Topics in Inverse Problems

Publicações Matemáticas

Topics in Inverse Problems

Johann Baumeister
Universität Frankfurt

Antonio Leitão
UFSC

impa

25º Colóquio Brasileiro de Matemática

Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

25º Colóquio Brasileiro de Matemática

- A Short Introduction to Numerical Analysis of Stochastic Differential Equations - Luis José Roman
- An Introduction to Gauge Theory and its Applications - Marcos Jardim
- Aplicações da Análise Combinatória à Mecânica Estatística - Domingos H. U. Marchetti
- Dynamics of Infinite-dimensional Groups and Ramsey-type Phenomena - Vladimir Pestov
- Elementos de Estatística Computacional usando Plataformas de Software Livre/Gratuito - Alejandro C. Frery e Francisco Cribari-Neto
- Espaços de Hardy no Disco Unitário - Gustavo Hoepfner e Jorge Hounie
- Fotografia 3D - Paulo Cezar Carvalho, Luiz Velho, Anselmo Antunes Montenegro, Adelailson Peixoto, Asla Sá e Esdras Soares
- Introdução à Teoria da Escolha - Luciano I. de Castro e José Heleno Faro
- Introdução à Dinâmica de Aplicações do Tipo Twist - Clodoaldo G. Ragazzo, Mário J. Dias Carneiro e Salvador Addas-Zanata
- Schubert Calculus: an Algebraic Introduction - Letterio Gatto
- Surface Subgroups and Subgroup Separability in 3-manifold Topology - Darren Long and Alan W. Reid
- Tópicos em Processos Estocásticos: Eventos Raros, Tempos Exponenciais e Metaestabilidade - Adilson Simonis e Cláudia Peixoto
- **Topics in Inverse Problems - Johann Baumeister and Antonio Leitão**
- Um Primeiro Curso sobre Teoria Ergódica com Aplicações - Krerley Oliveira
- Uma Introdução à Simetrização em Análise e Geometria - Renato H. L. Pedrosa

Distribuição:
IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br - http://www.impa.br
ISBN: 85-244-0224-5

Preface

The demands of natural science and technology have brought to the fore many problems that are inverse to the classical direct problems, that is, problems which may be interpreted as finding the cause of a given effect. Inverse problems are characterized by the fact that they are usually much harder to solve than their direct counterparts since they are usually associated to ill-posed problems. As a result a very exiting and important area of research has been developed in the last decades. The combination of classical analysis, linear algebra, applied functional and numerical analysis is one of the fascinating features of this relatively new research area.

This monograph will not give an extensive survey of papers on inverse problems. The goal of the notes is to present the main ideas in treating inverse problems and to make clear the progress of the theory of ill-posed problems. The monograph arose from a booklet [5], courses and lectures given by the authors.
The presentation is intended to be accessible to students whose mathematical background include basic courses in advanced calculus, linear algebra and functional analysis. The monograph can be used as the backbone for a lecture on inverse and ill-posed problems.

We would like to thank the SBM to make possible these notes. Moreover, both authors are grateful to the CAPES, CNPq and DAAD for the support during the preparation of the manuscript.

April 2005 Johann Baumeister
Frankfurt/Main

Antonio Leitão
Florianópolis

Contents

1 Introduction **1**

1.1 Inverse Problems 1

1.2 Ill-posedness 5

1.3 Contents 7

1.4 Bibliographical comments 8

1.5 Exercises 8

2 Basic concepts **10**

2.1 Ill-posedness in linear problems 10

2.1.1 Statement of the problem 10

2.1.2 Restoration of continuity 12

2.1.3 Compact operators 16

2.1.4 A first analysis of the method of Tikhonov . . 19

2.2 Regularization of ill-posed problems 23

2.2.1 The idea of regularization 23

2.2.2 A-priori regularizing strategies 26

2.2.3 L–curve for Tikhonov's method 30

2.3 Regularization in Hilbert scales 33

2.3.1 The method of Tikhonov in Hilbert scales . . . 33

2.3.2 Regularization by smoothing in the solution space 38

2.3.3 Regularization by smoothing in the data space . . 38

2.3.4 Severely ill-posed problems 41

2.3.5 Reconstruction of a functional 44

2.3.6 A-posteriori regularizing strategy 45

2.3.7 Regularization by discretization 47

2.4 Finite dimensional problems 49

2.4.1 Ill-conditioned problems 50
2.4.2 Least squares 53
2.5 Bibliographical comments 56
2.6 Exercises 56

3 Iterative methods 59
3.1 Introduction 60
3.2 Landweber iteration for linear equations 62
3.3 The nonlinear Landweber iteration 67
3.4 A modified Landweber iteration 73
3.5 Asymptotical regularization 79
3.6 Bibliographical comments 85
3.7 Exercises 85

4 Some inverse problems of convolution type 87
4.1 Deconvolution 87
4.1.1 Introduction 87
4.1.2 Stability and Regularization in the state space 91
4.1.3 Iterative deconvolution 95
4.1.4 On discretization in the state space 97
4.2 Convolution in Fourier space 99
4.2.1 Some results for the Fourier transform 99
4.2.2 Stability and regularization in the Fourier space 102
4.3 Autoconvolution and autocorrelation 107
4.4 Blind deconvolution 110
4.5 Bibliographical comments 112
4.6 Exercises 113

5 Tomography Problems 115
5.1 Computerized Tomography 115
5.1.1 Problem formulation 115
5.1.2 Computational aspects: the Fourier technique . 120
5.1.3 Computational aspects: filtered backprojection 123
5.1.4 Computational aspects: the ART-algorithm . . 124
5.2 Features in images 127
5.2.1 The Radon transform for shape detecting . . . 128
5.2.2 The Hough accumulator for detecting lines . . 130
5.3 Discrete Tomography 133

5.3.1 Statement of the problem 133
5.3.2 A stability result 136
5.4 Bibliographical comments 137
5.5 Exercises 138

6 Level set methods 140
6.1 Introduction 141
6.2 First level set approach for inverse problems 143
6.3 Level sets and asymptotic regularization 148
6.4 Level sets and constraint optimization 152
6.4.1 Introduction 152
6.4.2 Derivation of the Level Set Method 154
6.4.3 Relation to Shape Optimization 159
6.5 Applications 162
6.5.1 The inverse potential problem 162
6.5.2 Identification of doping profiles 167
6.6 Bibliographical comments 173
6.7 Exercises 174

Appendix: Basic facts in functional analysis 176
A.1 The Schwartz space 176
A.2 Hilbert spaces 177
A.3 Hilbert scales 178
A.4 Frechét derivative 183

Bibliography 185

List of Figures

1.1 A process 2
1.2 Error balance 5

2.1 L–curve 31

5.1 The beam geometry 116
5.2 A lattice in $\mathbb{Z}^2$ 134

6.1 Infinitesimal variation of the level set curve $\phi(x) = 0$. . 144
6.2 Inverse potential problem: Initial condition and exact solution. 165
6.3 Inverse potential problem: Evolution of the level set method. 166
6.4 Inverse potential problem: Evolution of the level set function. 167
6.5 Inverse doping profile: Example of a p-n diode. 169
6.6 Inverse doping profile: Setup of the numerical experiment. 172
6.7 Inverse doping profile: Error evolution for the level set method. 173

List of Tables

6.1 Inverse potential problem: Algorithm for one iterative step of the level set method. 164
6.2 Inverse doping profile: Problem formulation. 171

Chapter 1

Introduction

In this introduction we illustrate the wide range of inverse problems and give a first insight into the problems of solving inverse problems.

1.1 Inverse Problems

The problem which may be considered as one of the oldest inverse problem is the computation of the diameter of the earth by Eratosthenes in 200 b. Chr.. For many centuries people are searching for hiding places by tapping walls and analyzing echo; this is a particular case of an inverse problem. It was Heisenberg who conjectured that quantum interaction was totally characterized by its scattering matrix which collects information of the interaction at infinity. The discovery of neutrinos by measuring consequences of its existence is in the spirit of inverse problems too.

Over the past 30 years, the number of publications on inverse problems has grown rapidly. The following list of inverse problems gives a good impression of the wide variety of applications:

- the inverse problem of geomagnetic induction;
- X-ray tomography, ultrasound tomography, laser tomography;
- acoustic scattering, scattering in quantum mechanics;

- radio-astronomical imaging, image analysis;
- locating cracks or mines by electrical prospecting;
- seismic exploration, seismic tomography;
- the use of electrocardiography and magneto-cardiography;
- evolution backwards in time, inverse heat conduction;
- the inverse problem of potential theory;
- "can you hear the shape of a drum/manifold?"
- deconvolution, reconstruction of truncated signals;
- compartmental analysis, parameter identification;
- data assimilation;
- determing the volatility in models for financial markets;
- Discrete tomography, shape from probing.

Suppose that we have a mathematical model of a physical process. We assume that this model gives a description of the system behind the process and its operating conditions and explains the principal quantities of the model:

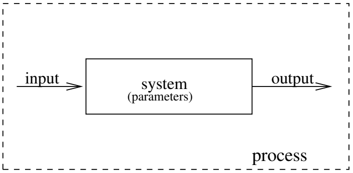

Figure 1.1: A process

input, system parameters, output

In most cases the description of the system is given in terms of a set of equations (ordinary and/or partial differential equations, integral equations, ...), containing certain parameters. The analysis of the given physical process via the mathematical model may be separated into three distinct types of problems; see Figure 1.1.

(A) **The direct problem**. Given the input and the system parameter, find out the output of the model.

(B) **The reconstruction problem**. Given the system parameters and the output, find out which input has led to this output.

(C) **The identification problem**. Given the input and the output, determine the system parameters which are in agreement with the relation between input and output.

We call a problem of type (A) a direct (or forward) problem since it is oriented towards a cause-effect sequence. In this sense problems of type (B) and (C) are called inverse problems because they are problems of finding out unknown causes of known consequences. It is immediately clear that the solution of one of the problems above involves a treatment of the other problems as well. A complete discussion of the model by solving the inverse problems is the main objective of inverse modelling.

Let us give a mathematical description of the input, the output and the systems in functional analytic terms.

X	: space of input quantities;
Y	: space of output quantities;
$\mathbb{P}$	: space of system parameters;
$A(p)$	: system operator from X into Y associated to $p \in \mathbb{P}$.

In these terms we may formulate the problems above in the following way:

(A) Given $x \in X$ and $p \in \mathbb{P}$, find $y := A(p)x$.

(B) Given $y \in Y$ and $p \in \mathbb{P}$, find $x \in X$ such that $A(p)x = y$.

(C) Given $y \in Y$ and $x \in X$, find $p \in \mathbb{P}$ such that $A(p)x = y$.

At first glance, the direct problem (A) seems to be solved much easier than the inverse problems (B), (C). However, for the computation of $y := A(p)x$ it may be necessary to solve a differential or integral equation, a task which may be of the same complexity as the solution of the equations in the inverse problems.

Example 1.1.1 (Differentiation of data). *We consider the problem of finding the integral of a given function. This is done analytically and numerically in a very stable way. When this problem is* *considered as a direct (forward) problem then to differentiate a given function is the related inverse problem. A mathematical description is given as follows:*

> **Direct Problem:** *With a continuous function $x : [0,1] \longrightarrow \mathbb{R}$ compute*
>
> $$y(t) := \int_0^t x(s)ds\,,\; t \in [0,1]\,.$$
>
> **Inverse Problem:** *Given a differentiable function $y : [0,1] \longrightarrow \mathbb{R}$ determine $x := y'$.*

We are interested in the inverse problem. Since y should be considered as the result of measurements then the data y are noisy and we may not expect that the noisy data $\tilde{y}$ are continuously differentiable. Therefore, the inverse problem has no obvious solution. Moreover, the problem should not be formulated in the space of continuous functions since perturbations due to noise lead to functions which are not continuous.

The differentiation of (measured) data is involved in many inverse problems. In a mechanical system one may ask for hidden forces. Since Newton's law relates forces to velocities and accelerations one has to differentiate observed data. We will see that in the problem of X-ray tomography differentiation is implicitly present too. □

In certain simple examples inverse problems can be converted formally into a direct problem. For example, if A has a known inverse then the reconstruction problem is solved by $x := A^{-1}y$. However, the explicit determination of the inverse does not help if the output y is not in the domain of definition of A^{-1}. This situation is typical in applications due to the fact that the output may only be imprecisely known and/or distorted by noise.

In the linear case, that is if $A(p)$ is a linear map for every $p \in \mathbb{P}$, problem (B) has been studied extensively and its theory is well-developed. The situation in the nonlinear case is somewhat less satisfactory. Linearization is very successful to find an acceptable solution to a nonlinear problem but in general, this principle provides only a partial answer.

The identification problem (C) in a general setting is rather difficult since it is in almost all cases a (highly) nonlinear problem with many local solutions. Moreover, the input signal may only be available imcompletely only.

1.2 Ill-posedness

Inverse modelling involves the estimation of the solution of an equation from a set of observed data. The theory falls into two distinct parts. One deals with the ideal case in which the data are supposed to be known exactly and completly (*perfect data*). The other treats the practical problems that are created by incomplete and imprecise data (*imperfect data*). It might be thought that an exact solution to an inverse problem with perfect data would prove also useful for the practical case. But it turns out in inverse problems that the solution obtained by the analytic formula is very sensitive to the way in which the data set is completed and to errors in it.

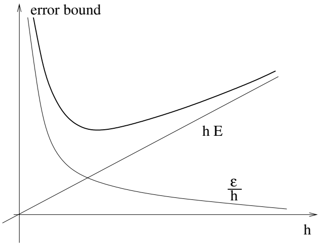

Figure 1.2: Error balance

In a complete solution of inverse problems the questions of **existence, uniqueness, stability** and **construction** are to be considered. The question of existence and uniqueness is of great importance in testing the assumption behind any mathematical model. If the answer in the uniqueness question is no, then we know that even perfect data do not contain enough information to recover the physical quantity to be estimated. In the question of stability we have to decide wether the solution depends continuously on the data. Stability is necessary if we want to be sure that a variation of the given data in a sufficiently small range leads to an arbitrarily small change in the solution. This concept was introduced by *Hadamard* in 1902 in connection with the study of boundary value problems for partial differential equations and he designated unstable problems *ill-posed*[1]. The nature of inverse problems (irreversibility, causality, unmodelled structures, ...) leads to ill-posedness as a characteristic property of these problems.

When solving ill-posed problems numerically, we must certainly expect some difficulties, since any errors act as a perturbation on the original equation and so may cause arbitrarily large variations in the solution. Observational errors have the same effect. Since errors cannot be completely avoided, there may be a range of plausible solutions and we have to find out a reasonable solution. These ambiguities in the solution of inverse problems which are unstable can be reduced by incorporating some sort of *a-priori* information that limits the class of allowable solutions. By a-priori information we mean an information which has been obtained independently of the observed values of the data. This a-priori information may be given as a deterministic or a statistical information. We shall restrict ourselves to deterministic considerations.

Let us present a first example of an ill-posed problem. This example will be considered again and again in this monograph.

Example 1.2.1 (Differentiation of data). *Suppose that we have for the continuous function* $y : [0,1] \longrightarrow \mathbb{R}$ *a measured function* $y^\varepsilon : [0,1] \longrightarrow \mathbb{R}$ *which is contaminated by noise in the following sense:*

$$|y^\varepsilon(t) - y(t)| \le \varepsilon \textit{ for all } t \in [0,1]\,.$$

It is reasonable to try to reconstruct the derivative $x := y'$ *of* y *at* $\tau \in (0,1)$ *by*

$$x^{\varepsilon,h}(\tau) := \frac{y^\varepsilon(\tau+h) - y^\varepsilon(\tau-h)}{2h}\,.$$

We obtain

$$\begin{aligned} |x^{\varepsilon,h}(\tau) - x(\tau)| \quad &\le \quad |\frac{y(\tau+h)-y(\tau-h)}{2h} - x(\tau)| \\ &\qquad + |\frac{(y^\varepsilon - y)(\tau+h) - (y^\varepsilon - y)(\tau-h)}{2h}| \end{aligned}$$

[1] Hadamard believed – many mathematicians still do – that ill-posed problems are actually incorrectely posed and artificial in that they would not describe physical systems. He was wrong!

If we know a bound

$$|x'(t)| \leq E \textit{ for all } t \in [0,1]\,,$$

then we get, roughly estimating,

$$|x^{\varepsilon,h}(\tau) - x(\tau)| \leq hE + \frac{\varepsilon}{h}\,. \tag{1.1}$$

Now it is clear that the best what we can do is to balance the terms on the right hand side of the bound:

$$h(\varepsilon) := E^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\,.$$

(It is assumed that $\tau \pm h) \in [0,1]$*.) This gives*

$$|x^{\varepsilon,h(\varepsilon)}(\tau) - x(\tau)| \leq 2E^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\,. \tag{1.2}$$

The diagram 1.2 which is a graphical presentation of the bound (1.1) *is typical for approximations in ill-posed problems. There are two terms in the error estimates: a term due to approximation of the inverse mapping and a term due to measurement error. The balance of these two terms gives an "optimal" reconstruction result. Thus, in contrast to well-posed problems, it is not the best to discretize finer and finer. One may consider ill-posed problems under the motto* "When the imprecise is preciser".[2] □

1.3 Contents

The lecture notes is organized as follows: In Chapter 1 we begin with a study of the basic concepts for stability and regularization. Here the Tikhonov regularization is a central theme. Chapter 2 is devoted to the (iterative) Landweber method and its applications. Chapter 3 deals with inverse problems of convolution type. Here we are confronted for the first time with nonlinear problems. In Chapter 4 we study some inverse problems in tomography. The subject of Chapter 5 are level set methods which have become important in solving problems where the boundary of sets has be reconstructed.

[2] This is the title of [52].

1.4 Bibliographical comments

In the 1970's, the monograph of Tikhonov, Arsenin [91] is in some sense the starting point of a systematic study of inverse problems. Nowadays, there exists a tremendous amount of literature on several aspects of inverse problems and ill-posedness. Instead of giving a complete list of relevant contributions we mention only some monographs [5, 23, 31, 53, 74] and survey articles [37, 95].

1.5 Exercises

1.1. Find a polynomial p with coefficients in $\mathbb{C}$ with given zeros $\xi_1, \dots, \xi_n$. When this problem is considered as an inverse problem, what is the formulation of the direct problem?

1.2. The problem of computing the eigenvalues of given matrix is well known. If this problem is considered as a direct problem what can be the formulation of the inverse problem?

1.3. Show that under the assumption

"$|x''(t)| \leq E$ for all $t \in [0,1]$"

the inequality (1.1) can be improved and an estimate

$$|x^{\varepsilon,h(\varepsilon)}(\tau) - x(\tau)| \leq cE^{1/3}\varepsilon^{2/3}$$

is possible (c is a constant independent of ε, E).

1.4. A model for the growth of a population is the law

$$u' = qu$$

where u represents the size of the population and q is a growth coefficient. Find a method to reconstruct q from the observation $u : [0,1] \longrightarrow \mathbb{R}$ when q is a time dependent function from $[0,1]$ into $\mathbb{R}$.

1.5. *Can you hear the length of string?*
Consider the boundary value problem

$$u'' = f\,,\ u(0) = u(l) = 0$$

where $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a given continuous function. Suppose that the solution u and f are known. Find the length l of the interval.

1.6. Consider the boundary value problem

$$u'' + qu = f\ ,\ u(0) = u(l) = 0$$

where $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a given continuous function. Find sufficient conditions on f such that q can be computed from an observation $u(\tau)$ for some point $\tau \in (0, l)$.

Chapter 2

Basic concepts

This chapter is intended to describe the basic concepts of solving inverse problems in a stable way. The method of Tikhonov is one of the central themes, compact operators are discussed since they are involved in the main applications.

2.1 Ill-posedness in linear problems

2.1.1 Statement of the problem

Let X, Y be Hilbert spaces[1] endowed with inner products $\langle\cdot,\cdot\rangle_X$ and $\langle\cdot,\cdot\rangle_Y$ respectively; the resulting norms in X and Y are denoted by $\|\cdot\|_X$ and $\|\cdot\|_Y$. Let $A : X \longrightarrow Y$ be a linear mapping with adjoint mapping A^* . We consider the linear equation

$$Ax = y \tag{2.1}$$

as a model equation for a linear inverse problem. As it is well known, A has a bounded inverse when A is bijective. Therefore we should not assume that A is bijective when we want to study the problems related to ill-posedness inherent in inverse problems. The assumption

[1] Some steps of our considerations could be done also in the context of Banach spaces which are no Hilbert spaces. But the Hilbert space setting is necessary when we use differentiability of the norm, orthogonality, projections,. . .

that A is injective is not a serious restriction since we can always factor out in X the null space of A. As a consequence we should give up the condition that A is surjective. Therefore we should operate under the following assumption:

Assumption A0

> $A : X \longrightarrow Y$, $A^* : Y \longrightarrow X$ linear, injective, bounded;
> range(A) dense in Y, range(A) $\neq Y$;
> range(A^*) dense in X, range(A^*) $\neq X$.

Indeed, under the assumption **A0** the inverse A^{-1} : range(A) $\longrightarrow X$ cannot be continuous since the following situation can be realized:

> For every $y \in Y\backslash$range(A) there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X with $\lim_n Ax_n = y$ and $(x_n)_{n\in\mathbb{N}}$ is divergent.

Clearly, assumption **A0** makes sense for the case $\dim \text{range}(A) = \infty$ only; the remaining case will be considered in Section 2.4. Notice that the list of assumptions in **A0** contain redundant informations. For example injectivity of A^* follows when range(A) is dense in Y. Especially, when A is a compact operator (see below) then the condition "$A : X \longrightarrow Y$ linear, injective, bounded, range(A) dense in Y" implies the other conditions.

Example 2.1.1 (Differentiation of data). *We set:*

$$\begin{aligned} X, Y &:= L_2[0,1] \textit{ endowed with the usual inner product};\\ (Ax)(t) &:= \int_0^t x(s)ds\,,\ t\in[0,1]\,,\ x\in L_2[0,1]\,. \end{aligned}$$

We want to solve the equation

$$\int_0^t x(s)ds = y(t)\,,\ t\in[0,1]\,. \qquad (2.2)$$

Clearly, every function in the range of A is continuous. Therefore range(A) $\nsubseteq Y$. Since every continuously differentiable function y is in *the range of A, we obtain that range(A) is dense in $L_2[0,1]$. The difficulty to solve the equation in a stable way becomes clear if one considers y^ε defined by $y^\varepsilon(t) := \varepsilon\sin(t\varepsilon^{-2})$. Then $x^\varepsilon(t) := \varepsilon^{-1}\cos(t\varepsilon^{-2})$ solves the equation and when ε is small then y^ε is small in norm and x^ε is large in norm.*

The equation (2.2) *is an example of an* integral equation of the first kind*, equations which are – considered on the interval* $[0,1]$ *– of the following form:*

$$\int_0^1 \kappa(t,s)x(s)ds = y(t)\,,\, t \in [0,1]\,; \tag{2.3}$$

κ is called a kernel function*. For* (2.2) *the kernel κ is given by*

$$\kappa(t,s) := \begin{cases} 1 & , \textit{if } s \le t \\ 0 & , \textit{if } s > t \end{cases}.$$

We see that the operator A is smoothing: each L_2–function becomes continuous (actually differentiable in a weak sense). This fact indicates the difficulties to solve integral equations of the first kind when the kernel is sufficiently smooth: the range of the integral operator is a "small" subset in the image space. Integral equations of the second kind like

$$x(t) + \int_0^1 \kappa(t,s)x(s)ds = y(t)\,,\, t \in [0,1]\,,$$

don't have this smoothing property when κ is not degenerate. □

2.1.2 Restoration of continuity

Suppose that **A0** holds. Let x^0 be the (unique) solution of (2.1) with right hand side y^0 :

$$Ax^0 = y^0\,. \tag{2.4}$$

In practice, y^0 is never known exactly but only up to an error. In a simple additive model for the perturbation of the data we may assume that for y^0 a distorted data $y^\varepsilon \in Y$ is available , satisfying

$$y^\varepsilon = y^0 + w^\varepsilon, w^\varepsilon \in Y, \|w^\varepsilon\|_Y \le \varepsilon \tag{2.5}$$

where $\varepsilon \geq 0$ is the so-called *noise level.* The problem consists in finding a reasonable approximation x^ε for x^0 using the data y^ε. Since we may not assume $y^\varepsilon \in \mathrm{range}(A)$ we cannot define x^ε as the solution of (2.1) for $y := y^\varepsilon$. But it is reasonable to reformulate the reconstruction problem in the following way:

$$\text{Find } x^\varepsilon \in X \text{ which satisfies } \|Ax^\varepsilon - y^\varepsilon\|_Y \leq \varepsilon \tag{2.6}$$

If $x^\varepsilon \in X$ satisfies (2.6) we see that the defect $Ax^\varepsilon - y^\varepsilon$ has the same order as the error $y^\varepsilon - y^0$ but the goal is to find a good approximation of x^0 . However, since A^{-1} is unbounded, the set

$$M(\varepsilon) := \{x \in X \mid \|Ax - y^\varepsilon\|_Y \leq \varepsilon\}$$

is not necessarily bounded and we may expect that an element x^ε which satisfies (2.6) may be no good approximation for x^0. In order to shrink the “solution set” $M(\varepsilon)$ we introduce the restriction that a solution should belong to a given subset K of X and K represents some a-priori information. Such a set is called a *source set* and “$x \in K$” is called a *source condition.* Obviously, such a restriction set K should have the property that the mapping $A^{-1} : A(K) \longrightarrow X$ is continuous and that $x^0 \in K$. If such a set K is chosen then we may reformulate the problem (2.6) in the following way:

$$\text{Find } x^\varepsilon \in M_K(\varepsilon) := \{x \in X | x \in K, \|Ax - y^\varepsilon\|_Y \leq \varepsilon\}\,. \tag{2.7}$$

We are interested in an estimation of $\|x^\varepsilon - x^0\|_X$, where $x^\varepsilon \in M_K(\varepsilon)$ is arbitrary. The worst case is described by

$$\|x^\varepsilon - x^0\|_X \leq \sigma_K(\varepsilon) := \sup\{\|x^1 - x^2\|_X \mid x^1, x^2 \in M_K(\varepsilon)\}$$

where $\sigma_K(\varepsilon)$ is the diameter of $M_K(\varepsilon)$.

A rather general choice for K is given in the following way:[2]

Choose a linear closed operator $B : \mathcal{D}_B \longrightarrow Z$ where $\mathcal{D}_B$ is a dense subset of X and Z is a Hilbert space. Define

$$\begin{aligned} K &:= K_E := \{v \in \mathcal{D}_B \mid \|Bv\|_Z \leq E\} \;\; (E \geq 0)\,, \\ M(\varepsilon, E) &:= \{x \in \mathcal{D}_B \mid \|Ax - y^\varepsilon\|_Y \leq \varepsilon, \|Bx\|_Z \leq E\}\,. \end{aligned}$$

[2] A linear mapping $B : \mathcal{D}_B \longrightarrow Z, \mathcal{D}_B \subset X$, is closed when its graph is closed in $X \times Z$.

Many different choices are possible for B, according to the prior knowledge. The simplest one is $B = I_X :=$ *identity* on X. Another usual choice is to take B as a differential operator and then the bound $\|Bx\|_Z \leq E$ is a smoothness requirement on the solution; see example 2.1.2 below.

The following quantities are of interest when we want to estimate the error x^ε for a particular choice $x^\varepsilon \in M(\varepsilon, E)$. We define

$$\begin{aligned}
\sigma(\varepsilon, E) &:= \sigma_{K_E}(\varepsilon)\,, \\
\omega(\varepsilon, E) &:= \sup\{\|x\|_X \mid z \in \mathcal{D}_B, \|Ax\|_Y \leq \varepsilon, \|Bz\|_Z \leq E\}\,, \\
\nu(\tau, B) &:= \sup\{\|x\|_X \mid z \in \mathcal{D}_B, \|Ax\|_Y \leq \tau, \|Bz\|_Z \leq 1\}\,.
\end{aligned}$$

By simple arguments we obtain

$$\sigma(\varepsilon, E) \leq 2\omega(\varepsilon, E) = 2E\nu(\frac{\varepsilon}{E}, B)\,. \tag{2.8}$$

The quantity

$$\text{SNR} := \frac{E}{\varepsilon}$$

which shows up in (2.8) is called the *signal to noise–ratio.*

Example 2.1.2 (Differentiation of data). *We want to solve the equation*

$$(Ax)(t) := \int_0^t x(s)ds = y(t)\,,\ t \in [0,1]\,, \tag{2.9}$$

for $x, y \in X := Y := L_2[0,1]$. *Our choice of the a-priori information is a bound on the norm of the first derivative of the function to be found:*

$$\int_0^1 |x'(s)|^2 ds \leq E^2\,.$$

Therefore we set[3]

$$\begin{aligned}
\mathcal{D}_B &:= H_0^1[0,1] := \{v \in AC[0,1] | v(1) = 0, v' \in L_2[0,1]\}\,, \\
Z &:= X, \\
(Bx)(t) &:= x'(t)\,,\ t \in [0,1]\,,\ x \in \mathcal{D}_B\,.
\end{aligned}$$

[3] $AC[0,1]$ denotes the space of absolutely continuous functions on $[0,1]$.

Clearly, the adjoint mapping is given by

$$A^* : L_2[0,1] \ni y \longmapsto (A^*y)(\cdot) := -\int_{\cdot}^{1} y(s)ds \in L_2[0,1]\,.$$

It is easy to give an realistic estimate for the key quantity $\nu(\cdot, B)$. *Let* $x \in \mathcal{D}_B$ *with* $\|Ax\|_Y \le \tau, \|Bx\|_Z \le 1$, $\tau \ge 0$. *Then*

$$\begin{aligned} \|x\|_X^2 &= \int_0^1 x(t)x(t)dt \;=\; -\int_0^1 x'(t)\int_0^t x(s)ds\,dt \\ &= -\int_0^1 x'(t)Ax(t)dt \le \|Bx\|_X\|Ax\|_Y \le \tau\,. \end{aligned}$$

This shows $\nu(\tau, B) = \tau^{\frac{1}{2}}$ *for all* $\tau \ge 0$.
Notice that the source condition "$v \in \mathcal{D}_B, \|Bx\|_X \le 1$*" can be summarized by the property*

$$v \in \mathcal{D}_B,\ \text{there exists } y \in Y \text{ with } x = A^*y, \|y\|_Y \le 1\,.$$

The following lemma shows that such a formulation of a source condition is helpful in our general context. □

Theorem 2.1.3. *Let* $A : X \longrightarrow Y$ *be a linear bounded operator.*

(a) If $x = A^*y$ *with* $\|y\|_Y \le E$ *and* $\|Ax\| \le \tau$ *then*

$$\|x\| \le E^{\frac{1}{2}}\,\tau^{\frac{1}{2}}\,. \tag{2.10}$$

(b) If $x = A^*Az$ *with* $\|z\|_X \le E$ *and* $\|Ax\| \le \tau$ *then*

$$\|x\| \le E^{1/3}\tau^{2/3}\,. \tag{2.11}$$

Proof:
Ad (a). Follows from

$$\|x\|_X^2 = \langle x, A^*y\rangle = \langle Ax, y\rangle \le \|Ax\|_X\|y\|_Y \le \tau E\,.$$

Ad (b). We have

$$\begin{aligned} \|x\|_X^2 &= \langle x, A^*Az\rangle \;=\; \langle Ax, Az\rangle \\ &\le \tau\|Az\| \;=\; \tau\langle Az, Az\rangle^{\frac{1}{2}} \;\le\; \tau\langle z, x\rangle^{\frac{1}{2}} \le \tau E^{\frac{1}{2}}\|x\|_X^{\frac{1}{2}}\,. \end{aligned}$$

This proves (b). ■

The interpretation of (a) in Theorem 2.1.3 is that the inverse of

$$A\,|_{M_E}\colon M_E \longrightarrow A(M_E)$$

is continuous in $y = \theta$ where $M_E := \{x \in X \mid x = A^*z, \|z\| \leq E\}$. Notice that the bound in (b) is better due to a more stringent assumption on x .

Remark 2.1.4. *The restriction set $K = K_E$ is for a reasonable chosen B a bounded set in X . In some cases it is reasonable to choose K as a cone which describes restrictions like $f \geq 0, f' \geq 0, \dots$. Such restrictions are called* descriptive constraints. *They are not so easy to handle since a cone is far from being a bounded set.* □

2.1.3 Compact operators

There is class of problems which can be considered as the generic case of an ill-posed problem, namely the solution of linear equations which are governed by compact operators. Since equations governed by compact operators are very important in applications we introduce some facts concerning these operators.

Definition 2.1.5. *Let $A : X \longrightarrow Y$ be a linear operator between infinite dimensional Hilbert spaces X, Y . Then A is called a* compact operator *if A maps the unit ball B_1^X in X into the subset $A(B_1^X)$ of Y whose closure is compact.* □

As a rule, integral operators with a smooth kernel function and defined on functions of finite support are compact operators. For a compact operator one has a very powerful "normal form", as we will show next; for the proof of this normal form we refer to the literature.

Theorem 2.1.6 (Singular value decomposition). *Let $A : X \to Y$ be an injective compact operator and assume that X is infinite dimensional. Then there exist sequences $(e^j)_{j\in\mathbb{N}}, (f^j)_{j\in\mathbb{N}}, (\sigma_j)_{j\in\mathbb{N}}$, called a* singular system, *such that the following assertions hold:*

(a) $e^j \in X, f^j \in Y$ *for all* $j \in \mathbb{N}$;

(b) $\sigma_j \in \mathbb{R}, 0 < \sigma_{j+1} < \sigma_j$ *for all* $j \in \mathbb{N}$, $\lim_j \sigma_j = 0$;

(c) $\langle e^j, e^k \rangle = 0, \langle f^j, f^k \rangle = 0$ *for all* $j, k \in \mathbb{N}, j \neq k$;

(d) $Ae^j = \sigma_j f^j, A^* f^j = \sigma_j e^j$ *for all* $j \in \mathbb{N}$;

(e) $Ax = \sum_{j=1}^{\infty} \sigma_j \langle x, e^j \rangle f^j$ *for all* $x \in X$,
$A^* y = \sum_{j=1}^{\infty} \sigma_j \langle y, f^j \rangle e^j$ *for all* $y \in Y$.

The singular value decomposition reflects the ill-posedness when solving $Ax = y$ with a compact operator A:

$$\lim_j \|\sigma_j^{-\frac{1}{2}} e^j\|_X = \infty \, , \; \lim_j \|A(\sigma_j^{-\frac{1}{2}} e^j)\|_Y = 0 \, .$$

The decay of the singular values σ_j is some measure of the degree of ill-posedness; we come back to this question.

Remark 2.1.7. *Notice that a bounded linear operator is compact when a system* $(e^j, f^j, \sigma_j)_{j \in \mathbb{N}}$ *exists such that the assertions* $(a) \ldots (e)$ *in Theorem 2.1.6 hold true.* □

Using the singular value decomposition one obtains a very useful characterization of the range of a compact operator.

Lemma 2.1.8 (Picard's criterion). *Let* $A : X \longrightarrow Y$ *be an injective compact operator with a singular system* $(e^j, f^j, \sigma_j)_{j \in \mathbb{N}}$. *Then the equation* $Ax = y$ *is solvable if and only if*

$$\sum_{j=1}^{\infty} \sigma_j^{-2} |\langle y, f^j \rangle|^2 < \infty \, , \tag{2.12}$$

in which case the solution is given by

$$x = A^{-1} y = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle y, f^j \rangle e^j \, . \tag{2.13}$$

Proof:
This follows immediately from Theorem 2.1.6. ■

Example 2.1.9 (Differentiation of data). *We consider the mapping*

$$A : L_2[0,1] \ni x \longmapsto (Ax)(\cdot) := \int_0^{\cdot} x(s)ds \in L_2[0,1]\,. \qquad (2.14)$$

This is a compact operator by the theorem of Arzela-Ascoli.
The equation $A^*Ae = \sigma e$ is equivalent with the boundary value problem

$$\sigma e'' + e = 0\,,\; e(1) = 0, e'(0) = 0\,,$$

and we obtain for A the singular system $(e^j, f^j, \sigma_j)_{j\in N}$ where

$$\sigma_j := \frac{2}{(2j-1)\pi}\,,\; e^j(s) := \sqrt{2}\cos(\sigma_j^{-1}s)\,,\; f^j(s) := \sqrt{2}\sin(\sigma_j^{-1}s)\,,$$

with $s \in [0,1]$. ■

Remark 2.1.10. *Let $A : X \longrightarrow Y$ be an injective compact operator with a singular system $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$. When $y = u + v \in range(A) + range(A)^\perp$ then*

$$\sum_{j=1}^{\infty} \sigma_j^{-1}\langle y, f^j\rangle e^j = \sum_{j=1}^{\infty} \sigma_j^{-1}\langle u, f^j\rangle e^j$$

is well defined. Therefore we may consider the mapping

$$A^- : range(A) + range(A)^\perp \ni u + v \longmapsto \sum_{j=1}^{\infty} \sigma_j^{-1}\langle u, f^j\rangle e^j\,.$$

This mapping is called the pseudoinverse *of A and we set $A^\dagger := A^-$. One may generalize the definition of the pseudoinverse to the case of an arbitrary linear bounded operator $A : X \longrightarrow Y$. The first observation is that the equation*

$$A^*Ax = A^*y$$

has a uniquely defined solution $x^\dagger \in range(A^*)$ for each $y \in range(A)$ $+range(A)^\perp$ which is called the minimal norm solution *of $Ax = y$. This defines the pseudoinverse $A^\dagger$ as the mapping*

$$A^\dagger : range(A) + range(A)^\perp \ni y \longmapsto x^\dagger \in X\,;$$

see also Subsection 2.4.2 □

2.1.4 A first analysis of the method of Tikhonov

To solve the problem (2.7) we have to specify a method which chooses an element $x^\varepsilon \in X$ satisfying $x^\varepsilon \in K, \|Ax^\varepsilon - y^\varepsilon\|_Y \le \varepsilon$. Two methods come into mind immediately:

The method of residuals:

Minimize $\|Bx\|_Z$ **subject to** $x \in \mathcal{D}_B, \|Ax - y^\varepsilon\|_Y \le \varepsilon.$

The method of quasisolutions:

Minimize $\|Ax - y^\varepsilon\|_Y$ **subject to** $x \in \mathcal{D}_B$ **and** $\|Bx\|_Z \le E\,.$

Due to the fact that $x^0 \in M(\varepsilon, E)$ we have for a solution $x^{E,\mathrm{rs}}$ of the method of residuals

$$\|Bx^{E,\mathrm{rs}}\|_Z \le \|Bx^0\|_Z \le E, \|Ax^{E,\mathrm{rs}} - y^\varepsilon\|_Y \le \varepsilon, x^{E,\mathrm{rs}} \in M(\varepsilon, E)\,. \tag{2.15}$$

Analogous we have for a solution $x^{\varepsilon,\mathrm{qs}}$ of the method of quasisolutions

$$\|Bx^{\varepsilon,\mathrm{qs}}\|_Z \le E, \|Ax^{\varepsilon,\mathrm{qs}} - y^\varepsilon\|_Y \le \|Ax^0 - y^\varepsilon\|_Y \le \varepsilon, x^{\varepsilon,\mathrm{qs}} \in M(\varepsilon, E)\,. \tag{2.16}$$

If we consider the methods above as optimization problems the theory of Langrangian multipliers leads us to the following compromise between these methods:

The method of Tikhonov:

Minimize $F(x) := \|Ax - y^\varepsilon\|_Y^2 + \dfrac{\varepsilon^2}{E^2}\|Bx\|_Z^2$ **subject to** $x \in \mathcal{D}_B\,.$

The case “$\mathcal{D}_B = X, B = I_X :=$ *identity* on X” is called the *classical method of Tikhonov.*

If x^ε is a solution of the method of Tikhonov then due to $x^0 \in$

$M(\varepsilon, E)$ we obtain by evaluating F in x^0

$$\frac{\varepsilon^2}{E^2}\|Bx^\varepsilon\|_Z^2 \le F(x^\varepsilon) \le F(x^0) \le 2\varepsilon^2,$$

$$\|Ax^\varepsilon - y^\varepsilon\|_Y^2 \le F(x^\varepsilon) \le F(x^0) \le 2\varepsilon^2\,,$$

and we see

$$\|Bx^\varepsilon\|_Z \le \sqrt{2}E, \|Ax^\varepsilon - y^\varepsilon\|_Y \le \sqrt{2}\varepsilon, x^\varepsilon \in M(\sqrt{2}\varepsilon, \sqrt{2}E)\,. \quad (2.17)$$

The consequence is that we are sure that we lose at most a factor of $\sqrt{2}$ if we replace the method of residuals or the method of quasisolutions by Tikhonov's method.

If we don't know the number $\lambda := \varepsilon/E$ as it is typically the case in practice we may modify the method of Tikhonov in the following way:

The generalized method of Tikhonov:

Minimize $F_t(x) := \|Ax - y^\varepsilon\|_Y^2 + t\|Bx\|_Z^2$ **subject to** $x \in \mathcal{D}_B$.

Here t is a given positive number.

Let $A^* : Y \longrightarrow X$ and $B^* : Z \longrightarrow X$ be the adjoint operators of A and B respectively. The mapping $F_t : \mathcal{D}_B \longrightarrow \mathbb{R}$ is differentiable in each $x \in V$ and

$$F_t'(x)(h) = 2\langle Ax, Ah\rangle_Y + 2t\langle Bx, Bh\rangle_Z\,,\ h \in \mathcal{D}_B\,. \quad (2.18)$$

Lemma 2.1.11. *Let $x^{\varepsilon,t} \in \mathcal{D}_B$. Then following conditions are equivalent:*

(a) $F_t(x^{\varepsilon,t}) := \inf_{x\in\mathcal{D}_B} F_t(x)\,.$

(b) $x^{\varepsilon,t}$ *solves*

$$(A^*A + tB^*B)x^{\varepsilon,t} = A^*y^\varepsilon\,. \quad (2.19)$$

Proof:
$(a) \implies (b)$. Equation (2.19) is a consequence of $F_t'(x)(h) = 0$ for all $h \in \mathcal{D}_B$; see (2.18).
$(b) \implies (a)$. Follows from the identity

$$F_t(x) - F_t(x^{\varepsilon,t}) = 2\|Ax - Ax^{\varepsilon,t}\|_Y^2 + 2t\|Bx - Bx^{\varepsilon,t}\|_Z^2 \,,\ x \in \mathcal{D}_B \,, \tag{2.20}$$

which we conclude from (2.19). ■

Next, we have to ask the question whether a solution $x^{\varepsilon,t}$ exists. The following assumption which is motivated by (2.20) is helpful for a positive answer.

Assumption A1

There exists $c > 0$ such that

$$c\|x\|_X^2 \le \|Ax\|_Y^2 + \|Bx\|_Z^2 \,,\ x \in \mathcal{D}_B \,.$$

Notice that in the case of the classical method of Tikhonov this assumption is satisfied. The proof of the following lemma is left to the reader.

Lemma 2.1.12. *Under the assumption* **A1** *there exists for each $t > 0$ a uniquely determined solution $x^{\varepsilon,t}$ of the generalized method of Tikhonov.*

Theorem 2.1.13. *Suppose that assumption* **A1** *is satisfied and let $x^{\varepsilon,t}$ be the solution of the generalized method of Tikhonov. Then we have for each $t > 0$:*

$$\|x^{\varepsilon,t} - x^0\| \le \left(\frac{\varepsilon}{\sqrt{t}} + \|Bx^0\|\right)\ \nu(\sqrt{t}, B)\,. \tag{2.21}$$

Proof:
It is easy to verify that $x^{\varepsilon,t} - x^0 = u - tv$ with

$$u := (A^*A + tB^*B)^{-1}A^*(y^\varepsilon - y^0)\,,\ v := (A^*A + tB^*B)^{-1}B^*Bx^0\,.$$

We obtain

$$\begin{aligned}\langle y^\varepsilon - y^0, Au\rangle &= \langle A^*(y^\varepsilon - y^0), u\rangle \\ &= \langle (A^*A + tB^*B)u, u\rangle \ = \ \|Au\|_Y^2 + t\|Bu\|_Z^2\end{aligned}$$

and therefore[4]

$$\|Au\|_Y^2 + t\|Bu\|_X^2 \le \varepsilon\|Au\|_Y \le \frac{\varepsilon^2}{2} + \frac{1}{2}\|Au\|_X^2$$

which implies

$$\|Au\|_Y \le \varepsilon\,,\ \|Bu\|_Z \le \frac{\varepsilon}{\sqrt{t}}\,,\ \|u\|_X \le \frac{\varepsilon}{\sqrt{t}}\nu(\sqrt{t}, B)\,.$$

We have

$$\langle Bx^0, Bv\rangle = \langle B^*Bx^0, v\rangle = \langle (A^*A + tB^*B)v, v\rangle = \|Av\|_Y^2 + t\|Bv\|_Z^2$$

and therefore

$$\|Av\|_Y^2 + t\|Bv\|_Z^2 \le \|Bx^0\|_Z\|Bv\|_X \le \frac{1}{2t}\|Bx^0\|_Z^2 + \frac{t}{2}\|Bv\|_Z^2$$

which shows

$$\|Av\|_Y \le \frac{1}{t}\|Bx^0\|_Z,\ \|Bv\|_Z \le \frac{1}{t}\|Bx^0\|_Z,\ \|v\|_X \le \frac{\nu}{t}(\sqrt{t}, B)\|Bx^0\|_Z.$$

Now the inequality (2.21) is proved. ■

Corollary 2.1.14. *Suppose that assumption* **A1** *is satisfied and let* $\|Bx^0\|_Z \le E$. *Then we obtain with the (a-priori) parameter strategy* $t(\varepsilon) = \dfrac{\varepsilon^2}{E^2}$

$$\|x^{\varepsilon,t(\varepsilon)} - x^0\| \le 2\,E\,\nu\left(\frac{\varepsilon}{E}, B\right)\,. \tag{2.22}$$

Proof:
This follows immediately from (2.21).

To make the estimate in Theorem 2.1.13 and Corollary 2.1.14 more applicable we need to obtain an estimate of $\nu(\cdot, B)$. This is done in a specified situation in section 2.3.1.

[4] We use frequently the inequality $ab \le (1/2\mu)a^2 + (\mu/2)b^2$ which holds for $a, b \ge 0, \mu > 0$.

2.2 Regularization of ill-posed problems

Here we discuss the methods which are used to solve an ill-posed problem in a stable way. we do this again using assumption **A0**.

2.2.1 The idea of regularization

Definition 2.2.1. *Any family $(R_t)_{t>0}$ of mappings from X into Y is called a* recovery family. *A family $(R_t)_{t>0}$ of linear bounded operators from Y into X is called a* regularizing family *for A if*

$$\lim_{t\downarrow 0} R_t A x = x \textit{ for all } x \in X\,. \tag{2.23}$$

□

Obviously, a regularizing family $(R_t)_{t>0}$ is a family which should approximate A^{-1} . Since A^{-1} is an unbounded operator R_tA does not converge to the identity in the operator norm by the theorem of Banach–Steinhaus if t goes to zero. Moreover, the family $(\|R_t\|)_{t>0}$ cannot be bounded.[5]

Suppose that x^0, y^0, y^ε are given as in (2.93), (2.5). With a regularizing family $(R_t)_{t>0}$ for A we define

$$x^{\varepsilon,t} := R_t y^\varepsilon\,,\; x^{\varepsilon,0} := R_t y^0\,,\; t > 0\,. \tag{2.24}$$

We want to find the parameter t such that $R_t y^\varepsilon$ deals with the noise ε in an optimal fashion. Since the reconstruction $x^{\varepsilon,t} - x^0$ error can be decomposed as

$$\begin{aligned} \|x^{\varepsilon,t} - x^0\|_X &\leq \|R_t y^\varepsilon - R_t y^0\|_X + \|R_t A x^0 - x^0\|_X \\ &\leq \varepsilon\|R_t\| + \|R_t A x^0 - x^0\|_X \end{aligned} \tag{2.25}$$

we observe that two competing effects enter (2.25). The first one is the ill-posedness effect: as t goes to 0 the norm $\|R_t\|$ tends to ∞ ; so t should not be chosen too small. The second one is the regularizing effect: as t increases, R_tA becomes a less accurate approximation of the identity; so t should not be chosen too large. Only properly chosen

[5] Any operator norm is denoted by $\|\cdot\|$.

values of t will provide an optimal reconstruction. We discuss this problem in the next subsection. A convenient method to construct

a regularizing family is given by filtering when we have a singular system $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$ of a compact operator A. By Picard's lemma 2.1.8, for every $y \in \text{range}(A)$ $x := A^{-1}y$ can be written as

$$x = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle y, f^j \rangle e^j \,. \tag{2.26}$$

Since for $y \notin \text{range}(A)$ the series above does not converge due to Picard's lemma or the fact that $\lim_j \sigma_j = 0$ we have to introduce a *filter* for the small singular values. This can be done by a mapping $q : (0,\infty) \times (0,\sigma_1] \longrightarrow \mathbb{R}$ which damps out the contribution of small singular values in the series (2.26). With such a filter q we define a potential candidate for a regularizing family in the following way:

$$R_t y := \sum_{j=1}^{\infty} q(t,\sigma) \sigma_j^{-1} \langle y, f^j \rangle e^j \,,\; y \in Y \,. \tag{2.27}$$

The assumptions which the filter q should satisfy can be read off from the following estimates:

$$\begin{aligned} \|R_t y\|_X^2 &= \sum_{j=1}^{\infty} |q(t,\sigma_j)|^2 \sigma_j^{-2} |\langle y, f^j \rangle|^2 \\ &\leq \sup_{\sigma\in(0,\sigma_1]} |q(t,\sigma)\sigma^{-1}|^2 \sum_{j=1}^{\infty} |\langle y, f^j \rangle|^2 \\ &= \sup_{\sigma\in(0,\sigma_1]} |q(t,\sigma)\sigma^{-1}|^2 \|y\|_Y^2 \qquad (2.28) \\ \|R_t A x - x\|_X^2 &= \sum_{j=1}^{\infty} |q(t,\sigma_j) - 1|^2 |\langle x, e^j \rangle|^2 \,. \qquad (2.29) \end{aligned}$$

Theorem 2.2.2. *Let $A : X \longrightarrow Y$ be an injective compact operator with singular system $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$ and let $q : (0,\infty) \times (0,\sigma_1] \longrightarrow \mathbb{R}$ be a filter function which satisfies the following conditions:*

F1) $|q(t,\sigma)| \leq 1$ *for all* $t > 0, \sigma \in (0,\sigma_1]$.

