Angewandte Mathematik

# Long-Time Behaviour of Nonlinear Fokker-Planck Equations

Angewandte Mathematik

# Long-Time Behaviour of Nonlinear Fokker-Planck Equations

Diplomarbeit
vorgelegt von
Jan-Frederik Pietschmann

Westfälische Wilhelms-Universität Münster
Institut für Numerische und Angewandte Mathematik

— November 2008 —

# Contents

# Chapter 1

# Introduction

This thesis deals with the convergence of solution of a system of non-linear Fokker-Planck equations

$$\rho_t = \operatorname{div}(D(\rho)(\nabla u'(\rho) + \nabla V(x))), \quad x \in \mathbb{R}^N,\ t \geq 0 \tag{1.1}$$

to a equilibrium or stationary state. These are defined as solutions which do not depend on time anymore. The Fokker-Planck equation plays a key role in many areas of physics. The linear Fokker-Planck equation

$$\rho_t = \operatorname{div}(\nabla \rho + \rho \nabla V(x)), \quad x \in \mathbb{R}^N, t \geq 0$$

comes from the underlying stochastic differential equation

$$dX_t = dW_t - \nabla V(X_t)dt,$$

where $W_t$ is a Brownian motion (also called standard Wiener process). This representation makes clear that Fokker-Planck equations describe the dynamics of a set of particles that are influenced by both diffusion and drift. For a comprehensive review see [Risken89].
For the linear equation a stationary state, which is defined as a solution which does not change in time, is given by

$$\rho = e^{-V}.$$

Given this state, it can be asked if all solutions of the equation converge to this state as $t \to \infty$ and if they do, at which rate.

To explore this behaviour for equation (1.1), two techniques are used.
The first one is the so-called entropy dissipation method. Here, an entropy functional is definied and then used to controll the distance between a solution at time $t$ and the steady state. By studying the derivatives of this functional and using log-sobolev inequalities as well as Gronwall's inequality, it is possible to derive rates for the convergence in terms of relative Entropy. These can be translated to rates in $L^1$ using Czisar-Kullback type inequalities.
The second aprroach is to see the Fokker-Planck equation as a Gradient flow, that means as a dynamical system that evolves according to

$$\rho_t = -\mathrm{grad} E(\rho).$$

The idea is to define a manifold such that the Fokker-Planck equation can be understood as such a system on this manifold. If it is furthermore possible to show that the functional $E$ is convex, a contraction principle follows. The proper sense of convexity here is the so-called displacement convexity which reveals to connection to the Monge-Kantorovich problem of optimal transport.

## 1.1 Acknowledgements

First of all I would like to thank Prof. Martin Burger and Prof. Peter Markowich for giving me the possibility to write a thesis on this interesting topic and for many helpful discussions. Furthermore, I would like to thank the Wolfgang Pauli Institut (WPI) in Vienna for providing financial support for a four week stay at Cambridge this February. Also, I would like to thank the Institute for Pure and Applied Mathematics (IPAM), which allowed me to participate in the Optimal Transport long-term program and to stay in Los Angeles for three and a half months. During this time, I had a number of useful discussions and I especially want to thank José Antonio Carrillo de la Plata and Yann Brenier for their help. Also, I want to thank Alessio Figalli and Marcus Wunsch for reading parts of this thesis and Christian Döring for delivering it. Finally, I'd like to thank my parents and of course Maja for their support over the years.

## 1.2 Notation

Throughout this thesis, we will use the following notations. Let $v \in \mathbb{R}^N$ be a real-valued vector and $A \in \mathbb{R}^{N \times N}$ be a real-valued matrix. We will use the following notations

- By $A^i$, we denote the i-th row, by $A_k$ the k-th column of A.
- The (vector-valued) divergence operator will act row-wise on matrices, i.e.

$$\mathrm{div} A(x) = (\mathrm{div} A^1, \ldots, \mathrm{div} A^N).$$

- For the scalar product of two matrices we write

$$A : B = tr(AB^T) = \sum_{i=1}^{N} A^i \cdot B^i,$$

  where $\cdot$ denotes the scalar product in $\mathbb{R}^N$.

- The gradient of a matrix is defined as the matrix which has as entries the gradient of each component of the original matrix, i.e.

$$\nabla A = \begin{pmatrix} \nabla A_{11} & \cdots & \nabla A_{1N} \\ \vdots & \ddots & \vdots \\ \nabla A_{N1} & \cdots & \nabla A_{NN} \end{pmatrix}$$

- The Jacobian matrix of a vector map $x \mapsto v(x)$, $x \in \mathbb{R}^N$ has gradients on its rows, i.e.

$$\nabla v(x) = \left[ \nabla v_1(x), \ldots, \nabla v_N(x) \right].$$

Furthermore, we recall the following theorem:

**Theorem 1.2.1 (A Green's formula for matrices)** *Let $\mathbb{R}^N \ni x \mapsto V(x) \in \mathbb{R}^N$ and $\mathbb{R}^N \ni x \mapsto A(x) \in \mathbb{R}^{N \times N}$ be smooth maps. Let $\Omega$ be a bounded subdomain of $\mathbb{R}^N$ with smooth boundary $\partial\Omega$ having $x \mapsto \nu(x)$ as normal unit vector. Then,*

$$\int_\omega V(x)^T \mathrm{div} A(x) dx = \int_{\partial\Omega} V(x)^T A(x) \nu(x) d\sigma - \int_\Omega \nabla V(x) : A(x) dx.$$

# Chapter 2

# The Optimal Transportation Problem

The problem of optimal transportation was originally proposed by Monge in 1781 in "mèmoire sur la thèorie des dèblais et des remblais". Today there exists a vast literature about this topic. We particularly mention the surveys by Evans [Evans01] and Ambrosio [Ambrosio00] as well as the book by Villani [Villani03].

## 2.1 Monge's Problem

The original problem of Monge was how to transfer a pile of sand into a hole with the lowest amount of work possible. To express this in mathematical formulae, we consider two separable metric space $X$ and $Y$ and denote by $\mathcal{P}(X)$, $\mathcal{P}(Y)$ the set of all probability measures on $X$, $Y$. We then take two probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, both having the same finite mass,

$$\mu(X) = \nu(Y) < \infty. \tag{2.1}$$

Furthermore, we make the following definition [Ambrosio05].

**Definition** (push-forward) Let $X$, $Y$ be separable metric spaces, $\mu \in \mathcal{P}(X)$, and $r : X \to Y$ is a $\mu$-measurable map denoted by $r_{\#}\mu \in \mathcal{P}(Y)$. The push-forward of $\mu$ through $r$ is defined by

$$r_{\#}\mu(B) := \mu(r^{-1}(B)) \quad \forall B \in \mathcal{B}(Y), \tag{2.2}$$

where $\mathcal{B}(Y)$ is the family of all Borel subsets of $Y$. More generally we have

$$\int_X f(r(x))d\mu(x) = \int_Y f(y)dr_{\#}\mu(y) \tag{2.3}$$

for every bounded (or $r_{\#}\mu$-integrable) Borel function $f : X \to \mathbb{R}$.

Given a Borel cost function $c : X \times Y \to [0, +\infty]$ that describes the cost of moving mass from location $x$ to location $y$ as $c(x, y)$, the Monge problem is given by

$$\inf_t \left\{ \int_X c(x, t(x))d\mu(x) : t_{\#}\mu = \nu \right\}. \tag{2.4}$$

In the original formulation, Monge was interested in the case that the work is proportional to the distance, i.e. $c(x, y) = |x - y|$. Other examples would be the distance squared ($c(x, y) = |x-y|^2$) which leads to interesting theoretical results or the relativistic case which is

$$c(x-y) = \begin{cases} 1 - \sqrt{1 - |x-y|^2} & \text{if } |x-y| \leq< 1 \\ +\infty & \text{else.} \end{cases}$$

**Existence and Uniqueness of Solutions**

In the following, we state two examples from [Ambrosio00] that show that neither existence nor uniqueness for the general problem can be expected.

**Example** (Non-existence) Let $\mu = \delta_0$ be the Dirac delta-measure located at zero and $\nu = (\delta_{-1/2} + \delta_{+1/2})/2$. Then there exists no Solution to (2.4), because there is no map $t$ such that $t_{\#}\mu = \nu$. In other words: In Monge's Problem does not allow the splitting of mass. □

To illustrate non-uniqueness, the cost function $c(x, y) = |x-y|$ is considered. In this case, it can be proven that the infimum Monge's problem is always greater than

$$\sum \left\{ \int_X ud(\nu - \mu) : u \in \text{Lip}_1(X) \right\}. \tag{2.5}$$

We use this for the following example.

**Example** (Non-uniqueness) Let $n \geq 1$ be an integer and $\mu = \chi_{[0,n]}\mathcal{L}^1$ and $\nu = \chi_{[1,n+1]}\mathcal{L}^1$. Then the map $\psi(t) = t + 1$ is optimal. Indeed, the cost relative to $\psi$ is $n$ and choosing the 1-Lipschitz function $u(t) = t$ in (2.5), we obtain that the supremum is at least $n$, whence the optimality of $\psi$ follows. But since the optimal cost is $n$, if $n > 1$ another optimal map $\psi$ is given by

$$\psi(t) = \begin{cases} t + n & \text{on } [0, 1] \\ t & \text{on } [1, n]. \quad \square \end{cases} \tag{2.6}$$

However, under additional assumptions on the cost function $c$, existence and uniqueness of an optimal map can be ensured. For example, if $\mu, \nu \in \mathcal{P}(\mathbb{R}^N)$ and $c$ is of the form $c(x, y) = h(x-y)$ with $h : \mathbb{R}^N \to [0, +\infty)$ strictly convex, then there exist a unique solution to (2.4). (See [McCann95].)

## 2.2 Kantorovich's Approach

Kantorovich's approach is a relaxed formulation of Monge's Problem. Instead of looking for transport maps, we are interested in transport plans which are probability measures on the product space $X \times Y$. We are only using transport plans $\pi \in \mathcal{P}(X \times Y)$ whose marginals are the measures $\mu$ and $\nu$, i.e.

$$\pi[A \times Y] = \mu[A], \quad \pi[X \times B] = \nu[B], \tag{2.7}$$

for all measurable subsets $A \subset X$ and $B \subset Y$. We denote the set of all such transport plans by $\Pi(\mu, \nu)$. Then the Kantorovich problem is given by

$$\min_{\pi} I[\pi], \tag{2.8}$$

with

$$I[\pi] := \min_{\pi} \left\{ \int_{X \times X} c(x, y) d\pi(x, y) : \pi \in \Pi(\mu, \nu) \right\} \tag{2.9}$$

## 2.3 Duality formulation

It is known that a linear minimization problem with convex constraints has a dual formulation. For the optimal transport problem in Kantorovich formulation this was introduced by Kantorovich himself in 1942.

**Theorem 2.3.1** *Under the assumption that c is lower semi-continous, the minimum of the Kantorovich problem is equal to*

$$\sup J(\varphi, \psi), \tag{2.10}$$

*with*

$$J(\varphi, \psi) := \sup \left\{ \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y) \right\}, \tag{2.11}$$

*where the supremum is taken over all pairs* $(\varphi, \psi) \in C_b^0(X) \times C_b^0(Y)$ *such that* $\varphi(x) + \psi(y) \leq c(x, y)$*. The set of all* $(\varphi, \psi)$ *fullfilling this condition is denoted by* $\Phi_c$*.*

**Proof** *(formal)* To justifiy this theorem, we state a formal proof from Villani [Villani03]. We write the constrained infimum problem as a inf sup problem and then use some minmax principle to replace "inf sup" by "sup inf". First we introduce the **indicator function** of Π as the function

$$\mathbf{1}_\Pi(\pi) = \begin{cases} 0 & \text{if } \pi \in \Pi(\mu, \nu) \\ +\infty & \text{else} \end{cases} \tag{2.12}$$

Since all the constraints defining Π are linear, this can be written as the solution of a supremum problem involving only linear functionals. This is

$$\mathbf{1}_\Pi(\pi) = \sup_{(\phi,\psi)} \left[ \int \phi d\mu + \int \psi d\nu - \int [\phi(x) + \psi(y)] d\pi(x, y) \right],$$

for all $(\phi, \psi) \in C_b(X) \times C_b(Y)$. Using the following formulation of the Kantorovich problem,

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \inf_{\pi \in \mathcal{P}(\mu,\nu)} (I[\pi] + \mathbf{1}_\Pi(\pi)), \tag{2.13}$$

we get

$$\begin{aligned} \inf_{\pi \in \Pi(\mu,\nu)} I[\pi] &= \inf_{\pi \in \mathcal{P}(\mu,\nu)} \sup_{(\phi,\psi)} \left\{ \int_{X \times Y} c(x, y) d\pi(x, y) \right. \\ &+ \left. \left[ \int \phi d\mu + \int \psi d\nu - \int [\phi(x) + \psi(y)] d\pi(x, y) \right] \right\}, \end{aligned}$$

because the first integral does not depend on $\varphi$ and $\psi$. If now we assume a minmax principle that allows us to interchange sup and inf, then

$$\begin{aligned} \inf_{\pi \in \Pi(\mu,\nu)} I[\pi] &= \sup_{\phi,\psi} \left\{ \int \phi d\mu + \int \psi d\nu \right. \\ &- \left. \sup_{\pi \in \mathcal{P}(\mu,\nu)} \int_{X \times Y} [\phi(x) + \psi(y) - c(x, y)] \, d\pi(x, y) \right\}. \end{aligned}$$

We now want to calculate the second supremum over all $\pi \in \mathcal{P}(\mu, \nu)$ . Therefore we define $\xi(x,y) := \phi(x) + \psi(x) - c(x,y)$ and consider two cases: If $\xi$ has a positive value at some point $(x_0, y_0)$, then by choosing $\pi = \lambda\delta_{(x_0,y_0)}$ and letting $\lambda \to +\infty$, we see that the supremum in infinity. If, on the other hand, $\xi \leq 0$ for all $(x,y)$ ($d\mu \otimes d\nu$-everywhere), when the supremum is obtained for $\pi = 0$. Therefore,

$$\sup_{\pi \in \mathcal{P}(\mu,\nu)} \int_{X \times Y} [\phi(x) + \psi(y) - c(x,y)]\, d\pi(x,y) d\pi(x,y) = \begin{cases} 0 & \text{if } (\phi, \psi) \in \Phi_c \\ +\infty & \text{else} \end{cases} \tag{2.14}$$

holds. By substituting this formula into (2.14), we finally obtain

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup_{\phi,\psi \in \Phi_c} \left\{ \int \phi d\mu + \int \psi d\nu \right\} = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi). \tag{2.15}$$

□

## 2.4 Quadratic Costs

For the special case of a cost function of the form $c(x,y) = |x-y|^2$ there exist additional results. Especially popular is the following one, which in [Villani03] is referred as Brenier's theorem.

