j+1}\frac{\omega-\omega_j}{\omega_{j+1}-\omega_j}\right)d\omega}_{\Phi_{j+1}}\right) \\
&+ \underbrace{\int_{\omega_{N-1}}^{\omega_N} F\left(u_{N-1}\frac{\omega_N-\omega}{\omega_N-\omega_{N-1}}\right)d\omega}_{\Phi_N}.
\end{aligned}
$$

For the porous medium equation, i.e. $F(u) = \frac{1}{m+1}u^{m+1}$ we get the coefficients:

$$
\begin{aligned}
\Phi_1 &= \frac{u_1^{m+1}(\omega_1-\omega_0)}{(m+1)(m+2)}, \\
\Phi_{j+1} &= \frac{(u_{j+1}^{m+2}-u_j^{m+2})(\omega_{j+1}-\omega_j)}{(m+1)(m+2)(u_{j+1}-u_j)}, \\
\Phi_N &= \frac{u_{N-1}^{m+1}(\omega_N-\omega_{N-1})}{(m+1)(m+2)}.
\end{aligned}
$$

The calculations can be found in in the Appendix B.

Minimizing the discrete functional

Summing up the results above we are given the discrete functional $\Psi(u)$ by

$$\Psi = \frac{1}{2\tau} a^T M a + \frac{\beta_k}{\alpha_k} \frac{u_1^{m+1}(\omega_1 - \omega_0)}{(m+1)(m+2)} + \frac{\beta_k}{\alpha_k} \sum_{i=1}^{N-2} \frac{(u_{i+1}^{m+2} - u_i^{m+2})(\omega_{i+1} - \omega_i)}{(u_{i+1} - u_i)(m+1)(m+2)}$$
$$+ \frac{\beta_k}{\alpha_k} \frac{u_{N-1}^{m+1}(\omega_N - \omega_{N-1})}{(m+1)(m+2)}.$$

Recall that $a_i = u(\omega_i) + \eta(\omega_i)$. As Ψ is smooth we find the minimum by finding the zero of the gradient:

$$\nabla\Psi(u) = \frac{1}{\tau} M a + \nabla\Phi(u).$$

The entries of $\nabla\Phi$ for general mesh and one-leg scheme can be found in the Appendix B. For the BDF-2 scheme with uniform grid the entries are given by:

$$\frac{\partial\Phi}{\partial u_1} = \frac{2h}{3}\left(\frac{u_1^m}{m+2} - \frac{u_1^{m+1}}{(m+1)(u_2 - u_1)} + \frac{u_2^{m+2} - u_1^{m+2}}{(m+1)(m+2)(u_2 - u_1)^2}\right),$$
$$\frac{\partial\Phi}{\partial u_l} = \frac{2h}{3(m+1)}\left(-\frac{u_l^{m+1}}{u_{l+1} - u_l} + \frac{u_{l+1}^{m+2} - u_l^{m+2}}{(m+2)(u_{l+1} - u_l)^2} + \frac{u_l^{m+1}}{u_l - u_{l-1}}\right)$$
$$- \frac{2h}{3(m+1)}\left(\frac{u_l^{m+2} - u_{l-1}^{m+2}}{(m+2)(u_l - u_{l-1})^2}\right),$$
$$\frac{\partial\Phi}{\partial u_{N-1}} = \frac{2h}{3}\left(\frac{u_{N-1}^{m+1}}{(m+1)(u_{N-1} - u_{N-2})} - \frac{u_{N-1}^{m+2} - u_{N-2}^{m+2}}{(m+1)(m+2)(u_{N-1} - u_{N-2})^2}\right)$$
$$+ \frac{2h}{3}\left(\frac{u_{N-1}^m}{(m+2)}\right).$$

Here h is the mesh size. The gradient of Φ is not linear (except for $m = 1$). To find the zero of

$$\nabla\Psi = \frac{1}{\tau} M a + \nabla\Phi = 0$$

we use Newton's method. Therefore we need the Hessian of Ψ:

$$\text{Hess}(\Psi) = \frac{1}{\tau} M + \text{Hess}(\Phi).$$

Find again the coefficients for the general case in the Appendix B. For the BDF-2 scheme with uniform grid we get:

$$\begin{aligned}
\frac{\partial\Phi}{\partial u_1\partial u_1} &= \frac{2h}{3}\left(\frac{mu_1^{m-1}}{m+2} - \frac{u_1^m}{(u_2-u_1)} - \frac{u_1^{m+1}}{(m+1)(u_2-u_1)^2}\right)\\
&+ \frac{2h}{3}\left(-\frac{u_1^{m+1}}{(m+1)(u_2-u_1)^2}\frac{2(u_2^{m+2}-u_1^{m+2})}{(m+1)(m+2)(u_2-u_1)^3}\right),\\
\frac{\partial\Phi}{\partial u_2\partial u_1} &= \frac{2h}{3(m+1)}\left(\frac{(u_1^{m+1}+u_2^{m+1})}{(u_2-u_1)^2} - \frac{2(u_2^{m+2}-u_1^{m+2})}{(m+2)(u_2-u_1)^3}\right),\\
\frac{\partial\Phi}{\partial u_{l-1}\partial u_l} &= \frac{2h}{3(m+1)}\left(\frac{(u_l^{m+1}+u_{l-1}^{m+1})}{(u_l-u_{l-1})^2} - \frac{2(u_l^{m+2}-u_{l-1}^{m+2})}{(m+2)(u_l-u_{l-1})^3}\right),\\
\frac{\partial\Phi}{\partial u_{l+1}\partial u_l} &= \frac{2h}{3(m+1)}\left(\frac{(u_l^{m+1}+u_{l+1}^{m+1})}{(u_{l+1}-u_l)^2} - \frac{2(u_{l+1}^{m+2}-u_l^{m+2})}{(m+2)(u_{l+1}-u_l)^3}\right),\\
\frac{\partial\Phi}{\partial u_l\partial u_l} &= \frac{2h}{3}\left(-\frac{u_l^m}{(u_{l+1}-u_l)} - \frac{2u_l^{m+1}}{(m+1)(u_{l+1}-u_l)^2}\right),\\
&+ \frac{2h}{3}\left(\frac{u_l^m}{(u_l-u_{l-1})} - \frac{2u_l^{m+1}}{(m+1)(u_l-u_{l-1})^2}\right)\\
&+ \frac{2h}{3}\left(\frac{2(u_{l+1}^{m+2}-u_l^{m+2})}{(m^2+3m+2)(u_{l+1}-u_l)^3} + \frac{2(u_l^{m+2}-u_{l-1}^{m+2})}{(m^2+3m+2(u_l-u_{l-1})^3}\right)\\
\frac{\partial\Phi}{\partial u_{N-2}\partial u_{N-1}} &= \frac{2h}{3(m+1)}\left(\frac{(u_{N-1}^{m+1}+u_{N-2}^{m+1})}{(u_{N-1}-u_{N-2})^2} - \frac{2(u_{N-1}^{m+2}-u_{N-2}^{m+2})}{(m+2)(u_{N-1}-u_{N-2})^3}\right),\\
\frac{\partial\Phi}{\partial u_{N-1}\partial u_{N-1}} &= \frac{2h}{3}\left(\frac{u_{N-1}^m}{(u_{N-1}-u_{N-2})} - \frac{2u_{N-1}^{m+1}}{(m+1)(u_{N-1}-u_{N-2})^2}\right),\\
&+ \frac{2h}{3}\left(\frac{2(u_{N-1}^{m+2}-u_{N-2}^{m+2})}{(m^2+3m+2)(u_{N-1}-u_{N-2})^3} + \frac{mu_{N-1}^{m-1}}{m+2}\right).
\end{aligned}$$

3.3.3 Implementation

Let the solution u at the nth time step be given and let $u^{(0)} := u$. The iteration is as follows:

$$u^{(s+1)} := u^{(s)} + (\delta u)^{(s+1)},$$

where $(\delta u)^{(s+1)}$ is the solution to the linear system

$$H(u^{(s)})(\delta u)^{(s+1)} = -G((u^{(s)})),$$

where $H(u^{(s)})$ is $\mathrm{Hess}(\Psi(u^{(s)}))$ and $G(u^{(s)})$ is $\nabla\Psi((u^{(s)}))$. The iteration is stopped if the norm of $(u^{(s+1)})$ is smaller than a certain threshold (see Section 3.4 for details), We assume that the solution $u \geq 0$. If a weight in the Newton scheme is calculated to be negative $u(\omega_i) < 0$ we set the weight to zero $u(\omega_i) := 0$, because in this case the global minimum is out of the domain of our problem and hence the minimum within the domain is on the boundary (recall that the functional to be minimized in every time step is convex and hence the minimum is unique).

3.4 Numerical experiments

3.4.1 The porous medium equation for $m = 5/3$

In this section we want to present some numerical results for the porous medium equation with $m = 5/3$, i.e.

$$\begin{cases} u_t = \Delta u^{5/3}, & x \in [0,1];\ t \in (0,T), \\ u(t,0) = u(t,1) = 0, & t \in [0,T], \\ u(0,x) = u_0, & x \in [0,1]. \end{cases}$$

by employing the BDF-2 scheme with linear ansatz functions. This equation arises as a high-density limit for a semiconductor drift diffusion model. For more on the model see [35]. An exact solution to the porous medium equation is given by the well known Barenblatt solution:

$$B(t) := \frac{1}{t^\alpha}\left(\left(C - \frac{m-1}{2m}\beta\frac{|x|^2}{t^{2\beta}}\right)^+\right)^{\frac{1}{m-1}},$$

where

$$\alpha := \frac{n}{n(m-1)+2} \qquad \beta := \frac{1}{n(m-1)+2.}$$

Here n is the space dimension. For $n = 1$ this yields

$$B(t) = \frac{1}{\sqrt[8]{t^3}}\left(\left(C - \frac{3}{40}\frac{|x|^2}{\sqrt[4]{t^3}}\right)^+\right)^{3/2}.$$

Note that for $t < \frac{1}{\sqrt[3]{(\frac{160C}{3})^4}} \approx 1.04$ the Barenblatt profile corresponds to the homogeneous Dirichlet boundary conditions. Figures 3.1 and 3.2 below illustrate the temporal evolution of the numerical solution with starting profile $u(0,x) = B(1)$ and end time $T = 1$, i.e. $u(1,x) = B(2)$, where we set $C = 0.008$.

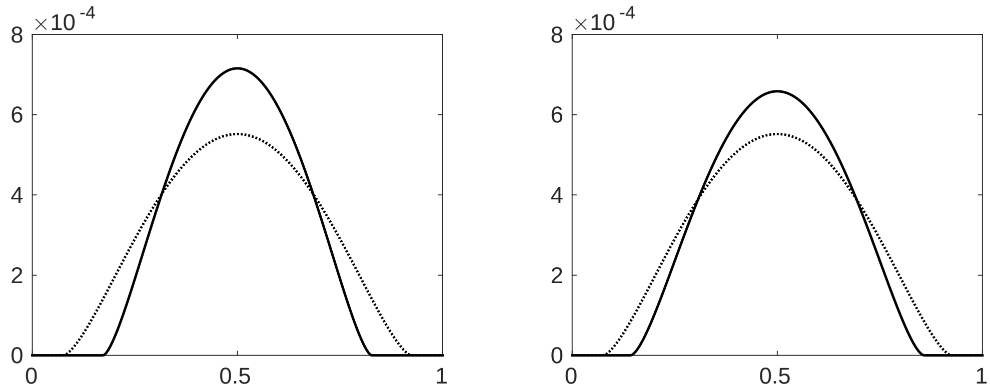

Figure 3.1: Evolution of the numerical solutions $u_n(0,x)$ left, and $u_n(0.25,x)$ right, compared with the exact solution $u(T,x) = B(2)$ (dotted line).

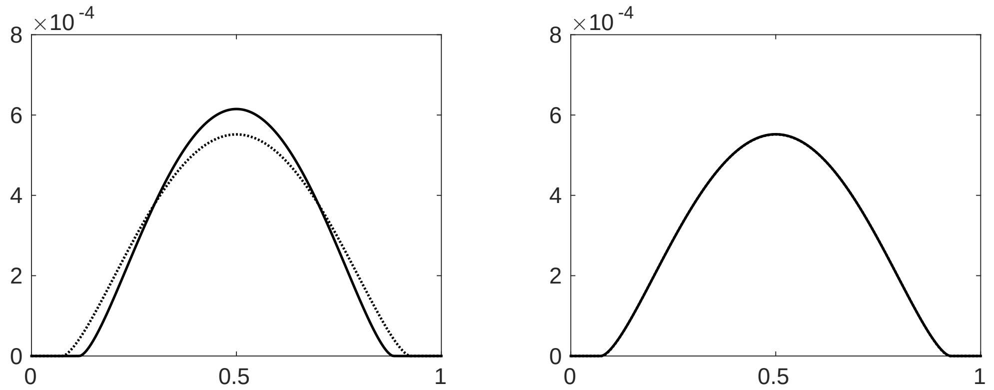

Figure 3.2: Evolution of the numerical solutions $u_n(0.5, x)$ left, and $u_n(1, x)$ compared with the exact $u(T, x) = B(2)$ (dotted line). On the right the dotted line is invisible as the numerical solution covers graphically the analytical solution $B(2)$.

Figure 3.3 shows the relative l^2-error of the scheme with respect to the time step size τ and the space mesh size h. We observe a second order decay in time as predicted and first order decay in space.

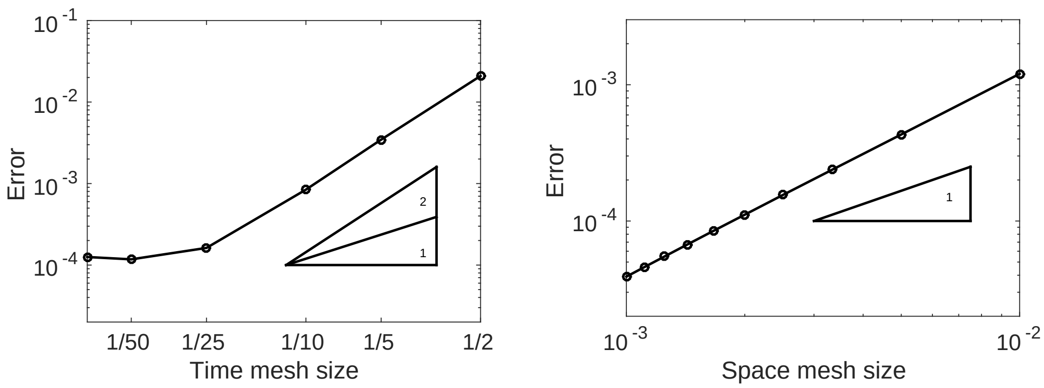

Figure 3.3: Left: Relative l^2-error with respect to τ. Due to the dominance of the space error under a certain threshold of τ the curve becomes flat. Right: Relative l^2-error with respect to h.

We observed the numerical evolution with a Barenblatt starting profile to verify the convergence order of the scheme. More interesting from the applied side is to do numerical simulations of an evolution starting with some initial $u(0, x)$ which can not be solved exactly. In Figure 3.4 on the left the evolution with initial condition

$$\begin{cases} u(0,x) = \max(\sin(2\pi x)^2 - \sin(2\pi 0.2)) & \text{for } x \in (0.2, 0.8), \\ u(0,x) = 0 & \text{else.} \end{cases}$$

The plot on the right-hand side shows the evolution with initial condition $u(0,x) = \sin(\pi x^4)$.

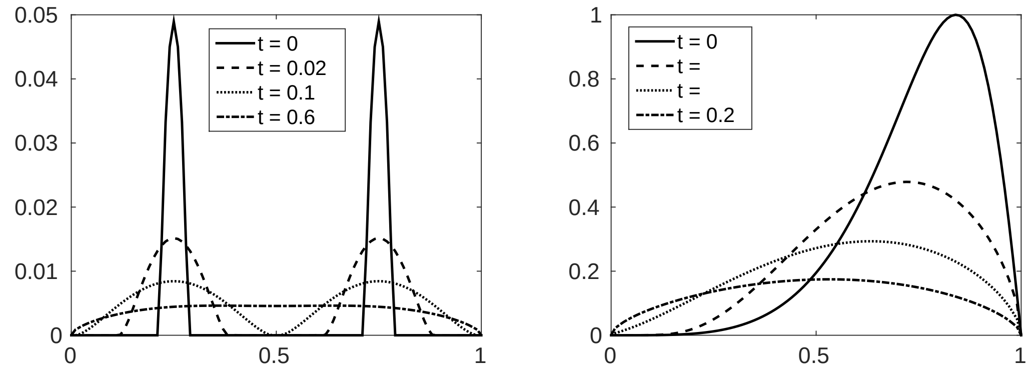

Figure 3.4: Simulation for symmetric and asymmetric initial conditions.

Recall that we showed that in general the entropy ϕ will not be dissipated by a higher order scheme. However this phenomenon seems to occur in more or less pathological situations. In the numerical analysis of the porous medium equation we found that the entropy is numerically decayed in any setting we studied. Figure 3.5 shows the decay of the entropy on the left and the decay of the G-norm on the right. The initial condition was $u(0,x) = \sin(\pi x^4)$, and end time was set to $T = 10$.

