Split Bregman method for higher order functionals

Christina Auer

Research Statement for the Marshall Plan scholarship from the field
'Inverse Problems'
University of California, Los Angeles, WS 2011/12.
Submitted on June 6, 2012

Name:	Christina Auer
Matr.Nr.:	0556540
StKz.:	403

Contents

Chapter 1

Introduction

Many models of the real world may be formulated as L^1 regularization problems. Plasticity, image processing and the application in material sciences are a few examples. Especially in the field of image processing the class of L^1 regularization problems has received much attention recently. Nevertheless, many L^1- regularization problems remained difficult to solve and some of them also require techniques that are very problem-specific.

In 2009 Tom Goldstein and Prof. Stanley Osher invented at UCLA a new technique based on the Bregman iteration for solving a very broad class of L^1- regularized problems. Originally, it was invented to solve problems in image processing. In the first article [8] it is applied, amongst others, to the first order ROF-model from Rudin, Osher and Fatemi. The application to higher order models have beed studied very little so far, although they quite often appear in continuums mechanic, especially in plasticity.

1.1 Applications of image processing

In our present day life, images are omnipresent, no matter if we consider mechanical engineering, remote sensing or everyday life. One the one hand they are an easy, compact and widespread way, to represent the physical world, on the other hand they are also a powerful and widely used medium of communication. According to Aubert Gilles and Pierre Kornprobst [7], '*it it very surprising indeed (maybe frightening) to realize just how much images are omnipresent in our lives, and how much we rely upon them*'. Digital television, luggage surveillance at the airport, or medical images such as radiograph and ultrasonography scan are example of images that constantly surrounded us.

Image processing covers any digital technique for signal processing where

the in- and output involves a matrix of gray values representing an image. Since the early sixties image processing gain more and more importance. One of the reasons is the development of computers powerful enough to cope with the large dimensionality of image data and the algorithm of high complexity to operate on them. Especially in astronomy and medicine, image processing has a great importance, nevertheless it has its applications in nearly ever science.

In medicine, image processing is extensively used form the earliest day. Many devices used in medicine are based on ultrasounds, X-rays, scanners, etc. The goal is to create images of the human body which can be read and interpreted more easily by the doctors or are used in medical sciences, as anatomy and physiology.

Another important application is remote sensing, as mentioned before. It is about the analysis, measurement or interpretation of scenes at at distance. Besides the defense and video surveillance and road traffic analysis, the observation of the earth's resources are important parts of this domain. With image processing one can keep an eye on and measure changes in forests, water supplies, urbanization, pollution, etc. It also provides tools for weather forecasting to analyze great amounts of data.

1.2 Overview of the Work

This work deals with the Split Bregman method and its application on different partial differential equations. The method was first applied by Dr. Tom Goldstein and Prof. Stanley Osher at a first order partial differential equation in image processing. The main focus of this work is in solving different second order equations with its origin in physics.

In Chapter 2, we introduce the general theory for image processing. We exemplify the rather new Split-Bregman method for solving one of the most important functionals in image processing, the ROF-functional. Therefore we will discuss the Bregman iteration and the augmented Lagrangian method.

Chapter 3 concerns the motivation for this work. There some insight into the deformation of a thin film and into plasticity, will be given. The models needed for the main chapter in this work represented by the fourth chapter are derived in this chapter.

Chapter 4 deals with the implementation of the Split-Bregman method on the higher order models discussed in Chapter 3, where also a discretization of the differential operators via finite differences is given.

Chapter 2

Theory

2.1 The ROF-model

This first chapter is mainly based on [7, 11, 15, 10, 5].

In the last couple of years there has been a lot of improvement in the field of image processing. These days, noise or blur in images is unavoidable. No matter if it occurs at the image formation process, at image recording or at image transmission. There are different types of noise such as white noise, salt and pepper noise, Poisson noise or speckle noise [7]. The last one appears in radar images and is described by the gamma distribution. The most common one and the one that we will consider is noise with a Gaussian distribution. In many cases regularization can be used for image restoration tasks such as denoising and deblurring.

Let $\Omega \subset \mathbb{R}^n$ be a bounded open subset. As we are speaking of pictures it is most of the time a rectangle in $\mathbb{R}^2$. Suppose that the intensity function $f : \Omega \to \mathbb{R}$ describes the noisy image (the data), meaning that in each point $x, y \in \Omega$, $f(x, y)$ is the pixel value of the observed image. Say $u : \Omega \to \mathbb{R}$ is the original image (the unknown) to which a random Gaussian noise $n(x, y)$ has been added [11]:

$$f(x, y) = u(x, y) + n(x, y).$$

The goal to reconstruct u from our given data f is not an easy task, because we only know the variance σ and the mean of the noise. Since n is a random variable the following theory is developed using probability. The maximum likelihood principle leads to a least squares (L^2) fit. The first ones, who used this framework in the one-dimensional case were Phillips [6], Tikhonov [17] and later Twomey [13, 14]. In our two-dimensional case we can find an approximation of our image u by solving a constrained minimization

problem. The resulting problem is presented using some smoothing term $R(u)$ [15]:

$$\begin{aligned} &\min_u R(u) \\ &\text{s.t} \int_\Omega (u-f)^2 = \sigma^2. \end{aligned} \tag{2.1}$$

The energy method now follows by introducing a Lagrange multiplier μ. One can write (2.1) as an unconstrained problem by minimizing the following functional:

$$J(u) = R(u) + \frac{\mu}{2}\int_\Omega (u-f)^2 dx.$$

$R(u)$ can be chosen in various ways. For instance, Tikhonov and Arsenin [17] in 1977 used the L^2-norm of the gradient of the unknown image.

$$\begin{aligned} J_T(u) &:= \int_\Omega |\nabla u|^2 dx + \frac{\mu}{2}\int_\Omega (u-f)^2 dx = \\ &= \underbrace{\|\nabla u\|^2_{L^2}}_{\text{regularity term}} + \frac{\mu}{2}\underbrace{\|u-f\|^2_{L^2}}_{\text{fidelity term}}. \end{aligned}$$

The first term is called regularization or penalty term. The second term, the data fidelity term, measures the fidelity to the given data. By minimizing this functional we want the resulting image u to be as close as possible to the given image f but without the noise. The regularization term is responsible for removing noise. By keeping the gradient low, this term penalizes high oscillations and points of discontinuity. Therefore, the new function u will be smoother than f. The positive regularization parameter μ weights the two parts. If the parameter is big, the similarity of the noisy image f and u is more important and otherwise the main focus is to get a smooth result.