**Remark** Here $\mu$ *does not give mass to small sets* means that $\mu$ does not give mass to sets of at most Hausdorff dimension $n-1$.

**Theorem 2.4.1** *Let $\mu$, $\nu$ be probability measures on $\mathbb{R}^N$ with finite second moments. We consider the Monge-Kantorovich problem with the quadratic cost function. Then, if $\mu$ does not give mass to small sets, there is a unique optimal $\pi$ which is*

$$d\pi(x,y) = d\mu(x)\delta_{\nabla\varphi(x)}(y), \tag{2.16}$$

*or equivalently,*

$$\pi = (\mathrm{Id} \times \nabla\varphi)\#\mu, \tag{2.17}$$

*where $\nabla\varphi$ is the unique (i.e. uniquely determined $d\mu$-almost everywhere) gradient of a convex function which pushes $\mu$ forward to $\nu$. Moreover,*

$$\mathrm{Supp}(\nu) = \overline{\nabla\varphi(\mathrm{Supp}(\mu))}. \tag{2.18}$$

Under certain assumptions, it is also possible to derive a differential formulation of Monge's Problem. We consider two measures $\mu$, $\nu \in \mathcal{P}(\mathbb{R}^N)$, absolutely continous with respect to the Lebesgue measure. Then, according to the Radon-Nikodym theorem, we can write

$$d\mu = f(x)dx, \quad d\nu = g(x)dx, \tag{2.19}$$

for some densities $f$, $g \in L^1(\mathbb{R}^N)$. Using the definition of the push-forward, the mass transportation problem is

$$\int \xi(y)g(y)dy = \int \xi(\nabla\varphi(x))f(x)dx, \tag{2.20}$$

for all test functions $\xi \in C_b(\mathbb{R}^N)$. If we further assume that $\nabla\varphi$ is smooth and one-to-one, we can apply the change of variables formula with $y = \nabla\varphi(x)$

$$\int \xi(y)g(y)dy = \int \xi(\nabla\varphi(x))g(\nabla\varphi(x))|\det(D^2\phi(x))|dx. \tag{2.21}$$

Therefore, as $\xi$ is an arbitrary test function, we get (neglecting the absolute value because $D^2\varphi > 0$ due to the convexity of $\varphi$)

$$f(x) = g(\nabla\varphi(x))\det(D^2\varphi(x)) \tag{2.22}$$

or, if $g$ is positive,

$$\det(D^2\varphi(x)) = \frac{f(x)}{g(\nabla\varphi(x))}. \tag{2.23}$$

This equation is usually called Monge-Ampère equation.

## 2.5 Kantorovich-Wasserstein distances

**Definition** *(Kantorovich-Wasserstein distances)* Given a metric space $(X, d)$ the p-Kantorovich-Wasserstein distance between two measures in the space $\mathcal{P}(X)$ is definied as

$$W_p(\mu, \nu) = \left( \inf_{\pi\in\Pi(\mu,\nu)} \int_X d(x,y)^p d\pi(x,y) \right)^{1/p} \tag{2.24}$$

### 2.5.1 Benamou-Brenier Reformulation

In their paper [Benamou00], Benamou and Brenier provide an alternative formulation for the $L^2$ Kantorovich-Wasserstein distance, in order to provide a numerical solution. Given a fixed time intervall $[0, T]$, they transfer the problem to a continuum mechanical framework. Considering everything smooth enough, time-dependent density and velocity fields $\rho(x,t) \geq 1$, $v(x,t) \in \mathbb{R}^N$, $t \in [0,T]$ and $x \in \mathbb{R}^N$ subject to the continuity equation

$$\rho_t + \operatorname{div}(\rho v) = 0, \tag{2.25}$$

and the initial and final conditions

$$\rho(0,\cdot) = \rho_0 \qquad \rho(T,\cdot) = \rho_T. \tag{2.26}$$

In this setting, they prove the following proposition.

**Proposition 2.5.1** *The square of the $L^2$ Kantorovich distance is equal to the infimum of*

$$T \int_{\mathbb{R}^N} \int_0^T \rho(t,x)|v(t,x)|^2 dx dt, \tag{2.27}$$

*among all $(\rho, v)$ satisfying (2.25) and (2.26).*

**Proof** The proof of this proposition is obtained using Lagrangian coordinates. We assume that $\rho_0$ and $\rho_T$ are compactly supported in $\mathbb{R}^N$ and bounded. Furthermore, we consider sufficiently smooth field $\rho$, $v$ satisfying (2.25) and (2.26). This again makes this proof only formal. It can be made rigorous using for example approximation arguments. We will state part of the proof from [Villani03, Section 8.1]. Another version can be found in [Ambrosio05, Section 8.1].
To introduce Lagrangian coordinates, we define $X(t,x)$ by

$$X(0,x) = x, \quad \frac{d}{dt}X(t,x) = v(t, X(t,x)), \tag{2.28}$$

and assume that $(X(t,\cdot))_{0\leq t\leq T}$ is a locally Lipschitz family of diffeomorphisms of $\mathbb{R}^N$.

**Step I:**

First of all, we show that the push forward $(X(t,\cdot))\#\rho_0$ is a solution of the

continuity equation in the distributional sense. Again, we neglect all issues like existence of derivatives, etc.
Let $\mu$ be a probability measure on $\mathbb{R}^N$, and $\rho(t,\cdot) = X(t,\cdot)\#\mu$. We will show that this is a solution of (2.25) in $\mathcal{C}([0,T];P(\mathbb{R}^N))$, where $P(\mathbb{R}^N)$ is equipped with the weak topology. We choose a test function $\varphi \in \mathcal{D}(\mathbb{R}^N)$. If we neglect all smoothness and existence issues, we can simply write

$$\begin{aligned} \frac{d}{dt}\int \varphi(x)\rho(t,x)dx &= \frac{d}{dt}\int \varphi(X(t,x))\rho_0(x)dx \\ &= \int \nabla\varphi(X(t,x))\frac{d}{dt}(X(t,x))\rho_0(x)dx \\ &= \int \nabla\varphi(x)v(t,x)\rho(t,x)dx. \end{aligned}$$

Therefore, $\rho(t,\cdot)$ satisfies the continuity equation in the distributional sense.

**Step II:**

We now show that the Wasserstein distance is a lower bound for our functional. To do so, we can write

$$\begin{aligned} T\int_{\mathbb{R}^N}\int_0^T \rho(t,x)|v(t,x)|^2 dxdt &= T\int_{\mathbb{R}^N}\int_0^T \rho_0(x)|v(t,X(t,x))|^2 dxdt \\ &= T\int_{\mathbb{R}^N}\int_0^T \rho_0(x)|\frac{d}{dt}X(t,x)|^2 dxdt \end{aligned}$$

Using Jensen's equation, this leads to

$$\begin{aligned} T\int_{\mathbb{R}^N}\int_0^T \rho_0(x)|\frac{d}{dt}X(t,x)|^2 dxdt &\geq \int_{\mathbb{R}^N} \rho_0(x)|X(T,x)-X(0,x)|^2 dx \\ &= \int_{\mathbb{R}^N} \rho_0(x)|X(T,x)-x|^2 dx \\ &\geq W_2^2(\rho_0,\rho_T). \end{aligned}$$

The last inequality is due to the fact that $\nabla\psi$ as well as $X(t,x)$ fullfill condition 2.3 and $\nabla\psi$ is the optimal map. This already shows that the new formulation of the distance is bounded below by the original Wasserstein distance.

<u>**Step III:**</u>

Next, we see that the optimal choise of $X(t,x)$ is

$$X(t,x) = x + \frac{t}{T}\left(\nabla\varphi(x) - x\right), \tag{2.29}$$

which impies that there actually exist a combination of $(\rho, v)$ that achieve the infimum, defined by

$$\int \varphi(x)\rho(t,x)dxdt = \int \varphi(x + \frac{t}{T}\left(\nabla\varphi(x) - x\right))\rho_0(x)dxdt, \tag{2.30}$$

Assuming that $X(t,\cdot)$ is invertable we define the velocity field by

$$v(t,x) = \left(\frac{d}{dt}X(t,\cdot)\right) \circ X(t,\cdot)^{-1} = (T - \mathrm{Id}) \circ X(t,\cdot)^{-1}. \tag{2.31}$$

We furthermore define $v = 0$ whenever $\rho$ is zero. As in step I, one can show that $(\rho_t, v_t)$ also solve the continuity equation. Now choose a nonnegative, measurable function $\Phi$, we have

$$\int \rho(t,x)\Phi(v(t,x))dx = \int \rho_0(x)\Phi(T(x) - x)dx. \tag{2.32}$$

Choosing $\Phi(v) = |v|^2$, this leads to

$$\int \rho_t|v_t|^2 dx = \int \rho_0(x)|T(x) - x|^2 dx. \tag{2.33}$$

This ensures the existiance of a minizing pair $(\rho, v)$ and therefore concludes the proof. □

Formally, the optimality conditions of this minimization problem can be obtained using the methods of Lagrange multipliers. Here, they turn out to be

$$v(t,x) = \nabla\phi(t,x), \tag{2.34}$$

where $\phi$ is the Lagrange multiplier of the constraints (2.25) and (2.26), and $\phi$ satisfies the Hamilton-Jacobi equation

$$\phi_t = -\frac{1}{2}|\nabla\phi|^2. \tag{2.35}$$

**Remark** i) In terms of fluid mechanics, this corresponds to a pressureless potential flow.

ii) Considering a metric space endowed with the $L^2$ Kantorovich-Wasserstein distance, the optimality conditions describe, by definition, the geodesics in this space.

## 2.6 Displacement Convexity

In this section, we will introduce the concept of diplacement convexity which was introduced by McCann in his PhD thesis and in [McCann97] in 1997. In our discussion, we will mostly follow the presentation in [Villani03, Chapter 5]. Futhermore, we will again consider only the quadratic cost function

$$c(x, y) = |x - y|^2.$$

We start with the following definition.

**Definition** Let $\mu$, $\nu$ be probability measures on $\mathbb{R}^N$ and assume that they do not give mass to small sets. Then, according to Brenier's theorem, there exists a gradient of a convex function $\nabla\varphi$ such that $\nabla\varphi_{\#}\mu = \nu$. Define the displacement interpolation as the family of probability measures

$$\rho_t := [\mu, \nu]_t = [(1-t)\mathrm{Id} + t\nabla\varphi]\,\#\mu. \tag{2.36}$$

We will now use this interpolation between the probability measures $\mu$ and $\nu$ to define a new form of convexity for functionals on the space of probability measures.

**Definition** *(Displacement Convexity)*

i) A subset $\mathcal{P}$ of $\mathcal{P}_{\mathrm{ac}}$ is called displacement convex if for all $\mu$, $\nu$ in $\mathcal{P}$ and for all $t \in [0, 1]$, the displacement interpolant $\rho_t$ is still in $\mathcal{P}$.

ii) If $E$ is a functional on the space $\mathcal{P}$, then $E$ is diplacement convex, if, whenever $\mu$ and $\nu$ are given elements of $\mathcal{P}$ and $\rho_t$ is the displacement interpolation between them, $t \mapsto E(\rho_t)$ is convex on $[0, 1]$. F is said to be strictly displacement convex, if $t \mapsto E(\rho_t)$ is strictly convex.

iii) A functional is said to be uniformly displacement convex if, for some $\alpha > 0$,

$$\frac{d^2}{dt^2}E(\rho_t) \geq \alpha W_2(\rho_0, \rho_1), \quad t \in (0,1). \tag{2.37}$$

It is said to be semi-displacement-convex, if, for some $\lambda \geq 0$

$$\frac{d^2}{dt^2}E(\rho_t) \geq -\lambda W_2(\rho_0, \rho_1), \quad t \in (0,1). \tag{2.38}$$

The concept of displacement convexity will be usefull later interpretated as convexity along geodesics in a space more complicated that $(\mathcal{P}, W_2)$.

# Chapter 3

# Gradient Flow Formulations of PDEs

## 3.1 Introduction to Gradient Flows

To introduce gradient flows, we start with an example in the Euclidian space. We fix some entropy functional $E \in \mathcal{C}^2(\mathbb{R}^N)$ und study the solution of the ordinary differential equation

$$\frac{dx(t)}{dt} = -\text{grad}E(x(t)), \quad x \in \mathbb{R}^N, \quad t \in [0, \infty). \tag{3.1}$$

This equation is called **steepest descent** or **gradient flow** on the entropy (energy) landscape created by $E$. We will now investigate proporties of solutions of this gradient flow under the assumption that $E$ is uniformly convex, following [Carillo06].