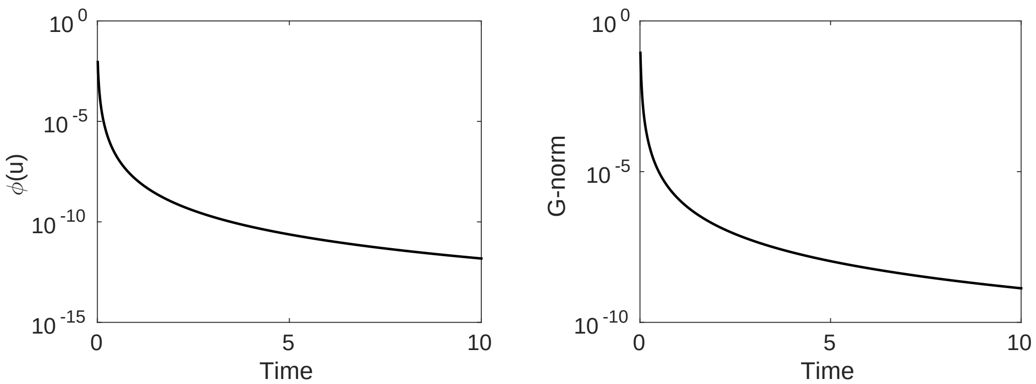

Figure 3.5: Decay of the entropy $\phi(u)$ compared with the decay of the G-norm.

3.4.2 Numerical violation of entropy decay

Finally we want to illustrate the violation of dissipation numerically for some functional $\phi : L^2([0,1]) \to \mathbb{R}$. Let $f \in L^2([0,1])$ and define the functional ϕ by:

$$\phi(f) := a \int_{f>0} f(x)dx - b \int_{f<0} f(x)dx; \qquad a, b > 0. \tag{3.25}$$

This entropy pushes f to 0, but with different speed depending on whether $f < 0$ or $f > 0$. It is a generalization of the 1-d problem we discussed in the first part. We first show that ϕ is convex.

Lemma 3.4.1. *Let $\phi : L^2([0,1]) \to \mathbb{R}$ be defined by*

$$\phi(f) := a \int_{f>0} f(x)dx - b \int_{f<0} f(x)dx$$

where $a, b \in \mathbb{R}_+$, then ϕ is convex.

Proof. We show that the convexity inequality holds ture. We start with

$$\phi(tf_1 + (1-t)f_2) = a \int_{tf_1>(t-1)f_2} tf_1 + (1-t)f_2 dx$$
$$- b \int_{tf_1<(t-1)f_2} tf_1 + (1-t)f_2 dx$$

We now split the integrals to get:

$$= ta \int_{tf_1>(t-1)f_2} f_1 dx + (1-t)a \int_{tf_1>(t-1)f_2} f_2 dx$$
$$- tb \int_{tf_1<(t-1)f_2} f_1 dx - (1-t)b \int_{tf_1<(t-1)f_2} f_2 dx.$$

Next we split the sets over which we take the integrals in the intersection of the set with the quadrants:

$$\begin{aligned}\{tf_1 > (t-1)f_2\} =& \{f_1, f_2 > 0, f_1 > \tfrac{t-1}{t} f_2\} \cup \{f_1 > 0, f_2 < 0, f_1 > \tfrac{t-1}{t} f_2\} \\ &\cup \{f_1 < 0, f_2 > 0, f_1 > \tfrac{t-1}{t} f_2\} \cup \underbrace{\{f_1, f_2 < 0, f_1 > \tfrac{t-1}{t} f_2\}}_{A} \\ \{tf_1 < (t-1)f_2\} =& \{f_1, f_2 < 0, f_1 < \tfrac{t-1}{t} f_2\} \cup \{f_1 > 0, f_2 < 0, f_1 < \tfrac{t-1}{t} f_2\} \\ &\cup \{f_1 < 0, f_2 > 0, f_1 < \tfrac{t-1}{t} f_2\} \cup \underbrace{\{f_1, f_2 > 0, f_1 < \tfrac{t-1}{t} f_2\}}_{B}\end{aligned}$$

As $t < 1$ the sets A and B are empty. We now split the integrals with respect to to the remaining sets:

$$
\begin{aligned}
&= ta \int_{f_1,f_2>0} f_1 dx + ta \int_{f_1>0,f_2<0,f_1>\frac{t-1}{t}f_2} f_1 dx + \underbrace{ta \int_{f_1<0,f_2>0,f_1>\frac{t-1}{t}f_2} f_1 dx}_{<0} \\
&+ (1-t)a \int_{f_1,f_2>0} f_2 dx + (1-t)a \int_{f_1<0,f_2>0;f_1>\frac{t-1}{t}f_2} f_2 dx \\
&+ \underbrace{(1-t)a \int_{f_1>0,f_2<0,f_1>\frac{t-1}{t}f_2} f_2 dx}_{<0} - tb \int_{f_1,f_2<0} f_1 dx - tb \int_{f_1<0,f_2>0;f_1<\frac{t-1}{t}f_2} f_1 dx \\
&\underbrace{- tb \int_{f_1>0,f_2>0 f_1<\frac{t-1}{t}f_2} f_1 dx}_{>0} - (1-t)b \int_{f_1,f_2<0} f_2 dx \\
&- (1-t)b \int_{f_1>0,f_2<0;f_1<\frac{t-1}{t}f_2} f_2 dx - \underbrace{(1-t)b \int_{f_1<0,f_2>0;f_1>\frac{t-1}{t}f_2} f_2 dx}_{>0}
\end{aligned}
$$

We drop the negative terms and arrive at:

$$
\begin{aligned}
&\leq ta \int_{f_1,f_2>0} f_1 dx + ta \int_{f_1>0,f_2<0;f_1>\frac{t-1}{t}f_2} f_1 dx + (1-t)a \int_{f_1,f_2>0} f_2 dx \\
&+ (1-t)a \int_{f_1<0,f_2>0;f_1>\frac{t-1}{t}f_2} f_2 dx - tb \int_{f_1,f_2<0} f_1 dx - tb \int_{f_1<0,f_2>0;f_1<\frac{t-1}{t}f_2} f_1 dx \\
&- (1-t)b \int_{f_1,f_2<0} f_2 dx - (1-t)b \int_{f_1>0,f_2<0;f_1<\frac{t-1}{t}f_2} f_2 dx.
\end{aligned}
$$

The sum above already looks familiar to $t\phi(f_1) + (1-t)\phi(f_2)$. However the domains of integration should be of the from $f_i > 0 (f_i < 0$ respectively). It is easy to see, that if we change the domains the sum only can get larger. (replace for example $\{f_1, f_2 > 0\}$ by$\{f_1 > 0\}$). We arrive at:

$$
\begin{aligned}
&\leq ta \int_{f_1>0} f_1 dx + (1-t)a \int_{f_2>0} f_2 dx - tb \int_{f_1<0} f_1 dx - (1-t)b \int_{f_2<0} f_2 dx \\
&= t \left(a \int_{f_1>0} f_1 dx - b \int_{f_1<0} f_1 dx \right) + (1-t) \left(a \int_{f_2>0} f_2 dx - b \int_{f_2<0} f_2 dx \right) \\
&= t\phi(f_1) + (1-t)\phi(f_2).
\end{aligned}
$$

□

Space discretization

We again use an equidistant mesh $\Omega_N := \{\omega_0, ..., \omega_N; \omega_i = \frac{2\pi i}{N}\}$ but for simplicity we use a set of piecewise constant ansatz functions $\mathcal{A} := \{\varphi_i; 1 \leq i \leq N\}$ where

$\varphi_i := 1_{[\omega_{i-1},\omega_i)}$, with equidistant grid, i.e. $\omega_{i+1} - \omega_i = h$. The discrete function is then given by $\sum_{i=1}^N u^i \varphi_i$. where $u^i := f(\frac{\omega_{i-1}+\omega_i}{2})$. The calculation of the L^2-norm yields

$$\|u\|^2 = \|\sum_{i=1}^N u^i \varphi_i\|^2 = \int_0^1 \left(\sum_{i=1}^N u^i \varphi_i\right)^2 d\omega = \sum_{i=1}^N (u^i)^2 \int_{\omega_{i-1}}^{\omega_i} \varphi^2 d\omega = h \sum_{i=1}^N (u^i)^2.$$

We calculate the entropy to be:

$$\phi(u) = a \int_{\omega_0}^{\omega_N} \sum_{u^i>0} u^i \varphi_i d\omega - b \int_{\omega_0}^{\omega_N} \sum_{u^i<0} u^i \varphi_i d\omega = ah \sum_{u^i>0} u^i - bh \sum_{u^i<0} u^i.$$

We again employ the BDF-2 scheme to solve the problem. Therefore the functional to be minimized is given by

$$\Psi(u) := \frac{\|u+\eta\|^2}{2} + \frac{2}{3}\tau\phi(u) = h\left(\sum_{i=1}^N (u^i+\eta^i)^2 + \frac{2}{3}\tau\left(a\sum_{u^i>0} u^i - b\sum_{u^i<0} u^i\right)\right)$$

$$= h\left(\sum_{i=1}^N \left((u^i+\eta^i)^2 + \frac{2}{3}\tau \max\{au_i, -bu_i\}\right)\right). \tag{3.26}$$

Where $\eta := \frac{1}{3}u_n - \frac{4}{3}u_{n-1}$. It is clear that the minimum in (3.26) is minimal if every single summand is minimal. Therefore the problem is to solve $N-1$ independent minimization problems exactly like in the counter example in section 3.1. We chose $a = 12, b = 1$ (this is the limit relation derived in Remark 10) $dt = 1/100, T = 0.5, N = 100$. The initial condition is defined by $u(0,x) = \sin(x)$, for $x \in (0, 2\pi)$.

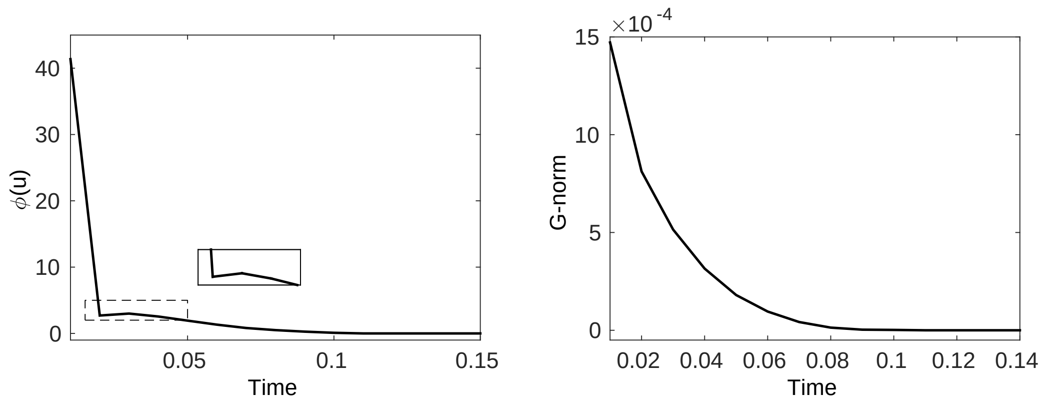

Figure 3.6: Illustriation of violation of dissipation of the entropy in contrast to the dissipation of the G-norm.

Chapter 4

Higher order Wasserstein gradient flow schemes in one space dimension

In this chapter we present a second order in time and space discretization in one space dimension for a non-linear diffusion equation interpreted as a Wasserstein gradient flow.

The Chapter is organized as follows. The global existence result and the exponential decay of the solutions to the constant steady state as well as some basic properties of the Wasserstein distance for periodic functions are stated in Section 4.1. Section 4.2 is devoted to the description of the numerical scheme, and some numerical experiments are presented in Section 4.3. The appendix contains the calculations of the coefficients of the matrix M_w and the Hessian of the discrete entropy.

4.1 Prerequisites

4.1.1 Existence of solutions and large-time asymptotics

Let us briefly review the literature for the diffusion equation

$$\partial_t u = \alpha^{-1}\Delta(u^\alpha) \quad \text{in } \mathscr{T}^d,\ t>0, \quad u(0)=u^0, \tag{4.1}$$

with $\alpha < 0$. Equation (4.1) with $\alpha = -1$ appears in the modeling of heat conduction in solid Helium, where the solution u corresponds to the inverse temperature [56]. When this equation is considered in the whole space or in a bounded domain with homogeneous Dirichlet boundary conditions, it is sometimes called the super-fast diffusion equation [60, Chap. 9]. The critical exponent of this equation in one space dimension is $\alpha = -1$. For $\alpha > -1$ and $u^0 \in L^1(\mathbb{R})$, we have a smoothing property, namely $u(t) \in L^\infty(\mathbb{R})$ for any $t > 0$ [60, Section 9.1]. For

$\alpha \leq -1$, no solutions exist with data in $L^1(\mathbb{R})$. The non-existence range in dimensions $d \geq 2$ contains even all negative exponents, $\alpha < 0$ [59]. However, if $d = 1$ and $\alpha \leq -1$, there is a weak smoothing effect. Indeed, given $u^0 \in L^p_{\text{loc}}(\mathbb{R})$ for $p = (1-\alpha)/2$, the solution u exists and is locally bounded in $\mathbb{R} \times (0, \infty)$. Furthermore, if $u^0 \in L^p(\mathbb{R})$ then there is instantaneous extinction, i.e. $u(t) = 0$ in $\mathbb{R}$ for all $t > 0$ [60, Theorem 9.3],

Clearly, such results cannot be expected when the super-fast diffusion equation is considered on the torus. We expect that global-in-time weak solutions exist, which converge to the constant steady state as $t \to \infty$. Since we could not find any results on the existence and large-time asymptotics in the literature in that situation and since the use of negative exponents is less standard, we provide a (short) proof for completeness.

Equation (4.1) on the torus $\mathscr{T}^d$ does not possess the non-existence or instantaneous extinction properties of the super-fast diffusion equation in the whole space since mass cannot get lost. In fact, we expect that for any $\alpha < 0$, there exists a global weak solution. If the initial datum u^0 is nonnegative only, (4.1) is still a singular diffusion equation. However, because of the fast diffusion, the solution becomes positive for all positive times and, by parabolic regularity theory, also smooth.

Theorem 4.1.1 (Existence of weak solutions)**.** *Let $\alpha < 0$ and let $u^0 \in L^\infty(\mathscr{T}^d)$ satisfy $u^0 \geq 0$. If $\alpha = -1$, we assume additionally that $\int_{\mathscr{T}^d} \log u^0 dx > -\infty$. Then there exists a unique weak solution to (4.1) satisfying $u^\alpha \in L^2(0, T; H^1(\mathscr{T}^d))$, $\partial_t u \in L^2(0, T; H^1(\mathscr{T}^d)')$ for all $T > 0$, and $0 \leq u(x,t) \leq \sup_{\mathscr{T}^d} u^0$ for $x \in \mathscr{T}^d$, $t \geq 0$.*

The proof of this theorem is based on a standard regularization procedure but we need to distinguish carefully the cases $-1 < \alpha < 0$, $\alpha = -1$, and $\alpha < -1$.

Proof. The existence proof is based on a regularization of the initial datum and a fixed-point argument. Let $0 < \varepsilon < 1$ and $u_{0,\varepsilon} = u^0 + \varepsilon$. Let $Q_T = \mathscr{T}^d \times (0, T)$ and $M = \sup_{\mathscr{T}^d} u^0$, Set

$$K = \{u \in L^2(Q_T) : \varepsilon \leq u \leq M,\ \|u\|_{L^2(0,T;H^1(\mathscr{T}^d))} + \|\partial_t u\|_{L^2(0,T;H^1(\mathscr{T}^d)')} \leq C\},$$

where $C > 0$ will be determined later. The set K is convex and, by Aubin's lemma, compact in $L^2(Q_T)$. Let $v \in K$ and let $u \in L^2(Q_T)$ be the weak solution to

$$\partial_t u = \operatorname{div}(v^{\alpha-1}\nabla u) \quad \text{in } \mathscr{T}^d,\ t > 0, \quad u(0) = u_{0,\varepsilon}. \tag{4.2}$$

This defines the fixed-point operator $Z : K \to L^2(Q_T)$, $v \mapsto u$. Standard arguments show that Z is continuous. We verify that $Z(K) \subset K$. By the maximum principle, $\varepsilon \leq u \leq M$. Using u as a test function in the weak formulation of (4.2) shows that $\|u\|_{L^2(0,T;H^1(\mathscr{T}^d))} \leq C_1(\varepsilon)$, where $C_1(\varepsilon) > 0$ is some constant depending on ε. Moreover, $\|\partial_t u\|_{L^2(0,T;H^1(\mathscr{T}^d)')} \leq \|v^{\alpha-1}\nabla u\|_{L^2(Q_T)} \leq C_2(\varepsilon)$. Thus, setting, $C := C_1(\varepsilon) + C_2(\varepsilon)$, we infer that $u \in K$. By the fixed-point theorem of Schauder, there exists a fixed point u_ε of Z.

