Leonie Knittelfelder, BSc.

Multiphase Segmentation with Convex Relaxation of Phase Functions

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Industrial Mathematics

submitted to

Graz University of Technology

Supervisor

a.o. Univ.-Prof. Mag. Dr. Stephen L. Keeling

Karl Franzens University of Graz, Institute for Mathematics and Scientific Computing

Graz, May 2016

AFFIDAVIT

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.

18.05.2016

Date

Signature

Abstract

This master's thesis presents a multiphase functional to segment and approximate a given gray scale image where the resulting image is smooth on the support of each characteristic function. Therefore, unique existence of the model functions is shown and then existence of a minimum of the functional with respect to functions in a convex relaxation of a set of characteristic functions is shown, i.e. $\mathrm{BV}(\Omega, \Delta_l)$.

This research puts a focus on image denoising and image segmentation simultaneously and uses both approaches in one functional. Moreover, since it is unsure if the functional is convex, a semi-gradient descent approach is established in a spatially continuous setting and later on in a finite dimensional setting. In order to prove certain features, a mollifying operator was introduced. Furthermore, a mapping that binds the update of the optimization process to have range in $[0, 1]$ was used.

The segmentation was performed with up to four characteristic functions, whereas segmentation with only two showed the best result, since $\chi_1 + \chi_2 = 1$. The work presented here has profound implications for future studies of a concurrent algorithm of image segmentation and denoising.

Contents

Chapter 1

Introduction

Think of your favourite image. It is easy for you to distinguish areas that have different colours or gray values. Even if it is disturbed by noise, in your head the disconnected pieces form together and you can still recognize what the colours tell you.

The two processes that were briefly described are image segmentation and image denoising - two important fields in the scientific world of image processing. Image segmentation is the process of combining the pixels of connected components. The number of applications considering this scientific tool is enormous. For instance, it is mainly used in medicine for automatic segmentation in computer tomography and magnet resonance tomography, but also for segmenting geological data such as satellite imagery. In addition, it is also needed in face recognition, inspection of work pieces and character recognition.

Whereas image denoising describes the procedure of reducing noise in a given image, i.e. the resulting one appears somewhat blurred – mathematically speaking, it shows signs of a smooth function that describes the underlying image.

This master thesis aims to combine image segmentation with image denoising. Therefore, a functional is proposed that depends on the model functions and on the characteristic functions, which together form the piecewise smooth approximation of the given image. The idea is to establish an algorithm, where model functions and characteristic functions are simultaneously obtained. In order to address this problem, in the first part a continuous dependence of the model functions on the characteristic functions is derived and other important features of the functional are discussed. The second part of this work consists in designing a proper algorithm. There are many different approaches, like primal dual methods with various regularizers, see [8].

Alas, for this particular functional with respect to characteristic functions it is not clear if it is convex, whereas it is with respect to the model functions, hence another optimizing strategy has to be investigated. However, a convex relaxation on the minimizing set is performed, which means, that minimization is not performed over the set of characteristic functions, but rather over a set of vector-valued functions that have range in $[0, 1]^l$ with the additional property that the sum of the function's components is one, where l denotes the number of segments of the corresponding characteristic functions. After the optimization process is completed, the algorithm relies on a heuristic rounding scheme that transforms the calculated relaxed functions into characteristic functions.

The rest of the thesis is organized as follows: In the second chapter, the basic concepts of image segmentation and image denoising are further elaborated, to fully grasp the meaning of a mathematical image. In addition, certain spaces and penalty terms are introduced which play an important role in image processing and some examples of established methods are presented.

Chapter three introduces a multiphase functional which combines image segmentation and image denoising of a gray scale image, such that out of the segmentation process the corresponding partitions are deduced. Furthermore, a proper algorithm is proposed to solve the image denoising and segmentation problem.

Chapter four analyses the proposed functional and gives proof of the existing minimizers for the relaxed constraint set. Therefore, continuity with respect to the model functions of the solution of the segmentational problem is proven. Moreover, an optimality system is deduced and so under certain conditions a simplified gradient of the functional is established. As a result a semi-implicit gradient descent algorithm is proposed and hence the existence of a fixed point is proven. All these results are given in a spatially continuous setting.

The fifth chapter deals with the discretization of the proposed method and shows that the spline approximations of the model functions and characteristic functions are consistent. At the end of this chapter, the discretized algorithm is presented.

In the sixth chapter, the numerical results are introduced. In addition, advantages and disadvantages of the method are elaborated and discussed.

Finally, in the Appendix, mathematical facts and important results that are used frequently throughout the Master Thesis are summarized.

Chapter 2

The Image Segmentation and Denoising Problem

2.1 Definition of a Mathematical Image

To begin with, to properly understand the term *image*, one has to comprehend that what is an easy task for the human brain, needs many difficult calculations for a computer. All objects on earth reflect light, which is projected onto the retina of the human eye. This data is processed by our brain immediately and we see an image without knowing the processes in the background. But what is an image in the mathematical sense? The following definition gives the first outlook into the world of image processing.

Definition 1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. A continuous image $u : \Omega \to \mathcal{F}$ is a bounded Lebesgue-measurable mapping in a colour space $\mathcal{F}$. For an image with continuous gray scales the corresponding colour space $\mathcal{F}$ is the interval $[0,1]$ or the space of real numbers $\mathbb{R}$, whereas for an image with continuous colours it is either $[0,1]^3$ or $\mathbb{R}^3$.*

Note that in this thesis only images in a spatially continuous setting are considered. Although, methods in Digital Image Processing rather deal with discrete images with discrete colour space. A more sophisticated way in mathematics is to tinker with spatially continuous images, like L^p-measurable functions.

Lemma 2.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let $1 \leq p \leq \infty$ and $u : \Omega \to \mathbb{R}$ be an image. Then $u \in L^p(\Omega)$.*

Proof. The image u is a bounded mapping, i.e., there exists a constant $M > 0$ with $|u(x)| \leq M$, for all $x \in \Omega$. So for $p = \infty$ follows immediately $\|u\|_{L^\infty(\Omega)} \leq M$ and for $1 \leq p < \infty$ the following holds

$$\int_\Omega |u|^p \ dx \leq |\Omega| \max_{x \in \Omega} |u(x)|^p \leq |\Omega| M^p.$$

□

Typically, $\Omega := (0,1)^2$ is the domain most often used in Mathematical Image Analysis. It will be also used in this thesis. An image $u \in L^p(\Omega)$ can even have more properties like the following one which will play an important role in this thesis.

Definition 2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let $u : \Omega \to \mathbb{R}$ be an image. Then the* Total Variation *of u is defined as follows:*

$$\mathrm{TV}(u) := \int_\Omega |\nabla u| \ dx = \sup \Big\{ \int_\Omega u \ \mathrm{div}\varphi \ dx : \ \|\varphi\|_{L^\infty} \leq 1, \ \varphi \in \mathcal{D}(\Omega, \mathbb{R}^n) \Big\}, \tag{2.1}$$

with the space of test functions $\mathcal{D}(\Omega, \mathbb{R}^n)$ mapping from the domain Ω to $\mathbb{R}^n$.

Lemma 2.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. The space containing functions with bounded Total Variation*

$$\mathrm{BV}(\Omega) := \left\{ u \in L^1(\Omega) : \ \mathrm{TV}(u) < \infty \right\}.$$

provided with the norm

$$\|u\|_{\mathrm{BV}(\Omega)} := \|u\|_{L^1(\Omega)} + \mathrm{TV}(u)$$

is a Banach space.

For the proof the reader is referred to the book of K. Bredies and D. Lorenz [4], Lemma 6.105. Note that the property of Total Variation admits discontinuities of the function u and that the gradient is understood more like a measure than a function.

2.2 A Brief Summary of Image Segmentation and Denoising

The purpose of segmentation is to decompose objects into certain parts. For example partition an image into foreground and background. One method to do so is approximating a given raw image whilst minimizing a functional. The other would be detecting edges appearing in the image, but this will not be discussed in this thesis. Therefore, the reader is referred to St. Fürtinger's Dissertation [3]. The following definition declares the meaning of segmentation in this thesis.

Definition 3. *Let $\tilde{u} : \Omega \to [0, 1]$ be a gray scale image. The segmentation is the approximation of $\tilde{u}$ with a function u, i.e.,*

$$\tilde{u}(x) \approx u(x) = \sum_{k=1}^{l} u_k(x)\chi_k(x), \quad x \in \Omega \tag{2.2}$$

where $l \in \mathbb{N}$ defines the number of phases. A segment of the image u is a connected component of the support of a given χ_k. Each $u_k : \Omega \to [0, 1]$ is a model function which smoothly approximates $\tilde{u}$ on the support of the characteristic function $\chi_k : \Omega \to \{0, 1\}$, which describes the segmented parts $\Omega_k := \{x \in \Omega : \chi_k(x) = 1\}$ of the domain. For all $i \neq k$ it follows $\Omega_k \cap \Omega_i = \emptyset$, so the characteristic functions $\{\chi_k\}$ have disjoint supports.

Now, a brief explanation of some functionals will give an overview of the segmentation idea in Mathematical Image Processing. Again, let $\tilde{u} : \Omega \to [0, 1]$ denote the given raw image. A computationally easy and established approach for segmentation is the *K-Means Clustering Algorithm*, which involves minimizing the following

$$\min_{p_k, \chi_k} \Big\{ \sum_{k=1}^{l} \int_\Omega |p_k \chi_k - \tilde{u}|^2 : \ \{p_k\} \in \mathcal{P}^0, \ \chi_k : \Omega \to \{0, 1\} \Big\},$$

where $\mathcal{P}^m$ denotes the space of polynomials with degree m. *K-Means* partitions the given image into l disjoint phases Ω_k, i.e. $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. In other words one pixel of the raw image is assigned to only one intensity cluster. For each cluster the intensity centroid is computed and the distance between a single pixel and the centroids determines whether the pixel belongs to a certain cluster or not. This easily leads to unnaturally disconnected segments and is unwanted at this point.

Figure 2.1: Segmentation done with the k-means algorithm. The input image is at the top left, followed by the resulting three characteristic functions.

The functional that serves as Status Quo in edge detection is the *Mumford-Shah Functional*, which involves minimizing the following

$$\min_{u,\Gamma} \Big\{ \int_\Omega |u - \tilde{u}|^2 + \ \delta^{-1} \int_{\Omega\setminus\Gamma} |\nabla u|^2 + \ \beta \mathcal{H}(\Gamma) \Big\},$$

where Γ denotes the one-dimensional contour set and $\mathcal{H}$ the one-dimensional Hausdorff measure. The segmentation is implicitly provided by Γ, which is a set of Lebesgue measure zero, i.e. $|\Gamma| = 0$. Note that in this functional u is already a piecewise smooth approximation of $\tilde{u}$ due to the penalty term $\delta^{-1} \int_{\Omega\setminus\Gamma} |\nabla u|^2$. The contour set Γ is excluded from the domain of integration, thus singularities of u are possible on Γ and hence u can be discontinuous on the edge set. Unfortunately, the two variables u and Γ are of a different kind, i.e. u is an element of one Banach space, whereas for Γ the structure of the other appropriate Banach space is unknown. Hence, existence of a minimizing pair (u^*, Γ^*) cannot be guaranteed.

A more general approach was done by J. Lellmann, [8], who introduced a variational convex formulation for multi-class labelling with different relaxations and regularizers like length-based, isotropic or separable ones, whereas minimizing the following is involved,

$$\inf_{u\in \mathrm{BV}(\Omega,\mathcal{E})} \int_\Omega \langle u(x), s(x)\rangle \ dx + J(u), \tag{2.3}$$

where $s \in L^\infty(\Omega)^l$ denotes the data term, J is the regularizer and $\mathcal{E} := \{e^1, ..., e^l\}$ denotes the set of unit vectors. This constraint set forces the solution u to have only discrete values and hence the solution attains a proper set of labelling functions. Unfortunately, due to the constraint set the problem is not convex, hence he proposes a relaxed constraint set, namely BV (Ω, Δ_l), where Δ_l indicates the unit simplex in $\mathbb{R}^l$, which will be used later in this work. Alas, this functional is not useful for this work because it only accomplishes the multi-class labelling part, whereas model functions with higher regularity than just constants are desired.

As for image denoising, often the minimization of the L^q-$H^{1,p}$-Denoising Functional for $1 < p \leq q < \infty$ is considered,

$$\min_{u\in L^q(\Omega)} \frac{1}{q} \int_\Omega |u - \tilde{u}|^q \ dx + \frac{\lambda}{p} \int_\Omega |\nabla u|^p \ dx. \tag{2.4}$$

This problem has a unique solution, see [4], Theorem 6.84. If $n = 2$ and $\Omega \in \mathbb{R}^n$ the solution u is already continuous in the interior of Ω by the *Sobolev Imbedding Theorem*, see Appendix (A.10).

Hence, if the reconstruction of images with discontinuities is desired, the above functional is not optimal. However, taking the limit $p \to 1$ the reconstructed images are less blurred. In this case using the *Total Variation* as the penalty term seems to be the better choice.

Figure 2.2: Noice reduction of a given image computed with the functional (2.4) for $q, p = 2$ and $\lambda = 0.5e^{-4}$.

Nevertheless, in this thesis the focus lies not on the proper reconstruction of a noisy image, but rather on the segmentation and therefore generating a piecewise smooth approximation of a given raw image. Thus, we will further explore the following functional, which has already been used to some extent and in a slightly different form in my Bachelor's Thesis [15].

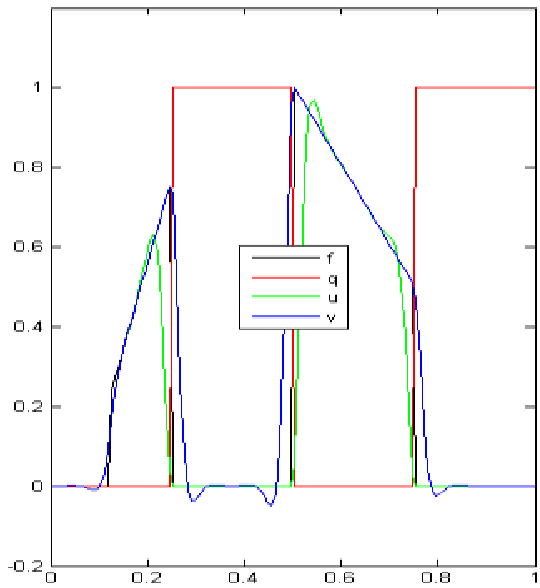

Figure 2.3: A multiphase segmentation of a one dimensional signal, see [15], where f denotes the input signal, q is the computed characteristic function and u and v are the corresponding model functions.

$$\mathcal{J}(u,\chi) = \begin{cases} \sum_{k=1}^{l} \int_\Omega \left[|u_k - K_\eta \tilde{u}|^2 (\beta K_\eta \chi_k + \delta) \right. \\ \qquad \left. + |\nabla^m u_k|^2 (\alpha K_\eta \chi_k + \varepsilon) \right] \, dx, & (u,\chi) \in H^m(\Omega)^l \times L^\infty(\Omega)^l, \\ \qquad \infty, & \text{otherwise} \end{cases} \tag{2.5}$$

where $u := (u_1, u_2, ..., u_l)^\top \in H^m(\Omega)^l$ denotes the vector of model functions and $\chi := (\chi_1, ..., \chi_l)^\top \in L^\infty(\Omega)^l$ the vector of multiple characteristic functions, with parameters $0 < \varepsilon, \delta \ll 1$ and $\alpha, \beta > 0$ and $m = 1, 2$ and K_η denotes a mollifying operator, which is only a technical necessity. Details considering this operator will be presented in the upcoming chapter. Together they build the smoothed approximation of the raw image, i.e. $\tilde{u} \approx \sum_{k=1}^{l} u_k \chi_k$.

Because of the parameters ε and δ the model function u_k is extended naturally outside of the support of χ_k. So these parameters serve to avoid falling into unwanted local minima. The variable m gives the order of regularity for the model functions u_k. Note that in most cases m will be set to one in this thesis.

In the following, the functional will be reformulated, such that it depends only on the set of characteristic functions, i.e. a model function u_k will also depend on the characteristic function χ_k.

Chapter 3

An Approach to Combine Multiphase Segmentation with Image Denoising

After revisiting some established models in the world of image segmentation, this chapter is devoted to analysing the proposed functional for $u \in H^1(\Omega)^l$ and $\chi \in L^\infty(\Omega)^l,\ \chi \leq 0$ and give rise to a proper algorithm,

$$\mathcal{J}(u,\chi) = \sum_{k=1}^{l} \int_\Omega \Big[|u_k - K_\eta \tilde{u}|^2 (\beta K_\eta \chi_k + \delta) + |\nabla u_k|^2 (\alpha K_\eta \chi_k + \varepsilon) \Big] \ dx. \tag{3.1}$$

Unfortunately, a minimizer for the combined problem,

$$\min_{(u,\chi)\in H^1(\Omega)^l \times L^\infty(\Omega)^l} \mathcal{J}(u,\chi), \tag{3.2}$$

cannot be guaranteed with standard mathematical arguments. Hence, the goal in this chapter is to combine multiphase segmentation with image denoising, i.e. reduce the dependence of the functional to only one variable,

$$\mathcal{J}\left(u,\chi\right) = \mathcal{J}\left(u(\chi),\chi\right) = \mathcal{J}\left(\chi\right).$$

In order to fulfil this task, a unique solution of model functions $u = (u_1, ..., u_l)^\top$ with the help of the *Lax-Milgram Lemma*, see Appendix (A.4) for an arbitrary fixed vector of characteristic functions $\chi \in L^\infty(\Omega)^l$ is computed. For the sake of brevity, the following assumption is outlined.

Assumption 3.1. *Let* $\Omega = (0,1)^2$, $0 < \alpha, \beta$ *and* $0 < \delta, \varepsilon \ll 1$ *and* $\tilde{u} \in L^\infty(\Omega)$. *In addition, let* $0 \leq \tilde{u}(x) \leq 1$ *for almost every* $x \in \Omega$.

Note that the second condition on the raw image may seem rather bold, but recall that this thesis is anchored in the space of gray scale images, which can either have the colour space $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = [0,1]$. At the end of this chapter, we will be ready to propose an algorithm to compute χ.

3.1 Introduction to the Mollifying Operator

In this subchapter, we take a look at the proposed operator K_η. This technical necessity was introduced in order to establish a proper algorithm, since in upcoming proofs certain features are needed, e.g. that the gradient of χ is bounded and together with the mollified data $K_\eta\tilde{u}$, the model functions u_k have more regularity than H^1. The parameter η will be chosen very close to 0, such that we are working with an accurate approximation to χ. This is purely theoretical and will not affect the implementation of the algorithm. Hence, denote the following characterization of a mollifier. This idea was used in another way in [3] and [11].

Definition 4. *Let $f \in L^p(\Omega)$, $1 \le p < \infty$ and let $f(x) = 0$ for $x \notin \Omega$. Define*

$$\varphi_\eta(x) := \frac{1}{4\pi\eta} e^{-\frac{|x|^2}{4\eta}} \tag{3.3}$$

with the property

$$\int_{\mathbb{R}^2} \varphi_\eta(x)\ dx = 1 \tag{3.4}$$

and let

$$f^\eta(x) := (f * \varphi_\eta)(x) = \int_{\mathbb{R}^2} \varphi_\eta(x-y) f(y)\ dy.$$

Lemma 3.2. *The function resulting from the convolution f^η is real analytic on Ω. In addition, $f^\eta \in L^p(\Omega)$ for $1 \le p < \infty$ and $f^\eta \to f$ almost everywhere for $\eta \to 0$. Moreover, if $0 \le f(x) \le 1$ for all $x \in \Omega$, then $0 \le f^\eta(x) \le 1$ [3].*

For the readers' interest on other properties, the proof can be found in [6], p. 30.

Proof. The function $\varphi_\eta(x-y)$ is infinitely differentiable in x and vanishes if $|y-x| \ge \eta$. Let α denote the multi-index and so it follows for an integrable function f

$$D^\alpha(\varphi_\eta * f)(x) = \int_{\mathbb{R}^2} D_x^\alpha \varphi_\eta(x-y) f(y)\ dy.$$

So the first conclusion is valid.

Now let $1/p + 1/q = 1$. Applying Hölder's Inequality and (3.4) the following is obtained

$$\begin{aligned}
|\varphi_\eta * f(x)| &= \left| \int_{\mathbb{R}^2} \varphi_\eta(x-y) f(y)\ dy \right| \\
&\le \left(\int_{\mathbb{R}^2} \varphi_\eta(x-y)\ dy \right)^{1/q} \left(\int_{\mathbb{R}^2} \varphi_\eta(x-y) |f(y)|^p\ dy \right)^{1/p} \\
&= \left(\int_{\mathbb{R}^2} \varphi_\eta(x-y) |f(y)|^p\ dy \right)^{1/p}.
\end{aligned}$$

Thus, by Fubini's Theorem and (3.4) we get $f * \varphi_\eta \in L^p(\Omega)$.

$$\begin{aligned}
\int_\Omega |\varphi_\eta * f|^p\ dx &\le \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \varphi_\eta(x-y) |f(y)|^p\ dy\ dx \\
&= \int_{\mathbb{R}^2} |f(y)|^p\ dy \int_{\mathbb{R}^2} \varphi_\eta(x-y)\ dx = \|f\|^p_{L^p(\Omega)}.
\end{aligned} \tag{3.5}$$

Hence, we have shown that $f^\eta \in L^p(\Omega)$. Now let $\varepsilon > 0$, since $C_0(\Omega)$ is dense in $L^p(\Omega)$ for $1 \le p < \infty$, there exists $g \in C_0(\Omega)$ such that $\|f - g\|_{L^p(\Omega)} < \varepsilon/3$ and by the previous calculations we got $\|\varphi_\eta * f - \varphi_\eta * g\|_{L^p(\Omega)} < \varepsilon/3$. So

$$|\varphi_\eta * g(x) - g(x)| = \left| \int_{\mathbb{R}^2} \varphi_\eta(x-y)(g(y) - g(x))\ dy \right| \le \sup_{|y-x|<\eta} |g(y) - g(x)|.$$

The right-hand side tends to 0 for $\eta \to 0$ since g is uniformly continuous on Ω. The last term $\|\varphi_\eta * g - g\|_{L^p(\Omega)} < \varepsilon/3$ follows from compactness of supp g and choosing η sufficiently small. Hence,

$$
\begin{aligned}
\|\varphi_\eta * f - f\|_{L^p(\Omega)} &\leq \|\varphi_\eta * f - \varphi_\eta * g\|_{L^p(\Omega)} + \|\varphi_\eta * g - g\|_{L^p(\Omega)} + \|f - g\|_{L^p(\Omega)} \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}
$$

Now, assume $0 \leq f(x) \leq 1$ for all $x \in \Omega$. Since $\varphi_\eta \geq 0$, then also $f^\eta \geq 0$ holds. Furthermore, using (3.4) concludes the proof.

$$
|f^\eta(x)| \leq \int_\Omega |\varphi_\eta(y) f(x-y)| \; dx \leq \int_\Omega |\varphi_\eta(y)| \; dx = 1.
$$

□

Lemma 3.3. *Let $f \in L^p(\Omega)$ for $1 \leq p < \infty$, then*

$$
\partial^\alpha (f * \varphi_\eta) = (\partial^\alpha f) * \varphi_\eta \tag{3.6}
$$

for $|\alpha| \leq 1$. Moreover, $\partial^\alpha (f * \varphi_\eta) \to \partial^\alpha f$ almost everywhere as $\eta \to 0$.

Proof. The proof for the first assertion follows easily by applying integration by parts,

$$
\begin{aligned}
\partial^\alpha (f * \varphi_\eta) =& \partial^\alpha \int_{\mathbb{R}^2} \varphi_\eta(x-y) f(y) \; dy \\
=& \int_{\mathbb{R}^2} \partial_x^\alpha \varphi_\eta(x-y) f(y) \; dy \\
=& (-1)^{|\alpha|} \int_{\mathbb{R}^2} \partial_y^\alpha \varphi_\eta(x-y) f(y) \; dy \\
=& \int_{\mathbb{R}^2} \varphi_\eta(x-y) \partial^\alpha f(y) \; dy \\
=& (\partial^\alpha f) * \varphi_\eta.
\end{aligned}
$$

The second claim follows from Lemma (3.2). □

Lemma 3.4. *Suppose $1 \leq p < \infty$ and $1 \leq r \leq \infty$ such that $\frac{1}{r} + 1 - \frac{1}{p} \in [0,1]$. Then the operator mapping an element $f \in L^p(\Omega)$ to $L^r(\Omega)$ is defined as follows*

$$
K_\eta f := f * \varphi_\eta.
$$

Then K_η is continuous and injective, see [3].

Proof. Let $1 \leq p < \infty$ and $1 \leq q \leq \infty$. We already know from Lemma (3.2) that φ_η is real analytic. So as a consequence, it is an element in $L^p(\Omega)$. This operator is linear, since the convolution of two functions is linear and hence, it suffices to show boundedness to guarantee continuity. Therefore, we use *Young's Inequality for Convolutions*, see Appendix (A.1), take $1 \leq r \leq \infty$ such that $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$ and thus for $f \in L^p(\Omega)$

$$
\|K_\eta f\|_{L^r(\Omega)} \leq \|f\|_{L^p(\Omega)} \|\varphi_\eta\|_{L^q(\Omega)}.
$$

Finally, from $\int_{\mathbb{R}^2} \varphi_\eta(x) \; dx = 1$ it follows that the operator is injective. □

Remark 3.5. *For the sake of brevity in writing, we denote the mollified data with $\tilde{u}_\eta := K_\eta \tilde{u}$.*

3.2 The Optimality System for Model Functions

In this section we want to deduce the necessary optimality conditions for minimizing the functional $\mathcal{J}$, recall (3.1), with respect to the model functions $u \in H^1(\Omega)^l$. Thus, consider the minimization problem for a fixed characteristic function $\chi \in L^\infty(\Omega)^l$

$$\min_{u\in H^1(\Omega)^l} \mathcal{J}_\chi(u) := \min_{u\in H^1(\Omega)^l} \mathcal{J}(u,\chi). \tag{3.7}$$

Note that we are deriving the necessary optimality condition for each model function u_k separately. As a consequence u_k depends only on the k-th characteristic function, i.e. $u_k(\chi_k)$. In other words u_k is modelled solely with the help of the k-th phase's support.

