JOHANNES KEPLER
UNIVERSITÄT LINZ | JKU

TNF

Technisch-Naturwissenschaftliche
Fakultät

Double Regularised Total Least Squares Method

DISSERTATION

zur Erlangung des akademischen Grades

Doktor

im Doktoratsstudium der

Technischen Wissenschaften

Eingereicht von:
Ismael Rodrigo Bleyer

Angefertigt am:
Doctoral Program "Computational Mathematics"

Beurteilung:
Univ.-Prof. Dr. Ronny Ramlau (Betreuung)
Univ.-Prof. Dr. Bernd Hofmann

Linz, 20th January 2014

To my beloved parents *Murillo* and *Zerinete*.

"It is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly..."

The Man in the Arena - April 23, 1910
Theodore Roosevelt

Abstract

Over the past years, many researchers in the field of inverse problems have concentrated on the efficiency of solving *ill-posed* problems both in sciences and industry. The main goal is to develop so-called *regularisation methods* to cope with numerical instabilities due to lack of measurements or unavoidable noise in practical problems. However, the operator is often also not known exactly, e.g., due to the discretisation error or the approximation of the mathematical model.

Therefore, to achieve reasonable results in the case where both the data and the operator are contaminated by some noise, one may be interested in reconstructing the *total least squares* (TLS) solution, which is a successful approach for well-posed linear problems. Additionally, many *regularised TLS* approaches have been considered to stabilise ill-posed problems.

In this thesis we consider operator equations in the infinite-dimensional setting where the operator can be characterised mainly by a function. For the stable reconstruction we propose the use of a *Tikhonov-type* functional with a generalised misfit term based on TLS and one additional penalty term which promotes sparsity. We refer to the novel technique as *double regularised total least squares*, or shortly, dbl-RTLS. Using an appropriate parameter choice rule for the two regularisation parameters we are able to derive convergence rates not only for the function, but also for the operator.

Moreover, we discuss computational aspects and we focus on the efficient numerical implementation with particular emphasis on the alternating minimisation strategy for solving not only the proposed method, but a vast class of optimisation problems: the minimisation of a bilinear nonconvex functional over two variables. The performance of this approach is illustrated for convolution problems.

Keywords: ill-posed problems, noisy operator, noisy data, regularised total least squares, alternating minimisation, wavelets, soft-shrinkage operator, subgradients, integral equation, convolution.

Kurzfassung

Sowohl in Wissenschaft als auch Industrie haben in letzter Zeit viele Forschende im Gebiet der "Inversen Probleme" ihr Hauptaugenmerk auf die effiziente Lösung schlecht gestellter Probleme gelegt. Das Hauptziel ist, so genannte Regularisierungsverfahren zu entwickeln, welche es ermöglichen, numerische Instabilitäten zu beherrschen, zum Beispiel bedingt durch unzureichende Messwerte oder in der Praxis unvermeidliches Rauschen. Oft ist jedoch auch der Operator nicht exakt bekannt, beispielsweise durch Diskretisierungsfehler oder der Approximation des Problems durch das mathematische Modell.

Sind sowohl Daten als auch der Operator selbst verrauscht, stellt die Methode der "Totalen kleinsten Quadrate" eine Möglichkeit dar, sinnvolle Ergebnisse zu erhalten. Dies wurde bereits für gut gestellte Probleme angewendet. Zudem wurden regularisierte "TLS-"Varianten bereits in Betracht gezogen um schlecht gestellte Probleme zu stabilisieren.

In dieser Dissertation betrachten wir Operatorgleichungen in unendlichdimensionalen Räumen, bei denen der Operator hauptsächlich durch eine Funktion charakterisiert werden kann. Zur stabilen Rekonstruktion nutzen wir ein Tikhonov-Funktional mit auf TLS basierendem, verallgemeinertem Fehlerterm und einem zusätzlichem Strafterm, der eine sogenannte "sparse" Lösung begünstigt. Wir bezeichnen dieses neuartige Verfahren "double regularised TLS", kurz "dbl-RTLS". Mittels geeigneter Methoden zur Wahl der beiden Regularisierungsparameter sind wir im Stande Konvergenzraten nicht nur für die Funktion, sondern auch für den Operator herzuleiten.

Überdies und mit Fokus auf effiziente numerische Implementation. Besondere Betonung liegt auf der alternierenden Minimierungs-Strategie, die nicht nur zur Lösung des vorgestellten Problems, sondern auch einer weiten Klasse von Optimierungsproblemen dient, der Minimierung bilinearer konvexer Funktionale über zwei Variablen. Die Leistungsfähigkeit dieses Vorgehens wird an Faltungsproblemen verdeutlicht.

Acknowledgements

First of all I offer my gratitude to my supervisor, Dr. Ronny Ramlau, who selected me among other candidates to join his project and team about four years ago. His support, guidance, advice throughout the research project, as well as his patience are greatly appreciated. Indeed, without him I would not be able to put the topic together and complete this thesis. One could not wish for a kinder supervisor.

Dr. Antônio Leitão, my former supervisor, for introducing me to inverse problems, for all his support and decision to study abroad, for believing on me, for encouraging me, and foremost for being a friend. I have no more words, "muito obrigado!" (in Portuguese).

Of course this project would not have been possible without the support provided by the Doctoral Program Computational Mathematics and Austrian Science Funds FWF (*Fonds zur Förderung der wissenschaftlichen Forschung*). I would like to say thank you to our director Dr. Peter Paule, who always had a smile in his face to share with us, and our secretary Gabriela Hahn, who always help me and listened to my struggles whenever I needed someone to talk to. Additionally, I should mention all the help from the lovely secretaries Monika Bayer and Doris Nikolaus.

I also recognise the following outstanding researches for being a great host, for their time, valuable discussion, and their departments to provide local support during my scientific stay; specifically, Dr. Shuai Lu (Fudan University, China), Dr. Bernd Hofmann (Chemnitz University of Technology, Germany), Dr. Antônio Leitão (Federal University of Santa Catarina, Brazil) and Dr. Samuli Siltanen (University of Helsinki, Finland).

I have been blessed in my daily work with a friendly group of student fellows, officemates and workmates. I surely will keep in mind not only our valuable mathematical discussions but also the fun and good time we spent in conferences, workshops and during our coffee breaks, lunch, dinner parties and out for a beer. Thanks for being there: Stephan Anzengruber, Mariya Zahriy, Valeriya Naumova, Irina Nikolova, Daniel Gerth, Stefan Takacs, Clemens Ho-

freither, and many others.

Beyond mathematics I have to acknowledge the great companion I found also outside of my work environment. Mainly during German courses and where I used to live - in special to all residents from *Raab-Heim* and *KHG-Heim*, each of you had somewhat an influence on me. I am afraid of either running out of space or forgetting some names, therefore I will not list names here. Nevertheless I would like to say how important it was for me to share my time and stories with you, and mainly to call you as my friend.

Last but not least, I would like to thank God for bringing the right people into my life and for my parents for supporting me throughout my life and all my studies, even when we live more than 10 000 kilometres apart, for believing on me and for loving me.

Contents

Chapter 1

A Tour on Inverse Problems

> "As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality."
>
> *Albert Einstein*

In this chapter we invite you to join our tour on inverse problems. We introduce a few names and their contribution to the nowadays called "Inverse Problems" research field, as well as the milestones in mathematics history.

Elementary definitions and the basic idea behind regularisation techniques are found in the fist part of this chapter, as a matter of introduction.

The reader familiar with the basic concepts can jump to Section 1.3. Here we summarise what has been done and what is missing in the literature, in order to allocate our work among the competitive approaches. Subsequently we list the main contribution found in this thesis and how the results are organised by chapters.

1.1 Inverse Problems

The field of inverse problems was first discovered and introduced by Soviet-Armenian physicist, and one of the founders of theoretical astrophysics, Viktor Amazaspovich Hambardzumyan (1908-1996), according to [17].

While still a student, Hambardzumyan thoroughly studied the theory of atomic structure, the formation of energy levels, and the Schrödinger equation and its properties. When he mastered the theory of eigenvalues of differential equations, he pointed out the apparent analogy between discrete energy levels and the eigenvalues of differential equations. He then asked

> "Given a family of eigenvalues, is it possible to find the form of the equations whose eigenvalues they are?"

Figure 1.1: Viktor Hambardzumyan

Essentially Hambardzumyan was examining the *inverse* Sturm-Liouville *problem*, which dealt with determining the equations of a vibrating string. This paper was published in 1929 in the German physics journal *Zeitschrift für Physik* and remained in oblivion for a rather long time.

Furthermore, the usage of the term *inverse problem* leads to a natural question "inverse of what?". Quoting [49], one calls two problems *inverse* to each other if the formulation of each of them requires full or partial knowledge of the other. For mostly historic reasons, one might call one of the problems (usually the simpler one or the one which was studied earlier) the *direct problem*, whereas the other one is the *inverse problem*.

There are two different motivations, pointed in [28], for studying such inverse problems: first, one wants to know past states or parameters of a physical system. Second, one wants to find out how to influence a system via its present state or via parameters in order to steer it to a desired state in the future. In summary, one might say that

> *inverse problems are concerned with determining causes for a desired or an observed effect.*

Another name to play an important role here is Hadamard. The major contribution in our field from the French mathematician Jacques Salomon Hadamard (1865 - 1963) is found in [39]. Hadamard introduced the concept

of a *well-posed problem*, originally called "correct set" as the discussions in Chapter I of his Lectures on *Cauchy's Problem in Linear Partial Differential Equations*. It represented a significant step forward not only in the classification of problems associated with differential equation, singling out those with sufficient general properties of existence, uniqueness and (by implication) stability of solutions. Hadamard observes:

> "But it is remarkable, on the other hand, that a sure guide is found in physical interpretation: an analytic problem always being *correctly set*, in our use of the phrase, when it is the translation of some mechanical or physical question."

Figure 1.2: Jacques Hadamard

The concept of well-posedness have been extensively studied over the last years (see e.g., [91, 28, 51]). To be more accurate, we consider an forward operator F (linear or non-linear) defined between two metric[1] spaces $\mathcal{U}$ and $\mathcal{H}$, so that the concept of *solution* of any quantitative problem usually ends in finding the "solution" u from given "initial data" g which belongs to the range of $F(u)$. The fundamental terminology of determining the solution is said to

[1] this definition remains valid for topological spaces.

be **well-posed on the pair of metric spaces** $(\mathcal{U}, \mathcal{H})$ if the following three conditions are satisfied:

existence: for every element $g \in \mathcal{H}$, there exists a solution u in the space $\mathcal{U}$;

uniqueness: for all admissible data, the solution is unique;

stability: the problem is stable on the spaces $(\mathcal{U}, \mathcal{H})$, i.e., the solution depends continuously on the data.

Problems that do not satisfy (at least one of) the statements listed above are said to be **ill-posed**.

It should be pointed out that the definition of an ill-posed problem is stated *with respect to a given pair of metric spaces* $(\mathcal{U}, \mathcal{H})$ since the same problem may be well-posed in other metrics.

The major issue raises whenever the underlying problem is not exactly known, e.g., we only know approximately the right-hand side g. This is often the case in practical problems, where the initial data is obtained from measurements, which either lacks in precision or may contain additionally undesirable noise.

The authors of [28] add the following remark for ill-posed problems:

> One is usually not too much concerned with the violation of *existence*, although of course also existence of a solution (for exact data) is an important requirement. It can usually be enforced by relaxing the notion of a solution at least for exact data, while for perturbed data, the problem has to be "regularised" and hence changed anyway.
>
> Violation of *uniqueness* is considered to be a little more serious. If a problem has several solutions, one either has to decide which one is of interest (e.g., the one with smallest norm, which is appropriate for some, but not all application). Even if this property is fulfilled when the data are measured "everywhere", non-uniqueness is usually introduced by the need for discretisation.
>
> For restoring *stability*, however, one has to change the topology of the spaces, which is in many cases impossible because of the presence of measurement. At first glance, it seems to be impossible to compute the solution of a problem numerically if the solution of the problem does not depend continuously on the data, then one has to expect that the numerical method (as one would use for a well-posed problem) becomes unstable. One has to keep in mind that no mathematical trick can make an inherently unstable problem stable. All that a *regularisation method* can do is to recover partial information about the solution as stably as possible, providing the right compromise between accuracy and stability.

It is very important to mention that many interesting and important *inverse* problems in science lead to *ill-posed* problems, while the corresponding direct problems are well-posed.

Examples of archetypal well-posed problems include the Dirichlet problem for Laplace's equation, and the heat equation with specified initial conditions. These might be regarded as 'natural' problems in that there are physical processes that solve these problems. By contrast the inverse heat equation, deducing a previous distribution of temperature from final data is not well-posed in that the solution is highly sensitive to changes in the final data. More examples can be found in classical books as [91, 70, 28, 51].

1.2 Regularisation Methods

The general methods of mathematical analysis were best adapted to the solution of well-posed problems and they are no longer meaningful in most applications in the sense of ill-posed problems. One of the earliest works in this field and the most outstanding was done by Andrey Nikolayevich Tikhonov (1906-1993). He succeeded in giving a precise mathematical definition of *approximated solution* for general classes of such problems and in constructing "optimal" solutions.

Figure 1.3: Andrey Tikhonov

Tikhonov was a Soviet and Russian mathematician. He made important contributions in a number of different fields in mathematics, e.g., in topology, functional analysis, mathematical physics, and certain classes of ill-posed problems. Certainly, *Tikhonov regularisation*, the most widely used method to solve ill-posed inverse problems, is named in his honour.

Nevertheless, we should make a note that Tikhonov regularisation has been invented independently in many different contexts. It became widely known from its application to integral equations from the work of Tikhonov [90] and David L. Phillips [75]. Some authors use the term Tikhonov-Phillips regularisation. The finite dimensional case was expounded by Arthur E. Hoerl [43], who took a statistical approach, and by Manus Foster [31], who interpreted this method as a Wiener-Kolmogorov filter. Following Hoerl, it is known in the statistical literature as *ridge regression*.

Obviously, the equation

$$F(u) = g \tag{1.1}$$

has a solution on $\mathcal{U}$ only for those elements g that belong to the set $\mathscr{R}(F)$. In the case that we only know an approximation or measurement g_δ instead of g, we can only speak of finding an approximate solution for $F(u) \approx g_\delta$.

For now we consider only noise on the right-hand side and we call such problems *genuinely ill-posed problems*. Moreover, for theoretical results, we assume that the measurement g_δ differs from the exact initial data g by no more than δ, that is,

$$\|g - g_\delta\| \leq \delta \tag{1.2}$$

and we call the numerical parameter **noise level**.

Generally the measurement g_δ does not belong to the range of F. Even if does, due the ill-posedness of the problem, the *generalised*[2] *inverse* operator $F^\dagger$ is unbounded and therefore $F^\dagger(g_\delta)$ is not a good approximation of the *best-approximated solution* $u^\dagger := F^\dagger(g)$. Under these conditions we have to build an approximation to the (generalised) inverse operator which is stable under small perturbation on the initial data and still provides a reasonable approximated solution.

In mathematical terms, a *regularisation* of $F^\dagger$ is the approximation of an ill-posed problem by a parameter-dependent family $\{\mathcal{R}_\alpha\}$ of neighbouring well-posed problems and we replace, intuitively, the approximation given from the unbounded operator $F^\dagger$ by the *regularised solution* $u_\delta^\alpha := \mathcal{R}_\alpha(g_\delta)$. Further we shall exemplify such family.

The parameter α introduced above is called *regularisation parameter*. It has a crucial play on the regularisation method, the "art" of finding the right compromise between accuracy and stability. If it is properly chosen, we can

[2] also called pseudo-inverse, e.g., the most widely known type of matrix pseudo-inverse is the Moore-Penrose.

guarantee the most desirable convergence result: as the noise level δ decreases to zero, u_δ^α tends to $u^\dagger$. On the upcoming Section 1.2.1 we shall give more details about how to choose it "appropriately".

In order to give a proper definition of *regularisation method*, as in [28] we restrict to the case of a linear operator A and we want to solve the linear system $Au = g$ from noisy data g_δ.

Definition 1.2.1. *Let $A : \mathcal{U} \to \mathcal{H}$ be a bounded linear operator between the Hilbert spaces $\mathcal{U}$ and $\mathcal{H}$, $\alpha_0 \in (0, \infty)$, let*

$$\mathcal{R}_\alpha : \mathcal{H} \to \mathcal{U}$$

be a continuous (not necessarily linear) operator. The family $\{\mathcal{R}_\alpha\}$ is called a ***regularisation*** *or a regularisation operator (for $A^\dagger$), if, for all $g \in \mathscr{D}(A^\dagger)$, there exists a parameter choice rule $\alpha = \alpha(\delta, g_\delta)$ such that*

$$\limsup_{\delta \to 0} \left\{ \left\| \mathcal{R}_{\alpha(\delta, g_\delta)} g_\delta - A^\dagger g \right\| \mid g_\delta \in \mathcal{H}, \|g - g_\delta\| \leq \delta \right\} = 0 \tag{1.3}$$

holds. Here,

$$\alpha : \mathbb{R}^+ \times \mathcal{H} \to (0, \alpha_0)$$

in such that

$$\limsup_{\delta \to 0} \{\alpha(\delta, g_\delta) \mid g_\delta \in \mathcal{H}, \|g - g_\delta\| \leq \delta\} = 0. \tag{1.4}$$

For specific $g \in \mathscr{D}(A^\dagger)$, a pair $(\mathcal{R}_\alpha, \alpha)$ is called a (convergent) ***regularisation method*** *(for solving $Au = g$) if (1.3) and (1.4) hold.*

We want the convergence $u^\dagger$ as $\mathcal{R}_\alpha(g_\delta) \to A^\dagger g$ for a specific parameter choice, which will be described with more details shortly. On the following we shall give an example of a regularisation method, namely, the Tikhonov regularisation.

The two most popular regularisation methods are *spectral cut-off regularisation* (also called truncated singular value decomposition) and *Tikhonov regularisation* (also called ridge regression or Wiener filtering in certain contexts). The latter method essentially consists of two parts: one *discrepancy* term which minimises the residue (e.g., based on least squares method) and one *regularisation* term which selects among all possible solutions one with desirable properties while adding stabilisation to the procedure. More precisely, the Tikhonov regularised solution u_δ^α is given by the minimiser of the following functional

$$J_\alpha^\delta(u) := \frac{1}{2} \|Au - g_\delta\|^2 + \frac{\alpha}{2} \|Lu\|^2 \tag{1.5}$$

where L is a continuous and boundedly invertible operator. Often L is assumed to be the identity operator.

The first order optimality condition reads as

$$A^*(Au - g_\delta) + \alpha L^* L u = 0. \tag{1.6}$$

This formula can be recast as a normal equation (of second kind) and therefore the solution u_δ^α has the close form

$$u_\delta^\alpha := (A^*A + \alpha L^* L)^{-1} A^* g_\delta \tag{1.7}$$

or more explicitly, this regularisation method is defined as $\mathcal{R}_\alpha := (A^*A + \alpha L^* L)^{-1} A^*$ for a given parameter choice rule.

By now two questions would naturally arise:

- How accurate is the regularised solution?
- How fast will it converge towards the limit of the sequence of solutions?

and we shall answer them on the upcoming sections, when we study two important topics: *parameter choice* and *regularisation term*.

1.2.1 Parameter Choice

In the literature on regularisation, many different parameter choice methods have been proposed in both stochastic and deterministic settings, we focus on the latter. However, based on the available information, it is not always easy to know how well a particular method will perform in a given situation and how it compares to other methods.

We can classify the parameter choice principle essentially into two groups, those depending (heavily) on the *noisy level* or *noise variance* and those without taking it into account. The latter, called *heuristic*, *noise level free* or *quasi-optimality* criterion, $\alpha := \alpha(g_\delta)$ depends only on the measurement; it has been studied primarily in [91, 89, 3, 50].

The noise free criterion might look a very appealing choice for practical problems, as it does not use any knowledge on the solution and the noise level which cannot be computed from the data. However, the main problem with noise level free parameter choice rules is that it can be proven, that they never yield a convergent regularisation method in the *worst case* for ill-posed problems. This result goes back to [3] and we recall as the follows

Theorem 1.2.2 (Bakushinskii veto)**.** *If the problem* (1.1) *is ill-posed then any noise level free parameter choice rule cannot give rise to convergence in the worst case.*

In our study we assume the noise level to be available. The parameter choice rules in this group can be classified as

- *a priori*, i.e. $\alpha := \alpha(\delta)$ using the noise level and information about the *a priori* smoothness of the solution;
- *a posteriori*, i.e., $\alpha := \alpha(\delta, g_\delta)$ using both noise level and noisy data.

One can derive theoretical results easily choosing an **a priori** parameter. But in the other hand they are not much practical if one needs to obtain convergence rates, because they need some information about the true solution $\overline{u}$. This assumption is called *source condition* and it is generally not known.

Classical results found in [28, thm 5.2] guarantee, if (1.1) is solvable and if the regularisation parameter $\alpha = \alpha(\delta)$ satisfies $\alpha \to 0$ and that $\delta^2/\alpha \to 0$ as $\delta \to 0$, then the regularised solution u_δ^α converges to a solution of (1.1). In general, this convergence can be arbitrarily slow [85].

One can answer the question raised in the previous section, in another words, derive convergence rates by assuming the following source condition:

$$\overline{u} \in \mathscr{R}((A^*A)^\mu) \quad \text{or equivalently,} \quad \overline{u} = (A^*A)^\mu \omega \tag{1.8}$$

where μ is the smoothness parameter and $\|\omega\| \leq \rho$. Therefore, the best possible convergence rate obtainable with this choice is that for $\mu = 1$, where

$$\alpha \sim \left(\frac{\delta}{\rho}\right)^{2/3} \qquad \text{or} \qquad \|u_\delta^\alpha - \overline{u}\| = \mathcal{O}\left(\delta^{2/3}\right).$$

and for the case of data free error the same assumption yields the convergence rate $\|u^\alpha - \overline{u}\| = \mathcal{O}(\alpha)$, as [28, thm 4.11].

For the practical point of view **a posteriori** parameter choice is more desirable. One example is the discrepancy principle of Morozov [69]. Here we are interested in choosing $\alpha = \alpha(\delta, g_\delta)$ which incorporates the data (available with noise) such that

$$\tau_1 \delta \leq \left\|Au_\delta^\alpha - g_\delta\right\| \leq \tau_2 \delta \tag{1.9}$$

for constants $1 < \tau_1 \leq \tau_2$. In other words, we select a regularised solution that, on the one side, the error of the defect is the same order as the noise level δ and, on the other side, α is not too small.

This approach is a variation of the original problem of determining α such that the equality $\|Au_\delta^\alpha - g_\delta\| = \delta$ is satisfied. Moreover, the author [51] commented this equation has a unique solution, provided $\|g - g_\delta\| \leq \delta < \|g_\delta\|$. Furthermore, the function $\alpha \mapsto \|Au_\delta^\alpha - g_\delta\|$ is continuous and strictly increasing, since by [51, thm 2.16]

$$\lim_{\alpha \to \infty} \|Au_\delta^\alpha - g_\delta\| = \|g_\delta\| > \delta$$

and

$$\lim_{\alpha \to 0} \|Au_\delta^\alpha - g_\delta\| = 0 < \delta.$$

We shortly introduced the Morozov discrepancy for linear operators, though it has been extended into non-linear operators in a more general setting. For such generalisation and recent achievement we refer [1, 2] and references therein. We also recommend to the reader [4] for an extensive comparison among several parameter choice methods.

1.2.2 Regularisation Term Choice

The main drawback concerning the quadratic regularisation term on the classical Tikhonov method is to select a solution which is rather (over-) smooth, what is not desirable for image processing. Observing this many researches have drawn their attention towards norms or functionals with desirable properties to guarantee the existence of a solution awhile imposing stability. In this context the quadratic term have been replaced by a general *convex*, lower semi-continuous and coercive functional. This work was popularised under the papers [12, 79, 80].

The two most successful frameworks in the nineties featuring sharp edges are Mumford and Shah [71] and ROF [83]. The first one usually leads to various difficulties in the analysis and numerical realisation due to the explicit treatment of edges and arising non-convexity (cf. [68]). The second one consists in minimising total variation (TV) among all functions within a variance bound; see [13].

As motivation we follows [95] defining of the *total variation* of a function u defined on the interval $[0, 1]$:

$$TV(u) = \sup \sum_i |u(x_i) - u(x_{i-1})| \tag{1.10}$$

where the supremum is taken over all partitions $0 = x_0 < x_1 < \cdots < x_n = 1$ of the interval. If u is piecewise constant with a finite number of jump discontinuities, then $TV(u)$ gives the sum of magnitudes of the jumps. If u is smooth, one can multiply and dived the right-hand side of (1.10) by $\Delta x_i := x_i - x_{i-1}$ and take the limit as the $\Delta x_i \to 0$ to obtain the representation

$$TV(u) = \int_0^1 \left| \frac{du}{dx} \right| dx. \tag{1.11}$$

An obvious generalisation into two space dimension is

$$TV(u) = \int_0^1 \int_0^1 |\nabla u| \, dxdy$$

where ∇u is the standard gradient definition.

$TV(u)$ can be interpreted geometrically as the lateral surface area of the graph of u. If u has many large amplitude oscillations, then it has large lateral

surface area, and hence $TV(u)$ is large. This is a property that TV shares with the more standard Sobolev H^1 "squared norm of the gradient" regularisation functionals. Unlike H^1 functional, with total variation one can effectively reconstruct functions with jump discontinuities (cf. [95]).

An extension of this representation valid even when u is not smooth needs a little more mathematical background in measure theory and we recommend to the reader the book [29, Chapter 5]. We start introducing $BV(\Omega)$ to denote the space of functions of bounded variation: a function $u \in L^1(\Omega)$ has *bounded variation* in Ω if

$$\sup \left\{ \int_\Omega u \operatorname{div} \varphi \;\mid\; \varphi \in C_0^\infty(\Omega)^d, \|\varphi\| \leq 1 \right\} < \infty.$$

So, a more rigorous definition of TV is based on the dual form (i.e., essentially the weakest measure theoretic sense in which a function can be differentiable)

$$TV(u) = \sup_{\varphi \in C_0^\infty(\Omega)^d} \int_\Omega u \operatorname{div} \varphi.$$

Finally, the TV regularisation method is defined

$$\min_u \|Au - g_\delta\|^2 + \alpha TV(u).$$

A new trend in our community is *sparsity*. It refers, usually, to the expansion of a solution u^α with respect to some given orthonormal basis $(\Psi_\gamma)_{\gamma\in\Gamma}$ which only finitely many coefficients are different from zero.

To be more precise, the milestone article [23] has shifted our attention towards the penalisation term with a weighted ℓ_p-norm for the case $1 \leq p < 2$, namely,

$$\|u\|_{\omega,p} = \sum_{\gamma\in\Gamma} \omega_\gamma |\langle u, \Psi_\gamma\rangle|^p.$$

Daubechies *et al* also showed that the minimisation of the Tikhonov-type functional with weighted ℓ_1 penalisation promotes sparsity. According to [77] the functional

$$\min_u \|Au - g_\delta\|^2 + \alpha \sum_{\gamma\in\Gamma} \omega_\gamma |\langle u, \Psi_\gamma\rangle|$$

yields a regularisation method.

A heuristic explanation is that this penalisation term give a higher weight to small coefficients and a lower weight to large coefficients. Moreover, error estimates on [23] were derived in a particular *wavelet* setting [22, 63]. A broader analysis of error estimates and convergence rates under different source conditions have been derived few years later [36, 55].

One common algorithm applied to find the (candidates) minimisers of the Tikhonov-type functional considered is done iteratively. The iterative

thresholding procedure turns out to be defined through the so called *soft-shrinkage operator* [23, 47].

Beyond the far most analysed regularisation with promotes sparsity, i.e., with weighted ℓ_1-norm, is the case $0 \leq p < 1$. For such choice of p, the regularisation is no longer convex and therefore the lack of well-established results and definitions still makes its analysis and converge rates a challenge, see [100, 56] and references therein.

Either approaches introduced above are examples of convex regularisation. In the upcoming Chapter 2 we shall summarise the qualitative and quantitative results for a broader class of regularisation strategies in both linear and non-linear setting, which supplies convergence theorems and convergence rates missing in this section.

1.3 Motivation

From the point of view of Tikhonov and Arsenin's book [91] in many applied problems we have to get along *without a precise knowledge of the causes*, and in others we are really trying to find "causes" that will produce a desired effect, i.e., we are then led to *ill-posed* problems.

Furthermore, in practical problems, we often know only approximately the right-hand side g_0 and the elements of the matrix[3] A_0, that is, the coefficients in the system

$$A_0 f = g_0 \,. \tag{1.12}$$

In such cases, we are dealing not with the system (1.12) but with some other system $A_\epsilon f \approx g_\delta$ such that $\|A_0 - A_\epsilon\| \leq \epsilon$ and $\|g_0 - g_\delta\| \leq \delta$, where the meaning of the norm is usually determined by the nature of the problem; more details shall be given in the Chapter 3. Having the matrix A_ϵ instead of the matrix A_0, we are even less able than before to draw a definite conclusion as to whether the system (1.12) is singular or not, which is a condition for existence and uniqueness of solution.

All we know regarding to the exact system is the approximation (A_ϵ, g_δ) and noise levels (ϵ, δ). But the approximate system may be unsolvable - independently if the original one is solvable (stable and well-posed) or not. The question then arises as to what we are to understand by an *approximate solution* of the underlying equation. It must also be stable under small changes in the pair (A, g).

To the best of our knowledge there are only two books considering the original setting exposed by Tikhonov. Both references [70, 92] have appeared about the same time. The first one took ten years to be translated into English and we can find two sections with results and proofs for linear case. The latter

[3] Tikhonov work takes place in the finite dimensional setup.

has an entire chapter devoted to solve the underlying problem in Hilbert spaces, but unfortunately is only available in Russian and thus is not easily accessible.

Additionally, over the last decades several approaches have been proposed that consider the inversion of an equation (1.12) with both noise in the data and in the operator. Most of the papers published in journals focus on the finite dimensional setup; to list only a few [57, 60].

In the other hand there are two approaches to the solution of infinite-dimensional optimisation problems: *discretise-then-optimise* and *optimise-then-discretise*. Each approach has advantages and disadvantages. We can observe that the first one has been taken as standard in the literature cited above.

The lack of methods which take advantage of the problem in its original formulation has driven our attention to the study, for instance, of integral equations whether the kernel comes from some noisy measurement (like the right hand-side function of (1.12)) or some of the parameters which describe the kernel function are not precisely known, as if we could only guess the belonging interval or roughly expect its values.

A typical example from imaging is a *deconvolution* problem with approximately known or unknown convolution kernel, as, e.g., it was the case for early Hubble images [14, 48, 15, 8]. Another example is connected to inverse scattering, where the *linear sampling method* involves the solution of an integral equation with approximately known kernel, see [16] and references therein.

1.4 Contribution of our Work

In our approach, we would like to restrict our attention to linear operators that can be mainly characterised by a function, as it is, e.g., the case for linear integral operators, where the kernel function determines the behaviour of the operator. Moreover, we will assume that the noise in the operator is due to an incorrect characterising function. This approach will allow us to treat the problem of finding a solution of (1.12) from incorrect data and operator in the framework of Tikhonov regularisation rather than as a constrained minimisation problem.

We introduce a new approach and formulation for the underlying problem with noisy data and incorrectly operator. Moreover, we propose a new method for solving the unconstrained minimisation problem, namely, *double regularised total least squares* (dbl-RTLS). One of the most important result achieved due the novel approach is an extended convergence rate: for both operator and solution.

The main results of this thesis is a collection of the work published in [6, 7] and presented in conferences worldwide.

1.5 Organisation

The rest of the thesis is organised as follows: Chapter 2 contains a summary of the convergence analysis available in the literature for Tikhonov-type methods. This is the most general formulation found and it is described in details. In addition we organize the chapter into two parts: linear and non-linear problems. For each convergence rates are given in two types of source condition. Chapter 3 contains a background on the regularised least squares problems, the main technique which inspired our work. Moreover, we describe shortly some variances of the original R-TLS method and the main results of each on convergence rates.

Our main contribution to the inverse problems community is presented in the last two chapters. In the Chapter 4 we introduce and propose the new *double regularised total least squares* approach used to solve a class of bilinear operators, recovering the characterising function and finding the approximated pair solution. We prove its well-posedness and give convergence rates within this chapter. The second part of the novelty in this thesis is given in the Chapter 5. In this last part we focus on minimising the functional proposed previously (Chapter 4) and introduce an algorithm based on alternating minimisation strategy. Its convergence proof is given and additionally we illustrate the theoretical achievements with numerical experiments in one and two dimensional problems.

Some mathematical preliminaries and background in functional analysis, optimisation and non-smooth analysis can be found in the Appendices A and B.

Chapter 2

Tikhonov Heritage

> "There is nothing more practical than a good theory."
>
> *Kurt Lewin*

In this chapter we present the Tikhonov-type regularisation method and we summarise the main convergence results available in the literature. The adjective "type" refers to the extension of the classical Tikhonov method mainly by setting the penalisation term to be a general convex functional (instead of the usual quadratic norm) while the discrepancy term base on LS is preserved.

This variation allow us not only to reconstruct a solution with special properties, but also to extend theoretical results for both linear and non-linear operators defined between general topological spaces, e.g., Banach spaces. In the other hand we need to be acquainted with more sophisticated concepts and tools brought from non-smooth optimisation and functional analysis. For a review we recommend the reader to survey the Appendices A and B.

On the following we shall display a collection of results from [12, 79, 80, 45, 5], organised in a schematic way.

2.1 Tikhonov-type Methods

We focus on the non-quadratic regularisation methods for solving ill-posed operator equations of the form

$$F(u) = g \ , \tag{2.1}$$

where $F : \mathscr{D}(F) \subset \mathcal{U} \to \mathcal{H}$ is an operator between infinite dimensional Banach spaces. Both linear and non-linear problems are considered.

The Tikhonov-type regularisation consists of minimising

$$J_\alpha^\delta(u) = \frac{1}{2}\|F(u) - g_\delta\|^2 + \alpha\mathcal{R}(u) \ , \tag{2.2}$$

where $\alpha \in \mathbb{R}_+$ is the regularisation parameter and $\mathcal{R}$ is a proper convex functional. Moreover, we assume the noisy data g_δ is available under the deterministic assumption

$$\|g - g_\delta\| \leq \delta \ . \tag{2.3}$$

If the underlying equation has (infinite) many solutions, we select one among all *admissible solutions* which minimises the functional $\mathcal{R}$; we call it the **$\mathcal{R}$-minimising solution**.

The functional J_α^δ presented above represents a generalisation of the classical Tikhonov regularisation [90, 37]. Consequently, the following questions should be considered on the new approach:

- For $\alpha > 0$, does a solution of (2.2) exist? Does the solution depends continuously on the data g_δ?
- Is the method convergent? (i.e., if the data g is exact and $\alpha \to 0$, do the minimisers of (2.2) converge to a solution of (2.1)?)
- Is the method stable in the following sense: if $\alpha = \alpha(\delta)$ is chosen appropriately, do the minimisers of (2.2) converge to a solution of (2.1) as $\delta \to 0$?
- What is the rate of convergence? How should the parameter $\alpha = \alpha(\delta)$ be chosen in order to get optimal convergence rates?

Existence and stability results can be found in the original articles cited above. In this chapter we focus on the last question and we repeat theorems (combined with a short proof) of error estimates and convergence rates.

