Pius Sonnberger BSc.

A Functional Lifting Approach for spatio-temporal Image Processing Problems and Applications to Image-Sequence Reconstruction

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Technical Mathematics

submitted to

Graz University of Technology

Supervisor:
Univ.-Prof. Dipl.-Math. Dr. Kristian Bredies

Institute for Mathematics and Scientific Computing
University of Graz

Graz, May 2016

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Datum/Date	Unterschrift/Signature

Abstract

In this thesis we would like to investigate a higher order regulariser with application on image sequences using a functional lifting approach. The presented methods do not only take the image gradient into account, but also curvature information of the level sets. The theory developed operates on continuous time-dependent problems and is independent of the problem dimension. Existence theorems in the resulting continuous function spaces for the derived approach are presented as well. Since the derived theory works on continuous domains, we also investigate a discretisation of the introduced approaches. A primal-dual algorithm is used to solve the upcoming minimisation problems. Further we give results for image sequences concerning denoising and inpainting problems.

Contents

1. Introduction

In image processing it has to be dealt with degraded images that are biased by noise, blurring, data-losses or other effects.
In this work we would like to discuss well-known and frequently-used approaches and their theories to overcome these issues. Further, a recent method will be introduced that showed very promising results in 2D image reconstruction [4]. The novelty in this work is that this theory is extended to higher dimensional problems (dimension greater or equal 3), to make this idea applicable on image sequences (e.g. videos) or moving 3D image data. We will treat the time dependence of these problems with special care. This will lead to an improvement of the results in comparison to standard approaches that treat the time dependence in the same way as the spatial dependence of the problem. Further some discretisation steps are remodelled in comparison to [4] to make the introduced approach working in higher dimensions.
The task of image processing tools is, to improve the "quality" of degraded images (e.g. deblurring, denoising, inpainting) or extract information from images (e.g. segmentation, registration). An often used approach to deal with these tasks is to measure how natural an image is, in some sense. Obviously, the answer to the question if an image is natural is neither easy nor unique. This leads to a wide range of different image processing approaches.
To make any kind of image-processing tool work, we first have to discuss how images can be modelled in an abstract mathematical framework. Therefore we define an image as a function, mapping a certain domain to a colour space [3].
The domain can be a continuous or a discrete set. A typically example for a continuous domain would be some bounded subset of $\mathbb{R}^2$ or $\mathbb{R}^3$. Examples for discrete domains would be a 2D domain that is composed of a finite amount of pixels or a 3D domain that is composed of a finite amount of voxels.
The colour space is also either continuous or discrete. An often used colour space is the grey-value space, where the image maps into the interval $[0, 1]$ or a reasonable discretisation of it. Another colour space is the RGB space, where each of the colours red, green and blue is represented as a value in the interval $[0, 1]$ or a discretisation

of this interval. This would mean, that the image maps to the set $[0,1]^3$. It has to be mentioned, that there are a lot more possibilities to represent colours (e.g. CMYK-space).
Anyhow, this work will focus only on images mapping on the grey-value space. Although it is fairly easy to come up with ideas to extend approaches for grey-valued images to coloured images, the results may not be as good as expected, since it is hard to model the colour perception of the human eye.
Further it has to be mentioned, that all theory developed in this work is done for a continuous domain and is then transferred to discrete settings using reasonable discretisations.
With this given basic setting of an image we can again ask the question, how to measure the naturalness or regularity of an image. In the history of image processing first-order regularisation was the common way to treat image processing problems [2, 11, 14, 17]. First-order regularisation means that the image gradient is penalised in some sense. Therefore, regions with the same intensity values would be preferred. For example, Total Variation regularisation (TV) [17] is a common first-order approach in image processing. Deblurring and denoising results obtained by TV-Regularisation, using a Primal-Dual Algorithm [6], can be seen in Figure 1.1 and Figure 1.2.

(a) given image (ground truth) (b) blurred image (c) Reconstruction

Figure 1.1.: Reconstruction of a blurred image using a TV-Regulariser.

As it can be seen first-order-regularisation has its drawbacks, since it leads to the so-called staircasing effect in the reconstructed image. Other first-order approaches overcome the problem of this artefacts, at the cost of blurry edges at object boundaries.
This gives the intuition, that first-order approaches may not be the best way to

(a) given image (ground truth) (b) noisy image (c) Reconstruction

Figure 1.2.: Reconstruction of a noisy image using a TV-Regulariser.

treat imaging problems. Some publications (e.g. [13]) are also giving evidence that curvature information is important for the human perception, what would lead to second-order approaches, which makes the study of higher-order regularisers relevant.

In this work we would like to discuss a class of functionals that deals with the curvature of object boundaries using a functional lifting approach. The idea of this functional lifting approach is presented in [7] for two dimensional images. There, the direction of the absolute image gradient is lifted to the so-called three dimensional rototranslation space. This means, that the gradient direction of an image is "lifted" in the third dimension. Having the lifted gradient provides not only its absolute value but also its direction in every point. This can be interpreted by the transformation of the image gradient to the polar coordinate system, where the polar angle of the absolute gradient is represented as the third dimension of the rototranslation space.

We want to adapt this idea to higher dimensional problems. For example, the lifting of three dimensional problems will lead to objects in a five dimensional rototranslation space, since two further dimensions are necessary to encode the direction of the derivative.

In a recent paper [4] a curvature-penalising functional was derived for two dimensional images using functional lifting. The continuation of this theory in higher dimensions was mentioned as an open problem. Here we want to derive a theory that is independent of the problem dimension, using the ideas presented in [4]. Furthermore we want to develop reasonable numerical approaches and give results for three dimensional objects as well.

In Chapter 2, the needed mathematical concepts are explained to get a basic under-

standing of mathematical image processing. This Chapter should provide a fundamental knowledge to get easier access to the content of the further Chapters. The required theroy for first-order-regularisers and tools for verifying solvability and convergence properties of the used numerical methods are presented as well. In Chapter 3, a class of curvature penalising functionals is derived which works in higher dimensions. Since the focus is on image sequences, the time-axis is always one dimension of the problem and is treated with special care. We want to point out again that the investigation of these functionals for higher dimensional problems, as well as the special treatment of the time dependence of the problems are the main novelties in this work. In Chapter 4, a discretisation approach for the continuous theory developed in Cahpter 3 is presented. Further numerical methods to solve the resulting minimisation problems are derived in this Chapter. The numerics is done for image sequences and has its own specialties at some point. In the Appendix we will discuss discretisation methods for two dimensional, time-depended problems. This means that the considered problems have one spatial and one time component. Further we present an approach that allows refinement of the discretisation for the two dimensional case.

2. Mathematical Fundamentals of Image Processing

In this chapter we will provide the mathematical background to handle a wide range of image processing problems. First we will again discuss the model of an image in a mathematical sense and how regularisers could look like, which operate on this images. We will see, that variational methods are the key to discuss existence of solutions for the upcoming equations. The fundamental ideas to model first-order-regularisers for continuous domains will be given and therefore the required tools of Functional Analysis and Measure Theory will be introduced.
Further theory about Duality and Saddle-Points will be presented, what provides an appropriate way to make the derived formulations computable. Convergence theorems for the derived algorithms will be given as well.

2.1. Mathematical Model of Images and Image Regulariser

To deal with images in a mathematical sense let us first discuss the mathematical model of an image [3]. An image can be interpreted as a function u out of some function space X that is defined on a domain Ω. The domain Ω is usually a bounded subset of $\mathbb{R}^n$ or an appropriate discretisation of those subsets. This discretisations lead to the concepts of pixels for 2D images and voxels for 3D images or image sequences, respectively.
The image function u maps to some colour space $\mathcal{F}$, following the notation presented in [3]. As already mentioned in the introduction, the colour space in this work will always be the interval $[0, 1]$, since we are only dealing with grey-scale images. It has to be mentioned that other subsets of $\mathbb{R}$ are also common to describe grey-scale

images, e.g. the interval $[0, 256]$.
Given this setting an image $u \in X$ can be written as

$$u : \Omega \to \mathcal{F},$$

what simplifies for the special case of grey-scale images to

$$u : \Omega \to [0, 1].$$

We want to motivate the upcoming theory on the example of noisy images. Let us assume that noise is purely additive and uniformly distributed on the whole image domain Ω. For a ground truth image $u^\dagger$ and a noise function η the resulting noisy image u^0 is given as

$$u^0 = u^\dagger + \eta.$$

The expectations that

- $u^\dagger$ has some kind of structure, like connected regions with the same grey value and
- η is a function with no structure,

motivates the properties of regularisers.
The idea of variational denoising approaches is to quantify the local noise information on the one hand and the global regularity on the other hand, using real valued functions.
Therefore we introduce a real-valued function ϕ that measures the local noise. For example, such a function can be given as $\phi : X \to \mathbb{R}_\infty$,

$$\phi(u - u^0) = \int_\Omega |u - u^0|^q dx, \tag{2.1}$$

for some $q > 1$ with $\mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$. The function ϕ will be denoted as the data term. It is clear that the data term has to change, if the problem type changes. Examples for different problem types are deblurring problems where the given image is out of focus or inpainting problems where several image data is lost.

To measure the regularity of an image, a real-valued function $\psi : X \to \mathbb{R}_\infty$ is introduced that takes the structure of the image into account. A reasonable example for this function is

$$\psi(u) = \int_\Omega |\nabla u|^p dx, \tag{2.2}$$

for some $p > 1$. The function ψ will be denoted as the regulariser.
Adding up these two functionals using a weight $\lambda > 0$ lead to the formulation

$$F(u) := \lambda \cdot \phi(u - u^0) + \psi(u).$$

The idea is to search for the image $u^* \in X$ that minimises this formulation leading to the minimising problem

$$\min_{u \in X} F(u). \tag{2.3}$$

In the next section we want to present the needed conditions that ensure existence of a minimiser $u^* \in X$, fulfilling

$$F(u^*) = \min_{u \in X} F(u).$$

2.2. The Direct Method

A well established method to prove the existence of a minimiser for (2.3) is the direct method. Subsequently an overview of the required definitions and theorems is given. For further details see [3, 20].

Definition 1. *A Banach Space X is called **reflexive** if its canonical mapping into the bidual space is surjective. The canonical mapping ι is defined as*

$$\iota(x) : X^* \mapsto \mathbb{K}, \qquad (\iota(x))(x^*) = x^*(x)$$

where X^* is the dual space of X.

In this work $\mathbb{K}$ will always be $\mathbb{R}$.

Remark 2. *The canonical mapping ι is an element of the bidual space X^{**}.*

Definition 3. *Let X be a Banach Space. A sequence $\{u_n\}_n$ in X* ***converges in a weak sense*** *if*

$$< \xi, u_n > \rightarrow < \xi, u >, \quad \forall \xi \in X^*.$$

A weakly convergent sequence is denoted by $u_n \rightharpoonup u$.

Definition 4. *Let X be a Banach Space. A sequence $\{u_n\}_n$ in X^** ***converges in a weak* sense*** *if*

$$< \xi, u_n > \rightarrow < \xi, u >, \quad \forall \xi \in X.$$

A weakly convergent sequence is denoted by $u_n \overset{*}{\rightharpoonup} u$.*

Definition 5. *Let X be a Banach Space, $F : X \rightarrow \mathbb{R}_\infty$ be a functional.*

(a) *The* ***domain*** *of F is defined as*

$$\mathrm{dom}(F) = \{u \in X | F(u) < \infty\}.$$

(b) *F is called* ***proper****, if* $\mathrm{dom}(F) \neq \emptyset$.
(c) *F is called* ***coercive*** *if $F(u) \rightarrow +\infty$ holds for $\|u\|_X \rightarrow \infty$.*
(d) *F is called* ***lower semi-continuous****, if*

$$F(u) \leq \liminf_{n \rightarrow \infty} F(u_n)$$

holds for every sequence $\{u_n\}_n$ in X with $u_n \rightarrow u$ as $n \rightarrow \infty$.
(e) *F is called* ***weak lower semi-continuous****, if*

$$F(u) \leq \liminf_{n \rightarrow \infty} F(u_n)$$

holds for every sequence $\{u_n\}_n$ in X with $u_n \rightharpoonup u$ as $n \rightarrow \infty$.
(f) *Let $G : X^* \rightarrow \mathbb{R}_\infty$ be a functional. G is called* ***weak* lower semi-continuous****, if*

$$G(u) \leq \liminf_{n \rightarrow \infty} G(u_n)$$

holds for every sequence $\{u_n\}_n$ in X^* with $u_n \overset{*}{\rightharpoonup} u$ as $n \rightarrow \infty$.

Theorem 6 (Theorem of Eberlein-Šmulyan)**.** *A Banach space X is reflexive if and only if every bounded sequence $\{u_n\}_n$ in X has a weak convergent subsequence.*

Proof. see [8]. □

Theorem 7 (Direct Method)**.** *Let X be a reflexive Banach Space. The functional $F : X \to \mathbb{R}_\infty$ is a proper, coercive, lower semi-continuous functional that is bounded from below. Then, F has a minimiser in X.*

Sketch of the proof. Since F is bounded from below it follows that there exists a sequence $\{u_n\}_n$ with $F(u_n) \to \inf_{u\in X} F(u)$.
Using the coercivity of F it is clear that a bounded subsequence $\{u_{n_k}\}_k$ can be chosen. Otherwise $\|u_n\|_X \to \infty$ would lead to $F(u_n) \to \infty$.
According to Theorem 6, there exists a weakly convergent subsequence $\{u_{n_{k_l}}\}_l$ and an element $u^* \in X$ that fulfils $u_{n_{k_l}} \rightharpoonup u^*$. Weak lower semi-continuity gives the desired minimiser. □

Remark 8. *To prove existence of a minimiser for the problem given at* (2.3) *one has to make sure that the conditions of Theorem 7 are satisfied. To show lower semi-continuity it is sufficient to investigate data term and regulariser separately since*

$$\liminf_{n\to\infty} F(u_n) + G(u_n) \leq \liminf_{n\to\infty} F(u_n) + \liminf_{n\to\infty} G(u_n)$$

holds for F and G lower semi-continuous.

2.3. Sobolev Spaces

In this section we want to discuss the needed function spaces to apply the direct method. As already indicated in (2.2), we want to integrate over the derivative of an image. This leads to the concept of weak derivatives and the theory of Sobolev Spaces. The application on the denoising problem is shown as well.

Definition 9. *The space $L^1_{\text{loc}}(\Omega)$ is defined as*

$$L^1_{\text{loc}}(\Omega) = \left\{ u \in L^0(\Omega) \,\middle|\, \int_\Omega u \cdot \phi \,\mathrm{d}x < \infty,\ \forall \phi \in D(\Omega) \right\}$$

where $L^0(\Omega)$ is the equivalence class of all measurable functions $u : \Omega \to \mathbb{R}$ that are identical almost everywhere and $D(\Omega) \cong C_C^\infty(\Omega)$ is the set of all test functions in Ω.

For further informations about the space $L^1_{\text{loc}}(\Omega)$ see [10].

Lemma 10 (Fundamental Lemma of Variational Calculus)**.** *For $u \in L^1_{\text{loc}}(\Omega)$ and $\int_\Omega u \cdot \phi \, \mathrm{d}x = 0$ for all $\phi \in D(\Omega)$ it follows that*

$$u = 0$$

almost everywhere in Ω.

Proof. see [12]. □

With Lemma 10 it is possible to define derivatives in the weak sense.

Definition 11. *Let $u \in L^1_{\text{loc}}(\Omega)$, $\alpha \in \mathbb{N}^d$ be a multi-index. Then $v \in L^1_{\text{loc}}(\Omega)$ is called the **weak derivative** of u if*

$$\int_\Omega v \cdot \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \cdot \partial^\alpha \phi \, \mathrm{d}x,$$

for all $\phi \in D(\Omega)$. The weak derivative is denoted as $v := \partial^\alpha u$.

Lemma 12. *The weak derivative is unique almost everywhere.*

Proof. Let $u \in L^1_{\text{loc}}(\Omega)$, $v_1 = \partial^\alpha u$ and $v_2 = \partial^\alpha u$. Applying the definition for the weak derivative leads to

$$\int_\Omega v_1 \cdot \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \cdot \partial^\alpha \phi \, \mathrm{d}x$$

and

$$\int_\Omega v_2 \cdot \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u \cdot \partial^\alpha \phi \, \mathrm{d}x.$$

This gives

$$\int_\Omega v_2 \cdot \phi \, \mathrm{d}x = \int_\Omega v_1 \cdot \phi \, \mathrm{d}x,$$

what leads to

$$\int_\Omega (v_2 - v_1) \cdot \phi \, \mathrm{d}x = 0.$$

With the Fundamental Lemma of Variational Calculus (Lemma 10) $v_1 = v_2$ holds almost everywhere. □

Definition 13. *Let* $1 \leq p \leq \infty$ *and* $m \in \mathbb{N}$. *The* ***Sobolev Space*** *is defined as*

$$H^{m,p}(\Omega) = \left\{ u \in L^p(\Omega) \middle| \partial^\alpha u \in L^p(\Omega) \text{ for } |\alpha| \leq m \right\}.$$

With the norm

$$\|u\|_{m,p} = \left(\sum_{|\alpha| \leq m} \|\partial^\alpha u\|_p^p \right)^{1/p},$$

$H^{m,p}(\Omega)$ *becomes a Banach Space.*

Theorem 14. *For* $1 < p < \infty$, *the Sobolev Space* $H^{m,p}(\Omega)$ *is reflexive.*

Proof. see [1]. □

This setting delivers the right function space to solve problems of type (2.3). For denoising problems this can be summarised with the following theorem.

Theorem 15. *Let* $1 < p \leq q < \infty$, $\lambda > 0$ *and* $u^0 \in L^q(\Omega)$. *Then the functional* $F : L^q(\Omega) \to \mathbb{R}_\infty$

$$F(u) = \begin{cases} \lambda \int_\Omega |u - u^0|^q \, \mathrm{d}x + \int_\Omega |\nabla u|^p \, \mathrm{d}x & \text{if } u \in H^{1,p}(\Omega) \\ \infty & \text{else} \end{cases}$$

has a minimiser.

Proof. Since $H^{1,p}(\Omega)$ is reflexive according to Theorem 14 it has to be shown that F is proper, coercive and lower semi-continuous (see [3]). Consequently, Theorem 7 provides a minimiser $u^* \in H^{1,p}(\Omega)$. □

However it is well known, that a regularisation with $p > 1$ leads to blurry edges in the solution, which is not wanted in most applications. Accordingly, one might wish for a minimiser in the space $H^{1,1}(\Omega)$. The problem here is that the space $H^{1,1}(\Omega)$ is not reflexive anymore. Another problem is, that $u \mapsto |u|_{1,1} = \int_\Omega |\nabla u| \, \mathrm{d}x$ is not (weak) lower semi-continuous in $L^q(\Omega)$, as it is shown in Example 16. Therefore, the Direct Method (Theorem 7) can not be applied and existence of a minimiser can not be achieved (see [3]).

Example 16. Let $\Omega =]0,1[$. Set

$$\psi(u) = \begin{cases} \int_\Omega |\nabla u| \, \mathrm{d}x & \text{if } u \in H^{1,1}(\Omega) \\ \infty & \text{else} \end{cases}$$

and $u^n(x) = \max(\min(n(x-\frac{1}{2}),1),-1)$ which leads to the weak derivative according to Definition 11

$$(u^n)'(x) = \begin{cases} n & \text{if } |x-\frac{1}{2}| \leq \frac{1}{n} \\ 0 & \text{else.} \end{cases}$$

The limit $u := \lim_{n\to\infty} u^n$ is given as

$$u(x) = \begin{cases} 1 & \text{if } x \geq -\frac{1}{2} \\ -1 & \text{else.} \end{cases}$$

Therefore $u \notin H^{1,1}(\Omega)$, what is leading to $\psi(u) = \infty$. This contradicts with the lower semi continuity, since $\psi(u^n) \leq 2$, giving $\psi(u) > \liminf_{n\to\infty} \psi(u^n)$.

2.4. Bounded Variation

To cope with this problem the space $H^{1,1}$ is extended in a way that the weak derivative of a functional may be a measure as well. This will lead to functions with bounded variation and the total variation (TV) measure. TV is a reasonable and often used first-order regulariser, that makes the existence of sharp edges possible. The new approaches developed in Chapter 3 and 4 will be tested against results achieved with TV-regularisation. Since the theory used for the method introduced in Chapter 3 has its starting point in measure theory as well, some deeper concepts will be discussed in subsection 2.4.2.

Definition 17. *Let X be a set. The system $\mathcal{A} \subset \mathcal{P}(\mathcal{A})$ is called a* ***σ-Algebra of*** *X if*

1. $X \in \mathcal{A}$,
2. $E \in \mathcal{A} \;\Rightarrow\; X \backslash E \in \mathcal{A}$,
3. $E_n \in \mathcal{A},\ n = 1,2,\ldots \;\Rightarrow\; \bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$.

Definition 18. *The **Borel σ-Algebra** is the smallest σ-Algebra containing all open sets of a set Ω. The Borel σ-Algebra of Ω will be denoted as $\mathcal{B}(\Omega)$.*

Definition 19. *Let $\Omega \subset \mathbb{R}^d$ be open. A function $\mu : \mathcal{B}(\Omega) \to \mathbb{R}^n$ is called **vector-valued Radon Measure** if*

1. *$\mu(\emptyset) = 0$.*
2. *For pairwise disjoint sets $E_i \in \mathcal{B}(\Omega)$, $i \in \mathbb{N}$ the equality $\mu\left(\bigcup_{i\in\mathbb{N}} E_i\right) = \sum_{i\in\mathbb{N}} \mu(E_i)$ holds.*

Definition 20. *The **total variation of a measure** $|\mu| : \mathcal{B} \to \mathbb{R}$ of a Radon Measure μ is defined as*

$$|\mu|(A) = \sup\left\{\sum_{i\in\mathbb{N}} \left|\mu(A_i)\right| \,\middle|\, A = \bigcup_{i\in\mathbb{N}} A_i,\ A_i \in \mathcal{B}(\Omega) \text{ disjoint}\right\},$$

where the supremum is taken over all decompositions $\bigcup_{i\in\mathbb{N}} A_i = A$, with $A_i \in \mathcal{B}(\Omega)$ being disjoint.

Remark 21. *The set $\mathcal{M}(\Omega, \mathbb{R}^n)$ of all $\mathbb{R}^n$-valued Radon Measures, with the norm*

$$\|\mu\|_{\mathcal{M}} = |\mu|(\Omega)$$

is a Banach Space (see [18]).

2.4.1. Function of Bounded Variation

It will turn out, that the space of functions with bounded variation is the correct function space to search for solutions of the minimisation problem (2.3), if TV-regularisation is used. It will also be the function space used for the new concepts discussed in Chapter 3.

Utilising the Riesz Representation Formula (see [19]) we get $C_0(\Omega, \mathbb{R}^n)^* = \mathcal{M}(\Omega, \mathbb{R}^n)$ (see [3]), with

$$C_0(\Omega, \mathbb{R}^n) = \Big\{u : \Omega \to \mathbb{R}^n \,\Big|\, u \in C(\Omega, \mathbb{R}^n), \\ \forall \varepsilon > 0,\ \exists K \subset \Omega \text{ compact, s.t. } |u| \leq \varepsilon \text{ on } \Omega \setminus K\Big\}.$$

Definition 22. *Let $\Omega \subset \mathbb{R}^d$ be open and non empty. The measure $\mu \in \mathcal{M}(\Omega, \mathbb{R}^d)$ is called the **weak derivative** of $u \in L^1_{\text{loc}}(\Omega)$ if*

$$\int_\Omega u \cdot \operatorname{div}\phi \, \mathrm{d}x = -\int_\Omega \phi \, \mathrm{d}\mu$$

holds for all test functions $\phi \in D(\Omega, \mathbb{R}^d)$. The weak derivative μ is then written as $\mu = \nabla u$.
*The **total variation** of u is denoted by*

$$\mathrm{TV}(u) = \|\nabla u\|_{\mathcal{M}}.$$

Lemma 23. *The weak derivative in $\mathcal{M}(\Omega, \mathbb{R}^d)$ is unique and it holds that*

$$\mathrm{TV}(u) = \sup\left\{ \int_\Omega u \operatorname{div}\phi \, \mathrm{d}x \,\middle|\, \phi \in D(\Omega, \mathbb{R}^d), \|\phi\|_\infty \leq 1 \right\}.$$

Proof. see [3]. □

Definition 24. *Let $\Omega \in \mathbb{R}^d$ be bounded. Then*

$$\mathrm{BV}(\Omega) = \left\{ u \in L^1(\Omega) \middle| \nabla u \in \mathcal{M}(\Omega, \mathbb{R}^d) \right\}$$

*is called the space of **Functions with Bounded Variation**, with the norm*

$$\|u\|_{\mathrm{BV}} = \|u\|_1 + \mathrm{TV}(u).$$

Lemma 25.

1. $\mathrm{BV}(\Omega)$ *is a Banach Space.*
2. TV *is proper, convex and weak lower semi-continuous on $L^q(\Omega)$ for $1 \leq q < \infty$.*

Definition 26 (Hausdorff measure)**.** *Let $k \in \mathbb{R}$, $k \geq 0$ and denote by ω_k the volume of the k dimensional unit ball.*
Define

$$\begin{aligned} \mathrm{Diam}(A) &= \sup\left\{ |x - y| \,\middle|\, x, y \in A \right\} \\ \mathrm{Diam}(\emptyset) &= 0. \end{aligned}$$

*The **k-dimensional Hausdorff measure** of a set $A \subset \mathbb{R}^d$ is defined as*

$$\mathcal{H}(A) = \lim_{\delta \to 0} \frac{\omega_k}{2^k} \inf\left\{ \sum_{i=1}^{\infty} \mathrm{Diam}(A_i)^k \,\middle|\, A \subset \bigcup_{i=1}^{\infty} A_i,\ \mathrm{Diam}(A_i) < \delta \right\}.$$

Remark 27. *With the Hausdroff measure it is possible to measure k-dimensional objects in d-dimensional spaces for $k \leq d$. For example, for a closed curve in 2D the Hausdorff measure will give the perimeter of the curve. In case $k = d$, the Hausdorff measure and Lebesgue measure deliver the same value.*

Example 28. Let $\Omega' \subset \Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $f \in L^1(\partial\Omega')$ with respect to $\mathcal{H}^{d-1}$. Then,

$$\mu(A) = \int_{A \cup \partial\Omega'} f \, \mathrm{d}\mathcal{H}^{d-1}$$

is a Radon measure.
Notation: $\mu = f\mathcal{H}^{d-1} \llcorner \partial\Omega'$.

Example 29. Let $\Omega' \subset \Omega$ be a Lipschitz domain and $u = \chi_{\Omega'}$ be the characteristic function of Ω'. Then

$$\int_\Omega u \, \mathrm{div}\phi \, \mathrm{d}x = \int_{\Omega'} \mathrm{div}\phi \, \mathrm{d}x = \int_{\partial\Omega'} \phi \cdot \nu \, \mathrm{d}\mathcal{H}^{d-1},$$

where ν is the outer normal of Ω', what leads to

$$\nabla u = -\nu \left(\mathcal{H}^{d-1} \llcorner \partial\Omega' \right).$$

For $u \in \mathrm{BV}(\Omega)$, the total variation $\mathrm{TV}(u) = \int_{\partial\Omega'} 1 \, \mathrm{d}\mathcal{H}^{d-1} = \mathrm{Per}(\Omega')$ gives the perimeter of Ω'.

2.4.2. Absolute Continuity of Measures

In this section, the concept of absolute continuous measures, the Lebesgue Decomposition Theorem, Theorem of Radon Nikodym and Polar Decomposition are discussed. We will make heavy use of these theorems in Chapter 3. A very deep look into these and other ideas of measure theory are given in [18].

Definition 30. *Let $\mu : \mathcal{M} \to [0, \infty]$ be a positive measure on some σ-Algebra $\mathcal{M}$ and λ be a positive or vector-valued measure on $\mathcal{M}$. Then*

1. *λ is called **absolutely continuous** with respect to μ if*

$$\mu(E) = 0 \Rightarrow \lambda(E) = 0, \, \forall E \in \mathcal{M}.$$

Notation: $\lambda \ll \mu$.

2. λ *is called* ***concentrated*** *on* $A \in \mathcal{M}$*, if* $\lambda(E) = \lambda(E \cap A)$, $\forall E \in \mathcal{M}$.
3. λ_1 *is called* ***singular*** *with respect to* λ_2 *if there exist* $A, B \in \mathcal{M}$ *disjoint, such that* λ_1 *is concentrated on A and* λ_2 *is concentrated on B. Notation:* $\lambda_1 \perp \lambda_2$.

Proposition 31. *Let* μ *be a positive measure on* $\mathcal{M}$ *and* $\lambda, \lambda_1, \lambda_2$ *be measures on* $\mathcal{M}$. *Then*

1. λ *concentrated on* $A \Rightarrow |\lambda|$ *concentrated on A.*
2. $\lambda_1 \perp \lambda_2 \Rightarrow |\lambda_1| \perp |\lambda_2|$.
3. $\lambda_1 \perp \mu$ *and* $\lambda_2 \perp \mu \Rightarrow \lambda_1 + \lambda_2 \perp \mu$.
4. $\lambda_1 \ll \mu$ *and* $\lambda_2 \ll \mu \Rightarrow \lambda_1 + \lambda_2 \ll \mu$.
5. $\lambda \ll \mu \Rightarrow |\lambda| \ll \mu$.
6. $\lambda_1 \ll \mu$ *and* $\lambda_2 \perp \mu \Rightarrow \lambda_1 \perp \lambda_2$.
7. $\lambda \ll \mu$ *and* $\lambda \perp \mu \Rightarrow \lambda = 0$.

Theorem 32 (Lebesgue's decomposition theorem)**.** *Let* μ *be a positive,* σ*-finite measure and* λ *be a complex measure on* $\mathcal{M}$.
Then there exist unique complex measures λ_a *and* λ_s *on* $\mathcal{M}$ *such that* $\lambda = \lambda_a + \lambda_s$, $\lambda_a \ll \mu$ *and* $\lambda_s \perp \mu$.

