
A Non-Linear Diffusion Filtering Method for Reconstruction Problems

MASTER OF SCIENCE THESIS BY
MARIE FOGED SCHMIDT



TECHNICAL UNIVERSITY OF DENMARK, DTU COMPUTE,
DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER
SCIENCE

 **DTU Compute**
Department of Applied Mathematics and Computer Science

By:

Marie Foged SCHMIDT, s082746

Supervisor:

Kim KNUDSEN

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Matematiktorvet, Building 303B,
2800 Kgs. Lyngby, Denmark.
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk

Project title: A Non-Linear Diffusion Filtering Method for Reconstruction Problems

Project period: 5. May - 3. November, 2014

Workload: 30 Honors ECTS

Degree: Master of Science

Programme: Mathematical Modelling and Computation (Honors)

Copyright: © Marie Foged Schmidt, 2014

Preface

This thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark (DTU) in the fulfilment of the requirements for acquiring an M.Sc. in Mathematical Modelling and Computation. It represents the completion of my honors master programme at DTU and the workload corresponds to 30 honors ECTS. The work was conducted from May 2014 to November 2014, under the supervision of Associate Professor Kim Knudsen at DTU Compute.

The thesis deals with the solution of reconstruction problems such as medical imaging through either regularization methods or non-linear diffusion processes and the relation between those solutions. The main theorems and proofs about existence of solutions are based on [1] and [2].

The prerequisites for reading this thesis is a basic understanding of the theory of inverse problems, minimization problems, and partial differential equations. A more profound understanding of functional analysis is required. Familiarity with Sobolev spaces and their role with regard to partial differential equations and their weak form is recommended. In Chapter 2 we introduce the space of functions of bounded variation as it will play an important role in the study of minimization problems. Finally, knowledge about measure theory and semi-groups might be helpful, but is not strictly necessary in order to understand the material. Appendix A includes some useful definitions and results regarding Sobolev spaces and their duals, minimization problems and some properties of subdifferentials.

Marie Foged Schmidt
November, 2014

Acknowledgements

First and foremost I would like to thank my supervisor, Associate Professor Kim Knudsen from DTU Compute, for introducing me to the subject of this thesis and for taking the time to discuss the topic with me. Kim has been very interested in the work I have been doing and during our weekly meetings he has been involved in discussions about the issues of the thesis. Kim has been my co-mentor on the honors programme and has been involved in many special courses developing my theoretical background.

I will also thank Associate Professor Christian Henriksen from DTU Compute who has been my honors mentor and has been offering me a lot of special courses from the beginning of my studies. Christian has always been up for half day discussions at the blackboard and has invited me to various conferences, which has encouraged me in my studies and broadened my international network.

Finally, I would like to thank Assistant Professor Martin S. Andersen from DTU Compute for taking the time to have a meeting with me on the subject of this thesis, and my fellow student, Sidsel P. Nielsen, who works with image analysis, for providing me with articles, which apply a different method to the thesis subject. My friends and family deserve special thanks for supporting me and in particular Denis Tcherniak and Axel Krämer for reading my thesis and giving me feedback.

Abstract

The purpose of this thesis is to investigate two different approaches for solving reconstruction problems. Reconstruction problems arise in, for example, medical imaging, satellite imaging, data compression, fingerprint analysis, and much more. A reconstruction problem may be seen as an inverse problem for which regularization methods are applied in order to obtain a reasonable and satisfactory reconstruction. This approach leads to the formulation of a minimization problem for which existence and uniqueness results are proven. The minimization problem turns out to be associated with an Euler-Lagrange equation in distributional sense for the minimizer. This Euler-Lagrange equation is turned into a non-linear diffusion problem for which the existence of a solution to the problem in its strong formulation is studied.

The thesis gives an introduction to regularization methods and minimization problems, followed by a study of the related non-linear diffusion problems and their solutions. The proofs of the existence results for the minimization problem and the non-linear diffusion problem are based on [1] and [2], respectively.

In order to investigate the diffusion problems in greater detail, numerical experiments are performed. The discretized iteration schemes for the diffusion problems are implemented in `MatLab` [33]. The experiments show that one needs an optimal way of choosing the time step length for the discretized diffusion problem and an optimal stopping criteria for the iteration process.

List of Figures

1.1	Degradation of image.	2
1.2	Computed tomography sinograms.	3
1.3	Shepp-Logan phantom.	4
2.1	x^* is a density point of the union of the two open white disks, but is a rarefaction point of the complementary of the union.	14
7.1	Denoising 1D: Data for reconstruction.	96
7.2	Denoising 1D: Standard Tikhonov reconstructions at different times. Blue (clean data), green (iteration 1), black (iteration 10), magenta (iteration 100).	97
7.3	Denoising 2D: Data for reconstruction.	98
7.4	Denoising 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 10^{-7}$	99
7.5	Denoising 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 5 \cdot 10^{-7}$	100
7.6	Denoising 1D: Total variation reconstructions at different times. Blue (clean data), green (iteration 1), black (iteration 10), magenta (iteration 100).	101
7.7	Denoising 2D: Total variation reconstructions at different times with $\Delta t = 1/2\Delta x$	102
7.8	Denoising 2D: Total variation reconstructions at different times with $\Delta t = \Delta x$	103
7.9	Deblurring 1D: Data for reconstruction.	103

7.10	Deblurring 1D: Standard Tikhonov reconstructions at different times with $\Delta t = 6 \cdot 10^{-6}$. Blue (actual function), green (iteration 50000), black (iteration 200000), magenta (iteration 1100000).	104
7.11	Deblurring 1D: Standard Tikhonov reconstructions at different times with $\Delta t = 5 \cdot 10^{-4}$. Blue (actual function), green (iteration 100), black (iteration 7000), magenta (iteration 20000).	105
7.12	Deblurring 2D: Data for reconstruction.	105
7.13	Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 0.05$	106
7.14	Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 1$	107
7.15	CT: X-ray going straight through domain with attenuation constant $\mu(\eta) = \mu$	108
7.16	CT: Shepp-Logan phantom.	108
7.17	CT: Data for reconstruction.	109
7.18	CT: Standard Tikhonov reconstructions at different times with $\Delta t = 2 \cdot 10^{-5}$	110
7.19	CT: Standard Tikhonov reconstructions at different times with $\Delta t = 4 \cdot 10^{-5}$	111
7.20	Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 0.2$	112

Notation and Symbols

Throughout the report the notation and symbols defined in this section will be used.

Function and measure spaces

For an open and bounded subset $\Omega \subset \mathbb{R}^n$ we define the real-valued function spaces:

$BV(\Omega)$	Space of functions of bounded variation.
$C_c^p(\Omega)$	Space of real-valued functions, p continuously differentiable with compact support.
$C_c^\infty(\Omega)$	Space of real-valued functions, infinitely continuously differentiable with compact support.
$L^p(\Omega)$	Space of Lebesgue measurable functions for $1 \leq p \leq \infty$.
$\mathcal{M}(\Omega)$	Space of Radon measures.
$W^{1,p}(\Omega)$	Sobolev space for $1 \leq p \leq \infty$.
$B(X, Y)$	Space of bounded linear operators from X to Y .

The spaces defined above can be regarded as vector-valued spaces as well, and we will write for example $W^{1,p}(\Omega, \mathbb{R}^n)$ meaning the vector valued Sobolev space in \mathbb{R}^n .

For a functional $F : X \rightarrow]-\infty, +\infty]$ where X is a Banach space we define:

$\operatorname{argmin} F$	$\{u \in X : F(u) = \inf_X F(v)\}.$
l.s.c.	Lower semi-continuous

Measures

For a Radon measure μ we define:

μ -a.e. x	For almost every x regarding the measure μ .
$ \mu $	Total variation of the measure μ . If μ is vector-valued, then $ \mu = \mu_1 + \mu_2 + \dots + \mu_n $.
dx	Lebesgue measure in \mathbb{R}^n .
$\mathcal{L}(\Omega)$	One-dimensional Lebesgue measure of Ω .
$\mathcal{L}^n(\Omega)$	n -dimensional Lebesgue measure of Ω .

Functions

For a function $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ and a sequence of functions $(f_n)_{n \in \mathbb{N}}$, $f_n : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ we define:

$\text{supp}(f)$	The support of f .
Df	Distributional gradient of f .
χ_E	Characteristic function on the set E .

Convergences

Let X be a normed space. We define:

$x_n \rightarrow x$ in X	Strong (norm) convergence on X .
$x_n \rightharpoonup x$ in X	Weak convergence on X ($f(x_n) \rightarrow f(x)$ for all $f \in X'$).
$x'_n \xrightarrow{*} x'$ in X'	Weak* convergence on X' ($x'_n(x) \rightarrow x'(x)$ for all $x \in X$).

Miscellaneous notation

$ \cdot $	Euclidean norm in \mathbb{R}^n .
$B_r(x) \subset \mathbb{R}^n$	Ball of center x and radius r in \mathbb{R}^n .
\mathbb{R}_+	Positive real axis.
$\mathbb{R}_{\geq 0}$	$\mathbb{R}_+ \cup \{0\}$.
$(\cdot, \cdot)_{X, X'}$	Pairing between X' and X . Often we leave out the lower-case text X, X' .
$G^{-1}(D)$	Pre-image of D under G unless otherwise stated.
$*$	Convolution.
I	Identity operator.
2^U	The set of all subsets of the set U .

Contents

1	Introduction and Motivation	1
1.1	Deriving a Mathematical Model	3
1.2	Problem Formulation	7
2	BV Space	9
2.1	Definition and properties	9
2.2	Structure of BV functions	13
2.3	Minimization Problems in BV	16
3	Variational Formulation	19
3.1	Existence	20
3.2	Uniqueness	28
3.3	Stability	30
3.4	Examples	30
	3.4.1 Standard Tikhonov Regularization	30
	3.4.2 Total Variation Regularization	32
	3.4.3 Non-Convex Regularization	33
4	An associated Diffusion Filtering Method	37
4.1	Derivatives of Operators and the Subdifferential	37
4.2	Optimality Condition and Diffusion Filtering	40
4.3	Examples	46
	4.3.1 Denoising	46
	4.3.2 Bounded Linear Operator	49

5	Non-linear Diffusion Methods	51
5.1	Semi-Group Theory	52
5.2	Extension of the Hille-Yosida Theorem	55
5.2.1	The Exponential Formula	56
5.2.2	The Cauchy Problem	68
5.3	Examples	84
6	Non-linear Diffusion Methods in Dual Space	87
6.1	Properties of the Subdifferential	88
6.2	Cauchy Problem in Dual Space	92
7	Numerical Experiments	95
7.1	Denoising with Standard Tikhonov Regularization	95
7.2	Denoising with Total Variation Regularization	98
7.3	Deblurring with Standard Tikhonov Regularization	101
7.4	Computed Tomography with Standard Tikhonov Regularization	107
8	Discussion and Conclusion	113
8.1	Future Work	114
A	Theory	121
A.1	Sobolev Spaces and their Duals	121
A.2	Minimization Problems	123
A.2.1	Minimization Problems	123
A.2.2	Lower Semi-Continuous Functionals	123
A.2.3	Inf-Compact Functionals and Coercivity	125
A.2.4	Minimization Theorems	126
A.2.5	Relaxed Minimization Problems	128
A.2.6	The Direct Method of the Calculus of Variations	130
A.3	Properties of Subdifferentials	131
B	MatLab Implementation	133
B.1	Denoising 1D	133
B.2	Denoising 2D	135
B.3	Deblurring 1D	139
B.4	Deblurring 2D and Computed Tomography	141

1

Introduction and Motivation

Reconstruction problems arise in many different industry areas. For example, they are widely implemented in image analysis. Images are used to express physical situations worldwide. They may be used in medical imaging, satellite images, old movie restoration, 3D reconstructions of scenes or objects from images, robotics, character recognition, data compression, industrial quality control, fluids motion analysis, fingerprint analysis, and much more. The huge range of applications makes it important to be able to produce good quality images. That is, to be able to recover details in images from possibly noisy data containing information about the image. In medical imaging or tomography one wants to produce an image of the inner structure of the body using e.g. ultrasound, X-rays, or electrical currents or voltages. Sending such signals through the body and measuring the resulting signals gives information about the body's inner structure due to, for example, different attenuation coefficients for different tissues. The recovered image will be deteriorated due to possible defects of the imaging system and noise coming from any signal transmission. That is, the data for the reconstruction is not exact and leads to artifacts in the reconstructed image. Since even tiny details in, for example, medical imaging, are important, it is of great interest to remove noise from the given data while reconstructing the image of the body's inner structure.

The image recovering problems may be formulated as reconstruction problems in a continuous setting. Three examples of reconstruction problems are denoising, deblurring, and computed tomography. In a denoising problem we are given noisy

data, for example, a noisy image as seen in Figure 1.1b. From the noisy image we then want to reconstruct the clean image, seen in Figure 1.1a as good as possible. That is, we want to remove the noise. There are more questions arising when trying to solve such a problem. Since in practice we do not know the clean image, we cannot really tell from the noisy image which pixel values are correct and which are not. Hence in order to remove the noise we would somehow like to smoothen the image. But if we just smoothen in all areas of the image, we will destroy the edges in the image. Another approach is to just smoothen in isotropic areas and not across edges. All such considerations arise when trying to reconstruct the image.

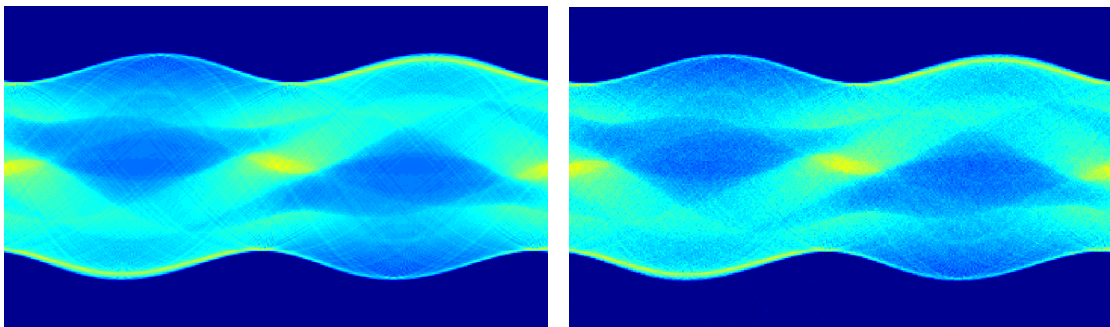


Fig. 1.1: Degradation of image.

Another kind of problem is a deblurring problem. For this we are given, for example, a blurred image as seen in Figure 1.1c and it may even be noisy too, as seen in Figure 1.1d. Now there are two things we need to do in order to get a good reconstruction of the original clean image. We need to remove noise, but we also

need to deblur the image, that is, to remove the blurring and make the edges sharp. Of course this is a more complex problem than the denoising problem and again more questions arise when trying to solve this kind of problem. Can we deblur the image and remove noise simultaneously? What knowledge do we need to obtain in order to deblur the image?

Finally, in a computed tomography problem, the available data is a sinogram, as seen in Figure 1.2a. The sinogram stems from measurements with a computed tomography imaging system. By sending X-rays through a body at different angles and measuring the corresponding signals, that is, the projected intensity data along straight lines through the body, one can obtain the sinogram. The first axis represents the angles at which the measurements are made and the second axis represents the projection displacement. Due to possible defects of the imaging system and noise coming from the signal transmission, the sinogram may contain noise and even missing angles. In this thesis, though, only the case of a sinogram with Gaussian additive noise is considered (see Figure 1.2b).



(a) Clean data.

(b) Noisy data (5% noise).

Fig. 1.2: Computed tomography sinograms.

From the noisy sinogram we should be able to reconstruct the inner structure of the body from which the measurements come. The true inner structure we should obtain is seen in Figure 1.3. In order to be able to reconstruct this phantom from the data in Figure 1.2b, we need some kind of model for the problem. In the next section we will see how we can model problems like denoising, deblurring, and computerized tomography.

1.1 Deriving a Mathematical Model

In order to solve problems like denoising, deblurring, and computerized tomography, we need a mathematical model describing these. The reconstruction problems are formulated in a continuous setting as follows: For the denoising problem, let u



Fig. 1.3: Shepp-Logan phantom.

denote the clean data and let δ denote the added noise. Then the noisy data is given by

$$u_\delta = u + \delta.$$

The problem of reconstructing the original clean data then reads: *From knowledge of u_δ , reconstruct u (without knowing δ).* This is what we call an *inverse problem*. Similarly for the deblurring problem, let u denote the function of interest, i.e. the function we want to reconstruct. Let \mathcal{K} denote an operator representing the blur (this could, for example, be a convolution operator) and let δ denote the added noise. The noisy data is then given by

$$u_\delta = \mathcal{K}u + \delta.$$

Again the inverse problem of reconstructing the original function is then: *From knowledge of u_δ (and \mathcal{K}), reconstruct u (without knowing δ).* Finally for the computed tomography problem, let u denote a function representing the inner structure of a body. Then the sinogram data is obtained by $\mathcal{R}u$, where \mathcal{R} is the Radon transform which exactly represents the intensity projection of u along lines. Let δ denote the noise in the sinogram data. The given data u_δ is then

$$u_\delta = \mathcal{R}u + \delta.$$

The inverse problem of reconstructing the image of the body's inner structure is then: *From knowledge of u_δ (and \mathcal{R}), reconstruct u (without knowing δ).*

From the above, we see that in general we can formulate reconstruction problems as inverse problems:

From knowledge of u_δ and the model $u_\delta = \mathcal{F}u + \delta$, reconstruct u .

Here \mathcal{F} is a known operator and δ is the unknown added noise. How should we approach such a problem? If we had no noise we would try to minimize the norm

$$\|u_\delta - \mathcal{F}u\|_X, \quad (1.1.1)$$

where the normed space X is chosen in a feasible way. If \mathcal{F} is linear then a minimizer u satisfies the normal equation

$$\mathcal{F}^\# u_\delta - \mathcal{F}^\# \mathcal{F}u = 0$$

where $\mathcal{F}^\#$ is the adjoint operator of \mathcal{F} . Hence we could try to find a solution to the normal equation in order to find a minimizer of (1.1.1). But since $\mathcal{F}^\# \mathcal{F}$ is not always injective, we cannot be sure that the above equation admits a unique solution. Furthermore if, for example, \mathcal{F} is a convolution operator, then $\mathcal{F}^\# \mathcal{F}$ admits small eigenvalues causing numerical instabilities. Finally, if noise was actually present, we would fit the solution to noise, which we definitely do not want to do. The problem is what we call *ill-posed* and needs some kind of regularization. That is, we add a regularizing term to the functional in (1.1.1), which should make the minimization problem well-posed such that it admits a unique solution which is stable with respect to small perturbations in the initial data. In this thesis, only a generalized Tikhonov regularization will be considered. That is, we want to minimize

$$\tau_{u_\delta, \alpha}(u) = \|u_\delta - \mathcal{F}u\|_X^2 + \alpha \|g(x, u, Du)\|_Y^2, \quad \alpha > 0.$$

Here the spaces X and Y and the function g need to be chosen according to knowledge about the noise model and a priori information.

The first term

$$\|u_\delta - \mathcal{F}u\|_X^2$$

of the Tikhonov functional is called the *fidelity term* and tells something about the noise in the given data. Hence the space X should reflect the noise model. In the discrete case where u_δ is actually sampled in a finite number of points N , we restrict ourselves to a Gaussian additive noise model. Then δ will consist of N realizations of independent Gaussian distributed random variables. It can be shown that for such a noise model, X should be $L^2(\Omega)$ in the continuous setting (see [18, pp. 46-47]). Hence we will use an L^2 -norm for the fidelity term:

$$\frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}u(x)|^2 dx,$$

where $\Omega \subset \mathbb{R}^n$ is the domain of u and u_δ .

The second term

$$\|g(x, u, Du)\|_Y^2$$

is called the *penalizing term* and should reflect the prior knowledge. Hence g and Y should be chosen according to prior information. Here we restrict ourselves to terms of the form

$$\int_{\Omega} g(x, u(x), Du(x)) \, dx,$$

where g then reflects the prior information, that could be e.g. that u is smooth or u has jumps. When g satisfies certain conditions, the minimization of the final generalized Tikhonov functional

$$\tau_{u_\delta, \alpha}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}u(x)|^2 \, dx + \alpha \int_{\Omega} g(x, u(x), Du(x)) \, dx$$

is then well-posed and admits a unique solution which will be our guess for the original function or image u . But in which space should we look for a solution? In Section 3 we show that it is important that the solution space is a reflexive Banach space. It will turn out that a natural choice of solution spaces is $W^{1,p}(\Omega)$ for $1 < p < \infty$. In some problems we would like to set $p = 1$. Since $W^{1,1}(\Omega)$ is not reflexive we introduce in Chapter 2 an extension of $W^{1,1}(\Omega)$, namely the space of functions of bounded variation $BV(\Omega)$, which is a Banach space and satisfies the needed properties of a reflexive space. The chapter is self-contained and may be skipped if the reader is already familiar with the BV space. In Chapter 3 the minimization of the Tikhonov functional will be discussed for $1 \leq p < \infty$. It will be proved that there exists a unique minimizer under certain restrictions on g and some well-known examples will be proved to satisfy these restrictions.

The formulation of the reconstruction problems given above is called the *variational formulation* and is well-known in, for example, tomography. However, it turns out that the reconstruction problems can be formulated in a completely different setting. Writing down what it means for a function u to be a minimizer of the Tikhonov functional $\tau_{u_\delta, \alpha}$ it turns out that u actually satisfies an Euler-Lagrange equation in distributional sense, which can be turned into a non-linear diffusion problem as shown in Chapter 4. This opens up for a completely different way to analyse the reconstruction problems. In some cases, we can even make sense of the diffusion problem in its strong formulation. In Chapter 5 and 6 it will be proven under which conditions the non-linear diffusion problem admits a strong solution, that is, a solution to the strong formulation of the diffusion problem. It will also be discussed how the conditions for the existence of a solution to a non-linear diffusion problem for a reconstruction problem relates to the conditions for existence of a minimizer for the corresponding variational formulation. The

variational formulation can now be set aside, as the reconstruction problems can at this point be solved using non-linear diffusion processes. Solving the non-linear diffusion problem for the function u will lead to the construction of a family of restored functions $\{u(t, x)\}_{t>0}$. The idea is that as t increases, the function $u(t, x)$ should be a more and more simplified version of u_δ but should still preserve certain structures, such as jumps, i.e. edges in an image. Moreover, no new structures should be created. For this reason t may be regarded as a *scale variable* instead of as a time variable. Choosing homogeneous Neumann boundary conditions implies that no new information is put into the system and no information leaves the system.

The diffusion process formulation for the reconstruction problems leads to possibly new “regularization” techniques which may not have a corresponding variational formulation. Keeping in mind that a diffusion process is smoothing, then for the denoising and deblurring problems for instance, we would like diffusion to take place in isotropic areas only and not across edges. These thoughts can be built into a non-linear diffusion process which then serves as the regularization of the problems.

In order to solve a non-linear diffusion problem, the partial differential equation (PDE) needs to be discretized. The most straightforward discretization model is the finite difference method, which is used in this thesis. Now the question of how to choose the time step size arises. If we choose the step size too small, then the iteration process is very slow, which means that the reconstruction process may require a huge amount of iterations. If we choose the time step size too large, we may skip the steps where noise is actually removed or diminished and then we will end up fitting data to noise. Hence there should be a balanced way of choosing the time step size. In Chapter 7 these considerations are discussed through numerical experiments. Still keeping in mind that the diffusion process is smoothing and as time evolves we will obtain a more and more simplified version of u_δ we have to stop the diffusion process before the details in u_δ that we actually need to keep are removed or diminished. This suggests that there is an optimal stopping criteria for the iteration process. In the numerical experiments of this thesis, however, the stopping time is chosen by trial and error.

1.2 Problem Formulation

The main goal of this thesis is to study the variational formulation and the non-linear diffusion method and their relation for general reconstruction problems. The noise model will be assumed to be additive and in the discrete case restricted to Gaussian additive noise. For the variational formulation the existence and uniqueness of a minimizer will be investigated and for the non-linear diffusion method only the existence will be examined. Different diffusion filtering methods and their strengths will be examined through numerical experiments for denoising,

deblurring and computed tomography examples.

2

BV Space

¹As described in Chapter 1, sometimes we want to be able to preserve features such as jumps in function values under the reconstruction of a function. For an image that would be edges. Hence the solution space of our problem must permit discontinuous solutions. The first distributional derivative of a function can be regarded as a measure which may charge zero Lebesgue measure sets. For example an edge in a two dimensional image will have 0 Lebesgue measure, but we may want to assign a “size” to it through the distributional derivative measure in order to distinguish it from isotropic areas. The solution of such a problem cannot be found in classical Sobolev spaces. Thus we introduce the new space of functions of bounded variation $BV(\Omega)$ and show that it is an extension of the Sobolev space $W^{1,1}(\Omega)$. The $BV(\Omega)$ space then completes the classical theory of Sobolev spaces.

2.1 Definition and properties

In this first section we define the BV space and equip it with a norm making it a Banach space. Then we show that the first distributional derivative of a BV -function can be regarded as a measure and show that the Sobolev space $W^{1,1}(\Omega)$ is a subset of $BV(\Omega)$. Finally we state an important lemma replacing the reflexivity of the Sobolev spaces $W^{1,p}(\Omega)$ for $1 < p < \infty$. This lemma makes the BV space

¹This Chapter is based on [20, ch. 10-11].

useful in minimization problems.

Throughout the section we let Ω denote a bounded open subset of \mathbb{R}^n . For a function $u \in L^1(\Omega)$ we define the total mass $|Du|(\Omega)$ by

$$|Du|(\Omega) = \sup \left\{ \int_{\Omega} u \operatorname{div} \phi \, dx : \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\},$$

where dx is the Lebesgue measure. Then the space of bounded variation is defined by

Definition 2.1.1 ($BV(\Omega)$). The *space of functions of bounded variation* $BV(\Omega)$ is defined by

$$BV(\Omega) = \{u \in L^1(\Omega) : |Du|(\Omega) < \infty\}.$$

In order to be able to define convergence properties in $BV(\Omega)$ we need to define a norm. $BV(\Omega)$ is equipped with the norm given by

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega).$$

It can be shown that the BV space equipped with this norm is a Banach space (see [20, Thm. 10.1.1]).

We now show that if $u \in BV(\Omega)$ then the distributional derivative Du of u can be identified with a vector-valued Radon measure². This allows for giving a possibly zero Lebesgue measure set a size using the distributional derivative measure instead. Letting $L : C_c^1(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$ be the functional defined by

$$L(\phi) = \int_{\Omega} u \operatorname{div} \phi \, dx.$$

Then L is clearly linear and since $u \in BV(\Omega)$ then

$$\sup \{L(\phi) : \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1\} = c < \infty,$$

where c is a constant only depending on Ω and u . Hence for all $\phi \in C_c^1(\Omega, \mathbb{R}^n)$ we have

$$|L(\phi)| \leq c \|\phi\|_{L^\infty(\Omega)}. \tag{2.1.1}$$

²A Radon measure on \mathbb{R}^n is a measure that is finite in each compact set $K \subset \mathbb{R}^n$.

Let $K \subset \Omega$ be a compact set and $\phi \in C_c(\Omega, \mathbb{R}^n)$ with $\text{supp}(\phi) \subset K$. Then for $\phi_\varepsilon = \eta_\varepsilon * \phi$ where η_ε is the standard mollifier (see [27, pp. 713-716]) we have $\phi_\varepsilon \in C_c^1(\Omega, \mathbb{R}^n)$ and

$$\begin{aligned} \phi_\varepsilon &\rightarrow \phi \quad \text{uniformly as } \varepsilon \rightarrow 0, \\ \|\phi_\varepsilon\|_{L^\infty(\Omega)} &\leq \|\phi\|_{L^\infty(\Omega)}, \quad \forall \varepsilon. \end{aligned}$$

From inequality (2.1.1) and the above we see that the sequence $(|L(\phi_\varepsilon)|)$ is uniformly bounded and hence the limit $\bar{L}(\phi) = \lim_{\varepsilon \rightarrow 0} L(\phi_\varepsilon)$ exists, and is independent of the choice of the sequence ϕ_ε . Then L uniquely extends to a linear bounded functional

$$\bar{L} : C_c(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}.$$

From the Riesz representation theorem (see [31, Section 1.8, Thm. 1, Cor. 1]) there exists a Radon measure μ and a μ -measurable function σ such that

$$\begin{aligned} |\sigma(x)| &= 1 \quad \mu - \text{a.e. } x, \\ \int_{\Omega} u \operatorname{div} \phi \, dx &= - \int_{\Omega} \sigma \cdot \phi \, d\mu, \quad \forall \phi \in C_c^1(\Omega, \mathbb{R}^n). \end{aligned}$$

This means that $Du = \sigma \, d\mu$ is a vector-valued Radon measure. For this reason we also denote $|Du|(\Omega)$ by $\int_{\Omega} |Du|$.

Using that for $u \in BV(\Omega)$ then Du is a Radon measure we can show that $W^{1,1}(\Omega)$ is a subset of $BV(\Omega)$: Since Du is a vector-valued Radon measure, there exists a Lebesgue decomposition (see [31, Section 1.6, Thm. 3]):

$$Du = \nabla u \mathcal{L}^n|_{\Omega} + D_s u,$$

where $\nabla u \in L^1(\Omega, \mathbb{R}^n)$ and $D_s u$ is singular with respect to the n -dimensional Lebesgue measure $\mathcal{L}^n|_{\Omega}$ restricted to Ω . This shows that $W^{1,1}(\Omega)$ is a subspace of $BV(\Omega)$ since if $u \in W^{1,1}(\Omega)$ then $u \in L^1(\Omega)$ and $Du \in L^1(\Omega, \mathbb{R}^n)$. Hence $u \in BV(\Omega)$ since for $D_s u = 0$ we have $Du = \nabla u \in L^1(\Omega, \mathbb{R}^n)$.

For $u \in BV(\Omega)$ we can now interpret Du as a Radon measure which leads to the following concept of *weak convergence* in the $BV(\Omega)$ space:

Definition 2.1.2 (Weak convergence). A sequence $(u_n)_{n \in \mathbb{N}} \subset BV(\Omega)$ converges weakly to $u \in BV(\Omega)$ written $u_n \rightharpoonup u$ if and only if the two following convergences hold:

- 1) $u_n \rightarrow u$ strongly in $L^1(\Omega)$
- 2) $Du_n \rightharpoonup Du$ weakly in $\mathcal{M}(\Omega, \mathbb{R}^n)$ (the space of all \mathbb{R}^n -valued Borel measures).

Remark 2.1.3. In Definition 2.1.2, $Du_n \rightharpoonup Du$ in $\mathcal{M}(\Omega, \mathbb{R}^n)$ means that $\int_{\Omega} u_n \operatorname{div} \phi \rightarrow \int_{\Omega} u \operatorname{div} \phi$ for all ϕ in $C_c(\Omega, \mathbb{R}^n)$.

The weak convergence leads to the following lower semi-continuity result (see definitions in appendix A.2):

Proposition 2.1.4. *Let $(u_k)_{k \in \mathbb{N}}$ be a sequence in $BV(\Omega)$ converging strongly to some u in $L^1(\Omega)$ and satisfying $\sup_{k \in \mathbb{N}} \int_{\Omega} |Du_k| < +\infty$. Then*

- i) $u \in BV(\Omega)$ and $|Du|(\Omega) = \int_{\Omega} |Du| \leq \liminf_{k \rightarrow +\infty} \int_{\Omega} |Du_k|$ (l.s.c. property).*
- ii) u_n weakly converges to u in $BV(\Omega)$.*

Proof. Since $u_k \in BV(\Omega)$ then for all $\phi \in C_c^1(\Omega, \mathbb{R}^n)$ with $\|\phi\|_{L^\infty(\Omega)} \leq 1$ we have

$$\begin{aligned} \int_{\Omega} u \operatorname{div} \phi \, dx &= \int_{\Omega} \lim_{k \rightarrow +\infty} u_k \operatorname{div} \phi \, dx = \lim_{k \rightarrow +\infty} \int_{\Omega} u_k \operatorname{div} \phi \, dx \\ &\leq \lim_{k \rightarrow +\infty} \int_{\Omega} |Du_k| = \liminf_{k \rightarrow +\infty} \int_{\Omega} |Du_k|. \end{aligned}$$

Taking the supremum on the left over all $\phi \in C_c^1(\Omega, \mathbb{R}^n)$ with $\|\phi\|_{L^\infty(\Omega)} \leq 1$ we obtain

$$\int_{\Omega} |Du| \leq \liminf_{k \rightarrow +\infty} \int_{\Omega} |Du_k| < \infty$$

which proves assertion (i).

In order to prove assertion (ii) we use the strong convergence of u_n to u in $L^1(\Omega)$. The strong convergence implies that Du_k converges to Du in distributional sense, i.e.

$$(\phi, Du_k) = - \int_{\Omega} u_k \operatorname{div} \phi \, dx \rightarrow - \int_{\Omega} u \operatorname{div} \phi \, dx = (\phi, Du), \quad \forall \phi \in C_c^\infty(\Omega, \mathbb{R}^n).$$

Using that $C_c^\infty(\Omega, \mathbb{R}^n)$ is dense in $C_c(\Omega, \mathbb{R}^n)$ for the L^∞ -norm and using the assumption of boundedness of $(Du_k)_{k \in \mathbb{N}}$ we conclude that the sequence $(Du_k)_{k \in \mathbb{N}}$ converges weakly to Du in $M(\Omega, \mathbb{R}^n)$. \square

Using the lower semi-continuity property from proposition 2.1.4 the following compactness result for the BV space can be proved:

Lemma 2.1.5. *Let $\Omega \subset \mathbb{R}^n$ be bounded, open, connected, and Lipschitz. Then every sequence $(u_k)_{k \in \mathbb{N}}$ in $BV(\Omega)$ satisfying $\sup_k \|u_k\|_{BV(\Omega)} < +\infty$ has a subsequence weakly converging to some function $u \in BV(\Omega)$.*

Proof. See [28, Prop. 3.13]. □

The above lemma is essential for reconstruction problems in $W^{1,1}(\Omega) \subset BV(\Omega)$ as described in Chapter 1. In Chapter 3 it is shown that we need bounded sets in the solution space of the reconstruction problem to be pre-compact. This property is satisfied by reflexive Banach spaces and therefore we can use $W^{1,p}(\Omega)$, $1 < p < \infty$, for the reconstruction problems. However, if we want to use $W^{1,1}(\Omega)$ for a reconstruction problem we need to extend the space to $BV(\Omega)$ since $W^{1,1}(\Omega)$ is not reflexive. Lemma 2.1.5 then replaces the reflexivity of the Sobolev spaces $W^{1,p}(\Omega)$ for $1 < p < \infty$ since it states that all bounded sets in $BV(\Omega)$ are pre-compact.

2.2 Structure of BV functions

The structure of a BV function u is naturally inherited from the level sets $[u > t]$, $t \in \mathbb{R}$. In order to motivate the structure theorem for BV functions we first look at a real-valued function u of bounded variation on an interval $I \subset \mathbb{R}$. It can be shown that such a function is the difference between two monotonous functions and therefore possesses two limits $\lim_{\delta \rightarrow 0} u(x_0 + \delta)$ and $\lim_{\delta \rightarrow 0} u(x_0 - \delta)$ at each point $x_0 \in I$ (see [20, Section 10.3]). We define the *jump set* to be the set of all points where the two limits are different, $S_u := \{x \in I : \lim_{\delta \rightarrow 0} u(x_0 + \delta) \neq \lim_{\delta \rightarrow 0} u(x_0 - \delta)\}$. In this section we show that this notion of a jump set can be generalized to any dimension. This is exactly the property we want in order to operate with jumps in function values and in particular edges in images.

In order to be able to define the jump set in any dimension, we need the notion of *density points*, *rarefaction points*, and *approximate limits*. The density and rarefaction points are generalizations of the well-known interior and exterior points of subsets in \mathbb{R}^n :

Definition 2.2.1 (Density Points and Rarefaction Points). Let E be a Borel subset of \mathbb{R}^n . A point $x_0 \in \mathbb{R}^n$ is a *density point* of E if and only if

$$\lim_{\rho \rightarrow 0} \frac{\mathcal{L}^n(B_\rho(x_0) \cap E)}{\mathcal{L}^n(B_\rho(x_0))} = 1.$$

A point x_0 is a *rarefaction point* of E if and only if

$$\lim_{\rho \rightarrow 0} \frac{\mathcal{L}^n(B_\rho(x_0) \cap E)}{\mathcal{L}^n(B_\rho(x_0))} = 0.$$

The set of all density points and all rarefaction points of E are respectively called *measure theoretical interior* and *measure theoretical exterior* of E and denoted by E_* and E^* .

In order to get an intuitive idea of the above definitions we give an example. Consider the example in Figure 2.1. Let E be the Borel subset of \mathbb{R}^2 consisting of the union of the two open white disks. Then all points in E are density points of E and all points in $\mathbb{R}^2 \setminus \overline{E}$ are rarefaction points of E . But anything can happen at the boundary ∂E . Take for example x^* seen in the figure. x^* is a density point of E but is a rarefaction point of E^c (the complementary of E).