In order to perform the limit $\varepsilon \to 0$, we need to derive ε-independent estimates for u_ε. To this end, we need to distinguish several cases. First, let $\alpha = -1$. Employing the test function $1 - 1/u_\varepsilon$ in the weak formulation of (4.2) with $u = v = u_\varepsilon$, we find that

$$\int_{\mathcal{T}^d} (u_\varepsilon(t) - \log u_\varepsilon(t))\, dx + \int_0^t \int_{\mathcal{T}^d} |\nabla u_\varepsilon^{-1}|^2\, dxds = \int_{\mathcal{T}^d} (u^0 + \varepsilon - \log(u^0 + \varepsilon))\, dx.$$

The right side is uniformly bounded as we assumed that $-\int_{\mathcal{T}^d} \log u^0\, dx < \infty$. Since $|\nabla u_\varepsilon^{-1}|^2 \geq M^{-4}|\nabla u_\varepsilon|^2$, we infer uniform estimates for u_ε in $L^2(0,T;H^1(\mathcal{T}^d))$ and also in $H^1(0,T;H^1(\mathcal{T}^d)')$.

Next, let $\alpha \neq -1$. The test function u_ε^α in the weak formulation of (4.2) gives

$$\frac{1}{\alpha+1} \int_{\mathcal{T}^d} u_\varepsilon(t)^{\alpha+1}\, dx + \frac{1}{\alpha} \int_0^t \int_{\mathcal{T}^d} |\nabla u_\varepsilon^\alpha|^2\, dxds = \frac{1}{\alpha+1} \int_{\mathcal{T}^d} (u^0 + \varepsilon)^{\alpha+1}\, dx,$$

If $-1 < \alpha < 0$, we write this equation as

$$\begin{aligned} \int_0^t \int_{\mathcal{T}^d} |\nabla u_\varepsilon^\alpha|^2\, dxds &= -\frac{\alpha}{\alpha+1} \int_{\mathcal{T}^d} u_\varepsilon(t)^{\alpha+1}\, dx + \frac{\alpha}{\alpha+1} \int_{\mathcal{T}^d} (u^0 + \varepsilon)^{\alpha+1}\, dx \\ &\leq -\frac{\alpha}{\alpha+1} \int_{\mathcal{T}^d} M^{\alpha+1}\, dx. \end{aligned}$$

If $\alpha < -1$, we obtain

$$\frac{1}{-\alpha-1} \int_{\mathcal{T}^d} u_\varepsilon(t)^{\alpha+1}\, dx + \frac{1}{-\alpha} \int_0^t \int_{\mathcal{T}^d} |\nabla u_\varepsilon^\alpha|^2\, dxds = \frac{1}{-\alpha-1} \int_{\mathcal{T}^d} (u^0 + \varepsilon)^{\alpha+1}\, dx.$$

In both cases, since u^0 is assumed to be bounded, we infer a uniform bound for u_ε^α in $L^2(0,T;H^1(\mathcal{T}^d))$ and consequently also for $\partial_t u_\varepsilon$ in $L^2(0,T;H^1(\mathcal{T}^d)')$. Moreover, in view of $|\nabla u_\varepsilon^\alpha|^2 = \alpha^2 u_\varepsilon^{2(\alpha-1)}|\nabla u_\varepsilon|^2 \geq \alpha^2 M^{2(\alpha-1)}|\nabla u_\varepsilon|^2$, it follows that (u_ε) is bounded in $L^2(0,T;H^1(\mathcal{T}^d))$.

We infer for all $\alpha < 0$ the following bounds:

$$\begin{aligned} &\|u_\varepsilon\|_{L^\infty(0,T;L^\infty(\mathcal{T}^d))} + \|u_\varepsilon^\alpha\|_{L^2(0,T;H^1(\mathcal{T}^d))} + \|u_\varepsilon\|_{L^2(0,T;H^1(\mathcal{T}^d))} + \|\partial_t u_\varepsilon\|_{L^2(0,T;H^1(\mathcal{T}^d)')} \\ &\quad \leq C_3. \end{aligned}$$

By Aubin's lemma, there exists a subsequence which is not relabeled such that $u_\varepsilon \to u$ strongly in $L^2(Q_T)$. Moreover, $u_\varepsilon^\alpha \rightharpoonup u^\alpha$ weakly in $L^2(0,T;H^1(\mathcal{T}^d))$ and $\partial_t u_\varepsilon \rightharpoonup \partial_t u$ weakly in $L^2(0,T;H^1(\mathcal{T}^d)')$. Thus, we may pass to the limit $\varepsilon \to 0$ in the weak formulation which shows that u solves (4.1). □

For $t \to \infty$, the (smooth) solution converges to the constant steady state. Since this constant is positive, diffusion slows down when time increases. Therefore, we cannot expect instantaneous extinction phenomena. Still, we are able to prove that the convergence is exponentially fast with respect to the L^1-norm. We introduce the entropy

$$H_\beta[u] = \int_{\mathscr{T}^d} u^\beta \, dx - \left(\int_{\mathscr{T}^d} u \, dx \right)^\beta, \quad \beta > 1.$$

The steady state of (4.1) is given by $u_\infty = \int_{\mathscr{T}^d} u^0 \, dx$ if $\mathrm{vol}(\mathscr{T}^d) = 1$.

Theorem 4.1.2 (Exponential decay)**.** *Let u be a smooth positive solution to* (4.1) *and let* $\mathrm{vol}(\mathscr{T}^d) = 1$*. Then*

$$\|u(t) - u_\infty\|_{L^1(\mathscr{T}^d)} \le C_\beta H_\beta[u^0]^{1/\beta} \|u^0\|_{L^1(\mathscr{T}^d)}^{1/2} e^{-\lambda t},$$

where $C_\beta > 0$ and for $\alpha < 0$, $1 < \beta \le 2$,

$$\lambda = \frac{2(\beta - 1)}{\beta C_B} (\sup_{\mathscr{T}^d} u^0)^{\alpha - 1},$$

and $C_B > 0$ is the constant in the Beckner inequality (4.3)*; for $-1 \le \alpha < 0$, $\beta = 2(1 - \alpha)$,*

$$\lambda = \frac{2(1 - 2\alpha)}{(1 - \alpha)\|u^0\|_{L^1(\mathscr{T}^d)}^{1-\alpha}}.$$

In the first result, the decay rate λ depends on $\sup_{\mathscr{T}^d} u^0$, which seems to be not optimal. The decay rate in the second result depends on the L^1-norm of u^0 only but we need a particular value of β. The proof is based on the entropy method. Stronger decay results have been derived for the fast-diffusion equation in the whole space or in bounded domains; see, e.g., [13, 60]. However, our proof is very elementary and just an illustration for the qualitative behavior of the solutions to (4.1).

Proof. Employing (4.1) and integration by parts, we find that

$$\frac{dH_\beta}{dt} = -\beta(\beta - 1) \int_{\mathscr{T}^d} u^{\alpha + \beta - 3} |\nabla u|^2 \, dx = -\frac{4}{\beta}(\beta - 1) \int_{\mathscr{T}^d} u^{\alpha - 1} |\nabla u^{\beta/2}|^2 \, dx.$$

We employ the bound $u \le M = \sup_{\mathscr{T}^d} u^0$ and the Beckner inequality [11] (note that we assumed that $\mathrm{vol}(\mathscr{T}^d) = 1$),

$$\int_{\mathscr{T}^d} u^\beta \, dx - \left(\int_{\mathscr{T}^d} u \, dx \right)^\beta \le C_B \int_{\mathscr{T}^d} |\nabla u^{\beta/2}|^2 \, dx \quad \text{for } u^{\beta/2} \in H^1(\Omega), \ 1 < \beta \le 2, \tag{4.3}$$

to obtain

$$\frac{dH_\beta}{dt} \le -\frac{4(\beta - 1)}{\beta C_B} M^{\alpha - 1} \left(\int_{\mathscr{T}^d} u^\beta \, dx - \left(\int_{\mathscr{T}^d} u \, dx \right)^\beta \right) = -\frac{4(\beta - 1)}{\beta C_B} M^{\alpha - 1} H_\beta.$$

Then by Gronwall's lemma $H[u(t)] \leq H[u^0]e^{-\lambda t}$ with $\lambda = -4(\beta-1)M^{\alpha-1}/(\beta C_B)$ for $1 < \beta \leq 2$.

For the second result, let $-1 \leq \alpha < 0$ and $\beta = 2(1-\alpha)$. Similarly as above, we find that

$$\begin{aligned}\frac{dH_\beta}{dt} &= -\frac{4\beta(\beta-1)}{(\alpha+\beta-1)^2}\int_{\mathscr{T}^d} |\nabla u^{(\alpha+\beta-1)/2}|^2\,dx \\ &\leq -\frac{8(1-2\alpha)}{1-\alpha}\int_{\mathscr{T}^d}\left(\int_{\mathscr{T}^d} u^{\alpha+\beta-1}\,dx - \left(\int_{\mathscr{T}^d} u\,dx\right)^{\alpha+\beta-1}\right) \\ &= -\frac{8(1-2\alpha)}{1-\alpha}H_{\beta/2}[u],\end{aligned}$$

since $\alpha+\beta-1 = \beta/2 \in (1,2]$. Using the inequalities $\|u^{\beta/2}\|_{L^2(\mathscr{T}^d)} \geq \|u^{\beta/2}\|_{L^1(\mathscr{T}^d)}$ and $\|u^{\beta/2}\|_{L^2(\mathscr{T}^d)} = \|u\|_{L^\beta(\mathscr{T}^d)}^{\beta/2} \geq \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}$ (again we employ $\mathrm{vol}(\mathscr{T}^d) = 1$ here), it follows that

$$\begin{aligned}H_\beta[u] &= \|u^{\beta/2}\|_{L^2(\mathscr{T}^d)}^2 - \|u\|_{L^1(\mathscr{T}^d)}^\beta \\ &= \big(\|u^{\beta/2}\|_{L^2(\mathscr{T}^d)} + \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}\big)\big(\|u^{\beta/2}\|_{L^2(\mathscr{T}^d)} - \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}\big) \\ &\geq \big(\|u^{\beta/2}\|_{L^2(\mathscr{T}^d)} + \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}\big)\big(\|u^{\beta/2}\|_{L^1(\mathscr{T}^d)} - \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}\big) \\ &\geq \big(\|u^{\beta/2}\|_{L^1(\mathscr{T}^d)} + \|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}\big)H_{\beta/2}[u] \geq 2\|u\|_{L^1(\mathscr{T}^d)}^{\beta/2}H_{\beta/2}[u].\end{aligned}$$

Since the solution to (4.1) conserves mass, $\|u(t)\|_{L^1(\mathscr{T}^d)} = \|u^0\|_{L^1(\mathscr{T}^d)}$, and we end up with

$$\frac{dH_\beta}{dt} \leq -\frac{4(1-2\alpha)}{(1-\alpha)\|u^0\|_{L^1(\mathscr{T}^d)}^{\beta/2}}H_\beta[u].$$

Then Gronwall's lemma shows that $H_\beta[u(t)] \leq H_\beta[u^0]e^{-\lambda t}$ with $\lambda = 4(1-2\alpha)/((1-\alpha)\|u^0\|_{L^1(\mathscr{T}^d)}^{1-\alpha})$.

Finally, the statement of the theorem follows after applying the generalized Csiszár-Kullback inequality in the form

$$\|u-v\|_{L^1(\mathscr{T}^d)}^2 \leq C_\beta\|v\|_{L^1(\mathscr{T}^d)}\left(\int_{\mathscr{T}^d} u^\beta - \left(\int_{\mathscr{T}^d} u\,dx\right)^\beta\right), \quad 1 < \beta \leq 2.$$

for functions $u, v \in L^\beta(\mathscr{T}^d)$ such that $\int_{\mathscr{T}^d} u\,dx = \int_{\mathscr{T}^d} v\,dx$. The proof is a slight generalization of the proof of Theorem 1.4 in [32] taking $\varphi(t) = t^\beta$. □

4.1.2 The Wasserstein distance for periodic functions

In the introduction 1.2 we defined the Wasserstein metric for general metric spaces. On $\mathbb{R}$, there exists an explicit formula to compute W. Let $X = (a,b) \subset \mathbb{R}$

be a (possibly infinite) interval, and let $\mu_1,\ \mu_2 \in \mathscr{P}(X)$ be two measures. We define their distribution functions

$$U_i : (a,b) \to [0,M], \quad U_i(x) = \mu_i((a,x]), \quad i = 1,2.$$

As these functions are right-continuous and monotonically increasing, they possess right-continuous increasing pseudo-inverse functions $G_i : [0,M] \to [a,b]$, given by

$$G_i(\omega) = \inf\{x \in (a,b) : U_i(x) > \omega\}, \quad i = 1,2.$$

Then [61],

$$W[u_1,u_2]^2 = \int_0^M (G_1(\omega) - G_2(\omega))^2 d\omega. \tag{4.4}$$

This formula does not extend to $X = \mathscr{T} \simeq (0,1)$ because of the topology induced by the periodic boundary conditions, $d(x,y) = \min\{|x-y|, 1-|x-y|\}$. The reason is that mass can be transported either clock- or counter-clockwise (see [22] for details). However, if the densities u_1 and u_2 are point-symmetric, (4.4) still holds. More precisely, let $u_i(x) = u_i(1-x)$ for $x \in (0,1)$ and $i = 1,2$. Then (4.4) holds, where $G_i : [0,M] \to [0,1]$ is the inverse function of $U_i(x) = \int_0^x u_i(y)dy$ [22, Lemma 2.2].

4.2 Time discretization and Lagrangian coordinates

4.2.1 The semi-discrete BDF scheme

We introduce the second-order minimizing movement scheme. First, we explain the underlying idea for the finite-dimensional gradient flow

$$\dot{x} = -\nabla\phi(x), \quad t > 0, \quad x(0) = x_0, \tag{4.5}$$

where $\phi : \mathbb{R}^d \to \mathbb{R}$ is a smooth potential. This equation can be approximated by the following minimization problem:

$$x^{n+1} = \mathrm{argmin}_{x\in\mathbb{R}^d}\, \Phi(x), \quad \Phi(x) = \frac{1}{2\tau}\|x - x^n\|^2 + \phi(x),$$

where $\tau > 0$ is the time step size and x^n is an approximation of $x(n\tau)$. The minimizer x^{n+1} is a critical point and thus,

$$0 = \nabla\Phi(x^{n+1}) = \frac{1}{\tau}(x^{n+1} - x^n) + \nabla\phi(x^{n+1}),$$

which corresponds to the implicit Euler scheme.

Instead of the Euler scheme, we wish to discretize (4.1) by a multistep method. As an example, consider the two-step BDF (or BDF-2) method,

$$\frac{1}{\tau}\left(\frac{3}{2}x^{n+2} - 2x^{n+1} + \frac{1}{2}x^n\right) = -\nabla\phi(x^{n+2}),$$

where x^n and x^{n+1} are given. Writing this scheme as

$$\frac{1}{2}(x^n - x^{n+2}) - 2(x^{n+1} - x^{n+2}) = -\tau\nabla\phi(x^{n+1}),$$

we see that x^{n+2} is a critical point of the functional

$$\Phi(x) = -\frac{1}{4\tau}\|x^n - x\|^2 + \frac{1}{\tau}\|x^{n+1} - x\|^2 + \phi(x).$$

More generally, the BDF-k approximation of (4.5),

$$\sum_{i=0}^{k} a_i x^{n+i} = -\tau\nabla\phi(x^{n+k}),$$

for given $x^n, \dots, x^{n+k-1}$ can be formulated as

$$-\tau\nabla\phi(x^{n+k}) = \sum_{i=0}^{k-1} a_i x^{n+i} + a_k x^{n+k} = \sum_{i=0}^{k-1} a_i(x^{n+i} - x^{n+k}),$$

since $\sum_{i=0}^{k} a_i = 0$, or as the minimization problem

$$x^{n+k} = \operatorname{argmin}_{x\in\mathbb{R}^d} \Phi(x), \quad \Phi(x) = -\frac{1}{2\tau}\sum_{i=0}^{k-1} a_i\|x^{n+i} - x\|^2 + \phi(x), \tag{4.6}$$

In a similar way, we may formulate general multistep methods. We recall that BDF-k schemes are consistent if $\sum_{i=0}^{k} a_i = 0$, $\sum_{i=1}^{k} ia_i = 1$ and zero-stable if and only if $k \le 5$ [54, Section 11.5].

The same idea as above is applicable to gradient flows in the Wasserstein distance. For this, we replace the L^2-norm by the Wasserstein distance. For equation (4.1), scheme (4.6) turns into

$$u^{n+k} = \operatorname{argmin}_{u\in\mathscr{P}(\mathscr{T})} \Phi(u), \quad \Phi(u) = -\frac{1}{2\tau}\sum_{i=0}^{k-1} a_i W[u^{n+i}, u]^2 + S[u], \tag{4.7}$$

where $S[u] = (\alpha(\alpha-1))^{-1}\int_{\mathscr{T}} u^\alpha\,dx$. Scheme (4.7) can be interpreted as a BDF-k minimizing movement scheme.