Theorem 33 (Theorem of Radon-Nikodym)**.** *Let* μ *be a positive,* σ*-finite measure and* λ *be a complex measure on* $\mathcal{M}$. *Let* λ_a *and* λ_s *be defined as in Theorem 32.*
Than there exists a unique $h \in L^1(\mu)$ *such that* $\lambda_a(E) = \int_E h \, \mathrm{d}\mu$, $\forall E \in \mathcal{M}$.

Proof. (Proof of both Theorems by J. von Neumann). see [15]. □

Corollary 34. *Let* μ *be a positive,* σ*-finite measure,* $\lambda \ll \mu$. *Then there exists a unique* $h \in L^1(\mu)$ *with* $\lambda(E) = \int_E h \, \mathrm{d}\mu$.

Theorem 35 (Polar Decomposition)**.** *Let* μ *be a complex measure. There exists a measurable function h with* $|h(x)| = 1$ *for all* $x \in X$, *such that*

$$\mathrm{d}\mu = h \, \mathrm{d}|\mu|$$

holds.

Proof. Since $\mu \ll |\mu|$ holds Theorem 33 provides $h \in L^1(|\mu|)$ such that

$$\mathrm{d}\mu = h \, \mathrm{d}|\mu|$$

is fulfilled. Define the set $A_r = \left\{x \in X \,\middle|\, |h(x)| < r\right\}$ for $r > 0$ arbitrary but fix. Let $\{E_j\}$ be an arbitrary partitioning of A_r. Than

$$\sum_j |\mu(E_j)| = \sum_j \left| \int_{E_j} h \, \mathrm{d}|\mu| \right| \leq \sum_j r|\mu|(E_j) = r|\mu|(A_r).$$

Taking the supremum over all partitions of A_r returns the total variation measure, which leads to

$$|\mu|(A_r) = \sup_{\{E_j\}\text{partition of } A_r} \sum_j \left|\mu(E_j)\right| \leq r \cdot |\mu|(A_r).$$

Accordingly for $r < 1$ follows that $|\mu|(A_r) = 0$. This means $|h| \geq 1$, $|\mu|-$ almost everywhere.

On the other hand, for $|\mu|(E) > 0$

$$\left| \frac{1}{|\mu|(E)} \int h \, \mathrm{d}|\mu| \right| = \left| \frac{\mu(E)}{|\mu|(E)} \right| \leq 1$$

can be obtained for all $E \in \mathcal{M}$. This leads to $h \leq 1$, $|\mu|$ – a.e..
Overall it follows $|h| = 1$, $|\mu|$ – a.e..
Since $\left\{x \in X \,\middle|\, h(x) \neq 1\right\}$ is a set of $|\mu|$-measure zero, we can set $h(x) = 1$ everywhere on X. □

Remark 36. *Theorem 35 provides existence of a function $h \in L^1(|\mu|)$ such that*

$$\mu(E) = \int_E h \, \mathrm{d}|\mu|$$

holds for all $E \in \mathcal{M}(X)$.

Remark 37. *The results of Theorem 33 and Theorem 35 can be applied to vector valued measures as well, if the Radon-Nikodym-Property is fulfilled. Since the considered functions have bounded variations, this property holds.*

2.5. Subdifferential, Convex Duality and Saddle-Point Problems

In this section we want to derive the needed theory that rewrites the minimisation problem (2.3) as a saddle point problem, to ease computation of solutions. Very detailed theory is given in [3].

Definition 38. *Let X be a normed space and $F : X \to \mathbb{R}_\infty$ be convex. Then $\xi \in X^*$ is called **subgradient** of F at $u \in X$, if*

$$F(u) + \langle \xi, v - u \rangle \leq F(v) \quad \forall v \in X \tag{2.4}$$

holds.
*The mapping $\partial F : X \to \mathcal{P}(X^*)$ that maps $u \in X$ onto the subgradient is called **subdifferential**.*

Remark 39. *It has to be mentioned that Definition 38 is a generalisation of the derivative and therefore ∂F is a multivalued function.*

Theorem 40. *Let X be normed and $F : X \to \mathbb{R}_\infty$ be a convex mapping. Then u^* is a minimiser of F if*

$$0 \in \partial F(u^*)$$

holds.

Proof.

$$\begin{aligned} u^* \text{is a minimiser} &\iff F(u^*) \leq F(u) \quad \forall u \in X \\ &\iff F(u^*) + \langle 0, u^* - u \rangle \leq F(u) \quad \forall u \in X \\ &\iff 0 \in \partial F(u^*). \end{aligned}$$

□

Lemma 41. *Let X, Y be normed spaces, F, $G : X \to \mathbb{R}_\infty$ be proper and convex and let $A : Y \to X$ be linear and continuous. Then*

1. $\partial(\lambda F) = \lambda \partial F \quad \forall \lambda > 0.$
2. $\partial\Big(F(\cdot + u^0)\Big)(u) = (\partial F)(u + u^0) \quad \forall u, u^0 \in X.$

3. $\partial F + \partial G \subset \partial(F+G)$.
   If there exists $u^0 \in \mathrm{dom}F \cap \mathrm{dom}G$ *where* F *is continuous, then*

$$\partial F + \partial G = \partial(F+G)$$

   holds.
4. $A^* \circ \partial F \circ A \subset \partial(F \circ A)$.
   If there exists $u^0 \in \mathrm{dom}F \cap \mathrm{range}(A)$, *then*

$$A^* \circ \partial F \circ A = \partial(F \circ A).$$

   holds.

Proof. see [3], Satz 6.51. □

Definition 42. *Let* X, Y *be vector spaces and* $C \subset X$, $K \subset Y$ *be subsets of these vector spaces. Let* $L : C \times K \to \mathbb{R}$ *be a mapping.*
The pair $(u^*, \phi^*) \in C \times K$ *is called a saddle point if*

$$u^* \in \arg\min_{u \in C} L(u, \phi^*),$$
$$\phi^* \in \arg\max_{\phi \in K} L(u^*, \phi)$$

holds.

Remark 43. *An equivalent formulation of* (u^*, ϕ^*) *being a saddle point, is given by*

$$L(u^*, \phi) \le L(u^*, \phi^*) \le L(u, \phi^*) \quad \forall (u, \phi) \in C \times K.$$

Lemma 44. *The pair* (u^*, ϕ^*) *is a saddle point of* L *if and only if*

$$\min_{u \in C} \sup_{\phi \in K} L(u, \phi) = \max_{\phi \in K} \inf_{u \in C} L(u, \phi) = L(u^*, \phi^*).$$

Proof. see [3], Bemerkung 6.72. □

Remark 45. *According to Lemma 44, for a given saddle point problem we define* $F(u) = \sup_{\phi \in K} L(u, \phi)$ *and* $G(\phi) = \inf_{u \in C} L(u, \phi)$. *Accordingly we define*

$$\min_{u \in C} F(u)$$

*as the **primal problem** and*

$$\max_{\phi \in K} G(\phi)$$

*as the **dual problem** of the saddle point problem according to L.*
We further get the equivalence

$$(u^*, \phi^*) \text{ is a saddle point } \iff \begin{cases} u^* = \arg\min_{u \in C} F(u) \text{ and} \\ \phi^* = \arg\max_{\phi \in K} G(\phi). \end{cases}$$

Definition 46. *Let X be a Banach space and $F : X \to \mathbb{R}_\infty$ be proper. Then the **Fenchel Conjugate** $F^* : X^* \to \mathbb{R}_\infty$ is defined as*

$$F^*(\xi) = \sup_{u \in X} \left(\langle \xi, u \rangle_{X^* \times X} - F(u) \right).$$

Further let $G : X^* \to \mathbb{R}_\infty$ be proper. The Fenchel Conjugate $G^* : X \to \mathbb{R}_\infty$ is defined as

$$G^*(u) = \sup_{\xi \in X^*} \left(\langle \xi, u \rangle_{X^* \times X} - G(\xi) \right).$$

Lemma 47 (Fenchel inequality)**.** *For $u \in X$ and $\xi \in X^*$ we get the inequality*

$$\langle \xi, u \rangle \leq F^*(\xi) + F(u)$$

Proof. see [3], Bemerkung 6.61. □

Lemma 48 (Fenchel equality)**.** *Let $F : X \to \mathbb{R}_\infty$ be proper, convex and lower semi continuous. Then*

$$\begin{aligned} \xi \in \partial F(u) &\iff \langle \xi, u \rangle = F(u) + F^*(\xi) \\ &\iff u \in \partial F^*(\xi). \end{aligned}$$

Proof. If $\langle \xi, u \rangle = F(u) + F^*(\xi)$ holds, we get for $v \in \operatorname{dom} F$ according to Definition 46

$$\begin{aligned} \langle \xi, u \rangle &\geq F(u) + \langle \xi, v \rangle - F(v) \\ &\Rightarrow F(u) + \langle \xi, v - u \rangle \leq F(v) \\ &\Rightarrow \xi \in \partial F(u). \end{aligned}$$

Vice versa for $\xi \in \partial F(u)$ we get

$$\begin{aligned} &\Rightarrow \langle \xi, u \rangle \geq F(u) + \langle \xi, v \rangle - F(v) \;\; v \in \operatorname{dom} F \\ &\Rightarrow \langle \xi, u \rangle \geq F(u) + F^*(\xi) \end{aligned}$$

Nevertheless the Fenchel inequality (Lemma 47) still holds. Therefore, we get

$$\langle \xi, u \rangle = F(u) + F^*(\xi).$$

The second equivalence

$$\langle \xi, u \rangle = F(u) + F^*(\xi) \iff u \in \partial F^*(\xi)$$

follows analogously. □

Lemma 49. *Let X be a Banach space and $F : X \to \mathbb{R}_\infty$ be proper. Then*

1. *For $\lambda \in \mathbb{R}$:*
   $(F + \lambda)^*(\xi) = F^*(\xi) - \lambda$.
2. *For $\lambda > 0$:*
   $(\lambda F)^*(\xi) = \lambda F^*\left(\frac{\xi}{\lambda}\right)$.
3. *For $u^0 \in X$, $\xi^0 \in X^*$*
   $\left(F(\cdot + u^0) + \langle \xi^0, \cdot \rangle\right)^*(\xi) = F^*(\xi - \xi^0) - \langle \xi - \xi^0, u^0 \rangle$.
4. *Let Y be a Banach space and $A : Y \to X$ be linear, continuous and continuous invertible with $A^{-*} = (A^{-1})^*$. Then*
   $(F \circ A)^*(\xi) = F^*(A^{-*}\xi)$.
5. *Let Y be a Banach space and $G : Y \to \mathbb{R}_\infty$. For $H : X \times Y \to \mathbb{R}_\infty$ with $H(u, v) = F(u) + G(v)$ we get*
   $H^*(\xi, \zeta) = F^*(\xi) + G^*(\zeta)$ *for* $(\xi, \zeta) \in X^* \times Y^*$.

Proof. see [3], Lemma 6.65. □

For the further considerations let us specify the minimisation problem (2.3).
Let X, Y be Banach spaces, $A : X \to Y$ be linear and continuous. $F : X \to \mathbb{R}_\infty$, $G : Y \to \mathbb{R}_\infty$ should be proper, convex and lower semi continuous. In imaging the resulting minimisation problem can often be rewritten as

$$\min_{u \in X} F(u) + G(Au). \tag{2.5}$$

Example 50. Given the theory of Fenchel Conjugates the discrete version for TV-Denoising problems can be reconsidered. Therefore let A be a discretisation of ∇. According to equation (2.5) the mappings F and G are given as

$$F(u) = \frac{\lambda}{2}\|u - u^0\|^2,$$
$$G(v) = \|v\|_1.$$

Utilising the Fenchel Conjugate, the Operator G is rewritten as

$$G(Au) = \sup_{\xi \in Y^*} \langle Au, \xi \rangle - G^*(\xi),$$

which transforms (2.5) to the saddle point formulation

$$\min_{u \in X} \sup_{\xi \in Y^*} \langle Au, \xi \rangle + F(u) - G^*(\xi).$$

Let us assume that we are allowed to interchange min and sup, we get for equation (2.5)

$$\begin{aligned}
&\min_{u \in X} F(u) + G(Au) && \text{(primal problem)} \\
=&\min_{u \in X} \sup_{\xi \in Y^*} \langle Au, \xi \rangle + F(u) - G^*(\xi) \\
=&\max_{\xi \in Y^*} \inf_{u \in X} \langle Au, \xi \rangle + F(u) - G^*(\xi) \\
=&\max_{\xi \in Y^*} -F^*(-A^*\xi) - G^*(\xi). && \text{(dual problem)}
\end{aligned}$$

The needed tool to make this interchange feasible is provided in the following theorem.

Theorem 51 (Fenchel-Rockafellar-Duality)**.** *Let X, Y be Banach spaces, $F : X \to \mathbb{R}_\infty$, $G : Y \to \mathbb{R}_\infty$ be proper, convex, lower semi-continuous and $A : X \to Y$ be linear and continuous.*
Let $u^* \in X$ be a solution of

$$\min_{u \in X} F(u) + G(Au).$$

If there exists $u^0 \in \mathrm{dom} F$ with $G(Au^0) < \infty$ and G continuous at Au^0, then

$$\min_{u \in X} F(u) + G(Au) = \max_{\xi \in Y^*} -F^*(-A^*\xi) - G^*(\xi)$$

holds and the maximum is achieved.
A pairing $(u^*, \xi^*) \in X \times Y^*$ *solves the primal-dual problem if, and only if, the conditions*

$$\begin{aligned} -A^*\xi^* &\in \partial F(u^*), \\ \xi^* &\in \partial G(Au^*) \end{aligned}$$

hold.

Proof. Since u^* is a solution of the primal problem, Theorem 40 provides

$$0 \in \partial \left(F + G \circ A\right)(u^*).$$

Lemma 41 further gives

$$\partial \left(F + G \circ A\right)(u^*) = \partial F + A^* \circ \partial G \partial A.$$

Therefore there exists $\xi^* \in Y^*$ with $\xi^* \in \partial G(Au^*)$ and $-A^*\xi^* \in \partial F(u^*)$.
The subdifferential inequality (2.4) provides

$$\begin{aligned} F(u^*) - \langle A^*\xi^*, v - u^* \rangle &\leq F(v) && \forall v \in X \\ \langle -A^*\xi^*, u^* \rangle - F(u^*) &\geq -\langle A^*\xi^*, v \rangle - F(v) && \forall v \in X. \end{aligned}$$

By taking the supremum over all $v \in X$ we get

$$\langle -A^*\xi^*, u^* \rangle - F(u^*) \geq F^*(-A^*\xi^*). \tag{2.6}$$

Analogous arguments for $\xi^* \in \partial G(Au^*)$ give

$$\langle \xi^*, Au^* \rangle - G(Au^*) \geq G^*(\xi^*). \tag{2.7}$$

Adding up the inequalities (2.6) and (2.7) lead to

$$-F(u^*) - G(Au^*) \geq F^*(-A^*\xi^*) + G^*(\xi^*)$$

which implies

$$\min_{u \in X} F(u) + G(Au) \leq -F^*(-A^*\xi^*) - G^*(\xi^*).$$

On the other hand, with Lemma 44 we have

$$\min_{u\in X} F(u) + G(Au) \leq -F^*(-A^*\xi^*) - G^*(\xi^*)$$

and therefore ξ^* solves the dual problem.

For the second part of the proof we make use of the Fenchel equality (Theorem 48) and obtain

$$-A^*\xi^* \in \partial F(u^*) \iff \langle -A^*\xi^*, u^*\rangle = F(u^*) + F^*(-A^*\xi^*),$$

as well as

$$\xi^* \in \partial G(Au^*) \iff \langle \xi^*, Au^*\rangle = G(Au^*) + G^*(\xi^*).$$

Adding these two equations leads to

$$F(u^*) + G(Au^*) = -F^*(-A^*\xi^*) - G^*(\xi^*)$$

and therefore, the pair (u^*, ξ^*) solves the primal-dual problem. □

2.6. A Primal-Dual Algorithm

In this section we want to briefly discuss a recent method to solve minimisation problems as given in (2.5). A detailed presentation of this method is given in [6]. Instead of the standard minimisation formulation

$$\min_{x\in X} F(x) + G(Kx)$$

for F, G convex and K linear, primal-dual algorithms operate on the resulting saddle point formulation

$$\min_{x\in X}\max_{y\in Y} L(x,y) = \min_{x\in X}\max_{y\in Y} \langle Kx, y\rangle + F(x) - G^*(y)$$

as discussed in the previous Subsection.

Definition 52. *Let $L : C_1 \times C_2 \to \mathbb{R}$ be a given saddle point formulation for $C_1 \subset X$, $C_2 \subset Y$ convex.*
For arbitrary $x_0 \in X$, $y_0 \in Y$ define

$$L_{y_0} : X \to \mathbb{R}_\infty, \quad L_{y_0}(x) = \begin{cases} L(x, y_0) \text{ if } x \in C_1, \\ \infty \ else. \end{cases}$$

$$L_{x_0} : Y \to [-\infty, \infty[, \quad L_{x_0}(y) = \begin{cases} L(x_0, y) \text{ if } y \in C_2, \\ -\infty \ else. \end{cases}$$

If L_{y_0} is convex for all $y_0 \in C_2$ and L_{x_0} is concave for all $x_0 \in C_1$ we call

$$\min_{x \in C_1} \max_{y \in C_2} L(x, y)$$

*a **convex-concave saddle point problem**.*

Definition 53. *Let X be a Hilbert space. Let $F : X \to \mathbb{R}_\infty$ be proper, convex and lower semi-continuous.*
The mapping $prox_F : X \to X$ is defined as

$$prox_F(\bar{u}) = \arg\min_{u \in X} \frac{\|u - \bar{u}\|^2}{2} + F(u)$$

*and is called **proximity operator**.*

Remark 54. *$prox_F$ corresponds to the resolvent $(I + \partial F)^{-1}$.*

Lemma 55. *Let L be a convex-concave saddle point problem. Then $(x^*, y^*) \in C_1 \times C_2$ is a saddle point if, and only if*

(i):

$$x^* \in \arg\min_{x \in C_1} L_{y^*}(x),$$
$$y^* \in \arg\min_{y \in C_2} L_{x^*}(y)$$

and

(ii): if $-L_{x^*}$, L_{y^*} are proper, convex and lower semi continuous. For σ, $\tau > 0$

$$\begin{cases} x^* = prox_{\sigma \cdot L_{y^*}}(x^*), \\ y^* = prox_{-\tau \cdot L_{x^*}}(y^*) \end{cases}$$

holds.

Proof. (i): Let (x^*, y^*) be a saddle point. Applying Remark 43 leads to the equivalent formulation

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall (x, y) \in C_1 \times C_2. \tag{2.8}$$

With Definition 52 we get

$$\begin{cases} y^* \in \arg\min_{y \in C_2} -L_{x^*}(y), \\ x^* \in \arg\min_{x \in C_1} L_{y^*}(x), \end{cases}$$

being equivalent to (2.8).

(ii): Since x^* minimises $L_{y^*}(x)$ and y^* minimises $-L_{x^*}(y)$ respectively, we get $0 \in \partial L_{y^*}(x^*)$ and $0 \in \partial\left(-L_{x^*}\right)(y^*)$ as an equivalent formulation, according to Theorem 40. This can be rewritten as

$$\begin{aligned} & \begin{cases} 0 \in \partial L_{y^*}(x^*), \\ 0 \in \partial\left(-L_{x^*}\right)(y^*). \end{cases} \\ \iff & \begin{cases} x^* \in x^* + \sigma \partial L_{y^*}(x^*), \\ y^* \in y^* + \tau \partial\left(-L_{x^*}\right)(y^*). \end{cases} \\ \iff & \begin{cases} x^* \in \left(I + \sigma \partial L_{y^*}\right)(x^*), \\ y^* \in \left(I + \tau \partial\left(-L_{x^*}\right)\right)(y^*). \end{cases} \end{aligned}$$

Finally Remark 54 delivers the equivalent formulation

$$\begin{cases} x^* = prox_{\sigma \cdot L_{y^*}}(x^*), \\ y^* = prox_{-\tau \cdot L_{x^*}}(y^*). \end{cases}$$

□

Corollary 56. *Let X, Y be Banach spaces. Let $F : X \to \mathbb{R}_\infty$, $G : Y \to \mathbb{R}_\infty$ be proper, convex and lower semi-continuous, $K : X \to Y$ be a linear and continuous mapping and let the Fenchel-Rockafellar-Duality (Theorem 51) hold, i.e.*

$$\min_{x \in X} F(x) + G(Ax) = \max_{y \in Y^*} -F^*(-A^*y) - G^*(y).$$

Then $(x^*, y^*) \in \mathrm{dom}F \times \mathrm{dom}G^*$ is a saddle point of $L : \mathrm{dom}F \times \mathrm{dom}G^* \to \mathbb{R}$

$$L(x, y) = \langle Kx, y \rangle + F(x) - G^*(y),$$

if and only if

$$\begin{cases} x^* = prox_{\sigma F}(x^* - \sigma K^* y^*), \\ y^* = prox_{\tau G^*}(y^* + \tau K x^*) \end{cases}$$

for some σ, $\tau > 0$.

Proof. Since $\partial L_{y^*}(x) = K^* y^* + \partial F(x)$ holds according to Lemma 41 we get

$$\begin{aligned} &\iff 0 \in K^* y^* + \partial F(x^*) \\ &\iff x^* \in x^* + \sigma K^* y^* + \sigma \partial F(x^*) \\ &\iff x^* - \sigma K^* y^* \in (I + \sigma \partial F)(x^*) \\ &\iff x^* = prox_{\sigma F}(x^* - \sigma K^* y^*). \end{aligned}$$

With similar arguments we get

$$\partial(-L_{x^*})(y) = -Kx^* + \partial G^*(y) \iff y^* = prox_{\tau G^*}(y^* + \tau K x^*).$$

□

Corollary 56 motivates an iterative way to calculate a minimiser by using the proximity operator as an update function. This leads to the so called Arrow-Hurwicz-Method:

$$\begin{cases} x^{n+1} = prox_{\sigma F}(x^n - \sigma K^* y^n), \\ y^{n+1} = prox_{\tau G^*}(y^n + \tau K x^n). \end{cases} \tag{2.9}$$

Additionally we can make use of x^{n+1} that is calculated in the first step of (2.9). If an extrapolation step is applied, we get the Arrow-Hurwicz-Method with extrapolation:

$$\begin{cases} x^{n+1} = prox_{\sigma F}(x^n - \sigma K^* y^n), \\ \bar{x}^{n+1} = 2x^{n+1} - x^n, \\ y^{n+1} = prox_{\tau G^*}(y^n + \tau K \bar{x}^{n+1}). \end{cases} \tag{2.10}$$

For the Arrow-Hurwicz-Method with extrapolation given in (2.10) a convergence Theorem can be shown.

Theorem 57. *Let X, Y be a Hilbert space, $K : X \to Y$ linear and continuous and $F : X \to \mathbb{R}_\infty$, $G : Y \to \mathbb{R}_\infty$ proper, convex and lower semi continuous.*
Let $L : \mathrm{dom}F \times \mathrm{dom}G^* \to \mathbb{R}$

$$L(x, y) = \langle Kx, y \rangle + F(x) - G^*(y)$$

have atleast one saddle point.
Then Iteration (2.10) *converges for each pair $(x^0, y^0) \in X \times Y^*$ to a saddle point of L in a weak sense, if $\sigma\tau\|K\|^2 < 1$ holds.*

Proof. see [3], Satz 6.141. □

3. Functional Lifting in the Context of Images and Image Sequences

In this chapter we introduce a regularisation approach that operates on curvature information of the image gradient. The fundamentals of this idea are given by Citti and Sarti [7] and were applied to 2D imaging problems in [4], as already mentioned in Chapter 1.
The idea, presented in [7], is to lift the two dimensional image gradient to the three-dimensional so-called rototranslation space, where the third dimension corresponds to the direction of the image gradient. A visualisation of this idea is given in Figure 3.1.

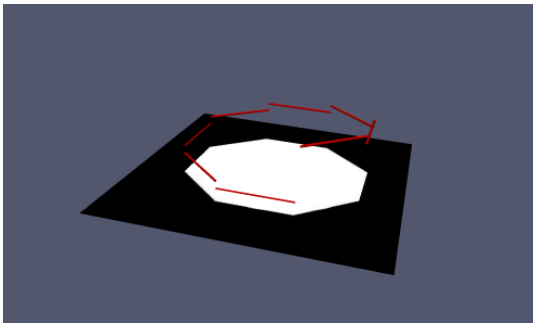


Figure 3.1.: Lifting of a binary octagon to the 3D rototranslation space, mentioning that the third dimension represents the gradient direction. Picture from [4].

In this chapter we want to extend this lifting approach to images in higher dimensions. Further we want to treat moving images, which implies that time is always one dimension of the problem. Therefore image sequences and moving 3D images are objects of interest. It has to be mentioned that time dependent problems have to be treated with special care, since object behaviour is different along the time axis in comparison to its spatial dimensions.

Very often we will refer to the ideas given in [4], in order to mention which concepts are the same and where we have to adapt the existing theory to make this approach working for higher dimensional problems.

3.1. Introducing a Functional Lifting approach concerning Velocities

First let us introduce the space-time unit sphere

$$S^n := \left\{ x = (\tilde{x}, x_t) \in \mathbb{R}^{n-1} \times \mathbb{R} \middle| \|\tilde{x}\|_2^2 + x_t^2 = 1 \right\}.$$

Since we are dealing with space-time problems it is useful to split the unit sphere in its positive and negative time components. To deal with the case $t = 0$, we define this split in a recursive manner.

$$\begin{aligned} S_+^1 &:= \{1\}, \\ S_+^n &:= \left\{ x = (\tilde{x}, x_t) \in S^n \middle| x_t > 0 \vee \tilde{x} \in S_+^{n-1} \right\}, \\ S_-^n &:= S^n \setminus S_+^n. \end{aligned}$$

Remark 58. *By definition we have for $\tilde{x} \in S_+^{n-1}$ and $x \in S_+^n$ that $x_t = 0$.*

For an example of these sets see Figure 3.2.

Here we deviate from the ideas presented in [4]. There the gradient of an image supported on $\Omega \subset \mathbb{R}^2$, was lifted in the space $\Omega \times S^1$. The straight-forward approach for images supported on $\Omega \subset \mathbb{R}^n$, with $n \geq 3$ would be, to lift the gradient in the space $\Omega \times S^{n-1}$. A more sensible way to perform functional lifting for time dependent imaging problems can be achieved by introducing the quotient space

$$\mathbb{P}^n := S^n / \{-1, 1\} = \left\{ \{x, -x\} \middle| x \in S_+^n \right\}.$$

By using this quotient space it is possible to identify edge movement in one direction independent from the contrast.
Consequently we define a mapping between S^n and $\mathbb{P}^n$ and vice versa.

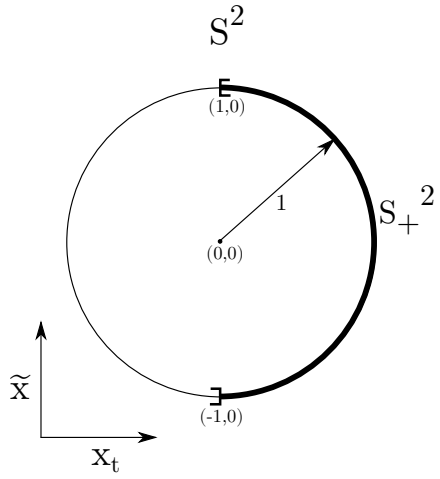


Figure 3.2.: Plot of the unit sphere S^2 and the subset S^2_+

Definition 59. *Let S^n and $\mathbb{P}^n$ be defined as above. As a mapping between these sets we define*

$$\begin{aligned} j :& S^n \to \mathbb{P}^n \\ & x \mapsto \{x, -x\} \end{aligned}$$

as well as its right inverse

$$\begin{aligned} j^+ :& \mathbb{P}^n \to S^n_+ \\ & \{x, -x\} \mapsto \begin{cases} x & \text{if } x \in S^n_+ \\ -x & \text{if } x \in S^n_-. \end{cases} \end{aligned} \quad .$$

Remark 60. *With this setting we have $j \circ j^+ = id_{\mathbb{P}^n}$.*

Further we need to introduce a location function to decide on which sub sphere of the space-time sphere a vector $x \in S^n$ is located. For this purpose, we define

$$\begin{aligned} \text{sign} :& S^n \to \{-1, 1\} \\ & x \mapsto \begin{cases} 1 & \text{if } x \in S^n_+ \\ -1 & \text{if } x \in S^n_-. \end{cases} \end{aligned} \tag{3.1}$$

Lemma 61. *For each $x \in S^n$ the property*

$$\operatorname{sign}(x) \cdot j^+(j(x)) = x$$

can be verified.

Proof. If $x \in S^n_+$ we get

$$\operatorname{sign}(x) \cdot j^+(j(x)) = 1 \cdot j^+(\{x,-x\}) = 1 \cdot x = x.$$

On the other hand $x \in S^n_-$ leads to

$$\operatorname{sign}(x) \cdot j^+(j(x)) = -1 \cdot j^+(\{x,-x\}) = -1 \cdot -x = x.$$

□

Definition 62. *Let $u \in BV(\Omega)$, $|\nabla u| \in \mathcal{M}(\Omega)$ be the total variation of $\nabla u \in \mathcal{M}(\Omega, \mathbb{R}^n)$ and $\sigma \in L^\infty_{|\nabla u|}(\Omega, \mathbb{R}^n)$ be the density of ∇u w.r.t. $|\nabla u|$. The measure $\mu \in \mathcal{M}(\Omega \times \mathbb{P}^n)$ is the **Functional Lifting** for ∇u if*

$$\int_{\Omega \times \mathbb{P}^n} \varphi(x,\tau)\, d\mu(x,\tau) = \int_\Omega \operatorname{sign}(\sigma(x)) \cdot \varphi\left(x, j(\sigma(x))\right)\, d\,|\nabla u|$$

holds for all $\varphi \in C_0(\Omega \times \mathbb{P}^n)$ with $C_0(\Omega \times \mathbb{P}^n)^* = \mathcal{M}(\Omega \times \mathbb{P}^n)$.

