Linear Scale-Spaces in Image Processing: Drift-Diffusion and Connections to Mathematical Morphology

Dissertation zur Erlangung des Grades
„Doktor der Naturwissenschaften“
der Fakultät für Mathematik und Informatik der
Universität des Saarlandes

vorgelegt von

Martin Schmidt

Saarbrücken
2017

Tag des Kolloquiums:	22. Januar 2018
Dekan der Fakultät:	Prof. Dr. Frank-Olaf Schreyer
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Frank-Olaf Schreyer
Erstgutachter:	Prof. Dr. Joachim Weickert
Zweitgutachter:	Prof. Dr. Remco Duits
Akademischer Beisitzer:	Dr. Matthias Augustin

Kurzzusammenfassung

Auf Skalenräumen basierende Ideen sind aus dem heutigen Alltag nicht mehr wegzudenken. Wir beginnen mit einem auf der homogenen Diffusionsgleichung aufbauenden Skalenraum und verfolgen zwei Strategien zur Konstruktion neuer Skalenräume. Als erstes beweisen wir, dass der lineare Osmosefilter, welcher auf einer Drift-Diffusionsgleichung beruht, eine Reihe von wichtigen Skalenraumeigenschaften erfüllt. Der zusätzliche Driftterm ermöglicht einen großen Freiraum in der Modellierung und hat sich bereits als vielversprechend in der Bildverarbeitung etabliert. Allerdings sorgt er auch dafür, dass der stationäre Zustand nicht konstant ist, im Gegensatz zu bisher untersuchten Skalenräumen. Bei dem Beweis von Vereinfachungseigenschaften im Sinne von Lyapunov-Funktionalen führt dies zu einer Reihe von Problemen.

Während der erste Teil der Arbeit einen neuen Skalenraum einführt, werden wir uns im zweiten Teil den beiden meist studierten Klassen von Skalenräumen widmen: den linearen shift-invarianten und den morphologischen Skalenräumen. Mithilfe der neu eingeführten Cramer-Fourier-Transformation zeigen wir, wie sich beide Klassen sowohl auf struktureller Ebene als auch auf der Ebene der Evolutionsgleichungen verbinden lassen. Dieses Resultat erweitert ein Ergebnis über die strukturelle Gleichheit des Gaußschen Skalenraumes mit seinem morphologischen Gegenstück. Weiterhin beweisen wir, dass die entscheidenden Eigenschaften der bisher verwendeten Cramer-Transformation erhalten bleiben.

Short Abstract

Scale-space ideas are ubiquitous and indispensable for modern image analysis. Starting from a linear scale-space based on a homogeneous diffusion equation we pursue two strategies to create new scale-spaces. First, we rigorously prove that the linear osmosis filtering, which is based on a drift-diffusion equation, fulfils several important scale-space properties. The additional drift term introduces a modelling choice that has proved valuable in the past for image processing applications. However, in contrast to previously analysed scale-spaces, the steady state is non-constant. This leads to a number of challenges when aiming for image simplification properties in terms of Lyapunov functionals.

Whereas we analyse a new scale-space in the first part, the second part picks up the two most studied classes of scale-spaces: linear shift-invariant and morphological scale-spaces. By introducing the Cramer-Fourier transform, we can connect these classes both on a structural level and on the level of evolution equations. This extends a structural similarity result between the Gaussian scale-space and its morphological counterpart. While the decisive properties of the previously used Cramer transform are preserved, our new transformation has many benefits in an image processing context. We use the Cramer-Fourier transform to construct not yet discovered scale-spaces.

Abstract

Scale-space ideas are ubiquitous and indispensable for modern image analysis. They combine efficiency with simplicity and enable a wide range of image processing applications. In this work, we review existing scale-spaces, construct new ones and analyse the relations between different classes of scale-spaces. Starting from a linear scale-space based on a homogeneous diffusion equation we pursue two strategies to create new scale-spaces.

First, we analyse the linear osmosis filtering. This image filter relies on a drift-diffusion equation where the additional drift term is used to steer the evolution towards a desired steady state. As previous results show, this has applications for seamless image cloning, shadow removal and image compression. For a given drift vector, we develop a complete scale-space theory for the linear osmosis filtering. Of particular importance is the existence of a family of Lyapunov functionals. Its existence guarantees simplification of the image evolution over time. Out of this family, one Lyapunov functional has an information theoretic interpretation and is related to the relative entropy. This explains how a scale-space can have reasonable simplification properties and still approaches a non-constant steady state. Furthermore we show convergence of the parabolic drift-diffusion equation towards a specific solution of the elliptic steady state equation. Both, existence of Lyapunov functionals and convergence result rely on having a positive lower bound of the steady state. We also prove such a statement.

Whereas we analyse a new scale-space in the first part, the second part picks up the two most studied classes of scale-spaces: linear shift-invariant and morphological scale-spaces. Previous research suggests a structural similarity between the Gaussian scale-space and the morphological scale-space with quadratic structuring function. However, the Cramér transform, that is used to analyse this relation, is limited to this specific example. We give an extensive overview over previous approaches to connect linear and morphological scale-spaces and show that a Fourier-based transform is much more natural in an image processing context. This leads us to the introduction of the novel Cramér-Fourier transform. We show that all decisive properties of the Cramér transform are preserved and the only known correspondence is recovered by our new transform. Furthermore, we extend the previously known result and connect the linear

shift-invariant and morphological scale-spaces both on a structural level and on the level of evolution equations. In particular we show that the morphological counterpart of the Poisson scale-space is the morphological scale-space with a flat structuring function. Apart from obtaining many new correspondences, we also construct not yet discovered scale-spaces. At the end, we also outline a way to unify both approaches based on the Cramér and the Cramér-Fourier transform. This allows us to sketch an extension of the obtained results to discrete filters.

To Clara

Acknowledgments

First of all, I would like to thank Joachim Weickert. He invited me to work in his mathematical image analysis group on a number of interesting and challenging problems. He also proposed and initiated to look into the topics discussed in this thesis. Without his supervision, this PhD thesis would have not been possible. I would also like to thank Remco Duits for agreeing to review this thesis and for discussions on possible extensions of the presented results. David Hafner and Pascal Peter deserve special thanks for proofreading and providing valuable feedback on a preliminary version of this work.

Furthermore, my thanks go to Darya Apushkinskaya for providing references for and key insights into a number of relevant research results related to the well-posedness of PDE evolution as well as Luis Alvarez for providing alternative proofs for the existence of weak solutions based on Galerkin approximations.

Next, I would like to express my gratitude to all other current and previous members of our group, in particular I want to thank Sarah Andris, Matthias Augustin, Leif Bergerhoff, Marcelo Cárdenas, Oliver Demetz, Sven Grewenig, Laurent Hoeltgen, Sebastian Hoffmann, Markus Mainberger, Sabine Müller, Nico Persch, Christian Schmaltz, Christopher Schroers and Simon Setzer for numerous and fruitful discussions. Furthermore, my gratitude goes to our secretary Ellen Wintringer and our system administrator Peter Franke who supported me in all non-scientific aspects related to working as research assistant at Saarland University.

Also my family, first and foremost my fiancée Clara, contributed to making this work a success. Their unlimited support always encouraged me.

Contents

Chapter 1

Introduction

> Mathematics is the art of giving the same name to different things.
>
> Henri Poincaré

1.1 Motivation

Over the last decades, the influence of image processing applications on our daily life has been rapidly increasing. Digital cameras have not only replaced film cameras, but nowadays they are also integrated in almost every smartphone. These mobile devices are omnipresent and allow to modify, improve and alter images in seconds or even real-time. For the purpose of designing the underlying algorithms two properties are vital: efficiency and simplicity. The designed algorithms and methods should have as few parameters as possible with sane defaults that work out of the box with no need of tweaking. Such algorithms are the daily driver of an increasing number of people publishing more and more images on the world wide web. However, there is also another class of methods whose advance may completely change the world we know today. These are the highly specialised algorithms that enable, among other things, driver assistance systems, self-driving cars and all kinds of virtual and augmented reality applications.

What do all of these image processing applications have in common? Without the emerge and prevalence of digital computers in the first half of the 20th century none of the above examples would be conceivable. The almost exponential increase in computing power over the last decades is only one reason for the current

overwhelming success of image processing. The main advancements came from an improved algorithm design and better understanding of drawbacks and advantages of the employed methods. This is most apparent when keeping in mind that the first image processing algorithms where investigated long before enough computing power was available to evaluate the algorithm's result numerically.

On its core, many image processing techniques rely implicitly or explicitly on the simplification of images. These simpler versions of the image are essential in many regards. The first preprocessing step after the acquisition of an image usually consists of carving out the details required for further analysis while eliminating unwanted image features. The important aspect here is that this preprocessing step does not need to accurately preserve all image details. The desired goal for which this image was taken decides on the meaningfulness of specific image features. As a simple example let us imagine a noisy image. In order to generate an visually present image it would be enough to remove the noise while preserving as many details as possible. However, for segmentation or object-detection algorithms, too many small-scale structures might deteriorate the results.

In this work we focus exclusively on simplification methods based on partial differential equations (PDEs) and, more generally, pseudodifferential equations. When starting with a single image, these methods have the advantage of generating a family of images indexed by a single time parameter. This time parameter is used to decide on the scale of an image and on the level of details. Such a family of images is known as a scale-space representation of an image. The scale-space concept is essential for this work. From a mathematical point of view scale-space approaches based on (pseudo-)differential equations offer several modelling choices. First of all, and this is the most influential decision - we need to pick a (pseudo-)differential operator. Since we are interested in creating time evolutions of images, we will mostly stick to hyperbolic and parabolic differential operators. However, there are also elliptic differential operators known to create scale spaces [e.g. Burgeth et al., 2005b]. The next choice lies in defining suitable boundary conditions. Here, the desired goal should be taken into account. When the purpose is to analyse the differential operator in question, is best to work on unbounded domains with no boundary condition that might influence the evolution of the image. For real world examples however, relying on an unbounded domain is an unrealistic assumption for practical applications. Therefore, we use homogeneous Neumann boundary conditions when discussing concrete algorithms that are designed to obtain a desired result. Neumann boundary conditions guarantee a no-flux condition over the boundary and help to preserve the average grey value of the image evolution. In contrast, we take an unbounded domain for granted when we

analyse the relations and connections between different scale-spaces.

The goal of this thesis is twofold. First, we develop a scale-space theory for the linear osmosis filter. This recently introduced image filter is based on a drift-diffusion equation and employs Neumann boundary conditions. We discuss not only the parabolic differential operator, we also analyse the accompanying elliptic steady state operator in detail. In particular, we construct a family of Lyapunov functionals and use it to prove convergence of the parabolic PDE towards a specific solution of the elliptic PDE.

Afterwards, we restrict our investigation to a sub-class of the linear scale-spaces: These are the linear shift-invariant (LSI) scale-spaces. Based on their representation in the Fourier domain we introduce the novel Cramér–Fourier transform and show how it connects these LSI scale-spaces to morphological scale-spaces.

However, before we go into more details the following sections give a overview of the different classes of scale-spaces that we use throughout this work. The end of this chapter concludes with an outline of the thesis.

1.2 Gaussian Scale-Space

Scale-spaces take a key part in this work. They offer a mathematical and precise description of an image that changes over time and adheres to a number of properties that can in many cases categorised as [Alvarez et al., 1993]

- architectural properties,
- information-reducing properties or
- invariance requirements .

First, however, let us start by fixing the notation used in connection with images. An *image* is represented by a function $f\colon \Omega \to \mathbb{R}_+$ mapping coordinates from the *image domain* $\Omega \subset \mathbb{R}^2$ to positive grey values. The image domain is always an open subset of an Euclidean space. Additional requirements are stated in the sections where they are needed. The grey values are real numbers with small and high values representing dark and light colours respectively. Although we restrict our analysis in this thesis to two-dimensional grey scale images, many of the results have analogues in higher dimensions.

Although it is also interesting to analyse images on its own, we want to study the family of images

$$\{T_t f \mid t \geqslant 0\} \tag{1.1}$$

indexed by a time parameter t. Such a family of images is an *scale-space representation* of an image f if $T_0 f$ coincides with f and if $T_t f$ represents increasingly simpler versions of f for growing t. However, there is no standardized definition of a scale-space. In most cases, images are considered to be simpler if they contain less structure and information. In this work we will also see an example where simplicity is defined relative to a given image.

One of the first studied scale-spaces is obtained by computing the convolution of an *initial image* f with a Gaussian G_{0,σ^2} with mean zero and increasing variance given by $\sigma^2 = 2t$, i.e.

$$u(x,t) = (f * G_{0,2t})(x)\,. \tag{1.2}$$

These Gaussian kernels also lend their name to this scale-space, it is widely known as *Gaussian scale-space*. Several authors assume a different set of requirements to single out this scale-space [see Weickert, 1998, Section 1.2.2]. The first ones where Iijima [1962], Witkin [1983] and Koenderink [1984].

Instead of given an explicit description of the scale-space as seen in (1.2), in many cases it is easier to construct and model scale-spaces as solution of parabolic PDEs. This leads to an alternative description of the Gaussian scale-space as the solution of the initial value problem

$$\partial_t u = \Delta u, \qquad \text{on } \Omega \times (0, T] \tag{1.3a}$$

$$u(.,0) = f(.) \qquad \text{on } \Omega \tag{1.3b}$$

where $T > 0$ and $\Delta = \left(\partial_{xx} + \partial_{yy}\right)$ denotes the spatial *Laplacian*.

The two equations (1.3a)–(1.3b) are the starting point of this thesis. By adapting and modifying this initial value problem, numerous scale-spaces have been proposed and applied to problems in image analysis.

Most importantly for practical applications: When working on real-world problems, the image domain is always bounded. Since the boundary highly influences the evolution and therefore the scale-space description of an image, it is desired to incorporate these boundaries into the theoretical investigations. This allows better predictions of the properties of discrete implementations and allows a more precise modelling of scale-spaces to achieve a specific purpose. In Chapter 2 we study a scale-space theory for the linear osmosis filtering. This image filter supplements the linear diffusion equation (1.3a) with an additive drift term. This method is useful for several image processing tasks such as seamless image cloning, shadow removal and even image compression. The point is that it is an already established initial boundary value problem where we cannot (and do not want to) change the underlying behaviour by modifying or even removing the boundary conditions.

The second important topic of this work is based on the known connection between the Gaussian scale-space and specific morphological scale-spaces. Analysing this connection in Chapter 3 allows us to define a new transform that extends previously known results.

1.3 Nonlinear Anisotropic Diffusion Scale-Spaces

Although the linear osmosis filter is a linear approach, many of its scale-space properties will turn out to resemble results that are obtained for nonlinear anisotropic scale-space approaches. For this reason, we give a short overview about nonlinear anisotropic diffusion scale-spaces.

The name anisotropic already suggests that some kind of directional information influences the evolution of the diffusion. These directional information are encoded in the *structure tensor*. It is obtained by

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \, (\nabla u_\sigma)^T) \qquad \rho \geqslant 0, \sigma > 0 \tag{1.4}$$

where u_σ is a presmoothed version of the image u and the Gaussian K_ρ is used to assemble directional information from a small (weighted) neighbourhood. If all pixels in this neighbourhood agree on the direction of the image gradient the structure tensor will have one large and one small eigenvalue. Therefore, the structure tensor can be used to detect edges and edge directions in images. In a diffusion context, the structure tensor is used to prevent smoothing across edges: Weickert [1998] analyses the following initial boundary value problem

$$\partial_t u = \operatorname{div}(D(J_\rho(\nabla u_\sigma))\nabla u) \qquad \text{on } \Omega \times (0, T] \tag{1.5a}$$
$$u(.,0) = f(.) \qquad \text{on } \Omega \tag{1.5b}$$
$$\langle D(J_\rho(\nabla u_\sigma))\nabla u, \mathbf{n}\rangle = 0 \qquad \text{on } \partial\Omega \times (0, T] \tag{1.5c}$$

where $f \in L^\infty(\Omega)$ and the diffusion tensor $D \in C^\infty(\mathbb{R}^{2\times 2}, \mathbb{R}^{2\times 2})$ is symmetry preserving with $D(J_\rho(\nabla u_\sigma))$ being uniformly positive definite.

The smoothing properties of (1.5a)–(1.5c) are carried over from the *linear homogeneous diffusion* initial value problem

$$\partial_t u = \operatorname{div}(\nabla u) \qquad \text{on } \Omega \times (0, T] \tag{1.6a}$$
$$u(.,0) = f(.) \qquad \text{on } \Omega \tag{1.6b}$$

whose solution is given by a convolution of the initial value f with a Gaussian with variance σ increasing quadratically. This convolution ensures that the evolution of (1.6a)–(1.6b) is smooth for every $t > 0$. This holds even in cases where the initial value admits only lesser smoothness assumptions.

Although the solution of (1.5a)–(1.5c) is not given by a convolution, similar smoothing properties apply and the unique solution u is a smooth function in $C^\infty(\overline{\Omega} \times (0, T])$ [Weickert, 1998, Theorem 1].

Furthermore, this solution u also admits an extremum principle which allows bounding its value by the essential infimum and essential supremum of the initial value f, i.e. the following inequality holds true for all $(\mathbf{x}, t) \in \Omega \times (0, T]$

$$\operatorname*{ess\,inf}_{\Omega} f \leqslant u(\mathbf{x}, t) \leqslant \operatorname*{ess\,sup}_{\Omega} f\,. \tag{1.7}$$

This inequality guaranties positivity preservation of the image evolution u if the initial value f is positive. Furthermore, it ensures that the solution cannot become arbitrarily large. However, not only upper and lower bounds are known a priori for the evolution u but also its average grey value. The combination of a differential operator in divergence formulation with homogeneous Neumann boundary conditions ensures that the average grey value stays constant over time. Incorporating the initial condition, we conclude that the average grey value of f is also the average grey value of $u(\cdot, t)$ for all $t \geqslant 0$. We use $\overline{f}$ to denote the average grey value of f, i.e.

$$\overline{f} := \frac{1}{|\Omega|} \int_{\Omega} f(\mathbf{x})\ \mathrm{d}\mathbf{x} \tag{1.8}$$

where

$$|\Omega| = \int_{\Omega} 1\ \mathrm{d}\mathbf{x}\,. \tag{1.9}$$

Furthermore, we use M to denote the operator

$$Mf(\mathbf{x}) := \overline{f} \qquad \text{for all } \mathbf{x} \in \Omega \tag{1.10}$$

turning an image into a constant image while preserving the average grey value.

Apart from studying these statistical properties of the evolution, also the long time behaviour is of interest. One of the important theorems is the following:

Theorem 1.3.1 (Weickert, 1998, Theorem 3)**.** *The solution of* (1.5a)–(1.5c) *has the following properties*

(a) (Lyapunov functionals)
For all $r \in C^2(\mathbb{R})$ with $r'' \geqslant 0$ the function

$$V(t) := \Phi(u(\cdot, t)) := \int_{\Omega} r(u(\mathbf{x}, t))\ \mathrm{d}\mathbf{x} \tag{1.11}$$

is a Lyapunov functional:

(i) $\Phi(u(\cdot,t)) \leqslant \Phi(Mf)$ *for all* $t \leqslant 0$.

(ii) $V \in C(R_{\geqslant 0}) \cap C^1(\mathbb{R})$ *and* $V' \leqslant 0$ *for all* $t > 0$.

Moreover, if $r'' > 0$ *on* $\mathbb{R}$, *then* $V(t) = \Phi(u(\cdot,t))$ *is a strict Lyapunov functional:*

(iii) $\Phi(u(\cdot,t)) = \Phi(Mf) \iff \begin{cases} u(\cdot,t) = Mf & \text{on } \overline{\Omega} \quad (\text{if } t > 0) \\ u(\cdot,t) = Mf & \text{a.e. on } \Omega \quad (\text{if } t = 0) \end{cases}$

(iv) If $t > 0$, *then* $V'(t) = 0$ *if and only if* $u(\cdot,t) = Mf$ *on* $\overline{\Omega}$.

(v) $V(0) = V(T)$ *for* $T > 0 \iff \begin{cases} f = Mf & \text{a.e. on } \Omega \quad \text{and} \\ u(\cdot,t) = Mf & \text{on } \overline{\Omega} \times (0,T] \end{cases}$

(b) (Convergence)

(i) $\lim_{t\to\infty} \|u(\cdot,t) - Mf\|_{L^p(\Omega)} = 0$ *for* $p \in \mathbb{R}_{\geqslant 1}$.

(b) In the 1D case, the convergence $\lim_{t\to\infty} u(x,t) = Mf$ *is uniform on* $\overline{\Omega}$.

The Lyapunov functionals play the role of a potential function and measure the progress of the evolution towards its final steady state. For two given times $t_1 < t_2$ the grey value at a single point $\mathbf{x}$ may violate

$$|u(\mathbf{x},t_1) - \overline{f}| \geqslant |u(\mathbf{x},t_2) - \overline{f}| \tag{1.12}$$

even though u converges to a constant image with value $\overline{f}$. This means that temporarily at a given point the distance of its grey value to its steady state value may increase. Therefore, instead of looking at individual pixels, the Lyapunov functional defines a potential functions that can be used to measure the progress of the evolution as a whole.

In the above nonlinear anisotropic diffusion setting, the Lyapunov functional is also used to verify L^p-convergence of the evolution to the average grey value of the initial image.

The method of Lyapunov functionals is closely related to the potential method [Cormen et al., 2001, Section 17.3] in computational complexity theory.

1.4 Morphological Scale-Spaces

This section is based on Schmidt and Weickert [2015, Chapter 3] and introduces another class of scale-spaces, called morphological scale-spaces. For their definition, we need to discuss the basic concepts of mathematical morphology.

Mathematical morphology is a system theory where the classical algebra $(\mathbb{R}, \times, +)$ that is used within linear system theory is replaced by the morphological max-plus algebra $\mathbb{R}_{\max} := (\mathbb{R} \cup \{-\infty\}, +, \max)$ or min-plus algebra $\mathbb{R}_{\min} := (\mathbb{R} \cup \{+\infty\}, +, \min)$. In the last decades, these morphological algebras have become very fruitful tools in applications such as discrete event systems [Baccelli et al., 1992]. From a more theoretical perspective, they have been studied in fields like tropical geometry [Maclagan and Sturmfels, 2015].

Morphological systems are based on the concepts of dilation and erosion. The dilation $\oplus$ resp. erosion $\ominus$ of an image f with some concave *structuring function* $s : \mathbb{R}^2 \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$(f \oplus s)(\mathbf{x}) := \sup_{\mathbf{y} \in \mathbb{R}^2} \{f(\mathbf{y}) + s(\mathbf{x}-\mathbf{y})\}, \tag{1.13}$$

$$(f \ominus s)(\mathbf{x}) := \inf_{\mathbf{y} \in \mathbb{R}^2} \{f(\mathbf{y}) - s(\mathbf{y}-\mathbf{x})\}. \tag{1.14}$$

In the following, we only focus on dilation for our derivations. Similar definitions in terms of erosions are also possible.

For the purpose of scale-space creation the width of the structuring function has to be increased over time. Commonly, an operation called *umbral scaling* is used:

$$s_t(\mathbf{x}) := t\, s\left(\frac{\mathbf{x}}{t}\right) \qquad \text{for } t > 0, \tag{1.15a}$$

$$s_0(\mathbf{x}) := s(\mathbf{x}). \tag{1.15b}$$

With $u(., 0) := f$, the dilation scale-space evolution $\{u(., t) \mid t \geqslant 0\}$ of f is given by

$$u(., t) = f \oplus s_t. \tag{1.16}$$

It is possible to derive PDE formulations for such scale-space evolutions, if one considers the *slope transform* of s [Maragos, 1994, Dorst and van den Boomgaard, 1994]:

$$\mathcal{S}[s](\mathbf{w}) := \operatorname*{stat}_{\mathbf{x} \in \mathbb{R}^2} \{s(\mathbf{x}) - \langle \mathbf{w}, \mathbf{x} \rangle\}, \tag{1.17}$$

where the *stationary values* $\operatorname{stat}_{\mathbf{x}} \{h(\mathbf{x})\}$ denote the set of function values for which the gradient is zero:

$$\operatorname*{stat}_{\mathbf{x} \in \mathbb{R}^2} \{h(\mathbf{x})\} := \{h(\mathbf{x}) \mid \nabla h(\mathbf{x}) = \mathbf{0}\}. \tag{1.18}$$

With these definitions, van den Boomgaard and Dorst [1997] have shown that $u(\mathbf{x},t) = (f \oplus s_t)(\mathbf{x})$ is the solution of

$$\partial_t u = \mathcal{S}[s](\boldsymbol{\nabla} u) \qquad \text{on } \mathbb{R}^2 \times (0,\infty), \tag{1.19a}$$
$$u(.,0) = f(.) \qquad \text{on } \mathbb{R}^2. \tag{1.19b}$$

For instance, choosing

$$s(\mathbf{x}) = -\frac{1}{4}|\mathbf{x}|^2 \tag{1.20}$$

as structuring function gives $\mathcal{S}[s](\mathbf{w}) = \mathbf{w}^2$. Thus, (1.19a) becomes

$$\partial_t u = |\boldsymbol{\nabla} u|^2. \tag{1.21}$$

An interesting equivalence between the Gaussian scale-space and morphological scale-spaces with a quadratic structuring function has been discovered by van den Boomgaard [1992b]: While Gaussians are the only separable and rotationally invariant convolution kernels [Otsu, 1981], quadratic functions are the only separable and rotationally invariant structuring functions. This has also triggered Florack et al. [1999] and Welk [2003] to consider evolutions that combine both scale-spaces.

If one uses as structuring function a flat disc

$$s(\mathbf{x}) \;=\; \begin{cases} 0 & (|\mathbf{x}| \leqslant 1), \\ -\infty & (\text{else}), \end{cases} \tag{1.22}$$

it has been shown [Brockett and Maragos, 1992] that one arrives at

$$\partial_t u = |\boldsymbol{\nabla} u|. \tag{1.23}$$

So far, it was an open question if this equation has a corresponding linear scale-space. We will answer this in Chapter 3.

1.5 Outline

The goal of this work is to extend the toolbox for analysing and designing scale-spaces. We focus on PDE-based approaches and two different strategies.

In Chapter 2 we present a detailed continuous scale-space theory for the linear osmosis filtering. The underlying drift-diffusion model was already introduced to

the image processing community by Weickert et al. [2013] and Vogel et al. [2013]. The authors also discuss several ways to construct useful drift vector fields that lead to impressive results. Furthermore, they depict essential continuous scale properties. This work expands their model mathematically. We do not only discuss and prove all properties stated in Weickert et al. [2013], we also show well-definedness of the approach and give more precise statements about prerequisites and results. We also prove that both the image evolution and the steady state are always strictly positive. In particular, this shows that the conjecture in Weickert et al. [2013] about the preservation of positivity holds true.

Whereas Section 2.1 discusses some definitions needed to state sharp requirements, the linear osmosis filtering itself is introduced in Section 2.2. Although we focus on the continuous formulation, we review some previously obtained results for discrete images using the linear osmosis filtering in Section 2.3. In Section 2.4 the existence of solutions is discussed before we analyse grey-value invariance in Section 2.5 and positivity results in Section 2.6. Section 2.7 introduces the elliptic model needed in Section 2.8 for the construction of Lyapunov functions which are essential to the convergence results in Section 2.9. Finally, we sum up the results of Chapter 2 in Section 2.10.