But this method has its disadvantages as it is described in [7] and [10]. In the study of this problem one can see, that the space of the solution, where both terms are well defined, is the Sobolev space H^1.

The L^p norm with $p = 2$ of the gradient allows us to remove the noise. Unfortunately the derivative of u is an element of the H^1-space and an over-smoothing of the solution follows. Moreover, lines of discontinuity which describe the edges can not exist in this space. This is a disadvantage because

the edges are the most important part for the brain to understand and perceive the picture. It follows that the previous method simply smoothes too much and the resulting picture is too blurry.

Therefore, we are searching for a norm that on the one hand is not too strong and it should allow regions of discontinuity. On the other hand it should not be too weak, because we want to get rid of the noise. Rudin, Osher and Fatemi [11] were the first who recently and successfully proposed to decrease p in order to preserve the edges as much as possible. A good compromise to balance the two aims is the space of functions of the bounded variation:

Definition 2.1 *(BV-seminorm)* Let $\Omega \subset \mathbb{R}^n$ be a given domain and let u be a function in $L^1(\Omega)$, then the *BV-seminorm* is defined as:

$$|u|_{BV} := \int_\Omega |Du| = \sup_{\substack{\varphi \in C_0^1(\Omega)^n, \\ \|\varphi\|_{L^\infty}(\Omega) \leq 1}} \int_\Omega u \mathrm{div}\varphi \, \mathrm{d}x.$$

This norm is basically the L^1 norm of the gradient but it also contains a sufficient amount of discontinuous functions. Therefore a very successful edge-preserving image restoration method is to minimize the *ROF* (Rudin Osher Fatemi) - functional:

$$J_{ROF} := |u|_{BV} + \frac{\mu}{2}\|u - f\|_{L^2}^2. \tag{2.2}$$

2.2 The Bregman method

The Bregman method is an iterative algorithm originated in functional analysis for solving problems in convex programming. It is a concept for finding extrema of convex functionals.

Recently the Bregman method became also important in image processing because its basic idea is used in a new method to solve the *ROF* model faster.

2.2.1 The Bregman distance

In 1967, L.M. Bregman introduced in [4] the concept of the Bregman distance. In order to define it for convex, not necessary differentiable functionals, we need a general notation of derivatives. In the following we introduce the definition of sub-differentiable (c.f [16, 5]):

Definition 2.2 *(sub-differentiable)* Let X^* be the dual space of a Banachspace X. A convex functional $J : X \to \mathbb{R}$ is said to be *sub-differentiable* at $u \in X$ if there exists a $p \in X^*$ such that

$$J(v) \geq J(u) + \langle p, v - u \rangle \ , \ \forall v \in X.$$

The general derivative p is then called the *sub-gradient* of J at the position u. The set of all sub-gradients at u

$$\partial J(u) := \{p \in X^* : J(v) \geq J(u) + \langle p, v - u \rangle \, , \, \forall v \in X\} \subseteq X^*$$

is called the *sub-differential* of J at the point u.

For continuously differentiable functionals the sub-gradient is unique.

As we set the foundation we can begin with the concept of the Bregman-distance, see ([8, 5, 4]):

Definition 2.3 *(Bregman distance)* Let $J : X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex functional and $\partial J(u)$ the non-empty sub-differential. Then the *Bregman distance* at a point $u \in X$ is defined as:

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v \rangle$$

where $p \in \partial J(u)$ is a subgradient.

As for continuously differentiable functionals only one sub-gradient exists, therefore also the Bregman distance is unique. In this case this distance is nothing else but the difference at a point u from $J(\cdot)$ to its first order approximation at a point v. Even for continuously differentiable functionals one can see that the Bregman distance is not a distance in the usual way since it is not symmetric and the triangle inequality does not hold. However, it is a measure for closeness in the sense, that $D_J^p(u, v) \geq 0$ and $D_J^p(u, v) = 0$ if $u = v$.

2.2.2 The Bregman iteration

With the Bregman iteration, a wide variety of convex optimization problems can be solved. For showing the general formulation of this technique we consider the following minimization problem:

$$(P) \qquad \min_u E(u) + \lambda H(u).$$

where E and H are two convex energy functionals and H is differentiable.

Then the Bregman iteration for solving this unconstrained problem (P) has the following form:

Algorithm 2.4 *(Bregman iteration)*
Let $u^0 = 0$ and $p^0 = 0$. For $k = 0, 1, 2, \ldots$ compute

$$\begin{aligned} u^{k+1} &= \min_u D_E^p(u, u^k) + \lambda H(u) = && (2.3)\\ &= \min_u E(u) - \langle p, u - u^k \rangle + \lambda H(u) && (2.4)\\ p^{k+1} &= p^k - \lambda H(u^{k+1}) && (2.5) \end{aligned}$$

with $p \in \partial E(u^k)$.

The equality from (2.3) to (2.4) can be easily observed by inserting the definition of the Bregman distance and considering that we minimize over u. The update rule of the dual variables, which are represented by the subgradients, derives from the first order optimality condition of the minimizer (2.4), which reads as following:

$$0 = (p^{k+1} - p^k) + \lambda H(u).$$

2.3 The augmented Lagrangian

The augmented Lagrangian method was first used in 1969 by Hestenes and Powell who called it the 'method of multipliers'. The algorithm has many advantages over penalty methods. The augmented Lagrangian method is used for solving constrained optimization problems mostly of the form

$$\min_{u \in X} J(u) \text{ such that } Ku = f,$$

where $J : X \to \mathbb{R}$ is a functional and $K : X \to Y$ is a linear operator between Banach spaces X and Y.

To solve this problem, we will convert it into an unconstrained problem. We want to approach the augmented Lagrangian method by first applying the classical penalty method. Therefore we add a quadratic penalty function:

$$L^\lambda(u; p) = J(u) + \frac{\lambda}{2} \|Ku - f\|_2^2.$$

But for small λ the penalty function does not accurately enforce the constraint. For that reason the conventional solution is to let λ tends to infinity. But this causes numerical instabilities. The goal of the augmented Lagrangian method is that one can obtain an exact solution although having the value of λ remain constant. Therefore we introduce a Lagrange multiplier p for the constraint. More precisely we define the Lagrangian by

$$L^\lambda(u;p) = J(u) + \langle p, f - Ku\rangle + \frac{\lambda}{2}\|Ku - f\|_2^2.$$

Adding the term $\langle p, f - Ku\rangle$ helped us to make the method numerically more stable.