Remark 63. *In this setting μ is not necessarily a positive measure, what deviates from the Definition given in [4].*

Remark 64. *With Theorem 35 it follows that $\sigma(x) \in S^n$, almost everywhere on $|\nabla u|$.*

Remark 65. *Let μ be the Functional Lifting of ∇u. Then μ indeed preserves the gradient of u. This can be verified by setting $\varphi(x,\tau) = \psi(x) \cdot j^+(\tau)$ for $\psi \in C_0(\Omega \times \mathbb{P}^n, \mathbb{R}^n)$. Definition 62 leads to*

$$\begin{aligned}
\int_{\Omega \times \mathbb{P}^n} \varphi(x,\tau)\, d\mu(x,\tau) &= \int_\Omega \operatorname{sign}(\sigma(x)) \cdot \varphi\left(x, j(\sigma(x))\right)\, d\,|\nabla u| \\
&= \int_\Omega \operatorname{sign}(\sigma(x)) \cdot \psi(x) \cdot j^+(j(\sigma(x)))\, d\,|\nabla u| \\
&= \int_\Omega \psi(x) \cdot \sigma(x) d\,|\nabla u| \\
&= \int_\Omega \psi(x)\, d\,\nabla u.
\end{aligned}$$

3.2. Vertex Penalisation

To penalise changes in the direction along the gradient of u we have to think of a way to calculate a directional derivative in the lifted setting. The directional derivative has to be calculated at each point of the domain Ω in every possible direction of $\mathbb{P}^n$. We introduce the tangential plane

$$T_\vartheta := \left\{ t \in \mathbb{R}^n \,\middle|\, t \perp j^+(\vartheta) \right\},$$

for each direction $\vartheta \in \mathbb{P}^n$ (see Figure 3.3). This hyperplane is of dimension $n-1$ and can be described locally for a fixed ϑ with an orthonormal basis $t_1, \ldots, t_{n-1} \in \mathbb{R}^n$. Accordingly $(j^+(\vartheta), t_1, \ldots, t_{n-1})$ forms a orthonormal basis on $\mathbb{R}^n$.
The effort to describe this tangential plane is one of the main novelties compared to the work done in [4] on two dimensional objects, where a simple counterclockwise rotation of ϑ is enough.

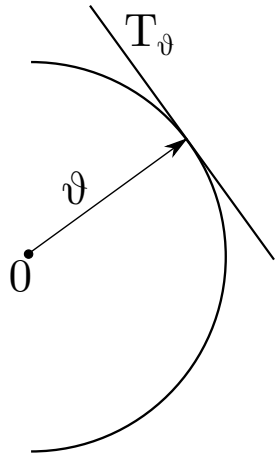


Figure 3.3.: Tangential plane T_ϑ for an arbitrary direction $\vartheta \in \mathbb{P}^2$

Now we are able to define the derivative of a function $\varphi \in C_0(\Omega \times \mathbb{P}^n)$ in a fixed point (x, ϑ) and a fixed direction t_i in the tangential hyperplane.

Definition 66. *Let $(x, \vartheta) \in (\Omega \times \mathbb{P}^n)$ be fixed, $t_1, \ldots, t_{n-1} \in \mathbb{R}^n$ be an orthonormal basis of the hyperplane T_ϑ and $\{j^+(\vartheta), t_1, \ldots, t_{n-1}\}$ be the used orthonormal basis of $\mathbb{R}^n$. Further let $R_\vartheta : \mathbb{R}^n \to \mathbb{R}^n$ be a rotation matrix that maps the vector $(1, 0, \ldots, 0)^t$*

onto $j^+(\vartheta)$. *Then the directional derivative for elements* $\varphi \in C_0(\Omega \times \mathbb{P}^n)$ *in direction* t_i *can be written as*

$$\frac{\partial \varphi}{\partial t_i}(x,\vartheta) = \lim_{s \to 0} \frac{\varphi(x, j(R_\vartheta\, \hat{t}(s,i))) - \varphi(x,\vartheta)}{s},$$

where $\hat{t}(s,i) \in \mathbb{R}^n$ *is defined as*

$$\hat{t}_k(s,i) := \begin{cases} \cos(s) & \text{if } k = 1 \\ \sin(s) & \text{if } k = i \\ 0 & \text{else.} \end{cases}$$

A graphical interpretation of Definition 66 is shown in Figure 3.4.

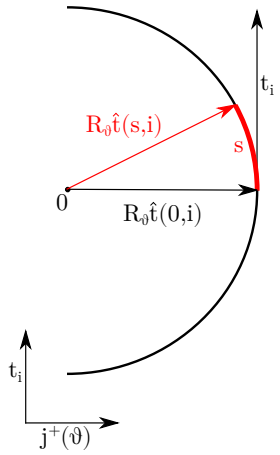


Figure 3.4.: Interpretation of the directional derivative along the vector t_i.

Definition 67. *Using Definition 66 we can introduce the directional derivative* $\nabla_{T_\vartheta}\varphi(x,\vartheta)$, *for* $\varphi : \Omega \times \mathbb{P}^n \mapsto \mathbb{R}$, *as a linear functional mapping*

$$\nabla_{T_\vartheta}\varphi(x,\vartheta) : T_\vartheta \mapsto \mathbb{R}$$

for a fixed point $(x,\vartheta) \in \Omega \times \mathbb{P}^n$. *In terms of the orthonormal basis vectors we introduce*

$$\nabla_{T_\vartheta}\varphi(x,\vartheta)\,(v) = \sum_{i=1}^{n-1} \frac{\partial \varphi}{\partial t_i}(x,\vartheta)(t_i(\vartheta) \cdot v)$$

for $\nu \in T_\vartheta$.
Furthermore, we need a corresponding way to calculate the divergence in this tangential plane, since differentiation in a weak sense is needed, when variational methods are used. For this purpose see the theory derived in Chapter 2.
For a vectorfield $\psi : \Omega \times \mathbb{P}^n \to \mathbb{R}^n$ s.t. $\psi(x,\vartheta) \in T_\vartheta$ for all $(x,\vartheta) \in \Omega \times \mathbb{P}^n$ we define

$$\operatorname{div}_{T_\vartheta}\psi(x,\vartheta) := \sum_{i=1}^{n-1} \frac{\partial \psi}{\partial t_i}(x,\vartheta) \cdot t_i. \tag{3.2}$$

Consequently the derivative $\frac{\partial}{\partial t_i}$ is applied in each component of ψ, thus Definition 66 still makes sense.

Next we show that the definitions of gradient and divergence are consistent by showing that a divergence theorem can be proven in this setting. Since we integrate on the domain $\Omega \times \mathbb{P}^n$ we declare

$$\int_{\Omega\times\mathbb{P}^n} \varphi(x,\vartheta)\,\mathrm{d}(\lambda^n \otimes \mathcal{H}^{n-1}) := \frac{1}{2}\int_{\Omega\times S^n} \varphi(x,j(t))\,\mathrm{d}(\lambda^n \otimes \mathcal{H}^{n-1})$$

for an integrable function φ.

Lemma 68. *For $\varphi \in C^1_C(\Omega \times \mathbb{P}^n)$ and $\psi \in C^1(\Omega \times \mathbb{P}^n, \mathbb{R}^n)$ s.t. $\psi(x,\vartheta) \in T_\vartheta$ the divergence theorem can be rewritten as*

$$\int_{\Omega\times\mathbb{P}^n} \nabla_{T_\vartheta}\varphi(x,\vartheta)\cdot\psi(x,\vartheta)\,\mathrm{d}(x,\vartheta) = -\int_{\Omega\times\mathbb{P}^n} \operatorname{div}_{T_\vartheta}\psi(x,\vartheta)\cdot\varphi(x,\vartheta)\,\mathrm{d}(x,\vartheta)$$

Proof. We compute

$$\begin{aligned}
\int_{\Omega\times\mathbb{P}^n} \nabla_{T_\vartheta}\varphi(x,\vartheta)\cdot\psi(x,\vartheta)\mathrm{d}(x,\vartheta) &= \int_{\Omega\times\mathbb{P}^n} \sum_{i=1}^{n-1} \frac{\partial\varphi}{\partial t_i}(x,\vartheta)(t_i(\vartheta)\cdot\psi(x,\vartheta))\mathrm{d}(x,\vartheta) \\
&= \sum_{i=1}^{n-1}\int_{\Omega\times\mathbb{P}^n} \frac{\partial\varphi}{\partial t_i}(x,\vartheta)(t_i(\vartheta)\cdot\psi(x,\vartheta))\mathrm{d}(x,\vartheta) \\
&= -\sum_{i=1}^{n-1}\int_{\Omega\times\mathbb{P}^n} \frac{\partial\psi}{\partial t_i}(x,\vartheta)(t_i(\vartheta)\cdot\varphi(x,\vartheta))\mathrm{d}(x,\vartheta)
\end{aligned}$$

$$= -\int_{\Omega\times\mathbb{P}^n} \sum_{i=1}^{n-1} \frac{\partial\psi}{\partial t_i}(x,\vartheta)(t_i(\vartheta)\cdot\varphi(x,\vartheta))\mathrm{d}(x,\vartheta)$$
$$= -\int_{\Omega\times\mathbb{P}^n} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\cdot\varphi(x,\vartheta)\mathrm{d}(x,\vartheta).$$

□

The next step is to apply Lemma 68 to the lifted measure μ and define the directional derivative in a weak sense. This will make the curvature information accessible in the lifted setting.

Definition 69. *Let $\mu \in \mathcal{M}(\Omega\times\mathbb{P}^n)$ be the Functional Lifting of ∇u. We call $\eta \in \mathcal{M}(\Omega\times\mathbb{P}^n,\mathbb{R}^n)$ the **directional derivative of μ in the weak sense** if the polar decomposition $\eta = \sigma_\eta\cdot|\eta|$ obeys $\sigma_\eta(x,\vartheta)\in T_\vartheta$, $|\eta|-$ a.e. and*

$$\int_{\Omega\times\mathbb{P}^n} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\mathrm{d}\mu = -\int_{\Omega\times\mathbb{P}^n} \psi(x,\vartheta)\cdot\sigma_\eta(x,\vartheta)\mathrm{d}|\eta|(x,\vartheta)$$

for all $\psi(x,\vartheta)\in C_C^1(\Omega\times\mathbb{P}^n,\mathbb{R}^n)$ s.t. $\psi(x,\vartheta)\in T_\vartheta$.
We will denote the weak directional derivative with $\eta = \nabla_{T_\vartheta}\mu$.

To get an idea how the weak derivative on the functional lifting works, we look at a basic 3D example.
Define a binary image where a vertical line distinguishes regions that are equal to zero and equal to one, respectively. With increasing time, this line moves right with a constant velocity v_1. At the time $x_t = 0$ the velocity should change to v_2. Illustrations of this example are given in Figure 3.5 and more detailed in Figure 3.6.

Example 70. Let the domain be the unit cube $\Omega = [-1,1]^3$ and the image be $u(x_x,x_y,x_t) = u_1(x_x,x_y,x_t) + u_2(x_x,x_y,x_t)$ with

$$u_1(x_x,x_y,x_t) = \begin{cases} 1 & \text{if } x_x \geq v_1\cdot x_t \ \wedge \ x_t \leq 0 \\ 0 & \text{else} \end{cases}$$

$$u_2(x_x,x_y,x_t) = \begin{cases} 1 & \text{if } x_x \geq v_2\cdot x_t \ \wedge \ x_t > 0 \\ 0 & \text{else} \end{cases}$$

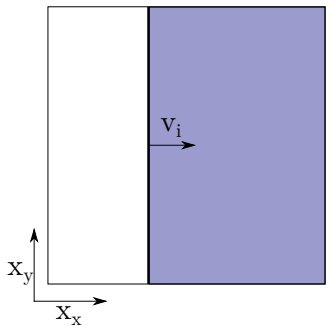


Figure 3.5.: Sketch of a simple binary image sequence. Velocity changes at $x_t = 0$.

and $v_1, v_2 > 0$ (see Figure 3.6).

This image leads to the normal vector

$$\sigma = \begin{cases} \sigma_1 & \text{if } x_t \leq 0 \\ \sigma_2 & \text{else} \end{cases}$$

with $\sigma_1 := \frac{1}{\sqrt{1+v_1^2}}(1,0,-v_1)^t$ and $\sigma_2 := \frac{1}{\sqrt{1+v_2^2}}(1,0,-v_2)^t$. Further we want to define

$$\vartheta_1 := j(\sigma_1) = \left\{ \frac{1}{\sqrt{1+v_1^2}}(-1,0,v_1)^t, \frac{-1}{\sqrt{1+v_1^2}}(-1,0,v_1)^t \right\}$$

and

$$\vartheta_2 := j(\sigma_2) = \left\{ \frac{1}{\sqrt{1+v_2^2}}(-1,0,v_2)^t, \frac{-1}{\sqrt{1+v_2^2}}(-1,0,v_2)^t \right\}$$

as the corresponding objects in the quotient space $\mathbb{P}^n$.

The Lifting μ of ∇u can then be rewritten as

$$\begin{aligned} \int_{\Omega\times\mathbb{P}^3} \varphi \,\mathrm{d}\mu = &\int_{-1}^{1}\int_{-1}^{0} -\varphi\Big((v_1 x_t, x_y, x_t), \vartheta_1\Big) \cdot \sqrt{1+v_1^2}\,\mathrm{d}x_t\,\mathrm{d}x_y \\ &+ \int_{-1}^{1}\int_{0}^{1} -\varphi\Big((v_2 x_t, x_y, x_t), \vartheta_2\Big) \cdot \sqrt{1+v_2^2}\,\mathrm{d}x_t\,\mathrm{d}x_y \end{aligned}$$

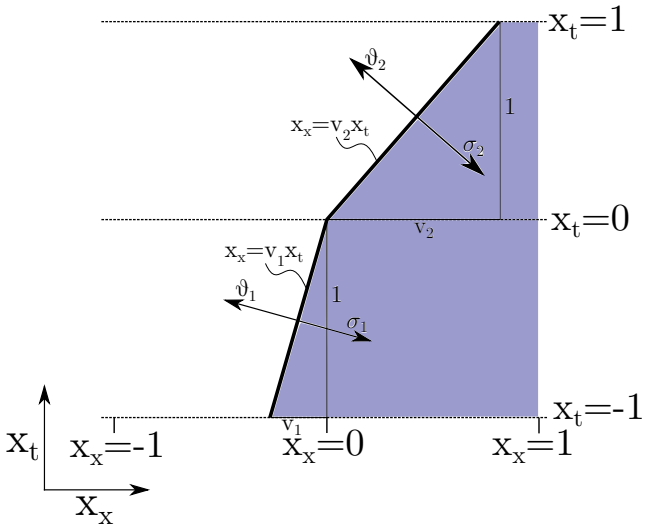


Figure 3.6.: Sketch of u, used at Example 70, in the $x_x - x_t$ plane (shaded area is equal to 1). With this setting we have $sign(\sigma_i) = -1$ with $i = \{1, 2\}$.

using the Jacobian determinant of the transformation $f : (x_y, x_t) \mapsto (v_i x_t, x_y, x_t)$.

$$\mathcal{J}f = \sqrt{\det\left((Df)^t \cdot Df\right)} = \sqrt{\det\begin{pmatrix} 1 & v_i^2 + 1 \\ 0 & v_i^2 + 1 \end{pmatrix}} = \sqrt{1 + v_i^2}.$$

For $\psi \in C_C^1(\Omega \times \mathbb{P}^3, \mathbb{R}^3)$ the directional derivative of μ along the tangential plane T_ϑ gives

$$\begin{aligned} \int_{\Omega\times\mathbb{P}^3} \psi(x,\vartheta)\, \mathrm{d}\nabla_{T_\vartheta}\mu &\overset{\text{Def.69}}{=} -\int_{\Omega\times\mathbb{P}^3} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\, \mathrm{d}\mu \\ &\overset{(3.2)}{=} -\int_{\Omega\times\mathbb{P}^3} \sum_{i=1}^{n-1} \frac{\partial\psi}{\partial t_i} \cdot t_i \, \mathrm{d}\mu \end{aligned}$$

and leads to

$$= \sqrt{1+v_1^2} \int_{-1}^{1} \int_{-1}^{0} \frac{\partial \psi}{\partial t_1^1}\Big((v_1 x_t, x_y, x_t), \vartheta_1\Big) \cdot t_1^1 + \frac{\partial \psi}{\partial t_1^2}\Big((v_1 x_t, x_y, x_t), \vartheta_1\Big) \cdot t_2^1 \, \mathrm{d}x_t \, \mathrm{d}x_y$$
$$+ \sqrt{1+v_2^2} \int_{-1}^{1} \int_{0}^{1} \frac{\partial \psi}{\partial t_2^1}\Big((v_2 x_t, x_y, x_t), \vartheta_2\Big) \cdot t_1^2 + \frac{\partial \psi}{\partial t_2^2}\Big((v_2 x_t, x_y, x_t), \vartheta_2\Big) \cdot t_2^2 \, \mathrm{d}x_t \, \mathrm{d}x_y.$$

Note that the subscript of the vectors t_i^j refers to the two different velocities and the superscript identifies the two basis vectors of the tangential planes T_{ϑ_i}.

We set $t_i^1 = (0,1,0)^t$ and $t_i^2 = \frac{1}{\sqrt{1+v_i^2}}(v_i, 0, 1)^t$ as ONB of the tangential plane. Using the theorem of Fubini, fundamental theorem of calculus and $\psi \in C_C^1(\Omega \times \mathbb{P}^3, \mathbb{R}^3)$ we get

$$\int_{-1}^{1} \frac{\partial \psi}{\partial t_1^1}\Big((v_1 x_t, x_y, x_t), \vartheta_1\Big) \cdot t_1^1 \mathrm{d}x_y = 0.$$

With $\tilde{\psi}(x_t, x_y) := \psi\left(\left(\frac{v_1}{\sqrt{1+v_1^2}} x_t, x_y, \frac{1}{\sqrt{1+v_1^2}} x_t\right), \vartheta_1\right) \cdot t_1^2$ and a transformation of variables we get for the second term

$$\int_{-1}^{1} \int_{-1}^{0} \frac{\partial \psi}{\partial t_1^2}\Big((v_1 x_t, x_y, x_t), \vartheta_1\Big) \cdot t_2^1 \cdot \sqrt{1+v_1^2} \, \mathrm{d}x_t \, \mathrm{d}x_y$$
$$= \int_{-1}^{1} \int_{-\sqrt{1+v_1^2}}^{0} \frac{\partial \psi}{\partial t_1^2}\left(\left(\frac{v_1}{\sqrt{1+v_1^2}} x_t, x_y, \frac{1}{\sqrt{1+v_1^2}} x_t\right), \vartheta_1\right) \cdot t_1^2 \, \mathrm{d}x_t \, \mathrm{d}x_y$$
$$= \int_{-1}^{1} \int_{-\sqrt{1+v_1^2}}^{0} \frac{\partial \tilde{\psi}}{\partial x_t}(x_t, x_y) \, \mathrm{d}x_t \, \mathrm{d}x_y.$$

Using again the fundamental theorem of calculus and that $\tilde{\psi}$ is continuous and compactly supported on the transformed domain we observe

$$= \int_{-1}^{1} \tilde{\psi}(0, x_y) \, \mathrm{d}x_y = \int_{-1}^{1} \psi\Big((0, x_y, 0), \vartheta_1\Big) \cdot t_1^2 \, \mathrm{d}x_y.$$

The calculation of $\int_{-1}^{1}\int_{-1}^{0} \frac{\partial\psi}{\partial t_1^2} t_1^2 + \frac{\partial\psi}{\partial t_2^2} t_2^2 \,\mathrm{d}t\,\mathrm{d}y$ follows analogously.
Summing up finally leads to

$$\int_{\Omega\times\mathbb{P}^3} \psi(x,\vartheta)\,\mathrm{d}\nabla_{T_\vartheta}\mu = \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_1\Big)\cdot t_1^2\,\mathrm{d}x_y - \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_2\Big)\cdot t_2^2\,\mathrm{d}x_y.$$

In view of the polar decomposition $\mathrm{d}\nabla_{T_\vartheta}\mu = \sigma_\eta \mathrm{d}|\eta|$ we obtain for this example

$$\begin{aligned} |\eta| =& \mathcal{H}^1 \llcorner \Big(\{0\}\times\{[-1,1]\}\times\{0\}\times\{\vartheta_1,\vartheta_2\}\Big) \\ \sigma_\eta((x_x,x_y,x_t),\vartheta) =& \begin{cases} \frac{1}{\sqrt{1+v_1^2}}(v_1,0,1)^t \;\text{ for }\; x_x=0,\; x_y\in[-1,1],\; x_t=0,\; \vartheta=\vartheta_1 \\ \frac{1}{\sqrt{1+v_2^2}}(v_2,0,1)^t \;\text{ for }\; x_x=0,\; x_y\in[-1,1],\; x_t=0,\; \vartheta=\vartheta_2. \end{cases} \end{aligned}$$

Example 70 leads to the result, that the directional derivative is only supported along the edge where the vector ϑ changes its direction. Since this is the behaviour we desired, we are able to think of a method to penalise the directional derivative. The idea is to take the supremum of a subset of all test functions $\psi \in C_C^1(\Omega\times\mathbb{P}^n,\mathbb{R}^n)$ such that the result of

$$\int_{\Omega\times\mathbb{P}^n} \psi(x,\vartheta)\,\mathrm{d}\nabla_{T_\vartheta}\mu$$

measures the change of ϑ of the lifted image.

Definition 71.

$$T_\rho(\mu) = \sup_{\psi\in M_\rho(\Omega)} \int_{\Omega\times\mathbb{P}^n} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\mathrm{d}\mu$$

with

$$\begin{aligned} M_\rho(\Omega) =& \Big\{\psi\in C_C^1(\Omega\times\mathbb{P}^n,\mathbb{R}^n)\,\Big|\,\mathrm{div}_{T_\vartheta}\psi\in C_C(\Omega\times\mathbb{P}^n),\psi(x,\cdot)\in C_\rho\;\;\forall x\in\Omega \\ & \psi(x,\vartheta)\in T_\vartheta,\;\forall\vartheta\in\mathbb{P}^n\Big\} \end{aligned}$$

Definition 71 is, up to the different test space of ψ and the resulting application of the divergence, consistent with the theory presented in [4].
How the directional derivative gets penalised is defined by the set C_ρ. In [4] the set C_ρ is generated by a lower semi-continuous metric ρ and rewrites as

$$C_\rho = \Big\{\varphi\in C(\mathbb{P}^n)\,\Big|\,\varphi(\vartheta_1)-\varphi(\vartheta_2)\le\rho(\vartheta_1,\vartheta_2),\;\forall(\vartheta_1,\vartheta_2)\in\mathbb{P}^n\times\mathbb{P}^n\Big\}.$$

This approach is not applicable in the higher dimensional setting any more since here the elements of C_ρ are mapping into the space $\mathbb{R}^n$. In [4] two special cases of the used metric ρ are investigated. They show that for these special cases it is possible to rephrase the set C_ρ in a way that only a special norm of φ has to be restricted in the right manner. This motivates to extend the norms used in [4] to vector valued functions.

Before we investigate these two special cases for C_ρ let us give some further definitions since they will be used in the upcoming sections.

Definition 72. *For any function $\zeta : \mathbb{P}^n \to \mathbb{R}^n$ we define the norm*

$$\|\zeta\|_{\infty,p} = \sup_{\vartheta \in \mathbb{P}^n} |\zeta(\vartheta)|_p.$$

Definition 73. *For any two elements $\vartheta_1, \vartheta_2 \in \mathbb{P}^n$ we define its distance as*

$$d_{\mathbb{P}^n}(\vartheta_1, \vartheta_2) = \min_{\tilde{\vartheta}_1 \in \vartheta_1, \tilde{\vartheta}_2 \in \vartheta_2} \|\tilde{\vartheta}_1 - \tilde{\vartheta}_2\|_{\mathbb{R}^n}$$

Definition 74. *For $\vartheta_1, \vartheta_2 \in \mathbb{P}^n$ we define the geodesic metric as*

$$\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) = \inf_{\vartheta \in \Theta} \int_0^1 \left|\vartheta'(s)\right| \, \mathrm{d}s$$

for elements $(\vartheta_1, \vartheta_2) \in \mathbb{P}^n \times \mathbb{P}^n$ and

$$\Theta := \{\vartheta(s) : [0,1] \to \mathbb{P}^n \,|\, \vartheta(0) = \vartheta_1 \,,\, \vartheta(1) = \vartheta_2 \,, \vartheta \in C \,, \vartheta' \in C \text{ a.e.}\}\,.$$

This is actually a metric, which is verified in the next Lemma.
To get a clearer structure we will use the sets Θ_{12}, Θ_{23} and Θ_{13} which have the same properties as Θ, but with suitable boundary conditions for $\vartheta(0)$ and $\vartheta(1)$.

Lemma 75. *The mapping $\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) = \inf_{\vartheta \in \Theta_{12}} \int_0^1 \left|\vartheta'(s)\right| \, \mathrm{d}s$ is a metric on $\mathbb{P}^n$.*

Proof. We verify
(1) Positive Definiteness:
For any pair $(\vartheta_1, \vartheta_2) \in \mathbb{P}^n \times \mathbb{P}^n$ we have

$$\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) = \inf_{\vartheta \in \Theta_{12}} \int_0^1 \underbrace{\left|\vartheta'(s)\right|}_{\geq 0} \, \mathrm{d}s \geq 0.$$

For $\vartheta_1 = \vartheta_2$ we can choose $\tilde{\vartheta}(s) \equiv \vartheta_1 \in \Theta_{12}$ for all $s \in [0,1]$ and we get

$$0 \leq \mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_1) = \inf_{\vartheta \in \Theta_{12}} \int_0^1 \left|\vartheta'(s)\right| \mathrm{d}s \leq \int_0^1 \underbrace{\left|\tilde{\vartheta}'(s)\right|}_{=0} \mathrm{d}s = 0,$$

providing $\vartheta_1 = \vartheta_2 \Rightarrow \mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_1) = 0$.
Vice versa we claim $\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) = 0$ and take for this pair $(\vartheta_1, \vartheta_2)$ a sequence $\vartheta_n \in \Theta_{12}$ with

$$\lim_{n\to\infty} \int_0^1 |\vartheta_n'(s)| \mathrm{d}s = \inf_{\vartheta \in \Theta} \int_0^1 |\vartheta'(s)| \mathrm{d}s = 0.$$

By making use of

$$\left| \int_0^1 \vartheta_n'(s) \mathrm{d}s \right| \leq \int_0^1 |\vartheta_n'(s)| \mathrm{d}s \to 0,$$

we get $\lim_{n\to\infty} \int_0^1 \vartheta_n'(s) \mathrm{d}s = 0$ and obtain that the fundamental theorem of calculus leads to

$$\vartheta(1) - \vartheta(0) = \int_0^1 \vartheta_n'(s) \mathrm{d}s$$

which can be rewritten as

$$\vartheta_2 = \vartheta_1 + \int_0^1 \vartheta_n'(s) \mathrm{d}s.$$

By taking the limit we get

$$\vartheta_2 = \vartheta_1 + \underbrace{\lim_{n\to\infty} \int_0^1 \vartheta_n'(s) \mathrm{d}s}_{=0} = \vartheta_1.$$

(2) Symmetry:
Symmetry follows immediately, since for every $\vartheta \in \Theta_{12}$ we can define $\tilde{\vartheta}(s) := \vartheta(1-s)$ with $\tilde{\vartheta} \in \Theta_{21}$. Since the equality

$$\int_0^1 |\vartheta'(s)| \mathrm{d}s = \int_0^1 |\tilde{\vartheta}'(s)| \mathrm{d}s.$$

holds, we have $\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) = \mathrm{d}_{\mathrm{geo}}(\vartheta_2, \vartheta_1)$.
(3) Triangle Inequality:
Let $\vartheta_{12}^n \in \Theta_{12}$ be a sequence with

$$\lim_{n\to\infty} \int_0^1 |\vartheta_{12}^{n'}(s)| \, \mathrm{d}s = \mathrm{d}(\vartheta_1, \vartheta_2)$$

and choose $\vartheta_{23}^n \in \Theta_{23}$ analogously.
Now we define $\vartheta_{13}^n \in \Theta_{13}$ as

$$\vartheta_{13}^n(s) := \begin{cases} \vartheta_{12}^n(2s) & \text{for } s \in [0, \frac{1}{2}] \\ \vartheta_{23}^n(2s-1) & \text{for } s \in [\frac{1}{2}, 1]. \end{cases}$$

With this setting we can verify the inequality

$$\begin{aligned} \mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_3) &= \inf_{\vartheta \in \Theta_{13}} \int_0^1 \left|\vartheta'(s)\right| \, \mathrm{d}s \\ &\leq \lim_{n\to\infty} \int_0^1 \left|\vartheta_{13}^{n'}(s)\right| \, \mathrm{d}s \\ &= \lim_{n\to\infty} \int_0^1 |\vartheta_{12}^{n'}(s)| \, \mathrm{d}s + \lim_{n\to\infty} \int_0^1 |\vartheta_{23}^{n'}(s)| \, \mathrm{d}s \\ &= \mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) + \mathrm{d}_{\mathrm{geo}}(\vartheta_2, \vartheta_3). \end{aligned}$$

□

Now we want to look at some variations of C_ρ and calculate how they apply to Example 70.

