TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Entropy-preserving numerical schemes for nonlinear diffusion equations

Translating from the continuous to the discrete case

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

Prof. Dr. Ansgar Jüngel
E101 – Institut für Analysis und Scientific Computing – TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation
von

Dipl.Ing. Stefan Schuchnigg
Matrikelnummer: 0426505
Buchengasse 155/150, 1100 Wien

Diese Dissertation haben begutachtet:

1. **Prof. Dr. Etienne Emmrich**
   Institut für Mathematik, Technische Universität Berlin
2. **Prof. Dr. Francis Filbet**
   Institut de Mathématiques de Toulouse
3. **Prof. Dr. Ansgar Jüngel**
   Institut für Analysis und Scientific Computing, TU Wien

Wien, am 1. März 2017

Abstract

The behavior of certain entropy or energy functionals of solutions of partial differential equations is widely regarded as an intrinsic property of the underlying equations. To this end it is of great importance to find numerical schemes that offer the same entropic structures as their continuous counterpart. Due to the often nonlinear nature of both the entropies and the partial differential equations, it is a highly nontrivial task to "translate" these properties and methods to a discrete level. In this thesis we present some examples of such translations where we are able to develop tools on a discrete level that allow us to achieve similar results as are known on the continuous level. The first main result are the discrete Beckner inequalities together with a discrete nonlinear integration-by-parts formula that allow us to mimic the entropy decay of the porous-medium equation for a finite volume scheme for a wide set of parameters. Regarding the time (semi)discretization, we prove conditions on the abstract Cauchy operator under which solutions to the associated Cauchy problem feature the same entropic behavior for Runge-Kutta schemes in time as in the continuous case. This very general approach can be used on various problems, for example the porous-medium equation, a linear diffusion system and the Derrida-Lebowitz-Speer-Spohn equation. Finally, by making use of a carefully constructed discrete scheme and the already mentioned discrete nonlinear integration-by-parts formula, we give a discrete analogon to the continuous Bakry-emery approach for the porous-medium equation.

Kurzfassung

Das Verhalten von bestimmten Entropie- beziehungsweise Energiefunktionalen von partiellen Differntialgleichungen wird üblicherweise als eine intrinsische Eigenschaft der zugrundeliegenden Gleichung angesehen. Aus diesem Grund ist es von großem Interesse, numerische Verfahren zu entwickeln, die die selben Entropiestrukturen aufweisen wie ihre kontinuierlichen Pendants. Aufgrund der zumeist nichtlinearen Natur sowohl der Entropien als auch der partiellen Differentialgleichungen, ist es eine nichttriviale Aufgabe, diese Eigenschaften auf ein diskretes Level zu "übersetzen". In dieser Arbeit präsenteren wir einige Beispiele von derartigen Übersetzungen, im Rahmen derer wir diskrete Werkzeuge entwickeln können, die es uns erlauben, die aus dem kontiuierlichen bekannten Resultate nachzuahmen. Das erste Hauptresultat sind dabei diskrete Beckner Ungleichungen, die es uns zusammen mit einer nichtlinearen partiellen Integrationsformel ermöglichen, die Entropiestruktur der porösen-Medien-Gleichung für einen großen Parameterbereich zu erhalten. Bezüglich der Zeit(semi)diskretisierung zeigen wir Bedingungen des abstraken Cauchyoperators, unter denen die Lösungen des zugehörigen Cauchy-Problems das selbe Entropieverhalten aufweisen wie die Runge-Kutta Semidiskretisierung. Dieser sehr allgemeine Zugang kann auf verschiedene Probleme angewandt werden, zum Beispiel die poröse-Medien-Gleichung, ein lineares Diffusionssystem und die Derrida-Lebowitz-Speer-Spohn-Gleichung. Das letzte präsentierte Resultat zeigt unter Ausnützung eines sorgfältig konstruierten numerischen Schemas und der schon erwähnten diskreten nichtlinearen partiellen Integrationsformel ein diskretes Analogon zu dem aus dem kontinuierlichen bekannten Bakry-Emery-Zugang für die poröse-Medien-Gleichung.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Prof. Ansgar Jüngel, for his mathematical guidance. I am deeply thankful for his knowledge, patience and invaluable advise.

I would like to thank Etienne Emmrich and Francis Filbet for their reports on this thesis and their interest in my work.

I appreciate the mathematical collaboration with Claire Chainais-Hillairet and would like to thank her for it, as well as for being a welcoming host during my stay in Lille.

I value the companionship of my all of my colleagues at the TU Vienna and University of Vienna, especially my fellow members of Prof. Jüngel's research group. I will always fondly remember my years at the Technical University.

Moreover, the author acknowledges partial support from the Austrian Science Fund (FWF), grants P22108, P24304, P27352, and W1245 and from the Austrian-French Project Amadeé of the Austrian Exchange Service (ÖAD).

Furthermore, I would like to express my gratitude to my parents for their continuing support and encouragement.

Finally, I wish to thank my beloved wife Ismene for making my life both steady and adventurous, both simple and rich. This thesis is dedicated to her and everything that lies before us.

Stefan Schuchnigg

Contents

Chapter 0

Overview and structure of the thesis

This doctoral thesis contributes to the broad topic of entropy-preserving numerical analysis and mainly discusses the results published in [27] (C. Chainais-Hillairet, A. Jüngel, S. Schuchnigg), [64] (A. Jüngel, S. Schuchnigg) and [66] (A. Jüngel, S. Schuchnigg). In addition some additional remarks and graphs are given to further illustrate the topic at hand.
In particular this thesis is concerned with the entropy-preserving nature of discretizations of nonlinear partial differential evolution equations. We mainly consider entropy functionals of the following form

$$E_\alpha[u] = \frac{1}{\alpha+1}\left(\int_\Omega u^{\alpha+1}dx - \left(\int_\Omega udx\right)^{\alpha+1}\right),$$
$$F_\alpha[u] = \frac{1}{2}\int_\Omega |\nabla u^{\alpha/2}|^2 dx, \quad \alpha > 0.$$

(which are referred to as entropies of zeroth and first order throughout this thesis) as well as their discrete counterparts. A recurrent toy and test equation is the nonlinear diffusion (porous-medium) equation

$$u_t = \Delta(u^\beta) \quad \text{in } \Omega,\ t > 0, \quad u(\cdot, 0) = u_0 \quad \text{in } \Omega,$$

which is studied in various different cases both analytically and numerically. In the presented numerical experiments we use the Barenblatt solution as inital datum, since it allows us to compare our numerical schemes to the explicitly given solution of the continuous equation.

The thesis is structured into three main parts, which are shortly outlined below, a detailed introduction into each of the chapters as well as a discussion of state of the art research is given at the beginning of each chapter.

In the first chapter, we show the time decay of the discrete entropies for fully discrete finite-volume approximations of porous-medium and fast-diffusion equations with Neumann or periodic boundary conditions. Here, the algebraic or exponential decay rates are computed explicitly. In particular, the numerical scheme dissipates all zeroth-order entropies which are also dissipated by the continuous equation. The proofs presented are based on continuous and discrete generalized Beckner inequalities. Furthermore, the exponential decay of some first-order entropies is proven in the continuous and discrete case using systematic integration by parts. Numerical experiments in one and two space dimensions illustrate the theoretical results of this chapter and indicate that some restrictions on the parameters seem to be only technical.

In the second chapter, results with a focus on time discretization are presented. The very general approach for nonlinear diffusion equations of parabolic type leads to conditions under which the schemes dissipate the discrete entropy locally. The dissipation property is a consequence of the concavity of the difference of the entropies at two consecutive time steps. The concavity property is shown to be related to the Bakry-Emery approach and the geodesic convexity of the entropy. The abstract conditions are verified for quasilinear parabolic equations (including the porous-medium equation), a linear diffusion system, and the fourth-order quantum diffusion equation. Numerical experiments for various Runge-Kutta finite-difference discretizations of the one-dimensional porous-medium equation show that the entropy-dissipation property is in fact global.

Finally, in the third chapter, the exponential decay of the relative entropy associated with a fully discrete porous-medium equation in one space dimension is shown by means of a discrete Bakry-Emery approach. The first ingredient of the proof is an abstract discrete Bakry-Emery method, which states conditions on a sequence under which the exponential decay of the discrete entropy follows. The second ingredient is a new nonlinear summation-by-parts formula which is inspired by systematic integration by parts developed by Matthes and Jüngel. Numerical simulations illustrate the exponential decay of the entropy for various time and space step sizes.

Appendix A contains the crucial nonlinear integration-by-parts formula needed in the first and third chapter and Appendix B features auxiliary inequalities needed throughout the thesis.

Chapter 1

Entropy preservation of the finite volume discretization of the porous-medium equation

1.1 Introduction

This chapter is concerned with the time decay of fully discrete finite-volume solutions to the nonlinear diffusion equation

$$u_t = \Delta(u^\beta) \quad \text{in } \Omega,\ t > 0, \quad u(\cdot, 0) = u_0 \quad \text{in } \Omega, \tag{1.1}$$

and with the relation to discrete generalized Beckner inequalities. Here, $\beta > 0$ and $\Omega \subset \mathbb{R}^d$ ($d \geq 1$) is a bounded domain. When $\beta > 1$, (1.1) is called the porous-medium equation, describing the flow of an isentropic gas through a porous medium [86]. Equation (1.1) with $\beta < 1$ is referred to as the fast-diffusion equation, which appears, for instance, in plasma physics with $\beta = \frac{1}{2}$ [10] or in semiconductor theory with $0 < \beta < 1$ [67]. We impose homogeneous Neumann boundary conditions

$$\nabla(u^\beta) \cdot \nu = 0 \quad \text{on } \partial\Omega,\ t > 0, \tag{1.2}$$

where ν denotes the unit normal exterior vector to $\partial\Omega$, or multiperiodic boundary conditions (i.e. Ω equals the torus $\mathbb{T}^d$). Let us denote by m the Lebesgue measure in $\mathbb{R}^d$ or $\mathbb{R}^{d-1}$; we assume for simplicity that $\mathrm{m}(\Omega) = 1$.
For existence and uniqueness results for the porous-medium equation in the whole space or under suitable boundary conditions, we refer to the monograph by Vázquez [86]. There are much less results for fast-diffusion equations (see [85]), and usually they hold for the whole-space problem. In particular, we are not aware of an existence result for fast-diffusion equations in bounded

domains with homogeneous Neumann boundary conditions, but such a result can be easily established since there is a maximum principle.[1]

There exist also many results on the time decay of the *continuous* porous-medium or fast-diffusion equation, with optimal decay rates or in strong norms. For instance, by using invariance principles, the sharp decay rate $t^{-1/(\beta-1)}$ in the L^∞ norm was shown in [1]. Spectral methods applied to (1.1) with confinement were used in [33] for $\beta \in ((d-2)/d, 1)$ and in [80] for $\beta > 1$. It seems to be difficult to "translate" these techniques to the discrete case. Sharp time-decay results in L^∞ for the solutions to the porous-medium equation with homogeneous Neumann boundary conditions were shown in [15, 51, 52], based on regular Sobolev inequalities. The connection between logarithmic Sobolev inequalities and ultracontractivity-like bounds was investigated in [15], also proving short- and long-time asymptotics. These results agree the results of the presented work in the *continuous* setting (in fact, the results of [15, 51, 52] are more general) but not in the *discrete* case. Optimal convergence rates to Barenblatt self-similar profiles for the fast-diffusion equation were derived in [14], employing entropy methods and Hardy-Poincaré inequalities. However, in this contributions it is unclear to what extent the mentioned techniques can be translated to the discrete case, partially because certain Sobolev inequalities (like Gagliardo-Nirenberg inequalities) seem not to be available. We refer to [11] for special discrete Gagliardo-Nirenberg inequalities.

In the literature, there exist many numerical schemes for nonlinear diffusion equations related to (1.1). Numerical techniques include (mixed) finite-element methods [3, 37, 78], finite-volume approximations [44, 77, 12], high-order relaxation ENO-WENO schemes [26], or particle methods [72]. In these references, stability and numerical convergence properties are proven. Another recent contribution [47] tackles the problem from the viewpoint of the discretization of the steady equation. Also the preservation of the structure of diffusion equations is a very important property of a numerical scheme. For instance, ideas employed for hyperbolic conservation laws were extended to degenerate diffusion equations, like the porous-medium equation, which may behave like hyperbolic equations in the regions of degeneracy [76]. Positivity-preserving schemes for nonlinear fourth-order equations were thoroughly investigated in the context of lubrication-type equations [7, 88] and quantum diffusion equations [60]. Entropy-consistent finite-volume finite-element schemes for the fourth-order thin-film equation were suggested by Grün and Rumpf [54]. For quantum diffusion models, an entropy-dissipative relaxation-type

[1] First, take strictly positive initial data u_0. By the maximum principle, any solution to the fast-diffusion equation is strictly positive. Thus, the equation is no longer singular, and the existence of weak solutions follows by a standard procedure. For nonnegative functions u_0, we take $u_0 + \varepsilon$ for $\varepsilon > 0$ as initial data, apply the first step, and pass to the limit $\varepsilon \to 0$.

finite-difference discretization was investigated by Carrillo et al. [23]. Furthermore, entropy-dissipative schemes for electro-reaction-diffusion systems were derived by Glitzky and Gärtner [48].
We aim to provide some results on the time decay of discrete solutions to (1.1)-(1.2) and to give estimates on the decay rates. To this end, we adapt the proofs for the continuous case to the discrete situation. The scheme under investigation is a backward Euler scheme in time and a finite-volume scheme in space, defined in (1.7). Only those proofs are chosen which can be directly translated in a finite-volume context.
Our main objective is to prove that the finite volume scheme for (1.1)-(1.2) dissipates the discrete versions of the functionals

$$E_\alpha[u] = \frac{1}{\alpha+1}\left(\int_\Omega u^{\alpha+1}dx - \left(\int_\Omega u dx\right)^{\alpha+1}\right), \tag{1.3}$$

$$F_\alpha[u] = \frac{1}{2}\int_\Omega |\nabla u^{\alpha/2}|^2 dx, \quad \alpha > 0. \tag{1.4}$$

In fact, we prove (algebraic or exponential) convergence rates at which the discrete functionals converge to zero as $t \to \infty$. We call E_α a zeroth-order entropy and F_α a first-order entropy. The functional F_1 is known as the Fisher information, used in mathematical statistics and information theory [38]. Our analysis of the decay rates of the entropies will be guided by the entropy-dissipation method. An essential ingredient of this technique is a functional inequality relating the entropy to the entropy dissipation [4, 22]. For the diffusion equation (1.1), this relation is realized by the Beckner inequality [8].
The entropy-dissipation method was applied to (1.1) in the whole space to prove the decay of the solutions to the asymptotic self-similar profile in, e.g., [24, 32]. The convergence towards the constant steady state on the one-dimensional torus was proven in [20]. However, we are not aware of general entropy decay estimates for solutions to (1.1)-(1.2) to the constant steady state, even in the continuous case. The reason might be that generalizations to the Beckner inequality, which is needed to relate the entropy dissipation to the entropy, have not been introduced before. To this end, we propose new Beckner-type inequalities which fill this gap. Moreover, our proof can be translated to the discrete case. These results are presented in Section 1.3.

The proof of discrete time decay for solutions to the finite-volume approximation of (1.1) is inspired by entropy decay estimates in the continuous case, which we review first. Differentiating $E_\alpha[u(t)]$ with respect to time and employing a Beckner inequality, we show for $\beta > 1$ that

$$\frac{dE_\alpha}{dt}[u(t)] \leq CE_\alpha[u(t)]^{(\alpha+\beta)/(\alpha+1)}, \quad t > 0,$$

where $C > 0$ depends on α, β, and $C_B(p, q)$. By a nonlinear Gronwall inequality, this implies the algebraic decay of $u(t)$ to equilibrium in the entropy sense; see Theorem 9. If the solution is positive and $0 < \alpha \leq 1$, the above inequality becomes

$$\frac{dE_\alpha}{dt}[u(t)] \leq C(u_0)E_\alpha[u(t)], \quad t > 0,$$

which results in an exponential decay rate; see Theorem 10. We obtain similar results for a discrete version of E_α in Theorems 11 (algebraic decay) and 12 (exponential decay).

The first-order entropies $F_\alpha[u(t)]$ decay exponentially (for positive solutions) for all (α, β) lying in the strip $-2 \leq \alpha - 2\beta \leq 1$ (one-dimensional case) or in the region M_d, which is illustrated in Figure 1.1 below (multi-dimensional case); see Theorems 13 and 14. The proof is based on systematic integration by parts [58]. In order to avoid boundary integrals arising from the iterated integrations by parts, these results are valid only if $\Omega = \mathbb{T}^d$. It is very difficult to translate the iterated integrations by parts to iterated summations by parts since there is no discrete nonlinear chain rule. For the zeroth-order entropies, this is done by exploiting the convexity of the mapping $x \mapsto x^{\alpha+1}$. For the first-order entropies, we employ the concavity of the discrete version of dF_α/dt with respect to the time approximation parameter. We prove in Theorem 16 that for $1 \leq \alpha \leq 2$ and $-2 < \alpha - 2\beta < 1$, the discrete first-order entropy is monotone (multi-dimensional case) and decays exponentially (one-dimensional case).

Throughout this chapter, we assume that the solutions to (1.1) are smooth and positive, such that we can perform all the computations and integrations by parts. In particular, we avoid any technicalities due to the degeneracy ($\beta > 1$) or singularity ($\beta < 1$) in (1.1). Most of our results presented in this chapter can be easily generalized to nonnegative weak solutions by using a suitable approximation scheme, but details are omitted here for the sake of readability. In addition, for reasons of simplicity, we restrict ourselves to a uniform time step size, although a generalisation of the nonlinear Gronwall's lemma (Corollary 41 in the Appendix B) for variable stepsizes (similar to [40]) seems feasible.

We stress the fact that we do not develop an efficient implementation and we do not perform a convergence analysis, since the scheme is rather standard. Our aim is of more theoretical interest. In fact, our results on the discrete decay rates contribute to the aim of developing and analyzing *structure-preserving* numerical schemes and this is the main originality of the work presented in this chapter.

This chapter is organized as follows. Section 1.2 is devoted to the finite-volume setting: We introduce the numerical scheme under investigation and define discrete norms and discrete entropies. Then we prove some novel generalized Beckner inequalities in Section 1.3, at the continuous and discrete level. The algebraic/exponential decay of $E_\alpha[u]$ is studied in Section 1.4. We first prove the results at the continuous level and then deduce similar results for the numerical scheme. Section 1.5 is devoted to the study of the exponential decay of the first-order entropies $F_\alpha[u]$. In Section 1.6, we illustrate the theoretical results by numerical experiments in one and two space dimensions. They indicate that some of the restrictions on the parameters (α, β) seem to be only technical.

1.2 The finite-volume setting

1.2.1 Notations and finite-volume scheme

Let Ω be an open bounded polyhedral subset of $\mathbb{R}^d$ ($d \geq 2$) with Lipschitz boundary and $\mathrm{m}(\Omega) = 1$. An admissible mesh of Ω is given by a family $\mathcal{T}$ of control volumes (open and convex polyhedral subsets of Ω with positive measure); a family $\mathcal{E}$ of relatively open parts of hyperplanes in $\mathbb{R}^d$ which represent the faces of the control volumes; and a family of points $(x_K)_{K\in\mathcal{T}}$ which satisfy Definition 9.1 in [42]. This definition implies that the straight line between two neighboring centers of cells (x_K, x_L) is orthogonal to the edge $\sigma = K|L$ between the two control volume K and L. For instance, triangular meshes in $\mathbb{R}^2$ satisfy the admissibility condition if all angles of the triangles are smaller than $\pi/2$ [42, Examples 9.1]. Voronoi meshes in $\mathbb{R}^d$ are also admissible meshes [42, Examples 9.2].

We distinguish the interior faces $\sigma \in \mathcal{E}_{\mathrm{int}}$ and the boundary faces $\sigma \in \mathcal{E}_{\mathrm{ext}}$. Then the union $\mathcal{E}_{\mathrm{int}} \cup \mathcal{E}_{\mathrm{ext}}$ equals the set of all faces $\mathcal{E}$. For a control volume $K \in \mathcal{T}$, we denote by $\mathcal{E}_K$ the set of its faces, by $\mathcal{E}_{\mathrm{int},K}$ the set of its interior faces, and by $\mathcal{E}_{\mathrm{ext},K}$ the set of edges of K included in $\partial\Omega$.

Let d be the distance in $\mathbb{R}^d$. We assume that the family of meshes satisfies the following regularity requirement: There exists $\xi > 0$ such that for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_{\mathrm{int},K}$ with $\sigma = K|L$, it holds

$$\mathrm{d}(x_K, \sigma) \geq \xi \mathrm{d}(x_K, x_L). \tag{1.5}$$

This hypothesis is needed to apply a discrete Poincaré inequality; see Lemma 2. Introducing for $\sigma \in \mathcal{E}$ the notation

$$d_\sigma = \begin{cases} \mathrm{d}(x_K, x_L) & \text{if } \sigma \in \mathcal{E}_{\mathrm{int}},\ \sigma = K|L, \\ \mathrm{d}(x_K, \sigma) & \text{if } \sigma \in \mathcal{E}_{\mathrm{ext},K}, \end{cases}$$

we define the transmissibility coefficient

$$\tau_\sigma = \frac{\mathrm{m}(\sigma)}{d_\sigma}, \quad \sigma \in \mathcal{E}.$$

The size of the mesh is denoted by $\triangle x = \max_{K\in\mathcal{T}} \mathrm{diam}(K)$. Let $T > 0$ be some final time and M_T the number of time steps. Then the time step size and the time points are given by, respectively, $\triangle t = T/M_T$, $t^k = k\triangle t$ for $0 \le k \le M_T$. We denote by $\mathcal{D}$ an admissible space-time discretization of $\Omega_T = \Omega \times (0,T)$ composed of an admissible mesh $\mathcal{T}$ of Ω and the values $\triangle t$ and M_T.
We are now in the position to define the finite-volume scheme of (1.1)-(1.2) on $\mathcal{D}$. The initial datum is approximated by its L^2 projection on control volumes:

$$u^0 = \sum_{K\in\mathcal{T}} u_K^0 \mathbf{1}_K, \quad \text{where } u_K^0 = \frac{1}{\mathrm{m}(K)}\int_K u_0(x)dx, \tag{1.6}$$

and $\mathbf{1}_K$ is the characteristic function on K. Then it holds $\sum_{K\in\mathcal{T}} \mathrm{m}(K)u_K^0 = \int_\Omega u_0 dx$.
The numerical scheme reads as follows:

$$\mathrm{m}(K)\frac{u_K^{k+1} - u_K^k}{\triangle t} + \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma\big((u_K^{k+1})^\beta - (u_L^{k+1})^\beta\big) = 0, \tag{1.7}$$

for all $K \in \mathcal{T}$ and $k = 0, \ldots, M_T - 1$. This scheme is based on a fully implicit Euler discretization in time and a finite-volume approach for the volume variable. The Neumann boundary conditions (1.2) are taken into account as the sum in (1.7) applies only on the interior edges. The implicit scheme allows us to establish discrete entropy-dissipation estimates which would not be possible with an explicit scheme.
In the following proposition, we summarize the classical results of existence, uniqueness and stability of the solution to the finite-volume scheme (1.6)-(1.7).

Proposition 1. *Let $u_0 \in L^\infty(\Omega)$, $m \ge 0$, $M \ge 0$ such that $m \le u_0 \le M$ in Ω. Let $\mathcal{T}$ be an admissible mesh of Ω. Then the scheme* (1.6)-(1.7) *admits a unique solution $(u_K^k)_{K\in\mathcal{T},\, 0\le k\le M_T}$ satisfying*

$$m \le u_K^k \le M, \quad \textit{for all } K \in \mathcal{T},\, 0 \le k \le M_T,$$
$$\sum_{K\in\mathcal{T}} \mathrm{m}(K)u_K^k = \|u_0\|_{L^1(\Omega)}, \quad \textit{for all } 0 \le k \le M_T.$$

We refer, for instance, to [42] and [43] for the proof of this proposition.

1.2.2 Discrete entropies

A finite-volume scheme provides an approximate solution which is constant on each cell of the mesh and on each time interval. Let $X(\mathcal{T})$ be the linear space of functions $\Omega \to \mathbb{R}$ which are constant on each cell $K \in \mathcal{T}$:

$$X(\mathcal{T}) = \left\{ u = \sum_{K\in\mathcal{T}} u_K \mathbf{1}_K \right\}.$$

The set $X(\mathcal{T})$ is included in $L^p(\Omega)$ for $1 \le p \le \infty$ and

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u|^p dx \right)^{1/p} = \left(\sum_{K\in\mathcal{T}} \mathrm{m}(K)|u_K|^p \right)^{1/p} \qquad \begin{array}{l} \forall u \in X(\mathcal{T}), \\ \forall 1 \le p < +\infty. \end{array}$$

Clearly, the set $X(\mathcal{T})$ is not included in $W^{1,p}(\Omega)$. However, for $1 \le p < +\infty$, we can define a discrete $W^{1,p}$ seminorm and a discrete $W^{1,p}$ norm by, respectively,

$$|u|_{1,p,\mathcal{T}} = \left(\sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}}, \\ \sigma=K|L}} \frac{\mathrm{m}(\sigma)}{d_\sigma^{p-1}} |u_K - u_L|^p \right)^{1/p} \qquad \forall u \in X(\mathcal{T}),$$

$$\|u\|_{1,p,\mathcal{T}} = \|u\|_{L^p(\Omega)} + |u|_{1,p,\mathcal{T}} \quad \forall u \in X(\mathcal{T}).$$

The zeroth-order entropies defined by (1.3) can be rewritten for $u \in X(\mathcal{T})$ as

$$E_\alpha[u] = \frac{1}{\alpha+1} \left(\sum_{K\in\mathcal{T}} \mathrm{m}(K) u_K^{\alpha+1} - \left(\sum_{K\in\mathcal{T}} \mathrm{m}(K) u_K \right)^{\alpha+1} \right). \tag{1.8}$$

Finally, we define the discrete first-order entropies, corresponding to (1.4), by

$$F_\alpha^d[u] = \frac{1}{2} |u^{\alpha/2}|_{1,2,\mathcal{T}}^2. \tag{1.9}$$

1.3 Generalized Beckner inequalities

The decay properties of the zeroth-order entropies rely on generalized Beckner inequalities which follow from the Poincaré-Wirtinger inequality. This section is devoted to the proof of these Beckner inequalities in the functional space $H^1(\Omega)$ and of their discrete counterpart in the functional space $X(\mathcal{T})$.

1.3.1 Poincaré-Wirtinger inequalities

We assume that $\Omega \subset \mathbb{R}^d$ ($d \geq 1$) is a bounded domain such that the Poincaré-Wirtinger inequality

$$\|f - \bar{f}\|_{L^2(\Omega)} \leq C_P \|\nabla f\|_{L^2(\Omega)} \tag{1.10}$$

holds for all $f \in H^1(\Omega)$, where $\bar{f} = \mathrm{m}(\Omega)^{-1} \int_\Omega f dx$ and $C_P > 0$ only depends on d and Ω. This is the case if, for instance, Ω has the cone property [70, Theorem 8.11] or if $\partial\Omega$ is locally Lipschitz continuous [87, Theorem 1.3.4]. We recall that $\mathrm{m}(\Omega) = 1$ in the following of this chapter (to shorten the proof). The discrete counterpart of (1.10) is stated in the following Lemma (see for instance [11, Theorem 5]):

Lemma 2 (Discrete Poincaré-Wirtinger inequality). *Let $\Omega \subset \mathbb{R}^d$ be an open bounded polyhedral set and let $\mathcal{T}$ be an admissible mesh satisfying the regularity constraint* (1.5)*. Then there exists a constant $C_p > 0$, only depending on d and Ω, such that for all $f \in X(\mathcal{T})$,*

$$\|f - \bar{f}\|_{L^2(\Omega)} \leq \frac{C_p}{\xi^{1/2}} |f|_{1,2,\mathcal{T}}, \tag{1.11}$$

where $\bar{f} = \int_\Omega f dx$ (recall that $\mathrm{m}(\Omega) = 1$) and ξ is defined in (1.5)*.*

We now present a new inequality which can be regarded as a generalized Poincaré inequality.

Lemma 3 (Generalized Poincaré-Wirtinger inequality). *Let $0 < q \leq 2$ and $f \in H^1(\Omega)$. Then*

$$\|f\|^q_{L^2(\Omega)} \leq C_P^q \|\nabla f\|^q_{L^2(\Omega)} + \|f\|^q_{L^q(\Omega)} \tag{1.12}$$

holds, where $C_P > 0$ is the constant of the Poincaré-Wirtinger inequality (1.10)*.*

Proof. Let first $1 \leq q \leq 2$. The Poincaré-Wirtinger inequality (1.10) rewrites as

$$\|f\|^2_{L^2(\Omega)} - \|f\|^2_{L^1(\Omega)} = \|f - \bar{f}\|^2_{L^2(\Omega)} \leq C_P^2 \|\nabla f\|^2_{L^2(\Omega)} \tag{1.13}$$

and together with the Hölder inequality leads to

$$\|f\|^2_{L^2(\Omega)} \leq C_P^2 \|\nabla f\|^2_{L^2(\Omega)} + \|f\|^2_{L^q(\Omega)}. \tag{1.14}$$

Here we use the assumption $\mathrm{m}(\Omega) = 1$. Since $q/2 \leq 1$, it follows that

$$\|f\|^q_{L^2(\Omega)} \leq \left(C_P^2 \|\nabla f\|^2_{L^2(\Omega)} + \|f\|^2_{L^q(\Omega)}\right)^{q/2} \leq C_P^q \|\nabla f\|^q_{L^2(\Omega)} + \|f\|^q_{L^q(\Omega)},$$

which equals (1.12).
Next, let $0 < q < 1$. We claim that

$$a^{q/2} - a^{q-1}b^{1-q/2} \leq (a-b)^{q/2} \quad \text{for all } a \geq b > 0. \tag{1.15}$$

Indeed, setting $c = b/a$, this inequality is equivalent to

$$1 - c^{1-q/2} \leq (1-c)^{q/2} \quad \text{for all } 0 < c \leq 1.$$

The function $g(c) = 1 - c^{1-q/2} - (1-c)^{q/2}$ for $c \in [0,1]$ satisfies $g(0) = g(1) = 0$ and $g''(c) = (q/2)(1-q/2)(c^{-1-q/2} + (1-c)^{q/2-2}) \geq 0$ for $c \in (0,1)$, which implies that $g(c) \leq 0$, proving (1.15). Applying (1.15) to $a = \|f\|^2_{L^2(\Omega)}$ and $b = \|f\|^2_{L^1(\Omega)}$ and using (1.13), we find that

$$\|f\|^q_{L^2(\Omega)} - \|f\|^{2(q-1)}_{L^2(\Omega)}\|f\|^{2-q}_{L^1(\Omega)} \leq \left(\|f\|^2_{L^2(\Omega)} - \|f\|^2_{L^1(\Omega)}\right)^{q/2} \leq C_P^q \|\nabla f\|^q_{L^2(\Omega)}. \tag{1.16}$$

In order to get rid of the L^1 norm, we employ the interpolation inequality

$$\|f\|_{L^1(\Omega)} = \int_\Omega |f|^\theta |f|^{1-\theta} dx \leq \|f\|^\theta_{L^q(\Omega)} \|f\|^{1-\theta}_{L^2(\Omega)}, \tag{1.17}$$

where $\theta = q/(2-q) < 1$. Since $(2-q)\theta = q$ and $(2-q)(1-\theta) = 2(1-q)$, (1.16) becomes

$$\|f\|^q_{L^2(\Omega)} - \|f\|^q_{L^q(\Omega)} \leq C_P^q \|\nabla f\|^q_{L^2(\Omega)},$$

which is the desired inequality. □

Starting from the discrete Poincaré-Wirtinger inequality (1.11) instead of (1.10), we obtain the discrete analogon of (1.13):

$$\|f\|^2_{L^2(\Omega)} - \|f\|^2_{L^1(\Omega)} = \|f - \bar{f}\|^2_{L^2(\Omega)} \leq C_p^2 \xi^{-1} |f|^2_{1,2,\mathcal{T}} \quad \text{for all } f \in X(\mathcal{T}).$$

Then, following the lines of the proof of Lemma 3, we obtain the discrete counterpart of the generalized Poincaré-Wirtinger inequality (1.12)

$$: \|f\|^q_{L^2(\Omega)} \leq C_p^q \xi^{-q/2} |f|^q_{1,2,\mathcal{T}} + \|f\|^q_{L^q(\Omega)} \quad \text{for all } f \in X(\mathcal{T}), \tag{1.18}$$

under the hypotheses of Lemma 2.

1.3.2 First generalization of the Beckner inequality

For the proof of the algebraic decay of the zeroth-order entropies, we need the following variant of the Beckner inequality.

Lemma 4 (Generalized Beckner inequality I)**.** *Let $d \geq 1$ and either $0 < q < 2$, $pq \geq 1$ or $q = 2$, $\frac{1}{2} - \frac{1}{d} \leq p \leq 1$ ($0 < p \leq 1$ if $d \leq 2$), and let $f \in H^1(\Omega)$. Then the generalized Beckner inequality*

$$\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{1/p} dx \right)^{pq} \leq C_B(p,q) \|\nabla f\|^q_{L^2(\Omega)} \tag{1.19}$$

holds, where

$$C_B(p,q) = \frac{2(pq-1)C_P^q}{2-q} \quad \text{if } q < 2, \quad C_B(p,2) = C_P^2 \quad \text{if } q = 2,$$

and $C_P > 0$ is the constant of the Poincaré-Wirtinger inequality (1.10)*.*

Remark 5. The case $q = 2$ corresponds to the usual Beckner inequality [8]

$$\int_\Omega |f|^2 dx - \left(\int_\Omega |f|^{2/r} dx \right)^r \leq C_B(p,2) \|\nabla f\|^2_{L^2(\Omega)},$$

where $1 \leq r = 2p \leq 2$. It is shown in [36] that the constant $C_B(p,2)$ can be related to the lowest positive eigenvalue of a Schrödinger operator if Ω is convex. On the one-dimensional torus, the generalized Beckner inequality (1.19) for $p > 0$ and $0 < q < 2$ has been derived in [20]. In the multi-dimensional situation, the special case $p = 2/q$ was proven in [35]. In this work, it was also shown that (1.19) with $q > 2$ and $p = 2/q$ cannot hold true unless the Lebesgue measure dx is replaced by the Dirac measure. In the limit $pq \to 1$, (1.19) leads to a generalized logarithmic Sobolev inequality (see (1.21) below). If $q = 2$ in this limit, the usual logarithmic Sobolev inequality [53] is obtained. $\square$

Proof of Lemma 4. Let first $q = 2$. Then the Beckner inequality is a consequence of the Poincaré-Wirtinger inequality (1.10) and the Jensen inequality:

$$C_P^2 \|\nabla f\|^2_{L^2(\Omega)} \geq \|f - \bar{f}\|^2_{L^2(\Omega)} = \|f\|^2_{L^2(\Omega)} - \|f\|^2_{L^1(\Omega)} \geq \int_\Omega f^2 dx - \left(\int_\Omega |f|^{2/r} dx \right)^r,$$

where $1 - \frac{2}{d} \leq r \leq 2$ ($0 < r \leq 2$ if $d \leq 2$). The lower bound for r ensures that the embedding $H^1(\Omega) \hookrightarrow L^{2/r}(\Omega)$ is continuous. The choice $p = r/2 \in [\frac{1}{2} - \frac{1}{d}, 1]$ yields the formulation (1.19).

Next, let $0 < q < 2$. The first part of the proof is inspired by the proof of Proposition 2.2 in [35]. Taking the logarithm of the interpolation inequality

$$\|f\|_{L^r(\Omega)} \leq \|f\|^{\theta(r)}_{L^q(\Omega)} \|f\|^{1-\theta(r)}_{L^2(\Omega)},$$

where $q \leq r \leq 2$ and $\theta(r) = q(2-r)/(r(2-q))$, gives

$$F(r) := \frac{1}{r} \log \int_\Omega |f|^r dx - \frac{\theta(r)}{q} \log \int_\Omega |f|^q dx - \frac{1-\theta(r)}{2} \log \int_\Omega |f|^2 dx \leq 0.$$

The function $F : [q, 2] \to \mathbb{R}$ is nonpositive, differentiable and $F(q) = 0$. Therefore, $F'(q) \le 0$, which equals

$$-\frac{1}{q^2}\log\int_\Omega |f|^q dx + \frac{1}{q}\left(\int_\Omega |f|^q dx\right)^{-1}\int_\Omega |f|^q \log|f| dx \\ + \theta'(q)\left(\frac{1}{2}\log\int_\Omega |f|^2 dx - \frac{1}{q}\log\int_\Omega |f|^q dx\right) \le 0.$$

We multiply this inequality by $q^2\int_\Omega |f|^q dx$ to obtain

$$\int_\Omega |f|^q \log\frac{|f|^q}{\|f\|^q_{L^q(\Omega)}} dx \le \frac{2}{2-q}\|f\|^q_{L^q(\Omega)}\log\frac{\|f\|^q_{L^2(\Omega)}}{\|f\|^q_{L^q(\Omega)}}. \tag{1.20}$$

Then, we employ Lemma 3 and the inequality $\log(x+1) \le x$ for $x \ge 0$ to infer that

$$\|f\|^q_{L^q(\Omega)}\log\frac{\|f\|^q_{L^2(\Omega)}}{\|f\|^q_{L^q(\Omega)}} \le \|f\|^q_{L^q(\Omega)}\log\left(\frac{C_P^q\|\nabla f\|^q_{L^2(\Omega)}}{\|f\|^q_{L^q(\Omega)}} + 1\right) \le C_P^q\|\nabla f\|^q_{L^2(\Omega)}.$$

Combining this inequality and (1.20), we conclude the generalized logarithmic Sobolev inequality

$$\int_\Omega |f|^q \log\frac{|f|^q}{\|f\|^q_{L^q(\Omega)}} dx \le \frac{2C_P^q}{2-q}\|\nabla f\|^q_{L^2(\Omega)}. \tag{1.21}$$

The generalized Beckner inequality (1.19) is derived by slightly extending the proof of [68, Corollary 1]. Let

$$G(r) = r\log\int_\Omega |f|^{q/r} dx, \quad r \ge 1.$$

The function G is twice differentiable with

$$G'(r) = \left(\int_\Omega |f|^{q/r} dx\right)^{-1}\left(\int_\Omega |f|^{q/r} dx \log\int_\Omega |f|^{q/r} dx - \frac{q}{r}\int_\Omega |f|^{q/r}\log|f| dx\right),$$

$$G''(r) = \frac{q^2}{r^3}\left(\int_\Omega |f|^{q/r} dx\right)^{-2} \\ \times\left(\int_\Omega |f|^{q/r} dx \int_\Omega |f|^{q/r}(\log|f|)^2 dx - \left(\int_\Omega |f|^{q/r}\log|f| dx\right)^2\right).$$

The Cauchy-Schwarz inequality shows that $G''(r) \ge 0$, i.e., G is convex. Consequently, $r \mapsto e^{G(r)}$ is also convex and $r \mapsto H(r) = -(e^{G(r)} - e^{G(1)})/(r-1)$ is nonincreasing on $(1, \infty)$, which implies that

$$H(r) \le \lim_{t\to 1} H(t) = -e^{G(1)}G'(1) = \int_\Omega |f|^q \log\frac{|f|^q}{\|f\|^q_{L^q(\Omega)}} dx.$$

This inequality is equivalent to

$$\frac{1}{r-1}\left(\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{q/r} dx\right)^r\right) \le \int_\Omega |f|^q \log \frac{|f|^q}{\|f\|^q_{L^q(\Omega)}} dx. \qquad (1.22)$$

Combining this inequality and the generalized logarithmic Sobolev inequality (1.21), it follows that

$$\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{q/r} dx\right)^r \le \frac{2(r-1)C_P^q}{2-q}\|\nabla f\|^q_{L^2(\Omega)}$$

for all $0 < q < 2$ and $r \ge 1$. Setting $p := r/q$, this proves (1.19) for all $pq = r \ge 1$. $\square$

Lemma 6 (Discrete generalized Beckner inequality I). *Let* $0 < q < 2$*,* $pq > 1$ *or* $q = 2$ *and* $0 < p \le 1$*, and* $f \in X(\mathcal{T})$*. Then*

$$\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{1/p} dx\right)^{pq} \le C_b(p,q)|f|^q_{1,2,\mathcal{T}}$$

holds, where

$$C_b(p,q) = \frac{2(pq-1)C_p^q}{(2-q)\xi^{q/2}} \quad \text{if } q < 2, \quad C_b(p,2) = \frac{C_p^2}{\xi} \quad \text{if } q = 2.$$

C_p *is the constant in the discrete Poincaré-Wirtinger inequality, and* ξ *is defined in* (1.5).