F2) For all $t > 0$ there exists a constant $c(t)$ such that for all $\sigma \in (0, \sigma_1]$ $|q(t, \sigma)| \le c(t)\sigma$.

F3) $\lim_{t\to 0} q(t, \sigma) = 1$ *for every $\sigma \in (0, \sigma_1]$.*

Then the family $(R_t)_{t>0}$, defined in (2.27), is a regularizing family. Additionally, we have with $x^{\varepsilon,t} := \sum_{j=1}^{\infty} q(t,\sigma)\sigma_j^{-1}\langle y^\varepsilon, f^j\rangle e^j$:

$$\|x^{\varepsilon,t} - x^0\|^2 \;\le\; \|x^0\|_X^2 \sup_{\sigma\in(0,\sigma_1]} |q(t,\sigma) - 1|^2 + \varepsilon^2 c(t)^2 \,. \qquad (2.30)$$

Proof:
With the condition F2) we conclude from (2.28) $\|R_t\| \le c(t), t > 0$. Let $x \in X$ and let $\varepsilon > 0$. Then there exists a $N \in \mathbb{N}$ with

$$\sum_{j=N+1}^{\infty} |\langle x, e^j\rangle|^2 < \varepsilon^2 \,.$$

According to the condition F3) we can choose a constant $t_0 > 0$ such that

$$|q(t,\sigma_j) - 1|^2 < \varepsilon^2 \text{ for all } j = 1, \ldots, N \text{ and } 0 < t \le t_0 \,.$$

With the condition F1) we obtain for $t \in (0, t_0]$

$$\begin{aligned}\|R_t Ax - x\|_X^2 &= \sum_{j=1}^{N} |q(t,\sigma_j) - 1|^2 |\langle x, e^j\rangle|^2 \\ &\quad + \sum_{j=N+1}^{\infty} |q(t,\sigma_j) - 1|^2 |\langle x, e^j\rangle|^2 \\ &\le \varepsilon^2 \sum_{j=1}^{N} |\langle x, e^j\rangle|^2 + 4\varepsilon^2 \le \varepsilon^2 \|x\|_X^2 + 4\varepsilon^2 \,.\end{aligned}$$

This shows $\lim_{t\to 0} R_t Ax = x$. The estimate (2.30) follows immediately from (2.25), (2.28), (2.29). ■

In the most interesting cases the quantity $\sup_{\sigma\in(0,\sigma_1]} |q(t,\sigma) - 1|$ can be expressed in a more compact form, independent of the (unknown) number σ_1 .

Example 2.2.3.
Truncation*:*

$$q(t,\sigma) := \begin{cases} 1 & ,\sigma^2 \geq t \\ 0 & ,\sigma^2 \leq t \end{cases} .$$

This is a filter which truncates the contribution of singular values larger than the threshold parameter $\sqrt{t}$ *. Here we can verify* $c(t) = 1/\sqrt{t}$ *. Moreover, the condition* $F3)$ *holds in a stronger form:*

$$F3') \qquad |q(t,\sigma)-1| \leq \frac{\sqrt{t}}{\sigma} \text{ or } |q(t,\sigma)-1| \leq \frac{t}{\sigma^2} \text{ for all } t>0, \sigma>0 .$$

Tikhonov*:*

$$q(t,\sigma) := \frac{\sigma^2}{\sigma^2+t} .$$

This is the filter which models the classical method of Tikhonov. Here we have $c(t) = 1/(2\sqrt{t})$ *and the condition* $F3')$ *holds again in the stronger form* $F3')$ *.*[6]

□

Under the assumption **A1** the mapping

$$R_t : Y \ni y \longmapsto (A^*A + tB^*B)^{-1}A^*y^\varepsilon \in X$$

is well defined. In the special case

$$A \text{ compact}, \mathcal{D}_B := X, B := identity$$

the method of Tikhonov is a regularizing family. This follows with Example 2.2.3 from the fact that with a singular system $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$ the operator R_t is defined by

$$R_t y := \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\sigma_j^2+t}\sigma_j^{-1}\langle y, f^j\rangle e^j \, , \, y \in Y .$$

2.2.2 A-priori regularizing strategies

Suppose we have a regularizing family $(R_t)_{t>0}$ for A. Then the candidates for a solution of $Ax = y^\varepsilon$ are defined by

$$x^{\varepsilon,t} := R_t y^\varepsilon \, , \, t > 0 . \tag{2.31}$$

[6] To simplify some expressions we do not always use the best possible constants.

The main problem is now to choose a parameter $t = t(\varepsilon)$ in such a way that under appropriate conditions $x^{\varepsilon,t}$ is the "best" approximation of x^0 . We may distinguish two main different strategies:

a-priori-strategy — a-posteriori-strategy

A-posteriori strategies try to find the "best" regularization parameter from the results $x^{\varepsilon,t}$ by using the given data y^ε (and the noise level ε). Methods will be discussed in Subsection 2.2.3 and 2.3.6.
An a-priori parameter choice strategy tries to determine the regularization parameter before doing numerical computation. Such an a-priori strategy usually starts from the estimate (2.25) for the error $\|x^{\varepsilon,t} - x^0\|_X$:

$$\|x^{\varepsilon,t} - x^0\|_X \le \varepsilon\|R_t\| + \|R_tAx^0 - x^0\|_X \, . \tag{2.32}$$

The best we could do is to minimize the right hand side in (2.32) with respect to t . In general, it is enough to balance the competing terms in (2.32): we choose $t(\varepsilon)$ such that

$$\|R_{t(\varepsilon)}Ax^0 - x^0\|_X = \varepsilon\|R_{t(\varepsilon)}\| \tag{2.33}$$

We have already used this idea when we derived the error estimate in Example 2.1.1.

Example 2.2.4.
In the case that A is a compact operator we obtained in Theorem 2.2.2 the following estimate

$$\|x^{\varepsilon,t} - x^0\|^2 \;\le\; \|x^0\|_X^2 \sup_{\sigma\in(0,\sigma_1]} |q(t,\sigma) - 1|^2 + \varepsilon^2 c(t)^2 \tag{2.34}$$

with a filter function q . In Example 2.2.3 we could specify this estimate for the regularizing family related to the idea of truncation in the following form:

$$\|x^{\varepsilon,t} - x^0\|^2 \;\le\; \|x^0\|_X^2 \left(\sup_{\sigma\in(0,\sigma_1]} \frac{t}{\sigma^2}\right) + \frac{\varepsilon^2}{t} \, , \, t > 0 \, .$$

Since $\sup_{\sigma\in(0,\sigma_1]} \frac{t}{\sigma^2} = \infty$ *for all $t > 0$ it is not possible to balance the terms. We see in Theorem 2.2.6 below what kind of a-priori information helps to avoid this situation.* □

Definition 2.2.5. *Let $(R_t)_{t>0}$ be a regularizing family for A. We say that a choice $t = t(\varepsilon)$ leads to a* regularizing scheme for x^0 *if we can prove:*

$$\lim_{\varepsilon \to 0} R_{t(\varepsilon)} w^\varepsilon = x^0 \textit{ for every } w^\varepsilon \in Y \textit{ with } \|w^\varepsilon - Ax^0\|_Y \le \varepsilon \,. \tag{2.35}$$

As a rule, in order to prove that a particular regularizing family is a regularizing scheme for x^0 one has to introduce a source condition on x^0.

Theorem 2.2.6. *Let A be an injective compact operator with singular system $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$ and let $q : (0,\infty) \times (0,\sigma_1] \longrightarrow \mathbb{R}$ be a filter function. Consider the regularizing family $(R_t)_{t>0}$ defined in* (2.27)*.*

(a) Under the assumptions

$$\begin{array}{rcl} |q(t,\sigma)| & \le & \dfrac{\sigma}{\sqrt{t}}, |q(t,\sigma) - 1| \le \dfrac{\sqrt{t}}{\sigma}, t > 0, \sigma > 0, \\ x^0 & = & A^* z \textit{ with } z \in Y, \end{array}$$

we choose (without loss of generality $z \ne \theta$)

$$t(\varepsilon) := \varepsilon \, \|z\|_Y^{-1} \tag{2.36}$$

and have

$$\|R_{t(\varepsilon)} y^\varepsilon - x^0\|_X \le \sqrt{2} \|z\|_Y^{\frac{1}{2}} \, \varepsilon^{\frac{1}{2}} \,. \tag{2.37}$$

(b) Under the assumption

$$\begin{array}{rcl} |q(t,\sigma)| & \le & \dfrac{\sigma}{\sqrt{t}}, |q(t,\sigma) - 1| \le \dfrac{t}{\sigma^2}, t > 0, \sigma > 0, \\ x^0 & = & A^* A z \textit{ with } z \in X, \end{array}$$

we choose (without loss of generality $z \ne \theta$)

$$t(\varepsilon) := \varepsilon^{\frac{2}{3}} \, \|z\|_X^{-\frac{2}{3}} \tag{2.38}$$

and have

$$\|R_{t(\varepsilon)} y^\varepsilon - x^0\|_X \le c \|z\|_X^{\frac{1}{3}} \, \varepsilon^{\frac{2}{3}} \,. \tag{2.39}$$

Additionally, in each case the given parameter choice strategy leads to a regularization scheme for x^0 .

Proof:
Ad (a). We have

$$\begin{aligned} \|R_t y^\varepsilon - x^0\|_X^2 &\leq \sum_{j=0}^\infty |q(t,\sigma_j)-1|^2 |\langle x^0, e^j\rangle|^2 + \|R_t\|^2\varepsilon^2 \\ &\leq \sum_{j=0}^\infty |q(t,\sigma_j)-1|^2 \sigma_j^2 |\langle z, f^j\rangle|^2 + t^{-1}\varepsilon^2 \\ &\leq \left(\sup_{\sigma>0} |q(t,\sigma)-1|^2|\sigma|^2\right) \|z\|_X^2 + t^{-1}\varepsilon^2 \\ &\leq t\|z\|_X^2 + t^{-1}\varepsilon^2 \,. \end{aligned}$$

With the choice (2.36) we obtain (2.37).
Ad (b). Follows by a similar argumentation.
By the estimates (2.37), (2.39) we see that each parameter choice strategy leads to regularization scheme for x^0 . ■

In Theorem 2.2.6 we have focus on the filter functions q of Example 2.2.3. One might expect that by improving the source condition better estimates are possible. Let us demonstrate that this is not the case with the filter function modelling the method of Tikhonov. We have here for $k \in \mathbb{N}$ and for $t > 0$:

$$\sup_{\sigma>0} |q(t,\sigma)-1||\sigma^k| = \sup_{\sigma>0} \frac{t\sigma^k}{\sigma^2+t} = \infty \text{ if } k > 2 \,.$$

Therefore a source condition

$$x^0 = (A^*A)^k z^0, k > 1,$$

leads to the same error estimate as in $b)$ of Theorem 2.2.6. This fact is called *order–nonoptimality* of this method.
On the other hand, for the filter function q which models the truncation method a source condition

$$x^0 = (A^*A)^k z^0, k > 1,$$

leads with the parameter choice

$$t(\varepsilon) := \frac{\varepsilon^{2/(k+1)}}{\|z^0\|_X^{2/(k+1)}}$$

to an estimate

$$\|R_{t(\varepsilon)} - x^0\|_X \leq c\varepsilon^{1/(k+1)}\|z^0\|_X^{1-1/(k+1)} .$$

For $k \to \infty$ one obtains asymptotically the order of a well-posed problem.

Remark 2.2.7. *The quality of a regularization family has to be compared with the quantity*

$$\Omega(\varepsilon, E) := \inf_{R:Y\to X} \sup\{\|x - Ry\|_X \mid x \in \mathcal{D}_B, \\ \|Bx\|_Z \leq E, \|Ax - y\|_Y \leq \varepsilon\} .$$

Since one can prove

$$\omega(\varepsilon, E) \leq \Omega(\varepsilon, E) \leq 2\omega(\varepsilon, E) ,$$

a reconstruction map $R : Y \longrightarrow X$ is called order–optimal *with respect to B if one has*

$$\|Ry - x^0\|_X = O(\omega(\varepsilon, E))$$

for all x^0, y with $x^0 \in \mathcal{D}_B, \|Bx^0\|_Z \leq E, \|Ax^0 - y\|_Y \leq \varepsilon$. □

2.2.3 L–curve for Tikhonov's method

Here we sketch a first idea of an a-posteriori parameter choice strategy. Let $(x^{\varepsilon,t})_{t>0}$ be the family of solutions of the generalized method of Tikhonov and set

$$u(t) := \|Ax^{\varepsilon,t} - y^\varepsilon\|_Y^2 \, , \, v(t) := \|Bx^{\varepsilon,t}\|_Z^2 \, .$$

Then it is easy to verify that $x^{\varepsilon,t}$ is a solution of the method of residuals with $\varepsilon = u(t)^{\frac{1}{2}}$ and $x^{\varepsilon,t}$ solves the method of quasisolutions with $E = v(t)^{\frac{1}{2}}$. Define

$$C := \{(a, b) \in \mathbb{R}^2 | \exists x \in \mathcal{D}_B \text{ with } \|Ax - y^\varepsilon\|_Y \leq a, \|Bx\|_Z \leq b\} .$$

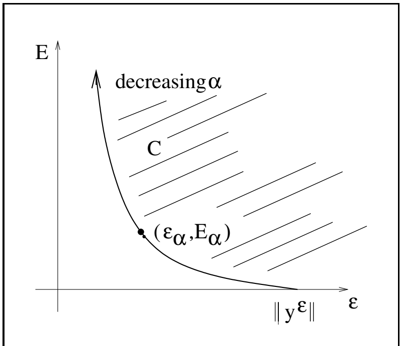

Figure 2.1: L–curve

Then it can be shown that $t \longmapsto \varepsilon_t$ is increasing, $t \longmapsto E_t$ is decreasing, C is a convex set and the curve $t \longmapsto (\varepsilon_t, E_t)$ is the boundary of C; see Figure 2.1.

If we do not know the number εE^{-1} we have to specify a method which determines t in an optimal way using the numbers $u(t), v(t)$. The L–curve selection criterion consists in locating the t–value which maximizes the curvature in the typical L-shaped plot of the curve

$$\Lambda : (0,\infty) \ni t \longmapsto (\ln(u(t)), \ln(v(t)) \in \mathbb{R}\,.$$

The motivation for doing so lies in the observation that the steep, almost vertical portion of the plot for very small values of t corresponds to rapidly varying, under-regularized solutions with very little change in $u(t)$, while the horizontal portion of larger values of t corresponds to over-regularized solutions where the plot is flat or slowly decreasing. The *L-curve corner* marks a natural transition point linking these two regions; we come back to this fact for the finite dimensional case from the computational point of view. Here we collect some results concerning the curve

$$L : (0,\infty) \ni t \longmapsto (u(t), v(t)) = (\|Ax^{\varepsilon,t} - y^\varepsilon\|_Y^2, \|Bx^{\varepsilon,t}\|_Z^2) \in \mathbb{R}^2\,.$$

To compute the derivative of the mappings u, v we start from

$$(A^*A + tB^*B)x^{\varepsilon,t} = A^*y^\varepsilon, t > 0\,. \tag{2.40}$$

We set

$$z^{\varepsilon,t} := \frac{d}{dt}x^{\varepsilon,t}\,,\ w^{\varepsilon,t} := \frac{d}{dt}z^{\varepsilon,t}\,,$$

and obtain in an obvious way:

$$(A^*A + tB^*B)z^{\varepsilon,t} + B^*Bx^{\varepsilon,t} = \theta, t > 0\,, \tag{2.41}$$

$$(A^*A + tB^*B)w^{\varepsilon,t} + 2B^*Bz^{\varepsilon,t} = \theta, t > 0\,. \tag{2.42}$$

Notice that (2.40), (2.41), (2.42) may be combined in a system for $x^{\varepsilon,t}, z^{\varepsilon,t}, w^{\varepsilon,t}$.

From these identities for $x^{\varepsilon,t}, z^{\varepsilon,t}, w^{\varepsilon,t}$ we conclude by using (2.40), (2.41), (2.42)

$$\begin{aligned}
u(t) &= \|Ax^{\varepsilon,t} - y^{\varepsilon}\|_Y^2 \,, t > 0 \,, && (2.43)\\
u'(t) &= -2t\langle B^*Bx^{\varepsilon,t}, z^{\varepsilon,t}\rangle && (2.44)\\
&= 2t\|(A^*A + tB^*B)^{\frac{1}{2}} z^{\varepsilon,t}\|_X^2 \,, t > 0 \,, &&\\
u''(t) &= -2\langle B^*Bx^{\varepsilon,t}, z^{\varepsilon,t}\rangle - 2t\langle B^*Bz^{\varepsilon,t}, z^{\varepsilon,t}\rangle && (2.45)\\
&\quad -2t\langle B^*Bx^{\varepsilon,t}, w^{\varepsilon,t}\rangle &&\\
&= 2\|(A^*A + tB^*B)^{\frac{1}{2}} z^{\varepsilon,t}\|_X^2 - 6t\|Bz^{\varepsilon,t}\|_Z^2 \,, t > 0, && (2.46)
\end{aligned}$$

and

$$\begin{aligned}
v(t) &= \|Bx^{\varepsilon,t}\|_Z^2, t > 0 \,, && (2.47)\\
v'(t) &= 2\langle B^*Bx^{\varepsilon,t}, z^{\varepsilon,t}\rangle \;=\; -\frac{1}{t}u'(t) \,, t > 0 \,, && (2.48)\\
v''(t) &= \frac{1}{t^2}u'(t) - \frac{1}{t}u''(t) \,, t > 0 \,. && (2.49)
\end{aligned}$$

Due to (2.44) and (2.48) $u'(t)$ is nonnegative and $v'(t)$ is nonpositive for all $t > 0$. Therefore u is monotone nondecreasing, v is monotone nonincreasing. $t > 0$. Therefore it makes no sense to optimize u or v without restriction in order to find a "best" parameter. Let us define

$$w(s) := u(\frac{1}{s}) \,,\ s > 0 \,.$$

Then it is easy to verify that

$$w''(s) = \frac{6}{s^4}\|Az^{\varepsilon,1/s}\|_Y^2 \,,\ s > 0 \,. \qquad (2.50)$$

This shows that w is strictly convex when $z^{\varepsilon,1/s} \neq 0$, a property which u does not have.

Now we want to compute the curvature κ of the curve L. We have

$$\kappa(t) = \frac{u'(t)v''(t) - u''(t)v'(t)}{(u'(t)^2 + v'(t)^2)^{\frac{3}{2}}} = \frac{u'(t)^2}{t^2(u'(t)^2 + v'(t)^2)^{\frac{3}{2}}}$$

and, therefore,

$$\begin{aligned}\kappa(t) &= -\frac{1}{v'(t)(t^2+1)^{\frac{3}{2}}}\,,\, t>0 & (2.51)\\ \kappa'(t) &= -\frac{v''(t)}{v'^2(t)(t^2+1)^{\frac{3}{2}}} + \frac{3t}{v'(t)(t^2+1)}\,,\ t>0\,, & (2.52)\end{aligned}$$

Due to the nonpositivity of $v'(t)$ we obtain that the curvature is positive. From (2.52) we conclude that $\kappa'(t)$ is nonnegative for small values of t. Since $v'(t)$ is nonpositive for all $t > 0$ we read off that $\kappa'(t)$ becomes negative for large values of t. Thus, there should be a point t_L where $\kappa(t)$ has its maximal value; $(u(t_L), v(t_L))$ is called the *"corner"* of the curve L. The denotation L comes from the fact that the curve L is L–shaped with the corner in $(u(t_L), v(t_L))$.

A potential candidate for such a corner point can be computed by setting $\kappa'(t) = 0$ which leads to the equation

$$\ln\left(\frac{v''(t)}{v'(t)}\right) = \int_0^t 3s(s^2+1)^{\frac{1}{2}}ds = (t^2+1)^{\frac{3}{2}} - 1$$

and we obtain for the candidate t_L the condition

$$v''(t_L) = v'(t_L)e^{(t_L^2+1)^{\frac{3}{2}}-1}\,. \qquad (2.53)$$

This observation suggests to find the parameter t_L as a solution of (2.53).

2.3 Regularization in Hilbert scales

In this section we exploit functional analytic tools to design regularization methods which can be used for a wide spectrum of problems.

2.3.1 The method of Tikhonov in Hilbert scales

As a consequence of the assumption **A0** an estimate

$$a\|Ax\|_Y \geq \|x\|_X\,,\, x\in X \qquad (2.54)$$

with $a > 0$ cannot hold.[7] We may see this as a consequence of the fact that the norm $\| \cdot \|_X$ in X is too strong in order to allow a continuous inverse A^{-1} . Therefore we are searching for a weaker norm – actually for a larger space endowed with a weaker norm – such that an inequality of the kind (2.54) holds.

Let V be a Hilbert space which is densely embedded in X . As it is well known, X may be considered as a dense subspace of the dual space V^* of V when we identify X with its dual space X^* . This leads to a Hilbert space triple $V \subset X \subset V^*$; such a triple is called a *Gelfand triple*. In general, in such a Gelfand triple the following interpolation inequality holds:

$$\|x\|_X \leq \|x\|_V^{\frac{1}{2}} \, \|x\|_{V^*}^{\frac{1}{2}} \tag{2.55}$$

Assumption A2'

Let $V \subset X \subset V^*$ be a Gelfand triple with

(a) $x^0 \in V$,

(b) there exists $a > 0$ with $a\|Ax\|_Y \geq \|x\|_{V^*}$, $x \in V^*$.

The assumption **A2'** contains two *basic ingredients*: (a) is a *source condition* and describes the "smoothness" of x^0 ; (b) says that the continuity of inverse A^{-1} is continuous as a mapping from the range of A into V^* .

Under the assumption **A2'** the method of Tikhonov should be changed into

The method of Tikhonov revisited:

Minimize $G_t(x) := \|Ax - y^\varepsilon\|_Y^2 + t\|x\|_V^2$ **subject to** $x \in V$.

Here t is a given positive number. Along the arguments in the proof of Lemma 2.1.12 we obtain existence and uniqueness of a minimizer $x^{\varepsilon,t} \in V$ of G_t for each $t > 0$ (set $\mathcal{D}_B := V, Z := V, Bv := v, v \in V$).

[7] If A is a compact operator the assertion is a generic one.

This method is applicable if we can find a Hilbert space V satisfying assumption **A2'**.

Here is the good news: there exists in all cases (under the assumption **A0**) such a space. It is constructed along the following steps:

- $X \ni x \longmapsto \|Ax\|_Y \in \mathbb{R}$ defines a norm $\|\cdot\|_\sim$ in X ;
- define W as the completion of X in the norm $\|\cdot\|_\sim$;
- W is a Hilbert space with a scalar product $\langle\cdot,\cdot\rangle_W$;
- we have $\langle x, x'\rangle_W = \langle Ax, Ax'\rangle_Y$ for all $x, x' \in X$;
- we set $V := W^*$ where W^* is the dual space of W ;
- $\|Ax\|_Y = \|x\|_{V^*}, x \in X$;
- one has $V \subset X \subset V^* = W$, $\|x\|_X \leq \|x\|_V^{\frac{1}{2}} \|x\|_{V^*}^{\frac{1}{2}}$ for all $x \in V$.

We sketch the proof of the interpolation inequality in the last step. Let $x \in X$ and suppose for the moment $x = A^*z$. Then

$$\begin{aligned}
\|x\|_X^2 &= \langle A^*z, A^*z\rangle \;\leq\; \|AA^*z\|_Y\|z\|_Y \\
&= \|A^*z\|_W\|z\|_Y \;=\; \|x\|_W\|z\|_Y \,, \\
\|z\|_Y &= \sup\{\langle u, z\rangle \mid u \in Y, \|u\|_Y \leq 1\} \\
&= \sup\{\langle Av, z\rangle \mid v \in X, \|Av\|_Y \leq 1\} \\
&= \sup\{\langle v, A^*z\rangle \mid v \in X, \|v\|_W \leq 1\} \;=\; \|A^*z\|_V \;=\; \|x\|_V \,.
\end{aligned}$$

Since range(A^*) is dense in X, the interpolation inequality holds for all $x \in X$.

Notation: We denote the space V by $H_X(A)$ and V^* by $H_X(A)^*$.

Example 2.3.1. *Let $A : X \longrightarrow Y$ be an injective compact operator* *with singular system* $(e^j, f^j, \sigma_j)_{j\in\mathbb{N}}$ *. Then we have*[8]

$$\begin{aligned} H_X(A) &= \{z \mid \sum_{j=0}^{\infty} \sigma_j^{-2} |\langle z, e^j\rangle|^2 < \infty\}\,, \\ H_X(A)^* &= \{z \mid \sum_{j=0}^{\infty} \sigma_j^{2} |\langle z, e^j\rangle|^2 < \infty\} \\ \|Ax\|_Y^2 &= \sum_{j=0}^{\infty} \sigma_j^{2} |\langle z, e^j\rangle|^2 \;=\; \|x\|^2_{H_X(A)^*}, x \in X\,. \end{aligned}$$

□

Remark 2.3.2. *By completing the space* Y *in the norm*

$$\|\cdot\|_\sim : Y \ni y \longmapsto \|A^*y\|_X \in \mathbb{R}$$

one obtains a space W *which is actually a Hilbert space. Then by*

$$H_Y(A) := W^* =: U \subset Y \subset U^* = W =: H_Y(A)^*$$

a Gelfand triple is defined. U *is now the space in which the equation* $Ax = y$ *can be solved. This principle was introduced by Lions as the so called HUM-method; see [63].* □

Here is the bad news: the Gelfand triple $H(A) \subset X \subset H(A)^*$, introduced above, is given in a way, that is not very constructive. Moreover, well known spaces like Sobolev spaces are defined very differently and the question arises: can well known Hilbert spaces be related to this triple? We consider the case of a Hilbert scale $(H_s)_{s\in\mathbb{R}}$; see appendix.

Assumption A2:

Let $(H_s)_{s\in\mathbb{R}}$ be a Hilbert scale with $X = H_0$ and
(a) there exists $q > 0$ such that $\mathcal{D}_B \subset H_q$ and $x^0 \in \mathcal{D}_B$;
(b) there exists $b > 0$ such that $b\|Bx\|_Z \geq \|x\|_q$, $x \in \mathcal{D}_B$;
(c) there exists $a > 0$ such that $a\|Ax\|_Y \geq \|x\|_{-p}$, $x \in X$.

[8] The definition of $H_X(A)^*$ is a little bit sloppy. Actually, $H_X(A)^*$ is the completion of X in the norm $z \longmapsto (\sum_{j=0}^{\infty} \sigma_j^2 |\langle z, e^j\rangle|^2)^{\frac{1}{2}}$.

We say that A is *related to the scale* $(H_s)_{s\in\mathbb{R}}$ if the condition (c) of assumption **A2** holds for no $p' < p$. If this is the case we say that *the degree of ill-posedness* in solving $Ax = y$ is p with respect to the scale $(H_s)_{s\in\mathbb{R}}$. If it is not possible to find such a parameter p then one says that the problem to solve $Ax = y$ is *severely ill-posed* with respect to the scale $(H_s)_{s\in\mathbb{R}}$. In Subsection 2.3.4 we will discuss such problems.

Under the assumption **A2** the generalized method of Tikhonov can be used; see Subsection 2.1.4. Due to the inequality

$$\|x\|_X \le \|x\|_{-p}^{\frac{q}{p+q}} \|x\|_q^{\frac{p}{p+q}} \,,\, x \in \mathcal{D}_B \,, \tag{2.56}$$

the assumption **A1** can be replaced by

$$\|x\|_X \le (a\|Ax\|_Y)^{\frac{q}{p+q}} (b\|Bx\|_Z)^{\frac{p}{p+q}} \,,\, x \in \mathcal{D}_B \,, \tag{2.57}$$

Therefore for each $t > 0$ a solution $x^{\varepsilon,t} \in \mathcal{D}_B$ exists with

$$x^{\varepsilon,t} = (A^*A + tB^*B)^{-1}A^*y^\varepsilon, t > 0\,. \tag{2.58}$$

Theorem 2.3.3. *Let the assumption* **A2** *be true and let us consider* $x^{\varepsilon,t} := (A^*A + B^*B)^{-1}A^*y^\varepsilon, t > 0$. *Then we have with* $\|Bx\|_Z \le E$*:*

$$\|x^{\varepsilon,t(\varepsilon)} - x^0\| \le 2\, a^{\frac{q}{p+q}}\, b^{\frac{p}{p+q}}\, E^{\frac{p}{p+q}}\, \varepsilon^{\frac{q}{p+q}} \quad \textit{for } t(\varepsilon) := \frac{\varepsilon^2}{E^2}\,. \tag{2.59}$$

Proof:
Due to Theorem 2.1.13 it is enough to estimate

$$\nu(\tau, B) := \sup\{\|x\|_X \mid x \in \mathcal{D}_B, \|Ax\|_Y \le \tau, \|Bx\|_Z \le 1\}\,.$$

Let $x \in \mathcal{D}_B$ with $\|Ax\|_Y \le \tau, \|Bx\|_Z \le 1$. Then

$$\|u\|_X \le \|u\|_q^{\frac{p}{p+q}} \|u\|_{-p}^{\frac{q}{p+q}} \le (b\|Bu\|_Z)^{\frac{p}{p+q}} (a\|Au\|_Y)^{\frac{q}{p+q}}$$
$$\le a^{\frac{q}{p+q}} b^{\frac{p}{p+q}}\, \tau^{\frac{q}{p+q}}\,.$$

■

2.3.2 Regularization by smoothing in the solution space

Now we want to look on the method of Tikhonov from a different point of view. We do this in the framework of Hilbert scales. Suppose that assumption **A2** holds. Then we can continue the inverse A^{-1} from the range(A) to Y by the following definition

$$A^- : Y \longrightarrow H_{-p}, A^- y := \lim_n A^{-1} y_n \text{ if } y = \lim_n y_n, y_n \in \text{range}(A).$$

Hence, we may define $\tilde{x} := A^- y^\varepsilon$ and consider $\tilde{x}$ as an approximation for x^0 . But $\tilde{x}$ may be contained in $H_{-p} \backslash X$. We can repair this defect by introducing a linear bounded "smoothing" map $Q_t : H_{-p} \longrightarrow X$. Then with

$$x^{\varepsilon,t} := Q_t A^- y^\varepsilon \, , \, x^{0,t} := Q_t A^- y^0$$

we obtain in the usual way an error estimate

$$\|x^{\varepsilon,t} - x^0\|_X \quad \leq \quad \|Q_t^X A^-\| \varepsilon + \|(Q_t^X - I)x^0\|_X$$

The composition $Q_t A^-$ may be considered as an *approximate inverse.* Since the family $(Q_t^X)_{t>0}$ is considered as a family of smoothing operators we should require that $\lim_{t \to 0} a(x^0; t) = 0$ where $a(x^0; t)$ $:= \|(Q_t^X - I)x^0\|_X$. Then, since A^{-1} : range(A) $\longrightarrow$ X is unbounded, we conclude $\lim_{t\to 0} c(t) = \infty$ where $c(t) := \|Q_t^X A^-\|$.

How to find such a smoothing operator? When we use the Hilbert scale $(H_s)_{s\in\mathbb{R}}$ with the generator $(A^*A)^{-1}$(see appendix) then we may use:

$$p = \frac{1}{2}, A^- = (A^*A)^{-1}A^*, Q_t := (t(A^*A)^{-1} + I)^{-1} \, ;$$

$(Q_t)_{t>0}$ is the family of resolvents of the unbounded operator $(A^*A)^{-1}$ and A^- is a left-inverse of A. In this context, the smoothing method is just the classical method of Tikhonov as we conclude from

$$(A^*A)^{-1}(t(A^*A)^{-1} + I)^{-1}A^*y = (tI + A^*A)^{-1}A^*y \, , \, y \in Y \, . \quad (2.60)$$

2.3.3 Regularization by smoothing in the data space

Consider the solution of equation $Ax = y$ and look at this equation in the following form:

$$AA^*z = y \, , \, x := A^*z \, .$$

This leads to the regularizing family $(R_t)_{t>0}$ where R_t is defined as follows:

$$R_t y := A^*(tI + AA^*)^{-1} y = A^*(AA^*)^{-1}(t(AA^*)^{-1} + I)^{-1} y \,,\; y \in Y\,. \tag{2.61}$$

R_t is a composition of the smoothing operator $(t(AA^*)^{-1} + I)^{-1}$ and the right-inverse $A^*(AA^*)^{-1}$.

Again this method is not related to a Hilbert scale which may be appropriate for applications. Therefore one should generalize the idea of smoothing the data as follows:

Smoothing in the data space

Take a family of mappings $Q_t : Y \longrightarrow \text{range}(A), t > 0$, and consider the family

$$R_t : Y \ni y \longmapsto x^{\varepsilon,t} := A^{-1} Q_t y^\varepsilon \in X, t > 0\,.$$

In this setting this method is called usually a *mollification method*. Data smoothing is an important tool in the context of partial differential equations. Here locally integrable functions are mollified to C^∞–functions independent of the domain. In general, mollification operators are of convolution type; see Chapter 4. Again, the composition $A^{-1}Q_t$ may be considered as an *approximate inverse*.

In order to implement such a method a very detailed study of the analytical properties of the operator A is necessary since we have to know the characteristic properties of the range of A; Lemma 2.1.8 (Picard's lemma) may be helpful. With

$$x^{\varepsilon,t} := A^{-1} Q_t y^\varepsilon \,,\; x^{\varepsilon,0} := A^{-1} Q_t y^0$$

we obtain an error estimate with two competing terms:

$$\|x^{\varepsilon,t} - x^0\|_X \le \|A^{-1}Q_t\|\varepsilon + \|A^{-1}(Q_t - I)Ax^0\|_X\,.$$

Example 2.3.4 (Differentiation of data). *Once again we consider the problem of differentiation of data. For convenience we do this in the interval* $[-\pi, \pi]$*. Set* $X := Y := L_2[-\pi, \pi]$ *and consider*

$$A : X \longrightarrow Y\,,\; Ax(t) := \int_{-\pi}^{t} x(s)ds, t \in [-\pi, \pi]\,. \tag{2.62}$$

We mollify the data y^ε by the Vallée Poussin kernel:

$$y^{\varepsilon,n}(t) := \tfrac{1}{2\pi n}\int_{-\pi}^{\pi} \frac{\cos((n+1)(t-s)) - \cos((2n+1)(t-s))}{\sin^2((t-s)/2)} y^\varepsilon(s)ds\,. \tag{2.63}$$

The function $y^{\varepsilon,n}$ is a trigonometric polynomial of order $2n$ and therefore the calculation of its derivative is very simple and can be done in a stable way; especially $y^{\varepsilon,n}(\cdot)$ is in the domain of A^{-1}.
Here the mollification parameter is discrete. But this is no serious drawback as we see in the following estimation of the error. We set

$$x^{\varepsilon,n} := A^{-1}y^{\varepsilon,n} = y^{\varepsilon,n\,\prime}\,,\ x^{0,n} := A^{-1}y^{0,n} = y^{0,n\,\prime}$$

and have

$$\|x^{\varepsilon,n} - x^0\|_X \le \|y^{\varepsilon,n\,\prime} - y^{0,n\,\prime}\|_X + \|y^{0,n\,\prime} - y^0\|_X\,.$$

From deep results concerning the approximation properties of the Vallée Poussin kernel one obtains the estimate

$$\|x^{\varepsilon,n} - x^0\|_X \le c_1 n\varepsilon + c_2 E n^{-1} \tag{2.64}$$

where c_1, c_2 are constants independent of n, x^0, ε, E and E comes from the a-priori information

$$x^0 \in AC[-\pi,\pi], x^{0\,\prime} \in X, \|x^{0\,\prime}\|_X \le E\,.$$

The choice $n(\varepsilon) := \lfloor \varepsilon^{-\frac{1}{2}} \rfloor$ yields the result

$$\|x^{\varepsilon,n} - x^0\|_X \le (c_1 E + c_2)\sqrt{\varepsilon}\,.$$

(When E is known the parameter n should be chosen such that the competing terms in the estimate (2.64) are in balance.) ■

Remark 2.3.5. *It is important to realize that a numerical realization $\widetilde{Q_t y^\varepsilon}$ of $Q_t y^\varepsilon$ has to take into account, that we have to stay in range(A).* □

2.3.4 Severely ill-posed problems

What can we do when an estimate

$$a\|Ax\| \geq \|x\|_{-p}, x \in H_{-p}$$

is not possible for any $p > 0$? For example, consider the integral equation

$$(Ax)(t) := \int_0^1 e^{-st}x(s)ds = y(t), t \in [0,1]\,,$$

for $X := Y := L_2[0,1]$. To solve this integral equation of the first kind is severely ill-posed with respect to the usual scale of Sobolev spaces. This is indicated by the fact that the operator A is smoothing of infinite type: Ax is infinitely differentiable for each $x \in L_2[0,1]$.

Example 2.3.6. *Consider with functions φ, ψ*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} \ , t \in \mathbb{R}, x \in (0,1), \\ u(0,t) &= \varphi(t), \ \frac{\partial u}{\partial x}(0,t) \ = \ 0, \ u(1,t) = \psi(t) \ , t \in \mathbb{R}\,. \end{aligned}$$

To solve the forward problem we want to compute φ from the data ψ : $\varphi := A\psi$. The inverse problem which is called the non-characteristic Cauchy problem for the heat equation *consists in finding ψ from the data φ : $\psi := A^{-1}\varphi$. The solution of the noncharacteristic Cauchy problem can be written formally as*

$$u(x,t) = [\cosh(x\sqrt{\frac{\partial}{\partial t}})\varphi](t) = \sum_{k=0}^{\infty} \frac{x^{2n}}{(2n!)} \frac{d^n \varphi}{dt^n}(t)\,,\ x \in (0,1), t > 0\,.$$

To compute ψ we have to use all derivatives of φ :

$$A^{-1}\varphi = \sum_{k=0}^{\infty} \frac{1}{(2n!)} \frac{d^n \varphi}{dt^n}\,.$$

□

Consider A where A^{-1} has a presentation by a formal power series

$$A^{-1} = \sum_{k=0}^{\infty} \alpha_k T^k \,, \tag{2.65}$$

and T is the generator of the scale $(H_s)_{s\in\mathbb{R}}$; see Example 2.3.6. In this case the domain of definition of A^{-1} should be contained in $H_\infty := \cap_{s>0} H_s$ and we should try to find a subspace of V of H_∞ such that $a\|Ax\| \geq \|x\|_{V^*}, x \in V^*$, holds. To find such a subspace we choose a sequence $\alpha := (\alpha_k)_{k\in\mathbb{N}_0}$ with

$$\alpha_k \geq 0, k \in \mathbb{N}_0 \,, \sum_{k=0}^{\infty} \alpha_k \leq 1 \,, \limsup_k \alpha_k^{1/k} = 0 \,. \tag{2.66}$$

We identify with such a sequence α the power series

$$\sum_{k=0}^{\infty} \alpha_k r^k \tag{2.67}$$

which converges for all $r \in \mathbb{R}$ due to assumption (2.66). Consider

$$Z(\alpha) \;:=\; \{x \in H_\infty \mid \sum_{k=0}^{\infty} \alpha_k \|x\|_k^2 < \infty\} \,. \tag{2.68}$$

$$\|x\|_\alpha^2 \;:=\; \sum_{k=0}^{\infty} \alpha_k \|x\|_k^2 < \infty \,. \tag{2.69}$$

$$\langle x, x\rangle_\alpha \;=\; \sum_{k=0}^{\infty} \alpha_k \langle x, x'\rangle_k \,. \tag{2.70}$$

$\langle\cdot,\cdot\rangle_\alpha$ is an inner product in $Z(\alpha)$ where $\langle\cdot,\cdot\rangle_k$ is the inner product in H_k . Since $Te^j = \lambda_j e^j, j = 1, \ldots,$ we obtain

$$\sum_{k=0}^{\infty} \alpha_k \|e^j\|_k^2 = \sum_{k=0}^{\infty} \alpha_k \|T^k e^j\|_0^2 = \sum_{k=0}^{\infty} \alpha_k \lambda_j^{2k} < \infty \,.$$

Therefore $\lambda_j \in Z(\alpha)$ for all $j \in \mathbb{N}_0$ and $Z(\alpha)$ is not empty; actually $Z(\alpha)$ is dense in $X = H_0$. We identify H_0 with its dual space H_0^* and obtain a Gelfand triple

$$V = Z(\alpha) \subset H_0 = X \subset Z(\alpha)^* =: V^* \,.$$

Now A^{-1} is well defined on $Z(\alpha)$ and

$$\begin{aligned} \|Ax\| &\geq \|x\|_{V^*}, x \in V^*, &(2.71)\\ \|x\|_0 &\leq \|x\|_{Z(\alpha)}^{\frac{1}{2}} \|x\|_{Z(\alpha)^*}^{\frac{1}{2}}, x \in Z(\alpha), &(2.72) \end{aligned}$$

can be verified. The last assertion follows from

$$\begin{aligned} \|x\|_0^2 &= \langle Ax, A^{-1}x\rangle = \sum_{k=0}^{\infty} \alpha_k \langle Ax, T^k x\rangle \\ &\leq \|Ax\|_Y (\sum_{k=0}^{\infty} \alpha_k^{\frac{1}{2}} \alpha_k^{\frac{1}{2}} \|T^k x\|_0) \\ &\leq \|Ax\|_Y (\sum_{k=0}^{\infty} \alpha_k)^{\frac{1}{2}} \|x\|_{Z(\alpha)} . \end{aligned}$$

The appropriate source condition is

$$x \in Z(\alpha), \|z\|_{Z(\alpha)} \leq E$$

and we have with $B : Z(\alpha) \ni x \longmapsto x \in Z(\alpha)$

$$\nu(\tau, B) := \{x \in X \mid x \in Z(\alpha), \|z\|_{Z(\alpha)} \leq E, \|Ax\|_Y \leq \tau\} \leq \tau^{\frac{1}{2}} .$$

Example 2.3.7. *The sequence*

$$\alpha \textit{ with } \alpha_k := \frac{1}{(2k)!}, k \in \mathbb{N}_0$$

satisfies the assumptions (2.66) *and the space $Z(\alpha)$ is appropriate for the solution of the non-characteristic Cauchy problem; see Example 2.3.6. Unfortunately the resulting source condition is very strong. It results from the fact that we consider classical solutions of the initial-boundary value problem with the heat operator. For example, a function*

$$\varphi(t) := \begin{cases} 1 & , \textit{ if } t \in [a,b] \\ 0 & , \textit{ if } t \notin [a,b] \end{cases}$$

$(0 < a < b)$ cannot be treated. □

2.3.5 Reconstruction of a functional

In some applications one might be satisfied to know just a functional of the solution x^0 . Due to the Riesz mapping the dual space X^* can be identified with X .

Recovery of a functional

> Given a functional $\mu \in X$. Find an approximation ξ^ε for $\xi^0 := \langle \mu, x^0 \rangle$ using the data y^ε .

Suppose we know ϕ with

$$A^*\phi = \mu\,. \tag{2.73}$$

Then

$$\xi^0 = \langle \mu, x^0 \rangle = \langle A^*\phi, x^0 \rangle = \langle \phi, Ax^0 \rangle = \langle \phi, y^0 \rangle$$

and it is reasonable to use

$$\xi^\varepsilon := \langle \phi, y^\varepsilon \rangle$$

as an approximation for ξ^0 . Since the right hand side μ is independent of the data y^ε the solution ϕ of (2.73) can be precomputed and ξ^ε is found by computing the scalar product $\langle \phi, y^\varepsilon \rangle$.

Unfortunately, the solution of the equation

$$A^*\phi = \mu$$

is again an ill-posed problem and all the machinery of regularization has to be used in order to compute a stable approximation ϕ^η when μ is given by an approximation μ^η only. For instance, such a case is given when the functional μ represents the measurement of an observed quantity of x^0 at some time τ which is corrupted by noise.

In many cases one tries to evaluate the unknown solution x^0 by a functional μ which does not belong to the dual space X^* which is identified here with X . Dirac type functionals are of this kind and a method which is related to these functionals is called the method of *Backus–Gilbert.* Therefore there is no chance to solve the equation (2.73). In the framework of Hilbert scales $H_s)_{s\in\mathbb{R}}$ we may consider the following problem:

Given $\mu \in H_{-p}$, find $\psi \in Y$ with
$\|A^*\psi - \mu\|_{-p} = \min_{w\in Y} \|A^*w - \mu\|_{-p}$.

Then with the solution ψ we define again $\xi^\varepsilon := \langle \psi, y^\varepsilon\rangle, \xi^0 := \langle \psi, y^0\rangle$ and obtain

$$\begin{aligned}|\xi^\varepsilon - \xi^0| &\leq |\langle \psi, y^\varepsilon - y^0\rangle| + |\langle A^*\psi - \mu, x^0\rangle|\psi, y^\varepsilon - y^0\rangle| \\ &\leq \|\psi\|\varepsilon + \|A^*\psi - \mu\|_{-p}\|x^0\|_p\end{aligned}$$

when we know that $x^0 \in H_p$. Define

$$\delta_R := \sup\{\|A^*\psi - \mu\|_{-p} \mid \|\psi\|_Y \leq R\}\,,\ R > 0\,.$$

Then

$$|\xi^\varepsilon - \xi^0| \leq R\varepsilon + \delta_R\|x^0\|_p\,,\ R > 0\,.$$

Since range(A^*) is dense in H_{-p} we obtain $\lim_{R\to\infty} \delta_R = 0$. Thus, two competing terms are involved in the error estimate for $|\xi^\varepsilon - \xi^0|$.

There is an important case when the computational effort of the reconstruction of functionals can be decreased when $\mu = \mu_t$ is the evaluation functional in a point t and A is an operator of convolution type; see Chapter 4. The property which is important in this context is that A commutes with translations.

2.3.6 A-posteriori regularizing strategy

Let $(R_t)_{t>0}$ be a regularization family. An a-posteriori-strategy starts from the requirement that the error $\|AR_ty^\varepsilon - y^\varepsilon\|_Y$ should be of the order of the noise level ε. In a rather general consideration it is reasonable to state the problem in the following way:

Determine $t(\varepsilon)$ such that

$$\|AR_{t(\varepsilon)}y^\varepsilon - y^\varepsilon\|_Y = \rho(\varepsilon, t(\varepsilon))\,. \qquad (2.74)$$

Such a parameter choice is called a *discrepancy principle.* Here we restrict our considerations to the regularization by Tikhonov's method and a special choice for ρ. Consider the set

$$N := \{Ax \mid x \in \mathcal{D}_B, Bx = 0\}$$

and let P denote the orthogonal projection of Y onto the closure of N.

Assumption A3:

Suppose that the following inequalities hold:

$$\|(I-P)y^0\|_Y > 0\,,\ \|(I-P)y^\varepsilon\|_Y > \varepsilon\,.$$

When $\mathcal{D}_B = X, B = \textit{identity}$, then we have $P = \Theta$ and $\|y^0\|_Y > 2\varepsilon$ implies the condition in assumption **A3**.

In the sequel we address an a-posteriori strategy with the function $\rho(\varepsilon, t(\varepsilon)) := \varepsilon$; see (2.74).

Morozov's a-posteriori strategy:

Let $x^{\varepsilon,t} := (A^*A+tB^*B)^{-1}A^*y^\varepsilon, t > 0$. Find $t = t(\varepsilon) > 0$ with

$$\|Ax^{\varepsilon,t} - y^\varepsilon\| = \varepsilon\,. \tag{2.75}$$

In Section 2.2.3 we obtained that the function

$$(0,\infty) \ni s \ \longmapsto\ \|Ax^{\varepsilon,1/s} - y^\varepsilon\|_Y^2 \in \mathbb{R}$$

is a convex function. Therefore the equation (2.75) is uniquely solvable and the solution can be found by Newton's method.

Theorem 2.3.8. *Suppose that the assumptions* **A0, A1, A3** *hold, let* $\|Bx^0\| \le E$, *and let* $t = t(\varepsilon)$ *be chosen according to the strategy* (2.75). *Then*

$$\|x^{\varepsilon,t(\varepsilon)} - x^0\| \le 2E\nu(\frac{\varepsilon}{E})\,. \tag{2.76}$$

Proof:
Let $t := t(\varepsilon)$ be chosen according to (2.75). Then

$$\begin{aligned} \varepsilon^2 + t\|Bx^{\varepsilon,t}\|_Z^2 &= \|Ax^{\varepsilon,t} - y^\varepsilon\|_Y + t\|Bx^{\varepsilon,t}\|_Z^2 \\ &\le \|Ax^0 - y^\varepsilon\|_Y + t\|Bx^0\|_Z^2 \ \le\ \varepsilon^2 + tE^2\,. \end{aligned}$$

This shows

$$\|Bx^{\varepsilon,t}\|_Z \le E\,.$$

Using the parallelogram identity we obtain

$$\|B(x^{\varepsilon,t}-x^0)\|_Z^2 = 2\|Bx^{\varepsilon,t}\|_Z^2 + 2\|Bx^0\|_Z^2 - \|B(x^{\varepsilon,t}+x^0)\|_Z^2 \le 4E^2\,,$$

i. e.

$$\|B(x^{\varepsilon,t}-x^0)\|_Z \le 2E\,. \tag{2.77}$$

From

$$Ax^0 = y^0, \|y^0 - y^\varepsilon\|_Y \le \varepsilon, \|Ax^{\varepsilon,t} - y^\varepsilon\|_Y = \varepsilon\,,$$

we obtain

$$\begin{aligned} \|x^{\varepsilon,t} - x^0\|_X &\le \sup\{\|u\|_X \mid u \in \mathcal{D}_B, \|Au\|_Y \le 2\varepsilon, \|Bu\|_Y \le 2E\} \\ &= 2E\nu(\frac{\varepsilon}{E})\,. \end{aligned}$$

■

The result of Theorem 2.3.8 is that the a-posteriori strategy (2.75) leads to the same error estimate as the a-priori strategy $t(\varepsilon) := \varepsilon^2 E^{-2}$; see Corollary 2.1.14.