3.2.1. Restriction to a Set of Functionals that measures Edge Length

In this subsection we want to extend the ideas presented in [4, Example 3.7], where the discrete metric was used to generate the set C_ρ, to a higher dimensional setting. As a reasonable generalisation to vector valued functions we define

$$C_0 := \left\{ \psi(x, \cdot) \in C(\mathbb{P}^n, \mathbb{R}^n) \,\middle|\, \|\psi(x, \cdot)\|_{\infty,2} \leq \frac{1}{2} \right\}.$$

With this setting we get for Example 70

$$\begin{aligned} T_0(\mu) &= \sup_{\psi \in M_0(\Omega)} \int_{\Omega \times \mathbb{P}^n} \operatorname{div}_{T_\vartheta} \psi(x,\vartheta) \mathrm{d}\mu \\ &= \sup_{\psi \in M_0(\Omega)} \left\{ \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_1\Big) \cdot t_1^2 \, \mathrm{d}x_y - \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_2\Big) \cdot t_2^2 \, \mathrm{d}x_y \right\} \\ &\leq \int_{-1}^{1} \frac{1}{2} \cdot \underbrace{\|t_1^2\|_2}_{=1} \, \mathrm{d}x_y + \int_{-1}^{1} \frac{1}{2} \cdot \underbrace{\|t_2^2\|_2}_{=1} \, \mathrm{d}x_y = 2. \end{aligned}$$

In the last line $\psi(x,\vartheta) \in M_0(\Omega)$ and the Hölder inequality were used. In order to get equality here, we search for $\psi_n(x,\vartheta) \in M_0(\Omega)$ which fulfils

$$\lim_{n\to\infty} \int_{\Omega \times \mathbb{P}^n} \operatorname{div}_{T_\vartheta} \psi_n(x,\vartheta) \mathrm{d}\mu = 2.$$

Therefore we look at the function

$$\psi_n(x,\vartheta) = \frac{1}{2} \xi_n(x) \left[\eta_1(\vartheta) \cdot s(\vartheta) + \eta_2(\vartheta) \cdot s(\vartheta) \right] \tag{3.3}$$

where $s(\vartheta)$ is a smooth function that fulfils $s(\vartheta) \in T_\vartheta$ with $s(\vartheta_1) = t_1^2$ and $s(\vartheta_2) = t_2^2$ respectively. Let ξ_n be a smooth cutoff function in space, η_1 and η_2 are smooth localisers at ϑ_1 and ϑ_2.

$$\begin{aligned} s &: \mathbb{P}^n \to \mathbb{R}^n, \\ s(\vartheta) &= \sum_{i=1}^{n-1} c_i(\vartheta) t_i(\vartheta), \end{aligned}$$

where $t_i(\vartheta)$ are basis vectors of T_ϑ. Since the $\tilde{t}_i$ forms a basis the smooth weights $c_i(\vartheta)$ can be easily chosen in a way that $s(\vartheta_1) = t_1^2$ and $s(\vartheta_2) = t_2^2$ hold.

In order to describe η_1 and η_2 we introduce the smooth cutoff function

$$\begin{aligned} &\eta_0 : \mathbb{R} \to \mathbb{R} \\ &\eta_0(x) = \begin{cases} k \cdot \exp\left(-\frac{1}{1-(2x)^2}\right) & \text{if } \ |x| < \frac{1}{2} \\ 0 & \text{else.} \end{cases} \end{aligned}$$

So we have a function compactly supported in $]-1,1[$ that fulfils $\eta_0(0)=1$ for the right constant $k=e^1$.
Using η_0 and Definition 73 we can define

$$\begin{aligned}&\eta_i : \mathbb{P}^n \to \mathbb{R}\\ &\eta_i := \eta_0\left(\frac{d_{\mathbb{P}^n}(\vartheta,\vartheta_i)^2}{c^2}\right) \qquad \text{for } i=1,2.\end{aligned}$$

If $c>0$ is properly chosen the property $\operatorname{supp}\eta_1 \cap \operatorname{supp}\eta_2 = \emptyset$ can be easily verified.

Further we get the smooth cutoff function $\xi_n : \Omega \to \mathbb{R}$ needed for (3.3) by applying the theorem about partition of unity on the compact set $K_n = \left\{x \in \Omega \,|\, \operatorname{dist}(\partial\Omega, x) \geq \frac{1}{n}\right\}$, leading to

$$\lim_{n\to\infty} \xi_n\left((0,x_y,0)^t\right) = 1 \quad \forall x_y \in (-1,1)$$

Therefore we obtain the desired equality

$$\begin{aligned}&\lim_{n\to\infty}\int_{\Omega\times\mathbb{P}^n} \operatorname{div}_{T_\vartheta}\psi_n(x,\vartheta)\mathrm{d}\mu\\ =&\lim_{n\to\infty}\left\{\int_{-1}^1 \psi_n\Big((0,x_y,0),\vartheta_1\Big)\cdot t_1^2\,\mathrm{d}x_y - \int_{-1}^1 \psi_n\Big((0,x_y,0),\vartheta_2\Big)\cdot t_2^2\,\mathrm{d}x_y\right\}\\ =&\lim_{n\to\infty}\left\{\int_{-1}^1 \frac{1}{2}\xi_n\left((0,x_y,0)^t\right)\cdot \underbrace{t_1^2\cdot t_1^2}_{=1}\,\mathrm{d}x_y + \int_{-1}^1 \frac{1}{2}\xi_n\left((0,x_y,0)^t\right)\cdot \underbrace{t_2^2\cdot t_2^2}_{=1}\,\mathrm{d}x_y\right\}\\ =&\frac{1}{2}\int_{-1}^1 \lim_{n\to\infty}\xi_n\left((0,x_y,0)^t\right)\,\mathrm{d}x_y + \frac{1}{2}\int_{-1}^1 \lim_{n\to\infty}\xi_n\left((0,x_y,0)^t\right)\,\mathrm{d}x_y = 2.\end{aligned}$$

This leads to the interpretation, that $T_0(\mu)$ penalises the edge length of objects. Therefore the used definition of the set C_0 seems reasonable.

3.2.2. Restriction to a Set of Functionals that measure Curvature

In this subsection we want to extend the ideas presented in [4, Example 3.8] to a higher dimensional setting. There the geodesic distance was used to generate the

set C_ρ what led to a restriction of $\|\nabla_\vartheta \varphi\|_\infty$. As a reasonable generalisation for vector valued functions we define the set C_1 as

$$C_1 := \left\{ \psi(x,\cdot) \in C(\mathbb{P}^n, \mathbb{R}^n) \,\middle|\, \|\psi(x,\cdot)\|_{\infty,2} \leq \alpha,\ \|\nabla_\vartheta \psi(x,\cdot)\|_{\infty,2} \leq \beta \right\}.$$

With this setting we get for Example 70

$$\begin{aligned} T_1(\mu) &= \sup_{\psi \in M_1(\Omega)} \int_{\Omega \times \mathbb{P}^n} \operatorname{div}_{T_\vartheta} \psi(x,\vartheta) \mathrm{d}\mu \\ &= \sup_{\psi \in M_1(\Omega)} \left\{ \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_1\Big) \cdot t_1^2 \,\mathrm{d}x_y - \int_{-1}^{1} \psi\Big((0,x_y,0),\vartheta_2\Big) \cdot t_2^2 \,\mathrm{d}x_y \right\}. \end{aligned}$$

As an upper estimate we get

$$\begin{aligned} \left|\psi(\vartheta_1)\cdot t_1^2 - \psi(\vartheta_2)\cdot t_2^2\right| &= \left|\psi(\vartheta_1)\cdot t_1^2 - \psi(\vartheta_1)\cdot t_2^2 + \psi(\vartheta_1)\cdot t_2^2 - \psi(\vartheta_2)\cdot t_2^2\right| \\ &\leq \left|\psi(\vartheta_1)\cdot t_1^2 - \psi(\vartheta_1)\cdot t_2^2\right| + \left|\psi(\vartheta_1)\cdot t_2^2 - \psi(\vartheta_2)\cdot t_2^2\right| \\ &\leq \left|\psi(\vartheta_1)\cdot (t_1^2 - t_2^2)\right| + \left|(\psi(\vartheta_1) - \psi(\vartheta_2))\cdot t_2^2\right|. \end{aligned}$$

Now define the set

$$\Theta := \{\vartheta(s) : [0,1] \to \mathbb{P}^n \,|\, \vartheta(0) = \vartheta_1\,,\ \vartheta(1) = \vartheta_2\,, \vartheta \in C\,, \vartheta' \in C \text{ a.e.}\}$$

and apply the fundamental theorem of calculus. We obtain for every curve $\vartheta(s) \in \Theta$

$$\begin{aligned} \left|\psi(\vartheta_1)\cdot t_1^2 - \psi(\vartheta_2)\cdot t_2^2\right| &\leq \left|\psi(\vartheta_1)\cdot (t_1^2 - t_2^2)\right| + \left| \int_0^1 \nabla\psi(\vartheta(s))\cdot\vartheta'(s)\,\mathrm{d}s \cdot t_2^2 \right| \\ &\leq \left|\psi(\vartheta_1)\cdot (t_1^2 - t_2^2)\right| + \left| \int_0^1 \nabla\psi(\vartheta(s))\cdot\vartheta'(s)\,\mathrm{d}s \right| \cdot \underbrace{\left|t_2^2\right|}_{=1} \\ &\leq \left|\psi(\vartheta_1)\cdot (t_1^2 - t_2^2)\right| + \int_0^1 \left|\nabla\psi(\vartheta(s))\cdot\vartheta'(s)\right| \mathrm{d}s \\ &\leq \|\psi\| \left|t_1^2 - t_2^2\right| + \|\nabla_\vartheta\psi\| \int_0^1 \left|\vartheta'(s)\right| \mathrm{d}s \\ &\leq \alpha \cdot \left|t_1^2 - t_2^2\right| + \beta \cdot \int_0^1 \left|\vartheta'(s)\right| \mathrm{d}s. \end{aligned}$$

Taking the Infimum over all curves $\vartheta \in \Theta$ we obtain

$$\begin{aligned}\left|\psi(\vartheta_1)\cdot t_1^2 - \psi(\vartheta_2)\cdot t_2^2\right| \leq & \alpha\cdot\left|t_1^2 - t_2^2\right| + \beta\cdot\inf_{\vartheta\in\Theta}\int_0^1 \left|\vartheta'(s)\right|\,\mathrm{d}s \\ \leq & \alpha\cdot\left|t_1^2 - t_2^2\right| + \beta\cdot \mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2),\end{aligned}$$

where $\mathrm{d}(\cdot,\cdot)$ is the geodesic metric on $\mathbb{P}^n\times\mathbb{P}^n$ as defined in Definition 74. This leads to

$$T_1(\mu) \leq 2\cdot\left(\alpha\cdot\left|t_1^2 - t_2^2\right| + \beta\cdot\mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2)\right) \tag{3.4}$$

To get a lower estimate as well, we choose a special function $\psi \in M_1(\Omega)$ instead of taking the supremum. For ease of notation we will calculate this estimate for the special case $n = 3$. Furthermore we will denote the vector t_1^2 by t_1. The idea is to construct a function $\psi \in M_1(\Omega)$ with the properties $\psi(x,\vartheta_1) = \mathrm{const}\cdot t_1$, $\psi(x,\vartheta_2) = 0$ and $\|\nabla_\vartheta\psi(x,\cdot)\|_{\infty,2} \leq \beta$.
For all calculations on the unit ball it is useful to introduce the Spherical Coordinates. We will call u the degree of longitude and v will denote the degree of latitude. Therefore we define the mapping

$$\begin{aligned}\tau :[-\pi,\pi]\times[-\frac{\pi}{2},\frac{\pi}{2}] &\to S^3 \\ \begin{pmatrix}u\\v\end{pmatrix} &\mapsto \begin{pmatrix}\cos v\cdot\cos u\\ \cos v\cdot\sin u\\ \sin v\end{pmatrix}\end{aligned}$$

which maps Spherical Coordinates to Cartesian Coordinates.

Using this we can define the equator as a subset of $\mathbb{P}^3$ as

$$\mathrm{Eq} := \left\{\vartheta\in\mathbb{P}^3\,\middle|\,\exists u\in[-\pi,\pi[\text{ s.t. } j\left(\tau\begin{pmatrix}u\\0\end{pmatrix}\right) = \vartheta\right\} \tag{3.5}$$

Further we rotate the coordinate system, that $j\left(\tau\begin{pmatrix}0\\0\end{pmatrix}\right) = \vartheta_1$ holds and $j^+(\vartheta_2)$ lays on the equator. The mappings j and j^+ are given in Definition 59.

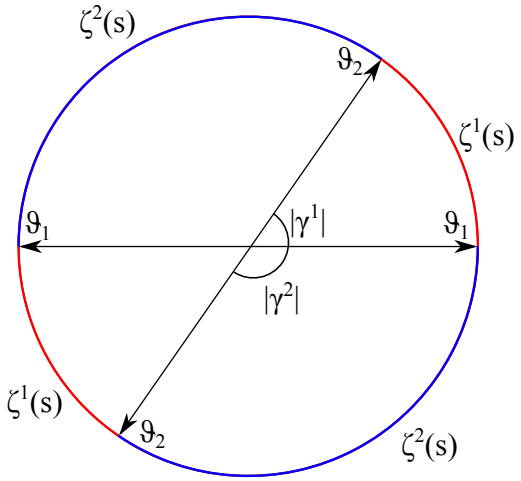


Figure 3.7.: The equator is parametrised by the paths $\zeta^1(s)$ and $\zeta^2(s)$

Now we want to parametrise the equator of $\mathbb{P}^3$. This is equivalent to parametrising the shortest and the second shortest path between ϑ_1 and ϑ_2 as shown in Figure 3.7.

In order to define these paths we will first investigate the lengths of these paths. For the shortest path we define γ^1 as the minimiser of

$$\begin{aligned} &\min |\gamma| \\ &\text{s.t. } j\left(\tau\begin{pmatrix}\gamma\\0\end{pmatrix}\right) = \vartheta_2. \end{aligned}$$

Clearly this is the correct way to calculate the length of the shortest path since it measures the length on the great circle between $j^+(\vartheta_1)$ and $j^+(\vartheta_2)$. This minimisation step is necessary since the mapping $j\left(\tau\begin{pmatrix}\cdot\\0\end{pmatrix}\right)$ is π-periodic.

If this solution is unique (keep in mind that there are two solutions in the case $j^+(\vartheta_1) \perp j^+(\vartheta_2)$) the length of the second shortest path will be denoted by γ^2 and is the minimiser of

$$\begin{aligned} &\min |\gamma| \\ &\text{s.t. } \begin{cases} j\left(\tau\begin{pmatrix}\gamma\\0\end{pmatrix}\right) = \vartheta_2, \\ \gamma \neq \gamma^1. \end{cases} \end{aligned}$$

The parametrisation of the shortest and second shortest path is given by $\zeta^1 : [0,1] \mapsto \mathbb{P}^3$ and $\zeta^2 : [0,1] \mapsto \mathbb{P}^3$ with

$$\zeta^1(s) = j\left(\tau\begin{pmatrix} s\gamma^1 \\ 0 \end{pmatrix}\right) \tag{3.6}$$

and

$$\zeta^2(s) = j\left(\tau\begin{pmatrix} s\gamma^2 \\ 0 \end{pmatrix}\right). \tag{3.7}$$

These two parametrisations have the properties

$$\zeta^i(0) = \vartheta_1\,,\ \zeta^i(1) = \vartheta_2 \quad \text{for } i = 1,2.$$

It is clear that

$$|\gamma_1| = \mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2) \tag{3.8}$$

holds, where $\mathrm{d}_{\mathrm{geo}}(\vartheta_1, \vartheta_2)$ is the geodesic distance on $\mathbb{P}^n$ as defined in Definition 74. For the absolute value of ψ on the equator we want to have a piecewise linear structure like it is shown in Figure 3.8.

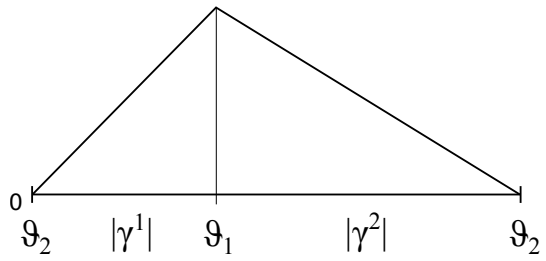


Figure 3.8.: Absolute value of ψ on the equator.

Since $\psi(\vartheta) \in T_\vartheta$ has to hold we look at the vectors t_1 and t_2 that characterise the tangential plane T_ϑ for any $\vartheta \in \mathbb{P}^3$. Let us take an arbitrary vector on the unit ball

$$\tau\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos v \cdot \cos u \\ \cos v \cdot \sin u \\ \sin v \end{pmatrix}$$

and derive it in direction u and v to get the desired vectors in the tangential plane. For $c_1, c_2 \neq 0$ we get

$$t_1 = c_1(u,v) \cdot \frac{\partial \tau}{\partial u} = c_1(u,v) \cdot \begin{pmatrix} -\cos v \cdot \sin u \\ \cos v \cdot \cos u \\ 0 \end{pmatrix}, \tag{3.9}$$

$$t_2 = c_2(u,v) \cdot \frac{\partial \tau}{\partial v} = c_2(u,v) \cdot \begin{pmatrix} -\sin v \cdot \cos u \\ -\sin v \cdot \sin u \\ \cos v \end{pmatrix}. \tag{3.10}$$

We can easily verify that $\tau\begin{pmatrix} u \\ v \end{pmatrix} \perp t_1 \perp t_2$. To make sure t_1 and t_2 form an orthonormal basis we have to choose $c_1(u,v) := \frac{1}{\cos v}$ and $c_2(u,v) := 1$. With this setting we have

$$T_\vartheta = \operatorname{span}\{t_1, t_2\} = \operatorname{span}\left\{\begin{pmatrix} -\sin u \\ \cos u \\ 0 \end{pmatrix}, \begin{pmatrix} -\sin v \cdot \cos u \\ -\sin v \cdot \sin u \\ \cos v \end{pmatrix}\right\} \text{ if } \vartheta = j\left(\tau\begin{pmatrix} u \\ v \end{pmatrix}\right).$$

Using the parametrisations ζ^1 and ζ^2 as defined in (3.6) and (3.7) and their inverses on the domain of definition we define a function ψ_0 on the equator (according to (3.5)).

$$\begin{aligned} \psi_0 :& \mathrm{Eq} \to \mathbb{R}^3 \\ & \vartheta \mapsto \text{const} \cdot \left(1 - (\zeta^k)^{-1}(\vartheta)\right) \cdot t_1(\vartheta) \quad \text{if } \vartheta \in \zeta^k([0,1]) \text{ for } k \in \{1,2\}. \end{aligned}$$

The constant in the definition of ψ_0 will be defined later. $(\zeta^k)^{-1}$ can be easily calculated using

$$(\zeta^k)^{-1} = \frac{u}{\gamma^k}, \text{ if } j\left(\tau\begin{pmatrix} u \\ 0 \end{pmatrix}\right) \in \zeta^k([0,1]) \text{ for } k \in \{1,2\}.$$

In order to extend this idea to the whole $\mathbb{P}^3$ we introduce a projection from $\mathbb{P}^3$ onto the equator.

$$\begin{aligned} \mathrm{P}_{\mathrm{Eq}} :& \mathbb{P}^3 \to \mathrm{Eq} \\ & \vartheta \mapsto j\left(\tau\begin{pmatrix} u \\ 0 \end{pmatrix}\right), \text{ with } \vartheta = j\left(\tau\begin{pmatrix} u \\ v \end{pmatrix}\right). \end{aligned}$$

To make sure that ψ is continuous on $\mathbb{P}^3$ we have to define a decay function as well.

$$\begin{aligned} f :&\mathbb{P}^3 \to [0,1] \\ &\vartheta \mapsto \cos v \ \text{ with } \vartheta = j\left(\tau\begin{pmatrix}u\\v\end{pmatrix}\right). \end{aligned}$$

Now we can define $\psi_n : \Omega \times \mathbb{P}^3 \to \mathbb{R}^3$ by

$$\psi_n(x,\vartheta) := \xi_n(x) \cdot \left[f(\vartheta) \cdot \psi_0(\mathrm{P}_{\mathrm{eq}}(\vartheta))\right],$$

where ξ_n is chosen in the same way as in Section 3.2.1.
In order to compute $\|\nabla_\vartheta \psi(x,\cdot)\|_{\infty,2}$ we have to think of how to differentiate in the directions t_1 and t_2. Since $\{c_1 t_1, c_2 t_2\}$ form an ONB of the Tangential plane, as we can see in Equation (3.9) and (3.10), the tangential directions can be identified with the derivatives in direction u and v.
Therefore we get for the following derivatives as the components of $\nabla_\vartheta \psi_n$.

$$\begin{aligned}
\frac{\partial}{\partial t_1}\left(\psi_n \cdot t_1\right)(x,\vartheta) &= \mathrm{const} \cdot \xi_n(x) \cdot \underbrace{c_1(u,v)}_{=\frac{1}{\cos v}} \cdot \frac{\partial}{\partial u}\left[\cos v \cdot \left(1 - \frac{u}{\gamma_k}\right)\underbrace{t_1 \cdot t_1}_{=1}\right] \\
&= -\mathrm{const} \cdot \xi_n(x) \cdot \frac{1}{\gamma_k} \ \text{ for } \ \mathrm{P}_{\mathrm{Eq}}(\vartheta) \in \vartheta_k^{-1}([0,1]). \\
\frac{\partial}{\partial t_2}(\psi_n \cdot t_1)(x,\vartheta) &= \mathrm{const} \cdot \xi_n(x) \cdot \frac{\partial}{\partial v} \cdot \underbrace{c_2(u,v)}_{=1}\left[\cos v \cdot \left(1 - \frac{u}{\gamma_k}\right)\underbrace{t_1 \cdot t_1}_{=1}\right] \\
&= -\mathrm{const} \cdot \xi_n(x) \cdot \sin v\left(1 - \frac{u}{\gamma_k}\right) \ \text{ for } \ \mathrm{P}_{\mathrm{Eq}}(\vartheta) \in \vartheta_k^{-1}([0,1]). \\
\frac{\partial}{\partial t_1}(\psi_n \cdot t_2)(x,\vartheta) &= \mathrm{const} \cdot \xi_n(x) \cdot \frac{\partial}{\partial u}\left[\cos v \cdot c_1(u,v) \cdot \left(1 - \frac{u}{\gamma_k}\right)\underbrace{t_1 \cdot t_2}_{=0}\right] = 0. \\
\frac{\partial}{\partial t_2}(\psi_n \cdot t_2)(x,\vartheta) &= \mathrm{const} \cdot \xi_n(x) \cdot \frac{\partial}{\partial v}\left[\cos v \cdot c_2(u,v) \cdot \left(1 - \frac{u}{\gamma_k}\right)\underbrace{t_1 \cdot t_2}_{=0}\right] = 0.
\end{aligned}$$

This leads to

$$\|\nabla_\vartheta \psi_n\|_{\infty,2} = \sup_{(x,\vartheta)\in\Omega\times\mathbb{P}^n} \sqrt{\text{const}^2 \cdot \underbrace{\xi_n(x)^2}_{\leq 1} \cdot \left[\underbrace{\frac{1}{\gamma_k^2}}_{\leq \frac{1}{\gamma_1^2}} + \underbrace{(\sin v)^2}_{\leq 1} \cdot \underbrace{\left(1 - \frac{u}{\gamma_k}\right)^2}_{\leq 1} \right]}$$
$$\leq \text{const} \cdot \sqrt{\frac{1}{\gamma_k^2} + 1}$$

and therefore we choose

$$\text{const} \leq \frac{\beta}{\sqrt{\frac{1}{\gamma_1^2} + 1}}.$$

This can be estimated with

$$\begin{aligned} \frac{\beta}{\sqrt{\frac{1}{\gamma_1^2} + 1}} &\geq \frac{\beta}{\frac{1}{|\gamma_1|} + 1} \\ &= \beta \cdot \frac{|\gamma_1|}{1 + \underbrace{|\gamma_1|}_{\leq \pi/2}} \\ &\overset{(3.8)}{\geq} \beta \cdot \frac{1}{1 + \pi/2} \cdot \mathrm{d}_{\text{geo}}(\vartheta_1, \vartheta_2). \end{aligned}$$

Now we can set

$$\text{const} = \frac{\beta}{1 + \pi/2} \cdot \mathrm{d}_{\text{geo}}(\vartheta_1, \vartheta_2),$$

and if the condition

$$\frac{\beta}{1 + \pi/2} \cdot \mathrm{d}_{\text{geo}}(\vartheta_1, \vartheta_2) \leq \frac{\beta}{1 + \pi/2} \cdot \frac{\pi}{2} \leq \alpha$$

is fulfilled we have $\|\psi_n\|_\infty \leq \alpha$ and get the estimate

$$\begin{aligned} T_1(\mu) &\geq \lim_{n\to\infty} \int_{-1}^{1} \psi_n((0,x_y,0),\vartheta_1)\cdot t_1 \,\mathrm{d}x_y - \int_{-1}^{1} \psi_n((0,x_y,0),\vartheta_2)\cdot t_2 \,\mathrm{d}x_y \\ &\geq \lim_{n\to\infty} \int_{-1}^{1} \xi_n(x)\cdot \frac{\beta}{1+\pi/2}\cdot \mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2)\cdot \underbrace{t_1\cdot t_1}_{=1}\cdot \mathrm{d}x_y \\ &= \frac{2\cdot\beta}{1+\pi/2}\cdot \mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2). \end{aligned}$$

Together with (3.4) we have

$$\frac{2\cdot\beta}{1+\pi/2}\cdot \mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2) \leq T_1(\mu) \leq 2\cdot\Big(\alpha\cdot\left|t_1^2-t_2^2\right| + \beta\cdot \mathrm{d}_{\mathrm{geo}}(\vartheta_1,\vartheta_2)\Big), \tag{3.11}$$

what leads to the intuition that $T_1(\mu)$ penalises quantitatively the geodesic distance between ϑ_1 and ϑ_2.

Remark 76. *We have to mention here, that ψ can be differentiated only almost everywhere. If $\mathrm{P}_{\mathrm{Eq}}(\vartheta)=\vartheta_1$ or $\mathrm{P}_{\mathrm{Eq}}(\vartheta)=\vartheta_2$ holds, ψ is continuous but not differentiable anymore. Anyway we can get rid of this problem by using suitable mollifiers.*

3.3. Relaxation

In this section we want to use the theory derived in Section 3.2 to set up regularisers that are applicable to a larger class of images. We will see that the resulting formulations are non convex and therefore we need to come up with a convex relaxed expression to make the numerical solution accessible. Later in this section we want to discuss inpainting and denoising problems using the derived relaxed regulariser. Using (3.11) and the results of Sections 3.2.1 and 3.2.2 we see how $T_0(\mu)$ and $T_1(\mu)$ work on the special case shown in Figure 3.6. The idea is to transfer this estimate to the general case of characteristic functions with piecewise linear boundary and further to extend this theory on images $u \in \mathrm{BV}(\Omega)$. Anyway, the analysis of this generalisation is beyond the scope of this work, but we are optimistic that this generalisation is possible. Since it was shown for two dimensional problems in [4], we are confident to say that this assumption is valid. Therefore we want to work with

functions $u \in \mathrm{BV}(\Omega)$ from now on.
For functions $u \in \mathrm{BV}(\Omega)$ we want to operate on the sublevel sets of u given by the characteristic function $\chi_{\{u<s\}}$ for an arbitrary $s \in \mathbb{R}$. Consequently we will denote the functional lifting of $\nabla(\chi_{\{u<s\}})$ by μ_s. Further we want to make use of Definition 62 and the coarea formula (see [3]) to observe

$$\begin{aligned}\int_{\Omega\times\mathbb{P}^n} \varphi \,\mathrm{d}\mu &= \int_{\Omega} \operatorname{sign}(\sigma(x)) \cdot \varphi\left(x, j(\sigma(x))\right) \,\mathrm{d}\,|\nabla u| \\ &= \int_{\mathbb{R}} \int_{\partial\{u<s\}} \operatorname{sign}(\sigma(x)) \cdot \varphi\left(x, j(\sigma(x))\right) \,\mathrm{d}\mathcal{H}^{n-1}\,\mathrm{d}s \\ &= \int_{\mathbb{R}} \int_{\Omega\times\mathbb{P}^n} \varphi \,\mathrm{d}\mu_s \,\mathrm{d}s.\end{aligned}$$

With the work done so far in this chapter, we are now able to introduce a reasonable regulariser for functions $u \in \mathrm{BV}(\Omega)$. We define

$$R_\rho^{\alpha,\beta}(u) := \int_{\mathbb{R}} \alpha\|\mu_s\|_{\mathcal{M}} + \beta T_\rho(\mu_s) \,\mathrm{d}s, \tag{3.12}$$

with $\alpha, \beta > 0$ and $T_\rho(\mu_s)$ as defined in Definition 71. The norm $\|\mu_s\|_{\mathcal{M}}$ corresponds to the perimeter of the sublevel sets induced by $\{u < s\}$ as discussed in Example 29. Due to the non-convexity of the extraction of the sublevel sets $\{u < s\}$ and the lifting operator $u \mapsto \mu(\nabla u)$ we have to come up with reasonable convex relaxations to make the introduced ideas computable.
Given (3.12) and Definition 71 we observe for the functional lifting μ of ∇u and an arbitrary $\psi \in M_\rho(\Omega)$ that

$$\begin{aligned}\int_{\Omega\times\mathbb{P}^n} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\mathrm{d}\mu &= \int_{\mathbb{R}} \int_{\Omega\times\mathbb{P}^n} \mathrm{div}_{T_\vartheta}\psi(x,\vartheta)\mathrm{d}\mu_s \,\mathrm{d}s \\ &\leq \int_{\mathbb{R}} T_\rho(\mu_s) \,\mathrm{d}s\end{aligned}$$

holds. On the other hand we get according to Remark 21 that

$$\|\mu\|_{\mathcal{M}} = \int_{\Omega\times\mathbb{P}^n} \mathbf{1} \,\mathrm{d}\mu = \int_{\mathbb{R}} \int_{\Omega\times\mathbb{P}^n} \mathbf{1} \,\mathrm{d}\mu_s \,\mathrm{d}s = \int_{\mathbb{R}} \|\mu_s\|_{\mathcal{M}} \,\mathrm{d}s,$$

which leads to

$$\alpha\|\mu\|_{\mathcal{M}} + \beta T_\rho(\mu) \leq R_\rho^{\alpha,\beta}(u). \tag{3.13}$$

Define $G_\nabla := \{(u,\mu) \in BV(\Omega) \times \mathcal{M}(\Omega \times \mathbb{P}^n) \,|\, \mu \text{ is the Functional Lifting of } \nabla u\}$ as the set of all pairs $(u,\mu) \in BV(\Omega) \times \mathcal{M}(\Omega \times \mathbb{P}^n)$ where $u \mapsto \mu(\nabla u)$ holds.
To make G_∇ computable we have to think of a proper way to relax this set. Therefore we make use of the necessary condition

$$\int_{\Omega \times \mathbb{P}^n} \varphi(x) \cdot j^+(\tau) \, \mathrm{d}\mu = \int_\Omega \varphi \, \mathrm{d}\nabla u = -\int_\Omega u \cdot \mathrm{div}\varphi \, \mathrm{d}x$$

that was obtained in Remark 65.
This leads to the set

$$M_\nabla = \Bigg\{ (u,\mu) \in L^1(\Omega) \times \mathcal{M}(\Omega \times \mathbb{P}^n) \,\Bigg|\, \int_{\Omega \times \mathbb{P}^n} \varphi(x) \cdot j^+(\tau) \, \mathrm{d}\mu + \int_\Omega u \cdot \mathrm{div}\varphi \, \mathrm{d}x = 0, \ \forall \varphi \in C_C^\infty(\Omega, \mathbb{R}^n) \Bigg\} \tag{3.14}$$

as a reasonable relaxation of G_∇.
Now we are able to give a reasonable regulariser of (3.12) that is presented in Definition 77.