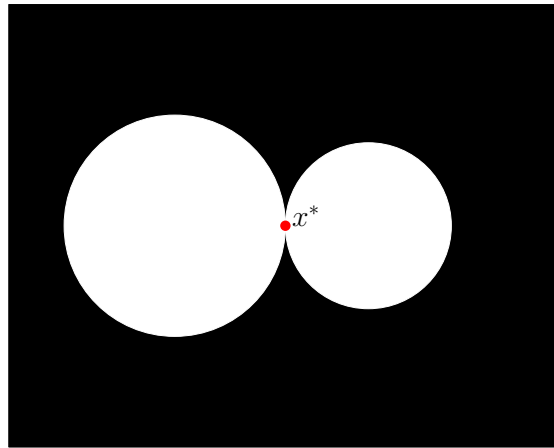


Fig. 2.1: x^* is a density point of the union of the two open white disks, but is a rarefaction point of the complementary of the union.

Next we define the *approximate limit*, the *approximate limit superior*, and the *approximate limit inferior* for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

Definition 2.2.2 (Approximate Limits). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a measurable function and $x_0 \in \mathbb{R}^n$. A real number α is the *approximate limit* of f at x_0 if and only if

$$\forall \varepsilon > 0, x_0 \text{ is a density point of the set } [|f - \alpha| < \varepsilon]$$

or equivalently

$$\forall \varepsilon > 0, x_0 \text{ is a rarefaction point of the set } [|f - \alpha| > \varepsilon].$$

We then write $\alpha = \text{ap lim}_{x \rightarrow x_0} f(x)$.

More generally we define in \mathbb{R} the *approximate limit supremum* and the *approximate limit infimum* of f at x_0 by

$$\text{ap lim sup}_{x \rightarrow x_0} f(x) = \inf \left\{ t \in \mathbb{R} : \lim_{\rho \rightarrow 0} \frac{\mathcal{L}^n(B_\rho(x_0) \cap [f > t])}{\mathcal{L}^n(B_\rho(x_0))} = 0 \right\}$$

and

$$\text{ap lim inf}_{x \rightarrow x_0} f(x) = \sup \left\{ t \in \mathbb{R} : \lim_{\rho \rightarrow 0} \frac{\mathcal{L}^n(B_\rho(x_0) \cap [f < t])}{\mathcal{L}^n(B_\rho(x_0))} = 0 \right\}.$$

Again looking at the example in Figure 2.1 and defining $f_1 = \chi_E$ where χ is the characteristic function, then for all $0 < \varepsilon < 1$ we have $[|f_1 - 1| < \varepsilon] = E$. Hence 1 is the approximate limit of f_1 for all points in $E \cup \{x^*\}$. On the other hand, defining $f_2 = \chi_{E^c}$ then for all $0 < \varepsilon < 1$ we have $[|f_2 - 0| > \varepsilon] = E^c$. Hence 0 is the approximate limit of f_2 for all points in $E \cup \{x^*\}$.

Before we define the generalized jump set we need the notion of a representative of a function $f \in L^1(\Omega)$. For the representative of f , still denoted f , we choose the one that satisfies $f(x_0) = \text{ap lim}_{x \rightarrow x_0} f(x)$ at every point $x_0 \in \Omega$ of approximate limit. Using the representative described above for every function $u \in BV(\Omega)$ we can define the generalized jump set in any dimension:

Definition 2.2.3 (Jump set). The jump set of a function $u \in BV(\Omega)$ whose representative satisfies the above convention is defined by

$$S_u = \{x \in \Omega : u^-(x) < u^+(x)\},$$

where $u^-(x) = \text{ap lim inf}_{y \rightarrow x} u(y)$ and $u^+(x) = \text{ap lim sup}_{y \rightarrow x} u(y)$.

We consider again the example in Figure 2.1 with $u = \chi_{E \cup \{x^*\}}$. Here we have

defined u to be 1 at the point x^* because then it satisfies the convention that $u(x_0) = \text{ap lim}_{x \rightarrow x_0} u(x)$ at every point x_0 of approximate limit. Then for every point $x_0 \in E$ we have

$$\begin{aligned} u^+(x_0) &= \text{ap lim sup}_{x \rightarrow x_0} u(x) = \inf\{t \in \mathbb{R} : t \geq 1\} = 1, \\ u^-(x_0) &= \text{ap lim inf}_{x \rightarrow x_0} u(x) = \sup\{t \in \mathbb{R} : t \leq 1\} = 1. \end{aligned}$$

This shows that no point in E is part of the jump set. Next, for every point $x_0 \in R^2 \setminus \overline{E}$ we have

$$\begin{aligned} u^+(x_0) &= \text{ap lim sup}_{x \rightarrow x_0} u(x) = \inf\{t \in \mathbb{R} : t \geq 0\} = 0, \\ u^-(x_0) &= \text{ap lim inf}_{x \rightarrow x_0} u(x) = \sup\{t \in \mathbb{R} : t \leq 0\} = 0. \end{aligned}$$

This shows that no point in $R^2 \setminus \overline{E}$ is part of the jump set. The only points left are those on the boundary of E . For $x_0 \in \partial E \setminus \{x^*\}$ we have

$$\begin{aligned} u^+(x_0) &= \text{ap lim sup}_{x \rightarrow x_0} u(x) = \inf\{t \in \mathbb{R} : t \geq 1\} = 1, \\ u^-(x_0) &= \text{ap lim inf}_{x \rightarrow x_0} u(x) = \sup\{t \in \mathbb{R} : t \leq 0\} = 0. \end{aligned}$$

Hence every point in $\partial E \setminus \{x^*\}$ is part of the jump set. Finally for x^* we have

$$u^+(x^*) = u^-(x^*) = \text{ap lim}_{x \rightarrow x^*} u(x) = 1$$

which shows that x^* is not part of the jump set. If we regard Figure 2.1 as an image, then we would call ∂E for an edge except maybe at the point x^* where we do not have an actual edge. Hence the jump set describes exactly what would be considered as edges in the image. This makes the $BV(\Omega)$ space good for the representation of images with sharp edges.

2.3 Minimization Problems in BV

Recovering of a function from noisy data can be formulated as a minimization problem as described in Chapter 1;

$$\inf\{F_0(u) : u \in C\}.$$

Here C is a constraint under which we want to minimize the functional $F_0 : X \rightarrow \mathbb{R}$. That is, C is either X or a subset hereof. Introducing a metric $d : X \times X \rightarrow \mathbb{R}_{\geq 0}$ on the space X we can rewrite the minimization problem as

$$\inf\{F_0(u) + kd(u, C) : u \in X, k \in \mathbb{R}_+\},$$

where

$$d(u, C) = \inf\{d(u, v) : v \in C\}.$$

Introducing $F_k(u) = F_0(u) + kd(u, C)$, the minimization problem can be written

$$\inf\{F_k(u) : u \in X, k \in \mathbb{R}_+\}.$$

Letting $k \rightarrow \infty$, the sequence of functionals $(F_k)_{k \in \mathbb{N}}$ increases to the functional $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$F(u) = \begin{cases} F_0(u), & u \in C, \\ +\infty, & \text{otherwise.} \end{cases}$$

Now F is an extended real-valued functional and the minimization of F_0 under the constraint C is equivalent to the minimization problem

$$\inf\{F(u) : u \in X\}$$

without any constraint. For minimization problems in $BV(\Omega)$ we often want to minimize a functional F_0 under the constraint $u \in W^{1,1}(\Omega)$. This leads to the minimization problem

$$\inf\{F(u) : u \in X\}$$

where

$$F(u) = \begin{cases} F_0(u), & u \in W^{1,1}(\Omega), \\ +\infty, & u \in BV(\Omega) \setminus W^{1,1}(\Omega). \end{cases}$$

In the next chapter we will analyse minimization problems in both $W^{1,p}(\Omega)$ for $1 < p < \infty$ and in $BV(\Omega)$. We will see that we actually cannot just use $W^{1,1}(\Omega)$ in order to say something about the existence and uniqueness of a minimizer and we will need the extension to $BV(\Omega)$.

3

Variational Formulation

¹In this Chapter, reconstruction problems such as denoising, deblurring, and CT are formulated as inverse problems. As described in Chapter 1 the problems then read: From noisy data u_δ , reconstruct u knowing the model

$$u_\delta = \mathcal{F}u + \delta,$$

where \mathcal{F} is a known operator and δ is the noise. To obtain a satisfactory reconstruction of u we then apply a generalized Tikhonov regularization. In fact, we restrict ourselves to the minimization of functionals of the form

$$\tau_{\alpha, u_\delta}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}(u)(x)|^2 dx + \alpha \int_{\Omega} g(x, u(x), Du(x)) dx. \quad (3.0.1)$$

Here $\Omega \subset \mathbb{R}^n$ is the domain of u . The above functional is of the form

$$F(u) = \int_{\Omega} f(x, u(x), Du(x)) dx \quad (3.0.2)$$

where $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a function. Here f is an extended real-valued function due to the possibility of minimizing F under a constraint as explained in Section 2.3. To begin with we therefore analyse the problem of minimizing F in (3.0.2). It turns out that we will need F to be defined on a reflexive space an

¹This Chapter is based on [1, pp. 3-6], [25, pp. 74-76], and [27, pp. 465-472].

we will need $x \mapsto (u(x), Du(x))$ to be Lebesgue measurable. A natural choice of solution space for the minimization problem is therefore $W^{1,p}(\Omega)$. These spaces are reflexive for $1 < p < \infty$ and if $u \in W^{1,p}(\Omega)$ then $x \mapsto (u(x), Du(x))$ is measurable. For $W^{1,1}(\Omega)$ though we will need to use the $BV(\Omega)$ space in order to analyse the problem since $W^{1,1}(\Omega)$ is not reflexive. In the following we will give results for the well-posedness of the minimization problem

$$\inf\{F(v) : v \in W^{1,p}(\Omega)\} \quad (3.0.3)$$

for different values of $1 \leq p < \infty$. For $p = 1$ we will need to extend the problem to $BV(\Omega)$ as described in Section 2.3.

According to Hadamard, the well-posedness of a problem requires the following three conditions:

- 1) *existence* of a solution,
- 2) *uniqueness* of the solution,
- 3) *stability* of the solution regarding small perturbations in the data.

In the following the two first items are treated for the problem in (3.0.3).

3.1 Existence

Before considering the minimization problem in (3.0.3) for the integral functional in (3.0.2), we first consider the minimization of a general functional $G : X \rightarrow \mathbb{R} \cup \{+\infty\}$ on a normed space $(X, \|\cdot\|)$. Then we can use the results for G to give restrictions on f in (3.0.2) in order for F to attain its minimum. To motivate the resulting theorem for general functionals, we begin with a simple example explaining what is needed to have existence of a minimizer. First note that even for a function $h : \mathbb{R} \rightarrow \mathbb{R}$ that is bounded from below we cannot be sure that it attains its minimum. Consider, for example, the function

$$h(y) = e^y, \quad y \in \mathbb{R}.$$

Then h is bounded from below by 0 but the infimum 0 is not attained for $y \in \mathbb{R}$. Hence we need to control the function values as $|y| \rightarrow +\infty$. Here we require h to be *sequentially coercive*. That is,

$$\lim_{|y| \rightarrow \infty} h(y) = +\infty.$$

Now a coercive function $h : \mathbb{R} \rightarrow \mathbb{R}$ which is bounded from below will indeed attain its minimum, but in general this is not the case. Consider the function

$$h(y) = e^y, \quad y \in (0, +\infty).$$

Then h is bounded from below and coercive. But the infimum

$$\inf\{h(y) : y \in (0, +\infty)\} = 1$$

is obtained as $y \rightarrow 0$, but 0 in the domain of definition. This suggests that we need some kind of compactness result. We require h to be *inf-compact*, i.e. we require that the level sets of h are relatively compact in the domain of definition. However, this is still not enough. Consider the function

$$h(y) = \begin{cases} e^y, & y \in (0, \infty), \\ 2, & y = 0. \end{cases}$$

Then h is bounded from below, coercive, and inf-compact. But still the infimum is not attained since

$$\inf\{h(y) : y \in [0, \infty)\} = 1$$

is obtained as $y \rightarrow 0$, but $h(0) = 2$. This suggests that we need some kind of continuity result. We do not need h to be continuous though since if we choose $h(0) = 0$ in the above instead then h attains its minimum. We therefore require h to be lower semi-continuous, i.e. for any sequence $(y_n)_{n \in \mathbb{N}}$ in the domain of h converging to y we have

$$h(y) \leq \liminf_{n \rightarrow \infty} h(y_n).$$

The considerations above are gathered in Theorem 3.1.1 below which gives the least requirements for a proper functional $G : X \rightarrow \mathbb{R} \cup \{+\infty\}$ to attain its minimum (here proper means that there exists an element $u \in X$ such that $G(u) < +\infty$):

Theorem 3.1.1. *Let $G : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous, inf-compact, and proper functional. Then G attains a minimizer in X .*

Proof. Define

$$m = \inf\{G(v) : v \in X\}.$$

We need to show that there exists an element $u \in X$ such that $G(u) = m$. Since G is proper we know that there exists at least one element in X for which the value of G is less than $+\infty$. This implies that $m < +\infty$. Also since G is inf-compact then $m > -\infty$. Together these two properties imply that there exists a minimizing sequence $(u_k)_{k \in \mathbb{N}} \subset \mathcal{D}(G)$ for G , i.e.

$$\lim_{k \rightarrow \infty} G(u_k) = m.$$

Since G is inf-compact then by definition every level sets of G are sequentially pre-compact. Since $(u_k)_{k \in \mathbb{N}}$ is a minimizing sequence then $\sup_k G(u_k) = a_1 < +\infty$ and therefore $(u_k)_{k \in \mathbb{N}}$ is part of the level set $\text{level}_{a_1}(G)$. Since every level set is pre-compact there exists a subsequence $(u_{k_n})_{n \in \mathbb{N}}$ converging to some $u \in X$. The lower semi-continuity of G then implies

$$G(u) \leq \liminf_{n \rightarrow \infty} G(u_{k_n}) = m.$$

By the definition of m we can conclude $G(u) = m$. This shows that $u \in X$ is a minimizer of G and thus G attains a minimizer in X . \square

Instead of general functionals we now consider integral functionals $F : W^{1,p}(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$, $1 \leq p < \infty$, of the form

$$F(u) = \int_{\Omega} f(x, u(x), Du(x)) \, dx, \quad (3.1.1)$$

where $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a function. In order to be sure of the existence of a minimizer we need F to be inf-compact and lower semi-continuous. In the following we give necessary and sufficient restrictions on the function f such that the integral functional F is well-defined and attains a minimizer.

We define two different integrand types. The first type is called *normal integrand* and ensures that the integral functional in (3.1.1) is well-defined:

Definition 3.1.2 (Normal Functional). Let $\Omega \subset \mathbb{R}^n$ be an open set, and let $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a functional. Then f is called *normal* if

- (i) $f(x, \cdot, \cdot)$ is lower semi-continuous for all $x \in \Omega$.
- (ii) $f(\cdot, u, \xi)$ is Borel measurable for all $(u, \xi) \in \mathbb{R} \times \mathbb{R}^n$.

Assuming that the integrand f is normal ensures that the integral functional defined in (3.1.1) is well-defined. Here we give a proof in the case when $f(x, \cdot, \cdot)$ is continuous and refer to a proof in [4] in the case when $f(x, \cdot, \cdot)$ is lower semi-continuous:

Theorem 3.1.3. *Let $\Omega \subset \mathbb{R}^n$ be open, $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a normal functional and $u : \Omega \rightarrow \mathbb{R}$ and $v : \Omega \rightarrow \mathbb{R}^n$ be Lebesgue measurable. Then the function $g : \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by*

$$g(x) = f(x, u(x), v(x))$$

is Lebesgue measurable. In particular, if $u \in W^{1,p}(\Omega)$ with $p \geq 1$ then the functional F defined in (3.1.1) is well-defined.

Proof. We prove the theorem only in the case when $f(x, \cdot, \cdot)$ is continuous. We start by proving the result for simple functions u and v . Hence let u and v be given by

$$u(x) = \sum_{i=1}^n \alpha_i \chi_{A_i}(x) \quad \text{and} \quad v(x) = \sum_{j=1}^m \beta_j \chi_{B_j}(x)$$

where $\alpha_i \in \mathbb{R}$, $\beta_j \in \mathbb{R}^n$, $A_i \subset \Omega$ are mutually disjoint, $B_j \subset \Omega$ are mutually disjoint and

$$\bigcup_{i=1}^n A_i = \Omega \quad \text{and} \quad \bigcup_{j=1}^m B_j = \Omega.$$

The function $\chi_C(x)$ for a set $C \subset \mathbb{R}^N$ is the characteristic on C . Now let $a \in \mathbb{R}$. Then

$$\{x \in \Omega : g(x) < a\} = \bigcup_{j=1}^m \bigcup_{i=1}^n \{x \in A_i \cap B_j : f(x, \alpha_i, \beta_j) < a\}.$$

Since $f(\cdot, u, \xi)$ is Borel measurable for each fixed $(u, \xi) \in \mathbb{R} \times \mathbb{R}^n$, then each of the sets

$$\{x \in A_i \cap B_j : f(x, \alpha_i, \beta_j) < a\}$$

are measurable and hence also the finite union

$$\bigcup_{j=1}^m \bigcup_{i=1}^n \{x \in A_i \cap B_j : f(x, \alpha_i, \beta_j) < a\}$$

is measurable. That is,

$$\{x \in \Omega : g(x) < a\}$$

is measurable for each $a \in \mathbb{R}$. Hence $g^{-1}(D)$ (the pre-image) is measurable for all sets D of the form $D = (-\infty, a)$, $a \in \mathbb{R}$, and since the set of all such sets is a

generator for the Borel-algebra, we conclude that g is measurable by [23, Lemma 4.7].

Now we move on to the case in which u and v are not simple functions. Since every measurable function is the pointwise limit of simple functions, we can then find sequences $(u_k)_{k \in \mathbb{N}}$ and $(v_l)_{l \in \mathbb{N}}$ such that $u_k(x) \rightarrow u(x)$ as $k \rightarrow \infty$ and $v_l(x) \rightarrow v(x)$ as $l \rightarrow \infty$ pointwise (see [23, Cor. 5.17]) By assumption $f(x, \cdot, \cdot)$ is continuous and hence

$$g(x) = f(x, u(x), v(x)) = \liminf_{k \rightarrow \infty} f(x, u_k(x), v_k(x)).$$

Since for each $k \in \mathbb{N}$ the functions u_k and v_k are simple, then by the above $f(x, u_k(x), v_k(x))$ is measurable. By [23, Lemma 5.5] also the function

$$x \mapsto \liminf_{k \rightarrow \infty} f(x, u_k(x), v_k(x))$$

is measurable. This proves the result in the case where $f(x, \cdot, \cdot)$ is continuous. For the proof when $f(x, \cdot, \cdot)$ is lower semi-continuous see [4, Corollary 2B].

In particular, for $u \in W^{1,p}(\Omega)$, $1 \leq p < \infty$, the functions $x \mapsto u(x)$ and $x \mapsto Du(x)$ are measurable and hence the function $g(x) = f(x, u(x), Du(x))$ is measurable by the above. By [23, Lemma 7.2] the integral functional in (3.1.1) is well-defined under the convention that if $|f|$ is not summable then we set $F(u) = +\infty$. \square

In order for F in (3.1.1) to attain a minimizer we need it to be lower semi-continuous in some topology. It turns out that the weak topology is favourable. In order for F to be weakly lower semi-continuous we need the integrand f to be weakly lower semi-continuous. Therefore we require $f(x, \cdot, \cdot)$ to be convex since this together with the lower semi-continuity of $f(x, \cdot, \cdot)$ implies the weak lower semi-continuity of $f(x, \cdot, \cdot)$ (see Theorem A.2.11 in Appendix A.2):

Definition 3.1.4 (Convex Integrand). Let $\Omega \subset \mathbb{R}^n$ be open. We say that $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a *convex integrand* if for every $x \in \Omega$ the function $(u, \xi) \mapsto f(x, u, \xi)$ is convex.

With the definitions above we are ready to state the following theorem about the lower semi-continuity of integral functionals of the form (3.1.1):

Theorem 3.1.5. *Let $\Omega \subset \mathbb{R}^n$ be open and let $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ be a normal and convex integrand. Then for all sequences $(u_k)_{k \in \mathbb{N}} \subset L^q(\Omega)$ norm converging to u and $(v_k)_{k \in \mathbb{N}} \subset L^r(\Omega, \mathbb{R}^n)$ weakly converging to v with $q \geq 1$ and $1 \leq r \leq p$, we have*

$$\liminf_{k \rightarrow \infty} \int_{\Omega} f(x, u_k(x), v_k(x)) \, dx \geq \int_{\Omega} f(x, u(x), v(x)) \, dx. \quad (3.1.2)$$

In particular for $1 \leq p \leq n$ and $1 \leq q < np/(n-p)$ the functional F defined in (3.1.1) is weakly sequentially lower semi-continuous in $W^{1,p}(\Omega)$. Similarly for $n < p < \infty$ and $1 \leq q < \infty$ the functional F is weakly sequentially lower semi-continuous in $W^{1,p}(\Omega)$.

Proof. By assumption $f(x, \cdot, \cdot)$ is lower semi-continuous and convex and by Theorem A.2.11 in Appendix A.2 this implies that $f(x, \cdot, \cdot)$ is weakly lower semi-continuous. Since u_k converges to u in the norm topology on $L^q(\Omega)$ then also u_k converges weakly to u in $L^q(\Omega)$ as well as v_k converges weakly to v in $L^r(\Omega, \mathbb{R}^n)$ and we obtain

$$\liminf_{k \rightarrow \infty} \int_{\Omega} f(x, u_k(x), v_k(x)) \, dx \geq \int_{\Omega} f(x, u(x), v(x)) \, dx. \quad (3.1.3)$$

Since f is normal then by Theorem 3.1.3 the functions

$$x \mapsto f(x, u_k(x), v_k(x)), \quad k \in \mathbb{N} \quad \text{and} \quad x \mapsto f(x, u(x), v(x))$$

are all measurable and therefore integrable under the convention that if the integral of the absolute value is not finite we set the integral equal to $+\infty$. Hence by Fatou's Lemma (see [23, Lemma 6.25]) and (3.1.3) we get

$$\begin{aligned} \int_{\Omega} f(x, u(x), v(x)) \, dx &\leq \int_{\Omega} \liminf_{k \rightarrow \infty} f(x, u_k(x), v_k(x)) \, dx \\ &\leq \liminf_{k \rightarrow \infty} \int_{\Omega} f(x, u_k(x), v_k(x)) \, dx. \end{aligned}$$

Regarding the last part of the theorem, let $(u_k)_{k \in \mathbb{N}} \subset W^{1,p}(\Omega)$ be a weakly convergent sequence with $u_k \rightharpoonup u \in W^{1,p}(\Omega)$ as $k \rightarrow \infty$. Then for $1 \leq p \leq n$ the mapping $i : W^{1,p}(\Omega) \rightarrow L^q(\Omega)$ defined by $i(u) = u$ is a compact imbedding for $1 \leq q < np/(n-p)$ and likewise for $n < p < \infty$ the mapping is a compact imbedding for $1 \leq q < \infty$. Hence i is continuous with respect to the weak topology on $W^{1,p}(\Omega)$ and the norm topology on $L^q(\Omega)$ so that

$$i(u_k) \rightarrow i(u) \text{ in } L^q(\Omega).$$

That is, $u_k \rightarrow u$ in $L^q(\Omega)$. Furthermore since $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$ then by the characterization of the dual space $(W^{1,p}(\Omega))'$ given in Theorem A.1.2 in Appendix

A.1 we conclude that $Du_k \rightharpoonup Du$ in $L^p(\Omega, \mathbb{R}^n)$. Hence because $(L^r(\Omega, \mathbb{R}^n))' \subseteq (L^p(\Omega, \mathbb{R}^n))'$ for all $1 \leq r \leq p$ then $Du_k \rightharpoonup Du$ in $L^r(\Omega, \mathbb{R}^n)$. So the conditions for using inequality (3.1.2) are satisfied and we obtain

$$\liminf_{k \rightarrow \infty} F(u_k) = \liminf_{k \rightarrow \infty} \int_{\Omega} f(x, u_k(x), Du_k(x)) \geq \int_{\Omega} f(x, u(x), Du(x)) = F(u).$$

This shows that F is weakly sequentially lower semi-continuous in $W^{1,p}(\Omega)$. \square

Now that we know that F is weakly sequentially lower semi-continuous, we just have to prove that F is weakly inf-compact in order to use Theorem 3.1.1 to prove the existence of a minimizer. In order to show that F is weakly inf-compact we need a lower bound condition on the integrand f which controls the growth of F as $\|u\|_{W^{1,p}(\Omega)} \rightarrow \infty$:

Corollary 3.1.6. *Let $1 < p < \infty$. Let $\Omega \subset \mathbb{R}^n$ be open, bounded and connected and assume that $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ is a normal and convex integrand satisfying the lower bound condition*

$$f(x, u, \xi) \geq c_1 + c_2(|u|^p + |\xi|^p)$$

for some $c_1 \in \mathbb{R}$ and $c_2 > 0$. Let $F : W^{1,p}(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ be defined as in (3.1.1), and let $X \subseteq W^{1,p}(\Omega)$ be a closed and convex subset. If there exists $u \in X$ with $F(u) < +\infty$, then $F|_X$ attains a minimizer in X .

Proof. In order to be able to use Theorem 3.1.1 we need to show that F is lower semi-continuous and inf-compact in some topology on X . Since we can use Theorem 3.1.5 to conclude that F is weakly sequentially lower semi-continuous in $W^{1,p}(\Omega)$, we show that it is weakly inf-compact. That is, we need to show that every level set $\text{level}_a(F|_X)$ is weakly sequentially pre-compact in X . From the lower bound assumption it follows that

$$F(u) \geq c_1 \mathcal{L}^n(\Omega) + c_2 \int_{\Omega} (|u(x)|^p + |Du(x)|^p) dx = c_1 \mathcal{L}^n(\Omega) + c_2 \|u\|_{W^{1,p}(\Omega)}^p.$$

Hence for $a \in \mathbb{R}$ the inequality

$$F(u) \leq a$$

implies

$$c_1 \mathcal{L}^n(\Omega) + c_2 \|u\|_{W^{1,p}(\Omega)}^p \leq a.$$

By the definition of the level sets we then have

$$\text{level}_a(F|_X) \subset X \cap \left\{ u \in W^{1,p}(\Omega) : \|u\|_{W^{1,p}(\Omega)}^p \leq \frac{a - c_1 \mathcal{L}^n(\Omega)}{c_2} \right\}.$$

Since X is a convex and closed subset of $W^{1,p}(\Omega)$ it is also weakly closed, see [20, theorem 3.3.2]. On the other hand for $p > 1$, $W^{1,p}(\Omega)$ is a reflexive and separable Banach space and hence the unit ball in $W^{1,p}(\Omega)$ is weakly sequentially compact, see [20, theorem 2.4.2]. Since the set

$$\left\{ u \in W^{1,p}(\Omega) : \|u\|_{W^{1,p}(\Omega)}^p \leq \frac{a - c_1 \mathcal{L}^n(\Omega)}{c_2} \right\}$$

is simply a scaled unit ball in $W^{1,p}(\Omega)$ it is also weakly sequentially compact. Hence the level set $\text{level}_a(F|_X)$ is contained in the intersection between a weakly closed set and a weakly sequentially compact set. This implies that the level set itself is weakly sequentially pre-compact and hence F is weakly inf-compact.

Since f is a non-negative normal and convex integrand, then f satisfies the assumptions of Theorem 3.1.5. Hence it follows that F is weakly sequentially lower semi-continuous. By assumption $F|_X$ is also proper and then by Theorem 3.1.1 there exists a minimizer of $F|_X$ in X . \square

In the proof of Theorem 3.1.6 we needed the space $W^{1,p}(\Omega)$ to be reflexive. Since the space $W^{1,1}(\Omega)$ is not reflexive, Theorem 3.1.6 does not hold for $p = 1$. In order to extend the theorem to $p = 1$ we need to extend the functional $F : W^{1,1}(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ defined in (3.1.1) to the Banach space $BV(\Omega) \supset W^{1,1}(\Omega)$ in which bounded sets are weakly sequentially pre-compact by Lemma 2.1.5. Note also that the space $BV(\Omega)$ allows discontinuities across edges in images which is actually a good a priori model.

The extension of F to $BV(\Omega)$ is carried out as described in Section 2.3. Since we want to minimize F in $BV(\Omega)$ under the constraint $u \in W^{1,1}(\Omega)$, F is extended in the following way:

$$\bar{F}(u) = \begin{cases} F(u), & u \in W^{1,1}(\Omega), \\ +\infty, & u \in BV(\Omega) \setminus W^{1,1}(\Omega). \end{cases} \quad (3.1.4)$$

Since Theorem 3.1.5 holds for $p = 1$ we can still conclude that $F(u)$ is weakly sequentially lower semi-continuous in $W^{1,1}(\Omega)$ and we reach the final theorem for $p = 1$:

Corollary 3.1.7. *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, connected and Lipschitz, and assume that $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ is a normal and convex integrand satisfying the lower bound condition*

$$f(x, u, \xi) \geq c_1 + c_2(|u| + |\xi|)$$

for some $c_1 \in \mathbb{R}$ and $c_2 > 0$. Let $\bar{F} : BV(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ be defined as in (3.1.4), and let $X \subseteq BV(\Omega)$ be a closed and convex subset. If there exists $u \in X$ with $\bar{F}(u) < +\infty$, then $\bar{F}|_X$ attains a minimizer in X .

Proof. Since \bar{F} is proper then any minimizing sequence will eventually enter $W^{1,1}(\Omega)$ on which \bar{F} is definitely lower semi-continuous by Theorem 3.1.5 which applies to $p = 1$. As in the proof of Corollary 3.1.6 we then want to show that the level sets $\text{level}_a(\bar{F}|_X)$ are weakly sequentially pre-compact in X . But since every minimizing sequence eventually enters $W^{1,1}(\Omega)$ we just need to show that the level sets $\text{level}_a(\bar{F}|_{X \cap W^{1,1}(\Omega)})$ are weakly sequentially pre-compact in X . From the lower bound assumption we obtain

$$\text{level}_a(\bar{F}|_{X \cap W^{1,1}(\Omega)}) \subset X \cap \left\{ u \in BV(\Omega) : \|u\|_{BV(\Omega)} \leq \frac{a - c_1 \mathcal{L}^n(\Omega)}{c_2} \right\}.$$

Again since X is a convex and closed subset of $BV(\Omega)$, it is weakly closed. Now since the set

$$\left\{ u \in BV(\Omega) : \|u\|_{BV(\Omega)} \leq \frac{a - c_1 \mathcal{L}^n(\Omega)}{c_2} \right\}$$

is a closed ball in $BV(\Omega)$ it is bounded and then by Lemma 2.1.5 it is weakly sequentially pre-compact. Hence the level set $\text{level}_a(\bar{F}|_{X \cap W^{1,1}(\Omega)})$ is contained in the intersection between a weakly closed set and a weakly sequentially pre-compact set. This implies that the level set itself is weakly sequentially pre-compact and hence \bar{F} is weakly inf-compact on $W^{1,1}(\Omega)$.

By Theorem 3.1.1 there exists a minimizer of $\bar{F}|_{X \cap W^{1,1}(\Omega)}$ in $X \cap W^{1,1}(\Omega)$. This implies that there exists a minimizer of $\bar{F}|_X$ in X . \square

3.2 Uniqueness

For the uniqueness of solutions to the minimization problem we need a slightly more strict condition for the integrand f . We need it to be *strictly* convex instead of just convex in the last two components. This is reasonable since considering the

generalized Tikhonov functional in (3.0.1) we have

$$f(x, u(x), Du(x)) = \frac{1}{2}|u_\delta(x) - \mathcal{F}(u)(x)|^2 + \alpha g(x, u(x), Du(x)).$$

If we want to penalise for solutions having large L^2 -norm, i.e. we want a smooth solution, we will choose $g(x, u(x), Du(x)) = \frac{1}{2}|u(x)|^2$. If we want to penalise for solutions having large oscillations we will choose $g(x, u(x), Du(x)) = \frac{1}{2}|Du(x)|^2$. These are two often used regularization methods and since $|\cdot|^2$ is strictly convex then f would be strictly convex if the operator \mathcal{F} is linear and bounded. Requiring that f is strictly convex we obtain the following uniqueness result:

Corollary 3.2.1. *Let $\Omega \subset \mathbb{R}^n$ be bounded, open, connected, and Lipschitz and assume that $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ is a normal integrand and that $(u, \xi) \mapsto f(x, u, \xi)$ is strictly convex. Assume further that f satisfies the lower bound condition*

$$f(x, u, \xi) \geq c_1 + c_2(|u|^p + |\xi|^p)$$

for some $c_1 \in \mathbb{R}$ and $c_2 > 0$. Let $F : W^{1,p}(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ and $F : BV(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ be defined as in (3.1.1) and (3.1.4), respectively, and let $X \subseteq W^{1,p}(\Omega)$ respectively $X \subseteq BV(\Omega)$ be closed and convex. If there exists $u \in X$ with $F(u) < +\infty$, then there exists a unique minimizer of $F|_X$ in X .

Proof. The existence follows from Corollaries 3.1.6 and 3.1.7, respectively. Regarding the proof of uniqueness, assume that u and v are two minimizers of F in X with $u \neq v$. Observe that for the BV -functional we can have neither $u \in BV(\Omega) \setminus W^{1,1}(\Omega)$ nor $v \in BV(\Omega) \setminus W^{1,1}(\Omega)$ since then u and v could not be minimizers.

Since u and v are both minimizers of F then $F(u) = F(v)$ and by the strict convexity of $(u, \xi) \mapsto f(x, u, \xi)$ it follows that

$$F\left(\frac{1}{2}u + \left(1 - \frac{1}{2}\right)v\right) < \frac{1}{2}F(u) + \frac{1}{2}F(v) = \frac{1}{2}(F(u) + F(v)) = F(u) = F(v)$$

This leads to a contradiction, since u and v are minimizers and hence

$$F(u) = F(v) \leq F\left(\frac{1}{2}u + \frac{1}{2}v\right).$$

Hence the assumption that $u \neq v$ must be wrong and we conclude $u = v$. \square

3.3 Stability

Regarding the stability there is a result in [18, Thm. 3.23]. This is outside the scope of this thesis and will not be discussed further.

3.4 Examples

In this section we give three examples of generalized Tikhonov regularization methods and show that all of them satisfy the conditions for the existence of a (unique) minimizer to the corresponding optimization problems. The first method is the standard Tikhonov regularization which penalises for functions having large L^2 -norm. This is a smoothing regularization method. We can as well use the L^2 -norm of the distributional derivative instead, in order to penalize for solutions having large oscillations. In the second example, we consider the total variation regularization which penalises for big total variations. This method reduces unwanted details in the data while preserving important details as for example jumps seen as edges in an image. This will also be seen from numerical experiments in Chapter 7. In the third and final example we consider a slightly changed Tikhonov functional penalising for solutions having large oscillations.

3.4.1 Standard Tikhonov Regularization

Let $\Omega \subset \mathbb{R}^n$ be bounded, open, connected, and Lipschitz and let $\mathcal{F} : L^2(\Omega) \rightarrow L^2(\Omega)$ be linear, bounded, and injective. \mathcal{F} could for example be a convolution operator or the Radon transform in the case of computed tomography data. The standard Tikhonov functional $\tau_{\alpha, u_\delta}^{\text{Tikh}} : L^2(\Omega) \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ is defined by

$$\tau_{\alpha, u_\delta}^{\text{Tikh}}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}u(x)|^2 dx + \alpha \frac{1}{2} \int_{\Omega} |u(x)|^2 dx.$$

We define $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ by

$$f(x, u(t)) = \frac{1}{2} |u_\delta(t) - \mathcal{F}u(t)|^2 + \alpha \frac{1}{2} |u(t)|^2.$$

By the continuity of \mathcal{F} and $|\cdot|$, $f(x, \cdot)$ is continuous and in particular lower semi-continuous. Furthermore $f(\cdot, u(t))$ is a Borel function since it is actually constant for fixed $t \in \Omega$ and hence f is a normal integrand.

We show that $f(x, \cdot)$ is strictly convex. Hence let $0 < \lambda < 1$ and $u_1, u_2 \in L^2(\Omega)$

with $u_1 \neq u_2$. Then

$$\begin{aligned}
f(x, \lambda u_1(t) + (1 - \lambda)u_2(t)) &= \frac{1}{2}|u_\delta(t) - \mathcal{F}(\lambda u_1(t) + (1 - \lambda)u_2(t))|^2 \\
&\quad + \alpha \frac{1}{2}|\lambda u_1(t) + (1 - \lambda)u_2(t)|^2 \\
&= \frac{1}{2}|\lambda(u_\delta(t) - \mathcal{F}u_1(t)) + (1 - \lambda)(u_\delta(t) - \mathcal{F}u_2(t))|^2 \\
&\quad + \alpha \frac{1}{2}|\lambda u_1(t) + (1 - \lambda)u_2(t)|^2. \tag{3.4.1}
\end{aligned}$$

Using that $|\cdot|^2$ is strictly convex, $u_1 \neq u_2$, and $\mathcal{F}u_1 \neq \mathcal{F}u_2$ since \mathcal{F} is injective, we obtain

$$\begin{aligned}
f(x, \lambda u_1(t) + (1 - \lambda)u_2(t)) &< \lambda \frac{1}{2}|u_\delta(t) - \mathcal{F}u_1(t)|^2 + (1 - \lambda) \frac{1}{2}|u_\delta(t) - \mathcal{F}u_2(t)|^2 \\
&\quad + \lambda \alpha \frac{1}{2}|u_1(t)|^2 + (1 - \lambda) \alpha \frac{1}{2}|u_2(t)|^2 \\
&= \lambda f(x, u_1(t)) + (1 - \lambda) f(x, u_2(t)).
\end{aligned}$$

This shows that $f(x, \cdot)$ is strictly convex. Finally, f satisfies the lower bound condition since

$$f(x, u(t)) = \frac{1}{2}|u_\delta(t) - \mathcal{F}u(t)|^2 + \alpha \frac{1}{2}|u(t)|^2 \geq \alpha \frac{1}{2}|u(t)|^2.$$

If we leave out the dependence on the third variable in the proofs of Theorem 3.1.5, Corollary 3.1.6 and Corollary 3.2.1, and use the space $L^2(\Omega)$ which is a Hilbert space and therefore reflexive, we can conclude that the standard Tikhonov functional attains a minimizer on $X \subseteq L^2(\Omega)$ whenever X is closed and convex and $u_\delta \in L^2(\Omega)$.

