DE GRUYTER

Michael V. Klibanov, Jingzhi Li

INVERSE PROBLEMS AND CARLEMAN ESTIMATES

GLOBAL UNIQUENESS, GLOBAL CONVERGENCE AND EXPERIMENTAL DATA

INVERSE AND ILL-POSED PROBLEMS SERIES 63

Michael V. Klibanov, Jingzhi Li

Inverse Problems and Carleman Estimates

Inverse and Ill-Posed Problems Series

Edited by
Sergey I. Kabanikhin, Novosibirsk, Russia;
Almaty, Kazakhstan

Volume 63

Michael V. Klibanov, Jingzhi Li

Inverse Problems and Carleman Estimates

Global Uniqueness, Global Convergence and Experimental Data

DE GRUYTER

Mathematics Subject Classification 2010
35R30, 65N21

Authors
Prof. Dr. Michael V. Klibanov
University of North Carolina at Charlotte
Department of Mathematics and Statistics
9201 University City Blvd.
Charlotte, NC 28223, USA
mklibanv@uncc.edu

Prof. Jingzhi Li
Southern University of Science and Technology
Department of Mathematics
1088 Xueyuan Ave
Shenzhen 518005, P.R. China
li.jz@sustech.edu.cn

ISBN 978-3-11-074541-2
e-ISBN (PDF) 978-3-11-074548-1
e-ISBN (EPUB) 978-3-11-074555-9
ISSN 1381-4524

Library of Congress Control Number: 2021940405

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Michael Victor Klibanov dedicates this book to his family: Ada (mother), Victor (father), Vera (wife), Olga and Gregory (children), Luba (sister), and Leah and Victor (grandchildren). They have been always very supportive of him and his mathematical effort. Jingzhi Li dedicates this book to his family: Fengying (mother), Shide (father), Ning (wife), Mengya and Mengqi (children). They have been always very supportive of him and his mathematical interest.

Preface

In 1981, almost 40 years ago, A. L. Bukhgeim and M. V. Klibanov introduced, *for the first time*, the powerful tool of Carleman estimates in the field of Inverse Problems [51]. The technique of [51] is now called the "Bukhgeim–Klibanov method" (BK); see [48, 49, 120, 122] for the first detailed proofs of theorems of [51]. According to Google Scholar, the paper [51] currently (2021) has more than 570 citations; see https://scholar.google.com/citations?user=pFmp7LMAAAAJ&hl=en.

While initially the BK method was thought only as a tool for proofs of global uniqueness and stability theorems for coefficient inverse problems, it was discovered recently that the ideas of BK generate a powerful numerical method for these problems, the so-called convexification method.

This book summarizes the main analytical and numerical results of M. V. Klibanov and J. Li about the technique of Carleman estimates, which they have obtained since the publication [51].

Given a Partial Differential Equation (PDE), a Coefficient Inverse Problem (CIP) for it is the problem of finding either one or several coefficients of that equation from additional boundary measurements. CIPs have a rapidly growing number of applications in many fields; see Chapters 7–12.

The following four topics are discussed in this book:

1. *Topic 1*: Derivation of Carleman estimates and conditional stability estimates for some ill-posed Cauchy problems, Chapter 2.
2. *Topic 2*: Global uniqueness for multidimensional CIPs on the basis of the BK method, Chapter 3.
3. *Topic 3*: Carleman estimates for numerical methods for ill-posed Cauchy problems for PDEs, Chapters 4 and 5.
4. *Topic 4*: The convexification globally convergent numerical concept for CIPs: a far reaching consequence of the BK method, Chapters 6–12.

https://doi.org/10.1515/9783110745481-201

Acknowledgments

Michael Victor Klibanov was supported by two US Army Research Laboratory and US Army Research Office grants W911NF-15-1-0233 and W911NF-19-1-0044. M. V. Klibanov heartily appreciates the support and understanding of Dr. Joseph D. Myers, a Program Manager of The US Army Research Office.

Jingzhi Li was partially supported by the NSF of China No. 11971221, Shenzhen Sci-Tech Fund Nos. JCYJ20190809150413261 and JCYJ20170818153840322, and Guangdong Provincial Key Laboratory of Computational Science and Material Design No. 2019B030301001.

https://doi.org/10.1515/9783110745481-202

Contents

1 Topics of this book

1.1 Topic 1: Derivation of Carleman estimates and stability estimates for some ill-posed Cauchy problems, Chapter 2

The first thing to do for this paper is to derive Carleman estimates and demonstrate their usefulness. So, in Chapter 2 we derive Carleman estimates for three main types of Partial Differential Operators (PDOs): parabolic, elliptic, and hyperbolic ones. In parabolic and elliptic cases, we follow Section 4.1 of [184], and in the hyperbolic case, we follow Section 1.10.2 of [22]; see more details in the first paragraph of Chapter 2. We use the so-called "pointwise" Carleman estimates. There are also integral Carleman estimates; see [93]. However, we believe that the pointwise case provides more details about boundary terms, which is important sometimes.

As soon as Carleman estimates are proven, we prove next Hölder stability estimates for ill-posed Cauchy problems for parabolic, elliptic, and hyperbolic PDEs and, more generally, inequalities. These estimates were actually known from [184].

Next, however, we prove a stronger Lipschitz stability estimate for the Cauchy problem with lateral data for a hyperbolic equation and, more generally, hyperbolic inequality. The initial data are not given in this case. The first such estimate was proven by Lop Fat Ho by the so-called method of multipliers [77]. However, it was clear from [77] that this method is sensitive to lower order terms of the hyperbolic operator. On the other hand, it is well known that Carleman estimates are independent on low order terms of PDOs; see, for example, Lemma 2.1.1. Thus, [153] was the first paper where a Carleman estimate was applied to get Lipschitz stability estimate for a hyperbolic PDE with lower order terms. Next, publications [64, 114, 132, 165] followed.

Furthermore, we use this idea in Section 2.7 to prove Lipschitz stability for a hyperbolic CIPs. Imanuvilov and Yamamoto were the first ones who has proved Lipschitz stability for hyperbolic CIPs [95–99]; also, see [103]. They have combined the idea of the BK method with the idea of [114, 153]. The idea of Section 2.7 is slightly different from the ones of these publications.

1.2 Topic 2: Global uniqueness theorems for multidimensional CIPs on the basis of the BK method, Chapter 3

First, we bring in some historical notes. The scientific career of the first author started in 1973 when he became a Ph. D. graduate student of the Computing Center of the Soviet Academy of Science (currently Russian Academy of Science) in Novosibirsk, Russian Federation. The thesis advisor of Klibanov was Mikhail M. Lavrent'ev, who was a Member of the Soviet Academy of Science. Lavrent'ev (1932–2010) was one of the founders of the theory of ill-posed and inverse problems; see, for example, the funda-

https://doi.org/10.1515/9783110745481-001

mental book [184]. He has led a large, creative, and very energetic team of experts in this field.

Definition 1.2.1. We call a Coefficient Inverse Problem (CIP) multidimensional if the unknown coefficient of the corresponding Partial Differential Equation (PDE) depends on $n \geq 2$ variables.

Definition 1.2.2. We call a Coefficient Inverse Problem (CIP) overdetermined if the number m of free variables in the boundary data exceeds the number n of free variables in the unknown coefficient, $m > n$. If, however, $m = n$, then we call such a CIP non-overdetermined.

The Lavrent'ev's group was focused on non-overdetermined multidimensional CIPs. Indeed, the non-overdetermined case is the most economical case of data gathering. In about 1978, it became clear to many experts that studied inverse problems that this field has entered an ideological crisis from the analytical standpoint; see [184] for the main results obtained by that time. Indeed, although many uniqueness theorems for non-overdetermined CIPs were proven by that time, people knew that those theorems were not completely satisfactory. The reason was that it was assumed in each of those theorems that the unknown coefficient belongs to a restrictive functional class, such as, for example, piecewise analytic functions; functions whose certain norm is sufficiently small, functions represented via truncated Fourier series, etc. We call such results *local uniqueness theorems*. However, many researchers dreamed to prove such uniqueness theorems for CIPs, which would basically impose only one condition on the unknown coefficient: that it belongs to one of main function spaces, such as, for example, C^k, H^k. We call such results *global uniqueness theorems*.

In 1979, Klibanov had the same dream and has worked tirelessly during 1979–1980 to prove global uniqueness theorems for multidimensional CIPs with non-overdetermined data. But nothing worked out in 2 years. Suddenly, however, after an infinite number of failed attempts, he got the right idea while vacationing in a Black Sea resort in August 1980. He figured out that the very powerful and sophisticated tool of Carleman estimates can be successfully combined with his own new ideas to prove the commonly dreamed global uniqueness theorems for multidimensional CIPs. Furthermore, to his great surprise, the resulting method did not depend on a specific PDE operator, unlike all previous publications. Rather, global uniqueness theorems were proven in a unified manner. Roughly speaking, as soon as the Carleman estimate is valid for a PDE operator, the global uniqueness theorem for a corresponding CIP is valid. However, since Carleman estimates are valid for all three main types of PDE operators, parabolic, elliptic, and hyperbolic ones (see Chapter 4 of [184] and Chapter 2 of this book), then this method is a very general one.

A famous Swedish mathematician, Torsten Carleman, has introduced in his 8-page paper [58] (1939) a ground breaking tool of weighted estimates for PDE operators, which currently carry Carleman's name. Since then, that idea of Carleman was not explored further for about 19 years until the work of Calderon in 1958 [55]. Since then, Carleman estimates were used by many mathematicians for proofs of uniqueness theorems for ill-posed problems for PDEs; see, for example, books [93, 184]. For a long time, an impression was that there is no connection between Carleman estimates and CIPs. It was discovered in [51], however, that a fundamental connection exists.

It turned out that A. L. Bukhgeim has also discovered that connection, independently and simultaneously with M. V. Klibanov. Thus, that idea about this connection was first published in a joint paper of Bukhgeim and Klibanov [51] with first complete proofs in publications [48, 49, 120, 122] of these two authors. Furthermore, actually a modified technique of [51], being combined with the technique of Chapter 4 of [184], enables one to prove conditional Hölder stability estimates for those CIPs. The following statement of the first paragraph of [51] well describes the state-of-the-art in the field of inverse problems prior the publication [51] "*Uniqueness theorems for multidimensional inverse problems have at present been obtained mainly in classes of piecewise analytic functions and similar classes or locally... Moreover, the technique of investigating these problems has, as a rule depended in an essential way on the type of the differential equation. In this note, a new method of investigating inverse problems is proposed that is based on weighted a priori estimates. This method makes it possible to consider in a unified way a broad class of inverse problems for those equations $Pu = f$ for which the solution of the Cauchy problem admits a Carleman estimate... The theorems of § 1 were proved by M. V. Klibanov and those of § 2 by A. L. Buhgeìm. They were obtained simultaneously and independently.*"

Currently, 40 years later, the BK method remains the single one allowing to prove global uniqueness and conditional stability results for multidimensional CIPs with non-overdetermined data. Many publications of many authors are devoted to the BK method. The main follow-up publications of Klibanov regarding the BK method, which are devoted to proofs of global uniqueness theorems for those CIPs, are [22, 73, 113, 121, 123–126, 128, 129, 131, 132, 165, 167]. There are also plenty of publications of many other authors regarding various versions of the BK method. Since we do not intend to provide a survey of this method in this book, then we list now only some of them: [13–15, 18, 19, 28–35, 39, 50, 59, 62] as well as [67, 78, 94–103, 181, 191–195, 197, 198, 220, 227, 245, 251–255]. We refer to [132] for a survey of the BK method as of 2013. Even though the main focus of this book is on Coefficient Inverse Problems, we also refer to, e. g. the book [54] for an important topic of inverse problems of shape reconstruction.

1.3 Topic 3: Carleman estimates for numerical methods for ill-posed Cauchy problems for PDEs, Chapters 4, 5, and 12

The problem of the reconstruction of the solution of a PDE of the second order using Dirichlet and Neumann data on a part of the boundary, that is, Cauchy data on that part, has a long history of interest starting from the famous Hadamard example of the

Cauchy problem for the Laplace equation; see, for example [184]. A variety of numerical methods are proposed for this problem. In this regard, we refer to, for example, [8, 37, 72, 74, 75, 90–92, 169]. However, these methods depend on the type of the PDE one is working with.

On the other hand, R. Lattes and J.-L. Lions have proposed in 1969 a universal method for solving these problems, which they named the "Quasi-Reversibility Method" (QRM) [182]. One of peculiarities of the QRM is the convergence rate of regularized solutions. So, while convergence of those solutions was proven in [182], convergence rates were not established. It was proposed in 1991 in [153, 161] to use Carleman estimates for proofs of convergence rates of regularized solutions of the QRM. Since then, this tool is widely used for the QRM. In this regard, we refer to works [57, 64, 134, 147, 148, 152, 156, 158, 165, 166, 190, 207–210, 236]. We also refer to works of L. Bourgeois and J. Dardé with coauthors in which the idea of the QRM is elegantly applied to a number of problems and Carleman estimates are used [40–47, 68, 69]; also, see references cited therein.

So, we discuss in Chapter 4 the idea of using Carleman estimates for proofs of convergence rates of various versions of the QRM for linear PDEs.

However, the nonlinear case was not considered in publications cited above in Section 1.3. In [135], a unified approach for construction of globally convergent numerical methods for ill-posed Cauchy problems for quasilinear PDEs was proposed. In fact, this is a prelude to the convexification method for CIPs (Chapters 7–11). This idea was developed further in [9, 146] with some numerical results. We present this idea in Chapter 5.

Finally, in Chapter 12, we show how the QRM can be applied to the linearized problem of travel time tomography with incomplete data [148]; also, see [236] for a similar result for an inverse source problem for the transport equation.

1.4 Topic 4: The convexification globally convergent numerical concept for CIPs: A far reaching consequence of the BK method, Chapters 6–11

So, we arrive now at numerical studies.

Definition 1.4.1. Given a coefficient inverse problem, we call a numerical method for it *locally* convergent if one can prove its convergence to the correct solution only if the starting point of iterations is located in a sufficiently small neighborhood of the exact solution. In other words, this is a version of the small perturbation approach.

Definition 1.4.2. Given a coefficient inverse problem, we call a numerical method for it *globally* convergent if there exists a theorem, which claims that iterations of this method lead to a sufficiently small neighborhood of the exact solution regardless on any advanced knowledge of this neighborhood.

It is clear that locally convergent numerical methods are unstable and unreliable since a good guess about the exact solution is rarely available. It is also clear that globally convergent numerical methods are by far more attractive than locally convergent ones.

Coefficient inverse problems are applied ones. Therefore, it is far insufficient to work on their theory only. Rather, reliable numerical methods for CIPs are indeed *paramount* for the entire field of inverse problems. It turns out that the idea of the BK method leads to the so-called *convexification* concept of globally convergent methods for CIPs.

CIPs are not only ill posed but highly nonlinear as well. Here is a simple example of the nonlinearity. Let $c = \text{const}$. Consider the Cauchy problem for an elementary ordinary differential equation:

$$\frac{du}{dt} = cu,$$
$$u(0) = 1.$$

Its solution is $u(t) = e^{ct}$. Thus, the function $u(t, c)$ depends highly nonlinearly on the coefficient c.

The vast majority of numerical methods for multidimensional CIPs are based on the minimization of least squares misfit cost functionals. As some of many examples, we refer here to [60, 81–83, 175, 222]. However, these functionals are nonconvex. As a rule, they have plenty of local minima. In this regard, we refer to, for example, [229] for a numerical example of the phenomenon of local minima. On the other hand, any minimization procedure is based on a version of the gradient method. The gradient method stops at such a point where the Fréchet derivative attains its zero value. Hence, it can stop at any point of a local minimum. In fact, convergence of this method can sometimes be guaranteed only if its starting point is located in a sufficiently small neighborhood of the exact solution [10, 11]. The latter means that conventional numerical methods for CIPs are locally convergent ones; see Definition 1.4.1. The authors are unaware about least squares minimization methods, which would satisfy Definition 1.4.2.

As to the globally convergent numerical methods for multidimensional CIPs with overdetermined data, we refer to methods of M. I. Belishev [26, 27, 71] and S. I. Kabanikhin [108–112]. There are also globally convergent numerical methods, which were developed by the group R. G. Novikov since 1988; see [211] for the first result as well as, for example, [1, 3, 52, 53, 212–216]. The statements of CIPs in these publications are different from ours. These reconstruction techniques are also different from the convexification. Nevertheless, these methods are globally convergent ones, as per the above Definition 1.4.2. Furthermore, an interesting feature of [1, 215] is that these publications consider the case of the non-overdetermined data for the reconstruction of the potential of the Schrödinger equation at the high values of the wavenumber.

Another noticeable feature of [1] is that the data there are phaseless. Corresponding numerical results can be found in [1, 3].

The research group of the first author has developed two globally convergent numerical methods for CIPs with non-overdetermined data. We call the first one the "tail function method." This method is basically due to works of L. Beilina and M. V. Klibanov; see [21] for the first publication and [24, 25, 63, 139, 154, 170, 171, 204, 205, 242, 243] for some follow-up results, many of them are the ones for the backscattering experimental data. The book [22] summarizes many important details about this method. Also, see, for example, [163, 196, 218, 234, 240] for another version of this method.

Chapters 6–11 are devoted to the second globally convergent numerical method for CIPs with non-overdetermined data. This is the so-called "convexification" method, and it has deep roots in the BK method. The convexification is not a ready-to-use algorithm, but rather a concept. The first author has started to work on the convexification in 1995 [140] with a coauthor. Some initial follow-up results of the first author with coauthors were in [23, 127, 141, 164, 165]. However, those results were mostly theoretical ones, although with some limited exceptions when numerical studies were also presented [164, 165]. The reason of this was that some important points for the numerical implementation were not addressed analytically in these works. The first work were these points were addressed was [9]. Since then both analytical and numerical results on the convexification of the research group of the first author started to flourish [115–117, 137, 138, 142–146, 150, 151, 237, 238].

The convexification concept allows one to get the global convergence property. In the convexification, one constructs a weighted cost functional J_λ, where $\lambda \geq 1$ is the parameter of the Carleman Weight Function (CWF), that is, the function involved in the Carleman estimate for the corresponding PDE operator. This functional is minimized on a convex bounded set $B(d) \subset H^k$, where $d > 0$ is the diameter of this set. The main theorem states that there exists a number $\lambda(d)$ such that for all $\lambda \geq \lambda(d)$ the functional J_λ is strictly convex on $B(d)$. We prove in this chapter that this strict convexity property implies convergence of the gradient projection method being applied to J_λ to the exact solution if starting from an arbitrary point of $B(d)$. An important point is that $d > 0$ is an arbitrary number here. Therefore, this is *global convergence*; see Definition 1.4.2. Of course, a concern can be raised that the parameter λ should be sufficiently large. However, our vast computational experience tells us that $\lambda \in [1, 3]$ is sufficient. Besides, such concerns can be quite rightfully raised about all asymptotic methods for many problems, and usually the computational answers are positive.

2 Carleman estimates and Hölder stability for ill-posed Cauchy problems

In this chapter, we follow publications [9, 22, 132, 134, 135, 165]. Permissions for republishing from corresponding publishers are obtained. In terms of Section 2.3, we refer to Section 4.1 of the book [184]. Full credit is given to the American Mathematical Society(AMS) publication in which the material was originally published by AMS. We emphasize that the permission for this part has been granted by AMS. The material to be used in our book is without credit or acknowledgment to another source. The material is not available, in whole or in part, on a standalone basis, or in any way exclusive of the book as a whole.

First, we introduce a definition of the Carleman estimate for a general linear Partial Differential Operator (PDO) of the second order. Next, we show how to obtain Hölder stability estimates for corresponding ill-posed Cauchy problems for these operators. Next, we derive Carleman estimates for three main types of Partial Differential Operators (PDOs) of the second order: parabolic, elliptic, and hyperbolic ones. Next, we specify Hölder stability estimates for corresponding ill-posed Cauchy problems for these operators. Finally, we show that, in the case of a hyperbolic operator, one can obtain even stronger Lipschitz stability estimate for the case when the Cauchy data are given on the lateral boundary of the time cylinder and initial data at $\{t = 0\}$ are not given.

We now remind some statements of Section 1.1. As to the Hölder stability estimates for ill-posed Cauchy problems for parabolic, elliptic, and hyperbolic equations on the basis of Carleman estimates, we refer to Chapter 4 of the book of Lavrentiev, Romanov, and Shishatskii [184]. As to the stronger Lipschitz stability estimate for the hyperbolic case, the first result was obtained by Lop Fat Ho [77] using the so-called method of multipliers; see, for example, the book of Isakov [102] for this method. However, the paper [77] works only for the purely wave operator $\partial_t^2 - \Delta_x$ without lower order terms. The *first work* where lower order terms were incorporated was the paper of Klibanov and Malinsky [153]. This is the first publication, where the apparatus of Carleman estimates was applied to this problem. It is the independence of Carleman estimates on low order terms of operators (Lemma 2.1.1 in Section 2.1), which has allowed to use them to obtain the Lipschitz stability estimate for a more general hyperbolic operator $\partial_t^2 - \Delta_x + \text{lot}$, where "lot" stands for "low order terms." The result of [153] was generalized by Kazemi and Klibanov [114]. Next, this idea has found applications in the control theory [176–180]. Naturally, since the Carleman estimate is valid for the hyperbolic operator $c(x)\partial_t^2 - \Delta_x$ with an appropriate coefficient $c(x)$ (see Section 2.3), then the idea of [153] was extended to the case of this operator, see, for example, [165].

https://doi.org/10.1515/9783110745481-002

2.1 What is the Carleman estimate

For further convenience, we consider in this section a general PDO of the second order. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ be a multiindex with nonnegative integer components and $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$. We remind that for any appropriate function $u(x)$,

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial_{x_n}^{\alpha_n} \cdots \partial_{x_1}^{\alpha_1}}.$$

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a piecewise smooth boundary $\partial\Omega$. Consider the function $\psi \in C^2(\overline{G})$ such that $|\nabla\psi| \neq 0$ in $\overline{\Omega}$. For a number $h \geq 0$, denote

$$\psi_h = \{x \in \overline{\Omega} : \psi(x) = h\}, \quad \Omega_h = \{x \in \Omega : \psi(x) > h\}. \tag{2.1}$$

Let the domain $\Omega_h \neq \varnothing$. Consider a part Γ_h of $\partial\Omega$ defined as

$$\Gamma_h = \{x \in \partial\Omega : \psi(x) \geq h\}. \tag{2.2}$$

Then the boundary $\partial\Omega_h$ of Ω_h is

$$\partial\Omega_h = \partial_1\Omega_h \cup \partial_2\Omega_h, \tag{2.3}$$

$$\partial_1\Omega_h = \psi_h, \quad \partial_2\Omega_h = \Gamma_h. \tag{2.4}$$

Let $\lambda \geq 1$ be a parameter, which is defined later. Consider the function $\varphi(x)$,

$$\varphi(x) = \exp(\lambda\psi(x)). \tag{2.5}$$

It follows from (2.2)–(2.5) that

$$\min_{\overline{\Omega}_h} \varphi(x) = \varphi(x)|_{\psi_h} \equiv e^{\lambda h}. \tag{2.6}$$

Consider a linear PDO $A(x, D)$ of the second order with real valued coefficients in Ω,

$$A(x, D)u = \sum_{|\alpha| \leq 2} a_\alpha(x) D^\alpha u. \tag{2.7}$$

The principal part of the operator $A(x, D)$ is the operator $A_0(x, D)$,

$$A_0(x, D)u = \sum_{|\alpha| = 2} a_\alpha(x) D^\alpha u. \tag{2.8}$$

We assume that coefficients of the operator $A(x, D)$ are such that

$$a_\alpha \in C^1(\overline{\Omega}) \quad \text{for } |\alpha| = 2, \quad K := \max_{|\alpha|=2}(\|a_\alpha\|_{C^1(\overline{\Omega})}); \quad a_\alpha \in C(\overline{\Omega}) \quad \text{for } |\alpha| = 0, 1. \tag{2.9}$$

Definition 2.1.1. Let $\Omega_h \neq \varnothing$ and (2.7)–(2.9) hold. The operator $A_0(x, D)$ in (2.8) admits pointwise Carleman estimate in the domain Ω_h if there exist constants $\lambda_0(\Omega_h, K) \geq 1$, $C(\Omega_h, K) > 0$ depending only on the domain Ω_h and the number K, such that the following a priori estimate holds:

$$(A_0 u)^2 \varphi^2(x) \geq C\lambda(\nabla u)^2 \varphi^2(x) + C\lambda^3 u^2 \varphi_\lambda^2(x) + \operatorname{div} U, \tag{2.10}$$

$$\forall \lambda \geq \lambda_0, \forall u \in C^2(\overline{\Omega}_h), \forall x \in \Omega_h. \tag{2.11}$$

The divergence term in (2.10) should satisfy the following estimate:

$$|U| \leq C\lambda^3 \left[(\nabla u)^2 + u^2\right] \varphi^2(x). \tag{2.12}$$

In this case, the function $\varphi(x)$ is called the Carleman Weight Function (CWF) for the operator $A_0(x, D)$.

Lemma 2.1.1. *Suppose that conditions* (2.9)–(2.12) *hold. Then conditions* (2.10)–(2.12) *are also valid for the operator* $A(x, D)$*, although with a different constant* λ_0*. In other words, the Carleman estimate depends only on the principal part of the operator.*

Proof. We have

$$(Au)^2 \varphi^2(x) \geq (A_0 u)^2 \varphi^2(x) - M\left[(\nabla u)^2 + u^2\right] \varphi^2(x), \quad \forall x \in \Omega_h, \tag{2.13}$$

where $M > 0$ is a constant depending only on the maximum of norms $\|a_\alpha\|_{C(\overline{\Omega})}$, $|\alpha| = 0, 1$. Substituting (2.13) in (2.10) and taking λ sufficiently large, we again obtain (2.10). □

2.2 Hölder stability for ill-posed Cauchy problems

We show in this section how the Carleman estimates enable one to obtain Hölder stability estimates for ill-posed Cauchy problems for PDEs. Rather then considering a PDE, we consider now a more general case Cauchy problem for a differential inequality,

$$|A_0 u| \leq B\big(|\nabla u| + |u| + |f|\big), \quad \forall x \in \Omega_h, \tag{2.14}$$

$$u|_{\Gamma_h} = g_0(x), \quad \partial_n u|_{\Gamma_h} = g_1(x). \tag{2.15}$$

In (2.14), $B = \text{const.} > 0$ and $f \in L_2(\Omega_h)$ is a function. Functions g_0, g_1 in (2.15) are the Cauchy data for the function u. In particular, equation $Au = f$ with the boundary data (2.15) can be reduced to the problem (2.14), (2.15). We assume that functions $g_0 \in H^1(\Gamma_h)$, $g_1 \in L_2(\Gamma_h)$. In Theorem 2.2.1, we estimate the function u via functions g_0, g_1, f.

Theorem 2.2.1 (Hölder stability estimate). *Assume that conditions* (2.9) *hold and that the operator* A_0 *satisfies the Carleman estimate of Definition* 2.1.1. *Let* $\varepsilon > 0$ *be a sufficiently small number such that* $\Omega_{h+3\varepsilon} \neq \varnothing$. *Let* $m = \max_{\overline{\Omega}_h} \psi(x)$. *Consider the number* $\beta = 2\varepsilon/(3m + 2\varepsilon) \in (0,1)$. *Assume that functions* g_0, g_1, f *are such that* $g_0 \in H^1(\Gamma_h)$, $g_1 \in L_2(\Gamma_h), f \in L_2(\Omega_h)$. *Suppose that the function* $u \in C^2(\overline{\Omega}_h)$ *satisfies conditions* (2.14), (2.15). *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, m, B, K, \Omega_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, m, B, K, \Omega_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0)$ *and*

$$\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\Gamma_h)} + \|g_1\|_{L_2(\Gamma_h)} \leq \delta, \tag{2.16}$$

then the following Hölder stability estimate holds:

$$\|u\|_{H^1(\Omega_{h+3\varepsilon})} \leq C_1\big(1 + \|u\|_{H^1(\Omega_h)}\big)\delta^{\beta}, \quad \forall \delta \in (0, \delta_0). \tag{2.17}$$

Proof. In this proof, $C = C(\varepsilon, K, \Omega_h)$ and $C_1 = C_1(\varepsilon, m, B, K, \Omega_h)$ denote different positive constants depending only on listed parameters. Since $\Omega_{h+3\varepsilon} \subset \Omega_{h+2\varepsilon} \subset \Omega_{h+\varepsilon} \subset \Omega_h$ and $\Omega_{h+3\varepsilon} \neq \varnothing$, then $\Omega_{h+2\varepsilon}, \Omega_{h+\varepsilon}, \Omega_h \neq \varnothing$. Consider a function $\chi(x)$ such that

$$\chi \in C^2(\overline{\Omega}_h), \chi(x) = \begin{cases} 1, & x \in \Omega_{h+2\varepsilon}, \\ 0, & x \in \Omega_h \diagdown \Omega_{h+\varepsilon}, \\ \in [0,1], & x \in \Omega_{h+\varepsilon} \diagdown \Omega_{h+2\varepsilon}. \end{cases} \tag{2.18}$$

The existence of such functions is well known from the real analysis course. Let the function v be

$$v = \chi u. \tag{2.19}$$

Using (2.14), (2.15), and (2.18), we obtain

$$|A_0 v| \leq C_1\left[|\nabla v| + |v| + |\nabla \chi||\nabla u| + \left(\sum_{|\alpha|=2} |D^{\alpha}\chi|\right)|u| + |f|\right], \quad \forall x \in \Omega_h, \tag{2.20}$$

$$v|_{\Gamma_h} = \chi g_0, \quad \partial_n v|_{\Gamma_h} = g_0 \partial_n \chi + \chi g_1, \tag{2.21}$$

$$v(x) = 0, \quad x \in \Omega_h \diagdown \Omega_{h+\varepsilon}. \tag{2.22}$$

Square both sides of (2.20). Next, multiply by $\varphi^2(x)$ and apply (2.10). We obtain

$$\begin{aligned} &C_1 f^2 \varphi^2(x) + C_1 |\nabla \chi|^2 |\nabla u|^2 \varphi^2(x) + C_1\left(\sum_{|\alpha|=2} |D^{\alpha}\chi|^2\right)|u|^2 \varphi^2(x) - \operatorname{div} U \\ &\geq C\lambda\left(1 - \frac{C_1}{\lambda}\right)(\nabla v)^2 \varphi^2(x) + C\lambda^3\left(1 - \frac{C_1}{\lambda^3}\right)v^2 \varphi^2(x), \\ &\quad \forall \lambda > \lambda_0, \forall x \in \Omega_h. \end{aligned}$$

Choose $\lambda_1 \geq \max(\lambda_0, 2C_1)$ so large that $C_1/\lambda_1 < 1/2$. We obtain

$$C_1 f^2\varphi^2(x) + C_1|\nabla\chi|^2\varphi^2(x)|\nabla u|^2 + C_1\left(\sum_{|\alpha|=2}|D^\alpha\chi|^2\right)\varphi^2(x)|u|^2 - \operatorname{div} U$$
$$\geq C\lambda(\nabla v)^2\varphi^2(x) + C\lambda^3 v^2\varphi^2(x), \quad \forall\lambda > \lambda_1, \forall x \in \Omega_h.$$

Integrate this inequality over Ω_h. Then the Gauss formula, (2.3), (2.6), (2.12), (2.18), (2.21), and (2.22) give

$$\begin{aligned} & C_1 e^{2\lambda m}\int_{\Omega_h} f^2 dx + C_1\lambda^3 e^{2\lambda m}\int_{\Gamma_h}[(\nabla g_0)^2 + g_1^2]dS_x \\ &\quad + C_1\exp[2\lambda(h+2\varepsilon)]\int_{\Omega_{h+\varepsilon}\setminus\Omega_{h+2\varepsilon}}(|\nabla u|^2 + u^2)dx \\ &\geq \lambda\int_{\Omega_{h+\varepsilon}}(\nabla v)^2\varphi^2 dx + \lambda^3\int_{\Omega_{h+\varepsilon}} v^2\varphi^2 dx. \end{aligned} \tag{2.23}$$

Since $\Omega_{h+3\varepsilon} \subset \Omega_{h+2\varepsilon} \subset \Omega_h$, then (2.23), (2.18), and (2.19) lead to

$$\begin{aligned} & C_1 e^{2\lambda m}\int_{\Omega_h} f^2 dx + C_1\lambda^3 e^{2\lambda m}\int_{\Gamma_h}[(\nabla g_0)^2 + g_1^2]dS_x \\ &\quad + C_1\exp[2\lambda(h+2\varepsilon)]\int_{\Omega_{h+\varepsilon}\setminus\Omega_{h+2\varepsilon}}(|\nabla u|^2 + u^2)dx \\ &\geq \lambda\int_{\Omega_{h+3\varepsilon}}(\nabla u)^2\varphi^2 dx + \lambda^3\int_{\Omega_{h+3\varepsilon}} u^2\varphi^2 dx \\ &\geq \lambda\exp[2\lambda(h+3\varepsilon)]\int_{\Omega_{h+3\varepsilon}}[(\nabla u)^2 + u^2]dx. \end{aligned}$$

Hence, we have established that

$$\begin{aligned} & C_1 e^{2\lambda m}\int_{\Omega_h} f^2 dx + C_1\lambda^3 e^{2\lambda m}\int_{\Gamma_h}[(\nabla g_0)^2 + g_1^2]dS_x \\ &\quad + C_1\exp[2\lambda(h+2\varepsilon)]\|u\|^2_{H^1(\Omega_h)} \\ &\geq \lambda\exp[2\lambda(h+3\varepsilon)]\|u\|^2_{H^1(\Omega_{h+3\varepsilon})}. \end{aligned}$$

Divide both sides of this inequality by $\lambda\exp[2\lambda(h+3\varepsilon)]$. Hence, there exists a number $\lambda_2 = \lambda_2(\varepsilon, m, B, K, G_c) > \lambda_1$ such that

$$\begin{aligned} & \left[\int_{\Omega_h} f^2 dx + \int_{\Gamma_h}[(\nabla g_0)^2 + g_1^2]dS_x\right]C_1 e^{3\lambda m} + C_1\exp[-2\lambda\varepsilon]\|u\|^2_{H^1(\Omega_h)} \\ &\geq \|u\|^2_{H^1(\Omega_{h+3\varepsilon})}, \quad \forall\lambda > \lambda_2. \end{aligned} \tag{2.24}$$

Using (2.16) and (2.24), we obtain

$$\|u\|^2_{H^1(\Omega_{h+3\varepsilon})} \le C_1\big(\delta^2 e^{3\lambda m} + e^{-2\lambda\varepsilon}\|u\|^2_{H^1(\Omega_h)}\big). \tag{2.25}$$

The idea now is to balance two terms in the right-hand side of (2.25). To do so, we choose $\lambda = \lambda(\delta)$ such that

$$\delta^2 e^{3\lambda m} = e^{-2\lambda\varepsilon}.$$

Hence,

$$\lambda = \ln\big(\delta^{-2(3m+2\varepsilon)^{-1}}\big). \tag{2.26}$$

Choose the number $\delta_0 = \delta_0(\varepsilon, m, B, K, G_c)$ so small that $\ln(\delta_0^{-2(3m+2\varepsilon)^{-1}}) > \lambda_2$. Then (2.25) and (2.26) imply (2.17). □

Theorem 2.2.2 (uniqueness). *Assume that conditions of Theorem* 2.2.1 *are valid, in* (2.15) $g_0(x) \equiv g_1(x) \equiv 0$, $x \in \Gamma_h$ *and also* $f(x) \equiv 0$. *Then* $u(x) \equiv 0$ *for* $x \in \Omega_h$.

This theorem follows immediately from Theorem 2.2.1 if setting in it $\delta = 0$.

We now replace the pointwise inequality (2.14) with the following integral inequality:

$$\int_{\Omega_h} (Au)^2 dx \le S^2, \tag{2.27}$$

where S is a certain number.

Theorem 2.2.3. *Let the function* $u \in H^2(\Omega_h)$ *satisfy inequality* (2.27) *and* $u|_{\Gamma_h} = \partial_n u|_{\Gamma_h} = 0$. *Assume that conditions* (2.9) *hold and that the Carleman estimate of Definition* 2.1.1 *is valid. Suppose that there exists a sufficiently small number* $\varepsilon > 0$ *such that the domain* $\Omega_{h+3\varepsilon} \neq \varnothing$. *Denote* $m = \max_{\overline{\Omega}_h} \psi(x)$. *Define the number* $\beta = 2\varepsilon/(3m+2\varepsilon) \in (0,1)$. *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, m, A, \Omega_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, m, A, \Omega_h) > 0$ *such that if* $\delta \in (0, \delta_0)$ *and* $S \in (0, \delta)$. *Then the following Hölder stability estimate holds:*

$$\|u\|_{H^1(\Omega_{h+3\varepsilon})} \le C_1\big(1 + \|u\|_{H^1(\Omega_h)}\big)\delta^\beta, \quad \forall \delta \in (0, \delta_0).$$

Proof. Assume first that the function $u \in C^2(\overline{\Omega}_h)$. We have

$$S^2 e^{2\lambda m} \ge \int_{\Omega_h} (Au)^2 \varphi^2(x) dx \ge \int_{\Omega_h} (A_0 u)^2 \varphi^2(x) dx - C_1 \int_{\Omega_h} \big((\nabla u)^2 + u^2\big)\varphi^2(x) dx.$$

This is equivalent with

$$S^2 e^{2\lambda m} + C_1 \int_{\Omega_h} \big((\nabla u)^2 + u^2\big)\varphi^2(x) dx \ge \int_{\Omega_h} (A_0 u)^2 \varphi^2(x) dx.$$

The rest of the proof is similar with the proof of Theorem 2.2.1. To replace $u \in C^2(\overline{\Omega}_h)$ with $u \in H^2(\Omega_h)$, density arguments should be used. □

2.3 Carleman estimate for the parabolic operator

For any $x \in \mathbb{R}^n$, denote $\overline{x} = (x_2, \dots, x_n)$ and $u_i = \partial_{x_i} u$. Let numbers $\alpha, h \in (0,1)$, and $\alpha < h$. Let $X, T > 0$ be two numbers. Consider the function $\psi(x,t)$,

$$\psi(x,t) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha. \tag{2.28}$$

Slightly abusing notation of Section 2.1, define the domain Ω_h as

$$\begin{aligned} \Omega_h &= \{(x,t) : \psi(x,t) < h, x_1 > 0\} \\ &= \left\{x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h, x_1 > 0\right\}. \end{aligned} \tag{2.29}$$

Let $\lambda, \nu > 1$ be two large parameters, which we will choose later. Consider the function $\varphi(x,t)$,

$$\varphi(x,t) = \exp(\lambda \psi^{-\nu}). \tag{2.30}$$

Here, $\varphi(x,t)$ is the CWF for our parabolic operator L defined below. To simplify notation, we use the notation $\varphi(x,t)$ instead of $\varphi_{\lambda,\nu}(x,t)$. Hence, the boundary of the domain Ω_h consists of a piece of the hyperplane $\{x_1 = 0\}$ and a piece of the paraboloid $\{\psi(x,t) = h, x_1 > 0\}$,

$$\partial\Omega_h = \partial_1\Omega_h \cup \partial_2\Omega_h, \tag{2.31}$$

$$\partial_1\Omega_h = \left\{x_1 = 0, \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h\right\}, \tag{2.32}$$

$$\partial_2\Omega_h = \left\{x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha = h\right\}. \tag{2.33}$$

Since by Lemma 2.1.1 the Carleman estimate for a PDO depends only on the principal part of this operator, then we consider only the principal part of an arbitrary elliptic operator L of the second order in the domain Ω_h,

$$Lu = u_t - L^{\text{ell}} u = u_t - \sum_{i,j=1}^{n} a^{ij}(x,t) u_{x_i x_j}. \tag{2.34}$$

To ensure the ellipticity of the operator L^{ell} in (2.34), we assume that

$$a^{ij}(x,t) = a^{ji}(x,t) \tag{2.35}$$

and also that there exist two numbers $\mu_1, \mu_2 > 0$ such that

$$\mu_1 |\xi|^2 \le \sum_{i,j=1}^{n} a^{ij}(x,t)\xi_i \xi_j \le \mu_2 |\xi|^2, \quad \forall (x,t) \in \overline{\Omega}_h, \forall \xi \in \mathbb{R}^n. \tag{2.36}$$

We also assume that

$$a^{ij} \in C^1(\overline{\Omega}_h), \quad K_0 = \max_{i,j} \|a_{ij}\|_{C^1(\overline{\Omega}_h)}. \tag{2.37}$$

By Definition 2.1.1, we need now to estimate $(Lu)^2\varphi^2$. Below $O(1/\lambda)$, $O(1/\nu)$ denote different $C^1(\overline{\Omega}_h)$-functions, which are independent on the function u, and such that

$$\left|O\left(\frac{1}{\lambda}\right)\right| \le \frac{C}{\lambda}, \quad \left|O\left(\frac{1}{\nu}\right)\right| \le \frac{C}{\nu}, \quad \forall \lambda, \nu \ge 1.$$

Here and below in this section, $C = C(\mu_1, \mu_2, K_0, \Omega_h)$ denotes different positive constants depending only on listed parameters. Below in this chapter, $f_i = f_{x_i}$ for any appropriate function f.

Lemma 2.3.1. *Suppose that conditions* (2.35)–(2.37) *are satisfied. Then there exist sufficiently large numbers* $\lambda_0 = \lambda_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$, $\nu_0 = \nu_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$ *depending only on listed parameters such that for every function* $u \in C^{2,1}(\overline{\Omega}_h)$ *the following estimate holds for all* $\lambda \ge \lambda_0$, $\nu \ge \nu_0$, $(x,t) \in \Omega_h$:

$$(Lu)^2\psi^{\nu+2}\varphi^2 \ge -C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div}\widetilde{U} + (\widetilde{V})_t, \tag{2.38}$$

$$|\widetilde{U}| + |\widetilde{V}| \le C\lambda^3\nu^3\psi^{-2\nu-2}((\nabla u)^2 + u^2)\varphi^2, \tag{2.39}$$

where the constant $C = C(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 0$ *depends only on listed parameters.*

Proof. Introduce the new function $v = u\varphi$ and express derivatives of the function u via derivatives of the function v, using (2.28) and (2.30). We have $u = v\exp(-\lambda\psi^{-\nu})$. Hence,

$$u_t = \left(v_t + \frac{t}{T^2}\lambda\nu\psi^{-\nu-1}v\right)\exp(-\lambda\psi^{-\nu}),$$

$$u_i = (v_i + \lambda\nu\psi^{-\nu-1}\psi_i v)\exp(-\lambda\psi^{-\nu}),$$

$$u_{ij} = \left[v_{ij} + \lambda\nu\psi^{-\nu-1}\psi_i v_j + \lambda\nu\psi^{-\nu-1}\psi_j v_i + \lambda^2\nu^2\psi^{-2\nu-2}\left(\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right)v\right].$$

Hence,

$$(Lu)^2\psi^{\nu+2}\varphi^2 = \left\{v_t - \sum_{i,j=1}^n a^{ij}v_{ij} - 2\lambda\nu\psi^{-\nu-1}\sum_{i,j=1}^n a^{ij}\psi_j v_i - \lambda^2\nu^2\psi^{-2\nu-2}\sum_{i,j=1}^n\left[\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]a^{ij}v\right\}^2\psi^{\nu+2}.$$

Denote

$$y_1 = v_t,$$

$$y_2 = -\sum_{i,j=1}^n a^{ij}v_{ij},$$

$$y_3 = -2\lambda\nu\psi^{-\nu-1}\sum_{i,j=1}^{n} a^{ij}\psi_j v_i,$$
$$y_4 = -\lambda^2\nu^2\psi^{-2\nu-2}\sum_{i,j=1}^{n}\left[\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]a^{ij}v.$$

Then

$$\begin{aligned}(Lu)^2\psi^{\nu+2}\varphi^2 &= (y_1+y_2+y_3+y_4)^2\psi^{\nu+2} \\ &\geq (y_1^2+y_3^2+2y_1y_2+2y_1y_3)\psi^{\nu+2} + 2y_2y_3\psi^{\nu+2} + 2y_3y_4\psi^{\nu+2} + 2y_1y_4\psi^{\nu+2}.\end{aligned} \tag{2.40}$$

We will estimate from the below terms in the second line of (2.40) in several steps.

Step 1. Estimate $2y_1y_2\psi^{\nu+2}$ below:

$$\begin{aligned}2y_1y_2\psi^{\nu+2} &= -2\sum_{i,j=1}^{n} a^{ij}v_{ij}v_t\psi^{\nu+2} = -\sum_{i,j=1}^{n}(a^{ij}v_{ij}v_t + a^{ij}v_{ji}v_t)\psi^{\nu+2} \\ &= \sum_{i,j=1}^{n}[(-a^{ij}\psi^{\nu+2}v_iv_t)_j + (-a^{ij}\psi^{\nu+2}v_jv_t)_i] \\ &\quad + \sum_{i,j=1}^{n}[a^{ij}\psi^{\nu+2}(v_iv_{tj}+v_jv_{ti})] + \sum_{i,j=1}^{n}[(a^{ij}\psi^{\nu+2})_jv_i + (2a^{ij}\psi^{\nu+2})_iv_j]v_t \\ &= \sum_{i,j=1}^{n} a^{ij}\psi^{\nu+2}(v_iv_j)_t + \sum_{i,j=1}^{n}[(a^{ij}\psi^{\nu+2})_jv_i + (2a^{ij}\psi^{\nu+2})_iv_j]v_t + \operatorname{div} U_1 \\ &= \left(\sum_{i,j=1}^{n} a^{ij}\psi^{\nu+2}v_iv_j\right)_t - \sum_{i,j=1}^{n}(a^{ij}\psi^{\nu+2})_tv_iv_j \\ &\quad + y_1\sum_{i,j=1}^{n}[(a^{ij}\psi^{\nu+2})_jv_i + (2a^{ij}\psi^{\nu+2})_iv_j] + \operatorname{div} U_1.\end{aligned}$$

Thus,

$$\begin{aligned}2y_1y_2\psi^{\nu+2} &= -\sum_{i,j=1}^{n}(a^{ij}\psi^{\nu+2})_tv_iv_j + y_1\sum_{i,j=1}^{n}[(a^{ij}\psi^{\nu+2})_jv_i + (2a^{ij}\psi^{\nu+2})_iv_j] \\ &\quad + \operatorname{div} U_1 + (V_1)_t,\end{aligned} \tag{2.41}$$

where

$$\operatorname{div} U_1 = \sum_{i,j=1}^{n}[(-a^{ij}\psi^{\nu+2}v_iv_t)_j + (-a^{ij}\psi^{\nu+2}v_jv_t)_i], \tag{2.42}$$

$$(V_1)_t = \left(\sum_{i,j=1}^{n} a^{ij}\psi^{\nu+2}v_iv_j\right)_t. \tag{2.43}$$

By (2.28),

$$\psi_1 = 1, \tag{2.44}$$

$$\psi_i = \frac{x_i}{X^2}, \quad \psi_t = \frac{t}{T^2}. \tag{2.45}$$

Hence,

$$-\sum_{i,j=1}^{n} (a^{ij}\psi^{\nu+2})_t v_i v_j \geq -C\nu|\nabla v|^2.$$

Next,

$$2y_1 \sum_{i,j=1}^{n} (a^{ij}\psi^{\nu+2})_j v_i = 2y_1(\nu+2) \sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{\nu+1} + O\left(\frac{1}{\nu}\right)\right] v_i.$$

Thus,

$$2y_1y_2\psi^{\nu+2} \geq -C\nu|\nabla v|^2 + 2y_1(\nu+2)\psi^{\nu+2} \sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right] v_i + \operatorname{div} U_1 + (V_1)_t. \tag{2.46}$$

Step 2. Using (2.46), estimate $(y_1^2 + 2y_1y_2 + 2y_1y_3 + y_3^2)\psi^{\nu+2}$.

$$\begin{aligned} &(y_1^2 + 2y_1y_2 + 2y_1y_3 + y_3^2)\psi^{\nu+2} \\ &\quad \geq (y_1^2 + y_3^2)\psi^{\nu+2} - C\nu|\nabla v|^2 \\ &\quad\quad + 2y_1(\nu+2)\psi^{\nu+2}\left\{\sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right] v_i + \frac{y_3}{\nu+2}\right\} \\ &\quad\quad + \operatorname{div} U_1 + (V_1)_t. \end{aligned} \tag{2.47}$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned} &2y_1(\nu+2)\psi^{\nu+2}\left\{\sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right] v_i + \frac{y_3}{\nu+2}\right\} \\ &\quad \geq -y_1^2\psi^{\nu+2} - (\nu+2)^2\psi^{\nu+2}\left\{\sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right] v_i + \frac{y_3}{\nu+2}\right\}^2 \\ &\quad = -(y_1^2 + y_3^2)\psi^{\nu+2} - (\nu+2)^2\psi^{\nu+2}\left\{\sum_{i,j=1}^{n} \left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right] v_i\right\}^2 \\ &\quad\quad - 2(\nu+2)\psi^{\nu+1}y_3 \sum_{i,j=1}^{n} \left[a^{ij}\psi_j + O\left(\frac{1}{\nu}\right)\right] v_i \end{aligned} \tag{2.48}$$

$$
\begin{aligned}
&= -(y_1^2 + y_3^2)\psi^{\nu+2} - (\nu+2)^2\psi^{\nu+2}\left\{\sum_{i,j=1}^{n}\left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right]v_i\right\}^2 \\
&\quad + 4\lambda\nu(\nu+2)\sum_{i,j=1}^{n} a^{ij}\psi_j v_i\left(\sum_{i,j=1}^{n}\left[a^{ij}\psi_j + O\left(\frac{1}{\nu}\right)\right]v_i\right) \\
&\geq -(y_1^2 + y_3^2)\psi^{\nu+2} + 3\lambda\nu(\nu+2)\left(\sum_{i,j=1}^{n} a^{ij}\psi_j v_i\right)^2 - C\lambda\nu(\nabla v)^2.
\end{aligned}
$$

We have used here the fact that $\lambda\nu(\nu+2) > (\nu+2)^2$ for sufficiently large λ. Also,

$$
4\lambda\nu(\nu+2)\sum_{i,j=1}^{n} a^{ij}\psi_j v_i \sum_{i,j=1}^{n} O\left(\frac{1}{\nu}\right)v_i \geq -C\lambda\nu(\nabla v)^2.
$$

Thus, by (2.48), we have obtained that

$$
\begin{aligned}
&2y_1(\nu+2)\psi^{\nu+2}\left\{\sum_{i,j=1}^{n}\left[a^{ij}\psi_j\psi^{-1} + O\left(\frac{1}{\nu}\right)\right]v_i + \frac{y_3}{\nu+2}\right\} \\
&\quad \geq -(y_1^2 + y_3^2)\psi^{\nu+2} + 3\lambda\nu(\nu+2)\left(\sum_{i,j=1}^{n} a^{ij}\psi_j v_i\right)^2 - C\lambda\nu(\nabla v)^2.
\end{aligned}
$$

Substituting this in (2.47), we obtain

$$
(y_1^2 + 2y_1y_2 + 2y_1y_3 + y_3^2)\psi^{\nu+2} \geq -C\lambda\nu(\nabla v)^2 + \operatorname{div} U_1 + (V_1)_t. \tag{2.49}
$$

Step 3. Estimate $2y_2y_3\psi^{\nu+2}$,

$$
\begin{aligned}
2y_2y_3\psi^{\nu+2} &= 4\lambda\nu\psi\left(\sum_{k,l=1}^{n} a^{kl}\psi_l v_k\right)\left(\sum_{i,j=1}^{n} a^{ij}v_{ij}\right) \\
&= 2\lambda\nu\psi\left(\sum_{k,l=1}^{n} a^{kl}\psi_l v_k\right)\left(\sum_{i,j=1}^{n} a^{ij}(v_{ij} + v_{ji})\right) \\
&= 2\lambda\nu\sum_{k,l=1}^{n}\sum_{i,j=1}^{n} a^{kl}a^{ij}\psi\psi_l(v_{ij}v_k + v_{ji}v_k).
\end{aligned} \tag{2.50}
$$

We have

$$
\begin{aligned}
v_{ij}v_k + v_{ji}v_k &= (v_iv_k)_j + (v_jv_k)_i - v_iv_{kj} - v_{ki}v_j \\
&= (v_iv_k)_j + (v_jv_k)_i + (-v_iv_j)_k.
\end{aligned}
$$

Hence, the term in the last line of (2.50) can be evaluated as

$$
\begin{aligned}
&a^{kl}a^{ij}\psi\psi_l(v_{ij}v_k + v_{ji}v_k) \\
&\quad = (a^{kl}a^{ij}\psi\psi_l v_i v_k)_j - (a^{kl}a^{ij}\psi\psi_l)_j v_i v_k
\end{aligned}
$$

$$+ (a^{kl}a^{ij}\psi\psi_l v_j v_k)_i - (a^{kl}a^{ij}\psi\psi_l)_i v_j v_k + (-a^{kl}a^{ij}\psi\psi_l v_i v_j)_k + (a^{kl}a^{ij}\psi\psi_l)v_i v_j$$
$$\geq -C(\nabla v)^2 + \operatorname{div} U_2.$$

Thus, (2.50) leads to

$$2y_2 y_3 \psi^{\nu+2} \geq -C\lambda\nu(\nabla v)^2 + \operatorname{div} U_2, \tag{2.51}$$

$$\operatorname{div} U_2 = 2\lambda\nu \sum_{k,l=1}^{n}\sum_{i,j=1}^{n}((a^{kl}a^{ij}\psi\psi_l v_i v_k)_j + (a^{kl}a^{ij}\psi\psi_l v_j v_k)_i). \tag{2.52}$$

Step 4. We now estimate $2y_3 y_4 \psi^{\nu+2}$ in (2.40),

$$\begin{aligned} 2y_3 y_4 \psi^{\nu+2} &= 4\lambda^3\nu^3\psi^{-2\nu-1}\left(\sum_{k,l=1}^{n} a^{kl}\psi_l v_k\right)\sum_{i,j=1}^{n}\left[a^{ij}\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]v \\ &= \left(2\lambda^3\nu^3\sum_{i,j=1}^{n}\left[a^{ij}\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]\cdot\sum_{k,l=1}^{n}\psi^{-2\nu-1}a^{kl}\psi_l(v^2)_k\right) \\ &= \sum_{k,l=1}^{n}\left(\left[2\lambda^3\nu^3\sum_{i,j=1}^{n}\left[a^{ij}\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]\cdot\psi^{-2\nu-1}a^{kl}\psi_l\right]v^2\right)_k \\ &\quad + 2\lambda^3\nu^3(2\nu+1)\psi^{-2\nu-2}v^2\sum_{i,j=1}^{n}\left[a^{ij}\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]\sum_{k,l=1}^{n}\left(a^{kl}\psi_k\psi_l + O\left(\frac{1}{\nu}\right)\right). \end{aligned} \tag{2.53}$$

It follows from (2.28) and (2.36) that

$$\sum_{i,j=1}^{n} a^{ij}\psi_i\psi_j \geq \mu_1|\nabla\psi|^2 \geq \mu_1.$$

Hence, (2.53) leads to

$$2y_3 y_4 \psi^{\nu+2} \geq C\lambda^3\nu^4\psi^{-2\nu-2}v^2 + \operatorname{div} U_3, \tag{2.54}$$

$$\operatorname{div} U_3 = \sum_{k,l=1}^{n}\left(\left[2\lambda^3\nu^3\sum_{i,j=1}^{n}\left(a^{ij}\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right)\cdot\psi^{-2\nu-1}a^{kl}\psi_l\right]v^2\right)_k. \tag{2.55}$$

Step 5. Similarly, with (2.55) we obtain

$$2y_1 y_4 \psi^{\nu+2} \geq -C\lambda^2\nu^3\psi^{-2\nu-2}v^2 + (V_2)_t, \tag{2.56}$$

$$V_2 = -2\lambda^2\nu^2\psi^{-\nu}\sum_{i,j=1}^{n} a^{ij}\left[\psi_i\psi_j + O\left(\frac{1}{\lambda}\right)\right]v^2. \tag{2.57}$$

Since $\nu \geq \nu_0(\mu_1, \mu_2, K_0, \Omega_h) > 1$, $\lambda \geq \lambda_0(\mu_1, \mu_2, K_0, \Omega_h) > 1$, where numbers ν_0 and λ_0 are sufficiently large, then $\nu^4 > \nu^3$ and $\lambda^3 > \lambda^2$. Hence, (2.54)–(2.57) imply that

$$2y_3 y_4 \psi^{\nu+2} + 2y_1 y_4 \psi^{\nu+2} \geq C\lambda^3\nu^4\psi^{-2\nu-2}v^2 + \operatorname{div} U_3 + (V_2)_t. \tag{2.58}$$

Finally, summing up (2.49) and (2.58), replacing v with $u = v\varphi^{-1}$ and using (2.42), (2.43), (2.55), (2.52), and (2.57) for estimating functions under signs of div and ∂_t, we obtain (2.38) and (2.39). □

Lemma 2.3.2. *Suppose that conditions* (2.35)–(2.37) *are satisfied. Then there exist sufficiently large numbers* $\lambda_0 = \lambda_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$, $\nu_0 = \nu_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$ *depending only on listed parameters such that for every function* $u \in C^{2,1}(\overline{\Omega}_h)$ *the following estimate holds for all* $\lambda \geq \lambda_0$, $\nu \geq \nu_0$, $(x, t) \in \Omega_h$:

$$(Lu)u\varphi^2 \geq \mu_1(\nabla u)^2\varphi^2 - C\lambda^2\nu^2\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U_4 + (V_3)_t, \tag{2.59}$$

$$|U_4| + |V_3| \leq C\lambda\nu\psi^{-\nu-1}((\nabla u)^2 + u^2), \tag{2.60}$$

Proof. We have

$$\begin{aligned}
(Lu)u\varphi^2 = (u_t - L^{\mathrm{ell}}u)u\varphi^2 &= u_t u \exp(2\lambda\psi^{-\nu}) - \sum_{i,j=1}^{n} a^{ij}u_{ij}u\exp(2\lambda\psi^{-\nu}) \\
&= \left(\frac{1}{2}u^2\exp(2\lambda\psi^{-\nu})\right)_t + \lambda\nu\psi_t\psi^{-\nu-1}u^2\varphi^2 \\
&\quad + \left(-\sum_{i,j=1}^{n} a^{ij}u_i u\exp(2\lambda\psi^{-\nu})\right)_j + \sum_{i,j=1}^{n} a^{ij}u_i u_j\varphi^2 \\
&\quad - 2\lambda\nu\psi^{-\nu-1}\sum_{i,j=1}^{n} a^{ij}u_i u\left(\psi_j + O\left(\frac{1}{\lambda}\right)\right)\exp(2\lambda\psi^{-\nu}) \\
&\geq \mu_1|\nabla u|^2\varphi^2 + \left(\frac{1}{2}u^2\exp(2\lambda\psi^{-\nu})\right)_t + \lambda\nu\psi_t\psi^{-\nu-1}u^2\varphi^2 \\
&\quad + \left(-\lambda\nu\psi^{-\nu-1}\sum_{i,j=1}^{n} a^{ij}u^2\left(\psi_j + O\left(\frac{1}{\lambda}\right)\right)\exp(2\lambda\psi^{-\nu})\right)_i \\
&\quad - 2\lambda^2\nu^2\psi^{-2\nu-2}\sum_{i,j=1}^{n} a^{ij}u^2\psi_i\left(\psi_j + O\left(\frac{1}{\lambda}\right)\right) \\
&\quad - \lambda\nu\sum_{i,j=1}^{n}\left(\psi^{-\nu-1}a^{ij}\left(\psi_j + O\left(\frac{1}{\lambda}\right)\right)\right)_i u^2 \\
&\geq \mu_1|\nabla u|^2\varphi^2 - C\lambda^2\nu^2\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U_4 + (V_3)_t,
\end{aligned}$$

which is (2.59). Estimates (2.60) easily follow from the above formulae. □

Lemmata 2.3.1 and 2.3.2 enable us to prove Theorem 2.3.1, which is the Carleman estimate for the operator $\partial_t - L_0^{\mathrm{ell}}$.

Theorem 2.3.1 (Carleman estimate for the parabolic operator). *Suppose that conditions* (2.35)–(2.37) *are satisfied. Then one can choose sufficiently large numbers* $\lambda_0 = \lambda_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$, $\nu_0 = \nu_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$ *such that for every function*

$u \in C^{2,1}(\overline{\Omega}_h)$ the following pointwise Carleman estimate holds for all $\lambda \geq \lambda_0$, $\nu \geq \nu_0$, $(x,t) \in \Omega_h$:

$$(L_0u)^2\varphi^2 \geq C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U + (V)_t, \tag{2.61}$$

$$|U| + |V| \leq C\lambda^3\nu^3\psi^{-2\nu-2}((\nabla u)^2 + u^2)\varphi^2, \tag{2.62}$$

where the constant $C = C(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 0$ depends only on listed parameters.

Proof. Multiply (2.59) by $2C/\mu_1$ and sum up with (2.38). We obtain

$$\begin{aligned}
&\frac{2C}{\mu_1}\lambda\nu(Lu)u\varphi^2 + (Lu)^2\psi^{\nu+2}\varphi^2\\
&\quad \geq C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 - 2C^2\mu_1^{-1}\lambda^3\nu^3\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U + V_t\\
&\quad = C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}\left(1 + O\left(\frac{1}{\nu}\right)\right)u^2\varphi^2 + \operatorname{div} U + V_t\\
&\quad \geq C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U + V_t.
\end{aligned} \tag{2.63}$$

Next, since $\psi^{\nu+2} < 1$, then

$$\frac{2C}{\mu_1}\lambda\nu(Lu)u\varphi^2 + (Lu)^2\psi^{\nu+2}\varphi^2 \leq C(Lu)^2\varphi^2 + \lambda^2\nu^2u^2\varphi^2.$$

Comparing this with (2.63), we obtain

$$C(L_0u)^2\varphi^2 + \lambda^2\nu^2u^2\varphi^2 \geq C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U + V_t. \tag{2.64}$$

Since

$$\lambda^3\nu^4\psi^{-2\nu-2} - \lambda^2\nu^2 = \lambda^3\nu^4\psi^{-2\nu-2}\left(1 + O\left(\frac{1}{\lambda}\right)O\left(\frac{1}{\nu^2}\right)\right) \geq \frac{1}{2}\lambda^3\nu^4\psi^{-2\nu-2},$$

then (2.64) implies (2.61). Estimate (2.62) follows from estimates (2.39) and (2.60). □

2.4 Carleman estimate for the elliptic operator

We consider the same case as in Section 2.3 with the only difference that functions are independent on t in this section.

Let numbers $\alpha, h \in (0,1)$, $\alpha < h$ and let $X > 0$ be a number. Consider the function $\psi(x)$,

$$\psi(x) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha.$$

The domain Ω_h now is

$$\Omega_h = \{x : \psi(x) < h, x_1 > 0\} = \left\{x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha < h, x_1 > 0\right\}. \tag{2.65}$$

Again, for two large parameters $\lambda, \nu > 1$ consider the function $\varphi(x)$,

$$\varphi(x) = \exp(\lambda\psi^{-\nu}), \tag{2.66}$$

where $\varphi(x)$ is the CWF in the elliptic case.

Again, the boundary of the domain Ω_h consists of a piece of the hyperplane $\{x_1 = 0\}$ and a piece of the paraboloid $\{\psi(x) = h, x_1 > 0\}$,

$$\partial\Omega_h = \partial_1\Omega_h \cup \partial_2\Omega_h, \tag{2.67}$$

$$\partial_1\Omega_h = \left\{x_1 = 0, \frac{|\overline{x}|^2}{2X^2} + \alpha < h\right\}, \tag{2.68}$$

$$\partial_2\Omega_h = \left\{x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha = h\right\}. \tag{2.69}$$

Consider the elliptic operator L,

$$Lu = \sum_{i,j=1}^{n} a^{ij}(x)u_{ij}, \tag{2.70}$$

$$a^{ij}(x) = a^{ji}(x). \tag{2.71}$$

Just as in (2.36), (2.37), we assume that there exist two numbers $\mu_1, \mu_2 > 0$ such that

$$\mu_1|\xi|^2 \leq \sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \leq \mu_2|\xi|^2, \quad \forall x \in \overline{\Omega}_h, \forall \xi \in \mathbb{R}^n \tag{2.72}$$

and also

$$a^{ij} \in C^1(\overline{\Omega}_h), \quad K_0 = \max_{i,j} \|a_{ij}\|_{C^1(\overline{\Omega}_h)}. \tag{2.73}$$

Theorem 2.4.1 (Carleman estimate for the elliptic operator). *Suppose that conditions (2.70)–(2.73) are satisfied. Then there exist sufficiently large numbers $\lambda_0 = \lambda_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$ and $\nu_0 = \nu_0(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 1$ depending only on listed parameters such that for every function $u \in C^{2,1}(\overline{\Omega}_h)$ the following pointwise Carleman estimate holds for all $\lambda \geq \lambda_0$, $\nu \geq \nu_0$, $x \in \Omega_h$:*

$$(L_0^{\text{ell}}u)^2\varphi^2 \geq C\lambda\nu(\nabla u)^2\varphi^2 + C\lambda^3\nu^4\psi^{-2\nu-2}u^2\varphi^2 + \operatorname{div} U, \tag{2.74}$$

$$|U| \leq C\lambda^3\nu^3\psi^{-2\nu-2}((\nabla u)^2 + u^2)\varphi^2, \tag{2.75}$$

where the constant $C = C(\mu_1, \mu_2, K_0, X, T, \Omega_h) > 0$ depends only on listed parameters.

Proof. In Theorem 2.3.1, set $u_t \equiv 0$ and all functions to be independent on t. Then use (2.67)–(2.73). Then we obtain (2.74)–(2.75). □

2.5 Carleman estimate for a hyperbolic operator

While we have derived above Carleman estimates for arbitrary parabolic and elliptic operators of the second order, in the case of a hyperbolic operator, conditions imposed on its coefficient are more restrictive; see (2.81)–(2.84). Nevertheless, a sufficiently large class of hyperbolic operators is covered.

Our derivation here is similar with the one of Section 1.10.2 of the book [22]. For simplicity, we assume here that $\Omega = \{|x| < R\} \subset \mathbb{R}^n$, although the case of a general convex domain also works. Let $T = \text{const.} > 0$. Let $x_0 \in \Omega$, $\eta \in (0,1)$. Introduce the function $\psi(x,t)$ as

$$\psi(x,t) = |x - x_0|^2 - \eta t^2. \tag{2.76}$$

The Carleman Weight Function (CWF) is

$$\varphi(x,t) = \exp[\lambda\psi(x,t)], \tag{2.77}$$

where $\lambda > 1$ is a large parameter, which we will be specified later. For $h > 0$, consider the hyperboloid

$$\psi_h = \{|x - x_0|^2 - \eta t^2 = h\}.$$

Then ψ_h is a level surface of the function $\psi(x,t)$. Obviously, $\nabla_x \psi(x,t) \neq 0$ in $\overline{G}_h$. For $h \in (0, R^2)$, consider the domain G_h,

$$G_h = \{(x,t) : x \in \Omega, |x - x_0|^2 - \eta t^2 > h\}. \tag{2.78}$$

We now prove that $G_h \neq \varnothing$. It is sufficient to prove that

$$\{x \in \Omega : |x - x_0| > \sqrt{h}\} \neq \varnothing. \tag{2.79}$$

Indeed, without a loss of generality, set $x_0 = (x_{01}, 0, \dots, 0)$, where $x_{01} \geq 0$. For a sufficiently small $\varepsilon > 0$, consider the point $y = (-R + \varepsilon, 0, \dots, 0)$. Then $y \in \Omega$. Hence, $|y - x_0| = R + x_{01} - \varepsilon$. Choose ε so small that $R + x_{01} - \varepsilon > \sqrt{h}$. Then the point y belongs to the domain (2.79). Thus,

$$G_h \neq \varnothing, \quad \forall h \in (0, R^2). \tag{2.80}$$

At this point of time, the Carleman estimate is known only for a special form of the hyperbolic operator,

$$Lu = c(x)u_{tt} - \Delta u. \tag{2.81}$$

The Carleman estimate for the operator L is established in Theorem 2.5.1.

Theorem 2.5.1. *Let* $\Omega = \{|x| < R\} \subset \mathbb{R}^n$*. Let* L *be the hyperbolic operator defined in* (2.81) *and* $\varphi(x,t)$ *be the function defined in* (2.77)*. Assume that the coefficient* $c(x)$ *is such that*

$$c(x) \in [1, \overline{c}], \quad \text{where } \overline{c} = \text{const.} \geq 1, \tag{2.82}$$

$$c \in C^1(\overline{\Omega}). \tag{2.83}$$

In addition, assume that there exists a point $x_0 \in \Omega$ *such that*

$$\big(x - x_0, \nabla c(x)\big) \geq 0, \quad \forall x \in \overline{\Omega}, \tag{2.84}$$

where $(\cdot,\cdot)$ *denotes the scalar product in* $\mathbb{R}^n$*. Then there exist a sufficiently small number* $\eta_0 = \eta_0(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R) \in (0,1)$ *such that for any* $\eta \in (0, \eta_0]$ *there exists a sufficiently large number* $\lambda_0 = \lambda_0(\Omega, \eta, c, x_0) > 1$ *and a number* $C = C(\Omega, \eta, c, x_0) > 0$ *such that for all* $u \in C^2(\overline{G}_h)$ *and all values of the parameter* $\lambda \geq \lambda_0$ *the following pointwise Carleman estimate holds:*

$$(L_0 u)^2 \varphi^2 \geq C\lambda(|\nabla u|^2 + u_t^2 + \lambda^2 u^2)\varphi^2 + \operatorname{div} U + V_t, \quad \text{in } G_h, \tag{2.85}$$

where components of the vector function (U, V) *can be estimated in* G_h *as*

$$|U| \leq C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \tag{2.86}$$

$$|V| \leq C\lambda^3\big[|t|(u_t^2 + |\nabla u|^2 + u^2) + (|\nabla u| + |u|)|u_t|\big]\varphi^2. \tag{2.87}$$

Corollary 2.5.1. *In particular,* (2.87) *implies that if either* $u(x,0) = 0$ *or* $u_t(x,0) = 0$*, then*

$$V(x,0) = 0. \tag{2.88}$$

Hence, the Carleman estimate (2.85) *in this case can be considered only in the domain* $G_h^+ = G_h \cap \{t > 0\}$.

Corollary 2.5.2. *Assume now that* $n \geq 2$ *and in* (2.81) *the operator* $Lu = u_{tt} - \Delta u$*. Then condition* (2.84) *holds automatically and one can choose* $\eta_0 = 1$ *in Theorem* 2.5.1.

Remark 2.5.1. A careful analysis of the proof of Theorem 2.5.1 shows that it is valid not only for the case when $x_0 \in \Omega$ but also in the case when $x_0 \in \mathbb{R}^n \setminus \Omega$.

Proof of Theorem 2.5.1. In this proof, $(x,t) \in G_h$ and C denotes different positive constants depending on the same parameters as indicated in the conditions of this theorem. Also, just as in the previous section, $O(1/\lambda)$ denotes different $C^1(\overline{Q}_T^{\pm})$-functions, such that

$$\left\| O\left(\frac{1}{\lambda}\right) \right\|_{C^1(\overline{Q}_T^{\pm})} \leq \frac{C}{\lambda}, \quad \forall \lambda > 1. \tag{2.89}$$

Denote $v = u \cdot \varphi$. Expressing derivatives of u via derivatives of v, we obtain

$$u = v \cdot \exp[\lambda(\eta t^2 - |x - x_0|^2)],$$

$$\begin{aligned}
u_t &= (v_t + 2\lambda\eta t \cdot v) \exp[\lambda(\eta t^2 - |x - x_0|^2)],\\
u_{tt} &= \left(v_{tt} + 4\lambda\eta t \cdot v_t + 4\lambda^2\left(\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right)v\right) \exp[\lambda(\eta t^2 - |x - x_0|^2)],\\
u_i &= [v_i - 2\lambda(x_i - x_{0i})v] \exp[\lambda(\eta t^2 - |x - x_0|^2)],\\
u_{ii} &= \left[v_{ii} - 4\lambda(x_i - x_{0i})v_i + 4\lambda^2\left(|x - x_0|^2 + O\left(\frac{1}{\lambda}\right)\right)v\right] \exp[\lambda(\eta t^2 - |x - x_0|^2)].
\end{aligned}$$

Hence,

$$\begin{aligned}
(Lu)^2\varphi^2 &= (c(x)u_{tt} - \Delta u)^2\varphi^2\\
&= \left\{\begin{array}{l} [c(x)v_{tt} - \Delta v - 4\lambda^2(|x - x_0|^2 - c\eta^2 t^2 + O(1/\lambda))v] + 4\lambda c\eta t v_t\\ \quad + 4\lambda \sum_{i=1}^{n}(x_i - x_{0i})v_i \end{array}\right\}^2.
\end{aligned}$$

Denote

$$\begin{aligned}
y_1 &= cv_{tt} - \Delta v - 4\lambda^2\left(|x - x_0|^2 - c\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right)v,\\
y_2 &= 4\lambda c\eta t \cdot v_t,\\
y_3 &= 4\lambda \sum_{i=1}^{n}(x_i - x_{0i})v_i.
\end{aligned}$$

Then $(Lu)^2\varphi^2 = (y_1 + y_2 + y_3)^2$. Hence,

$$(Lu)^2\varphi^2 \geq y_1^2 + 2y_1y_2 + 2y_1y_3. \tag{2.90}$$

In the following, each term in the inequality (2.90) is estimated from the below separately. We do this in five steps.