2.3.7 Regularization by discretization

A projection method for the solution of an equation

$$Ax = y \tag{2.78}$$

is defined as follows:

Projection method

Given families $(X_h)_{h>0}$ and $(Y_h^*)_{h>0}$ of subspaces of X and Y respectively.
Find $x^h \in X_h$ such that

$$\langle Ax^h - y, w\rangle = 0 \text{ for all } w \in Y_h^*\,. \tag{2.79}$$

We assume

$$\dim X_h = \dim Y_h^* = n_h \in \mathbb{N}, h > 0\,.$$

Then by choosing bases of X_h and Y_h^* respectively, the solution (2.78) can be found by solving a linear system of equations. We introduce linear operators

$$P_h : X \longrightarrow X_h\,,\, R_h : y \longrightarrow X_h$$

by the definition

$$\langle AP_h u - Au, w\rangle = 0\,,\, \langle AR_h z - z, w\rangle = 0 \text{ for all } w \in Y_h^*\,.$$

Since

$$P_h = R_h A\,,\, P_h P_h = P_h$$

the term "projection method" becomes clear. Let

$$d(z, X_h) := \inf\{\|z - u\|_X \mid u \in X_h\}\,,\, z \in X\,.$$

d is a measure how well elements in X can be approximated by elements in X_h .

Coming back to our general setting, we can use a projection method to find an approximation of x^0 using the data y^ε in the following way:

$$x^{\varepsilon,h} := R_h y^\varepsilon\,,\, h > 0\,. \tag{2.80}$$

Theorem 2.3.9. *We have*

$$\|x^{\varepsilon,h} - x^0\|_X \le \|R_h\|\varepsilon + (1 + \|P_h\|)d(x^0, X_h)\,,\, h > 0\,. \tag{2.81}$$

Proof:
Since $\dim X_h < \infty$, there exists $z \in X_h$ with $\|z - x^0\|_X = d(x^0, X_h)$. We have

$$\begin{aligned}
\|x^{\varepsilon,h} - x^0\|_X &\le \|P_h x^0 - z\|_X + \|z - x^0\|_X + \|P_h x^0 - x^{\varepsilon,h}\|_X \\
&\le \|P_h(x^0 - z)\|_X + \|z - x^0\|_X + \|R_h A x^0 - R_h y^\varepsilon\|_X \\
&\le \|P_h\|\|x^0 - z\|_X + \|z - x^0\|_X + \|R_h\|\|Ay^0 - y^\varepsilon\|_X \\
&\le \|R_h\|\varepsilon + (1 + \|P_h\|)d(x^0, X_h)\,.
\end{aligned}$$

■

The error estimate in (2.81) indicates which quantities have to be discussed for a further analysis of a projection method.

The projection method is called *quasi-optimal* when the family $(P_h)_{h>0}$ is uniformly bounded and $\lim_{h\to 0} d(x, X_h) = 0$ for all $x \in X$. The projection method is called *robust* when the family $(\rho_h^{-1} P_h)_{h>0}$ is uniformly bounded where

$$\rho_h := \sup\{\|u\|_X \|Au\|_Y^{-1} \mid u \neq \theta, u \in X_h\}\,,\ h > 0$$

is the modulus of continuity of $A\ |_{X_h}^{-1}$. It is easy to prove that $\|R_h\| \geq \rho_h$. Therefore robustness means that the family $\|P_h\|)_{h>0}$ has the same asymptotic as $(\rho_h)_{h>0}$. Notice that $\lim_{h\to 0} \rho_h = \infty$ if $\lim_{h\to 0} d(x, X_h) = 0$ for all $x \in X$ since A^{-1} is unbounded.

Corollary 2.3.10. *Suppose that the projection method is quasi-optimal and robust. Then there exists a constant $c \geq 0$ such that*

$$\|x^{\varepsilon,h} - x^0\|_X \leq c(\rho_h \varepsilon + d(x^0, X_h))\,,\ h > 0\,. \tag{2.82}$$

Corollary 2.3.10 follows from Theorem 2.3.9. From the error estimate (2.82) one reads off the necessity to implement regularizing strategies in order to balance the competing terms. Again, the discussion of a-priori and a-posteriori strategies can be analyzed in the framework of Hilbert scales. Specific methods are:

- **Least squares method**: Choose X_h in X and set $Y_h^* := AX_h$.
- **Ritz method**: Choose X_h in X and set $Y_h^* := X_h$. Here it is assumed $X = Y, A = A^*$.
- **Generalized least squares method**: Choose X_h in X and set $Y_h^* := B^*BAX_h$. Here B is a linear mapping from $\mathcal{D}_B \longrightarrow Z$; see above.

We don't go into the analysis of these methods. In each case one can give conditions which imply quasi-optimality and robustness.

2.4 Finite dimensional problems

Here we take a short look at the problems which result when ill-posed problems are "projected" down to a finite dimensional situation.

2.4.1 Ill-conditioned problems

The discretization of linear ill-posed problems gives rise to linear systems of equations

$$Ax = y \quad (A \in \mathbb{R}^{m,n}, x \in \mathbb{R}^n, y \in \mathbb{R}^m) . \tag{2.83}$$

We discuss this system under the assumption $n \leq m$ which means that the system is *overdetermined.* Again, the singular value decomposition is the main tool to understand the problems in solving the equations in a stable way.

Let A^t denote the transposed matrix of A. The euclidian norm in $\mathbb{R}^n$ is denoted by $\| \cdot \|_2$.

Theorem 2.4.1 (Singular value decomposition). *Let $A \in \mathbb{R}^{m,n}$ with $\leq m$. Then there exist $U \in \mathbb{R}^{m,m}, V \in \mathbb{R}^{n,n}$ and a diagonal matrix Σ with entries $\sigma_1, \ldots, \sigma_n \geq 0$ in the diagonal such that*

$$A = U\Sigma V^t \, , \, U^tU = UU^t = E, V^tV = VV^t = E$$

holds. $\sigma_1, \ldots, \sigma_n$ are called the **singular values** *of A.*

The decomposition above is analogous to the case in the infinite dimensional situation. When $A = U\Sigma V^t$ is a singular value decomposition with singular values $\sigma_1, \ldots, \sigma_n$, and matrices $U = (u^1|\ldots|u^m), V = (v^1|\ldots|v^n)$ then

$$Av^i \;=\; \sigma_i u^i \, , \, A^t u^i \;=\; \sigma_i v^i \, , \, i = 1, \ldots, n \, . \tag{2.84}$$

The system $(u^1, \ldots, u^m, v^1, \ldots, v^n, \sigma_1, \ldots, \sigma_n)$ is called a *singular system.* Notice that the squares of singular values are the eigenvalues of the matrix A^tA. Notice too that the singular values of the matrix A are uniquely determined but not the matrices U, V due to the fact that in general a basis is not uniquely determined. Without loss of generality we may assume

$$\sigma_1 \geq \ldots \geq \sigma_r > 0 \, , \, \sigma_{r+1} = \cdots = \sigma_n = 0 \text{ and } \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \, .$$

Now we have

$$\begin{aligned}
\mathrm{rank}(A) &= r; && (2.85)\\
\mathrm{null}(A) &= \mathrm{span}(v^{r+1}, \dots, v^n); && (2.86)\\
\mathrm{range}(A) &= \mathrm{span}(u^1, \dots, u^r); && (2.87)\\
\|A\|_2 &= \sigma_1; && (2.88)\\
A &= \sum_{i=1}^{r} \sigma_i u^i (v^i)^t\,. && (2.89)
\end{aligned}$$

The identity in (2.89) shows that A may be decomposed into a sum of r matrices with rank equal to 1.

When the system (2.83) comes from an infinite dimensional ill-posed problem we may expect that the matrix A has many "tiny" singular values, some of which may be vanishing. Following [37], we refer to such linear systems as *discrete ill-posed problems*. It is clear that the separation between ill-conditioned and well-conditioned problems is more vague than the concept of well-posed problems.

Consider the conditions

$$\begin{aligned}
AXA &= A\,,\ XAX \;=\; A\,. && (2.90)\\
(AX)^t &= AX\,,\ (XA)^t = XA\,. && (2.91)
\end{aligned}$$

It is easy to verify that a matrix X which satisfies the conditions (2.90), (2.91) is uniquely defined.

Definition 2.4.2. *Let $A \in \mathbb{R}^{m,n}$. A matrix $X \in \mathbb{R}^{n,m}$ is called a* (Moore–Penrose) pseudoinverse *of A, if X is a solution of the equations* (2.90),(2.91). *We denote the (Moore–Penrose) pseudoinverse of A by $A^\dagger := X$.* □

Notation: Let $\sigma \in \mathbb{R}$. Define $\sigma^- := \sigma^{-1}$ if $\sigma \neq 0$ and $\sigma^- := 0$ if $\sigma = 0$.

Let $A = U\Sigma V^t$ be a singular value decomposition, let r be the rank of A. If

$$\Sigma = \begin{pmatrix} \Sigma_r & \Theta \\ \Theta & \Theta \end{pmatrix} \text{ with } \Sigma_r = \mathrm{diag}(\sigma_1, \dots, \sigma_r) \in \Sigma_r \in \mathbb{R}^{r,r}$$

then we set

$$\Sigma^\dagger := \begin{pmatrix} \Sigma_r^- & \Theta \\ \Theta & \Theta \end{pmatrix} \text{ with } \Sigma_r^- := \operatorname{diag}(\sigma_1^-, \dots, \sigma_k^-)\,.$$

Corollary 2.4.3. *Let $A \in \mathbb{R}^{m,n}$ and let $A = U\Sigma V^t$ be a singular value decomposition of A. Then $X := V\Sigma^\dagger U^t$ is the pseudo inverse of A.*

Proof:
It is easy to verify that $X := V\Sigma^\dagger U^t$ solves equations (2.90),(2.91). ■

Example 2.4.4. *Let*

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\,,\ \Delta(\varepsilon) := \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \\ 0 & 0 \end{pmatrix}\,.$$

Then

$$A^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,,\ (A + \Delta(\varepsilon))^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1/\varepsilon & 0 \end{pmatrix}\,.$$

Therefore $\lim_{\varepsilon\to 0}(A + \Delta(\varepsilon)) = \Theta$, *but* $\lim_{\varepsilon\to 0}(A + \Delta(\varepsilon))^\dagger$ *does not exist. This shows that the mapping* $A \longmapsto A^\dagger$ *is not a continuous one.* □

Remark 2.4.5. *The definition of the pseudoinverse of a matrix A is compatible with the definition of the pseudoinverse of a compact injective operator in Remark 2.1.10.*

Remark 2.4.6. *Suppose that we have a matrix $B \in \mathbb{R}^{n,l}$. Then the method of Tikhonov in the finite dimensional context is to determine the solution $x^{\varepsilon,t}$ of the associated normal equations:*

$$(A^tA + tB^tB)x^{\varepsilon,t} = y^{\varepsilon,t}\,. \tag{2.92}$$

Again t is a regularization parameter which has to be determined in a proper way. We set

$$u(t) := \|Ax^{\varepsilon,t} - y^{\varepsilon,t}\|_2^2\,,\ v(t) := \|Bx^{\varepsilon,t}\|_2^2\,,\ t > 0\,.$$

This curve displays the trade-off between minimizing the residuals $\|Ax - y^\varepsilon\|_2$ *and* $\|Bx\|_2$ *. The L–curve selection criterion consists in locating the* t*–value which maximizes the curvature in the typical L-shaped plot of the curve*

$$\Lambda : (0, \infty) \ni t \ \longmapsto \ (\ln(u(t)), \ln(v(t)) \in \mathbb{R}\,.$$

This "L-curve corner" *marks a natural transition point linking these two regions. The location of the parameter* t_L*, corresponding to this corner point, may be derived by using the singular value decomposition. But for large scale problems where the cost of the singular value decomposition of* A *is very costly, the curvature* κ *should be expressed directly by* A *.* □

2.4.2 Least squares

Let x^0 be the (unique) solution of (2.83) with right hand side y^0 :

$$Ax^0 = y^0\,. \tag{2.93}$$

Again, we assume that the right hand side y^0 is contaminated by noise:

$$y^\varepsilon = y^0 + w^\varepsilon, w^\varepsilon \in Y, \|w^\varepsilon\| \le \varepsilon \tag{2.94}$$

where $\varepsilon \ge 0$ is the so-called *noise level.* We want to solve the equation

$$Ax = y \tag{2.95}$$

for y^ε . Since this systems is overdetermined there is no chance to solve this system. Therefore we search for a solution of the linear least squares problem

$$\text{Minimize } \|Ax - y^\varepsilon\|_2^2\,. \tag{2.96}$$

The existence of a solution of (2.96) is easy to prove. Let $r := \text{rank}(A) \le n \le m$. Then a solution of (2.96) is uniquely determined if and only if $r = n$. If $r < n$, then there exists in $\mathbb{R}^n$ a $(n - r)$–dimensional subspace of solutions. In particular, let x^* denote the shortest solution (with respect to the euclidian norm) then the general solution can be written as

$$x = x^* + z$$

with an arbitrary $z \in \text{null}(A)$. Now we can verify that this solution x^* can be represented by the (Moore-Penrose)–pseudoinverse, namely $x^* = A^\dagger y$. For the right hand side $y := y^\varepsilon$ we set

$$x^\varepsilon := A^\dagger y^\varepsilon \,.$$

When $A = U\Sigma V^t$ is a singular value decomposition of A with rang(A) $= r, U = (u^1|\dots|u^m)$ then we obtain

$$\|AA^\dagger x^\varepsilon - y^\varepsilon\|^2 = \sum_{i=r+1}^{m} \langle u^i, y^\varepsilon\rangle^2 \,.$$

In order to find $x^\varepsilon = A^\dagger y^\varepsilon$ the singular value decomposition is not the method of choice. Essentially, there are four basic approaches to compute the solution x^ε :

- Normal equations solution
- QR-decomposition
- Augmented system solution
- Krylow subspace methods

The first approach is the one originally derived by C. F. Gauss. It consists in solving the normal equations

$$A^t Ax = A^t y^\varepsilon$$

by a decomposition method like the Cholesky method.

The QR–decomposition method uses the fact that the euclidian norm is invariant under an orthogonal transformation Q . There are several methods to find an orthogonal matrix Q with

$$QA = \begin{pmatrix} R \\ \Theta \end{pmatrix} \,,\; Qy^\varepsilon = \begin{pmatrix} u \\ v \end{pmatrix}$$

where R is an upper tridiagonal matrix. Once Q, R are found one has to solve the equation $Rx = u$.

In the third approach one introduces as an additional unknown $d := y^\varepsilon - Ax$ and solves the normal equations in the following augmented form:

$$\begin{pmatrix} I & A \\ A^t & \Theta \end{pmatrix} \begin{pmatrix} d \\ x \end{pmatrix} = \begin{pmatrix} y^\varepsilon \\ \theta \end{pmatrix} .$$

Krylow subspace methods are iterative methods for solving (large) systems of linear equations. The conjugate gradient method is for symmetric problems and GMRES is a well known Krylov method for the nonsymmetric case. The Krylow subspaces with respect to A and y^ε are defined by

$$\mathcal{K}_j(A, y^\varepsilon) := \operatorname{span}(\{y^\varepsilon, Ay^\varepsilon, \dots, A^{j-1}y^\varepsilon\}\, , \, j = 1, 2, \dots . \tag{2.97}$$

The GMRES method determines iterates $x^{\varepsilon,j} \in \mathcal{K}_j(A, y^\varepsilon)$, $j \in \mathbb{N}$, that satisfy

$$\|Ax^{\varepsilon,j} - y^\varepsilon\| = \min_{x \in \mathcal{K}_j(A, y^\varepsilon)} \|Ax - y^\varepsilon\| . \tag{2.98}$$

```
Input             Matrix $A \in \mathbb{R}^{n,n}$, righthand side $y$,
                  initial guess $v^0$.

Initialization:   $r^0 := Av^0 - b, \beta := \|r^0\|, v^1 := \beta^{-1}v^0$,
                  $V_1 := (v^1)$.

          For $i = 1, \dots, i_{\max}$ do:
Normalization:    $v^{i+1} := \hat{v}^{i+1}\|\hat{v}^{i+1}\|^{-1}$.
Update:           $V_{i+1} := (V_i|v^{i+1}), H_i := \begin{pmatrix} H_{i-1} & h_i \\ 0 & \|\hat{v}^{i+1}\| \end{pmatrix}$.
Least squares:    $\|\beta e^1 - H_i z^i\| = \min_z \|\beta e^1 - H_i z\|$.
          end do

Output             $v^i = V_i z^i + v^0$, $i = 0, \dots, i_{\max}$.
```

It follows from (2.98) and the relation $\mathcal{K}_{j-1}(A, y^\varepsilon) \subset \mathcal{K}_j(A, y^\varepsilon)$ that the inequalities

$$\|r_0^\varepsilon\| \geq \|r_1^\varepsilon\| \geq \|r_2^\varepsilon\| \geq \cdots \tag{2.99}$$

hold for the residual vectors

$$r_j^\varepsilon := Ax_j^\varepsilon - y^\varepsilon, j = 1, 2, \ldots .$$

However the GMRES iterates are not guaranteed to satisfy the inequalities

$$\|x_0^\varepsilon\| \leq \|x_1^\varepsilon\| \leq \|x_2^\varepsilon\| \leq \ldots , \tag{2.100}$$

as it is the case when we apply the cg-method to solve the problem. This fact can be used to develop a criterion for an early termination of the iteration of the GMRES method. This termination criterion is based on the condition number of the matrices $A_i := V_{i+1}H_iV_i^t$; see [15].

2.5 Bibliographical comments

The results in Sections 2.1, 2.2, 2.3 can be found in nearly all monographs on ill-posed problems; see [5, 23, 53]. In Subsection 2.3.4 we are inspired by [68]. More on a-posteriori parameter choice principles can be found in [73]. The reconstruction of functionals is treated a little bit different from the literature; see [66]. The problems of numerical algebra in solving ill-conditioned problems are discussed very detailed in [37] under the aspect of using matlab.

2.6 Exercises

2.1. Prove the inequality (2.8).

2.2. Let V be Hilbert space which is densely imbedded in X and suppose that the linear bounded operator $B : V \longrightarrow X$ has a bounded inverse $B^{-1} : X \longrightarrow V$. Show that there exists a constant $c > 0$ such that

$$c^{-1}\|v\|_V \leq \|Bv\|_X \leq c\|v\|_V \text{ for all } v \in V .$$

2.3. Let U, V, W be Hilbert spaces and let $Q : U \longrightarrow W, R : V \longrightarrow W$ be linear bounded operators. Prove the equivalence of the following conditions:

(a) There exists a linear bounded operator $S : V \longrightarrow U$ with $QS = R$.

(b) range$(R) \subset$ range(Q).

(c) There exists $\lambda > 0$ such that $RR^* \leq \lambda QQ^*$.

2.4. Consider the filter $q : (0,\infty)\times(0,\infty) \ni (t,\sigma) \longmapsto 1-(1-a\sigma^2)^{\frac{1}{t}}$ where a is a positive number. Prove:

$$|q(t,\sigma)| \leq \sqrt{\frac{a}{t}}\,,\ |q(t,\sigma)-1| \leq \frac{t}{a\sigma^2}\,,\ (t,\sigma) \in (0,\infty)\times(0,\infty)\,.$$

2.5. Consider the operator $A : L_2(0,1) \longrightarrow L_2(0,1)$, defined by

$$Ax(t) := \int_0^t x(s)ds\,,\ t \in [0,1]\,.$$

Prove that A is a compact operator with norm 1.
Hint: Use the theorem of Arzela-Ascoli.

2.6. Give an estimate for

$$\sup\{\|x\| \mid \|Ax\| \leq \tau, \|Bx\| \leq 1, x \in H_0^2[0,1]\}$$

where A is the integral operator in the problem 2.14, $H_0^2[0,1]$ is the space

$$\{x \in L_2[0,2] \| x, x' \in AC[0,1], x'' \in L_2[0,1], x(1) = 0, x'(0) = 0\}$$

and Bx is defined as x''.

2.7. Compute the adjoint operator A^* of the operator A in the exercise above.

2.8. Let $x \in L_2(-\pi,\pi)$ be given and consider x as a function on the circle with radius $r = 1$. According to Poisson's formula, if a harmonic function equals x on the unit circle, then in the unit disk it is given by

$$u(r,\phi) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1+r^2-2r\cos(\phi-\alpha)} x(\alpha)d\alpha\,,$$

for $r \in (0,1)$, $\phi \in (-\pi,\pi)$. The problem of harmonic continuation consists in finding x from the restriction $y := u(r,\cdot)$ where $r \in (0,1)$ is given. Show that A, defined by

$$Ax(\phi) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1+r^2-2r\cos(\phi-\alpha)} x(\alpha)d\alpha\,,\ -\pi < \phi \le \pi\,,$$

is a linear bounded operator from $L_2(-\pi,\pi)$ into $L_2(-\pi,\pi)$.
Hint: $\dfrac{1-r^2}{1+r^2-2r\cos(\phi-\alpha)} \le \dfrac{1+r}{1-r}$.

2.9. Consider the operator A in the exercise above and let $x_n := \cos(n\cdot), n \in \mathbb{N}$. Compute $\|x_n\|_{L_2(-\pi,\pi)}, \|Ax_n\|_{L_2(-\pi,\pi)} n \in \mathbb{N}$, and $\lim_n Ax_n$.

2.10. Consider the integral equation

$$\int_{-\infty}^{\infty} \frac{1}{1+(t-s)^2} x(s)ds = y(t)\,,\ t \in \mathbb{R}\,.$$

Set $y(t) := \tau\sin(nt), x(s) := \tau e^n \sin(ns), t, s \in \mathbb{R}$. Show that x is a solution.

2.11. Consider with $X := Y := L_2[0,1]$ the operator

$$J : X \longrightarrow X\,,\ J(x)(t) := \int_0^t x(s)ds, t \in [0,1],$$

and the multiplication operator

$$M : X \ni x \longmapsto mx \in Y$$

where m is a measurable function with $0 < |m(t)| \le \mu, t \in [0,1]$. Show that $M \circ J : X \longrightarrow Y$ is compact and that the asymptotic of the singular values of $M \circ J$ is $O(n^{-1})$.

Chapter 3

Iterative methods

This chapter is devoted to the analysis of iterative regularization methods. In particular we shall focus our attention on the Landweber type methods. What concerns linear equations, iterative methods for approximating the generalized inverse (i.e. the least square solution of minimum norm) are based on algorithms for solving fixed point equations related to the normal equation (see [23, 31, 32] for corresponding definitions). The regularization character of these methods (in the case of noisy data) is related to an early termination of the iteration, and a corresponding stopping rule is determined by an *a posteriori* evaluation of the iteration residual.

We address the Landweber iteration for both linear and nonlinear equations. Further, we investigate a modified Landweber iteration and also a continuous version of the Landweber iteration (the so called *asymptotical regularization*). For all these methods we prove convergence results for exact as well as for noisy data. Moreover, under additional regularity assumptions on the solution of the inverse problem, it is possible to obtain convergence rates, i.e. estimates to the number of iterative steps required in order to reach the stopping criterion as well as estimates to the iteration error.

It is worth noticing that all these iterative methods are adjoint type methods, i.e. the iteration is governed by an operator which is adjoint to the operator which models the inverse problem.

3.1 Introduction

We are concerned with operator equations of the type

$$F(x) \; = \; y \, , \tag{3.1}$$

where $F : D(F) \subset X \to Y$ and X, Y are Hilbert spaces. If F is linear and bounded, we denote by $x^\dagger$ the (generalized) solution of (3.1), i.e. $x^\dagger = F^\dagger y$ is the least square solution of (3.1) which has minimum norm; see Remark 2.1.10. In practical applications the available data is infected by measurement errors (not to mention the modeling errors). Therefore, only approximate data y^ε with $\|y - y^\varepsilon\| \leq \varepsilon$ is available. We shall refer to $\varepsilon > 0$ as *noise level*.

It is a well known fact that $x^\dagger$ is a solution of the normal equation

$$F^* F x \; = \; F^* y \, , \tag{3.2}$$

where F^* is the Hilbert adjoint of F (actually the unique solution of (3.2) lying in $\mathrm{null}(F)^\perp$; see [31] for details).

In the linear case, the iterative regularization methods for approximating $x^\dagger = F^\dagger y$ are based on fixed point equations related to (3.2), like

$$x \; = \; x - F^*(Fx - y) \; = \; (I - F^*F)x + F^* y \, .$$

Notice that if F is continuous, then F^*F is a bounded self-adjoint nonnegative operator. Further, if $\|F\|^2 < 2$, then the fixed point operator $(I - F^*F)$ is nonexpansive (i.e., $\|I - F^*F\| \leq 1$). Although this operator is not necessarily a contraction, convergence results of the fixed point iteration can be obtained (cf. [8]). If problem (3.1) is ill-posed (e.g., when F is compact), then $(I - F^*F)$ is not a contraction, since zero belongs to the continuous spectrum of F.

The fixed point iteration above suggests the explicit iteration:

$$x_{k+1} \; = \; x_k - F^*(Fx_k - y) \, , \; k \geq 0 \tag{3.3}$$

($x_0 \in D(F)$), which corresponds to the Landweber iteration.The strong convergence of (3.3) in the case of F compact and $y \in D(F^\dagger)$ was proved by Landweber in [57]. Independently, Fridman analyzed in [25] the same iteration for F compact self-adjoint and positive.[1]There are other names associated with the analysis of equivalent

[1] Iteration (3.3) is also called Landweber-Fridman iteration.

iterations (specially in the engineering literature). The reader is referred to [23] for further historical references.

If the operator F in (3.1) is nonlinear, there are several ways to generalize the Landweber iteration (3.3). If the Fréchet derivative $F'(\cdot)$ of F is locally uniformly bounded, a possible alternative is to consider the iteration:

$$x_{k+1} \;=\; x_k - F'(x_k)^*(F(x_k) - y)\,,\ k \geq 0\,, \tag{3.4}$$

where $x_0 \in D(F)$ is some initial guess. Convergence rates results for iteration (3.4) are proven under the sourcewise representation

$$x_* - x_0 \;=\; (F'(x_*)^* F'(x_*))^\nu w\,,\ \nu > 0\ ,\ w \in X\,, \tag{3.5}$$

where $x_* \in D(F)$ is a solution of (3.1). Further, one needs (locally) a representation condition on F':

$$F'(x) \;=\; R_x F'(x_*)\,,\ x \in B_\rho(x_0)\,, \tag{3.6}$$

where $\{R_x\}_{x \in B_\rho(x_0)}$ is a family of bounded linear operators $R_x : Y \to Y$ with

$$\|R_x - I\| \;\leq C\|x - x_*\|\,,\ x \in B_\rho(x_0)\,. \tag{3.7}$$

See Section 3.3 for a detailed exposition.

An alternative method for solving (3.1) in the nonlinear case is the modified Landweber iteration

$$x_{k+1} \;=\; x_k - F'(x_k)^*(F(x_k) - y) - \alpha_k(x_k - \zeta)\,, \tag{3.8}$$

where $0 \leq \alpha_k \leq 1$ and $\zeta \in D(F)$. The advantage of the modified Landweber iteration resides in the fact that one can prove a convergence rates result under the usual source condition (3.5), without requiring any additional representation condition on F' (as the one in (3.6)). From the point of convergence analysis, this method is optimal, since convergence, stability and convergence rates can be guaranteed under minimal assumptions. This iterative method is considered in Section 3.4.

In Section 3.5 we consider a continuous version of the Landweber iteration. It corresponds to the so called asymptotic regularization.

In this method a regularized approximation $x(T)$ of a solution x_* of (3.1) is obtained by solving the initial value problem:

$$x'(t) \;=\; F'(x(t))^*(y - F(x(t)))\,,\; t \in (0,T] \quad x(0) \;=\; x_0\,. \qquad (3.9)$$

Convergence for this method can be proved and, assuming a source condition of type (3.5), it is also possible to obtain stability estimates.

3.2 Landweber iteration for linear equations

We consider in this section operator equations of the type (3.1) for linear compact operators F. What concerns the Landweber iteration (3.3) we shall adopt the notation

$$\begin{aligned} x_{k+1} &= x_k - F^*(Fx_k - y) && (3.10)\\ x^\varepsilon_{k+1} &= x^\varepsilon_k - F^*(Fx^\varepsilon_k - y^\varepsilon)\,, && (3.11)\end{aligned}$$

where y^ε and ε have the same meaning as in Section 3.1. Notice that $x_0 = x^\varepsilon_0 \in X$. For simplicity of the presentation we assume $\|F\| \le 1$. If this were not the case, we could always introduce a relaxation parameter $\lambda \in (0, \|F\|^{-2}]$ in the normal equation (3.2) and rewrite the iteration as $x_{k+1} = x_k - \lambda F^*(Fx_k - y)$. In order to simplify the notation we use as initial guess $x_0 = 0$. The results presented in this section are in no way affected by this particular choice.

Our first result concerns convergence for exact data.

Lemma 3.2.1. *Let $y \in D(F^\dagger) = range(F) + null(F^*)$ and $\{x_k\}$ be the sequence defined in (3.10). Then $x_k \to F^\dagger y$ as $k \to \infty$.*

Proof:
Since $y \in D(F^\dagger)$, we have $F^*y = F^*FF^\dagger y$. Thus, from the definition of x_k follows $F^\dagger y - x_k = (I - F^*F)^k F^\dagger y$. Since $\|F\| \le 1$, it follows from the spectral theory for linear bounded self-adjoint operators (cf., e.g., [56]) that $(I - F^*F)^k F^\dagger y \to P(F^\dagger y)$, $k \to \infty$, where P is the orthogonal projector onto null(F). Now, from $F^\dagger y \in$ null$(F)^\perp$ follows $x_k \to F^\dagger y$. ■

The assumption $y \in D(F^\dagger)$ is not only sufficient but also necessary for the convergence of the sequence x_k (even in weak sense). Actually, one can prove that if $y \notin D(F^\dagger)$, then $\|x_k\|$ is unbounded (see Exercise 3.1). This is a classical result in the regularization theory and we refer the reader to [23, Theorem 4.1] or [31, 32].

The next result gives an estimate for the error propagation in the Landweber iteration.

Lemma 3.2.2. *Let $y \in D(F^\dagger)$ and $y^\varepsilon \in Y$ be such that $\|y - y^\varepsilon\| \leq \varepsilon$. Further, let $\{x_k\}$, $\{x_k^\varepsilon\}$ be the sequences defined in (3.10), (3.11). Then we have the estimate*

$$\|x_k - x_k^\varepsilon\| \leq \sqrt{k}\varepsilon\,.$$

Proof:
A simple algebraic calculation shows that $x_k - x_k^\varepsilon = R_k(y - y^\varepsilon)$, where

$$R_k \;:=\; \sum_{j=0}^{k-1} (I - F^*F)^j F^*\,.$$

Now, the lemma follows from $\|I - F^*F\| \leq 1$ together with the estimate $\|R_k\|^2 = \|R_k R_k^*\| \leq \|\sum_{j=0}^{k-1}(I - F^*F)^j\|$. ■

From the above lemmas we conclude that, in the case of noisy data $y^\varepsilon \notin D(F^\dagger)$ with $\|y - y^\varepsilon\| \leq \varepsilon$, we have

$$\|F^\dagger y - x_k^\varepsilon\| \leq \|F^\dagger y - x_k\| + \|x_k - x_k^\varepsilon\|.$$

Notice that the total error can be divided in two components: the first one is the *approximation error* and converges to zero as $k \to \infty$. The second one is the *propagated data error* and has the order of $\sqrt{k}\varepsilon$, which becomes unbounded as $k \to \infty$. As one iterates, the propagated data error increases and, when $\sqrt{k}\varepsilon$ becomes larger than the approximation error, the approximations x_k^ε become worst (i.e. $\|F^\dagger y - x_k^\varepsilon\|$ starts increasing). Notice that, since $F^\dagger$ is unbounded, the pre-image of the ball $B_\varepsilon(y^\varepsilon) \subset Y$ with respect to F is unbounded in $\mathrm{null}(F)^\perp \subset X$.

Another important property of the Landweber iteration concerns the evolution of the residual $y^\varepsilon - Fx_k^\varepsilon$. Since

$$y^\varepsilon - Fx_{k+1}^\varepsilon \;=\; (I - FF^*)(y^\varepsilon - Fx_k^\varepsilon)$$

and $\|I - F^*F\| \leq 1$, it follows that the norm of the residual decreases monotonically during the iteration.

The necessity of a stopping criterion for the Landweber iteration in case of noisy data is now clear. A common parameter choice rule is the one described by the discrepancy principle, according to which the iteration should be terminated at the step $k = k(\varepsilon, y^\varepsilon)$ when

$$\|y^\varepsilon - Fx^\varepsilon_{k(\varepsilon,y^\varepsilon)}\| \leq \tau\varepsilon \,, \tag{3.12}$$

for the first time (here $\tau > 1$ is fixed). The next result guarantees a monotony property of the monotony of the iteration error as long as the discrepancy principle is not reached.

Lemma 3.2.3. *Let $y = Fx_* \in range(F)$, and $y^\varepsilon \in Y$ be such that $\|y - y^\varepsilon\| \leq \varepsilon$. Further, let $\{x^\varepsilon_k\}$ be the sequence defined in (3.11). If $\|y^\varepsilon - Fx^\varepsilon_k\| > 2\varepsilon$ then*

$$\|x_* - x^\varepsilon_{k+1}\| \leq \|x_* - x^\varepsilon_k\| \,.$$

Proof:
Notice that

$$\begin{aligned}\|x_* - x^\varepsilon_{k+1}\|^2 = \|x_* - x^\varepsilon_k\|^2 - 2\langle y - y^\varepsilon, y^\varepsilon - Fx^\varepsilon_k\rangle - \|y^\varepsilon - Fx^\varepsilon_k\|^2 \\ + \langle y^\varepsilon - Fx^\varepsilon_k, (FF^* - I)(y^\varepsilon - Fx^\varepsilon_k)\rangle\end{aligned}$$

(see Exercise 3.2). Since $FF^* - I$ is nonpositive, it follows

$$\|x_* - x^\varepsilon_k\|^2 - \|x_* - x^\varepsilon_{k+1}\|^2 \geq \|y^\varepsilon - Fx^\varepsilon_k\|(\|y^\varepsilon - Fx^\varepsilon_k\| - 2\varepsilon)$$

completing the proof. ■

The last lemma tell us that the Landweber iteration (3.11) should not be stopped before the discrepancy principle (3.12) with $\tau = 2$. It is worth noticing that no $\tau \geq 1$ should be employed in (3.12), otherwise the residual may never reach the tolerance given by the discrepancy principle.

The next lemma gives us an estimate for the stopping index $k(\varepsilon, y_\varepsilon)$ in (3.12).

Lemma 3.2.4. *If we choose $\tau > 1$ in (3.12) then the stopping index given by the discrepancy principle can be estimated by*

$$k(\varepsilon, y_\varepsilon) = O(1/\varepsilon^2)\,.$$

Proof:
Let $y = Fx_*$ and $\{x_k\}$, $\{x_k^\varepsilon\}$ be the sequences defined in (3.10), (3.11). Arguing as in the previous lemma we obtain

$$\|x_* - x_k\|^2 - \|x_* - x_{k+1}\|^2 \geq \|y - Fx_k\|^2\,.$$

Therefore

$$\|x_* - x_1\|^2 - \|x_* - x_{k+1}\|^2 \geq k\|y - Fx_k\|^2\,.$$

Now, from $y - Fx_k = (I - FF^*)^k(y - Fx_0)$ follows

$$\|(I - FF^*)^k(y - Fx_0)\| \leq k^{-1/2}\|x_* - x_1\|$$

Finnaly we obtain the estimate

$$\|y^\varepsilon - Fx_k^\varepsilon\| \ = \ \|(I - FF^*)^k(y^\varepsilon - Fx_0)\| \ \leq \ \varepsilon + k^{-1/2}\|x_* - x_1\|$$

and conclude that the right hand side is smaller than $\tau\varepsilon$ when k becomes larger than $(\tau - 1)^{-2}\|x_* - x_1\|^2\varepsilon^{-2}$. Hence, $k(\varepsilon, y^\varepsilon) = c\varepsilon^{-2}$ and the lemma is proved. ■

The next result gives an estimate for $k(\varepsilon, y^\varepsilon)$ and $\|F^\dagger y - x^\varepsilon_{k(\varepsilon,y^\varepsilon)}\|$, when we consider the Landweber iteration with the discrepancy principle (3.12).

Lemma 3.2.5. *If $y \in$ range(F) and $F^\dagger y = (F^*F)^\nu w$, holds for some $w \in X$ and $\nu > 0$, then the stopping rule defined by (3.12) with $\tau > 1$ satisfies*

$$k(\varepsilon, y^\varepsilon) \ = \ O\big(\varepsilon^{-2/(2\nu+1)}\big)\,. \tag{3.13}$$

Moreover,

$$\|F^\dagger y - x^\varepsilon_{k(\varepsilon,y^\varepsilon)}\| \ = \ O\big(\varepsilon^{2\nu/(2\nu+1)}\big)\,. \tag{3.14}$$

Sketch of the proof:
For a complete proof we refer the reader to [23, Theorem 6.5] or [65]. A proof for the special case $\nu = 1$ can be found in [32, 53].

The natural way of proving this result is to consider the Landweber iteration within the framework of general regularization theory (see, e.g., [23, 32, 53, 65]). Then the estimates (3.13), (3.14) follow from a main result of this theory in a straightforward way. In the sequel we briefly introduce some relevant concepts of general regularization theory and describe the steps of the proof of Lemma 3.2.5.

It is a well known fact (see Section 2.2 and, e.g., [53, Theorem 2.6]) that the filter functions $q(\alpha, \mu) := 1 - (1 - \mu^2)^{1/\alpha}$ define a family of linear operators $R_\alpha : Y \to X$, $\alpha > 0$,

$$R_\alpha : y \ \longmapsto \ \sum_{k=1}^{\infty} \frac{q(\alpha, \nu)}{\nu} \langle y, y_k \rangle x_k \, , \ y \in Y \, , \tag{3.15}$$

(here $(\mu_k; x_k, y_k)$ denotes the singular system for F), such that $\{R_\alpha\}$ is a *regularization strategy*, i.e. $R_\alpha F x \to x$, as $\alpha \to 0$, for all $x \in X$. Note that each x_k defined in (3.10) satisfies $x_k = R_\alpha y$, with $\alpha = 1/k$. Therefore, the Landweber iteration can be interpreted as a regularization strategy $\{R_k\}$ with discrete regularization parameter (cf., e.g., [53, Theorem 2.15]).[2]

The notion of regularization strategy is based on exact data, i.e. we apply the operators R_α to $y \in D(F^\dagger)$ and obtain $R_\alpha y \to F^\dagger y$ as $\alpha \to 0$. In the presence of noise, i.e. when only a garbled version y^ε of the data y is available, one has also to define a *parameter choice rule* $\alpha = \alpha(\varepsilon, y^\varepsilon)$ such that

$$\lim_{\varepsilon \to 0} \sup\{\|R_{\alpha(\varepsilon, y^\varepsilon)} y^\varepsilon - F^\dagger y\| \, ; \ y^\varepsilon \in Y, \ \|y - y^\varepsilon\| \le \varepsilon\} \ = \ 0 \, ,$$

and

$$\lim_{\varepsilon \to 0} \sup\{\alpha(\varepsilon, y^\varepsilon); \mid y^\varepsilon \in Y, \ \|y - y^\varepsilon\| \le \varepsilon\} = 0 \text{ for all } y \in D(F^\dagger).$$

If the above conditions are fulfilled, the pair (R_α, α) is called a (convergent) *regularization method* for solving $Fx = y$. The regularization strategy $\{R_k\}$ defined by the Landweber iteration together with

[2] The operators R_k were already defined in the proof of Lemma 3.2.2.

the parameter choice rule $k(\varepsilon, y^\varepsilon)$ given by the discrepancy principle (3.12) generate a regularization method $(R_k, k(\varepsilon, y^\varepsilon))$ (see, e.g., [53, Theorem 2.19]).

Convergence rates can only be given on subsets of $D(F^\dagger)$ (or, equivalently, on subsets of X), e.g., $\mathcal{X}_\nu := \mathcal{R}((\mathcal{F}^*\mathcal{F})^\nu) \subset \mathcal{X}$, for $\nu > 0$. For a general regularization method (R_α, α), where α is defined via the discrepancy principle, one can prove that, under the assumption $F^\dagger y \in \mathcal{X}_\nu$, for some $\nu \in (0, \nu_0/2]$, the rate of convergence $\|F^\dagger y - R^\varepsilon_{k(\varepsilon,y^\varepsilon)} y^\varepsilon\| = O(\varepsilon^{2\nu/(2\nu+1)})$ hold (cf. [23, Theorem 4.17]). The constant ν_0 is called *qualification* of the regularization $\{R_\alpha\}$ (cf. [23, Chapter 4]). Thus, in order to prove Lemma 3.2.5, it is enough to verify that the regularization method $(R_k, k(\varepsilon, y^\varepsilon))$ has qualification $\mu_0 = \infty$. This is actually the main task in [23, Theorem 6.5]. ■

The numerical cost of implementing the Landweber iteration is very high, since it usually requires a very large number of iterative steps until the stopping criterion (3.12) is achieved. When combined with accelerating semiiterative methods, the so called *accelerated Landweber methods* become an attractive alternative to Tikhonov regularization. We refer the reader to [35, 23] for details.

Using similar arguments as in the proof of the lemma above, it can be shown that the exponent in the estimate (3.13) cannot be improved in general. For details we refer to [23, Theorem 6.9].

3.3 The nonlinear Landweber iteration

In this section we consider the nonlinear Landweber method in (3.4). As in (3.10), (3.11) we shall denote by $\{x_k\}$, $\{x_k^\varepsilon\}$ be the sequences defined by (3.4) when we use the exact and perturbed data y, y^ε respectively.

Notice that this method can be considered as a fixed point iteration $x_{k+1} = \psi(x_k)$ with the fixed point operator $\psi(x) := x - F'(x)^*(F(x) - y)$. Note that ϕ need not to be contractive (e.g., if F is compact and twice continuous Fréchet differentiable, and X is infinite dimensional, then 1 belongs to the spectrum of $\phi'(x_*)$).

Iterative methods for solving fixed point equations for nonexpansive operators (i.e. $\|\phi(x) - \phi(\tilde{x})\| \le \|x - \tilde{x}\|$, for all $x, \tilde{x} \in X$) have been considered in the literature (see [8] for a survey). However, in many practical applications it is virtually impossible to determine whether the operator ϕ is nonexpansive or not.

In [36] the nonexpansivity condition on ϕ is replaced by the following *tangential cone condition* (see also [84])

$$\|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\| \le \eta \|F(x) - F(\tilde{x})\| \, , \ \eta < \tfrac{1}{2} \, , \quad (3.16)$$

for all $x, \tilde{x}$ in $B_\rho(x_0) \subset D(F)$. This strong condition on the nonlinearity of F guarantees local convergence of the iteration x_k to a solution $x_* \in B_{\rho/2}(x_0)$ of (3.1). Further, it also guarantees that all iterates x_k^ε remain in $D(F)$ as long as $k < k(\varepsilon, y^\varepsilon)$, the index defined by the discrepancy principle

$$\|y^\varepsilon - F(x^\varepsilon_{k(\varepsilon,y^\varepsilon)})\| \ \le \ \tau\varepsilon \ < \ \|y^\varepsilon - F(x_k^\varepsilon))\| \, , \quad (3.17)$$

for $0 \le k \le k(\varepsilon, y^\varepsilon)$, where τ satisfies

$$\tau \ > \ 2(1+\eta)(1-2\eta)^{-1} \, . \quad (3.18)$$

Note that the right hand side of (3.18) is strictly greater than 2.

We present a first result concerning the characterization of the solutions of equation (3.1).

Lemma 3.3.1. *Let the tangential cone condition (3.16) be fulfilled and $x_* \in B_\rho(x_0)$ be a solution of (3.1). Then, $\tilde{x}_* \in B_\rho(x_0)$ is another solution iff $x_* - \tilde{x}_* \in null(F'(x_*))$.*

Proof:
Notice that (3.16) implies the inequalities

$$\tfrac{1}{1+\eta} \|F'(x)(x-\tilde{x})\| \ \le \ \|F(x) - F(\tilde{x})\| \ \le \ \tfrac{1}{1-\eta} \|F'(x)(x-\tilde{x})\| \, ,$$

for $x, \tilde{x} \in B_\rho(x_0)$. The assertion follows now from (3.3). ■

As in the linear case, we make an scaling assumption. In the sequel we shall assume

$$\|F'(x)\| \ \le \ 1 \, , \ x \in B_\rho(x_0) \, . \quad (3.19)$$

The next result is the nonlinear analogon of Lemma 3.2.3.

Lemma 3.3.2. *Let $y = Fx_*$ for some $x_* \in B_{\frac{\rho}{2}}(x_0)$, $y^\varepsilon \in Y$ be such that $\|y - y^\varepsilon\| \leq \varepsilon$, and $k(\varepsilon, y^\varepsilon)$ the stopping index defined by the discrepancy principle (3.17), (3.18). If (3.16) and (3.19) hold, then we have*

$$\|x_* - x_{k+1}^\varepsilon\| \leq \|x_* - x_k^\varepsilon\|\,,\ 0 \leq k \leq k(\varepsilon, y^\varepsilon).$$

Moreover, if the data is exact (i.e. $\varepsilon = 0$) then $\sum\limits_{j=0}^{\infty} \|y - F(x_k)\|^2 < \infty$.

Proof:
Notice that, from (3.16) and (3.19) follows that $x_k^\varepsilon \in B_{\frac{\rho}{2}}(x_0)$ for $0 \leq k < k(\varepsilon, y^\varepsilon)$. Moreover,

$$\begin{aligned} &\|x_* - x_{k+1}^\varepsilon\|^2 - \|x_* - x_k^\varepsilon\|^2 \leq \\ &\qquad \|y^\varepsilon - Fx_k^\varepsilon\| \big\{(2\eta - 1)\|y^\varepsilon - Fx_k^\varepsilon\| + 2(1+\eta)\varepsilon\big\} \end{aligned} \tag{3.20}$$

(see Exercise 3.5). Note that the right hand side of (3.20) is nonnegative for $k < k(\varepsilon, y^\varepsilon)$ due to (3.17), proving the first assertion. Notice that, if $\varepsilon = 0$, we obtain instead of (3.20) the sharper estimate

$$\|x_* - x_{k+1}\|^2 + (1 - 2\eta)\|y - F(x_k)\|^2 \ \leq\ \|x_* - x_k\|^2\,,$$

for all $k \geq 0$. The second assertion follows now from the inequality $\sum_{j=0}^{\infty} \|y - F(x_k)\|^2 < \frac{1}{1-2\eta}\|x_* - x_0\|^2$. ■

If $\varepsilon > 0$, one can argue as in Lemma 3.3.2 and prove that

$$\sum_{j=0}^{k(\varepsilon, y^\varepsilon)-1} \|y - F(x_k^\varepsilon)\|^2 \ \leq\ \frac{\tau}{(1-2\eta)\tau - 2(1+\eta)} \|x_* - x_0\|^2\,. \tag{3.21}$$

The next result guarantees convergence for exact data and is the nonlinear analogon of Lemma 3.2.1.

Lemma 3.3.3. *Let (3.1) be solvable in $B_{\frac{\rho}{2}}(x_0)$. If (3.16) and (3.19) hold, then x_k converges to a solution $x_* \in B_{\frac{\rho}{2}}(x_0)$.*

Proof:
Let $\tilde{x}_* \in B_{\frac{\rho}{2}}(x_0)$ be any solution of (3.1). Arguing with (3.3) one

can prove that $x_k - \tilde{x}_*$ is a Cauchy sequence in X. Thus, x_k is also a Cauchy sequence and we denote it's limit by x_*. Since $y - F(x_k) \to 0$ as $k \to \infty$ (cf. second assertion of Lemma 3.3.2) we conclude that $F(x_*) = y$. ■

If, additionally to the assumptions of Lemma 3.3.3, the condition $\text{null}(F'(x^\dagger)) \subset \text{null}(F'(x))$, for all $x \in B_\rho(x_0)$ is fulfilled (here $x^\dagger$ is the (unique) solution of (3.1) of minimal distance to x_0), then x_k converges to $x^\dagger$ as $k \to \infty$.

The next result guarantees convergence for noisy data and characterizes the Landweber iteration as a regularization method for (3.1).

Lemma 3.3.4. *Let $y^\varepsilon \in Y$ be such that $\|y - y^\varepsilon\| \le \varepsilon$ and let $k(\varepsilon, y^\varepsilon)$ be chosen according to (3.17), (3.18). Under the same assumptions of Lemma 3.3.3 we have $x^\varepsilon_{k(\varepsilon,y^\varepsilon)} \to x_*$ as $\varepsilon \to 0$.*

Proof:
Let ε_n be a sequence converging to zero as $n \to \infty$, and let y^{ε_n} be a corresponding sequence of perturbed data. Now take $k_n := k(\varepsilon_n, y^{\varepsilon_n})$ the index defined by the discrepancy principle. In the sequel we consider two cases:
1) If k_n has a finite accumulation point $\bar{k}$, we can assume $k_n = \bar{k}$, $n \in \mathbb{N}$). Therefore, since $\bar{k}$ is fixed, $x^{\varepsilon_n}_{\bar{k}}$ converges to $x_{\bar{k}}$ (the iterate with exact data; see Exercise 3.6) and $F(x^{\varepsilon_n}_{\bar{k}}) \to F(x_{\bar{k}})$ as $n \to \infty$. Now, taking the limit $n \to \infty$ in $\|y^{\varepsilon_n} - F(x^{\varepsilon_n}_{\bar{k}})\| \le \tau\varepsilon_n$, we conclude that $F(x_{\bar{k}}) = y$. Thus, $x_{\bar{k}} = x_*$ and $x^{\varepsilon_n}_{k_n} \to x_*$ as $n \to \infty$.
2) If $k_n \to \infty$ as $n \to \infty$, we can assume k_n is monotonically increasing. From the first part of Lemma 3.3.2 follows $\|x^{\varepsilon_n}_{k_n} - x_*\| \le \|x^{\varepsilon_n}_{k_m} - x_{k_m}\| + \|x_{k_m} - x_*\|$ for $n > m$. The last term on the right hand side of this estimate can be made small by choosing m large. Now, with k_m fixed, we can make the first term small by choosing n appropriately. Thus, $\|x^{\varepsilon_n}_{k_n} - x_*\| \to 0$ as $n \to \infty$. ■

In general, the convergence of $x^\varepsilon_{k(\varepsilon,y^\varepsilon)} \to x_*$ as $\varepsilon \to 0$ can be arbitrarily slow. Examples in the linear case can be found in [23, 31], where y^ε is chosen according to the singular system for the compact operator F. In the sequel we shall use the source condition (3.5) and the representation condition (3.6) in order to prove convergence rates.