In Chapter 3 on the other hand, we focus on the connections between linear shift invariant and morphological scale-spaces. The underlying classes of scale-spaces are introduced and discussed in detail in the Section 3.1 and Section 3.2 respectively. Section 3.3 gives an overview over several previous approaches that connects linear to morphological scale-spaces before we introduce the novel Cramér–Fourier transform to in Section 3.4. Although it is obtained by modifying a transformation called Cramér transform and we show that both transformations share the same characteristic properties. However, our approach has the benefit of being applicable to a wider range of scale-spaces and allows us to connect linear and morphological scale-spaces both on the level of convolutions/infimal convolutions and on the level of evolution equations. This leads us to the definition of the morphological counterpart of a linear shift-invariant scale-space is Section 3.5. To visualise the results, we conduct several experiments in Section 3.7. Section 3.8 is dedicated to a comparison of both transformations with the goal to find a unifying method generalising both, the Cramér and Cramér–Fourier transform. The results are used to sketch a possible application of our analysis to discrete filters. Section 3.9 sums up the results of Chapter 3.

In Chapter 4 we conclude this thesis and give an outlook.

Chapter 2

Scale-Space Theory for Linear Osmosis

> In contrast with nature, where the nonlinear limiting of concentrations is caused by the depletion of morphogens, engineering systems require dedicated mechanisms for preventing overflow.
>
> Alexander S. Sherstinsky and Rosalind W. Picard

In this chapter we develop a scale-space theory for the linear osmosis filtering. The evolution equation (without boundary conditions and initial values) underlying the linear osmosis filtering is given by

$$\partial_t u = \operatorname{div}(\nabla u - \mathbf{d}u) \tag{2.1}$$

where the vector field $\mathbf{d}$ is called drift vector field. This PDE extends the linear diffusion equation which is recovered for a vanishing drift vector field. In an image processing context, $\mathbf{d}$ can be used to create a local grey value imbalance which is resolved globally by the above evolution equation. In a discrete setting we can think of having a semi-permeable membrane between neighbouring pixels. This also explains the name osmosis. The goal of this chapter is to establish a scale-space theory for the linear osmosis filtering. However, before we can give a precise definition of the linear osmosis filtering and the necessary requirements we have to discuss a few prerequisites. Therefore, we review a few definitions.

2.1 Mathematical Background

One of the key concepts needed in this chapter is Hölder continuity.
Let the image domain Ω be an open and bounded subset of the Euclidean space $\mathbb{R}^2$. A function $g\colon \Omega \to \mathbb{R}$ is called *Hölder continuous* with exponent $\alpha \in (0,1]$ if it fulfils the *Hölder condition*

$$|g(\mathbf{x}) - g(\mathbf{y})| \leqslant C|\mathbf{x}-\mathbf{y}|^{\alpha} \tag{2.2}$$

for all $x, y \in \Omega$ and some constant $C \geqslant 0$. This condition defines a seminorm

$$[g]_{\alpha} := \sup\left\{ \frac{|g(\mathbf{x}) - g(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^{\alpha}} \;\middle|\; \mathbf{x},\mathbf{y} \in \Omega, \mathbf{x} \neq \mathbf{y} \right\}, \tag{2.3}$$

that is, it satisfies the seminorm properties

1. $[g]_{\alpha} \geqslant 0$,
2. $[cg]_{\alpha} = |c|[g]_{\alpha}$ and
3. $[f+g]_{\alpha} \leqslant [f]_{\alpha} + [g]_{\alpha}$

for real values $c \in \mathbb{R}$ and functions $f, g\colon \Omega \to \mathbb{R}$. For being a norm, only the condition $[g]_{\alpha} = 0$ if and only if $g \equiv 0$ is missing. Therefore, we can make it a norm by adding the infinity norm

$$\|g\|_{\infty} := \sup\{g(\mathbf{x}) \mid \mathbf{x} \in \Omega\}\,. \tag{2.4}$$

This leads to the definition of the *Hölder space* $C^{0+\alpha}(\overline{\Omega})$ as the set of functions g on Ω that have a finite norm

$$\|g\|_{0+\alpha} := \|g\|_{\infty} + [g]_{\alpha}\,. \tag{2.5}$$

More generally, the *Hölder spaces* $C^{k+\alpha}(\overline{\Omega})$ are defined as the set of k-times differentiable functions on Ω for which the norm $\|g\|_{k+\alpha}$ given by

$$\|g\|_{k+\alpha} := \sum_{|\ell| \leqslant k} \|\nabla^{\ell} g\|_{\infty} + \sum_{|\ell|=k} \left[\nabla^{\ell} g\right]_{\alpha} \tag{2.6}$$

is finite. We use the following multi-index notations:

$$|\ell| := \ell_1 + \ell_2 \qquad \text{and} \qquad \nabla^{\ell} := \partial_x^{\ell_1}\, \partial_y^{\ell_2}. \tag{2.7}$$

This means the Hölder space $C^{k+\alpha}(\overline{\Omega})$ consists of k-times differentiable functions whose partial derivatives up to order k are bounded and fulfil a Hölder condition. Later on, we also use that Hölder spaces are Banach spaces, i.e. every Cauchy sequence in $C^{k+\alpha}(\overline{\Omega})$ has a limit in $C^{k+\alpha}(\overline{\Omega})$.

Now that we know the definition of Hölder spaces we turn our attention to the definition of the Hölder class of a domain. Gilbarg and Trudinger [2001, Chapter 6.2] define that the boundary $\partial\Omega$ of a bounded domain $\Omega \subset \mathbb{R}^2$ is of class $C^{k+\alpha}$, if at each point $\mathbf{x}_0 \in \partial\Omega$ there is a ball $B = B(\mathbf{x}_0)$ and a one-to-one mapping ψ of B onto $D \subset \mathbb{R}^2$ such that

1. $\psi(B \cap \Omega) \subset \mathbb{R}^2$,
2. $\psi(B \cap \partial\Omega) \subset \partial\mathbb{R}^n_+$ and
3. $\psi \in C^{k+\alpha}(B), \psi^{-1} \in C^{k+\alpha}(D)$.

Locally, we can regard such a boundary as the graph of a function in $C^{k+\alpha}(\overline{\Omega})$. Having such a condition on the boundary is very typical for existence results of solutions of partial differential equations. It is needed to extend results obtained on the open set Ω to its completion $\overline{\Omega}$ [e.g. Gilbarg and Trudinger, 2001, Chapter 6.2].

2.2 Linear Osmosis Model

The linear osmosis filtering was introduced to the image processing community by Weickert et al. [2013] and Vogel et al. [2013]. It is based on an initial boundary value problem that is obtained by incorporating an additional term - called drift term - into the homogeneous linear diffusion equation with Neumann boundary conditions.

For the following discussion we will always assume that the image domain Ω is a bounded, simply connected and open subset of the Euclidean space $\mathbb{R}^2$ with boundary $\partial\Omega$ of class $C^{2+\alpha}$. The outer normal vector on $\partial\Omega$ is denoted by $\mathbf{n}$. Due to the requirements on the boundary it is well-defined on the whole boundary. We also require the positive initial image $f\colon \overline{\Omega} \to \mathbb{R}_{>0}$ to be a function in $C^{2+\alpha}(\overline{\Omega})$. Furthermore, its gradient on the boundary is constrained to vanish in outer normal direction, i.e.

$$\langle \nabla f, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega\,. \tag{2.8}$$

One goal to modify the diffusion equation was to introduce an additional parameter that influences the steady state of the evolution. This new parameter defines

the long term solution (almost) independent of the initial image as we will see in Section 2.9. Since it forces the evolution to deviate from its path to a constant image, this parameter is typically called *drift vector field* and denoted by $\mathbf{d}$. It is a smooth vector field $\mathbf{d}\colon \overline{\Omega} \to \mathbb{R}^2$ with $\mathbf{d} \in \left(C^\infty\left(\overline{\Omega}\right)\right)^2$. Similarly to the initial image, its outer normal derivative on the boundary is required to be zero, i.e.

$$\langle \mathbf{d}, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega\,. \tag{2.9}$$

With these conditions we are now ready to define the linear osmosis initial boundary value problem.

Definition 2.2.1 (Linear Osmosis)**.** *Let f and $\mathbf{d}$ be given as described above. The linear osmosis initial boundary value problem is given by the parabolic PDE*

$$\partial_t u = \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) \qquad \text{on } \Omega \times (0, T] \tag{2.10a}$$

with initial condition

$$u(.,0) = f(.) \qquad \text{on } \overline{\Omega} \tag{2.10b}$$

and homogeneous Neumann boundary condition

$$\langle \boldsymbol{\nabla} u, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega \times [0, T]. \tag{2.10c}$$

Similar to the structure tensor introduced in Section 1.3 the drift vector field $\mathbf{d}$ usually incorporates and encodes directional information. Conceptionally, however, there are several differences between nonlinear diffusion and linear osmosis. First of all, the model stays linear after introducing the drift term whereas using a time-dependent structure tensor leads to an nonlinear anisotropic PDE. Second, since evolutions of diffusion equations converge towards a constant steady state, the long time behaviour is known in advance. This means that the real value of this kind of diffusion equations come from analysing and studying the evolution itself. In contrast, the distinct property of the linear osmosis model is the dependence of the steady state on the drift term. In fact, linear osmosis was even designed to allow to use the drift vector field as a model parameter to steer the evolution towards a desired steady state. This means that instead of analysing the evolution itself, understanding the long term behaviour is much more important for the linear osmosis model.

Throughout this chapter we will use L to denote the elliptic second-order partial differential operator

$$Lu := -\operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) \tag{2.11}$$

appearing (up to a minus sign) at on the right-hand side of (2.10a). It is a partial differential operator in divergence form, i.e. L has the form

$$Lu := -\operatorname{div}(\mathbf{A}\nabla u) + \langle \mathbf{b}, \nabla u \rangle + cu \tag{2.12}$$

with $\mathbf{A}(\mathbf{x}) = (a_{ij}(\mathbf{x}))_{i,j=1,2}$ being the constant identity matrix in $\mathbb{R}^{2\times 2}$ for all $\mathbf{x} \in \Omega$ and

$$\mathbf{b}(\mathbf{x}) = -\mathbf{d}(\mathbf{x}), \qquad c(\mathbf{x}) = -\operatorname{div}(\mathbf{d}(\mathbf{x})) \tag{2.13}$$

for all $\mathbf{x} \in \Omega$.

Equation (2.12) is a standard form of an elliptic differential operator. This form is very useful when applying well-known results from classical PDE theory to our setting. One important prerequisite needed for standard methods is uniform ellipticity of L, i.e. the existence of constants A_0 and A_1 is required such that L satisfies

$$A_0|\xi|^2 \leqslant \sum_{i,j=1}^{2} a_{ij}(\mathbf{x})\xi_i\xi_j \leqslant A_1|\xi|^2 \tag{2.14}$$

for all $\xi \in \mathbb{R}^2, \mathbf{x} \in \Omega$. Since $a_{ij}(\mathbf{x})$ is either zero or one this conditions is always trivially fulfilled and only mentioned for completeness.

2.3 Related Work

The evolution equation

$$\partial_t u = \operatorname{div}(\nabla u - \mathbf{d}u) \tag{2.15}$$

is known under several names. In the literature *Fokker-Planck equation* [Jordan et al., 1998, Risken, 1984], *Smoluchowski equation* [van Kampen, 2007, Dhont, 1996] and *Kolmogorov forward equation* or *second Kolmogorov equation* [Gnedenko, 1998, Freidlin and Wentzell, 2012] are common. Physically, equation (2.15) is motivated in cases where we have a particle moving under external forces. The above evolution equation describes the time-evolution of the probability density function of the velocity or position of this particle. In these cases, however, the drift vector field $\mathbf{d}$ is known to be a gradient vector field, i.e.

$$\mathbf{d} = -\nabla F\,. \tag{2.16}$$

This simplifies the solution enormously and is violated in all cases of interest in our image processing setting.

Biology is another research area where equation (2.15) is used. It describes the interaction between two (neuron) populations [Carrillo et al., 2011, Cantrell and Cosner, 2003].

In an image processing context, equation (2.15) was introduced by Weickert et al. [2013] and Vogel et al. [2013]. The authors supplemented the evolution equation (2.15) with homogeneous Neumann boundary conditions and a given initial image. In the following we want to review some of their applications. We start by visualising the image evolution starting with a constant image in Figure 2.1.

In this case the drift vector field $\mathbf{d}$ is the so-called *canonical drift vector field*. For a given positive image v it is given by

$$\mathbf{d} := \frac{\nabla v}{v}. \tag{2.17}$$

With this choice the elliptic steady state equation reads

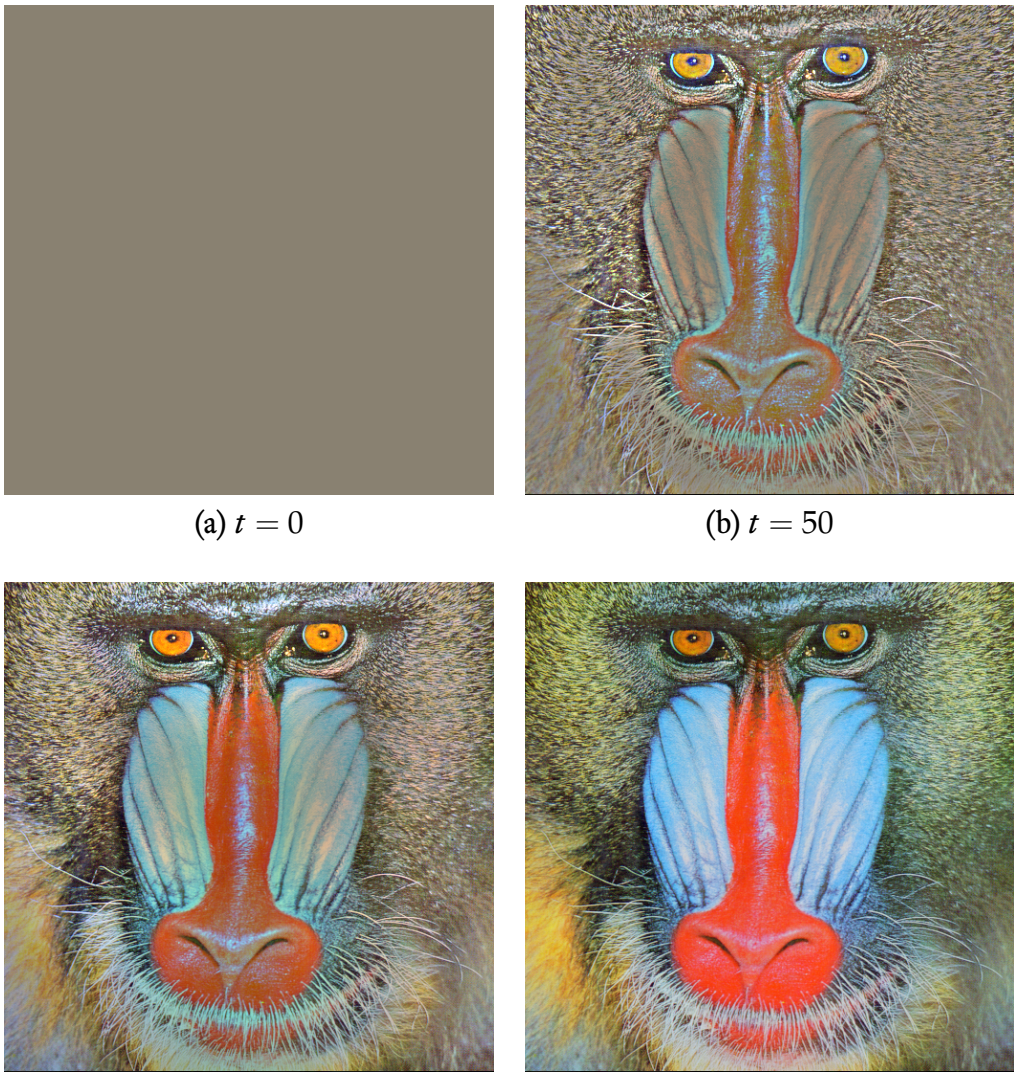
$$0 = \operatorname{div}\left(\nabla u - \frac{\nabla v}{v} u\right) \qquad \text{on } \Omega \tag{2.18a}$$

$$\nabla u \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega \tag{2.18b}$$

and is solved by v if v satisfies the boundary condition

$$\nabla v \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega\,. \tag{2.19}$$

This is the same condition that we also assume for the initial image f. Although this elliptic problem has more than one solution we prove in Section 2.9 that the evolution obtained from the parabolic problem (2.10a)–(2.10c) converges towards v up to a scaling constant which preserve the average grey value of the initial image. This is also what we observe in Figure 2.1. In Figure 2.1 the canonical drift vector field was computed from a mandrill image and the images (a)–(d) show how the constant initial image evolves over time and converges to this mandrill image again. Furthermore, we observe the multiplicative invariance of the canonical drift vector field: under multiplicative rescalings of v the canonical drift vector field stays the same. This property is one of the distinct advantages of the linear osmosis model for practical applications. With regard to the implementation, Weickert et al. [2013] employ a simple finite difference scheme. As can be seen in Figure 2.1, most of the details of the image are already visible after a several thousand iterations. This discretisation of the linear osmosis model can also be used to construct discretisations of hyperbolic PDEs [Hagenburg et al., 2012]. However, to obtain a high quality reconstruction of the original image much more iterations are needed. Vogel et al. [2013] have tested several implicit schemes and



(a) $t = 0$ (b) $t = 50$

(c) $t = 1,000$ (d) $t = 250,000$

Figure 2.1: Image evolution for different values of t [Weickert et al., 2013]

propose to use the biconjugate gradient stabilized method (BiCGStab) [van der Vorst, 1992]. This scheme is particularly designed to solve the large unsymmetric system of equations that is obtained for the linear osmosis model.

In the next example, we show a result for seamless image cloning (Figure 2.2).

(a) Euler

(b) Lagrange with selection boundary

(c) fused image

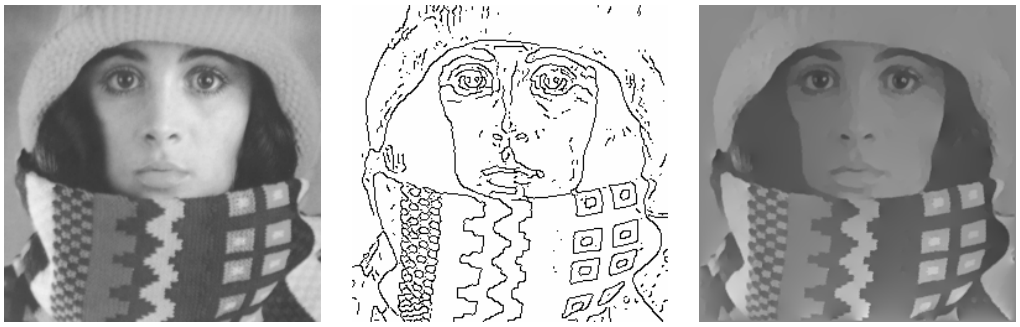
Figure 2.2: Seamless image cloning [Weickert et al., 2013]

For seamless image cloning we start with two images f_1 and f_2. For both we compute their canonical drift vector fields $\mathbf{d}_1$ and $\mathbf{d}_2$ respectively. Then if we want to seamlessly replace f_1 by f_2 in a (open) subdomain Γ of the image domain Ω we define

$$\mathbf{d}(\mathbf{x}) = \begin{cases} \mathbf{d_2}(\mathbf{x}) & \text{on } \Gamma \\ \frac{1}{2}\left(\mathbf{d_1}(\mathbf{x}) + \mathbf{d_1}(\mathbf{x})\right) & \text{on } \partial\Gamma \\ \mathbf{d_1}(\mathbf{x}) & \text{on } \Omega \setminus \bar{\Gamma} \end{cases} \tag{2.20}$$

and use f_1 as initial image. Although such a drift vector field would violate the assumed smoothness assumptions of our continuous setting the discretisation still satisfies the discrete stability criterion [Vogel et al., 2013].

For last example let us look at a proof-of-concept for image compression. Exploiting the fact that the canonical drift vector field is large at edge locations, an edge detector is used to find such locations. Afterwards, the original image is reconstructed f from the canonical drift vectors at these edge locations K. For all other points not in K the drift vector field is set to zero. Apart from the selected drift vector fields at location in K also the average grey value of f has to be stored. Figure 2.3(c) shows a reconstruction obtained from the drift vector field at locations marked in black in (b). The reconstruction is far from perfect but only a mere fraction of the whole drift vector field is used for the reconstruction.



(a) original image (b) selected pixel at edge locations (c) reconstructed image

Figure 2.3: Image compression [Weickert et al., 2013]

2.4 Existence and Regularity

In the previous section we have seen that the linear osmosis filtering is a valuable tool in every image processing toolbox. From a mathematical point of view, however, many questions are open. In the following section we want to develop a complete scale-space framework for the linear osmosis filtering.

We start by discussing existence and regularity results of solutions of the linear osmosis initial boundary value problem introduced in Definition 2.2.1. In this section, the main goal is to show that we have indeed a solution and that this solution is at least twice continuously differentiable. Furthermore, we are also interested in a uniqueness result. The boundary conditions and the initial value play a key role in establishing these objectives.

Only if the boundary conditions and the initial value reduce the solution space of the differential operator sufficiently, the possibility of having multiple solutions or even a whole family of solutions can be averted. On the other hand, care should also be taken to avoid having contradicting requirements. However, not only prerequisite from the employed model influence the number of solutions. Also the requirements on the solution itself play a crucial rule.

Our model (2.10a)–(2.10c) uses derivatives of order one in time and of order up to two in space. Naturally, we expect a solution to be differentiable in time and twice differentiable in spacial variables. Unfortunately, starting with a twice differentiable initial value $f \in C^2(\overline{\Omega})$ with Ω of class C^2 is not enough to deduce such a conclusion. Even more importantly, the Hölder spaces introduced in Section 2.1 are not suited to analyse evolutions of parabolic PDEs which depend on a time parameter. In order to incorporate this time dependence into the definition of a suitable solution space, the Hölder spaces $H^{k+\alpha}(\overline{\Omega})$ are adapted to a parabolic

setting. Following Lieberman [1988], we use $Q = \Omega \times (0, T)$ and define

$$|u|_0 = \sup_{\Omega} |u|, \tag{2.21}$$

$$[u]_\alpha = \sup \left\{ \frac{|u(X) - u(Y)|}{|X - Y|^\alpha} \,\middle|\, X \neq Y \text{ on } Q \right\} \text{ and} \tag{2.22}$$

$$\langle u \rangle_\beta = \sup \left\{ \frac{|u(x,t) - u(x,t+k)|}{k^{\beta/2}} \,\middle|\, (x,t), (x,t+k) \in Q, k > 0 \right\} . \tag{2.23}$$

These values form the foundation of the parabolic Hölder spaces $H_{k+\alpha}(Q)$. Since we only need these spaces for $k \leqslant 2$ we only give the definition for these spaces. For $0 < \alpha \leqslant 1$ we have the following definitions:

- H_α as the space of functions with finite norm

$$|u|_\alpha = |u|_0 + [u]_\alpha . \tag{2.24}$$

- $H_{1+\alpha}$ as the space of functions with finite norm

$$|u|_{1+\alpha} = |u|_0 + |\nabla u|_\alpha + \langle u \rangle_{1+\alpha} . \tag{2.25}$$

- $H_{2+\alpha}$ as the space of functions with finite norm

$$|u|_{1+\alpha} = |u|_0 + |\nabla u|_\alpha + |\nabla^2 u|_\alpha + |u_t|_\alpha + \langle \nabla u \rangle_{1+\alpha} . \tag{2.26}$$

To sum up the definitions, the norms do not only rely on the functions and its space derivatives, but also include time derivatives. Since for second order parabolic PDEs, the highest order derivative in space is twice as high as the highest order time derivative, this imbalance is taken care of in the definition. The time derivative of u only appears in the definition of $H_{2+\alpha}$ were also the second derivatives in space appear first.

Ladyzhenskaya et al. [1968], Nazarov and Uraltseva [1995] and Lieberman [1988] use these spaces to study the solvability of parabolic PDEs with Neumann boundary conditions. Although they do not study the osmosis model (2.10a)–(2.10c) explicitly, they developed and applied much more general methods in order to prove existence and regularity of a much broader class of problems.

In this section we want to follow Lieberman [1988] and sketch his results. He employs several methods originally developed for nonlinear elliptic PDEs from Gilbarg and Trudinger [2001] and adapts them to parabolic PDEs [e.g. Lieberman, 1986, 1996].

Based on the method of continuity, he uses a function $\psi \in H_{2+\beta}$ with $\beta \in (0, \alpha)$ satisfying the initial condition and the boundary conditions at time zero

$$\nabla\psi(.,0) \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega \tag{2.27a}$$
$$\psi(.,0) = f \qquad \text{on } \Omega \tag{2.27b}$$

to define the set E as the set of solutions $u \in H_{2+\alpha}$ of

$$Lu - \partial_t u = \sigma(L\psi - \partial_t\psi) \qquad \text{on } \Omega \times (0,T) \tag{2.28a}$$
$$\nabla u \cdot \mathbf{n} = \sigma\nabla\psi \cdot \mathbf{n} \qquad \text{on } \partial\Omega \times (0,T) \tag{2.28b}$$
$$u(.,0) = f \qquad \text{on } \Omega \tag{2.28c}$$

for some $\sigma \in [0,1]$.

Lieberman [1988] uses a parabolic version of Gilbarg and Trudinger [2001, Theorem 17.28] which comes down to showing boundedness of E.

Such boundedness results can be obtained by parabolic analogues of the Leray-Schauder Theory for elliptic PDEs [Gilbarg and Trudinger, 2001, Chapter 11]. In order to apply them, a priori estimates are needed. Such a priori estimates are established by both Ladyzhenskaya et al. [1968] and Lieberman [1988].

This allows us to formulate the following theorem.

Theorem 2.4.1. *The osmosis initial boundary value problem* (2.10a)–(2.10c) *has a unique solution in* $H_{2+\beta}$.

This guarantees that we have a unique classical solution of the osmosis model meaning a solution that is at least twice differentiable in the spatial variables and at least once differentiable in time.

2.5 Grey Value Invariance

For diffusion filters, one of their most compelling features is to preserve the average grey value. This allows image manipulations that do not alter the overall brightness of an image. Essential for such an invariance are the boundary conditions. One has to ensure that no mass leaves or enters the image domain. This comes down to a zero flux condition on the boundary which is expressed in terms of the outer normal vector $\mathbf{n}$ as follows

$$\langle \nabla u - \mathbf{d}u, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega\,. \tag{2.29}$$

As we can see, not only the evolution u is involved when looking at boundary conditions but also the drift vector field $\mathbf{d}$. While requiring a zero flux condition of the drift vector field itself

$$\langle \mathbf{d}, n \rangle = 0 \qquad \text{on } \partial\Omega \tag{2.30}$$

simplifies the boundary condition (2.29), it potentially decreases the modelling choices for $\mathbf{d}$ as well. However, currently most applications of the discrete osmosis filtering [Vogel et al., 2013, Weickert et al., 2013] use canonical drift vector fields at the boundary that also satisfy (2.30).

Theorem 2.5.1 (Grey value invariance)**.** *Let u be the unique solution of the osmosis evolution equation. Then we have average grey value stays constant over time, i.e.*

$$\int_\Omega u(\mathbf{x}, t)\, \mathrm{d}\mathbf{x} = \int_\Omega f\, \mathrm{d}\mathbf{x} \qquad \text{for all } t \geqslant 0. \tag{2.31}$$

While this theorem was previously proved by Weickert et al. [2013, Proposition 1], we want to give a proof for the sake of completeness.

Proof of Theorem 2.5.1. For $t = 0$, $u(\cdot, t)$ and f coincide by the initial condition (2.10b) and for $t > 0$ we compute the derivative of the average grey values

$$\begin{aligned}
\frac{\partial}{\partial t} \int_\Omega u(\mathbf{x}, t)\, \mathrm{d}\mathbf{x} &= \int_\Omega u_t(t)\, \mathrm{d}\mathbf{x} && \text{(2.32a)}\\
&= \int_\Omega \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u)\, \mathrm{d}\mathbf{x} && \text{(2.32b)}\\
&= \int_{\partial\Omega} (\boldsymbol{\nabla} u - \mathbf{d}u) \cdot \mathbf{n}\, \mathrm{d}S && \text{(2.32c)}\\
&= 0 && \text{(2.32d)}
\end{aligned}$$

where the last two equalities are due to the divergence theorem and the assumed boundary conditions respectively. As a result, the average grey value does not change. □

2.6 Positivity Preservation

Whereas nonnegativity of the image evolution was already discussed and proved in a continuous setting by Weickert et al. [2013, Proposition 1], Vogel et al. [2013,

Proposition 1] showed that in a discrete setting even positivity is preserved. In this section we extend the result of Weickert et al. and show that positivity is preserved also in a continuous setting.