**Proposition 3.1.1** (Contraction / expansion bounds in a semi-convex valley) *Fix $k \in \mathbb{R}$. If $E \in \mathcal{C}^2(\mathbb{R}^N)$ is uniformly convex, i.e. $D^2E(X) \geq k\text{Id}$ for all $x \in \mathbb{R}^N$, and the curve $x(t)$, $y(t) \in \mathbb{R}^N$, $t \in [0, \infty)$ both solve (3.1), then*

$$|x(t + t_0) - y(t + t_0)| \leq e^{-kt}|x(t_0) - y(t_0)|. \tag{3.2}$$

**Proof** To proof this, we first define the linear interpolation between $x(t)$ and $y(t)$ by

$$g(s) = (1 - s)x(t) + ty(t). \tag{3.3}$$

With $f(t) = |x(t) - y(t)|^2/2$ it follows

$$\begin{aligned} f'(t) &= -\langle x(t) - y(t), \nabla E(x(t)) - \nabla E(y(t))\rangle \\ &= -\left\langle x(t) - y(t), \int_0^1 \frac{d}{ds}\left(\nabla E(g(s))\right) ds \right\rangle \\ &= -\left\langle x(t) - y(t), \int_0^1 D^2 E(g(s)) \frac{dg(s)}{ds} ds \right\rangle \\ &\leq -\left\langle x(t) - y(t), \int_0^1 k\mathrm{Id}(x(t) - y(t)) ds \right\rangle \\ &= -2kf(t). \end{aligned}$$

Using Gronwalls inequality [Gronwall19], this leads to $f(t+t_0) \leq e^{-2kt} f(t_0)$.
□

If $k > 0$, the above proposition shows that the solution of the map

$$x(0) \in \mathbb{R}^N \to X_t(x(0)) = x(t)$$

of the initial value problem of (3.1) defines a uniform contraction on $\mathbb{R}^N$ for each $t \leq 0$. The fact that this map is well-defined follows locally in space and time from the $\mathcal{C}^2$-smoothness of $E$. To show that it is globally well-defined for all future times, we make use of the fact that $E$ is coercive which follows from its uniform convexity and smoothness. Using Taylors formula we get

$$E(x) \geq E(x(0)) + < \nabla E(x(0)), x - x(0) > + k\frac{|x - x(0)|^2}{2}. \tag{3.4}$$

This proporty ensures the compactness of the level sets

$$\mathcal{S}_{E(x(0))} = \{x | E(x) \leq E(x(0))\}, \tag{3.5}$$

which contain $x(t)$ for all $t$.
Furthermore, since $\mathbb{R}^N$ is complete, from the contraction mapping principle follows the existance of a unique fixpoint $x_\infty$ of the map $X_t$ as well as the fact that every solution of (3.1) converges to $x_\infty$ for $t \to \infty$.

## 3.2 Otto's Approach for the porous medium equation

In his by now famous paper [Otto01], Otto formally derives a gradient flow formulation for the porous medium equation

$$\rho_t - \Delta(\rho^m) = 0, \tag{3.6}$$

where $\rho \geq 0$ is a density function on $\mathbb{R}^N$, $t \in [0, \infty)$ is the time and $x \in \mathbb{R}^N$. Furthermore, $\Delta$ is the laplace operator with respect to the spatial variable $x$.

**Remark** i) For $m = 1$ this is the heat equation. For $m > 1$, the equation is called slow diffusion equation or for $0 < m < 1$ fast diffusion equation respectivly.

ii) In the following, for technical reasons, we restrict the values of $m$ via

$$m > \frac{N}{N+2} \quad \text{and} \quad m \geq 1 - \frac{1}{N}. \tag{3.7}$$

As a consequence of the gradient flow interpretation of this equation, Otto (formally) maintains convergence rates to the asymptotic solutions of the porous medium equation by standard Riemannian calculus.

**Remark** In case of the porous medium equation, the stationary solutions are the so-called Prattle-Barenblatt [Pattle59] solutions. These are self-similar solutions of the form

$$\rho_\infty(t,x) = \frac{1}{t^{N\alpha}} \hat{\rho}_\infty\left(\frac{x}{t^\alpha}\right), \tag{3.8}$$

where the function $\hat{\rho}_\infty$ is implicitly given by

$$e'(\hat{\rho}_\infty(y)) = \left\{ \begin{array}{lll} \frac{m}{m-1}\hat{\rho}_\infty(y)^{m-1} & = \max\{\lambda - \alpha\frac{1}{2}|y|^2, 0\} & \text{for } m > 1 \\ \ln \hat{\rho}_\infty(y) + 1 & = \lambda - \alpha\frac{1}{2}|y|^2 & \text{for } m = 1 \\ \frac{m}{m-1}\hat{\rho}_\infty(y)^{m-1} & = \lambda - \alpha\frac{1}{2}|y|^2 & \text{for } m < 1 \end{array} \right\}. \tag{3.9}$$

Furthermore, $\alpha$ is given by

$$\alpha = \frac{1}{N(m-1)+2} \tag{3.10}$$

and $\lambda$ is chosen such that

$$\int \hat{\rho}_\infty(y)dy = 1. \tag{3.11}$$

The Bahrenplatt-Prattle profiles are stationary solutions of the porous medium equation in the sense that, rescaling time and space acccording to

$$x = t^\alpha y \qquad t = \exp(\tau),$$

the rescaled solution (which we will call $\hat{\rho}$ in the following) approaches the Bahrenblatt-Prattle profiles for large times.

### 3.2.1 Gradient Flow Formuation

The setting in which the gradient flow will be established is the following. First, a differentiable manifold $\mathcal{M}$ is defined by

$$\mathcal{M} = \left\{ \text{non-negative functions } \rho \text{ on } \mathbb{R}^N \text{ with } \int \rho = 1 \right\}. \tag{3.12}$$

Furthermore, the tangent space is given by

$$T_\rho\mathcal{M} = \left\{ \text{functions } s \text{ on } \mathbb{R}^N \text{ with } \int s = 0 \right\}. \tag{3.13}$$

To define a metric tensor, we use the following identification of the trangent space.

$$T_\rho\mathcal{M} = \left\{ \text{functions } p \text{ on } \mathbb{R}^N \ \right\} / \sim \tag{3.14}$$

where the identification between $p$ and $s$ is via

$$-\mathrm{div}(\rho \nabla p) = s. \tag{3.15}$$

The $\sim$ means that $p$'s, which differ only by a constant are indentified. Then, the metric tensor is defined as

$$g_\rho(s_1, s_2) = \int \rho \nabla p_1 \nabla p_2. \tag{3.16}$$

We define as a gradient flow on this manifold $\mathcal{M}$ the dynamical system given by

$$\frac{d\rho}{dt} = -\text{grad}E_{|\rho} \tag{3.17}$$

Now, we want to represent the porous medium equation as a gradient flow on this manifold. Therefore, we define the Energy functional

$$E(\rho) = \left\{ \begin{array}{ll} \frac{1}{m-1}\int \rho^m & \text{for } m \neq 1 \\ \int \rho \ln(\rho) & \text{for } m = 1 \end{array} \right\}. \tag{3.18}$$

Thus the differential is given by

$$\text{diff}E(\rho) \cdot s = \left\{ \begin{array}{ll} \int \frac{m}{m-1}\rho^{m-1}s & \text{for } m \neq 1 \\ \int (\ln(\rho)+1)s & \text{for } m = 1 \end{array} \right\} \tag{3.19}$$

Using the definition of the gradient flow on $\mathcal{M}$, we get

$$\left\{ \begin{array}{ll} \int \rho_t p + \int \frac{m}{m-1}\rho^{m-1}s = 0 & \text{for } m \neq 1 \\ \int \rho_t p + \int (\ln(\rho)+1)s = 0 & \text{for } m = 1 \end{array} \right\} \tag{3.20}$$

using equation (3.15), this leads to

$$\left\{ \begin{array}{ll} \int \rho_t p + \int \frac{m}{m-1}\rho^{m-1}\nabla \cdot (\rho \nabla p) = 0 & \text{for } m \neq 1 \\ \int \rho_t p + \int (\ln(\rho)+1)s\nabla \cdot (\rho \nabla p) = 0 & \text{for } m = 1 \end{array} \right\} \tag{3.21}$$

and finally, integrating by parts

$$\int (\rho_t - \Delta \rho^m)p = 0. \tag{3.22}$$

As $p$ is arbitrary, we recover the porous medium equation.
This calculation shows that the porous mediums equation can indeed be understood as a gradient flow on a metric space $\mathcal{M}$ endowed with the metric tensor as defined in 3.16.

### 3.2.2 Formal Derivation of Convergence Rates to Stationary Solutions

In [Otto01, Section 3.2] Otto states three inequalities that show convergence to the stationary state, which he later proofs by formal Riemannian calculus. To do so, we assume the following conditions:

i) $\hat{\rho}$ evolves according to the gradient flow

$$\frac{d\hat{\rho}}{d\tau} = \text{grad}F_{|\hat{\rho}} \tag{3.23}$$

on $(\mathcal{M}, g)$, where the augmented functional $F$ is given by

$$F(\hat{\rho}) = E(\hat{\rho}) + \alpha \int \frac{1}{2}|y|^2 \hat{\rho}(y)dy. \tag{3.24}$$

ii) $\hat{\rho}_*$ is a minimizer of $F$ on $\mathcal{M}$, i.e.

$$F(\hat{\rho}) \geq F(\hat{\rho}_*) \quad \forall \hat{\rho} \in \mathcal{M}. \tag{3.25}$$

This implies also that

$$-\text{grad}F_{|\hat{\rho}_*} = 0. \tag{3.26}$$

iii) $F$ is uniformly strictly (wirklich) convex on $(\mathcal{M}, g)$, i.e.

$$\text{Hess}F_{|\hat{\rho}} \leq \alpha \text{Id} \quad \forall \hat{\rho} \in \mathcal{M}, \tag{3.27}$$

in the sense of

$$\left\langle s, \text{Hess}(F_{|\hat{\rho}})s \right\rangle \geq \alpha |s|^2 \quad \forall s \in T_{\hat{\rho}}\mathcal{M}, \quad \forall \hat{\rho} \in \mathcal{M}. \tag{3.28}$$

This follows from

$$\text{Hess}E_{|\hat{\rho}} \geq 0 \text{ and } \text{Hess}M_{|\hat{\rho}} = \text{Id} \quad \forall \hat{\rho} \in \mathcal{M}. \tag{3.29}$$

Note that the ealier condition $m \geq 1 - \frac{1}{N}$ ensures the convexity of $E$ whereas the condition $m \geq \frac{N}{N+2}$ ensures that both $E(\hat{\rho}_*)$ and $M(\hat{\rho}_*)$ are well-defined and finite.

Using these assumptions, Otto derives

$$\frac{d}{d\tau}\left(\exp(2\alpha\tau)|\text{grad}F_{|\hat{\rho}}|^2\right) \leq 0, \tag{3.30}$$

$$\frac{d}{d\tau}\left(\exp(2\alpha\tau)(F(\hat{\rho}) - F(\hat{\rho}_\infty))\right) \leq 0, \tag{3.31}$$

$$\frac{d}{d\tau}\left(\exp(2\alpha\tau)d(\hat{\rho}, \hat{\rho}_\infty)^2\right) \leq 0. \tag{3.32}$$

All these equation mean that the rate of convergence towards $\hat{\rho}_\infty$ is exponential with a rate $\alpha$ with respect to $\tau$ and with a polynomial rate with respect to $t$. In the first equation, the distance between the stationary solution and the solution and $\hat{\rho}$ is measured. In the second equation, the distance of $\hat{\rho}$ from a minimizer of $F$ is measured and in the third the distance between $\hat{\rho}$ and $\hat{\rho}_*$.
The important fact is, that Otto was able to derive all these equations by relatively simple calculus on the Riemannian manifold $M$. We illustrate this for the first equation, which is the easiest, only.

$$\begin{aligned}
\frac{d}{d\tau}|\mathrm{grad}F_{|\hat{\rho}}|^2 &= 2\left\langle \mathrm{grad}F_{|\hat{\rho}}, \frac{D}{D\tau}\mathrm{grad}F_{|\hat{\rho}} \right\rangle \\
&= -2\left\langle \mathrm{grad}F_{|\hat{\rho}}, \mathrm{Hess}F_{|\hat{\rho}}\mathrm{grad}F_{|\hat{\rho}} \right\rangle \\
&\leq -2\alpha|\mathrm{grad}F_{|\hat{\rho}}|^2.
\end{aligned}$$

Using some sort of Gronwall-type inequality, this directly leads to 3.30.

## 3.3 Gradient flow for Systems

In this chapter, we try to generalize Otto's ideas to systems of diffusive equations. Therefore, we first show that the equation (or the system of equations) we consider can also be seen as a gradient flow. However, we need to define a more complicated metric tensor on this manifold. Then, we try to use displacement convexity to prove convergence to stationary solutions.

### 3.3.1 Introduction

In the following, we will use the methods described above to analyse the equation

$$\rho_t = \mathrm{div}(D(\rho)(\nabla u(\rho) + \nabla V(x))), \quad i = 1, \dots, n \quad x \in \mathbb{R}^N, t > 0 \qquad (3.33)$$

with initial condition $\rho(0) = \rho_0$.
We will distinguish between the scalar and and the vector valued case. In the scalar case, $\rho(x,t) \in \mathbb{R}$ and $D(\rho)$ is a real valued function of $\rho$. In the vector valued case, $\rho = (\rho_0, \dots, \rho_n) \in \mathbb{R}^N$ and $D$ is a matrix-valued function of $\rho$. In this case $D$ is assumed to be symmetric and positive definite. The

functions $u$ and $\rho$ are linked by the algebraic relation $\rho = \rho(u)$. $u$ and the potential $V(x)$ are also either scalar or vector-valued.

#### Examples

Setting $n = 1$, $D(\rho) = \rho$, $u(\rho) = \log \rho$, the system turns into the heat equation

$$\rho_t = \Delta \rho. \tag{3.34}$$

For $n = 2$, $\rho = (\rho_1, \rho_2)$ and $u = (u_1, u_2)$ are related by

$$\rho_1 = (-u_2)^{-3/2} e^{u_1} \qquad \rho_2 = -\frac{3}{2}\frac{\rho_1}{u_2} \tag{3.35}$$

the system turns into the energy-transport-equations.

#### The scalar case and the Otto formulation

We show that in the scalar case without potential, equation (3.33) is equivalent to the porous medium equation in Otto's formulation.

$$\rho_t = \operatorname{div}(\nabla f(\rho)) = \nabla \cdot (f'(\rho)\nabla \rho). \tag{3.36}$$

For the system formulation

$$\rho_t = \operatorname{div}(D(\rho)\nabla h'(\rho)) = \nabla \cdot (D(\rho)h''(\rho)\nabla \rho), \qquad u = h' \tag{3.37}$$

holds. Setting

$$f'(\rho) = D(\rho)u''(\rho) \tag{3.38}$$

the two formulations are equal for the scalar case.

**Remark** Assuming $D$ to be positive, (3.38) means that monotonicity of $f$ ($f' > 0$) and convexity of $u$ ($u'' > 0$) are equivalent.

### 3.3.2 Gradient flow interpretation

In this section, we show that in the vector valued case without a potential, the equation can still be considered as a gradient flow.