Lemma 3.3.5. *Assume problem (3.1) is solvable in $B_{\frac{\varrho}{2}}(x_0)$ and let $y^\varepsilon \in Y$ be such that $\|y - y^\varepsilon\| \le \varepsilon$. Moreover, assume that the operator F fulfills (3.16), (3.19) and the representation condition (3.6), (3.7). If $x^\dagger - x_0$ satisfies the source condition (3.5) with $\nu \le 1/2$ and $\|w\|$ sufficiently small, then there exists a constant $c = c(\nu)$ such that*

$$\begin{aligned} \|x^\dagger - x_k^\varepsilon\| &\le c(\nu)\|w\|(k+1)^{-\nu} \\ \|y^\varepsilon - F(x_k^\varepsilon)\| &\le 4c(\nu)\|w\|(k+1)^{-\nu-1/2} \end{aligned}$$

for $0 \le k < k(\varepsilon, y^\varepsilon)$, where $k(\varepsilon, y^\varepsilon)$ is the index defined by the discrepancy principle (3.17), (3.18). Here $x^\dagger$ denotes the (unique) solution of (3.1) of minimal norm.

Sketch of the proof:
From Lemma 3.3.2 we conclude that the iteration x_k^ε is well defined in $B_{\frac{\varrho}{2}}(x_0) \subset D(F)$ for $0 \le k \le k(\varepsilon, y^\varepsilon)$. Moreover, from (3.21) follows $k(\varepsilon, y^\varepsilon) < \infty$ for $\varepsilon > 0$. Next one defines the error $e_k := x^\dagger - x_k^\varepsilon$ and uses (3.16), (3.7) to obtain the expression

$$\begin{aligned} e_k = {} & (I - K^*K)^k e_0 \\ & + \sum_{j=0}^{k-1} (I - K^*K)^j K^* z_{k-j-1} + \left[\sum_{j=0}^{k-1} (I - K^*K)^j \right] (y - y^\varepsilon)\,, \end{aligned}$$

where $K := F'(x^\dagger)$ and the norm of the z_k can be estimated (up to a constant) by $\|e_k\|\|Ke_k\|$, for $0 \le k \le k(\varepsilon, y^\varepsilon)$. To prove the Lemma, it is enough to obtain as adequate estimate for $\|e_k\|$ and $\|Ke_k\|$. From the above representation of e_k with (3.16), (3.17) and (3.18) one obtains

$$\|e_k\| \;\le\; c_1(\eta)c_2(\nu)\|w\|(k+1)^{-\nu}\,,$$

which proves the first part of the lemma, since $c_2(\nu)$ can be made smaller than 2 if $\|w\|$ is sufficiently small. The second part of the lemma follows from the estimate

$$\|Ke_k\| \;\le\; c_1(\eta)c_2(\nu)\|w\|(k+1)^{-\nu-1/2}$$

and an analog argumentation. ■

In the sequel we prove the main result of this section, obtaining estimates for $k(\varepsilon, y^\varepsilon)$ as well as for $\|x^\dagger - x^\varepsilon_{k(\varepsilon,y^\varepsilon)}\|$ in terms of ε and ν (compare with Lemma 3.2.5).

Lemma 3.3.6. *Under the assumptions of Lemma 3.3.5, we have*

$$\begin{aligned} k(\varepsilon, y^\varepsilon) \ &\le\ c_1(\|w\|/\varepsilon)^{2/(2\nu+1)} \\ \|x^\dagger - x^\varepsilon_{k(\varepsilon,y^\varepsilon)}\| \ &\le\ c_2\|w\|^{1/(2\nu+1)}\varepsilon^{2\nu/(2\nu+1)} \end{aligned}$$

with $c_j = c_j(\nu)$, $j = 1, 2$.

Sketch of the proof:
We use the same notation as in Lemma 3.3.5. To simplify the notation we write $k_\varepsilon = k(\varepsilon, y^\varepsilon)$. As in the proof of Lemma 3.3.5 we obtain

$$e_{k_\varepsilon} = (I - K^*K)^\nu w_{k_\varepsilon} + \sum_{j=0}^{k_\varepsilon - 1} (I - K^*K)^j K^* \ (y - y^\varepsilon) \tag{3.22}$$

where

$$w_{k_\varepsilon} = (I - K^*K)^{k_\varepsilon} w + \sum_{j=0}^{k_\varepsilon - 1} (I - K^*K)^j (K^*K)^{1/2-\nu} \tilde{z}_{k_\varepsilon - j - 1} \,,$$

and $\|\tilde{z}_j\| = \|z_j\|$. Therefore, we can estimate $\|w_{k_\varepsilon}\| \le (1 + c(\nu))\|w\|$ and, consequently, $\|K(K^*K)^\nu w_{k_\varepsilon}\| \le c\|w\|^{1/(2\nu+1)}\varepsilon^{2\nu/(2\nu+1)}$. Using these estimates and (3.22) we obtain

$$\|e_{k_\varepsilon}\| \le \|(K^*K)^\nu w_{k_\varepsilon}\| + \sqrt{k_\varepsilon}\varepsilon \tag{3.23}$$

proving the lemma for $k_\varepsilon = 0$. Otherwise, an estimate analog to the one used in the proof of Lemma 3.3.5 gives

$$k_\varepsilon^{\nu+1/2} \le c(\nu)\|w\|/\varepsilon$$

(here $c(\nu)$ is the same constant as in Lemma 3.3.5). The second assertion of the lemma follows now from this inequality. Further, the first assertion (error estimate) follows if we substitute the last inequality in (3.23). ■

Notice that (3.6) implies $\mathrm{null}(F'(x_*)) \subset \mathrm{null}(F(x))$ for all $x \in B_\rho(x_0)$. Further, one can prove that the tangential cone condition (3.16) with $\tilde{x} = x_*$ follows from (3.6) and (3.7) (cf. Exercise 3.7).

3.4 A modified Landweber iteration

In the sequel we shall consider the same nonlinear problem as in Section 3.3. An alternative to the iteration defined in (3.4) is investigated, namely

$$x_{k+1} \;=\; x_k - F'(x_k)^*(F(x_k) - y) - \alpha_k(x_k - \xi)\,, \tag{3.24}$$

where $0 \le \alpha_k \le 1$ and $\xi \in B_\rho(x_0) \subset D(F)$. A remarkable advantage of this iteration is the fact that one can prove convergence rates results (under the usual source condition (3.5)) without requiring a representation condition on F' (as in (3.6)).

If the iteration (3.24) is applied to noisy data, we write x_k^ε instead of x_k. As in the previous sections, we assume $x_0^\varepsilon = x_0$. Further, it is assumed that

$$\|F'(x)\| \;\le\; L\,,\ x \in B_\rho(x_0) \tag{3.25}$$

(compare with (3.19) in Section 3.3).

The first result concerns a monotonicity property of the modified Landweber iteration.

Lemma 3.4.1. *Let F satisfy (3.16), (3.25) and the sequence $\{\alpha_k\}$ be chosen as above. Further, let x_* be a solution of (3.1) in $B_{\rho/8}(x_0) \cap B_{\rho/8}(\xi)$. The we have:*

a) *Denote by k_ε the stopping index defined by the discrepancy principle (3.17) with τ satisfying*

$$(1-\alpha_k)\big(1 - \eta - \frac{1+\eta}{\tau}\big) - L^2 \ge D > 0\,,\ 0 \le k < k_\varepsilon\,.$$

Then, for $0 \le k < k_\varepsilon$, we have $x_{k+1}^\varepsilon \in B_\rho(x_0)$ and

$$\|x_* - x_{k+1}^\varepsilon\| \le \|x_* - x_k^\varepsilon\|(1-\alpha_k) + \tfrac{\rho}{2}\alpha_k \le \tfrac{\rho}{2}\,.$$

b) *If $\varepsilon/\alpha_k \le C$ for $0 \le k \le N_0$, with $C \le \rho/6$, and if $(1-\alpha_k)(1-\eta) - L^2 \ge E > 0$ for $0 \le k \le N_0$, then, for $0 \le k \le N_0$, we have $x_{k+1}^\varepsilon \in B_\rho(x_0)$ and*

$$\|x_* - x_{k+1}^\varepsilon\| \le \|x_* - x_k^\varepsilon\|(1-\alpha_k) + \tfrac{\rho}{2}\alpha_k \le \tfrac{\rho}{2}\,.$$

In Lemma 3.4.1, D and E are fixed positive constants. Notice that item a) furnishes an *a posteriori* stopping rule, while item b) corresponds to an *a priori* stopping rule.

Sketch of the proof:
To prove a) the first step is to derive from (3.24) the inequality

$$\begin{aligned}\|x_* - x_{k+1}^\varepsilon\|^2 \ \le\ & (1-\alpha_k)^2\|x_* - x_k^\varepsilon\|^2 + 2\alpha_k^2\|x_* - \xi\|^2 \\ & + 2\|F(x_k^\varepsilon) - y^\varepsilon\|^2(L^2 - (1-\alpha_k)(1-\eta)) \\ & + 2\alpha_k(1-\alpha_k)\|x_* - x_k^\varepsilon\|\|x_* - \xi\| \\ & + 2\varepsilon\|F(x_k^\varepsilon) - y^\varepsilon\|(1-\alpha_k)(1+\eta)\,. \qquad (3.26)\end{aligned}$$

From (3.17) and the choice of τ in a) follows

$$\|x_* - x_{k+1}^\varepsilon\| \le \|x_* - x_k^\varepsilon\|(1-\alpha_k) + \tfrac{\rho}{4}\alpha_k\,.$$

Using this last inequality inductively for $0 \le k < k_\varepsilon$ together with Exercise 3.8, we obtain $\|x_* - x_{k+1}^\varepsilon\| \le \rho/2$. From this and the triangle inequality follows $\|x_{k+1}^\varepsilon - x_0\| \le \rho$, proving assertion a). The proof of assertion b) follows from an analog argumentation. It is worth noticing that the estimate

$$\begin{aligned}\|x_* - x_{k+1}^\varepsilon\|^2 \ \le\ & (1-\alpha_k)^2\|x_* - x_k^\varepsilon\|^2 \\ & + 2\|F(x_k^\varepsilon) - y^\varepsilon\|^2(L^2 - (1-\alpha_k)(1-\eta)) \\ & + 2\alpha_k(1-\alpha_k)\|x_* - x_k^\varepsilon\|\big(\|x_* - \xi\| + CL(1+\eta)\big) \\ & + 2\alpha_k^2\left[(1-\alpha_k)(1+\eta) + \|x_* - \xi\|\right] \qquad (3.27)\end{aligned}$$

replaces (3.26), used in the proof of the first assertion. ■

Notice that the *a posteriori* stopping rule defined in item a) of Lemma 3.4.1 can be applied to the Landweber iteration. However, the *a priori* stopping rule in item b) cannot (since $\alpha_k = 0$ for the Landweber iteration).

The next result generalizes the second assertion of Lemma 3.3.2 as well as (3.21) for the modified Landweber iteration.

Lemma 3.4.2. *Under the assumptions of Lemma 3.4.1, if* $\sum_{k=0}^\infty \alpha_k < \infty$ *then*

a) *If (3.24) is stopped according to (3.17), with constants τ, D as in assertion a) of Lemma 3.4.1, then*

$$\sum_{k=0}^{k_\varepsilon-1} \|F(x_k^\varepsilon) - y^\varepsilon\| \ \le\ \tfrac{\rho^2}{D\tau\varepsilon}\Big(\tfrac{1}{64} + \sum_{k=0}^{k_\varepsilon-1} \alpha_k\Big)\,.$$

b) *If (3.24) is stopped according to assertion b) in Lemma 3.4.1, then*

$$\sum_{k=0}^{N_0} \|F(x_k^\varepsilon) - y^\varepsilon\|^2 \ \le\ \tfrac{\rho^2}{E}\Big(\tfrac{1}{64} + \sum_{k=0}^{N_0} \alpha_k\Big)\,.$$

c) *If* $(1-\alpha_k)(1-\eta) - L^2 \ge E > 0$ *for* $k \ge 0$, *and* $\varepsilon = 0$, *then*

$$\sum_{k=0}^{\infty} \|F(x_k) - y\|^2 \ <\ \infty\,.$$

Sketch of the proof:
First we prove a). From Lemma 3.4.1 follows $\|x_* - x_k^\varepsilon\| \le \rho/2$, $0 \le k < k_\varepsilon$, and we can estimate

$$\|x_* - x_{k+1}^\varepsilon\|^2 \ \le\ \|x_* - x_k^\varepsilon\|^2 + \rho^2\alpha_k\,.$$

Assertion a) follows now arguing with Lemma 3.4.1 and (3.17). To prove b) one proceeds as in the proof of Lemma 3.4.1 b) and obtains from (3.27) the estimate

$$\|x_* - x_{k+1}^\varepsilon\|^2 + E\|F(x_k^\varepsilon) - y^\varepsilon\|^2 \ \le\ \|x_* - x_k^\varepsilon\|^2(1-\alpha_k)^2 + \rho^2\alpha_k\,.$$

Assertion b) follows now from this last inequality. In order to prove c) one should notice that, in the noise free case ($\varepsilon = 0$), the estimate (3.26) implies

$$\begin{aligned}&\|x_* - x_{k+1}\|^2 + E\|F(x_k) - y\|^2\\ &\le\ (1-\alpha_k)^2\|x_*-x_k\|^2 + 2\alpha_k^2\|x_*-\xi\|^2 + 2\alpha_k(1-\alpha_k)\|x_*-x_k\|\|x_*-\xi\|\,.\end{aligned}$$

Therefore, $x_k \in B_\rho(x_0)$, $k \ge 0$, and consequently

$$\sum_{k=0}^{\infty} \|F(x_k) - y\|^2 \ \le\ \tfrac{\rho^2}{E}\Big(\tfrac{1}{64} + \sum_{k=0}^{\infty} \alpha_k\Big)\,,$$

completing the proof. ■

Convergence of the modified Landweber iteration for exact data can be proved under additional assumptions on $\{\alpha_k\}$.

Lemma 3.4.3. *Assume the data is exact, i.e. $\varepsilon = 0$, and α_k satisfy $0 \leq \alpha_k \leq 1$, $\sum_{k=0}^{\infty} \alpha_k < \infty$. Let the operator F satisfies (3.16) and (3.25) with L satisfying $(1-\alpha_k)(1-\eta) - L^2 \geq E > 0$ for $k \geq 0$. Moreover, assume (3.1) is solvable in $B_{\rho/8}(x_0) \cap B_{\rho/8}(\xi)$. Then x_k defined by (3.24) converges to a solution $x_* \in B_\rho(x_0)$. Further, if one choose $\xi = x_0$ in (3.25) and null$(F'(x^\dagger)) \subset$ null$(F'(x))$ for all $x \in B_\rho(x_0)$, then $x_k \to x^\dagger$ as $k \to \infty$.*[3]

Sketch of the proof:
Let $\tilde{x}_* \in B_{\rho/8}(x_0) \cap B_{\rho/8}(\xi)$ be any solution of (3.1) and define $e_k := x_k - \tilde{x}_*$. First one proves that $\{\|e_k\|\}$ is a convergence sequence. This is an advanced analysis exercise, that can be solved using the assertions in Exercise 3.8. To prove of the first assertion, it is enough to prove that $\{e_k\}$ is a Cauchy sequence. This follows from the convergence of $\{\|e_k\|\}$ together with Lemma 3.4.1 a) and Exercise 3.8.
To prove the second assertion, notice that $x^\dagger - x_0 \in \text{null}(F'(x^\dagger))^\perp$. Since $\text{null}(F'(x^\dagger)) \subset \text{null}(F'(x_k))$, for all k, then $x_k - x_0 \in N(F'(x^\dagger))^\perp$. Therefore, $x^\dagger - x_* \in N(F'(x^\dagger))^\perp$. If $x^\dagger \neq x_*$, then follows from (3.16)

$$\|F'(x^\dagger)(x_* - x^\dagger)\| \ \leq \ (1+\eta)\|F(x^\dagger) - F(x_*)\| \ \leq \ 0 \,.$$

Thus, $x^\dagger - x_*$ also belongs to $N(F'(x^\dagger))$ and $x^\dagger = x_*$ follows. ■

From Lemma 3.4.2 follows that, in the case of noisy data ($\varepsilon > 0$), the discrepancy principle (3.17) with τ as in Lemma 3.4.1 a) determines a well-defined and finite stopping index k_ε. The next result characterizes the modified Landweber iteration with this stopping criterion as a regularization method.

Lemma 3.4.4. *Let α_k satisfy $0 \leq \alpha_k \leq 1$, $\sum_{k=0}^{\infty} \alpha_k < \infty$. Further, let F satisfy (3.16), (3.25) with L satisfying $(1-\alpha_k)(1-\eta) - L^2 \geq E > 0$ for $k \geq 0$. Moreover, assume (3.1) is solvable in $B_{\rho/8}(x_0) \cap B_{\rho/8}(\xi)$. Then we have*
a) *If the iteration (3.24) is stopped at the index k_ε defined by (3.17) with τ, D as in assertion a) of Lemma 3.4.1, then $x_{k_\varepsilon}^\varepsilon \to x_*$ as $\varepsilon \to 0$.*

[3] Here $x^\dagger$ denotes the (unique) solution of (3.1) of minimal distance to x_0. Compare with Lemma 3.3.5.

b) *If the iteration (3.24) is stopped at the index N_0^ε defined by assertion b) of Lemma 3.4.1, and $\{\alpha_k\}$ is a strictly monotonically decreasing sequence with $\alpha_k > 0$, then $x^\varepsilon_{N_0^\varepsilon} \to x_*$ as $\varepsilon \to 0$.*

Sketch of the proof:
We consider assertion a) first. Let $\varepsilon_n \geq 0$, $n \in \mathbb{N}$, be a sequence converging to zero and $y_n := y^{\varepsilon_n}$ be a corresponding sequence of perturbed data. We denote by k_n the index defined by assertion a) of Lemma 3.4.1 for each pair (ε_n, y_n). We consider two cases:
Case I: the sequence $\{k_n\}$ has a finite accumulation point; we can assume, without loss of generality, that $k_n = k$ constant for all $n \in \mathbb{N}$. Thus, $\|y_n - F(x_k^{\varepsilon_n})\| \leq \tau\varepsilon_n$, and from the continuous dependence of x_k^ε on y_ε (k is now fixed) follows $x_k^{\varepsilon_n} \to x_k$ and $F(x_k^{\varepsilon_n}) \to F(x_k)$ as $n \to \infty$. Since $\tau\varepsilon_n \to 0$ as $n \to \infty$, it follows from the above estimate that $F(x_k) = y$. Therefore, $x_k = x_*$ and the assertion follows.
Case II: the sequence $\{k_n\}$ has no finite accumulation point, i.e. $k_n \to \infty$, as $n \to \infty$. We can assume, without loss of generality, that k_n is monotone. Then, for $n > m$ it follows from Lemma 3.4.1 that

$$\|x_{k_n}^{\varepsilon_n} - x_*\| \leq \|x_{k_m}^{\varepsilon_n} - x_{k_m}\| + \|x_{k_m} - x_*\| + \tfrac{\rho}{2} \sum_{j=k_m}^{\infty} \alpha_j \, .$$

First we fix m such that the last two terms on the right hand side become small. Since, with m fixed, we have $x_{k_m}^{\varepsilon_n} \to x_{k_m}$ as $\varepsilon_n \to 0$, then the right hand side must go to zero as $n \to \infty$ an the proof of assertion a) is complete.

Now we prove assertion b). Let $\varepsilon_n \geq 0$, $n \in \mathbb{N}$, be a (strictly monotone) sequence converging to zero and $y_n := y^{\varepsilon_n}$ be a corresponding sequence of perturbed data. We denote by $N_0^{\varepsilon_n}$ the index defined by assertion b) of Lemma 3.4.1 for each pair (ε_n, y_n). Therefore, $\varepsilon_n < C\alpha_k$, for $0 \leq k \leq N_0^{\varepsilon_n}$, and $\varepsilon_n \geq C\alpha_{N_0^{\varepsilon_n}+1}$. Since both α_k and ε_n are strictly monotone sequences, $N_0(\varepsilon_n)$ is strictly monotone in n and $N_0^{\varepsilon_n} \to infty$ as $n \to \infty$. Thus, for $n > m$ we have

$$\|x^{\varepsilon_n}_{N_0^{\varepsilon_n}} - x_*\| \leq \|x^{\varepsilon_n}_{N_0^{\varepsilon_m}} - x_{N_0^{\varepsilon_m}}\| + \|x_{N_0^{\varepsilon_m}} - x_*\| + \tfrac{\rho}{2} \sum_{j=N_0^{\varepsilon_m}}^{\infty} \alpha_j \, .$$

The rest of the proof is analogous to the proof of assertion a). ■

Notice that Lemma 3.4.4 is the analogon of Lemma 3.3.4 for the modified Landweber iteration (3.24).

The next result gives convergence rates for the modified Landweber iteration. As in the previous sections, such rates are obtained assuming a sourcewise representation of the exact solution of (3.1). The reader should compare with the results in Lemmas 3.2.5 and 3.3.6.

Lemma 3.4.5. *Let x_* be a solution of (3.1) in $B_{\rho/2}(x_0)$ and assume that the Fréchet derivative of the operator F satisfy*

$$\|F'(x)\| \ \leq L\,, \quad \|F'(x) - F'(\tilde{x})\| \ \leq \ \hat{L}\|x - \tilde{x}\|\,, \ x, \tilde{x} \in B_\rho(x_0)\,,$$

where max$\{L, \hat{L}\} \leq 1/4$. Further, assume that the source condition $x_* - \xi = F'(x_*)^* w$ is fulfilled. Moreover, assume that $\alpha_k = (k + l_0)^{-\psi}$ for $k \geq 0$, where $0 < \psi < 1$ is fixed and $l_0 \in \mathbb{N}$ is sufficiently large (it is enough to take $l_0^{-\psi} \leq 1/8$). Further we require

$$\Phi(k) \ := \ \frac{1}{(1+1/k)^\psi} \frac{1-(1+1/k)^\psi}{1/k} \frac{1}{k^{1-\psi}} + 1 \ \geq \ \hat{L}^2\,,$$

and

$$\alpha_0^{-1}\|x_0 - x_*\|^2 \ \leq \hat{C} \ < \ \min\left\{1, \ \rho^2/(4\alpha_0)\right\}.$$

a) *If the iteration (3.24) is stopped at the index N_0^ε defined by* a priori *stopping rule*

$$\frac{\varepsilon}{\alpha_k} \ \leq \ C\,, \ 0 \leq k \leq N_0^\varepsilon\,, \quad \frac{\varepsilon}{\alpha_{N_0^\varepsilon + 1}} \ > \ C\,,$$

and $2\|x_* - \xi\|^2 + 17C^2 + 2\|w\|^2 \leq \hat{L}^2\hat{C}/2$, then

$$\|x^\varepsilon_{N_0^\varepsilon+1} - x_*\| = O(\sqrt{\varepsilon}).$$

b) *If the iteration (3.24) is stopped at the index k_ε defined by (3.17) with τ satisfying $1 + 2L^2 - 7/8(3/2 - \tau/2) + 1/\tau^2 \ \leq \ E \ < \ 0$ and $2\|x_* - \xi\|^2 + 2\|w\|^2 \leq \hat{L}^2\hat{C}/2$, then*

$$\|x^\varepsilon_{k_\varepsilon} - x_*\| = O(\sqrt{\varepsilon}).$$

c) *In the case $\varepsilon = 0$, if $2\|x_* - \xi\|^2 + \|w\|^2 \leq \hat{L}^2\hat{C}/2$, then*

$$\|x_k - x_*\| = O(k^{-\psi/2}).$$

Sketch of the proof:
The proof is long and requires the derivation of several inequalities (the main ones are presented in this text). The convergence rates results follow than from an inductive argument. The first step of the proof is the derivation of the estimate

$$\begin{aligned}\|x_* - x_{k+1}^\varepsilon\|^2 \ \leq\ & (1-\alpha_k)^2\|x_* - x_k^\varepsilon\|^2 + 2\alpha_k^2\|x_* - \xi\|^2 \\ & + 2\|F'(x_k^\varepsilon)^*(F(x_k^\varepsilon) - y^\varepsilon)\|^2 \\ & - 2\alpha_k(1-\alpha_k)\langle F'(x_*)(x_k^\varepsilon - x_*), w\rangle \\ & - 2(1-\alpha_k)\langle F(x_k^\varepsilon) - y^\varepsilon, F'(x_k^\varepsilon)(x_k^\varepsilon - x_*)\rangle\,.\end{aligned}$$

Using the assumptions this estimate can be improved, and we obtain

$$\begin{aligned}\|x_* - x_{k+1}^\varepsilon\|^2 \ \leq\ & \|x_* - x_k^\varepsilon\|^2(1-\alpha_k)(1-\alpha_k(1-\hat{L}\|w\|)) + \tfrac{\hat{L}^2}{2}\|x_* - x_k^\varepsilon\|^4 \\ & + 2\alpha_k[\alpha_k\|x_* - \xi\|^2 + (1-\alpha_k)\|F(x_k^\varepsilon) - y^\varepsilon\|\|w\|] \\ & + \|F(x_k^\varepsilon) - y^\varepsilon\|^2(2L^2 - \tfrac{3}{2}(1-\alpha_k)) \\ & + 2(1-\alpha_k)\varepsilon(\|F(x_k^\varepsilon) - y^\varepsilon\| + \alpha_k\|w\|)\,.\end{aligned}$$

In order to prove assertion a) one uses the above estimate together with the assumptions in a) and an inductive argument. The proof of assertion b) is analogous. It follows basically from the above estimate, the assumptions in b) and (again) an inductive argument. Assertion c) follows from the above estimate and $1 + 2L^2 - 3(1-\alpha_k)/2 < 0$. Indeed, combining these two inequalities we obtain

$$\begin{aligned}\|x_* - x_{k+1}\|^2 \ \leq\ & \|x_* - x_k\|^2(1-\alpha_k) + \tfrac{\hat{L}^2}{2}\|x_* - x_k\|^4 \\ & + \alpha_k^2[2\|x_* - \xi\|^2 + (1-\alpha_k)^2\|w\|^2] \\ & + \|F(x_k) - y^0\|^2(1 + 2L^2 - \tfrac{3}{2}(1-\alpha_k))\,,\end{aligned}$$

and the proof follows analogous to the proof of the previous assertions. ■

3.5 Asymptotical regularization

In this section we consider the method of asymptotical regularization for solving the inverse problem (3.1). This method is the continuous analogon of method (3.4). In this method an approximation $x(T)$ of a solution x_* of (3.1) is obtained by solving the initial value problem (3.9). The *final time* plays the rule of the regularization parameter. From the theory of ordinary differential equations (see, e.g., [46]) one knows that, for finite $T < \infty$, problem (3.9) admits a unique solution in $C(0,T;X)$ if the operator $G(x) = F'(x)^*(y - F(x))$ is locally Lipschitz continuous in X.

We shall discuss some well known properties of the asymptotical regularization for both linear and nonlinear problems. The results related to the linear theory can be found in [92] and also in the more modern textbook [23]. Results related to the nonlinear theory can be found in [89, 90].

We start with the linear inverse problems, i.e. we assume that the operator F in (3.1) is a linear bounded operator between the infinite dimensional Hilbert spaces X and Y with non closed range. If the evolution (3.9) is applied with noisy data $y^\varepsilon \in Y$ (with $\|y - y^\varepsilon\| \le \varepsilon$) instead of y, then we write $x^\varepsilon(t)$ instead of $x(t)$. Moreover, we assume that problem (3.1) has a solution x_*, which need not to be unique (see Section 3.1).

The next two lemmas summarize some relevant properties selected from [92, 23] of the asymptotical regularization for linear equations.

Lemma 3.5.1. *Let F be a linear bounded operator with non closed range range$(F) \subset Y$ and $F^\dagger y$ be the unique solution of (3.1) with minimal distance to $x_0 \in X$. Then,*
a) *$x(T) \to F^\dagger y$ as $T \to \infty$ (convergence for exact data);*

b) *$x^\varepsilon(T) \to F^\dagger y$ for $T \to \infty$ such that $\varepsilon^2 T \to 0$ (convergence);*

c) *Let $x_* - x_0 = (F^*F)^\nu w$ for some $w \in X$ and $\nu > 0$. If T is chosen according to the* a priori *parameter choice $T = C(\varepsilon/\|w\|)^{-2/(2\nu+1)}$ with $C > 0$, then we have the error estimate*

$$\|x^\varepsilon(T) - F^\dagger y\| \;\le\; c\|w\|^{1/(2\nu+1)}\varepsilon^{2\nu/(2\nu+1)},$$

where the constant $c > 0$ depends only on C and ν (convergence rates under source conditions).

The next result concerns the asymptotical regularization with the stopping rule given by the discrepancy principle. According to this principle, the evolution should be terminated at the time $T_\varepsilon = T(\varepsilon, y^\varepsilon)$ such that

$$\|Fx^\varepsilon(t) - y^\varepsilon\| > \tau\varepsilon\,,\ t \in [0, T_\varepsilon) \qquad \|Fx^\varepsilon(T_\varepsilon) - y^\varepsilon\| = \tau\varepsilon \tag{3.28}$$

holds with a constant $\tau \geq 1$. The assumption $\|Fx_0 - y^\varepsilon\| > \tau\varepsilon$ guarantees that the time stopping time T_ε in (3.28) is well defined.

Lemma 3.5.2. *Let T_ε be given by the discrepancy principle (3.28) with $\tau \geq 1$ and $F^\dagger y$ be the unique solution of (3.1) with minimal distance to $x_0 \in X$. Then we have*
a) *$x^\varepsilon(T_\varepsilon) \to F^\dagger y$ as $\varepsilon \to 0$;*

b) *If $x_* - x_0 = (F^*F)^\nu w$ for some $w \in X$ and $\nu > 0$, then we have the error estimate $\|x^\varepsilon(T_\varepsilon) - F^\dagger y\| \leq c\|w\|^{1/(2\nu+1)}((\tau+1)\varepsilon)^{2\nu/(2\nu+1)}$, where the constant $c > 0$ is independent of $\|w\|$ and ε.*

In the sequel we devote our attention to convergence properties of the asymptotical regularization method for nonlinear ill-posed problems (3.1). The basic tools needed for the proof of the convergence results have been developed in [36] and are discussed in Section 3.3.

The first result concerns a monotony property related to the solution of (3.9).

Lemma 3.5.3. *Let $x^\varepsilon(t)$ be the solution of (3.9) with $y = y^\varepsilon$. Then, for $t > 0$ we have*

$$\frac{d}{dt}\|F(x^\varepsilon(t)) - y^\varepsilon\|^2 \ = \ -2\|F'(x^\varepsilon(t))^*(F(x^\varepsilon(t)) - y^\varepsilon)\|^2\,.$$

Moreover, if $x_* \in B_\rho(x_0) \subset D(F)$ is a solution of (3.1) and F satisfies (3.16) with $\eta < 1$, then we have the estimate

$$\frac{d}{dt}\|x^\varepsilon(t) - x_*\|^2 \leq -2\|F(x^\varepsilon(t)) - y^\varepsilon\|\big\{(1-\eta)\|F(x^\varepsilon(t)) - y^\varepsilon\| - (1+\eta)\varepsilon\big\},$$

and, if the data is exact (i.e. $\varepsilon = 0$) then

$$\int_0^\infty \|F(x(t)) - y\|^2\, dt \ \leq \ \tfrac{1}{2(1-\eta)}\|x_* - x_0\|^2\,.$$

Sketch of the proof:
The proof is completely analogous to the proof of Lemma 3.3.2. The first assertion follows from the estimate

$$\frac{d}{dt}\|F(x^\varepsilon(t)) - y^\varepsilon\|^2 \; = \\ - 2\langle F'(x^\varepsilon(t))F'(x^\varepsilon(t))^*[F(x^\varepsilon(t)) - y^\varepsilon], F(x^\varepsilon(t)) - y^\varepsilon\rangle \, .$$

The second assertion follows basically from the estimate

$$\frac{d}{dt}\|x^\varepsilon(t) - x_*\|^2 \; \leq \\ 2\|F(x^\varepsilon(t)) - y^\varepsilon\|\big\{\eta(\|F(x^\varepsilon(t)) - y^\varepsilon\| + \varepsilon) - \|F(x^\varepsilon(t)) - y^\varepsilon\| + \varepsilon\big\} \, .$$

The last assertion ($\varepsilon = 0$) follows when we integrate both sides of the second assertion at $[0, \infty)$. ■

The next result gives sufficient conditions to guarantee that the equation defining the discrepancy principle (i.e. $\|F(x^\varepsilon(t)) - y^\varepsilon\| = \tau\varepsilon$) has a unique solution $t = T_\varepsilon$ in $(0, \infty)$.

Lemma 3.5.4. *Let $x^\varepsilon(t)$ be the solution of (3.9) with $y = y^\varepsilon$ and let $x_* \in B_\rho(x_0) \subset D(F)$ be a solution of (3.1). Furthermore, let the tangential cone condition (3.16) be fulfilled with $\eta < 1$. If $\|F(x_0) - y^\varepsilon\| > \tau\varepsilon > 0$, and $\tau > (1+\eta)/(1-\eta)$ hold, then there is a unique $T_\varepsilon \in (0, \infty)$ satisfying $\|F(x^\varepsilon(T_\varepsilon)) - y^\varepsilon\| = \tau\varepsilon$.*

The main argument in the proof of Lemma 3.5.4 is the monotonicity of the application $t \mapsto \|F(x^\varepsilon(t)) - y^\varepsilon\| - \tau\varepsilon$, which follows from Lemma 3.5.3. An important sub-product of the proof of Lemma 3.5.4 is the inequality $\|x^\varepsilon(t) - x_*\| \leq \|x_* - x_0\|$, $t \leq T_\varepsilon$. This inequality together with the triangle inequality $\|x^\varepsilon(t) - x_0\| \leq \|x^\varepsilon(t) - x_*\| + \|x_* - x_0\|$ results in $x^\varepsilon(t) \in B_\rho(x_0)$ for $t \leq T_\varepsilon$ and $\rho = 2\|x_* - x_0\|$.

The next lemma proves convergence of the asymptotical regularization method for exact data (i.e. $\varepsilon = 0$).

Lemma 3.5.5. *Let problem (3.1) be solvable in $B_\rho(x_0) \subset D(F)$ and assume that F satisfies the tangential cone condition (3.16) with $\eta < 1$. Then we have $x(t) \to x_*$ as $t \to \infty$, where $x_* \in B_\rho(x_0)$ is a solution of (3.1).*

Sketch of the proof:
The proof uses the same ideas as the proof of Lemma 3.3.3. Let $\tilde{x}_* \in B_\rho(x_0)$ be any solution of (3.1). Define $e(t) := \tilde{x}_* - x(t)$, for $t \geq 0$ and obtain

$$\|e(t) - e(s)\|^2 \;=\; 2\langle e(s) - e(t), e(s)\rangle + \|e(t)\|^2 - \|e(s)\|^2 \, .$$

From Lemma 3.5.3 follows the existence of the limit $\lim_{t\to\infty} \|e(t)\|$. Therefore, the sum of the last two terms in (3.5) becomes arbitrarily small if s and t are large. Moreover, since

$$|\langle e(s) - e(t), e(s)\rangle| \;\leq\; 3(1+\eta) \int_s^t \|F(x(r)) - y\|^2 dr \, ,$$

it follows from Lemma 3.5.3 that $\langle e(s) - e(t), e(s)\rangle$ becomes arbitrarily small if s and t are large. Consequently, $\lim_{t\to\infty} e(t)$ exists and, therefore, $\lim_{t\to\infty} x(t)$ also exists. Denoting this limit by x_*, we conclude (since $\lim_{t\to\infty} \|F(x(t)) - y\| = 0$ from Lemma 3.5.3) that $F(x_*) = y$, i.e. x_* is a solution of (3.1). ■

If, additionally to the assumptions of Lemma 3.5.5, the condition $\text{null}(F'(x^\dagger)) \subset \text{null}(F'(x))$ for all $x \in B_\rho(X_0)$, is fulfilled (here $x^\dagger$ is the (unique) solution of (3.1) of minimal distance to x_0), then $x(t)$ converges to $x^\dagger$ as $k \to \infty$ (compare with Section 3.3).

The next result guarantees the convergence of the asymptotical regularization method for noisy data, if the discrepancy principle is used as stopping rule.

Lemma 3.5.6. *Let $x^\varepsilon(t)$ be the solution of (3.9) with $y = y^\varepsilon$ and let problem (3.1) be solvable in $B_\rho(x_0) \subset D(F)$. Furthermore, let the tangential cone condition (3.16) be fulfilled with $\eta < 1$, and assume that $\|F(x_0) - y^\varepsilon\| > \tau\varepsilon > 0$, and $\tau > (1+\eta)/(1-\eta)$ hold. Moreover, let $T_\varepsilon \in (0, \infty)$ be defined by the discrepancy principle as in Lemma 3.5.4. Then, $x^\varepsilon(T_\varepsilon) \to x_*$ as $\varepsilon \to 0$.*

Sketch of the proof:
The proof can be carried out using the same method of proof of Lemma 3.3.4, where a corresponding result for the nonlinear Landweber iteration is proved. ■

The next results give rates of convergence for the asymptotical regularization method. For this purpose we need some extra assumptions. We assume that, instead of the sourcewise representation (eq:src-cond) we have

$$x^\dagger - x_0 \ = \ (F'(x^\dagger)^* F'(x^\dagger))^\nu w \, , \ 0 < \nu \leq \tfrac{1}{2} \ , \ w \in X \, , \qquad (3.29)$$

where $x^\dagger \in D(F)$ denotes the solution of (3.1) with minimal distance to x_0. Further, we assume (locally) a representation condition on F':

$$F'(x) \ = \ R_x F'(x^\dagger) \, , \ x \in B_\rho(x_0) \, , \qquad (3.30)$$

where $\{R_x\}_{x \in B_\rho(x_0)}$ is a family of bounded linear operators $R_x : Y \to Y$ with

$$\|R_x - I\| \ \leq C\|x - x^\dagger\| \, , \ x \in B_\rho(x_0) \, . \qquad (3.31)$$

Before stating the convergence rates results, we present a useful lemma, which gives a representation for the error $x^\varepsilon(t) - x^\dagger$. This result certainly looks familiar to those who are acquainted with representation theory for the solutions of ordinary differential equations (cf., e.g., [46]).

Lemma 3.5.7. *Let $x^\varepsilon(t)$ be the solution of (3.9) with $y = y^\varepsilon$ and let $x^\dagger \in D(F)$ be the (unique) solution of (3.1) with minimal distance to x_0. Then we have*

$$\begin{aligned} x^\varepsilon(t) - x^\dagger \ &= \ e^{-F'^* F' T}(x_0 - x^\dagger) \\ &+ \int_0^t e^{-F'^* F'(t-s)} F'^*(y^\varepsilon - y) \, ds + \int_0^t e^{-F'^* F'(t-s)} w(s) \, ds \, , \end{aligned}$$

where $F' := F'(x^\dagger)$, $e^{-F'^* F' T} = I + \sum\limits_{k=1}^{\infty} (-1)^k T^k (F'^* F')^k / k!$ and

$$w(s) := F'^* F'(x^\varepsilon(s) - x^\dagger) - F'(x^\varepsilon(s))^* \big(F(x^\varepsilon(s)) - y^\varepsilon\big) + F'^*(y - y^\varepsilon).$$

Now we are ready to present the main result of this section. For simplicity of the presentation we assume (without loss of generality) that $\|F'(x^\dagger)\| \leq 1$.

Lemma 3.5.8. *Let (3.16) be fulfilled with $\eta < 1$, and assume that (3.29), (3.30) and (3.31) hold true. Further, assume that $\|F(x_0) - y^\varepsilon\| > \tau\varepsilon > 0$, with $\tau > (1+\eta)/(1-\eta)$, and $B_\rho(x_0) \in \operatorname{int}(D(F))$. Moreover, let $x^\dagger$ be the (unique) solution of (3.1) with minimal distance to x_0 and let $x^\varepsilon(t)$, $0 \le t \le T_\varepsilon$, be the solution of (3.9) with $y = y^\varepsilon$, where T_ε is defined by the discrepancy principle as in Lemma 3.5.4. If $\tau > (2-\eta)/(1-\eta)$ and $\|w\|$ is sufficiently small, then*

$$\|x^\varepsilon(T_\varepsilon) - x^\dagger\| \ \le \ c\|w\|^{1/(2\nu+1)}\varepsilon^{2\nu/(2\nu+1)} \,,$$

where the constant $c > 0$ is independent of $\|w\|$ and ε.

The proof of Lemma 3.5.8 is rather long and technical. Several auxiliary estimates are required as preparation for the main proof. Although this lemma is the analogon of Lemma 3.3.6 for the nonlinear Landweber iteration, there are substantial differences in the technicalities of both proofs. We refer the reader to [90] for a detailed proof of Lemma 3.5.8.

3.6 Bibliographical comments

Most of the classical results discussed in Section 3.2 can be found in [23, 31]. For details on the Landweber iteration for linear equations the reader can also consult [32, 34, 44, 53, 57, 65, 71, 72]. The Landweber iteration for nonlinear operators is considered in [23, 36]. The results reported in Section 3.4 can be found in [85]. Related results can also be found in [3, 4]. What concerns the asymptotical regularization method, results related to the linear theory can be found in [92, 23]. Results related to the nonlinear theory can be found in [89, 90].

3.7 Exercises

3.1. Let $F : X \to Y$ be a linear compact operator. Given $y \notin D(F^\dagger)$, let $\{x_k\}$ be the sequence defined in (3.10). Prove that every subsequence $\{x_{k_j}\}$ has no weak convergent subsequence.

3.2. Complete the details of the proof of Lemma 3.2.3.

3.3. Obtain an analytical expression for the positive constant c in the proof of Lemma 3.2.4. Conclude that c does not depend on ε.

3.4. Prove that 1 belongs to the spectrum of the operator $\psi(x) = x - F'(x)^*(F(x) - y)$ if F is compact and twice continuous Fréchet differentiable, and X is infinite dimensional. A fixed point of ϕ is any $x_* \in X$ with $F(x_*) = y$.

3.5. Prove the inequality (3.20).

3.6. Prove that, for $\bar{k} \in \mathbb{N}$ fixed, then $x^\varepsilon_{\bar{k}}$ depends continuously on y^ε. Here $\{x^\varepsilon_{\bar{k}}\}$ is the sequence obtained by the nonlinear Landweber iteration.

3.7. Prove that (3.16) with $\tilde{x} = x_*$ can be obtained from (3.6) and (3.7). Hint: First prove that

$$\|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\| \le C\|x - \tilde{x}\|\|F(x) - F(\tilde{x})\| \,.$$

Then use the identity:

$$\|F(x) - F(x_*) - F'(x)(x - x_*)\| = \left\| \int_0^1 (F'(z_t) - F'(x))(x - x_*)\, dt \right\| ,$$

where $z_t = tx + (1-t)x_*$, $t \in [0, 1]$.

3.8. Let $l, k \in \mathbb{N}_0$ with $l < k$ and $\{\alpha_k\}$ be chosen as in Section 3.4. Prove that

$$1 - \prod_{s=l}^{k}(1 - \alpha_s) \;=\; \prod_{j=l}^{k} \alpha_j \prod_{s=j+1}^{k} (1 - \alpha_s) \;\le\; 1\,.$$

Moreover, if $\sum_{k=0}^{\infty} \alpha_k < \infty$, then $\prod_{k=0}^{\infty}(1 - \alpha_k)$ is convergent and thus $\lim\limits_{l \to \infty} \prod_{k=l}^{\infty}(1 - \alpha_k) = 1$.

Chapter 4

Some inverse problems of convolution type

Convolution is both a mathematical concept and an important tool in data processing, in particular in digital signal and image processing. Correlation is a technique that is very similar in mechanism to convolution. Deconvolution is the inverse problem to convolution. In this chapter we discuss several variants of deconvolution.

4.1 Deconvolution

4.1.1 Introduction

Given a blurred photograph, or the result of passing a signal through a medium which acts as a filter, how can we reconstruct an unblurred version of the photograph, or the original signal before the filtering occurred?

Consider a plane image characterized by its intensity distribution I, corresponding to the observation of a “real image ” O through an optical system. If the imaging system is linear and shift-invariant, the relation between the data and the image is a convolution

$$I(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(u-u',v-v')O(u,v)dudv =: (P*O)(u,v) \tag{4.1}$$

where P is the *point spread function/psf* of the imaging system and the operation $*$ is called the *convolution-operator.* The point spread function is also called a *kernel function* or an *impulse response* since, in a very informal consideration, the output signal I is P if the input signal O is a Dirac distribution. Usually, a psf is nonnegative and its integral equals 1; this refers to conservation of energy in the imaging process. In a discrete consideration an image I is given by its pixel function. Convolution is a process in which each pixel is averaged with its neighbors using the kernel P as a multiplier to determine each of the neighbors contribution or weight.

Convolution and related operations are found in many applications: in statistics (moving average, correlation), in optics (blurring, atmospheric degradation), in acoustics (echo), in engineering (input–output mapping), in physics (superposition principle).

Consider the equation (4.1). The **forward problem** in the "**state space**" is:

Given O and P, find I .

Our interest is in the **inverse problem**:

Given I and P, find O .

Unfortunately, in practice I is perturbated by noise. Then the mathematical formulation of (4.1) is

$$\begin{aligned} I(u,v) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(u-u',v-v')O(u,v)dudv + N(u,v) \\ &= (P*O)(u,v) + N(u,v) \end{aligned}$$

where N is an additive noise.

This is the formulation of the problems in the state space. A very efficient tool in analyzing convolution problems is the *Fourier* *transform*; see Subsection 4.2.1. Due to the convolution theorem we obtain in the "*Fourier space/frequency space*" from (4.1)

$$\mathcal{F}(I) = \sqrt{2\pi}\mathcal{F}(P)\mathcal{F}(O)$$

where $\mathcal{F}$ is the Fourier transform. In order to find I, i. e. to solve the forward problem, we have just to multiply the Fourier transform of P and O and to apply the inverse transform $\mathcal{F}^{-1}$. The solution of the inverse problem (with noise) in the Fourier space can be obtained by computing the Fourier transform of the deconvolved object O by a simple division between the image $\mathcal{F}(I)$ and the point spread function $\mathcal{F}(P)$:

$$\mathcal{F}(\tilde{O}) = \sqrt{2\pi}\frac{\mathcal{F}(I)}{\mathcal{F}(P)} = \sqrt{2\pi}\,\mathcal{F}(O) + \sqrt{2\pi}\,\frac{\mathcal{F}(N)}{\mathcal{F}(P)}\,.$$

As we will see in Subsection 4.2.1 $|\mathcal{F}(P)|$ has small values and the error $\mathcal{F}(N)$ is amplified.

The solution of the inverse problem in the state space or in the frequency space is called *deconvolution*. Deconvolution is an ill-posed problem due to the fact that the error in the frequency space is amplified. As a consequence deconvolution becomes a difficult problem especially when noise is present.

Convolution in the context of image processing is two-dimensional. To simplify our considerations we describe in the following mostly the one-dimensional situation. Consider the equation

$$g * x = y \tag{4.2}$$

where

$$(g * x)(t) := \int_{-\infty}^{\infty} g(t-s)x(s)ds, t \in \mathbb{R}\,,$$

is the convolution integral with kernel g . A problem in a finite interval $[0,T]$ results when we consider a signal x which vanishes outside $[0,T]$. Another special case (Volterra-type convolution) is obtained when the kernel g satisfies $g(r) = 0$ if $r < 0$. The problem of differentiation of data considered from various aspects of view in Chapter 2 may formally be considered as a deconvolution problem. The kernel is given as follows:

$$g(r) := \begin{cases} 1 & \text{, if } t \geq s \\ 0 & \text{, if } t < s \end{cases} .$$

Example 4.1.1. *Let us present a very famous integral equation of convolution type, namely Abel's (singular) integral equation. This is one of the first integral equations ever treated.*

In the vertical x-y–plane find a curve $\mathcal{C}$ that is the graph of an increasing function $[0,H] \ni x \longmapsto \psi(x) \in [0,\infty)$ such that the falling time of a particle under the gravity force along this curve is equal to the value a given function τ in every moment. In absence of friction the problem is that of solving the equation

$$\int_0^y (y-z)^{-\frac{1}{2}} u(z)dz = \sqrt{2g}\,\tau(y)\,,\ y \in [0,H]\,, \tag{4.3}$$

where $u(z) = \sqrt{1+\psi'(z)^2}$. With the operator J of integration, given by

$$Jg(x) := \int_0^x g(t)dt\,,\ x > 0\,,$$

we may consider the equation (4.3) *as a special case of the following family of equations:*

$$(J^\alpha u)(x) := \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} u(t)dt = f(x)\,, \, x > 0\,. \tag{4.4}$$

Here $\Gamma(\alpha)$ is the gamma function. When α is a positive integer, J^α is nothing but the α–fold integral of u. These and other properties justify the term fractional integral operator *for J^α if $\alpha \in (0, \frac{1}{2})$, and* **fractional derivative operator** *for the inverse of J^α:*

$$D^\alpha u(x) := \frac{d}{dx} J^{1-\alpha} u(x)\,.$$

□

When the kernel function in equation (4.1) or (4.2) is not known, then "parallel" to the reconstruction of the input signals O and x

respectively one has to find out the kernel function. This problem is called *blind deconvolution.*

Correlation is the close mathematical cousin of convolution. The *correlation integral*

$$c(t) := \int_{-\infty}^{\infty} u(t+s)v(s)ds, t \in \mathbb{R}$$

is used to measure the similarity between two signals u, v: a large value $|c(t)|$ represents a strong similarity between the two signals. The correlation with itself is called the *autocorrelation.* Related to the problem of autocorrelation is the problem of **autoconvolution** where the *autoconvolution integral* is given as

$$a(t) := \int_{-\infty}^{\infty} u(t-s)u(s)ds, t \in \mathbb{R}$$

for a given function $u : \mathbb{R} \longrightarrow \mathbb{R}$. Autocorrelation and blind deconvolution will be considered in sections below.

4.1.2 Stability and Regularization in the state space

Here we consider the convolution equation

$$\int_0^T g(t-s)x(s)ds = y(t)\,,\; t \in [0,T]\,. \tag{4.5}$$

in the state space, i. e. in the space of functions $x : [0,T] \longrightarrow \mathbb{R}$. The assumption that the finite interval $[0,T]$ is the same with respect to t and s is no a serious assumption, it can be achieved in other cases by a simple affine transform.