Step 1. Estimate the term $2y_1y_2$ in (2.90),

$$\begin{aligned}
2y_1y_2 &= 8\lambda c\eta t \cdot v_t\left[cv_{tt} - \Delta v - 4\lambda^2\left(|x - x_0|^2 - c\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right)v\right]\\
&= [4\lambda c^2\eta t \cdot v_t^2]_t - 4\lambda c^2\eta v_t^2\\
&\quad + \sum_{i=1}^{n}(-8\lambda c\eta t \cdot v_t v_i)_i + \sum_{i=1}^{n} 8\lambda c\eta t \cdot v_{it}v_i\\
&\quad + 8\lambda\eta t \cdot v_t \sum_{i=1}^{n} c_i v_i + \left[-16\lambda^3 c\eta\left(t|x - x_0|^2 - c\eta^2 t^3 + tO\left(\frac{1}{\lambda}\right)\right)v^2\right]_t\\
&\quad + 16\lambda^3 c\eta\left(|x - x_0|^2 - 3c\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right)v^2\\
&= -4\lambda c^2\eta v_t^2 + \left(4\lambda c^2\eta t \cdot v_t^2 + \sum_{i=1}^{n} 4\lambda c\eta t v_i^2\right)_t - 4\lambda c\eta|\nabla v|^2 + 8\lambda\eta t \cdot v_t \sum_{i=1}^{n} c_i v_i
\end{aligned}$$

$$+16\lambda^3 c\eta\left[|x-x_0|^2 - 3c\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2$$
$$+\operatorname{div} U_1 + \left[4\lambda c^2\eta t v_t^2 - 16\lambda^3 c\eta\left(t|x-x_0|^2 - c\eta^2 t^2 + tO\left(\frac{1}{\lambda}\right)\right)v^2\right]_t.$$

Thus, the above means that

$$2y_1y_2 = -4\lambda c\eta(cv_t^2 + |\nabla v|^2) + 8\lambda\eta t\cdot v_t\sum_{i=1}^n c_i v_i$$
$$+16\lambda^3 c\eta\left[|x-x_0|^2 - 3c\eta^2 t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \operatorname{div} U_1 + (V_1)_t, \tag{2.91}$$

where the following estimates hold for the vector function (U_1, V_1) :

$$|U_1| \le C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \tag{2.92}$$
$$|V_1| \le C\lambda^3|t|(u_t^2 + |\nabla u|^2 + u^2)\varphi^2. \tag{2.93}$$

To include the function u in the estimate for $|U_1|, |V_1|$, we have replaced in (2.91) v with $u = v\cdot\varphi^{-1}$.

Step 2. Estimate the term $2y_1y_3$ in (2.90). We have

$$\begin{aligned}
2y_1y_3 &= 8\lambda\sum_{i=1}^n(x_i - x_{0i})v_i\left[cv_{tt} - \Delta v - 4\lambda^2\left(|x-x_0|^2 - c\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right)v\right]\\
&= \left(\sum_{i=1}^n 8c\lambda(x_i - x_{0i})v_iv_t\right)_t - \sum_{i=1}^n 8\lambda(x_i - x_{0i})cv_{it}v_t\\
&\quad - \sum_{j=1}^n\sum_{i=1}^n 8\lambda(x_i - x_{0i})v_iv_{jj}\\
&\quad + \sum_{i=1}^n\left[-16\lambda^3(x_i - x_{0i})\left(|x-x_0|^2 - c\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right)v^2\right]_i\\
&\quad + 16\lambda^3\left[(n+2)|x-x_0|^2 - nc\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2\\
&= \sum_{i=1}^n(-4\lambda(x_i - x_{0i})cv_t^2)_i + 4\lambda[nc + (x - x_0, \nabla c)]v_t^2\\
&\quad + \sum_{j=1}^n\left[\sum_{i=1}^n(-8\lambda(x_i - x_{0i})v_iv_j)\right]_j + 8\lambda|\nabla v|^2 + \sum_{j=1}^n\sum_{i=1}^n 8\lambda(x_i - x_{0_i})v_{ij}v_j\\
&\quad + 16\lambda^3\left[(n+2)|x-x_0|^2 - nc\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \left(8c\lambda v_t\sum_{i=1}^n(x_i - x_{0i})v_i\right)_t\\
&= 4\lambda[nc + (x - x_0, \nabla c)]v_t^2 + 8\lambda|\nabla v|^2\\
&\quad + \sum_{i=1}^n\left[\sum_{j=1}^n 4\lambda(x_i - x_{0_i})v_j^2\right]_i - 4\lambda|\nabla v|^2
\end{aligned}$$

$$+16\lambda^3\left[(n+2)|x-x_0|^2 - c\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \nabla\cdot U_2 + (V_2)_t.$$

Here is the place where we use condition (2.84). This condition, combined with (2.82), implies that $nc + (x - x_0, \nabla c) \geq nc \geq n$. We obtain

$$\begin{aligned}
2y_1y_3 \geq{}& 4\lambda n v_t^2 + 4\lambda|\nabla v|^2 \\
&+ 16\lambda^3\left[(n+2)|x-x_0|^2 - nc\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 \\
&+ \operatorname{div} U_2 + (V_2)_t,
\end{aligned} \tag{2.94}$$

$$|U_2| \leq C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \tag{2.95}$$

$$|V_2| \leq C\lambda^3\big[|t|(|\nabla u|^2 + |u|^2) + (|\nabla u| + |u|)|u_t|\big]\varphi^2. \tag{2.96}$$

Step 3. We now estimate the term $2y_1y_2 + 2y_1y_3$. By the triangle inequality,

$$|x - x_0| \leq |x| + |x_0| < 2R, \quad \forall x \in \Omega.$$

On the other hand, since $|x-x_0|^2 - \eta t^2 > h > 0$ in G_h and $\eta \in (0,1)$, then $\eta|t| \leq 2R\sqrt{\eta}$ in G_h. This estimate combined with the Cauchy–Schwarz inequality leads to

$$\begin{aligned}
8\lambda\eta t\cdot v_t\sum_{i=1}^n c_iv_i &= -8\lambda\eta t\cdot v_t(\nabla c, \nabla v) \geq -8\lambda\eta|t|\cdot|v_t|\cdot|\nabla v|\cdot\|\nabla c\|_{C(\overline{\Omega})} \\
&\geq -8\lambda\sqrt{\eta}R\|\nabla c\|_{C(\overline{\Omega})}(v_t^2 + |\nabla v|^2).
\end{aligned} \tag{2.97}$$

By (2.82), $c, \overline{c} \geq 1$. Hence, (2.91) and (2.97) imply that

$$\begin{aligned}
2y_1y_2 \geq{}& -4\lambda\sqrt{\eta}(\overline{c}^2 + 2R\|\nabla c\|_{C(\overline{\Omega})})(v_t^2 + |\nabla v|^2) \\
&+ 16\lambda^3\eta\left[|x-x_0|^2 - 3\overline{c}\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \operatorname{div} U_1 + (V_1)_t.
\end{aligned} \tag{2.98}$$

Denote $U_3 = U_1 + U_2$, $V_3 = V_1 + V_2$. Hence, using (2.94)–(2.98), we obtain

$$\begin{aligned}
2y_1y_2 + 2y_1y_3 \geq{}& 4\lambda[1 - \sqrt{\eta}(\overline{c}^2 + 2R\|\nabla c\|_{C(\overline{\Omega})})](v_t^2 + |\nabla v|^2) \\
&+ 16\lambda^3\left[(n+2+\eta)|x-x_0|^2 - (n+3\eta)\overline{c}\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 \\
&+ \operatorname{div} U_3 + (V_3)_t,
\end{aligned} \tag{2.99}$$

$$|U_3| \leq C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \tag{2.100}$$

$$|V_3| \leq C\lambda^3\big[|t|(u_t^2 + |\nabla u|^2 + u^2) + (|\nabla u| + |u|)|u_t|\big]\varphi^2. \tag{2.101}$$

For sufficiently small $\eta_0 \in (0,1)$ and for $\eta \in (0, \eta_0)$,

$$1 - \sqrt{\eta}(\overline{c}^2 + 2R\|\nabla c\|_{C(\overline{\Omega})}) \geq \frac{1}{2}.$$

Hence, (2.99) implies for $\lambda \ge \lambda_0$,

$$2y_1y_2 + 2y_1y_3 \ge 2\lambda(v_t^2 + |\nabla v|^2) \\ + 8\lambda^3[(n+2)|x-x_0|^2 - (n+3\eta)\bar{c}\eta^2t^2]v^2 + \operatorname{div} U_3 + (V_3)_t.$$

Returning to the function u, we obtain (2.85)–(2.87). □

Proof of Corollary 2.5.2. We now assume that $c(x) \equiv 1$ and $\eta \in (0,1)$. We estimate the term y_1^2 in (2.90) from the below. The equality (2.91) becomes

$$2y_1y_2 = -4\lambda\eta(v_t^2 + |\nabla v|^2) \\ + 16\lambda^3\eta\left[|x-x_0|^2 - 3\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \operatorname{div} U_1 + (V_1)_t.$$

Next, (2.94) becomes

$$2y_1y_3 \ge 4\lambda n v_t^2 + 4\lambda|\nabla v|^2 \\ + 16\lambda^3\left[(n+2)|x-x_0|^2 - n\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 \\ + \operatorname{div} U_2 + (V_2)_t.$$

Hence, we now have

$$2y_1y_2 + 2y_1y_3 \ge 4\lambda(1-\eta)(v_t^2 + |\nabla v|^2) \\ + 16\lambda^3\left[(n+2+\eta)|x-x_0|^2 - (3\eta+n)\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \operatorname{div} U_3 + (V_3)_t. \tag{2.102}$$

Let $b > 0$ be a number, which we will be defined later. Then

$$y_1^2 = \left[v_{tt} - \Delta v - 4\lambda^2\left(|x-x_0|^2 - \eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right)v + \lambda b v\right]^2 \\ = (2\lambda b v v_t)_t - 2\lambda b v_t^2 + \sum_{i=1}^{n}(-2\lambda b v v_i)_i \\ + 2\lambda b|\nabla v|^2 - 8\lambda^3 b\left[|x-x_0|^2 - c\eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2.$$

Thus,

$$y_1^2 \ge 2\lambda b|\nabla v|^2 - 2\lambda b v_t^2 \\ - 8\lambda^3 b\left[|x-x_0|^2 - \eta^2t^2 + O\left(\frac{1}{\lambda}\right)\right]v^2 + \operatorname{div} U_4 + (V_4)_t, \tag{2.103}$$

$$|U_4| \le C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \tag{2.104}$$

$$|V_4| \le C\lambda^3(|t|u^2 + |u_t|\cdot|u|)\varphi^2. \tag{2.105}$$

Finally, we estimate the term $y_1^2 + 2y_1y_2 + 2y_1y_3$ in (2.90). Summing up (2.102) and (2.103) and taking into account (2.100), (2.101), (2.104), and (2.105), we obtain

$$\begin{aligned} y_1^2 + 2y_1y_2 + 2y_1y_3 \geq 4\lambda\Big(1-\eta-\frac{b}{2}\Big)(v_t^2+|\nabla v|^2) \\ + 16\lambda^3\Big[\Big(n+2+\eta-\frac{b}{2}\Big)|x-x_0|^2 - \Big(n+3\eta+\frac{b}{2}\Big)\eta^2t^2 + O\Big(\frac{1}{\lambda}\Big)\Big]v^2 \\ + \operatorname{div} U_5 + (V_5)_t, \\ |U_5| \leq C\lambda^3(|\nabla u|^2 + u_t^2 + u^2)\varphi^2, \\ |V_5| \leq C\lambda^3[|t|(u_t^2 + |\nabla u|^2 + u^2) + (|\nabla u| + |u|)|u_t|]\varphi^2. \end{aligned}$$

Note that $n+2+\eta > n+3\eta$ for any $\eta \in (0,1)$. Hence, for any fixed $\eta \in (0,1)$ we can choose $b > 0$ so small

$$\begin{aligned} y_1^2 + 2y_1y_2 + 2y_1y_3 \geq d_1\lambda(v_t^2 + |\nabla v|^2) + d_2\lambda^3v^2 \\ + \operatorname{div} U_5 + (V_5)_t \quad \text{in } G_h,\ \forall\lambda \geq \lambda_0, \end{aligned}$$

where $d_1, d_2 > 0$ are two constants. □

2.6 Specifying Hölder stability estimates for ill-posed Cauchy problems

Using specific Carleman estimates for parabolic, elliptic and hyperbolic operators, we specify in this section Hölder stability estimates of Section 2.2 for ill-posed Cauchy problems for these operators.

2.6.1 The parabolic operator

Let $x = (x_1, \overline{x}) \in \mathbb{R}^n$, where $\overline{x} = (x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$. Let $G \subset \mathbb{R}^n$ be a bounded domain with a piecewise smooth boundary ∂G. Let $\Gamma \subset \partial G$ and $\Gamma \in C^2$. For $T > 0$, denote

$$Q_T^{\pm} = G \times (-T, T), \quad \Gamma_T^{\pm} = \Gamma \times (-T, T).$$

Let L be the parabolic operator (2.34) in $G_T^{\pm}$.

Ill-posed Cauchy problem 1. Assume that the function $u \in C^{2,1}(\overline{Q_T^{\pm}})$ satisfies the following conditions:

$$Lu = f, \quad (x,t) \in Q_T^{\pm}, \tag{2.106}$$

$$u|_{\Gamma_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{\Gamma_T^{\pm}} = g_1(x,t), \tag{2.107}$$

Estimate the function u via functions f, g_0, g_1 in a subdomain of the domain $G_T^{\pm}$. And also determine the function u in the entire domain $G_T^{\pm}$.

Let $F(x) = 0$ be the equation of a part of the hypersurface Γ, where x belongs to a certain bounded domain and $|\nabla F| \neq 0$ in that domain. Without loss of generality, we assume that $F_{x_1}(x) \neq 0$ in that domain. Then equation $F(x) = 0$ can be rewritten in an equivalent form as $x_1 = \widetilde{F}(\overline{x})$, $\widetilde{F} \in C^2$. Change variables $x \Leftrightarrow x' = (x_1', \overline{x})$, where $x_1' = x_1 - \widetilde{F}(\overline{x})$ and keep previous notations for brevity. Then the operator L is changed accordingly, although properties (2.35)–(2.37) will be kept. Hence, we have obtained that at least a part of Γ is a part of the hyperplane $\{x_1 = 0\}$. For brevity, we assume that the entire hypersurface $\Gamma \subset \{x_1 = 0\}$,

$$\Gamma = \{x_1 = 0, |\overline{x}| < d\}, \tag{2.108}$$

where $d > 0$ is a certain number. Choose a number $h \in (0, \min(1, d))$ and let the number $\alpha \in (0, h)$. Define the domain Ω_h the same way as in (2.29),

$$\begin{aligned}\Omega_h &= \{(x,t) : \psi(x,t) < h, x_1 > 0\} \\ &= \left\{x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h, x_1 > 0\right\}.\end{aligned} \tag{2.109}$$

Note that $\max_{\overline{\Omega}_h} \psi^{-\nu}(x,t) = \alpha^{-\nu}$. Similarly with (2.31)–(2.33)

$$\partial\Omega_h = \partial_1\Omega_h \cup \partial_2\Omega_h, \tag{2.110}$$

$$\partial_1\Omega_h = \left\{x_1 = 0, \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h\right\} \subset \Gamma_T^{\pm}, \tag{2.111}$$

$$\partial_2\Omega_h = \left\{x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha = h\right\}. \tag{2.112}$$

Since $\Gamma \subset \partial G$, then (2.108) and (2.109) imply that

$$\Omega_h \subset Q_T^{\pm}, \tag{2.113}$$

as long as the positive number $h - \alpha$ is sufficiently small. We assume that functions

$$g_0 \in H^1(\partial_1\Omega_h), \quad g_1 \in L_2(\partial_1\Omega_h), \quad f \in L_2(\Omega_h). \tag{2.114}$$

Theorem 2.6.1 follows immediately from Theorem 2.2.1 and Theorem 2.3.1.

Theorem 2.6.1 (Hölder stability estimate for problem (2.106), (2.107)). *Let the domain Ω_h be as in* (2.109). *Suppose that conditions* (2.35)–(2.37) *are satisfied. Let $\varepsilon > 0$ be a sufficiently small number such that $\Omega_{h-3\varepsilon} \neq \varnothing$, i. e. $h - 3\varepsilon > \alpha$. Consider the number $\beta = 2\varepsilon/(3\alpha^{-\nu_0} + 2\varepsilon) \in (0,1)$, where $\nu_0 = \nu_0(\mu_1, K_0, \Omega_h) > 1$ is the number of Theorem* 2.3.1. *Suppose that the function $u \in C^{2,1}(\overline{\Omega}_h)$ satisfies conditions* (2.106), (2.107), *and also that conditions* (2.114) *are in place. Then there exists a sufficiently small number $\delta_0 = \delta_0(\varepsilon, \mu_1, K, K_1, \Omega_h) \in (0,1)$ and a constant $C_1 = C_1(\varepsilon, \mu_1, K, K_1, \Omega_h) > 0$ depending only on listed parameters such that if $\delta \in (0, \delta_0)$ and*

$$\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)} \leq \delta, \tag{2.115}$$

then the following Hölder stability estimate holds:

$$\|u\|_{H^1(\Omega_{h-3\varepsilon})} \le C_1(1 + \|u\|_{H^1(\Omega_h)})\delta^\beta, \quad \forall\delta \in (0, \delta_0). \tag{2.116}$$

While Theorem 2.6.1 is concerned only with a subdomain of the domain $Q_T^\pm$, a natural question to ask here is about the uniqueness of the problem (2.106), (2.107) in the entire domain $Q_T^\pm$. This question is positively addressed in Theorem 2.6.2. Theorem 2.6.2 is also called sometimes “unique continuation” theorem for the parabolic operator. This is because the Cauchy data g_0, g_1 on the boundary are “continued” in the entire domain $Q_T^\pm$.

Theorem 2.6.2 (uniqueness). *Suppose that conditions* (2.35)–(2.37) *are satisfied where* Ω_h *is replaced with* $Q_T^\pm$. *Then there exists at most one function* $u \in C^{2,1}(\overline{Q_T^\pm})$ *satisfying conditions* (2.106), (2.107).

Proof. We need to prove that if functions $g_0 = g_1 = 0, f = 0$, then $u = 0$ in $G_T^\pm$. By (2.115) and (2.116), $u = 0$ in $\Omega_{h-3\varepsilon}$. Since $\varepsilon > 0$ is an arbitrary sufficiently small number, then $u = 0$ in Ω_h. Consider the domain $\widetilde{\Omega}_h = \Omega_h \cap \{t = 0\} \subset \mathbb{R}^n$. It follows from (2.113) that $\widetilde{\Omega}_h \subset G$. Consider a piece P of a hyperplane such that $P \subset \widetilde{\Omega}_h$. Rotating and moving the coordinate system in $\mathbb{R}^n$, we can assume, without any loss of generality, that $P \subset \{x_1 = 0\}$. Hence, we construct an analog Ω_h' of the domain Ω_h in which Γ is replaced with P. Since $u = \partial_n u = 0$ on $\partial_1\Omega_h'$, which is the analog of the hypersurface $\partial_1\Omega_h$ in (2.111). Then by (2.115) and (2.116) $u = 0$ in Ω_h'. One can always choose P in such a way that $\Omega_h' \setminus \Omega_h \neq \varnothing$. It is clear, therefore, that we can cover the entire domain $G_T^\pm$ with domains like Ω_h'. Thus, $u(x,t) = 0$ in $Q_T^\pm$. □

Next, we formulate an analog of Theorem 2.6.1 for the case of the parabolic inequality. It was shown in Theorem 2.2.1 that it does not make much difference in this regard whether one considers equation or inequality.

Ill-posed Cauchy problem 2. Assume that the function $u \in C^{2,1}(\overline{Q_T^\pm})$ satisfies the following conditions:

$$|L_0 u| \le B(|\nabla u| + |u| + |f|), \quad \forall x \in \Omega_h, \tag{2.117}$$

$$u|_{\partial_1\Omega_h} = g_0(x,t), \quad \partial_n u|_{\partial_1\Omega_h} = g_1(x,t), \tag{2.118}$$

where $B > 0$ is a certain constant. Estimate the function u via functions f, g_0, g_1 in the subdomain Ω_h of the domain $Q_T^\pm$.

Theorem 2.6.3 is an analog of Theorem 2.6.1 and it follows immediately from Theorem 2.2.1 and Theorem 2.3.1.

Theorem 2.6.3 (Hölder stability estimate for the parabolic inequality). *Let the domain* Ω_h *be as in* (2.109). *Suppose that conditions* (2.35)–(2.37) *are satisfied. Let* $\varepsilon > 0$ *be a sufficiently small number such that* $\Omega_{h-3\varepsilon} \neq \varnothing$, *that is,* $h - 3\varepsilon > \alpha$. *Consider the* *number* $\beta = 2\varepsilon/(3\alpha^{-\nu_0} + 2\varepsilon) \in (0,1)$, *where* $\nu_0 = \nu_0(\mu_1, K, \Omega_h) > 1$ *is the number of Theorem* 2.3.1. *Suppose that the function* $u \in C^{2,1}(\overline{\Omega}_h)$ *satisfies conditions* (2.117), (2.118), *and also that conditions* (2.114) *are in place. Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, \mu_1, K, B, \Omega_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, \mu_1, K, B, \Omega_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0)$ *and*

$$\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)} \leq \delta,$$

then for any $\delta \in (0, \delta_0)$ *the following Hölder stability estimate holds:*

$$\|u\|_{H^1(\Omega_{h-3\varepsilon})} \leq C_1\big(1 + \|u\|_{H^1(\Omega_h)}\big)\big(\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)}\big)^{\beta}.$$

2.6.2 The elliptic operator

Here, we keep notation of Section 2.4. As in Section 2.6.1, let $G \subset \mathbb{R}^n$ be a bounded domain with a piecewise smooth boundary ∂G. Let $\Gamma \subset \partial G$ and $\Gamma \in C^2$.

Ill-posed Cauchy problem 3. Assume that the function $u \in C^2(\overline{G})$ satisfies the following conditions:

$$Lu = f, \quad x \in G, \tag{2.119}$$

$$u|_\Gamma = g_0(x), \quad \partial_n u|_\Gamma = g_1(x). \tag{2.120}$$

Estimate the function u via functions f, g_0, g_1 in a subdomain of the domain G. And also determine the function u in the entire domain G.

Just as in Section 2.6.1, we transform a part of the hypersurface Γ in a part of the hyperplane $\{x_1 = 0\}$ and assume for brevity that the entire $\Gamma \subset \{x_1 = 0\}$. In other words, we assume the validity of (2.108). Theorem 2.6.4 follows immediately from Theorem 2.2.1 and Theorem 2.4.1.

Theorem 2.6.4 (Hölder stability estimate for problem (2.119), (2.120)). *Let the domain* Ω_h *be as in* (2.65). *Suppose that conditions* (2.70)–(2.73) *are satisfied. Let* $\varepsilon > 0$ *be a sufficiently small number such that* $\Omega_{h-3\varepsilon} \neq \varnothing$, *i. e.* $h - 3\varepsilon > \alpha$. *Let* $\partial_1\Omega_h \subset \Gamma$, *where* $\partial_1\Omega_h$ *is defined in* (2.68). *Consider the number* $\beta = 2\varepsilon/(3\alpha^{-\nu_0} + 2\varepsilon) \in (0,1)$, *where* $\nu_0 = \nu_0(\mu_1, K_0, \Omega_h) > 1$ *is the number of Theorem* 2.4.1. *Suppose that the function* $u \in C^2(\overline{\Omega}_h)$ *satisfies conditions* (2.119), (2.120), *and also that conditions*

$$g_0 \in H^1(\partial_1\Omega_h), \quad g_1 \in L_2(\partial_1\Omega_h), \quad f \in L_2(\Omega_h)$$

are in place, where the hypersurface $\partial_1\Omega_h$ *is defined in* (2.68). *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, \mu_1, K, K_1, \Omega_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, \mu_1, K, K_1, \Omega_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0)$ *and*

$$\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)} \leq \delta, \tag{2.121}$$

then the following Hölder stability estimate holds for all $\delta \in (0, \delta_0)$:

$$\|u\|_{H^1(\Omega_{h-3\varepsilon})} \le C_1\big(1 + \|u\|_{H^1(\Omega_h)}\big)\big(\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)}\big)^{\beta}. \tag{2.122}$$

The proof of the uniqueness Theorem 2.6.5 is similar with the proof of Theorem 2.6.2. Therefore, we omit this proof.

Theorem 2.6.5 (uniqueness). *Suppose that conditions* (2.70)–(2.73) *are satisfied. Then there exists at most one function* $u \in C^{2,1}(\overline{G})$ *satisfying conditions* (2.119), (2.120).

We now consider the case of the elliptic inequality, which is more general than equation (2.119).

Ill-posed Cauchy problem 4. Assume the domain Ω_h is as in (2.67) and that the function $u \in C^{2,1}(\overline{G})$ satisfies the following conditions:

$$|L_0 u| \le B\big(|\nabla u| + |u| + |f|\big), \quad \forall x \in \Omega_h, \tag{2.123}$$

$$u|_{\partial_1\Omega_h} = g_0(x), \quad \partial_n u|_{\partial_1\Omega_h} = g_1(x), \tag{2.124}$$

where $B > 0$ is a certain constant. Estimate the function u via functions f, g_0, g_1 in the subdomain Ω_h of the domain G.

Theorem 2.6.6 is the Hölder stability estimate for this problem.

Theorem 2.6.6 (Hölder stability estimate for problem (2.123), (2.124)). *Let the domain* Ω_h *be as in* (2.67). *Suppose that conditions* (2.70)–(2.73) *are satisfied. Let* $\varepsilon > 0$ *be a sufficiently small number such that* $\Omega_{h-3\varepsilon} \neq \varnothing$, *i. e.* $h - 3\varepsilon > \alpha$. *Consider the number* $\beta = 2\varepsilon/(3\alpha^{-\nu_0} + 2\varepsilon) \in (0,1)$, *where* $\nu_0 = \nu_0(\mu_1, K_0, \Omega_h) > 1$ *is the number of Theorem* 2.4.1. *Suppose that the function* $u \in C^2(\overline{\Omega}_h)$ *satisfies conditions* (2.123), (2.124). *Let* $\partial_1\Omega_h$ *be the hypersurface defined in* (2.68). *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, \mu_1, K, B, \Omega_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, \mu_1, K, B, \Omega_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0)$ *and*

$$\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)} \le \delta,$$

then the following Hölder stability estimate holds for all $\delta \in (0, \delta_0)$:

$$\|u\|_{H^1(\Omega_{h-3\varepsilon})} \le C_1\big(1 + \|u\|_{H^1(\Omega_h)}\big)\big(\|f\|_{L_2(\Omega_h)} + \|g_0\|_{H^1(\partial_1\Omega_h)} + \|g_1\|_{L_2(\partial_1\Omega_h)}\big)^{\beta}.$$

2.6.3 A hyperbolic operator

A Carleman estimate for a general hyperbolic operator of the second order is unknown. All what we can do is to work with operators whose principal is the operator L defined in (2.81). The function $c(x)$ also cannot be an arbitrary positive function. Rather, it should satisfy conditions (2.82)–(2.84).

In this section, the domain $\Omega = \{|x| < R\}$, the point $x_0 \in \Omega$, the number $h \in (0, R^2)$, and the domain G_h is defined in (2.78) is

$$G_h = \{(x,t) : x \in \Omega, |x - x_0|^2 - \eta t^2 > h\}. \tag{2.125}$$

By (2.80), $G_h \neq \varnothing$. Hence, the boundary of this domain consists of two parts:

$$\partial G_h = \partial_1 G_h \cup \partial_2 G_h, \tag{2.126}$$

$$\partial_1 G_h = \overline{G}_h \cap \partial\Omega = \{(x,t) : |x| = R, |x - x_0|^2 - \eta t^2 > h\}, \tag{2.127}$$

$$\partial_2 G_h = \{(x,t) : x \in \Omega, |x - x_0|^2 - \eta t^2 = h\}. \tag{2.128}$$

We consider here the hyperbolic operator of the form

$$Lu = c(x)u_{tt} - \Delta u - \sum_{j=1}^{n} b^j(x,t)u_{x_j} - d(x,t)u. \tag{2.129}$$

We assume that

$$b^j, d \in C(\overline{G_h}), \quad K = \max\Big(\|d\|_{C(\overline{G_h})}, \max_j \|b^j\|_{C(\overline{G_h})}\Big). \tag{2.130}$$

Ill-posed Cauchy problem 5. Assume that the function $u \in C^2(\overline{G}_h)$ satisfies the following conditions:

$$Lu = f, \quad x \in G_h, \tag{2.131}$$

$$u|_{\partial_1 G_h} = g_0(x,t), \quad \partial_n u|_{\partial_1 G_h} = g_1(x,t). \tag{2.132}$$

Estimate the function u via functions f, g_0, g_1 in a subdomain of the domain G_h. And also determine the function u in the domain G_h.

Theorem 2.6.7 follows immediately from Lemma 2.1.1, Theorem 2.2.1, and Theorem 2.5.1.

Theorem 2.6.7. *Let the domain G_h be as in* (2.78) *where $h \in (0, R^2)$. Let L be the hyperbolic operator defined in* (2.129) *and let conditions* (2.130) *be satisfied. Assume that the function $c(x)$ satisfies conditions* (2.82)–(2.84). *Let $\varepsilon > 0$ be a sufficiently small number such that $G_{h+3\varepsilon} \neq \varnothing$, that is, $h + 3\varepsilon < R^2$. Suppose that the function $u \in C^2(\overline{G}_h)$ satisfies conditions* (2.131), (2.132). *Let $\partial_1 G_h$ be the hypersurface defined in* (2.127). *Then there exists a sufficiently small number $\delta_0 = \delta_0(\varepsilon, c, K, G_h) \in (0,1)$ and a constant $C_1 = C_1(\varepsilon, c, K, G_h) > 0$ depending only on listed parameters such that if $\delta \in (0, \delta_0)$ and*

$$\|f\|_{L_2(G_h)} + \|g_0\|_{H^1(\partial_1 G_h)} + \|g_1\|_{L_2(\partial_1 G_h)} \leq \delta,$$

then the following Hölder stability estimate holds for all $\delta \in (0, \delta_0)$:

$$\|u\|_{H^1(G_{h+3\varepsilon})} \leq C_1\big(1 + \|u\|_{H^1(G_h)}\big)\big(\|f\|_{L_2(G_h)} + \|g_0\|_{H^1(\partial_1 G_h)} + \|g_1\|_{L_2(\partial_1 G_h)}\big)^{\beta}. \tag{2.133}$$

Theorem 2.6.8 (uniqueness). *The function* $u \in C^2(\overline{G_h})$ *is determined uniquely in the domain* G_h *from conditions* (2.131), (2.132).

Proof. Setting in (2.131), (2.132) $f = 0$, $g_0 = g_1 = 0$, and applying (2.133), we obtain $u = 0$ in $\overline{G}_h$. □

Remark 2.6.1. Note that uniqueness in Theorem 2.6.7 is claimed only in the domain G_h. This is unlike parabolic and elliptic cases, where uniqueness is actually claimed in an appropriate arbitrary domain. To claim uniqueness in a whole time cylinder $\Omega \times (0, T)$, we prove the Lipschitz stability estimate in the next section.

We now consider Hölder stability estimate for the case of a hyperbolic inequality, as opposed to hyperbolic equation (2.131).

Ill-posed Cauchy problem 6. Let L_0 be the principal part of the hyperbolic operator defined in (2.81), $L_0 = c(x)\partial_t^2 - \Delta$. Assume that the function $u \in C^2(\overline{G}_h)$ satisfies the following conditions:

$$|L_0 u| \leq B\big(|\nabla u| + |u| + |f|\big), \quad \forall x \in G_h, \tag{2.134}$$

$$u|_{\partial_1 G_h} = g_0(x,t), \quad \partial_n u|_{\partial_1 G_h} = g_1(x,t), \tag{2.135}$$

where $B > 0$ is a certain constant. Estimate the function u via functions f, g_0, g_1 in the domain G_h.

Similarly, with the above considered parabolic and elliptic operators, this Hölder stability estimate of Theorem 2.6.9 follows immediately from Theorem 2.2.1 and Theorem 2.5.1.

Theorem 2.6.9. *Let the domain* G_h *be as in* (2.78) *where* $h \in (0, R^2)$. *Assume that the function* $c(x)$ *satisfies conditions* (2.82)–(2.84). *Let* $\varepsilon > 0$ *be a sufficiently small number such that* $G_{h+3\varepsilon} \neq \varnothing$, *that is,* $h + 3\varepsilon < R^2$. *Suppose that the function* $u \in C^2(\overline{G}_h)$ *satisfies conditions* (2.134), (2.135). *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, c, B, G_h) \in (0,1)$ *and a constant* $C_1 = C_1(\varepsilon, c, B, G_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0)$ *and*

$$\|f\|_{L_2(G_h)} + \|g_0\|_{H^1(\partial_1 G_h)} + \|g_1\|_{L_2(\partial_1 G_h)} \leq \delta,$$

then the following Hölder stability estimate holds for all $\delta \in (0, \delta_0)$:

$$\|u\|_{H^1(G_{h+3\varepsilon})} \leq C_1\big(1 + \|u\|_{H^1(G_h)}\big)\big(\|f\|_{L_2(G_h)} + \|g_0\|_{H^1(\partial_1 G_h)} + \|g_1\|_{L_2(\partial_1 G_h)}\big)^{\beta}.$$

2.7 Lipschitz stability estimate for an ill-posed problem for a hyperbolic equation

In this section, we follow the paper [132]. Historical notes about the result of this section can be found in the beginning of Chapter 2.

Just as in Sections 2.5 and 2.6.3, we consider the simplest case when the domain of interest Ω is a ball, $\Omega = \{|x| < R\} \subset \mathbb{R}^n$, where $R = \text{const.} > 0$. For $T = \text{const.} > 0$ denote, replacing in previous notation G with Ω,

$$Q_T^{\pm} = \Omega \times (-T, T), \quad S_T^{\pm} = \partial\Omega \times (-T, T).$$

We consider the same operator L as the one in (2.129),

$$Lu = c(x)u_{tt} - \Delta u - \sum_{j=1}^{n} b^j(x,t)u_{x_j} - d(x,t)u, \tag{2.136}$$

$$Lu = L_0 u - \sum_{j=1}^{n} b^j(x,t)u_{x_j} - d(x,t)u,$$

$$L_0 u = c(x)u_{tt} - \Delta u. \tag{2.137}$$

Just as in (2.130), we assume that

$$b^j, d \in C(\overline{Q_T^{\pm}}), \quad K_1 = \max\Big(\|d\|_{C(\overline{Q_T^{\pm}})}, \max_j \|b^j\|_{C(\overline{Q_T^{\pm}})}\Big). \tag{2.138}$$

We consider Dirichlet and Neumann lateral boundary data at $S_T^{\pm}$. These are lateral Cauchy data. However, we do not assume a knowledge of any function at any hyperplane $\{t = \text{const.}\} \cap Q_T^{\pm}$.

2.7.1 The pointwise case

Ill-posed Cauchy problem 7. Assume that the function $u \in C^2(\overline{Q_T^{\pm}})$ satisfies the following conditions:

$$Lu = f, \quad (x,t) \in Q_T^{\pm}, \quad f \in L_2(Q_T^{\pm}), \tag{2.139}$$

$$u|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{S_T^{\pm}} = g_1(x,t). \tag{2.140}$$

Estimate the function u via functions f, g_0, g_1 in the time cylinder $Q_T^{\pm}$. And also determine the function u in $Q_T^{\pm}$.

Equation (2.139) can be reduced to inequality (2.141) with the constant B depending on the constant K_1 in (2.138). So, the method of this section works for a more general case of a hyperbolic inequality.

Ill-posed Cauchy problem 8. Let the function $u \in C^2(\overline{Q_T^{\pm}})$ satisfy the following conditions:

$$|L_0 u| \leq B(|\nabla u| + |u_t| + |u| + |f|), \quad (x,t) \in Q_T^{\pm}, \tag{2.141}$$

$$u|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{S_T^{\pm}} = g_1(x,t). \tag{2.142}$$

Estimate the function u via functions g_0, g_1, f in the time cylinder $Q_T^{\pm}$.

Theorem 2.7.1 provides the Lipschitz stability estimate for Ill-posed Cauchy problem 8. We impose on the function c the same conditions (2.82), (2.83):

$$c(x) \in [1, \overline{c}], \quad \text{where } \overline{c} = \text{const.} \geq 1, \tag{2.143}$$

$$c \in C^1(\overline{\Omega}). \tag{2.144}$$

As to the condition (2.84), we replace it with

$$\big(x, \nabla c(x)\big) \geq \alpha = \text{const.} > 0, \quad \forall x \in \overline{\Omega}, \tag{2.145}$$

where $(\cdot, \cdot)$ is the scalar product in $\mathbb{R}^n$.

Theorem 2.7.1. *Let the domain* $\Omega = \{|x| < R\} \subset \mathbb{R}^n$. *Assume that conditions* (2.143)–(2.145) *hold. Let the function* $u \in H^2(\overline{Q}_T^{\pm})$ *satisfies inequality* (2.141) *with the lateral Cauchy data* (2.142). *Then there exists a constant* $\eta_0 = \eta_0(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R) \in (0, 1]$ *depending only on listed parameters such that if*

$$T > \frac{R}{\sqrt{\eta_0}}, \tag{2.146}$$

then the following Lipschitz stability estimate holds for the function u:

$$\|u\|_{H^1(Q_T^{\pm})} \leq C_1\big[\|g_0\|_{H^1(S_T^{\pm})} + \|g_1\|_{L_2(S_T^{\pm})} + \|f\|_{L_2(Q_T^{\pm})}\big], \tag{2.147}$$

where the constant $C_1 = C_1(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R, T, B) > 0$ *depends only on listed parameters. In particular, if* $c(x) \equiv 1$, *then one can take* $\eta_0 = 1$ *and in* (2.146) $T > R$.

Proof. We prove this theorem only for functions $u \in C^2(\overline{Q}_T^{\pm})$. The case $u \in H^2(\overline{Q}_T^{\pm})$ can be obtained using density arguments. In this proof, $C_1 = C_1(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R, T) > 0$ denotes different positive constants depending on listed parameters. Since we consider two different points $x_0 \in \{|x_0| < \varepsilon\}$ in this proof, we use here the notation $G_h(x_0)$ for the domain G_h in (2.125), and respectively for parts of its boundary in (2.126)–(2.128). So, the domain $G_h(x_0)$ is the same here as in (2.125) and the structure (2.126)–(2.128) of its boundary remains.

As in Theorem 2.5.1, we choose the number $\eta_0 = \eta_0(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R) \in (0, 1)$. By (2.146), we can choose $\varepsilon \in (0, 3R/4) > 0$ so small that

$$\frac{(R - 4\varepsilon/3)^2}{T^2} < \eta_0.$$

Hence, we choose η such that

$$\eta \in \left(\frac{(R - 4\varepsilon/3)^2}{T^2}, \eta_0\right) \subset (0, 1), \tag{2.148}$$

We choose

$$h \in \left(0, \frac{\varepsilon^2}{9}\right). \tag{2.149}$$

By (2.148) and (2.149),

$$\overline{G}_h(x_0) \subset \{|t| < T\}, \quad \forall x_0 \in \{|x_0| < \varepsilon\}. \tag{2.150}$$

By (2.149), we can choose a sufficiently small number δ such that

$$h + 3\delta \in (0, \varepsilon^2/9). \tag{2.151}$$

Hence, using (2.150), we obtain for all $x_0 \in \{|x_0| < \varepsilon\}$,

$$G_{h+3\delta}(x_0) \neq \varnothing \quad \text{and} \quad G_{h+3\delta}(x_0) \subset G_{h+2\delta}(x_0) \subset G_{h+\delta}(x_0) \subset G_h(x_0) \subset \{|t| < T\}. \tag{2.152}$$

Consider a function $\chi_\delta(x,t)$ such that

$$\chi_\delta(x,t) \in C^2(\overline{Q}_T^\pm), \quad \chi_\delta(x,t) = \begin{cases} 1, & (x,t) \in G_{h+2\delta}(0), \\ 0, & (x,t) \in Q_T^\pm \setminus G_{h+\delta}(0), \\ \text{between 0 and 1} & \text{otherwise.} \end{cases} \tag{2.153}$$

Introduce the function $v(x,t)$ by

$$v(x,t) = u(x,t)\chi_\delta(x,t). \tag{2.154}$$

Multiplying both sides of (2.141) by χ_δ and using (2.142), (2.150), (2.153), and (2.154), we obtain for $(x,t) \in G_h(0)$,

$$\left|c(x)v_{tt} - \Delta v\right| \leq C_1\big(|\nabla v| + |v_t| + |v| + |f|\big) + C_1(1-\chi_\delta)\big(|\nabla u| + |u_t| + |u|\big), \tag{2.155}$$

$$v|_{S_T^\pm} = \chi_\delta g_0, \quad \partial_n v|_{S_T^\pm} = \chi_\delta g_1 + g_0 \partial_n \chi_\delta. \tag{2.156}$$

Squaring both sides of (2.155) and using Theorem 2.5.1 for $x_0 = 0$, we obtain

$$\begin{aligned} &C_1(|\nabla v|^2 + v_t^2 + v^2 + f^2)\varphi^2 + C_1(1-\chi_\delta)\big(|\nabla u|^2 + u_t^2 + u^2\big)\varphi^2 \\ &\quad \geq \lambda\big(|\nabla v|^2 + v_t^2\big)\varphi^2 + \lambda^3 v^2 \varphi^2 + \operatorname{div} U + V_t, \quad \text{in } G_h(0), \forall \lambda \geq \lambda_0, \end{aligned}$$

where the vector function (U,V) satisfies conditions (2.86), (2.87) with the replacement of u by v. Hence, (2.153) and (2.128) imply that $U = V = 0$ on $\partial_2 G_h(0)$. Hence, integrating the latter inequality over G_h and using Gauss' formula and (2.156), we obtain

$$\begin{aligned} &\int\limits_{G_h(0)} \lambda\big(|\nabla v|^2 + v_t^2\big)\varphi^2 dxdt + \lambda^3 \int\limits_{G_h(0)} v^2\varphi^2 dxdt \\ &\leq C_1 \int\limits_{G_h(0)} \big(|\nabla v|^2 + v_t^2 + v^2 + g^2\big)\varphi^2 dxdt \\ &\quad + C_1 \exp[2\lambda(h+2\delta)]\|u\|^2_{H^1(Q_T^\pm)} + C_1 e^{2\lambda R^2}\big(\|g_0\|^2_{H^1(S_T^\pm)} + \|g_1\|^2_{L_2(S_T^\pm)} + \|f\|^2_{L_2(Q_T^\pm)}\big). \end{aligned}$$

Let $\lambda_0 > 1$ be the number of Theorem 2.5.1. There exists a number $\lambda_1 = \lambda_1(\lambda_0, C_1) \geq \lambda_0$ such that $\lambda_1 > C_1/2$. Hence,

$$\lambda \int\limits_{G_h(0)} (|\nabla v|^2 + v_t^2 + v^2)\varphi^2 dxdt \tag{2.157}$$
$$\leq C_1 \exp[2\lambda(h + 2\delta)] \|u\|^2_{H^1(Q_T^{\pm})}$$
$$+ C_1 e^{2\lambda R^2} [\|g_0\|^2_{H^1(S_T^{\pm})} + \|g_1\|^2_{L_2(S_T^{\pm})} + \|f\|^2_{L_2(Q_T^{\pm})}], \quad \forall \lambda > \lambda_1.$$

By (2.152) and (2.153),

$$\lambda \int\limits_{G_h(0)} (|\nabla v|^2 + v_t^2 + v^2)\varphi^2 dxdt$$
$$\geq \lambda \int\limits_{G_{h+3\delta}(0)} (|\nabla v|^2 + v_t^2 + v^2)\varphi^2 dxdt$$
$$= \lambda \int\limits_{G_{h+3\delta}(0)} (|\nabla u|^2 + u_t^2 + u^2)\varphi^2 dxdt \geq \exp[2\lambda(h + 3\delta)] \|u\|^2_{H^1(G_{h+3\delta}(0))}.$$

Hence, using (2.157), we obtain

$$\|u\|^2_{H^1(G_{h+3\delta}(0))} \leq C_1 \exp(-2\lambda\delta) \|u\|^2_{H^1(Q_T^{\pm})} \tag{2.158}$$
$$+ C_1 e^{2\lambda R^2} [\|g_0\|^2_{H^1(S_T^{\pm})} + \|g_1\|^2_{L_2(S_T^{\pm})} + \|f\|^2_{L_2(Q_T^{\pm})}].$$

Note that by (2.125) and (2.151)

$$\left\{x : |x| \in \left(\frac{\varepsilon}{3}, R\right)\right\} \subset G_{h+3\delta}(0) \cap \{t = 0\} = \{x : |x| \in (\sqrt{h + 3\delta}, R)\}. \tag{2.159}$$

Choose a point x_0 such that $|x_0| = 3\sqrt{h + 3\delta}$. Then (2.151) implies that $|x_0| < \varepsilon$. Hence, using (2.145), we can assume that

$$(x - x_0, \nabla c(x)) \geq \frac{\alpha}{2}, \quad \forall x \in \overline{\Omega}. \tag{2.160}$$

Consider now an arbitrary point $y \in \{|x| \leq \sqrt{h + 3\delta}\}$. Then

$$|y - x_0| \geq |x_0| - |y| = 3\sqrt{h + 3\delta} - |y|$$
$$\geq 3\sqrt{h + 3\delta} - \sqrt{h + 3\delta} = 2\sqrt{h + 3\delta} > \sqrt{h + 3\delta}.$$

Hence,

$$\{|x| \leq \sqrt{h + 3\delta}\} \subset \{|y - x_0| > \sqrt{h + 3\delta}\}. \tag{2.161}$$

It follows from (2.159) and (2.161) that there exists a sufficiently small number $\sigma = \sigma(\varepsilon)$ such that

$$\{t \in (0,\sigma)\} \subset [G_{h+3\delta}(0) \cup G_{h+3\delta}(x_0)]. \tag{2.162}$$

Since $h+3\delta \in (0,\varepsilon^2/9)$ and $|x_0| = 3\sqrt{h+3\delta}$, then (2.150) implies that $\overline{G}_h(x_0) \subset \{|t| < T\}$. Next, we use (2.160) and (2.162) to obtain, similarly with (2.158),

$$\begin{aligned}\|u\|^2_{H^1(G_{h+3\delta}(x_0))} \le C_1 \exp(-2\lambda\delta)\|u\|^2_{H^1(Q_T^\pm)}\\ + C_1 e^{2\lambda R^2}\big[\|g_0\|^2_{H^1(S_T^\pm)} + \|g_1\|^2_{L_2(S_T^\pm)} + \|f\|^2_{L_2(Q_T^\pm)}\big].\end{aligned}$$

Combining this with (2.158), we obtain

$$\begin{aligned}\|u\|^2_{H^1(G_{\gamma+3\delta}(0)\cup G_{\gamma+3\delta}(x_0))} \le C_1 \exp(-2\lambda\delta)\|u\|^2_{H^1(Q_T^\pm)}\\ + C_1 e^{2\lambda R^2}\big[\|g_0\|^2_{H^1(S_T^\pm)} + \|g_1\|^2_{L_2(S_T^\pm)} + \|f\|^2_{L_2(Q_T^\pm)}\big].\end{aligned}$$

This, (2.162) and the mean value theorem imply that there exists a number $t_0 \in [0,\sigma]$ such that

$$\begin{aligned}&\|u(x,t_0)\|^2_{H^1(\Omega)} + \|u_t(x,t_0)\|^2_{L_2(\Omega)}\\ &\quad\le C_1 \exp(-2\lambda\delta)\|u\|^2_{H^1(Q_T^\pm)}\\ &\qquad + C_1 e^{2\lambda R^2}\big[\|g_0\|^2_{H^1(S_T^\pm)} + \|g_1\|^2_{L_2(S_T^\pm)} + \|f\|^2_{L_2(Q_T^\pm)}\big].\end{aligned} \tag{2.163}$$

Let $c(x)u_{tt} - \Delta u := Z(x,t)$. Then inequality (2.141) implies

$$|Z| \le B\big(|\nabla u| + |u_t| + |u| + |f|\big), \quad \forall (x,t) \in Q_T^\pm. \tag{2.164}$$

Consider the initial boundary value problem with the reversed time

$$\begin{aligned}c(x)u_{tt} - \Delta u = Z(x,t), \quad &\forall (x,t) \in \{x \in \Omega, t \in (-T,t_0)\},\\ u(x,t_0) = u_0(x), \quad &u_t(x,t_0) = u_1(x),\\ u|_{(x,t)\in\partial\Omega\times(-T,t_0)} = g_0(x,t).&\end{aligned}$$

Next, consider the same initial boundary value problem but in the time cylinder $(x,t) \in \{x \in \Omega, t \in (t_0,T)\}$. Recall that the hyperbolic equation can be solved in both positive and negative directions of time. Hence, the standard method of energy estimates being applied to two latter problems combined with inequalities (2.163) and (2.164) leads to

$$\begin{aligned}\|u\|^2_{H^1(Q_T^\pm)} \le C_1 \exp(-2\lambda\delta)\|u\|^2_{H^1(Q_T^\pm)}\\ + C_1 e^{2\lambda R^2}\big[\|g_0\|^2_{H^1(S_T^\pm)} + \|g_1\|^2_{L_2(S_T^\pm)} + \|f\|^2_{L_2(Q_T^\pm)}\big].\end{aligned} \tag{2.165}$$

Choosing $\lambda = \lambda(C_1,\delta)$ so large that $C_1 \exp(-2\lambda\delta) \le 1/2$, we obtain from (2.165) the target estimate (2.147). □

We now formulate the uniqueness result for Ill-posed Cauchy problem 7.

Theorem 2.7.2. *Let* $\eta_0 = \eta_0(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R, \varepsilon) \in (0,1]$ *be the number of Theorem* 2.7.1 *and let inequality* (2.146) *hold. Let* L *be the hyperbolic operator satisfying conditions* (2.136)–(2.138), *where the function* $c(x)$ *satisfies conditions* (2.143)–(2.145). *Then there exists at most one solution* $u \in H^2(Q_T^{\pm})$ *of the problem* (2.129)–(2.140).

Proof. Rewrite equation (2.139) in a more general form (2.141) where $B = K$. Set in (2.139) and (2.140) $f = 0$, $g_0 = g_1 = 0$. Next, apply (2.147). We obtain $u = 0$ in $Q_T^{\pm}$. □

2.7.2 The integral inequality

We now replace the pointwise inequality (2.141) with an integral inequality, which is similar with (2.27),

$$\int_{Q_T^{\pm}} (Lu)^2 dxdt \leq S^2, \tag{2.166}$$

where L is the hyperbolic operator in (2.136) and $S > 0$ is a number. We again assume that the function $u \in H^2(Q_T^{\pm})$ satisfies boundary conditions (2.142)

$$u|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{S_T^{\pm}} = g_1(x,t). \tag{2.167}$$

Theorem 2.7.3 generalized Theorem 2.7.1 to the case of the following problem.

Theorem 2.7.3. *Let the domain* $\Omega = \{|x| < R\} \subset \mathbb{R}^n$. *Assume that conditions* (2.143)–(2.145) *hold. Let the function* $u \in H^2(\overline{Q}_T^{\pm})$ *satisfies integral inequality* (2.166) *with the lateral Cauchy data* (2.167). *Then there exists a constant* $\eta_0 = \eta_0(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R) \in (0,1]$ *depending only on listed parameters such that if*

$$T > \frac{R}{\sqrt{\eta_0}}, \tag{2.168}$$

then the following Lipschitz stability estimate holds for the function u:

$$\|u\|_{H^1(Q_T^{\pm})} \leq C_1 [\|g_0\|_{H^1(S_T^{\pm})} + \|g_1\|_{L_2(S_T^{\pm})} + S],$$

where the constant $C_1 = C_1(\overline{c}, \|\nabla c\|_{C(\overline{\Omega})}, R, T, B) > 0$ *depends only on listed parameters. In particular, if* $c(x) \equiv 1$, *then one can take* $\eta_0 = 1$ *and in* (2.168) $T > R$.

We omit the proof of Theorem 2.7.3 since it is quite similar with the proof of Theorem 2.7.1.

3 Global uniqueness for coefficient inverse problems and Lipschitz stability for a hyperbolic CIP

With the exception of Sections 3.3 and 3.6, the material of this chapter is basically republished from a part of Section 1.10 of [22]. Permission for republishing is obtained from the publisher of [22].

In this chapter, the Bukhgeim–Klibanov method (BK) [51] is presented. Historical remarks can be found in Chapter 1. In principle, this method can be formulated for a general PDE operator of the second order for which the Carleman estimate is valid, and this was done in [122, 126]. However, it is better to demonstrate how this method actually works on specific examples. So, we formulate the BK method here for Coefficient Inverse Problems (CIPs) for hyperbolic, parabolic, and elliptic PDEs. Hölder stability estimates for these CIPs can be obtained by simple combinations of the method of this chapter with the one of Chapter 2. Hence, we are not proving these estimates here. Instead, however, we prove a stronger Lipschitz stability estimate for a CIP for a hyperbolic equation. These Lipschitz stability estimates were first established by Imanuvilov and Yamamoto [95–99, 103] via a combination of the BK method with the idea of the proof of Theorem 2.7.1. The methods we use in Sections 3.3 and 3.6 are different from the one of Imanuvilov and Yamamoto.

Unless stated otherwise, everywhere in this chapter, $\Omega \subset \mathbb{R}^n$ is a domain with a piecewise smooth boundary $\partial\Omega$. In most cases, the domain Ω is bounded, but it is unbounded sometimes. For any $T > 0$, denote

$$Q_T = \Omega \times (0,T), \quad Q_T^{\pm} = \Omega \times (-T,T), \quad S_T = \partial\Omega \times (0,T), \quad S_T^{\pm} = \partial\Omega \times (T,T).$$

We remind that for any multiindex $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ with nonnegative integer coordinates,

$$D^{\alpha} = \frac{\partial^{|\alpha|}}{\partial^{\alpha_n} \dots \partial^{\alpha_1}}, \quad |\alpha| = \alpha_1 + \dots + \alpha_n.$$

Everywhere below, we do not discuss conditions imposed on the coefficients and initial/boundary data of PDEs, which would guarantee the smoothness we need of the solutions of forward problems. These conditions are well known from the classical results of the theory of PDEs; see, for example, classical books [79, 173, 174]. Usually we require the solution of the forward problem $u \in C^{3+k}$, where the integer $k \geq 0$ depends on the unknown coefficient of our interest.

3.1 Estimating an integral

Lemma 3.1.1 is a very important element of the BK method; also, see a little bit more general result in Lemma 1.10.3 of [22] and Lemma 3.1.1 of [165].

https://doi.org/10.1515/9783110745481-003

Lemma 3.1.1. *Let the number $a > 0$. Let $\lambda > 0$ be a parameter. Then for all functions $p \in L_2(-a, a)$,*

$$\int_{-a}^{a} \left(\int_{0}^{t} p(\tau)d\tau \right)^2 \exp[-2\lambda t^2]dt \le \frac{1}{4\lambda} \int_{-a}^{a} p^2(t) \exp[-2\lambda t^2]dt. \tag{3.1}$$

Proof. We have

$$\begin{aligned}\int_{0}^{a} \left(\int_{0}^{t} p(\tau)d\tau \right)^2 \exp(-2\lambda t^2)dt &\le \int_{0}^{a} \exp(-2\lambda t^2)t\left(\int_{0}^{t} p^2(\tau)d\tau \right)dt \\ &= \frac{1}{4\lambda} \int_{0}^{a} \frac{d}{dt}[-\exp(-2\lambda t^2)]\left(\int_{0}^{t} p^2(\tau)d\tau \right)dt \\ &= -\frac{1}{4\lambda} \exp(-2\lambda a^2) \int_{0}^{a} p^2(\tau)d\tau + \frac{1}{4\lambda} \int_{0}^{a} p^2(\tau) \exp(-2\lambda t^2)dt \\ &\le \frac{1}{4\lambda} \int_{0}^{a} p^2(\tau) \exp(-2\lambda t^2)dt.\end{aligned}$$

Hence, we have proved that

$$\int_{0}^{a} \exp(-2\lambda t^2)\left(\int_{0}^{t} p(\tau)d\tau \right)^2 dt \le \frac{1}{4\lambda} \int_{0}^{a} p^2(t) \exp(-2\lambda t^2)dt. \tag{3.2}$$

Similarly,

$$\int_{-a}^{0} \exp(-2\lambda t^2)\left(\int_{0}^{t} p(\tau)d\tau \right)^2 dt \le \frac{1}{4\lambda} \int_{-a}^{0} p^2(t) \exp(-2\lambda t^2)dt. \tag{3.3}$$

Summing up (3.2) and (3.3), we obtain (3.1). □

3.2 Hyperbolic equation

In this section, $\Omega = \{x \in \mathbb{R}^n : |x| < R\}$. We assume that the coefficient $c(x)$ in the principal part of our hyperbolic operator satisfies conditions (2.132), (2.133) of Chapter 2, that is,

$$c(x) \in [1, \overline{c}], \quad \text{where } \overline{c} = \text{const.} \ge 1, \tag{3.4}$$

$$c \in C^1(\overline{\Omega}). \tag{3.5}$$

In addition, we assume that $c(x)$ satisfies the following analog of condition (2.134):

$$(\nabla c, x) \geq 0, \quad \forall x \in \overline{\Omega}. \tag{3.6}$$

First, we consider the forward problem. In this problem, the integer $k \geq 0$ will be specified below in Sections 3.2, 3.3, depending on the unknown coefficient of our concern.

Forward problem 3.2. Find the solution $u \in C^{2+k}(\overline{Q_T})$ of the following initial boundary value problem:

$$c(x)u_{tt} = \Delta u + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha u, \quad \text{in } Q_T, \tag{3.7}$$

$$u(x,0) = g_0(x), \quad u_t(x,0) = g_1(x), \tag{3.8}$$

$$u|_{S_T} = p_0(x,t). \tag{3.9}$$

Coefficient Inverse Problem 3.2 (CIP 3.2). Assume that the normal derivative of the function u is known on the boundary $\partial\Omega$ of the domain Ω,

$$\left.\frac{\partial u}{\partial n}\right|_{S_T} = p_1(x,t). \tag{3.10}$$

Determine one of coefficients of equation (3.7).

Since only a single pair (f_0, f_1) of initial conditions is known here, then the CIP 3.2 is the problem with the single measurement data, because only a single pair

Theorem 3.2.1. *Let the coefficient $c(x)$ in* (3.7) *satisfies conditions* (3.4)–(3.6). *In addition, let coefficients $a_\alpha \in C(\overline{\Omega})$. Assume that the function $c(x)$ is unknown while coefficients $a_\alpha(x)$ are known. Also, suppose that*

$$\Delta g_0(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g_0(x) \neq 0, \quad \forall x \in \overline{\Omega}. \tag{3.11}$$

Then there exists a sufficiently large number $T > 0$ such that conditions (3.7)–(3.10) *are satisfied for no more than one vector function (u, c) with $u \in C^5(\overline{Q_T})$.*

Theorem 3.2.2. *Let the coefficient $c(x)$ in* (3.7) *satisfies conditions* (3.4)–(3.6). *In addition, let all coefficients $a_\alpha \in C(\overline{\Omega})$. Fix an index α_0 with $|\alpha_0| \leq 1$. Assume that the coefficient $a_{\alpha_0}(x)$ is unknown while all other coefficients of equation* (3.7) *are known. Also, let*

$$D_x^{\alpha_0} g_0(x) \neq 0 \quad \textit{for } x \in \overline{\Omega}.$$

Then there exists a sufficiently large number $T > 0$ such that conditions (3.7)–(3.10) *are satisfied for no more than one vector function (u, a_{α_0}) with $u \in C^{3+|\alpha_0|}(\overline{Q_T})$.*

Remark 3.2.1. Suppose that the function $g_0(x) \equiv 0$. Then one should consider the function $\widehat{u}(x,t) = u_t(x,t)$ implying $\widehat{u}(x,0) = g_1(x)$ and $\widehat{u}_t(x,0) = 0$. Hence, Theorems 3.2.1, 3.2.2 can be reformulated for this case with the replacement of $g_0(x)$ with $g_1(x)$ and the increase of the required smoothness of the function $u(x,t)$ by one.

We prove only Theorem 3.2.1 since the proof of Theorem 3.2.2 is completely similar.

Proof of Theorem 3.2.1. Unlike the proof of Theorem 2.2.1, we are not introducing a cut-off function here. Consider two possible pairs of functions (u_1, c_1) and (u_2, c_2). Let $v = u_1 - u_2$, $\widetilde{c} = c_1 - c_2$. Obviously,

$$c_1 u_{1tt} - c_2 u_{2tt} = c_1 u_{1tt} - c_1 u_{2tt} + (c_1 - c_2) u_{2tt} = c_1 v_{tt} + b u_{2tt}.$$

Using (3.7)–(3.11), we obtain

$$Mv = c_1(x) v_{tt} - \Delta v - \sum_{j=1}^{n} a_\alpha(x) D_x^\alpha v = -b(x) f(x,t), \quad \text{in } Q_T, \tag{3.12}$$

$$v(x,0) = 0, \quad v_t(x,0) = 0, \tag{3.13}$$

$$v|_{S_T} = \left.\frac{\partial v}{\partial n}\right|_{S_T} = 0, \tag{3.14}$$

$$f(x,t) := u_{2tt}(x,t). \tag{3.15}$$

By (3.7) and (3.15),

$$f(x,t) = \frac{1}{c_2(x)} \left(\Delta u_2 + \sum_{|\alpha| \le 1} a_\alpha(x) D_x^\alpha u_2 \right). \tag{3.16}$$

Hence, it follows from (3.11) and (3.16) that $f(x,0) \neq 0$ in $\overline{\Omega}$. More precisely,

$$f(x,0) = \frac{1}{c_2(x)} \left(\Delta g_0(x) + \sum_{|\alpha| \le 1} a_\alpha(x) D_x^\alpha g_0(x) \right) \neq 0, \quad \forall x \in \overline{\Omega}. \tag{3.17}$$

Hence, we have for a sufficiently small $\delta > 0$:

$$f(x,t) \neq 0, \quad \forall (x,t) \in \overline{Q_\delta} = \overline{\Omega} \times [0,\delta]. \tag{3.18}$$

The form of equation (3.12) is inconvenient since it is one equation with two unknown functions $v(x,t)$ and $b(x)$. Hence, the first step of the BK method, which is used not only in the uniqueness issue but in numerical methods as well, is to eliminate the unknown function $b(x)$ from (3.12) using the fact that $b(x)$ is independent on the variable t. By (3.12),

$$b(x) = -\frac{Mv}{f}(x,t), \quad \forall (x,t) \in \overline{Q_\delta}.$$

Hence,

$$0 = \partial_t\big(-b(x)\big) = \partial_t\left(\frac{Mv}{f}(x,t)\right) \quad \text{for } (x,t) \in \overline{Q_\delta}. \tag{3.19}$$

Hence,

$$Mv_t = \left(\frac{f_t}{f}\right)Mv \quad \text{for } (x,t) \in \overline{Q_\delta}. \tag{3.20}$$

Introduce the function $q(x,t)$,

$$q(x,t) = \frac{f_t}{f}(x,t) \in C^2(\overline{Q_\delta}). \tag{3.21}$$

Consider the function $w(x,t)$,

$$w(x,t) = v_t(x,t) - q(x,t)v(x,t). \tag{3.22}$$

Since by (3.13) $v(x,0) = 0$, then (3.21) and (3.22) imply that

$$v(x,t) = \int_0^t K(x,t,\tau)w(x,\tau)d\tau, \tag{3.23}$$

$$K(x,t,\tau) = \frac{f(x,t)}{f(x,\tau)} \in C^2\big(\overline{\Omega} \times [0,\delta] \times [0,\delta]\big). \tag{3.24}$$

In addition it follows from the second condition (3.13) and (3.22) that

$$w(x,0) = 0. \tag{3.25}$$

We now rewrite the expression in the right-hand side of (3.20). We have

$$qv_{tt} = (qv)_{tt} - 2q_t v_t - q_{tt}v, \tag{3.26}$$

$$q\Delta v = \Delta(qv) - 2\nabla q\nabla v - v\Delta q, \tag{3.27}$$

$$qD^\alpha v = D^\alpha(qv) - vD^\alpha q, \quad |\alpha| = 1. \tag{3.28}$$

Hence,

$$\left(\frac{f_t}{f}\right)Mv = qMv = M(qv) + M_1\big(D^\alpha v\big), \quad |\alpha| \le 1,$$

where $M_1(D^\alpha v)$ is a linear operator with respect to x, t-derivatives of the first and zero order of the function v. Hence, using (3.20) and (3.22), we obtain

$$\begin{aligned} 0 &= Mv_t - qMv = Mv_t - M(qv) - M_1(v) \\ &= M(v_t - qv) - M_1(v) = Mw - M_1(v). \end{aligned}$$

Hence,

$$Mw = M_1(D^\alpha v), \quad |\alpha| \le 1, \quad (x,t) \in \overline{Q_\delta}. \tag{3.29}$$

Next, by (3.23) and (3.24)

$$\partial_{x_i} v(x,t) = \int_0^t \partial_{x_i} K(x,t,\tau) w(x,\tau) d\tau + \int_0^t K(x,t,\tau) \partial_{x_i} w(x,\tau) d\tau, \quad i = 1,\dots,n, \tag{3.30}$$

$$\partial_t v(x,t) = w(x,t) + \int_0^t \partial_t K(x,t,\tau) w(x,\tau) d\tau. \tag{3.31}$$

We saw in Section 2.2 of Chapter 2 that Carleman estimates can handle not only equations but inequalities as well. This is an important property of Carleman estimates. This property, combined with (3.30) and (3.31) enables us to rewrite equation (3.29) in a more general form as an integral differential inequality,

$$\left|c_1(x) w_{tt} - \Delta w\right| \le A\left[|\nabla w| + |w_t| + |w| + \int_0^t (|\nabla w| + |w|)(x,\tau) d\tau \right], \quad (x,t) \in \overline{Q_\delta}. \tag{3.32}$$

Here and below in this chapter $A = \text{const.} > 0$ is independent on w, x, t and is used for notation of different constants. Recall that $S_\delta = \partial\Omega \times (0,\delta)$. Hence, using (3.14), (3.22), and (3.25), we obtain in addition to (3.11):

$$w|_{S_\delta} = \partial_n w|_{S_\delta} = 0, \tag{3.33}$$

$$w(x,0) = 0. \tag{3.34}$$

We now apply the Carleman estimate of Theorem 2.5.1. Consider the functions $\psi(x,t)$, $\varphi(x,t)$,

$$\psi(x,t) = |x|^2 - \xi t^2, \tag{3.35}$$

$$\varphi(x,t) = \exp(\lambda \psi(x,t)), \tag{3.36}$$

where $\lambda > 0$ is a parameter. Let $\xi_0 = \xi_0(\overline{c}, R, \|\nabla c\|_{C(\overline{\Omega})}) \in (0,1)$ be the number considered in that theorem and this number depends only on listed parameters. Let the number $\xi \in (0,\xi_0)$. Define the domain $H_{\xi\delta^2}$ as

$$H_{\xi\delta^2} = \{(x,t) : |x|^2 - \xi t^2 > R^2 - \xi\delta^2, t > 0, |x| < R\}. \tag{3.37}$$

Hence, $H_{\xi\delta^2} \subset \overline{Q_\delta}$. Hence, for $\xi \in (0,\xi_0)$, we can apply Theorem 2.5.1. Let $\lambda_0 = \lambda_0(\Omega, \xi, \overline{c}, \|\nabla c\|_{C(\overline{\Omega})}) > 1$ be the number chosen in that theorem and let $\lambda \ge \lambda_0$. Then

$$(c_1(x) w_{tt} - \Delta w)^2 \varphi^2 \ge C\lambda(|\nabla w|^2 + w_t^2 + \lambda^2 w^2)\varphi^2 + \operatorname{div} W_1 + \partial_t W_2, \tag{3.38}$$

for $(x,t) \in H_{\xi_0\delta^2}$, where the vector function (W_1, W_2) can be estimated as

$$|W_1| \le C\lambda^3(|\nabla w|^2 + w_t^2 + w^2)\varphi^2, \tag{3.39}$$

$$|W_2| \le C\lambda^3[t(w_t^2 + |\nabla w|^2 + w^2) + (|\nabla w| + |w|)|w_t|]\varphi^2. \tag{3.40}$$

Here, the constant $C = C(\Omega, \xi, \overline{c}, \|\nabla c\|_{C(\overline{\Omega})}) > 0$ depends only on listed parameters. Using (3.34) and (3.40), we obtain

$$W_2(x,0) = 0. \tag{3.41}$$

The boundary $\partial H_{\xi\delta^2}$ consists of three parts:

$$\partial H_{\xi\delta^2} = \partial_1 H_{\xi_0\delta^2} \cup \partial_2 H_{\xi_0\delta^2} \cup \partial_3 H_{\xi_0\delta^2}, \tag{3.42}$$

$$\partial_1 H_{\xi\delta^2} = \{(x,t)\} \in \{|x| = R\} \cap \overline{H_{\xi_0\delta^2}}, \tag{3.43}$$

$$\partial_2 H_{\xi\delta^2} = \{|x|^2 - \xi t^2 = R^2 - \xi\delta^2, t > 0, |x| < R\}, \tag{3.44}$$

$$\partial_3 H_{\xi\delta^2} = \{\sqrt{R^2 - \xi\delta^2} < |x| < R, t = 0\}. \tag{3.45}$$

By (3.33) and (3.34),

$$w|_{\partial_1 H_{\xi\delta^2}} = \partial_n w|_{\partial_1 H_{\xi\delta^2}} = w|_{\partial_3 H_{\xi\delta^2}} = 0. \tag{3.46}$$

Integrate both parts of (3.38) over the domain $H_{\xi\delta^2}$ and use Gauss' formula as well as (3.38)–(3.46). Observe that by (3.41) and (3.46) the resulting integral over $\partial_1 H_{\xi\delta^2} \cup \partial_3 H_{\xi\delta^2}$ equals zero. Hence,

$$\begin{aligned} &\int\limits_{H_{\xi\delta^2}} (c_1(x)w_{tt} - \Delta w)^2 \varphi^2 dxdt \\ &\ge C\lambda \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2)\varphi^2 dxdt + C\lambda^3 \int\limits_{H_{\xi\delta^2}} w^2\varphi^2 dxdt \\ &\quad - C\lambda^3 \exp[2\lambda(R^2 - \xi\delta^2)] \int\limits_{\partial_2 H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dS. \end{aligned} \tag{3.47}$$

Square both sides of (3.32). Then multiply by φ^2 and integrate over $H_{\xi\delta^2}$. We obtain

$$\begin{aligned} \int\limits_{H_{\xi\delta^2}} (c_1(x)w_{tt} - \Delta w)^2 \varphi^2 dxdt \le A \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt \\ + A \int\limits_{H_{\xi\delta^2}} \left(\int\limits_0^t (|\nabla w| + |w|)(x,\tau) d\tau \right)^2 \varphi^2 dxdt. \end{aligned} \tag{3.48}$$

Now we need to apply Lemma 3.1.1. Note that

$$H_{\xi\delta^2} \cap \{t = 0\} = H^0_{\xi\delta^2} = \left\{x : \sqrt{R^2 - \xi\delta^2} < |x| < R\right\}. \tag{3.49}$$

Denote

$$t(x) = \frac{\sqrt{|x|^2 - (R^2 - \xi\delta^2)}}{\sqrt{\xi}}. \tag{3.50}$$

Then applying Lemma 3.1.1, we obtain

$$\begin{aligned}
&\int\limits_{H_{\xi\delta^2}} \left(\int\limits_0^t (|\nabla w| + |w|)(x,\tau)d\tau \right)^2 \varphi^2 dxdt \\
&\quad = \int\limits_{H^0_{\xi\delta^2}} \left[\int\limits_{-t(x)}^{t(x)} \left(\int\limits_0^t (|\nabla w| + |w|)(x,\tau)d\tau \right)^2 \varphi^2 dt \right] dx \\
&\quad \le \frac{C}{\lambda} \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w^2)\varphi^2 dxdt.
\end{aligned} \tag{3.51}$$

Combining (3.48) and (3.51), we obtain

$$\int\limits_{H_{\xi\delta^2}} (c_1(x)w_{tt} - \Delta w)^2 \varphi^2 dxdt \le A \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt.$$

Hence, (3.47) implies that

$$\begin{aligned}
&A \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt \\
&\quad \ge C\lambda \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2)\varphi^2 dxdt \\
&\qquad + C\lambda^3 \int\limits_{H_{\xi\delta^2}} w^2\varphi^2 dxdt \\
&\qquad - C\lambda^3 \exp[2\lambda(R^2 - \xi\delta^2)] \int\limits_{\partial_2 H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)dS, \quad \forall \lambda \ge \lambda_0.
\end{aligned} \tag{3.52}$$

Choose $\lambda_1 \ge \lambda_0$ so large that $C\lambda_1 > 2A$. Then (3.52) implies that with a different constant A

$$\begin{aligned}
&A\lambda^2 \exp[2\lambda(R^2 - \xi\delta^2)] \int\limits_{\partial_2 H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)dS \\
&\quad \ge \int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt.
\end{aligned} \tag{3.53}$$

Let $\varepsilon \in (0, \delta^2)$ be an arbitrary number. Consider a new domain $H^{\varepsilon}_{\xi\delta^2}$,

$$H^{\varepsilon}_{\xi\delta^2} = \{(x,t) : |x|^2 - \xi t^2 > R^2 - \xi\delta^2 + \xi\varepsilon, t > 0, |x| < R\}.$$

Then

$$H^{\varepsilon}_{\xi\delta^2} \subset H_{\xi\delta^2}, \quad \varphi^2 > \exp[2\lambda(R^2 - \xi\delta^2 + \xi\varepsilon)] \quad \text{in } H^{\varepsilon}_{\xi\delta^2}.$$

Hence,

$$\begin{aligned}\int\limits_{H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt &\geq \int\limits_{H^{\varepsilon}_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2)\varphi^2 dxdt \\ &\geq \exp[2\lambda(R^2 - \xi\delta^2 + \xi\varepsilon)] \int\limits_{H^{\varepsilon}_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dxdt.\end{aligned}$$

Hence, using (3.53), we obtain

$$\begin{aligned}&A\lambda^2 \exp[2\lambda(R^2 - \xi\delta^2)] \int\limits_{\partial_2 H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dS \\ &\quad \geq \exp[2\lambda(R^2 - \xi\delta^2 + \xi\varepsilon)] \int\limits_{H^{\varepsilon}_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dxdt.\end{aligned}$$

Dividing this estimate by $\exp[2\lambda(R^2 - \xi\delta^2 + \xi\varepsilon)]$, we obtain

$$\int\limits_{H^{\varepsilon}_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dxdt \leq A\lambda^2 \exp(-2\lambda\xi\varepsilon) \int\limits_{\partial_2 H_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dS. \tag{3.54}$$

Setting in (3.54) $\lambda \to \infty$, we obtain

$$\int\limits_{H^{\varepsilon}_{\xi\delta^2}} (|\nabla w|^2 + w_t^2 + w^2) dxdt = 0.$$

Hence, $w(x,t) = 0$ in $H^{\varepsilon}_{\xi\delta^2}$. Since $\varepsilon \in (0, \delta^2)$ and $\xi \in (0, \xi_0)$ are arbitrary numbers, then $w(x,t) = 0$ in $H_{\xi_0\delta^2}$. Hence, (3.23) implies that

$$v(x,t) = 0 \quad \text{in } H_{\xi_0\delta^2}. \tag{3.55}$$

Substituting (3.55) in (3.12) and using (3.18) and (3.37), we obtain

$$b(x) = 0 \quad \text{for } x \in \{\sqrt{R^2 - \xi_0\delta^2} < |x| < R\}. \tag{3.56}$$

In fact, (3.56) is the *MAIN* step of the proof. The rest of the proof is easier.