Proof. The proof follows the lines of the proof of Lemma 4, noting that in the discrete (finite-dimensional) setting, we do no longer need the lower bound on p. Indeed, if $q = 2$, the conclusion results from the discrete Poincaré-Wirtinger inequality (1.11) and the Jensen inequality. If $q < 2$, we first remark that (1.20) and (1.22) still holds for $f \in X(\mathcal{T})$, leading to

$$\begin{aligned}\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{1/p} dx\right)^{pq} &\le (pq-1)\int_\Omega |f|^q \log \frac{|f|^q}{\|f\|^q_{L^q(\Omega)}} dx \\ &\le \frac{2(pq-1)}{2-q}\|f\|^q_{L^q(\Omega)} \log \frac{\|f\|^q_{L^2(\Omega)}}{\|f\|^q_{L^q(\Omega)}}. \qquad (1.23)\end{aligned}$$

Then, inserting the discrete Poincaré-Wirtinger inequality (1.18) instead of (1.12) into (1.23) to replace $\|f\|_{L^2(\Omega)}$ and using $\log(x+1) \le x$ for $x \ge 0$, the lemma follows. $\square$

1.3.3 Second generalization of the Beckner inequality

For the proof of exponential decay rates, we need the following variant of the Beckner inequality.

Lemma 7 (Generalized Beckner inequality II). *Let* $0 < q < 2$, $pq \geq 1$ *and* $f \in H^1(\Omega)$. *Then*

$$\|f\|_{L^q(\Omega)}^{2-q}\left(\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{1/p} dx\right)^{pq}\right) \leq C_B'(p,q)\|\nabla f\|_{L^2(\Omega)}^2, \tag{1.24}$$

where

$$C_B'(p,q) = \begin{cases} \dfrac{q(pq-1)C_P^2}{2-q} & \text{if } 1 \leq q < 2, \\ (pq-1)C_P^2 & \text{if } 0 < q < 1. \end{cases}$$

Proof. By (1.20), it holds that for all $0 < q < 2$,

$$\int_\Omega |f|^q \log \frac{|f|^q}{\|f\|_{L^q(\Omega)}^q} dx \leq \frac{q}{2-q}\|f\|_{L^q(\Omega)}^q \log \frac{\|f\|_{L^2(\Omega)}^2}{\|f\|_{L^q(\Omega)}^2}.$$

Then, for $q > 1$, the Poincaré-Wirtinger inequality in the version (1.14) and the inequality $\log(x+1) \leq x$ for $x \geq 0$ yield

$$\begin{aligned} \|f\|_{L^q(\Omega)}^q \log \frac{\|f\|_{L^2(\Omega)}^2}{\|f\|_{L^q(\Omega)}^2} &\leq \|f\|_{L^q(\Omega)}^q \log\left(C_P^2 \frac{\|\nabla f\|_{L^2(\Omega)}^2}{\|f\|_{L^q(\Omega)}^2} + 1\right) \\ &\leq C_P^2 \|f\|_{L^q(\Omega)}^{q-2}\|\nabla f\|_{L^2(\Omega)}^2. \end{aligned} \tag{1.25}$$

Taking into account (1.22), the conclusion follows for $q > 1$. Let $0 < q \leq 1$. Suppose that the following inequality holds:

$$\|f\|_{L^q(\Omega)}^2 + \frac{2-q}{q} C_P^2 \|\nabla f\|_{L^2(\Omega)}^2 - \|f\|_{L^2(\Omega)}^2 \geq 0. \tag{1.26}$$

This implies that, by (1.22) and for $r = pq$,

$$\begin{aligned} \int_\Omega |f|^q dx - \left(\int_\Omega |f|^{q/r} dx\right)^r &\leq \frac{(pq-1)q}{2-q}\|f\|_{L^q(\Omega)}^q \log \frac{\|f\|_{L^2(\Omega)}^2}{\|f\|_{L^q(\Omega)}^2} \\ &\leq \frac{(pq-1)q}{2-q}\|f\|_{L^q(\Omega)}^q \log\left(\frac{(2-q)C_P^2}{q}\frac{\|\nabla f\|_{L^2(\Omega)}^2}{\|f\|_{L^q(\Omega)}^2} + 1\right) \\ &\leq (pq-1)C_P^2\|\nabla f\|_{L^2(\Omega)}^2 \|f\|_{L^q(\Omega)}^{q-2}, \end{aligned}$$

which shows the desired Beckner inequality.
It remains to prove (1.26). For this, we employ the Poincaré-Wirtinger inequality (1.13)

$$C_P^2 \|\nabla f\|_{L^2(\Omega)}^2 \geq \|f\|_{L^2(\Omega)}^2 - \|f\|_{L^1(\Omega)}^2$$

and the interpolation inequality (1.17) in the form

$$\|f\|_{L^q(\Omega)}^2 \geq \|f\|_{L^1(\Omega)}^{2/\theta} \|f\|_{L^2(\Omega)}^{2(\theta-1)/\theta}, \quad \theta = \frac{q}{2-q} \leq 1,$$

to obtain

$$\begin{aligned}\|f\|_{L^q(\Omega)}^2 &+ \frac{2-q}{q} C_P^2 \|\nabla f\|_{L^2(\Omega)}^2 - \|f\|_{L^2(\Omega)}^2 \\ &\geq \|f\|_{L^1(\Omega)}^{2/\theta} \|f\|_{L^2(\Omega)}^{2(\theta-1)/\theta} + \left(\frac{2-q}{q} - 1\right) \|f\|_{L^2(\Omega)}^2 - \frac{2-q}{q} \|f\|_{L^1(\Omega)}^2.\end{aligned}$$

We interpret the right-hand side as a function G of $\|f\|_{L^1(\Omega)}^2$. Then, setting $A = \|f\|_{L^2(\Omega)}^2$,

$$\begin{aligned}G(y) &= y^{1/\theta} A^{1-1/\theta} + \frac{2(1-q)}{q} A - \frac{2-q}{q} y, \\ G'(y) &= \frac{1}{\theta} y^{1/\theta-1} A^{1-1/\theta} - \frac{2-q}{q}, \\ G''(y) &= \frac{1}{\theta}\left(\frac{1}{\theta} - 1\right) y^{1/\theta-2} A^{1-1/\theta} \geq 0,\end{aligned}$$

Therefore, G is a convex function which satisfies $G(A) = 0$ and $G'(A) = 0$. This implies that G is a nonnegative function on $\mathbb{R}^+$ and in particular, $G(\|f\|_{L^1(\Omega)}^2) \geq 0$. This proves (1.26), completing the proof. □

The adaptation of the proof of Lemma 7 is straightforward, using the discrete Poincaré-Wirtinger inequality (1.11) instead of (1.10). This yields the following result.

Lemma 8 (Discrete generalized Beckner inequality II). *Let $0 < q < 2$, $pq \geq 1$, and $f \in X(\mathcal{T})$. Then*

$$\|f\|_{L^q(\Omega)}^{2-q} \left(\int_\Omega |f|^q dx - \left(\int_\Omega |f|^{1/p} dx\right)^{pq}\right) \leq C_b'(p,q) |f|_{1,2,\mathcal{T}}^2$$

holds, where

$$C_b'(p,q) = \begin{cases} \dfrac{q(pq-1)C_p^2}{(2-q)\xi} & \text{if } 1 \leq q < 2, \\ \dfrac{(pq-1)C_p^2}{\xi} & \text{if } 0 < q < 1, \end{cases}$$

C_p is the constant in the discrete Poincaré-Wirtinger inequality, and ξ is defined in (1.5).

1.4 Zeroth-order entropies: from the continuous to the discrete level

In this section, we prove the algebraic or exponential decay of the zeroth-order entropies. We first study the continuous case and then show how to extend the obtained result to the numerical scheme.

1.4.1 The continuous case

Let u be a smooth solution to (1.1)-(1.2) and let $u_0 \in L^\infty(\Omega)$, $\inf_\Omega u_0 \geq 0$ in Ω. By the maximum principle, $0 \leq \inf_\Omega u_0 \leq u(t) \leq \sup_\Omega u_0$ in Ω for $t \geq 0$. First, we prove algebraic decay rates for $E_\alpha[u]$, defined in (1.3).

Theorem 9 (Polynomial decay for E_α)**.** *Let $\alpha > 0$ and $\beta > 1$. Let u be a smooth solution to* (1.1)-(1.2) *and $u_0 \in L^\infty(\Omega)$ with $\inf_\Omega u_0 \geq 0$. Then*

$$E_\alpha[u(t)] \leq \frac{1}{(C_1 t + C_2)^{(\alpha+1)/(\beta-1)}}, \quad t \geq 0,$$

where

$$C_1 = \frac{4\alpha\beta(\beta-1)}{(\alpha+1)(\alpha+\beta)^2}\left(\frac{\alpha+1}{C_B(p,q)}\right)^{(\alpha+\beta)/(\alpha+1)}, \quad C_2 = E_\alpha[u_0]^{-(\beta-1)/(\alpha+1)},$$

and $C_B(p,q) > 0$ is the constant in the Beckner inequality for $p = (\alpha+\beta)/2$ and $q = 2(\alpha+1)/(\alpha+\beta)$.

Proof. We apply Lemma 4 with $p = (\alpha+\beta)/2$ and $q = 2(\alpha+1)/(\alpha+\beta)$. The assumptions on α and β guarantee that $0 < q < 2$ and $pq > 1$. Then, with $f = u^{(\alpha+\beta)/2}$,

$$\begin{aligned} E_\alpha[u] &= \frac{1}{\alpha+1}\left(\int_\Omega u^{\alpha+1}dx - \left(\int_\Omega u dx\right)^{\alpha+1}\right) \\ &\leq \frac{C_B(p,q)}{\alpha+1}\left(\int_\Omega |\nabla u^{(\alpha+\beta)/2}|^2 dx\right)^{(\alpha+1)/(\alpha+\beta)}. \end{aligned}$$

Now, computing the derivative,

$$\frac{dE_\alpha}{dt} = -\int_\Omega \nabla u^\alpha \cdot \nabla u^\beta dx = -\frac{4\alpha\beta}{(\alpha+\beta)^2}\int_\Omega |\nabla u^{(\alpha+\beta)/2}|^2 dx \tag{1.27}$$

$$\leq -\frac{4\alpha\beta}{(\alpha+\beta)^2}\left(\frac{\alpha+1}{C_B(p,q)}\right)^{(\alpha+\beta)/(\alpha+1)} E_\alpha[u]^{(\alpha+\beta)/(\alpha+1)}. \tag{1.28}$$

An integration of this inequality gives the assertion. ☐

Next, we show exponential decay rates.

Theorem 10 (Exponential decay for E_α). *Let u be a smooth solution to (1.1)-(1.2), $0 < \alpha \le 1$, $\beta > 0$, $u_0 \in L^\infty(\Omega)$ with $\inf_\Omega u_0 \ge 0$. Then*

$$E_\alpha[u(t)] \le E_\alpha[u_0] e^{-\Lambda t}, \quad t \ge 0.$$

The constant Λ is given by

$$\Lambda = \frac{4\alpha\beta}{C_B(\frac{1}{2}(\alpha+1), 2)(\alpha+1)} \inf_\Omega \left(u_0^{\beta-1}\right) \ge 0,$$

for $\beta > 0$ and

$$\Lambda = \frac{4\alpha\beta(\alpha+1)}{C_B'(p,q)(\alpha+\beta)^2} \|u_0\|_{L^1(\Omega)}^{\beta-1},$$

for $\beta > 1$. Here, $C_B(\frac{1}{2}(\alpha+1), 2)$ and $C_B'(p,q)$ are the constants in the Beckner inequalities (1.19) and (1.24), respectively, with $p = (\alpha+\beta)/2$ and $q = 2(\alpha+1)/(\alpha+\beta)$.

Proof. Let $\beta > 0$. We compute

$$\begin{aligned} \frac{dE_\alpha}{dt} &= -\frac{4\alpha\beta}{(\alpha+1)^2} \int_\Omega u^{\beta-1} |\nabla u^{(\alpha+1)/2}|^2 dx \\ &\le -\frac{4\alpha\beta}{(\alpha+1)^2} \inf_\Omega (u_0^{\beta-1}) \int_\Omega |\nabla u^{(\alpha+1)/2}|^2 dx. \end{aligned} \tag{1.29}$$

By means of the Beckner inequality (1.19) with $p = (\alpha+1)/2$, $q = 2$, and $f = u^{(\alpha+1)/2}$, we find that

$$\frac{dE_\alpha}{dt} \le -\frac{4\alpha\beta}{C_B(p,2)(\alpha+1)} \inf_\Omega (u_0^{\beta-1}) E_\alpha, \tag{1.30}$$

and Gronwall's lemma proves the claim. The restriction $p \le 1$ in Lemma 4 is equivalent to $\alpha \le 1$. Next, let $\beta > 1$. By means of Lemma 7, with $p = (\alpha+\beta)/2$, $q = 2(\alpha+1)/(\alpha+\beta)$, and $f = u^{(\alpha+\beta)/2}$, it follows that

$$\|u\|_{L^{\alpha+1}(\Omega)}^{\beta-1} \left(\int_\Omega u^{\alpha+1} dx - \left(\int_\Omega u dx \right)^{\alpha+1} \right) \le C_B'(p,q) \int_\Omega |\nabla u^{(\alpha+\beta)/2}|^2 dx.$$

Hence, we can estimate

$$\begin{aligned} \frac{dE_\alpha}{dt} = -\frac{4\alpha\beta}{(\alpha+\beta)^2} \int_\Omega |\nabla u^{(\alpha+\beta)/2}|^2 dx &\le -\frac{4\alpha\beta(\alpha+1)}{(\alpha+\beta)^2} \frac{\|u\|_{L^{\alpha+1}(\Omega)}^{\beta-1}}{C_B'(p,q)} E_\alpha[u] \\ &\le -\frac{4\alpha\beta(\alpha+1)}{(\alpha+\beta)^2} \frac{\|u_0\|_{L^1(\Omega)}^{\beta-1}}{C_B'(p,q)} E_\alpha[u], \end{aligned}$$

and Gronwall's lemma gives the conclusion. Note that in the last step of the inequality we used $\|u\|_{L^{\alpha+1}(\Omega)} \ge \|u\|_{L^1(\Omega)} = \|u_0\|_{L^1(\Omega)}$. $\square$

1.4.2 The discrete case

We prove a result which is the discrete analogon of Theorem 9. The finite-volume scheme (1.7) permits to define uniquely a piecewise constant solution at each time step: $u^k = \sum_{K\in\mathcal{T}} u_K^k \mathbf{1}_K$. Then the discrete entropies at each time step $E_\alpha[u^k]$ are defined in (1.8).

Theorem 11 (Polynomial decay)**.** *Let $\alpha > 0$ and $\beta > 1$. Let $(u_K^k)_{K\in\mathcal{T},k\geq 0}$ be the solution to the finite-volume scheme (1.7) with $\inf_{K\in\mathcal{T}} u_K^0 \geq 0$. Then*

$$E_\alpha[u^k] \leq \frac{1}{(c_1 t^k + c_2)^{(\alpha+1)/(\beta-1)}}, \quad k \geq 0,$$

where

$$c_1 = \frac{\beta-1}{\alpha+\beta}\left(\frac{(\alpha+1)(\alpha+\beta)}{4\alpha\beta}\left(\frac{C_b(p,q)}{\alpha+1}\right)^{(\alpha+\beta)/(\alpha+1)} + \triangle t E_\alpha[u^0]^{(\alpha+1)/(\beta-1)}\right)^{-1},$$
$$c_2 = E_\alpha[u^0]^{-(\beta-1)/(\alpha+1)},$$

and the constant $C_b(p,q)$ for $p = (\alpha+\beta)/2$ and $q = 2(\alpha+1)/(\alpha+\beta)$ is defined in Lemma 6.

Proof. In order to prove 9, we translate the proof of Theorem 9 to the discrete case. To this end, we use the elementary inequality $y^{\alpha+1} - x^{\alpha+1} \leq (\alpha+1)y^\alpha(y-x)$, which follows from the convexity of the mapping $x \mapsto x^{\alpha+1}$. Using the scheme (1.7), we obtain

$$\begin{aligned}
E_\alpha[u^{k+1}] - E_\alpha[u^k] &= \frac{1}{\alpha+1}\sum_{K\in\mathcal{T}} \mathrm{m}(K)\big((u_K^{k+1})^{\alpha+1} - (u_K^k)^{\alpha+1}\big) \\
&\leq \sum_{K\in\mathcal{T}} \mathrm{m}(K)(u_K^{k+1})^\alpha(u_K^{k+1} - u_K^k) \\
&\leq -\triangle t \sum_{K\in\mathcal{T}} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma (u_K^{k+1})^\alpha\big((u_K^{k+1})^\beta - (u_L^{k+1})^\beta\big).
\end{aligned}$$

Rearranging the sum leads to the discrete counterpart of (1.28):

$$E_\alpha[u^{k+1}] - E_\alpha[u^k] \leq -\triangle t \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma\big((u_K^{k+1})^\alpha - (u_L^{k+1})^\alpha\big)\big((u_K^{k+1})^\beta - (u_L^{k+1})^\beta\big). \tag{1.31}$$

Then, employing the inequality in Lemma 42 (see Appendix B), we deduce the discrete version of (1.28):

$$
\begin{aligned}
E_\alpha[u^{k+1}] - E_\alpha[u^k] &\le -\frac{4\alpha\beta\triangle t}{(\alpha+\beta)^2} \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1})^{(\alpha+\beta)/2} - (u_L^{k+1})^{(\alpha+\beta)/2} \big)^2 \\
&\le -\frac{4\alpha\beta\triangle t}{(\alpha+\beta)^2} |(u^{k+1})^{(\alpha+\beta)/2}|_{1,2,\mathcal{T}}^2.
\end{aligned}
$$

Applying the discrete Beckner inequality given in Lemma 6 with $p = (\alpha+\beta)/2$, $q = 2(\alpha+1)/(\alpha+\beta)$, and $f = (u^{k+1})^{(\alpha+\beta)/2}$, we obtain the discrete counterpart of (1.28):

$$
E_\alpha[u^{k+1}] - E_\alpha[u^k] \le -\frac{4\alpha\beta\triangle t}{(\alpha+\beta)^2} \left(\frac{\alpha+1}{C_b(p,q)} \right)^{(\alpha+\beta)/(\alpha+1)} E_\alpha[u^{k+1}]^{(\alpha+\beta)/(\alpha+1)}.
$$

The discrete nonlinear Gronwall lemma (see Corollary 41 in Appendix B) with

$$
\tau = \frac{4\alpha\beta\triangle t}{(\alpha+\beta)^2} \left(\frac{\alpha+1}{C_b(p,q)} \right)^{(\alpha+\beta)/(\alpha+1)}, \quad \gamma = \frac{\alpha+\beta}{\alpha+1} > 1,
$$

implies that

$$
E_\alpha[u^k] \le \frac{1}{(E_\alpha[u^0]^{1-\gamma} + c_1 t^k)^{1/(\gamma-1)}}, \quad k \ge 0,
$$

where $c_1 = (\gamma-1)/(1+\gamma\tau E_\alpha[u^0]^{\gamma-1})$. Finally, computing c_1 shows the result. □

The discrete analogon to Theorem 10 reads as follows.

Theorem 12 (Exponential decay for E_α)**.** *Let $(u_K^k)_{K\in\mathcal{T},k\ge0}$ be a solution to the finite-volume scheme* (1.7) *and let $0 < \alpha \le 1$, $\beta > 0$, $\inf_{K\in\mathcal{T}} u_K^0 \ge 0$. Then*

$$
E_\alpha[u^k] \le E_\alpha[u^0] e^{-\lambda t^k}, \quad k \ge 0.
$$

The constant λ is given by

$$
\lambda = \frac{4\alpha\beta}{C_b(\frac{1}{2}(\alpha+1),2)(\alpha+1)} \inf_{K\in\mathcal{T}} \big((u_K^0)^{\beta-1} \big) \ge 0,
$$

for $\beta > 0$, and

$$
\lambda = \frac{4\alpha\beta(\alpha+1)}{C_b'(p,q)(\alpha+\beta)^2} \|u_0\|_{L^1(\Omega)}^{\beta-1}
$$

for $\beta > 1$. Here $C_b'(p,q) > 0$ is the constant from Lemma 8 with $p = (\alpha+\beta)/2$ and $q = 2(\alpha+1)/(\alpha+\beta)$.

Proof. Let $\alpha \leq 1$ and $\beta > 0$. As in the proof of Theorem 11, we find that (see (1.31))

$$E_\alpha[u^{k+1}] - E_\alpha[u^k] \leq -\triangle t \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1})^\alpha - (u_L^{k+1})^\alpha\big)\big((u_K^{k+1})^\beta - (u_L^{k+1})^\beta\big).$$

Employing Corollary 43 (see Appendix B), we obtain

$$\begin{aligned} E_\alpha[u^{k+1}] - E_\alpha[u^k] &\leq -\frac{4\alpha\beta\triangle t}{(\alpha+1)^2} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \min\big\{(u_K^{k+1})^{\beta-1}, (u_L^{k+1})^{\beta-1}\big\} \\ &\quad \times \big((u_K^{k+1})^{(\alpha+1)/2} - (u_L^{k+1})^{(\alpha+1)/2}\big)^2 \\ &\leq -\frac{4\alpha\beta\triangle t}{(\alpha+1)^2} \inf_{K\in\mathcal{T}} (u_K^{k+1})^{\beta-1} |(u^{k+1})^{(\alpha+1)/2}|_{1,2,\mathcal{T}}^2, \end{aligned}$$

which is the discrete counterpart of (1.29). Then, applying the discrete Beckner inequality given in Lemma 6 with $p = (\alpha+1)/2$, $q = 2$, and $f = u^{(\alpha+1)/2}$, we obtain the analogon of (1.30):

$$E_\alpha[u^{k+1}] - E_\alpha[u^k] \leq -\frac{4\alpha\beta\triangle t}{C_b(\frac{1}{2}(\alpha+1),2)(\alpha+1)} \inf_{K\in\mathcal{T}} (u_K^0)^{\beta-1} E_\alpha[u^{k+1}],$$

and the Gronwall lemma shows the claim.

Next, let $\beta > 1$. As in the proof of Theorem 11, we find that

$$E_\alpha[u^{k+1}] - E_\alpha[u^k] \leq -\frac{4\alpha\beta\triangle t}{(\alpha+\beta)^2} |(u^{k+1})^{(\alpha+1)/2}|_{1,2,\mathcal{T}}^2.$$

We apply Lemma 8 with $p = (\alpha+\beta)/2$, $q = 2(\alpha+1)/(\alpha+\beta)$, and $f = u^{(\alpha+\beta)/2}$ to obtain

$$\begin{aligned} E_\alpha[u^{k+1}] - E_\alpha[u^k] &\leq -\frac{4\alpha\beta(\alpha+1)\triangle t}{(\alpha+\beta)^2} \frac{\|u^{k+1}\|_{L^{\alpha+1}(\Omega)}^{\beta-1}}{C_b'(p,q)} E_\alpha[u^{k+1}] \\ &\leq -\frac{4\alpha\beta(\alpha+1)\triangle t}{(\alpha+\beta)^2} \frac{\|u^0\|_{L^1(\Omega)}^{\beta-1}}{C_b'(p,q)} E_\alpha[u^{k+1}]. \end{aligned}$$

Then Gronwall's lemma finishes the proof. $\square$

1.5 First-order entropies: from the continuous to the discrete level

In this section, we consider the diffusion equation (1.1) on the torus $\Omega = \mathbb{T}^d$ and we first prove the exponential decay of the first-order entropies.
In the discrete setting, we consider the diffusion equation (1.1) on the half open unit cube $[0,1)^d \subset \mathbb{R}^d$ with multiperiodic boundary conditions (this is topologically equivalent to the torus $\mathbb{T}^d$). By identifying "opposite" faces on $\partial\Omega$, we can construct a family of control volumes and a family of edges in such a way that every face is an interior face. Then cells with such identified faces are neighboring cells. The numerical scheme we consider is similar to (1.7).

1.5.1 The continuous case

The exponential decay for the first-order entropies (1.4) is given, for the one-dimensional case, in the following theorem.

Theorem 13 (Exponential decay of F_α in one space dimension)**.** *Let u be a smooth solution to* (1.1) *on the one-dimensional torus $\Omega = \mathbb{T}$. Let $u_0 \in L^\infty(\Omega)$ with $\inf_\Omega u_0 \geq 0$ and let $\alpha, \beta > 0$ satisfy $-2 \leq \alpha - 2\beta < 1$. Then*

$$F_\alpha[u(t)] \leq F_\alpha[u_0]e^{-\Lambda t}, \quad 0 \leq t \leq T,$$

where

$$\Lambda = \frac{2\beta}{C_P^2} \inf_\Omega (u_0^{\alpha+\beta-\gamma-1}) \inf_\Omega (u_0^{\gamma-\alpha}) \geq 0, \quad \gamma = \frac{2}{3}(\alpha+\beta-1),$$

where $C_P > 0$ is the Poincaré constant in (1.10)*.*

Proof. We slightly extend the entropy construction method of [58]. The time derivative of the entropy reads as

$$\begin{aligned}\frac{dF_\alpha}{dt} &= \frac{\alpha}{2}\int_\Omega (u^{\alpha/2})_x (u^{\alpha/2-1}u_t)_x dx = -\frac{\alpha}{2}\int_\Omega (u^{\alpha/2})_{xx} u^{\alpha/2-1}(u^\beta)_{xx} dx \\ &= -\frac{\alpha^2\beta}{4}\int_\Omega u^{\alpha+\beta-1}\left(\left(\frac{\alpha}{2}-1\right)(\beta-1)\xi_G^4 + \left(\frac{\alpha}{2}+\beta-2\right)\xi_G^2\xi_L + \xi_L^2\right)dx,\end{aligned}$$

where we introduced

$$\xi_G = \frac{u_x}{u}, \quad \xi_L = \frac{u_{xx}}{u}.$$

This integral is compared to

$$\int_\Omega u^{\alpha+\beta-\gamma-1}(u^{\gamma/2})_{xx}^2 dx = \frac{\gamma^2}{4}\int_\Omega u^{\alpha+\beta-1}\left(\left(\frac{\gamma}{2}-1\right)^2\xi_G^4 + (\gamma-2)\xi_G^2\xi_L + \xi_L^2\right)dx,$$

where, in contrast to the method of [58], $\gamma \neq 0$ gives an additional degree of freedom in the calculations. In the one-dimensional situation, there is only one relevant integration-by-parts rule:

$$0 = \int_\Omega (u^{\alpha+\beta-4} u_x^3)_x dx = \int_\Omega u^{\alpha+\beta-1} \big((\alpha+\beta-4)\xi_G^4 + 3\xi_G^2 \xi_L \big) dx.$$

We introduce the polynomials

$$S_0(\xi) = \left(\frac{\alpha}{2} - 1\right)(\beta - 1)\xi_G^4 + \left(\frac{\alpha}{2} + \beta - 2\right)\xi_G^2 \xi_L + \xi_L^2, \tag{1.32}$$

$$D_0(\xi) = \left(\frac{\gamma}{2} - 1\right)^2 \xi_G^4 + (\gamma - 2)\xi_G^2 \xi_L + \xi_L^2, \tag{1.33}$$

$$T(\xi) = (\alpha + \beta - 4)\xi_G^4 + 3\xi_G^2 \xi_L,$$

where $\xi = (\xi_G, \xi_L)$. We wish to show that there exist numbers $c, \gamma \in \mathbb{R}$ ($\gamma \neq 0$) and $\kappa > 0$ such that

$$S(\xi) = S_0(\xi) + cT(\xi) - \kappa D_0(\xi) \geq 0 \quad \text{for all } \xi \in \mathbb{R}^2.$$

The polynomial S can be written as $S(\xi) = a_1 \xi_G^4 + a_2 \xi_G^2 \xi_L + (1 - \kappa)\xi_L^2$, where

$$a_1 = -\frac{1}{4}(\gamma - 2)^2 \kappa + (\alpha + \beta - 4)c + \frac{1}{2}(\alpha - 2)(\beta - 1),$$

$$a_2 = -(\gamma - 2)\kappa + 3c + \frac{1}{2}(\alpha + 2\beta - 4).$$

Therefore, the maximal value for κ is $\kappa = 1$. Let $\kappa = 1$. Then we need to eliminate the mixed term $\xi_G^2 \xi_L$. The solution of $a_2 = 0$ is given by $c = -\frac{1}{6}(\alpha + 2\beta - 2\gamma)$, which leads to

$$a_1 = -\frac{1}{4}\left(\gamma - \frac{2}{3}(\alpha + \beta - 1)\right)^2 - \frac{1}{18}(\alpha - 2\beta - 1)(\alpha - 2\beta + 2).$$

Choosing $\gamma = \frac{2}{3}(\alpha + \beta - 1)$ to maximize a_1, we find that $a_1 \geq 0$ and hence $S(\xi) \geq 0$ if and only if $-2 \leq \alpha - 2\beta \leq 1$.

Using the Poincaré inequality (1.10) and the maximum principle, we obtain

$$\begin{aligned}
\frac{dF_\alpha}{dt} &= -\frac{\alpha^2 \beta}{4} \int_\Omega u^{\alpha+\beta-1} S_0(\xi) dx = -\frac{\alpha^2 \beta}{4} \int_\Omega u^{\alpha+\beta-1} (S_0(\xi) + cT(\xi)) dx \\
&\leq -\frac{\alpha^2 \beta}{4} \int_\Omega u^{\alpha+\beta-1} D_0(\xi) dx = -\frac{\alpha^2 \beta}{\gamma^2} \int_\Omega u^{\alpha+\beta-\gamma-1} (u^{\gamma/2})_{xx}^2 dx \\
&\leq -\frac{\alpha^2 \beta}{\gamma^2} \inf_{\Omega \times (0,\infty)} (u^{\alpha+\beta-\gamma-1}) \int_\Omega (u^{\gamma/2})_{xx}^2 dx \\
&\leq -\frac{\alpha^2 \beta}{\gamma^2 C_P^2} \inf_\Omega (u_0^{\alpha+\beta-\gamma-1}) \int_\Omega (u^{\gamma/2})_x^2 dx.
\end{aligned}$$

And therefore

$$\frac{dF_\alpha}{dt} \leq -\frac{2\beta}{C_P^2} \inf_\Omega(u_0^{\alpha+\beta-\gamma-1}) \inf_\Omega(u_0^{\gamma-\alpha}) F_\alpha.$$

For the last inequality, we use the identity $(u^{\gamma/2})_x = \frac{\gamma}{\alpha} u^{(\gamma-\alpha)/2}(u^{\alpha/2})_x$, which cancels out the ratio α^2/γ^2. An application of Gronwall's lemma finishes the proof. $\square$

We turn to the multi-dimensional case.

Theorem 14 (Exponential decay of F_α in several space dimensions)**.** *Let u be a smooth solution to* (1.1) *on the torus $\Omega = \mathbb{T}^d$. Let $u_0 \in L^\infty(\Omega)$ with $\inf_\Omega u_0 > 0$ and let*

$$\begin{aligned}(\alpha, \beta) \in M_d = \big\{(\alpha,\beta) \in \mathbb{R}^2 :\ &(\alpha - 2\beta - 1)(\alpha - 2\beta + 2) < 0 \textit{ and} \\ &(2 - 2\alpha + 2\beta - d + \alpha d)(4 - 4\beta - 2d + \alpha d + 2\beta + 2\beta d) > 0\big\}\end{aligned}$$

(see Figure 1.1). Then there exists $\Lambda > 0$, depending on α, β, d, u_0, and Ω such that

$$F_\alpha[u(t)] \leq F_\alpha[u_0]e^{-\Lambda t}, \quad t \geq 0.$$

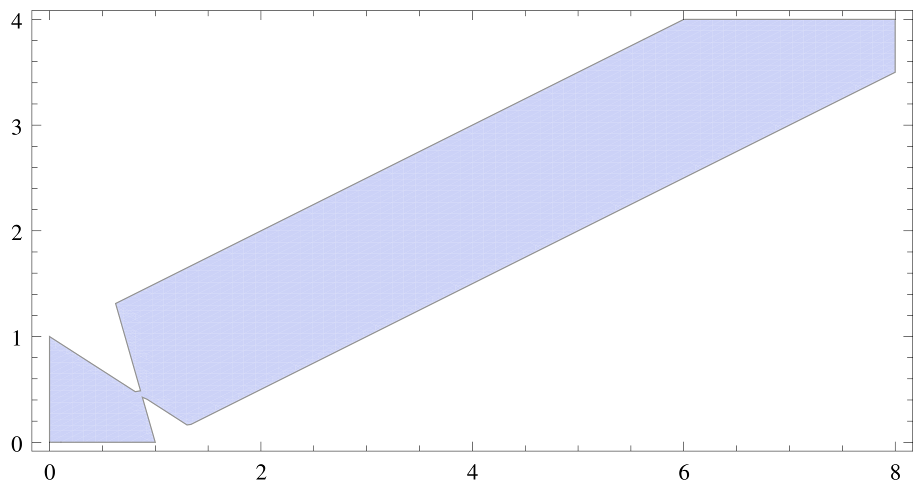

Figure 1.1: Illustration of the set M_d, defined in Theorem 14, for $d = 9$.