To accomplish our task we assume throughout this chapter the following assumptions:

Assumption A.

(A1) Given the Banach spaces $\mathcal{U}$ and $\mathcal{H}$ one associates the topologies $\tau_\mathcal{U}$ and $\tau_\mathcal{H}$, respectively, which are weaker than the norm topologies;

(A2) The topological duals of $\mathcal{U}$ and $\mathcal{H}$ are denoted by $\mathcal{U}^*$ and $\mathcal{H}^*$, respectively;

(A3) The norm $\|\cdot\|_\mathcal{U}$ is sequentially lower semi-continuous with respect to $\tau_\mathcal{H}$, i.e., for $u_k \to u$ with respect to the $\tau_\mathcal{U}$ topology, $\mathcal{R}(u) \leq \liminf_k \mathcal{R}(u_k)$;

(A4) $\mathscr{D}(F)$ has non-empty interior with respect to the norm topology and is $\tau_\mathcal{U}$-closed. Moreover[1], $\mathscr{D}(F) \cap dom\ \mathcal{R} \neq \emptyset$;

[1] on the following *dom* $\mathcal{R}$ denotes the *effective domain*, i.e., the set of elements where the functional $\mathcal{R}$ is bounded.

(A5) $F : \mathscr{D}(F) \subseteq \mathcal{U} \to \mathcal{H}$ is continuous from $(\mathcal{U}, \tau_{\mathcal{U}})$ to $(\mathcal{H}, \tau_{\mathcal{H}})$;

(A6) The functional $\mathcal{R} : \mathcal{U} \to [0, +\infty]$ is proper, convex, bounded from below and $\tau_{\mathcal{U}}$ lower semi-continuous;

(A7) For every $M > 0$, $\alpha > 0$, the sets

$$\mathcal{M}_\alpha(M) = \left\{ u \in \mathcal{U} \mid J_\alpha^\delta(u) \leq M \right\}$$

are $\tau_{\mathcal{U}}$ compact, i.e. every sequence (u_k) in $\mathcal{M}_\alpha(M)$ has a subsequence, which is convergent in $\mathcal{U}$ with respect to the $\tau_{\mathcal{U}}$ topology.

Convergence rates and error estimates with respect to the generalised *Bregman distances* were derived originally introduced in [11]; further details can be found on Appendix B.3. Even though this tool does not satisfy neither symmetry nor triangle inequality, it is still the key ingredient whenever we consider convex penalisation.

2.2 Collection of Convergence Rates for Linear Problems

In this section we consider the linear case. Therefore the Equation (2.1) shall be denoted by $Fu = g$, where the operator F is defined from a Banach space into a Hilbert space. The main results of this section were proposed originally in [12, 79].

2.2.1 Rates of Convergence for SC of Type I

First of all we have to decide which "solution" we aim to recover for the underlying problem. Therefore in this section we assume that the noise free data g is *attainable*, i.e., $g \in \mathscr{R}(F)$ and so we define u an *admissible solution* if u satisfies

$$Fu = g. \tag{2.4}$$

In particular, among all admissible solutions, we denote $\overline{u}$ the $\mathcal{R}$-minimising solution of (2.4).

Secondly, error estimates between the regularised solution u_δ^α and $\overline{u}$ can be obtained only under additional smoothness assumption. This assumption, also called source condition, can be stated in the following (slightly) different ways:

1. there exist at least one element ξ in $\partial\mathcal{R}(\overline{u})$ which belongs to the range of the adjoint operator of F;

2. there exists an element $\omega \in \mathcal{H}$ such that

$$F^*\omega =: \xi \in \partial\mathcal{R}(\overline{u}). \tag{2.5}$$

In summary we say the *Source Condition of type I* (SC-I) is satisfied if there is an element $\xi \in \partial\mathcal{R}(\overline{u}) \subseteq \mathcal{U}^*$ in the range of the operator F^*, i.e.,

$$\mathscr{R}(F^*) \cap \partial\mathcal{R}(\overline{u}) \neq \varnothing. \tag{2.6}$$

This assumption enable us to derive the upcoming stability result.

Theorem 2.2.1 ([12, Thm 2])**.** *Let (2.3) hold and let $\overline{u}$ be a $\mathcal{R}$-minimising solution of (2.1) such that the source condition (2.6) and (2.4) are satisfied. Then, for each minimiser u_δ^α of (2.2) the estimate*

$$D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right) \leq \frac{1}{2\alpha}\left(\alpha\|\omega\| + \delta\right)^2 \tag{2.7}$$

holds for $\alpha > 0$. In particular, if $\alpha \sim \delta$, then $D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right) = \mathcal{O}(\delta)$.

Proof. We note that $\left\|F\overline{u} - g_\delta\right\| \leq \delta^2$, by (2.4) and (2.3). Since u_δ^α is a minimiser of the regularised problem (2.2), we have

$$\frac{1}{2}\left\|Fu_\delta^\alpha - g_\delta\right\| + \alpha\mathcal{R}(u_\delta^\alpha) \;\leq\; \frac{\delta^2}{2} + \alpha\mathcal{R}(\overline{u}).$$

Let $D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right)$ the Bregman distance between u_δ^α and $\overline{u}$, so the above inequality becomes

$$\frac{1}{2}\left\|Fu_\delta^\alpha - g_\delta\right\| + \alpha\left(D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right) + \left\langle F^*\omega\ ,\ u_\delta^\alpha - \overline{u}\right\rangle\right) \;\leq\; \frac{\delta^2}{2}.$$

Hence, using (2.3) and Cauchy-Schwarz inequality we can derive the estimate

$$\frac{1}{2}\left\|Fu_\delta^\alpha - g_\delta\right\| + \langle\alpha\omega\ ,\ Fu_\delta^\alpha - g_\delta\rangle_{\mathcal{H}} + \alpha D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right) \;\leq\; \frac{\delta^2}{2} + \alpha\|\omega\|\,\delta.$$

Using the the equality $\|a+b\| = \|a\| + 2\langle a\ ,\ b\rangle + \|b\|$, it is easy to see that

$$\frac{1}{2}\left\|Fu_\delta^\alpha - g_\delta + \alpha\omega\right\| + \alpha D_{\mathcal{R}}^{F^*\omega}\left(u_\delta^\alpha, \overline{u}\right) \;\leq\; \frac{\alpha^2}{2}\|\omega\| + \alpha\delta\|\omega\| + \frac{\delta^2}{2},$$

which yields (2.7) for $\alpha > 0$. □

Theorem 2.2.2 ([12, Thm 1])**.** *If $\overline{u}$ is a $\mathcal{R}$-minimising solution of (2.1) such that the source condition (2.6) and (2.4) are satisfied, then for each minimiser u^α of (2.2) with exact data, the estimate*

$$D_{\mathcal{R}}^{F^*\omega}\left(u^\alpha, \overline{u}\right) \leq \frac{\alpha}{2}\|\omega\|^2$$

holds true.

Proof. The proof is analogous to the proof of Theorem 2.2.1, taking $\delta = 0$. □

2.2.2 Rates of Convergence for SC of Type II

In this section we use another type of source condition, which is stronger than the one assumed in previous subsection. We relax the definition of *admissible solution*, where it is understood in the context of least-squares[2], i.e.,

$$F^*Fu = F^*g\,. \tag{2.8}$$

Note that we do not require $g \in \mathscr{R}(F)$. Moreover, we still denote $\overline{u}$ the $\mathcal{R}$-minimising solution, but instead with respect to (2.8).

Likewise in the previous section, we introduce the *Source Condition of type II* (SC-II)[3] as follows: there exists one element $\xi \in \partial\mathcal{R}\left(\overline{u}\right) \subset \mathcal{U}^*$ in the range of the operator F^*F,

$$\xi \in \mathscr{R}(F^*F) \cap \partial\mathcal{R}\left(\overline{u}\right) \neq \varnothing\,. \tag{2.9}$$

This condition is equivalent to the existence of $\omega \in \mathcal{U}\backslash\{0\}$ such that $\xi = F^*F\omega$, where F^* is the adjoint operator of F and $F^*F : \mathcal{U} \to \mathcal{U}^*$.

Theorem 2.2.3 ([79, Thm 2.2])**.** *Let (2.3) hold and let $\overline{u}$ be a $\mathcal{R}$-minimising solution of (2.1) such that the source condition (2.9) as well as (2.8) are satisfied. Then the following inequalities hold for any $\alpha > 0$:*

$$D_{\mathcal{R}}^{F^*F\omega}\big(u_\delta^\alpha, \overline{u}\big) \leq D_{\mathcal{R}}^{F^*F\omega}\big(\overline{u} - \alpha\omega, \overline{u}\big) + \frac{\delta^2}{\alpha} + \frac{\delta}{\alpha}\sqrt{\delta^2 + 2\alpha D_{\mathcal{R}}^{F^*F\omega}\big(\overline{u} - \alpha\omega, \overline{u}\big)}, \tag{2.10}$$

$$\left\|Fu_\delta^\alpha - F\overline{u}\right\| \leq \alpha\left\|F\omega\right\| + \delta + \sqrt{\delta^2 + 2\alpha D_{\mathcal{R}}^{F^*F\omega}\big(\overline{u} - \alpha\omega, \overline{u}\big)}\,. \tag{2.11}$$

Proof. Since u_δ^α is a minimiser of (2.2), it follows from algebraic manipulation and from the definition of Bregman distance that

$$\begin{aligned} 0 \;&\geq\; \frac{1}{2}\left[\left\|Fu_\delta^\alpha - g_\delta\right\| - \left\|Fu - g_\delta\right\|\right] + \alpha\mathcal{R}(u_\delta^\alpha) - \alpha\mathcal{R}(u) \\ &=\; \frac{1}{2}\left[\left\|Fu_\delta^\alpha\right\| - \left\|Fu\right\|\right] - \left\langle F\left(u_\delta^\alpha - u\right)\,,\, g_\delta\right\rangle_{\mathcal{H}} - \alpha D_{\mathcal{R}}^{F^*F\omega}\big(u, \overline{u}\big) \\ &\quad + \alpha\left\langle F\omega\,,\, F\left(u_\delta^\alpha - u\right)\right\rangle_{\mathcal{H}} + \alpha D_{\mathcal{R}}^{F^*F\omega}\big(u_\delta^\alpha, \overline{u}\big)\,. \end{aligned} \tag{2.12}$$

Notice that

$$\begin{aligned} \left\|Fu_\delta^\alpha\right\| - \left\|Fu\right\| \;&=\; \left\|F\left(u_\delta^\alpha - \overline{u} + \alpha\omega\right)\right\| - \left\|F\left(u - \overline{u} + \alpha\omega\right)\right\| \\ &\quad + 2\left\langle Fu_\delta^\alpha - Fu\,,\, F\overline{u} - \alpha F\omega\right\rangle_{\mathcal{H}}\,. \end{aligned}$$

[2] in the literature this definition of generalised solution is also known as *best-approximate solution*.

[3] also called *source condition of second kind*.

Moreover, by (2.8), we have $\langle F\left(u_\delta^\alpha - u\right)\ ,\ g_\delta - F\overline{u}\rangle_{\mathcal{H}} = \langle F\left(u_\delta^\alpha - u\right)\ ,\ g_\delta - g\rangle_{\mathcal{H}}$. Therefore, it follows from (2.12) that

$$\begin{aligned}&\frac{1}{2}\left\|F\left(u_\delta^\alpha - \overline{u} + \alpha\omega\right)\right\| + \alpha D_{\mathcal{R}}^{F^*F\omega}\left(u_\delta^\alpha, \overline{u}\right)\\ \leq\ &\langle F\left(u_\delta^\alpha - u\right)\ ,\ g_\delta - g\rangle_{\mathcal{H}} + \alpha D_{\mathcal{R}}^{F^*F\omega}\left(u, \overline{u}\right) + \frac{1}{2}\left\|F\left(u - \overline{u} + \alpha\omega\right)\right\|\end{aligned}$$

for every $u \in \mathcal{U}$, $\alpha \geq 0$ and $\delta \geq 0$.

Replacing u by $\overline{u} - \alpha\omega$ in the last inequality, using (2.3), relations $\left\langle a\ ,\ b\right\rangle \leq \left|\left\langle a\ ,\ b\right\rangle\right| \leq \|a\|\,\|b\|$, and defining $\gamma = \left\|F\left(u_\delta^\alpha - \overline{u} + \alpha\omega\right)\right\|$ we obtain

$$\frac{1}{2}\gamma^2 + \alpha D_{\mathcal{R}}^{F^*F\omega}\left(u_\delta^\alpha, \overline{u}\right)\ \leq\ \delta\gamma + \alpha D_{\mathcal{R}}^{F^*F\omega}\left(\overline{u} - \alpha\omega, \overline{u}\right).$$

We estimate separately each term on the left-hand side by right-hand side. One of the estimates is an inequality in the form of a polynomial of the second degree for γ, which gives us the inequality

$$\gamma \leq \delta + \sqrt{\delta^2 + 2\alpha D_{\mathcal{R}}^{F^*F\omega}\left(\overline{u} - \alpha\omega, \overline{u}\right)}\,.$$

This inequality together with the other estimate, gives us (2.10). Now, (2.11) follows from the fact that $\left\|F\left(u_\delta^\alpha - \overline{u}\right)\right\| \leq \gamma + \alpha\,\|F\omega\|$. □

Theorem 2.2.4 ([79, Thm 2.1])**.** *Let $\alpha \geq 0$ be given. If $\overline{u}$ is a $\mathcal{R}$-minimising solution of (2.1) satisfying the source condition (2.9) as well as (2.8), then the following inequalities hold true:*

$$D_{\mathcal{R}}^{F^*F\omega}\left(u^\alpha, \overline{u}\right) \leq D_{\mathcal{R}}^{F^*F\omega}\left(\overline{u} - \alpha\omega, \overline{u}\right),$$

$$\left\|Fu^\alpha - F\overline{u}\right\| \leq \alpha\,\|F\omega\| + \sqrt{2\alpha D_{\mathcal{R}}^{F^*F\omega}\left(\overline{u} - \alpha\omega, \overline{u}\right)}\,.$$

Proof. The proof is analogous to the proof of Theorem 2.2.3, taking $\delta = 0$. Notice that here α can be taken equal to zero. □

Corollary 2.2.5 ([79])**.** *Let the assumptions of the Theorem 2.2.3 hold true. Further, assume that $\mathcal{R}$ is twice differentiable in a neighbourhood U of $\overline{u}$ and there exists a number $M > 0$ such that for any $v \in \mathcal{U}$ and $u \in U$ the inequality*

$$\left\langle \mathcal{R}''(u)v\ ,\ v\right\rangle \leq M\,\|v\|^2 \tag{2.13}$$

hold true. Then, for the parameter choice $\alpha \sim \delta^{\frac{2}{3}}$ we have $D_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right) = \mathcal{O}\left(\delta^{\frac{4}{3}}\right)$. Moreover, for exact data we have $D_{\mathcal{R}}^{\xi}\left(u^\alpha, \overline{u}\right) = \mathcal{O}\left(\alpha^2\right)$.

Proof. Using Taylor's expansion at the element $\overline{u}$ we obtain

$$\mathcal{R}(u) \;=\; \mathcal{R}(\overline{u}) + \left\langle \mathcal{R}'(\overline{u}) \;,\, u-\overline{u} \right\rangle + \frac{1}{2}\left\langle \mathcal{R}''(\mu)(u-\overline{u}) \;,\, u-\overline{u} \right\rangle$$

for some $\mu \in [u, \overline{u}]$. Let $u = \overline{u} - \alpha\omega$ in the above equality. For sufficiently small α, it follows from assumption (2.13) and the definition of the Bregman distance, with $\xi = \mathcal{R}'(\overline{u})$, that

$$\begin{aligned} D_{\mathcal{R}}^{\xi}\left(\overline{u}-\alpha\omega, \overline{u}\right) &= \frac{1}{2}\left\langle \mathcal{R}''(\mu)(-\alpha\omega) \;,\; -\alpha\omega \right\rangle \\ &\leq \alpha^2 \frac{M}{2} \left\|\omega\right\|_{\mathcal{U}}^2 . \end{aligned}$$

Note that $D_{\mathcal{R}}^{\xi}\left(\overline{u}-\alpha\omega, \overline{u}\right) = \mathcal{O}\left(\alpha^2\right)$, so the desired rates of convergence follow from Theorems 2.2.3 and 2.2.4. □

2.3 Collection of Convergence Rates for Non-linear Problems

This section displays a collection the convergence analysis for the non-linear problems. In contrast with other classical conditions, the following analysis covers the case when both $\mathcal{U}$ and $\mathcal{H}$ are Banach spaces.

Back to [27] we learn through two examples of non-linear problems the interesting effect: ill-posedness of a non-linear problem need not imply ill-posedness of its linearisation. Also that the converse implication need not be true. A well-posed non-linear problem may have ill-posed linearisation. Hence we need additional assumptions concerning both operator and its linearisation.

This assumption is known as *non-linearity condition* and it is based on first-order Taylor expansion of the operator F around $\overline{u}$. The non-linearity condition assumed in this section is given originally in [80] and stated as follows.

Assumption B. Assume that a $\mathcal{R}$-minimising solution $\overline{u}$ of (2.1) exists and that the operator $F : \mathscr{D}(F) \subseteq \mathcal{U} \to \mathcal{H}$ is Gâteaux differentiable. Moreover, we assume that there exists $\rho > 0$ such that, for every $u \in \mathscr{D}(F) \cap \mathcal{B}_{\rho}\left(\overline{u}\right)$

$$\left\|F\left(u\right) - F\left(\overline{u}\right) - F'\left(\overline{u}\right)\left(u-\overline{u}\right)\right\| \leq c D_{\mathcal{R}}^{\xi}\left(u, \overline{u}\right) \;,\; c > 0 \tag{2.14}$$

and $\xi \in \partial\mathcal{R}\left(\overline{u}\right)$.

2.3.1 Rates of Convergence for SC of Type I

In comparison with the source condition (2.6) introduced on previous section, the extension of the *Source Condition of type I* to non-linear problems are done with respect to the linearisation of the operator and its adjoint. Namely, we assume that

$$\xi \in \mathscr{R}(F'(\overline{u})^*) \cap \partial\mathcal{R}(\overline{u}) \neq \varnothing \tag{2.15}$$

where $\overline{u}$ is a $\mathcal{R}$-minimising solution of (2.1).

Note that the derivative of operator F is defined between the Banach space $\mathcal{U}$ and $\mathscr{L}(\mathcal{U}, \mathcal{H})$, the space of the linear transformations from $\mathcal{U}$ into $\mathcal{H}$. When we apply the derivative at $\overline{u} \in \mathcal{U}$ we have a linear operator $F'(\overline{u}) : \mathcal{U} \to \mathcal{H}$ and so we define its adjoint $F'(\overline{u})^* : \mathcal{H}^* \to \mathcal{U}^*$.

The source condition (2.15) is stated equivalently as follows: there exists an element $\omega \in \mathcal{H}^*$ such that

$$\xi = F'(\overline{u})^* \omega \in \partial\mathcal{R}(\overline{u}) \ . \tag{2.16}$$

Theorem 2.3.1 ([80, Thm 3.2])**.** *Let the Assumptions A, B and relation (2.3) hold true. Moreover, assume that there exists $\omega \in \mathcal{H}^*$ such that (2.16) is satisfied and $c\|\omega\|_{\mathcal{H}^*} < 1$. Then, the following estimates hold:*

$$\|F(u_\delta^\alpha) - F(\overline{u})\| \leq 2\alpha\|\omega\|_{\mathcal{H}^*} + 2\left(\alpha^2\|\omega\|_{\mathcal{H}^*}^2 + \delta^2\right)^{\frac{1}{2}} ,$$

$$D_{\mathcal{R}}^{F'(\overline{u})^*\omega}(u_\delta^\alpha, \overline{u}) \leq \frac{2}{1 - c\|\omega\|_{\mathcal{H}^*}}\left[\frac{\delta^2}{2\alpha} + \alpha\|\omega\|_{\mathcal{H}^*}^2 + \|\omega\|_{\mathcal{H}^*}\left(\alpha^2\|\omega\|_{\mathcal{H}^*}^2 + \delta^2\right)^{\frac{1}{2}}\right] .$$

In particular, if $\alpha \sim \delta$, then $\|F(u_\delta^\alpha) - F(\overline{u})\| = \mathcal{O}(\delta)$ and $D_{\mathcal{R}}^{F'(\overline{u})^*\omega}(u_\delta^\alpha, \overline{u}) = \mathcal{O}(\delta)$.

Proof. Since u_δ^α is the minimiser of (2.2), it follows from the definition of the Bregman distance that

$$\frac{1}{2}\left\|F(u_\delta^\alpha) - g_\delta\right\| \ \leq \ \frac{1}{2}\delta^2 - \alpha\left(D_{\mathcal{R}}^{F'(\overline{u})^*\omega}(u_\delta^\alpha, \overline{u}) + \left\langle F'(\overline{u})^*\omega \ , \ u_\delta^\alpha - \overline{u}\right\rangle\right) .$$

By using (2.3) and (2.1) we obtain

$$\frac{1}{2}\left\|F(u_\delta^\alpha) - F(\overline{u})\right\| \ \leq \ \left\|F(u_\delta^\alpha) - g_\delta\right\| + \delta^2 .$$

Now, using the last two inequalities above, the definition of Bregman distance, the non-linearity condition and the assumption $(c\|\omega\|_{\mathcal{H}^*} - 1) < 0$, we

obtain

$$\begin{aligned}
\frac{1}{4}\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)\right\| &\leq \frac{1}{2}\left(\left\|F\left(u_\delta^\alpha\right)-g_\delta\right\|+\delta^2\right)\\
&\leq \delta^2-\alpha D_{\mathcal{R}}^{F'(\overline{u})^*\omega}\left(u_\delta^\alpha,\overline{u}\right)+\alpha\left\langle\omega\ ,\ -F'\left(\overline{u}\right)\left(u_\delta^\alpha-\overline{u}\right)\right\rangle\\
&\leq \delta^2-\alpha D_{\mathcal{R}}^{F'(\overline{u})^*\omega}\left(u_\delta^\alpha,\overline{u}\right)+\alpha\left\|\omega\right\|_{\mathcal{H}^*}\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)\right\|\\
&\quad+\alpha\left\|\omega\right\|_{\mathcal{H}^*}\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)-F'\left(\overline{u}\right)\left(u_\delta^\alpha-\overline{u}\right)\right\|\\
&= \delta^2+\alpha\left(c\left\|\omega\right\|_{\mathcal{H}^*}-1\right)D_{\mathcal{R}}^{F'(\overline{u})^*\omega}\left(u_\delta^\alpha,\overline{u}\right)\\
&\quad+\alpha\left\|\omega\right\|_{\mathcal{H}^*}\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)\right\| \qquad (2.17)\\
&\leq \delta^2+\alpha\left\|\omega\right\|_{\mathcal{H}^*}\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)\right\| \qquad (2.18)
\end{aligned}$$

From (2.18) we obtain an inequality in the form of a polynomial of second degree for the variable $\gamma=\left\|F\left(u_\delta^\alpha\right)-F\left(\overline{u}\right)\right\|$. This gives us the first estimate stated by the theorem. For the second estimate we use (2.17) and the previous estimate for γ. □

Theorem 2.3.2. *Let the Assumptions A and B hold true. Moreover, assume the existence of $\omega\in\mathcal{H}^*$ such that (2.16) is satisfied and $c\left\|\omega\right\|_{\mathcal{H}^*}<1$. Then, the following estimates hold:*

$$\left\|F\left(u^\alpha\right)-F\left(\overline{u}\right)\right\|\leq 4\alpha\left\|\omega\right\|_{\mathcal{H}^*}\ ,$$

$$D_{\mathcal{R}}^{F'(\overline{u})^*\omega}\left(u^\alpha,\overline{u}\right)\leq\frac{4\alpha\left\|\omega\right\|_{\mathcal{H}^*}^2}{1-c\left\|\omega\right\|_{\mathcal{H}^*}}\,.$$

Proof. The proof is analogous to the proof of Theorem 2.3.1, taking $\delta=0$. □

2.3.2 Rates of Convergence for SC of Type II

Similarly as in the previous subsection, the extension of the *Source Condition of type II* (2.9) to non-linear problems is given as:

$$\xi\in\mathscr{R}(F'\left(\overline{u}\right)^*F'\left(\overline{u}\right))\cap\partial\mathcal{R}\left(\overline{u}\right)\neq\varnothing$$

where $\overline{u}$ is a $\mathcal{R}$-minimising solution of (2.1).

The assumption above has the following equivalent formulation: there exists an element $\omega\in\mathcal{U}$ such that

$$\xi=F'\left(\overline{u}\right)^*F'\left(\overline{u}\right)\omega\in\partial\mathcal{R}\left(\overline{u}\right)\ . \qquad (2.19)$$

Theorem 2.3.3 ([80, Thm 3.4])**.** *Let the Assumptions A, B hold as well as estimate (2.3). Moreover, let $\mathcal{H}$ be a Hilbert space and assume the existence of a $\mathcal{R}$-minimising solution $\overline{u}$ of (2.1) in the interior of $\mathscr{D}(F)$. Assume also the*

existence of $\omega \in \mathcal{U}$ *such that* (2.19) *is satisfied and* $c\,\|F'(\overline{u})\,\omega\| < 1$. *Then, for* α *sufficiently small the following estimates hold:*

$$\|F(u_\delta^\alpha) - F(\overline{u})\| \leq \alpha\,\|F'(\overline{u})\,\omega\| + h(\alpha,\delta)\,,$$

$$D_{\mathcal{R}}^{\xi}\big(u_\delta^\alpha, \overline{u}\big) \leq \frac{\alpha s + (cs)^2/2 + \delta h(\alpha,\delta) + cs\,(\delta + \alpha\,\|F'(\overline{u})\,\omega\|)}{\alpha\,(1 - c\,\|F'(\overline{u})\,\omega\|)}\,, \tag{2.20}$$

where $h(\alpha,\delta) := \delta + \sqrt{(\delta + cs)^2 + 2\alpha s\,(1 + c\,\|F'(\overline{u})\,\omega\|)}$ *and* $s = D_{\mathcal{R}}^{\xi}\big(\overline{u} - \alpha\omega, \overline{u}\big)$.

Proof. Since u_δ^α is the minimiser of (2.2), it follows that

$$\begin{aligned} 0 &\geq \frac{1}{2}\big\|F(u_\delta^\alpha) - g_\delta\big\| - \frac{1}{2}\big\|F(u) - g_\delta\big\| + \alpha\,(\mathcal{R}(u_\delta^\alpha) - \mathcal{R}(u)) \\ &= \frac{1}{2}\big\|F(u_\delta^\alpha)\big\| - \frac{1}{2}\big\|F(u)\big\| + \langle F(u) - F(u_\delta^\alpha)\ ,\ g_\delta\rangle_{\mathcal{H}} \\ &\quad + \alpha\,(\mathcal{R}(u_\delta^\alpha) - \mathcal{R}(u)) \\ &= \varrho(u_\delta^\alpha) - \varrho(u)\,. \end{aligned} \tag{2.21}$$

where $\varrho(u) = \frac{1}{2}\big\|F(u) - q\big\| + \alpha D_{\mathcal{R}}^{\xi}\big(u, \overline{u}\big) - \langle F(u)\ ,\ g_\delta - q\rangle_{\mathcal{H}} + \alpha\big\langle \xi\ ,\ u\big\rangle$, $q = F(\overline{u}) - \alpha F'(\overline{u})\,\omega$ and ξ is given by source condition (2.19).

From (2.21) we have $\varrho(u_\delta^\alpha) \leq \varrho(u)$. By the definition of $\varrho(\cdot)$, taking $u = \overline{u} - \alpha\omega$ and setting $v = F(u_\delta^\alpha) - F(\overline{u}) + \alpha F'(\overline{u})\,\omega$ we obtain

$$\frac{1}{2}\,\|v\| + \alpha D_{\mathcal{R}}^{\xi}\big(u_\delta^\alpha, \overline{u}\big) \quad \leq \quad \alpha s + T_1 + T_2 + T_3\,, \tag{2.22}$$

where s is given in the theorem, and

$$T_1 = \frac{1}{2}\big\|F(\overline{u} - \alpha\omega) - F(\overline{u}) + \alpha F'(\overline{u})\,\omega\big\|\,,$$

$$T_2 = |\langle F(u_\delta^\alpha) - F(\overline{u} - \alpha\omega)\ ,\ g_\delta - g\rangle_{\mathcal{H}}|\,,$$

$$T_3 = \alpha\,\langle F'(\overline{u})\,\omega\ ,\ F(u_\delta^\alpha) - F(\overline{u} - \alpha\omega) - F'(\overline{u})\,(u_\delta^\alpha - (\overline{u} - \alpha\omega))\rangle_{\mathcal{H}}\,.$$

The next step is to estimate each one of the constants T_j above, $j = 1, 2$ and 3. We use the non-linear condition (2.14), Cauchy-Schwarz, and some algebraic manipulation to obtain $T_1 \leq \frac{c^2 s^2}{2}$,

$$\begin{aligned} T_2 &\leq |\langle v\ ,\ g_\delta - g\rangle_{\mathcal{H}}| + |\langle F(\overline{u} - \alpha\omega) - F(\overline{u}) + \alpha F'(\overline{u})\,\omega - \ ,\ g_\delta - g\rangle_{\mathcal{H}}| \\ &\leq \|v\|\,\|g_\delta - g\| + cD_{\mathcal{R}}^{\xi}\big(\overline{u} - \alpha\omega, \overline{u}\big)\,\|g_\delta - g\| \\ &\leq \delta\,\|v\| + \delta cs\,, \end{aligned}$$

and

$$\begin{aligned} T_3 &= \alpha \left\langle F'\left(\overline{u}\right)\omega \ , \ F\left(u_\delta^\alpha\right) - F\left(\overline{u}\right) - F'\left(\overline{u}\right)\left(u_\delta^\alpha - \overline{u}\right)\right\rangle_{\mathcal{H}} \\ &\quad + \alpha \left\langle F'\left(\overline{u}\right)\omega \ , \ -\left(F\left(\overline{u} - \alpha\omega\right) - F\left(\overline{u}\right) + \alpha F'\left(\overline{u}\right)\omega\right)\right\rangle_{\mathcal{H}} \\ &\leq \alpha \left\|F'\left(\overline{u}\right)\omega\right\| \left\|F\left(u_\delta^\alpha\right) - F\left(\overline{u}\right) - F'\left(\overline{u}\right)\left(u_\delta^\alpha - \overline{u}\right)\right\| \\ &\quad + \alpha \left\|F'\left(\overline{u}\right)\omega\right\| \left\|F\left(\overline{u} - \alpha\omega\right) - F\left(\overline{u}\right) + \alpha F'\left(\overline{u}\right)\omega\right\| \\ &\leq \alpha \left\|F'\left(\overline{u}\right)\omega\right\| cD_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right) + \alpha \left\|F'\left(\overline{u}\right)\omega\right\| cD_{\mathcal{R}}^{\xi}\left(\overline{u} - \alpha\omega, \overline{u}\right) \\ &= \alpha c \left\|F'\left(\overline{u}\right)\omega\right\| D_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right) + \alpha cs \left\|F'\left(\overline{u}\right)\omega\right\| \ . \end{aligned}$$

Using these estimates in (2.22), we obtain

$$\begin{aligned} \|v\| + 2\alpha D_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right)\left[1 - c\left\|F'\left(\overline{u}\right)\omega\right\|\right] &\leq 2\delta\|v\| + 2\alpha s + (cs)^2 \\ &\quad + 2\delta cs + 2\alpha cs\left\|F'\left(\overline{u}\right)\omega\right\| \ . \end{aligned}$$

Analogously as in the proof of Theorem 2.2.3, each term on the left-hand side of the last inequality is estimated separately by the right-hand side. This allows the derivation of an inequality described by a polynomial of second degree. From this inequality, the theorem follows. □

Theorem 2.3.4. *Let Assumptions A, B hold and assume $\mathcal{H}$ to be a Hilbert space. Moreover, assume the existence of a $\mathcal{R}$-minimising solution $\overline{u}$ of (2.1) in the interior of $\mathscr{D}(F)$, also the existence of $\omega \in \mathcal{U}$ such that (2.19) is satisfied and $c\left\|F'\left(\overline{u}\right)\omega\right\| < 1$. Then, for α sufficiently small the following estimates hold:*

$$\left\|F\left(u^\alpha\right) - F\left(\overline{u}\right)\right\| \leq \alpha\left\|F'\left(\overline{u}\right)\omega\right\| + \sqrt{(cs)^2 + 2\alpha s\left(1 + c\left\|F'\left(\overline{u}\right)\omega\right\|\right)} \, ,$$

$$D_{\mathcal{R}}^{\xi}\left(u^\alpha, \overline{u}\right) \leq \frac{\alpha s + (cs)^2/2 + \alpha cs\left\|F'\left(\overline{u}\right)\omega\right\|_{\mathcal{H}}}{\alpha\left(1 - c\left\|F'\left(\overline{u}\right)\omega\right\|_{\mathcal{H}}\right)} \, , \tag{2.23}$$

where $s = D_{\mathcal{R}}^{\xi}\left(\overline{u} - \alpha\omega, \overline{u}\right)$.

Proof. The proof is analogous to the proof of Theorem 2.3.3, taking $\delta = 0$. □

Corollary 2.3.5 ([80, Prop 3.5])**.** *Let assumptions of the Theorem 2.3.3 hold true. Moreover, assume that $\mathcal{R}$ is twice differentiable in a neighbourhood U of $\overline{u}$, and that there exists a number $M > 0$ such that for all $u \in U$ and for all $v \in \mathcal{U}$, the inequality $\left\langle \mathcal{R}''(u)v \ , \ v\right\rangle \leq M\|v\|$ holds. Then, for the choice of parameter $\alpha \sim \delta^{\frac{2}{3}}$ we have $D_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right) = \mathcal{O}\left(\delta^{\frac{4}{3}}\right)$, while for exact data we obtain $D_{\mathcal{R}}^{\xi}\left(u_\delta^\alpha, \overline{u}\right) = \mathcal{O}\left(\alpha^2\right)$.*

Proof. The proof is similar to the proof of Corollary 2.2.5 and is based on Theorems 2.3.3 and 2.3.4. □

2.4 Advances in Convergence Rates

We briefly comment on two new trends for deriving convergences rates, namely, *variational inequalities* and *approximated source condition*.

Since the first convergence rates results for non-linear problems given in [27] until the results [12, 79, 80] presented previously, the results of Engl and co-workers seems to be fully generalised. Nevertheless another paper concerning convergence rates came out [45] bringing new insights. The authors observed the following:

> In all these papers relatively strong regularity assumptions are made. However, it has been observed numerically that violations of the smoothness assumptions of the operator do not necessarily affect the convergence rate negatively. We take this observation and weaken the smoothness assumptions on the operator and prove a novel convergence rate result. The most significant difference in this result from the previous ones is that the source condition is formulated as a *variational inequality* and not as an equation as previously.

We display the **variational inequality** (VI) proposed in [45, Assumption 4.1], regardless auxiliary assumptions found in the paper.

Assumption C. There exist numbers $c_1, c_2 \in [0, \infty)$, where $c_1 < 1$, and $\xi \in \partial\mathcal{R}(\overline{u})$ such that

$$\langle \xi \ , \ u - \overline{u} \rangle \leq c_1 D_{\mathcal{R}}^{\xi}(u, \overline{u}) + c_2 \left\| F(u) - F(\overline{u}) \right\|$$

for all $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$ where $\rho > \alpha_{\max}\left(\mathcal{R}(\overline{u}) + \frac{\delta^2}{\alpha}\right)$.