We now are left to prove that

$$b(x) = 0, \quad \forall x \in \Omega = \{|x| < R\}. \tag{3.57}$$

Consider the time cylinder $Q_T^{\xi_0\delta^2} \subset Q_T$,

$$Q_T^{\xi_0\delta^2} = \{(x,t) : \sqrt{R^2 - \xi_0\delta^2} < |x| < R, t \in (0,T)\}.$$

It follows from (3.57) that in the time cylinder $Q_T^{\xi_0\delta^2}$ conditions (3.12)–(3.14) become

$$Mv = c_1(x)v_{tt} - \Delta v - \sum_{j=1}^{n} a_\alpha(x) D_x^\alpha v = 0, \quad (x,t) \in Q_T^{\xi_0\delta^2}, \tag{3.58}$$

$$v|_{S_T} = \frac{\partial v}{\partial n}\Big|_{S_T} = 0. \tag{3.59}$$

$$v(x,0) = 0, \quad v_t(x,0) = 0. \tag{3.60}$$

Consider a number $t_0 \in (0, T - \delta)$. For $\xi \in (0, \xi_0)$, consider the domain $H_{\xi\delta^2}(t_0)$,

$$H_{\xi\delta^2}(t_0) = \{(x,t) : |x|^2 - \xi(t - t_0)^2 > R^2 - \xi\delta^2, t > 0, |x| < R\}.$$

Since $t_0 \in (0, T - \delta)$, then $H_{\xi\delta^2}(t_0) \subset Q_T^{\xi\delta^2}$. Hence, applying a slightly modified discussion of the above part of the proof to (3.58)–(3.60), we obtain the following analog of (3.55): $v(x,t) = 0$ in $H_{\xi_0\delta^2}(t_0)$. In particular, the latter means that

$$v(x, t_0) = 0 \quad \text{for } x \in \{\sqrt{R^2 - \xi_0\delta^2} < |x| < R\}.$$

Hence, varying the number t_0 in the interval $t_0 \in (0, T - \delta)$, we obtain $v(x,t) = 0$ in $Q_{T-\delta}^{\xi_0\delta^2}$. In particular, this implies that

$$v = \frac{\partial v}{\partial n} = 0 \quad \text{for } (x,t) \in \{|x| = \sqrt{R^2 - \xi_0\delta^2}, t \in (0, T - \delta)\}.$$

Hence, we can repeat the above process now in the domain

$$\{(x,t) : |x| < \sqrt{R^2 - \xi_0\delta^2}, t \in (0, T - \delta)\}.$$

Since δ is a sufficiently small number, we can choose it in such a way that $m\xi_0\delta^2 = (R^2 - \sigma)$, where $m \geq 1$ is an integer and $\sigma > 0$ is an arbitrary sufficiently small number. Hence, making m steps as ones above and assuming that $T > R^2/(\xi_0\delta)$, we obtain (3.57). Hence, the right- hand side of equation (3.12) equals zero for $(x,t) \in Q_T$. Finally, the uniqueness theorem for equation (3.12) with one of boundary conditions (3.14) and initial conditions (3.13) implies that $v(x,t) = 0$ in Q_T. □

3.3 Hyperbolic equation when one of initial conditions equals zero

A slightly inconvenient point of Theorem 3.2.1 is that the observation time T is assumed to be sufficiently large. In this section, we consider the case when one of initial conditions in identically equals zero. Then the BK method can significantly relax the condition of a sufficiently large observation time T. This observation was made in [95–99, 103]. The proof of Theorem 3.3 partially uses arguments of [95–99, 103]. For brevity, we consider here only an analog of Theorem 3.2.1. A corresponding analog of Theorem 3.2.2 is formulated and proved similarly.

Consider slightly modified versions of the Forward problem 3.2 and CIP 3.2.

Forward problem 3.3. Find the solution $u \in C^4(\overline{Q_T})$ of the following initial boundary value problem:

$$c(x)u_{tt} = \Delta u + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha u, \text{ in } Q_T, \tag{3.61}$$

$$u(x,0) = g(x), \quad u_t(x,0) = 0, \tag{3.62}$$

$$u|_{S_T} = p_0(x,t). \tag{3.63}$$

Coefficient Inverse Problem 3.3 (CIP 3.3). Assume that the normal derivative of the function u is known on the boundary $\partial\Omega$ of the domain Ω,

$$\frac{\partial u}{\partial n}\bigg|_{S_T} = p_1(x,t). \tag{3.64}$$

Determine the coefficient $c(x)$ of equation (3.61).

Theorem 3.3. *Suppose that conditions* (3.61)–(3.63) *are satisfied. Let the coefficient* $c(x)$ *satisfy conditions* (3.4)–(3.6). *In addition, let coefficients* $a_\alpha \in C(\overline{\Omega})$. *Assume that the function* $c(x)$ *is unknown while coefficients* $a_\alpha(x)$ *are known. Also, assume that*

$$\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x) \neq 0, \quad \forall x \in \overline{\Omega}$$

and

$$T > \frac{R}{\sqrt{\xi_0}}, \tag{3.65}$$

where $\xi_0 = \xi_0(\overline{c}, R, \|\nabla c\|_{C(\overline{\Omega})}) \in (0,1)$ *is the number of Theorem* 2.5.1 *and it depends only on listed parameters. Then there exists at most one vector function* (u,c) *with* $u \in C^4(\overline{Q_T})$ *satisfying* (3.61)–(3.64). *In particular, if* $c(x) \equiv 1$, *then* $\xi_0 = 1$ (*Corollary* 2.5.2) *and* (3.65) *becomes* $T > R$.

Remark 3.3. Even though we consider here only the case when $u(x,0) \neq 0$, $u_t(x,0) = 0$, the case $u(x,0) = 0$, $u_t(x,0) \neq 0$ can be considered similarly. To do this, one needs to differentiate equation (3.61) once with respect to t and then consider the function $\widehat{u}(x,t) = u_t(x,t)$ instead of the function $u(x,t)$. Then $u \in C^4(\overline{Q_T})$ in Theorem 3.3 will be replaced with $u \in C^5(\overline{Q_T})$.

Proof. Consider two possible pairs of functions (u_1, c_1) and (u_2, c_2). Let $v = u_{1tt} - u_{2tt}$, $\widetilde{c} = c_1 - c_2$. Then, using (3.61)–(3.64), we obtain

$$c_1(x)v_{tt} - \Delta v - \sum_{j=1}^{n} a_\alpha(x) D_x^\alpha v = -\widetilde{c}(x)\partial_t^4 u_2, \quad \text{in } Q_T, \tag{3.66}$$

$$v(x,0) = -\widetilde{c}(x) q(x), \tag{3.67}$$

$$v_t(x,0) = 0, \tag{3.68}$$

$$v|_{S_T} = \frac{\partial v}{\partial n}\Big|_{S_T} = 0, \tag{3.69}$$

$$q(x) = \frac{1}{c_2(x)}\Big(\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)\Big) \neq 0, \quad \forall x \in \overline{\Omega}. \tag{3.70}$$

By (3.67) and (3.70),

$$-\widetilde{c}(x) = \frac{v(x,0)}{q(x)}, \quad \forall x \in \overline{\Omega}.$$

Hence,

$$-\widetilde{c}(x) = \frac{1}{q(x)}\Big(v(x,t) - \int_0^t v_t(x,\tau)d\tau\Big), \quad \forall (x,t) \in \overline{Q}_T. \tag{3.71}$$

Substituting in (3.66) and using (3.68) and (3.69), we obtain

$$\big|c_1(x)v_{tt} - \Delta v\big| \leq M\Big(|\nabla v| + |v| + \int_0^t \big|v_t(x,\tau)\big| d\tau\Big), \quad \forall (x,t) \in \overline{Q}_T, \tag{3.72}$$

$$v_t(x,0) = 0, \tag{3.73}$$

$$v|_{S_T} = \frac{\partial v}{\partial n}\Big|_{S_T} = 0. \tag{3.74}$$

Here and below in this proof, $M > 0$ denotes different positive constants depending only on listed parameters,

$$M = M\Big(\min_{\overline{\Omega}}\Big(\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)\Big), \max_\alpha \|a_\alpha\|_{C(\overline{\Omega})}, \big\|\partial_t^4 u_2\big\|_{C(\overline{Q}_T)}\Big) > 0. \tag{3.75}$$

Let $\varepsilon > 0$ be a sufficiently small number, which we will choose later. Let the number $\xi \in (0, \xi_0)$. Consider the domain $D_{\xi,\varepsilon}$,

$$D_{\xi,\varepsilon} = \big\{(x,t) : |x|^2 - \xi t^2 > \varepsilon, |x| < R, t > 0\big\}. \tag{3.76}$$

We need

$$D_{\xi,\varepsilon} \cap \{t = T\} = \varnothing. \tag{3.77}$$

It follows from (3.65) that if we choose ξ such that

$$\xi_0\left(1 - \frac{\varepsilon}{R^2}\right) < \xi < \xi_0,$$

then (3.77) holds. The boundary $\partial D_{\xi,\varepsilon}$ consists of three parts,

$$\partial D_{\xi,\varepsilon} = \partial_1 D_{\xi,\varepsilon} \cup \partial_2 D_{\xi,\varepsilon} \cup \partial_3 D_{\xi,\varepsilon}, \tag{3.78}$$

$$\partial_1 D_{\xi,\varepsilon} = \{(x,t) : |x| = R\} \cap \overline{\partial D_{\xi,\varepsilon}}, \tag{3.79}$$

$$\partial_2 D_{\xi,\varepsilon} = \{(x,t) : |x|^2 - \xi t^2 = \varepsilon, |x| < R, t > 0\}, \tag{3.80}$$

$$\partial_3 D_{\xi,\varepsilon} = \{(x,t) : \sqrt{\varepsilon} < |x| < R, t = 0\}. \tag{3.81}$$

Let $\lambda_0 = \lambda_0(\Omega, \xi, \overline{c}, \|\nabla c\|_{C(\overline{\Omega})}) > 1$ be the number chosen in Theorem 2.5.1 and let $\lambda \geq \lambda_0$. Square both sides of (3.72) and multiply by the function $\varphi(x,t)$ defined in (3.35), (3.36). Then integrate the obtained inequality over the subdomain $D_{\xi,\varepsilon}$. In doing so, take into account the Carleman estimate (3.38)–(3.40) and the Gauss' formula. We obtain

$$\begin{aligned} M & \int\limits_{D_{\xi,\varepsilon}} \left((\nabla v)^2 + v^2 + \left(\int\limits_0^t |v_t(x,\tau)| d\tau \right)^2 \right) \varphi dx dt \\ & \geq \int\limits_{D_{\xi,\varepsilon}} \left(c_1(x) v_{tt} - \Delta v\right)^2 \varphi dx dt \\ & \geq C\lambda \int\limits_{D_{\xi,\varepsilon}} ((\nabla v)^2 + v_t^2) \varphi dx dt + C\lambda^3 \int\limits_{D_{\xi,\varepsilon}} v^2 \varphi dx dt \\ & \quad + \int\limits_{\partial_1 D_{\xi,\varepsilon}} (W_1, \nu_x) dS + \int\limits_{\partial_2 D_{\xi,\varepsilon}} (W_1, \nu_x) dS + \int\limits_{\partial_2 D_{\xi,\varepsilon}} W_2 \cos(\nu, t) dS - \int\limits_{\partial_3 D_{\xi,\varepsilon}} W_2(x,0) dx. \end{aligned} \tag{3.82}$$

In (3.82), $\nu_x = (\nu_x^1, \nu_x^2, \ldots \nu_x^n, 0)$ is the unit outer normal vector at an arbitrary point (x,t) of either $\{(x,t) : |x| = R, t > 0\} \cap \overline{\partial D_{\xi,\varepsilon}}$ or of $\partial_2 D_{\xi,\varepsilon}$ and (W_1, ν_x) is the scalar product in $\mathbb{R}^n$ of these two vectors. Next, $\nu = (\nu_x, \nu_t^{n+1}) \in \mathbb{R}^{n+1}$ is the unit outer normal vector at an arbitrary point $(x,t) \in \partial_2 D_{\xi,\varepsilon}$.

It follows from (3.39), (3.74), and (3.79) that

$$\int\limits_{\partial_1 D_{\xi,\varepsilon}} (W_1, \nu_x) dS = 0. \tag{3.83}$$

Next, using (3.40), (3.68), and (3.81), we obtain $W_2(x,0)$, $(x,0) \in \partial_3 D_{\xi,\varepsilon}$. Hence,

$$\int_{\partial_3 D_{\xi,\varepsilon}} W_2(x,0)dx = 0. \tag{3.84}$$

Next, since by (3.35), (3.36), and (3.80)

$$\varphi(x,t) = e^{2\lambda\varepsilon}, \quad (x,t) \in \partial_2 D_{\xi,\varepsilon},$$

then (3.39) and (3.40) imply that

$$\begin{aligned} &\int_{\partial_2 D_{\xi,\varepsilon}} (W_1, \nu_x)dS + \int_{\partial_2 D_{\xi,\varepsilon}} W_2 \cos(\nu,t)dS \\ &\quad \geq C\lambda^3 e^{2\lambda\varepsilon} \int_{\partial_2 D_{\xi,\varepsilon}} (|\nabla v|^2 + v_t^2 + v^2)dS. \end{aligned} \tag{3.85}$$

Hence, substituting (3.83)–(3.85) in (3.82), we obtain

$$\begin{aligned} &C\lambda^3 e^{2\lambda\varepsilon} \int_{\partial_2 D_{\xi,\varepsilon}} (|\nabla v|^2 + v_t^2 + v^2)dS \\ &\quad + M \int_{D_{\xi,\varepsilon}} \left((\nabla v)^2 + v^2 + \left(\int_0^t |v_t(x,\tau)| d\tau \right)^2 \right) \varphi dxdt \\ &\quad \geq C\lambda \int_{D_{\xi,\varepsilon}} ((\nabla v)^2 + v_t^2)\varphi dxdt + C\lambda^3 \int_{D_{\xi,\varepsilon}} v^2 \varphi dxdt. \end{aligned} \tag{3.86}$$

Using (3.35) and (3.36), acting similarly with (3.49)–(3.51) and applying Lemma 3.1.1, we obtain

$$\int_{D_{\xi,\varepsilon}} \left(\int_0^t |v_t(x,\tau)| d\tau \right)^2 \varphi dxdt \leq \frac{1}{4\xi\lambda} \int_{D_{\xi,\varepsilon}} v_t^2 \varphi dxdt.$$

Hence, the second line of (3.86) can be estimated as

$$\begin{aligned} &M \int_{D_{\xi,\varepsilon}} \left((\nabla v)^2 + v^2 + \left(\int_0^t |v_t(x,\tau)| d\tau \right)^2 \right) \varphi dxdt \\ &\quad \leq M' \int_{D_{\xi,\varepsilon}} ((\nabla v)^2 + v_t^2 + v^2)\varphi dxdt, \end{aligned} \tag{3.87}$$

where the constant $M' > 0$ depends on the same parameter as the constant M in (3.75). Choose $\lambda_1 > \lambda_0$ so large that $C\lambda_1 > 2M'$. Then, using (3.86) and (3.87), we obtain with a new constant M,

$$M\lambda^3 e^{2\lambda\varepsilon} \int\limits_{\partial_2 D_{\xi,\varepsilon}} (|\nabla v|^2 + v_t^2 + v^2) dS \geq \int\limits_{D_{\xi,\varepsilon}} ((\nabla v)^2 + v_t^2 + v^2)\varphi dxdt, \quad \forall \lambda \geq \lambda_1. \tag{3.88}$$

Choose a sufficiently small number $\delta > 0$ such that the domain $D_{\xi,\varepsilon+\delta} \neq \varnothing$; see (3.77). Then $\varphi(x,t) \geq e^{2\lambda(\varepsilon+\delta)}$ in $D_{\xi,\varepsilon+\delta}$ and obviously $D_{\xi,\varepsilon+\delta} \subset D_{\xi,\varepsilon}$. Hence, $\forall \lambda \geq \lambda_1$,

$$\begin{aligned} \int\limits_{D_{\xi,\varepsilon}} ((\nabla v)^2 + v_t^2 + v^2)\varphi dxdt &\geq \int\limits_{D_{\xi,\varepsilon+\delta}} ((\nabla v)^2 + v_t^2 + v^2)\varphi dxdt \\ &\geq e^{2\lambda(\varepsilon+\delta)} \int\limits_{D_{\xi,\varepsilon+\delta}} ((\nabla v)^2 + v_t^2 + v^2) dxdt, \quad \forall \lambda \geq \lambda_1. \end{aligned}$$

Comparing this with (3.88), we obtain

$$M\lambda^3 e^{-2\lambda\delta} \int\limits_{\partial_2 D_{\xi,\varepsilon}} (|\nabla v|^2 + v_t^2 + v^2) dS \geq \int\limits_{D_{\xi,\varepsilon+\delta}} ((\nabla v)^2 + v_t^2 + v^2) dxdt, \quad \forall \lambda \geq \lambda_1. \tag{3.89}$$

Setting in (3.89) $\lambda \to \infty$, we obtain $v(x,t) = 0$ in $D_{\xi,\varepsilon+\delta}$. Since $\delta, \varepsilon > 0$ are arbitrary sufficiently small number, then $v(x,t) = 0$ in $D_{\xi,0}$. Hence, (3.67) and (3.70) imply that $\tilde{c}(x) = 0$ for $x \in \Omega$. In turn, the latter immediately implies that $u_1(x,t) = u_2(x,t) = 0$ for $(x,t) \in Q_T$. □

3.4 Parabolic equations

The forward problem for any parabolic equation of the second order is considered on the time interval $t \in (0,T)$ with the initial condition at $\{t = 0\}$. In the CIP, the Cauchy data are given on the lateral surface of the time cylinder. However, it turns out that the BK method does not work for this case, at least directly. The reason is a technical one and it is still unknown how to handle this case.

So, BK method works in three cases of CIPs for parabolic equations:

1. *Case 1.* When the inverse operator of the so-called Reznickaya transform [221] is applicable to the solution of that forward problem. In this case, the CIP for the parabolic equation is reduced to a CIP for an associated hyperbolic equation. Then the methods of Section 3.3 is applicable; also see [22, 165, 184] for the Reznickaya transform. The Reznickaya transform is an analog of the Laplace transform. The inversion of the Laplace transform is a very unstable procedure. The same is true for the inversion of the Reznickaya transform.
2. *Case 2.* When the data are given at $\{t = t_0 \in (0,T)\}$, instead of $\{t = 0\}$.

3. *Case 3.* When one considers the CIP with the final over determination, that is, when the data are given at $\{t = T\}$. If, in addition, one assumes that the unknown coefficient is known in a small subdomain, then this case can be reduced to Case 2.

Therefore, we consider in this section the above cases 1–3.

3.4.1 Case 1: Parabolic and hyperbolic equations

Assume again that the domain $\Omega = \{|x| < R\}$, functions

$$c(x) \in C^1(\mathbb{R}^n), \quad a_\alpha(x) \in C^\beta(\mathbb{R}^n), \tag{3.90}$$

$$c(x) \in [1, \overline{c}], \quad \forall x \in \mathbb{R}^n, \text{ where } \overline{c} = \text{const.} \geq 1, \tag{3.91}$$

$$(\nabla c, x) \geq 0, \quad \forall x \in \overline{\Omega}. \tag{3.92}$$

Here and everywhere below, $C^{k+\beta}$, $C^{2k+\beta/2}$ are Hölder spaces with integers $k \geq 0$ and $\beta \in (0,1)$ [80, 174]. Also, denote $D_T^{n+1} = \mathbb{R}^n \times (0,T)$.

Forward problem 3.4.1. This is the Cauchy problem:

$$c(x)u_t = \Delta u + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha u, \quad (x,t) \in D_T^{n+1}, \tag{3.93}$$

$$u(x,0) = g(x), \tag{3.94}$$

$$g \in C^{2+\beta}(\mathbb{R}^n). \tag{3.95}$$

Given conditions (3.90), (3.91), the problems (3.93)–(3.95) have a unique solution [79, 174],

$$u \in C^{2+\beta,1+\beta/2}(\overline{D}_T^{n+1}).$$

Coefficient Inverse Problem 3.4.1 (CIP 3.4.1). Assume that all coefficients of equation (3.93) are known, except of one. Let that coefficient be unknown inside of the domain Ω and is known in $\mathbb{R}^n \setminus \Omega$. Let $\Gamma \subseteq \partial\Omega$ be a part of the boundary $\partial\Omega$ of the domain Ω and let $\Gamma_T = \Gamma \times (0,T)$. Determine that unknown coefficient inside of Ω, assuming that the following functions $p(x,t)$ and $q(x,t)$ are known:

$$u|_{\Gamma_T} = p(x,t), \quad \left.\frac{\partial u}{\partial n}\right|_{\Gamma_T} = q(x,t). \tag{3.96}$$

Theorem 2.6.2 implies that, given functions $p(x,t)$ and $q(x,t)$ in (3.96), the function $u(x,t)$ can be uniquely determined in the domain $(\mathbb{R}^n \setminus \Omega) \times (0,T)$. Hence, recalling that $S_T = \partial\Omega \times (0,T)$, we can assume now that functions $p(x,t)$ and $q(x,t)$ are known for $(x,t) \in S_T$. Thus, we replace (3.96) with

$$u|_{S_T} = p(x,t), \quad \left.\frac{\partial u}{\partial n}\right|_{S_T} = q(x,t). \tag{3.97}$$

Consider the Cauchy problem for a hyperbolic equation:

$$c(x)v_{tt} = \Delta v + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha v \quad \text{in } D_\infty^{n+1} = \mathbb{R}^n \times (0,\infty), \tag{3.98}$$

$$v|_{t=0} = g(x), \quad v_t|_{t=0} = 0. \tag{3.99}$$

In addition to (3.90)–(3.92) and (3.95), we assume that the coefficients $c(x)$, $a_\alpha(x)$ and the initial condition $g(x)$ are so smooth that the solution v of problems (3.98), (3.99) is such that $v \in C^4(\overline{D}_T^{n+1})$, for every $T > 0$. First, we establish conditions guaranteeing this smoothness.

Theorem 3.4.1 was proved in [133] via a combination of results of Chapter 4 of the book of Ladyzhenskaya [173] with embedding theorems. Below

$$\left[\frac{n+1}{2}\right] = \begin{cases} (n+1)/2 & \text{if } n \text{ is an odd number,} \\ n/2 & \text{if } n \text{ is an even number.} \end{cases}$$

Theorem 3.4.1 ([133]). *Assume that all coefficients of the hyperbolic operator* (3.98) *belong to the space $C^{[(n+1)/2]+3}(\mathbb{R}^n)$ and the initial condition $g \in H^{[(n+1)/2]+5}(\mathbb{R}^n)$. Then for every $T > 0$, there exists unique solution $v \in H^2(\mathbb{R}^n \times (0,T))$ of problems* (3.98), (3.99). *Furthermore,*

$$v \in H^{[(n+1)/2]+5}(\mathbb{R}^n \times (0,T)) \subset C^4(\mathbb{R}^n \times [0,T])$$

and the following estimate holds with an arbitrary number $\rho > 0$ for all $T > 0$:

$$\|v\|_{C^4(\mathbb{R}^n \times [0,T])} \leq B \exp(\rho T^2) \|g\|_{H^{[(n+1)/2]+5}(\mathbb{R}^n)},$$

where the constant $B > 0$ depends only on ρ and on $C^{[(n+1)/2]+3}(\mathbb{R}^n)$-norms of coefficients of the hyperbolic operator (3.98).

Remark 3.4.1. A direct analog of Theorem 3.4.1 is valid for the hyperbolic operator $\partial^2 - L$ with an arbitrary operator L, uniformly elliptic in $\mathbb{R}^n$, whose coefficients depend only on x [133].

Consider now an interesting Laplace-like transform, which was proposed, for the first time, by Reznickaya (1973) [221]; also, see [22, 165, 184]. Assume that conditions of Theorem 3.4.1 hold. Choose an arbitrary number $\rho > 0$ and let $t \in (0, 1/(4\rho))$. Consider

$$(Rv)(x,t) = \widehat{v}(x,t) = \frac{1}{\sqrt{\pi t}} \int_0^\infty \exp\left[-\frac{\tau^2}{4t}\right] v(x,\tau)\, d\tau. \tag{3.100}$$

One can directly verify that

$$\frac{\partial}{\partial t}\left(\frac{1}{\sqrt{\pi t}} \exp\left[-\frac{\tau^2}{4t}\right]\right) = \frac{\partial^2}{\partial \tau^2}\left(\frac{1}{2\sqrt{\pi t}} \exp\left[-\frac{\tau^2}{4t}\right]\right), \tag{3.101}$$

$$\frac{\partial}{\partial\tau}\left(\frac{1}{2\sqrt{\pi t}}\exp\left[-\frac{\tau^2}{4t}\right]\right)(t,0)=0. \tag{3.102}$$

Using (3.99), (3.101), and (3.102), we obtain

$$\begin{aligned}\hat{v}_t(x,t)&=\int_0^\infty\frac{\partial^2}{\partial\tau^2}\left(\frac{1}{\sqrt{\pi t}}\exp\left[-\frac{\tau^2}{4t}\right]\right)v(x,\tau)d\tau\\&=-\frac{\partial}{\partial\tau}\left(\frac{1}{\sqrt{\pi t}}\exp\left[-\frac{\tau^2}{4t}\right]\right)(t,0)v(x,0)\\&\quad-\int_0^\infty\frac{\partial}{\partial\tau}\left(\frac{1}{\sqrt{\pi t}}\exp\left[-\frac{\tau^2}{4t}\right]\right)v_\tau(x,\tau)d\tau\\&=\frac{1}{\sqrt{\pi t}}v_\tau(x,0)+\frac{1}{\sqrt{\pi t}}\int_0^\infty\exp\left[-\frac{\tau^2}{4t}\right]v_{\tau\tau}(x,\tau)d\tau\\&=\frac{1}{\sqrt{\pi t}}\int_0^\infty\exp\left[-\frac{\tau^2}{4t}\right]v_{\tau\tau}(x,\tau)d\tau.\end{aligned}$$

Thus,

$$\hat{v}_t(x,t)=\frac{1}{\sqrt{\pi t}}\int_0^\infty\exp\left[-\frac{\tau^2}{4t}\right]v_{\tau\tau}(x,\tau)d\tau=R(v_{\tau\tau})(x,t). \tag{3.103}$$

Next,

$$\hat{v}(x,t)=\frac{2}{\sqrt{\pi}}\int_0^\infty v(x,2\sqrt{t}z)e^{-z^2}dz.$$

Hence,

$$\lim_{t\to 0^+}\hat{v}(x,t)=v(x,0)=g(x). \tag{3.104}$$

Hence, it follows from (3.100)–(3.104) and Theorem 3.4.1 that solutions u and v of forward problems (3.93), (3.94) and (3.98), (3.99) are connected via

$$u(x,t)=(Rv)(x,t)=\hat{v}(x,t). \tag{3.105}$$

Changing variables in (3.100) $\tau^2=z$, $1/(4t)=s$, we obtain

$$\hat{v}\left(x,\frac{1}{4s}\right)=\frac{\sqrt{s}}{\sqrt{\pi}}\int_0^\infty\frac{v(x,\sqrt{z})}{\sqrt{z}}e^{-sz}dz,\quad s>\rho. \tag{3.106}$$

Hence, (3.100) is an analog of the Laplace transform. Since this transform is one-to-one for $s > \rho$, then (3.97) and (3.105) imply that functions $\overline{p}(x,t) = R^{-1}(p(x,t))$ and $\overline{q}(x,t) = R^{-1}(q(x,t))$ are known, where

$$v|_{S_\infty} = \overline{p}(x,t), \quad \left.\frac{\partial v}{\partial n}\right|_{S_\infty} = \overline{q}(x,t). \tag{3.107}$$

Therefore, uniqueness theorem for CIP 3.4.1 is equivalent to the uniqueness theorem for the hyperbolic CIP (3.98), (3.99), (3.107). Therefore, we can now directly apply uniqueness theorems for the hyperbolic CIPs for the parabolic CIP 3.4.1. Thus, using Theorem 3.3, we arrive at Theorem 3.4.1.

Theorem 3.4.2. *Assume that conditions* (3.90)–(3.92) *as well as conditions of Theorem* 3.4.1 *hold. In the case when the coefficient* $c(x)$ *in* (3.93) *is unknown, assume that*

$$\Delta g(x) + \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha g(x) \neq 0, \quad \forall x \in \overline{\Omega}.$$

And in the case when a coefficient $a_{\alpha_0}(x)$ *with a fixed* α_0 *is unknown, assume that*

$$D_x^{\alpha_0} g(x) \neq 0, \quad \forall x \in \overline{\Omega}.$$

Then CIP 3.4.1 *has at most one solution.*

3.4.2 Case 2: The data at $\{t = t_0 \in (0,T)\}$ as well as the lateral Cauchy data

In this section, we assume that we have lateral Cauchy data for a parabolic equation as well as the data at the moment of time $\{t = t_0 \in (0,T)\}$. These assumptions allow us to consider a general elliptic operator in the parabolic equation rather than the one in (3.93) with condition (3.92). Furthermore, combining the technique of this section with the technique of Theorem 2.6.3, one can obtain Hölder stability estimate for the CIP considered in this section, although we do not do this here. Even more: Imanuvilov and Yamamoto have obtained Lipschitz stability estimate for this problem [96, 252]. However, we do not obtain stability estimates in this section. We point out that the assumption of this section that the data are given at $\{t = t_0 \in (0,T)\}$ are used in all works devoted to applications of the BK method to parabolic CIPs; see, for example, [32, 39, 67, 96, 252].

In this section, $\Omega \subset \mathbb{R}^n$ is a bounded domain, and its boundary $\partial\Omega$ is piecewise smooth, $\Gamma \in C^2$, $\Gamma \subseteq \partial\Omega$ is a part of $\partial\Omega$ and $T = \text{const} > 0$. Recall that $Q_T = \Omega \times (0,T)$, $\Gamma_T = \Gamma \times (0,T)$. We denote in this section $x = (x_1, y) = (x_1, x_2, \dots, x_n)$, where $y = (x_2, \dots, x_n)$. Let L be the following elliptic operator in Ω:

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x) u_{x_i x_j} + \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha u, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \Omega, \tag{3.108}$$

$$a_{ij} \in C^1(\overline{\Omega}), \quad a_\alpha \in C(\overline{\Omega}), \tag{3.109}$$

$$\mu_1|\xi|^2 \le \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \le \mu_2|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \overline{\Omega}, \tag{3.110}$$

$$\mu_1, \mu_2 = \text{const.} > 0, \quad \mu_1 \le \mu_2. \tag{3.111}$$

We represent the operator L as

$$Lu = L_0 u + L_1 u,$$
$$L_0 u = \sum_{i,j=1}^{n} a_{ij}(x) u_{x_i x_j},$$
$$L_1 u = \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha u.$$

Thus, L_c is the principal part of the operator L and L_1 is the sum of its low order terms.

Coefficient Inverse Problem 3.4.2 (CIP 3.4.2). Let the function $u \in C^{4,2}(\overline{Q}_T)$ satisfies the parabolic equation

$$u_t = Lu + F(x,t) \quad \text{in } Q_T. \tag{3.112}$$

Let the number $t_0 \in (0, T)$. Determine the unknown coefficient of the operator L for $x \in \Omega$ assuming that the function $F(x,t)$ is known in Q_T and that the following functions $f(x)$, $p(x,t)$, and $q(x,t)$ are known as well:

$$u(x, t_0) = f(x), \quad x \in \Omega, \tag{3.113}$$

$$u|_{\Gamma_T} = p(x,t), \quad \left.\frac{\partial u}{\partial n}\right|_{\Gamma_T} = q(x,t). \tag{3.114}$$

Theorem 3.4.3. *Assume that conditions* (3.108)–(3.114) *hold. Also, assume that:*
1. *If that unknown coefficient is $a_{i_0 j_0}(x)$, then $\partial^2_{x_{j_0} x_{i0}} f(x) \ne 0$ in $\overline{\Omega}$.*
2. *If that unknown coefficient is $a_{\alpha_0}(x)$, then $D^{\alpha_0} f(x) \ne 0$ in $\overline{\Omega}$, where α_0 is a fixed multiindex.*

Then the CIP 3.4.2 *has at most one solution.*

Proof. Without loss of generality, we can assume that

$$\Gamma = \{(x_1, y) : x_1 = g(y), |y| < r\},$$

where $r > 0$ is a certain number and the function $g \in C^2(|y| \le r)$. Assume, for example, that the coefficient $a_{i_0 j_0}(x)$ is unknown. Suppose that there exists two pairs of functions $(a^1_{i_0 j_0}, u_1)$ and $(a^2_{i_0 j_0}, u_2)$ satisfying conditions (3.108)–(3.114). Denote

$$b(x) = a^1_{i_0 j_0} - a^2_{i_0 j_0}, \quad \tilde{u}(x,t) = u_1(x,t) - u_2(x,t), \quad R(x,t) = \partial^2_{x_{i_0} x_{j_0}} u_2(x,t). \tag{3.115}$$

Using (3.112)–(3.115), we obtain

$$\tilde{u}_t - L^{(1)}\tilde{u} = b(x)R(x,t), \quad (x,t) \in Q_T, \tag{3.116}$$

$$\tilde{u}(x,t_0) = 0, \quad x \in \Omega, \tag{3.117}$$

$$\tilde{u}|_{\Gamma_T} = \left.\frac{\partial \tilde{u}}{\partial n}\right|_{\Gamma_T} = 0, \tag{3.118}$$

where $L^{(1)}$ is the operator L in which the coefficient $a_{i_0 j_0}(x)$ is replaced with $a^1_{i_0 j_0}(x)$. Denote $L_0^{(1)}$ the principal part of the operator $L^{(1)}$.

Since $\partial^2_{x_{j_0} x_{i_0}} f(x) \neq 0$ in $\overline{\Omega}$, then by (3.113) and (3.115) there exists a sufficiently small number $\sigma > 0$ such that

$$R(x,t) \neq 0 \quad \text{for } (x,t) \in \overline{Q}_{t_0,\sigma} = \overline{\Omega} \times [t_0 - \sigma, t_0 + \sigma] \subset Q_T. \tag{3.119}$$

Divide both sides of equation (3.116) by $R(x,t)$ and denote

$$v(x,t) = \frac{\tilde{u}(x,t)}{R(x,t)}.$$

We obtain

$$v_t - \hat{L}v = b(x), \quad (x,t) \in Q_{t_0,\sigma}, \tag{3.120}$$

$$v(x,0) = 0, \tag{3.121}$$

$$v|_{\Gamma_T} = \left.\frac{\partial v}{\partial n}\right|_{\Gamma_T} = 0, \tag{3.122}$$

where $\hat{L}$ is another elliptic operator, whose principal part is the same as in $L_0^{(1)}$, although coefficients at lower order derivatives depend on both x and t. Differentiate both sides of (3.120) with respect to t and denote $w(x,t) = v_t(x,t)$. Then (3.121) implies

$$v(x,t) = \int_0^t w(x,\tau)d\tau. \tag{3.123}$$

Since $\partial_t b(x) \equiv 0$, then (3.120) implies $w_t - \partial_t(\hat{L}v) = 0$ in $Q_{t_0,\sigma}$. Hence, (3.120)–(3.123) lead to

$$\left|w_t - L_0^{(1)}w\right| \leq A\left(|\nabla w| + |w| + \left|\int_{t_0}^t (|\nabla w| + |w|)(x,\tau)d\tau\right|\right), \quad (x,t) \in Q_{t_0,\sigma}, \tag{3.124}$$

$$w = \frac{\partial w}{\partial n} = 0, \quad (x,t) \in \Gamma \times (t_0 - \sigma, t_0 + \sigma) \tag{3.125}$$

with a positive constant $A > 0$ which is independent on the function w. Below in this section, $A > 0$ denotes different positive constants independent on the function w as well as on parameters of the Carleman estimate.

To apply the Carleman estimate of Theorem 2.3.1, we arrange the part Γ of the boundary $\partial\Omega$ to be a part of the hyperplane $\{x_1 = 0\}$. To do this, we change variables as $(x_1, y) \Leftrightarrow (x_1', y) = (x_1 - g(y), y)$. For brevity, we keep the same notation for new variables and the principal part L_c of the elliptic operator: this new operator still holds properties (3.109)–(3.111). Also, we assume that a part $\Omega'' \subset \Omega'$ of the transformed domain Ω is such that $\Omega'' \subset \{x_1 > 0, |y| < r\}$ and there exists a part $\partial_1\Omega'' \subset \partial\Omega''$ of the boundary $\partial\Omega''$ such that $\partial_1\Omega'' \subset \{x_1 = 0, |y| < r\}$. We keep the same notation

$$Q_{t_0,\sigma} = (x,t) \in \Omega'' \times (t_0 - \sigma, t_0 + \sigma).$$

Thus, (3.124), (3.125) become

$$\left|w_t - L_0^{(1)} w\right| \le A\left(|\nabla w| + |w| + \left|\int_{t_0}^{t} (|\nabla w| + |w|)(x,\tau) d\tau\right|\right), \quad (x,t) \in Q_{t_0,\sigma}, \tag{3.126}$$

$$w(0,y,t) = w_{x_1}(0,y,t) = 0, \quad (0,y) \in \partial_1\Omega'', t \in (t_0 - \sigma, t_0 + \sigma). \tag{3.127}$$

Consider the function $\psi(x,t)$,

$$\psi(x,t) = x_1 + y^2 + \frac{(t-t_0)^2}{\sigma^2} + \alpha, \tag{3.128}$$

where $\alpha > 0$ is a parameter. Let $h > \alpha$ be another parameter. Denote

$$P_h = \{(x,t) : \psi(x,t) < h, x_1 > 0\} = \left\{x_1 + y^2 + \frac{(t-t_0)^2}{\sigma^2} + \alpha < h, x_1 > 0\right\}. \tag{3.129}$$

Choose parameters α and h so small that $P_h \subset Q_{t_0,\sigma}$ and also

$$\{(x,t) \in \partial P_h : x_1 = 0\} \subset \left(\partial_1\Omega'' \times (t_0 - \sigma, t_0 + \sigma)\right). \tag{3.130}$$

It follows from (3.129) that boundary ∂P_h of the domain P_h consists of two parts,

$$\begin{aligned}
\partial P_h &= \partial_1 P_h \cup \partial_2 P_h, \\
\partial_1 P_h &= \left\{x_1 = 0, y^2 + \frac{(t-t_0)^2}{\sigma^2} + \alpha < h\right\}, \\
\partial_2 P_h &= \left\{x_1 + y^2 + \frac{(t-t_0)^2}{\sigma^2} + \alpha = h, x_1 > 0\right\}.
\end{aligned}$$

By (3.127) and (3.130),

$$w(x,t) = w_{x_1}(x,t) = 0, \quad (x,t) \in \partial_1 P_h. \tag{3.131}$$

For parameters $\lambda, \nu > 1$, denote

$$\varphi(x,t) = \exp(2\lambda\psi^{-\nu}).$$

Then, using (3.128) and (3.129), we obtain

$$\min_{\overline{P}_h} \varphi(x,t) = \varphi(x,t)|_{\partial_2 P_h} = \exp(2\lambda h^{-\nu}). \tag{3.132}$$

It follows from Theorem 2.3.1 that the following Carleman estimate holds for any function $z \in C^{2,1}(\overline{P}_h)$:

$$(z_t - L_0^{(1)} z)^2 \varphi \geq C\lambda\nu(\nabla z)^2\varphi + C\lambda^3\nu^4\psi^{-2\nu-2}z^2\varphi + \operatorname{div} U + V_t, \tag{3.133}$$

$$|U| + |V| \leq C\lambda^3\nu^3\psi^{-2\nu-2}((\nabla z)^2 + z^2)\varphi, \tag{3.134}$$

for all $\lambda \geq \lambda_0 > 1$ and all $\nu \geq \nu_0 > 1$, where λ_0, ν_0 and $C > 0$ are certain numbers independent on z, and C is also independent on λ and ν. Here, $(x,t) \in \overline{P}_h$.

Square both sides of (3.126), multiply by the function $\varphi(x,t)$ and apply (3.133) and (3.134). We obtain for $(x,t) \in \overline{P}_h$,

$$A\left(|\nabla w|^2 + |w|^2 + \left(\int_{t_0}^{t} (|\nabla w| + |w|)(x,\tau)d\tau \right)^2 \right) \tag{3.135}$$
$$\geq \lambda(\nabla w)^2\varphi + \lambda^3 w^2\varphi + \operatorname{div} U + V_t,$$

$$|U| + |V| \leq C\lambda^3\nu^3\psi^{-2\nu-2}((\nabla w)^2 + w^2)\varphi. \tag{3.136}$$

Integrate (3.135) over P_h using Gauss' formula and taking into account (3.131), (3.132), and (3.136). We obtain

$$C\lambda^3\nu^3 h^{-2\nu-2}\exp(2\lambda h^{-\nu}) \int_{\partial_2 P_h} ((\nabla w)^2 + w^2)dS$$
$$+ A\int_{P_h}\left(|\nabla w|^2 + |w|^2 + \left(\int_{t_0}^{t} (|\nabla w| + |w|)(x,\tau)d\tau \right)^2 \right)dxdt \tag{3.137}$$
$$\geq \lambda \int_{P_h} (\nabla w)^2\varphi dxdt + \lambda^3 \int_{P_h} w^2\varphi dxdt.$$

Fix a number $\nu \geq \nu_0$. We now use Lemma 3.1.1, similarly with (3.49)–(3.51). Choosing a sufficiently large $\lambda_1 \geq \lambda_0$, we obtain from (3.137) for all $\lambda \geq \lambda_1$,

$$A\lambda^3\nu^3 h^{-2\nu-2}\exp(2\lambda h^{-\nu}) \int_{\partial_2 P_h} ((\nabla w)^2 + w^2)dS \tag{3.138}$$
$$\geq \lambda \int_{P_h} (\nabla w)^2\varphi dxdt + \lambda^3 \int_{P_h} w^2\varphi dxdt.$$

Choose a sufficiently small number $\varepsilon > 0$ such that

$$P_{h-\varepsilon} = \left\{ x_1 + y^2 + \frac{(t-t_0)^2}{\sigma^2} + \alpha < h - \varepsilon, x_1 > 0 \right\} \neq \varnothing,$$

i. e. $h - \varepsilon > \alpha$. Clearly $P_{h-\varepsilon} \subset P_h$ and

$$\varphi(x,t) > \exp(2\lambda(h-\varepsilon)^{-\nu}) > \exp(2\lambda h^{-\nu}), \quad (x,t) \in P_{h-\varepsilon}.$$

Hence, using (3.138), we obtain

$$A\lambda^3\nu^3 h^{-2\nu-2}\exp(2\lambda h^{-\nu}) \int\limits_{\partial_2 P_h} ((\nabla w)^2 + w^2) dS$$
$$\geq \int\limits_{P_{h-\varepsilon}} ((\nabla w)^2 + w^2)\varphi dxdt \geq \exp(2\lambda(h-\varepsilon)^{-\nu}) \int\limits_{P_{h-\varepsilon}} ((\nabla w)^2 + w^2) dxdt.$$

Or

$$A\lambda^3\nu^3 h^{-2\nu-2}\exp[-2\lambda((h-\varepsilon)^{-\nu} - h^{-\nu})] \int\limits_{\partial_2 P_h} ((\nabla w)^2 + w^2) dS \tag{3.139}$$
$$\geq \int\limits_{P_{h-\varepsilon}} ((\nabla w)^2 + w^2) dxdt.$$

Setting in (3.139) $\lambda \to \infty$, we obtain

$$\int\limits_{P_{h-\varepsilon}} ((\nabla w)^2 + w^2) dxdt = 0.$$

Hence, $w(x,t) = 0$ in $P_{h-\varepsilon}$. Since $\varepsilon > 0$ is an arbitrary sufficiently small number, then $w(x,t) = 0$ in P_h. This, (3.120), (3.123), and (3.129) imply that

$$b(x) = 0 \quad \text{for } x \in \{x_1 + y^2 < h - \alpha, x_1 > 0\}.$$

Hence, using (3.116), (3.118), and Theorem 2.6.2, we obtain

$$\widetilde{u}(x,t) = 0 \quad \text{for } (x,t) \in \{x_1 + y^2 < h - \alpha, x_1 > 0\} \times (0,T).$$

It is clear that, continuing this way, we will arrive at $b(x) = 0$ in Ω and $\widetilde{u}(x,t) = 0$ in Q_T. □

3.4.3 Case 3: Final overdetermination

Let the elliptic operator L have the same form as in (3.108)–(3.111), where, however, the bounded domain Ω is replaced with the entire space $\mathbb{R}^n$. Thus,

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x) u_{x_i x_j} + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha u, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \mathbb{R}^n, \tag{3.140}$$

$$\mu_1|\xi|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leq \mu_2|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \mathbb{R}^n, \tag{3.141}$$

$$\mu_1, \mu_2 = \text{const.} > 0, \quad \mu_1 \leq \mu_2. \tag{3.142}$$

Let the number $\beta \in (0,1)$ and let $k \geq 0$ be an integer. Following Theorem 3.4.1, we assume that in (3.140),

$$a_{ij}, a_\alpha \in C^{[(n+1)/2]+3}(\mathbb{R}^n). \tag{3.143}$$

Let the number $T = \text{const.} > 0$. Denote $D_T^{n+1} = \mathbb{R}^n \times (0,T)$. Consider the Cauchy problem

$$u_t = Lu, \quad (x,t) \in D_T^{n+1}, \tag{3.144}$$

$$u|_{t=0} = f(x), \tag{3.145}$$

where the function

$$f \in H^{[(n+1)/2]+3}(\mathbb{R}^n). \tag{3.146}$$

Given conditions (3.140)–(3.143) and (3.146), the problems (3.144), (3.145) have a unique solution $u \in C^{2+\beta,1+\beta/2}(\overline{D}_T^{n+1})$ and actually the function $u(x,t)$ has a better smoothness of course [174].

Coefficient Inverse Problem 3.4.3 (CIP 3.4.3). Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Suppose that one of coefficients of the operator L in (3.140) is known inside of Ω and is unknown outside of Ω while all other coefficients of L are known everywhere. Assume that the initial condition $f(x)$ is also unknown. Determine both that coefficient for $x \in \mathbb{R}^n \smallsetminus \Omega$ and the initial condition $f(x)$ for $x \in \mathbb{R}^n$, assuming that the following function $F(x)$ is known:

$$F(x) = u(x,T), \quad x \in \mathbb{R}^n. \tag{3.147}$$

Theorem 3.4.4. *Assume that conditions* (3.140)–(3.142), (3.146) *hold. Also, assume that all coefficients of the operator L are such that*

$$a_{ij}, a_\alpha \in C^\infty(\Omega). \tag{3.148}$$

In addition, let

$$\frac{\partial^2}{\partial x_{j_0} \partial x_{i_0}} F(x) \neq 0 \quad \text{in } \mathbb{R}^n \smallsetminus \Omega$$

if the coefficient $a_{i_0 j_0}(x)$ is unknown in $\mathbb{R}^n \smallsetminus \Omega$ and let

$$D^{\alpha_0} F(x) \neq 0 \quad \text{in } \mathbb{R}^n \smallsetminus \Omega$$

if the coefficient $a_{\alpha_0}(x)$ is unknown in $\mathbb{R}^n \smallsetminus \Omega$, where α_0 is a fixed multiindex. Then there exists at most one solution of CIP 3.4.3.

Proof. Consider the Cauchy problem for the following hyperbolic PDE:

$$v_{tt} = Lv \quad \text{in } D_{\infty}^{n+1},$$
$$v(x,0) = f(x), \quad v_t(x,0) = 0.$$

By Theorem 3.4.1 and Remark 3.4.1 for any number $T' = \text{const.} > 0$, this problem has unique solution $v \in H^2(\mathbb{R}^n \times (0,T'))$. This function

$$v \in H^{[(n+1)/2]+5}(\mathbb{R}^n \times (0,T)) \subset C^4(\mathbb{R}^n \times [0,T])$$

and also the following estimate holds with an arbitrary number $\rho > 0$ for all $T' > 0$:

$$\|v\|_{C^4(\mathbb{R}^n \times [0,T'])} \leq B \exp(\rho T^2) \|f\|_{H^{[(n+1)/2]+5}(\mathbb{R}^n)}.$$

Let $T < 1/(8\rho)$. Hence, considering the Reznickaya transform (3.100), we obtain for $t \in (0,2T)$,

$$u(x,t) = \frac{1}{\sqrt{\pi t}} \int_0^{\infty} \exp\left[-\frac{\tau^2}{4t}\right] v(x,\tau)\, d\tau. \tag{3.149}$$

It follows from (3.106) with $s = 1/(4t)$ and (3.149) that for any given point $x \in \mathbb{R}^n$ the function $u(x,t)$ is analytic with respect to $t \in (0,2T)$ as the function of the real variable t. Also, it follows from (3.148) that the function $u \in C^{\infty}(\Omega \times (0,2T))$ [79]. Hence, using (3.147), we obtain

$$u_t(x,T) = L(F)(x), \quad \partial_t^2 u(x,T) = L^2(F)(x), \ldots, \partial_t^{k+1} u(x,T) = L^k(F)(x), \tag{3.150}$$
$$x \in \Omega, \quad k = 2,3,\ldots. \tag{3.151}$$

Since the function $u(x,t)$ is analytic with respect to $t \in (0,2T)$, then (3.150) and (3.151) imply that the function $u(x,t)$ is uniquely determined for $(x,t) \in \Omega \times (0,2T)$. Given this, the assertion of Theorem 3.4.4 about the unknown coefficient follows immediately from Theorem 3.4.2.

Next, since the unknown coefficient is determined uniquely, so as the function $u(x,t)$ for $(x,t) \in \Omega \times (0,2T)$, then by Theorem 2.6.2 the function $u(x,t)$ is also uniquely determined in Q_{2T}. Hence, the initial condition $f(x) = u(x,0)$ is also determined uniquely. □

Some other uniqueness theorems for parabolic CIPs with final overdetermination can be found in [100, 102].

3.5 A coefficient inverse problem for an elliptic equation

Consider a convex domain $\Omega \subset \mathbb{R}^n$ whose boundary $\partial\Omega$ is piecewise smooth. Consider a piece $\Gamma \subset \partial\Omega$ of the boundary and let $\Gamma \in C^2$. Let the number $T > 0$. Recall that

$$Q_T^{\pm} = \Omega \times (-T, T), \quad \Gamma_T^{\pm} = \Gamma \times (-T, T).$$

Consider the same elliptic operator L in Ω as the one in Section 3.4.2,

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j} + \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha u, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \Omega, \tag{3.152}$$

$$a_{ij} \in C^1(\overline{\Omega}), \quad a_\alpha \in C(\overline{\Omega}), \tag{3.153}$$

$$\mu_1 |\xi|^2 \le \sum_{i,j=1}^{n} a_{ij}(x)\xi_i \xi_j \le \mu_2 |\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \forall x \in \overline{\Omega}, \tag{3.154}$$

$$\mu_1, \mu_2 = \text{const.} > 0. \tag{3.155}$$

Coefficient Inverse Problem 3.5.1 (CIP 3.5.1). Let conditions (3.152)–(3.155) hold. Let the function $u \in C^3(\overline{Q}_{\pm T})$ satisfy the following conditions:

$$u_{tt} + Lu = H(x,t) \quad \text{in } Q_T^{\pm}, \tag{3.156}$$

$$u(x,0) = p(x) \quad \text{in } \Omega, \tag{3.157}$$

$$u|_{\Gamma_T^{\pm}} = q_0(x,t), \quad \partial_n u|_{\Gamma_T^{\pm}} = q_1(x,t). \tag{3.158}$$

Assume that one of coefficients of the operator L in (3.152) is unknown while all other coefficients of L are known. Determine that coefficient, assuming that conditions (3.156)–(3.158) with $u \in C^3(\overline{Q}_{\pm T})$ are satisfied and functions $H(x,t)$, $p(x)$, $q_0(x,t)$, $q_1(x,t)$ are known.

Theorem 3.5.1. *In the case when the coefficient $a_{i_0 j_0}(x)$ is unknown, assume that*

$$\frac{\partial^2}{\partial x_{j_0} \partial x_{i_0}} p(x) \ne 0 \quad \text{in } \overline{\Omega}.$$

And in the case when the coefficient $a_{\alpha_0}(x)$ is unknown, assume that $D^{\alpha_0} p(x) \ne 0$ in $\overline{\Omega}$, where α_0 is a fixed multiindex. Then CIP 3.5.1 *has at most one solution.*

Proof. Recall that Theorem 2.4.1 provides the Carleman estimate for the elliptic operator. Therefore, the proof of Theorem 3.5.1 is completely similar with the proof of Theorem 3.4.2. □

3.6 Lipschitz stability estimate of a CIP for a hyperbolic equation

In this section, we derive a Lipschitz stability estimate for an analog of the CIP 3.3. The uniqueness Theorem 3.3 was proved for the latter. It is convenient for us to work here within the framework of Theorem 3.4.1. Hence, we consider the Cauchy problem as the forward problem rather than the initial boundary value problem of Section 3.3. More precisely, the forward problem now is

$$c(x)u_{tt} = \Delta u + \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha u, \quad \text{in } D_T^{n+1}, \tag{3.159}$$

$$u(x,0) = g(x), \quad u_t(x,0) = 0, \quad x \in \mathbb{R}^n. \tag{3.160}$$

We assume that

$$c, a_\alpha \in C^{[(n+1)/2]+3}(\mathbb{R}^n), \quad g \in H^{[(n+1)/2]+5}(\mathbb{R}^n), \tag{3.161}$$

$$\max_\alpha \|a_\alpha\|_{C^{[(n+1)/2]+3}(\mathbb{R}^n)}, \|c\|_{C^{[(n+1)/2]+3}(\mathbb{R}^n)}, \|g\|_{H^{[(n+1)/2]+5}(\mathbb{R}^n)} < M, \tag{3.162}$$

$$c(x) \in [1, M], \quad \forall x \in \mathbb{R}^n, \tag{3.163}$$

where $M > 0$ is a known number. Then by Theorem 3.4.1, for each $T > 0$, there exists unique solution $u \in H^2(D_T^{n+1})$ of problems (3.159)–(3.163). Furthermore,

$$u \in C^4(\overline{D}_T^{n+1}) \tag{3.164}$$

and the following estimate holds for all $T > 0$:

$$\|u\|_{C^4(\overline{D}_T^{n+1})} \le B \exp(T^2) \|g\|_{H^{[(n+1)/2]+5}(\mathbb{R}^n)}. \tag{3.165}$$

In (3.165), we take $\exp(T^2)$ instead of $\exp(\rho T^2)$ with an arbitrary $\rho > 0$ of Theorem 3.4.1 for the sake of convenience. In (3.165), the constant $B = B(M) > 0$ depends only on M.

Just like in Sections 3.2, 3.3, in this section $\Omega = \{x \in \mathbb{R}^n : |x| < R\}$. Let the number $d > 0$. Let the point x_0 be such that $x_0 \in \mathbb{R}^3 \setminus \overline{\Omega}$. Let the number $h > 0$ be such that

$$|x_0| - R > 2\sqrt{h}. \tag{3.166}$$

To apply the Carleman estimate of Theorem 2.5.1, we impose the following analog of condition (3.6):

$$(x - x_0, \nabla c) \ge d > 0, \quad \forall x \in \overline{\Omega}. \tag{3.167}$$

Coefficient Inverse Problem 3.6 (CIP 3.6). Let the function $u(x,t)$ be the solution of the problems (3.159)–(3.164). Assume that the coefficient $c(x)$ is unknown in the domain Ω and is known outside of it. Determine the coefficient $c(x)$ for $x \in \Omega$, assuming that the following functions $p_0(x,t)$ and $p_1(x,t)$ are known:

$$u|_{S_T} = p_0(x,t), \quad \left.\frac{\partial u}{\partial n}\right|_{S_T} = p_1(x,t). \tag{3.168}$$

Note that in real measurements only the function $p_0(x,t)$ is known. To find the Neumann boundary condition $p_1(x,t)$, one can solve equation (3.159) in the domain $(\mathbb{R}^n \setminus \Omega) \times (0,T)$ with the initial conditions (3.160) and the Dirichlet boundary condition $p_0(x,t)$.

Theorem 3.6. *Assume that conditions* (3.159)–(3.163) *and* (3.166) *hold. Suppose that there exist two functions* $c_1(x)$, $c_2(x)$ *satisfying conditions* (3.162), (3.163), (3.167), *which generate two solutions* $u_1(x,t), u_2(x,t) \quad \in\in \quad H^2(D_T^{n+1})$ *of problems* (3.159), (3.160). *By* (3.168), *let*

$$u_1|_{S_T} = p_{0,1}(x,t), \quad \left.\frac{\partial u_1}{\partial n}\right|_{S_T} = p_{1,1}(x,t),$$
$$u_2|_{S_T} = p_{0,2}(x,t), \quad \left.\frac{\partial u_2}{\partial n}\right|_{S_T} = p_{1,2}(x,t).$$

Also assume that

$$\frac{1}{c}\left|\Delta g(x) + \sum_{|\alpha|\le 1} a_\alpha(x) D_x^\alpha g(x)\right| \ge r_0 = \text{const.} > 0, \quad x \in \overline{\Omega}, \tag{3.169}$$

$$T > \frac{\sqrt{(R+|x_0|)(R+|x_0|+h)}}{\sqrt{\xi_0}}, \tag{3.170}$$

where r_0 *is a certain number,* h *is as in* (3.166), *and the number* $\xi_0 = \xi_0(M,x_0) \in (0,1)$ *is the number of Theorem* 2.5.1. *Then there exists a constant*

$$C = C(M,d,r_0,T,x_0) > 0 \tag{3.171}$$

depending only on listed parameters such that the following Lipschitz stability estimate holds:

$$\begin{aligned} &\|c_1 - c_2\|_{L_2(\Omega)} \\ &\quad \le C\big(\|\partial_t^2 p_{0,1} - \partial_t^2 p_{0,2}\|_{H^1(S_T)} + \|\partial_t^2 p_{1,1} - \partial_t^2 p_{1,2}\|_{L_2(S_T)}\big). \end{aligned} \tag{3.172}$$

Proof. This proof combines ideas of proofs of Theorems 2.7.1 and 3.3. In this proof, $C > 0$ denotes different constants depending only on parameters listed in (3.171).

Denote

$$\widetilde{c} = c_1 - c_2, \quad \widetilde{u} = u_1 - u_2, \quad \widetilde{p}_0 = p_{0,1} - p_{0,2}, \quad \widetilde{p}_1 = p_{1,1} - p_{1,2}.$$

We obtain

$$c_1(x)\widetilde{u}_{tt} - \Delta\widetilde{u} - \sum_{j=1}^{n} a_\alpha(x) D_x^\alpha \widetilde{u} = -\widetilde{c}(x) u_{2tt}, \quad \text{in } Q_T, \tag{3.173}$$

$$\widetilde{u}(x,0) = \widetilde{u}_t(x,0) = 0, \tag{3.174}$$

$$\tilde{u}|_{S_T} = \tilde{p}_0, \quad \left.\frac{\partial \tilde{u}}{\partial n}\right|_{S_T} = \tilde{p}_1. \tag{3.175}$$

It follows from (3.159), (3.160), (3.163), and (3.169) that

$$|u_{2tt}(x,0)| \geq \frac{1}{\overline{c}}\left|\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)\right| \geq r_0, \quad \forall x \in \overline{\Omega}. \tag{3.176}$$

Differentiate (3.173) and (3.175) twice with respect to t and denote

$$v(x,t) = \tilde{u}_{tt}(x,t). \tag{3.177}$$

Using (3.164) and (3.174), we obtain

$$c_1(x)v_{tt} - \Delta v - \sum_{|\alpha|\leq 1} a_\alpha(x) D^\alpha v = -\tilde{c}(x)\partial_t^4 u_2 \quad \text{in } Q_T, \tag{3.178}$$

$$v(x,0) = \tilde{c}(x) f(x), \tag{3.179}$$

$$v_t(x,0) = 0, \tag{3.180}$$

$$v|_{S_T} = \tilde{p}_{0tt}(x,t), \quad \left.\frac{\partial u}{\partial n}\right|_{S_T} = \tilde{p}_{1tt}(x,t), \tag{3.181}$$

$$\begin{aligned} f(x) &= -\frac{u_{2tt}(x,0)}{c_1(x)} \\ &= -\frac{1}{c_1(x)c_2(x)}\left(\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)\right). \end{aligned} \tag{3.182}$$

By (3.173), (3.177), (3.179), and (3.182)

$$\begin{aligned} -\tilde{c}(x) &= \frac{c_1(x)}{u_{2tt}(x,0)} v(x,0) \\ &= \frac{c_1(x)c_2(x)}{\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)}\left(v(x,t) - \int_0^t v_t(x,\tau)d\tau\right). \end{aligned} \tag{3.183}$$

Denote

$$P(x,t) = \frac{c_1(x)c_2(x)}{\Delta g(x) + \sum_{|\alpha|\leq 1} a_\alpha(x) D_x^\alpha g(x)} \partial_t^4 u_2(x,t). \tag{3.184}$$

Then, using (3.162)–(3.176), we obtain

$$\|P\|_{C(\overline{Q}_T)} \leq C. \tag{3.185}$$

Substituting (3.183) and (3.184) in (3.178), we replace (3.178)–(3.181) with

$$c_1(x)v_{tt} - \Delta v - \sum_{|\alpha|\leq 1} a_\alpha(x) D^\alpha v = P(x,t)\left(v(x,t) - \int_0^t v_t(x,\tau)d\tau\right), \tag{3.186}$$

$$v_t(x,0) = 0, \tag{3.187}$$

$$v|_{S_T} = \widetilde{p}_{0tt}(x,t), \quad \left.\frac{\partial v}{\partial n}\right|_{S_T} = \widetilde{p}_{1tt}(x,t). \tag{3.188}$$

Prior applying the Carleman estimate of Theorem 2.5.1 and Corollary 2.5.1, we now figure out some subdomains of the domain Q_T. By (3.166), there exists a number $\xi_1 \in (0, \xi_0)$ such that

$$\xi T^2 > (R + |x_0|)(R + |x_0| + 2\sqrt{h}), \quad \forall \xi \in (\xi_1, \xi_0). \tag{3.189}$$

We now show that

$$\Omega \subset \{|x - x_0| > 2\sqrt{h}\}. \tag{3.190}$$

Indeed, by the triangle inequality and (3.166)

$$|x - x_0| \geq |x_0| - |x| > |x_0| - R > 2\sqrt{h}, \quad \forall x \in \Omega = \{|x| < R\},$$

which proves (3.190).

Let the number $\xi \in (\xi_1, \xi_0)$. Introduce the function $\psi(x,t)$ depending on ξ as on a parameter,

$$\psi(x,t) = |x - x_0|^2 - \xi t^2.$$

The Carleman Weight Function (CWF) is

$$\varphi(x,t) = \exp[2\lambda\psi(x,t)],$$

where $\lambda > 1$ is a large parameter, which we will be specified later.

Consider the hypersurface ψ_h in Q_T defined as

$$\psi_h = \{(x,t) : |x - x_0|^2 - \xi t^2 = h, x \in \Omega, t > 0\}. \tag{3.191}$$

Then ψ_h is a level surface of the function $\psi(x,t)$. Also, consider the interior G_h of that hypersurface

$$G_h = \{(x,t) : |x - x_0|^2 - \xi t^2 > h, x \in \Omega, t > 0\}. \tag{3.192}$$

It follows from (3.189)–(3.192) that

$$G_{3h} \subset G_{3h} \subset G_{2h} \subset G_h, \tag{3.193}$$

$$\overline{G_h} \cap \{t = T\} = \varnothing, \tag{3.194}$$

$$\overline{G_{4h}} \cap \{t = 0\} = \Omega, \tag{3.195}$$

$$\partial G_h = \partial_1 G_h \cup \partial_2 G_h, \tag{3.196}$$

$$\partial_1 G_h = \{(x,t) : |x| = R, |x-x_0|^2 - \xi t^2 > h\}, \tag{3.197}$$

$$\partial_2 G_h = \{(x,t) : x \in \Omega, |x-x_0|^2 - \xi t^2 = h\}. \tag{3.198}$$

Furthermore, it follows from (3.189) and (3.195) that there exists a number $\sigma > 0$ such that

$$\{(x,t) \in \Omega \times (0,\sigma)\} = Q_\sigma \subset G_{4h}. \tag{3.199}$$

Now we are ready to apply the Carleman estimate of Theorem 2.5.1 under the condition of Corollary 2.5.1. Keeping in mind (3.193), consider the cut-off function $\chi(x,t)$,

$$\begin{cases} \chi \in C^2(\overline{G_h}), \\ \chi(x,t) = \begin{cases} 1, & (x,t) \in G_{3h}, \\ 0, & (x,t) \in G_h(0) \diagdown G_{2h}, \\ \text{between 0 and 1} & \text{for } (x,t) \in G_{2h} \diagdown G_{3h}. \end{cases} \end{cases} \tag{3.200}$$

In particular, by (3.193), (3.199), and (3.200)

$$\chi(x,t) = 1 \quad \text{for } (x,t) \in Q_\sigma \subset G_{4h}. \tag{3.201}$$

Introduce the function $w(x,t)$ as

$$w(x,t) = \chi(x,t) v(x,t). \tag{3.202}$$

We have

$$\begin{aligned} \chi(c_1 v_{tt} - \Delta v) = c_1 w_{tt} - \Delta w & \\ - 2c_1 \chi_t v_t - c_1 v \chi_{tt} + 2\nabla\lambda \nabla v + v \Delta\chi. & \end{aligned} \tag{3.203}$$

Multiply both sides of equation (3.186) by the function $\chi(x,t)$. Using (3.185), (3.201), (3.202), and (3.203), we obtain

$$\begin{aligned} |c_1 w_{tt} - \Delta w| \le C(|\nabla w| + |w|) & \\ + C \int_0^t |v_t(x,\tau)| d\tau + C(|\nabla v| + |v_t| + |v|), \quad \text{in } Q_T, & \end{aligned} \tag{3.204}$$

$$w_t(x,0) = 0, \tag{3.205}$$

$$w|_{S_T} = \chi \widetilde{p}_{0tt}(x,t), \quad \frac{\partial w}{\partial n}\Big|_{S_T} = \chi \widetilde{p}_{1tt}(x,t) + \frac{\partial \chi}{\partial n} \widetilde{p}_{0tt}(x,t). \tag{3.206}$$

Denote

$$m = m(x_0, R, M) = \max_{(x,t)\in \overline{G_h}} \max_{\xi \in [\xi_1, \xi_0]} \psi(x,t). \tag{3.207}$$

Square both sides of (3.204). Then multiply by the CWF $\varphi(x,t)$ with $\xi \in (\xi_1, \xi_0)$, integrate over the subdomain G_h of the time cylinder Q_T and use the Carleman estimate of Theorem 2.5.1 in combination with Corollary 2.5.1. In addition, use (3.193)–(3.198) and (3.205)–(3.207). We obtain

$$\begin{aligned}
&\lambda^3 e^{2\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}) \\
&\quad + \int_{G_h} \left(\int_0^t |v_t(x,\tau)| d\tau \right)^2 \varphi dxdt + \int_{G_h} (|\nabla v|^2 + v_t^2 + v^2)\varphi dxdt \\
&\quad + \int_{G_h} (|\nabla w|^2 + w^2)\varphi dxdt \\
&\geq C\lambda \int_{G_h} (|\nabla w|^2 + w_t^2)\varphi dxdt + C\lambda^3 \int_{G_h} w^2 \varphi dxdt, \quad \forall \lambda \geq \lambda_0,
\end{aligned} \tag{3.208}$$

where $\lambda_0 = \lambda_0(\Omega, \xi_0, x_0) > 1$ is the number of Theorem 2.5.1. Similarly, with (3.51)

$$\int_{G_h} \left(\int_0^t |v_t(x,\tau)| d\tau \right)^2 \varphi dxdt \leq \frac{C}{\lambda} \int_{G_h} v_t^2 \varphi dxdt. \tag{3.209}$$

Hence, choosing a sufficiently large $\lambda_1 = \lambda_1(M, d, T) \geq \lambda_0$ and using (3.208), we obtain for all $\lambda \geq \lambda_0$,

$$\begin{aligned}
&\lambda^3 e^{2\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}) \\
&\quad + \int_{G_h} (|\nabla v|^2 + v_t^2 + v^2)\varphi dxdt \\
&\geq C\lambda \int_{G_h} (|\nabla w|^2 + w_t^2)\varphi dxdt + C\lambda^3 \int_{G_h} w^2 \varphi dxdt, \quad \forall \lambda \geq \lambda_1.
\end{aligned} \tag{3.210}$$

Next since by (3.193) $G_{3h} \subset G_h$, then (3.193) and (3.210) imply that

$$\begin{aligned}
&\lambda^3 e^{2\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}) \\
&\quad + \int_{G_h \setminus G_{3h}} (|\nabla v|^2 + v_t^2 + v^2)\varphi dxdt + \int_{G_{3h}} (|\nabla v|^2 + v_t^2 + v^2)\varphi dxdt \\
&\geq C\lambda \int_{G_{3h}} (|\nabla w|^2 + w_t^2)\varphi dxdt + C\lambda^3 \int_{G_{3h}} w^2 \varphi dxdt, \quad \forall \lambda \geq \lambda_1.
\end{aligned} \tag{3.211}$$

Since by (3.192) $\varphi(x,t) \in (e^{2\lambda h}, e^{6\lambda h}]$ for $(x,t) \in G_h \setminus G_{3h}$ and since by (3.200) and (3.202) $w(x,t) = v(x,t)$ for $(x,t) \in G_{3h}$, then (3.211) implies for all sufficiently large values of

$\lambda \geq \lambda_2 = \lambda_2(M, d, T) \geq \lambda_1$,

$$\begin{aligned}
&\lambda^3 e^{2\lambda m}\left(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}\right) \\
&\quad + e^{6\lambda h} \int\limits_{G_h \setminus G_{3h}} \left(|\nabla v|^2 + v_t^2 + v^2\right)\varphi dxdt \\
&\geq C\lambda \int\limits_{G_{3h}} (|\nabla v|^2 + v_t^2)\varphi dxdt + C\lambda^3 \int\limits_{G_{3h}} v^2 \varphi dxdt \\
&\geq \lambda^3 e^{2\lambda m}\left(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}\right) \\
&\quad + e^{6\lambda h} \int\limits_{G_h \setminus G_{3h}} \left(|\nabla v|^2 + v_t^2 + v^2\right) dxdt \\
&\geq C\lambda \int\limits_{G_{4h}} (|\nabla v|^2 + v_t^2)\varphi dxdt + C\lambda^3 \int\limits_{G_{4h}} v^2 \varphi dxdt \\
&\geq Ce^{8\lambda h} \int\limits_{G_{4h}} \left(|\nabla v|^2 + v_t^2 + v^2\right) dxdt.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int\limits_{G_{4h}} \left(|\nabla v|^2 + v_t^2 + v^2\right) dxdt &\leq Ce^{-2\lambda h} \|v\|^2_{H^1(Q_T)} \\
&\quad + Ce^{3\lambda m}\left(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}\right), \\
&\quad \forall \lambda \geq \lambda_2 = \lambda_2(M, d, T).
\end{aligned}$$

Combining the latter estimate with (3.199), we obtain

$$\begin{aligned}
\int\limits_{Q_\sigma} \left(|\nabla v|^2 + v_t^2 + v^2\right) dxdt &\leq Ce^{-2\lambda h} \|v\|^2_{H^1(Q_T)} \\
&\quad + Ce^{3\lambda m}\left(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}\right), \\
&\quad \forall \lambda \geq \lambda_2 = \lambda_2(M, d, T).
\end{aligned} \tag{3.212}$$

By implying that there exists a number $t_0 \in [0, \sigma]$ such that

$$\int\limits_{\Omega} \left(|\nabla v|^2 + v_t^2 + v^2\right)(x, t_0) dx = \frac{1}{\sigma} \int\limits_{Q_\sigma} \left(|\nabla v|^2 + v_t^2 + v^2\right) dxdt.$$

Hence, using (3.212), we obtain

$$\begin{aligned}
\int\limits_{\Omega} \left(|\nabla v|^2 + v_t^2 + v^2\right)(x, t_0) dx &\leq Ce^{-2\lambda h} \|v\|^2_{H^1(Q_T)} \\
&\quad + Ce^{3\lambda m}\left(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}\right), \\
&\quad \forall \lambda \geq \lambda_2 = \lambda_2(M, d, T).
\end{aligned} \tag{3.213}$$

In addition, since $Q_{t_0} \subset Q_\sigma$, then (3.212) implies that for all $\lambda \geq \lambda_2$,

$$\int_{Q_{t_0}} \left(\int_0^t v_t(x,\tau) d\tau \right)^2 dxdt \leq C\|v\|^2_{H^1(Q_{t_0})} \tag{3.214}$$
$$\leq Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}).$$

And, more generally,

$$\|v\|^2_{H^1(Q_{t_0})} \leq Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}), \quad \forall \lambda \geq \lambda_2 = \lambda_2(M, d, T). \tag{3.215}$$

Consider now equation (3.186) with the Dirichlet boundary data of (3.188) in the time cylinder $Q_{t_0,T} = \Omega \times (t_0, T)$. We can consider the following initial boundary value problem with the initial data at $\{t = t_0\}$:

$$c_1(x)v_{tt} - \Delta v - \sum_{|\alpha|\leq 1} a_\alpha(x) D^\alpha v \tag{3.216}$$
$$= P(x,t)\left(v(x,t) - \int_{t_0}^t v_t(x,\tau) d\tau \right) + P(x,t) \int_0^{t_0} v_t(x,\tau)d\tau, \quad \text{in } Q_{t_0,T},$$

$$v(x, t_0) = f_0(x), \quad v_t(x, t_0) = f_1(x), \tag{3.217}$$

$$v|_{S_{t_0,T}} = \widetilde{p}_{0tt}(x,t), \quad \frac{\partial v}{\partial n}\Big|_{S_{t_0,T}} = \widetilde{p}_{1tt}(x,t) \tag{3.218}$$

where $S_{t_0,T} = \Omega \times (t_0, T)$. Hence, applying the standard method of energy estimates to the problems (3.216)–(3.218) and taking into account (3.214), we obtain

$$\|v\|^2_{H^1(Q_{t_0,T})} \leq C(\|f_0\|^2_{H^1(\Omega)} + \|f_1\|^2_{L_2(\Omega)}) \tag{3.219}$$
$$\leq Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}).$$

Summing up (3.215) and (3.219), we obtain for all $\lambda \geq \lambda_2$,

$$\|v\|^2_{H^1(Q_T)} \leq C(\|f_0\|^2_{H^1(\Omega)} + \|f_1\|^2_{L_2(\Omega)}) \tag{3.220}$$
$$\leq Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}).$$

We now estimate from the above the term $\|f_0\|^2_{H^1(\Omega)} + \|f_1\|^2_{L_2(\Omega)}$ in (3.220). It follows from (3.213) and (3.217) that

$$\|f_0\|^2_{H^1(\Omega)} + \|f_1\|^2_{L_2(\Omega)} \leq Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}(\|\widetilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\widetilde{p}_{1tt}\|^2_{L_2(S_T)}), \quad \forall \lambda \geq \lambda_2 = \lambda_2(M, R, d, T).$$

Substituting this in (3.220), we obtain

$$\|v\|^2_{H^1(Q_T)} \le Ce^{-2\lambda h}\|v\|^2_{H^1(Q_T)} + Ce^{3\lambda m}\big(\|\tilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\tilde{p}_{1tt}\|^2_{L_2(S_T)}\big), \quad \forall\lambda \ge \lambda_2 = \lambda_2(M,R,d,T). \tag{3.221}$$

Choose $\lambda_3 = \lambda_3(M,R,d,T) \ge \lambda_2$ so large that $Ce^{-2\lambda_3 h} \le 1/2$. And set in (3.221) $\lambda = \lambda_3$. We obtain

$$\|v\|^2_{H^1(Q_T)} \le C\big(\|\tilde{p}_{0tt}\|^2_{H^1(S_T)} + \|\tilde{p}_{1tt}\|^2_{L_2(S_T)}\big).$$

Hence, by the trace theorem

$$\big\|v(x,0)\big\|_{L_2(\Omega)} \le C\big(\|\tilde{p}_{0tt}\|_{H^1(S_T)} + \|\tilde{p}_{1tt}\|_{L_2(S_T)}\big). \tag{3.222}$$

Finally, the target estimate (3.172) of this theorem follows from estimate (3.222) combined with (3.176), (3.179), and (3.182). □

4 The quasi-reversibility numerical method for ill-posed Cauchy problems for linear PDEs

4.1 Introduction

In this chapter, we republish some parts of [9, 132, 134, 135, 165], although with some deviations. Permissions for republications are obtained from the publishers.