Definition 77. *For $\alpha, \beta > 0$ and $u \in L^1(\Omega)$ we define the regulariser*

$$\overline{R}_\rho^{\alpha,\beta}(u) = \inf_{(u,\mu) \in M_\nabla} \alpha \cdot \|\mu\|_{\mathcal{M}} + \beta \cdot T_\rho(\mu).$$

Remark 78. *For the Functional $T_\rho(\mu)$ we examined the special cases $T_0(\mu)$ and $T_1(\mu)$ in the Sections 3.2.1 and 3.2.2.*
We will denote the regulariser $\overline{R}_{\rho_0}^{\alpha,\beta}(u)$ that is induced by $T_0(\mu)$ with TVX_0 and the regulariser $\overline{R}_{\rho_1}^{\alpha,\beta}(u)$ that is induced by $T_1(\mu)$ with TVX_1 to stay consistent with [4].

3.4. Application on imaging problems

In order to apply the developed regulariser to imaging problems we have to examine if the resulting formulations will have a solution. To show solvability we want to use the direct method that was presented in Section 2.2. The following proofs can be

found in [4] for the two dimensional case.
First we want to discuss the differentiation operator of μ that was declared in Definition 71 and is given by

$$T_\rho(\mu) = \sup_{\psi \in M_\rho(\Omega)} \int_{\Omega \times \mathbb{P}^n} \mathrm{div}_{T_\vartheta} \psi(x,\vartheta) \mathrm{d}\mu$$

with

$$M_\rho(\Omega) = \Big\{ \psi \in C_C(\Omega \times \mathbb{P}^n, \mathbb{R}^n) \,\Big|\, \mathrm{div}_{T_\vartheta}\psi \in C_C(\Omega \times \mathbb{P}^n), \psi(x,\cdot) \in C_\rho \;\; \forall x \in \Omega \\ \psi(x,\vartheta) \in T_\vartheta, \; \forall \vartheta \in \mathbb{P}^n \Big\}$$

At first we mention that $T_\rho(\mu)$ can be interpreted as the duality product

$$T_\rho(\mu) = \sup_{\psi \in M_\rho(\Omega)} \langle \mu, \mathrm{div}_{T_\vartheta}\psi \rangle. \tag{3.15}$$

In the next Lemma we want to investigate some key properties of T_ρ, that will be needed to prove existence of the resulting imaging formulations.

Lemma 79. *$T_\rho : \mathcal{M}(\Omega \times \mathbb{P}^n) \to [0,\infty]$ is weak* lower semi-continuous, positively homogeneous and satisfies the triangle inequality.*

Proof. (see [4, Prop.3.12]) First we want to show that $T_\rho(\mu) \geq 0$ holds for all $\mu \in \mathcal{M}(\Omega \times \mathbb{P}^n)$.
With the symmetry of $M_\rho(\Omega)$ we have for $\psi \in M_\rho(\Omega)$ that $-\psi \in M_\rho(\Omega)$ holds as well. Therefore we get with equation (3.15)

$$T_\rho(\mu) = \sup_{\psi \in M_\rho(\Omega)} |\langle \mu, \mathrm{div}_{T_\vartheta}\psi \rangle| \geq 0,$$

what provides the non negativity.
The weak* lower semi-continuity of T_ρ follows immediately, since it is the pointwise supremum of functions out of $C_C(\Omega \times \mathbb{P}^n)$.
Using the symmetry of M_ρ again, we get for $\lambda \in \mathbb{R}$ the positive homogeneity with

$$\begin{aligned} T_\rho(\lambda\mu) &= \sup_{\psi \in M_\rho(\Omega)} \langle |\lambda| \mu, \mathrm{sgn}(\lambda) \mathrm{div}_{T_\vartheta}\psi \rangle \\ &= |\lambda| \sup_{\psi \in \mathrm{sgn}(\lambda) M_\rho(\Omega)} \langle \mu, \mathrm{div}_{T_\vartheta}\psi \rangle \\ &= |\lambda| \, T_\rho(\mu). \end{aligned}$$

To prove the triangle inequality we observe for $\mu_1, \mu_2 \in \mathcal{M}(\Omega \times \mathbb{P}^n)$

$$\langle \mu_1 + \mu_2, \operatorname{div}_{T_\vartheta} \psi \rangle \leq \sup_{\psi \in M_\rho(\Omega)} \langle \mu_1, \operatorname{div}_{T_\vartheta} \psi \rangle + \sup_{\psi \in M_\rho(\Omega)} \langle \mu_2, \operatorname{div}_{T_\vartheta} \psi \rangle = T_\rho(\mu_1) + T_\rho(\mu_2).$$

By taking the supremum on the left hand side we get

$$T_\rho(\mu_1 + \mu_2) \leq T_\rho(\mu_1) + T_\rho(\mu_2).$$

□

Remark 80. *Convexity of T_ρ is an immediate consequence of the positive homogeneity and the triangle inequality. For $\lambda \in [0,1]$ and $\mu_1, \mu_2 \in \mathcal{M}(\Omega \times \mathbb{P}^n)$ we get*

$$\begin{aligned} T_\rho(\lambda \mu_1 + (1-\lambda)\mu_2) &\leq T_\rho(\lambda \mu_1) + T_\rho((1-\lambda)\mu_2) \\ &= \lambda T_\rho(\mu_1) + (1-\lambda) T_\rho(\mu_2). \end{aligned}$$

In the following Lemma we give properties for the relaxed set M_∇ given in (3.14). In [4, Proposition 4.1] these properties were shown for the two dimensional case.

Lemma 81. *The relaxed set M_∇ derived in* (3.14) *is non empty and convex. Further it is sequentially closed with weak convergence in $L^1(\Omega)$ and weak* convergence in $\mathcal{M}(\Omega \times \mathbb{P}^n)$.*

Proof. With $(\mathbf{0}, \mathbf{0}) \in M_\nabla$, the set is clearly non empty.
The convexity follows directly by using the linearity of the constraint. For $(u_1, \mu_1) \in M_\nabla$, $(u_2, \mu_2) \in M_\nabla$ and $\lambda \in [0,1]$ we get

$$\begin{aligned} &\int_{\Omega \times \mathbb{P}^n} \varphi(x) \cdot j^+(\tau) \, \mathrm{d}(\lambda\mu_1 + (1-\lambda)\mu_2) + \int_\Omega (\lambda u_1 + (1-\lambda)u_2) \cdot \operatorname{div}\varphi \, \mathrm{d}x \\ &= \lambda \underbrace{\left[\int_{\Omega \times \mathbb{P}^n} \varphi(x) \cdot j^+(\tau) \, \mathrm{d}\mu_1 + \int_\Omega \lambda u_1 \cdot \operatorname{div}\varphi \, \mathrm{d}x \right]}_{=0} \\ &+ (1-\lambda) \underbrace{\left[\int_{\Omega \times \mathbb{P}^n} \varphi(x) \cdot j^+(\tau) \, \mathrm{d}\mu_2 + \int_\Omega u_2 \cdot \operatorname{div}\varphi \, \mathrm{d}x \right]}_{=0} = 0. \end{aligned}$$

Therefore we have $(\lambda u_1 + (1-\lambda)u_2, \lambda\mu_1 + (1-\lambda)\mu_2) \in M_\nabla$, which gives the convexity of M_∇.
Now consider a sequence $(u^n, \mu^n) \in M_\nabla$ where $\{u^n\}_n$ converges weak in $L^1(\Omega)$ and $\{\mu^n\}_n$ converges weakly* in $\mathcal{M}(\Omega \times \mathbb{P}^n)$. With (3.14) we get immediately

$$\begin{aligned}&\int_{\Omega\times\mathbb{P}^n} \varphi(x)\cdot j^+(\tau)\,\mathrm{d}\mu + \int_\Omega u\cdot\mathrm{div}\varphi\,\mathrm{d}x\\ &= \lim_{n\to\infty}\int_{\Omega\times\mathbb{P}^n}\varphi(x)\cdot j^+(\tau)\,\mathrm{d}\mu^n + \int_\Omega u^n\cdot\mathrm{div}\varphi\,\mathrm{d}x = 0\end{aligned}$$

for all $\varphi \in C_C^\infty(\Omega, \mathbb{R}^n)$. □

Proposition 82. *For elements $(u,\mu) \in M_\nabla$, we get that $u \in \mathrm{BV}(\Omega)$ holds.*

Proof. This property follows, by restricting the set of all test functions $\varphi \in C_C^\infty(\Omega, \mathbb{R}^n)$ to $\|\varphi\|_\infty \le 1$. With respect to Remark 21 we get

$$\int_\Omega u\cdot\mathrm{div}\varphi\,\mathrm{d}x \le \int_{\Omega\times\mathbb{P}^n} \underbrace{|\varphi(x)\cdot j^+(\tau)|}_{\le 1}\,\mathrm{d}\mu \le \|\mu\|_{\mathcal{M}} < \infty.$$

□

The next step is a deeper investigation of the relaxed functional $\overline{R}_\rho^{\alpha,\beta}$, defined in Definition 77, like it was done in [4, Proposition 4.4] for the two dimensional case.

Lemma 83. *The functional $\overline{R}_\rho^{\alpha,\beta} : L^1(\Omega) \to \mathbb{R}_\infty$ is proper, bounded form below, lower semi-continuous and convex. Further we get the properties*

$$\overline{R}_\rho^{\alpha,\beta}(\lambda u) = \lambda\overline{R}_\rho^{\alpha,\beta}(u)$$

for $\lambda \ge 0$ and

$$\alpha\mathrm{TV} \le \overline{R}_\rho^{\alpha,\beta} \le R_\rho^{\alpha,\beta}.$$

Proof. First is has to be mentioned that $(\mathbf{0},\mathbf{0}) \in M_\nabla$ provides immediately that $\overline{R}_\rho^{\alpha,\beta}$ is proper, since $\|\mathbf{0}\|_{\mathcal{M}} = 0$ and $T_\rho(\mathbf{0}) = 0$ lead to $\overline{R}_\rho^{\alpha,\beta}(\mathbf{0}) = 0$.

To show that $\overline{R}_\rho^{\alpha,\beta}(u) = \inf_{(u,\mu)\in M_\nabla} \alpha\cdot\|\mu\|_{\mathcal{M}} + \beta\cdot T_\rho(\mu)$ is bounded from below we utilise Lemma 79 where $T_\rho(\mu) \geq 0$ was shown and get

$$\alpha\cdot\|\mu\|_{\mathcal{M}} + \beta\cdot T_\rho(\mu) \geq 0$$

for all $\mu \in \mathcal{M}(\Omega\times\mathbb{P}^n)$.
To prove the lower semi-continuity we consider a sequence $u^n \to u$ in $L^1(\Omega)$. Further we choose for each $\{u^n\}_n$ a minimising sequence $(\mu^{n_m}) \in \mathcal{M}(\Omega\times\mathbb{P}^n)$ such that $(u^n,\mu^{n_m}) \in M_\nabla$ holds. Since $\{\mu^{n_m}\}_m$ is a minimising sequence we get

$$\overline{R}_\rho^{\alpha,\beta}(u^n) = \alpha\cdot\left\|\mu^{n_m}\right\|_{\mathcal{M}} + \beta\cdot T_\rho(\mu^{n_m}).$$

Without loss of generality we assume that the series $\overline{R}_\rho^{\alpha,\beta}(u^n)$ converges to a finite value. Further we consider a diagonal sequence $\{\tilde{\mu}^n\}_n$ of $\{\mu^{n_m}\}_m$, that fulfils

$$\alpha\|\tilde{\mu}^n\|_{\mathcal{M}} + \beta T_\rho(\tilde{\mu}^n) \geq \overline{R}_\rho^{\alpha,\beta}(u^n) + \frac{1}{n} \tag{3.16}$$

for all $n \geq 1$.
With $T_\rho(\tilde{\mu}^n) \geq 0$ we can show the inequality

$$\begin{aligned}\|\tilde{\mu}^n\|_{\mathcal{M}} &\leq \|\tilde{\mu}^n\|_{\mathcal{M}} + \frac{\beta}{\alpha}T_\rho(\tilde{\mu}^n)\\ &\overset{(3.16)}{\leq} \frac{1}{\alpha}\left(\sup_{n\geq 1}\overline{R}_\rho^{\alpha,\beta}(u^n)+1\right) < \infty,\end{aligned}$$

which provides that the sequence $\{\tilde{\mu}^n\}_n$ is bounded in $\mathcal{M}(\Omega\times\mathbb{P}^n)$. This proves the existence of a subsequence $\{\tilde{\mu}^{n_k}\}_k$ of $\{\tilde{\mu}^n\}_n$ and an element $\mu \in \mathcal{M}(\Omega\times\mathbb{P}^n)$ with $\tilde{\mu}^{n_k} \overset{*}{\rightharpoonup} \mu$. Let us denote this subsequence with $\hat{\mu}^n$.
Since Lemma 81 showed that M_∇ is closed we get $(u,\mu) \in M_\nabla$.
With the weak* lower semi-continuity of $\|\cdot\|_{\mathcal{M}}$ and T_ρ that was shown in Lemma 79 we get the lower semi-continuity of $\overline{R}_\rho^{\alpha,\beta}$ with

$$\begin{aligned}\overline{R}_\rho^{\alpha,\beta}(u) &= \alpha\cdot\|\mu\|_{\mathcal{M}} + \beta\cdot T_\rho(\mu)\\ &\leq \liminf_{n\to\infty} \alpha\|\hat{\mu}^n\|_{\mathcal{M}} + \beta T_\rho(\bar{\mu}^n) = \lim_{n\to\infty}\overline{R}_\rho^{\alpha,\beta}(u^n).\end{aligned}$$

To show convexity of $\overline{R}_\rho^{\alpha,\beta}$, let u^1, u^2 be in $L^1(\Omega)$ with $\overline{R}_\rho^{\alpha,\beta}(u^1)$, $\overline{R}_\rho^{\alpha,\beta}(u^2) < \infty$ and let $(\mu^1)^n$, $(\mu^2)^n$ be corresponding minimising sequences as it was already used above.

The convexity of M_∇ and T_ρ were shown in Lemma 81 and Remark 80. Therefore we have for $\lambda \in [0,1]$

$$\begin{aligned}
&\overline{R}_\rho^{\alpha,\beta}\left(\lambda u^1 + (1-\lambda)u^2\right)\\
&\leq \liminf_{n\to\infty} \alpha \left\|\lambda\,(\mu^1)^n + (1-\lambda)\,(\mu^2)^n\right\|_{\mathcal{M}} + \beta T_\rho\left(\lambda\,(\mu^1)^n + (1-\lambda)\,(\mu^2)^n\right)\\
&\leq \lambda\left(\lim_{n\to\infty} \alpha \left\|(\mu^1)^n\right\|_{\mathcal{M}} + \beta\, T_\rho\left((\mu^1)^n\right)\right)\\
&+ (1-\lambda)\left(\lim_{n\to\infty} \alpha \left\|(\mu^2)^n\right\|_{\mathcal{M}} + \beta\, T_\rho\left((\mu^2)^n\right)\right)\\
&= \lambda\,\overline{R}_\rho^{\alpha,\beta}\left(u^1\right) + (1-\lambda)\,\overline{R}_\rho^{\alpha,\beta}\left(u^2\right).
\end{aligned}$$

The positive one-homogeneity of $\overline{R}_\rho^{\alpha,\beta}$ is an immediate consequence of the positive homogeneity of $\|\cdot\|_{\mathcal{M}}$ and T_ρ, which was shown in Lemma 79, since $(u,\mu) \in M_\nabla$ is equivalent to $(\lambda u, \lambda\mu) \in M_\nabla$ for $\lambda > 0$.

Finally for $u \in \mathrm{BV}(\Omega)$ and $\mu \in \mathcal{M}(\Omega \times \mathbb{P}^n)$ with $(u,\mu) \in M_\nabla$ and test functions $\varphi \in C_C^\infty(\Omega, \mathbb{R}^n)$ with $\|\varphi\|_\infty \leq 1$ we get in view of Proposition 82

$$\alpha \int_\Omega u \cdot \mathrm{div}\varphi \,\mathrm{d}x \leq \alpha \int_{\Omega\times\mathbb{P}^n} \underbrace{|\varphi(x)\cdot j^+(\tau)|}_{\leq 1} \,\mathrm{d}\mu \leq \alpha\|\mu\|_{\mathcal{M}} + \beta T_\rho(\mu),$$

since $T_\rho \geq 0$. By taking the supremum over all φ on the left hand side and the infimum over all μ with $(u,\mu) \in M_\nabla$ on the right hand side we get

$$\alpha \mathrm{TV} \leq \overline{R}_\rho^{\alpha,\beta}$$

for $u \in \mathrm{BV}(\Omega)$.
The second inequality $\overline{R}_\rho^{\alpha,\beta} \leq R_\rho^{\alpha,\beta}$ is an immediate consequence of equation (3.13). It hast to be mentioned that all functionals are equal to infinity if $u \notin \mathrm{BV}(\Omega)$ holds. □

We are now able to define the final minimisation problem that can be applied on d-dimensional imaging problems. For a given data term $G : L^{\frac{d}{d-1}}(\Omega) \to]-\infty,\infty[$ that is bounded from below, convex and lower semi continuous the minimisation problem is given by

$$\min_{u\in L^{\frac{d}{d-1}}(\Omega)} \quad G(u) + \overline{R}_\rho^{\alpha,\beta}(u). \tag{3.17}$$

To prove existence of the minimiser given in (3.17) we want to apply the theory provided in section 2.2. Therefore we rephrase [4, Theorem 4.7] to prove existence of (3.17).

Theorem 84. *Let $G : L^{\frac{d}{d-1}}(\Omega) \to]-\infty, \infty[$ be a data term that is bounded from below, convex and lower semi-continuous. Further $G(u^n) \to \infty$ should hold if*

$$\begin{cases} \left|\int_\Omega u^n \,\mathrm{d}x\right| \to \infty \text{ and} \\ \left\|u^n - |\Omega|^{-1} \int_\Omega u^n \,\mathrm{d}x\right\|_{\frac{d}{d-1}} \text{ bounded.} \end{cases}$$

The minimisation problem (3.17) *has a solution $u^* \in L^{\frac{d}{d-1}}(\Omega)$ for all $\alpha, \beta > 0$.*

Proof. For the proof we want to apply the direct method that was presented in Section 2.2 on the resulting objective function $F = G + \overline{R}_\rho^{\alpha,\beta}$.
Without loss of generality we want to assume that F is proper.
By assumption F is bounded from below, since G is bounded from below and Lemma 83 shows that $\overline{R}_\rho^{\alpha,\beta}$ is bounded form below as well.
In the proof of the direct method (Theorem 7) coercivity was used to get a bounded sequence. Here we make use of the assumptions on G and the Poincaré-Friedrichs inequality (see [9]) to show existence of a bounded subsequence that converges to u^* in a weak sense.
With Lemma 83 we get

$$\alpha \mathrm{TV} \leq \overline{R}_\rho^{\alpha,\beta} \leq F - \inf_{u \in L^2(\Omega)} G(u).$$

as well as the boundedness of $\{\mathrm{TV}(u^n)\}_n$. With the Poincaré-Friedrichs inequality we get the existence of $C < \infty$ such that

$$\left\|u^n - |\Omega|^{-1} \int_\Omega u^n \,\mathrm{d}x\right\|_{\frac{d}{d-1}} \leq C \sup_{n \in \mathbb{N}} \mathrm{TV}(u^n) < \infty.$$

Since G is bounded, the last inequality implies that $\int_\Omega u^n \,\mathrm{d}x$ has to be bounded as well.
Making use of the triangle inequality gives

$$\|u^n\|_{\frac{d}{d-1}} \leq \sup_{n \in \mathbb{N}} |\Omega|^{-\frac{d-1}{d}} \left|\int_\Omega u^n \,\mathrm{d}x\right| + \left\|u^n - |\Omega|^{-1} \int_\Omega u^n \,\mathrm{d}x\right\|_{\frac{d}{d-1}} < \infty.$$

This provides a weakly convergent subsequence that converges to $u^* \in L^{\frac{d}{d-1}}(\Omega)$.
Finally the lower semi-continuity of G and $\overline{R}_\rho^{\alpha,\beta}$ implies the lower semi-continuity of F as shown in Remark 8.
Together, Theorem 7 can be applied which proves existence of a solution $u^* \in L^{\frac{d}{d-1}}(\Omega)$.
□

Remark 85. *If the data term $G : L^{\frac{d}{d-1}}(\Omega) \to]-\infty, \infty[$ is strictly convex, the solution $u^* \in L^2(\Omega)$ is unique.*

In the following we want to take a closer look on data terms that arise for denoising and inpainting problems like it is shown in [4, Example 4.9] and [4, Example 4.10]. We will give numerical results for these two problems in the end of Chapter 4. The mathematical model of the denoising problem was already given in Section 2.1.

Theorem 86. *Let $u_0 \in L^p(\Omega)$ be given. The denoising problem*

$$\min_{u \in L^p} \frac{\lambda}{p} \int_\Omega |u - u_0|^p \, \mathrm{d}x + \overline{R}_\rho^{\alpha,\beta}(u)$$

for $1 \leq p \leq 2$ and $\lambda, \alpha, \beta > 0$ has a minimiser.

Proof. Define the data term

$$G(u) = \frac{\lambda}{p} \int_\Omega |u - u_0|^p \, \mathrm{d}x.$$

This functional is clearly bounded from below, convex and lower semi-continuous. Further $|\int_\Omega u^n \, \mathrm{d}x| \to \infty$ implies $G(u^n) \to \infty$, because

$$\|u^n - u_0\|_p^p \geq 2^{1-p} \|u^n\|_p^p - \|u_0\|_p^p \geq (2|\Omega|)^{1-p} \left| \int_\Omega u^n \, \mathrm{d}x \right|^p - \|u_0\|_p^p$$

holds.
Therefore existence of a solution of the denoising problem utilising Theorem 84 is given. □

Remark 87. *With Remark 85 the denoising problem is unique for $p > 1$.*

In the next theorem we want to show existence of the inpainting problem.

Inpainting problems deal with given image data u_0 on a non empty set $\Omega' \subset \Omega$. The task is to extend this given image data u_0 on the whole domain Ω.
For instance, this kind of problems arise if we want to construct an additional video frame between two given frames (Interpolation). Another application is image restoration, where an image supported on Ω is damaged and has to be reconstructed on the domain $\Omega' \setminus \Omega$.
The idea of inpainting methods is to fix the given image on the domain where the data is known. On the domain $\Omega' \setminus \Omega$ where no image data is given we want to use a regulariser to calculate a "natural" prolongation of the given image u_0.
This motivates the data term

$$G(u) = \begin{cases} 0 \text{ if } u|_{\Omega'} = f, \\ \infty \text{ else.} \end{cases} \tag{3.18}$$

Obviously, this data term fits our model in the context of the minimisation problem (3.17). Therefore the minimisation problem rewrites as

$$\min_{\substack{u \in L^{\frac{d}{d-1}}(\Omega) \\ u|_{\Omega'} = f}} \overline{R}_\rho^{\alpha,\beta}(u).$$

Theorem 88. *Let $f \in L^{\frac{d}{d-1}}(\Omega)$. The inpainting problem*

$$\min_{\substack{u \in L^{\frac{d}{d-1}}(\Omega) \\ u|_{\Omega'} = f}} \overline{R}_\rho^{\alpha,\beta}(u).$$

for $\alpha, \beta > 0$ has a minimiser.

Proof. The data term given in (3.18) is clearly proper, bounded from below and convex. Since $u(x) = f(x)$ on Ω' forms a closed subset of $L^{\frac{d}{d-1}}(\Omega)$ we get the lower semi-continuity.
To make use of Theorem 84 we have to show that $G(u^n)$ diverges if

$$\begin{cases} \left|\int_\Omega u^n \, \mathrm{d}x\right| \to \infty \text{ and} \\ \left\|u^n - |\Omega|^{-1} \int_\Omega u^n \, \mathrm{d}x\right\|_{\frac{d}{d-1}} \text{ is bounded.} \end{cases} \tag{3.19}$$

Let u^n be a given sequence where (3.19) and $u|_{\Omega'} = f$ holds for infinitely n. W.l.o.g. we assume that this holds for all $n \in \mathbb{N}$. Now we define $v^n := u^n - |\Omega|^{-1} \int_\Omega u^n \,\mathrm{d}x$ that has a bounded norm according to (3.19). Therefore we get with $\int_\Omega u^n \,\mathrm{d}x = |\Omega| \,(u^n - v^n)$ the inequality

$$\begin{aligned}
\sup_{n\in\mathbb{N}} \left| \int_\Omega u^n \,\mathrm{d}x \right| &= \sup_{n\in\mathbb{N}} \left| \frac{1}{|\Omega'|} \int_{\Omega'} \int_\Omega u^n \,\mathrm{d}x \,\mathrm{d}y \right| \\
&= \sup_{n\in\mathbb{N}} \left| \frac{1}{|\Omega'|} \int_{\Omega'} |\Omega| \,(u^n(y) - v^n(y)) \,\mathrm{d}y \right| \\
&= \sup_{n\in\mathbb{N}} \left| \frac{|\Omega|}{|\Omega'|} \int_{\Omega'} (f(y) - v^n(y)) \,\mathrm{d}y \right| \\
&\leq C \sup_{n\in\mathbb{N}} (\|f\|_{\frac{d}{d-1}} + \|v^n\|_{\frac{d}{d-1}}) < \infty.
\end{aligned}$$

This is a contradiction to $\left| \int_\Omega u^n \,\mathrm{d}x \right| \to \infty$ and therefore $u|_{\Omega'} = f$ can only hold for a finite number of n. This implies $G(u^n) \to \infty$. Therefore Theorem 84 can be applied, which provides a minimiser of the inpainting problem. □

4. Numerical Realisation

To apply the developed theory to real image data we have to think of a way to discretise this functional lifting approach, since Chapter 3 deals with continuous problems only. We will see that conventional discretisation ideas fail at some points due to the special structure of our model. Therefore we will provide different approaches to overcome this problems. Further we will rewrite the upcoming formulation to saddle-point problems to be able to apply the methods derived in Section 2.6. At the end of this chapter we will apply the derived discretisation to denoising and inpainting imaging problems. The results achieved using the functional lifting approach will be compared to results obtained using a TV-regulariser.
In this chapter all considerations are done for the dimension $n = 3$, if not mentioned differently.
Discretisations for the case $n = 2$ are given in the Appendix. There, we will also discuss further ideas that will make the given approaches more sensible according to the discretisation of the different gradient directions.
In a discrete setting Definition 77 in combination with a data term $G(u)$ can be rewritten as

$$\min_{(u,\mu)\in M_\nabla} \max_{\psi\in M_\rho} G(u) + \alpha \sum_{\Omega\times\mathbb{P}^n} |\mu| + \beta\langle B\mu, \psi\rangle, \tag{4.1}$$

where B is a discrete version of ∇_{T_ϑ} established in Definition 67. In the following sections we will discuss discrete versions of the sets M_∇, M_ρ and the operator ∇_{T_ϑ}.

4.1. Discretisation of M_∇

The set

$$M_\nabla = \Bigg\{ (u,\mu) \in L^1(\Omega) \times \mathcal{M}(\Omega \times \mathbb{P}^n) \,\Bigg|\, \int_{\Omega\times\mathbb{P}^n} \varphi(x)\cdot j^+(\tau)\,\mathrm{d}\mu + \int_\Omega u \cdot \mathrm{div}\varphi\,\mathrm{d}x = 0,\ \forall \varphi \in C_C^\infty(\Omega,\mathbb{R}^n) \Bigg\},$$

as derived in (3.14), can be interpreted as an additional term in the cost function by using the indicator functional. The indicator functional $I : X \to \{0,\ \infty\}$ of a subset $C \subset X$ is defined as

$$I_C(u) = \begin{cases} 0 & \text{if } u \in C \\ \infty & \text{else.} \end{cases}$$

This leads to

$$\begin{aligned} \min_{(u,\mu)\in L^1(\Omega)\times\mathcal{M}(\Omega\times\mathbb{P}^n)} F\begin{pmatrix} u \\ \mu \end{pmatrix} &= I_{\left\{\int_{\Omega\times\mathbb{P}^n} j^+(\tau)\,\mathrm{d}\mu - \int_\Omega \nabla u\,\mathrm{d}x = 0\right\}}(u,\mu) \\ &= I_{\left\{\int_{\Omega\times\mathbb{P}^n} j^+(\tau)\mu_0(x,\tau)\,\mathrm{d}(x,\tau) - \int_\Omega \nabla u\,\mathrm{d}x = 0\right\}}(u,\mu). \end{aligned}$$

This Functional interpreted in a discrete setting leads to

$$\min_{(u,\mu)\in\mathbb{R}^{M\cdot N\cdot T}\times\mathbb{R}^{M\cdot N\cdot T\cdot R}} F\begin{pmatrix} u \\ \mu \end{pmatrix} = I_{\{Au - D\mu = 0\}}(u,\mu), \tag{4.2}$$

and can be rewritten as a saddle-point formulation using Lagrange multipliers,

$$\min_{(u,\mu)\in\mathbb{R}^{M\cdot N\cdot T}\times\mathbb{R}^{M\cdot N\cdot T\cdot R}} \max_{\varphi\in\mathbb{R}^{M\cdot N\cdot T\cdot 3}} L\left(\begin{pmatrix} u \\ \mu \end{pmatrix}, \varphi\right) = \langle Au - D\mu, \varphi\rangle. \tag{4.3}$$

Here, the operator A is a discrete version of ∇u and D is a discrete version of $\int_{\mathbb{P}^n} j^+(\tau)\mu_0(x,\tau)\,\mathrm{d}\tau$.