Lemma 3.6. *Given Assumption (3.1), let $\chi_k \in L^\infty(\Omega)$ be a characteristic function. Then the necessary optimality system for $u_k \in H^2(\Omega)$, $k = 1, ..., l$ reads as follows:*

$$\begin{aligned} -\nabla\cdot\left[(\alpha K_\eta\chi_k+\varepsilon)\nabla u_k\right] + (\beta K_\eta\chi_k+\delta)\, u_k &= (\beta K_\eta\chi_k+\delta)\,\tilde{u}_\eta, \qquad \text{in } \Omega,\\ \frac{\partial u_k}{\partial n} &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{3.8}$$

Proof. Let $v \in C_0^\infty(\bar{\Omega})$ be an arbitrary perturbation. Due to the following calculations, i.e. the first variation $\frac{\delta\mathcal{J}_\chi(u_k;v)}{\delta u_k} := \frac{d}{dt}\mathcal{J}_\chi(u_k+tv)\big|_{t=0}$ exists for every $v \in C_0^\infty\left(\bar{\Omega}\right)$, we observe that $\mathcal{J}_\chi$ is everywhere Gâteaux differentiable.

$$\begin{aligned} \frac{\delta\mathcal{J}_\chi}{\delta u_k}(u_k;v) =& \frac{d}{dt}\frac{1}{2}\sum_{k=1}^{l}\int_\Omega (\beta K_\eta\chi_k+\delta)\,|u_k+tv-\tilde{u}_\eta|^2 + (\alpha K_\eta\chi_k+\varepsilon)\,|\nabla(u_k+tv)|^2\Big|_{t=0}\, dx\\ =& \int_\Omega (\beta K_\eta\chi_k+\delta)(u_k+tv-\tilde{u}_\eta)\, v + (\alpha K_\eta\chi_k+\varepsilon)(\nabla(u_k+tv))\cdot\nabla v\Big|_{t=0}\, dx\\ =& \int_\Omega (\beta K_\eta\chi_k+\delta)(u_k-\tilde{u}_\eta)\, v + (\alpha K_\eta\chi_k+\varepsilon)\nabla u_k\cdot\nabla v\, dx, \quad \forall k=1,...,l. \end{aligned} \tag{3.9}$$

By using *Green's Formula*, see Appendix (A.2), and the assumption $u_k \in H^2(\Omega)$ it follows that

$$0 = \int_\Omega (\beta K_\eta\chi_k+\delta)(u_k-\tilde{u}_\eta)\, v - \nabla\cdot\left[(\alpha K_\eta\chi_k+\varepsilon)\nabla u_k\right] v\, dx + \int_{\partial\Omega}\frac{\partial u_k}{\partial n} v\, ds_x.$$

Letting v be concentrated on $\partial\Omega$ gives the boundary condition $\frac{\partial u_k}{\partial n} = 0$, and thus the boundary integral cancels out. Moreover, v was chosen arbitrarily, so after applying the *Fundamental Lemma of Variational Calculus*, see Appendix (A.3), the necessary optimality condition for $u_k \in H^1(\Omega)$ is obtained and thus completes the proof. □

Hence, we are now able to prove unique existence of a minimizer u of $\mathcal{J}$ for an arbitrary fixed set of characteristic functions and so we are one step closer to gaining our final cost functional.

Theorem 3.7. *Given Assumption (3.1), the necessary optimality system (3.8) has a unique weak solution $u_k \in H^1(\Omega)$ for an arbitrary fixed characteristic function $\chi_k \in L^\infty(\Omega)$, $0 \le \chi_k \le 1$ almost everywhere for all $k = 1, ..., l$. Moreover,*

$$\|u_k\|_{H^1(\Omega)} \le (\beta+\delta)\frac{1}{\min(\varepsilon,\delta)}\|\tilde{u}_\eta\|_{L^\infty(\Omega)}. \tag{3.10}$$

Proof. Multiply (3.8) with $v \in H^1(\Omega)$ and apply *Green's Formula* such that the following holds:

$$a_k(u_k, v) = l_k(v), \quad \forall v \in H^1(\Omega),\ k = 1, ..., l, \tag{3.11}$$

where $a_k : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ is a bilinear form, defined by

$$a_k(u_k, v) := \int_\Omega (\beta K_\eta \chi_k + \delta) u_k v + (\alpha K_\eta \chi_k + \varepsilon) \nabla u_k \cdot \nabla v \ dx, \quad \text{for } k = 1, ..., l \tag{3.12}$$

and l_k is a linear functional mapping from the Sobolev space $H^1(\Omega)$ onto $\mathbb{R}$, defined by

$$l_k(v) := \int_\Omega (\beta K_\eta \chi_k + \delta) \tilde{u}_\eta v \ dx, \quad \text{for } k = 1, ..., l \tag{3.13}$$

Note further, since $0 \leq (K_\eta \chi_k)(x) \leq 1$ for $x \in \Omega$ almost everywhere,

$$\begin{aligned} &\delta \leq \|\beta K_\eta \chi + \delta\|_{L^\infty(\Omega)} \leq \beta \|K_\eta \chi_k\|_{L^\infty(\Omega)} + \delta \leq \beta + \delta, \\ &\varepsilon \leq \|\alpha K_\eta \chi + \varepsilon\|_{L^\infty(\Omega)} \leq \alpha + \varepsilon. \end{aligned} \tag{3.14}$$

To guarantee uniqueness *Lax-Milgram* is applied to the necessary optimality condition (3.11). Firstly, boundedness of the linear functional l_k is shown:

$$\begin{aligned} |l_k(v)| &\leq \int_\Omega |\beta K_\eta \chi_k + \delta||\tilde{u}_\eta v| \ dx \leq \|\beta K_\eta \chi_k + \delta\|_{L^\infty(\Omega)} \|\tilde{u}_\eta\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} \\ &\leq (\beta + \delta) \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \|v\|_{H^1(\Omega)}. \end{aligned}$$

Secondly, boundedness of the bilinear form a_k is proven:

$$\begin{aligned} |a_k(u_k, v)| &\leq \|\beta K_\eta \chi_k + \delta\|_{L^\infty(\Omega)} \int_\Omega |u_k v| \ dx + \|\alpha K_\eta \chi_k + \varepsilon\|_{L^\infty(\Omega)} \int_\Omega |\nabla u_k||\nabla v| \ dx \\ &\leq (\beta + \delta) \|u_k\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + (\alpha + \varepsilon) \|\nabla u_k\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \\ &\leq \max\{(\beta + \delta), (\alpha + \varepsilon)\} \left(\|u_k\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|\nabla u_k\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \right) \\ &\leq 2 \max\{(\beta + \delta), (\alpha + \varepsilon)\} \|u_k\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}. \end{aligned}$$

Last but not least, ellipticity of the bilinear form a_k is obtained:

$$\begin{aligned} a_k(u_k, u_k) &= \int_\Omega (\beta K_\eta \chi_k + \delta) |u_k|^2 + (\alpha K_\eta \chi_k + \varepsilon) |\nabla u_k|^2 \ dx \\ &\geq \delta \int_\Omega |u_k|^2 \ dx + \varepsilon \int_\Omega |\nabla u_k|^2 \ dx \\ &\geq \min(\delta, \varepsilon) \left(\|u_k\|^2_{L^2(\Omega)} + \|\nabla u_k\|^2_{L^2(\Omega)} \right) \\ &\geq \min(\delta, \varepsilon) \|u_k\|^2_{H^1(\Omega)}. \end{aligned}$$

Thus, the *Lax-Milgram* guarantees the existence of a unique minimizer $u_k^* \in H^1(\Omega)$, $\forall k = 1, ..., l$. Furthermore, *Lax-Milgram* gives a bound for the solution, i.e.,

$$\|u_k\|_{H^1(\Omega)} \leq \frac{1}{\min(\varepsilon, \delta)} \|l_k\| \leq (\beta + \delta) \frac{1}{\min(\varepsilon, \delta)} \|\tilde{u}_\eta\|_{L^\infty(\Omega)}, \qquad \forall \chi_k \in \mathrm{BV}(\Omega, [0, 1]).$$

□

Remark 3.8. *Since the mollification operator K_η was already used on χ and the data $\tilde{u}$, note that existence of $u \in H^1(\Omega)^l$ is also ensured for $\chi \in L^\infty(\Omega)$, $\chi \geq 0$ and $\tilde{u} \in L^\infty(\Omega)$. But in later calculations problems arise that demand higher regularity on the model functions u_k, $k = 1, ..., l$, namely at least $u_k \in H^4(\Omega)$, which cannot be guaranteed.*

Proposition 3.9. *Given Assumption (3.1), then the unique solution $u \in H^1(\Omega)^l$ from Theorem (3.7) is an element in $C^\infty(\bar{\Omega})^l$.*

Proof. Theorem 3, Section 6.3 in [7] states, that if the coefficients and the right-hand side are in $C^\infty(\bar{\Omega})$, then so is the solution. In this case $\beta K_\eta \chi_k + \delta$, $\alpha K_\eta \chi_k + \varepsilon$ and $(\beta K_\eta \chi_k + \delta)\tilde{u}_\eta$ are elements in $C^\infty(\bar{\Omega})$ for all $k = 1, ..., l$ due to the mollifying operator K_η and hence $u \in C^\infty(\bar{\Omega})^l$. □

3.3 Expressing the Model Functions in terms of a Characteristic Function

As a consequence of the last section we are able to reveal that the model functions u_k depend implicitly on its corresponding characteristic function χ_k, $k = 1, ..., l$. Recall that there exists a corollary of *Lax-Milgram* which guarantees the existence of a unique solution operator $L(\chi_k) \in \mathcal{L}(H^1(\Omega))$, for which the following holds:

$$\begin{aligned}(L(\chi_k)\, u_k, v)_{H^1(\Omega)} =& a_k(u_k, v) = \int_\Omega (\beta K_\eta \chi_k + \delta)\, u_k v + (\alpha K_\eta \chi_k + \varepsilon)\, \nabla u_k \cdot \nabla v \; dx, \\ & \text{for } k = 1, ..., l.\end{aligned} \tag{3.15}$$

Furthermore, the operator is bounded for all χ_k due to its characterization of *Lax-Milgram*,

$$\|L(\chi_k)\| \leq 2\max\{(\alpha + \varepsilon), (\beta + \delta)\} \quad \text{and} \left\|L(\chi_k)^{-1}\right\| \leq \frac{1}{\min(\varepsilon, \delta)}. \tag{3.16}$$

Now let $f(\chi_k) \in H^1(\Omega)$ be the representing element of

$$(f(\chi_k), v)_{H^1(\Omega)} = l_k(v) = \int_\Omega (\beta K_\eta \chi_k + \delta)\, \tilde{u}_\eta v \; dx, \quad \text{for } k = 1, ..., l, \tag{3.17}$$

with

$$\|f(\chi_k)\|_{H^m(\Omega)} = \|l_k\|, \tag{3.18}$$

where $\|\cdot\|$ denotes the corresponding operator norm. Note that its unique existence is ensured due to the *Riesz' Representation Theorem*, see Appendix (A.6). So, from

$$(L(\chi_k)\, u_k, v)_{H^1(\Omega)} = (f(\chi_k), v)_{H^1(\Omega)}, \tag{3.19}$$

the unique solution $u_k \in H^1(\Omega)$ of (3.8) has the representation

$$u_k(\chi_k) := L(\chi_k)^{-1} f(\chi_k), \tag{3.20}$$

So u_k is well-defined and the k-th model function depends on the k-th characteristic function, creating the k-th phase of the image.

3.4 The Resulting Cost Functional

Before the final functional is introduced, some further definitions will be investigated, which will be needed in the later chapter. For brevity in writing the following set is introduced.

Definition 5. *Let the space of functions that have range in* $[0,1]$ *and bounded variation denoted by*

$$\mathrm{BV}(\Omega,[0,1]) := \{\chi \in \mathrm{BV}(\Omega) : \chi(x) \in [0,1] \text{ for almost every } x \in \Omega\}. \tag{3.21}$$

This introduction might seem out of the blue, but to give a sneak peak for upcoming proofs, we will investigate minimizing the resulting cost functional with respect to functions with bounded variation and their range in $[0,1]$. Concerning the details, they follow in the upcoming chapter. Note that Theorem (3.7) still holds for $\chi_k \in \mathrm{BV}(\Omega,[0,1])$ because $0 \leq \chi_k(x) \leq 1$ for almost every $x \in \Omega$ and thus, $\chi_k \in L^\infty(\Omega)$.

Hence, we introduce the following mappings.

Definition 6. *The map that assigns an element* $\chi \in \mathrm{BV}(\Omega,[0,1])$ *to a linear operator is defined as follows*

$$L := \begin{cases} \mathrm{BV}(\Omega,[0,1]) \to \mathcal{L}\left(H^1(\Omega)\right), \\ \chi \mapsto L(\chi), \end{cases}$$

and the map that assigns it to an element in $H^1(\Omega)$.

$$f := \begin{cases} \mathrm{BV}(\Omega,[0,1]) \to H^1(\Omega). \\ \chi \mapsto f(\chi). \end{cases}$$

Later a continuous dependence on $\chi \in \mathrm{BV}(\Omega,[0,1])$ will be shown for u_k for $k = 1,...,l$. This particular result will be needed later on. Nevertheless, for this work minimizing the upcoming functional is involved:

$$\mathcal{J}(\chi) = \frac{1}{2}\sum_{k=1}^{l}\int_\Omega |u_k(\chi_k) - \tilde{u}_\eta|^2 (K_\eta\chi_k\beta + \delta) + |\nabla u_k(\chi_k)|^2 (K_\eta\chi_k\alpha + \varepsilon)\ dx. \tag{3.22}$$

In this case we will not further investigate minimizing over $L^\infty(\Omega)^l$, but rather over the constraint set

$$\mathrm{BV}(\Omega,\mathcal{E}) := \{\chi \in \mathrm{BV}(\Omega)^l :\ \chi(x) \in \mathcal{E} \ for\ x \in \Omega\ a.e.\} \subset L^1(\Omega)^l, \tag{3.23}$$

where $\mathcal{E} = \{e_1,...,e_l\}$ is the set of unit vectors, which was introduced in J. Lellmann's Docotoral Thesis [8]. This set describes the nature of characteristic functions in image segmentation. So for an arbitrary $x \in \Omega$ take a vector-valued $\chi \in \mathrm{BV}(\Omega,\mathcal{E})$ and thus, $\chi(x) = (\chi_1(x),...,\chi_l(x))^\top = e_k$, for $k \in \{1,...,l\}$. In other words, $x \in \Omega$ belongs to a single phase of the partition resulting from the segmentation process. But this constraint set is actually too strict to find a minimum or even a unique one. Thus, J. Lellmann has introduced the relaxed constraint set,

$$\mathrm{BV}(\Omega,\Delta_l) := \{\chi \in \mathrm{BV}(\Omega)^l :\ \chi(x) \in \Delta_l \ for\ x \in \Omega\ a.e.\}, \tag{3.24}$$

where Δ_l denotes the unit simplex in $\mathbb{R}^l$. This means, that $\chi_k(x)$ is allowed to attain values between $[0,1]$, and that $\sum_{k=1}^{l}\chi_k(x) = 1$ for almost every $x \in \Omega$. It is comparable to the soft clustering methods, where every data point can belong to several clusters by a certain percentage.

In addition, $\mathrm{BV}(\Omega, \Delta_l)$ is convex, i.e. for $\chi, \tilde{\chi} \in \mathrm{BV}(\Omega, \Delta_l)$ and $\lambda \in [0,1]$

$$\sum_{k=1}^{l} \lambda \chi_k(x) + (1-\lambda)\tilde{\chi}_k(x) = \lambda \sum_{k=1}^{l} \chi_k(x) + (1-\lambda) \sum_{k=1}^{l} \tilde{\chi}_k(x) = \lambda + 1 - \lambda = 1,$$

and closed. If a sequence $\left(\chi^{(n)}\right)$ converges to some χ in $\mathrm{BV}(\Omega, \Delta_l)$, then it converges in $L^2(\Omega)$ since $\Omega \subset \mathbb{R}^n$ and is bounded, see Appendix (A.8). Then there exists subsequence $\left(\chi^{(n_m)}\right) \subset \left(\chi^{(n)}\right)$ that converges point-wise almost everywhere to χ. Then it follows that

$$\sum_{k=1}^{l} \chi_k(x) = \lim_{m\to\infty} \sum_{k=1}^{l} \chi_k^{(n_m)}(x) = 1.$$

Note that in contrast $\mathrm{BV}(\Omega, [0,1])$ was only introduced for a single-valued characteristic function and will solely be used for some technical results.

However to obtain a suitable result in a space of bounded variation theoretically and computationally, a certain penalty term is needed which was already introduced in Chapter 2, namely the *Total Variation*. Thus, the functional reads as follows

$$\begin{aligned} \mathcal{J}(\chi) + \gamma \mathrm{TV}(\chi) =& \frac{1}{2} \sum_{k=1}^{l} \int_{\Omega} |u_k(\chi_k) - \tilde{u}_\eta|^2 (K_\eta \chi_k \beta + \delta) \\ &+ |\nabla u_k(\chi_k)|^2 (K_\eta \chi_k \alpha + \varepsilon)\ dx + \gamma \mathrm{TV}(\chi_k), \end{aligned} \tag{3.25}$$

where $\gamma > 0$.

So we are able to propose the following algorithm. It is the first version of our actual working scheme. The parameters are chosen according to *Assumption* (3.1). The k-th summand of $\mathcal{J}$ is denoted by $\mathcal{J}_k$ because χ_k only depends on this part of the sum. Note further that for now it is not clear, what the initial χ_0 will be.

Algorithm 1 Image Multiclass Labelling and Denoising

1: **Input:** $\chi_0, \tilde{u}, \alpha, \beta, \varepsilon, \delta, \gamma, l$
2: **Output:** χ, u
3: **while** χ changes **do**
4:     **for** $k = 1, ..., l$ **do**
5:         **if** $\chi_k(x) \leq 0$ or $\chi_k(x) \geq 1$ for some $x \in \Omega$ **then**
6:             Cut χ_k such that it has range in $[0,1]$.
7:         Compute u_k satisfying

$$-\nabla \cdot [(K_\eta \chi_k \alpha + \varepsilon) \nabla u_k] + (K_\eta \beta \chi_k + \delta) u_k = (K_\eta \beta \chi_k + \delta) \tilde{u}_\eta,$$

8:         Compute $\chi_k = \arg\ \min \mathcal{J}_k(\chi_k) + \gamma \mathrm{TV}(\chi_k)$.

Chapter 4

Analysis of the Proposed Method

This chapter focuses fully on analysing the cost functional and refining the algorithmic strategy of Chapter 3. Thus, existence of a minimum for $\chi \in \mathrm{BV}(\Omega, \Delta_l)$ will be proven and then a more in-depth algorithm will be presented. Convergence of the algorithm will be shown with *Schauder's Fixed Point Theorem Version II*, see Appendix (A.9). Note that the technical necessity of K_η was cautiously introduced in order to obtain the desired fixed point. At the end of this chapter, the strategy to map elements from $\mathrm{BV}(\Omega, \Delta_l)$ to $\mathrm{BV}(\Omega, \mathcal{E})$ will be elaborated.

4.1 Existence of Minimum

4.1.1 Preliminary Results

Firstly, some precursory results are needed, e.g. continuous dependence of the model function u on χ or boundedness of $\mathcal{J}(\chi)$, because in order to be able to prove some important features of the functional like lower semi-continuity later on one has to know how the solution function $u(\chi) \in H^1(\Omega)^l$ and $\mathcal{J}$ behaves in dependence of χ.

Lemma 4.1. *Let Assumption (3.1) hold, then u_k for all $k = 1, ..., l$ is non-expansive and thus continuous in* $\mathrm{BV}(\Omega, [0,1])$, *i.e.,*

$$\lim_{n\to\infty} \left\| u_k\left(\chi^{(n)}\right) - u_k(\chi) \right\|_{H^1(\Omega)} = 0, \tag{4.1}$$

for a sequence $\left(\chi^{(n)}\right)$ that converges to some χ in $\mathrm{BV}(\Omega, [0,1])$.

Proof. Let $\left(\chi^{(n)}\right) \subset \mathrm{BV}(\Omega, [0,1])$ be sequence converging to some $\chi \in \mathrm{BV}(\Omega, [0,1])$ (and thus L^1 convergence). Subtracting $\left(\frac{\partial}{\partial u_k}\mathcal{J}(u_k; v)\right)(\chi)$ from $\left(\frac{\partial}{\partial u_k}\mathcal{J}(u_k; v)\right)(\chi^{(n)})$ and linearity of K_η

gives

$$
\begin{aligned}
0 &= \left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k; v\right)\right)\left(\chi^{(n)}\right) - \left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k; v\right)\right)(\chi) \\
\Leftrightarrow & \int_\Omega \left(\beta K_\eta \chi^{(n)} + \delta\right) u_k\left(\chi^{(n)}\right) v + \left(\alpha K_\eta \chi^{(n)} + \varepsilon\right) \nabla u_k\left(\chi^{(n)}\right) \cdot \nabla v \; dx \\
& - \int_\Omega \left(\beta K_\eta \chi + \delta\right) u_k(\chi)\, v + \left(\alpha K_\eta \chi + \varepsilon\right) \nabla u_k(\chi) \cdot \nabla v \; dx \\
& = \int_\Omega \left(\beta K_\eta \chi^{(n)} + \delta\right) \tilde{u}_\eta v \; dx - \int_\Omega \left(\beta K_\eta \chi + \delta\right) \tilde{u}_\eta v \; dx \\
\Leftrightarrow & \int_\Omega \left(\beta K_\eta \chi + \delta\right)\left(u_k\left(\chi^{(n)}\right) - u_k(\chi)\right) v + \left(\alpha K_\eta \chi + \varepsilon\right)\left(\nabla u_k\left(\chi^{(n)}\right) - \nabla u_k(\chi)\right) \cdot \nabla v \; dx \\
& = \int_\Omega \beta K_\eta\left(\chi^{(n)} - \chi\right)\left(\tilde{u}_\eta - u_k\left(\chi^{(n)}\right)\right) v - \alpha K_\eta\left(\chi^{(n)} - \chi\right) \nabla u_k\left(\chi^{(n)}\right) \cdot \nabla v \; dx
\end{aligned}
\tag{4.2}
$$

Comparing the last equation of (4.2) to a_k, see (3.12), we see that they coincide and so only boundedness of the right-hand side is left to show. Using *Young's Inequality for Convolutions* on $\left\|K_\eta\left(\chi^{(n)} - \chi\right)\right\|_{L^r(\Omega)} \le \|\varphi_\eta\|_{L^q(\Omega)} \left\|\chi^{(n)} - \chi\right\|_{L^p(\Omega)}$ for $r, q = \infty$ and $p = 1$ such that $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$ and (3.10) gives

$$
\begin{aligned}
& \left| \int_\Omega \beta K_\eta\left(\chi^{(n)} - \chi\right)\left(\tilde{u}_\eta - u_k\left(\chi^{(n)}\right)\right) v - \alpha K_\eta\left(\chi^{(n)} - \chi\right) \nabla u_k\left(\chi^{(n)}\right) \cdot \nabla v \; dx \right| \\
\le & \left\|K_\eta\left(\chi^{(n)} - \chi\right)\right\|_{L^\infty(\Omega)} \left(\beta \left\|\tilde{u}_\eta - u_k\left(\chi^{(n)}\right)\right\|_{L^2(\Omega)} + \alpha \left\|\nabla u_k\left(\chi^{(n)}\right)\right\|_{L^2(\Omega)}\right) \|v\|_{H^1(\Omega)} \\
\le & \|\varphi_\eta\|_{L^\infty(\Omega)} \left\|\left(\chi^{(n)} - \chi\right)\right\|_{L^1(\Omega)} \left(\beta \|\tilde{u}_\eta\|_{L^\infty(\Omega)} + (\beta + \alpha) \left\|u_k\left(\chi^{(n)}\right)\right\|_{H^1(\Omega)}\right) \|v\|_{H^1(\Omega)} \\
\le & \|\varphi_\eta\|_{L^\infty(\Omega)} \left\|\left(\chi^{(n)} - \chi\right)\right\|_{L^1(\Omega)} \left(\beta \|\tilde{u}_\eta\|_{L^\infty(\Omega)} + (\beta + \alpha) \frac{\beta + \delta}{\min(\varepsilon, \delta)} \|\tilde{u}_\eta\|_{L^\infty(\Omega)}\right) \|v\|_{H^1(\Omega)} .
\end{aligned}
\tag{4.3}
$$

The Corollary of *Lax-Milgram* implies the estimate

$$
\begin{aligned}
\left\|u_k(\chi^{(n)}) - u_k(\chi)\right\|_{H^1(\Omega)} \le & \frac{1}{\min(\varepsilon, \delta)} \|\varphi_\eta\|_{L^\infty(\Omega)} \left\|\chi^{(n)} - \chi\right\|_{L^1(\Omega)} \\
& \cdot \left\{\beta + (\beta + \alpha) \frac{\beta + \delta}{\min(\varepsilon, \delta)}\right\} \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \\
& \le C_2 \left\|\chi^{(n)} - \chi\right\|_{L^1(\Omega)}
\end{aligned}
\tag{4.4}
$$

Recall that $u_k\left(\chi^{(n)}\right) \in H^1(\Omega)$ for all n, since Theorem (3.7) also holds for functions in $\mathrm{BV}(\Omega, [0,1])$. Taking the limit n to ∞ completes the proof. □

Next, recall the functional $\mathcal{J}$ (3.1), so before the proof of existence of minimum is presented, some features of the cost functional $\mathcal{J}$ have to be ensured, i.e. boundedness from below of the functional that guarantees existence of a minimizing sequence and continuity, such that together with the penalty term lower semi-continuity of the cost is ensured. These are all necessary for using the direct method, see [4] p. 250, for proving existence of a minimum.