We describe here those numerical methods for ill-posed Cauchy problems for linear PDEs, which are based on the Quasi Reversibility Method (QRM). Indeed, while Hölder stability results were obtained for these problems in Chapter 2, the next natural question is: *How to solve these problems numerically?* It is well known that these problems are simpler than coefficient inverse problems. In fact, there are many publications devoted to various numerical methods for ill-posed Cauchy problems for PDE. We now cite some of them and there are many more [8, 72, 74, 75, 90–92, 107, 169]. However, each of these works uses its own idea for handling the problem under study.

The QRM was first proposed in the pioneering book of French mathematicians Lattes and Lions in 1969 [182]. The first numerical results on this method were also presented in [182]. However, Lattes and Lions have in their computations the FDM for the strong formulations with equations of the fourth order with the operator A^*A. Carleman estimates were not used and convergence rates of the regularized solutions were not established in [182].

We formulate the QRM as the problem of the minimization of a Tikhonov-like functional with an unbounded partial\differential operator in it. The first step of the convergence analysis is to prove existence and uniqueness of the minimizer of this functional. In fact, this is rather easy to do if using the classical Riesz theorem. What is not trivial is the second step: to establish the convergence rate of minimizers to the exact solution as the noise level in the data tends to zero. Recall that in the theory of ill-posed problems, such minimizers are called *regularized solutions*; see, for example, [22, 244]. The latter proof is obtained via an application of a Carleman estimate. We demonstrate in this chapter that the QRM can be applied to a wide class of ill-posed Cauchy problems for linear PDEs, that is, the QRM is a universal regularization method. It works for those PDEs, for which Carleman estimates hold.

The idea of applications of Carleman estimates for establishing the convergence rate of regularized solutions of the QRM as the level of the noise in the data tends to zero was originated in 1991 in works [153, 161]. Next, this idea was explored in a number of publications; see works [57, 64, 134, 147, 148, 156, 158, 161, 166, 236]. Furthermore, it was discovered in [156] that the QRM can be applied to the inversion of the Radon transform with incomplete data. Indeed, it was shown in, for example, the book of Hasanoglu and Romanov [87] that the inversion of the Radon transform is equivalent to the solution of a linear inverse source problem for a transport equation. Thus, a new numerical method, which is based on the QRM, was developed in [156]

https://doi.org/10.1515/9783110745481-004

for the latter inverse problem. Furthermore, the data in [156] are incomplete. Next, two modifications of the method of [156] were applied to various inverse source problems with incomplete data [148, 236], including even the linearized travel time tomography problem [148]. The latter problem is considered in detail in Chapter 12.

We also draw attention of the reader to our work [152] in which the famous linear integral equation derived by Mikhail Mikhailovich Lavrent'ev in 1964 [183] is applied to solve a highly nonlinear coefficient inverse problem for the equation $c(x)u_{tt} = \Delta u$. Surprisingly, it works well. The Lavrent'ev equation is solved in [152] via applying the QRM to a system of elliptic partial differential equations. A member of the Russian Academy of Science, M. M. Lavrent'ev (1932–2010), was the thesis advisor of M. V. Klibanov in 1973–1977. Lavrent'ev was one of founders of the theory of inverse and ill-posed problems. That integral equation is actually the formula (7.18) of the book [184].

We point out that the conventional regularization theory for linear ill-posed problems uses only bounded linear operators for the Tikhonov functional; see, for example, [184]. Unlike this, the QRM uses unbounded PDE operators for this purpose. Actually, this is the underlying reason of why do we need Carleman estimates to estimate convergence rates of regularized solutions.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $A(x, D)$ be a linear Partial Differential Operator (PDO) of the second order acting in Ω. We assume that this operator admits a Carleman estimate, that is, conditions formulated in the beginning of Section 2.1 of Chapter 2 are in place. Therefore, results of this chapter are very general ones. Consider the Partial Differential Equation (PDE) $A(x, D)u = f$, $x \in \Omega$ and an ill-posed Cauchy problem for it. First, we present below the QRM below for a general operator $A(x, D)$ and establish the convergence rate of the QRM as the noise level in the data tends to zero. More precisely, we establish the convergence rate of the regularized solutions. Next, we specify our method for three main classes of ill-posed Cauchy problems: for elliptic, hyperbolic, and parabolic PDEs. Such rates were also established for the initial boundary value problem for the parabolic PDE with the reversed time [166].

The material of this chapter is a purely analytical one; also, see Chapter 12 for the QRM for a linearized coefficient inverse problem with a numerical result. Numerical studies of a variety of versions of the QRM can be found in papers of Bourgeois with his coauthors Dardé and Ponomarev [40–43, 45–47], Dardé with his coauthors Hannukainen and Hyvönen [68, 69] and Klibanov with his coauthors [57, 64, 147, 148, 156, 158, 161, 236].

Remark 4.1.1. A natural question is on how to minimize functionals of QRM numerically. The answer is as follows. First, the partial differential operator of this functional should be written in finite differences. Next, the minimization should be done with respect to the values of the unknown function at grid points. There are two ways of calculating the gradient of the functional in this case:

1. First, suppose that we have numbered grid points are $\{x_j\}_{j=1}^n$. Let $v_j = v(x_j)$ be the values of the target function v at grid points. Let $J(v) = J(v_1, \dots, v_n)$ be the func-

tional we minimizer in the QRM. Then the vector of the gradient $\nabla J(v)$ at the point v is formed using

$$\frac{\partial J(v)}{\partial v_j} \approx \frac{J(v_1,\dots,v_{j-1},v_j+h,v_{j+1},\dots,v_n) - J(v_1,\dots,v_{j-1},v_j,v_{j+1},\dots,v_n)}{h},$$

where $h > 0$ is an appropriate sufficiently small number.

2. Second, one could calculate the gradient via explicit, although long, formulas using the Kronecker symbol; see some details in section 6 of the work [147]. In [57, 64, 147, 148, 156, 158, 161, 236], the second approach was used. The same is true for all works on the numerical issues of the convexification.

4.2 The Quasi-Reversibility Method (QRM)

4.2.1 The Carleman estimate

For reader's convenience, we now recall a general Carleman estimate of Section 2.1 of Chapter 2. Consider the function $\psi \in C^2(\overline{\Omega})$ such that $|\nabla\psi| \neq 0$ in $\overline{\Omega}$. For a number $h \geq 0$, denote

$$\psi_h = \{x \in \overline{\Omega} : \psi(x) = h\}, \quad \Omega_h = \{x \in \Omega : \psi(x) > h\}. \tag{4.1}$$

Let the domain $\Omega_h \neq \varnothing$. Consider a part Γ_h of $\partial\Omega$ defined as

$$\Gamma_h = \{x \in \partial\Omega : \psi(x) > h\}. \tag{4.2}$$

Then the boundary $\partial\Omega_h$ of Ω_h is

$$\partial\Omega_h = \partial_1\Omega_h \cup \partial_2\Omega_h, \tag{4.3}$$

$$\partial_1\Omega_h = \psi_h, \quad \partial_2\Omega_h = \Gamma_h. \tag{4.4}$$

Let $\lambda > 1$ be a sufficiently large parameter, which we will specify later. Consider the function $\varphi(x)$,

$$\varphi(x) = \exp(\lambda\psi(x)). \tag{4.5}$$

It follows from (4.1) and (4.5) that

$$\min_{\overline{\Omega}_h} \varphi(x) = \varphi(x)|_{\psi_h} \equiv e^{\lambda h}. \tag{4.6}$$

Consider a linear partial differential operator $A(x,D)$ of the second order with real valued coefficients in Ω,

$$A(x,D)u = \sum_{|\alpha|\leq 2} a_\alpha(x) D^\alpha u, \quad x \in \Omega. \tag{4.7}$$

Its principal part $A_0(x,D)$ is

$$A_0(x,D)u = \sum_{|\alpha|=2} a_\alpha(x) D^\alpha u. \tag{4.8}$$

We assume that coefficients $a_\alpha(x)$ of the operator $A(x,D)$ satisfy the following conditions:

$$a_\alpha \in C^1(\overline{\Omega}) \quad \text{for } |\alpha| = 2, \tag{4.9}$$

$$a_\alpha \in C(\overline{\Omega}) \quad \text{for } |\alpha| = 0, 1. \tag{4.10}$$

We now recall the Definition 2.1.1 of the pointwise Carleman estimate for the operator $A_0(x,D)$.

Definition 4.2.1. For the coefficients $a_\alpha(x)$ of the operator $A_0(x,D)$ in (4.8), let $K = \max_{|\alpha|=2} \|a_\alpha\|_{C^1(\overline{\Omega})}$. The operator $A_0(x,D)$ admits pointwise Carleman estimate in the domain Ω_h if there exist constants $\lambda_0(\Omega_h, K) > 1$, $C(\Omega_h, K) > 0$ depending only on the domain Ω_h and the number K, such that the following a priori estimate holds:

$$(A_0 u)^2 \varphi^2(x) \geq C\lambda(\nabla u)^2 \varphi^2(x) + C\lambda^3 u^2 \varphi_\lambda^2(x) + \operatorname{div} U, \tag{4.11}$$

$$\forall \lambda \geq \lambda_0, \forall u \in C^2(\overline{\Omega}_h), \forall x \in \Omega_h, \tag{4.12}$$

where following estimate holds for the divergence term $\operatorname{div} U$:

$$|U| \leq C\lambda^3 [(\nabla u)^2 + u^2] \varphi^2(x). \tag{4.13}$$

In this case, the function $\varphi(x)$ is called the Carleman Weight Function (CWF) for the operator $A(x,D)$.

We use in the last sentence $A(x,D)$ instead of $A_0(x,D)$ because by Lemma 2.1.1 the Carleman estimate for the operator $A(x,D)$ in (4.7) is independent on low order terms of this operator.

4.2.2 The Quasi-Reversibility Method (QRM)

Let the function $f \in L_2(\Omega_h)$. Assume that conditions of Section 4.2.1 hold.

Cauchy problem. Suppose that the function $u \in H^2(\Omega_h)$ solves the equation

$$A(x,D)u = f \quad \text{in } \Omega_h \tag{4.14}$$

and satisfies the following Cauchy boundary conditions at Γ_h :

$$u|_{\Gamma_h} = p_0(x), \quad \partial_n u|_{\Gamma_h} = p_1(x). \tag{4.15}$$

Find the function $u(x)$ for $x \in \Omega_h$.

We assume that there exists a function $F \in H^2(\Omega_h)$ satisfying Cauchy boundary conditions (4.15), that is,

$$F|_{\Gamma_h} = p_0(x), \quad \partial_n F|_{\Gamma_h} = p_1(x). \tag{4.16}$$

In fact, it is unlikely that the existence theorem holds for problem (4.14), (4.15). Thus, we find an approximate solution of this problem in the least squares sense, that is, as a minimizer of the Tikhonov functional with the regularization parameter $\alpha \in (0,1)$. This functional is

$$J_\alpha(u) = \|A(x,D)u - f\|^2_{L_2(\Omega_h)} + \alpha\|u - F\|^2_{H^2(\Omega_h)}. \tag{4.17}$$

Minimization problem of the QRM. Find a minimizer $u_\alpha \in H^2(\Omega_h)$ of the functional $J_\alpha(u)$ in (4.17) satisfying Cauchy boundary conditions (4.15).

For brevity, $A := A(x,D)$ below. As stated in Section 4.1, in the regularization theory, such a minimizer is called the regularized solution. As the first step, we prove the existence and uniqueness of the regularized solution. This is rather easy to do if applying Riesz theorem.

Theorem 4.2.1 (existence and uniqueness of the minimizer). *For every value of the regularization parameter $\alpha \in (0,1)$, there exists unique minimizer $u_\alpha \in H^2(\Omega_h)$ of the functional $J_\alpha(u)$. Furthermore, there exists a constant $C = C(\Omega_h, A(x,D)) > 0$ depending only on listed parameters such that the following estimate holds:*

$$\|u_\alpha\|_{H^2(\Omega_h)} \le \frac{C}{\sqrt{\alpha}}\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right). \tag{4.18}$$

Proof. In this proof, $C > 0$ denotes different constants depending only on above listed parameters. Introduce the subspace $H_0^2(\Omega_h)$ of the space $H^2(\Omega_h)$ as

$$H_0^2(\Omega_h) = \{w \in H^2(\Omega_h) : w|_{\Gamma_h} = \partial_n w|_{\Gamma_h} = 0\}.$$

Let $w = u - F$. By (4.15) and (4.16), the function $w \in H_0^2(\Omega_h)$. The functional $J_\alpha(u)$ becomes

$$I_\alpha(w) = \|Aw + (AF - f)\|^2_{L_2(\Omega_h)} + \alpha\|w\|^2_{H^2(\Omega_h)}, \quad w \in H_0^2(\Omega_h). \tag{4.19}$$

The minimization of functional (4.19) is equivalent with the minimization of functional (4.17), subject to boundary conditions (4.15). Let w_α be a minimizer of the functional $I_\alpha(w)$. By the variational principle, the following identity holds:

$$(Aw_\alpha, Av) + \alpha[w_\alpha, v] = (Av, f - AF), \quad \forall v \in H_0^2(\Omega_h), \tag{4.20}$$

where $(\, , \,)$ and $[\, , \,]$ are scalar products in $L_2(\Omega_h)$ and $H^2(\Omega_h)$, respectively. Denote

$$\{p, v\}_\alpha = (Ap, Av) + \alpha[p, v], \quad \forall p, v \in H_0^2(\Omega_h).$$

Hence, $\{p, v\}_\alpha$ defines a new scalar product in the Hilbert space $H_0^2(\Omega_h)$ and the corresponding norm $\{v\}_\alpha = \sqrt{\{v, v\}_\alpha}$ satisfies

$$\sqrt{\alpha}\|v\|_{H^2(\Omega_h)} \leq \{v\}_\alpha \leq C\|v\|_{H^2(\Omega_h)}, \quad \forall v \in H_0^2(\Omega_h). \tag{4.21}$$

Thus, the new norm $\{v\}_\alpha$ is equivalent with the standard norm $\|v\|_{H^2(\Omega_h)}$. Hence, we rewrite (4.20) in the equivalent form,

$$\{w_\alpha, v\}_\alpha = (Av, f - AF), \quad \forall v \in H_0^2(\Omega_h). \tag{4.22}$$

By (4.21),

$$\left|(Av, f - AF)\right| \leq C\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right)\{v\}_\alpha, \quad \forall v \in H_0^2(\Omega_h). \tag{4.23}$$

This means that the right-hand side of (4.22) can be considered as a bounded linear functional $\Phi(v) : H_0^2(\Omega_h) \to \mathbb{R}$. Hence, Riesz theorem implies the existence and uniqueness of a function $q_\alpha \in H_0^2(\Omega_h)$ depending on the function $f - AF$, $q_\alpha = q_\alpha(f - AF)$ such that

$$(Av, f - AF) = \{q_\alpha, v\}_\alpha, \quad \forall v \in H_0^2(\Omega_h). \tag{4.24}$$

It follows from (4.22) and (4.24) that

$$\{w_\alpha, v\}_\alpha = \{q_\alpha, v\}_\alpha, \quad \forall v \in H_0^2(\Omega_h).$$

Furthermore, since by Riesz theorem $\{q_\alpha\}_\alpha = \|\Phi\|$, then (4.23) implies that

$$\{q_\alpha\}_\alpha \leq C\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right).$$

Hence, the minimizer w_α exists, is unique, $w_\alpha = q_\alpha$ and

$$\{w_\alpha\}_\alpha \leq C\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right). \tag{4.25}$$

It follows from (4.21) and (4.25) that

$$\|w_\alpha\|_{H^2(\Omega_h)} \leq \frac{C}{\sqrt{\alpha}}\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right). \tag{4.26}$$

Finally, setting $u_\alpha = w_\alpha + F$, recalling that $\alpha \in (0, 1)$ and using (4.26), we obtain (4.18). □

A Carleman estimate was not used in the proof of Theorem 4.2.1. In fact, the proof of Theorem 4.2.1 is rather simple since it is based on the Riesz theorem. We now prove a more sophisticated result. More precisely, we establish the convergence rate of regularized solutions to the exact solution when the level of noise in the data tends to zero. By one of the fundamental concepts of the theory of ill-posed problems [22, 244], we assume that there exists an exact solution $u^* \in H^2(\Omega_h)$ of Cauchy problem (4.14), (4.15) with the exact, that is, noiseless data

$$f^* \in L_2(\Omega_h), \quad u^*|_{\Gamma_h} = p_0^* \in H^1(\Gamma_h), \quad \partial_n u^*|_{\Gamma_h} = p_1^* \in L_2(\Gamma_h). \tag{4.27}$$

In other words, the function u^* is an ideal solution of problem (4.14), (4.15). By Theorem 2.2.1, the exact solution $u^* \in H^2(\Omega_h)$ is unique.

The existence of the exact solution $u^* \in H^2(\Omega_h)$ implies the existence of a function $F^* \in H^2(\Omega_h)$ satisfying boundary conditions (4.27), that is,

$$F^*|_{\Gamma_h} = p_0^*, \quad \partial_n F^*|_{\Gamma_h} = p_1^*. \tag{4.28}$$

We now describe an example of the function F^*. Consider the function $\chi \in C^2(\overline{\Omega}_h)$ such that $\chi(x) = 1$ in a small neighborhood $Nb_\rho(\Gamma_h) = \{x \in \Omega_c : \operatorname{dist}(x, \Gamma_h) < \rho\}$ and $\chi(x) = 0$ for $x \in \Omega_h \diagdown Nb_{2\rho}(\Gamma_h)$, where $\rho > 0$ is a sufficiently small number. We set $F^*(x) = \chi(x)u^*(x)$.

Let $\delta > 0$ be a sufficiently small number, which we call the level of the noise in the data. Suppose that there exists the function $F \in H^2(\Omega_h)$ satisfying boundary conditions (4.16). We assume that

$$\|f^* - f\|_{L_2(\Omega_h)}, \|F^* - F\|_{H^2(\Omega_h)} \le \delta. \tag{4.29}$$

Theorem 4.2.2 (convergence rate of regularized solutions). *Assume that the Carleman estimate of Definition* 4.2.1 *is valid. Also, suppose that there exist functions* $F \in H^2(\Omega_h)$ *and* $F^* \in H^2(\Omega_h)$ *satisfying conditions* (4.16) *and* (4.28)*, respectively. In addition, assume that conditions* (4.29) *hold. Let the regularization parameter* $\alpha = \alpha(\delta) = \delta^{2\beta}$, *where* $\beta = \text{const.} \in (0, 1]$. *Suppose that there exists a sufficiently small number* $\varepsilon > 0$ *such that* $\Omega_{h+3\varepsilon} \neq \varnothing$ *and* $\Gamma_{h+3\varepsilon} \neq \varnothing$. *Let* $m = \max_{\overline{\Omega}_h} \psi(x)$ *and* $\gamma = 2\varepsilon/(3m + 2\varepsilon)$. *Then there exists a sufficiently small number* $\delta_0 = \delta_0(\varepsilon, m, A, \Omega_h) \in (0, 1)$ *and a constant* $C = C(\varepsilon, m, A, \Omega_h) > 0$ *depending only on listed parameters such that if* $\delta \in (0, \delta_0^{1/\beta})$, *then the following convergence rate of regularized solutions holds:*

$$\|u_{\alpha(\delta)} - u^*\|_{H^1(\Omega_{h+3\varepsilon})} \le C(1 + \|u^*\|_{H^2(\Omega_h)})\delta^{\beta\gamma}, \quad \forall \delta \in (0, \delta_0), \tag{4.30}$$

where $u_\alpha \in H^2(\Omega_h)$ *is the minimizer of the functional* $J_\alpha(u)$ *in* (4.17) *satisfying Cauchy boundary conditions* (4.15). *The existence and uniqueness of this minimizer is guaranteed by Theorem* 4.2.1.

Proof. In this proof, $C = C(\varepsilon, m, A, \Omega_h) > 0$ denotes different positive constants depending only on listed parameters. Let $w^* = u^* - F^*$. Then $w^* \in H_0^2(\Omega_h)$ and $Aw^* = f^* - AF^*$. Hence,

$$(Aw^*, Av) + \alpha[w^*, v] = (Av, f^* - AF^*) + \alpha[w^*, v], \quad \forall v \in H_0^2(\Omega_h). \tag{4.31}$$

Subtract identity (4.20) from (4.31). Denote $\widetilde{w}_\alpha = w^* - w_\alpha$, $\widetilde{f} = f^* - f$, $\widetilde{F} = F^* - F$. We obtain

$$(A\widetilde{w}_\alpha, Av) + \alpha[\widetilde{w}_\alpha, v] = (Av, \widetilde{f} - A\widetilde{F}) + \alpha[w^*, v], \quad \forall v \in H_0^2(\Omega_h). \tag{4.32}$$

Set in (4.32) here $v := \widetilde{w}_\alpha$. We obtain

$$\|A\widetilde{w}_\alpha\|^2_{L_2(\Omega_h)} + \alpha\|\widetilde{w}_\alpha\|^2_{H^2(\Omega_h)} = (A\widetilde{w}_\alpha, \widetilde{f} - A\widetilde{F}) + \alpha[w^*, \widetilde{w}_\alpha]. \tag{4.33}$$

Apply the Cauchy–Schwarz inequality to (4.33). We obtain

$$\begin{aligned}&\|A\widetilde{w}_\alpha\|^2_{L_2(\Omega_h)} + \alpha\|\widetilde{w}_\alpha\|^2_{H^2(\Omega_h)}\\ &\quad\le \frac{1}{2}\|A\widetilde{w}_\alpha\|^2_{L_2(\Omega_h)} + \frac{1}{2}\|\widetilde{f} - A\widetilde{F}\|^2_{L_2(\Omega_h)} + \frac{\alpha}{2}\|w^*\|^2_{H^2(\Omega_h)} + \frac{\alpha}{2}\|\widetilde{w}_\alpha\|^2_{H^2(\Omega_h)}.\end{aligned}$$

Hence, by (4.29)

$$\|A\widetilde{w}_\alpha\|^2_{L_2(\Omega_h)} + \alpha\|\widetilde{w}_\alpha\|^2_{H^2(\Omega_h)} \le C\delta^2 + \alpha\|w^*\|^2_{H^2(\Omega_h)}. \tag{4.34}$$

Since $\alpha = \alpha(\delta) = \delta^{2\beta}$, where $\beta \in (0,1]$, then $\delta^2 \le \alpha$. Hence, it follows from (4.34) that

$$\|A\widetilde{w}_\alpha\|^2_{L_2(\Omega_h)} \le C(1 + \|w^*\|_{H^2(\Omega_h)})\delta^{2\beta}. \tag{4.35}$$

Applying Theorem 2.2.3 to (4.35), we obtain

$$\|\widetilde{w}_\alpha\|_{H^1(\Omega_{h+3\varepsilon})} \le C(1 + \|v^*\|_{H^2(\Omega_h)})\delta^{\beta\gamma}, \quad \forall \delta \in (0, \delta_0). \tag{4.36}$$

Recall that $\widetilde{w}_\alpha = (u_\alpha - u^*) + (F^* - F)$. Also, by (4.29) $\|F^* - F\|_{H^1(\Omega_{h+3\varepsilon})} \le \delta$. Hence, using the triangle inequality, we obtain

$$\begin{aligned}\|\widetilde{w}_\alpha\|_{H^1(\Omega_{h+3\varepsilon})} &\ge \|u_\alpha - u^*\|_{H^1(\Omega_{h+3\varepsilon})} - \|F^* - F\|_{H^1(\Omega_{h+3\varepsilon})}\\ &\ge \|u_\alpha - u^*\|_{H^1(\Omega_{h+3\varepsilon})} - \delta.\end{aligned} \tag{4.37}$$

Since numbers $\beta, \gamma, \delta \in (0,1)$, then $\delta^{\beta\gamma} > \delta$. Thus, (4.36) and (4.37) imply the target estimate (4.30) of this theorem. □

4.3 Elliptic equation

Let L be an elliptic operator of the second order in the bounded domain $\Omega \subset \mathbb{R}^n$ with its principal part L_0,

$$Lu = \sum_{i,j=1}^{n} a_{i,j}(x)u_{x_i x_j} + \sum_{j=1}^{n} b_j(x)u_{x_j} + c(x)u, \quad x \in \Omega, \tag{4.38}$$

$$L_0 u = \sum_{i,j=1}^{n} a_{i,j}(x) u_{x_i x_j}, \quad x \in \Omega, \tag{4.39}$$

where $a_{i,j}(x) = a_{j,i}(x)$, $\forall i,j$. Just as in (4.9) and (4.10), we assume that

$$a_{i,j} \in C^1(\overline{\Omega}); \quad b_j, c \in C(\overline{\Omega}). \tag{4.40}$$

To ensure the ellipticity of the operator L_0, we assume that there exist two constants $\mu_1, \mu_2 > 0$, $\mu_1 \le \mu_2$ such that

$$\mu_1 |\xi|^2 \le \sum_{i,j=1}^{n} a_{i,j}(x)\xi_i \xi_j \le \mu_2 |\xi|^2, \quad \forall x \in \overline{\Omega}, \forall \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n. \tag{4.41}$$

Let $Y \subset \partial\Omega$ be the part of the boundary $\partial\Omega$, where the Cauchy data are given. Let $Y' \subset \mathbb{R}^{n-1}$ be a bounded domain. Denote $\overline{x} = (x_2, \dots, x_n)$. We assume that

$$Y = \{x \in \mathbb{R}^n : x_1 = z(\overline{x}), \overline{x} \in Y'\},$$

where the function $z \in C^2(\overline{Y}')$. Change variables as $x = (x_1, \overline{x}) \Leftrightarrow (x_1', \overline{x})$, where $x_1' = x_1 - z(\overline{x})$. For brevity, we do not change notations. We obtain that in new variables the domain Y becomes

$$Y = \{x \in \mathbb{R}^n : x_1 = 0, \overline{x} \in Y'\}.$$

The operator L still remains elliptic after this change of variables. Let $X > 0$ be a number. Without loss of generality, we assume that

$$\Omega \subset \{x_1 > 0\}, \quad Y = \{x \in \mathbb{R}^n : x_1 = 0, |\overline{x}| < X\} \subset \partial\Omega. \tag{4.42}$$

Cauchy problem for the elliptic equation. Suppose that conditions (4.38)–(4.42) hold. Find such a function $u \in H^2(\Omega)$ that satisfying the following conditions:

$$Lu = f \quad \text{in } \Omega, \tag{4.43}$$

$$u|_{x \in Y} = p_0(\overline{x}), \quad u_{x_1}|_{x \in Y} = p_1(\overline{x}), \tag{4.44}$$

where functions g_0, g_1 are the Cauchy data and the function $f \in L_2(\Omega)$.

These are incomplete Cauchy data, since they are given only at a part of the boundary of the domain Ω rather than at the whole boundary. We now remind the Carleman estimate of Theorem 2.4.1. Let $\lambda > 1$ and $\nu > 1$ be two large parameters, which we will specify later. Consider two arbitrary numbers $\alpha, h = \text{const.} \in (0,1)$, where $\alpha < h$. Denote

$$\psi(x) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha, \quad \varphi(x) = \exp(\lambda \psi^{-\nu}). \tag{4.45}$$

The domain Ω_h now is similar with the one in (2.65),

$$\Omega_h = \left\{x : x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha < h\right\}. \tag{4.46}$$

Following (4.46), we define Γ_h as

$$\Gamma_h = \{x : x_1 = 0, |\overline{x}| < (h - \alpha)^{1/2}\sqrt{2}X\}. \tag{4.47}$$

Let ψ_h be the part of the level hypersurface of the level h of the function ψ, which is contained in the half-space $\{x_1 > 0\}$,

$$\psi_h = \left\{x : x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha = h\right\}.$$

Then

$$\partial\Omega_h = \Gamma_h \cup \psi_h.$$

We assume that $\overline{\Omega}_h \subseteq \Omega$. For a sufficiently small number $\varepsilon > 0$ and a number $\nu > 0$ define the subdomain $\Omega_{h^{-\nu}+3\varepsilon} \subset \Omega_h$ as

$$\Omega_{h^{-\nu}+3\varepsilon} = \left\{x : x_1 > 0, \left(x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha\right)^{-\nu} > h^{-\nu} + 3\varepsilon\right\} \tag{4.48}$$

and assume that $\Omega_{h^{-\nu}+3\varepsilon} \neq \varnothing$, which is true if ε is sufficiently small.

The Carleman estimate of Theorem 2.4.1 allows us to construct the QRM for problem (4.43), (4.44). First, we construct an example of the function $F \in H^2(\Omega_h)$ satisfying boundary conditions (4.15). Let functions

$$p_0, p_1 \in H^2(\Gamma_h). \tag{4.49}$$

Let the number $\sigma > 0$ be sufficiently small. Consider the function $\chi(x_1)$ satisfying

$$\chi \in C^\infty[0, h - \alpha], \quad \chi(x_1) = \begin{cases} 1, & x_1 \in (0, \sigma), \\ 0, & x_1 \in (2\sigma, h - \alpha). \end{cases} \tag{4.50}$$

Set

$$F(x) = \chi(x_1)p_0(\overline{x}) - \chi(x_1)x_1p_1(\overline{x}). \tag{4.51}$$

It follows from (4.49)–(4.51) that the function $F \in H^2(\Omega_h)$ and also that F satisfies boundary conditions (4.15). Assume now that there exists the exact solution $u^* \in H^2(\Omega)$ of problem (4.43), (4.44) with the exact Cauchy data $p_0^*, p_1^* \in H^2(\Gamma_h)$ and the exact function $f^* \in L_2(\Omega)$. Then, replacing p_0, p_1 with p_0^*, p_1^*, we construct the function $F^* \in H^2(\Omega_h)$ as in (4.51). We assume that

$$\|F - F^*\|_{H^2(\Omega_h)} < \delta, \tag{4.52}$$

where $\delta \in (0, 1)$ is the level of noise in the data.

The Tikhonov functional for problem (4.43), (4.44) is

$$J_\beta(u) = \|Lu - f\|^2_{L_2(\Omega_h)} + \beta\|u - F\|^2_{H^2(\Omega_h)}. \tag{4.53}$$

Minimization problem for functional (4.53)**.** Find a minimizer $u_\beta \in H^2(\Omega_h)$ of the functional $J_\alpha(u)$ in (4.53) satisfying boundary conditions (4.44).

Theorem 4.3.1 follows immediately from Theorems 2.4.1, 4.2.1, and 4.2.2.

Theorem 4.3.1. *For $\beta \in (0,1)$ there exists unique minimizer $u_\beta \in H^2(\Omega_h)$ of functional* (4.52) *satisfying boundary conditions* (4.44) *and the following estimate holds with the constant $C > 0$ independent on β*:

$$\|u_\beta\|_{H^2(\Omega_h)} \le \frac{C}{\sqrt{\beta}}\left(\|F\|_{H^2(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\right).$$

Furthermore, suppose that (4.52)*. Choose $\beta = \beta(\delta) = \delta^{2\rho}$, where the number $\rho =$ const. $\in (0,1]$. Then there exists a number $\gamma = \gamma(\varepsilon) \in (0,1)$ and a sufficiently small number $\delta_0 \in (0,1)$ such that if $\delta \in (0, \delta_0^{1/\rho})$, then the following convergence rate of regularized solutions $u_{\alpha(\delta)}$ holds:*

$$\left\|u_{\beta(\delta)} - u^*\right\|_{H^1(\Omega_{h^{-\nu_0}+3\varepsilon})} \le C\left(1 + \left\|u^*\right\|_{H^2(\Omega_h)}\right)\delta^{\beta\gamma}, \quad \forall \delta \in (0, \delta_0),$$

where the subdomain $\Omega_{h^{-\nu_0}+3\varepsilon} \ne \varnothing$ of the domain Ω_h is defined in (4.48) *and the number ν_0 is defined in Theorem* 2.4.1.

4.4 Parabolic equation with the lateral Cauchy data

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a piecewise smooth boundary $\partial\Omega$. As in (4.42) of the previous section, we assume that

$$\Omega \subset \{x_1 > 0\}, \quad Y = \left\{x \in \mathbb{R}^n : x_1 = 0, |\overline{x}| < \sqrt{2}X\right\} \subset \partial\Omega, \tag{4.54}$$

where $\overline{x} = (x_2, \dots, x_n)$. For $T > 0$, denote

$$\Omega_T^{\pm} = \Omega \times (-T, T).$$

Let numbers α, X, ν, λ be the same as ones in Section 4.3. Consider functions $\psi(x,t)$ and $\varphi(x,t)$ which are similar to those defined in (2.28) and (2.30), respectively,

$$\psi(x,t) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha, \quad \varphi(x,t) = \exp(\lambda\psi^{-\nu}).$$

For a number $h \in (0,1)$, we define the domain Ω_h and a part of its boundary Γ_h similarly with (4.46) and (4.47), respectively,

$$\Omega_{Th}^{\pm} = \left\{(x,t) : x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h\right\},$$

$$\Gamma_{Th}^{\pm} = \left\{ (x,t) : x_1 = 0, \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha < h \right\} \subset \partial\Omega_h.$$

Let

$$\psi_h = \left\{ x : x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha = h \right\}.$$

Hence,

$$\partial\Omega_{Th}^{\pm} = \Gamma_{Th}^{\pm} \cup \psi_h.$$

We again assume that $\overline{\Omega_{Th}^{\pm}} \subseteq \Omega$. Choose a sufficiently small number $\varepsilon > 0$. Define the subdomain $\Omega_{Th}^{\pm}(h^{-\nu} + 3\varepsilon)$ as

$$\Omega_{Th}^{\pm}(h^{-\nu} + 3\varepsilon) = \left\{ x : x_1 > 0, \left(x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha \right)^{-\nu} > h^{-\nu} + 3\varepsilon \right\}. \tag{4.55}$$

We assume that ε is so small that $\Omega_{Th}^{\pm}(h^{-\nu} + 3\varepsilon) \neq \varnothing$.

Consider now the parabolic operator L in $\Omega_T^{\pm}$,

$$Lu = u_t - \sum_{i,j=1}^{n} a^{ij}(x,t) u_{x_i x_j} - \sum_{j=1}^{n} b^j(x,t) u_{x_j} - c(x,t)u.$$

The ellipticity of the operator $\partial_t - L$ means that $a^{ij}(x,t) = a^{ji}(x,t)$ and also that there exist two numbers $\mu_1, \mu_2 > 0$ such that

$$\mu_1 |\xi|^2 \leq \sum_{i,j=1}^{n} a^{ij}(x,t)\xi_i\xi_j \leq \mu_2|\xi|^2, \quad \forall (x,t) \in \overline{\Omega}_h, \forall \xi \in \mathbb{R}^n.$$

We also assume that

$$a^{ij} \in C^1(\overline{\Omega_{Th}^{\pm}}), \quad K_0 = \max_{i,j} \|a_{ij}\|_{C^1(\overline{\Omega_{Th}^{\pm}})},$$
$$b^j, c \in C(\overline{\Omega_{Th}^{\pm}}), \quad K_1 = \max\left\{ \max_j \|b_j\|_{C(\overline{\Omega_{Th}^{\pm}})}, \|c\|_{C(\overline{\Omega_{Th}^{\pm}})} \right\}.$$

Cauchy problem for the parabolic equation with lateral Cauchy data. Let the function $f(x,t) \in L_2(\Omega_{Th}^{\pm})$. Find the function $u \in H^{2,1}(\Omega_{Th}^{\pm})$ satisfying the following conditions:

$$Lu = f \quad \text{in } \Omega_{Th}^{\pm}, \tag{4.56}$$

$$u|_{\Gamma_{Th}^{\pm}} = p_0(x,t), \quad \partial_n u|_{\Gamma_{Th}^{\pm}} = p_1(x,t), \tag{4.57}$$

where functions f, g_0, g_1 are given.

Assume that there exists a function $F \in H^{2,1}(\Omega_{Th}^{\pm})$ satisfying boundary conditions (4.57), that is,

$$F|_{\Gamma_{Th}^{\pm}} = p_0(x,t), \quad \partial_n u|_{\Gamma_{Th}^{\pm}} = p_1(x,t). \tag{4.58}$$

An example of this function can be constructed similarly with the one of Section 4.3. Assume now that there exists the exact solution $u^* \in H^{2,1}(\Omega_{Th}^{\pm})$ of problem (4.56), (4.57) with the exact Cauchy data $p_0^*, p_1^* \in H^{2,1}(\Gamma_{Th}^{\pm})$ in (4.57) and the exact function $f^* \in L_2(\Omega_{Th}^{\pm})$ in (4.56). Then there exists a function $F^* \in H^{2,1}(\Omega_{Th}^{\pm})$ satisfying

$$F^*|_{\Gamma_{Th}^{\pm}} = p_0^*(x,t), \quad \partial_n F^*|_{\Gamma_{Th}^{\pm}} = p_1^*(x,t). \tag{4.59}$$

Just as in (4.52), we assume that

$$\|F - F^*\|_{H^2(\Omega_h)} < \delta. \tag{4.60}$$

Consider the following functional of the QRM:

$$J_\beta(u) = \|Lu - f\|^2_{L_2(\Omega_{Th}^{\pm})} + \beta \|u - F\|^2_{H^{2,1}(\Omega_{Th}^{\pm})}. \tag{4.61}$$

Minimization problem for functional (4.61). Find a minimizer $u_\beta \in H^{2,1}(\Omega_{Th}^{\pm})$ of the functional $J_\beta(u)$ in (4.61) satisfying boundary conditions (4.57).

Theorem 4.4.1 follows immediately from Theorems 2.3.1, 4.2.1, and 4.2.2.

Theorem 4.4.1. *Assume that there exists a function $F \in H^{2,1}(\Omega_{Th}^{\pm})$ satisfying boundary conditions* (4.58). *Let $F^* \in H^{2,1}(\Omega_{Th}^{\pm})$ be a function satisfying conditions* (4.59) *and let* (4.60) *holds, where $\delta \in (0,1)$ is the level of noise. Then for any $\beta \in (0,1)$ there exists unique minimizer $u_\alpha \in H^{2,1}(\Omega_h)$ of functional* (4.61) *satisfying boundary conditions* (4.57) *and the following estimate holds with the constant $C > 0$ independent on β*:

$$\|u_\beta\|_{H^2(\Omega_h)} \le \frac{C}{\sqrt{\beta}} \big(\|F\|_{H^{2,1}(\Omega_h)} + \|f\|_{L_2(\Omega_h)}\big).$$

Furthermore, choose $\beta = \beta(\delta) = \delta^{2\rho}$, where $\beta = const. \in (0,1)$. Then there exists a number $\gamma = \gamma(\varepsilon) \in (0,1)$ and a sufficiently small number $\delta_0 \in (0,1)$ such that if $\delta \in (0, \delta_0^{1/\rho})$, then the following convergence rate of regularized solutions $u_{\alpha(\delta)}$ holds:

$$\|u_{\beta(\delta)} - u^*\|_{H^{1,0}(\Omega_{Th}^{\pm}(h^{-\nu}+3\varepsilon))} \le C\big(1 + \|u^*\|_{H^{2,1}(\Omega_{Th}^{\pm})}\big)\delta^{\beta\gamma}, \quad \forall \delta \in (0,\delta_0),$$

where the subdomain $\Omega_{Th}^{\pm}(h^{-\nu}+3\varepsilon)$ of the domain $\Omega_{Th}^{\pm}$ is defined in (4.55) *and the number ν_0 is defined in Theorem* 2.3.1.

4.5 Hyperbolic equation with lateral Cauchy data

Results of this section were originated in the work of Klibanov and Malinsky [153] and were developed further in works of Klibanov with coauthors [64, 132, 147, 158, 165]. Let $\Omega = \{|x| < R\} \subset \mathbb{R}^n$ and $T = \text{const.} > 0$. Denote

$$Q_T^{\pm} = \Omega \times (-T, T), \quad S_T^{\pm} = \partial\Omega \times (-T, T).$$

In this section, we obtain the Lipschitz type convergence rate (Theorem 4.5.1) in the whole time cylinder $Q_T^{\pm}$ rather than weaker Hölder type estimates in subdomains, as in previous sections. Corresponding numerical studies of [64, 147, 158] have demonstrated a good performance.

We repeat the statement of the Cauchy problem of Section 2.7. For the reader's convenience, we copy formulas (2.136)–(2.138) of that section,

$$Lu = c(x)u_{tt} - \Delta u - \sum_{j=1}^{n} b^j(x,t)u_{x_j} - d(x,t)u, \quad (x,t) \in Q_T^{\pm}, \tag{4.62}$$

$$Lu = L_0 u - \sum_{j=1}^{n} b^j(x,t)u_{x_j} - d(x,t)u,$$

$$L_0 u = c(x)u_{tt} - \Delta u, \tag{4.63}$$

$$b^j, d \in C(\overline{Q}_T^{\pm}), \quad K_1 = \max\Big(\|d\|_{C(\overline{Q}_T^{\pm})}, \max_j \|b^j\|_{C(\overline{Q}_T^{\pm})}\Big). \tag{4.64}$$

We impose the same condition on the function $c(x)$ as ones in (2.143)–(2.145)

$$c(x) \in [1, \overline{c}], \quad \text{where } \overline{c} = \text{const.} \geq 1, \tag{4.65}$$

$$c \in C^1(\overline{\Omega}), \tag{4.66}$$

$$\big(x, \nabla c(x)\big) \geq \alpha = \text{const.} > 0, \quad \forall x \in \overline{\Omega}. \tag{4.67}$$

Cauchy problem. Let L be the hyperbolic operator in (4.62) and let the function $f \in L_2(Q_T^{\pm})$. Find the solution $u \in H^2(Q_T^{\pm})$ of the equation

$$Lu = f \tag{4.68}$$

with the Cauchy data

$$u|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{S_T^{\pm}} = g_1(x,t). \tag{4.69}$$

Just as before, we assume that there exists a function $F \in H^2(Q_T^{\pm})$ satisfying Cauchy boundary conditions (4.69),

$$F|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n F|_{S_T^{\pm}} = g_1(x,t). \tag{4.70}$$

Consider the following functional of the QRM:

$$J_\beta(u) = \|Lu - f\|^2_{L_2(Q_T^{\pm})} + \beta\|u - F\|^2_{H^2(Q_T^{\pm})}. \tag{4.71}$$

Minimization problem for functional (4.71). Find a minimizer $u_\beta \in H^2(Q_T^\pm)$ of the functional $J_\beta(u)$ in (4.71) satisfying boundary conditions (4.69).

Suppose that there exists the exact solution $u^* \in H^2(Q_T^\pm)$ of problem (4.68), (4.69) with the exact Cauchy data $g_0^* \in H^1(S_T^\pm)$, $g_1^* \in L_2(S_T^\pm)$ in (4.69) and the exact function $f^* \in L_2(Q_T^\pm)$ in (4.68). Then there exists a function $F^* \in H^2(Q_T^\pm)$ satisfying

$$F^*|_{S_T^\pm} = g_0^*(x,t), \quad \partial_n F^*|_{S_T^\pm} = g_1^*(x,t). \tag{4.72}$$

Theorem 4.5.1. *Suppose that conditions* (4.62)–(4.67), (4.70)*, and* (4.72) *are satisfied. Then for every* $\beta > 0$ *there exists unique minimizer* $u_\beta \in H^2(Q_T^\pm)$ *of the functional* $J_\beta(u)$ *in* (4.71) *and*

$$\|u_\beta\|_{H^2(Q_T^\pm)} \le \frac{C}{\sqrt{\beta}}\left(\|F\|_{H^2(Q_T^\pm)} + \|f\|_{L_2(Q_T^\pm)}\right). \tag{4.73}$$

Furthermore, the following convergence rate of regularized solutions holds:

$$\left\|u_\beta - u^*\right\|_{H^1(Q_T^\pm)} \le C\left(\left\|F - F^*\right\|_{H^2(Q_T^\pm)} + \left\|f - f^*\right\|_{L_2(Q_T^\pm)} + \sqrt{\beta}\left\|u^*\right\|_{H^2(Q_T^\pm)}\right). \tag{4.74}$$

Proof. Existence and uniqueness of the minimizer u_β as well as estimate (4.73) follow immediately from Theorem 4.2.1. Thus, we now focus on the proof of the convergence rate (4.74). Denote

$$w_\beta = u_\beta - F, \quad w^* = u^* - F^*, \quad \widetilde{w} = w_\beta - w^*, \tag{4.75}$$

$$\widetilde{F} = F - F^*, \quad \widetilde{f} = f - f^*. \tag{4.76}$$

Let [,] and { , } be scalar products in $L_2(Q_T^\pm)$ and $H^2(Q_T^\pm)$, respectively. Denote

$$H_0^2(Q_T^\pm) = \{u \in H^2(Q_T^\pm) : u|_{S_T^\pm} = \partial_n u|_{S_T^\pm} = 0\}.$$

We obtain

$$[Lw_\beta, Lh] + \beta\{w_\beta, h\} = [f - LF, Lh], \quad \forall h \in H_0^2(Q_T^\pm), \tag{4.77}$$

$$\left[Lw^*, Lh\right] + \beta\left\{w^*, h\right\} = \left[f^* - LF^*, Lh\right] + \beta\left\{w^*, h\right\}, \quad \forall h \in H_0^2(Q_T^\pm). \tag{4.78}$$

Subtract (4.78) from (4.77). Taking into account (4.75) and (4.76), we obtain

$$[L\widetilde{w}, Lh] + \beta\{\widetilde{w}, h\} = [\widetilde{f} - L\widetilde{F}, Lh] - \beta\{w^*, h\}, \quad \forall h \in H_0^2(Q_T^\pm). \tag{4.79}$$

Setting in (4.79) $h = \widetilde{w}$ and using the Cauchy–Schwarz inequality, we obtain

$$\|L\widetilde{w}\|_{L_2(Q_T^\pm)}^2 + \beta\|\widetilde{w}\|_{H^2(Q_T^\pm)}^2 \le \|\widetilde{f} - L\widetilde{F}\|_{L_2(Q_T^\pm)}^2 + \beta\left\|w^*\right\|_{H^2(Q_T^\pm)}^2.$$

This implies that

$$\int\limits_{Q_T^\pm} (L\widetilde{w})^2 dxdt \le \|\widetilde{f} - L\widetilde{F}\|_{L_2(Q_T^\pm)}^2 + \beta\left\|w^*\right\|_{H^2(Q_T^\pm)}^2. \tag{4.80}$$

The target estimate (4.74) follows immediately from (4.80) and Theorem 2.7.3. □

5 Convexification for ill-posed Cauchy problems for quasi-linear PDEs

This chapter follows the work [9]. In addition, the short Subsection 5.3.2 uses the material of [132, 134, 135, 165]. Permissions for republishing are obtained from publishers.

5.1 Introduction

In this chapter, we start the presentation of the convexification method of Klibanov. We work in this chapter on the convexification for ill-posed Cauchy problems for quasilinear PDEs. These problems are simpler than CIPs. Many numerical methods have been developed for ill-posed Cauchy problems for linear PDEs; see, for example, the beginning of Chapter 4. However, prior to the work [135], rigorously justified numerical methods for ill-posed Cauchy problems for quasi-linear PDEs did not exist.

In the meantime, these problems have broad applications in processes involving high temperatures [4, 5]. In such a process, one can measure both the temperature and the heat flux on one side of the boundary. However, it is impossible to measure these quantities on the rest of the boundary. Still, one is required to compute the temperature in at least a part of the domain of interest. The underlying PDE, which governs the process of the propagation of this temperature, is a quasi-linear parabolic PDE. This equation is quasi-linear rather than linear because of high temperatures. Effective numerical methods for ill-posed Cauchy problems for quasi-linear parabolic equations were developed in [4, 5]. They work quite well numerically. These methods are based on the least square minimization, and the problem of local minima is not rigorously addressed in those works.

As it was pointed out in Chapter 1, conventional numerical methods for nonlinear ill-posed problems are based on the minimization procedure of the least squares cost functionals. The major drawback of this procedure, however, is that those functionals are typically nonconvex. Thus, they typically suffer from the phenomenon of local minima and ravines. This, in turn means that convergence to the exact solution of the underlying problem of any gradient-like numerical method for the minimization of such a functional can be rigorously guaranteed sometimes only if its starting point is located in an ε-neighborhood of the exact solution, where $\varepsilon > 0$ is a sufficiently small number [10, 11]. This is *local convergence*; see Definition 1.4.1 in Chapter 1. The main problem with the local convergence is that it is unclear how to get *a priori* a point in a sufficiently small neighborhood of the exact solution. This, in turn makes locally convergent numerical methods both unstable and unreliable. See Chapter 1 for the definitions of locally and globally convergent numerical methods.

Unlike the above, the convexification has the global convergence property. As to the definition of the global convergence, see Definition 1.4.2 in Chapter 1. Initial publi-

https://doi.org/10.1515/9783110745481-005

cations about the convexification were the ones in 1995 and 1997 [127, 140] for CIPs for hyperbolic PDEs. The convexification is the main topic of follow up Chapters 6–11 of the current book. Later, however, it was noticed that the important topic of numerical solutions of ill-posed Cauchy problems for quasi-linear PDEs remained not to be investigated. Thus, in [135] the convexification is extended to this case. The current chapter discusses this topic. See Section 1.4 for a brief description of the convexification.

5.2 Some facts of the convex analysis

Results of this section are known and can be found in Chapters 4 and 5 of the book of Vasiliev [248]. Still, we prove below Lemmata 5.2.1, 5.2.3, and Theorem 5.2.1 for the convenience of the reader. Even though all results of this section are formulated for a strictly convex functional, some of them are valid under less restrictive condition, which we do not list here for brevity.

Let H be a Hilbert space of real valued functions. Below in this section $\|\cdot\|$ and $(,)$ denote the norm and the scalar product in this space, respectively. Let $B(R) = \{x \in H : \|x\| < R\} \subset H$ be the ball of the radius R with the center at $\{0\}$. Even though results of this section can be easily extended to the case when $B(R)$ is a convex bounded set, we are not doing this here for brevity. Let $\delta > 0$ be a sufficiently small number. Let $J : B(R+\delta) \to \mathbb{R}$ be a functional, which has the Fréchet derivative $J'(x)$, $\forall x \in \overline{B}(R)$. Below we sometimes denote the action of the functional $J'(x)$ at the point x on any element $h \in H$ as $J'(x)(h)$. But sometimes we also denote this action as $(J'(x), h)$. This difference will not lead to a misunderstanding. The Fréchet derivative $J'(x)$ at a point $x \in \{\|x\| = R\}$ is understood as

$$J(y) - J(x) = J'(x)(y-x) + o(\|x-y\|), \quad \|x-y\| \to 0, y \in B(R+\delta).$$

We assume that this Fréchet derivative satisfies the Lipschitz continuity condition,

$$\|J'(x) - J'(y)\| \le L\|x-y\|, \quad \forall x, y \in \overline{B}(R), \tag{5.1}$$

with a certain constant $L > 0$. In addition, we assume that the functional $J(x)$ is strictly convex on the set $\overline{B}(R)$,

$$J(y) - J(x) - J'(x)(y-x) \ge \varkappa\|x-y\|^2, \quad \forall x, y \in \overline{B}(R), \tag{5.2}$$

where $\varkappa = const. > 0$. The strict convexity of $J(x)$ on $\overline{B}(R)$ implies

$$\left(J'(x) - J'(y), x-y\right) \ge 2\varkappa\|x-y\|^2, \quad \forall x, y \in \overline{B}(R). \tag{5.3}$$

The Weierstrass theorem implies existence of a point of a relative minimum of the functional $J(x)$ on the closed ball $\overline{B}(R)$.

Lemma 5.2.1. *A point $x_{\min} \in \overline{B}(R)$ is a point of a relative minimum of the functional $J(x)$ on the set $\overline{B}(R)$ if and only if*

$$\left(J'(x_{\min}), x_{\min} - y\right) \leq 0, \quad \forall y \in \overline{B}(R). \tag{5.4}$$

If a point $x_{\min} \in \overline{B}(R)$ is a point of a relative minimum of the functional $J(x)$ on the set $\overline{B}(R)$, then this point is unique and it is, therefore, the point of the unique global minimum of $J(x)$ on the set $\overline{B}(R)$.

Note that if $x_{\min}$ is an interior point of $B(R)$, then in (5.4) "≤" must be replaced with "=" and the assertion of this lemma becomes obvious. However, this assertion is not immediately obvious if $x_{\min}$ belongs to the boundary of the closed ball $\overline{B}(R)$.

Proof. Suppose that $x_{\min}$ is a point of a relative minimum of $J(x)$ on $\overline{B}(R)$. Assume to the contrary: that there exists a point $y \in \overline{B}(R)$ such that $(J'(x_{\min}), x_{\min} - y) > 0$. Let $h = y - x_{\min}$. Then

$$\left(J'(x_{\min}), \xi h\right) < 0, \quad \forall \xi > 0 \tag{5.5}$$

for any number $\xi > 0$. Since the set $\overline{B}(R)$ is convex, then $\{x_{\min} + \xi h, \xi \in [0,1]\} \subset \overline{B}(R)$. We have

$$J(x_{\min} + \xi h) = J(x_{\min}) + \xi\left[\left(J'(x_{\min}), h\right) + o(1)\right], \quad \xi \to 0^+. \tag{5.6}$$

By (5.5), $(J'(x_{\min}), h) + o(1) < 0$ for sufficiently small values of $\xi > 0$. Hence, (5.6) implies that $J(x_{\min} + \xi h) < J(x_{\min})$ for sufficiently small ξ. The latter contradicts the assumption that $x_{\min}$ is a point of a relative minimum of the functional $J(x)$ on the set $\overline{B}(R)$.

Assume now the reverse: that the inequality (5.4) is valid for a certain point $x_{\min} \in \overline{B}(R)$. We prove below that $x_{\min}$ is a point of a relative minimum of the functional $J(x)$ on the set $\overline{B}(R)$. Indeed, let $y \in \overline{B}(R)$ be an arbitrary point and let $y \neq x$. By (5.4) $J'(x_{\min})(y - x_{\min}) \geq 0$. Hence, (5.2) implies that

$$J(y) \geq J(x_{\min}) + J'(x_{\min})(y - x_{\min}) + \varkappa\|x - y\|^2 > J(x_{\min}). \tag{5.7}$$

Hence, the functional $J(x)$ attains its minimal value at $x = x_{\min}$. Hence, $x_{\min}$ is indeed the point of a relative minimum of the functional $J(x)$ on the set $\overline{B}(R)$.

We now prove uniqueness of the point of a relative minimum. Indeed, assume that there are two points $x_{\min}$ and $y_{\min}$ of relative minima of the functional $J(x)$ on the set $\overline{B}(R)$. We have

$$J(y_{\min}) - J(x_{\min}) - J'(x_{\min})(y_{\min} - x_{\min}) \geq \varkappa\|x_{\min} - y_{\min}\|^2, \tag{5.8}$$

$$J(x_{\min}) - J(y_{\min}) - J'(y_{\min})(x_{\min} - y_{\min}) \geq \varkappa\|x_{\min} - y_{\min}\|^2. \tag{5.9}$$

Summing up (5.8) and (5.9), we obtain

$$-J'(x_{\min})(y_{\min} - x_{\min}) - J'(y_{\min})(x_{\min} - y_{\min}) \geq 2\varkappa\|x_{\min} - y_{\min}\|^2. \tag{5.10}$$

However, by (5.4)

$$-J'(x_{\min})(y_{\min}-x_{\min})-J'(y_{\min})(x_{\min}-y_{\min})\le 0. \tag{5.11}$$

Hence, (5.10) and (5.11) imply that $x_{\min}=y_{\min}$. □

Let $y\in H$ be an arbitrary point. The point $\overline{y}$ is called projection of the point y on the set $\overline{B}(R)$ if

$$\|y-\overline{y}\|=\inf_{v\in\overline{B}(R)}\|y-v\|.$$

Lemma 5.2.2. *Each point $y\in H$ has unique projection $\overline{y}$ on the set $\overline{B}(R)$. Furthermore, the point $\overline{y}\in\overline{B}(R)$ is the projection of the point y on the set $\overline{B}(R)$ if and only if*

$$(\overline{y}-y,v-\overline{y})\ge 0,\quad \forall v\in\overline{B}(R). \tag{5.12}$$

For the proof of this lemma, we refer to Theorem 1 of Section 4 of Chapter 4 of [248]. Denote the projection operator of the space H on the set $\overline{B}(R)$ as $P_{\overline{B}(R)}:H\to\overline{B}(R)$. Then (see Theorem 2 of Section 4 of Chapter 4 of [248]),

$$\left\|P_{\overline{B}(R)}(u)-P_{\overline{B}(R)}(v)\right\|\le\|u-v\|,\quad \forall u,v\in H. \tag{5.13}$$

Lemma 5.2.3. *The point $x_{\min}\in\overline{B}(R)$ is the point of the unique global minimum of the functional $J(x)$ on the set $\overline{B}(R)$ if and only if there exits a number $\gamma>0$ such that*

$$x_{\min}=P_{\overline{B}(R)}\big(x_{\min}-\gamma J'(x_{\min})\big). \tag{5.14}$$

If (5.14) *is valid for one number γ, then it is also valid for all $\gamma>0$.*

Proof. Uniqueness of the global minimum, and the absence of other relative minima, follow from Lemma 5.2.1. By (5.4),

$$\big(\gamma J'(x_{\min}),v-x_{\min}\big)\ge 0,\quad \forall v\in\overline{B}(R),\forall\gamma>0.$$

This is equivalent with

$$\big(x_{\min}-\big(x_{\min}-\gamma J'(x_{\min})\big),v-x_{\min}\big)\ge 0,\quad \forall v\in\overline{B}(R),\forall\gamma>0. \tag{5.15}$$

Hence, Lemma 5.2.2 and (5.15) imply (5.14). □

Consider now the gradient projection method to find the minimum of the functional $J(x)$ on the set $\overline{B}(R)$. Let $x_0\in\overline{B}(R)$ be an arbitrary point. We construct the following sequence:

$$x_{n+1}=P_{\overline{B}(R)}\big(x_n-\gamma J'(x_n)\big),\quad n=0,1,2,\dots. \tag{5.16}$$

Theorem 5.2.1. *Assume that the functional $J(x)$ is strictly convex on the closed ball $\overline{B}(R)$ and let condition (5.1) hold. Then there exists unique point of the relative minimum $x_{\min}$ of this functional on the set $\overline{B}(R)$. Furthermore, $x_{\min}$ is the unique point of the global minimum of $J(x)$ on $\overline{B}(R)$. Let L and $\varkappa$ be numbers in (5.1) and (5.2), respectively, and let $\gamma \in (0, L]$. Let the number γ in (5.16) be so small that*

$$0 < \gamma < 2\varkappa L^{-2}. \tag{5.17}$$

Let $q(\alpha) = (1 - 2\gamma\varkappa + \alpha^2 L^2)^{1/2}$. Then the sequence (5.16) converges to the point $x_{\min}$ and

$$\|x_n - x_{\min}\| \le q^n(\gamma)\|x_0 - x_{\min}\|. \tag{5.18}$$

Proof. We note first that by (5.17) the number $q(\gamma) \in (0,1)$. The idea of the proof is to show that the operator in the right-hand side of (5.16) is contraction mapping, as long as (5.17) holds. Denote $D(x) = P_{\overline{B}(R)}(x - \gamma J'(x))$, $x \in \overline{B}(R)$. Then the operator $D : \overline{B}(R) \to \overline{B}(R)$. Let x and y be two arbitrary points of $\overline{B}(R)$. Using (5.13), we obtain

$$\begin{aligned} \left\|D(x) - D(y)\right\|^2 &\le \left\|(x - \gamma J'(x)) - (y - \gamma J'(y))\right\|^2 \\ &= \left\|(x - y) - \gamma(J'(x) - J'(y))\right\|^2 \\ &= \|x - y\|^2 + \gamma^2\left\|J'(x) - J'(y)\right\|^2 - 2\gamma(J'(x) - J'(y), x - y). \end{aligned} \tag{5.19}$$

By (5.1), $\gamma^2\|J'(x) - J'(y)\|^2 \le \gamma^2 L^2\|x - y\|^2$. Next, by (5.3),

$$-2\gamma(J'(x) - J'(y), x - y) \le -2\gamma\varkappa\|x - y\|^2.$$

Hence, (5.19) leads to

$$\left\|D(x) - D(y)\right\|^2 \le (1 - 2\gamma\varkappa + \gamma^2 L^2)\|x - y\|^2 = q^2(\gamma)\|x - y\|^2.$$

Hence, the operator D is a contraction mapping of the set $\overline{B}(R)$. The rest of the proof follows immediately from Lemmata 5.2.1 and 5.2.3. □

5.3 The general scheme of the method

5.3.1 The Cauchy problem

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Let A be a quasilinear Partial Differential Operator of the second order in Ω with its linear principal part A_0,

$$A(u) = \sum_{|\alpha|=2} a_\alpha(x) D^\alpha u + A_1(x, \nabla u, u), \tag{5.20}$$

$$A_0 u = \sum_{|\alpha|=2} a_\alpha(x) D^\alpha u, \tag{5.21}$$

$$a_\alpha \in C^1(\overline{\Omega}), \tag{5.22}$$

$$A_1(x,y) \in C^3\{(x,y) : x \in \overline{\Omega}, y \in \mathbb{R}^{n+1}\}. \tag{5.23}$$

Denote $k = [n/2] + 2$, where $[n/2]$ is the largest integer, which does not exceed the number $n/2$. By theembedding theorem

$$H^k(\Omega) \subset C^1(\overline{\Omega}) \quad \text{and} \quad \|f\|_{C^1(\overline{\Omega})} \leq C\|f\|_{H^k(\Omega)}, \quad \forall f \in H^k(\Omega), \tag{5.24}$$

where the constant $C = C(\Omega) > 0$ depends only on listed parameters. Let $\Gamma \subseteq \partial\Omega$, $\Gamma \in C^\infty$ be a part of the boundary of the domain Ω. We assume that Γ is not a part of the characteristic hypersurface of the operator A_0.

Cauchy problem 1. Consider the following Cauchy problem for the operator A:

$$A(u) = 0 \quad \text{in } \Omega, \tag{5.25}$$

$$u|_\Gamma = g_0(x), \quad \partial_n u|_\Gamma = g_1(x). \tag{5.26}$$

Find the solution $u \in H^k(\Omega)$ of the problem (5.25), (5.26) either in the entire domain Ω or at least in its subdomain.

The Cauchy–Kowalewski uniqueness theorem is inapplicable here since we do not impose the analyticity assumption on coefficients $a_\alpha(x)$ of the principal part A_0 of the operator A and also since A is not a linear operator. Still, Theorem 5.3.2 guarantees uniqueness of this problem in the domain Ω_c defined in Subsection 5.3.1.

Suppose that there exists a function $F \in H^{k+1}(\Omega)$ such that

$$F|_\Gamma = g_0(x), \quad \partial_n F|_\Gamma = g_1(x). \tag{5.27}$$

Consider the function $v(x) = u(x) - F(x)$. Here is an example of the function $F(x)$. Suppose that $\Omega = \{|x| < 1\} \subset \mathbb{R}^3$. Let $\Gamma = \{|x| = 1\}$. Assume that functions $g_0, g_1 \in C^{k+1}(\Gamma)$. Let the function $\chi(x) \in C^{k+1}(\overline{\Omega})$ be such that

$$\chi(x) = \begin{cases} 1, & |x| \in [3/4, 1], \\ \text{between 0 and 1} & \text{for } x \in (1/2, 3/4), \\ 0 & \text{for } x \in (0, 1/2). \end{cases}$$

The existence of such functions $\chi(x)$ is well known from the real analysis course. Then the function $F(x)$ can be constructed as $F(x) = \chi(x)[g_0(x) + (|x| - 1)g_1(x)]$.

Define the subspace $H_0^k(\Omega)$ of the Hilbert space of real valued functions $H^k(\Omega)$ as

$$H_0^k(\Omega) = \{f \in H^k(\Omega) : f|_\Gamma = 0, \partial_n f|_\Gamma = 0\}.$$

Hence, we come up with the following Cauchy problem.

Cauchy problem 2. Determine the function $v \in H_0^k(\Omega)$ such that

$$A(v + F) = 0 \quad \text{in } \Omega. \tag{5.28}$$

Note that the function $A(F) \in H^{k-1}(\Omega)$. By the embedding theorem, the latter means that $A(F) \in C(\overline{\Omega})$. In the realistic case, the Cauchy data $g_0(x)$, $g_1(x)$ are given with a random noise. On the other hand, by (5.27) one should have at least the following smoothness $g_0 \in H^k(\Gamma)$, $g_1 \in H^{k-1}(\Gamma)$. Hence, a data smoothing procedure might be applied to these functions in a data preprocessing procedure. A specific form of a smoothing procedure depends on a specific problem under the consideration. As a result, one would obtain the Cauchy data with a smooth error. A smoothing procedure is outside of the scope of this publication. Still, we work with noisy data in our computations; see Section 5.9.

5.3.2 The pointwise Carleman estimate

In this subsection, we briefly repeat our general scheme of the pointwise Carleman estimate, which was presented above in Sections 2.1 and 4.2.1. Let the function $\psi \in C^\infty(\overline{\Omega})$ and $|\nabla \psi| \neq 0$ in $\overline{\Omega}$. For a number $\alpha > 0$, denote

$$\psi_\alpha = \{x \in \overline{\Omega} : \psi(x) = \alpha\}, \quad \Omega_\alpha = \{x \in \Omega : \psi(x) > \alpha\}. \tag{5.29}$$

Hence, a part of the boundary $\partial\Omega_\alpha$ of the domain Ω_α is the level hypersurface ψ_α of the function ψ. We assume that $\Omega_\alpha \neq \varnothing$. Obviously, $\Omega_\omega \subset \Omega_\alpha$ if $\omega > \alpha$. Choose a sufficiently small number $\varepsilon > 0$ such that $\Omega_{\alpha+2\varepsilon} \neq \varnothing$. Denote $\Gamma_\alpha = \Gamma \cap \overline{\Omega}_\alpha$ and assume that $\Gamma_\alpha \neq \varnothing$. Hence, the boundary $\partial\Omega_\alpha$ of the domain Ω_α is

$$\partial\Omega_\alpha = \partial_1\Omega_\alpha \cup \partial_2\Omega_\alpha, \tag{5.30}$$

$$\partial_1\Omega_\alpha = \psi_\alpha, \quad \partial_2\Omega_\alpha = \Gamma_\alpha. \tag{5.31}$$

Let $\lambda > 1$ be a large parameter. Consider the function $\varphi_\lambda(x)$,

$$\varphi_\lambda(x) = \exp[\lambda\psi(x)]. \tag{5.32}$$

By (5.30)–(5.32),

$$\min_{\overline{\Omega}_\alpha} \varphi_\lambda(x) = \varphi_\lambda(x)|_{\psi_\alpha} = e^{\lambda\alpha}. \tag{5.33}$$

Let

$$m = \max_{\overline{\Omega}_\alpha} \psi(x). \tag{5.34}$$

Then

$$\max_{\overline{\Omega}_\alpha} \varphi_\lambda(x) = e^{\lambda m}. \tag{5.35}$$

Assume that the following pointwise estimate is valid for the principal part A_0 of the operator A:

$$(A_0 u)^2 \varphi_\lambda^2(x) \geq C_1 \lambda (\nabla u)^2 \varphi_\lambda^2(x) + C_1 \lambda^3 u^2 \varphi_\lambda^2(x) + \operatorname{div} U, \tag{5.36}$$

$$U = (U_1, \dots, U_n), \quad |U(x)| \leq C_1 \lambda^3 [(\nabla u)^2 + u^2] \varphi_\lambda^2(x), \tag{5.37}$$

$$\forall \lambda \geq \lambda_0, \forall x \in \Omega_\alpha, \forall u \in C^2(\overline{\Omega}_\alpha), \tag{5.38}$$

where constants $\lambda_0 = \lambda_0(A_0, \Omega) > 1$, $C_1 = C_1(A_0, \Omega) > 0$ depend only on listed parameters. Then the estimate (5.36) together with (5.37) and (5.38) is called a pointwise Carleman estimate for the operator A_0 with the CWF $\varphi_\lambda^2(x)$ in the domain Ω_α.