If we consider instead the functional $\tau_{\alpha, u_\delta}^{\text{Tikh}2} : W^{1,2}(\Omega) \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ defined by

$$\tau_{\alpha, u_\delta}^{\text{Tikh}2}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}u(x)|^2 \, dx + \alpha \frac{1}{2} \int_{\Omega} |Du(x)|^2 \, dx,$$

then we can use the same argumentation as above to conclude that the function $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ defined by

$$f(x, u(x), Du(x)) = \frac{1}{2}|u_\delta(x) - \mathcal{F}u(x)|^2 + \alpha \frac{1}{2}|Du(x)|^2$$

is a normal integrand and that it is strictly convex. Furthermore we see that f satisfies the lower bound condition

$$f(x, u(x), Du(x)) = \frac{1}{2}|u_\delta(x) - \mathcal{F}u(x)|^2 + \alpha \frac{1}{2}|Du(x)|^2 \geq \frac{1}{2} \alpha |Du(x)|^2.$$

In the proof of Corollary 3.1.6 we can use Poincaré's inequality

$$\|u - (u)_\Omega\|_{L^2(\Omega)} \leq C\|Du\|_{L^2(\Omega)}, \quad (u)_\Omega = \int_{\Omega} u(y) \, dy = \frac{1}{\mathcal{L}^n(\Omega)} \int_{\Omega} u(y) \, dy$$

to obtain

$$\|u\|_{L^2(\Omega)} \leq c_1 + C\|Du\|_{L^2(\Omega)}.$$

Here $c_1 \in \mathbb{R}$ stems from the average constant $(u)_\Omega$ that we subtract from u . Then

$$\|u\|_{W^{1,2}(\Omega)} = \|u\|_{L^2(\Omega)} + \|Du\|_{L^2(\Omega)} \leq c_1 + (1 + C)\|Du\|_{L^2(\Omega)}$$

for $c_1 \in \mathbb{R}$ and $C > 0$ and then we see that the above lower bound condition for f is enough to ensure the existence of a unique minimizer in $X \subseteq W^{1,2}(\Omega)$ whenever X is closed and convex and $u_\delta \in L^2(\Omega)$.

3.4.2 Total Variation Regularization

Let $\Omega \subset \mathbb{R}^n$ be bounded, open, connected and Lipschitz and let $\mathcal{F} : BV(\Omega) \rightarrow L^2(\Omega)$ be linear, bounded, and injective. For example the embedding $I : BV(\Omega) \rightarrow L^p(\Omega)$ is continuous for $1 \leq p \leq n/(n-1)$. So in two dimensions for example $I : BV(\Omega) \rightarrow L^2(\Omega)$ is bounded. The generalized Tikhonov functional $\tau_{\alpha, u_\delta}^{\text{TV}} : BV(\Omega) \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ with a total variation regularization penalization term is defined by

$$\tau_{\alpha, u_\delta}^{\text{TV}}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}u(x)|^2 \, dx + \alpha|Du|(\Omega).$$

We see that the functional is not of the form (3.0.1) and we cannot define our integrand f as in the standard Tikhonov functional example. Instead we show directly that $\tau_{\alpha, u_\delta}^{\text{TV}}$ satisfies the conditions of Theorem 3.1.1. By Proposition 2.1.4 in Chapter 2, we know that $|Du|(\Omega)$ is lower semi continuous. Using the definition of the total mass $|Du|(\Omega)$, we can show that it is also convex: For $\lambda \in [0, 1]$ and $u, v \in BV(\Omega)$ we have

$$\begin{aligned} & |D(\lambda u + (1 - \lambda)v)|(\Omega) \\ &= \sup \left\{ \int_{\Omega} (\lambda u + (1 - \lambda)v) \operatorname{div} \phi \, dx, \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\} \\ &= \lambda \sup \left\{ \int_{\Omega} u \operatorname{div} \phi \, dx, \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\} \\ &\quad + (1 - \lambda) \sup \left\{ \int_{\Omega} v \operatorname{div} \phi \, dx, \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\} \\ &= \lambda|Du|(\Omega) + (1 - \lambda)|Dv|(\Omega). \end{aligned}$$

Since $|Du|(\Omega)$ is both lower semi-continuous and convex then by Theorem A.2.11 in Appendix A.2, $|Du|(\Omega)$ is weakly lower semi-continuous.

We show directly that $\tau_{\alpha, u_\delta}^{\text{TV}}$ is inf-compact. Using an extended Poincaré inequality for $BV(\Omega)$ (see [17, Thm. 3.2]) we get for $u \in BV(\Omega)$:

$$\|u - (u)_\Omega\|_{L^1(\Omega)} \leq C|Du|(\Omega), \quad (u)_\Omega = \int_{\Omega} u(y) \, dy = \frac{1}{\mathcal{L}^n(\Omega)} \int_{\Omega} u(y) \, dy.$$

From this we obtain

$$\|u\|_{L^1(\Omega)} \leq c_1 + C|Du|(\Omega),$$

where $c_1 \in \mathbb{R}$ stems from the average constant $(u)_\Omega$ that we subtract from u . This implies

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega) \leq c_1 + (1 + C)|Du|(\Omega).$$

Since for $a \in \mathbb{R}$

$$a \geq \tau_{\alpha, u_\delta}^{\text{TV}}(u) \geq \alpha|Du|(\Omega) \geq \alpha \frac{\|u\|_{BV(\Omega)} - c_1}{1 + C}$$

implies

$$\|u\|_{BV(\Omega)} \leq \alpha^{-1}(1 + C)a + c_1$$

we can use this to show that $\tau_{\alpha, u_\delta}^{\text{TV}}$ is weakly inf-compact as in the proof of corollary 3.1.7. Since $\tau_{\alpha, u_\delta}^{\text{TV}}$ is also proper (take for example $u = 0$) then by Theorem 3.1.1 the total variation functional attains a minimizer in a closed and convex subset $X \subseteq BV(\Omega)$ whenever $u_\delta \in L^2(\Omega)$.

3.4.3 Non-Convex Regularization

This final example deals with a slightly changed generalized Tikhonov functional used when the noise model is a sampling error model, see [18, pp. 47-48]. Let $\Omega \subset \mathbb{R}^n$ be bounded, open, connected, and Lipschitz and let $\mathcal{F} : W^{1,2}(\Omega) \rightarrow L^2(\Omega)$ be a linear, bounded, and injective operator. Then we want to minimize the functional $\tau_{\alpha, u_\delta}^{\text{NC}} : W^{1,2}(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\tau_{\alpha, u_\delta}^{\text{NC}}(u) = \int_{\Omega} \frac{|u_\delta(x) - \mathcal{F}u(x)|^2}{2|Du(x)|^2} \, dx + \alpha \frac{1}{2} \int_{\Omega} |Du(x)|^2 \, dx.$$

We observe that this functional is the standard Tikhonov functional where we regularize for large oscillations except that we divide by the length of the squared weak derivative of u in the fidelity term. Why would we want to do this? The

fidelity term stems from a different discrete noise model (see [18, pp. 47-49]) called *sampling errors*. Intuitively we see that when $|Du(x)|$ is much larger than $|u_\delta(x) - \mathcal{F}u(x)|$ we see that the fidelity term becomes small. This leaves the regularizing term as the leading term, when minimizing the functional. On the other hand, when $|Du(x)|$ is much smaller than $|u_\delta(x) - \mathcal{F}u(x)|$, the fidelity term becomes the leading term when minimizing $\tau_{\alpha, u_\delta}^{NC}$. Hence the change in the fidelity term ensures that we do not make any new jumps in $\mathcal{F}u$ that are not already in u_δ and the regularizing term still ensures that u does not have large oscillations.

We define the function $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$f(x, u(x), Du(x)) = \frac{|u_\delta(x) - \mathcal{F}u(x)|^2}{2|Du(x)|^2} + \alpha \frac{1}{2}|Du(x)|^2.$$

The first thing we notice is that f is not defined for $|Du(x)| = 0$. Hence we need to define the value in this case. The next thing we notice is that f is not convex in its third variable. Hence we cannot use the results in Corollaries 3.1.6 and 3.2.1. Instead we use a convexification of f , that is, the convex hull of f . We will not go into detail about what the convex hull of a function is, but just state that it is a convex function that somehow relates to f itself. The convexification of f is given by (see [18, lemma 5.12])

$$f_c(x, u(x), Du(x)) = \begin{cases} \frac{|u_\delta(x) - \mathcal{F}u(x)|^2}{2|Du(x)|^2} + \alpha \frac{1}{2}|Du(x)|^2, & \sqrt{\alpha}|Du(x)|^2 > |u_\delta(x) - \mathcal{F}u(x)|, \\ \sqrt{\alpha}|u_\delta(x) - \mathcal{F}u(x)|, & \sqrt{\alpha}|Du(x)|^2 \leq |u_\delta(x) - \mathcal{F}u(x)|. \end{cases}$$

Instead of integrating over f and minimizing the functional $\tau_{\alpha, u_\delta}^{NC}$ we integrate over f_c and minimize the approximated functional

$$\tau_{\alpha, u_\delta}^{NC, \text{app}}(u) = \int_{\Omega} f_c(x, u(x), Du(x)) \, dx.$$

We see that f_c is well-defined also when $|Du(x)| = 0$. It turns out that the function $f_c(x, \cdot, \cdot)$ is convex and continuously differentiable for almost every $x \in \Omega$, see [18, lemma 5.13]. This implies that f_c is a normal integrand almost everywhere which is sufficient, see [18, Chap. 5]. Observing that

$$f_c(x, u(x), Du(x)) \geq \alpha \frac{1}{2}|Du(x)|^2,$$

we see that f_c satisfies the conditions of Corollary 3.1.6, if we in the proof use that

$$\|u\|_{W^{1,2}(\Omega)} \leq c_1 + (C + 1)\|Du\|_{L^2(\Omega)}$$

for some constant $C > 0$ (as shown for the standard Tikhonov regularization). Hence the approximated functional $\tau_{\alpha, u_\delta}^{NC, \text{app}}$ attains a minimizer in $X \subset W^{1,2}(\Omega)$

whenever X is a closed and convex subset. The main result here is that every minimizer of the approximated functional $\tau_{\alpha, u_\delta}^{NC_{\text{app}}}$ is the limit of a minimizing sequence for the original functional $\tau_{\alpha, u_\delta}^{NC}$ (see Theorem A.2.16 in Appendix A.2). This suggests that instead of minimizing $\tau_{\alpha, u_\delta}^{NC}$, we minimize $\tau_{\alpha, u_\delta}^{NC_{\text{app}}}$.

4

An associated Diffusion Filtering Method

¹In this section we consider the optimization problem of minimizing the generalized Tikhonov functional $\tau_{\alpha, u_\delta} : X \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\tau_{\alpha, u_\delta}(u) = \frac{1}{2} \int_{\Omega} |u_\delta(x) - \mathcal{F}(u)(x)|^2 dx + \alpha \int_{\Omega} g(x, u(x), Du(x)) dx, \quad (4.0.1)$$

where $\mathcal{F} : X \rightarrow \mathbb{L}^2(\Omega)$ is an operator. The space X is either $W^{1,p}(\Omega)$ for $1 < p < \infty$ or $BV(\Omega)$. Assuming that there exists a minimizer of τ_{α, u_δ} , we can show that the minimization problem is associated with an optimality condition equation for u . In special cases the optimality condition can be turned into a non-linear diffusion problem for u . Solving this problem using a finite difference method corresponds to iteratively regularizing the reconstruction problem and solving minimization problems. In the first section below we introduce derivatives of operators, which will be the main ingredient in the formulation of the optimality condition.

4.1 Derivatives of Operators and the Subdifferential

In order to turn the minimization problem for the functional in (4.0.1) into a optimality condition for a minimizer u_α , we need to be able to take the derivative

¹This Chapter is based on [5], [18, Chap. 10], and [29, Chap. 4].

of τ_{α, u_δ} in some way. In this section we define different notions of derivatives of operators. First we define the *one-sided directional derivative* of an operator:

Definition 4.1.1 (One-sided Directional Derivative). Let $F : U \rightarrow V$ be an operator between normed spaces U and V . Then F admits a *one-sided directional derivative* $F'(u; h) \in V$ at $u \in U$ in direction $h \in U$ if

$$F'(u; h) = \lim_{t \searrow 0} \frac{F(u + th) - F(u)}{t}. \quad (4.1.1)$$

If $F : U \rightarrow \mathbb{R} \cup \{+\infty\}$ is an extended real-valued functional then we extend the definition of the one-sided directional derivative by replacing the limit above by a limit superior. Then the directional derivative exists for every $u \in \mathcal{D}(F)$ but may take the values $\pm\infty$. Next we define the *Gâteaux derivative* and the *Fréchet derivative* of an operator under the condition that the one-sided directional derivative exists for all elements in the domain:

Definition 4.1.2 (Gâteaux and Fréchet Derivatives). Let $F : U \rightarrow V$ be an operator between normed spaces U and V and assume that the one-sided directional derivative $F'(u; h)$ exists for $u \in U$ and for all $h \in U$. If there exists a linear and bounded operator denoted by $F'(u) \in B(U, V)$ such that

$$F'(u; h) = (h, F'(u)) = F'(u)(h), \quad \forall h \in U,$$

then F is *Gâteaux differentiable* at u and $F'(u)$ is called the *Gâteaux derivative* of F at u . If further the convergence in (4.1.1) is uniform with respect to $h \in B_\rho(0)$ for some $\rho > 0$ then F is *Fréchet differentiable* at u and $F'(u)$ is called the *Fréchet derivative* of F at u .

Remark 4.1.3. Note that the Gâteaux derivative do not need to be linear and bounded, but for the sake of convenience we use this definition here. Note also that the definition of the Fréchet derivative above is not the usual definition, but we do not need the usual definition here.

It is worth mentioning that if U , V , and W are Banach spaces and $F : U \rightarrow V$ and $G : V \rightarrow W$ are two Fréchet differentiable operators, then the composition $G \circ F : U \rightarrow W$ is Fréchet differentiable and

$$(G \circ F)'(u) = G'(F(u)) \circ F'(u). \quad (4.1.2)$$

Finally we define the subdifferential of a proper convex functional on a normed space X . The subdifferential replaces the Fréchet derivative of a functional and is defined by

Definition 4.1.4 (Subdifferential). Let $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper and convex functional on a normed space X . Then the *subdifferential* $\partial\phi(v)$ of ϕ at $v \in X$ is defined by

$$\partial\phi(v) = \{v' \in X' : \phi(v) \leq \phi(w) + (v - w, v'), \forall w \in X\}.$$

Furthermore an $v' \in \partial\phi(v)$ is called a *subgradient* of ϕ at v .

If $v' \in \partial\phi(v)$ is a subgradient of ϕ at v we see that

$$\phi(v + th) - \phi(v) \leq (v + th - v, v') = t(h, v'), \forall h \in X, \forall t > 0.$$

Assuming that ϕ is Gâteaux differentiable, we can divide by t and take the limit as $t \searrow 0$ in order to obtain

$$\phi'(v)h = \lim_{t \searrow 0} \frac{\phi(v + th) - \phi(v)}{t} \leq (h, v'), \forall h \in X.$$

where $\phi'(v)$ is the Gâteaux derivative of ϕ at v . Using $-h$ instead of h we obtain

$$\phi'(v)h \geq (h, v'), \forall h \in X.$$

In total we have

$$\phi'(v)h = (h, v'), \forall h \in X.$$

This shows that if ϕ is Gâteaux differentiable at v with Gâteaux derivative $\phi'(v)$ and $v' \in \partial\phi(v)$ then $v' = \phi'(v)$ in distributional sense so that the subdifferential contains only one element which is equal to the Gâteaux derivative.

In the following we define the *duality mapping* associated with a *weight function*. The duality can be used to express the subdifferential of integral functionals as we will see:

Definition 4.1.5 (Weight Function and associated Duality Mapping). Let U be a Banach space. A continuous and strictly increasing function $\phi : [0, \infty) \rightarrow [0, \infty)$ satisfying $\phi(0) = 0$ and $\lim_{t \rightarrow \infty} \phi(t) = \infty$ is called a *weight function*. The *duality mapping* according to the weight function ϕ is the set-valued mapping $\mathcal{J} : U \rightarrow 2^{U'}$ defined by

$$\mathcal{J}(u) = \{u' \in U' : (u, u')_{U, U'} = \|u\|_U \|u'\|_{U'}, \|U'\|_{U'} = \phi(\|U\|_U)\}.$$

The next theorem relates the subdifferential of the primitive of a weight function to the duality mapping according to the same weight function:

Theorem 4.1.6 (Asplund's Theorem). *Let U be a Banach space and $\phi : [0, \infty) \rightarrow [0, \infty)$ a weight function. Set $\Phi(t) = \int_0^t \phi(s) ds$. Then*

$$\mathcal{J}(u) = \partial\Phi(\|u\|_U), \quad u \in U,$$

where $\partial\Phi$ denotes the subdifferential of Φ .

Proof. See [26, Chap. 1, Thm. 4.4] □

Asplund's theorem will be important in the formulation of the diffusion filtering methods in the next sections. It will help us define the problems such that they make sense even in there strong formulation.

4.2 Optimality Condition and Diffusion Filtering

We turn our attention to the minimization of τ_{α, u_δ} defined in (4.0.1) on $X = W^{1,p}(\Omega)$, $1 < p < \infty$ or $X = BV(\Omega)$. Assume that there exists a minimizer u_α of τ_{α, u_δ} . Then

$$\tau_{\alpha, u_\delta}(u_\alpha) \leq \tau_{\alpha, u_\delta}(u_\alpha + th), \quad \forall h \in X.$$

This implies

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} (u_\delta - \mathcal{F}(u_\alpha + th))^2 - (u_\delta - \mathcal{F}(u_\alpha))^2 \, dx \\ & + \alpha \int_{\Omega} (g(x, u_\alpha + th, Du_\alpha + tDh) - g(x, u_\alpha, Du_\alpha)) \, dx \geq 0. \end{aligned} \quad (4.2.1)$$

Note that in this section we skip the dependence on x in for example $u(x)$. If $g(x, u, \xi)$ is twice differentiable in both u and ξ , then we have the Taylor series expansion

$$\begin{aligned} & g(x, u_\alpha + th, Du_\alpha + tDh) \\ & = g(x, u_\alpha, Du_\alpha) + g_u(x, u_\alpha, Du_\alpha) \cdot th + \nabla_\xi g(x, u_\alpha, Du_\alpha) \cdot tDh + O(t^2). \end{aligned}$$

Similarly if we assume that \mathcal{F} is twice Fréchet differentiable then we have the Taylor series expansion

$$(\mathcal{F}(u_\alpha + th) - u_\delta)^2 = (\mathcal{F}(u_\alpha) - u_\delta)^2 + 2t(\mathcal{F}'(u_\alpha)h) \cdot (\mathcal{F}(u_\alpha) - u_\delta) + O(t^2).$$

Using the Taylor series expansions and (4.2.1) we obtain

$$0 \leq \frac{1}{2} \int_{\Omega} (2t(\mathcal{F}'(u_{\alpha})h) \cdot (\mathcal{F}(u_{\alpha}) - u_{\delta}) + O(t^2)) \, dx \\ + \alpha \int_{\Omega} (g_u(x, u_{\alpha}, Du_{\alpha}) \cdot th + \nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha}) \cdot tDh + O(t^2)) \, dx.$$

Dividing by t and letting $t \searrow 0$ we get

$$0 \leq \int_{\Omega} ((\mathcal{F}'(u_{\alpha})h) \cdot (\mathcal{F}(u_{\alpha}) - u_{\delta})) \, dx \\ + \alpha \int_{\Omega} (g_u(x, u_{\alpha}, Du_{\alpha}) \cdot h + \nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha}) \cdot Dh) \, dx.$$

Repeating the above with $-t$ instead gives

$$0 \geq \int_{\Omega} ((\mathcal{F}'(u_{\alpha})h) \cdot (\mathcal{F}(u_{\alpha}) - u_{\delta})) \, dx \\ + \alpha \int_{\Omega} (g_u(x, u_{\alpha}, Du_{\alpha}) \cdot h + \nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha}) \cdot Dh) \, dx.$$

In total we therefore obtain

$$0 = \int_{\Omega} ((\mathcal{F}'(u_{\alpha})h) \cdot (\mathcal{F}(u_{\alpha}) - u_{\delta})) \, dx \\ + \alpha \int_{\Omega} (g_u(x, u_{\alpha}, Du_{\alpha}) \cdot h + \nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha}) \cdot Dh) \, dx.$$

Using integration by parts we finally get

$$0 = \int_{\Omega} \mathcal{F}'(u_{\alpha})^{\#}(\mathcal{F}(u_{\alpha}) - u_{\delta}) \cdot h \, dx + \alpha \int_{\Omega} h \cdot g_u(x, u_{\alpha}, Du_{\alpha}) \, dx \\ - \alpha \int_{\Omega} \nabla \cdot (\nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha})) \cdot h + \alpha \int_{\partial\Omega} h \cdot \nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha}) \cdot \nu, \quad \forall h \in X.$$

Here ν is the outward pointing unit normal vector to $\partial\Omega$ and $\mathcal{F}'(u_{\alpha})^{\#} : L^2(\Omega) \rightarrow X'$ denotes the dual adjoint of $\mathcal{F}'(u_{\alpha})$, i.e.

$$(\omega, \mathcal{F}'(u_{\alpha})^{\#}(v)) = (\mathcal{F}'(u_{\alpha})(\omega), v), \quad \forall v \in L^2(\Omega), \forall \omega \in X. \quad (4.2.2)$$

The above shows that we obtain a optimality condition for the minimizer u_{α} of $\tau_{\alpha, u_{\delta}}$ given by

$$\mathcal{F}'(u_{\alpha})^{\#}(\mathcal{F}(u_{\alpha}) - u_{\delta}) = \alpha (\nabla \cdot (\nabla_{\xi}g(x, u_{\alpha}, Du_{\alpha})) - g_u(x, u_{\alpha}, Du_{\alpha})) \quad \text{in } \Omega, \quad (4.2.3)$$

with boundary condition

$$\frac{\partial \nabla_{\xi} g(x, u_{\alpha}, Du_{\alpha})}{\partial \nu} = 0 \quad \text{on } \partial \Omega. \quad (4.2.4)$$

Observe that the optimality condition and the boundary condition is to be understood in a distributional sense. However, if we consider the optimality condition in its written form, we see that on the left hand side we obtain an element in the dual space X' . Hence the operator on the right hand side should be interpreted as an element in X' too in order to make sense to the PDE in its written form.

In the case of noise-free attainable data $u_{\delta} = u_0 = \mathcal{F}(u^{\dagger})$ the optimality condition reads as

$$\mathcal{F}'(u_{\alpha})^{\#} \left(\frac{\mathcal{F}(u_{\alpha}) - \mathcal{F}(u^{\dagger})}{\alpha} \right) = \nabla \cdot (\nabla_{\xi} g(x, u_{\alpha}, Du_{\alpha})) - g_u(x, u_{\alpha}, Du_{\alpha}).$$

Setting $\alpha = \Delta t$, $u(0) = u^{\dagger}$, and $u(\Delta t) = u_{\alpha}$ and using (4.1.2) we obtain

$$\mathcal{F}'(u)^{\#} \mathcal{F}'(u) \frac{\partial u}{\partial t} = \nabla \cdot (\nabla_{\xi} g(x, u, Du) - g_u(x, u, Du)).$$

Hence the solution to the optimality condition above together with its boundary condition is equivalent to the finite difference solution to the PDE problem

$$\begin{aligned} \mathcal{F}'(u)^{\#} \mathcal{F}'(u) \frac{\partial u}{\partial t} &= \nabla \cdot (\nabla_{\xi} g(x, u, Du) - g_u(x, u, Du)) \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial \nabla_{\xi} g(x, u, Du)}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial \Omega, \\ u(0) &= u^{\dagger} \quad \text{on } \Omega. \end{aligned}$$

In general though, we will not be able to solve the equation $\mathcal{F}(u^{\dagger}) = u_0$ since we will be given noisy data u_{δ} and not exact data u_0 . Hence the above considerations are not in general useful. Instead, if we use any Fréchet differentiable operator \mathcal{F} and set $g(x, u, Du) = \frac{1}{2}u^2$ then the optimality condition is

$$\mathcal{F}'(u_{\alpha})^{\#} (\mathcal{F}(u_{\alpha}) - u_{\delta}) = -\alpha u_{\alpha}.$$

That is,

$$\alpha u_{\alpha} = -\mathcal{F}'(u_{\alpha})^{\#} (\mathcal{F}(u_{\alpha}) - u_{\delta}).$$

Setting $\alpha = \frac{1}{\Delta t}$, $u(\Delta t) = u_{\alpha}$, and $u(0) = 0$ the solutions $(u_{\Delta t}, u_{2\Delta t}, \dots)$ to the above equation are equivalent to the finite difference (in time) solutions $(u(\Delta t), u(2\Delta t), \dots)$ to the PDE problem

$$\frac{\partial u}{\partial t} = -\mathcal{F}'(u)^{\#} (\mathcal{F}(u) - u_{\delta}), \quad (4.2.5)$$

$$u(0) = 0. \quad (4.2.6)$$

Observe that there is no easy way to directly interpret this equation in its written form, since on the right hand side we obtain an element in X' whereas that is not necessarily the case for the left hand side. Hence this equation is meant to be interpreted in a distributional sense as mentioned earlier.

The idea of solving the diffusion problem in (4.2.6) applying a finite difference method is actually equivalent to iteratively solving the minimization problems

$$P_k : \inf \left\{ \frac{1}{2} \int_{\Omega} |u_{\delta} - \mathcal{F}(u)|^2 dx + \alpha_k \frac{1}{2} \int_{\Omega} |u - u^{k-1}|^2 dx \right\}$$

for u , where $u^{(k)}$ is the solution to the P_k 'th minimization problem and $u^{(0)} = 0$. Letting $\Delta t_k = t_k - t_{k-1}$ be the time steps for the finite difference method then $\alpha_k = 1/(t_k - t_{k-1})$. The process is called *iterative Tikhonov-Morozov regularization* and could lead to a possibly better solution than the solution to the single minimization problem

$$\inf \left\{ \frac{1}{2} \int_{\Omega} |u_{\delta} - \mathcal{F}(u)|^2 dx + \alpha \frac{1}{2} \int_{\Omega} |u|^2 dx \right\}.$$

Of course there are more issues involved with the iterative regularization. We need a rule of how to choose α_k . Here we can use one of the already known rules, for example, the Morozov discrepancy principle, or other more complex methods. We also need a stopping rule for terminating the iteration process. It could also be asked if the α_k s converge to zero or if there is some other limit that we can determine. These questions will not be answered here, but are definitely questions that could be interesting to answer. Finally, it could be considered if the solutions of the iterative regularization process converge to a solution $u \in X$.

In the above calculations we assumed that g was twice differentiable in both u and ξ . This is not necessarily the case. If we instead let g be proper and convex in the last two components, then we can use the subdifferential of g in these two components to obtain similar results. We start by proving a general result for the subdifferential of a functional at its minimizer:

Lemma 4.2.1. *Let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be convex. Then u_{α} is a minimizer of F if and only if $0 \in \partial F(u_{\alpha})$.*

Proof. Proof of " \Rightarrow ": Let u_{α} be a minimizer of F . Then

$$F(u_{\alpha}) \leq F(v), \quad \forall v \in X. \tag{4.2.7}$$

By definition we have

$$\partial F(u_{\alpha}) = \{x' \in X' : F(u_{\alpha}) - F(v) - (u_{\alpha} - v, x') \leq 0, \forall v \in X\}.$$

Setting $x' = 0$ and using (4.2.7) we get

$$F(u_\alpha) - F(v) - (u_\alpha - v, 0) = F(u_\alpha) - F(v) \leq 0$$

for all $v \in X$. This shows that $0 \in \partial F(u_\alpha)$.

Proof of " \Leftarrow ": Let $0 \in \partial F(u_\alpha)$. By definition of $\partial F(u_\alpha)$ we then have

$$F(u_\alpha) \leq F(v) + (u_\alpha - v, 0) = F(v)$$

for all $v \in X$. This shows that u_α is a minimizer of F . \square

Next we consider the minimization of the functional $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$F(u) = \int_{\Omega} f(x, u(x), Du(x)) \, dx$$

where $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is a normal, proper, and convex integrand, and X is either $W^{1,p}(\Omega)$ for $1 < p < \infty$ or $BV(\Omega)$ for $p = 1$. In case $X = BV(\Omega)$ we extend the functional as described in Section 2.3 and we note that a minimizer is an element of $W^{1,1}(\Omega)$. Since f is a convex integrand, it is convex in the last two components, which implies that F is convex. Hence by Lemma 4.2.1 u_α is a minimizer of F if and only if $0 \in \partial F(u_\alpha)$. Denote by $j : X \rightarrow L^p(\Omega) \times L^p(\Omega; \mathbb{R}^n)$ the operator

$$j(u) = (u, Du).$$

Then

$$F = G \circ j$$

where

$$G(\tilde{v}, v) = \int_{\Omega} f(x, \tilde{v}(x), v(x)) \, dx$$

for $(\tilde{v}, v) \in L^p(\Omega) \times L^p(\Omega, \mathbb{R}^n)$. From Lemma A.3.1 in Appendix A.3 it follows that

$$\partial F(u_\alpha) = \partial(G \circ j)(u_\alpha) = j^\# \partial G(u_\alpha, Du_\alpha). \quad (4.2.8)$$

Using Theorem A.3.2 in Appendix A.3 we obtain

$$\partial G(u_\alpha, Du_\alpha) = (\partial_u f(x, u_\alpha, Du_\alpha), \partial_\xi f(x, u_\alpha, Du_\alpha)) \quad (4.2.9)$$

where ∂_u and ∂_ξ denotes the subdifferential with respect to the second variable u of f respectively the third variable ξ of f . The subdifferential ∂_ξ will be computed as a gradient. The adjoint $j^\# : (L^p(\Omega) \times L^p(\Omega, \mathbb{R}^n))' \rightarrow X'$ is defined by the equation

$$(u, j^\#(\tilde{v}, v)) = (j(u), (\tilde{v}, v))$$

where $u \in X$ and $(\tilde{v}, v) \in (L^p(\Omega) \times L^p(\Omega, \mathbb{R}^n))'$. By definition of the pairing we have

$$(j(u), (\tilde{v}, v)) = ((u, Du), (\tilde{v}, v)) = (u, \tilde{v}) + (Du, v)$$

and using Theorem A.1.2 we obtain

$$\begin{aligned} (u, j^\#(\tilde{v}, v))_{W^{1,p}(\Omega) \times (W^{1,p}(\Omega))'} \\ = (u, j^\#(\tilde{v}, v))_{L^p(\Omega), (L^p(\Omega))'} + (Du, Dj^\#(\tilde{v}, v))_{L^p(\Omega, \mathbb{R}^n), (L^p(\Omega, \mathbb{R}^n))'}. \end{aligned}$$

The two above equalities and the definition of $j^\#$ then implies

$$(u, j^\#(\tilde{v}, v))_{L^p(\Omega), (L^p(\Omega))'} + (Du, Dj^\#(\tilde{v}, v) - v)_{L^p(\Omega, \mathbb{R}^n), (L^p(\Omega, \mathbb{R}^n))'} = (u, \tilde{v})$$

Using integration by parts, this implies that $\omega = j^\#(\tilde{v}, v)$ is a solution to the equation

$$\omega - \nabla \cdot (D\omega - v) = \tilde{v} \text{ in } \Omega,$$

i.e.

$$\Delta\omega - \omega = \nabla \cdot (v) - \tilde{v} \text{ in } \Omega$$

with boundary condition

$$\frac{\partial}{\partial \nu} (D\omega - v) = 0 \text{ on } \partial\Omega.$$

Both the PDE and the boundary equation should be interpreted in distributional sense. Formally denoting $j^\#(\tilde{v}, v) = (\Delta - I)^{-1}(\nabla \cdot (v) - \tilde{v})$ we obtain from (4.2.8) and (4.2.9) that the optimality condition $0 \in \partial F(u_\alpha)$ reads as

$$0 \in (\Delta - I)^{-1} (\nabla \cdot (\partial_\xi f(x, u_\alpha, Du_\alpha)) - \partial_u f(x, u_\alpha, Du_\alpha))$$

or simplified as what we call the *Euler-Lagrange equation*:

$$\partial_u f(x, u_\alpha, Du_\alpha) \in \nabla \cdot (\partial_\xi f(x, u_\alpha, Du_\alpha)). \quad (4.2.10)$$

Now if we let $F = \tau_{\alpha, u_\delta}$ as defined in (4.0.1) in the above, then

$$f(x, u, \xi) = \frac{1}{2} |u_\delta - \mathcal{F}u|^2 + \alpha g(x, u, \xi).$$

Assuming that \mathcal{F} is linear and bounded and that u_α is a minimizer of \mathcal{F} and by using Lemma A.3.1 in Appendix A.3 again, we get

$$\partial_u f(x, u_\alpha, Du_\alpha) = \partial \mathcal{F}(u_\alpha)^\# (\mathcal{F}u_\alpha - u_\delta) + \alpha \partial_u g(x, u_\alpha, Du_\alpha).$$

and

$$\partial_\xi f(x, u_\alpha, Du_\alpha) = 0 + \alpha \partial_\xi g(x, u_\alpha, Du_\alpha).$$

The Euler-Lagrange equation (4.2.10) then reads as

$$\partial \mathcal{F}(u_\alpha)^\#(\mathcal{F}u_\alpha - u_\delta) + \alpha \partial_u g(x, u_\alpha, Du_\alpha) \in \alpha \nabla \cdot (\partial_\xi g(x, u_\alpha, Du_\alpha)).$$

Actually since \mathcal{F} is linear and bounded it is Gâteaux differentiable and hence $\partial \mathcal{F}(u_\alpha) = \mathcal{F}'(u_\alpha)$ contains only one element so that the above equation becomes

$$\mathcal{F}'(u_\alpha)^\#(\mathcal{F}u_\alpha - u_\delta) \in \alpha (\nabla \cdot (\partial_\xi g(x, u_\alpha, Du_\alpha)) - \partial_u g(x, u_\alpha, Du_\alpha)). \quad (4.2.11)$$

Comparing the above optimality condition (4.2.11) with the optimality condition in (4.2.3) we see that they are actually the same except that in the above we use the subdifferential of g instead of the strong derivatives of g . Hence g needs not be twice differentiable, but just proper and convex, in order to formulate an Euler-Lagrange equation associated with the minimizer of F . Note also that the above optimality condition can be analysed in its written form since both the left hand side and the right hand side are subsets of X' .

4.3 Examples

In this section we give concrete examples of what the optimality condition and the associated diffusion problem looks like for different choices of \mathcal{F} and g . We will observe that more of the diffusion problems make sense in their strong (written) formulation, i.e. they can be analysed directly in their written form. Therefore the next chapter will deal with non-linear diffusion problems and their strong solutions.

4.3.1 Denoising

When $\mathcal{F} : L^2(\Omega) \rightarrow L^2(\Omega)$ is the identity mapping

$$\mathcal{F}(u) = u$$

the problem of minimizing the functional τ_{α, u_δ} defined in (4.0.1) is called *denoising*. This is because we want to reconstruct u from noisy data u_δ . The way we compare these two quantities is through the identity mapping, so we actually want to remove the noise from u_δ . In this case the optimality condition in (4.2.11) reads as

$$\frac{u_\alpha - u_\delta}{\alpha} \in \nabla \cdot (\partial_\xi g(x, u_\alpha, Du_\alpha)) - \partial_u g(x, u_\alpha, Du_\alpha).$$

Letting $\alpha = \Delta t$, $u(\Delta t) = u_\alpha$ and $u(0) = u_\delta$ then the solutions $(u_{\Delta t}, u_{2\Delta t}, \dots)$ to the above equation are equivalent to the finite difference (in time) solutions $(u(\Delta t), u(2\Delta t), \dots)$ to the non-linear diffusion problem

$$\begin{aligned} \frac{\partial u}{\partial t} &\in \nabla \cdot (\partial_\xi g(x, u, Du)) - \partial_u g(x, u, Du) \quad \text{in } (0, \infty) \times \Omega, \\ u(0) &= u_\delta \quad \text{on } \Omega. \end{aligned}$$

If we want to penalise for solutions having large $L^2(\Omega)$ -norm we let $g(x, u, Du) = \frac{1}{2}|u|^2$ and $X = L^2(\Omega)$. This is the standard Tikhonov regularization. We replace the original boundary condition by a homogeneous Neumann boundary condition for u . This implies that no new information can be put into the system and no information leaves the system. This is a reasonable assumption for reconstruction problems. The choice of g leads to the ODE problem in $L^2(\Omega)$:

$$\begin{aligned} \frac{\partial u}{\partial t} + u &= 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= u_\delta \quad \text{on } \Omega. \end{aligned}$$

Here both $\frac{\partial u}{\partial t} \in L^2(\Omega)$ and $u \in L^2(\Omega)$ and in this case we can make sense to the diffusion problem in its written (strong) form.