As the Lagrange multiplier appears as an extra unknown we formulate a saddle point problem, for which (u, p) is a solution.

$$\min_u \max_p L^\lambda(u;p)$$

We obtain the augmented Lagrangian method, by applying the standard Uzawa algorithm:

Algorithm 2.5 *(Augmented Lagrangian Method)*
Let $u^0 = 0$ and $p^0 = 0$. For $k = 0, 1, 2, \ldots$ compute

$$\begin{aligned} u^{k+1} &= \min_u L^\lambda(u;p^k) = \\ &= \min_u J(u) + \langle p^k, f - Ku\rangle + \frac{\lambda}{2}\|Ku - f\|_2^2 \\ p^{k+1} &= p^k + \lambda(f - Ku^{k+1}). \end{aligned}$$

2.4 The Split-Bregman method

In this section, we are introducing the from the Bregman iteration and from the Augmented Lagrangian method resulting Split-Bregman method, which is a method to solve L^1-regularized optimization problems of the form:

$$\min_u |\Phi(u)|_1 + G(u).$$

The key idea of the method is to decouple the L^1 and the L^2 parts of the problem. Therefore we begin by introducing a new variable d to represent the L^1 term. This generates a clearly equivalent constrained problem:

$$\min_{u,d} |d|_1 + G(u) \text{ such that } d = \Phi(u).$$

By setting $J = |d|_1 + G(u)$ and $K = \Phi$ it is easy to see that this is a constrained optimization problem of the form (2.3) so that we can apply the augmented Lagrangian method.

As before we will convert this problem this into an unconstrained problem by adding the quadratic penalty function:

$$\min_{u,d} |d|_1 + G(u) + \frac{\lambda_k}{2}\|d - \Phi(u)\|_2^2. \tag{2.6}$$

And again, we do not want to make λ_k tend to infinity so we apply the augmented Lagrangian method by adding a Lagrange multiplier p for the constraint:

$$\begin{aligned}(u^{k+1}, d^{k+1}) &= \min_{u,d} |d|_1 + G(u) + \langle p, d - \Phi(u)\rangle + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \\ p^{k+1} &= p^k + \lambda(d - \Phi(u)).\end{aligned}$$

By taking the derivative with respect to u we can see that:

$$\begin{aligned}&\min_{u,d} \quad \langle p^k, d - \Phi(u)\rangle + \tfrac{\lambda}{2}\|\Phi(u) - d\|_2^2 \quad \Leftrightarrow \\ &\min_{u,d} \qquad \tfrac{\lambda}{2}\|d - \Phi(u) - \tfrac{1}{\lambda}p^k\|_2^2\end{aligned}$$

And by setting $b^k = \frac{1}{\lambda}p^k$ we obtain the algorithm:

Algorithm 2.6 *(Split Bregman)*
Let $u^0 = 0$ and $b^0 = 0$. For $k = 0, 1, 2, \ldots$ compute

$$(u^{k+1}, d^{k+1}) = \min_u |d|_1 + G(u) + \frac{\lambda}{2}\|d - \Phi(u) - b^k\|_2^2 \tag{2.7}$$

$$b^{k+1} = b^k + (\Phi(u^{k+1}) - d^{k+1}). \tag{2.8}$$

At the end of this chapter we want to show some results. The Split-Bregman method has been implemented numerically in Matlab. The first image is a picture where we added some noise and the second one shows the result of our algorithm:

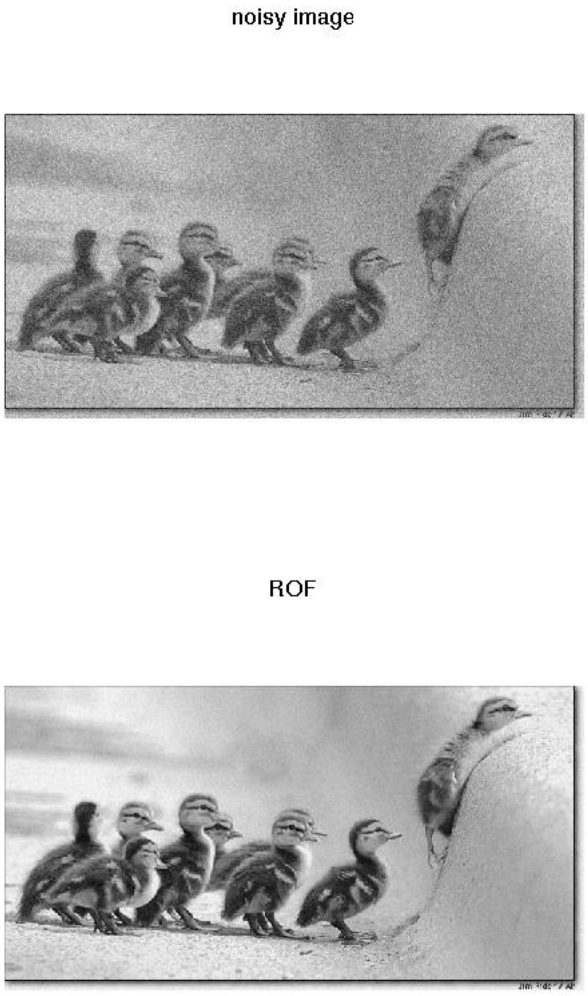

Figure 2.1: The Split Bregman method

Chapter 3

Higher order regularization models

The Split-Bregman method was first invented to solve mainly $L1$-regularized problems in image processing. The aim of this work is to apply this algorithm to higher order regularization models arising in image processing and continuums mechanics. The motivation to do so came from the article by Pavel Belik and Mitchell Luskin [2], which will be described in the following section. The motivation for my thesis and research area was the BV model in this article.

3.1 The thin film model of Belik and Luskin

This section deals with the mathematical model of the deformation of a thin film. As the motivation for my work we will introduce the Belik-Luskin paper [2]. Fist shortly an application area is mentioned as well as the definition of the BV-space for this problem. The main part will be to introduce and understand the model of the thin film.

3.1.1 The motivation

One application area where we can model a physical process by the deformation of a thin film is in crystals. More specific the microstructure of martensitic and ferromagnetic ones is modeled in this section.

Let us consider the phase transition of a crystal, which is the transformation from one region of space where all the physical parameters are homogenous to another region. During this procedure, the microstructure

of the crystal is deformed. The change from one of this piecewise constant variant phase to another piecewise constant variant phase unleashes a surface energy. This transformation is modeled as a deformation of a thin film.

We are interested in computing energy-minimizing deformations. In 1999 Bhottachara and James [3] derived a variational principle where the surface energy is modeled by

$$\kappa \int_{\Omega_h} |\nabla^2 u|^2 dx.$$

In this chapter let κ be a small positive constant, Ω_h be the reference configuration of the crystal with thickness h. The integral $\int_{\Omega_h} |\nabla^2 u|^2 dx$ denotes the square of the L^2-norm of the matrix of all the second derivatives of the deformation $u : \Omega_h \to \mathbb{R}^3$. Unless the constant κ is equal to zero we cannot have sharp interfaces between two compatible variants.