5.3.3 Theorems

Let $R > 0$ be an arbitrary number. We now specify the ball $B(R)$ as

$$B(R) = \{u \in H_0^k(\Omega) : \|u\|_{H^k(\Omega)} < R\}. \tag{5.39}$$

To solve the Cauchy problem 2, we take into account (5.28) and consider the following minimization problem.

Minimization problem. Assume that the operator A_0 satisfies conditions (5.36)–(5.38). Let $\beta \in (0,1)$ be the regularization parameter. Minimize with respect to the function $v \in \overline{B}(R)$ the functional $J_{\lambda,\beta}(v,F)$, where

$$J_{\lambda,\beta}(v,F) = e^{-2\lambda(\alpha+\varepsilon)} \int_\Omega [A(v+F)]^2 \varphi_\lambda^2 dx + \beta \|v\|_{H^k(\Omega)}^2. \tag{5.40}$$

The multiplier $e^{-2\lambda(\alpha+\varepsilon)}$ is introduced to balance two terms in the right-hand side of (5.40). Below “the Frechét derivative $J'_{\lambda,\beta}(v,F)$” means the Frechét derivative of the functional $J_{\lambda,\beta}(v,F)$ with respect to v. Also, below $[\,,\,]$ denotes the scalar product in $H^k(\Omega)$.

Theorem 5.3.1. *The functional $J_{\lambda,\beta}(v,F)$ has the Fréchet derivative $J'_{\lambda,\beta}(v,F) \in H_0^k(\Omega)$ for $v \in B(2R)$. This derivative satisfies the Lipschitz continuity condition*

$$\|J'_{\lambda,\beta}(v_1,F) - J'_{\lambda,\beta}(v_2,F)\|_{H^k(\Omega)} \leq L\|v_1 - v_2\|_{H^k(\Omega)}, \quad \forall v_1, v_2 \in \overline{B}(R), \tag{5.41}$$

where the constant $L = L(R, A, F, \Omega, \lambda, \alpha, \varepsilon, \beta) > 0$ depends only on listed parameters.

As to Theorem 5.3.2, we note that since $e^{-\lambda\varepsilon} \ll 1$ for sufficiently large λ, then the requirement of this theorem $\beta \in [e^{-\lambda\varepsilon}, 1)$ enables the regularization parameter β to change from being very small and up to the unity.

Theorem 5.3.2. *Assume that the operator A_0 admits the pointwise Carleman estimates* (5.36)–(5.38) *in the domain Ω_α. Then there exists a sufficiently large number $\lambda_1 = \lambda_1(R, A, F, \Omega) > \lambda_0(A_0, \Omega) > 1$ and a number $C_2 = C_2(R, A, F, \Omega) > 0$, both depending only on listed parameters, such that for all $\lambda \geq \lambda_1$ and for every $\beta \in [e^{-\lambda\varepsilon}, 1)$ the functional $J_{\lambda,\beta}(v, F)$ is strictly convex on the ball $\overline{B}(R)$,*

$$J_{\lambda,\beta}(v_2, F) - J_{\lambda,\beta}(v_1, F) - J'_{\lambda,\beta}(v_1, F)(v_2 - v_1) \tag{5.42}$$
$$\geq C_2 e^{2\lambda\varepsilon} \|v_2 - v_1\|^2_{H^1(\Omega_{\alpha+2\varepsilon})} + \frac{\beta}{2}\|v_2 - v_1\|^2_{H^k(\Omega)}, \quad \forall v_1, v_2 \in \overline{B}(R).$$

To minimize the functional (5.40) on the set $\overline{B}(R)$, we apply the gradient projection method. Let $P_{\overline{B}(R)} : H_0^k(\Omega) \to \overline{B}(R)$ be the projection operator of the space $H_0^k(\Omega)$ on the closed ball $\overline{B}(R)$ (Lemma 5.2.2). Let an arbitrary function $v_0 \in \overline{B}(R)$ be our starting point for iterations of this method. Let the step size of the gradient method be $\gamma > 0$. Consider the sequence $\{v_n\}_{n=0}^{\infty}$,

$$v_{n+1} = P_{\overline{B}(R)}\big(v_n - \gamma J'_{\lambda,\beta}(v_n, F)\big), \quad n = 0, 1, 2, \ldots. \tag{5.43}$$

For brevity, we do not indicate here the dependence of functions v_n on parameters λ, β, γ.

Theorem 5.3.3. *Suppose that all conditions of Theorem* 5.3.2 *are satisfied. Choose a number $\lambda \geq \lambda_1$. Let the regularization parameter $\beta \in [e^{-\lambda\varepsilon}, 1)$. Then there exists a point $v_{\min} \in \overline{B}(R)$ of the relative minimum of the functional $J_{\lambda,\beta}(v)$ on the set $\overline{B}(R)$. Furthermore, $v_{\min}$ is also the unique point of the global minimum of this functional on $\overline{B}(R)$. Consider the sequence* (5.43)*, where $v_0 \in \overline{B}(R)$ is an arbitrary point of the closed ball $\overline{B}(R)$. Then there exist a sufficiently small number $\gamma = \gamma(R, A, F, \Omega, \alpha, \varepsilon, \beta, \lambda) \in (0, 1)$ and a number $q(\gamma) \in (0, 1)$, both depending only on listed parameters, such that the sequence* (5.43) *converges to the point $v_{\min}$,*

$$\|v_{n+1} - v_{\min}\|_{H^k(\Omega)} \leq q^n(\gamma)\|v_0 - v_{\min}\|_{H^k(\Omega)}, \quad n = 0, 1, 2, \ldots. \tag{5.44}$$

Following the regularization theory [22, 76, 244], the next natural question to address is whether regularized solutions converge to the exact solution (if it exists) for some values of the parameter $\lambda = \lambda(\delta)$ if the level of the error δ in the Cauchy data g_0, g_1 tends to zero. Since functions g_0, g_1 generate the function F, we consider the error only in F. Following one of concepts of the regularization theory, we assume now the existence of the exact solution $v^* \in H_0^k(\Omega)$ of the problem (5.28), which satisfies the following conditions:

$$A\big(v^* + F^*\big) = 0, \tag{5.45}$$

$$v^* \in B(R), \tag{5.46}$$

where the function $F^* \in H^{k+1}(\Omega)$ is generated by the exact (i. e., noiseless) Cauchy data $g_0^*(x)$ and $g_1^*(x)$. We assume that

$$\|F - F^*\|_{H^{k+1}(\Omega)} \le \delta, \tag{5.47}$$

where $\delta \in (0,1)$ is a sufficiently small number characterizing the level of the error in the data. The construction (5.45)–(5.47) corresponds well with the regularization theory [22, 184, 244]. First, consider the case when the data are noiseless, that is, when $\delta = 0$.

Theorem 5.3.4. *Suppose that all conditions of Theorem* 5.3.2 *are satisfied. Choose a number* $\lambda^* = \lambda^*(R, A, F^*, \Omega) > \lambda_0$ *such that estimate* (5.42) *is valid for* $J_{\lambda,\beta}(v, F^*)$ *for all* $\lambda \ge \lambda^*$. *Let the level of the error in the data be* $\delta = 0$. *Choose* $\lambda \ge \lambda^*$ *and* $\beta = e^{-\lambda\varepsilon}$. *Let* $v_{\min} \in \overline{B}(R)$ *be the point of the unique global minimum on* $\overline{B}(R)$ *of the functional* $J_{\lambda,\beta}(v, F^*)$ (*Theorem* 5.3.3). *Then there exists a constant* $C_3 = C_3(R, A, F^*, \Omega) > 0$ *depending only on listed parameters such that*

$$\|v^* - v_{\min}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \le C_3 \exp(-3\lambda\varepsilon/2). \tag{5.48}$$

Furthermore, let $\{v_n\}_{n=0}^{\infty}$ *be the sequence* (5.43) *where the number* $\gamma = \gamma(R, A, F^*, \Omega, \alpha, \varepsilon, \beta, \lambda) \in (0,1)$ *is the same as in Theorem* 5.3.3. *Then with the same constant* $q(\gamma) \in (0,1)$ *as in Theorem* 5.3.3 *the following estimate holds:*

$$\|v^* - v_{n+1}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \le C_3 \exp(-3\lambda\varepsilon/2) + q^n(\gamma)\|v_0 - v_{\min}\|_{H^k(\Omega)}, \quad n = 0, 1, 2, \ldots. \tag{5.49}$$

Let m be the number in (5.34). Denote

$$\theta = \min\left(\frac{\varepsilon}{4m}, \frac{1}{2}\right). \tag{5.50}$$

Theorem 5.3.5 estimates the rate of convergence of minimizers $v_{\min}$ to the exact solution v^* in the norm of the space $H^1(\Omega_{\alpha+2\varepsilon})$.

Theorem 5.3.5. *Let all conditions of Theorem* 5.3.2 *hold. Let the number* $\lambda_1 = \lambda_1(R, A, F, \Omega) > \lambda_0$ *be the same as in Theorem* 5.3.2 *and let* θ *be the number defined in* (5.50). *Let the number* $\delta_0 \in (0,1)$ *be so small that* $\delta_0^{-1/(2m)} > e^{\lambda_1}$. *Let* $\delta \in (0, \delta_0)$ *be the level of the error in the function F, that is, let* (5.47) *be valid. Choose* $\lambda = \lambda(\delta) = \ln(\delta^{-1/(2m)}) > \lambda_1$ *and* $\beta = e^{-\lambda(\delta)\varepsilon}$. *Let* $v_{\min} \in \overline{B}(R)$ *be the point of the unique global minimum on* $\overline{B}(R)$ *of the functional* $J_{\lambda,\beta}(v, F)$ (*Theorem* 5.3.3). *Then there exists a constant* $C_4 = C_4(R, A, F, \Omega) > 0$ *depending only on listed parameters such that*

$$\|v^* - v_{\min}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \le C_4\delta^{\theta}. \tag{5.51}$$

Next, let $\{v_n\}_{n=0}^{\infty}$ *be the sequence* (5.43), *where the number* $\gamma = \gamma(R, A, F, \Omega, \alpha, \varepsilon, \beta, \delta) \in (0,1)$ *is the same as in Theorem* 5.3.3. *Then with the same constant* $q(\gamma) \in (0,1)$ *as in Theorem* 5.3.3 *the following estimate holds:*

$$\|v^* - v_{n+1}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \le C_4\delta^{\theta} + q^n(\gamma)\|v_0 - v_{\min}\|_{H^k(\Omega)}, \quad n = 0, 1, 2, \ldots. \tag{5.52}$$

Theorem 5.3.3 follows immediately from Theorems 5.2.1, 5.3.1, and 5.3.2. Hence, we do not prove Theorem 5.3.3 here.

5.4 Proof of Theorem 5.3.1

In this proof, $L = L(R, A, F, \Omega, \alpha, \varepsilon, \beta, \lambda) > 0$ denotes different numbers depending only on listed parameters. Let $v_1, v_2 \in B(2R)$ be two arbitrary functions. Denote $h = v_2 - v_1$. Hence, $h \in H_0^k(\Omega)$. Let

$$D = \left(A(v_2 + F)\right)^2 - \left(A(v_1 + F)\right)^2. \tag{5.53}$$

By the Lagrange formula,

$$f(y+z) = f(y) + f'(y)z + \frac{z^2}{2} f''(\eta), \quad \forall y, z \in \mathbb{R}, \forall f \in C^2(\mathbb{R}), \tag{5.54}$$

where $\eta = \eta(y, z)$ is a number located between numbers y and $y + z$. By (5.24),

$$\|h\|_{C^1(\overline{\Omega})} = \|v_2 - v_1\|_{C^1(\overline{\Omega})} \le 4CR. \tag{5.55}$$

Hence, using (5.20)–(5.23), (5.54), and (5.55), we obtain

$$\begin{aligned} A_1\big(x, \nabla(v_2 + F), v_2 + F\big) &= A_1\big(x, \nabla(v_1 + F + h), v_1 + F + h\big) \\ &= A_1(x, \nabla v_1 + \nabla F, v_1 + F) \\ &\quad + \sum_{i=1}^n \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \\ &\quad + P(x, \nabla v_1 + \nabla F, v_1 + F, \nabla h, h), \end{aligned}$$

where the function P satisfies the following estimate:

$$\big|P(x, \nabla v_1 + \nabla F, v_1 + F, \nabla h, h)\big| \le K\big((\nabla h)^2 + h^2\big), \quad \forall x \in \overline{\Omega}, \forall v_1 \in B(2R), \tag{5.56}$$

where the constant $K = K(R, F, \Omega) > 0$ depends only on listed parameters. Hence,

$$\begin{aligned} A(v_2 + F) &= A_0(v_1 + F + h) + A_1\big(x, \nabla(v_1 + F + h), v_1 + F + h\big) = A(v_1 + F) \\ &\quad + \left[A_0(h) + \sum_{i=1}^n \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h\right] \\ &\quad + P(x, \nabla u_1, u_1, h). \end{aligned}$$

Hence, by (5.53),

$$\begin{aligned} D &= 2A(v_1 + F) \\ &\quad \times \left[A_0(h) + \sum_{i=1}^n \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h\right] \end{aligned} \tag{5.57}$$

$$+ \left[A_0(h) + \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \right]^2$$
$$+ P^2.$$

The expression in the first two lines of (5.57) is linear with respect to h. We denote this expression as $Q(v_1 + F)(h)$,

$$Q(v_1 + F)(h) = 2A(v_1 + F) \times \left[A_0(h) + \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \right]. \tag{5.58}$$

Consider the linear functional acting on functions $h \in H_0^k(\Omega)$ as

$$\widetilde{J}(v_1, F)(h) = \int_{\Omega} Q(v_1 + F)(h) \varphi_\lambda^2 dx + 2\beta[v_1, h]. \tag{5.59}$$

Clearly, $\widetilde{J}(v_1, F)(h) : H_0^k(\Omega) \to \mathbb{R}$ is a bounded linear functional. Hence, by the Riesz theorem, there exists a single element $M(v_1) \in H_0^k(\Omega)$ such that

$$\widetilde{J}(v_1, F)(h) = [M(v_1, F), h], \quad \forall h \in H_0^k(\Omega). \tag{5.60}$$

Furthermore,

$$\|M(v_1, F)\|_{H^k(\Omega)} = \|\widetilde{J}(v_1, F)\|. \tag{5.61}$$

Next, since by (5.24) $\|h\|_{C^1(\overline{\Omega})} \leq C\|h\|_{H^k(\Omega)}$, then (5.40), (5.53), (5.57), and (5.59) imply that

$$J_{\lambda,\beta}(v_1 + h, F) - J_{\lambda,\beta}(v_1, F) - \widetilde{J}(v_1, F)(h) = O(\|h\|_{H^k(\Omega)}^2), \tag{5.62}$$

as $\|h\|_{H^k(\Omega)} \to 0$. The existence of the Fréchet derivative $J'_{\lambda,\beta}(v_1)$ follows from (5.58)–(5.62). Also, for all $h \in H_0^k(\Omega)$ and all $v \in B(2R)$,

$$J'_{\lambda,\beta}(v, F)(h) = \widetilde{J}(v, F)(h) = \int_{\Omega} Q(v + F)(h) \varphi_\lambda^2 dx + 2\beta[v, h], \tag{5.63}$$

$$J'_{\lambda,\beta}(v, F) = M(v, F) \in H_0^k(\Omega). \tag{5.64}$$

We now prove the Lipschitz continuity of the Fréchet derivative $J'_{\lambda,\beta}(v, F)$. By (5.57), (5.58), (5.59), (5.63), and (5.64) we should analyze the following expression for all $v_1, v_2 \in \overline{B}(R)$ and for all $h \in H_0^k(\Omega)$:

$$Y(v_1, h) - Y(v_2, h)$$
$$= 2A(v_1 + F)$$

$$\times \left[A_0(h) + \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \right] \quad (5.65)$$
$$- 2A(v_2 + F)$$
$$\times \left[A_0(h) + \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_2 + \nabla F, v_2 + F) h_{x_i} + \partial_v A_1(x, \nabla v_2 + \nabla F, v_2 + F) h \right].$$

We have

$$\begin{aligned} &Y(v_1, h) - Y(v_2, h) \\ &\quad = 2\big(A(v_1 + F) - A(v_2 + F)\big) \\ &\qquad \times \left[A_0(h) + \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \right] \\ &\qquad + 2A(v_2 + F) \qquad (5.66) \\ &\qquad \times \left[\sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} + \partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F) h \right] \\ &\qquad - 2A(v_2 + F) \\ &\qquad \times \left[\sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_2 + \nabla F, v_2 + F) h_{x_i} + \partial_v A_1(x, \nabla v_2 + \nabla F, v_2 + F) h \right]. \end{aligned}$$

First, using (5.20) and (5.54), we obtain

$$\begin{aligned} &2\big(A(v_1 + F) - A(v_2 + F)\big) \\ &\quad = 2A_0(v_1 - v_2) \qquad (5.67) \\ &\qquad + 2\left[\sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_2 + \nabla F, v_2 + F)(v_1 - v_2)_{x_i} + \partial_v A_1(x, \nabla v_2 + \nabla F, v_2 + F)(v_1 - v_2) \right] \\ &\qquad + Y_1(x, v_1, v_2), \end{aligned}$$

where

$$\big|Y_1(x, v_1, v_2)\big| \le L\|v_1 - v_2\|^2_{C^1(\overline{\Omega})} \le L\|v_1 - v_2\|_{H^k(\Omega)}, \quad \forall v_1, v_2 \in \overline{B}(R). \qquad (5.68)$$

Thus, (5.67) and (5.68) imply that the modulus of the expression in the first two lines of (5.66) can be estimated from the above via Y_2, where

$$Y_2 \le L\|v_1 - v_2\|_{H^k(\Omega)} \|h\|_{H^k(\Omega)}, \quad \forall v_1, v_2 \in \overline{B}(R), \forall h \in H_0^k(\Omega). \qquad (5.69)$$

Estimate now from the above the modulus of the expression in the line numbers 3–6 of (5.66). By (5.54),

$$\begin{aligned} &\partial_v A_1(x, \nabla v_1 + \nabla F, v_1 + F)h - \partial_v A_1(x, \nabla v_2 + \nabla F, v_2 + F)h \\ &\quad = \partial_v^2 A_1(x, \nabla v_2 + \nabla F, v_2 + F) h(v_1 - v_2) + \frac{(v_1 - v_2)^2}{2} h \partial_v^3 A_1\big(x, \nabla v_2 + \nabla F, \xi(x) + F\big), \end{aligned}$$

where the point $\xi(x)$ is located between points $v_1(x)$ and $v_2(x)$. Similar formulas are valid of course for terms

$$\sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_1 + \nabla F, v_1 + F) h_{x_i} - \sum_{i=1}^{n} \partial_{v_{x_i}} A_1(x, \nabla v_2 + \nabla F, v_2 + F) h_{x_i}.$$

Hence, the modulus of the expression in line numbers 3-6 of (5.66) can be estimated from the above similarly with (5.69) via Y_3, where

$$Y_3 \leq L \|v_1 - v_2\|_{H^k(\Omega)} \|h\|_{H^k(\Omega)}, \quad \forall v_1, v_2 \in \overline{B}(R), \forall h \in H_0^k(\Omega). \tag{5.70}$$

Thus, (5.58) and (5.63)–(5.70) imply that

$$\left| J'_{\lambda,\beta}(v_1, F)(h) - J'_{\lambda,\beta}(v_2, F)(h) \right| \leq L \|v_1 - v_2\|_{H^k(\Omega)} \|h\|_{H^k(\Omega)},$$

for all $v_1, v_2 \in \overline{B}(R)$ and for all $h \in H_0^k(\Omega)$. This in turn implies (5.41).

5.5 Proof of Theorem 5.3.2

In this proof, $C_2 = C_2(R, A, F, \Omega) > C_1 > 0$ denotes different constants depending only on listed parameters. Here, $C_1 = C_1(A_0, \Omega) > 0$ is the constant of the pointwise Carleman estimates (5.36)–(5.38). For two arbitrary points, $v_1, v_2 \in \overline{B}(R)$ let again $h = v_2 - v_1$ and let D be the same as in (5.53). Denote $S = D - Q(v_1 + F)(h)$, where $Q(v_1 + F)(h)$ is given in (5.58) and it is linear, with respect to h. Then, using (5.56)–(5.58) and the Cauchy–Schwarz inequality, we obtain

$$S \geq \frac{1}{2}(A_0 h)^2 - C_2((\nabla h)^2 + h^2), \quad \forall x \in \Omega.$$

Hence, using (5.62) and (5.63), we obtain

$$\begin{aligned} &J_{\lambda,\beta}(v_1 + h, F) - J_{\lambda,\beta}(v_1, F) - J'_{\lambda,\beta}(v_1, F)(h) \\ &\quad \geq \frac{1}{2} e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega} (A_0 h)^2 \varphi_\lambda^2 dx - C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega} ((\nabla h)^2 + h^2) \varphi_\lambda^2 dx + \beta \|h\|_{H^k(\Omega)}^2. \end{aligned} \tag{5.71}$$

Since $\Omega_\alpha \subset \Omega$, then

$$e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega} (A_0 h)^2 \varphi_\lambda^2 dx \geq e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (A_0 h)^2 \varphi_\lambda^2 dx. \tag{5.72}$$

Next,

$$\begin{aligned} -C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega} ((\nabla h)^2 + h^2) \varphi_\lambda^2 dx &= -C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} ((\nabla h)^2 + h^2) \varphi_\lambda^2 dx \\ &\quad - C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega \setminus \Omega_\alpha} ((\nabla h)^2 + h^2) \varphi_\lambda^2 dx. \end{aligned} \tag{5.73}$$

Since by (5.29) and (5.32) $\varphi_\lambda^2(x) < \exp(2\lambda\alpha)$ for $x \in \Omega \setminus \Omega_\alpha$, then

$$-C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega\setminus\Omega_\alpha} ((\nabla h)^2 + h^2)\varphi_\lambda^2 dx \geq -C_2 e^{-2\lambda\varepsilon} \int_{\Omega\setminus\Omega_\alpha} ((\nabla h)^2 + h^2) dx. \tag{5.74}$$

Integrate (5.36) over the domain Ω_α, using the Gauss' formula, (5.37) and (5.38). Next, replace u with h in the resulting formula. Even though there is no guarantee that $h \in C^2(\overline{\Omega}_\alpha)$, still density arguments ensure that the resulting inequality remains true. Hence, taking into account (5.29)–(5.33), (5.71), and (5.72), we obtain

$$\begin{aligned} \frac{1}{2} e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (A_0 h)^2 \varphi_\lambda^2 dx \geq C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (\lambda(\nabla h)^2 + \lambda^3 h^2)\varphi_\lambda^2 dx \\ - C_2 \lambda^3 e^{-2\lambda\varepsilon} \int_{\psi_\alpha} ((\nabla h)^2 + h^2) dS, \quad \forall \lambda \geq \lambda_0. \end{aligned} \tag{5.75}$$

Since $k \geq 2$, then the trace theorem implies that

$$C_2 \lambda^3 e^{-2\lambda\varepsilon} \int_{\psi_\alpha} ((\nabla h)^2 + h^2) dx \leq C_2 \lambda^3 e^{-2\lambda\varepsilon} \|h\|_{H^k(\Omega)}^2. \tag{5.76}$$

Also,

$$C_2 e^{-2\lambda\varepsilon} \int_{\Omega\setminus\Omega_\alpha} ((\nabla h)^2 + h^2) dx \leq C_2 e^{-2\lambda\varepsilon} \|h\|_{H^k(\Omega)}^2. \tag{5.77}$$

Since $\beta \geq e^{-\lambda\varepsilon}$, then (5.76) and (5.77) imply that for sufficiently large $\lambda_1 = \lambda_1(R, A, F, \Omega, \alpha, \varepsilon, \beta) > \lambda_0$ and for $\lambda \geq \lambda_1$,

$$-C_2 \lambda^3 e^{-2\lambda\varepsilon} \int_{\psi_\alpha} ((\nabla h)^2 + h^2) dx - C_2 e^{-2\lambda\varepsilon} \int_{\Omega\setminus\Omega_\alpha} ((\nabla h)^2 + h^2) dx \geq -\frac{\beta}{2} \|h\|_{H^k(\Omega)}^2. \tag{5.78}$$

Also, for $\lambda \geq \lambda_1$

$$\begin{aligned} C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (\lambda(\nabla h)^2 + \lambda^3 h^2)\varphi_\lambda^2 dx - C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} ((\nabla h)^2 + h^2)\varphi_\lambda^2 dx \\ \geq \frac{1}{2} C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (\lambda(\nabla h)^2 + \lambda^3 h^2)\varphi_\lambda^2 dx. \end{aligned} \tag{5.79}$$

Hence, using (5.71), (5.73), (5.75), (5.78), and (5.79), we obtain for $\lambda \geq \lambda_1$ with a new constant C_2,

$$\begin{aligned} J_{\lambda,\beta}(v_1 + h, F) - J_{\lambda,\beta}(v_1, F) - J'_{\lambda,\beta}(v_1, F)(h) \\ \geq C_2 e^{-2\lambda(\alpha+\varepsilon)} \int_{\Omega_\alpha} (\lambda(\nabla h)^2 + \lambda^3 h^2)\varphi_\lambda^2 dx + \frac{\beta}{2} \|h\|_{H^k(G_c)}^2. \end{aligned} \tag{5.80}$$

Next, since $\Omega_{\alpha+2\varepsilon} \subset \Omega_\alpha$ and $\varphi_\lambda^2(x) > e^{2\lambda(\alpha+2\varepsilon)}$ for $x \in \Omega_{\alpha+2\varepsilon}$, then (5.80) implies that for all $v_1, v_2 = v_1 + h \in \overline{B}(R)$,

$$J_{\lambda,\beta}(v_1 + h, F) - J_{\lambda,\beta}(v_1, F) - J'_{\lambda,\beta}(v_1, F)(h) \geq C_2 e^{2\lambda\varepsilon} \|h\|^2_{H^1(\Omega_{\alpha+2\varepsilon})} + \frac{\beta}{2} \|h\|^2_{H^k(\Omega)}.$$

5.6 Proof of Theorem 5.3.4

Recall that $\lambda \geq \lambda^*$. The existence and uniqueness of the point $v_{\min} \in \overline{B}(R)$ of the global minimum of the functional $J_{\lambda,\beta}(v, F^*)$ follows immediately from Theorems 5.2.1, 5.3.2, and 5.3.3. Since by (5.45) $A(v^* + F^*) = 0$ and by (5.46) $v^* \in B(R)$, then, using (5.40), we obtain

$$J_{\lambda,\beta}(v^*, F^*) = \beta \|v^*\|^2_{H^k(\Omega)} \leq \beta R^2. \tag{5.81}$$

Next, by (5.4)

$$-J'_{\lambda,\beta}(v_{\min}, F^*)(v^* - v_{\min}) \leq 0. \tag{5.82}$$

Hence, combining (5.81) and (5.82), we obtain

$$J_{\lambda,\beta}(v^*, F^*) - J_{\lambda,\beta}(v_{\min}, F^*) - J'_{\lambda,\beta}(v_{\min}, F^*)(v^* - v_{\min}) \leq \beta R^2. \tag{5.83}$$

Next, combining (5.83) with Theorem 5.3.2 and setting $\beta = e^{-\lambda\varepsilon}$, we obtain (5.48). Next, since

$$\|v^* - v_{n+1}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \leq \|v^* - v_{\min}\|_{H^1(\Omega_{\alpha+2\varepsilon})} + \|v_{\min} - v_{n+1}\|_{H^1(\Omega_{\alpha+2\varepsilon})}$$

and $\|v_{\min} - v_{n+1}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \leq \|v_{\min} - v_{n+1}\|_{H^k(\Omega)}$, then (5.49) follows from (5.48) and (5.44).

5.7 Proof of Theorem 5.3.5

In this proof, $C_4 = C_4(R, A, F, \Omega) > 0$ denotes different constants depending only on listed parameters. Since functions $F, F^* \in H^{k+1}(\Omega)$, then, as it was noticed in Subsection 5.3.1,

$$A(F), A(F^*) \in C(\overline{\Omega}). \tag{5.84}$$

It follows from (5.20), (5.45)–(5.47), (5.54)- and (5.84) that

$$A(v^* + F) = A(v^* + F^* + (F - F^*)) = A(v^* + F^*) + \widetilde{A}(v^*, F - F^*) = \widetilde{A}(v^*, F - F^*),$$

where $|\widetilde{A}(v^*, F - F^*)| \leq C_4 \delta$, $\forall x \in \overline{\Omega}$. Hence, recalling that $v^* \in B(R)$ and applying (5.35) and (5.40), we obtain

$$J_{\lambda,\beta}(v^*, F) \leq C_4(\delta^2 e^{2\lambda m} + \beta). \tag{5.85}$$

Recall that $\lambda \geq \lambda_1$. Let $v_{\min} \in \overline{B}(R)$ be the unique point of the global minimum of the functional $J_{\lambda,\beta}(v, F)$ on the set $\overline{B}(R)$. Since (5.82) is valid, then, using Theorem 5.3.2 and (5.85), we obtain

$$\|v^* - v_{\min}\|_{H^1(\Omega_{\alpha+2\varepsilon})} \leq C_4(\delta e^{\lambda m} + \sqrt{\beta}). \tag{5.86}$$

Choose $\lambda = \lambda(\delta)$ such that $\delta \exp(\lambda(\delta)m) = \sqrt{\delta}$. This means that

$$\lambda(\delta) = \ln\left(\frac{1}{\delta^{1/(2m)}}\right). \tag{5.87}$$

The choice (5.87) is possible since $\ln(\delta_0^{-1/(2m)}) > \lambda_1$ and, therefore, $\lambda(\delta) > \lambda_1$ for $\delta \in (0, \delta_0)$. Choose $\beta = e^{-\lambda(\delta)\varepsilon}$. Hence, taking into account (5.50), we obtain

$$\delta e^{\lambda(\delta)m} + e^{-\lambda(\delta)\varepsilon/2} = \sqrt{\delta} + \delta^{\varepsilon/(4m)} \leq 2\delta^{\theta}. \tag{5.88}$$

Thus, (5.86)–(5.88) imply (5.51). Next, (5.52) is established similarly with the part of the proof of Theorem 5.3.4 after (5.83).

5.8 Specifying equations

The scheme of Section 5.3 is a general one and it can be applied to all three main classes of partial differential equations of the second order: elliptic, parabolic, and hyperbolic ones. So, Theorems 5.2.1, 5.3.1–5.3.5 can be reformulated for all three Cauchy problems considered in this section.

5.8.1 Quasilinear elliptic equation

We now rewrite the operator A in (5.20) as

$$A_{\text{ell}}(u) = \sum_{i,j=1}^{n} a_{i,j}(x)u_{x_i x_j} + A_1(x, \nabla u, u), \quad x \in \Omega, \tag{5.89}$$

$$A_{0,\text{ell}}(u) = \sum_{i,j=1}^{n} a_{i,j}(x)u_{x_i x_j}, \tag{5.90}$$

$$a_{i,j} \in C^1(\overline{\Omega}), \tag{5.91}$$

where $a_{i,j}(x) = a_{j,i}(x)$, $\forall i,j = 1,\dots,n$ and A_0 is the principal part of the operator A_{ell}. Condition (5.22) becomes now condition (5.91). Also, we assume that condition (5.23) holds. The ellipticity of the operator $A_{0,ell}$ means that there exist two constants $\mu_1, \mu_2 > 0$, $\mu_1 \leq \mu_2$ such that

$$\mu_1|\eta|^2 \leq \sum_{i,j=1}^{n} a_{i,j}(x)\eta_i\eta_j \leq \mu_2|\eta|^2, \quad \forall x \in \overline{\Omega}, \forall \eta = (\eta_1, \dots \eta_n) \in \mathbb{R}^n. \tag{5.92}$$

As above, let $\Gamma \subset \partial\Omega$ be the part of the boundary $\partial\Omega$, where the Cauchy data are given. Assume that the equation of Γ is

$$\Gamma = \{x \in \mathbb{R}^n : x_1 = p(\overline{x}), \overline{x} = (x_2, \dots, x_n) \in \Gamma' \subset \mathbb{R}^{n-1}\}$$

and that the function $p \in C^2(\overline{\Gamma}')$. Here $\Gamma' \subset \mathbb{R}^{n-1}$ is a bounded domain. Changing variables $x = (x_1, \overline{x}) \Leftrightarrow (x_1', \overline{x})$, where $x_1' = x_1 - p(\overline{x})$ and keeping the same notation for x_1 for brevity, we obtain that in new variables

$$\Gamma = \{x \in \mathbb{R}^n : x_1 = 0, \overline{x} \in \Gamma'\}.$$

This change of variables does not affect the ellipticity property of the operator $A_{0,\mathrm{ell}}$. Let $X > 0$ be a certain number. Let $X > 0$ be a number. Without loss of generality, we assume that, as in (4.42),

$$\Omega \subset \{x_1 > 0\}, \quad Y = \{x \in \mathbb{R}^n : x_1 = 0, |\overline{x}| < 2X\} \subset \partial\Omega. \tag{5.93}$$

Cauchy problem for the quasi-linear elliptic equation.
Suppose that conditions (5.89)–(5.93) hold. Let the functions $g_0(\overline{x})$ and $g_1(\overline{x})$ be known for $\overline{x} \in \Gamma$. Find such a function $u \in H^k(\Omega)$ that satisfies the following conditions:

$$A_{\mathrm{ell}}(u) = 0, \tag{5.94}$$

$$u|_\Gamma = g_0(\overline{x}), \quad u_{x_1}|_\Gamma = g_1(\overline{x}). \tag{5.95}$$

Let $h, \alpha \in (0,1)$ be some numbers and $\alpha < h$. Similarly, with Section 4.3 we define

$$\Omega_h = \left\{x : x_1 > 0, x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha < h\right\},$$

and assume that $\overline{\Omega}_h \subset \Omega$. Next, as in (4.45), we define the CWF $\varphi_\lambda(x)$ for the operator $A_{0,\mathrm{ell}}$ as

$$\psi(x) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \alpha, \quad \varphi_\lambda(x) = \exp(\lambda\psi^{-\nu}).$$

Here, the number $\nu \geq \nu_0$, where $\nu_0 = \nu_0(\Omega, n, \rho, X, A_{0,\mathrm{ell}}) > 1$ is a certain number depending only on listed parameters. Then Theorems 5.2.1, 5.3.1–5.3.5 can be reformulated for problem (5.94), (5.95).

5.8.2 Quasilinear parabolic equation

Since in this and next subsections we work with the space $\mathbb{R}^{n+1} = \{(x,t), x \in \mathbb{R}^n, t \in \mathbb{R}\}$, then we replace the above number k with $k_{n+1} = [(n+1)/2] + 2$. Choose an arbitrary number $T = \text{const.} > 0$ and denote $Q_T = \Omega\times(-T, T)$. Let L_{par} be the quasi-linear elliptic operator of the second order in Q_T, which we define the same way as the operator A_{ell} in (5.89)–(5.91) with the only difference that now its coefficients depend on both x and t, and also the domain Ω is replaced with the domain Q_T. Let $L_{0,\text{par}}$ be the similarly defined principal part of the operator L_{par}; see (5.90). Next, we define the quasilinear parabolic operator as $A_{\text{par}} = \partial_t - L_{\text{par}}$. The principal part of A_{par} is $A_{0,\text{par}} = \partial_t - L_{0,\text{par}}$. Thus, in Q_T,

$$L_{\text{par}}(u) = \sum_{i,j=1}^{n} a_{i,j}(x,t)u_{x_i x_j} + A_1(x,t,\nabla u,u), \tag{5.96}$$

$$A_{\text{par}}(u) = u_t - L_{\text{par}}(u), \tag{5.97}$$

$$A_{0,\text{par}}(u) = u_t - L_{0,\text{par}}u = u_t - \sum_{i,j=1}^{n} a_{i,j}(x,t)u_{x_i x_j}, \tag{5.98}$$

$$a_{i,j} \in C^1(\overline{Q}_T), \tag{5.99}$$

$$A_1(x,t,y) \in C^3\{(x,t,y) : (x,t) \in \overline{Q}_T, y \in \mathbb{R}^{n+1}\}, \tag{5.100}$$

$$\mu_1|\eta|^2 \le \sum_{i,j=1}^{n} a_{i,j}(x,t)\eta_i\eta_j \le \mu_2|\eta|^2, \quad \forall (x,t) \in \overline{Q}_T, \forall \eta = (\eta_1,\dots,\eta_n) \in \mathbb{R}^n. \tag{5.101}$$

Let the domain Ω and the hypersurface $\Gamma \subset \partial\Omega$ be the same as in (5.93). Denote $\Gamma_T = \Gamma \times (-T, T)$. Consider the quasi-linear parabolic equation

$$A_{\text{par}}(u) = u_t - L_{\text{par}}(u) = 0 \quad \text{in } Q_T. \tag{5.102}$$

Cauchy problem with the lateral data for the quasilinear parabolic equation. Assume that conditions (5.96)–(5.101) hold. Find such a function $u \in H^{k_{n+1}}(Q_T)$, which satisfies equation (5.102) and has the following lateral Cauchy data g_0, g_1 at Γ_T:

$$u|_{\Gamma_T} = g_0(\overline{x},t), \quad u_{x_1}|_{\Gamma_T} = g_1(\overline{x},t). \tag{5.103}$$

Let $X > 0$ be a number. Similarly, with Section 4.4, we introduce the CWF $\varphi_\lambda(x,t)$ for the operator $A_{0,\text{par}}$ as

$$\psi(x,t) = x_1 + \frac{|\overline{x}|^2}{2X^2} + \frac{t^2}{2T^2} + \alpha, \quad \varphi_\lambda(x,t) = \exp(\lambda\psi^{-\nu}). \tag{5.104}$$

Here, the number $\nu \ge \nu_0$, where $\nu_0 = \nu_0(\Omega, n, \rho, X, T, L_{0,\text{par}}) > 1$ is a certain number depending only on listed parameters. Theorems 5.2.1, 5.3.1–5.3.5 can be reformulated for problem (5.102), (5.103).

5.8.3 Quasilinear hyperbolic equation

Let $\Omega = \{|x| < R\} \subset \mathbb{R}^n$ and $T = \text{const.} > 0$. Denote

$$Q_T^{\pm} = \Omega \times (-T, T), \quad S_T^{\pm} = \partial\Omega \times (-T, T).$$

Assume that the function A_1 satisfies condition (5.100), where $\overline{Q}_T$ is replaced with $\overline{Q}_T^{\pm}$. Consider the quasi-linear hyperbolic equation in the time cylinder $Q_T^{\pm}$ with the lateral Cauchy data $g_0(x,t)$, $g_1(x,t)$ at S_T,

$$L(u) = c(x)u_{tt} - \Delta u - A_1(x,t,\nabla u,u) = 0 \quad \text{in } Q_T^{\pm}, \tag{5.105}$$

$$u|_{S_T^{\pm}} = g_0(x,t), \quad \partial_n u|_{S_T^{\pm}} = g_1(x,t). \tag{5.106}$$

We impose the same condition on the function $c(x)$ as ones in (4.65) and (4.66),

$$c(x) \in [1,\overline{c}], \quad \text{where } \overline{c} = \text{const.} \geq 1, \tag{5.107}$$

$$c \in C^1(\overline{\Omega}). \tag{5.108}$$

Let the number $d > 0$. Let the point x_0 be such that $x_0 \in \mathbb{R}^3 \setminus \overline{\Omega}$. To apply the Carleman estimate of Theorem 2.5.1, we impose the following analog of condition (3.167):

$$(x - x_0, \nabla c) \geq d > 0, \quad \forall x \in \overline{\Omega}. \tag{5.109}$$

Cauchy problem. Find the solution $u \in H^{k_{n+1}}(Q_T^{\pm})$ of equation (5.105) with the lateral Cauchy data (5.106).

Let the number $\eta \in (0,1)$. Define functions $\psi(x,t)$ and $\varphi_\lambda(x,t)$ as

$$\psi(x,t) = |x - x_0|^2 - \eta t^2, \quad \varphi_\lambda(x,t) = \exp(\lambda\psi(x,t)). \tag{5.110}$$

Denote $L_{0,\text{hyp}}(u) = c(x)u_{tt} - \Delta u$. By Theorem 2.5.1, given conditions (5.107)–(5.109), the Carleman estimate for the operator $L_{0,\text{hyp}}$ holds with the CWF φ_λ from (5.110). Hence, Theorems 5.2.1, 5.3.1–5.3.5 can be reformulated for this Cauchy problem.

5.9 Numerical study

In this section, we study numerically a 1-D analog of the ill-posed Cauchy problem (5.102), (5.103) for the parabolic equation. The numerical study of this section is similar with the one of [146]. There are important differences, however. First, we obtain zero Dirichlet and Neumann boundary conditions on one edge of the interval, where the lateral Cauchy data are given. Second, the specific formulas for the quasi-linear part $S(u)$ of the parabolic operator considered below are different from ones of [146].

5.9.1 The forward problem

Here, $T = 1/2$ and

$$Q_{1/2}^{\pm} = \{(x,t) : x \in (0,1), t \in (-1/2,1/2)\}.$$

We consider the following forward problem:

$$u_t = u_{xx} + S(u) + G(x,t), \quad (x,t) \in Q^{\pm}_{1/2}, \tag{5.111}$$

$$u(x,-1/2) = f(x), \tag{5.112}$$

$$u(0,t) = g(t), \quad u(1,t) = p(t). \tag{5.113}$$

Our specific functions in (5.111)–(5.113) are:

$$S_1(u) = 10\cos(u + x + 2t), \quad S_2(u) = 10\frac{u^2}{1+u^2}, \tag{5.114}$$

$$G(x,t) = 10\sin\left[100\left((x-0.5)^2 + t^2\right)\right], \tag{5.115}$$

$$f(x) = 10(x - x^2), \tag{5.116}$$

$$g(t) = 10\sin\left[10(t-0.5)(t+0.5)\right], \quad p(t) = \sin\left[10(t+0.5)\right]. \tag{5.117}$$

Thus, due to the presence of the multiplier 10 in (5.114), the influence of the nonlinear term $S(u)$ on the solution u of the problems (5.111)–(5.113) is significant.

We use the FDM to solve the forward problems (5.111)–(5.113) numerically. Introduce the uniform mesh in the domain Ω_T,

$$\overline{M} = \left\{(x_i, t_j) : x_i = ih, t_j = -\frac{1}{2} + j\tau, i \in [0,N), j \in [0,M)\right\},$$

where $h = 1/N$ and $\tau = 1/M$ are grid step sizes in x and t directions, respectively. For generic functions $f^{(1)}(x,t)$, $f^{(2)}(x)$, $f^{(3)}(t)$ denote $f^{(1)}_{ij} = f^{(1)}(x_i, t_j)$, $f^{(2)}_i = f^{(2)}(x_i)$, $f^{(3)}_j = f^{(3)}(t_j)$. Let $\varphi_{ij} = S(u_{ij}) + G_{ij}$. We have solved the forward problem (5.111)–(5.113) using the implicit finite difference scheme,

$$\frac{u_{ij+1} - u_{ij}}{\tau} = \frac{1}{h^2}(u_{i-1j+1} - 2u_{ij+1} + u_{i+1j+1}) + \varphi_{ij}, \quad i \in [1, N-1), j \in [0, M-1),$$

$$u_{i0} = f_i, \quad u_{0j} = g_j, \quad u_{Nj} = p_j, \quad i \in [0,N), j \in [0,M).$$

In all our numerical tests, we have used $M = 32$, $N = 128$. Even though these numbers are the same both for the solution of the forward and inverse problems, the “inverse crime” was not committed since we have used noisy data and since we have used the minimization of a functional rather than solving a forward problem again.

Thus, solving the forward problem (5.111)–(5.113) with the input functions (5.114)–(5.117), we have computed the function $q_{\text{comp}}(t)$,

$$u_x(1,t) = q_{\text{comp}}(t), \quad t \in \left(-\frac{1}{2}, \frac{1}{2}\right). \tag{5.118}$$

5.9.2 The ill-posed Cauchy problem and noisy data

Our interest in this section is to solve numerically the following Cauchy problem:

1-D problem with lateral Cauchy data. Suppose that in (5.111)–(5.113) functions $f(x)$ and $g(t)$ are unknown whereas the functions $G(x,t)$, $p(t)$, and $S(u)$ are known. Let in the data simulation process functions F, S, f, g, p are the same as in (5.114)–(5.117). Determine the function $u(x,t)$ in at least a subdomain of the time cylinder $Q^{\pm}_{1/2}$, assuming that the function $q_{\text{comp}}(t)$ in (5.118) is known.

We have introduced 5 % level of random noise in the data. Let $\sigma \in [-1,1]$ be the random variable representing the white noise. Let $p^{(m)} = \max_j |p_j|$ and $q^{(m)} = \max_j |q_{\text{comp},j}|$. Then the noisy data, which we have used, were

$$\widetilde{u}_{N-1j} = p_j + 0.05p^{(m)}\sigma_j, \quad \widetilde{u}_{N-2j} = p_j - h(q_{\text{comp},j} + 0.05q^{(m)}\sigma_j). \tag{5.119}$$

Below we use functions $p'(t)$ and $q'(t)$. We have calculated derivatives $p'(t)$, $q'(t)$ of noisy functions via finite differences. Even though the differentiation of noisy functions is an ill-posed problem, we have not observed instabilities in our case. A more detailed study of this topic is outside of the scope of the current publication.

5.9.3 Specifying the functional $J_{\lambda,\beta}$

We introduce the function $F(x,t)$ as

$$F(x,t) = p(t) + (x-1)q(t).$$

Let $v = u - F$. Then $v(1,t) = v_x(1,t) = 0$ and

$$A(v+F) = v_t - v_{xx} - S(v+F) - G(x,t) + [p'(t) + (x-1)q'(t)].$$

By (5.40), the functional $J_{\lambda,\beta}(v,F)$ becomes

$$J_{\lambda,\beta}(v,F) = \int_{-1/2}^{1/2}\int_0^1 [A(v+F)]^2 \varphi_\lambda^2 dxdt + \beta\|v\|^2_{H^2(Q^{\pm}_{1/2})}. \tag{5.120}$$

Here, we use the same the CWF as the one in [146]

$$\varphi_\lambda(x,t) = \exp[\lambda(x^2 - t^2)]. \tag{5.121}$$

The Carleman estimate for the operator $A_0 = c(x,t)\partial_t - \partial_x^2$ with an appropriate strictly positive function $c(x,t)$ is proven in [146] and, in a more general form, in Theorem 9.4.1 in Section 9.4. The reason of the replacement of the CWF (5.104) with the CWF (5.121) this is that the rate of change of the CWF (5.104) is too large due to the presence of two large parameters λ and ν in (5.104). We have observed in our computations that the accuracy of results does not change much for β varying in a large interval. In all our numerical experiments below, $\beta = 0.00063$. The norm $\|v\|_{H^2(Q^{\pm}_{1/2})}$ is taken instead of $\|v\|_{H^3(Q^{\pm}_{1/2})}$ due to the convenience of computations. Note that since we do not use too many grid points when discretizing the functional $J_{\lambda,\beta}(v,F)$, then these two norms are basically equivalent in our computations, since all norms are equivalent in a finite dimensional space.

5.9.4 Minimization of $J_{\lambda,\beta}(v,F)$

To minimize the functional (5.120), we have attempted first to use the gradient projection method, as it was done in the above theoretical part. However, we have observed in our computations that just the conjugate gradient method (GCM) with the starting point $v_0 \equiv 0$ works well and much more rapidly. The latter is true for all our numerical studies of the convexification method. So, our results below are obtained via the GCM. We have written the functional $J_{\lambda,\beta}(v,F)$ in the discrete form $\bar{J}_{\lambda,\beta}(v,F)$ using finite differences. Next, we have minimized the functional $\bar{J}_{\lambda,\beta}(v,F)$ with respect to the values v_{ij} of the discrete function v at the grid points. Hence, we have calculated derivatives $\partial_{v_{ij}}\bar{J}_{\lambda,\beta}(v,F)$ via explicit formulas. The method of the calculation of these derivatives is described in [146].

Normally, for a quadratic functional the GCM reaches the minimum of this functional after $M \cdot N$ gradient steps with the automatic step choice. However, our computational experience tells us that we can obtain a better accuracy if using a small constant step in the GCM and a large number of iterations. Thus, we have used the step size $\gamma = 10^{-8}$ and 10,000 iterations of the GCM. It took 0.5 minutes of CPU Intel Core i7 to do these iterations.

5.9.5 Results

Let $v(x,t)$ be the numerical solution of the forward problem (5.111)–(5.113). Let $v_{\lambda\beta}(x,t)$ be the minimizer of the functional $\bar{J}_{\lambda,\beta}(v,F)$, which we have found via the GCM. Of course, $v(x,t)$ and $v_{\lambda\beta}(x,t)$ here are discrete functions defined on the above grid and norms used below are discrete norms. Recall that $u(x,t) = v(x,t) + F(x,t)$. Hence, denote $u_{\lambda\beta}(x,t) = v_{\lambda\beta}(x,t) + F(x,t)$. For each x of our grid, we define the “line error” $E(x)$ as

$$E(x) = \frac{\|u_{\lambda\beta}(x,t) - u(x,t)\|_{L_2(-1/2,1/2)}}{\|u(x,t)\|_{L_2(-1/2,1/2)}}. \tag{5.122}$$

We evaluate how the line error changes with the change of x, that is, how the computational error changes when the point x moves away from the edge $x = 1$ where the lateral Cauchy data are given. Naturally, it is anticipated that the function $E(x)$ should be decreasing.

Remark 5.9.1. It is clear, intuitively at least, that the further a point $x \in (0,1)$ is from the point $x = 1$ where the lateral Cauchy data are given, the less accuracy of solution at this point one should anticipate. So, we observe in graphs of line errors on Figures 5.1(a)–5.4(a) that the accuracy of the calculated solutions for $x \in (0, 0.6)$ is not as good as this accuracy for $x \in [0.6, 1]$. This is why we graph below only line errors and functions $u_{\lambda\beta}(0.6, t)$, superimposed with $u(0.6, t)$.

In the case of Figures 5.1 and 5.2, the starting function for iterations of the GCM was $v_0 \equiv 0$. We have tested three values of the parameter $\lambda : \lambda = 0, 1, 3$ in (5.120). We

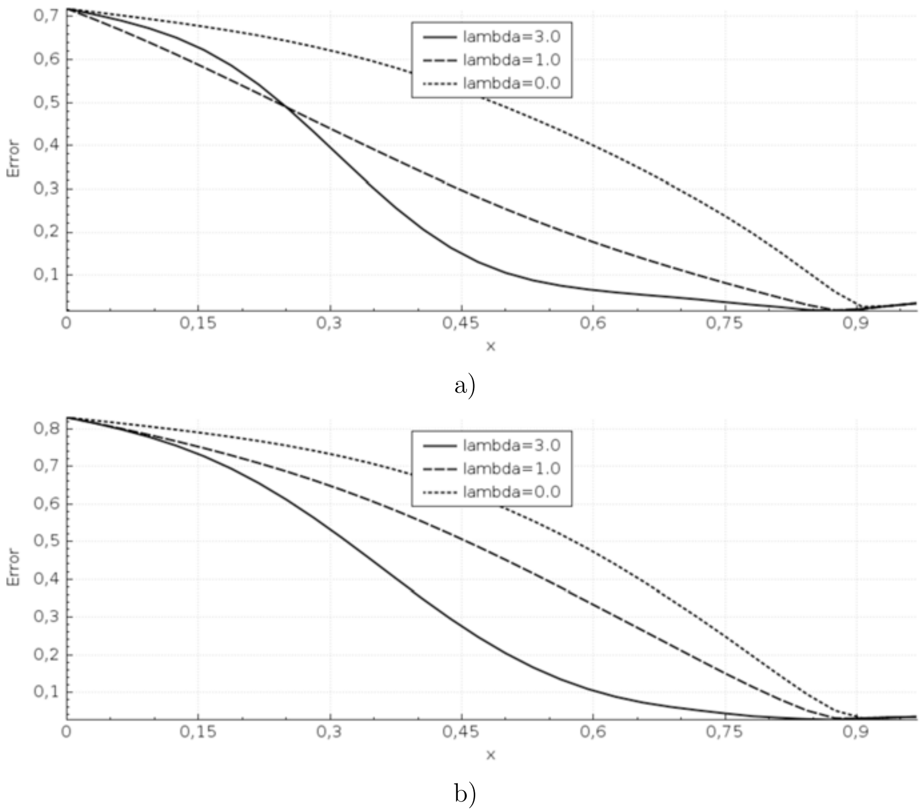
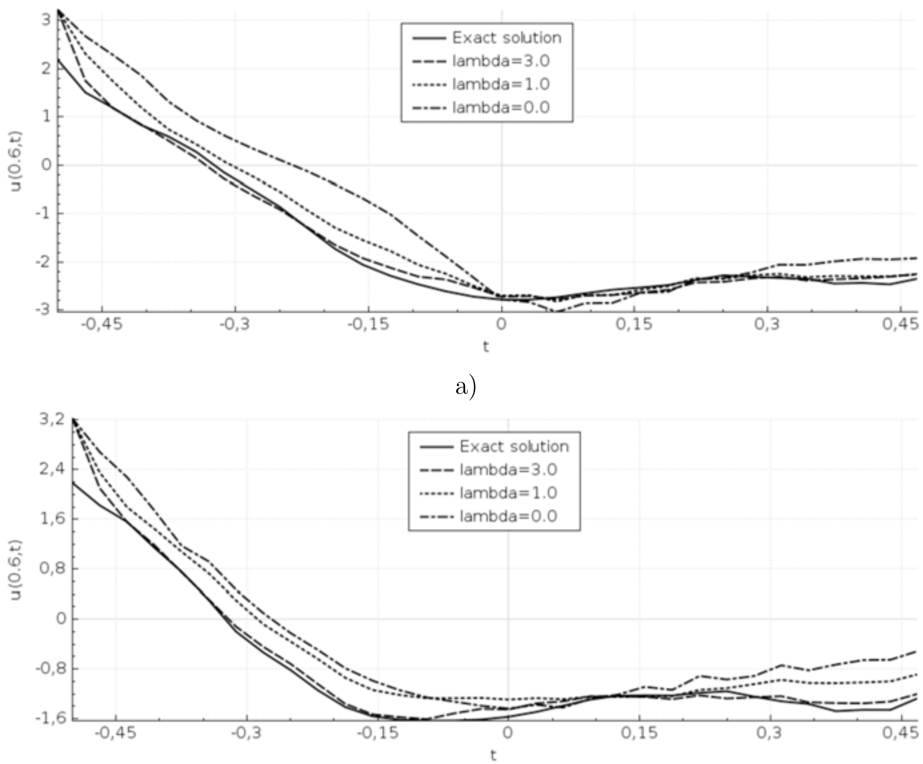


Figure 5.1: Distribution of error along the x-axis. (a) $\lambda = 0, 1, 3$ and $S(u) = \cos(u + x + 2t)$. (b) $\lambda = 0, 1, 3$ and $S(u) = u^2/(1 + u^2)$. Thus, the presence of the CWF in the functional (5.120) significantly improves the accuracy of the solution. One can observe that a rather accurate reconstruction is obtained on the interval $[0.6, 1]$; also, see Remark 5.9.1.



a)

b)

Figure 5.2: Superimposed graphs of functions $u_{0\beta}(0.6,t)$, $u_{1\beta}(0.6,t)$, $u_{3\beta}(0.6,t)$, and $u(0.6,t)$. (a) $S(u) = \cos(u + x + 2t)$. (b) $S(u) = u^2/(1 + u^2)$. Observe that the presence of the CWF with $\lambda = 3$ significantly improves the accuracy of the solution.

have found that $\lambda = 3$ is the best choice for those problems which we have studied. This is also clear from Figures 5.1. Note that the case $\lambda = 0$ provides a poor accuracy.

As one can see on Figures 5.1, the line error at $x = 0.6$ is between about 6 % and 10 % for $\lambda = 3$. Thus, we superimpose graphs of functions $u_{\lambda\beta}(0.6,t)$ with graphs of functions $u(0.6,t)$ (see Remark 5.9.1). Corresponding graphs are displayed on Figures 5.2. One can observe again that the computational accuracy with $\lambda = 3$ is the best and that the accuracy with $\lambda = 0$ is poor. Thus, we observe again that the presence of the CWF in the functional (5.120) significantly improves the accuracy of the solution. On the other hand, the accuracy at $t \approx \pm 1/2$ is not good on Figures 5.2. We explain this by the fact that Theorem 5.3.5 guarantees a good accuracy only in a subdomain $\Omega_{\alpha+2\varepsilon}$ of the domain Ω rather than in the entire domain Ω. The latter can be reformulated for our specific domain $Q_{1/2}$ [146].

To see how the starting function of the GCM affects the accuracy of our results, we took $S(u) = S_1(u)$ and have tested three starting functions for the GCM: $v_0(x,t) = 0$,

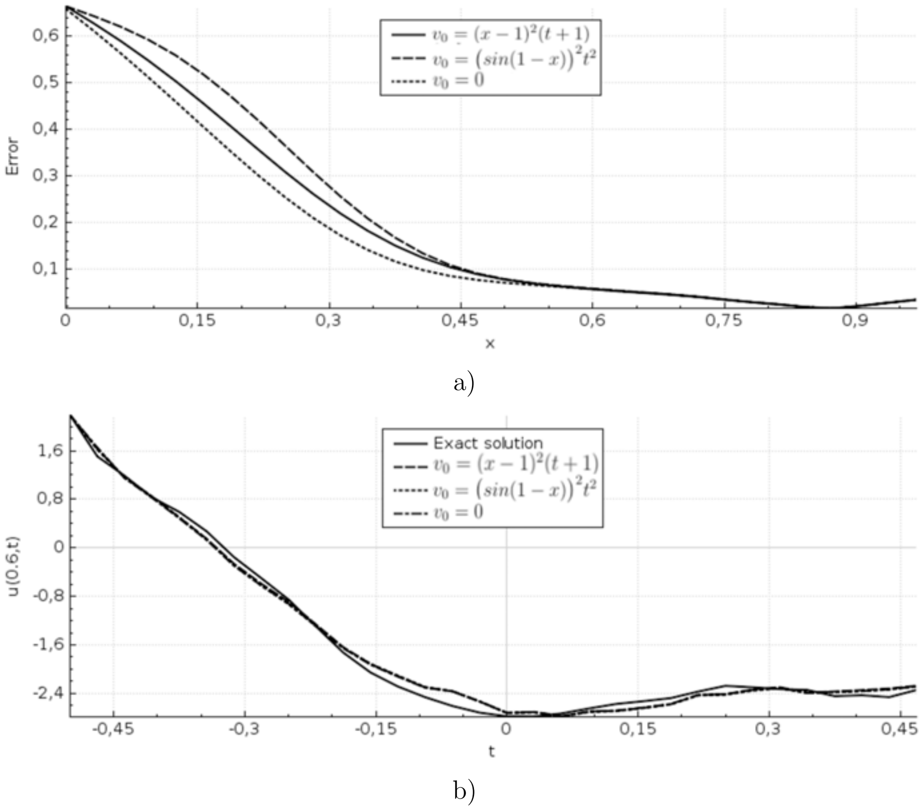


Figure 5.3: The influence of the choice of the starting function v_0 of the GCM. We have tested three starting functions: $v_0 = 0$, $v_0 = (x-1)^2(t+1)$ and $v_0 = (\sin(1-x))^2t^2$ We took $\lambda = 3$, $S(u) = 10\cos(u + x + 2t)$. (a) Superimposed line errors. (b) Superimposed functions $u(0.6, t)$ and $u_{3\beta}(0.6, t)$. These tests demonstrate that for $x \in [0.6, 1]$ our solution depends only very insignificantly from the choice of the starting function v_0 of the GCM: just as it was predicted by Theorems 5.3.3 and 5.3.5.

$v_0(x,t) = (x-1)^2(t+1)$ and $v_0(x,t) = (\sin(x-1))^2t^2$. Hence, for any of these three functions $v_0(x,t)$ we have $v_0(1,t) = \partial_x v_0(1,t) = 0$. Graphs of Figure 5.3(a) displays superimposed line errors and Figure 5.3(b) displays functions $u_{3\beta}(0.6,t)$ and $u(0.6,t)$ for these three cases (see Remark 5.9.1). One can see that for $x \in [0.6,1]$ results depend only very insignificantly on the starting point of the GCM: just as it was predicted by Theorems 5.3.3 and 5.3.5; also, see Remark 5.9.1.

It is interesting to see how the presence of the CWF affects the linear case. In this case, the above method with $\lambda = 0$ becomes the quasi-reversibility method; see Chapter 4. So, we have tested the case when $S(u) \equiv 0$ in (5.111), while functions $G(x,t)$, $f(x)$, $g(t)$ and $p(t)$ are the same as in (5.115)–(5.117). Results for $\lambda = 0, 1, 3$ are presented on Figures 5.4. One can observe that even in the linear case the presence of the CWF significantly improves the computational accuracy for $x \in [0.6, 1]$.

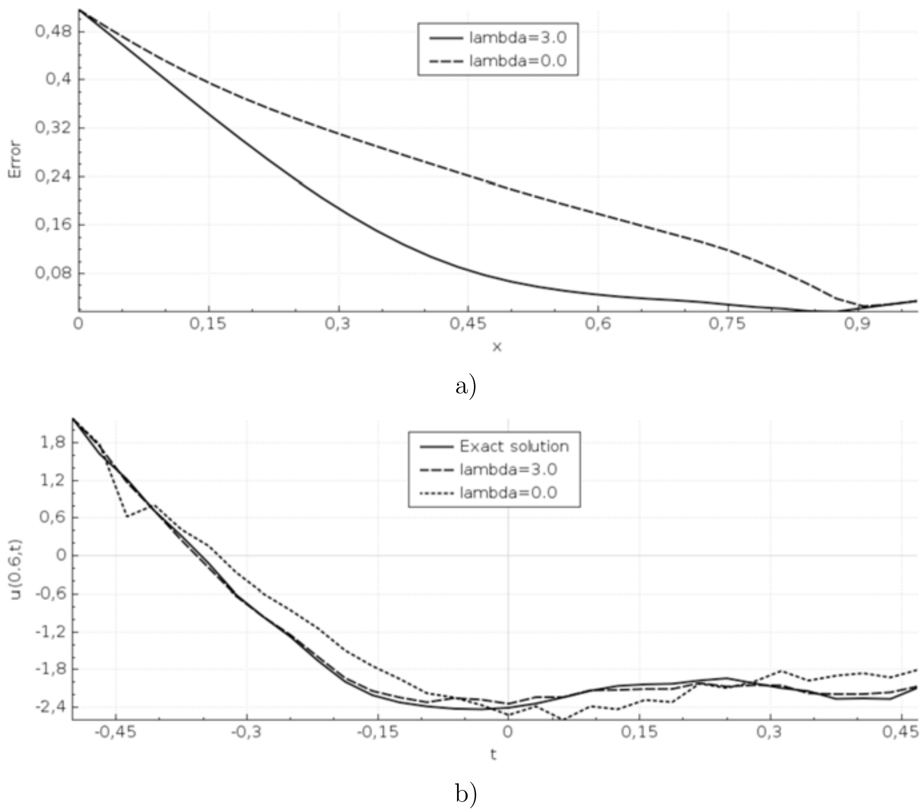


Figure 5.4: The linear case: when $S(u) \equiv 0$ in (5.111). (a) Line errors. (b) Functions $u(0.6,t)$ and $u_{3\beta}(0.6,t)$. One can observe that the presence of the CWF significantly improves the accuracy of the solution for $x \in [0.6, 1]$ even in the linear case.

5.10 Summary

In this chapter, we have presented some facts of the convex analysis in Section 5.2. Next, using, as an example, a general ill-posed Cauchy problem for a quasi-linear PDE of the second order, we have shown that these facts ensure the existence of the minimizer of our weighted Tikhonov-like functional with the CWF in it on any closed ball in a reasonable Hilbert space. This functional is strictly convex on that ball. The strict convexity is due to the presence of the CWF. Next, we have specified PDEs of the second order for which this construction works.

In addition, we have presented some numerical results for the side Cauchy problem for a 1-D quasilinear parabolic PDE. These results indicate that the presence of the CWF significantly improves the accuracy of the solution. Furthermore, this is also true even in the linear case. It was also demonstrated numerically that for $x \in [0.6, 1]$ our resulting solution depends on the starting function for the GCM only very insignificantly: just as it is predicted by our theory.

6 A special orthonormal basis in $L_2(a, b)$ for the convexification for CIPs without the initial conditions—restricted Dirichlet-to-Neumann map

In this chapter, we follow the work of Klibanov [136]. In addition, the Subsection 6.2.1 uses the material of [9, 132, 134, 135, 165]. Permissions for republishing are obtained from publishers.

6.1 Introduction

A special orthonormal basis in the $L_2(a, b)$ space introduced in this chapter plays the *pivotal* role in the convexification method for those CIPs, for which initial conditions are not given. Some important examples include electrical impedance tomography (Chapter 7 and [149]), CIP for the Helmholtz equation with the moving source (Chapter 10 and [115–117]), CIPs for the Helmholtz equation with a single source and varying frequency [145], travel time tomography problem (Chapter 11 and [137, 138, 148]) inversion of the Radon transform with incomplete data [156], numerical solution of the Lavrent'ev linear integral equation [152], and the inverse source problem for the full radiative transfer equation [236].

The conventional Dirichlet-to-Neumann map (DN) data for a Coefficient Inverse Problem (CIP) can be generated, at least sometimes, by the point source running along a hypersurface; see pages 10–14 in [165] for DN and [89] for the Neumann-to-Dirichlet map data. We define "restricted DN data" for a CIP as the ones, in which Dirichlet and Neumann boundary data are generated by a point source running along an interval of a straight line. These data are non-overdetermined in the n-D case with $n \geq 2$. On the other hand, the conventional DN data are overdetermined for $n \geq 3$; see Definition 1.2.2.

We show in this chapter how to construct the convexification for the restricted DN using that special orthonormal basis. The convexification is a *globally* convergent numerical method; see Chapter 1 for the definitions of locally and globally convergent numerical methods. In fact, we present here a *general* concept of constructing of the convexification for CIPs with restricted DN data. This concept also covers both Hölder stability and uniqueness results for the CIPs we consider. Our construction is independent on a specific PDE operator: It is the same for those PDEs of the second order, which admit Carleman estimates. In particular, it works for three main types of PDEs of the second order: elliptic, parabolic, and hyperbolic ones. The Dirichlet and Neumann data in elliptic and parabolic cases can be given on a part of the boundary.

The price we pay for our concept is a well acceptable one in the numerical analysis: We truncate a Fourier-like series with respect to that orthonormal basis. Next, to find spatially dependent coefficients of that truncated series, we construct a weighted

https://doi.org/10.1515/9783110745481-006

globally strictly convex Tikhonov-like functional with the Carleman Weight Function (CWF) in it. This is the function, which is involved in the Carleman estimate for the corresponding PDE operator. Also, just as in Theorem 5.3.4, we establish the global convergence of the gradient projection method to the exact solution under the natural condition that the noise in the data tends to zero.

As to the DN data, a very substantial number of works have been published. Since this book is not a survey of DN, we refer to only a very few of them for brevity, and the reader can find other references in these publications. Global uniqueness theorems for the elliptic case, that is, for the Calderon problem, were obtained in [203, 211, 241]. Some reconstruction procedures can be found in [175, 203, 211–213]. In the reconstruction procedure of [213], a certain infinite matrix is truncated, which is philosophically close to our truncation of that Fourier-like series. We refer to, for example, [2, 85, 86, 106] for some numerical studies of DN. In [26] and [111], reconstruction procedures for DN for hyperbolic PDEs were developed, and they were computationally tested in [27] and [71].

We point out that since our goal here is to present a new numerical concept for brevity, we are not concerned in this chapter with some issues related to solutions of forward problems, since they can be discussed in later publications. These issues are: the minimal smoothness assumptions, existence and uniqueness of the solutions of the forward problems under considerations, the positivity of those solutions, and also the continuous differentiability of those solutions with respect to the position of the point source.

6.2 A CIP with the restricted DN data

6.2.1 The Carleman estimate

Below all functions are real valued, unless stated otherwise. The material of Section 6.2.1 is a somewhat modified material of Section 2.1.2 of [165]. Below $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Also, below $\alpha = (\alpha_1, \dots, \alpha_n)$ is the multiindex with integer coordinates $\alpha_i \geq 0$ and with $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Consider a general partial differential operator of the second order

$$A(x,u) = \sum_{|\alpha|\leq 2} a_\alpha(x) D_x^\alpha u = A_0(x,u) + A_1(x,u), \quad x \in \mathbb{R}^n, \tag{6.1}$$

$$A_0(x,u) = \sum_{|\alpha|=2} a_\alpha(x) D_x^\alpha u, \quad A_1(x,u) = \sum_{|\alpha|=1} a_\alpha(x) D_x^\alpha u + a_0(x)u. \tag{6.2}$$

Thus, $A_0(x,u)$ is the principal part of the operator $A(x,u)$ and the operator $A_1(x,u)$ contains lower order terms. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a piecewise smooth boundary. Let $Z > 0$ be a given number. We assume that coefficients

$$a_\alpha(x) = \widehat{a}_\alpha = \text{const.} \quad \text{for } x \notin \Omega \text{ and for all } \alpha \text{ with } |\alpha| \leq 2, \tag{6.3}$$

$$a_\alpha \in C^1(\mathbb{R}^n) \quad \text{for } |\alpha| \leq 2, \tag{6.4}$$

$$\|a_\alpha\|_{C^1(\overline{\Omega})} \leq Z \quad \text{for } |\alpha| \leq 1. \tag{6.5}$$

Let $\Gamma \in C^2$, $\Gamma \subseteq \partial\Omega$ be a part of the boundary of the domain Ω. We assume that any part of Γ is not a characteristic surface of the operator $A_0(x,u)$. Let the function $\xi \in C^\infty(\overline{\Omega})$ and $|\nabla\xi| \neq 0$ in $\overline{\Omega}$. For a number $d > 0$, denote

$$\xi_d = \{x \in \Omega : \xi(x) = d\}, \quad \Omega_d = \{x \in \Omega : \xi(x) > d\}. \tag{6.6}$$

We assume below that $\Omega_d \neq \varnothing$ and that $(\overline{\Omega}_d \cap \partial\Omega) = \Gamma_d \subseteq \Gamma$. Hence,

$$\Gamma_d = \{x \in \Gamma : \xi(x) > d\}. \tag{6.7}$$

Hence, the boundary of the domain Ω_d consists of two parts,

$$\partial\Omega_d = \partial_1\Omega_d \cup \partial_2\Omega_d, \quad \partial_1\Omega_d = \xi_d, \partial_2\Omega_d = \Gamma_d. \tag{6.8}$$

We assume below that $\partial\Omega_d$ is piecewise smooth. Below $C_1 = C_1(A_0, \Omega_d) > 0$ denotes different constants depending only on the operator A_0 and the domain Ω. Let $\lambda > 1$ be a large parameter. Consider the function $\varphi_\lambda(x)$,

$$\varphi_\lambda(x) = \exp(\lambda\xi(x)). \tag{6.9}$$

It follows from (6.6)–(6.8) that

$$\min_{\overline{\Omega}_d} \varphi_\lambda(x) = \varphi_\lambda(x)|_{\xi_d} = e^{\lambda d}, \tag{6.10}$$

$$m = \max_{\overline{\Omega}_d} \xi(x) \Rightarrow \max_{\overline{\Omega}_d} \varphi_\lambda(x) = e^{\lambda m}. \tag{6.11}$$

Definition 6.2.1. We say that the operator A_0 with its coefficients $a_\alpha(x)$ satisfying conditions (6.2), (6.4) admits the pointwise Carleman estimate in the domain Ω_d with the CWF $\varphi_\lambda(x)$ if there exist constants $\lambda_0 = \lambda_0(A_0, \Omega_d) > 1$, $C_1 = C_1(A_0, \Omega_d) > 0$, depending only on listed parameters, such that the following estimates hold:

$$(A_0 u)^2 \varphi_\lambda^2(x) \geq C_1\lambda(\nabla u)^2\varphi_\lambda^2(x) + C_1\lambda^3 u^2 \varphi_\lambda^2(x) + \operatorname{div} U, \tag{6.12}$$

$$|U(x)| \leq C_1\lambda^3[(\nabla u)^2 + u^2]\varphi_\lambda^2(x), \tag{6.13}$$

$$\forall\lambda \geq \lambda_0, \forall x \in \overline{\Omega_d}, \forall u \in C^2(\overline{\Omega_d}). \tag{6.14}$$

6.2.2 Statement of the problem

Denote $\overline{x} = (x_2, \dots, x_n) \in \mathbb{R}^{n-1}$. Below $\overline{x}^0 \in \mathbb{R}^{n-1}$ is a fixed point of $\mathbb{R}^{n-1}$ and $x_0 \in [0,1]$ is a varying parameter. Consider an interval I of a straight line such that

$$I = \{x = (x_0, \overline{x}^0) : x_0 \in (0,1)\}, \tag{6.15}$$

$$I \cap \overline{\Omega} = \varnothing. \tag{6.16}$$

Consider the following equation:

$$A(u) = -\delta(x_1 - x_0, \overline{x} - \overline{x}^0), \quad x \in \mathbb{R}^n, \forall x_0 \in [0,1], \tag{6.17}$$

where $u = u(x, x_0)$ is a distribution with respect to x. Since we do not impose any condition at the infinity on the distribution u, equation (6.17) might have many solutions or even none. Suppose that it has a solution, which we still denote as $u(x, x_0)$. We assume that the following conditions are valid for this solution:

Condition 1. For each $x_0 \in [0,1]$, the function $u(x, x_0) \in C^2(\overline{\Omega})$.