Related to the equation (4.5) there is an operator equation of the type which is extensively studied in Chapter 2:

$$Ax = y \text{ where } A(x) := \int_0^T g(\cdot - s)x(s)ds\,. \tag{4.6}$$

We conclude from the convolution theorem (see subsection 4.2.1) that A can be considered as continuous mapping from $X := L_2[0,T]$ into $L_2[0,T]$ when the kernel g belongs to $L_1[0,T]$. If the kernel g is not degenerate, then the range of A is not closed and to solve the equation (4.5) is an ill-posed problem.

For sufficiently smooth g and y, we may differentiate equation (4.5) with respect to t to obtain

$$g(t)x(t) + \int_0^T g'(t-s)x(s)ds = y'(t)\,,\ t \in [0,T]\,. \tag{4.7}$$

If $g(t) \neq 0$, for $t \in [0,T]$, division of equation (4.20) by $g(t)$ yields a standard Volterra equation of the second kind which can be solved in a stable way (well-posedness of Volterra equation of the second kind). In particular the solution depends continuously on the right hand side y'. If $g(t) = 0, t \in [0,T]$, we may repeat the process by differentiating the equation once again.

We will say that the equation (4.5) is a *l-smoothing problem* if the kernel g is l-times continuously differentiable and

$$g^{(k)}(t) = 0, t \in [0,T], k = 0,\dots,l-1\,,\ g^{(l)}(t) \neq 0, t \in [0,T]\,.$$

The problem of differentiation of data is a 1-smoothing problem. As a rule, the asymptotic of the singular values of l–smoothing problems increases with increasing l.

We will say that the equation (4.5) is an *infinitely–smoothing problem* if the g is l-times continuously differentiable and

$$g^{(l)}(t) = 0, t \in [0,T], l \in \mathbb{N}_0\,.$$

Of course, not all equations of the form (4.20) fall into precisely one of the above classes of problems. A classic example of an infinitely smoothing problem is the sideways heat equation; here the kernel g is given by

$$g(r) = \frac{1}{2\sqrt{\pi} r^{\frac{3}{2}}}\, e^{-\dfrac{1}{4r}}\,,\ r \in \mathbb{R}\,.$$

The importance of the behavior of the kernel g near $t = 0$ shows up also in the following Theorem which is proved in [82].

Theorem 4.1.2 (Reverse convolution inequality).
Let $\delta > 0, 0 \leq \tau < T$, and let $f, g \in L_\infty(0,T)$ satisfy

$$0 \leq f(t), g(t) \leq M, t \in [0,T]\,. \tag{4.8}$$

Then

a) $\|f\|_{L_2[\tau,T]}\|g\|_{L_2[0,\delta]} \leq M\left(\int_\tau^{T+\delta}\left(\int_\tau^t g(t-s)f(s)ds\right)dt\right)^{\frac{1}{2}}.$

b) $\|f\|_{L_2[0,T]}\|g\|_{L_2[0,\delta]} \leq M\|g*f\|^{\frac{1}{2}}_{L_1[0,T+\delta]}.$

Corollary 4.1.3. *Let $\delta \in (0,T)$ and suppose $g:[0,T] \longrightarrow \mathbb{R}$ with*

$$g \geq 0, g \in H_1[0,T], \|g\|_{L_2[0,T]} \geq \gamma(\delta) > 0 \tag{4.9}$$

Suppose that $x \in H_2[0,T]$ is a solution of equation (4.5) and suppose that x has at most finitely many zeros in $[0,T-\delta]$. Then $y \in H_{2,0}[0,T]$ and there exists $c(\delta) > 0$ with

$$\|x\|_{L_2[0,T-\delta]} \leq 2\delta c(\delta)\|y''\|^{\frac{1}{4}}_{L_2[0,T]}\|y\|^{\frac{1}{4}}_{L_2[0,T]}. \tag{4.10}$$

Proof:
There exists $M \geq 0$ such that $0 \leq |g(x)| \leq M, x \in [0,T]$. Let $0 \leq t_1 < \cdots < t_n \leq T-\delta$ be the zeros of x. Assume $0 \leq x(s) \leq M, s \in [0,t_1]$. Then we have with Theorem 4.1.2

$$\|x\|_{L_2[0,t_1]} \leq c(\delta)^{-1}M\|y\|^{\frac{1}{2}}_{L_2[0,t_1+\delta]} \leq c(\delta)^{-1}M\delta^{\frac{2}{3}}\|y'\|^{\frac{1}{2}}_{L_2[0,t_1+\delta]}.$$

Since

$$\begin{aligned} y(t) &= \int_0^t g(t-s)x(s)ds \\ &= \int_0^{t_1} g(t-s)x(s)ds + \int_{t_1}^t g(t-s)x(s)ds \\ &= y(t_1) - \int_{t_1}^t g(t-s)(-x(s))ds \end{aligned}$$

we obtain with Theorem 4.1.2

$$\begin{aligned} \|x\|_{L_2[t_1,t_2]} &\leq c(\delta)^{-1}M(\int_{t_1}^{t_2+\delta}\int_{t_1}^t g(t-s)(-x(s))ds)dt^{\frac{1}{2}} \\ &= c(\delta)^{-1}M(\int_{t_1}^{t_2+\delta}(-y(t)+y(t_1))dt)^{\frac{1}{2}} \end{aligned}$$

and therefore

$$\|x\|_{L_2[t_1,t_2]} \le c(\delta)^{-1} M \delta^{\frac{2}{3}} \|y'\|_{L_2[t_1,t_2+\delta]} \,.$$

Proceeding in the same way in each interval $[t_{i-1}, t_i]$ and summing up we obtain the bound when we apply in addition the interpolation inequality $\|y'\|_{L_2[0,T]} \le \|y''\|^{\frac{1}{2}}_{L_2[0,T]} \|y\|^{\frac{1}{2}}_{L_2[0,T]}$. ■

The result in Corollary 4.1.3 leads to a stability estimate

$$\|x\|_{L_2[0,T-\delta]} \le c\|Ax\|^{\frac{1}{4}}_{L_2[0,T]} E^{\frac{1}{4}} \tag{4.11}$$

when we have an a-priori bound E for $\|y''\|_{L_2[0,T]}$ which can be derived from a bound for $\|x'\|_{L_2[0,T]}$.

Suppose there are given real valued functions

$$x^0, y^0, y^\varepsilon \in L_2(\mathbb{R}) \text{ with } g * x^0 = y^0, \|y^0 - y^\varepsilon\| \le \varepsilon \,. \tag{4.12}$$

We want to find an approximation x^ε of x^0 using the data y^ε. For the regularization of the equation we can apply all the methods studied in Chapter 2. The application of the method of Tikhonov is straight forward. But this method is not appropriate when the equation (4.5) is of Volterra type since then the operator A^*A comes in and this operator is not of Volterra type.

When the equation (4.5) is l–smoothing we may differentiate the equation l–times to obtain

$$g^{(l)}(t)x(t) + \int_0^T g^{(l-1)}(t-s)x(s) = y^{(l)}(t)\,,\ t \in [0,T]\,.$$

Since the righthand side y^ε may not be differentiable we have to apply the methods which are considered in Chapter 2. A related method is *Lavrentiev's method of singular perturbation.* Consider the equation

$$\alpha x(t) + \int_0^T g(t-s)x(s)ds = y(t)\,,\ t \in [0,T]\,. \tag{4.13}$$

where α is a small "regularization parameter". Equation (4.13) looks like an integral equation of the second kind. Indeed, when the kernel g

is of Volterra type, then the operator $\alpha I + A$ is invertible under certain circumstances and one may prove results similar to the situation $\alpha I + A^*A$. In the general case, A may destroy the invertibility of $\alpha I + A$ since, in contrast to A^*A, A may not be a "positive" operator. "Positivity" of A may be formulated as

$$\langle Au, u\rangle \geq 0, x \in X .$$

Such operators are called *accretive*. Accretive operators include those with kernels g that are positive, decreasing and convex. Even the generalized Abel integral operator can be considered as an accretive operator.

4.1.3 Iterative deconvolution

Consider again equation (4.2)

$$g * x = y \tag{4.14}$$

We restrict ourselves to the one-dimensional case, extensions to the two-dimensional case are immediate.

Assumption A5:

$$g \geq 0, g \in L_1(\mathbb{R}), \|g\|_{L_1(\mathbb{R})} = 1 .$$

Under this assumption we obtain for a pair (x, y) which solves (4.14)

$$\int_{-\infty}^{\infty} y(t)dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t-s)x(s)dsdt = \int_{-\infty}^{\infty} x(t)dt$$

(when g and x, y are appropriate "regular").

Let $A : L_2(\mathbb{R}) \ni x \longmapsto g * x \in L_2(\mathbb{R})$ and let A^* be the adjoint operator. Let y be a positive function. The following steps to derive iterative methods for the solution of (4.14) can be made rigorous.

Van Cittert method
This method starts from the identity

$$x = x + \lambda A^*(y - Ax)$$

and transforms this into the iteration

$$u^{n+1} = u^n + \lambda A^*(y - Au^n)\,,\ u^0 \text{ given } (u^0 \approx y^\varepsilon)\,. \tag{4.15}$$

Here $\lambda > 0$ is a constant that controls convergence. The iteration above may be considered as a method of steepest descent applied to the minimization of $\|Ax - y\|^2_{L_2(\mathbb{R}^2)}$. In the context of operator equations this method is also called the **Landweber method**; see Chapter 3.

Lucy–Richardson method
This method starts from the identities

$$1 = \frac{y}{Ax}\,,\ 1 = A^*(\frac{y}{Ax})\,,\ x = xA^*(\frac{y}{Ax})$$

and transforms these into the iteration

$$u^{n+1} = u^n A^*(\frac{y}{Au^n})\,,\ u^0 \text{ given } (u^0 \approx y^\varepsilon)\,. \tag{4.16}$$

This method is widely used in astronomical imaging.

Poisson MAP method
This method starts from the identities

$$x = x\exp(A^*(\frac{y}{Ax} - 1))$$

and transforms these into the iteration

$$u^{n+1} = u^n \exp(A^*(\frac{y}{Au^n} - 1))\,,\ u^0 \text{ given } (u^0 \approx y^\varepsilon)\,. \tag{4.17}$$

As we see, positivity is preserved during the iteration.

Remark 4.1.4. *The basis of the iteration-method above are "fixed point" identities for x with data g and y; no further assumptions are necessary.*
But when x, g, y are viewed as probability densities, not necessarily normalized, then the a-priori probability $p(x|y)$ and the a-posteriori probability $p(y|x)$ are defined and the fixed point identities may also be based on statistical considerations concerning $p(x|y)$ and $p(x|y)$ respectively. When one assumes that the image y is corrupted by

Poisson noise Bayes's theorem leads to the identity $x = xA^*(\frac{y}{Ax})$ *and then the iteration method* (4.16) *is called the* E-M method*. It is widely used in medical imaging.*
The identity $x = x\exp(A^*(\frac{y}{Ax} - 1))$ *may be based on statistical arguments concerning the probability* $p(x|y)$*.* □

4.1.4 On discretization in the state space

For the regularization of the equation (4.5) by discretization there is a variety of schemes available. Here we restrict ourselves to quadrature methods based on well-known quadrature rules. Recall that a quadrature rule for computing an approximation to an integral on the interval $[0,T]$ takes the following form:

$$\int_0^T \phi(t)dt = \sum_{j=1}^n w_j\phi(s_j)$$

where $s_1,\dots,s_n$ are the abscissas for the particular quadrature rule, and $w_1, \dots, w_n$ are the corresponding weights. Examples are the midpoint rule and the Simpson rule. Using a quadrature rule, we can approximate the integral in our equation (4.5) as follows:

$$\sum_{j=1}^n w_j g(t,s_j)x(s_j) \approx \int_0^T g(t-s)x(s)ds = y(t)\,,\ t\in[0,T]\,. \quad (4.18)$$

In order to obtain a system of linear equations, we can use *collocation* at given points $t_1,\dots,t_n$:

$$\sum_{j=1}^n w_j g(t_i,s_j)x(s_j) = y(t_i)\,,\ i=1,\dots,n\,. \quad (4.19)$$

This system of equations (4.19) is in matrix notation a quadratic system $Au = b$ where

$$A=(a_{ij}), a_{ij} := w_j g(t_i - s_j),\ b_i = y(t_i), u_j = x(s_j)\,,\ i,j=1,\dots,n\,.$$

As already mentioned in Chapter 2, we cannot expect that the solution of the system (4.19) without regularization leads to a meaningful approximation $x(t_1), \dots, x(t_n)$, especially when the system size n is chosen too large.

The key observation here is that for convolution problems the corresponding matrix A with entries $a_{ij} = w_j g(t_i - s_j)$ can be written in the form

$$A = GW\,, \tag{4.20}$$

where $W = \mathrm{diag}(w_1, \dots, w_n)$ is a diagonal matrix consisting of the quadrature weights and the entries of the matrix G are "samples" of g, i. e.

$$g_{ij} = g(t_i - s_j)\,,\ i, j = 1, \dots, n\,.$$

We simplify our consideration by choosing the discretization points s_j identically to the collocation points t_i and identically spaced:

$$s_j = t_j = hj\,,\ j = 0, \dots, n,$$

where $h := T/n$. Then the entries of the matrix G satisfy

$$g_{ij} = g((i-j)h)\,,\ i, j = 1, \dots, n\,.$$

This special structure of the coefficient matrix A can be used to derive very efficient algorithms to solve the related system of equations. A *Toeplitz matrix* $M \in \mathbb{R}^{n,n}$ is a matrix whose elements depend only on the difference $i - j$ between the indices, i. e. M can be written as

$$M = \begin{pmatrix} m_0 & m_{-1} & m_{-2} & \dots & m_{1-n} \\ m_1 & m_0 & m_{-1} & \dots & m_{2-n} \\ m_2 & m_1 & m_0 & \dots & m_{3-n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ m_{n-1} & m_{n-2} & m_{n-3} & \dots & m_0 \end{pmatrix}.$$

Obviously, the matrix A in (4.20) is a Toeplitz matrix. Toeplitz matrices are symmetric across the antidiagonal. Let J denote the matrix

$$J = \begin{pmatrix} & & & 1 \\ & & \cdot & \\ & \cdot & & \\ 1 & & & \end{pmatrix}.$$

Then for each Toeplitz matrix M one has the identities

$$M = JM^tJ\,,\ (MJ)^t = MJ \text{ and } M^{-1} = J(M^{-1})^tJ,$$

when M is invertible. These identities can be used to derive properties of the singular value decomposition of a Toeplitz matrix: the left and right singular vectors are related by the fact that the entries are identical except perhaps for a sign change.

A system $Ax = b$ with a Toeplitz matrix A can be solved iteratively in a very efficient way since a matrix-vector multiplication can be performed in $O(n\log_2 n)$ flops. The key idea is to embed the $n\times n$ Toeplitz matrix M in a larger $p\times p$ circulant matrix C and to use the fast Fourier transform to perform the matrix-vector multiplication with C.

4.2 Convolution in Fourier space

4.2.1 Some results for the Fourier transform

In this section we mention the necessary information from the theory of Fourier transforms; see [17].

Let $f:\mathbb{R}^n \longrightarrow \mathbb{C}$ be a function in $L_1(\mathbb{R}^n)$, the space of Lebesgue–integrable functions. We define $f^\wedge$ as follows:

$$f^\wedge(w) := (2\pi)^{-\frac{n}{2}}\int\limits_{\mathbb{R}^n} f(t)\exp(\langle -iw,x\rangle dx\,,\ w\in\mathbb{R}^n\,; \qquad (4.21)$$

here $\langle\cdot,\cdot\rangle$ denotes the euclidian inner product in $\mathbb{R}^n$. Notice that the integral exists if $f\in L_1(\mathbb{R}^n)$. In the same way we define $f^\vee$ as follows:

$$f^\vee(x) := (2\pi)^{-\frac{n}{2}}\int\limits_{\mathbb{R}^n} f(t)\exp(\langle +iw,x\rangle dw\,,\ x\in\mathbb{R}^n\,. \qquad (4.22)$$

A key result in understanding the ill-posedness of deconvolution in Fourier space is the **Lemma of Riemann-Lebesgue:**

If $f\in L_1(\mathbb{R}^n)$ then $f^\wedge\in C_0(\mathbb{R}^n)$.

Here $C_0(\mathbb{R}^n)$ is the space of uniformly continuous mappings $h : \mathbb{R}^n \longrightarrow \mathbb{C}$ with $\lim_{|t|\to\infty} h(t) = 0$. We conclude from this Lemma that "each" signal in the Fourier domain has small values. Therefore it is dangerous to divide with a signal in the Fourier domain.

We want to consider the transformations $^\wedge, ^\vee$ also on the space $L_2(\mathbb{R}^n)$ (with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$). How this is possible becomes clear from **Plancherel's theorem:**

> There exist uniquely determined bounded linear operators $\mathcal{F}, \mathcal{F}^{-1} : L_2(\mathbb{R}^n) \longrightarrow L_2(\mathbb{R}^n)$ such that the following assertions hold:
>
> (1) $\mathcal{F}(f) = f^\wedge, \mathcal{F}^{-1}(f) = f^\vee$ for all $f \in L_2(\mathbb{R}^n) \cap L_1(\mathbb{R}^n)$.
>
> (2) $\text{range}(\mathcal{F}) = \text{range}(\mathcal{F}^{-1}) = L_2(\mathbb{R}^n)$.
>
> (3) $\mathcal{F}\mathcal{F}^{-1} = \mathcal{F}^{-1}\mathcal{F} = I$
>
> (4) $\langle \mathcal{F}(f), \mathcal{F}(g)\rangle = \langle f, g\rangle$ for all $f, g \in L_2(\mathbb{R}^n)$.
>
> (5) $\|\mathcal{F}(f)\| = \|f\|, \|\mathcal{F}^{-1}(f)\| = \|f\|$ for all $f, g \in L_2(\mathbb{R}^n)$.

In the following $\mathcal{F}$ is called the *Fourier transform* and $\mathcal{F}^{-1}$ the **inverse Fourier transform**. It is useful to consider the transformation $\mathcal{F}$ as a transformation of the "state space with time t" into the "frequency space of spectral values ω".

The *convolution* $g * f$ of functions g, f on $\mathbb{R}^n$ is defined in a formal way by

$$(g * f)(t) := \int_{\mathbb{R}^n} g(t-s) f(s) ds \ , t \in \mathbb{R}^n .$$

A key result is the **convolution theorem:**

> Let $g \in L_1(\mathbb{R}^n)$ and $f \in L_2(\mathbb{R}^n)$. Then $g * f \in L_2(\mathbb{R}^n), \|g * f\|_2 \le \|g\|_1 \|f\|_2$ and $\mathcal{F}(g * f) = (2\pi)^{\frac{n}{2}} \mathcal{F}(g)\mathcal{F}(f)$.

Convolution operation has the same effect as a frequency filter, in that it enhances some frequencies in an image and suppresses others. Notice that the operator $A : L_2(\mathbb{R}^n) \longrightarrow L_2(\mathbb{R}^n), x \longmapsto g * x$, associated with the kernel $g \in L_1(\mathbb{R}^n)$ is not compact in general.

But the Fourier transform plays a similar role as the singular value decomposition for compact operators.

Let us consider the Fourier transform of special functions which are of some interest in the following.

Example 4.2.1.

1) The perfect lowpass filter*. We have the pair*

$$\begin{aligned} g_1(t) &:= \frac{\Omega}{\sqrt{2\pi}} \frac{\sin(\frac{1}{2}\Omega t)}{\frac{1}{2}\Omega t}, t \in \mathbb{R}\,, \\ h_1(\omega) &:= \mathcal{F}(g_1)(\omega) = \left\{ \begin{array}{ll} 1 & , |\omega| \le \Omega \\ 0 & , |\omega| > \Omega \end{array} \right. \end{aligned}$$

2) The triangle window*. We have the pair*

$$\begin{aligned} g_2(t) &:= \frac{1}{\sqrt{2\pi}} \left\{ \frac{\sin(\frac{1}{2}\Omega t)}{\frac{1}{2}\Omega t} \right\}^2, t \in \mathbb{R}\,, \\ h_2(\omega) &:= \mathcal{F}(h_2)(\omega) \;=\; \left\{ \begin{array}{ll} \frac{1}{\Omega}(1 - \frac{|\omega|}{\Omega}), & |\omega| \le \Omega \\ 0 & |\omega| > \Omega \end{array} \right. \end{aligned}$$

which becomes clear by the convolution theorem using the fact

$$g_2 = \sqrt{2\pi}\frac{1}{\Omega^2} g_1^2\,.$$

3) The Gaussian filter*. We have the pair*

$$g(t) := \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{t^2}{4\sigma^2}}, t \in \mathbb{R}\,,\; h(\omega) := \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{\omega^2}{4\sigma^2}}, \omega \in \mathbb{R}.$$

4) The Lorentzian filter*. We have the pair*

$$g_L(t) := \frac{a}{t^2 + a^2}, t \in \mathbb{R}\,,\; h_L(\omega) := \mathcal{F}(h)(\omega) = \sqrt{\frac{\pi}{2}} e^{-a\omega}, \omega \in \mathbb{R}.$$

4.2.2 Stability and regularization in the Fourier space

Consider a one-dimensional convolution equation:

$$g * x = y \tag{4.23}$$

where the kernel g is a given real valued function in $L_1(\mathbb{R})$. As we know from the convolution theorem, the convolution equation (4.23) is equivalent to the "algebraic" equation

$$\sqrt{2\pi}\mathcal{F}(g)\mathcal{F}(x) = \mathcal{F}(y).$$

If a solution x of (4.23) exists then $x = \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}(f)$ where $f = \mathcal{F}(y)\mathcal{F}(g)^{-1}$. Since $\lim_{|\omega|\to\infty} |\mathcal{F}(g)(\omega)| = 0$ due to the lemma of Riemann-Lebesgue the equation (4.23) is ill-posed (lack of stability):

> A small perturbation η in y whose transform $\mathcal{F}(\eta)$ does not decay faster than $\mathcal{F}(g)$ as $|\omega| \to \infty$ will result in a perturbation in $\mathcal{F}(y)\mathcal{F}(g)^{-1}$ which will grow without bound.

Therefore it is not reasonable to define the reconstruction x^ε of x^0 as

$$x^\varepsilon := \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}(f) \text{ where } f = \mathcal{F}(y^\varepsilon)\mathcal{F}(g)^{-1}.$$

We have to regularize the equation in order to solve it in a stable way. We do this by using a filter function (window function) h in the following way:

$$x_h = \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}(f_h) \text{ where } f_h = h\mathcal{F}(y)\mathcal{F}(g)^{-1}.$$

The choice of the *filter function* h has to ensure that $f_h = h\mathcal{F}(y)\mathcal{F}(g)^{-1}$ doesn't blow up for $|\omega| \to \infty$. In general, the filter function depends on a parameter which has to be chosen properly. If we consider the filter function $h = h_1$ (see example 4.2.1) $\mathcal{F}(x_h)$ has the truncated frequency spectrum

$$\mathcal{F}(x_h)(\omega) = \begin{cases} \mathcal{F}(x)(\omega) & , \quad |\omega| \le \Omega \\ 0 & , \quad |\omega| > \Omega \end{cases}$$

Here we consider the following family $(x^{\varepsilon,\alpha})_{\alpha>0}$ of regularized solutions of the reconstruction problem:

$$x^{\varepsilon,\alpha} := \frac{1}{\sqrt{2\pi}}\,\mathcal{F}^{-1}(f^{\varepsilon,\alpha})$$

where

$$f^{\varepsilon,\alpha}(\omega) := \frac{|\mathcal{F}(g)(\omega)|^2}{|\mathcal{F}(g)(\omega)|^2+\alpha(1+\omega^2)} \cdot \frac{\mathcal{F}(y^\varepsilon)(\omega)}{\mathcal{F}(g)(\omega)}, \omega \in \mathbb{R}.$$

This regularized solution $x^{\varepsilon,\alpha}$ may be defined equivalently as the minimizer of the Tikhonov-functional

$$x \longmapsto \|g * x - y^\varepsilon\|^2 + \alpha\|x\|_1^2$$

where $\|\ \|_1$ is the norm in the Sobolev space

$$H_1(\mathbb{R}) := \{x \in L_2(\mathbb{R}) | x \text{ is absolutely continuous, } x' \in L_2(\mathbb{R})\}\,.$$

In the following we shall give a bound for the reconstruction error $x^{\varepsilon,\alpha} - x^0$. By Plancherel's theorem it is sufficient to estimate $f^{\varepsilon,\alpha} - f^0$ where $f^0 = \mathcal{F}(x^0)$. Let

$$z(\omega,\alpha) := \frac{|\mathcal{F}(g)(\omega)|^2}{|\mathcal{F}(g)(\omega)|^2+\alpha(1+\omega^2)}\,,\ v(\omega;\alpha) := \frac{1+\omega^2}{|\mathcal{F}(g)(\omega)|^2+\alpha(1+\omega^2)},$$

and

$$f^{0,\alpha}(\omega) := z(\omega,\alpha)\frac{\mathcal{F}(y^0)(\omega)}{\mathcal{F}(g)(\omega)} = z(\omega;\alpha)\mathcal{F}(x^0)(\omega), \omega \in \mathbb{R}.$$

We estimate $f^{\varepsilon,\alpha} - f^0$ by giving bounds for each term on the righthand side of

$$\|f^{\varepsilon,\alpha} - f^0\| \le \|f^{\varepsilon,\alpha} - f^{0,\alpha}\| + \|f^{0,\alpha} - f^0\|.$$

Notice that the functions $\omega \longmapsto |\mathcal{F}(x^0)(\omega)|$, $\omega \longmapsto |\mathcal{F}(h)(\omega)|$ are even functions since x^0 and h are real valued.

Lemma 4.2.2. *We have*

$$\|f^{\varepsilon,\alpha} - f^{0,\alpha}\| \le \frac{\varepsilon}{\sqrt{\alpha}}.$$

Proof:

$$\begin{aligned}\|f^{\varepsilon,\alpha} - f^{0,\alpha}\|^2 &= \int\limits_{-\infty}^{\infty} z(\omega;\alpha)^2 \frac{1}{|\mathcal{F}(h)(\omega)|^2} |\mathcal{F}(y^\varepsilon)(\omega) - \mathcal{F}(y^0)(\omega)|^2 d\omega \\ &\le \frac{1}{\alpha} \int\limits_{-\infty}^{\infty} |\mathcal{F}(y^\varepsilon)(\omega) - \mathcal{F}(y^0)(\omega)|^2 ds \\ &= \frac{1}{\alpha} \int\limits_{-\infty}^{\infty} |y^\varepsilon(t) - y^0(t)|^2 dt \le \frac{\varepsilon^2}{\alpha}\,.\end{aligned}$$

■

As we already know, for a parameter choice strategy one needs two ingredients:

- an a-priori information concerning the “smoothness” of x^0;
- an information concerning the degree of ill-posedness of the equation (4.23).

Here we present just one possible combination of these ingredients.

Assumption A6:

$$\exists\, q > \frac{1}{2}\ \exists\, d > 0\ (|\mathcal{F}(x^0)(\omega)| \le d(1+\omega^2)^{-\frac{q}{2}}),\ \omega \in \mathbb{R}. \quad (4.24)$$

$$\exists\, a > 0\ \exists\, c > 0\ (|\mathcal{F}(g)(\omega)| \ge c_0 \exp(-a|\omega|)),\ \omega \in \mathbb{R}. (4.25)$$

A kernel which satisfies the assumption (4.25) in **A5** is given by the Lorentzian filter (see Example 4.2.1).

Theorem 4.2.3. *Let the assumption* **A6** *hold. Then*

$$\|x^{\varepsilon,\alpha(\varepsilon)} - x^0\| \le c(\ln\frac{1}{\varepsilon})^{-q+\frac{1}{2}}$$

where $\alpha(\varepsilon) = \varepsilon^2 \ln(\frac{1}{\varepsilon^2})^{-q+\frac{1}{2}}$ *and* c *is a constant independent of* q *and* ε.

Proof:
Let $M > 0$. We have

$$\begin{aligned}\frac{1}{2}\|f^{0,\alpha} - f^0\|^2 &= \alpha^2 \int_0^M v(\omega;\alpha)^2 |\mathcal{F}(x^0)(\omega)|^2 d\omega \\ &\quad +\alpha^2 \int_M^\infty v(\omega,\alpha)^2 |\mathcal{F}(x^0)(\omega)|^2 d\omega \\ &=: J_1(M) + J_2(M).\end{aligned}$$

From assumption **A6** we obtain immediately

$$J_2(M) \leq c\alpha^2 \varepsilon^{3aM}$$

for some constant c independent of α, M.

$$\begin{aligned}J_2(M) &\leq \alpha^2 \int_M^\infty \frac{(1+\omega^2)^2}{\alpha^2(1+\omega^2)^2} |\mathcal{F}(x^0)(\omega)|^2 d\omega \\ &\leq \int_M^\infty d^2(1+\omega^2)^{-q} d\omega \;\leq\; d^2 \int_M^\infty \omega^{-2q} d\omega\,.\end{aligned}$$

This implies

$$J_2(M) \leq c_2 M^{-2q+1}$$

and we have

$$\|f^{0,\alpha} - f^0\|^2 \leq c(M^{-2q+1} + \alpha^2 e^{3aM}) \tag{4.26}$$

where c is a constant independent of q and α. With the choice

$$M = \frac{2}{3a}\ln(\frac{1}{\alpha}) - \frac{2q-1}{3a}\ln(\frac{2}{3a}\ln(\frac{1}{\alpha}))$$

(as an approximation for the minimizer of the right-hand side in (4.26) with respect to M) we obtain

$$\|f^{0,\alpha} - f^0\| \leq c(\ln(\frac{1}{\alpha}))^{-q+\frac{1}{2}} \text{ and } \|x^{\varepsilon,\alpha} - x^0\| \leq c(\frac{\varepsilon}{\sqrt{\alpha}} + \ln(\frac{1}{\varepsilon})^{-q+\frac{1}{2}})$$

and the choice

$$\alpha(\varepsilon) := \varepsilon^2 (\ln(\frac{1}{\varepsilon}))^{2q-1}$$

leads to

$$\|x^{\varepsilon,\alpha} - x^0\| \le c(\ln(\frac{1}{\varepsilon}))^{-q+\frac{1}{2}}$$

where c is a constant independent of q and ε. ■

After discretization of a convolution equation

$$g * x = y\,. \tag{4.27}$$

within the space of trigonometric polynomials – with or without regularization – one is lead to the following problem:

Given $\omega(N) := \exp(-i\frac{2\pi}{N})$ and $z_0, \dots, z_{N-1}$.
Compute $Z_k := \sum\limits_{n=0}^{N-1} z_n \omega(N)^{kn}, 0 \le k \le N-1\,.$

The transformation $(z_0, \dots, z_{N-1}) \longmapsto (Z_0, \dots, Z_{N-1})$ is called the *discrete Fourier transform/DFT* . As it is easily seen, the computation of $Z_0, \dots, Z_{N-1}$ requires a number of arithmetic operations which is proportional to N^2. The fast Fourier transform (FFT) is a method which reduces the number of arithmetic operations to $N \log_2 N$ by using the following observation: A DFT of order N can be evaluated from two DFT of order $N/2$.
Let us give a short sketch of this method. Let $N = 2^\tau, \tau \ge 2$. We have with $N_1 := N/2$

$$\begin{aligned} Z_k &= \sum_{n=0}^{N-1} z_n \omega(N)^{kn} \\ &= \sum_{r=0}^{N-1} z_{2r}(N)^{2rk} + \sum_{r=0}^{N-1} z_{2r+1}\omega(N)^{(2r+1)k} \\ &= \sum_{r=0}^{N-1} z_{2r}\omega(N_1)^{rk} + \omega(N)^k \sum_{r=0}^{N-1} z_{2r+1}\omega(N_1)^{rk}\,, \end{aligned}$$

and therefore

$$Z_k = U_k + \omega(N)^k V_k, Z_{k+N-1} = U_k - \omega(N)^k, 0 \le k \le N_1 - 1 \tag{4.28}$$

where

$$U_k = \sum_{r=0}^{N-1} z_{2r}\omega(N_1)^{rk}, V_k = \sum_{r=0}^{N-1} z_{2r+1}\omega(N_1)^{rk}, 0 \le k \le N_1 - 1\,. \tag{4.29}$$

Obviously, $U_0, \ldots, U_{N_1-1}$ and $V_0, \ldots, V_{N_1-1}$ are discrete Fourier transforms of $z_0, z_2, \ldots, z_{2(N_1-1)}$ and $z_1, \ldots, z_{N-1}$ respectively. If we apply the same procedure to these two DFT of order $N_1 = N/2$ we have to compute four DFT of order $N_2 := N/4$. This decomposition process has $\tau = \log_2 N$ stages. Since each stage requires $\frac{1}{2}N$ complex multiplications and N complex additions the number of arithmetic operations needed to compute the DFT of the data $z_0, \ldots, z_{N-1}$ is proportional to $N \log_2 N$.

4.3 Autoconvolution and autocorrelation

Consider the equation

$$\int_{-\infty}^{\infty} p(t+s)p(s)ds = a(t)\,,\ -\infty < t < \infty\,. \tag{4.30}$$

This equation appears for example in probability theory in the following way. Let Z_1, Z_2 be two identically independent continuous random variables with density function p. Then the right-hand side describes the density function of the random variable $Z_1 - Z_2$. This is the forward problem connected with (4.30). The inverse problem is to find from the autocorrelation function a the density function p. As a necessary condition for the existence of a solution we have $a(t) = a(-t)$, for all $t \in \mathbb{R}$. Then, using the convolution theorem, we obtain

$$\begin{aligned} \mathcal{F}(a)(\omega) &= \sqrt{2\pi}\mathcal{F}(p)(-\omega)\,\mathcal{F}(p)(\omega) \\ &= \sqrt{2\pi}\,\overline{\mathcal{F}(p)(\omega)}\mathcal{F}(p)(\omega) \\ &= \sqrt{2\pi}|\mathcal{F}(p)(\omega)|^2\,,\ -\infty < \omega < \infty\,. \end{aligned}$$

This shows

$$|\mathcal{F}(p)(\omega)|^2 = f(\omega) \text{ where } f(\omega) = 2\int_0^{\infty} a(t)\cos(\omega t)dt\,,\ \omega \in \mathbb{R}\,. \tag{4.31}$$

Now, we conclude that every function F with

$$\mathcal{F}(\omega) := \sqrt{f(\omega)}e^{i\phi(\omega)} \text{ where } |\mathcal{F}(p)(\omega)|^2 = f(\omega)\,,\, \omega \in \mathbb{R}\,,$$

leads to a solution of equation (4.30). Hence the autocorrelation equation has infinitely many solutions.

An important problem in spectroscopy (measurement of laser pulses) consists in the solution of the equation

$$\int_{-\infty}^{\infty} p(s)p(s+t)p(s+\tau)ds = h(t,\tau)\,,\, -\infty < t,\tau < \infty\,, \qquad (4.32)$$

under the assumption $p \in L_1(\mathbb{R})$. This equation is called the *triple correlation equation.* It can be studied again by using the Fourier transform. Let P be the Fourier transform of p and let

$$P(\omega) = |P(\omega)|e^{i\phi(\omega)}\,,\ -\infty < \omega < \infty\,,$$

with the phase ϕ. Along the considerations above we obtain

$$|P(\omega)||P(\zeta)||P(\omega+\zeta)| = |f(\omega,\zeta)|\,,\ \omega,\zeta \in \mathbb{R}, \qquad (4.33)$$

$$\phi(\omega+\zeta) = \phi(\omega) + \phi(\zeta) - \gamma(\omega,\zeta)\,,\ \omega,\zeta \in \mathbb{R}, \quad (4.34)$$

where f is the two-dimensional Fourier transform of h and γ is the phase of f : $f(\omega,\zeta) = |f(\omega,\zeta)|e^{i\gamma(\omega,\zeta)}, \omega,\zeta \in \mathbb{R}$. We assume $f(0,0) > 0$. Then a necessary condition for solvability of (4.32) is given by

$$P(0)^3 = f(0,0)\,,\, |P(\omega)| = \frac{\sqrt{f(\omega,0)}}{f(0,0)^{\frac{1}{6}}}\,,\, \omega \in \mathbb{R}\,.$$

Consider equation (4.35) for $\zeta := \omega$

$$\phi(2\omega) = 2\phi(\omega) - \gamma(\omega,\omega)\,,\, \omega \geq 0\,,$$

when $f(0,0) \neq 0$, which is the generic case. The homogeneous part of this equation is solved by the family $\omega \longmapsto \alpha\omega, \alpha \in \mathbb{R}$, a special solution of the inhomogeneous equation is given by

$$\hat{\phi}(\omega) = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}}\gamma(2^n\omega, 2^n\omega)\,,\, \omega \in \mathbb{R}\,.$$

Thus, we obtain as a general solution of (4.32) in the frequency space

$$P(\omega) = e^{i\alpha\omega} \frac{\sqrt{f(\omega,0)}}{f(0,0)^{\frac{1}{6}}}\, e^{i\hat{\phi}(\omega)}\,,\ \omega \in \mathbb{R}\,;\ \alpha \in \mathbb{R}\,. \tag{4.35}$$

Next, we consider autoconvolution in a finite interval.

$$\int_0^t x(t-s)x(s)ds = y(t)\,,\ t \in [0,T]\,. \tag{4.36}$$

Let $Fx := x * x$ where $(x*x)(t) := \int_0^t x(t-s)x(s)ds\,,\ t \in [0,T]\,$. The mapping F is well-defined in $L_2[0,1]$ and we have

$$\|Fx\|_{L_2[0,T]} \le \|x\|^2_{L_2[0,T]}\,,\ x \in L_2[0,T]\,. \tag{4.37}$$

Moreover, since Fx is continuous, we have $Fx(0) = 0\,$. Therefore we should consider the equation in the scale of Hilbert spaces

$$X_s := H_s[0,T]\,,\ Y_s := H_{s,0}[0,T]\,,\ s \ge 0\,. \tag{4.38}$$

It is easy to verify

Theorem 4.3.1. *Let $n \ge 1\,$. Suppose that the right-hand side y in* (4.36) *satisfies*

$$y \in H_{2n}[0,T]\,,\ y(0) = y'(0) = \cdots = y^{(2n-2)}(0) = 0\,. \tag{4.39}$$

Then we have for a solution x of (4.36) *$x \in H_n[0,T]$ and*

$$x(0) = \cdots = x^{(n-2)}(0) \quad = \quad 0 \ \ (\textit{if}\ n \ge 2)\,, \tag{4.40}$$

$$x^{(n-1)}(0)^2 \quad = \quad y^{(2n-1)}(0)\,, \tag{4.41}$$

$$2x^{(n-1)}(0)x^{(n)} + x^{(n)} * x^{(n)} \quad = \quad y^{(2n)}\,. \tag{4.42}$$

Consider the equation (4.36) under the assumption of Lemma 4.3.1. When $y^{(2n-1)}(0) < 0$ then equation (4.36) has no (real-valued) solution, since (4.42) can have no solution. When $y^{(2n-1)}(0) > 0$ then equation (4.36) two solutions in $H_n[0,T]\,$. They are related to the cases

$$x^{(n-1)}(0) = +\sqrt{y^{(2n-1)}(0)}\,,\ x^{(n-1)}(0) = -\sqrt{y^{(2n-1)}(0)}\,.$$

Suppose $y^{(2n-1)}(0) > 0$. The equation (4.42) is a well-posed Volterra equation of the second kind. Therefore, under the assumption of Lemma 4.3.1, the solution of equation (4.36) is reduced to the problem to differentiate the right-hand side y n-times when $y^{(2n-1)}(0) > 0$. Here we can use known methods; see Chapter 2.

4.4 Blind deconvolution

Blind deconvolution is the identification of a point spread function and an input signal from observation of their convolution. The stable solution of this problem is of interest in many practical areas in signal and image processing.

Consider again the equation (4.23)

$$g * x = y \tag{4.43}$$

where the kernel is assumed to be in $L_1(\mathbb{R})$. Since we know y only we cannot identify g and x when g can be decomposed as $g = g_1 * g_2$. Thus, we see that nonuniqueness is compound with discontinuous dependence on data due to ill-posedness in the deconvolution problem.

When $g = x^-$ where $x^-(t) = x(-t), t \in \mathbb{R}$, then blind deconvolution in equation (4.43) is equivalent to the problem of recovering the image x from the modulus of its Fourier transform.

In practice, all blind deconvolution algorithms require some partial information to be known and some conditions to be satisfied. We require that the true image x and the point spread function g to be nonnegative. These and possible other a-priori conditions are incorporated in descriptive sets $\mathcal{K}, \mathcal{X}$. Then an iterative blind deconvolution method consists of the following steps:

```
INPUT         Start with an initial estimate $u^0 \in \mathcal{X}$.
               Set $k := 0$.

New kernel:   Form a new estimate $G^k$ from $\mathcal{F}(u^k)$, $\mathcal{F}(y)$.
              Compute $\tilde{g}^k := \mathcal{F}^{-1}(G^k)$.
              Realize an approximation $g^k \in \mathcal{K}$ for $\tilde{g}^k$.

New image:    Form a new estimate $U^k$ from $\mathcal{F}(g^k), \mathcal{F}(y)$.
              Compute $\tilde{u}^{k+1} := \mathcal{F}^{-1}(U^k)$.
              Realize an approximation$u^{k+1} \in \mathcal{X}$for $\tilde{u}^{k+1}$.

Update:       Set $k := k + 1$ and go to ``New kernel''.

OUTPUT        Sequences $(u^k)_{k\in\mathbb{N}}, (g^k)_{k\in\mathbb{N}}$.
```

Clearly, there are serious problems to handle since G^k and U^k are to be found by division of $\mathcal{F}(y)$ by $\mathcal{F}(u^k)$ and $\mathcal{F}(g^k)$ respectively. Moreover, the realization of the approximations depends heavily on properties of the sets $\mathcal{K}, \mathcal{X}$.

A *Lévy-distribution* $l_{\alpha,\beta}$ is a function on $\mathbb{R}^2$ which has a Fourier transform depending on the parameter α, β as follows:

$$\mathcal{F}(l_{\alpha,\beta})(\omega,\zeta) = \exp(-\alpha(\omega^2+\zeta^2)^\beta), \omega, \zeta \in \mathbb{R}\,.$$

Set

$$\mathcal{K} = \{P : \mathbb{R}^2 \longrightarrow \mathbb{R} \mid \mathcal{F}(P) = \sum_{i=1}^{k} \mathcal{F}(l_{\alpha_i,\beta_i}), \alpha_i \geq 0, 0 < \beta_i \leq 1, l \in \mathbb{N}\}\,.$$

The Gaussian case $(k = 1, \beta = 1)$ and the Lorentzian case $(k = 1, \beta = \frac{1}{2})$ are included.

Consider the Gaussian case. The blurred image y may be viewed as the solution of the heat equation with an appropriate constant diffusion coefficient $\lambda > 0$ at time $t = 1$. The desired blurred image

x is the initial data in this heat flow problem:

$$u_t = \lambda\Delta u, 0 < t \le 1, u(1) = y\,.$$

For pdf's in $\mathcal{K}$ the heat equation becomes an evolution equation with a pseudo–differential operator given by fractional powers of the Laplacian:

$$u_t = -\sum_{i=1}^{k} \lambda_i(-\Delta)^{\beta_i} u, 0 < t \le 1, u(1) = y\,,$$

with $\lambda_i = \alpha_i(4\pi^2)^{-\beta_i}$. If we would know the parameter α_i, β_i, k, then we could apply methods which are designed to solve evolution equations backward in time. Here is an idea to find the parameter α, β in the pure Lévy-case. We have

$$\mathcal{F}(g)(\omega,\zeta)\mathcal{F}(x)(\omega,\zeta) = \sqrt{2\pi}\mathcal{F}(y)(\omega,\zeta), \omega,\zeta \in \mathbb{R},$$

and therefore

$$\begin{aligned} \exp(-\alpha(\omega^2+\zeta^2)^\beta)|\mathcal{F}(x)(\omega,\zeta)| &= \sqrt{2\pi}|\mathcal{F}(y)(\omega,\zeta)|, \omega,\zeta \in \mathbb{R} \\ -\alpha|\omega|^{2\beta} + \ln(|\mathcal{F}(x)(\omega,0)|) &= \ln(\sqrt{2\pi}|\mathcal{F}(y)(\omega,0)|), \omega \in \mathbb{R}. \end{aligned}$$

We make the ansatz

$$\ln(|\mathcal{F}(x)(\omega,0)|) = -a|\omega|^b$$

and find a, b such that $\omega \longmapsto -\alpha|\omega|^{2\beta} - a|\omega|^b$ fits the function $\omega \longmapsto \ln(\sqrt{2\pi}|\mathcal{F}(y)(\omega,0)|)$. With these parameters α, β we solve the associated evolution equation backwards in time.

4.5 Bibliographical comments

The process of (numerical) differentiation is considered in almost all monographs on ill-posed problems. The generalization of the methods proposed in the context of fractional differentiation is discussed for instance in [40]. In Section 4.3 we follow mainly the results developed by Baumeister, Gorenflo, Hofmann, Janno and Wolfersdorf; see [6, 29, 94]. The reverse convolution inequality can be found in [82]. For blind deconvolution consult [12, 16].

4.6 Exercises

4.1. Show that an integral equation

$$\int_{a'}^{b'} \kappa(\frac{x}{x'})g(x')dx' = r(x)\,,\; x \in [c', d'],$$

with a kernel κ of the division type can be transformed to an integral equation of convolution type:

$$\int_a^b h(t-s)\tilde{g}(s)ds = \tilde{r}(t)\,,\; t \in [c, d]\,.$$

4.2. Consider for $\alpha, \beta > 0$, the convolution of the following two signals:

$$x(t) := \begin{cases} e^{-\alpha t} & , \text{if } t > 0 \\ 0 & , \text{if } t \leq 0 \end{cases}\,,\; x(t) := \begin{cases} e^{-\beta t} & , \text{if } t > 0 \\ 0 & , \text{if } t \leq 0 \end{cases}\,.$$

Compute the convolution $h * x$.

4.3. Consider the sideways heat equation in the quarter plane:

$$\begin{cases} u_{xx} = u_t & , x \geq 0, t \geq 0 \\ u(x, 0) = 0 & , x \geq 0 \\ u(1, t) = g(t) & , t \geq 0 \\ \|u(x, \cdot)\|_{L_2(0,\infty)} & \text{is bounded as } x \to \infty \end{cases}\,. \tag{4.44}$$

Physically, the problem corresponds to a situation in which the endpoint $x = 0$ is inaccessible, but for which one can make measurements at $x = 1$. Find an indication that this problem is ill-posed.

4.4. Consider again the sideways heat equation (4.44). Show that a solution is given by

$$\hat{u}(x, \xi) := e^{(1-x)\sqrt{i\xi}}\hat{g}(\xi) \text{ where } \hat{u}(x, \xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x, \xi)e^{-i\xi t}dt\,,\; \xi \in \mathbb{R}\,. \tag{4.45}$$

Here $\sqrt{i\xi} = (1 + \text{sign}(\xi)i)\sqrt{|\xi|/2}\,$.

4.5. Consider the problem of calculating the fractional derivative of a function f given in $L_2(\mathbb{R})$:

$$D^\alpha f(x) := \frac{1}{\Gamma(n+1-\alpha)} \frac{d^{n+1}}{dx^{n+1}} \int_{-\infty}^{x} \frac{f(t)}{(x-t)^{\alpha-n}} dt$$

for $n \in \mathbb{N}, n < \alpha < n+1$. Such problems are frequently encountered in many practical contexts. It is well known that if $0 < \alpha \leq 1$, then $D^\alpha f(x)$ is a formal solution of the Abel integral equation

$$I^\alpha u)(x) = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^{x} \frac{u(t)}{(x-t)^{1-\alpha}} dt = f(x), -\infty < x\infty\,.$$

Compute $D^\alpha f$ for

$$f(x) := e^{-x^2}\,,\ f(x) := \begin{cases} 0 & , x \leq -1 \\ 1+x & , -1 < x \leq 0 \\ 1-x & , 0 < x \leq 1 \\ 0 & 1 < x \end{cases}\,,$$

$$f(x) := \begin{cases} 0 & , x \leq -1 \\ 1 & , -1 < x \leq 1 \\ 0 & 1 < x \end{cases}\,,\ f(x) := 0\,.$$

Chapter 5

Tomography Problems

Tomography concerns recovering images from a number of projections: how can an image of an object be constructed, of which only the density distribution in a number of directions is known. The main application is the problem of locating tumors, various other applications of the Radon transform (partial differential equations, seismics,...) can be found in the literature.

The exposition is divided into three parts: tomography via the Radon transform, detecting of features in images, Discrete Tomography.

5.1 Computerized Tomography

5.1.1 Problem formulation

Let f describe the density of a medium in a region $\Omega \subset \mathbb{R}^2$. The combined effects of scattering and absorption result in a exponential attenuation of a beam of X-ray photons as it passes through the medium. If I_0 is the input intensity of the beam of X-ray photons the output intensity of the beam is given due to basic modelling by

$$I_0 \exp(-\int\limits_L f(z)dz)$$

where L is the beam path; see Figure 5.1. The line integral

$$\int_L f(z)dz$$

is called a projection. By moving the source of the beam and the detector around the medium it is possible to obtain a set of projections. Then an appropriate inversion algorithm is applied to recover an approximation to the density distribution f. By stacking several transverse sections of a body, the two-dimensional information may be converted to a three-dimensional information.

This method has been applied in various fields of applications: X-ray tomography in medicine (determination of tumors), geophysical tomography (determination of subsurface structure), optical tomography (flow field diagnostics).

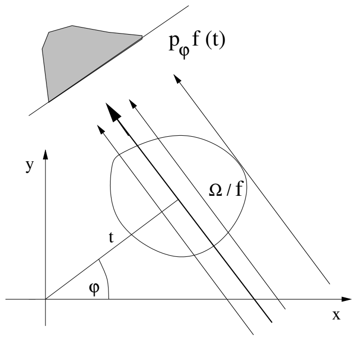

Figure 5.1: The beam geometry

Let us give a more mathematical formulation of the problem. A beam path L may be parameterized by an angle and a distance in the following way:

$$L = L_{t,\varphi} = \{z \in \mathbb{R}^2 | z = tu(\varphi) + su^\perp(\varphi), s \in \mathbb{R}\}$$

where $t \in (-\infty, \infty)$, $u(\varphi) = (\cos\varphi, \sin\varphi)$, $u^\perp(\varphi) = (-\sin\varphi, \cos\varphi)$, $\varphi \in [0, \pi)$. Then the line integral corresponding to $L_{t,\varphi}$ may be written as follows:

$$(p_\varphi f)(t) := \int_{-\infty}^{\infty} f(tu(\varphi) + su^\perp(\varphi))ds. \qquad (5.1)$$

We refer to $p_\varphi f$ as the *radiograph* or *projection* of f in the direction perpendicular to $u(\varphi)$. The function Rf, defined by

$$Rf(t,\varphi) := (p_\varphi f)(t)\,,\; t \in \mathbb{R}, \varphi \in [0,\pi) \tag{5.2}$$

is called the (two-dimensional) *Radon transform* of f; a density in the (x,y)-space is transformed into the "Radon space" (t,φ). Thus, the solution of the reconstruction problem by projections consists in an inversion of the operator R, defined on a suitable space of functions.