Let

$$M = \left\{ \rho_i \geq 0,\ i = 1 \ldots N \mid \int_{\mathbb{R}^d} \rho_i \, dx = 1 \right\} \tag{3.39}$$

be a manifold and

$$T_\rho M = \left\{ \rho_i \geq 0,\ i = 1 \dots N \mid \int_{\mathbb{R}^d} \rho_i \, dx = 0 \right\} \tag{3.40}$$

be the tangent space at $\rho$. For a given $\rho$ define the metric tensor

$$g_\rho(p, q) = \int_{\mathbb{R}^d} \nabla v^T D(\rho) \nabla u \, dx, \qquad p, q \in T_\rho M \tag{3.41}$$

where $u_i$ and $v_j$ are the solutions of

$$-\mathrm{div}(D(\rho)\nabla u) = p, \qquad -\mathrm{div}(D(\rho)\nabla v) = q,$$

$p = (p_1, \dots, p_n)$, $q = (q_1, \dots, q_n)$, etc.
We assume that the functional setting is such that the above elliptic systems are uniquely solvable. Note that the metric tensor can be rewritten as

$$g_\rho(p, q) = \int_{\mathbb{R}^d} p \cdot v dx = \int_{\mathbb{R}^d} q \cdot u \, dx \tag{3.42}$$

In the following lemma we prove that, using these definitions, (3.33) can in fact be interpreted as a gradient flow on $M$.

**Lemma 3.3.1** *Assume that the setting is such that the formula $g_\rho(\mathrm{grad}E, p) = dE \cdot p$ for all vector fields $p$ on $M$ holds. Furthermore, suppose that there exists a function $h : \mathbb{R}^n \to \mathbb{R}$ such that $u(\rho) = h'(\rho)$. Then (3.33) can be written as*

$$\rho_t = -\mathrm{grad}E|_\rho \tag{3.43}$$

*where the entropy $E$ is given by*

$$E(\rho) = \int_{\mathbb{R}^d} h(\rho) \, dx \tag{3.44}$$

**Proof** This is a slight generalization of Otto's argument [Otto01]. Let $p = (p_1, \dots, p_n) \in T_\rho M$ be a vector field and $v$ be a vector-valued function defined by $p = -\mathrm{div}(D(\rho)\nabla v)$.

Then

$$\begin{aligned}
g_\rho(\rho_t, p) + dE \cdot p &= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + h'(\rho) \cdot p\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v - h'(\rho)\nabla \cdot (D(\rho)\nabla v)\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + (\nabla h'(\rho)) : (D(\rho)\nabla v)\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + (D(\rho)\nabla h'(\rho)) : \nabla v\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v - \nabla \cdot (D(\rho)\nabla h'(\rho))v\right] \, dx \\
&= \int_{\mathbb{R}^d} \underbrace{[\rho_t - \nabla \cdot (D(\rho)\nabla u(\rho)]}_{=0,\ (5.15)} v \, dx \\
&= 0 \qquad \forall v \in \mathbb{R}^d
\end{aligned}$$

Thus, $g_\rho(\rho_t, p) = -g_\rho(\mathrm{grad} E|_\rho, p)$ for all $p$ and the conclusion follows.
□

**Remark** i) This calculation includes the scalar case for $d = 1$.

ii) Adding a potential, this calculation still works

$$\begin{aligned}
g_\rho(\rho_t, p) + dE \cdot p &= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + (u'(\rho) + V) \cdot p\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v - (u'(\rho) + V)\mathrm{div}(D(\rho)\nabla v)\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + \nabla(h'(\rho) + V) : (D(\rho)\nabla v)\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v + D(\rho)\nabla(h'(\rho) + V) : \nabla v\right] \, dx \\
&= \int_{\mathbb{R}^d} \left[\rho_t \cdot v - \mathrm{div}(D(\rho)\nabla(h'(\rho) + V))v\right] \, dx \\
&= \int_{\mathbb{R}^d} \underbrace{[\rho_t - \mathrm{div}(D(\rho)\,(\nabla u(\rho) + \nabla V))]}_{=0,\ (5.15)} v \, dx \\
&= 0 \qquad \forall v \in \mathbb{R}^d
\end{aligned}$$

A consequence of the above lemma is that the entropy is decreasing along the trajectories

$$\frac{dE}{dt} = dE|_\rho \cdot \rho_t = g_\rho(\mathrm{grad}E, \rho_t) = -g_\rho(\rho_t, \rho_t) \leq 0. \tag{3.45}$$

**Examples**

Let $n = 1$, $D_{ij}(\rho) = \rho$, and $u(\rho) = \frac{\alpha}{\alpha-1}\rho^{\alpha-1}$, where $\alpha > 0$ (this is the slow ($\alpha > 1$ or fast $\alpha \leq 1$ diffusion case). Then $h(\rho) = \frac{1}{\alpha-1}\rho^\alpha$ and we are in the setting of [Otto01].

### 3.3.3 Distance and Geodesics on $(\mathcal{M}, g)$

To get a distance on the Space $(\mathcal{M}, g)$, we generalize the idea of Benamou and Brenier (2.27) and define

$$d_\rho^2(\rho_0, \rho_T) = \frac{1}{2} \inf_{(\rho,v)} \int_0^T \int_{\mathbb{R}^N} \mathrm{tr}(v^T D(\rho) v)\, dxdt \tag{3.46}$$

under the constraints

$$\mathrm{div}(D(\rho)v) = \rho_t, \qquad \rho(0) = \rho_0, \quad \rho(T) = \rho_T \tag{3.47}$$

Again, we use the method of Lagrange multipliers to formally derive the optimality conditions for this minimization problem.

**Lemma 3.3.2** *The formal solution of the constrained minimization problem stated above is given by $v = -\nabla\lambda$ and*

$$\lambda_t = \frac{1}{2} \sum_{k,l,m} \nabla_\rho D_{kl}(\rho) v_{km} v_{lm}, \quad \lambda(0) = -\mu_0, \lambda(T) = \mu_T, \tag{3.48}$$

$$\mathrm{div}(D(\rho)v) = \rho_t, \quad \rho(0) = \rho_0, \quad \rho(T) = \rho_T. \tag{3.49}$$

**Proof** We start with the minimization problem. The Lagrangian is by definition given as

$$
\begin{aligned}
\mathcal{L} &= \frac{1}{2}\int_0^T \int_{\mathbb{R}^N} \operatorname{tr}(v^T D(\rho) v) + \lambda \cdot \rho_t - \lambda \cdot \operatorname{div}(D(\rho)v)\, dxdt \\
&+ \int_{\mathbb{R}^N} \mu_0(\rho(0) - \rho_0) + \mu_T(\rho(T) - \rho_T)\, dx \\
&= \frac{1}{2}\int_0^T \int_{\mathbb{R}^N} \sum_{k,l,m} D_{kl}(\rho) v_{km} v_{lm} + \lambda \cdot \rho_t - \lambda \cdot \operatorname{div}(D(\rho)v)\, dxdt \\
&+ \int_{\mathbb{R}^N} \mu_0(\rho(0) - \rho_0) + \mu_T(\rho(T) - \rho_T)\, dx \\
&=: \mathcal{L}_\lambda + \mathcal{L}_\mu.
\end{aligned}
$$

where $\lambda \in \mathbb{R}^N$ and $\mu = (\mu_0, \mu_T)$ are the Lagrange multiplier. We first calculate the derivative of $\mathcal{L}_\lambda$ with respect to $v$ into a arbitrary direction $w$.

$$
\begin{aligned}
\frac{\partial \mathcal{L}_\lambda}{\partial v} \cdot w &= \int_0^T \int_{\mathbb{R}^N} \sum_{k,l,m} D_{kl}(\rho) v_{km} w_{lm} + \sum_{k,l,m} D_{kl}(\rho) w_{lm} \frac{\partial \lambda_k}{\partial x_m}\, dxdt \\
&= \int_0^T \int_{\mathbb{R}^N} \sum_{k,l,m} D_{kl}(\rho)(v_{km} + \frac{\partial \lambda_k}{\partial x_m}) w_{lm}\, dxdt \quad \forall w.
\end{aligned}
$$

From this follows

$$
v = -\nabla \lambda. \tag{3.50}
$$

Next, the derivative with respect to $\rho$ into direction $g$ is given by

$$
\begin{aligned}
\frac{\partial \mathcal{L}_\lambda}{\partial \rho} \cdot g &= \frac{1}{2}\int_0^T \int_{\mathbb{R}^N} \sum_{k,l,m} (\nabla_\rho D_{kl}(\rho) \cdot g) v_{km} v_{lm} - \lambda_t \cdot g \\
&- \sum_{k,l,m} (\nabla_\rho D_{kl}(\rho) \cdot g) v_{km} \frac{\partial \lambda_l}{\partial x_m}\, dxdt \\
&= \int_0^T \int_{\mathbb{R}^N} -\frac{1}{2} \sum_{k,l,m} (\nabla_\rho D_{kl}(\rho) \cdot g) v_{km} v_{lm} - \lambda_t \cdot g\, dxdt \quad \forall g.
\end{aligned}
$$

From this follows, after partial integration with respect to time:

$$
\lambda_t = \frac{1}{2} \sum_{k,l,m} \nabla_\rho D_{kl}(\rho) v_{km} v_{lm}. \tag{3.51}
$$

For the constraint, we consider the derivative of $\mathcal{L}_\mu$. As the derivative with respect to $v$ is zero, we only look at

$$\begin{aligned}\left(\frac{\partial \mathcal{L}_\lambda}{\partial \rho} + \frac{\partial \mathcal{L}_\mu}{\partial \rho}\right) \cdot g &= \int_0^T \int_{\mathbb{R}^N} -\frac{1}{2} \sum_{k,l,m} (\nabla_\rho D_{kl}(\rho) \cdot g) v_{km} v_{lm} - \lambda_t \cdot g \, dxdt \\ &+ \int_{\mathbb{R}^N} (\mu_0 \cdot g(0) + \mu_T \cdot g(t)) \, dx.\end{aligned}$$

Partial integration with respect to time and (3.51) finally reads to

$$\lambda(0) = -\mu_0, \qquad \lambda(T) = \mu_T. \tag{3.52}$$

□

**Remark** In contrast to Benamou and Brenier, in this case the two equations (3.48) and (3.49) are coupled.

# Chapter 4

# Entropy Dissipation Methods

In this chapter, so-called entropy methods for PDEs will be discussed. Especially, their use to obtain convercence rates towords stationary sulutions will be regarded. We will mainly follow the discussion in [Carillo01]. For a more general review on the connection between PDEs and entropy as well as for the physical meaning of the entropy, see [Evans04].
The strategy will be the following. After defining an entropy functional, one shows that this functional is decreasing in time. Furthermore, it can be shown that the asymptotic solution is a minimizer of the entropy functional. Theorefore, one uses the relative entropy to measure the distance between a sulution at time $t$ and the stationary solution. Finally, one uses a Csiszar-Kullback [Csisz'ar67, Kullback59] type inequality to proof a convergence rate in $L^1$.

## 4.1 Principle

To demonstrate the method, we consider the scalar, nonlinear Fokker-Planck equation

$$u_t = \operatorname{div}(u\nabla V(x) + \nabla f(u)), \quad x \in \mathbb{R}^N, t > 0, \tag{4.1}$$
$$u(x, t = 0) = u_0(x) \geq 0, \quad x \in \mathbb{R}^N. \tag{4.2}$$

We assume that $f$ and $V$ are such that there exists a unique stationary solution $u_\infty$. Furthermore, we assume $f'$ and $u$ to be positive, so that the equation is parabolic. Note that the equation is mass preserving, i.e.

$$\int u(x,t)dx = \int u_0(x)dx =: M. \tag{4.3}$$

This is due to the fact that it is in divergence form. In the following, we only consider the simple case $f(u) = u$.

### Entropy

It can be checked that for every convex function $\phi$ in $\mathbb{R}$, the equation

$$E(u(t)) = \int \phi(u(t))dx \tag{4.4}$$

defines a Lyaponov functional along the solutions of the PDE, i.e. its derivative with respect to $t$ is the negative of a positive functional. Here, we define the entropy making the choice

$$\phi(x) := u(\log(u) - 1) + 1. \tag{4.5}$$

Then, the entropy itself is defined by

$$E(u(t)) = \int \phi(u(x,t)) + u(x,t)V(x)dx. \tag{4.6}$$

Assuming $V$, $u$ and $\phi(u)$ to be in $L^1(\mathbb{R}^N)$, we can compute the first derivative in time.

$$\begin{aligned}
\frac{dE(u(t))}{dt} &= \int u_t(\log(u) + V(x))dx \\
&= \int \operatorname{div}(u\nabla V(x) + \nabla u)(\log(u) + V(x))dx \\
&= -\int \left[u\nabla V(x) + \nabla u \left(\frac{\nabla u}{u}\right) + \nabla V(x)\right] dx \\
&= -\int u(x,t) \left|\nabla V(x) + \frac{\nabla u}{u}\right|^2 dx
\end{aligned}$$

This gives an expression for the dissipation of the entropy and

$$I(u(t)) = \int u(x,t) \left|\nabla V(x) + \frac{\nabla u}{u}\right|^2 dx \tag{4.7}$$

is called entropy production functional. This implies that the Entropy is a Lyaponov functional (along solutions) since $u \geq 0$. The next step in the stategy outlined above is to differentiate the entropy production functional. If the resulting expression is of the form such that for $\lambda \leq 0$,

$$\frac{dI(u(t))}{dt} = -\lambda I(u(t)) - R(t), \tag{4.8}$$

with some reminder $R(t) \geq 0$ on $\mathbb{R}^N$, using (4.7), we imediately get

$$I(u(t)) = \frac{dE(u(t))}{dt} = \frac{1}{\lambda}\left[\frac{dI(u(t))}{dt} + R(t)\right]. \tag{4.9}$$

Using the fact that $R(t)$ is positive, this leads to

$$\frac{dI(u(t))}{dt} \leq \lambda I(u(t)). \tag{4.10}$$

This implies (using Gronwall's inequality)

$$I(u(t)) \leq e^{-\lambda t} I(u(0)) \tag{4.11}$$

Because the stationary solution of is a minimizer of the entropy functional, the value of the entropy production functional at the stationary solution is zero. Therefore, integrating the previous expression from zero to infinity and using the fact that $R(t)$ is positive,

$$E(u_0) - E(u_\infty) \leq \frac{1}{\lambda} I(u_0) \tag{4.12}$$

and.

$$\frac{d[E(u(t)) - E(u_\infty)]}{dt} \leq -\lambda[E(u(t)) - E(u_\infty)] \tag{4.13}$$

hold. Introducing the relative entropy $E(u(t)|u_\infty)$ by

$$E(u(t)|u_\infty) = E(u(t)) - E(u_\infty), \tag{4.14}$$

the last equation can be written as

$$E(u(t)|u_\infty) \leq I(u(t)|u_\infty). \tag{4.15}$$

This equation is called log-Sobolev inequality. This is exponential convergence of the entropy functional towards its minumum at a rate $\lambda$, using Gronwalls lemma [Gronwall19] Using a Csiszar-Kullback type inquality, the final result is convergence in $L^1$ at a rate of $\lambda/2$.