Proof. The time derivative of the first-order entropy becomes

$$\frac{dF_\alpha}{dt} = -\frac{\alpha}{2}\int_\Omega u^{\alpha/2-1}\Delta(u^{\alpha/2})\Delta(u^\beta)dx = -\frac{\alpha^2\beta}{4}\int_\Omega u^{\alpha+\beta-1}S_0 dx, \qquad (1.34)$$

where S_0 is defined in (1.32) with the (scalar) variables $\xi_G = |\nabla u|/u$, $\xi_L = \Delta u/u$. We compare this integral to

$$\int_\Omega u^{\alpha+\beta-\gamma-1}(\Delta(u^{\gamma/2}))^2 dx = \frac{\gamma^2}{4}\int_\Omega u^{\alpha+\beta-1}D_0 dx,$$

where D_0 is as in (1.33) and $\gamma \neq 0$. In contrast to the one-dimensional case, we employ two integration-by-parts rules:

$$
\begin{aligned}
0 &= \int_\Omega \operatorname{div}\left(u^{\alpha+\beta-4}|\nabla u|^2 \nabla u\right) dx = \int_\Omega u^{\alpha+\beta-1} T_1 dx, \\
0 &= \int_\Omega \operatorname{div}\left(u^{\alpha+\beta-3}(\nabla^2 u - \Delta \mathbb{I}) \cdot \nabla u\right) dx = \int_\Omega u^{\alpha+\beta-1} T_2 dx,
\end{aligned}
$$

where

$$
\begin{aligned}
T_1 &= (\alpha+\beta-4)\xi_G^4 + 2\xi_{GHG} + \xi_G^2 \xi_L, \\
T_2 &= (\alpha+\beta-3)\xi_{GHG} - (\alpha+\beta-3)\xi_G^2\xi_L + \xi_H^2 - \xi_L^2,
\end{aligned}
$$

and $\xi_{GHG} = u^{-3}\nabla u^\top \nabla^2 u \nabla u$, $\xi_H = u^{-1}\|\nabla^2 u\|$. Here, $\|\nabla^2 u\|$ denotes the Frobenius norm of the Hessian.
In order to compare $\nabla^2 u$ and Δu, we employ Lemma 2.1 of [59]:

$$
\|\nabla^2 u\|^2 \geq \frac{1}{d}(\Delta u)^2 + \frac{d}{d-1}\left(\frac{\nabla u^\top \nabla^2 u \nabla u}{|\nabla u|^2} - \frac{\Delta u}{d}\right)^2 .
$$

Therefore, there exists $\xi_R \in \mathbb{R}$ such that

$$
\xi_H^2 = \frac{\xi_L^2}{d} + \frac{d}{d-1}\left(\frac{\xi_{GHG}}{\xi_G^2} - \frac{1}{d}\xi_L\right)^2 + \xi_R^2 = \frac{\xi_L^2}{d} + \frac{d}{d-1}\xi_S^2 + \xi_R^2,
$$

where we introduced $\xi_S = \xi_{GHG}/\xi_G^2 - \xi_L/d$. Rewriting the polynomials T_1 and T_2 in terms of $\xi = (\xi_G, \xi_L, \xi_S, \xi_R) \in \mathbb{R}^4$ leads to:

$$
\begin{aligned}
T_1(\xi) &= (\alpha+\beta-4)\xi_G^4 + \frac{2+d}{d}\xi_G^2\xi_L + 2\xi_G^2\xi_S, \\
T_2(\xi) &= \frac{1-d}{d}(\alpha+\beta-3)\xi_G^2\xi_L + \frac{1-d}{d}\xi_L^2 + \xi_S\xi_G^2(\alpha+\beta-3) + \frac{d}{d-1}\xi_S^2 + \xi_R^2.
\end{aligned}
$$

We wish to find c_1, c_2, $\gamma \in \mathbb{R}$ ($\gamma \neq 0$) and $\kappa > 0$ such that

$$
S(\xi) = S_0(\xi) + c_1 T_1(\xi) + c_2 T_2(\xi) - \kappa D_0(\xi) \geq 0 \quad \text{for all } \xi \in \mathbb{R}^4.
$$

The polynomial S can be written as

$$
\begin{aligned}
S(\xi) &= a_1\xi_G^4 + a_2\xi_G^2\xi_L + a_3\xi_L^2 + a_4\xi_G^2\xi_S + a_5\xi_S^2 + c_2\xi_R^2, \text{ where} \\
a_1 &= \left(\frac{\alpha}{2} - 1\right)(\beta - 1) + (\alpha+\beta-4)c_1 - \left(\frac{\gamma}{2} - 1\right)^2 \kappa, \\
a_2 &= \frac{\alpha}{2} + \beta - 2 + \left(\frac{2}{d} + 1\right)c_1 - (\alpha+\beta-3)\frac{d-1}{d}c_2 - (\gamma-2)\kappa, \\
a_3 &= 1 + \frac{1-d}{d}c_2 - \kappa,
\end{aligned}
$$

$$a_4 = 2c_1 + (\alpha+\beta-3)c_2,$$
$$a_5 = \frac{d}{d-1}c_2.$$

We remove the variable ξ_R by imposing the condition that $c_2 \geq 0$. The remaining polynomial can be reduced to a quadratic polynomial by setting $x = \xi_L/\xi_G^2$ and $y = \xi_S/\xi_G^2$:

$$S(x,y) \geq a_1 + a_2x + a_3x^2 + a_4y + a_5y^2 \geq 0 \quad \text{for all } x,\ y \in \mathbb{R}. \tag{1.35}$$

This quadratic decision problem can be solved by employing the computer algebra system `Mathematica`. The result of the command

```
Resolve[ForAll[{x, y}, Exists[{C1, C2, kappa, gamma},
a1 + a2*x + a3*x^2 + a4*y + a5*y^2 >= 0 && kappa > 0
&& gamma != 0]], Reals]
```

yields all $(\alpha,\beta) \in \mathbb{R}^2$ such that there exist c_1, c_2, $\gamma \in \mathbb{R}$ ($\gamma \neq 0$) and $\kappa > 0$ such that (1.35) holds. This region equals the set M_d, defined in the statement of this theorem.
Similar to the one-dimensional case, we infer that

$$\frac{dF_\alpha}{dt} \leq -\frac{\alpha^2\beta\kappa}{4}\int_\Omega u^{\alpha+\beta-1}D_0(\xi)dx = -\frac{\alpha^2\beta\kappa}{\gamma^2}\int_\Omega u^{\alpha+\beta-\gamma-1}(\Delta u^{\gamma/2})^2dx.$$

Thus, proceeding as in the proof of Theorem 13 and using the identity

$$\int_\Omega (\Delta f)^2 dx = \int_\Omega \|\nabla^2 f\|^2 dx$$

for smooth functions f (which can be derived by integrating by parts twice), we obtain

$$\frac{dF_\alpha}{dt} \leq -\frac{2\beta\kappa}{C_P^2}\inf_\Omega(u_0^{\alpha+\beta-\gamma-1})\inf_\Omega(u_0^{\gamma-\alpha})F_\alpha.$$

Gronwall's lemma concludes the proof. $\square$

Remark 15. Under the additional constraints $a_2 = a_3 = 0$, we are able to solve the decision problem (1.35) without the help of the computer algebra system. The solution set, however, is slightly smaller than M_d, which is obtained from `Mathematica` without these constraints. Indeed, we can compute c_1 and c_2 from the equations $a_2 = a_3 = 0$ to give

$$c_1 = \frac{d}{d+2}\left(\frac{\alpha}{2} - 1 + \kappa(1+\gamma-\alpha-\beta)\right), \quad c_2 = \frac{d(1-\kappa)}{d-1}.$$

The decision problem (1.35) reduces to

$$a_1 + a_4 y + a_5 y^2 \geq 0 \quad \text{for all } y \in \mathbb{R}.$$

If $\kappa < 1$, it holds $c_2 > 0$ and consequently, $a_5 > 0$. Therefore, the above polynomial is nonnegative for all $y \in \mathbb{R}$ if it has no real roots, i.e., if

$$0 \leq 4a_1 a_5 - a_4^2 = q_0 + q_1\gamma + q_2\gamma^2$$

for some $\gamma \neq 0$, where (for $d > 1$)

$$q_2 = -\frac{d^2\kappa}{(d+2)^2(d-1)^2}\big(3d(d-4)\kappa + (d+2)^2\big) < 0,$$

and q_0, q_1 are polynomials depending on d, α, β, and κ. The above problem is solvable if and only if there exist real roots, i.e. if

$$0 \leq q_1^2 - 4q_0q_2 = \frac{4\kappa(1-\kappa)}{(d+2)^2(d-1)^2}(s_0 + s_1\kappa + s_2\kappa^2),$$

where

$$\begin{aligned} s_0 &= -d(5d-8) + 6d(d-1)\alpha + 2d(d+2)\beta + 2(d+2)\alpha\beta - (2d^2+1)\alpha^2 \\ &\quad - (d+2)^2\beta^2, \\ s_1 &= 2d(3d-4) - 2d(4d-3)\alpha - 4d(d+1)\beta + 2d(3d-5)\alpha\beta + 2d(d+1)\alpha^2 \\ &\quad - 2d(d-6)\beta^2, \\ s_2 &= -d^2(\alpha+\beta-1)^2. \end{aligned}$$

We set $f(\kappa) = s_0 + s_1\kappa + s_2\kappa^2$. We have to find $0 < \kappa < 1$ such that $f(\kappa) \geq 0$. Since $s_2 \leq 0$, this is possible if $f(\kappa)$ possesses two (not necessarily distinct) real roots κ_0 and κ_1 and if at least one of those roots is between zero and one. Since $f(1) = -(d-1)^2(\alpha - 2\beta)^2 \leq 0$, there are only two possibilities for κ_0 and κ_1: either $\kappa_0 \leq 0 \leq \kappa_1 \leq 1$ or $0 \leq \kappa_0 \leq \kappa_1 \leq 1$. The first case holds if $f(0) = s_0 \geq 0$, the latter one if

$$f'(0) = s_1 \geq 0, \quad f'(1) = s_1 + 2s_2 \leq 0, \tag{1.36}$$

$$\begin{aligned} s_1^2 - 4s_0s_2 &= -4d^2(\alpha - 2\beta + 2)(\alpha - 2\beta - 1)(4 - 2d + d\alpha + 2d\beta) \\ &\quad \times (2 - d + (d-2)\alpha + 2\beta) \geq 0. \end{aligned} \tag{1.37}$$

The set of all $(\alpha, \beta) \in \mathbb{R}^2$ fulfilling these conditions is illustrated in Figure 1.2. □

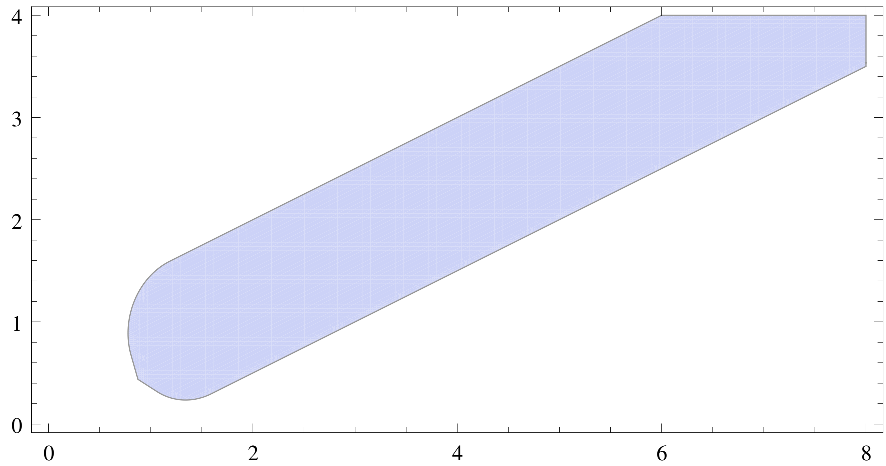

Figure 1.2: Set of all (α, β) fulfilling $s_0 \geq 0$, (1.36), and (1.37) for $d = 9$.

1.5.2 The discrete case

At the discrete level, we establish the decay of the first-order entropies in any dimension, with an exponential rate in one space dimension. We recall that the discrete first-order entropies are defined by (1.9).

Theorem 16 (Exponential decay of F_α^d)**.** *Let $(u_K^k)_{K\in\mathcal{T},\, k\geq 0}$ be the solution to the finite-volume scheme* (1.7) *with $\Omega = \mathbb{T}^d$ and $\inf_{K\in\mathcal{T}} u_K^0 \geq 0$. Then, for all $1 \leq \alpha \leq 2$, and $\alpha = 2\beta$ it follows that*

$$F_\alpha^d[u^{k+1}] \leq F_\alpha^d[u^k], \quad k \in \mathbb{N}.$$

Furthermore for all $1 \leq \alpha \leq 2$, $-2 < \alpha - 2\beta < 1$, if $d = 1$ and the grid is uniform with N subintervals, there exists $0 < \varepsilon \leq 1$ such that

$$F_\alpha^d[u^k] \leq F_\alpha^d[u_0] e^{-\lambda t^k},$$

where $\lambda = \frac{\varepsilon\alpha^2}{\beta} \sin^2 \frac{\pi}{N} \min_{i=1,\ldots,N} \left((u_i^0)^{\alpha-\beta-1}\right) \geq 0$.

Proof. The difference $G_\alpha = F_\alpha^d[u^{k+1}] - F_\alpha^d[u^k]$ can be written as

$$G_\alpha = \frac{1}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\text{int}},\\ \sigma=K|L}} \tau_\sigma \left(\left((u_K^{k+1})^{\alpha/2} - (u_L^{k+1})^{\alpha/2}\right)^2 - \left((u_K^k)^{\alpha/2} - (u_L^k)^{\alpha/2}\right)^2 \right).$$

Introducing $a_K = (u_K^{k+1} - u_K^k)/\tau$, we find that

$$G_\alpha = \frac{1}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\text{int}},\\ \sigma=K|L}} \tau_\sigma \Big(\left((u_K^{k+1})^{\alpha/2} - (u_L^{k+1})^{\alpha/2}\right)^2 \\ - \left((u_K^{k+1} - \tau a_K)^{\alpha/2} - (u_L^{k+1} - \tau a_L)^{\alpha/2}\right)^2 \Big).$$

We claim that G_α is concave with respect to τ. Indeed, we compute

$$\begin{aligned}
\frac{\partial G_\alpha}{\partial \tau} &= \frac{\alpha}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1}-\tau a_K)^{\alpha/2} - (u_L^{k+1}-\tau a_L)^{\alpha/2} \big) \\
&\quad \times \big((u_K^{k+1}-\tau a_K)^{\alpha/2-1} a_K - (u_L^{k+1}-\tau a_L)^{\alpha/2-1} a_L \big), \\
\frac{\partial^2 G_\alpha}{\partial \tau^2} &= -\frac{\alpha^2}{4} \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1}-\tau a_K)^{\alpha/2-1} a_K - (u_L^{k+1}-\tau a_L)^{\alpha/2-1} a_L \big)^2 \\
&\quad - \frac{\alpha}{2}\Big(\frac{\alpha}{2}-1\Big) \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1}-\tau a_K)^{\alpha/2} - (u_L^{k+1}-\tau a_L)^{\alpha/2} \big) \\
&\quad \times \big((u_K^{k+1}-\tau a_K)^{\alpha/2-2} a_K^2 - (u_L^{k+1}-\tau a_L)^{\alpha/2-2} a_L^2 \big).
\end{aligned}$$

Replacing $u_K^{k+1}-\tau a_K$, $u_L^{k+1}-\tau a_L$ by u_K^k, u_L^k, respectively, the second derivative becomes

$$\begin{aligned}
\frac{\partial^2 G_\alpha}{\partial \tau^2} &= -\frac{\alpha^2}{4} \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^k)^{\alpha/2-1} a_K - (u_L^k)^{\alpha/2-1} a_L \big)^2 \\
&\quad - \frac{\alpha}{2}\Big(\frac{\alpha}{2}-1\Big) \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma \big((u_K^k)^{\alpha/2} - (u_L^k)^{\alpha/2} \big)\big((u_K^k)^{\alpha/2-2} a_K^2 - (u_L^k)^{\alpha/2-2} a_L^2 \big) \\
&= -\frac{\alpha}{4} \sum_{\substack{\sigma\in\mathcal{E}_{\rm int},\\ \sigma=K|L}} \tau_\sigma (c_1 a_K^2 + c_2 a_K a_L + c_3 a_L^2),
\end{aligned}$$

where

$$\begin{aligned}
c_1 &= (\alpha-2)\big((u_K^k)^{\alpha/2} - (u_L^k)^{\alpha/2} \big)(u_K^k)^{\alpha/2-2} + \alpha (u_K^k)^{\alpha-2}, \\
c_2 &= -2\alpha (u_K^k)^{\alpha/2-1}(u_L^k)^{\alpha/2-1}, \\
c_3 &= -(\alpha-2)\big((u_K^k)^{\alpha/2} - (u_L^k)^{\alpha/2} \big)(u_L^k)^{\alpha/2-2} + \alpha (u_L^k)^{\alpha-2}.
\end{aligned}$$

We show that the quadratic polynomial in the variables a_K and a_L is nonnegative for all u_K^k and u_L^k. This is the case if and only if $c_1 \geq 0$ and $4c_1c_3 - c_2^2 \geq 0$. The former condition is equivalent to

$$2(\alpha-1)(u_K^k)^{\alpha-2} \geq (\alpha-2)(u_K^k)^{\alpha/2-2}(u_L^k)^{\alpha/2},$$

which is true for $1 \leq \alpha \leq 2$. After an elementary computation, the latter condition becomes

$$4c_1c_3 - c_2^2 = 8(\alpha-1)(2-\alpha)(u_K^k)^{\alpha/2-2}(u_L^k)^{\alpha/2-2}\big((u_K^k)^{\alpha/2} - (u_L^k)^{\alpha/2} \big)^2 \geq 0$$

for $1 \le \alpha \le 2$. This proves the concavity of $\tau \mapsto G_\alpha(\tau)$.
A Taylor expansion and $G_\alpha(0) = 0$ leads to

$$\begin{aligned}
G_\alpha(\tau) &\le G_\alpha(0) + \tau \frac{\partial G_\alpha}{\partial \tau}(0) \\
&= \frac{\alpha\tau}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1})^{\alpha/2} - (u_L^{k+1})^{\alpha/2}\big)\big((u_K^{k+1})^{\alpha/2-1} a_K - (u_L^{k+1})^{\alpha/2-1} a_L\big) \\
&= \frac{\alpha\tau}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \big((u_K^{k+1})^{\alpha/2} - (u_L^{k+1})^{\alpha/2}\big)(u_K^{k+1})^{\alpha/2-1} a_K \\
&\quad + \frac{\alpha\tau}{2} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \big((u_L^{k+1})^{\alpha/2} - (u_K^{k+1})^{\alpha/2}\big)(u_L^{k+1})^{\alpha/2-1} a_L.
\end{aligned}$$

Replacing a_K and a_L by scheme (1.7) and rearranging the terms, we infer that

$$\begin{aligned}
G_\alpha(\triangle t) = &-\frac{\alpha\triangle t}{2\mathrm{m}(K)} \sum_{K\in\mathcal{T}} \sum_{\substack{\sigma\in\mathcal{E}_{\mathrm{int}},\\ \sigma=K|L}} \tau_\sigma \sum_{\substack{\widetilde{\sigma}\in\mathcal{E}_{\mathrm{int}},\\ \widetilde{\sigma}'=K|M}} \tau_{\widetilde{\sigma}} (u_K^{k+1})^{\alpha/2-1} \\
&\times \big((u_K^{k+1})^\beta - (u_M^{k+1})^\beta\big)\big((u_K^{k+1})^{\alpha/2} - (u_L^{k+1})^{\alpha/2}\big). \qquad (1.38)
\end{aligned}$$

Note that the expression on the right-hand side is the discrete counterpart of the integral

$$-\frac{\alpha}{2}\int_\Omega u^{\alpha/2-1}(u^\beta)_{xx}(u^{\alpha/2})_{xx} dx,$$

appearing in (1.34). The condition $\alpha = 2\beta$ immediately implies the monotonicity of $k \mapsto F_\alpha^d[u^k]$.
For the proof of the second statement, we let $d = 1$ and decompose the interval Ω in N subintervals $K_1, \ldots, K_N$ of length $h > 0$. Because of the periodic boundary conditions, we may set $u_{N+1}^k = u_0^k$ and $u_{-1}^k = u_N^k$, where u_i^k is the approximation of the mean value of $u(\cdot, t^k)$ on the subinterval K_i, $i = 1, \ldots, N$. Then, by using the discrete integrations-by-parts formula (Appendix A) with

$$\begin{aligned}
A &= \frac{\alpha}{2\beta} \\
B &= \frac{\alpha/2-1}{\beta} \\
\Rightarrow (2A - B - 1)(A + B - 2) &= \frac{(\alpha - 2\beta + 2)(\alpha - 2\beta - 1)}{2\beta^2}
\end{aligned}$$

we can estimate (1.38) as:

$$
\begin{aligned}
G_\alpha(\tau) &= -\frac{\alpha\tau}{4h}\sum_{i=1}^{N}(u_i^{k+1})^{\alpha/2-1}\left((u_{i+1}^{k+1})^\beta + (u_{i-1}^{k+1})^\beta - 2(u_i^{k+1})^\beta\right) \\
&\quad \times \left((u_{i+1}^{k+1})^{\alpha/2} + (u_{i-1}^{k+1})^{\alpha/2} - 2(u_i^{k+1})^{\alpha/2}\right) \\
&\le -\frac{\varepsilon\alpha^2\tau}{8h\beta}\min_{i=1,\ldots,N}\left((u_i^{k+1})^{\alpha-\beta-1}\right)\sum_{i=1}^{N}(z_i - z_{i-1})^2,
\end{aligned}
$$

for some $0 < \varepsilon \le 1$ and $z_i = (u_i^{k+1})^\beta - (u_{i+1}^{k+1})^\beta$. The periodic boundary conditions imply that $\sum_{i=1}^N z_i = 0$. Hence, we can employ the discrete Wirtinger inequality in [81, Theorem 1] to obtain

$$
\begin{aligned}
G_\alpha(\tau) &\le -\frac{\varepsilon\alpha^2\tau}{2h\beta}\sin^2\frac{\pi}{N}\min_{i=1,\ldots,N}\left((u_i^{k})^{\alpha-\beta-1}\right)\sum_{i=1}^{N} z_i^2 \\
&= -\frac{\varepsilon\alpha^2\tau}{h\beta}\sin^2\frac{\pi}{N}\min_{i=1,\ldots,N}\left((u_i^{k})^{\alpha-\beta-1}\right)F_\alpha^d[u^{k+1}].
\end{aligned}
$$

From the discrete maximum principle, it follows that

$$
\max_i (u_i^{k+1})^{-\alpha+\beta+1} \le \max_i (u_i^0)^{-\alpha+\beta+1} \Leftrightarrow \min_i (u_i^{k+1})^{\alpha-\beta-1} \ge \min_i (u_i^0)^{\alpha-\beta-1}
$$

Therefore,

$$
F_\alpha^d[u^{k+1}] - F_\alpha^d[u^k] = G_\alpha(\triangle t) \le -\frac{\varepsilon\alpha^2\tau}{h\beta}\sin^2\frac{\pi}{N}\min_{i=1,\ldots,N}\left((u_i^0)^{\alpha-\beta-1}\right)F_\alpha^d[u^{k+1}],
$$

and Gronwall's lemma finishes the proof. □

Remark 17 (Special case $\alpha = 2\beta$)**.** In this case, equation (1.38) can immediately be simplified to

$$
G_{2\beta}(\tau) \le -\frac{\beta\tau}{2h}\sum_{i=1}^{N}\Bigg(\sum_{j\in\{i-1,i+1\}}(u_i^{k+1})^{\beta-1}\left((u_i^{k+1})^\beta - (u_j{}^{k+1})^\beta\right)\Bigg)^2
$$

and we do not need the discrete integration-by-parts formula (Appendix A). In this case $\varepsilon = 1$ becomes optimal and the constant λ simplifies to $\lambda = 4\beta\sin^2(\pi/N)\min_i((u_i^0)^{2(\beta-1)})$.

1.6 Numerical experiments

We illustrate the time decay of the solutions to the discretized porous-medium ($\beta = 2$) and fast-diffusion equation ($\beta = 1/2$) in one and two space dimensions. First, let $\beta = 2$. We recall that the Barenblatt profile

$$
u_B(x,t) = (t+t_0)^{-A}\Big(C - \frac{B(\beta-1)}{2\beta}\frac{|x-x_0|^2}{(t+t_0)^{2B}}\Big)_+^{1/(\beta-1)}
$$

is a special solution to the porous-medium equation in the whole space. (Here, z_+ denotes the positive part of a function $z_+ := \max\{0, z\}$.) The constants are given by

$$A = \frac{d}{d(\beta-1)+2}, \quad B = \frac{1}{d(\beta-1)+2},$$

and C is typically determined by the initial datum by setting the total mass $\int_\Omega u(x,t)dx = \int_\Omega u(x,0)dx$. We choose $C = B(\beta-1)(2\beta)^{-1}(t_1+t_0)^{-2B}|x_1 - x_0|^2$, where $t_1 > 0$ is the smallest time for which $u(x_1, t_1) = 0$.
In the one-dimensional situation, we choose $\Omega = (0,1)$ with homogeneous Neumann boundary conditions and a uniform grid $(x_i, t^j) \in [0,1] \times [0, 0.2]$ with $1 \leq i \leq 50$ and $0 \leq j \leq 1000$, i.e., the space grid size is $\triangle x = 0.02$ and the time step size equals $\triangle t = 2 \cdot 10^{-4}$. We have chosen a very small time step size for a smoother graphical representation of the solution, but the implicit scheme clearly also works for time step sizes of the order of $\triangle x$ and for smaller values of $\triangle x$. The initial datum is given by the Barenblatt profile $u_B(\cdot, 0)$ with $x_0 = 0.5$, $x_1 = 1$ and $t_0 = 0.01$. The constant C is computed by using $t_1 = 0.1$, which yields $C \approx 0.091$. For $0 \leq t \leq 0.1$, the analytical solution corresponds to the Barenblatt profile.
The time decay of the zeroth- and first-order entropies are depicted in Figure 1.3 using a semi-logarithmic scale for various values of α. The decay rates are exponential for sufficiently large times, even for $\alpha > 1$ (compare to Theorem 12) and for $\alpha \neq 2\beta$ (see Theorem 16), which indicates that the conditions imposed in these theorems are technical. For small times, the decay seems to be faster than the decay in the large-time regime. This fact has been already observed in [20, Remark 4]. There is a significant change in the decay rate of the first-order entropies F_α^d for times around $t_1 = 0.1$. Indeed, the positive part of the discrete solution, which approximates the Barenblatt profile u_B for $t < t_1$, reaches the boundary and does not approximate u_B anymore. The change is more apparent for $\alpha < 1$.
Next, we investigate the two-dimensional situation (still with $\beta = 2$). The domain $\Omega = (0,1)^2$ is divided into 144 quadratic cells each of which consists of four control volumes (see Figure 1.4). Again, we employ the Barenblatt profile as the initial datum, choosing $t_0 = 0.01$, $t_1 = 0.1$, and $x_0 = (0.5, 0.5)$, and impose homogeneous boundary conditions. The time step size equals $\triangle t = 8 \cdot 10^{-4}$.
In Figure 1.5, the time evolution of the (logarithmic) zeroth- and first-order entropies are presented. Again, the decay appears to be exponential for large times, even for values of α not covered by the theoretical results of this chapter. At time $t = t_1$, the profile reaches the boundary of the domain. Since the radially symmetric profile does not reach the boundary everywhere simultane-

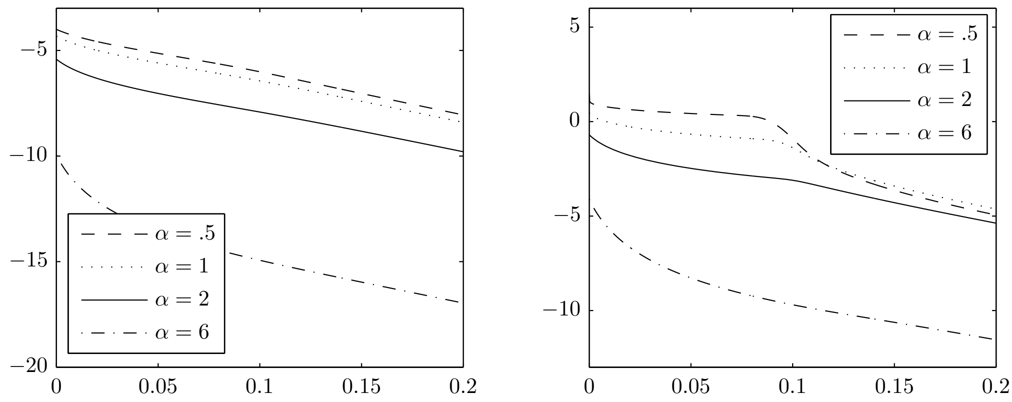

Figure 1.3: The natural logarithm of the entropies $\log(E_\alpha^d[u](t))$ (left) and $\log(F_\alpha^d[u](t))$ (right) versus time for different values of α ($\beta = 2$, $d = 1$).

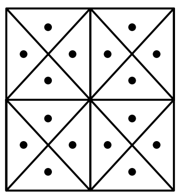

Figure 1.4: Four of the 144 cells used for the two-dimensional finite-volume scheme.

ously, the time decay rate of F_α^d does not change as distinctly as in Figure 1.3.

Let $\beta = 1/2$. The one-dimensional interval $\Omega = (0, 1)$ is discretized as before, using 51 grid points and the time step size is $\triangle t = 2 \cdot 10^{-4}$. We impose homogeneous Neumann boundary conditions. As initial datum, we choose the following truncated polynomial $u_0(x) = C((x_0 - x)(x - x_1))_+^2$, where $x_0 = 0.3$, $x_1 = 0.7$, and $C = 3000$. In the two-dimensional box $\Omega = (0, 1)^2$, we employ the discretization described above and the initial datum $u_0(x) = C(R^2 - |x - x_0|^2)_+^2$, where $R = 0.2$, $x_0 = (0.5, 0.5)$ and again $C = 3000$.
In the fast-diffusion case $\beta < 1$, we do not expect significant changes in the decay rate since the initial values propagate with infinite speed. This expectation is supported by the numerical results presented in Figures 1.6 and 1.7. For a large range of values of α, the decay rate is exponential, at least for large times. Interestingly, the rate seems to approach almost the same value for $\alpha \in \{0.5, 1, 2\}$ in Figure 1.7.

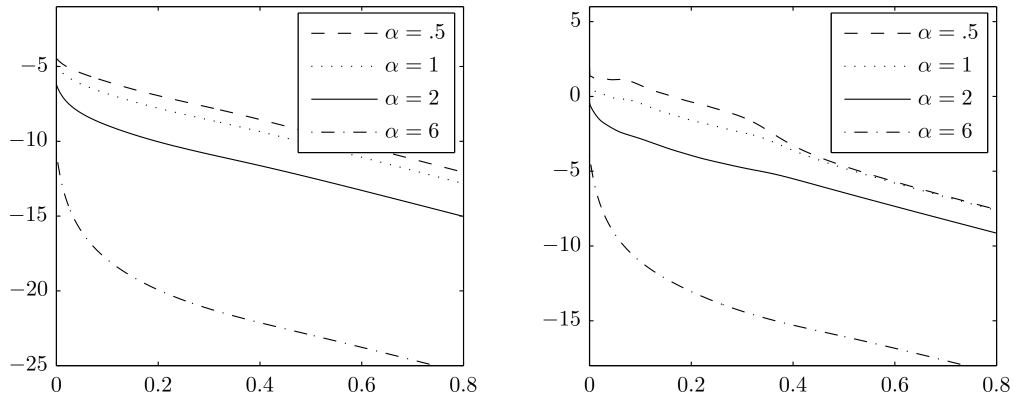

Figure 1.5: The natural logarithm of the entropies $\log(E_\alpha^d[u](t))$ (left) and $\log(F_\alpha^d[u](t))$ (right) versus time for different values of α ($\beta = 2$, $d = 2$).

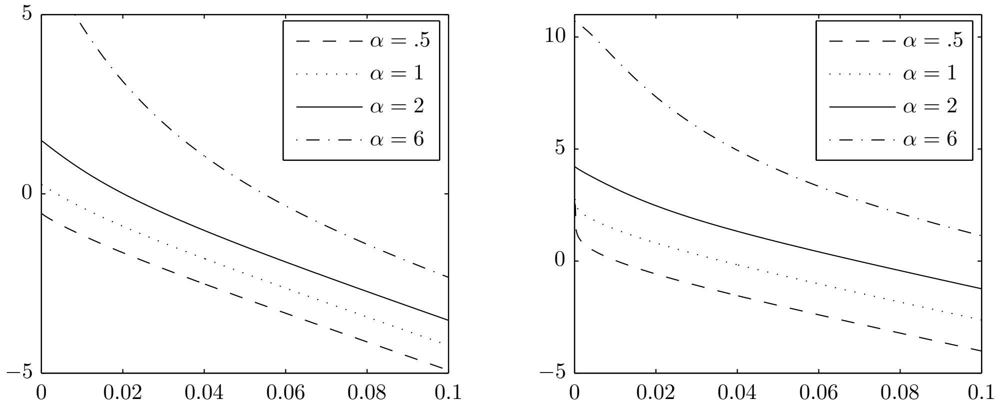

Figure 1.6: The natural logarithm of the entropies $\log(E_\alpha^d[u](t))$ (left) and $\log(F_\alpha^d[u](t))$ (right) versus time for different values of α ($\beta = 1/2$, $d = 1$).

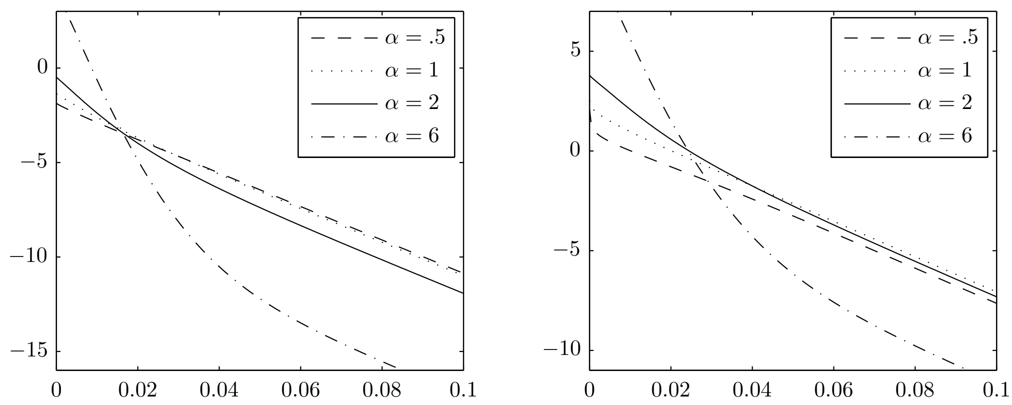

Figure 1.7: The natural logarithm of the entropies $\log(E^d_\alpha[u](t))$ (left) and $\log(F^d_\alpha[u](t))$ (right) versus time for different values of α ($\beta = 1/2$, $d = 2$).

Chapter 2

Runge-Kutta discretizations in time - A general approach

2.1 Introduction

In this chapter, we are mainly concerned with the entropy-preserving nature of the *time*-discretization of nonlinear evolution equations.
Evolution equations often contain some structural information reflecting inherent physical properties such as positivity of solutions, conservation laws, and entropy dissipation. Numerical schemes should be designed in such a way that these structural features are preserved on the discrete level in order to obtain accurate and stable algorithms. In the last decades, concepts of structure-preserving schemes, geometric integration, and compatible discretization have been developed [30], but much less is known about the preservation of entropy dissipation and large-time asymptotics.
Entropy-stable schemes were derived by Tadmor already in the 1980s [83] in the context of conservation laws, thus without (physical) diffusion. Later, entropy-dissipative schemes were developed for (finite-volume) discretizations of diffusion equations [13, 46, 48]. In [18], a finite-volume scheme which preserves the gradient-flow structure and hence the entropy structure is proposed. All these schemes are based on the implicit Euler method and are of first order (in time) only. Higher-order schemes which diminish the total variation were developed for hyperbolic conservation laws, and they are often based on flux or slope limiters [82]. More general approaches are known under the name of strong stability preserving schemes ensuring stability in the same norm as the forward Euler scheme. They are used, for instance, for method-of-lines approximations of partial differential equations. For Runge-Kutta discretizations with this property, we refer to [50, 57].

Further numerical approaches of higher-order entropy-dissipating schemes include the second-order predictor-corrector approximation of [73] and the higher-order semi-implicit Runge-Kutta (DIRK) method of [16], together with a spatial fourth-order central finite-difference discretization. In [17, 63], multistep time approximations were employed, but they can only be of second order at most and they dissipate only one specific entropy in comparison to all entropy functionals dissipated by the continuous equation. In this chapter, we remove these restrictions by investigating time-discrete Runge-Kutta schemes of order $p \geq 1$ for general diffusion equations.

We stress the fact that we are interested in the analysis of entropy-dissipating schemes by translating properties of the continuous equation to the semi-discrete level, i.e., we study the stability of the schemes. We do not investigate convergence, stiffness, or computational issues here (see e.g. [16]).

However, we consider time discretizations of the abstract Cauchy problem

$$\partial_t u(t) + A[u(t)] = 0, \quad t > 0, \quad u(0) = u^0, \tag{2.1}$$

where $A : D(A) \to X'$ is a (differential) operator defined on $D(A) \subset X$ and X is a Banach space with dual X'. In the work presented in this thesis, we restrict ourselves to diffusion operators $A[u]$ defined on some Sobolev space with solutions $u : \Omega \times (0, \infty) \to \mathbb{R}^n$, which may be vector-valued. A typical example is $A[u] = \operatorname{div}(a(u)\nabla u)$ defined on $X = L^2(\Omega)$ with domain $D(A) = H^2(\Omega)$, where $a : \mathbb{R} \to \mathbb{R}$ is a smooth function (see section 2.3). Equation (2.1) often possesses a Lyapunov functional $H[u] = \int_\Omega h(u)dx$ (called *entropy* in the following), where $h : \mathbb{R}^n \to \mathbb{R}$, such that

$$\frac{dH}{dt}[u] = \int_\Omega h'(u)\partial_t u dx = -\int_\Omega h'(u)A[u]dx \leq 0,$$

at least when the *entropy production* $\int_\Omega h'(u)A[u]dx$ is nonnegative. Here, h' is the derivative of h and $h'(u)A[u]$ is interpreted as the inner product of $h'(u)$ and $A[u]$ in $\mathbb{R}^n$. Furthermore, if h is convex, the convex Sobolev inequality $\int_\Omega h'(u)A[u]dx \geq \kappa H[u]$ for some $\kappa > 0$ may hold [27], which implies that $dH/dt \leq -\kappa H$ and hence exponential convergence of $H[u]$ to zero with rate κ. The aim is to design a higher-order time-discrete scheme which preserves this entropy-dissipation property.

To this end, we propose the following semi-discrete Runge-Kutta approximation of (2.1): Given $u^{k-1} \in X$, define

$$u^k = u^{k-1} + \tau \sum_{i=1}^{s} b_i K_i, \quad K_i = -A\Big[u^{k-1} + \tau \sum_{j=1}^{s} a_{ij} K_j\Big], \quad i = 1, \ldots, s, \tag{2.2}$$

where t^k are the time steps, $\tau = t^k - t^{k-1} > 0$ is the uniform time step size, u^k approximates $u(t^k)$, and $s \geq 1$ denotes the number of Runge-Kutta stages.

Since the Cauchy problem is autonomous, the knots $c_1, \dots, c_s$ are not needed here. In concrete examples (see below), u^k are functions from Ω to $\mathbb{R}^n$. If $a_{ij} = 0$ for $j \geq i$, the Runge-Kutta scheme is explicit, otherwise it is implicit and a nonlinear system of size s has to be solved to compute K_i. We assume that scheme (2.2) is solvable for $u^k : \Omega \to \mathbb{R}^n$.
Given $h : \mathbb{R}^n \to \mathbb{R}$, we wish to determine conditions under which the functional

$$H[u^k] = \int_\Omega h(u^k(x))dx \tag{2.3}$$

is dissipated by the numerical scheme (2.2),

$$H[u^k] + \tau \int_\Omega A[u^k]h'(u^k)dx \leq H[u^{k-1}], \quad k \in \mathbb{N}. \tag{2.4}$$

In many examples (see below), $\int_\Omega A[u^k]h'(u^k)dx \geq 0$ and thus, the function $k \mapsto H[u^k]$ is decreasing. Such a property is the first step in proving the preservation of the large-time asymptotics of the numerical scheme (see Remark 19). Our main results in this chapter, stated on an informal level, are as follows:

(i) We determine an abstract condition under which the discrete entropy-dissipation inequality (2.4) holds for sufficiently small $\tau > 0$. This condition is made explicit for special choices of A and h, yielding entropy-dissipative implicit or explicit Runge-Kutta schemes of any order.

(ii) Numerical experiments for the porous-medium equation indicate that τ may be chosen independent of the time step k, thus yielding discrete entropy dissipation for all discrete times.

(iii) We show that for Runge-Kutta schemes of order $p \geq 2$, the abstract condition in (i) is exactly the criterion of Liero and Mielke [71] to conclude geodesic 0-convexity of the entropy. In particular, it is related to the Bakry-Emery condition [5].

In the following, the main results of chapter 2 are described in more detail. We recall that the Runge-Kutta scheme (2.2) is consistent if $\sum_{j=1}^s a_{ij} = c_i$ and $\sum_{i=1}^s b_i = 1$. Furthermore, if $\sum_{i=1}^s b_i c_i = \frac{1}{2}$, it is at least of order two [55, Chap. II]. We introduce the number

$$C_{\mathrm{RK}} = 2\sum_{i=1}^s b_i(1 - c_i), \tag{2.5}$$

which takes only three values:

$$C_{\mathrm{RK}} = 0 \quad \text{for the implicit Euler scheme,}$$

$$
\begin{aligned}
C_{\mathrm{RK}} &= 1 \quad \text{for any Runge-Kutta scheme of order } p \geq 2,\\
C_{\mathrm{RK}} &= 2 \quad \text{for the explicit Euler scheme.}
\end{aligned}
$$

The *first main result* is an abstract entropy-dissipation property of scheme (2.2) for entropies of type (2.3).

Theorem 18 (Entropy-dissipation structure I)**.** *Let $h \in C^2(\mathbb{R}^n)$, let $A : D(A) \to X'$ be Fréchet differentiable with Fréchet derivative $DA[u] : X \to X'$ at $u \in D(A)$, and let (u^k) be the Runge-Kutta solution to* (2.2)*. Suppose that*

$$
I_0^k := \int_\Omega \big(C_{\mathrm{RK}} h'(u^k) DA[u^k](A[u^k]) + h''(u^k)(A[u^k])^2\big)dx > 0. \tag{2.6}
$$

Then there exists $\tau^k > 0$ such that for all $0 < \tau \leq \tau^k$,

$$
H[u^k] + \tau \int_\Omega A[u^k]h'(u^k)dx \leq H[u^{k-1}]. \tag{2.7}
$$

Compared to strong stability preserving Runge-Kutta schemes [57, 50], we obtain not only a time-discrete dissipation property, but also an estimate for $A[u^k]h'(u^k)$, which usually provides gradient bounds. Another difference is that we study semi-discrete problems, while the references [57, 50] the authors are concerned with ordinary differential equations derived from method-of-lines approximations.
We assume that the solutions to (2.2) are sufficiently regular such that the integral (2.6) can be defined. In the vector-valued case, $h''(u^k)$ is the Hessian matrix and we interpret $h''(u^k)(A[u^k])^2$ as the product $A[u^k]^\top h''(u^k)A[u^k]$. For Runge-Kutta schemes of order $p \geq 2$ (for which $C_{\mathrm{RK}} = 1$), the integral (2.6) corresponds exactly to the second-order time derivative of $H[u(t)]$ for solutions $u(t)$ to the *continuous* equation (2.1). Note that the entropy-dissipation estimate (2.7) is only *local*, since the time step restriction depends on the time step k. For implicit Euler schemes (and convex entropies h), it is known that τ^k can be chosen independent of k. For general Runge-Kutta methods, we cannot prove rigorously that τ^k stays bounded from below as $k \to \infty$. However, our numerical experiments in section 2.7 indicate that inequality (2.7) holds for sufficiently small $\tau > 0$ uniformly in k.