For this reason, we employ the *strong maximum principle*. To put it briefly, it states that when the solution of a partial differential equation has a maximum in the interior of its domain, then the solution is constant. The original version was developed by Hopf in 1927 [see Hopf et al., 2002] and only applies to a specific class of elliptic PDEs. Nirenberg [1953] extends these results to parabolic PDEs, and later on, Friedman [1964] prove a strong maximum principal for parabolic partial differential equations involving Neumann boundary conditions. Even more general boundary conditions are discussed in Lieberman [1996]. We will use the notation of Friedman [1964, Chapter 2]. Starting point is a strongly parabolic evolution equation

$$\partial_t u = \operatorname{div}(\mathbf{A}\boldsymbol{\nabla} u) + \mathbf{b} \cdot \boldsymbol{\nabla} u + cu \tag{2.33}$$

with bounded functions b_i and c. In (2.12) we have seen that the osmosis evolution equation has this form. However, the results of Friedman [1958, 1964] and Lieberman [1996] state as additional assumption the nonpositivity of c, i.e.

$$c \leqslant 0\,. \tag{2.34}$$

In our case, the problem is that the variable c is the divergence of the drift vector field $\mathbf{d}$ and only in trivial cases free of sign changes. As a remedy, the transformation

$$u \mapsto u \exp(\alpha t) \tag{2.35}$$

transforms equations of type (2.33) to equations of type (2.33):

$$\partial_t u = \operatorname{div}(\mathbf{A}\boldsymbol{\nabla} u) + \mathbf{b} \cdot \boldsymbol{\nabla} u + (c - \alpha)u\,. \tag{2.36}$$

The only difference is the shift of c by α. Hence, this change of variable transforms the osmosis evolution equation to an equation of form (2.36) with $(c - \alpha) < 0$ for α large enough. The transformation (2.35) is commonly used when studying partial differential equations of the form (2.33), see e.g. Smoller [1983], Cantrell and Cosner [2003] and Protter and Weinberger [1999].

These observations allow to apply classical results from the theory of partial differential equations. In particular, these results apply to the osmosis. Cantrell and Cosner [2003, Corollary 1.18] formulate the following conclusion:

Theorem 2.6.1. *If* $u \in C^{2,1}(\Omega \times (0,T]) \cap C(\overline{\Omega} \times [0,T])$ *satisfies*

$$u_t - Lu \geqslant 0 \qquad \text{on } \Omega \times (0,T] \tag{2.37a}$$
$$\langle \nabla u, \mathbf{n} \rangle \geqslant 0 \qquad \text{on } \partial\Omega \times [0,T] \tag{2.37b}$$

and $u(\mathbf{x},0) > 0$ *for some* $\mathbf{x} \in \Omega$*, then*

$$u(\mathbf{x},t) > 0 \qquad \text{on } \Omega \times (0,T]. \tag{2.38}$$

All these requirements are fulfilled by the unique solution obtained in Theorem 2.4.1. Therefore, the unique solution of the linear osmosis model (2.10a)–(2.10c)

$$\begin{aligned} \partial_t u = \operatorname{div}(\nabla u - \mathbf{d}u) &\qquad \text{on } \Omega \times (0,T] \\ u(.,0) = f(.) &\qquad \text{on } \overline{\Omega} \\ \langle \nabla u, \mathbf{n} \rangle = 0 &\qquad \text{on } \partial\Omega \times [0,T] \end{aligned}$$

is strictly positive in $\Omega \times (0,T]$.

2.7 Elliptic PDE and Steady State

When aiming at long-term solutions of parabolic PDEs, it can be useful to also look at the corresponding elliptic steady state PDE.

Complementing the operator L from Definition (2.12) with a homogeneous Neumann boundary condition yields the elliptic partial differential equation

$$\operatorname{div}(\nabla u - \mathbf{d}u) = 0 \qquad \text{on } \Omega \tag{2.40a}$$
$$\langle \nabla u, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega. \tag{2.40b}$$

In contrast to the previously parabolic PDE, this PDE does not include a time parameter. These kind of equations are interesting since they arise naturally when studying the long time behaviour of parabolic PDEs. In many cases it can be shown that the solution of the parabolic equation converges to a solution of the elliptic one when the time t approaches infinity. This is exactly what we do in the next few sections.

The most beautiful theory for this kind of differential equations with Neumann boundary condition is available in the case where the sign of $c \not\equiv 0$ does not change, e.g.

$$c \leqslant 0 \qquad \text{on } \Omega. \tag{2.41}$$

Not only does a unique solution of

$$Lu = f \qquad \text{on } \Omega \tag{2.42a}$$
$$\langle \nabla u, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega \tag{2.42b}$$

exist in $C^{2+\alpha}(\overline{\Omega})$, but it also satisfies a strong maximum-principle (Cantrell and Cosner [2003], Theorem 1.2 and 1.16).

Existence and unique solvability are proven by Gilbarg and Trudinger [2001, Theorem 6.31].

Theorem 2.7.1. *Let Ω be a domain with boundary of class $C^{2+\alpha}$ and let the coefficients of L be in $C^{\alpha}(\overline{\Omega})$ with $c \leqslant 0, c \not\equiv 0$. Then, the homogeneous Neumann boundary problem*

$$Lu = f \qquad \textit{on } \Omega \tag{2.43a}$$
$$\langle \nabla u, \mathbf{n} \rangle = 0 \qquad \textit{on } \partial\Omega \tag{2.43b}$$

has a unique solution $u \in C^{2+\alpha}(\overline{\Omega})$ for every $f \in C^{\alpha}(\overline{\Omega})$.

Unfortunately, the elliptic PDE (2.40a)–(2.40b) does not fulfil $c \leqslant 0$. Furthermore, also the transformation used in the previous section to shift c is not applicable without incorporating a time parameter.

Although we lose uniqueness of the solution, this is beneficial since we are not interested in the trivial solution $u \equiv 0$ which always solves (2.40a)–(2.40b) for $f \equiv 0$.

Our goal is to show that (2.40a)–(2.40b) has a positive, smooth solution that is unique up to a multiplicative constant.

The first step to achieve this is finding a *weak solution*. Weak solutions are generalised solutions that rely on an integral formulation employing the L^2-inner product. The differential equation is multiplied by a so-called *test function* v and afterwards integrated over the whole domain Ω. The space of test functions is the space $C_0^{\infty}(\Omega)$ of all smooth functions on $\overline{\Omega}$ with compact support in Ω.

In order to formalise the above descriptions and simplify them, a generalised notion of differentiability, called *weak differentiability* is needed. A function $g \in L^2(\Omega)$ is a weak derivative of order $\alpha \in \mathbb{N}^2$ of a function $f \in L^2(\Omega)$ if

$$\int_{\Omega} g v \, \mathrm{d}\mathbf{x} = (-1)^{|\alpha|} \int_{\Omega} f \, D^{\alpha} v \, \mathrm{d}\mathbf{x} \tag{2.44}$$

holds for all $v \in C_0^{\infty}(\Omega)$. The weak derivative g is then denoted by $D^{\alpha} f$.

If the function f is sufficiently often differentiable the weak derivative coincides with the usual derivative. In this case we could apply the divergence theorem (as a generalisation of integration by parts to multidimensional domains) $|\boldsymbol{\alpha}|$-times to obtain (2.44) since the boundary terms would vanish due to v having compact support in Ω and therefore being zero on the boundary. However, weak derivatives can be computed even in cases where f is merely integrable.

Starting from weak derivatives, it is also possible and straight forward to construct *weak gradients* and *weak divergences*.

To study the space of weakly differentiable functions *Sobolev spaces* are used. The Sobolev space

$$H^k(\Omega) := \{u \in L^2(\Omega) \mid D^\alpha u \in L^2(\Omega) \text{ for all } \alpha \text{ with } |\boldsymbol{\alpha}| \leqslant k\} \tag{2.45}$$

of k-times weakly differentiable functions with weak derivatives up to order k in $L^2(\Omega)$ is a Hilbert space with scalar product

$$\langle u, v\rangle_k = \int_\Omega \sum_{|\boldsymbol{\alpha}|\leqslant k} \langle D^\alpha u, D^\alpha v\rangle \, \mathrm{d}\mathbf{x}\,. \tag{2.46}$$

Now, having all the above definitions at hand, we can define what we mean when we talk about a weak solution of our elliptic partial differential equation. A function $u \in H^1(\Omega)$ is said to be a weak solution of (2.40a)–(2.40b) if

$$-\int_\Omega \left(\langle \boldsymbol{\nabla} u, \boldsymbol{\nabla}\varphi\rangle - \langle \mathbf{d}u, \boldsymbol{\nabla}\varphi\rangle\right) \mathrm{d}\mathbf{x} = 0 \qquad \text{for all } \varphi \in H^1(\Omega). \tag{2.47}$$

Similar to the definition of the weak derivative in (2.44), the above definition is the one one would obtain if all functions would be sufficiently smooth. Then the computations would involve multiplying (2.40a) by the test function φ and applying the divergence theorem to it. The boundary term would cancel due to the boundary condition (2.40b) and the boundary property (2.9) of the drift vector field $\mathbf{d}$.

For finding a unique (weak) solution of our elliptic PDE

$$Lu = -\operatorname{div}(\mathbf{A}\boldsymbol{\nabla} u) + \langle \mathbf{b}, \boldsymbol{\nabla} u\rangle + cu \tag{2.48}$$

Droniou and Vázquez [2009] heavily rely on the formal adjoint operator L^* given by

$$L^* v = -\operatorname{div}(\mathbf{A}\boldsymbol{\nabla} v) - \langle \mathbf{b}, \boldsymbol{\nabla} v\rangle\,. \tag{2.49}$$

They define the operator L_γ by $L_\gamma u := Lu + \gamma u$. This is an isomorphism between $H^1(\Omega)$ and its dual $\left(H^1(\Omega)\right)'$ for γ large enough since the second term dominates in this case. We can use this isomorphism to define a compact operator K on $L^2(\Omega)$ by

$$K := \gamma \left(L_\gamma\right)^{-1}\Big|_{L^2(\Omega)} \tag{2.50}$$

having the property

$$\ker(L) = \ker(\mathrm{Id}_{L^2} - K)\,. \tag{2.51}$$

Using the theory of compact operators in Hilbert spaces [see Evans, 1998, Section 6.2.3] we conclude that

$$\ker(L) = \ker(\mathrm{Id}_{L^2} - K) = \ker(\mathrm{Id}_{L^2} - K^*) = \ker(L^*)\,. \tag{2.52}$$

For the osmosis differential operator, the adjoint operator is

$$L^* = -\operatorname{div}(\mathbf{A}\nabla v) - \langle \mathbf{d}, \nabla v\rangle\ . \tag{2.53}$$

For this operator, the constant functions lie in its kernel. Therefore, $\ker(L^*) \geqslant 1$. To show $\ker(L^*) \leqslant 1$, Droniou and Vázquez [2009] have constructed the piecewise linear test function $\varphi_\varepsilon \in H^1(\Omega)$ by

$$\varphi_\varepsilon(r) := \begin{cases} 0 & \text{if } r \geqslant \varepsilon \\ r - \varepsilon & \text{if } 0 < r < \varepsilon \\ -\varepsilon & \text{if } r \geqslant 0 \end{cases}\ . \tag{2.54}$$

This allows them to prove the following theorem [Droniou and Vázquez, 2009, Theorem 1.1].

Theorem 2.7.2. *The elliptic partial differential equation* (2.40a)–(2.40b) *has a unique up to a multiplicative constant weak solution $w \in H^1(\Omega)$ with*

$$w > 0 \qquad \text{a.e. on } \Omega. \tag{2.55}$$

In order to reduce the solution space to a single solution, we use the multiplicative constant to fix the average grey value of the solution w to be the same as the average grey value of the initial image f used in the parabolic setting. The positivity constraint (2.55) makes this possible.

With the previous theorem, we have a candidate w which solves (2.40a)–(2.40b) in a weak sense. The next step is to show, that this candidate is indeed sufficient

often differentiable such that the derivatives in (2.40a)–(2.40b) make sense in the usual, classical meaning.

Regularity in the interior of Ω is a classic result and can be found in [Gilbarg and Trudinger, 2001, Theorem 8.11], or [Evans, 1998, Chapter 6.3, Theorem 3].

Theorem 2.7.3. *Suppose $u \in H^1(\Omega)$ is a weak solution of the elliptic PDE*

$$Lu = f \qquad \text{on } \Omega \tag{2.56}$$

where $f \in C^\infty(\Omega)$ and L has C^∞-coefficients. Then

$$u \in C^\infty(\Omega). \tag{2.57}$$

Alternatively, also the theory of hypoelliptic differential operators [Hörmander, 1983a,b, 1985] gives the same result since every elliptic operator is hypoelliptic.

It remains to show the regularity at the boundary. Both Gilbarg and Trudinger [2001] as well as Evans [1998], prove regularity results at the boundary only for Dirichlet boundary conditions. Fortunately, also regularity results for Neumann boundary are available in the literature. The following theorem is proved by Winkert [2010, Theorem 4.3].

Theorem 2.7.4. *Let $u \in H^1(\Omega)$ be a weak solution of*

$$Au = f(\mathbf{x}, u, \nabla u) \qquad \text{on } \Omega \tag{2.58a}$$
$$\langle \nabla u, \mathbf{n} \rangle = g(\mathbf{x}, u) \qquad \text{on } \partial\Omega. \tag{2.58b}$$

Under suitable assumptions, it holds

$$u \in C^{1+\beta}(\overline{\Omega}) \tag{2.59}$$

for some $\beta > 0$.

If we apply this theorem to

- $a_i(\mathbf{x}, s, \xi) = \xi_i$,
- $Au(\mathbf{x}) = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i} a_i(\mathbf{x}, u(\mathbf{x}), \nabla u(\mathbf{x}))$,
- $f(\mathbf{x}, s, \xi) = -\langle \xi, \mathbf{d} \rangle - s \operatorname{div}(\mathbf{d})$ and
- $g(\mathbf{x}, s) \equiv 0$,

the boundary value problem (2.58a)–(2.58b) boils down to (2.40a)–(2.40b), i.e.

$$\begin{aligned} \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) &= 0 && \text{on } \Omega \\ \langle \boldsymbol{\nabla} u, \mathbf{n} \rangle &= 0 && \text{on } \partial\Omega\,. \end{aligned}$$

Furthermore, the used functions A, a_i, f and g are easy enough to check the conditions given in Winkert [2010] without any problems. Together with Theorem 2.7.3, this shows

$$w \in C^{\infty}(\Omega) \cap C^{1+\beta}(\overline{\Omega}) \tag{2.60}$$

for the weak solution w from Theorem 2.7.2.

Next, we want to show positivity of w. Therefore, we use *Harnacks inequality*. It tells us that for any Ω' that is compactly embedded in Ω, there is a constant C such that

$$\sup_{\Omega'} w \leqslant C \inf_{\Omega'} w \tag{2.61}$$

holds true. A proof of this inequality can be found in any standard book on partial differential equations [e.g Gilbarg and Trudinger, 2001, Corollary 8.21].

Since we already know $w > 0$ a.e. on Ω, this shows by contradiction that

$$w > 0 \text{ on } \Omega\,. \tag{2.62}$$

If (2.62) would be violated, we could find a point $\mathbf{x}_0$ and a compact neighbourhood Ω' in Ω with $w(\mathbf{x}_0) \leqslant 0$. With Harnacks inequality (2.61) this would mean w is nonpositive in the whole set Ω' violating Theorem 2.7.2.

It remains to show positivity at the boundary. This is obtained by a theorem called *Hopf's boundary lemma* in the literature. The formulation used in this work is modified in such a way that $-u$ would fulfil the requirements given in Gilbarg and Trudinger [2001, Lemma 3.4].

Theorem 2.7.5 (Hopf's boundary lemma)**.** *Suppose that L is uniformly elliptic, $Lu \leqslant 0$ in Ω. Let $\mathbf{x}_0 \in \partial\Omega$ be such that*

- *u is continuous at $\mathbf{x}_0$*
- *$0 = u(\mathbf{x}_0) < u(\mathbf{x})$ for all $\mathbf{x} \in \Omega$*
- *$\partial\Omega$ satisfies an interior sphere condition at $\mathbf{x}_0$.*

Then

$$\langle \boldsymbol{\nabla} u(\mathbf{x}_0), \mathbf{n}(\mathbf{x}_0) \rangle < 0 \tag{2.63}$$

if the left hand side exists at $\mathbf{x}_0$.

Since this conclusion contradicts our boundary condition and the first and third assumption is fulfilled for w, the second condition must be violated for every point $\mathbf{x}_0$ at the boundary of Ω. Summing up, we have proven

Theorem 2.7.6. *Let* $\mathbf{d}$ *and* Ω *be as in Section 2.2. The elliptic partial differential equation*

$$\operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) = 0 \qquad \text{on } \Omega \tag{2.64a}$$

$$\langle \boldsymbol{\nabla} u, \mathbf{n} \rangle = 0 \qquad \text{on } \partial\Omega \tag{2.64b}$$

has a positive, smooth solution $w \in C^\infty(\Omega) \cap C^{1+\beta}(\overline{\Omega})$ *that is unique up to a multiplicative constant. In particular, there are constants* $\mathcal{E}, \varepsilon > 0$ *such that*

$$\mathcal{E} \geqslant w \geqslant \varepsilon > 0 \text{ on } \overline{\Omega} \tag{2.65}$$

since w *is continuous and positive on the closed domain* $\overline{\Omega}$ *and therefore takes its maximum* $\mathcal{E}$ *and minimum* ε *on* $\overline{\Omega}$ *as well.*

Especially, the lower bound is of utmost importance for the definition of the osmotic measure in the next section.

2.8 Lyapunov Functionals

In this section we study the simplification properties of the linear osmosis filtering.

It is also the basis for the convergence results that we establish in Section 2.9. We employ an entropy method that allows us to show that the evolution of the parabolic linear osmosis equation (2.10a)–(2.10c) approaches a solution of the elliptic equation analysed in Section 2.7 such that the multiplicative constant is uniquely determined by the initial value. For the theory of partial differential equations a particular entropy methods has proved itself useful. It is based on so-called Lyapunov functionals [Lyapunov, 1907]. These functionals play the role of the potential that is missing in cases where the drift vector field is not a gradient vector field. We show that the family of Lyapunov functions that we construct is strictly decreasing over time and has a minimum for a specific elliptic solution.

In contrast to physically motivated discussions (e.g. Kinderlehrer and Kowalczyk [2001], Bartier et al. [2007], Ge [2009]) the osmosis model was specifically designed to work in an image processing context. Noticeable differences are the non-integrability of the drift vector field and the use of Neumann boundary conditions.

Since we have already established regularity and uniqueness results for solutions for the parabolic and elliptic equations, we can prove similar results as in Bartier et al. [2007], however, in a much more rigorous way.

Let f be an initial image satisfying the conditions given in Section 2.2. Then, Theorem 2.7.6 guarantees that there is a unique positive solution $w(\mathbf{x})$ of the elliptic PDE

$$\operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) = 0 \tag{2.66}$$

satisfying the homogeneous Neumann boundary condition and having the same average grey value as f.

In contrast to results involving diffusion equations, we need the steady state of the elliptic PDE to formulate a Lyapunov functional for the parabolic PDE . However, as we have seen, the steady state solution only depends on the drift vector field $\mathbf{d}$.

Theorem 2.8.1 (Lyapunov Functionals)**.** *Let u be the unique solution of the linear osmosis model* (2.10a)–(2.10c)

$$\begin{aligned} \partial_t u &= \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) && \text{on } \Omega \times (0, T] \\ u(.,0) &= f(.) && \text{on } \overline{\Omega} \\ \langle \boldsymbol{\nabla} u, \mathbf{n} \rangle &= 0 && \text{on } \partial\Omega \times [0, T] \end{aligned}$$

from Theorem 2.4.1 and let w be the unique solution of the steady state equation (2.40a)–(2.40b)

$$\begin{aligned} \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d}u) &= 0 && \text{on } \Omega \\ \langle \boldsymbol{\nabla} u, \mathbf{n} \rangle &= 0 && \text{on } \partial\Omega \end{aligned}$$

from Theorem 2.7.6 that has the same average grey value as the initial image f.
For all convex C^2-functions $r : \mathbb{R} \to \mathbb{R}$,

$$V(t) = \Phi(u(\cdot, t)) = \int_\Omega w(\mathbf{x}) r\left(\frac{u(\mathbf{x}, t)}{w(\mathbf{x})}\right) \mathrm{d}\mathbf{x} \tag{2.68}$$

is a Lyapunov functional, i.e.

i) $\Phi(u(\cdot, t)) \geqslant \Phi(w)$ *for all* $t \geqslant 0$ *and*

ii) $V \in C[0, \infty) \cap C^1(0, \infty)$ *and* $V'(t) \leqslant 0$ *for all* $t > 0$.

In particular, we have

$$\frac{\partial}{\partial t} V(t) = -\int_\Omega r''\left(\frac{u(\mathbf{x}, t)}{w(\mathbf{x})}\right) w(\mathbf{x}) \left(\boldsymbol{\nabla}\left(\frac{u(\mathbf{x}, t)}{w(\mathbf{x})}\right)\right)^2 \mathrm{d}\mathbf{x}. \tag{2.69}$$

Moreover, if $r'' > 0$, then $V(t) = \Phi(u(\cdot, t))$ is a strict Lyapunov functional, i.e.

iii) $\Phi(u(\cdot,t)) = \Phi(w)$ *if and only if* $u(\cdot,t) = w$ *on* $\overline{\Omega}$ *for all* $t \geqslant 0$.

iv) If $t > 0$, *then* $V'(t) = 0$ *if and only if* $u(\cdot,t) = w$ *on* $\overline{\Omega}$.

v) $V(0) = V(T)$ *for* $T > 0$ *if and only if* $u(\cdot,t) = w$ *on* $\overline{\Omega}$ *for all* $t \in [0,T]$.

This theorem resembles the analysis of the nonlinear diffusion case in Section 1.3. Moreover, although the results look quite similar, there is a large conceptual difference: The Lyapunov function in Theorem 2.8.1 depends on the unknown solution of the elliptic evolution. Without the analysis carried out in Section 2.7, we would not be able to prove this theorem since we require sufficient smoothness that we get for free if we know that the solution is constant.

Since the Lyapunov functionals play the role of a potential function, it is not surprising that they have an information theoretic interpretation. In dependence of the convex function r which is denoted by ϕ in this context, the term

$$\int_{\Omega} w(\mathbf{x}) \phi\left(\frac{u(\mathbf{x},t)}{w(\mathbf{x})}\right) \mathrm{d}\mathbf{x} \tag{2.70}$$

is known as ϕ-divergence [Liese, 2012]. We discuss specific choice of r or ϕ and the consequences thereof at the end of this section.

In order to prove Theorem 2.8.1 we reformulate the results in a stochastic framework. This allows to incorporate the size of the image domain Ω as well as the solution w of the elliptic equation into a new measure. Its properties are crucial for the proof.

Let us start by reviewing the definition of a measurable space. A *measurable space* is a pair (X,Σ) where X is a set and Σ is a σ-algebra over X, i.e. Σ is a set of subsets of X that satisfies

- $\emptyset \in \Sigma$,
- $A \in \Sigma$ implies $X \setminus A \in \Sigma$ and
- $A_i \in \Sigma$ for $i \in \mathbb{N}$ implies $\bigcup_{i\in\mathbb{N}} A_i \in \Sigma$.

The subsets in Σ are the ones on which a measure can be defined. A *measure* μ on a measurable space (X,Σ) is a function $\mu\colon \Sigma \to \mathbb{R}$ with

- $\mu(E) \geqslant 0$ for all $E \in \Sigma$,
- $\mu(\emptyset) = 0$ and

- for a pairwise distinct family $(E_i)_{i\in\mathbb{N}}$ in Σ:

$$\mu\left(\bigcup_{i\in\mathbb{N}} E_i\right) = \sum_{i\in\mathbb{N}} \mu(E_i)\,. \tag{2.71}$$

The triple (X, Σ, μ) is then called a measure space. If in addition $\mu(X) = 1$, then it is called a probability space.
The most important measure is the Lebesgue measure λ on an open set Ω of an Euclidean space. The σ-algebra of all Lebesgue-measurable sets is denoted by $\mathcal{B}(\Omega)$. It contains every subset of every null set. The Lebesgue measure is the measure we usually use, it satisfies

$$\lambda(\mathcal{A}) = \int_{\mathcal{A}} 1 \,\mathrm{d}\mathbf{x} = |\mathcal{A}| = \int_{\mathcal{A}} 1 \,\mathrm{d}\lambda(\mathbf{x})\,. \tag{2.72}$$

In the following we construct a measure that is equivalent to the Lebesgue measure.

Definition/Theorem 2.8.2. *Let w be the unique positive solution of the elliptic PDE* (2.40a)–(2.40b) *with the same average grey value as the initial image f. The osmotic measure δ on the σ-algebra $\mathcal{B}(\Omega)$ of all Lebesgue-measurable subsets of Ω is defined by*

$$\delta : \mathcal{A} \mapsto \frac{1}{\overline{f}} \int_{\mathcal{A}} w(\mathbf{x}) \,\mathrm{d}\mathbf{x} \tag{2.73}$$

for all $\mathcal{A} \in \mathcal{B}(\Omega)$.
It has the following properties:

i) The osmotic measure is a probability measure, i.e.

$$\delta(\Omega) = 1 \tag{2.74}$$

and therefore $(\Omega, \mathcal{B}(\Omega), \delta)$ is a probability space.

ii) The osmotic measure δ and the Lebesgue measure λ are equivalent.

The equivalence of measures is based on absolute continuity and guarantees that both measures agree on which sets are null sets.

Definition 2.8.3. *Let two measures γ and σ on the same measurable space (X, Σ) be given.*

(a) *The measure γ is called absolute continuous with respect to the measure σ if every null set of σ is also a null set of γ.*

In this case we write $\gamma \ll \sigma$.

(b) *We say γ and σ are equivalent if $\gamma \ll \sigma$ and $\sigma \ll \gamma$.*

Proof of Theorem 2.8.2. i) This follows from the construction of δ and choice of w.

ii) Let $\mathcal{A} \in \mathcal{B}(\Omega)$. By definition, $\delta(\mathcal{A})$ is

$$\delta(\mathcal{A}) = \frac{1}{\overline{f}} \int_{\mathcal{A}} w(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \frac{1}{\overline{f}} \int_{\mathcal{A}} w(\mathbf{x}) \, \mathrm{d}\lambda(\mathbf{x}). \tag{2.75}$$

Hence, if $\mathcal{A}$ is a λ-null set it is also a δ-null set.

On the other hand, if $\mathcal{A}$ is not a λ-null set it is also not a δ-null set since $w \geqslant \varepsilon > 0$ on $\mathcal{A} \subset \Omega$.