Belik and Luskin [1] avoided this drawback in 2002 by modeling the interfacial energy with the total variation instead of the L^2 norm:

$$\kappa \int_{\Omega_h} |D(\nabla u)| dx.$$

As $\int_{\Omega_h} |D(\nabla u)| dx$ denotes the total variation of the gradient of u, this term can again be interpreted as a second derivative. This special norm allows deformations of finite energy to have sharp interfaces.

With this fitting norm Belik and Luskin found a good model for the martensitic phase transform of the microstructure from a crystal. In this model the energy is given by the sum of the surface energy and the elastic energy. Since we are interested in computing energy-minimizing deformations we want to find the deformation $u : \Omega_h \to \mathbb{R}^3$ that minimizes the following functional:

$$\mathcal{E}(u) = \underbrace{\kappa \int_{\Omega_h} |D(\nabla u)|}_{\text{surface energy}} + \underbrace{\int_{\Omega_h} \Phi(\nabla u, \theta)}_{\text{elastic energy}} \tag{3.1}$$

again for a fixed $\kappa > 0$. The function $\Phi(\nabla u, \theta)$ denotes the energy density for martensitic crystals. The second argument of the density function is a fixed temperature θ.

3.1.2 The model

In this subsection we want to explain in detail the model, described in the previous subsection. It is mainly based on [1].

In this subsection we denote by $S \subset \mathbb{R}^2$ an open and connected domain with a Lipschitz continuous boundary. Let Ω_h be the reference configuration of a thin film of martensitic material with thickness $0 < h \leq 1$:

$$\Omega_h = S \times (-h/2, h/2).$$

We are interested in functions u with the domain Ω_h and the range $\mathbb{R}^3$ and whose gradient ∇u belongs to $\mathbb{R}^{3\times 3}$. Let us define the total variation again for functions of those dimensions. Furthermore, we require for our function u that $\int_{\Omega_h} |Du| < \infty$, with

$$\begin{aligned} \int_\Omega |Dw| : \; &= \; \sup\Bigg\{ \sum_{i=1}^{m} \int_\Omega w_i(x) \mathrm{div}(\psi_i)(x) dx : \\ &\qquad \psi \in \mathcal{C}_c^1(\Omega; \mathbb{R}^{m\times n}), |\psi(x)| \leq 1 \text{ for all } x \in \Omega \Bigg\}. \end{aligned} \tag{3.2}$$

We assume that the thin film adheres to a rigid material on its edges

$$\Gamma_h = \partial S \times (-h/2, h/2).$$

Let us now define the space $\mathcal{A}$ of all admissible deformations of the thin film:

$$\mathcal{A}_h = \{u \in W^{1,p}(\Omega_h; \mathbb{R}^3) : \nabla u \in BV(\Omega_h), u = u_0 \text{ on } \Gamma_h\}$$

where $u_0 \in W^{1,p}(\Omega_h; \mathbb{R}^3)$ such that $\nabla u_0 \in BV(S)$.

With this information we can understand and interpret the definition of the energy-functional better:

$$\begin{aligned} \mathcal{E}(u) \; &= \; \underbrace{\kappa \int_{\Omega_h} |D(\nabla u)|}_{\text{surface energy}} + \underbrace{\int_{\Omega_h} \Phi(\nabla u, \theta)}_{\text{elastic energy}} = \qquad (3.3) \\ &= \; \kappa |\nabla u|_{BV} + \|\phi(\nabla u, \theta)\|_{L^2}^2. \qquad (3.4) \end{aligned}$$

At this point it remains to mention a property of the energy density function Φ and the temperature θ. In the article [2] and therefore also at this point we consider the temperature as a fixed measured variable and we will not explicitly denote the dependency of the energy $\mathcal{E}(u)$ on the temperature.

For the energy density $\Phi : \mathbb{R}^{3\times 3} \times \mathbb{R} \to [0, \infty)$ which is a continuous function the following growth condition is assumed:

$$c_1(|F|^p - 1) \leq \Phi(F, \theta) \leq c_2(|F|^p + 1) \text{ for all } F \in \mathbb{R}^{3\times 3} \text{ and } \theta \in \mathbb{R}, \quad (3.5)$$

where c_1 and c_2 are fixed positive constants. To ensure that deformations with finite energy are uniformly continuous we will assume that $p > 3$.

Remark 3.1 *[1, Remark 2.1]* Note that due to the growth condition (3.5), we have

$$\mathcal{A}_h = \{u : \Omega_h \to \mathbb{R}^3 : \mathcal{E}_h(u) < +\infty, u = u_0 \text{ on } \Gamma_h.\}$$

Since we chose $p > 3$ it follows from the Sobolev embedding theorem that $\mathcal{A}_h \subset \mathcal{C}(\Omega_h)$ and therefore we can ensure that there is no tear in the deformed configurations $u(\Omega_h)$ for $u \in \mathcal{A}$.

3.2 Our models

In this section we present the higher order toy models which we use in the main part of this work, represented in the next chapter. Originally, the Split-Bregman method, introduced in the previous chapter, was invented to solve problems in image processing. It is known to be a very fast and successful method on the first order ROF-functional. The main focus of attention is to expand this technique to different mathematical areas as material sciences and continuums mechanics.

The intention for my work was to apply the Split-Bregman method to the Belik-Luskin model for martensitic crystals. The motivation was that it is a higher order functional with a similar structure. Inspired by that Belik-Luskin functional we tried to create a similar but simpler model. To avoid a detailed modeling of the elastic behavior of the thin films and in view of the similarity with the ROF-model, we modeled the second term in (3.3) and defined:

Model 3.2 *The Belik-Luskin toy model*
Let f be a given function from $\mathbb{R}^2 \to \mathbb{R}^2$ and let u denote the wanted function from $\mathbb{R}^2 \to \mathbb{R}$:

$$\min_u \|\Delta u\|_1 + \frac{\mu}{2}\|\nabla u - f\|_2^2. \quad (3.6)$$

3.2.1 The plasticity model

Further interested was aroused by the book of Roger Teman [16]. The domain of plasticity was another motivation for this work. Also in plasticity we got similar structured problems which we also tried to solve with the Split-Bregman method.