# Chapter 5

# Application to Systems of Diffusive PDEs

In this Chapter both the entropy dissipation and the gradient flow method will be used to examine the long-time behaviour of nonlinear Fokker-Planck equations of the type

$$\rho_t = \operatorname{div}(D(\rho)\left(\nabla u'(\rho) + \nabla V\right)). \tag{5.1}$$

There $\rho = \rho(x,t)$, the potential $V$ depends only on $x$ and $'$ denotes the derivative with respect to $\rho$. First, the case of a scalar non-linearity $D(\rho)$ will be treated. This leads to different results for both methods. To explain these differences, both methods will be appled to the fully one-dimensional case.
Finally, system case will be treated but only for the example of energy transport.

## 5.1 Entropy Dissipation Methods

To use the entropy dissipation method we first define the entropy functional as

$$E(\rho) = \int u(\rho) + \rho V(x) dx. \tag{5.2}$$

We will examine the derivatives of this functional along the solutions of the PDE.

### 5.1.1 The Scalar Case

We start showing that the first derivative of $E$ (along solutions of (5.15)) is the negative of some entropy production functional.

$$\begin{aligned}\frac{dE}{dt} &= \int \rho_t(u'(\rho)+V(x))\,dx \\ &= -\int D(\rho)|\nabla u'(\rho)+\nabla V(x)|^2\,dx \\ &= -\int D(\rho)|\xi|^2\,dx,\end{aligned}$$

with

$$\xi = \nabla u'(\rho)+\nabla V.$$

Next, we study the derivative of this production functional (i.e. second derivative of the entropy).

$$\frac{dI}{dt}(u(t)) = -\frac{d^2E}{dt^2} = \int \rho_t D'(\rho)|\xi|^2 dx + 2\int D(\rho)\xi\cdot\frac{\partial \xi}{\partial t}dx =: I_1+I_2. \quad (5.3)$$

For the second term we get, after partial integration,

$$I_2 = -2\int \operatorname{div}(D(\rho)\xi)\frac{\partial u'}{\partial t}(\rho)dx = -2\int u''(\rho)(\operatorname{div}(D(\rho)\xi))^2 dx. \quad (5.4)$$

For the first term we get

$$\begin{aligned}I_1 &= \int \operatorname{div}(D(\rho)\xi)D'(\rho)|\xi|^2 dx = -\int D(\rho)\xi\cdot\nabla\left(D'(\rho)|\xi|^2\right)dx \\ &= -\int D(\rho)D''(\rho)|\xi|^2(\xi\cdot\nabla\rho)dx - 2\int D(\rho)D'(\rho)\xi^T\operatorname{Jacob}(\xi)\xi dx \\ &=: I_1^{(1)}+I_1^{(2)}.\end{aligned}$$

The second term can we written as

$$I_1^{(2)} = -2\int D(\rho)D'(\rho)\xi^T\operatorname{Hess}(V)\xi dx - 2\int D(\rho)D'(\rho)\xi^T\operatorname{Hess}(u'(\rho))\xi dx. \quad (5.5)$$

Our goal now is to apply Bochner's formula (see below) and therefore we will have to rewrite the second part of (5.5). Assuming $D(\rho)\neq 0$, we can do the following calculation.

$$
\begin{aligned}
- & \; 2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(u'(\rho))\xi dx \\
= & \; -2\sum_{i,j=1}^{N}\int D'(\rho)(D(\rho))^{-1}(D(\rho)\xi_i)(D(\rho)\xi_j)\frac{\partial^2 u'(\rho)}{\partial x_i \partial x_j} \\
= & \; -2\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}(D'(\rho))^2(D(\rho))^{-2}\frac{\partial \rho}{\partial x_j}(D(\rho)\xi_i)(D(\rho)\xi_j) \\
+ & \; 2\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}D'(\rho)(D(\rho))^{-1}\frac{\partial\left[(D(\rho)\xi_i)(D(\rho)\xi_j)\right]}{\partial x_j} \\
+ & \; \underbrace{2\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}\frac{\partial \rho}{\partial x_j}D''(\rho)(D(\rho))^{-1}(D(\rho)\xi_i)(D(\rho)\xi_j)}_{=:I_1^{(3)}} \\
= & \; I_1^{(3)} - 2\sum_{i,j=1}^{N}\int (D'(\rho))^2 u''(\rho)\frac{\partial \rho}{\partial x_i}\frac{\partial \rho}{\partial x_j}\xi_i\xi_j \\
+ & \; 2\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}D'(\rho)\left(\xi_i\frac{\partial(D(\rho)\xi_j)}{\partial x_j} + \xi_j\frac{\partial(D(\rho)\xi_i)}{\partial x_j}\right) \\
= & \; I_1^{(3)} - 2\sum_{i,j=1}^{N}\int (D'(\rho))^2 u''(\rho)\frac{\partial \rho}{\partial x_i}\frac{\partial \rho}{\partial x_j}\xi_i\xi_j \\
+ & \; 4\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}D'(\rho)\xi_i\frac{\partial(D(\rho)\xi_j)}{\partial x_j} \\
- & \; 2\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}D'(\rho)\left(\xi_i\frac{\partial(D(\rho)\xi_j)}{\partial x_j} - \xi_j\frac{\partial(D(\rho)\xi_i)}{\partial x_j}\right) \\
= & \; -2\sum_{i,j=1}^{N}\int (D'(\rho))^2 u''(\rho)\frac{\partial \rho}{\partial x_i}\frac{\partial \rho}{\partial x_j}\xi_i\xi_j \\
+ & \; 4\sum_{i,j=1}^{N}\int \frac{\partial u'(\rho)}{\partial x_i}D'(\rho)\xi_i\frac{\partial(D(\rho)\xi_j)}{\partial x_j} \\
- & \; 2\int D(\rho)D'(\rho)u''(\rho)\left[(\xi\cdot\nabla\rho)\,\mathrm{div}(\xi) - \frac{1}{2}\nabla(|\xi|^2)\cdot\nabla\rho\right]dx + I_1^{(3)}
\end{aligned}
$$

We now collect all the above terms. Furthermore, we define $f$ via $f' = D(\rho)u''(\rho)$. Hence,

$$\begin{aligned} I_1 + I_2 &= I_1^{(1)} + I_1^{(3)} - 2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx - 2\int D(\rho)f'(\rho)(\mathrm{div}(\xi))^2 dx \\ &- 2\int D'(\rho)f'(\rho)\left[(\xi\cdot\nabla\rho)\,\mathrm{div}(\xi) - \frac{1}{2}\nabla(|\xi|^2)\cdot\nabla\rho\right] dx. \end{aligned}$$

This can be rewritten as

$$\begin{aligned} I_1 + I_2 &= I_1^{(1)} + I_1^{(3)} - 2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx - 2\int D(\rho)f'(\rho)(\mathrm{div}(\xi))^2 dx \\ &- 2\int D'(\rho)\left[(\xi\cdot\nabla f(\rho))\,\mathrm{div}(\xi) - \frac{1}{2}\nabla(|\xi|^2)\cdot\nabla f(\rho)\right] dx. \end{aligned}$$

Then, partial integration leads to

$$\begin{aligned} I_1 + I_2 &= I_1^{(1)} + I_1^{(3)} - 2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx \\ &- 2\int (D(\rho)f'(\rho) - D'(\rho)f(\rho))(\mathrm{div}(\xi))^2 dx \\ &- 2\int D'(\rho)f(\rho)\left[\frac{1}{2}\Delta(|\xi|^2) - (\xi\cdot\nabla\mathrm{div}(\xi))\right] dx \\ &- 2\int D''(\rho)f(\rho)\left[\frac{1}{2}\nabla(|\xi|^2)\cdot\nabla\rho - (\xi\mathrm{div}(\xi)\cdot\nabla\rho)\right] dx. \end{aligned}$$

Now, we can really use Bochner's formula (for a space with Ricci curvature zero),

$$\frac{1}{2}\Delta(|\xi|^2) - (\xi\cdot\nabla\mathrm{div}(\xi)) = \mathrm{tr}(\nabla\xi)^2, \tag{5.6}$$

which results in

$$\begin{aligned} I_1 + I_2 &= -2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx - \int D(\rho)D''(\rho)|\xi|^2(\xi\cdot\nabla\rho) dx \\ &- 2\int (D(\rho)f'(\rho) - D'(\rho)f(\rho))(\mathrm{div}(\xi))^2 dx - 2\int D'(\rho)f(\rho)\mathrm{tr}(\nabla\xi)^2 dx \\ &- 2\int D''(\rho)f(\rho)\left[\frac{1}{2}\nabla(|\xi|^2)\cdot\nabla\rho - (\mathrm{div}(\xi)\xi\cdot\nabla\rho)\right] dx \\ &+ 2\int D(\rho)D''(\rho)u''(\rho)|\xi|^2|\nabla\rho|^2 dx. \end{aligned}$$

Furthermore

$$\begin{aligned}
I_1 + I_2 &= -2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx - \int D(\rho)D''(\rho)|\xi|^2(\xi\cdot\nabla\rho)dx \\
&- 2\int (D(\rho)f'(\rho) - D'(\rho)f(\rho))(\mathrm{div}(\xi))^2 dx - 2\int D'(\rho)f(\rho)\mathrm{tr}(\nabla\xi)^2 dx \\
&- 2\int D''(\rho)f(\rho)(\nabla u' + \nabla V)^T(\mathrm{Hess}(\frac{du}{d\rho}(\rho)) + \mathrm{Hess}(V))\nabla\rho dx \\
&+ 2\int D''(\rho)f(\rho)\mathrm{div}(\xi)\left(\nabla u'\cdot\nabla\rho + \nabla V\cdot\nabla\rho\right)dx \\
&+ 2\int D(\rho)D''(\rho)u''(\rho)|\xi|^2|\nabla\rho|^2 dx,
\end{aligned}$$

and

$$\begin{aligned}
I_1 + I_2 &= -2\int D(\rho)D'(\rho)\xi^T \mathrm{Hess}(V)\xi dx - \int D(\rho)D''(\rho)|\xi|^2(\xi\cdot\nabla\rho)dx \\
&- 2\int (D(\rho)f'(\rho) - D'(\rho)f(\rho))(\mathrm{div}(\xi))^2 dx - 2\int D'(\rho)f(\rho)\mathrm{tr}(\nabla\xi)^2 dx \\
&- 2\int D''(\rho)f(\rho)(u''\nabla\rho^T \mathrm{Hess}(u'(\rho))\nabla\rho + u''\nabla\rho^T\mathrm{Hess}(V)\nabla\rho)dx \\
&- 2\int D''(\rho)f(\rho)(\nabla\rho^T\mathrm{Hess}(u'(\rho))\nabla V + \nabla\rho^T\mathrm{Hess}(V)\nabla V)dx \\
&+ 2\int D''(\rho)f(\rho)\mathrm{div}(\xi)\left(u''|\nabla\rho|^2 + \nabla V\cdot\nabla\rho\right)dx \\
&+ 2\int D(\rho)D''(\rho)u''(\rho)|\xi|^2|\nabla\rho|^2 dx.
\end{aligned}$$

Setting $V=0$, which implies $\xi = \nabla u'$ we get

$$\begin{aligned}
I_1 + I_2 &= -2\int (D(\rho)f'(\rho) - D'(\rho)f(\rho))(\mathrm{div}(\nabla u'))^2 dx - 2\int D'(\rho)f(\rho)\mathrm{tr}(\nabla^2 u')^2 dx \\
&- 2\int D''(\rho)f(\rho)u''(\rho)(\nabla\rho^T\mathrm{Hess}(u'(\rho))\nabla\rho)dx \\
&+ 2\int D''(\rho)f(\rho)\mathrm{div}(\nabla u')u''|\nabla\rho|^2 dx \\
&+ \int D(\rho)D''(\rho)u''(\rho)|\xi|^2|\nabla\rho|^2 dx.
\end{aligned}$$

## 5.2 Gradient Flow and Displacement Convexity

We will now use the method of displacement convexity of examine the asymptotic behaviour of solutions of (5.15).

### 5.2.1 Linear Scalar Case

First, we consider the case of a scalar $\rho \in \mathbb{R}$ and $D(\rho) = \rho$. This means, we recover the porous medium equation. Following [Otto05], we write the equation in the form

$$\rho_t = \text{div}(\nabla u'(\rho) + \nabla V). \tag{5.7}$$

and define the Entropy by

$$E(\rho) = \int u(\rho) + \rho V(x) dx =: E_{int} + E_{pot}. \tag{5.8}$$

First, we take care of the internal energy. Recall that displacement convexity means convexity along geodesics which are parametrised by $s$ here. First derivative is

$$\frac{dE}{ds} = \int u'(\rho)\rho_s = -\int u'' \nabla\rho \cdot \rho v = -\int \nabla f(\rho) \cdot v, \tag{5.9}$$

when $f$ is defined via

$$f'(\rho) = \rho u''(\rho).$$

The Second derivative turns out to be

$$\begin{aligned}
\frac{d^2E}{ds^2} &= \int_{\mathbb{R}^N} f'(\rho)\text{div}(\rho v)\text{div}(v) + f(\rho)\text{div}(v_s) dx \\
&= \int_{\mathbb{R}^N} \rho f'(\rho)(\text{div}(v))^2 + \nabla f(\rho) v \text{div}(v) + f(\rho)\Delta\left(\frac{1}{2}|v|^2\right) dx \\
&= \int_{\mathbb{R}^N} \rho f'(\rho)(\text{div}(v))^2 + f(\rho)\left\{-\text{div}(v\text{div}(v)) + \Delta\left(\frac{1}{2}|v|^2\right)\right\} dx \\
&= \int_{\mathbb{R}^N} \rho f'(\rho)(\text{div}(v))^2 + f(\rho)\left\{-(\text{div}(v))^2 - v \cdot \nabla\text{div}(v) + \Delta\left(\frac{1}{2}|v|^2\right)\right\} dx \\
&= \int_{\mathbb{R}^N} \left(\rho f'(\rho) - f(\rho)\right)(\text{div}(v))^2 + f(\rho)\left\{-v \cdot \nabla\text{div}(v) + \Delta\left(\frac{1}{2}|v|^2\right)\right\} dx \\
&= \int_{\mathbb{R}^N} \left(\rho f'(\rho) - f(\rho)\right)(\text{div}(v))^2 + f(\rho) tr(Dv)^2 dx.
\end{aligned}$$

In this calculation we used

$$\operatorname{div}(v\operatorname{div}(v)) = (\operatorname{div}(v))^2 + v \cdot \nabla \operatorname{div}(v) \tag{5.10}$$

and (again) Bochner's formula for a space with Ricci curvature zero

$$\Delta \frac{1}{2}|v|^2 - v \cdot \nabla \operatorname{div}(v) = tr(Dv)^2. \tag{5.11}$$

Obviously, the second derivative is positive for $\rho \geq 0$ and

$$\rho f'(\rho) - \left(1 - \frac{1}{n}\right) f(\rho) \geq 0, \tag{5.12}$$

because of $\operatorname{div}(v)^2 \leq n\operatorname{tr}(Dv)^2$.