4.2.2 Lagrangian coordinates

Before introducing the spatial discretization, we rewrite the scheme (4.7) in a more explicit manner, using the inverse distribution functions G and G^* of u and u^*, respectively, which were introduced in Section 4.1.2. The numerical procedure is similar to that in [22] but our higher-order scheme introduces some changes. We call $\omega = U(x) \in [0, M]$ the Lagrangian coordinate, which is conjugate to the Eulerian coordinate $x \in \mathscr{T}$, and we refer to the inverse distribution function G as the associated Lagrangian map. For a consistent change of variables, we need to express the entropy $S[u]$ in terms of the Lagrangian coordinates. With the formula for the differential of an inverse function,

$$u(x) = \partial_x U(x) = \frac{1}{\partial_\omega G(\omega)},$$

and the change of unknowns $x = G(\omega)$ under the integral in $S[u]$, we obtain

$$S[u] = \frac{1}{\alpha(\alpha-1)} \int_{\mathscr{T}} \frac{u(x)\,dx}{u(x)^{1-\alpha}} = \frac{1}{\alpha(\alpha-1)} \int_{\mathscr{T}} g(\omega)^{1-\alpha}\,d\omega,$$

where $g(\omega) = \partial_\omega G(\omega)$. Note that the exponent in the integrand is positive since $\alpha < 0$. In terms of g, the expression for the Wasserstein distance in (4.4) becomes (see [22, Section 2.3])

$$\begin{aligned} W[u,u^*]^2 &= \int_0^M (G(\omega) - G^*(\omega))^2\,d\omega = \int_0^M \left(\int_0^\omega (g(\eta) - g^*(\eta))d\eta \right)^2 d\omega \\ &= \int_0^M \int_0^M (M - \max\{\eta, \eta'\})(g(\eta) - g^*(\eta))(g(\eta') - g^*(\eta'))\,d\eta\,d\eta'. \end{aligned}$$

This expression is simply a quadratic form in $g - g^*$ and thus easy to implement in the numerical scheme. We summarize our results which slightly generalize Lemma 2.3 in [22].

Lemma 4.2.1. *Let the initial datum $u^0 : \mathscr{T} \to \mathbb{R}$ be point-symmetric. Then the solution u^{n+k} to the BDF-k scheme (4.7) is in one-to-one correspondence to the solution g^{n+k} obtained from the inductive scheme*

$$g^{n+k} = \operatorname{argmin}_g \Psi(g), \tag{4.8}$$

with initial condition $g^0 = 1/u \circ G$ and given $g^1, \dots, g^{n+k-1}$ obtained from a lower-order scheme. The argmin has to be taken over all measurable functions $g : [0, M] \to (0, \infty)$ satisfying the mass constraint $\int_0^M g(\omega)\,d\omega = 1$, and the *function* $\Psi(g)$ *is given by*

$$\Psi(g) = \tag{4.9}$$

$$-\frac{1}{2\tau}\int_0^M \int_0^M (M - \max\{\eta,\eta'\}) \sum_{i=0}^{k-1} a_i(g(\eta) - g^{n+i}(\eta))(g(\eta') - g^{n+i}(\eta'))\, d\eta\, d\eta' \tag{4.10}$$

$$+\frac{1}{\alpha(1-\alpha)}\int_0^M g(\omega)^{1-\alpha}\, d\omega.$$

Moreover, $\Phi(u) = \Psi(g)$ *and the functions* u^n *and* g^n *are related by*

$$u^n(x_\omega) = \frac{1}{g^n(\omega)}, \quad x_\omega = \int_0^\omega g^n(\eta)\, d\eta. \tag{4.11}$$

In Lagrangian coordinates, the problem has become a minimization problem in $\Psi(g)$ which is the sum of a quadratic form and a convex functional, hence it is convex. In the special case $\alpha = -1$, the second integral in (4.9) is quadratic too which simplifies the numerical discretization. Therefore, we will consider mainly numerical examples with $\alpha = -1$ in Section 4.3.

4.2.3 Spatial discretization

We approximate the infinite-dimensional variational problem (4.8) by a finite-dimensional one. Minimization in (4.8) is performed over the finite-dimensional space of quadratic ansatz functions. This generalizes the approach in [22], where only linear ansatz functions were used. We define the ansatz space as follows.

Let $N \in \mathbb{N}$ and a mesh $\{x_0, \ldots, x_N\}$ on $[0,1]$ be given with $x_0 = 0$ and $x_N = 1$. Using (4.11), we construct the mesh $\Omega_N = \{\omega_0, \omega_1 \ldots, \omega_N\}$ of $[0, M]$. Then $\omega_0 = 0$, $\omega_N = M$, $\omega_i < \omega_{i+1}$, and $\omega_{N-i} = M - \omega_i$ (point-symmetry), where $i = 1, \ldots, N-1$. Since we wish to introduce quadratic ansatz functions, we add the grid points $\omega_{j+1/2} = (\omega_{j+1} + \omega_j)/2$ for $j = 0, \ldots, N-1$.

The basis functions $\phi_j : \mathscr{T} \to \mathbb{R}$ are defined by

$$\phi_j(\omega) = \begin{cases} \frac{\omega-\omega_{j-1}}{\omega_j-\omega_{j-1}} & \text{for } \omega \in [\omega_{j-1}, \omega_j], \\ \frac{\omega_{j+1}-\omega}{\omega_{j+1}-\omega_j} & \text{for } \omega \in [\omega_j, \omega_{j+1}], \\ 0 & \text{otherwise}, \end{cases} \qquad j = 1, \ldots, N-1,$$

$$\phi_N(\omega) = \begin{cases} \frac{\omega_1-\omega}{\omega_1} & \text{for } \omega \in [0, \omega_1], \\ \frac{\omega-\omega_{N-1}}{M-\omega_{N-1}} & \text{for } \omega \in [\omega_{N-1}, M], \\ 0 & \text{otherwise}. \end{cases}$$

This set of piecewise linear functions is supplemented by the following piecewise quadratic basis functions:

$$\phi_{N+j}(\omega) = \begin{cases} 1 - \left(\frac{2\omega-(\omega_{j-1}+\omega_j)}{\omega_j-\omega_{j-1}}\right)^2 & \text{for } \omega \in [\omega_{j-1}, \omega_j], \\ 0 & \text{otherwise}, \end{cases} \qquad j = 1, \ldots, N.$$

The ansatz space is the set of all positive, piecewise quadratic functions $g : [0, M] \to \mathbb{R}_+$ of the form

$$g(\omega) = \sum_{j=1}^{2N} g_j \phi_j(\omega). \tag{4.12}$$

We call $\mathbf{g} := (g_1, \ldots, g_{2N}) \in [0, \infty)^{2N}$ the associated weight vector. By definition of ϕ_j, we have $g(\omega_j) = g_j$ for $j = 1, \ldots, N$. Moreover, as g is point-symmetric, $g_0 = g_N$.

Now, for given mass $M > 0$ and grid $\Omega_N \subset [0, M]$, we define the set $\mathbb{G}_M^N \subset \mathbb{R}_+^{2N}$ as the set of weight vectors $\mathbf{g}$ for which the associated interpolation g from (4.12) satisfies the mass constraint,

$$1 = \int_0^M g(\omega)\, d\omega = \sum_{j=1}^{N} \left(\frac{g_{j-1} + g_j}{2} + \frac{2}{3} g_{N+j} \right) (\omega_j - \omega_{j-1}). \tag{4.13}$$

The Wasserstein metric for functions approximated in this way becomes

$$\begin{aligned} &W[u, u^*]^2 = \\ &\int_0^M \int_0^M (M - \max\{\eta, \eta'\}) \sum_{i=1}^{2N} (g_i - g_i^*) \phi_i(\eta) \sum_{j=1}^{2N} (g_j - g_j^*) \phi_j(\eta')\, d\eta\, d\eta' \\ &= \sum_{i,j=1}^{N} (g_i - g_i^*)(g_j - g_j^*) a_{ij} + \sum_{i,j=1}^{N} (g_{N+i} - g_{N+i}^*)(g_j - g_j^*) b_{ij} \\ &\quad + \sum_{i,j=1}^{N} (g_i - g_i^*)(g_{N+j} - g_{N+j}^*) b_{ji} + \sum_{i,j=1}^{N} (g_{N+i} - g_{N+i}^*)(g_{N+j} - g_{N+j}^*) c_{ij}, \end{aligned}$$

where

$$\begin{aligned} a_{ij} &= \int_0^M \int_0^M (M - \max\{\eta, \eta'\}) \phi_i(\eta) \phi_j(\eta')\, d\eta\, d\eta', \\ b_{ij} &= \int_0^M \int_0^M (M - \max\{\eta, \eta'\}) \phi_{N+i}(\eta) \phi_j(\eta')\, d\eta\, d\eta', \\ c_{ij} &= \int_0^M \int_0^M (M - \max\{\eta, \eta'\}) \phi_{N+i}(\eta) \phi_{N+j}(\eta')\, d\eta\, d\eta'. \end{aligned} \tag{4.14}$$

The coefficients a_{ij}, b_{ij}, and c_{ij} can be computed explicitly. The explicit expressions are given in Appendix C.1. Setting $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$, and defining the matrix

$$M_w = (M_{ij}) = \begin{pmatrix} A & B^\top \\ B & C \end{pmatrix}, \tag{4.15}$$

we can formulate the above sum as

$$W[u, u^*]^2 = \sum_{i,j=1}^{2N} M_{ij} (g_i - g_i^*)(g_j - g_j^*).$$

As the matrices A and C are symmetric, M_w is symmetric, too. The matrix A corresponds to the linear approximation considered in [22].

4.2.4 Minimization

The numerical scheme consists of the following finite-dimensional variational problem:

$$\mathbf{g}^{n+k} = \operatorname{argmin}_{\mathbf{g}\in\mathbb{G}_M^N} \Psi_N(\mathbf{g}), \tag{4.16}$$

$$\text{where} \quad \Psi_N(\mathbf{g}) = -\frac{1}{2\tau}\sum_{\ell=0}^{k-1} a_\ell \sum_{i,j=1}^{2N} M_{ij}(g_i - g_i^{n+\ell})(g_j - g_j^{n+\ell}) + S_N[\mathbf{g}],$$

and where $\mathbf{g} = (g_1, \ldots, g_{2N})$,

$$S_N[\mathbf{g}] = \frac{1}{\alpha(\alpha-1)} \sum_{i=0}^{N-1} \int_{\omega_i}^{\omega_{i+1}} \big(g_i\phi_i + g_{i+1}\phi_{i+1} + g_{N+i}\phi_{N+i}\big)^{1-\alpha}\, d\omega. \tag{4.17}$$

The functions $\Psi_N(\mathbf{g})$ and $\Psi(g)$ from (4.9) are related by $\Psi_N(\mathbf{g}) = \Psi(g)$ with a piecewise quadratic function g defined from $\mathbf{g}$ by (4.12). Since Ψ_N is convex for $\alpha < 0$ and the set $\mathbb{G}_M^N$ is convex, there exists a unique minimizer of (4.16).

4.2.5 Fully discrete Euler-Lagrange equations

The minimizer $\mathbf{g}_{n+k}$ in (4.16) is subject to the mass constraint (4.13), by definition of the set $\mathbb{G}_M^N$. Therefore, instead of working on the set $\mathbb{G}_M^N$, it is more convenient to consider (4.16) as a constrained minimization problem for $\mathbf{g}$ on the larger set $\mathbb{R}^{2N}$, which is solved by the method of Lagrange multipliers λ using the Lagrange functional

$$L(\mathbf{g}, \lambda) = \Psi_N(\mathbf{g}) - \lambda\Bigg(1 - \sum_{j=1}^{N}\left(\frac{g_{j-1}+g_j}{2} + \frac{2}{3}g_{N+j}\right)(\omega_j - \omega_{j-1})\Bigg).$$

A critical point of L satisfies the $2N$ conditions

$$0 = \mathbf{G}_j := \frac{\partial L}{\partial g_j} = -\frac{1}{\tau}\sum_{\ell=0}^{k-1} a_\ell \sum_{i=1}^{2N} M_{ij}(g_i - g_i^{n+\ell}) + \frac{\partial S_N}{\partial g_j}, \quad j = 1, \ldots, 2N.$$

The precise values for $\partial S_N/\partial g_j$ are given in Appendix C.2 for $\alpha = -1$. The condition for the constraint is recovered from

$$0 = \mathbf{G}_{2N+1} := \frac{\partial L}{\partial \lambda} = 1 - \sum_{j=1}^{N}\left(\frac{g_{j-1}+g_j}{2} + \frac{2}{3}g_{N+j}\right)(\omega_j - \omega_{j-1}).$$

The vector $\mathbf{G}[\mathbf{g}, \lambda] = (\mathbf{G}_1, \ldots, \mathbf{G}_{2N+1}) \in \mathbb{R}^{2N+1}$ is the gradient of $L(\mathbf{g}, \lambda)$ with respect to $(\mathbf{g}, \lambda)$. We approximate a critical point numerically by applying the Newton method to the first-order optimality condition $\mathbf{G}[\mathbf{g}, \lambda] = 0$. This leads to a sequential quadratic programming method, since at every Newton iteration step a quadratic subproblem has to be solved.

4.2.6 Implementation

Let the solution $\mathbf{g}$ at the nth time step be given and let $\mathbf{g}^{(0)} := \mathbf{g}$, $\lambda^{(0)} := 0$. The iteration is as follows:

$$\mathbf{g}^{(s+1)} := \mathbf{g}^{(s)} + (\delta\mathbf{g})^{(s+1)}, \quad \lambda^{(s+1)} := \lambda^{(s)} + (\delta\lambda)^{(s+1)},$$

where $((\delta\mathbf{g})^{(s+1)}, (\delta\lambda)^{(s+1)})$ is the solution to the linear system

$$H[\mathbf{g}^{(s)}, \lambda^{(s)}]((\delta\mathbf{g})^{(s+1)}, (\delta\lambda)^{(s+1)}) = -\mathbf{G}[\mathbf{g}^{(s)}, \lambda^{(s)}],$$

where $H[\mathbf{g}^{(s)}, \lambda^{(s)}]$ denotes the Hessian of Ψ_N, whose entries are given in Appendix C.2 for $\alpha = -1$. The iteration is stopped if the norm of $((\delta\mathbf{g})^{(s+1)}, (\delta\lambda)^{(s+1)})$ is smaller than a certain threshold (see Section 4.3 for details). In this case, we define $\mathbf{g}^{n+1} := \mathbf{g}^{(s+1)}$ and $\lambda^{n+1} := \lambda^{(s+1)}$ at the $(n+1)$th time step. For the BDF-k scheme, the values $\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}$ are computed from a lower-order scheme. In the numerical section below, we employ the BDF-2 scheme only such that $\mathbf{g}^1$ is calculated by the implicit Euler method.

Note that the constrained minimization problem is exactly mass conserving by construction, but the Newton iteration introduces a small error which depends on the tolerance imposed in the Newton method.

In each iteration step, we need to invert the dense matrix H which is the sum of M_w and the Hessian of S_N. This is not a numerical challenge in the one-dimensional case we consider but it may become critical in multi-dimensional discretizations on fine grids. For $\alpha = -1$, however, M_w and the Hessian of S_N are constant matrices which significantly simplifies the Newton scheme.

4.2.7 Choice of the initial condition

In order to compute the initial condition in Lagrangian coordinates, we need to make precise the values g_i^0 of the vector $\mathbf{g} \in \mathbb{G}_M^N$ and the points x_j^0 of the spatial lattice which is moving as the solution evolves. Let the mesh $\{x_0^0, \ldots, x_N^0\}$ be given and set $g^0(\omega_j) = 1/u^0(x_j)$. Approximating the initial function by a linear ansatz function, we obtain

$$x_j^0 = G^0(\omega_j) = \frac{1}{2}\sum_{i=1}^{j}(\omega_j - \omega_{j-1})(g_{i-1}^0 + g_i^0), \quad j = 1, \ldots, N. \tag{4.18}$$

This is a system of linear equations in $\omega_1, \ldots, \omega_N$. Choosing the uniform grid $x_j^0 = j/N$, we can solve this system explicitly. Indeed, since

$$\frac{1}{N} = x_{j+1}^0 - x_j^0 = \frac{1}{2}(\omega_{j+1} - \omega_j)(g_{j+1}^0 + g_j^0),$$

which can be solved for ω_{j+1}:

$$\omega_{j+1} = \omega_j + \frac{2}{N}(g_{j+1}^0 + g_j^0)^{-1}, \quad j = 0, \ldots, N-1.$$

As u^0 is assumed to be point-symmetric, this is true for (ω_i) too. Finally, we approximate $g^0(\omega_{j+1/2})$ by the arithmetic mean $\frac{1}{2}(g_{j-1}^0 + g_j^0)$, $j = 1, \ldots, N$. Consequently, the weights g_{N+i}^0 vanish for $i = 1, \ldots, N$ in the expansion $g^0(\omega) = \sum_{j=0}^{2N} g_j^0 \phi_j$, which is consistent with our approximation (4.18). We also refer to the discussion in [22, Section 2.8].