The word 'plastic' originates from the Greek and means 'to shape' [12]. Plasticity is the property of solid materials that its shape can be changed and retained upon removal of the forces. The irreversible deformation is initiated through the application of appropriate directed and intensified forces. It must be of sufficient intensity, at times the object has to be softened by the application of heat in order to be worked. For the object to attain a useful value it sometimes need to be hardened for example through exposure to air.

On the contrary an elastic material would recapture its original form. A brittle one would immediately break. Typical plastic materials are ductile metals, clay and putty and more daily used ones are toothpaste, mayonnaise and butter. The plastic deformation behavior depends amongst other things on the state of stress, the temperature, the nature of load and the rate of loading.

One of the pioneers in plasticity was Heinrich Hencky. The Hencky model is a static model from the 1920s, see [16]. It has no memory taking prior deformations into account. Although it has more in common with a model of nonlinear elasticity with threshold and does not represent the phenomena involved in plasticity very well one can think of it as one of the simplest plasticity model. Some of the research and techniques developed on this model represents an indispensable first step for the study of more complex ones. Inspired by this model we consider the following one:

Model 3.3 *The plasticity toy model*
Let f be a given function from $\mathbb{R}^2 \to \mathbb{R}$, u the wanted function from $\mathbb{R}^2 \to \mathbb{R}$ and let H denote the hessian:

$$\min_u \|H(u)\|_{L^1} - \frac{1}{\mu} \int f.u \, \mathrm{d}x \, \mathrm{d}y. \tag{3.7}$$

Chapter 4

Numerical Implementation

In the previous chapter, toy examples based on models from material sciences and plasticity have been formulated. The goal of this chapter, which is the main part of this thesis, is to give a brief overview of how one could implement the Split-Bregman method on those toy examples given in section 3.2. The guideline for this chapter was [8].

By applying the Split-Bregman method to our toy examples we will derive the second order functional $\mathcal{F}$. As the minimum is demanded the first aim in section 4.1.2 will be to find its optimality condition. In order to apply the Jacobi-method to solve this condition the reader will be presented with the discretization of the fourth order derivative $\Delta\Delta$.

4.1 Implementation of the thin film toy model

In this section, we will show how the Split-Bregman method is applied on the model (3.2) introduced in section 3.2. The toy model of a thin film for crystal microstructure is the starting point of this section. We will not only illustrate how to use the Split-Bregman framework on this problem, but also show how one could numerically discretize the equation. After applying the Split-Bregman method we have to minimize a functional with respect to two variables. The second one can be easily solved with a general shrinkage formula, which makes the algorithm very fast. At the end of this section we will apply the Jacobi method to solve the optimality condition with respect to the first variable. Let us begin by recalling the Model (3.2)

$$\min_u \|\Delta u\|_1 + \frac{\mu}{2}\|\nabla u - f\|_2^2.$$

4.1.1 Apply Split-Bregman

The first step solving the Problem (3.2) is to apply the Split-Bregman method. Following the description of this technique in Section 2.4 we should first replace Δ by d. We get the equivalent constrained problem

$$\min_u \|d\|_1 + \frac{\mu}{2}\|\nabla u - f\|_2^2 \text{ , s.t } d = \Delta u.$$

In order to obtain again an unconstraint problem we will first weakly and then strictly enforce the constraint. First weakly enforcing the constraint by adding a quadratic penalty function as it was done in (2.6), yields

$$\min_{u,d} \|d\|_1 + \frac{\mu}{2}\|\nabla u - f\|_2^2 + \frac{\lambda}{2}\|d - \Delta u\|_2^2.$$

In order to finally getting an unconstrained problem we strictly enforce the constraint by applying a Bregman iteration, therefore we get:

$$\begin{aligned}(u^{k+1}, d^{k+1}) &= \min_{u,d} \|d\|_1 + \frac{\mu}{2}\|\nabla u - f\|_2^2 + \frac{\lambda}{2}\|d - \Delta u - b^k\|_2^2 \\ b^{k+1} &= b^k + (\Delta u^{k+1} - d^{k+1}).\end{aligned}$$

The way how to chose proper values of b^k was described in the Bregman iteration 2.5. As a result of "splitting" the L^1 and the L^2 components of the functional, we are able to minimize with respect to u and d separately. In order to get a solution for the original problem we have to solve the following 3 subproblems.

$$\text{Step 1: } u^{k+1} = \min_u \frac{\mu}{2}\|\nabla u - f\|_2^2 + \frac{\lambda}{2}\|d - \Delta u - b^k\|_2^2 \tag{4.1}$$

$$\text{Step 2: } d^{k+1} = \min_d \|d\|_1 + \frac{\lambda}{2}\|d - \Delta u - b^k\|_2^2 \tag{4.2}$$

$$\text{Step 3: } b^{k+1} = b^k + (\Delta u^{k+1} - d^{k+1}). \tag{4.3}$$

Note that because u is now "decoupled" from the L^1-term, the optimization problem for u^k is differentiable. Therefore, the computation of the first step can be done by using one of the standard iterative methods such as the Gauss-Seidel or the conjugated gradient algorithm for solving a system of linear equations. It will turn out that due to the structure of the matrix, we will chose an additive decomposition for our matrix as in the article [8]. We chose to use the Jacobi-method.

One of the great advantages of this method is that the second step can be calculated extremely fast, with only a few operations per element. As there is no coupling between the elements of d we can use shrinkage operators for solving this equation.

$$d_j^{k+1} = shrink((\Delta u)_j + b_j^k, \frac{1}{\lambda})$$

where

$$shrink(x, \gamma) = \frac{x}{|x|} * max(|x| - \gamma, 0).$$

4.1.2 Numerics of the thin film toy model

In the previous subsection we applied the Split-Bregman method to our toy model of a thin film for crystal microstructures. As a result, we obtained three subproblems. The aim of this section is to describe how to solve the first equation numerically. At the starting point we want to find the optimality condition for that minimization problem. Then we want to discretize the resulting fourth order equation by using finite differences. Finally, we want to apply the Jacobi method on the discretized equation.

In order to find the optimality condition we will take the derivative of the functional we derived in the first step (4.1)

$$\mathcal{F}(u) := \frac{\mu}{2}\|\nabla u - f\|_2^2 + \frac{\lambda}{2}\|d - \Delta u - b^k\|_2^2$$

and set it equal to zero

$$\mu\nabla.f - \mu\Delta u + \lambda\Delta\Delta u - \lambda\Delta(d^k - b^k) = 0. \tag{4.4}$$

For the purpose of finding our minimal deformation u we will rearrange and discretize this optimality condition. One way of how the discretization can be done is given later.

$$\underbrace{(\mu\Delta - \lambda\Delta\Delta)}_{A}\underbrace{u^{k+1}}_{x} = \underbrace{\mu\nabla.f - \lambda\Delta(d^k - b^k)}_{b}. \tag{4.5}$$

Note that to calculate u^{k+1} the matrix A in general has to be inverted. In our case the matrix A is a band matrix. This encourage us to use the iterative Jacobi-method. A short description of the method follows.