**Example**

As a simple example we show that for the choice

$$u'(\rho) = \log(\rho),$$

the above condition is always fulfilled. We get

$$f'(\rho) = 1 \Rightarrow f(\rho) = \rho.$$

Inserting this into (5.12) leads to the new condition

$$\frac{1}{n}\rho \geq 0,$$

which is always fulfilled for non-negative $\rho$.

**Potential Energy**

Adding a Potential to the equation does not alter the situation as long as the potential is convex. This is shown by the following calculation.
The part of the entropy corresponding to the potential is given by

$$E(\rho) = \int \rho V(x)\, dx. \tag{5.13}$$

The derivatives of this entropy functional are

$$\frac{\partial E}{\partial s} = -\int \rho v \nabla V \; dx \tag{5.14}$$

and

$$\begin{aligned}
\frac{\partial^2 E}{\partial s^2} &= -\int \rho_s v \nabla V \; dx - \int \rho v_s \nabla V \; dx \\
&= -\int \operatorname{div}(\rho v) v \nabla V \; dx - \int \rho \nabla \left( \frac{1}{2} |v|^2 \right) \nabla V \; dx \\
&= \int \rho v \nabla (v \nabla V) \; dx - \int \rho \nabla v \cdot v \nabla V \; dx \\
&= \int \rho v \nabla v \nabla V \; dx + \int \rho v^T \mathrm{Hess}(V) v \; dx - \int \rho \nabla v \cdot v \nabla V \; dx \\
&= \int \rho v^T \mathrm{Hess}(V) v \; dx.
\end{aligned}$$

This shows that the sign the second derivative is really given by the convexity properties of the potential as stated above.

### 5.2.2 Nonlinear Scalar Case with Potential

We start with the equation in the form

$$\rho_t = \operatorname{div} \big( D(\rho) \big( \nabla u'(\rho) + \nabla V \big) \big) , \tag{5.15}$$

where $\rho = \rho(x,t)$, the potential $V$ depends only on $x$ and $'$ denotes the derivative with respect to $\rho$. The Entropy is defined as

$$E(\rho) = \int u(\rho) + \rho V(x) dx. \tag{5.16}$$

Again we study the derivatives of the functional $E(\rho(s))$ along the geodesics which are parametrised via the parameter $s$.

$$\frac{dE}{ds} = \int_{\mathbb{R}^N} \rho_s u'(\rho) \; dx = -\int_{\mathbb{R}^N} D(\rho) v \nabla u'(\rho) \; dx = -\int_{\mathbb{R}^N} D(\rho) v \nabla(\rho) u''(\rho) \; dx. \tag{5.17}$$

Defining $f(\rho)$ via $f'(\rho) = D(\rho) u''(\rho)$ we obtain

$$\frac{dE}{ds} = -\int_{\mathbb{R}^N} \nabla f(\rho) v \; dx. \tag{5.18}$$

For the second derivative we get

$$\begin{aligned}
\frac{d^2E}{ds^2} &= -\int_{\mathbb{R}^N} \nabla(\rho_s f'(\rho))\cdot v - \int_{\mathbb{R}^N} \nabla(f(\rho))v_s \, dx \\
&= \int_{\mathbb{R}^N} \mathrm{div}(D(\rho)v)f'(\rho)\mathrm{div}(v)\, dx - \frac{1}{2}\int_{\mathbb{R}^N} \nabla f(\rho)\nabla\left(D'(\rho)|v|^2\right)\, dx \\
&= \int_{\mathbb{R}^N} D(\rho)f'(\rho)(\mathrm{div}(v))^2\, dx + \int_{\mathbb{R}^N} D'(\rho)\nabla\rho f'(\rho)v\mathrm{div}(v)\, dx \\
&- \frac{1}{2}\int_{\mathbb{R}^N} \nabla f(\rho)\nabla\left(D'(\rho)\right)|v|^2\, dx - \int_{\mathbb{R}^N} D'(\rho)\nabla\rho f'(\rho)\nabla\left(|v|^2\right)\, dx
\end{aligned}$$

Now defining $g$ via $g'(\rho) = D(\rho)D'(\rho)u''(\rho)$ we can write

$$\begin{aligned}
\frac{d^2E}{ds^2} &= \int_{\mathbb{R}^N} D(\rho)f'(\rho)(\mathrm{div}(v))^2\, dx - \int_{\mathbb{R}^N} g(\rho)\mathrm{div}(v\mathrm{div}(v))\, dx \\
&- \frac{1}{2}\int_{\mathbb{R}^N} \nabla f(\rho)\nabla\left(D'(\rho)\right)|v|^2\, dx - \int_{\mathbb{R}^N} g(\rho)\Delta\left(|v|^2\right)\, dx.
\end{aligned}$$

Using

$$\mathrm{div}(v\mathrm{div}(v)) = (\mathrm{div}(v))^2 + v\cdot\nabla\mathrm{div}(v) \tag{5.19}$$

we get

$$\begin{aligned}
\frac{d^2E}{ds^2} &= \int_{\mathbb{R}^N} \left\{D(\rho)f'(\rho) - g(\rho)\right\}(\mathrm{div}(v))^2\, dx + \int_{\mathbb{R}^N} \left\{\frac{1}{2}\Delta\left(|v|^2\right) - v\cdot\nabla\mathrm{div}(v)\right\} g(\rho)\, dx \\
&- \frac{1}{2}\int_{\mathbb{R}^N} D(\rho)u''(\rho)D''(\rho)|\nabla\rho|^2|v|^2 dx
\end{aligned}$$

and with Bochner's formula (again for a space with Ricci curvature zero)

$$\Delta\frac{1}{2}|v|^2 - v\cdot\nabla\mathrm{div}(v) = tr(\nabla v)^2. \tag{5.20}$$

we finally obtain

$$\begin{aligned}
\frac{d^2E}{ds^2} &= \int_{\mathbb{R}^N} \left\{D(\rho)f'(\rho) - g(\rho)\right\}(\mathrm{div}(v))^2 dx + \int_{\mathbb{R}^N} g(\rho)\mathrm{tr}(\nabla v)^2\, dx \\
&- \frac{1}{2}\int_{\mathbb{R}^N} D(\rho)u''(\rho)D''(\rho)|\nabla\rho|^2|v|^2\, dx.
\end{aligned}$$

Because of $\mathrm{div}(v)^2 \le n\mathrm{tr}(\nabla v)^2$, the first two terms are positive if $\rho \ge 0$ and

$$D(\rho)f'(\rho) - \left(1 - \frac{1}{n}\right)g(\rho) \ge 0. \tag{5.21}$$

The second term is only positive, if $D''(\rho) \leq 0$, assuming $D(\rho)$ and $u''(\rho)$ to be nonnegative.

### Examples

We will examine when the conditions for (5.12) will be fulfilled for three typical examples.

i) To check if our results are compatible with the linear case, we write the linear equation in the form

$$\rho_t = \operatorname{div}\big(D(\rho)\nabla u'(\rho)\big)$$

with $u'(\rho)$ defined via

$$u''(\rho) = \frac{1}{D(\rho)}.$$

This leads to

$$g'(\rho) = D(\rho)D'(\rho)u''(\rho) = D'(\rho),$$

and therefore

$$\frac{1}{n}D(\rho) \geq 0.$$

Again we see that this condition is always fulfilled (as long as $D$ is positive).

ii) As a second example we choose

$$D(\rho) = \rho^\alpha, \quad u'(\rho) = \log(\rho), \quad \alpha \geq 0.$$

This leads to

$$f'(\rho) = D(\rho)u''(\rho) = \alpha\rho^{\alpha-1},$$

and by integrating

$$f(\rho) = \frac{1}{\alpha}\rho^\alpha.$$

Inserting these formula's into (5.12), we obtain

$$\rho^\alpha\left(1 - \frac{1}{\alpha} + \frac{1}{\alpha n}\right) \geq 0.$$

As $\rho$ is always nonnegative for nonnegative initial data, this leads to the condition

$$\alpha \geq \left(1 + \frac{1}{n}\right).$$

iii) For the third example we choose

$$D(\rho) = \rho(1-\rho), \quad u'(\rho) = \log(\rho).$$

As above we get

$$g'(\rho) = (1-\rho)(1-2\rho) \Rightarrow g(\rho) = \frac{2}{3}\rho^3 - \frac{3}{2}\rho^2 + \rho,$$

and

$$\frac{D(\rho)}{D'(\rho)} = \frac{\rho(1-\rho)}{1-2\rho}.$$

If we put these two results in (5.12), the result is

$$\begin{aligned} &\frac{\rho(1-\rho)}{1-2\rho}(1-\rho)(1-2\rho) - \left(1 - \frac{1}{n}\right)\left(\frac{2}{3}\rho^3 - \frac{3}{2}\rho^2 + \rho\right) \geq 0 \\ \Rightarrow \quad &\rho\left(\rho^2\left(\frac{1}{3} + \frac{2}{3n}\right) - \rho\left(\frac{1}{2} + \frac{3}{2n}\right) + \frac{1}{n}\right) \geq 0 \end{aligned}$$

For large $\rho$ (i.e. for $\rho$ close to 1) this leads to the condition

$$\frac{1}{3} + \frac{2}{3n} - \frac{1}{2} - \frac{3}{2n} + \frac{1}{n} = -\frac{1}{6} + \frac{4-9+6}{6n} = -\frac{1}{6}\left(1 - \frac{1}{n}\right).$$

This term is negative for every $n > 1$.

### Potential Energy Term

As above, we examine the effect of an additional potential. Again, the entropy is defined as

$$P(\rho) = \int \rho V(x)\, dx. \tag{5.22}$$

The first and second derivative are

$$\frac{\partial P}{\partial s} = -\int D(\rho v)\nabla V\, dx, \tag{5.23}$$

and

$$\begin{aligned}
\frac{\partial^2 P}{\partial s^2} &= -\int \rho_s D'(\rho) v \nabla V \, dx - \int D(\rho) v_s \nabla V \, dx \\
&= -\int \operatorname{div}(D(\rho)v) D'(\rho) v \nabla V \, dx - \int D(\rho) \nabla \left( \frac{1}{2} D'(\rho) \left| v \right|^2 \right) \nabla V \, dx \\
&= \int D(\rho) v \nabla (D'(\rho) v \nabla V) \, dx - \int D(\rho) D'(\rho) \mathrm{Jacob}(v) v \nabla V \, dx \\
&- \frac{1}{2} \int D(\rho) \nabla \rho D''(\rho) |v|^2 \nabla V \, dx \\
&= \int D(\rho) |v|^2 \nabla \rho D''(\rho) \nabla V \, dx + \int D(\rho) v D'(\rho) \mathrm{Jacob}(v) \nabla V \, dx \\
&+ \int D(\rho) |v|^2 D'(\rho) \mathrm{Hess}(V) \, dx - \int D(\rho) D'(\rho) \mathrm{Jacob}(v) v \nabla V \, dx \\
&- \frac{1}{2} \int D(\rho) \nabla \rho D''(\rho) |v|^2 \nabla V \, dx \\
&= \frac{1}{2} \int D(\rho) |v|^2 \nabla \rho D''(\rho) \nabla V \, dx + \int D(\rho) |v|^2 D'(\rho) \mathrm{Hess}(V) \, dx.
\end{aligned}$$

The second term is nonnegative, if $D'(\rho) \geq 0$ and $\mathrm{Hess}(V) \geq 0$. However, nothing can be said about the sign of the first term, as the sign of $\nabla \rho$ is unknown and cannot be controlled by any other quantity.

### 5.2.3 System Case without Potential

In this section, we attempt to examine the displacement convexity of the entropy functional in the case when $D(\rho)$ is matrix valued. We will call this the system case. The entropy is still defined via

$$E = \int u(\rho) \, dx, \tag{5.24}$$

and we start again by calculating the first derivative.

$$\begin{aligned}
\frac{dE}{ds} &= \int \rho_s \cdot \nabla_\rho (u(\rho)) \, dx = \int \operatorname{div}(D(\rho)v) \cdot \nabla_\rho u(\rho) \, dx \\
&= -\int \sum_{i,j,k} \frac{\partial^2 u(\rho)}{\partial \rho_i \partial \rho_k} D_{ij} v_j \cdot \nabla \rho_k
\end{aligned}$$

Again, we define a function $f$, this time via

$$\frac{\partial f_i}{\partial \rho_k} = \sum_i \frac{\partial^2 u(\rho)}{\partial \rho_i \partial \rho_k} D_{ij}. \tag{5.25}$$

Using this definition the above expression can be written as

$$\frac{dE}{ds} = \int \sum_j \nabla(f_j) \cdot v_j. \tag{5.26}$$

Next, we calculate the second derivative

$$\frac{d^2E}{ds^2} = -\int \sum_j \frac{\partial \rho_j}{\partial t} \text{div}(v_j)\, dx + \int \sum_j \nabla(F_j) \cdot \frac{\partial v_j}{\partial t}$$

Unfortunately, in this general setting, we were not able to obtain any general conditions ensuring that the entropy is displacement convex. To better understand the properties of these terms we consider the following example.