Additionally, it was proved that standard non-linearity conditions imply the new VI. Under this assumption one can derive the same rate of convergence obtained in Section 2.3. For more details see [45, 30, 46].

In [44] an alternative concept for proving convergence rates for linear problems in Hilbert spaces is presented, when the source condition

$$\overline{u} = F^*\omega, \quad \omega \in \mathcal{H}^* \tag{2.24}$$

is injured.

Instead we have an **approximated source condition** like $\overline{u} = F^*\omega + r$, where $r \in \mathcal{U}$. The theory is based on the decay rate of so-called *distance functions* which measures the degree of violation of the solution with respect to a prescribed benchmark source condition, e.g. (2.24). For the linear case the distance function is defined intuitively as

$$d(\rho) = \inf \left\{ \|\overline{u} - F^*\omega\| \ \mid \ \omega \in \mathcal{H}^*, \|\omega\| \leq \rho \right\}.$$

The article [41] points out that this approach can be generalised to Banach spaces, as well as to non-linear operators. Afterwards, with the aid of this distance functions, the authors of [42] presented error bounds and convergence rates for regularised solutions of non-linear problems for Tikhonov-type functionals when the reference source condition is not satisfied.

Chapter 3

Least Squares Revolution

> "Each problem that I solved became a rule, which served afterwards to solve other problems."
>
> *René Descartes*

In the classical Least Squares approach the system matrix is assumed to be free from error and all the errors are confined to the observation vector. However in many applications this assumption is often unrealistic. Therefore a new technique was proposed: *Total Least Squares*, or shortly, TLS[1]. This concept has been independently develop in various literatures, namely, *error-in-variables*, *rigorous least squares*, or (in a special case) *orthogonal regression*, listing only few in statistical literature. It also leads to a procedure investigated in this chapter named *regularised total least squares*.

In this chapter we shall introduce the TLS fitting technique and the *regularised TLS*. Additionally we compare them, respectively, with the least squares and standard Tikhonov regularisation. We conclude this chapter with a general overview on competitive approaches to related problems.

3.1 Total Least Squares

Gene Howard Golub (1932–2007) was an American mathematician with remarkable work in the field of numerical linear algebra; listing only a few topics: least-squares problems, singular value decomposition, domain decomposition, differentiation of pseudo-inverses, inverse eigenvalue problem, conjugate gradient method, Gauss quadrature.

In 1980 Golub and Van Loan [34] investigated a fitting technique based on the least squares problem for solving a matrix equation with incorrect matrix

[1] We hope to not mislead to *truncate least squares* notation.

Figure 3.1: Gene H. Golub

and data vector, named *total least squares* (TLS) method. On the following we presented the TLS method and we compare it briefly with another classical approach; more details can be found in [93, 65] and references therein.

Let A_0 be a matrix in $\mathbb{R}^{m\times n}$ and y_0 a vector in $\mathbb{R}^{m\times 1}$, obtained after the discretisation of the linear operator equation $Fu = g$, where $F : \mathcal{U} \to \mathcal{H}$ is a mapping between two Hilbert spaces. We then consider[2] solving the equation

$$A_0 x = y_0 \tag{3.1}$$

where both A_0 and y_0 information out of inaccurate measurement or inherent some roundoff errors. More precisely, it is available only the following pair

$$\|y_0 - y_\delta\|_2 \leq \delta \tag{3.2}$$

and

$$\|A_0 - A_\epsilon\|_F \leq \epsilon. \tag{3.3}$$

In particular the classical *least squares* (LS) approach, proposed by Carl Friedrich Gauss (1777-1855), the measurements A_0 are assumed to be free of error; hence, all the errors are confined to the observation vector y_δ. The LS solution is given by solving the following minimisation problem

$$\begin{array}{ll} \text{minimise}_y & \|y - y_\delta\|_2 \\ \text{subject to} & y \in \mathscr{R}(A_0) \end{array}$$

or equivalently

$$\underset{x}{\text{minimise}} \|A_0 x - y_\delta\|_2. \tag{3.4}$$

Solutions of the ordinary LS problem are characterised by the following theorem.

Theorem 3.1.1 ([93, Cor 2.1])**.** *If* $\text{rank}(A_0) = n$ *then* (3.4) *has a unique LS solution, given by*

$$x^{LS} = (A_0^T A_0)^{-1} A_0^T y_\delta. \tag{3.5}$$

the corresponding LS correction is given by the residual

$$r = y_\delta - A_0 x^{LS} = y_\delta - y^{LS}, \qquad y^{LS} = P_{A_0} y_\delta$$

where $P_{A_0} = A_0(A_0^T A_0)^{-1} A_0^T$ *is the orthogonal projector onto* $\mathscr{R}(A_0)$.

[2] We introduce the problem only for the one dimensional case $Ax = y$, i.e., when x and y are vectors. In the book [93, Chap 3] it is also considered the multidimensional case $AX = Y$, where all elements are matrices.

This approach is frequently unrealistic: sampling errors, human errors and instrument errors may imply inaccuracies of the data matrix A_0 as well (e.g., due discretisation, approximation of differential or integral models).

Therefore the need of an approach which amounts to fitting a "best" subspace to the measurement data (A_ϵ, y_δ) leads to the TLS approach. In comparison to LS method the new minimisation problem is with respect to the pair (A, y). The element paring $\tilde{A}x^{TLS} = \tilde{y}$ is then called the *total least squares solution*, where $\tilde{A}$ and $\tilde{y}$ are the arguments which minimises the following constrained problem[3]

$$\begin{array}{c} \text{minimise}_{(A,y)} \\ \text{subject to} \end{array} \quad \begin{array}{c} \left\| [A, y] - [A_\epsilon, y_\delta] \right\|_F \\ y \in \mathscr{R}(A) \end{array} . \tag{3.6}$$

The basic principle of TLS is that the noisy data $[A_\epsilon, y_\delta]$, while not satisfying a linear relation, are modified with minimal effort, as measured by the Frobenius norm, in a "nearby" matrix $[\tilde{A}, \tilde{y}]$ that is rank-deficient so that the set $\tilde{A}x = \tilde{y}$ is compatible. This matrix $[\tilde{A}, \tilde{y}]$ is a rank one modification of the data matrix $[A_\epsilon, y_\delta]$.

The foundation is the *singular value decomposition* (SVD), an important role in a number of matrix approximation problems [35]; see its definition in the upcoming theorem.

Theorem 3.1.2 ([35, Thm 2.5.1])**.** *If $A \in \mathbb{R}^{m\times n}$ then there exist orthonormal matrices $U = [u_1, \ldots, u_m] \in \mathbb{R}^{m\times m}$ and $V = [v_1, \ldots, v_n] \in \mathbb{R}^{n\times n}$ such that*

$$U^T A V = \Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_p), \quad \sigma_1 \geq \cdots \geq \sigma_p \geq 0$$

where $p = \min\{m, n\}$.

The triplet (u_i, σ_i, v_i) reveals a great deal about the structure of A. For instance, defining r as the number of nonzeros singular values, i.e., $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots \sigma_p = 0$ it is known that

$$\operatorname{rank}(A) = r \quad \text{and} \quad A = U_r \Sigma_r {V_r}^T = \sum_{i=1}^{r} \sigma_i u_i {v_i}^T \tag{3.7}$$

where U_r (equivalently Σ_r and V_r) denotes the first r columns of the matrix U (equivalently Σ and V). The Equation (3.7) displays the decomposition of the matrix A of rank r in a sum of r matrices of rank one.

Through SVD we can define the Frobenious norm of a matrix A as

$$\|A\|_F^2 := \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 = \sigma_1^2 + \cdots + \sigma_p^2, \quad p = \min\{m, n\}$$

[3] we use the same Matlab's notation to add the vector y as a new column to the matrix A and so create an extended matrix $[A, y] \in \mathbb{R}^{m\times(n+1)}$

while the 2-norm

$$\|A\|_2 := \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1.$$

The core of matrix approximation problem is stated by Eckart-Young ([25] with Frobenious norm) and Mirsky ([66] with 2-norm) and summarised on the next result, known as *Eckart-Young-Mirsky (matrix approximation) theorem.*

Theorem 3.1.3 ([93, Thm 2.3])**.** *Let the SVD of $A \in \mathbb{R}^{m\times n}$ be given by $A = \sum_{i=1}^{r} \sigma_i u_i {v_i}^T$ with $r = \text{rank}(A)$. If $k < r$ and $A_k = \sum_{i=1}^{k} \sigma_i u_i {v_i}^T$, then*

$$\min_{\text{rank}(D)=k} \|A - D\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

and

$$\min_{\text{rank}(D)=k} \|A - D\|_F = \|A - A_k\|_F = \sqrt{\sum_{i=k+1}^{p} \sigma_i^2}, \quad p = \min\{m, n\}$$

On the following we give a close form characterising the TLS solution, similar as (3.5) for the LS solution.

Theorem 3.1.4 ([93, Thm 2.7])**.** *Let $A_\epsilon = U'\Sigma' V'^T$ (respectively, $[A_\epsilon, y_\delta] = U\Sigma V^T$) be the SVD decomposition of A_ϵ (respectively, $[A_\epsilon, y_\delta]$). If $\sigma'_n > \sigma_{n+1}$, then*

$$x^{TLS} = (A_\epsilon^T A_\epsilon - \sigma_{n+1}^2 I)^{-1} A_\epsilon^T y_\delta \tag{3.8}$$

and

$$\sigma_{n+1}^2 \left[1 + \sum_{i=1}^{n} \frac{(u_i'^T y_\delta)^2}{\sigma_i'^2 - \sigma_{n+1}^2}\right] = \min \left\|A_\epsilon x - y_\delta\right\|_2^2 \tag{3.9}$$

In order to illustrate the effect of the use of TLS as opposed to LS, we consider here the simplest example of *parameter estimation in 1D.*

Example 1. Find the slope m of the linear equation

$$xm = y$$

for given a set of eight pairs measurements (x_i, y_i), where $x_i = y_i$ for $1 \leq i \leq 8$. It is easy to find out that the slop (solution) is $m = 1$. Although this example is straightforward and well-posed, we can learn the following geometric interpretation: the LS solution displayed on Figure 3.2 with measurements on the left-hand side fits the curve on the horizontal direction, since the axis y is free of noise; meanwhile, the LS solution displayed on Figure 3.3 with measurements on the right-hand side fits the curve on the vertical direction, since the axis x is fixed (free of noise).

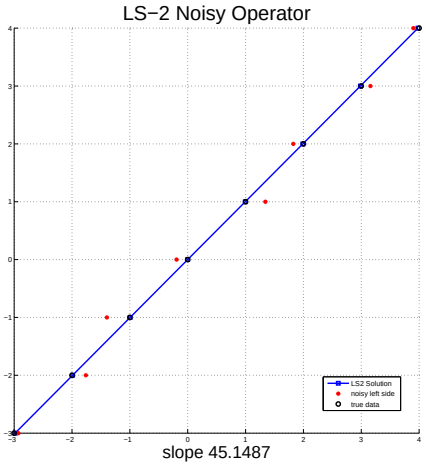

Figure 3.2: Solution for the data (x_ϵ, y_0), i.e., only noise on the left-hand side.

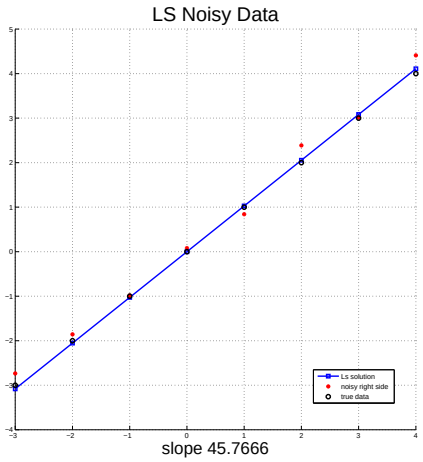

Figure 3.3: Solution for the data (x_0, y_δ), i.e., only noise on the right-hand side.

The TLS solution on Figure 3.4 illustrates the estimation with noise on both directions and now the deviations are orthogonal to the fitted line, i.e., it minimises the sum of squares of their lengths. Therefore, this estimation procedure is sometimes called as *orthogonal regression*.

Van Loan commented on her book [93] that in typical applications, gains of 10–15 percent in accuracy can be obtained using TLS over the standard LS method, almost at no extra computational cost. Moreover, it becomes more effective when more measurements can be made.

Another formulation for the TLS, investigate e.g. in [57], of the set $A_\epsilon x \approx y_\delta$ is given through the following constrained problem

$$\begin{array}{ll} \text{minimise} & \left\|A - A_\epsilon\right\|_F^2 + \left\|y - y_\delta\right\|_2^2 \\ \text{subject to} & Ax = y \end{array} . \tag{3.10}$$

This formulation emphasises the perpendicular distance by minimising the sum of squared misfit in each direction. One can also recast this constrained minimisation as an unconstrained problem, by replacing $y = Ax$ in the second term of Equation (3.10).

In the upcoming section we extend this approach to the regularised version, that is, adding a stabilisation term.

3.2 Regularised Total Least Squares

Since our focus is on very ill-posed problems the approach introduced in the previous section is no longer efficient. We can observe from the discretisation

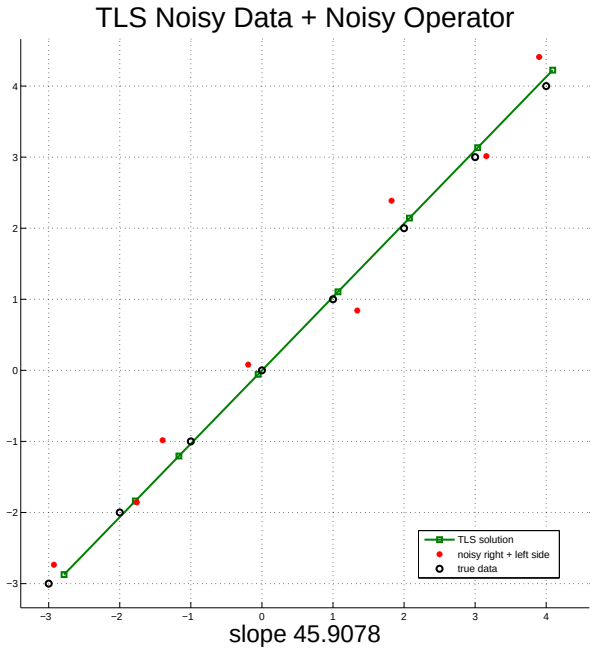

Figure 3.4: Solution for the data (x_ϵ, y_δ), i.e., noise on the both sides.

of ill-posed problems, such as integral equations of the first kind, that the singular values of the discrete operator decay gradually to zero. The need of a stabilisation term leads us to regularisation methods, e.g., the Tikhonov method already defined in (1.5). We introduce now one equivalent formulation (see commentaries on [33]) called *regularised least squares* problem, as the following constrained optimisation problem

$$\begin{array}{ll} \text{minimise} & \left\|A_0 x - y_\delta\right\|_2^2 \\ \text{subject to} & \left\|Lx\right\|^2 \leq M \end{array} . \tag{3.11}$$

This idea can be carried over when both sides of the underlying Equation (3.1) are contaminated with some noise, i.e., using TLS instead of the LS misfit term.

So was Tikhonov regularisation recast as a TLS formulation and the resulting was coined *regularised total least squares* method (R-TLS), see [34, 40, 33]. Intuitively it is added some constrained to the TLS problem (3.10). Consequently, in a finite dimensional setting[4], the R-TLS method can be formu-

[4]we keep the same notation as in the infinite dimensional setup.

lated as

$$\begin{array}{ll} \text{minimise} & \left\|A - A_\epsilon\right\|_F^2 + \left\|y - y_\delta\right\|_2^2 \\ \text{subject to} & \begin{cases} Ax = y \\ \left\|Lx\right\|_2^2 \leq M\,. \end{cases} \end{array} \tag{3.12}$$

The optimal pair $(\hat{A}, \hat{y})$ minimises the residual in the operator and in the data, measured by Frobenius and Euclidian norm, respectively. Moreover, the solution pair is connected via the equation $\hat{A}x = \hat{y}$, where the element x belongs to a ball in $\mathcal{V}$ of radius M. The "size" of the ball is measured by a linear and invertible operator L (often the identity). Any element x^R satisfying these constraineds defines a *R-TLS solution*.

The *Karush-Kuhn-Tucker* (KKT[5]) condition for the optimisation problem introduced in (3.12) are summarised in the upcoming result.

Theorem 3.2.1 ([33, Thm 2.1])**.** *If the inequality constrained is active, then*

$$\left(A_\epsilon^T A_\epsilon + \alpha L^T L + \beta I\right)x^R = A_\epsilon^T y_\delta \text{ and } \left\|Lx^R\right\| = M$$

with $\alpha = \mu(1 + \left\|x^R\right\|^2)$, $\beta = -\dfrac{\left\|A_\epsilon x^R - y_\delta\right\|^2}{1 + \left\|x^R\right\|^2}$ *and* $\mu > 0$ *is the Lagrange multiplier. The two parameters are related by*

$$\alpha M^2 = y_\delta^T(y_\delta - A_\epsilon x^R) + \beta.$$

Moreover, the TLS residual satisfies

$$\left\|[A_\epsilon, y_\delta] - [\hat{A}, \hat{y}]\right\|_F^2 = -\beta \tag{3.13}$$

The main drawback on this approach is the following: the method requires a reliable bound M for the norm $\left\|Lx^{true}\right\|^2$, where such estimation for the true solution is not known. In [57] there is an example showing the dependence and instability of the method for different values of M.

Observe that the R-TLS residual given in (3.13) is a *weighted LS* misfit term. In other words, it is minimised the LS error with weight

$$w(x) = \frac{1}{1 + \left\|x\right\|^2}\,. \tag{3.14}$$

Moreover, the solution of both problems are the same, as stated in the next theorem.

[5]KKT are first order necessary conditions for a solution in non-linear programming to be optimal, provided that some regularity conditions are satisfied; see more details in [73].

Theorem 3.2.2 ([57, Thm 2.3]). *The R-TLS problem solution of the problem* (3.12) *is the solution of the constrained minimisation problem*

$$\underset{x}{\text{minimise}} \frac{\left\|A_\epsilon x - y_\delta\right\|^2}{1 + \left\|x\right\|^2} \quad \text{subject to} \quad \left\|Lx\right\| \leq M \tag{3.15}$$

In the next section we comment on another approach to deal with this class of problems. Moreover, this approach leads to error bounds.

3.3 Dual Regularised Total Least Squares

The accuracy of the R-TLS depends heavily on the right choice of M, which is usually difficult to obtain, as commented previously.

An alternative is the *dual regularised total least square* (D-RTLS) method proposed few years ago [57, 61, 60]. When some reliable bounds for the noise levels δ and ϵ are known it makes sense to look for approximations $(\hat{A}, \hat{x}, \hat{y})$ which satisfy the side conditions

$$Ax = y, \quad \left\|y - y_\delta\right\| \leq \delta \quad \text{and} \quad \left\|A - A_\epsilon\right\| \leq \epsilon.$$

The solution set characterised by these three side conditions is non-empty, according to [57]. This is the major advantage of the dual method over the R-TLS, because we can avoid the dependence of the bound M.

Selecting from the solution set the element which minimises $\|Lx\|$ leads us to a problem in which some estimate $(\hat{A}, \hat{x}, \hat{y})$ for (A_0, x^{true}, y_0) is determined by solving the constrained minimisation problem

$$\begin{aligned} &\text{minimise} && \|Lx\|_2^2 \\ &\text{subject to} && \begin{cases} Ax = y \\ \|y - y_\delta\|_2^2 \leq \delta \\ \|A - A_\epsilon\|_F^2 \leq \epsilon\,, \end{cases} \end{aligned} \tag{3.16}$$

where $\|\cdot\|_F$ still denotes again the Frobenius norm. Please note that most of the available results on this method do again require a finite dimensional setup, see, e.g., [57, 61, 88].

Theorem 3.3.1 ([57, Thm 3.2]). *If the two inequalities constraineds are active, then the dual R-TLS solution x^D of the problem* (3.16) *is a solution of the equation*

$$\left(A_\epsilon^T A_\epsilon + \alpha L^T L + \beta I\right) x^D = A_\epsilon^T y_\delta$$

with $\alpha = \dfrac{\nu + \mu\left\|x^D\right\|^2}{\nu\mu}$, $\beta = -\dfrac{\mu\left\|A_\epsilon x^D - y_\delta\right\|^2}{\nu + \mu\left\|x^D\right\|^2}$ *and* $\nu, \mu > 0$ *are Langrange multipliers. Moreover,*

$$\left\|A_\epsilon x^D - y_\delta\right\| = \delta + \epsilon\left\|x^D\right\| \quad \text{and} \quad \beta = -\frac{\epsilon\left(\delta + \epsilon\left\|x^D\right\|\right)}{\left\|x^D\right\|} \tag{3.17}$$

As result of the above theorem (see [57, Remark 3.4]), if the two constraineds of the dual problem are active, then we can also characterise either by the *constrained* minimisation problem

$$\begin{array}{ll} \text{minimise} & \|Lx\| \\ \text{subject to} & \|A_\epsilon x - y_\delta\| = \delta + \epsilon\|x\| \end{array}$$

or by the *unconstrained* minimisation problem

$$\underset{x}{\text{minimise}}\, \|A_\epsilon x - y_\delta\|^2 + \alpha\|Lx\|^2 - (\delta + \epsilon\left\|x\right\|)^2$$

wiht α chosen by the nonlinear equation $\|A_\epsilon x - y_\delta\| = \delta + \epsilon\|x\|$.

The relation of constrained and unconstrained minimisation problems is essential for understanding the new regularisation method proposed in the upcoming Chapter 4.

Additionally to this short revision we list two important theorems concerning error bounds for both R-TLS and D-RTLS method, for the standard case $L = I$ (identity operator). As indicated in the article [57] these are the first results to prove order optimal error bounds so far given in the literature and they depend on the following classical source condition

$$x^\dagger = A_0^*\omega \quad \omega \in \mathcal{U}. \tag{3.18}$$

This SC-I is the same type assumed on the previous chapter, see (2.5) and (2.16), respectively, for the linear and non-linear case.

Theorem 3.3.2 ([57, Thm 6.2])**.** *Assume that the exact solution* $x^\dagger$ *of the problem* (3.1) *satisfies the SC* (3.18) *and let* x^D *be the D-RTLS solution of the problem* (3.16)*. Then*

$$\left\|x^D - x^\dagger\right\| \leq 2\|\omega\|^{1/2}\sqrt{\delta + \epsilon\|x^\dagger\|}.$$

In contrast we present also convergence rate for the R-TLS solution, that is to say, both of order $\mathcal{O}\left(\sqrt{\delta + \epsilon}\right)$.

Theorem 3.3.3 ([57, Thm 6.1])**.** *Assume that the exact solution* $x^\dagger$ *of the problem* (3.1) *satisfies the SC* (3.18) *and the side condition* $\left\|x^\dagger\right\| = M$*. Let in addition* x^R *be the R-TLS solution of the problem* (3.12)*, then*

$$\left\|x^R - x^\dagger\right\| \leq (2 + 2\sqrt{2})^{1/2}\|\omega\|\max\left\{1, M^{1/2}\right\}\sqrt{\delta + \epsilon}.$$

3.4 Comments on Related Problems

Heretofore we listed only few approaches to treat ill-posed problems with error in both operator and data, namely, the first regularised version of TLS (R-TLS) method proposed in 1999 and the D-RTLS, which was the first approach given with rates of convergence.

One efficient algorithm for solving the R-TLS problem was developed in [78], based on the minimisation of the *Rayleigh quotient*

$$\phi(x) := \frac{\left\|A_\epsilon x - y_\delta\right\|^2}{1 + \left\|x\right\|^2}.$$

To be more precise, it solves the equivalent problem (3.15), also known as *weighted LS* or *normalised residual* problem, instead of minimising the constrained functional (3.12). Usually one refers to this formulation as *regularised Rayleigh quotient form for total least squares* (RQ-RTLS).

Adding a quadratic constrained to the TLS minimisation problem can be solved via a *quadratic eigenvalue problem* [32]. It results in an iterative procedure for solving the R-TLS proposed in [86], named as *regularised total least squares solved by quadratic eigenvalue problem* (RTLSQEP). The authors of [54] also analysed the same problem, focusing in the efficiency of solving the R-TLS in mainly two different approaches: via a sequence of quadratic eigenvalue problems and via a sequence of linear eigenvalue problems.

A typical algorithm for solving the D-RTLS is based on *model function*, see e.g., [61, 60]. The D-RTLS solution x^D has a close form given in the Theorem 3.3.1, but it depends on two variables, i.e., $x^D = x(\alpha, \delta)$. The parameters α and β are found to be the solution of the (highly) non-linear system (3.17). The main idea is to approximate the unconstrained minimisation problem by a simple function which relates the derivatives of this functional with respect to each parameter. The "simple" function is called *model function* and it is denoted by $m(\alpha, \beta)$, to emphasise its parametrisation, and it should solve a differential equation. We skip further comments and formulas, recommending to the reader the article [61] and references therein for more details.

Finally we cite a very new approach towards non-linear operators [58]. In this article is considered a standard Tikhonov-type functional for solving a non-linear operator equation of the form $F_0(u) = g_0$, where additionally to the noisy data g_δ it is assumed the noisy operator F_ϵ holds

$$\sup_u \left\|F_0(u) - F_\epsilon(u)\right\| \leq \epsilon$$

with a known constant ϵ referring to the operator noise level. A regularised solution is obtained, as well as convergence rates.

Chapter 4

Double Regularised Total Least Squares

> "The mere formulation of a problem is far more often essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old problems from a new angle requires creative imagination and marks real advances in science."
>
> *Albert Einstein*

In our approach, we would like to restrict our attention to linear operators that can be mainly characterised by a function, as it is, e.g., the case for linear integral operators, where the kernel function determines the behaviour of the operator. Moreover, we will assume that the noise in the operator is due to an incorrect characterising function. This approach will allow us to treat the problem of finding a solution of an operator equation from incorrect data and operator in the framework of Tikhonov regularisation rather than as a constrained minimisation problem.

In this chapter we introduce the proposed method as well as its mathematical setting. We focus on analysing its regularisation properties: existence, stability and convergence. Additionally we study source condition and derive convergence rates with respect to Bregman distance.

4.1 Problem Formulation

We aim at the inversion of linear operator equation

$$A_0 f = g_0$$

from noisy data g_δ and incorrect operator A_ϵ. Additionally we assume that the operators $A_0, A_\epsilon : \mathcal{V} \to \mathcal{H}$, $\mathcal{V}, \mathcal{H}$ Hilbert spaces, can be characterised by

functions $k_0, k_\epsilon \in \mathcal{U}$. To be more specific, we consider operators

$$\begin{aligned} A_k : \quad \mathcal{V} &\longrightarrow \mathcal{H} \\ v &\longmapsto B(k, v) \ , \end{aligned}$$

i.e., $A_k v := B(k, v)$, where B is a *bilinear* operator

$$B : \mathcal{U} \times \mathcal{V} \to \mathcal{H}$$

fulfilling, for some $C > 0$,

$$\big\|B(k, f)\big\|_{\mathcal{H}} \leq C \big\|k\big\|_{\mathcal{U}} \big\|f\big\|_{\mathcal{V}}. \tag{4.1}$$

From (4.1) follows immediately

$$\big\|B(k, \cdot)\big\|_{\mathcal{V}\to\mathcal{H}} \leq C \big\|k\big\|_{\mathcal{U}}. \tag{4.2}$$

Associated with the bilinear operator B, we also define the linear operator

$$\begin{aligned} C_f : \quad \mathcal{U} &\longrightarrow \mathcal{H} \\ u &\longmapsto B(u, f) \ , \end{aligned}$$

i.e., $C_f u := B(u, f)$.

From now on, let us identify A_0 with A_{k_0} and A_ϵ with A_{k_ϵ}. From (4.2) we deduce immediately

$$\|A_0 - A_\epsilon\| \leq C \|k_0 - k_\epsilon\| \ , \tag{4.3}$$

i.e., the operator error norm is controlled by the error norm of the characterising functions. Now we can formulate our problem as follows:

$$\begin{aligned} \text{Solve} \qquad & A_0 f = g_0 && \text{(4.4a)} \\ \text{from noisy data } g_\delta \text{ with} \qquad & \|g_0 - g_\delta\| \leq \delta && \text{(4.4b)} \\ \text{and noisy function } k_\epsilon \text{ with} \qquad & \|k_0 - k_\epsilon\| \leq \epsilon \ . && \text{(4.4c)} \end{aligned}$$

Please note that the problem with explicitly known k_0 (or the operator A_0) is often ill-posed and needs regularisation for a stable inversion. Therefore we will also propose a regularising scheme for the problem (4.4a)-(4.4c). Now let us give some examples.

Example 2. Consider a linear integral operator A_0 defined through

$$\left(A_0 f\right)(s) := \int_\Omega k_0(s, t) f(t) dt = B(k_0, f)$$

with $\mathcal{V} = \mathcal{H} = L_2(\Omega)$ and let k_0 be a function in $\mathcal{U} = L_2(\Omega^2)$. Then the bilinear operator B yields

$$\|B(k_0, f)\| \leq \|k_0\|_{\mathcal{U}} \|f\|_{\mathcal{V}}.$$

The considered class of operators also contains deconvolution problems, which are important in imaging, as well as *blind deconvolution* problems [53, 14, 48], where it is assumed that also the exact convolution kernel is unknown.

Example 3. In medical imaging, the data of *Single Photon Emission Computed Tomography* (SPECT) is described by the attenuated Radon transform [72, 24, 76]:

$$Af(s,\omega) = \int_{\mathbb{R}} f(s\omega^{\perp} + t\omega) \cdot e^{-\int_t^{\infty} \mu(s\omega^{\perp}+\tau\omega)\, d\tau}\, dt \ .$$

The function μ is the density distribution of the body. In general, the density distribution is also unknown. Modern scanner, however, perform a CT scan in parallel. Due to measurement errors, the reconstructed density distribution is also incorrect. Setting

$$k_\epsilon(s,t,\omega) = e^{-\int_t^{\infty} \mu_\epsilon(s\omega^{\perp}+\tau\omega)\, d\tau} \ ,$$

we have

$$A_\epsilon f = B(k_\epsilon, f) \ ,$$

and similar estimates as in (4.1) can be obtained.

4.2 Proposed Method

Due to our assumptions on the structure of the operator A_0, the inverse problem of identifying the function f^{true} from noisy measurements g_δ and inexact operator A_ϵ can now be rewritten as the task of solving the inverse problem

$$B(k_0, f) = g_0 \tag{4.5}$$

from noisy measurements (k_ϵ, g_δ) fulfilling

$$\left\|g_0 - g_\delta\right\|_{\mathcal{H}} \leq \delta, \tag{4.6a}$$

and

$$\left\|k_0 - k_\epsilon\right\|_{\mathcal{U}} \leq \epsilon. \tag{4.6b}$$

In most applications, the "inversion" of B will be ill-posed (e.g., if B is defined via a Fredholm integral operator), and a regularisation strategy is needed for a stable solution of the problem (4.5).

The structure of our problem allows to reformulate (4.4a)-(4.4c) as an unconstrained Tikhonov-type problem:

$$\underset{(k,f)}{\text{minimise}} \; J^{\delta,\varepsilon}_{\alpha,\beta}(k,f) := \frac{1}{2}T^{\delta,\varepsilon}(k,f) + R_{\alpha,\beta}(k,f) \;, \tag{4.7a}$$

where

$$T^{\delta,\varepsilon}(k,f) = \left\|B(k,f) - g_\delta\right\|^2 + \gamma\left\|k - k_\epsilon\right\|^2 \tag{4.7b}$$

and

$$R_{\alpha,\beta}(k,f) = \frac{\alpha}{2}\left\|Lf\right\|^2 + \beta\mathcal{R}(k). \tag{4.7c}$$

Here, α and β are the regularisation parameters which have to be chosen properly, γ is a scaling parameter, L is a bounded linear and continuously invertible operator and $\mathcal{R} : X \subset \mathcal{U} \to [0, +\infty]$ is proper, convex and weakly lower semi-continuous functional . We wish to note that most of the available papers assume that L is a densely defined, unbounded self-adjoint and strictly positive operator, see, e.g. [57, 59]. For our analysis, however, boundedness is needed and it is an open question whether the analysis could be extended to cover unbounded operators, too.

We call this scheme the *double regularised total least squares method* (dbl-RTLS). Please note that the method is closely related to the total least squares method, as the term $\left\|k-k_\epsilon\right\|^2$ controls the error in the operator. The functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ is composed as the sum of two terms: one which measures the discrepancy of data and operator, and one which promotes stability. The functional $T^{\delta,\varepsilon}$ is a *data-fidelity* term based on the TLS technique, whereas the functional $R_{\alpha,\beta}$ acts as a *penalty* term which stabilizes the inversion with respect to the pair (k, f). As a consequence, we have two regularisation parameters, which also occurs in *double regularisation*, see, e.g., [98].

The domain of the functional $J^{\delta,\varepsilon}_{\alpha,\beta} : \big(\mathcal{U} \cap X\big) \times \mathcal{V} \longrightarrow \mathbb{R}$ can be extended over $\mathcal{U} \times \mathcal{V}$ by setting $\mathcal{R}(k) = +\infty$ whenever $k \in \mathcal{U} \setminus X$. Then $\mathcal{R}$ is proper, convex and weak lower semi-continuous functional in $\mathcal{U}$.

4.3 Regularisation Properties

In this section we shall analyse some analytical properties of the proposed dbl-RTLS method. In particular, we prove its well-posedness as a regularisation method, i.e., the minimisers of the regularisation functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ exist for every $\alpha, \beta > 0$, depend continuously on both g_δ and k_ϵ, and converge to a solution of $B(k_0, f) = g_0$ as both noise level approaches zero, provided the regularisation parameters α and β are chosen appropriately.

For the pair $(k, f) \in \mathcal{U} \times \mathcal{V}$ we use the canonical inner product

$$\langle (k_1, f_1), (k_2, f_2) \rangle_{\mathcal{U}\times\mathcal{V}} := \langle k_1, k_2 \rangle_{\mathcal{U}} + \langle f_1, f_2 \rangle_{\mathcal{V}} \ ,$$

i.e., convergence is defined componentwise. For the upcoming results, we need the following assumption on the operator B:

Assumption D.

(D1) B is strongly continuous, i.e., if $(k^n, f^n) \rightharpoonup (\bar{k}, \bar{f})$ then $B(k^n, f^n) \to B(\bar{k}, \bar{f})$.