□

For the above proof, it is essential that we not only have a positive function w but a positive function with a positive lower bound. The last ingredients that we need to prove Theorem 2.8.1 is the following formulation of the divergence theorem

$$\int_{\Omega} \nabla g \cdot F \, \mathrm{d}\mathbf{x} = \int_{\partial\Omega} g, F \cdot \mathbf{n} \, \mathrm{d}S - \int_{\Omega} g \, \mathrm{div}\, F \, \mathrm{d}\mathbf{x} \tag{2.76}$$

that can be obtained by applying the standard divergence theorem to g, F, and a stochastic version of Jensens' inequality [e.g Niculescu and Persson, 2005]:

Theorem 2.8.4. *(Jensen's inequality) Let (X, Σ, σ) be a measure space with $\sigma(X) = 1$ and φ a convex function on $\mathbb{R}$, then*

$$\varphi\left(\int_X g \, \mathrm{d}\sigma\right) \leqslant \int_X \varphi \circ g \, \mathrm{d}\sigma \tag{2.77}$$

for all σ-measurable functions $g : X \to \mathbb{R}$ for which $\varphi \circ g$ is σ-integrable.
If φ is strictly convex, the above inequality becomes an equality if and only if g is constant σ-almost everywhere.

Proof of Theorem 2.8.1.

i) Using the preservation of the average grey value, we can express the Lyapunov functional Φ in terms of the measure δ by

$$\Phi(u(\cdot,t)) = \overline{f} \int_\Omega r\left(\frac{u(\mathbf{x},t)}{w(\mathbf{x})}\right) \mathrm{d}\delta(\mathbf{x}) \tag{2.78}$$

$$\Phi(w) = \overline{f}\, r\left(\int_\Omega \frac{u(\mathbf{x},t)}{w(\mathbf{x})} \mathrm{d}\delta(\mathbf{x})\right) \tag{2.79}$$

and the probabilistic version of Jensen's inequality (Theorem 2.8.4) gives the desired result

$$\Phi(u(\cdot,t)) \geqslant \Phi(w)\,. \tag{2.80}$$

ii) Continuity and differentiability are clear since all occurring functions are at least C^2. It remains to show that the derivative of V is non-positive. Since w does not depend on t, we compute

$$\frac{\partial}{\partial t} V(t) = \int_\Omega w \frac{\partial}{\partial t} r\left(\frac{u}{w}\right) \mathrm{d}\mathbf{x} \tag{2.81a}$$

$$= \int_\Omega r'\left(\frac{u}{w}\right) \partial_t u \,\mathrm{d}\mathbf{x} \tag{2.81b}$$

$$= \int_\Omega r'\left(\frac{u}{w}\right) \operatorname{div}(\nabla u - \mathbf{d}u) \,\mathrm{d}\mathbf{x}\,. \tag{2.81c}$$

For

$$g = r'\left(\frac{u}{w}\right) \qquad \text{and} \qquad F = \nabla u - \mathbf{d}u \tag{2.82}$$

we have

$$\nabla g = r''\left(\frac{u}{w}\right) \nabla\left(\frac{u}{w}\right) \tag{2.83}$$

and since

$$F \cdot \mathbf{n} = (\nabla u - \mathbf{d}u) \cdot \mathbf{n} = 0 \tag{2.84}$$

by the boundary condition (2.10c), applying (2.76) (divergence theorem) yields

$$\frac{\partial}{\partial t}V(t) = -\int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot(\nabla u - \mathbf{d}u)\,\mathrm{d}\mathbf{x} \tag{2.85a}$$

$$= -\int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot\Big(\nabla u - \mathbf{d}u - w\nabla\left(\frac{u}{w}\right) + w\nabla\left(\frac{u}{w}\right)\Big)\,\mathrm{d}\mathbf{x} \tag{2.85b}$$

$$= -\int_\Omega r''\left(\frac{u}{w}\right) w\left(\nabla\left(\frac{u}{w}\right)\right)^2\,\mathrm{d}\mathbf{x} - \int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot\left(\nabla u - \mathbf{d}u - w\nabla\left(\frac{u}{w}\right)\right)\,\mathrm{d}\mathbf{x}. \tag{2.85c}$$

The second integral is zero as the following computation shows:

$$-\int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot\left(\nabla u - \mathbf{d}u - w\nabla\left(\frac{u}{w}\right)\right)\,\mathrm{d}\mathbf{x} \tag{2.86a}$$

$$= -\int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot\left(\nabla u - \mathbf{d}u - w\frac{w\nabla u - u\nabla w}{w^2}\right)\,\mathrm{d}\mathbf{x} \tag{2.86b}$$

$$= -\int_\Omega r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot\left(-\mathbf{d}u + \frac{u}{w}\nabla w\right)\,\mathrm{d}\mathbf{x} \tag{2.86c}$$

$$= -\int_\Omega r''\left(\frac{u}{w}\right)\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right)\cdot(\nabla w - \mathbf{d}w)\,\mathrm{d}\mathbf{x}. \tag{2.86d}$$

We apply (2.76) (divergence theorem) to

$$g = r'\left(\frac{u}{w}\right) \qquad \text{and} \qquad F = \left(\frac{u}{w}\right)(\nabla w - \mathbf{d}w)\,. \tag{2.87}$$

Then we have

$$\nabla g = r''\left(\frac{u}{w}\right)\nabla\left(\frac{u}{w}\right) \tag{2.88}$$

and

$$\operatorname{div} F = \nabla\left(\frac{u}{w}\right)\cdot(\nabla w - \mathbf{d}w) + \left(\frac{u}{w}\right)\operatorname{div}(\nabla w - \mathbf{d}w) \tag{2.89a}$$

$$= \nabla\left(\frac{u}{w}\right)\cdot(\nabla w - \mathbf{d}w) \tag{2.89b}$$

since w solves the PDE (2.66),

$$\operatorname{div}\left(\nabla w - \mathbf{d}w\right) = 0. \tag{2.90}$$

The scalar product of F with $\mathbf{n}$ is zero by the boundary condition (2.10c). Hence, the integral (2.86d) simplifies to

$$-\int_\Omega r'\left(\frac{u}{w}\right) \nabla\left(\frac{u}{w}\right) \cdot \left(\nabla w - \mathbf{d}w\right)\, \mathrm{d}\mathbf{x}. \tag{2.91}$$

Another application of the divergence theorem with

$$g = r\left(\frac{u}{w}\right) \qquad \text{and} \qquad F = \left(\nabla w - \mathbf{d}w\right) \tag{2.92}$$

yields

$$-\int_\Omega r\left(\frac{u}{w}\right) \operatorname{div}\left(\nabla w - \mathbf{d}w\right)\, \mathrm{d}\mathbf{x} \tag{2.93}$$

since

$$\nabla g = r'\left(\frac{u}{w}\right) \nabla\left(\frac{u}{w}\right) \qquad \text{and} \qquad F \cdot \mathbf{n} = 0. \tag{2.94}$$

Again, we use that w solves (2.66). Therefore, the integral is zero and we obtain

$$\frac{\partial}{\partial t} V(t) = -\int_\Omega r''\left(\frac{u}{w}\right) w \left(\nabla\left(\frac{u}{w}\right)\right)^2 \mathrm{d}\mathbf{x} \tag{2.95}$$

which is non-positive since r has a non-negative second derivative as a convex function, w is positive by assumption and the last factor is a quadratic term.

iii) In i) we showed

$$\Phi(u(\cdot,t)) \geqslant \Phi(w) \tag{2.96}$$

by applying Jensen's inequality. For a strictly convex function r we can utilize the second part of Jensen's inequality.

$$\Phi(u(\cdot,t)) = \Phi(w) \tag{2.97a}$$

$$\Leftrightarrow \quad \frac{u(\cdot,t)}{w} \text{ is constant } \delta\text{-a.e. on } \Omega \qquad \text{Jensen's inequality} \tag{2.97b}$$

$$\Leftrightarrow \quad \frac{u(\cdot,t)}{w} \text{ is constant } \lambda\text{-a.e. on } \Omega \qquad \text{equivalence of } \delta \text{ and } \lambda \tag{2.97c}$$

Furthermore, $u(\cdot,t)$ and w have the same average grey value, hence,

$$u(\cdot,t) = w \qquad \text{a.e. on } \Omega \tag{2.98}$$

for all $t \geqslant 0$. Additionally, we know that for $t \geqslant 0$ both, $u(\cdot,t)$ and w are continuous on $\overline{\Omega}$. This shows

$$\Phi(u(\cdot,t)) = \Phi(w) \qquad \Longleftrightarrow \qquad u(\cdot,t) = w \text{ on } \overline{\Omega} \tag{2.99}$$

for all $t \geqslant 0$.

iv) Let $t > 0$. In ii) we computed

$$\frac{\partial}{\partial t} V(t) = -\int_{\Omega} r''\left(\frac{u}{w}\right) w \left(\nabla\left(\frac{u}{w}\right)\right)^2 \, \mathrm{dx}. \tag{2.100}$$

Since both w and r'' are positive on Ω, we have

$$V'(t) = 0 \Longleftrightarrow \nabla\left(\frac{u(\cdot,t)}{w}\right) = 0 \text{ a.e. on } \Omega \tag{2.101a}$$

$$\Longleftrightarrow \frac{u(\cdot,t)}{w} \text{ is constant a.e. on } \Omega \tag{2.101b}$$

and the claim follows as in iii).

v) Let $V(0) = V(T)$ for some $T > 0$. Since $\frac{\partial}{\partial t} V(t) \leqslant 0$, it follows

$$V(t) = \textit{const}. \text{ on } [0,T]. \tag{2.102}$$

Hence, $V'(t) = 0$ on $(0,T)$ and by iv)

$$u(\cdot,t) = w \qquad \text{on } \overline{\Omega} \times (0,T). \tag{2.103}$$

By continuity (Theorem 2.4.1) this carries over to $t = 0, T$.

The other direction is quite simple since $V(0) = V(T)$ follows from

$$u(\cdot,t) = w \qquad \text{on } \overline{\Omega} \tag{2.104}$$

for all $t \in [0,T]$.

□

This proof shows that, indeed, the functional defined by

$$\Phi(u(\cdot,t)) = \int_\Omega w(\mathbf{x}) r\left(\frac{u(\mathbf{x},t)}{w(\mathbf{x})}\right) \mathrm{d}\mathbf{x} \tag{2.105}$$

is a strict Lyapunov functional for linear osmosis filter. As a consequence, the image evolution $u(.,t)$ fulfils a simplification property in term of Lyapunov functionals. For specific choices of the convex function this also allows an information theoretic interpretation. Especially, if we choose

$$r(s) = s\log(s) \tag{2.106}$$

the Lyapunov functional simplifies to

$$\int_\Omega u(\mathbf{x},t)\log\left(\frac{u(\mathbf{x},t)}{w(\mathbf{x})}\right) \mathrm{d}\mathbf{x}\,. \tag{2.107}$$

This term is known as *relative entropy* or *Kullback–Leibler distance* [see e.g. Cover and Thomas, 2006, Section 8.5].
In the next section, we use

$$r(s) = (s-1)^2 \tag{2.108}$$

to show convergence of $u(.,t)$ towards w in $L^2(\Omega)$.

2.9 Convergence

The next step after having found a Lyapunov functional is to use it to prove convergence of the time evolution.

Theorem 2.9.1 (Convergence)**.** *Using the notation from Theorem 2.8.1, we have*

$$\lim_{t\to\infty} \|u(\cdot,t) - w\|_{L^2(\Omega)} = 0. \tag{2.109}$$

For the proof we follow an idea of Illner and Neunzert [1993] where the authors used Lyapunov functionals to prove convergence for directed diffusion equation

$$\begin{aligned} \partial_t u &= w\Delta u - u\Delta w \\ &= \operatorname{div}\left(w\nabla u - u\nabla w\right)\,. \end{aligned} \tag{2.110}$$

Proof. From Theorem 2.8.1, we know that

$$V(t) = \int_{\Omega} w\, r\left(\frac{u(\cdot,t)}{w}\right)\, \mathrm{d}\mathbf{x} \tag{2.111}$$

is bounded from below and non-increasing with derivative

$$\frac{\partial}{\partial t}V(t) = -\int_{\Omega} r''\left(\frac{u}{w}\right) w \left(\boldsymbol{\nabla}\left(\frac{u}{w}\right)\right)^2 \mathrm{d}\mathbf{x}. \tag{2.112}$$

Hence, the limit $\lim_{t\to\infty} V(t)$ exists.
Now, we choose

$$r(x) = \frac{1}{2}(x-1)^2. \tag{2.113}$$

On the one hand, this gives

$$\frac{\partial}{\partial t}V(t) = -\int_{\Omega} w \left(\boldsymbol{\nabla}\left(\frac{u}{w}\right)\right)^2 \mathrm{d}\mathbf{x}, \tag{2.114}$$

meaning that V is the antiderivative of the right hand side. Therefore,

$$0 \leqslant \int_0^{\infty} \left(\int_{\Omega} w \left(\boldsymbol{\nabla}\left(\frac{u}{w}\right)\right)^2 \mathrm{d}\mathbf{x} \right) \mathrm{d}t = \lim_{t\to\infty} -[V(t) - V(0)] < \infty. \tag{2.115}$$

Since

$$\int_0^{\infty} \left(\int_{\Omega} w \left(\boldsymbol{\nabla}\left(\frac{u}{w}\right)\right)^2 \mathrm{d}\mathbf{x} \right) \mathrm{d}t < \infty \tag{2.116}$$

and $w \geqslant \varepsilon > 0$ on Ω, there exists a sequence $(t_i)_{i\in\mathbb{N}}$ in $\mathbb{R}$ such that

$$\lim_{i\to\infty} t_i = \infty \tag{2.117}$$

and

$$\lim_{i\to\infty} \left|\left| \boldsymbol{\nabla}\left(\frac{u(t_i)}{w}\right) \right|\right|_{L^2(\Omega)} = 0, \tag{2.118}$$

otherwise the integral in (2.116) would be infinite.

On the other hand, knowing that

$$V(t) = \int_{\Omega} w \left(\frac{u(\cdot, t)}{w} - 1 \right)^2 \, \mathrm{d}\mathbf{x} \tag{2.119}$$

is bounded from below and non-increasing, means $(V(t_i))_{i \in \mathbb{N}}$ is a convergent sequence in $\mathbb{R}$ and by the boundedness of w (Theorem 2.7.6),

$$\left(\frac{u(t_i)}{w} \right)_{i \in \mathbb{N}} \tag{2.120}$$

is a bounded sequence in $L^2(\Omega)$. By Alaoglu's theorem there is a weakly convergent subsequence, also denoted by $(t_i)_{i \in \mathbb{N}}$, with weak limit g in $L^2(\Omega)$. Next, we apply Rellich's Lemma (Richtmyer [1978] or Attouch et al. [1987, Theorem 5.4.2]), to see that there is another subsequence, also denoted by $(t_i)_{i \in \mathbb{N}}$ such that

$$\left(\frac{u(t_i)}{w} \right)_{i \in \mathbb{N}} \tag{2.121}$$

converges to a function g strongly in $L^2(\Omega)$.
It follows

$$\nabla \left(\frac{u(t_i)}{w} \right) \longrightarrow \nabla g \tag{2.122}$$

in the sense of distributions. We also know from (2.118)

$$\lim_{i \to \infty} \left\| \nabla \left(\frac{u(t_i)}{w} \right) \right\|_{L^2(\Omega)} = 0. \tag{2.123}$$

Hence, weakly we have

$$\nabla g = 0 \quad \text{on } \Omega\,. \tag{2.124}$$

This means g is constant and by the grey value invariance,

$$g = 1 \quad \text{on } \Omega. \tag{2.125}$$

Thus, we have

$$\frac{u(t_i)}{w} \longrightarrow 1 \quad \text{strongly in } L^2(\Omega) \tag{2.126}$$

for some sequence $(t_i) \longrightarrow \infty$ and therefore,

$$\lim_{t \to \infty} \|u(\cdot, t) - w\|_{L^2(\Omega)} = 0. \tag{2.127}$$

□

This theorem shows that we have the expected convergence of the parabolic solution towards the elliptic solution that has the same average grey value as the initial image of the parabolic problem.

A similar result also holds in a discrete setting [Vogel et al., 2013].

2.10 Summary

To sum up the results of this chapter: We have developed a complete scale-space theory for the linear osmosis filtering. We have shown the existence of a unique, classical solution of the parabolic PDE

$$\begin{aligned} \partial_t u &= \operatorname{div}(\boldsymbol{\nabla} u - \mathbf{d} u) && \text{on } \Omega \times (0, T] \\ u(.,0) &= f(.) && \text{on } \overline{\Omega} \\ \langle \boldsymbol{\nabla} u, \mathbf{n} \rangle &= 0 && \text{on } \partial\Omega \times [0, T]. \end{aligned}$$

Furthermore, we have seen, that the average grey value stays constant over the evolution and that positivity is presevered. By analysing the steady state equation, we were able to prove an information reducing property in terms of Lyapunov functionals. At the end, we have also seen a convergence result stating L^2-convergence of the parabolic solution towards a specific solution of the elliptic problem.

Chapter 3

Morphological Counterparts of Linear Scale-Spaces

> The mathematician's patterns, like the painter's or the poet's must be beautiful; the ideas like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics.
>
> Godfrey H. Hardy

We want to take a closer look at linear shift-invariant and morphological scale-spaces. Both classes of scale-spaces are well-known in the literature and there are several attempts to connect them. In this Chapter we review these approaches and introduce the novel Cramér–Fourier transform. This allows us to extend known results and connect the linear shift-invariant and morphological scale-spaces on both a structural level and on the level of evolution equations. We start by introducing the classes of scale-spaces under consideration in detail.

The Sections 3.1, 3.2 and 3.4–3.7 are based on Schmidt and Weickert [2016].

3.1 Linear Shift-Invariant Scale-Spaces

This section provides a general framework for the so-called linear shift-invariant (LSI) scale-spaces. While their image evolution is steered by a pseudodifferential operator the corresponding convolution kernel can be described in the Fourier

domain. Furthermore, this section introduces the *symbol* as an alternative description of the pseudodifferential operator. It reduces the analysis from a hard to handle pseudodifferential operator to a mere function in two variables.

3.1.1 Pseudodifferential Evolutions

Pseudodifferential operators are a generalisation of partial differential operators. In order to have a compact notation, we use the following multi-index notation: For a multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{N}^2$ and some vector $\mathbf{v} \in \mathbb{R}^2$ we define $|\boldsymbol{\alpha}| := \alpha_1 + \alpha_2$ and $\mathbf{v}^{\boldsymbol{\alpha}} := v_1^{\alpha_1} v_2^{\alpha_2}$.
A kth-order *partial differential operator* P is then given by

$$P(\mathbf{x}, \boldsymbol{\nabla}) = \sum_{|\boldsymbol{\alpha}| \leqslant k} c_{\boldsymbol{\alpha}}(\mathbf{x}) \boldsymbol{\nabla}^{\boldsymbol{\alpha}}, \tag{3.1}$$

where $\boldsymbol{\nabla} = \left(\partial_x, \partial_y\right)^\top$ denotes the *spatial gradient*, $c_{\boldsymbol{\alpha}}$ are real-valued functions on the image domain Ω and $\mathbf{x}$ is a point in Ω. The Fourier transform allows us to obtain a different characterisation of the differential operator (3.1). We use the following convention to define the *Fourier transform* of an integrable function u:

$$\hat{u}(\boldsymbol{\xi}) := \mathcal{F}[u](\boldsymbol{\xi}) := \int_{\mathbb{R}^2} u(\mathbf{x})\, e^{-2\pi \mathrm{i} \langle \boldsymbol{\xi}, \mathbf{x} \rangle} \, \mathrm{d}\mathbf{x} \tag{3.2}$$

where $\mathrm{i}^2 = -1$ and $\langle ., . \rangle$ denotes the Hermitian inner product. The above choices of constants make the Fourier transform an unitary transform with inverse given by

$$u(\mathbf{x}) = \mathcal{F}^{-1}[\,\hat{u}\,](\mathbf{x}) := \int_{\mathbb{R}^2} \hat{u}(\boldsymbol{\xi})\, e^{2\pi \mathrm{i} \langle \mathbf{x}, \boldsymbol{\xi} \rangle} \, \mathrm{d}\boldsymbol{\xi}\,. \tag{3.3}$$

The integral on the right-hand side is the L^2 scalar product of the Fourier transform of u and a basis function $e^{2\pi i \langle \mathbf{x}, \cdot \rangle}$. It can be regarded as a frequency decomposition of the original image u. For derivatives, this decomposition yields

$$\boldsymbol{\nabla}^{\boldsymbol{\alpha}} u(\mathbf{x}) = \int_{\mathbb{R}^2} (2\pi \mathrm{i}\, \boldsymbol{\xi})^{\boldsymbol{\alpha}}\, \hat{u}(\boldsymbol{\xi})\, e^{2\pi \mathrm{i} \langle \mathbf{x}, \boldsymbol{\xi} \rangle} \, \mathrm{d}\boldsymbol{\xi}\,. \tag{3.4}$$

Using the linearity of the Fourier transform we can also use the pseudodifferential operator P defined in (3.1) to obtain

$$P(\mathbf{x}, \boldsymbol{\nabla}) u(\mathbf{x}) = \int_{\mathbb{R}^2} p\,(\mathbf{x}, 2\pi\, \boldsymbol{\xi})\; \hat{u}(\boldsymbol{\xi})\, e^{2\pi \mathrm{i} \langle \mathbf{x}, \boldsymbol{\xi} \rangle} \, \mathrm{d}\boldsymbol{\xi} \tag{3.5}$$

where the polynomial

$$p(\mathbf{x}, \xi) = \sum_{|\alpha| \leqslant k} c_\alpha(\mathbf{x})(\mathrm{i}\,\xi)^\alpha \tag{3.6}$$

is called the *symbol* of $P(\mathbf{x}, \nabla)$.

Since the right-hand side of (3.5) also makes sense in cases where p is not a polynomial, we use this equation to define a so-called *pseudodifferential operator* denoted by $P(\mathbf{x}, \nabla)$. This definition is slightly different from the one used by Taylor [2010] due to a different convention for the Fourier transform. By restricting our attention to pseudodifferential equations with constant coefficients we can guarantee shift-invariance of the resulting system. This means that we have to restrict p to functions that are independent of the location $\mathbf{x}$. We drop these spatial parameters in the following discussion. The resulting pseudodifferential operator $P(\nabla)$ then defines a scale-space by the following initial value problem

$$\partial_t u = P(\nabla) u \tag{3.7a}$$

$$u(., 0) = f\,. \tag{3.7b}$$

We call linear shift-invariant (LSI) evolutions of this type *LSI scale-spaces*. They comprise many well-known linear scale-spaces. We give an overview in Section 3.1.3.

3.1.2 Interpretation as Convolution Scale-Spaces

Let us now interpret LSI scale-spaces in terms of convolutions with appropriate kernels. Since equation (3.7a) is linear, it can be described by a multiplication in the Fourier domain where the factor is given by the symbol. To see this, we apply the Fourier transform to (3.5) and obtain

$$\mathcal{F}\,[\,P(\nabla) u(\cdot, t)\,]\,(\xi) = p(2\pi\,\xi)\hat{u}(\xi, t)\,. \tag{3.8}$$

Therefore, (3.7a)–(3.7b) simplifies under the Fourier transform to the initial value problem

$$\partial_t \hat{u}(\xi, t) = p(2\pi\,\xi)\hat{u}(\xi, t)\,, \tag{3.9a}$$

$$\hat{u}(\xi, 0) = \hat{f}(\xi)\,. \tag{3.9b}$$

Its solution is given by

$$\hat{u}(\xi, t) = \hat{f}(\xi)\, \exp\left(p(2\pi\,\xi)\, t\right)\,. \tag{3.10}$$

Applying the inverse Fourier transform translates the result to the spatial domain again:

$$u(\mathbf{x},t) = \left(f * \mathcal{F}^{-1}[\exp\left(p(2\pi\cdot)\,t\right)]\right)(\mathbf{x})\,. \tag{3.11}$$

This shows the importance of the symbol p: We can use it to characterise the solution of (3.7a)–(3.7b) as a convolution of the initial image $f(\mathbf{x})$ with an appropriate kernel $k(\mathbf{x},t)$:

$$u(\mathbf{x},t) = (f * k(\cdot,t))(\mathbf{x}), \tag{3.12a}$$
$$k(\mathbf{x},t) = \mathcal{F}^{-1}[\exp\left(p(2\pi\cdot)\,t\right)](\mathbf{x}). \tag{3.12b}$$

3.1.3 Examples of LSI Scale-Spaces

In order to illustrate that the family of LSI scale-spaces is fairly rich, let us investigate five examples in more detail.

1. **Gaussian Scale-Space.** It computes smoothed versions $u(\mathbf{x},t)$ of $f(\mathbf{x})$ as solutions of the initial value problem

$$\partial_t u = \Delta u \qquad \text{on } \mathbb{R}^2 \times (0,\infty), \tag{3.13a}$$
$$u(.,0) = f \qquad \text{on } \mathbb{R}^2, \tag{3.13b}$$

   where $\Delta = \partial_{xx} + \partial_{yy}$ denotes the spatial Laplacian. Gaussian scale-space goes back to Iijima [1962] and Weickert et al. [1999]. It became popular in the western world by the work of Witkin [1983], Koenderink [1984], Lindeberg [1994], Florack [1997], and many others; see e.g. Sporring et al. [1997] and the references therein.

2. **α-Scale-Spaces.** These evolutions replace the homogeneous diffusion equation (3.13a) by the pseudodifferential equation

$$\partial_t u = -(-\Delta)^{\alpha} u \tag{3.14}$$

   with some parameter $\alpha \in (0,\infty)$. While such processes can already be found implicitly in Iijima's early work [Iijima, 1962] and more explicitly e.g. in a publication by Pauwels et al. [1995], they became popular as scale-spaces due to the work of Duits et al. [2004]. Gaussian scale-space is recovered for $\alpha = 1$, while $\alpha = \frac{1}{2}$ gives the so-called Poisson scale-space

$$\partial_t u = -\sqrt{-\Delta}\, u \tag{3.15}$$

of Felsberg and Sommer [2001]. If one renounces a maximum–minimum principle, one can also study scale-spaces for $\alpha > 1$, comprising e.g. the biharmonic scale-space for $\alpha = 2$ [Didas et al., 2005].

3. **Summed α-Scale-Spaces.** Didas et al. [2005] discuss finite linear combinations of fractional Laplacians

$$\partial_t u = -\sum_{k=1}^{m} \lambda_k (-\Delta)^{\alpha_k} u \tag{3.16}$$

with fractional derivative orders $\alpha_1, \ldots, \alpha_m > 0$ and weights $\lambda_1, \ldots, \lambda_m > 0$. Interestingly they can satisfy a maximum–minimum principle even if some terms with $\alpha > 1$ are present, as long as they are dominated by terms with $\alpha < 1$.

The special case of a linear combination of one Gaussian and one Poisson kernel is used in Kanters et al. [2007] to approximate α-scale-spaces.

4. **Relativistic Scale-Spaces.** Burgeth et al. [2005a] have advocated a generalisation of the Poisson scale-space by considering the evolution equations

$$\partial_t u = -\left(\sqrt{m^2 - \Delta} - m\right) u \tag{3.17}$$

with $m \geqslant 0$. We see that this family contains the Poisson scale-space for $m = 0$.

5. **Anisotropic Scale-Spaces.** Formally one can construct anisotropic versions of any of the preceding scale-spaces by replacing their Laplacian by $\nabla^\top \mathbf{D} \nabla$ with some symmetric positive definite matrix $\mathbf{D} \in \mathbb{R}^{2\times 2}$. In the case of Gaussian scale-space, this leads to the anisotropic Gaussian scale-spaces

$$\partial_t u = \operatorname{div}(\mathbf{D} \nabla u)\,. \tag{3.18}$$

They have been derived axiomatically by Iijima [1963, 1971] and later on by Lindeberg [1994, 2011].

We discussed scale-space properties of nonlinear variants where the diffusion tensor $\mathbf{D}$ is a function of the local structure of the evolving image in Section 1.3.