### 5.2.4 Chen Model for Energy Transport

The Chen-Model for energy transport is given by the energy

$$u(\rho_1, \rho_2) = \frac{5}{2}\rho_1 \log(\rho_1)) - \frac{3}{2}\rho_1 \log(\rho_2), \tag{5.27}$$

and the diffusion matrix

$$D(\rho_1, \rho_2) := \begin{pmatrix} \rho_1 & \rho_2 \\ \rho_2 & \frac{5}{3}\frac{\rho_2^2}{\rho_1} \end{pmatrix}. \tag{5.28}$$

With this model, eq. 5.25 reduces to

$$\frac{\partial f_{ik}}{\partial \rho_k} = \delta_{jk}, \tag{5.29}$$

where $\delta_{jk}$ denotes the Kronecker delta. This furthermore simplyfies the second derivative to

$$\frac{d^2E}{ds^2} = -\int \left( \frac{\partial \rho_1}{\partial t}\text{div}(v_1) + \frac{\partial \rho_2}{\partial t}\text{div}(v_2) \right) dx - \frac{1}{2}\int \sum_i \nabla \rho_i \cdot \nabla \left( \sum_{j,k} \frac{\partial D_{jk}}{\partial \rho_i} v_j \cdot v_k \right) dx.$$

Using the definition of the geodesic in the $n$-dimensional setting, eq. (3.48), we get

$$
\begin{aligned}
\frac{d^2E}{ds^2} &= -\int \left( \frac{\partial \rho_1}{\partial t}\mathrm{div}(v_1) + \frac{\partial \rho_2}{\partial t}\mathrm{div}(v_2) \right) dx \\
&- \int \left( \frac{1}{2}\nabla\rho_1 \cdot \nabla \left( v_1^2 - \frac{5}{3}\frac{g_2^2}{\rho_1^2}v_2^2 \right) - \frac{1}{2}\nabla\rho_2 \cdot \nabla \left( 2v_1 \cdot v_2 + \frac{10}{3}\frac{\rho_2}{\rho_1}v_2^2 \right) \right) dx \\
&= -\int \left( \frac{\partial \rho_1}{\partial t}\mathrm{div}(v_1) + \frac{\partial \rho_2}{\partial t}\mathrm{div}(v_2) \right) dx \\
&- \int \left( -\nabla\rho_1 \cdot v_1 - \nabla\rho_2 \cdot v_2 \right) \mathrm{div}(v_1) \\
&+ \left( \frac{5}{3}\frac{\rho_2^2}{\rho_1^2}\nabla\rho_1 \cdot v_2 - \frac{10}{3}\frac{\rho_2}{\rho_2}\nabla\rho_2 \cdot v_2 - \nabla\rho_2 \cdot v_1 \right) \mathrm{div}(v_2) \\
&+ \frac{5}{3}\nabla\left( \frac{\rho_2^2}{\rho_1^2} \right) \cdot \nabla\rho_1 v_2^2 - \frac{5}{3}\nabla\left( \frac{\rho_2}{\rho_1} \right) \cdot \nabla\rho_2 v_2^2 \, dx.
\end{aligned}
$$

Now we use the second PDE to the geodesics, eq. (3.49) to explicitly calculate $\frac{\partial \rho_1}{\partial s}$ and $\frac{\partial \rho_2}{\partial s}$.

$$
\begin{aligned}
\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_s &= \mathrm{div}(D(\rho)v) = \mathrm{div}\left( \begin{pmatrix} \rho_1 & \rho_2 \\ \rho_2 & \frac{5}{3}\frac{\rho_2^2}{\rho_1} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \right) \\
&= \mathrm{div}\left( \begin{pmatrix} \rho_1 v_{11} + \rho_2 v_{21} & \rho_1 v_{12} + \rho_2 v_{22} \\ \rho_2 v_{11} + \frac{5}{3}\frac{\rho_2^2}{\rho_1}v_{21} & \rho_2 v_{12} + \frac{5}{3}\frac{\rho_2^2}{\rho_1}v_{22} \end{pmatrix} \right) \\
&= \begin{pmatrix} (\rho_1 v_{11})_{x_1} + (\rho_2 v_{21})_{x_1} + (\rho_1 v_{12})_{x_2} + (\rho_2 v_{22})_{x_2} \\ (\rho_2 v_{11})_{x_1} + (\frac{5}{3}\frac{\rho_2^2}{\rho_1}v_{21})_{x_1} + (\rho_2 v_{12})_{x_2} + (\frac{5}{3}\frac{\rho_2^2}{\rho_1}v_{22})_{x_2} \end{pmatrix}
\end{aligned}
$$

Inserting this into the equation for the second derivative and using the product rule we finally obtain

$$
\begin{aligned}
\frac{d^2E}{ds^2} &= \int \left( \rho_1\mathrm{div}^2(v_1) + 2\rho_2\mathrm{div}(v_1)\mathrm{div}(v_2) + \frac{5}{3}\frac{\rho_2^2}{\rho_1}\mathrm{div}^2(v_2) \right) dx \\
&+ \int \left( \frac{5}{6}\nabla\left( \frac{\rho_2^2}{\rho_1^2} \right) \cdot \nabla\rho_1 v_2^2 - \frac{5}{3}\nabla\left( \frac{\rho_2}{\rho_1} \right) \cdot \nabla\rho_2 v_2^2 \right) dx \\
&= \int \left( \sqrt{\rho_1}\mathrm{div}(v_1) + \sqrt{\frac{\rho_2^2}{\rho_1}}\mathrm{div}(v_2) \right)^2 + \frac{2}{3}\frac{\rho_2^2}{\rho_1}\mathrm{div}^2(v_2) \, dx \\
-\tfrac{5}{3} &\int \left( \frac{\rho_2}{\rho_1^{3/2}}|\nabla\rho_1| + \frac{1}{\rho_1^{1/2}}|\nabla\rho_2| \right)^2 v_2^2 \, dx.
\end{aligned}
$$

Unfortunately, in this result, the last term is negative and therefore we don't know the sign of the difference of the two terms. However, as the Chen model without potential reduces to the heat equation, it is very likely that the entropy is displacement convex. Thus, this results should be investigated further in the future.

## 5.3 Purely One-Dimensional Calculations

In the above calculations it turned out that the entropy dissipation calculation gave different results that the displacement convexity calculation. Interestingly, this is not the case for a purely one-dimensional calculation.

### 5.3.1 Entropy Dissipation in One Dimension

Again, the entropy is given by

$$E(\rho) = \int u(\rho) + \rho V(x),$$

where $\rho$, $u$ and $V$ are now all functions from $\mathbb{R}$ to $\mathbb{R}$.
For the first derivative we obtain

$$\frac{dE}{dt} = \int \rho_t(u'(\rho) + V(x))\,dx = -\int D(\rho)\xi^2 \quad \text{with } \xi := (u')_x + V_x$$

The second derivative is

$$\begin{aligned}
-\frac{d^2E}{dt^2} &= \int \rho_t D'(\rho)\xi^2\,dx + 2\int D(\rho)\xi\xi_t\,dx \\
&= \int (D(\rho)\xi)_x D'(\rho)\xi^2\,dx - 2\int ((D(\rho)\xi)_x)^2\,dx \\
&= -\int D(\rho)\xi(D'(\rho)\xi^2)_x\,dx - 2\int ((D(\rho)\xi)_x)^2\,dx \\
&= -2\int D(\rho)D'(\rho)\xi^2\xi_x\,dx - \int D(\rho)D''(\rho)\xi^2\xi\rho_x\,dx - 2\int ((D(\rho)\xi)_x)^2\,dx \\
&= -2\int D(\rho)D'(\rho)\xi^2 V_{xx}\,dx - 2\int D(\rho)D'(\rho)\xi^2(u')_{xx}\,dx - \int D(\rho)D''(\rho)\xi^2\xi\rho_x\,dx \\
&- 2\int ((D(\rho)\xi)_x)^2\,dx \\
&=: (1) + (2) + (3) + (4).
\end{aligned}$$

We will now have a closer look at each of these terms,

$$\begin{aligned}(2) &= -2\int D(\rho)D'(\rho)\xi^2(u')_{xx}\,dx = 2\int (D(\rho)D'(\rho)\xi^2)_x(u')_x\,dx \\ &= 2\int (D'(\rho))^2\xi^2(\rho_x)^2u''(\rho)\,dx + 2\int D(\rho)D''(\rho)\xi^2(\rho_x)^2u''(\rho)\,dx \\ &+ 4\int D(\rho)D'(\rho)\xi\xi_x\rho_x u''(\rho)\,dx,\end{aligned}$$

and

$$\begin{aligned}(4) &= -2\int D^2(\rho)(\xi_x)^2u''(\rho)\,dx - 2\int (D'(\rho))^2\xi^2(\rho_x)^2u''(\rho)\,dx \\ &- 4\int D(\rho)D'(\rho)\xi_x\xi\rho_x u''(\rho)\,dx.\end{aligned}$$

Collecting terms this leads to

$$\begin{aligned}-\frac{d^2E}{dt^2} &= -2\int D(\rho)D'(\rho)\xi^2V_{xx}\,dx + \cancel{2\int (D'(\rho))^2\xi^2(\rho_x)^2u''(\rho)\,dx} \\ &+ 2\int D(\rho)D''(\rho)\xi^2(\rho_x)^2u''(\rho)\,dx + \cancel{4\int D(\rho)D'(\rho)\xi\xi_x\rho_x u''(\rho)\,dx} \\ &- \int D(\rho)D''(\rho)\xi^2\xi\rho_x\,dx - 2\int D^2(\rho)(\xi_x)^2u''(\rho)\,dx - \cancel{2\int (D'(\rho))^2\xi^2(\rho_x)^2u''(\rho)\,dx} \\ &- \cancel{4\int D(\rho)D'(\rho)\xi_x\xi\rho_x u''(\rho)\,dx}.\end{aligned}$$

The final result is then

$$\begin{aligned}-\frac{d^2E}{dt^2} &= -2\int D(\rho)D'(\rho)\xi^2V_{xx}\,dx + 2\int D(\rho)D''(\rho)\xi^2(\rho_x)^2u''(\rho)\,dx \\ &- \int D(\rho)D''(\rho)\xi^2\xi\rho_x\,dx - 2\int D^2(\rho)(\xi_x)^2u''(\rho)\,dx.\end{aligned}$$

Using the notation $f' = D(\rho)u''(\rho)$ this can be rewritten as

$$\begin{aligned}-\frac{d^2E}{dt^2} &= -2\int D(\rho)D'(\rho)\xi^2V_{xx}\,dx + 2\int f'(\rho)D''(\rho)\xi^2(\rho_x)^2\,dx \\ &- \int D(\rho)D''(\rho)\xi^2\xi\rho_x\,dx - 2\int D(\rho)f'(\rho)(\xi_x)^2\,dx\end{aligned}$$

### 5.3.2 Displacement Convexity in One Dimension

Here, the entropy is defined via

$$E(\rho) = \int u(\rho) + \rho V(X)\, dx.$$

We calculate the derivatives of this entropy functional along geodesics which are defined by the set of equations

$$\rho_s = (D(\rho)v)_x$$

$$v_s = -\frac{1}{2}(D'(\rho)v^2).$$

For the first derivative we obtain

$$\begin{aligned}\frac{dE}{ds} &= \int \rho_s(u'(\rho) + V(x))\, dx = -\int D(\rho)v(u')_x\, dx - \int D(\rho)vV_x\, dx \\ &= -\int f(\rho)_x v\, dx - \int D(\rho)vV_x\, dx.\end{aligned}$$

For the Second derivative the result is

$$\begin{aligned}\frac{d^2E}{ds^2} &= -\int (\rho_s f'(\rho))_x v\, dx - \int f(\rho)_x v_x\, dx - \int \rho_s D'(\rho)vV_x\, dx - \int D(\rho)v_x V_x\, dx \\ &= \int (D(\rho)v)_x f'(\rho)v_x\, dx - \frac{1}{2}\int f'(\rho)\rho_x(D'(\rho)v^2)_x\, dx + \int D(\rho)v(D'(\rho)vV_x)_x\, dx \\ &- \frac{1}{2}\int D(\rho)(D'(\rho)v^2)_x V_x\, dx.\end{aligned}$$

Using product rule this leads to

$$\begin{aligned}\frac{d^2E}{ds^2} &= \int D'(\rho)f'(\rho)\rho_x v v_x\, dx + \int D(\rho)f'(\rho)(v_x)^2 - \frac{1}{2}\int f'(\rho)D''(\rho)(\rho_x)^2(v_x)^2\, dx \\ &- \int D'(\rho)f'(\rho)v v_x \rho_x\, dx + \int D(\rho)D''(\rho)v^2\rho_x V_x\, dx + \int D(\rho)D'(\rho)v v_x V_x\, dx \\ &+ \int D(\rho)D'(\rho)v^2 V_{xx}\, dx - \frac{1}{2}\int D(\rho)D''(\rho)v^2\rho_x V_x - \int D(\rho)D'(\rho v v_x V_x\, dx.\end{aligned}$$

This leads to

$$\begin{aligned}\frac{d^2E}{ds^2} &= \int D(\rho)f'(\rho)(v_x)^2 - \frac{1}{2}\int f'(\rho)D''(\rho)(\rho_x)^2(v_x)^2\, dx \\ &+ \int D(\rho)D''(\rho)v^2\rho_x V_x\, dx + \int D(\rho)D'(\rho)v^2 V_{xx}\, dx - \frac{1}{2}\int D(\rho)D''(\rho)v^2\rho_x V_x\, dx\end{aligned}$$

and thus

$$\begin{aligned}\frac{d^2E}{ds^2} &= \int D(\rho)f'(\rho)(v_x)^2 - \frac{1}{2}\int f'(\rho)D''(\rho)(\rho_x)^2(v_x)^2\,dx \\ &+ \int D(\rho)D'(\rho)v^2V_{xx}\,dx + \frac{1}{2}\int D(\rho)D''(\rho)v^2\rho_xV_x\,dx.\end{aligned}$$

## 5.4 Results

If we finally compare the results of the two calculations in one dimension, namely for the entropy dissipation

$$\begin{aligned}\frac{d^2E}{dt^2} &= 2\int D(\rho)D'(\rho)\xi^2V_{xx}\,dx + \int D(\rho)D''(\rho)\xi^2\rho_xV_x\,dx \\ &+ 2\int D(\rho)f'(\rho)(\xi_x)^2\,dx - \int D''(\rho)f'(\rho)\xi^2(\rho_x)^2\,dx.\end{aligned}$$

and for the displacement convexity

$$\begin{aligned}\frac{d^2E}{ds^2} &= \int D(\rho)D'(\rho)v^2V_{xx}\,dx + \frac{1}{2}\int D(\rho)D''(\rho)v^2\rho_xV_x\,dx \\ &+ \int D(\rho)f'(\rho)(v_x)^2\,dx - \frac{1}{2}\int D''(\rho)f'(\rho)(v_x)^2(\rho_x)^2\,dx,\end{aligned}$$

we see that, despite a factor 1/2, both calculation agree.
The differences that occur in displacement and entropy convexity calculations in more than one dimension might be the consequence of the fact that every two points on the Riemannian manifold can always be connected via a geodesic while this is certainly not true for a solution of the PDE. Therefore, every curve defined by a solution could probably be approximated piecewise using geodesics. This would be done in such a way that the starting point is chosen on the solution and the end point of the geodesic is determined to that the slope of the geodesic and the slope of the solution agree for a short time interval. To explore this argument more carefully would be an interesting to understand the relation between the two types of convexity used in this thesis.