Proposition 4.3.1. *Let $J^{\delta,\varepsilon}_{\alpha,\beta}$ be the functional defined in (4.7). Assume that L is a bounded linear and continuously invertible operator and B fulfills Assumption D1. Then $J^{\delta,\varepsilon}_{\alpha,\beta}$ is a positive, weakly lower semi-continuous and coercive functional.*

Proof. By the definition of $T^{\delta,\varepsilon}$, $\mathcal{R}$ and Assumption D1, $J^{\delta,\varepsilon}_{\alpha,\beta}$ is positive and w-lsc. As the operator L is continuously invertible, there exists a constant $c > 0$ such that

$$c\|f\| \leq \|Lf\|$$

for all $f \in \mathscr{D}(L)$. We get

$$J^{\delta,\varepsilon}_{\alpha,\beta}(k, f) \geq \gamma\|k - k_\epsilon\|^2 + \frac{\alpha c}{2}\|f\|^2 \to \infty$$

as $\|(k, f)\|^2 := \|k\|^2 + \|f\|^2 \to \infty$ and therefore $J^{\delta,\varepsilon}_{\alpha,\beta}$ is coercive. □

We point out here that the problem (4.5) may not even have a solution for any given noisy measurements (k_ϵ, g_δ) whereas the regularised problem (4.7) does, as stated below:

Theorem 4.3.2 (Existence)**.** *Let the assumptions of Proposition 4.3.1 hold. Then the functional $J^{\delta,\varepsilon}_{\alpha,\beta}(k, f)$ has a global minimiser.*

Proof. By Proposition 4.3.1, $J^{\delta,\varepsilon}_{\alpha,\beta}(k, f)$ is positive, proper and coercive, i.e., there exists $(k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta})$ such that $J^{\delta,\varepsilon}_{\alpha,\beta}(k, f) < \infty$.

Let $\nu = \inf\{ J^{\delta,\varepsilon}_{\alpha,\beta}(k, f) \mid (k, f) \in \operatorname{dom} J^{\delta,\varepsilon}_{\alpha,\beta}\}$. Then, there exists $M > 0$ and a sequence $(k^j, f^j) \in \operatorname{dom} J^{\delta,\varepsilon}_{\alpha,\beta}$ such that $J(k^j, f^j) \to \nu$ and

$$J^{\delta,\varepsilon}_{\alpha,\beta}(k^j, f^j) \leq M \quad \forall j.$$

In particular we have

$$\frac{1}{2}\alpha\|Lf^j\|^2 \leq M \quad \text{and} \quad \frac{1}{2}\gamma\|k^j - k_\epsilon\|^2 \leq M.$$

Using

$$\|k^j\| - \|k_\epsilon\| \leq \|k^j - k_\epsilon\| \leq \left(\frac{2M}{\gamma}\right)^{1/2}$$

it follows

$$\|k^j\| \leq \left(\frac{2M}{\gamma}\right)^{1/2} + \|k_\epsilon\| \quad \text{and} \quad \|f^j\| \leq \left(\frac{2M}{\alpha c^2}\right)^{1/2},$$

i.e., the sequences (k^j) and (f^j) are bounded. Thus there exist subsequences of (k^j), (f^j) (for simplicity, again denoted by (k^j) and (f^j)) s.t.

$$k^j \rightharpoonup \bar{k} \quad \text{and} \quad f^j \rightharpoonup \bar{f},$$

and thus

$$(k^j, f^j) \rightharpoonup (\bar{k}, \bar{f}) \in (\mathcal{U} \cap X) \times \mathcal{V}.$$

By the w-lsc of the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ we obtain

$$\nu \leq J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big) \leq \liminf J^{\delta,\varepsilon}_{\alpha,\beta}\big(k^j, f^j\big) = \lim J^{\delta,\varepsilon}_{\alpha,\beta}\big(k^j, f^j\big) = \nu$$

Hence $\nu = J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big)$ is the minimum of the functional and $(\bar{k}, \bar{f})$ is a global minimiser,

$$(\bar{k}, \bar{f}) = \arg\min\{\ J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big)\ \mid\ (k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta})\}.$$

□

The stability property of the standard Tikhonov regularisation strategy for problems with noisy right hand side is well known. We next investigate this property for the Tikhonov-type regularisation scheme (4.7) for perturbations on both (k_ϵ, g_δ).

Theorem 4.3.3 (Stability)**.** *Let $\alpha, \beta > 0$ be fixed the regularisation parameters, L a bounded and continuously invertible operator and $(g_{\delta_j})_j$, $(k_{\epsilon_j})_j$ sequences with $g_{\delta_j} \to g_\delta$ and $k_{\epsilon_j} \to k_\epsilon$. If (k^j, f^j) denote minimisers of $J^{\delta_j,\varepsilon_j}_{\alpha,\beta}$ with data g_{δ_j} and characterising function k_{ϵ_j}, then there exists a convergent subsequence of $(k^j, f^j)_j$. The limit of every convergent subsequence is a minimiser of the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$.*

Proof. By the definition of (k^j, f^j) as minimisers of $J^{\delta_j,\varepsilon_j}_{\alpha,\beta}$ we have

$$J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\big(k^j, f^j\big) \leq J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\big(k, f\big) \quad \forall (k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta}), \tag{4.8}$$

With $(\tilde{k}, \tilde{f}) := (k^{\delta,\epsilon}_{\alpha,\beta}, f^{\delta,\epsilon}_{\alpha,\beta})$ we get $J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\big(\tilde{k}, \tilde{f}\big) \to J^{\delta,\varepsilon}_{\alpha,\beta}\big(\tilde{k}, \tilde{f}\big)$. Hence, there exists a $\tilde{c} > 0$ so that $J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\big(\tilde{k}, \tilde{f}\big) \leq \tilde{c}$ for j sufficiently large. In particular,

we observe with (4.8) that $\left(\left\|k^j - k_{\epsilon_j}\right\|\right)_j$ as well as $\left(\left\|Lf^j\right\|\right)_j$ are uniformly bounded.

Analogous to the proof of Theorem 4.3.2 we conclude that the sequence $(k^j, f^j)_j$ is uniformly bounded. Hence there exists a subsequence (for simplicity also denoted by$(k^j, f^j)_j$) such that

$$k^j \rightharpoonup \bar{k} \quad \text{and} \quad f^j \rightharpoonup \bar{f}.$$

By the weak lower semicontinuity (w-lsc) of the norm and continuity of B we have

$$\left\|B(\bar{k}, \bar{f}) - g_\delta\right\| \leq \liminf_j \left\|B(k^j, f^j) - g_{\delta_j}\right\|$$

and

$$\left\|\bar{k} - k_\epsilon\right\| \leq \liminf_j \left\|k^j - k_{\epsilon_j}\right\|.$$

Moreover, (4.8) implies

$$\begin{aligned} J^{\delta,\varepsilon}_{\alpha,\beta}\left(\bar{k}, \bar{f}\right) &\leq \liminf_j J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\left(k^j, f^j\right) \\ &\leq \limsup_j J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\left(k, f\right) \\ &= \lim_j J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\left(k, f\right) \\ &= J^{\delta,\varepsilon}_{\alpha,\beta}\left(k, f\right) \end{aligned}$$

for all $(k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta})$. In particular, $J^{\delta,\varepsilon}_{\alpha,\beta}\left(\bar{k}, \bar{f}\right) \leq J^{\delta,\varepsilon}_{\alpha,\beta}\left(\tilde{k}, \tilde{f}\right)$. Since $(\tilde{k}, \tilde{f})$ is by definition a minimiser of $J^{\delta,\varepsilon}_{\alpha,\beta}$, we conclude $J^{\delta,\varepsilon}_{\alpha,\beta}\left(\bar{k}, \bar{f}\right) = J^{\delta,\varepsilon}_{\alpha,\beta}\left(\tilde{k}, \tilde{f}\right)$ and thus

$$\lim_{j\to\infty} J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\left(k^j, f^j\right) = J^{\delta,\varepsilon}_{\alpha,\beta}\left(\bar{k}, \bar{f}\right). \tag{4.9}$$

It remains to show

$$k^j \to \bar{k} \quad \text{and} \quad f^j \to \bar{f}.$$

As the sequences are weakly convergent, convergence of the sequences holds if

$$\left\|k^j\right\| \to \left\|\bar{k}\right\| \quad \text{and} \quad \left\|f^j\right\| \to \left\|\bar{f}\right\|.$$

The norms on $\mathcal{U}$ and $\mathcal{V}$ are w-lsc, thus it is sufficient to show

$$\left\|\bar{k}\right\| \geq \limsup\left\|k^j\right\| \quad \text{and} \quad \left\|\bar{f}\right\| \geq \limsup\left\|f^j\right\|.$$

The operator L is bounded and continuously invertible, therefore $f^j \to \bar{f}$ if and only if $Lf^j \to L\bar{f}$. Therefore, we accomplish the prove for the sequence $(Lf^j)_j$. Now suppose there exists τ_1 as

$$\tau_1 := \limsup\left\|Lf^j\right\| > \left\|L\bar{f}\right\|$$

and there exists a subsequence $(f^n)_n$ of $(f^j)_j$ such that $Lf^n \rightharpoonup L\bar{f}$ and $\|Lf^n\| \to \tau_1$.
From the first part of this proof (4.9), it holds

$$\lim_{j\to\infty} J^{\delta_j,\varepsilon_j}_{\alpha,\beta}\big(k^j, f^j\big) = J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big).$$

Using (4.7) we observe

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{1}{2}\|B(k^n,f^n)-g_{\delta_n}\|^2+\frac{\gamma}{2}\|k^n-k_{\epsilon_n}\|^2+\beta\mathcal{R}(k^n)\right)\\
=\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\frac{\gamma}{2}\|\bar{k}-k_\epsilon\|^2+\beta\mathcal{R}(\bar{k})+\frac{\alpha}{2}\left(\|L\bar{f}\|^2-\lim_{n\to\infty}\|Lf^n\|^2\right)\\
=\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\frac{\gamma}{2}\|\bar{k}-k_\epsilon\|^2+\beta\mathcal{R}(\bar{k})+\frac{\alpha}{2}\left(\|L\bar{f}\|^2-{\tau_1}^2\right)\\
<\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\frac{\gamma}{2}\|\bar{k}-k_\epsilon\|^2+\beta\mathcal{R}(\bar{k})\ ,
\end{aligned}$$

which is a contradiction to the w-lsc property of the involved norms and the functional $\mathcal{R}$. Thus $Lf^j \to L\bar{f}$ and

$$f^j \to \bar{f}.$$

The same idea can be used in order to prove convergence of the characterising functions. Suppose there exists τ_2 s.t.

$$\tau_2 := \limsup\|k^j - k_\epsilon\| > \|\bar{k} - k_\epsilon\|$$

and there exists a subsequence $(k^n)_n$ of $(k^j)_j$ such that $(k^n - k_\epsilon) \rightharpoonup (\bar{k} - k_\epsilon)$ and $\|k^n - k_\epsilon\| \to \tau_2$.

By the triangle inequality we get

$$\|k^n - k_\epsilon\| - \|k_{\epsilon_n} - k_\epsilon\| \le \|k^n - k_{\epsilon_n}\| \le \|k^n - k_\epsilon\| + \|k_{\epsilon_n} - k_\epsilon\|,$$

and thus

$$\lim_{n\to\infty}\|k^n - k_{\epsilon_n}\| = \lim_{n\to\infty}\|k^n - k_\epsilon\|.$$

Therefore

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{1}{2}\|B(k^n,f^n)-g_{\delta_n}\|^2+\beta\mathcal{R}(k^n)\right)\\
=\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\frac{\gamma}{2}\left(\|\bar{k}-k_\epsilon\|^2-\lim_{n\to\infty}\|k^n-k_{\epsilon_n}\|^2\right)+\beta\mathcal{R}(\bar{k})\\
=\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\frac{\gamma}{2}\left(\|\bar{k}-k_\epsilon\|^2-{\tau_2}^2\right)+\beta\mathcal{R}(\bar{k})\\
<\;&\frac{1}{2}\|B(\bar{k},\bar{f})-g_\delta\|^2+\beta\mathcal{R}(\bar{k})\ ,
\end{aligned}$$

which is again a contradiction to the w-lsc of the involved norms and functionals. □

In the following, we investigate the regularisation property of our approach, i.e., we show, under an appropriate parameter choice rule, that the minimisers $(k_{\alpha,\beta}^{\delta,\epsilon}, f_{\alpha,\beta}^{\delta,\epsilon})$ of the functional (4.7) converge to an exact solution as the noise level (δ, ϵ) goes to zero.

Let us first clarify our notion of a solution. In principle, the equation

$$B(k, f) = g$$

might have different pairs (k, f) as solution. However, as $k_\epsilon \to k_0$ as $\epsilon \to 0$, we get k_0 for free in the limit, that is, we are interested in reconstructing solutions of the equation

$$B(k_0, f) = g.$$

In particular, we want to reconstruct a solution with minimal value of $\|Lf\|$, and therefore define:

Definition 4.3.4. *We call $f^\dagger$ a minimum-norm solution if*

$$f^\dagger = \arg\min_f \{\|Lf\| \mid B(k_0, f) = g_0\}.$$

The definition above is the standard *minimum-norm solution* for the classical Tikhonov regularisation (see for instance [28]).

Furthermore, we have to introduce a regularisation parameter choice which depends on both noise level, defined through (4.10) in the upcoming theorem.

Theorem 4.3.5 (convergence)**.** *Let the sequences of data g_{δ_j} and k_{ϵ_j} with $\|g_{\delta_j} - g_0\| \leq \delta_j$ and $\|k_{\epsilon_j} - k_0\| \leq \epsilon_j$ be given with $\epsilon_j \to 0$ and $\delta_j \to 0$. Assume that the regularisation parameters $\alpha_j = \alpha(\epsilon_j, \delta_j)$ and $\beta_j = \beta(\epsilon_j, \delta_j)$ fulfill $\alpha_j \to 0$, $\beta_j \to 0$, as well as*

$$\lim_{j\to\infty} \frac{\delta_j^2 + \gamma\epsilon_j^2}{\alpha_j} = 0 \quad \text{and} \quad \lim_{j\to\infty} \frac{\beta_j}{\alpha_j} = \eta \tag{4.10}$$

for some $0 < \eta < \infty$.

Let the sequence

$$(k^j, f^j)_j := \left(k_{\alpha_j,\beta_j}^{\delta_j,\epsilon_j}, f_{\alpha_j,\beta_j}^{\delta_j,\epsilon_j}\right)_j$$

be the minimiser of (4.7), obtained from the noisy data g_{δ_j} and k_{ϵ_j}, regularisation parameters α_j and β_j and scaling parameter γ.

Then there exists a convergent subsequence of $(k^j, f^j)_j$ with $k^j \to k_0$ and the limit of every convergent subsequence of $(f^j)_j$ is a minimum-norm solution of (4.5).

Proof. The minimising property of (k^j, f^j) guarantees

$$J^{\delta_j,\varepsilon_j}_{\alpha_j,\beta_j}(k^j, f^j) \leq J^{\delta_j,\varepsilon_j}_{\alpha_j,\beta_j}(k, f), \quad \forall (k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta}).$$

In particular,

$$0 \leq J^{\delta_j,\varepsilon_j}_{\alpha_j,\beta_j}(k^j, f^j) \leq J^{\delta_j,\varepsilon_j}_{\alpha_j,\beta_j}(k_0, f^\dagger) \leq \frac{\delta_j^2 + \gamma\epsilon_j^2}{2} + \frac{\alpha_j}{2}\left\|Lf^\dagger\right\|^2 + \beta_j \mathcal{R}(k_0), \quad (4.11)$$

where $f^\dagger$ denotes a minimum-norm solution of the equation $B(k_0, f) = g_0$, see Definition 4.3.4.

Combining this estimate with the assumptions on the regularisation parameters, we conclude that the sequences

$$\left\|B(k^j, f^j) - g_{\delta_j}\right\|^2, \left\|k^j - k_{\epsilon_j}\right\|^2, \left\|Lf^j\right\|^2, \mathcal{R}(k^j)$$

are uniformly bounded and by the invertibility of L, the sequence $(k^j, f^j)_j$ is uniformly bounded.

Therefore it exists a weakly convergent subsequence $(k^m, f^m)_m := (k^{j_m}, f^{j_m})_{j_m}$ of $(k^j, f^j)_j$ with

$$(k^m, f^m) \rightharpoonup (\bar{k}, \bar{f}).$$

In the following we will prove that for the weak limit $(\bar{k}, \bar{f})$ holds $\bar{k} = k_0$ and $\bar{f}$ is a minimum-norm solution.

By the weak lower semi-continuity of the norm we have

$$\begin{aligned}
0 &\leq \frac{1}{2}\left\|B(\bar{k}, \bar{f}) - g_0\right\|^2 + \frac{\gamma}{2}\left\|\bar{k} - k_0\right\|^2 \\
&\leq \liminf_{m\to\infty} \left\{\frac{1}{2}\left\|B(k^m, f^m) - g_{\delta_m}\right\|^2 + \frac{\gamma}{2}\left\|k^m - k_{\epsilon_m}\right\|^2\right\} \\
&\overset{(4.11)}{\leq} \liminf_{m\to\infty} \left\{\frac{\delta_m^2 + \gamma\epsilon_m^2}{2} + \frac{\alpha_m}{2}\left\|Lf^\dagger\right\|^2 + \beta_m \mathcal{R}(k_0)\right\} \\
&= 0,
\end{aligned}$$

where the last equality follows from the parameter choice rule.

In particular, we have

$$\bar{k} = k_0 \quad \text{and} \quad B(\bar{k}, \bar{f}) = g_0.$$

From (4.11) follows

$$\frac{1}{2}\left\|Lf^m\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k^m) \leq \frac{\delta_m^2 + \gamma\epsilon_m^2}{2\alpha_m} + \frac{1}{2}\left\|Lf^\dagger\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k_0).$$

Again, weak lower semi-continuity of the norm and the functional $\mathcal{R}$ result in

$$\begin{aligned}\frac{1}{2}\left\|L\bar{f}\right\|^2+\eta\mathcal{R}(\bar{k}) &\leq \liminf_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2+\eta\mathcal{R}(k^m)\right\}\\ &= \liminf_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2+\frac{\beta_m}{\alpha_m}\mathcal{R}(k^m)\right\}\\ &\leq \liminf_{m\to\infty}\left\{\frac{\delta_m^2+\gamma\epsilon_m^2}{2\alpha_m}+\frac{1}{2}\left\|Lf^\dagger\right\|^2+\frac{\beta_m}{\alpha_m}\mathcal{R}(k_0)\right\}\\ &\overset{(4.10)}{=} \frac{1}{2}\left\|Lf^\dagger\right\|^2+\eta\mathcal{R}(k_0)\ .\end{aligned}$$

As $\bar{k}=k_0$ we conclude that $\bar{f}$ is a minimum-norm solution and

$$\begin{aligned}\frac{1}{2}\left\|L\bar{f}\right\|^2+\eta\mathcal{R}(\bar{k}) &= \lim_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2+\frac{\beta_m}{\alpha_m}\mathcal{R}(k^m)\right\}\\ &= \frac{1}{2}\left\|Lf^\dagger\right\|^2+\eta\mathcal{R}(k_0).\end{aligned}\tag{4.12}$$

So far we showed the existence of a subsequence $(k^m,f^m)_m$ which converges weakly to $(k_0,\bar{f})$, where $\bar{f}$ is a minimising solution. It remains to show that the sequence also converges in the strong topology of $\mathcal{U}\times\mathcal{V}$.

In order to show $f^m\to\bar{f}$ in $\mathcal{V}$, we prove $Lf^m\to L\bar{f}$. Since is $Lf^m\rightharpoonup L\bar{f}$ it is sufficient to show

$$\left\|Lf^m\right\|\to\left\|L\bar{f}\right\|,$$

or, as the norm is w.-l.s.c.,

$$\limsup_{m\to\infty}\left\|Lf^m\right\|\leq\left\|L\bar{f}\right\|.$$

Assume that the above inequality does not hold. Then there exists a constant τ_1 such that

$$\tau_1:=\limsup_{m\to\infty}\left\|Lf^m\right\|^2>\left\|L\bar{f}\right\|^2$$

and there exists a subsequence of $(Lf^m)_m$ denoted by $(Lf^n)_n:=(Lf^{m_n})_{m_n}$ such that

$$Lf^n\rightharpoonup L\bar{f}\quad\text{ and }\quad\left\|Lf^n\right\|^2\to\tau_1.$$

From (4.12) and the hypothesis stated above

$$\begin{aligned}\limsup_{n\to\infty}\frac{\beta_n}{\alpha_n}\mathcal{R}(k^n) &= \eta\mathcal{R}(k_0)+\frac{1}{2}\left(\left\|L\bar{f}\right\|^2-\limsup_{n\to\infty}\left\|Lf^n\right\|^2\right)\\ &< \eta\mathcal{R}(k_0),\end{aligned}$$

which is a contradiction to the w.-l.s.c. of $\mathcal{R}$. Thus

$$\limsup_{m\to\infty}\left\|Lf^m\right\|\leq\left\|L\bar{f}\right\|,$$

i.e., $f^m \to \bar{f}$ in $\mathcal{V}$.

The convergence of the sequence $(k^m)_m$ in the topology of $\mathcal{U}$ follows straightforward by

$$\begin{aligned} \left\|k^m - k_0\right\| &\leq \left\|k^m - k_{\epsilon_m}\right\| + \left\|k_{\epsilon_m} - k_0\right\| \\ &\leq \left\|k^m - k_{\epsilon_m}\right\| + \epsilon_m \overset{(4.11)}{\longrightarrow} 0 \text{ as } m \to \infty. \end{aligned}$$

Moreover, if $f^\dagger$ is unique, the assertion about the convergence of the whole sequence $(k^j, f^j)_j$ follows from the fact that then every subsequence of the sequence converges towards the same limit $(k_0, f^\dagger)$. □

Remark 4.3.6. *Note that the easiest parameter choice rule fulfilling condition* (4.10) *is given by*

$$\beta = \eta\alpha, \qquad \eta > 0.$$

For this specific choice, we only have one regularisation parameter left, and the problem (4.7) *reduces to*

$$\underset{(k,f)}{\text{minimise}} \; J_\alpha\big(k, f\big) := \frac{1}{2} T^{\delta,\varepsilon}\left(k, f\right) + \alpha\Phi(k, f)\,, \tag{4.13}$$

where $T^{\delta,\varepsilon}$ *is defined in* (4.7b) *and*

$$\Phi(k, f) := \frac{1}{2}\left\|Lf\right\|^2 + \eta\mathcal{R}(k). \tag{4.14}$$

It is well known that, under the general assumptions, the rate of convergence of $(k^j, f^j)_j \to (k_0, f^\dagger)$ for $(\delta_j, \epsilon_j) \to 0$ can be in general arbitrarily slow. For linear and nonlinear inverse problems convergence rates were obtained if *source conditions* are satisfied (see [27, 28, 12, 79] and Chapter 2).

For our analysis, we will use the following source condition:

$$\mathscr{R}(B'(k_0, f^\dagger)^*) \cap \partial\Phi\big(k_0, f^\dagger\big) \neq \emptyset,$$

where $\partial\Phi$ denotes the subdifferential of the functional Φ defined in (4.14). This condition says there exists a subgradient $(\xi_{k_0}, \xi_{f^\dagger})$ of Φ s.t. $(\xi_{k_0}, \xi_{f^\dagger}) = B'(k_0, f^\dagger)^*\omega$, $\omega \in \mathcal{H}$.

Convergence rates are often given with respect to the *Bregman distance* generated by the regularisation functional Φ. In our setting, the distance is defined by

$$D_\Phi^{(\xi_{\bar{u}}, \xi_{\bar{v}})}\big((u, v), (\bar{u}, \bar{v})\big) = \Phi(u, v) - \Phi(\bar{u}, \bar{v}) - \left\langle (\xi_{\bar{u}}, \xi_{\bar{v}})\ ,\ (u, v) - (\bar{u}, \bar{v})\right\rangle \tag{4.15}$$

for $(\xi_{\bar{u}}, \xi_{\bar{v}}) \in \partial\Phi\big(\bar{u}, \bar{v}\big)$.

Lemma 4.3.7. *Let Φ be the functional defined in (4.14) with $L = I$. Then the Bregman distance is given by*

$$D_{\Phi}^{(\xi_{\bar{u}},\xi_{\bar{v}})}\big((u,v),(\bar{u},\bar{v})\big) = \frac{1}{2}\big\|v-\bar{v}\big\|^2 + \eta D_{\mathcal{R}}^{\zeta}\big(u,\bar{u}\big), \tag{4.16}$$

with $\zeta \in \partial\mathcal{R}\big(\bar{u}\big)$.

Proof. By definition of Bregman distance we have

$$\begin{aligned} D_{\Phi}^{(\xi_{\bar{u}},\xi_{\bar{v}})}\big((u,v),(\bar{u},\bar{v})\big) &= \left(\frac{1}{2}\left\|v\right\|^2 + \eta\mathcal{R}(u)\right) - \left(\frac{1}{2}\left\|\bar{v}\right\|^2 + \eta\mathcal{R}(\bar{u})\right) \\ &\quad - \left\langle (\xi_{\bar{u}},\xi_{\bar{v}})\ ,\ (u-\bar{u},v-\bar{v})\right\rangle \\ &= \frac{1}{2}\left\|v\right\|^2 - \frac{1}{2}\left\|\bar{v}\right\|^2 - \left\langle \xi_{\bar{v}}\ ,\ v-\bar{v}\right\rangle \\ &\quad + \eta\mathcal{R}(u) - \eta\mathcal{R}(\bar{u}) - \left\langle \xi_{\bar{u}}\ ,\ u-\bar{u}\right\rangle \\ &= \frac{1}{2}\left\|v-\bar{v}\right\|^2 + \eta D_{\mathcal{R}}^{\zeta}\big(u,\bar{u}\big) \end{aligned}$$

with $\zeta = \frac{1}{\eta}\xi_{\bar{u}}$. Note that the functional Φ is composed as a sum of a differentiable and a convex functional. Therefore, the subgradient of the first functional is an unitary set and it holds (see, e.g.,[19])

$$\begin{aligned} \partial\Phi\big(\bar{u},\bar{v}\big) &= \partial\left(\left\|\bar{v}\right\|^2 + \eta\mathcal{R}(\bar{u})\right) \\ &= \left\{(\xi_{\bar{u}},\xi_{\bar{v}}) \in \mathcal{U}^* \times \mathcal{V}^* \ \middle|\ \xi_{\bar{v}} \in \partial\left\|\bar{v}\right\|^2 \text{ and } \xi_{\bar{u}} \in \eta\partial\mathcal{R}\big(\bar{u}\big)\right\} \end{aligned}$$

□

For the convergence rate analysis, we need the following result:

Lemma 4.3.8. *Let $B : \mathcal{U} \times \mathcal{V} \to \mathcal{H}$ be a bilinear operator with $\|B(k,f)\| \leq C\left\|k\right\|\left\|f\right\|$. Then its Fréchet derivative at point (k,f) is given by*

$$B'(k,f)(u,v) = B(u,f) + B(k,v),$$

$(u,v) \in \mathcal{U} \times \mathcal{V}$. Moreover, the remainder of the Taylor expansion can be estimated by

$$\left\|B(k+u,f+v) - B(k,f) - B'(k,f)(u,v)\right\| \leq \frac{C}{2}\big\|(u,v)\big\|^2 . \tag{4.17}$$

Proof. The proof is straightforward and follows from the bilinearity of the operator and its boundedness. □

The following theorem gives an error estimate within an infinite dimensional setting, similar to the results found in [57, 88]. Please note that we have not only an error estimate for the solution f, but also for the characterising function k, i.e., we are able to derive convergence rate for the operator via (4.3).

Theorem 4.3.9 (Convergence rates). *Let $g_\delta \in \mathcal{H}$ with $\|g_0 - g_\delta\| \leq \delta$, $k_\epsilon \in \mathcal{U}$ with $\|k_0 - k_\epsilon\| \leq \epsilon$ and let $f^\dagger$ be a minimum norm solution. For the regularisation parameter $0 < \alpha < \infty$, let (k^α, f^α) denote the minimiser of (4.13) with $L = I$. Moreover, assume that the following conditions hold:*

(i) There exists $\omega \in \mathcal{H}$ satisfying

$$(\xi_{k_0}, \xi_{f^\dagger}) = B'(k_0, f^\dagger)^* \omega,$$

with $(\xi_{k_0}, \xi_{f^\dagger}) \in \partial\Phi\big(k_0, f^\dagger\big)$.

(ii) $C\|\omega\|_{\mathcal{H}} < \min\left\{1, \frac{\gamma}{2\alpha}\right\}$, where C is the constant in (4.17).

Then, for the parameter choice $\alpha \sim (\delta + \epsilon)$ holds

$$\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|_{\mathcal{H}} = \mathcal{O}\left(\delta + \epsilon\right)$$

and

$$D_\Phi^\xi\big((k^\alpha, f^\alpha), (k_0, f^\dagger)\big) = \mathcal{O}\left(\delta + \epsilon\right).$$

Proof. Since (k^α, f^α) is a minimiser of J_α, defined in (4.13), it follows

$$J_\alpha\big(k^\alpha, f^\alpha\big) \leq J_\alpha\big(k, f\big) \quad \forall (k, f) \in \mathcal{U} \times \mathcal{V}.$$

In particular,

$$\begin{aligned} J_\alpha\big(k^\alpha, f^\alpha\big) &\leq J_\alpha\big(k_0, f^\dagger\big) \\ &\leq \frac{\delta^2}{2} + \frac{\gamma\epsilon^2}{2} + \alpha\Phi(k_0, f^\dagger). \end{aligned} \tag{4.18}$$

Using the definition of the Bregman distance (at the subgradient $(\xi_{k_0}, \xi_{f^\dagger}) \in \partial\Phi\big(k_0, f^\dagger\big)$), we rewrite (4.18) as

$$\begin{aligned} &\frac{1}{2}\left\|B(k^\alpha, f^\alpha) - g_\delta\right\|^2 + \frac{\gamma}{2}\left\|k^\alpha - k_\epsilon\right\|^2 \qquad (4.19) \\ &\leq \frac{\delta^2 + \gamma\epsilon^2}{2} + \alpha\left(\Phi(k_0, f^\dagger) - \Phi(k^\alpha, f^\alpha)\right) \\ &= \frac{\delta^2 + \gamma\epsilon^2}{2} - \alpha\left[D_\Phi^{\xi^\dagger}\left((k^\alpha, f^\alpha), (k_0, f^\dagger)\right) + \left\langle (\xi_{k_0}, \xi_{f^\dagger}) \,,\, (k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\rangle\right]. \end{aligned}$$

Using

$$\begin{aligned} \frac{1}{2}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 &\leq \left\|B(k^\alpha, f^\alpha) - g_\delta\right\|^2 + \left\|g_\delta - g_0\right\|^2 \\ &\leq \left\|B(k^\alpha, f^\alpha) - g_\delta\right\|^2 + \delta^2 \end{aligned}$$

and

$$\begin{aligned}\frac{\gamma}{2}\left\|k^\alpha - k_0\right\|^2 &\leq \gamma\left\|k^\alpha - k_\epsilon\right\|^2 + \gamma\left\|k_\epsilon - k_0\right\|^2 \\ &\leq \gamma\left\|k^\alpha - k_\epsilon\right\|^2 + \gamma\epsilon^2,\end{aligned}$$

we get

$$\begin{aligned}&\frac{1}{4}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 + \frac{\gamma}{4}\left\|k^\alpha - k_0\right\|^2 \\ &\quad\leq \quad \frac{1}{2}\left\|B(k^\alpha, f^\alpha) - g_\delta\right\|^2 + \frac{\gamma}{2}\left\|k^\alpha - k_\epsilon\right\|^2 + \left(\frac{\delta^2 + \gamma\epsilon^2}{2}\right) \\ &\quad\overset{(4.19)}{\leq} \quad (\delta^2 + \gamma\epsilon^2) - \alpha\left[D_\Phi^{\xi^\dagger}\left((k^\alpha, f^\alpha), (k_0, f^\dagger)\right) + \left\langle(\xi_{k_0}, \xi_{f^\dagger})\ ,\ (k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\rangle\right].\end{aligned}$$

Denoting $r := B(k^\alpha, f^\alpha) - B(k_0, f^\dagger) - B'(k_0, f^\dagger)((k^\alpha, f^\alpha) - (k_0, f^\dagger))$ and using the source condition *(i)*, the last term in the above inequality can be estimated as

$$\begin{aligned}&-\left\langle(\xi_{k_0}, \xi_{f^\dagger})\ ,\ (k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\rangle \\ &\quad= \quad -\left\langle B'(k_0, f^\dagger)^*\omega\ ,\ (k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\rangle \\ &\quad= \quad \left\langle\omega\ ,\ -B'(k_0, f^\dagger)\left((k^\alpha, f^\alpha) - (k_0, f^\dagger)\right)\right\rangle \\ &\quad= \quad \left\langle\omega\ ,\ B(k_0, f^\dagger) - B(k^\alpha, f^\alpha) + r\right\rangle \\ &\quad\leq \quad \|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| + \|\omega\|\,\|r\| \\ &\quad\overset{(4.17)}{\leq} \quad \|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| + \frac{C}{2}\|\omega\|\left\|(k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\|^2.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}&\frac{1}{4}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 + \frac{\gamma}{4}\left\|k^\alpha - k_0\right\|^2 + \alpha D_\Phi^{\xi^\dagger}\left((k^\alpha, f^\alpha), (k_0, f^\dagger)\right) \qquad (4.20) \\ &\quad\leq \quad (\delta^2 + \gamma\epsilon^2) + \alpha\|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| + \alpha\tfrac{C}{2}\|\omega\|\left\|(k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\|^2.\end{aligned}$$

Using (4.16) and the definition of the norm on $\mathcal{U}\times\mathcal{V}$, (4.20) can be rewritten as

$$\begin{aligned}&\frac{1}{4}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 + \frac{\alpha}{2}\left(1 - C\|\omega\|\right)\left\|f^\alpha - f^\dagger\right\|^2 + \alpha\eta D_{\mathcal{R}}^{\zeta}\left(k^\alpha, k_0\right) \\ &\quad\leq \quad (\delta^2 + \gamma\epsilon^2) + \alpha\|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| + \tfrac{1}{2}\left(\alpha C\|\omega\| - \tfrac{\gamma}{2}\right)\left\|k^\alpha - k_0\right\|^2 \\ &\quad\leq \quad \left(\delta^2 + \gamma\epsilon^2\right) + \alpha\|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|, \qquad (4.21)\end{aligned}$$

as $\left(C\|\omega\| - \frac{\gamma}{2\alpha}\right) \leq 0$ according to *(ii)*. As $(1 - C\|\omega\|)$ as well as the Bregman distance are non-negative, we derive

$$\frac{1}{4}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 - \alpha \left\|\omega\right\| \left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| - \left(\delta^2 + \gamma\epsilon^2\right) \leq 0,$$

which only holds for

$$\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| \leq 2\alpha \left\|\omega\right\| + 2\sqrt{\alpha^2 \left\|\omega\right\|^2 + (\delta^2 + \gamma\epsilon^2)}.$$

Using the above inequality to estimate the right-hand side of (4.21) yields

$$\left\|f^\alpha - f^\dagger\right\|^2 \leq \frac{2}{1 - C\left\|\omega\right\|}\left(\frac{\delta^2 + \gamma\epsilon^2}{\alpha} + 2\alpha\left\|\omega\right\|^2 + 2\left\|\omega\right\|\sqrt{\alpha^2\left\|\omega\right\|^2 + (\delta^2 + \gamma\epsilon^2)}\right)$$

and

$$D_{\mathfrak{K}}^{\zeta}\left(k^\alpha, k_0\right) \leq \frac{\delta^2 + \gamma\epsilon^2}{\eta\alpha} + \frac{2\left\|\omega\right\|}{\eta}\left(\alpha\left\|\omega\right\| + \sqrt{\alpha^2\left\|\omega\right\|^2 + (\delta^2 + \gamma\epsilon^2)}\right),$$

and for the parameter choice $\alpha \sim (\delta + \epsilon)$ follows the convergence rate $\mathcal{O}(\delta + \epsilon)$. □

Remark 4.3.10. *The assumptions of Theorem 4.3.9 include the condition*

$$C\left\|\omega\right\|_{\mathcal{H}} < \min\left\{1, \frac{\gamma}{2\alpha}\right\} .$$

Note that $\frac{\gamma}{(2\alpha)} < 1$ for α small enough (i.e., for small noise level δ and ϵ), and thus (ii) reduces to the standard smallness assumption common for convergence rates for nonlinear ill-posed problems, see [28].