Therefore the minimisation problem (4.1) can be rewritten as

$$\min_{u,\mu} \max_{\varphi,\psi\in M_\rho} G(u) + \alpha \sum_{\tilde{\Omega}\times\tilde{\mathbb{P}}^n} |\mu| + \beta\langle B\mu, \psi\rangle + \langle Au - D\mu, \varphi\rangle \tag{4.4}$$

where $\tilde{\Omega}$ is a discrete version of the domain Ω and $\tilde{\mathbb{P}}^n$ is a discrete version of the quotient space $\mathbb{P}^n$.

The first thought that comes to mind is to choose A as a standard forward difference scheme

$$(Au)^j_{i_1,\dots,i_j,\dots,i_N} = \begin{cases} \frac{u_{i_1,\dots,i_j+1,\dots,i_n} - u_{i_1,\dots,i_j,\dots,i_n}}{h} & \text{if } i_j < N_j \\ 0 \text{ else,} \end{cases}$$

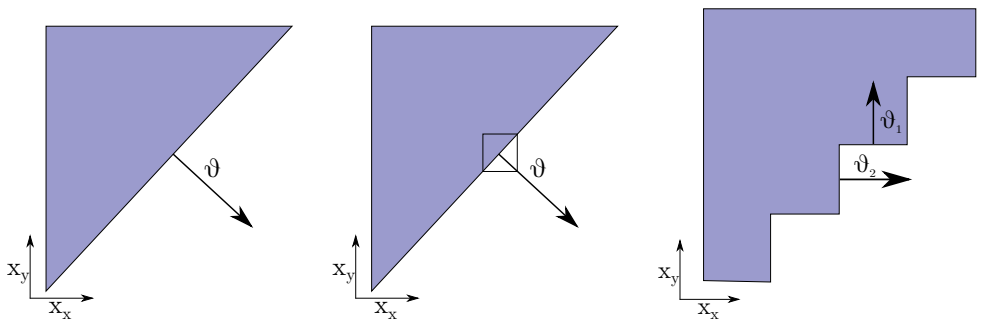
where h denotes the length of one voxel edge and $1 \leq j \leq 3$.
$D\mu$ can be realised by

$$\sum_{\tau \in \tilde{\mathbb{P}}^n} j^+(\tau) \cdot \mu_{\Omega,\tau}.$$

However even for simple images this approach leads to problems. Consider the 2D example

$$u_0 = \begin{cases} 1 \text{ if } x_y > x_x, \\ 0 \text{ else.} \end{cases}$$

This example would result in a lifting that is concentrated on the line $x_x = x_y$ with a constant direction. On the other hand, the discrete version of this problem would result in a step function which would lead to a detection of two different directions as shown in Figure 4.1.



(a) Continuous Ramp (b) Zoom (c) Discrete Version

Figure 4.1.: Continuous and Discrete 2D ramp example, showing the problems arising if the common forward difference scheme is used. Note that $j^+(\vartheta_1), j^+(\vartheta_2) \in S^n_+$ hold and therefore $\text{sign}(\vartheta_1) = -1$ according to (3.1) follows for this example.

This leads to differences in numerical results compared to continuous expectations.

4.1.1. Pattern-supported design of M_∇

As a way out we propose the use of patterns to verify the condition

$$\int_{\Omega\times\mathbb{P}^n} \varphi(x)\cdot j^+(\tau)\,\mathrm{d}\mu + \int_\Omega u\cdot \mathrm{div}\varphi\,\mathrm{d}x = 0,\ \forall\varphi\in C_C^\infty(\Omega,\mathbb{R}^n).$$

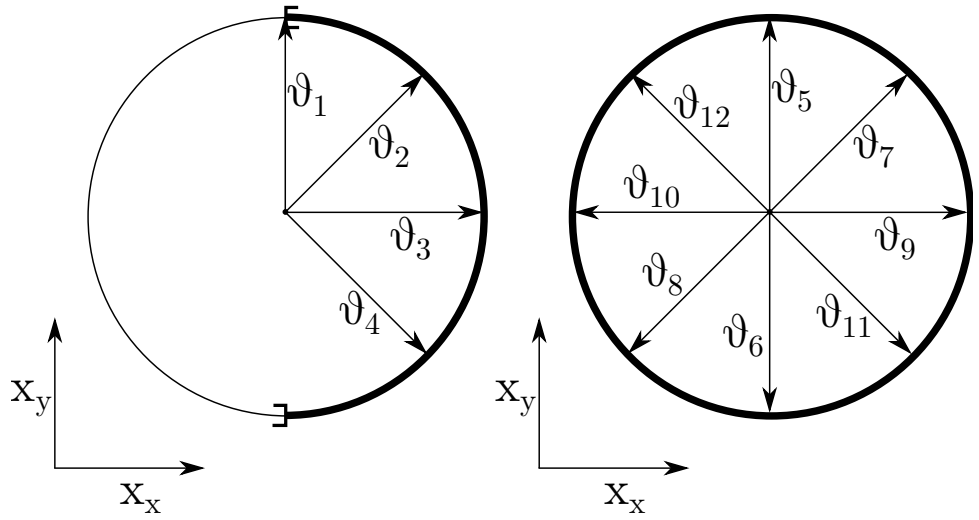
The idea is to design a small pattern for each discrete direction $\vartheta\in\mathbb{P}^n$ and match these patterns to the image u. It is useful to keep these patterns as small as possible in order to get a resulting u with sharp edges. Therefore we propose a pattern size of $2\times2\times2$ for problems in 3 dimensions. As a possible discretisation of $\mathbb{P}^n$ we choose two different velocities and a change in spatial direction of $\pi/4$.
The chosen discretisation of $\mathbb{P}^n$ will model the velocities 0 pixel shift per frame and 1 pixel shift per frame. We will only use this discretisation for the rest of this Chapter. An approach to get a generic refinement of the pattern-supported discretisation approach is missing jet and will not be discussed in this work.
It has to be mentioned that it is sufficient to discretise the set S^n_+ as defined in the beginning of Section 3.1 since the mapping $j: S^n\to\mathbb{P}^n$ is applied as defined in Definition 59.

Remark 89. *For ease of notation in the following formulations, $\vartheta\in\mathbb{P}^n$ will be often written as $\vartheta\in S^n_+$ which is the corresponding element induced by the mapping $j^+:\mathbb{P}^n\to S^n_+$.*

As a possible discretisation of $\vartheta\in S^n_+$ we give

$$\begin{array}{llll}
\vartheta_1=\begin{pmatrix}0\\1\\0\end{pmatrix} & \vartheta_2=\begin{pmatrix}1\\1\\0\end{pmatrix} & \vartheta_3=\begin{pmatrix}1\\0\\0\end{pmatrix} & \vartheta_4=\begin{pmatrix}-1\\1\\0\end{pmatrix}\\
\vartheta_5=\begin{pmatrix}0\\1\\1\end{pmatrix} & \vartheta_6=\begin{pmatrix}0\\-1\\1\end{pmatrix} & \vartheta_7=\begin{pmatrix}1\\1\\1\end{pmatrix} & \vartheta_8=\begin{pmatrix}-1\\-1\\1\end{pmatrix}\\
\vartheta_9=\begin{pmatrix}0\\1\\1\end{pmatrix} & \vartheta_{10}=\begin{pmatrix}0\\-1\\1\end{pmatrix} & \vartheta_{11}=\begin{pmatrix}1\\-1\\1\end{pmatrix} & \vartheta_{12}=\begin{pmatrix}-1\\1\\1\end{pmatrix}.
\end{array} \tag{4.5}$$

It has to be mentioned that these ϑ_i still needs to be normed to 1. A graphical interpretation of this discretisation is shown in Figure 4.2.

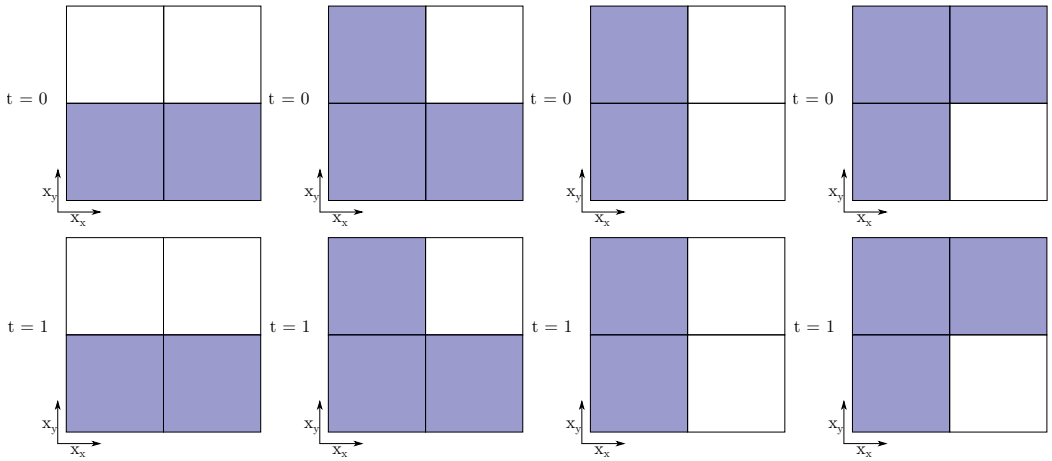


(a) Discretisation of S^n_+ for $x_t = 0$ (b) Discretisation of S^n_+ for $x_t = 1/\sqrt{2}$

Figure 4.2.: Discretisation of the half sphere S^n_+ that corresponds to the direction of the image gradient.

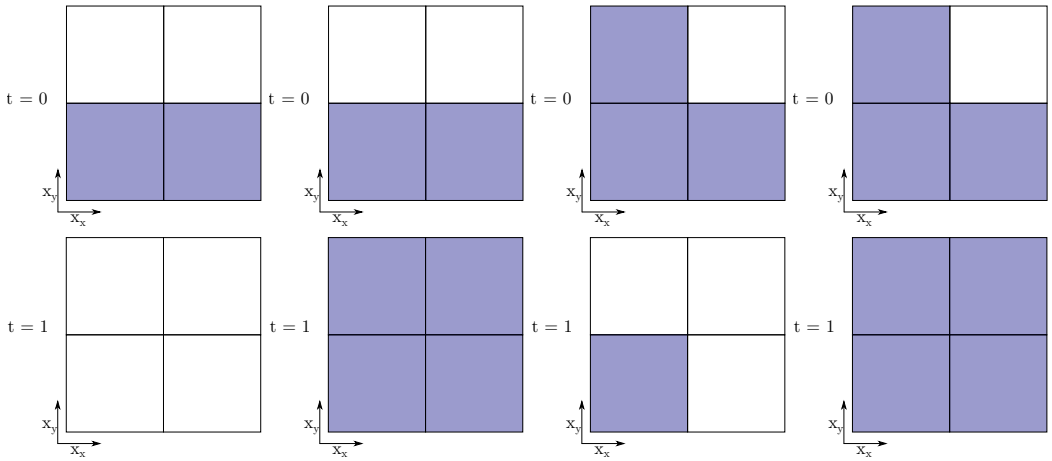
The corresponding Patterns are given in Figure 4.3, Figure 4.4 and Figure 4.5.

Now we want to use these patterns to build the desired image u. Therefore the operator A cuts, for each voxel of the given image u, the corresponding $2 \times 2 \times 2$ blocks. The operator D maps a linear combination of the defined patterns onto the blocks created from the operator A.



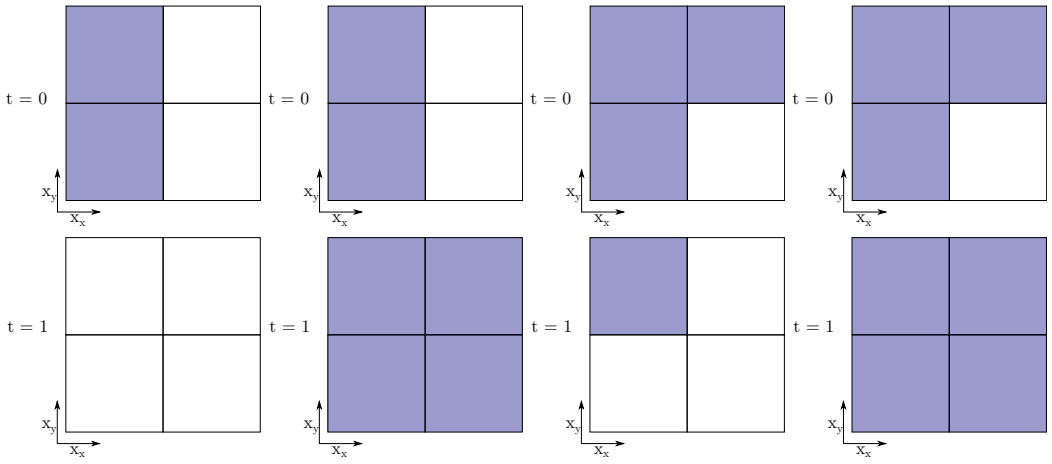
(a) ϑ_1 (b) ϑ_2 (c) ϑ_3 (d) ϑ_4

Figure 4.3.: Patterns corresponding to the directions ϑ_1 to ϑ_4, used for a pattern supported design of the set M_∇. Since those patterns are of size $2 \times 2 \times 2$ the two spatial coordinates at time step 1 are shown in the first line and the spatial coordinates at time step 2 are shown in the second line.



(a) ϑ_5 (b) ϑ_6 (c) ϑ_7 (d) ϑ_8

Figure 4.4.: Patterns corresponding to the directions ϑ_5 to ϑ_8. For a more detailed description see Figure 4.3.



(a) ϑ_9 (b) ϑ_{10} (c) ϑ_{11} (d) ϑ_{12}

Figure 4.5.: Patterns corresponding to the directions ϑ_9 to ϑ_{12}. For a more detailed description see Figure 4.3.

For a constant extension of u on the boundary this leads to the discrete operator $A \in \mathbb{R}^{8\cdot M\cdot N\cdot T \times M\cdot N\cdot T}$

$$
\begin{aligned}
(Au)^1_{i,j,t} &= u_{i,j,t}.\\
(Au)^2_{i,j,t} &= \begin{cases} u_{i+1,j,t} & \text{if } i < M \\ u_{i,j,t} & \text{else.} \end{cases}\\
&\vdots\\
(Au)^8_{i,j,t} &= \begin{cases} u_{i+1,j+1,t+1} & \text{if } i < M,\ j < N,\ t < T \\ u_{i,j+1,t+1} & \text{if } i = M,\ j < N,\ t < T \\ u_{i+1,j,t+1} & \text{if } i < M,\ j = N,\ t < T \\ u_{i,j,t+1} & \text{if } i = M,\ j = N,\ t < T \\ u_{i+1,j+1,t} & \text{if } i < M,\ j < N,\ t = T \\ u_{i,j+1,t} & \text{if } i = M,\ j < N,\ t = T \\ u_{i+1,j,t} & \text{if } i < M,\ j = N,\ t = T \\ u_{i,j,t} & \text{else.} \end{cases}
\end{aligned}
\tag{4.6}
$$

It hast to be mentioned that this careful treatment of the boundary is necessary to

minimise boundary effects. We chose a constant prolongation of the image on the boundary.
Further we get the operator $D \in \mathbb{R}^{8\cdot M\cdot N\cdot T\times M\cdot N\cdot T\cdot R}$ that maps μ according to the defined patterns onto the corresponding patch, where R is the number of discrete elements of $\mathbb{P}^n$ and is equal to the number of different patterns as well.
Let us denote the set of all patterns by Λ. The elements of Λ will be called p and have the index k which refers to the position within the $2\times 2\times 2$ block. The r-th pattern on position k is then written as $p_r(k)$, where r is the index of the corresponding pattern induced by the discretisation of $\vartheta \in \mathbb{P}^n$.

$$\begin{aligned}(D\mu)^1_{i,j,t} &= \sum_{r=1}^{R} p_r(1)\cdot \mu_{i,j,t,r}.\\ &\vdots\\ (D\mu)^8_{i,j,t} &= \sum_{r=1}^{R} p_r(8)\cdot \mu_{i,j,t,r}.\end{aligned}$$

To make this approach working we have to introduce an additional variable $\tilde{\mu} \in \mathbb{R}^{M\cdot N\cdot T}$ which corresponds to the mean value of u within a pattern. The corresponding pattern will be denoted by p_0 and has the constant value one (see Figure 4.6).

The needed operator, that maps the mean value on the $2\times 2\times 2$ blocks, will be denoted by $E \in \mathbb{R}^{8\cdot M\cdot N\cdot T\times M\cdot N\cdot T}$ and is defined as

$$(E\tilde{\mu})^l_{i,j,t} = p_0(l)\cdot \tilde{\mu}_{i,j,t} \qquad \text{for } 1 \le l \le 8. \tag{4.7}$$

Given this additional pattern $p_0(k)$, the patterns $p_r(k)$ for $r = 1,\dots,8$ are not influenced by the mean value of the cutted image and therefore only correspond to the edges. The separate treatment of the mean value is necessary, since we want penalise $|\mu|$ but not $|\tilde{\mu}|$.

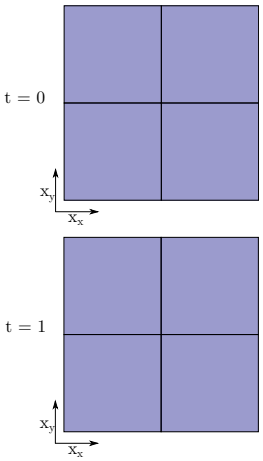


Figure 4.6.: Pattern for the average value.

4.1.2. Information Flow in this pattern-supported framework

First let us discuss which structure of μ we expect for a given image u. For ease of visualisation we want to discuss this on a non moving image that should only correspond to one pattern. Let u be given by

$$u = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

For an image like this, we want to have μ concentrated on the diagonal at the corresponding direction. For the discretisation described in Section 4.1.1 and in more detail in Figure 4.3 this would mean

$$\mu(:,:,2) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.8}$$

and zero everywhere else.
Anyway, the method introduced in section 4.1.1 will not jet deliver the desired result.

The patch that corresponds to $\mu(:,:,2)$ do fit nicely on the main diagonal, but the blocks on the secondary diagonal are not constant, as shown below.

$$u = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & \mathbf{1} & \mathbf{0} & 0 \\ 1 & 1 & \mathbf{1} & \mathbf{1} & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \qquad u = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & \mathbf{0} & \mathbf{0} & 0 \\ 1 & 1 & \mathbf{1} & \mathbf{0} & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

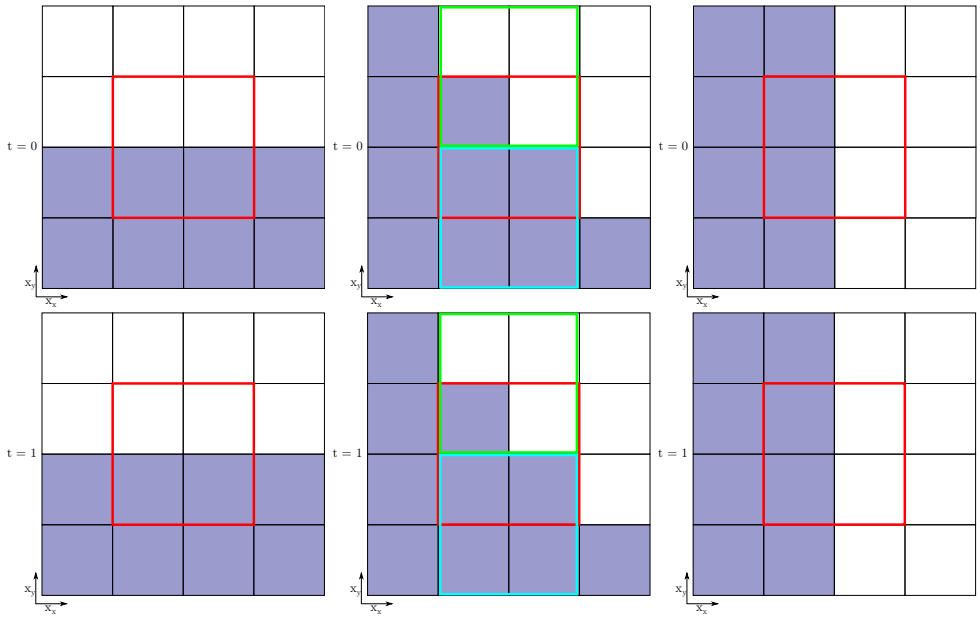
Since $Au = D\mu$, as described in (4.2), has to hold μ would not be zero on the secondary diagonal.
To get the desired μ, given in (4.8), a method to transport information in spatial direction has to be discussed. Therefore we introduce extended patterns that are enlarged in space. It has to be mentioned that extended patterns do not mean, that the block size of $2 \times 2 \times 2$ is changed. We just want to move the block in this extended pattern to get edge information from neighbouring μ.
Anyway, overlap effects will occur if the information is transported to all neighbouring μ. Therefore an additional mask χ has to be established. The extended patterns are shown in Figures 4.7, 4.8, 4.9 and 4.10. The mask is visualised with green, red and cyan coloured boxes.

Given this approach the operator D can be rewritten as

$$\begin{aligned} (D\mu)^1_{i,j,t} &= \sum_{r=1}^{R} \sum_{x_s=\{0,-1,1\}} \sum_{y_s=\{0,-1,1\}} \chi_r(x_s, y_s) \cdot p_r(1, x_s, y_s) \cdot \mu_{i+x_s, j+y_s, t, r}. \\ &\vdots \\ (D\mu)^8_{i,j,t} &= \sum_{r=1}^{R} \sum_{x_s=\{0,-1,1\}} \sum_{y_s=\{0,-1,1\}} \chi_r(x_s, y_s) \cdot p_r(8, x_s, y_s) \cdot \mu_{i+x_s, j+y_s, t, r}. \end{aligned} \tag{4.9}$$

with zero Dirichlet boundary conditions.
Given this approach and equation (4.7), the saddle point formulation (4.3) can be rewritten as

$$\min_{(u,\mu,\tilde{\mu}) \in \mathbb{R}^{M \cdot N \cdot T} \times \mathbb{R}^{M \cdot N \cdot T \cdot R} \times \mathbb{R}^{M \cdot N \cdot T}} \max_{\varphi \in \mathbb{R}^{M \cdot N \cdot T \cdot 8}} L \begin{pmatrix} u \\ \mu \\ \tilde{\mu} \end{pmatrix} = \langle Au - D\mu - E\tilde{\mu}, \varphi \rangle.$$



(a) Extended pattern that corresponds to ϑ_1.

(b) Extended pattern that corresponds to ϑ_2.

(c) Extended pattern that corresponds to ϑ_3.

Figure 4.7.: Extended patterns to ϑ_1, ϑ_2 and ϑ_3 with the corresponding masks χ. The first line represents the first time step and the second line represents the second timestep respectively. The masks are represented by the coloured boxes. Each colour represents one active mask. Which masks are active for a specific direction ϑ_r is decided by pure combinatorics.

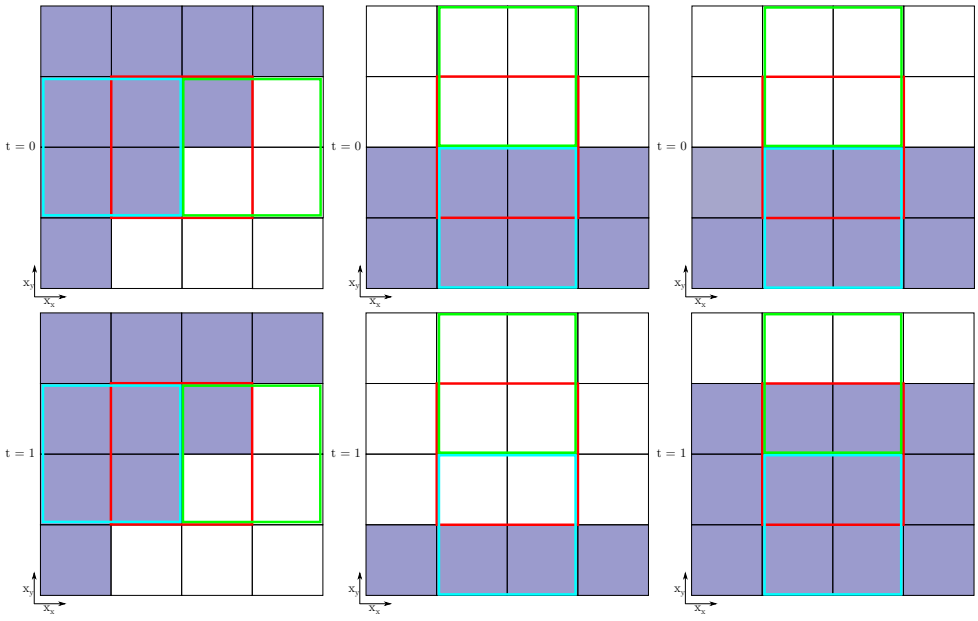
Further, the saddle point formulation (4.4) becomes

$$\min_{u,\mu,\tilde{\mu}} \max_{\varphi,\psi\in M_\rho} G(u) + \alpha \sum_{\tilde{\Omega}\times\tilde{\mathbb{P}}^n} |\mu| + \beta\langle B\mu, \psi\rangle + \langle Au - D\mu - E\tilde{\mu}, \varphi\rangle. \tag{4.10}$$

Numerical investigations have shown that a further relaxation of the set M_∇ leads to more stable results and faster convergence rates. Up to now we asked for strict equality of the necessary condition introduced in (3.14) which led to the indicator functional in (4.2) and to the saddle point formulations (4.3) and (4.10).
Better results were reached by relaxing the condition of M_∇ to

$$\|Au - D\mu - E\tilde{\mu}\|_{\infty,2} \leq C_{M_\nabla} \tag{4.11}$$

for some $C_{M_\nabla} > 0$ that is small enough.
The norm $\|\cdot\|_{\infty,2}$ is defined for K dimensional vectors x on the discrete image domain



(a) Extended pattern that corresponds to ϑ_4.

(b) Extended pattern that corresponds to ϑ_5.

(c) Extended pattern that corresponds to ϑ_6.

Figure 4.8.: Extended patterns to ϑ_4, ϑ_5 and ϑ_6 with the corresponding masks χ. For a detailed description see Figure 4.7.

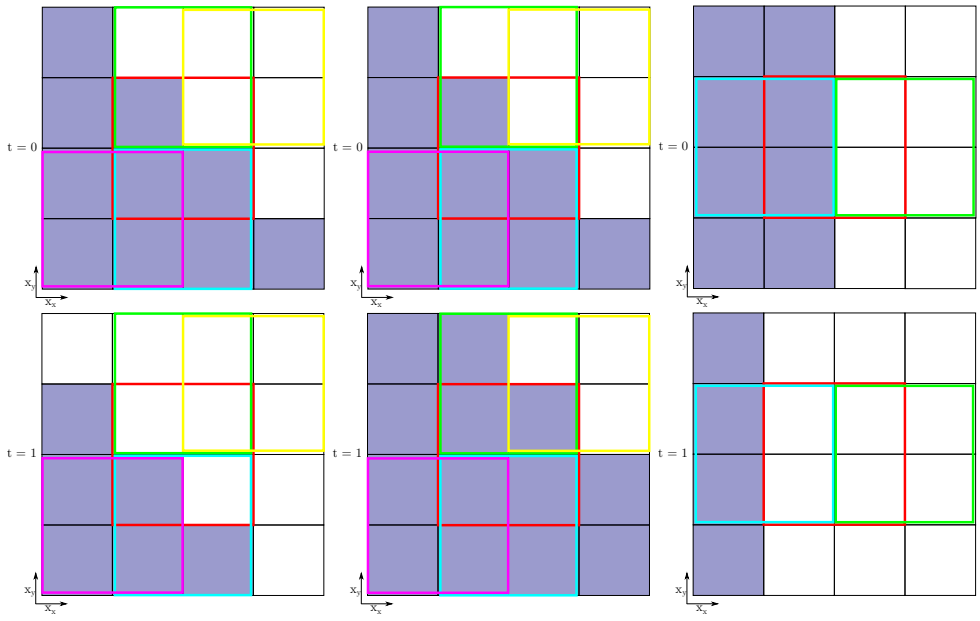
as

$$\|x\|_{\infty,2} = \left\| \sqrt{\sum_{i=1}^{K} x_i^2} \right\|_{\infty}.$$

Clearly, $K = 8$ is used in (4.11).

4.2. Discretisation of the directional derivative ∇_{T_ρ}

The next operator of (4.10) which has to be discussed is B as a suitable discretisation of the derivation along the tangential plane ∇_{T_ϑ}. Therefore we have to derive two linear independent vectors that span the tangential plane T_{ϑ_r} for $1 \leq r \leq R$. It has to be mentioned that the discrete directions ϑ_r where already fixed in (4.5).



(a) Extended pattern that corresponds to ϑ_7.