If we instead want to penalise for solutions having large oscillations then we can set $g(x, u, Du) = \frac{1}{2}|Du|^2$ and $X = W^{1,2}(\Omega)$. The corresponding PDE problem is then

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= u_\delta \quad \text{on } \Omega. \end{aligned}$$

In this case the diffusion problem can be analysed in its written form as well.

Finally, if we want to penalise for solutions having large total variation, that is, we want to have sharp edges in an image for instance, then we let the penalization term be $|Du|(\Omega) = \int_\Omega |Du|$. This looks like we could set $g(x, u, Du) = |Du|$. We let $X = BV(\Omega)$. The associated PDE problem reads as (see [24, Chap. 1])

$$\begin{aligned} \frac{\partial u}{\partial t} - \nabla \cdot \left(\frac{1}{|Du|} Du \right) &= 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= u_\delta \quad \text{on } \Omega. \end{aligned}$$

In contrary to the two other regularisation methods and their associated diffusion problems, the above problem cannot be analysed in its written (strong) form. Here we need to interpret it in its distributional sense.

Actually let us go one step back and try to calculate directly the subdifferential of τ_{α, u_δ} for $\mathcal{F} = I : W^{1,1}(\Omega) \rightarrow L^2(\Omega)$ (note that we need Ω to be at least two dimensional in order for \mathcal{F} to be bounded) and penalization term $|Du|(\Omega)$. Using Asplund's Theorem 4.1.6 we obtain

$$\begin{aligned} \partial\tau_{\alpha, u_\delta}(u) &= \partial_u \left(\frac{1}{2} \|u_\delta - u\|_{L^2(\Omega)}^2 \right) + \alpha \partial_u (|Du|(\Omega)) \\ &= \mathcal{J}(u - u_\delta) + \mathcal{J}_{1,1}(u), \end{aligned}$$

where $\mathcal{J} : W^{1,1}(\Omega) \rightarrow (W^{1,1}(\Omega))'$ is the duality mapping according to the weight function $\phi(t) = t$, i.e.

$$\mathcal{J}(u) = \{u' \in X' : (u, u') = \|u'\|^2 = \|u\|^2\}.$$

and $\mathcal{J}_{1,1} : W^{1,1}(\Omega) \rightarrow (W^{1,1}(\Omega))'$ is the duality mapping according to $\phi(t) = 1$ given by

$$\mathcal{J}_{1,1}(u) = -\nabla \cdot \left(\frac{1}{|Du|} Du \right),$$

see [18, pp. 280-281]. In this case the final optimality condition for a minimizer u_α of τ_{α, u_δ} is

$$0 \in \mathcal{J}(u_\alpha - u_\delta) - \alpha \nabla \cdot \left(\frac{1}{|Du_\alpha|} Du_\alpha \right).$$

Setting $\alpha = \Delta t$, $u(\Delta t) = u_\alpha$ and $u(0) = u_\delta$ and applying a Neumann boundary condition the diffusion filtering method is

$$\begin{aligned} \mathcal{J} \left(\frac{\partial u}{\partial t} \right) - \nabla \cdot \left(\frac{1}{|Du|} Du \right) &\ni 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= u_\delta \quad \text{on } \Omega. \end{aligned}$$

The above non-linear diffusion problem makes sense in its written form as both terms on the left hand side of the diffusion inclusion are subsets of $(W^{1,1}(\Omega))'$.

4.3.2 Bounded Linear Operator

Whenever $\mathcal{F} : X \rightarrow L^2(\Omega)$ is a bounded linear operator, it is Fréchet differentiable. Hence we can use the considerations of this chapter to write up an associated non-linear diffusion problem to the minimization problem of τ_{α, u_δ} defined in (4.0.1). If we use the standard Tikhonov regularization as described above and apply a homogeneous Neumann boundary condition for u , we get the associated diffusion problem

$$\begin{aligned}\frac{\partial u}{\partial t} &= -\mathcal{F}'(u)^\#(\mathcal{F}u - u_\delta) \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= 0 \quad \text{on } \Omega.\end{aligned}$$

This problem should be interpreted in a distributional sense, but we can still try to solve it directly using a finite difference method. As for the denoising problem we can try to calculate the subdifferential of τ_{α, u_δ} directly. Using Asplund's Theorem 4.1.6 again we reach at the non-linear diffusion filtering problem

$$\begin{aligned}\mathcal{J} \left(\frac{\partial u}{\partial t} \right) &\ni -\mathcal{F}'(u)^\#(\mathcal{F}u - u_\delta) \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= 0 \quad \text{on } \Omega,\end{aligned}$$

where $\mathcal{J} : X \rightarrow X'$ is the duality mapping defined by

$$\mathcal{J}(u) = \{u' \in X' : (u, u') = \|u'\|^2 = \|u\|^2\}.$$

This diffusion problem can be analysed in its written form both the left hand side and the right hand side of the PDE are subsets of X' .

A final example is that of a general linear, bounded operator $\mathcal{F} : W^{1,1} \rightarrow L^2(\Omega)$ and penalization terme $|Du|(\Omega)$. Using Asplund's Theorem to calculate the subdifferential of τ_{α, u_δ} we get

$$\begin{aligned}\partial\tau_{\alpha, u_\delta}(u) &= \partial_u \left(\frac{1}{2} \|u_\delta - \mathcal{F}u\|_{L^2(\Omega)}^2 \right) + \alpha \partial_u(|Du|(\Omega)) \\ &= \mathcal{F}'(u)^\#(\mathcal{F}u - u_\delta) + \alpha \mathcal{J}_{1,1}(u)\end{aligned}$$

where again $\mathcal{J}_{1,1} : W^{1,1} \rightarrow (W^{1,1}(\Omega))'$ is the duality mapping according to the weight function $\phi(t) = t$. The optimality condition for a minimizer u_α of τ_{α, u_δ} is therefore

$$0 \in \mathcal{F}'(u_\alpha)^\#(\mathcal{F}u_\alpha - u_\delta) + \alpha \mathcal{J}_{1,1}(u_\alpha).$$

Setting $\alpha = 1/\Delta t$, $u(\Delta t) = u_\alpha$ and $u(0) = 0$ and applying a Neumann boundary condition the corresponding diffusion filtering method is

$$\begin{aligned} \mathcal{F}'(u)^\#(\mathcal{F}u - u_\delta) + \alpha \mathcal{J}_{1,1} \left(\frac{\partial u}{\partial t} \right) &\ni 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= 0 \quad \text{on } \Omega \end{aligned}$$

or using the definition of $\mathcal{J}_{1,1}$:

$$\begin{aligned} \mathcal{F}'(u)^\#(\mathcal{F}u - u_\delta) - \alpha \nabla \cdot \left(\frac{1}{|Du_t|} Du_t \right) &\ni 0 \quad \text{in } (0, \infty) \times \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } (0, \infty) \times \partial\Omega, \\ u(0) &= 0 \quad \text{on } \Omega. \end{aligned}$$

This is a diffusion problem in the dual space X' . However, for this problem the time and spacial derivatives are mixed and the discretization of the problem would lead to an implicit scheme.

It is worth mentioning some examples of bounded linear operators that actually could occur in practice. If, for example, we use a camera for taking a picture and make a motion blur by moving the camera while taking the picture. Then we can set up a minimization problem where \mathcal{F} is a blurring operator, that could be a convolution, representing the motion blur. Reconstructing the picture is then called a *deblurring* process. The blur could also occur according to a lens out of focus. Another very interesting application is in tomography. In computed tomography, \mathcal{F} would be the Radon transform which is a bounded linear operator from $L^2(\Omega)$ to $L^2(\Omega)$. Numerical experiments for these problems will be carried out and discussed in Chapter 7.

We see that in all cases above, we need to analyse problems of the form

$$\frac{\partial u}{\partial t} + A(u) \ni 0, \quad u(0) = x.$$

where the time derivative u_t is regarded either as an element of X or as an element of X' depending on whether $A : X \rightarrow X$ or $A : X \rightarrow X'$. In Chapter 5 we deal with problems where $A : X \rightarrow X$. The problems are analysed in their written (strong) form and it is shown that there exists a solution under certain conditions on A . In Chapter 6 we return to the case in which $A : X \rightarrow X'$ and set up a conjecture for the existence of a solution, based on the results of Chapter 5.

5

Non-linear Diffusion Methods

¹In this chapter we analyse the non-linear diffusion filtering method for reconstruction problems. As we saw in Chapter 4, the method consists in solving a Cauchy problem:

$$\frac{\partial u}{\partial t} + A(u) \ni 0, u(0) = u_0, \quad (5.0.1)$$

where A is an operator on X and may be non-linear and unbounded. We use the ' \in ' symbol instead of '=' because A can be multivalued. As seen in Chapter 4 the diffusion problems of interest can either include an operator $A : X \rightarrow X$ or an operator $A : X \rightarrow X'$. In this section we treat only the case in which A maps X into X . In the case X is a Hilbert space we can of course include the problems in which A maps X into X' by identifying X' with X . Hille and Yosida described the case in which A is a linear operator; see [10] or [9]. Here we want to extend the case to non-linear operators on Banach spaces. This makes the situation much more complicated. The theory of semi-groups will be needed as we will show that a solution to the Cauchy problem can be written as the evaluation of a semi-group at time t operating on the initial condition. The semi-group theory will be covered in the first section of this chapter.

¹This chapter is based on [2] and [7].

5.1 Semi-Group Theory

In this section we set up the necessary definitions and state the necessary theorems in order to analyse the non-linear diffusion based methods. Throughout the chapter X will denote a Banach space. We begin by defining a semi-group on a Banach space:

Definition 5.1.1 (Semi-Group). Let $C \subset X$. A *semi-group* on C is a function S on $[0, \infty)$ such that $S(t)$ maps C into C for each $t \geq 0$ and the following two conditions hold:

- (i) $S(t + \tau) = S(t)S(\tau)$ for $t, \tau \geq 0$,
- (ii) $\lim_{t \searrow 0} S(t)v = S(0)v = v$ for $v \in C$.

Whenever the values of a semi-group S are linear bounded operators, then $S(t)u_0$ solves the Cauchy problem (5.0.1) if A is what we call the *infinitesimal generator* of S and $u_0 \in \mathcal{D}(A)$. The infinitesimal generator of a semi-group is defined as:

Definition 5.1.2 (Infinitesimal Generator). Let S be a semi-group on X for which the values are linear bounded operators. The infinitesimal generator of S is the operator $A : \mathcal{D}(A) \rightarrow X$ defined by

$$\mathcal{D}(A) = \left\{ v \in X : \lim_{t \searrow 0} \frac{S(t)v - v}{t} \in X \text{ exists} \right\},$$

and

$$Av = -\lim_{t \searrow 0} \frac{S(t)v - v}{t}, \quad v \in \mathcal{D}(A).$$

It turns out that every infinitesimal generator A of a semi-group S is a closed operator and has dense domain in X . Furthermore if we set $u(t) = S(t)u_0$ whenever $u_0 \in \mathcal{D}(A)$ then u solves the Cauchy problem in (5.0.1):

Proposition 5.1.3. *Let S be a semi-group on X for which the values are linear bounded operators and let $A : \mathcal{D}(A) \rightarrow X$ be its infinitesimal generator. Then the set $\mathcal{D}(A)$ is dense in X and $A : \mathcal{D}(A) \rightarrow X$ is closed. Furthermore for $u_0 \in \mathcal{D}(A)$ let $u(t) = S(t)u_0$. Then the function $u : [0, \infty) \rightarrow X$ is continuously differentiable and*

$$\frac{\partial u}{\partial t}(t) = -Au(t) = -S(t)Au_0.$$

Proof. See [30, Chap. 1, Prop. 3.1]. \square

The Hille-Yosida theorem characterizes those A s which are actually infinitesimal generators of uniquely defined semi-groups. That is, the theorem actually states under which conditions the Cauchy problem (5.0.1) has a solution whenever A is linear. The corresponding semi-group is given by the exponential formula

$$S(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} := e^{-tA}.$$

In this chapter we will extend the notion of infinitesimal generator to the non-linear case and in this way prove under which conditions the Cauchy problem in (5.0.1) has a solution.

We will sometimes need a semi-group to be Lipschitz continuous on its domain for each $t > 0$. We define the notion of such semi-groups:

Definition 5.1.4 ($Q_\rho(C)$). If S is a semi-group on C and there exists a real number ρ such that

$$\|S(t)v - S(t)w\| \leq e^{\rho t} \|v - w\|, \text{ for } t \geq 0, v, w \in C$$

we write $S \in Q_\rho(C)$.

Since we are going to describe situations in which the value $S(t)$ of a semi-group S at a point $t > 0$ is a non-linear operator we need the notion of *multivalued operators* $A : X \rightarrow X$. These can be viewed as subsets of $X \times X$, and we write $A \subset X \times X$. Then we can define

1. $A(v) = \{w : [v, w] \in A\}$
2. $\mathcal{D}(A) = \{v : A(v) \neq \emptyset\}$,
3. $\mathcal{R}(A) = \bigcup_{v \in \mathcal{D}(A)} A(v)$,
4. $A^{-1} = \{[w, v] : [v, w] \in A\}$.

Furthermore if both $A, B \subset X \times X$ and $\lambda \in \mathbb{R}$ then

5. $A + B = \{[v, w + z] : w \in A(v), z \in B(v)\}$,
6. $\lambda A = \{[v, \lambda w] : w \in A(v)\}$.

Finally we need the notion of an accretive operator:

Definition 5.1.5 (Accretive Operator or Set). Let $B \subset X \times X$. Then B is accretive if $(I + \lambda B)^{-1}$ is a single-valued operator for $\lambda > 0$ and

$$\|(I + \lambda B)^{-1}(v) - (I + \lambda B)^{-1}(w)\| \leq \|v - w\|, \text{ for } v, w \in \mathcal{D}((I + \lambda B)^{-1}).$$

Let us try to understand what it means for an operator to be accretive. Let $\lambda > 0$ and let B be linear and accretive. Then

$$(I + \lambda B)^{-1} = \left(\lambda \left(\frac{1}{\lambda} I + B \right) \right)^{-1} = \frac{1}{\lambda} \left(B - \frac{1}{-\lambda} I \right)^{-1}$$

exists. Also by the second condition of B being accretive we obtain

$$\|(I + \lambda B)^{-1}\| = \left\| \frac{1}{\lambda} \left(B - \frac{1}{-\lambda} I \right)^{-1} \right\| = \frac{1}{\lambda} \left\| \left(B - \frac{1}{-\lambda} I \right)^{-1} \right\| \leq 1$$

so that

$$\left\| \left(B - \frac{1}{-\lambda} I \right)^{-1} \right\| \leq \lambda$$

for each $\lambda > 0$. Hence the operators $(B - \frac{1}{-\lambda} I)^{-1}$ exist and are bounded. If further the domain $\mathcal{D}((I + \lambda B)^{-1})$ is dense in X then we see that $-\frac{1}{\lambda} < 0$ is exactly a regular value of the resolvent operator of B and this shows that the resolvent set of B contains the negative real axis.

A simple example of an accretive operator is $B = -\Delta : W^{1,2}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ defined on $H^2(\mathbb{R}^n) = W^{2,2}(\mathbb{R}^n)$. Since the resolvent set of $-\Delta$ is the negative real axis (see [14, Section 12.3]), then for $\rho < 0$ the resolvent operator

$$(B - \rho I)^{-1}$$

exists and is bounded. Rewriting the operator expression we get

$$(B - \rho I)^{-1} = \left(-\rho \left(I - \frac{1}{\rho} B \right) \right)^{-1} = -\frac{1}{\rho} \left(I + \frac{1}{-\rho} B \right)^{-1}$$

Letting $\lambda = \frac{1}{-\rho} > 0$ this shows that

$$(I + \lambda B)^{-1}$$

exists for all $\lambda > 0$. That is $(I + \lambda B)^{-1}$ is a single-valued operator for $\lambda > 0$. Now we need to show that the operator is bounded. For this we use the Fourier transform

to rewrite the expression for $\omega = (I + \lambda B)^{-1}v$ for $v \in \mathcal{D}((I + \lambda B)^{-1}) \subset L^2(\mathbb{R}^n)$. First we see that

$$(I + \lambda B)\omega = v.$$

Now $v \in L^2(\mathbb{R}^n)$. Letting \mathcal{F} or $\hat{\cdot}$ denote the Fourier transform we get

$$\mathcal{F}((I + \lambda B)\omega) = \mathcal{F}((I - \lambda \Delta)\omega) = \mathcal{F}(v).$$

This implies that

$$(1 + \lambda|\xi|^2)\hat{\omega}(\xi) = \hat{v}(\xi).$$

This leads to the expression for $(I + \lambda B)^{-1}v$:

$$(I + \lambda B)^{-1}v = \omega = \mathcal{F}^{-1} \left(\frac{\hat{v}(\xi)}{1 + \lambda|\xi|^2} \right).$$

This shows that $(I + \lambda B)^{-1}v$ is a bounded linear operator in $L^2(\rho)$ with

$$\begin{aligned} \|(I + \lambda B)^{-1}v\|_{L^2(\mathbb{R}^2)} &= \|\mathcal{F}^{-1} \left(\frac{\hat{v}(\xi)}{1 + \lambda|\xi|^2} \right)\| \\ &= \frac{1}{1 + \lambda|\xi|^2} \frac{1}{\sqrt{2\pi}} \|\hat{v}\|_{L^2(\mathbb{R}^2)} \\ &= \frac{1}{1 + \lambda|\xi|^2} \frac{\sqrt{2\pi}}{\sqrt{2\pi}} \|\hat{v}\|_{L^2(\mathbb{R}^2)} \\ &\leq \|v\|_{L^2(\rho)}. \end{aligned}$$

From this we conclude that B is accretive.

5.2 Extension of the Hille-Yosida Theorem

The Hille-Yosida theorem characterizes those $S \in Q_\rho(X)$ for which the values of S are linear operators and X is a Banach space, see [30, Thm. 3.1]. In this section we want to give a complete characterization of those $S \in Q_\rho(C)$ and their infinitesimal generators for which the values of S are non-linear operators, and C is a subset of a Banach space X . In a Hilbert space the situation is simpler because we have the notion of an inner product. A characterization in the non-linear case where X is a Hilbert space is given in the appendix of [15]. However, when X is a Banach space the situation is much more complex, and we need a similar notion of an inner product.

5.2.1 The Exponential Formula

In this section we show that if $A + \rho I$ is accretive for some $\rho \in \mathbb{R}$ and the domain of A is dense in the range of $I + \lambda A$ for sufficiently small $\lambda > 0$, then A is the “infinitesimal generator” of a semi-group S , which has an exponential formula. Note that here the definition of an infinitesimal generator is used even though A is non-linear. Throughout the rest of this chapter we fix $A \subset X \times X$ and $\rho \in \mathbb{R}$ such that $A + \rho I$ is accretive. In case A is linear, all $\mu < -\rho$ are part of the resolvent set of A as long as the domain of the resolvent operator of $A + \rho I$ is dense in X . Next we define the set

$$J_\lambda := (I + \lambda A)^{-1}$$

for all $\lambda \in \mathbb{R}$ and let $\mathcal{D}_\lambda = \mathcal{D}(J_\lambda)$ be its domain. Finally we define

$$|A(v)| = \inf\{\|w\| : w \in A(v)\}$$

for $v \in \mathcal{D}(A)$.

The first thing we need is some elementary facts about J_λ :

Lemma 5.2.1 (Elementary Facts about J_λ). *Take $\lambda \geq 0$ so that $\lambda\rho < 1$. Then the following four statements hold:*

(i) J_λ is a function and for $v, w \in \mathcal{D}_\lambda$ we have

$$\|J_\lambda(v) - J_\lambda(w)\| \leq (1 - \lambda\rho)^{-1}\|v - w\|.$$

(ii) For $v \in \mathcal{D}_\lambda \cap \mathcal{D}(A)$ we have

$$\|J_\lambda(v) - v\| \leq \lambda(1 - \lambda\rho)^{-1}|A(v)|.$$

(iii) If n is a positive integer, $v \in \mathcal{D}(J_\lambda^n)$ and $\lambda|\rho| < 1$ then

$$\|J_\lambda^n(v) - v\| \leq n(1 - \lambda|\rho|)^{-n+1}\|J_\lambda(v) - v\|.$$

(iv) If $v \in \mathcal{D}_\lambda$, $\lambda > 0$, and $\mu \in \mathbb{R}$, then

$$\frac{\mu}{\lambda}v + \frac{\lambda - \mu}{\lambda}J_\lambda(v) \in \mathcal{D}_\mu \quad \text{and} \quad J_\lambda v \in J_\mu \left(\frac{\mu}{\lambda}v + \frac{\lambda - \mu}{\lambda}J_\lambda(v) \right).$$

Proof. By assumption $A + \rho I$ is accretive so for $t \geq 0$ the operator $(I + t(A + \rho I))^{-1}$ is a single-valued operator and has Lipschitz constant 1. Now for $1 + t\rho \neq 0$ we

have

$$\left(I + \frac{t}{1+t\rho}A\right)^{-1} = \left(\frac{1}{1+t\rho}((1+t\rho)I + tA)\right)^{-1} = (1+t\rho)(I + t(A + \rho I))^{-1}.$$

This shows that

$$\left(I + \frac{t}{1+t\rho}A\right)^{-1}$$

is a single valued operator and has Lipschitz constant $|1+t\rho|$ for $t \geq 0$ and $1+t\rho \neq 0$. Setting $\lambda = t(1+t\rho)^{-1}$ and using the restrictions $t \geq 0$ and $\lambda\rho < 1$ we find that $J_\lambda = (I + \lambda A)^{-1}$ is a single-valued operator and has Lipschitz constant

$$\begin{aligned} |1+t\rho| &= |((1+t\rho)^{-1})^{-1}| = |((1+t\rho)(1+t\rho)^{-1} - t\rho(1+t\rho)^{-1})^{-1}| \\ &= |(1 - t(1+t\rho)^{-1}\rho)^{-1}| = |(1 - \lambda\rho)^{-1}| = (1 - \lambda\rho)^{-1}. \end{aligned}$$

This proves assertion (i).

To prove assertion (ii) take $[v_1, w_1] \in A$ and $[v, w] \in A$ such that $v_1 + \lambda w_1 = v$. This is possible simply because X is a vector space. Using the fact that J_λ is single-valued, and that $v + \lambda w \in J_\lambda^{-1}(v)$ we obtain

$$J_\lambda(v + \lambda w) = v.$$

Using assertion (i) we obtain

$$\begin{aligned} \|J_\lambda(v) - v\| &= \|J_\lambda(v_1 + \lambda w_1) - J_\lambda(v + \lambda w)\| \\ &\leq (1 - \lambda\rho)^{-1}\|(v_1 + \lambda w_1) - (v + \lambda w)\| \\ &= (1 - \lambda\rho)^{-1}\|v - v - \lambda w\| \\ &= \lambda(1 - \lambda\rho)^{-1}\|w\|. \end{aligned}$$

Since $w \in A(v)$ was arbitrary we can take the infimum over all $w \in A(v)$ on the right side of the inequality and obtain

$$\|J_\lambda(v) - v\| \leq \lambda(1 - \lambda\rho)^{-1}|A(v)|$$

which is assertion (ii).

In order to prove assertion (iii) we use assertion (i):

$$\begin{aligned}
\|J_\lambda^n(v) - v\| &= \left\| \sum_{i=0}^{n-1} (J_\lambda^{n-i}(v) - J_\lambda^{n-(i+1)}(v)) \right\| \leq \sum_{i=0}^{n-1} \|J_\lambda^{n-i}(v) - J_\lambda^{n-(i+1)}(v)\| \\
&\stackrel{(i)}{\leq} \sum_{i=0}^{n-1} (1 - \lambda\rho)^{-1} \|J_\lambda^{n-(i+1)}(v) - J_\lambda^{n-(i+2)}(v)\| \stackrel{(i)}{\leq} \dots \\
&\stackrel{(i)}{\leq} \sum_{i=0}^{n-1} (1 - \lambda\rho)^{-n+(i+1)} \|J_\lambda(v) - v\| \\
&\leq \sum_{i=0}^{n-1} (1 - \lambda|\rho|)^{-n+(i+1)} \|J_\lambda(v) - v\| \\
&\stackrel{0 < (1 - \lambda|\rho|) < 1}{\leq} \sum_{i=0}^{n-1} (1 - \lambda|\rho|)^{-n+1} \|J_\lambda(v) - v\| \\
&= n(1 - \lambda\rho)^{-n+1} \|J_\lambda(v) - v\|.
\end{aligned}$$

Finally we prove assertion (iv). If $v \in \mathcal{D}_\lambda$ then by definition there exists $[v_0, w_0] \in A$ such that $v_0 + \lambda w_0 = v$. Since $v_0 + \lambda w_0 \in J_\lambda^{-1}(v)$ and since J_λ is single-valued then

$$J_\lambda(v) = J_\lambda(v_0 + \lambda w_0) = v_0$$

and we get

$$\frac{\mu}{\lambda}v + \frac{\lambda - \mu}{\lambda}J_\lambda(v) = \frac{\mu}{\lambda}(v_0 + \lambda w_0) + \frac{\lambda - \mu}{\lambda}v_0 = v_0 + \mu w_0.$$

Since $[v_0 + \mu w_0, v_0] \in J_\mu$ by the definition of J_μ then

$$\frac{\mu}{\lambda}v + \frac{\lambda - \mu}{\lambda}J_\lambda(v) \in \mathcal{D}_\mu.$$

We also see that

$$J_\lambda(v) = v_0 \in J_\mu(v_0 + \mu w_0) = J_\mu\left(\frac{\mu}{\lambda}v + \frac{\lambda - \mu}{\lambda}J_\lambda(v)\right).$$

This proves assertion (iv). □

Since we actually want to say something about the existence of $\lim_{n \rightarrow \infty} J_{t/n}^n(v)$ for $v \in \mathcal{D}(A)$ we need an estimate for the norm-difference between $J_\mu^n(v)$ and $J_\lambda^m(v)$:

Lemma 5.2.2. *Let $\lambda \geq \mu > 0$, $\rho\lambda < 1$, and $v \in \mathcal{D}(J_\lambda^m) \cap \mathcal{D}(J_\mu^n)$, where m and n are positive integers satisfying $n \geq m$. Then*

$$\begin{aligned} \|J_\mu^n(v) - J_\lambda^m(v)\| &\leq (1 - \rho\mu)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) \|J_\lambda^{m-j}(v) - v\| \\ &\quad + \sum_{j=m}^n (1 - \rho\mu)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) \|J_\mu^{n-j}(v) - v\|, \end{aligned}$$

where $\alpha = \frac{\mu}{\lambda}$, $\beta = \frac{\lambda-\mu}{\lambda}$, and $B(k, l)$ are the binomial coefficients

$$B(k, l) = \frac{k!}{l!(k-l)!} \text{ for } 0 \leq l \leq k.$$

Proof. For integers j and k satisfying $0 \leq j \leq n$ and $0 \leq k \leq m$ we define the numbers

$$a_{k,j} = \|J_\mu^j(v) - J_\lambda^k(v)\|.$$

Then if $j, k > 0$ we can use Lemma 5.2.1 assertions (i) and (iv) to obtain

$$\begin{aligned} a_{k,j} &\stackrel{(iv)}{=} \left\| J_\mu^j(v) - J_\mu \left(\frac{\mu}{\lambda} J_\lambda^{k-1}(v) + \frac{\lambda-\mu}{\lambda} J_\lambda^k(v) \right) \right\| \\ &\stackrel{(i)}{\leq} (1 - \mu\rho)^{-1} \left\| J_\mu^{j-1}(v) - \left(\frac{\mu}{\lambda} J_\lambda^{k-1}(v) + \frac{\lambda-\mu}{\lambda} J_\lambda^k(v) \right) \right\| \\ &= (1 - \mu\rho)^{-1} \left\| \frac{\mu}{\lambda} J_\mu^{j-1}(v) + \frac{\lambda-\mu}{\lambda} J_\mu^{j-1}(v) - \left(\frac{\mu}{\lambda} J_\lambda^{k-1}(v) + \frac{\lambda-\mu}{\lambda} J_\lambda^k(v) \right) \right\| \\ &\leq (1 - \mu\rho)^{-1} \left(\frac{\mu}{\lambda} \|J_\mu^{j-1}(v) - J_\lambda^{k-1}(v)\| + \frac{\lambda-\mu}{\lambda} \|J_\mu^{j-1}(v) - J_\lambda^k(v)\| \right). \end{aligned}$$

This shows that for $j, k > 0$ we have

$$a_{k,j} \leq \alpha_1 a_{k-1,j-1} + \beta_1 a_{k,j-1} \tag{5.2.1}$$

with

$$\alpha_1 = (1 - \mu\rho)^{-1} \frac{\mu}{\lambda} \text{ and } \beta_1 = (1 - \mu\rho)^{-1} \frac{\lambda - \mu}{\lambda}.$$

Now using inequality (5.2.1) multiple times for $a_{m,n}$ we obtain

$$\begin{aligned}
a_{m,n} &\leq \alpha_1 a_{m-1,n-1} + \beta_1 a_{m,n-1} \\
&\leq \alpha_1 (\alpha_1 a_{m-2,n-2} + \beta_1 a_{m-1,n-2}) + \beta_1 (\alpha_1 a_{m-1,n-2} + \beta_1 a_{m,n-2}) \\
&= \alpha_1^2 a_{m-2,n-2} + 2\alpha_1 \beta_1 a_{m-1,n-2} + \beta_1^2 a_{m,n-2} \\
&\leq \alpha_1^2 (\alpha_1 a_{m-3,n-3} + \beta_1 a_{m-2,n-3}) + 2\alpha_1 \beta_1 (\alpha_1 a_{m-2,n-3} + \beta_1 a_{m-1,n-3}) \\
&\quad + \beta_1^2 (\alpha_1 a_{m-1,n-3} + \beta_1 a_{m,n-3}) \\
&= \alpha_1^3 a_{m-3,n-3} + 3\alpha_1^2 \beta_1 a_{m-2,n-3} + 3\alpha_1 \beta_1^2 a_{m-1,n-3} + \beta_1^3 a_{m,n-3} \\
&\leq \dots
\end{aligned}$$

Continuing this process we see that we will get a sum with $n + 1$ terms in all. In the first $n - m + 1$ we will drive m to zero so that we will have $n - m + 1$ terms with a constant times $a_{0,n-l}$ for some $l \in [m, n]$. In the last m terms we will drive n to zero before m so that we will have m terms with a constant times $a_{m-l,0}$ for some $l \in [0, m - 1]$. In total we get

$$a_{m,n} \leq \sum_{l=m}^n \alpha_1^m \beta_1^{l-m} B(l-1, m-1) a_{0,n-l} + \sum_{l=0}^{m-1} \alpha_1^l \beta_1^{n-l} B(n, l) a_{m-l,0}. \quad (5.2.2)$$

The binomial coefficients simply come from the way we combine the multiplications of α_1 and β_1 . In the first sum we have the first α_1 from the beginning, so we can take that out of the sum. Then we need to combine in total $l - 1$ multiplications. Since these multiplications have to appear before we have driven m to zero there are $m - 1$ ways to combine the $l - 1$ multiplications. In the second sum we have to combine n multiplications and since each $\alpha_1^k \beta_1^{j-k} a_{m-k,j}$ will never appear again when using inequality (5.2.1) for $a_{m-k,j}$ then the n multiplications can only appear l times.

Now inequality (5.2.2) is exactly the inequality we wanted to prove. \square

Finally we want to get rid of the binomial coefficients occurring in Lemma 5.2.2. For this we have the following lemma:

Lemma 5.2.3. *Let $n \geq m > 0$ be integers, and let α, β be positive numbers satisfying $\alpha + \beta = 1$. Then the two following inequalities are satisfied*

$$(i) \sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m - j) \leq \sqrt{(n\alpha - m)^2 + n\alpha\beta}$$

$$(ii) \sum_{j=m}^n B(j - 1, m - 1) \alpha^m \beta^{j-m} (n - j) \leq \sqrt{\frac{m\beta}{\alpha^2} + \left(\frac{m\beta}{\alpha} + m - n\right)^2}$$

Proof. To obtain assertion (i) we first use the Schwartz inequality to do a simple estimation:

$$\begin{aligned}
& \sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m - j) \\
& \leq \sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} |m - j| \leq \sum_{j=0}^n B(n, j) \alpha^j \beta^{n-j} |m - j| \\
& = \sum_{j=0}^n (B(n, j) \alpha^j \beta^{n-j})^{1/2} (B(n, j) \alpha^j \beta^{n-j})^{1/2} |m - j| \\
& \leq \left(\sum_{j=0}^n B(n, j) \alpha^j \beta^{n-j} \right)^{1/2} \left(\sum_{j=0}^n B(n, j) \alpha^j \beta^{n-j} (m - j)^2 \right)^{1/2}.
\end{aligned}$$

By definition of the binomial coefficients we have

$$\begin{aligned}
& \sum_{j=0}^n B(n, j) \alpha^j \beta^{n-j} = (\alpha + \beta)^n, \\
& \sum_{j=0}^n j B(n, j) \alpha^j \beta^{n-j} = n \alpha (\alpha + \beta)^{n-1}, \\
& \sum_{j=0}^n j^2 B(n, j) \alpha^j \beta^{n-j} = \alpha^2 n(n-1) (\alpha + \beta)^{n-2} + \alpha n (\alpha + \beta)^{n-1}.
\end{aligned}$$

Using the above considerations together with $\alpha + \beta = 1$ we obtain

$$\begin{aligned}
\sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m - j) & \leq 1 \cdot \left(\sum_{j=0}^n B(n, j) \alpha^j \beta^{n-j} (m^2 + j^2 - 2mj) \right)^{1/2} \\
& = \sqrt{m^2 + \alpha^2 n(n-1) + \alpha n - 2m n \alpha} \\
& = \sqrt{(n\alpha - m)^2 + n\alpha(1 - \alpha)} \\
& = \sqrt{(n\alpha - m)^2 + n\alpha\beta}.
\end{aligned}$$

This is the first estimate.

For the second assertion we use the the Schwartz inequality again to obtain

$$\begin{aligned} & \sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j) \\ & \leq \sum_{j=m}^{\infty} B(j-1, m-1) \alpha^m \beta^{j-m} |n-j| \\ & \leq \left(\sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} \right)^{1/2} \left(\sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j)^2 \right)^{1/2}. \end{aligned}$$

Again by definition of the binomial coefficients we have

$$\sum_{j=m}^{\infty} B(j-1, m-1) \beta^{j-m} = \frac{1}{(1-\beta)^m}$$

for $|\beta| < 1$. Differentiating the above we get

$$\sum_{j=m}^{\infty} B(j-1, m-1) \beta^{j-m-1} (j-m) = \frac{m}{(1-\beta)^{m+1}}.$$

Multiplying by β on both sides we then get

$$\sum_{j=m}^{\infty} B(j-1, m-1) \beta^{j-m} (j-m) = \frac{\beta m}{(1-\beta)^{m+1}}.$$

Differentiating again we obtain

$$\sum_{j=m}^{\infty} B(j-1, m-1) \beta^{j-m-1} (j-m)^2 = \frac{m}{(1-\beta)^{m+1}} + \beta \frac{m(m+1)}{(1-\beta)^{m+2}}.$$

Again multiplying by β we get the equality

$$\sum_{j=m}^{\infty} B(j-1, m-1) \beta^{j-m} (j-m)^2 = \frac{\beta m}{(1-\beta)^{m+1}} + \frac{\beta^2 m(m+1)}{(1-\beta)^{m+2}}.$$

Finally we will need the following rewriting

$$\begin{aligned} (n-j)^2 &= (n-m+m-j)^2 = (n-m-(j-m))^2 \\ &= (n-m)^2 + (j-m)^2 - 2(n-m)(j-m). \end{aligned}$$

Collecting the above considerations and using $\alpha + \beta = 1$, we obtain the second assertion by

$$\begin{aligned}
& \sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j) \\
& \leq \sqrt{\frac{\alpha^m}{\alpha^m} \left(\sum_{j=m}^{\infty} B(j-1, m-1) \alpha^m \beta^{j-m} \left((n-m)^2 \right. \right. \\
& \quad \left. \left. + (j-m)^2 - 2(n-m)(j-m) \right) \right)^{1/2}} \\
& = \sqrt{(n-m)^2 \frac{\alpha^m}{\alpha^m} + \frac{\beta m \alpha^m}{\alpha^{m+1}} + \frac{\beta^2 m(m+1) \alpha^m}{\alpha^{m+2}} - 2(n-m) \frac{\beta m \alpha^m}{\alpha^{m+1}}} \\
& = \sqrt{(n-m)^2 + \frac{\beta m}{\alpha} + \frac{\beta^2 m^2}{\alpha^2} + \frac{\beta^2 m}{\alpha^2} - 2(n-m) \frac{\beta m}{\alpha}} \\
& = \sqrt{\left(\frac{m\beta}{\alpha} - (n-m) \right)^2 + \frac{\beta^2 m + \alpha \beta m}{\alpha^2}} \\
& = \sqrt{\left(\frac{m\beta}{\alpha} + m - n \right)^2 + \frac{\beta m(\beta + \alpha)}{\alpha^2}} \\
& = \sqrt{\left(\frac{m\beta}{\alpha} + m - n \right)^2 + \frac{\beta m}{\alpha^2}}.
\end{aligned}$$

□

Now we are ready to prove the main result of this section, namely that if $A + \rho I$ is accretive for some $\rho \in \mathbb{R}$ and $\overline{\mathcal{D}(A)} \subset \mathcal{R}(I + \lambda A)$ for sufficiently small $\lambda > 0$ then A is the "infinitesimal generator" of a semi-group which has an exponential formula. Here the notion of an infinitesimal generator is extended to the case where A is non-linear. In order to motivate why the formula for the semi-group should be exponential, let first A be a square matrix. Then for $u_0 \in \mathcal{D}(A)$ the matrix exponential e^{-tA} satisfies

$$-\lim_{t \searrow 0} \frac{e^{-tA} u_0 - u_0}{t} = A u_0$$

so that A is the infinitesimal generator of $S(t) = e^{-tA}$. Moreover if we set $u(t) = S(t)u_0$ then

$$\frac{du}{dt}(t) = \frac{d}{dt}(e^{-tA} u_0) = -A e^{-tA} u_0 = -A u(t)$$

and

$$u(0) = u_0.$$

Hence in this case $S(t)u_0$ would satisfy the Cauchy problem (5.0.1). The idea is to prove in which cases the operator exponential

$$e^{-tA} = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n}$$

exists for general non-linear operators A :

Theorem 5.2.4. *Let $A \subset X \times X$ and $\rho \in \mathbb{R}$ such that $A + \rho I$ is accretive. If $\mathcal{R}(I + \lambda A) \supset \overline{\mathcal{D}(A)}$ for all sufficiently small positive λ , then*

$$\lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} (v) \tag{5.2.3}$$

exists for $v \in \overline{\mathcal{D}(A)}$ and $t > 0$. Moreover, if $S(t)v$ is defined as the limit in (5.2.3) then $S \in Q_\rho(\overline{\mathcal{D}(A)})$.