4.4 Numerical Example

In order to illustrate our analytical results we present first reconstructions from a convolution operator. That is, the kernel function is defined by $k_0(s,t) := k_0(s-t)$ over $\Omega = [0,1]$, see also Example 2 in Section 4.1, and we want to solve the integral equation

$$\int_\Omega k_0(s-t)f(t)dt = g_0(s)$$

from given measurements k_ϵ and g_δ satisfying (4.6). For our test, we defined k_0 and f as

$$k_0 = \begin{cases} 1 & x \in [0.1, 0.4] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f = \begin{cases} 1 - 5|t - 0.3| & t \in [0.1, 0.5] \\ 0 & \text{otherwise} \end{cases},$$

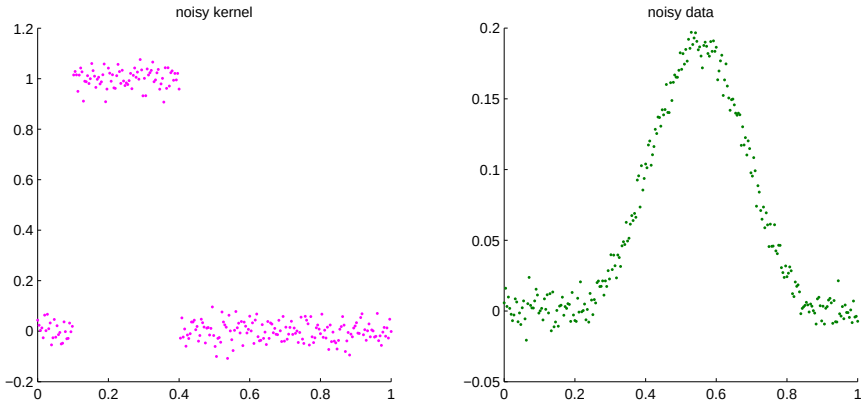

Figure 4.1: Simulated measurements for k_0 (left) and g_0 (right), both with 10% relative error.

respectively, the characteristic and the hat function. An example of noisy measurements k_ϵ and g_δ is displayed in Figure 4.1.

The functions k and f were expanded in a wavelet basis, as for example,

$$k = \sum_{l\in\mathbb{Z}} \langle k \ , \ \phi_{0,l}\rangle \phi_{0,l} + \sum_{j=0}^{\infty}\sum_{l\in\mathbb{Z}} \langle k \ , \ \psi_{j,l}\rangle \psi_{j,l} \ ,$$

where $\{\phi_\lambda\}_\lambda$ and $\{\psi_\lambda\}_\lambda$ are the pair of scaling and wavelet function associated to Haar wavelet basis. The convolution operator was implemented in terms of the wavelet basis as well. For our numerical tests, we used the Haar wavelet. The integration interval $\Omega = [0,1]$ was discretized into $N = 2^8$ points, the maximum level considered by the Haar wavelet is $J = 6$. The functional $\mathcal{R}$ was defined as

$$\mathcal{R}(k) := \|k\|_{\ell_1} = \sum_{\lambda\in\Lambda} |\langle k \ , \ \psi_\lambda\rangle|,$$

where $\Lambda = \{\{l\} \cup (j,l) \ | \ j \in \mathbb{N}_0, l \leq 2^j - 1\}$.

In order to find the optimal set of coefficients minimising (4.7) we used Matlab's internal function `fminsearch`.

Figure 5.4 displays the numerical solutions for three different (relative) error levels: 10%, 5% and 1%. The scaling parameter was set to $\gamma = 1$ and the regularisation parameters are chosen according to the noise level, i.e., $\alpha = 0.01(\delta+\varepsilon)$ and $\beta = 0.2(\delta+\varepsilon)$, $(\eta = 20)$ was chosen. Our numerical results confirm our analysis. In particular it is observed that the reconstruction quality increases with decreasing noise level, see also Table 4.1.

Please note that the optimisation with the `fminsearch` routine is by no means efficient. In the upcoming chapter we shall propose a fast iterative optimisation routine for the minimisation of (4.7).

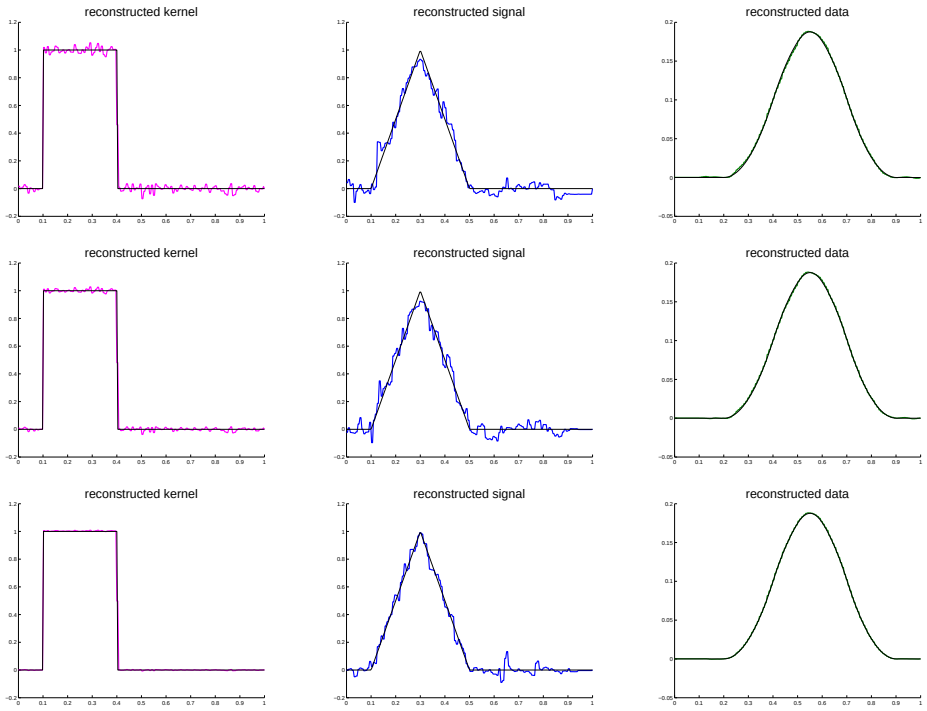

Figure 4.2: Reconstruction of the characterising function k_0, the signal f (solution) and the data g_0. From top to bottom: reconstruction with 10%, 5% and 1% relative error (both for g_δ and k_ϵ). The reconstructions are colored.

	$\left\|k^{\text{rec}} - k_0\right\|_1$	$\left\|f^{\text{rec}} - f^{\text{true}}\right\|_1$	$\left\|k^{\text{rec}} - k_0\right\|_2$	$\left\|f^{\text{rec}} - f^{\text{true}}\right\|_2$
10%	6.7543e-02	1.8733e-01	8.1216e-03	1.7436e-02
5%	4.0605e-02	1.7173e-01	6.9089e-03	1.5719e-02
1%	2.0139e-02	1.1345e-01	6.5219e-03	8.0168e-03

Table 4.1: Relative error measured by the L_1- and L_2-norm.

Chapter 5

An Alternating Minimisation Algorithm

> "The essence of mathematics is not to make simple things complicated, but to make complicated things simple."
>
> *Stanley Gudder*

The task of solving the optimisation problem proposed previously seems, at first sight, not a trivial task, since the regarded functional is most likely nonconvex and nonlinear. Therefore our main focus is on efficient numerical implementation with particular emphasis on alternating minimisation strategy. It solves not only the dbl-RTLS, but a vast class of optimisation problems: on the minimisation of a bilinear functional over two variables.

Initially we determine the optimality condition for the underlying problem. Subsequently we develop an algorithm based on an alternating minimisation strategy and we study its convergence properties. Finally, some numerical examples for the proposed algorithm are provided and the efficiency of the method is discussed.

5.1 Optimality Condition

The first-order necessary condition for critical points of the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ given in (4.7) requires in particular the derivative of the bilinear operator B.

It is well known that the study of local behaviour of nonsmooth functions can be achieved handled by the concept of *sub-differentiality* which replaces the classical derivative at non-differentiable points.

Therefore the first-order necessary condition based on sub-differentiability is stated as the following: if $(\bar{k}, \bar{f})$ minimises the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ then

$$(0,0) \in \partial J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big). \tag{5.1}$$

We denote the set of all sub-derivatives of the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ at (k,f) by $\partial J^{\delta,\varepsilon}_{\alpha,\beta}\big(k,f\big)$ and we name it the *sub-differential* of $J^{\delta,\varepsilon}_{\alpha,\beta}$ at (k,f). For a quick revision on sub-differentiability we refer to Appendix B.2.

The first result gives us the derivative of a bilinear operator B.

Lemma 5.1.1. *Let B be a bilinear operator and assume that (4.1) holds. Then the Fréchet derivative of B at $(k,f) \in \mathcal{U} \times \mathcal{V}$ is given by*

$$\begin{aligned} B'\big(k,f\big)(u,v) &= B(k,v) + B(u,f) \\ &= A_k v + C_f u. \end{aligned}$$

Moreover, the derivative is Lipschitz continuous with constant $\sqrt{2}C$.

Proof. We have to show

$$B(k+u,f+v) = B(k,f) + B'\big(k,f\big)(u,v) + o\left(\|(u,v)\|\right).$$

Since B is bilinear, we have

$$B(k+u,f+v) - B(k,f) = B(k,v) + B(u,f) + B(u,v),$$

and we observe $\|B(u,v)\| = o\left(\|(u,v)\|\right)$: As B fulfils (4.1), we have

$$\frac{\|B(u,v)\|}{\|(u,v)\|} \leq \frac{C\,\|u\|\,\|v\|}{(\|u\|^2 + \|v\|^2)^{1/2}} \leq \frac{C}{\sqrt{2}}(\|u\|\,\|u\|)^{1/2},$$

which converges to zero as $(u,v) \to 0$.

We further observe

$$\begin{aligned} B'\big(k,f\big)(u,v) - B'\big(\tilde{k},\tilde{f}\big)(u,v) &= B(k,v) + B(u,f) - \Big(B(\tilde{k},v) + B(u,\tilde{f})\Big) \\ &= B(u,f-\tilde{f}) + B(k-\tilde{k},v) \end{aligned}$$

which implies

$$\begin{aligned} \big\|B'\big(k,f\big)(u,v) - B'\big(\tilde{k},\tilde{f}\big)(u,v)\big\| &= \big\|B(u,f-\tilde{f}) + B(k-\tilde{k},v)\big\| \\ &\leq \big\|B(u,f-\tilde{f})\big\| + \big\|B(k-\tilde{k},v)\big\| \\ &\leq C\|u\|\|f-\tilde{f}\| + C\|k-\tilde{k}\|\|v\| \end{aligned}$$

Using the inequality $(a+b)^2 \leq 2(a^2+b^2)$ we get

$$\begin{aligned} \big\|B'\big(k,f\big)(u,v) - B'\big(\tilde{k},\tilde{f}\big)(u,v)\big\|^2 &\leq 2C^2\left(\|u\|^2\|f-\tilde{f}\|^2 + \|k-\tilde{k}\|^2\|v\|^2\right) \\ &\leq 2C^2\left(\|u\|^2 + \|v\|^2\right)\left(\|k-\tilde{k}\|^2 + \|f-\tilde{f}\|^2\right) \\ &= 2C^2\|(u,v)\|^2\|(k-\tilde{k},f-\tilde{f})\|^2 \end{aligned}$$

and thus

$$\begin{aligned}\big\|B'(k,f) - B'(\tilde{k},\tilde{f})\big\| &= \sup_{\|(u,v)\|=1} \big\|B'\big(k,f\big)(u,v) - B'\big(\tilde{k},\tilde{f}\big)(u,v)\big\| \\ &\leq \sqrt{2}C\big\|(k-\tilde{k}, f-\tilde{f})\big\|.\end{aligned}$$

□

Note that the adjoint operator $(B'(k,f))^*$ of the Frechét derivative $B'(k,f)$ exists and is a bounded linear operator whenever both $\mathcal{H}$ and $\mathcal{U}\times\mathcal{V}$ are Hilbert spaces.

In order to analyse the optimality condition (5.1) we shall compute the sub-differential of a functional over two variables. As pointed out in the book [19, Proposition 2.3.15] for a general function h the set-valued mapping $\partial h : \mathcal{U} \rightrightarrows \mathcal{U}^*$ the set $\partial h(x_1,x_2)$ and the product set $\partial_1 h(x_1,x_2) \times \partial_2 h(x_1,x_2)$ are not necessarily contained in each other. Here, $\partial_i h$ denotes the partial sub-gradient with respect to x_i for $i = 1,2$. However this is not the case for the functional we are interested in as will be shown in the following Theorem.

Theorem 5.1.2. *Let $J : \mathcal{U}\times\mathcal{V} \to \overline{\mathbb{R}}$ be a functional with the structure*

$$J(u,v) = \varphi(u) + Q(u,v) + \psi(v), \tag{5.2}$$

where Q is a (nonlinear) differentiable term and $\varphi : \mathcal{U} \to \overline{\mathbb{R}}$, $\psi : \mathcal{V} \to \overline{\mathbb{R}}$ are proper convex functions, $u \in \operatorname{dom}\varphi$ and $v \in \operatorname{dom}\psi$. Then

$$\begin{aligned}\partial J(u,v) &= \{\partial\varphi\big(u\big) + Q'_u(u,v)\} \times \{\partial\psi\big(v\big) + Q'_v(u,v)\} \\ &= \{\partial_u J(u,v)\} \times \{\partial_v J(u,v)\}.\end{aligned}$$

Proof. In general the sub-differential of a sum of functions does not equal the sum of its sub-differentials. However, if Q is differentiable, φ and ψ are convex some inclusions and even equalities hold true (combining [19, Prop 2.3.3; Cor 3; Prop 2.3.6]), as for instance,

$$\partial J(u,v) = \partial\left(\varphi(u) + \psi(v)\right) + \partial Q(u,v).$$

Since Q is differentiable, calling the previous results, the (partial) sub-derivative is unique [19, Prop 2.3.15] and therefore

$$\begin{aligned}\partial Q(u,v) &= \partial_u Q(u,v) \times \partial_v Q(u,v) \\ &= \left(Q'_u(u,v), Q'_v(u,v)\right).\end{aligned}$$

Note that for the special case where the functional $\psi(u) + \varphi(v)$, the sub-derivative of separable convex functions [99, Corollary 2.4.5] satisfies

$$\partial\left(\varphi(u) + \psi(v)\right) = \big(\partial\varphi\big(u\big), \partial\psi\big(v\big)\big)$$

Altogether, we can compute the sub-derivative as follows

$$\begin{aligned} \partial J(u,v) &= \big(\partial\varphi(u), \partial\psi(v)\big) + (Q'_u(u,v), Q'_v(u,v)) \\ &= \{\partial_u\varphi(u) + Q'_u(u,v)\} \times \{\partial_v\psi(v) + Q'_v(u,v)\}. \end{aligned} \tag{5.3}$$

The last implication of this theorem,

$$\partial J(u,v) = \{\partial_u J(u,v)\} \times \{\partial_v J(u,v)\}$$

follows straightforward by definition of partial sub-derivative and (5.3). □

Please note that the above proof holds for all definitions of sub-differential introduced in the Appendix B.2, as for convex functionals all the definitions are equivalent, and for differentiable (possibly nonlinear) terms the sub-differential is a unitary set and the sub-derivative equals the derivative. Based on Theorem 5.1.2 we can now calculate the derivative of the functional is the gist for building up the upcoming algorithm; please give heed to the structure of (5.2) and the proposed functional $J^{\delta,\varepsilon}_{\alpha,\beta}$:

Corollary 5.1.3. *Let $J^{\delta,\varepsilon}_{\alpha,\beta}$ the functional defined in (4.7), then*

$$\partial J^{\delta,\varepsilon}_{\alpha,\beta}\big(k,f\big) = \{C^*_f(C_f k - g_\delta) + \gamma(k - k_\epsilon) + \beta\zeta\} \times \{A^*_k(A_k f - g_\delta) + \alpha L^* L f\}$$

where $\zeta \in \partial\mathcal{R}(k)$.

Proof. The result follows straightforward from Lemma 5.1.1 and Theorem 5.1.2. Observe that the sum $C^*_f(C_f k - g_\delta) + \gamma(k - k_\epsilon) + \beta\zeta$ is well-defined in the Hilbert space $\mathcal{U}$, since the sub-derivative $\zeta\partial\mathcal{R}(k)$ is also an element of $\mathcal{U}$. □

Up to now, we did not specify the functional $\mathcal{R}$, it is only required to be convex and lower semi-continuous. We are in particular interested in, e.g., the L_p norm or the weighted ℓ_p norm, denoted by $\mathcal{R}(k) = \big\|k\big\|_{w,p}$. Its sub-differential is given in Section 5.3. An easy way to compute the sub-derivatives of functionals $\mathcal{R}$ with a specific structure is given by the following Lemma.

Lemma 5.1.4 ([10, Lemma 4.4])**.** *Let $\mathcal{H} = L_2(\Omega, d\mu)$ where Ω is a σ-finite measure space. Let $\mathcal{R} : \mathcal{H} \to (-\infty, +\infty]$ be defined by*

$$\mathcal{R}(u) = \begin{cases} \int_\Omega h(u)d\mu & \text{if the integral is finite} \\ \infty & \text{else,} \end{cases} \tag{5.4}$$

where $h : \mathbb{C} \to \mathbb{R}$ is a convex function. Then $\xi \in \mathcal{H}$ is an element of $\partial\mathcal{R}\big(u\big)$ if and only if $\xi(x) \in \partial h\big(u(x)\big)$ for almost every $x \in \Omega$ (with the identification $\mathbb{C}^2 = \mathbb{R}$).

5.2 Proposed Algorithm

The computation of a solution of dbl-RTLS is not straightforward, as the minimum of the functional (4.7) with respect to both parameters is a nonlinear and nonconvex problem over two variables. Nevertheless, there is a simple algorithm that has been successfully used for optimisation problems over two variables: *alternating minimisation* (AM). This procedure has been studied by several authors, see, e.g., [14, 98, 96].

In the following we shall denote the dbl-RTLS functional by J instead of $J^{\delta,\varepsilon}_{\alpha,\beta}$, as the parameters of the functionals are kept fix for the minimisation process.

In the AM algorithm, the functional is minimised iteratively with two alternating minimisation steps. Each step minimises the problem over one variable while keeping the second variable fixed:

$$f^{n+1} \in \operatorname*{arg\,min}_{f \in V} J(k, f|k^n) \tag{5.5a}$$

$$k^{n+1} \in \operatorname*{arg\,min}_{k \in U} J(k, f|f^{n+1}). \tag{5.5b}$$

The notation $J(k, f|u)$ means we minimise the function J with u fixed, where u can be either k or f. Thus we minimise in each cycle the functionals

$$J(k, f|k^n) = \left\|A_{k^n} f - g_\delta\right\|^2 + \alpha \left\|Lf\right\|^2,$$

and

$$J(k, f|f^{n+1}) = \left\|C_{f^{n+1}} k - g_\delta\right\|^2 + \gamma \left\|k - k_\epsilon\right\|^2 + \beta \mathcal{R}(k).$$

We highlight some important facts:

1. For each subproblem, the considered operators are linear, and the functional is convex. Thus a local minimum is global.

2. The first step is a standard quadratic minimisation problem.

First we will show a monotonicity result for the sequence $\{(k^n, f^n)\}_n$ of iterates:

Proposition 5.2.1. *The functional J is non-increasing on the AM iterates,*

$$J(k^{n+1}, f^{n+1}) \leq J(k^n, f^{n+1}) \leq J(k^n, f^n).$$

Proof. The iterates are defined as

$$f^{n+1} \in \operatorname*{arg\,min}_{f \in V} J(k, f|k^n)$$

and

$$k^{n+1} \in \arg\min_{k \in U} J(k, f|f^{n+1}).$$

Therefore,

$$J(k^n, f^{n+1}) \leq J(k^n, f) \qquad \forall f \in V$$

and

$$J(k^{n+1}, f^{n+1}) \leq J(k, f^{n+1}) \qquad \forall k \in U,$$

and in particular, setting $f = f^n$ and $k = k^n$,

$$\begin{aligned} J(k^n, f^{n+1}) &\leq J(k^n, f^n) \\ J(k^{n+1}, f^{n+1}) &\leq J(k^n, f^{n+1}), \end{aligned}$$

and

$$J(k^{n+1}, f^{n+1}) \leq J(k^n, f^{n+1}) \leq J(k^n, f^n).$$

□

The existence of minimiser of the the functional J has already been proven in [6, Thm 4.2]. The goal of the following results is to prove that the sequence generated by the alternating minimisation algorithm has at least a subsequence which converges towards to a critical point of the functional. Throughout this Section, let us take the following assumptions.

Assumption E.

(E1) B is strongly continuous, i.e., if $(k^n, f^n) \rightharpoonup (\bar{k}, \bar{f})$ then $B(k^n, f^n) \to B(\bar{k}, \bar{f})$.

(E2) The adjoint of the Fréchet derivative B' of B is strongly continuous, i.e., if $(k^n, f^n) \rightharpoonup (\bar{k}, \bar{f})$ then $B'(k^n, f^n)^* z \to B'(\bar{k}, \bar{f})^* z$, $\forall z \in \mathscr{D}(B')$

Additionally to the standard norm for the pair $(k, f) \in \mathcal{U} \times \mathcal{V}$

$$\|(k, f)\|^2 = \|k\|^2 + \|f\|^2$$

we define the weighted norm for given $\gamma > 0$ as

$$\|(k, f)\|_\gamma^2 = \gamma\|k\|^2 + \|f\|^2.$$

Proposition 5.2.2. *For given regularisation parameters $0 < \alpha$ and β, the sequence $\{(k^{n+1}, f^{n+1})\}_{n+1}$ of iterates generated by the AM algorithm has at least a weakly convergent subsequence $(k^{n_j+1}, f^{n_j+1}) \rightharpoonup (\bar{k}, \bar{f})$, and its limit fulfils*

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, f) \quad \text{and} \quad J(\bar{k}, \bar{f}) \leq J(k, \bar{f}) \tag{5.6}$$

for all $f \in \mathcal{V}$ and for all $k \in \mathcal{U}$.

Proof. As the iterates of the AM algorithm can be characterised as the minimisers of a reduced dbl-RTLS functional, see (5.5a), (5.5b) we observe

$$\begin{aligned}\alpha\left\|Lf^{n+1}\right\|^2+\gamma\left\|k^n-k_\epsilon\right\|^2+\beta\mathcal{R}(k^n) &\le J(k^n,f^{n+1})\\ &=\min_f J(k,f|k^n)\\ &\le J(k^n,0)\\ &=\left\|g_\delta\right\|^2+\gamma\left\|k^n-k_\epsilon\right\|^2+\beta\mathcal{R}(k^n)\end{aligned}$$

and

$$\begin{aligned}\alpha\left\|Lf^{n+1}\right\|^2+\gamma\left\|k^{n+1}-k_\epsilon\right\|^2 &\le J(k^{n+1},f^{n+1})\\ &=\min_k J(k,f|f^{n+1})\\ &\le J(0,f^{n+1})\\ &=\left\|g_\delta\right\|^2+\gamma\left\|k_\epsilon\right\|^2+\alpha\left\|Lf^{n+1}\right\|^2.\end{aligned}$$

Keeping in mind that the operator L is continuously invertible, the first inequality gives

$$\left\|f^{n+1}\right\|^2\le\frac{1}{\left\|L^{-1}\right\|^2\alpha}\left\|g_\delta\right\|^2.$$

Using the second estimate above and the standard inequality $\|a+b\|^2\le 2(\|a\|^2+\|b\|^2)$ we have

$$\gamma\left\|k^{n+1}\right\|^2\le 2\left\|g_\delta\right\|^2+4\gamma\left\|k_\epsilon\right\|^2.$$

Thus, the sequence $\{(k^{n+1},f^{n+1})\}_{n+1}$ is bounded

$$\begin{aligned}\left\|(k^{n+1},f^{n+1})\right\|_\gamma^2&=\gamma\left\|k^{n+1}\right\|^2+\left\|f^{n+1}\right\|^2\\ &\le 2\left\|g_\delta\right\|^2+4\gamma\left\|k_\epsilon\right\|^2+\frac{1}{c^2\alpha}\left\|g_\delta\right\|^2\\ &=\left(2+\frac{1}{\left\|L^{-1}\right\|^2\alpha}\right)\left\|g_\delta\right\|^2+4\gamma\left\|k_\epsilon\right\|^2\end{aligned}$$

and by Alaoglu's theorem, it has a weakly convergent subsequence $\{(k^{n_j+1},f^{n_j+1})\}_{n_j+1}\rightharpoonup(\bar{k},\bar{f})$.

Since f^{n_j+1} minimises the functional $J(k^{n_j},f)$ for a fixed k^{n_j}, it holds

$$J(k^{n_j},f^{n_j+1})\le J(k^{n_j},f)\quad\forall f\in\mathcal{V}$$

and thus

$$\left\|B(k^{n_j},f^{n_j+1})-g_\delta\right\|^2+\alpha\left\|Lf^{n_j+1}\right\|^2\le\left\|B(k^{n_j},f)-g_\delta\right\|^2+\alpha\left\|Lf\right\|^2.$$

Using the fact that J is w-lsc and the strong continuity of B, we observe

$$\begin{aligned}
&\big\|B(\bar{k},\bar{f})-g_\delta\big\|^2+\alpha\big\|L\bar{f}\big\|^2\\
&\qquad\le \liminf_{n_j\to\infty}\Big\{\big\|B(k^{n_j+1},f^{n_j+1})-g_\delta\big\|^2+\alpha\big\|Lf^{n_j+1}\big\|^2\Big\}\\
&\qquad\le \liminf_{n_j\to\infty}\Big\{\big\|B(k^{n_j},f^{n_j+1})-g_\delta\big\|^2+\alpha\big\|Lf^{n_j+1}\big\|^2\Big\}\\
&\qquad\le \liminf_{n_j\to\infty}\Big\{\big\|B(k^{n_j},f)-g_\delta\big\|^2+\alpha\big\|Lf\big\|^2\Big\}\\
&\qquad\le \limsup_{n_j\to\infty}\big\|B(k^{n_j},f)-g_\delta\big\|^2+\alpha\big\|Lf\big\|^2\\
&\qquad= \lim_{n_j\to\infty}\big\|B(k^{n_j},f)-g_\delta\big\|^2+\alpha\big\|Lf\big\|^2\\
&\qquad\overset{(E1)}{=} \big\|B(\bar{k},f)-g_\delta\big\|^2+\alpha\big\|Lf\big\|^2 \qquad (5.7)
\end{aligned}$$

Therefore,

$$J(\bar{k},\bar{f})\le J(\bar{k},f)\quad\forall f\in\mathcal{V}.$$

The second inequality in (5.6) is proven similarly: Since k^{n_j+1} minimises the functional $J(k,f^{n_j+1})$ for fixed f^{n_j+1} it is

$$J(k^{n_j+1},f^{n_j+1})\le J(k,f^{n_j+1})\quad\forall k\in\mathcal{U},$$

which is equivalent to

$$\begin{aligned}
&\big\|B(k^{n_j+1},f^{n_j+1})-g_\delta\big\|^2+\gamma\big\|k^{n_j+1}-k_\epsilon\big\|^2+\beta\mathcal{R}(k^{n_j+1})\\
&\qquad\le\big\|B(k,f^{n_j+1})-g_\delta\big\|^2+\gamma\big\|k-k_\epsilon\big\|^2+\beta\mathcal{R}(k).
\end{aligned}$$

Again, we observe

$$\begin{aligned}
&\big\|B(\bar{k},\bar{f})-g_\delta\big\|^2+\gamma\big\|\bar{k}-k_\epsilon\big\|^2+\beta\mathcal{R}(\bar{k})\\
&\qquad\le \liminf_{n_j\to\infty}\Big\{\big\|B(k^{n_j+1},f^{n_j+1})-g_\delta\big\|^2+\gamma\big\|k^{n_j+1}-k_\epsilon\big\|^2+\beta\mathcal{R}(k^{n_j+1})\Big\}\\
&\qquad\le \liminf_{n_j\to\infty}\big\|B(k,f^{n_j+1})-g_\delta\big\|^2+\gamma\big\|k-k_\epsilon\big\|^2+\beta\mathcal{R}(k)\\
&\qquad= \lim_{n_j\to\infty}\big\|B(k,f^{n_j+1})-g_\delta\big\|^2+\gamma\big\|k-k_\epsilon\big\|^2+\beta\mathcal{R}(k)\\
&\qquad= \big\|B(k,\bar{f})-g_\delta\big\|^2+\gamma\big\|k-k_\epsilon\big\|^2+\beta\mathcal{R}(k), \qquad (5.8)
\end{aligned}$$

and thus

$$J(\bar{k},\bar{f})\le J(k,\bar{f}),\quad\forall k\in\mathcal{U}.$$

□

In summary, the alternating minimisation (AM) algorithm yields a bounded sequence $\{(k^{n+1},f^{n+1})\}_n$ and hence a weakly convergent subsequence. The next result extends the convergence on the strong topology, for both $\{k^{n_j+1}\}_{n_j}$ and $\{f^{n_j+1}\}_{n_j}$.

Proposition 5.2.3. *Let $\{(k^{n_j+1}, f^{n_j+1})\}_{n_j}$ be a weakly convergent (sub-) sequence generated by the AM algorithm (5.5), where $k^{n_j+1} \rightharpoonup \bar{k}$ and $f^{n_j+1} \rightharpoonup \bar{f}$. Then there exists a subsequence $\{k^{n_{j_m}+1}\}_{n_{j_m}}$ of $\{k^{n_j+1}\}_{n_j}$ such that $k^{n_{j_m}+1} \to \bar{k}$ and $0 \in \partial_k J(\bar{k}, \bar{f})$.*

Proof. Inequalities (5.8) in the Proposition 5.2.2's proof reads

$$\begin{aligned}\liminf_{n_j\to\infty} \left\{ \left\|B(k^{n_j+1}, f^{n_j+1}) - g_\delta\right\|^2 + \gamma\left\|k^{n_j+1} - k_\epsilon\right\|^2 + \beta\mathcal{R}(k^{n_j+1})\right\} \\ = \left\|B(k, \bar{f}) - g_\delta\right\|^2 + \gamma\left\|k - k_\epsilon\right\|^2 + \beta\mathcal{R}(k).\end{aligned}$$

for any k. Setting $k = \bar{k}$ yields in particular

$$\begin{aligned}\liminf_{n_j\to\infty} \left\{ \left\|B(k^{n_j+1}, f^{n_j+1}) - g_\delta\right\|^2 + \gamma\left\|k^{n_j+1} - k_\epsilon\right\|^2 + \beta\mathcal{R}(k^{n_j+1})\right\} \\ = \left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \gamma\left\|\bar{k} - k_\epsilon\right\|^2 + \beta\mathcal{R}(\bar{k}).\end{aligned}$$

As the limes inferior exists, we can in particular extract a subsequence $(k^{n_{j_m}+1}, f^{n_{j_m}+1})_{n_{j_m}}$ of $(k^{n_j+1}, f^{n_j+1})_{n_j}$ such that

$$\begin{aligned}\lim_{n_{j_m}\to\infty} \left\{ \left\|B(k^{n_{j_m}+1}, f^{n_{j_m}+1}) - g_\delta\right\|^2 + \gamma\left\|k^{n_{j_m}+1} - k_\epsilon\right\|^2 + \beta\mathcal{R}(k^{n_{j_m}+1})\right\} \\ = \left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \gamma\left\|\bar{k} - k_\epsilon\right\|^2 + \beta\mathcal{R}(\bar{k}).\end{aligned} \tag{5.9}$$

For the sake of notation simplicity we denote for the remainder of the proof the index $n_{j_m} + 1$ by $m + 1$. By (E1) we observe

$$\lim_{m\to\infty} \left\|B(k^{m+1}, f^{m+1}) - g_\delta\right\|^2 \overset{(E1)}{=} \left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2$$

As all summands in (5.9) are positive, we have thus and

$$\begin{aligned}\lim_{m\to\infty} \left\{\gamma\left\|k^{m+1} - k_\epsilon\right\|^2 + \beta\mathcal{R}(k^{m+1})\right\} &= \gamma \lim_{m\to\infty}\left\|k^{m+1} - k_\epsilon\right\|^2 + \beta \lim_{m\to\infty} \mathcal{R}(k^{m+1}) \\ &= \gamma\left\|\bar{k} - k_\epsilon\right\|^2 + \beta\mathcal{R}(\bar{k}).\end{aligned} \tag{5.10}$$

Now let us show that k^{m+1} converges strongly. As the sequence converges weakly, it is enough to show

$$\lim_{m\to\infty}\left\|k^{m+1}\right\|^2 = \left\|\bar{k}\right\|^2$$

Equivalently, we can also show $\lim_{m\to\infty}\left\|k^{m+1} - k_\epsilon\right\|^2 = \left\|\bar{k} - k_\epsilon\right\|^2$. Again due to the weak convergence of k^{m+1} it is sufficient to prove

$$\limsup_{m\to\infty}\left\|k^{m+1} - k_\epsilon\right\|^2 \leq \left\|\bar{k} - k_\epsilon\right\|^2.$$

Let us assume that

$$\mu := \limsup_{m\to\infty} \left\|k^{m+1} - k_\epsilon\right\|^2 > \left\|\bar{k} - k_\epsilon\right\|^2.$$

holds. Rewriting (5.10) yields

$$\begin{aligned}\beta \limsup_{m\to\infty} \left\{\mathcal{R}(k^{m+1})\right\} &= \gamma \left(\left\|\bar{k} - k_\epsilon\right\|^2 - \limsup_{m\to\infty} \left\|k^{m+1} - k_\epsilon\right\|^2 \right) + \beta\mathcal{R}(\bar{k}) \\ &= \gamma \left(\left\|\bar{k} - k_\epsilon\right\|^2 - \mu \right) + \beta\mathcal{R}(\bar{k}) \\ &< \beta\ \mathcal{R}(\bar{k}). \end{aligned} \tag{5.11}$$

However, since $\mathcal{R}$ is w-lsc, we observe

$$\mathcal{R}(\bar{k}) \leq \liminf_{m\to\infty} \mathcal{R}(k^{m+1}) \leq \limsup_{m\to\infty} \mathcal{R}(k^{m+1}),$$

which is in contradiction to (5.11). Thus we have shown the convergence of k^{m+1} to $\bar{k}$ in norm.

The last part of this proof focus on the convergence of the partial sub-differential of J with respect to k.

Since k^{m+1} solves the sub-minimisation problem (5.5b), the optimality condition reads as $0 \in \partial_k J(k^{m+1}, f^{m+1})$, or equivalently, there exists an element

$$\xi_k^{m+1} := -\frac{1}{\beta} \left(C^*_{f^{m+1}} (C_{f^{m+1}} k^{m+1} - g_\delta) + \gamma(k^{m+1} - k_\epsilon) \right) \tag{5.12}$$

such that $\xi_k^{m+1} \in \partial\mathcal{R}\left(k^{m+1}\right) \subset \mathcal{U}$; see Corollary 5.1.3.