Remark 19 (Exponential decay of the discrete entropy)**.** If the convex Sobolev inequality $\int_\Omega A[u^k]h'(u^k)dx \geq \kappa H[u^k]$ holds for some constant $\kappa > 0$ and if there exists $\tau^* > 0$ such that $\tau^k \geq \tau^* > 0$ for all $k \in \mathbb{N}$, we infer from (2.7) that for $\tau := \tau^*$,

$$
H[u^k] \leq (1+\kappa\tau)^{-k}H[u^0] = \exp(-\eta\kappa t^k)H[u^0], \quad \text{where } \eta = \frac{\log(1+\kappa\tau)}{\kappa\tau} < 1,
$$

which implies exponential decay of the discrete entropy with rate $\eta\kappa$. This rate converges to the continuous rate κ as $\tau \to 0$ and therefore it is asymptotically sharp. $\square$

Theorem 18 can be generalized to a larger class of entropies, namely to so-called *first-order entropies*

$$F[u^k] = \int_\Omega |\nabla f(u^k)|^2 dx, \tag{2.8}$$

where, for simplicity, we consider only the scalar case with $f : \mathbb{R} \to \mathbb{R}$. An important example is the Fisher information with $f(u) = \sqrt{u}$.

Theorem 20 (Entropy-dissipating structure II)**.** *Let $f \in C^2(\mathbb{R})$, let $A : D(A) \to X'$ be Fréchet differentiable, and let (u^k) be the Runge-Kutta solution to (2.2). Assume that the boundary condition $\nabla f(u^k) \cdot \nu = 0$ on $\partial\Omega$ is satisfied. Furthermore, suppose that*

$$\begin{aligned} I_1^k := \int_\Omega \Big(&|\nabla(f'(u^k)A[u^k]|^2 - C_{\mathrm{RK}} \Delta f(u^k) f'(u^k) DA[u^k](A[u^k]) \\ &- \Delta f(u^k) f''(u^k)(A[u^k])^2 \Big) dx > 0. \end{aligned} \tag{2.9}$$

Then there exists $\tau^k > 0$ such that for all $0 < \tau \leq \tau^k$,

$$F[u^k] + \tau \int_\Omega A[u^k] f'(u^k) dx \leq F[u^{k-1}].$$

The key idea of the proof of Theorem 18 (and similarly for the proof of Theorem 20) is a concavity property of the difference of the entropies at two consecutive time steps with respect to the time step size τ. To explain this idea, let $u := u^k$ be fixed and introduce $v(\tau) := u^{k-1}$. Clearly, $v(0) = u$. A formal Taylor expansion of $G(\tau) := H[u] - H[v(\tau)]$ yields

$$H[u^k] - H[u^{k-1}] = G(\tau) = G(0) + \tau G'(0) + \frac{\tau^2}{2} G''(\xi^k),$$

where $0 < \xi^k < \tau$. A computation, presented explicit in section 2.2, shows that $G'(0) = \int_\Omega A[u^k] h'(u^k) dx$ and $G''(0) = -I_0^k$. Now, if $G''(0) < 0$, there exists $\tau^k > 0$ such that $G''(\tau) \leq 0$ for $\tau \in [0, \tau^k]$ and in particular $G''(\xi^k) \leq 0$. Consequently, $G(\tau) \leq \tau G'(0)$, which equals (2.4). The definition of $v(\tau)$ assumes implicitly that (2.2) is *backward* solvable. We prove in Proposition 22 below that this property holds if the operator A is a smooth self-mapping on X.

Remark 21 (Discussion of τ^k)**.** Since (u^k) is expected to converge to the stationary solution, $\lim_{k\to\infty} I_0^k = 0$. Thus, in principle, for larger values of k, we expect that τ^k becomes smaller and smaller, thus restricting the choice of time step sizes τ. However, practically, the situation more favorable. For instance, for the implicit Euler scheme, if h is convex, we obtain

$$H[u^k] - H[u^{k-1}] \le \int_\Omega h'(u^k)(u^k - u^{k-1})dx = -\tau \int_\Omega h'(u^k)A[u^k]dx$$

for *any* value of $\tau > 0$. Moreover, for other (higher-order) Runge-Kutta schemes, the numerical experiments in section 2.7 indicate that there exists $\tau^* > 0$ such that $G''(\tau) \le 0$ holds for all $\tau \in [0, \tau^*]$ uniformly in $k \in \mathbb{N}$. In this situation, inequality (2.7) holds for all $0 < \tau \le \tau^*$, and thus our estimate is global. In fact, numerically, the function G'' is even nonincreasing in some interval $[0, \tau^*]$ but we are not able to prove this analytically. $\square$

The *second main result* is the specification of the abstract conditions (2.6) and (2.9) for a number of examples: a quasilinear diffusion equation, porous-medium or fast-diffusion equations, a linear diffusion system, and the fourth-order Derrida-Lebowitz-Speer-Spohn equation (see sections 2.3-2.6 for details). For instance, for the porous-medium equation

$$\partial_t u = \Delta(u^\beta) \text{ in } \Omega,\ t > 0, \quad \nabla u^\beta \cdot \nu = 0 \text{ on } \partial\Omega, \quad u(0) = u^0,$$

we show that the Runge-Kutta scheme satisfies

$$H[u^k] + \tau\beta \int_\Omega (u^k)^{\alpha+\beta-2} |\nabla u^k|^2 dx \le H[u^{k-1}], \text{ where } H[u] = \frac{1}{\alpha(\alpha+1)} \int_\Omega u^{\alpha+1} dx,$$

for $0 < \tau \le \tau^k$ and all (α, β) belonging to some region in $[0, \infty)^2$ (see Figure 2.1 below). For $\alpha = 0$, we write $H[u] = \int_\Omega u(\log u - 1)dx$. In one space dimension and for Runge-Kutta schemes of order $p \ge 2$, this region becomes $-2 < \alpha - \beta < 1$, which is the same condition as for the continuous equation (except the boundary values). Furthermore, the first-order entropy (2.8) is dissipated for Runge-Kutta schemes of order $p \ge 2$, in one space dimension,

$$F[u^k] + \tau C_{\alpha,\beta} \int_\Omega (u^k)^{\alpha+\beta-2} (u^k)_{xx}^2 dx \le F[u^{k-1}], \quad \text{where } F[u] = \int_\Omega (u^{\alpha/2})_x^2 dx,$$

for $0 < \tau \le \tau^k$ and all (α, β) belonging to the region shown in Figure 2.2 below, and $C_{\alpha,\beta} > 0$ is some constant. This region is smaller than the region of admissible values (α, β) for the continuous entropy. The borders of that region are indicated in the Figure 2.2 by dashed lines.
The proof of the above results, specifically of $G''(0) < 0$, is based on systematic

integration by parts [58]. The idea of the method is to formulate integration by parts as manipulations with polynomials and to conclude the inequality $G''(0) < 0$ from a polynomial decision problem. This problem can be solved directly or by using computer algebra software.

Our *third main result* is the relation to geodesic 0-convexity of the entropy and the Bakry-Emery approach for the case $C_{\mathrm{RK}} = 1$ (Runge-Kutta scheme of order $p \geq 2$). Liero and Mielke [71] formulate the abstract Cauchy problem (2.1) as the gradient flow

$$\partial_t u = -K[u]DH[u], \quad t > 0, \quad u(0) = u^0,$$

where the Onsager operator $K[u]$ describes the sum of diffusion and reaction terms. For instance, if $A[u] = \operatorname{div}(a(u)\nabla u)$, we can rewrite the operator $A[u] = \operatorname{div}(a(u)h''(u)^{-1}\nabla h'(u))$ and thus, identifying $h'(u)$ and $DH[u]$, we have $K[u]\xi = \operatorname{div}(a(u)h''(u)^{-1}\nabla\xi)$. It is shown in [71] that the entropy H is geodesic λ-convex if the inequality

$$M(u,\xi) := \langle \xi, DA[u]K[u]\xi\rangle - \frac{1}{2}\langle \xi, DK[u]A[u]\xi\rangle \geq \lambda\langle \xi, K[u]\xi\rangle \tag{2.10}$$

holds for all suitable u and ξ. We prove in section 2.2 that

$$G''(0) = 2M(u^k, h'(u^k)).$$

Hence, if $G''(0) \leq 0$ then (2.10) with $\lambda = 0$ is satisfied for $u = u^k$ and $\xi = h'(u^k)$, yielding geodesic 0-convexity for the semi-discrete entropy. Moreover, if $G''(0) \leq -\lambda G'(0)$ we obtain geodesic λ-convexity. Since $G'(0) = -dH[u]/dt$ and $G''(0) = -d^2H[u]/dt^2$ in the continuous setting, the inequality $G''(0) \leq -\lambda G'(0)$ can be written as

$$\frac{d^2H}{dt^2}[u] \geq -\lambda\frac{dH}{dt}[u],$$

which corresponds to a variant of the Bakry-Emery condition [5], yielding exponential convergence of $H[u]$ (if $\tau^k \geq \tau^* > 0$ for all k). Thus, our results constitute a first step towards a *discrete Bakry-Emery approach*.

This chapter is organized as follows. The abstract method, i.e. the proof of backward solvability and of Theorems 18 and 20, is presented in section 2.2. The method is applied in the subsequent sections to a scalar diffusion equation (section 2.3), the porous-medium equation (section 2.4), a linear diffusion system (section 2.5), and the fourth-order Derrida-Lebowitz-Speer-Spohn equation (section 2.6). Finally, section 2.7 is devoted to some numerical experiments showing that G'' is negative in some interval $[0,\tau^*]$.

2.2 The abstract method

In this section, we show that the Runge-Kutta scheme is backward solvable if A is a self-mapping and we prove Theorems 18 and 20.

Proposition 22 (Backward solvability)**.** *Let $(\tau, u^k) \in [0,\infty) \times X$, where X is some Banach space, and let $A \in C^2(X,X)$ be a self-mapping. Then there exists $\tau_0 > 0$, a neighborhood $V \subset X$ of u^k, and a function $v \in C^2([0,\tau_0); X)$ such that* (2.2) *holds for $u^{k-1} := v(\tau)$. Moreover,*

$$v(0) = 0, \quad v'(0) = A[u], \quad \textit{and} \quad v''(0) = C_{\mathrm{RK}} DA[u](A[u]). \tag{2.11}$$

The self-mapping assumption is strong for differential operators A but it is natural in the context of Runge-Kutta methods and valid for smooth solutions.

Proof. The idea of the proof is to apply the implicit function theorem in Banach spaces (see [31, Corollary 15.1]). To this end, we set $u := u^k$ and define the mapping $J = (J_0, \dots, J_s) : \mathbb{R} \times X^{s+1} \to X^{s+1}$ by

$$\begin{aligned} J_0(\tau, y) &= v - u + \tau \sum_{i=1}^{s} b_i k_i, \quad \text{where } y = (k_1, \dots, k_s, v), \\ J_i(\tau, y) &= k_i + A\left[v + \tau \sum_{j=1}^{s} a_{ij} k_j\right], \quad i = 1, \dots, s. \end{aligned}$$

The Fréchet derivative of the mapping J in the direction of (τ_h, y_h), where $y_h = (k_{h1}, \dots, k_{hs}, v_h)$, reads as

$$\begin{aligned} DJ_0(\tau, y)(\tau_h, y_h) &= v_h + \tau_h \sum_{i=1}^{s} b_i k_i + \tau \sum_{i=1}^{s} b_i k_{hi}, \\ DJ_i(\tau, y)(\tau_h, y_h) &= k_{hi} + DA\left[v + \tau \sum_{j=1}^{s} a_{ij} k_j\right]\left(v_h + \tau_h \sum_{j=1}^{s} a_{ij} k_j + \tau \sum_{j=1}^{s} a_{ij} k_{hj}\right), \end{aligned}$$

where $i = 1, \dots, s$. Let $\tau_0 = 0$ and $y_0 = (-A[u], \dots, -A[u], u)$. Then $J(\tau_0, y_0) = 0$ and

$$DJ_0(\tau_0, y_0)(0, y_h) = v_h, \quad DJ_i(\tau_0, y_0)(0, y_h) = k_{hi} + DA[u](v_h), \quad i = 1, \dots, s.$$

The mapping $y_h \mapsto DJ(\tau_0, y_0)(0, y_h)$ is clearly an isomorphism from X^{s+1} onto X^{s+1}. By the implicit function theorem, there exist an interval $U \subset [0, \tau_0)$, a neighborhood $V \subset X^{s+1}$ of y_0, and a function $(k, v) \in C^2([0, \tau_0); V)$ such that

$(k,v)(0) = (-A[u],\dots,-A[u],u)$ and $J(\tau,k(\tau),v(\tau)) = 0$ for all $\tau \in [0,\tau_0)$. Implicit differentiation of $J(\tau,k(\tau),v(\tau)) = 0$ yields

$$
\begin{aligned}
0 &= v'(\tau) + \sum_{i=1}^{s} b_i k_i(\tau) + \tau \sum_{i=1}^{s} b_i k_i'(\tau), \\
0 &= k_i'(\tau) + DA\Big[v + \tau \sum_{j=1}^{s} a_{ij} k_j(\tau)\Big]\Big(v'(\tau) + \sum_{j=1}^{s} a_{ij} k_j(\tau) + \tau \sum_{j=1}^{s} a_{ij} k_j'(\tau)\Big),
\end{aligned}
$$

where $i = 1,\dots,s$ and $\tau \in [0,\tau_0)$. Using $\sum_{i=1}^{s} b_i = 1$ and $\sum_{j=1}^{s} a_{ij} = c_i$, we infer that

$$
\begin{aligned}
v'(0) &= -\sum_{i=1}^{s} b_i k_i(0) = \sum_{i=1}^{s} b_i A[u] = A[u], \\
k_i'(0) &= -DA[u]\Big(A[u] - \sum_{j=1}^{s} a_{ij} A[u]\Big) = -(1-c_i) DA[u](A[u]).
\end{aligned} \tag{2.12}
$$

Differentiating $J_0(\tau,k(\tau),v(\tau)) = 0$ twice leads to

$$
0 = v''(\tau) + 2\sum_{i=1}^{s} b_i k_i'(\tau) + \tau \sum_{i=1}^{s} b_i k_i''(\tau).
$$

Because of (2.12), this reads at $\tau = 0$ as

$$
v''(0) = -2\sum_{i=1}^{s} b_i k_i'(0) = 2\sum_{i=1}^{s} b_i (1-c_i) DA[u](A[u]) = C_{\mathrm{RK}} DA[u](A[u]).
$$

This finishes the proof. $\square$

We now prove Theorems 18 and 20.

Proof of Theorem 18. We set $u := u^k$. Following Proposition 22, there exists a backward solution $v \in C^2([0,\tau_0))$ such that $v(0) = u$, $v'(0) = A[u]$, and $v''(0) = C_{\mathrm{RK}} DA[u](A[u])$. Furthermore, the function $G(\tau) = \int_\Omega (h(u) - h(v(\tau)))dx$ satisfies $G(0) = 0$,

$$
\begin{aligned}
G'(0) &= -\int_\Omega h'(v(0))v'(0)dx = -\int_\Omega h'(u)A[u]dx, \\
G''(0) &= -\int_\Omega \big(h'(v(0))v''(0) + h''(v(0))v'(0)^2\big)dx \\
&= -\int_\Omega \big(C_{\mathrm{RK}} h'(u) DA[u](A[u]) + h''(u)(A[u])^2\big)dx = -I_0^k < 0,
\end{aligned}
$$

using the assumption. By continuity, there exists $0 < \tau^k < \tau_0$ such that $G''(\xi) \le 0$ for $0 \le \xi \le \tau^k$. Then the Taylor expansion $G(\tau) = G(0) + G'(0)\tau + \frac{1}{2}G''(\xi)\tau^2 \le G'(0)\tau$ concludes the proof. $\square$

Proof of Theorem 20. Following the lines of the previous proof, it is sufficient to compute $G'(0)$ and $G''(0)$, where now $G(\tau) = \int_\Omega (|\nabla f(u)|^2 - |\nabla f(v(\tau))|^2)dx$. Using integration by parts and the boundary condition $\nabla f(v)\cdot\nu = 0$ on $\partial\Omega$, we compute

$$G'(0) = -\int_\Omega \nabla f(v(0))\cdot\nabla\big(f'(v(0))v'(0)\big)dx = \int_\Omega \Delta f(u)f'(v(\tau))A[u]dx,$$

since $v(0) = u$ and $v'(0) = A[u]$. Furthermore, again integrating by parts,

$$\begin{aligned} G''(\tau) &= -\int_\Omega \Big(\big|\nabla\big(f'(v(\tau))v'(\tau)\big)\big|^2 + \nabla f(v(\tau))\cdot\nabla\big(f''(v(\tau))(v'(\tau))^2\big) \\ &\quad + \nabla f(v(\tau))\cdot\nabla\big(f'(v(\tau))v''(\tau)\big)\Big)dx \\ &= -\int_\Omega \Big(\big|\nabla\big(f'(v(\tau))v'(\tau)\big)\big|^2 - \Delta f(v(\tau))f''(v(\tau))(v'(\tau))^2 \\ &\quad - \Delta f(v(\tau))f'(v(\tau))v''(\tau)\Big)dx. \end{aligned}$$

Since $v''(0) = C_{\mathrm{RK}}DA[u](A[u])$, this reduces at $\tau = 0$ to

$$\begin{aligned} G''(0) = -\int_\Omega \Big(&|\nabla(f'(u)A[u])|^2 - \Delta f(u)f''(u)(A[u])^2 \\ &- C_{\mathrm{RK}}\Delta f(u)f'(u)DA[u](A[u])\Big)dx. \end{aligned}$$

This expression equals $-I_1^k$, and the result follows. □

Finally, we show that $G''(0)$ for entropies (2.3) is related to the geodesic convexity condition of [71].

Lemma 23. *Let $A[u] = K(u)DH[u]$ for some symmetric operator $K : D(A)\to X$ and Fréchet derivative $DH[u]$, let G be defined as in the proof of Theorem 18 for a solution u^k to the Runge-Kutta scheme* (2.2) *of order $p\geq 2$, and let $M(u,\xi)$ be given by* (2.10)*. Then*

$$G''(0) = -2M(u^k, DH[u^k]).$$

Proof. The proof is just a (formal) calculation. Recall that for Runge-Kutta schemes of order $p\geq 2$, we have $C_{\mathrm{RK}} = 1$. Set $u := u^k$ and identify $DH[u]$ with $\xi = h'(u)$. Inserting the expression $DA[u](v) = DK[u](v)h'(u) + K[u]h''(u)v$ into the definition of $G''(0)$, we find that

$$\begin{aligned} -G''(0) &= \langle\xi, DA[u](A[u])\rangle + \langle A[u], h''(u)A[u]\rangle \\ &= \big\langle\xi, DK[u](A[u])\xi + K[u]h''(u)A[u]\big\rangle + \langle A[u], h''(u)A[u]\rangle \\ &= \langle\xi, DK[u](K[u]\xi)\xi\rangle + \langle\xi, K[u]h''(u)K[u]\xi\rangle + \langle K[u]\xi, h''(u)K[u]\xi\rangle \end{aligned}$$

$$= \langle \xi, DK[u](K[u]\xi)\xi\rangle + 2\langle \xi, K[u]h''(u)K[u]\xi\rangle,$$

since $K[u]$ is assumed to be symmetric. Rearranging the terms, we obtain

$$\begin{aligned} -G''(0) &= 2\langle \xi, DK[u](K[u]\xi)\xi\rangle + 2\langle \xi, K[u]h''(u)K[u]\xi\rangle - \langle \xi, DK[u](K[u]\xi)\rangle \\ &= 2\langle \xi, DA[u](K[u]\xi)\xi\rangle - \langle \xi, DK[u](A[u])\rangle = 2M(u,\xi), \end{aligned}$$

which proves the claim. □

2.3 Scalar diffusion equation

In this section, we analyze time-discrete Runge-Kutta schemes of the diffusion equation

$$\partial_t u = \operatorname{div}(a(u)\nabla u), \quad t > 0, \quad u(0) = u^0, \tag{2.13}$$

with periodic or homogeneous Neumann boundary conditions. This equation, also including a drift term, was analyzed in [71] in the context of geodesic convexity. Our results are similar to those in [71] but we consider the time-discrete and not the continuous equation and we employ systematic integration by parts [58].

Setting $\mu(u) = a(u)/h''(u)$, we can write the diffusion equation as a formal gradient flow:

$$\partial_t u = -A[u] := \operatorname{div}(\mu(u)\nabla h'(u)), \quad t > 0.$$

We prove that the Runge-Kutta scheme (2.2) dissipates all convex entropies subject to some conditions on the functions μ and h.

Theorem 24. *Let $\Omega \subset \mathbb{R}^d$ be convex with smooth boundary. Let (u^k) be a sequence of (smooth) solutions to the Runge-Kutta scheme* (2.2) *of the diffusion equation* (2.13)*. Let $k \in \mathbb{N}$ be fixed and let u^k not be equal to the constant steady state of* (2.13)*. We suppose that for all admissible u, it holds that $a(u) \geq 0$, $h''(u) \geq 0$,*

$$b(u) := \frac{2}{3}(C_{\mathrm{RK}} + 1)\int_{u_0}^{u} \mu(v)\mu'(v)h''(v)dv \geq 0, \tag{2.14}$$

$$\frac{d-1}{d}b(u) \leq (C_{\mathrm{RK}} + 1)h''(u)\mu(u)^2, \tag{2.15}$$

$$(C_{\mathrm{RK}} + 2)\mu(u)\mu''(u) + (C_{\mathrm{RK}} - 1)\mu'(u)^2 < 0. \tag{2.16}$$

Then there exists $\tau^k > 0$ such that for all $0 < \tau < \tau^k$,

$$H[u^k] + \tau \int_\Omega h''(u^k)a(u^k)|\nabla u^k|^2 dx \leq H[u^{k-1}].$$

Conditions (2.14)-(2.15) correspond to (4.12) in [71]. Condition (2.16) is satisfied for concave functions μ, except for the explicit Euler scheme ($C_{\rm RK} = 2$) for which we additionally need that $4\mu\mu'' + (\mu')^2 < 0$. For the implicit Euler scheme, we may even allow for nonconcave mobilities μ, e.g. $\mu(u) = u^\gamma$ for $1 < \gamma < 2$.

Proof. According to Theorem 18, we only need to show that $I_0^k = -G''(0) > 0$. To simplify, we set $u := u^k$. First, we observe that the boundary condition $\nabla u \cdot \nu = 0$ on Ω implies that $0 = \partial_t \nabla u \cdot \nu = \nabla \partial_t u \cdot \nu = -\nabla A[u] \cdot \nu$ on $\partial\Omega$. Using $DA[u](A[u]) = \operatorname{div}(a'(u)A[u]\nabla u + a(u)\nabla A[u]) = \Delta(a(u)A[u])$, the abbreviation $\xi = h'(u)$, and integration by parts, we compute

$$\begin{aligned}
G''(0) &= -\int_\Omega \Big(C_{\rm RK} h'(u)\Delta(a(u)A[u]) + h''(u)\big(\operatorname{div}(\mu(u)\nabla h'(u))\big)^2\Big) dx \\
&= \int_\Omega \Big(C_{\rm RK} \nabla h'(u)\cdot\nabla(a(u)A[u]) - h''(u)\big(\mu'(u)\nabla u\cdot\nabla h'(u) \\
&\qquad + \mu(u)\Delta h'(u)\big)^2\Big) dx \\
&= -\int_\Omega \left(C_{\rm RK}\Delta\xi a(u)A[u] + h''(u)\left(\frac{\mu'(u)}{h''(u)}|\nabla\xi|^2 + \mu(u)\Delta\xi\right)^2\right) dx.
\end{aligned}$$

The boundary integrals vanish since $\nabla u \cdot \nu = \nabla A[u]\cdot\nu = 0$ on $\partial\Omega$. Replacing $A[u]$ by $\operatorname{div}(\mu(u)\nabla\xi) = \mu(u)\Delta\xi + \mu'(u)|\nabla\xi|^2/h''(u)$ and expanding the square, we arrive at

$$\begin{aligned}
G''(0) &= -\int_\Omega \Bigg(\big(C_{\rm RK} a(u)\mu(u) + h''(u)\mu(u)^2\big)(\Delta\xi)^2 \\
&\qquad + \left(C_{\rm RK} a(u)\frac{\mu'(u)}{h''(u)} + 2\mu(u)\mu'(u)\right)\Delta\xi|\nabla\xi|^2 + \frac{\mu'(u)^2}{h''(u)}|\nabla\xi|^4\Bigg) dx \qquad (2.17) \\
&= -\int_\Omega \Big((C_{\rm RK}+1)h''(u)\mu(u)^2\xi_L^2 + (C_{\rm RK}+2)\mu(u)\mu'(u)\xi_L\xi_G^2 \qquad (2.18) \\
&\qquad + \mu'(u)^2 h''(u)^{-1}\xi_G^4\Big) dx,
\end{aligned}$$

where we have employed the identity $a(u) = \mu(u)h''(u)$ and the abbreviations $\xi_G = |\nabla\xi|$ and $\xi_L = \Delta\xi$; see Table 2.3 for an overview of the various abbreviations. We now apply the method of systematic integration by parts [58]. The idea is to identify useful integration-by-parts formulas and to add them to $G''(0)$ without changing the sign of $G''(0)$. The first formula is given by

$$\int_\Omega \operatorname{div}\big(\Gamma_1(u)(\nabla^2\xi - \Delta\xi\mathbb{I})\cdot\nabla\xi\big) dx = \int_{\partial\Omega} \Gamma_1(u)\nabla\xi^\top(\nabla^2\xi - \Delta\xi\mathbb{I})\nu ds, \quad (2.19)$$

Abbrev.	Definition	Abbrev.	Definition
ξ	$h'(u)$		
ξ_L	$\Delta\xi$	ξ_G	$\lvert\nabla\xi\rvert$
ξ_H	$\lvert\nabla^2\xi\rvert$	ξ_{GHG}	$\nabla\xi^\top\nabla^2\xi\nabla\xi$
ξ_S	$(d-1)^{-1}\xi_G^{-2}(\xi_{GHG}-\xi_L\xi_G^2/d)$	ξ_R^2	$\xi_H^2-\xi_L^2/d-d(d-1)\xi_S^2$

Table 2.1: Overview of the abbreviations for the proof of Theorem 24.

where $\Gamma_1(u)\le 0$ is an arbitrary (smooth) scalar function which still needs to be chosen, and $\mathbb{I}$ is the unit matrix in $\mathbb{R}^{d\times d}$. Computing the divergence and using the property $\nabla u=\nabla\xi/h''(u)$, the left-hand side can be expanded as

$$\begin{aligned}
&\int_\Omega\Big(\Gamma_1'(u)\nabla u^\top(\nabla^2\xi-\Delta\xi\mathbb{I})\nabla\xi+\Gamma_1(u)(\nabla^2\xi-\Delta\xi\mathbb{I}):\nabla^2\xi\Big)dx\\
&\quad=\int_\Omega\left(\frac{\Gamma_1'(u)}{h''(u)}\nabla\xi^\top\nabla^2\xi\nabla\xi-\frac{\Gamma_1'(u)}{h''(u)}\Delta\xi|\nabla\xi|^2+\Gamma_1(u)|\nabla^2\xi|^2-\Gamma_1(u)(\Delta\xi)^2\right)dx\\
&\quad=\int_\Omega\left(\frac{\Gamma_1(u)}{h''(u)}\xi_{GHG}-\frac{\Gamma_1'(u)}{h''(u)}\xi_L\xi_G^2+\Gamma_1(u)\xi_H^2-\Gamma_1(u)\xi_L^2\right)dx,
\end{aligned}$$

where we have set $\xi_{GHG}=\nabla\xi^\top\nabla^2\xi\nabla\xi$ and $\xi_H=|\nabla^2\xi|$. The boundary integral in (2.19) becomes

$$\int_{\partial\Omega}\Gamma_1(u)\left(\frac{1}{2}\nabla(|\nabla\xi|^2)-\Delta\xi\nabla\xi\right)\cdot\nu ds=\frac{1}{2}\int_{\partial\Omega}\Gamma_1(u)\nabla(|\nabla\xi|^2)\cdot\nu ds\ge 0,$$

since $\Gamma_1(u)\le 0$, $\nabla\xi\cdot\nu=0$ on $\partial\Omega$, and it holds that $\nabla(|\nabla\xi|^2)\cdot\nu\le 0$ on $\partial\Omega$ for all smooth functions satisfying $\nabla\xi\cdot\nu=0$ on $\partial\Omega$ [71, Prop. 4.2]. Here we need the convexity of Ω. Thus, the first integration-by-parts formula becomes

$$J_1:=\int_\Omega\left(\frac{\Gamma_1'(u)}{h''(u)}\xi_{GHG}-\frac{\Gamma_1'(u)}{h''(u)}\xi_L\xi_G^2+\Gamma_1(u)\xi_H^2-\Gamma_1(u)\xi_L^2\right)dx\ge 0.\tag{2.20}$$

The second formula reads

$$\begin{aligned}
0&=\int_\Omega\operatorname{div}\big(\Gamma_2(u)|\nabla\xi|^2\nabla\xi\big)dx\qquad(2.21)\\
&=\int_\Omega\left(\frac{\Gamma_2'(u)}{h''(u)}\xi_G^4+2\Gamma_2(u)\xi_{GHG}+\Gamma_2(u)\xi_L\xi_G^2\right)dx=:J_2,
\end{aligned}$$

where Γ_2 is an arbitrary scalar function. The goal is to find functions $\Gamma_1(u)\le 0$ and $\Gamma_2(u)$ such that $G''(0)\le G''(0)+J_1+J_2<0$.

According to [59], the computations simplify if we introduce the variables ξ_R and ξ_S satisfying

$$(d-1)\xi_G^2\xi_S=\xi_{GHG}-\frac{1}{d}\xi_L\xi_G^2,\quad \xi_H^2=\frac{1}{d}\xi_L^2+d(d-1)\xi_S^2+\xi_R^2.$$

The existence of ξ_R follows from the inequality

$$\xi_H^2 = |\nabla^2\xi|^2 \geq \frac{1}{d}(\Delta\xi)^2 + \frac{d}{d-1}\left(\frac{\nabla\xi^\top\nabla^2\xi\nabla\xi}{\nabla\xi^2} - \frac{\Delta\xi}{d}\right)^2 = \frac{1}{d}\xi_L^2 + d(d-1)\xi_S^2,$$

which is proven in [59, Lemma 2.1]. Then

$$G''(0) \leq G''(0)+J_1+J_2 = -\int_\Omega \left(a_1\xi_L^2+a_2\xi_L\xi_G^2+a_3\xi_G^4+a_4\xi_S\xi_G^2+a_5\xi_R^2+a_6\xi_S^2\right)dx, \tag{2.22}$$

where

$$\begin{aligned}
a_1 &= (C_{\mathrm{RK}}+1)h''(u)\mu(u)^2 + \left(1-\frac{1}{d}\right)\Gamma_1(u), \\
a_2 &= (C_{\mathrm{RK}}+2)\mu(u)\mu'(u) + \left(1-\frac{1}{d}\right)\frac{\Gamma_1'(u)}{h''(u)} - \left(\frac{2}{d}+1\right)\Gamma_2(u), \\
a_3 &= \frac{\mu'(u)^2-\Gamma_2'(u)}{h''(u)}, \quad a_4 = -(d-1)\left(\frac{\Gamma_1'(u)}{h''(u)}+2\Gamma_2(u)\right), \\
a_5 &= -\Gamma_1(u), \quad a_6 = -d(d-1)\Gamma_1(u).
\end{aligned} \tag{2.23}$$

The aim now is to determine conditions on $a_1, \ldots, a_6$ such that the polynomial $P(\xi) = a_1\xi_L^2 + a_2\xi_L\xi_G^2 + a_3\xi_G^4 + a_4\xi_S\xi_G^2 + a_5\xi_R^2 + a_6\xi_S^2$ is nonnegative, as this implies that $G''(0) \leq 0$. In the general case, this leads to nonlinear ordinary differential equations for Γ_1 and Γ_2 which cannot be easily solved. A possible approach is to require that the coefficients of the mixed terms vanish, i.e. $a_2 = a_4 = 0$, and that the remaining coefficients are nonnegative. The case $d = 1$ being simpler than the general case (since J_1 is not necessary), we assume that $d > 1$. Then $a_4 = 0$ implies that $\Gamma_1'(u)/h''(u) = -2\Gamma_2(u)$. Replacing $\Gamma_1'(u)/h''(u)$ by $-2\Gamma_2(u)$ in $a_2 = 0$ gives

$$\Gamma_2(u) = \frac{C_{\mathrm{RK}}+2}{3}\mu(u)\mu'(u).$$

On the other hand, replacing $\Gamma_2(u)$ by $-\Gamma_1'(u)/(2h''(u))$ in $a_2 = 0$, we find that

$$\Gamma_1'(u) = -\frac{2}{3}(C_{\mathrm{RK}}+2)\mu(u)\mu'(u)h''(u)$$

or, after integration,

$$\Gamma_1(u) = -\frac{2}{3}(C_{\mathrm{RK}}+2)\int_{u_0}^u \mu(v)\mu'(v)h''(v)dv.$$

These functions have to satisfy the conditions

$$a_1 \geq 0 \quad \text{or} \quad \frac{d-1}{d}\Gamma_1(u) \geq -(C_{\mathrm{RK}}+1)h''(u)\mu(u)^2,$$

$$a_3 \geq 0 \quad \text{or} \quad (C_{\mathrm{RK}}+2)\mu(u)\mu''(u) + (C_{\mathrm{RK}}-1)\mu'(u)^2 \leq 0,$$
$$a_5 \geq 0 \quad \text{or} \quad \Gamma_1(u) \leq 0 \quad \text{for all } u,$$

Note that $a_1 \geq 0$ and $a_5 \geq 0$ correspond to (2.15) and (2.14), respectively. This shows that $P(\xi) \geq 0$ for all $\xi \in \mathbb{R}^4$ and $G''(0) \leq 0$.

If $G''(0) = 0$, the nonnegative polynomial P, which depends on $x \in \Omega$ via ξ, has to vanish. In particular, $a_3\xi_G^4 = a_3|\nabla u|^4 = 0$ in Ω. As $a_3 > 0$ by assumption, $u(x) = \text{const.}$ for $x \in \Omega$. This contradicts the hypothesis that u is not a steady state. Consequently, $G''(0) < 0$, and we finish the proof by setting $b(u) = -\Gamma_1(u)$. □

2.4 Porous-medium equation

The results of the previous section can be applied in principle to the Runge-Kutta scheme for the porous-medium or fast-diffusion equation

$$\partial_t u = \Delta(u^\beta) \quad \text{in } \Omega,\ t > 0, \quad \nabla u^\beta \cdot \nu = 0 \quad \text{on } \partial\Omega, \quad u(0) = u^0, \qquad (2.24)$$

where $\beta > 0$. It can be seen that conditions (2.14)-(2.16) are not optimal for particular entropies. This is not surprising since we have neglected the mixed terms in the polynomial in (2.22) (i.e. $a_2 = a_4 = 0$) which is not optimal. In this section, we apply a different approach by making an ansatz for the functions Γ_1 and Γ_2, considering both zeroth-order and first-order entropies.

2.4.1 Zeroth-order entropies

We prove the following result.

Theorem 25. *Let $\Omega \subset \mathbb{R}^d$ be convex with smooth boundary. Let (u^k) be a sequence of (smooth) solutions to the Runge-Kutta scheme (2.2) for (2.24). Let the entropy be given by $H[u] = \alpha^{-1}(\alpha+1)^{-1}\int_\Omega u^{\alpha+1}dx$ with $\alpha > 0$, let $k \in \mathbb{N}$, and let u^k not be the constant steady state of (2.24). There exists a nonempty region $R_0(d) \subset (0,\infty)^2$ and $\tau^k > 0$ such that for all $(\alpha,\beta) \in R_0(d)$ and $0 < \tau \leq \tau^k$,*

$$H[u^k] + \tau\beta\int_\Omega (u^k)^{\alpha+\beta-2}|\nabla u^k|^2 dx \leq H[u^{k-1}], \quad k \in \mathbb{N}.$$

In one space dimension, we have

implicit Euler:	$R_0(1) = (0,\infty)^2$,
Runge-Kutta of order $p \geq 2$:	$R_0(1) = \{(\alpha,\beta) \in (0,\infty)^2 : -2 < \alpha - \beta < 1\}$,
explicit Euler:	$R_0(1) = \{(\alpha,\beta) \in (0,\infty)^2 : -1 < \alpha - \beta < 1\}$.

For the implicit Euler scheme, the theorem shows that any positive values for (α, β) are admissible, which corresponds to the continuous situation. For the Runge-Kutta case with $C_{\mathrm{RK}} = 1$, our condition is more restrictive. As expected, the explicit Euler scheme requires the most restrictive condition. The set $R_0(d)$ is illustrated in Figure 2.1 for $d = 2$ and $d = 10$.

Proof. Since $k \in \mathbb{N}$ is fixed, we set $u := u^k$. We choose the functions

$$\Gamma_1(u) = c_1\beta^2 u^{2\beta-\alpha-1}, \quad \Gamma_2(u) = c_2\beta^2 u^{2\beta-2\alpha-1}.$$

It holds $h''(u) = u^{\alpha-1}$ and $\mu(u) = \beta u^{\beta-\alpha}$. Then the coefficients in (2.23) are as follows:

$$\begin{aligned}
a_1 &= \beta^2\big((C_{\mathrm{RK}}+1) + (1-\tfrac{1}{d})c_1\big)u^{2\beta-\alpha-1},\\
a_2 &= \beta^2\big((C_{\mathrm{RK}}+2)(\beta-\alpha) + (1-\tfrac{1}{d})(2\beta-\alpha-1)c_1 - (\tfrac{2}{d}+1)c_2\big)u^{2\beta-2\alpha-1},\\
a_3 &= \beta^2\big((\beta-\alpha)^2 - (2\beta-2\alpha-1)c_2\big)u^{2\beta-3\alpha-2},\\
a_4 &= -\beta^2(d-1)\big((2\beta-\alpha-1)c_1 + 2c_2\big)u^{2\beta-2\alpha-1},\\
a_5 &= -\beta^2 c_1 u^{2\beta-\alpha-1}, \quad a_6 = -\beta^2 d(d-1)c_1 u^{2\beta-\alpha-1}.
\end{aligned}$$

Introducing the variables $\eta_j = \xi_j/u^\alpha$ for $j \in \{G, L, R, S\}$, we can write (2.22) as

$$\begin{aligned}
&G''(0) \le G''(0) + J_1 + J_2 = -\beta^2 \int_\Omega u^{2\beta+\alpha-1}Q(\eta)dx,\\
&\text{where } Q(\eta) = b_1\eta_L^2 + b_2\eta_L\eta_G^2 + b_3\eta_G^4 + b_4\eta_S\eta_G^2 + b_5\eta_R^2 + b_6\eta_S^2
\end{aligned}$$

with coefficients

$$\begin{aligned}
b_1 &= (C_{\mathrm{RK}}+1) + (1-\tfrac{1}{d})c_1,\\
b_2 &= (C_{\mathrm{RK}}+2)(\beta-\alpha) + (1-\tfrac{1}{d})(2\beta-\alpha-1)c_1 - (\tfrac{2}{d}+1)c_2,\\
b_3 &= (\beta-\alpha)^2 - (2\beta-2\alpha-1)c_2,\\
b_4 &= -(d-1)\big((2\beta-\alpha-1)c_1 + 2c_2\big),\\
b_5 &= -c_1, \quad b_6 = -d(d-1)c_1.
\end{aligned}$$

We need to determine all (α, β) such that there exist $c_1 \le 0$, $c_2 \in \mathbb{R}$ such that $Q(\eta) \ge 0$ for all $\eta = (\eta_G, \eta_L, \eta_R, \eta_S)$. Without loss of generality, we exclude the cases $b_1 = b_2 = 0$ and $b_4 = b_6 = 0$ since they lead to parameters (α, β) included in the region calculated below. Thus, let $b_1 > 0$ and $b_6 > 0$. These inequalities give the bound $-(C_{\mathrm{RK}}+1)/(1-1/d) < c_1 < 0$. Thus, we may introduce the parameter $\lambda \in (0,1)$ by setting $c_1 = -\lambda(C_{\mathrm{RK}}+1)/(1-1/d)$.