An interesting special case consists of a radially symmetric density f and Ω being a circle. In this case is suffices to use a single direction $u(\varphi)$, e.g. $\varphi = \pi/2$, and moreover $f(tu(\varphi) + su(\varphi)^\perp) =: \hat{f}(r)$ can be written as a function of the radius $r = \sqrt{t^2+s^2}$. Using a transformation to polar coordinates, the Radon transform can be rewritten as

$$Rf(t,\pi/2) = 2\int_t^\rho \frac{r\hat{f}(r)}{\sqrt{r^2-t^2}}dr$$

with ρ sufficiently large such that $\hat{f}(r) = 0$ for $r > \rho$. With the notation $g(t) := \frac{1}{2}Rf(t,\pi/2)$, the Radon inversion in this special case can be written as the Abel integral equation

$$g(t) = \int_t^\rho \frac{r\hat{f}(r)}{\sqrt{r^2-t^2}}dr\,,\; 0 < t \le \rho\,.$$

It is possible to find an explicit inversion formula for the Abel integral equation, which yields

$$\hat{f}(r) = -\frac{2}{\pi}\int_t^\rho \frac{g'(t)}{\sqrt{t^2-r^2}}dt\,.$$

Note that in the inversion formula, the derivative g' appears and we have seen in Section 2 that differentiation of data is ill-posed. The differentiation is compensated partly by the additional integration, but one can show that the inversion of the Abel integral equation is still ill-posed; see Example 4.1.1.

For the general situation, an explicit (but more complicated) inversion formula exists, which involves differentiation of data too:

$$f(z) = \frac{1}{2\pi^2}\int_0^\pi \int_{-\infty}^\infty \frac{\partial_t p_\varphi(t,\varphi)}{\langle z, u(\varphi)\rangle - t}\,dt\,d\varphi\,,\; z \in \mathbb{R}^2\,. \tag{5.3}$$

But it is not clear a-priori which properties of the density f are needed in order to make the formula applicable.

The space of functions in which the Radon transform is accessible by simple arguments is the Schwartz space $\mathcal{S}(\mathbb{R}^2)$; see the Appendix A.1. Roughly speaking, the main properties of a function in $\mathcal{S}(\mathbb{R}^2)$ are that it is smooth and that it goes to zero as $|z| \to \infty$ faster than any negative power of $|z|$. In the Schwartz space we can give a second equivalent definition of the Radon transform which is under certain circumstances – at a first look – a little bit easier to handle. We put with $h > 0$

$$\delta_h(t) := \frac{1}{2\pi} \int_{-h}^{h} e^{i(t-s)} ds \, , \ t \in \mathbb{R} \, .$$

Then one can show that $\lim_{h\to\infty} \delta_h \longrightarrow \delta$ in the dual space $\mathcal{S}(\mathbb{R}^2)^*$ of $\mathcal{S}(\mathbb{R}^2)$. Here δ is the one-dimensional Dirac distribution which operates as a linear functional on smooth functions as follows:

$$\langle \delta, g \rangle = g(0) \, .$$

Hence, for $f \in \mathcal{S}(\mathbb{R}^2)$,

$$\lim_{h\to\infty} \int_{\mathbb{R}^2} f(z)\delta_h(t - \langle z, u(\varphi)\rangle)dz \quad = \quad \int_{\mathbb{R}^2} f(z)\delta(t - \langle z, u(\varphi)\rangle)dz$$

for all $t \in \mathbb{R}, \varphi \in [0, \pi)$. In this sense we write

$$Rf(t, \varphi) = \int_{\mathbb{R}^2} f(z)\delta(t - \langle z, u(\varphi)\rangle)dz \, , \ t \in \mathbb{R}, \varphi \in [0, \pi) \, . \qquad (5.4)$$

In the Schwartz space the properties of the mapping R (injectivity, continuity, range,...) can be discussed clearly and completley. The following *projection theorem*, also called the *central slice theorem*, prepares an answer to the uniqueness question. The proof of this theorem can be found in all textbooks on computer tomography. Here we give a sketch of the proof based on the definition (5.4).

Theorem 5.1.1 (Central slice theorem/Projection theorem).

Let $f \in \mathcal{S}(\mathbb{R}^2)$ and let $P_\varphi f$ be the one-dimensional Fourier transform of the projection $p_\varphi f$ for each $\varphi \in [0, \pi)$. Then we have

$$P_\varphi f(\omega) = \sqrt{2\pi}\mathcal{F}(f)(\omega u(\varphi))\,,\ \omega \in \mathbb{R}, \varphi \in [0, \pi)\,. \tag{5.5}$$

Sketch of a Proof:
Let $\omega \in \mathbb{R}, \varphi \in [0, \pi)$. Using the definition (5.4) we can conclude

$$\begin{aligned}
P_\varphi f(\omega) &= \mathcal{F}(p_\varphi f(\cdot))(\omega) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Rf(t, \varphi) e^{-i\omega t} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^2} f(z) \int_{-\infty}^{\infty} \delta(t - \langle z, u(\varphi)\rangle) e^{-i\omega t} dt dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^2} f(z) e^{-i\omega \langle z, u(\varphi)\rangle} dz \\
&= \sqrt{2\pi}\mathcal{F}(f)(\omega u(\varphi)))\,.
\end{aligned}$$

■

Thus, the Fourier transform of a radiograph is proportional to the spectrum of the original object on a beam normal to the direction of the projection beam. Since Rf yields the Fourier transform of f which uniquely determines f the uniqueness question is answered by the central slice theorem: $f \in \mathcal{S}(\mathbb{R}^2)$ *is uniquely determined by* Rf.

Remark 5.1.2. *The Radon transform described above generalizes in many directions.*

- *First, it generalizes by considering integrals over d–planes in $\mathbb{R}^n$ for $n > 2$.*

- *An alternative to the parallel mode of data collection is when data are collected for rays diverging from a single point:* fan beam scanning.

- *Another problem is to recover a function in the exterior of some ball from projections outside the ball. This problem is uniquely solvable, providing the function is decaying fast enough at infinity. The situation is much more difficult when the angles* *are restricted to lie in a strict subset of* $[0,\pi)$ *(*limited angle tomography*) as it is the case when a piece of metal blocks the radiation.*

□

So far, we have been working with R defined in the Schwartz space which is not sufficient for applications. It turns out that R extends in a natural way to other interesting spaces. Since we may prove an estimate

$$\|Rf\|_{L_1(\mathbb{R}\times[0,\pi))} \le c\|f\|_{L_1(\mathbb{R}^2)}\,,\ f\in\mathcal{S}(\mathbb{R}^2),$$

with a constant c we conclude that R can be extended by continuity to $L_1(\mathbb{R}^2)$. The extension to $L_2(\mathbb{R}^2)$ is done in a similar way when we consider densities of compact support. As we know from the general basics for ill-posed problems it is helpful to consider R in an appropriate scale of Hilbert spaces in order to obtain good results for regularization. In the domain of densities f with compact support it is natural to choose the usual scale $H_{s,0}(\Omega)_{s\in\mathbb{R}}$ of Sobolev spaces. In the data space the appropriate scale of Hilbert spaces is the family of Sobolev spaces on the "cylinder" $\mathcal{Z} := \mathbb{R}\times[0,\pi)$. When these spaces are denoted by $H_s(\mathcal{Z})_{s\in\mathbb{R}}$, then one can conclude that the inversion of the Radon transform has degree of ill-posedness $\frac{1}{2}$.

In the following subsections we consider computational schemes for inverting the Radon transform. Of course, Tikhonov's method would be the first choice for such a computation method. But we omit this method since the regularization depends heavily on the choice of an operator B which may be a little bit artificial for the Radon transform; see Chapter 2.

5.1.2 Computational aspects: the Fourier technique

We assume that the image function f is compactly supported in the square $[0,N]\times[0,N]$ and consequently, $p_\varphi(\cdot)$ has compact support in $[0,N]$ where $d \ge N\sqrt{2}$. When working with digital images, a discretized form of the Radon transform is required. A digital image

f is – without loss of generality – a $N \times N$ array of *pixels* $z = (x,y) \in \mathbb{Z}^2, 0 \leq x,y \leq N-1$, each representing the average gray level of a unit square in the image. The gray levels can be taken to be nonnegative real numbers. A line integral along $L_{t,\varphi}$ is approximated by a summation of the pixels lying in the one-pixel-wide strip $t - \frac{1}{2} \leq \langle z, u(\varphi)\rangle < t + \frac{1}{2}$. Since strips have unit width, t can be restricted to integer values, and for a given φ at most $\sqrt{2}N$ strips are needed. The number of angles is defined to be uniformly distributed between 0 and π.

In this discrete context the line integrals are computed as follows. For any given angle φ, each of the pixels lies in exactly one strip. Therefore, for each pixel we simply determine the strip to which it belongs (t relative to φ) and add the pixel's value to the current total for strip (t,φ). This procedure is repeated for each value of φ.

```
Input             Density $f$, mesh-parameter $M,N$.

Initialization:   $Rf(t,\varphi) := 0$ for all $(t,\varphi)$.
          For $\varphi = 0$ to $\frac{(M-1)\pi}{M}$ step $\frac{\pi}{M}$ do:
          For $x = 0$ to $N-1$, for $y = 0$ to $N-1$ do:
Summation:        $t := \lfloor \langle (x,y), u(\varphi)\rangle + \frac{1}{2} \rfloor$,
                  $R(t,\varphi) := R(t,\varphi) + f(x,y)$.

Output            $M$ radiographs.
```

The *complexity* of this method is $O(N^2M)$. Much research has been devoted to speed up the computation of the discrete Radon transform (parallel processing, computation via the central slice theorem, computation of segments which share different strips,...).

For the (approximate) inversion of the Radon transform we may use the central slice theorem.

Input	`Given a density` f `with radiographs` $p_1 f := p_{\varphi_1} f, \ldots, p_m f := p_{\varphi_m}$ `on a grid` $\omega_1, \ldots, \omega_n$
Fourier Transform:	`Compute the discrete Fourier transform` $P_{i,j} := \mathcal{DF}(p_i)(\omega_j), 1 \leq i \leq m, 1 \leq j \leq n$.
Interpolation:	`Place` P_{ij} `on a cartesian grid with` $(x,y) \in \mathbb{Z}^2, 0 \leq x \leq N-1, 0 \leq y \leq M-1$.
Fourier Transform:	`Use the inverse discrete Fourier` `transform to obtain` $\hat{f}$ `from the data` P_{ij}.
Output	`Approximation` $\hat{f}$ `for the density` f.

The approximate computation of f from Rf in this way is called a *Fourier technique.* The computational problem is that a two-dimensional inverse transform is required. In addition, various coordinate system shifts and interpolations that complicate the calculations further are needed. In practice, resampling from polar to rectangular coordinates involves considerable interpolation, which makes the resultant reconstruction of f noisy; see remark below. A way around this problem is to use the filtered backprojection method; see Section 5.1.3.

Remark 5.1.3. *When one computes a radiograph via the central slice theorem one has to interpolate the density f from a cartesian grid in the frequency space into a polar grid. Let f be a density whose support in the Fourier domain is contained in the square $[-Q,Q] \times [-Q,Q]$. Suppose that we have an equidistant cartesian grid $\mu_i = \nu_i = i\Delta, i = -I \ldots, I$, where $I\Delta := F$. Consider grid points φ_j with $\varphi_j = j\Delta_\varphi$, $j = 0, \ldots, J-1$, $\omega_l = l\Delta_\omega$, $l = -L, \ldots, L$, where $J\Delta_\varphi = \pi, L\Delta_\omega := Q$. Then the maximal distance of two points in the polar grid is given by $d := \sqrt{2}\Delta_\varphi Q$. d should not be bigger than Δ. This leads to the bound $\Delta_\varphi \leq \dfrac{\Delta}{\sqrt{2}Q}$. When the given density is sampled on an equidistant grid which fulfills the requirement of the Nyquist-sampling theorem then we obtain the sampling requirement for the Radon trans-*

form:

$$\Delta_\varphi \le \frac{1}{\sqrt{2}I}\frac{F}{Q} \textit{ with } \frac{F}{Q} \ge 1\,. \tag{5.6}$$

□

5.1.3 Computational aspects: filtered backprojection

The basis of the filtered backprojection technique is the following result.

Theorem 5.1.4 (Inversion formula). *Let $f \in \mathcal{S}(\mathbb{R}^2)$ and let $P_\varphi f$ be the one-dimensional Fourier transform of the projection $p_\varphi f$ for each $\varphi \in [0,\pi)$. Then we have*

$$f(z) = \frac{1}{2\pi}\int_0^\pi (\kappa * p_\varphi f)(\langle z, u(\varphi)\rangle)d\varphi\,,\ z \in \mathbb{R}^2, \tag{5.7}$$

where the Fourier transform K of the convolution kernel κ is given by $K(\omega) := |w|, \omega \in \mathbb{R}$.

Sketch of a Proof:
We have

$$\begin{aligned}
f(z) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} \mathcal{F}(\mu,\nu)e^{i(\mu x+\nu y)}d\mu d\nu \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty \omega\mathcal{F}(\omega u(\varphi))e^{i\omega\langle z,u(\varphi)\rangle}d\omega d\varphi \\
&= \frac{1}{2\pi}\int_0^{\pi}\int_{-\infty}^\infty |\omega|\mathcal{F}(\omega u(\varphi))e^{i\omega\langle z,u(\varphi)\rangle}d\omega d\varphi \\
&= (2\pi)^{-\frac{3}{2}}\int_0^{\pi}\int_{-\infty}^\infty |\omega|P_\varphi f(\omega)e^{i\omega\langle z,u(\varphi)\rangle}d\omega d\varphi \\
&= \frac{1}{2\pi}\int_0^\pi \mathcal{F}^{-1}(KP_\varphi f)(\langle z,u(\varphi)\rangle)d\varphi \\
&= (2\pi)^{-\frac{3}{2}}\int_0^\pi (\mathcal{F}^{-1}(K) * p_\varphi f)(\langle z,u(\varphi)\rangle)d\varphi \\
&= (2\pi)^{-\frac{3}{2}}\int_0^\pi (\kappa * p_\varphi f)(\langle z,u(\varphi)\rangle)d\varphi\,.
\end{aligned}$$

■

Notice that the kernel κ cannot be a $L_1(\mathbb{R})-$ nor a $L_2(\mathbb{R})$-function since the *ramp filter* K is an unbounded function. Thus, for the use of the Fourier transform and the convolution theorem in the sketch of the proof above one has to add some arguments.

Let $f \in \mathcal{S}(\mathbb{R}^2)$ and let $P_\varphi f$ be the one-dimensional Fourier transform of the projection $p_\varphi f$ for each $\varphi \in [0,\pi)$. We define

$$B_f(\xi,\varphi) := (2\pi)^{-\frac{3}{2}} \int_{-\infty}^{\infty} |\omega| P_\varphi f(\omega) e^{i\omega\xi} d\omega \, , \, \xi \in \mathbb{R}, \varphi \in [0,\pi) \, . \quad (5.8)$$

From the proof of the central slice theorem we conclude

$$f(z) = \int_0^\pi B_f(\langle z, u(\varphi)\rangle) d\varphi \, , \, z \in \mathbb{R}^2 \, , \quad (5.9)$$

and

$$B_f(\xi,\varphi) = (2\pi)^{-1}\mathcal{F}^{-1}(\mathcal{F}(\kappa)\mathcal{F}(p_\varphi(\cdot))(\xi) = (2\pi)^{-1}(\kappa * p_\varphi f(\cdot))(\xi) \, .$$

To get f from B_f via the identity (5.9) is called **backpropagation**. Therefore we have obtained that f is the backpropagation of filtered radiographs. The realization of this fact as a computational method is called *filtered backprojection.* Filtered backprojection is widely used in medicine. Here to reduce the number of projections is desirable in order to reduce the X-ray dose. But this is in conflict with the need to avoid aliasing as a consequence of insufficient sampling.

5.1.4 Computational aspects: the ART-algorithm

Suppose we have a system of linear equations

$$Ax = y \quad (5.10)$$

governed by the matrix $A \in \mathbb{R}^{m,n}$ with righthand side $y \in \mathbb{R}^m$, $y^t = (y_1, \ldots, y_m)$. Let $(a^i)^t \in \mathbb{R}^n, i = 1, \ldots, m,$ be the rows of the matrix A . Set

$$J(x) := \sum_{i=1}^m |(a^i)^t x - y_i|^2 \, , \, x \in \mathbb{R}^n \, .$$

We want to minimize the functional J. The method we chose is iterative and of adaptive type: in any step, an estimate of x at the next iteration step is constructed from that at the preceding step and from a "new observation" given by a pair (a^i, y_i); adaptation is done by moving a certain (small) step in the direction opposite to the current gradient of the objective function J . This procedure leads to the following form of a recursive algorithm:

$$x^{k+1} := x^k + \lambda_k(y_k - (a^k)^t x^k)a^k \, , \, k \in \mathbb{N} \, . \tag{5.11}$$

where x^1 is an initial guess and $(\lambda_k)_{k\in\mathbb{N}}$ is a sequence of relaxation parameters. Moreover, we use the data (a^i, y_i) in a cyclic order:

$$a^k = a^j, y_k = y_j, \text{ if } j = k \mod m \, .$$

In the literature an algorithm of this type is called an *ART-algorithm (algebraic reconstruction technique).* In our context of computerized tomography

y_i is the result of a projection,
a^i describes the geometry of the beam,
x is the unknown vector of the pixel density.

Thus, the ART-algorithm may be considered as "the most direct method" to invert the Radon transform in its discretized version.

To simplify in the sequel the computations we assume that the vectors a^i are normalized:

$$(a^i)^t a^i = 1 \, , \, i = 1, \ldots, m \, .$$

Moreover, we consider cyclic relaxation only, that is: λ_k is constant during a cycle.

Let $A^\dagger$ be the pseudoinverse of A, $x^\dagger := A^\dagger y$ and set $\alpha := |A^\dagger x^\dagger - y|^2$.

Theorem 5.1.5. *Let the sequence $(x^l)_{l\in\mathbb{N}}$ be determined by iteration and suppose that $x^1 \in range(A^t)$. Then we have:*

a) *If the system is consistent* ($\alpha = 0$) *and if the relaxation sequence* $(\lambda_k)_{k\in\mathbb{N}}$ *satisfies*

$$0 \le l_k \le 2, k \in \mathbb{N}, \sum_{k=1}^{\infty} \lambda_k(2-\lambda_k) = \infty,$$

then $x^\dagger = \lim_k x^{km}$.

b) *If the system is inconsistent* ($\alpha > 0$) *and if the relaxation sequence* (λ_k) *satisfies*

$$0 \le l_k, k \in \mathbb{N}, \sum_{k=1}^{\infty} \lambda_k = \infty, \sum_{k=1}^{\infty} \lambda_k^2 < \infty,$$

then $x^\dagger = \lim_k x^{km}$.

Sketch of a Proof:
The source condition $x^1 \in \text{range}(A^t)$ implies

$$x^l \in \text{null}(A)^\perp, v^l := x^l - A^\dagger y \in \text{null}(A)^\perp, l \in \mathbb{N}.$$

Put

$$b_k := |x^{km} - x^\dagger|^2, c_k := \sum_{j=1}^{m} |y_j - (a^j)^t x^{km+j-1}|^2, k \in \mathbb{N}.$$

Then one can verify

$$\lambda_k(\mu_k c_k - \alpha) \le b_k - b_{k+1}, k \in \mathbb{N},$$

where

$$\mu_k := \begin{cases} 2-\lambda_k & , \text{if } \alpha = 0 \\ 1-\lambda_k & , \text{if } \alpha > 0 \end{cases}.$$

Now one has to analyze the inequality above in order to obtain $\lim_k b_k = 0$. ■

The attractivity of the ART-algorithm comes from the following facts:

- the iteration step is easy to implement;

- no extra effort is necessary to add data (a^i, y_i).

The shortcomings of this type of computation scheme are

- that the convergence is slow in general;
- that it is difficult to implement a stopping rule for the iteration, especially when the data are corrupted by noise.

Remark 5.1.6. *Consider the Radon transform for densities in a Hilbert space X. Each radiograph p_φ defines a mapping from X into a Hilbert space, say Y. The inversion problem for finite many undiscretized radiographs can be stated as follows:*

> *Given $g_i := p_i f := p_{\varphi_i} f \in Y, i = 1, \ldots, m$.*
> *Find an approximation of f.*

Let p_i^* be the adjoint mapping of p_i from $X \to Y$ and consider the iteration

$$f^{i+1} := f^i + \lambda p_i^* (p_i p_i^*)^{-1} (g_i - p_i f^i), i \in \mathbb{N},$$

where f^1 is a given initial guess and $\lambda \in (0, 2)$ is a relaxation parameter. Again, the data g_i are used in a cyclic way. The limit of the sequence $(f^i)_{i \in \mathbb{N}}$ which can be shown to exist under minor assumptions is a solution of the reconstruction problem.

This is an infinite-dimensional variant of the ART-algorithm. It may be considered as a method of successive iteration of nonexpansive mappings. □

5.2 Features in images

Suppose we look at a two-dimensional image. Extraction of primitives "hidden" in the density, such as lines (airfield runways), wires (detection of mines) and curves, is often a key step in an image analysis procedure. The most popular technique for curve detection is based on the Hough transform – we don't present it as a transform – which is closely related to the Radon transform.

5.2.1 The Radon transform for shape detecting

Let I be a plane image. The image intensity can be regarded as a function $f(z)$ of the position $z \in \mathbb{R}^2$ in the image. The Radon transform according to definition (5.4) is given by

$$Rf(t,\varphi) = \int_{\mathbb{R}^2} f(z)\delta(t - \langle z, u(\varphi)\rangle)dz\,,\ t \in \mathbb{R}, \varphi \in [0,\pi)\,.$$

The projections are integrals along straight lines. Therefore the projections should enhance a detail in the image which is of the shape of a line.

Consider a line in normal form:

$$L_{\rho,\vartheta} : \langle z, u(\varphi)\rangle = \rho \ \ (\rho \in \mathbb{R}, \vartheta \in [0,\pi))$$

in $\mathbb{R}^2$. Modelling the density f on the line $L_{\rho,\vartheta}$ with a Dirac distribution δ gives certainly $Rf(t,\varphi) = 0$ when $\varphi = \vartheta$ and $t \neq \rho$. When $\varphi \neq \vartheta$ then we obtain

$$Rf(t,\varphi) = \int_{\mathbb{R}^2} \delta(\rho - \langle z, u(\vartheta)\rangle)\delta(t - \langle z, u(\varphi)\rangle)dz = -\frac{1}{\langle u(\varphi^\perp), u(\vartheta)\rangle}\,.$$

This implies in the case $t = \rho$ and $\varphi = \vartheta$ by a limit argument that a peak in (ρ, ϑ) results. Thus, by considering the Radon transform Rf it should be possible to detect a detail which is of the shape of a line.

Now, it is not difficult to generalize the Radon transform to other types of shapes. Suppose that a shape is given implicitely by an equation

$$\gamma_p : \Gamma(z,p) = 0\,,$$

where p is a parameter in a subset P of a *parameter space* $\mathbb{R}^d$. Examples are lines

$$t - \langle z, u(\varphi)\rangle = 0 \ \ (p := (t,\varphi) \in P := \mathbb{R} \times [0,\pi)),$$

and circles (extract a football in an image!)

$$|z - z^0|^2 - r = 0 \ \ (p := (z^0, r) \in P := \mathbb{R}^2 \times [0,\infty))\,.$$

Let γ_p be a family of given shapes. The generalization of the Radon transform is given by

$$R_{\gamma_p} f(t,\varphi) = \int_{\mathbb{R}^2} f(z)\delta(\Gamma(z,p))dx \tag{5.12}$$

where again δ is the Dirac distribution. Clearly, this is just a very informal definition but we omit the arguments to make this definition sound.

Now imagine that there is a shape in the image with parameter q. When $q \neq p$, the Radon transform will evaluate to some finite number which is proportional to the number of intersections between the shapes γ_q and γ_p. However, when $p = q$, the Radon transform yields a large response, a peak in the parameter space. We can now interpret the Radon transform as follows: it provides a mapping from image space to a parameter space. The mapping created in this way contains peaks for those p for which the corresponding shape $\gamma(p)$ is present in the image. Shape detection is reduced to the simpler problem of peak detection.

Let us go back to a very simple feature in an image and consider the Radon transform of a point source. Modelling the point source in $z^* = (x^*, y^*)$ by a Dirac distribution

$$f := \tilde{\delta}(\cdot - z^*) := \delta(\cdot - x^*)\delta(\cdot - y^*),$$

we obtain in a formal argumentation

$$Rf(t,\varphi) = \delta(t - \langle z^*, u(\varphi)\rangle\,,\, t \in \mathbb{R}, \varphi \in [0,\pi)\,.$$

For each pair (t,φ) with $t - \langle z^*, u(\varphi)\rangle$ we have the line

$$t - \langle z, u(\varphi)\rangle = 0\,,\; z \in \mathbb{R}^2\,.$$

In this way every point $z^* \in \mathbb{R}^2$ is mapped into a *sinusoid* σ_{z^*} with representation

$$\sigma_{z^*} : [0,\pi) \ni \varphi \;\longmapsto\; t := \langle z^*, u(\varphi)\rangle \in \mathbb{R}\,.$$

This sinusoid has its maximum value in (ρ, ϑ) with

$$\rho := |z^*|\,,\ \cos(\vartheta) = x^*/\rho\,.$$

Conversely, each point of the graph of such a sinusoid represents a point lying on a line through z^*.

Example 5.2.1. *Consider the points $A(0|2), B(1|1), C(2|0)$ in $\mathbb{R}^2$. These points are mapped into the following sinusoids in the parameter space:*

$$\rho = \cos(\vartheta) + \sin(\vartheta)\,,\ \rho = 2\sin(\vartheta)\,,\ \rho = 2\cos(\vartheta)\,,\ \vartheta \in [0, \pi)\,.$$

The intersection point $(\rho = 1, \vartheta = \pi/4)$ of these curves indicates that the points are lying on the line

$$x + y = 2\,.$$

□

5.2.2 The Hough accumulator for detecting lines

If an image is very sparse, e. g., a binary image with only a few nonzero pixels, most of the computational effort to evaluate a discrete Radon transform is summing up zeros that do not contribute to the value of the projection. In one of the most cited patents in the image processing Hough proposed a way to incorporate sparsity; see [45].

The *Hough transform* is a standard tool in image analysis that allows recognition of global patterns in an image space by recognition of local patterns (ideally a point) in a transformed parameter space. It is particularly useful when the patterns one is looking for are sparsely digitized and/or the pictures are noisy.

The basic idea of this technique is to find curves that can be parameterized like straight lines, polynomials, circles in a suitable parameter space (Hough space). The Hough transform is a mapping from the image space into the Hough space. We set up an d-dimensional accumulator array, each dimension corresponding to one of the parameters of the shape looked for. Each element of this array contains the number of votes for the presence of a shape with the parameter corresponding to that element. The votes themselves are obtained as follows. Consider each point (x, y) in the input image. Now we consider which shapes this point, with grey value $g(x, y)$, could potentially be a member of. We increment the vote for each of these shapes with $g(x, y)$. Of course, if a shape with parameter p is present in the image, all of the pixels that are part of it will vote for it, yielding a large peak in an accumulator array.

This method was originally defined to detect straight lines in binary images. Let

$$\begin{aligned}
x_m &= x_{\min} + m\Delta x\,,\ m = 0, \ldots, M-1\,, && (5.13)\\
y_n &= y_{\min} + n\Delta y\,,\ n = 0, \ldots, N-1\,, && (5.14)\\
\rho_l &= \rho_{\min} + l\Delta\rho\,,\ l = 0, \ldots, L-1\,, && (5.15)\\
\vartheta_k &= k\Delta\vartheta\,,\ k = 0, \ldots, K-1\ \ (\vartheta_K = \pi)\,, && (5.16)
\end{aligned}$$

be a discretization of the image and parameter space (Hough space), respectively. Let f be a given image with nonnegative “greyvalues” and let

$$g(m, n) = f(x_m, y_n)\,,\ m = 0, \ldots, M-1, n = 0, \ldots, N-1\,,$$

be the pixel image of f. Here is the computational scheme for the *Hough-accumulator*.

```
INPUT             Discretization $(5.13),\dots,(5.16)$;
                  Discretized image $g(m,n)$.

Initialization:   $h(l,k) = 0$ for all $k, l$.

            For $m = 0,\dots,M-1, n = 0,\dots,N-1$ do:

Contribution :    $g := g(m,n)$; if $g \neq 0$ do
                  for $k = 0,\dots,K-1$ do
                  $\rho := x_m \cos(\vartheta_k) + y_n \sin(\vartheta_k)$
                  $r :=$ round$(\rho - \rho_{min})/\Delta\rho$

Voting            if $r \geq 0$ and $r < L$ then $h(r,k) := h(r,k) + g$

            end do.

OUTPUT            Histogram $h(m,n)$ (Hough accumulator)
```

Local maxima of the histogram h identify straight line segments in the original image space. Ideally, the Hough space has to be searched for a maximum only once. In situations where a picture contains many patterns of different size, it may, however, be necessary to take out first those patterns in the original image space that correspond to clearly identifiable peaks in the Hough domain and to repeat the process. But a key question is what happens to peaks in the parameter space when the image is corrupted by noise. For a binary image the value of the accumulator h in a parameter (ρ, ϑ) is related to the number of points lying on the line given by (ρ, ϑ).

Particular care must be taken in choosing the angular resolution Δ_ϑ and the distance resolution Δ_ρ. The values should be such that collinear points in the image space do correspond to curves intersecting at the same Hough accumulator cell.

Remark 5.2.2. *Within the seismics, the parameterization of a line* *by the slope m and the line offset c is used:*

$$Rf(m,c) := \int_{-\infty}^{\infty} f(x, mx+c)dx\,,\ m,c \in \mathbb{R}\,; \tag{5.17}$$

the parameter space is $\mathbb{R}^2$. The Hough-algorithm may be adapted to this case. However, here we have a problem with using $y = mx + c$ to represent lines when the line is vertical. □

Let us sketch two ideas to detect circles in a given image.

Suppose we have a pixel image. Each pair of points in the image space defines a line in the image space and a line that perpendicular bisects this line. Since a line that perpendicular bisects a chord of a circle contains the center of the circle we are able to find eventually – again by voting in the parameter space – the center of a circle. Because we are dealing with digital circles the points of their circumference are affected by digitation and, therefore, do not exactly satisfy the standard circle equation.

When the image is binary (black/white) and when we try to find circles of given radius r we may use the fact that the center of circles with a intersection point x^0 lie themselves on a circle of radius r round the point x^0. Hence the computation goes as follows. For each black pixel (x_m, y_n) find all pixels $(x_{m'}, y_{n'})$ lying on the boundary of a circle with radius r and center (x_m, y_n); increment the vote for $(x_{m'}, y_{n'})$ by one. Highly voted pixels provide an indication of the existence of a circle with many black pixels on the boundary. A parallel implementation is possible.

5.3 Discrete Tomography

Here we have a short look on the reconstruction of binary images from the knowledge of their line sums.

5.3.1 Statement of the problem

A *lattice set* is a non–empty finite subset of the integer lattice $\mathbb{Z}^2$. A vector v in $\mathbb{Z}^2\backslash\{\theta\}$ is called a *lattice direction* and a projection of a lattice set F in a lattice direction v is the function p_vF giving the number of points in F on each line parallel to this direction; see Figure 5.2.

Discrete Tomography is concerned with the reconstruction of images from their projections (in a small number of directions). The reconstruction task is an ill-posed inverse problem touching all three Hadamard criteria for ill-posedness. In fact, for general data there need not exist a solution, if the data is consistent, the solution need not be uniquely determined, and even in the case of uniqueness the solution may change dramatically with small changes of the data.

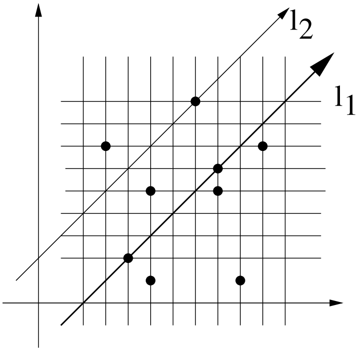

Figure 5.2: A lattice in $\mathbb{Z}^2$

As the name suggests, Discrete Tomography has its theory based on discrete mathematics [11,14]. In addition, it deals with many other fields of mathematics, namely combinatorics, functional analysis, geometry, coding theory and graph theory. It has been applied to diverse areas such as medical sciences, image processing, electron microscopy, scheduling, statistical data security, game theory and material sciences.

Remark 5.3.1. *We don't give the most general formulation of the problems in Discrete Tomography. Especially, we restrict ourselves on lattices in $\mathbb{Z}^2$. Moreover, the results in Subsection 5.3.2 hold in a more general setting.* □

Let us first give the basic notation and definitions. Lattice sets and lattice direction are already introduced. Let $\mathcal{E}$ be the family of lattice sets. Given a lattice direction v, let $\mathcal{A}(v) := \{w + v | w \in \mathbb{Z}^2\}$.

The projection of $F \in \mathcal{E}$ is the mapping

$$p_v F : \mathcal{A}(v) \ni l \longmapsto \#(F \cap l) = \sum_{x \in l} \chi_F(x) \in \mathbb{N}_0 := \mathbb{N} \cup \{0\} \,,$$

where χ_F is the characteristic function of F. Two lattice sets F and F' are said to be *tomographically equivalent* with respect to the lattice directions $v^1, \dots, v^m$ if we have

$$p_{v^k} F = p_{v^k} F' \,,\ k = 1, \dots, m \,.$$

Given m different lattice directions $v^1, \dots, v^m$, the basic questions are as follows. What kind of information about a lattice set F can be retrieved from its projections $p_{v^1}, \dots, p_{v^m}$? How difficult is the reconstruction algorithmically? How sensitive is the task to data errors? Here the data are given in terms of functions

$$g_k : \mathcal{A}(v^k) \longrightarrow \mathbb{N}_0 \,,\ k = 1, \dots, m \,.$$

Let us formulate these questions more technically.

Consistency

> Given data $g_k : \mathcal{A}(v^k) \longrightarrow \mathbb{N}_0, k = 1, \dots, m$, with finite support.
> Question: Does there exist an $F \in \mathcal{E}$ such that $p_{v^k} F = g_k, k = 1, \dots, m$.

Uniqueness

> Given any $F \in \mathcal{E}$.
> Question: Does there exist a subset F' different from F such that F and F' are tomographically equivalent with respect to the directions of $v^1, \dots, v^m$.

Reconstruction

> Given data $g_k : \mathcal{A}(v^k) \longrightarrow \mathbb{N}_0, k = 1, \dots, m$, with finite support.
> Task: Construct a subset $F \in \mathcal{E}$ such that $p_{v^k} F = g_k$ for all $k = 1, \dots, m$.

Example 5.3.2. *Consider in* $\mathbb{Z}^2$ *the directions* $v^1 := (1,0)$ *and* $v^2 := (0,1)$ *and let* $\{(1,0),(2,0),(2,1),(1,3)\}$ *be lattice directions. Here* $q = 5$ *is the number of lattice lines on which there is at least one element from the discrete set.* $n = 6$ *is the total number of points to be reconstructed. The aim is to find a binary vector which satisfies a matrix equation*

$$Bx = b \tag{5.18}$$

where $B \in \{0,1\}^{5,6}, b \in \mathbb{N}_0^6$ *are given as follows:*

$$B := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}.$$

The corresponding system of equations (5.18) *is uniquely solvable.* □

5.3.2 A stability result

This subsection should give just a first impression of the goals in Discrete Tomography. Especially, the stable reconstruction of special sets like the convex hull of a lattice set is in the focus of the present research in Discrete Tomography.

The distance of two functions $g,h : \mathcal{A}(v) \longrightarrow \mathbb{N}_0$ will be measured in terms of the l_1-norm:

$$\|g-h\|_1 := \sum_{l\in\mathcal{A}(v)} |\#g(l) - \#h(l)|\,.$$

Here is a positive result concerning uniqueness and stability.

Theorem 5.3.3 (Alpers/Gritzmann). *Let* $v_1,\dots,v_m$ *be pairwise different lattice directions and let* $F_1,F_2 \in \mathcal{E}$ *with* $\#F_1 = \#F_2$. *If*

$$\sum_{i=1}^{m} |p_{v^i}F_1 - p_{v^i}F_2| < 2(m-1)$$

then F_1 *and* F_2 *are tomographically equivalent.*

Sketch of a Proof:
The proof is a rather deep combination of combinatorical and algebraic arguments. ■

Corollary 5.3.4. *Given two sets $F_1, F_2 \in \mathcal{E}$ with $\#F_1 = \#F_2$, and $m \geq \#F_1 + 1$ pairwise different lattice directions $v_1, \dots, v_m$. If*

$$\sum_{i=1}^{m} |p_{S_i} F_1 - p_{S_i} F_2| < 2\#F_1,$$

then $F_1 = F_2$.

Sketch of a Proof:
By Theorem 5.3.3 F_1 and F_2 are tomographically equivalent. Due to Rényi's famous theorem we have that if the cardinality $\#F$ of a finite set is known uniqueness is guaranteed from projections taken in any $m \geq \#F + 1$ directions. ■

The interpretation of Corollary 5.3.4 is that error correction is possible: a total error smaller than $2n$ can be compensated without increasing the number of projections if the number n of elements in the original set F_1 is known.

5.4 Bibliographical comments

A rather complete development of the mathematics of computerized tomography can be found in [74]; see also [95]. The ART-algorithm, originally proposed by Kaczmarz was rediscovered several times. It may be considered as a special case for the successive iteration of nonexpansive mappings; see [7]. For a numerical analysis of the Fourier technique see for instance [30]. Results concerning Tikhonov's method in the context of computer tomography may be found in [2].

The use of the Radon transform for shape detection dates back to the sixties of the last century; see for instance [45] and [81]. For a survey of the Hough transform literature up to 1988 see [47]. The detection of circles is discussed in [48].

The name Discrete Tomography was given at 1994 by L. Shepp, the question of uniquely determining a planar convex object was already proposed by P.C.Hammer in 1963. For further information on the theory, algorithms and applications of Discrete Tomography see [1] and [42] and the references there in. The research in Discrete Tomography is in progress and many problems are open.

5.5 Exercises

5.1. Verify for functions $f, g \in \mathcal{S}(\mathbb{R}^2)$:

$$\begin{aligned} Rf(t,\varphi) &= Rf(-t,-\varphi) \\ R(f+g) &= Rf + Rg \\ p_\varphi \frac{\partial}{\partial x} f &= \varphi \frac{\partial}{\partial t} p_\varphi f \\ p_\varphi(f * g) &= p_\varphi f * p_\varphi g \\ \int_0^\pi \int_{-\infty}^{\infty} Rf(t,\varphi) g(t,\varphi) dt d\varphi &= \int_{\mathbb{R}^2} f(z) R^\# g(z) dz \end{aligned}$$

$$\text{where } R^\# g(z) = \int_0^\pi g(\langle z, u(\varphi)^\perp \rangle) dz\,.$$

5.2. Let $f : \Omega \longrightarrow \mathbb{R}$ be an image. Compute the result in the Radon transform of f when the image is rotated by an angle ψ.

5.3. Compute the Radon transform of the density

$$f(z) := \begin{cases} 1 & , \text{ if } |z| < 1 \\ 0 & , \text{ if } |z| \geq 1 \end{cases} \quad \text{(Shepp Logan Phantom)}\,.$$

5.4. Using the basic properties above compute the Radon transform of constant densities with support in circles and ellipses.

5.5. Compute the Radon transform of the density

$$f(x,y) := \begin{cases} 1 & , \text{ if } (x,y) \in [-1,1] \times [-1,1] \\ 0 & , \text{ else} \end{cases}.$$

5.6. Compute the Radon transform of Gaussian bell

$$f(z) := e^{-|z|^2}\,.$$

5.7. Consider a square pixel image with 4 pixels. Then we have 6 (meaningful) projections (2 horizontal, 2 vertical, 2 diagonal). Suppose the values of this projections are 12,8,11,9,13,7, respectively. Write down a system $\langle a_i^t, x\rangle = y^i, i = 1, \ldots, 6$, in order to model the reconstruction of this pixel image and compute a solution $x \in \mathbb{R}^6$ by the ART-algorithm.

5.8. Consider the matrix

$$A := \begin{pmatrix} 1 & 3 & 2 & -1 \\ 1 & 2 & -1 & -2 \\ 1 & -1 & 2 & 3 \\ 2 & 1 & 1 & 1 \\ 5 & 5 & 4 & 1 \\ 4 & -1 & 5 & 7 \end{pmatrix}$$

and let $y^t = \begin{pmatrix} 5 & 0 & 5 & 5 & 15 & 15 \end{pmatrix}$.

a) Compute $\text{rank}(A)$.

b) Compute the manifold of the solutions of

$$Ax = y\,. \tag{5.19}$$

5.9. Consider the system (5.19).

a) Compute the first 4 iterations of the ART-algorithm with relaxation factor $\lambda = 1$.

b) Compute the first 4 iterations of the ART-algorithm with relaxation factor $\lambda = 0.5$.

5.10. Suppose that the support of an image is contained in $I := [a, b] \times [c, d]$. Find the parameterization of a line through I from the points where the line enters the set I and where the line exists the set I .

Chapter 6

Level set methods

In this chapter, a recently developed methodology for solving inverse problems involving obstacles is investigated. This approach is based on the so called *level set methods*, which has been shown to be effective in treating problems of moving boundaries, particularly those that involve topological changes in the geometry. These methods can be applied to a particular class of inverse problems where the desired unknown is a region in $\mathbb{R}^n$. The region is possibly multiply connected or consisting of several subregions. A classical example is the inverse scattering problem for an obstacle (see, e.g., [18]).

We shall concentrate on three different level set approaches for inverse problems. The first one was suggested by Santosa in 1996 (see [83]), who introduced level set theory into the context of inverse problems. The second one corresponds to the results obtained by Burger in 2001 (see [9]) and contains a first formal mathematical analysis (focusing on regularization theory) of a level set method. The last approach was introduced by Leitão and Scherzer in 2003 (see [59]) and makes a correspondence between level set theory and constraint optimization.

6.1 Introduction

We start this chapter by introducing the inverse problems which can be handled by level set type methods, the so called *inverse problems involving obstacles*. This particular family of inverse problems is characterized by the fact that the desired unknown is a subset $D \subset \mathbb{R}^n$. Alternatively, one can think on the determination of the characteristic function of an unknown set D.

As usual in the framework of inverse problems, we assume that only indirect data is available for the determination of the unknown set D. Abstractly, we can formulate the problem as follows:

Let $\Omega \subseteq \mathbb{R}^n$ be a given (fixed) set, X, Y Hilbert spaces, and $F : X \to Y$ a Fréchet differentiable operator. Find $D \subset \mathrm{int}(\Omega)$ in the equation

$$F(u) = g\,, \tag{6.1}$$

where

$$u = \begin{cases} u_{int}, & x \in D \\ u_{ext}, & x \in \Omega / D \end{cases} .$$

Here u_{int}, $u_{ext} \in \mathbb{R}$ are given constants. The function g represents the problem data. The set D represents the (unknown) model parameters. The operator F is the parameter to output operator, i.e. a map from the model to the data.

Some possible applications are: inverse scattering, mine detection, inverse potential problem, deblurring. In the sequel we briefly discuss the level set approach for each one of these problems.

Inverse scattering by an obstacle

The operator F represents the map to the far field pattern from a scatterer D, for a given set of incident waves. For this example u_{ext} is the sound speed of the exterior propagating medium. Instead of defining u_{int}, boundary conditions (sound-soft or sound-hard) on the wave field on ∂D are prescribed (see [18]).

Mine detection

For the mine detection problem, u_{int} is the conductivity of the mine while u_{ext} represents the conductivity of the soil. The fixed region to be analyzed is denoted by $\Omega \subset \mathbb{R}^2$ and the set $D \subset \mathrm{int}(\Omega)$ represents the position of the mines. In this application, the set D is obviously disconnected. For more details we refer to [26].

Inverse potential problem

In this inverse problem $\Omega \subset \mathbb{R}^2$, or $\mathbb{R}^3$ is known and $D \subset \mathrm{int}(\Omega)$ has to be reconstructed from (partial) knowledge of the function U, which solves

$$\begin{cases} \Delta U = \chi_D & , \text{ in } \Omega \\ U = 0 & , \text{ on } \partial\Omega \end{cases}$$

(here χ_D denotes the characteristic function of the set D). The function U is the potential corresponding to the unknown source χ_D. Two variants of this inverse problem can be considered. In the fist one, the parameter to output operator is given by $F_1 : L^2(\Omega) \to H^1(\Omega)$, $F_1(\chi_D) = U$. In the second problem, the model operator is defined by $F_2 : L^2(\Omega) \to H^{-1/2}(\partial\Omega)$, $F_2(v) = (\partial U/\partial\nu)|_{\partial\Omega}$, where ν is the outer normal vector to $\partial\Omega$.

The first problem is simpler, since F_1 is the inverse of the Laplace operator with homogeneous boundary conditions (F_1 is compact). The analysis of the second problem is more demanding. However, F_2 has the nice property of being a linear operator, what is very uncommon for parameter reconstruction problems (see [23]). For a detailed analysis of the inverse problems for F_1 and F_2 we refer to [23] and [43] respectively (see also Exercises 6.2 to 6.5).

Deblurring (deconvolution)

A simple deconvolution in two dimensions is modeled by the linear operator $F : L^2(\Omega) \to L^2(\Omega)$,

$$F(u)(x) \ = \ \int_\Omega k(x-y)u(y)\ dy\,,\ x \in \Omega\,.$$

We assume that the kernel k is defined by a Gaussian: $k(x) := \exp(-\sigma|x|^2)$. Therefore, the operator F is compact and selfadjoint. Implicit in this model is the assumption that u is a characteristic function, i.e. it satisfies $u = \chi_D$, for some $D \subset \mathrm{int}(\Omega)$ (for this application $u_{int} = 1$, $u_{ext} = 0$).

This is a classical inverse problem and is usually presented as a tutorial problem. In particular, the ill-conditioning of the operator F can be clearly observed in its numerical discretizations. For details we refer to [5, 23, 53].

It is worth mentioning that, recently, level set methods have been successfully applied for the solution of several other inverse problems (see, e.g., [9, 20, 50, 64, 77, 78, 83]).

6.2 First level set approach for inverse problems

Level set methods were originally developed by Osher and Sethian [76, 88] for problems involving the motion of curves and surfaces. A particular advantage of this approach is the ability of the method to track the motion through topological changes. An other attribute of the method is that it gives a natural way of describing closed curves, specially, those that sequentially change following a certain rule.

A first attempt to introduce the level set approach to inverse problems was presented by Santosa in [83]. In this section we shall focus on this approach. It is important to remark that, up to now, no rigorous analysis of the method investigated in this section has been developed.

For simplicity, let's consider a two dimensional problem, where $\Omega \subset \mathbb{R}^2$ is known and the set of interest is $D \subset \mathrm{int}(\Omega)$.

The boundary of D is described by a function $\phi : \Omega \to \mathbb{R}$, i.e.

$$\partial D \; = \; \{x \in \Omega\, ; \;\phi(x) = 0\} \,.$$

The function ϕ is called *level set function* and the level set approach consists of generating a sequence of functions $\phi_k : \Omega \to \mathbb{R}$ such that $D_k \to D$, where $\partial D_k = \{x \in \Omega\, ; \; \phi_k(x) = 0\}$. Notice that k, the evolution parameter, may be considered continuous as well.

In terms of the function ϕ, we obtain a level set representation of the characteristic function u, namely

$$u(x) \; = \; \left\{ \begin{array}{l} u_{int}, \; \{x \in \Omega\, ; \; \phi(x) < 0\} \\ u_{ext}, \; \{x \in \Omega\, ; \; \phi(x) > 0\} \end{array} \right.$$

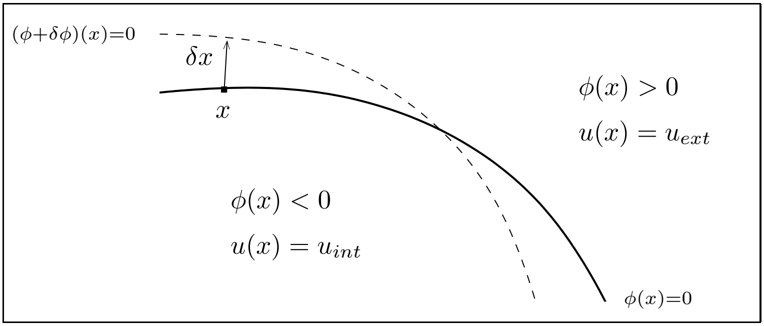

Figure 6.1: Infinitesimal variation of the level set curve $\phi(x) = 0$.

Immediately one observes the following characteristics of this approach:
1) The level set representation of a given $D \subset \mathrm{int}(\Omega)$ is not unique. Indeed, if a function ϕ gives a level set representation of D, then $\psi(x) = c\phi(x)$ with $c > 0$ is also a level set function for D (actually, if ϕ_1 and ϕ_2 give a level set representation of D, then any linear combination $a\phi_1 + b\phi_2$, with $a, b > 0$ also does).
2) No *a priori* assumptions on the topology of D is required, i.e. D could be made up of several disconnected subregions.
3) The dependence of u on ϕ is nonlinear, therefore the inverse problem becomes nonlinear, even if F is a linear operator.

In the sequel we derive an evolution rule for the level set function ϕ. A first goal is to determine the dependence of the forward map with respect to small changes on the obstacle boundary. Therefore, we need to calculate the variation of u caused by a variation in ϕ.

Let x be a point on the surface $\partial D = \{x \in \Omega\,;\ \phi(x) = 0\}$ and suppose that the level set function ϕ is perturbed by a small variation $\delta\phi$ (see Figure 6.1). We denote by δx the resulting variation of the point x and by D' the new region originated from D after the perturbation. Finally, δu is the corresponding variation of the function u (to be computed).