Proof. Let $v \in \mathcal{D}(A)$ and assume that $\lambda \geq \mu > 0$, $n \geq m$ and $\lambda|\rho| < 1$. By assumption $\overline{\mathcal{D}(A)} \subset \mathcal{R}(I + \lambda A)$ and by Lemma 5.2.1(iv) we then conclude $v \in$

$\mathcal{D}(J_\lambda^m) \cap \mathcal{D}(J_\mu^n)$. Then Lemma 5.2.1(ii) and (iii) and Lemma 5.2.2 implies

$$\begin{aligned}
& \|J_\mu^n(v) - J_\lambda^m(v)\| \tag{5.2.4} \\
& \stackrel{5.2.2}{\leq} (1 - \rho\mu)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) \|J_\lambda^{m-j}(v) - v\| \\
& + \sum_{j=m}^n (1 - \rho\mu)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) \|J_\mu^{n-j}(v) - v\| \\
& \stackrel{5.2.1(iii)}{\leq} (1 - \rho\mu)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) (m-j) (1 - \lambda|\rho|)^{-(m-j)+1} \|J_\lambda(v) - v\| \\
& + \sum_{j=m}^n (1 - \rho\mu)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) (n-j) (1 - \mu|\rho|)^{-(n-j)+1} \|J_\mu(v) - v\| \\
& \stackrel{5.2.1(ii)}{\leq} (1 - \rho\mu)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) (m-j) (1 - \lambda|\rho|)^{-(m-j)+1} \lambda (1 - \lambda\rho)^{-1} |A(v)| \\
& + \sum_{j=m}^n (1 - \rho\mu)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) (n-j) (1 - \mu|\rho|)^{-(n-j)+1} \mu (1 - \mu\rho)^{-1} |A(v)| \\
& \leq (1 - \mu|\rho|)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) (m-j) (1 - \lambda|\rho|)^{-m+1} \lambda (1 - \lambda|\rho|)^{-1} |A(v)| \\
& + \sum_{j=m}^n (1 - \mu|\rho|)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) (n-j) (1 - \mu|\rho|)^{-(2n-j)+1} \mu (1 - \mu|\rho|)^{-1} |A(v)| \\
& = \left[(1 - \mu|\rho|)^{-n} (1 - \lambda|\rho|)^{-m} \lambda \sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m-j) \right. \\
& \quad \left. + (1 - \mu|\rho|)^{-2n} \mu \sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j) \right] |A(v)| \tag{5.2.5}
\end{aligned}$$

where $\alpha = \frac{\mu}{\lambda}$ and $\beta = \frac{\lambda - \mu}{\lambda}$. We want to approximate the constants $(1 - \mu|\rho|)^{-n}$, $(1 - \lambda|\rho|)^{-m}$, and $(1 - \mu|\rho|)^{-2n}$ by exponential functions: For $t \in [0, \frac{1}{2}]$ and $f(t) = e^{2t}(1-t) - 1$ we have

$$f'(t) = 2e^{2t}(1-t) - e^{2t} = e^{2t}(1-2t) \geq 0.$$

Hence for $t \in [0, \frac{1}{2}]$ we have

$$0 = f(0) \leq f(t) = e^{2t}(1-t) - 1 \Leftrightarrow (1-t)^{-1} \leq e^{2t} \Leftrightarrow (1-t)^{-n} \leq e^{2nt}.$$

If $\mu|\rho| \leq \lambda|\rho| \leq \frac{1}{2}$ and since $\alpha + \beta = 1$ then inequality (5.2.5) and Lemma 5.2.3

imply

$$\begin{aligned} \|J_\mu^n(v) - J_\lambda^m(v)\| &\leq \left[e^{2n|\rho|\mu} e^{2m|\rho|\lambda} \lambda \sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m-j) \right. \\ &\quad \left. + e^{4n|\rho|\mu} \mu \sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j) \right] |A(v)| \\ &\stackrel{5.2.3}{\leq} \left[((n\alpha - m)^2 + n\alpha\beta)^{1/2} \lambda e^{2|\rho|(n\mu+m\lambda)} \right. \\ &\quad \left. + \left(\frac{m\beta}{\alpha^2} + \left(\frac{m\beta}{\alpha} + m - n \right)^2 \right)^{1/2} \mu e^{4|\rho|n\mu} \right] |A(v)|. \end{aligned}$$

Inserting $\alpha = \frac{\mu}{\lambda}$ and $\beta = \frac{\lambda - \mu}{\lambda}$ and rewriting we then aim at

$$\begin{aligned} \|J_\mu^n(v) - J_\lambda^m(v)\| &\leq \left[((n\mu - \lambda m)^2 + n\mu(\lambda - \mu))^{1/2} e^{2|\rho|(n\mu+m\lambda)} \right. \\ &\quad \left. + (m\lambda(\lambda - \mu) + (m\lambda - n\mu)^2)^{1/2} e^{4|\rho|n\mu} \right] |A(v)|. \end{aligned} \quad (5.2.6)$$

Setting $\mu = t/n$ and $\lambda = t/m$ where n and m are integers with $n \geq m$ in the above inequality we obtain

$$\begin{aligned} \|J_{t/n}^n(v) - J_{t/m}^m(v)\| &\leq \left[\left((t-t)^2 + t^2 \left(\frac{1}{m} - \frac{1}{n} \right) \right)^{1/2} e^{2|\rho| \cdot 2t} \right. \\ &\quad \left. + \left(t^2 \left(\frac{1}{m} - \frac{1}{n} \right) + (t-t)^2 \right)^{1/2} e^{4|\rho|t} \right] |A(v)| \\ &= 2te^{4|\rho|t} \left(\frac{1}{m} - \frac{1}{n} \right)^{1/2} |A(v)|. \end{aligned} \quad (5.2.7)$$

This shows that $(J_{t/n}^n(v))_{n \in \mathbb{N}}$ is a Cauchy sequence in X and hence converges in the Banach space X , so that $\lim_{n \rightarrow \infty} J_{t/n}^n(v)$ exists. So now we have shown the existence for $v \in \mathcal{D}(A)$. Next we will extend the result to $v \in \overline{\mathcal{D}(A)}$.

By Lemma 5.2.1(i), $J_{t/n}^n$ has $(1 - \rho t/n)^{-n}$ as Lipschitz constant. Since

$$\lim_{n \rightarrow \infty} \left(1 - \frac{t\rho}{n} \right)^n = e^{-t\rho}$$

then

$$\lim_{n \rightarrow \infty} \left(1 - \frac{t\rho}{n} \right)^{-n} = e^{t\rho}$$

and we see that $S(t)(v) = \lim_{n \rightarrow \infty} J_{t/n}^n(v)$ is Lipschitz continuous in v and has Lipschitz constant $e^{\rho t}$. Hence we can extend $S(t)$ to an operator with Lipschitz constant $e^{\rho t}$ on $\overline{\mathcal{D}(A)}$ in the following way: Let $(v_n)_{n \in \mathbb{N}} \subset \mathcal{D}(A)$ be a sequence converging to $v \in \partial \mathcal{D}(A)$ and let $(\tilde{v}_n)_{n \in \mathbb{N}} \subset \mathcal{D}(A)$ be another sequence converging to the same v . Since $S(t)$ is Lipschitz continuous on $\mathcal{D}(A)$ then $(S(t)(v_n))_{n \in \mathbb{N}} \subset X$ and $(S(t)(\tilde{v}_n))_{n \in \mathbb{N}} \subset X$ are both Cauchy sequences in X and hence converges. Let $S(t)(v_n) \rightarrow w$ and $S(t)(\tilde{v}_n) \rightarrow \tilde{w}$. We show that $w = \tilde{w}$:

$$\begin{aligned} \|w - \tilde{w}\| &= \|w - S(t)(v_n) + S(t)(v_n) - S(t)(\tilde{v}_n) + S(t)(\tilde{v}_n) - \tilde{w}\| \\ &\leq \|w - S(t)(v_n)\| + \|S(t)(v_n) - S(t)(\tilde{v}_n)\| + \|S(t)(\tilde{v}_n) - \tilde{w}\| \\ &\leq \|w - S(t)(v_n)\| + e^{\rho t} \|v_n - \tilde{v}_n\| + \|S(t)(\tilde{v}_n) - \tilde{w}\| \\ &\leq \|w - S(t)(v_n)\| + e^{\rho t} (\|v_n - v\| + \|v - \tilde{v}_n\|) + \|S(t)(\tilde{v}_n) - \tilde{w}\|. \end{aligned}$$

Since $S(t)(v_n) \rightarrow w$, $v_n \rightarrow v$, $\tilde{v}_n \rightarrow v$, and $S(t)(\tilde{v}_n) \rightarrow \tilde{w}$ we conclude from the above that $w = \tilde{w}$. Hence defining $S(t)(v) = w$ we have extended $S(t)$ to an operator with Lipschitz constant $e^{\rho t}$ on $\overline{\mathcal{D}(A)}$. Now we have shown the existence of the limit $\lim_{n \rightarrow \infty} J_{t/n}^n(v)$ for all $v \in \overline{\mathcal{D}(A)}$ and $t > 0$. Now we have to show that $S \in Q_\rho(\overline{\mathcal{D}(A)})$.

We have already shown that $S(t)$ has Lipschitz constant $e^{\rho t}$. So we just need to show that S satisfies the semi-group properties. First of all it is clear that $S(t)$ maps $\overline{\mathcal{D}(A)}$ into $\overline{\mathcal{D}(A)}$ since $J_{t/n}^n$ leaves $\mathcal{D}(A)$ invariant. Let $\tau > t \geq 0$. Then if we put $n = m$, $\mu = t/n$ and $\lambda = \tau/n$ in 5.2.6 we obtain

$$\begin{aligned} \|J_{t/n}^n(v) - J_{\tau/n}^n(v)\| &\leq \left[\left((t - \tau)^2 + \frac{t}{n}(\tau - t) \right)^{1/2} e^{2|\rho|(t+\tau)} \right. \\ &\quad \left. + \left(\frac{\tau}{n}(\tau - t) + (\tau - t)^2 \right)^{1/2} e^{4|\rho|t} \right] |A(v)|. \end{aligned}$$

Taking the limit as $n \rightarrow \infty$ on both sides in the above inequality we obtain

$$\|S(t)(v) - S(\tau)(v)\| \leq (e^{2|\rho|(t+\tau)} + e^{4|\rho|t}) |A(v)|(\tau - t).$$

This shows that $S(t)(v)$ is Lipschitz continuous in t on bounded t -sets for $v \in \overline{\mathcal{D}(A)}$ and hence $S(t)(v)$ is continuous in t . Setting $t = 0$ in the above and letting $\tau \rightarrow 0$ we obtain $\lim_{\tau \searrow 0} S(\tau)(v) = S(0)(v) = v$.

Finally we have to show that $S(t + \tau) = S(t)S(\tau)$ for $t, \tau \geq 0$. We use the operator convergence

$$S(t) = \lim_{n \rightarrow \infty} J_{t/n}^n$$

and the uniform Lipschitz continuity of $(J_{t/n^n})_{n>N}$ for sufficiently large N to conclude that

$$(S(t))^m = \left(\lim_{n \rightarrow \infty} J_{t/n^n} \right)^m = \lim_{n \rightarrow \infty} ((J_{t/n^n})^m) = \lim_{n \rightarrow \infty} ((J_{t/n^m})^n).$$

This implies that for an integer m

$$S(mt) = \lim_{n \rightarrow \infty} (J_{mt/n^n}) = \lim_{k \rightarrow \infty} (J_{mt/(mk)^{mk}}) = \lim_{k \rightarrow \infty} ((J_{t/k^m})^k) = (S(t))^m.$$

Now if we let $l, k, r,$ and s be positive integers. Then using the above we obtain

$$\begin{aligned} S\left(\frac{l}{k} + \frac{r}{s}\right) &= S\left(\frac{ls + rk}{ks}\right) = \left(S\left(\frac{1}{ks}\right)\right)^{ls+rk} = \left(S\left(\frac{1}{ks}\right)\right)^{ls} \left(S\left(\frac{1}{ks}\right)\right)^{rk} \\ &= S\left(\frac{l}{k}\right) S\left(\frac{r}{s}\right). \end{aligned}$$

This shows that the property $S(t + \tau) = S(t)S(\tau)$ holds for rational t and τ . Now for general $t > 0$ and $\tau > 0$ we can approximate both t and τ by sequences of rationals. Hence let $(t_n)_{n \in \mathbb{N}} \subset \mathbb{Q}$ and $(\tau_n)_{n \in \mathbb{N}} \subset \mathbb{Q}$ be such that $t_n \rightarrow t$ and $\tau_n \rightarrow \tau$. Then since $S(\cdot)v$ is continuous for all $v \in \overline{\mathcal{D}(A)}$ we obtain

$$S(t + \tau)v = \lim_{n \rightarrow \infty} S(t_n + \tau_n)v = \lim_{n \rightarrow \infty} S(t_n)S(\tau_n)v = S(t)S(\tau)v$$

for all $v \in \overline{\mathcal{D}(A)}$. This shows that $S(t + \tau) = S(t)S(\tau)$ for all $t > 0, \tau > 0$. \square

5.2.2 The Cauchy Problem

In this section we consider the existence of a solution to the Cauchy problem

$$0 \in \frac{\partial u}{\partial t} + Au, \quad u(0) = u_0 \tag{5.2.8}$$

for an operator $A \subset X \times X$ and a given $u_0 \in \mathcal{D}(A)$. In order to discuss the existence of solutions, a reasonable notion of a solution is required. Here we motivate the definition of a *strong solution* to the problem: First of all, a solution to the Cauchy problem should of course satisfy $u(0) = u_0$ and the differential equation should be satisfied at least for almost every $t \in [0, T)$ for some $T > 0$. Finally it would be nice to be able to write an expression for u . Therefore we require u to be able to be written as an integral expression and furthermore to be continuous. Note that since u takes on values in the Banach space X then writing u as an integral expression requires the ability to integrate over values in X . Here we use the Bochner integral which is defined in the same way as the Lebesgue integral but for

values in a Banach space instead. It can be shown that a function $v : [0, T) \rightarrow X$ is Bochner integrable (strongly integrable) if and only if

$$\int_0^T \|v(t)\|_X dt < +\infty,$$

see [29, Section 3.3.1]. Hence we require u to be able to be written as an integral over a Bochner integrable function:

Definition 5.2.5 (Strong Solution). Let $0 < T \leq \infty$ and $u : [0, T) \rightarrow X$. Then u is a *strong solution* of the Cauchy problem (5.2.8) on $[0, T)$ if

- (i) u is continuous,
- (ii) u is the indefinite integral of a function which is Bochner integrable (strongly integrable) on compact subsets of $(0, T)$,
- (iii) $u(0) = u_0$,
- (iv) $u'(t) \in -A(u)(t)$ for almost all $t \in [0, T)$.

For the rest of this section fix $A \subset X \times X$ and $\rho \in \mathbb{R}$ such that $A + I\rho$ is accretive. Then the following equivalent formulation of a strong solution can be proved:

Lemma 5.2.6 (Strong Solution). Let $A \subset X \times X$ and $\rho \in \mathbb{R}$ be such that $A + I\rho$ is accretive and let $u_0 \in \mathcal{D}(A)$. Then a function $u : [0, T) \rightarrow X$ is a *strong solution* of the Cauchy problem (5.2.8) if and only if u is Lipschitz continuous on compact subsets of $[0, T)$, u is differentiable almost everywhere on $[0, T)$ and (5.2.8) is satisfied almost everywhere.

Proof. See [12, lemma 6.2]. □

In order to prove the existence of a strong solution to the Cauchy problem (5.2.8) we need some preliminary results. Since X is a Banach space it does not necessarily have an inner product. But we need some kind of “inner product” notion. In order to be able to define such a notion on a Banach space we first define the *duality mapping*. The duality mapping $\mathcal{J} : X \rightarrow X'$ associates with each $v \in X$ one or more duals $v' \in X'$ in the following way

$$\mathcal{J}(v) = \{v' \in X' : (v, v') = \|v\|^2 = \|v'\|^2\}$$

where the pairing (v, v') is the value of $v' \in X'$ at $v \in X$. Now if X is a Hilbert space we can identify X' with X and then the pairing (v, v') can be identified with an inner product between v and v' on X . Hence for a Hilbert space the duality

mapping \mathcal{J} is essentially the identity mapping. In a general Banach space X we know that $\mathcal{J}(v)$ is non-empty for each $v \in X$ by the geometric Hahn-Banach Theorem (see [22, Thm. 4.3-3]): For any $v \in X$ with $v \neq 0$ there exists $w' \in X'$ such that $\|w'\| = 1$ and $(v, w') = \|v\|$. Let $v' = \|v\|w' \in X'$. Then

$$\|v'\| = \|v\|\|w'\| = \|v\|$$

and

$$(v, v') = (v, \|v\|w') = \|v\|(v, w') = \|v\|^2.$$

This shows that $v' \in \mathcal{J}(v)$. For $v = 0$ we have $0 \in \mathcal{J}(v)$. This shows that $\mathcal{J}(v)$ is non-empty for each $v \in X$. We also see that $\mathcal{J}(v)$ is bounded since

$$\begin{aligned} \text{diam}(\mathcal{J}(v)) &= \sup_{v'_1, v'_2 \in \mathcal{J}(v)} \|v'_1 - v'_2\| \leq \sup_{v'_1, v'_2 \in \mathcal{J}(v)} \|v'_1\| + \|v'_2\| \\ &= \sup_{v'_1, v'_2 \in \mathcal{J}(v)} 2\|v\| = 2\|v\| < \infty. \end{aligned}$$

Finally we show that $\mathcal{J}(v)$ is weak* compact: Take a sequence $(v_n')_{n \in \mathbb{N}} \subset \mathcal{J}(v)$ and assume that $v_n' \xrightarrow{*} v' \in X'$ as $n \rightarrow \infty$, i.e.

$$(v, v_n') \rightarrow (v, v'), \quad \forall v \in X.$$

Then by the lower semi-continuity of the norm on X' (see [20, Prop. 2.4.12]), and using $v_n' \in \mathcal{J}(v)$ we get

$$\|v'\| \leq \liminf_{n \rightarrow \infty} \|v_n'\| = \|v\|.$$

On the other hand

$$\|v\|^2 = (v, v_n') \rightarrow (v, v') \leq \|v\|\|v'\|.$$

Using the above and $v_n' \in \mathcal{J}(v)$ we then have

$$(v, v') = \|v'\|^2 = \|v\|^2.$$

This implies that $v' \in \mathcal{J}(v)$ and hence $\mathcal{J}(v)$ is closed in X' for the weak* topology. Finally, since the unit ball in X' is weak* compact (see Banach-Alaoglu-Bourbaki Theorem in [20, Thm. 2.4.7]), then so is $\mathcal{J}(v)$ because $\mathcal{J}(v)$ is contained in a scaled unit ball and is closed in the weak* topology.

We are now ready to define what should be the alternative to an inner product on a Banach space X . With the above observations about the duality mapping \mathcal{J} the functions $\langle \cdot, \cdot \rangle_i : X \times X \rightarrow \mathbb{R}$ and $\langle \cdot, \cdot \rangle_s : X \times X \rightarrow \mathbb{R}$ defined below are both well-defined and finite-valued. For $v, w \in X$ we define

$$\langle v, w \rangle_i = \inf\{(v, w') : w' \in \mathcal{J}(w)\} \quad \text{and} \quad \langle v, w \rangle_s = \sup\{(v, w') : w' \in \mathcal{J}(w)\}.$$

By the definition of supremum and infimum we have the relation

$$\langle v, w \rangle_s = -\langle -v, w \rangle_i.$$

First we list and prove some elementary properties of the functions $\langle \cdot, \cdot \rangle_i$ and $\langle \cdot, \cdot \rangle_s$. For the proofs of property (e) and (f), though, we just refer to an article containing the proofs:

Lemma 5.2.7 (Elementary properties of $\langle \cdot, \cdot \rangle_i$ and $\langle \cdot, \cdot \rangle_s$). *Let $v, w \in X$. Then*

(a) $\langle \alpha w + v, w \rangle_j = \alpha \|w\|^2 + \langle v, w \rangle_j$ for $\alpha \in \mathbb{R}$ and $j = i$ or $j = s$.

(b) $\langle \beta v, \gamma w \rangle_j = \gamma \beta \langle v, w \rangle_j$ for $\gamma \beta \geq 0$ and $j = i$ or $j = s$.

(c) $\langle z + v, w \rangle_j \leq \|z\| \|w\| + \langle v, w \rangle_j$ for $z \in X$ and $j = i$ or $j = s$.

(d) $\langle \cdot, \cdot \rangle_s : X \times X \rightarrow \mathbb{R}$ is upper semi-continuous.

(e) $B \subset X \times X$ is accretive if and only if

$$\langle w_1 - w_2, v_1 - v_2 \rangle_s \geq 0 \text{ for } [v_i, w_i] \in B, i = 1, 2.$$

(f) If $v : (a, b) \rightarrow X$ then

$$\left. \frac{d}{dt} \|v(t)\|^2 \right|_{t=t_0} = 2 \langle Dv(t_0), v(t_0) \rangle_s = 2 \langle Dv(t_0), v(t_0) \rangle_i$$

at each point t_0 at which $\|v(t)\|^2$ is differentiable and $v(t)$ is weakly differentiable.

Proof. Proof of (a): Let $v, w \in X$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} \langle \alpha w + v, w \rangle_i &= \inf\{(\alpha w + v, w') : w' \in \mathcal{J}(w)\} \\ &= \inf\{\alpha(w, w') + (v, w') : w' \in \mathcal{J}(w)\} \\ &= \inf\{\alpha \|w\|^2 + (v, w') : w' \in \mathcal{J}(w)\} \\ &= \alpha \|w\|^2 + \inf\{(v, w') : w' \in \mathcal{J}(w)\} \\ &= \alpha \|w\|^2 + \langle v, w \rangle_i. \end{aligned}$$

The same proof goes for $j = s$.

Proof of (b): Let $v, w \in X$ and $\gamma\beta \geq 0$. Then

$$\begin{aligned} \langle \beta v, \gamma w \rangle_i &= \inf\{(\beta v, w') : w' \in \mathcal{J}(\gamma w)\} = \inf\{\beta(v, w') : w' \in \mathcal{J}(\gamma w)\} \\ &= \inf\{\beta(v, \gamma w') : w' \in \mathcal{J}(w)\} = \inf\{\beta\gamma(v, w') : w' \in \mathcal{J}(w)\} \\ &= \beta\gamma \inf\{(v, w') : w' \in \mathcal{J}(w)\} = \beta\gamma \langle v, w \rangle_i. \end{aligned}$$

The same proof goes for $j = s$.

Proof of (c): Let $v, w, z \in X$. Then

$$\begin{aligned} \langle z + v, w \rangle_i &= \inf\{(z + v, w') : w' \in \mathcal{J}(w)\} \\ &= \inf\{(z, w') + (v, w') : w' \in \mathcal{J}(w)\} \\ &\leq \inf\{\|z\|\|w'\| + (v, w') : w' \in \mathcal{J}(w)\} \\ &= \inf\{\|z\|\|w\| + (v, w') : w' \in \mathcal{J}(w)\} \\ &= \|z\|\|w\| + \inf\{(v, w') : w' \in \mathcal{J}(w)\} = \|z\|\|w\| + \langle v, w \rangle_i. \end{aligned}$$

The same proof goes for $j = s$.

Proof of (d): We want to show that if $v_n \rightarrow v_0$ and $w_n \rightarrow w_0$ then

$$\limsup_{n \rightarrow \infty} \langle v_n, w_n \rangle_s \leq \langle v_0, w_0 \rangle_s.$$

Hence let $v_n \rightarrow v_0$ and $w_n \rightarrow w_0$. Since $\mathcal{J}(w_n)$ is weak* compact there exists a $\zeta_n' \in \mathcal{J}(w_n)$ such that

$$\langle v_n, w_n \rangle_s = (v_n, \zeta_n').$$

Let

$$\beta_n = \sup_{N > n} \langle v_N, w_N \rangle_s.$$

Then

$$\lim_{n \rightarrow \infty} \beta_n = \lim_{n \rightarrow \infty} \sup_{N > n} \langle v_N, w_N \rangle_s = \lim_{n \rightarrow \infty} \sup_{N > n} (v_N, \zeta_N') = \lim_{n \rightarrow \infty} \sup_{N > n} (v_N, \zeta_N')$$

exists. Since $\|\zeta_n'\| = \|w_n\| \rightarrow \|w_0\|$ as $n \rightarrow \infty$, the sequence $(\|\zeta_n'\|)_{n \in \mathbb{N}}$ is bounded, now which implies that (ζ_n') is weak* convergent. Hence let ζ_0' be any weak* limit point of $(\zeta_n')_{n \in \mathbb{N}}$. Since the norm on X' is weak* lower semi-continuous (see [20, Prop. 2.4.12]) we obtain

$$\|\zeta_0'\| \leq \liminf_{n \rightarrow \infty} \|\zeta_n'\| = \liminf_{n \rightarrow \infty} \|w_n\| = \|w_0\|$$

and using that ζ_n' is bounded for each n we also have

$$\|w_0\|^2 = \lim_{n \rightarrow \infty} \|w_n\|^2 = \lim_{n \rightarrow \infty} (w_n, \zeta_n') = (w_0, \zeta_0').$$

since $w_n \rightarrow w_0$ in the norm topology. Finally we obtain from the above

$$\|w_0\|^2 = (w_0, \zeta_0') \leq \|w_0\| \|\zeta_0'\|.$$

Together the three above (in)equalities imply

$$(w_0, \zeta_0') = \|w_0\|^2 = \|\zeta_0'\|^2$$

so that $\zeta_0' \in \mathcal{J}(w_0)$. Now, since $v_n \rightarrow v_0$ in the norm topology then

$$\lim_{n \rightarrow \infty} \beta_n = \limsup_{n \rightarrow \infty} \langle v_n, w_n \rangle_s = \limsup_{n \rightarrow \infty} (v_n, \zeta_n') = (v_0, \zeta_0') \leq \langle v_0, w_0 \rangle_s.$$

This was exactly what we wanted to show.

The proofs of (e) and (f) can be found in [11]. \square

Remark 5.2.8. Note that assertions (e) and (f) of Lemma 5.2.7 are well-known results in Hilbert spaces. An accretive operator in a Hilbert space is a monotonic operator which is exactly defined by the property in (e). For assertion (f) the definition of $\langle \cdot, \cdot \rangle_s$ is essentially equal to the inner product on a Hilbert space and here we have seen the result before.

In the end we would like to show that

$$S(t)(u_0) = \lim_{n \rightarrow \infty} J_{t/n}^n(u_0)$$

is a strong solution to the Cauchy problem in (5.2.8). In order to show this we first approximate the Cauchy problem by some similar problems. Hence let S_λ be the semi-groups on $\overline{\mathcal{D}(A)}$ defined by

$$\frac{d}{dt} S_\lambda(t)(v) + A_\lambda S_\lambda(t)(v) = 0 \quad (5.2.9)$$

for $t \geq 0$ and $0 < \lambda \leq \lambda_0$, where

$$A_\lambda(v) = \lambda^{-1}(v - J_\lambda(v)).$$

Note that here A_λ is single-valued and hence we can use the equality sign in (5.2.9). The existence of S_λ satisfying the approximate Cauchy problem (5.2.9) is a consequence of the assumption

$$\mathcal{R}(I + \lambda A) \supset \overline{\mathcal{D}(A)}$$

and is proved in [7, pp. 246-248]. The next thing we need to know is how well $S_\lambda(t)$ approximates $S(t)$. For this we have the result:

Lemma 5.2.9. *Let $A \subset X \times X$ and $\rho \in \mathbb{R}$ be such that $A + \rho I$ is accretive. Assume that $\mathcal{R}(I + \lambda A) \supset \overline{\mathcal{D}(A)}$ for sufficiently small $\lambda > 0$, and that A is a closed subset of $X \times X$. Take $v \in \mathcal{D}(A)$, $0 < \lambda \leq \lambda_0$, and $\lambda|\rho| < 1/2$. Then*

$$\|S_\lambda(t)(v) - S(t)(v)\| \leq \left[e^{2t|\rho|} \lambda (1 - \lambda\rho)^{-1} + 2e^{4t|\rho|} \sqrt{t\lambda} + e^{4t|\rho|} (4(t|\rho|)^2 \lambda + 2t|\rho|\lambda + t)^{1/2} \sqrt{\lambda} (1 - \lambda\rho)^{-1} + e^{t|\rho|} \lambda \right] |A(v)|.$$

Proof. Choosing the integer m such that $t = m\lambda + \delta$ and $0 \leq \delta < \lambda$ we use the triangle inequality to estimate

$$\begin{aligned} \|S_\lambda(t)(v) - S(t)(v)\| &\leq \|S_\lambda(t)(v) - S_\lambda(m\lambda)(v)\| + \|S_\lambda(m\lambda)(v) - J_\lambda^m(v)\| \\ &\quad + \|J_\lambda^m(v) - S(m\lambda)(v)\| + \|S(m\lambda)(v) - S(t)(v)\|. \end{aligned} \tag{5.2.10}$$

Now we estimate each of the four norms in the inequality above. Using $S \in Q_\rho(\overline{\mathcal{D}(A)})$ and using properties (ii) and (iii) in Lemma 5.2.1 the last term can be estimated by

$$\begin{aligned} \|S(m\lambda)(v) - S(t)(v)\| &= \|S(m\lambda)(v) - S(m\lambda + \delta)(v)\| \\ &= \|S(m\lambda)(v) - S(m\lambda)S(\delta)(v)\| \\ &\leq e^{\rho m\lambda} \|v - S(\delta)(v)\| = e^{\rho m\lambda} \|v - \lim_{n \rightarrow \infty} J_{\delta/n}^n(v)\| \\ &= e^{\rho m\lambda} \lim_{n \rightarrow \infty} \|v - J_{\delta/n}^n(v)\| \\ &\stackrel{(iii)}{\leq} e^{\rho m\lambda} \lim_{n \rightarrow \infty} \left(n \left(1 - \frac{\delta}{n} |\rho| \right)^{-n+1} \|J_{\delta/n}^n(v) - v\| \right) \\ &\stackrel{(ii)}{\leq} e^{\rho m\lambda} \lim_{n \rightarrow \infty} \left(n \left(1 - \frac{\delta}{n} |\rho| \right)^{-n+1} \frac{\delta}{n} \left(1 - \frac{\delta}{n} \rho \right)^{-1} |A(v)| \right) \\ &\leq e^{\rho m\lambda} \delta |A(v)| \lim_{n \rightarrow \infty} \left(\left(1 - \frac{\delta}{n} |\rho| \right)^{-n+1} \left(1 - \frac{\delta}{n} |\rho| \right)^{-1} \right) \\ &= e^{\rho m\lambda} \delta |A(v)| \lim_{n \rightarrow \infty} \left(1 - \frac{\delta}{n} |\rho| \right)^{-n} = e^{\rho m\lambda} \delta |A(v)| e^{\delta|\rho|} \\ &\leq e^{|\rho|(m\lambda + \delta)} \delta |A(v)| < e^{t|\rho|} \lambda |A(v)|. \end{aligned}$$

The first term admits a similar estimate. Using Lemma 5.2.7 and Lemma 5.2.1 we

can estimate

$$\begin{aligned}
& \frac{d}{dt} \|S_\lambda(t)(z) - S_\lambda(t)(w)\|^2 \\
& \stackrel{5.2.7(f)}{=} 2 \left\langle \frac{d}{dt} S_\lambda(t)(z) - \frac{d}{dt} S_\lambda(t)(w), S_\lambda(t)(z) - S_\lambda(t)(w) \right\rangle_s \\
& = 2 \langle A_\lambda S_\lambda(t)(w) - A_\lambda S_\lambda(t)(z), S_\lambda(t)(z) - S_\lambda(t)(w) \rangle_s \\
& = 2 \langle \lambda^{-1}(S_\lambda(t)(w) - J_\lambda S_\lambda(t)(w)) - \lambda^{-1}(S_\lambda(t)(z) - J_\lambda S_\lambda(t)(z)), \\
& \quad S_\lambda(t)(z) - S_\lambda(t)(w) \rangle_s \\
& \stackrel{5.2.7(b)}{=} 2\lambda^{-1} \langle (S_\lambda(t)(w) - S_\lambda(t)(z)) + (J_\lambda S_\lambda(t)(z) - J_\lambda S_\lambda(t)(w)), \\
& \quad S_\lambda(t)(z) - S_\lambda(t)(w) \rangle_s \\
& \stackrel{5.2.7(c)}{=} 2\lambda^{-1} (\|J_\lambda S_\lambda(t)(z) - J_\lambda S_\lambda(t)(w)\| \|S_\lambda(t)(z) - S_\lambda(t)(w)\| \\
& \quad \langle -(S_\lambda(t)(z) - S_\lambda(t)(w)), S_\lambda(t)(z) - S_\lambda(t)(w) \rangle_s) \\
& \stackrel{5.2.7(a)}{=} 2\lambda^{-1} (\|J_\lambda S_\lambda(t)(z) - J_\lambda S_\lambda(t)(w)\| \|S_\lambda(t)(z) - S_\lambda(t)(w)\| \\
& \quad - \|S_\lambda(t)(z) - S_\lambda(t)(w)\|^2) \\
& = 2\lambda^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| (\|J_\lambda S_\lambda(t)(z) - J_\lambda S_\lambda(t)(w)\| \\
& \quad - \|S_\lambda(t)(z) - S_\lambda(t)(w)\|) \\
& \stackrel{5.2.1(i)}{\leq} 2\lambda^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| ((1 - \lambda\rho)^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| \\
& \quad - \|S_\lambda(t)(z) - S_\lambda(t)(w)\|) \\
& = 2\lambda^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\|^2 ((1 - \lambda\rho)^{-1} - 1).
\end{aligned}$$

By the above inequality we get

$$\begin{aligned}
2\|S_\lambda(t)(z) - S_\lambda(t)(w)\| \frac{d}{dt} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| \\
& = \frac{d}{dt} \|S_\lambda(t)(z) - S_\lambda(t)(w)\|^2 \\
& \leq 2\lambda^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\|^2 ((1 - \lambda\rho)^{-1} - 1)
\end{aligned}$$

Rearranging and using $\lambda|\rho| < 1/2$ we reach at

$$\begin{aligned}
\frac{d}{dt} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| & \leq \lambda^{-1} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| ((1 - \lambda\rho)^{-1} - 1) \\
& \leq 2\rho \|S_\lambda(t)(z) - S_\lambda(t)(w)\|.
\end{aligned}$$

Multiplying both sides of the above inequality by $e^{-2\rho t}$, we obtain

$$\begin{aligned} & \frac{d}{dt} (e^{-2\rho t} \|S_\lambda(t)(z) - S_\lambda(t)(w)\|) \\ &= e^{-2\rho t} \frac{d}{dt} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| - 2\rho e^{-2\rho t} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| \leq 0. \end{aligned}$$

Integrating from 0 to t we then get

$$e^{-2\rho t} \|S_\lambda(t)(z) - S_\lambda(t)(w)\| - \|z - w\| \leq 0$$

and hence

$$\|S_\lambda(t)(z) - S_\lambda(t)(w)\| \leq e^{2\rho t} \|z - w\|.$$

So $S_\lambda \in Q_{2\rho}(\overline{\mathcal{D}(A)})$. Also using Lemma 5.2.1(ii) we have

$$\begin{aligned} \left\| \frac{d}{dt} S_\lambda(t)(v) \Big|_{t=0} \right\| &= \|-A_\lambda S_\lambda(0)(v)\| = \|A_\lambda(v)\| = \|\lambda^{-1}(v - J_\lambda(v))\| \\ &\stackrel{5.2.1(ii)}{\leq} \frac{1}{\lambda} \lambda(1 - \lambda\rho)^{-1} |A(v)| = (1 - \lambda\rho)^{-1} |A(v)|. \end{aligned}$$