Now, on the limit, $0 \in \partial_k J(\bar{k}, \bar{f})$, means that

$$\bar{\xi} := -\frac{1}{\beta} \left(C^*_{\bar{f}} (C_{\bar{f}} \bar{k} - g_\delta) + \gamma(\bar{k} - k_\epsilon) \right) \quad \text{and} \quad \bar{\xi} \in \partial\mathcal{R}\left(\bar{k}\right)$$

holds, i.e., the right hand-side of (5.12) converges and the limit of the sequence of sub-derivatives belongs also to the sub-differential set $\partial\mathcal{R}\left(\bar{k}\right)$.

The first part of the statement above can be seeing by using condition (E2). Whereas the second part is obtained by the assumption that $\mathcal{R}$ is a convex functional, because in this case the Fenchel sub-differential coincides with the limiting sub-differential, which is a strong-weakly closed mapping (see Appendix B.2). □

Proposition 5.2.4. *Let $\{m\}$ be a subsequence of $\mathbb{N}$ such that the (sub-) sequence $\{(k^{m+1}, f^{m+1})\}_m$ generated by AM algorithm (5.5) satisfies $k^{m+1} \to \bar{k}$ and $f^{m+1} \rightharpoonup \bar{f}$. Then there is a subsequence of $\{f^{m+1}\}_m$ such that $f^{m_j+1} \to \bar{f}$ and $0 \in \partial_f J(\bar{k}, \bar{f})$.*

Proof. Similarly as the previous theorem, by setting $f = \bar{f}$ at (5.7) in the Proposition 5.2.2's proof we obtain

$$\begin{aligned}\liminf_{m\to\infty} \left\{ \left\|B(k^{m+1}, f^{m+1}) - g_\delta\right\|^2 + \alpha \left\|Lf^{m+1}\right\|^2 \right\} \\ = \left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \alpha \left\|L\bar{f}\right\|^2.\end{aligned}$$

As the limes inferior exists, we can in particular extract a subsequence $(k^{m_j+1}, f^{m_j+1})_{m_j}$ of $(k^{m+1}, f^{m+1})_m$ such that

$$\begin{aligned}\lim_{m_j\to\infty} \left\{ \left\|B(k^{m_j+1}, f^{m_j+1}) - g_\delta\right\|^2 + \alpha \left\|Lf^{m_j+1}\right\|^2 \right\} \\ = \left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \alpha \left\|L\bar{f}\right\|^2.\end{aligned}$$

Since both summands in the limit above are positive and due to (E1), we conclude that

$$\lim_{m_j\to\infty} \left\|Lf^{m_j+1}\right\|^2 = \left\|L\bar{f}\right\|^2.$$

Moreover, as L is a bounded and continuously invertible operator we have

$$\lim_{m_j\to\infty} \left\|f^{m_j+1}\right\|^2 = \left\|\bar{f}\right\|^2,$$

which in combination with the weak convergence of the subsequence gives its strong convergence $f^{m_j+1} \to \bar{f}$.

The second half of this proof refers to the convergence of the partial sub-differential of J with respect to f and its limit.

Since f^{m+1} solves the sub-minimisation problem (5.5a), the optimality condition reads as $0 \in \partial_f J(k^m, f^{m+1})$. However we are interested on the partial sub-derivate at the pair (k^{m_j+1}, f^{m_j+1}). Namely, with help of Corollary 5.1.3 the sub-derivative (which is a unique element) $\xi_f^{m_j+1} \in \partial_f J(k^{m_j+1}, f^{m_j+1})$ is computed[1] as

$$\xi_f^{m+1} := A^*_{k^{m+1}}(A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1},$$

which may not be necessarily null for each cycle of the AM algorithm (5.5), otherwise the stoping criteria would be satisfied and nothing would be left to be proven. Therefore we shall prove that it converges towards zero.

So far we have strong convergence of both sequences $\{k^{m+1}\}_m$ and $\{f^{m+1}\}_m$. Additionally, the Assumption E implies that both linear operators A_k and A_k^* are also strongly continuous, therefore

$$\begin{aligned}\lim_{m\to\infty} \xi_f^{m+1} &= \lim_{m\to\infty} \{A^*_{k^{m+1}}(A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1}\} \\ &= A^*_{\bar{k}}(A_{\bar{k}}\bar{f} - g_\delta) + \alpha L^* L \bar{f}. \qquad (5.13)\end{aligned}$$

[1] For sake of notation we continue to denote the subsequence's indices by $m+1$ instead of $m_j + 1$.

Our goal is to show that the limit given in (5.13) is zero. Let's suppose by contradiction that $0 \notin \partial_f J(\bar{k}, \bar{f})$. Since this set is unitary we conclude that

$$A_{\bar{k}}^*(A_{\bar{k}}\bar{f} - g_\delta) + \alpha L^* L\bar{f} \neq 0.$$

This means that $\bar{f}$ does not fulfil the normal equation associated to the standard Tikhonov problem

$$\underset{f}{\text{minimise}} \left\|A_{\bar{k}} f - g_\delta\right\|^2 + \alpha \left\|Lf\right\|^2,$$

which is a necessary condition to be a minimiser candidate to the underlying functional.

Therefore the functional $J(\bar{k}, \cdot)$ for a given fixed $\bar{k}$ does not attain its minimum value at $\bar{f}$ and there is at least one element f such that $J(\bar{k}, f) < J(\bar{k}, \bar{f})$.

Moreover this functional is convex and it has a global solution, here denoted by $\tilde{f}$. By definition

$$J(\bar{k}, \tilde{f}) \leq J(\bar{k}, f)$$

for all $f \in V$.

In particular, since $\bar{f}$ is not a minimiser for $J(\bar{k}, \cdot)$, the inequality above is strict,

$$J(\bar{k}, \tilde{f}) < J(\bar{k}, \bar{f}). \tag{5.14}$$

On the other hand, from Propostion 5.2.2 it also holds

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, f)$$

for all $f \in V$. Setting $f := \tilde{f}$ in this inequality we get

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, \tilde{f}),$$

which leads to an absurd to (5.14).

Therefore for $\bar{f}$ the optimality condition holds true, i.e., in the limit the source condition is fulfilled and the limit of the partial sub-derivative sequence is zero, i.e., $0 \in \partial_f J(\bar{k}, \bar{f})$, which completes the proof. □

Remark 5.2.5. *One alternative proof would be assuming that the sequence* $\{k^{m+1}\}_m$ *fulfils*

$$\left\|k^{m+1} - k^m\right\| \to 0. \tag{5.15}$$

More specifically, we have

$$A_{k^m}^*(A_{k^m} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1} = 0$$

from the optimality condition, but we would like to show

$$\lim_{m\to\infty} \{A_{k^{m+1}}^*(A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1}\} = 0.$$

Subtracting the latter expression from the first one, we get

$$(A^*_{k^m}A_{k^m} - A^*_{k^{m+1}}A_{k^{m+1}})f^{m+1} + (A^*_{k^m} - A^*_{k^{m+1}})g_\delta.$$

Note that by assuming the condition (5.15) *the expression above converges to zero and the proof would be complete. Nevertheless we cannot guarantee that subsequent elements of the original sequence will be selected for the subsequence. As an alternative one can verify numerically if the sequence provided from the AM algorithm satisfies this assumption. Moreover, if we restrict the problem to the simple case that the characterising function is known, then the assumption* (5.15) *is trivial, the problem becomes the standard Tikhonov regularisation and the theory is carried on.*

The forthcoming and most substantial result within this section shows that the limit $(\bar{k}, \bar{f})$ of the sequence generated by the AM algorithm is a critical point (pair) of the functional J.

Theorem 5.2.6 (Main result)**.** *Let* $\{m\}$ *a index set of* $\mathbb{N}$ *such that the sequence generated by AM algorithm* $\{(k^{m+1}, f^{m+1})\}_m \to (\bar{k}, \bar{f})$ *and* $(\xi_k^{m+1}, \xi_f^{m+1}) \rightharpoonup (0,0)$*. Then there is subsequence converging towards to a critical point of* J*, i.e.,*

$$(0,0) \in \partial J\big(\bar{k}, \bar{f}\big).$$

Proof. The Proposition 5.2.3 guarantees that $k^{m+1} \to \bar{k}$ and $\xi_{k^{m+1}} \in \partial\mathcal{R}(k^{m+1})$ (or equivalently, $0 \in \partial_k J(k^{m+1}, f^{m+1})$) such that $0 \in \partial_k J(\bar{k}, \bar{f})$. Likewise, Proposition 5.2.4 guarantees that the sequence $f^{m+1} \to \bar{f}$ and $\xi_{f^{m+1}} \in \partial J(k^{m+1}, f^{m+1})$ such that $0 \in \partial_f J(\bar{k}, \bar{f})$. Combining this with the strong-weakly closedness property of the sub-derivative (see Appendix B.2) and Theorem 5.1.2 we have

$$(0,0) \in \partial J\big(\bar{k}, \bar{f}\big) = \partial_k J(\bar{k}, \bar{f}) \times \partial_f J(\bar{k}, \bar{f})$$

on the limit. □

5.3 Computational Remarks

On the previous section we proposed an algorithm to minimise the functional J over two variables. Each cycle of the alternating minimisation problem (5.5) consists of two steps. In each step we solve instead a linear and convex minimisation over one variable, while the other one is fixed. In this section we discuss few ideas for a practical implementation.

Within an extensive choices for the regularisation term $\mathcal{R}$, we elect the weighted ℓ_p norm of the coefficients of k with respect to an orthonormal basis $\{\phi_\lambda\}_\lambda$ of $\mathcal{U}$, so

$$\big\|k\big\|^p_{w,p} := \sum_\lambda w_\lambda |k_\lambda|^p, \tag{5.16}$$

where $k_\lambda = |\langle k \ , \ \phi_\lambda \rangle|$.

It is well known [23] this choice promotes sparsity for $p = 1$. Furthermore, this functional it is nondifferentiable and its subdifferential is the the set-value function $\mathrm{Sgn}(k)$, i.e.,

$$\begin{aligned} \partial\mathcal{R}\big(k\big) &= \mathrm{Sgn}(k) \\ &= \{\xi \in L_2(\Omega^2) \ | \ \xi(s,t) \in \mathrm{sgn}(k(s,t)) \text{ a.e. } (s,t) \in \Omega^2\}, \end{aligned}$$

where the set-value sign-function $\mathrm{sgn}(k(s,t))$ is the subgradient of the function $z \mapsto |z|$ at $z = k(s,t)$. We can see that $\partial|z| = \mathrm{sgn}(z)$, where

$$\mathrm{sgn}(z) = \begin{cases} \left\{ \dfrac{z}{|z|} \right\} \ , & \text{if } z \neq 0 \ , \\ \{\zeta \in \mathbb{C} \ | \ |\zeta| \leq 1\} \ , & \text{otherwise .} \end{cases}$$

The first step consists on solving (5.5a), which is a linear, quadratic and convex minimisation problem. We skip further comments, since it can be easily done, recalling many techniques well known for solving a classical Tikhonov regularisation.

The second step is the minimisation on k of (5.5b): given f^{n+1} from the previous step we solve

$$\underset{k}{\text{minimise}} \big\| C_{f^{n+1}} k - g_\delta \big\|^2_{L_2(\Omega)} + \gamma \big\| k - k_\epsilon \big\|^2_{L_2(\Omega^2)} + \beta \big\| k \big\|^p_{w,p}.$$

This problem can be recast as a Tikhonov type functional with augmented misfit (discrepancy) term. For this purpose we define a γ-norm as

$$\big\| (x,y) \big\|^2_\gamma = \big\| x \big\|^2 + \gamma \big\| y \big\|^2 \tag{5.17}$$

for a given $\gamma > 0$, related with the inner product

$$\big\langle (x_1, y_1) \ , \ (x_2, y_2) \big\rangle_\gamma = \big\langle x_1 \ , \ x_2 \big\rangle + \gamma \big\langle y_1 \ , \ y_2 \big\rangle \tag{5.18}$$

We also define the operator

$$\begin{array}{rccc} \tilde{B} & : & \mathcal{U} \times \mathcal{V} & \longrightarrow & L_2(\Omega) \\ & & (k, f) & \longmapsto & (B(k,f), k) \end{array}$$

and the data $z_{\delta,\epsilon} = (g_\delta, k_\epsilon)$.

Under this notation we rewrite the augmented discrepancy term as

$$\begin{aligned} \big\| \tilde{B}(k,f) - z_{\delta,\epsilon} \big\|^2_\gamma &= \big\| (B(k,f), k) - (g_\delta, k_\epsilon) \big\|^2_\gamma \\ &= \big\| B(k,f) - g_\delta \big\|^2 + \gamma \big\| k - k_\epsilon \big\|^2. \end{aligned}$$

For a fixed f we can straightforward define $\tilde{C}_f$ as $\tilde{B}$ for a fixed f. Therefore we look at the following minimisation problem

$$\mathcal{J}(k) = \underset{k}{\text{minimise}} \left\| \tilde{C}_f k - z_{\delta,\epsilon} \right\|_\gamma^2 + \beta \left\| k \right\|_{w,p}^p. \tag{5.19}$$

For solving (5.19) under regularisation choice as (5.16) we construct a surrogate functional that removes the nonlinear term ${C_f}^* C_f k$. We follow the ideas of [23], adding a functional which depends of an auxiliary element u,

$$\Xi(k;u) = \eta \left\| k - u \right\|^2 - \left\| \tilde{C}_f k - \tilde{C}_f u \right\|_\gamma^2.$$

For a suitable choice of $\eta > 0$, discussed later on, the whole functional is strictly convex. Therefore the **surrogate functional** - extended functional is

$$\begin{aligned}
\mathcal{J}^{\text{Sur}}(k;u) &= \mathcal{J}(k) + \Xi(k;u) \\
&= \left\| \tilde{C}_f k - z_{\delta,\epsilon} \right\|_\gamma^2 + \beta \left\| k \right\|_{w,p}^p + \eta \left\| k - u \right\|^2 - \left\| \tilde{C}_f k - \tilde{C}_f u \right\|_\gamma^2 \\
&= \left\| \tilde{C}_f k \right\|_\gamma^2 + \left\| z_{\delta,\epsilon} \right\|_\gamma^2 - 2 \left\langle \tilde{C}_f k \ , \ z_{\delta,\epsilon} \right\rangle_\gamma + \beta \left\| k \right\|_{w,p}^p + \eta \left\| k \right\|^2 + \eta \left\| u \right\|^2 \\
&\quad -2\eta \left\langle k \ , \ u \right\rangle - \left\| \tilde{C}_f k \right\|_\gamma^2 - \left\| \tilde{C}_f u \right\|_\gamma^2 + 2 \left\langle \tilde{C}_f k \ , \ \tilde{C}_f u \right\rangle_\gamma.
\end{aligned}$$

Defining the constants $c_1 := \left\| (g_\delta, k_\epsilon) \right\|_\gamma^2$, $c_2 := \eta \left\| u \right\|^2 - \left\| \tilde{C}_f u \right\|_\gamma^2$ and $c_3 := c_1 + c_2$, applying (5.17) and (5.18)

$$\begin{aligned}
\mathcal{J}^{\text{Sur}}(k;u) &= \eta \left\| k \right\|^2 - 2 \left\langle C_f k \ , \ g_\delta \right\rangle - 2\gamma \left\langle k \ , \ k_\epsilon \right\rangle - 2\eta \left\langle k \ , \ u \right\rangle + 2 \left\langle C_f k \ , \ C_f u \right\rangle \\
&\quad +2\gamma \left\langle k \ , \ u \right\rangle + \beta \left\| k \right\|_{w,p}^p + c_1 + c_2 \\
&= \eta \left\| k \right\|^2 - 2 \left\langle k \ , \ \eta u - \gamma(u - k_\epsilon) - C_f^*(C_f u - g_\delta) \right\rangle + \beta \left\| k \right\|_{w,p}^p + c_3.
\end{aligned}$$

Under (5.16) and writing k as a linear combination of an ONB $\{\phi_\lambda\}_\lambda$

$$\begin{aligned}
\mathcal{J}^{\text{Sur}}(k;u) &= \sum_\lambda \eta (k_\lambda)^2 - 2k_\lambda \left(\eta u - \gamma(u - k_\epsilon) - C_f^*(C_f u - g_\delta) \right)_\lambda \\
&\quad +\beta w_\lambda |k_\lambda|^p + c_3.
\end{aligned}$$

We can explicitly compute the minimiser of $\mathcal{J}^{\text{Sur}}(k;u)$ with respect to k for a given auxiliary element u computing its derivative. For a choice $p = 1$ the optimality condition is translated as

$$2\eta k_\lambda = 2 \left(\eta u - \gamma(u - k_\epsilon) - C_f^*(C_f u - g_\delta) \right)_\lambda - \beta w_\lambda \operatorname{sgn}(k_\lambda).$$

Under definition of **soft-shrinkage** operator

$$\mathcal{S}_\beta(x) = \max\{\|x\| - \beta, 0\} \frac{x}{\|x\|},$$

or equivalently,

$$\mathcal{S}_\beta(x) = \begin{cases} x - \beta\frac{x}{\|x\|} & \text{if } \|x\| > \beta \\ 0 & \text{if } \|x\| \leq \beta \end{cases},$$

we end up with the explicit expression

$$k_\lambda = \mathcal{S}_{\frac{w_\lambda}{\eta}\frac{\beta}{2}}\left(u - \frac{\gamma}{\eta}(u-k_\epsilon)_\lambda - \frac{1}{\eta}[C_f^*(C_f u - g_\delta)]_\lambda\right).$$

An iterative approach can be done setting $u = k^n$ and so

$$k^{n+1} = \arg\min_k \mathcal{J}^{\mathrm{Sur}}(k; k^n)$$

for a initial guess k^0.

Therefore the minimisation subproblem on k can be solved by an iterative soft-shrinkage. It is done in two steps: first we update k with the negative direction of the gradient from the augmented discrepancy term and then we shrinkage it in the wavelet domain. More precisely

$$k_\lambda^{n+1} = \mathcal{S}_{\frac{w_\lambda}{\eta}\frac{\beta}{2}}\left(k_\lambda^n - \frac{\gamma}{\eta}(k^n-k_\epsilon)_\lambda - \frac{1}{\eta}[C_f^*(C_f k^n - g_\delta)]_\lambda\right). \tag{5.20}$$

In the following we show how to choose the constant η in order to guarantee that Ξ is strictly positive on k and so $\mathcal{J}^{\mathrm{Sur}}(k;u)$.

Lemma 5.3.1. *Let $\eta = 2(\gamma + \|f\|^2)$, where f is a fixed function in this step. Then*

$$\Xi(k;u) \geq 0 \quad \textit{and} \quad \mathcal{J}^{\mathrm{Sur}}(k;u) \geq \mathcal{J}(k)$$

Proof. It is easy to see that

$$\Xi(k;u) = (\eta-\gamma)\|k-u\|^2 - \|C_f k - C_f u\|^2.$$

Since F is linear and bounded

$$\|C_f k - C_f u\|^2 = \|C_f(k-u)\|^2 \leq \|f\|^2\|k-u\|^2.$$

Therefore,

$$\begin{aligned}(\eta-\gamma)\|k-u\|^2 - \|C_f k - C_f u\|^2 &\geq (\eta-\gamma)\|k-u\|^2 - \|f\|^2\|k-u\|^2 \\ &= \eta\|k-u\|^2 - (\gamma+\|f\|^2)\|k-u\|^2 \\ &= \eta\|k-u\|^2 - \frac{\eta}{2}\|k-u\|^2 \\ &= \frac{\eta}{2}\|k-u\|^2.\end{aligned}$$

This concludes the first part, i.e. $\Xi(k;u) \geq 0$.

Consequently, the second estimate follows easily,

$$\begin{aligned} \mathcal{J}^{\mathrm{Sur}}(k;u) &= \mathcal{J}(k) + \Xi(k;u) \\ &\geq \mathcal{J}(k) + \frac{\eta}{2}\|k-u\|^2 \\ &\geq \mathcal{J}(k). \end{aligned}$$

□

We conclude this chapter and thesis with a numerical example in order to illustrate the proposed method and algorithm.

5.4 Numerical Example

In this section we shall test the performance of the proposed method and AM algorithm through the two dimensional convolution operator equation. More precisely we convolve an image[2] composed by three levels of grey with a blurring kernel described by a Gaussian function (see the Figure 5.1 for more details).

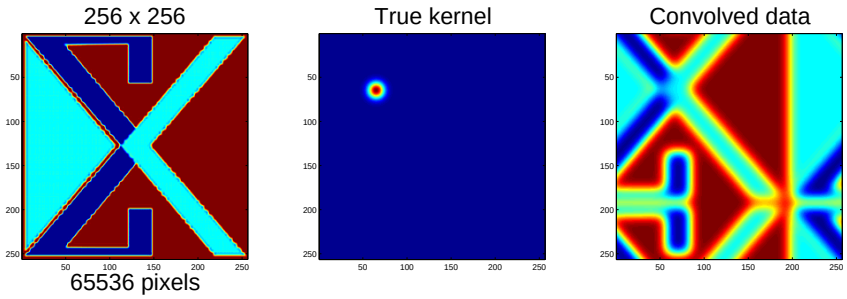

Figure 5.1: From left to right: true image f^{true}, blurring Gaussian kernel k_0 and convolved data g_0.

One cycle of the alternating minimisation problem (5.5) consists of two steps, each one solves instead a linear and convex minimisation over one variable, while the other one is fixed. Firstly, solving (5.5a) we fix k^n and find the solution f^{n+1} through, e.g., a conjugate gradient method. Secondly, solving (5.5b) we fix f^{n+1} from the previous step and solve the Shrinkage minimisation problem described on [23] and we get k^{n+1}. We shortly remark that this optimisation problem has to be first recast in a Tikhonov-type with an augmented misfit (discrepancy) term, so we can construct a surrogate functional

[2] DK Computational Mathematics' logo from JKU Linz.

to remove some nonlinear term. The algorithm starts with an initial guess k^0 and one cycle ends when we have the pair solution (k^{n+1}, f^{n+1}).

Numerical experiments are performed from given measurements not only for the data, but also for the kernel. An example of the initial noisy data and noisy kernel is illustrated on Figure 5.2, where we add 8% of white noise.

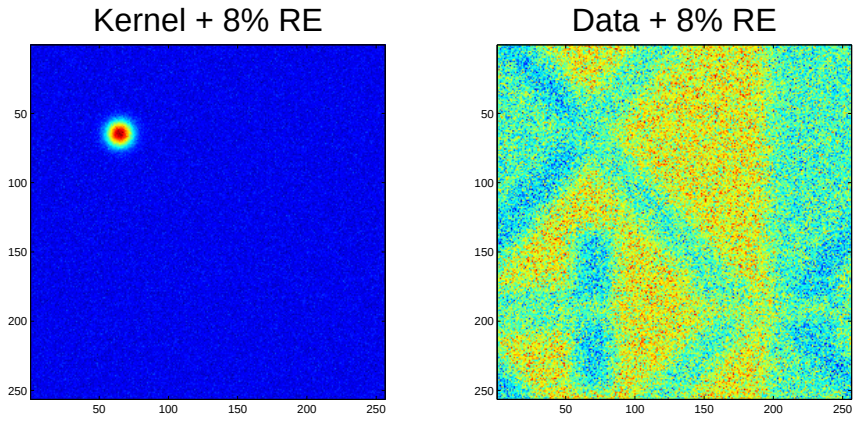

Figure 5.2: Measurements: noisy kernel (left) and noisy data (right), both with 8% relative white noise error.

The numerical results are given in the Figure 5.4, which displays in each row three graphics: the approximated image, the reconstructed kernel and its convolution. It plots a collection of numerical solutions computed from four samples with 8%, 4%, 2% and 1% noise level on both measurements, respectively in each row from top to bottom. Moreover, we compare the numerical reconstruction with the true image and kernel; the errors are displayed in the Table 5.1. Either numerically or visually one can conclude that dbl-RTLS is indeed a regularisation method, since its reconstruction and computed data improve as the noise level decreases.

The Figure 5.3 illustrates the significant improvement from the initial given noisy data (with 8% relative noise) compared to the one obtained from the dbl-RTLS solution. We also remark that for higher noise levels the dbl-RTLS reconstruction gives more than 10% accuracy than the standard Tikhonov reconstruction. On the other hand, for small noise levels, numerical experiments suggest that the improvement obtained from the dbl-RTLS method maynot payoff its computational cost.

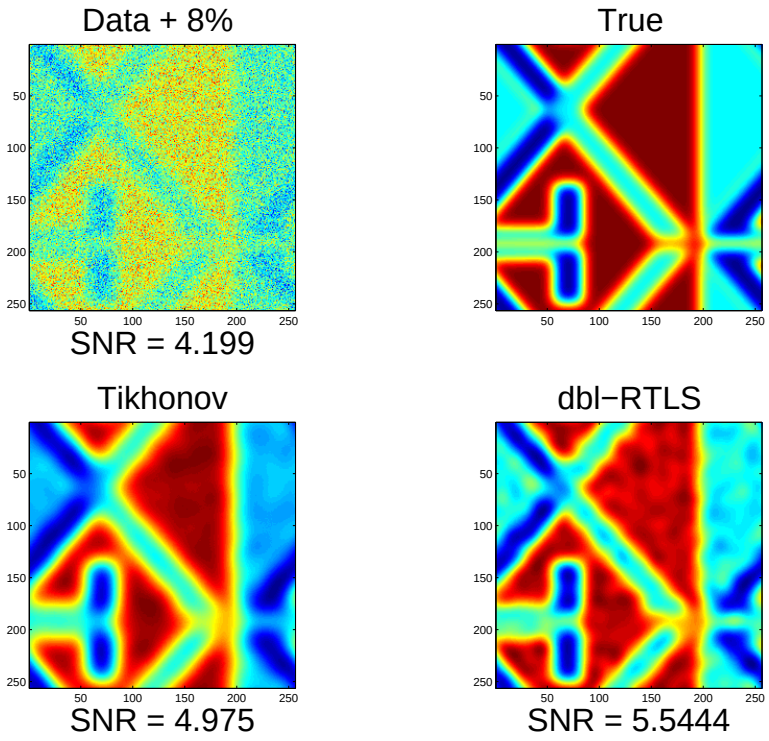

Figure 5.3: Firs row: noisy data (left) and true data (right). Second row: data attainability obtained from Tikhonov method (left) and dbl-RTLS method (right).

ε	δ	$\left\|k_n - \bar{k}\right\|_2$	$\left\|f_n - \bar{f}\right\|_2$	SNR f_n	SNR k_n	β	α
8%	8%	3.64384e-01	1.73112e-01	8.62762	10.5621	0.45254	0.12466
4%	4%	2.41851e-01	1.50364e-01	12.1164	12.2723	0.22627	0.07841
2%	2%	2.15457e-01	1.36483e-01	13.0996	13.1291	0.11313	0.04937
1%	1%	1.67541e-01	1.25965e-01	15.1905	13.6879	0.05656	0.03109

Table 5.1: Error with L_2-norm and SNR (signal-to-noise ratio).

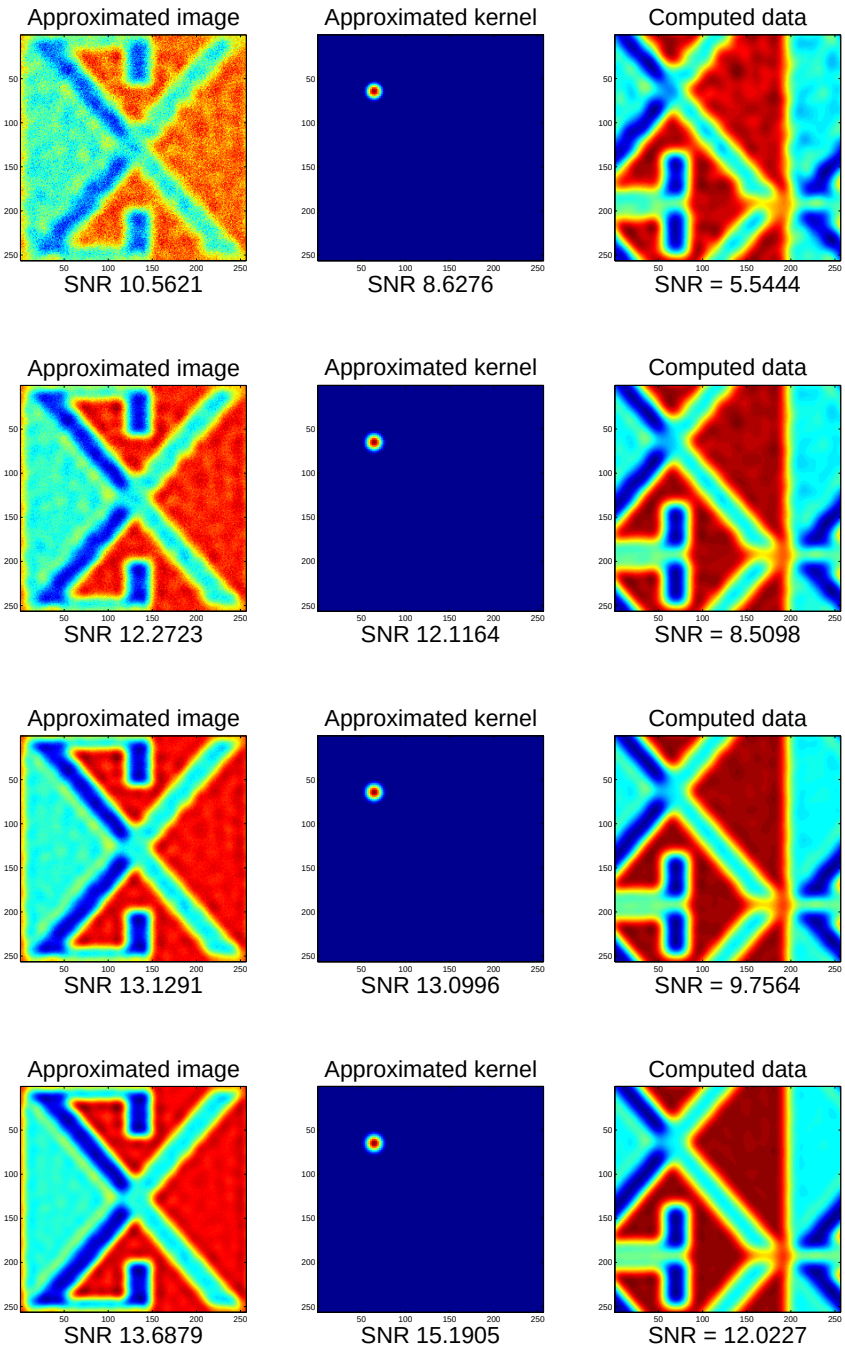

Figure 5.4: From left to right columns: deconvolution solution f^n, the reconstruction of the characterising function k^n and the attained data g^n. From the top to bottom each row is the solution given by the AM algorithm initiated with 8%, 4%, 2% and 1% relative error for both g_δ and k_ϵ.

Conclusions and Future Work

In this thesis we have explored an inverse problem in a different level, that is to say, instead of considering only noise on the data we also took into account the case where also the operator, or the function which characterises the operator, is contaminated with some noise.

Our approach to the problem combined the essential elements of both deterministic and theoretical interpretation of Tikhonov regularisation. Specifically, we designed a method not only to find a stable solution for the underlying ill-posed problem, but also to capture important features of the characterising function to be recovered. Observe that most of the approaches available in the literature also modify the discrepancy term in oder to improve the quality of the pair data and operator. Nonetheless our strategy stands up by adding an additional regularisation term to the characterising function and so we are able to give a complete qualitative and quantitative convergence results.

In detail, the first part of this thesis is intended for an overview about inverse problems which covers the most fundamental definitions and concepts. Following, we added a survey on Tikhonov-type methods also found in the literature, but here they were conveniently classified according the nature of the operator and source condition. Finally, we reviewed the most influential technique which enlightens our work, namely, the *regularised total least squares* problems.

In the second part of this thesis, the core of this dissertation, was dedicated to introduce the new methodology to solve the underlying ill-posed problem, the so called *double regularised total least squares*.

This novel scheme was presented in the Chapter 4 and also published in the most influential journal for the inverse problems community, see [6]. More explicitly we provided convergence and stability results which classify the **dbl-RTLS** as a regularisation method. In particular we would like to point out once more the remarkable result presented in the same chapter, namely, the convergence rates for both solution and operator reconstructions. Note that similar methods found in the literature can provide at most convergence results only for the regularised solution. The rates of convergence was achieved

by combining key ingredients as *source condition* and *sub-differentiability* with the concept of *Bregman distances*. Nonetheless, the outcomes are still comprehensible and comparable with advanced results well known, also summarised for convenience of the reader in the Chapter 2.

This dissertation would not be complete without a detailed study on numerical implementation. One of main drawbacks to our approach is that the minimisation problem to be solved becomes nonlinear even when the operator is linear. Jointly with non-convexity and non-differentiability that may appear, depending of the regularisation term, the most natural strategy to tackle this problem was applying an alternating minimisation algorithm.

Experimental results showed that our algorithm brought out restoration to the characterising function while finding a stable solution for the problem. Once compared against traditional techniques, where is sought only a reconstruction of the solution, the total signal-to-noise ratio (SNR) for the pair is improved. Although the computational efforts associated to the alternating minimisation algorithm are relatively low, since we solved in each step a linear problem, numerical experiments are more likely to payoff for problems with larger noise levels than for problems with relative low noise in the characterising functional. In the latter case, similar results are also given by Tikhonov-type regularisation method; which is common feature found in the classical RTLS. Indeed, as the noise level decreases, the dbl-RTLS solution converges to the (standard) regularised solution.

The work proposed in this thesis provided not only new perspectives on finding a stable solution while dealing with the instability issues mentioned, but also gave the theoretical and numerical tools required. Even though our main target problem has been solved, the main issues we dealt with now suggest numerous venues for possible extensions and future work. Therefore we list a few interesting future directions that require further investigation:

- extending the theory to a general class of nonlinear operators;
- extending the theory for Banach spaces and topological spaces;
- learn from the finite dimensional case and extend the idea of using the weighted least squares term, namely, $\dfrac{\|Af - g_\delta\|^2}{1 + \|f\|^2}$, as a generalised misfit (discrepancy) term;
- extending the regularisation properties for f, e.g., deriving rates of convergence for the non-identity operator in case of quadratic regularisation;
- extending the regularisation functional of f to a general convex and weakly lower semi-continuous functional;

- studying and adapting the discrepancy principle to choose the regularisation parameter for multiple parameter case;
- studying the variational inequality condition (VI) for the case of augmented (non-standard) discrepancy term and deriving also rates of convergence;
- finding new techniques to minimise the dbl-RTLS functional and compare it against the AM algorithm.

Bibliography

[1] Stephan W. Anzengruber and Ronny Ramlau. Morozov's discrepancy principle for Tikhonov-type functionals with nonlinear operators. *Inverse Problems*, 26(2):025001, 17, 2010.

[2] Stephan W Anzengruber and Ronny Ramlau. Convergence rates for morozov's discrepancy principle using variational inequalities. *Inverse Problems*, 27(10):105007, 2011.

[3] A. B. Bakushinskiĭ. Remarks on the choice of regularization parameter from quasioptimality and relation tests. *Zh. Vychisl. Mat. i Mat. Fiz.*, 24(8):1258–1259, 1984.

[4] Frank Bauer and Mark A. Lukas. Comparing parameter choice methods for regularization of ill-posed problems. *Math. Comput. Simulation*, 81(9):1795–1841, 2011.

[5] I. R. Bleyer and A. Leitão. On Tikhonov functionals penalized by Bregman distances. *Cubo*, 11(5):99–115, 2009.