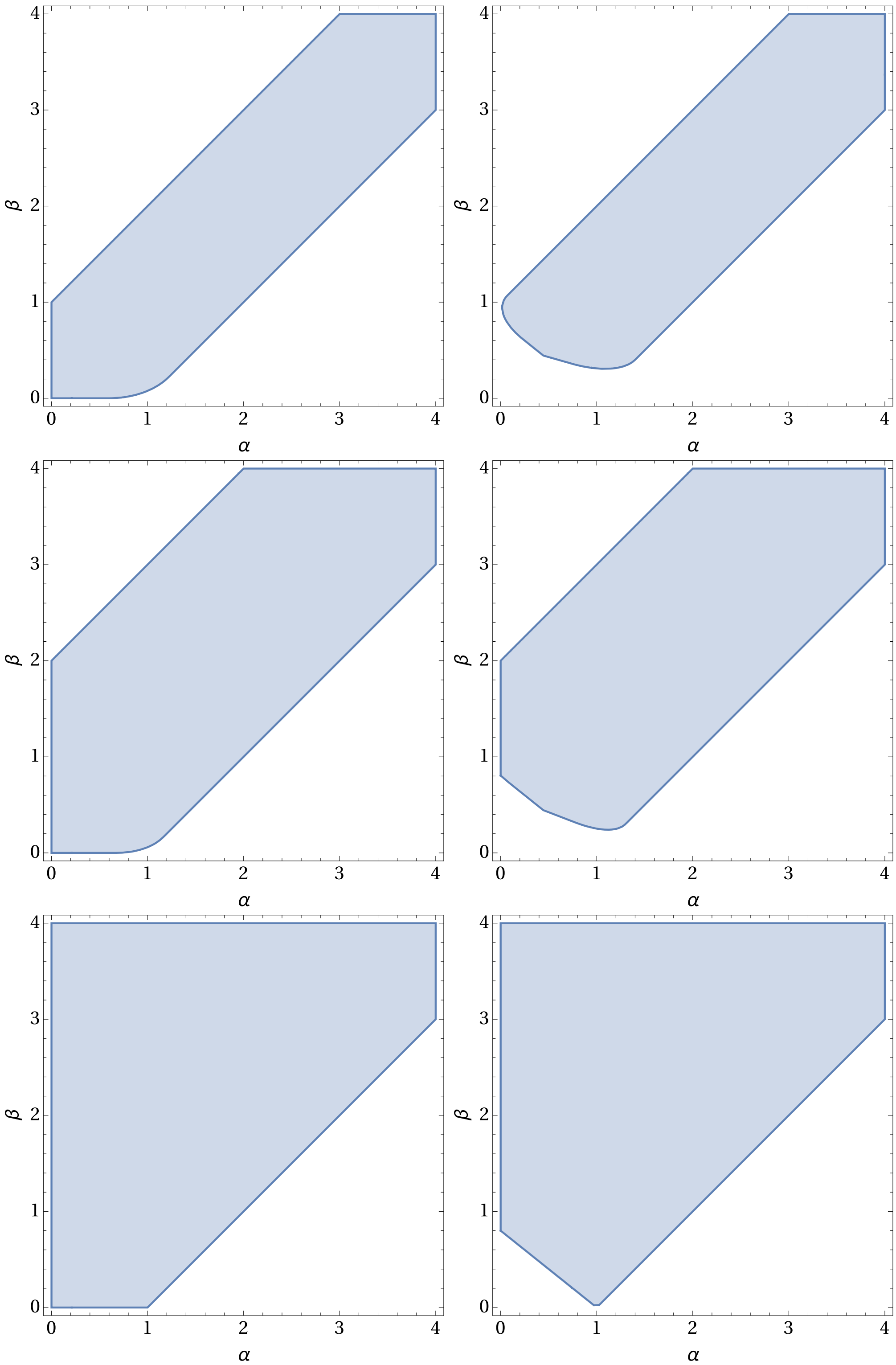

Figure 2.1: Set $R_0(d)$ of all (α, β) for which the zeroth-order entropy is dissipating. Left column: $d = 2$, right column: $d = 10$. Top row: explicit Euler scheme with $C_{\mathrm{RK}} = 2$, middle row: implicit Euler scheme with $C_{\mathrm{RK}} = 1$, bottom row: Runge-Kutta scheme of order $p \geq 2$ with $C_{\mathrm{RK}} = 0$.

The polynomial $Q(\eta)$ can be rewritten as

$$\begin{aligned}Q(\eta) &= b_1\left(\eta_L + \frac{b_2}{2b_1}\eta_G^2\right)^2 + b_6\left(\eta_S + \frac{b_4}{2b_6}\eta_G^2\right)^2 + b_5\eta_R^2 + \eta_G^4\left(b_3 - \frac{b_2^2}{4b_1} - \frac{b_4^2}{4b_6}\right)\\ &\geq \eta_G^4\left(b_3 - \frac{b_4^2}{4b_6} - \frac{b_2^2}{4b_1}\right) =: \frac{\eta_G^4(C_{\mathrm{RK}}+1)}{4b_1b_6}R(c_2;\lambda,\alpha,\beta),\end{aligned}$$

where $R(c_2;\lambda,\alpha,\beta)$ is a quadratic polynomial in c_2 with the nonpositive leading term $-d^2(4-3\lambda)+4(2-3\lambda)d-4$. The polynomial $R(c_2;\lambda,\alpha,\beta)$ is nonnegative for some c_2 if and only if its discriminant $4d^2\lambda(1-\lambda)S(\lambda;\alpha,\beta)$ is nonnegative. Here, $S(\lambda;\alpha,\beta)$ is a quadratic polynomial in λ. In order to derive the conditions on (α,β) such that $S(\lambda;\alpha,\beta)\geq 0$ for some $\lambda\in(0,1)$, we employ the computer-algebra system `Mathematica`. The result of the command

```
Resolve[Exists[LAMBDA, S[LAMBDA] >= 0 && LAMBDA > 0
&& LAMBDA < 1], Reals]
```

gives all $(\alpha,\beta)\in\mathbb{R}^2$ such that there exist $c_1\leq 0$, $c_2\in\mathbb{R}$ such that $Q(\eta)\geq 0$. The interior of this region equals the set $R_0(d)$, defined in the statement of the theorem. This shows that $G''(0)\leq 0$ for all $(\alpha,\beta)\in R_0(d)$.

If $G''(0)=0$, the nonnegative polynomial Q has to vanish. In particular, $b_1\eta_L^2=0$. If $\eta_L=0$ in Ω, the boundary conditions imply that u is constant, which contradicts our assumption that u is not the steady state. Thus $b_1=0$. Similarly, $b_2=b_3=b_4=0$. This gives a system of four inhomogeneous linear equations for (c_1,c_2) which is unsolvable. Consequently, $G''(0)<0$.

The set $R_0(d)$ is nonempty since, e.g., $(1,1)\in R_0(d)$. Indeed, choosing $c_1=-1$ and $c_2=0$, we find that $Q(\eta)=(C_{\mathrm{RK}}+\frac{1}{d})\eta_L^2+\eta_R^2+d(d-1)\eta_S^2\geq 0$.

In one space dimension, the situation simplifies since the Laplacian coincides with the Hessian and thus, the integration-by-parts formula (2.20) is not needed. Then (see (2.21))

$$G''(0) = G''(0) + J_1 = -\beta^2\int_\Omega u^{2\beta+\alpha-1}\big(a_1\xi_L^2 + a_2\xi_L\xi_G^2 + a_3\xi_G^4\big)dx,$$

where

$$a_1 = C_{\mathrm{RK}}+1,\quad a_2 = (C_{\mathrm{RK}}+2)(\beta-\alpha)-3c_2,\quad a_3 = (\beta-\alpha)^2-(2\beta-2\alpha-1)c_2.$$

The polynomial $P(\xi)=\xi_G^4(a_1y^2+a_2y+a_3)$ with $y=\xi_L/\xi_G^2$ is nonnegative if and only if $a_1\geq 0$ and $4a_1a_3-a_2^2\geq 0$, which is equivalent to

$$-9c_2^2 + 2\big((C_{\mathrm{RK}}-2)(\alpha-\beta)+2(C_{\mathrm{RK}}+1)\big)c_2 - C_{\mathrm{RK}}^2(\alpha-\beta)^2 \geq 0. \qquad (2.25)$$

This inequality has a solution $c_2 \in \mathbb{R}$ if and only if the quadratic polynomial has real roots, i.e. if its discriminant is nonnegative,

$$\begin{aligned}0 &\leq \big((C_{\mathrm{RK}}-2)(\alpha-\beta)+2(C_{\mathrm{RK}}+1)\big)^2 - 9C_{\mathrm{RK}}^2(\alpha-\beta)^2 \\ &= 4(C_{\mathrm{RK}}+1)\left(-(2C_{\mathrm{RK}}-1)(\alpha-\beta)^2+(C_{\mathrm{RK}}-2)(\alpha-\beta)+(C_{\mathrm{RK}}+1)\right).\end{aligned}$$

The polynomial $-(2C_{\mathrm{RK}}-1)z^2+(C_{\mathrm{RK}}-2)z+(C_{\mathrm{RK}}+1)$ with $z=\alpha-\beta$ is always nonnegative if $C_{\mathrm{RK}}=0$ (implicit Euler). For $C_{\mathrm{RK}}=1$ and $C_{\mathrm{RK}}=2$, this property holds if and only if $-(C_{\mathrm{RK}}+1)/(2C_{\mathrm{RK}}-1) \leq \alpha-\beta \leq 1$. This concludes the proof. □

2.4.2 First-order entropies

We consider the one-dimensional case and first-order entropies with $f(u) = u^{\alpha/2}$, $\alpha > 0$.

Theorem 26. *Let $\Omega \subset \mathbb{R}$ be a bounded interval. Let (u^k) be a sequence of (smooth) solutions to the Runge-Kutta scheme* (2.2) *of order $p \geq 2$ for* (2.24) *in one space dimension. Let the entropy be given by $F[u] = \int_\Omega (u^{\alpha/2})_x^2 dx$ with $\alpha > 0$, let $k \in \mathbb{N}$ be fixed, and let u^k not be the constant steady state of* (2.24). *There exists a nonempty region $R_1 \in [0,\infty)^2$ and $\tau^k > 0$ such that for all $(\alpha,\beta) \in R_1$, there is a constant $C_{\alpha,\beta} > 0$ such that for all $0 < \tau \leq \tau^k$,*

$$F[u^k] + \tau C_{\alpha,\beta}\int_\Omega (u^k)^{\alpha+\beta-3}(u^k_{xx})^2 dx \leq F[u^{k-1}], \quad k \in \mathbb{N}.$$

Figure 2.2 illustrates the set R_1. The set of admissible values (α,β) for the continuous equation is given by $\{-2 \leq \alpha - 2\beta < 1\}$ (the borders of this set are depicted in the figure by dashed lines).

Proof. First, we compute $G'(0)$ according to Theorem 20:

$$G'(0) = -\alpha\int_\Omega u^{\alpha/2-1}(u^{\alpha/2})_{xx}(u^\beta)_{xx}dx.$$

We show that $G'(0)$ is nonpositive in a certain range of values (α,β). We formulate $G'(0)$ as

$$G'(0) = -\frac{\alpha^2\beta}{4}\int_\Omega u^{\alpha+\beta-1}\big((\alpha-2)(\beta-1)\xi_1^4+(\alpha+2\beta-4)\xi_1^2\xi_2+2\xi_2^2\big)dx,$$

where $\xi_1 = u_x/u$, $\xi_2 = u_{xx}/u$. We employ the integration-by-parts formula

$$0 = \int_\Omega (u^{\alpha+\beta-4}u_x^3)_x dx = \int_\Omega u^{\alpha+\beta-1}\big((\alpha+\beta-4)\xi_1^4+3\xi_1^2\xi_2\big)dx =: J.$$

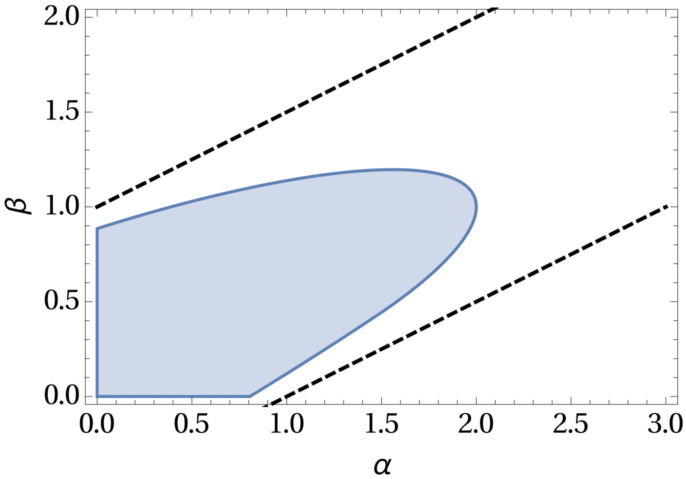

Figure 2.2: Set of all (α, β) for which the discrete first-order entropy for solutions to the one-dimensional porous-medium equation is dissipating. The continuous first-order entropy is dissipated for $-2 \leq \alpha - 2\beta < 1$. The borders of this set are indicated in the figure by dashed lines.

Therefore,

$$G'(0) = G'(0) - \frac{\alpha^2\beta}{4}cJ = -\frac{\alpha^2\beta}{4}\int_\Omega u^{\alpha+\beta-1}P(\xi)dx,$$

where

$$P(\xi) = \big((\alpha-2)(\beta-1) + (\alpha+\beta-4)c\big)\xi_1^4 + \big(\alpha+2\beta-4+3c\big)\xi_1^2\xi_2 + 2\xi_2^2.$$

This polynomial is nonnegative if and only if

$$8\big((\alpha-2)(\beta-1) + (\alpha+\beta-4)c\big) - (\alpha+2\beta-4+3c)^2 \geq 0,$$

which is equivalent to

$$g(c) := -9c^2 + 2(\alpha-2\beta-4)c - (\alpha-2\beta)^2 \geq 0.$$

The maximizing value $c^* = (\alpha - 2\beta - 4)/9$, obtained from $g'(c) = 0$, yields

$$g(c^*) = -\frac{8}{9}(\alpha-2\beta-1)(\alpha-2\beta+2) \geq 0$$

and consequently $G'(0) \leq 0$ if $-2 \leq \alpha - 2\beta \leq 1$. This condition is the same as in [27, Theorem 13] for the continuous equation.

Next, we turn to the proof of $G''(0) < 0$. The proof of Theorem 20 shows that

$$\begin{aligned} G''(0) = &-\frac{\alpha}{2}\int_\Omega \bigg(\frac{\alpha}{2}\big(u^{\alpha/2-1}(u^\beta)_{xx}\big)_x^2 - \Big(\frac{\alpha}{2}-1\Big)u^{\alpha/2-2}(u^{\alpha/2})_{xx}(u^\beta)_{xx}^2 \\ &- \beta C_{\mathrm{RK}}u^{\alpha/2-1}(u^{\alpha/2})_{xx}\big(u^{\beta-1}(u^\beta)_{xx}\big)_{xx}\bigg)dx. \end{aligned}$$

We integrate by parts in the last term and use $(\beta u^{\beta-1}(u^\beta)_{xx})_x = 0$ on $\partial\Omega$:

$$\begin{aligned} G''(0) = &-\frac{1}{8}\alpha^2\beta^2\int_\Omega u^{\alpha+2\beta-2} \\ &\times\big(a_1\xi_1^6 + a_2\xi_1^4\xi_2 + a_3\xi_1^3\xi_3 + a_4\xi_1^2\xi_2^2 + a_5\xi_1\xi_2\xi_3 + a_6\xi_2^3 + a_7\xi_3^2\big)dx, \end{aligned}$$

where $\xi_1 = u_x/u$, $\xi_2 = u_{xx}/u$, $\xi_3 = u_{xxx}/u$, and

$$\begin{aligned} a_1 &= (\beta-1)\big(2C_{\mathrm{RK}}\alpha^2\beta - 3C_{\mathrm{RK}}\alpha^2 + 2\alpha\beta^2 - 2(5C_{\mathrm{RK}}+3)\alpha\beta + (15C_{\mathrm{RK}}+4)\alpha \\ &\quad + 2\beta^3 - 14\beta^2 + 4(3C_{\mathrm{RK}}+7)\beta - 2(9C_{\mathrm{RK}}+8)\big), \\ a_2 &= (\beta-1)\big(4C_{\mathrm{RK}}\alpha^2 + (8C_{\mathrm{RK}}+7)\alpha\beta - (32C_{\mathrm{RK}}+9)\alpha + 12\beta^2 \\ &\quad - 2(8C_{\mathrm{RK}}+25)\beta + 6(8C_{\mathrm{RK}}+7)\big), \\ a_3 &= C_{\mathrm{RK}}\alpha^2 + 2\alpha\beta - (5C_{\mathrm{RK}}+2)\alpha + 4(C_{\mathrm{RK}}+1)\beta^2 - 2(5C_{\mathrm{RK}}+8)\beta \\ &\quad + 12(C_{\mathrm{RK}}+1), \\ a_4 &= 2(\beta-1)\big(2(4C_{\mathrm{RK}}+1)\alpha + 9\beta - (16C_{\mathrm{RK}}+13)\big), \\ a_5 &= 2(2C_{\mathrm{RK}}+1)\alpha + 4(2C_{\mathrm{RK}}+3)\beta - 16(C_{\mathrm{RK}}+1), \\ a_6 &= 2-\alpha, \quad a_7 = 2(C_{\mathrm{RK}}+1). \end{aligned}$$

We employ three integration-by-parts formulas:

$$\begin{aligned} 0 &= \int_\Omega \big(u^{\alpha+2\beta-5}u_{xx}^2u_x\big)_x dx \\ &= \int_\Omega u^{\alpha+2\beta-2}\big((\alpha+2\beta-5)\xi_1^2\xi_2^2 + 2\xi_1\xi_2\xi_3 + \xi_2^3\big)dx =: J_1, \\ 0 &= \int_\Omega \big(u^{\alpha+2\beta-6}u_{xx}u_x^3\big)_x dx \\ &= \int_\Omega u^{\alpha+2\beta-2}\big((\alpha+2\beta-6)\xi_1^4\xi_2 + \xi_1^3\xi_3 + 3\xi_1^2\xi_2^2\big)dx =: J_2, \\ 0 &= \int_\Omega \big(u^{\alpha+2\beta-7}u_x^5\big)_x dx = \int_\Omega u^{\alpha+2\beta-2}\big((\alpha+2\beta-7)\xi_1^6 + 5\xi_1^4\xi_2\big)dx =: J_3. \end{aligned}$$

Then

$$G''(0) = G''(0) - \frac{1}{8}\alpha^2\beta^2(c_1J_1 + c_2J_2 + c_3J_3) = -\frac{1}{8}\alpha^2\beta^2\int_\Omega u^{\alpha+2\beta-2}P(\xi)dx,$$

where $P(\xi) = b_1\xi_1^6 + b_2\xi_1^4\xi_2 + b_3\xi_1^3\xi_3 + b_4\xi_1^2\xi_2^2 + b_5\xi_1\xi_2\xi_3 + b_6\xi_2^3 + b_7\xi_3^2$, and the coefficients are given by

$$
\begin{aligned}
b_1 &= a_1 + (\alpha + 2\beta - 7)c_3, & b_2 &= a_2 + (\alpha + 2\beta - 6)c_2 + 5c_3,\\
b_3 &= a_3 + c_2, & b_4 &= a_4 + (\alpha + 2\beta - 5)c_1 + 3c_2,\\
b_5 &= a_5 + 2c_1, & b_6 &= a_6 + c_1,\\
b_7 &= a_7.
\end{aligned}
$$

Choosing $c_1 = -a_6$, we eliminate the cubic term ξ_2^3. Furthermore, setting, $x = \xi_2/\xi_1^2$ and $y = \xi_3/\xi_1^3$, we can write the polynomial P as a quadratic polynomial in (x, y):

$$Q(x, y) = \xi_1^6 P(\xi) = b_1 + b_2 x + b_3 y + b_4 x^2 + b_5 xy + b_7 y^2.$$

The following lemma is a consequence of the proof of Lemma 2.2 in [62].

Lemma 27. *The polynomial $p(x, y) = A + Bx + Cy + Dx^2 + Exy + Fy^2$ with $F > 0$ is nonnegative for all $(x, y) \in \mathbb{R}^2$ if and only if*

(i) $4DF - E^2 > 0$ *and* $A(4DF - E^2) - B^2F - C^2D + BCE \geq 0$, *or*
(ii) $4DF - E^2 = 0$ *and* $2BF - CE = 0$ *and* $4AF - C^2 \geq 0$.

Note that in case $4DF - E^2 = 0$ and $E \neq 0$, we may replace $2BF - CE = 0$ by the condition $2BEF = CE^2 = 4CDF$ or (since $F > 0$) $BE = 2CD$.
The first inequality in case (i),

$$
\begin{aligned}
0 &< 4b_4b_7 - b_5^2\\
&= -(C_{\mathrm{RK}} + 1)(2C_{\mathrm{RK}} + 1)\alpha^2 + (2C_{\mathrm{RK}} + 2)(4C_{\mathrm{RK}} - 3)\alpha\beta + (9C_{\mathrm{RK}} + 9)\alpha\\
&\quad - 2C_{\mathrm{RK}}(4C_{\mathrm{RK}} + 3)\beta^2 + (8C_{\mathrm{RK}} + 12)\beta + (3C_{\mathrm{RK}} + 3)c_2 - (12C_{\mathrm{RK}} + 14),
\end{aligned}
$$

is linear in c_2 and provides a lower bound for c_2:

$$
\begin{aligned}
c_2 > \frac{1}{3(C_{\mathrm{RK}} + 1)}\Big(&(C_{\mathrm{RK}} + 1)(2C_{\mathrm{RK}} + 1)\alpha^2 - (2C_{\mathrm{RK}} + 2)(4C_{\mathrm{RK}} - 3)\alpha\beta\\
&- (9C_{\mathrm{RK}} + 9)\alpha + 2C_{\mathrm{RK}}(4C_{\mathrm{RK}} + 3)\beta^2 - (8C_{\mathrm{RK}} + 12)\beta + (12C_{\mathrm{RK}} + 14)\Big) =: c_2^*.
\end{aligned}
$$

The second inequality in case (i) becomes

$$
\begin{aligned}
0 &\leq b_1(4b_4b_7 - b_5^2) - b_2^2 b_7 - b_3^2 b_4 + b_2 b_3 b_5\\
&= -50(C_{\mathrm{RK}} + 1)c_3^2 + p_1(\alpha, \beta, c_2)c_3 + p_2(\alpha, \beta, c_2),
\end{aligned}
$$

where p_1 and p_2 are polynomials in α, β, and c_2. This quadratic expression in c_3 is nonnegative if and only if its discriminant is nonnegative,

$$
\begin{aligned}
0 &\leq -200(C_{\mathrm{RK}}+1)p_2(\alpha,\beta,c_2) - p_1(\alpha,\beta,c_2)^2 \\
&= -8\big(4b_4b_7 - b_5^2\big)\big(25c_2^2 + p_3(\alpha,\beta)c_2 + p_4(\alpha,\beta)\big),
\end{aligned}
$$

where $p_3(\alpha,\beta)$ and $p_4(\alpha,\beta)$ are polynomials in α and β. The factor $4b_4b_7 - b_5^2$ is positive, so we have to ensure that $R_{\alpha,\beta}(c_2) = 25c_2^2 + p_3(\alpha,\beta)c_2 + p_4(\alpha,\beta) \leq 0$ for some $c_2 > c_2^*$. Therefore we must ensure that the rightmost root of $R_{\alpha,\beta}(c_2)$ is larger or equal than the lower bound for c_2, i.e., $-p_3(\alpha,\beta) + \sqrt{p_3^2(\alpha,\beta) - 100p_4(\alpha,\beta)} \geq 50c_2^*$. For $C_{\mathrm{RK}} = 1$, the values (α,β) for which there exists $c_2 > c_2^*$ such that $R_{\alpha,\beta}(c_2) \leq 0$ are depicted in Figure 2.2. In case (ii), we may immediately calculate c_2 and c_3, but this yields a region which is already contained in the first one. This shows that $G''(0) \leq 0$.
If $G''(0) = 0$, the polynomial Q vanishes. Thus, either $u_x/u = \xi_1 = 0$ or $P(\xi) = 0$ in Ω. The first case is impossible since u is not constant in Ω. As $b_7 = a_7 = 2(C_{\mathrm{RK}}+1) > 0$, the second case $P(\xi) = 0$ implies that $\xi_3 = 0$. Hence, u is a quadratic polynomial. In view of the boundary conditions, u must be constant, but this contradicts our assumption. Hence, $G''(0) < 0$. $\square$

2.5 Linear diffusion system

We consider the following linear diffusion system:

$$
\partial_t u_1 - \rho_1 \Delta u_1 = \mu(u_2 - u_1), \quad \partial_t u_2 - \rho_2 \Delta u_2 = \mu(u_1 - u_2), \tag{2.26}
$$

with initial and homogeneous Neumann boundary conditions, $\rho_1,\ \rho_2,\ \mu > 0$, and the entropy

$$
H[u] = \int_\Omega h(u)dx = \int_\Omega \sum_{i=1}^{2} u_i(\log u_i - 1)dx, \tag{2.27}
$$

where $u = (u_1, u_2)$. If the initial data is nonnegative, the maximum principle shows that the solutions to (2.26) are nonnegative too.

Theorem 28. *Let (u^k) be a sequence of (smooth) nonnegative solutions to the Runge-Kutta scheme* (2.2) *for* (2.26) *with $C_{\mathrm{RK}} = 1$ and $\rho := \rho_1 = \rho_2$. Let the entropy H be given by* (2.27). *Let $k \in \mathbb{N}$ be fixed and let u^k not be the steady state of* (2.2). *Then there exists $\tau^k > 0$ such that for all $0 < \tau < \tau^k$,*

$$
H[u^k] + \tau \int_\Omega \left(\rho \sum_{i=1}^{2} \frac{|\nabla u_i^k|^2}{u_i^k} + \mu(\log u_1^k - \log u_2^k)(u_1^k - u_2^k) \right) dx \leq H[u^{k-1}].
$$

Note that we need equal diffusivities $\rho_1 = \rho_2$ and higher-order schemes ($C_{\mathrm{RK}} = 1$). These conditions are in accordance of [71], where the continuous equation was studied. In order to highlight the step where these conditions are needed, the following proof is slightly more general than actually needed.

Proof. We fix $k \in \mathbb{N}$ and set $u := u^k$. Let $A[u] = (A_1[u], A_2[u]) = (\rho_1 \Delta u_1 + \mu(u_2 - u_1), \rho_2 \Delta u_2 + \mu(u_1 - u_2))$. Since A is linear, $DA[u](h) = A[h]$. Thus,

$$G''(0) = -\int_\Omega \big(C_{\mathrm{RK}} h'(u)^\top A[A[u]] + A[u]^\top h''(u) A[u]\big) dx = -G_1 - G_2.$$

In the following, we set $\partial_i h = \partial h / \partial u_i$ for $i = 1, 2$. We integrate by parts twice, using the boundary conditions $\nabla u_i \cdot \nu = 0$ and $\nabla A_i[u] \cdot \nu = 0$ on $\partial\Omega$, and rearrange the terms:

$$\begin{aligned}
G_1 &= C_{\mathrm{RK}} \int_\Omega \Big(\partial_1 h(u) \big(\rho_1 \Delta A_1[u] + \mu(A_2[u] - A_1[u])\big) \\
&\quad + \partial_2 h(u) \big(\rho_2 \Delta A_2[u] + \mu(A_1[u] - A_2[u])\big)\Big) dx \\
&= C_{\mathrm{RK}} \int_\Omega \Big(\rho_1 \Delta \partial_1 h(u) A_1[u] + \rho_2 \Delta \partial_2 h(u) A_2[u] \\
&\quad + \mu(\partial_1 h(u) - \partial_2 h(u))(A_2[u] - A_1[u])\Big) dx \\
&= C_{\mathrm{RK}} \int_\Omega \Big(\rho_1 \big(\partial_1^2 h(u) \Delta u_1 + \partial_1^3 h(u) |\nabla u_1|^2\big)\big(\rho_1 \Delta u_1 + \mu(u_2 - u_1)\big) \\
&\quad + \rho_2 \big(\partial_2^2 h(u) \Delta u_2 + \partial_2^3 h(u) |\nabla u_2|^2\big)\big(\rho_2 \Delta u_2 + \mu(u_1 - u_2)\big) \\
&\quad + \mu(\partial_2 h(u) - \partial_1 h(u))\big(\rho_1 \Delta u_1 - \rho_2 \Delta u_2 + 2\mu(u_2 - u_1)\big)\Big) dx \\
&= C_{\mathrm{RK}} \int_\Omega \Big(\rho_1^2 \partial_1^2 h(u) (\Delta u_1)^2 + \rho_2^2 \partial_2^2 h(u) (\Delta u_2)^2 + \rho_1^2 \partial_1^3 h(u) \Delta u_1 |\nabla u_1|^2 \\
&\quad + \rho_2^2 \partial_2^3 h(u) \Delta u_2 |\nabla u_2|^2 + \rho_1 \mu \big(\partial_1^2 h(u)(u_2 - u_1) + \partial_2 h(u) - \partial_1 h(u)\big) \Delta u_1 \\
&\quad + \rho_2 \mu \big(\partial_2^2 h(u)(u_1 - u_2) + \partial_1 h(u) - \partial_2 h(u)\big) \Delta u_2 + \rho_1 \mu \partial_1^3 h(u)(u_2 - u_1)|\nabla u_1|^2 \\
&\quad + \rho_2 \mu \partial_2^3 h(u)(u_1 - u_2)|\nabla u_2|^2 + 2\mu^2 (\partial_2 h(u) - \partial_1 h(u))(u_2 - u_1)\Big) dx.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
G_2 &= \int_\Omega \Big(\partial_1^2 h(u) \big(\rho_1 \Delta u_1 + \mu(u_2 - u_1)\big)^2 + \partial_2^2 h(u) \big(\rho_2 \Delta u_2 + \mu(u_1 - u_2)\big)^2 \Big) dx \\
&= \int_\Omega \Big(\rho_1^2 \partial_1^2 h(u) (\Delta u_1)^2 + \rho_2^2 \partial_2^2 h(u) (\Delta u_2)^2 + 2\rho_1 \mu \partial_1^2 h(u)(u_2 - u_1) \Delta u_1 \\
&\qquad + 2\rho_2 \mu \partial_2^2 h(u)(u_1 - u_2) \Delta u_2 + \mu^2 (\partial_1^2 h(u) + \partial_2^2 h(u))(u_1 - u_2)^2 \Big) dx.
\end{aligned}$$

Adding G_1 and G_2, we arrive at

$$\begin{aligned}
G''(0) &= -\sum_{i=1}^{2}\int_\Omega \Big(\rho_i^2(C_{\mathrm{RK}}+1)\partial_i^2 h(u)(\Delta u_i)^2 + \rho_i^2 C_{\mathrm{RK}}\partial_i^3 h(u)\Delta u_i|\nabla u_i|^2\Big)dx \\
&\quad - \int_\Omega \Big(\rho_1\mu\big((C_{\mathrm{RK}}+2)\partial_1^2 h(u)(u_2-u_1) + C_{\mathrm{RK}}(\partial_2 h(u)-\partial_1 h(u))\big)\Delta u_1 \\
&\quad + \rho_2\mu\big((C_{\mathrm{RK}}+2)\partial_2^2 h(u)(u_1-u_2) + C_{\mathrm{RK}}(\partial_1 h(u)-\partial_2 h(u))\big)\Delta u_2 \\
&\quad + \rho_1\mu C_{\mathrm{RK}}\partial_1^3 h(u)(u_2-u_1)|\nabla u_1|^2 + \rho_2\mu C_{\mathrm{RK}}\partial_2^3 h(u)(u_1-u_2)|\nabla u_2|^2\Big)dx \\
&\quad - \int_\Omega \mu^2\Big(2(\partial_1 h(u)-\partial_2 h(u)) + (\partial_1^2 h(u)+\partial_2^2 h(u))(u_1-u_2)\Big)(u_1-u_2)dx \\
&= -I_2 - I_1 - I_0.
\end{aligned}$$

The idea of [71] is to show that each integral I_i, involving only derivatives of order i, is nonnegative. In contrast to [71], we employ systematic integration by parts, which allows for a simpler and more general proof in our case. For the term I_2, we use the following integration-by-parts formula:

$$0 = \int_\Omega \operatorname{div}\big(u_i^{-2}|\nabla u_i|^3\big)dx = \int_\Omega \big(-2u_i^{-3}|\nabla u_i|^4 + 3u_i^{-2}\Delta u_i|\nabla u_i|^2\big)dx =: J_i.$$

Then, for $\varepsilon > 0$,

$$\begin{aligned}
I_2 - c\sum_{i=1}^{2}\rho_i^2 J_i - \varepsilon\sum_{i=1}^{2} u_i^{-3}|\nabla u_i|^4 dx = \sum_{i=1}^{2}\rho_i^2\int_\Omega \Big((C_{\mathrm{RK}}+1)u_i^{-1}(\Delta u_i)^2 \\
-(3c+C_{\mathrm{RK}})u_i^{-2}\Delta u_i|\nabla u_i|^2 + (2c-\varepsilon)u_i^{-3}|\nabla u_i|^4\Big)dx.
\end{aligned}$$

The integrand defines a quadratic polynomial in the variables Δu_i and $|\nabla u_i|^2$ and is nonnegative if its discriminant satisfies $4(2c-\varepsilon)(C_{\mathrm{RK}}+1)-(3c+C_{\mathrm{RK}})^2 \geq 0$. It turns out that this inequality holds for $C_{\mathrm{RK}} \in \{0,1\}$ if we choose $c = 2/3$ and $\varepsilon > 0$ sufficiently small. When $C_{\mathrm{RK}} = 2$, we can only show that $I_2 \geq 0$ which is not sufficient to prove that $G''(0) < 0$ (see below). We conclude that

$$I_2 \geq \varepsilon\sum_{i=1}^{2}\int_\Omega u_i^{-3}|\nabla u_i|^4 dx. \tag{2.28}$$

Integrating by parts in I_1 in order to obtain only first-order derivatives, we find after some rearrangements that

$$\begin{aligned}
I_1 &= \mu\int_\Omega \big(a_1|\nabla\log u_1|^2 + a_2\nabla\log u_1\cdot\nabla\log u_2 + a_3|\nabla\log u_2|^2\big)dx, \quad \text{where} \\
a_1 &= 2\rho_1(C_{\mathrm{RK}}u_1+u_2), \quad a_3 = 2\rho_2(C_{\mathrm{RK}}u_2+u_1), \\
a_2 &= -(C_{\mathrm{RK}}(\rho_1+\rho_2)+2\rho_2)u_1 - (C_{\mathrm{RK}}(\rho_1+\rho_2)+2\rho_1)u_2.
\end{aligned}$$

The integrand is nonnegative if and only if $4a_1a_3 - a_2^2 \geq 0$ for all (u_1, u_2). We compute:

$C_{\mathrm{RK}} = 0$: $4a_1a_3 - a_2^2 = -4(\rho_1 u_2 - \rho_2 u_1)^2$,
$C_{\mathrm{RK}} = 1$: $4a_1a_3 - a_2^2 = (\rho_1 - \rho_2)\big(\rho_1(u_1^2 + 6u_1u_2 + 9u_2^2) - \rho_2(9u_1^2 + 6u_1u_2 + u_2^2)\big)$,
$C_{\mathrm{RK}} = 2$: $4a_1a_3 - a_2^2 = -4\big(\rho_1(u_1 + 2u_2) - \rho_2(2u_1 + u_2)\big)$.

Thus, $4a_1a_3 - a_2^2 \geq 0$ is possible only if $\rho_1 = \rho_2$ and $C_{\mathrm{RK}} = 1$.
Finally, we immediately see that the remaining term

$$I_0 = \mu^2 \int_\Omega \left(2(\log u_1 - \log u_2)(u_1 - u_2) + \left(\frac{1}{u_1} + \frac{1}{u_2}\right)(u_1 - u_2)^2 \right) dx$$

is nonnegative. This shows that $G''(0) \leq 0$. If $G''(0) = 0$, we infer from (2.28) that $u_i = \text{const.}$, but this contradicts our hypothesis that u_i is not a steady state. $\square$

2.6 The Derrida-Lebowitz-Speer-Spohn equation

Consider the one-dimensional fourth-order equation

$$\partial_t u = -(u(\log u)_{xx})_{xx} \quad \text{in } \Omega,\ t > o, \quad u(0) = u^0 \tag{2.29}$$

with periodic boundary conditions. This equation appears as a scaling limit of the so-called (time-discrete) Toom model, which describes interface fluctuations in a two-dimensional spin system [34]. The variable u is the limit of a random variable related to the deviation of the spin interface from a straight line. The multi-dimensional version of (2.29) models the electron density u in a quantum semiconductor, the equation is the zero-temperature, zero-field approximation of the quantum drift-diffusion model [61]. For existence results for (2.29), we refer to [59] and the references therein.
To simplify our calculations, we analyze only the logarithmic entropy $H[u] = \int_\Omega u(\log u - 1)dx$. It is also possible to verify condition (2.6) for entropies of the form $\int_\Omega u^\alpha dx$, but it turns out that only sufficiently small $\alpha > 0$ are admissible (about $0 < \alpha < 0.15\ldots$) and the computations are very tedious. Therefore, we restrict ourselves to the case $\alpha = 0$.