So that $(1 - \lambda\rho)^{-1} |A(v)|$ is a Lipschitz constant (in t) for $S_\lambda(t)(v)$. The above considerations lead to the final estimate for the first norm term in (5.2.10):

$$\begin{aligned} \|S_\lambda(t)(v) - S_\lambda(m\lambda)(v)\| &= \|S_\lambda(m\lambda + \delta)(v) - S_\lambda(m\lambda)(v)\| \\ &= \|S_\lambda(m\lambda)S_\lambda(\delta)(v) - S_\lambda(m\lambda)(v)\| \\ &\leq e^{2\rho m\lambda} \|S_\lambda(\delta)(v) - v\| = e^{2\rho m\lambda} \|S_\lambda(\delta)(v) - S_\lambda(0)(v)\| \\ &\leq e^{2\rho m\lambda} (1 - \lambda\rho)^{-1} |A(v)| \delta \leq e^{2|\rho|(m\lambda + \delta)} (1 - \lambda\rho)^{-1} |A(v)| \lambda \\ &= e^{2t|\rho|} \lambda (1 - \lambda\rho)^{-1} |A(v)|. \end{aligned}$$

The third norm term is estimated via (5.2.7) in the proof of Theorem 5.2.4:

$$\|J_{t/n}^n(v) - J_{t/m}^m(v)\| \leq 2te^{4|\rho|t} \left(\frac{1}{m} - \frac{1}{n} \right)^{1/2} |A(v)|.$$

This gives

$$\begin{aligned}
\|J_\lambda^m(v) - S(m\lambda)(v)\| &= \|J_\lambda^m(v) - \lim_{n \rightarrow \infty} J_{m\lambda/n}^n(v)\| \\
&= \|\lim_{n \rightarrow \infty} (J_\lambda^m(v) - J_{m\lambda/n}^n(v))\| \\
&= \lim_{n \rightarrow \infty} \|J_\lambda^m(v) - J_{m\lambda/n}^n(v)\| \\
&\leq \lim_{n \rightarrow \infty} \left(2m\lambda e^{4|\rho|m\lambda} \left(\frac{1}{m} - \frac{1}{n} \right)^{1/2} |A(v)| \right) \\
&= 2m\lambda e^{4|\rho|m\lambda} (1/\sqrt{m}) |A(v)| = 2\sqrt{m}\lambda e^{4|\rho|m\lambda} |A(v)| \\
&= 2\sqrt{\lambda(\lambda m)} e^{4|\rho|m\lambda} |A(v)| \leq 2\sqrt{\lambda(m\lambda + \delta)} e^{4|\rho|(m\lambda + \delta)} |A(v)| \\
&= 2\sqrt{\lambda t} e^{4|\rho|t} |A(v)|.
\end{aligned}$$

Finally, we estimate the second norm term by using Lemma 4' in the appendix of [8]. The lemma states that

$$\begin{aligned}
\|S_\lambda(m\lambda)(v) - J_\lambda^m(v)\| &\leq (1 - \lambda\rho)^{-m} e^{m((1-\lambda\rho)^{-1}-1)} \left[m^2 ((1 - \lambda\rho)^{-1} - 1)^2 \right. \\
&\quad \left. + m (((1 - \lambda\rho)^{-1} - 1) + m)^{1/2} \right] \|J_\lambda(v) - v\|.
\end{aligned} \tag{5.2.11}$$

Using $\lambda|\rho| < 1/2$ we observe that

$$\begin{aligned}
m((1 - \lambda\rho)^{-1} - 1) &= m \left(\frac{1 - (1 - \lambda\rho)}{1 - \lambda\rho} \right) = \frac{m\lambda}{1 - \lambda\rho} \\
&\leq \frac{m\lambda + \delta}{1 - \lambda|\rho|} |\rho| = \frac{1}{1 - \lambda|\rho|} |\rho|t \\
&< 2|\rho|t.
\end{aligned}$$

As shown in the proof of Theorem 5.2.4, for $\lambda\rho \leq \lambda|\rho| < 1/2$ we have

$$(1 - \lambda\rho)^{-m} \leq e^{2m\lambda\rho} \leq e^{2(m\lambda + \delta)|\rho|} = e^{2t|\rho|}.$$

Applying the two above considerations and Lemma 5.2.1(ii) in (5.2.11) the following estimate for the second norm term is obtained:

$$\begin{aligned}
\|S_\lambda(m\lambda)(v) - J_\lambda^m(v)\| &\leq e^{2t|\rho|} e^{2t|\rho|} (4(t|\rho|)^2 + 2t|\rho| + m)^{1/2} \lambda(1 - \lambda\rho)^{-1} |A(v)| \\
&= e^{4t|\rho|} (4(t|\rho|)^2 \lambda + 2t|\rho|\lambda + m\lambda)^{1/2} \sqrt{\lambda}(1 - \lambda\rho)^{-1} |A(v)| \\
&\leq e^{4t|\rho|} (4(t|\rho|)^2 \lambda + 2t|\rho|\lambda + t)^{1/2} \sqrt{\lambda}(1 - \lambda\rho)^{-1} |A(v)|.
\end{aligned}$$

Collecting the four norm estimates gives us exactly the final desired estimate. \square

From Lemma 5.2.9 we see that $S_\lambda(t)(v) \rightarrow S(t)(v)$ as $\lambda \searrow 0$. The next and final lemma of this section will be the main step in the proof of the existence of a solution to the Cauchy problem and will be used to show that $S(t)(u_0)$ satisfies the Cauchy problem whenever it is strongly differentiable:

Lemma 5.2.10. *Let $A \subset X \times X$ and $\rho \in \mathbb{R}$ be such that $A + \rho I$ is accretive. Assume that $\mathcal{R}(I + \lambda A) \supset \overline{\mathcal{D}(A)}$ for sufficiently small $\lambda > 0$ with $\lambda\rho \leq 1$, and that A is a closed subset of $X \times X$. Put*

$$S(t)v = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} v$$

for $t \geq 0$ and $v \in \overline{\mathcal{D}(A)}$. If $v \in \overline{\mathcal{D}(A)}$ and $[v_0, w_0] \in A$ then

$$\sup_{\zeta' \in \mathcal{J}(v-v_0)} \limsup_{t \searrow 0} \left(\frac{S(t)v - v}{t} + \rho(v_0 - v), \zeta' \right) \leq \langle w_0, v_0 - v \rangle_s. \quad (5.2.12)$$

Proof. Having chosen $[v_0, w_0] \in A$ we set

$$v_\lambda = v_0 + \lambda w_0.$$

Then using the properties in Lemma 5.2.7 we obtain

$$\begin{aligned} \frac{d}{dt} \|S_\lambda(t)(v) - v_\lambda\|^2 &= 2 \left\langle \frac{d}{dt} S_\lambda(t)(v), S_\lambda(t)(v) - v_\lambda \right\rangle_i \\ &= 2 \langle -A_\lambda S_\lambda(t)(v), S_\lambda(t)(v) - v_\lambda \rangle_i \\ &= 2 \langle -A_\lambda S_\lambda(t)(v), S_\lambda(t)(v) - v_\lambda \rangle_i + 2 \langle A_\lambda(v_\lambda), S_\lambda(t)(v) - v_\lambda \rangle_i \\ &\quad - 2 \langle A_\lambda(v_\lambda), S_\lambda(t)(v) - v_\lambda \rangle_i \\ &= -2 \langle A_\lambda S_\lambda(t)(v), S_\lambda(t)(v) - v_\lambda \rangle_s \\ &\quad - 2 \langle -A_\lambda(v_\lambda), S_\lambda(t)(v) - v_\lambda \rangle_s + 2 \langle -A_\lambda(v_\lambda), S_\lambda(t)(v) - v_\lambda \rangle_s \\ &= -2 \langle A_\lambda S_\lambda(t)(v) - A_\lambda(v_\lambda), S_\lambda(t)(v) - v_\lambda \rangle_s \\ &\quad + 2 \langle A_\lambda(v_\lambda), v_\lambda - S_\lambda(t)(v) \rangle_s. \end{aligned}$$

Recalling that $A_\lambda = \lambda^{-1}(I - J_\lambda)$ and that J_λ has $(1 - \lambda\rho)^{-1}$ as Lipschitz constant

and using $\lambda\rho \leq 1$ we obtain

$$\begin{aligned}
\langle A_\lambda(v) - A_\lambda(w), v - w \rangle_s &\geq \langle A_\lambda(v) - A_\lambda(w), v - w \rangle_i \\
&= \langle \lambda^{-1}(v - J_\lambda(v)) - \lambda^{-1}(w - J_\lambda(w)), v - w \rangle_i \\
&= \lambda^{-1} \langle (v - w) - (J_\lambda(v) - J_\lambda(w)), v - w \rangle_i \\
&\geq \lambda^{-1} (\langle v - w, v - w \rangle_i - \|J_\lambda(v) - J_\lambda(w)\| \|v - w\|) \\
&\geq \lambda^{-1} (\|v - w\|^2 - (1 - \lambda\rho)^{-1} \|v - w\|^2) \\
&= \lambda^{-1} \|v - w\|^2 \left(\frac{1 - \lambda\rho - 1}{1 - \lambda\rho} \right) \\
&= -\rho(1 - \lambda\rho)^{-1} \|v - w\|^2.
\end{aligned}$$

Using the above and

$$\begin{aligned}
A_\lambda(v_\lambda) &= \lambda^{-1}(I - J_\lambda)(v_0 + \lambda w_0) = \lambda^{-1}(v_0 + \lambda w_0) - \lambda^{-1}J_\lambda(v_0 + \lambda w_0) \\
&= \lambda^{-1}(v_0 + \lambda w_0) - \lambda^{-1}v_0 = w_0
\end{aligned}$$

we then obtain

$$\frac{d}{dt} \|S_\lambda(t)(v) - v_\lambda\|^2 \leq 2\rho(1 - \lambda\rho)^{-1} \|S_\lambda(t)(v) - v_\lambda\|^2 + 2 \langle w_0, v_\lambda - S_\lambda(t)(v) \rangle_s.$$

Multiplying both sides of the above inequality by $e^{-2\rho t/(1-\lambda\rho)}$, we obtain

$$\begin{aligned}
&\frac{d}{dt} (e^{-2\rho t/(1-\lambda\rho)} \|S_\lambda(t)(v) - v_\lambda\|^2) \\
&= e^{-2\rho t/(1-\lambda\rho)} \frac{d}{dt} \|S_\lambda(t)(v) - v_\lambda\|^2 - 2\rho(1 - \lambda\rho)^{-1} e^{-2\rho t/(1-\lambda\rho)} \|S_\lambda(t)(v) - v_\lambda\|^2 \\
&\leq 2e^{-2\rho t/(1-\lambda\rho)} \langle w_0, v_\lambda - S_\lambda(t)(v) \rangle_s.
\end{aligned}$$

Integrating from 0 to t we get

$$\begin{aligned}
e^{-2\rho t/(1-\lambda\rho)} \|S_\lambda(t)(v) - v_\lambda\|^2 - \|v - v_\lambda\|^2 \\
\leq 2 \int_0^t \langle w_0, v_\lambda - S_\lambda(\tau)(v) \rangle_s e^{-2\rho\tau/(1-\lambda\rho)} d\tau.
\end{aligned}$$

Note that the integral on the right hand side above is well-defined since the integrand is upper semi-continuous by Lemma 5.2.7(d) and hence measurable and integrable. Now $v_\lambda = v_0 + \lambda w_0 \rightarrow v_0$ as $\lambda \searrow 0$ and Lemma 5.2.9 implies that $S_\lambda(t)(v) \rightarrow S(t)(v)$ as $\lambda \searrow 0$ (both convergences in the norm topology). Using Lemma 5.2.7(d) and the reverse Fatou's Lemma, we can take the limit superior on

both sides as $\lambda \searrow 0$ in the above inequality and obtain

$$\begin{aligned}
e^{-2\rho t} \|S(t)(v) - v_0\|^2 - \|v - v_0\|^2 & \\
& \leq \limsup_{\lambda \searrow 0} \left(2 \int_0^t \langle w_0, v_\lambda - S_\lambda(\tau)(v) \rangle_s e^{-2\rho\tau/(1-\lambda\rho)} d\tau \right) \\
& \leq 2 \int_0^t \langle w_0, v_0 - S(\tau)(v) \rangle_s e^{-2\rho\tau} d\tau \\
& = 2t \int_0^1 \langle w_0, v_0 - S(t\tau)(v) \rangle_s e^{-2\rho t\tau} d\tau. \tag{5.2.13}
\end{aligned}$$

Finally if $\zeta' \in \mathcal{J}(v - v_0)$ then

$$\begin{aligned}
2(S(t)(v) - v, \zeta') & = 2(S(t)(v) - v_0 + v_0 - v, \zeta') \\
& = 2(S(t)(v) - v_0, \zeta') - 2(v - v_0, \zeta') \\
& \leq 2\|S(t)(v) - v_0\| \|\zeta'\| - 2\|v - v_0\|^2 \\
& = 2\|S(t)(v) - v_0\| \|v - v_0\| - 2\|v - v_0\|^2 \\
& \leq \|S(t)(v) - v_0\|^2 + \|v - v_0\|^2 - 2\|v - v_0\|^2 \\
& = \|S(t)(v) - v_0\|^2 - \|v - v_0\|^2
\end{aligned}$$

so that

$$\|S(t)(v) - v_0\|^2 \geq \|v - v_0\|^2 + 2(S(t)(v) - v, \zeta').$$

Using the above inequality and inequality (5.2.13) we obtain

$$\begin{aligned}
& \left(\frac{e^{-2\rho t} - 1}{t} (v - v_0) + 2e^{-2\rho t} \frac{S(t)(v) - v}{t}, \zeta' \right) \\
& = \frac{e^{-2\rho t} - 1}{t} (v - v_0, \zeta') + \frac{e^{-2\rho t}}{t} 2(S(t)(v) - v, \zeta') \\
& = \frac{e^{-2\rho t}}{t} (\|v - v_0\|^2 + 2(S(t)(v) - v, \zeta')) - \frac{1}{t} \|v - v_0\|^2 \\
& \leq \frac{e^{-2\rho t}}{t} \|S(t)(v) - v_0\|^2 - \frac{1}{t} \|v - v_0\|^2 \\
& = \frac{1}{t} (e^{-2\rho t} \|S(t)(v) - v_0\|^2 - \|v - v_0\|^2) \\
& \leq 2 \int_0^1 \langle w_0, v_0 - S(t\tau)(v) \rangle_s e^{-2\rho t\tau} d\tau.
\end{aligned}$$

Taking the limit superior on both sides as $t \searrow 0$ in the above inequality and

applying Lemma 5.2.7(d) again gives

$$\begin{aligned}
& \limsup_{t \searrow 0} \left(\frac{e^{-2\rho t} - 1}{t} (v - v_0) + 2e^{-2\rho t} \frac{S(t)(v) - v}{t}, \zeta' \right) \\
&= \limsup_{t \searrow 0} \left(2 \frac{S(t)(v) - v}{t} - 2\rho(v - v_0), \zeta' \right) \\
&= 2 \limsup_{t \searrow 0} \left(\frac{S(t)(v) - v}{t} - \rho(v - v_0), \zeta' \right) \\
&\leq \limsup_{t \searrow 0} 2 \int_0^1 \langle w_0, v_0 - S(t\tau)(v) \rangle_s e^{-2\rho t\tau} \, d\tau \\
&\stackrel{5.2.7(d)}{\leq} 2 \int_0^1 \langle w_0, v_0 - v \rangle_s \, d\tau \\
&= 2 \langle w_0, v_0 - v \rangle_s.
\end{aligned}$$

Finally taking the supremum over all $\zeta' \in \mathcal{J}(v - v_0)$ on the left side of the above inequality we obtain the desired inequality

$$\sup_{\zeta' \in \mathcal{J}(v - v_0)} \limsup_{t \searrow 0} \left(\frac{S(t)(v) - v}{t} + \rho(v_0 - v), \zeta' \right) \leq \langle w_0, v_0 - v \rangle_s.$$

□

We are now ready to prove the following main result of this section concerning the existence of a solution to the Cauchy problem (5.2.8):

Theorem 5.2.11. *Assume that X is a real Banach space. Let $A \subset X \times X$ and $\rho \in \mathbb{R}$ be such that $A + \rho I$ is accretive. Assume that $\mathcal{R}(I + \lambda A) \supset \mathcal{D}(A)$ for sufficiently small $\lambda > 0$, and that A is a closed subset of $X \times X$. If $u_0 \in \mathcal{D}(A)$ and $0 < T \leq \infty$ then condition (i) below implies condition (ii) on a function $u : [0, T) \rightarrow X$:*

(i) $u(t) = \lim_{n \rightarrow \infty} (I + \frac{t}{n}A)^{-n} (u_0)$ for $t \in [0, T)$ is strongly differentiable almost everywhere

(ii) u is a strong solution to the Cauchy problem

$$0 \in \frac{\partial u}{\partial t} + A(u), \quad u(0) = u_0$$

on $[0, T)$.

Proof. Suppose

$$S(t)(u_0) = u(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} (u_0)$$

for $t \in [0, T]$ and that $u(t)$ is strongly differentiable almost everywhere. Now take $z \in \overline{\mathcal{D}(A)}$, and assume that $S(t)(z)$ is strongly differentiable at $t_0 > 0$ so that

$$S(t_0 + h)(z) = S(t_0)(z) + hw + o(h) \text{ as } h \rightarrow 0,$$

for $h > -t_0$ where $w = \left. \frac{d}{dt} S(t)z \right|_{t=t_0}$. The aim is to show that $[S(t_0)(z), -w] \in A$. This will prove that u is a strong solution since we have already shown in the proof of Theorem 5.2.4 that $S(t)z$ is Lipschitz continuous on bounded t -sets for $z \in \overline{\mathcal{D}(A)}$ and $[S(t_0)(z), -w] \in A$ implies that the Cauchy problem is satisfied almost everywhere.

By assumption we observe that $S(t) : \overline{\mathcal{D}(A)} \rightarrow \overline{\mathcal{D}(A)} \subset \mathcal{R}(I + \lambda A)$. Hence if $0 < \lambda < t_0$, there is a point $[v_\lambda, w_\lambda] \in A$ such that

$$v_\lambda + \lambda w_\lambda = S(t_0 - \lambda)(z) = S(t_0)(z) - \lambda w + o(\lambda) \text{ as } \lambda \rightarrow 0. \quad (5.2.14)$$

If we can prove that

$$v_\lambda \rightarrow S(t_0)(z) \text{ and } w_\lambda \rightarrow -w$$

as $\lambda \searrow 0$ then we have proved that $[S(t_0)(z), -w] \in A$ since A is closed. To show the above put $[v_0, w_0] = [v_\lambda, w_\lambda]$ and $v = S(t_0)(z)$ in inequality (5.2.12) of Lemma 5.2.10. Since $\mathcal{J}(v_\lambda - S(t_0)(z))$ is weak* compact in X' , we can replace the supremum on the right hand side of (5.2.12) by a maximum:

$$\begin{aligned} \sup_{\zeta' \in \mathcal{J}(S(t_0)(z) - v_\lambda)} \limsup_{t \searrow 0} & \left(\frac{S(t)S(t_0)(z) - S(t_0)(z)}{t} + \rho(v_\lambda - S(t_0)(z)), \zeta' \right) \\ & \leq \langle w_\lambda, v_\lambda - S(t_0)(z) \rangle_s \\ & = \max\{(w_\lambda, w') : w' \in \mathcal{J}(v_\lambda - S(t_0)(z))\}. \end{aligned}$$

This yields the existence of an $\eta' \in \mathcal{J}(v_\lambda - S(t_0)(z))$ such that

$$(w + \rho(v_\lambda - S(t_0)(z)), \zeta') \leq (w_\lambda, \eta') \quad (5.2.15)$$

for all $\zeta' \in \mathcal{J}(S(t_0)(z) - v_\lambda)$. Letting $\zeta' = -\eta' \in \mathcal{J}(S(t_0)(z) - v_\lambda)$ in the above and using (5.2.14) and (5.2.15) we obtain

$$\begin{aligned} (S(t_0)(z) - v_\lambda + \lambda \rho(v_\lambda - S(t_0)(z)) + o(\lambda), \eta') \\ & \stackrel{(5.2.14)}{=} (\lambda w_\lambda + \lambda w + \lambda \rho(v_\lambda - S(t_0)(z)), \eta') \\ & = \lambda(w_\lambda, \eta') - \lambda(w + \rho(v_\lambda - S(t_0)(z)), \zeta') \\ & \stackrel{(5.2.15)}{\geq} \lambda(w_\lambda, \eta') - \lambda(w_\lambda, \eta') = 0 \end{aligned}$$

as $\lambda \searrow 0$. Since $\eta' \in \mathcal{J}(v_\lambda - S(t_0)(z))$ we get

$$\begin{aligned} 0 &\leq (S(t_0)(z) - v_\lambda + \lambda\rho(v_\lambda - S(t_0)(z)) + o(\lambda), \eta') \\ &= -(v_\lambda - S(t_0)(z), \eta') + \lambda\rho(v_\lambda - S(t_0)(z), \eta') + (o(\lambda), \eta') \\ &= -\|v_\lambda - S(t_0)(z)\|^2 + \lambda\rho\|v_\lambda - S(t_0)(z)\|^2 + (o(\lambda), \eta') \\ &\leq (\lambda\rho - 1)\|v_\lambda - S(t_0)(z)\|^2 + o(\lambda)\|\eta'\| \\ &= (\lambda\rho - 1)\|v_\lambda - S(t_0)(z)\|^2 + o(\lambda)\|v_\lambda - S(t_0)(z)\|. \end{aligned}$$

which implies

$$(1 - \lambda\rho)\|v_\lambda - S(t_0)(z)\|^2 \leq o(\lambda)\|v_\lambda - S(t_0)(z)\|$$

as $\lambda \searrow 0$. Letting $\lambda \searrow 0$ we see that

$$\lim_{\lambda \searrow 0} \|S(t_0)(z) - v_\lambda\| = 0 \quad \text{and} \quad \lim_{\lambda \searrow 0} \frac{S(t_0)(z) - v_\lambda}{\lambda} = 0.$$

Returning to (5.2.14) again we see that

$$\lim_{\lambda \searrow 0} \|w_\lambda + w\| = \lim_{\lambda \searrow 0} \left\| \frac{S(t_0)(z) - v_\lambda + o(\lambda)}{\lambda} \right\| = 0.$$

Hence we have shown the desired,

$$v_\lambda \rightarrow S(t_0)(z) \quad \text{and} \quad w_\lambda \rightarrow -w$$

as $\lambda \searrow 0$. □

In order to use Theorem 5.2.11 we need to know that the function

$$u(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} (u_0), \quad t \in [0, T],$$

is strongly differentiable almost everywhere. Here we note that if X is reflexive then each Lipschitz continuous X -valued function is strongly differentiable almost everywhere, see [2, p. 282]². Since for $u_0 \in \overline{\mathcal{D}(A)}$ the function $S(t)(u_0) = u(t)$ is Lipschitz continuous on bounded t -sets as shown in the proof of Theorem 5.2.4, then we see that if we assume that X is reflexive, then u as defined above is a strong solution to the Cauchy problem (5.2.8).

²Note that the reference does not give a proof. It is, though, the only reference found yet that states this.

5.3 Examples

We look at two examples of diffusion problems for which we can use the above results to conclude that there exists a solution. The first example is the diffusion problem on $X = L^2(\Omega)$ given by

$$u_t + u = 0, \quad u(0) = u_\delta. \quad (5.3.1)$$

This problem stems from a denoising problem with standard Tikhonov regularization function $g(x, u, Du) = \frac{1}{2}|u|^2$ and initial noisy data u_δ . Using the notation of this chapter we then have $A = I : L^2(\Omega) \rightarrow L^2(\Omega)$. Clearly I is accretive since for $\lambda > 0$ the operator

$$(I + \lambda I)^{-1} = ((1 + \lambda)I)^{-1} = \frac{1}{1 + \lambda}I$$

is single-valued and

$$\|(I + \lambda I)^{-1}(v - w)\| = \frac{1}{1 + \lambda}\|v - w\| \leq \|v - w\|.$$

The identity operator is also closed and satisfies the range condition

$$L^2(\Omega) = \overline{\mathcal{D}(I)} \subset \mathcal{R}(I + \lambda I) = L^2(\Omega)$$

for all $\lambda > 0$. Hence $A = I : L^2(\Omega) \rightarrow L^2(\Omega)$ satisfies all the conditions of Theorem 5.2.11 and since $L^2(\Omega)$ is reflexive, the diffusion problem in (5.3.1) has a strong solution given by

$$u(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n}I \right)^{-n} u_\delta.$$

The second example is the Cauchy problem

$$u_t - \Delta u = 0, \quad u(0) = u_\delta \quad (5.3.2)$$

on $L^2(\Omega)$. The problem stems from a denoising problem with standard Tikhonov regularization function $g(x, u, Du) = \frac{1}{2}|Du|^2$ and initial noisy data u_δ . We have already shown as an introductory example that the operator $A : W^{1,2}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ defined by $A = -\Delta$ is accretive. If we restrict A to a closed subset C of $W^{1,2}(\mathbb{R}^n)$ then A is closed by [22, Lemma 4.13-5] and

$$C \subset \mathcal{R}(I - \lambda \Delta).$$

for sufficiently small $\lambda > 0$. Hence $A = -\Delta$ on $L^2(\mathbb{R}^n)$ satisfies all the conditions of Theorem 5.2.11 and since $L^2(\mathbb{R}^n)$ is a Hilbert space it is reflexive and the diffusion problem in (5.3.2) admits a strong solution given by

$$u(t) = \lim_{n \rightarrow \infty} \left(I - \frac{t}{n}\Delta \right)^{-n} u_\delta.$$

Note that we can restrict the spacial domain to an open, bounded, connected and Lipschitz domain $\rho \subset \mathbb{R}^n$ by setting $u_\delta = 0$ outside the domain.

In Chapter 6 we will return to the general Cauchy problem in which A maps X into X' and the time derivative u_t is interpreted as an element of X' .

6

Non-linear Diffusion Methods in Dual Space

¹In this chapter we return to the general Cauchy problem

$$\mathcal{J} \left(\frac{\partial u}{\partial t} \right) + A(u) \ni 0, \quad u(0) = u_0, \quad (6.0.1)$$

in which $A : X \rightarrow X'$ is a possibly non-linear operator, X is a Banach space and \mathcal{J} is the duality mapping according to the weight function $\phi(t) = t$ as described in Chapter 4. We set up a conjecture about the existence of a solution to this kind of Cauchy problem based on the results in Chapter 5. Then we show that all A operators stemming from a minimization problem as described in Chapter 4 satisfies the conditions of the conjecture and in particular the conditions of Theorem 5.2.11 whenever X is a Hilbert space and we can identify X' with X .

All the operators $A : X \rightarrow X'$ of interest are actually the subdifferential of lower semi-continuous and convex functionals, as seen in Chapter 4 (we will return to this statement). The first section below is therefore concerned with properties of such kind of subdifferentials. We will see that the properties somehow relates to the conditions of Theorem 5.2.11 about the existence of a solution to the Cauchy

¹This chapter is based on [19, Chap. 2] and [18, Chap. 7].

problem when A maps X into X . This leads to the formulation of the conjecture regarding the existence of a solution to the Cauchy problem (6.0.1).

6.1 Properties of the Subdifferential

In this section we prove some general properties of the subdifferential of a lower semi-continuous, convex, and proper functional on a Banach space X . The main result is that the subdifferential of such a functional is what we call a *maximal monotone operator*.

The first important thing we observe for the subdifferential $\partial\phi$ of a function $\phi : X \rightarrow Y$, is that it is what we call *monotone*. That is, for $v_1, v_2 \in \mathcal{D}(\partial\phi)$ and $z_1 \in \partial\phi(v_1)$, $z_2 \in \partial\phi(v_2)$ we have

$$\begin{aligned} (v_1 - v_2, z_1 - z_2) &= (v_1 - v_2, z_1) - (v_1 - v_2, z_2) \\ &= (v_1 - v_2, z_1) + (v_2 - v_1, z_2) \\ &\geq \phi(v_1) - \phi(v_2) + \phi(v_2) - \phi(v_1) \\ &= 0. \end{aligned} \tag{6.1.1}$$

The set of all monotone operators can be ordered by graph inclusion. That is, for two monotone operators A_1 and A_2 on X we say that $A_1 \leq A_2$ if $A_1(v) \subset A_2(v)$ for all $v \in X$. By Zorn's lemma (see [22, 4.1-6]) we then obtain at least one maximal element, which we will call a *maximal monotone operator* on X . In practice, to show that an operator is maximal monotone it is easier to use the following characterization of maximal monotone operators:

Theorem 6.1.1 (Maximal Monotone Operator). *Let X be a reflexive Banach space and let X be strictly convex. Then a monotone operator $A : X \rightarrow X'$ is maximal monotone if and only if for any $\lambda > 0$ (equivalently, for some $\lambda > 0$) we have*

$$\mathcal{R}(A + \lambda\mathcal{J}) = X',$$

where \mathcal{J} is the duality mapping of X .

Proof. See [19, theorem 2.2]. □

Sometimes we also consider an operator $A : X \rightarrow X'$ as a subset of $X \times X'$ and if A is a maximal monotone operator then we call $A \subset X \times X'$ a *maximal monotone set*. Since we have shown that the subdifferential $\partial\phi : X \rightarrow X'$ of a functional $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is monotone, we would like to show that in special cases it is actually maximal monotone:

Theorem 6.1.2. *Let X be a reflexive real Banach space and let $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous, convex and proper functional. Then $\partial\phi : X \rightarrow X'$ is a maximal monotone operator.*

Proof. We have already shown that $\partial\phi$ is monotone in (6.1.1). In order to prove that $\partial\phi$ is maximal monotone, we fix $z \in X'$ and consider the equation

$$\mathcal{J}(v) + \partial\phi(v) \ni z.$$

Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be defined by

$$f(v) = \frac{1}{2}\|v\|^2 + \phi(v) - (v, z).$$

Since $\|\cdot\|^2$, $\phi(\cdot)$ and (\cdot, z) are all lower semi-continuous and convex functions then also f is lower semi-continuous and convex. By Proposition A.2.5 in Appendix A.2 used for ϕ we get

$$\begin{aligned} f(v) &\geq \frac{1}{2}\|v\|^2 + (v, v') + \beta - (v, z) = \frac{1}{2}\|v\|^2 + \beta - (v, z - v') \\ &\geq \frac{1}{2}\|v\|^2 + \beta - \|v\|\|z - v'\|, \end{aligned}$$

where $\beta \in \mathbb{R}$ and $v' \in X'$. Hence if we let $\|v\| \rightarrow \infty$ we obtain

$$\lim_{\|v\| \rightarrow \infty} f(v) = +\infty.$$

This shows that f is coercive and by Theorem A.2.12 in Appendix A.2 there exists a minimizer $v_0 \in X$ of f , that is,

$$f(v_0) = \inf\{f(v) : v \in X\}.$$

Since v_0 is a minimizer of f then we have the following inequality:

$$\frac{1}{2}\|v_0\|^2 + \phi(v_0) - (v_0, z) = f(v_0) \leq f(v) = \frac{1}{2}\|v\|^2 + \phi(v) - (v, z), \quad \forall v \in X.$$

This implies

$$\phi(v_0) - \phi(v) \leq (v_0 - v, z) + \frac{1}{2}(\|v\|^2 - \|v_0\|^2). \quad (6.1.2)$$

For $\omega \in \mathcal{J}(v)$ we can estimate the last term by

$$\begin{aligned} (v - v_0, \omega) &= (v, \omega) - (v_0, \omega) \geq \|v\|^2 - \|v_0\|\|\omega\| \\ &= \|v\|^2 - \|v_0\|\|v\| \geq \|v\|^2 - \frac{1}{2}(\|v_0\|^2 + \|v\|^2) \\ &= \frac{1}{2}(\|v\|^2 - \|v_0\|^2). \end{aligned}$$

Using this in (6.1.2) we obtain

$$\phi(v_0) - \phi(v) \leq (v_0 - v, z) + (v - v_0, \omega). \quad (6.1.3)$$

Set $v = tv_0 + (1 - t)u$ for $0 < t < 1$ and $u \in X$. Then by the convexity of ϕ we obtain

$$\begin{aligned} \phi(v_0) - \phi(tv_0 + (1 - t)u) &\geq \phi(v_0) - t\phi(v_0) - (1 - t)\phi(u) \\ &= (1 - t)(\phi(v_0) - \phi(u)). \end{aligned}$$

By inequality (6.1.3) we then obtain

$$\begin{aligned} \phi(v_0) - \phi(u) &\leq \frac{1}{1 - t} ((v_0 - (tv_0 + (1 - t)u), z) + (tv_0 + (1 - t)u - v_0, \omega_t)) \\ &= \frac{1}{1 - t} ((1 - t)(v_0 - u, z) + (1 - t)(u - v_0, \omega_t)) \\ &= (v_0 - u, z) + (u - v_0, \omega_t), \end{aligned}$$

where $\omega_t \in \mathcal{J}(tv_0 + (1 - t)u)$. Letting $t \rightarrow 1$ then $tv_0 + (1 - t)u \rightarrow v_0$ so that

$$\|\omega_t\| = \|tv_0 + (1 - t)u\| \rightarrow \|v_0\|.$$

This shows that the sequence of norms $(\|\omega_t\|)$ is bounded. This implies that ω_t weakly* converges in X' . Since $\mathcal{J}(v_0)$ is weak* compact then $\omega_t \rightarrow \omega_0 \in \mathcal{J}(v_0)$ as $t \rightarrow 1$. Hence we obtain

$$\begin{aligned} \phi(v_0) - \phi(u) &\leq (v_0 - u, z) + (u - v_0, \omega_0) \\ &= (v_0 - u, z - \omega_0), \quad \forall u \in X. \end{aligned}$$

This inequality shows that $z - \omega_0 \in \partial\phi(v_0)$, that is, v_0 is solution to

$$\mathcal{J}(v) + \partial\phi(v) \ni z.$$

Since $z \in X'$ was arbitrary we have thus shown that for each such z we can find a solution to the above equation. This implies that

$$\mathcal{R}(\mathcal{J} + \partial\phi) = X'.$$

Since X is a reflexive Banach space then by [13, Corollary 1(i)], X has an equivalent strictly convex norm. This implies that we can use Theorem 6.1.1 with $\lambda = 1$ to conclude that $\partial\phi$ is maximal monotone. \square

We have not yet discussed the domain on which $\partial\phi$ is defined. It turns out that in special cases $\partial\phi$ is actually densely defined on $\mathcal{D}(\phi)$:

Proposition 6.1.3. *Let $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous, convex, and proper functional. Then $\mathcal{D}(\partial\phi)$ is a dense subset of $\mathcal{D}(\phi)$.*

Proof. Let $v \in \mathcal{D}(\phi)$ and let v_λ be the solution to the equation

$$\mathcal{J}(v_\lambda - v) + \lambda \partial\phi(v_\lambda) \ni 0.$$

Pairing the equation with $v_\lambda - v$ (and by misuse of notation) we obtain

$$\begin{aligned} 0 &\in (v_\lambda - v, \mathcal{J}(v_\lambda - v)) + \lambda (v_\lambda - v, \partial\phi(v_\lambda)) \\ &= \|v_\lambda - v\|^2 + \lambda (v_\lambda - v, \partial\phi(v_\lambda)) \\ &\geq \|v_\lambda - v\|^2 + \lambda (\phi(v_\lambda) - \phi(v)), \quad \forall \lambda > 0. \end{aligned}$$

This implies

$$\|v_\lambda - v\|^2 + \lambda (\phi(v_\lambda) - \phi(v)) \leq 0, \quad \forall \lambda > 0.$$

By Proposition A.2.5 in Appendix A.2 we then obtain

$$\|v_\lambda - v\|^2 + \lambda ((v_\lambda, v') + \beta - \phi(v)) \leq 0, \quad \forall \lambda > 0,$$

where $v' \in X'$ and $\beta \in \mathbb{R}$. Since $v \in \mathcal{D}(\phi)$ and hence $\phi(v) < \infty$ then the above shows that

$$\lim_{\lambda \rightarrow 0} v_\lambda = v.$$

As $v \in \mathcal{D}(\phi)$ was arbitrarily chosen and $v_\lambda \in \mathcal{D}(\partial\phi)$ we have thus shown

$$\overline{\mathcal{D}(\partial\phi)} = \overline{\mathcal{D}(\phi)}.$$

□

Finally we show that in the case of a lower semi-continuous functional ϕ the subdifferential $\partial\phi$ is a closed subset of $X \times X'$:

Proposition 6.1.4. *Let X be a reflexive Banach space and let $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous, convex, and proper functional. Then $\partial\phi$ is a weak-weak* closed subset of $X \times X'$.*

Proof. We have to show that if $[v_n, w_n] \in \partial\phi$ are such that $v_n \rightharpoonup v$ in X and $w_n \xrightarrow{*} w$ in X' as $n \rightarrow \infty$ then $[v, w] \in \partial\phi$. By Theorem 6.1.2, $\partial\phi$ is maximal monotone. By the monotonicity we have

$$(v_n - y, w_n - z) \geq 0, \quad \forall [y, z] \in \partial\phi.$$

Using $v_n \rightharpoonup v$ and $w_n \xrightarrow{*} w$ we then obtain

$$(v - y, w - z) \geq 0, \quad \forall [y, z] \in \partial\phi.$$

Since $\partial\phi$ is maximal monotone then we must have $[v, w] \in \partial\phi$. □

6.2 Cauchy Problem in Dual Space

In the above we have shown that the subdifferential $\partial\phi$ of a lower semi-continuous, convex and proper functional $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ defined on a reflexive Banach space X is maximal monotone, densely defined on $\mathcal{D}(\phi)$ and closed. If we let X be a Hilbert space then we can identify X' with X . Using Lemma 5.2.7(e) and the monotonicity of $\partial\phi$, we conclude that $\partial\phi$ is accretive. Since $\partial\phi$ is maximal monotone then

$$\begin{aligned} \mathcal{D}(\partial\phi) \subset X \simeq X' &= \mathcal{R}(\lambda I + \partial\phi) = \mathcal{R}(\lambda(I + \lambda^{-1}\partial\phi)) \\ &= \lambda\mathcal{R}(I + \lambda^{-1}\partial\phi) \subset \mathcal{R}(I + \beta\partial\phi) \end{aligned}$$

for sufficiently small $\beta > 0$. Here \simeq emphasizes that we have identified X and X' . Finally, $\partial\phi$ is a weak-weak* closed subset of $X \times X' \simeq X \times X$. The conditions of Theorems 5.2.4 and 5.2.11 are therefore satisfied for $A = \partial\phi$ and since X is reflexive then

$$u(t) = \lim_{n \rightarrow \infty} \left(\mathcal{J} + \frac{t}{n} \partial\phi \right)^{-n} (u_0), \quad t \in [0, T)$$

is a strong solution to the Cauchy problem

$$0 \in \frac{\partial u}{\partial t} + \partial\phi(u), \quad u(0) = u_0 \tag{6.2.1}$$

for $u_0 \in \mathcal{D}(\partial\phi)$ and $0 < T \leq \infty$.