Lemma 4.2. *Given Assumption (3.1). Then $\mathcal{J}$ is bounded from below and continuous for $\chi \in \mathrm{BV}(\Omega, \Delta_l)$.*

Proof. For boundedness from below, by definition $\mathcal{J} \geq 0$. For continuity, take a sequence $\left(\chi^{(n)}\right) \in \mathrm{BV}(\Omega, \Delta_l)$ that converges to χ in BV and hence in L^1. Linearity of K_η gives

$$\begin{aligned} \left|\mathcal{J}\left(\chi^{(n)}\right) - \mathcal{J}(\chi)\right| = & \Bigg| \frac{1}{2} \sum_{k=1}^{l} \int_\Omega \left(\beta K_\eta \chi_k^{(n)} + \delta\right) \left|u_k\left(\chi_k^{(n)}\right) - \tilde{u}_\eta\right|^2 \\ & + \left(\alpha K_\eta \chi_k^{(n)} + \varepsilon\right) \left|\nabla u_k\left(\chi_k^{(n)}\right)\right|^2 - (\beta K_\eta \chi_k + \delta)|u_k(\chi_k) - \tilde{u}_\eta|^2 \\ & - (\alpha K_\eta \chi_k + \varepsilon)|\nabla u_k(\chi_k)|^2 \, dx \Bigg| \end{aligned} \tag{4.5}$$

By adding the terms $\pm\left(\beta K_\eta \chi_k^{(n)} + \delta\right)|u_k(\chi_k) - \tilde{u}_\eta|^2$ and $\pm\left(\alpha K_\eta \chi_k^{(n)} + \varepsilon\right)|\nabla u_k(\chi_k)|^2$ gives

$$\begin{aligned} \mathcal{J}_1\left(\chi^{(n)}\right) - \mathcal{J}_1(\chi) := & \frac{1}{2} \sum_{k=1}^{l} \int_\Omega \beta K_\eta \left(\chi_k^{(n)} - \chi_k\right) |u_k(\chi_k) - \tilde{u}_\eta|^2 \\ & + \alpha K_\eta \left(\chi_k^{(n)} - \chi_k\right) |\nabla u_k(\chi_k)|^2 \, dx, \\ \mathcal{J}_2\left(\chi^{(n)}\right) - \mathcal{J}_2(\chi) := & \frac{1}{2} \sum_{k=1}^{l} \int_\Omega \left(\beta K_\eta \chi_k^{(n)} + \delta\right) \left(\left|u_k\left(\chi_k^{(n)}\right) - \tilde{u}_\eta\right|^2 - |u_k(\chi_k) - \tilde{u}_\eta|^2\right) dx, \\ \mathcal{J}_3\left(\chi^{(n)}\right) - \mathcal{J}_3(\chi) := & \frac{1}{2} \sum_{k=1}^{l} \int_\Omega \left(\alpha K_\eta \chi_k^{(n)} + \varepsilon\right) \left(\left|\nabla u_k(\chi_k^{(n)})\right|^2 - |\nabla u_k(\chi_k)|^2\right) dx. \end{aligned} \tag{4.6}$$

With the help of the triangle inequality, (3.10) and *Young's Inequality for Convolutions* for $r, q = \infty$ and $p = 1$

$$
\begin{aligned}
\left|\mathcal{J}_1\left(\chi^{(n)}\right)-\mathcal{J}_1(\chi)\right| &\leq \frac{1}{2}\sum_{k=1}^{l}\left\|K_\eta\left(\chi_k^{(n)}-\chi_k\right)\right\|_{L^\infty(\Omega)} \\
&\cdot\left(\beta\left\|u_k(\chi_k)-\tilde{u}_\eta\right\|_{L^2(\Omega)}^2+\alpha\left\|\nabla u_k(\chi_k)\right\|_{L^2(\Omega)}^2\right) \\
&\leq \frac{1}{2}\sum_{k=1}^{l}\left\|\varphi_\eta\right\|_{L^\infty(\Omega)}\left\|\chi_k^{(n)}-\chi_k\right\|_{L^1(\Omega)} \\
&\cdot\left((\beta+\alpha)\left\|u_k(\chi_k)\right\|_{H^1(\Omega)}^2+\beta\left\|\tilde{u}_\eta\right\|_{L^2(\Omega)}^2\right) \\
&\leq \frac{1}{2}\sum_{k=1}^{l}\left\|\varphi_\eta\right\|_{L^\infty(\Omega)}\left\|\chi_k^{(n)}-\chi_k\right\|_{L^1(\Omega)} \\
&\cdot\left\{(\beta+\alpha)\left(\frac{\beta+\varepsilon}{\min(\varepsilon,\delta)}\right)^2\left\|\tilde{u}_\eta\right\|_{L^\infty(\Omega)}^2+\beta\left\|\tilde{u}_\eta\right\|_{L^2(\Omega)}^2\right\}.
\end{aligned}
$$

The differences $\mathcal{J}_i\left(\chi^{(n)}\right)-\mathcal{J}_i(\chi)$ for $i=2,3$ have the form $a^2-b^2=(a+b)(a-b)$ and together with the *Cauchy-Schwarz Inequality*, see Appendix (A.17), (3.10) and $\left\|K_\eta\chi_k^{(n)}\right\|_{L^\infty(\Omega)}\leq 1$ for $k=1,...,l$ and for all $n\in\mathbb{N}$ and Lemma (4.1) implies

$$
\begin{aligned}
\left|\mathcal{J}_2\left(\chi^{(n)}\right)-\mathcal{J}_2(\chi)\right| &\leq \frac{1}{2}\sum_{k=1}^{l}(\beta+\delta)\left\|u_k(\chi_k)-u_k\left(\chi_k^{(n)}\right)\right\|_{H^1(\Omega)}\left\{\left(\left\|u_k(\chi_k)\right\|_{H^1(\Omega)}\right.\right. \\
&\left.+\left\|u_k\left(\chi_k^{(n)}\right)\right\|_{H^1(\Omega)}+2\left\|\tilde{u}_\eta\right\|_{L^\infty(\Omega)}\right\} \\
&\leq \frac{1}{2}\sum_{k=1}^{l}(\beta+\delta)C_2\left\|\chi_k-\chi_k^{(n)}\right\|_{L^1(\Omega)} \\
&\cdot 2\left(\frac{\beta+\delta}{\min(\varepsilon,\delta)}\left\|\tilde{u}_\eta\right\|_{L^\infty(\Omega)}+\left\|\tilde{u}_\eta\right\|_{L^\infty(\Omega)}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\left|\mathcal{J}_3\left(\chi^{(n)}\right)-\mathcal{J}_3(\chi)\right| &\leq \frac{1}{2}\sum_{k=1}^{l}(\alpha+\varepsilon)\left\|u_k(\chi_k)-u_k\left(\chi_k^{(n)}\right)\right\|_{H^1(\Omega)} \\
&\cdot\left(\left\|u_k(\chi_k)\right\|_{H^1(\Omega)}+\left\|u_k\left(\chi_k^{(n)}\right)\right\|_{H^1(\Omega)}\right) \\
&\leq \frac{1}{2}\sum_{k=1}^{l}C_2\left\|\chi_k-\chi_k^{(n)}\right\|_{L^1(\Omega)}2\frac{\beta+\delta}{\min(\varepsilon,\delta)}\left\|\tilde{u}_\eta\right\|_{L^\infty(\Omega)}.
\end{aligned}
$$

Finally, adding these three estimates leads to

$$|\mathcal{J}\left(\chi^{(n)}\right) - \mathcal{J}(\chi)| \leq \frac{1}{2}\sum_{k=1}^{l} C \left\|\chi_k^{(n)} - \chi_k\right\|_{L^1(\Omega)} \tag{4.7}$$

and thus $n \to \infty$ completes the proof. □

At the end of *Chapter 3* the *Total Variation* was introduced as the penalty term such that a minimum in a space of bounded variation can be obtained. But to avoid non-differentiability of the penalty term at $\nabla\chi = 0$, a more general approach is used, i.e.,

$$J_\tau(\chi) = \int_\Omega \sqrt{|\nabla\chi|^2 + \tau}\, dx. \tag{4.8}$$

where $0 < \tau \ll 1$. This term is well-defined for $\chi \in W^{1,1}(\Omega)$, see [12], where $W^{1,1}(\Omega)$ denotes the corresponding Sobolev space. In this paper it was also proven, that the effective domain of J_τ is indeed $\mathrm{BV}(\Omega)$.

Theorem 4.3. *Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$. For any $0 < \tau \ll 1$ and $\chi \in L^1(\Omega)$, $\mathrm{TV}(\chi) < \infty$ holds if and only if $J_\tau(\chi) < \infty$ [12].*

Proof. Let $\chi \in L^1(\Omega)$ and $v \in \mathcal{V} := \{v \in C_0^1(\Omega;\mathbb{R}^n) : \ |v(x)| \leq 1 \quad \forall x \in \Omega\}$, then it holds

$$\int_\Omega (-\chi \text{ div } v)\, dx \leq \int_\Omega \left(-\chi \text{ div } v + \sqrt{\tau\,(1-|v|^2)}\right)\, dx \leq \int_\Omega \left(-\chi \text{ div } v + \sqrt{\tau}\right)\, dx.$$

By taking the supremum over $v \in \mathcal{V}$, it follows that

$$\mathrm{TV}(\chi) \leq J_\tau(\chi) \leq \mathrm{TV}(\chi) + \sqrt{\tau}|\Omega|. \tag{4.9}$$

The proof completes due to boundedness of Ω. □

Theorem 4.4. *Let $0 < \tau \ll 1$, then J_τ is lower semi-continuous with respect to the L^p-topology [12].*

Proof. Let (u_n) be a sequence that converges weakly to some $\bar{u}$ in $L^p(\Omega)$. Taking $v \in \mathcal{V}$, where $\mathcal{V}$ is the same space as in the proof of Theorem (4.3), so $\text{div } v \in C(\Omega)$ and thus,

$$\begin{aligned}\int_\Omega \left((-\bar{u} \text{ div } v) + \sqrt{\tau\,(1-|v|^2)}\right)\, dx &= \lim_{n\to\infty} \int_\Omega \left((-u_n \text{ div } v) + \sqrt{\tau\,(1-|v|^2)}\right)\, dx \\ &= \liminf_{n\to\infty} \int_\Omega \left((-\bar{u} \text{ div } v) + \sqrt{\tau\,(1-|v|^2)}\right)\, dx \\ &\leq \liminf_{n\to\infty} J_\tau(u_n).\end{aligned} \tag{4.10}$$

Taking the supremum over $v \in \mathcal{V}$ gives $J_\tau(\bar{u}) \leq \liminf_{n\to\infty}(u_n)$. □

4.1.2 Proof of Existence of Minimizer χ

After gathering all the necessary preliminaries, we are now able to prove existence of a minimizer. Note that uniqueness cannot be guaranteed because of the apparent lack of convexity with respect to χ. Therefore, considering the algorithmic strategy we will start sufficiently close to a minimum and then apply the proposed algorithm.

Theorem 4.5. *Given Assumption* (3.1)*,*$\gamma > 0$ *and let* $\tilde{u} \neq 0$*. Then the functional* $\mathcal{J} + \gamma J_\tau$*,* $0 < \tau \ll 1$ *has at least one minimizer in* $\mathrm{BV}(\Omega, \Delta_l)$*.*

Proof. First of all, Lemma (4.2) implies that $\mathcal{J}$ is bounded from below and so is J_τ because of (4.9)

$$0 \leq \mathrm{TV}(\chi) \leq J_\tau(\chi),$$

and so $\mathcal{J} + \gamma J_\tau$ is bounded from below. Thus, a minimizing sequence $\left(\chi^{(n)}\right) \in \mathrm{BV}(\Omega, \Delta_l)$ exists. Furthermore, together with the lower semi-continuity of J_τ the cost functional $\mathcal{J}(\chi) + \gamma J_\tau(\chi)$ for $\chi \in \mathrm{BV}(\Omega, \Delta_l)$ is lower semi-continuous because $\mathcal{J}$ is continuous and J_τ is lower semi-continuous. The next important step is to show that this sequence lies in a sequentially compact set. It seems natural to choose the L^1-topology but *J. Lellmann* stated in his thesis [8] that this is too strong to actually find a minimum. So we choose the weak*-topology. Coercivity of $\mathcal{J} + \gamma J \tau$ with respect to the BV-norm will give the necessary upper uniform bound for the minimizing sequence: Let $\left(\chi^{(n)}\right) \in \mathrm{BV}(\Omega, \Delta_l)$ with $\left\|\chi^{(n)}\right\|_1 + \mathrm{TV}\left(\chi^{(n)}\right) \to \infty$ for $n \to \infty$. Since χ is bounded, so is $\|\chi\|_{L^1(\Omega)^l} < \infty$ and therefore, it follows that $\mathrm{TV}\left(\chi^{(n)}\right) \to \infty$. Lemma (4.2), (3.10) and $\left\|K_\eta \chi^{(n)}\right\|_{L^\infty(\Omega)} \leq 1$ for all $n \in \mathbb{N}$ show that $\mathcal{J}$ is bounded, i.e.,

$$0 \leq \mathcal{J}(\chi) \leq C \|\tilde{u}\|^2_{L^2(\Omega)} \tag{4.11}$$

and

$$J_\tau(\chi^{(n)}) \geq \mathrm{TV}(\chi^{(n)}), \tag{4.12}$$

so $\mathcal{J}(\chi^{(n)}) + \gamma J_\tau(\chi^{(n)}) \to \infty$. Thus, it is coercive. Moreover, the minimizing sequence $\left(\chi^{(n)}\right)$ is bounded in the BV-norm. Proposition (A.13) implies the existence of a weak*-convergent subsequence $\left(\chi^{(n_m)}\right) \subset \left(\chi^{(n)}\right)$, such that the corresponding limit χ^* lies in $\mathrm{BV}(\Omega, \Delta_l)$. Since $\mathcal{J} + \gamma J_\tau$ is lower semi-continuous and $\mathrm{BV}(\Omega, \Delta_l)$ is closed with respect to L^1-convergence, it follows that

$$\begin{aligned} \inf_{\chi \in \mathrm{BV}(\Omega, \Delta_l)} \mathcal{J}(\chi) + \gamma J_\tau(\chi) \leq& \mathcal{J}(\chi^*) + \gamma J_\tau(\chi^*) = \liminf_{m \to \infty} \mathcal{J}\left(\chi^{(n_m)}\right) + \gamma J_\tau\left(\chi^{(n_m)}\right) \\ =& \inf_{\chi \in \mathrm{BV}(\Omega, \Delta_l)} \mathcal{J}(\chi) + \gamma J_\tau(\chi). \end{aligned} \tag{4.13}$$

Thus, $\chi^* \in \mathrm{BV}(\Omega, \Delta_l)$ is a minimizer of the proposed functional.

□

Observe that the assumption $\tilde{u} \neq 0$ was made because a simple black image is not of practical significance, as the unwanted global minimizer is $\chi = 0$ and thus $\mathcal{J}(\chi) = 0$ and so the iteration process terminates immediately. We will later see, that the algorithm will strive for the global minimum if the initial χ_0 is not chosen properly.

4.2 The Proposed Algorithm

Finally, having proven existence of a minimum, it is now time to establish an algorithmic strategy to actually compute it. Thus, a semi-implicit gradient descent procedure will be derived. First of all, the gradient of the cost functional $\mathcal{J}$ with respect of χ will be computed. Therefore, another result, namely $\frac{\partial u_k}{\partial \chi_k}(\chi_k; \delta\chi) \in H^1(\Omega),\ k = 1, ..., l$, is needed. Secondly, other important mappings will be established that form the algorithm.

Finally, convergence of the iterative scheme will be proven with *Schauder's Fixed Point Theorem Version II*. Therefore, Lemma (4.1) will play an important role, since it showed continuity of the model functions u_k with respect to $\chi \in \mathrm{BV}(\Omega, [0,1])$.

4.2.1 The Gradient of $\mathcal{J}$ with respect to χ

This subchapter is devoted to calculating an explicit formulation for the gradient of $\mathcal{J}(\chi)$, since it is essential for the gradient descent step. In the following Theorem it is shown, that luckily the gradient has a rather simple structure and furthermore is positive.

Theorem 4.6. *Let Assumption (3.1) hold. Then the gradient of $\mathcal{J}$ with respect to χ_k for $k = 1, ...l$ reads as follows*

$$\nabla \mathcal{J}\left(\chi_k\right) = \frac{1}{2}|u_k\left(\chi_k\right) - \tilde{u}_\eta|^2\beta + \frac{1}{2}|\nabla u_k\left(\chi_k\right)|^2\alpha, \tag{4.14}$$

provided that the directional derivative $\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right) \in H^1(\Omega)$ for all $k = 1, ..., l$ exists and satisfies

$$\int_\Omega \left(\beta K_\eta \chi_k + \delta\right)\left(u_k - \tilde{u}_\eta\right)\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right) + \left(\alpha K_\eta \chi_k + \varepsilon\right)\nabla u_k \cdot \nabla\left(\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right)\right)\, dx = 0. \tag{4.15}$$

Proof. First of all, the first variation of $\mathcal{J}$ with respect to χ_k for all $k = 1, \ldots, l$ and an arbitrary perturbation $\delta\chi \in L^\infty(\Omega)$, $\delta\chi \geq 0$ is computed. Thus, $\mathcal{J}$ is everywhere Gâteaux-differentiable. Note that $K_\eta \delta\chi \in C^\infty(\Omega)$. Linearity of K_η gives

$$\begin{aligned}
\frac{\partial}{\partial \chi_k}\mathcal{J}\left(\chi_k; \delta\chi\right) =& \frac{d}{dt}\mathcal{J}\left(\chi_k + t\delta\chi\right)|_{t=0} \\
=& \frac{d}{dt}\frac{1}{2}\sum_{k=1}^{l}\int_\Omega |u_k\left(\chi_k + t\delta\chi\right) - \tilde{u}_\eta|^2\left(K_\eta\left(\chi_k + t\delta\chi\right)\beta + \delta\right) \\
&+ |\nabla u_k\left(\chi_k + t\delta\chi\right)|^2\left(K_\eta\left(\chi_k + t\delta\chi\right)\alpha + \varepsilon\right)\, dx\Big|_{t=0} \\
=& \int_\Omega \left(u_k\left(\chi_k + t\delta\chi\right) - \tilde{u}_\eta\right)\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right)\left(K_\eta\left(\chi_k + t\delta\chi\right)\beta + \delta\right) \\
&+ \frac{1}{2}|u_k\left(\chi_k\right) - \tilde{u}|^2\beta K_\eta\delta\chi + \frac{1}{2}|\nabla u_k\left(\chi_k\right)|^2\alpha K_\eta\delta\chi \\
&+ \left(K_\eta\left(\chi_k + t\delta\chi\right)\alpha + \varepsilon\right)\nabla u_k\left(\chi_k + t\delta\chi\right)\cdot\nabla\left(\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right)\right)\, dx\Big|_{t=0} \\
=& \int_\Omega \left(u_k\left(\chi_k\right) - \tilde{u}_\eta\right)\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right)\left(K_\eta\chi_k\beta + \delta\right) + \frac{1}{2}|u_k\left(\chi_k\right) - \tilde{u}_\eta|^2\beta K_\eta\delta\chi \\
&+ \left(K_\eta\chi_k\alpha + \varepsilon\right)\nabla u_k\left(\chi_k\right)\cdot\nabla\left(\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right)\right) + \frac{1}{2}|\nabla u_k\left(\chi_k\right)|^2\alpha K_\eta\delta\chi\, dx.
\end{aligned}$$

This equation can be simplified by using the condition (4.15), which is well-posed, provided that $\frac{\partial u_k}{\partial \chi_k}\left(\chi_k; \delta\chi\right) \in H^1(\Omega)$ and so the following holds

$$\frac{\partial}{\partial \chi_k}\mathcal{J}\left(\chi_k; \delta\chi\right) = \int_\Omega \frac{1}{2}|u_k\left(\chi_k\right) - \tilde{u}_\eta|^2\beta K_\eta\delta\chi + \frac{1}{2}|\nabla u_k\left(\chi_k\right)|^2\alpha K_\eta\delta\chi\, dx.$$

Finally, applying the *Riesz' Representation Theorem* completes the proof. □

Before we continue with other important results, we perform some calculations, that will be significant in proving some upcoming facts. Here, again we subtract $\left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k;v\right)\right)(\chi_k)$ from $\left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k;v\right)\right)(\chi_k+t\delta\chi)$. Notice, that the same procedure was done in Lemma (4.1). Thus, the calculation below will not be shown in full detail. Note further, that the following proofs will all advance in the same pattern, namely with the help of *Lax-Milgram* and the already established bilinear form a_k, see (3.12).

$$
\begin{aligned}
0 &= \left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k;v\right)\right)(\chi_k+t\delta\chi) - \left(\frac{\partial}{\partial u_k}\mathcal{J}\left(u_k;v\right)\right)(\chi_k) \\
\Leftrightarrow & \int_\Omega \left(\beta K_\eta \chi_k + \delta\right)\left(u_k\left(\chi_k+t\delta\chi\right) - u_k\left(\chi_k\right)\right) v \\
&+ \left(\alpha K_\eta \chi_k + \varepsilon\right)\left[\nabla u_k\left(\chi_k+t\delta\chi\right) - \nabla u_k\left(\chi_k\right)\right]\cdot\nabla v \; dx \\
&= \int_\Omega \beta t K_\eta \delta\chi\left(\tilde{u}_\eta - u_k\left(\chi_k+t\delta\chi\right)\right) v - \alpha t K_\eta \delta\chi \nabla u_k\left(\chi_k+t\delta\chi\right)\cdot\nabla v \; dx
\end{aligned}
\tag{4.16}
$$

Now we divide (4.16) by $t \neq 0$ and the equation below holds,

$$
\begin{aligned}
&\int_\Omega \left(\beta K_\eta \chi_k + \delta\right)\frac{1}{t}\left(u_k\left(\chi_k+t\delta\chi\right) - u_k\left(\chi_k\right)\right) v \\
&+ \left(\alpha K_\eta \chi_k + \varepsilon\right)\left[\nabla\left(\frac{1}{t}\left(u_k\left(\chi_k+t\delta\chi\right) - u_k\left(\chi_k\right)\right)\right)\right]\cdot\nabla v \; dx \\
&= \int_\Omega \beta K_\eta \delta\chi(\tilde{u}_\eta - u_k(\chi_k+t\delta\chi))v - \alpha K_\eta \delta\chi \nabla u_k\left(\chi_k+t\delta\chi\right)\cdot\nabla v \; dx.
\end{aligned}
\tag{4.17}
$$

We will show, that it actually holds $\exists \frac{\partial u_k}{\partial \chi_k}(\chi_k;\delta\chi) \in H^1(\Omega)$ for all $k = 1, ..., l$. So firstly, we will derive a suitable candidate for the directional derivative and secondly, show that these two correspond. Moreover, some preliminary form has to be established, that helps finding the candidate $D_{u_k}(\chi_k;\delta\chi)$.

So we use a new system with the bilinear form a_k,

$$
\begin{aligned}
a_k\left(D_{u_k}(\chi_k;\delta\chi), v\right) := &\int_\Omega \left(\beta K_\eta \chi_k + \delta\right) D_{u_k}(\chi_k;\delta\chi) v \\
&+ \left(\alpha K_\eta \chi_k + \varepsilon\right)\left(\nabla\left(D_{u_k}(\chi_k;\delta\chi)\right)\right)\cdot\nabla v \; dx
\end{aligned}
\tag{4.18}
$$

and a right-hand side b_k, for all $k = 1, ..., l$,

$$
b_k(v) := \int_\Omega \beta K_\eta \delta\chi\left(\tilde{u}_\eta - u_k\left(\chi_k\right)\right) v - \alpha K_\eta \delta\chi \nabla u_k\left(\chi_k\right)\nabla v \; dx. \tag{4.19}
$$

Note that $\frac{\partial u_k}{\partial \chi_k}(\chi_k;\delta\chi)$ cannot be calculated directly. Hence, its candidate will be determined in a weak sense. Due to the fact, that the bilinear form is the same as in (3.12), certain features are already known, i.e., it is a bounded and elliptic bilinear form.What is left to prove, is the continuity of the right-hand side, which is linear in v.

Lemma 4.7. *Given Assumption (3.1), there exists $D_{u_k}(\chi_k;\delta\chi) \in H^1(\Omega)$ for an arbitrary $\delta\chi \in \mathrm{BV}\left(\Omega,[0,1]\right)$ satisfying*

$$
a_k\left(D_{u_k}(\chi_k;\delta\chi), v\right) = b_k(v) \quad \forall v \in H^1(\Omega). \tag{4.20}
$$

Proof. Again, with the help of *Lax-Milgram* the solution of (4.20) is determined. As mentioned above, the bilinear form $a_k(\cdot,\cdot)$ is elliptic and bounded, which can be reread in Chapter 3. Thus, with the help of (3.10), it follows that

$$
\begin{aligned}
|b_k(v)| =& \left| \int_\Omega \beta K_\eta \delta\chi \left(\tilde{u}_\eta - u_k\left(\chi_k\right)\right) v - \alpha K_\eta \delta\chi \nabla u_k\left(\chi\right) \cdot \nabla v \; dx \right| \\
\leq& \beta \left\| K_\eta \delta\chi \right\|_{L^\infty(\Omega)} \left\| \tilde{u}_\eta - u_k\left(\chi_k\right) \right\|_{L^2(\Omega)} \left\| v \right\|_{L^2(\Omega)} \\
&+ \alpha \left\| K_\eta \delta\chi \right\|_{L^\infty(\Omega)} \left\| \nabla u_k\left(\chi_k\right) \right\|_{L^2(\Omega)} \left\| \nabla v \right\|_{L^2(\Omega)} \\
\leq& \left\| K_\eta \delta\chi \right\|_{L^\infty(\Omega)} \left(\beta \left\| \tilde{u}_\eta \right\|_{L^\infty(\Omega)} + (\beta+\alpha) \left\| u_k(\chi_k) \right\|_{H^1(\Omega)} \right) \left\| v \right\|_{H^1(\Omega)} \\
\leq& \tilde{c} \left\| v \right\|_{H^1(\Omega)},
\end{aligned}
$$

with $\tilde{c} = \left\| K_\eta \delta\chi \right\|_{L^\infty(\Omega)} \left(\beta + (\beta+\alpha) \frac{(\beta+\delta)}{\min(\varepsilon,\delta)} \right) \left\| \tilde{u}_\eta \right\|_{L^2(\Omega)}$.
Hence, *Lax-Milgram* guarantees existence and uniqueness of $D_{u_k}(\chi_k;\delta\chi) \in H^1(\Omega)$. □

Thus, we have derived a proper candidate for the directional derivative of u_k for all $k = 1, \ldots l$. So we are able to show the following.

Lemma 4.8. *Let Assumption (3.1) hold, then for all* $k = 1, \ldots, l$

$$
\lim_{t\to 0} \left\| \frac{u_k(\chi_k + t\delta\chi) - u_k(\chi_k)}{t} - D_{u_k}(\chi_k;\delta\chi) \right\|_{H^1(\Omega)} = 0, \tag{4.21}
$$

and thus $\frac{\partial u_k}{\partial \chi_k}(\chi_k;\delta_k)$ *identifies with* $D_{u_k}(\chi_k;\delta\chi)$ *and so* $\frac{\partial u_k}{\partial \chi_k}(\chi_k;\delta_k) \in H^1(\Omega)$.

Proof. Therefore, we begin by subtracting (4.20) from (4.17). Hence, it follows

$$
\begin{aligned}
&\int_\Omega \left(\beta K_\eta \chi_k + \delta\right) \left(\frac{1}{t} \left(u_k\left(\chi_k + t\delta\chi\right) - u_k\left(\chi_k\right) \right) - D_{u_k}(\chi_k;\delta\chi) \right) v \\
&+ \left(\alpha K_\eta \chi_k + \varepsilon\right) \nabla \left(\frac{1}{t} \left(u_k\left(\chi_k + t\delta\chi\right) - u_k\left(\chi_k\right) \right) - D_{u_k}(\chi_k;\delta\chi) \right) \cdot \nabla v \; dx \\
=& \int_\Omega \beta K_\eta \delta\chi (u_k(\chi_k + t\delta\chi) - u_k(\chi_k)) v \\
&- \alpha K_\eta \delta\chi \nabla \left(u_k\left(\chi_k + t\delta\chi\right) - u_k(\chi_k) \right) \cdot \nabla v \; dx =: \tilde{b}_k(v).
\end{aligned} \tag{4.22}
$$

The right-hand side can be estimated using the *Cauchy-Schwarz Inequality*, *Young's Inequality for Convolutions* (again taking $r, q = \infty$ and $p = 1$) and (4.4) with $\chi_k + t\delta\chi$ instead of $\chi^{(n)}$ and χ_k instead of χ,

$$
\begin{aligned}
|\tilde{b}_k(v)| \leq& \left\| \varphi_\eta \right\|_{L^\infty(\Omega)} \left\| \delta\chi \right\|_{L^1(\Omega)} \max(\beta,\alpha) \left\| u_k\left(\chi_k + t\delta\chi\right) - u_k(\chi_k) \right\|_{H^1(\Omega)} \left\| v \right\|_{H^1(\Omega)} \\
\leq& t C_2 \left\| \delta\chi \right\|^2_{L^1(\Omega)} \left\| \varphi_\eta \right\|^2_{L^\infty(\Omega)} \max(\beta,\alpha) \left\| v \right\|_{H^1(\Omega)}.
\end{aligned} \tag{4.23}
$$

Thus, the final estimate from *Lax-Milgram* (A.5) implies the following estimate,

$$\left\| \frac{u_k(\chi_k + t\delta\chi) - u_k(\chi_k)}{t} - D_{u_k}(\chi_k; \delta\chi) \right\|_{H^1(\Omega)} \leq tC_2 \left\|\varphi_\eta\right\|^2_{L^\infty(\Omega)} \left\|\delta\chi\right\|^2_{L^1(\Omega)} \max(\beta, \alpha)\frac{1}{c},$$

where c denotes the ellipticity constant of the bilinear form a_k. Letting t converge to 0 completes the proof. □

4.2.2 Establishing a Proper Algorithmic Strategy

As the title suggests, we will finally merge certain pieces to create the iterative scheme for the gradient descent approach. It is actually constructed of three mappings, which will be introduced in the following. Now we take a step back and consider the minimization problem

$$\min_{\chi \in \mathrm{BV}(\Omega, \Delta_l)} \mathcal{J}(\chi) + \gamma J_\tau(\chi). \tag{4.24}$$

An explicit formulation of $\nabla\mathcal{J}(\chi)$ was already established, so our interest focuses on a representation for a gradient of the penalty term J_τ.