[6] Ismael Rodrigo Bleyer and Ronny Ramlau. A double regularization approach for inverse problems with noisy data and inexact operator. *Inverse Problems*, 29(2):025004, 16, 2013.

[7] Ismael Rodrigo Bleyer and Ronny Ramlau. An efficient algorithm for solving the dbl-RTLS problem. DK Report, 2013.

[8] A. F. Boden, D. C. Redding, R. J. Hanisch, and J. Mo. Massively parallel spatially variant maximum-likelihood restoration of hubble space telescope imagery. *J. Opt. Soc. Am. A*, 13(7):1537–1545, Jul 1996.

[9] R.N. Bracewell. *The Fourier Transform and Its Applications*. Electrical engineering series. McGraw-Hill Higher Education, 2000.

[10] Kristian Bredies, Dirk A. Lorenz, and Peter Maass. Mathematical concepts of multiscale smoothing. *Appl. Comput. Harmon. Anal.*, 19(2):141–161, 2005.

[11] Lev Bregman. The relaxation method for finding the common point of convex sets and its applications to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7:200–217, 1967.

[12] Martin Burger and Stanley Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411–1421, 2004.

[13] Martin Burger and Stanley Osher. A guide to the tv zoo i: Models and analysis. preprint, April 2010.

[14] Martin Burger and Otmar Scherzer. Regularization methods for blind deconvolution and blind source separation problems. *Math. Control Signals Systems*, 14(4):358–383, 2001.

[15] C. J. Burrows, J. A. Holtzman, S. M. Faber, P. Y. Bely, H. Hasan, C. R. Lynds, and D. Schroeder. The imaging performance of the hubble space telescope. *Astrophysical Journal*, 369:L21–L25, March 1991.

[16] Fioralba Cakoni and David Colton. *Qualitative methods in inverse scattering theory.* Interaction of Mechanics and Mathematics. Springer-Verlag, Berlin, 2006. An introduction.

[17] K. Chadan and P.C. Sabatier. *Inverse problems in quantum scattering theory.* Texts and monographs in physics. Springer-Verlag, 1977.

[18] Nguyen Huy Chieu. The Fréchet and limiting subdifferentials of integral functionals on the spaces $L_1(\Omega, E)$. *J. Math. Anal. Appl.*, 360(2):704–710, 2009.

[19] Frank H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics.* SIAM, Philadelphia, 2 edition, 1990.

[20] J. Cooley, P. Lewis, and P. Welch. Historical notes on the fast fourier transform. *Audio and Electroacoustics, IEEE Transactions on*, 15(2):76–79, Jun 1967.

[21] R. Correa, A. Jofré, and L. Thibault. Characterization of lower semicontinuous convex functions. *Proc. Amer. Math. Soc.*, 116(1):67–72, 1992.

[22] Ingrid Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[23] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11):1413–1457, 2004.

[24] V. Dicken. A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems*, 15(4):931–960, 1999.

[25] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, September 1936.

[26] Ivar Ekeland and Roger Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[27] H W Engl, K Kunisch, and A Neubauer. Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems*, 5(4):523, 1989.

[28] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996.

[29] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[30] Jens Flemming and Bernd Hofmann. A new approach to source conditions in regularization with general residual term. *Numer. Funct. Anal. Optim.*, 31(1-3):254–284, 2010.

[31] Manus Foster. An application of the Wiener-Kolmogorov smoothing theory to matrix inversion. *J. Soc. Indust. Appl. Math.*, 9:387–392, 1961.

[32] Walter Gander, Gene H. Golub, and Urs von Matt. A constrained eigenvalue problem. *Linear Algebra Appl.*, 114/115:815–839, 1989.

[33] Gene H. Golub, Per Christian Hansen, and Dianne P. O'leary. Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl*, 21:185–194, 1999.

[34] Gene H. Golub and Charles F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17(6):883–893, 1980.

[35] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[36] Markus Grasmair, Markus Haltmeier, and Otmar Scherzer. Sparse regularization with l^q penalty term. *Inverse Problems*, 24(5):055020, 13, 2008.

[37] Charles W. Groetsch. *The theory of Tikhonov regularization for Fredholm equation of the first kind*. Pitman, Boston, 1984.

[38] Alfred Haar. Zur Theorie der orthogonalen Funktionensysteme. *Math. Ann.*, 69(3):331–371, 1910.

[39] Jaques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.

[40] Per Christian Hansen and Dianne P. O'leary. Regularization algorithms based on total least squares. In *in S. Van Huffel (Ed.), Recent Advances in Total Least Squares Techniques and Errors-inVariables Modeling, SIAM*, pages 127–137, 1996.

[41] Torsten Hein. Convergence rates for regularization of ill-posed problems in Banach spaces by approximate source conditions. *Inverse Problems*, 24(4):045007, 10, 2008.

[42] Torsten Hein and Bernd Hofmann. Approximate source conditions for nonlinear ill-posed problems—chances and limitations. *Inverse Problems*, 25(3):035003, 16, 2009.

[43] A. E. Hoerl. Application of ridge analysis to regression problems. *Chemical Engineering Progress*, 58(3):54–59, 1962.

[44] Bernd Hofmann. Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Methods Appl. Sci.*, 29(3):351–371, 2006.

[45] Bernd Hofmann, Barbara Kaltenbacher, Christiane Pöschl, and Otmar Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23:987–1010, 2007.

[46] Bernd Hofmann and Masahiro Yamamoto. On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems. *Appl. Anal.*, 89(11):1705–1727, 2010.

[47] L. Justen and R. Ramlau. A general framework for soft-shrinkage with applications to blind deconvolution and wavelet denoising. *Appl. Comput. Harmon. Anal.*, 26(1):43–63, 2009.

[48] L. Justen and R. Ramlau. A general framework for soft-shrinkage with applications to blind deconvolution and wavelet denoising. *Applied and Computational Harmonic Analysis*, 26(1):43 – 63, 2009.

[49] J. B. Keller. Inverse problems. *Am. Math. Mon.*, 83(2):107–118, 1976.

[50] S. Kindermann and A. Neubauer. On the convergence of the quasi-optimality criterion for (iterated) Tikhonov regularization. *Inv. Prob. Imag.*, 2(2):291–299, 2008.

[51] Andreas Kirsch. *An introduction to the mathematical theory of inverse problems*, volume 120 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.

[52] Erwin Kreyszig. *Introductory functional analysis with applications*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1989.

[53] D. Kundur and D. Hatzinakos. Blind image deconvolution. *Signal Processing Magazine, IEEE*, 13(3):43 –64, may 1996.

[54] Jörg Lampe and Heinrich Voss. Solving regularized total least squares problems based on eigenproblems. *Taiwanese J. Math.*, 14(3A):885–909, 2010.

[55] D. A. Lorenz. Convergence rates and source conditions for Tikhonov regularization with sparsity constraints. *J. Inverse Ill-Posed Probl.*, 16(5):463–478, 2008.

[56] D A Lorenz, S Schiffler, and D Trede. Beyond convergence rates: exact recovery with the tikhonov regularization with sparsity constraints. *Inverse Problems*, 27(8):085009, 2011.

[57] S. Lu, S. Pereverzev, and U. Tautenhahn. Regularized total least squares: Computational aspects and error bounds. *SIAM Journal on Matrix Analysis and Applications*, 31(3):918–941, 2010.

[58] Shuai Lu and Jens Flemming. Convergence rate analysis of tikhonov regularization for nonlinear ill-posed problems with noisy operators. *Inverse Problems*, 28(10):104003, 2012.

[59] Shuai Lu and Sergei V. Pereverzev. Multi-parameter regularization and its numerical realization. *Numer. Math.*, 118(1):1–31, 2011.

[60] Shuai Lu, Sergei V. Pereverzev, and Ulrich Tautenhahn. Dual regularized total least squeres and multi-parameter regularization. *Computational methods in applied mathematics*, 8(3):253–262, 2008.

[61] Shuai Lu, Sergei V. Pereverzev, and Ulrich Tautenhahn. A model function method in regularized total least squares. *Appl. Anal.*, 89(11):1693–1703, 2010.

[62] David G. Luenberger. *Optimization by vector space methods*. John Wiley & Sons Inc., New York, 1969.

[63] Stéphane Mallat. *A wavelet tour of signal processing*. Academic Press Inc., San Diego, CA, 1998.

[64] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier/Academic Press, Amsterdam, third edition, 2009. The sparse way, With contributions from Gabriel Peyré.

[65] I. Markovsky and S. Van Huffel. Overview of total least squares methods. *Signal Processing*, 87:2283–2302, 2007.

[66] L. Mirsky. Symmetric gauge functions and unitarily invariant norms. *Quart. J. Math. Oxford Ser. (2)*, 11:50–59, 1960.

[67] Boris S. Mordukhovich and Yong Heng Shao. On nonconvex subdifferential calculus in Banach spaces. *J. Convex Anal.*, 2(1-2):211–227, 1995.

[68] Jean Michel Morel and Sergio Solimini. *Variational methods in image segmentation*. Birkhauser Boston Inc., Cambridge, MA, USA, 1995.

[69] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 7:414–417, 1966.

[70] V. A. Morozov. *Methods for solving incorrectly posed problems*. Springer-Verlag, New York, 1984. Translated from the Russian by A. B. Aries, Translation edited by Z. Nashed.

[71] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42(5):577–685, 1989.

[72] F. Natterer. *The Mathematics of Computerized Tomography*. B.G. Teubner, Stuttgart, 1986.

[73] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.

[74] Stanley Osher and Ronald Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2003.

[75] David L. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.*, 9:84–97, 1962.

[76] R. Ramlau. TIGRA–an iterative algorithm for regularizing nonlinear ill–posed problems. *Inverse Problems*, 19(2):433–467, 2003.

[77] Ronny Ramlau. Regularization properties of Tikhonov regularization with sparsity constraints. *Electron. Trans. Numer. Anal.*, 30:54–74, 2008.

[78] Rosemary A. Renaut and Hongbin Guo. Efficient algorithms for solution of regularized total least squares. *SIAM J. Matrix Anal. Appl.*, 26(2):457–476, 2004/05.

[79] Elena Resmerita. Regularization of ill-posed problems in Banach spaces: convergence rates. *Inverse Problems*, 21(4):1303–1314, 2005.

[80] Elena Resmerita and Otmar Scherzer. Error estimates for non-quadratic regularization and the relation to enhancement. *Inverse Problems*, 22:801–814, 2006.

[81] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17:497–510, 1966.

[82] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[83] L.T Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithm. *Physica D*, 60:259–268, 1992.

[84] W. Rudin. *Functional Analysis*. MacGraw-Hill Series in Higher Mathematics. McGraw-Hill, 1977.

[85] Eberhard Schock. Approximate solution of ill-posed equations: Arbitrarily slow convergence vs. superconvergence. In G. Hämmerlin and K.-H. Hoffmann, editors, *Constructive Methods for the Practical Treatment of Integral Equations*, volume 73 of *International Series of Numerical Mathematics*, pages 234–243. Birkhäuser Basel, 1985.

[86] Diana M. Sima, Sabine Van Huffel, and Gene H. Golub. Regularized total least squares based on quadratic eigenvalue problem solvers. *BIT*, 44(4):793–812, 2004.

[87] V.S. Sunder. *Functional Analysis: Spectral Theory*. Birkhäuser Advanced Texts. Birkhäuser, 1997.

[88] U. Tautenhahn. Regularization of linear ill-posed problems with noisy right hand side and noisy operator. *Journal of Inverse and Ill-posed Problems*, 16(5):507–523, 2008.

[89] A. N. Tihonov, V. B. Glasko, and Ju. A. Kriksin. On the question of quasi-optimal choice of a regularized approximation. *Dokl. Akad. Nauk SSSR*, 248(3):531–535, 1979.

[90] A. N. Tikhonov. On the solution of incorrectly put problems and the regularisation method. In *Outlines Joint Sympos. Partial Differential Equations (Novosibirsk, 1963)*, pages 261–265. Acad. Sci. USSR Siberian Branch, Moscow, 1963.

[91] Andrei Nikolaevich Tikhonov and V. A. Arsenin. *Solution of Ill-posed Problems*. Winston & Sons, Washington, 1977.

[92] G. M. Vaĭnikko and A. Yu. Veretennikov. *Iteratsionnye protsedury v nekorrektnykh zadachakh (Extreme Methods for Solving Ill-Posed Problems)*. "Nauka", Moscow, 1986.

[93] Sabine Van Huffel and Joos Vandewalle. *The total least squares problem*, volume 9 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1991. Computational aspects and analysis, With a foreword by Gene H. Golub.

[94] C. Van Loan. *Computational Frameworks for the Fast Fourier Transform*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 1992.

[95] Curtis R. Vogel. *Computational methods for inverse problems*, volume 23 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. With a foreword by H. T. Banks.

[96] Yilun Wang, Junfeng Yang, Wotao Yin, and Yin Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.*, 1(3):248–272, 2008.

[97] K. Yosida. *Functional Analysis*. Classics in Mathematics Series. Cambridge University Press, 1980.

[98] Yu-Li You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *Image Processing, IEEE Transactions on*, 5(3):416 –428, mar 1996.

[99] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific, 2002.

[100] Clemens A. Zarzer. On Tikhonov regularization with non-convex sparsity constraints. *Inverse Problems*, 25(2):025006, 13, 2009.

Appendices

Appendix A

Preliminaries and Related Topics in Functional Analysis

> "A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. One can imagine that the ultimate mathematician is one who can see analogies between analogies."
>
> *Stefan Banach*

Functional analysis plays an important role in the applied sciences as well as in the inverse problems itself. Therefore we aim in this chapter to collect few essential concepts and results from classical literature; for a complete survey we recommend the following books [84, 97, 52, 87].

A.1 Normed and Banach Spaces

A normed space is a vector space with a metric defined by a norm (generalisation of the elementary concept of the length of a vector). A Banach space is a complete normed space.

A mapping from a normed space X into a normed space Y is called an *operator*. Of particular importance are so-called bounded linear operators since they are continuous and take advantage of the vector space structure.

Definition A.1.1. *Let X and Y be normed spaces and $L : \mathscr{D}(L) \to Y$ a linear operator, where $\mathscr{D}(L) \subset X$. The operator L is said to be bounded if there is a real number c such that for all $x \in \mathscr{D}(L)$,*

$$\|Lx\| \leq c\|x\|$$

A mapping from X into the scaler filed $\mathbb{R}$ or $\mathbb{C}$ is called a functional. In particular they are operators and all definitions and theorems for linear operators still hold.

It is of basic importance that the set of all linear functionals defined on a vector space X can itself be made into a vector space. This space is denoted by X^* and is called the *algebraic dual space* of X.

We can go a step further and consider the algebraic dual $(X^*)^*$ of X^*, whose elements are the linear functionals defined on X^*. We denote by X^{**} and call it the *second algebraic dual space* of X.

The main point to consider X^{**} is that we can obtain an interesting and important relation between X and X^{**}. First mind the notation

Space	General element	Value at a point
X	x	-
X^*	f	$f(x)$
X^{**}	g	$g(f)$

To each $x \in X$ there corresponds a $g_x \in X^{**}$, defined as $g_x(f) = f(x)$ for $f \in X^*$ variable. This defines a mapping $C : X \to X^{**}$ as $x \mapsto g_x$, called the *canonical mapping.*

It can be shown that C is injective, see [52]. Since C is linear, it is an isomorphism of X onto the range $\mathscr{R}(C) \subset X^{**}$. If C is surjective (hence bijective), so that $\mathscr{R}(C) = X^{**}$, then X is said to be *algebraically reflexive.*

Some desirable properties of finite dimensional normed spaces are related to the concept of compactness.

Definition A.1.2. *A metric space X is said to be compact if every sequence in X has a convergent subsequence. A subset M of X is said to be compact if M is compact considered as a subsapce of X, that is, if every sequence in M has a convergent subsequence whose limit is an element of M.*

Compacts sets are important since they are "well-behaved", i.e., they have several basic properties similar to those of finite sets and not shared by non-compact sets.

Theorem A.1.3 ([52, Thm 2.5-6])**.** *Let X and Y be metric spaces and $T : X \to Y$ a continuous mapping. Then the image of a compact subset M of X under T is compact*

Corollary A.1.4 ([52, Cor 2.5-7])**.** *A continuous mapping T of a compact subset M of a metric space X into $\mathbb{R}$ assumes a maximum and a minimum at some point of M.*

Let us restrict to the case of *bounded* linear operators from X into Y (both real or both complex normed spaces) and we denote this set by $B(X,Y)$. By defining the norm

$$\|T\| = \sup_{x\in X, x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{x\in X, \|x\|=1} \|Tx\|$$

$B(X,Y)$ is a normed space. Moreover, it is a Banach space if Y is a Banach space.

Comparing with the *algebraic* dual space defined previously, we define have the following similar definition.

Definition A.1.5. *Let X be a normed space. Then the set of all bounded linear functionals on X constitutes a normed space with norm defined by*

$$\|f\| = \sup_{x\in X, x\neq 0} \frac{|f(x)|}{\|x\|} = \sup_{x\in X, \|x\|=1} |f(x)|$$

which is called the dual space of X and is denoted by X'.

Since a linear functional on X maps X into $\mathbb{R}$ or $\mathbb{C}$(which are complete with the usual metric), we see that X' is $B(X,Y)$, with $Y = \mathbb{R}$ or $Y = \mathbb{C}$. Hence the following result is basic.

Theorem A.1.6 ([52, Thm 2.10-4])**.** *The dual space X' of a normed space X is a Banach space (whether or not X is).*

We can defined, as previously, X'' the *second dual space* of X or *bidual space* of X. Moreover we also define C the *canonical embedding* of X into X'' in order to define reflexive spaces, but mind that now the element f is bounded (not only linear) and therefore g_x is also bounded.

Definition A.1.7. *A normed space X is said to be reflexive if $\mathscr{R}(C) = X''$ where $C : X \to X''$ is the canonical mapping given by $x \mapsto g_x$ and $g_x(f) = f(x)$, $f \in X'$ variable.*

Usually we refer to spaces with infinite dimension, because it is known that every finite dimensional (vector) normed space is reflexive, [52, Thm 4.6-5].

A.2 Hilbert Spaces

In a general normed space is missing some condition for orthogonality (perpendicularity) which is important tool in many applications. Therefore in this section we shall review *Hilbert spaces*, i.e., a *complete inner product space*

An inner product on X is a mapping of $X \times X$ into the scalar field K of X, i.e., every pair of vectors x and y is associated a scalar denoted by $\langle x \ , \ y \rangle$. This mapping is conjugate symmetric, linear in the first argument and positive definite.

They are the most natural generalisation of Euclidean space, also a special normed spaces. For instance, an inner product on X defines a *norm* on X given by

$$\|x\| = \sqrt{\langle x \ , \ x \rangle}$$

and a *metric* on X given by

$$d(x, y) = \|x - y\| .$$

Not all normed spaces are inner product spaces. However, a norm on an inner space satisfies the *parallelogram equality*

$$\|x + y\|^2 + \|x - y\|^2 = 2 \left(\|x\|^2 + \|y\|^2 \right) . \tag{A.1}$$

We can also add the Apollonius' identity

$$\|z - x\|^2 + \|z - y\|^2 = \frac{1}{2} \|x - y\|^2 + 2 \left\| z - \frac{1}{2}(x + y) \right\|^2 . \tag{A.2}$$

An element $x \in X$ is said to be *orthogonal* to a element $y \in X$ if $\langle x \ , \ y \rangle = 0$ and we write $x \perp y$. Note that under this assumption we can define the dual pairing for $(\psi, u) \in \mathcal{U}^* \times \mathcal{U}$, where $\psi \in \mathscr{R}(F^*)$ as

$$\langle \psi \ , \ u \rangle = \langle F^* \nu \ , \ u \rangle := \langle \nu \ , \ Fu \rangle_{\mathcal{H}} \ ,$$

for some $\nu \in \mathcal{H}$.

A.3 Fourier Transform

The *continuous Fourier Transform* (FT), more precisely non-unitary definition, for a given function f and its *inverse Fourier Transform* (IFT) are given respectively as follows

$$\hat{f}(w) = \int_{-\infty}^{+\infty} f(t) \exp(-iwt) dt \tag{A.3}$$

and

$$f(t) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \hat{f}(w) \exp(iwt) dw.$$

Since the numerical experiments given in this thesis are conducted via MATLAB, we should compare its internal discrete function with the FT and IFT given above, namely:

$$F[l] = \sum_{n=1}^{N} f[n] \exp\left(-\frac{2\pi}{N} i(l-1)(n-1)\right) \tag{A.4}$$

and

$$f[n] = \frac{1}{N} \sum_{l=1}^{N} F[l] \exp\left(\frac{2\pi}{N} i(l-1)(n-1)\right) \tag{A.5}$$

where $1 \leq l \leq N$ and $1 \leq n \leq N$.

An exact relationship between the FT and the *discrete Fourier Transform* (DFT) was established by Cooley, Lewis and Welch in 1967 [20]: let $f(x)$ and $\hat{f}(w)$ be the FT pair and let $f_X(x)$ and $\hat{f}_\Omega(w)$ be their periodic versions with period $2X$ and 2Ω respectively; then, if one takes N equispaced samples of these functions in the interval $[-X, X]$ and $[-\Omega, \Omega]$, these samples form a DFT pair provided that

$$\Omega X = \pi \frac{N}{2}.$$

Following ideas given in [20] and [94] we shall find an explicit formula to relate the continuous definition with the discrete one, respectively for both FT and IFT.

A.3.1 Computational Remarks

Given a function $f : [a, b] \to \mathbb{R}$, which is periodic of the interval $[a, b]$ one can transform the definition

$$\begin{aligned}
\hat{f}(w) &= \int_{-\infty}^{+\infty} f(t) \exp(-iwt) dt \\
&= \int_{a}^{b} f(t) \exp(-iwt) dt \\
&= \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} f\left(x + \frac{a+b}{2}\right) \exp\left(-iw\left(x + \frac{a+b}{2}\right)\right) dx \\
&= \exp\left(-iw\left(\frac{a+b}{2}\right)\right) \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} f\left(x + \frac{a+b}{2}\right) \exp\left(-iwx\right) dx
\end{aligned}$$

with the change of variable $x = t - (a+b)/2$.

Therefore we have a function $f : [-X, X] \to \mathbb{R}$, where $X = (b-a)/2$. Now we can define the compute a periodic function $\hat{f}(w)$ choosing $w \in [-\Omega, \Omega]$, where

$$\Omega = \pi \frac{N}{2X} = \pi \frac{N}{b-a} = \frac{\pi}{\Delta x}.$$

We have N equispaced samples of these functions in the intervals $[-X, X]$ and $[-\Omega, \Omega]$, with

$$X = \frac{b-a}{2} \quad \text{and} \quad \Omega = \frac{\pi N}{b-a}.$$

Then we shall approximate the integral by the trapezoidal rule using the following sampling points:

$$x_n = -X + (n-1)\Delta x : n = 1, \ldots, N \text{ and } \Delta x = \frac{2X}{N} = \frac{b-a}{N} \tag{A.6}$$

with the notation $x_n := x[n]$ and

$$w_l = -\Omega + (l-1)\Delta w : l = 1, \ldots, N \text{ and } \Delta w = \frac{2\Omega}{N} = \frac{2\pi}{b-a} \tag{A.7}$$

with the notation $w_l := w[l]$.

Let us remind that the trapezoidal rule for a function $\phi(x)$ given $N+1$ points

$$\begin{aligned} \int \phi(x)dx &= \sum_{n=1}^{N} \frac{\phi(x_n) + \phi(x_{n+1})}{2} \Delta x \\ &= \left(\frac{\phi(x_1)}{2} + \sum_{n=2}^{N} \phi(x_n) + \frac{\phi(x_{N+1})}{2} \right) \Delta x \\ &= \Delta x \sum_{n=1}^{N} \phi(x_n) \end{aligned}$$

where we assume without loss of generality $\phi(x_1) = \phi(x_{N+1})$.

Then, we can compute $\hat{f}(w_l)$, i.e., on the N discretisation samples

$$\begin{aligned} \hat{f}(w_l) &= \exp\left(-iw_l\left(\frac{a+b}{2}\right)\right) \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} f\left(x + \frac{a+b}{2}\right) \exp\left(-iw_l x\right) dx \\ &\approx \exp\left(-iw_l\left(\frac{a+b}{2}\right)\right) \Delta x \sum_{n=1}^{N} f\left(x_n + \frac{a+b}{2}\right) \exp\left(-iw_l x_n\right) \end{aligned}$$

for $1 \leq l \leq N$.

Let us comment on the previous equation, namely we shall divide into three parts.

First, due the discretisation (A.6) and (A.7), the term $\exp(-iw_l x_n)$ becomes

$$\begin{aligned}
& \exp\left(-i\left(-\Omega+(l-1)\Delta w\right)\left(-X+(n-1)\Delta x\right)\right) \\
= \;& \exp\left(-i(\Omega X-\Omega(n-1)\Delta x-(l-1)X\Delta w+(l-1)(n-1)\Delta x\Delta w)\right) \\
= \;& \exp\left(-i\left(N\frac{\pi}{2}-(n-1)\pi-(l-1)\pi+(l-1)(n-1)\frac{2\pi}{N}\right)\right) \\
= \;& \exp\left(-i\pi\frac{N}{2}\right)\exp\left(i(n-1)\pi+(l-1)\pi\right)\exp\left(-i(l-1)(n-1)\frac{2\pi}{N}\right)
\end{aligned}$$

for $1 \leq l \leq N$ and $1 \leq n \leq N$.

Which leads to the following exponentials, with $N = 2^p$ where $p \geq 2$

$$\exp\left(-i\pi\frac{N}{2}\right) = \exp\left(-i\pi 2^{p-1}\right) = cos(\pi 2^{p-1}) - i\sin(\pi 2^{p-1}) = 1$$

for all $p \geq 2$.

$$\exp\left(i(n-1)\pi+(l-1)\pi\right) = \exp\left(i(n-1)\pi\right)\exp\left(i(l-1)\pi\right) = (-1)^{l-1}(-1)^{n-1}$$

Therefore, we summarize

$$\exp(-iw_l x_n) = (-1)^{l-1}(-1)^{n-1}\exp\left(-i(l-1)(n-1)\frac{2\pi}{N}\right) \tag{A.8}$$

Second remark,

$$\begin{aligned}
f\left(x_n+\frac{a+b}{2}\right) &= f\left((-X+(n-1)\Delta x)+\frac{a+b}{2}\right) \\
&= f\left(-\frac{b-a}{2}+(n-1)\Delta x+\frac{a+b}{2}\right) \\
&= f\left(a+(n-1)\Delta x\right) \\
&= f[n]
\end{aligned}$$

it is the function on the original sampling within the interval $[a, b]$, that is, $f[n] := f(t_n)$ where $t_n = a + (n-1)\Delta t$, $\Delta t = \frac{b-a}{N} = \Delta x$, for $1 \leq n \leq N$.

All together we have

$$\begin{aligned}
\hat{f}(w_l) &= \exp\left(-iw_l\left(\frac{a+b}{2}\right)\right)\int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} f\left(x+\frac{a+b}{2}\right)\exp(-iw_l x)\,dx \\
&\approx \exp\left(-iw_l\left(\frac{a+b}{2}\right)\right)\Delta x(-1)^{l-1} \\
&\quad \sum_{n=1}^{N} f[n](-1)^{n-1}\exp\left(-i(l-1)(n-1)\frac{2\pi}{N}\right)
\end{aligned}$$

where $1 \leq l \leq N$.

We can compute the Fourier transform $F[l]$ of the function $f[n](-1)^{n-1}$ through (A.4) and therefore the approximation of the continuous Fourier transform is given by

$$\hat{f}(w_l) \approx \exp\left(-iw_l\left(\frac{a+b}{2}\right)\right)(-1)^{l-1}\Delta x F[l]$$

where $1 \leq l \leq N$.

Following the previous idea, we can reconstruct the function f through its inverse Fourier transform

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(w)\exp(iwt)dw \\ &= \frac{1}{2\pi}\int_{-\Omega}^{+\Omega}\hat{f}(w)\exp(iwt)dw \\ &\approx \frac{1}{2\pi}\sum_{l=1}^{N}\hat{f}(w_l)\exp(iw_l t)\Delta w. \end{aligned}$$

Applying the same change of variable before, namely $x = t - (a+b)/2$, we can compute this approximation at each discrete point t_n for $1 \leq n \leq N$

$$\begin{aligned} f(t_n) &= f\left(x_n + \frac{a+b}{2}\right) \\ &\approx \frac{1}{2\pi}\sum_{l=1}^{N}\hat{f}(w_l)\exp\left(iw_l\left(x_n+\frac{a+b}{2}\right)\right)\Delta w \\ &= \frac{1}{2\pi}\sum_{l=1}^{N}\hat{f}(w_l)\exp\left(iw_l\left(\frac{a+b}{2}\right)\right)\exp(iw_l x_n)\Delta w \\ &\overset{(A.8)}{=} \frac{1}{2\pi}\Delta w\sum_{l=1}^{N}\hat{f}(w_l)\exp\left(iw_l\left(\frac{a+b}{2}\right)\right)(-1)^{n-1}(-1)^{l-1}\exp\left(i(l-1)(n-1)\tfrac{2\pi}{N}\right) \\ &= (-1)^{n-1}\frac{1}{\Delta x}\frac{1}{N}\sum_{l=1}^{N}G[l]\exp\left(i(l-1)(n-1)\frac{2\pi}{N}\right) \\ &\overset{(A.5)}{=} (-1)^{n-1}\frac{1}{\Delta x}g[n] \end{aligned}$$

where

$$G[l] := \hat{f}(w_l)(-1)^{l-1}\exp\left(iw_l\left(\frac{a+b}{2}\right)\right)$$

and $g[n]$ is the DFT[1] of the function G defined above.

[1]The MATLAB's routine is know as `fft`.

A.3.2 Numerical Example

On the following we illustrate (see Figure A.1) the FT of the *rectangular function*, also called *unit pulse*,

$$f(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 & \text{elsewhere} \end{cases} \tag{A.9}$$

and its known IFT, called *Sinc function*

$$\hat{f}(w) = \frac{2\sin(w)}{w}. \tag{A.10}$$

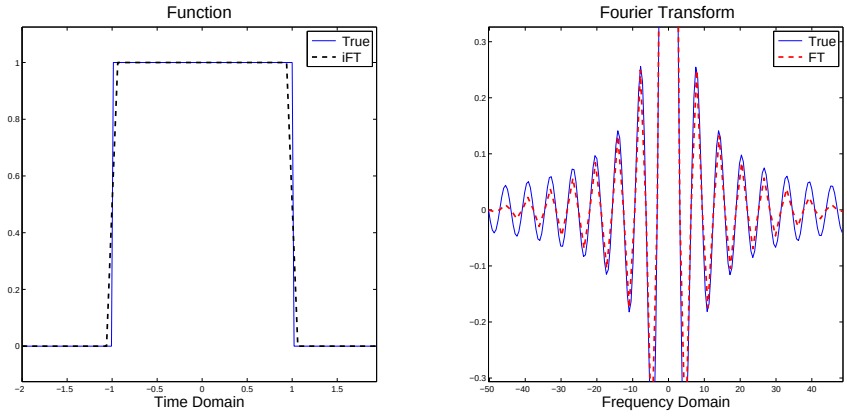

Figure A.1: (Left) Rectangular function and its reconstruction with IFT. (Right) Sinc function and FT.

A.4 Convolution Operator

One of the most important concepts in Fourier theory is related to the convolution operator. Convolutions arise in many applications. Because of a mathematical property of the Fourier transform, referred to as the convolution theorem, it is convenient to carry out calculations involving convolutions.

For a $k \in L$ and $f \in L^2$ one can define the convolution operator as

$$(k * f)(x) := \int_{\mathbb{R}} k(x-t)f(t)dt \tag{A.11}$$

where $k * f \in L^2$. Moreover it holds $\|k * f\| \leq \|k\|\|f\|$.

Theorem A.4.1. *The Fourier transform of a convolution is the product of the Fourier transforms, i.e.,*

$$\widehat{k * f} = \hat{k}\hat{f}$$

Proof. The proof follows straightforward as

$$\begin{aligned}
\widehat{(k * f)}(w) &\overset{(A.3)}{=} \int_{\mathbb{R}} (k * f)(x) \exp(-iwx) dx \\
&\overset{(A.11)}{=} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} k(x-t) f(t) dt \right) \exp(-iwx) dx \\
&= \int_{\mathbb{R}} f(t) \int_{\mathbb{R}} k(x-t) \exp(-iwx) dx dt \\
&= \int_{\mathbb{R}} f(t) \int_{\mathbb{R}} k(y) \exp(-iw(y+t)) dy dt \\
&= \int_{\mathbb{R}} f(t) \exp(-iwt) \int_{\mathbb{R}} k(y) \exp(-iwy) dy dt \\
&= \hat{k}\hat{f}
\end{aligned}$$

□

For further information on its theory and applications we recommend to the reader the book [9].

A.5 Wavelets and Multi-Resolution Analysis

For the matter of brevity, in this appendix we comment on Haar wavelet on 1 dimensional case, historically, the first orthonormal wavelet basis, constructed long before the term "wavelet" was coined, [22].

The key here is *multi-resolution analysis* (MRA). The theory was introduced in the eighties by Stephane Mallat and Yves Meyer. The basic idea consists a sequence of nested subspaces of $L_2(\mathbb{R})$ such that

1. $\{0\} \subset \ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots \subset L_2(\mathbb{R})$,
2. $\overline{\bigcup_j V_j} = L_2(\mathbb{R})$,
3. $\overline{\bigcap_j V_j} = \{0\}$,
4. $\phi(\cdot) \in V_j$ if and only if $\phi(2^{-j}\cdot) \in V_0$,
5. there exists a function $\phi(\cdot) \in V_0$, such that the system $\{\phi(\cdot - l)\}_{l \in \mathbb{Z}}$ is an orthonormal basis in V_0.

There are many sequences of subspaces which satisfies the first three conditions and they are not MRA. We emphasize last two statements. First one, all subspaces are scaled version of V_0. Second, V_0 is invariant under integer translation and there exists an orthonormal basis (ONB) in V_0.

Given a function $\phi \in L_2(\mathbb{R})$ we define

$$\phi_{j,l}(t) = 2^{j/2}\phi(2^j t - l) \qquad j, l \in \mathbb{Z}$$

where j is the scaling (or dilation) parameter and l is the shift (or translation) parameter.

One can show if $\{\phi_{0,l}\}$ is an ONB for V_0, so $\{\phi_{j,l}\}$ is an ONB for V_j, ie, $V_j = \overline{\text{span}\{\phi_{j,l} \mid l \in \mathbb{Z}\}}$.

The most important remark comes from the nested subspaces assumption

$$V_j \subset V_{j+1}.$$

For a fixed level j one element on the refined space V_{j+1} can be decomposed into two parts, one in the coarse space V_j and then in an orthogonal subspace W_j. Therefore,

$$V_{j+1} = V_j \oplus W_j. \tag{A.12}$$

In this decomposition, from a coarse to a refined level j, we keep information on the previous subspace and add new details are attached on W_j. The space W_j is called *detail space* or *wave space*.