The formal variation of the equation $\phi(x) = 0$ gives us

$$\delta\phi + \nabla\phi \cdot \delta x \ = \ 0. \tag{6.2}$$

Notice that, in Figure 6.1, for each point between x and $x+\delta x$ we have

$u = u_{ext}$ and $(u + \delta u) = u_{int}$. Therefore, for all these points we have $\delta u = u_{int} - u_{ext}$. From an analogous argument, one observes that for points x at the lower part of the picture the equality $\delta u = u_{ext} - u_{int}$ holds.

Next, we test the variation δu with an arbitrary test function $f \in L^2$, obtaining

$$\langle \delta u, f \rangle \ = \ \int_\Omega \delta u(x) f(x) \; dx \ = \ \int_{D \Delta D'} \delta u(x) f(x) \; dx \, ,$$

where $D \Delta D'$ is the symmetric difference between the sets D and D'. Notice that, up to a sign, $\delta u(x) = (u_{ext} - u_{int})$, $x \in D \Delta D'$. Moreover, $\delta u(x) = 0$, $x \in \Omega / D \Delta D'$. Assuming δx to be infinitesimal, it follows

$$\langle \delta u, f \rangle \ = \ \int_{\partial D} (u_{int} - u_{ext}) \, \delta x \cdot n(x) \, f(x) ds(x) \, ,$$

where $ds(x)$ is the arclength and $n(x) = \nabla \phi(x) / |\nabla \phi(x)|$ is the unit normal vector to the curve $\phi(x) = 0$. Here we used the fact that the inner product $\delta x \cdot n(x)$ gives the correct sign to $(u_{int} - u_{ext})$. We can now determine u from the last expression:

$$\delta u \ = \ (u_{int} - u_{ext}) \frac{\nabla \phi(x)}{|\nabla \phi(x)|} \cdot \delta x \Big|_{x \in \partial D} . \tag{6.3}$$

Now we are ready to derive an evolution equation for the level set function $\phi(x)$. Let the free variable t represent (an artificial) time variable. The level set function depends actually on both variables t and x, i.e. $\phi = \phi(x, t)$. We adopt the notation: $\partial D(t) = \{x \in \Omega \, ; \ \phi(x, t) = 0\}$.

We shall search for a least square solution of the inverse problem, i.e. a minimizer of

$$J(u) \ := \ \tfrac{1}{2} \| F(u) - g \|^2 .$$

The derivative $\partial \phi / \partial t$ should be chosen such that $J(u(t))$ is a decreasing function of t. At this point we make the assumption that each point $x \in \partial D(t)$ moves perpendicular to the surface, i.e.

$$\delta x \ = \ v(x, t) \frac{\nabla \phi}{|\nabla \phi|} \tag{6.4}$$

(the value $v(x,t)$ is called *velocity* of the surface $\partial D(T)$ at the point x and time t). Substituting this expression in (6.3), it follows

$$\delta u \;=\; (u_{int} - u_{ext})\; v(x,t)\big|_{x \in \partial D(t)} \,. \tag{6.5}$$

The next step is the computation of $\delta J(u;\delta u)$, the Gateaux derivative of J at direction δu. Using (6.5) we obtain

$$\begin{aligned} \delta J(u;\delta u) \;&=\; \langle F'(u)^*(F(u)-g), \delta u\rangle \\ &=\; \int_{\partial D(t)} [F'(u)^*(F(u)-g)]\,(u_{int}-u_{ext})\,v(x,t)\,ds(x) \end{aligned} \tag{6.6}$$

Now, making the non restrictive assumption $u_{int} > u_{ext}$, we arrive at a natural choice of v (remember that we want δJ to become negative or, at least, non positive)

$$v(x,t) = -F'(u)^*(F(u)-g)\,,\; x \in \partial D(t)\,. \tag{6.7}$$

Notice that $v(x,t)$ remains to be defined for $x \in \Omega/\partial D(t)$. Any function v satisfying (6.7) will generate a δu such that the corresponding $\delta J(u;\delta u)$ is non positive. Therefore, Santosa chose for simplicity

$$v(x,t) = -F'(u)^*(F(u)-g)\,,\; x \in \Omega\,. \tag{6.8}$$

From (6.2), (6.4) and (6.8) we conclude that the corresponding variation of ϕ is given by

$$\begin{aligned} \delta\phi(x,t) \;&=\; -\nabla\phi\cdot\delta x \\ &=\; -\nabla\phi\left(v(x,t)\frac{\nabla\phi}{|\nabla\phi|}\right) \\ &=\; -v(x,t)\,|\nabla\phi| \\ &=\; [F'(u)^*(F(u)-g)]|\nabla\phi|\,. \end{aligned}$$

Thus, we have obtained an initial value problem for the evolution of the level set function, namely

$$\left\{\begin{aligned} \frac{\partial\phi}{\partial t} \;&=\; [F'(u)^*(F(u)-g)]|\nabla\phi|\,,\; x\in\Omega\,,\; t\geq 0 \\ \phi(x,0) \;&=\; \phi_0(x)\,,\; x\in\Omega \end{aligned}\right. \tag{6.9}$$

Remark 6.2.1. *It's worth mentioning that this problem corresponds to the Hamilton–Jacobi equation:*

$$\frac{\partial \phi}{\partial t} + V|\nabla \phi| \;=\; 0\,,$$

with $V(x,t)$ *given by*

$$V \;=\; -F'(u)^*(F(u)-g)\,,\; x \in \Omega\,.$$

Due to the derivation of the evolution above, some properties of Santosa's level set method are obvious:

a) With the choice of velocity (6.8), $J(u(t))$ is a non increasing function of the time variable t;

b) If $\bar{u}$ is a solution of $F(u) = g$ and $\phi(x,\tau) = \bar{u}(x)$, $x \in \Omega$, for some $\tau \geq 0$, then $\frac{\partial \phi}{\partial t}(x,\tau) = 0$. In other words, $\bar{u}$ is a stationary point of the dynamical system (6.9).

c) This evolution corresponds to a (continuous) steepest descent method for the least square functional J. In inverse problems theory this method is also known as *asymptotical regularization* (see [89, 90] for details).

Least square approaches for inverse problems are very common in the literature (see, e.g., [5, 23, 34, 44, 53]). When one tries to apply this technique to inverse problems involving obstacles, it is immediate to observe that the variation of the least square functional J is given by a functional of the residual $F(u) - g$ evaluated along the level set curve (see (6.6) above). Thus, if $\partial D(\tau) = \emptyset$ for some $\tau > 0$, the identity $v(x,t) \equiv 0$, for $t \geq \tau$ immediately follows. A consequence of this fact is that, for numerical implementations, it may be necessary to scale the velocity v along the evolution, in order to avoid the vanishing of $D(t)$.

The particular structure of $\delta J(u;\delta u)$ in (6.6) allows the determination of $v(x,t)$ only at $x \in D(t)$. The *velocity* v, however, must be defined at all Ω. The intuitive extension made by Santosa (see (6.8)) actually does the job, but it does not allow a rigorous analysis of his level set method. For instance, it is not possible to prove that

$D(t) \to D$, a solution of (6.1). As a matter of fact, there is no reasonable metric to measure the convergence of $D(t)$ to D or, alternatively, of $u(t)$ to $\bar{u}$. At this point, a rigorous analysis of level set methods in terms of *regularization theory* was needed (cf. [23]). Other authors proposed alternative least square approaches to problem (6.1), which allowed the first convergence results as well as the verification of regularization properties for level set methods. In the following sections we shall focus on the approaches introduced by Burger in [9] and by Leitão and Scherzer in [59, 27].

6.3 Level sets and asymptotic regularization

In this section we investigate the alternative level set approach for inverse obstacle problems proposed by Burger in [9]. The basic idea is to develop an iterative method related to the well known *asymptotic regularization method* (see [89, 90]), which consists of solving the differential equation

$$\frac{\partial}{\partial t} u(x,t) = -F'(u)^*(F(u) - g)\,,\ t > 0\,, \quad u(x,0) = u_0(x)$$

in order to approximate the solution of (6.1) – we use the same notation of Section 6.1. The guideline for the construction of this level set method is a basic property of the asymptotic regularization, namely

$$\frac{\partial}{\partial t} \|u(x,t) - \bar{u}(x)\|^2 = -2\|F(u) - g\|^2\,,\ t > 0\,, \tag{6.10}$$

where $\bar{u}$ is a solution of (6.1). This identity holds for the evolution of the distance between $u(t)$ and $\bar{u}$ and is fundamental in the derivation of convergence rates for the asymptotic regularization. The same identity was used by Burger in [9] to analyze the convergence of a level set type method.

Notice that Santosa's level set method, discussed in Section 6.2, uses a velocity that lead to a steepest descent flow with respect to the residual $\|F(u) - g\|^2$. With his approach, Burger manages to write down the idea of a descent flow for the level set method in a formal functional analytic framework. In addition, the regularizing properties of the asymptotic regularization can be translated to the level set method in the case of noisy data, i.e. if one only knows a perturbation g^ε of the exact data g satisfying $\|g - g^\varepsilon\| \le \varepsilon$. Just like in the asymptotic regularization, the regularization effect of this level set method comes from an early termination of the evolution at some stopping time $T = T(\varepsilon, g^\varepsilon)$, given by the discrepancy principle, i.e. the minimal time such that the is less than the noise level.

In the sequel we shall consider $F : X = L^2(\Omega) \to Y$ to be a bounded linear operator with unbounded generalized inverse. Furthermore, we assume that there exists $\bar{u} = \chi_D$, for some $D \subset \operatorname{int}(\Omega)$, which solves (6.1). The distance between the evolving characteristic function $u(\cdot, t)$ and $\bar{u}$ is measured by the error functional:

$$E(t) \ = \ \|u(\cdot,t) - \bar{u}\|^2_{L^2(\Omega)} = \int_{D(t)\Delta D} 1 \; dx \ = \ d_S(D(t), D)$$

Here, $d_S(A,B) := |A - B| + |B - A|$ denotes the symmetric difference, $u(\cdot, t) = \chi_{D(t)}$ and the sets $D(t)$, $t \ge 0$, are defined by

$$D(t) := \{x \in \Omega; \ \phi(x) < 0\}$$

where ϕ is a level set function solving the Hamilton-Jacobi equation

$$\frac{\partial \phi}{\partial t} + V \nabla \phi \ = \ 0 \tag{6.11}$$

for a given *velocity*

$$V(x,t) \ := \ v(x,t) \frac{\nabla \phi}{|\nabla \phi|} \, .$$

The next result allows us to compute the Fréchet derivative of the functional E.

Lemma 6.3.1 ([9, Proposition 3.3]). *Let $V \in L^\infty(0,T; L^2(\mathbb{R}^n))^n$ with $\operatorname{div} V \in L^1(0,T; L^\infty(\mathbb{R}^n)) \cap L^\infty(0,T; L^2(\mathbb{R}^n))$. Further, let ϕ be a level set function satisfying (6.11) and $u(\cdot)$ and $D(\cdot)$ be defined as above. Then, the derivative of the error functional E is given by*

$$\frac{\partial E}{\partial t}(t) \ = \ \int_\Omega (u(x,t) - \bar{u}(x))\, h(x,t)\, dx \, , \ \ t > 0 \, ,$$

where $h(x,t) \ := \ (-1 + 2u(x,t)) \operatorname{div} V(x,t)$, for $x \in \Omega$, $t \ge 0$.

From Lemma 6.3.1 it is immediate to conclude that, in order to obtain a flow satisfying (6.10), one has to choose the velocity V such that

$$\operatorname{div} V = -F^*(Fu-g)\,(-1+2u),\ x \in \Omega\,.$$

Since one expects the corresponding sets $D(t)$ to satisfy $D(t) \subset \operatorname{int}(\Omega)$ for $t \geq 0$, it is convenient to impose the condition $\operatorname{div} V \equiv 0,\ \ x \in \mathbb{R}^n/\Omega$. Therefore, one can write

$$-\operatorname{div} V = P_\Omega\big(F^*(Fu-g)\big)(-1+2u)\,,\ x \in \Omega\,, \tag{6.12}$$

where $P_\Omega : L^2(\Omega) \to L^2(\mathbb{R}^n)$ is the extension operator defined by

$$(P_\Omega(v))(x) := \left\{ \begin{array}{r} v(x),\ x \in \Omega \\ 0,\ x \in \mathbb{R}^n/\Omega \end{array} \right.$$

It is worth noticing that a function V satisfying (6.12) always exists. Indeed, by choosing $\nabla\psi - V = 0$ in $\mathbb{R}^n$, with the decay condition $\psi(x) \to 0$ as $|x| \to \infty$, it becomes clear that ψ is the unique solution of the Poisson equation (in $\mathbb{R}^n$) with righthand side as in (6.12) and Dirichlet boundary condition. The velocity V is the gradient of ψ.

One should notice that the regularity assumptions of Lemma 6.3.1 can be verified if the adjoint operator F^* maps Y continuously to $L^\infty(\Omega)$. Indeed, in this case one can define

$$-\operatorname{div} V_n = h_n * (F^*(Fu-g)(-1+2u))\,,$$

where '$*$' denotes the Fourier convolution (with respect to x) and h_n is a sequence of smooth, nonnegative convolution kernels such that $h_n * v \to v$ for all $v \in L^\infty(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ as $n \to \infty$. If one chooses $V_n = \nabla\psi_n$, then ψ_n solves the Poisson equation with continuous righthand side and, from classical elliptic theory, it follows that $\psi_n \in C(0,T;C^2(\mathbb{R}^n))$. Consequently, $V_n = \nabla\psi_n \in C(0,T;C^1(\mathbb{R}^n))$ and

$$\operatorname{div} V_n \to \operatorname{div} V \quad \text{in} \quad L^\infty(\mathbb{R}^n \times [0,T]) \hookrightarrow L^1(0,T;L^\infty(\mathbb{R}^n))\,.$$

It is immediate to observe that (6.12) yields the estimate

$$\frac{\partial E}{\partial t}(t) = -\|Fu(t)-g\|^2\,,\ t > 0\,,$$

which, up to a constant, corresponds to the estimate (6.10) for the asymptotic regularization.

In the case of nonlinear operators F which are continuously Fréchet differentiable in $L^2(\Omega)$, the extension of formula (6.12) is immediate. In this case the velocity V should be chosen as the solution of

$$-\operatorname{div} V = P_\Omega\big(F'(u)^*(F(u)-g)\big)(-1+2u)\,,\ x \in \Omega\,,$$

where F' denotes the Fréchet derivative of F. Notice that, if F satisfies the so called *tangential cone condition* (see [84])

$$\|F(v) - F(\bar{u}) - F'(\bar{u})(v-\bar{u})\| \le \eta \|F(v) - F(\bar{u})\| \qquad (6.13)$$

with $\eta < \frac{1}{2}$ for all v in a neighborhood of the solution $\bar{u}$, then it follows

$$\frac{\partial E}{\partial t}(t) = -\langle F'(u)(\bar{u}-u), g - F(u)\rangle \le -(1-\eta)\,\|F(u)-g\|^2$$

and the method can be analyzed analogously as in the linear case (see Exercise 6.1).

In the sequel we devote our attention to the convergence issue of this level set method. The first result concerns the monotonicity of both the iteration error and the residual.

Lemma 6.3.2 ([9, Propositions 4.1, 4.3]). *If $t > 0$ is such that $\|Fu(t) - g^\varepsilon\| > \varepsilon$, then $dE^\varepsilon(t)/dt < 0$, i.e. the iteration error decreases. Furthermore, the function $t \mapsto \|Fu(t) - g^\varepsilon\|$ is monotonically decreasing. Moreover, in the particular case $\varepsilon = 0$, we have*

$$\int_0^\infty \|Fu(t) - F\bar{u}\|^2\, dt \ < \ \infty\,.$$

Using the results above, one can prove a first convergence result, concerning the exact data case.

Lemma 6.3.3 ([9, Proposition 4.4]). *Assume one has exact data, i.e. $\varepsilon = 0$. Further, let the velocity V be chosen according to (6.12). Then, $\|u(t) - \bar{u}\| \to 0$ as $t \to \infty$, where $\bar{u}$ solves (6.1).*

In order to present a convergence result for noisy data, we need to recall the *generalized discrepancy principle.* According to this stopping rule, the iteration should be stopped at the time $T = T(\delta, g^\varepsilon)$

when $\|Fu(T) - g^\varepsilon\| \leq \tau\varepsilon$ for the first time (here τ is a positive constant).

Lemma 6.3.4 ([9, Propositions 4.5, 4.6]). *Let $\varepsilon > 0$ and $\tau > 1$. The stopping rule $T(\varepsilon, g^\varepsilon)$ defined by the discrepancy principle is finite. Moreover, $\|u^\varepsilon(T(\varepsilon, g^\varepsilon)) - \bar{u}\| \to 0$ as $\varepsilon \to 0$.*

Lemma 6.3.4 means that $D^\varepsilon(T(\varepsilon, g^\varepsilon)) \to \bar{D}$ in the symmetric difference metric as $\varepsilon \to 0$, where $\bar{D}$ is the set corresponding to $\bar{u}$.

Remark 6.3.5. *Equation (6.12) is a Hamilton-Jacobi type equation of the form*

$$\frac{\partial \phi}{\partial t} + V|\nabla\phi| = 0\,.$$

with $V(x,t)$ given by

$$V \;=\; \operatorname{div}^{-1}\left(F'(u)^*(F(u) - g)\,(-1 + 2u)\right)\frac{1}{|\nabla\phi|}\,.$$

Compare with Remarks 6.2.1 and 6.4.4.

6.4 Level sets and constraint optimization

In this section the alternative level set approach for inverse obstacle problems, proposed by Leitão and Scherzer (see [59]) is presented. The level set methods are interpreted as constraint regularization methods based on the coupling of Tikhonov regularization and projection strategies.

6.4.1 Introduction

The general context is to solve the constraint ill-posed operator equation:

$$F(u) = g\,, \tag{6.14}$$

where u is in the admissible class

$$U \;:=\; \{u : u = P(\phi) \;\text{ and }\; \phi \in \mathcal{D}(P)\}\,.$$

The constraint equation can be formulated as an unconstrained equation

$$F(P(\phi)) = g\,. \tag{6.15}$$

Assuming that the operator equation is ill-posed it has to be regularized for a stable solution. Classical results on convergence and stability of regularization (see e.g. [23, 71, 72]) such as

1. existence of a regularized solution

2. stability of the regularized approximations

3. approximation properties of the regularized solutions

are applicable if P is either bounded and linear or nonlinear, continuous, and weakly closed.

In order to link constraint regularization methods and level sets, discontinuous operators P are required, and thus the classical framework of regularization theory is not applicable yet.

Tikhonov regularization for solving the unconstrained equation (6.14) consists in approximation the solution of (6.14) by the minimizer u_α of the functional

$$\|F(u) - g\|^2 + \alpha \|u - u_*\|^2\,.$$

If F is Fréchet differentiable, then

$$F'(u_\alpha)^*(F(u_\alpha) - y) + \alpha(u_\alpha - u_*) = 0\,, \tag{6.16}$$

where $F'(u_\alpha)^*$ denotes the adjoint of the derivative of F at u_α. Notice that (6.16) is the optimality condition for a minimizer of the Tikhonov functional. Using the formal setting $\Delta t := 1/\alpha$, $u(\Delta t) := u_\alpha$, and $u(0) := u_*$ one finds

$$F'(u(\Delta t))^*(F(u(\Delta t)) - g) + \frac{u(\Delta t) - u(0)}{\Delta t} = 0\,.$$

Thus $u_\alpha = u(\Delta t)$ can be considered as the solution of one implicit time step with step-length $\Delta t = \frac{1}{\alpha}$ for solving

$$\frac{\partial u}{\partial t} = -F'(u)^*(F(u) - g) \tag{6.17}$$

and one ends up with the *inverse scale-space method* (see, e.g., [33, 86]). Note that the inverse scale-space method corresponds to the *asymptotic regularization method* as introduced by Tautenhahn [89, 90]. The terminology "inverse scale-space" is motivated from *scale-space* theory in *computer vision*: images contain structures at a variety of scales. Any feature can optimally be recognized at a particular scale. If the optimal scale is not available a-priori, it is desirable to have an image representation at multiple scales. For more background on the topic of scale-space theory we refer to [62, 75, 51].

A consequence of the approach presented in this section is that the inverse scale-space method for the constrained inverse problem (6.15) with appropriate P is a *level set method.* In this notes, however, we will not go any further into this discussion.

6.4.2 Derivation of the Level Set Method

In the sequel we consider the constraint optimization problem of solving (6.14) on the set of piecewise constant functions which attain two values, which we fix for the sake of simplicity of presentation to 0 and 1. Let $\Omega \subseteq \mathbb{R}^n$ $(n = 1, 2)$ be bounded with boundary $\partial\Omega$ Lipschitz. Set

$$\mathcal{P} := \{u : u = \chi_{\tilde{\Omega}} : \tilde{\Omega} \subseteq \Omega\} \cap L^2(\Omega)\,,$$

then the unconstrained inverse problem consists in solving (6.15) with

$$\begin{aligned} P : H^1(\Omega) &\to \mathcal{P}\,. \\ \phi &\mapsto \tfrac{1}{2} + \tfrac{1}{2}\mathrm{sgn}(\phi) =: \tfrac{1}{2} + \tfrac{1}{2}\left\{ \begin{array}{r} 1 \text{ for } \phi \geq 0 \\ -1 \text{ for } \phi < 0 \end{array} \right. \end{aligned}$$

Moreover, let for the sake of simplicity of presentation,

$$F : L^2(\Omega) \to L^2(\Omega)$$

be Fréchet-differentiable. It is as well possible to consider the operator F in various Hilbert space settings such as for instance $F : H^1(\Omega) \to L^2(\partial\Omega)$. Since it does not make any methodological differences we shall concentrate on an operator on $L^2(\Omega)$. Also the space $H^1(\Omega)$ is chosen more or less arbitrarily; these spaces were selected in such a way that the typical distance functions for smooth domains are contained in $H^1(\Omega)$.

Tikhonov regularization for this problem consists in minimizing the functional

$$\int_\Omega (F(P(\phi)) - g)^2 + \alpha \int_\Omega \left((\phi - \phi_*)^2 + |\nabla(\phi - \phi_*)|^2\right) . \qquad (6.18)$$

Since the functional does not attain a minimum, the "minimizer" ϕ_α is considered as

$$\phi_\alpha := \lim_{\epsilon \to 0+} \phi_{\epsilon,\alpha} \, ,$$

where $\phi_{\epsilon,\alpha}$ minimizes the functional

$$\int_\Omega (F(P_\epsilon(\phi)) - g)^2 + \alpha \int_\Omega \left((\phi - \phi_*)^2 + |\nabla(\phi - \phi_*)|^2\right) \ . \qquad (6.19)$$

The operators

$$P_\epsilon(t) := \begin{cases} 0 & \text{for} \quad t < -\epsilon \, , \\ 1 + \frac{t}{\epsilon} & \text{for} \quad t \in [-\epsilon, 0] \ , \\ 1 & \text{for} \quad t > 0 \, , \end{cases}$$

are used for approximating P as $\epsilon \to 0^+$. In this case we have

$$P'(t) = \lim_{\epsilon \to 0+} P'_\epsilon(t) = \delta(t) \, .$$

Here and in the following $\delta(t)$ denotes the one-dimensional δ-distribution. Moreover, we denote

$$u_\alpha := \lim_{\epsilon \to 0+} P_\epsilon(\phi_{\alpha,\epsilon}) \, .$$

Notice that $u_\alpha = P(\phi_\alpha)$ is *not* required. The proposed methodology to define generalized solutions $u_\alpha = \lim_{\epsilon \to 0+} P(\phi_{\epsilon,\alpha})$ is a standard way in *phase transitions*.

In the following an optimality condition for a minimizer of (6.18) is derived, which is considered the limit $\epsilon \to 0+$ of the minimizers of the functionals (6.19). For this purpose it is convenient to recall some basic results from *Morse theory* of surfaces. The particular results are collected from [24]. It is worth emphasizing that, here, the Morse theory is only applied to compact, smooth subset of $\mathbb{R}^2$, which of course can be considered as surfaces.

Lemma 6.4.1 ([59, Proposition 2.1]). *Let ϕ be a smooth function on a compact smooth surface M, and $\phi^{-1}[a,b] \subseteq M$ contain no critical point of ϕ. Then,*

1. *the level sets $\phi^{-1}(b)$ and $\phi^{-1}(a)$ are diffeomorphic (in particular they consist of the same number of smooth circles diffeomorphic to a standard circle). In particular the Hausdorff measure of $\phi^{-1}(t), t \in [a,b]$ changes continuously.*
2. *Moreover, for any $\rho \in [a,b]$, $\phi^{-1}(\rho)$ is a smooth compact 1-manifold. In particular $\phi^{-1}(\rho)$ can be parameterized by finitely many disjoint curves.*

The following lemma is central to derive the optimality condition for a minimizer of (6.18).

Lemma 6.4.2 ([59, Lemma 2.2]). *Let ϕ be a smooth function, having no critical points in a compact neighborhood M of the level set $\phi^{-1}(0)$. Then,*

$$\lim_{\epsilon \to 0+} P'_\epsilon(\phi) = \frac{1}{|\nabla \phi|}\delta(\phi)\,.$$

where $\delta(\phi)$ is the one-dimensional δ-distribution centered at the level line in normal direction.

From the definition of a minimizer of (6.19) it follows that for all $h \in H^1(\Omega)$

$$\int_\Omega (F(u_{\epsilon,\alpha}) - g) F'(u_{\epsilon,\alpha}) P'_\epsilon(\phi_{\epsilon,\alpha}) h \\ + \alpha \int_\Omega ((\phi_{\epsilon,\alpha} - \phi_*)h + \nabla(\phi_{\epsilon,\alpha} - \phi_*)\nabla h) \;=\; 0\,.$$

We denote by $F'(u)^*, P'_\epsilon(\phi)^*$ the L^2-adjoints of $F'(u)$, $P'_\epsilon(\phi)$ respectively, i.e.,

$$\int_\Omega w(F'(u)v) = \int_\Omega (F'(u)^* w)v \;\text{ and }\; \int_\Omega w(P'_\epsilon(\phi)v) = \int_\Omega (P'_\epsilon(\phi)^* w)v\,,$$

for all test functions $v, w \in L^2(\Omega)$. Since $P'_\epsilon(\phi)$ is self-adjoint, i.e., $P'_\epsilon(\phi)^* = P'_\epsilon(\phi)$, it follows that

$$\begin{aligned} P'_\epsilon(\phi_{\epsilon,\alpha})F'(u_{\epsilon,\alpha})^*(F(u_{\epsilon,\alpha}) - g) + \alpha(I - \Delta)(\phi_{\epsilon,\alpha} - \phi_*) &= 0 \text{ on } \Omega\,, \\ \alpha\frac{\partial(\phi_{\epsilon,\alpha} - \phi_*)}{\partial n} &= 0 \text{ at } \partial\Omega. \end{aligned}$$

Thus, $u_\alpha = \lim_{\epsilon\to 0+} u_{\epsilon,\alpha}$ and $\phi_\alpha = \lim_{\epsilon\to 0+} \phi_{\epsilon,\alpha}$ satisfy

$$\delta(\phi_\alpha)\frac{F'(u_\alpha)^*(F(u_\alpha) - g)}{|\nabla\phi_\alpha|} + \alpha(I - \Delta)(\phi_\alpha - \phi_*) = 0\,. \qquad (6.20)$$

For the sake of simplicity of presentation the operator F is assumed to be of such quality that $F'(u)^*(F(u)-g)$ is continuous on Ω. Note that in general this may not be the case since $F'(u)^*(F(u) - g) \in H^1(\Omega)$. Therefore, it follows from (6.20) that

$$(I - \Delta)^{-1}\left(\delta(\phi_\alpha)\frac{F'(u_\alpha)^*(F(u_\alpha) - g)}{|\nabla\phi_\alpha|}\right) + \alpha(\phi_\alpha - \phi_*) = 0\,.$$

Set $\alpha := \frac{1}{\Delta t}$ and set $\phi_\alpha := \phi(t)$, $\phi_* := \phi(0)$ and accordingly $u(t) := P(\phi(t))$. Then, by taking the formal limit $\Delta t \to 0+$ the asymptotic regularization method follows:

$$\frac{\partial\phi}{\partial t} = -(I - \Delta)^{-1}\left(\delta(\phi(t))\frac{F'(u(t))^*(F(u(t)) - g)}{|\nabla\phi(t)|}\right)\,. \qquad (6.21)$$

The right hand side v of (6.21) solves the equation

$$\begin{aligned} (I - \Delta)v &= -\delta(\phi(t))\frac{F'(u(t))^*(F(u(t)) - g)}{|\nabla\phi(t)|} \text{ on } \Omega \\ \frac{\partial v}{\partial n} &= 0 \text{ at } \partial\Omega\,. \end{aligned} \qquad (6.22)$$

Using potential theory (see, e.g., [22, 55]), a solution v_1 of the problem

$$\Delta v_1(t) = \delta(\phi(t))\frac{F'(u(t))^*(F(u(t)) - g)}{|\nabla\phi(t)|}$$

with homogeneous (Dirichlet) boundary conditions is given by the *single layer potential*

$$v_1(x) = -\int_{\phi(t)^{-1}(0)} \frac{F'(u(t))^*(F(u(t)) - g)(z)\gamma(x,z)}{|\nabla\phi(t)(z)|}\,dz\,,$$

where

$$\gamma(x,y) = \begin{cases} \frac{1}{2\pi}\ln\left(\frac{1}{|x-y|}\right) & \text{in } \mathbb{R}^2\,, \\ \frac{1}{4\pi}\frac{1}{|x-y|} & \text{in } \mathbb{R}^3 \end{cases} \tag{6.23}$$

is the *single layer potential.* Then, $v = v_1 + v_2$ solves (6.22), where v_2 solves

$$\begin{aligned} v_2 - \Delta v_2 &= -v_1 \text{ on } \Omega \\ \frac{\partial v_2}{\partial n} &= -\frac{\partial v_1}{\partial n} \text{ at } \partial\Omega\,. \end{aligned}$$

Equation (6.21) represents a *level set method* describing the evolution of the level set function ϕ. The zero level set of ϕ, i.e., the set $\{\phi = 0\}$, describes the boundary of the inclusions to be recovered.

Remark 6.4.3. *An adequate approximation of P is central in this considerations. The family of functions*

$$Q_\epsilon(t) := \begin{cases} 0 & \textit{for} \quad t < -\epsilon\,, \\ \frac{t+\epsilon}{2\epsilon} & \textit{for} \quad t \in [-\epsilon, \epsilon]\,, \\ 1 & \textit{for} \quad t > \epsilon\,, \end{cases}$$

approximates the δ-distribution too. Since the point-wise limit of Q_ϵ is

$$P(t) := \begin{cases} 0 & \textit{for } t < 0\,, \\ \frac{1}{2} & \textit{for } t = 0\,, \\ 1 & \textit{for } t > 0\,, \end{cases}$$

which is not in $\mathcal{P}$ if the n-dimensional Lebesgue measure of $\phi^{-1}(0)$ is greater than zero. This would not be appropriate for our problem setting.

In this section we have elaborated on the interaction between constraint regularization methods and level set methods. We observed that the level set method in [59] can be considered as an inverse scale-space method, respectively asymptotic regularization method. In contrast to standard results on asymptotic regularization methods and inverse scale-space methods (see [89, 90, 33]), here the situation is more involved, since the regularizer of the underlying regularization functional (6.18) is considered as approximation of the minimizers of the functional (6.19).

One of the most significant advantages of level set methods is that the topology of the zero–level set may change over time. This situation has not been covered by the present derivation of level set methods, where the authors essentially relied on Lemmas 6.4.1 and 6.4.2. In case a topology change occurs the Morse index of the level set function ϕ changes and Lemma 6.4.1 (and consequently Lemma 6.4.2) are not applicable. Moreover, in this case the single layer potential representations (6.23) are no longer valid (cf., e.g., [18, 55]), since the topology changes results in domain with cusps. The effect of topology changes on the level set methods are status of ongoing research.

Remark 6.4.4. *Equation (6.21) is a Hamilton-Jacobi type equation of the form*

$$\frac{\partial \phi}{\partial t} + V|\nabla \phi| = 0\,.$$

with $V(x,t)$ given by

$$V \;=\; (I-\Delta)^{-1}\left(\delta(\phi)\frac{F'(u(t))^*(F(u(t))-g)}{|\nabla\phi(t)|}\right)\frac{1}{|\nabla\phi|}\,.$$

Compare with Remarks 6.2.1 and 6.3.5.

The numerical solution of (6.21) is similar to the implementation of well-established level set methods, like the ones considered in Sections 6.2 and 6.3. The differential equation

$$\frac{\partial \phi}{\partial t} = F'(u(t))^*(F(u(t))-g)|\nabla\phi(t)|$$

is solved explicit in time, which results in

$$\frac{\phi(t+\Delta t)-\phi(t)}{\Delta t} = F'(u(t))^*(F(u(t))-g)|\nabla\phi(t)|\,.$$

After several numerical time-steps the iterates are *updated*. In the present level-set approach such an update is inherent, since in each step the data is normalized by the operator $(I-\Delta)^{-1}$.

6.4.3 Relation to Shape Optimization

In the sequel we show that the term

$$\delta(\phi)\frac{F'(u)^*(F(u)-g)}{|\nabla\phi|}$$

is the steepest descent direction of the functional $\|F(u) - g\|^2$ with respect to the *shape* of the level set $\phi^{-1}(0)$.

It is much more illustrative to show this relation exemplary. To this end we consider the *inverse potential problem* of recovery of an object $D \subseteq \mathbb{R}^2$ in

$$\Delta v = \chi(D) \text{ in } \Omega \text{ with } v = 0 \text{ on } \partial\Omega$$

(see Section 6.1). In this context

$$\begin{array}{rcl} F : L^2(\Omega) & \to & L^2(\Omega) \; . \\ f & \mapsto & \Delta^{-1} f \text{ with homogeneous Dirichlet data} \end{array}$$

The numerical recovery of shape of the inclusion D from Neumann boundary measurements was considered in [43]. For the sake of simplicity of presentation, here we are interested in the shape derivative of F, while Hettlich and Rundell considered the operator $T \circ F$, where T is the Neumann trace operator. Since T is linear the shape derivative of $T \circ F$ is completely determined by the shape derivative of F, and thus we do not impose any restriction on the consideration by considering the simpler problem.

The operator F is linear and thus the Gateaux-derivative of F at u in direction h satisfies $F'(u)h = F(h)$. Thus the *level set derivative* is given by

$$v := F'(u)P'(\phi)h = F(P'(\phi)h) = \Delta^{-1}\left(\delta(\phi)\frac{h}{|\nabla\phi|}\right) . \tag{6.24}$$

Let v_1 be the single layer potential according to h on $\phi^{-1}(0)$, i.e.,

$$v_1(x) = -\int_{\phi^{-1}(0)} \frac{1}{2\pi} \ln \frac{1}{|x-y|} \frac{h}{|\nabla\phi|}(y)\, dy \, .$$

This function satisfies

$$\Delta v_1 = \delta(\phi)\frac{h}{|\nabla\phi|} \text{ on } \Omega \, .$$

Let v_2 be the solution of

$$\Delta v_2 = 0 \text{ on } \Omega \quad \text{and} \quad v_1 = -v_2 \text{ at } \partial\Omega \, .$$

Then $v = v_1 + v_2$ solves

$$\Delta v = \delta(\phi) \frac{h}{|\nabla \phi|} \text{ on } \Omega \quad \text{and} \quad v = 0 \text{ at } \partial\Omega\,.$$

Moreover, the single layer potential satisfies on the zero level set

$$\begin{array}{rcl} \left(\frac{\partial v_1}{\partial n}\right)_+ - \left(\frac{\partial v_1}{\partial n}\right)_- & = & \frac{h}{|\nabla\phi|}\,, \\ (v_1)_+ & = & (v_1)_-\,. \end{array}$$

Here $(\cdot)_+$, $(\cdot)_-$ denote the limits from outside, inside of the domain bounded by the zero level curves, respectively.

Recall that h is considered a perturbation of the level set *function.* A change in the level set function implies a change in the zero level set, which eventually turns out to be the shape derivative.

To make this concrete, let s_{th} the parameterizations of $(\phi + th)^{-1}(0)$, i.e.,
$(\phi + th)(s_{th}) = 0$. We make a Taylor Ansatz with respect to the parameterization

$$s_{th} \;=\; s + t\tilde{h} + O(t^2)\,, \tag{6.25}$$

and a series expansion for ϕ and h, which gives

$$0 \;=\; (\phi + th)(s_{th}) \;=\; t\nabla\phi\tilde{h} + th(s) + O(t^2)\;.$$

This shows that on the zero level set we have

$$\frac{h}{|\nabla\phi|} \;=\; -\frac{\nabla\phi}{|\nabla\phi|}\cdot\tilde{h} \;=\; n\cdot\tilde{h}\,.$$

Thus, v satisfies the differential equation

$$\left\{ \begin{array}{rcl} \Delta v & = & 0 \text{ on } \Omega\backslash\phi^{-1}(0)\,, \\ v & = & 0 \text{ on } \partial\Omega\,; \end{array} \right.$$

$$\left(\frac{\partial v}{\partial n}\right)_+ - \left(\frac{\partial v}{\partial n}\right)_- = \tilde{h}\cdot n \text{ on } \phi^{-1}(0)\,,$$

$$(v)_+ = (v)_- \text{ on } \phi^{-1}(0)\,.$$

This is the shape derivative $F'(D)(\tilde{h})$ of F at $D = \{x : P(\phi) > 0\}$ in direction $\tilde{h}$ as calculated by Hettlich and Rundell in [43]. The calculations above show the level set derivative $v := F'(u)P'(\phi)h$ can be computed from the shape derivative. Now, we point out that the converse is evenly true. This is nontrivial since the arguments $\tilde{h}$ appearing in the shape derivative are multidimensional functions, while the argument h in the level set derivative is one-dimensional.

Let $\tilde{h}$ be expressed in terms of the local coordinate system n and τ, where n, τ are the normal, respectively tangential vectors on the zero level set, i.e.,

$$\tilde{h} = hn + h_\tau \tau\,.$$

The shape derivative is independent of the tangential component, which in particular implies that the shape derivative gradient descent deforms the shapes in normal direction to the level curve. Thus, from (6.24) we find that

$$F'(D)(\tilde{h}) = F'(D)(hn) = F'(P\phi)h\,. \tag{6.26}$$

Summarizing, by (6.26) the level set derivative $F'(u)P'(\phi)h = F(P'(\phi))h$ is uniquely determined from the shape derivative and vice versa. Therefore, we see that the level set derivative moves the zero level set in direction of the shape derivative.

6.5 Applications

In this section we consider two distinct applications of level set methods to parameter identification problems modeled by partial differential equations. The first problem is the *inverse potential problem* (for the operator F_2) introduced in Section 6.1. The second application is related to the *inverse doping profile problem.* It is a technological application and concerns the identification of doping profiles in semiconductor devices (see [10, 11]).

6.5.1 The inverse potential problem

We consider the inverse potential problem of recovering the shape of a domain D using the knowledge of its (constant) density and the measurements of the Cauchy data of the corresponding potential on the boundary of a fixed Lipschitz domain $\Omega \subset \mathbb{R}^2$, which contains

$\overline{D}$. This is the same problem as considered by Hettlich and Rundell [43] (see Section 6.1), which used iterative methods for recovering a single star-shaped object.

To achieve an analogous problem, a certain definition of the operator F is necessary:

$$\begin{aligned} F : L^2(\Omega) \to &\quad L^2(\partial\Omega) \\ \chi_D \to &\quad F(\chi_D) \end{aligned}$$

This is possible, because we consider only characteristic functions χ_D. The $L^2(\Omega)$-norm is then equivalent to the $L^1(\Omega)$-norm of χ_D. Therefore the necessary properties are retained.

The problem introduced above can mathematically be described as follows:

$$\Delta u = \chi_D \, , \text{ in } \Omega \, ; \quad u|_{\partial\Omega} = 0 \, , \tag{6.27}$$

where χ_D is the characteristic function of the domain $D \subset \Omega$, which has to be reconstructed. Since $\chi_D \in L^2(\Omega)$, the Dirichlet boundary value problem in (6.27) has a unique solution, the potential $u \in H^2(\Omega) \cap H_0^1(\Omega)$. Here $H_0^1(\Omega)$ is defined as the closure with respect to $H^1(\Omega)$ of functions in $C^\infty(\Omega)$ with compact support in Ω.

The inverse problem we are concerned with, consists in determining the shape of D from measurements of the Neumann trace of u at $\partial\Omega$, i.e. from $[\partial u/\partial\nu]_{\partial\Omega}$, where ν represents the outer normal vector to $\partial\Omega$.

Notice that this problem can be considered in the framework of an inverse problem for the *Dirichlet to Neumann map.* For given $h \in L^2(\Omega)$, the Dirichlet to Neumann operator maps a Dirichlet boundary data onto the Neumann trace of the potential, i.e., $\Lambda : H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega)$, $\Lambda(\varphi) := [\partial\tilde{u}/\partial\nu]_{\partial\Omega}$, where $\tilde{u}$ solves

$$\Delta\tilde{u} = h \, , \text{ in } \Omega \, ; \quad \tilde{u}|_{\partial\Omega} = \varphi \, .$$

The inverse problem for the Λ operator consists in determining the unknown parameter (i.e., the function h) from different pairs of Dirichlet, Neumann boundary data. The general case with $h \in L^2(\Omega)$ has already been considered by many authors, among them we mention [13, 80], which introduced numerical methods based on Tikhonov regularization, and [43] with iterative regularization methods.

```
1. Evaluate the residual $r_k := F(P_\epsilon(\phi_k)) - y^\delta = \frac{\partial u_k}{\partial \nu} - y^\delta$,
   where $u_k$ solves
       $\Delta u_k = P_\epsilon(\phi_k)$, in $\Omega$;    $u_k|_{\partial\Omega} = 0$.
2. Evaluate $v_k := F'(P_\epsilon(\phi_k))^*(r_k) \in L^2(\Omega)$, solving
       $\Delta v_k = 0$, in $\Omega$;  $v_k|_{\partial\Omega} = r_k$.
3. Evaluate $w_k \in H^1(\Omega)$, satisfying
       $(I - \Delta) w_k = -P'_\epsilon(\phi_k)\, v_k$, in $\Omega$;
       $\frac{\partial w_k}{\partial \nu} = 0$, at $\partial\Omega$.
4. Update the level set function $\phi_{k+1} = \phi_k + \frac{1}{\alpha}\, w_k$.
```

Table 6.1: Algorithm for one iterative step of the level set method (cf. [59]) for the inverse potential problem.

Hettlich and Rundell [43] observe that, in the particular case $h = \chi_D$, one pair of measurement data of Dirichlet–Neumann data furnishes as many information as the full Dirichlet–Neumann operator, i.e., it is sufficient to consider only one pair of Cauchy data for the inverse problem. Therefore, no further information on D can be gained by using various pairs of Dirichlet–Neumann data, since we can always reduce the reconstruction problem to the homogeneous Dirichlet case.

For the particular case $h = \chi_D$, it has been observed by Hettlich and Rundell [43] that the Cauchy data may not furnish enough information to reconstruct the boundary of D, e.g., if D is not simply connected. On the other hand, Isakov observed in [49] that star like domains D are uniquely determined by their potentials.

The inverse potential problem is discussed within the general framework introduced in Section 6.4. In particular, we allow domains, that consists of a number of connected inclusions. For this general

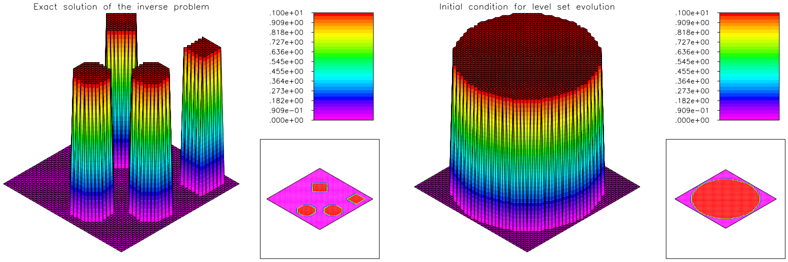

Figure 6.2: The picture on the left hand side shows the coefficient to be reconstructed. On the other picture, the initial condition for the level set method.

class we have not unique identifiability and we restrict attention to "minimum-norm solutions". Recall that in this case a minimum-norm-solution is a level set function ϕ, where $P(\phi)$ determines the inclusion. A minimum norm solution satisfies that it minimizes the functional $\rho(z, \phi)$ in the class of level set functions such that the according Neumann boundary values $\frac{\partial u}{\partial \nu}$ fit the data y^δ.

In the following we describe the level set regularization algorithm of [59, 27]. The complexity of the algorithm is as follows: at each iteration of the level set method, three elliptic boundary value problems are solved (two of Dirichlet type and one of Neumann type). The iterative procedure corresponding to the evolution equation (6.21) is outlined in Table 6.1.

The algorithm can be implemented using finite element codes (as we did) or finite difference methods for the solution of partial differential equations.

In this experiment we consider the inverse problem of reconstructing the right hand side χ_D in (6.27) from the knowledge of a single pair of boundary data $(u, \Lambda u) = (0, y^\delta)$ at $\partial\Omega$, where $\Omega = (0,1)^2 \subset \mathbb{R}^2$. $\chi_D \in L^2(\Omega)$ is the characteristic function as represented in Figure 6.2.

The overdetermined boundary measurement data y^δ for solving the inverse problem, is obtained by solving the elliptic boundary value problem (of Dirichlet type) in (6.27). Notice that χ_D corresponds to the characteristic function of a not-connected proper subset of Ω.

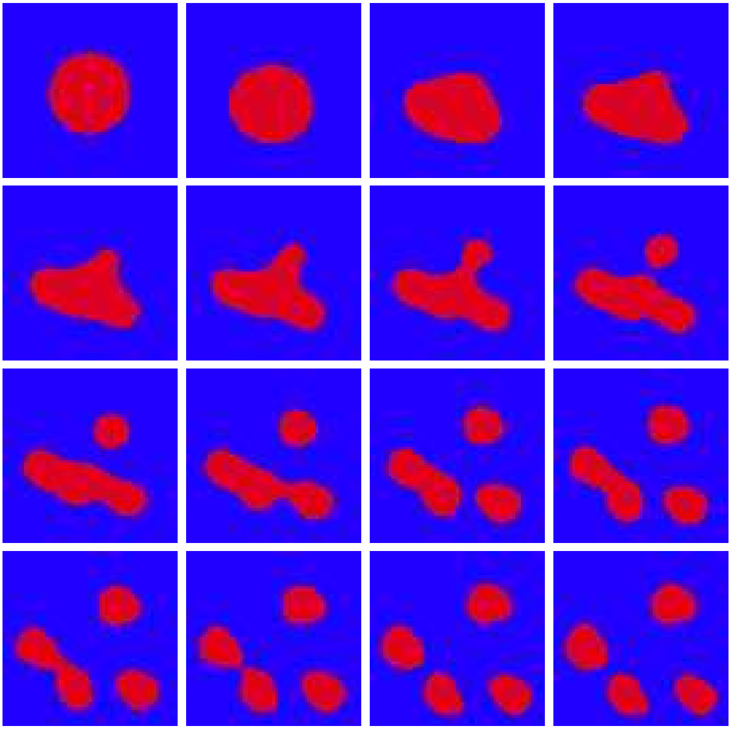

Figure 6.3: Evolution of the level set method for the inverse potential problem. Pictures after $10, 100, 1000, 2000,$ $3000, 4000, 5000, 5500,$ $6000, 7000, 8000, 9000,$ $10000, 11000, 15000, 20000$ iterative steps.

The initial condition for the level set function is shown in Figure 6.2.

For this experiment we used the operator P_ϵ defined in Section 6.4 with $\epsilon = 1/8$. This seams to be compatible with the size of our mesh, since the diameter of the triangles in the uniform grid (used in the finite element method) is approximately $\sqrt{2}/32$.

In Figure 6.3 we present the evolution of the level set method for given exact data for the first 20000 iterative steps. As one can see in this figure, the original level set splits into two connected components after approximately 5000 steps and after 11000 steps the four connected components of the solution can be recognized. After 15000

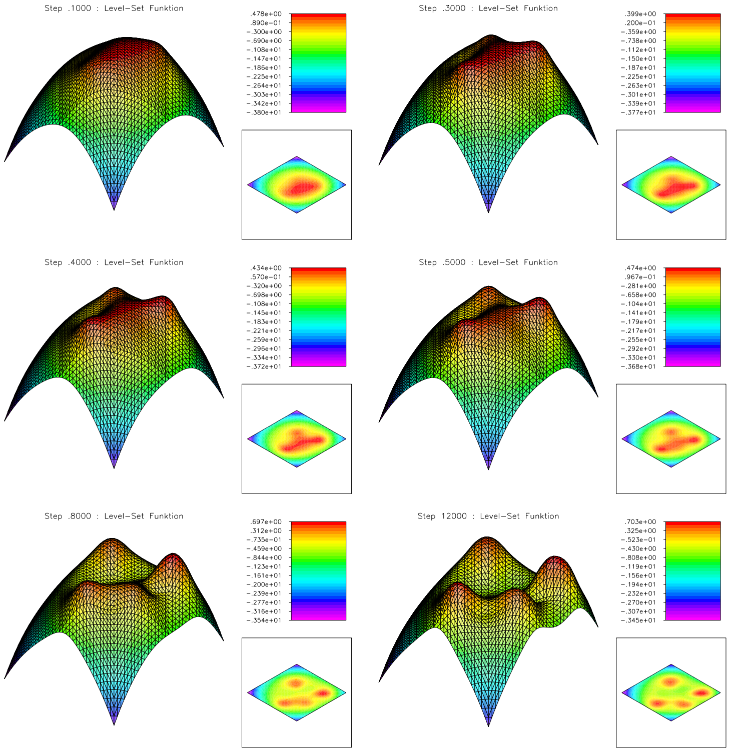

Figure 6.4: Evolution of the level set function for the inverse potential problem.

iterations the level set function still changes, but very slowly. We performed similar tests for different initial conditions and observed that, after 1000 iterations, the corresponding pictures look very much alike. In Figure 6.4 the corresponding evolution of the level set function is shown.