For an invertible matrix in $\mathbb{R}^{n\times n}$, there are matrices L and D in $\mathbb{R}^{n\times n}$, where L is a lower triangular with zeros in the diagonal and D is a diagonal matrix, such that

$$A = L + D + U.$$

As it can be found in countless books, e.g [9, Chapter 5, Section 2] the Jacobi method has the following form:

$$x^{k+1} = D^{-1}(b - (L + U)x^k).$$

Discretization of the differential operators

As mentioned before in order to be able to apply the Jacobi method one has to discretize Equation (4.4). In this part we want to talk about the numerical approximation of the differential operators $\nabla\cdot$, Δ and $\Delta\Delta$. One possibility of discretizing the differential operators is to approximate the derivatives by the finite differences.

At the starting point we will discretize the simplest version of our toy Example (3.2) where u and f are functions from $\mathbb{R} \to \mathbb{R}$. Therefore derivatives in one direction are needed. We will first discretize the differential operator ∇. Note, that in the one dimensional case there is no difference between the gradient and the divergence. The first derivative with respect to only one variable can be approximated by the forward difference quotient

$$\nabla_x^+ v := \frac{1}{h}(v_{i+1} - v_i).$$

In order to be able to understand the connection with the Jacobi-method, the discrete gradient can be written in a matrix form. Note that one can interpret the discrete gradient of the function u as matrix-vector multiplication where the matrix has the following form:

$$\nabla_x = \frac{1}{h}\begin{pmatrix} -1 & 1 & & 0 \\ 0 & -1 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 0 & -1 \end{pmatrix}.$$

Now we want to discuss the discretization of the second order Laplace operator which is also needed in Equation (4.4). It is a linear differential operator. Therefore we will need the similar backward difference quotient

$$\nabla_x^- v := \frac{1}{h}(v_i - v_{i-1}).$$

Higher derivatives can be approximated analogously. We will apply the backward difference quotient on the forward difference quotient . Therefore, we get the second order central difference quotient

$$\Delta^c v = \frac{1}{h^2}(v_{i+1} - 2v_i + v_{i-1}). \tag{4.6}$$

For Equation (4.5) we can rewrite the discrete second order Laplacian once again as a matrix

$$\Delta = \frac{1}{h^2}\begin{pmatrix} -2 & 1 & & 0 \\ 1 & -2 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & -2 \end{pmatrix}$$

Finally we want to introduce the discretization of the fourth order operator $\Delta\Delta$. Therefore we will apply the discrete second order operator on itself. As a result we get

$$\Delta\Delta^c v = \frac{1}{h^4}(v_{i+2} - 4v_{i+1} + 6v_i - 4v_{i-1} + v_{i-2}) \tag{4.7}$$

and therefore

$$\Delta\Delta = \frac{1}{h^4}\begin{pmatrix} 6 & -4 & 1 & & & 0 \\ -4 & 6 & -4 & \ddots & & \\ 1 & -4 & 6 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & -4 & 1 \\ & & \ddots & -4 & 6 & -4 \\ 0 & & & 1 & -4 & 6 \end{pmatrix}.$$

As we discretized all the operation needed we can plug in the obtained matrices into Equation (4.5). By simple algebraic matrix calculations we get our matrix A and our left hand side vector b. We will chose h often as 1 and therefore we will substitute the constants $\mu_1 = \frac{\mu}{h^2}$, $\lambda_1 = \frac{\lambda}{h^4}$, $\mu_2 = \frac{\mu}{h}$ and $\lambda_2 = \frac{\lambda}{h^2}$. Now it can be easily seen that applying the Jacobi algorithm leads to the iteration

$$\begin{aligned} u_i^{k+1} = & - \frac{1}{2\mu_1 + 6\lambda_1}\Big[\mu_2(-f_i + f_{i+1}) - \\ & - \lambda_2\Big((d^k - b^k)_{i-1} - 2(d^k - b^k)_i + (d^k - b^k)_{i+1}\Big) + \\ & + \lambda_1 u_{i-2} - (\mu_1 + 4\lambda_1)u_{i-1} - (\mu_1 + 4\lambda_1)u_{i+1} + \lambda_1 u_{i+2}\Big]. \end{aligned}$$

The 2D case:

Previously we assumed that the function in Problem (3.2) are from $\mathbb{R} \to \mathbb{R}$. In order to approach the model form article [2] we will review the problem with the given function f from $\mathbb{R}^2 \to \mathbb{R}^2$ and the wanted function $u : \mathbb{R}^2 \to \mathbb{R}$. We want to show one way of discretizing the differential operators in the 2D case. Starting again with the divergence for the vector field $f : \mathbb{R}^2 \to \mathbb{R}^2$ on the right hand side. As we know that the divergence is just the sum of the first component differentiated with respect to x and the second component differentiated with respect to y

$$\nabla . f = \nabla_x f_1 + \nabla_y f_2$$

where ∇_y has an equivalent form ∇_x. Therefore all the information needed were already derived before.

In this part we want to talk about the Laplace operator for our given function u from $\mathbb{R}^2 \to \mathbb{R}$. As we know the Laplacian is the sum of the second derivatives with respect to x and y. Therefore we only have to sum up the central difference quotient from the first part of (4.6) for x and for y and we get

$$\Delta^c v = \frac{1}{h^2}(u_{i+1,j} + u_{i-1,j} - 4u_{i,j} + u_{i,j-1} + u_{i,j+1}). \tag{4.8}$$

As we have now mixed derivatives of a function this discretization in 2D is a bit more of a challenge. In the previous paragraph the discretized function u was a vector and therefore the discretization of the Laplacian was easily described as a matrix. In this case the discrete function u is a matrix $u_{i,j} \in \mathbb{R}^{m \times n}$ and the operator would have the form of a cuboid. To bring this in a mathematical structure with which one can do basic algebraic operations we will convert the Laplacian cuboid into a matrix. Therefore we will create a new index k which will be a combination of the indices i, j. The matrix $u_{i,j}$ will be rewritten into a vector u_k of the size $m.n$ where each column is below the other. We will not explain this conversion in detail because of the way we did the implementation we only need to know how the discretization stencil looks

$$\Delta = \frac{1}{h^2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

We will continue with the discretization of the two dimensional fourth

order operator again for our wanted function $u : \mathbb{R}^2 \to \mathbb{R}^2$. Therefore we have to apply the Laplace operator on itself which lead us to

$$\Delta\Delta = \Delta_{xx}\Delta_{xx} + 2\Delta_{xy}\Delta_{xy} + \Delta_{yy}\Delta_{yy}$$