Theorem 4.9. *Given Assumption (3.1). The gradient of J_τ, $0 < \tau \ll 1$ with respect to χ_k for $k = 1, ..., l$ reads as follows,*

$$\nabla J_\tau(\chi_k) = -\nabla \cdot \left(\frac{\nabla\chi_k}{\sqrt{|\nabla\chi_k|^2 + \tau}} \right), \tag{4.25}$$

if the right-hand side of (4.25) is in $L^2(\Omega)$ and

$$\frac{\partial\chi_k}{\partial n} = 0, \quad \text{on } \partial\Omega \tag{4.26}$$

Proof. Let $\delta\chi \in C_0^\infty(\bar{\Omega})$ and integration by parts provides

$$\begin{aligned}
\frac{\partial J_\tau}{\partial\chi_k}(\chi_k; \delta\chi) &= \frac{d}{dt} \int_\Omega \sqrt{|\nabla(\chi_k + t\delta\chi)|^2 + \tau}\; dx \bigg|_{t=0} \\
&= \int_\Omega \frac{2\nabla\chi_k \cdot \nabla\delta\chi}{2\sqrt{|\nabla(\chi_k + t\delta\chi)|^2 + \tau}}\; dx \bigg|_{t=0} \\
&= -\int_\Omega \nabla \cdot \left(\frac{\nabla\chi_k}{\sqrt{|\nabla(\chi_k)|^2 + \tau}} \right) \delta\chi\; dx + \int_{\partial\Omega} \frac{\partial\chi_k}{\partial n} \delta\chi\; ds_x,
\end{aligned}$$

where $\frac{\partial\chi_k}{\partial n} = 0$ on $\partial\Omega$ means the second term vanishes and thus, using *Riesz' Representation Theorem* gives the gradient of J_τ. □

Remark 4.10. *What was not emphasized, was the fact that the gradient of J_τ does not necessarily exist for any $\chi_k \in \mathrm{BV}(\Omega, [0, 1])$. Since it is not possible to deduce $\nabla\chi_k \in L^\infty(\Omega)$ with standard regularity estimates.*

Now to derive a proper iterative scheme, combine Theorem (4.9) and Theorem (4.6)to get the necessary optimality condition a minimum has to satisfy, multiply it with a parameter $\omega \in (0,1)$ and reformulate it.

$$\begin{aligned} 0 &= \nabla \mathcal{J}(\chi^*) + \gamma \nabla J_\tau(\chi^*) \\ \Leftrightarrow \quad -\omega \nabla \mathcal{J}(\chi^*) &= \omega\gamma \nabla J_\tau(\chi^*) \\ \Leftrightarrow \quad \chi^* - \omega \nabla \mathcal{J}(\chi^*) &= (\mathrm{id} + \gamma\omega \nabla J_\tau)(\chi^*) \\ \Leftrightarrow \quad \chi^* &= (\mathrm{id} + \gamma\omega\nabla J_\tau)^{-1}(\chi^* - \omega\nabla\mathcal{J}(\chi^*)). \end{aligned} \tag{4.27}$$

As already mentioned, a semi-implicit gradient descent strategy will be performed. Thus, the basic iterative scheme reads as follows

$$\chi^{(n+1)} - \omega\gamma\nabla \cdot \left(\frac{\nabla \chi^{(n+1)}}{\sqrt{|\nabla \chi^{(n+1)}|^2 + \tau}} \right) = \chi^{(n)} - \frac{\omega}{2}\left(\beta \left| u\left(\chi^{(n)}\right) - \tilde{u}_\eta \right|^2 + \alpha \left| \nabla u\left(\chi^{(n)}\right) \right|^2 \right). \tag{4.28}$$

Since we are interested in deriving an algorithm in a continuous setting and recall Remark (4.10), we will slightly change the formulation above. Therefore, the mollifier K_η is applied to $\chi^{(n)}$ in the denominator of ∇J_τ. Later on we will see that this even gives solvability in $H^1(\Omega)$ for the update. So we have

$$\chi^{(n+1)} - \omega\gamma\nabla \cdot \left(\frac{\nabla \chi^{(n+1)}}{\sqrt{|\nabla K_\eta \chi^{(n)}|^2 + \tau}} \right) = \chi^{(n)} - \frac{\omega}{2}\left(\beta \left| u\left(\chi^{(n)}\right) - \tilde{u}_\eta \right|^2 + \alpha \left| \nabla u\left(\chi^{(n)}\right) \right|^2 \right). \tag{4.29}$$

For the sake of notational brevity, we define the following function

$$G := \begin{cases} L^2(\Omega) & \to L^2(\Omega), \\ \chi & \mapsto \chi - \frac{\omega}{2}\left(\beta |u(\chi) - \tilde{u}_\eta|^2 + \alpha |\nabla u(\chi)|^2 \right), \end{cases}$$

and the operator $F(\chi): \ L^2(\Omega) \to H^1(\Omega)$, which satisfies

$$a_1(F(\chi)g, v; \chi) = (g, v)_{L^2(\Omega)}, \quad \forall v \in H^1(\Omega), \ g \in L^2(\Omega), \tag{4.30}$$

where a_1 denotes a bilinear form which will be defined later in (4.42). So the minimum has to satisfy the condition (4.27). However this does not necessarily mean that χ^* is an element of $\mathrm{BV}(\Omega, \Delta_l)$. In computational reality $\chi^*(x) \leq 0$ and $\chi^*(x) \geq 1$ for some $x \in \Omega$ can hold, since $\omega \in (0,1)$ cannot be chosen small enough, and thus there exists $x \in \Omega$ such that $G(\chi^*(x)) < 0$. The actual problem that arises concerns violating ellipticity of (3.12) and thus existence and uniqueness of the model functions u_k for $k = 1, ..., l$ cannot be guaranteed. Therefore, we introduce the mapping T that cuts G such that it has range in $[0,1]$. It is defined as follows

$$T := \begin{cases} L^2(\Omega) & \to & L^2(\Omega), \\ v & \mapsto & 1 - \max(1 - \max(v, 0), 0). \end{cases} \tag{4.31}$$

Hence, the fixed point mapping that is described by a semi-implicit gradient descent step is defined as follows

$$\chi^{(n+1)} = \Phi\left(\chi^{(n)}\right) := \left[F\left(K_\varsigma \chi^{(n)}\right) \circ T \circ G \right]\left(\chi^{(n)}\right). \tag{4.32}$$

So we are able to propose an algorithmic strategy. Choose the input parameters according to *Assumption* (3.1), $0 < \tau \ll 1$ and $\omega \in (0,1)$. Note that this was designed in a more general way, since the update is not an element of $\mathrm{BV}(\Omega, \mathcal{E})$, i. e. characteristic functions will not be computed right away. The final refinement such that the outcome will in fact be in $\mathrm{BV}(\Omega, \mathcal{E})$ will be introduced at the end of this chapter, namely the *Modified First Max* approach.

Algorithm 2 Semi-Implicit Gradient Descent Method for Functions with Range $[0,1]$

Input: $\chi_0, \tilde{u}, \alpha, \beta, \varepsilon, \delta, \gamma, \rho, n_{max}, \omega, l$
Output: χ, u
Set $n = 1$
while $\left\|\chi_k^{(n+1)} - \chi_k^{(n)}\right\|_{L^1(\Omega)^l} \geq \rho$ and $n \leq n_{max}$ **do**
  for $k = 1, ..., l$ **do**
    Calculate u_k satisfying

$$-\nabla \cdot \left[\left(K_\eta \chi_k^{(n)} \alpha + \varepsilon\right) \nabla u_k\right] + \left(\beta K_\eta \chi_k^{(n)} + \delta\right) u_k = \left(\beta \chi_k^{(n)} + \delta\right) \tilde{u}_\eta,$$

    Compute $\chi_k^{(n+1)} = \left[F\left(K_\eta \chi_k^{(n)}\right) \circ T \circ G\right]\left(\chi_k^{(n)}\right)$.

4.2.3 Existence of a Fixed Point

In the last section, the iterative scheme was presented which describes the update of χ. Therefore existence of a fixed point of

$$\Phi(\chi) = \left[F\left(K_\eta \chi\right) \circ T \circ G\right](\chi)$$

will be shown in this subsection. For that reason continuity of this composition will be proven at first.

Theorem 4.11. *Let Assumption* (3.1) *hold,* $\omega \in (0,1)$ *and recall*

$$G(\chi_k) = \chi_k - \omega\left(\frac{1}{2}|u_k(\chi_k) - \tilde{u}_\eta|^2 \beta + \frac{1}{2}|\nabla u_k(\chi_k)|^2 \alpha\right). \tag{4.33}$$

Then G *is non-expansive and continuous, i.e.,*

$$\lim_{t \to 0} \|G(\chi_k + t\delta\chi) - G(\chi_k)\|_{L^2(\Omega)} = 0. \tag{4.34}$$

Proof. Denote

$$G_1(\chi_k) := |u_k(\chi_k) - \tilde{u}_\eta|^2,$$

$$G_2(\chi_k) := |\nabla u_k(\chi_k)|^2$$

Let $\chi_k, \chi_k + t\delta\chi \in \mathrm{BV}(\Omega, [0,1])$ for $t > 0$ and some perturbation $\delta\chi \in L^\infty(\Omega)$. Denote that $\mathrm{BV}(\Omega, [0,1])$ is continuously embedded in $L^2(\Omega)$, see Appendix (A.8). So by the triangle inequality, the following is obtained

$$\begin{aligned}\|G(\chi_k + t\delta\chi) - G(\chi_k)\|_{L^2(\Omega)} \leq t\|\delta\chi\|_{L^\infty\Omega)} &+ \frac{\omega\beta}{2}\|G_1(\chi_k + t\delta\chi) - G_1(\chi_k)\|_{L^2(\Omega)} \\ &+ \frac{\alpha\omega}{2}\|G_2(\chi_k + t\delta\chi) - G_2(\chi_k)\|_{L^2(\Omega)}\end{aligned}$$

Both $G_i(\chi_k + t\delta\chi) - G_i(\chi_k),\ i = 1, 2$ have the form $a^2 - b^2 = (a-b)(a+b)$ so we are able to reformulate them such that

$$
\begin{aligned}
\|G_1(\chi_k + t\delta\chi) - G_1(\chi_k)\|_{L^2(\Omega)} \leq& \|u_k(\chi_k + t\delta\chi) + u_k(\chi_k) - 2\tilde{u}_\eta\|_{L^\infty(\Omega)} \\
&\cdot \|u_k(\chi_k + t\delta\chi) - u_k(\chi_k)\|_{L^2(\Omega)} \\
\leq& \left(\|u_k(\chi_k + t\delta\chi)\|_{L^\infty(\Omega)} + \|u_k(\chi_k)\|_{L^\infty(\Omega)} + 2\,\|\tilde{u}_\eta\|_{L^\infty(\Omega)}\right) \\
&\cdot \|u_k(\chi_k + t\delta\chi) - u_k(\chi_k)\|_{H^1(\Omega)},
\end{aligned}
$$

and

$$
\begin{aligned}
\|G_2(\chi_k + t\delta\chi) - G_2(\chi_k)\|_{L^2(\Omega)} \leq& \frac{\alpha}{2}\,\|\nabla u_k(\chi_k + t\delta\chi) + \nabla u_k(\chi_k)\|_{L^\infty(\Omega)} \\
&\cdot \|\nabla\left(u_k(\chi_k + t\delta\chi) - u_k(\chi_k)\right)\|_{L^2(\Omega)} \\
\leq& \frac{\alpha}{2}\left(\|\nabla u_k(\chi_k + t\delta\chi)\|_{L^\infty(\Omega)} + \|\nabla u_k(\chi_k)\|_{L^\infty(\Omega)}\right) \\
&\cdot \|u_k(\chi_k + t\delta\chi) - u_k(\chi_k)\|_{H^1(\Omega)},
\end{aligned}
$$

Since $u_k \in C^\infty(\bar{\Omega})$ for all $k = 1, ..., l$ and $\tilde{u}_\eta \in L^\infty(\Omega)$ the estimates hold for all $\chi_k \in \mathrm{BV}\,(\Omega)$

$$
\begin{aligned}
\|\tilde{u}_\eta\|_{L^\infty(\Omega)} &\leq 1, \\
\|u_k(\chi_k)\|_{L^\infty(\Omega)} &< \infty, \\
\|\nabla u_k(\chi_k)\|_{L^\infty(\Omega)} &< \infty.
\end{aligned}
$$

Finally, with the help of (4.4), with $\chi_k + t\delta\chi$ instead of $\chi^{(n)}$ and χ_k instead χ, the assertion is valid for some $c_1, c_2 > 0$ and for the limit $t \to 0$,

$$
\begin{aligned}
\|G(\chi_k + t\delta\chi) - G(\chi_k)\|_{L^2(\Omega)} &\leq t\,\|\delta\chi\|_{L^\infty(\Omega)} + \frac{\omega}{2}\,(\beta c_1 + \alpha c_2)\,\|u_k(\chi_k + t\delta\chi) - u_k(\chi_k)\|_{H^1(\Omega)} \\
&\leq t\,\|\delta\chi\|_{L^\infty(\Omega)}\left(1 + C_1\frac{\omega}{2}\,(\beta c_1 + \alpha c_2)\right).
\end{aligned}
$$

□

Theorem 4.12. *The map T is non-expansive and thus continuous in $L^2(\Omega)$.*

Proof. Consider the map $h(v) := \max(v, 0)$ for $v \in L^2(\Omega)$. Firstly, we show

$$
\|h(v_1) - h(v_2)\|_{L^2(\Omega)} \leq \|v_1 - v_2\|_{L^2(\Omega)}, \quad v_1, v_2 \in L^2(\Omega). \tag{4.35}
$$

This will be achieved for two cases.

1. Case: $v_1(x)v_2(x) < 0$ for some $x \in \Omega$, i.e., $v_1(x) < 0$ and $v_2(x) > 0$ (or $v_1(x) > 0$ and $v_2(x) < 0$). Then,

$$
\begin{gathered}
\max(v_1(x), 0) - \max(v_2(x), 0) = -v_2(x) \\
(\text{or } \max(v_1(x), 0) - \max(v_2(x), 0) = v_1(x)).
\end{gathered}
$$

Since $v_1(x) < 0$ and $v_2(x) > 0$, the difference $|v_1(x) - v_2(x)|$ is greater than $|-v_2(x)|$. Also for $v_2(x) < 0$ and $v_1(x) > 0$, $|v_1(x)| \leq |v_1(x) - v_2(x)|$ holds and thus,

$$\|h(v_1) - h(v_2)\|_{L^2(\Omega)} \leq \|v_1 - v_2\|_{L^2(\Omega)}.$$

2. Case: $v_1(x)v_2(x) > 0$ for some $x \in \Omega$, i.e., $v_1(x) > 0$ and $v_2(x) > 0$ (or $v_1(x) < 0$ and $v_2(x) < 0$). Then,

$$\max(v_1(x), 0) - \max(v_2(x), 0) = v_1(x) - v_2(x)$$

$$(\text{or } \max(v_1(x), 0) - \max(v_2(x), 0) = 0 - 0 = 0),$$

and thus,

$$\|h(v_1) - h(v_2)\|_{L^2(\Omega)} \leq \|v_1 - v_2\|_{L^2(\Omega)}.$$

Then it follows,

$$\begin{aligned}\|T(v_1) - T(v_2)\|_{L^2(\Omega)} &= \|1 - \max(1 - \max(v_1, 0)) - (1 - \max(1 - \max(v_2, 0)))\|_{L^2(\Omega)} \\ &\leq \|1 - \max(v_2, 0) - (1 - \max(v_1, 0))\|_{L^2(\Omega)} \\ &\leq \|v_1 - v_2\|_{L^2(\Omega)}.\end{aligned} \tag{4.36}$$

Thus, T is non-expansive. Now, take a sequence (v_n) which converges to some v in $L^2(\Omega)$, since (4.36) holds, continuity follows immediately. □

Now, define the following set

$$\mathcal{K} := \left\{\chi \in L^2(\Omega) : \; 0 \leq \chi(x) \leq 1 \text{ for almost every } x \in \Omega\right\}. \tag{4.37}$$

Recall (4.29), it is easy to see that applying $F(K_\varsigma \chi)$ to $T(G(\chi))$ is equivalent to solving the upcoming minimizing problem, which will be proven in the following Lemma,

$$\min_{\theta \in H^1(\Omega)} \frac{1}{2} \int_\Omega |\theta - T(G(\chi))|^2 + \gamma\omega \frac{|\nabla\theta|^2}{\sqrt{|\nabla(K_\varsigma \chi)|^2 + \tau}} \, dx \tag{4.38}$$

with the necessary optimality condition

$$0 = \int_\Omega (\theta - T(G(\chi)))v + \gamma\omega \frac{\nabla\theta \cdot \nabla v}{\sqrt{|\nabla(K_\varsigma \chi)|^2 + \tau}} \, dx, \quad \forall v \in H^1(\Omega). \tag{4.39}$$

Further note, that in the following proofs the index k will be dropped since the various χ_k do not depend on each other and the procedure for computing χ_k is the same for all $k = 1, ..., l$.

Lemma 4.13. *Given Assumption (3.1) and let $0 < \gamma \ll 1$, $\omega \in (0, 1)$ and assume $\chi \in \mathcal{K}$. Then (4.38) has a unique solution in $H^1(\Omega)$ and we write the solution operator as $\theta = F(K_\varsigma \chi)T(G(\chi))$.*

Proof. First of all, since $K_\varsigma \chi \in C^\infty(\bar{\Omega})$, it holds that $|\nabla(K_\varsigma \chi)|$ is bounded. So we estimate

$$\sqrt{\tau} \leq \sqrt{|\nabla(K_\varsigma \chi)|^2 + \tau} \leq \sqrt{\|\nabla(K_\varsigma \chi)\|^2_{L^\infty(\Omega)} + \tau} =: \kappa \tag{4.40}$$

and thus,

$$\frac{1}{\kappa} \leq \frac{1}{\sqrt{|\nabla(K_\varsigma\chi)|^2+\tau}} \leq \frac{1}{\sqrt{\tau}} \tag{4.41}$$

and define a bilinear form a_1 and a linear functional b

$$\begin{aligned} a_1(\theta, v; \chi) &:= \int_\Omega \theta v + \gamma\omega \frac{\nabla\theta\cdot\nabla v}{\sqrt{|\nabla\chi|^2+\tau}}\, dx, \\ b(v) &:= \int_\Omega T(G(\chi))v\; dx. \end{aligned} \tag{4.42}$$

So we apply *Lax-Milgram* to

$$a_1(\theta, v; K_\varsigma\chi) = b(v), \quad \forall v \in H^1(\Omega). \tag{4.43}$$

We start with showing boundedness and ellipticity of the bilinear form a_1 using (4.41),

$$|a_1(\theta, v; K_\varsigma\chi)| \leq \|\theta\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + \frac{\omega\gamma}{\sqrt{\tau}}\|\nabla\theta\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} \leq \left(1+\frac{\omega\gamma}{\sqrt{\tau}}\right)\|\theta\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \tag{4.44}$$

and

$$a_1(\theta,\theta;K_\varsigma\chi) \geq \int_\Omega |\theta|^2 + \frac{\omega\gamma}{\kappa}|\nabla\theta|^2\; dx \geq \min\left(1, \frac{\omega\gamma}{\kappa}\right)\|\theta\|^2_{H^1(\Omega)}. \tag{4.45}$$

Concerning boundedness of the right-hand side of (4.43), recall $T(G(\chi)) \in L^2(\Omega)$ for all $\chi \in \mathcal{K}$, since $u_k \in C^\infty(\bar{\Omega})$ holds and thus $T(G(\chi)) \in L^2(\Omega)$

$$|b(v)| \leq \|T(G(\chi))\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \leq \|T(G(\chi))\|_{L^2(\Omega)}\|v\|_{H^1(\Omega)}.$$

Hence, there exists a unique $\theta \in H^1(\Omega)$ with

$$\|\theta\|_{H^1(\Omega)} \leq \frac{1}{\min(1, \gamma\omega/\kappa)}\|T(G(\chi))\|_{L^2(\Omega)}. \tag{4.46}$$

So, the corresponding operator is derived as follows. The minimum has to satisfy the necessary optimality condition (4.39)

$$0 = \int_\Omega (\theta - T(G(\chi)))v + \gamma\omega\frac{\nabla\theta\cdot\nabla v}{\sqrt{|\nabla(K_\varsigma\chi)|^2+\tau}}\, dx, \quad \forall v\in H^1(\Omega).$$

Integration by parts leads to

$$\int_\Omega T(G(\chi))v\; dx = \int_\Omega \theta v - \gamma\omega\nabla\cdot\left(\frac{\nabla\theta}{\sqrt{|\nabla(K_\varsigma\chi)|^2+\tau}}\right) v\; dx, \quad \forall v\in H^1(\Omega).$$

Recall (4.30) and thus we have

$$(T(G(\chi)), v)_{L^2(\Omega)} = \left(F(K_\varsigma\chi)^{-1}\theta, v\right)_{L^2(\Omega)}.$$

Now, using *Fundamental Lemma of Variational Calculus* (A.3) gives

$$\theta = F(K_\varsigma\chi)T(G(\chi)).$$

Due to a Corollary of *Lax-Milgram* this operator on the left-hand side is unique and $F(K_\varsigma\chi)$ maps onto the unique solution.

□

Theorem 4.14. *Given Assumption (3.1), then* Φ *is continuous, i.e.,*

$$\lim_{n\to\infty} \left\| \Phi\left(\chi^{(n)}\right) - \Phi(\chi) \right\|_{L^2(\Omega)} = 0, \tag{4.47}$$

where $\chi^{(n)}, \chi \in \mathcal{K}$ *for all* $n \in \mathbb{N}$ *and* $\chi^{(n)} \to \chi$ *in* L^2 *for* $n \to \infty$.

Proof. Recall that $\Phi(\chi) = F(K_\varsigma \chi) T(G(\chi))$ and thus, it is a solution of (4.39). Therefore, define $\theta_n := F\left(K_\varsigma \chi^{(n)}\right) T\left(G\left(\chi^{(n)}\right)\right)$ and $\theta := F(K_\varsigma \chi) T(G(\chi))$. The proof will follow the same pattern as Theorem (4.1), so we will show that $\|\theta_n - \theta\|_{L^2(\Omega)} \to 0$. Considering (4.42), we simplify the following

$$0 = \int_\Omega v(\theta_n - \theta) - v\left(T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right)\, dx$$

$$+\gamma\omega \int_\Omega \left(\frac{\nabla \theta_n}{\sqrt{|\nabla K_\varsigma \chi^{(n)}|^2 + \tau}} - \frac{\nabla \theta}{\sqrt{|\nabla K_\varsigma \chi|^2 + \tau}} \right) \cdot \nabla v \, dx$$

By adding zero, we get

$$\begin{aligned} &\int_\Omega v(\theta_n - \theta) + \gamma\omega \int_\Omega \frac{(\nabla\theta_n - \nabla\theta)\cdot\nabla v}{\sqrt{|\nabla K_\varsigma \chi^{(n)}|^2 + \tau}}\, dx \\ =& \int_\Omega v\left(T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right) + \omega\gamma A_n \nabla\theta\cdot\nabla v \, dx =: \tilde{b}(v), \end{aligned} \tag{4.48}$$

where

$$A_n(x) := \frac{1}{\sqrt{|\nabla\left(K_\varsigma\chi\right)|^2 + \tau}} - \frac{1}{\sqrt{|\nabla\left(K_\varsigma\chi^{(n)}\right)|^2 + \tau}}.$$

To determine, if A_n converges, (4.40) gives

$$\begin{aligned} A_n(x) \leq& \frac{\left|\nabla\left(K_\varsigma\chi^{(n)}\right)\right|^2 - |\nabla\left(K_\varsigma\chi\right)|^2}{\left(\sqrt{|\nabla\left(K_\varsigma\chi^{(n)}\right)|^2 + \tau} + \sqrt{|\nabla\left(K_\varsigma\chi\right)|^2 + \tau}\right)\sqrt{|\nabla\left(K_\varsigma\chi^{(n)}\right)|^2 + \tau}\sqrt{|\nabla\left(K_\varsigma\chi\right)|^2 + \tau}} \\ \leq& \frac{1}{2\sqrt{\tau}\tau}\left(\left|\nabla\left(K_\varsigma\chi^{(n)}\right)\right| - |\nabla\left(K_\varsigma\chi\right)|\right)\left(\left|\nabla\left(K_\varsigma\chi^{(n)}\right)\right| + |\nabla\left(K_\varsigma\chi\right)|\right). \end{aligned}$$

Now applying *Young's Inequality for Convolutions* for $r = \infty$ and $p, q = 2$ and $\chi, \chi^{(n)} \leq 1$ almost everywhere leads to

$$\|A_n\|_{L^\infty(\Omega)} \leq \frac{1}{\sqrt{\tau}\tau} |\Omega|^{1/2} \|\nabla\varphi_\varsigma\|^2_{L^2(\Omega)} \left\|\chi^{(n)} - \chi\right\|_{L^2(\Omega)}. \tag{4.49}$$

Since $\left\|\chi^{(n)} - \chi\right\|_{L^2(\Omega)} \to 0$ for $n \to \infty$, it holds

$$\lim_{n\to\infty} \|A_n\|_{L^\infty(\Omega)} = 0. \tag{4.50}$$

The left-hand side of (4.48) resembles the bilinear form $a_1(\theta_n - \theta, v; \chi^{(n)})$, see (4.42), and thus it is continuous. Concerning ellipticity, the constant should be independent of n. Therefore, we take a look at the following

$$\frac{1}{\sqrt{|\nabla K_\varsigma \chi^{(n)}|^2 + \tau}} = \frac{1}{\sqrt{|\nabla K_\varsigma \chi|^2 + \tau}} - A_n(x)$$
$$\geq \frac{1}{\kappa} - \|A_n\|_{L^\infty(\Omega)} .$$

We already know that A_n converges to 0 in the L^∞-norm and thus for a sufficiently big $n \in \mathbb{N}$ it holds for $\tilde{\kappa} \geq \kappa$.

$$\frac{1}{\sqrt{|\nabla K_\varsigma \chi^{(n)}|^2 + \tau}} \geq \frac{1}{\tilde{\kappa}}$$

and hence, uniform bound holds, i.e.,

$$a_1\left(v, v; \chi^{(n)}\right) \geq \min\left(1, \frac{\gamma\omega}{\tilde{\kappa}}\right) \|v\|_{H^1(\Omega)} .$$

What is left to prove, is uniform boundedness of the right-hand side, together with $\|T(G(\chi))\|_{L^2(\Omega)} \leq |\Omega|^{1/2}$ leads to

$$|\tilde{b}(v)| \leq \|v\|_{H^1(\Omega)} \left\{ \left\|T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right\|_{L^2(\Omega)} + \omega\gamma \|A_n\|_{L^\infty(\Omega)} \|\theta\|_{H^1(\Omega)} \right\}$$
$$\leq \|v\|_{H^1(\Omega)} \left\{ \left\|T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right\|_{L^2(\Omega)}\right.$$
$$\left. + \omega\gamma \|A_n\|_{L^\infty(\Omega)} \frac{1}{\min\left(1, \gamma\omega/\kappa\right)} \|T(G(\chi))\|_{L^2(\Omega)} \right\}$$
$$\leq \|v\|_{H^1(\Omega)} \left\{ \left\|T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right\|_{L^2(\Omega)}\right.$$
$$\left. + \omega\gamma \|A_n\|_{L^\infty(\Omega)} \frac{1}{\min\left(1, \gamma\omega/\kappa\right)} |\Omega^{1/2}| \right\},$$

where (4.46) was used. From the estimates above it follows that

$$\begin{aligned} \|\theta_n - \theta\|^2_{H^1(\Omega)} &\leq \frac{1}{\min\left(1, \gamma\omega/\tilde{\kappa}\right)} \sup_{v \in H^1(\Omega)} a_1\left(\theta_n - \theta, \theta_n - \theta; \chi^{(n)}\right) \\ &\leq \frac{1}{\min\left(1, \gamma\omega/\tilde{\kappa}\right)} \left\{ \left\|T\left(G\left(\chi^{(n)}\right)\right) - T\left(G\left(\chi\right)\right)\right\|_{L^2(\Omega)} \|\theta_n - \theta\|_{H^1(\Omega)} \right. \\ &\left. + \omega\gamma \|A_n\|_{L^\infty(\Omega)} \frac{1}{\min\left(1, \gamma\omega/\kappa\right)} |\Omega|^{1/2} 6 \right\} \|\theta_n - \theta\|_{H^1(\Omega)} . \end{aligned} \tag{4.51}$$

Now we divide by $\|\theta_n - \theta\|_{H^1(\Omega)}$ and since $\|\theta_n - \theta\|_{L^2(\Omega)} \leq \|\theta_n - \theta\|_{H^1(\Omega)}$, (4.50) holds and $T \circ G$ is continuous in $L^2(\Omega)$ the right-hand side of (4.51) converges to 0 and thus, the proof is complete. □

Now we are able to give a full proof concerning convergence of our fixed point operator Φ in the following Theorem.