Theorem 29. *Let (u^k) be a sequence of (smooth) solutions to the Runge-Kutta scheme (2.2) with $C_{\mathrm{RK}} = 1$ for (2.29). Let the entropy be given by $H[u] = \int_\Omega u(\log u - 1)dx$, let $k \in \mathbb{N}$ be fixed, and let u^k not be a steady state. Then there exists $\tau^k > 0$ such that for all $0 < \tau < \tau^k$,*

$$H[u^k] + \tau q \int_\Omega u(\log u)_x^8 dx + \tau \int_\Omega u(\log u)_{xx}^2 dx \leq H[u^{k-1}], \quad q \approx 0.0045.$$

Proof. First, we observe that

$$G'(0) = -\int_\Omega (u(\log u)_{xx})_{xx} \log u dx = -\int_\Omega u(\log u)_{xx}^2 dx.$$

With $A[u] = (u(\log u)_{xx})_{xx}$ and $DA[u](h) = \big(h_{xx} - 2(\log u)_x h_x + (\log u)_x^2 h\big)_{xx}$, we can write $G''(0) = -I_0^k$ according to (2.6) as

$$\begin{aligned} G''(0) &= -\int_\Omega \Big(\log u\big(A[u]_{xx} - 2(\log u)_x A[u]_x + (\log u)_x^2 A[u]\big)_{xx} + \frac{1}{u} A[u]^2 \Big) dx \\ &= -\int_\Omega \Big((\log u)_{xx}\big(A[u]_{xx} - 2(\log u)_x A[u]_x + (\log u)_x^2 A[u]\big) + \frac{1}{u} A[u]^2 \Big) dx \\ &= -\int_\Omega \Big(\big(v_{xxxx} + 2(v_x v_{xx})_x + v_x^2 v_{xx}\big) A[u] + \frac{1}{u} A[u]^2 \Big) dx, \end{aligned}$$

where we have integrated by parts several times and have set $v = \log u$. Then $A[u] = u(v_x^2 v_{xx} + 2 v_x v_{xxx} + v_{xx}^2 + v_{xxxx})$ and, with the abbreviations $\xi_1 = v_x, \ldots, \xi_4 = v_{xxxx}$,

$$\begin{aligned} G''(0) = -\int_\Omega u\Big(& 2\xi_1^4\xi_2^2 + 8\xi_1^3\xi_2\xi_3 + 5\xi_1^2\xi_2^3 + 4\xi_1^2\xi_2\xi_4 + 8\xi_1^2\xi_3^2 + 10\xi_1\xi_2^2\xi_3 \\ & + 8\xi_1\xi_3\xi_4 + 3\xi_2^4 + 5\xi_2^2\xi_4 + 2\xi_4^2\Big) dx. \end{aligned}$$

We employ the following integration-by-parts formulas:

$$\begin{aligned} 0 &= \int_\Omega (u v_x^7)_x dx = \int_\Omega u(\xi_1^8 + 7\xi_1^6\xi_2) dx =: J_1, \\ 0 &= \int_\Omega (u v_{xx} v_x^5)_x dx = \int_\Omega u(\xi_1^6\xi_2 + \xi_1^5\xi_3 + 5\xi_1^4\xi_2^2) dx =: J_2, \\ 0 &= \int_\Omega (u v_{xxx} v_x^4)_x dx = \int_\Omega u(\xi_1^5\xi_3 + \xi_1^4\xi_4 + 4\xi_1^3\xi_2\xi_3) dx =: J_3, \\ 0 &= \int_\Omega (u v_{xx}^2 v_x^3)_x dx = \int_\Omega u(\xi_1^4\xi_2^2 + 2\xi_1^3\xi_2\xi_3 + 3\xi_1^2\xi_2^3) dx =: J_4, \\ 0 &= \int_\Omega (u v_{xx} v_{xxx} v_x^2)_x dx = \int_\Omega u(\xi_1^3\xi_2\xi_3 + \xi_1^2\xi_2\xi_4 + \xi_1^2\xi_3^2 + 2\xi_1\xi_2^2\xi_3) dx =: J_5, \\ 0 &= \int_\Omega (u v_{xxx}^2 v_x)_x dx = \int_\Omega u(\xi_1^2\xi_3^2 + 2\xi_1\xi_3\xi_4 + \xi_2\xi_3^2) dx =: J_6, \\ 0 &= \int_\Omega (u v_{xx}^3 v_x)_x dx = \int_\Omega u(\xi_1^2\xi_2^3 + 3\xi_1\xi_2^2\xi_3 + \xi_2^4) dx =: J_7, \\ 0 &= \int_\Omega (u v_{xxx} v_{xx}^2)_x dx = \int_\Omega u(\xi_1\xi_2^2\xi_3 + 2\xi_2\xi_3^2 + \xi_2^2\xi_4) dx =: J_8. \end{aligned}$$

Then

$$G''(0) = G''(0) - 4\sum_{i=1}^{8} c_i J_i = -\int_\Omega u\Big(a_1\xi_1^8 + a_2\xi_1^6\xi_2 + a_3\xi_1^5\xi_3 + a_4\xi_1^4\xi_2^2 + a_5\xi_1^4\xi_4$$
$$+ a_6\xi_1^3\xi_2\xi_3 + a_7\xi_1^2\xi_2^3 + a_8\xi_1^2\xi_2\xi_4 + a_9\xi_1^2\xi_3^2 + a_{10}\xi_1\xi_2^2\xi_3 + a_{11}\xi_1\xi_3\xi_4$$
$$+ a_{12}\xi_2^4 + a_{13}\xi_2^2\xi_4 + a_{14}\xi_2\xi_3^2 + a_{15}\xi_4^2\Big)\,dx,$$

where

$$\begin{aligned}
&a_1 = 4c_1, && a_2 = 28c_1 + 4c_2, && a_3 = 4c_2 + 4c_3,\\
&a_4 = 2 + 20c_2 + 4c_4, && a_5 = 4c_3, && a_6 = 8 + 16c_3 + 8c_4 + 4c_5,\\
&a_7 = 5 + 12c_4 + 4c_7, && a_8 = 4 + 4c_5, && a_9 = 8 + 4c_5 + 4c_6,\\
&a_{10} = 10 + 8c_5 + 12c_7 + 4c_8, && a_{11} = 8 + 8c_6, && a_{12} = 3 + 4c_7,\\
&a_{13} = 5 + 4c_8, && a_{14} = 4c_6 + 8c_8, && a_{15} = 2.
\end{aligned}$$

Next, we eliminate all terms involving ξ_4 by formulating the following square:

$$\begin{aligned}
G''(0) = -\int_\Omega u\bigg[&a_{15}\left(\xi_4 + \frac{a_5}{2a_{15}}\xi_1^4 + \frac{a_8}{2a_{15}}\xi_1^2\xi_2 + \frac{a_{11}}{2a_{15}}\xi_1\xi_3 + \frac{a_{13}}{2a_{15}}\xi_2^2\right)^2\\
&+ \left(a_1 - \frac{a_5^2}{4a_{15}}\right)\xi_1^8 + \left(a_2 - \frac{a_5a_8}{2a_{15}}\right)\xi_1^6\xi_2 + \left(a_3 - \frac{a_5a_{11}}{2a_{15}}\right)\xi_1^5\xi_3\\
&+ \left(a_4 - \frac{a_8^2}{4a_{15}} - \frac{a_5a_{13}}{2a_{15}}\right)\xi_1^4\xi_2^2 + \left(a_6 - \frac{a_8a_{11}}{2a_{15}}\right)\xi_1^3\xi_2\xi_3 + \left(a_7 - \frac{a_8a_{13}}{2a_{15}}\right)\xi_1^2\xi_2^3\\
&+ \left(a_9 - \frac{a_{11}^2}{4a_{15}}\right)\xi_1^2\xi_3^2 + \left(a_{10} - \frac{a_{11}a_{13}}{2a_{15}}\right)\xi_1\xi_2^2\xi_3 + \left(a_{12} - \frac{a_{13}^2}{4a_{15}}\right)\xi_2^4 + a_{14}\xi_2\xi_3^2\bigg]\,dx.
\end{aligned}$$

We eliminate all terms involving ξ_3 and set the corresponding coefficients to zero. From $a_{14} = 0$ we conclude that $c_6 = -2c_8$. Furthermore,

$$\begin{aligned}
a_9 - \frac{a_{11}^2}{4a_{15}} &= 0 &&\text{gives} && c_5 = 8c_8^2 - 6c_8,\\
a_{10} - \frac{a_{11}a_{13}}{2a_{15}} &= 0 &&\text{gives} && c_7 = -\frac{20}{3}c_8^2 + \frac{8}{3}c_8,\\
a_6 - \frac{a_8a_{11}}{2a_{15}} &= 0 &&\text{gives} && c_4 = -2c_3 - 16c_8^3 + 16c_8^2 - 5c_8,\\
a_3 - \frac{a_5a_{11}}{2a_{15}} &= 0 &&\text{gives} && c_2 = c_3 - 4c_3c_8.
\end{aligned}$$

Following these choices, we obtain

$$b_{12} := a_{12} - \frac{a_{11}^2}{4a_{15}} = -\frac{86}{3}c_8^2 + \frac{17}{3}c_8 - \frac{1}{8}.$$

This quadratic polynomial in c_8 takes on its maximal value at $c_8^* = 17/172$ with value $b_{12} = 20/129$. The integral can now be written as

$$G''(0) \leq -\int_\Omega u\big(b_1\xi_1^8 + b_2\xi_1^6\xi_2 + b_4\xi_1^4\xi_2^2 + b_7\xi_1^2\xi_2^3 + b_{12}\xi_2^4\big)dx,$$

where

$$\begin{aligned}
b_1 &= a_1 - \frac{a_5^2}{4a_{15}} = 4c_1 - 2c_3^2, \\
b_2 &= a_2 - \frac{a_5 a_8}{2a_{15}} = 28c_1 - 32c_3c_8^2 + 8c_3c_8, \\
b_4 &= a_4 - \frac{a_8^2}{4a_{15}} - \frac{a_5a_{13}}{2a_{15}} = 7c_3 - 84c_3c_8 - 128c_8^4 + 128c_8^3 - 40c_8^2 + 4c_8, \\
b_7 &= a_7 - \frac{a_8a_{13}}{2a_{15}} = -24c_3 - 244c_8^3 + \frac{448}{3}c_8^2 - \frac{70}{3}c_8.
\end{aligned}$$

If $b_4 = 2b_2b_{12}/b_7 + b_7^2/(4b_{12})$, we can write the integral as the sum of two squares, noting that b_{12} is positive,

$$G''(0) \leq -\int_\Omega u\left(b_{12}\left(\xi_2^2 + \frac{b_7}{2b_{12}}\xi_1^2\xi_2 + \frac{b_2}{b_7}\xi_1^4\right)^2 + \left(b_1 - \frac{b_2^2b_{12}}{b_7^2}\right)\xi_1^8\right)dx.$$

The expression $b_4b_7 - 2b_2b_{12} - b_7^3/(4b_{12}) = 0$ defines a polynomial in (c_1, c_3) which is linear in c_1. Solving it for c_1 gives

$$c_1 = \frac{449307}{175}c_3^3 + \frac{741681}{2150}c_3^2 + \frac{35780649411}{2393160700}c_3 + \frac{34135130165539}{163091166664200}.$$

It remains to show that $p(c_3) := b_1 - b_2^2b_{12}/b_7^2$, which is a polynomial of fourth order in c_3, is positive. Choosing $c_3^* = -0.029$, we find that $p(c_3^*) \approx 0.0045 > 0$. This shows that

$$G''(0) \leq -q(c_3^*)\int_\Omega u\xi_1^8 dx = -q(c_3^*)\int_\Omega u(\log u)_x^8 dx \leq 0.$$

Finally, if $G''(0) = 0$, we infer that u is constant which is contradicts the assumption that u is not the steady state. Therefore, $G''(0) < 0$, which ends the proof. □

2.7 Numerical examples

The aim of this section is to explore the numerical behavior of the second-order derivative of the function $G(\tau)$, defined in the introduction, for the porous-medium equation (2.24) in one space dimension. The equation is discretized by standard finite differences, and we employ periodic boundary conditions. The discrete solution u_i^k approximates the solution $u(x_i, t^k)$ to (2.24) with $x_i = i\triangle x$, $t^k = k\tau$, where $\triangle x$, τ are the space and time step sizes, respectively. As in the previous chapter, we choose the Barenblatt profile which reads in one space dimension

$$u^0(x) = t_0^{-1/(\beta+1)} \max\left(0, C - \frac{\beta-1}{2\beta(\beta+1)} \frac{(x-1/2)^2}{t_0^{2/(\beta+1)}}\right)^{1/(\beta-1)}, \quad 0 \le x \le 1, \tag{2.30}$$

where

$$t_0 = 0.01, \quad C = \frac{\beta-1}{2\beta(\beta+1)} \frac{(x_R-1/2)^2}{t_0^{2/(\beta+1)}}, \quad x_R = \frac{1}{4},$$

as the initial datum. Its support is contained in $[\frac{1}{2} - x_R, \frac{1}{2} + x_R]$; see Figure 2.3 (left). We choose the exponent $\beta = 2$. The semi-logarithmic plot of the discrete entropy $H_d[u^k] = \sum_{i=0}^N (u_i^k)^\alpha \triangle x$ with $\alpha = 5$ versus time is illustrated in Figure 2.3 (right), using the implicit Euler scheme with parameters $\tau = 10^{-4}$ and the number of grid points $N = 1/\triangle x = 64$. The decay is exponential for "large" times. The nonlinear discrete system is solved by Newton's method with the tolerance `tol` $= 10^{-15}$. We have highlighted four time steps t_i at which we compute the function $G(\tau)$ numerically for the following Runge-Kutta schemes:

explicit Euler scheme:	$u^k - u^{k-1} = -\tau A[u^{k-1}]$,
implicit Euler scheme:	$u^k - u^{k-1} = -\tau A[u^k]$,
second-order trapezoidal rule:	$u^k - u^{k-1} = -\frac{\tau}{2}(A[u^k] + A[u^{k-1}])$,
third-order Simpson rule:	$u^k - u^{k-1} = -\frac{\tau}{6}\big(A[u^k] + 4A[(u^k + u^{k-1})/2] + A[u^{k-1}]\big)$.

We set as before $u := u^k$, $v(\tau) := u^{k-1}$ and compute $G(\tau) = H_d[u] - H_d[v(\tau)]$ and the discrete second-order derivative $\partial^2 G$ of G (using central differences). The result is presented in Figure 2.4. As expected, the discrete derivative $\partial^2 G$ is negative on a (small) interval for all times t_i, $i = 1, 2, 3$. We observe that $\partial^2 G$ is even slightly decreasing, but we expect that it becomes positive for sufficiently large values of τ. Clearly, the values for $\partial^2 G$ tend to zero as we approach the steady state (see Remark 21). This experiment indicates that τ^k from Theorem 18 is bounded from below by $\tau^* = 3 \cdot 10^{-4}$, for instance.
In order to understand the behavior of $G(\tau)$ in a better way, it is convenient to study the discrete version of the quotient

$$Q(\tau) := \frac{G''(\tau)}{\|u^{\alpha+2\beta-2} u_x^4\|_{L^1}}. \tag{2.31}$$

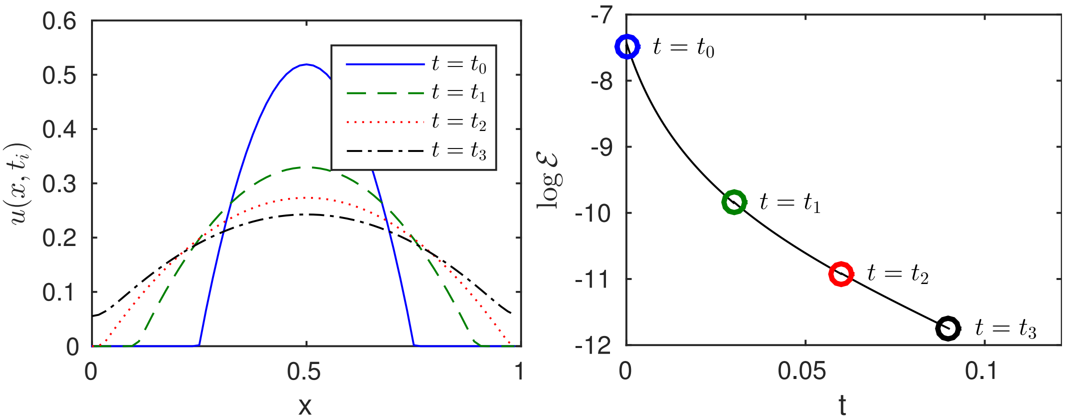

Figure 2.3: Left: Evolution of the initial datum (2.30) for $\beta = 2$ at various time steps t_i, $i = 0, 1, 2, 3$. Right: Semi-logarithmic plot of the discrete entropy $H_d[u^k]$ versus time.

Indeed, the analysis in Section 2.4 gives an estimate of the type $G''(0) \leq -C \int_\Omega u^{2\beta+\alpha-5} u_x^4 dx$ for some constant $C > 0$. Thus, we expect that for sufficiently small $\tau > 0$, $Q(\tau)$ is bounded from above by some negative constant. This expectation is confirmed in Figure 2.5. In the depicted examples, $Q(\tau)$ is a decreasing function of τ, and $Q(0)$ is decreasing with increasing time.
All these results indicate that the threshold parameter τ^k in Theorem 18 can be chosen independently of the time step k.

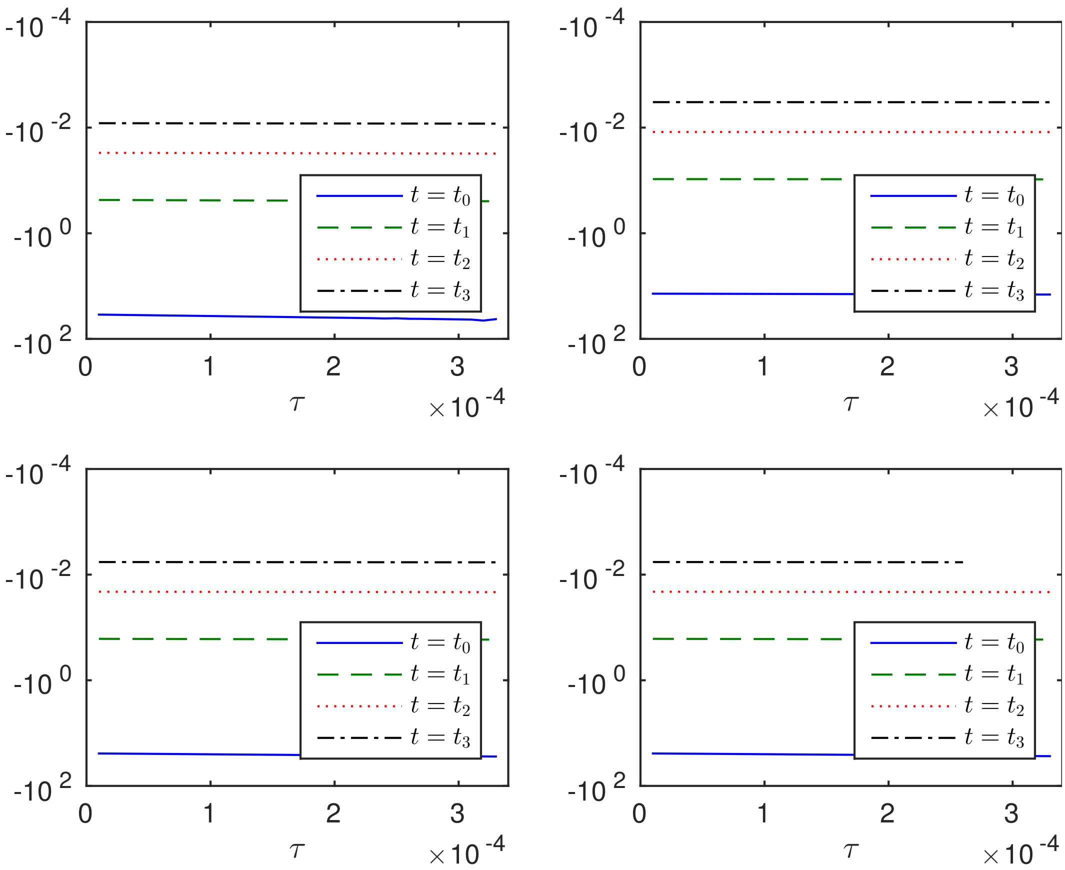

Figure 2.4: Numerical evaluation of the discrete version of $G''(\tau)$ for various Runge-Kutta schemes at the time steps t_i: Explicit Euler scheme (top left); Implicit Euler scheme (top right); Implicit trapezoidal rule (bottom left); Simpson rule (bottom right).

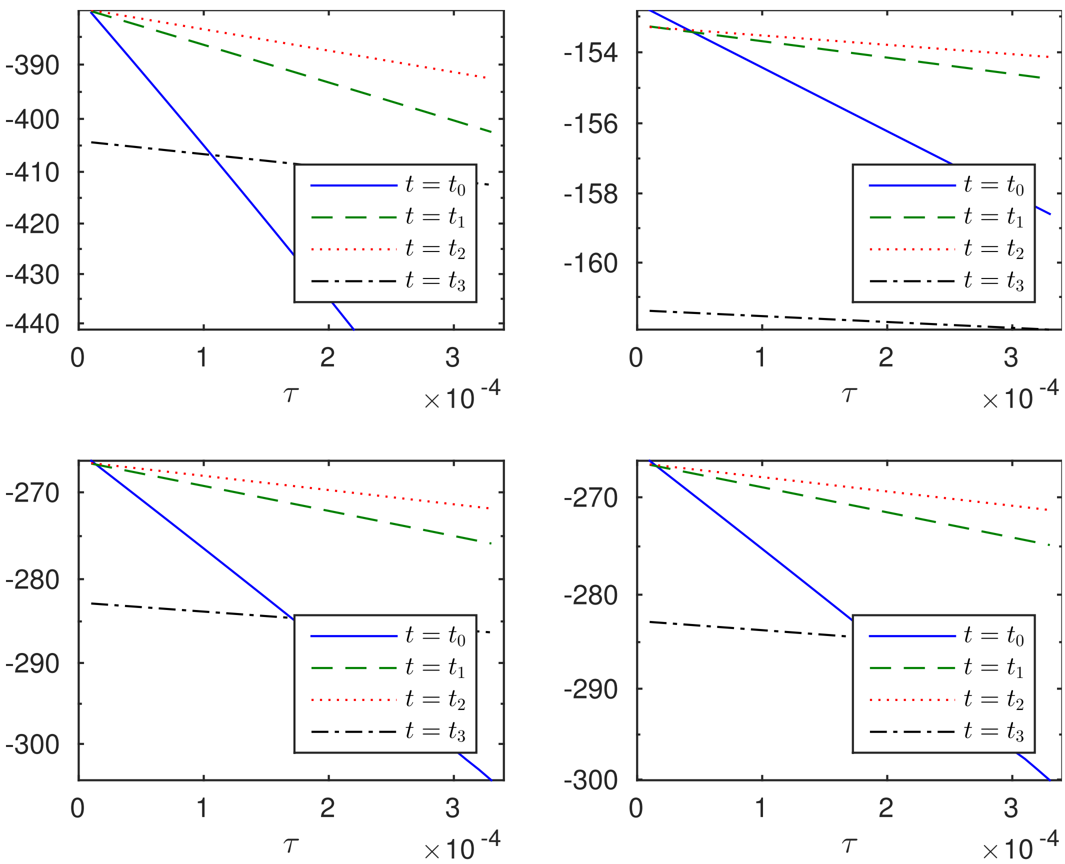

Figure 2.5: Numerical evaluation of the discrete version of $Q(\tau)$, defined in (2.31), for various Runge-Kutta schemes at the time steps t_i: Explicit Euler scheme (top left); Implicit Euler scheme (top right); Implicit trapezoidal rule (bottom left); Simpson rule (bottom right).

Chapter 3

A discrete Bakry-emery approach

3.1 Introduction

The Bakry-Emery method allows one to establish convex Sobolev inequalities and to compute exponential decay rates towards equilibrium for solutions to diffusion equations [5, 6]. The key idea of Bakry and Emery is to differentiate a so-called entropy functional twice with respect to time and to relate the second-order derivative to the entropy production and in this chapter, we develop a discrete version of this technique.

The study of discrete Bakry-Emery methods and related topics is rather recent. Caputo et al. [19] computed exponential decay rates for time-continuous Markov processes, using the Bochner-Bakry-Emery method. Given a stochastic process with density $u(t)$ and the entropy functional $H_c(u(t))$, the main aim of the Bakry-Emery approach is to find a constant $\lambda > 0$ such that the inequality $d^2H_c/dt^2 \geq -\lambda dH_c/dt$ holds for all times. Integrating this inequality, one may show that $dH_c/dt \leq -\lambda H_c$ which implies that $H_c(u(t)) \leq e^{-\lambda t}H_c(u(0))$ for all $t > 0$, i.e., the entropy decays exponentially along $u(t)$. The relation between d^2H_c/dt^2 and dH_c/dt is achieved in [19] by employing a discrete Bochner-type identity which replaces the Bochner identity of the continuous case. The Bochner-Bakry-Emery method was extended by Fathi and Maas in [45] in the context of Ricci curvature bounds and used by the authors of [65] to derive discrete Beckner inequalities.

Another approach has been suggested by Mielke [75]. He investigated geodesic convexity properties of nonlocal transportation distances on probability spaces such that continuous-time Markov chains can be formulated as gradient flows. Related results have been obtained independently by Chow et al. [29] and Maas [74]. The geodesic convexity property implies exponential decay rates [2]. Mielke showed that the inequality $d^2H_c/dt^2 \geq -\lambda dH_c/dt$ is equivalent

to the positive semi-definiteness of a certain matrix such that matrix algebra can be applied. This idea was extended recently to certain nonlinear Fokker-Planck equations [25].
All these examples involve spatial semi-discretizations of diffusion equations. Temporal semi-discretizations often employ the implicit Euler scheme since it gives entropy dissipation, $dH_c/dt \le 0$, under rather general conditions; see, e.g., the implicit Euler finite-volume approximations in [27, 49]. Entropy-dissipating higher-order semi-discretizations have been analyzed in [39, 63, 64]. However, there seem to be no results for fully discrete schemes using the Bakry-Emery approach. In this thesis, we make a first step to fill this gap.
In order to understand the mathematical difficulty of fully discrete schemes, consider the abstract Cauchy problem

$$\partial_t u + A(u) = 0, \quad t > 0, \quad u(0) = u^0, \tag{3.1}$$

where $A : D(A) \to X'$ is a (nonlinear) operator defined on its domain $D(A) \subset X$ of the Banach space X with dual X'. If the dual product $\langle A(u), H_c'(u)\rangle$ is nonnegative, where $H_c'(u)$ is the (Fréchet) derivative of the entropy and $u(t)$ a solution to (3.1), then

$$\frac{dH_c}{dt} = \langle \partial_t u, H_c'(u)\rangle = -\langle A(u), H_c'(u)\rangle \le 0,$$

showing entropy dissipation. Next, consider the implicit Euler scheme

$$\tau^{-1}(u^k - u^{k-1}) + A_h(u^k) = 0, \quad k \in \mathbb{N}, \ \tau > 0,$$

where u^k is an approximation of $u(k\tau)$ and A_h is an approximation of A still satisfying $\langle A_h(u^k), H'(u^k)\rangle \ge 0$. Here, $H(u^k)$ is the discrete entropy, which is supposed to be convex. Then entropy dissipation is preserved by the scheme since

$$H(u^k) - H(u^{k-1}) \le \langle u^k - u^{k-1}, H'(u^k)\rangle = -\tau\langle A_h(u^k), H'(u^k)\rangle \le 0. \tag{3.2}$$

The problem is to estimate the discrete analog of d^2H_c/dt^2. It turns out that the inequality in (3.2) is too weak, we need an equation; see Section 3.2.1 for details. We overcome this difficulty by developing two ideas.
The *first idea* is to identify the elements which are necessary to build an abstract discrete Bakry-Emery method. Unlike in the continuous case, we distinguish between the discrete entropy production $P := -\tau^{-1}(H(u^k) - H(u^{k-1}))$ and the Fisher information $F := \langle A_h(u^k), H'(u^k)\rangle$. We explain this difference in Section 3.2.

The Bakry-Emery method relies on an estimate of $\tau^{-1}(F(u^k)-F(u^{k-1}))$, which approximates d^2H_c/dt^2. For this estimate, discrete versions of suitable integrations by parts and chain rules are necessary. Our *second idea* is to translate a nonlinear integration-by-parts formula to the discrete case, using the systematic integration-by-parts method of [58]. This leads to a new inequality for numerical three-point schemes as is explained below.

Again, consider first the continuous case. We show in Lemma 37 that for all $(A,B) \in R_c := \{(A,B) \in \mathbb{R}^2 : (2A-B-1)(A+B-2) < 0\}$ and all smooth positive functions w,

$$\int_{\mathbb{T}} w_{xx}(w^A)_{xx} w^B dx \geq \kappa_c \int_{\mathbb{T}} w^{A+B-1} w_{xx}^2 dx,$$

where the constant $\kappa_c > 0$ depends on A and B; see (A.6) below. The proof is based on systematic integration by parts [58]. The discrete counterpart is the following inequality: For any $0 < \varepsilon \leq 1$, there exists a region R of admissible values (A,B), containing the line $A = 1$, such that for all $w_0, \ldots, w_{N+1} \geq 0$ with $w_N = w_0$, $w_{N+1} = w_1$,

$$\begin{aligned}\sum_{i=1}^N &(w_{i+1} - 2w_i + w_{i-1})(w_{i+1}^A - 2w_i^A + w_{i-1}^A) w_i^B \\ &\geq \kappa \sum_{i=1}^N \min_{j=i,i\pm 1} w_j^{A+B-1}(w_{i+1} - 2w_i + w_{i-1})^2, \end{aligned} \tag{3.3}$$

where $\kappa = \varepsilon A$; see Lemma 38. Interestingly, the inequality does not hold for each term independently but only for the sum. The admissible set R for (3.3) is generally smaller than R_c; see Section A. We conjecture that $R = R_c$ for $\kappa = 0$.

Inequality (3.3) is the first nonlinear summation-by-parts formula derived from a systematic method. We believe that this idea will lead to a whole family of new finite-difference inequalities useful in numerical analysis.

We apply the abstract discrete Bakry-Emery method in Section 3.3 to an implicit Euler finite-difference approximation of the porous-medium equation

$$\partial_t u = (u^\beta)_{xx} \quad \text{in } \mathbb{T},\ t > 0, \quad u(0) = u^0 \geq 0,$$

where $\beta > 1$ and $\mathbb{T}$ is the one-dimensional torus. We assume, for simplicity, that $\mathrm{meas}(\mathbb{T}) = 1$ and identify $\mathbb{T}$ with $[0,1]$. The entropy functional is $H_c(u) = \int_{\mathbb{T}}(u^\alpha - \overline{u}^\alpha)dx/(\alpha-1)$, where $\alpha > 0$ and $\overline{u} = \int_{\mathbb{T}} u^0 dx$ is the constant steady state. We show in Proposition 36 that $H_c(u(t))$ decays exponentially to zero for all $(\alpha,\beta) \in S_c$, where

$$S_c = \{(\alpha,\beta) \in \mathbb{R}_+^2 : \alpha+\beta > 1,\ -2 < \alpha - \beta < 1\}, \tag{3.4}$$

with a decay rate depending on (α, β) and $\min_{\mathbb{T}} u^{\beta-1}$.
To overcome the difficulty with the entropy production inequality, we introduce the new variable $v = u^\alpha$ and write the porous-medium equation in the form

$$\partial_t v = \alpha u^{\alpha-1}\partial_t u = \alpha v^{(\alpha-1)/\alpha}(v^{\beta/\alpha})_{xx}. \tag{3.5}$$

The advantage of this formulation is that the entropy becomes *linear* in the variable v. thus avoiding inequality (3.2).
We discretize (3.5) by an implicit Euler finite-difference scheme. Let $\tau > 0$ be the time step, $h > 0$ the space step, and let $v_i^k = (u_i^k)^\alpha$ be an approximation of $(h^{-1}\int_{(i-1)h}^{ih} u(x, k\tau)dx)^\alpha$, $i = 1, \dots, N$. The iterative scheme reads as

$$v_i^k - v_i^{k-1} = \tau h^{-2}\alpha(v_i^k)^{(\alpha-1)/\alpha}\big((v_{i+1}^k)^{\beta/\alpha} - 2(v_i^k)^{\beta/\alpha} + (v_{i-1}^k)^{\beta/\alpha}\big), \tag{3.6}$$

where $i = 1, \dots, N$, $k \in \mathbb{N}$, and $v_N^k = v_0^k$, $v_{N+1}^k = v_1^k$. In Lemma 35 we show the existence of solutions to (3.6) as well as the preservation of nonnegativity. However, the total mass $h\sum_{i=1}^N u_i^k = h\sum_{i=1}^N (v_i^k)^{1/\alpha}$ is not conserved, which is the price to pay for estimating the entropy production. We discuss this point in Section 3.4. This chapter's main result reads as follows.

Theorem 30. *Let $v^k = (v_i^k)$ be a nonnegative solution to (3.6) and set $u_i^k = (v_i^k)^{1/\alpha}$. Let $0 < \varepsilon < 1$. Then there exist a region $S \subset (0, \infty)^2$, containing the line $\alpha - \beta = 1$, and a number $U > 0$ such that all $(\alpha, \beta) \in S$ with $\alpha > 1$ and $\beta \geq 1$. It holds that*

$$\mathcal{H}(u^k) \leq \mathcal{H}(u^0)e^{-\eta\lambda k\tau}, \quad k \in \mathbb{N},$$

where

$$\mathcal{H}(u^k) = \frac{h}{\alpha - 1}\sum_{i=1}^N \big((u_i^k)^\alpha - U^\alpha\big)dx$$

is the discrete (relative) entropy,

$$\eta = \frac{\log(1 + k\tau)}{k\tau}, \quad \lambda = \frac{8\varepsilon(\alpha-1)\beta^2}{C_p(\alpha+\beta-1)^2}\min_{i=1,\dots,N} u_i^{\beta-1},$$

and $C_p = h^2/(4\sin^2(h\pi)) \geq 1/(4\pi^2)$ the the discrete Poincaré constant. Moreover, the total mass $h\sum_{i=1}^N u_i^k$ is increasing in k and converges to U as $k \to \infty$.

Remark 31 (Exponential versus algebraic decay)**.** The exponential decay rate depends on the minimum of the solution, which is not surprising. Indeed, because of the degeneracy, we cannot generally expect exponential decay; an example is the Barenblatt solution. Algebraic decay rates for implicit Euler finite-volume schemes have been derived in, e.g., [27]. When the minimum is positive, the equation is no longer degenerate, and exponential decay follows. □

Remark 32 (Shannon entropy)**.** Unfortunately, the theorem does not apply to the Shannon entropy $h\sum_i u_i \log u_i$, corresponding to $\alpha \to 1$, since $\lambda \to 0$ as $\alpha \to 1$. The reason is that for $\alpha \to 1$, the entropy production P cannot be bounded from above by the Fisher information F and so, Assumption A1 of our abstract Bakry-Emery method does not hold; see Section 3.2.2. $\square$

Remark 33 (Discrete gradient flow)**.** Erbar and Maas [41] showed that the gradient flow of the Shannon entropy with respect to a nonlocal transportation measure equals the discrete porous-medium equation in one space dimension. The porous-medium equation in several space dimensions was solved by Benamou et al. [9] by providing a spatial discretization of this equation as a convex optimization problem. In both references, no decay rates have been derived. $\square$

The set S is illustrated in Figure 3.1 for two different values of ε. Numerical computations indicate that S approaches the set S_c defined in (3.4) if $\varepsilon \to 0$ but for fixed $\varepsilon > 0$, S is strictly contained in S_c.

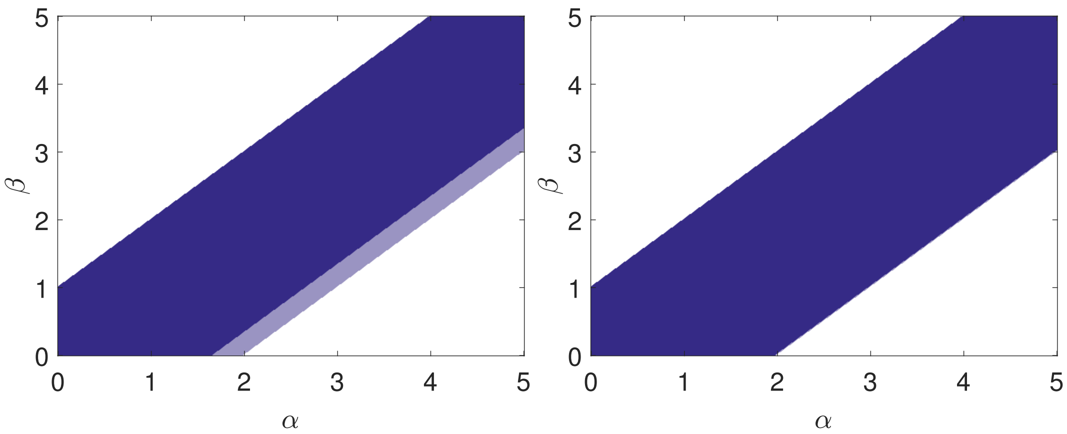

Figure 3.1: Admissible region S for $\varepsilon = 1/4$ (left) and $\varepsilon = 1/100$ (right). The set S_c, defined by $-1 < \alpha - \beta < 2$, is shown in light blue for comparison; it contains the dark blue region S.

This chapter is organized as follows. The abstract Bakry-Emery result is presented in Section 3.2. Theorem 30 is proven in Section 3.3. Numerical examples are presented in Section 3.4, and some auxiliary inequalities and the crucial discrete integration-by-parts estimate can be found in Appendix A and Appendix B.

3.2 An abstract Bakry-Emery method

In this section, we present our abstract result. In order to identify the key ingredients of the Bakry-Emery method, we recall the basic ideas for continuous evolution equations.

3.2.1 The continuous Bakry-Emery method

Let us first consider the abstract Cauchy problem

$$\partial_t u + A(u) = 0, \quad t > 0, \quad u(0) = u^0. \tag{3.7}$$

The nonlinear operator A is defined on some domain $D(A)$ of a Banach space X. We do not specify the properties of A nor its domain since they are not needed in the following. As mentioned in the introduction, the idea of the Bakry-Emery method is to differentiate the entropy functional $H_c : D(A) \to [0, \infty)$ twice with respect to time along solutions to (3.7). We define the entropy production $P_c(u(t)) := -\frac{d}{dt} H_c(u(t))$. If $\langle A(u), H_c'(u)\rangle \geq 0$ holds for all $u \in D(A)$ ($\langle \cdot, \cdot \rangle$ is the dual product in X) then

$$P_c(u) = -\langle \partial_t u, H_c'(u)\rangle = \langle A(u), H_c'(u)\rangle \geq 0,$$

i.e., the entropy production is nonnegative and the entropy is nonincreasing along solutions to (3.7). We call $F_c(u) := \langle A(u), H_c'(u)\rangle$ the generalized Fisher information since if $A(u) = -\Delta u$ on $\mathbb{T}^d$ and $H_c(u) = \int_{\mathbb{T}^d} u(\log u - 1)dx$, we obtain the Fisher information functional $F_c(u) = 4\int_{\mathbb{T}^d} |\nabla \sqrt{u}|^2 dx$. Clearly, $P_c(u(t)) = F_c(u(t))$ along solutions $u(t)$ to (3.7).
Differentiating F_c gives

$$\begin{aligned} \frac{dF_c}{dt} &= \langle A'(u)[\partial_t u], H_c'(u)\rangle + \langle A(u), H_c''(u)\partial_t u\rangle \\ &= -\langle A'(u)[A(u)], H_c'(u)\rangle - \langle A(u), H_c''(u)A(u)\rangle, \end{aligned}$$

where $A'(u)$ is the Fréchet derivative of A at u. If the functional inequality

$$\langle A'(u)[A(u)], H_c'(u)\rangle + \langle A(u), H_c''(u)A(u)\rangle \geq \lambda_c \langle A(u), H_c'(u)\rangle \tag{3.8}$$

holds for some $\lambda_c > 0$ then

$$\frac{dF_c}{dt} \leq -\lambda_c \langle A(u), H_c'(u)\rangle = -\lambda_c F_c, \tag{3.9}$$

and we conclude exponential decay of $t \mapsto F_c(u(t))$ with rate $\lambda_c > 0$. In particular, $\lim_{t\to\infty} F_c(u(t)) = 0$. Then, integrating the previous inequality over (t, ∞), it follows that

$$\frac{dH_c}{dt}(u(t)) = -F_c(u(t)) \leq -\lambda_c \int_t^\infty F_c(u(s))ds = \lambda_c \int_t^\infty \frac{dH_c}{dt}(u(s))ds.$$

Assuming that also

$$\lim_{t\to\infty} H_c(u(t)) = 0, \tag{3.10}$$

we conclude that

$$\frac{dH_c}{dt}(u(t)) \le -\lambda_c H_c(u(t)), \quad t > 0,$$

and by Gronwall's lemma, $t \mapsto H_c(u(t))$ converges exponentially to zero with rate λ_c.
We see that two assumptions are essential: the functional inequality (3.8) and the limit (3.10). On the discrete level, we need to distinguish between the (discrete) entropy production and the (discrete) Fisher information since dH_c/dt and $\langle A(u), H_c'(u)\rangle$ may differ on the discrete level. We assume that both functionals can be estimated by each other. Instead of the functional inequality (3.8) we assume a discrete version of inequality (3.9). Finally, a discrete version of (3.10) is required.