The fist part and the last part, the fourth order derivatives with respect to one variable where already discussed earlier in this section. One has to sum up (4.7) once with respect to variable x and once again with respect to variable y. The mixed derivative can also be easily derived by applying formula (4.6) on itself but with respect to the other index j. Therefore we get

$$\begin{aligned}\Delta\Delta = \frac{1}{h^4} \quad & (u_{i+2,j} + u_{i-2,j} - 8u_{i+1,j} - 8u_{i-1,j} + 20u_{i,j} - \\ & -8u_{i,j+1} - 8u_{i,j-1} + u_{i,j+2} + u_{i,j-2}).\end{aligned}$$

which gives us the fourth order stencil

$$\Delta\Delta = \frac{1}{h^4}\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Putting all the things together and again substituting some constants $\mu_1 = \frac{\mu}{h^2}$, $\lambda_1 = \frac{\lambda}{h^4}$, $\mu_2 = \frac{\mu}{h}$ and $\lambda_2 = \frac{\lambda}{h^2}$, the Jacobi iteration has the following form:

$$\begin{aligned} u_i^{k+1} = \; & - \; \frac{1}{4\mu_1 + 16\lambda_1}\Big[\mu_2(f_{i-1} - 2f_i + f_{i+1}) - \\ & - \; \lambda_2\Big((d^k - b^k)_{i-2} + (d^k - b^k)_{i-1} - 4(d^k - b^k)_i + \\ & + \; (d^k - b^k)_{i+1} + (d^k - b^k)_{i+2}\Big) + \\ & + \; \lambda_1 u_{i-6} + \lambda_1 u_{i-5} + \lambda_1 u_{i-4} + \lambda_1 u_{i-3} - \\ & - \; (\mu_1 + 6\lambda_1)u_{i-2} - (\mu_1 + 6\lambda_1)u_{i-1} - (\mu_1 + 6\lambda_1)u_{i+1} - (\mu_1 + 6\lambda_1)u_{i+2} + \\ & + \; \lambda_1 u_{i+3} + \lambda_1 u_{i+4} + \lambda_1 u_{i+5} + \lambda_1 u_{i+6}\Big]. \end{aligned}$$

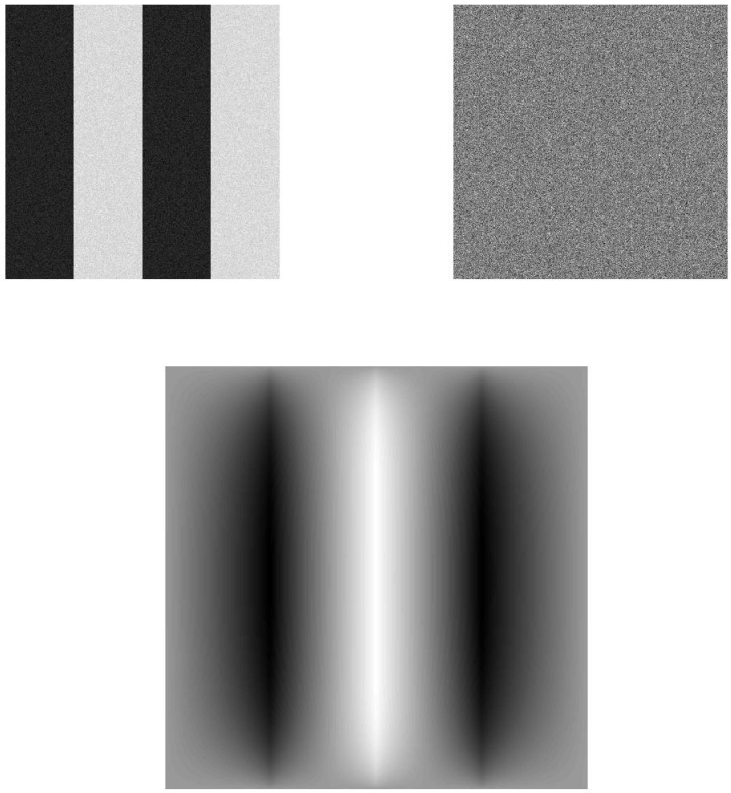

Figure 4.1: The Thin Film deformation in 2D

At the end of this chapter we want to show some outputs of the algorithm. The first two picture represent the noisy- x and y components of the input function f. After applying the Split-Bregman method as explained in 4.1.2 we get the following result shown in figure 4.1

4.2 Implementation of the plasticity toy model

In this section we want to talk about the implementation of the Split-Bregman method on the plasticity toy model (3.3) we discussed in Section 3.2. As in Section 4.1.2 we are also going to discuss the discretization and the computation of the solution via a numerical method. In this case, we have to distinguish between the anisotropic and the isotropic case. Because this model is from the area of plasticity it is more authentic to use the isotropic

case. Therefore, first we are going to rewrite the problem

$$\min_u \|H(u)\|_1 - \frac{1}{\mu}\iint f.u \,\mathrm{d}x\,\mathrm{d}y$$
$$\Leftrightarrow \min_u \iint \sqrt{\Delta_{xx}u^2 + 2\Delta_{xy}u^2 + \Delta_{yy}u^2}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{\mu}\iint f.u \,\mathrm{d}x\,\mathrm{d}y$$

As in the previous section our first step for solving the problem is to apply the Split-Bregman method to our problem. In this isotropic case we will substitute each derivative with respect to the directions separately. Replace the second order derivative Δ_{xx} by the variable d_{xx}, analogously $\Delta_{xy}u$ by d_{xy} and $\Delta_{yy}u$ by d_{yy}. Therefore, we get the following constrained problem:

$$\begin{aligned}
\min_u \quad & \iint \sqrt{d_{xx}^2 + 2d_{xy}^2 + d_{yy}u^2}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{\mu}\iint f.u \,\mathrm{d}x\,\mathrm{d}y \\
\text{s.t} \quad & d_{xx} = \Delta_{xx}u \\
& d_{xy} = \Delta_{xy}u \\
& d_{yy} = \Delta_{yy}u.
\end{aligned}$$