In the case X is not a Hilbert space we would like to say something about the existence of a solution to the Cauchy problem

$$\mathcal{J} \left(\frac{\partial u}{\partial t} \right) + A(u) \ni 0, \quad u(0) = u_0,$$

where again \mathcal{J} is the duality mapping according to $\phi(t) = t$. Here we give the conjecture:

Conjecture 6.2.1. *Assume that X is a real and reflexive Banach space. Let $A \subset X \times X'$ and $\rho \in \mathbb{R}$ be such that $A + \rho I$ is maximal monotone. Assume that $\mathcal{R}(\mathcal{J} + \lambda A) \supset \overline{\mathcal{R}(A)}$ for sufficiently small $\lambda > 0$, and that A is a closed subset of $X \times X'$. If $u_0 \in \mathcal{D}(A)$ and $0 < T \leq \infty$ then $u : [0, T) \rightarrow X$ defined by*

$$u(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} (u_0')$$

for $t \in [0, T)$ and $u_0' \in J(u_0)$ is a strong solution to the Cauchy problem

$$0 \in \mathcal{J} \left(\frac{\partial u}{\partial t} \right) + A(u), \quad u(0) = u_0$$

on $[0, T)$.

We have shown that the subdifferential $\partial\phi$ of a lower semi-continuous, convex, and proper functional ϕ is maximal monotone, weak-weak* closed and satisfies

$$\mathcal{R}(\partial\phi) \subset X' = \mathcal{R}(\partial\phi + \lambda\mathcal{J}) = \lambda\mathcal{R}(\mathcal{J} + \lambda^{-1}\partial\phi) \subset \mathcal{R}(\mathcal{J} + \beta\partial\phi)$$

for sufficiently small $\beta > 0$. The subdifferential therefore satisfies the conditions of Conjecture 6.2.1 and according to this there should exist a solution to the Cauchy problem in the dual space X' .

Returning to section 4 we see that the diffusion problems of interest deals with two kinds of operators. Either

$$A(u) = \partial_u \left(\int_{\Omega} u - u_{\delta} \, dx \right)$$

or

$$A(u) = \partial_u \left(\int_{\Omega} g(x, u, Du) \, dx \right)$$

where $g : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ is a normal and convex integrand. In either case A is the subdifferential of a lower semi-continuous, convex and proper functional if $X = W^{1,p}(\Omega)$ for $1 < p < \infty$ or $X = BV(\Omega)$ (see Theorem 3.1.5) and therefore satisfies the conditions of Conjecture 6.2.1. Note that since $BV(\Omega)$ is not reflexive then we it should be proven that every Lipschitz continuous $BV(\Omega)$ -valued function is strongly differentiable almost everywhere. The relations between the conditions of Corollaries 3.1.7 and 3.1.7 and Theorem 5.2.11 and Conjecture 6.2.1 has become clear through this chapter by relating the subdifferential of a lower semi-continuous, convex and proper functional to maximal monotone and closed

operators.

In Chapter 7 numerical experiments are performed in order to investigate the convergence and the effect of the diffusion filtering method for specific examples.

7

Numerical Experiments

In this chapter numerical experiments for the diffusion methods, discussed in Chapter 4, are set up and solved using finite difference methods in `MatLab`. Some of the problems have not been proved to admit a solution as discussed in Chapter 6. Here we try to solve them anyway using a finite difference method. We will see that there is a difference in how ill-posed the problems are comparing denoising problems, deblurring problems and computed tomography problems. It will also be seen that the time step parameter for the finite difference methods has a huge impact on whether we obtain a reasonable solution or not.

7.1 Denoising with Standard Tikhonov Regularization

Let $\Omega = (0, 1) \subset \mathbb{R}$. We want to reconstruct the function given by

$$u_{\text{exact}}(x) = \begin{cases} 1, & 0.1 \leq x \leq 0.4 \text{ or } 0.6 \leq x \leq 0.9, \\ 0, & \text{else,} \end{cases}$$

from noisy given data. Dividing Ω into 100 equidistant points, u_{exact} can be represented by the function values at these points. To all the function values 7% Gaussian distributed additive noise is added in order to obtain simulated noisy data. In Figure 7.1 the discretized clean data $u_{\text{exact}}(x)$ is seen in blue to the left and the noisy data $u_{\delta}(x)$ in red to the right.

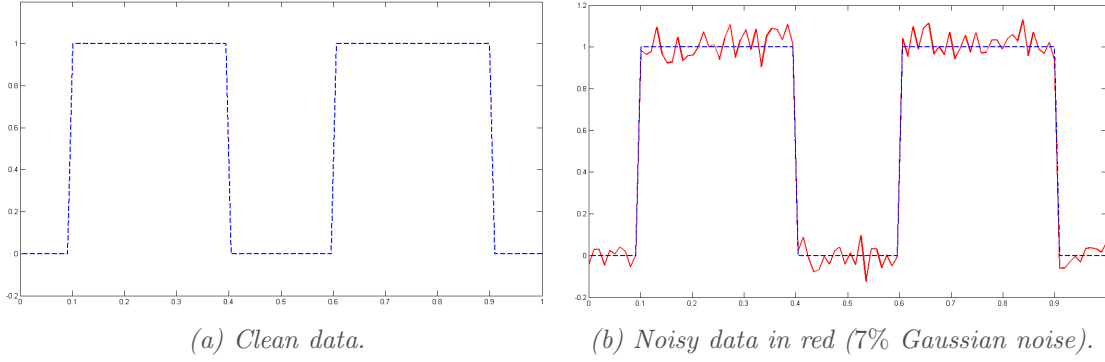


Fig. 7.1: Denoising 1D: Data for reconstruction.

Now we want to use a standard Tikhonov regularization with regularization function $g(x, u, Du) = \frac{1}{2}|Du|^2$ for this problem. The corresponding diffusion problem is

$$u_t = u_{xx}, \quad u(0) = u_\delta.$$

As described in section 4 a homogeneous Neumann boundary condition is applied. We use a finite difference method with time step Δt and discretize the problem using a spacial step length $\Delta x = \frac{1}{100}$. Letting u_j^n denote the value of u at time step $t = n\Delta t$ and point $x = j\Delta x$ and using a forward difference approximation of u_t and a second order central difference approximation of u_{xx} we then obtain the finite difference scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2}, \quad j = 2, \dots, 99$$

or simplified as

$$u_j^{n+1} = (1 - 2r)u_j^n + r(u_{j+1}^n + u_{j-1}^n), \quad j = 2, \dots, 99$$

where $r = \Delta t/\Delta x^2$. Using the homogeneous Neumann boundary condition we obtain

$$u_0^n = u_1^n \quad \text{and} \quad u_{101}^n = u_{100}^n$$

and hence at the endpoints we have

$$u_1^{n+1} = (1 - r)u_1^n + ru_2^n \quad \text{and} \quad u_{100}^{n+1} = (1 - r)u_{100}^n + ru_{99}^n.$$

Implementing the iteration scheme in `MatLab`, see Appendix B, and choosing $\Delta t = 10^{-5}$ by trial and error we obtain the reconstructions at different time steps seen in Figure 7.2. Here the dashed blue line is the original exact data. The green curve is the reconstruction at time $t = \Delta t$, the black at time $t = 10\Delta t$ and the magenta at time $t = 100\Delta t$. We see that the reconstructions become smoother and

smoother as time evolves. Hence if we continue the process for too long we will just get approximately a constant reconstruction which is equal to the average value of u_{exact} on Ω . This suggests that we should have a stopping criteria. Furthermore Δt is just chosen by trial and error which is not optimal. Finally we could actually change the time step length Δt in each iteration in order to optimize the algorithm. We can imagine that smaller time steps are needed at the beginning in order to obtain an approximate reconstruction and then we can use greater and greater time step lengths.

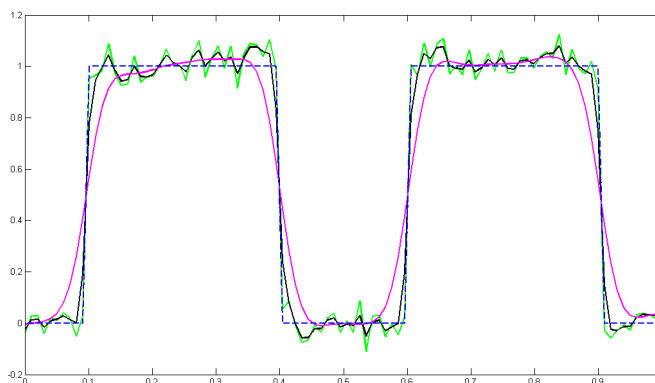


Fig. 7.2: Denoising 1D: Standard Tikhonov reconstructions at different times. Blue (clean data), green (iteration 1), black (iteration 10), magenta (iteration 100).

Now we move on to a two-dimensional case. In this case we want to reconstruct an image from a degraded one. The degradation is due to Gaussian additive noise. In Figure 7.3 the clean image and a noisy image with 2% additive Gaussian noise are seen.

The idea is now to use the diffusion problem

$$u_t = \Delta u, \quad u(0) = u_\delta$$

with homogeneous Neumann boundary conditions to remove the noise and recover the clean image. As in the one dimensional case, we use a finite difference method with time step Δt and discretize the problem using the spacial step lengths Δx in the x -direction and Δy in the y -direction. Again letting $u_{j,k}^n$ denote the value of u at time step $t = n\Delta t$ and point $(x, y) = (j\Delta x, k\Delta y)$ we use a forward difference approximation of u_t and a second order central difference approximation of u_{xx} and u_{yy} . Like in the one dimensional case, we fix the boundary values using the homogeneous Neumann boundary condition. Since the image we use is a 512×512 pixel image, then $\Delta x = \Delta y = \frac{1}{512}$. Implementing the iteration scheme in `MatLab`, see Appendix B, and choosing $\Delta t = 10^{-7}$ by trial and error we obtain the reconstructions at different time steps seen in Figure 7.4. Iteration n corresponds



(a) Clean image.

(b) Noisy image (2% Gaussian noise)

Fig. 7.3: Denoising 2D: Data for reconstruction.

to time $t = n\Delta t$. We see that at early time steps there is still noise in the image. As time evolves the noise is removed by smoothening the data. At iteration 100 we see that the picture is actually smeared because of too much smoothening. Again there should be a stopping rule for when to stop the iteration.

Trying to increase the time step length to $\Delta t = 5 \cdot 10^{-7}$ we obtain the reconstructions in Figure 7.5. Due to the larger time step the smoothening went much faster for this iteration process. Choosing the time step parameter too large we may obtain no good reconstruction because of a too fast smoothening process.

Denoising with a standard Tikhonov regularization is not the best way to obtain good results for functions or images with sharp edges. In the next section we use a total variation regularization and it is seen that we obtain much better results for the two examples above.

7.2 Denoising with Total Variation Regularization

We use the same clean data and noisy data as in Figure 7.1. Instead of using a standard Tikhonov regularization we use a total variation regularization term given by $|Du|(\Omega)$. The corresponding diffusion problem is

$$u_t = \nabla \cdot \left(\frac{1}{|u_x|} u_x \right)$$

with a homogeneous Neumann boundary condition. We have actually not established an existence result for a solution to this problem but here we try to solve it using a finite difference method anyway. And we see that we obtain good results.



Fig. 7.4: Denoising 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 10^{-7}$.

For the finite difference method (in time) we use the time step Δt and discretize the problem using the spacial step length $\Delta x = \frac{1}{100}$. As above we let u_j^n denote the value of u at time step $t = n\Delta t$ and point $x = j\Delta x$. This time we use first a central difference approximation for u_x :

$$\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}.$$

Then using a forward difference approximation of u_t and a central difference approximation for $\nabla \cdot (\cdot)$ we arrive at the finite difference scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\text{sgn}(u_{j+2}^n - u_j^n) - \text{sgn}(u_j^n - u_{j-2}^n)}{2\Delta x}$$

or simplified as

$$u_j^{n+1} = u_j^n + r(\text{sgn}(u_{j+2}^n - u_j^n) - \text{sgn}(u_j^n - u_{j-2}^n))$$

where $r = \Delta t / (2\Delta x)$. This scheme suffers from an odd-even decoupling. We remove this by decreasing the width of the differences:

$$u_j^{n+1} = u_j^n + r(\text{sgn}(u_{j+1}^n - u_j^n) - \text{sgn}(u_j^n - u_{j-1}^n))$$



Fig. 7.5: Denoising 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 5 \cdot 10^{-7}$.

where $r = \Delta t / \Delta x$. Using the homogeneous Neumann boundary conditions we again obtain

$$u_0^n = u_1^n \quad \text{and} \quad u_{101}^n = u_{100}^n$$

and hence at the endpoints we have

$$u_1^{n+1} = u_1^n + r \cdot \text{sgn}(u_2^n - u_1^n) \quad \text{and} \quad u_{100}^{n+1} = u_{100}^n - r \cdot \text{sgn}(u_{100}^n - u_{99}^n).$$

Implementing this iteration scheme in MatLab, see Appendix B, and choosing $\Delta t = 10^{-5}$ by trial and error we obtain the reconstructions seen in Figure 7.6. Again the dashed blue line is the original exact data, the green curve is the reconstruction at iteration 1, the black at iteration 10, and the magenta at iteration 100. We see that the recovering of u_{exact} is much better than for the standard Tikhonov regularization. This is because the total variation regularization tends to preserve sharp edges. We also see the drawback from the total variation method; the staircase effect. Because TV regularization insists on preserving the sharp edges, it sometimes makes new edges based on the noise in the data. This is seen in Figure 7.6 in the reconstructions for example between 0.3 and 0.4 at the x -axis.

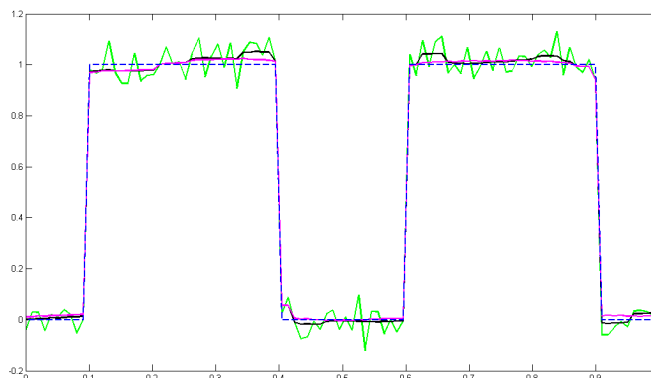


Fig. 7.6: Denoising 1D: Total variation reconstructions at different times. Blue (clean data), green (iteration 1), black (iteration 10), magenta (iteration 100).

In order to investigate the method in further detail we move on to the two dimensional example seen in Figure 7.3. As for the one dimensional case we use a finite difference scheme implemented in `MatLab`, see Appendix B, in order to obtain the results in Figure 7.7. The time step size is set to $\Delta t = 1/2\Delta x$ by trial and error. We see that at early stages there is still noise in the picture. During the iteration process the edges are kept sharp and a smoothing between edges occurs. Turning our attention to iteration 150 we see that we obtain an almost cartoon like figure. Hence the iteration process should have been stopped before this stage.

Increasing the time step size parameter to $\Delta t = \Delta x$ we see that the smoothing in isotropic areas becomes too dominating (see Figure 7.8).

7.3 Deblurring with Standard Tikhonov Regularization

In this section we add a blur to the problems of section 7.1. In one dimension we use again the clean data from Figure 7.1. This time we blur the data using the convolution operator

$$\mathcal{K}u(x) = \int_{\Omega} k(x-y)u(y) \, dy, \quad k(x-y) = \frac{1}{\sqrt{2\pi}\beta} \exp\left(-\frac{(x-y)^2}{2\beta^2}\right)$$

where $\beta \in \mathbb{R}$. In practice we multiply the data vector

$$U = (u_{\text{exact}}(0), u_{\text{exact}}(\Delta x), \dots, u_{\text{exact}}(100\Delta x))'$$

by the matrix A stemming from a convolution operator and given by

$$A_{ij} = \frac{\Delta x}{\sqrt{2\pi}\beta} \exp\left(-\frac{(i\Delta x - j\Delta x)^2}{2\beta^2}\right)$$



Fig. 7.7: Denoising 2D: Total variation reconstructions at different times with $\Delta t = 1/2\Delta x$.

where $\beta = 0.05$. To the blurred version of u_{exact} we then add 2% Gaussian additive noise in order to obtain simulated noisy data. The clean data is the blurred version of u_{exact} and the noisy data is the blurred version with 2% Gaussian additive noise. In Figure 7.9 the cyan clean data is seen to the left and the noisy red data is seen to the right.

Let \mathcal{K} denote the convolution operator that gives rise to the matrix A . Then the diffusion problem associated with the standard Tikhonov regularization for this problem reads as

$$u_t = -\mathcal{K}'(u)^\#(\mathcal{K}u - u_\delta), \quad u(0) = 0$$

with homogeneous Neumann boundary conditions (see Chapter 4). Here u_t should be regarded as an element in the dual. We have not yet proved the existence of a solution for this problem but here we solve it using a finite difference method (in time) anyway. Again u_δ is the noisy given data. Discretizing the problem as for the denoising problem we let

$$U^n = (u_1^n, u_2^n, \dots, u_{100}^n).$$



Fig. 7.8: Denoising 2D: Total variation reconstructions at different times with $\Delta t = \Delta x$.

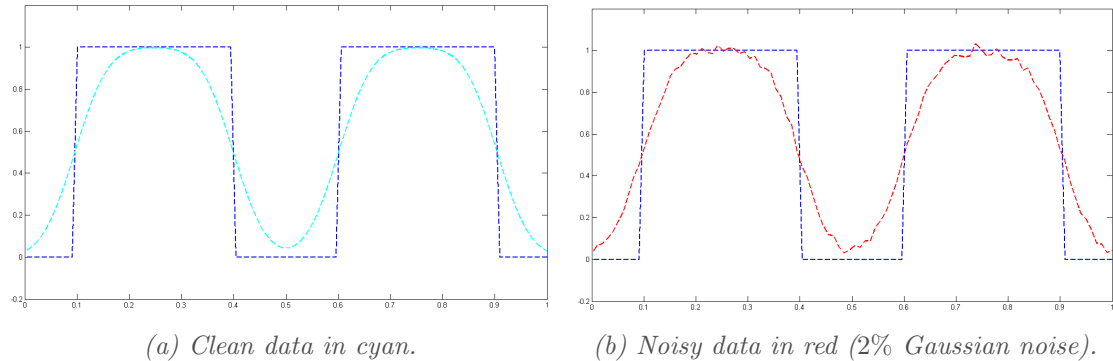


Fig. 7.9: Deblurring 1D: Data for reconstruction.

Observing that the discrete version of $\mathcal{F}'(u)^\#$ is simply A^T and using a forward finite difference approximation of u_t we obtain the finite difference scheme

$$U^{n+1} = U^n - \Delta t A^T (AU^n - u_\delta), \quad U^1 = (0, 0, \dots, 0). \quad (7.3.1)$$

Using the Neumann boundary conditions we set

$$u_1^{n+1} = u_2^{n+1} \quad \text{and} \quad u_{100}^{n+1} = u_{99}^{n+1}.$$

Implementing the scheme in `MatLab`, see Appendix B, and choosing $\Delta t = 6 \cdot 10^{-6}$ by trial and error we obtain the reconstructions seen in Figure 7.10. Here the blue dashed line is the actual function we want to reconstruct. Then the green is the reconstruction after iteration 50000, the black after iteration 200000, and the magenta after iteration 1100000. We see that we need quite a lot of iterations in order to obtain a reconstruction that looks like the actual function u_{exact} . There are several reasons for this. For the reconstruction method our initial guess is 0. This is far from the true solution. Hence we need a lot of iterations in order to obtain just a rather rough reconstruction. Furthermore the problem is extremely ill-posed which makes it even more difficult to obtain a good reconstruction. And finally the standard Tikhonov regularization method is not the most optimal method for this kind of problem since it is a smoothing reconstruction method and we are trying to reconstruct sharp edges. So far though this is the only regularization method for which we can formulate an easy solvable diffusion problem.

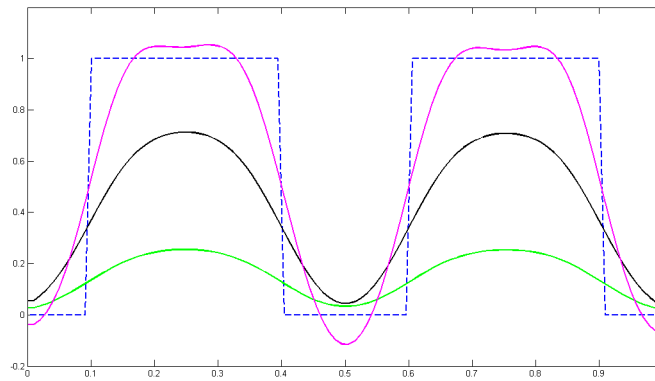


Fig. 7.10: Deblurring 1D: Standard Tikhonov reconstructions at different times with $\Delta t = 6 \cdot 10^{-6}$. Blue (actual function), green (iteration 50000), black (iteration 200000), magenta (iteration 1100000).

Increasing the time step size to $\Delta t = 5 \cdot 10^{-4}$ we obtain the reconstructions seen in Figure 7.11. It is seen that we actually obtain as good results with this choice of time step length. The process is just running faster.

For the deblurring problem in one dimension we have now seen that it is difficult to obtain a good reconstruction. This is due to the extreme ill-posedness of the deblurring problem, but it might also be due to a poor choice of time step parameter. Moving on to a two dimensional problem we will see the same difficulties in obtaining a good reconstruction. Once again we use the clean image in Figure 7.3 to the left. Instead of just adding noise we first do a motion blur to the image using the motion blur matrix generated by the `mblur`-function in the `AIR tools` package in `MatLab`, see [34]. To the blurred image 1% Gaussian additive noise is added in order to construct simulated noisy data. The clean data (the blurred

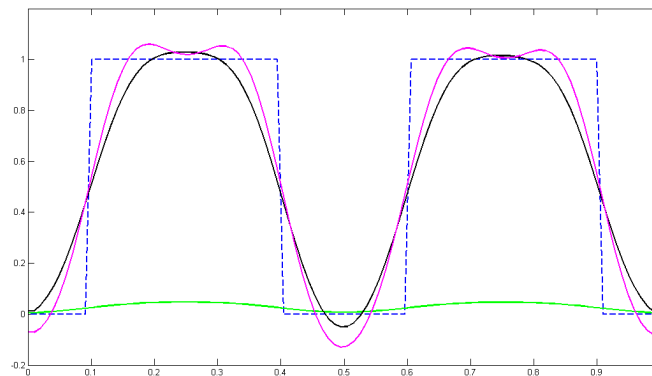


Fig. 7.11: Deblurring 1D: Standard Tikhonov reconstructions at different times with $\Delta t = 5 \cdot 10^{-4}$. Blue (actual function), green (iteration 100), black (iteration 7000), magenta (iteration 20000).

image) and the noisy data (the blurred image with noise) are seen in Figure 7.12.



(a) Clean data.

(b) Noisy data (1% Gaussian noise).

Fig. 7.12: Deblurring 2D: Data for reconstruction.

We use again the finite difference scheme in (7.3.1) where now the matrix A is the motion blur matrix. For the deblurring problem it is very difficult to choose the right time step size. Due to the extreme ill-posedness of the problem we need a large Δt in order to deblur the given blurred data. On the other hand the time step size should be small in order to remove noise. Choosing $\Delta t = 0.2$ by trial and error we obtain the reconstructions seen in Figure 7.20 (the figure is made the last figure of this Chapter since it is a one page figure). Again we see that it is difficult to obtain good reconstructions. At the beginning the noise is removed, but the blur is not removed. At some point the blur is removed and then we

start fitting to noise. Again we need a criteria of how to choose the time step size and when to stop the iteration process. Setting Δt to a large value means that we do not really want to use the regularizing term, since then the regularization parameter $\alpha = 1/\Delta t$ becomes small (see Chapter 4). In order to remove the noise even better it may be necessary to set a smaller value for Δt but because of the extreme ill-posedness of the deblurring problem we would then have to run a huge amount of iterations, which is time consuming. This is not optimal in practice.

Using a too small time step parameter $\Delta t = 0.05$ leads to the results in Figure 7.13. We see that we obtain no good results in this case. The noise is removed, but we never obtain the deblur of the image.



Fig. 7.13: Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 0.05$.

Using a too large time step parameter $\Delta t = 1$ leads to the results in Figure 7.14. We see that we obtain no good results in this case either. Now the blur in the image is removed but we fit to noise and the noise is never removed.



Fig. 7.14: Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 1$.

7.4 Computed Tomography with Standard Tikhonov Regularization

In computed tomography (CT) one wants to reconstruct an image of the body's inner structure. Sending X-rays through the body at different angles one can measure the corresponding damped signals on the opposite side of the body assuming that the X-rays go straight through the body. The damping is due to an attenuation coefficient $\mu(\eta)$ depending on the travelled distance η along the straight line, see Figure 7.15. The given data for CT problems are the projections of the signal intensity along each line for different angles. Due to the difference in attenuation coefficient for different tissues we are then able to reconstruct an image of the body's inner structure by reconstructing the attenuation coefficient.

To obtain an analytical reconstruction formula for CT scanning we start by parametrising each line L along which the intensity is projected. Using the

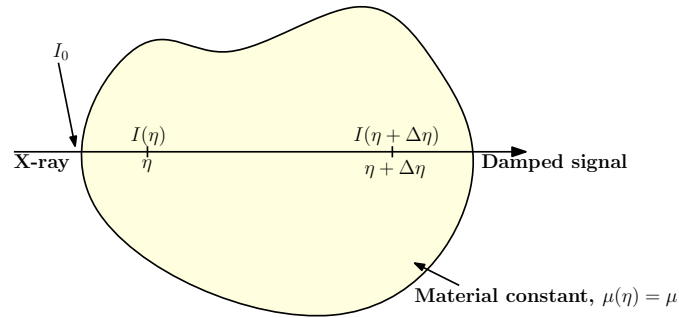


Fig. 7.15: CT: X-ray going straight through domain with attenuation constant $\mu(\eta) = \mu$.

coordinates ρ (distance) and θ (angle) the parametrisation of the line L is

$$L = \{\rho v + s v^\perp \mid s \in \mathbb{R}\}, \quad v = (\cos(\theta), \sin(\theta)), \quad v^\perp = (-\sin(\theta), \cos(\theta)).$$

The *Radon transform* of an attenuation coefficient represented by f is defined by

$$Rf(\rho, \theta) = \int_{\mathbb{R}} f(\rho v + s v^\perp) ds \quad \text{or simply} \quad Rf(L) = \int_L f ds.$$

Hence $Rf(L)$ exactly represents the intensity projection of f along the line L . That is, $Rf(L)$ is the data from which we want to reconstruct the body's inner structure, or simply the attenuation coefficient function f . Assuming that there is also Gaussian additive noise δ in the data we have the following problem

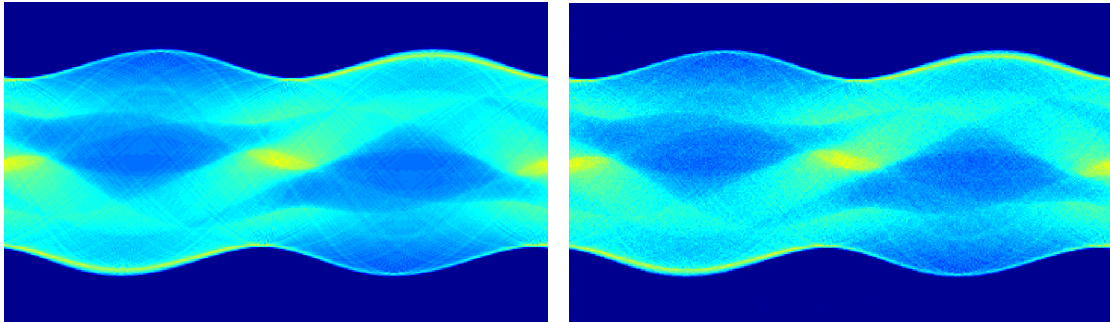
$$f_\delta = Rf + \delta.$$

The known data is f_δ from which we want to reconstruct f . Using the `fanbeamtomo`-function in the `AIR tools` package in `MatLab` we can generate simulated data for the CT problem for a reconstruction of the inner structure seen in Figure 7.16. The image is a 150×150 pixel image and hence in this case $\Delta x = \Delta y = \frac{1}{150}$.



Fig. 7.16: CT: Shepp-Logan phantom.

The `fanbeamtomo`-function generates a matrix representing the discretized Radon transform and a vector containing the projected data. To the data vector we add 2% Gaussian additive noise. The data can be shown in a sinogram as seen in Figure 7.17. Here the first-axis represents the angles and the second-axis represents the travelled distance. In this case we have used 360 different angles.



(a) Clean data.

(b) Noisy data (2% Gaussian additive noise).

Fig. 7.17: CT: Data for reconstruction.

Again we use the iteration scheme in (7.3.1) where now the matrix A is replaced by the matrix representing the Radon transform. Choosing $\Delta t = 2 \cdot 10^{-5}$ by trial and error we obtain the reconstructions of the phantom seen in Figure 7.18. We see that we obtain a quite good reconstruction even of the small details in the phantom. We also notice that we begin fitting to noise at the 200th iteration.

Especially for this problem we see the importance of the choice of time step size. Increasing the size of the time step to $\Delta t = 4 \cdot 10^{-5}$ we obtain the completely unreasonable results seen in Figure 7.19.

We complete this chapter with the open question of how to choose the time step parameter and when to stop the iteration.

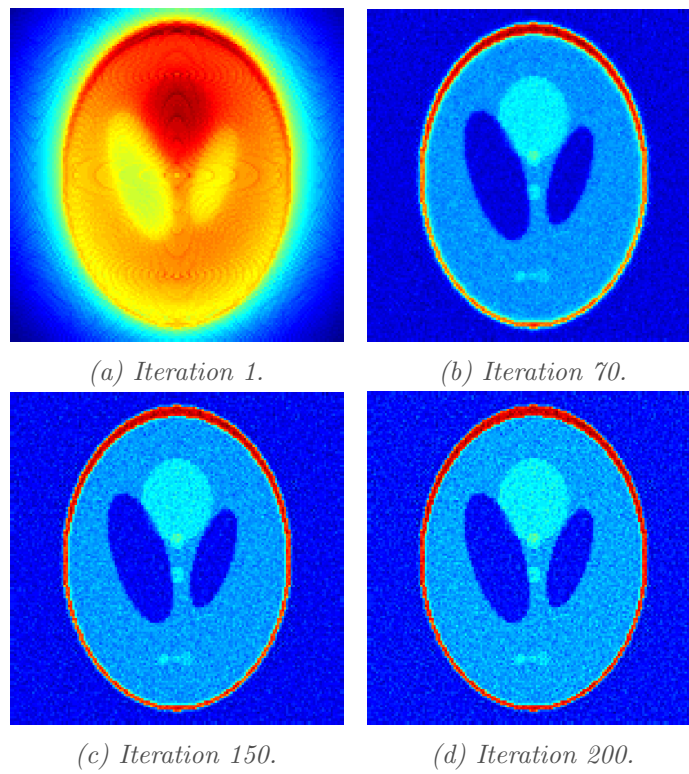


Fig. 7.18: CT: Standard Tikhonov reconstructions at different times with $\Delta t = 2 \cdot 10^{-5}$.

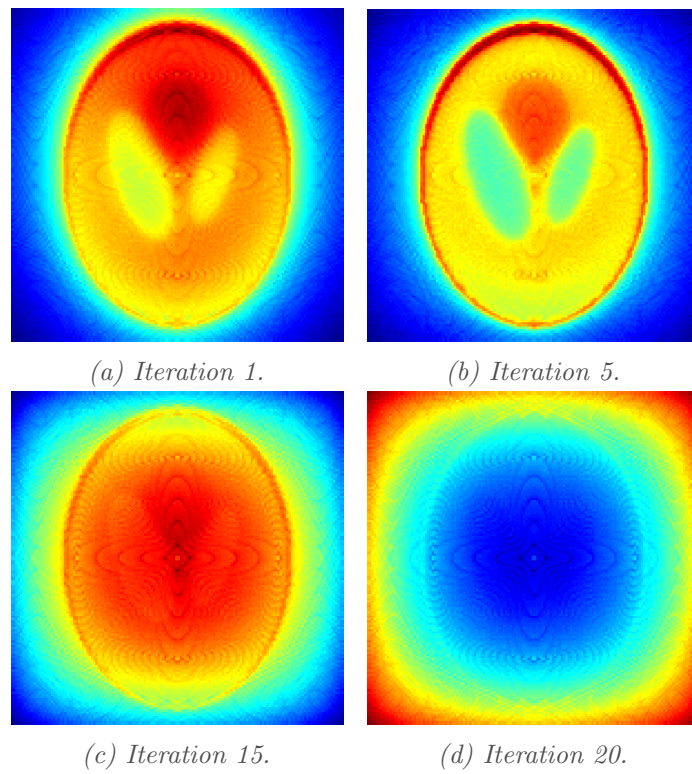


Fig. 7.19: CT: Standard Tikhonov reconstructions at different times with $\Delta t = 4 \cdot 10^{-5}$.



Fig. 7.20: Deblurring 2D: Standard Tikhonov reconstructions at different times with $\Delta t = 0.2$.

Discussion and Conclusion

Reconstruction problems like denoising, deblurring, computed tomography and other such issues can be analysed and solved through two different settings; the variational formulation or the non-linear diffusion method. For the variational formulation the reconstruction problem is seen as an inverse problem which is solved using a regularization method. In order to obtain a satisfactory solution we have to solve a minimization problem for the regularizing functional. We have proven a general theorem about the existence and uniqueness of such a minimizer.

Observing that the subdifferential of a functional at its minimizer contains the null element, the minimization problem can be turned into an Euler-Lagrange equation or an optimality condition for which the minimizer is a solution in distributional sense. The Euler-Lagrange equation can then, in the case of either denoising or in the case of standard Tikhonov regularization, be turned into a non-linear diffusion problem for which the solutions at different times correspond to the minimizers of an associated iterative regularization for the variational formulation. It can be discussed why we would like to use one or the other method for obtaining a reconstruction. The minimization problems obtained through the variational formulation are well-known and there exist lots of methods for solving such problems. Different penalization terms are well-studied and many regularization functions are already known to provide good reconstruction results. On the other hand, the non-linear diffusion process is very easy to implement using a finite difference scheme (in time). The PDE formulation might also give rise to new and better

reconstruction methods and regularization techniques. For this method we can think of the reconstruction process as a smoothing process retaining specified features and structures in the given data. For example we could specify that we want diffusion to take place in isotropic areas and not across edges in images for denoising problems.

In the case of a standard Tikhonov regularization, we ended up with a non-linear diffusion equation in the dual space of the solution space. There is still an open question about how to solve this problem. A conjecture about the existence of a solution has been stated and may be investigated further. For a general regularization function and a general reconstruction problem we have not yet been able to turn the optimality condition for a minimizer for the variational formulation into a diffusion process. Maybe it can be turned into a diffusion process for an operator evaluated at the sought solution or the evaluated at the time derivative of the sought solution from which the reconstruction can be generated. This is also still an open question.

For the implementation of the non-linear diffusion processes a finite difference method has been used. For this method we have seen that the time step size plays an important role. An optimal way of choosing this parameter should therefore be developed. Since there is a direct relation between the time step parameter for the discretized diffusion process and the regularization parameter for the variational formulation it could be possible to use methods like Morozov's discrepancy principle to choose the time step size. This is open for further investigation. In the numerical experiments the time step parameter is set to the same value in all iterations. This works fine, but is not optimal regarding time consumption. In tomography, for example, we would like to obtain a good reconstruction as fast as possible, so the time step parameter should be chosen optimally in each iteration. At first the parameter should be chosen small, in order to give priority to removing noise from the data. But as soon as almost all noise is removed and a more or less good reconstruction is obtained, the time step parameter should be increased in order to speed up the reconstruction process. This leads to the question of when to change the size of the time step size. Finally, we have seen that in order to obtain a good reconstruction the diffusion process has to be stopped at some kind of optimal time. If the diffusion process is continued for too long we obtain an extremely simplified version of the actual function we want to reconstruct. Therefore the diffusion process requires an optimal stopping criteria.