4.3 Numerical experiments

In this section, we present some numerical results for (4.1) with $\alpha = -1$, by employing the BDF-2 method with quadratic ansatz functions. We choose a uniform grid for $x \in [0, 1]$ with $N = 100$ grid points, and the time step size $\tau = 10^{-5}$. The Newton iterations are stopped if both the relative ℓ^∞ error in the g-variables and the ℓ^2-norm of $\mathbf{G}[\mathbf{g}^{(s)}, \lambda^{(s)}]$ are smaller than 10^{-8}.

Figure 4.1 illustrates the temporal evolution of the solution $u(x,t)$ with the initial conditions $u^0(x) = \cos(2\pi x)^2 + 0.01$(left) and $u^0(x) = \sqrt[5]{|x - 0.5| + 0.0001} - 0.1$ (right). We observe that, as expected, the solutions converge to the constant steady state. Because of the negative exponent α, very small initial values increase quickly in time.

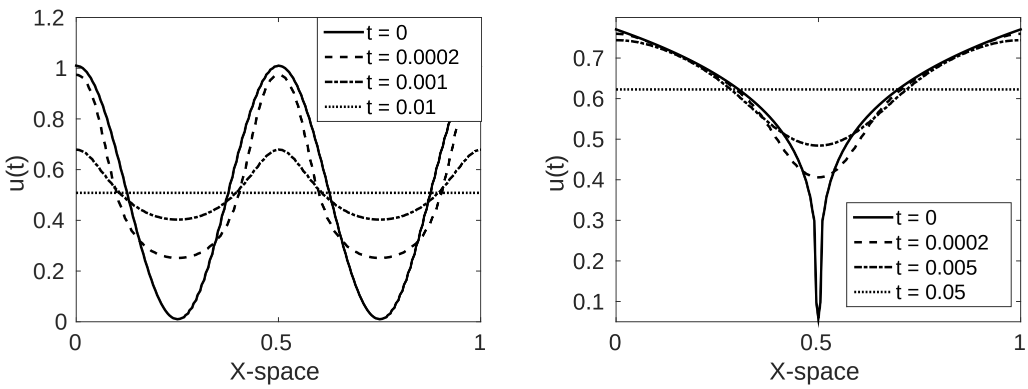

Figure 4.1: Time evolution of the solution to the diffusion equation (4.1) with $\alpha = -1$ for two different initial conditions.

A nice feature of the Wasserstein gradient flow scheme is that we may interpret the evolution as a process of redistribution of particles with spatio-temporal density $u(x,t)$ on $\mathscr{T}$ under the influence of a nonlinear particle interaction, which is described by S. The way in which the initial density u^0 is "deformed" during the time evolution is illustrated in Figure 4.2. We have chosen 50 "test particles" for the solutions to (4.1) for the initial conditions chosen above. We stress the fact that the density of trajectories can generally be not identified with the density u of the solution.

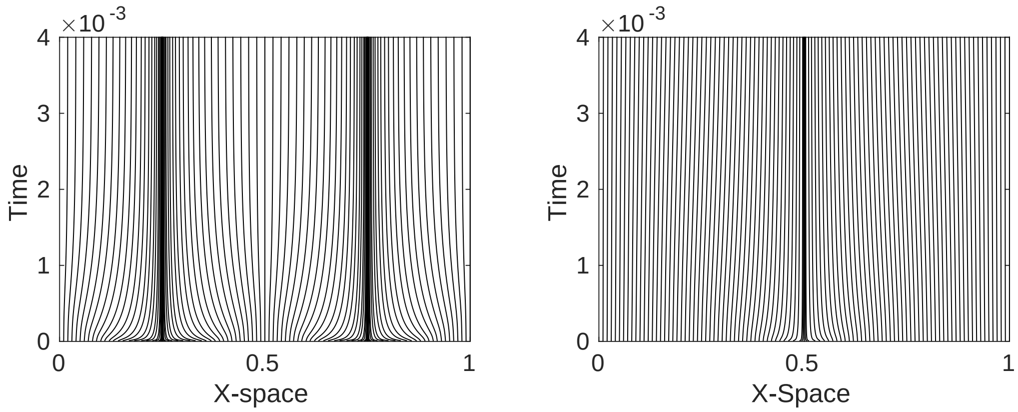

Figure 4.2: Particle trajectories in the Wasserstein gradient flow scheme, corresponding to the solutions of Figure 4.1 with $N = 50$.

We verify that the discretization is indeed of second order. Figure 4.3 shows the ℓ^∞-error for various numbers of grid points N. We have chosen the initial datum $u_0(x) = \cos(2\pi x)^2 + 0.1$, the end time $T = 0.004$, and the time step size $\tau = 10^{-7}$. The reference solution is computed by using $N = 500$, and $\tau = 10^{-7}$. The differences $g(\cdot,T) - g_{\mathrm{ref}}(\cdot,T)$ and $u(\cdot,T) - u_{\mathrm{ref}}(\cdot,T)$ in the ℓ^∞-norm feature the expected second-order dependence on N.

Next, we fix the number of grid points $N = 100$ and compute the $L^\infty(\tau^*,T; L^2(\mathscr{T}))$ error for varying time step sizes τ; see Figure 4.4. Because of the approximation of the initial datum as detailed in Section 4.2.7, the error is not of second order initially. Therefore, we compute the error in the interval (τ^*,T) with $\tau^* = 10^{-4}$. The errors are of second order, as expected.

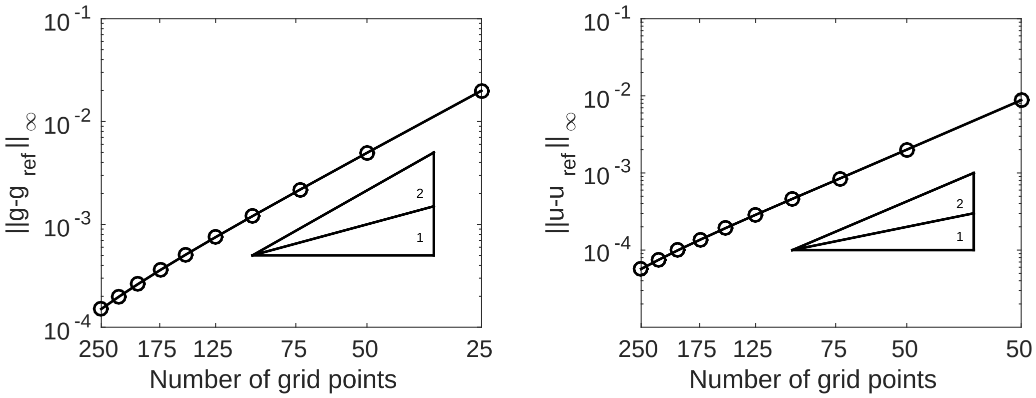

Figure 4.3: ℓ^∞-error for $(g - g_{\rm ref})(T)$ (left) and $(u - u_{\rm ref})(T)$ (right) at $T = 0.004$ for various numbers of grid points.

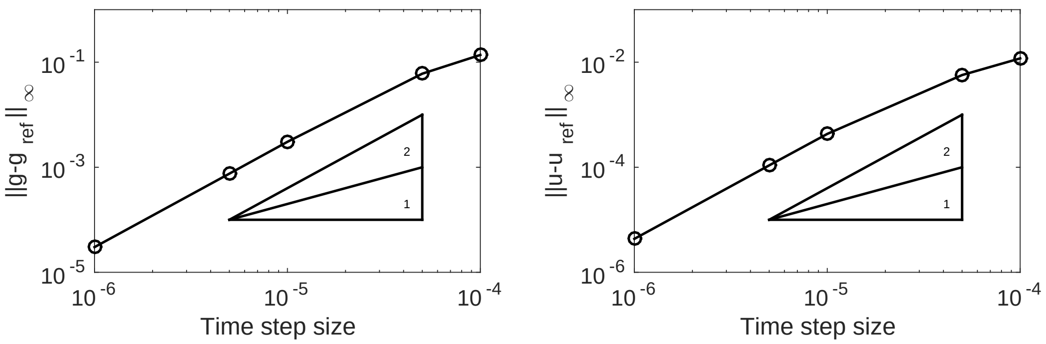

Figure 4.4: $L^\infty(\tau^*, T; L^2(\mathscr{T}))$-error for $g - g_{\rm ref}$ (left) and $u - u_{\rm ref}$ (right) for various time step sizes, with $\tau^* = 10^{-4}$.

The time decay of the discrete version of the relative entropy $S[u] - S[u_\infty]$ is presented in Figure 4.5 (left) for various grid numbers. We observe that the decay is exponential until saturation. The saturation comes from the spatial error and the error from the Newton iteration. The decay rate is estimated in the linear regime from the difference quotient

$$\lambda \approx \frac{1}{\tau}\big(\log S[u(t+\tau)] - \log S[u(t)]\big).$$

The numerical decay rates for $\alpha = -1, -2$ are shown in Figure 4.6. The rate for $\alpha = -2$ (right) is much larger than the corresponding one for $\alpha = -1$ (left), since a smaller exponent yields a larger diffusion coefficient (if $u < 1$) and thus, diffusion becomes faster. We also see that the decay rates become larger on a finer spatial grid. This behavior seems to confirm recent analytical results for spatial discretizations of Fokker-Planck equations; see [45, Section 5]. One may ask if a similar behavior can be observed for the decay rate as a function of the

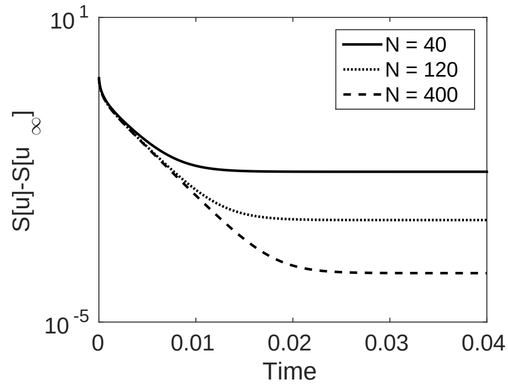

Figure 4.5: Discrete relative entropy $S[u(t)] - S[u_\infty]$ versus time with $\alpha = -1$ for various N.

time step size. However, our numerical experiments do not show a monotonic dependence (figures not presented); rather the decay rates vary in a small range which seems to be determined by the other numerical error parts.

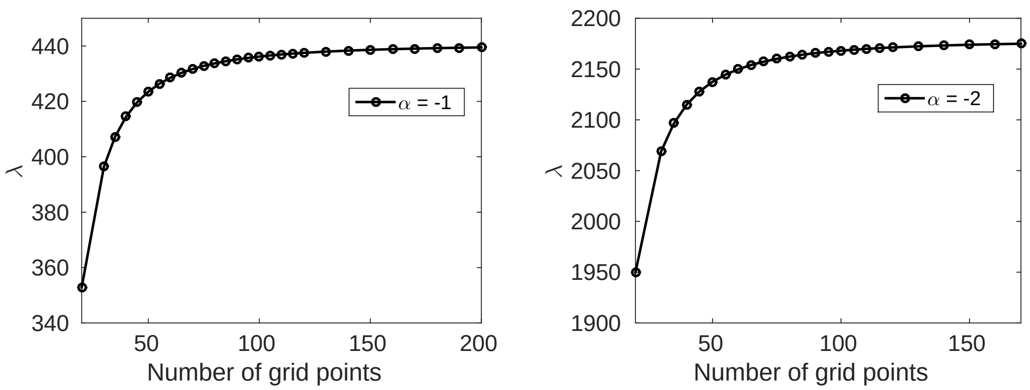

Figure 4.6: Estimated decay rate of the entropy versus number of grid points N for $\alpha = -1$ (left) and $\alpha = -2$ (right).

In Figure 4.7, the decay of the square of the relative G-norm is presented. The G-norm is calculated according to (1.4), where the argument is given by $\mathbf{g} - \mathbf{g}_\infty$ and $\mathbf{g}_\infty$ is the weight vector corresponding to the constant steady state. Again, the decay is much faster for $\alpha = -2$ because of the faster diffusion.

Finally, we present some results on the time decay of the discrete variance of u^n and g^n at time τn, defined by

$$\mathrm{Var}(u^n)^2 = \sum_{i=1}^{N+1} (u_i^n - E)^2 (x_i - x_{i-1}),$$

where E is the expectation value of u^n (which equals the mass and is therefore

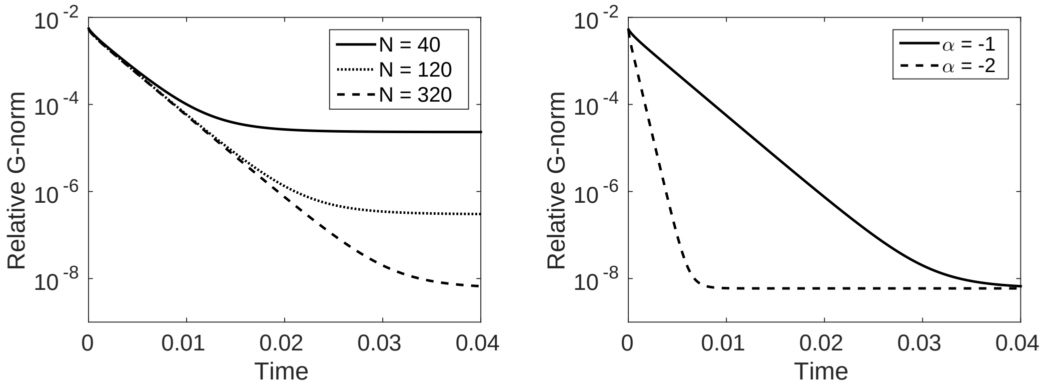

Figure 4.7: Relative G-norm versus time for various grid numbers N (left) and for two values of the exponent α (right).

constant in time). The discrete variance of g^n is defined in a similar way. Interestingly, the variances are exponentially decaying (Figure 4.8), although it is not clear how to prove this property analytically.

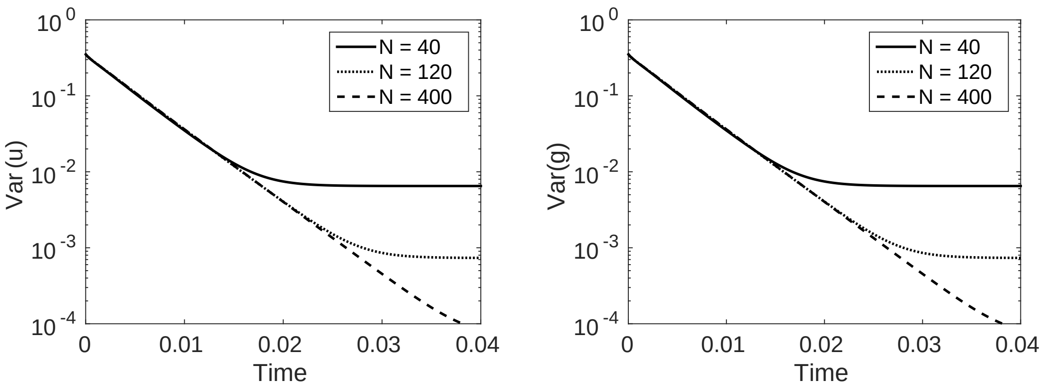

Figure 4.8: Discrete variance of u^n (left) and g^n (right) versus time for various grid numbers N.

Appendix A

Wasserstein gradient

In Chapter 3 we investigate numerical schemes for gradient flows on the Wasserstein space. To give a meaning to a gradient flow, we need a definition of a gradient. To this end we give here a (formal) derivation of the Wasserstein gradient. Remember that on a manifold with inner product $\langle \cdot, \cdot \rangle_M$ the gradient of some function $f : M \to \mathbb{R}$ in direction v is defined via

$$\langle \nabla^M f, v \rangle_M = Df(v).$$

In the same manner we want to define the Wasserstein gradient $\nabla^W F$ of some functional $F : \mathscr{P}(M) \to \mathbb{R}$. To this end, we have to define an inner product on the tangent space of $\mathscr{P}(X)$. Following the derivation of the norm of a velocity vector $\partial_t \mu$ (associating a velocity field ∇u to each velocity vector $\partial_t \mu$) in the introduction 1.2 we set

$$\langle \partial_t \mu, \partial_t \mu^* \rangle_W = \int \mu \langle \nabla u, \nabla u^* \rangle_{L^2} dx,$$

where u, u^* solve

$$\partial_t \mu + \operatorname{div}(\mu \nabla u) = 0, \quad \partial_t \mu^* + \operatorname{div}(\mu \nabla u^*) = 0.$$

We now turn to the calculation of the gradient with respect to this inner product. Recall that the infinitesimal variation of a functional F in direction $\partial_t \mu$ is given by

$$\delta F(\partial_t \mu) = \int DF(\mu) \partial_t \mu \, dx,$$

where DF is the Fréchet derivative. We want to define ∇^W so that the identity

$$\langle \nabla^W F, \partial_t \mu \rangle_W = \int DF(\mu) \partial_t \mu \, dx$$

holds true. The gradient is then obtained by a straight forward calculation:

$$\begin{aligned} \int DF(\mu) \partial_t \mu dx &= - \int DF(\mu) \operatorname{div}(\mu \nabla u) dx = \int \mu \nabla DF(\mu) \nabla u dx \\ &= \langle - \operatorname{div}(\mu \nabla DF), \partial_t \mu \rangle_W dx = \langle \nabla^W F, \partial_t \mu \rangle_W dx \end{aligned}$$

In the first line we used the identification $\partial_t \mu = -\operatorname{div}(\mu \nabla u^*)$. The second line already is in the form of a Wasserstein inner product for some vector $\eta = -\operatorname{div}(\mu \nabla DF(\mu))$. Consequently, we set $\nabla^W F = \eta = -\operatorname{div}(\mu \nabla DF(\mu))$.