Theorem 4.15. *Given Assumption (3.1) and let* $0 < \tau \ll 1$, $\omega \in (0,1)$ *and*

$$\mathcal{K} := \left\{ \chi \in L^2(\Omega) : \ 0 \leq \chi \leq 1 \textit{ a.e. in } \Omega \right\}.$$

Then Φ *has a fixed point in* $\mathcal{K}$*.*

Proof. This proof will be done with *Schauder's Fixed Point Theorem Version II.* Therefore, it is necessary that Φ is continuous, which was already proven in Theorem (4.14). Furthermore, the set $\mathcal{K}$ is indeed convex, since for $\lambda \in [0,1]$ and $\chi_1, \chi_2 \in \mathcal{K}$

$$0 \leq \lambda\chi_1(x) + (1-\lambda)\chi_2(x) \leq \lambda + (1-\lambda) = 1$$

holds almost everywhere in Ω. In addition, $\mathcal{K}$ is closed. Take a sequence $\left(\chi^{(n)}\right) \subset \mathcal{K}$ that converges to some χ in $L^2(\Omega)$. Convergence in L^2 implies existence of a point-wise almost everywhere converging subsequence $(\chi^{n_l}) \subset (\chi^{(n)})$. Hence, it follows that

$$|\chi(x)| \leq |\chi(x) - \chi_{n_l}(x)| + |\chi_{n_l}(x)| \leq |\chi(x) - \chi_{n_l}(x)| + 1,$$

taking the limit $l \to \infty$ gives $|\chi(x)| \leq 1$ and of course, if $\chi_{n_l}(x) \geq 0$ for all $l \in \mathbb{N}$ then $\chi(x) \geq 0$.

Moreover, we have to show that Φ satisfies $\Phi(\mathcal{K}) \subset \mathcal{K}$. We already know that the problem (4.38) has a unique solution in $H^1(\Omega)$ and thus, $\Phi(\mathcal{K}) \subset L^2(\Omega)$ holds. Take $\chi \in \mathcal{K}$, then $\theta = \Phi(\chi)$ is a minimizer of problem (4.38). Since for $T(G(\chi))$ it holds $0 \leq T(G(\chi)) \leq 1$ almost everywhere, it is shown that also θ satisfies $0 \leq \theta \leq 1$ and thus $\theta \in \mathcal{K}$. For a similar result, see Theorem 6.95 in [4].

Define $\theta^* := \min(1, \max(0,\theta))$. Let $x \in \Omega$, then for $\theta(x) \geq 1$ it follows that

$$|\theta^*(x) - T(G(\chi))(x)| = 1 - T(G(\chi))(x) \leq |\theta(x) - T(G(\chi))(x)|.$$

Also for $\theta(x) \leq 0$ we get $|\theta^*(x) - T(G(\chi))(x)| \leq |\theta(x) - T(G(\chi))(x)|$ and hence, these estimates imply

$$\frac{1}{2}\int_\Omega |\theta^*(x) - T(G(\chi))(x)|^2 \, dx \leq \frac{1}{2}\int_\Omega |\theta(x) - T(G(\chi))(x)|^2 \, dx. \tag{4.52}$$

Furthermore, since $\theta \in H^1(\Omega)$ it also gives $\nabla\theta^* = \nabla\theta$ almost everywhere in $\{0 \leq \theta^* \leq 1\}$ and $\nabla\theta^* = 0$ almost everywhere outside of $\{0 \leq \theta^* \leq 1\}$. Hence, the following holds

$$\frac{1}{2}\int_\Omega \frac{|\nabla\theta^*|^2}{\sqrt{|\nabla(K_\varsigma\chi)|^2 + \tau}} \, dx \leq \frac{1}{2}\int_\Omega \frac{|\nabla\theta|^2}{\sqrt{|\nabla(K_\varsigma\chi)|^2 + \tau}} \, dx \tag{4.53}$$

and thus we get

$$\begin{aligned} &\frac{1}{2}\int_\Omega |\theta^*(x) - T(G(\chi))(x)|^2 + \gamma\omega\frac{|\nabla\theta^*|^2}{\sqrt{|\nabla(K_\varsigma\chi)|^2 + \tau}} \, dx \\ \leq &\frac{1}{2}\int_\Omega |\theta(x) - T(G(\chi))(x)|^2 + \gamma\omega\frac{|\nabla\theta|^2}{\sqrt{|\nabla(K_\varsigma\chi)|^2 + \tau}} \, dx. \end{aligned} \tag{4.54}$$

Because θ was proven to be the unique minimizer in $H^1(\Omega)$, it follows that $\theta^* = \theta$ almost everywhere. The statement follows from the definition of θ^* and hence $\theta = \Phi(\chi)$ is an element of $\mathcal{K}$.

Last but not least, the following condition has to be satisfied, i.e. $\Phi(\mathcal{K})$ has to be relatively compact. Therefore, for $\chi \in \mathcal{K}$ and $\theta = \Phi(\chi)$ we already know the following, since $\|T(G(\chi))\|_{L^2(\Omega)} \leq |\Omega|^{1/2}$,

$$\|\theta\|_{H^1(\Omega)} \leq \frac{1}{\min(1, \omega\gamma/\kappa)} \|T(G(\chi))\|_{L^2(\Omega)} \leq \frac{1}{\min(1, \omega\gamma/\kappa)}.$$

The above result implies boundedness in $H^1(\Omega)$ of the sequence $\left\{\Phi\left(\chi^{(n)}\right)\right\}_{n=1}^{\infty}$. Since $\Omega \subset \mathbb{R}^2$ is bounded, $H^1(\Omega)$ is compactly embedded in $L^2(\Omega)$ and thus, $\left\{\Phi\left(\chi^{(n)}\right)\right\}_{n=1}^{\infty}$ has a convergent subsequence. So $\Phi(\mathcal{K})$ is compact in $\mathcal{K}$ and hence, there exists a unique $\chi^* \in \mathcal{K}$ such that $\Phi(\chi^*) = \chi^*$.

□

Concerning, if the fixed point is actually a minimizer. We consider it the other way round in a heuristical way. Assume $\chi_k^* \in \mathrm{BV}(\Omega, [0,1])$ is a minimizer that satisfies the optimality condition

$$\nabla\mathcal{J}(\chi_k^*) + \gamma\nabla J_\tau(\chi_k^*) = 0,$$

which leads to

$$F(\chi_k^*)^{-1}\chi_k^* = G(\chi_k^*)$$

and that $G(\chi_k^*)$ has range in $[0,1]$ such that the mapping T can be omitted. Plugging this equation into the fixed point mapping Φ for $T(G(\chi_k^*))$ gives

$$\Phi(\chi_k^*) = F(K_\zeta\chi_k^*)F(\chi_k^*)^{-1}(\chi_k^*).$$

For $\zeta \to 0$ it holds $\|K_\zeta\chi_k^* - \chi_k^*\|_{L^2(\Omega)}$ converges to 0, see [6] p. 30, and thus χ_k^* satisfies $\Phi(\chi_k^*) = \chi_k^*$, i.e. the minimizer is a fixed point.

4.2.4 Modified First-Max

Unfortunately, the computed minimizer χ^* cannot likely be considered a set of suitable characteristic functions. A minimum was proven in the relaxed set $\mathrm{BV}(\Omega, \Delta_l)$, where χ_k was allowed to have range in $[0,1]$ for all $k = 1, ..., l$ with $\sum_{k=1}^{l}\chi_k(x) = 1$ for all $x \in \Omega$. But how do we transform χ^* to a set of proper characteristic functions?

Therefore, J. Lellmann proposed a method called *Modified First-Max* to map the minimizer from $\mathrm{BV}(\Omega, \Delta_l)$ to $\mathrm{BV}(\Omega, \mathcal{E})$. This heuristic rounding scheme selects the label k at the point $x \in \Omega$ corresponding to the nearest unit vector with respect to the chosen norm,

$$k(x) = \min\left\{\arg\min_{k' \in \{1,\dots,l\}} \left\|\chi(x) - e^{k'}\right\|_2\right\}. \tag{4.55}$$

Moreover, this formula can be generalized by using non-uniform distances or choosing any normalized vector instead of e^k. It performs better in practice than the *First-Max* rounding scheme, see [8] p.117.

So for the algorithm we will start sufficiently close to a minimum and after it has finished, the Modified First-Max strategy (4.55) is applied to get a suitable set of characteristic functions and then recalculate the various model function of u to get the approximation of the raw image $\tilde{u}$.

Algorithm 3 Semi-Implicit Gradient Descent Method to Compute Characteristic Functions

Input: $\chi_0, \tilde{u}, \alpha, \beta, \varepsilon, \delta, \gamma, \rho, n_{max}, \omega, l$

Output: χ, u

Set $n = 1$

while $\left\|\chi_k^{(n+1)} - \chi_k^{(n)}\right\|_{L^1(\Omega)^l} \geq \rho$ and $n \leq n_{max}$ **do**

  for $k = 1, ..., l$ **do**

    Calculate u_k satisfying

$$-\nabla \cdot \left[\left(K_\eta \chi_k^{(n)} \alpha + \varepsilon\right) \nabla u_k\right] + \left(\beta K_\eta \chi_k^{(n)} + \delta\right) u_k = \left(\beta \chi_k^{(n)} + \delta\right) \tilde{u}_\eta,$$

    Compute $\chi_k^{(n+1)} = \left[F\left(K_\eta \chi_k^{(n)}\right) \circ T \circ G\right] \left(\chi_k^{(n)}\right).$

Compute $k(x) = \min\left\{\arg\min_{k\in\{1,...,l\}} \left\|\chi(x) - e^k\right\|_2\right\}.$

Chapter 5

Numerical Approximation

This chapter focuses on the numerical consistency of the proposed method. Therefore, the discretization of Algorithm (3) is developed relying on the *Finite Element Method*. However, a suitable approximation space has to be defined first. Secondly, we establish a unique discretized solution for the model function u_k and the update of χ_k, namely θ for $k = 1, ..., l$. Thirdly, we prove that these numerical approximations are consistent. To give a sneak peak the most important tool to achieve this will be the *First Strang Lemma*. For the sake of simplicity only quadratic images with $N \times N$, where $N = 256$, are considered. Also as Chapter 2 already pointed out, only grey scale images will be of interest.

To begin with, we use the space of splines as our approximation space. Thus, it is initially defined in one dimension and extended to two dimensions using the tensor product. Let I be an interval, for our sake $I := (0, 1)$, and partition it into equidistant intervals, called $I_i := (x_{i-1}, x_i)$ with nodes $x_i := ih$ for $i = 0, ..., N$ and stepsize $h = 1/N$, which also form the grid for the finite element method. Hence, the space of splines of order q reads as follows

$$\mathcal{S}_h^q(I) := \left\{ s \in \mathcal{P}^q([x_{i-1}, x_i]) : \; s \in C^{q-1}(I), \; i = 1, 2, ..., N \right\}, \quad q = 0, 1, ..., \tag{5.1}$$

where $\mathcal{P}^q([x_{i-1}, x_i])$ denotes the space of polynomials with degree at most q defined on $[x_{i-1}, x_i]$. Now denote the canonical splines of degree q, which are defined by the convolution, i.e.

$$\pi_q(x) = (\pi_{q-1} * \pi_0)(x),$$

where

$$\pi_0(x) := \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

Thus, the splines that form the basis of $\mathcal{S}_h^q(I)$ read as follows, see also [3],

$$\bar{s}_{1+n+i}^{(q)}(x) := \pi_q\left(\frac{x - x_i}{h}\right),$$

for $i = -q, -q + 1, ..., N - 1$. Regarding the two dimensional setting, note that our domain $\Omega = (0, 1)^2$ can be rewritten as a tensor product, i.e. $\Omega = (0, 1) \otimes (0, 1)$. Therefore, the same is done concerning the space $\mathcal{S}_h^q(\Omega)$ with its base splines

$$\left\{ s_{ij}^{(q)} \right\}_{i,j=-n}^{N-1} = \left\{ \pi_q\left(\frac{x - x_i}{h}\right) \pi_q\left(\frac{y - y_j}{h}\right) \right\}_{i,j=-q}^{N-1}.$$

Note that $\dim\left(\mathcal{S}_h^q(I)\right) = N + q$ and thus, $\dim\left(\mathcal{S}_h^q(\Omega)\right) = (N+q)^2$. We rearrange the basis splines such that they have lexicographic ordering, i.e. the base of $\mathcal{S}_h^q(\Omega)$ reads $\left\{s_{ij}^{(q)}\right\}_{i,j=1}^{(N+q)^2}$. Now we are finally able to define the spline representation of any real-valued function f. Denote by $\{f_{ij}\}_{i,j=1}^{(N+q)^2}$ the values of the function f on the various nodes, then

$$f_h := \sum_{i,j=1}^{N+q} f_{ij} s_{ij}^{(q)}.$$

To approximate the model functions u_k the space of linear splines is chosen, i.e., $\mathcal{S}_h^1(\Omega)$. As a side note, the canonical linear splines have the following form

$$\pi_1(x) := \begin{cases} x, & 0 \le x \le 1, \\ 2 - x, & 1 < x \le 2, \\ 0, & \text{else}, \end{cases}.$$

Furthermore, the space of constant splines $\mathcal{S}_h^0(\Omega)$ is used to approximate the characteristic functions χ_k and the data $\tilde{u}$. Thus, the spline representations and their coefficients read as follows,

$$\begin{aligned} u_h^k &= \sum_{i,j=1}^{N+1} u_{ij}^{k,h} s_{ij}^{(1)}, & \mathbf{u_k} &= \{u_{ij}^{k,h}\}_{i,j=1}^{N+1}, \\ \chi_h^k &= \sum_{i,j=1}^{N} \chi_{ij}^{k,h} s_{ij}^{(0)}, & \mathbf{q_k} &= \{\chi_{ij}^{k,h}\}_{i,j=1}^{N}, \\ \tilde{u}_h &= \sum_{i,j=1}^{N} \tilde{u}_{ij}^{h} s_{ij}^{(0)}, & \tilde{\mathbf{u}} &= \{\tilde{u}_{ij}^{h}\}_{i,j=1}^{N}. \end{aligned} \tag{5.2}$$

The following Lemma will accompany us throughout this chapter, compare [11]. The results are presented without proof, but for the readers interest these can be found in [17] and [18].

Lemma 5.1. *1. Let $i,j \in \{1, ..., N+q\}$, then $s_{ij}^{(q)}(x) > 0$ for all $x \in (x_{i-1}, x_i) \otimes (y_{j-1}, y_j)$ and $s_{ij}^{(q)}(x) = 0$ outside of $(x_{i-1}, x_i) \otimes (y_{j-1}, y_j)$. Moreover, $\sum_{i,j=1}^{N+q} s_{ij}^{(q)}(x) = 1$ for all $x \in (x_{i-1}, x_i) \otimes (y_{j-1}, y_j)$.*

2. Let $u \in H^m(\Omega)$ and u_h be its spline interpolation. Then there exists a constant $C > 0$, that only depends on m, such that the following estimate holds

$$\|u - u_h\|_{L^\infty(\Omega)} \le C h^{m-1} \left(\sum_{i=1}^{2} \left\| \frac{\partial^m u}{\partial x_i^m} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}. \tag{5.3}$$

3. Let $q \in \{0, 1, ...\}$, then the splines $\{\mathcal{S}_h^q(\Omega)\}_h$ are dense in $H^q(\Omega)$. The space $H^0(\Omega)$ is associated with $L^2(\Omega)$.

Observe that $0 \le \chi_h^k \le 1$ almost everywhere in Ω still holds by Lemma (5.1) since $\chi_h^k(x) = \sum_{i,j=1}^{N} \chi_{ij}^{k,h} s_{ij}^{(0)}(x) \le \sum_{i,j=1}^{N} s_{ij}^{(0)}(x) = 1$ for all $x \in (x_{i-1}, x_i) \otimes (y_{j-1}, y_j)$ for $i,j = 1, ..., N$ and $\chi_h^k(x) \ge 0$ since $\chi_k(x), s_{ij}^{(0)}(x) \ge 0$ for all $x \in \Omega$. In addition, $|\Omega \backslash \bigcup_{i,j=1}^{N} (x_{i-1}, x_i) \otimes (y_{j-1}, y_j)| = 0$, i.e. its measure is zero, holds and thus, $0 \le \chi_h^k \le 1$ almost everywhere in Ω.

Lemma 5.2. *Let Assumption* (3.1) *hold, then*

$$\lim_{h\to 0} \left\| \chi_h^k - \chi_k \right\|_{L^2(\Omega)} = 0, \tag{5.4}$$

$$\lim_{h\to 0} \left\| \tilde{u}_h^\eta - \tilde{u}_\eta \right\|_{L^\infty(\Omega)} = 0, \tag{5.5}$$

$$\lim_{h\to 0} \left\| u_h^k - u_k \right\|_{L^\infty(\Omega)} = 0. \tag{5.6}$$

Proof. Applying Lemma (5.1) gives the desired results. Note that we used it on the mollified data $\tilde{u}_h^\eta$ instead of $\tilde{u}_h$. □

Having gathered the essential results from spline theory, we continue analyzing Algorithm (3). Recall that we are interested in solving the optimality systems (3.8) and (4.39). Therefore, we show well-posedness and convergence of the FEM-models.

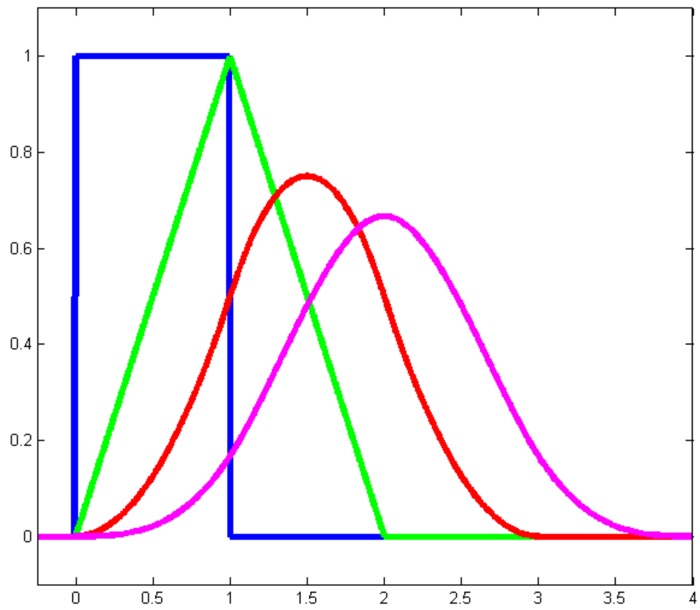

Figure 5.1: The canonical B-splines: The constant ($q = 0$, blue), linear ($q = 1$, green), quadratic ($q = 2$, red) and the cubic splines ($q = 3$, pink) are presented.

5.1 Analysis of the Finite Element Method for the Proposed Algorithm

Starting with the discretized version of the model functions u_k for $k = 1, ..., l$, we prove that there exists a unique solution $u_h^k \in \mathcal{S}_h^1(\Omega)$ and then show that the spline approximation is consistent. Therefore, we use the *First Strang Lemma*. Also numerical consistency of the gradient descent mapping G will be proven as well as for the mapping T. Finally, in this section we will also show that there exists a unique discretized solution $\theta_h \in \mathcal{S}_h^1(\Omega)$ of (4.39) and also that the approximation is consistent. Once more, this will be achieved by using the *First Strang Lemma*. Then the numerical updating process of χ will be completed.

Theorem 5.3. *Given Assumption (3.1) and let $k = 1, ..., l$. Then there exists a unique solution $u_h^k \in \mathcal{S}_h^1(\Omega)$ satisfying*

$$a_h^k(u_h, v_h) = l_h^k(v_h), \quad \forall v_h \in \mathcal{S}_h^1(\Omega), \tag{5.7}$$

with the bilinear form

$$a_h^k(u_h, v_h) := \int_\Omega (\alpha K_\eta \chi_h^k + \delta)\nabla u_h \cdot \nabla v_h + (\beta K_\eta \chi_h^k + \delta) u_h v_h \; dx \tag{5.8}$$

and the right-hand side

$$l_h^k(v_h) := \int_\Omega (\beta K_\eta \chi_h^k + \delta)\tilde{u}_h^\eta v_h \; dx. \tag{5.9}$$

Proof. This proof will be concluded with *Lax-Milgram* and proceeds analogously to Theorem (3.7). Therefore, boundedness follows from

$$a_h^k(u_h, v_h) \leq \max\{(\beta + \delta), (\alpha + \varepsilon)\} \left\| u_h^k \right\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)}$$

and ellipticity from

$$a_h^k(v_h, v_h) \geq \min(\delta, \varepsilon) \|v_h\|_{H^1(\Omega)}.$$

Now boundedness for the right-hand side holds since

$$|l_h^k(v_h)| \leq (\beta + \delta) \|\tilde{u}_h^\eta\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)}.$$

Thus, *Lax-Milgram* guarantees existence and uniqueness of the solution $u_h^k \in \mathcal{S}_h^1(\Omega)$. □

Before we proceed with the other discretized variational models, we present the Lemma that will play the most important role in two of the upcoming proofs. Since our finite element approach is non-consistent, i.e. $a_h \neq a$, this result is the best we can get.

Lemma 5.4 (First Strang Lemma [19])**.** *Let $u \in H^1(\Omega)$ be the solution of $a(u, v) = b(v)$ for all $v \in H^1(\Omega)$ and u_h the solution of $a_h(u_h, v_h) = b_h(v_h)$ for all $v_h \in \mathcal{S}_h^1(\Omega)$. Let $w_h \in \mathcal{S}_h^1(\Omega)$ be arbitrary and suppose there exist constants $C_1, C_2 > 0$ that are independent of h such that*

$$C_1 \|u_h - w_h\|_{H^1(\Omega)} \leq \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{a_h(u_h - w_h, v_h)}{\|v_h\|_{H^1(\Omega)}} \tag{5.10}$$

and

$$a(u, v_h) \leq C_2 \|u\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)} \tag{5.11}$$

for all $u \in H^1(\Omega)$ and $v_h \in \mathcal{S}_h^1(\Omega)$. Then u and v_h satisfy

$$\begin{aligned} \|u - u_h\|_{H^1(\Omega)} \leq & \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|b_h(v_h) - b(v_h)|}{\|v_h\|_{H^1(\Omega)}} \\ & + \inf_{w_h \in \mathcal{S}_h^1(\Omega)} \left\{ \left(1 + \frac{C_2}{C_1}\right) \|u - w_h\|_{H^1(\Omega)} \right. \\ & \left. + \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a(w_h, v_h) - a_h(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}} \right\}. \end{aligned} \tag{5.12}$$

Proof. Let $w_h \in \mathcal{S}_h^1(\Omega)$. Since $\mathcal{S}_h^1(\Omega) \subset H^1(\Omega)$, the equation $a(u, v_h) = b(v_h)$ holds for all $v_h \in S_h^1(\Omega)$. Furthermore, $a_h(u_h, v_h) = b_h(v_h)$ for all $v_h \in \mathcal{S}_h^1(\Omega)$. Thus,

$$a_h(u_h - w_h, v_h) = a(u - w_h, v_h) + a(w_h, v_h) - a_h(w_h, v_h) + b_h(v_h) - b(v_h).$$

Now by dividing by $\|v_h\|_{H^1(\Omega)}$ and taking the supremum over all $v_h \in \mathcal{S}_h^1(\Omega)$ leads to

$$\sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{a_h(u_h - w_h, v_h)}{\|v_h\|_{H^1(\Omega)}} \leq \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{a_h(u - w_h, v_h)}{\|v_h\|_{H^1(\Omega)}} + \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|b_h(v_h) - b(v_h)|}{\|v_h\|_{H^1(\Omega)}} + \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a(w_h, v_h) - a_h(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}}.$$

Then the assumptions (5.10) and (5.11) give

$$C_1 \|u_h - w_h\|_{H^1(\Omega)} \leq C_2 \|u - w_h\|_{H^1(\Omega)} + \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|b_h(v_h) - b(v_h)|}{\|v_h\|_{H^1(\Omega)}} + \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a(w_h, v_h) - a_h(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}}. \tag{5.13}$$

Combining the above estimate (5.13) with the following which is obtained with the help of the triangle inequality

$$\|u - u_h\|_{H^1(\Omega)} \leq \|u - w_h\|_{H^1(\Omega)} + \|u_h - w_h\|_{H^1(\Omega)}$$

leads to

$$\|u - u_h\|_{H^1(\Omega)} \leq \|u - w_h\|_{H^1(\Omega)} + \frac{C_2}{C_1} \|u - w_h\|_{H^1(\Omega)} + \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|b_h(v_h) - b(v_h)|}{\|v_h\|_{H^1(\Omega)}} + \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a(w_h, v_h) - a_h(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}}$$

So, taking the infimum over all $w_h \in \mathcal{S}_h^1(\Omega)$ proves the assertion. □

Theorem 5.5. *Given Assumption* (3.1) *and let* $k = 1, ..., l,$ *then*

$$\lim_{h \to 0} \left\| u_h^k - u^k \right\|_{H^1(\Omega)} = 0. \tag{5.14}$$

Proof. Let $u_h^k \in \mathcal{S}_h^1(\Omega)$ be the solution of the variational equation $a_h^k(u_h, v_h) = l_h^k(v_h)$ for all $v_h \in \mathcal{S}_h^1(\Omega)$, recall (5.7), and u_k the solution of $a_k(u^k, v) = l_k(v)$ for all $v \in H^1(\Omega)$. So to begin with, the two requirements of the *First Strang Lemma* will be proven. The first one is already valid considering that the bilinear form a_k is elliptic, i.e.,

$$C \|u_h - w_h\|_{H^1(\Omega)} \leq \frac{a_h(u_h^k - w_h, u_h^k - w_h)}{\|u_h^k - w_h\|_{H^1(\Omega)}} \leq \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{a_h(u_h^k - w_h, v_h)}{\|v_h\|_{H^1(\Omega)}}$$

where $C > 0$ is the ellipticity constant of (3.12). The second condition follows from

$$a_k(u_k, v_h) \leq |a_k(u_k, v_h)| \leq \max\{(\beta + \delta), (\alpha + \varepsilon)\} \|u_k\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)},$$

since $\max\{(\beta+\delta),(\alpha+\varepsilon)\}$ is clearly a constant independent of h. Thus, the *Strang Lemma* holds, i.e. there exist constants $C_1, C_2 > 0$ that are independent of h such that

$$\begin{aligned}\left\|u_h^k - u_k\right\|_{H^1(\Omega)} \leq & \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|l_h^k(v_h) - l_k(v_h)|}{\|v_h\|_{H^1(\Omega)}} \\ & + \inf_{w_h \in \mathcal{S}_h^1(\Omega)} \left\{ \left(1 + \frac{C_2}{C_1}\right) \|u_k - w_h\|_{H^1(\Omega)} \right. \\ & \left. + \frac{1}{C_1} \sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a_k(w_h, v_h) - a_h^k(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}} \right\}.\end{aligned} \tag{5.15}$$