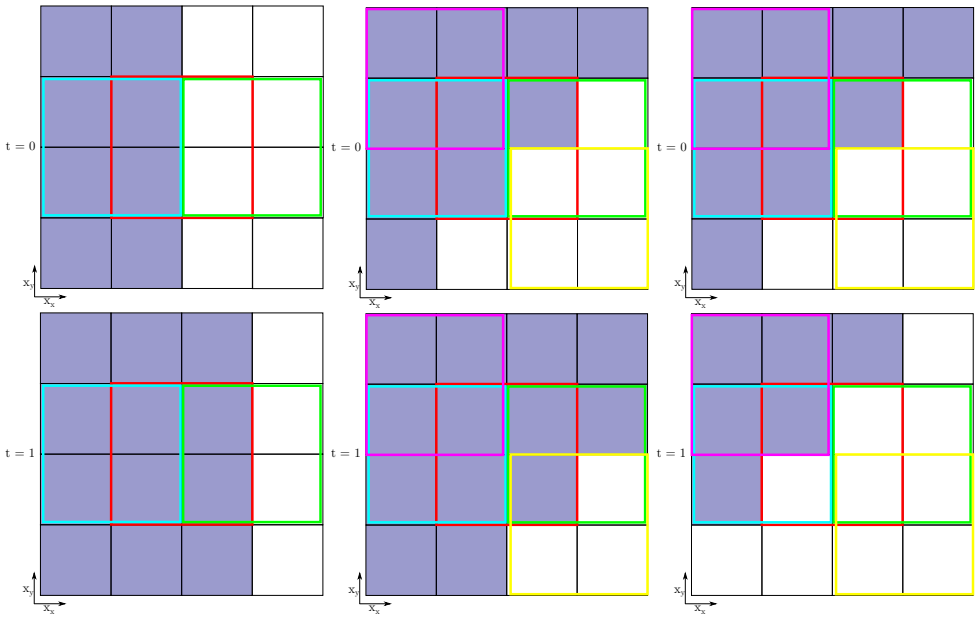
(b) Extended pattern that corresponds to ϑ_8.

(c) Extended pattern that corresponds to ϑ_9.

Figure 4.9.: Extended patterns to ϑ_7, ϑ_8 and ϑ_9 with the corresponding masks χ. For a detailed description see Figure 4.7.

Let us denote these two vectors by $t_{\vartheta_r,1}$ and $t_{\vartheta_r,2}$. Since $t_{\vartheta_r,1}$ and $t_{\vartheta_r,2}$ are not uniquely defined we want them to fulfil $t_{\vartheta_r,1} = (\cdot,\cdot,0)^T$ and $t_{\vartheta_r,2} = (\cdot,\cdot,1)^T$ as well as the necessary conditions $t_{\vartheta_r,1} \perp \vartheta_r$ and $t_{\vartheta_r,2} \perp \vartheta_r$.
For the discretisation of ϑ introduced in (4.5) we set

$$
\begin{aligned}
t_{\vartheta_1,1} &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad t_{\vartheta_2,1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad t_{\vartheta_3,1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad t_{\vartheta_4,1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\
t_{\vartheta_5,1} &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad t_{\vartheta_6,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad t_{\vartheta_7,1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad t_{\vartheta_8,1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\
t_{\vartheta_9,1} &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad t_{\vartheta_{10},1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad t_{\vartheta_{11},1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad t_{\vartheta_{12},1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},
\end{aligned} \tag{4.12}
$$



(a) Extended pattern that corresponds to ϑ_{10}.

(b) Extended pattern that corresponds to ϑ_{11}.

(c) Extended pattern that corresponds to ϑ_{12}.

Figure 4.10.: Extended patterns to ϑ_{10}, ϑ_{11} and ϑ_{12} with the corresponding masks χ. For a detailed description see Figure 4.7.

$$
\begin{aligned}
t_{\vartheta_1,2} &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad t_{\vartheta_2,2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad t_{\vartheta_3,2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad t_{\vartheta_4,2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
t_{\vartheta_5,2} &= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad t_{\vartheta_6,2} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad t_{\vartheta_7,2} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad t_{\vartheta_8,2} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\
t_{\vartheta_9,2} &= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad t_{\vartheta_{10},2} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad t_{\vartheta_{11},2} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad t_{\vartheta_{12},2} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.
\end{aligned} \tag{4.13}
$$

It can be argued that a restriction like $t_{\vartheta_r,1} = (\cdot,\cdot,0)^T$ with $t_{\vartheta_r,1} \perp t_{\vartheta_r,2} \perp \vartheta_r$ is even more consistent, but results showed that the necessary averaging of voxel-values lead

to loss of sharp edges.

The resulting operator $B \in \mathbb{R}^{2\cdot M\cdot N\cdot T\cdot R\times M\cdot N\cdot T\cdot R}$ is given by

$$\begin{aligned}(B\mu)^1_{i,j,t,r} &= \frac{\mu_{i+t_{\vartheta_r,1}(1),j+t_{\vartheta_r,1}(2),t+t_{\vartheta_r,1}(3),r} - \mu_{i,j,t,r}}{|t_{\vartheta_r,1}|}\\ (B\mu)^2_{i,j,t,r} &= \frac{\mu_{i+t_{\vartheta_r,2}(1),j+t_{\vartheta_r,2}(2),t+t_{\vartheta_r,2}(3),r} - \mu_{i,j,t,r}}{|t_{\vartheta_r,2}|}\end{aligned} \tag{4.14}$$

with Neumann boundary conditions.

4.3. Discretisation of M_ρ

The set M_ρ for $\rho = \rho_0$ according to Section 3.2.1 is given as

$$M_{\rho_0} = \left\{\psi \in C_C(\Omega\times\mathbb{P}^n,\mathbb{R}^n)\,\middle|\,\|\psi\|_{\infty,2} \le 1/2\right\}. \tag{4.15}$$

With the measure $\rho = \rho_1$ introduced in Section 3.2.2 the set M_ρ can be rewritten as

$$M_{\rho_1} = \left\{\psi \in C_C(\Omega\times\mathbb{P}^n,\mathbb{R}^n)\,\middle|\,\|\psi\|_{\infty,2} \le \alpha \wedge \|C\psi\|_{\infty,2} \le \beta\right\}, \tag{4.16}$$

where C is a suitable discretisation of ∇_ϑ.
In the following section we set $\alpha = \frac{1}{2}$ and $\beta = 1$.
To discretise the differentiation in label direction ∇_ϑ on $\mathbb{P}^n$ in a consistent manner we introduce a list that describes in which label directions the derivation is done. See therefore Table 4.1 and Figure 4.11.

Let $I(i,r)$ for $i = 1,2$ and $r = 1,\dots,R$ be the index mapping induced by Table 4.1. The operator $C \in \mathbb{R}^{2\cdot M\cdot N\cdot T\cdot R\times 2\cdot M\cdot N\cdot T\cdot R}$, as a discrete version of ∇_ϑ, can be written as

Origin	ϑ_1	ϑ_2	ϑ_3	ϑ_4	ϑ_5	ϑ_6	ϑ_7	ϑ_8	ϑ_9	ϑ_{10}	ϑ_{11}	ϑ_{12}
Direction 1	ϑ_5	ϑ_7	ϑ_9	ϑ_{11}	ϑ_7	ϑ_8	ϑ_9	ϑ_{10}	ϑ_{11}	ϑ_{12}	ϑ_6	ϑ_5
Direction 2	ϑ_2	ϑ_3	ϑ_4	ϑ_1	ϑ_1	ϑ_1	ϑ_2	ϑ_2	ϑ_3	ϑ_3	ϑ_4	ϑ_4

Table 4.1.: Table of derivation directions that are needed to model the discrete operator corresponding to the derivation in label direction.

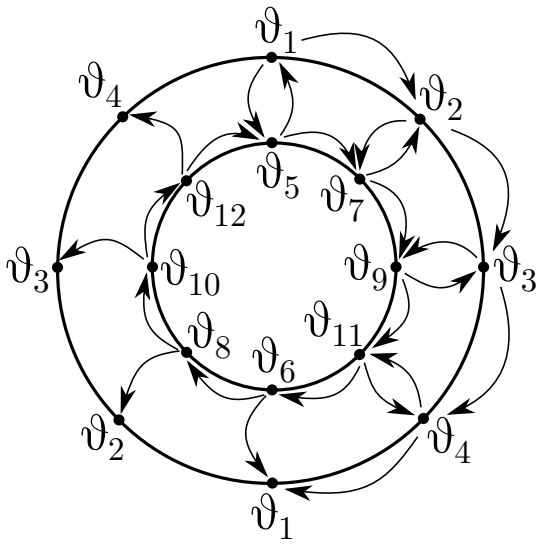


Figure 4.11.: Figure of derivation directions that are needed to model the discrete operator corresponding to the derivation in label direction.

$$
\begin{aligned}
(C\psi)^1_{i,j,t,r} &= \frac{\psi_{i,j,t,I(1,r)} - \psi_{i,j,t,r}}{\omega(1,r)}, \\
(C\psi)^2_{i,j,t,r} &= \frac{\psi_{i,j,t,I(2,r)} - \psi_{i,j,t,r}}{\omega(2,r)},
\end{aligned} \tag{4.17}
$$

where the weights $\omega(i,r)$ are calculated with

$$
\omega(i,r) = \arccos\left(\frac{\vartheta_r \cdot \vartheta_{I(i,r)}}{|\vartheta_r| \cdot |\vartheta_{I(i,r)}|}\right)
$$

as a distance measure between two elements $\vartheta_i \in S^n_+$.

4.4. Numerical Optimisation of the resulting Saddle-Point Formulation

Since every operator of (4.10) is discussed, we want to rewrite this formulation as

$$\min_{x\in X}\max_{y\in Y}\langle Kx,y\rangle+\mathcal{F}(x)-\mathcal{G}^*(y) \tag{4.18}$$

to make the theory established in Section 2.6 applicable. The same approach is used in [4]. In [16] a diagonal preconditioning method is presented that overcomes the need of calculating the norm $\|K\|$ as shown in Theorem 57.
According to [16] it is sufficient to computed the diagonal preconditioning matrices $T\in\mathbb{R}^{m\times m}$ and $S\in\mathbb{R}^{n\times n}$, induced by $K\in\mathbb{R}^{n\times m}$, that are given as

$$\mathcal{T}_{b,b}=\frac{1}{\sum_{a=1}^{n}|K_{a,b}|},$$

and

$$\mathcal{S}_{a,a}=\frac{1}{\sum_{b=1}^{m}|K_{a,b}|}.$$

If the denominator is equal to zero the corresponding value of the preconditioner is set to one. It has to be mentioned that these entries do not effect the result of the minimiser.
Further it has been shown in [5] that the convergence rate of (2.10) can be further improved by applying an additional overrelaxation step at the end. Therefore the implemented fixed point algorithm is given as

$$\begin{cases} x^{n+\frac{1}{2}} = prox_{\mathcal{T},\mathcal{F}}(x^n-\mathcal{T}K^Ty^n),\\ \hat{x}^{n+1} = 2x^{n+1}-x^n,\\ y^{n+\frac{1}{2}} = prox_{\mathcal{S},\mathcal{G}^*}(y^n+\mathcal{S}K\hat{x}^{n+1}),\\ x^{n+1} = x^{n+\frac{1}{2}}+\gamma\cdot\left(x^{n+\frac{1}{2}}-x^n\right),\\ y^{n+1} = y^{n+\frac{1}{2}}+\gamma\cdot\left(y^{n+\frac{1}{2}}-y^n\right), \end{cases} \tag{4.19}$$

with $\gamma \in [0, 1[$.
The perturbed proximity operators used in algorithm (4.19) are given as

$$prox_{T,\mathcal{F}}(\bar{x}) = \arg\min_{x \in X} \frac{1}{2}\langle \mathcal{T}^{-1}(x - \bar{x}), x - \bar{x}\rangle + \mathcal{F}(x), \tag{4.20}$$
$$prox_{S,\mathcal{G}^*}(\bar{y}) = \arg\min_{y \in Y^*} \frac{1}{2}\langle \mathcal{S}^{-1}(y - \bar{y}), y - \bar{y}\rangle + \mathcal{G}^*(y).$$

4.4.1. Saddle Point Formulation with $\mathrm{TVX}_0^{\alpha,\beta}$

The discrete version of the saddle point formulation (4.10) according to the functional $\mathrm{TVX}_0^{\alpha,\beta}$ is given by restricting M_ρ to M_{ρ_0} as shown in (4.15).
Therefore the saddle point formulation can be rewritten as

$$\min_{u,\mu,\tilde{\mu}} \max_{\varphi,\psi} G(u) + \alpha \sum_{\Omega \times \mathbb{P}^n} |\mu| + \beta\langle B\mu, \psi\rangle + \langle Au - D\mu - E\tilde{\mu}, \varphi\rangle, \tag{4.21}$$
$$\text{s.t.} \quad \|\psi_{i,j,t,r,\cdot}\|_2 \leq \frac{1}{2}.$$

The dimensions of the variables, for an image sequence of size $M \times N \times T$ and R discrete directions of $\mathbb{P}^n$, are $u \in \mathbb{R}^{MNT}$, $\mu \in \mathbb{R}^{MNTR}$, $\tilde{\mu} \in \mathbb{R}^{MNT}$, $\varphi \in \mathbb{R}^{8MNT}$ and $\psi \in \mathbb{R}^{2MNTR}$.
Since the algorithm introduced in section 2.6 operates on saddle point formulations of form (4.18) we define

$$x := (\mu, u, \tilde{\mu})^T$$

and

$$y := (\psi, \varphi)^T,$$

as well as the operator

$$K \in \mathbb{R}^{n \times m},$$

with $n = 8MNT + 2MNTR$ and $m = MNT + MNTR + MNT$ that is given as

$$K = \begin{pmatrix} \beta B & 0 & 0 \\ -D & A & -E \end{pmatrix}.$$

The Functionals $\mathcal{F}(x)$ and $\mathcal{G}^*(y)$ result in

$$\mathcal{F}(x) = \alpha \sum_{i,j,t,r} |\mu| + G(u),$$
$$\mathcal{G}^*(y) = \sum_{m,n,t,r=1}^{M,N,T,R} I_{\left\{\|\cdot\|_2 \leq \frac{1}{2}\right\}}\left(\psi_{m,n,t,r,\cdot}\right).$$

The perturbed proximity operators used in Algorithm (4.19) follow, by solving for the minimising function in (4.20), with

$$x^{n+\frac{1}{2}} = prox_{\mathcal{T},\mathcal{F}}(\bar{x}^n) \iff \begin{cases} \mu_{i,j,t,r}^{n+\frac{1}{2}} = \max\left(0,\ |\bar{\mu}_{i,j,t,r}^n| - \alpha \cdot (\mathcal{T}^\mu \cdot \mathbf{1})_{i,j,t,r}\right) \cdot \operatorname{sign}(\bar{\mu}_{i,j,t,r}^n) \\ u_{i,j,t}^{n+\frac{1}{2}} = prox_{\mathcal{T}^u,G}(\bar{u}_{i,j,t}^n) \\ \tilde{\mu}_{i,j,t}^{n+\frac{1}{2}} = \bar{\tilde{\mu}}_{i,j,t}^n \end{cases}$$

and

$$y^{n+\frac{1}{2}} = prox_{\mathcal{S},\mathcal{G}^*}(\bar{y}^n) \iff \begin{cases} \psi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \frac{\bar{\psi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ 2\cdot\|\bar{\psi}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \varphi_{i,j,t,l}^{n+\frac{1}{2}} = \max\left(-\frac{1}{2}, \min\left(\frac{1}{2}, \bar{\varphi}_{i,j,t,l}^n\right)\right). \end{cases}$$

The occurring variables are defined as $\bar{x}^n := x^n - \mathcal{T}K^T y^n$ and $\bar{y}^n := y^n + \mathcal{S}K\hat{x}^{n+1}$. The operator $\mathcal{T}^\mu$ corresponds to the induced block matrices of $\mathcal{T}$. Mention here again, that the preconditioning matrices $\mathcal{T}$ and $\mathcal{S}$ are diagonal matrices since this property was used to calculate the proximity operators.
Clearly, the proximity operator $prox_{\mathcal{T}^u,G}(\bar{u}_{i,j,t}^n)$ depends on the considered imaging problem. Examples for G will be given later.
If the relaxation step presented in (4.11) is implemented as well, the update of the dual variable changes to

$$y^{n+\frac{1}{2}} = prox_{\mathcal{S},\mathcal{G}^*}(\bar{y}^n) \iff \begin{cases} \psi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \frac{\bar{\psi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ 2\cdot\|\bar{\psi}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \varphi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \frac{\bar{\varphi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ \frac{\|\bar{\varphi}_{i,j,t,r,\cdot}^n\|_2}{C_{M_\nabla}}\right)}. \end{cases}$$

4.4.2. Saddle Point Formulation with $\mathrm{TVX}_1^{\alpha,\beta}$

The discrete version of the saddle point formulation (4.10) according to the functional $\mathrm{TVX}_1^{\alpha,\beta}$ is given by restricting M_ρ to M_{ρ_1} as shown in (4.16).
Therefore the saddle point formulation can be rewritten as

$$\min_{u,\mu,\tilde{\mu},\nu} \max_{\varphi,\psi,\zeta} G(u) + \alpha \sum_{\Omega\times\mathbb{P}^n} |\mu| + \beta\langle B\mu,\psi\rangle + \langle Au - D\mu - E\tilde{\mu},\varphi\rangle + \langle C\psi - \zeta,\nu\rangle, \tag{4.22}$$

$$\text{s.t.} \quad \|\psi_{i,j,t,r,\cdot}\|_2 \leq \frac{1}{2}, \quad \|\zeta_{i,j,t,r,\cdot}\|_2 \leq 1.$$

Here the dimensions of the variables u, μ, $\tilde{\mu}$, ψ and φ stay the same as described in section 4.4.1. The dimensions of the newly introduced variables are given by $\zeta \in \mathbb{R}^{2MNTR}$ and $\nu \in \mathbb{R}^{2MNTR}$.
Further we define

$$x := (\mu, \nu, u, \tilde{\mu})^T$$

and

$$y := (\psi, \zeta, \varphi)^T,$$

as well as the operator

$$K \in \mathbb{R}^{n\times m},$$

with $n = 8MNT + 2MNTR + 2MNTR$ and $m = MNT + 2MNTR + MNTR + MNT$ which is given as

$$K = \begin{pmatrix} \beta B & C^T & 0 & 0 \\ 0 & -I & 0 & 0 \\ -D & 0 & A & -E \end{pmatrix},$$

again following the scheme presented in [4].

The Functionals $\mathcal{F}(x)$ and $\mathcal{G}^*(y)$ of saddle point formulation (4.18) can be rewritten as

$$\mathcal{F}(x) = \alpha \sum_{i,j,t,r} |\mu| + G(u),$$
$$\mathcal{G}^*(y) = \sum_{m,n,t,r=1}^{M,N,T,R} I_{\left\{\|\cdot\|_2 \leq \frac{1}{2}\right\}}\left(\psi_{m,n,t,r,\cdot}\right) + \sum_{m,n,t,r=1}^{M,N,T,R} I_{\{\|\cdot\|_2 \leq 1\}}\left(\zeta_{m,n,t,r,\cdot}\right).$$

The perturbed proximity operators used in Algorithm (4.19) are given as

$$x^{n+\frac{1}{2}} = prox_{\mathcal{T},\mathcal{F}}(\bar{x}^n) \iff \begin{cases} \mu_{i,j,t,r}^{n+\frac{1}{2}} = \max\left(0,\ |\bar{\mu}_{i,j,t,r}^n| - \alpha \cdot (\mathcal{T}^\mu \cdot \mathbf{1})_{i,j,t,r}\right) \cdot \operatorname{sign}(\bar{\mu}_{i,j,t,r}^n) \\ \nu_{i,j,t,r,l}^{n+\frac{1}{2}} = \bar{\nu}_{i,j,t,r,l}^n \\ u_{i,j,t}^{n+\frac{1}{2}} = prox_{\mathcal{T}^u,G}(\bar{u}_{i,j,t}^n) \\ \tilde{\mu}_{i,j,t}^{n+\frac{1}{2}} = \bar{\tilde{\mu}}_{i,j,t}^n \end{cases}$$

and

$$y^{n+\frac{1}{2}} = prox_{\mathcal{S},\mathcal{G}^*}(\bar{y}^n) \iff \begin{cases} \psi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \dfrac{\bar{\psi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ 2 \cdot \|\bar{\psi}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \zeta_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \dfrac{\bar{\zeta}_{i,j,t,r,\cdot}^n}{\max\left(1,\ \|\bar{\zeta}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \varphi_{i,j,t,l}^{n+\frac{1}{2}} = \max\left(-\frac{1}{2}, \min\left(\frac{1}{2}, \bar{\varphi}_{i,j,t,l}^n\right)\right) \end{cases}$$

with $\bar{x}^n := x^n - \mathcal{T} K^T y^n$ and $\bar{y}^n := y^n + \mathcal{S} K \hat{x}^{n+1}$ as already given in the previous section.
If the relaxation step presented in (4.11) is implemented as well, the update of the dual variable changes to

$$y^{n+\frac{1}{2}} = prox_{\mathcal{S},\mathcal{G}^*}(\bar{y}^n) \iff \begin{cases} \psi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \dfrac{\bar{\psi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ 2 \cdot \|\bar{\psi}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \zeta_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \dfrac{\bar{\zeta}_{i,j,t,r,\cdot}^n}{\max\left(1,\ \|\bar{\zeta}_{i,j,t,r,\cdot}^n\|_2\right)} \\ \varphi_{i,j,t,r,\cdot}^{n+\frac{1}{2}} = \dfrac{\bar{\varphi}_{i,j,t,r,\cdot}^n}{\max\left(1,\ \frac{\|\bar{\varphi}_{i,j,t,r,\cdot}^n\|_2}{C_{M_\nabla}}\right)}. \end{cases}$$

4.5. Application to Image-Sequence Reconstruction

In this section we want to apply the introduced algorithm on denoising and inpainting imaging problems in the case of image sequences. For all examples we set $\alpha = 0.1$ and $\beta = 1$. We also used the relaxation of the set M_∇ as described in (4.11) with a constant $C_{M_\nabla} = 1$. The discretisation of the directions ϑ is done as described above.

In the case of inpainting problems we deal with a given image $f \in \mathbb{R}^{MNT}$ that is only given on some subset of $\Omega' \subset \Omega$ and unknown on $\Omega \setminus \Omega'$. Therefore we want to fix the image on the set Ω' and apply the regulariser on the set $\Omega \setminus \Omega'$. This leads to the data term

$$G(u) = \sum_{(i,j,t)\in\Omega'} I_{\{f_{i,j,t}\}}(u_{i,j,t}),$$

where I is again the indicator functional.
The proximity operator $prox_{\mathcal{T}^u,G}(\bar{u}^n_{i,j,t})$ is given as

$$u^{n+\frac{1}{2}}_{i,j,t} = \begin{cases} \bar{u}^n_{i,j,t} & \text{if } (i,j,t) \in \Omega \setminus \Omega', \\ f_{i,j,t} & \text{else.} \end{cases}$$

In Figure 4.12 and Figure 4.14 we see the results of the presented algorithms on inpainting examples as well as the solution achieved by a standard Total Variation regulariser like it was derived in Chapter 2. For a better impression of the temporal evolution of the image sequence we present an animated version of the results in Figure 4.13 and Figure 4.15.
We can see a clear improvement in terms of sharp edges, if the proposed method is used in comparison to the results of TV - inpainting. The results achieved by TVX_0 and TVX_1 are more or less comparable.

In the case of denoising problems we deal with a given image $f \in \mathbb{R}^{MNT}$ that is the sum of the ground truth f_0 and some noise function η and can be written as $f = f_0 + \eta$. In this example we want η to be Gaussian noise, what makes the squared l_2 norm a suitable data term. Thus, we set

$$G(u) = \frac{\lambda}{2}\|u - f\|_2^2.$$

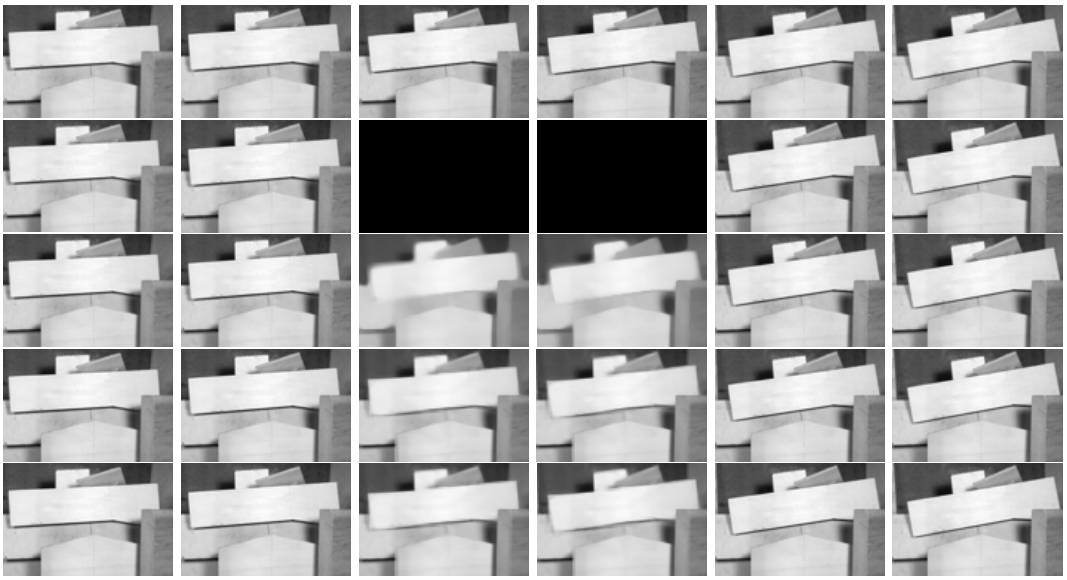


Figure 4.12.: **First row**: Ground truth from "http://vision.middlebury.edu/flow/data/". File: Wooden. **Second row**: Given data, where two central frames are lost. **Third row**: Inpainting using TV-regularisation. PSNR = 28.19 dB **Fourth row**: Inpainting using TVX_0. PSNR = 33.40 dB **Fifth row**: Inpainting using TVX_1. PSNR = 33.03 dB

The proximity operator $prox_{\mathcal{T}^u,G}(\bar{u}^n_{i,j,t})$ is given as

$$u^{n+\frac{1}{2}}_{i,j,t} = \frac{\bar{u}^n_{i,j,t} + \lambda \mathcal{T}^u_{i,j,t} f_{i,j,t}}{1 + \lambda \mathcal{T}^u_{i,j,t}}.$$

In Figure 4.16 and Figure 4.18 we present results on denoising problems. For a better impression of the temporal evolution of the image sequence we present an animated version of the results in Figure 4.17 and Figure 4.19.
We see again, that our proposed method is able to outmatch a usual TV regulariser. The data-term-constant λ was set to 15 for the TV results and $\lambda = 6$ was used for the TVX approaches. Especially, on thin contours our approach delivers way better results as a TV regulariser.

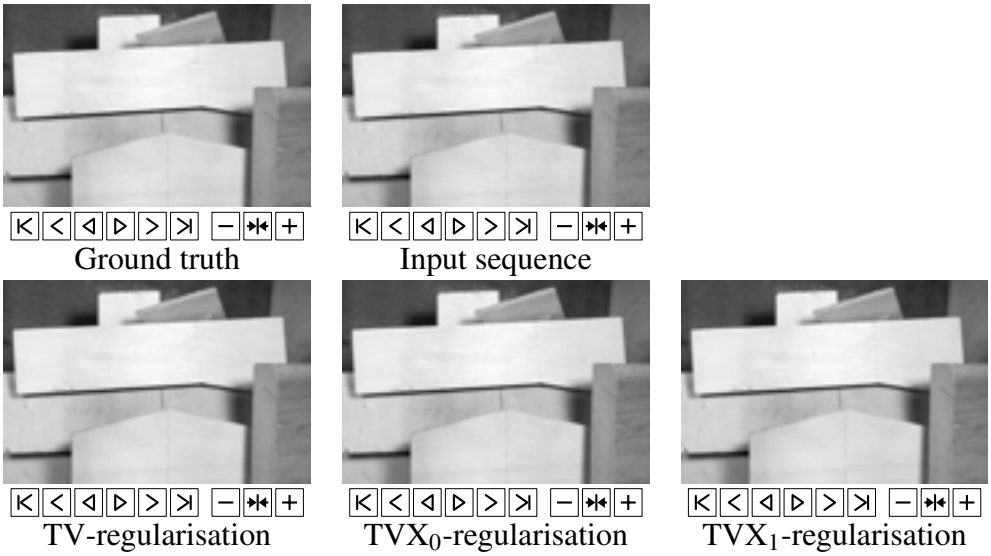


Figure 4.13.: Animated version of Figure 4.12, that can be played with a recent pdf viewer.