Appendix B

Discretizations needed in Chapter 2

B.1 The H^{-1}-norm

We calculate the entries for the matrix of the discrete H^{-1}-norm to be

$$
\begin{aligned}
A_{j-1,j} &= \int_{\omega_{j-1}}^{\omega_j} \varphi_{j-1}(\omega)\varphi_j(\omega)d\omega = \int_{\omega_{j-1}}^{\omega_j} \frac{\omega_j - \omega}{\omega_j - \omega_{j-1}} \frac{\omega - \omega_{j-1}}{\omega_j - \omega_{j-1}} d\omega \\
&= \frac{\omega_j - \omega_{j-1}}{6}. \\
A_{j,j} &= \int_{\omega_{j-1}}^{\omega_{j+1}} \varphi_j(\omega)\varphi_j(\omega)d\omega \\
&= \int_{\omega_{j-1}}^{\omega_j} \left(\frac{\omega - \omega_{j-1}}{\omega_j - \omega_{j-1}}\right)^2 d\omega + \int_{\omega_j}^{\omega_{j+1}} \left(\frac{\omega_{j+1} - \omega}{\omega_{j+1} - \omega_j}\right)^2 d\omega \\
&= \frac{w_{j+1} - w_{j-1}}{3}. \\
A_{j-1,j} &= \int_{\omega_j}^{\omega_{j+1}} \varphi_{j+1}(\omega)\varphi_j(\omega)d\omega = \int_{\omega_j}^{\omega_{j+1}} \frac{\omega - \omega_j}{\omega_{j+1} - \omega_j} \frac{\omega_{j+1} - \omega}{\omega_{j+1} - \omega_j} d\omega \\
&= \frac{\omega_{j+1} - \omega_j}{6}. \\
B_{j-1,j} &= \int_{\omega_{j-1}}^{\omega_j} \nabla\varphi_{j-1}(\omega)\nabla\varphi_j(\omega)d\omega = \int_{\omega_{j-1}}^{\omega_j} \frac{-1}{\omega_j - \omega_{j-1}} \frac{1}{\omega_j - \omega_{j-1}} d\omega \\
&= -\frac{1}{\omega_j - \omega_{j-1}}.
\end{aligned}
$$

$$\begin{aligned}
B_{j,j} &= \int_{\omega_{j-1}}^{\omega_{j+1}} \varphi_j(\omega)\varphi_j(\omega)d\omega \\
&= \int_{\omega_{j-1}}^{\omega_j} \left(\frac{1}{\omega_j-\omega_{j-1}}\right)^2 d\omega + \int_{\omega_j}^{\omega_{j+1}} \left(\frac{-1}{\omega_{j+1}-\omega_j}\right)^2 d\omega \\
&= \frac{1}{\omega_j-\omega_{j-1}} + \frac{1}{\omega_{j+1}-\omega_j}. \\
B_{j-1,j} &= \int_{\omega_j}^{\omega_{j+1}} \varphi_{j+1}(\omega)\varphi_j(\omega)d\omega = \int_{\omega_j}^{\omega_{j+1}} \frac{1}{\omega_{j+1}-\omega_j}\frac{-1}{\omega_{j+1}-\omega_j}d\omega \\
&= -\frac{1}{\omega_{j+1}-\omega_j}.
\end{aligned}$$

B.2 The entropy Φ

We calculate Φ_1, Φ_{j+} and Φ_N starting with

$$\begin{aligned}
\Phi_1 &= \frac{1}{m+1}\int_{\omega_0}^{\omega_1} \left(u_1\frac{\omega-\omega_0}{\omega_1-\omega_0}\right)^{m+1} d\omega \\
&= \underbrace{\frac{1}{m+1}\left(\frac{1}{\omega_1-\omega_0}\right)^{m+1}}_{c_1} \int_{\omega_0}^{\omega_1} \left(\underbrace{u_1}_{c_3}\omega \underbrace{-u_1\omega_0}_{c_2}\right)^{m+1} d\omega \\
&= c_1\int_{\omega_j}^{\omega_{j+1}} (c_2+c_3\omega)^{m+1}\, d\omega.
\end{aligned}$$

Setting $y = c_2 + c_3\omega$ yields

$$\begin{aligned}
\Phi_1 &= \frac{c_1}{c_3}\int_{c_2+c_3\omega_0}^{c_2+c_3\omega_1} y^{m+1}dy = \frac{c_1}{c_3(m+2)} y^{m+2}|_{c_2+c_3\omega_0}^{c_2+c_3\omega_1} \\
&= \frac{(u_1\omega_1-u_1\omega_0)^{m+2}}{(\omega_1-\omega_0)^{m+1}(u_1)(m+1)(m+2)} = \frac{u_1^{m+2}(\omega_1-\omega_0)}{(m+1)(m+2)u_1} \\
&= \frac{u_1^{m+1}(\omega_1-\omega_0)}{(m+1)(m+2)}.
\end{aligned}$$

We calculate Φ_{j+1} in the same way:

$$\begin{aligned}
\Phi_{j+1} &= \frac{1}{m+1}\int_{\omega_j}^{\omega_{j+1}} \left(u_j \frac{\omega_{j+1}-\omega}{\omega_{j+1}-\omega_j} + u_{j+1}\frac{\omega-\omega_j}{\omega_{j+1}-\omega_j}\right)^{m+1} d\omega \\
&= \underbrace{\frac{1}{(m+1)(\omega_{j+1}-\omega_j)^{m+1}}}_{c_1} \int_{\omega_j}^{\omega_{j+1}} (\underbrace{u_j\omega_{j+1}-u_{j+1}\omega_j}_{c_2} + \underbrace{(u_{j+1}-u_j)}_{c_3}\omega)^{m+1} d\omega \\
&= c_1 \int_{\omega_j}^{\omega_{j+1}} (c_2+c_3\omega)^{m+1}\, d\omega.
\end{aligned}$$

Setting $y = c_2 + c_3\omega$ yields

$$\begin{aligned}
\Phi_{j+1} &= \frac{c_1}{c_3}\int_{c_2+c_3\omega_j}^{c_2+c_3\omega_{j+1}} y^{m+1} dy = \frac{c_1}{c_3(m+2)} y^{m+2}\big|_{c_2+c_3\omega_j}^{c_2+c_3\omega_{j+1}} \\
&= \frac{(u_{j+1}^{m+2}-u_j^{m+2})(\omega_{j+1}-\omega_j)}{(m+1)(m+2)(u_{j+1}-u_j)}.
\end{aligned}$$

Finally we calculate Φ_N

$$\begin{aligned}
\Phi_N &= \frac{1}{m+1}\int_{\omega_{N-1}}^{\omega_N} \left(u_{N-1}\frac{\omega_N-\omega}{\omega_N-\omega_{N-1}}\right)^{m+1} d\omega \\
&= \underbrace{\frac{1}{m+1}\left(\frac{1}{\omega_N-\omega_{N-1}}\right)^{m+1}}_{c_1} \int_{\omega_{N-1}}^{\omega_N} \left(\underbrace{u_{N-1}\omega_N}_{c_2} \underbrace{-u_{N-1}}_{c_3}\omega \right)^{m+1} d\omega \\
&= c_1 \int_{\omega_{N-1}}^{\omega_N} (c_2+c_3\omega)^{m+1}\, d\omega.
\end{aligned}$$

Set $y = c_2 + c_3\omega$

$$\begin{aligned}
\Phi_N &= \frac{c_1}{c_3}\int_{c_2+c_3\omega_{N-1}}^{c_2+c_3\omega_N} y^{m+1} dy = \frac{c_1}{c_3(m+1)} y^{m+2}\big|_{c_2+c_3\omega_{N-1}}^{c_2+c_3\omega_N} \\
&= \frac{u_{N-1}^{m+1}(\omega_N-\omega_{N-1})}{(m+1)(m+2)}.
\end{aligned}$$

B.3 Gradient of Φ

We turn to the calculation of the gradient. Therefore note that only Φ_1 and Φ_2 depend on u_1. We sum up the two terms and calculate the derivative as follows:

$$\frac{\partial \Phi}{\partial u_1}$$

$$\begin{aligned}
&= \frac{\partial}{\partial u_1}\frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m+1}(\omega_1-\omega_0)}{(m+1)(m+2)}\right)\\
&+ \frac{\partial}{\partial u_1}\frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m+1}(\omega_N-\omega_{N-1})}{(m+1)(m+2)}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m}(\omega_1-\omega_0)}{(m+2)} - \frac{u_1^{m+1}(m+1)(\omega_1-\omega_0)(u_2-u_1)(m+1)(m+2)}{(u_2-u_1)^2(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)(m+1)(m+2)}{(u_2-u_1)^2(m+1)^2(m+2)^2}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m}(\omega_1-\omega_0)}{(m+2)} - \frac{u_1^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)(m+1)} + \frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)(m+2)}\right)
\end{aligned}$$

In the same way as only Φ_1 and Φ_2 depend on u_1, only Φ_{l-1} and Φ_l depend on u_l. We sum up the terms and calculate

$$\frac{\partial \Phi}{\partial u_l}$$

$$\begin{aligned}
&= \frac{\partial}{\partial u_l}\frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)(m+1)(m+2)} + \frac{(u_l^{m+2}-u_{l-1}^{m+2})(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})(m+1)(m+2)}\right)\\
&= -\frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+2}(m+1)(\omega_{l+1}-\omega_l)(u_{l+1}-u_l)(m+1)(m+2)}{(u_{l+1}-u_l)^2(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)(m+1)(m+2)}{(u_{l+1}-u_l)^2(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(m+2)(\omega_l-\omega_{l-1})(u_l-u_{l-1})(m+1)(m+2)}{(u_l-u_{l-1})^2(m+1)^2(m+2)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_l^{m+2}-u_{l-1}^{m+2})(\omega_l-\omega_{l-1})(m+1)(m+2)}{(u_l-u_{l-1})^2(m+1)^2(m+2)^2}\right)\\
&= -\frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)(m+1)} + \frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)(m+2)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})(m+1)} - \frac{(u_l^{m+2}-u_{l-1}^{m+2})(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^2(m+1)(m+2)}\right)
\end{aligned}$$

We sum up Φ_{N-1} and Φ_N and calculate:

$$\frac{\partial \Phi}{\partial u_{N-1}}$$

$$= \frac{\partial}{\partial u_{N-1}} \left(\frac{\beta_k}{\alpha_k} \frac{(u_{N-1}^{m+2} - u_{N-2}^{m+2})(\omega_{N-1} - \omega_{N-2})}{(u_{N-1} - u_{N-2})(m+1)(m+2)} + \frac{u_{N-1}^{m+1}(\omega_N - \omega_{N-1})}{(m+1)(m+2)} \right)$$
$$= \frac{\beta_k}{\alpha_k} \left(\frac{(m+2)u_{n-1}^{m+1}(\omega_{N-1} - \omega_{N-2})(m+1)(m+2)(u_{N-1} - u_{N-2}}{(u_{N-1} - u_{N-2})^2(m+1)^2(m+2)^2} \right)$$
$$- \frac{\beta_k}{\alpha_k} \left(\frac{(u_{N-1}^{m+2} - u_{N-2}^{m+2})(m+1)(m+2)(\omega_{N-1} - \omega_{N-2})}{(u_{N-1} - u_{N-2})^2(m+1)^2(m+2)^2} \right)$$
$$+ \frac{\beta_k}{\alpha_k} \left(\frac{(m+1)u_{N-1}^{m}(\omega_N - \omega_{N-1})}{(m+1)(m+2)} \right)$$
$$= \frac{\beta_k}{\alpha_k} \left(\frac{u_{N-1}^{m+1}(\omega_{N-1} - \omega_{N-2})}{(u_{N-1} - u_{N-2})(m+1)} - \frac{(u_{N-1}^{m+2} - u_{N-2}^{m+2})(\omega_{N-1} - \omega_{N-2})}{(u_{N-1} - u_{N-2})^2(m+1)(m+2)} \right)$$
$$+ \frac{\beta_k}{\alpha_k} \left(\frac{u_{N-1}^{m}(\omega_N - \omega_{N-1})}{(m+2)} \right)$$

B.4 Hessian of Φ

The calculation of the gradient above showed that we only have to collect the terms depending on certain variables. Summing up the terms for the different derivatives we calculate:

$$\frac{\partial \Phi}{\partial u_{l-1} \partial u_l}$$

$$= \frac{\beta_k}{\alpha_k} \frac{\partial}{\partial u_{l-1}} \left(\frac{u_l^{m+1}(\omega_l - \omega_{l-1})}{(u_l - u_{l-1})(m+1)} - \frac{(u_l^{m+2} - u_{l-1}^{m+2})(\omega_l - \omega_{l-1})}{(u_l - u_{l-1})^2(m+1)(m+2)} \right)$$
$$= \frac{\beta_k}{\alpha_k} \left(\frac{u_l^{m+1}(\omega_l - \omega_{l-1})(m+1)}{(u_l - u_{l-1})^2(m+1)^2} \right)$$
$$+ \frac{\beta_k}{\alpha_k} \left(\frac{(m+2)u_l^{m+1}(\omega_l - \omega_{l-1})(u_l - u_{l-1})^2(m+1)(m+2)}{(u_l - u_{l-1})^4(m+1)^2(m+2)^2} \right)$$
$$- \frac{\beta_k}{\alpha_k} \left(\frac{(u_l^{m+2} - u_{l-1}^{m+2})(\omega_l - \omega_{l-1})2(u_l - u_{l-1})(m+1)(m+2)}{(u_l - u_{l-1})^2(m+1)^2(m+2)^2} \right)$$
$$= \frac{\beta_k}{\alpha_k} \left(\frac{u_l^{m+1}(\omega_l - \omega_{l-1})}{(u_l - u_{l-1})^2(m+1)} + \frac{u_l^{m+1}(\omega_l - \omega_{l-1})}{(u_l - u_{l-1})^2(m+1)} \right)$$
$$- \frac{\beta_k}{\alpha_k} \left(\frac{(u_l^{m+2} - u_{l-1}^{m+2})(\omega_l - \omega_{l-1})2}{(u_l - u_{l-1})(m+1)(m+2)} \right)$$

$$\frac{\partial \Phi}{\partial u_l \partial u_l}$$

$$
\begin{aligned}
&= -\frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_l}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)(m+1)} + \frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)(m+2)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_l}\left(\frac{u_l^{m+1}(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})(m+1)} - \frac{(u_l^{m+2}-u_{l-1}^{m+2})(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^2(m+1)(m+2)}\right)\\
&= -\frac{\beta_k}{\alpha_k}\left(\frac{(m+1)u_l^m(\omega_{l+1}-\omega_l)(u_{l+1}-u_l)(m+1)}{(u_{l+1}-u_l)^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)(m+1)}{(u_{l+1}-u_l)^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_l^{m+1}(\omega_{l+1}-\omega_l)(u_{l+1}-u_l)^2(m+1)(m+2)}{(u_{l+1}-u_l)^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)2(u_{l+1}-u_l)(m+1)(m+2)}{(u_{l+1}-u_l)^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(m+1)u_l^m(\omega_l-\omega_{l-1})(u_l-u_{l-1})(m+1)}{(u_l-u_{l-1})^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_l-\omega_{l-1})(m+1)}{(u_l-u_{l-1})^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_l^{m+1}(\omega_l-\omega_{l-1}(u_l-u_{l-1})^2(m+1)(m+2))}{(u_l-u_{l-1})^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_l^{m+2}-u_{l-1}^{m+2}2(u_l-u_{l-1})(m+1)(m+2))(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^4(m+1)^2(m+2)^2}\right)\\
&= -\frac{\beta_k}{\alpha_k}\left(\frac{u_l^m(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)} + \frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)} + \frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)2}{(u_{l+1}-u_l)^3(m+1)(m+2)} + \frac{u_l^m(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^2(m+1)} + \frac{u_l^{m+1}(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^2(m+1)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_l^{m+2}-u_{l-1}^{m+2})2(\omega_l-\omega_{l-1})}{(u_l-u_{l-1})^3(m+1)(m+2)}\right)
\end{aligned}
$$