Although these scale-spaces differ w.r.t. decay behaviour in Fourier space, separability, extremum principle, nonenhancement of local extrema and scale invariance, the pseudodifferential operators $P(\nabla)$ and their corresponding kernels

$k(\mathbf{x}, t)$ can be computed following the strategy in Subsections 3.1.1 and 3.1.2. The results are summarised in Table 3.1. Note that the symbol representation allows simple formulas even in those cases where the corresponding kernels do not have a closed form representation. Therefore, we will also use it later on for establishing correspondences to morphological scale-spaces.

In most cases, working with the symbol is much easier compared to using the kernel.

Although the above examples give a compact overview of known linear scale-spaces, there are many more approaches to construct scale-spaces. For example, regularisation methods and related concepts can be interpreted as scale-spaces by considering their Euler-Lagrange equations, both in the linear and the nonlinear setting [see e.g. Poggio et al., 1988, Nielsen et al., 1997, Scherzer and Weickert, 2000, Burgeth et al., 2005b, Demetz et al., 2012]. Since Gaussian scale-space can be described by a linear diffusion equation, it is natural to generalise it also to nonlinear diffusion scale-spaces [Perona and Malik, 1990, Weickert, 1998]. On the morphological side, continuous-scale versions of erosions are given by hyperbolic PDEs [Alvarez et al., 1993, Arehart et al., 1993, Brockett and Maragos, 1992, van den Boomgaard and Smeulders, 1994]. Parabolic morphological PDEs comprise mean curvature motion [Alvarez et al., 1992, Kimia and Siddiqi, 1996] which can be derived from iterated median filtering [Guichard and Morel, 1997]. It is also possible to construct affine invariant morphological scale-spaces [Alvarez et al., 1993, Sapiro and Tannenbaum, 1993]. Moreover, morphological variants of linear and nonlinear diffusion scale-spaces can be created by embedding these scale-spaces into a counter-harmonic framework [Angulo, 2010].

Table 3.1: Specific LSI scale-spaces, their evolution equations, symbols, and convolution kernels. The gamma function is denoted by Γ, and K_ν refers to the modified Bessel function of the third kind [Abramowitz and Stegun, 1974].

LSI scale-space	evolution equation	symbol	kernel
alpha	$\partial_t u = -(-\Delta)^\alpha u$	$p(\xi) = -\lvert\xi\rvert^{2\alpha}$	no closed formula
Gaussian	$\partial_t u = \Delta u$	$p(\xi) = -\lvert\xi\rvert^2$	$k(\mathbf{x}) = \frac{1}{4\pi} \exp\left(-\frac{\lvert\mathbf{x}\rvert^2}{4}\right)$
Poisson	$\partial_t u = -\sqrt{-\Delta}\, u$	$p(\xi) = -\lvert\xi\rvert$	$k(\mathbf{x}) = \frac{\Gamma(\frac{3}{2})}{\pi^{3/2}} \frac{1}{(1+\lvert\mathbf{x}\rvert^2)^{3/2}}$
summed alpha	$\partial_t u = -\sum_{k=1}^{m} \lambda_k (-\Delta)^{\alpha_k} u$	$p(\xi) = -\sum_{k=1}^{m} \lambda_k \lvert\xi\rvert^{2\alpha_k}$	no closed formula
relativistic	$\partial_t u = \left(m - \sqrt{m^2 - \Delta}\right) u$	$p(\xi) = m - \sqrt{m^2 + \lvert\xi\rvert^2}$	$k(\mathbf{x}) = \left(\frac{m}{2\pi}\right)^{3/2} \frac{2e^m}{(1+\lvert\mathbf{x}\rvert^2)^{3/4}} K_{\frac{3}{2}}\left(m\sqrt{1+\lvert\mathbf{x}\rvert^2}\right)$
anis. Gaussian	$\partial_t u = \operatorname{div}(\mathbf{D}\,\nabla u)$	$p(\xi) = -\langle \xi, \mathbf{D}\,\xi \rangle$	$k(\mathbf{x}) = \frac{1}{4\pi \det \mathbf{D}} \exp\left(-\frac{\mathbf{x}^\top \mathbf{D}^{-1} \mathbf{x}}{4}\right)$

3.2 Morphological Scale-Spaces Revisited

In this section we want to revive the discussion of the morphological scale-spaces in Section 1.4. We will use an alternative description and follow the general structure of Section 3.1 as close as possible. This emphasises the similarities and difference between linear shift-invariant and morphological scale-spaces.

3.2.1 Hamilton-Jacobi Equations

Morphological scale-spaces are given by *Hamilton–Jacobi equations* of type

$$\partial_t v = -H(\nabla v), \tag{3.19a}$$
$$v(\mathbf{x}, 0) = f(\mathbf{x}). \tag{3.19b}$$

In such a setting, many of the concepts from the Section 1.4 can be reformulated in the language of convex analysis.

In the sequel, we will assume that f is bounded and *lower semi-continuous*, i.e.

$$f(\mathbf{x}_0) \leqslant \liminf_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) \tag{3.20}$$

for all points $\mathbf{x}_0$ and that the function H is convex and *coercive*, i.e.

$$\lim_{|\mathbf{x}|\to\infty} \frac{H(\mathbf{x})}{|\mathbf{x}|} = \infty. \tag{3.21}$$

Unfortunately, the concept of weak or generalised solutions is not suitable to uniquely describe a solution of the Hamilton–Jacobi equation. As an example, let us consider the one-dimensional PDE

$$\partial_t u = -(u_x)^2 \qquad \text{on } \mathbb{R}\times(0,T) \tag{3.22a}$$
$$u(.,0) = 0 \qquad \text{on } \mathbb{R}. \tag{3.22b}$$

Then there are two solutions: Besides the constant function $u \equiv 0$ also the piecewise linear function

$$u(x,t) = \begin{cases} 0 & \text{for } |x| \geqslant t \geqslant 0 \\ -t + |x| & \text{for } t \geqslant |x|. \end{cases} \tag{3.23}$$

solves (3.22a)–(3.22b) with the exception of the lower dimensional subspaces given by $t = |x|$ and $x = 0$.

In order to enforce uniqueness of a solution, additional constraints are needed. This is were *viscosity solutions* come into play. They are defined as follows [see Evans, 1998, Section 10.1.1]:

Definition 3.2.1. *A bounded, uniformly continuous function u is called a viscosity solution of the Hamilton–Jacobi equation* (3.19a)–(3.19b)

$$\begin{aligned} \partial_t v &= -H(\nabla v), \\ v(\mathbf{x}, 0) &= f(\mathbf{x}) \end{aligned}$$

provided that $u(\cdot, t) = f$ on $\mathbb{R}^2$ and that for each $v \in C^\infty(\mathbb{R}^2 \times (0, T))$ both

$$\begin{cases} \text{if } u - v \text{ has a local maximum at a point } (\mathbf{x}_0, t_0) \in \mathbb{R}^2 \times (0, \infty) \\ \text{then } v_t(\mathbf{x}_0, t_0) + H(\nabla v(\mathbf{x}_0, t_0)) \leqslant 0\,, \end{cases}$$

and

$$\begin{cases} \text{if } u - v \text{ has a local minimum at a point } (\mathbf{x}_0, t_0) \in \mathbb{R}^2 \times (0, \infty) \\ \text{then } v_t(\mathbf{x}_0, t_0) + H(\nabla v(\mathbf{x}_0, t_0)) \geqslant 0 \end{cases}$$

holds true.

This definition is motivated by the method of vanishing viscosity [Crandall and Lions, 1983]. The idea is to add a second order term $\varepsilon\Delta$ with a small $\varepsilon > 0$ which converts the nonlinear first-order PDE into a second-order linear PDE. In many cases the second-order equation is easier to solve or even allows explicit solutions. Afterwards, the limit $\varepsilon \to 0$ is considered. The additional constraints in the definition of viscosity solutions ensure that this limit is well-defined. The name of the method originated from fluid dynamics where ε is the viscosity of a fluid.

Viscosity solutions have a number of theoretical properties which render them perfectly for solving hyperbolic transport equations. For an overview and more details we refer to Crandall et al. [1992].

3.2.2 Interpretation as Infimal Convolution Scale-Spaces

The unique viscosity solution of the Hamilton–Jacobi equation (3.19a)–(3.19b) is given by the *Hopf-Lax formula* [e.g. Heijmans, 2002] as

$$v(\mathbf{x}, t) = \min_{\mathbf{y} \in \mathbb{R}^2} \left\{ f(\mathbf{y}) + t\, H^* \left(\frac{\mathbf{x} - \mathbf{y}}{t} \right) \right\} \tag{3.25}$$

where H^* denotes the the convex conjugate of H, i.e.

$$H^*(\mathbf{x}) := \sup_{\mathbf{y} \in \mathbb{R}^2} \{ \langle \mathbf{y}, \mathbf{x} \rangle - H(\mathbf{y}) \}. \tag{3.26}$$

We define the *structuring function* (SF)

$$s(\mathbf{x}, t) := t\, H^*\left(\frac{\mathbf{x}}{t}\right) = (tH)^*(\mathbf{x})\,. \tag{3.27}$$

Up to a sign difference, this definition matches the one given in Section 1.4. Furthermore, dilations and erosions are replaced by an operation called *infimal convolution*

$$(f \,\square\, g)(\mathbf{x}) = \inf_{\mathbf{y}\in\mathbb{R}^2} \{f(\mathbf{y}) + g(\mathbf{x}-\mathbf{y})\}\,. \tag{3.28}$$

Then the solution of our morphological evolution is given by

$$v(\mathbf{x}, t) = (f \,\square\, s(\cdot, t))(\mathbf{x})\,, \tag{3.29a}$$
$$s(\mathbf{x}, t) = (tH)^*(\mathbf{x})\,. \tag{3.29b}$$

3.2.3 Examples of Morphological Scale-Spaces

In a similar way as our LSI framework covers a large family of linear scale spaces, the Hamilton–Jacobi formulation comprises also many morphological scale-spaces. We illustrate this by a number of examples.

1. **Dilation and Erosion Scale-Spaces.**

   As we have seen in Section 1.4, mathematical morphology is usually expressed in terms of dilations and erosions. The dilation and erosion were defined in (1.13) and (1.14) respectively by

   $$(f \oplus b)(\mathbf{x}) := \sup_{\mathbf{y}\in\mathbb{R}^2} \{f(\mathbf{y}) + b(\mathbf{x}-\mathbf{y})\},$$
   $$(f \ominus b)(\mathbf{x}) := \inf_{\mathbf{y}\in\mathbb{R}^2} \{f(\mathbf{y}) + \bar{b}(\mathbf{x}-\mathbf{y})\}$$

   They relate to the infimal convolution by

   $$f \ominus b \;=\; f \;\square\; \bar{b}, \tag{3.30}$$
   $$-(f \oplus b) = (-f) \,\square\, (-b)\,. \tag{3.31}$$

   Therefore, infimal convolutions behave essentially like erosions. However, instead of requiring concavity of the structuring functions, infimal convolutions expect structuring functions to be convex. Dilations can be obtained by applying an infimal convolution with the negative of the structuring function to the negative of the signal.

   For these reasons, results for dilations and erosions are equivalent to results for infimal convolutions.

2. **Quadratic Structuring Function Scale-Space.** Taking $s(\mathbf{x}) = \frac{1}{4}|\mathbf{x}|^2$ as structuring function, equation (3.26) implies that $H(\mathbf{x}) = |\mathbf{x}|^2$. In this case, the infimal convolution

$$v(\mathbf{x}, t) = (f \,\square\, s(\cdot, t))\,(\mathbf{x}) \tag{3.32}$$

is the viscosity solution of the evolution process

$$\partial_t v = -|\nabla v|^2\,, \tag{3.33a}$$

$$v(0) = f\,. \tag{3.33b}$$

van den Boomgaard [1992b] has shown that quadratic structuring functions such as

$$s(\mathbf{x}) = \tfrac{1}{4}|\mathbf{x}|^2 \tag{3.34}$$

are the only structuring functions that are rotationally invariant and separable. This has motivated him to regard (3.33a)–(3.33b) as the morphological equivalent of Gaussian scale-space, since the latter one is the only scale-space with a rotationally invariant and separable convolution kernel [Otsu, 1981, Weickert et al., 1999].

3. **Scale-Spaces with Flat Disc Structuring Functions.** If one uses as structuring function a flat disc

$$s(\mathbf{x}) \;=\; \begin{cases} 0 & \text{for } |\mathbf{x}| \leqslant 1, \\ \infty & \text{else,} \end{cases} \tag{3.35}$$

it has been shown in Alvarez et al. [1993], Arehart et al. [1993], Brockett and Maragos [1992] that one arrives at

$$\partial_t v = -|\nabla v|. \tag{3.36}$$

Evolutions of this type can be interpreted in many ways as scale-spaces; see Alvarez et al. [1993], van den Boomgaard and Smeulders [1994], Jackway and Deriche [1996] for more details.

4. **Structuring Functions of Arbitrary Power.** Jackway [1994] as well as Diop and Angulo [2014] have investigated morphological processes that can be described by evolution equations of type

$$\partial_t v = -|\nabla v|^{\beta} \tag{3.37}$$

with arbitrary powers $\beta > 1$. Their corresponding structuring functions are given by the poweroids

$$s(\mathbf{x}) = (\beta - 1)\,|\mathbf{x}/\beta|^{\beta/(\beta-1)}\,. \tag{3.38}$$

5. **Anisotropic Structuring Functions.** So far, all our morphological scale-spaces use isotropic structuring functions that do not favour specific directions. Depending on the application, it can be beneficial to consider also anisotropic structuring functions that are adapted to directions of special interest.
   An early anisotropic morphological PDE model goes back to Arehart et al. [1993]: They have used ellipse-shaped flat structuring functions

$$s(\mathbf{x}) = \begin{cases} 0 & \mathbf{x}^\top \mathbf{D}^{-1} \mathbf{x} \leqslant 1, \\ \infty & \text{else} \end{cases} \tag{3.39}$$

   with some positive definite symmetric matrix $\mathbf{D}$. This leads to evolutions of type

$$\partial_t v = -|\mathbf{D}\,\boldsymbol{\nabla} v|\,. \tag{3.40}$$

   Later on, Breuß et al. [2007] have adapted $\mathbf{D}$ to the underlying local image structure.
   Evolutions with anisotropic quadratic structuring functions

$$s(\mathbf{x}) = \frac{1}{4} \mathbf{x}^\top \mathbf{D}^{-1} \mathbf{x} \tag{3.41}$$

   can be described by

$$\partial_t v = -\left(\boldsymbol{\nabla} v\right)^\top \mathbf{D}\,\boldsymbol{\nabla} v. \tag{3.42}$$

   Such processes go back to van den Boomgaard [1992a] and Jackway [1995] in a space-invariant setting. More recently, Landström [2015] has considered space-adaptive generalisations.

Table 3.2 gives a compact representation of the morphological examples that we have discussed.

Table 3.2: Specific morphological scale-spaces of Hamilton–Jacobi type and their structuring functions (SFs).

morphological scale-space	evolution equation	structuring function
quadratic SF	$\partial_t v = -\lvert\nabla v\rvert^2$	$s(\mathbf{x},t) = \frac{1}{4t}\lvert\mathbf{x}\rvert^2$
flat disk SF	$\partial_t v = -\lvert\nabla v\rvert$	$s(\mathbf{x},t) = \begin{cases} 0 & \lvert\mathbf{x}\rvert \leqslant t \\ \infty & \text{else} \end{cases}$
poweroid SF	$\partial_t v = -\lvert\nabla v\rvert^\beta$	$s(\mathbf{x},t) = t\,(\beta-1)\left\lvert\frac{\mathbf{x}}{t\,\beta}\right\rvert^{\beta/(\beta-1)}$
flat ellipse-shaped SF	$\partial_t v = -\lvert\mathbf{D}\,\nabla v\rvert$	$s(\mathbf{x},t) = \begin{cases} 0 & \mathbf{x}^\top\mathbf{D}^{-1}\mathbf{x} \leqslant t^2 \\ \infty & \text{else} \end{cases}$
anisotropic quadratic SF	$\partial_t v = -\left(\nabla v\right)^\top \mathbf{D}\,\nabla v$	$s(\mathbf{x},t) = \frac{1}{4t}\mathbf{x}^\top\mathbf{D}^{-1}\mathbf{x}$

3.3 Related Work

Both morphological and linear shift-invariant PDEs are well-studied topics with many interconnections. This section gives an overview over previous approaches.

3.3.1 Hopf-Cole Transform

Hopf [1950] and Cole [1951] transformed the nonlinear Burgers' equation

$$\partial_t u + u u_x = \varepsilon u_{xx} \tag{3.43}$$

into a linear heat equation with known solution. This allows them to give an explicit solution of (3.43). Their method can be a generalised to PDEs of the form

$$\partial_t u - a\Delta u + b\|\nabla u\|^2 = 0\,. \tag{3.44}$$

The transform

$$w = \exp\left(-\frac{b}{a}u\right) \tag{3.45}$$

converts (3.44) into a linear PDE [see Evans, 1998]. It is named after them and nowadays known as *Hopf-Cole transform*.

The same concepts are also able to solve the morphological initial value problem

$$\partial_t u = |\nabla u|^2, \tag{3.46a}$$
$$u(.,0) = f \tag{3.46b}$$

where the Hamiltonian $H(p) = p^2$ is used. In this case we have already seen that the solution is given by an dilation with $-H^*(p) = -\frac{1}{4}p^2$ as structuring function. This gives

$$\begin{aligned} u(\mathbf{x},t) &= (f \oplus (-H^*))(\mathbf{x}) \\ &= \sup_{\mathbf{y}\in\mathbb{R}^2} \left\{ f(\mathbf{y}) - \tfrac{1}{4t}|\mathbf{x}-\mathbf{y}|^2 \right\} \end{aligned} \tag{3.47}$$

as solution of (3.46a)–(3.46b).

The second possibility for solving (3.46a)–(3.46b) is to introduce a small perturbation to the partial differential operator. Adding a small multiple of the Laplacian yields a new initial value problem of type (3.44):

$$\partial_t u_\varepsilon = \varepsilon \Delta u_\varepsilon + |\nabla u_\varepsilon|^2 \tag{3.48a}$$
$$u_\varepsilon(.,0) = f\,. \tag{3.48b}$$

We use u_ε to emphasise the dependence of the problem on the choice of ε. Using the Hopf-Cole transform

$$w_\varepsilon = \exp\left(\tfrac{1}{\varepsilon} u_\varepsilon\right) \tag{3.49}$$

turns (3.48a)–(3.48b) into the linear heat equation

$$\partial_t w_\varepsilon = \varepsilon \Delta w_\varepsilon \tag{3.50a}$$
$$w_\varepsilon(\mathbf{x},0) = \exp\left(\tfrac{1}{\varepsilon} f(\mathbf{x})\right)\,. \tag{3.50b}$$

Its solution is given by a convolution of the initial value $\exp\left(\frac{1}{\varepsilon} f(\mathbf{x})\right)$ with a Gaussian (see Section 1.2). By inverting the Hopf-Cole transform we obtain

$$u_\varepsilon(\mathbf{x},t) = \varepsilon \log\left(\int_{\mathbb{R}^2} \frac{1}{4\pi\varepsilon t} e^{-\frac{|\mathbf{x}-\mathbf{y}|^2}{4\varepsilon t}} e^{\frac{1}{\varepsilon} f(\mathbf{x})}\, d\mathbf{y} \right) \tag{3.51}$$

as the solution of the perturbed problem (3.48a)–(3.48b). We also have pointwise convergence for $\varepsilon \to 0$, i.e.

$$\lim_{\varepsilon\to 0} \varepsilon \log\left(\int_{\mathbb{R}^2} \frac{1}{4\pi\varepsilon t} e^{-\frac{|\mathbf{x}-\mathbf{y}|^2}{4\varepsilon t}} e^{\frac{1}{\varepsilon} f(\mathbf{x})}\, d\mathbf{y} \right) = \sup_{\mathbf{y}\in\mathbb{R}^2} \left\{ f(\mathbf{y}) - \tfrac{1}{4t}|\mathbf{x}-\mathbf{y}|^2 \right\} \tag{3.52}$$

for every $\mathbf{x} \in \mathbb{R}^2$ [e.g. Dolcetta, 2003].

The Hopf-Cole transform gives already a first hint for a logarithmic connection between morphological and LSI scale-spaces.

3.3.2 Non-Newtonian Calculi

The Hopf-Cole transform has an interpretation in terms of non-Newtonian calculi. The non-Newtonian calculi were introduced by Grossman and Katz [1972] and play an increasing role in current image processing applications. They allow to design problem-tailored algorithms. Instead of solving a given task directly, it is solved in an alternative domain where known or desired properties can be exploited.

The non-Newtonian calculi are based on a set A together with four operations $+_\alpha, -_\alpha, \cdot_\alpha, /_\alpha$ and an ordering $<_\alpha$. The function $\alpha\colon A \to A$ is called the *generator* of the calculus and defines the operations as follows:

$$x +_\alpha y := \alpha\left(\alpha^{-1}(x) + \alpha^{-1}(y)\right), \tag{3.53a}$$

$$x -_\alpha y := \alpha\left(\alpha^{-1}(x) - \alpha^{-1}(y)\right), \tag{3.53b}$$

$$x \cdot_\alpha y := \alpha\left(\alpha^{-1}(x) \cdot \alpha^{-1}(y)\right), \tag{3.53c}$$

$$x /_\alpha y := \alpha\left(\alpha^{-1}(x) / \alpha^{-1}(y)\right) \text{ and} \tag{3.53d}$$

$$x <_\alpha y \text{ if and only if } \alpha^{-1}(x) < \alpha^{-1}(y). \tag{3.53e}$$

Since the generator α has a crucial influence on the arithmetic operations, the previous definitions are referred to as α-*arithmetics*. Also integration of a function f in the α-calculus is possible by evaluating

$$\int_\alpha f \, d\mu := \alpha\left(\int (\alpha^{-1} \circ f) \, d\mu\right). \tag{3.54}$$

Whereas the classical or Newtonian calculus is recovered when α is the identity function, other choices of α lead to the so-called *non-Newtonian calculi*. The best known example is the geometric or multiplicative calculus where α is taken to be the exponential function. It is used by Fattal et al. [2002] to reflect the human's visual perception as described by the Weber-Fechner law. This law states a logarithmic connection between the perceived colour values and their actual intensities. A more formal justification is given by Georgiev [2006]. In Florack and van Assen [2012] the geometric calculus is advocated to be used for biomedical imaging applications. Many definitions from Newtonian calculus carry over to the geometric calculus and allow to simplify derivations that are cumbersome in the classical setting [Bashirov et al., 2008].

The Hopf-Cole transformation $w \mapsto \exp(\frac{1}{\varepsilon} w)$ defined in (3.49) can be seen as an non-Newtonian calculus where we take α_n to be the family of scaled logarithms

$$\alpha_n(x) := \tfrac{1}{n} \log(x) \qquad\qquad n \in \mathbb{R} \setminus \{0\} \tag{3.55}$$

with inverse

$$\alpha_n^{-1}(x) := \exp(n\,x) \qquad\qquad n \in \mathbb{R} \setminus \{0\}\,. \tag{3.56}$$

We are interested in the limit $n \to \infty$ where $\frac{1}{n}$ plays the role of ε. These limits are closely related to mathematical morphology (see Section 1.4) since

$$x \cdot_{\alpha_n} y = \tfrac{1}{n} \log\left(\exp(n\,x) \cdot \exp(n\,y)\right)) = x + y \tag{3.57a}$$

$$\lim_{n\to\infty} x +_{\alpha_n} y = \tfrac{1}{n} \log\left(\exp(n\,x) + \exp(n\,y)\right) = \max(x, y)\,. \tag{3.57b}$$

3.3.3 Large Deviations Theory

The theory of large deviations is concerned with quantifying the rate of convergence of a stochastic system [e.g. Varadhan, 1984, Dembo and Zeitouni, 2009, Ellis, 1985]. In many cases the law of large numbers predicts convergence towards the expected value and the probability of different outcomes decays to zero exponentially. In order to specify the rate of the exponential decay, rare events are of particular interest. These are events with very small probability but significant impact if they occur.

As a first example, let us look at the empirical mean

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{3.58}$$

of a family of independent Gaussian distributed random variables $\{X_i\}_{i\in\mathbb{N}}$ with mean 0 and variance 1. Since Z_n is again Gaussian distributed, the central limit theorem tells us:

$$\lim_{n\to\infty} \Pr\left(|\sqrt{n} Z_n| < \delta\right) = \frac{1}{\sqrt{2\pi}} \int_{-\delta}^{\delta} e^{-\frac{x^2}{2}}\,\mathrm{d}x\,. \tag{3.59}$$

Therefore, we have

$$\Pr\left(|Z_n| \geqslant \delta\right) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} e^{-\frac{x^2}{2}}\,\mathrm{d}x \tag{3.60}$$

and

$$\lim_{n\to\infty} \Pr\left(|Z_n| \geqslant \delta\right) = 0\,. \tag{3.61}$$

For these kinds of estimates it is more natural to look at

$$\frac{1}{n}\log \Pr\left(|Z_n| \geqslant \delta\right) . \tag{3.62}$$

In the case of Gaussian distributed random variables this gives [e.g. Dembo and Zeitouni, 2009, Section 1.1]

$$\lim_{n\to\infty}\frac{1}{n}\log \Pr\left(|Z_n| \geqslant \delta\right) = -\frac{\delta^2}{2} . \tag{3.63}$$

The limit depends on the underlying distribution of the involved random variables. One of the main objectives of the large deviations theory is to find an exact statement for the above limit. Several results for more general distributions or families of distributions with relaxed requirements on the dependence of the random variables are available [Ellis, 1985, Dembo and Zeitouni, 2009].

Let $(\mathbf{X}_i)_{i\in\mathbb{N}}$ be a family of independent identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ and define

$$\mathbf{Z}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i . \tag{3.64}$$

We want to focus on rare events and look at

$$\frac{1}{n}\log \mathbb{E}\left[e^{nf(\mathbf{Z}_n)}\right] . \tag{3.65}$$

This term can be evaluated using a result from the large deviations theory. For $n \gg 0$, only rare events with large values contribute to (3.65).

Under appropriate conditions on the probability measure P_n of the random variable $\mathbf{Z}_n$ (see Appendix A) the following formula known as Varadhan's integral lemma [see Dembo and Zeitouni, 2009, Section 4.3] holds true

$$\lim_{n\to\infty}\frac{1}{n}\log\left[\int_{\mathbb{R}^d}\exp(nf(x))\,\mathrm{d}P_n(x)\right] = \sup_{\mathbf{y}\in\mathbb{R}^d}(f(\mathbf{y}) - I(\mathbf{y})) \tag{3.66}$$

where I is a so-called rate function that only depends on the family of measures P_n. In our setting the rate function is given as the convex conjugate of the logarithmic moment generating function [Dembo and Zeitouni, 2009, Theorem 2.2.30, Cramér's Theorem], i.e.

$$I(\mathbf{x}) := \sup_{\mathbf{t}\in\mathbb{R}^d}\{\langle\mathbf{x},\mathbf{t}\rangle - \log m(\mathbf{t})\}, \tag{3.67}$$

$$m(\mathbf{t}) := \mathbb{E}\left[\exp\left(\langle\mathbf{t},\mathbf{X}_1\rangle\right)\right] . \tag{3.68}$$

When $\mathbf{X}_1$ is Gaussian distributed with mean $\mathbf{x}$ and covariance matrix $2t\,\mathbf{I}_{2\times 2}$, (3.66) simplifies to Eq. (3.52) obtained by the the Hopf-Cole transform in Section 3.3.1:

$$\lim_{n\to\infty} \varepsilon \log \left(\int_{\mathbb{R}^2} \frac{1}{4\pi\varepsilon t} e^{-\frac{|\mathbf{x}-\mathbf{y}|^2}{4\varepsilon t}} e^{\frac{1}{\varepsilon} f(\mathbf{x})}\, \mathrm{d}\mathbf{y} \right) = \sup_{\mathbf{y}\in\mathbb{R}^2} \left\{ f(\mathbf{y}) - \tfrac{1}{4t}|\mathbf{x}-\mathbf{y}|^2 \right\} .$$

This means that Varadhan's integral lemma (3.66) generalises the ideas that lead to the Hopf-Cole transform. Instead of using Gaussian kernels more general densities can be used. For example, if we can think of a multi-dimensional Gaussian distribution with mean $\boldsymbol{\mu}$ and positive definite covariance matrix $\boldsymbol{\Sigma}$, then $\mathbf{Z}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i$ is also Gaussian distributed with mean $\boldsymbol{\mu}$ but covariance matrix $\frac{1}{n}\boldsymbol{\Sigma}$ and the rate function is given by

$$I_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = \frac{1}{2} \left\langle \mathbf{x}-\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\rangle . \tag{3.69}$$

The rate function is known for a lot more distributions, but many of them are not useful in the context of scale-space creation.