Condition 2. For each $x \in \overline{\Omega}$, the functions $D_x^\alpha u(x, x_0)$, are differentiable with respect to $x_0 \in (0,1)$ and functions $\partial_{x_0}^k D_x^\alpha u(x, x_0) \in C(\overline{\Omega} \times [0,1])$ for $k = 0,1$; $|\alpha| \le 2$.

Condition 3. $u(x, x_0) \ge \beta = \text{const.} > 0$, $\forall (x, x_0) \in \overline{\Omega} \times [0,1]$; see Remark 6.2.1.

Condition 4. The following Dirichlet and Neumann boundary conditions are given for the function $u(x, x_0)$:

$$u(x, x_0)|_{x \in \Gamma, x_0 \in [0,1]} = g_0(x, x_0), \quad \partial_n u(x, x_0)|_{x \in \Gamma, x_0 \in [0,1]} = g_1(x, x_0), \tag{6.18}$$

where $g_0(x, x_0)$ and $g_1(x, x_0)$ are two given functions of $(x, x_0) \in \Gamma \times [0,1]$.

We call the Dirichlet and Neumann boundary data (6.18) "restricted DN data."

Coefficient Inverse Problem 1 (CIP 1). Suppose that for each value $x_0 \in [0,1]$ of the parameter x_0 there exists a distribution $u(x, x_0)$ satisfying equation (6.17) and Conditions 1–4. Determine the coefficient $a_0(x)$ in (6.2) from functions $g_0(x, x_0)$ and $g_1(x, x_0)$ in (6.18).

Remark 6.2.1. Thus, (6.17) and (6.18) mean that the source $(x_0, \overline{x}^0)$ runs along the interval I. In the cases of elliptic and parabolic PDEs Condition 3 can often be established via the maximum principle [79, 80].

Sometimes it is hard to prove the validity of Conditions 1–3 in the case when the fundamental solution (6.17) of the operator A is considered. Hence, to avoid dealing with singularities, we replace the δ-function in (6.17) with a delta-like function. Let $\varepsilon > 0$ be a sufficiently small number. Let the functions $f \in C^\infty(\mathbb{R})$ and $\chi(\overline{x}) \in C^\infty(\mathbb{R}^{n-1})$ be such that $f(0)\chi(0) \ne 0$ and also $f(z) = 0$ for $|z| > \varepsilon$ as well as $\chi(y) = 0$ for $y \in \{|y| > \varepsilon\}$. Let

$$I_\varepsilon = \{x \in \mathbb{R}^n : \text{dist}(x, I) < \varepsilon\},$$

where $\text{dist}(x, I)$ is the Hausdorff distance between the point x and the interval I. Let $G \subset \mathbb{R}^n$ be a bounded domain with its boundary $\partial G \in C^1$ and such that $\Omega \subset G$, $\partial\Omega \cap \partial G = \varnothing$. We assume that $I_\varepsilon \subset (G \setminus \overline{\Omega})$.

We now replace (6.17) with

$$A(u) = f(x_1 - x_0)\chi(\overline{x} - \overline{x}^0), \quad \forall x_0 \in [0,1], \tag{6.19}$$

$$\widetilde{u}|_{x \in \partial G} = 0, \quad \forall x_0 \in [0,1]. \tag{6.20}$$

Coefficient Inverse Problem 2 (CIP 2). Assume that the function $u(x, x_0)$ satisfies Conditions 1–4, (6.19) and (6.20). Determine the coefficient $a_0(x)$ in (6.2) from functions $g_0(x, x_0)$ and $g_1(x, x_0)$ in (6.18).

Both CIP 1 and CIP 2 are non-overdetermined. Indeed the number n of free variables in the data (6.18) coincides with the number of free variables in the unknown coefficient. Since our method of the numerical solution of CIP 2 is exactly the same as the one of CIP 1, we consider below CIP 1 in most cases.

6.2.3 A special orthonormal basis in $L_2(0,1)$

We need to construct such an orthonormal basis in the space $L_2(0,1)$ of functions depending on x_0 that the first derivative with respect to x_0 of any element of this basis is not identically zero. In addition, this derivative should be a linear combination of a finite number of elements of this basis. Neither the basis of trigonometric functions nor the basis of standard orthonormal polynomials are not suitable for this goal. Therefore, we construct a new basis. Our basis is similar with Laguerre functions, which, however, form an orthonormal basis in $L_2(0,\infty)$ rather than in $L_2(0,1)$.

For $x_0 \in (0,1)$, consider the set of functions $\{x_0^k e^{x_0}\}_{k=0}^{\infty}$. Clearly, these functions are linearly independent and form a complete set in $L_2(0,1)$. We apply the classical Gram–Schmidt orthonormalization procedure to this set. We start from e^{x_0}. Then we take $x_0 e^{x_0}$, then $x_0^2 e^{x_0}$, etc. As a result, we obtain an orthonormal basis in $L_2(0,1)$, which consists of functions $\{P_m(x_0)e^{x_0}\}_{m=0}^{\infty} = \{\psi_m(x_0)\}_{m=0}^{\infty}$, where $P_m(x_0)$ is a polynomial of the degree m. Denote $[\,,\,]$ the scalar product in $L_2(0,1)$. Let $Q_s(x_0)$ be an arbitrary polynomial of the degree $s \geq 0$. By the construction of functions $\psi_m(x_0)$, there exists numbers $b_j = b_j(Q_s)$ such that

$$Q_s(x_0) = \sum_{j=0}^{s} b_j(Q_s)P_j(x_0). \tag{6.21}$$

Theorem 6.2.1. *We have*

$$a_{mk} = [\psi_k', \psi_m] = \begin{cases} 1 & \text{if } k = m, \\ 0 & \text{if } k < m. \end{cases} \tag{6.22}$$

Let $N > 1$ be an integer. Consider the $N \times N$ matrix $M_N = (a_{mk})_{(k,m)=(0,0)}^{(N-1,N-1)}$. Then (6.22) *implies that* $\det(M_N) = 1$, *which means that there exists the inverse matrix M_N^{-1}.*

Proof. We have $\psi_k'(x_0) = P_k(x_0)e^{x_0} + P_k'(x_0)e^{x_0} = \psi_k(x_0) + P_k'(x_0)e^{x_0}$. Since the degree of the polynomial $P_k'(x_0)$ is less than k, then (6.21) implies that the function $P_k'(x_0)e^{x_0}$ is a linear combination of functions $\psi_j(x_0)$ with $j \le k-1$. Hence,

$$\psi_k'(x_0) = \psi_k(x_0) + \sum_{j=0}^{k-1} b_{jk}\psi_j(x_0). \tag{6.23}$$

First, let $m = k$. Since $[\psi_j, \psi_m] = 0$ for $j < m$, then (6.23) implies that $[\psi_m'(x_0), \psi_m(x_0)] = 1$. Consider now the case $m > k$. Then we obtain similarly from (6.23) that $[\psi_k'(x_0), \psi_m(x_0)] = 0$. Thus, (6.22) is established. □

6.3 An ill-posed problem for a coupled system of quasilinear PDEs

If we say below that a certain vector function belongs to a functional space, then this means that each component of this function belongs to this space. The norm of that vector function in that space is defined as the square root of the sum of squares of norms of its components.

It follows from Condition 3 of Section 6.2.2 that we can consider the function $v(x,x_0) = \ln u(x,x_0)$ for $x \in \overline{\Omega}$. Substituting $u = e^v$ in (6.19) for $x \in \Omega$ and using (6.1)–(6.5), (6.15), (6.16), and (6.18), we obtain

$$A_0(x,v) + F_1(x,\nabla v) = -a_0(x), \quad x \in \Omega, x_0 \in [0,1], \tag{6.24}$$

$$v(x,x_0)|_{x\in\Gamma,x_0\in[0,1]} = \widetilde{g}_0(x,x_0), \quad \partial_n v(x,x_0)|_{x\in\Gamma,x_0\in[0,1]} = \widetilde{g}_1(x,x_0), \tag{6.25}$$

$$\widetilde{g}_0(x,x_0) = \ln g_0(x,x_0), \quad \widetilde{g}_1(x,x_0) = g_1(x,x_0)/g_0(x,x_0),$$

where the function $F_1 \in C^1(\mathbb{R}^{2n})$, and it is quadratic with respect to derivatives $\partial_{x_k} v$. Denote $v_{x_0}(x,x_0) = \partial_{x_0} v(x,x_0)$. Differentiate both sides of (6.24) with respect to x_0. Since $\partial_{x_0}(a_0(x)) \equiv 0$, then using (6.25), we obtain

$$A_0(x,v_{x_0}) + F_2(x,\nabla v,\nabla v_{x_0}) = 0, \quad x \in \Omega, x_0 \in [0,1], \tag{6.26}$$

$$v_{x_0}(x,x_0)|_{x\in\Gamma,x_0\in[0,1]} = \partial_{x_0}\widetilde{g}_0(x,x_0), \quad \partial_n v_{x_0}(x,x_0)|_{x\in\Gamma,x_0\in[0,1]} = \partial_{x_0}\widetilde{g}_1(x,x_0), \tag{6.27}$$

where the function $F_2 \in C^1(\mathbb{R}^{3n})$ is quadratic with respect to derivatives $\partial_{x_k} v$, $\partial_{x_k} v_{x_0}$.

It follows from Conditions 1–3 of Section 6.2.2 that, for each $x \in \overline{\Omega}$, the function $v(x,x_0)$ can be represented as a Fourier-like series with respect to the orthonormal basis $\{\psi_m(x_0)\}_{m=0}^{\infty}$. Coefficients of this series depend on x. We, however, assume that the function $v(x,x_0)$ can be represented as a truncated series,

$$v(x,x_0) = \sum_{k=0}^{N-1} v_k(x)\psi_k(x_0), \quad \forall x \in \overline{\Omega}, \forall x_0 \in [0,1], \tag{6.28}$$

where coefficients $v_k(x) \in C^2(\overline{\Omega})$ and $N \geq 2$ is an integer of ones choice. Substituting (6.28) in (6.26), we obtain

$$\sum_{k=0}^{N-1} A_0(x, v_k)\psi_k'(x_0) + F_2\left(x, \sum_{m=0}^{N-1} \nabla v_m(x)\psi_m(x_0), \sum_{k=0}^{N-1} \nabla v_k(x)\psi_k'(x_0)\right) = 0, \tag{6.29}$$

where $x \in \Omega$, $x_0 \in [0,1]$. Let the integer $m \in [0, N-1]$. Multiply both sides of (6.29) by the function $\psi_m(x_0)$ and integrate with respect to $x_0 \in (0,1)$. We obtain

$$\begin{aligned}\sum_{k=0}^{N-1} A_0(x, v_k)[\psi_k'(x_0), \psi_m(x_0)]\\ = -\left[F_2\left(x, \sum_{k=0}^{N-1} \nabla v_k(x)\psi_k(x_0), \sum_{k=0}^{N-1} \nabla v_k(x)\psi_k'(x_0)\right), \psi_m(x_0)\right],\end{aligned} \tag{6.30}$$

where $x \in \Omega$, $m \in [0, N-1]$. Denote

$$V(x) = \left(v_0(x), \dots, v_{N-1}(x)\right)^T, \quad A_0(x, V) = \left(A_0(x, v_0), \dots, A_0(x, v_{N-1})\right)^T. \tag{6.31}$$

Also, let $F(x, \nabla V) = (F_{2,0}(x, \nabla V), \dots, F_{2,N-1}(x, \nabla V))^T$ be the vector of right-hand sides of equations (6.30). Then (6.30) can be rewritten as

$$M_N A_0(x, V) = F(x, \nabla V), \tag{6.32}$$

where M_N is the matrix of Theorem 6.2.1. Applying Theorem 6.2.1 to (6.32) and denoting $P(x, \nabla V) = M_N^{-1} F(x, \nabla V)$, we obtain

$$A_0(x, V) - P(x, \nabla V) = 0, \tag{6.33}$$

$$V|_\Gamma = p_0(x), \quad \partial_n V|_\Gamma = p_1(x), \tag{6.34}$$

where vector functions $p_0(x)$ and $p_1(x)$ are obtained from functions $\partial_{x_0}\tilde{g}_0(x, x_0)$ and $\partial_{x_0}\tilde{g}_1(x, x_0)$ of (6.27) in an obviously similar manner, the N-dimensional vector function $P \in C^1(\mathbb{R}^{s_1})$, $s_1 = n(N+1)$, and each component of P is a quadratic function of the first derivatives $\partial_{x_k} v_i(x)$, where $k = 1, \dots, n$, and $i = 0, \dots, N-1$.

Remark 6.3. In fact, the truncated series (6.28) represents our approximate mathematical model; see Remarks 7.3 for more details.

Equalities (6.33), (6.34) form an ill-posed problem for the coupled system of quasilinear equations. A similar problem was considered in Chapter 5 for the case of a single quasi-linear PDE. Thus, we proceed below similarly with Chapter 5. It follows from (6.24), (6.28), and (6.31) that, given the vector function $V(x)$, we can find the unknown coefficient $a_0(x)$. However, since only u and ∇u are involved in the Carleman estimate (6.12), while the second derivatives $u_{x_i x_j}$ are not involved, we formulate all theorems below in terms of the vector functions V, ∇V rather than in terms of the unknown coefficient $a_0(x)$. At the same time, it is well known that in the case of parabolic and elliptic operators (unlike hyperbolic ones) derivatives involved in their principal parts can be incorporated in corresponding Carleman estimates; see, for example, Theorem 2.5 in [132]. Hence, the Hölder stability result of Theorem 6.3.1 as well as the global convergence result (Theorem 6.4.4 below) can be reformulated in terms of $a_0(x)$ in these cases. We are not doing this here for brevity.

Theorem 6.3.1 (Hölder stability and uniqueness). *Suppose that there exist two vector functions $V^{(1)}, V^{(2)} \in C^2(\overline{\Omega})$ satisfying equation* (6.33) *and with two pairs of boundary conditions* (6.34)*, $V^{(i)}|_\Gamma = p_0^{(i)}(x)$ and $\partial_n V^{(i)}|_\Gamma = p_1^{(i)}(x)$, $i = 1,2$. Let $K > 0$ be such a number that $\|V^{(i)}\|_{C^1(\overline{\Omega})} \le K$. Let $Z > 0$ be the number defined in* (6.5)*. Let $\sigma \in (0,1)$ be the level of the error in the data* (6.34)*, that is,*

$$\|p_0^{(1)} - p_0^{(2)}\|_{H^1(\Gamma)} \le \sigma, \quad \|p_1^{(1)} - p_1^{(2)}\|_{L_2(\Gamma)} \le \sigma. \tag{6.35}$$

Choose a number $c > 0$ such that $\Omega_{d+c} \neq \varnothing$. Then there exists a sufficiently small constant $\sigma_0 = \sigma_0(\Omega, K, Z, \xi, m, c) \in (0,1)$ and a constant $C_2 = C_2(\Omega, K, Z, \xi, m, c) > 0$, both depending only on listed parameters, such that for all $\sigma \in (0, \sigma_0)$ the following Hölder stability estimate is valid:

$$\|V^{(1)} - V^{(2)}\|_{H^1(\Omega_{d+c})} \le C_2(1 + \|V^{(1)} - V^{(2)}\|_{H^2(\Omega)})\sigma^\rho, \quad \rho = c/(m+c). \tag{6.36}$$

In particular, if $p_0^{(1)} = p_0^{(2)}$ and $p_1^{(1)} = p_1^{(2)}$, that is, if $\sigma = 0$, then $V^{(1)}(x) = V^{(2)}(x)$ in Ω_d, which means that uniqueness of the problem (6.33), (6.34) *holds in the domain Ω_d.*

Proof. Uniqueness in the domain Ω_d follows from (3.13) immediately. In this proof, $C_2 = C_2(\Omega, K, Z, \xi, m, c) > 0$ denotes different positive constants depending only on listed parameters. Consider the set of vector functions $Y = Y(K) = \{V \in C^2(\overline{\Omega}) : \|V\|_{C^1(\overline{\Omega})} \le K\}$. Denote $\widetilde{V}(x) = V^{(1)}(x) - V^{(2)}(x)$. Then $\widetilde{V}(x) = (\widetilde{v}_0(x), \dots, \widetilde{v}_{N-1}(x))^T$. Since each component of the vector function $P(x, \nabla V)$ is a quadratic function with respect to the first derivatives $\partial_{x_k} v_i(x)$, then

$$P(x, \nabla V^{(1)}) - P(x, \nabla V^{(2)}) = \widehat{P}(x, \nabla V^{(1)}, \nabla V^{(2)}) \nabla \widetilde{V}(x), \tag{6.37}$$

where the matrix $\widehat{P}(x, \nabla V^{(1)}, \nabla V^{(2)})$ is such that

$$\max_{V^{(1)}, V^{(2)} \in Y} \|\widehat{P}(x, \nabla V^{(1)}, \nabla V^{(2)})\|_{C(\overline{\Omega})} \le C_2. \tag{6.38}$$

We obtain from (6.33), (6.34), (6.37), and (6.38)

$$|A_0(x, \widetilde{V})| \le C_2 |\nabla \widetilde{V}(x)|, \quad \forall x \in \Omega, \tag{6.39}$$

$$\widetilde{V}|_\Gamma = \widetilde{p}_0(x), \quad \partial_n \widetilde{V}|_\Gamma = \widetilde{p}_1(x), \tag{6.40}$$

where $\widetilde{p}_0(x) = (p_0^{(1)} - p_0^{(2)})(x)$, and $\widetilde{p}_1(x) = (p_1^{(1)} - p_1^{(2)})(x)$. Square both sides of (6.39), sum up with respect to $i = 0, \ldots, N-1$, multiply by the function $\varphi_\lambda^2(x)$ defined in (6.9), integrate over the domain Ω_d, and then apply (6.8), (6.12)–(6.14) as well as the Gauss' formula. Also, use (6.10), (6.11) and (6.35). Since $\|\widetilde{V}\|_{L_2(\xi_d)}, \|\nabla\widetilde{V}\|_{L_2(\xi_d)} \leq C_1\|\widetilde{V}\|_{H^2(\Omega)}$, then we obtain for $\lambda \geq \lambda_0$,

$$C_1\lambda \int\limits_{\Omega_d} \left|\nabla\widetilde{V}(x)\right|^2 \varphi_\lambda^2 dx + C_1\lambda^3 \int\limits_{\Omega_d} \left|\widetilde{V}(x)\right|^2 \varphi_\lambda^2 dx \leq C_1\lambda^3 e^{2\lambda m}\sigma^2 + C_1\lambda^3 e^{2\lambda d}\|\widetilde{V}\|_{H^2(\Omega)}^2 + C_2 \int\limits_{\Omega_d} \left|\nabla\widetilde{V}(x)\right|^2 \varphi_\lambda^2 dx. \tag{6.41}$$

Choose $\lambda_1 = \lambda_1(\Omega, K, Z, c) \geq \lambda_0$ so large that $C_2 < C_1\lambda_1/2$. Then we obtain from (6.41) for $\lambda \geq \lambda_1$,

$$\lambda \int\limits_{\Omega_d} \left|\nabla\widetilde{V}(x)\right|^2 \varphi_\lambda^2 dx + \lambda^3 \int\limits_{\Omega_d} \left|\widetilde{V}(x)\right|^2 \varphi_\lambda^2 dx \leq C_1\lambda^3 e^{2\lambda m}\sigma^2 + C_1\lambda^3 e^{2\lambda d}\|\widetilde{V}\|_{H^2(\Omega)}^2. \tag{6.42}$$

Since $\Omega_{d+c} \subset \Omega_d$, $\Omega_{d+c} \neq \varnothing$ and also since

$$\varphi_\lambda^2(x) > e^{2\lambda(d+c)} \quad \text{for } x \in \Omega_{d+c}, \tag{6.43}$$

we obtain from (6.42)

$$\|\widetilde{V}\|_{H^1(\Omega_{d+c})}^2 \leq C_2 e^{2\lambda m}\sigma^2 + C_2 e^{-2\lambda c}\|\widetilde{V}\|_{H^2(\Omega)}^2, \quad \forall \lambda \geq \lambda_1. \tag{6.44}$$

Choose $\lambda = \lambda(\sigma, m, c)$ such that $e^{2\lambda m}\sigma^2 = e^{-2\lambda c}$. Hence, $\lambda = \ln \sigma^{-1/(m+c)}$. We assume that the number σ_0 is so small that $\ln \sigma_0^{-1/(m+c)} > \lambda_1$. Hence, by (6.44) for $\sigma \in (0, \sigma_0)$

$$\|\widetilde{V}\|_{H^1(\Omega_{d+c})}^2 \leq C_2\left(1 + \|\widetilde{V}\|_{H^2(\Omega)}^2\right)\sigma^{2\rho}, \quad \rho = c/(m+c). \tag{6.45}$$

□

6.4 Convexification

6.4.1 Weighted Tikhonov-like functional

Assume that there exists a vector function $p \in C^2(\overline{\Omega})$ such that

$$p|_\Gamma = p_0(x), \quad \partial_n p|_\Gamma = p_1(x), \tag{6.46}$$

where functions p_0, p_1 are defined in (6.34). Consider the vector function $W(x)$,

$$W(x) = (w_0, w_1, \ldots, w_{N-1})^T(x) = V(x) - p(x). \tag{6.47}$$

Then the problem (6.33), (6.34) becomes

$$L(x,p,W) := A_0 W - Q(x,\nabla p,\nabla W) + A_0 p = 0, \tag{6.48}$$

$$W|_\Gamma = \partial_n W|_\Gamma = 0, \tag{6.49}$$

where the N-Dimensional vector function $Q \in C^1(\mathbb{R}^{s_2})$, $s_2 = n(2N+1)$ and each component of Q is a quadratic function with respect to first derivatives $\partial_{x_k} w_i(x)$, $\partial_{x_k} p_i(x)$, where $k = 1,\dots,n$ and $i = 0,\dots,N-1$.

Let $s = [n/2]+2$, where $[n/2]$ is the largest integer, which does not exceed $n/2$. Consider the space $H^s(\Omega)$. By the embedding theorem $H^s(\Omega) \subset C^1(\overline{\Omega})$ and with a generic constant $C > 0$,

$$\|f\|_{C^1(\overline{\Omega})} \le C\|f\|_{H^s(\Omega)}, \quad \forall f \in H^s(\Omega). \tag{6.50}$$

Introduce the space $H^s_{0,\Gamma}(\Omega)$ of N-dimensional vector functions $W(x)$ as

$$H^s_{0,\Gamma}(\Omega) = \{W \in H^s(\Omega) : W|_\Gamma = \partial_n W|_\Gamma = 0\}.$$

Let $R > 0$ be an arbitrary number. Denote

$$B(R) = \{W \in H^s_{0,\Gamma}(\Omega) : \|W\|_{H^s(\Omega)} < R\}.$$

As in Theorem 6.3.1, choose a number $c > 0$ such that $\Omega_{d+c} \neq \varnothing$. Obviously, $\Omega_{d+c} \subset \Omega_d$. To solve the problem (6.48), (6.49) numerically, consider the following weighted Tikhonov-like functional with the CWF $\varphi_\lambda^2(x)$ in it:

$$J_{\lambda,\gamma}(W) = e^{-2\lambda(d+c)} \int_\Omega [L(x,p,W)]^2 \varphi_\lambda^2(x)dx + \gamma\|W\|^2_{H^s(\Omega)}, \tag{6.51}$$

where $\gamma > 0$ is the regularization parameter and the multiplier $e^{-2\lambda(d+c)}$ is introduced here in order to balance first and second terms in the right-hand side of (6.51).

Minimization problem. Minimize the functional $J_{\lambda,\gamma}(W)$ on the closed ball $\overline{B(R)}$.

The second term in the right-hand side of (6.51) is taken in the norm of the space $H^s(\Omega)$ in order to make sure that the iterative terms of the gradient projection method applied to the functional $J_{\lambda,\gamma}(W)$ belong to the space $C^1(\overline{\Omega})$; see (6.50).

Theorem 6.4.1. *The functional $J_{\lambda,\gamma}(W)$ has the Frechét derivative $J'_{\lambda,\gamma}(W)$ at every point $W \in H^s_{0,\Gamma}(\Omega)$. This derivative satisfies the Lipschitz condition in $\overline{B(R)}$, that is, there exists a constant* $\mathrm{Lip} = \mathrm{Lip}(\lambda,\gamma,Z,R) > 0$ *depending only on listed parameters such that for all $\lambda,\gamma > 0$,*

$$\|J'_{\lambda,\gamma}(W_1) - J'_{\lambda,\gamma}(W_2)\|_{H^s(\Omega)} \le \mathrm{Lip}\,\|W_1 - W_2\|_{H^s(\Omega)}, \quad \forall W_1, W_2 \in \overline{B(R)}.$$

Theorem 6.4.2 (global strict convexity). *Choose a number $D > 0$ such that $\|p\|_{C^2(\overline{\Omega})} \leq D$. There exists a sufficiently large number $\lambda_2 = \lambda_2(\Omega, R, Z, d, \xi, c) \geq 1$ depending only on listed parameters such that for all $\lambda \geq \lambda_2$ and for $\gamma \in [e^{-\lambda c}, 1)$ the functional $J_{\lambda,\gamma}(W)$ is strictly convex on $B(R)$, that is,*

$$J_{\lambda,\gamma}(W_2) - J_{\lambda,\gamma}(W_1) - J'_{\lambda,\gamma}(W_1)(W_2 - W_1) \geq C_1\|W_2 - W_1\|^2_{H^1(\Omega_{d+c})} + \frac{\gamma}{2}\|W_2 - W_1\|^2_{H^s(\Omega)}, \quad \forall W_1, W_2 \in \overline{B(R)}. \tag{6.52}$$

Remark 6.4.1. Since the regularization parameter $\gamma \in (e^{-\lambda c}, 1)$, then this allows values of γ to be small. Also, the presence of the first term in the right-hand side of (6.52) indicates that the stable reconstruction should be expected in the subdomain Ω_{d+c} rather than in the whole domain Ω. Theorem 6.4.4 confirms the latter.

Let $P_{\overline{B(R)}} : H^s_{0,\Gamma}(\Omega) \to \overline{B(R)}$ be the projection operator of the Hilbert space $H^s_{0,\Gamma}(\Omega)$ on the closed ball $\overline{B(R)}$. Let $\rho \in (0, 1)$ be a number, which we will choose later. Let $W_0 \in B(R)$ be an arbitrary point. The gradient projection method of the minimization of the functional $J_{\lambda,\gamma}(W)$ on the set $\overline{B(R)}$ is defined by the following sequence:

$$W_n = P_{\overline{B(R)}}(W_{n-1} - \rho J'_{\lambda,\gamma}(W_{n-1})), \quad n = 1, 2, \ldots . \tag{6.53}$$

Theorem 6.4.3. *Let $\lambda_2 = \lambda_2(\Omega, R, Z, d, \xi)$ be the number introduced in Theorem* 6.4.2. *Fix a number $\lambda \geq \lambda_2$ and let the regularization parameter $\gamma \in [e^{-\lambda c}, 1)$. Then there exists unique minimizer $W_{\min} \in \overline{B(R)}$ of the functional $J_{\lambda,\gamma}(W)$ on the set $\overline{B(R)}$. Furthermore, there exists a sufficiently small number $\rho_0 = \rho_0(\Omega, R, Z, d, \xi, \lambda, c) \in (0, 1)$ depending only on listed parameters such that for every $\rho \in (0, \rho_0)$ there exists a number $q = q(\rho) \in (0, 1)$ such that the sequence* (6.53) *converges to $W_{\min}$,*

$$\|W_n - W_{\min}\|_{H^s(\Omega)} \leq q^n\|W_0 - W_{\min}\|_{H^s(\Omega)}, \quad n = 1, 2, \ldots .$$

Consider now the question of the convergence of the sequence (6.53) to the exact solution W^* of the problem (6.48), (6.49).

Theorem 6.4.4. *Assume that there exists exact solution $W^* \in B(R)$ of the problem* (6.48), (6.49) *with the exact data $p^* \in C^2(\overline{\Omega})$. Let $p \in C^2(\overline{\Omega})$ be the noisy data. Assume that $\|p - p^*\|_{C^2(\overline{\Omega})} \leq \sigma$, where $\sigma \in (0, 1)$ is the level of the error in the data. Also, assume that the $C^2(\overline{\Omega})$-norm of the exact data p^* is bounded by an a priori given constant M^*, that is, $\|p^*\|_{C^2(\overline{\Omega})} \leq M^*$ (then $\|p\|_{C^2(\overline{\Omega})} \leq M^* + 1$). Let $\lambda_2 = \lambda_2(\Omega, R, Z, d, \xi)$ be the number of Theorem* 6.4.2. *Then there exists a number $\lambda_3 = \lambda_3(\Omega, R, Z, d, \xi, c, M^*) > \lambda_2$, a sufficiently small number $\sigma_1 = \sigma_1(\Omega, R, Z, c, d, \xi, M^*) \in (0, 1)$ and a number $\theta = c/(8m) \in (0, 1)$, all depending only on listed parameters, such that if $\ln \sigma_1^{-2\theta/c} > \lambda_3$, then if for any $\sigma \in (0, \sigma_1)$ one chooses $\lambda = \ln \sigma^{-2\theta/c}$ and $\gamma = e^{-\lambda c} = \sigma^{2\theta}$, then the following convergence estimate holds for the sequence* (6.53)*:*

$$\|W^* - W_n\|_{H^1(\Omega_{d+c})} \leq C_4\sigma^\theta + q^n\|W_0 - W_{\min}\|_{H^s(\Omega)}, \quad n = 1, 2, \ldots , \tag{6.54}$$

where the number $q = q(\rho) \in (0,1)$ *and* $\rho \in (0,\rho_1)$, *where* $\rho_1 = \rho_1(\Omega, R, Z, c, d, \xi, M^*) \in (0,1)$ *is a sufficiently small number. In* (6.54), $C_4 = C_4(\Omega, R, Z, D, c, d, \xi, M^*) = \text{const.} > 0$. *All numbers here depend only on listed parameters.*

Theorem 6.4.2 is the central one among Theorems 6.4.1–6.4.4. Thus, we prove Theorem 6.4.2 below. As to the rest of theorems of this section, we omit their proofs referring the reader to proofs of similar theorems in Chapter 5 for the case of a single quasilinear PDE.

Remark 6.4.2. Unlike (6.54), in the case of nonconvex functionals, there is no guarantee that a gradient-like method converges to the exact solution starting from an arbitrary point. Since the starting point $W_0 \in B(R)$ of the iterative process, (6.53) is an arbitrary one and since smallness restrictions on the radius R are not imposed. Then convergence estimate (6.54) means the global convergence in the space $H^1(\Omega_{d+c})$; see Definition 1.4.2.

Proof of Theorem 6.4.2. In this proof, $C_3 = C_3(\Omega, R, Z, D, c, d, \xi) > 0$ denotes different constants depending only on listed parameters. Let $W_1, W_2 \in \overline{B(R)}$ be two arbitrary functions. Denote $W_2 - W_1 = h = (h_0(x), \dots, h_{N-1}(x))^T$. Since each component of the vector function $Q(x, \nabla p, \nabla W)$ in (6.48) is a quadratic function with respect to first derivatives $\partial_{x_k} w_i(x), \partial_{x_k} p_i(x)$, we have

$$
\begin{aligned}
&Q(x, \nabla p, \nabla W_1 + \nabla h) \\
&\quad = Q(x, \nabla p, \nabla W_1) + Q^{(1)}(x, \nabla p, \nabla W_1, \nabla h) + Q^{(2)}(x, \nabla p, \nabla W_1, \nabla h).
\end{aligned} \tag{6.55}
$$

Here, each component of the vector function $Q^{(1)}$ is linear with respect to derivatives $\partial_{x_k} h_i$ and each component of the vector function $Q^{(2)}$ contains only quadratic terms $(\partial_{x_k} h_i) \cdot (\partial_{x_l} h_j)$. Hence, the following estimates hold for all $x \in \overline{\Omega}$:

$$
\left|Q^{(1)}(x, \nabla p, \nabla W_1, \nabla h)\right| \le C_3 |\nabla h|, \quad \left|Q^{(2)}(x, \nabla p, \nabla W_1, \nabla h)\right| \le C_3 |\nabla h|^2. \tag{6.56}
$$

By (6.48) and (6.55),

$$
\begin{aligned}
&[L(x, p, W_1 + h)]^2 - [L(x, p, W_1 + h)]^2 \\
&\quad = \text{Lin}(x, p, h) \\
&\qquad + (A_0(h))^2 + 2A_0(h)[Q^{(1)}(x, \nabla p, \nabla W_1, \nabla h) + Q^{(2)}(x, \nabla p, \nabla W_1, \nabla h)] \\
&\qquad + [Q^{(1)}(x, \nabla p, \nabla W_1, \nabla h) + Q^{(2)}(x, \nabla p, \nabla W_1, \nabla h)]^2,
\end{aligned} \tag{6.57}
$$

where the functional $\text{Lin}(x, p, h)$ depends linearly on h. Combining (6.57) with the Cauchy–Schwarz inequality and as well as with (6.56), we obtain

$$
[L(x, p, W_1 + h)]^2 - [L(x, p, W_1 + h)]^2 - \text{Lin}(x, p, h) \ge \frac{1}{2}(A_0(h))^2 - C_3 (\nabla h)^2. \tag{6.58}
$$

Let $\{,\}$ be the scalar product in the space of such real valued N-dimensional vector functions whose components belong to $H^s(\Omega)$. Then (6.51) and (6.58) imply that

$$J_{\lambda,\gamma}(W_1+h)-J_{\lambda,\gamma}(W_1)-e^{-2\lambda(d+c)}\int_\Omega \operatorname{Lin}(x,p,h)\varphi_\lambda^2dx-2\gamma\{W,h\}$$
$$\geq \frac{1}{2}\int_\Omega (A_0(h))^2\varphi_\lambda^2dx - C_3\int_\Omega(\nabla h)^2\varphi_\lambda^2dx+\gamma\|h\|^2_{H^s(\Omega)}. \tag{6.59}$$

It easily follows from the proof of Theorem 3.1 of [9], which is a close analog of Theorem 6.4.1, that

$$J'_{\lambda,\gamma}(W_1)(h)=e^{-2\lambda(d+c)}\int_\Omega \operatorname{Lin}(x,p,h)\varphi_\lambda^2dx+2\gamma\{W,h\}. \tag{6.60}$$

Applying the Carleman estimate (6.12)–(6.14) to the right-hand side of (6.59), we obtain for $\lambda\geq\lambda_0$:

$$\frac{e^{-2\lambda(d+c)}}{2}\int_\Omega (A_0(h))^2\varphi_\lambda^2dx - C_3e^{-2\lambda(d+c)}\int_\Omega(\nabla h)^2\varphi_\lambda^2dx+\gamma\|h\|^2_{H^s(\Omega)}$$
$$\geq \frac{e^{-2\lambda(d+c)}}{2}\int_{\Omega_d} (A_0(h))^2\varphi_\lambda^2dx - C_3e^{-2\lambda(d+c)}\int_{\Omega_d}(\nabla h)^2\varphi_\lambda^2dx$$
$$-C_3\int_{\Omega\setminus\Omega_d}(\nabla h)^2\varphi_\lambda^2dx+\gamma\|h\|^2_{H^s(\Omega)}$$
$$\geq C_1e^{-2\lambda(d+c)}\lambda\int_{\Omega_d}(\nabla h)^2\varphi_\lambda^2dx+C_1e^{-2\lambda(d+c)}\lambda^3\int_{\Omega_d}h^2\varphi_\lambda^2dx-C_3e^{-2\lambda(d+c)}\int_{\Omega_d}(\nabla h)^2\varphi_\lambda^2dx$$
$$-C_3e^{-2\lambda(d+c)}\int_{\Omega\setminus\Omega_d}(\nabla h)^2\varphi_\lambda^2dx-C_1\lambda^3e^{-2\lambda c}\int_{\xi_d}((\nabla h)^2+h^2)dS+\gamma\|h\|^2_{H^s(\Omega)}. \tag{6.61}$$

Choose $\lambda_2=\lambda_2(\Omega,R,Z,D,d,\xi)\geq\lambda_0$ so large that $C_1\lambda_2/2>C_3$. Also, observe that

$$\varphi_\lambda^2(x)\leq e^{2\lambda d},\quad \forall x\in\Omega\setminus\Omega_d \quad\text{and}\quad \|\nabla h\|^2_{L_2(\xi_d)}+\|h\|^2_{L_2(\xi_d)}\leq C_3\|h\|^2_{H^s(\Omega)}.$$

Hence, taking into account (6.43), we obtain from (6.61),

$$\frac{e^{-2\lambda(d+c)}}{2}\int_\Omega (A_0(h))^2\varphi_\lambda^2dx - C_3e^{-2\lambda(d+c)}\int_\Omega(\nabla h)^2\varphi_\lambda^2dx+\gamma\|h\|^2_{H^s(\Omega)} \tag{6.62}$$
$$\geq C_1\|h\|^2_{H^1(\Omega_{d+c})}+(\gamma-C_3e^{-2\lambda c})\|h\|^2_{H^s(\Omega)},\quad \forall\lambda\geq\lambda_2.$$

Since $\gamma\in[e^{-\lambda c},1)$, then (6.59), (6.60), and (6.62) imply (6.52). □

6.4.2 Numerical scheme

The numerical scheme for the above technique is as follows:

Step 1. Using the Gram–Schmidt orthonormalization procedure in $L_2(0,1)$, obtain functions $\{\psi_m(x_0)\}_{m=0}^{N-1}$, $x_0 \in (0,1)$ from functions $\{x_0^m e^{x_0}\}_{m=0}^{N-1}$ for a reasonable integer $N \geq 2$.

Step 2. Sequentially obtain problems (6.26), (6.27), then (6.33), (6.34), and then (6.48), (6.49) for the specific operator A.

Step 3. Minimize the functional (6.51) on the set $\overline{B(R)}$ using the gradient projection method.

Step 4. Let $W_{\min}(x)$ be the minimizer of the functional (6.51) on the set $\overline{B(R)}$ (Theorem 6.4.3). Set $V_{\min}(x) = W_{\min}(x) + p(x)$. Next, use the first formula (6.31), then use (6.28) and finally use (6.24).

6.5 Two specific examples

The goal of Section 6.5 is to provide some specific examples of CIPs for parabolic and hyperbolic equations for which the above technique works. The case of CIPs for elliptic PDEs is considered in Chapters 7 and 10. In Chapters 11 and 12, we consider the travel time tomography problem. In these chapters, the case of a moving source is considered. In fact, however, in the case of Helmholtz equation, the source can be fixed while the frequency can be varied, and still the technique of this chapter works; see references in the beginning of this chapter. Basically, it does not really matter for the method of this chapter which parameter is varied: source position, frequency, angle of the incident plane wave, etc. The only important factor is that the unknown coefficient should not depend on this parameter.

6.5.1 Parabolic equation

Let $T > 1$ be an arbitrary number. Denote $D_T^{n+1} = \mathbb{R}^n \times (0,T)$. Consider the parabolic operator in D_T^{n+1},

$$Au = u_t - \sum_{i,j=1}^{n} a_{i,j}(x,t)u_{x_i x_j} - \sum_{j=1}^{n} b_j(x,t)u_{x_j} + a_0(x,t)u, \tag{6.63}$$

$$A_0 u = u_t - \sum_{i,j=1}^{n} a_{i,j}(x,t)u_{x_i x_j}, \tag{6.64}$$

where $a_0(x,t)$ is the unknown coefficient and $a_{i,j}(x,t) = a_{j,i}(x,t)$, $\forall i,j$. We assume that all coefficients of the operator (6.63) belong to $C^1(\overline{D_T^{n+1}})$ and also that the obvious analog of the ellipticity condition holds,

$$\mu_1|\xi|^2 \le \sum_{i,j=1}^{n} a_{i,j}(x,t)\xi_i\xi_j \le \mu_2|\xi|^2, \quad (\xi,x,t) \in \mathbb{R}^n \times \overline{D_T^{n+1}},$$

where $\mu_1, \mu_2 = \text{const.}$, $0 < \mu_1 < \mu_2$.

Let $\overline{x}^0 = (0,\dots,0,-1)$ and let the domain $\Omega_1 \subset \{x_n > 0\}$. We assume that $Au = u_t - \Delta u$ for $x \in \mathbb{R}^n \setminus \Omega_1$. In the case of the parabolic equation, the point source runs over $I_1 = I \times \{t = 0\}$, where the interval I is defined in (6.15). Consider the fundamental solution $u(x,t,x_0)$ of the operator A,

$$Au = \delta(x_1 - x_0, \overline{x} - \overline{x}^0, t), \quad (x,t) \in D_T^{n+1}, \tag{6.65}$$

$$u(x,0,x_0) = 0, \quad \forall x_0 \in [0,1]. \tag{6.66}$$

It is well known that there exists unique solution $u(x,t,x_0) \in C^{2+\alpha,1+\alpha/2}(\overline{D_T^{n+1}} \setminus I_{1\varepsilon})$, $\forall \varepsilon > 0$, $\forall x_0 \in [0,1]$ of the problem (6.65), (6.66); see Chapter 4 of [173]. Here, $I_{1\varepsilon} \subset \overline{D_T^{n+1}}$ is defined similarly with I_ε in Section 6.2.2. Furthermore, $u(x,t,x_0) > 0$ for $t > 0$; see Theorem 11 in Chapter 2 of [79]. Hence (see Condition 3 in Section 6.2.2), we define $\Omega = \Omega_1 \times (\zeta, T)$, where $\zeta \in (0, (T-1)/2)$ is a sufficiently small number.

Choose a number $\omega > 1$ and set

$$\Gamma_0 = \left\{ x \in \mathbb{R}^n : x_n = 0, (x_1 - 1/2)^2/\omega^2 + \sum_{k=2}^{n-1} x_k^2 < 1/4 \right\}, \quad \Gamma = \Gamma_0 \times (\zeta, T),$$

$$\xi(x,t) = \left[x_n + (x_1 - 1/2)^2/\omega^2 + \sum_{k=2}^{n-1} x_k^2 + (t - T/2)^2 + 1/4 \right]^{-\nu}, \tag{6.67}$$

and $\varphi_\lambda(x,t) = \exp(\lambda\xi(x,t))$, where $\nu = \nu(\omega) > 1$ is a parameter depending on ω. We assume that $\Gamma_0 \subset \partial\Omega_1$. Hence, $\Gamma \subset \partial\Omega$. In addition, we assume that the domain

$$\{\xi(x,t) > 2^\nu, x_n > 0\} = \Omega_{2^\nu} \subset \Omega.$$

Note that

$$\Gamma_{2^\nu} = \left\{ x_n = 0, (x_1 - 1/2)^2/\omega^2 + \sum_{k=2}^{n-1} x_k^2 + (t - T/2)^2 < 1/4 \right\} \subset \Gamma.$$

Let the restricted DN data be given on Γ,

$$u(x,t,x_0) = g_0(x,t,x_0), \quad \partial_n u(x,t,x_0) = g_1(x,t,x_0), \quad \forall (x,t,x_0) \in \Gamma \times [0,1]. \tag{6.68}$$

A direct analog of the Carleman estimate of Theorem 2.3.1 (6.12)–(6.14) is valid for the operator A_0 in (6.64). Hence, the above construction works in this case. The unknown coefficient $a_0(x,t)$ can be Hölder-stable reconstructed numerically by the above method in the domain $\Omega_{2^\nu+c}$ for any $c > 0$ such that $\Omega_{2^\nu+c} \neq \varnothing$. Uniqueness, of the corresponding CIP holds for the entire domain $\Omega_1 \times (0,T)$.

6.5.2 Hyperbolic equation

Let the domains $\Omega_k \subset \mathbb{R}^3$ be defined as $\Omega_k = \{|x| < k\}$, $k = 1,2,3$. Let I_ε be the set defined in Section 6.2.2. We assume that

$$I_\varepsilon \subset (\Omega_3 \setminus \Omega_2). \tag{6.69}$$

Let the function $a(x,t) \in C(\overline{D_T^4})$ be such that

$$a(x,t) \geq 0 \quad \text{in } D_T^4, \quad a(x,t) = 0 \quad \text{for } x \in \mathbb{R}^3 \setminus \Omega_2. \tag{6.70}$$

Let $f(z)$, $z \in \mathbb{R}$ and $\chi(\overline{x})$, $\overline{x} \in \mathbb{R}^{n-1}$ be functions defined in Section 6.2.2. Recall that $f(0)\chi(0) \neq 0$. We assume that

$$f(z) \geq 0, \quad \forall z \in \mathbb{R} \quad \text{and} \quad \chi(\overline{x}) \geq 0, \quad \forall \overline{x} \in \mathbb{R}^{n-1}. \tag{6.71}$$

Consider the following Cauchy problem for the function $u(x,t,x_0)$:

$$u_{tt} = \Delta u + a(x,t)u + f(x_1 - x_0)\chi(\overline{x}), \quad (x,t) \in D_T^4, \tag{6.72}$$

$$u(x,0,x_0) = u_t(x,0,x_0) = 0, \tag{6.73}$$

where $x_0 \in [0,1]$ is a parameter. Then the problem (6.72), (6.73) is equivalent with

$$u(x,t,x_0) = \int\limits_{|x-\eta|<t} \frac{f(\eta_1 - x_0)\chi(\overline{\eta} - \overline{x}^0)}{4\pi|x-\eta|} d\eta + \int\limits_{|x-\eta|<t} \frac{(au)(\eta, t - |x-\eta|)}{4\pi|x-\eta|} d\eta. \tag{6.74}$$

On can prove (see [184] for a similar result) that the integral equation (6.74) can be rewritten as Volterra integral equation, whose solution can be represented as a series, which converges absolutely and uniformly in any subdomain $(G \times (0,T)) \subset D_T^4$ and for any $x_0 \in [0,1]$, where $G \subset \mathbb{R}^3$ is an arbitrary bounded domain. This series is

$$u = \sum_{n=0}^{\infty} u_n,$$

$$u_0 = \int\limits_{|x-\eta|<t} \frac{f(\eta_1 - x_0)\chi(\overline{\eta})}{4\pi|x-\eta|} d\eta, \quad \eta = (\eta_1, \overline{\eta}), \tag{6.75}$$

$$u_n = \int\limits_{|x-\eta|<t} \frac{(au_{n-1})(\eta, t - |x-\eta|)}{4\pi|x-\eta|} d\eta, \quad n \geq 1. \tag{6.76}$$

We now prove that

$$u(x,t,x_0) \geq C_5 T, \quad \forall x \in \Omega_1, \forall x_0 \in [0,1], \forall t \in (T/4, T), \forall T > 20, \tag{6.77}$$

where the constant $C_5 = C_5(I_\varepsilon, f, \chi) > 0$ depends only on listed parameters and is independent on T.

Indeed, let $x \in \Omega_1$ and $\eta \in \Omega_3$ be two arbitrary points. Let $t \in (T/4, T)$ and $T > 20$. Then

$$|x - \eta| \le |x| + |\eta| < 4 < t. \tag{6.78}$$

Since by (6.71) $f(0)\chi(0) > 0$, then (6.77) follows from (6.69)–(6.71), (6.75), (6.76), and (6.78).

We now set $\Omega = \Omega_1 \times (T/4, T)$, where $T > 20$. Next, let the point $y \in \mathbb{R}^3$ be such that $|y| > 3$. Define the Carleman weight function $\varphi_\lambda(x,t)$ as

$$\varphi_\lambda(x,t) = \exp(\lambda\xi(x,t)), \quad \xi(x,t) = |x - y|^2 - \varrho^2(t - T/2)^2. \tag{6.79}$$

Choose any number $d \in (0,1)$. Next, choose $\varrho \in (4\sqrt{1-d}/T, 1)$. Let $\Omega_d = \{(x,t) : x \in \Omega_1, \xi(x,t) > d\}$. Then

$$\Omega_d \subset \Omega, \quad \Omega_d \cap \{t = T/4\} = \Omega_d \cap \{t = T\} = \varnothing. \tag{6.80}$$

Hence, we define Γ and Γ_d as

$$\Gamma = \{(x,t) : |x| = 1, t \in (T/4, T)\}, \quad \Gamma_d = \{(x,t) : |x| = 1, \xi(x,t) > d\}. \tag{6.81}$$

It follows from (6.79)–(6.81) that $\Gamma_d \subset \Gamma$.

Similarly, with (6.18) we define the CIP in this case as the problem of determining the unknown coefficient $a(x,t) \in C(\overline{D_T^4})$ satisfying conditions (6.70), given functions $g_0(x,t,x_0)$, $g_1(x,t,x_0)$, where

$$u(x,t,x_0) = g_0(x,t,x_0), \quad \partial_n u(x,t,x_0) = g_1(x,t,x_0), \quad \forall (x,t,x_0) \in \Gamma \times [0,1].$$

By Theorem 2.5.1, the Carleman estimate is valid for the operator $\partial_t^2 - \Delta$ with the CWF $\varphi_\lambda(x,t)$ given in (6.79). Therefore, the above construction works for this CIP. The function $a(x,t)$ can be reconstructed numerically by the above method in Ω_{d+c} for any $c \in (0, 1-d)$.

7 Convexification of electrical impedance tomography with restricted Dirichlet-to-Neumann map data

In this chapter, we follow [149]. Permission for republication has been obtained from the publisher.

7.1 Introduction

We use in this chapter the idea of Chapter 6 to develop a globally convergent numerical method of the reconstruction of the internal electrical conductivity in the inverse problem of electrical impedance tomography (EIT). We "convexify" the problem. More precisely, we construct a weighted least squares Tikhonov-like functional. The weight of this functional is the CWF. Its presence is the *key element* of that functional. The CWF is the function, which is involved in the Carleman estimate for the Laplace operator. The presence of the CWF ensures the strict convexity of this functional on any a priori chosen ball of an arbitrary radius $R > 0$ in an appropriate Hilbert space. The latter guarantees the global convergence of the gradient projection method of the minimization of this functional to the exact solution of the original inverse EIT problem. We remind (see Definition 1.4.2) that we call a numerical method for a coefficient inverse problem (CIP) *globally convergent* if there is a theorem, which guarantees that this method delivers at least one point in a sufficiently small neighborhood of the exact solution of that CIP without any advanced knowledge of this neighborhood. We point out that the numerical method of this paper *converges globally.*

EIT is a noninvasive and diffusive imaging method to recover the electrical conductivity distribution inside an object of interest by using the DtN map on the boundary. This modality is safe, portable, and also has many clinical imaging applications. There is a vast number of research papers discussing EIT. It has been analytically proven that the interior electrical conducting is uniquely determined by the Dirichlet-to-Neumann map on the boundary [56, 203, 241].

In the past three decades, there were numerous studies on the EIT imaging method with quite many publications. We now provide a far incomplete list of references on this topic; also see references cited therein: [6, 7, 38, 89, 105, 106, 200, 231, 232, 235]. Harrach [85, 86] has developed two globally convergent numerical methods for EIT.

In a typical EIT experiment, constant electrical currents are applied to the electrodes on the boundary of the object to image. Then the electrical potentials are measured on the boundary. This gives the DtN map data. The EIT problem is to recover the internal electric conductivity from these DtN measurements. This problem is essentially ill posed.

https://doi.org/10.1515/9783110745481-007

The key element of our method is the construction of a weighted Tikhonov-like functional which is strictly convex on any a priori chosen ball of an arbitrary radius $R > 0$ in an appropriate Hilbert space. In other words, we "convexify" the problem. The main ingredient of that Tikhonov-like functional is the presence of the CWF in it. If the exact solution belongs to that ball (as it should be assumed in the framework of the regularization theory [22, 76, 244]), then the convergence of the gradient projection method to the exact solution is guaranteed if starting from an arbitrary point of this ball. Hence, this is *global convergence*. On the other hand, recall that convergence of any gradient-like method to the exact solution for a nonconvex functional might be guaranteed only if its starting point is located in a small neighborhood of this solution.

To minimize the above mentioned weighted Tikhonov-like functional, we propose a multi-level method, which is somewhat similar with the adaptivity method; see, for example, [22] for a detailed theory of the adaptivity. However, we do not extend to our case the theory of the adaptivity presented in [22], that is, we restrict our attention only to the numerical aspect of the adaptivity. Thus, we minimize that functional on a coarse mesh first and use the solution achieved on the coarse mesh (first level) as the starting point for a finer mesh (second level). We repeat this process until we get a solution on K_{th} level. We have found that we get a rough image on the coarse mesh (e. g., support, shape) of the internal conductivity much faster than on a finer mesh, while on the finer mesh with the starting point from the solution on the coarse mesh, the solution is corrected in details (e. g., amplitude and shape).

7.2 EIT with restricted Dirichlet-to-Neumann (DtN) data

7.2.1 The mathematical model

In this section, we formulate the restricted DtN for the inverse EIT problem. First, we recall the traditional DtN for EIT. Let Ω be an open bounded domain in $\mathbb{R}^d$ $(d = 2, 3)$ to be imaged with a smooth boundary $\partial\Omega$. The EIT forward problem is formulated as: For any given input, current

$$g_1 \in L_0^2(\partial\Omega) := \left\{ g \in L^2(\Omega) : \int_{\partial\Omega} g \, ds = 0 \right\}$$

and the conductivity distribution $\sigma(x)$, find the function $u(x) \in H^1(\Omega)$ such that

$$\begin{cases} \nabla \cdot (\sigma(x) \nabla u(x)) = 0 & \text{in } \Omega, \\ \sigma(x) \frac{\partial u}{\partial \nu} = g_1(x) & \text{on } \partial\Omega, \\ \int_{\partial\Omega} u(x) \, ds = 0, \end{cases} \tag{7.1}$$

where ν is the outward unit normal vector on $\partial\Omega$. Denote $g_0(x) = u|_{\partial\Omega}$. Then the inverse EIT problem is to recover the internal conductivity function $\sigma(x)$ from the DtN map $\Lambda : g_0 \to g_1$.

We consider the EIT problem with the source outside the domain of interest and the restricted DtN data measured on the boundary of the domain of interest, as described below.

To avoid working with singularities and also to simplify the presentation, we model the point source here by a δ-like function instead of the δ-function. Let $\varepsilon > 0$ be a sufficiently small number. Let the source function $f(x)$ be such that

$$f(x) \in C^{\infty}(\mathbb{R}^n), \quad f(0) \neq 0, \quad f(x) \geq 0, \quad \forall x \in \mathbb{R}^d, \quad f(x) = 0 \quad \text{for } |x| > \varepsilon. \tag{7.2}$$

Let $G \subset \mathbb{R}^n$ be a bounded domain with its boundary $\partial G \in C^1$, $\Omega \subset G$, and $\partial\Omega \cap \partial G = \varnothing$. Let $\overline{x} \in \mathbb{R}^{d-1}$ be a fixed point. For $s \in [0,1]$, denote $x_s = (x_{1s}, \overline{x})$ the position of the point source. Let $I = \{x_s = (x_{1s}, \overline{x}) : s \in [0,1]\}$ be the interval of the straight line $\{x = (x_1, \overline{x}), x_1 \in \mathbb{R}\}$. Let $I_\varepsilon = \{x \in \mathbb{R}^d : \text{dist}(x, I) < \varepsilon\}$, where $\text{dist}(x, I)$ is the Hausdorff distance between the point x and the set I. We also assume that $I_\varepsilon \subset (G \setminus \overline{\Omega})$, which means that the support of the source function is outside of the domain Ω.

Let the function

$$\sigma(x) \in C^{2+\alpha}(\overline{G}), \quad \sigma(x) = 1 \quad \text{for } x \in G \setminus \Omega \quad \text{and} \quad \sigma(x) \geq \sigma_0 = \text{const.} > 0. \tag{7.3}$$

Here, $\alpha = \text{const.} \in (0,1)$ and $C^{k+\alpha}(\overline{G})$ be the Hölder space, where $k \geq 0$ is an integer. Assume first that $\sigma(x)$ is known. For each source position $x_s \in I$, we define the forward boundary value problem for EIT as the problem of finding the function $u(x,s)$ such that

$$\begin{cases} \nabla \cdot (\sigma(x) \nabla u(x,s)) = -f(x - x_s), & x \in G, \forall x_s \in \overline{I}, \\ u(x,s)|_{x \in \partial G} = 0, & \forall x_s \in \overline{I}. \end{cases} \tag{7.4}$$

It is well known that for each $x_s \in I$ the problem (7.4) has a unique solution

$$u(x,s) \in C^{3+\alpha}(\overline{G}), \quad \forall x_s \in \overline{I}; \tag{7.5}$$

see, for example, [80]. We measure both Dirichlet and Neumann boundary conditions of the function u on a part $\Gamma \subseteq \partial\Omega$ of the boundary $\partial\Omega$,

$$u(x,s)|_{x \in \Gamma, x_s \in \overline{I}} = g_0(x,s) \quad \text{and} \quad \partial_\nu u(x,s)|_{x \in \Gamma, x_s \in \overline{I}} = g_1(x,s). \tag{7.6}$$

We call the Dirichlet and Neumann boundary data (7.6) “restricted DtN data”.

If the coefficient $\sigma(x)$ is known, then having the solution of the forward problem (7.4), one can easily compute functions $g_0(x,s)$ and $g_1(x,s)$. Suppose now that the function $\sigma(x)$ is unknown. Then we arrive at the following inverse problem.

Coefficient Inverse Problem (CIP). Assume that the function $\sigma(x)$ is unknown for $x \in \Omega$ and also that conditions (7.2), (7.3) hold. Also, assume that functions $g_0(x,s)$ and $g_1(x,s)$ in (7.6) are known for all $x \in \Gamma$, $x_s \in \overline{I}$. Determine the function $\sigma(x)$.

Note that in this CIP the number d of free variables in the data equals the number of free variables in the unknown coefficient.

7.2.2 An equivalent problem

In this subsection, we transform the above CIP to an inverse problem for a quasilinear PDE. First, introduce the well-known change of variables

$$u_1 = \sqrt{\sigma} u, \tag{7.7}$$

where $u(x,s)$ is the solution of problem (7.4). Then

$$\begin{cases} \Delta u_1(x,s) + a_0(x) u_1(x,s) = -f(x - x_s), & \forall x_s \in \bar{I}, \\ u_1(x,s)|_{x \in \partial G} = 0, & \forall x_s \in \bar{I}, \end{cases} \tag{7.8}$$

where

$$a_0(x) = -\frac{\Delta(\sqrt{\sigma(x)})}{\sqrt{\sigma(x)}}. \tag{7.9}$$

Recalling that $\sigma = 1$ on $\partial\Omega$, we obtain from (7.6),

$$u_1(x,s)|_{x \in \Gamma, s \in [0,1]} = g_0(x,s) \quad \text{and} \quad \partial_n u_1(x,s)|_{x \in \Gamma, s \in [0,1]} = g_1(x,s). \tag{7.10}$$

If we would recover the function $a_0(x)$ for $x \in \Omega$ from the conditions (7.8) and (7.10), then assuming that 0 is not an eigenvalue of the elliptic operator $\Delta + a_0(x)$ with the Dirichlet boundary condition either on $\partial\Omega$ or on ∂G, we would recover the function $\sigma(x)$ via solving the elliptic equation (7.9) either in the domain Ω with the Dirichlet boundary condition $\sigma|_{\partial\Omega} = 1$, or in the domain G with the Dirichlet boundary condition $\sigma|_{\partial G} = 1$. Hence, we focus below on the recovery of the function $a_0(x)$ for $x \in \Omega$ from conditions (7.8), (7.10).

It follows from (7.2), (7.4), (7.7), and the maximum principle for elliptic equations [80] that $u_1(x,s) > 0$ for all $x \in \overline{\Omega}$ and all $s \in [0,1]$. Hence, we can consider the function $v(x,s)$,

$$v(x,s) = \ln u_1(x,s). \tag{7.11}$$

Then $u_1(x,s) = e^{v(x,s)}$ and (7.8) imply that

$$\Delta v(x,s) + \left(\nabla v(x,s)\right)^2 = -a_0(x), \quad x \in \Omega, \forall s \in [0,1]. \tag{7.12}$$

Here, we use (7.2) and the fact that $I_\varepsilon \subset (G \setminus \overline{\Omega})$. In addition, using (7.10), we obtain

$$v(x,s)|_{x \in \Gamma, s \in [0,1]} = \tilde{g}_0(x,s) \quad \text{and} \quad \partial_\nu v(x,s)|_{x \in \Gamma, s \in [0,1]} = \tilde{g}_1(x,s), \tag{7.13}$$

where

$$\tilde{g}_0(x,s) = \ln g_0(x,s) \quad \text{and} \quad \tilde{g}_1(x,s) = \frac{g_1(x,s)}{g_0(x,s)}.$$

Differentiating equation (7.12) with respect to s and noting that the function $a_0(x)$ is independent on s, we obtain

$$\Delta v_s + 2\nabla v_s \cdot \nabla v = 0, \quad x \in \Omega, \forall s \in [0,1]. \tag{7.14}$$

Now the above CIP is reduced to the following problem.

Reduced problem. Recover the function $v(x,s)$ from the equation (7.14), given the boundary measurements $\tilde{g}_0(x,s)$ and $\tilde{g}_1(x,s)$ in (7.13).

If the function $v(x,s)$ is approximated, then the approximate coefficient $a_0(x)$ can be found via (7.12). Thus, our focus below is on the solution of the reduced problem.

7.3 Cauchy problem for a system of coupled quasilinear elliptic equations

To solve the above reduced problem, we obtain in this section the Cauchy problem for a system of coupled quasi-linear elliptic equations.

7.3.1 The orthonormal basis of Section 6.2.3

Let [,] denotes the scalar product in $L_2(0,1)$. We use the orthonormal basis in $L_2(0,1)$ of real valued functions $\{\psi_n(s)\}_{n=0}^{\infty}$ which was constructed in Section 6.2.3. Recall that the function $\psi_n(s)$ has the form $\psi_n(s) = P_n(s)e^s$, where $P_n(s)$ is the polynomial of the degree n. Let $a_{mn} = [\psi_n', \psi_m]$. Consider the $N \times N$ matrix $M_N = (a_{mk})_{(k,m)=(0,0)}^{(N-1,N-1)}$. Theorem 6.2.1 claims that this matrix is invertible, that is, there exists the matrix M_N^{-1}.

7.3.2 Cauchy problem for a system of coupled quasilinear elliptic equations

Fix an integer $N \geq 1$. Denote $\Psi(N) = \{\psi_n(s)\}_{n=0}^{N-1}$. We approximate the function $v(x,s)$ in (7.11) via the truncated Fourier-like series with respect to the orthonormal basis of functions $\psi_n(s)$,

$$v(x,s) = \sum_{n=0}^{N-1} v_n(x)\psi_n(s), \quad x \in \Omega, \forall s \in [0,1]. \tag{7.15}$$

Here and below, we use "=" instead of "≈" for convenience. Note that the functions $v_n(x)$ are unknown and should be determined and they are targets of our considerations below.

Remarks 7.3.
1. We assume that the approximation (7.15) of the true function $v(x,s)$ satisfies equation (7.12). This is our *approximate mathematical model* mentioned in item 1 of Section 7.1. Once again: we can use this model since we work with a numerical method. The number N terms in (7.15) should be chosen in numerical experiments; see Section 7.6.1. Due to the ill posedness of the inverse EIT problem, the proof of convergence of our solutions to the exact one as $N \to \infty$ is a very challenging one and is, therefore, omitted here. In fact, it is well known that proofs of such results are quite challenging ones for many ill-posed problems. On the other hand, it is also well known that numerical methods developed for the cases of truncated Fourier-like series usually work quite well computationally; see, for example, publications [84, 108–112, 256].
2. The same as in item 1, is true for the approximate mathematical model (6.28) of Chapter 6, for approximate mathematical models of Chapters 11 and 12 as well as for [115–117, 145, 150, 152, 156, 236].
3. We refer to publications [117, 150] for more detailed discussions of the issue of approximate mathematical models.

The derivative $v_s(x,s)$ of the function $v(x,s)$ in (7.15) is

$$v_s(x,s) = \sum_{n=0}^{N-1} v_n(x)\psi_n'(s), \quad x \in \Omega, \forall s \in [0,1]. \tag{7.16}$$

By (7.3), (7.5), and (7.7) it is reasonable to assume that the functions

$$v_n \in C^3(\overline{\Omega}), \quad n = 0,\dots,N-1. \tag{7.17}$$

It is likely that (7.17) can be proven using the classical theory of elliptic PDEs [80]. However, we are not doing this here for brevity.

Substituting (7.15) and (7.16) in (7.14), we obtain

$$\sum_{n=0}^{N-1} \Delta v_n(x)\psi_n'(s) + \sum_{n,k=0}^{N-1} \nabla v_n(x)\nabla v_k(x)\psi_n'(s)\psi_k(s) = 0, \quad x \in \Omega, \forall s \in [0,1]. \tag{7.18}$$

Consider the vector function of unknown coefficient $v_n(x)$ in the expansion (7.15),

$$V(x) = \left(v_0(x),\dots,v_{N-1}(x)\right)^T. \tag{7.19}$$

For $m = 0,\dots,N-1$ multiply both sides of (7.18) by the function $\psi_m(s)$ and then integrate with respect to $s \in (0,1)$. Using (7.17) and (7.19), we obtain

$$M_N \Delta V - \widetilde{F}(\nabla V) = 0, \quad x \in \Omega, V \in C^3(\overline{\Omega}), \tag{7.20}$$

where the N-dimensional vector function $\widetilde{F}$ is quadratic with respect to the first derivatives $\partial_{x_j} v_k(x), j = 1,\dots,d; k = 0,\dots,N-1$. Multiplying both sides of (7.20) by the inverse matrix M_N^{-1} (Theorem 6.2.1), we obtain a system of coupled quasilinear elliptic equations,

$$\Delta V - F(\nabla V) = 0, \quad x \in \Omega, V \in C^3(\overline{\Omega}), \tag{7.21}$$

$$F(\nabla V) = M_N^{-1}\widetilde{F}(\nabla V). \tag{7.22}$$

Since the vector function $\widetilde{F}$ is quadratic with respect to the first derivatives $\partial_{x_j} v_k(x)$, then (7.22) implies that the vector function F is also quadratic. In addition, using (7.13), we obtain Cauchy data for the vector function $V(x)$ on Γ,

$$V(x)|_\Gamma = p_0(x), \quad \partial_\nu V(x)|_\Gamma = p_1(x). \tag{7.23}$$

If we would solve the Cauchy problem (7.21), (7.23), then we would find the coefficients $v_n(x)$ in (7.15). Next, we would substitute (7.15) in (7.12) and obtain the following approximate formula for the function $a_0(x)$:

$$a_0(x) = -\sum_{n=0}^{N-1} \Delta v_n(x)\psi_n(s) + \left(\sum_{n=0}^{N-1} \nabla v_n(x)\psi_n(s)\right)^2, \quad x \in \Omega, s \in (0,1). \tag{7.24}$$

As to the value of the parameter s for which the function $a_0(x)$ should be calculated in (7.24), it should be chosen numerically. Hence, we develop below a numerical method for solving problem (7.21), (7.23).

7.3.3 Two new Carleman estimates

Since in our numerical examples the domain $\Omega \subset \mathbb{R}^2$ is a disk, we prove in this subsection a new Carleman estimate for the Laplace operator, which is specifically used for the disk in the 2D case and for the ball in the 3D case. We work with the case when $\Gamma = \partial\Omega$ since this is done in our numerical experiments. In principle, Carleman estimates are known for this kind of domains; see, for example, [127]. However, the CWF in [127] has a rather complicated form and changes too rapidly. On the other hand, our extensive numerical experience with the convexification for CIPs [115–117, 142–146, 150, 151, 164] tells us that one should use a CWF of the simplest possible form. This is the reason of presenting here two new Carleman estimates with a simple CWF.

The 3D case

We derive in this section a new Carleman estimate for the 3D case when the domain Ω is a ball of the radius ρ,

$$\Omega = \{x \in \mathbb{R}^3 : |x| < \rho\}. \tag{7.25}$$

Let $\mu \in (0,\rho)$ be a number. Define the domain Ω_μ as

$$\Omega_\mu = \{x \in \mathbb{R}^3 : \mu < |x| < \rho\} \subset \Omega. \tag{7.26}$$

Consider spherical coordinates

$$r = |x| \in (\mu,\rho), \quad \varphi \in (0,2\pi), \quad \theta \in (0,\pi),$$
$$x_1 = r\cos\varphi\sin\theta, \quad x_2 = r\sin\varphi\sin\theta, \quad x_3 = r\cos\theta.$$

Also, denote

$$S_\rho = \{r = \rho\}, \quad S_\mu = \{r = \mu\}.$$

The Laplace operator in the spherical coordinates is

$$\Delta_{\mathrm{sp}} w = w_{rr} + \frac{1}{r^2\sin^2\theta} w_{\varphi\varphi} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta w_\theta) + \frac{2}{r}w_r = \widehat{\Delta}_{\mathrm{sp}} w + \frac{2}{r}w_r, \tag{7.27}$$

$$\widehat{\Delta}_{\mathrm{sp}} w = w_{rr} + \frac{1}{r^2\sin^2\theta} w_{\varphi\varphi} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta w_\theta), \tag{7.28}$$

for an arbitrary function $w \in C^2(\overline{\Omega}_\mu)$. We single out the operator $\widehat{\Delta}_{\mathrm{sp}}$ in (7.27), (7.28) since any Carleman estimate is independent on the low order derivatives of an operator, and also since we work in Ω_μ where $r > \mu > 0$. Everywhere below $C = C(\Omega_\mu) > 0$ denotes different constants depending only on the domain Ω_μ. Let

$$\nabla w = (w_{x_1}, w_{x_2}, w_{x_3})^T \quad \text{and} \quad \nabla_{\mathrm{sp}} w = \left(w_r, \frac{w_\varphi}{r\sin\theta}, \frac{w_\theta}{r}\right)^T.$$

Note that since $w_\varphi = -w_{x_1}\sin\varphi\sin\theta + w_{x_2}\cos\varphi\sin\theta$, then the function $w_\varphi/\sin\theta$ does not have a singularity. It is well known that

$$|\nabla w| \le C|\nabla_{\mathrm{sp}} w| \quad \text{in } \overline{\Omega}_\mu, \tag{7.29}$$

$$|\nabla_{\mathrm{sp}} w| \le C|\nabla w| \quad \text{in } \overline{\Omega}_\mu. \tag{7.30}$$

Introduce the subspace $H_0^m(\Omega_\mu)$ of the Hilbert space $H^m(\Omega_\mu)$ as

$$H_0^m(\Omega_\mu) = \{u \in H^m(\Omega_\mu) : u|_{S_\rho} = u_r|_{S_\rho} = 0\}, \quad m = 2, 3.$$

We include the term with $(\Delta w)^2$ in the right-hand side of the Carleman estimate (7.31) since we will need to estimate not only the convergence for the vector function $W(x)$ (Theorem 7.5.4), but also to estimate the convergence for the target coefficient $a_0(x)$ (Theorem 7.5.5). For this purpose, we will need to use equation (7.12) in which the Laplace operator is involved.