$$\frac{\partial\Phi}{\partial u_{l+1}\partial u_l}$$

$$\begin{aligned}
&= -\frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_{l+1}}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)(m+1)} - \frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)(m+2)}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)(m+1)}{(u_{l+1}-u_l)^2(m+1)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_{l+1}^{m+1}(\omega_{l+1}-\omega_l)(u_{l+1}-u_l)^2(m+1)(m+2)}{(u_{l+1}-u_l)^4(m+1)^2(m+2)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)2(u_{l+1}-u_l)(m+1)(m+2)}{(u_{l+1}-u_l)^4(m+1)^2(m+2)^2}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_l^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)} + \frac{u_{l+1}^{m+1}(\omega_{l+1}-\omega_l)}{(u_{l+1}-u_l)^2(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_{l+1}^{m+2}-u_l^{m+2})(\omega_{l+1}-\omega_l)2}{(u_{l+1}-u_l)^3(m+1)(m+2)}\right)
\end{aligned}$$

$$\frac{\partial\Phi}{\partial u_1\partial u_1}$$

$$\begin{aligned}
&= \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_1}\left(\frac{u_1^m(\omega_1-\omega_0)}{(m+2)} - \frac{u_1^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)(m+1)} + \frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)(m+2)}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{mu_1^{m-1}(\omega_1-\omega_0)}{(m+2)} - \frac{u_1^{m+1}(\omega_1-\omega_0)(m+1)}{(u_2-u_1)^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(m+1)u_1^m(\omega_1-\omega_0)(u_2-u_1)(m+1)}{(u_2-u_1)^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)(u_1^{m+1})(\omega_1-\omega_0)(u_2-u_1)^2(m+1)(m+2)}{(u_2-u_1)^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)2(u_2-u_1)(m+1)(m+2)}{(u_2-u_1)^4(m+1)^2(m+2)^2}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{mu_1^{m-1}(\omega_1-\omega_0)}{(m+2)} - \frac{u_1^m(\omega_1-\omega_0)}{(u_2-u_1)} - \frac{u_1^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m+1})(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)} - \frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)2}{(u_2-u_1)^3(m+1)(m+2)}\right)
\end{aligned}$$

$$\frac{\partial\Phi}{\partial u_2 \partial u_1}$$

$$\begin{aligned}
&= -\frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_2}\left(\frac{u_1^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)(m+1)} - \frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)(m+2)}\right)\\
&= +\frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m+1}(\omega_1-\omega_0)(m+1)}{(u_2-u_1)^2(m+1)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_2^{m+1}(\omega_1-\omega_0)(u_2-u_1)^2(m+1)(m+2)}{(u_2-u_1)^4(m+1)^2(m+2)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)(u_2-u_1)(m+1)(m+2)}{(u_2-u_1)^4(m+1)^2(m+2)^2}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_1^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)} + \frac{u_2^{m+1}(\omega_1-\omega_0)}{(u_2-u_1)^2(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_2^{m+2}-u_1^{m+2})(\omega_1-\omega_0)2}{(u_2-u_1)^3(m+1)(m+2)}\right)
\end{aligned}$$

$$\frac{\partial\Phi}{\partial u_{N-2} \partial u_{N-1}}$$

$$\begin{aligned}
&= \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_{N-2}}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})(m+1)} - \frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)(m+2)}\right)\\
&= +\frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})(m+1)}{(u_{N-1}-u_{N-2})^2(m+1)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_{N-2}^{m+1}(\omega_{N-1}-\omega_{N-2})(u_{N-1}-u_{N-2})^2(m+1)(m+2)}{(u_{N-1}-u_{N-2})^4(m+1)^2(m+2)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})2(u_{N-1}-u_{N-2})(m+1)(m+2)}{(u_{N-1}-u_{N-2})^4(m+1)^2(m+2)^2}\right)\\
&= +\frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)} + \frac{u_{N-2}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})2}{(u_{N-1}-u_{N-2})^3(m+1)(m+2)}\right)
\end{aligned}$$

$$\frac{\partial\Phi}{\partial u_{N-1}\partial u_{N-1}}$$

$$
\begin{aligned}
&= \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_{N-1}}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_{N-1}}\left(\frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)(m+2)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\frac{\partial}{\partial u_{N-1}}\left(\frac{u_{N-1}^{m}(\omega_{N}-\omega_{N-1})}{(m+2)}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{(m+1)u_{N-1}^{m}(\omega_{N-1}-\omega_{N-2})(u_{N-1}-u_{N-2})(m+1)}{(u_{N-1}-u_{N-2})^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})(m+1)}{(u_{N-1}-u_{N-2})^2(m+1)^2}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{(m+2)u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})(u_{N-1}-u_{N-2})^2(m+1)(m+2)}{(u_{N-1}-u_{N-2})^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})2(u_{N-1}-u_{N-2})(m+1)(m+2)}{(u_{N-1}-u_{N-2})^4(m+1)^2(m+2)^2}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{mu_{N-1}^{m-1}(\omega_{N}-\omega_{N-1})}{(m+2)}\right)\\
&= \frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})} - \frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)}\right)\\
&- \frac{\beta_k}{\alpha_k}\left(\frac{u_{N-1}^{m+1}(\omega_{N-1}-\omega_{N-2})}{(u_{N-1}-u_{N-2})^2(m+1)} - \frac{(u_{N-1}^{m+2}-u_{N-2}^{m+2})(\omega_{N-1}-\omega_{N-2})2}{(u_{N-1}-u_{N-2})^3(m+1)(m+2)}\right)\\
&+ \frac{\beta_k}{\alpha_k}\left(\frac{mu_{N-1}^{m-1}(\omega_{N}-\omega_{N-1})}{(m+2)}\right)
\end{aligned}
$$

Appendix C

Discretizations needed in Chapter 3

In this appendix, we detail the calculations for the coefficients of the matrix (4.14), and the Hessian of the discrete entropy (4.17), both in the case $\alpha = -1$.

C.1 Computation of the coefficients M_{ij}

We compute the coefficients of the matrix (4.15), i.e. the coefficients a_{ij}, b_{ij}, and c_{ij} defined in (4.14). In the following, we set

$$\delta_j = \omega_j - \omega_{j-1}, \quad \Delta_j = \frac{1}{2}(\omega_{j+1} - \omega_{j-1}), \quad \sigma_j = \frac{1}{3}(\omega_{j+1} + \omega_j + \omega_{j-1}).$$

Lemma C.1.1 (Coefficients a_{ij})**.** *The coefficients of the symmetric matrix $A = (a_{ij})$, defined in* (4.14)*, read as*

$$\begin{aligned}
a_{jj} &= \Delta_j^2(M - \sigma_j) - \frac{\Delta_j}{60}(12\Delta_j^1 + \delta_j^2 + \delta_{j+1}^2), \quad 1 \le j \le N-1,\\
a_{j,j+1} &= \Delta_j \Delta_{j+1}(M - \sigma_{j+1}) - \frac{\delta_j^3}{120}, \quad 1 \le j \le N-1,\\
a_{jk} &= \Delta_j \Delta_k (M - \sigma_k), \quad j+2 \le k \le N-1,\\
a_{1N} &= \frac{1}{2}\Delta_1 \Delta_N \left(M - \frac{\omega_2}{3} \right) - \frac{\Delta_N^3}{120},\\
a_{jN} &= \frac{1}{2}\Delta_j \Delta_N \left(M - \sigma_j + \frac{\delta_N}{3} \right), \quad 2 \le j \le N-2,\\
a_{N-1,N} &= \frac{1}{2}\Delta_{N-1}\Delta_N \left(M - \frac{1}{3}(\omega_{N-2} + 2\omega_{N-1}) \right) - \frac{\Delta_N^3}{120},\\
a_{NN} &= \frac{M}{4}\Delta_N^2 + \frac{\Delta_N^3}{10}.
\end{aligned}$$

Proof. We reformulate the integral in (4.14):

$$a_{jk} = M \int_0^M \phi_j(\eta)\, d\eta \int_0^M \phi_k(\eta')\, d\eta' - \int_0^M \phi_j(\eta)\, d\eta \int_0^M \eta' \phi_k(\eta')\, d\eta' - J_{jk},$$
$$\text{where } J_{jk} = \int_0^M \int_0^M (\eta - \eta')_+ \phi_j(\eta) \phi_k(\eta')\, d\eta d\eta',$$

where $(\eta - \eta')_+ := \max\{0, \eta - \eta'\}$. The first two integrals become

$$\int_0^m \phi_j(\eta)\, d\eta = \delta_j, \qquad \int_0^M \eta' \phi_k(\eta')\, d\eta' = \Delta_k \sigma_k.$$

Since A is symmetric, it is sufficient to consider $1 \le j \le k$. If $j+2 \le k < N$, the support of $\phi_j(\eta)\phi_k(\eta')$ is contained in $[\omega_{j-1}, \omega_{j+1}] \times [\omega_{k-1} \times \omega_{k+1}]$. Hence, the support is non-vanishing if $\eta \le \omega_{j+1} \le \omega_{k-1} \le \eta'$, but then $(\eta - \eta')_+ = 0$ except for $\eta = \eta'$. We conclude that $J_{jk} = 0$ and it is sufficient to compute only J_{jj} and $J_{j,j+1}$:

$$J_{jj} = \int_{\omega_{j-1}}^{\omega_{j+1}} \phi_j(\eta) \left(\int_{\omega_{j-1}}^{\eta} (\eta - \eta') \phi_j(\eta')\, d\eta' \right) d\eta = \frac{\Delta_j}{60}(12\Delta_j^2 + \delta_j^2 + \delta_{j+1}^2),$$
$$J_{j,j+1} = \int_{\omega_{j-1}}^{\omega_{j+1}} \phi_j(\eta) \left(\int_{\omega_{j-1}}^{\max\{\eta, \omega_j\}} (\eta - \eta') \phi_j(\eta')\, d\eta' \right) d\eta = \frac{\delta_j^3}{120}.$$

Next, let $k = N$. Then the support of ϕ_N is contained in $[0, \omega_1] \cup [\omega_{N-1}, M]$, and we compute:

$$\begin{aligned} a_{jN} &= M \int_0^M \phi_j(\eta)\, d\eta \int_0^M \phi_N(\eta')\, d\eta' - \int_0^M \eta \phi_j(\eta)\, d\eta \int_0^{\omega_1} \phi_N(\eta')\, d\eta' \\ &\quad - \int_0^M \phi_j(\eta)\, d\eta \int_{\omega_{N-1}}^M \eta' \phi_N(\eta')\, d\eta' - K_j^+ - K_j^- \\ &= \frac{1}{2} \Delta_j \Delta_N \left(M - \sigma_j + \frac{\delta_N}{3} \right) - K_j^+ - K_j^-, \end{aligned}$$

where

$$K_j^+ := \int_0^M \int_\eta^{\omega_1} (\eta' - \eta)_+ \phi_j(\eta) \phi_N(\eta')\, d\eta d\eta',$$
$$K_j^- := \int_0^M \int_{\omega_{N-1}}^M (\eta - \eta')_+ \phi_j(\eta) \phi_N(\eta')\, d\eta d\eta'.$$

For $2 \le j \le N-2$, the supports of ϕ_j and ϕ_N do not intersect such that $K_j^\pm = 0$. For $j = 1$, we have $K_1^- = 0$ and $K_1^+ = \Delta_N^3/120$, whereas for $j = N-1$,

$K^+_{N-1} = 0$ and $K^-_{N-1} = \Delta_N^3/120$. Furthermore, $K^\pm_N = \Delta_N^3/30$. Moreover, since $\delta_N = \omega_N - \omega_{N-1} = (M - \omega_0) - (M - \omega_1) = \omega_1$,

$$M - \sigma_1 + \frac{\delta_N}{3} = M - \frac{1}{3}(\omega_0 + \omega_1 + \omega_2) + \frac{1}{3}\omega_1 = M - \frac{\omega_2}{3}.$$

Collecting these results, the lemma follows. □

Lemma C.1.2 (Coefficients b_{ij}). *The coefficients of the matrix $B = (b_{ij})$, defined in* (4.14), *read as*

$$\begin{aligned}
b_{jj} &= \frac{2}{3}\delta_j(M\Delta_j - \delta_j\sigma_j) + \beta_{jj}, \quad 1 \le j \le N,\\
b_{j+1,j} &= \frac{1}{3}\big(2M\delta_{j+1}\Delta_j - (\omega_{j+2}^2 - \omega_{j+1}^2)\Delta_j\big) + \beta_{j+1,j}, \quad 1 \le j \le N-1,\\
b_{jk} &= \frac{2}{3}(M\delta_j\Delta_k - \delta_j\delta_k\sigma_k), \quad 1 \le j < k \le N,\\
b_{jk} &= \frac{1}{3}\big(2M\delta_j\Delta_k - (\omega_{j+1}^2 - \omega_j^2)\Delta_k\big), \quad j \ge k-2,\\
b_{1N} &= \frac{2}{3}\delta_1\Delta_N - \frac{1}{6}(\omega_2^2 - \omega_1^2)\delta_N - \frac{1}{9}(2\omega_N^2 - \omega_{N-1}^2 - \omega_{N-1}\omega_N)\delta_1 - \beta_{1N},\\
b_{jN} &= \frac{2}{3}\delta_j\Delta_N - \frac{1}{6}(\omega_{j+1}^2 - \omega_j^2)\delta_N\\
&\quad - \frac{1}{9}(2\omega_N^2 - \omega_{N-1}^2 - \omega_{N-1}\omega_N)\delta_j, \quad 2 \le j \le N-1,\\
b_{NN} &= \frac{2}{3}\delta_N\Delta_N - \frac{1}{6}(\omega_N^2 - \omega_{N-1}^2)\delta_N - \frac{1}{9}(2\omega_N^2 - \omega_{N-1}^2 - \omega_{N-1}\omega_N)\delta_N - \beta_{NN},
\end{aligned}$$

where

$$\begin{aligned}
\beta_{jj} &= -\frac{1}{45}\omega_{j+1}^2 + \frac{1}{90}\omega_{j+1}^2\omega_j + \frac{1}{18}\omega_{j+1}^2\omega_{j+2} - \frac{1}{15}\omega_{j+1}\omega_j^2 + \frac{1}{9}\omega_{j+2}\omega_{j+1}\omega_j\\
&\quad - \frac{1}{9}\omega_{j+2}^2\omega_{j+1} + \frac{7}{90}\omega_j^3 - \frac{1}{6}\omega_j^2\omega_{j+2} + \frac{1}{9}\omega_j\omega_{j+2}^2,\\
\beta_{j+1,j} &= \frac{1}{45}\left(\omega_{j+2}^3 - \omega_{j+1}^3 + 3(\omega_{j+1}^2\omega_{j+2} - \omega_{j+2}^2\omega_{j+1})\right),\\
\beta_{1N} &= \frac{1}{45}\big(\omega_2^2 - \omega_1^2 + 3(\omega_2\omega_1^2 - \omega_2^2\omega_1)\big),\\
\beta_{NN} &= \frac{1}{45}\big(\omega_{N+1}^3 - \omega_N^3 + 3(\omega_N^2\omega_{N+1} - \omega_{N+1}^2\omega_N)\big).
\end{aligned}$$

Proof. The computation is similar to the previous proof. We write the integral for b_{jk} with $j,\, k \le N-1$ and for $j < k$ as

$$b_{jk} = M\int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \phi_k(\eta')\,d\eta' - \int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \eta'\phi_k(\eta')\,d\eta' - J^-_{jk},$$

and for $j > k$ as

$$b_{jk} = M\int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \phi_k(\eta')\,d\eta' - \int_0^M \eta\phi_{N+j}(\eta)\,d\eta \int_0^M \phi_k(\eta)\,d\eta' - J_{jk}^+,$$

where

$$J_{jk}^- = \int_0^M \int_0^M (\eta-\eta')_+ \phi_{N+j}(\eta)\phi_k(\eta')\,d\eta d\eta',$$
$$J_{jk}^+ = \int_0^M \int_0^M (\eta'-\eta)_+ \phi_{N+j}(\eta)\phi_k(\eta')\,d\eta d\eta'.$$

We compute

$$\int_0^M \phi_{N+j}(\eta)\,d\eta = \frac{2}{3}\delta_j, \quad \int_0^M \eta\phi_{N+j}(\eta)\,d\eta = \frac{1}{3}(\omega_{j+1}^2 - \omega_j^2).$$

The integrals $J_{jk}^{\pm}$ vanish if $k \neq j-1, j$ since $(\eta-\eta')_+$ and $(\eta'-\eta)_+$ vanish. This proves the expressions for b_{jk} with $j \leq k-1$ and $j \geq k+2$. The coefficients b_{jj} and $b_{j+1,j}$ are calculated in the same way.