# Chapter 6

# Summary

This thesis addresses the asymptotic (or long-time) behaviour of solutions to non-linear Fokker-Planck type PDEs. This is done using two different methods. Both techniques exploit convexity of an entropy functional which can naturally be defined for the given type of equation.

The first idea is to calculate the derivatives of the entropy functional along solutions of the equation. It is observed that the first derivative is the negative of some entropy production functional (which means that the entropy is a Lyaponov functional with respect to solutions of the PDE). This already shows that the entropy is non-increasing along solutions. Looking at the second derivative it can be shown that the functional is even uniformly convex which ensures the existence of a minimum of the entropy and thus the existence of a stationary state. However, this is only archived under certain assumptions on the non-linearity in the PDE which are further studied using some common examples.

The second method is based on the observation that the PDE can be interpreted as a gradient flow on a certain (at least formally) Riemannian manifold, with respect to the entropy functional. Therefore it is natural to examine the convexity properties of the entropy along geodesics in this manifold. In the context of optimal transport, convexity along geodesics is called displacement convexity. To clearify the connection to optimal transport a short introduction is given in chapter 2. Again, it is shown that the entropy defines a Lyaponov functional and that, again under certain assumptions, it is uniformly displacement convex.

However, it is observed that the conditions for both methods do not coincide, except for the one-dimensional case. The reason for this is not yet completely understood even though the fact that between any two points on the manifold a geodesic connecting them can be found. This is certainly not true for solutions of the PDE.

Finally, an attempt has been made to examine the PDE for a matrix-value nonlinearity. However, in this general setting, no statements could be made.

# Appendix A

# Some facts from Measure Theory

We state some basic definitions and theorems concerning measure theory.

**Definition** *(Metric space)* A metric space is a set $X$ together with a distance function $d : X \times X \to [0, \infty)$, having the following proporties:

$$d(x,y) = 0 \quad \Leftrightarrow \quad x = y \tag{A.1}$$
$$d(x,y) \quad = \quad d(y,x) \tag{A.2}$$
$$d(x,y) \quad \leq \quad d(x,z) + d(z,y) \quad \forall z \in X. \tag{A.3}$$
$$\tag{A.4}$$

**Definition** *(Separable Metric space)* A Metric Space $(X, d)$ is called seperable if it contains a countable dense subset.

**Definition** *(Summable Functions)* Let f be measurable and nonnegative. Then, $f$ is called summable (or integrable) function if

$$\int f d\mu < \infty.$$

**Definition** *($L^p$ spaces)* Let $\Omega$ be a measureable space with a (positive) measure $\mu$ and let $1 \leq p < \infty$. Then the Space $L^p(\Omega, d\mu)$ is defined by

$$L^p(\Omega, d\mu) := \{f : f : \Omega \to (C), \text{f is } \mu\text{-measurable and } |f|^p \text{ is } \mu\text{-summable}\}. \tag{A.5}$$

Furthermore, we define

$$L^\infty(\Omega, d\mu) := \{f : f : \Omega \to (C), \text{f is } \mu\text{-measurable and there exists a finite constant } K \text{ such that } |f(x)| \leq K \text{ for } \mu - a.e.\ x \in \mathbb{R}^N\}.$$

**Theorem A.0.1** (Rademacher's Theorem) *If $\Omega \subset \mathbb{R}^N$ is open and $f : \Omega \to \mathbf{R}^m$ is Lipschitz, then $f$ is differentiable almost everywhere.*

**Definition** *(Absolute Continous Measures)* Let $\mu$ and $\nu$ be two measures on $(X, \mathcal{F})$. Then $\nu$ is absolutely continous with respect to $\mu$ (write $\nu << \mu$) if

$$\mu(S) = 0 \Rightarrow \nu(S) = 0$$

for all $S \subset \mathcal{F}$.

**Theorem A.0.2** (Radon-Nikodym Theorem) *Let $\mu$ and $\nu$ be two finite measures on $(X, \mathcal{F})$. If $\nu$ is abolutely continous with respect to $\mu$, then*

$$\int_X F(x) d\nu(x) = \int_X F(x) h(x) d\mu(x) \tag{A.6}$$

*holds for some nonnegative $h \in L^1(X, \mu)$ and every measurable $F$.*

## A.1 Convex Analysis

**Definition** *(Lower semi-continuity)* A function on a topological space $X$ is called lower semi-continous, if it satisfies

$$F(x) \leq \liminf_{y \to x} F(y). \tag{A.7}$$

**Theorem A.1.1** (Jenson's inequality) *Assume $f : \mathbb{R} \to \mathbb{R}$ is convex, and $U \subset \mathbb{R}^N$ open, bounded. Let $u : U \to \mathbb{R}$ be summable. Then*

$$f\left(\frac{1}{|U|} \int_U u dx\right) \leq \frac{1}{|U|} \int_U f(u) dx. \tag{A.8}$$

# Appendix B

# Some facts from Differential Geometry

Here, we will recall some facts from differential geometry, mainly following [Bär06] and [Burns05].
First, we introduce the concept of a topological Mannifold.

**Definition** *(Topology)*Let $M$ be a set. A collection $\mathcal{O}$ of subsets of $M$ is called topology if,

i) $\emptyset$, $M \in \mathcal{O}$

ii) If $U_i \in \mathcal{O}$, $i \in I$, then
$$\cup_{i \in I} U_i \in \mathcal{O}$$

iii) If $U_1$, $U_2 \in \mathcal{O}$, then $U_1 \cap U_2 \in \mathcal{O}$.

A Pair $(M, \mathcal{O})$ is called a topological space. A set $U \subset M$ is called open in M, if $U \in \mathcal{O}$. A subset is called closed, if $M - A \in \mathcal{O}$.
We now consider two topological spaces $M$ and $N$ and look at the map $f : M \to N$. We call $f$ continous, if

$$f^{-1}(V) \in \mathcal{O}_M \quad \text{for all } V \in \mathcal{O}_N. \tag{B.1}$$

If $f$ is one to one and $f^{-1}$ is continous as well , we call $f$ a homoomorphism. Furthermore, we call two topological spaces homeomorph, if there exists a homeomorphism between them.
No, we can define the central concept of a Mannifold.

**Definition** *(Topological Mannifold)*Let $(M, \mathcal{O})$ be a topological space. $M$ is called $n$-dimensional topological Mannifold if

i) $M$ is hausdorff, i.e. for all $p, q \in \mathcal{M}$ with $p \neq q$there exists open sets $U, V \subset M$ with $p \in U$, $q \in V$ and $U \cap V = \emptyset$.

ii) The topology on $M$ has a countable basis, i.e. a countable subset $\mathcal{B} \in \mathcal{O}$, such that for every $U \in \mathcal{O}$ there exist $B_i \in \mathcal{B}$, $i \in I$ with

$$U = \cup_{i \in I} B_i.$$

iii) $M$ is locally homeomorph to $\mathbb{R}^N$, i.e. for all $p \in M$ there exist an open subset $U \subset M$ with $p \in U$, an open subset $V \subset \mathbb{R}^N$ and a homeomorphism $x : U \to V$.

**Definition** *(Tangential Vector)* Let $M$ be a differentiable Mannifold and $p \in M$. A tangential vector in the point $p$ is definied as the aquivalence class of all differentiable curves $c : (-\epsilon, \epsilon) \to M$ such that $\epsilon \geq 0$ and $c(0) = p$. The aquivalence relation for two curves $c_1 : (-\epsilon_1, \epsilon_1) \to M$ and $c : (-\epsilon_2, \epsilon_2) \to M$ is defined by

$$\frac{d}{dt}(x \circ c_1)_{|t=0} = \frac{d}{dt}(x \circ c_2)_{|t=0}, \tag{B.2}$$

for a given map $x : U \to V$.

The next step is to introduce differentiation on Manifolds. The simplest idea would be, given a point $p \in M$ to say that a function $f : M \to \mathbb{R}$ is differentiable in $p$, if the composition $f \circ x^{-1}$ is differentiable, i.e. to use the well-known differentiation on $\mathbb{R}$. However, it is a priori not clear that this definition does not depend on the choise of the map $x$.

**Definition** Let $M$ be a $n$-dimensional topological Mannifold. Two maps $x : U \to V$ and $y : \tilde{U} \to \tilde{V}$ a called $\mathcal{C}^\infty$-compatible, if

$$y \circ x^{-1} : x(U \cap \tilde{U}) \to y(U \cap \tilde{U})$$

is a $\mathcal{C}^\infty$ diffeomorphism. Next, we call the set of maps $x_\alpha : U_\alpha \to V_\alpha$, $\alpha \in A$, an atlas, if

$$\cup_{\alpha \in A} U_\alpha = M$$

. Finally, an atlas is called $\mathcal{C}^\infty$-atlas if every two maps in it are $\mathcal{C}^\infty$-compatible.

This directly leads to the definition of a differential structure.

**Definition** A $\mathcal{C}^\infty$-atlas $\mathcal{A}_{max}$ is called maximal (or differential structure), if every map that is $\mathcal{C}^\infty$-compatible with $\mathcal{A}_{max}$ is already in it.
Furthermore, a pair $(M, \mathcal{A}_{max})$, where $M$ is a $n$-dimensional Mannifold is called $n$-dimensional differentiable Mannifold.

## B.1 The Tangent Space

We know want to define the derivative of a differentiable map between two Mannifolds. This will lead to the concept of the tangent space. Furthermore, we call the set $T_pM := \{\dot{c}(0)|c : (-\epsilon, \epsilon) \to M\}$ differentiable with $c(0) = p$ the tangent space of $M$ at the point $p$.

## B.2 Riemannian Geometry

**Definition** *(Metric Tensor)* A Metric tensor is a bilinear form definied on the tangent space $T_pM$

$$g_p(q, s) : T_pM \times T_pM \to \mathbb{R}$$

which smoothly varies with the base point $p$.

**Definition** A Riemannian metric an a differentiable manifold $M$ is given by a scalar product on each tangent space $T_pM$ which depends smoothly on the base point $p$. A Riemannian manifold is a differential manifold, equipped with a Riemannian metric.

# Bibliography


[Ambrosio00] L. Ambrosio. Lecture notes on optimal transport problems, 2000. Lecture Notes in Mathematics "Mathematical aspects of evolving interfaces" (CIME Series, Madeira (PT), 2000) **1812**, P. Colli and J.F. Rodrigues Eds., 1–52, 2003.

[Ambrosio05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savare. *Gradient Flows in Metric Spaces and in the Space of Probability Measures.* Lectures in Mathematics ETH Zuerich. Birkhaeuser, Basel, Boston, Berlin, 2005. ISBN 3-7643-2428-7.

[Bär06] C. Bär. Differentialgeometrie, lecture notes. 2006.

[Benamou00] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. 84(3):375–393, October 2000. ISSN 0029-599X (print), 0945-3245 (electronic).

[Burns05] Keith Burns and Marian Gidea. *Differential Geometry and Topology: With a View to Dynamical Systems (Studies in Advanced Mathematics).* Chapman and Hall, 2005.

[Carillo01] José A. Carillo, Ansgar Jüngel, Peter A. Markovich, Giuseppe Toscani, and Andreas Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized sobolev inequalities. *Monatsh. Math.*, 133:1–82, 2001.

[Carillo06] José A. Carillo, Robert J. McCann, and Cedric Villani. Contractions in the 2-wasserstein length space and thermalization of granular media. *Archive for Rational Mechanics and Analysis*, 179:217–263, 2006.

[Csisz'ar67] I. Csisz'ar. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.

[Evans01] Lawrence C. Evans. Partial differential equations and mongekantorovich mass transfer, 2001. Lecture Notes in Mathematics.

[Evans04] Lawrence C. Evans. A survey of entropy methods for partial differential equations. *Bull. Amer. Math. Soc.*, 41:409–438, 2004.

[Gronwall19] T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *The Annals of Mathematics*, 20(4):292–296, 1919. ISSN 0003486X.

[Kullback59] S. Kullback. *Information theory and statistics.* John Wiley and Sons., New York, 1959.

[McCann95] Robert McCann and Wilfrid Gangbo. Optimal maps in monge's mass transport problem. *C.R. Acad. Sci. Paris. Ser. I. Math.*, 325:1653–1658, 1995.

[McCann97] Robert McCann. A convexity principle for interacting gases. *Adv. Math.*, 128:153–179, 1997.

[Otto01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001. ISSN 0360-5302.

[Otto05] Felix Otto and Michael Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255 (electronic), 2005. ISSN 0036-1410.

[Pattle59] R. E. Pattle. Diffusion from an instantaneous point source with a concentration-dependent coefficient. *Quart. J. Mech. Appl. Math.*, 12:407–409, 1959.

[Risken89] Hannes Risken. *The FokkerPlanck Equation: Methods of Solutions and Applications, second ed.* Springer Verlag, Berlin, 1989.

[Villani03] Cedric Villani. *Topics in Optimal Transportation (Graduate Studies in Mathematics, Vol. 58).* American Mathematical Society, March 2003.

## Eigenständigkeitserklärung

Ich versichere, diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Münster, 4. Oktober 2008

**Jan-Frederik Pietschmann**