3.2.2 A discrete Bakry-Emery method

We consider two functions $H : \mathbb{R}^N \to [0,\infty)$ and $F : \mathbb{R}^N \to [0,\infty)$ and define $P(v) := P(v;w) = -\tau^{-1}(H(v) - H(w))$, where v, $w \in \mathbb{R}^N$ and $\tau > 0$. We call H an entropy, F the Fisher information, and P the entropy production. Not that the following result is not only applicable to solutions of discrete problems, but holds more generally.

Proposition 34. *Let $(v^k) \subset \mathbb{R}^N$ be any sequence. We assume that*

A1 *There exist C_m, $C_M > 0$ such that $C_m F(v^k) \le P(v^k) \le C_M F(v^k)$ for all $k \in \mathbb{N}$.*

A2 *There exists $\kappa > 0$ such that $F(v^k) - F(v^{k-1}) \le -\tau\kappa F(v^k)$ for all $k \in \mathbb{N}$.*

A3 $\lim_{k\to\infty} H(v^k) = 0$.

Then

$$H(v^k) \le e^{-\eta\lambda k\tau} H(v^0), \quad k \in \mathbb{N},$$

where $\lambda = (C_m/C_M)\kappa$ and $\eta = \log(1+\tau\lambda)/(\tau\lambda) \in (0,1)$.

The discrete decay rate λ is generally smaller than the decay rate κ of the Fisher information, since $\eta < 1$ and we may have $C_m < C_M$. If the entropy production and the Fisher information coincide, i.e. $C_m = C_M = 1$, then $\lambda = \kappa$.

Proof. From assumption A2, it follows that $\lim_{k\to\infty} F(v^k) = 0$. From assumption A2 again and the second inequality in assumption A1, we obtain

$$F(v^k) - F(v^{k-1}) \leq -\tau\kappa F(v^k) \leq -\tau\kappa C_M^{-1} P(v^k) = \kappa C_M^{-1}\big(H(v^k) - H(v^{k-1})\big).$$

Taking the sum from $k = \ell + 1$ to $k = m > \ell + 1$, we find that

$$F(v^m) - F(v^\ell) \leq \kappa C_M^{-1}(H(v^m) - H(v^\ell)).$$

Passing to the limit $m \to \infty$, observing that $\lim_{m\to\infty} F(v^m) = 0$ and, from assumption A3, $\lim_{m\to\infty} H(v^m) = 0$, we deduce that

$$F(v^\ell) \geq \kappa C_M^{-1} H(v^\ell),$$

which holds for all $\ell \in \mathbb{N}$. We now use the first inequality in assumption A1 to conclude that

$$H(v^k)-H(v^{k-1}) = -\tau P(v^k) \leq -\tau C_m F(v^k) \leq -\tau C_m \kappa C_M^{-1} H(v^k) = -\tau\lambda H(v^k).$$

We deduce that $H(v^k) \leq (1 + \lambda\tau)^{-k} H(v^0) = \exp(-\eta\lambda k\tau) H(v^0)$, finishing the proof. $\square$

3.3 Discrete porous-medium equation

To apply the abstract Bakry-Emery method to the porous-medium equation, we need to verify the assumptions of Proposition 34. The key condition is assumption A2. To verify it, we need to translate an integrations-by-parts rule to the discrete level, similarly to 1.5.2. Once again we refer to Appendix A.
We apply the abstract Bakry-Emery method to a finite-difference approximation of the porous-medium equation, i.e., we choose $A(u) = -(u^\beta)_{xx}$ on $\mathbb{T}$ for suitable functions u. Let $\tau > 0$ be the time step and $h > 0$ the space step. A natural scheme would be

$$u_i^k - u_i^{k-1} = \tau h^{-2}\big((u_{i+1}^k)^\beta - 2(u_i^k)^\beta + (u_{i-1}^k)^\beta\big),$$

for all $i = 1, \ldots, N$, $k \in \mathbb{N}$, and $u_N^k = u_0^k$, $u_{N+1}^k = u_1^k$. The corresponding discrete entropy is $H(u^k) = h\sum_{i=1}^N ((u_i^k)^\alpha - \overline{u}^\alpha)/(\alpha - 1)$ and $\overline{u} = h\sum_{i=1}^N u_i^0$ is the constant steady state. We choose $\alpha > 1$ and $\beta > 1$.
Unfortunately, the abstract Bakry-Emery method cannot be applied to this scheme. The problem is the second inequality in assumption A1. Indeed, using the inequality $y^\alpha - z^\alpha \geq \alpha z^{\alpha-1}(y - z)$ for all y, $z \geq 0$, which follows

from the convexity of $z \mapsto z^\alpha$ for $\alpha > 1$, inserting the numerical scheme and then summing by parts, we find that

$$
\begin{aligned}
-\tau P = H(u^k) - H(u^{k-1}) &= h \sum_{i=0}^{N} \big((u_i^k)^\alpha - (u_i^{k-1})^\alpha \big) \\
&\geq \alpha h \sum_{i=1}^{N} (u_i^{k-1})^{\alpha-1} (u_i^k - u_i^{k-1}) \\
&= \alpha h^{-1} \tau \sum_{i=1}^{N} (u_i^{k-1})^{\alpha-1} \Big(\big((u_{i+1}^k)^\beta - (u_i^k)^\beta \big) - \big((u_i^k)^\beta - (u_{i-1}^k)^\beta \big) \Big) \\
&= -\alpha h^{-1} \sum_{i=1}^{N} \big((u_{i+1}^{k-1})^{\alpha-1} - (u_i^{k-1})^{\alpha-1} \big) \big((u_{i+1}^k)^\beta - (u_i^k)^\beta \big).
\end{aligned}
$$

This expression cannot be estimated further; it may even have the wrong sign. We need a scheme that avoids the use of the inequality $y^\alpha - z^\alpha \geq \alpha z^{\alpha-1}(y-z)$. We stress the fact that this problem does not occur in the semi-discrete scheme

$$
\partial_t u_i = h^{-2} \big((u_{i+1}^k)^\beta - 2(u_i^k)^\beta + (u_{i-1}^k)^\beta \big),
$$

since then

$$
\begin{aligned}
\frac{dH}{dt} &= \frac{\alpha h}{\alpha - 1} \sum_{i=0}^{N} u_i^{\alpha-1} \partial_t u_i = \frac{\alpha}{(\alpha-1)h} \sum_{i=0}^{N} u_i^{\alpha-1} \big((u_{i+1}^k)^\beta - 2(u_i^k)^\beta + (u_{i-1}^k)^\beta \big) \\
&= -\frac{\alpha}{(\alpha-1)h} \sum_{i=0}^{N} \big((u_{i+1}^k)^{\alpha-1} - (u_i^k)^{\alpha-1} \big) \big((u_{i+1}^k)^\beta - (u_i^k)^\beta \big),
\end{aligned}
$$

and this expression is nonpositive (since $\alpha > 1$).

Our idea is to make the entropy production *linear* in its argument. To this end, we introduce the new variable $v_i^k = (u_i^k)^\alpha$. In the (continuous) variable $v = u^\alpha$, the evolution equation transforms to $\partial_t v = -v^{(\alpha-1)/\alpha} \Delta(v^{\beta/\alpha})$, which inspires the numerical scheme

$$
v_i^k - v_i^{k-1} = \alpha \tau h^{-2} (v_i^k)^{(\alpha-1)/\alpha} \big((v_{i+1}^k)^{\beta/\alpha} - 2(v_i^k)^{\beta/\alpha} + (v_{i-1}^k)^{\beta/\alpha} \big), \tag{3.11}
$$

for all $i = 1, \ldots, N$, $k \in \mathbb{N}$, and $v_N^k = v_0^k$, $v_{N+1}^k = v_1^k$. The discrete entropy and Fisher information become

$$
H(v^k) = \frac{h}{\alpha - 1} \sum_{i=1}^{N} (v_i^k - V), \quad F(v^k) = \frac{1}{h} \sum_{i=1}^{N} \big((v_{i+1}^k)^\gamma - (v_i^k)^\gamma \big)^2,
$$

where $V > 0$ has to be determined and $\gamma = (\alpha + \beta - 1)/(2\alpha)$.

The entropy production can be estimated, using summation by parts, as

$$
\begin{aligned}
-\tau P(v^k) &= H(v^k) - H(v^{k-1}) = \frac{h}{\alpha-1}\sum_{i=1}^{N}(v_i^k - v_i^{k-1}) \\
&= \frac{\alpha\tau}{(\alpha-1)h}\sum_{i=1}^{N}(v_i^k)^{(\alpha-1)/\alpha}\big((v_{i+1}^k)^{\beta/\alpha} - 2(v_i^k)^{\beta/\alpha} + (v_{i-1}^k)^{\beta/\alpha}\big) \\
&= -\frac{\alpha\tau}{(\alpha-1)h}\sum_{i=1}^{N}\big((v_{i+1}^k)^{(\alpha-1)/\alpha} - (v_i^k)^{(\alpha-1)/\alpha}\big)\big((v_{i+1}^k)^{\beta/\alpha} - (v_i^k)^{\beta/\alpha}\big) \\
&\le 0.
\end{aligned} \tag{3.12}
$$

According to Lemma 44, the entropy production can be estimated from below *and* above in terms of the Fisher information.
After this motivation, we prove the existence of solutions to (3.11).

Lemma 35. *For given $v_i^{k-1} \ge 0$, $i = 1, \ldots, N$, there exists a solution $v_i^k \ge 0$, $i = 1, \ldots, N$, to* (3.11).

Proof. We give only a sketch of the proof since the existence of solutions follows from a standard fixed-point theorem. We only provide the a priori estimates needed for this argument. First multiply (3.11) by $(v_i^k)_- = \min\{v_i^k, 0\}$ and sum over $i = 1, \ldots, N$. Since $v_i^k(v_i^k)_- = (v_i^k)_-^2$, we obtain

$$
\begin{aligned}
\sum_{i=1}^{N}(v_i^k)_-^2 &= \sum_{i=1}^{N} v_i^{k-1}(v_i^k)_- \\
&\quad + \alpha\tau h^{-2}\sum_{i=1}^{N}(v_i^k)_-^{2-1/\alpha}\big((v_{i+1}^k)^{\beta/\alpha} - 2(v_i^k)^{\beta/\alpha} + (v_{i-1}^k)^{\beta/\alpha}\big) \\
&\le \alpha\tau h^{-2}\sum_{i=1}^{N}(v_i^k)_-^{2-1/\alpha}\big(((v_{i+1}^k)^{\beta/\alpha} - (v_i^k)^{\beta/\alpha}) - ((v_i^k)^{\beta/\alpha} - (v_{i-1}^k)^{\beta/\alpha})\big).
\end{aligned}
$$

By summation by parts, this becomes

$$
\sum_{i=1}^{N}(v_i^k)_-^2 \le -\alpha\tau h^{-2}\sum_{i=1}^{N}\big((v_{i+1}^k)_-^{2-1/\alpha} - (v_i^k)_-^{2-1/\alpha}\big)\big((v_{i+1}^k)^{\beta/\alpha} - (v_i^k)^{\beta/\alpha}\big) \le 0,
$$

since $z \mapsto z_-^{2-1/\alpha}$ is nondecreasing. We infer that $(v_i^k)_- = 0$ and hence $v_i^k \ge 0$. Next, by (3.12),

$$
\sum_{i=1}^{N} v_i^k \le \sum_{i=1}^{N} v_i^{k-1} \le \sum_{i=1}^{N} v_i^0,
$$

and this is the desired a priori estimate. □

Next, we turn to the proof of our main result.

Proof of Theorem 30. We verify the assumptions of Proposition 34. For this, we continue our estimates for P. Applying Lemma 44 with $a = (\alpha-1)/\alpha$ and $b = \beta/\alpha$ to (3.12), we obtain the inequalities

$$P(v^k) \le \frac{\alpha}{\alpha-1}F(v^k), \quad F(v^k) \le \frac{\alpha\gamma^2}{\beta}P(v^k) = \frac{(\alpha+\beta-1)^2}{4\alpha\beta}P(v^k).$$

Thus, assumption A1 is satisfied with $C_m = 4\alpha\beta/(\alpha+\beta-1)^2$ and $C_M = \alpha/(\alpha-1)$.

Next, we estimate the difference $F(v^k) - F(v^{k-1})$. To this end, we set $v_i := v_i^k$, $\overline{v}_i := v_i^{k-1}$, $a_i := (v_i - \overline{v}_i)/\tau$ and write

$$\begin{aligned}
F(v) - F(\overline{v}) &= \frac{1}{h}\sum_{i=1}^{N}\big((v_{i+1}^\gamma - v_i^\gamma)^2 - (\overline{v}_{i+1}^\gamma - \overline{v}_i^\gamma)^2\big) \\
&= \frac{1}{h}\sum_{i=1}^{N}\Big((v_{i+1}^\gamma - v_i^\gamma)^2 - \big((v_{i+1} - \tau a_{i+1})^\gamma - (v_i - \tau a_i)^\gamma\big)^2\Big) \\
&=: G(\tau).
\end{aligned}$$

The idea of the proof is to expand $G(\tau)$ around zero:

$$F(v) - F(\overline{v}) = G(0) + G'(0)\tau + \frac{1}{2}G''(\xi)\tau^2$$

for some $\xi \in (0,\tau)$. We show that the right-hand side can be bounded from above by $-\tau K F(v)$ for some $K > 0$, which verifies assumption A2. This idea has first been employed in [27]. Clearly, we have $G(0) = 0$. The first derivatives of G equal

$$\begin{aligned}
G'(\tau) &= 2\gamma h^{-1}\sum_{i=1}^{N}\big((v_{i+1} - \tau a_{i+1})^\gamma - (v_i - \tau a_i)^\gamma\big) \\
&\quad \times \big((v_{i+1} - \tau a_{i+1})^{\gamma-1}a_{i+1} - (v_i - \tau a_i)^{\gamma-1}a_i\big), \\
G''(\tau) &= -2\gamma h^{-1}\sum_{i=1}^{N}\Big(\gamma\big((v_{i+1} - \tau a_{i+1})^{\gamma-1}a_{i+1} - (v_i - \tau a_i)^{\gamma-1}a_i\big)^2 \\
&\qquad + (\gamma-1)\big((v_{i+1} - \tau a_{i+1})^\gamma - (v_i - \tau a_i)^\gamma\big) \\
&\quad \times \big((v_{i+1} - \tau a_{i+1})^{\gamma-2}a_{i+1}^2 - (v_i - \tau a_i)^{\gamma-2}a_i^2\big)\Big).
\end{aligned}$$

First, we claim that $G''(\tau) \le 0$ for any $\tau > 0$. Indeed, we replace $v_i - \tau a_i$ by $\overline{v}_i$ and obtain

$$G''(\tau) = -2\gamma h^{-1}\sum_{i=0}^{N}(c_1 a_{i+1}^2 + c_2 a_{i+1}a_i + c_3 a_i^2), \quad \text{where}$$

$$
\begin{aligned}
c_1 &= \gamma \overline{v}_{i+1}^{2\gamma-2} + (\gamma-1)\overline{v}_{i+1}^{\gamma-2}(\overline{v}_{i+1}^{\gamma} - \overline{v}_i^{\gamma}) = (2\gamma-1)\overline{v}_{i+1}^{2\gamma-2} - (\gamma-1)\overline{v}_{i+1}^{\gamma-2}\overline{v}_i^{\gamma}, \\
c_2 &= -2\gamma \overline{v}_{i+1}^{\gamma-1}\overline{v}_i^{\gamma-1}, \\
c_3 &= \gamma \overline{v}_i^{2\gamma-2} - (\gamma-1)\overline{v}_i^{\gamma-2}(\overline{v}_{i+1}^{\gamma} - \overline{v}_i^{\gamma}) = (2\gamma-1)\overline{v}_i^{2\gamma-2} - (\gamma-1)\overline{v}_i^{\gamma-2}\overline{v}_{i+1}^{\gamma}.
\end{aligned}
$$

It holds that $c_1 \geq 0$, since this inequality is equivalent to $(2\gamma-1)\overline{v}_{i+1}^{\gamma} \geq (\gamma-1)\overline{v}_i^{\gamma}$, and this is true for $1/2 \leq \gamma \leq 1$ (which is equivalent to $\beta \geq 1$ and $\alpha - \beta \geq -1$). Moreover, the discriminant $4c_1c_3 - c_2^2 \geq 0$ is equivalent to

$$
4(2\gamma-1)(1-\gamma)(\overline{v}_{i+1}\overline{v}_i)^{\gamma-2}(\overline{v}_{i+1}^{\gamma} - \overline{v}_i^{\gamma})^2 \geq 0,
$$

which also holds true for $1/2 \leq \gamma \leq 1$. This shows that $G''(\tau) \leq 0$ and consequently,

$$
F(v) - F(\overline{v}) = G(\tau) = G(0) + \tau G'(0) + \frac{\tau^2}{2}G''(\xi) \leq \tau G'(0).
$$

Let us now compute $G'(0)$. Inserting the definition of a_i, we find that

$$
\begin{aligned}
G'(0) &= 2\gamma h^{-1}\sum_{i=1}^{N}(v_{i+1}^{\gamma} - v_i^{\gamma})(v_{i+1}^{\gamma-1}a_{i+1} - v_i^{\gamma-1}a_i) \\
&= -2\gamma h^{-1}\sum_{i=1}^{N} v_i^{\gamma-1}a_i(v_{i+1}^{\gamma} - 2v_i^{\gamma} + v_{i-1}^{\gamma}) \\
&= -2\alpha\gamma h^{-3}\sum_{i=1}^{N} v_i^{(\alpha+\beta-3)/(2\alpha)}(v_{i+1}^{\beta/\alpha} - 2v_i^{\beta/\alpha} + v_{i-1}^{\beta/\alpha})(v_{i+1}^{\gamma} - 2v_i^{\gamma} + v_{i-1}^{\gamma}),
\end{aligned}
$$

since $v_i^{\gamma-1}v_i^{(\alpha-1)/\alpha} = v_i^{(\alpha+\beta-3)/(2\alpha)}$.
We apply Lemma 38 with $w_i = v_i^{\gamma}$, $A = 2\beta/(\alpha+\beta-1)$ and $B = (\alpha+\beta-3)/(\alpha+\beta-1)$ and infer that

$$
G'(0) \leq -2\alpha\gamma\kappa h^{-3}\sum_{i=1}^{N}\min_{j=i,i\pm1} v_j^{(\beta-1)/\alpha}(v_{i+1}^{\gamma} - 2v_i^{\gamma} + v_{i-1}^{\gamma})^2.
$$

By the discrete Poincaré-Wirtinger inequality (Lemma 46), applied with $z_i = v_{i+1}^{\gamma} - v_i^{\gamma}$, it follows that

$$
\begin{aligned}
G'(0) &\leq -2C_p^{-1}\alpha\gamma\kappa h^{-1}\min_{i=1,\ldots,N} v_i^{(\beta-1)/\alpha}\sum_{i=1}^{N}(v_{i+1}^{\gamma} - v_i^{\gamma})^2 \\
&= -2C_p^{-1}\alpha\gamma\kappa\min_{i=1,\ldots,N} v_i^{(\beta-1)/\alpha}F(v),
\end{aligned}
$$

and hence, with $\kappa_0 = 2C_p^{-1}\alpha\gamma\kappa\min_{i=1,\ldots,N} v_i^{(\beta-1)/\alpha}$,

$$
F(v) - F(\overline{v}) \leq -\tau\kappa_0 F(v).
$$

This shows assumption A2 of Proposition 34 and, in particular, after applying Gronwall's lemma, $\lim_{k\to\infty} F(v^k) = 0$.
It remains to prove that assumption A3, i.e. $\lim_{k\to\infty} H(v^k) = 0$, holds. We know that

$$v_i^k \le \sum_{j=1}^N v_j^k \le \sum_{j=1}^N v_j^0 < \infty,$$

so, for any fixed $i = 1, \dots, N$, (v_i^k) is bounded. Therefore, there exists a sequence $k_j \to \infty$ such that $v_i^{k_j} \to y_i$ for some $y_i \ge 0$. By the discrete Poincaré-Wirtinger inequality (Lemma 46), applied to $z_i = (v_i^k)^\gamma - (V^k)^\gamma$, where $(V^k)^\gamma := h\sum_{i=1}^N (v_i^k)^\gamma$, it follows that

$$\sum_{i=1}^N \big((v_i^k)^\gamma - (V^k)^\gamma\big)^2 \le C_p h^{-2} \sum_{i=1}^N \big((v_{i+1}^k)^\gamma - (v_i^k)^\gamma\big)^2 = C_p h^{-1} F(v^k).$$

Since $\lim_{k\to\infty} F(v^k) = 0$, we deduce that $(v_i^{k_j})$ and V^{k_j} have the same limit, say $y := y_i$. Set $U := y^{1/\alpha}$. This defines the entropy

$$\mathcal{H}(u^k) = \frac{h}{\alpha - 1} \sum_{i=1}^N \big((u_i^k)^\alpha - U^\alpha\big)$$

for $u_i^k := (v_i^k)^{1/\alpha}$. It holds that $\mathcal{H}(u^{k_j}) \to 0$ as $k_j \to \infty$. But $k \mapsto \mathcal{H}(u^k)$ is nonincreasing, from which we deduce that $H(v^k) = \mathcal{H}(u^k) \to 0$ for any sequence $k \to \infty$.
According to Proposition 34, the discrete entropy converges exponentially with decay rate

$$\lambda = \frac{C_m}{C_M}\kappa_0 = \frac{4(\alpha-1)\beta}{\alpha+\beta-1}\frac{\kappa}{C_p} \min_{i=1,\dots,N} u_i^{\beta-1} = \frac{8\varepsilon(\alpha-1)\beta^2}{C_p(\alpha+\beta-1)^2} \min_{i=1,\dots,N} u_i^{\beta-1}.$$

Next, we claim that the total mass $h\sum_{i=1}^N u_i^k$ is nondecreasing in k. Indeed, by the concavity of $z \mapsto z^{1/\alpha}$ (recall that $\alpha > 1$), we have $y^{1/\alpha} - z^{1/\alpha} \ge (1/\alpha) y^{(1-\alpha)/\alpha}(y - z)$ for all y, $z \ge 0$ and hence,

$$\sum_{i=1}^N (u_i^k - u_i^{k-1}) = \sum_{i=1}^N \big((v_i^k)^{1/\alpha} - (v_i^{k-1})^{1/\alpha}\big) \ge \frac{1}{\alpha}\sum_{i=1}^N (v_i^k)^{(1-\alpha)/\alpha}(v_i^k - v_i^{k-1}).$$

Inserting scheme (3.6), we find that

$$\sum_{i=1}^N (u_i^k - u_i^{k-1}) \ge \frac{\tau}{h^2}\sum_{i=1}^N \big((v_{i+1}^k)^{\beta/\alpha} - 2(v_i^k)^{\beta/\alpha} + (v_{i-1}^k)^{\beta/\alpha}\big) = 0,$$

since v_i^k satisfies periodic boundary conditions. This shows the claim.
The monotonicity of the total mass and the convergence property $h\sum_{i=1}^N u_i^{k_j} \to y^{1/\alpha} = U$ as $k_j \to \infty$ imply that $h\sum_{i=1}^N u_i^k \to U$ for $k \to \infty$, and the convergence is monotone. This finishes the proof. $\square$

3.4 Numerical examples

We present some numerical results for the porous-medium equation discretized in the previous section. Similarly to the previous chapters, we choose the Barenblatt profile as initial datum

$$u^0(x) = \frac{1}{t_0^{1/(\beta+1)}}\left(C - \frac{\beta-1}{2\beta}\frac{|x-x_0|^2}{t_0^{2/(\beta+1)}}\right)_+^{1/(\beta-1)},$$

where $z_+ = \max\{0,z\}$. We consider two cases. For the slow diffusion case $\beta = 4$, we choose $x_0 = 0.5$, $t_0 = 10^{-4}$, and

$$C = \frac{\beta-1}{2\beta}\frac{|x_0|^2}{(t_{\mathrm{end}}+t_0)^{2/(\beta+1)}}, \quad t_{\mathrm{end}} = 5\cdot 10^{-4}.$$

The profile reaches the boundary of $\Omega = (0,1)$ at time t_{end}. For the fast diffusion case $\beta = 0.5$, we take $x_0 = 0.5$, $t_0 = 10^{-2}$, and $C = t_0^{(\beta-1)/(\beta+1)}$ such that the maximum of the initial profile equals 1.
Figure 3.2 illustrates the evolution of the total mass for $\alpha = 2$, $\beta = 0.5$ (left) and $\alpha = 3$, $\beta = 4$ (right). As predicted in Theorem 30, the total mass is indeed increasing in time. The mass defect scales well with both the time step τ and the grid size h, where the influence of τ is more prevalent.

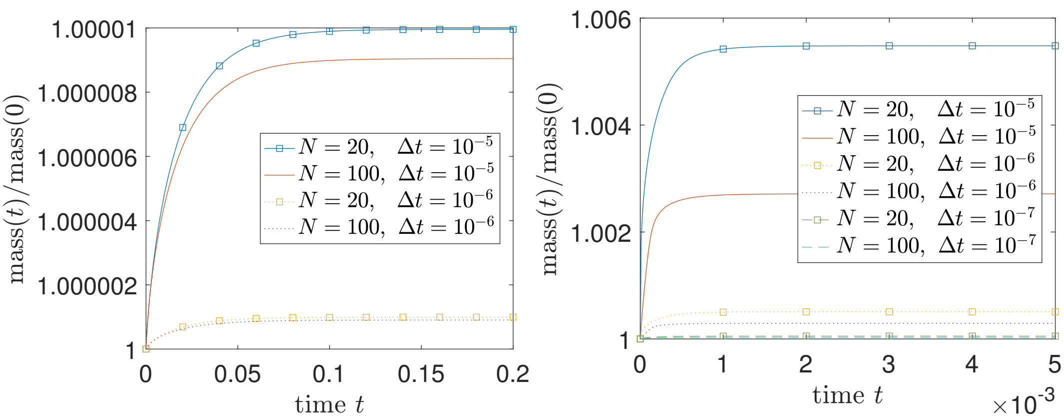

Figure 3.2: Evolution of the total mass for two test scenarios: $\alpha = 2$, $\beta = 0.5$ (left) and $\alpha = 3$, $\beta = 4$ (right).

The time decay of the (relative) entropy $\mathcal{H}$ is shown in Figure 3.3 for various space and time steps. We observe that the decay is indeed exponential. Here, the steady state u_∞ (which is needed to define the relative entropy) is given by $u_\infty = h\sum_{i=0}^N u_i^{k_{\max}}$, where $k_{\max}$ is the final time step. This choice clearly depends on the scheme since the mass is not conserved. The relative entropy converges exponentially even when (α,β) is chosen outside of the admissible region; see Figure 3.4.

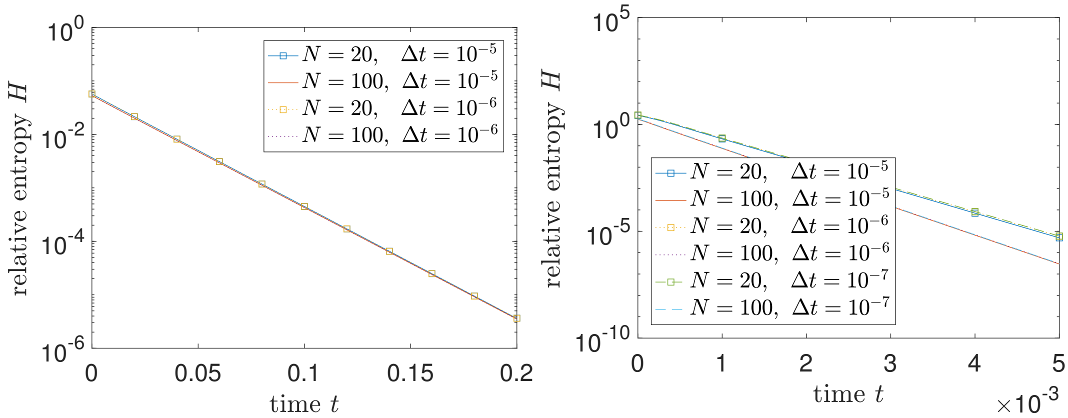

Figure 3.3: Evolution of the relative entropy for two test scenarios in the admissible region: $\alpha = 2$, $\beta = 0.5$ (left) and $\alpha = 3$, $\beta = 4$ (right).

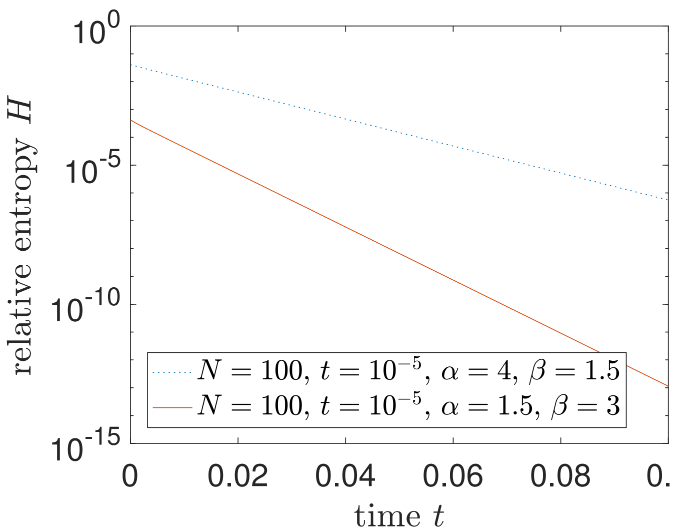

Figure 3.4: Evolution of the relative entropies for (α, β) outside of the admissible region.

Appendix A

A nonlinear discrete integration-by-parts formula

In this section we develop a nonlinear discrete integration-by-parts formula. Since it is convenient to investigate the continuous situation first. Hence, we start by looking at the nonlinear diffusion equation

$$\partial_t u = (u^\beta)_{xx}, \quad t > 0, \quad u(0) = u^0 \geq 0 \quad \text{in } \mathbb{T}, \tag{A.1}$$

where $\beta > 0$, and introduce the (relative) entropy

$$H_c(u) = \frac{1}{\alpha - 1}\int_{\mathbb{T}} (u^\alpha - \overline{u}^\alpha)dx, \quad \alpha > 0.$$

Here, $\overline{u} = \int_{\mathbb{T}} u^0(x)dx$ is the constant steady state. (Recall that $\operatorname{meas}(\mathbb{T}) = 1$.)

Proposition 36. *Let $\beta \neq 1$, $\alpha + \beta - 1 > 0$, and $-1 < \alpha - \beta < 2$. Then, for any positive smooth solution to* (A.1)*,*

$$H_c(u(t)) \leq H_c(u^0)e^{-\lambda_c t}, \quad t > 0,$$

where

$$\lambda_c = \frac{16\pi^2\alpha\beta\kappa_c}{\alpha + \beta - 1}\min_{\mathbb{T}} u^{\beta - 1} \geq 0, \quad \kappa_c = -\frac{4\beta(\alpha - \beta - 2)}{(\alpha + \beta - 1)(\alpha - \beta + 1)} > 0.$$

Proof. Integrating by parts, the time derivatives of $H_c(u(t))$ become

$$\begin{aligned}
\frac{dH_c}{dt} &= \frac{\alpha}{\alpha - 1}\int_{\mathbb{T}} u^{\alpha - 1}(u^\beta)_{xx}dx = -\frac{4\alpha\beta}{(\alpha + \beta - 1)^2}\int_{\mathbb{T}} \big(u^{(\alpha+\beta-1)/2}\big)_x^2 dx, \\
\frac{d^2H_c}{dt^2} &= -\frac{8\alpha\beta}{(\alpha + \beta - 1)^2}\int_{\mathbb{T}} \big(u^{(\alpha+\beta-1)/2}\big)_x \left(\frac{\alpha + \beta - 1}{2} u^{(\alpha+\beta-3)/2}\partial_t u\right)_x dx \\
&= \frac{4\alpha\beta}{\alpha + \beta - 1}\int_{\mathbb{T}} u^{(\alpha+\beta-3)/2}\big(u^{(\alpha+\beta-1)/2}\big)_{xx}(u^\beta)_{xx}dx.
\end{aligned}$$

We wish to estimate the second time derivative. To this end, we set $w = u^{(\alpha+\beta-1)/2}$, $A = 2\beta/(\alpha+\beta-1)$, and $B = (\alpha+\beta-3)/(\alpha+\beta-1)$. Then the derivatives can be written as

$$\frac{dH_c}{dt} = -\frac{4\alpha\beta}{(\alpha+\beta-1)^2}\int_{\mathbb{T}} w_x^2 dx, \quad \frac{d^2H_c}{dt^2} = \frac{4\alpha\beta}{\alpha+\beta-1}\int_{\mathbb{T}}(w^A)_{xx}w_{xx}w^B dx. \tag{A.2}$$

In Lemma 37 below we show that there exists $\kappa_c > 0$ such that

$$\int_{\mathbb{T}}(w^A)_{xx}w_{xx}w^B dx \geq \kappa_c \int_{\mathbb{T}} w^{A+B-1}w_{xx}^2 dx$$

if the assumption $(2A-B-1)(A+B-2) < 0$ holds. (Note that $\beta \neq 1$ is equivalent to $A+B-2 \neq 0$.) This condition is actually satisfied since

$$(2A-B-1)(A+B-2) = \frac{2(\alpha-\beta-2)(\alpha-\beta+1)}{(\alpha+\beta-1)^2} < 0,$$

and we infer that

$$\frac{d^2H_c}{dt^2} \geq \frac{4\alpha\beta\kappa_c}{\alpha+\beta-1}\int_{\mathbb{T}} w^{A+B-1}w_{xx}^2 dx = \frac{4\alpha\beta\kappa_c}{\alpha+\beta-1}\int_{\mathbb{T}} u^{\beta-1}\big(u^{(\alpha+\beta-1)/2}\big)_{xx}^2 dx.$$

Furthermore, by the Poincaré inequality applied to w_x (see Lemma 45),

$$\begin{aligned}\int_{\mathbb{T}} u^{\beta-1}\big(u^{(\alpha+\beta-1)/2}\big)_{xx}^2 dx &\geq \min_{\mathbb{T}} u^{\beta-1}\int_{\mathbb{T}}\big(u^{(\alpha+\beta-1)/2}\big)_{xx}^2 dx \\ &\geq 4\pi^2 \min_{\mathbb{T}} u^{\beta-1}\int_{\mathbb{T}}\big(u^{(\alpha+\beta-1)/2}\big)_{x}^2 dx = 4\pi^2 \min_{\mathbb{T}} u^{\beta-1}\int_{\mathbb{T}} w_x^2 dx,\end{aligned}$$

and it follows that

$$\frac{d^2H_c}{dt^2} \geq \frac{16\pi^2\alpha\beta\kappa_c}{\alpha+\beta-1}\min_{\mathbb{T}} u^{\beta-1}\int_{\mathbb{T}} w_x^2 dx = -\lambda_c\frac{dH_c}{dt}. \tag{A.3}$$

Denoting by $P_c = -dH_c/dt$ the entropy production, this inequality can be formulated as $dP_c/dt \leq -\lambda_c P_c$. Gronwall's lemma then implies that $P_c(u(t)) \leq P_c(u_0)e^{-\lambda_c t}$ for $t > 0$ and in particular $\lim_{t\to\infty} P_c(u(t)) = 0$.
Integrating (A.3) over (t,s) with $t < s$ and passing to the limit $s \to \infty$, we see that

$$\int_0^\infty \int_{\mathbb{T}} w_x^2 dx \leq H_c(u_0) < \infty.$$

Thus, there exists a sequence $t_j \to \infty$ such that $\|w_x(t_j)\|_{L^2(\mathbb{T})} \to 0$. Following the arguments of [20, Prop. 1ii],[1] it follows that $\lim_{t_j\to\infty} H_c(u(t_j)) = 0$, and since $t \mapsto H_c(u(t))$ is nonincreasing, any sequence converges, it follows that $\lim_{t\to\infty} H_c(u(t)) = 0$.

[1] Also see the erratum www.asc.tuwien.ac.at/∼juengel/publications/pdf/errata05carri.pdf.

Therefore, we proceed by integrating inequality (A.3) over (t,∞) (and using $\lim_{t\to\infty}(dH_c/dt)(u(t)) = \lim_{t\to\infty} H_c(u(t)) = 0$ from above):

$$-\frac{dH_c}{dt}(u(t)) \geq \lambda_c H_c(u(t)), \quad t > 0.$$

Thus, another application of Gronwall's lemma gives the conclusion. $\square$

It remains to prove Lemma 37. Set

$$R_c = \{(A,B) \in \mathbb{R}^2 : A > 0,\ (2A-B-1)(A+B-2) < 0\}. \tag{A.4}$$

Lemma 37. *Let $(A,B) \in R_c$. Then for all smooth positive functions w,*

$$\int_{\mathbb{T}} w_{xx}(w^A)_{xx} w^B dx \geq \kappa_c \int_{\mathbb{T}} w_{xx}^2 w^{A+B-1} dx, \tag{A.5}$$

where

$$\kappa_c = \begin{cases} -A(2A-B-1)/(A+B-2) > 0 & \text{if } A+B-2 \neq 0, \\ A & \text{if } A+B-2 = 0. \end{cases} \tag{A.6}$$

Proof. The idea of the proof is to employ systematic integration by parts [58]. Since

$$\int_{\mathbb{T}} (w_x^3 w^{A+B-2})_x dx = 0, \tag{A.7}$$

we can formulate (A.5) as the following problem: Find $c \in \mathbb{R}$ and $\kappa_c > 0$ such that for all smooth positive functions w,

$$\int_{\mathbb{T}} \big(w_{xx}(w^A)_{xx} w^B + c(w_x^3 w^{A+B-3})_x - \kappa_c w_{xx}^2 w^{A+B-1}\big) dx \geq 0.$$

Calculating the derivatives and setting $\xi_1 = w_x/w$, $\xi_2 = w_{xx}/w$, this inequality is equivalent to

$$\int_{\mathbb{T}} w^{A+B-1}\Big((A-\kappa_c)\xi_2^2 + (A^2 - A + 3c)\xi_2\xi_1^2 + c(A+B-2)\xi_1^4\Big) dx \geq 0.$$

The idea is to interpret the integrand as a polynomial in the variables ξ_1, ξ_2 and to solve the following polynomial decision problem: Find $c \in \mathbb{R}$ and $\kappa_c > 0$ such that for all $(\xi_1, \xi_2) \in \mathbb{R}^2$,

$$(A-\kappa_c)\xi_2^2 + (A^2 - A + 3c)\xi_2\xi_1^2 + c(A+B-2)\xi_1^4 \geq 0. \tag{A.8}$$

This problem can be solved explicitly. Clearly, it must hold that $A \geq \kappa_c > 0$. We distinguish two cases: $\kappa_c = A$ and $\kappa_c < A$.