Varadhan's integral lemma offers also much more insides into why we see the term $\frac{1}{4t}|\mathbf{x}-\mathbf{y}|^2$ on the right-hand side of (3.52). It shows a direct connection between families of distributions that are used in scale-space recreation and the structuring functions employed in mathematical morphology.

The family in (3.65) is also used by Florack et al. [1999] to construct an interpolation between a morphological and a linear PDE. Varadhan's integral lemma gives an elegant alternative to their analysis. Compared to the reparametrisation argument they are using, a stochastic setting seems to be more natural.

3.3.4 Cramér Transform

Whereas the definition of the rate function I in the previous section is based on the convex conjugate of the logarithmic moment-generating function of a random variable X, a slightly different approach is common in the image processing literature.

For conservative image processing applications, images can be represented by densities. Therefore, working with densities is favoured over using random variables. Let us consider a random variable X with density p. The (double sided) *Laplace transform*

$$\mathcal{L}\left[p\right](\mathbf{x}) = \int_{\mathbb{R}^2} p(\mathbf{y}) e^{\langle \mathbf{x},\mathbf{y}\rangle}\, \mathrm{d}\mathbf{y} \tag{3.70}$$

allows an alternative definition of the moment-generating function since

$$m(\mathbf{t}) = \mathbb{E}\left[e^{\langle \mathbf{t}, \mathbf{X}_1 \rangle}\right] = \int e^{\langle \mathbf{t}, \mathbf{x} \rangle} p(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \mathcal{L}\left[p\right](\mathbf{t}) \,. \tag{3.71}$$

Based on this observation the *Cramér transform* is defined as the convex conjugate of the logarithm of the Laplace transform, i.e.

$$\mathcal{C}\left[f\right] := \left(\log \mathcal{L}\left[f\right]\right)^* . \tag{3.72}$$

In cases where f is a probability density, it can be regarded as a rate function, but the above definition is not restricted to probability densities.

Burgeth and Weickert [2005] have used the definition (3.72) as the starting point for their analysis of the logarithmic connections between linear and morphological systems.

Furthermore, the Cramér transform plays also an important role for discrete event systems. Akian et al. [1998] state results that are quite similar to some of our own findings in Section 3.6. In their language, the connections between linear and morphological scale-spaces is a duality between probability and optimization.

3.4 The Cramér-Fourier Transform

In Section 3.1.2 we have seen that linear shift-invariant scale-spaces can be obtained by convolving an initial image with an convolution kernel flattening over time. To preserve the average grey value, these kernels are normalized in such a way that integrating over them yields one. Thus, they represent probability densities. This limits the applicability of the Cramér transform since we have to ensure that the corresponding moment-generating function is finite. In other words, if the convolution kernel does not decay exponentially fast the double-sided Laplace transform evaluates to infinity.

One of the main contributions of this work is the construction of a Fourier-based Cramér transform. For image processing applications, the Fourier transform is a much more natural choice. We call our novel transformation the Cramér–Fourier transform and define it as follows:

Definition 3.4.1. *Let f be a function with a real-valued and nonnegative Fourier transform. Then its* Cramér–Fourier transform *is given by*

$$\mathcal{C}_{\mathcal{F}}\left[f\right] := \left(-\log \mathcal{F}\left[f\right]\left(\tfrac{\cdot}{2\pi}\right)\right)^* . \tag{3.73}$$

The constants are chosen in such a way that the Cramér and Cramér–Fourier transform coincide for Gaussian kernels. An in-depth comparison of both transforms is conducted in Section 3.8.
The definition fits also well to our kernels in Table 3.1: Their Fourier transforms are real-valued (due to the kernel symmetry) and positive.

First we prove that the Cramér–Fourier transform benefits from the same key property as the classical Cramér transform considered in Burgeth and Weickert [2005]: It maps convolutions to infimal convolutions.

Theorem 3.4.2 (Convolution Property of the Cramér–Fourier transform)**.** *Assume that two functions f and g are proper, lower semi-continuous, and have convex Cramér–Fourier transforms. Then the following holds true:*

$$\mathcal{C}_{\mathcal{F}}[f * g] \;=\; \mathcal{C}_{\mathcal{F}}[f] \,\square\, \mathcal{C}_{\mathcal{F}}[g]\,. \tag{3.74}$$

Proof. Our proof uses several results from convex analysis (see e.g. Rockafellar [1970]). Since $\mathcal{C}_{\mathcal{F}}[f]$ and $\mathcal{C}_{\mathcal{F}}[g]$ are lower semi-continuous, proper and convex, also their convex conjugates $\mathcal{C}_{\mathcal{F}}[f]^*$ and $\mathcal{C}_{\mathcal{F}}[g]^*$ share these properties. Moreover, it follows that $\mathcal{C}_{\mathcal{F}}[f]^* = -\log \mathcal{F}[f]$. Therefore, a direct computation gives

$$\begin{aligned}
&\mathcal{C}_{\mathcal{F}}[f * g] = \left(-\log \mathcal{F}[f * g]\left(\tfrac{\cdot}{2\pi}\right)\right)^* \\
&= \left((-\log \mathcal{F}[f]\left(\tfrac{\cdot}{2\pi}\right)) + (-\log \mathcal{F}[g]\left(\tfrac{\cdot}{2\pi}\right))\right)^* \\
&= (-\log \mathcal{F}[f]\left(\tfrac{\cdot}{2\pi}\right))^* \,\square\, (-\log \mathcal{F}[g]\left(\tfrac{\cdot}{2\pi}\right))^* \\
&= \mathcal{C}_{\mathcal{F}}[f] \,\square\, \mathcal{C}_{\mathcal{F}}[g]
\end{aligned} \tag{3.75}$$

where we have applied the convolution theorem

$$\mathcal{F}[f * g] = \mathcal{F}[f]\;\mathcal{F}[g] \tag{3.76}$$

for the Fourier transform and the well known property

$$(f + g)^* = f^* \,\square\, g^* \tag{3.77}$$

of the convex conjugate. □

Taking a *delta peak*

$$\delta(\mathbf{x}) = \begin{cases} 1 & \text{for } \mathbf{x} = 0 \\ 0 & \text{else} \end{cases} \tag{3.78}$$

as f, and the convolution kernel k of an LSI space-space as g, Theorem 3.4.2 shows that we obtain a morphological scale-space with structuring function $\mathcal{C}_{\mathcal{F}}[k]$ and *morphological delta peak*

$$\chi(\mathbf{x}) = \begin{cases} 0 & \text{for } \mathbf{x} = 0 \\ \infty & \text{else} \end{cases} \tag{3.79}$$

as initial value. The morphological delta peak is the neutral element for infimal convolutions in the same way as a delta peak is the neutral element for convolutions.

For applications in image processing, we use the convolution kernels and these obtained structuring functions to create linear and morphological scale-spaces with the same initial image f.

Definition 3.4.3. *For some LSI scale-space*

$$u(\mathbf{x}, t) = (f * k(\cdot, t))(\mathbf{x}) \tag{3.80}$$

with convolution kernel

$$k(\mathbf{x}, t) \;=\; \mathcal{F}^{-1}[\exp(p(2\pi\cdot)\, t)](\mathbf{x}) \tag{3.81}$$

and initial image f, the structuring function $s(\mathbf{x}, t)$ for a corresponding morphological scale-space

$$v(\mathbf{x}, t) = (f \,\square\, s(\cdot, t))(\mathbf{x}) \tag{3.82}$$

is obtained by applying the Cramér–Fourier transform to $k(\mathbf{x}, t)$:

$$s(\mathbf{x}, t) \;=\; \mathcal{C}_{\mathcal{F}}[k(\cdot, t)](\mathbf{x}) \;=\; (-t\, p(\cdot))^{*}(\mathbf{x}). \tag{3.83}$$

3.5 Connections between Linear and Morphological Evolutions

Interestingly, we can also use the Cramér–Fourier transform on the level of evolution equations. This enables us to prove the following theorem which constitutes one of our main results.

Theorem 3.5.1. *Let $u(\mathbf{x}, t)$ be the solution of the LSI scale-space evolution*

$$\partial_t u(\mathbf{x}, t) = p(\boldsymbol{\nabla})\, u(\mathbf{x}, t) \qquad \text{on } \mathbb{R}^2 \times (0, \infty) \tag{3.84a}$$

$$u(\mathbf{x}, 0) = \delta(\mathbf{x}) \qquad \text{on } \mathbb{R}^2 \tag{3.84b}$$

where $p(\xi)$ denotes the symbol of the pseudodifferential operator $P(\nabla)$ with constant coefficients.

If p is proper, lower semi-continuous and convex, the Cramér–Fourier transform of u, denoted by v, is the unique viscosity solution of the morphological scale-space evolution

$$\partial_t v(\mathbf{x}, t) = p\left(\nabla v(\mathbf{x}, t)\right) \qquad \text{on } \mathbb{R}^2 \times (0, \infty) \tag{3.85a}$$

$$v(\mathbf{x}, 0) = \chi(\mathbf{x}) \qquad \text{on } \mathbb{R}^2 . \tag{3.85b}$$

A similar theorem using the Cramér transform is proved by Akian et al. [1994]. While their proof could be modified to our setting, using the previous obtained results is much simpler.

Proof. The discussion in Section 3.1.2 shows that the solution of (3.84a)–(3.84b) is given by

$$u(\mathbf{x}, t) = (\delta * k(\cdot, t))(\mathbf{x}) \tag{3.86}$$

with $k(\mathbf{x}, t) = \mathcal{F}^{-1}[\exp(p(2\pi\cdot)t)\,](\mathbf{x})$. Since p is proper, lower semi-continuous and convex, applying Theorem 3.4.2 shows that $v(\mathbf{x}, t) = \mathcal{C}_{\mathcal{F}}\,[\,u(\cdot, t)\,]\,(\mathbf{x})$ is given by

$$v(\mathbf{x}, t) = (\chi \,\square\, s(\cdot, t))(\mathbf{x}) \tag{3.87a}$$

with

$$s(\mathbf{x}, t) = (-t\, p(\cdot))^*(\mathbf{x}) . \tag{3.87b}$$

Following the discussion in Section 3.2 with $H = -p$, this is the Hopf-Lax formula for the unique viscosity solution of (3.85a)–(3.85b).

□

This motivates us to introduce the following definition that is the counterpart of Definition 3.4.3 in terms of evolution equations:

Definition 3.5.2. *Let an LSI scale-space evolution be given by*

$$\partial_t u = P(\nabla) u , \tag{3.88a}$$

$$u(., 0) = f . \tag{3.88b}$$

Then its corresponding morphological scale-space *evolution satisfies*

$$\partial_t v = p\left(\nabla v\right) , \tag{3.89a}$$

$$v(\mathbf{x}, 0) = f(\mathbf{x}) . \tag{3.89b}$$

Note that for computing the corresponding morphological scale-space, only the symbol p of the LSI scale-space is required. In particular, no closed-form kernel representation is necessary.

Figure 3.1 summarises our theoretical findings. We observe that we have obtained a simple dictionary that allows to translate results between linear and morphological scale-spaces, both in terms of convolutions / infimal convolutions and evolution equations.

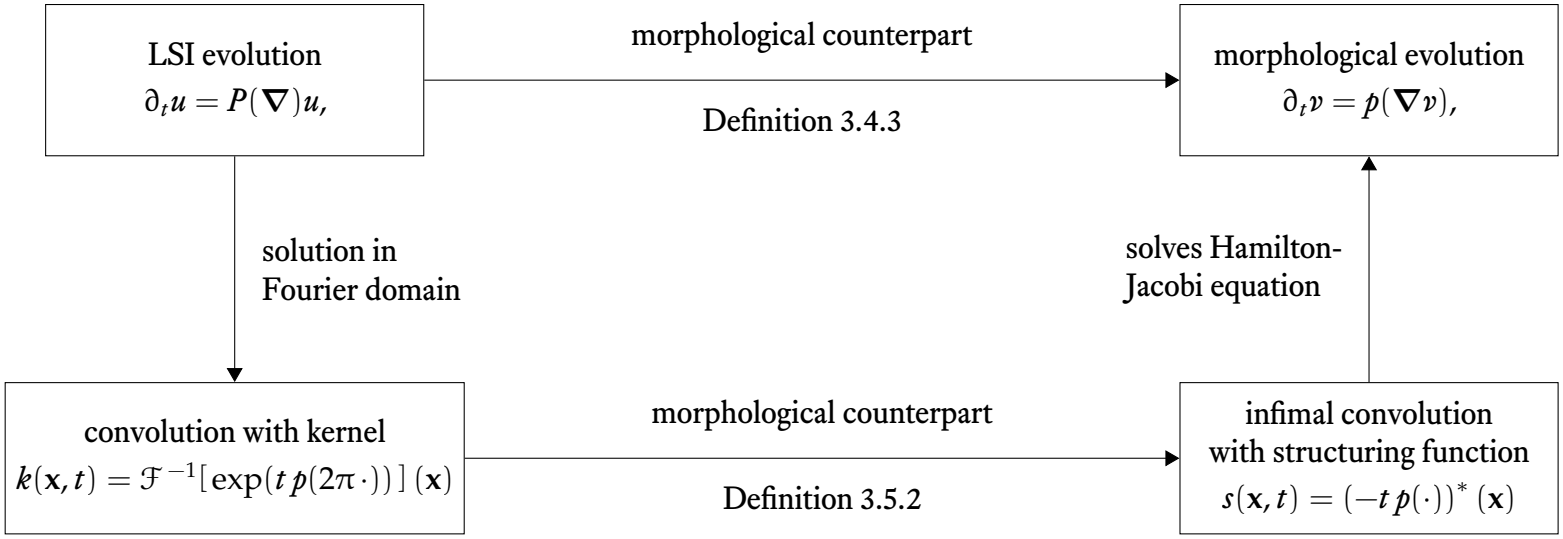


Figure 3.1: General dictionary that allows to translate results between LSI scale-spaces and morphological scale-spaces, both on the level of evolution equations (top) and the level of convolutions / infimal convolutions (bottom).

3.6 Application to Specific Scale-Spaces

Now we are in a position to apply our theory to a number of linear scale-spaces in order to derive their morphological counterparts.

1. **Gaussian Scale-Space.** Table 3.1 specifies the symbol of Gaussian scale-space as
$$p(\xi) = -|\xi|^2 . \tag{3.90}$$
According to Definition 3.5.2, its morphological counterpart is given by
$$\partial_t v = -|\nabla v|^2 , \tag{3.91}$$
which coincides with van den Boomgaard's result [van den Boomgaard, 1992b]. According to our framework, the corresponding structuring function can be computed as
$$s(\mathbf{x}) = \left(-p\left(\cdot\right)\right)^* \left(\mathbf{x}\right) = \tfrac{1}{4}|\mathbf{x}|^2, \tag{3.92}$$
which again confirms van den Boomgaard's result. This shows that our framework reproduces the only connection between linear and morphological scale-spaces that is known so far. Thus, we can focus now on establishing novel connections.

2. **α-Scale-Spaces.** In the same way as above, one can show that the morphological equivalents for the α-scale-spaces are given by
$$\partial_t v = -|\nabla v|^{2\alpha}. \tag{3.93}$$
We observe that this is exactly the class of morphological evolutions that are studied by Jackway [1994] and Diop and Angulo [2014].

   Interestingly, (3.93) also proves that for $\alpha = \frac{1}{2}$, the linear counterpart of the widely-used morphological scale-space
$$\partial_t v = -|\nabla v|, \tag{3.94}$$
which describe erosion with a flat disc of radius t, is given by the Poisson scale-space
$$\partial_t u = -\sqrt{-\Delta}\, u . \tag{3.95}$$
To our knowledge, this connection has not been stated before.

As a didactic example, let us now confirm that our computations also reproduce the structuring functions of Diop and Angulo [2014]. Knowing the symbol $p(\mathbf{x}) = -|\mathbf{x}|^{2\alpha}$, we can use (3.26) again to compute s_α :

$$\begin{aligned} s_\alpha(\mathbf{x}) &= \left(-p\left(\cdot\right)\right)^*(\mathbf{x}) = \left(|\cdot|^{2\alpha}\right)^*(\mathbf{x}) \\ &= (2\alpha - 1)\left|\frac{\mathbf{x}}{2\alpha}\right|^{\frac{2\alpha}{2\alpha-1}}, \end{aligned} \tag{3.96}$$

since [see e.g Rockafellar, 1970, p. 106]

$$\left(\tfrac{1}{b}|\cdot|^b\right)^*(\mathbf{x}) = \tfrac{b-1}{b}|\mathbf{x}|^{\frac{b}{b-1}} \qquad \text{for } b > 1\,. \tag{3.97}$$

This coincides with the result from Diop and Angulo [2014] stated in (3.38). Although this formula only holds for $\alpha > \frac{1}{2}$, we can compute the pointwise limit

$$\lim_{\alpha \to \frac{1}{2}^+} s_\alpha(\mathbf{x}) = \begin{cases} 0 & |\mathbf{x}| \leqslant 1, \\ \infty & \text{else.} \end{cases} \tag{3.98}$$

As expected, this is a flat disc of radius 1.

3. **Summed α-Scale-Spaces.** We know that summed α-scale-spaces have the symbol

$$p(\boldsymbol{\xi}) = -\sum_{k=1}^{m} \lambda_k |\boldsymbol{\xi}|^{2\alpha_k}\,. \tag{3.99}$$

This yields

$$\partial_t u = -\sum_{k=1}^{m} \lambda_k |\boldsymbol{\nabla} u|^{2\alpha_k} \tag{3.100}$$

as morphological counterpart of

$$\partial_t u = -\sum_{k=1}^{m} \lambda_k (-\Delta)^{\alpha_k} u\,. \tag{3.101}$$

In a similar way as before, its structuring function can be derived as

$$s(\mathbf{x}) = \square_{k=1}^{m} \lambda_k (2\alpha_k - 1)\left|\frac{\mathbf{x}}{2\alpha_k \lambda_k}\right|^{\frac{2\alpha_k}{2\alpha_k - 1}}. \tag{3.102}$$

4. **Relativistic Scale-Spaces.** From Table 3.1 we see that relativistic scale-spaces are characterised by the symbol

$$p(\xi) = m - \sqrt{|\xi|^2 + m^2}\,. \tag{3.103}$$

This gives

$$\partial_t v = m - \sqrt{|\nabla v|^2 + m^2} \tag{3.104}$$

as morphological counterparts. The structuring function $s_{r,m}$ can be computed as before as the convex conjugate of the negative symbol:

$$s_{r,m}(\mathbf{x}) = \left(\sqrt{|\cdot|^2 + m^2} - m\right)^* (\mathbf{x}) \tag{3.105}$$

$$= \sup_{\mathbf{y}\in\mathbb{R}^2} \left(\langle \mathbf{x}, \mathbf{y}\rangle + m - \sqrt{|y|^2 + m^2}\right). \tag{3.106}$$

If $|\mathbf{x}| \leqslant 1$, the solution for $\mathbf{y}$ is given by

$$\mathbf{y} = \frac{\mathbf{x}\, m}{1 - |\mathbf{x}|^2}\,. \tag{3.107}$$

Thus, it follows that

$$s_{r,m}(\mathbf{x}) = \begin{cases} m\left(1 - \sqrt{1 - |\mathbf{x}|^2}\right) & |\mathbf{x}| \leqslant 1, \\ \infty & \text{else.} \end{cases} \tag{3.108}$$

For $m \to 0$, the structuring function $s_{r,m}$ converges to a flat disc of radius 1. This is expected from the results from the last section, since the relativistic scale-spaces converge to the Poisson scale-space for $m \to 0$.

5. **Anisotropic Scale-Spaces.** The symbol for anisotropic Gaussian scale-space is

$$p(\xi) = -\langle \xi, \mathbf{D}\,\xi\rangle\,. \tag{3.109}$$

This allows to compute the morphological counterpart of

$$\partial_t u = \operatorname{div}(\mathbf{D}\,\nabla u) \tag{3.110}$$

as

$$\partial_t v = -\langle \nabla v, \mathbf{D}\,\nabla v\rangle\,. \tag{3.111}$$

As already mentioned, this morphological evolution has been studied by van den Boomgaard [1992a] and by Jackway [1995].

So far we have always started with LSI scale-spaces and derived their corresponding morphological scale-space. The only morphological evolution that we could not derive in this way was the anisotropic differential equation of Arehart et al. [1993]:

$$\partial_t v \;=\; -|\mathbf{D}\boldsymbol{\nabla} v| \;=\; -\sqrt{\boldsymbol{\nabla} v^\top \mathbf{D}^2\, \boldsymbol{\nabla} v}\,. \tag{3.112}$$

This is a good opportunity to show that our theoretical framework provides us with a dictionary that can be used also in the reverse direction. Obviously (3.112) can be expressed as

$$\partial_t v = p(\boldsymbol{\nabla} v) \tag{3.113}$$

with symbol

$$p(\xi) = -\sqrt{\xi^\top \mathbf{D}^2\, \xi}\,. \tag{3.114}$$

This gives rise to an anisotropic Poisson scale-space

$$\partial_t u = -\sqrt{-\boldsymbol{\nabla}^\top \mathbf{D}^2\, \boldsymbol{\nabla}}\, u \tag{3.115}$$

that has not been described in the literature before.

Table 3.3 summarises the results of this Section 3.5. We observe that we have derived many correspondences between known LSI scale-spaces and morphological ones. Moreover, we have also managed to come up with novel scale-spaces.

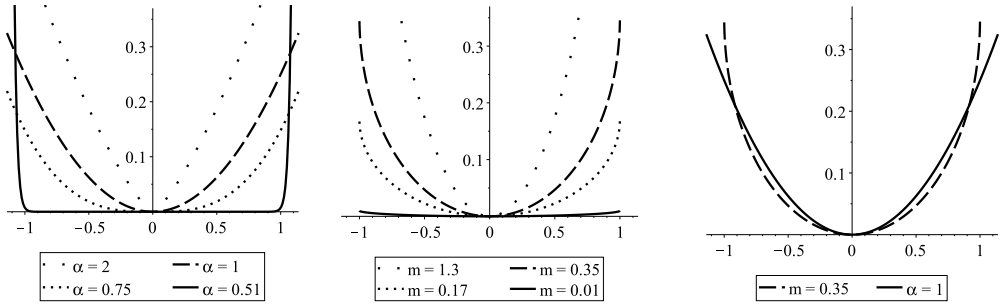
Table 3.3: Specific LSI scale-spaces and their morphological equivalents.

LSI scale-space	LSI evolution	morphological evolution	morphological scale-space
Gaussian	$\partial_t u = \Delta u$	$\partial_t v = -\lvert\boldsymbol{\nabla} v\rvert^2$	quadratic SF
Poisson	$\partial_t u = -\sqrt{-\Delta}\, u$	$\partial_t v = -\lvert\boldsymbol{\nabla} v\rvert$	flat disk SF
alpha	$\partial_t u = -(-\Delta)^{\alpha} u$	$\partial_t v = -\lvert\boldsymbol{\nabla} v\rvert^{2\alpha}$	poweroid SF
summed alpha	$\partial_t u = -\sum_{k=1}^{m} \lambda_k (-\Delta)^{\alpha_k} u$	$\partial_t v = -\sum_{k=1}^{m} \lambda_k \lvert\boldsymbol{\nabla} v\rvert^{2\alpha_k}$	morphological summed alpha
relativistic	$\partial_t u = \left(m - \sqrt{m^2 - \Delta}\right) u$	$\partial_t v = m - \sqrt{m^2 + \lvert\boldsymbol{\nabla} v\rvert^2}$	morphological relativistic
anisotropic Gaussian	$\partial_t u = \operatorname{div}(\mathbf{D}\,\boldsymbol{\nabla} u)$	$\partial_t v = -\boldsymbol{\nabla}^{\top} v\, \mathbf{D}\, \boldsymbol{\nabla} v$	anisotropic quadratic SF
anisotropic Poisson	$\partial_t u = -\sqrt{-\boldsymbol{\nabla}^{\top} \mathbf{D}^2\, \boldsymbol{\nabla}}\, u$	$\partial_t v = -\lvert\mathbf{D}\boldsymbol{\nabla} v\rvert$	flat ellipse-shaped SF

3.7 Experiments

Although our results are of theoretical nature, we would like to illustrate some of the discussed scale-spaces and their correspondences by experiments. The implementation for the linear scale-spaces uses a multiplication in the Fourier domain. For the morphological scale-spaces, we compute the infimal convolution over the image domain.

In Figure 3.2, we plot structuring functions of the morphological counterparts of various alpha- and relativistic scale-spaces. First of all, we observe the convexity of all structuring functions. Fig. 3.2(a) shows that for $\alpha \to 0.5$, the structuring function of the morphological α-scale-space converges to a flat structuring function. A similar behaviour can be observed for morphological relativistic scale-spaces when $m \to 0$; see Fig. 3.2(b). On the other hand, Fig. 3.2(c) shows that for $m = 0.4$, the morphological relativistic scale-space gives a good approximation to the morphological equivalent of Gaussian scale-space.



(a) morphological α-scale-space

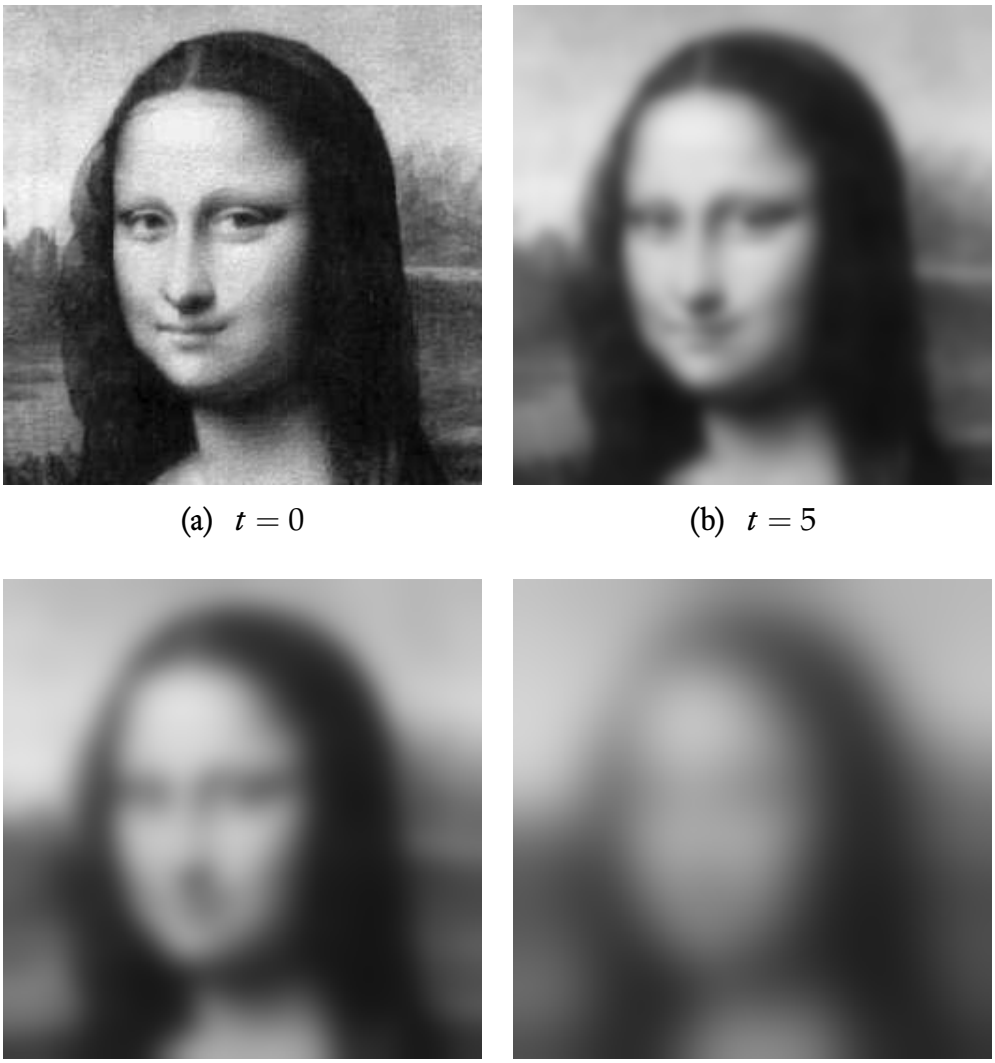
(b) morphological relativistic scale-space

(c) morphological Gaussian and relativistic scale-space

Figure 3.2: Structuring functions for one-dimensional morphological scale-spaces.