From second assumption of MRA, given $f \in L_2(\mathbb{R})$ it is easy to see

$$\lim_{j\to\infty} P_j f = f$$

where P_j is the projection operator onto V_j. Let J be the level which we want to approximate some function, by the refinement equation (A.12)

$$\begin{aligned} V_J &= V_{J-1} \oplus W_{J-1} \\ &= (V_{J-2} \oplus W_{J-2}) \oplus W_{J-1} \\ &= (V_{J-3} \oplus W_{J-3}) \oplus W_{J-2} \oplus W_{J-1} \\ &\vdots \\ &= V_0 \oplus \bigoplus_{j=0}^{J-1} W_j. \end{aligned}$$

Therefore, taking the limit for J going to infinite, we decompose

$$L_2(\mathbb{R}) = V_0 \oplus \bigoplus_{j=0}^{\infty} W_j. \tag{A.13}$$

We can imagine with this decomposition one starts at the coarse level 0 and one improves the result adding successively a finer level and so adding more details.

One possible choice for a function which satisfies MRA is

$$\phi(t) = \begin{cases} 1 & \text{if} \quad 0 \leq t < 1 \\ 0 & \text{otherwise} \ . \end{cases} \tag{A.14}$$

This function is called father wavelet, generating function or usually *scaling* function, Figure A.2.

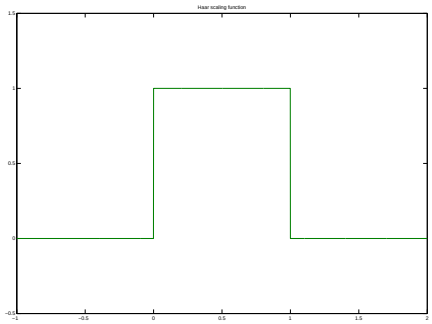

Figure A.2: Haar scaling.

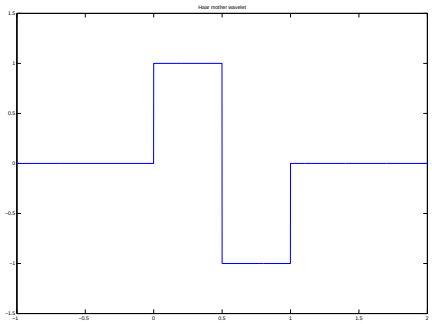

Figure A.3: Haar wavelet.

Associated with the scaling function, we define the mother *Haar wavelet* function (Figure A.3)

$$\psi(t) = \begin{cases} 1 & \text{if} \quad 0 \leq t < \frac{1}{2} \\ -1 & \text{if} \quad \frac{1}{2} \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

and the Haar wavelets (see Figure A.4)

$$\psi_{j,l}(t) = 2^{j/2}\psi(2^j t - l) \quad j, l \in \mathbb{Z}. \tag{A.15}$$

This sequence was proposed in 1909 by Alfréd Haar, [38].

By the orthogonal decomposition of L_2 in (A.13), given any function $f \in L_2(\mathbb{R})$

$$f = \sum_{l\in\mathbb{Z}} \langle f \ , \ \phi_{0,l}\rangle \phi_{0,l} + \sum_{j=0}^{\infty}\sum_{l\in\mathbb{Z}} \langle f \ , \ \psi_{j,l}\rangle \psi_{j,l} \tag{A.16}$$

We highlight this decomposition holds in general MRA in 1D case for a pair of scaling and wavelet family function, for instance, Daubechies wavelets

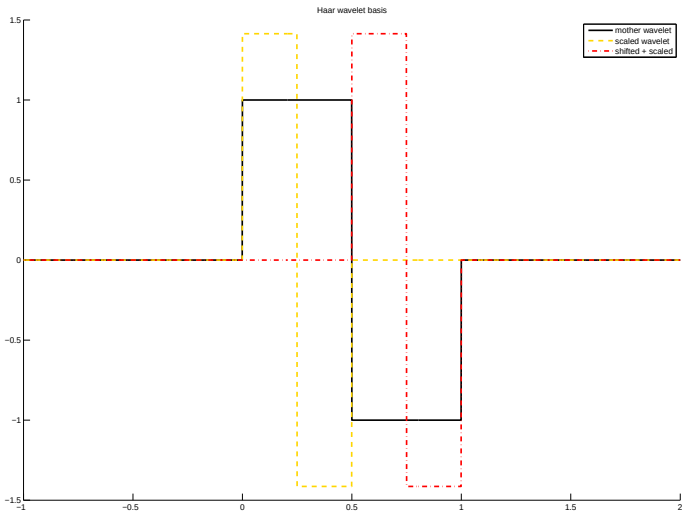

Figure A.4: Example of the Haar family with $J = 1$

[22]. For upward dimensions it also holds adapting the formula in a proper way, see the books [22, 64] for further details.

In our particular case, discrete Haar wavelet are considered on the interval $[0, 1]$. Considering the support of $\phi_{j,l}$ and $\psi_{j,l}$, namely $0 \leq 2^j t - l < 1$, it is straightforward to conclude that is enough to select the shifts between 0 and $2^j - 1$. In the Equation (A.16) the summation wrt l for wavelet function regards

$$0 \leq l < 2^j - 1.$$

It is clear the first summation where $j = 0$, the Haar scaling function when shifted is entirely outside of the interval $[0, 1]$. Therefore we do not take any translations from scaling functions, $l = 0$.

A.5.1 Wavelet Representation

From now on we setup the following notation: the family $\{\varphi_\lambda\}_{\lambda \in \Lambda}$ constitutes an orthonormal basis of Hilbert space X, where

$$\Lambda = \{1\} \cup \left\{(j, l) \mid j \in \mathbb{N}_0,\ 0 \leq 2^j - 1\right\}$$

and

$$\varphi = \begin{cases} \phi & \text{if} \quad \lambda = 1 \\ \psi_{j,l} & \text{if} \quad \lambda = (j, l). \end{cases}$$

Given a signal represented as a function $f \in L_2(\mathbb{R})$, we can decompose as

$$f = \sum_{\lambda \in \Lambda} \langle f \ , \varphi_\lambda \rangle \varphi_\lambda, \quad \text{where} \quad \langle f \ , \varphi_\lambda \rangle = \int_0^1 \varphi_\lambda(s) f(s) ds.$$

Denoting $F(f)$ the coefficient sequence of $f \in X$ with respect to the Haar wavelets, i.e.,

$$\begin{array}{rccl} F & : & X & \longrightarrow & \ell_2 \\ & & f & \longmapsto & F(f) = \{\langle f \ , \varphi_\lambda \rangle\}_{\lambda \in \Lambda}, \end{array} \tag{A.17}$$

we can represent the signal f by the sequence $x = F(f)$ which belongs to ℓ_2.

On account of numerical reasons we have to truncate the summation on j up to a certain fixed index J, called *maximal level* for wavelet. Ideally J should be taken as large as possible, however it is possible only in theory.

For instance, the function given as discrete vector with $N = 2^{13}$ samples (Figure A.5 left), the coefficient representation has length 512 (Figure A.5 right), which 216 entries are nonzero, this means a *sparse* representation on the wavelet basis.

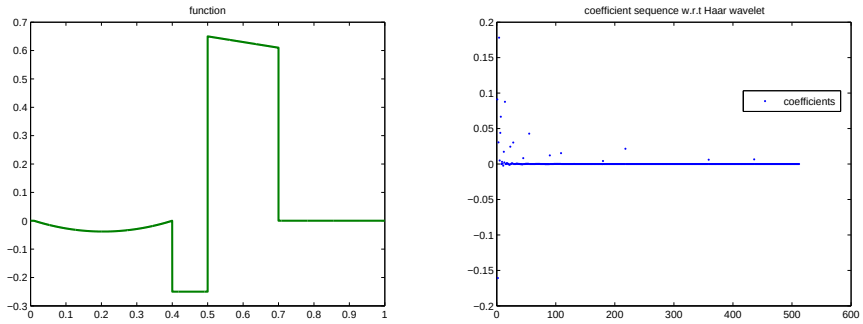

Figure A.5: Function in time domain (left) and the coefficient of its wavelet decomposition (right).

On this example we picked out $J = 8$. By combining the indices l and j for the wavelet family up to maximal index J, plus the first entry of the coefficients vector, one can compute the length of the coefficients, namely 2^{J+1}.

An important relation between the coefficients $x := F(f)$ and the function f is

$$\|f\|_{L_2}^2 = \|x\|_{\ell_2}^2.$$

With the coefficients x we can reconstruct the function via

$$f = \sum_{\lambda \in \Lambda} x_\lambda \varphi_\lambda.$$

This theory is extremely useful for solving inverse problems whenever one is interested in a sparse reconstruction. Also it has desirable properties for imaging analysis.

A.6 Derivatives

In the theory of nonlinear problems the assumptions are with respect to the linearisation of the operator, or more precisely, to its derivative. On the following we define the Gâteaux and Frechét derivates.

Definition A.6.1. *Let $u \in \Omega \subset \mathcal{U}$ and d be arbitrary in $\mathcal{U}$. If the limit*

$$dF(u;d) = \lim_{t \to 0} \frac{1}{t} \left[F(u + td) - F(u) \right]$$

exists, it is called the Gâteaux differential of F at u with increment d. If the limit exists for each $d \in \mathcal{U}$, the transformation F is said to be Gâteaux differentiable at u.

Definition A.6.2. *Let F be a transformation defined on an open domain Ω in a normed space $\mathcal{U}$ and having range in a normed space $\mathcal{H}$. If for fixed $u \in \Omega$ and each $d \in \mathcal{U}$ there exists $DF(u;d) \in \mathcal{U}$ which is linear and continuous with respect to d such that*

$$\lim_{\|d\| \to 0} \frac{\|F(u + d) - F(u) - DF(u;d)\|}{\|d\|} = 0$$

then F is said to be Frechét differential at u and $DF(u;d)$ is said to be the Frechét differential of F at u with increment d.

Commonly we use the same symbol $F'(u;d)$ for the Frechét and Gâteaux differential since generally it is apparent from the context which is meant.

Proposition A.6.3 ([62, Prop 1, §7.2])**.** *If the transformation F has a Frechét differential, it is unique.*

Proposition A.6.4 ([62, Prop 2, §7.2])**.** *If the Frechét differential of F exists at u, then the Gâteaux differential exists at u and they are equal.*

It is relatively simple to apply the concepts of Gâteaux and Frechét differential to the task of minimising a functional on a linear space, which is a topic in the Appendix B.

Appendix B

Optimisation and Non-smooth Analysis

> "Nothing at all takes place in the universe in which some rule of maximum or minimum does not appear."
>
> *Leonhard Euler*

In this chapter we consider optimisation of more general objective functionals. However, much of the theory and geometry insight are based on elements of functional analysis, provided on the previous chapter, and classical results on smooth functions. In the literature we find mostly two common geometric representation of non-linear functionals. The first one, and the most obvious, is in terms of its graph. Suppose the functional $h : \mathcal{U} \to \mathbb{R}$, so we look at elements of the space $\mathcal{U} \times \mathbb{R}$ consisting of ordered pairs (u, r). The **graph** of h is the surface in $\mathcal{U} \times \mathbb{R}$ consisting of the points $(u, h(u))$ with $u \in \mathcal{U}$.

The second representation is an extension of representing a linear functional by a hyperplane. The functional is described by its **contours** in the space $\mathcal{U}$. A *contour line* (also called isoline) of a function of two variables is a curve along which the function has a constant value. More generally, a contour line for a function of two variables is a curve connecting points where the function has the same particular value. The gradient of the function is always perpendicular to the contour lines. When the lines are close together the magnitude of the gradient is large: the variation is steep. A *level set* is a generalisation of a contour line for functions of any number of variables, see more details in [74].

Unfortunately we do not cover both representation here, for further details we recommend to the reader to check out the books [82, 19, 26]. On the following we display a brief collection of classical theorems and definitions.

B.1 General Theory

Mathematically speaking, optimisation is the minimisation or maximisation of a function over its variables, which may be subject to some constraints.

Since in the appendix we aim only a short revision for a general setting we should first of all introduce the following notation:

- u is the vector of *variables*, also called *unknowns* or *parameters*;
- h is the *objective function*, a functional of u that we want to maximise or minimise;
- c_i are *constraint* functions of u, usually they are separated into equations (where we denote by $i \in \mathcal{E}$) and inequalities (where we denote by $i \in \mathcal{I}$)

Therefore a general minimisation problem can be written as follows:

$$\begin{array}{ll} \text{minimise} & h(u) \\ \text{subject to} & c_i(u) = 0,\ i \in \mathcal{E} \\ & c_i(u) \geq 0,\ i \in \mathcal{I} \end{array} \tag{B.1}$$

Without loss of generality we restrict to minimisation problems, since we maximisation can be obtained by *minimising* $-h$.

Once the model is defined as above, one algorithm can be used to find the variables that optimise the objective functional. After succeeding in the task of finding a solution we have to check if it is indeed a solution of the problem, which leads us to *optimality condition*.

The general problem (B.1) can be classified as linear, non-linear, convex, differentiable or non-differentiable depending on the objective function and constraints.

If $\mathcal{E} = \mathcal{I} = \emptyset$ then (B.1) is called *unconstrained problem*, otherwise it is called *constrained problem*. It is also very important to highlight that unconstrained problems arise also as reformulations of constrained optimisation problems, in which the constraints are replaced by penalisation terms added to objective function that have the effect of discouraging constraint violations, e.g., the Tikhonov functional is mostly considered as an unconstrained problem [33].

In general, for non-linear problems, both constrained and unconstrained, may possess local solutions that are not global solutions. Algorithms may not always find global solutions. Therefore *convex problems* are becoming very popular, in particular linear problems, where local solutions are also global solutions.

B.1.1 Unconstrained Optimisation

For a general unconstrained optimisation problem we refer to

$$\underset{u}{\text{minimise}} \; h(u) \tag{B.2}$$

where u belongs to the entire space $\mathcal{U}$ and h is assumed to be differentiable until further remark.

We say $\overline{u}$ is a *global solution* of h if $h(\overline{u}) \leq h(u)$ for all $u \in \mathcal{U}$. Since the task of finding a global solution is not always an easy task, we can accept a local solution. Namely, $\overline{u}$ is a *local solution* if there is a neighbourhood Ω of $\overline{u}$ such that $h(\overline{u}) \leq h(u)$ for all $u \in \Omega$.

On the following we cite the so called, respectively, First-Order Necessary Conditions, Second-Order Necessary Conditions and Second-Order Sufficient Conditions.

Theorem B.1.1 ([73, Thm 2.2]). *If $\overline{u}$ is a local minimiser and h is continuously differentiable in an open neighbourhood of $\overline{u}$, then $h'(\overline{u}) = 0$.*

Theorem B.1.2 ([73, Thm 2.3]). *If $\overline{u}$ is a local minimiser of h and h'' exists and is continuous in an open neighbourhood of $\overline{u}$, then $h'(\overline{u}) = 0$ and $h''(\overline{u})$ is positive semidefinite.*

Theorem B.1.3 ([73, Thm 2.4]). *Suppose that h'' is continuous in an open neighbourhood of $\overline{u}$ and that $h'(\overline{u}) = 0$ and $h''(\overline{u})$ is positive definite. Then $\overline{u}$ is a strict local minimiser of h.*

The necessary and sufficient condition are essential tools for designing an efficient algorithm, however the smoothness assumption is not always fulfilled. Therefore we shall see in Section B.2 a more general definition of subderivatives and generalisation of the results given above.

B.1.2 Constrained Optimisation

A general formulation for this problem is given as (B.1), where $\mathcal{E}$ and $\mathcal{I}$ are two finite sets of indices. We define the *feasible set* Ω to be the set of points u that satisfy the constraints; that is

$$\Omega = \{u \mid c_i(u) = 0, i \in \mathcal{E}; c_i(u) \geq 0, i \in \mathcal{I}\}$$

As in the previous subsection, we also discuss necessary and sufficient conditions. For that we need another definition, and one of the most important terminology, is the following:

The *active set* $\mathcal{A}(u)$ at any feasible u consists of the equality constraint induces from $\mathcal{E}$ together with the indices of the inequality constraints i for which $c_i(u) = 0$; that is,

$$\mathcal{A}(u) = \mathcal{E} \cup \{i \in \mathcal{I} \mid c_i(u) = 0\}$$

Definition B.1.4 (LICQ)**.** *Given the point u and the active set $\mathcal{A}(u)$ defined above, we say theta the linear independence constraint qualification (LICQ) holds if the set of active constraint gradients $\{\nabla c_i(u), i \in \mathcal{A}(u)\}$ is linearly independent.*

As a preliminary to stating the necessary conditions, we have to define the *Lagrangian function* for the general problem (B.1)

$$\mathcal{L}(u, \lambda) = h(u) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(u)$$

The necessary condition defined In the following are called *first-order necessary conditions*, because they are concerned with properties of the gradients (first derivatives) of the objective and constraint functions.

Theorem B.1.5 ([73, Thm 12.1])**.** *Suppose that $\overline{u}$ is a local solution of (B.1), that the functions h and c_i are continuously differentiable, and the LICQ holds at $\overline{u}$. Then there is a Lagrange multiplier vector $\bar{\lambda}$ with components $i \in \mathcal{E} \cup \mathcal{I}$, such that the following conditions are satisfied at $(\overline{u}, \bar{\lambda})$*

$$\nabla_u \mathcal{L}(\overline{u}, \bar{\lambda}) = 0, \tag{B.3a}$$
$$c_i(\overline{u}) = 0, \quad \forall i \in \mathcal{E} \tag{B.3b}$$
$$c_i(\overline{u}) \geq 0, \quad \forall i \in \mathcal{I} \tag{B.3c}$$
$$\bar{\lambda}_i \geq 0, \quad \forall i \in \mathcal{I} \tag{B.3d}$$
$$\bar{\lambda}_i c_i(\overline{u}) = 0, \quad \forall i \in \mathcal{E} \cup \mathcal{I} \tag{B.3e}$$

The conditions (B.3) are often known as the *Karush-Kuhn-Tacker condition* or *KKT condition* for short.

There are more results which depend on new definitions and special subspaces called *cones*, also further qualification conditions. However we left them to the read to check them in the book [73]. We continue in the next Section with the case where smoothness is not required.

B.2 Generalised Derivatives

The definition of subdifferential for convex functions appeared first in 1960 by Moreau and Rockafellar [81]. The *Fenchel subdifferential* of a functional $h : \mathcal{U} \to \overline{\mathbb{R}}$ (or $[-\infty, +\infty]$) at $\bar{u} \in \mathcal{U}$ is the set

$$\partial^F h(\bar{u}) = \{\xi \in \mathcal{U}^* \mid h(\bar{u} + d) - h(\bar{u}) \geq \langle \xi , d \rangle \ \forall d \in \mathcal{U}\}.$$

For nonconvex function the extension of this definition is due Frank Clarke on 1973. It is based on generalised directional derivative for locally Lipschitzian

functions on Banach spaces. He performed pioneering work in the area of nonsmooth analysis spread far beyond the scope of convexity, [19]. The *Clark subdifferential* of h at $\bar{u}$ is defined by

$$\partial^C h(\bar{u}) = \{\xi \in \mathcal{U}^* \mid h^\circ(\bar{u}; d) \geq \langle \xi , d \rangle \ \forall d \in \mathcal{U}\}$$

where

$$h^\circ(\bar{u}; d) = \limsup_{\substack{u \to \bar{u} \\ t \downarrow 0}} \frac{h(u + td) - h(u)}{t}$$

is the *generalised directional derivative*.

We add to this list two more definitions of subdifferentials. As preliminary, for a set-valued mapping $G : \mathcal{U} \rightrightarrows \mathcal{U}^*$ between a Banach space $\mathcal{U}$ and its topological dual $\mathcal{U}^*$, the set

$$\operatorname*{Lim\,sup}_{u \to \bar{u}} G(\bar{u}) = \{\xi \in \mathcal{U}^* \mid \exists\, u^n \to \bar{u} \text{ and } \xi^n \overset{*}{\rightharpoonup} \xi \text{ with } \xi^n \in G(u^n)\ \forall n \in \mathbb{N}\}$$

denotes the *sequential Painlevé-Kuratowski upper/outer limit* of a set-value mapping. In another words, it is the set of limits on $\mathcal{U}^*$ with desirable convergence property.

On between the previous two definitions is the *Fréchet subdifferential*. Given a lower semicontinuous function h, the ε-Fréchet subdifferential of h at $\bar{u}$ is defined by

$$\hat{\partial}_\varepsilon h(\bar{u}) = \left\{ \xi \in \mathcal{U}^* \mid \liminf_{\|d\| \to 0} \frac{h(\bar{u} + d) - h(\bar{u}) - \langle \xi , d \rangle}{\|d\|} \geq \varepsilon \right\}.$$

If $|h(\bar{u})| = \infty$ then $\hat{\partial}_\varepsilon h(\bar{u}) = \emptyset$. When $\varepsilon = 0$ the set $\hat{\partial}_0 h(\bar{u})$ will be denoted by $\hat{\partial} h(\bar{u})$.

The *limiting subdifferential* or *Mordukhovich subdifferential* of h at $\bar{u}$ is defined as

$$\hat{\partial} h(\bar{u}) = \operatorname*{Lim\,sup}_{\substack{u \overset{h}{\to} \bar{u} \\ \varepsilon \downarrow 0}} \hat{\partial}_\varepsilon h(\bar{u})$$

where the notation $u \overset{h}{\to} \bar{u}$ means $u \to \bar{u}$ with $h(u) \to h(\bar{u})$. This subdifferential corresponds to the collection of weak-star sequential limiting points of the so-called ε-Fréchet subdifferential.

On [21] the authors emphasize the following inclusion property between the sets

$$\partial^F h(\bar{u}) \subset \hat{\partial} h(\bar{u}) \subset \partial^C h(\bar{u}).$$

The intermediary set of subgradients $\hat{\partial} h(\bar{u})$ can be nonconvex and therefore it is hard for some analysis, while Clark subdifferential is always nonempty convex subset of $\mathcal{U}^*$ whenever $\bar{u} \in \operatorname{dom} h$. A very important remark found in

[18] says *the subdifferential definitions generate the same set if the function is convex*. Some basic proprieties of differentiable functions in the classical sense, like linearity, may be not hold in the subdifferentiable case. More details can be found in [19, 67].

Another very important fact pointed out in the book [19, Proposition 2.3.15] is: for a general function neither of the set $\partial h(x_1, x_2)$ and the product set $\partial_1 h(x_1, x_2) \times \partial_2 h(x_1, x_2)$ need be contained in the other, where $\partial_i h$ denotes the partial subderivative with respect to x_i for $i = 1, 2$.

Proposition B.2.1. *Optimality condition: If $\bar{u}$ minimises h then*

$$0 \in \partial^F h(\bar{u})$$

Example 4. Consider the function $\mathcal{R}(u) = |u|$

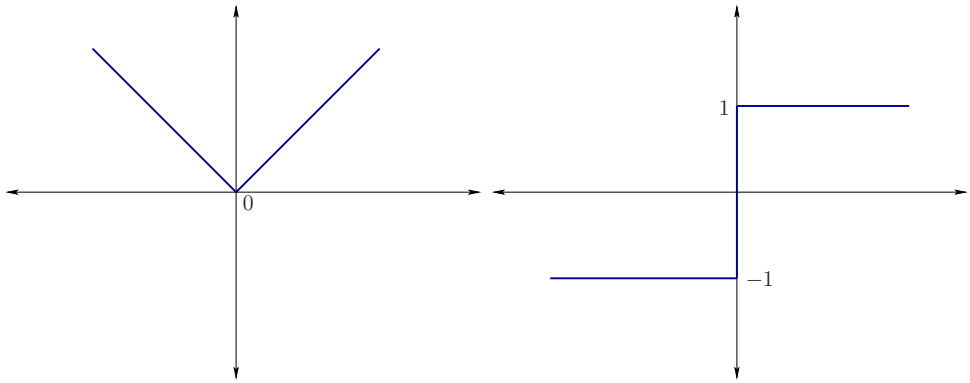

Figure B.1: Function (left) and its subdifferential (right).

B.2.1 Basic Properties

Let M be a subset of the Banach space $\mathcal{U}$. A function $h : M \to \mathbb{R}$ is said to satisfy a Lipschitz condition (on M) provided that for some non-negative scalar K, one has

$$|h(u) - h(\tilde{u})| \leq K \|u - \tilde{u}\| \tag{B.4}$$

Proposition B.2.2 ([19, Prop 2.1.1])**.** *Let h be Lipschitz of rank K near u. Then*

(a) then function $d \to h^\circ(u; d)$ is finite, positively homogeneous, subadditive on $\mathcal{U}$ and satisfies

$$|h^\circ(u; d)| \leq K \|d\|$$

(b) $h^\circ(u; d)$ is upper semi-continuous as a function of (u, d) and, as a function of d alone, is Lipschitz of rank K on $\mathcal{U}$.

(c) $h^\circ(u;-d) = (-h)^\circ(u;d)$

$$\|\xi\|_* = \sup\left\{\langle \xi \ , \ u\rangle \ \mid \ u \in \mathcal{U}, \|u\| \leq 1\right\} \tag{B.5}$$

Proposition B.2.3 ([19, Prop 2.1.2 and 2.1.5])**.** *Let h be Lipschitz of rank K near u. Then*

(a) $\partial^C h\,(u)$ is non-empty, convex, weak-compact subset of $\mathcal{U}^*$ and $\|\xi\|_* \leq K$ for every $\xi \in \partial^C h\,(u)$.*

(b) for every $u \in \mathcal{U}$, one has

$$h^\circ(u;d) = \max\left\{\langle \xi \ , \ u\rangle \ \mid \ \xi \in \partial^C h\,(u)\right\}$$

(c) Let u_i and ξ_i be sequences in $\mathcal{U}$ and $\mathcal{U}^*$ such that $\xi_i \in \partial^C h\,(u_i)$. Suppose that $u_i \to \bar{u}$ and that $\bar{\xi}$ is a cluster point of ξ_i in the weak topology. then one has $\bar{\xi} \in \partial^C h\,(\bar{u})$*

The last item states the multifunctional $\partial^C h\,(u)$ is weak*-closed.

Now we introduce the most natural concept of differentiability linked to the theory of this chapter: *strictly differentiability* (Bourbaki).

Definition B.2.4. *Let $F : \mathcal{U} \to \mathcal{V}$ be map between Banach spaces. F admits a strict derivative at $\bar{u}$, an element of $\mathscr{L}(\mathcal{U},\mathcal{V})$ denoted $D_s F(\bar{u})$, provided that for each d it holds*

$$\lim_{\substack{u\to\bar{u}\\ t\downarrow 0}} \frac{F(u+td)-F(u)}{t} = \langle D_s F(\bar{u}) \ , \ d\rangle \tag{B.6}$$

Proposition B.2.5 ([19, Prop 2.2.4])**.** *If h is strictly differentiable at u, then h is Lipschitz near u and $\partial^C h\,(u) = \{D_s h(u)\}$. Conversely, if h is Lipschitz near u and $\partial^C h\,(u)$ reduces to a singleton $\{\xi\}$, then h is strictly differentiable at u and $D_s h(u) = \xi$.*

Monotonicity of sub-gradient: add some comments pointed by Resmerita

Proposition B.2.6 ([19, Prop 2.2.9])**.** *Let h be Lipschitz near each point of an open convex subset X of $\mathcal{U}$. Then h is convex on $\mathcal{U}$ iff the multifunction $\partial^C h\,(\cdot)$ is monotone on $\mathcal{U}$; that is, iff*

$$\left\langle u-\tilde{u} \ , \ \xi-\tilde{\xi}\right\rangle \geq 0 \quad \text{for all } u,\tilde{u} \in \mathcal{U}, \xi \in \partial^C h\,(u)\,, \tilde{\xi} \in \partial^C h\,(\tilde{u}) \tag{B.7}$$

B.2.2 Basic Calculus

Proposition B.2.7 ([19, Prop 2.3.1]). *For any scalar s, one has*

$$\partial^C\left(sh\right)(u) = s\partial^C\left(h\right)(u) \tag{B.8}$$

Proposition B.2.8 ([19, Prop 2.3.2]). *If h attains a local minimum or maximum at $\bar{u}$, then $0 \in \partial^C h\,(\bar{u})$.*

If h_i for $i = 1, 2, \ldots, n$ is a finite family of functions each of which is Lipschitz near u, it follows easily that their sum $h = \sum h_i$ is also Lipschitz near u. Moreover the following two results hold true.

Proposition B.2.9 ([19, Prop 2.3.3]).

$$\partial^C\left(\sum h_i\right)(u) \subset \sum \partial^C h_i\,(u)$$

Corollary B.2.10 ([19, Cor 1]). *Equality holds in Proposition B.2.13 if all but at most one of the functional h_i are strictly differentiable at u.*

It is often the case that calculus formulas for generalised gradients involve inclusions, such as in Proposition B.2.13. For instance, equality certainly holds if all the functions in question are continuously differentiable. However, one would wish for a less extreme condition: one that would cover the convex case. A class of functions is

Definition B.2.11. *h is said to be regular at u provided*

(a) For all d, the usual one-sided directional derivative $h'(u;d)$ exists.

(b) Fora all d, $h'(u;d) = h^\circ(u;d)$

This class of function is very useful and mainly contributes as stated in the following result

Corollary B.2.12 ([19, Cor 3]). *If each h_i is regular at u, equality holds in Proposition B.2.13.*

Proposition B.2.13 ([19, Prop 2.3.6]). *Let h be Lipschitz near u*

(a) If h is strictly differentiable at u, then h is regular at u.

(b) If h is convex, then h is regular at u.

(c) A finite linear combination (by non-negative scalars) of functions regular at u is regular at u.

(d) If h admits a Gâteaux derivative $Dh(u)$ and is regular at u, then $\partial^C h\,(u) = \{Dh(u)\}$

Now we introduce the *partial generalised gradients* or *partial sub-gradients*. If $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$, are Banach spaces, an let $h(u_1, u_2)$ on $\mathcal{U}$ be Lipschitz near (u_1, u_2). We denote by $\partial_1 h(u_1, u_2)$ the partial generalised gradient of $h(\cdot, u_2)$ at u_1, and respectively, $\partial_2 h(u_1, u_2)$ the partial generalised gradient of $h(u_1, \cdot)$ at u_2.

It is a fact that in general neither of the sets $\partial h(u_1, u_2)$ and $\partial_1 h(u_1, u_2) \times \partial_2 h(u_1, u_2)$ need to be contained in the other. For regular functions, however, a general relationship does hold between these sets, as presented in the following.

Proposition B.2.14 ([19, Prop 2.3.15])**.** *If h is regular at $h(u_1, u_2)$, then*

$$\partial h(u_1, u_2) \subset \partial_1 h(u_1, u_2) \times \partial_2 h(u_1, u_2)$$

For non-regular functions we need to define a projection operator, however we will not cover this case; for more details [19, Sec 2.3].

B.3 Bregman Distance

In this last Section we introduce quickly the so called *Bregman distance*.

Definition B.3.1. *Let $\Omega \subset \mathcal{U}$ a convex set from a Banach space $\mathcal{U}$ and $h : \Omega \to \mathbb{R}_+$ a convex functional, the generalised Bregman distance of h between the elements $u, v \in \Omega$ is*

$$D_h\big(v, u\big) = \Big\{ D_h^{\xi}\big(v, u\big) \;\mid\; \xi \in \partial h\big(u\big) \Big\}$$

where $D_h^{\xi}\big(v, u\big) := h(v) - h(u) - \big\langle \xi \;,\; v - u \big\rangle$.

For a visual illustration of the Bregman distance see the Figure B.2

Example 5. Let h be a differentiable functional defined as $h(u) = \frac{1}{2}\|u\|^2$. Therefore the sub-derivative set is unitary $\xi = \{u\}$ and so the Bregman distance coincides with the standard metric distance

$$D_h\big(u, \bar{u}\big) = \frac{1}{2}\|u - \bar{u}\|^2 .$$

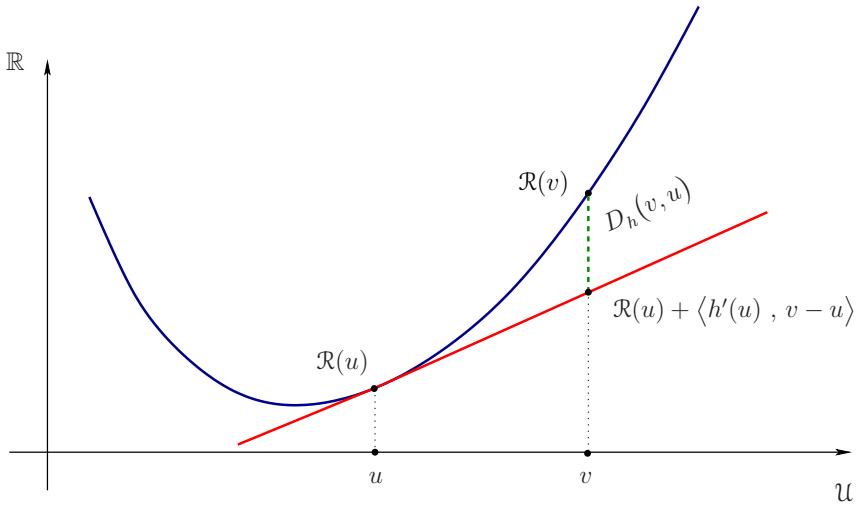

Figure B.2: Bregman distance illustration.

Curriculum Vitae

Personal Information

Surname, name	BLEYER, Ismael Rodrigo
Date of Birth	April 11, 1985
Birthplace	Florianópolis, Brazil
Citenzenship	Brazilian
Sex	Male

University Education

2009 – 2014	Doctor in Mathematics. Johannes Kepler University (Austria). Thesis title: *Double Regularised Total Least Squares Method.* Supervisor: Prof. Dr. Ronny Ramlau.
2007 – 2008	M.Sc. in Mathematics. Major: Honor Magister in Applied Mathematics. Federal University of Santa Catarina (Brazil). Thesis title: *Tikhonov Functional and Penalisation with Bregman Distances.* Supervisor: Prof. Dr. Antônio Carlos Gardel Leitão.
2003 – 2007	B.Sc. in Mathematics and Scientific Computing. Federal University of Santa Catarina (Brazil). Thesis title: *An Interior Point Algorithm for the Tangential Subproblem in Filter Methods for Nonlinear Programming.* Supervisor: Prof. Dr. Clovis Caesar Gonzaga.

Research Interests

Inverse problems, regularisation theory, iterative methods, optimisation methods.

Scientific Production

Journals	Bleyer, I. and Ramlau, R. *A Double Regularisation Approach for Inverse Problems with Noisy Data and Inexact Operator.* Inverse Problems, Vol 29 (2): 025002, 2013.
Submitted	Bleyer, I. and Ramlau, R. *An Efficient Algorithm for Solving the dbl-RTLS Problem.* TBA

Scientific Stays

Oct – Nov, 2012	Fudan University, China. visiting Prof. Shuai Lu.
May – Jul, 2012	Chemnitz University of Technology, Germany. visiting Prof. Bernd Hofmann.
Sep, 2011	Federal University of Santa Catarina. Florianópolis, Brazil. visiting Prof. Antônio Leitão.
Feb – May, 2011	University of Helsinki, Finland. visiting Prof. Samuli Siltanen.
Jul, 2010	University of Frankfurt, Germany. visiting Prof. Antônio Leitão.

Sworn Declaration

"I hereby declare under oath that the submitted Doctor's thesis has been written solely by me without any outside assistance, information other than provided sources or aids have not been used and those used have been fully documented. The Doctor's thesis here present is identical to the electronically transmitted text document."

German: **Eidesstattliche Erklärung.**

"Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Die vorliegende Dissertation ist mit dem elektronisch übermittelten Textdokument identisch."

Ismael Rodrigo Bleyer