First let $0 < \kappa_c < A$. Then (A.8) is valid if the discriminant is nonpositive,

$$\begin{aligned}
0 &\geq (A^2 - A + 3c)^2 - 4c(A - \kappa_c)(A + B - 2) \\
&= \Big(3c + A(A-1) - \frac{2}{3}(A - \kappa_c)(A + B - 2)\Big)^2 \\
&\quad - \frac{4}{9}(A - \kappa_c)^2(A + B - 2)^2 + \frac{4}{3}A(A-1)(A - \kappa_c)(A + B - 2).
\end{aligned}$$

Choosing the minimizing value

$$\begin{aligned}
c &= -\frac{1}{3}\Big(A(A-1) - \frac{2}{3}(A - \kappa_c)(A + B - 2)\Big) \\
&= -\frac{A}{9}(A - 2B + 1) - \frac{2}{9}\kappa_c(A + B - 2),
\end{aligned} \tag{A.9}$$

the discriminant is nonpositive if and only if

$$\begin{aligned}
0 &\geq -\frac{4}{9}(A - \kappa_c)^2(A + B - 2)^2 + \frac{4}{3}A(A-1)(A - \kappa_c)(A + B - 2) \\
&= \frac{4}{9}(A - \kappa_c)(A + B - 2)\big(\kappa_c(A + B - 2) + A(2A - B - 1)\big).
\end{aligned}$$

Set $\kappa_c = \varepsilon A$ for $0 < \varepsilon < 1$. Then the previous inequality is true if and only if

$$A(A + B - 2)\big(\varepsilon(A + B - 2) + 2A - B - 1\big) \leq 0. \tag{A.10}$$

We infer that if

$$\varepsilon = -\frac{2A - B - 1}{A + B - 2} > 0$$

then (A.8) holds. This implies that $\kappa_c = \varepsilon A = -A(2A-B-1)/(A+B-2) > 0$ and we need to choose $A > 0$ and $(2A - B - 1)(A + B - 2) < 0$.
Next, let $\kappa_c = A$. Then the quadratic term in ξ_2 in (A.8) vanishes and the mixed term must vanish too, i.e. $c = -A(A-1)/3$. Hence, the coefficient of the remaining term has to be nonnegative, i.e. $-A(A-1)(A+B-2) \geq 0$. If $A = 1$, inequality (A.5) becomes trivial. The set of all (A, B) such that $A > 0$ and $(A-1)(A+B-2) < 0$ is contained in the set of all (A, B) satisfying $A > 0$ and $(2A - B - 1)(A + B - 2) < 0$. This finishes the proof. $\square$

We now state a discrete version of inequality (A.5).

Lemma 38. *Let $w_0, \dots, w_{N+1} \in \mathbb{R}$ satisfy $w_N = w_0$, $w_{N+1} = w_1$ and let $0 < \varepsilon \le 1$. There exists a region $R \subset \mathbb{R}^2$, containing the line $A = 1$, such that for all $(A, B) \in R$,*

$$\sum_{i=1}^{N} (w_{i+1} - 2w_i + w_{i-1})(w_{i+1}^A - 2w_i^A + w_{i-1}^A) w_i^B$$
$$\ge \kappa \sum_{i=1}^{N} \min_{j=i,i\pm 1} w_j^{A+B-1} (w_{i+1} - 2w_i + w_{i-1})^2, \tag{A.11}$$

where $\kappa = \varepsilon A > 0$.

As stated above the lemma is trivial since (A.11) clearly holds for $R = \{(A, B) : A = 1\}$ with $\kappa = 1$. Figure A.1 illustrates the numerically admissible regions for (A, B) for two different values of ε. The admissible region R is smaller than the region R_c for the continuous case but it approaches the latter region as $\kappa \to 0$.

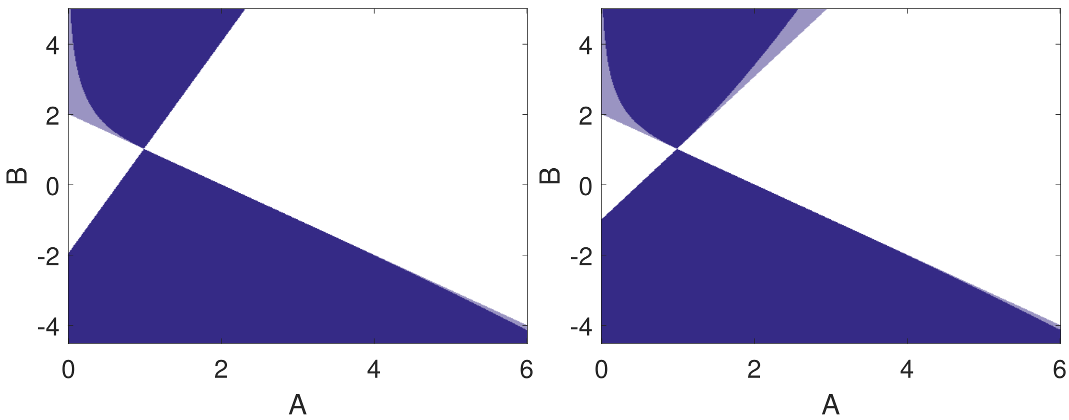

Figure A.1: The regions of admissible (A, B) such that $T(X, Y) \ge 0$ for all X, $Y \ge 0$ using c as in (A.9) with $\kappa_c = \kappa$ and $\kappa = A/4$ (left), $\kappa = A/100$ (right). The set R is depicted in dark blue, $R_c \supset R$ in light blue.

The idea of the proof of (A.11) is to add the following discrete version of the integration-by-parts formula (A.7)

$$\frac{1}{\rho^3} \sum_{i=1}^{N} \Big(M(w_{i+1}, w_i)^{A+B+1-3\rho} (w_{i+1}^\rho - w_i^\rho)^3 - M(w_i, w_{i-1})^{A+B+1-3\rho} (w_i^\rho - w_{i-1}^\rho)^3 \Big).$$

The sum vanishes because of the periodic boundary conditions. Here $\rho > 0$ is a free parameter, and the function $M(x, y)$ is a symmetric mean value, i.e., it satisfies

$$M(x, y) = M(y, x), \quad M(\lambda x, \lambda y) = \lambda M(x, y), \quad M(x, x) = x \tag{A.12}$$

for all x, y, $\lambda \geq 0$. For the numerical simulations below, we choose $\rho = (A+B+1)/3$ such that the mean function does not need to be specified. Then (A.11) holds if we can show the following inequality for all admissible (A, B) and $w_i \neq 0$:

$$\sum_{i=1}^{N} w_i^{A+B+1} \Bigg\{ \bigg(\Big(\frac{w_{i+1}}{w_i}\Big)^A + \Big(\frac{w_{i-1}}{w_i}\Big)^A - 2 \bigg) \bigg(\frac{w_{i+1}}{w_i} + \frac{w_{i-1}}{w_i} - 2 \bigg)$$
$$- \kappa \min_{j=i,i\pm1} \Big(\frac{w_j}{w_i}\Big)^{A+B-1} \bigg(\frac{w_{i+1}}{w_i} + \frac{w_{i-1}}{w_i} - 2 \bigg)^2$$
$$+ \frac{c}{\rho^3} \bigg(M\Big(\frac{w_{i+1}}{w_i}, 1\Big)^{A+B+1-3\rho} \bigg(\Big(\frac{w_{i+1}}{w_i}\Big)^\rho - 1 \bigg)^3$$
$$- M\Big(\frac{w_{i-1}}{w_i}, 1\Big)^{A+B+1-3\rho} \bigg(1 - \Big(\frac{w_{i-1}}{w_i}\Big)^\rho \bigg)^3 \bigg) \Bigg\} \geq 0.$$

We verify this inequality pointwise, i.e. setting $X = w_{i+1}/w_i$ and $Y = w_{i-1}/w_i$, we wish to find $c \in \mathbb{R}$, $\kappa > 0$ such that for all X, $Y > 0$,

$$\begin{aligned} T(X,Y) := {} & (X^A + Y^A - 2)(X + Y - 2) \\ & + \frac{c}{\rho^3} \Big(M(X,1)^{A+B+1-3\rho}(X^\rho - 1)^3 + M(Y,1)^{A+B+1-3\rho}(Y^\rho - 1)^3 \Big) \\ & - \kappa \min\{1, X^{A+B-1}, Y^{A+B-1}\}(X + Y - 2)^2 \geq 0. \end{aligned} \tag{A.13}$$

The first term $(X^A + Y^A - 2)(X + Y - 2)$ becomes negative in certain regions; see Figure A.2. It is compensated by the second term (shift term) on the right-hand side of (A.13) if we choose the constant c according to (A.9) with $\kappa = \kappa_c$ as in (A.6).

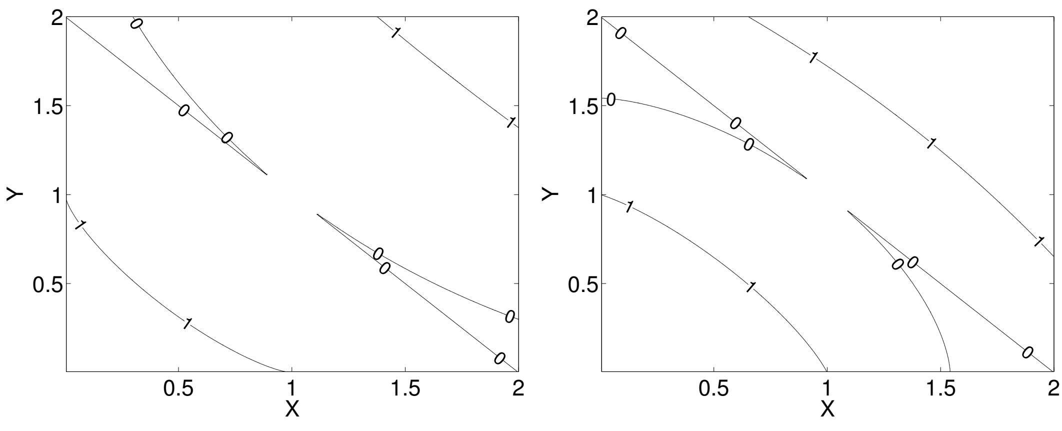

Figure A.2: Level sets $(X^A + Y^A - 2)(X + Y - 2) = 0$ and $(X^A + Y^A - 2)(X + Y - 2) = 1$ for $A = 0.6$, $B = 4$ (left) and $A = 1.6$, $B = 2.5$ (right). We have chosen $\kappa = \kappa_0 = A/200$ and c as in (A.9).

Unfortunately, it is rather difficult to prove (A.13) analytically in full generality. Note that polynomial quantifier elimination does not apply if A and B are non-integers, since the function $T(X,Y)$ generally is *not* a polynomial. Instead, we verify (A.13) analytically for all $(A,B) \in R_c$ and all (X,Y) in some neighborhood of $(1,1)$.

Lemma 39. *Let T be given by* (A.13) *and let $(A,B) \in R_c$, where R_c is defined in* (A.4)*. Then there exists a neighborhood W of $(1,1)$ such that for all $(X,Y) \in W$,*

$$T(X,Y) \geq 0$$

holds for c as in (A.9) *and with $\kappa_c = \kappa$ as in* (A.6)*.*

If the step size $h > 0$ is small enough, we expect that the quotients w_{i+1}/w_i are close to one for all $i = 0,\dots,N-1$. This means that (X,Y) lies in a neighborhood of $(1,1)$, and the lemma applies.

Proof. We use the local coordinates $u = (X+Y-2)/h^2$ and $v = (X-Y)/(2h)$, which correspond to (central) second-order and first-order derivatives. Then $X = 1+hv+h^2u/2$ and $Y = 1-hv+h^2u/2$. We develop T as a function of h at $h=0$. For this, we observe that $M(1,1) = 1$ and $M_X(1,1) = M_Y(1,1) = 1/2$. Indeed, we infer from the properties (A.12) that

$$\begin{aligned}
M_X(1,1) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\big(M(1+\varepsilon,1) - M(1,1)\big) = \lim_{\varepsilon\to 0}\left(\frac{1+\varepsilon}{\varepsilon} M\left(1, \frac{1}{1+\varepsilon}\right) - \frac{1}{\varepsilon}\right) \\
&= \lim_{\varepsilon\to 0}\left(\frac{1+\varepsilon}{\varepsilon} M\left(1, 1-\frac{\varepsilon}{1+\varepsilon}\right) - \frac{1+\varepsilon}{\varepsilon} + 1\right) \\
&= \lim_{\varepsilon\to 0}\left(\frac{1+\varepsilon}{\varepsilon} M\left(1-\frac{\varepsilon}{1+\varepsilon}, 1\right) - \frac{1+\varepsilon}{\varepsilon} + 1\right) \\
&= \lim_{\delta\to 0} \frac{1}{\delta}\big(M(1-\delta,1) - M(1,1)\big) + 1 \\
&= -M_X(1,1) + 1,
\end{aligned}$$

which implies that $M_X(1,1) = 1/2$.
Calculating the Taylor series of the terms in T with respect to h at $h=0$ leads to

$$\begin{aligned}
&(X^A + Y^A - 2)(X+Y-2) = Au\big((A-1)v^2 + u\big)h^4 + O(h^6), \\
&\frac{c}{\rho^3} M(X,1)^{A+B+1-3\rho}(X^\rho - 1)^3 = cv^3h^3 + \frac{c}{2}\big((A+B-2)v^2 + 3u\big)v^2h^4 + O(h^5), \\
&\frac{c}{\rho^3} M(Y,1)^{A+B+1-3\rho}(Y^\rho - 1)^3 = -cv^3h^3 + \frac{c}{2}\big((A+B-2)v^2 + 3u\big)v^2h^4 + O(h^5).
\end{aligned}$$

In particular, as expected, the explicit choices of both ρ and $M(x,y)$ do not change the behavior of the shift term locally around the equilibrium $w_{i-1} =$

$w_i = w_{i+1}$ or $h = 0$. Moreover, $\min\{1, X^{A+B-1}, Y^{A+B-1}\} = 1 + O(h)$ and $(X + Y - 2)^2 = u^2 h^4$. Combining these expressions gives

$$T(X, Y) = h^4\Big((A - \kappa)u^2 + \big(A(A-1) + 3c\big)uv^2 + c(A + B - 2)v^4\Big) + O(h^5).$$

The polynomial

$$(u, v) \mapsto (A - \kappa)u^2 + \big(A(A-1) + 3c\big)uv^2 + c(A + B - 2)v^4$$

is the same as in (A.8). The proof of Lemma 37 shows that it is nonnegative for all $(A, B) \in R_c$ with c as in (A.9) and κ_c as in (A.6). We deduce that $T(X, Y) \geq 0$ holds for all $(A, B) \in R_c$ if $h \in \mathbb{R}$ is sufficiently small. This proves the lemma. $\square$

Appendix B

Auxiliary inequalities

Discrete Gronwall lemmas

First, we prove a rather general discrete nonlinear Gronwall lemma.

Lemma 40 (Discrete nonlinear Gronwall lemma). *Let $f \in C^1([0,\infty))$ be a positive, nondecreasing, and convex function such that $1/f$ is locally integrable. Define*

$$w(x) = \int_1^x \frac{dz}{f(z)}, \quad x \geq 0.$$

Let (x_n) be a sequence of nonnegative numbers such that $x_{n+1}-x_n+f(x_{n+1}) \leq 0$ for $n \in \mathbb{N}_0$. Then

$$x_n \leq w^{-1}\left(w(x_0) - \frac{n}{1+f'(x_0)}\right), \quad n \in \mathbb{N}.$$

Note that the function w is strictly increasing such that its inverse is well defined.

Proof. Since f is nondecreasing and (x_n) is nonincreasing, we obtain

$$w(x_{n+1}) - w(x_n) = \int_{x_n}^{x_{n+1}} \frac{dz}{f(z)} \leq \frac{x_{n+1}-x_n}{f(x_n)}.$$

The sequence (x_n) satisfies $f(x_{n+1})/(x_{n+1}-x_n) \geq -1$. Therefore,

$$\begin{aligned} w(x_{n+1}) - w(x_n) &\leq \Big(\frac{f(x_{n+1})}{x_{n+1}-x_n} + \frac{f(x_n)-f(x_{n+1})}{x_{n+1}-x_n}\Big)^{-1} \\ &\leq \Big(-1 - \frac{f(x_n)-f(x_{n+1})}{x_n - x_{n+1}}\Big)^{-1}. \end{aligned}$$

From the convexity of f, it follows that $f(x_n)-f(x_{n+1}) \leq f'(x_n)(x_n-x_{n+1}) \leq f'(x_0)(x_n - x_{n+1})$, which implies that

$$w(x_{n+1}) - w(x_n) \leq (-1 - f'(x_0))^{-1}.$$

Summing this inequality from $n = 0$ to $N-1$, where $N \in \mathbb{N}$, yields

$$w(x_N) \leq w(x_0) - \frac{N}{1+f'(x_0)}.$$

Applying the inverse function of w shows the lemma. $\square$

The choice $f(x) = \tau K x^\gamma$ for some $\gamma > 1$ in Lemma 40 leads to the following result.

Corollary 41. *Let (x_n) be a sequence of nonnegative numbers satisfying*

$$x_{n+1} - x_n + \tau x_{n+1}^\gamma \leq 0, \quad n \in \mathbb{N},$$

where $K > 0$ and $\gamma > 1$. Then

$$x_n \leq \frac{1}{\left(x_0^{1-\gamma} + c\tau n\right)^{1/(\gamma-1)}}, \quad n \in \mathbb{N},$$

where $c = (\gamma-1)/(1+\gamma\tau x_0^{\gamma-1})$.

Discrete gradient inequalities

We show some inequalities in two variables.

Lemma 42. *Let α, $\beta > 0$. Then, for all x, $y \geq 0$,*

$$(y^\alpha - x^\alpha)(y^\beta - x^\beta) \geq \frac{4\alpha\beta}{(\alpha+\beta)^2}(y^{(\alpha+\beta)/2} - x^{(\alpha+\beta)/2})^2. \tag{B.1}$$

Proof. If $y = 0$, inequality (B.1) holds. Let $y \neq 0$ and set $z = (x/y)^\beta$. Then the inequality is proven if for all $z \geq 0$,

$$f(z) = (1-z^{\alpha/\beta})(1-z) - \frac{4\alpha\beta}{(\alpha+\beta)^2}(1-z^{(\alpha+\beta)/2\beta})^2 \geq 0.$$

We differentiate f twice:

$$f'(z) = -1 - \frac{\alpha}{\beta}z^{\alpha/\beta-1} + \frac{(\alpha-\beta)^2}{\beta(\alpha+\beta)}z^{\alpha/\beta} + \frac{4\alpha}{\alpha+\beta}z^{(\alpha+\beta)/2\beta},$$

$$f''(z) = \frac{\alpha(\alpha-\beta)}{\beta} z^{\alpha/2\beta-3/2}\Big(-\frac{1}{\beta}z^{\alpha/2\beta-1/2} + \frac{\alpha-\beta}{\beta(\alpha+\beta)}z^{\alpha/2\beta+1/2} + \frac{2}{\alpha+\beta}\Big).$$

Then $f(1) = 0$ and $f'(1) = 0$. Thus, if we show that f is convex, the assertion follows. In order to prove the convexity of f, we define

$$g(z) = -\frac{1}{\beta}z^{\alpha/2\beta-1/2} + \frac{\alpha-\beta}{\beta(\alpha+\beta)}z^{\alpha/2\beta+1/2} + \frac{2}{\alpha+\beta}.$$

Then $g(1) = 0$ and it holds

$$g'(z) = \frac{\alpha-\beta}{2\beta^2}z^{\alpha/2\beta-3/2}(-1+z),$$

and therefore, $g'(1) = 0$. Now, if $\alpha > \beta$, $g(0) = 2/(\alpha+\beta) > 0$, and g is decreasing in $[0, 1]$ and increasing in $[1, \infty)$. Thus, $g(z) \geq 0$ for all $z \geq 0$. If $\alpha < \beta$ then $g(0+) = -\infty$, and g is increasing in $[0, 1]$ and decreasing in $[1, \infty)$. Hence, $g(z) \leq 0$ for $z \geq 0$. Independent of the sign of $\alpha - \beta$, we obtain

$$f''(z) = \frac{\alpha(\alpha-\beta)}{\beta}z^{\alpha/2\beta-3/2}g(z) \geq 0$$

for all $z \geq 0$, which shows the convexity of f. □

Corollary 43. *Let α, $\beta > 0$. Then, for all x, $y \geq 0$,*

$$(y^\beta - x^\beta)(y^\alpha - x^\alpha) \geq \frac{4\alpha\beta}{(\alpha+1)^2}\min\{x^{\beta-1}, y^{\beta-1}\}(y^{(\alpha+1)/2} - x^{(\alpha+1)/2})^2.$$

Proof. We assume without restriction that $y > x$. Then we apply Lemma 42 to $\beta = 1$:

$$\begin{aligned}(y^\beta - x^\beta)(y^\alpha - x^\alpha) &= \frac{y^\beta - x^\beta}{y - x}(y^\alpha - x^\alpha)(y - x)\\ &\geq \frac{4\alpha}{(\alpha+1)^2}\frac{y^\beta - x^\beta}{y - x}(y^{(\alpha+1)/2} - x^{(\alpha+1)/2})^2.\end{aligned}$$

Since

$$y^\beta - x^\beta = \beta\int_x^y t^{\beta-1}dt \geq \beta\min\{x^{\beta-1}, y^{\beta-1}\}(y - x),$$

the conclusion follows. □

Corollary 44. *Let a, $b > 0$ and x, $y \geq 0$. Then*

$$(x^a - y^a)(x^b - y^b) \leq (x^{(a+b)/2} - y^{(a+b)/2})^2 \leq \frac{(a+b)^2}{4ab}(x^a - y^a)(x^b - y^b).$$

Proof. The second inequality is already proven in 42. For the proof of the first inequality, we divide it by y^{a+b} and set $z = x/y$. Then the inequality is equivalent to

$$(z^a - 1)(z^b - 1) \leq (z^{(a+b)/2} - 1)^2,$$

which after expansion can be equivalently written as $(z^{a/2} - z^{b/2})^2 \geq 0$, and this is true. □

Poincaré-Wirtinger inequalities

Lemma 45 (Poincaré-Wirtinger inequality). *Let* $\operatorname{meas}(\mathbb{T}) = 1$*. It holds for all* $v \in H^1(\mathbb{T})$ *satisfying* $\int_{\mathbb{T}} u dx = 0$ *that*

$$\int_{\mathbb{T}} u^2 dx \leq C_P \int_{\mathbb{T}} u_x^2 dx,$$

and the constant $C_P = 1/(4\pi^2)$ *is sharp.*

Lemma 46 (Discrete Poincaré-Wirtinger inequality). *Let* $N \in \mathbb{N}$*,* $h = 1/N$*,* $z_0, \ldots, z_N \in \mathbb{R}$ *satisfying* $z_N = z_0$ *and* $\sum_{i=0}^{N} z_i = 0$*. Then*

$$h \sum_{i=0}^{N-1} z_i^2 \leq C_p h^{-1} \sum_{i=0}^{N-1} (z_{i+1} - z_i)^2,$$

where $C_p = h^2/(4\sin^2(h\pi)) \geq 1/(4\pi^2)$*. This constant is sharp.*

These lemmas are stated in [81, Theorem 1]; for proofs see [56, p. 185] (Lemma 45) and [79, Theorem 1] (Lemma 46).

Appendix C

Bibliography

[1] N.D. Alikakos and R. Rostamian. Large time behavior of solutions of Neumann boundary value problem for the porous medium equation. *Indiana Univ. Math. J.* 30 (1981), 749-785.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures.* Lect. Math., Birkhäuser, Basel, 2005.

[3] T. Arbogast, M. Wheeler, and N.-Y. Zhang. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous medium. *SIAM J. Numer. Anal.* 33 (1996), 1669-1687.

[4] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Commun. Part. Diff. Eqs. 26* (2001), 43-100.

[5] D. Bakry and M. Emery. Diffusions hypercontractives. In: *Séminaire de probabilités XIX*, 1983/84, *Lect. Notes Math.* 1123 (1985), 177-206.

[6] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators.* Springer, Cham, 2014.

[7] J. Barrett, J. Blowey, and H. Garcke. Finite element approximation of a fourth order nonlinear degenerate parabolic equation. *Numer. Math.* 80 (1998), 525-556.

[8] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.* 105 (1989), 397-400.

[9] J.-D. Benamou, G. Carlier, Q. Mérigot, and E. Oudet. Discretization of functionals involving the Monge–Ampère operator. *Numer. Math.* 164 (2016). 611-636.

[10] J. Berryman. Extinction time for fast diffusion. *Phys. Lett. A* 72 (1979), 107-110.

[11] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 2014. `DOI:10.1093/imanum/dru032.`

[12] M. Bessemoulin-Chatard and F. Filbet. A Finite Volume Scheme for Nonlinear Degenerate Parabolic Equations *SIAM J. Sci. Comput.* 34 (2012), 559–583.

Read More: http://epubs.siam.org/doi/abs/10.1137/110853807

[13] M. Bessemoulin-Chatard. A finite volume scheme for convection-diffusion equations with nonlinear diffusion derived from the Scharfetter-Gummel scheme. *Numer. Math.* 121 (2012), 637-670.

[14] 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.

[15] M. Bonforte and G. Grillo. Asymptotics of the porous media equations via Sobolev inequalities. *J. Funct. Anal.* 225 (2005), 33-62.

[16] S. Boscarino, F. Filbet, and G. Russo. High order semi-implicit schemes for time dependent partial differential equations. *J. Sci. Comput.* 68 (2016), 975–1001.

[17] M. Bukal, E. Emmrich, and A. Jüngel. Entropy-stable and entropy-dissipative approximations of a fourth-order quantum diffusion equation. *Numer. Math.* 127 (2014), 365-396.

[18] C. Cancès and C. Guichard. Numerical analysis of a robust entropy-diminishing finite-volume scheme for parabolic equations with gradient structure. Preprint, 2015. `arXiv:1503.05649.`

[19] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the Bochner-Bakry-Emery approach. *Ann. Inst. H. Poincaré Prob. Stat.* 45 (2009), 734-753.

[20] J.A. Carrillo, J. Dolbeault, I. Gentil, and A. Jüngel. Entropy-energy inequalities and improved convergence rates for nonlinear parabolic equations. *Discr. Cont. Dyn. Sys. B* 6 (2006), 1027-1050.

[21] J. F. Carrillo and K. Fellner. Long-time asymptotics via entropy methods for diffusion dominated equations. *Asympt. Anal.* 42 (2005), 29-54.

[22] J. A. Carrillo, A. Jüngel, P. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.* 133 (2001), 1-82.

[23] J.A. Carrillo, A. Jüngel, and S. Tang. Positive entropic schemes for a nonlinear fourth-order equation. *Discrete Contin. Dyn. Sys. B* 3 (2003), 1-20.

[24] J. A. Carrillo and G. Toscani. Asymptotic L^1-decay of solutions of the porous medium equation to selfsimilarity. *Indiana Univ. Math. J.* 49 (2000), 113-141.

[25] J. A. Carrillo, A. Jüngel, and M. C. Santos. Displacement convexity for the entropy in semidiscrete nonlinear Fokker-Planck equations. Submitted for publication, 2016. `arXiv:1611.04716`.

[26] F. Cavalli, G. Naldi, G. Puppo, and M. Semplice. High-order relaxation schemes for nonlinear degenerate diffusion problems. *SIAM J. Numer. Anal.* 45 (2007), 2098-2119.

[27] C. Chainais-Hillairet, A. Jüngel, and S. Schuchnigg. Entropy-dissipative discretization of nonlinear diffusion equations and discrete Beckner inequalities. *Math. Model. Numer. Anal.* 50 (2016), 135-162.

[28] C. Chainais-Hillairet, A. Jüngel, and P. Shpartko. A finite-volume scheme for a spinorial matrix drift-diffusion model for semiconductors. *Numer. Meth. Partial Diff. Eqs.* 32 (2016), 819-846.

[29] S. Chow, W. Huang, Y. Li, and H. Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Rational Mech. Anal.* 203 (2012), 969-1008.

[30] S. Christiansen, H. Munthe-Kaas, and B. Owren. Topics in structure-preserving discretization. *Acta Numerica* 20 (2011), 1-119.

[31] K. Deimling. *Nonlinear Functional Analysis.* Springer, Berlin, 1985.

[32] M. Del Pino and J. Dolbeault. Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. *J. Math. Pures Appl.* 81 (2002), 847-875.

[33] J. Denzler and R. McCann. Fast diffusion to self-similarity: complete spectrum, long-time asymptotics, and numerology. *Arch. Ration. Mech. Anal.* 175 (2005), 301-342.

[34] B. Derrida, J. Lebowitz, E. Speer, and H. Spohn. Fluctuations of a stationary nonequilibrium interface. *Phys. Rev. Lett.* 67 (1991), 165-168.

[35] J. Dolbeault, I. Gentil, A. Guillin, and F.-Y. Wang. L^q-functional inequalities and weighted porous media equations. *Potential Anal.* 28 (2008), 35-59.

[36] J. Dolbeault, B. Nazaret, and G. Savaré. On the Bakry-Emery criterion for linear diffusions and weighted porous media equations. *Commun. Math. Sci.* 6 (2008), 477-494.

[37] C. Ebmayer and W. W. Liu. Finite element approximation of the fast diffusion and the porous medium equations. *SIAM J. Numer. Anal.* 46 (2008), 2393-2410.

[38] F. Edgeworth. On the probable errors of frequency-constants (contd.). *J. Roy. Stat. Soc.* 71 (1908), 499-512.

[39] E. Emmrich. Variable time-step θ-scheme for nonlinear evolution equations governed by a monotone operator. *Calcolo* 46 (2009), 187-210.

[40] E. Emmrich. Discrete versions of Gronwall's lemma and their application to the numerical analysis of parabolic problems. Preprint `http://www3.math.tu-berlin.de/preprints/files/Preprint-637-1999.pdf`

[41] M. Erbar and J. Maas. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Sys.* 34 (2014), 1355-1374.

[42] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In: P. G. Ciarlet and J. L. Lions (eds.). *Handbook of Numerical Analysis*, Vol. 7. North-Holland, Amsterdam (2000), 713-1020.

[43] R. Eymard, T. Gallouët, R. Herbin, and A. Michel. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.* 92 (2002) 41-82.

[44] R. Eymard, M. Gutnic, and D. Hilhorst. The finite volume method for an elliptic-parabolic equation. Proceedings of the Algoritmy'97 Conference on Scientific Computing (Zuberec). *Acta Math. Univ. Comenian.* (N.S.) 67 (1998), 181-195.

[45] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Prob.* 26 (2016), 1774-1806.

[46] F. Filbet. An asymptotically stable scheme for diffusive coagulation-fragmentation models. *Commun. Math. Sci.* 6 (2008), 257-280.

[47] F. Filbet and M. Herda. A finite volume method for driven-boundary convection-diffusion equations with relative entropy structure. Preprint `arXiv:1606.01104.`

[48] A. Glitzky and K. Gärtner. Energy estimates for continuous and discretized electro-reaction-diffusion systems. *Nonlin. Anal.* 70 (2009), 788-805.

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

[50] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.* 43 (2001), 89-112.

[51] G. Grillo and M. Muratori. Sharp short and long-time asymptotics for solutions to the Neumann problem for the porous media equation. *J. Diff. Eqs.* 254 (2013), 2261-2288.

[52] G. Grillo and M. Muratori. Sharp asymptotics for the porous media equation in low dimensions via Gagliardo-Nirenberg inequalities. *Riv. Mat. Univ. Parma* 5 (2014), 15-38.

[53] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* 97 (1975), 1061-1083.

[54] G. Grün and M. Rumpf. Nonnegativity preserving convergent schemes for the thin film equation. *Numer. Math.* 87 (2000), 113-152.

[55] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I.* Springer, Berlin, 1993.

[56] G. Hardy, J. Littlewood, and G. Pólya. *Inequalities.* Second edition. Cambridge University Press, Cambridge, 1952.

[57] I. Higueras. Representations of Runge-Kutta methods and strong stability preserving methods. *SIAM J. Numer. Anal.* 43 (2005), 924-948.

[58] A. Jüngel and D. Matthes. An algorithmic construction of entropies in higher-order nonlinear PDEs. *Nonlinearity* 19 (2006), 633-659.

[59] A. Jüngel and D. Matthes. The Derrida-Lebowitz-Speer-Spohn equation: existence, non-uniqueness, and decay rates of the solutions. *SIAM J. Math. Anal.* 39 (2008), 1996-2015.

[60] A. Jüngel and R. Pinnau. A positivity preserving numerical scheme for a nonlinear fourth-order parabolic equation. *SIAM J. Numer. Anal.* 39 (2001), 385-406.

[61] A. Jüngel. *Transport Equations for Semiconductors.* Lect. Notes Phys. 773, Springer, Berlin, 2009.

[62] A. Jüngel and J.-P. Milišić. A sixth-order nonlinear parabolic equation for quantum systems. *SIAM J. Math. Anal.* 41 (2009), 1472-1490.

[63] 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.

[64] A. Jüngel and S. Schuchnigg. Entropy-dissipating semi-discrete Runge-Kutta schemes for nonlinear diffusion equations. *Commun. Math. Sci.* 15 (2017), 27-54.

[65] A. Jüngel and W. Yue. Discrete Beckner inequalities via the Bochner-Bakry-Emery approach for Markov chains. To appear in *Ann. Appl. Prob.*, 2017.

[66] A. Jüngel and S. Schuchnigg. A discrete Bakry-Emery method and its application to the porous-medium equation. Preprint, 2017. `arXiv:1702.03780`.

[67] J. King. Interacting dopant diffusions in crystalline silicon. *SIAM J. Appl. Math.* 48 (1988), 405-415.

[68] R. Latala and K. Oleszkiewicz. Between Sobolev and Poincaré. *Lect. Notes Math.* 1745 (2000), 147-168.

[69] M. Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM Prob. Stat.* 1 (1996), 63-87.

[70] E. Lieb and M. Loss. *Analysis.* Second edition. Amer. Math. Soc., Providence, 2010.

[71] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. Royal Soc. A.* 371 (2013), 20120346 (28 pages), 2013.

[72] P.-L. Lions and S. Mas-Gallic. Une méthode particulaire déterministe pour des équations diffusives non linéaires. *C. R. Acad. Sci. Paris Sér. I Math* 332 (2001), 369-376.

[73] H. Liu and H. Yu. Entropy/energy stable schemes for evolutionary dispersal models. *J. Comput. Phys.* 256 (2014), 656-677.

[74] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* 261 (2011), 2250-2292.

[75] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Diff. Eqs.* 48 (2013), 1-31.

[76] A. Michel and J. Vovelle. Entropy formulation for parabolic degenerate equations with general Dirichlet boundary conditions and application to the convergence of FV methods. *SIAM J. Numer. Anal.* 41 (2003), 2262-2293.

[77] M. Ohlberger. A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations. *Math. Mod. Numer. Anal.* 35 (2001), 355-387.

[78] M. Rose. Numerical methods for flows through porous media I. *Math. Comput.* 40 (1983), 435-467.

[79] I. Schoenberg. The finite Fourier series and elementary geometry. *Amer. Math. Monthly* 57 (1950), 390-404.

[80] C. Seis. Long-time asymptotics for the porous medium equation: the spectrum of the linearized operator. *J. Diff. Eqs.* 256 (2014), 1191-1223.

[81] O. Shisha. On the discrete version of Wirtinger's inequality. *Amer. Math. Monthly* 80 (1973), 755-760.

[82] C.-W. Shu. Total-variation-diminishing time discretizations. *SIAM J. Sci. Stat. Comput.* 9 (1988), 1073-1084.

[83] E. Tadmor. Numerical viscosity of entropy stable schemes for systems of conservation laws I. *Math. Comp.* 49 (1987), 91-103.

[84] J. L. Vázquez. Asymptotic behaviour for the porous medium equation posed in the whole space. *J. Evol. Eqs.* 3 (2003), 67-118.

[85] J. L. Vázquez. *Smoothing and Decay Estimates for Nonlinear Parabolic Equations of Porous Medium Type.* Oxford Lecture Series in Mathematics and Its Applications 33, Oxford University Press, 2006.

[86] J. L. Vázquez. *The Porous Medium Equation. Mathematical Theory.* Oxford University Press, Oxford, 2007.

[87] Z. Wu, J. Yin, and C. Wang. *Elliptic and Parabolic Equations.* World Scientific, Singapore, 2006.

[88] L. Zhornitskaya and A. Bertozzi. Positivity-preserving numerical schemes for lubrication-type equations. *SIAM J. Numer. Anal.* 37 (2000), 523-555.

Appendix D

Curriculum Vitae

Buchengasse 155/150, 1100 Vienna
0676/ 63 82 132
stefan.schuchnigg@tuwien.ac.at

Born on July 21st 1986 in Vienna
Austrian citizenship

Educational background

Feb. 2017	Submitted PhD thesis *Entropy-preserving numerical schemes for nonlinear diffusion equations*
2011 – 2017	PhD studies in the field of PDEs at the Technical University (TU) Vienna, Institute for Analysis and Scientific Computing
Juni 2011	Degree "Dipl.-Ing." (equivalent to MSc) in *Technical Mathematics*
2010 – 2011	Diploma thesis *The Keller-Segel Model in* R^d*: Global Existence in the Case of Linear and Non-Linear Diffusion*
Sept. 2009	Summerschool *MathNanoSci*, University of L'Aquila
2008 –	Studies for Bachelor's program in *Technical Physics*, TU Vienna
2005 – 2007	Studies for Diploma program in *Music Education*, University of Music and Performing Arts Vienna
2004 – 2011	Studies for Diploma program in *Technical Mathematics* with a focus on Mathematics in Science, TU Vienna
2004	"Matura" (general qualification for university entrance), Don-Bosco-Gymnasium, Unterwaltersdorf

Professional experience

03/17 – 06/17	Research assistant at the Institute for Analysis and Scientific Computing, TU Vienna
03/15 – 02/17	Scientific manager of the doctoral program *Dissipation and dispersion in nonlinear partial differential equations*, TU Vienna
09/11 – 02/15	Research assistant at the Institute for Analysis and Scientific Computing, TU Vienna
07/04, 07/05	Programmer for Siemens, I&S department

Teaching experience (TU Vienna)

10/13 – 01/14	Analysis *(Analysis 1 UE)*
03/14 – 06/14	Nonlinear partial differential equations *(Nichtlineare partielle Differentialgleichungen UE)*
10/13 – 01/14	Partial differential equations *(Partielle Differentialgleichungen UE)*
03/13 – 06/13	Calculus of variations *(Variationsrechnung UE)*
03/12 – 06/12	Mathematics course for students of physics *(Konversatorium zur Mathematik für TPH)*
03/12 – 06/12	Numerical mathematics *(Numerische Mathematik für LA)*
10/11 – 01/12	Analysis for students of physics *(Analysis I für TPH)*
10/10 – 06/11	Introductory physics *(Grundlagen der Physik I/II UE)*
10/10 – 01/11	Partial differential equations *(Partielle Differentialgl. UE)*
03/10 – 06/10	Statistical physics *(Statistische Physik I VU)*
10/09 – 06/10	Mathematical methods in theoretical physics *(Mathematische Methoden der theoretischen Physik UE)*
10/07 – 10/08	Mathematics for students of computer science *(Mathematik 1 für Informatik und Wirtschaftsinformatik UE)*

Publications

2017, Preprint	A. Jüngel and S. Schuchnigg. A discrete Bakry-Emery method and its application to the porous-medium equation. Preprint, 2017. `arXiv:1702.03780`.
2016	A. Jüngel and S. Schuchnigg. Entropy-dissipating semi-discrete Runge-Kutta schemes for nonlinear diffusion equations. *Commun. Math. Sci.* 15 (2017), 27-54.
2016	C. Chainais-Hillairet, A. Jüngel, and S. Schuchnigg. Entropy-dissipative discretization of nonlinear diffusion equations and discrete Beckner inequalities. *Math. Model. Numer. Anal.* 50 (2016), 135-162.

Civilian national service and avocational activities

2009 –	Member of students association's theater group
2007 – 2015	Volunteer for the Red Cross (Mödling)
2006 –	Drummer/percussionist for the Vienna University Orchestra
2006	Emergency medical assistant and driver, Red Cross (Mödling)

Language proficiency

German (native)

English (fluent)

French (basic)

IT skills

Operating systems	Linux, Windows
Languages	C/C++, C$^\sharp$, HTML
Mathematics	MATLAB, MAPLE, MATHEMATICA
Other	UNITY3D, LaTeX