Now we continue by estimating the following by adding zero

$$\begin{aligned}|l_h^k(v_h) - l_k(v_h)| &= \left| \int_\Omega \left[(\beta K_\eta \chi_h^k + \delta)\tilde{u}_h^\eta - (\beta K_\eta \chi_k + \delta)\tilde{u}_\eta \right] v_h \; dx \right| \\ &\leq \int_\Omega |(\beta K_\eta \chi_h^k + \delta)(\tilde{u}_h^\eta - \tilde{u}_\eta) v_h| + |\tilde{u}_\eta v_h \beta (K_\eta \chi_h^k - K_\eta \chi_k)| \; dx \\ &\leq (\beta + \delta) \|\tilde{u}_\eta - \tilde{u}_h^\eta\|_{L^\infty(\Omega)} \|v_h\|_{H^1(\Omega)} \\ &+ \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \beta \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \|v_h\|_{H^1(\Omega)}.\end{aligned}$$

We divide by $\|v_h\|_{H^1(\Omega)}$ and taking the supremum over all v_h leads to the estimate

$$\sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|l_h^k(v_h) - l_k(v_h)|}{\|v_h\|_{H^1(\Omega)}} \leq (\beta+\delta) \|\tilde{u}_\eta - \tilde{u}_h^\eta\|_{L^\infty(\Omega)} + \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \beta \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)}. \tag{5.16}$$

Considering the third term, it holds

$$\begin{aligned}|a_k(w_h, v_h) - a_h^k(w_h, v_h)| &\leq \int_\Omega \left| \beta \left(K_\eta \chi_h^k - K_\eta \chi_k \right) w_h v_h \right| \\ &+ \left| \alpha \left(K_\eta \chi_h^k - K_\eta \chi_k \right) \nabla w_h \cdot \nabla v_h \right| \; dx \\ &\leq \max(\beta, \alpha) \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \|w_h\|_{H^1(\Omega)} \|v_h\|_{H^1(\Omega)}.\end{aligned}$$

Analogously to the above result by dividing and taking the supremum over all v_h we get

$$\sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a_k(w_h, v_h) - a_h^k(w_h, v_h)|}{\|v_h\|_{H^1(\Omega)}} \leq \max(\beta, \alpha) \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \|w_h\|_{H^1(\Omega)}. \tag{5.17}$$

So (5.15) can be expressed as

$$\begin{aligned}\left\|u_h^k - u_k\right\|_{H^1(\Omega)} \leq & \frac{1}{C_1} \Big\{ (\beta + \delta) \|\tilde{u}_\eta - \tilde{u}_h^\eta\|_{L^\infty(\Omega)} \\ & + \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \beta \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \Big\} + \mathcal{F}(h),\end{aligned} \tag{5.18}$$

where

$$\mathcal{F}(h) := \inf_{w_h \in \mathcal{S}_h^1(\Omega)} \left\{ \left(1 + \frac{C_2}{C_1}\right) \|u_k - w_h\|_{H^1(\Omega)} \right.$$
$$\left. + \frac{1}{C_1} \max(\beta, \alpha) \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \|w_h\|_{H^1(\Omega)} \right\}. \tag{5.19}$$

Since $u_k \in H1(\Omega)$, there exists an interpolation $\mathcal{I}_h u_k \in \mathcal{S}_h^1(\Omega)$. Note that $\mathcal{I}_h u_k$ differs from u_h^k, since u_h^k is the solution of the discretized variational formulation $a_h^k(u_h^k, v_h) = l_h^k(v_h)$ whereas $\mathcal{I}_h u_k$ is the spline interpolation of u_k. Hence,

$$\mathcal{F}(h) \leq \left\{ \left(1 + \frac{C_2}{C_1}\right) \|u_k - \mathcal{I}_h u_k\|_{H^1(\Omega)} + \frac{1}{C_1} \max(\beta, \alpha) \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \|\mathcal{I}_h u_k\|_{H^1(\Omega)} \right\} \tag{5.20}$$

holds and with the use of the triangle inequality

$$\|\mathcal{I}_h u_k\|_{H^1(\Omega)} \leq \|\mathcal{I}_h u_k - u_k\|_{H^1(\Omega)} + \|u_k\|_{H^1(\Omega)} \tag{5.21}$$

and applying the second assertion of Lemma (5.1) gives

$$\mathcal{F}(h) \leq \left(1 + \frac{C_2}{C_1}\right) C_3 h \left(\sum_{i=1}^{2} \left\| \frac{\partial^2 u_k}{\partial x_i^2} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}$$
$$+ \frac{1}{C_1} \max(\beta, \alpha) \left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \left\{ C_3 h \left(\sum_{i=1}^{2} \left\| \frac{\partial^2 u_k}{\partial x_i^2} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + \|u_k\|_{H^1(\Omega)} \right\}. \tag{5.22}$$

Then *Young's Inequality for Convolution* for $r = \infty$ and $p, q = 2$ gives

$$\left\| K_\eta \chi_h^k - K_\eta \chi_k \right\|_{L^\infty(\Omega)} \leq \|\varphi_\eta\|_{L^2(\Omega)} \left\| \chi_h^k - \chi_k \right\|_{L^2(\Omega)}.$$

Thus, (5.18) is estimated

$$\left\| u_h^k - u_k \right\|_{H^1(\Omega)} \leq \frac{1}{C_1} \left\{ (\beta + \delta) \|\tilde{u}_\eta - \tilde{u}_h^\eta\|_{L^\infty(\Omega)} + \|\tilde{u}_\eta\|_{L^\infty(\Omega)} \beta \|\varphi_\eta\|_{L^2(\Omega)} \left\| \chi_h^k - \chi_k \right\|_{L^2(\Omega)} \right\}$$
$$+ \left(1 + \frac{C_2}{C_1}\right) C_3 h \left(\sum_{i=1}^{2} \left\| \frac{\partial^2 u_k}{\partial x_i^2} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}$$
$$+ \frac{1}{C_1} \max(\beta, \alpha) \|\varphi_\eta\|_{L^2(\Omega)} \left\| \chi_h^k - \chi_k \right\|_{L^2(\Omega)}$$
$$\cdot \left\{ C_3 h \left(\sum_{i=1}^{2} \left\| \frac{\partial^2 u_k}{\partial x_i^2} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + \|u_k\|_{H^1(\Omega)} \right\}. \tag{5.23}$$

By letting h go to 0 the propositions of Lemma (5.2) complete the proof. □

Now, we consider the update of χ_h^k. Therefore, we start with proving the following Lemma which will be needed for $G(\chi_h)$ and thus, for the finite element approach of θ_h.

Lemma 5.6. *Given Assumption* (3.1)*, then*

$$\lim_{h\to 0} \left\| \nabla u_h^k \left(\chi_h^k\right) \right\|_{L^\infty(\Omega)} < \infty. \tag{5.24}$$

Proof. This proof will be done by contradiction. Assume $\lim_{h\to 0} \left\| \nabla u_h^k(\chi_h^k) \right\|_{L^\infty(\Omega)} = \infty$, i.e. there exists $R > 0$ such that for h small enough

$$\left\| \nabla u_h^k \left(\chi_h^k\right) \right\|_{L^\infty(\Omega)} > R. \tag{5.25}$$

Again, the optimality system for u_k, recall (3.8) will be used, i.e. for all $v_h \in \mathcal{S}_h^1(\Omega)$ holds

$$\int_\Omega \left(\alpha K_\eta \chi_h^k + \varepsilon\right) \nabla u_h^k \left(\chi_h^k\right) \cdot \nabla v_h + \left(\beta K_\eta \chi_h^k + \delta\right) u_h^k \left(\chi_h^k\right) v_h \; dx = \int_\Omega \tilde{u}_h^\eta v_h \left(\beta K_\eta \chi_h^k + \delta\right).$$

Inserting $u_h^k \left(\chi_h^k\right)$ for v_h gives

$$\int_\Omega \left(\alpha K_\eta \chi_h^k + \varepsilon\right) \left|\nabla u_h^k \left(\chi_h^k\right)\right|^2 \; dx = \int_\Omega \left(\tilde{u}_h^\eta - u_h^k \left(\chi_h^k\right)\right) u_h^k \left(\chi_h^k\right) \left(\beta K_\eta \chi_h^k + \delta\right).$$

The left-hand side can be estimated using (5.25), such that

$$\varepsilon |\Omega| R^2 < \int_\Omega \left(\tilde{u}_h^\eta - u_h^k \left(\chi_h^k\right)\right) u_h^k \left(\chi_h^k\right) \left(\beta K_\eta \chi_h^k + \delta\right).$$

Now adding zero and using (3.10), the right-hand side leads to the following

$$\begin{aligned}
&\int_\Omega \left(\tilde{u}_h^\eta - u_h^k \left(\chi_h^k\right)\right) u_h^k \left(\chi_h^k\right) \left(\beta K_\eta \chi_h^k + \delta\right) \\
&\leq (\beta + \delta)\Big(\|\tilde{u}_h^\eta - \tilde{u}_\eta\|_{L^2(\Omega)} + \|\tilde{u}_\eta - u_k(\chi_k)\|_{L^2(\Omega)} \\
&+ \left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{L^2(\Omega)}\Big) \left(\left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{L^2(\Omega)} + \|u_k(\chi_k)\|_{L^2(\Omega)}\right) \\
&\leq (\beta + \delta)\Big[\|\tilde{u}_h^\eta - \tilde{u}_\eta\|_{L^\infty(\Omega)} + \|\tilde{u}_\eta - u_k(\chi_k)\|_{L^2(\Omega)} \\
&+ \left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{H^1(\Omega)}\Big] \left[\left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{H^1(\Omega)} + \|u_k(\chi_k)\|_{L^2(\Omega)}\right]. \\
&\leq (\beta + \delta) \Bigg[\|\tilde{u}_h^\eta - \tilde{u}_\eta\|_{L^\infty(\Omega)} + \left(1 + \frac{\beta + \delta}{\min(\varepsilon, \delta)}\right) \|\tilde{u}_\eta\|_{L^2(\Omega)} \\
&+ \left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{H^1(\Omega)}\Bigg] \left[\left\|u_k(\chi_k) - u_h^k \left(\chi_h^k\right)\right\|_{H^1(\Omega)} + \frac{\beta + \delta}{\min(\varepsilon, \delta)} \|\tilde{u}_\eta\|_{L^2(\Omega)}\right].
\end{aligned}$$

Letting h tend to 0, implies that the right-hand side is bounded, since $\tilde{u}_\eta - u_k(\chi_k), u_k(\chi_k) \in L^2(\Omega)$ and Theorem (5.5) and Lemma (5.1) hold. However this contradicts our assumption $\lim_{h\to 0} \left\| \nabla u_h^k(\chi_h^k) \right\|_{L^\infty(\Omega)} = \infty$ at the beginning of the proof. Therefore, we get

$$\lim_{h\to 0} \left\| \nabla u_h^k(\chi_h^k) \right\|_{L^\infty(\Omega)} < \infty$$

which proves the assertion.

□

Theorem 5.7. *Given Assumption (3.1), then*

$$\lim_{h\to 0}\left\|G(\chi_h^k)-G(\chi_k)\right\|_{L^2(\Omega)}=0. \tag{5.26}$$

Proof. We will proceed in a similar way than in Theorem (4.11), so again we denote the following for G

$$G_1(\chi_k):=|u_k(\chi_k)-\tilde{u}_\eta|^2,$$
$$G_2(\chi_k):=|\nabla u_k(\chi_k)|^2,$$

as well as

$$G_1(\chi_h^k):=|u_h^k(\chi_h^k)-\tilde{u}_h^\eta|^2,$$
$$G_2(\chi_h^k):=|\nabla u_h^k(\chi_h^k)|^2.$$

So by the triangle inequality, the following is obtained

$$\begin{aligned}\left\|G(\chi_h^k)-G(\chi_k)\right\|_{L^2(\Omega)}\leq\left\|\chi_h^k-\chi_k\right\|_{L^2(\Omega)}+\frac{\omega\beta}{2}\left\|G_1(\chi_h^k)-G_1(\chi_k)\right\|_{L^2(\Omega)}\\+\frac{\alpha\omega}{2}\left\|G_2(\chi_h^k)-G_2(\chi_k)\right\|_{L^2(\Omega)}\end{aligned}$$

As in Theorem (4.11), both $G_i(\chi_h^k)-G_i(\chi_k),\ i=1,2$ have the form $a^2-b^2=(a-b)(a+b)$ which again leads to

$$\begin{aligned}\left\|G_1(\chi_h^k)-G_1(\chi_k)\right\|_{L^2(\Omega)}&\leq\left\|u_h^k(\chi_h^k)+u_k(\chi_k)-2\tilde{u}_\eta\right\|_{L^\infty(\Omega)}\left\|u_h^k(\chi_h^k)-u_k(\chi_k)\right\|_{L^2(\Omega)}\\&\leq\left(\left\|u_h^k(\chi_h^k)\right\|_{L^\infty(\Omega)}+\|u_k(\chi_k)\|_{L^\infty(\Omega)}+2\left\|\tilde{u}\right\|_{L^\infty(\Omega)}\right)\\&\quad\cdot\left\|u_h^k(\chi_h^k)-u_k(\chi_k)\right\|_{H^1(\Omega)},\end{aligned}$$

and

$$\begin{aligned}\left\|G_2(\chi_h^k)-G_2(\chi_k)\right\|_{L^2(\Omega)}\leq&\frac{\alpha}{2}\left\|\nabla u_h^k(\chi_h^k)+\nabla u_k(\chi_k)\right\|_{L^\infty(\Omega)}\\&\cdot\left\|\nabla\left(u_h^k(\chi_h^k)-u_k(\chi_k)\right)\right\|_{L^2(\Omega)}\\\leq&\frac{\alpha}{2}\left(\left\|\nabla u_h^k(\chi_h^k)\right\|_{L^\infty(\Omega)}+\|\nabla u_k(\chi_k)\|_{L^\infty(\Omega)}\right)\\&\cdot\left\|u_h^k(\chi_h^k)-u_k(\chi_k)\right\|_{H^1(\Omega)},\end{aligned}$$

We know $\lim_{h\to 0}\left\|u_h^k(\chi_h^k)\right\|_{L^\infty(\Omega)}<\infty$ holds since Lemma (5.1), $u_k\in C^\infty(\Omega)$ for all $k=1,...,l$ and

$$\left\|u_h^k\right\|_{L^\infty(\Omega)}\leq\left\|u_h^k-u_k\right\|_{L^\infty(\Omega)}+\|u_k\|_{L^\infty(\Omega)} \tag{5.27}$$

give an upper bound. Also Lemma (5.6) guarantees $\lim_{h\to 0}\left\|\nabla u_h^k(\chi_h^k)\right\|_{L^\infty(\Omega)} < \infty$. Also the discretized data term is bounded for $h \to 0$, i.e. $\lim_{h\to 0}\|\tilde{u}_h^\eta\|_{L^\infty(\Omega)} < \infty$, which follows from the same argument as (5.27). For some $\tilde{c}_1, \tilde{c}_2 > 0$ holds

$$\left\|G(\chi_h^k) - G(\chi_k)\right\|_{L^2(\Omega)} \leq \left\|\chi_h^k - \chi_k\right\|_{L^2(\Omega)} + \frac{\omega}{2}\left(\beta\tilde{c}_1 + \alpha\tilde{c}_2\right)\left\|u_h^k(\chi_h^k) - u_k(\chi_k)\right\|_{H^1(\Omega)}.$$

Together with (5.5) and letting h tend to 0, the assertion holds true. □

Theorem 5.8. *Given Assumption (3.1), then*

$$\lim_{h\to 0}\|T(\chi_h) - T(\chi)\|_{L^2(\Omega)} = 0. \tag{5.28}$$

Proof. Let χ_h the spline approximation of $\chi \in L^2(\Omega)$. Since $\{S_h^0(\Omega)\}_h$ is dense in $L^2(\Omega)$, χ_h is an element in $L^2(\Omega)$ as well. Hence, we apply Theorem (4.12)

$$\|T(\chi_h) - T(\chi)\|_{L^2(\Omega)} \leq \|\chi_h - \chi\|_{L^2(\Omega)}.$$

Letting h tend to 0 implies the desired since $\{S_h^0(\Omega)\}_h$ is dense in $L^2(\Omega)$. □

Finally, we can prove existence of the discretized solution of (4.39) and that the numerical approximation is convergent.

Theorem 5.9. *Given Assumption (3.1) and $\chi_h \in \mathcal{K}$, then there exists a unique solution $\theta_h \in \mathcal{S}_h^1(\Omega)$ which satisfies*

$$\int_\Omega \theta_h v_h + \gamma\omega\frac{\nabla\theta_h\cdot\nabla v_h}{\sqrt{|\nabla(K_\eta\chi_h)|^2+\tau}}\,dx = \int_\Omega T(G(\chi_h))v_h\,dx, \quad \forall v_h \in \mathcal{S}_h^1(\Omega). \tag{5.29}$$

Proof. The conditions of *Lax-Milgram* will be verified to obtain existence and uniqueness of a solution that suffices (5.29). Therefore, we use the following estimate

$$\frac{1}{\kappa_h} \leq \frac{1}{\sqrt{|\nabla(K_\eta\chi_h)|^2+\tau}} \leq \frac{1}{\sqrt{\tau}},$$

where

$$\sqrt{|\nabla(K_\eta\chi_h)|^2+\tau} \leq \sqrt{\|\nabla(K_\eta\chi_h)\|_{L^\infty(\Omega)}^2+\tau} =: \kappa_h$$

Letting $v_h \in \mathcal{S}_h^1(\Omega)$, starting with boundedness of the bilinear form gives

$$|a_1(\theta_h, v_h; K_\zeta\chi_h)| := \left|\int_\Omega \theta_h v_h + \gamma\omega\frac{\nabla\theta_h\cdot\nabla v_h}{\sqrt{|\nabla(K_\eta\chi_h)|^2+\tau}}\,dx\right|$$

$$\leq \max\left(1, \frac{\omega\gamma}{\sqrt{\tau}}\right)\|\theta_h\|_{H^1(\Omega)}\|v_h\|_{H^1(\Omega)}$$

Ellipticity is achieved by

$$a_1(v_h, v_h; K_\zeta\chi_h) \geq \min\left(1, \frac{\omega\gamma}{\kappa_h}\right)\|v_h\|_{H^1(\Omega)}^2.$$

Last but not least, boundedness of the right-hand side is guaranteed since $T(G(\chi_h)) \in L^2(\Omega)$,

$$|b_h(v_h)| := \left|\int_\Omega T(G(\chi_h))v_h\,dx\right| \leq \|T(G(\chi_h))\|_{L^2(\Omega)}\|v_h\|_{H^1(\Omega)}.$$

Thus, *Lax-Milgram* guarantees existence and uniqueness of $\theta_h \in \mathcal{S}_h^1(\Omega)$. □

Theorem 5.10. *Given Assumption (3.1), then*

$$\lim_{h\to 0}\|\theta_h-\theta\|_{H^1(\Omega)}=0, \tag{5.30}$$

where θ solves (4.39).

Proof. Let $\theta_h\in\mathcal{S}_h^1(\Omega)$ be the solution of (5.29). First of all, the two conditions of the *First Strang Lemma* will be proven. To verify the first one, consider

$$\frac{1}{\sqrt{|\nabla\left(K_\zeta\chi_h\right)|^2+\tau}}=\frac{1}{\sqrt{|\nabla\left(K_\zeta\chi\right)|^2+\tau}}+B(h)$$

with

$$B(h):=\frac{1}{\sqrt{|\nabla\left(K_\eta\chi_h\right)|^2+\tau}}-\frac{1}{\sqrt{|\nabla\left(K_\eta\chi\right)|^2+\tau}}.$$

As in (4.49), we get the following estimate

$$\|B(h)\|_{L^\infty(\Omega)}\leq\frac{1}{\sqrt{\tau}\tau}|\Omega|^{1/2}\|\nabla\varphi_\zeta\|^2_{L^2(\Omega)}\|\chi_h-\chi\|_{L^2(\Omega)}$$

and so Lemma (5.2) leads to

$$\lim_{h\to 0}\|B(h)\|_{L^\infty(\Omega)}=0. \tag{5.31}$$

Considering (5.31), h can be chosen sufficiently small such that for a $\bar{\kappa}\geq\kappa$

$$\frac{1}{\sqrt{|\nabla\left(K_\zeta\chi_h\right)|^2+\tau}}\geq\frac{1}{\kappa}-\|B(h)\|_{L^\infty(\Omega)}\geq\frac{1}{\bar{\kappa}},$$

compare Theorem (4.14). Then it follows from ellipticity of the bilinear form that

$$\min\left(1,\frac{\gamma\omega}{\bar{\kappa}}\right)\|\theta_h-w_h\|_{H^1(\Omega)}\leq\frac{a_1(\theta_h-w_h,\theta_h-w_h;K_\zeta\chi_h)}{\|\theta_h-w_h\|_{H^1(\Omega)}}\leq\sup_{v_h\in\mathcal{S}_h^1(\Omega)}\frac{a_1(\theta_h-w_h,v_h;K_\zeta\chi_h)}{\|v_h\|_{H^1(\Omega)}}.$$

The second condition follows from

$$a_1(\theta,v_h;K_\zeta\chi)\leq|a_1(\theta,v_h;K_\zeta\chi)|\leq\max\left(1,\frac{\omega\gamma}{\sqrt{\tau}}\right)\|\theta\|_{H^1(\Omega)}\|v_h\|_{H^1(\Omega)},$$

where the constant $\max\left(1,\frac{\omega\gamma}{\sqrt{\tau}}\right)$ is clearly independent of h. Thus, the *Strang Lemma* holds, i.e. there exist constants $C_1,C_2>0$ that are independent of h such that

$$\begin{aligned}\|\theta_h-\theta\|_{H^1(\Omega)}\leq&\frac{1}{C_1}\sup_{v_h\in\mathcal{S}_h^1(\Omega)}\frac{|b_h(v_h)-b(v_h)|}{\|v_h\|_{H^1(\Omega)}}\\&+\inf_{w_h\in\mathcal{S}_h^1(\Omega)}\left\{\left(1+\frac{C_2}{C_1}\right)\|\theta-w_h\|_{H^1(\Omega)}\right.\\&\left.+\frac{1}{C_1}\sup_{v_h\in\mathcal{S}_h^1(\Omega)}\frac{|a_1(w_h,v_h;K_\zeta\chi)-a_1(w_h,v_h;K_\zeta\chi_h)|}{\|v_h\|_{H^1(\Omega)}}\right\}.\end{aligned} \tag{5.32}$$

The first term can be estimated with (5.8)

$$\sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|b_h(v_h) - b(v_h)|}{\|v_h\|_{H^1(\Omega)}} \leq \|T(G(\chi_h)) - T(G(\chi))\|_{L^2(\Omega)} \leq \|G(\chi_h) - G(\chi)\|_{L^2(\Omega)} . \tag{5.33}$$

Considering the third term, it holds

$$\sup_{v_h \in \mathcal{S}_h^1(\Omega)} \frac{|a_1(w_h, v_h; K_\zeta \chi) - a_1(w_h, v_h; K_\zeta \chi_h)|}{\|v_h\|_{H^1(\Omega)}} \leq \omega\gamma \|B(h)\|_{L^\infty(\Omega)} \|w_h\|_{H^1(\Omega)} . \tag{5.34}$$

So (5.32) can be expressed as

$$\|\theta_h - \theta\|_{H^1(\Omega)} \leq \frac{1}{C_1} \|G(\chi_h) - G(\chi)\|_{L^2(\Omega)} + \mathcal{W}(h), \tag{5.35}$$

where

$$\mathcal{W}(h) := \inf_{w_h \in \mathcal{S}_h^1(\Omega)} \left\{ \left(1 + \frac{C_2}{C_1}\right) \|\theta - w_h\|_{H^1(\Omega)} + \frac{1}{C_1} \omega\gamma \|B(h)\|_{L^\infty(\Omega)} \|w_h\|_{H^1(\Omega)} \right\} . \tag{5.36}$$

Since $\theta \in H^1(\Omega)$ there exists an interpolation $\mathcal{I}_h\theta \in \mathcal{S}_h^1(\Omega)$. Hence,

$$\mathcal{W}(h) \leq \left\{ \left(1 + \frac{C_2}{C_1}\right) \|\theta - \mathcal{I}_h\theta\|_{H^1(\Omega)} + \frac{1}{C_1} \omega\gamma \|B(h)\|_{L^\infty(\Omega)} \|\mathcal{I}_h\theta\|_{H^1(\Omega)} \right\} \tag{5.37}$$

holds and again with the use of the triangle inequality

$$\|\mathcal{I}_h\theta\|_{H^1(\Omega)} \leq \|\mathcal{I}_h\theta - \theta\|_{H^1(\Omega)} + \|\theta\|_{H^1(\Omega)} . \tag{5.38}$$

Inserting the above estimates to (5.35) implies

$$\begin{aligned} \|\theta_h - \theta\|_{H^1(\Omega)} \leq & \frac{1}{C_1} \|G(\chi_h) - G(\chi)\|_{L^2(\Omega)} + \left(1 + \frac{C_2}{C_1}\right) \|\theta - \mathcal{I}_h\theta\|_{H^1(\Omega)} \\ & + \frac{1}{C_1} \omega\gamma \|B(h)\|_{L^\infty(\Omega)} \left(\|\mathcal{I}_h\theta - \theta\|_{H^1(\Omega)} + \|\theta\|_{H^1(\Omega)} \right) . \end{aligned} \tag{5.39}$$

Finally, (5.31), Lemma (5.1) and Theorem (5.7) ensures that (5.39) converges to 0 for the limit $h \to 0$. □

5.2 Discretization of the Proposed Method

This subchapter is devoted to the discretization of the finite element approach. We have already shown, that the numerical approximations are consistent and thus, we proceed to rewrite Algorithm (3) such that it can be implemented numerically. Remark, since we stated that our mollifying operator is just a technical necessity, it will not emerge in this subchapter. Now recall that gray scale images with resolution $N \times N$, where $N = 256$, are considered and thus, the step size is $h = 1/N$. So, further recall our initial approximations and their coefficients in (5.2),

$$u_h^k = \sum_{i,j=1}^{N+1} u_{ij}^{k,h} s_{ij}^{(1)}, \qquad \mathbf{u_k} = \{u_{ij}^{k,h}\}_{i,j=1}^{N+1} \in \mathbb{R}^{(N+1)^2},$$

$$\chi_h^k = \sum_{i,j=1}^{N} \chi_{ij}^{k,h} s_{ij}^{(0)}, \qquad \mathbf{q_k} = \{\chi_{ij}^{k,h}\}_{i,j=1}^{N} \in \mathbb{R}^{N^2},$$

$$\tilde{u}_h = \sum_{i,j=1}^{N} \tilde{u}_{ij}^{h} s_{ij}^{(0)}, \qquad \tilde{\mathbf{u}} = \{\tilde{u}_{ij}^{h}\}_{i,j=1}^{N} \in \mathbb{R}^{N^2}.$$

Since $\{s_{ij}^{(q)}\}_{i,j=1}^{N+q}$ is the basis of $\mathcal{S}_h^q(\Omega)$, it suffices to calculate the various coefficients of the discrete model functions u_h^k and the discrete characteristic functions χ_h^k for $k = 1, ..., l$. Note further that due to this representation, a two dimensional discrete image can be identified with a matrix.