Theorem 7.3.1 (Carleman estimate). *There exists a number $\lambda_0 = \lambda_0(\Omega_\mu) \geq 1$ and a number $C = C(\Omega_\mu) > 0$, both depending only on the domain Ω_μ, such that for any function $w \in H^2(\Omega_\mu)$ and for all $\lambda \geq \lambda_0$ the following Carleman estimate with the CWF $e^{2\lambda r}$ holds:*

$$\int\limits_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx \geq \frac{1}{2} \int\limits_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx + C\lambda \int\limits_{\Omega_\mu} (\nabla w)^2 e^{2\lambda r} dx + C\lambda^3 \int\limits_{\Omega_\mu} w^2 e^{2\lambda r} dx \\ - C\lambda e^{2\lambda\rho} \int\limits_{S_\rho} w_r^2 dS - C\lambda^3 e^{2\lambda\rho} \int\limits_{S_\rho} w^2 dS - C\lambda^3 e^{2\lambda\mu} \|w\|^2_{H^2(\Omega_\mu)}. \tag{7.31}$$

In particular, if $w \in H_0^2(\Omega_\mu)$, then

$$\int\limits_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx \geq \frac{1}{2} \int\limits_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx + C\lambda \int\limits_{\Omega_\mu} (\nabla w)^2 e^{2\lambda r} dx + C\lambda^3 \int\limits_{\Omega_\mu} w^2 e^{2\lambda r} dx \\ - C\lambda^3 e^{2\lambda\mu} \|w\|^2_{H^2(\Omega_\mu)}. \tag{7.32}$$

Proof. We assume that $w \in C^2(\overline{\Omega}_\mu)$ since the case $w \in H^2(\Omega_\mu)$ can be handled automatically later via density arguments. Introduce the new function $q = we^{\lambda r}$. Then

$$w = qe^{-\lambda r}, \quad w_{rr} = (q_{rr} - 2\lambda q_r + \lambda^2 q)e^{-\lambda r}.$$

By (7.28),

$$\begin{aligned}
&(\widehat{\Delta}_{\rm sp} w)^2 e^{2\lambda r} r \sin\theta \\
&\quad = \left[(\widehat{\Delta}_{\rm sp} q + \lambda^2 q) - 2\lambda q_r\right]^2 r \sin\theta \\
&\quad \geq -4\lambda q_r \left(r\sin\theta q_{rr} + \frac{1}{r\sin\theta} q_{\varphi\varphi} + \frac{1}{r}\frac{\partial}{\partial\theta}(\sin\theta q_\theta) + \lambda^2 r \sin\theta q \right) \\
&\quad = \partial_r(-2\lambda r\sin\theta q_r^2) + 2\lambda\sin\theta q_r^2 + \partial_\varphi\left(-4\lambda\frac{q_r q_\varphi}{r\sin\theta}\right) + 4\lambda\frac{1}{r\sin\theta} q_{r\varphi} q_\varphi \\
&\qquad + \partial_\theta\left(-4\lambda\sin\theta\frac{q_r q_\theta}{r}\right) + 4\lambda\sin\theta\frac{q_{r\theta} q_\theta}{r} + \partial_r(-2\lambda^3 r\sin\theta q^2) + 2\lambda^3 r\sin\theta q^2 \\
&\quad = \partial_r\left(-2\lambda r\sin\theta q_r^2 - 2\lambda^3 r\sin\theta q^2 + 2\lambda\frac{1}{r\sin\theta} q_\varphi^2 + 2\lambda\frac{\sin\theta}{r} q_\theta^2\right) \\
&\qquad + \partial_\varphi\left(-4\lambda\frac{q_r q_\varphi}{r\sin\theta}\right) + \partial_\theta\left(-4\lambda\sin\theta\frac{q_r q_\theta}{r}\right) \\
&\qquad + 2\lambda\left(\sin\theta q_r^2 + \frac{q_\varphi^2}{\sin\theta r^2} + \sin\theta\frac{q_\theta^2}{r^2}\right) + \lambda^3 r\sin\theta q^2.
\end{aligned}$$

Hence, we have proven that

$$\begin{aligned}
&(\widehat{\Delta}_{\rm sp} w)^2 e^{2\lambda r} r \sin\theta \\
&\quad \geq 2\lambda\left(\sin\theta q_r^2 + \frac{q_\varphi^2}{\sin\theta r^2} + \sin\theta\frac{q_\theta^2}{r^2}\right) + 2\lambda^3 r\sin\theta q^2
\end{aligned}$$

$$+ \partial_r\left(-2\lambda r\sin\theta q_r^2 - 2\lambda^3 r\sin\theta q^2 + 2\lambda\frac{1}{r\sin\theta}q_\varphi^2 + 2\lambda\frac{\sin\theta}{r}q_\theta^2\right)$$
$$+ \partial_\varphi\left(-4\lambda\frac{q_r q_\varphi}{r\sin\theta}\right) + \partial_\theta\left(-4\lambda\sin\theta\frac{q_r q_\theta}{r}\right).$$

Integrate this inequality over Ω_μ while keeping in mind that the function $q(r,\varphi,\theta)$ is periodic with respect to φ with the period 2π and that $\sin 0 = \sin\pi = 0$ and also that $dx = r^2\sin\theta dr d\varphi d\theta$. We obtain

$$\begin{aligned}
&\int_{\Omega_\mu}(\widehat{\Delta}_{\mathrm{sp}}w)^2 e^{2\lambda r}dx \\
&\geq C\int_{\Omega_\mu}(\widehat{\Delta}_{\mathrm{sp}}w)^2\frac{e^{2\lambda r}}{r}dx = \int_{\Omega_\mu}(\widehat{\Delta}_{\mathrm{sp}}w)^2 e^{2\lambda r}r\sin\theta dr d\varphi d\theta \\
&\geq 2\lambda\int_{\Omega_\mu}\left(q_r^2 + \frac{q_\varphi^2}{\sin^2\theta r^2} + \frac{q_\theta^2}{r^2}\right)\sin\theta dr d\varphi d\theta + 2\lambda^3\int_{\Omega_\mu}q^2 r\sin\theta dr d\varphi d\theta \\
&\quad - \int_{S_\rho}(2\lambda q_r^2 + 2\lambda^3 q^2)dS - C\lambda e^{2\lambda\mu}\int_{S_\mu}(\nabla w)^2 dS.
\end{aligned} \tag{7.33}$$

Change variables back from q to w. Since $q = we^{\lambda r}$, then

$$\begin{aligned}
q_r^2 &= (w_r + \lambda w)^2 e^{2\lambda r} = w_r^2 e^{2\lambda r} + 2\lambda w_r w e^{2\lambda r} + \lambda^2 w^2 e^{2\lambda r} \\
&= w_r^2 e^{2\lambda r} + \partial_r(\lambda w^2 e^{2\lambda r}) - 2\lambda^2 w^2 e^{2\lambda r} + \lambda^2 w^2 e^{2\lambda r} \\
&= w_r^2 e^{2\lambda r} - \lambda^2 w^2 e^{2\lambda r} + \partial_r(\lambda w^2 e^{2\lambda r}).
\end{aligned}$$

Let the number $a = \min(\mu/2, 1)$. Then in the first line of (7.33),

$$\begin{aligned}
2\lambda q_r^2 + 2\lambda^3 q^2 r &\geq 2\lambda a q_r^2 + 2\lambda^3 q^2\mu \\
&\geq 2\lambda a w_r^2 e^{2\lambda r} + 2\lambda^3\left(\mu - \frac{\mu}{2}\right)w^2 e^{2\lambda r} + \partial_r(2\lambda^2 a w^2 e^{2\lambda r}) \\
&\geq C\lambda w_r^2 e^{2\lambda r} + C\lambda^3 w^2 e^{2\lambda r} + \partial_r(2\lambda^2 a w^2 e^{2\lambda r}).
\end{aligned} \tag{7.34}$$

Hence, using (7.29), (7.33), and (7.34), we obtain

$$\begin{aligned}
\int_{\Omega_\mu}(\widehat{\Delta}_{\mathrm{sp}}w)^2 e^{2\lambda r}dx &\geq C\lambda\int_{\Omega_\mu}(\nabla w)^2 e^{2\lambda r}dx + C\lambda^3\int_{\Omega_\mu}w^2 e^{2\lambda r}dx \\
&\quad - C\lambda e^{2\lambda\rho}\int_{S_\rho}w_r^2 dS - C\lambda^3 e^{2\lambda\rho}\int_{S_\rho}w^2 dS - C\lambda^3 e^{2\lambda\mu}\int_{S_\mu}((\nabla w)^2 + w^2)dS.
\end{aligned} \tag{7.35}$$

Noticing that by (7.27), we have

$$(\Delta w)^2 = (\widehat{\Delta}_{\mathrm{sp}}w)^2 + 4(\widehat{\Delta}_{\mathrm{sp}}w)\frac{w_r}{r} + 4\left(\frac{w_r}{r}\right)^2 \geq \frac{1}{2}(\widehat{\Delta}_{\mathrm{sp}}w)^2 - Cw_r^2,$$

and also that

$$\int_{S_\mu} ((\nabla w)^2 + w^2) dS \le C\|w\|^2_{H^2(\Omega_\mu)},$$

and then using (7.35), we obtain

$$\int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx \ge C\lambda \int_{\Omega_\mu} (\nabla w)^2 e^{2\lambda r} dx + C\lambda^3 \int_{\Omega_\mu} w^2 e^{2\lambda r} dx - C\lambda e^{2\lambda\rho} \int_{S_\rho} w_r^2 dS - C\lambda^3 e^{2\lambda\rho} \int_{S_\rho} w^2 dS - C\lambda^3 e^{2\lambda\mu} \|w\|^2_{H^2(\Omega_\mu)}. \tag{7.36}$$

Obviously,

$$\int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx = \int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx. \tag{7.37}$$

Summing up (7.36) with (7.37) and then dividing the resulting estimate by 2, we obtain (7.31). □

The 2D case

In this case we keep the same notations for the domains Ω, Ω_μ as those in Subsection 7.3.3. We assume that they are the domains in $\mathbb{R}^2$. Polar coordinates are

$$r = |x| \in (\mu, \rho), \quad \varphi \in (0, 2\pi),$$
$$x_1 = r\cos\varphi, \quad x_2 = r\sin\varphi.$$

The Laplace operator in polar coordinates is

$$\Delta_p w = w_{rr} + \frac{1}{r^2} w_{\varphi\varphi} + \frac{1}{r} w_r = \widehat{\Delta}_p w + \frac{1}{r} w_r.$$

Theorem 7.3.2 (Carleman estimate). *There exists a number $\lambda_0 = \lambda_0(\Omega_\mu) \ge 1$ depending only on the domain Ω_μ such that for any function $w \in H^2(\Omega_\mu)$ and for all $\lambda \ge \lambda_0$ the following Carleman estimate holds:*

$$\int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx \ge \frac{1}{2} \int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx + C\lambda \int_{\Omega_\mu} (\nabla w)^2 e^{2\lambda r} dx + C\lambda^3 \int_{\Omega_\mu} w^2 e^{2\lambda r} dx - C\lambda e^{2\lambda\rho} \int_{S_\rho} w_r^2 dS - C\lambda^3 e^{2\lambda\rho} \int_{S_\rho} w^2 dS - C\lambda e^{2\lambda\mu} \|w\|^2_{H^2(\Omega_\mu)}.$$

In particular, if $w \in H_0^2(\Omega_\mu)$, then

$$\int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx \ge \frac{1}{2} \int_{\Omega_\mu} (\Delta w)^2 e^{2\lambda r} dx + C\lambda \int_{\Omega_\mu} (\nabla w)^2 e^{2\lambda r} dx + C\lambda^3 \int_{\Omega_\mu} w^2 e^{2\lambda r} dx - C\lambda^3 e^{2\lambda\mu} \|w\|^2_{H^2(\Omega_\mu)}.$$

The proof of this theorem is omitted since it is very similar with the proof of Theorem 7.3.1.

7.3.4 Hölder stability and uniqueness of the Cauchy problem (7.21), (7.23)

We establish in this subsection the Hölder stability estimate for problem (7.21), (7.23). Uniqueness follows immediately from this estimate. We work here only with the 3D case. Theorem 7.3.2 implies that the 2D case can be handled almost exactly the same way. Thus, in this subsection the domain Ω is as in (7.25), and in (7.23) $\Gamma = \partial\Omega = \{r = \rho\}$. Everywhere below we often work with N-dimensional vector functions, like, for example, $V(x)$. Norms in standard functional spaces of such vector functions are defined in the natural well-known way via corresponding norms of their components. The same about scalar products. It is always clear from the context whether we work with regular functions or with those N-dimensional vector functions.

Suppose that there exist two solutions of problem (7.21), (7.23), $V_1, V_2 \in H^2(\Omega) \cap C^1(\overline{\Omega})$ such that

$$V_1|_{S_\rho} = p_0, \quad V_2|_{S_\rho} = p_{0,\delta}, \quad V_{1,r}|_{S_\rho} = p_1, \quad V_{2,r}|_{S_\rho} = p_{1,\delta}, \tag{7.38}$$

where

$$\|p_{0,\delta} - p_0\|_{H^1(S_\rho)} \le \delta, \quad \|p_{1,\delta} - p_1\|_{L^2(S_\rho)} \le \delta, \tag{7.39}$$

where $\delta \in (0,1)$ is a sufficiently small number which is interpreted as the level of the noise in the data. Denote

$$\widetilde{V} = V_1 - V_2, \quad \widetilde{p} = p_0 - p_{0,\delta}, \quad \widetilde{p}_1 = p_1 - p_{1,\delta}. \tag{7.40}$$

Recalling that the function F in (7.21) is quadratic with respect to the derivatives $\partial_{x_j} v_k(x)$, we obtain from (7.21), (7.23), and (7.38)–(7.40),

$$\Delta\widetilde{V} = Q(\nabla V_1, \nabla V_2)\nabla\widetilde{V}, \quad x \in \Omega, \widetilde{V} \in H^2(\Omega) \cap C^1(\overline{\Omega}), \tag{7.41}$$

$$\widetilde{V}|_{S_\rho} = \widetilde{p}_0, \quad \widetilde{V}_r|_{S_\rho} = \widetilde{p}_1, \tag{7.42}$$

$$\|\widetilde{p}_0\|_{H^1(S_\rho)} \le \delta, \quad \|\widetilde{p}_1\|_{L^2(S_\rho)} \le \delta, \tag{7.43}$$

where the vector function $Q(\nabla V_1, \nabla V_2)$ is linear with respect to components of the vector functions ∇V_1, ∇V_2.

Theorem 7.3.3 (Hölder stability estimate). *For any two vector functions*

$$V_1, V_2 \in H^2(\Omega) \cap C^1(\overline{\Omega})$$

introduced in this section, let $\|V_1\|_{C^1(\overline{\Omega})}, \|V_2\|_{C^1(\overline{\Omega})} \le A$, *where* $A = \text{const.} > 0$. *Let* (7.38)–(7.40) *hold. Choose a number* $\eta \in (0, \rho-\mu)$ *and let* $\Omega_{\mu+\eta} = \{x : \mu+\eta < |x| < \rho\} \subset \Omega_\mu$.

Then there exists a number $C_1 = C_1(\Omega_\mu, \eta, F, \Psi(N), A) > 0$ *and a sufficiently small number* $\delta_0 = \delta_0(\Omega_\mu, \eta, F, \Psi(N), A) \in (0,1)$ *such that, for all* $\delta \in (0, \delta_0)$ *the following Hölder stability estimate holds:*

$$\|\widetilde{V}\|_{H^1(\Omega_{\mu+\eta})} \le C_1 \delta^\gamma, \quad \text{where } \gamma = \eta/(4\rho). \tag{7.44}$$

Proof. In this proof, $C_1 = C_1(\Omega_\mu, \eta, F, \Psi(N), A) > 0$ denotes different positive constants depending only on listed parameters. A careful analysis of the proof of Theorem 7.3.1, more precisely of the last term in the second line of (7.35), shows that the term $\|w\|^2_{H^2(\Omega_\mu)}$ in (7.31) can be replaced with the term $\|w\|^2_{C^1(S_\mu)}$. Note that $|Q(\nabla V_1, \nabla V_2)| \le C_1$. Hence, squaring both sides of (7.41), replacing the equality with the inequality, we obtain

$$(\Delta \widetilde{V})^2 \le C_1 (\nabla \widetilde{V})^2, \quad x \in \Omega_\mu. \tag{7.45}$$

The rest of the proof uses (7.42), (7.43), (7.45), the Carleman estimate of Theorem 7.3.1 in the way and proceeds very similarly with proofs of the Hölder stability estimates in Sections 2.2, 2.6.2, and 6.3. Thus, we omit the rest of the proof for brevity. □

7.4 Convexification

To solve the Cauchy problem (7.21), (7.23) numerically, we construct in this section a weighted Tikhonov-like functional with the CWF $e^{2\lambda r}$ in it and prove necessary theorems. For brevity, we construct the Tikhonov-like functional only for the 3D case. So, in Sections 7.4 and 7.5 we work only with the 3D case. The 2D case is completely similar and direct analogs of Theorems 7.5.1–7.5.4 (below) are valid in 2D.

We assume that in (7.23)

$$\Gamma = \partial\Omega = S_\rho; \quad p_0, p_1 \in C^3(S_\rho). \tag{7.46}$$

We now arrange zero Dirichlet and Neumann boundary conditions for a new vector function W, which is associated with the vector function V. We are doing so since we use below some theorems of [9], which are applicable only in the case of zero boundary conditions.

Denote

$$P(r, \varphi, \theta) = p_0(r, \varphi, \theta) + (r - \rho) p_1(r, \varphi, \theta), \tag{7.47}$$

$$W(r, \varphi, \theta) = V(r, \varphi, \theta) - P(r, \varphi, \theta); \quad W(r, \varphi, \theta) = (W_0, \dots, W_{N-1})^T (r, \varphi, \theta). \tag{7.48}$$

Then by (7.46) $P \in C^3(\overline{\Omega}_\mu)$. Hence, (7.21), (7.23), (7.47), and (7.48) imply that

$$\Delta W + \Delta P - F(\nabla W + \nabla P) = 0, \tag{7.49}$$

$$W \in H_0^3(\Omega_\mu). \tag{7.50}$$

Note that by the embedding theorem

$$H^3(\Omega_\mu) \subset C^1(\overline{\Omega}_\mu), \quad \|f\|_{C^1(\overline{\Omega}_\mu)} \leq C\|f\|_{H^3(\Omega_\mu)}. \tag{7.51}$$

Let $\eta \in (0, \rho - \mu)$ be the number which was chosen in Theorem 7.3.3. Our weighted Tikhonov-like functional is

$$J_{\lambda,\beta}(W) \\ = e^{-2\lambda(\mu+\eta)} \int\limits_{\Omega_\mu} [\Delta W + \Delta P - F(\nabla W + \nabla P)]^2 e^{2\lambda r} dx + \beta\|W + P\|^2_{H^3(\Omega_\mu)}. \tag{7.52}$$

Here, $\beta \in (0,1)$ is the regularization parameter and the multiplier $e^{-2\lambda(\mu+\eta)}$ is introduced to balance two terms in the right-hand side of (7.52). Let $R > 0$ be an arbitrary number. Let $B(R) \subset H_0^3(\Omega_\mu)$ be the ball of the radius R with the center at $\{0\}$,

$$B(R) = \{W \in H_0^3(\Omega_\mu) : \|W\|_{H^3(\Omega_\mu)} < R\}. \tag{7.53}$$

We consider the following minimization problem:

Minimization problem. Minimize the functional $J_{\lambda,\beta}(W)$ on the closed ball $\overline{B(R)}$.

7.5 Theorems

In this section, we formulate and prove some theorems about the above minimization problem.

7.5.1 Formulations of theorems

The central analytical result of this chapter is Theorem 7.5.1.

Theorem 7.5.1. *The functional $J_{\lambda,\beta}(W)$ has the Frechét derivative $J'_{\lambda,\beta}(W)$ at every point $W \in H_0^3(\Omega_\mu)$. Furthermore, there exist numbers $\lambda_2 = \lambda_2(\mu, \eta, F, \Psi(N), P, R) \geq \lambda_0 > 0$ and $C_2 = C_2(\mu, \eta, F, \Psi(N), P, R) > 0$ depending only on listed parameters such that $2e^{-\lambda_2\eta} < 1$ and for all $\lambda \geq \lambda_2$ the functional $J_{\lambda,\beta}(W)$ is strictly convex on $\overline{B(R)}$ for the choice of β as*

$$\beta \in (2e^{-\lambda\eta}, 1). \tag{7.54}$$

More precisely, the following inequality holds:

$$J_{\lambda,\beta}(W_2) - J_{\lambda,\beta}(W_1) - J'_{\lambda,\beta}(W_1)(W_2 - W_1) \\ \geq C_2\|\Delta(W_2 - W_1)\|_{L^2(\Omega_{\mu+\eta})} + C_2\|W_2 - W_1\|^2_{H^1(\Omega_{\mu+\eta})} + \frac{\beta}{2}\|W_2 - W_1\|^2_{H^3(\Omega_\mu)}, \tag{7.55} \\ \forall W_1, W_2 \in \overline{B(R)}.$$

Note that, allowing the regularization parameter $\beta \in (2e^{-\lambda\eta}, 1)$, we actually allow β to be sufficiently small. We now formulate the theorem about the Lipschitz continuity condition of the Frechét derivative $J'_{\lambda,\beta}(W)$.

Theorem 7.5.2. *For any numbers $\widetilde{R}, \lambda > 0, \beta \in (0,1)$ the Frechét derivative $J'_{\lambda,\beta}(W)$ of the functional $J_{\lambda,\beta}(W)$ satisfies the Lipschitz continuity condition in the ball $B(\widetilde{R})$. In other words, there exists a number $Z = Z(\Omega_\mu, F, \Psi(N), \widetilde{R}, \lambda) > 0$ depending only on the listed parameters such that*

$$\left|J'_{\lambda,\beta}(W_2) - J'_{\lambda,\beta}(W_1)\right| \le Z\|W_2 - W_1\|_{H^3(\Omega_\mu)}, \quad \forall W_1, W_2 \in B(\widetilde{R}).$$

Consider now the gradient projection method of the minimization of the functional $J_{\lambda,\beta}$ on the closed ball $\overline{B(R)}$. Let $P_B : H_0^3(\Omega_\mu) \to \overline{B(R)}$ be the projection operator of the space $H_0^3(\Omega_\mu)$ onto the closed ball $\overline{B(R)} \subset H_0^3(\Omega_\mu)$. Let $W_0 \in B(R)$ be an arbitrary point. The sequence $\{W_n\}_{n=1}^{\infty}$ of the gradient projection method is defined as

$$W_n = P_B(W_{n-1} - \zeta J'_{\lambda,\beta}(W_{n-1})), \quad n = 1, 2, \ldots, \tag{7.56}$$

where $\zeta \in (0,1)$ is a sufficiently small number. Below $[\cdot,\cdot]$ denotes the scalar product in the space of real valued N-D vector functions $H^3(\Omega_\mu)$.

Theorem 7.5.3. *Let $\lambda_2 = \lambda_2(\mu, \eta, F, N, P, R) \ge \lambda_0 > 0$ be the number of Theorem 7.5.1 and let the regularization parameter $\beta \in (2e^{-\lambda\eta}, 1)$. Then for every $\lambda \ge \lambda_2$ there exists a unique minimizer $W_{\min,\lambda,\beta} \in \overline{B(R)}$ of the functional $J_{\lambda,\beta}(W)$ on the closed ball $\overline{B(R)}$. Furthermore, the following inequality holds:*

$$\left[J'_{\lambda,\beta}(W_{\min,\lambda,\beta}), W - W_{\min,\lambda,\beta}\right] \ge 0, \quad \forall W \in \overline{B(R)}. \tag{7.57}$$

In addition, there exists a sufficiently small number $\zeta_0 = \zeta_0(\mu, \eta, F, \Psi(N), P, R, \lambda, \beta) \in (0,1)$ depending only on listed parameters such that for every $\zeta \in (0, \zeta_0)$ the sequence (7.56) converges to the minimizer $W_{\min,\lambda,\beta}$ and the following estimate of the convergence rate holds:

$$\|W_n - W_{\min,\lambda,\beta}\|_{H^3(\Omega)} \le \omega^n \|W_0 - W_{\min,\lambda,\beta}\|_{H^3(\Omega)}, \quad n = 1, 2, \ldots, \tag{7.58}$$

where the number $\omega = \omega(\zeta) \in (0,1)$ depends only on the parameter ζ.

Even though Theorem 7.5.3 guarantees the convergence of the gradient projection method to the unique minimizer of the functional (7.52), it is not yet clear how far this minimizer is from the exact solution. To address this question, we assume, as it is commonly accepted in the theory of ill-posed problems [244], that there exists an exact solution $W^* \in B(R)$ of the problem (7.49), (7.50), that is, solution with the noiseless data.

Let $\delta \in (0,1)$ be a sufficiently small number characterizing the level of the noise in the data. Let W^* be the exact solution of problem (7.49), (7.50) with the noiseless data $P^* \in C^3(\overline{\Omega}_\mu)$,

$$\Delta W^* + \Delta P^* - F(\nabla W^* + \nabla P^*) = 0, \tag{7.59}$$

$$W^* \in H_0^3(\Omega_\mu). \tag{7.60}$$

Let $P \in C^3(\overline{\Omega}_\mu)$ be the noisy data. Denote $\widetilde{P} = P - P^*$. We assume that

$$\|\widetilde{P}\|_{H^3(\overline{\Omega}_\mu)} \leq \delta. \tag{7.61}$$

Theorem 7.5.4. *Let $\lambda_2 \geq \lambda_0 > 0$ and $C_2 > 0$ be numbers of Theorem 7.5.1. Choose the number $\delta_1 > 0$ so small that $\delta_1 < \min(e^{-4\rho\lambda_2}, 3^{-4\rho/\eta})$ and let $\delta \in (0,\delta_1)$. Set $\lambda = \lambda(\delta) = \ln \delta^{-1/(4\rho)}$, $\beta = \beta(\delta) = 3\delta^{\eta/(4\rho)}$. Let (7.61) be true. Also, assume that the vector function $W^* \in B(R)$. Let $W_{\min,\lambda(\delta),\beta(\delta)} \in \overline{B(R)}$ be the minimizer of the functional (7.52), which is guaranteed by Theorem 7.5.3. Also, let the number $\zeta \in (0,\zeta_0)$ in (7.56) be the same as in Theorem 7.5.3, so as the number $\omega \in (0,1)$. Then the following estimates hold:*

$$\|W^* - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^1(\Omega_{\mu+\eta})} \leq C_2\delta^{\eta/(8\rho)}, \tag{7.62}$$

$$\|\Delta W^* - \Delta W_n\|_{L^2(\Omega_{\mu+\eta})} \leq C_2\delta^{\eta/(8\rho)}, \tag{7.63}$$

$$\|W^* - W_n\|_{H^1(\Omega_{\mu+\eta})} \leq C_2\delta^{\eta/(8\rho)} + \omega^n\|W_0 - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^3(\Omega)}, \quad n = 1,2,\ldots, \tag{7.64}$$

$$\|\Delta W^* - \Delta W_n\|_{L^2(\Omega_{\mu+\eta})} \leq C_2\delta^{\eta/(8\rho)} + \omega^n\|W_0 - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^3(\Omega)}, \quad n = 1,2,\ldots. \tag{7.65}$$

In the case of noiseless data with $\delta = 0$ one should replace in (7.62) and (7.64) $\delta^{\eta/(8\rho)}$ with $\sqrt{\beta}$, where $\beta = 3e^{-\lambda\eta}$ and $\lambda \geq \lambda_2$.

While (7.62)–(7.65) are convergence estimates for the vector function $W^*(x)$, we still need to obtain a convergence estimate for our target coefficient $a_0(x)$ in equation (7.12). This is done in Theorem 7.5.5. Let $V^*(x) = W^*(x) + P^*(x)$. Then $V^*(x) = (v_0^*(x),\ldots,v_{N-1}^*(x))^T$. Let $a_0^*(x)$ be the exact coefficient $a_0(x)$ which corresponds to $V^*(x)$ via (7.24), that is,

$$a_0^*(x) = -\sum_{k=0}^{N-1} \Delta v_k^*(x)\psi_k(s) + \left(\sum_{k=0}^{N-1} \nabla v_k^*(x)\psi_k(s)\right)^2, \quad x \in \Omega, s \in (0,1). \tag{7.66}$$

Next, let

$$V_{\min,\lambda(\delta),\beta(\delta)}(x) = W_{\min,\lambda(\delta),\beta(\delta)}(x) + P(x) = (v_{0,\min,\lambda(\delta),\beta(\delta)}(x),\ldots,v_{N-1,\min,\lambda(\delta),\beta(\delta)}(x))^T$$

and let

$$a_{0,\min,\lambda(\delta),\beta(\delta)}(x) = -\sum_{k=0}^{N-1} \Delta v_{k,\min,\lambda(\delta),\beta(\delta)}(x)\psi_k(\overline{s}) + \left(\sum_{k=0}^{N-1} \nabla v_{k,\min,\lambda(\delta),\beta(\delta)}(x)\psi_k(\overline{s})\right)^2, \quad x \in \Omega, \overline{s} \in (0,1). \tag{7.67}$$

Let $V_n(x) = W_n(x) + P(x)$, where the sequence $\{W_n\}_{n=0}^{\infty}$ is defined in (7.56). Then $V_n(x) = (v_0^{(n)}(x), \dots, v_{N-1}^{(n)}(x))^T$. Define the function $a_{0,n}(x)$ as

$$a_{0,n}(x) = -\sum_{k=0}^{N-1} \Delta v_k^{(n)}(x)\psi_k(\overline{s}) + \left(\sum_{k=0}^{N-1} \nabla v_k^{(n)}(x)\psi_k(\overline{s})\right)^2, \quad x \in \Omega, \overline{s} \in (0,1), \tag{7.68}$$

where $\overline{s}$ is a certain fixed number.

Theorem 7.5.5. *Assume that the conditions of Theorem* 7.5.4 *hold. Then the following analogs of estimates* (7.62)–(7.65) *are in place:*

$$\|a_0^* - a_{0,\min,\lambda(\delta),\beta(\delta)}\|_{L^2(\Omega_{\mu+\eta})} \le C_2 \delta^{\eta/(8\rho)}, \tag{7.69}$$

$$\|a_0^* - a_{0,n}\|_{L^2(\Omega_{\mu+\eta})} \le C_2 \delta^{\eta/(8\rho)} + \omega^n \|W_0 - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^3(\Omega)}, \quad n = 1, 2, \dots, \tag{7.70}$$

where the functions a_0^*, $a_{0,\min,\lambda(\delta),\beta(\delta)}$ and $a_{0,n}$ are defined in (7.66)–(7.68).

Remarks 7.5.

1. Theorems 7.5.4 and 7.5.5 guarantee that a sufficiently small neighborhood of the exact solution is reached if the gradient projection method starts from an arbitrary point of the ball $B(R)$. Since the radius R of this ball is an arbitrary one, then this is the *global convergence*; see Definition 1.4.2.
2. The proof of Theorem 7.5.2 is quite similar with the proof of Theorem 5.3.1. Theorem 7.5.3 follows immediately from a combination of Theorems 7.5.1 and 7.5.2 with Lemma 5.2.1 and Theorem 5.2.1. Thus, we omit proofs of Theorems 7.5.2 and 7.5.3 and focus only on Theorems 7.5.1, 7.5.4, and 7.5.5. In the proofs below, $C_2 = C_2(\mu, \eta, F, \Psi(N), P, R) > 0$ denotes different constants depending only on the listed parameters.

7.5.2 Proof of Theorem 7.5.1

Let $W_1, W_2 \in \overline{B(R)}$ be two arbitrary points. Denote $h = W_2 - W_1$. Hence, $W_2 = W_1 + h$. By the triangle inequality and (7.53)

$$\|h\|_{H^3(\Omega_\mu)} \le 2R. \tag{7.71}$$

We have

$$
\begin{aligned}
&[\Delta W_1 + \Delta h - F(\nabla W_1 + \nabla P + \nabla h) + \Delta P]^2 - [\Delta W_1 - F(\nabla W_1 + \nabla P) + \Delta P]^2 \\
&\quad = [\Delta h - (F(\nabla W_1 + \nabla P + \nabla h) - F(\nabla W_1 + \nabla P))] \\
&\qquad \times [\Delta h + 2\Delta W_1 - F(\nabla W_1 + \nabla P + \nabla h) - F(\nabla W_1 + \nabla P) + 2\Delta P].
\end{aligned} \tag{7.72}
$$

Recall that the vector function $F(\nabla W + \nabla P)$ is quadratic with respect to the derivatives $\partial_{x_j} W_k(x)$, $j = 1, 2, 3$; $k = 0, \dots, N-1$. Hence, (7.72) implies that

$$
\begin{aligned}
&[\Delta W_2 - F(\nabla W_2 + \nabla P) + \Delta P]^2 - [\Delta W_1 - F(\nabla W_1 + \nabla P) + \Delta P]^2 \\
&\quad = \Delta h[Q_1(\nabla W_1 + \nabla P) + 2(\Delta W_1 + \Delta P)] + \nabla h[Q_2(\nabla W_1 + \nabla P, \Delta W_1 + \Delta P)] \\
&\qquad + (\Delta h)^2 + \Delta h D_1(\nabla W_1 + \nabla P, \nabla h) + D_2(\nabla W_1 + \nabla P, \nabla h).
\end{aligned} \tag{7.73}
$$

In (7.73), vector functions Q_1, Q_2, D_1, D_2 are continuous with respect to their indicated variables. In addition, (7.51) and (7.71) imply that the following estimates are valid for vector functions $D_1(\nabla W_1 + \nabla P, \nabla h)$, $D_2(\nabla W_1 + \nabla P, \nabla h)$:

$$
|D_1(\nabla W_1 + \nabla P, \nabla h)| \le C_2(|\nabla h| + |\nabla h|^2), \tag{7.74}
$$

$$
|D_2(\nabla W_1 + \nabla P, \nabla h)| \le C_2|\nabla h|^2, \quad j = 1, 2. \tag{7.75}
$$

In the second line of (7.73), we single out the part which is linear with respect to h. On the other hand, using (7.51), (7.74), (7.75), and the Cauchy–Schwarz inequality, we obtain the following estimate for the expression in the third line of (7.73):

$$
(\Delta h)^2 + \Delta h D_1(\nabla W_1 + \nabla P, \nabla h) + D_2(\nabla W_1 + \nabla P, \nabla h) \ge \frac{1}{2}(\Delta h)^2 - C_2(\nabla h)^2. \tag{7.76}
$$

In addition, the following estimate from the above follows from (7.51), (7.73), (7.74), and (7.75):

$$
\begin{aligned}
&|(\Delta h)^2 + \Delta h D_1(\nabla W_1 + \nabla P, \nabla h) + D_2(\nabla W_1 + \nabla P, \nabla h)| \\
&\quad \le C_2[(\Delta h)^2 + (\nabla h)^2].
\end{aligned} \tag{7.77}
$$

Thus, (7.52) and (7.73) imply that

$$
\begin{aligned}
&J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) \\
&\quad = e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} \{\Delta h[Q_1(\nabla W_1 + \nabla P) + 2\Delta P] + \nabla h[Q_2(\nabla W_1 + \nabla P, \Delta P)]\} e^{2\lambda r} dx \\
&\qquad + 2\beta[h, W_1] \\
&\qquad + e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} [(\Delta h)^2 + \Delta h D_1(\nabla W_1 + \nabla P, \nabla h) + D_2(\nabla W_1 + \nabla P, \nabla h)] e^{2\lambda r} dx \\
&\qquad + \beta \|h\|^2_{H^3(\Omega_\mu)}.
\end{aligned} \tag{7.78}
$$

The expression in the second line of (7.78) is generated by the second line of (7.73), and it is linear with respect to h. Actually, the sum of the second and third lines of (7.78) is a linear functional with respect to h, and we denote it $\mathrm{Lin}(W_1)(h)$. In addition, the following estimate holds:

$$\left|\mathrm{Lin}(W_1)(h)\right| \le C_2 \exp\left(2\lambda(\rho - \mu - \eta)\right)\|h\|_{H^3(\Omega_\mu)}.$$

Hence, $\mathrm{Lin}(W_1)(h) : H_0^3(\Omega_\mu) \to \mathbb{R}$ is a bounded linear functional with respect to h. Hence, by Riesz theorem there exists a vector function $Y(x) \in H_0^3(\Omega_\mu)$ such that

$$\mathrm{Lin}(W_1)(h) = [Y, h]. \tag{7.79}$$

Also, it follows from (7.74) and (7.78) that if $\|h\|_{H^3(\Omega_\mu)} < 1$, then the following estimate holds:

$$\left|J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) - \mathrm{Lin}(W_1)(h)\right| \le C_2 \exp\left(2\lambda(\rho - \mu - \eta)\right)\|h\|^2_{H^3(\Omega_\mu)}. \tag{7.80}$$

Thus, using (7.79) and (7.80), we obtain that the Frechét derivative $J'_{\lambda,\beta}(W_1)$ of the functional $J_{\lambda,\beta}(W)$ exists at the point W_1 and $J'_{\lambda,\beta}(W_1) = Y(x)$. Even though the existence of the Frechét derivative $J'_{\lambda,\beta}(W_1)$ is proved here only for the case when W_1 is an interior point of the ball $B(R)$, still since $R > 0$ is an arbitrary number, then actually this existence is proved for an arbitrary point $W_1 \in H_0^3(\Omega_\mu)$.

We now need to prove the strict convexity estimate (7.55). To do this, we will use the Carleman estimate of Theorem 7.3.1. Using (7.76) and (7.78), we obtain

$$\begin{aligned} &J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) - J'_{\lambda,\beta}(W_1)(h) \\ &\quad \ge \frac{1}{2} e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\Delta h)^2 e^{2\lambda r}\, dx - C_2 e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\nabla h)^2 e^{2\lambda r}\, dx + \beta\|h\|^2_{H^3(\Omega_\mu)}. \end{aligned} \tag{7.81}$$

Next, using (7.32), we obtain from (7.81)

$$\begin{aligned} &J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) - J'_{\lambda,\beta}(W_1)(h) \\ &\quad \ge \frac{1}{4} e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\Delta h)^2 e^{2\lambda r}\, dx + C\lambda e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\nabla h)^2 e^{2\lambda r}\, dx \\ &\quad - C_2 e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\nabla h)^2 e^{2\lambda r}\, dx + C\lambda^3 e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} h^2 e^{2\lambda r}\, dx \\ &\quad - C\lambda^3 e^{-2\lambda\eta}\|h\|^2_{H^2(\Omega_\mu)} + \beta\|h\|^2_{H^3(\Omega_\mu)}. \end{aligned} \tag{7.82}$$

Choose $\lambda_2 = \lambda_2(\mu, \eta, F, N, P, R) \ge \lambda_0 > 0$ so large that $C\lambda_2 > 2C_2$ and also that $C\lambda^3 e^{-2\lambda\eta} < e^{-\lambda\eta}$, $\forall \lambda \ge \lambda_2$. Recalling (7.54) and using $\Omega_{\mu+\eta} \subset \Omega_\mu$, we obtain from (7.82)

$$\begin{aligned} &J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) - J'_{\lambda,\beta}(W_1)(h) \\ &\quad \ge \frac{1}{4} e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\Delta h)^2 e^{2\lambda r}\, dx \end{aligned}$$

$$+ C_2 e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\nabla h)^2 e^{2\lambda r} dx + C\lambda^3 e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} h^2 e^{2\lambda r} dx + \frac{\beta}{2} \|h\|^2_{H^3(\Omega_\mu)} \tag{7.83}$$
$$\geq \frac{1}{4} e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\Delta h)^2 e^{2\lambda r} dx + C_2 e^{-2\lambda(\mu+\eta)} \int_{\Omega_{\mu+\eta}} [(\nabla h)^2 + h^2] e^{2\lambda r} dx + \frac{\beta}{2} \|h\|^2_{H^3(\Omega_\mu)}.$$

Next, $e^{2\lambda r} \geq e^{2\lambda(\mu+\eta)}$ for $x \in \Omega_{\mu+\eta}$. Hence,

$$\frac{1}{4} e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} (\Delta h)^2 e^{2\lambda r} dx + e^{-2\lambda(\mu+\eta)} \int_{\Omega_{\mu+\eta}} [(\nabla h)^2 + h^2] e^{2\lambda r} dx \tag{7.84}$$
$$\geq \frac{1}{4} \int_{\Omega_{\mu+\eta}} (\Delta h)^2 dx + \int_{\Omega_{\mu+\eta}} [(\nabla h)^2 + h^2] dx.$$

Thus, (7.83) and (7.84) imply that

$$J_{\lambda,\beta}(W_1 + h) - J_{\lambda,\beta}(W_1) - J'_{\lambda,\beta}(W_1)(h) \tag{7.85}$$
$$\geq \frac{1}{4} \int_{\Omega_{\mu+\eta}} (\Delta h)^2 dx + C_2 \int_{\Omega_{\mu+\eta}} [(\nabla h)^2 + h^2] dx + \frac{\beta}{2} \|h\|^2_{H^3(\Omega_\mu)}.$$

7.5.3 Proof of Theorem 7.5.4

Temporary change notation for the functional (7.52) as

$$J_{\lambda(\delta),\beta(\delta)}(W + P) = e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} [\Delta W + \Delta P - F(\nabla W + \nabla P)]^2 e^{2\lambda r} dx \tag{7.86}$$
$$+ \beta \|W + P\|^2_{H^3(\Omega_\mu)}.$$

Obviously

$$e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} [\Delta W^* + \Delta P^* - F(\nabla W^* + \nabla P^*)]^2 e^{2\lambda r} dx = 0.$$

Hence, by (7.61) and (7.86)

$$J_{\lambda(\delta),\beta(\delta)}(W^* + P^*) = \beta(\delta) \|W^* + P^*\|^2_{H^3(\Omega_\mu)} \leq C_2 \beta(\delta). \tag{7.87}$$

By (7.86),

$$J_{\lambda(\delta),\beta(\delta)}(W^* + P) - J_{\lambda(\delta),\beta(\delta)}(W^* + P^*)$$
$$= e^{-2\lambda(\mu+\eta)} \int_{\Omega_\mu} [\Delta W^* + \Delta P^* + \Delta\widetilde{P} - F(\nabla W^* + \nabla P^* + \nabla\widetilde{P})]^2 e^{2\lambda r} dx \tag{7.88}$$

$$- e^{-2\lambda(\mu+\eta)} \int\limits_{\Omega_\mu} [\Delta W^* + \Delta P^* - F(\nabla W^* + \nabla P^*)]^2 e^{2\lambda r} dx$$
$$+ \beta[\widetilde{P}, 2W^* + P + P^*].$$

Recall that $F(\nabla V)$ is a quadratic vector function with respect to the derivatives $\partial_{x_j} v_k(x)$. Hence, (7.61) and (7.88) imply that

$$|J_{\lambda(\delta),\beta(\delta)}(W^* + P) - J_{\lambda(\delta),\beta(\delta)}(W^* + P^*)| \leq C_2 \delta e^{2\lambda\rho} + C_2 \delta\beta. \tag{7.89}$$

Next,

$$|J_{\lambda(\delta),\beta(\delta)}(W^* + P) - J_{\lambda(\delta),\beta(\delta)}(W^* + P^*)| \geq J_{\lambda(\delta),\beta(\delta)}(W^* + P) - J_{\lambda(\delta),\beta(\delta)}(W^*, P^*).$$

Hence, using (7.87) and (7.89) and keeping in mind that $C_2\delta\beta < C_2\beta$, we obtain

$$J_{\lambda(\delta),\beta(\delta)}(W^* + P) \leq C_2 \delta e^{2\lambda\rho} + C_2 \beta. \tag{7.90}$$

Since $\lambda(\delta) = \ln \delta^{-1/(4\rho)}$ and $\delta < \delta_1 < \min(e^{-4\rho\lambda_2}, 3^{-4\rho/\eta})$, then $\lambda(\delta) > \lambda_2$ and also $\delta e^{2\lambda\rho} = \sqrt{\delta}$. Next, since $\beta = 3\delta^{\eta/(4\rho)}$, then condition (7.54) is fulfilled. Also, since $3\delta^{\eta/(4\rho)} > \sqrt{\delta}$, then $\delta e^{2\lambda\rho} + \beta \leq 2\delta^{\eta/(4\rho)}$.

Hence, using (7.90), we obtain

$$J_{\lambda,\beta}(W^* + P) \leq C_2 \delta^{\eta/(4\rho)}. \tag{7.91}$$

Since by (7.57) $[J'_{\lambda(\delta),\beta(\delta)}(W_{\min,\lambda(\delta),\beta(\delta)}), W^* - W_{\min,\lambda(\delta),\beta(\delta)}] \geq 0$, then using (7.91), we obtain

$$J_{\lambda(\delta),\beta(\delta)}(W^*) - J_{\lambda(\delta),\beta(\delta)}(W_{\min,\lambda(\delta),\beta(\delta)}) - J'_{\lambda(\delta),\beta(\delta)}(W_{\min,\lambda(\delta),\beta(\delta)})(W^* - W_{\min,\lambda}) \leq C_2 \delta^{\eta/(4\rho)}.$$

Hence, by (7.55)

$$\|\Delta W^* - \Delta W_{\min,\lambda(\delta),\beta(\delta)}\|^2_{L^2(\Omega_{\mu+\eta})} + \|W^* - W_{\min,\lambda(\delta),\beta(\delta)}\|^2_{H^1(\Omega_{\mu+\eta})} \leq C_2 \delta^{\eta/(4\rho)},$$

from which (7.62) and (7.63) follow.

We now prove (7.64). Using triangle inequality (7.58) and (7.62), we obtain for $n \geq 1$,

$$\begin{aligned} \|W^* - W_n\|_{H^1(\Omega_{\mu+\eta})} &\leq \|W^* - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^1(\Omega_{\mu+\eta})} + \|W_{\min,\lambda(\delta),\beta(\delta)} - W_n\|_{H^1(\Omega_{\mu+\eta})} \\ &\leq C_2 \delta^{\eta/(8\rho)} + \omega^n \|W_0 - W_{\min,\lambda(\delta),\beta(\delta)}\|_{H^3(\Omega)}, \end{aligned}$$

which proves (7.64). The proof of (7.65) is completely similar.

7.5.4 Proof of Theorem 7.5.5

Subtracting (7.67) from (7.66), we obtain

$$\begin{aligned} &\left|a_0^*(x) - a_{0,\min,\lambda(\delta),\beta(\delta)}(x)\right| \\ &\quad \le C_2 \sum_{k=0}^{N-1} \left|\Delta v_k^*(x) - \Delta v_{k,\min,\lambda(\delta),\beta(\delta)}(x)\right| \\ &\qquad + C_2 \sum_{k=0}^{N-1} \left|\nabla v_k^*(x) - \nabla v_{k,\min,\lambda(\delta),\beta(\delta)}(x)\right| \left|\nabla v_k^*(x) + \nabla v_{k,\min,\lambda(\delta),\beta(\delta)}(x)\right|. \end{aligned} \tag{7.92}$$

Since vector functions $W_{\min,\lambda(\delta),\beta(\delta)}, W^* \in \overline{B(R)}$, then (7.53) implies that $|\nabla v_k^*(x) + \nabla v_{k,\min,\lambda(\delta),\beta(\delta)}(x)| \le C_2$. Hence, (7.69) follows from (7.62), (7.63), and (7.92). The proof of (7.70) is completely similar.

7.6 Numerical studies

We have applied the above technique to numerical studies of the inverse EIT problem in the 2D case. Recall that even though Theorems 7.5.1–7.5.4 are formulated only in the 3D case, their direct analogs are also valid in the 2D case due to the Carleman estimate of Theorem 7.3.2; see the beginning of Section 7.4. In this section, we describe our numerical results. Hence, in this section

$$\Omega = \{r \in (0,\rho)\} \subset \mathbb{R}^2, \quad \Omega_\mu = \{r \in (\mu,\rho)\} \subset \Omega.$$

We have found in our computations that the influence of the regularization parameter β in (7.52) is not essential. Hence, we set $\beta := 0$ in our computational examples.

7.6.1 Some details of the numerical implementation

In all our numerical examples,

$$G = \{x_1^2 + x_2^2 < 5\}, \quad \Omega = \{x_1^2 + x_2^2 \le 1\} \quad \text{and} \quad \Omega_\mu = \{r \in (0.01,1)\} \subset \Omega.$$

We measure the data on the whole boundary $\partial\Omega = S_1$. The source runs over the circle $C^{(s)} = \{x_1^2 + x_2^2 = 4\}$. In other words, in polar coordinates

$$x_s = (r,s) = (2,s), \quad s = \varphi \in (0,2\pi), \quad x_s \in C^{(s)}. \tag{7.93}$$

However, when constructing the required orthonormal basis $\{\psi_n(s)\}_{n=0}^{\infty}$, we still have used functions $\{s^n e^s\}_{n=0}^{\infty}$, that is, we did not impose the periodicity condition on

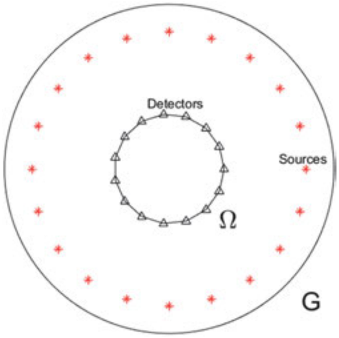


Figure 7.1: A schematic diagram of domains G, Ω, sources, and detectors.

this basis. The source function $f(x)$ in our case is the bump function below:

$$f(x - x^{(s)}) = \begin{cases} \frac{1}{\varepsilon}\exp(-\frac{1}{1-|x-x_s|^2/\varepsilon}), & \text{if } (x - x_s)^2 < \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

We have chosen $\varepsilon = 0.01$.

We use 32 sources and 32 detectors, see Figure 7.1. In Tests 1–5, both sources and detectors are uniformly distributed over the whole circle $\{x_1^2 + x_2^2 = 4\}$ and the whole circle $S_1 = \{x_1^2 + x_2^2 = 1\}$, respectively. However, this changes in Test 6 (see below).

To solve the forward problem (7.4), we have used the standard FEM. However, to minimize functional (7.52), we have written the differential operators in it via finite differences. Thus, we have not committed “inverse crime.” To use the finite differences, we have discretized the domain Ω_μ in polar coordinates using the uniform finite difference mesh. Next, we have used the gradient descent method to minimize functional (7.52) with respect to the values of the vector function $W(r, \varphi)$ at grid points. As the basis ψ_k is not periodic over $[0, 2\pi]$, we treat numerically $s = 0$ and $s = 2\pi$ as two different discrete points.

As to the choice of the parameter λ, even though the above theory works only for sufficiently large values of λ, we have established in our computational experiments that the choice

$$\lambda = 1 \tag{7.94}$$

is sufficient for all six numerical tests we have performed. We have also tested three different values of the number N terms in the series (7.15):

$$N = 4, 6, 8. \tag{7.95}$$

Our computational results indicate that $N = 8$ is the best choice out of these three.

Remark 7.6.1. The choice (7.94) of the parameter λ corresponds well with the observations of all publications on numerical studies of the convexification method [9, 115–117, 142–145, 145, 146, 150, 151, 164]. This observation is that not large values of λ can be chosen in computations; also, see Section 5.9.5.

7.6.2 A multilevel method of the minimization of functional (7.52)

We have found in our computational experiments that the gradient descent method for our weighted Tikhonov-like functional (7.52) converges rapidly on a coarse mesh. This provides us with a rough image. Hence, we have implemented a multilevel method [189]. Let $M_{h_1} \subset M_{h_2} \subset \cdots \subset M_{h_K}$ be nested finite difference meshes, that is, M_{h_k} is a refinement of $M_{h_{k-1}}$ for $k \le K$. Let P_{h_k} be the corresponding finite difference functional space. One the first level M_{h_1}, we solve the discrete optimization problem. In other words, let $V_{h_1,\min}$ be the minimizer of the following functional, which is found via the gradient descent method:

$$J_\lambda^{(h_1)}(W_{h_1}) = e^{-2\lambda(\mu+\eta)} \int\limits_{\Omega_\mu} [\Delta W_{h_1} + \Delta P - F(\nabla W_{h_1} + \nabla P)]^2 e^{2\lambda r} dx, \tag{7.96}$$

where the integral is understood in the discrete sense. Then we interpolate the minimizer $W_{h_1,\min,\lambda}$ on the finer mesh M_{h_2} and take the resulting vector function $W_{h_2,\text{int}}$ as the starting point of the gradient descent method of the optimization of the direct analog of functional (7.86) in which h_1 is replaced with h_2 and W_{h_1} is replaced with W_{h_2}. This process was repeated until we got the minimizer $W_{h_K,\min}$ on the K_{th} level on the mesh M_{h_K}.

Since $(r,\varphi) \in (0,1) \times (0,2\pi)$, then our first level M_{h_1} is set to be the uniform mesh with the mesh size in the r direction to be $1/4$ and the mesh size in the φ direction to be $2\pi/8$. For each mesh refinement, we will refine the mesh in both r direction and φ direction in a way that we set the mesh size of the refined mesh in both directions to be 1/2 of the previous mesh sizes. On each level M_{h_k}, as soon as we see that $\|\nabla J_\lambda^{(h_k)}(W_{h_k})\| < 2 \times 10^{-2}$, we refine the mesh and compute the solution on the next level $M_{h_{k+1}}$. In the end, we compute $a_0(x)$ using the relation (7.12) with $s = 0$.

Our starting point $W^{(0)}(r,\varphi)$ for the vector function $W(r,\varphi)$ for the gradient descent method on the coarse mesh M_{h_1} is set to be the background solution $W^{(0)}(r,\varphi,1)$ which corresponds to the solution of the problem (7.4) with $\sigma(x) \equiv 1$. Hence, our starting point is not located in a small neighborhood of the exact solution.

7.6.3 Numerical testing

In the tests of this section, we demonstrate the efficiency of our numerical method for imaging of small inclusions as well as for imaging of a smoothly varying function $\sigma(x)$, that is, a "stretched" inclusion with a wide range of change of the conductivity inside of it. In particular, we test the case of a rather high contrast 5:1 of the inclusion. In all tests, the background value of the conductivity is $\sigma_{\text{bkgr}} = 1$. In addition, we test the influence of the number N in (7.95). We also test the effects of both: the data given only on a part of the boundary and the source running only along a part of the circle

$\{r = 2\}$. In Tests 1–6, we have stopped on the 3rd mesh refinement for all three values of N listed in (7.95) (except for Test 4 where $N = 8$). The reason of stopping on the 3rd mesh refinement is that images were changing very insignificantly when on the 3rd mesh refinement, as compared with the second.

All necessary derivatives of the data were calculated using finite differences, just as in previous above cited publications about the convexification [9, 115–117, 142–145, 145, 146, 150, 151, 164], including even the ones with noisy experimental data [115–117, 143–145]; also see Chapter 10. Just as in those works, we have not observed instabilities due to the differentiation, most likely because the step sizes of finite differences were not too small.

Test 1. First, we test the reconstruction by our method of a single inclusion depicted on Figure 7.2(a). Here, $\sigma = 2$ inside of this inclusion and $\sigma = 1$ outside. Hence, the

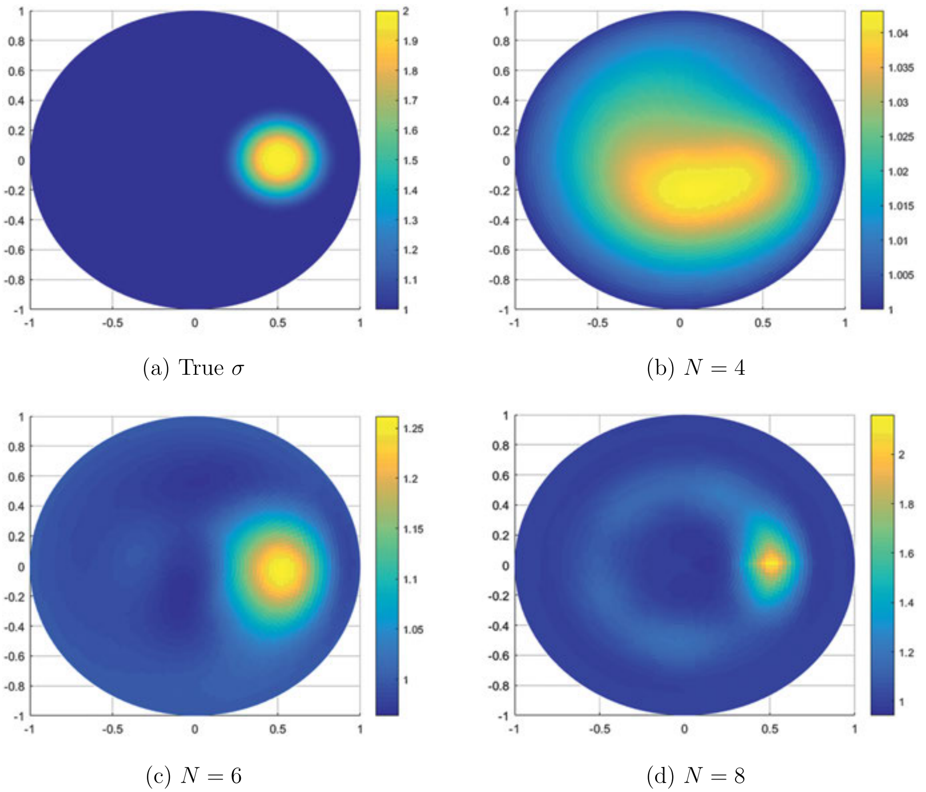
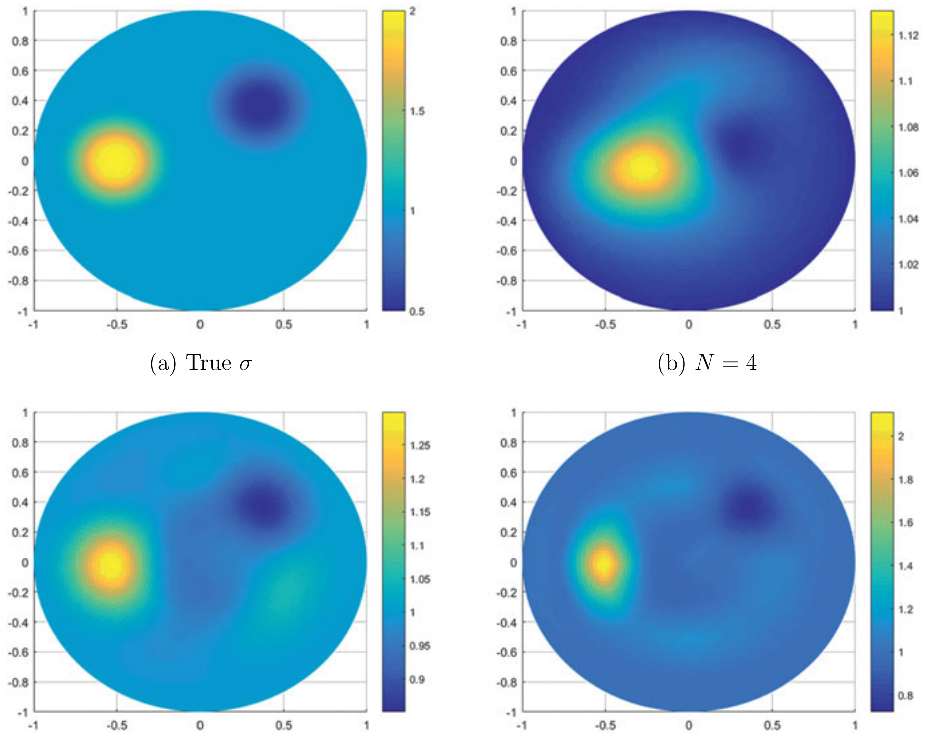


Figure 7.2: Results of Test 1. Imaging of one inclusion with $\sigma = 2$ in it and $\sigma = 1$ outside. Hence, the inclusion/background contrast is 2:1. We have stopped at the 3rd mesh refinement for all three values of N listed in (7.95). (a) Correct image. (b) Computed image for $N = 4$. (c) Computed image for $N = 6$. (d) Computed image for $N = 8$. Both the correct contrast and correct location are achieved at $N = 8$.



(a) True σ (b) $N = 4$

(c) $N = 6$ (d) $N = 8$

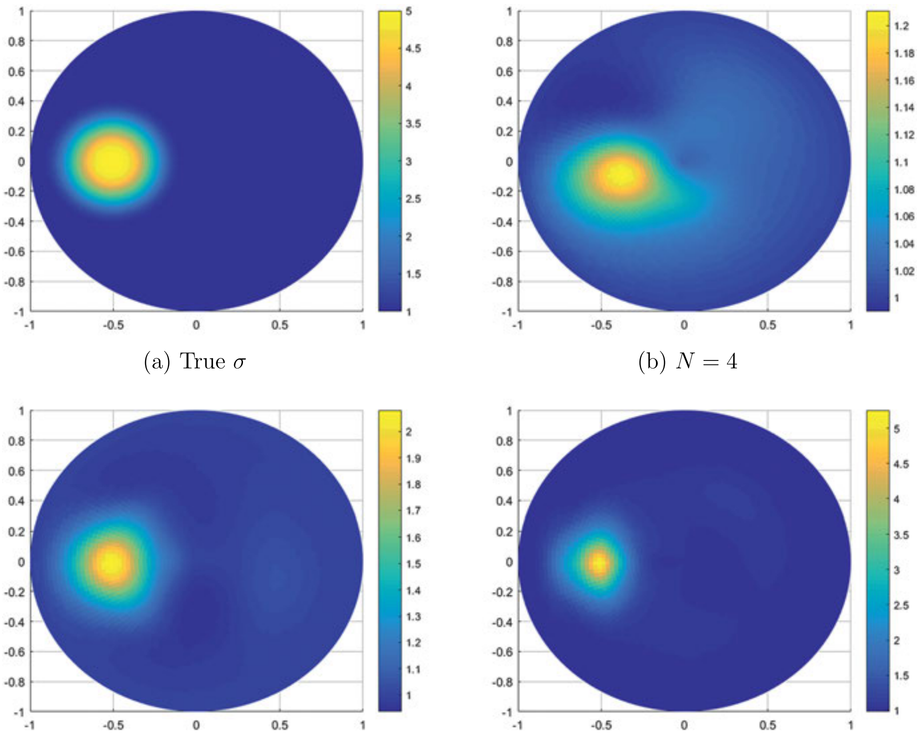
Figure 7.3: Results of Test 2. Imaging of two inclusions. Here, $\sigma = 2$ in the left inclusion, $\sigma = 0.5$ in the right inclusion and $\sigma = 1$ otherwise. Hence, the inclusion/background contrast is 2:1 in the left inclusion and is 0.5:1 in the right inclusion. This means that the electric conductivity of the left inclusion is higher than the one of the background and it is lower of the right inclusion. We have stopped on the 3rd mesh refinement for all three values of N listed in (7.95). (a) Correct image. (b) Computed image for $N = 4$. (c) Computed image for $N = 6$. (d) Computed image for $N = 8$, which is the best one out of three.

inclusion/background contrast is 2:1. The best result is achieved at $N = 8$; see Figures 7.2.

Test 2. We test now the performance of our method for imaging of two inclusions depicted on Figure 7.3(a): $\sigma = 2$ inside of each inclusion and $\sigma = 1$ outside of these inclusions. See Figures 7.3 for results.

Test 3. In this example, we test the reconstruction method for a single inclusion with a rather high inclusion/background contrast 5:1. The results are shown on Figure 7.4.

Test 4. We now test our method for the case when the function $\sigma(x)$ is smoothly varying within an abnormality and with a wide range of variations between 0.4 and 1.6. The results are shown in Figure 7.5. Again $N = 8$ is the best value out of three listed in (7.95).



(a) True σ (b) $N = 4$ (c) $N = 6$ (d) $N = 8$

Figure 7.4: Results of Test 3. Imaging of a single inclusion with a high inclusion/background contrast 5:1. Here, $\sigma = 5$ inside the inclusion and $\sigma = 1$ outside. (a) Correct image. (b) Computed image for $N = 4$. (c) Computed image for $N = 6$. (d) Computed image for $N = 8$, the best one out of three.

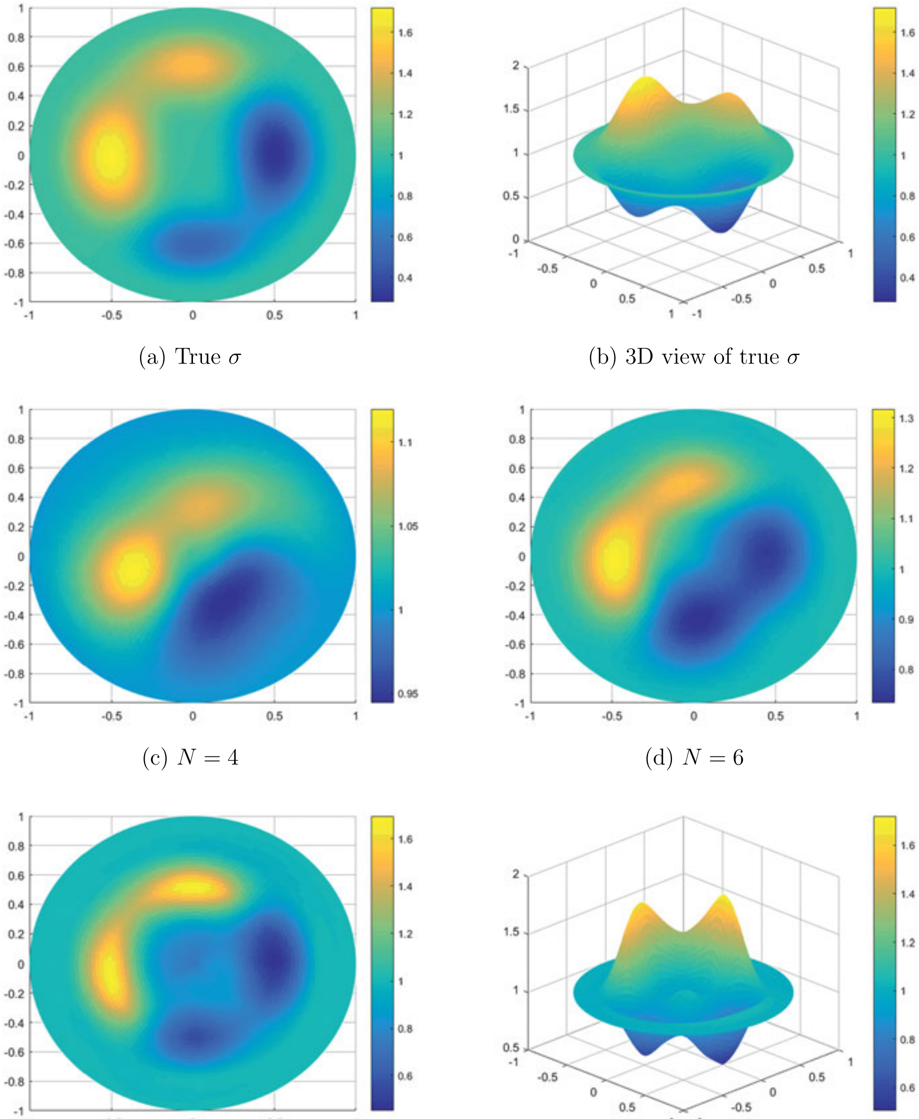
Thus, our method can accurately image not only “sharp” inclusions as in Tests 1–3, but smoothly varying functions as well.

Test 5. Here, we give a test for the case when the function $\sigma(x)$ is in the form of void shape. The results are shown in Figure 7.6. $N = 8$ is the best value out of three listed in (7.95).

Test 6. In this example, we test the stability of the algorithm with respect to the random noise in the data. We test the most challenging case among ones above: the case of the function $\sigma(x)$ of test 4. We set $N = 8$. The noise is added for $x \in S_1$ and for the source s as in (7.93), $s \in [0, 2\pi]$:

$$g_{0,\text{noise}}(x,s) = g_0(x,s)(1 + \epsilon\xi_s) \quad \text{and} \quad g_{1,\text{noise}}(x,s) = g_1(x,s)(1 + \epsilon\xi_s),$$

where ε is the noise level and ξ_s is the independent random variable depending only on the source position s and uniformly distributed on $[-1, 1]$. The computational results are displayed on Figure 7.7 for the levels of noise of 1 % and 10 %.

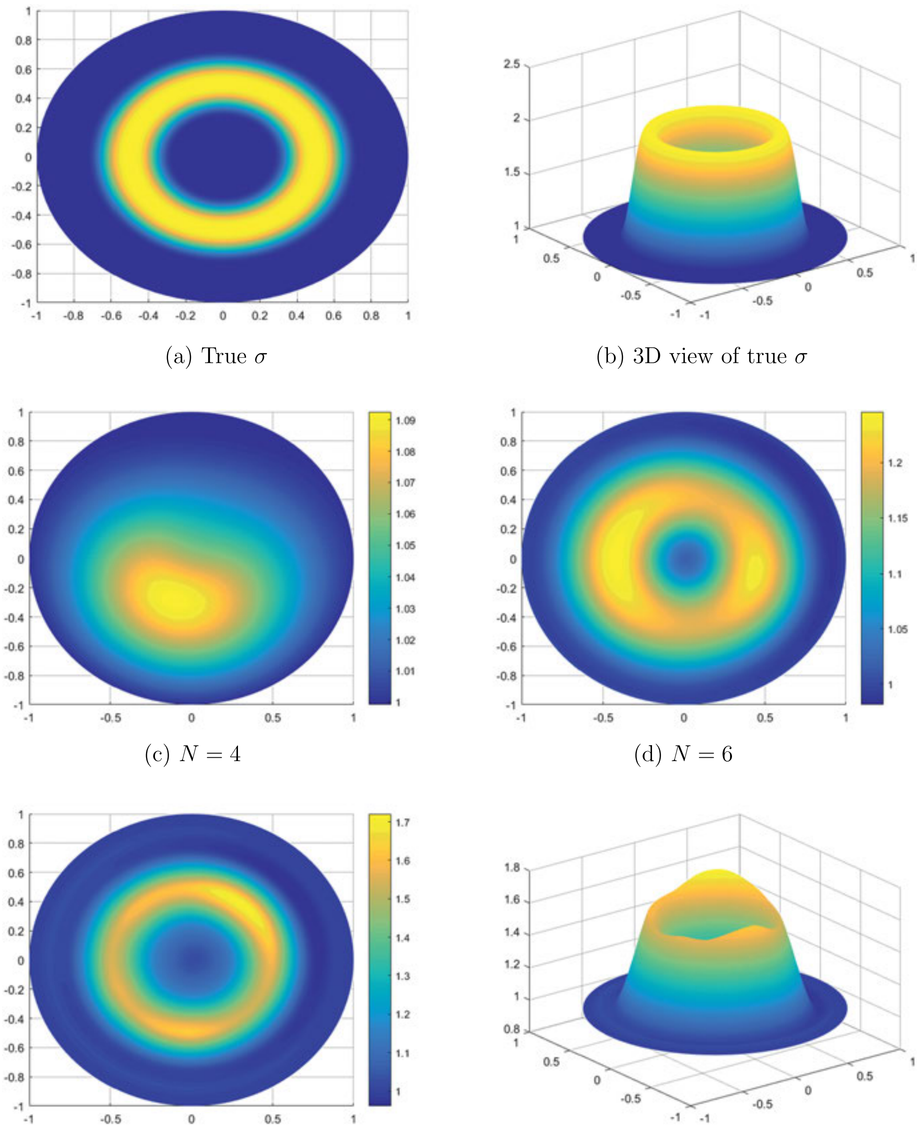


(a) True σ (b) 3D view of true σ

(c) $N = 4$ (d) $N = 6$

(e) $N = 8$ (f) 3D view of the result $N = 8$

Figure 7.5: Results of Test 4. We now test imaging of a smoothly varying conductivity rather than of inclusions above. The values of (x) inside of the inhomogeneity vary in a wide range $\sigma_{\min} \approx 0.3$ and $\sigma_{\max} \approx 1.7$. And $\sigma = 1$ in the homogeneous part of this disk. Here, we have stopped on the 3rd mesh refinement. (a) Correct 2D image. (b) 3D presentation of (a). (c) Computed image for $N = 4$. (d) Computed image for $N = 6$. (e) Computed 2D image for $N = 8$. (f) 3D presentation of (e). Thus, we can accurately image not only “sharp” inclusions but smoothly varying functions as well.

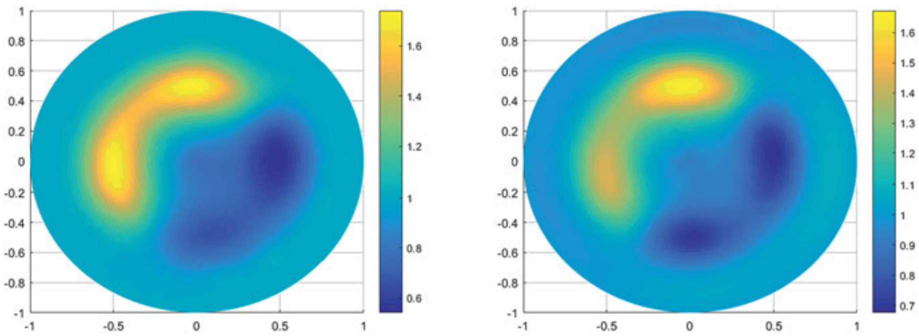


(a) True σ (b) 3D view of true σ

(c) $N = 4$ (d) $N = 6$

(e) $N = 8$ (f) 3D view of the result $N = 8$

Figure 7.6: Results of Test 5. We test imaging of conductivity with void shape. Here, we have stopped on the 3rd mesh refinement. (a) Correct 2D image. (b) 3D presentation of (a). (c) Computed image for $N = 4$. (d) Computed image for $N = 6$. (e) Computed 2D image for $N = 8$. (f) 3D presentation of (e).