6.5.2 Identification of doping profiles

We consider the problem of identifying discontinuous doping profiles in semiconductor devices, where the data is obtained by a voltage-current map. The underlying mathematical model is the unipolar system, a system derived from the drift diffusion equations. The related inverse problem corresponds to an inverse conductivity problem with partial data.

The *drift diffusion* equations are the most widely used model to describe semiconductor devices. The *basic semiconductor device equations* where first presented, in the level of completeness discussed here, by W.R. van Roosbroeck in [93]. Since then it has been subject of intensive mathematical and numerical investigation. Recent detailed expositions of the subject of modeling, analysis and simulation of semiconductor equations can be found, e.g., in [69, 70].

The stationary drift diffusion equations consist of the Poisson equation (6.28a) for the electrostatic potential V and the continuity equations (6.28b) and (6.28c) for the electron density n and the hole density p respectively (notice that $-\nabla V$ is the electric field, while n and p are the concentration of free carriers of negative charge and positive charge respectively).

$$\begin{aligned}
\mathrm{div}(\epsilon\nabla V) &= q(n-p-C) \text{ in } \Omega && (6.28a)\\
\mathrm{div}(D_n\nabla n - \mu_n n\nabla V) &= R \qquad\qquad \text{in } \Omega && (6.28b)\\
\mathrm{div}(D_p\nabla p - \mu_p p\nabla V) &= R \qquad\qquad \text{in } \Omega. && (6.28c)
\end{aligned}$$

A word on notation: The domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ represents the semiconductor device. ϵ, q, μ_n, μ_p, D_n, D_p are physical constants. $R = R(n, p, x)$ denotes the recombination-generation rate, which is typically a function of the type: $R = \mathcal{R}(n, p, x)(np - n_i^2)$, where n_i denotes the intrinsic density. The function $C = C(x)$ denotes the doping concentration, which is produced by diffusion of different materials into the silicon crystal and by implantation with an ion beam.

In many technological applications, the *doping profile* C is the parameter which has to be identified. After the manufacturing process of the semiconductor device, it is necessary to test whether the doping has been correctly implanted. The inverse problem we are concerned with is related to a non destructive identification procedure, based on experiments modeled by the *voltage to current* operator.

For our numerical experiment, we consider a very simple semiconductor device, namely a P-N diode (see Figure 6.5). The boundary of

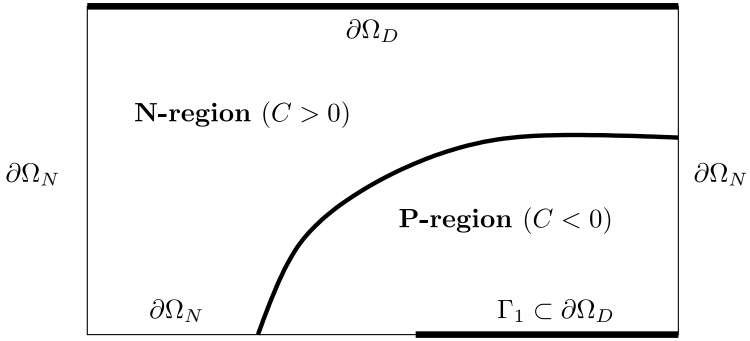

Figure 6.5: The domain $\Omega \subset \mathbb{R}^2$ represents a p-n diode. The P-region corresponds to the subregion of Ω, where the pre-concentration of negative ions predominate (i.e., $C < 0$). The N-region is defined analogously. The curve between these regions is called *p-n junction.*

Ω is assumed to be divided in two nonempty parts: $\partial\Omega = \partial\Omega_N \cup \partial\Omega_D$. The semiconductor contacts correspond to $\partial\Omega_D$, the part of the boundary where Dirichlet boundary conditions for system (6.28) are prescribed. The Neumann part of the boundary $\partial\Omega_N = \partial\Omega - \partial\Omega_D$ models insulating or artificial surfaces. Therefore, a zero current flow and a zero electric field in the normal direction are prescribed, i.e. homogeneous boundary conditions, in terms of the current densities[1] J_n and J_p. Therefore, the boundary conditions for system (6.28) are

$$\begin{aligned} V &= V_D(x) := U(x) + U_T \ln(n_D(x)/n_i) && \text{on } \partial\Omega_\mathrm{D} \,(6.28\mathrm{d})\\ n &= n_D(x) := \tfrac{1}{2}\big(C(x) + \sqrt{C(x)^2 + 4n_i^2}\big) && \text{on } \partial\Omega_\mathrm{D} \,(6.28\mathrm{e})\\ p &= p_D(x) := \tfrac{1}{2}\big(-C(x) + \sqrt{C(x)^2 + 4n_i^2}\big) && \text{on } \partial\Omega_\mathrm{D} \,(6.28\mathrm{f})\\ \nabla V \cdot \nu &= J_n \cdot \nu \;=\; J_p \cdot \nu \;=\; 0 && \text{on } \partial\Omega_\mathrm{N} \,(6.28\mathrm{g}) \end{aligned}$$

(the constant U_T denotes the thermal voltage).

[1] The densities of the electron and hole current J_n and J_p satisfy the current relations:

$$J_n = q(D_n \nabla n - \mu_n n \nabla V)\,, \quad J_p = q(-D_p \nabla p - \mu_p p \nabla V)\,, \quad \text{in } \Omega\,.$$

Next we define the *voltage-current* (V-C) map:

$$\begin{array}{rcl} \Sigma_C : H^{3/2}(\partial\Omega_D) & \to & H^{1/2}(\Gamma_1) \\ U & \mapsto & (J_n + J_p) \cdot \nu|_{\Gamma_1} \, , \end{array}$$

where $\Gamma_1 \subset \partial\Omega_D$ corresponds to the part of the boundary (a contact) where measurements are taken (see Figure 6.5). Notice that, due to the nature of the physical problem related to the semiconductor modeling, we can only consider as *inputs* for the DtN map functions of the type: $\{U \in H^{3/2}(\partial\Omega_D);\ U|_{\Gamma_1} = 0\}$. In practical applications, the function $U \in H^{3/2}(\partial\Omega_D)$ modeling the voltage input in (6.28) is piecewise constant in the contacts. The map Σ_C takes the applied voltage U into the corresponding current density.

For this application, instead of using the drift diffusion equations, we shall consider a simpler model, the so called *linearized unipolar case*

$$\left\{ \begin{array}{rll} \lambda^2 \Delta V^0 = e^{V^0} - C & \text{in } \Omega \\ V^0 = V_{\text{bi}} & \text{on } \Omega_D \\ \nabla V^0 \cdot \nu = 0 & \text{on } \Omega_N \end{array} \right. \qquad \left\{ \begin{array}{rl} \operatorname{div}(e^{V^0} \nabla u) = 0 & \text{in } \Omega \\ u = U & \text{on } \Omega_D \\ J_n \cdot \nu = 0 & \text{on } \Omega_N \end{array} \right. \tag{6.29}$$

The linearized unipolar case (close to equilibrium) corresponds to the model obtained from the drift diffusion equations (6.28) by linearizing the V-C map at $U = 0$. This simplification is motivated by the fact that, due to hysteresis effects for large applied voltage, the V-C map can only be defined in a neighborhood of $U = 0$. Furthermore, the following assumptions are taken into account in the derivation of (6.29):

A1) The concentration of holes satisfy $p = 0$;

A2) No recombination-generation rate is present, i.e. $\mathcal{R} = 0$; (for details see, e.g., [10, 11, 70]).

The inverse problem of identifying the doping profile in the linearized unipolar model (6.29) corresponds to the identification of $C(x)$ from the map

$$\begin{array}{rrcl} F : & H^2(\Omega) & \to & \mathcal{L}(H^{3/2}(\Omega_D); H^{1/2}(\Gamma_1)) \\ & C & \mapsto & \Lambda_C \end{array}$$

where Λ_C is the map that takes U into $(J_n \cdot \nu)|_{\Gamma_1}$, by solving the decoupled system (6.29). Notice that Λ_C derives from $\Sigma_C'(0)$ if we take into account the simplifying assumptions *A1)* and *A2)*.

1. Define $\gamma := e^{V^0}$ and identify γ from the DtN map $\Lambda_\gamma : U \mapsto \gamma u_\nu|_{\Gamma_1}$, where u solves

$$\mathrm{div}(\gamma \nabla u) = 0 \text{ in } \Omega\,, \ \ u = U \text{ on } \partial\Omega_D\,, \ \ u_\nu = 0 \text{ on } \partial\Omega_N;$$

2. Obtain the doping profile from:

$$C(x) = \gamma(x) - \lambda^2 \Delta(\ln \gamma(x))\,.$$

Table 6.2: Formulation of the inverse doping profile problem in the linearized unipolar case (close to equilibrium).

Since $V(x)$ is known at $\partial\Omega_D$, the current data $J_n \cdot \nu = \mu_n e^{V^0} u_\nu$ at Γ_1 (output) can be directly replaced by the Neumann data u_ν. Therefore, the inverse problem can be splited in two distinct steps, as shown in Table 6.2.

For our numerical implementation the level set method we consider the situation where only a single measurement of the DtN map is available, i.e. instead of knowing the operator Λ_C, we know only the pair of functions $(U, \Lambda_C(U)) \in H^{3/2}(\Omega_D) \times H^{1/2}(\Gamma_1)$. The choice of the *source* function $U(x)$ corresponds to a typical voltage input used in practical experiments.

The setup of the problem is shown in Figure 6.6. The doping profile to be identified is shown in picture (a) – the p-n junction is a straight line. The solution of the direct problem for a typical source U is shown in picture (b). In this figure, as well as in the forthcoming ones, Γ_1 corresponds to the lower edge and the contact $\partial\Omega_D/\Gamma_1$ to the top edge (the origin corresponds to the upper right corner).

The evolution of the level set method for the problem stated above is shown in Figure 6.7. The picture on the top left shows the error for the initial condition. The subsequent pictures show the iteration error for the level set method after 3, 5, 10, 100 and 1000 steps respectively.

One word about the quantity of information used in the identification. In [10, 60] a Landweber-Kaczmarz method was used for solving this inverse doping profile problem. In [60] the authors im-

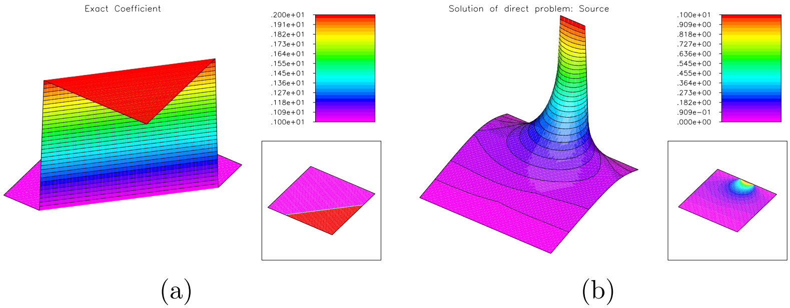

Figure 6.6: Picture (a) shows the doping profile to be reconstructed. On picture (b) the problem data is shown: A typical source $U(x)$ appears as Dirichlet boundary condition at $y = 0$ (upper right edge). The corresponding current is measured at the contact Γ_1 (lower left edge).

plemented the Landweber-Kaczmarz method using different amount of data, i.e. different number of (voltage, current) pairs. In one of the experiments a single pair of data was used and, in this case, the Landweber-Kaczmarz method reduces to the ordinary Landweber iteration.

In [60] the authors observed that the amount of available data strongly influences the quality of the reconstruction. However, no matter how many pairs of (voltage, current) data one uses in the implementation of the Landweber-Kaczmarz method, it does not allow a proper determination of the p-n junction. The observation was that, after a certain number of data pairs, the quality of the reconstruction does not improve any further.

A possible explanation for the poor quality of the results obtained by the Landweber-Kaczmarz method is the fact that this method does not incorporate the assumption that the doping profile is a piecewise C^0 function. This method tries actually to identify a real function defined in Ω, which is a much more complicated object than the original unknown curve (the p-n junction).

In [60] a comparison between the performances of the Landweber method and the level set method (for the inverse doping profile problem) was investigated. Due to the nature of the level set approach, it incorporates in a natural way the assumption that γ is piecewise C^0

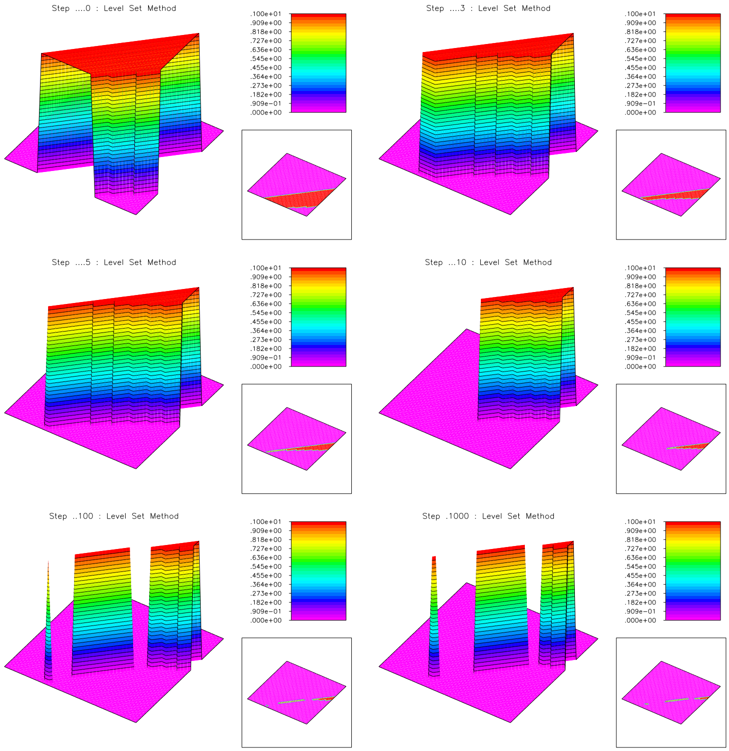

Figure 6.7: Evolution of the level set method for the inverse doping profile problem. Iteration error is shown after 0, 3, 5, 10, 100 and 1000 steps.

in Ω (actually, without this assumption the level set method could not be applied at all). The reconstruction results are much better, although the level set method only uses one pair of (voltage, current) data.

6.6 Bibliographical comments

A comprehensive presentation of level set methods can be found, e.g., in [64, 76, 88], and applications to various inverse problems can be found in [9, 20, 27, 50, 64, 60, 77, 78, 83].

More details on the approach discussed in Section 6.2 are available in [83]. For a detailed discussion of the results presented in Section 6.3 we refer the reader to [9]. The standard references for the results presented in Section 6.4 are [59, 27].

The theoretical background for the inverse potential problem discussed in Subsection 6.5.1 can be found, e.g., in [49, 43]. A level set approach for this inverse problem is treated in [27]. Detailed expositions of semiconductor equations, discussed in Subsection 6.5.2, can be found in [69, 70, 87, 93]. Inverse doping problems for semiconductor equations are discussed in [10, 11]. For a level set approach of this inverse problem we refer the reader to [60].

6.7 Exercises

6.1. Assume that F is Fréchet differentiable and satisfies the tangential cone condition (6.13) in a neighborhood of a solution $\bar{u}$ of (6.1). Prove the estimate

$$\frac{\partial E}{\partial t}(t) \ \leq \ -(1-\eta)\, \|F(u) - g\|^2 \,,$$

for the derivative of the iteration error $E(t)$ in the noise free case (i.e. $\varepsilon = 0$).

6.2. Consider the inverse potential problem discussed in Subsection 6.5.1. Prove that no further on the information on the unknown domain D can be gained by inputting different Dirichlet boundary values in (6.27) and measuring the corresponding Neumann ones.

6.3. Consider the inverse problem in Exercise 6.2. Prove that the integral around $\partial\Omega$ of $\partial u / \partial \nu$ is equal to the area of the unknown domain D.

6.4. Consider again the inverse problem in Exercise 6.2. Prove that it is not possible to recover non-simply connected domains D. (Hint: Let Ω be the unit disc in $\mathbb{R}^2$ and choose Ω to be the annular region $\{(r,\theta);\ 0 < R_1 < r < R_2 < 1,\ 0 \leq \theta < 2\pi\}$. Then, using a symmetry argument, conclude that the measured Neumann data must reduce to a constant value $c \in \mathbb{R}$ and $R_2^2 - R_1^2 = c$.)

6.5. Consider once more the inverse problem in Exercise 6.2. Prove that a starlike domain D with respect to the center of gravity is uniquely determined by its potential.
(Hint: See [49] and the references therein.)

6.6. Let the velocity function $V \in L^1(0,T;L^2(\mathbb{R}^n))^n$ be such that $\operatorname{div} V \in L^1(0,T;L^\infty(\mathbb{R}^n))$. Prove that the initial value problem

$$\frac{\partial \phi}{\partial t} + V \cdot \nabla \phi = 0 \qquad \phi(0) = \phi_0\,, \tag{6.30}$$

with $\phi_0 \in L^2(\mathbb{R}^n)$, has a weak solution $\phi \in C(0,T;L^2(\mathbb{R}^n))$ in the finite-time interval $(0,T)$.
(Hint: Use the vanishing viscosity method, cf. [19].)

6.7. Additionally to the assumptions of Exercise 6.6, let the velocity function satisfy $V \in C(0,\infty;C^{0,1}(\mathbb{R}^n))^n$, i.e. V is continuous on $\mathbb{R}^n \times (0,\infty)$ and Lipschitz continuous with respect to x. Moreover, let the initial condition satisfy $\phi_0 \in L^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$.
a) Prove that the solution ϕ of the initial value problem (6.30) is unique (actually ϕ is a L^∞-function on $\mathbb{R}^n \times (0,T)$).
b) If in addition ϕ_0 is Lipschitz continuous, prove that ϕ is also Lipschitz continuous on $\mathbb{R}^n \times (0,T)$.
(Hint: Use the Picard-Lindelöf theorem, cf. [46].)

6.8. Let $u(x,t)$, $E(t)$ and $\bar{u}(x)$ be defined as in Section 6.3. Prove that, for $t > 0$, the condition $\|Fu(t) - g^\varepsilon\| > \varepsilon$ implies $dE^\varepsilon(t)/dt < 0$. Moreover, if $\varepsilon = 0$, prove that $\int_0^\infty \|Fu(t) - F\bar{u}\|^2\, dt < \infty$.
(Hint: Use Lemma 6.3.1.)

Appendix:

Basic facts in functional analysis

A.1 The Schwartz space

A tuple $k = (k_1, \ldots, k_n) \in \mathbb{Z}^n$ is called a *multiindex with length l* if $k_i \geq 0, 1 \leq i \leq n$, and $l = k_1 + \cdot + k_n$. For a point $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we define with a multiindex $k = (k_1, \ldots, k_n)$

$$x^k := \prod_{i=1}^{n} x_i^{k_i} .$$

Also, if $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a smooth function and $k = (k_1, \ldots, k_n)$ is a multiindex of length l then

$$D^l := \partial^l f := \frac{\partial^l}{\partial^{k_1} x_1 \cdots \partial^{k_n} x_n} .$$

$\mathcal{S}(\mathbb{R}^n)$ is the linear space of those C^∞–functions f on $\mathbb{R}^n$ for which

$$|f|_{k,l} := sup_{u \in \mathbb{R}^n} |u^k D^l f(u)|$$

is finite for all multiindices $k, l \in \mathbb{Z}^2$. $\mathcal{S}(\mathbb{R}^n)$ is a locally convex linear topological space which is called the *Schwartz space*. The continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$ constitute the dual space. This dual

space is denoted by $\mathcal{S}(\mathbb{R}^n)'$ (*tempered distributions*).
Clearly, the Fourier transform may be defined on $\mathcal{S}(\mathbb{R}^n)$ since every function in $\mathcal{S}(\mathbb{R}^n)$ is a $L_1(\mathbb{R}^n)$-function; see Subsection 4.2.1. Now, the Fourier transform can be extended to the space $\mathcal{S}(\mathbb{R}^n)'$ with values in $\mathcal{S}(\mathbb{R}^n)'$ by duality:

$$\mathcal{F}(\lambda)(f) := \lambda(\mathcal{F}(f))\,,\ \lambda \in \mathcal{S}(\mathbb{R}^n)', f \in \mathcal{S}(\mathbb{R}^n)\,.$$

A.2 Hilbert spaces

A *pre-Hilbert* space is a linear space[2] H which is endowed with an *inner product* $\langle\cdot,\cdot\rangle : H \times H \longrightarrow \mathbb{K}$ where $\mathbb{K}$ is a field; for convenience we restrict ourselves to the real case $\mathbb{K} = \mathbb{R}$. This inner product defines a *norm* in H via $\|x\| := \langle x, x\rangle^{\frac{1}{2}}$. When the topological space H endowed with this norm is complete (convergence of Cauchy sequences) then we say that H is a *Hilbert space.*

Given two Hilbert spaces G, H we have the family of linear continuous mappings T from $H \longrightarrow G$. The *operator norm* is given by $\|T\| := \sup_{x\in H}\langle Tx, Tx\rangle_G$ where $\langle\cdot,\cdot\rangle_G, \langle\cdot,\cdot\rangle_H$ are the inner products in G, H respectively. In the special case when $G = \mathbb{R}$ then we set $H^* := \{\lambda : H \longrightarrow \mathbb{R} \mid \lambda \text{ linear, continuous}\}$ and call H^* the *dual space.* H^* is itself a Hilbert space and we know from the Riesz representation theorem that H^* can be identified isometrically with H.

Given a Hilbert spaces H and a linear closed subspace of H we may decompose H as follows:

$$H = U \oplus U^\perp \text{ where } U^\perp := \{v \in H| \mid \langle v, u\rangle = 0 \text{ for all } u \in U\}\,.$$

This result is called the *projection theorem* since each $x \in H$ has a uniquely determined projection $x_U \in U$ with

$$x = x_U + (x - x_U) \text{ and } x - x_U \in U^\perp\,.$$

This result makes it so easy to introduce geometric concepts in Hilbert spaces. Especially, the pseudoinverse of a linear bounded operator can be defined in a straightforward manner.

[2] The null vector in a linear space will be denoted by θ.

When in a Hilbert space H there exists a dense and countable subset then this Hilbert space is called a *separable space.* A countable set of elements $O := \{x^k \mid k \in \mathbb{N}\}$ is called an *orthonormal system* when we have[3]

$$\langle x^k, x^j \rangle = \delta_{ij}\,,\, i,j \in \mathbb{N}\,.$$

An orthonormal set is called an *orthonormal basis* if no orthonormal set of H contains O as a proper subset.

A.3 Hilbert scales

Let H be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_H$ and orthonormal basis $(e_j)_{j\in\mathbb{N}}$ and let $(\alpha_j)_{j\in\mathbb{N}}$ be a sequence with

$$0 < \alpha_{j+1} \le \alpha_j \le 1, j \in \mathbb{N}, \lim_j a_j = 0.$$

Let M be the set of elements in H representable by a finite linear combination of the elements $(e_j)_{j\in\mathbb{N}}$. Then we define on M for each $s \in \mathbb{N}$ an inner product $\langle\cdot,\cdot\rangle_s$:

$$\langle x, y\rangle_s := \sum_{j=1}^{\infty} a_j^{-2s} \langle x, e_j\rangle_H \langle y, e_j\rangle_H .$$

(Notice that the series above is actally a finite sum). The completions H_s of M in the norm $\|\cdot\|_s := \langle\cdot,\cdot\rangle_s^{\frac{1}{2}}$ is a Hilbert space by definition whose inner product is denoted again by $\langle\cdot,\cdot\rangle_s$.

Lemma A.3.1. *We have*

1) $H_0 = H, \langle\cdot,\cdot\rangle_0 = \langle\cdot,\cdot\rangle_H$.

2) *We have for each* $s \ge 0$:

$$\begin{aligned} H_s &= \{x \in H_0 | \sum_{j=1}^{\infty} \alpha_j^{-2s} |\langle x, e_j\rangle_0|^2 < \infty\}\,; \\ \|x\|_s^2 &= \sum_{j=1}^{\infty} \alpha_j^{-2s} |\langle x, e_j\rangle_0|^2, x \in H_s\,. \end{aligned}$$

[3] δ_{ij} is the Kronecker symbol.

3) $H_s \subset H_t \subset H_0$ *for* $s \geq t \geq 0$.

4) *If* $s > t \geq 0$ *then the embedding of* H_s *into* H_t *is dense and compact.*

5) $\|x\|_r \leq \|x\|_s^{1-a} \|x\|_t^a$ *for all* $x \in H_s$ *if* $s \geq r \geq t \geq 0, s \neq t,$ *and* $a = (s-r)(s-t)^{-1}$.

Proof:
The assertions 1), 2) and 3) are simple consequences of the definition of H_s and $\langle \cdot, \cdot \rangle_s$. Let us define $T_N : H_0 \longrightarrow H_0$ by

$$T_N x := \sum_{j=1}^{N} \langle x, e_j \rangle_0 e_j, \; x \in H_0.$$

Then we have

$$\begin{aligned}\|T_N x - x\|_t^2 = \sum_{j=N+1}^{\infty} \alpha_j^{-2t} |(x, e_j)_0\|^2 \leq \sup_{j \geq N+1} \alpha_j^{2(s-t)} \|x\|_s^2 \\ \leq \alpha_{N+1}^{2(s-t)} \|x\|_s^2, \; x \in H_s,\end{aligned}$$

which shows that $(T_N)_{N \in \mathbb{N}}$ converges uniformly to the embedding of H_s into H_t. This implies that this embedding is compact since each T_N is compact and a limit in the operator norm of a sequence of compact operators is compact.
Let $x \in H_s$. Then

$$\|x\|_r^2 = \sum_{j=1}^{\infty} \alpha_j^{-2r} |\langle x, e_j \rangle_0|^2 =$$

$$\sum_{j=1}^{\infty} (\alpha_j^{-2ta} |\langle x, e_j \rangle_0|^{2a})(\alpha_j^{-2s(1-a)} |\langle x, e_j \rangle_0|^{2(1-a)})$$

and by Hölder's inequality

$$\begin{aligned}\|x\|_r^2 &\leq (\sum_{j=1}^{\infty} \alpha_j^{-2t} |\langle x, e_j \rangle_0|^2)^a (\sum_{j=1}^{\infty} \alpha_j^{-2s} |\langle x, e_j \rangle_0|^2)^{(1-a)} \\ &= \|x\|_t^{2a} \|x\|_s^{2(1-a)}.\end{aligned}$$

This gives the result 5.

■

The spaces $H_s, s > 0$, contain generalized (ideal) elements. We clarify the structure of these spaces.
Let $r \in \mathbb{R}$. We define a map $\tilde{D_r} : M \longrightarrow M$ by

$$\tilde{D_r}x := \sum_{j=1}^{\infty} \alpha_j^{-r} |\langle x, e_j\rangle_0 e_j \text{ if } x = \sum_{j=1}^{\infty} \langle x, e_j\rangle_0 e_j$$

Notice that this definition makes sense since $\sum_{j=1}^{\infty} \langle x, e_j\rangle_0 e_j$ is a finite sum if $x \in M$. Since $\|\tilde{D_r}x\| = \|x\|_r$ for all $x \in M$ and since M is dense in H_r the map $\tilde{D_r}$ may be extended to a map $F_r : H_r \longrightarrow H_0$. The following properties of D_r are simple consequences of the definition of $\tilde{D_r}$ and the construction of D_r :

$$D_r(M) = M. \tag{A.31}$$

$$\|D_r x\|_0 = \|x\|_r \text{ for all } x \in H_r. \tag{A.32}$$

$$D_r \text{ is bijective.} \tag{A.33}$$

$$D_r^{-1} : H_0 \longrightarrow H_r \quad , \quad D_r^{-1}x = \sum\nolimits_{j=1}^{\infty} \alpha_j^r \langle x, e_j\rangle_0 e_j;\ r \geq 0. \tag{A.34}$$

$$\langle D_r^{-1}u, x\rangle_0 \quad = \quad \langle u, D_{-r}x\rangle_0 \text{ for all } u \in H, x \in M;\ r \geq 0. \tag{A.35}$$

$$\langle D_{2r}x, u\rangle_0 \quad = \quad \langle D_r, x, D_r u\rangle_0 \text{ for all } u \in H, x \in M;\ r \geq 0 \tag{A.36}$$

Theorem A.3.2. *The dual space H_s^* of H_s is isometric isomorph to H_{-s} for all $s \geq 0$.*

Proof:
Let $s > 0$ (for $s = 0$ nothing has to be proved) and set $t := -s$. Let $\phi \in H_T^*$. If $x \in M$ we have

$$\begin{aligned} \|\langle \phi, x\rangle_t| \quad &\leq \quad \|\phi\|_{H_t^*} \|x\| \ = \ \|\phi\|_{H_t^*} \|D_t x\|_0, \\ \|\langle \phi, D_s x\rangle_t| \quad &\leq \quad \|\phi\|_{H_t^*} \|D_s x\|_t \ = \ \|\phi\|_{H_t^*} \|x\|_0, \end{aligned}$$

where $< \cdot, \cdot >_t$ is the canonic bilinear form on $H_t^* \times H_t$. This shows that the linear functional

$$M \ni x \longmapsto < \phi, D_s x >_t \in \mathbb{R}$$

is bounded in the norm of H_0 on the set $D_s(M)$. Since $D_s(M) = M$ and since M is dense in H_0 there exists by the Riesz representation theorem an element $u \in H_0$ with

$$< \phi, D_s x >_t = \langle u, x \rangle_0 \text{ for all } x \in M.$$

This implies (see (A.35))

$$\langle \varphi, x \rangle_t = \langle u, D_t x \rangle_0 = \langle D_s^{-1} u, x \rangle_0 = \langle z, x \rangle_0 = \langle z, x \langle_0 \text{for all } x \in M \tag{A.37}$$

where $z := D_s^{-1} u \in H_s$. On the other hand, for every $z \in H_s$ the linear form

$$M \ni x \ \longmapsto \ \langle z, x \rangle_0 \in \mathbb{R}$$

defines a linear functional on M, bounded in the norm of H_t :

$$< z, x >= | \ (D_s z, D_t x)_0 \| \leq \| D_s z \|_0 \| D_t x \|_0 = \| z \|_s \| x \|_t .$$

This functional can be extended by continuity to a functional φ in H_t^* with $\|\varphi\|_{H_t^*} \leq \|z\|_s$. We show that we have actually equality: $\|\phi\|_{H_t^*} = \|z\|_s$.

Since $z \in H_s = H_{-t}$ there exist a sequence $(z_n)_{n\in\mathbb{N}}$ of elements in M such that $\lim\limits_n \|z - z_n\|_s = 0$. We set $u_n := D_{2s} z_n, n \in \mathbb{N}$, and have (see (A.36))

$$\langle u_n, z \rangle_0 = \langle D_{2s} z_n, z \rangle_0 = \langle D_s z_n, D_s z \rangle_0$$

and therefore

$$\lim_n \langle u_n, z \rangle_0 = \| D_s z \|_0^2 = \| z \|_s^2 .$$

For any $\varepsilon > 0$ and sufficiently large $n \in \mathbb{N}$ we then obtain

$$\begin{aligned} < \phi, u_n >_t &= \langle u_n, z \rangle_0 \geq (1-\varepsilon) \| z \|_s^2 \\ &\geq (1-2\varepsilon) \| z_n \|_s \| z \|_s = (1-2\varepsilon) \| u_n \|_t \| z \|_s . \end{aligned}$$

This implies the desired inequality $\|\phi\|_{H_t^*} \geq \|z\|_s$.

Thus we have proved that H_t^* is isometric to H_s and an isometry is given by the map $D_t \circ J_0$ where J_0 is the Riesz mapping from H_t^* into H_0. Since Hilbert spaces are reflexive, $H*_t$ is isometric isomorph to H_{-s}. ■

From Lemma A.3.1 and Theorem A.3.2 we obtain by identifying the spaces H_s^* with $H_{-s}, s \geq 0$, the following result.

Theorem A.3.3. *Let $r, t, s \in \mathbb{R}$. We have:*

1) *$H_t \subset H_s$ if $t \geq s$.*

2) *The imbedding of H_t into H_s is dense and compact if $t > s$.*

3) *$\|x\|_r \leq \|x\|_s^{1-a} \|x\|_t^a$ for all $x \in H_s$ if $s \geq r \geq t, s \neq t$, where $a = (s-r)(s-t)^{-1}$.*

4) *$H_t^* = H_{-s}$ for all $s \in \mathbb{R}$.*

5) *$\|u\|_s = \sup\{|\langle u, v\rangle_0 \mid v \in H_{-s}, \|v\|_{-s} \leq 1\}, u \in H_s$, for all $s \in \mathbb{R}$.*

Definition A.3.4. *A family $(H_s)_{j\in\mathbb{R}}$ of separable Hilbert spaces, with inner products $\langle\cdot,\cdot\rangle_s)$, is called a* Hilbert scale *if and only if the following properties hold:*

i) *$H_s \subset H_0 \subset H_t$ with dense and continuous imbeddings; $s \geq 0 \geq t$.*

ii) *For all $s \in \mathbb{R}$ we have $H_s^* = H_{-s}$ and*

$$\|\langle u, v\rangle_0\| \leq \langle u, u\rangle_s^{\frac{1}{2}} \langle v, v\rangle_{-s}^{\frac{1}{2}} \text{ for all } u \in H_s, v \in H_{-s}\,.$$

iii) *$\langle u, u\rangle_r \leq \langle u, u\rangle_s^{2(1-a)} \langle u, u\rangle_t^{2a}$ for all $u \in H_s$; $s \geq r \geq t$, $s \neq t, a := (s-r)(s-t)^{-1}$.*

□

We give an example how to construct specific scales of Hilbert spaces.

Let T be a compact injective operator from a Hilbert space H_0 into another Hilbert space. Let $(e_j, f_j, \sigma_j)_{j\in\mathbb{N}}$ be a singular system of T. Then with the pair $((e_j)_{j\in\mathbb{N}}, (\sigma_j)_{j\in\mathbb{N}})$ a scale of Hilbert spaces can be constructed; we denote this scale by $(H_s(T))_{s\in\mathbb{R}}$ and call T the *generator of the scale.* If we apply this construction to the case considered in 2.1.9 a scale of spaces of Sobolev type results.

The following assertions hold for each pair $T := K, Z := X$ and $T := L, Z := Y$

1. D_T is a dense subspace of Z;
2. T is an unbounded, selfadjoint, positive and closed operator;
3. $\|Tz\| \geq \|z\|$ for all $z \in D_T$, $T^{-1} : Z \longrightarrow D_T$ exists and is bounded;
4. $a := \sup_{t>0} \|(tT+I)^{-1}\| < \infty$ and $b := \sup_{t>0} \|tT(tT+I)^{-1}\| < \infty$;
5. $tT + I : D_T \longrightarrow Z, t > 0$, has a bounded inverse $(tT + I)^{-1}$, the **resolvent**;
6. the fractional powers $T^s : D_{T^s} \longrightarrow Z, s \geq 0$, are welldefined;
7. D_{T^s} can be endowed with the inner product
$$\langle z, \tilde{z}\rangle_s := \langle T^s z, T^s \tilde{z}\rangle, z, \tilde{z} \in D_{T^s}, s \geq 0\,;$$
8. D_{T^s} becomes with $\langle \cdot, \cdot\rangle_s$ a Hilbert space which we denote by $H_s(T), s \geq 0$;
9. the dual space $H_s(T)^*$ is denoted by $H_{-s}(T), s \geq 0$;
10. $H_r(T) \subset Z = H_0(T) \subset H_s(T)\,, \, r < 0 \leq s$;
11. $H_s(T)_{s\in\mathbb{R}}$ is a scale of Hilbert spaces with the following interpolation property:
$$\|z\|_s \leq \|z\|_r^{\frac{t-s}{t-r}} \|z\|_t^{\frac{s-r}{t-r}}, z \in H_t(T), r \leq s \leq t, r \neq t\,;$$
12. we have $T^s : H_l(T) \longrightarrow H_{l-s}(T)$ for all $l, s \in \mathbb{R}$.

All these results are based on the spectral decomposition of T.

A.4 Frechét derivative

Let X, Y be Hilbert spaces and let $F : X \longrightarrow Y$ be a mapping with domain of definition $D(F)$. F is called *Frechét differentiable* in an interior point $x \in D(F)$, if there exists a linear bounded operator $A : X \longrightarrow Y$ such that

$$\lim_{h \rightarrow \theta} \|h\|^{-1} \|F(x+h) - F(x) - Ah\| = 0\,.$$

We write $F'(x) = A$. Notice that $F'(x)$ is uniquely determined. $F'(x)$ is called the *Frechét derivative* in x.

Bibliography

[1] A. Alpers and P. Gritzmann. *On stability, error correction and noise compensation in discrete tomography.* Preprint, München, 2004.

[2] J. August. *Decoupling the equations of regularized tomography.* In: Proceedings 2002 IEEE International Symposium on Biomedical Imaging: Macro to Nano, Washington, D.C., July 7-10, 2002.

[3] A.B. Bakushinsky. *On a convergence problem of the iterative-regularized Gauss-Newton method.* Comput. Math. Math. Phys., 32:1353–1359, 1992.

[4] A.B. Bakushinsky and A. Goncharsky. *Ill-posed problems: theory and applications.* Kluwer, Dordrecht, 1994.

[5] J. Baumeister. *Stable Solution of Inverse Problems.* Fried. Vieweg & Sohn, Braunschweig, 1987.

[6] J. Baumeister. *Deconvolution of appearance potential spectra.* In: Kleinmann, R., et al.: Direct and Inverse Boundary Value Problems. Proc. of Conf. Oberwolfach, Lang-Verlag, Frankfurt/Main, 1991.

[7] J. Baumeister and A. Leitão. *On iterative methods for solving ill-posed problems modeled by PDE's.* Journal of Inverse and Ill-Posed Problems, 9:13–29, 2001.

[8] F. Browder and W. Petryshyn *Construction of fixed points of nonlinear mappings in Hilbert space.* J. Math. Anal. Appl., 20:197–228, 1967.

[9] M. Burger. *A level set method for inverse problems.* Inverse Problems, 17:1327–1355, 2001.

[10] M. Burger, H. Engl, A. Leitão and P. Markowich. *On inverse problems for semiconductor equations.* Milan Journal of Mathematics, 72:273–314, 2004.

[11] M. Burger, H.W. Engl, P.A. Markowich, P. Pietra. *Identification of doping profiles in semiconductor devices.* Inverse Problems, 17:1765–1795, 2001.

[12] M. Burger and O. Scherzer. *Regularization methods for blind deconvolution and blind source separation problems.* Mathematics of Control, Signals and Systems, 14:358-383, 2001.

[13] H. Cabayan and G. Belford. *On computing a stable least squares solution to the inverse problem for a planar Newtonian potential.* SIAM J. Appl. Math., 20:51–61, 1971.

[14] D. Calvetti, P.C. Hansen and L. Reichel. *L-curve curvature bounds via Lanczos bidiagonalization.* Electronic Trans. Numer. Anal. 14:134-149, 2002.

[15] D. Calvetti, B. Lewis and L. Reichel. *GMRES, L-Curves, and discrete ill-posed problems.* BIT, 42:44-65, 2002 .

[16] A.S. Carasso. *Direct blind deconvolution.* SIAM J. Appl. Math., 61:1980-2007, 2001.

[17] K. Chandrasekharan. *Classical Fourier Transform.* Springer, Berlin, 1989.

[18] D. Colton and R. Kress. *Integral Equation Methods in Scattering Theory.* Wiley, New York, 1983.

[19] M.G. Crandall and J.L. Lions. *On existence and uniqueness of solutions of Hamilton-Jacobi equations.* Nonlinear Anal., 10:353–370, 1986.

[20] O. Dorn, E.L. Miller and C.M. Rappaport. *A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets.* Inverse Problems, 16:1119–1156, 2000.

[21] H. Egger. *Semiiterative regularization in Hilbert scales.* SFB-Report 2004-26, Linz, 2004.

[22] H.W. Engl. *Integralgleichungen.* Springer, Vienna, 1997 (in German).

[23] H.W. Engl, M. Hanke and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, Dordrecht, 1996.

[24] A.T. Fomenko and T.L. Kunii. *Topological Modeling and Visualization.* Springer, New York, 1997.

[25] V. Fridman. *Methods of successive approximation for Fredholm integral equations of the first kind.* Uspekhi Mat. Nauk, 11:233–234, 1956 (in Russian).

[26] A. Friedman. *Detection of mines by electric measurements.* SIAM J. Applied Math., 47:201–212, 1987.

[27] F. Frühauf, O. Scherzer and A. Leitão. *Analysis of regularization methods for the solution of ill–posed problems involving discontinuous operators.* SIAM J. Numerical Analysis, to appear, 2005.

[28] G.H. Golub and C.F. Van Loan. *Matrix Computations.* John Hopkins, Baltimore, 1996.

[29] R. Gorenflo and B. Hofmann. *On autoconvolution problems.* Inverse Problems, 10:353–373, 1994.

[30] D. Gottlieb, B. Gustaffson, and P. Forssén. *On the direct Fourier method for computer tomography.* IEEE Trans. on medical imaging 19: 223-232, 2000.

[31] C.W. Groetsch. *Generalized Inverses of Linear Operators.* Dekker, New York, 1977.

[32] C.W. Groetsch. *The theory of Tikhonov regularisation for Fredholm equations of the first kind.* Pittman Publishing, Boston, 1984.

[33] C.W. Groetsch and O. Scherzer. *Nonstationary iterated Tikhonov-Morozov method and third order differential equations for the evaluation of unbounded operators.* Math. Meth. Appl. Sci., 23:1287–1300, 2000.

[34] C.W. Groetsch. *Inverse Problems in Mathematical Sciences.* Fried. Vieweg & Sohn, Braunschweig, 1993.

[35] M. Hanke. *Accelerated Landweber iterations for the solution of ill-posed equations.* Numerische Mathematik, 60:341–373, 1991.

[36] M. Hanke, A. Neubauer and O. Scherzer. *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems.* Numerische Mathematik, 72:21–37, 1995.

[37] P.C. Hansen. *Regularization Tools: A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems.* Numerical Algorithms, 6:1-35, 1994.

[38] P.C. Hansen. *Deconvolution and regularization with Toeplitz matrices.* Numerical Algorithms, 29:323-378, 2002.

[39] D.N. Hao. *A mollification method for ill-posed problems.* Numer. Math., 68:469-506, 1994.

[40] D.N. Hao, H.-J. Reinhardt and A. Schneider. *Stable approximation of fractional derivatives of rough functions.* BIT, 35:488-503, 1995

[41] M. Hegland and R.S. Anderssen. *A mollification framework for imprperly posed problems.* Numerische Mathematik, 78:549-575, 1998.

[42] G.T. Herman and A. Kuba (eds.). *Discrete tomography: Foundations, algorithms, and applications.* Birhäuser, Boston, 1999.

[43] F. Hettlich and W. Rundell. *Iterative methods for the reconstruction of an inverse potential problem.* Inverse Problems, 12:251–266, 1996.

[44] B. Hofmann. *Regularisation for Applied Inverse and Ill–Posed Problems.* Teubner, Stuttgart, 1986.

[45] P.C.V. Hough. *Method and means for recognizing complex patterns.* US patent nr. 3069654, 1962

[46] W. Hurewicz. *Lectures on ordinary differential equations.* Dover, New York, 1990.

[47] J. Illingworth and J. Kittler. *A survey of the Hough transform.* Computer Vision, Graphics and Image Processing, 44:87-116, 1988.

[48] D. Ioannou, W. Huda and A.F. Laine. *Circle recognition through a 2D Hough transform and radius histogramming.* Iamge and Vision Computing, 17:15-26, 1999.

[49] V. Isakov. *Inverse Source Problems.* American Mathematical Society, Providence, Rhode Island, 1990.

[50] K. Ito, K. Kunisch and Z. Li. *Level set function approach to an inverse interface problem.* Inverse Problems, 17:1225–1242, 2001.

[51] M. Kerckhove, editor. *Scale-Space and Morphology in Computer Vision.* Springer, New York, 2001.

[52] A. Kirchgraber, D. Stoffer and A. Kirsch. *Schlecht gestellte Probleme – Oder Wenn das Ungenaue genauer ist.* Mathematische Semesterberichte 51:175–206, 2005.

[53] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Springer, New York, 1996.

[54] S.G. Krein, Ju.I. Petunin and E.M. Demenov. *Interpolation of Linear Operators.* Transl. of Math. Monographs, AMS, Vol. 54, 1982.

[55] R. Kress. *Linear Integral Equations.* 2nd ed Springer, Berlin, 1999.

[56] R. Kreyzig. *Introductory Functional Analysis with Applications.* John Wiley & Sons, New York, 1978.

[57] L. Landweber. *An iteration formula for Fredholm integral euqations of the first kind.* Amer. J. Math., 73:615–624, 1951.

[58] A. Leitão. *An iterative method for solving elliptic Cauchy problems.* Numerical Functional Analysis and Optimization, 21:715–742, 2000.

[59] A. Leitão and O. Scherzer. *On the relation between constraint regularization, level sets and shape optimization.* Inverse Problems, 19:L1–L11, 2003.

[60] A. Leitão, P. Markowich and J.P. Zubelli. *On inverse dopping profile problems for the voltage-current map.* submitted, 2005.

[61] J. Liesen, P Tichý. *The worst-case GMRES for normal matrices.* BIT Numerical MathematicsAnalysis, 44:79–98, 2004.

[62] T. Lindeberg. *Scale-Space Theory in Computer Vision.* Kluwer, Boston, 1994.

[63] J.L. Lions. *Exact controllability, stabilization and perturbations for distributed systems.* SIAM Review, 30:1–68, 1988.

[64] A. Litman, D. Lesselier and F. Santosa. *Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set.* Inverse Problems, 14:685–706, 1998.

[65] A.K. Louis. *Inverse und schlecht gestellte Probleme.* Teubner, Stuttgart, 1989 (in German).

[66] A.K. Louis. *A unified approach to regularization methods for linear ill-posed problems.* Inverse Problems, 15:489–498, 1999.

[67] B.A. Mair. *Tikhonov reularization for finitely and infinitely smoothing operators.* SIAM J. Math. Anal., 25:135–147, 1994.

[68] B.A. Mair. *Remarks on a non-well posed problem.* Proc. of the Royal Society of Edinbourgh 102A:131–140, 1986.

[69] P.A. Markowich. *The Stationary Semiconductor Device Equations.* Springer, Vienna, 1986.

[70] P.A. Markowich, C.A. Ringhofer and C. Schmeiser. *Semiconductor Equations.* Springer, Vienna 1990.

[71] V.A. Morozov. *Methods for Solving Incorrectly Posed Problems.* Springer, New York, 1984.

[72] V.A. Morozov. *Regularization Methods for Ill–Posed Problems.* CRC Press, Boca Raton, 1993.

[73] M.T. Nair, M Hegland and R.S. Anderssen. *The trade-off between regularity and stability in Tikhonov regularization.* Math. of Computations, 66:193–206, 1997.

[74] F. Natterer. *The Mathematics of Computerized Tomography.* John Wiley, New York, 1986.

[75] M. Nielsen, P. Johansen, O.F. Olsen and J. Weickert, editors. *Scale-Space Theories in Computer Vision.* Lecture Notes in Computer Science Vol. 1683, Springer, 1999. Proceedings of the Second International Conference, Scale-Space'99, Corfu, Greece, 1999.

[76] S. Osher and J.A. Sethian. *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations.* J. Comput. Phys., 79:12–49, 1988.

[77] C. Ramananjaona, M. Lambert and D. Lesselier. *Shape inversion from TM and TE real data by controlled evolution of level sets.* Inverse Problems, 17:1585–1595, 2001. Special section: Testing inversion algorithms against experimental data.

[78] C. Ramananjaona, M. Lambert, D. Lesselier and J.-P. Zolésio. *Shape reconstruction of buried obstacles by controlled*

evolution of a level set: from a min-max formulation to numerical experimentation. Inverse Problems, 17:1087–1111, 2001.

[79] H.-J. Reinhardt, D.N. Háo and F. Seiffarth. *Stable numerical fractional differentiation by mollification.* Numer. Funct. Anal. and Optimiz., 15:635–659, 1994.

[80] W. Ring. *Identification of a core from boundary data.* SIAM J. Appl. Math., 55:677–706, 1995.

[81] A. Rosenfeld. *Picture Processing by Computer.* Academic Press, 1969.

[82] S. Saitoh, V.K. Tuan and M. Yamamoto. *Reverse Convolution inequalities and applications to inverse heat conduction source problems.* Journal of Inequalities in Pure and Applied Mathematics, 3, 2002.

[83] F. Santosa. *A level set approach for inverse problems involving obstacles.* ESAIM Contrôle Optim. Calc. Var., 1:17–33 (electronic), 1995/96.

[84] O. Scherzer. *Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems.* J. Math. Anal. Appl., 194:911–933, 1995.

[85] O. Scherzer. *A modified Landweber iteration for solving parameter estimation problems.* Appl. Math. Optim., 38:45–68, 1998.

[86] O. Scherzer and C.W. Groetsch. *Inverse scale space theory for inverse problems.* In *[51]*, pages 317–325, 2001.

[87] S. Selberherr. *Analysis and Simulation of Semiconductor Devices.* Springer, Vienna, New York, 1984.

[88] J.A. Sethian. *Level set methods and fast marching methods.* 2nd ed Cambridge University Press, Cambridge, 1999.

[89] U. Tautenhahn. *On the asymptotical regularization of nonlinear ill-posed problems.* Inverse Problems, 10:1405–1418, 1994.

[90] U. Tautenhahn. *On the asymptotical regularization method for nonlinear ill-posed problems.* In Dang Dinh Ang (ed.) et al., editor, *Inverse Problems and Applications to Geophysics, Industry, Medicine and Technology*, 158–169. Vietnam Mathematical Society, 1995. Proceedings of the international workshop on inverse problems, 17–19 January 1995, HoChiMinh City, Vietnam.

[91] A.N. Tikhonov and V.Y. Arsenin. *Solutions of Ill-Posed Problems.* Winston & Sons, Washington, D.C., 1977.

[92] G.M. Vainikko and A.Y. Veretennikov. *Iteration procedures in ill-posed problems.* Nauka, Moscow, 1986 (in Russian).

[93] W.R. van Roosbroeck. *Theory of flow of electrons and holes in germanium and other semiconductors.*

[94] L.v. Wolfersdorf and L. Janno. *On a class of nonlinear convolution equations.* Z. Anal. Anw., 14:497-508, 1995.

[95] J.P. Zubelli. *An Introduction to Inverse Problems: Examples, Methods and Questions.* IMPA (22$^{\text{nd}}$ Brazilian Mathematics Colloquium), Rio de Janeiro, 1999.