To begin with, u_h^k satisfies

$$\int_\Omega (\alpha K_\eta \chi_h^k + \delta)\nabla u_h \cdot \nabla v_h + (\beta K_\eta \chi_h^k + \delta) u_h v_h \; dx = \int_\Omega (\beta K_\eta \chi_h^k + \delta)\tilde{u}_h^\eta v_h \; dx,$$

by using the spline representations of u_h^k and substituting $v_h = s_{\nu\mu}^{(1)} \in \mathcal{S}_h^1(\Omega)$, the following holds

$$\sum_{i,j,\nu,\mu=1}^{N+1} u_{ij}^{k,h} \int_\Omega (\alpha\chi_h^k + \delta)\nabla_{ij}^{(1)} \cdot \nabla s_{\nu\mu}^{(1)} + (\beta K_\eta \chi_h^k + \delta) s_{ij}^{(1)} s_{\nu\mu}^{(1)} \; dx = \int_\Omega (\beta\chi_h^k + \delta)\tilde{u}_h s_{\nu\mu}^{(1)} \; dx. \quad (5.40)$$

Hence, we introduce the following matrix for the two dimensional linear splines,

$$\mathcal{G}(\mathbf{q_k}) = \left\{ \frac{1}{h} \int_\Omega \chi_h^k s_{ij}^{(1)} s_{\nu\mu}^{(1)} \right\}_{i,j,\nu,\mu=1}^{N+1} = \frac{1}{36} \{\mathcal{G}_{mn}\}_{m,n=1}^{(N+1)^2}, \quad (5.41)$$

where $m, n = 1, ..., (N+1)^2$ are calculated by $m = (j-1)(N+1) + i$ and $n = (\mu-1)(N+1) + \nu$. Thus, denote the non-zero coefficients of $\mathcal{G}$

$$\mathcal{G}_{m,m-N-2} = \chi_{i-1,j-1}^{k,h},$$

$$\mathcal{G}_{m,m-N-1} = 2\left(\chi_{i-1,j-1}^{k,h} + \chi_{i,j-1}^{k,h}\right),$$

$$\mathcal{G}_{m,m-N} = \chi_{i,j-1}^{k,h},$$

$$\mathcal{G}_{m,m-1} = 2\left(\chi_{i-1,j-1}^{k,h} + \chi_{i-1,j}^{k,h}\right),$$

$$\mathcal{G}_{m,m} = 4\left(\chi_{i-1,j-1}^{k,h} + \chi_{i-1,j}^{k,h} + \chi_{i,j-1}^{k,h} + \chi_{i,j}^{k,h}\right),$$

$$\mathcal{G}_{m,m+1} = 2\left(\chi_{i,j-1}^{k,h} + \chi_{i,j}^{k,h}\right),$$

$$\mathcal{G}_{m,m+N} = \chi_{i-1,j}^{k,h},$$

$$\mathcal{G}_{m,m+N+1} = 2\left(\chi_{i-1,j}^{k,h} + \chi_{i,j}^{k,h}\right),$$

$$\mathcal{G}_{m,m+N+2} = \chi_{i,j}^{k,h}.$$

Also the stiffness matrix reads as follows

$$\mathcal{A}(\mathbf{q_k}) = \left\{ \frac{1}{h} \int_\Omega \chi_h^k \nabla s_{ij}^{(1)} \cdot \nabla s_{\nu\mu}^{(1)} \right\}_{i,j,\nu,\mu=1}^{N+1} = \frac{1}{6h^2} \{\mathcal{A}_{mn}\}_{m,n=1}^{(N+1)^2}, \quad (5.42)$$

and thus only the following entries are non-trivial,

$$\mathcal{A}_{m,m-N-2} = -2\chi^{k,h}_{i-1,j-1}$$

$$\mathcal{A}_{m,m-N-1} = -\left(\chi^{k,h}_{i-1,j-1} + \chi^{k,h}_{i,j-1}\right)$$

$$\mathcal{A}_{m,m-N} = -2\chi^{k,h}_{i,j-1}$$

$$\mathcal{A}_{m,m-1} = -\left(\chi^{k,h}_{i-1,j-1} + \chi^{k,h}_{i-1,j}\right)$$

$$\mathcal{A}_{m,m} = 4\left(\chi^{k,h}_{i-1,j-1} + \chi^{k,h}_{i-1,j} + \chi^{k,h}_{i,j-1} + \chi^{k,h}_{i,j}\right)$$

$$\mathcal{A}_{m,m+1} = -\left(\chi^{k,h}_{i,j-1} + \chi^{k,h}_{i,j}\right)$$

$$\mathcal{A}_{m,m+N} = -2\chi^{k,h}_{i-1,j}$$

$$\mathcal{A}_{m,m+N+1} = -\left(\chi^{k,h}_{i-1,j} + \chi^{k,h}_{i,j}\right)$$

$$\mathcal{A}_{m,m+N+2} = -2\chi^{k,h}_{i,j}$$

For both matrices if the indices i, j of $\chi^{k,h}_{ij}$ are not in the range $1, ..., N$, we take $\chi^{k,h}_{ij} = 0$. Note further that $\mathcal{G}(\mathbf{1})$ is the Gram matrix for linear splines and $\mathcal{A}(\mathbf{1})$ is the finite difference approximation to the Laplacian with natural boundary conditions. Moreover, we define a projection matrix P from the linear spline coefficients to piecewise constants

$$\left\{\frac{1}{h}\int_\Omega \chi^k_h s^{(1)}_{ij}\right\}_{i,j=1}^{N+1} = \mathbf{P}^\top \mathbf{q_k}, \tag{5.43}$$

where

$$\mathbf{P} = \begin{bmatrix} P_1 & P_1 & & \\ & \ddots & \ddots & \\ & & P_1 & P_1 \end{bmatrix} \in \mathbb{R}^{N^2\times(N+1)^2}, \quad P_1 = \begin{bmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix} \in \mathbb{R}^{N\times(N+1)}.$$

In addition, the diagonal matrix is defined in terms of the lexicographic ordering of $\mathbf{q_k}$,

$$\mathcal{D}(\mathbf{q_k}) = \mathrm{diag}\{\chi^{k,h}_{ij}\}^N_{i,j=1}. \tag{5.44}$$

The combination of the matrices (5.41), (5.42), (5.43) and (5.44) leads back to the discretized optimality system for u^k_h (5.40) and thus, they give the system

$$\mathbf{K}(\mathbf{q_k})\mathbf{u_k} = \left[\mathcal{G}(\beta\mathbf{q_k} + \delta\mathbf{1}) + \mathcal{A}(\alpha\mathbf{q_k} + \varepsilon\mathbf{1})\right]\mathbf{u_k} = \mathbf{P}^\top\mathcal{D}\left(\beta\mathbf{q_k} + \delta\mathbf{1}\right)\tilde{\mathbf{u}} =: \mathbf{L}(\mathbf{q_k}). \tag{5.45}$$

In order to compute the various model functions u^k_h for $k = 1, ..., l$, the pixels are rearranged into a $(N+1)^2 \times 1$-vector, in the lexicographic ordering, i.e., complete increments in the index i along the x-axis are carried out after every step in the index j along the y-axis. Now that we have established a way of calculating the model functions u_h^k, we deploy a discretized version of the semi-implicit gradient descent step. Therefore, recall the mapping

$$G(\chi) = \chi - \omega\nabla\mathcal{J}(\chi).$$

First of all, we deduce a discretized version of the gradient of $\mathcal{J}$. The cost functional $\mathcal{J}$ itself reads as follows

$$\begin{aligned} h^{-1}\mathcal{J}(\mathbf{q_k}) &= \frac{1}{2}\sum_{k=1}^{l} \mathbf{u_k}(\mathbf{q_k})^\top \mathcal{G}(\beta\mathbf{q_k}+\delta\mathbf{1})\mathbf{u_k}(\mathbf{q_k}) - 2\mathbf{u_k}(\mathbf{q_k})^\top \mathbf{P}^\top \mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}} \\ &\quad + \tilde{\mathbf{u}}^\top \mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}} + \mathbf{u_k}(\mathbf{q_k})^\top \mathcal{A}(\alpha\mathbf{q_k}+\varepsilon\mathbf{1})\mathbf{u_k}(\mathbf{q_k}) \\ &= \frac{1}{2}\sum_{k=1}^{l} \mathbf{u_k}(\mathbf{q_k})^\top \mathbf{K}(\mathbf{q_k})\mathbf{u_k}(\mathbf{q_k}) - 2\mathbf{u_k}(\mathbf{q_k})^\top \mathbf{L}(\mathbf{q_k}) + \tilde{\mathbf{u}}^\top \mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}}. \end{aligned}$$

Since $\mathbf{K}(\mathbf{q_k})\mathbf{u_k}(\mathbf{q}_k) - \mathbf{L}(\mathbf{q_k}) = 0$ and thus

$$h^{-1}\mathcal{J}(\mathbf{q_k}) = \frac{1}{2}\sum_{k=1}^{l} \tilde{\mathbf{u}}^\top \mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}} - \mathbf{u_k}(\mathbf{q_k})^\top \mathbf{L}(\mathbf{q_k}). \tag{5.46}$$

The discrete gradient of $\mathcal{J}$ with respect to χ_k is determined by adding $-\mathbf{u_k}^\top(\mathbf{q_k})\mathbf{L}(\mathbf{q_k}) + \mathbf{u_k}(\mathbf{q_k})^\top \mathbf{K}(\mathbf{q_k})\mathbf{u_k}(\mathbf{q_k})$, so that we can insert the operator $\mathbf{K}(\mathbf{q_k})$,

$$\begin{aligned} D_{\mathbf{q_k}}\mathcal{J}(\mathbf{q_k}) = \frac{1}{2h}\sum_{k=1}^{l} \Big[&\tilde{\mathbf{u}}^\top D_{\mathbf{q_k}}\mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}} - 2\mathbf{u_k}(\mathbf{q}_k)^\top D_{\mathbf{q_k}}\mathbf{L}(\mathbf{q_k}) \\ &+ \mathbf{u_k}(\mathbf{q_k})^\top D_{\mathbf{q_k}}\mathbf{K}(\mathbf{q_k})\mathbf{u_k}(\mathbf{q_k})\Big]. \end{aligned} \tag{5.47}$$

Denote with $e_i \in \mathbb{R}^N$ the i^{th} unit vector, such that

$$\begin{aligned} \tilde{\mathbf{u}}^\top D_{\mathbf{q_k}}\mathcal{D}(\beta\mathbf{q_k}+\delta\mathbf{1})\tilde{\mathbf{u}} &= \beta\left\{\tilde{\mathbf{u}}^\top \mathcal{D}(e_i)\tilde{\mathbf{u}}\right\}_{i=1}^{N} = \beta\tilde{\mathbf{u}}^\top\tilde{\mathbf{u}}, \\ -2\mathbf{u_k}(\mathbf{q_k})^\top D_{\mathbf{q_k}}\mathbf{L}(\mathbf{q_k}) &= -2\beta\left\{\mathbf{u_k}(\mathbf{q_k})^\top \mathbf{P}^\top \mathcal{D}(e_i)\tilde{\mathbf{u}}\right\}_{i=1}^{N} = -2\beta\mathbf{u_k}(\mathbf{q_k})^\top \mathbf{P}^\top \mathcal{D}(\tilde{\mathbf{u}}), \\ \mathbf{u_k}(\mathbf{q_k})^\top D_{\mathbf{q_k}}\mathbf{K}(\mathbf{q_k})\mathbf{u_k}(\mathbf{q_k}) &= \left\{\mathbf{u_k}(\mathbf{q_k})^\top[\alpha\mathcal{A}(e_i) + \beta\mathcal{G}(e_i)]\mathbf{u_k}(\mathbf{q_k})\right\}_{i=1}^{N}. \end{aligned}$$

Thus,

$$G(\mathbf{q_k}) = \mathbf{q_k} - \omega D_{\mathbf{q_k}}\mathcal{J}(\mathbf{q_k}). \tag{5.48}$$

Note that the mapping T will not be discussed in this section since its representation can be easily implemented. Finally, we build the disrete version of the operator $F(\chi_k)$. Therefore, denote that the discrete version of $J_\tau(\chi_k) = \int_\Omega \sqrt{|\nabla\chi_k|^2 + \tau}\; dx$ reads as follows

$$h\mathbf{D}(\mathbf{q_k}) := \sum_{j=1}^{N}\sum_{i=2}^{N} |\chi_{i,j}^{k,h} - \chi_{i-1,j}^{k,h}|_\tau + \sum_{i=1}^{N}\sum_{j=2}^{N} |\chi_{i,j}^{k,h} - \chi_{i,j-1}^{k,h}|_\tau, \tag{5.49}$$

where $|x|_\tau = \sqrt{x^2 + \tau^2}$.

$$
\begin{aligned}
&D_{\mathbf{q_k}} \mathbf{D}(\mathbf{q_k}) = \mathbf{H}(\mathbf{q_k})\mathbf{q_k}, \quad \mathbf{H}(\mathbf{q_k}) = D_x^\top \mathcal{D}(\delta_x) D_x + D_y^\top \mathcal{D}(\delta_y) D_y, \\
&\delta_x = \left\{ \frac{h}{|\chi_{i,j}^{k,h} - \chi_{i-1,j}^{k,h}|_\tau} \right\}_{i=2,j=1}^{N}, \quad \delta_y = \left\{ \frac{h}{|\chi_{i,j}^{k,h} - q_{i,j-1}^{k,h}|_\tau} \right\}_{i=1,j=2}^{N}, \\
&D_x = \begin{bmatrix} \tilde{D}_x & & \\ & \ddots & \\ & & \tilde{D}_x \end{bmatrix} \in \mathbb{R}^{N(N-1)\times N^2}, \quad D_x = \frac{1}{h} \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(N-1)\times N} \\
&D_y = \frac{1}{h} \begin{bmatrix} -\tilde{D}_y & \tilde{D}_y & & \\ & \ddots & \ddots & \\ & & -\tilde{D}_y & \tilde{D}_y \end{bmatrix} \in \mathbb{R}^{N(N-1)\times N^2}, \quad \tilde{D}_y = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in \mathbb{R}^{N\times N}.
\end{aligned}
\tag{5.50}
$$

So the operator $F(\chi_k)$ has the following discrete form $\mathbf{F}(\mathbf{q_k}) := [I + \omega\gamma\mathbf{H}(\mathbf{q_k})]^{-1}$. Thus, denote the coefficients of the update of $\mathbf{q_k}$ with $\tilde{\mathbf{q}}_\mathbf{k}$, and so it is determined by the following system

$$
\tilde{\mathbf{q}}_\mathbf{k} = \mathbf{F}(\mathbf{q_k}) T \left(\mathbf{q_k} - \omega D_{\mathbf{q_k}} \mathcal{J}(\mathbf{q_k}) \right), \tag{5.51}
$$

where $I \in \mathbb{R}^{N^2}$ Now, we are able to denote the final algorithm that will be implemented. Again, choose the parameters according to Assumption (3.1). Also the number of model functions or characteristic functions have to chosen by hand. Note further, that $\mathbf{q}$ contains all $\mathbf{q_k}$ in vector form, i.e. the matrices $\mathbf{q_k}$ rearranged to vectors in lexicographic ordering, $\mathbf{q}(j,:)$ denotes the j-th row vector of the matrix $\mathbf{q}$ and $\mathbf{e^k}$ denotes the k-th unit row vector.

Algorithm 4 Discrete Semi-Implicit Gradient Descent Method to Compute Characteristic Functions

Input: $\mathbf{q_0}, \tilde{\mathbf{u}}, \alpha, \beta, \varepsilon, \delta, \gamma, \rho, n_{max}, \omega, l$
Output: $\tilde{\mathbf{q}}, \mathbf{u}$
Set $n = 1$
while $\left\| \mathbf{q_k}^{(n+1)} - \mathbf{q_k}^{(n)} \right\|_2 \geq \rho$ and $n \leq n_{max}$ **do**
  for $k = 1, ..., l$ **do**
    Calculate $\mathbf{u_k}$ satisfying

$$
\mathbf{K}(\mathbf{q_k})\mathbf{u_k} = \mathbf{L}(\mathbf{q_k}),
$$

    Compute $\mathbf{q_k}^{(n+1)} = \mathbf{F}\left(\mathbf{q_k}^{(n)}\right) T \left(\mathbf{q_k}^{(n)} - \omega D_{\mathbf{q_k}} \mathcal{J}\left(\mathbf{q_k}^{(n)}\right)\right)$.
for $j = 1, ..., N^2$ **do**
  Compute $\tilde{\mathbf{q}}(j,:) = \min \left\{ \arg\min_{k\in\{1,...,l\}} \left\| \mathbf{q}(j,:) - \mathbf{e^k} \right\|_2 \right\}$

To plot the results, particularly the model functions $\mathbf{u_k}$, and calculating the approximated raw image, the model functions $\mathbf{u_k}$ have to have the same size as the characteristic functions $\mathbf{q_k}$. Therefore, we make use of the projection matrix $\mathbf{P}$, which projects the linear spline coefficients to piecewise constant ones, i.e. we use $\mathbf{P}\mathbf{u_k}$.

Chapter 6

Numerical Results

Finally, this chapter is devoted to present the numerical results obtained with Algorithm (4). Computations were performed on a Samsung R540 PC with the operating system Microsoft Windows 7. All codes were written in Matlab R2012b (MathWorks, Natick, MA). Recall the bounded image domain $\Omega := [0,1]^2$. All images considered in this thesis are digital gray scale images of $N \times N$ arrays of pixels with $N = 256$. First of all, an artificial image is observed, then we regard a drawing and a photograph. Finally, two medical images are considered.

For the sake of notational brevity, the h indicating the discrete form of the functions is omitted, e.g. the resulting discrete characteristic function will be denoted with χ instead of χ_h. Concerning choosing the parameters from Assumption (3.1), it was stated earlier that ε and δ are safeguards for computing the model functions u_k for $k = 1, ..., l$ if the function χ_k does not behave properly. But the farther away they are chosen from 0 the more blurred the resulting χ becomes and ultimately converges to $\chi = 0$, which is not desired in this work. Concerning the parameters α and β, whereas β is mostly chosen equal to 1, the dimensionless quantity α serves as the smoothing parameter and thus, a large value of α penalizes the gradient of the model function, i.e. $|\nabla u_k|^2$ for all $k = 1, ..., l$. So for noisy images it is necessary to choose α relatively large. However, also for raw data with less noise it can be helpful to use a comparably large value of α at the beginning and then reduce it gradually [15]. In addition, γ determines the smoothness of the function χ, but also it accelerates convergence to $\chi = 0$ if it is selected too high. Unfortunately, there is another factor which encourages convergence to $\chi = 0$, namely if the optimization process takes too long, e.g. the number of iterations is not fixed, then at some point $\chi = 0$ is reached. One reason for this seems to be that the χ_k are calculated separately from χ_j for $j \neq k$ and thus, have no connection with each other during the optimization process. This can be considered the greatest disadvantage of the method. Therefore, it gives better results to terminate the optimization by hand, than wait till it has completed. Nevertheless, we will later see that the functional $\mathcal{J}$ is still being minimized in the optimization process.

Considering $\chi_0 := 0$, computations indicate that χ_0 is the global minimizer. To prove this theoretically, recall that $u(\chi)$ are the corresponding model functions of the phase functions χ and that $u(\chi)$ minimizes J in (2.5) with respect to the model functions for fixed χ. Hence, $J(u(\chi), \chi) \leq J(u, \chi)$ for arbitrary u. On the other hand, $J(u, \chi) = J_{\epsilon,\delta}(u) + J_{\alpha,\beta}(u, \chi)$, where $J_{\epsilon,\delta}(u) \geq 0$ depends only upon u, ϵ, δ but not upon χ, α, β and $J_{\alpha,\beta}(u, \chi) \geq 0$ depends only upon u, χ, α, β but not upon ϵ, δ. Hence, $J(u(\chi), \chi) \geq J_{\epsilon,\delta}(u(\chi)) = J(u(\chi), \chi_0)$. With these inequalities, it follows that for any $\chi \in \mathcal{K}$,

$$J(u(\chi), \chi) \geq J(u(\chi), \chi_0) \geq J(u(\chi_0), \chi_0) \tag{6.1}$$

and hence χ_0 is the global unconstrained minimizer in $\mathcal{K}$ of (4.37). Yet, χ_0 is not the minimizer in the constrained set $\mathrm{BV}(\Omega, \Delta_l)$ of Theorem (4.5). This fact will be seen among the computations below in which the constraint is implemented. Other computations appearing below demonstrate positive results obtained by seeking a local minimum in $\mathcal{K}$ obtained by avoiding the global minimizer χ_0.

Concerning the parameter $\omega \in (0, 1)$, it was chosen rather close to 1 since changes in χ would appear too slow otherwise. Recall the function T which guarantees the gradient is always in the constraint $[0, 1]$. Note further that the number of phases which also corresponds to the number of characteristic functions and model functions has to be chosen by hand. However, the *Four Color Theorem* assures that choosing $l = 4$ as the number of phases suffices to partition the given image properly. For further details see [20].

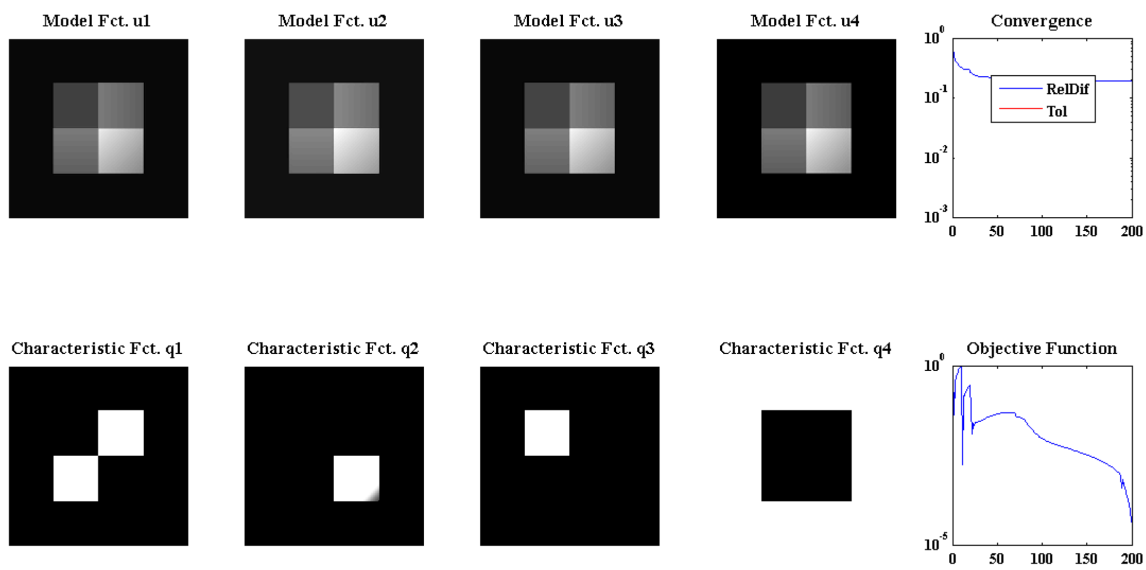

Figure 6.1: Applying the Algorithm to an artificial image. The optimization process was terminated after 300 iterations. Parameters: $\alpha = 10^{-7}$, $\gamma = 10^{-4}$, $\varepsilon = \delta = 10^{-10}$, $\beta = 1$, $\tau = 10^{-3}$, $\omega = 0.9$.

Starting off with an artificial image. The segmentation shows an intuitively good result. Figure (6.1) shows the result before Modified First-Max approach was used. The final set of characteristic functions seems rather clear here. We will later see a before-after example of the Modified First-Max. Note further that the objective functional is indeed minimized, which can be seen in the down right window of Figure (6.1), where the graph of the relative values of $\mathcal{J}$ are mapped.

Concerning a realistic image, i.e. a photograph, the optimization process in Figure (6.2) was terminated by hand after 150 iterations. All χ_k for $k = 1, 2, 3$ have values in $[0, 1]$, so the Modified First-Max needs to be applied to receive a suitable set of characteristic functions. The difference between before and after using the Modified First-Max is clearly visible, see Figure (6.3). Note further that the objective function $\mathcal{J}$ was minimized although the difference between the updates is still significantly high such that the optimization will not complete by its own. Concerning the resulting phases in Figure (6.3), intuitively one would think that parts of the blouse, teeth and the white spot at the right upper corner would be captured in one phase, whereas the dark parts described by χ_2 seem correctly combined in one phase. However,

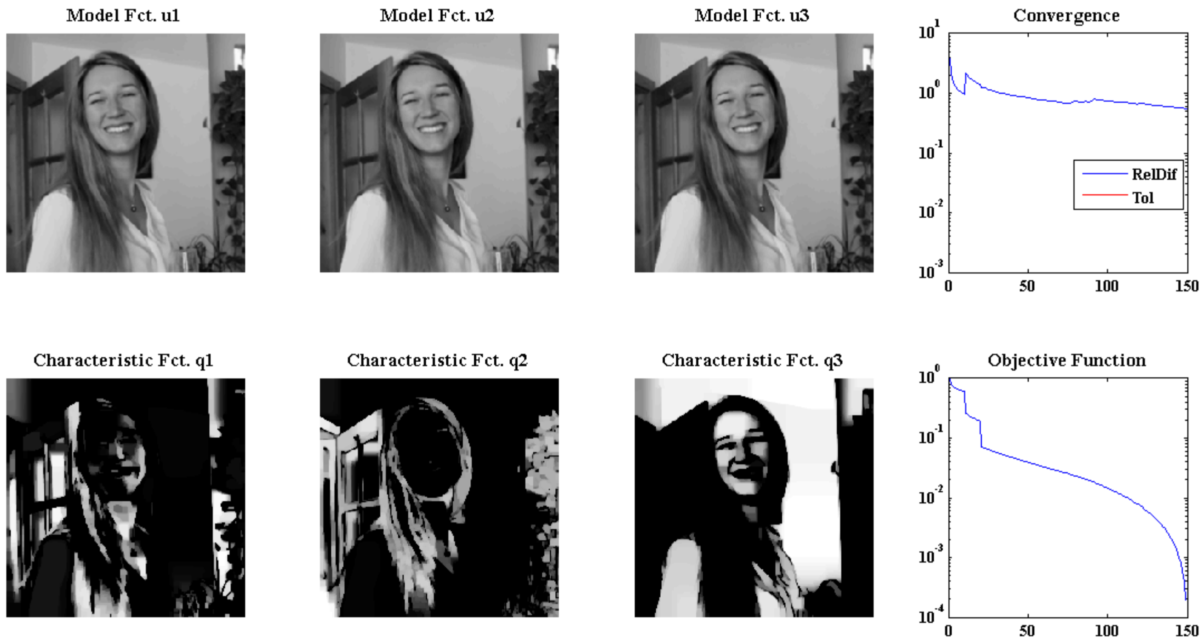

Figure 6.2: Parameters: $\alpha = 10^{-6},\ \gamma = 10^{-5},\ \varepsilon = \delta = 10^{-10},\ \beta = 1,\ \tau = 10^{-3},\ \omega = 0.9$.