8.1 Future Work

As there are a lot of open questions in relation to this thesis, there is room for future work on the subject. The future work could consist of formulating non-linear diffusion processes for general reconstruction problems and penalization terms and

investigating the solvability of these. From the formulation of such non-linear diffusion processes, the aim is to be able to formulate new and possibly better regularization techniques. For the discretized non-linear diffusion processes it is important to understand how the time step parameter influences the iteration process and when the iteration process should be stopped in order to obtain an optimal numerical solution as fast as possible. For this, an algorithm for choosing an optimal time step parameter and an optimal stopping criteria should be suggested.

Bibliography

Articles and papers

- [1] M. Grasmir, *Relaxation of Nonlocal Integrals with Rational Integrand*s, ch. 4. Ph.D. Thesis, Leopold Franzens University Innsbruck, June 2006
- [2] M. G. Crandall and T. M. Liggett, *Generation of Semi-Groups of Nonlinear Transformations on General Banach Spaces*. American Journal of Mathematics, Vol. 93, No. 2 (Apr. 1971), pp. 265-298, 1971.
- [3] A. Dold and B. Eckmann, *Nonlinear Operators and the Calculus of Variations, Bruxelles 1975*. Lecture Notes in Mathematics, Vol. 543. Springer Verlag, Berlin, Heidelberg, New York, 1976.
- [4] R. T. Rockafellar, *Integral functionals, normal integrands and measurable selections*. In [3], pages 157-207, 1976.
- [5] O. Scherzer, *Scale Space Methods and Regularization for Denoising and Inverse Problems*. Advances in Imaging and Electron Physics, Vol. 128, pp. 445-530 (2003).
- [6] O. Scherzer and J. Weickert, *Relations Between Regularization and Diffusion Filtering*. Journal of Mathematical Imaging and Vision, Vol. 12, 43-63 (2000).
- [7] I. Miyadera, *Some remarks on semi-groups of nonlinear operators*. Tôhoku Mathematical Journal, Vol. 23 (1971), pp. 245-258, 1971.
- [8] I. Miyadera and S. Ôharu, *Approximation of semi-groups of nonlinear operators*. Tôhoku Mathematical Journal, Vol. 22 (1970), pp. 24-27, 1970.

-
- [9] K. Yosida, *On the differentiability and the representation of one-parameter semi-group of linear operators*. J. Math. Soc. Japan, 1 (1948), pp. 15-21, 1948.
- [10] E. Hille, *Functional analysis and semi-groups*. Amer. Math. Soc. Coll. Publ., New York, Vol. 31, 1948.
- [11] T. Kato, *Nonlinear semi-groups and evolution equations*. Journal of the Mathematical Society of Japan, Vol. 19 (1967), pp. 508-520, 1967.
- [12] T. Kato, *Accretive operators and nonlinear evolution equations in Banach spaces*. Proc. Symp. in Pure Math., Vol. 18, Part I, Amer. Math. Soc., Providence, Rhode Island, pp. 138-161, 1970.
- [13] J. Lindenstrauss, *On nonseparable reflexive Banach spaces*. Communicated by V. Klee, April 21, 1966.
- [14] B. Helffer, *Spectral theory and applications. An elementary introduction course. Bucarest version 2010.*. Université Paris-Sud, Département de Mathématiques, March 26, 2010.
- [15] M. Crandall and A. Pazy, *Semi-groups of nonlinear contractions and dissipative sets*. Journal of Functional Analysis, Vol. 3 (1996), pp. 345-353, 1996.
- [16] M. Grasmir and F. Lenzen, *Anisotropic Total Variation Filtering*. Springer Science+Business Media, LLC, Published online 15 May 2010.
- [17] M. Bergouioux, *On Pointcaré-Wirtinger inequalities in spaces of functions of bounded variation*. HAL archives-ouvertes, HAL Id: hal-00515451, <http://hal.archives-ouvertes.fr/hal-00515451v1> September 7, 2010.

Books

- [18] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*. Springer Science+Business Media, LLC, 2009.
- [19] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer Science+Business Media, LLC, 2010.
- [20] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces - Applications to PDEs and Optimization*. Society for Industrial and Applied Mathematics and the Mathematical Programming Society, 2006.
- [21] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing - Partial Differential Equations and the Calculus of Variations*. Springer Science+Business Media, LLC, Second Edition, 2006.

-
- [22] E. Kreyszig, *Introductory Functional Analysis with Applications*. John Wiley & Sons Inc., 1978.
- [23] E. Hansen, *Measure theory*. Department of Mathematical Sciences, University of Copenhagen, Fourth edition, 2009.
- [24] F. Andreau-Vaillo, V. Caselles, and J. M. Mazón, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*. Springer Basel AG, 2004.
- [25] B. Dacorogna, *Direct Methods in the Calculus of Variations*. Springer Science+Business Media, LLC, 2008.
- [26] I. Ciaranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*. Kluwer, Dordrecht, Mathematics and its Applications, Vol. 62, 1990.
- [27] L. C. Evans, *Partial Differential Equations*. American Mathematical Society, 2. edition, 2010.
- [28] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, New York, 2000.
- [29] A. J. Kurdila and M. Zabrankin, *Convex Functional Analysis*. Birkhäuser Verlag, 2005.
- [30] V. Barbu, *Nonlinear semigroups and differential equations in Banach spaces*. Editura Academiei Bucuresti România, Noordhoff International Publishing, 1976.
- [31] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, 1992.
- [32] R. A. Adams and J. F. Fournier, *Sobolev Spaces*. Elsevier Ltd., Second Edition, 2003.

Software

- [33] MatLab R2012b (8.0.0.783), August 22, 2012.
- [34] AIR tools package in MatLab, www2.compute.dtu.dk/~pcha/AIRtools/.

A

Theory

A.1 Sobolev Spaces and their Duals

In this section we characterize the dual spaces $(W^{1,p}(\Omega))'$. The section is based on [32, pp. 62-63]. First we introduce the notion of multi-indices. Given integers $n \geq 1$ and $m \geq 0$, let $N = N(n, m)$ be the number of multi-indices $\alpha = (\alpha_1, \dots, \alpha_n)$ such that $|\alpha| = \alpha_1 + \dots + \alpha_n \leq m$. For each multi-index α let Ω_α be a copy of $\Omega \subset \mathbb{R}^n$ in a different copy of \mathbb{R}^n such that each of the N domains Ω_α are mutually disjoint. Finally let $\Omega^{(m)}$ be the union of these N domains; $\Omega^{(m)} = \bigcup_{|\alpha| \leq m} \Omega_\alpha$. We further need the definition of the pairing:

$$(u, v) = \int_{\Omega} u(x)v(x) \, dx$$

for any functions u and v for which the integral makes sense. Note that the pairing makes sense for $u \in L^p(\Omega)$ and $v \in (L^p(\Omega))' = L^{p'}(\Omega)$ where

$$p' = \begin{cases} \infty, & p = 1, \\ p/(p-1), & 1 < p < \infty, \\ 1, & p = \infty. \end{cases}$$

Now we first give a characterization of the dual spaces $(L^p(\Omega^{(m)}))'$:

Theorem A.1.1 (Characterization of $(L^p(\Omega^{(m)}))'$). *To every $L \in (L^p(\Omega^{(m)}))'$, where $1 \leq p < \infty$, there corresponds a unique $v \in L^{p'}(\Omega^{(m)})$ such that for every $u \in L^p(\Omega^{(m)})$ we have*

$$L(u) = \int_{\Omega^{(m)}} u(x)v(x) \, dx = \sum_{|\alpha| \leq m} \int_{\Omega_\alpha} u_\alpha(x)v_\alpha(x) \, dx = \sum_{|\alpha| \leq m} (u_\alpha, v_\alpha),$$

where u_α and v_α are the restrictions of u and v , respectively, to Ω_α .

Proof. See [32, Thm. 2.44]. □

Finally we give the characterization of the dual spaces $(W^{1,p}(\Omega))'$:

Theorem A.1.2. *Let $1 \leq p < \infty$. For every $L \in (W^{1,p}(\Omega))'$ there exist elements $v \in L^{p'}(\Omega^{(m)})$ such that if the restriction of v to Ω_α is v_α , we have for all $u \in W^{1,p}(\Omega)$*

$$L(u) = \sum_{0 \leq |\alpha| \leq m} (D^\alpha u, v_\alpha).$$

Proof. See [32, Thm. 3.9]. □

A.2 Minimization Problems

In this section we introduce minimization problems and give necessary and sufficient conditions for the existence of solutions. The section is based on [20, Chap. 3], [21, pp. 30-40] and [18, Chap. 10]. All proofs of the theorems, which are not given here, can be found in these books.

A.2.1 Minimization Problems

Any minimization problem can be formulated as

$$\inf\{F(u) : u \in X\}$$

where X is the space in which we want to minimize F and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is an extended real-valued functional. The solution set for a given minimization problem for a functional $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is defined by

$$\arg \min F = \{\bar{u} \in X : F(\bar{u}) = \inf_X F(u)\}.$$

A.2.2 Lower Semi-Continuous Functionals

In the problem of minimizing an extended real-valued functional the concept of lower semi-continuity plays an important role. In order to introduce the concept of lower semi-continuous functionals and their properties we need to introduce the epigraph and lower level sets of extended real-valued functionals.

The definition of the epigraph and the lower level sets of an extended real valued functional is given by

Definition A.2.1 (Epigraph and Lower Level Sets). Let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be an extended real-valued functional. The *epigraph* of F is defined by

$$\text{epi}F = \{(u, \lambda) \in X \times \mathbb{R} : \lambda \geq F(u)\}$$

and the *lower γ -level set* is defined by

$$\text{lev}_\gamma F = \{u \in X : F(u) \leq \gamma\}.$$

Note that the solution set of a minimization problem can be expressed by the lower level sets as

$$\arg \min F = \bigcap_{\gamma > \inf_X F} \text{lev}_\gamma F$$

and the lower level sets can be expressed by the epigraph as

$$\text{lev}_\gamma F \times \{\gamma\} = \text{epi}F \cap (X \times \{\gamma\}).$$

Hence the epigraph plays an important role in minimization problems.

Now we introduce the concept of lower semi-continuity of an extended real-valued functional:

Definition A.2.2 (Lower Semi-Continuity). Let (X, τ) be a topological space and denote by $\mathcal{N}_\tau(u)$ the family of neighbourhoods of u for the topology τ . The functional $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be τ -lower semi-continuous (τ -lsc) at u if

$$\forall \lambda < F(u) \exists N_\lambda \in \mathcal{N}_\tau(u) : F(v) > \lambda \forall v \in N_\lambda.$$

If F is τ -lsc at every point of X , then F is said to be τ -lower semi-continuous on X .

In order to show that a functional is lower semi-continuous we can use the following proposition:

Proposition A.2.3. Let (X, τ) be a topological space and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ an extended real-valued functional. Then the following statements are equivalent:

- (i) F is τ -lower semi-continuous.
- (ii) $\text{epi}F$ is closed in $X \times \mathbb{R}$ (where $X \times \mathbb{R}$ is equipped with the product topology of τ on X and of the usual topology on \mathbb{R}).
- (iii) $\forall \gamma \in \mathbb{R}$, $\text{lev}_\gamma F$ is closed in (X, τ) .
- (iv) $\forall \gamma \in \mathbb{R}$, $\{u \in X : F(u) > \gamma\}$ is open in (X, τ) .
- (v) $\forall u \in X$, $F(u) \leq \liminf_{v \rightarrow u} F(v) := \sup \inf_{N \in \mathcal{N}_\tau(u), v \in N} F(v)$.

Proof. See [20, Prop. 3.2.2]. □

For a functional consisting of more terms, we can use the following proposition

Proposition A.2.4. Let $F, G : (X, \tau) \rightarrow \mathbb{R} \cup \{+\infty\}$ be two lower semi-continuous functionals. Then $F + G$ is still lower semi-continuous.

Proof. See [20, Prop. 3.2.4] □

Finally a lower semi-continuous functional is bounded from below in the following way:

Proposition A.2.5. *Let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper and lower semi-continuous functional. Then F is bounded from below by an affine functional. That is, there are $u' \in X'$ and $\beta \in \mathbb{R}$ such that*

$$\phi(u) \geq (u, u') + \beta, \forall u \in X.$$

Proof. See [30, Prop. 1.1]. □

A.2.3 Inf-Compact Functionals and Coercivity

The second important concept in minimization problems is the inf-compactness property which can be shown to be equivalent to the coercivity of a functional.

The definition of an inf-compact functional is:

Definition A.2.6 (Inf-Compactness). Let (X, τ) be a topological space and let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be an extended real-valued functional. The functional F is said to be τ -inf-compact if for any $\gamma \in \mathbb{R}$

$$\text{lev}_\gamma F = \{u \in X : F(u) \leq \gamma\}$$

is relatively compact in X for the topology τ .

A coercive functional is defined by

Definition A.2.7 (Coercivity). Let X be a normed linear space. A functional $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be *coercive* if $\lim_{\|u\| \rightarrow +\infty} F(u) = +\infty$.

Now the relation between inf-compactness and coercivity can be explained through the following proposition

Proposition A.2.8. *Let X be a normed space and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$. Then the following conditions are equivalent:*

- (i) F is coercive.
- (ii) For any $\gamma \in \mathbb{R}$, $\text{lev}_\gamma F$ is bounded.

Proof. See [20, Prop. 3.2.8]. □

Recalling that in a normed space the bounded sets are relatively compact if and only if the space is finite dimensional, we see that in the finite dimensional case coercivity and inf-compactness are equivalent conditions. In the infinite dimensional case bounded sets are relatively compact for the weak topology if X is a reflexive Banach space. Hence in this case coercivity is equivalent to weak inf-compactness.

A.2.4 Minimization Theorems

With the above definitions we are ready to state Weierstrass theorem which gives necessary and sufficient conditions for the existence of solutions to minimization problems:

Theorem A.2.9 (Weierstrass theorem). *Let (X, τ) be a topological space and let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be an extended real-valued functional which is τ -lower semi-continuous and τ -inf compact. Then, $\inf_X F > -\infty$ and there exists some $\bar{u} \in X$ which minimizes F on X :*

$$F(\bar{u}) \leq F(u), \forall u \in X.$$

Proof. See [20, Thm. 3.2.1]. □

According to proposition A.2.8, in the finite dimensional case we can just choose the usual topology on $X = \mathbb{R}^n$ and then show that F is lower semi-continuous and coercive. In the infinite dimensional case we can use the weak topology if X is a reflexive Banach space and then show that F is coercive and weakly lower semi-continuous. In order to show that F is weakly lower semi-continuous we need the concept of convexity. A convex functional is defined as:

Definition A.2.10 (Convexity). Let X be a linear space and let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$. Then F is said to be *convex* if for each $u, v \in X$ and each $\lambda \in [0, 1]$ we have

$$F(\lambda u + (1 - \lambda)v) \leq \lambda F(u) + (1 - \lambda)F(v).$$

It can be shown that for a convex functional, lower semi-continuity and weak lower semi-continuity are equivalent conditions:

Theorem A.2.11. *Let X be a normed linear space and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ a convex proper functional. Then the following conditions are equivalent:*

- (i) F is lower semi-continuous for the norm topology on X .
- (ii) F is lower semi-continuous for the weak topology on X .

Proof. See [20, Thm. 3.3.3]. □

This equivalence and Weierstrass theorem leads to the following convex minimization theorem in reflexive Banach spaces:

Theorem A.2.12 (Convex Minimization Theorem). *Let $(X, \|\cdot\|)$ be a reflexive Banach space and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ a convex, lower semi-continuous, and coercive functional. Then $\inf_X F > -\infty$ and there exists some $\bar{u} \in X$ which minimizes F on X :*

$$F(\bar{u}) \leq F(u), \forall u \in X.$$

Proof. See [20, Thm. 3.3.4]. □

In the case that F is strictly convex the solution of the minimization problem is actually unique.

Sometimes we can express a functional as the supremum over a sequence of functionals for which we can prove lower semi-continuity and convexity results. In this case we have the following lemma:

Lemma A.2.13. *Let X be a normed space and $(F_n)_{n \in \mathbb{N}}$ a family of functionals on X with values in $\mathbb{R} \cup \{+\infty\}$. Denote by $F = \sup_{n \in \mathbb{N}} F_n$ their pointwise supremum. Then the following two assertions hold*

- 1) *If every functional F_n is (strictly) convex, then F is (strictly) convex.*
- 2) *If every functional F_n is lower semi-continuous, then F is lower semi-continuous.*

Proof. See [20, Prop. 3.2.3]. □

A.2.5 Relaxed Minimization Problems

In general, inf-compactness and lower semi-continuity are two properties which are antagonist. That is, if $\tau_1 > \tau_2$, i.e. the topology τ_1 is stronger than the topology τ_2 , then

$$F \text{ is } \tau_1 - \text{inf-compact} \Rightarrow F \text{ is } \tau_2 - \text{inf-compact}$$

while

$$F \text{ is } \tau_2 - \text{lsc} \Rightarrow F \text{ is } \tau_1 - \text{lsc}.$$

Hence it can be difficult to choose a "good" topology for a minimization problem. A good topology might even not exist and in this situation we consider instead a relaxed minimization problem.

If the functional F to minimize fails to be lower semi-continuous for a topology τ which makes F τ -inf-compact, then we introduce instead the τ -relaxed functional $\mathcal{R}_\tau F$ which leads to a relaxed minimization problem:

Definition A.2.14 (Lower Envelope). Given the topological space (X, τ) and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$, the τ -relaxed functional of F is the greatest τ -lsc functional which minorizes F :

$$\mathcal{R}_\tau F = \sup\{G : X \rightarrow \mathbb{R} \cup \{+\infty\} : G \text{ is } \tau - \text{lsc and } g \leq F\}.$$

It can be proven that $\mathcal{R}_\tau F$ has the following sequential formulation if (X, τ) is metrizable:

Proposition A.2.15. *Let (X, τ) be a metrizable topological space and $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$. Then for any $u \in X$*

$$(\mathcal{R}_\tau F)(u) = \liminf_{v \rightarrow u} F(v) = \min\{\liminf_n F(u_n) : u_n \rightarrow u \text{ as } n \rightarrow \infty\}.$$

Proof. See [20, Prop. 3.2.6]. □

It can be shown that the relaxed functional satisfies the following properties:

Theorem A.2.16. *Let $F : (X, \tau) \rightarrow \mathbb{R} \cup \{+\infty\}$ be an extended real-valued functional. Then*

$$\inf_X F = \inf_X \mathcal{R}_\tau F.$$

More generally for any τ -open subset C of X

$$\inf_C F = \inf_C \mathcal{R}_\tau F.$$

Moreover,

$$\arg \min F \subset \arg \min \mathcal{R}_\tau F.$$

And finally if $(u_n)_{n \in \mathbb{N}}$ is a minimizing sequence for F containing a convergent subsequence u_{n_k} τ -converging to some $\bar{u} \in X$ then

$$(\mathcal{R}_\tau F)(\bar{u}) \leq (\mathcal{R}_\tau F)(u), \forall u \in X,$$

that is, \bar{u} is a minimum point for $\mathcal{R}_\tau F$.

Assume further that X is a reflexive Banach space and F is coercive. Then

- (i) $\mathcal{R}_\tau F$ is coercive*
- (ii) $\mathcal{R}_\tau F$ has a minimum point in X*
- (iii) Every minimum point for $\mathcal{R}_\tau F$ is the limit of a minimizing sequence for F .*

Proof. See [20, Prop. 3.2.7] and [21, Thm. 2.1.6]. □

Theorem A.2.16 suggests that instead of considering the original minimization problem

$$\inf\{F(u) : u \in X\}$$

we should consider the *relaxed minimization problem*

$$\min\{\mathcal{R}F(u) : u \in X\}.$$

This relaxed problem satisfies the conditions for the existence of a minimizer and therefore admits a solution even in the case where F is not lower semi-continuous. In case F is actually lower semi-continuous the relaxed functional coincides with F .

A.2.6 The Direct Method of the Calculus of Variations

The direct method of the calculus of variations is a method for solving minimization problems. Considering the minimization of $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ the steps are:

- 1) Construct a *minimizing sequence* $(u_n)_{n \in \mathbb{N}} \subset X$, i.e. a sequence such that $\lim_{n \rightarrow \infty} F(u_n) = \inf_{u \in X} F(u)$.
- 2) Establish that the sequence $(u_n)_{n \in \mathbb{N}}$ is relative compact w.r.t. some topology τ . Here the topology τ is chosen as strong as possible. The relative compactness follows if F is coercive. From the relative compactness of (u_n) it follows that there exists a subsequence u_{n_k} τ -converging to some $\bar{u} \in X$.
- 3) Prove that \bar{u} is a minimum point of F . This follows if F is τ -lower semi-continuous.

Hence as soon as we have established the τ -inf-compactness and τ -lower semi-continuity of F we only need to find a minimizing sequence of F in order to obtain a subsequence converging to a minimum point of F .

A.3 Properties of Subdifferentials

In this section we state some properties of the subdifferential of an operator. The section is based on [18, Chap. 10]. The first lemma states some calculation rules for the subdifferential:

Lemma A.3.1. *Let $F : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be convex and $\lambda > 0$. Then*

$$\partial(\lambda F)(u) = \lambda \partial F(u), \quad u \in X.$$

Let $F, G : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be convex and assume that there exists $v \in \mathcal{D}(F) \cap \mathcal{D}(G)$ such that F is continuous in v . Then

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

Let X and Y be locally convex spaces, $L \in B(X, Y)$, and $G : Y \rightarrow \mathbb{R} \cup \{+\infty\}$ convex. Assume that there exists $v \in Y$ such that $G(v) < \infty$ and G is continuous in v . Then

$$\partial(G \circ L)(u) = L^\#(\partial G(Lu)), \quad u \in X.$$

Here $L^\# : Y' \rightarrow X'$ is the dual-adjoint of L .

The next theorem gives a characterization of the subdifferential in special cases:

Theorem A.3.2. *Let f be a normal integrand on $\Omega \times \mathbb{R}^m$ such that $F : L^p(\Omega; \mathbb{R}^m) \rightarrow \mathbb{R} \cup \{+\infty\}$, $1 \leq p < \infty$, defined by*

$$F(v) = \int_{\Omega} f(x, v(x)) \, dx$$

is proper. Then

$$\partial F(v) = \{v' \in L^q(\Omega; \mathbb{R}^m) : v'(x) \in \partial f(x, v(x)) \text{ a.e.}\}.$$

Here $1/p + 1/q = 1$ and the subdifferential of f is understood to be computed only with respect to $v(x)$.

B

MatLab Implementation

In this chapter the code for the implementation of the diffusion processes for the numerical experiments is seen. The code is written in MatLab [33].

B.1 Denoising 1D

```
1 clear all; close all; clc;
2
3 % Initializing random generators
4
5 randn('state',sum(100*clock))
6 rand('state',sum(100*clock))
7
8 % Generating true unknown f
9
10 J = 100;
11 a = 0;
12 b = 1;
13 x = linspace(a,b,J)';
14 f_exact = double(x >= 0.1 & x <= 0.4) + double(x >= 0.6 & x <= 0.9);
15
16 % Plotting true unknown
17
18 figure(1)
```

```

19 plot(x,f_exact,'-b','linewidth',2)
20 axis([0 1 -0.2 1.2]);
21
22 % Generating noisy data
23
24 eta = 0.07;
25 e = randn(size(f_exact));
26 e = eta*norm(f_exact)*e/norm(e);
27 f_noise = f_exact + e;
28
29 % Plotting noisy data
30
31 figure(2)
32 plot(x,f_noise,'-r','LineWidth',2)
33 hold on
34 plot(x,f_exact,'-b','linewidth',1.5)
35 hold off
36
37 %% FDM for standard Tikhonov regularization with  $g(x,u,Du) = 1/2|Du|^2$ .
38
39 % Initialization
40
41 Delta_t = 0.000003;
42 N = 100; % Number of time steps
43 h = (b-a)/J;
44 r = Delta_t/h^2;
45 F = zeros(J,N+1);
46
47 F(:,1) = f_noise;
48
49 for n = 1:N
50
51     % Fixing boundary conditions
52
53     F(1,n+1) = r*F(2,n) + (1-r)*F(1,n);
54     F(J,n+1) = r*F(J-1,n) + (1-r)*F(J,n);
55
56     % Reconstruction
57
58     for j = 2:J-1
59         F(j,n+1) = (1-2*r)*F(j,n) + r*F(j-1,n) + r*F(j+1,n);
60     end
61
62 end
63
64 % Plotting reconstructions
65
66 figure(3)
67 plot(x,F(:,2),'-g','LineWidth',2)
68 hold on
69 plot(x,F(:,11),'-k','LineWidth',2)
70 hold on
71 plot(x,F(:,101),'-m','LineWidth',2)

```

```

72 hold on
73 plot(x, f_exact, '—b', 'LineWidth', 1.5)
74 hold off
75
76 %% FDM for TV regularization with  $g(x, u, Du) = |Du|(\omega)$ 
77
78 % Initialization
79
80 Delta_t = 0.00001;
81 N = 150; % Number of time steps
82 h = (b-a)/J;
83 r = Delta_t/h;
84 F = zeros(J, N+1);
85
86 F(:, 1) = f_noise;
87
88 for n = 1:N
89
90     % Fixing boundary conditions
91
92     F(1, n+1) = F(1, n) + r*csgn(F(2, n)-F(1, n));
93     F(J, n+1) = F(J, n) + r*(-csgn(F(J, n)-F(J-1, n)));
94
95     % Reconstruction
96
97     for j = 2:J-1
98         F(j, n+1) = F(j, n) + r*(csgn(F(j+1, n)-F(j, n))-csgn(F(j, n)-F(j-1, n)));
99     end
100
101 end
102
103 % Plotting reconstructions
104
105 figure(4)
106 plot(x, F(:, 2), '-g', 'LineWidth', 2)
107 hold on
108 plot(x, F(:, 61), '-k', 'LineWidth', 2)
109 hold on
110 plot(x, F(:, 151), '-m', 'LineWidth', 2)
111 hold on
112 plot(x, f_exact, '—b', 'LineWidth', 1.5)
113 hold off

```

B.2 Denoising 2D

```

1 clear all; close all; clc;
2
3 % Initializing random generators
4
5 randn('state', sum(100*clock));
6 rand('state', sum(100*clock));

```

```

7
8 % Importing image
9
10 load('lena512.mat')
11 f_exact = lena512;
12
13 % Plotting image
14
15 figure(1)
16 imagesc(f_exact);
17 colormap(gray);
18 axis image off
19 axis equal
20
21 % Adding 2% Gaussian noise
22
23 eta = 0.02;
24 e = randn(size(lena512));
25 e = eta*norm(f_exact)*e/norm(e);
26 f_noise = f_exact + e;
27
28 % Plotting noisy image
29
30 figure(2)
31 imagesc(f_noise)
32 colormap(gray)
33 axis image off
34 axis equal
35
36 %% FDM for standard Tikhonov regularization with  $g(x,u,Du) = 1/2|Du|^2$ .
37
38 % Initialization
39
40 J = size(f_noise,1);
41 K = size(f_noise,2);
42 N = 100; % Number of time steps
43 h1 = 1/J;
44 h2 = 1/K;
45 Delta_t = 0.0000001;
46 r1 = Delta_t/h1^2;
47 r2 = Delta_t/h2^2;
48 F = zeros(J,K,N+1);
49
50 F(:, :, 1) = f_noise;
51
52 for n = 1:N
53
54     % Fixing boundary conditions
55
56     % Corners
57
58     F(1,1,n+1) = r1*(F(2,1,n)+F(1,1,n))+r2*(F(1,2,n)+F(1,1,n))...
59         +(1-2*r1-2*r2)*F(1,1,n);

```

```

60     F(1,K,n+1) = r1*(F(2,K,n)+F(1,K,n))+r2*(F(1,K,n)+F(1,K-1,n))...
61         +(1-2*r1-2*r2)*F(1,K,n);
62     F(J,1,n+1) = r1*(F(J,1,n)+F(J-1,1,n))+r2*(F(J,2,n)+F(J,1,n))...
63         +(1-2*r1-2*r2)*F(J,1,n);
64     F(J,K,n+1) = r1*(F(J,K,n)+F(J-1,K,n))+r2*(F(J,K,n)+F(J,K-1,n))...
65         +(1-2*r1-2*r2)*F(J,K,n);
66
67     % Edges
68
69     for j = 2:J-1
70
71         F(j,1,n+1) = r1*(F(j+1,1,n)+F(j-1,1,n))+r2*(F(j,2,n)+F(j,1,n))...
72             +(1-2*r1-2*r2)*F(j,1,n);
73         F(j,K,n+1) = r1*(F(j+1,K,n)+F(j-1,K,n))+r2*(F(j,K,n)+F(j,K-1,n))...
74             +(1-2*r1-2*r2)*F(j,K,n);
75
76     end
77
78     for k = 2:K-1
79
80         F(1,k,n+1) = r1*(F(2,k,n)+F(1,k,n))+r2*(F(1,k+1,n)+F(1,k-1,n))...
81             +(1-2*r1-2*r2)*F(1,k,n);
82         F(J,k,n+1) = r1*(F(J,k,n)+F(J-1,k,n))+r2*(F(J,k+1,n)+F(J,k-1,n))...
83             +(1-2*r1-2*r2)*F(J,k,n);
84
85     end
86
87     % Reconstruction
88
89     for k = 2:K-1
90         for j = 2:J-1
91             F(j,k,n+1) = r1*(F(j+1,k,n)+F(j-1,k,n))...
92                 +r2*(F(j,k+1,n)+F(j,k-1,n))+(1-2*r1-2*r2)*F(j,k,n);
93         end
94     end
95
96 end
97
98 % Plotting reconstruction
99
100 figure(3)
101 imagesc(F(:, :, 101))
102 colormap(gray);
103 axis image off
104 axis equal
105
106 %% FDM for TV regularization  $g(x,u,Du) = |Du|(\omega)$ 
107
108 % Initialization
109
110 J = size(f.noise,1);
111 K = size(f.noise,2);
112 N = 150; % Number of time steps

```



```
166         end
167     end
168
169 end
170
171 % Plotting reconstruction
172
173 figure(7)
174 imagesc(F(:, :, 101))
175 colormap(gray);
176 axis image off
177 axis equal
```

B.3 Deblurring 1D

```
1 clear all; close all; clc;
2
3 % Initializing random generators
4
5 randn('state', sum(100*clock))
6 rand('state', sum(100*clock))
7
8 % Generating true unknown f
9
10 J = 100;
11 a = 0;
12 b = 1;
13 x = linspace(a,b,J)';
14 f_exact = double(x >= 0.1 & x <= 0.4) + double(x >= 0.6 & x <= 0.9);
15
16 % Plotting true unknown
17
18 figure(1)
19 plot(x, f_exact, '—b', 'linewidth', 2)
20 axis([0 1 -0.2 1.2])
21
22 % Generating blurred f
23
24 beta = 0.05;
25 A = zeros(J,J);
26 h = x(2)-x(1);
27
28 for k = 1:J,
29
30     A(:,k) = h/sqrt(2*pi*beta^2)*exp(-1/(2*beta^2)*(x - x(k)).^2);
31
32 end
33
34 f_blur = A*f_exact;
35
36 % Plotting blurred f
```

```

37
38 figure(2)
39 plot(x,f_exact,'-b','linewidth',2)
40 hold on
41 plot(x,f_blur,'-c','linewidth',2)
42 axis([0 1 -0.2 1.2])
43 hold off
44
45 % Adding 2% Gaussian noise
46
47 eta = 0.02;
48
49 e = randn(size(f_blur));
50 e = eta*norm(f_blur)*e/norm(e);
51 f_blur_noise = f_blur + e;
52
53 % Plotting noisy data
54
55 figure(3)
56 plot(x,f_exact,'-b','linewidth',2)
57 hold on
58 plot(x,f_blur_noise,'-r','linewidth',2)
59 axis([0 1 -0.2 1.2])
60 hold off
61
62 %% Asymptotic Tikhonov-Morozov filtering technique
63
64 Delta_t = 0.005;
65 N = 20000; % Number of time steps
66 F_blur = zeros(J,N+1);
67
68 for n = 1:N
69
70     % Reconstruction
71
72     F_blur(:,n+1) = F_blur(:,n)-Delta_t*A'*...
73         (A*F_blur(:,n)-f_blur_noise);
74
75     % Fixing boundary conditions
76
77     F_blur(1,n+1) = F_blur(2,n+1);
78     F_blur(J,n+1) = F_blur(J-1,n+1);
79
80 end
81
82 % Plotting reconstructions
83
84 figure(4)
85 plot(x,f_exact,'-b','linewidth',2)
86 hold on
87 plot(x,F_blur(:,501),'-g','linewidth',2)
88 hold on
89 plot(x,F_blur(:,7001),'-k','linewidth',2)

```

```
90 hold on
91 plot(x,F_blur(:,20001),'-m','linewidth',2)
92 axis([0 1 -0.2 1.2])
93 hold off
```

B.4 Deblurring 2D and Computed Tomography

```
1 clear all; close all; clc;
2
3 % Initializing random generators
4
5 randn('state',sum(100*clock));
6 rand('state',sum(100*clock));
7
8 % Importing Lena image
9
10 load('lena512.mat')
11 f_exact = lena512;
12
13 % Plotting image
14
15 figure(1)
16 imagesc(f_exact);
17 colormap(gray);
18 axis image off
19 axis equal
20
21 % Generating blurred image
22
23 N = size(f_exact,1);
24 Tm = mblur(N,6); % motion blur matrix
25
26 f_m = Tm*f_exact(:);
27 f_m = reshape(f_m,N,N);
28
29 % Plotting blurred image
30
31 figure(3)
32 imagesc(f_m)
33 colormap(gray)
34 axis image off
35
36 % Generating computed tomography data
37
38 [A b tomo theta] = fanbeamtomo(150); % tomography test problem
39
40 % Plotting sinogram
41
42 figure(7)
43 image(reshape(b,size(b,1)/size(theta,2),size(theta,2)));
44 axis image off
```

```

45
46 % Adding 2% Gaussian noise
47
48 eta = 0.05;
49
50 e_m = randn(size(f_m));
51 e_m = eta*norm(f_m)*e_m/norm(e_m);
52 f_m_noise = f_m + e_m;
53
54 e_tomo = randn(size(b));
55 e_tomo = eta*norm(b)*e_tomo/norm(e_tomo);
56 tomo_noise = b + e_tomo;
57
58 % Plotting noisy data
59
60 figure(5)
61 imagesc(f_m_noise)
62 colormap(gray)
63 axis image off
64
65 figure(8)
66 image(reshape(tomo_noise,size(tomo_noise,1)/size(theta,2),size(theta,2)));
67 axis image off
68
69 %% Asymptotic Tikhonov–Morozov filtering technique
70
71 % Initialization
72
73 J = size(f_m_noise,1);
74 K = size(f_m_noise,2);
75 Delta_t_tomo = 0.00004;
76 Delta_t_blur = 0.5;
77 N = 200; % Number of time steps
78 F_m_vector = zeros(J*K,N+1);
79 F_m = zeros(J,K,N+1);
80 f_m_noise_vector = f_m_noise(:);
81 Tomo_vector = zeros(150*150,N+1);
82 Tomo = zeros(150,150,N+1);
83 tomo_noise_vector = tomo_noise(:);
84
85 for n = 1:N
86
87     % Reconstruction
88
89     F_m_vector(:,n+1) = F_m_vector(:,n)-Delta_t_blur*Tm'*...
90         (Tm*F_m_vector(:,n)-f_m_noise_vector);
91     F_m(:, :,n+1) = reshape(F_m_vector(:,n+1),J,K);
92
93     Tomo_vector(:,n+1) = Tomo_vector(:,n)-Delta_t_tomo*A'*...
94         (A*Tomo_vector(:,n)-tomo_noise_vector);
95     Tomo(:, :,n+1) = reshape(Tomo_vector(:,n+1),150,150);
96
97     % Fixing boundary conditions

```

```
98
99     F_m(1, :, n+1) = F_m(2, :, n+1);
100    F_m(J, :, n+1) = F_m(J-1, :, n+1);
101    F_m(:, 1, n+1) = F_m(:, 2, n+1);
102    F_m(:, K, n+1) = F_m(:, K-1, n+1);
103
104    Tomo(1, :, n+1) = Tomo(2, :, n+1);
105    Tomo(150, :, n+1) = Tomo(150-1, :, n+1);
106    Tomo(:, 1, n+1) = Tomo(:, 2, n+1);
107    Tomo(:, 150, n+1) = Tomo(:, 150-1, n+1);
108
109 end
110
111 % Calculating minimum error
112
113 error_m = zeros(1, N+1);
114 error_tomo = zeros(1, N+1);
115
116 for j = 1:N+1
117     error_m(j) = norm(f_exact - F_m(:, :, j), 2);
118     error_tomo(j) = norm(reshape(tomo, 150, 150) - Tomo(:, :, j), 2);
119
120 end
121
122 % Determining time step of minimum error
123
124 index_m = find(error_m == min(error_m));
125 index_tomo = find(error_tomo == min(error_tomo));
126
127 % Plotting minimum error image reconstruction
128
129 figure(12)
130 imagesc(F_m(:, :, index_m))
131 colormap(gray);
132 axis image off
133 axis equal
134
135 % Plotting minimum error tomography reconstruction
136
137 figure(13)
138 imagesc(Tomo(:, :, index_tomo))
139 axis image off
140 axis equal
```