It remains to compute the matrix coefficients coming from the boundary elements. The computation of b_{jN} is straightforward as the support of ϕ_{2N} is contained on the single subinterval $[\omega_{N-1}, M]$. For the boundary elements b_{jN}, we take into account that the support of ϕ_N is contained in $[\omega_{N-1}, M]$ and $[0, \omega_1]$, which yields

$$b_{jN} = M\int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \phi_N(\eta')\,d\eta' - \int_0^M \eta\phi_{N+j}(\eta)\,d\eta \int_0^{\omega_1} \phi_N(\eta')\,d\eta'$$
$$- \beta_{1N} - \beta_{NN},$$

where

$$\beta_{1N} = \int_0^M \int_0^{\omega_1} (\eta'-\eta)_+ \phi_{N+j}(\eta)\phi_N(\eta')\,d\eta d\eta',$$
$$\beta_{NN} = \int_0^M \int_{\omega_{N-1}}^M (\eta-\eta')_+ \phi_j(\eta)\phi_N(\eta')\,d\eta d\eta',$$

and the computation is as before. □

Lemma C.1.3 (Coefficients c_{ij}). *The coefficients of the symmetric matrix $C = (c_{ij})$, defined in* (4.14)*, read as*

$$c_{jj} = \frac{8}{9}\delta_j^2 - \frac{2}{9}\delta_j(\omega_{j+1}^2 - \omega_j^2)$$
$$- \frac{2}{35}\big(\omega_{j+1}^3 - \omega_j^3 - 3\omega_{j+1}\omega_j(\omega_{j+1}-\omega_j)\big), \quad 1 \leq j \leq N,$$
$$c_{jk} = \frac{8}{9}\delta_j\delta_k - \frac{2}{9}\delta_j(\omega_{k+1}^2 - \omega_k^2), \quad 1 \leq j < k \leq N.$$

Proof. We compute

$$c_{jk} = M\int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \phi_{N+k}(\eta')\,d\eta' - \int_0^M \phi_{N+j}(\eta)\,d\eta \int_0^M \eta'\phi_{N+k}(\eta')\,d\eta' - \gamma_{jk},$$

$$\text{where } \gamma_{jk} = \int_0^M\int_0^M (\eta-\eta')_+\phi_{N+j}(\eta)\phi_{N+k}(\eta')\,d\eta d\eta'.$$

As before, we find that $\gamma_{jk}=0$ for all $j\neq k$. Moreover,

$$\begin{aligned}\gamma_{jj} &= \int_{\omega_{j-1}}^{\omega_j} \phi_{N+j}(\eta)\,d\eta \int_{\omega_j}^{\eta}(\eta-\eta')\phi_{N+j}(\eta')\,d\eta' \\ &= \frac{2}{35}\big(\omega_{j+1}^3-\omega_j^3-3(\omega_{j+1}\omega_j(\omega_{j+1}-\omega_j))\big),\end{aligned}$$

which finishes the proof. □

C.2 Computation of the coefficients of the Hessian of S_N

We compute the gradient and Hessian of the discrete entropy (4.17) for the case $\alpha=-1$. We set for $k=0,\dots,N-1$:

$$S_{N,k}[\mathbf{g}] = \frac{1}{2}\int_{\omega_k}^{\omega_{k+1}} \big(g_k\phi_k(\omega)+g_{k+1}\phi_{k+1}(\omega)+g_{N+k}\phi_{N+k}(\omega)\big)\,d\omega,$$

where $\mathbf{g}=(g_1,\dots,g_{2N})\in\mathbb{G}_M^N$. Furthermore, we abbreviate $\partial_k S_{N,j}=\partial S_{N,j}/\partial g_k$ and $\partial_{j,k}S_{N,\ell}=\partial S_{N,\ell}/\partial g_j\partial_k g$. A computation shows that

$$\begin{aligned}\partial_k S_{N,k}[\mathbf{g}] &= \frac{\delta_{k+3}}{3}\big(2(g_{N+k}+g_k)+g_{k+1}\big),\\ \partial_k S_{N,k-1}[\mathbf{g}] &= \frac{2}{3}\delta_k(g_k-g_{k-1}-g_{N+k-1}),\\ \partial_{N+k}S_{N,k}[\mathbf{g}] &= \delta_{k+1}\left((g_{k+1}+g_k)+\frac{8}{5}g_{N+k}\right).\end{aligned}$$

As $S_{N,k}$ and $S_{N,k-1}$ depend on g_k, we obtain (recall (4.17))

$$\partial_k S_N = \partial_k S_{N,k}+\partial_k S_{N,k-1}, \quad \partial_{N+k}S_N = \partial_{N+k}S_{N,N+k}.$$

The second-order derivatives become

$$\begin{aligned}&\partial_{k,k-1}S_{N,k-1} = -\frac{2}{3}\delta_k, \qquad \partial_{k,k}S_{N,k-1} = \frac{2}{3}\delta_k, \qquad \partial_{k,k}S_{N,k} = -\frac{2}{3}\delta_{k+1},\\ &\partial_{k,k+1}S_{N,k} = \frac{1}{3}\delta_{k+1}, \quad \partial_{k,N+k-1}S_{N,k-1} = -\frac{2}{3}\delta_k, \quad \partial_{k,N+k}S_{N,k} = \frac{2}{3}\delta_{k+1},\\ &\partial_{N+k,k+1}S_{N,k} = \frac{2}{3}\delta_{k+1}, \qquad \partial_{N+k,N+k}S_{N,k} = \frac{16}{15}\delta_{k+1}.\end{aligned}$$

Then the elements of the Hessian of S_N read as

$$\begin{aligned}
\partial_{k,k-1}S_N &= \partial_{k,k-1}S_{N,k-1}, & \partial_{k,k}S_N &= \partial_{k,k}S_{N,k} + \partial_{k,k}S_{N,k-1}, \\
\partial_{k,k+1}S_N &= \partial_{k,k+1}S_{N,k}, & \partial_{k,N+k}S_N &= \partial_{k,N+k}S_{N,k}, \\
\partial_{k,N+k-1}S_N &= \partial_{k,N+k-1}S_{N,k-1}, & \partial_{N+k,k+1}S_N &= \partial_{N+k,k+1}S_{N,k}, \\
\partial_{N+k,N+k}S_N &= \partial_{N+k,N+k}S_{N,k}.
\end{aligned}$$

Bibliography

Bibliography

[1] M. Agueh and M. Bowles. One-dimensional numerical algorithms for gradient flows in the p-Wasserstein spaces. *Acta Appl. Math.* 125 (2013), 121-134.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures,* Second edition, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 2008.

[3] A. Arnold, *Mathematical properties of quantum evolution equations*, in "Quantum Transport Modelling, Analysis and Asymptotics" (eds. G. Allaire, A. Arnold, P. Degond and T. Y. Hou), Lecture Notes in Mathematics, **1946**, Springer, Berlin, (2008), 45–109.

[4] V. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. Inst. Fourier (Grenoble), **16** (1966), 319–361.

[5] Claudio Baiocchi, Michel Crouzeix. On the equivalence of a-stability and g-stability. [Research Report] RR-0609, 1987. ¡inria-00075945¿

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

[7] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations.*. Springer, New York, 2011.

[8] H. Brezis, *Operateur maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert,* Math. Studies, North-Holland, Amsterdam, 1973.

[9] S. Brull and F. Méhats, *Derivation of viscous correction terms for the isothermal quantum Euler model*, Z. Angew. Math. Mech., **90** (2010), 219–230.

[10] B. van Brunt, "The Calculus of Variations," Universitext, Springer-Verlag, New York, 2004.

[11] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.* 105 (1989), 397-400.

[12] 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 (2008), 691-721.

[13] M. Bonforte, J. Dolbeault, G. Grillo, and J. L. Vázquez. Sharp rates of decay of solutions to the nonlinear fast diffusion equation via functional inequalities. *Proc. Natl. Acad. Sci. USA* 107 (2010), 16459-16464.

[14] C. Budd, M. Cullen, and E. Walsh. Monge-Ampère based moving mesh methods for numerical weather prediction, with applications to the Eady problem. *J. Comput. Phys.* 236 (2013), 247-270.

[15] M. Burger, J. A. Carrillo, and M.-T. Wolfram. A mixed finite element method for nonlinear diffusion equations. *Kinetic Related Models* 3 (2010), 5983.

[16] M. Burger, M. Franeka, and C.-B. Schönlieb. Regularised regression and density estimation based on optimal transport. *Appl. Math. Res. Express* 2 (2012), 209-253.

[17] J. A. Carrillo, A. Chertock, and Y. Huang. A finite-volume method for non-linear nonlocal equations with a gradient flow structure. *Commun. Comput. Phys.* 17 (2015), 233-258.

[18] J. A. Carrillo, R. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Rational Mech. Anal.* 179 (2006), 217-263.

[19] P. Dirac, *The Lagrangian in quantum mechanics*, Phys. Z. Sowjet., **3** (1933), 64–72.

[20] (MR0403371) Dj. Djukic and B. Vujanović, *Noether's theory in classical nonconservative mechanics*, Acta Mech., **23** (1975), 17–27.

[21] E. De Giorgi, New problems on minimizing movements. In: C. Baiocchi and J.-L. Lions (eds.), *Boundary Value Problems for PDE and Applications*, pp. 81-98. Masson, Paris, 1993.

[22] B. Düring, D. Matthes, and J.-P. Milišić. A gradient flow scheme for nonlinear fourth order equations. *Discrete Cont. Dyn. Sys. B* 14 (2010), 935-959.

[23] Detlef Dürr, Stefan Teufel. *Bohmian Mechanics.* American Mathematical Society, 1998.

[24] Lawrence C. Evans. *Partial Differential Equations.* Springer, Berlin, 2009.

[25] E. Feireisl, "Dynamics of Viscous Compressible Fluids," Oxford Lecture Series in Mathematics and its Applications, **26**, Oxford University Press, Oxford, 2004.

[26] J. Feng and T. Nguyen, *Hamilton-Jacobi equations in space of measures associated with a system of conservation laws*, J. Math. Pures Appl. (9), **97** (2012), 318–390.

[27] L. Brown, ed., *"Feynman's Thesis. A New Approach to Quantum Theory,"* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005.

[28] T. Frankel, "The Geometry of Physics. An Introduction," Cambridge University Press, Cambridge, 1997.

[29] G. Frederico and D. Torres, *Nonconservative Noether's theorem in optimal control*, Intern. J. Tomogr. Stat., **5** (2007), 109–114.

[30] J.-L. Fu and L.-Q. Chen, *Non-Noether symmetries and conserved quantities of nonconservative dynamical systems*, Phys. Lett. A, **317** (2003), 255–259.

[31] U. Gianazza, G. Savaré, and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Rational Mech. Anal.* 194 (2009), 133-220.

[32] N. Gozlan and C. Léonard. Transport inequalities. A survey. *Markov Processes Related Fields* 16 (2010), 635-736.

[33] W.H. Hundsdorfer & B.I. Steininger. Convergence of linear multistep methods and one-leg methods for stiff nonlinear initial value problems. *BTI.* vol. 31, p 124-143.

[34] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II* Springer, Berlin, 2002.

[35] A. Jüngel, *"Transport Equations for Semiconductors,"* Lecture Notes in Physics, **773**, Springer-Verlag, Berlin, 2009.

[36] A. Jüngel, *Global weak solutions to compressible Navier-Stokes equations for quantum fluids*, SIAM J. Math. Anal., **42** (2010), 1025–1045.

[37] A. Jüngel and J.-P. Milišić. Entropy dissipative one-leg multistep time approximations of nonlinear diffusive equations. *Numer. Meth. Part. Diff. Eqs.* 31 (2015), 1119-1149.

[38] A. Jüngel, J. L. López and J. Montejo-Gámez, *A new derivation of the quantum Navier-Stokes equations in the Wigner-Fokker-Planck approach*, J. Stat. Phys., **145** (2011), 1661–1673.

[39] A. Jüngel and J.-P. Milišić, *Full compressible Navier-Stokes equations for quantum fluids: Derivation and numerical solution*, Kinetic Related Models, **4** (2011), 785–807.

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

[41] D. Kinderlehrer and N. Walkington. Approximation of parabolic equations using the Wasserstein metric. *ESAIM Math. Model. Numer. Anal.* 33 (1999), 837-852.

[42] J. Lafferty, *The density manifold and configuration space quantization*, Trans. Amer. Math. Soc., **305** (1988), 699–741.

[43] J. Lott, *Some geometric calculations on Wasserstein space*, Commun. Math. Phys., **277** (2008), 423–437.

[44] D. Matthes and H. Osberger. Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.* 48 (2014), 697-726.

[45] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var.* 48 (2013), 1-31.

[46] E. Madelung, *Quantentheorie in hydrodynamischer Form*, Z. Phys., **40** (1926), 322–326.

[47] P. Markowich, T. Paul and C. Sparber, *Bohmian measures and their classical limit*, J. Funct. Anal., **259** (2010), 1542–1576.

[48] R. McCann, *Polar factorization of maps on Riemannian manifolds*, GAFA Geom. Funct. Anal., **11** (2001), 589–608.

[49] E. Nelson, *Derivation of the Schrödinger equation from Newtonian mechanics*, Phys. Rev., **150** (1966), 1079–1085.

[50] H. Osberger. Long-time behaviour of a fully discrete Lagrangian scheme for a family of fourth order. Preprint, 2015. `arXiv:1501.04800`.

[51] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Commun. Part. Diff. Eqs.* 26 (2001), 101-174.

[52] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., **173** (2000), 361–400.

[53] G. Peyré. Entropic Wasserstein gradient flows. Preprint, 2015. `arXiv:1502.06216.`

[54] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics.* Second edition. Springer, Berlin, 2007.

[55] M.-K. von Renesse, *On optimal transport view on Schrödinger's equation*, Canad. Math. Bull., **55** (2012), 858–869.

[56] G. Rosen. Nonlinear heat conduction in solid H_2. *Phys. Rev. B* 19 (1979), 2398-2399.

[57] W. Sarlett and F. Cantrijn, *Generalization of Noether's Theorem in classical mechanics*, SIAM Review, **23** (1981), 467–494.

[58] R. Talman, "Geometric Mechanics," Wiley, New York, 2000.

[59] J. L. Vazquez. Nonexistence of solutions for heat nonlinear equations of fast-diffusion type. *J. Math. Pures Appl.* 71 (1992), 503-526.

[60] J. L. Vazquez. *Smooting and Decay Estimates for Nonlinear Parabolic Equations of Porous Medium Type.* Oxford Lecture Series in Mathematics and Its Applications 33, Oxford University Press, Oxford, 2006.

[61] C. Villani. *Topics in Optimal Transportation.* Graduate Studies in Mathematics 58, American Mathematical Society, Providence, RI, 2003.

[62] C. Villani. *Optimal Transport. Old and New.* Springer, Berlin, 2009.

[63] Robert E. Wayet. *Quantum Dynamics with Trajectories.* Springer, New York, 2005.

[64] M. Westdickenberg and J. Wilkening. Variational particle schemes for the porous medium equation and for the system of isentropic Euler equations. *ESAIM Math. Model. Numer. Anal.* 44 (2010), 133-166.

Acknowledgements

I would like to express my gratitude to all those, who supported me over the last years and helped me writing this thesis. First of all, I want to thank Professor Ansgar Jüngel for his supervision, his professional, financial and scientific support and the pleasant working atmosphere. I would also like to thank Bertram Düring for the fruitful collaboration and the scientific and mental support within the last years. Furthermore, I want to thank my colleagues for several hours of discussion, partly far beyond mathematics. Also, I would like to thank the Vienna University of Technology and the Austrian Science Fund FWF for financing the projects I was working on. Special thanks go to my parents Marianne and Josef who provided all the opportunities that paved they way to write this thesis, and to my girlfriend Lena who was encouraging and supportive every single day the last years.

Curriculum vitae

Personal information

Name	**Philipp Fuchs**
Address	**Grundsteingasse 5/15,**
	1160 Vienna, Austria
Date of birth	**August 24th, 1982**
Place of birth	**Braunau am Inn, Austria**
Citizenship	**Austrian**

Education

since 10/2010	**PhD student**, Vienna University of Technology
10/2003 – 06/2010	**Student of mathematics**, University of Vienna diploma '*with distinction*'
10/2002 – 10/2003	**Studies of mathematics and physics**, University of Innsbruck
06/2001	**Matura**, Handelsakademie Braunau

Civilian service

10/2001 – 10/2002	**Krankenhaus St. Josef**, Braunau

Occupation

18/2009 – 07/2010	**University of Vienna**, student assistant
10/2010 – 10/2011	**Vienna University of Technology**, project assistant, Institute for Analysis and Scientific Computing
10/2011 – 10/2015	**Vienna University of Technology**, university assistant, Institute for Analysis and Scientific Computing

Teaching experience

10/2011 – 10/2015	**Exercise courses in:** *Partial Differential Equations, Analysis II, Analysis III, Calculus of Variations,* Vienna University of Technology

Talks

11/2011	**Ludwig Maximilians University München** Perspectives in Optimal Transport Invited Speaker: On the Lagrangian Structure of Quantum Fluid Models

Skills

Languages	**English**(full professional proficiency), **French** (fluently)
Programming	**Matlab, Mathematica, Maple, LaTex, C/C++**