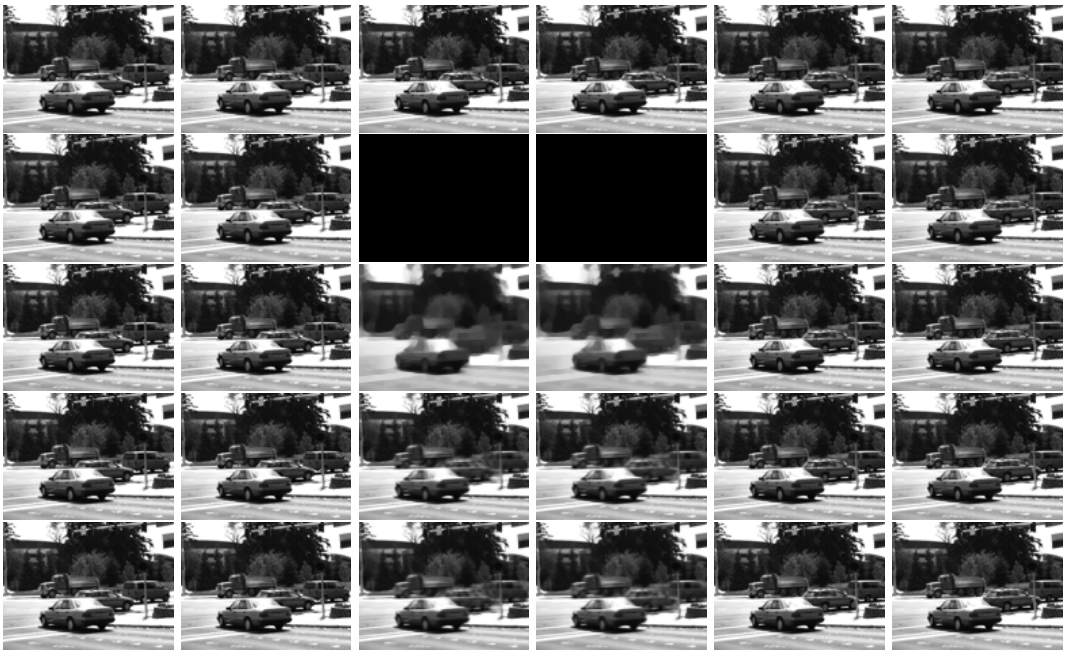


Figure 4.14.: **First row**: Ground truth from "http://vision.middlebury.edu/flow/data/". File: Dumptruck. **Second row**: Given data, where two central frames are lost. **Third row**: Inpainting using TV-regularisation. PSNR = 25.32 dB **Fourth row**: Inpainting using TVX_0. PSNR = 32.38 dB **Fifth row**: Inpainting using TVX_1. PSNR = 32.09 dB

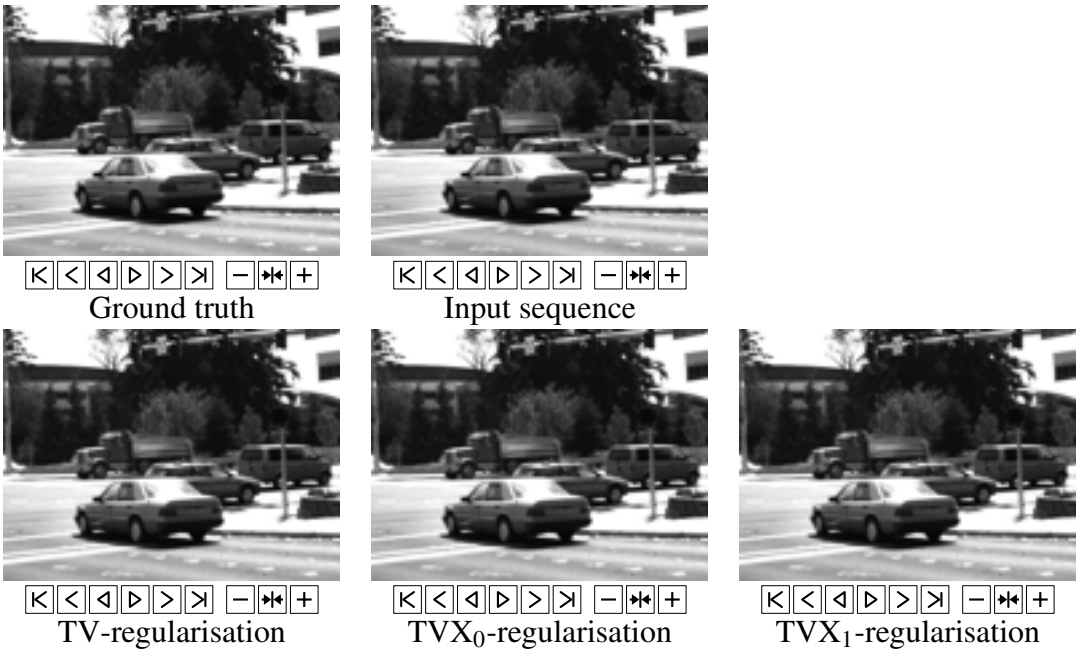


Figure 4.15.: Animated version of Figure 4.14, that can be played with a recent pdf viewer.

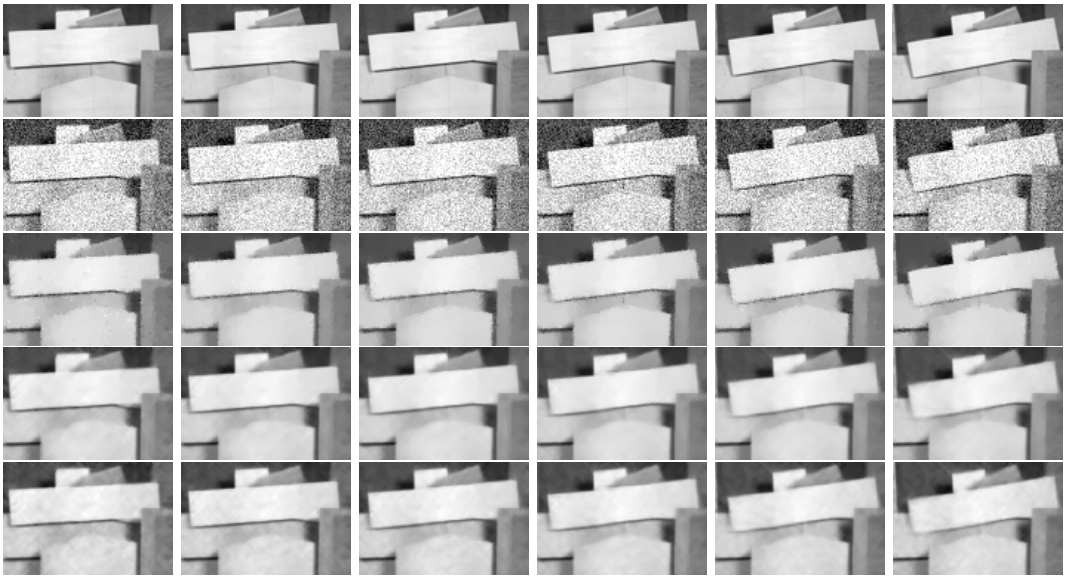


Figure 4.16.: **First row**: Ground truth from "http://vision.middlebury.edu/flow/data/". File: Wooden. **Second row**: Given data, with added Gausian noise with a standard deviation of $\sigma = 0.1$. PSNR = 20 dB**Third row**: Denoising using TV-regularisation. PSNR = 29.08 dB **Fourth row**: Denoising using TVX_0. PSNR = 29.30 dB **Fifth row**: Denoising using TVX_1. PSNR = 29.21 dB

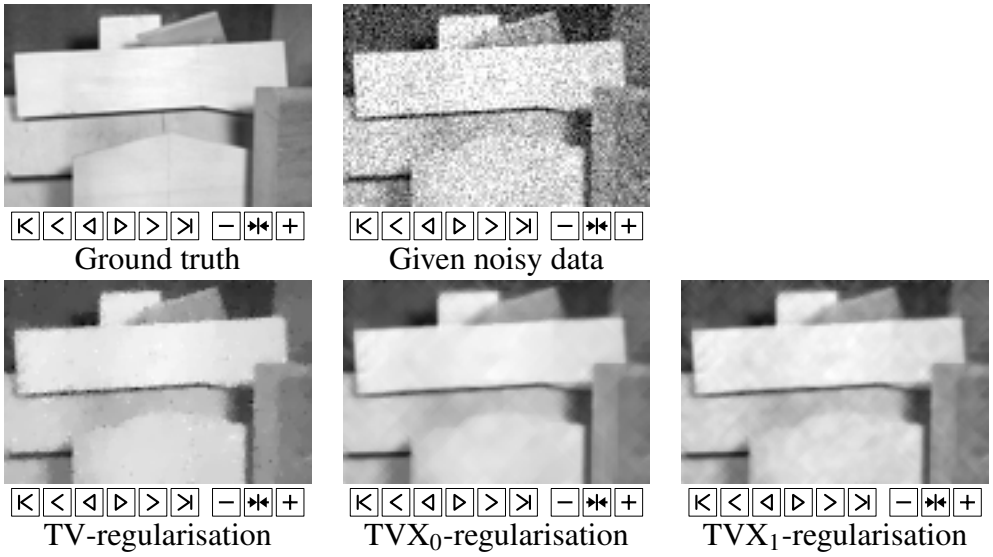


Figure 4.17.: Animated version of Figure 4.16, that can be played with a recent pdf viewer.

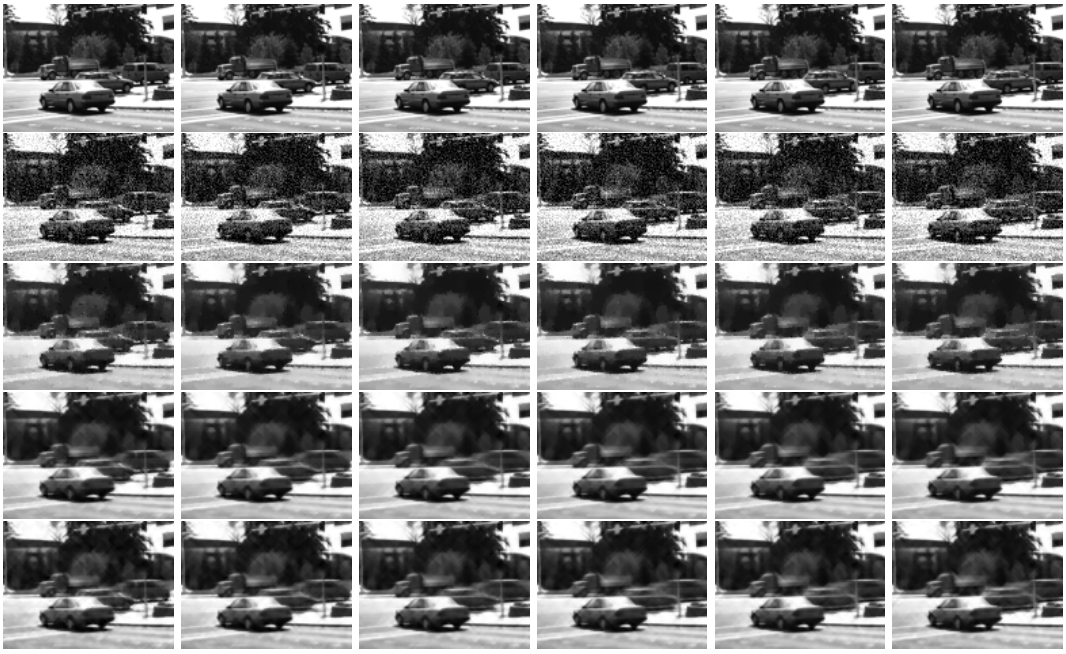


Figure 4.18.: **First row**: Ground truth from "http://vision.middlebury.edu/flow/data/". File: Dumptruck. **Second row**: Given data, with added Gausian noise with a standard deviation of $\sigma = 0.1$. PSNR = 20 dB**Third row**: Denoising using TV-regularisation. PSNR = 25.47 dB **Fourth row**: Denoising using TVX_0. PSNR = 25.64 dB **Fifth row**: Denoising using TVX_1. PSNR = 25.86 dB

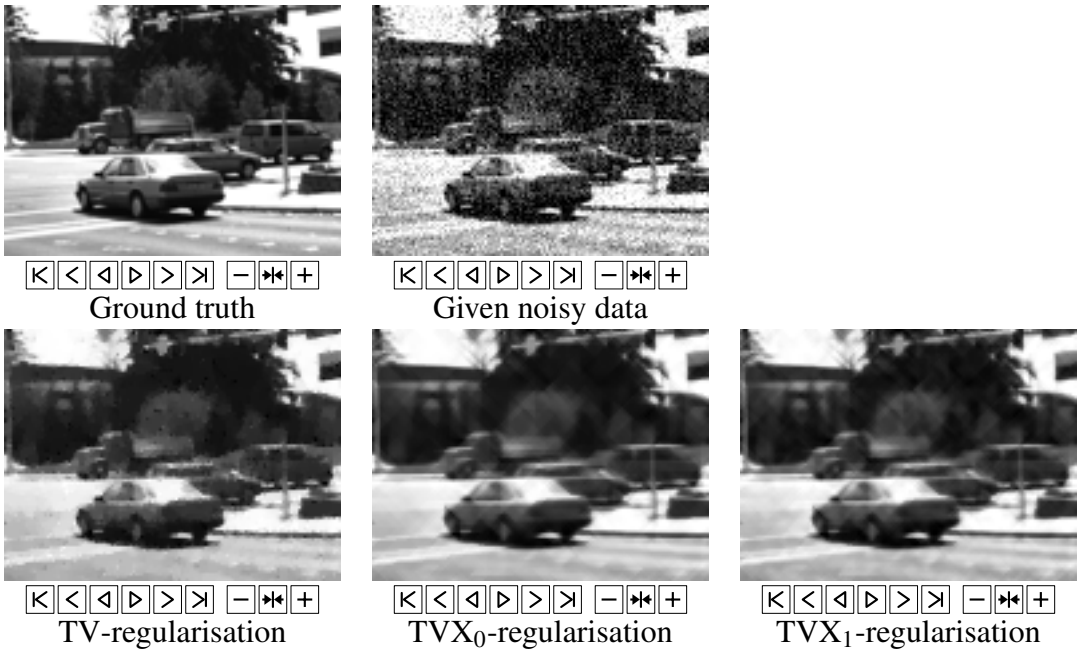


Figure 4.19.: Animated version of Figure 4.18, that can be played with a recent pdf viewer.

Appendix

Appendix A.

Discretisation in 2D

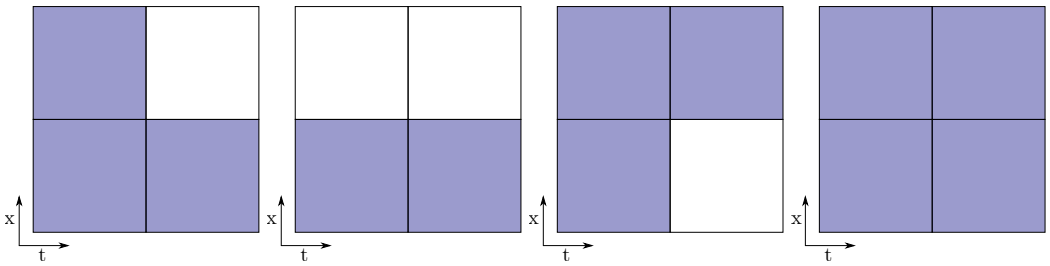
Here we want to provide methods to discretise the method derived in Chapter 3 also for the 2D case, since all numerics done in Chapter 4 are only applicable on 3D image data (image sequences). It has to be mentioned that we still discuss problems where one dimension is reserved for time, leading to the situation that 2D data has one spatial dimension and one time dimension. We want to point out again, that not only the higher order regularisation but also the special treatment of the temporal component of the problem is the main novelty in comparison to common approaches. We will see that the main parts of the numerical theory derived in Chapter 4 will stay more or less the same.
In the second part of the Appendix we want to discuss ideas how the pattern supported lifting approach can be adapted to deal with a finer resolution of the discrete directions of $\vartheta \in \mathbb{P}^n$. This finer discretisation methods will be derived for the 2D case.

A.1. Pattern-supported design of M_∇ in 2D

As already obtained in Section 4.1, numerical difficulties arise when standard difference schemes are used to describe the set M_∇. A graphical interpretation of this issue was already given in Figure 4.1.
As a way out, pattern-supported design of M_∇ was derived in Subsection 4.1.1. For the 2D case 2×2 patterns are used, to stay consistent with the theory of Subsection 4.1.1. In Subsection 4.1.2 the need of information flow in the spatial direction was discussed. As a sensible way to treat the flow of information, a mask was used to

describe the arising operator in an instructive way as you can see in equation (4.9). All these ideas work for the two dimensional case as well and will be presented in this section.

The patterns needed to cover the velocities $v = -1$, $v = 0$ and $v = 1$, as well as the additional pattern p_0, which deals with the mean value of a pattern, are given in Figure A.1.



(a) Pattern that corresponds to $v = -1$

(b) Pattern that corresponds to $v = 0$

(c) Pattern that corresponds to $v = 1$

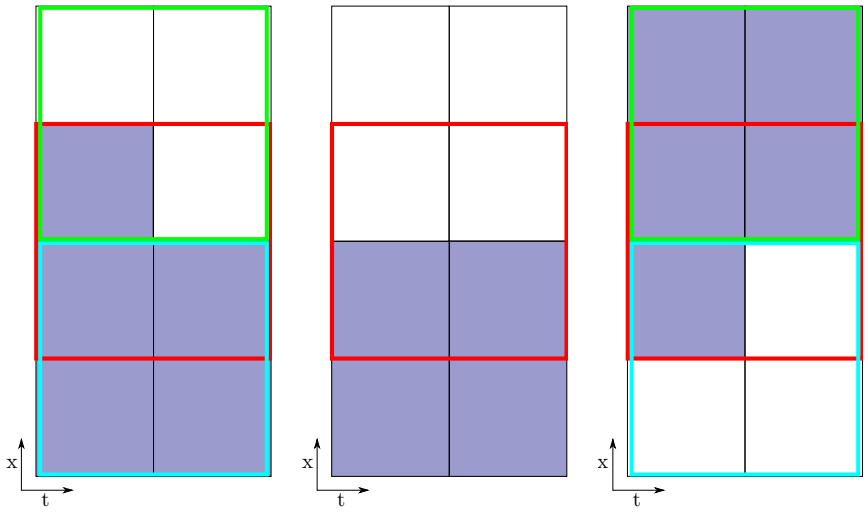
(d) Pattern that corresponds to the mean value

Figure A.1.: Visualisation of patterns used to model the set M_∇ in a 2D setting. The velocities modelled with the shown patterns are $v = \{-1, 0, 1\}$. The pattern that models the mean value is given as well.

The methods of spatial information flow, that were presented in the beginning of Subsection 4.1.2, have to be applied in the 2D case as well. Therefore we introduce the extended patterns and the resulting masks analogously to the 3D case. The extended patterns are given in Figure A.2. The masks are given with red, green and cyan coloured boxes.

It has to be mentioned that the green and cyan coloured masks for the pattern given in Figure A.2(b) can be neglected since they do not contain edge information.

Now we can give the operators A, B, C, D and E, for the 2D case. The corresponding 3D operators are given at (4.6), (4.14), (4.17), (4.9) and (4.7).
Let the image sequence u be of size $N \times T$ and let R be the number of discrete directions.



(a) Extended pattern and masks that correspond to $v = -1$

(b) Extended pattern and masks that correspond to $v = 0$

(c) Extended pattern and masks that correspond to $v = 1$

Figure A.2.: Visualisation of the extended patterns corresponding to $v = \{-1, 0, 1\}$. Further the masks, which are needed to control the information flow, are shown in red, green and cyan colour.

The operator $A \in \mathbb{R}^{4NT \times NT}$ that extracts the 2×2 patches out of the image can be rewritten as

$$
\begin{aligned}
(Au)^1_{i,t} &= u_{i,j,t}. \\
(Au)^2_{i,t} &= \begin{cases} u_{i+1,t} & \text{if } i < N \\ u_{i,t} & \text{else.} \end{cases} \\
(Au)^3_{i,t} &= \begin{cases} u_{i,t+1} & \text{if } t < T \\ u_{i,t} & \text{else.} \end{cases} \\
(Au)^4_{i,t} &= \begin{cases} u_{i+1,t+1} & \text{if } i < N,\ t < T \\ u_{i,t+1} & \text{if } i = N,\ t < T \\ u_{i+1,t} & \text{if } i < N,\ t = T \\ u_{i,t} & \text{else.} \end{cases}
\end{aligned}
$$

The operator $D \in \mathbb{R}^{4NT \times NTR}$ that maps the patterns onto the extracted 2×2 patches

is given as

$$(D\mu)^1_{i,t} = \sum_{r=1}^{R} \sum_{x_{\text{shift}}=\{0,-1,1\}} \chi_r(x_{\text{shift}}) \cdot p_r(1, x_{\text{shift}}) \cdot \mu_{i+x_{\text{shift}},t,r}.$$

$$\vdots$$

$$(D\mu)^4_{i,t} = \sum_{r=1}^{R} \sum_{x_{\text{shift}}=\{0,-1,1\}} \chi_r(x_{\text{shift}}) \cdot p_r(4, x_{\text{shift}}) \cdot \mu_{i+x_{\text{shift}},t,r},$$

with Dirichlet boundary conditions.
The operator $E \in \mathbb{R}^{4NT \times NT}$ which maps the mean value $\tilde{\mu}$ onto the 2×2 patches consequently can be rewritten as

$$(E\tilde{\mu})^l_{i,t} = p_0(l) \cdot \tilde{\mu}_{i,t} \qquad \text{for } 1 < l < 4.$$

In the two dimensional case the implementation of the operator B, which corresponds to the directional derivative, is far easier. Here the tangential plane presented in Chapter 3 is a one dimensional object that is simply given by $\vartheta^{\perp}$, like it was used in [4]. We will give a discrete version of the directional derivative for the discretisation presented in Figure A.1. The three resulting derivative directions $t_r = (x,t)^t$ are given as

$$t_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad t_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad t_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The operator $B \in \mathbb{R}^{NTR \times NTR}$ can be rewritten as

$$(B\mu)_{i,t,r} = \frac{\mu_{i+t_r(1),t+t_r(2),r} - \mu_{i,t,r}}{|t_r|}.$$

Finally we define the operator $C \in \mathbb{R}^{NTR \times NTR}$ that corresponds to the derivative in label direction and is needed to set up the regulariser TVX_1.

$$(C\mu)_{i,t,r} = \begin{cases} \mu_{i,t,r+1} - \mu_{i,t,r} & \text{if } r < R, \\ \frac{\mu_{i,t,1} - \mu_{i,t,R}}{R-1} & \text{else.} \end{cases}$$

Now the theory presented in Section 4.4 can be applied.

A.2. Refinement of the directional discretisation in 2D using patterns

All the patterns presented in Sections 4.1 and A.1 deal with integer pixel shifts only. Here we want to derive the needed theory to get a finer resolutions of the discrete directions in the two dimensional case using our pattern-supported approach. We will see how this refinement step changes the structure of the discrete operators and how they have to be adapted to make the presented methods working.

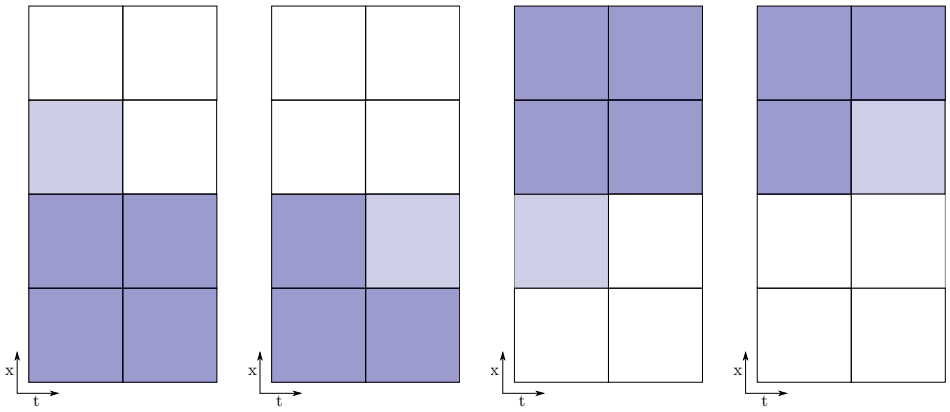
First let us give an example of an image u were an edge moves with the speed $v = -\frac{1}{2}$ and the desired discrete lifting μ.

$$u = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} \end{pmatrix} \qquad \mu_{v=-\frac{1}{2}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Now let us take a closer look on the patterns that appear, if 2×2 patches are extracted on the positions that are suggested by the given $\mu_{v=-\frac{1}{2}}$.

$$u = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \mathbf{\frac{1}{2}} & \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \mathbf{1} & \mathbf{1} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} \end{pmatrix} \qquad u = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & \mathbf{1} & \mathbf{\frac{1}{2}} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{1}{2} \end{pmatrix} \tag{A.1}$$

Here we see that two different patterns are needed to design the image created by a speed of $v = -\frac{1}{2}$. These two patterns are given in Figure A.3. In this Figure the extended patterns are already shown. Since the topic of spatial information flow stays the same, the theory of section A.1 has to be applied.



(a) First pattern that corresponds to $v = -\frac{1}{2}$.

(b) Second pattern that corresponds to $v = -\frac{1}{2}$.

(c) First pattern that corresponds to $v = \frac{1}{2}$.

(d) Second pattern that corresponds to $v = \frac{1}{2}$.

Figure A.3.: Visualisation of the extended patterns corresponding to $v = \left\{-\frac{1}{2}, \frac{1}{2}\right\}$.

Since there are two patterns corresponding to one speed, we have to remodel the linear operators B and D.
First let us mention, that the need of two patterns that correspond to only one speed leads to a dimensional increase of the lifted measure μ. For the case of half-valued speeds we get $\mu \in \mathbb{R}^{2NTR}$ instead of $\mu \in \mathbb{R}^{NTR}$.
The operator D that maps the patterns onto the 2×2 patches can be rewritten as

$$(D\mu)^1_{i,t} = \sum_{g=1}^{2} \sum_{r=1}^{R} \sum_{x_s=\{0,-1,1\}} \chi_{r,g}(x_s) \cdot p_{r,g}(1, x_s) \cdot \mu_{i+x_s,t,r,g}.$$

$$\vdots$$

$$(D\mu)^4_{i,t} = \sum_{g=1}^{2} \sum_{r=1}^{R} \sum_{x_s=\{0,-1,1\}} \chi_{r,g}(x_s) \cdot p_{r,g}(4, x_s) \cdot \mu_{i+x_s,t,r,g}.$$

Here the variable g iterates over the different patterns corresponding to the same speed. Since only one pattern is needed for integer speeds, we recommend to set the second pattern to constant zero. It has to be mentioned, that any other constant-valued

pattern will work as well.

In view of Figure A.3 we now want to derive the changes that have to be made in the calculation of the directional derivative (operator B). Since there are two patterns corresponding to the same speed we need two different differentiation directions for one speed.
For the speed $v = -\frac{1}{2}$, equation (A.1) and Figure A.3 suggest that the differentiation direction that corresponds to the first pattern, shown in Figure A.3(a), should have the direction $t = (0,1)^T$ and point on the second μ that corresponds to the speed $v = -\frac{1}{2}$. Vice versa the differentiation direction corresponding to the pattern shown in Figure A.3(b) has to be $t = (-1,1)^T$ and point to the first μ that corresponds to the speed $v = -\frac{1}{2}$.
Analogously the directional derivatives corresponding to the speed $v = \frac{1}{2}$ are given as $t = (0,1)^T$ and $t = (1,1)^T$. Therefore the operator $B \in \mathbb{R}^{NTR \times 2NTR}$ is given as

$$\begin{aligned}(B\mu)_{i,t,r} = &\frac{\mu_{i+t_{r,1}(1),t+t_r(2),r,2} - \mu_{i,t,r,1}}{|t_r|} \\ &+ \frac{\mu_{i+t_{r,2}(1),t+t_r(2),r,1} - \mu_{i,t,r,2}}{|t_r|}\end{aligned}$$

for non integer speeds and as

$$(B\mu)_{i,t,r} = \frac{\mu_{i+t_{r,1}(1),t+t_r(2),r,1} - \mu_{i,t,r,1}}{|t_r|}$$

for integer speeds. Mention that $t_{r,2}$ for integer valued speeds is not defined, since it is not used in the calculations.
Accordingly the theory presented in Section 4.4 can be applied to calculate a solution. In Figure A.4 we present results for the denoising problem with a finer discretisation of the directional patterns.

(a) Given noisy image (b) Result using a TV regularisation term (c) Result using a TVX_0 regularisation term

Figure A.4.: Results of a two dimensional denoising problem using non integer discretisation of the directional space.

List of Figures

Bibliography

[1] R. A. Adams and J. J. F. Fournier. *Sequences and Series in Banach Spaces*. Elsevier Ltd., 2^{nd} edition, 2003.

[2] A. Blake and A. Zisserman. Visual Reconstruction. *MIT Press*, 1987.

[3] K. Bredies and D. Lorenz. *Mathematische Bildverarbeitung*. Vieweg + Teubner Verlag, 1^{st} edition, 2011.

[4] K. Bredies, T. Pock, and B. Wirth. Convex Relaxation of a Class of Vertex Penalizing Functionals. *Journal of Mathematical Imaging and Vision*, 47:278 – 302, 11 2013.

[5] X. Cai, G. Gu, B. He, and X. Yuan. A relaxed customized proximal point algorithm for separable convex programming. Technical report, Hong Kong Baptist University, 2011.

[6] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.

[7] G. Citti and A. Sarti. A Cortical Based Model of Perceptual Completion in the Roto-Translation Space. *Journal of Mathematical Imaging and Vision*, 24(3):307–326, 2006.

[8] J. Diestel. *Sequences and Series in Banach Spaces*. Springer-Verlag, 1984.

[9] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. CRC Press, 1992.

[10] O. Forster. *Analysis 3. Integralrechnung im* $\mathbb{R}^n$ *mit Anwendungen*. Vieweg-Verlag, 4^{th} edition, 2007.

[11] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans Pattern Anal Mach Intell.*, 6:721 – 741, 1984.

[12] M. Giaquinta and S. Hildebrandt. *Calculus of Variations I.* Springer, 1996.

[13] G. Kanizsa. *Organization in Vision.* Praeger, New York, 1979.

[14] D. Mumford and J. Shah. Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems. *Commun. Pure Appl. Math.*, 42:577 – 685, 1989.

[15] A. Munk. *Maß- und Integrationstheorie*. Universitätsverlag Göttingen, 2011.

[16] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms. *International Conference of Computer Vision (ICCV 2011)*, pages 1762–1769, 2011.

[17] L. Rudin, S. J. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259 – 268, 1992.

[18] W. Rudin. *Real and Complex Analysis*. McGraw - Hill Inc., 1^{st} edition, 1970.

[19] D. Werner. *Funktionalanalysis*. Springer-Verlag, 6^{th} edition, 2007.

[20] E. Zeidler. *Applied Functional Analysis, Applications to Mathematical Physics*, volume 108. Springer-Verlag, 1995.