Figure 6.3: Here the Modified First-Max approach was used on the results of Figure (6.2)

considering Figure (6.2) one can see that the blouse belongs to the phase function χ_3 with a lower percentage than the pixels describing the wall in the background, which is also part of the third characteristic function. An idea would be to choose four phases such that the brighter shades of gray are again divided into two phases.

Comparing the performance for this input image to other raw data it appears that the difference between the updates remains at a certain value. As mentioned earlier, sooner or later the algorithm proceeds to strive for $\chi = 0$ and thus, the algorithm terminates on its own. The results imply that terminating early in the optimization process and then applying the Modified First-Max approach to be an effective strategy to find a multiphase segmentation and thus, an appropriate piecewise smooth approximation of the given image.

Now, we briefly put our focus on the initial χ_0. As explained in previous chapters, in order to get a segmentation that is considered good, we have to start relatively close to a minimum since the segmentation mostly strives for $\chi = 0$. Considering the example below, Figure (6.4), a segmentation was performed for $\chi_1^0 = \mathbf{1}{\cdot}0.75$ and only two model functions, such that $\chi_2^0 = \mathbf{1} - \chi_1^0$. Immediately, the algorithm calculates for both $k = 1, 2$ the same characteristic function, which eventually converges to 0. Again, the reason for this seems to be the lack of connection between χ_1 and χ_2.

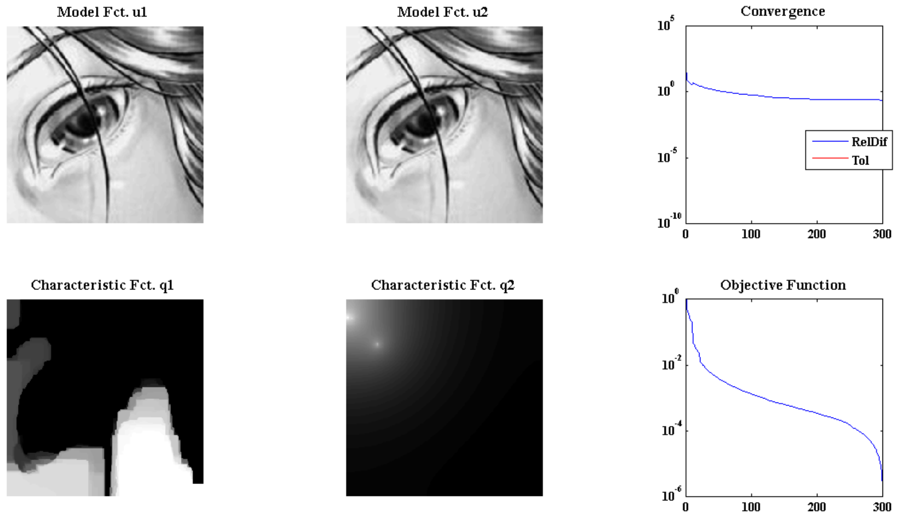

Figure 6.4: Here, the initial χ_0 was not chosen sufficiently close to a local minimum. The optimization process was terminated after 300 iterations. Parameters: $\alpha = 10^{-5}$, $\gamma = 10^{-4}$, $\varepsilon = \delta = 10^{-10}$, $\beta = 1$, $\tau = 10^{-3}$, $\omega = 0.9$.

In Figure (6.4) the segmentation immediately arrived at what looked like an edge map and then gradually blurred and darkened. However, if a certain connection between χ_1 and χ_2 is established, i.e. introduce a new dimensionless variable $\mu \in (0, 1]$ such that

$$\chi_1^{(n+1)} = F\left(\chi_1^{(n)}\right) T\left[\chi_1^{(n+1)} - \omega\left(\nabla\mathcal{J}\left(\chi_1^{(n)}\right) - \mu\nabla\mathcal{J}\left(\chi_2^{(n)}\right)\right)\right],$$

where $\nabla\mathcal{J}\left(\chi_2^{(n)}\right)$ is indeed associated with the model function u_2. However, $\chi_2 = 1 - \chi_1$ holds always, i.e. there is no separate optimization process for χ_2, see Figure (6.5).

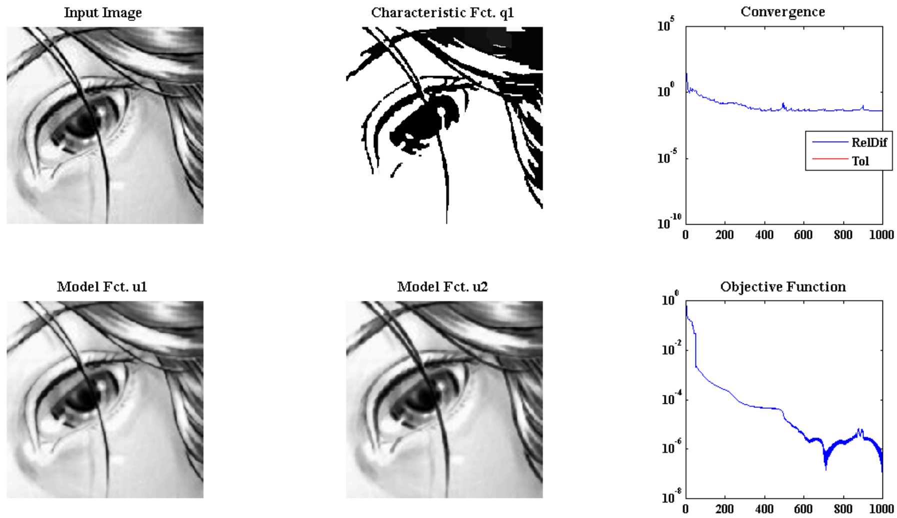

Figure 6.5: Here, χ_1 and χ_2 share the connection $\chi_1 + \chi_2 = 1$. Parameters: $\alpha = 10^{-5}$, $\gamma = 10^{-4}$, $\varepsilon = \delta = 10^{-10}$, $\beta = 1$, $\tau = 10^{-3}$, $\omega = 0.9$.

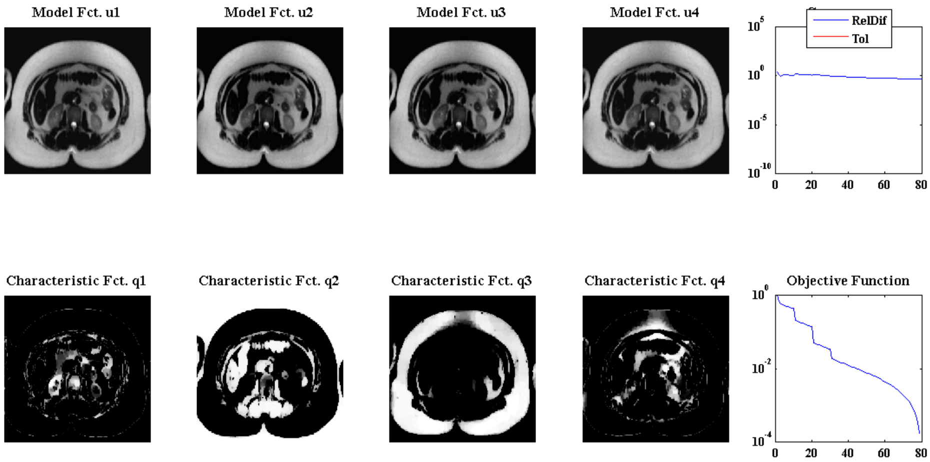

Figure 6.6: Parameters: $\alpha = 10^{-7}$, $\gamma = 10^{-5}$, $\varepsilon = \delta = 10^{-10}$, $\beta = 1$, $\tau = 10^{-3}$, $\omega = 0.9$.

Finally, we consider the application of the algorithm to some medical images, such as ones resulting from MRT. In the following images, it is important to know where the fat of the body is stored, it is divided into the fat around the muscles directly under the skin, visceral fat which lies in between the organs in the abdomen, essential and intramuscular fat. Concerning the upcoming images the first two types are of interest in this segmentation. Of course, for the segmentation to become accurate, it is important that the pixels that represent fat have roughly the same intensity values. Starting sufficiently close to a minimum, the algorithm captures the the visceral fat, the subcutaneous fat and the organs separately. However, in Figure (6.6) and (6.7) some organs do not display a significant difference in shades of gray to the surrounding fat and thus, the algorithm captures the organ together with the visceral fat. Again, the updates show no reduction in relative difference, see in Figure (6.6) top right panel. As mentioned earlier, the optimization process gives better results if it is stopped at a certain point. Note further that the functional is still relatively minimized. Here, the optimization process was terminated after 80 iterations.

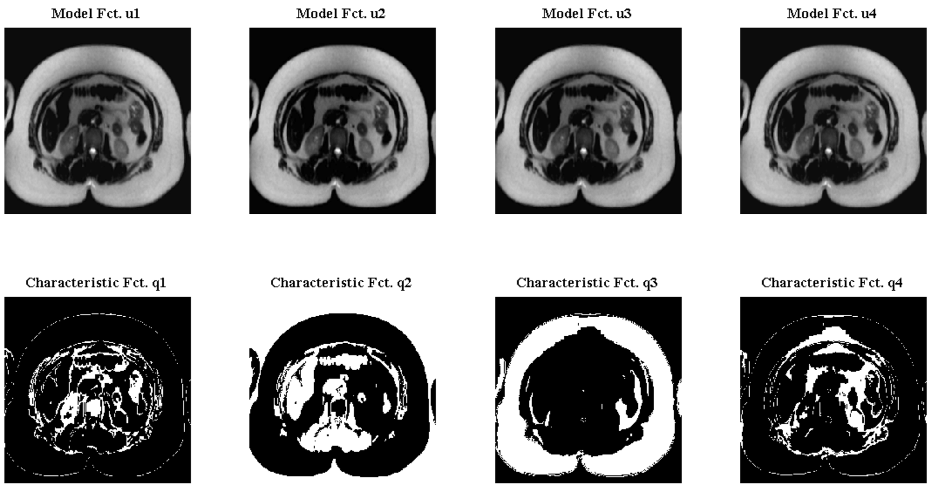

Figure 6.7: Parameters: $\alpha = 10^{-7},\ \gamma = 10^{-5},\ \varepsilon = \delta = 10^{-10},\ \beta = 1,\ \tau = 10^{-3},\ \omega = 0.9$

In the end, let us reconsider what one actually gets from the algorithm. Therefore, we briefly summarize what it does in practice. So starting sufficiently close to a local minimum with respect to phase functions, the optimization process extends the support of the current phase functions and refines it. Recall that the phase functions have range in $[0, 1]$ at this point. Then, after a fixed number of iterations, the process is stopped and the Modified First-Max strategy is applied such that the result is a set of phase functions which satisfies $\sum_{k=1}^{l} \chi_k = 1$ and each χ_k has range in $\{0, 1\}$, i.e. each one is transformed into a characteristic function. Simultaneously, it also calculates the corresponding model functions u_k for all $k = 1, ..., l$. So a piecewise smooth reconstruction of the input image is obtained.

Chapter 7

Conclusion

In the preceding chapters we proposed a hybrid model to combine image segmentation with image denoising with the goal to compute model functions in dependence of their corresponding characteristic functions and the segmentation of the raw image simultaneously. Existence and uniqueness of model functions $u_k \in H^1(\Omega)$ for $k = 1, ..., l$ that minimize the cost functional $\mathcal{J}$ were proven, and also certain features were established which provide existence of a minimum of $\mathcal{J} + \gamma J_\tau$ with respect to relaxed characteristic functions $\chi \in \mathrm{BV}(\Omega, \Delta_l)$. Then a procedure to calculate the minimizer χ^* was proposed, analysed and eventually existence of a fixed point of the iterative scheme was proposed. Finally, to obtain a suitable set of characteristic functions the heuristic rounding scheme Modified First-Max was presented. Concerning the numerical approximations the finite element method is the procedure of choice and thus, in the last chapter the results were presented.

Problems arose since the functional $\mathcal{J}$ is convex with respect to the model functions but evidently not with respect to the characteristic functions. Therefore, we could not rely on primal-dual methods, but instead used a semi-implicit gradient descent procedure. However, since our gradient is always positive, the resulting function from the gradient descent step can have values outside $[0, 1]$ and the same can be the case for the final update. Therefore, we introduced the mapping T that puts the constraint of having range in $[0, 1]$ on the update.

The results show that the procedure converges to an effective local minimum. Thus, the optimization was terminated by hand. The reason for this, as mentioned in the previous chapter, is the lack of connection between the χ_k in the updating process. Nevertheless, following this idea gives good results when starting close to a local minimum.

However, the method is far from perfect. First of all, one can investigate convexity for $\mathcal{J}$ with respect to characteristic functions. However, the conjecture is that convexity does not holds. Furthermore, creating more efficiency since evaluating the discrete gradient of $\mathcal{J}$ is computationally very expensive, is another aspect that deserves more consideration. Moreover, whereas the procedure enhances massively if the image should only be partitioned into two segments since a connection between χ_1 and χ_2 is established, this link is still missing for more than two characteristic functions concerning the optimization process. Unfortunately, $\sum_{k=1}^{l} \chi_k = 1$ is not enough to establish this link, since it only states that the last characteristic function can be calculated in terms of the others, i.e. $\chi_l = 1 - \sum_{k=1}^{l-1} \chi_k$, so one segment is always neglected. In addition, as in my bachelor's thesis, the focus here was only on gray scale images. Hence, one's research can be expanded to include color images with range $[0, 1]^3$.

Finally, although there are some aspects that haven't been considered in this thesis, it gives an idea on how to overcome optimization of a functional with respect to characteristic functions.

Appendix A

Cited Analysis Results

Theorem A.1 (Young's Inequality for Convolution [6]). *Let* $1 \leq p, q, r \leq \infty$ *such that* $1/p + 1/q = 1 + 1/r$ *and let* $f \in L^p(\Omega)$ *and* $g \in L^q(\Omega)$*, then*

$$\|f * g\|_{L^r(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} . \tag{A.1}$$

Theorem A.2 (Green's Formula [7]). *Let* $\Omega \subset \mathbb{R}^n$ *be an open, bounded, non-empty set and* $u, v \in H^2(\Omega)$*. This implies*

$$\int_\Omega v\Delta u + \nabla u \cdot \nabla v \; dx = \int_{\partial\Omega} v \frac{\partial u}{\partial n} \; ds_x$$

Lemma A.3 (Fundamental Lemma of Calculus of Variation [5]). *Let* $\Omega \subset \mathbb{R}^n$ *be an open, bounded, non-empty set,* $u \in L^2(\Omega)$ *and let the following hold*

$$\int_\Omega u\varphi \; dx = 0 \quad \forall \varphi \in C_0^\infty(\Omega).$$

Then $u(x) = 0$ *holds for almost every* $x \in \Omega$*.*

Theorem A.4 (Theorem of Lax-Milgram [5]). *Let* X *be a Hilbert space over* $\mathbb{K}$*, let* $a : X \times X \to \mathbb{K}$ *be sesquilinear and there exist constants* c_1 *and* c_2 *with* $0 < c_1 \leq c_2 < \infty$*, such that for all* $x, y \in X$ *it follows*

- $|a(x, y)| \leq c_2 \|x\|_X \|y\|_X$
- $\operatorname{Re} a(x, x) \geq c_1 \|x\|_X^2$.

Then there exists a unique mapping $A : X \to X$ *with*

$$a(x, y) = (y, Ax)_X \quad \textit{for all } x, y \in X$$

Furthermore, it follows that $A \in \mathcal{L}(X)$ *is an invertible operator with*

$$\|A\| \leq c_2 \quad \text{and} \quad \left\|A^{-1}\right\| \leq \frac{1}{c_1}$$

Lemma A.5 (Lemma of Lax-Milgram [7]). *Let* H *be a Hilbert space over* $\mathbb{R}$*, let* $B : H \times H \to \mathbb{R}$ *be a bilinear mapping and there exist constants* $\alpha, \beta > 0$*, such that for all* $u, v \in H$

$$|B(u, v)| \leq \alpha \|u\| \|v\| \tag{A.2}$$

and

$$\beta \left\|u\right\|^2 \leq B(u,v). \tag{A.3}$$

Finally, let $f : H \to \mathbb{R}$ be a bounded linear functional on H. Then there exists a unique element $u \in H$ such that

$$B(u,v) = f(v), \quad \forall v \in H. \tag{A.4}$$

In addition, the following estimate holds

$$\left\|u\right\| \leq \frac{1}{\beta} \left\|f\right\|. \tag{A.5}$$

Theorem A.6 (Riesz Representation Theorem [7])**.** *Let X be a Hilbert space over $\mathbb{K}$. A linear functional x' of X belongs to X' if and only if there exists a unique $x \in X$ such that for every $y \in X$ follows*

$$x'(y) = (y,x)_X,$$

and so

$$\left\|x'\right\|_{X'} = \left\|x\right\|_X.$$

Theorem A.7 (Young's Inequality [5])**.** *Let $\alpha, \beta \in \mathbb{R}$. If $\alpha, \beta \geq 0$, then for any $\varepsilon > 0$ follows $2\alpha\beta \leq \varepsilon\alpha^2 + \beta^2/\varepsilon$.*

Lemma A.8. *See [4], let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz-Domain and $1 \leq q \leq n/(n-1)$ with $n/(n-1) = \infty$ for $n = 1$. Then it holds:*

1. *There exists a continuous embedding $\mathrm{BV}(\Omega) \hookrightarrow L^q(\Omega)$. If $q < n/(n-1)$ holds, it is compact.*
2. *There exists a constant $C > 0$, such that for all $u \in \mathrm{BV}(\Omega)$ the following Poincaré-Wirtinger inequality is true:*

$$\left\|P_1 u\right\|_q = \left\|u - \frac{1}{|\Omega|}\int_\Omega u \, dx\right\|_q \leq C\mathrm{TV}(u).$$

Theorem A.9 (Schauder's Fixed Point Theorem [16])**.** *Let M be a convex and closed subset of a Banachspace, let f be a continuous mapping of M into itself, and suppose that $f(M)$ is compact in M. Then f has at least one fixed point.*

Theorem A.10 (General Sobolev Imbedding Theorem [7])**.** *Let $u \in W^{k,p}(\Omega)$.*

1. *If $k < \frac{n}{p}$ and $\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$, then $u \in L^q(\Omega)$. Moreover, there exists a constant $C > 0$, which depends on Ω, k, p and n, such that*

$$\left\|u\right\|_{L^q(\Omega)} \leq C \left\|u\right\|_{W^{k,p}(\Omega)}.$$

2. *If $k > \frac{n}{p}$ and*

$$\gamma = \begin{cases} \left[\frac{n}{p}\right] + 1 - \frac{n}{p}, & \text{if } \frac{n}{p} \text{ is not an integer,} \\ \text{any positive number} < 1, & \text{otherwise,} \end{cases}$$

then $u \in C^{k-\left[\frac{n}{p}\right]-1,\gamma}(\bar{\Omega})$. In addition, there exists a constant $C > 0$, which only depends on Ω, k, p, γ and n, such that

$$\left\|u\right\|_{C^{k-\left[\frac{n}{p}\right]-1,\gamma}(\bar{\Omega})} \leq C \left\|u\right\|_{W^{k,p}(\Omega)}.$$

Proposition A.11. *See [8], the total variation has the following properties:*

1. TV *is convex, i.e.*

$$\mathrm{TV}(\lambda u + \lambda v) \leq \lambda \mathrm{TV}(u) + (1-\lambda)\mathrm{TV}(v), \quad \forall u, v \in L^1(\Omega),\ \forall \lambda \in [0,1].$$

2. TV *is positively homogeneous:*

$$\mathrm{TV}(\alpha u) = \alpha \mathrm{TV}(u), \quad \forall u \in L^1(\Omega),\ \forall \lambda \in \mathbb{R}.$$

3. TV *is lower semi-continuous in* $\mathrm{BV}(\Omega)^l$ *with respect to the* $L^1(\Omega)^l$ *topology, i.e. for all sequences* $(u^{(k)}) \subset \mathrm{BV}(\Omega)^l$ *converging (in the* L^1*-sense) to some* $u \in \mathrm{BV}(\Omega)^l$,

$$\liminf_{k\to\infty} \mathrm{TV}(u^{(k)}) \geq \mathrm{TV}(u).$$

Theorem A.12. *See [4], the space* $\mathrm{BV}(\Omega)^l$ *with the corresponding norm*

$$\|u\|_{\mathrm{BV}} := \left(\sum_{k=1}^{l} \int_\Omega |u_k(x)|^2 \, dx \right)^{1/2} + \mathrm{TV}(u)$$

is a Banach space.

Definition 7. *See [8], the sequence* $(\chi^{(n)})$ *converges weakly* to an element* χ *in* $\mathrm{BV}(\Omega)^l$ *if and only if*

- $\chi \in \mathrm{BV}(\Omega)^l,\ \chi^{(n)} \in \mathrm{BV}(\Omega)^l\ \forall n \in \mathbb{N}$,
- $\chi^{(n)} \to \chi$ *in* $L^1(\Omega)^l$, *and*
- $(D\chi^{(n)}) \to D\chi$ *weakly* in measure, i.e.,*

$$\forall v \in C_0(\Omega) : \lim_{n\to\infty} \int_\Omega v \; dD\chi^{(n)} = \int_\Omega v \; dD\chi. \tag{A.6}$$

Proposition A.13. *See [8], the sequence* $(\chi^{(n)}) \subset \mathrm{BV}(\Omega)^l$ *weakly* converges to some* $\chi \in \mathrm{BV}(\Omega)^l$ *if and only if*

1. $\chi^{(n)} \to \chi$ *in* $L^1(\Omega)^l$ *and*
2. *the sequence* $(\chi^{(n)})$ *is uniformly bounded in* $\mathrm{BV}(\Omega)^l$, *i.e.,*

$$\exists C < \infty\ \forall n \in \mathbb{N} : \left\| \chi^{(n)} \right\|_{BV} \leq C.$$

Theorem A.14. *See [8], let* $(\chi^{(n)}) \subset \mathrm{BV}(\Omega)^l$ *be uniformly bounded in* $\mathrm{BV}(\Omega)^l$. *Then* $(\chi^{(n)})$ *contains a subsequence weakly* converging to* $u \in \mathrm{BV}(\Omega)^l$.

Theorem A.15 (Minkowski's inequality [2])**.** *Let* $1 \leq p < \infty$ *and* $u, v \in L^p(\Omega)$. *Then,*

$$\|u+v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}. \tag{A.7}$$

Theorem A.16 (Hölder's Inequality [6])**.** *Let* $1 \leq p, q \leq \infty$ *sucht that* $1/p + 1/q = 1$. *Then for any* $f \in L^p(\Omega)$ *and* $g \in L^q(\Omega)$,

$$\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}. \tag{A.8}$$

Theorem A.17 (Cauchy-Schwarz Inequality [6])**.** *Let H be a Hilbert space and $f, g \in H$. Then*

$$|(f,g)_H| \leq \|f\|_H \, \|g\|_H \tag{A.9}$$

Theorem A.18 (Fubini's Theorem [2])**.** *Let $f : \mathbb{R}^n \to \mathbb{K}$ be a measurable function. Then,*

$$\int_{\mathbb{R}^n} f(x) \ dx = \int_{\mathbb{R}} \ ... \int_{\mathbb{R}} f(x_1, x_2, ..., x_n) \ dx_1 \ dx_2 \ ... \ dx_n. \tag{A.10}$$

Definition 8 (Directional Derivative [14])**.** *Let $F : X \to Y$ and $u, \delta u \in X$, then*

$$\frac{\partial F}{\partial u}(u; \delta u) := \lim_{t \to 0} \frac{F(u + t\delta u) - F(u)}{t} \tag{A.11}$$

is called Directional Derivative of F in direction δu. Moreover, if $\frac{\partial F}{\partial u}(u; \delta u)$ exists for all $\delta u \in X$ and is bounded linear operator from X to Y. Then F is Gâteaux differentiable with its Gâteaux derivative $\frac{\partial F}{\partial u}$.

Definition 9 (Gâteaux Differentiability [4])**.** *Let X, Y be two normed spaces and $U \subset X$ non-empty. A mapping $F : U \to Y$ is called Gâteaux differentiable in $x \in U$ if there exists $DF(x) \in \mathcal{L}(X, Y)$ such that for all $y \in X$ the mapping $F_{x,y} : \lambda \mapsto F(x + \lambda y)$ defined on an open neighbourhood around 0 is differentiable at $\lambda = 0$ and $DF_{x,y} = DF(x)y$. Analogously F is Gâteaux differentiable if it is for all $x \in U$.*

Theorem A.19. *See [4], let $F : U \to \mathbb{R}$ be a functional defined on an open subset U of a convex set K in a real normed space X and let it be Gâteaux differentiable. Then F is convex in K if and only if*

$$F(u) + (DF(u), v - u) \leq F(v) \quad \forall u, v \in K. \tag{A.12}$$

In addition, if u lies in the interior of K, then $w = DF(u)$ is the unique element in X^* such that

$$F(u) + (w, v - u) \leq F(v) \quad \forall v \in K. \tag{A.13}$$

Bibliography

[1] Tony F. Chan, Jianhong Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, Siam, 2005.

[2] Dirk Werner, *Funktionalanalysis*, Springer-Lehrbuch, Springer-Verlag, Berlin Heidelberg, 2011.

[3] Stefan Fürtinger, *An Approach to Computing Binary Edge Maps of the Purpose of Registering Intensity Modulated Images*, Dissertation, Graz 2012.

[4] Kristian Bredies, Dirk Lorenz, *Mathematische Bildverarbeitung*, Springer Fachmedien Wiesbaden GmbH, 2011.

[5] Hans W. Alt, *Lineare Funktionalanalysis*, Springer-Verlag, Berlin Heidelberg, 2011.

[6] Robert A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[7] Lawrence C. Evans, *Partial Differential Equations*, Volume 19 of *Graduate Studies in Mathematics*, AMS, 2008.

[8] Jan Lellmann, *Nonsmooth Convex Variational Approaches to Image Analysis*, Inaugural-Dissertation, 2011.

[9] G. Aubert, P. Kornprobst. *Mathematical Problems In Image Processing*, Volume 147 of *Applied Mathematical Sciences.* Springer, 2nd Edition, 2006.

[10] M. Hintermüller, W. Ring, *An Inexact Newton-CG-Type Active Contour Approach for the Minimization of the Mumford-Shah Functional*, Journal of Mathematical Imaging and Vision 20: 19-42, 2004.

[11] Renier Mendoza, *A Multi-Phase Segmentation Approach to the Electrical Impedance Tomography Problem*, Dissertation, Graz 2013.

[12] R. Acar, C. R. Vogel, *Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems*, Inverse Problems 10 (1994) 1217-1229, United Kingdom.

[13] D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, Reprint of the 1998 Edition.

[14] M. Gerdts, *Optimal Control of ODEs and DAEs*, De Gruyter, 2012.

[15] L. Knittelfelder, *A Multiphase Segmentation Approach with Higher Order Regularization*, Bachelor's Thesis, Graz 2013.

[16] J. Schauder, *Der Fixpunktsatz in Funktionalräumen*, Studia Mathematica, (2): 171-180, 1930.

[17] L. Schumaker, *Spline Functions: Basic Theory*, Cambridge University Press, 2007.

[18] S. Brenner, R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, 2008.

[19] P. Ciarlet, *The Finite Element Method or Elliptic Problems*, SIAM Classics in Applied Mathematics, 2002.

[20] R. Fritsch, G. Fritsch, *The Four-Color Theorem: History, Topological Foundations, and Idea of Proof*, Springer, New York, 1998.