Whereas Figure 3.3 shows the two-dimensional linear α-scale-space for $\alpha = 0.75$, Figures 3.4 shows its morphological infimal convolution counterpart which corresponds to an erosion process. To enable comparisons, we have chosen the same Mona Lisa image as in Burgeth and Weickert [2005]. Moreover, for the sake of completeness, we also depict the corresponding dilation scale-space in Figure 3.5.

Figures 3.6 and 3.7 compare the linear and morphological relativistic scale-space for $m = 0.1$. Since this m value is fairly close to the limit $m \to 0$, the linear evolution resembles Poisson scale-space, and its morphological counterpart approximates erosion with disc-shaped structuring functions. The latter is well visible.



(a) $t = 0$ (b) $t = 5$

(c) $t = 15$ (d) $t = 50$

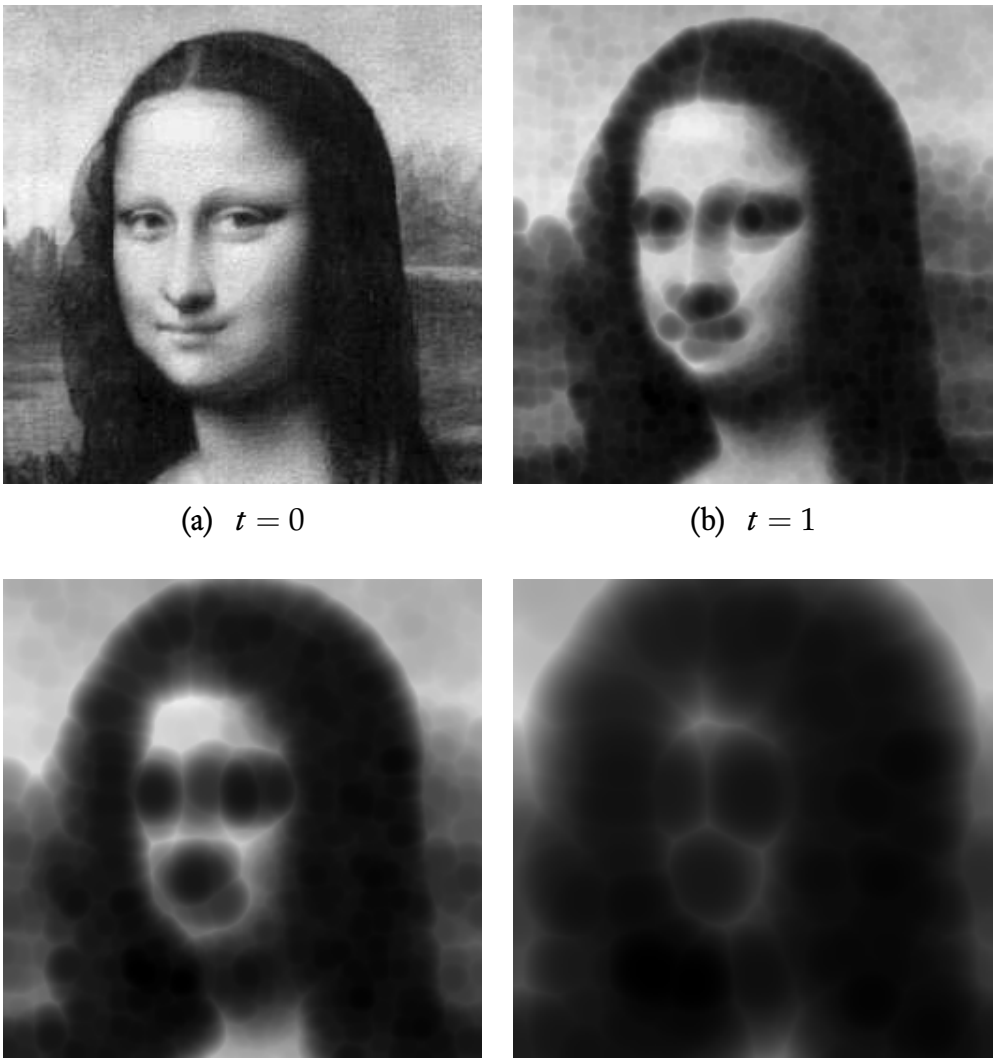
Figure 3.3: Linear α-scale-space with $\alpha = 0.75$.

The last example in Figure 3.8 shows that the anisotropy of the convolution kernel carries over to the structuring function. For this experiment we take the matrix

$$\mathbf{D} = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix} \tag{3.116}$$

and compare the convolution kernel of the anisotropic Poisson scale-space to its corresponding flat, ellipse-shaped structuring function (3.39) with the inverse matrix

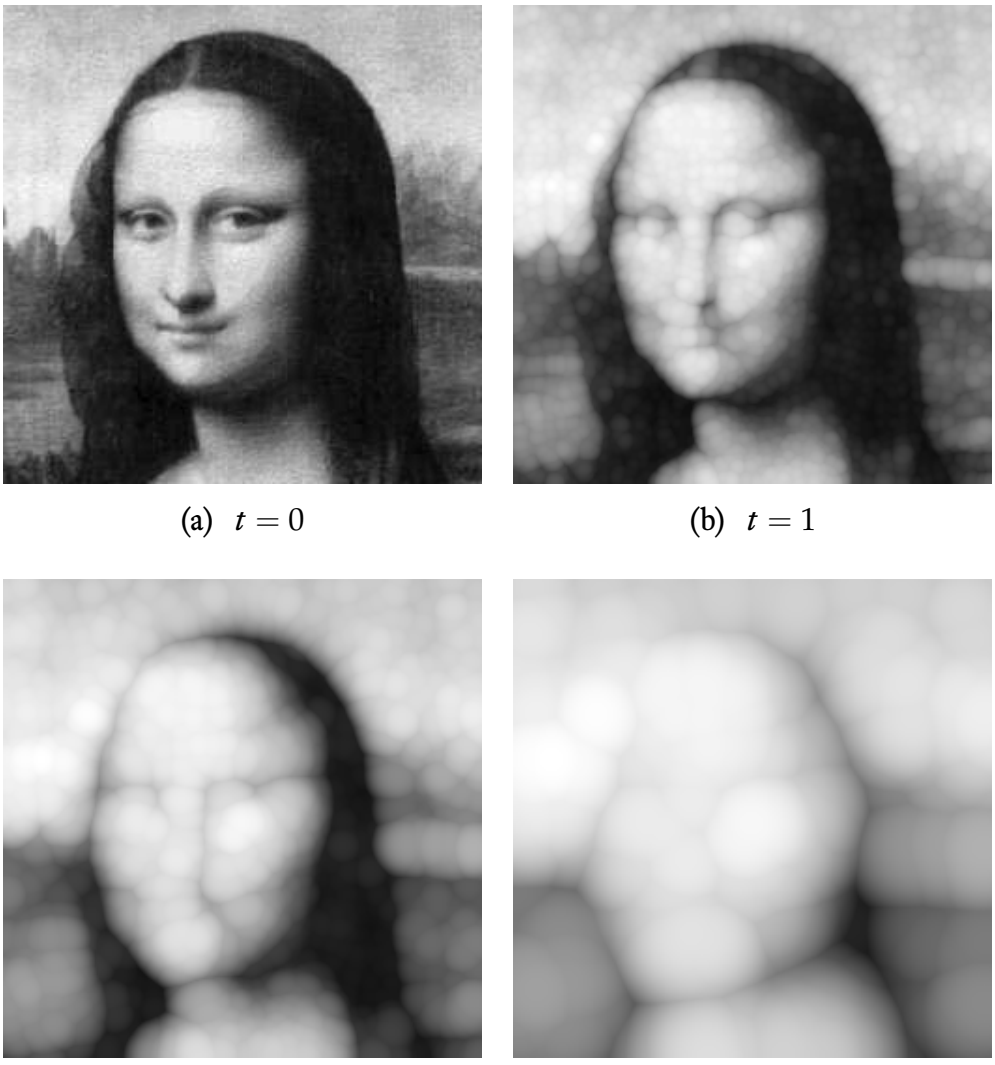
$$\mathbf{D}^{-1} = \frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}. \tag{3.117}$$



(a) $t = 0$ (b) $t = 1$

(c) $t = 3$ (d) $t = 10$

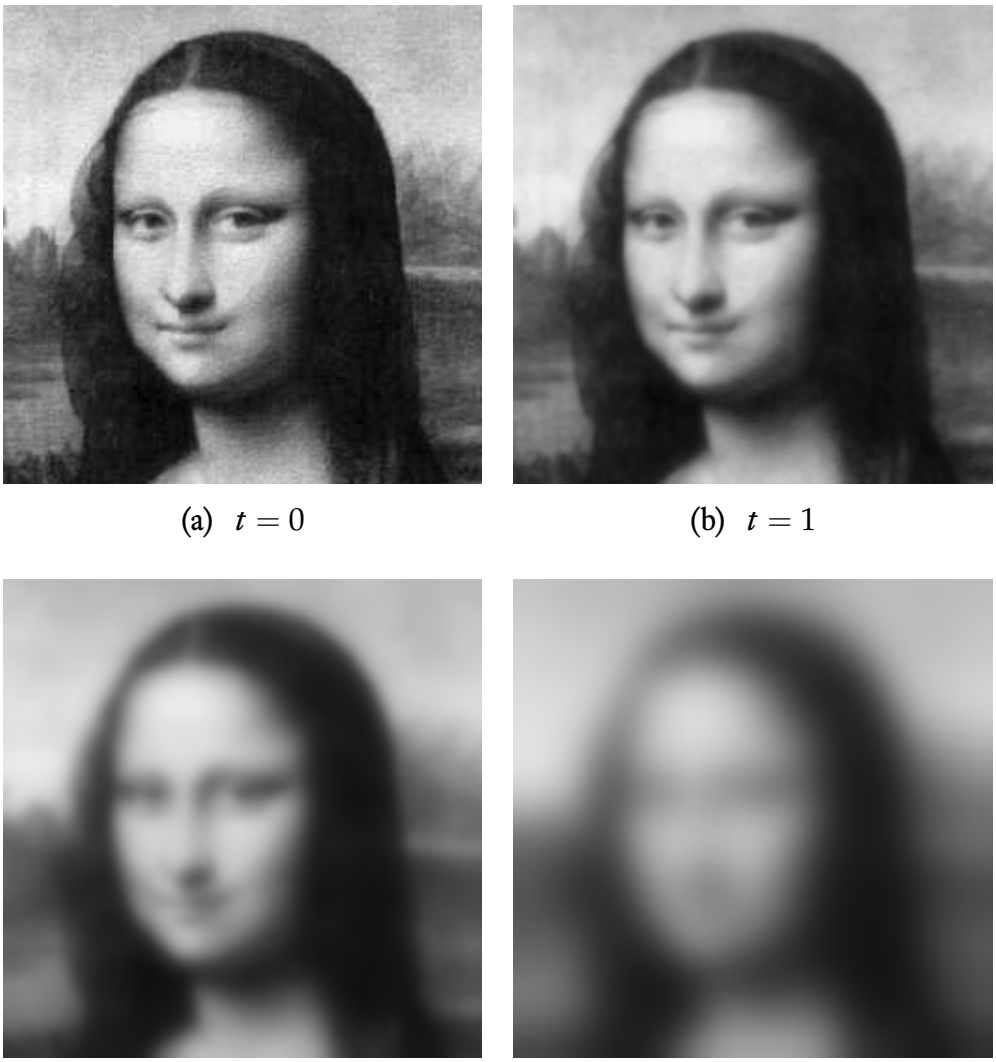
Figure 3.4: Morphological α-scale-space with $\alpha = 0.75$.



(a) $t = 0$ (b) $t = 1$

(c) $t = 3$ (d) $t = 10$

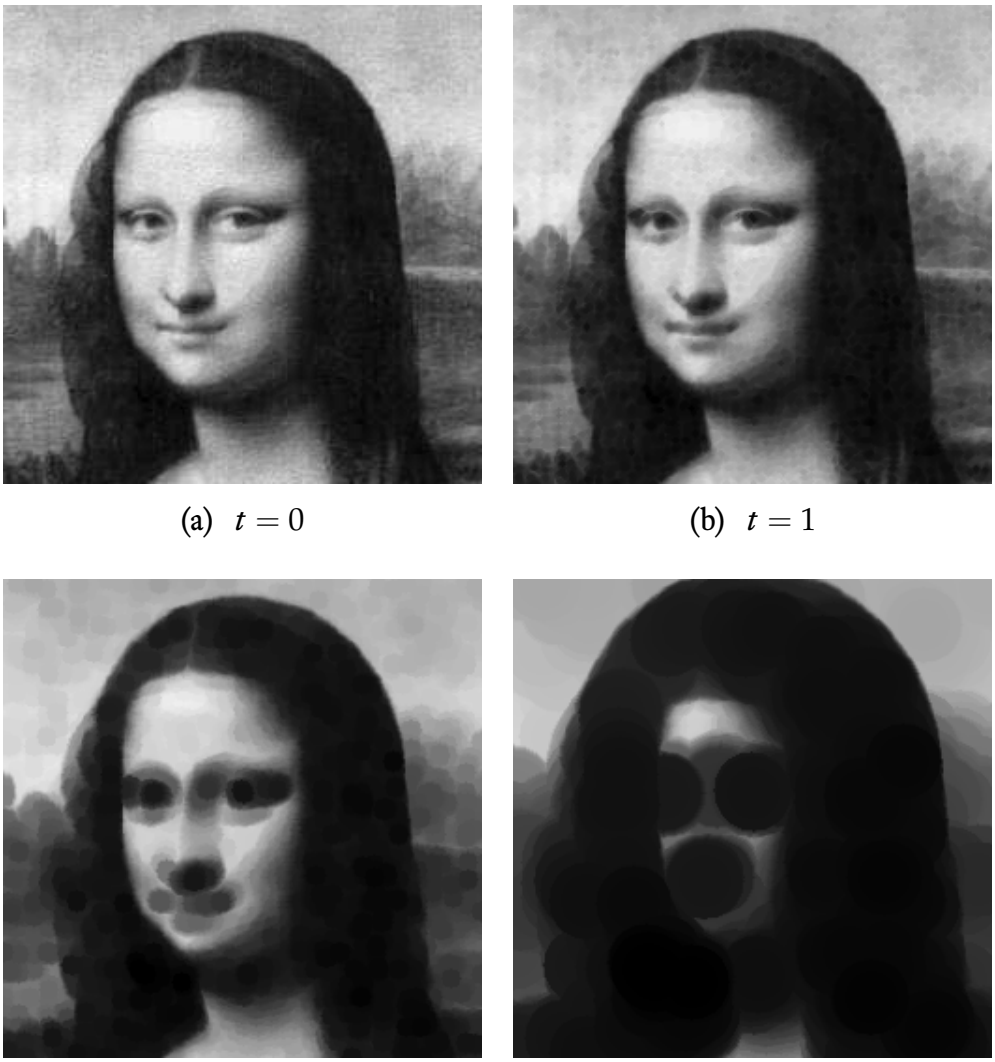
Figure 3.5: Corresponding dilation scale-space.



(a) $t = 0$ (b) $t = 1$

(c) $t = 5$ (d) $t = 20$

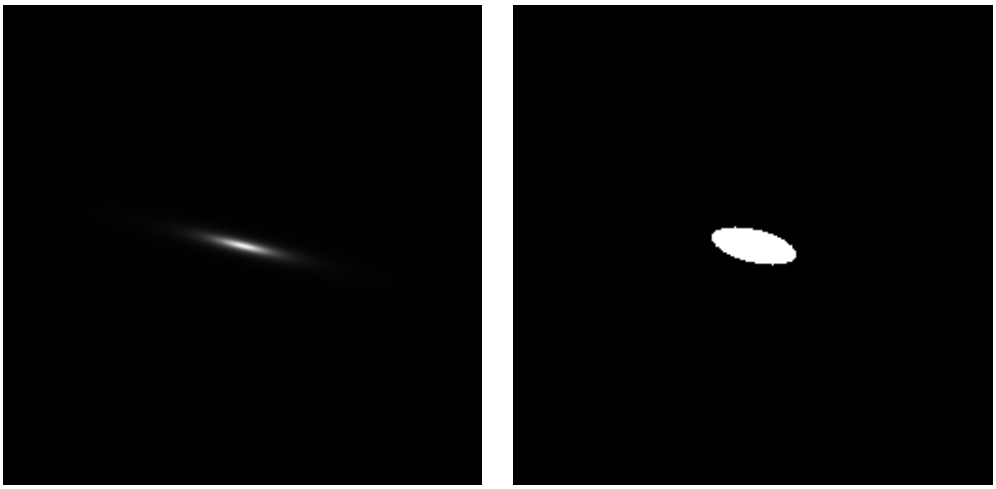
Figure 3.6: Linear relativistic scale-space with $m = 0.1$.



(a) $t = 0$ (b) $t = 1$

(c) $t = 5$ (d) $t = 20$

Figure 3.7: Morphological relativistic scale-space with $m = 0.1$.



(a) $t = 3$ (b) $t = 10$

Figure 3.8: **Left:** Convolution kernel for linear anisotropic Poisson scale-space. **Right:** Corresponding structuring function.

3.8 Comparison to the Cramér Transform

Most applications of the large deviations theory rely on the interpretation that the rate function describes the tail probability of a distribution. However, this is different from the use case in Section 3.3.3 We are more interested is the structural change that takes place by replacing the classical algebra by the max-plus-algebra $\mathbb{R}_{\max}$. This transition is made explicit by the Cramér transform which needs the moment generating function (Laplace transform of the density) to be finite. In probability theory, the preferred way of describing a random variable X (with density p) is the characteristic function defined by

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}\left[e^{\mathrm{i}\langle \mathbf{t},\mathbf{X}\rangle}\right] = \int_{\mathbb{R}^d} p(\mathbf{x})\, e^{\mathrm{i}\langle \mathbf{t},\mathbf{x}\rangle}\, \mathrm{d}\mathbf{x}\,. \tag{3.118}$$

Its main advantage compared to the moment generating function

$$m_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}\left[e^{\langle \mathbf{t},\mathbf{X}\rangle}\right] \tag{3.119}$$

is the fact that the characteristic function always exists, for every distribution. By definition, one expects similar properties of both characteristic and moment generating function.

We used this observation in Section 3.4 to construct the Cramér–Fourier transform. The definition can be rewritten in term of the characteristic function as follows:

$$\mathcal{C}_{\mathcal{F}}[p](\mathbf{x}) = \left(-\log\left(\mathcal{F}[p]\left(\tfrac{\cdot}{2\pi}\right)\right)\right)^*(\mathbf{x}) \tag{3.120a}$$
$$= (-\log \varphi_{\mathbf{X}})^*(\mathbf{x}) \tag{3.120b}$$

with p being the density of $\mathbf{X}$. Although the Cramér and the Cramér–Fourier transform do not coincide, we can show the following lemma.

Lemma 3.8.1. *Let X be a random variable with symmetric probability density function and finite moment generating function. Then, the leading terms of the Taylor expansions of $\log \mathcal{L}[X] = \log m_X$ and $-\log \mathcal{F}[X]\left(\frac{\cdot}{2\pi}\right) = -\log \varphi_X$ are equal.*

The preconditions are chosen in such a way to make sure that both transforms exist.

Proof. Since all odd moments of X vanish, we have

$$m_X(t) = \sum_{i=1}^{\infty} \frac{t^n}{n!}\mathbb{E}[X^n] \tag{3.121a}$$
$$= 1 + \frac{1}{2}t^2\left(\int_{\mathbb{R}} p(x)\, x^2 \,\mathrm{d}x\right) + \mathcal{O}(t^4). \tag{3.121b}$$

For the characteristic function we obtain

$$\varphi_X(t) = \int_{\mathbb{R}} p(x) e^{itx} \,\mathrm{d}x \tag{3.122a}$$
$$= \int_{\mathbb{R}} p(x)(\cos(tx) + i\,\sin(tx)) \,\mathrm{d}x \tag{3.122b}$$
$$= \int_{\mathbb{R}} p(x)\cos(tx) \,\mathrm{d}x \qquad \text{since } p \text{ is even} \tag{3.122c}$$
$$= \int_{\mathbb{R}} \left(p(x)\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(tx)^{2n}\right) \mathrm{d}x \tag{3.122d}$$
$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, t^{2n} \left(\int_{\mathbb{R}} p(x)\, x^{2n} \,\mathrm{d}x\right) \tag{3.122e}$$

$$= 1 - \frac{1}{2}t^2 \left(\int_{\mathbb{R}} p(x)\, x^2 \, \mathrm{d}x \right) + \mathcal{O}(t^4)\,. \tag{3.122f}$$

The integral and the sum can be exchanged by Fubini's theorem because the sum can be interpreted as an integral with a discrete measure and we know that the integral of the absolute value of the integrand is finite. Keep in mind that we use the existence of the moment generating function of X at this point. The lemma follows by observing

$$\log(1+z) = z - \frac{z^2}{2} + \mathcal{O}(z^3). \tag{3.123}$$

□

Although we cannot prove it at this point, we also expect the following to hold true.

Conjecture 3.8.2. *Let X be a random variable with density p. The leading terms of the Taylor expansions of the Cramér transform*

$$\mathcal{C}\,[p\,] = (\log m_X)^* \tag{3.124}$$

and the Cramér–Fourier transform

$$\mathcal{C}_{\mathcal{F}}\,[p\,] = (-\log \varphi_X)^* \tag{3.125}$$

coincide if both transforms exist.

Let us look at three examples to understand this conjecture in more details:

- **Gaussian distribution.** For a Gaussian distributed random variable X with mean 0 we have

  $$-\log \varphi_X \equiv \log m_X. \tag{3.126}$$

  In this case, even

  $$\mathcal{C}\,[p\,] \equiv \mathcal{C}_{\mathcal{F}}\,[p\,]\,. \tag{3.127}$$

  holds true. This is expected, since the Cramér–Fourier transform is designed to coincide with the Cramér transform for Gaussian kernels.

- **Uniform distribution.** Take a continuous random variable X that is uniformly distributed on the interval $[a,b] = [-1,1]$. Then it is known:

$$m_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)} = \frac{\sinh(t)}{t} = 1 + \frac{t^2}{3!} + \frac{t^4}{5!} + \mathcal{O}(t^6) \tag{3.128a}$$

$$\varphi_X(t) = \frac{e^{\mathrm{i}bt} - e^{\mathrm{i}at}}{\mathrm{i}t(b-a)} = \frac{\sin(t)}{t} = 1 - \frac{t^2}{3!} + \frac{t^4}{5!} + \mathcal{O}(t^6) \tag{3.128b}$$

and we have

$$\log(m_X(t)) = \frac{t^2}{6} - \frac{t^4}{180} + \mathcal{O}(t^6) \tag{3.129a}$$

$$-\log(\varphi_X(t)) = \frac{t^2}{6} + \frac{t^4}{180} + \mathcal{O}(t^6). \tag{3.129b}$$

This last example shows that the Cramér and Cramér–Fourier transform can be different.

- **Discrete Distributions.** If we assume that the conjecture holds true, the following extension to discrete filters would be possible.

Let $\mathbf{X}$ be a discrete random variable that takes the value $\mathbf{x}_i \in \mathbb{R}^d$ with probability p_i with $1 \leqslant i \leqslant n$ and $\sum_{i=1}^n p_i = 1$. The moment generating function of $\mathbf{X}$ is given by

$$m_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}\left[e^{\langle \mathbf{t}, \mathbf{X}\rangle}\right] = \sum_{i=1}^{n} e^{\langle \mathbf{t}, \mathbf{x}_i\rangle} p_i \tag{3.130}$$

and its Cramér transform is

$$\mathcal{C}\left[\mathbf{X}\right](\mathbf{t}) = \left(\log m_{\mathbf{X}}\right)^*(\mathbf{t}) \tag{3.131a}$$

$$= \sup_{\mathbf{s}\in\mathbb{R}^d} \left\{ \underbrace{\langle \mathbf{s}, \mathbf{t}\rangle - \log\left(\sum_{i=1}^{n} e^{\langle \mathbf{s}, \mathbf{x}_i\rangle} p_i\right)}_{h(\mathbf{s})} \right\} \tag{3.131b}$$

To compute it explicitly we need the gradient of h

$$\partial_{s_j} h(\mathbf{s}) = t_j - \frac{\sum_{i=1}^n x_{i,j}\, e^{\langle \mathbf{s}, \mathbf{x}_i\rangle} p_i}{\sum_{i=1}^n e^{\langle \mathbf{s}, \mathbf{x}_i\rangle} p_i} \qquad j = 1, \ldots, d \tag{3.132}$$

to be zero. The notation $x_{i,j}$ is used to denote the jth component of $\mathbf{x}_i$. If we assume $\mathbf{s} = \log(\mathbf{k})$ (defined componentwise) we obtain

$$\sum_{i=1}^{n} p_i \left(x_{i,j} - t_j\right) \mathbf{k}^{\mathbf{x}_i} \stackrel{!}{=} 0 \qquad k > 0, j = 1, \ldots, d\,. \tag{3.133}$$

Thus, we need to solve a system of polynomial equations to evaluate the Cramér transform of a general discrete filter at a given argument $\mathbf{t}$.

In special cases, however, this is much easier. Let us consider the discrete linear diffusion equation given by the stencil k

$$\begin{array}{|c|c|c|} \hline \frac{\tau}{h^2} & 1-2\frac{\tau}{h^2} & \frac{\tau}{h^2} \\ \hline \end{array} \tag{3.134}$$

with grid size $h > 0$ and time step size $\tau > 0$.

For every grid point j this kernel defines a probability distribution $\mathbf{X}_j$ by:

$$\Pr(\mathbf{X}_j = i) \begin{cases} \frac{\tau}{h^2} & \text{for } i = j \pm 1 \\ 1-2\frac{\tau}{h^2} & \text{for } i = j \\ 0 & \text{else} \end{cases} . \tag{3.135}$$

The convolution of k with a discrete signal f at point j can be expressed as expectation of $\mathbf{X}_j$ in the following way:

$$(k * f)_j = \mathbb{E}[f(\mathbf{X}_j)] = \frac{\tau}{h^2} f_{j-1} + \left(1-2\frac{\tau}{h^2}\right) f_j + \frac{\tau}{h^2} f_{j+1} . \tag{3.136}$$

The moment generating function is

$$m_{\mathbf{X}_j}(t) = \frac{\tau}{h^2}\, e^{(j-1)\,t} + \left(1-2\frac{\tau}{h^2}\right)\, e^{j\,t} + \frac{\tau}{h^2}\, e^{(j+1)\,t} \tag{3.137}$$

and to compute $(\log m_{\mathbf{X}_j})^*$ we have to solve (3.133), i.e.

$$\frac{\tau}{h^2}\,(j-1-t)\,k^{j-1} + \left(1-2\frac{\tau}{h^2}\right)\,(j-t)\,k^j + \frac{\tau}{h^2}\,(j+1-t)\,k^{j+1} \stackrel{!}{=} 0 \tag{3.138}$$

for $k > 0$.