As before we want to weakly and strongly enforce the constraints. This is done again by adding a penalty function term and applying the Brgeman iteration which leads us to the unconstrained problem

$$\begin{aligned}
\min_{u,d_{xx},d_{xy},d_{yy}} \quad & \iint \sqrt{d_{xx}^2 + 2d_{xy}^2 + d_{yy}u^2}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{\mu}\iint f.u \,\mathrm{d}x\,\mathrm{d}y + \\
+ \quad & \frac{\lambda}{2}\|d_{xx} - \Delta_{xx}u - b_{xx}^k\|_2^2 + \lambda\|d_{xy} - \Delta_{xy}u - b_{xy}^k\|_2^2 + \\
+ \quad & \frac{\lambda}{2}\|d_{yy} - \Delta_{yy}u - b_{yy}^k\|_2^2. \qquad (4.9)
\end{aligned}$$

How to calculate the variables b_{xx}, b_{xy} and b_{yy} is again derived in (2.5) from Chapter 2 and in details explained in the Split-Bregman algorithm 2.6:

$$\begin{aligned}
b_{xx}^{k+1} &= b_{xx}^k + (\Delta_{xx}u^{k+1} - d_{xx}^{k+1}) \\
b_{xy}^{k+1} &= b_{xy}^k + (\Delta_{xy}u^{k+1} - d_{xy}^{k+1}) \\
b_{yy}^{k+1} &= b_{yy}^k + (\Delta_{yy}u^{k+1} - d_{yy}^{k+1}).
\end{aligned}$$

In order to solve the equation 4.9 from above and therefore the plasticity toy model two steps must be performed. One has to minimize again with respect to u and the variables d_{xx}, d_{xy} and d_{yy} separately

$$
\begin{aligned}
u^{k+1} = \min_u \quad & - \quad \frac{1}{\mu}\iint f.u \,\mathrm{d}x\,\mathrm{d}y + \\
& + \quad \frac{\lambda}{2}\|d_{xx} - \Delta_{xx}u - b_{xx}^k\|_2^2 + \lambda\|d_{xy} - \Delta_{xy}u - b_{xy}^k\|_2^2 + \\
& + \quad \frac{\lambda}{2}\|d_{yy} - \Delta_{yy}u - b_{yy}^k\|_2^2. \qquad (4.10)
\end{aligned}
$$

First we want to mention that the minimization over u can be solved equivalently as in the case for the thin film model numerically with the Jacobi method. Note that in this application the variables d_{xx}, d_{xy} and d_{yy} do not decouple as they did in the previous section 4.1. In order to apply the iterative minimization procedure to this problem, we have to solve the following subproblem

$$
\begin{aligned}
\min_{d_{xx},d_{xy},d_{yy}} \quad & \iint \sqrt{d_{xx}^2 + 2d_{xy}^2 + d_{yy}u^2}\,\mathrm{d}x\,\mathrm{d}y + \\
& + \quad \frac{\lambda}{2}\|d_{xx} - \Delta_{xx}u - b_{xx}^k\|_2^2 + \lambda\|d_{xy} - \Delta_{xy}u - b_{xy}^k\|_2^2 + \\
& + \quad \frac{\lambda}{2}\|d_{yy} - \Delta_{yy}u - b_{yy}^k\|_2^2.
\end{aligned}
$$

In this case minimizing over the variables d_{xx}, d_{xy} and d_{yy} changes a bit because they do not decouple. Despite this fact, we can still explicitly solve the minimization problem using a generalized shrinkage formula [18]

$$
\begin{aligned}
d_{xx}^{k+1} &= \max(s^k - 1, 0)\frac{\Delta_{xx}u^k + b_{xx}^k}{s^k}) \\
d_{xy}^{k+1} &= \max(s^k - 1, 0)\frac{\Delta_{xy}u^k + b_{xy}^k}{s^k}) \\
d_{yy}^{k+1} &= \max(s^k - 1, 0)\frac{\Delta_{yy}u^k + b_{yy}^k}{s^k})
\end{aligned}
$$

where

$$
s^k = \sqrt{|\Delta_{xx}u^k + b_{xx}^k|^2 + 2|\Delta_{xy}u^k + b_{xy}^k|^2 + |\Delta_{yy}u^k + b_{yy}^k|^2}.
$$

The numerics of this section are analogue to the numerics in Section 4.1.2. Therefore, we will not discuss them in this work.

Bibliography

[1] Pavel Belik and Michell Luskin. A total-variation surface energy model for the thin films of martensitic crystals. *Interfaces and Free Boundaries*, 4:71 – 88, 2002.

[2] Pavel Belik and Michell Luskin. Approximation by piecewise constant functions in a BV metric. *Mathematical Models and Methods in Applied Sciences*, 13:373 – 393, 2003.

[3] Kaushik Bhattacharya and Richard D. James. A theory of thin films of martensitic materials with applications to microactuators. *J. Mech. Phys. Solids*, 47(3):531 – 576, 1999.

[4] L.M. Bregman. The relaxation method of finding the common point convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. and Math. Phys.*, 7:200 – 217, 1967.

[5] Christoph Brune. 4D Imaging in Tomography and Optical Nanoscopy. Ph.d. thesis, Westfälische Wilhelms- Universität Münster, 2010.

[6] D.L. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J.ACM*, 9:84, 1962.

[7] Gilles Aubert and Pierre Kornprobst. *Mathematical Problems in Image Processing*. Springer-Verlag, New York, 2002.

[8] T. Goldstein and S. Osher. The Split Bregman Method for L1-Regularized Problems. *SIAM J. Imaging Sci.*, 2(2):323 – 343, 2009.

[9] Michael Jung and Ulrich Langer. *Methode der finiten Elemente für Ingenieure*. Teubner, Stuttgart/Leipzig/Wiesbaden, 2001.

[10] Stefan Kindermann. Signal- und Bildverarbeitung. Skript zur Vorlesung, Universität Linz, 2007.

[11] Leonid I.Rudin, Stanley Osher and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259 – 268, 1992.

[12] Jacob Lubliner. *Plasticity Theory*. Dover Publications, 2008.

[13] S. Twomey. On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. *J.ACM*, 10:97, 1963.

[14] S. Twomey. The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements. *J. Franklin Inst.*, 297:95, 1965.

[15] Xue-Cheng Tai and Chunlin Wu. Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model. In *SSVM*, volume 5567, pages 502–513. Springer, 2009.

[16] R. Teman. *Mathematical Problems in Plasticity*. Gautgier-villars, Paris, 1983.

[17] A.N. Tikhonov and V.Y. Arsenin. *Solution of Ill-Posed Problems*. Wongston and Sons, Washington D.C., 1977.

[18] W Yin Y Wang and Y Zhang. A fast algorithm for image deblurring with total variation regularization. *CAAM Technical Reports*, 2007.