Since k is positive, we can divide by k^{j-1} and obtain a quadratic polynomial. Furthermore, we may assume j to be zero since we can substitute t by $t-j$. For $\frac{\tau}{h^2} < \frac{1}{4}$ this gives the positive solution

$$k = -\frac{1}{2}\frac{h^2 t - 2\tau t + \sqrt{h^4 t^2 - 4\tau h^2 t^2 + 4\tau^2}}{\tau(t-1)} \tag{3.139}$$

with $-1 \leqslant t \leqslant 1$.

Reintroducing j by translating t, we get an explicit representation of the structuring function. As expected, its Taylor expansion at $h = 0$ is given by

$$\frac{1}{4\,\frac{\tau}{h^2}}\,t^2 + O(h^4) \tag{3.140}$$

which approximates the structuring function

$$\frac{1}{4\,t}\|x\|^2 \tag{3.141}$$

of a continuous Gaussian.

Conjecture 3.8.2 tells us that the result for the Cramér–Fourier transform would be similar.

3.9 Summary

In this chapter we have established a mathematical dictionary that allows to translate any linear shift-invariant scale-space evolution into its morphological counterpart of Hamilton–Jacobi type and vice versa. In contrast to previous work on structural similarities between linear and morphological systems, we have achieved these equivalences in the terminology of differential or pseudodifferential operators. It turned out that the symbol p is a very simple and powerful concept: It allows to transform the linear evolution equation

$$\partial_t u = P(\boldsymbol{\nabla})u \tag{3.142}$$

into its morphological counterpart

$$\partial_t v = p(\boldsymbol{\nabla}v)\,. \tag{3.143}$$

By considering specific examples of linear or morphological scale-spaces we have discovered hitherto unexplored relations between known scale-spaces, such as the Poisson scale-space and morphology with a disc-shaped structuring element of increasing size. Moreover, novel scale-spaces have been introduced that have not been studied before, e.g. anisotropic Poisson scale-spaces and morphological relativistic scale-spaces.

Furthermore, we depicted a possible unification of Cramér and Cramér–Fourier transform by only considering the leading order term of both transforms. We stated the conjecture that these leading order terms always coincide (under the assumption that both transforms exist).

Chapter 4

Conclusions and Outlook

> Study nature, love nature, stay close to nature.
> It will never fail you.
>
> Frank Lloyd Wright

4.1 Conclusions

In this thesis we have seen several new scale-space aspects in image processing. First, we have analysed the continuous formulation of the linear osmosis filtering. This is an existing framework for seamless image manipulation based on a drift-diffusion equation where the desired outcome depends on the modelling of a drift vector field. The employed differential operator is well studied in physics and describes the evolution of the density of a moving particle in a conservative force field. Based on the scalar potential function of this force field, the evolution of the density can be written down explicitly. In our notation, the drift vector field would be the gradient of the scalar potential function. However, in practical applications, the constructed drift vector fields will not admit a scalar potential function. This lead us to considering not only the parabolic but also the elliptic formulation of the problem. We showed that the elliptic boundary value problem allows a one-dimensional solution space. Furthermore, we proved convergence for the evolution of the parabolic problem to a specific solution of the elliptic problem. This specific solution is singled out by the average grey value preservation of the parabolic image evolution - it is the solution that shares the average grey value with the initial value. We need this solution as well when

defining a class of Lyapunov functionals. They show the simplification of the image evolution in a weighted L^2-space. To sum up, we provided a complete scale-space characterisation of the linear osmosis filtering.

Second, we introduced the Cramér–Fourier transform. Apart from connecting linear shift-invariant and morphological scale-spaces on a structural and PDE level it also offers a way to construct new scale-spaces. The established dictionary allows to translate evolution equations between LSI and morphological settings. The Cramér–Fourier transform generalises previous research on the logarithmic connection between Gaussian and dilation/erosion scale-spaces. We have shown that the Fourier transform is the more natural choice to establish the above relationships. For the known correspondences, the results coincide. However, the Cramér–Fourier transform extends to the whole class of LSI scale-spaces.

4.2 Outlook

In this work we have been focussing solely on linear scale-space ideas: In Chapter 2 we have analysed a scale-space theory for the linear osmosis filter and in Chapter 3 we introduced the Cramér–Fourier transform to connect linear shift-invariant scale-spaces to morphological scale-spaces. While we already depicted a way to bring both the Cramér and Cramér–Fourier transform together again in Section 3.8, there are many more open challenges.

The linear osmosis filter was obtained by adding a drift term to a homogeneous linear diffusion equation. Although we have seen several applications in Section 2.3, we feel that the full potential of the osmosis model is not yet utilised. So far, neither the evolution itself has been used in practical applications nor drift vector fields depending on evolution itself have been studied.

It is also possible to introduce a diffusivity function $g\colon \Omega \to \mathbb{R}_{>0}$ into the evolution equation. This leads to a generalised osmosis evolution equation

$$\partial_t u = \operatorname{div}\left(g\nabla u - \mathbf{d}u\right) . \tag{4.1}$$

In a discrete setting, first results are already available: Baumann [2014] showed that there is a one-to-one correspondence between a specific, but straightforward space-discretisation of (4.1) and continuous-time Markov chains. He uses (4.1) to rank football teams by modelling both diffusivities and drift vectors based on given game results. However, in a continuous setting, no analysis was conducted yet and many of the employed methods in this work do not extend to this case. An importance special case is obtained when $g \equiv \varepsilon$ for a small constant $\varepsilon > 0$.

This leads to a similar situation as discussed in Section 3.3.1: The ansatz

$$u = a \exp\left(-\frac{1}{\varepsilon} w\right) \tag{4.2}$$

transforms

$$\partial_t u = \operatorname{div}\left(\varepsilon \nabla u - \mathbf{d} u\right) \tag{4.3}$$

into [Risken, 1984, Section 6.6.7]

$$\partial_t w = -|\nabla w|^2 - \mathbf{d} \cdot \nabla w + \mathcal{O}(\varepsilon)\,. \tag{4.4}$$

It would be interesting to see whether the above results combined with ideas from Section 3.8 could lead to a morphological counterpart of a linear osmosis filter. Furthermore, it would be desirable to include scale-spaces of nonlinear nature such as nonlinear diffusion [Perona and Malik, 1990, Weickert, 1998] and curvature-based morphological evolutions [Alvarez et al., 1993, Sapiro and Tannenbaum, 1993] into the theory.

Overall, there are a lot of interesting research questions and we hope that this work inspires more people to study linear scale-space approaches and extend the presented results in many unexpected ways.

Appendix A

Large Deviations Theory

> We all matter - maybe less then a lot but always more than none.
>
> John Green

Before we introduce the exact statements, we need some notations. From now on, let (Ω, B) be a measure space. A *rate function* is a lower semi-continuous function $I\colon \Omega \to \mathbb{R}_{>0}$ whose level sets are closed in Ω.

Definition A.1 (Large Deviation Principe)**.** *A family of probability measures* $\{\mu_n\}_{n\in\mathbb{N}}$ *on* (Ω, B) *satisfies the large deviation principle (LDP) with a rate function I if for all* $\Gamma \in B$

$$\liminf_{n\to\infty} \tfrac{1}{n} \log \mu_n(\Gamma) \geqslant - \inf_{x\in\Gamma^\circ} I(\mathbf{x}), \tag{A.1}$$

$$\limsup_{n\to\infty} \tfrac{1}{n} \log \mu_n(\Gamma) \leqslant - \inf_{x\in\bar{\Gamma}} I(\mathbf{x}) \tag{A.2}$$

where Γ° *and* $\bar{\Gamma}$ *denote the interior and closure of* Γ*, respectively.*

This definition is the most important one in large deviation theory and a lot of effort has been spent into proving that a LDP holds for specific families of probability measures. The first proven LDP goes back to Harald Cramér and is nowadays (together with all of its extensions) known as Cramér's Theorem. Cramér proved the following theorem in one dimension (see Dembo and Zeitouni [2009] for a proof and a historic overview).

Theorem A.2 (Cramér's Theorem in $\mathbb{R}^d$)**.** *Let $\{\mathbf{X}_i\}_{i\in I}$ be a family of real-valued independent identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with finite moment-generating function*

$$m(\mathbf{t}) = \mathbb{E}\left[e^{\langle \mathbf{t}, \mathbf{X}_1\rangle}\right]. \tag{A.3}$$

Let $\mathbf{Z}_n := \frac{1}{n}\sum_i \mathbf{X}_i$ be its empirical mean with distribution P_n. Then $\{P_n\}$ fulfils the LDP with rate function

$$I(\mathbf{x}) := \sup_{\mathbf{t}\in\mathbb{R}^d}\{\langle \mathbf{x}, \mathbf{t}\rangle - \log m(\mathbf{t})\}. \tag{A.4}$$

Theorem A.3 (Varadhan's Integral Lemma)**.** *Suppose that $\{\mu_\varepsilon\}$ satisfies the LDP with a good rate function $I\colon X \to [0,\infty]$, and let $\Phi\colon X \to \mathbb{R}$ be any continuous function. Assume further either the tail condition*

$$\lim_{M\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[e^{\Phi(Z_\varepsilon)/\varepsilon} 1_{\{\Phi(Z_\varepsilon)\geqslant M\}}\right] = -\infty \tag{A.5}$$

or the following moment condition for some $\gamma > 1$,

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[e^{\gamma\Phi(Z_\varepsilon)/\varepsilon}\right] < \infty. \tag{A.6}$$

Then

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[e^{\Phi(Z_\varepsilon)/\varepsilon}\right] = \sup_{x\in X}\{\Phi(\mathbf{x}) - I(\mathbf{x})\}. \tag{A.7}$$

Index

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1974.

M. Akian, J.-P. Quadrat, and M. Viot. Bellman processes. In *ICAOS '94: Discrete Event Systems*, volume 199 of *Lecture Notes in Control and Information Sciences*, pages 302–311. Springer, London, 1994.

M. Akian, J.-P. Quadrat, and M. Viot. Duality between probability and optimization. In Jeremy Gunawardena, editor, *Idempotency*, pages 331–353. Cambridge University Press, 1998.

L. Alvarez, P.-L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM Journal on Numerical Analysis*, 29: 845–866, 1992.

L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel. Axioms and fundamental equations in image processing. *Archive for Rational Mechanics and Analysis*, 123: 199–257, 1993.

J. Angulo. Pseudo-morphological image diffusion using the counter-harmonic paradigm. In Jacques Blanc-Talon, Don Bone, Wilfried Philips, Dan Popescu, and Paul Scheunders, editors, *Advanced Concepts for Intelligent Vision Systems*, volume 6474 of *Lecture Notes in Computer Science*, pages 426–437. Springer, Berlin, 2010.

A. B. Arehart, L. Vincent, and B. B. Kimia. Mathematical morphology: The Hamilton–Jacobi connection. In *Proc. Fourth International Conference on Computer Vision*, pages 215–219, Berlin, May 1993. IEEE Computer Society Press.

H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 1987.

F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity: An Algebra for Discrete Event Systems*. Wiley, Chichester, 1992.

J.-P. Bartier, J. Dolbeault, R. Illner, and M. Kowalczyk. A qualitative study of linear drift-diffusion equations with time-dependent or degenerate coefficients. *Mathematical Models and Methods in Applied Sciences*, 17(3):327–362, 2007.

A. E. Bashirov, E. M. Kurpınar, and A. Özyapıcı. Multiplicative calculus and its applications. *Journal of Mathematical Analysis and Applications*, 337(1):36 – 48, 2008.

C. K. Baumann. A Helmholtz-like decomposition for discrete osmosis and Markov chains. Bachelor's thesis, Saarland University, Saarbrücken, April 2014.

M. Breuß, B. Burgeth, and J. Weickert. Anisotropic continuous-scale morphology. In J. Martí, J. M. Benedí, A. M. Mendonça, and J. Serrat, editors, *Pattern Recognition and Image Analysis*, volume 4478 of *Lecture Notes in Computer Science*, pages 515–522. Springer, Berlin, 2007.

R. W. Brockett and P. Maragos. Evolution equations for continuous-scale morphology. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages 125–128, San Francisco, CA, March 1992.

B. Burgeth and J. Weickert. An explanation for the logarithmic connection between linear and morphological system theory. *International Journal of Computer Vision*, 64(2/3):157–169, September 2005.

B. Burgeth, S. Didas, and J. Weickert. Relativistic scale-spaces. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 1–12. Springer, Berlin, 2005a.

B. Burgeth, S. Didas, and J. Weickert. The Bessel scale-space. In O. F. Olsen, L. Florack, and A. Kuijper, editors, *Deep Structure, Singularities, and Computer Vision*, volume 3753 of *Lecture Notes in Computer Science*, pages 84–95. Springer, Berlin, 2005b.

R. S. Cantrell and C. Cosner. *Spatial Ecology via Reaction-Diffusion Equations*. Wiley Series in Mathematical & Computational Biology. Wiley, 2003.

J. A. Carrillo, S. Cordier, and S. Mancini. A decision-making Fokker–Planck model in computational neuroscience. *Journal of Mathematical Biology*, 63(5): 801–830, Nov 2011.

J. D. Cole. On a quasi-linear parabolic equation occurring in aerodynamics. *Quarterly of Applied Mathematics*, 9(3):225–236, 1951.

T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction To Algorithms*. MIT Press, 2001.

T. M. Cover and J. A. Thomas. *Elements of Information Theory*. A Wiley-Interscience publication. Wiley, 2006.

M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.*, 277:1 – 42, 1983.

M. G. Crandall, H. Ishii, and P. L. Lions. *User's Guide to Viscosity Solutions of Second Order Partial Differential Equations*. American Mathematical Society, 1992.

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer, 2009.

O. Demetz, J. Weickert, A. Bruhn, and H. Zimmer. Optic flow scale-space. In A. M. Bruckstein, B. ter Haar Romeny, A. M. Bronstein, and M. M. Bronstein, editors, *Scale Space and Variational Methods in Computer Vision*, volume 6667 of *Lecture Notes in Computer Science*, pages 713–724. Springer, Berlin, 2012.

J. K. G. Dhont. *An Introduction to Dynamics of Colloids*. Studies in Interface Science. Elsevier Science, 1996. ISBN 9780080535074.

S. Didas, B. Burgeth, A. Imiya, and J. Weickert. Regularity and scale-space properties of fractional high order linear filtering. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 13–25. Springer, Berlin, 2005.

E. H. Diop and J. Angulo. Multiscale image analysis based on robust and adaptive morphological scale-spaces. *Image Analysis & Stereology*, 34(1):39–50, 2014.

I. C. Dolcetta. Representations of solutions of Hamilton–Jacobi equations. In Daniela Lupo, CarloD. Pagani, and Bernhard Ruf, editors, *Nonlinear Equations: Methods, Models and Applications*, volume 54 of *Progress in Nonlinear Differential Equations and Their Applications*, pages 79–90. Birkhäuser Basel, 2003.

L. Dorst and R. van den Boomgaard. Morphological signal processing and the slope transform. *Signal Processing*, 38:79–98, 1994.

J. Droniou and J.-L. Vázquez. Noncoercive convection–diffusion elliptic problems with Neumann boundary conditions. *Calculus of Variations and Partial Differential Equations*, 34:413–434, 2009.

R. Duits, L. Florack, J. de Graaf, and B. ter Haar Romeny. On the axioms of scale space theory. *Journal of Mathematical Imaging and Vision*, 20:267–298, 2004.

R. S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics*, volume 271 of *Grundlehren der Mathematischen Wissenschaften*. Springer, New York, 1985.

L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 1998.

R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In *Proc. SIGGRAPH 2002*, pages 249–256, San Antonio, TX, July 2002.

M. Felsberg and G. Sommer. Scale-adaptive filtering derived from the Laplace equation. In B. Radig and S. Florczyk, editors, *Pattern Recognition*, volume 2032 of *Lecture Notes in Computer Science*, pages 95–106. Springer, Berlin, 2001.

L. Florack. *Image Structure*, volume 10 of *Computational Imaging and Vision*. Kluwer, Dordrecht, 1997.

L. Florack and H. van Assen. Multiplicative calculus in biomedical image analysis. *Journal of Mathematical Imaging and Vision*, 42(1):64–75, 2012.

L. Florack, R. Maas, and W. Niessen. Pseudo-linear scale-space theory. *International Journal of Computer Vision*, 31(2/3):247–259, April 1999.

M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012.

A. Friedman. Remarks on the maximum principle for parabolic equations and its applications. *Pacific J. Math.*, 8(2):201–211, 1958.

A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964.

H. Ge. Extended forms of the second law for general time-dependent stochastic processes. *Phys. Rev. E*, 80:021137, Aug 2009.

T. Georgiev. Covariant derivatives and vision. In H. Bischof, A. Leonardis, and A. Pinz, editors, *Computer Vision – ECCV 2006, Part IV*, volume 3954 of *Lecture Notes in Computer Science*, pages 56–69. Springer, Berlin, 2006.

D. Gilbarg and N. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin, reprint edition, 2001. Paperback edition.

B. V. Gnedenko. *Theory of Probability*. Taylor & Francis, 1998.

M. Grossman and R. Katz. *Non-Newtonian Calculus*. 1972.

F. Guichard and J.-M. Morel. Partial differential equations and image iterative filtering. In I. S. Duff and G. A. Watson, editors, *The State of the Art in Numerical Analysis*, number 63 in IMA Conference Series (New Series), pages 525–562. Clarendon Press, Oxford, 1997.

K. Hagenburg, M. Breuß, J. Weickert, and O. Vogel. Novel schemes for hyperbolic PDEs using osmosis filters from visual computing. In A. M. Bruckstein, B. ter Haar Romeny, A. M. Bronstein, and M. M. Bronstein, editors, *Scale Space and Variational Methods in Computer Vision*, volume 6667 of *Lecture Notes in Computer Science*, pages 532–543. Springer, Berlin, 2012.

H. J. A. M. Heijmans. Morphological scale-spaces, scale-invariance and Lie groups. In Hugues Talbot and Richard Beare, editors, *International Symposium on Mathematical Morphology*, pages 253–264, apr 2002.

E. Hopf. The partial differential equation $u_t + uu_x = \mu_{xx}$. *Communications on Pure and Applied Mathematics*, 3(3):201–230, 1950.

E. Hopf, C. S. Morawetz, J. Serrin, and I. A. G. Sina. *Selected Works of Eberhard Hopf: With Commentaries*. Collected works series: American Mathematical Society. American Mathematical Society, 2002.

L. Hörmander. *The Analysis of Linear Partial Differential Operators: Distribution Theory and Fourier Analysis*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, 1983a.

L. Hörmander. *The Analysis of Linear Partial Differential Operators II: Differential Operators with Constant Coefficients*, volume 274 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, 1983b.

L. Hörmander. *The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators*, volume 257 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, 1985.

T. Iijima. Basic theory on normalization of pattern (in case of typical one-dimensional pattern). *Bulletin of the Electrotechnical Laboratory*, 26:368–388, 1962. In Japanese.

T. Iijima. Basic theory on normalization of two-dimensional visual pattern. *Studies on Information and Control (IECE, Japan)*, (1):15–22, 1963. Pattern Recognition Issue. In Japanese.

T. Iijima. Basic equation of figure and observational transformation. *Systems, Computers, Controls*, 2(4):70–77, 1971. In English.

R. Illner and H. Neunzert. Relative entropy maximization and directed diffusion equations. *Mathematical Methods in the Applied Sciences*, 16:545–554, 1993.

P. T. Jackway. Properties of multiscale morphological smoothing by poweroids. *Pattern Recognition Letters*, 15(2):135 – 140, 1994.

P. T. Jackway. On dimensionality in multiscale morphological scale-space with elliptic poweroid structuring functions. *Journal of Visual Communication and Image Representation*, 6(2):189 – 195, 1995.

P. T. Jackway and M. Deriche. Scale-space properties of the multiscale morphological dilation–erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:38–51, 1996.

R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

F. Kanters, L. Florack, R. Duits, B. Platel, and B. Haar Romeny. ScaleSpaceViz: α-scale spaces in practice. *Pattern Recognition and Image Analysis*, 17(1): 106–116, 2007.

B. B. Kimia and K. Siddiqi. Geometric heat equation and non-linear diffusion of shapes and images. *Computer Vision and Image Understanding*, 64:305–322, 1996.

D. Kinderlehrer and M. Kowalczyk. Diffusion mediated transport and the flashing ratchet, 2001.

J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uraltseva. *Linear and Quasi-Linear Equations of Parabolic Type*. American Mathematical Society, translations of mathematical monographs. American Mathematical Society, 1968.

A. Landström. An approach to adaptive quadratic structuring functions based on the local structure tensor. In Jón Atli Benediktsson, Jocelyn Chanussot, Laurent Najman, and Hugues Talbot, editors, *Mathematical Morphology and Its Applications to Signal and Image Processing*, volume 9082 of *Lecture Notes in Computer Science*, pages 729–740. Springer, 2015.

G. M. Lieberman. Intermediate Schauder theory for second order parabolic equations. i. estimates. *Journal of Differential Equations*, 63(1):1 – 31, 1986.

G. M. Lieberman. The conormal derivative problem for non-uniformly parabolic equations. *Indiana Univ. Math. J.*, 37:23–72, 1988.

G. M. Lieberman. *Second Order Parabolic Differential Equations*. World Scientific, 1996.

F. Liese. ϕ-divergences, sufficiency, Bayes sufficiency, and deficiency. *Kybernetika*, (4):690–713, 2012.

T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer, Boston, 1994.

T. Lindeberg. Generalized Gaussian scale-space axiomatics comprising linear scale-space, affine scale-space and spatio-temporal scale-space. *Journal of Mathematical Imaging and Vision*, 40:36–81, 2011.

A. M. Lyapunov. Problème général de la stabilité du mouvement. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 9:203–474, 1907.

D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry*, volume 161 of *Graduate Studies in Mathematics*. AMS, Providence, 2015.

P. Maragos. Morphological systems: Slope transforms and max-min difference and differential equations. *Signal Processing*, 38(1):57–77, 1994.

A. I. Nazarov and N. N. Uraltseva. The oblique boundary-value problem for a quasilinear parabolic equation. *Journal of Mathematical Sciences*, 77: 3212–3220, 1995.

C. Niculescu and L. E. Persson. *Convex Functions and their Applications: A Contemporary Approach*. Number Bd. 13 in CMS Books in Mathematics. Springer, 2005.

M. Nielsen, L. Florack, and R. Deriche. Regularization, scale-space and edge detection filters. *Journal of Mathematical Imaging and Vision*, 7:291–307, 1997.

L. Nirenberg. A strong maximum principle for parabolic equations. *Communications on Pure and Applied Mathematics*, 6(2):167–177, 1953.

N. Otsu. Mathematical studies on feature extraction in pattern recognition. Technical Report 818 (PhD Thesis), Electrotechnical Laboratory, Tsukuba, Japan, July 1981. In Japanese.

E. J. Pauwels, L. J. Van Gool, P. Fiddelaers, and T. Moons. An extended class of scale-invariant and recursive scale space filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:691–701, 1995.

P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

T. Poggio, H. Voorhees, and A. Yuille. A regularized solution to edge detection. *Journal of Complexity*, 4(2):106 – 123, 1988.

M. H. Protter and H. F. Weinberger. *Maximum Principles in Differential Equations*. Springer, Berlin, 1999.

R. D. Richtmyer. *Principles of advanced mathematical physics*. Number Bd. 1 in Texts and monographs in physics. Springer Verlag, 1978.

H. Risken. *The Fokker–Planck Equation*. Springer, New York, 1984.

R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

G. Sapiro and A. Tannenbaum. Affine invariant scale-space. *International Journal of Computer Vision*, 11:25–44, 1993.

O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *Journal of Mathematical Imaging and Vision*, 12(1):43–63, February 2000.

M. Schmidt and J. Weickert. The morphological equivalents of relativistic and alpha-scale-spaces. In J.-F. Aujol, M. Nikolova, and N. Papadakis, editors, *Scale Space and Variational Methods in Computer Vision*, volume 9087 of *Lecture Notes in Computer Science*, pages 28–39. Springer, Berlin, 2015.

M. Schmidt and J. Weickert. Morphological counterparts of linear shift-invariant scale-spaces. *Journal of Mathematical Imaging and Vision*, 56(2):352–366, 2016.

J. Smoller. *Shock Waves and Reaction—Diffusion Equations*. Grundlehren der mathematischen Wissenschaften 258. Springer US, 1983.

J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors. *Gaussian Scale-Space Theory*, volume 8 of *Computational Imaging and Vision*. Kluwer, Dordrecht, 1997.

M. Taylor. *Partial Differential Equations II: Qualitative Studies of Linear Equations*. Applied Mathematical Sciences. Springer, New York, 2010.

R. van den Boomgaard. *Mathematical Morphology: Extensions Towards Computer Vision*. PhD thesis, University of Amsterdam, The Netherlands, 1992a.

R. van den Boomgaard. The morphological equivalent of the Gauss convolution. *Nieuw Archief Voor Wiskunde*, 10(3):219–236, November 1992b.

R. van den Boomgaard and L. Dorst. The morphological equivalent of Gaussian scale-space. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory*, volume 8 of *Computational Imaging and Vision*, pages 203–220. Kluwer, Dordrecht, 1997.

R. van den Boomgaard and A. Smeulders. The morphological structure of images: The differential equations of morphological scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:1101–1113, 1994.

H. A. van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 13(2):631–644, 1992.

N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, third edition, 2007.

S. Varadhan. *Large Deviations and Applications*. Society for Industrial and Applied Mathematics, 1984.

O. Vogel, K. Hagenburg, J. Weickert, and S. Setzer. A fully discrete theory for linear osmosis filtering. In A. Kuijper, K. Bredies, T. Pock, and H. Bischof, editors, *Scale Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 368–379. Springer, Berlin, 2013.

J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart, 1998.

J. Weickert, S. Ishikawa, and A. Imiya. Linear scale-space has first been proposed in Japan. *Journal of Mathematical Imaging and Vision*, 10(3):237–252, May 1999.

J. Weickert, K. Hagenburg, M. Breuß, and O. Vogel. Linear osmosis models for visual computing. In A. Heyden, F. Kahl, C. Olsson, M. Oskarsson, and X.-C. Tai, editors, *Energy Minimisation Methods in Computer Vision and Pattern Recognition*, volume 8081 of *Lecture Notes in Computer Science*, pages 26–39. Springer, Berlin, 2013.

M. Welk. Families of generalised morphological scale spaces. In L.D. Griffin and M. Lillholm, editors, *Scale Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 770–784. Springer, Berlin, 2003.

P. Winkert. L^∞-estimates for nonlinear elliptic Neumann boundary value problems. *Nonlinear Differential Equations and Applications NoDEA*, 17(3):289–302, 2010.

A. P. Witkin. Scale-space filtering. In *Proc. Eighth International Joint Conference on Artificial Intelligence*, volume 2, pages 945–951, Karlsruhe, West Germany, August 1983.

Own Publications

M. Schmidt and J. Weickert. The morphological equivalents of relativistic and alpha-scale-spaces. In J.-F. Aujol, M. Nikolova, and N. Papadakis, editors, *Scale Space and Variational Methods in Computer Vision*, volume 9087 of *Lecture Notes in Computer Science*, pages 28–39. Springer, Berlin, 2015.

M. Schmidt and J. Weickert. Morphological counterparts of linear shift-invariant scale-spaces. *Journal of Mathematical Imaging and Vision*, 56(2):352–366, 2016.