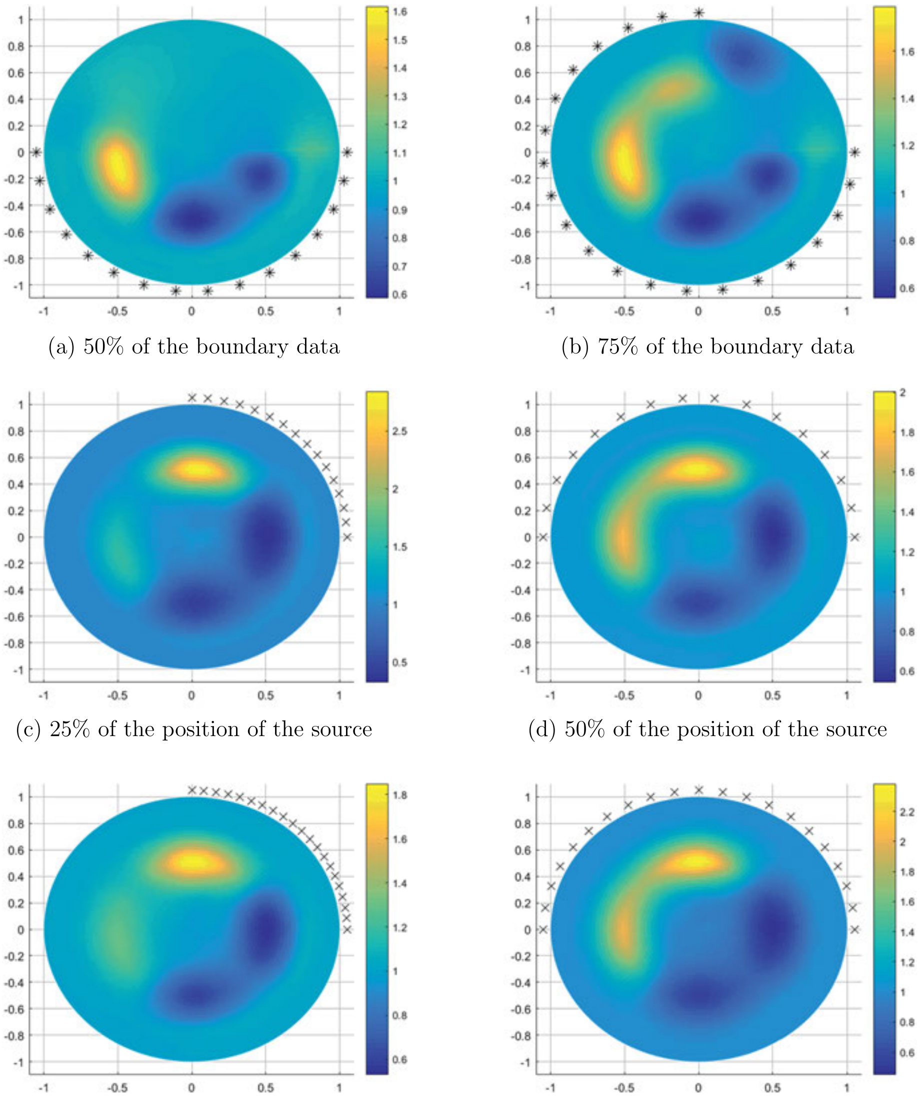
(a) Noise level 1%   (b) Noise level 10%.

Figure 7.7: Results of Test 6. In this test, we have introduced random noise in the data of test 4. Here, $N = 8$. (a) Computed image with 1% noise. (b) Computed image with 10% noise.

This example indicates that our method is quite stable with respect to the noise in the measured data.

Test 7. In all the above Tests 1–6, we have used the Dirichlet and Neumann data on the entire boundary S_1 of our disk Ω. Also, the source was running along the entire circle $C^{(s)}$ as in (7.93). In this test, however, we study the case of incomplete data. First, we work with the case when the source runs over the entire circle (7.93) while the data $g_0(x,s)$ and $g_0(x,s)$ are measured only on a part of the circle S_1. Next, we study the case when the source runs only along a part of the circle $C^{(s)}$ in (7.93) while the data are measured on the entire circle S_1. We again use $N = 8$ and the same function $\sigma(x)$ as in Test 4. We also test the robustness against the noise for the second case. The noise is added as we did in Test 6. The computational results are displayed on Figure 7.8 for the levels of noise of 1 %.

Figures 7.8 display results of Test 7. Comparing with the correct image of Figure 7.5, one can observe that, using 50 % of the measured boundary data, one loses about 50 % of the internal information. On the other hand, using 50 % of the positions of the source, one can still recover the internal conductivity with a rather good accuracy. Hence, it seems to be more important to measure at the entire boundary than to use the entire circle $C^{(s)}$ for the positions of the source.



(a) 50% of the boundary data
(b) 75% of the boundary data
(c) 25% of the position of the source
(d) 50% of the position of the source
(e) 25% of the source with 1% noise
(f) 50% of the source with 1% noise

Figure 7.8: Results of Test 7. While the structure to be imaged is the same as the one of Figure 7.4(a), a lesser amount of data is used here. In (a) and (b), we use incomplete boundary data on the circle $S_1 = \{r = 1\}$, while the source is still running as in (7.93): $s \in (0, 2\pi)$, that is, over the entire circle $C^{(s)} = \{r = 4\}$. On the other hand, in (c) and (d) the boundary data are measured at the entire circle $S_1 = \{r = 1\}$, while the source is running over only a part of the circle $C^{(s)} = \{r = 4\}$. In (a) and (b), $*$ indicates the part of the circle $S_1 = \{r = 1\}$ where the data are measured. In (c) and (d), $\times$ indicates the part of the circle $\{r = 4\}$ where the source runs. The part with $\times$ is depicted on $\{r = 1\}$ rather than on $\{r = 4\}$ only for the convenience of the presentation. (e) and (f) show the recovery results with 1% noise level for incomplete source case, the noise is added as we did in Test 6.

8 Convexification for a coefficient inverse problem for a hyperbolic equation with a single location of the point source

In this chapter, we follow our publication [150]. Permission for republishing is obtained from the publisher.

8.1 Introduction

In this chapter, we develop analytically and test numerically a version of the convexification numerical method for a Coefficient Inverse Problem (CIP) for the hyperbolic equation:

$$c(\mathbf{x})u_{tt} = \Delta u, \quad \mathbf{x} \in \mathbb{R}^3, t > 0, \tag{8.1}$$

$$u(\mathbf{x}, 0) = 0, \quad u_t(\mathbf{x}, 0) = \delta(\mathbf{x} - \mathbf{x}_0) \tag{8.2}$$

with a single position of the point source $\mathbf{x}_0$. We consider the case of transmitted data. This CIP has applications in, for example, problem of the noninvasive inspections of buildings using measurements of the propagated electric field [162, 206]. In this case, $c(\mathbf{x})$ is the spatially distributed dielectric constant. In addition, this CIP has applications in acoustical testing of the medium, in which case $1/\sqrt{c(\mathbf{x})}$ is the speed of sound. Just as in Chapters 6 and 7, we use a truncated Fourier series here with N terms. But we do not prove convergence of our method as $N \to \infty$. In other words, we work within a framework of an approximate mathematical model. We refer to Remarks 7.3 and 10.3.1 as well as to [117, 150] for further discussions of this issue.

In the case of a nonvanishing initial condition, a different version of the convexification was recently proposed in [16] for the PDE $u_{tt} = \Delta u + a(x)u$ with the unknown coefficient $a(x)$. Next, this method was extended in [17] to the case of equation (8.1). The case of a nonvanishing initial condition is less challenging one than our case of the δ-function in the initial condition. Publications [16, 17] work exactly within the framework of the Bukhgeim–Klibanov method. We, however, go beyond this method, since our initial conditions (8.2) vanishing. The idea of [16] was explored further in [185] to develop globally convergent numerical methods for some inverse problems for quasi-linear parabolic PDEs.

While we consider here only the case when the unknown coefficient of a hyperbolic PDE depends only on spatial variables, various cases of its dependence on both spatial variables in time were considered in works of [118, 119]; also see some follow up works of these authors.

https://doi.org/10.1515/9783110745481-008

8.2 Statement of the inverse problem

Below $\mathbf{x} = (x, y, z) = (x_1, x_2, x_3) \in \mathbb{R}^3$. Let $A > 0$ be a number. Since the domain of interest Ω is a cube in our computations, then it is convenient to set $\Omega \subset \mathbb{R}^3$ as

$$\Omega = \{\mathbf{x} = (x, y, z) : -A/2 < x, y < A/2, z \in (0, A)\}. \tag{8.3}$$

Let the number $a > 0$. We set the single point source we use as $\mathbf{x}_0 = (0, 0, -a)$. Hence, this source is located below the domain Ω. Let Γ_0 be the upper boundary of Ω and Γ_1 be the rest of this boundary,

$$\Gamma_0 = \{\mathbf{x} = (x, y, z) : -A/2 < x, y < A/2, z = A\}, \quad \Gamma_1 = \partial\Omega \setminus \Gamma_0. \tag{8.4}$$

Thus, Γ_0 is the “transmitted” side of Ω. Let the function $c(\mathbf{x})$ be such that

$$c \in C^{13}(\mathbb{R}^3), \tag{8.5}$$

$$c(\mathbf{x}) \geq c_0 = \text{const.} > 0 \quad \text{in } \overline{\Omega}, \tag{8.6}$$

$$c(\mathbf{x}) = 1, \quad \forall \mathbf{x} \in \mathbb{R}^3 \setminus \Omega. \tag{8.7}$$

Remark 8.2.1. We assume in (8.3) the C^{13}-smoothness of the function $c(\mathbf{x})$ since the representation (8.13) of the solution of the Cauchy problem (8.11), (8.12) works only under this assumption. Indeed, this smoothness was carefully calculated in Theorem 4.1 of the book [224].

The physical meaning of the function $c(\mathbf{x})$ is that $1/\sqrt{c(\mathbf{x})}$ is the speed of sound. Consider the conformal Riemannian metric generated by the function $c(\mathbf{x})$,

$$d\tau = \sqrt{c(\mathbf{x})}\sqrt{(dx)^2 + (dy)^2 + (dz)^2}. \tag{8.8}$$

The metric (8.8) generates geodesic lines $\Gamma(\mathbf{x}, \mathbf{x}_0)$, $\mathbf{x} \in \mathbb{R}^3$. Let $\tau(\mathbf{x})$ is the travel time along the geodesic line $\Gamma(\mathbf{x}, \mathbf{x}_0)$. Then [224] the function $\tau(\mathbf{x})$ is the solution of the eikonal equation

$$|\nabla\tau(\mathbf{x})|^2 = c(\mathbf{x}) \tag{8.9}$$

with the condition $\tau(\mathbf{x}) = O(|\mathbf{x} - \mathbf{x}_0|)$ as $|\mathbf{x} - \mathbf{x}_0| \to 0$. Furthermore,

$$\tau(\mathbf{x}) = \int\limits_{\Gamma(\mathbf{x}, \mathbf{x}_0)} \sqrt{c(\mathbf{y})}\, d\sigma.$$

Everywhere below we rely on the following assumption without further comments [224].

Regularity assumption. For the above specific point source $\mathbf{x}_0$, geodesic lines generated by the function $c(\mathbf{x})$ are regular. In other words, for any points $\mathbf{x} \in \mathbb{R}^3$ there exists unique geodesic line $\Gamma(\mathbf{x}, \mathbf{x}_0)$ connecting points $\mathbf{x}$ and $\mathbf{x}_0$.

A sufficient condition of the regularity of geodesic lines was derived in [225]. This condition is

$$\sum_{i,j=1}^{3} \frac{\partial^2 \ln c(\mathbf{x})}{\partial x_i \partial x_j} \xi_i \xi_j \geq 0, \quad \forall \mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^3. \tag{8.10}$$

As the forward problem, we consider the following Cauchy problem for the acoustic equation [66] of the hyperbolic type for the function $u(\mathbf{x}, \mathbf{x}_0, t)$:

$$c(\mathbf{x}) u_{tt} = \Delta u, \quad \mathbf{x} \in \mathbb{R}^3, t > 0, \tag{8.11}$$

$$u(\mathbf{x}, 0) = 0, \quad u_t(\mathbf{x}, 0) = \delta(\mathbf{x} - \mathbf{x}_0). \tag{8.12}$$

It was proven in Theorem 4.1 of [224] that, given the above conditions (8.5)–(8.7) as well as the regularity assumption, the solution of problem (8.11), (8.12) is a sum of the singular part and the regular part. The singular part is $A(\mathbf{x})\delta(t - \tau(\mathbf{x}))$, The regular part equals zero for $t < \tau(\mathbf{x})$. And, with a certain function $\widehat{u}(\mathbf{x}, t) \in C^2(t \geq \tau(\mathbf{x}))$, the regular part is $H(t - \tau(\mathbf{x}))\widehat{u}(\mathbf{x}, t)$. To summarize, the solution $u(\mathbf{x}, t)$ of problem (8.11), (8.12) has the form:

$$u(\mathbf{x}, t) = A(\mathbf{x})\delta\big(t - \tau(\mathbf{x})\big) + H\big(t - \tau(\mathbf{x})\big)\widehat{u}(\mathbf{x}, t). \tag{8.13}$$

In (8.13), functions $\tau(\mathbf{x}), A(\mathbf{x}) \in C^{12}(\mathbb{R}^3)$, the function $A(\mathbf{x}) > 0$, $\forall \mathbf{x} \in \mathbb{R}^3$, and $H(z)$ is the Heaviside function,

$$H(z) = \begin{cases} 1, & z > 0, \\ 0, & z < 0. \end{cases}$$

Let the number $T > 0$. Denote $S_T = \partial\Omega \times (0, T)$ and $\Gamma_{0,T} = \Gamma_0 \times (0, T)$. Since we work with only a single position $\mathbf{x}_0$ of the point source, we will omit below indications of dependencies on $\mathbf{x}_0$.

Coefficient Inverse Problem (CIP). Let the domain Ω be as in (8.3). Suppose that the following two functions are given:

$$u(\mathbf{x}, t)|_{(\mathbf{x},t) \in S_T} = f_0(\mathbf{x}, t), \quad \partial_z u(\mathbf{x}, t)|_{(\mathbf{x},t) \in \Gamma_{0,T}} = f_1(\mathbf{x}, t). \tag{8.14}$$

Determine the function $c(\mathbf{x})$ for $\mathbf{x} \in \Omega$.

We now explain how to obtain the normal derivative

$$\partial_z u(\mathbf{x}, t)|_{(\mathbf{x},t) \in \Gamma_{0,T}} = \partial_z u(x, y, A, t) = f_1(\mathbf{x}, t), \quad t \in (0, T), \tag{8.15}$$

in (8.14). Suppose that measurements $\varphi(x, y, t)$ of the amplitude $u(\mathbf{x}, t)$ of acoustic waves are conducted on the surface $\Gamma_1 = \partial\Omega \setminus \{z = A\}$ as well as on the full plane $\{z = A\}$. Using (8.7), (8.11), and (8.12), we obtain in the half-space $\{z > A\}$:

$$u_{tt} = \Delta u \quad \text{in } \big\{(\mathbf{x}, t) = (x, y, z, t) : z > A, t \in (0, T)\big\}, \tag{8.16}$$

$$u(\mathbf{x},0) = u_t(\mathbf{x},0) = 0, \tag{8.17}$$

$$u(x,y,A,t) = \varphi(x,y,t), \quad (x,y,t) \in \mathbb{R}^2 \times (0,T). \tag{8.18}$$

Solving initial boundary value problem (8.16)–(8.18) in the domain indicated in (8.16), one can uniquely obtain the function $u(x,y,z,t)$ in that domain. Hence, the normal derivative $\partial_z u(x,y,A,t)$ in (8.15) can be calculated using the knowledge of the function $u(x,y,z,t)$ for $z > A$.

8.3 A system of coupled quasi-linear elliptic equations

In this section, we reduce the CIP (8.11)–(8.14) to the Cauchy problem for a system of coupled elliptic PDEs.

8.3.1 The function $w(\mathbf{x},t)$

Integrate equation (8.11) twice with respect to t for points $\mathbf{x} \in \Omega$. Hence, we consider the function $p(\mathbf{x},t)$,

$$p(\mathbf{x},t) = \int_0^t dy \int_0^y u(\mathbf{x},s)ds, \quad \mathbf{x} \in \Omega. \tag{8.19}$$

Hence, $p_{tt}(\mathbf{x},t) = u(\mathbf{x},t)$ for $\mathbf{x} \in \Omega$. Since $\mathbf{x}_0 \notin \overline{\Omega}$, then $\delta(\mathbf{x}-\mathbf{x}_0) = 0$ for $\mathbf{x} \in \Omega$. Hence, by (8.12)

$$\int_0^t dy \int_0^y u_{tt}(\mathbf{x},s)ds = \int_0^t (u_t(\mathbf{x},y) - u_t(\mathbf{x},0))dy = u(\mathbf{x},t) = p_{tt}(\mathbf{x},t), \quad \mathbf{x} \in \Omega. \tag{8.20}$$

Next, by (8.11) and (8.19),

$$c(\mathbf{x}) \int_0^t dy \int_0^y u_{tt}(\mathbf{x},s)ds = \int_0^t dy \int_0^y \Delta u(\mathbf{x},s)ds = \Delta p(\mathbf{x},t), \quad \mathbf{x} \in \Omega. \tag{8.21}$$

Comparing (8.20) and (8.21), we obtain

$$c(\mathbf{x})p_{tt}(\mathbf{x},t) = \Delta p(\mathbf{x},t) \quad \text{for } (\mathbf{x},t) \in \Omega \times (0,T). \tag{8.22}$$

Next, by (8.13) and (8.19),

$$p(\mathbf{x},t) = A(\mathbf{x})(t-\tau(\mathbf{x}))H(t-\tau(\mathbf{x})) + O((t-\tau(\mathbf{x}))^2)H(t-\tau^0(\mathbf{x})), \tag{8.23}$$

where

$$\left|O\left((t-\tau(\mathbf{x}))^2\right)\right| \le B\left(t-\tau(\mathbf{x})\right)^2 \quad \text{as } t \to \tau^+(\mathbf{x}), \tag{8.24}$$

$$\left|\partial_t O\left((t-\tau(\mathbf{x}))^2\right)\right| \le B\left(t-\tau(\mathbf{x})\right) \quad \text{as } t \to \tau^+(\mathbf{x}) \tag{8.25}$$

with a certain constant $B > 0$ independent on $(\mathbf{x}, t) \in \Omega \times (0, T)$. In (8.24), (8.25) "$t \to \tau^+(\mathbf{x})$" means that t approaches $\tau(\mathbf{x})$ from the right.

Consider the function $w(\mathbf{x}, t)$ defined as

$$w(\mathbf{x}, t) = p\left(\mathbf{x}, t + \tau(\mathbf{x})\right), \quad \text{for } (\mathbf{x}, t) \in \Omega \times (0, T). \tag{8.26}$$

Then it follows from the above that $w \in C^2(\overline{\Omega} \times [0, T])$ and by (8.23)–(8.25),

$$w(\mathbf{x}, 0) = 0, \tag{8.27}$$

$$w_t(\mathbf{x}, 0) = A(\mathbf{x}) > 0. \tag{8.28}$$

We now want to derive a PDE for the function w. By (8.26),

$$w_t = p_t, w_{tt} = p_{tt}, \quad w_{x_i} = p_{x_i} + p_t \tau_{x_i},$$
$$w_{x_i t} = p_{x_i t} + p_{tt} \tau_{x_i}.$$

Hence, $2p_{x_i t}\tau_{x_i} = 2w_{x_i t}\tau_{x_i} - 2p_{tt}\tau_{x_i}^2$. Hence,

$$\begin{aligned} w_{x_i x_i} &= p_{x_i x_i} + 2p_{x_i t}\tau_{x_i} + p_{tt}\tau_{x_i}^2 + p_t \tau_{x_i x_i} \\ &= p_{x_i x_i} + 2w_{x_i t}\tau_{x_i} - p_{tt}\tau_{x_i}^2 + w_t \tau_{x_i x_i}. \end{aligned}$$

Hence,

$$\Delta w - 2\sum_{i=1}^{3} w_{x_i t}\tau_{x_i} - w_t \Delta\tau = \Delta p - p_{tt}(\nabla\tau)^2. \tag{8.29}$$

Next, by (8.9) $|\nabla\tau(\mathbf{x})|^2 = c(\mathbf{x})$. Hence, using (8.22), we obtain $\Delta p - p_{tt}(\nabla\tau)^2 = 0$. Hence, (8.29) implies that the following equation is valid for the function $w(\mathbf{x}, t)$ defined in (8.26):

$$\Delta w - 2\sum_{i=1}^{3} w_{x_i t}\tau_{x_i} - w_t \Delta\tau = 0, \quad \mathbf{x} \in \Omega, t \in (0, T_1), \tag{8.30}$$

$$T_1 = T - \max_{\overline{\Omega}} \tau(\mathbf{x}). \tag{8.31}$$

To explain (8.31), we note that since in the data (8.14) $t \in (0, T)$, then (8.26) implies that we should consider the function $w(\mathbf{x}, t)$ only for such values of $t > 0$ that $t + \tau(\mathbf{x}) < T$. Hence, to have a uniform with respect to $x \in \overline{\Omega}$ upper bound for t, we should have $t + \max_{\overline{\Omega}} \tau(\mathbf{x}) < T$, which explains (8.31).

Denote $\widetilde{f}_0(\mathbf{x},t) = f_0(\mathbf{x},t+\tau(\mathbf{x}))$ and $\widetilde{f}_1(\mathbf{x},t) = f_1(\mathbf{x},t+\tau(\mathbf{x}))$. Then by (8.14)

$$w(\mathbf{x},t)|_{(\mathbf{x},t)\in S_{T_1}} = \widetilde{f}_0(\mathbf{x},t), \quad \partial_z w(\mathbf{x},t)|_{(\mathbf{x},t)\in\Gamma_{0,T_1}} = \widetilde{f}_1(\mathbf{x},t). \tag{8.32}$$

Thus, our goal below is to construct a numerical method, which would approximately find the functions $w(\mathbf{x},t)$, $\tau(\mathbf{x})$ for $\mathbf{x} \in \Omega$, $t \in (0,T_1)$ from conditions (8.27)–(8.32). Suppose that these two functions are approximated. Then the corresponding approximation for the target coefficient $c(\mathbf{x})$ can be easily found via the backwards calculation,

$$c(\mathbf{x}) = |\nabla\tau(\mathbf{x})|^2. \tag{8.33}$$

8.3.2 The system of coupled quasi-linear elliptic PDEs

Lemma 8.3.1. *Consider the set of functions*

$$\{t, t^2, \dots, t^n, \dots\} = \{t^n\}_{n=1}^{\infty}. \tag{8.34}$$

The set (8.34) *is complete in* $L_2(0,T_1)$.

Proof. Let a function $f(t) \in L_2(0,T_1)$ be such that

$$\int_0^{T_1} f(t)t^n dt = 0, \quad n = 1,2,\dots.$$

Consider the function $\widetilde{f}(t) = f(t)t$. Then

$$\int_0^{T_1} \widetilde{f}(t)t^m dt = 0, \quad m = 0,1,2,\dots. \tag{8.35}$$

It is well known that (8.35) implies that $\widetilde{f}(t) \equiv 0$. □

Orthonormalize the set (8.34) using the Gram–Schmidt orthonormalization procedure. Then Lemma 8.3.1 implies that we obtain a basis $\{P_n(t)\}_{n=1}^{\infty}$ of orthonormal polynomials in $L_2(0,T_1)$ such that

$$P_n(0) = 0, \quad \forall n = 1,2,\dots. \tag{8.36}$$

By (8.36), this is not a set of standard orthonormal polynomials.

Let the integer $N \geq 1$. Approximate the function $w(\mathbf{x},t)$ as

$$w(\mathbf{x},t) = \sum_{n=1}^{N} w_n(\mathbf{x})P_n(t). \tag{8.37}$$

Here and below, we use "=" instead of "≈" for convenience. Substitute (8.37) in the left- hand side of (8.30) and assume that the resulting left-hand side equals zero.

We obtain for $\mathbf{x} \in \Omega$:

$$\sum_{n=1}^{N} \Delta w_n(\mathbf{x}) P_n(t) - 2 \sum_{i=1}^{3} \tau_{x_i}(\mathbf{x}) \sum_{n=1}^{N} P_n'(t) \partial_{x_i} w_n(\mathbf{x}) \qquad (8.38)$$
$$- \Delta\tau(\mathbf{x}) \sum_{n=1}^{N} P_n'(t) w_n(x) = 0.$$

By (8.28) and (8.37), we can assume that

$$\sum_{n=1}^{N} P_n'(0) w_n(\mathbf{x}) = A(\mathbf{x}) > 0, \quad \forall \mathbf{x} \in \overline{\Omega}. \qquad (8.39)$$

Set in (8.38) $t = 0$. Hence, we obtain the first elliptic equation,

$$\Delta\tau(\mathbf{x}) + 2 \left[\sum_{i=1}^{3} \tau_{x_i} \sum_{n=1}^{N} P_n'(0) \partial_{x_i} w_n(\mathbf{x}) \right] \left[\sum_{n=1}^{N} P_n'(0) w_n(\mathbf{x}) \right]^{-1} = 0, \quad \mathbf{x} \in \Omega.$$

We rewrite this equation as

$$\Delta\tau = F_1(\nabla\tau, \nabla\widetilde{W}, \widetilde{W}), \quad \mathbf{x} \in \Omega, \qquad (8.40)$$

where $\widetilde{W}(x) = (w_1(\mathbf{x}), \dots, w_N(\mathbf{x}))^T$. Next, for $n = 1, \dots, N$ multiply both sides of (8.38) by $P_n(t)$ and integrate with respect to $t \in (0, T_1)$. Replace in the resulting equation $\Delta\tau$ with the right-hand side of (8.40). We obtain

$$\Delta\widetilde{W} = F_2(\nabla\tau, \nabla\widetilde{W}, \widetilde{W}), \quad \mathbf{x} \in \Omega. \qquad (8.41)$$

Consider the $(N+1)$-dimensional vector function

$$W(\mathbf{x}) = \left(\tau(\mathbf{x}), \widetilde{W}(\mathbf{x})\right)^T. \qquad (8.42)$$

Thus, (8.32), (8.40), (8.41), and (8.42) lead to the following Cauchy problem for a system of coupled quasi-linear elliptic equations:

$$\Delta W + F(\nabla W, W) = 0, \quad \mathbf{x} \in \Omega, \qquad (8.43)$$
$$W|_{\partial\Omega} = q^0(\mathbf{x}), \quad \partial_z W|_{\Gamma_0} = q^1(\mathbf{x}), \qquad (8.44)$$
$$q^0(\mathbf{x}) = \left(\tau(\mathbf{x}), q_1^0(\mathbf{x}), \dots, q_N^0(\mathbf{x})\right)^T,$$
$$q_n^0(\mathbf{x}) = \int_0^{T_1} \widetilde{f}_0(\mathbf{x}, t) P_n(t) dt, \quad \mathbf{x} \in \partial\Omega, \qquad (8.45)$$

$$q^1(\mathbf{x}) = \left(\partial_z \tau(\mathbf{x}), q_1^1(\mathbf{x}), \dots, q_N^1(\mathbf{x})\right)^T,$$
$$q_n^1(\mathbf{x}) = \int_0^{T_1} \widetilde{f}_1(\mathbf{x}, t) P_n(t) dt, \quad \mathbf{x} \in \Gamma_0. \tag{8.46}$$

In (8.45) and (8.46), $n = 1, \dots, N$. In (8.43) the $(N+1)$-dimensional vector function $F \in C^1(\mathbb{R}^{3N+5})$. Thus, we have obtained the system (8.43) of coupled quasi-linear elliptic PDEs with the Cauchy data (8.44)–(8.46). Unknowns in this problem are the function $\tau(\mathbf{x})$ and Fourier coefficients $w_n(\mathbf{x})$ of the function $w(\mathbf{x}, t)$ in (8.37). Therefore, we solve below the problem (8.43)–(8.46) of finding the $(N+1)$-dimensional vector function $W \in C^2(\overline{\Omega})$. In fact, however, we find below $W \in H^3(\Omega)$.

8.4 Globally strictly convex Tikhonov-like functional

8.4.1 The functional

All Banach spaces considered below are spaces of real valued functions. If we say below that a vector function belongs to a certain Banach space, then this means that all its components belong to this space, and the norm of this function in that space is defined as the square root of the sum of squares of norms of its components.

To arrange a certain orthogonal projection operator for the gradient projection method below, the best way is to have zero Cauchy data. Hence, we assume that there exists an $(N+1)$-dimensional vector function $G = (g_0(\mathbf{x}), \dots, g_N(\mathbf{x}))^T \in H^3(\Omega)$ satisfying boundary conditions (8.44), that is, such that

$$G|_{\partial\Omega} = q^0(\mathbf{x}), \quad \partial_z G|_{\Gamma_0} = q^1(\mathbf{x}). \tag{8.47}$$

Let

$$W - G = Q \in H^3(\Omega), \quad Q(\mathbf{x}) = \left(q_0(\mathbf{x}), \dots, q_N(\mathbf{x})\right)^T. \tag{8.48}$$

Then (8.43), (8.44), and (8.47) imply that

$$\Delta Q + \Delta G + F\left(Q + G, \nabla(Q + G)\right) = 0, \quad \mathbf{x} \in \Omega, \tag{8.49}$$
$$Q|_{\partial\Omega} = 0, \quad \partial_z Q|_{\Gamma_0} = 0. \tag{8.50}$$

Let $H_0^3(\Omega)$ be the subspace of the space $H^3(\Omega)$ defined as

$$H_0^3(\Omega) = \left\{v \in H^3(\Omega) : v|_{\partial\Omega} = 0, \partial_z v|_{\Gamma_0} = 0\right\}.$$

Choose an arbitrary number $R > 0$ and also choose another number $m \in (0, R)$, which is independent on R. Consider the set $B(m, R)$ of $(N+1)$-dimensional vector functions

$Z(\mathbf{x}) = (z_0(\mathbf{x}), \dots, z_N(\mathbf{x}))^T$ such that

$$B(m,R) = \begin{cases} Z \in H_0^3(\Omega), \quad \|Z\|_{H^3(\Omega)} < R, \\ \sum_{n=1}^N P_n'(0)(z_n(\mathbf{x}) + g_n(\mathbf{x})) > m, \quad \forall \mathbf{x} \in \overline{\Omega}. \end{cases} \tag{8.51}$$

The second condition (8.51) is generated by (8.39). By embedding theorem $H^3(\Omega) \subset C^1(\overline{\Omega})$. This implies that $\overline{B(m,R)} \subset C^1(\overline{\Omega})$ and also that there exist numbers $D_1(R) > 0$ and $D_2(G) > 0$ depending only on listed parameters such that

$$\|Z\|_{C^1(\overline{\Omega})} \le D_1(R), \quad \forall Z \in \overline{B(m,R)}, \tag{8.52}$$

$$\|G\|_{C^1(\overline{\Omega})} \le D_2(G). \tag{8.53}$$

Temporarily replace the vector functions $Q(\mathbf{x}) = (q_0(\mathbf{x}), \dots, q_N(\mathbf{x}))^T$ and $\nabla Q(\mathbf{x}) = (\nabla q_0(\mathbf{x}), \dots, \nabla q_N(\mathbf{x}))^T$ with the vector of real variables $(y_0, y_1, \dots, y_{4N+3})^T = y \in \mathbb{R}^{4N+4}$. Consider the set $Y \subset \mathbb{R}^{4N+4}$,

$$Y = \left\{ y \in \mathbb{R}^{4N+4} : \sum_{n=1}^N P_n'(0)\big(y_n + g_{n-1}(\mathbf{x})\big) > m, \forall \mathbf{x} \in \overline{\Omega} \right\}.$$

Obviously, Y is an open set in $\mathbb{R}^{4N+4}$. Denote $p_1 = (y_0, \dots, y_N)$, $p_2 = (y_{0,1}, y_{0,2}, y_{0,3}, y_{1,1}, \dots, y_{N,3})$. Then $y = (p_1, p_2)^T \in \mathbb{R}^{4N+4}$. It follows from (8.40)–(8.43) that, as the function of y,

$$F\big(p_1 + G(\mathbf{x}), p_2 + \nabla G(\mathbf{x})\big) \in C^2(\overline{Y}), \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{8.54}$$

Lemma 8.4.1. *The set $B(m,R)$ is convex.*

Proof. Let the number $\alpha \in (0,1)$ and vector functions $Z, Y \in B(m,R)$. Consider the function $\alpha Z + (1-\alpha)Y$. Then

$$\big\|\alpha Z + (1-\alpha)Y\big\|_{H^3(\Omega)} \le \alpha\|Z\|_{H^3(\Omega)} + (1-\alpha)\|Y\|_{H^3(\Omega)} < \alpha R + (1-\alpha)R = R.$$

Next, let $Z(\mathbf{x}) = (\tau_1(\mathbf{x}), z_1(\mathbf{x}), \dots, z_N(\mathbf{x}))^T$, $Y(\mathbf{x}) = (\tau_2(\mathbf{x}), y_1(\mathbf{x}), \dots, y_N(\mathbf{x}))^T$. Then

$$\alpha \sum_{n=1}^N P_n'(0) z_n(\mathbf{x}) + (1-\alpha) \sum_{n=1}^N P_n'(0) y_n(\mathbf{x}) > \alpha m + (1-\alpha)m = m. \qquad \square$$

Our weighted Tikhonov-like cost functional is

$$J_{\lambda,\beta}(Q+G) = e^{-2\lambda b^2} \int_\Omega \big(\Delta Q + \Delta G + F\big(\nabla(Q+G), Q+G\big)\big)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \tag{8.55}$$
$$+ \beta\|Q+G\|^2_{H^3(\Omega)}.$$

In (8.55), the numbers $\lambda \ge 1$, $b > 0$, $\beta \in (0,1)$. Here, λ is the parameter of our CWF $e^{2\lambda(z+b)^2}$ and β is the regularization parameter. The multiplier $e^{-2\lambda b^2}$ is introduced to balance two terms in the right-hand side of (8.55). Indeed, by (8.3),

$$\min_{\mathbf{x} \in \overline{\Omega}} \big(e^{-2\lambda b^2} e^{2\lambda(z+b)^2}\big) = 1. \tag{8.56}$$

Remark 8.4.1. The CWF $e^{2\lambda(z+b)^2}$ is simpler than the CWF of Theorem 2.4.1, since the function $e^{2\lambda(z+b)^2}$ depends on one large parameter λ, whereas the CWF of Theorem 2.4.1 depends on two large parameters λ and ν. The main reason of such a simplification is our numerical observation that one should use the simplest possible CWFs in numerical implementations of the convexification.

Minimization problem. Minimize the functional $J_{\lambda,\beta}(Q)$ in (8.55) on the set $B(m,R)$ defined in (8.51).

8.4.2 Theorems

Theorem 8.4.1. *There exists a sufficiently large number $\lambda_0 = \lambda_0(\Omega, b) \geq 1$ and a constant $C_1 = C_1(\Omega, b) > 0$, both depending only on Ω and b, such that for all $\lambda \geq \lambda_0$ and for all functions $u \in H^2(\Omega)$ such that $u|_{\partial\Omega} = u_z|_{\Gamma_0} = 0$ the following Carleman estimate holds:*

$$\int_\Omega (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \geq \frac{C_1}{\lambda} \sum_{i,j=1}^{3} \int_\Omega u_{x_i x_j}^2 e^{2\lambda(z+b)^2} d\mathbf{x} + C_1 \lambda \int_\Omega ((\nabla u)^2 + \lambda^2 u^2) e^{2\lambda(z+b)^2} d\mathbf{x}. \tag{8.57}$$

Below $C_2 = C_2(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b) > 0$ denotes different constants depending only on listed parameters.

Theorem 8.4.2 (global strict convexity). *For all $Q \in B(m, 2R)$, $\lambda, \beta > 0$ there exists the Fréchet derivative $J'_{\lambda,\beta}(Q + G) \in H_0^3(\Omega)$. Let λ_0 be the number of Theorem* 8.4.1. *There exists a number $\lambda_1 = \lambda_1(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b) \geq \lambda_0$ depending only on listed parameters such that for any $\lambda \geq \lambda_1$ and any $\beta > 0$ the functional $J_{\lambda,\beta}(Q)$ is strictly convex on $B(m, R)$, that is, the following estimate holds for all $Q_1, Q_2 \in B(m, R)$:*

$$\begin{aligned} &J_{\lambda,\beta}(Q_2 + G) - J_{\lambda,\beta}(Q_1 + G) - J'_{\lambda,\beta}(Q_1 + G)(Q_2 - Q_1) \\ &\quad \geq \frac{C_2}{\lambda} \sum_{i,j=1}^{3} \|(Q_2 - Q_1)_{x_i x_j}\|_{L_2(\Omega)}^2 + C_2 \lambda \|Q_2 - Q_1\|_{H^1(\Omega)}^2 + \beta \|Q_2 - Q_1\|_{H^3(\Omega)}^2. \end{aligned} \tag{8.58}$$

Theorem 8.4.3. *The Fréchet derivative $J'_{\lambda,\beta}(Q + G) \in H_0^3(\Omega)$ of the functional $J_{\lambda,\beta}(Q)$ satisfies the Lipschitz continuity condition in $B(m, 2R)$ for all $\lambda, \beta > 0$. In other words, there exists a number $L = L(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b, \lambda, \beta)$ depending only on listed parameters such that*

$$\|J'_{\lambda,\beta}(Q_2 + G) - J'_{\lambda,\beta}(Q_1 + G)\|_{H^3(\Omega)} \leq L \|Q_2 - Q_1\|_{H^3(\Omega)}, \quad \forall Q_1, Q_2 \in B(m, 2R).$$

Theorem 8.4.4. *For each pair $\lambda \geq \lambda_1$, $\beta > 0$ there exists unique minimizer $Q_{\min,\lambda,\beta} \in \overline{B(m, R)}$ of the functional $J_{\lambda,\beta}(Q)$ on the set $\overline{B(m, R)}$. Furthermore,*

$$J'_{\lambda,\beta}(Q_{\min,\lambda,\beta} + G)(Q_{\min,\lambda,\beta} - p) \leq 0, \quad \forall p \in H_0^3(\Omega). \tag{8.59}$$

We now arrange the gradient projection method of the minimization of the functional $J_{\lambda,\beta}(Q+G)$ on the set $\overline{B(m,R)}$. Let the number $\gamma \in (0,1)$. Let $P_{\overline{B}} : H_0^3(\Omega) \to \overline{B(m,R)}$ be the projection operator of the space H_0^3 on the set $\overline{B(m,R)}$. Let $Q_0 \in B(m,R)$ be an arbitrary point of this set. The gradient projection method amounts to the following sequence:

$$Q_n = P_{\overline{B}}(Q_{n-1} - \gamma J'_{\lambda,\beta}(Q_{n-1} + G)), \quad n = 1, 2, \ldots. \tag{8.60}$$

Theorem 8.4.5. *Let λ_1 and β be parameters of Theorem* 8.4.2. *Choose a number $\lambda \geq \lambda_1$. Let $Q_{\min,\lambda,\beta} \in \overline{B(m,R)}$ be the unique minimizer of the functional $J_{\lambda,\beta}(Q)$ on the set $\overline{B(m,R)}$* (*Theorem* 8.4.3). *Then there exists a sufficiently small number $\gamma_0 = \gamma_0(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b, \beta) \in (0,1)$ depending only on listed parameters such that the sequence* (8.60) *converges to $Q_{\min,\lambda,\beta}$ in the space $H^3(\Omega)$. More precisely, there exists a number $\theta = \theta(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b, \beta) \in (0,1)$ such that the following estimate holds:*

$$\|Q_n - Q_{\min,\lambda,\beta}\|_{H^3(\Omega)} \leq \theta^n \|Q_{\min,\lambda,\beta} - Q_0\|_{H^3(\Omega)}. \tag{8.61}$$

Following the Tikhonov concept for ill-posed problems [22, 244], we now assume the existence of the exact solution $Q^*(\mathbf{x}) = (q_0^*(\mathbf{x}), \ldots, q_N^*(\mathbf{x}))^T \in B(m,R)$ of the problem (8.49), (8.50) with the noiseless data $G^*(\mathbf{x}) = (g_0^*(\mathbf{x}), \ldots, g_N^*(\mathbf{x}))^T \in H^3(\Omega)$. In particular, this means that

$$\sum_{n=1}^{N} P'_n(0)(q_n^*(\mathbf{x}) + g_n^*(\mathbf{x})) > m, \quad \forall \mathbf{x} \in \overline{\Omega}.$$

Also, let the number $\delta \in (0,1)$ be the level of the error in the data G, that is,

$$\|G - G^*\|_{H^3(\Omega)} < \delta. \tag{8.62}$$

Since $\delta \in (0,1)$, then (8.62) implies that we can regard in Theorem 8.4.6 that constants λ_1, C_2, γ_0, θ introduced above depend on $\|G^*\|_{H^3(\Omega)}$ rather than on $\|G\|_{H^3(\Omega)}$. We are doing so both in the formulation and in the proof of Theorem 8.4.6.

Theorem 8.4.6 (error estimates and global convergence). *Let $\lambda_1 = \lambda_1(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b)$ be the number of Theorem* 8.4.2. *Define the number η as $\eta = [4(A+b)^2]^{-1}$. Choose a sufficiently small number $\delta_0 \in (0,1)$ such that $\ln \delta_0^{-\eta} \geq \lambda_1$. Let in* (8.62) *$\delta \in (0, \delta_0)$. Choose $\lambda = \lambda(\delta) = \ln \delta^{-\eta} > \lambda_1$ implying that $\exp[2\lambda(\delta)(A+b)^2] = 1/\sqrt{\delta}$. Let $Q_{\min,\lambda,\beta} \in \overline{B(m,R)}$ be the unique minimizer of the functional $J_{\lambda,\beta}(Q)$ on the set $\overline{B(m,R)}$, the existence of which is guaranteed by Theorem* 8.4.4. *Let $\{Q_n\}_{n=0}^{\infty} \subset \overline{B(m,R)}$ be the sequence of the gradient projection method* (8.60) *with an arbitrary starting point $Q_0 \in \overline{B(m,R)}$. Then the following estimates hold for all $\beta \in (0,1)$:*

$$\|Q^* - Q_{\min,\lambda,\beta}\|_{H^1(\Omega)} \leq C_2(\delta^{\eta/2+1/4} + \sqrt{\beta}\delta^{\eta/2}), \tag{8.63}$$

$$\|Q^* - Q_{\min,\lambda,\beta}\|_{H^2(\Omega)} \le C_2(\delta^{1/4} + \sqrt{\beta})\sqrt{\ln \delta^{-\eta}}, \quad (8.64)$$

$$\|Q^* - Q_n\|_{H^1(\Omega)} \le C_2(\delta^{\eta/2+1/4} + \sqrt{\beta}\delta^{\eta/2}) + \theta^n \|Q_{\min,\lambda,\beta} - Q_0\|_{H^3(\Omega)}, \quad (8.65)$$

$$\|Q^* - Q_n\|_{H^2(\Omega)} \le C_2(\delta^{1/4} + \sqrt{\beta})\sqrt{\ln \delta^{-\eta}} + \theta^n \|Q_{\min,\lambda,\beta} - Q_0\|_{H^3(\Omega)}, \quad (8.66)$$

$$\|c^* - c_n\|_{L_2(\Omega)} \le C_2(\delta^{1/4} + \sqrt{\beta})\sqrt{\ln \delta^{-\eta}} + \theta^n \|Q_{\min,\lambda,\beta} - Q_0\|_{H^3(\Omega)}. \quad (8.67)$$

In particular, if the regularization parameter $\beta = \sqrt{\delta}$, as required by the regularization theory [244]*, then estimates* (8.63)–(8.66) *become*

$$\|Q^* - Q_{\min,\lambda,\beta}\|_{H^1(\Omega)} \le C_2\delta^{\eta/2+1/4},$$
$$\|Q^* - Q_{\min,\lambda,\beta}\|_{H^2(\Omega)} \le C_2\delta^{1/4}\sqrt{\ln \delta^{-\eta}},$$
$$\|Q^* - Q_n\|_{H^1(\Omega)} \le C_2\delta^{\eta/2+1/4} + \theta^n \|Q_{\min,\lambda,\sqrt{\delta}} - Q_0\|_{H^3(\Omega)},$$
$$\|Q^* - Q_n\|_{H^2(\Omega)} \le C_2\delta^{1/4}\sqrt{\ln \delta^{-\eta}} + \theta^n \|Q_{\min,\lambda,\sqrt{\delta}} - Q_0\|_{H^3(\Omega)},$$
$$\|c^* - c_n\|_{L_2(\Omega)} \le C_2\delta^{1/4}\sqrt{\ln \delta^{-\eta}} + \theta^n \|Q_{\min,\lambda,\sqrt{\delta}} - Q_0\|_{H^3(\Omega)}.$$

Here, $c_n(\mathbf{x})$ is defined via (8.48) *with $Q = Q_n$, next* (8.42)*, and next* (8.33)*. Further, $c^*(\mathbf{x})$ is defined as the exact target coefficient, which corresponds to the noiseless data G^* with Q^* and $W^* = Q^* + G^*$ in* (8.48)*, and next similarly as for $c_n(\mathbf{x})$.*

The presence of the regularization term $\beta\|Q + G\|^2_{H^3(\Omega)}$ in the functional $J_{\lambda,\beta}(Q + G)$ is important since this term ensures that in the gradient projection method (8.60) all functions $Q_n \in H^3_0(\Omega)$. Since $H^3(\Omega) \subset C^1(\overline{\Omega})$, and since we use estimates of $C^1(\overline{\Omega})$-norms of some functions in the proof of Theorem 8.4.2, then we indeed need $Q_n \in H^3_0(\Omega)$.

Remark 8.4.2. Since $Q_0 \in B(m, R)$ is an arbitrary point and $R > 0$ is an arbitrary number, then Theorem 8.4.6 implies the *global* convergence of the gradient projection method (8.60); see Definition 1.4.2.

8.5 Proofs

The proof of Theorem 8.4.3 is very similar with the proof of Theorem 5.3.1. As soon as Theorems 8.4.2 and 8.4.3 are proven, the proof of Theorem 8.4.4 is quite similar with the proof of Lemma 5.2.1. Next, as soon as Theorems 8.4.2 and 8.4.4 are proven, the proof of Theorem 8.4.5 is again quite similar with the proof of Theorem 5.2.1. Hence, we prove here only Theorems 8.4.1, 8.4.2, and 8.4.6.

8.5.1 Proof of Theorem 8.4.1

In this proof, the function $u \in C^3(\overline{\Omega})$. The case $u \in H^2(\Omega)$ follows from density arguments. Consider the function $v = ue^{\lambda(z+b)^2}$. Then $u = ve^{-\lambda(z+b)^2}$. Hence, $u_{xx} =$

$v_{xx}e^{-\lambda(z+b)^2}$, $u_{yy} = v_{yy}e^{-\lambda(z+b)^2}$,

$$u_{zz} = \left(v_{zz} - 2\lambda(z+b)v_z + 4\lambda^2(z+b)^2(1+O(1/\lambda))v\right)e^{-\lambda(z+b)^2}.$$

In this proof, $C_1 = C_1(\Omega, b) > 0$ denotes different constants depending only on Ω and b and $O(1/\lambda)$ denotes different z-dependent functions satisfying $|O(1/\lambda)|, |\nabla O(1/\lambda)| \le C_1/\lambda$, Hence,

$$\begin{aligned}(\Delta u)^2 e^{2\lambda(z+b)^2} &= \left[\left(v_{xx} + v_{yy} + v_{zz} + 4\lambda^2(z+b)^2(1+O(1/\lambda))v\right) - 2\lambda(z+b)v_z\right]^2 \\ &\ge -4\lambda(z+b)v_z\left(v_{xx} + v_{yy} + v_{zz} + 4\lambda^2(z+b)^2(1+O(1/\lambda))v\right) \\ &= \left(-4\lambda(z+b)v_zv_x\right)_x + 4\lambda(z+b)v_{zx}v_x + \left(-4\lambda(z+b)v_zv_y\right)_y + 4\lambda(z+b)v_{zy}v_y \\ &\quad + \left(-2\lambda(z+b)v_z^2\right)_z + 2\lambda v_z^2 + \left(-8\lambda^3(z+b)^3(1+O(1/\lambda))v^2\right)_z \\ &\quad + 24\lambda^3(z+b)^2(1+O(1/\lambda))v^2 \\ &= -2\lambda\left(v_x^2 + v_y^2\right) + 2\lambda v_z^2 + 24\lambda^3(z+b)^2(1+O(1/\lambda))v^2 \\ &\quad + \left(-2\lambda(z+b)v_z^2 + 2\lambda(z+b)v_x^2 + 2\lambda(z+b)v_y^2 - 8\lambda^3(z+b)^3(1+O(1/\lambda))v^2\right)_z \\ &\quad + \left(-4\lambda(z+b)v_zv_x\right)_x + \left(-2\lambda(z+b)v_zv_y\right)_y.\end{aligned}$$

Since $v|_{\partial\Omega} = v_z|_{\Gamma_0} = 0$ and $2\lambda v_z^2 \ge 0$, then integrating the above over Ω, going back from v to u and using Gauss' formula, we obtain for sufficiently large $\lambda \ge C_1$,

$$\int_\Omega (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \ge -2\lambda \int_\Omega \left(u_x^2 + u_y^2\right)^2 e^{2\lambda(z+b)^2} d\mathbf{x} + 23\lambda^3 \int_\Omega u^2 e^{2\lambda(z+b)^2} d\mathbf{x}. \tag{8.68}$$

Next,

$$\begin{aligned}-u\Delta u e^{2\lambda(z+b)^2} &= \left(-u_x u e^{2\lambda(z+b)^2}\right)_x + u_x^2 e^{2\lambda(z+b)^2} + \left(-u_y u e^{2\lambda(z+b)^2}\right)_y + u_y^2 e^{2\lambda(z+b)^2} \\ &\quad + \left(-u_z u e^{2\lambda(z+b)^2}\right)_z + u_z^2 e^{2\lambda(z+b)^2} + 4\lambda(z+b)u_z u e^{2\lambda(z+b)^2} \\ &= \left(u_x^2 + u_y^2 + u_z^2\right)e^{2\lambda(z+b)^2} + \left(2\lambda(z+b)u^2 e^{2\lambda(z+b)^2}\right)_z \\ &\quad - 8\lambda^2(z+b)^2(1+O(1/\lambda))u^2 e^{2\lambda(z+b)^2} \\ &\quad + \left(-u_x u e^{2\lambda(z+b)^2}\right)_x + \left(-u_y u e^{2\lambda(z+b)^2}\right)_y.\end{aligned}$$

Integrating this over Ω and using Gauss' formula, we obtain for sufficiently large $\lambda \ge C_1$,

$$\begin{aligned}-\int_\Omega u\Delta u e^{2\lambda(z+b)^2} d\mathbf{x} &= \int_\Omega \left(u_x^2 + u_y^2 + u_z^2\right)e^{2\lambda(z+b)^2} d\mathbf{x} \\ &\quad - 9\lambda^2 \int_\Omega (z+b)^2 u^2 e^{2\lambda(z+b)^2} d\mathbf{x}.\end{aligned} \tag{8.69}$$

Multiply (8.69) by $5\lambda/2$ and sum up with (8.68). Since $23\lambda^3 - (9 \cdot 5/2)\lambda^3 = 3\lambda^3/2$, then

$$-\frac{5}{2}\lambda \int_{\Omega} u\Delta u e^{2\lambda(z+b)^2} d\mathbf{x} + \int_{\Omega} (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \geq \frac{1}{2}\lambda \int_{\Omega} (u_x^2 + u_y^2 + u_z^2) e^{2\lambda(z+b)^2} d\mathbf{x} + \frac{3}{2}\lambda^3 \int_{\Omega} u^2 e^{2\lambda(z+b)^2} d\mathbf{x}. \quad (8.70)$$

Next, applying Cauchy–Schwarz inequality, we obtain

$$-\frac{5}{2}\lambda \int_{\Omega} u\Delta u e^{2\lambda(z+b)^2} d\mathbf{x} + \int_{\Omega} (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \leq \frac{25}{4}\lambda^2 \int_{\Omega} u^2 e^{2\lambda(z+b)^2} d\mathbf{x} + \frac{1}{2} \int_{\Omega} (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x}.$$

Combining this with (8.70), we obtain for sufficiently large $\lambda_0 = \lambda_0(\Omega, b) > 0$ and for $\lambda \geq \lambda_0$,

$$\int_{\Omega} (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \geq C_1 \lambda \int_{\Omega} ((\nabla u)^2 + \lambda^2 u^2) e^{2\lambda(z+b)^2} d\mathbf{x}. \quad (8.71)$$

The next step is to incorporate the term with second derivatives in (8.57). We have

$$\begin{aligned}(\Delta u)^2 e^{2\lambda(z+b)^2} &= (u_{xx} + u_{yy} + u_{zz})^2 e^{2\lambda(z+b)^2} \\ &= (u_{xx}^2 + u_{yy}^2 + u_{zz}^2) e^{2\lambda(z+b)^2} \\ &\quad + 2(u_{xx}u_{yy} + u_{xx}u_{zz} + u_{yy}u_{zz}) e^{2\lambda(z+b)^2}.\end{aligned} \quad (8.72)$$

The second line of (8.57) gives:

$$\begin{aligned}&(2u_{xx}u_{yy} + 2u_{xx}u_{zz} + 2u_{yy}u_{zz}) e^{2\lambda(z+b)^2} \\ &\quad = \left(2u_{xx}u_y e^{2\lambda(z+b)^2}\right)_y - 2u_{xxy}u_y e^{2\lambda(z+b)^2} \\ &\quad\quad + \left(2u_{xx}u_z e^{2\lambda(z+b)^2}\right)_z - 2u_{xxz}u_z e^{2\lambda(z+b)^2} - 8\lambda(z+b)u_{xx}u_z e^{2\lambda(z+b)^2} \\ &\quad\quad + \left(2u_{yy}u_z e^{2\lambda(z+b)^2}\right)_z - 2u_{yyz}u_z e^{2\lambda(z+b)^2} - 8\lambda(z+b)u_{yy}u_z e^{2\lambda(z+b)^2} \\ &\quad = 2(u_{xy}^2 + u_{xz}^2 + u_{yz}^2) e^{2\lambda(z+b)^2} + \left[2(u_{xy}u_y - u_{xz}u_z) e^{2\lambda(z+b)^2}\right]_x \\ &\quad\quad + \left[2(u_{xx}u_y - u_{xz}u_z) e^{2\lambda(z+b)^2}\right]_y - 8\lambda(z+b)u_{xx}u_z e^{2\lambda(z+b)^2} - 8\lambda(z+b)u_{yy}u_z e^{2\lambda(z+b)^2}.\end{aligned} \quad (8.73)$$

Using the Cauchy–Schwarz inequality, we estimate from the below the last line of (8.73) as

$$\begin{aligned}&- 8\lambda(z+b)u_{xx}u_z e^{2\lambda(z+b)^2} - 8\lambda(z+b)u_{yy}u_z e^{2\lambda(z+b)^2} \\ &\quad \geq -\frac{1}{2}(u_{xx}^2 + u_{yy}^2) e^{2\lambda(z+b)^2} - 64\lambda^2(z+b)^2 u_z^2 e^{2\lambda(z+b)^2}.\end{aligned} \quad (8.74)$$

Combining (8.73)–(8.74), we obtain

$$\int_\Omega (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \geq \frac{1}{2}\sum_{i,j=1}^{3}\int_\Omega u_{x_i x_j}^2 e^{2\lambda(z+b)^2} d\mathbf{x} - 64\lambda^2 \int_\Omega (z+b)^2 u_z^2 e^{2\lambda(z+b)^2} d\mathbf{x}.$$

Multiply this estimate by $C_1/(128\lambda)$ and sum up with (8.71). We obtain

$$\begin{aligned}\left(1+\frac{C_1}{128\lambda}\right)\int_\Omega (\Delta u)^2 e^{2\lambda(z+b)^2} d\mathbf{x} \geq \frac{C_1}{256\lambda}\sum_{i,j=1}^{3}\int_\Omega u_{x_i x_j}^2 e^{2\lambda(z+b)^2} d\mathbf{x} \\ + \frac{C}{2}\lambda \int_\Omega ((\nabla u)^2 + \lambda^2 u^2) e^{2\lambda(z+b)^2} d\mathbf{x}.\end{aligned} \tag{8.75}$$

Since $C_1 > 0$ denotes different constants, then the target estimate (8.57) follows from (8.59) immediately.

8.5.2 Proof of Theorem 8.4.2

Denote $h = Q_2 - Q_1$ implying that $Q_2 = Q_1 + h$. Also, $h \in H_0^3(\Omega)$, $\|h\|_{H^3(\Omega)} < 2R$. Hence, by (8.52)

$$\|h\|_{C^1(\overline{\Omega})} < D_1(2R). \tag{8.76}$$

Using the multidimensional analog of Taylor formula (see, e. g., [247] for this formula) and (8.54), we obtain

$$\begin{aligned}&\Delta h + (\Delta Q_1 + \Delta G) + F\big(h + Q_1 + G, \nabla(h + Q_1 + G)\big) \\ &\quad = \big[\Delta h + F^{(1)}(Q_1 + G, \nabla(Q_1 + G))h + F^{(2)}(Q_1 + G, \nabla(Q_1 + G))\nabla h\big] \\ &\qquad + F_{\text{nonlin}}\big(h, \nabla h, Q_1 + G, \nabla(Q_1 + G)\big) + \big[(\Delta Q_1 + \Delta G) + F\big(Q_1 + G, \nabla(Q_1 + G)\big)\big],\end{aligned} \tag{8.77}$$

where elements of $(N+1)\times(N+1)$ matrix $F^{(1)}$ and $(3N+3)\times(3N+3)$ matrix are bounded for $\mathbf{x} \in \overline{\Omega}$, that is,

$$\big|F_{i,j}^{(1)}(Q_1 + G, \nabla(Q_1 + G))\big|, \big|F_{i,j}^{(2)}(Q_1 + G, \nabla(Q_1 + G))\big| \leq C_2, \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{8.78}$$

where the subscript "i, j" denotes an arbitrary element of the corresponding matrix indexed as (i, j). Next, the $(N+1)$-dimensional vector function F_{nonlin} depends nonlinearly on h, ∇h. Furthermore, the following estimate follows from (8.52)–(8.54):

$$\big|F_{\text{nonlin}}\big(h, \nabla h, Q_1 + G, \nabla(Q_1 + G)\big)\big| \leq C_2\big(|h|^2 + |\nabla h|^2\big), \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{8.79}$$

Next, (8.76) and (8.79) imply with a different constant C_2,

$$\big|F_{\text{nonlin}}\big(h, \nabla h, Q_1 + G, \nabla(Q_1 + G)\big)\big| \leq C_2\big(|h| + |\nabla h|\big), \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{8.80}$$

It follows from (8.77)–(8.80) that

$$\begin{aligned}&[\Delta h + (\Delta Q_1 + \Delta G) + F(h + Q_1 + G, \nabla(h + Q_1 + G))]^2\\&\quad - [(\Delta Q_1 + \Delta G) + F(Q_1 + G, \nabla(Q_1 + G))]^2\\&= \mathrm{Lin}_1(\Delta h) + \mathrm{Lin}_2(\nabla h) + \mathrm{Lin}_3(h)\\&\quad + (\Delta h)^2 + M_1(h, \nabla h, Q_1 + G, \nabla(Q_1 + G))\Delta h + M_2(h, \nabla h, Q_1 + G, \nabla(Q_1 + G)),\end{aligned} \tag{8.81}$$

where expressions $\mathrm{Lin}_1(\Delta h)$, $\mathrm{Lin}_2(\nabla h)$ and $\mathrm{Lin}_3(h)$ are linear with respect to Δh, ∇h, and h, respectively, and

$$\left|\mathrm{Lin}_1(\Delta h) + \mathrm{Lin}_2(\nabla h) + \mathrm{Lin}_3(h)\right| \leq C_2(|\Delta h| + |\nabla h| + |h|), \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{8.82}$$

Next, the following estimates are valid for M_1 and M_2:

$$\left|M_1(h, \nabla h, Q_1 + G, \nabla(Q_1 + G))\right| \leq C_2(|\nabla h| + |h|), \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{8.83}$$

$$\left|M_2(h, \nabla h, Q_1 + G, \nabla(Q_1 + G))\right| \leq C_2(|\nabla h|^2 + |h|^2), \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{8.84}$$

In particular, (8.83), (8.84), and the Cauchy–Schwarz inequality imply

$$\begin{aligned}&(\Delta h)^2 + M_1(h, \nabla h, Q_1 + G, \nabla(Q_1 + G))\Delta h + M_2(h, \nabla h, Q_1 + G, \nabla(Q_1 + G))\\&\quad \geq \frac{1}{2}(\Delta h)^2 - C_2(|\nabla h|^2 + |h|^2), \quad \forall \mathbf{x} \in \overline{\Omega}.\end{aligned} \tag{8.85}$$

Using (8.55) and (8.81)–(8.84), we obtain

$$J_{\lambda,\beta}(Q_1 + h) - J_{\lambda,\beta}(Q_1) = X_{\mathrm{lin}}(h) + X_{\mathrm{nonlin}}(h), \tag{8.86}$$

where $X_{\mathrm{lin}}(h)$ can be extended from $\{\|h\|_{H^3(\Omega)} < 2R\} \subset H_0^3(\Omega)$ to the entire space $H^3(\Omega)$ as a bounded linear functional,

$$X_{\mathrm{lin}}(h) = e^{-2\lambda b^2} \int_{\Omega} (\mathrm{Lin}_1(\Delta h) + \mathrm{Lin}_2(\nabla h) + \mathrm{Lin}_3(h))(\mathbf{x}) e^{2\lambda(z+b)^2} d\mathbf{x} + 2\beta[h, Q_1 + G], \tag{8.87}$$

where $[\cdot,\cdot]$ is the scalar product in $H^3(\Omega)$. As to $X_{\mathrm{nonlin}}(h)$ in (8.86), it follows from (8.55), (8.81), (8.83), and (8.84) that

$$\lim_{\|h\|_{H^3(\Omega)} \to 0} \frac{X_{\mathrm{nonlin}}(h)}{\|h\|_{H^3(\Omega)}} = 0. \tag{8.88}$$

Using (8.82) and (8.86)–(8.88), we obtain that $X_{\mathrm{lin}}(h)$ is the Frechét derivative $J'_{\lambda,\beta}(Q)$ of the functional $J_{\lambda,\beta}(Q)$ at the point Q, that is, $X_{\mathrm{lin}}(h) = J'_{\lambda,\beta}(Q_1)(h)$. Thus, the existence of the Fréchet derivative is established.

Next, using (8.55) and (8.81)–(8.87), we obtain

$$J_{\lambda,\beta}(Q_1 + G + h) - J_{\lambda,\beta}(Q_1 + G) - J'_{\lambda,\beta}(Q_1 + G)(h) \tag{8.89}$$
$$\geq \frac{1}{2}e^{-2\lambda b^2}\int_\Omega (\Delta h)^2 e^{2\lambda(z+b)^2}d\mathbf{x} - C_2 e^{-2\lambda b^2}\int_\Omega (|\nabla h|^2 + |h|^2)e^{2\lambda(z+b)^2}d\mathbf{x} + \beta\|h\|^2_{H^3(\Omega)}.$$

We now apply Carleman estimate (8.57), assuming that $\lambda \geq \lambda_0$,

$$\frac{1}{2}e^{-2\lambda b^2}\int_\Omega (\Delta h)^2 e^{2\lambda(z+b)^2}d\mathbf{x} - C_2 e^{-2\lambda b^2}\int_\Omega (|\nabla h|^2 + |h|^2)e^{2\lambda(z+b)^2}d\mathbf{x} + \beta\|h\|^2_{H^3(\Omega)}$$
$$\geq \frac{C_1}{\lambda}\sum_{i,j=1}^{3} e^{-2\lambda b^2}\int_\Omega h^2_{x_i x_j} e^{2\lambda(z+b)^2}d\mathbf{x} + C_1\lambda e^{-2\lambda b^2}\int_\Omega ((\nabla h)^2 + \lambda^2 h^2)e^{2\lambda(z+b)^2}d\mathbf{x}$$
$$- C_2 e^{-2\lambda b^2}\int_\Omega (|\nabla h|^2 + |h|^2)e^{2\lambda(z+b)^2}d\mathbf{x} + \beta\|h\|^2_{H^3(\Omega)}.$$

Choosing sufficiently large $\lambda_1 = \lambda_1(F, \|G\|_{H^3(\Omega)}, m, R, \Omega, b) \geq \lambda_0$ and letting $\lambda \geq \lambda_1$, we obtain with a different constant C_2,

$$\frac{1}{2}e^{-2\lambda b^2}\int_\Omega (\Delta h)^2 e^{2\lambda(z+b)^2}d\mathbf{x} - C_2 e^{-2\lambda b^2}\int_\Omega (|\nabla h|^2 + |h|^2)e^{2\lambda(z+b)^2}d\mathbf{x} + \beta\|h\|^2_{H^3(\Omega)}$$
$$\geq \frac{C_2}{\lambda}\sum_{i,j=1}^{3} e^{-2\lambda b^2}\int_\Omega h^2_{x_i x_j} e^{2\lambda(z+b)^2}d\mathbf{x} + C_2\lambda e^{-2\lambda b^2}\int_\Omega ((\nabla h)^2 + \lambda^2 h^2)e^{2\lambda(z+b)^2}d\mathbf{x} + \beta\|h\|^2_{H^3(\Omega)}.$$

This, (8.89), and (8.56) imply (8.58).

8.5.3 Proof of Theorem 8.4.6

We rewrite the functional $J_{\lambda,\beta}(Q)$ in (8.55) as

$$J_{\lambda,\beta}(Q + G) = J^0_{\lambda,\beta}(Q + G) + \beta\|Q + G\|^2_{H^3(\Omega)}. \tag{8.90}$$

Since the vector function $Q^* \in B(m, R)$ is the exact solution of the problem (8.49), (8.50) with the noiseless data G^*, then $J^0_{\lambda,\beta}(Q^* + G^*) = 0$. Hence,

$$J_{\lambda,\beta}(Q^* + G^*) \leq C_2\beta. \tag{8.91}$$

Next, $J_{\lambda,\beta}(Q^* + G) = (J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q^* + G^*)) + J_{\lambda,\beta}(Q^* + G^*)$. Hence, applying (8.91), we obtain

$$J_{\lambda,\beta}(Q^* + G) \leq \left|J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q^* + G^*)\right| + C_2\beta. \tag{8.92}$$

Using (8.62) and (8.90), we estimate the first term in the right-hand side of (8.92),

$$\begin{aligned}\left|J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q^* + G^*)\right| &\le \left|J^0_{\lambda,\beta}(Q^* + G) - J^0_{\lambda,\beta}(Q^* + G^*)\right| + C_2\beta\delta \\ &\le C_2\delta\exp(2\lambda(A+b)^2) + C_2\beta\delta. \end{aligned} \tag{8.93}$$

Recall that due to our choice $\lambda = \lambda(\delta) = \ln\delta^{-\eta}$, where $\eta = [4(A+b)^2]^{-1}$, we have $\delta\exp(2\lambda(A+b)^2) = 1/\sqrt{\delta}$. Hence, (8.93) implies

$$\left|J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q^* + G^*)\right| \le C_2\sqrt{\delta}, \quad \forall\beta \in (0,1).$$

Combining this with (8.91), we obtain

$$J_{\lambda,\beta}(Q^* + G) \le C_2(\sqrt{\delta} + \beta). \tag{8.94}$$

Until now, we have not used in this proof the strict convexity of the functional $J_{\lambda,\beta}(Q+G)$ for $Q \in B(m,R)$. But we will use this property in the rest of the proof. Recall that by Theorem 8.4.3 $Q_{\min,\lambda,\beta} \in \overline{B(m,R)}$ is the unique minimizer on the set $\overline{B(m,R)}$ of the functional $J_{\lambda,\beta}(Q+G)$ on the set $\overline{B(m,R)}$. By Theorem 8.4.2,

$$\begin{aligned}&J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q_{\min,\lambda,\beta} + G) - J'_{\lambda,\beta}(Q_{\min,\lambda,\beta})(Q^* - Q_{\min,\lambda,\beta}) \\ &\quad\ge \frac{C_2}{\lambda}\sum_{i,j=1}^{2}\|(Q^* - Q_{\min,\lambda,\beta})_{x_ix_j}\|^2_{L_2(\Omega)} + C_2\lambda\|Q^* - Q_{\min,\lambda,\beta}\|^2_{H^1(\Omega)} + \frac{\beta}{2}\|Q^* - Q_{\min,\lambda,\beta}\|^2_{H^3(\Omega)}.\end{aligned} \tag{8.95}$$

Since by (8.59) $-J'_{\lambda,\beta}(Q_{\min,\lambda,\beta})(Q^* - Q_{\min,\lambda,\beta}) \le 0$, then (8.94) implies that the left-hand side of (8.95) can be estimated as

$$J_{\lambda,\beta}(Q^* + G) - J_{\lambda,\beta}(Q_{\min,\lambda,\beta} + G) - J'_{\lambda,\beta}(Q_{\min,\lambda,\beta} + G)(Q^* - Q_{\min,\lambda,\beta}) \le C_2(\sqrt{\delta} + \beta).$$

Hence, using our choice of $\lambda = \lambda(\delta) = \ln\delta^{-\eta}$ and (8.95), we obtain estimates (8.63) and (8.64). Estimates (8.65) and (8.66) are obtained from (8.63) and (8.64), respectively, using (8.61) and the triangle inequality. Estimate (8.67) obviously follows from estimate (8.66).

8.6 Numerical studies

The single point source is now $\mathbf{x}_0 = (0,0,-5)$. We choose in (8.3) the numbers $A = 1$. Hence, below

$$\begin{aligned}&\Omega = \{-1/2 < x, y < 1/2, z \in (0,1)\}, \\ &\Gamma_0 = \{\mathbf{x} = (x,y,z) : -1/2 < x, y < 1/2, z = 1\}, \quad \Gamma_1 = \partial\Omega \setminus \Gamma_0.\end{aligned} \tag{8.96}$$

We have introduced the vector function G in Section 8.4.1, and thus obtained the problem (8.49), (8.50) for the vector function $Q = W - G$ from the problem (8.43), (8.44) for the vector function W in order to obtain zero boundary conditions (8.50) for Q.

The latter was convenient for proofs of Theorems 8.4.3–8.4.6. However, it follows from Theorems 8.4.2–8.4.6 that their obvious analogs hold true for the functional

$$J_{\lambda,\beta}(W) = e^{-2\lambda b^2} \int_{\Omega} (\Delta W + F(\nabla W, W))^2 e^{2\lambda(z+b)^2} d\mathbf{x} + \beta \|W\|^2_{H^3(\Omega)}, \tag{8.97}$$

$$W \in B_W(m,R) = \{W : W = Q - G, \forall Q \in B(m,R)\}. \tag{8.98}$$

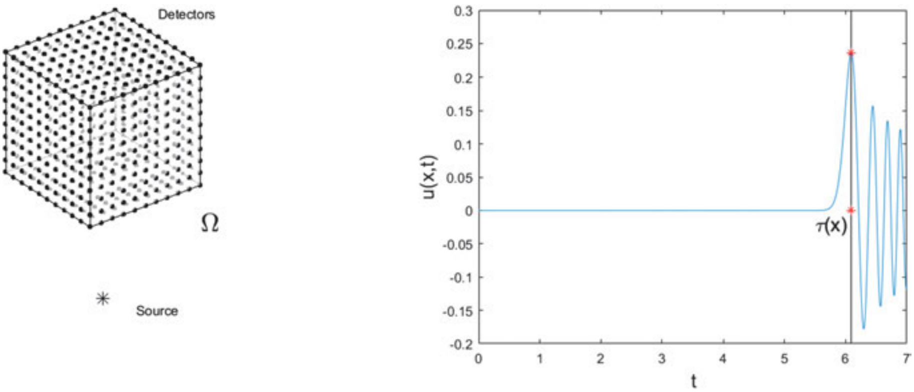
Furthermore, we use in (8.97) $\beta = 0$, $b = 0$. Therefore, we ignore the multiplier $e^{-2\lambda b^2}$, which was used above to balance first and second terms in the right-hand side of (8.55); see (8.56). Hence, we minimize the weighted cost functional

$$J_{\lambda}(W) = \int_{\Omega} (\Delta W + F(\nabla W, W))^2 e^{2\lambda z^2} d\mathbf{x} \tag{8.99}$$

on the set (8.98). We conjecture that the case $\beta = 0$ works probably because the minimal mesh size of 1/32 in the finite differences we use to minimize this functional is not too small, and all norms in finite dimensional spaces are equivalent; see item 3 of Remark 8.4.1. In addition, recall that, by the same item, one can choose any value of $\beta \in (0,1)$ in convergence estimates (8.63)–(8.66). Also, we use the gradient descent method (GD) instead of a more complicated gradient projection method. We have observed that GD works well for our computations. The latter coincides with observations in all earlier publications about numerical studies of the convexification [9, 115–117, 142–145, 145, 146, 150, 151, 164]. As to our choice $b = 0$, one can derive from the proof of Theorem 9.4.1 that a slightly modified Carleman estimate of this theorem works in a little bit smaller domain $\Omega' = \Omega \cap \{z > \varepsilon\}$ for any small number $\varepsilon > 0$. Finally, we believe that simplifications listed in this section work well numerically due to a commonly known observation that numerical studies are usually less pessimistic than the theory is.

8.6.1 Some details of the numerical implementation

To solve the inverse problem, we should first computationally simulate the data (8.14) at $\partial\Omega$ via the numerical solution of the forward problem (8.11), (8.12). To solve the problem (8.11), (8.12) computationally, we have used the standard finite difference method. To avoid the use of the infinite space $\mathbb{R}^3$ in the solution of the forward problem, we choose the cube $\Omega_f = \{-6.5 < x, y < 6.5, z \in (-6, 7)\}$. So that $\Omega \subset \Omega_f$, $\partial\Omega \cap \partial\Omega_f = \varnothing$ and $\mathbf{x}_0 = (0,0,-5) \in \Omega_f$. We choose a sufficiently large number $T_0 = 6.5$. Then we solve equation (8.11) with the initial condition (8.12) and zero Dirichlet boundary condition at $\partial\Omega_f$ for $(\mathbf{x}, t) \in \Omega_f \times (0, T_0)$ via finite differences. Indeed, the wave originated at $\mathbf{x}_0$ cannot reach neither vertical sides of Ω_f nor the upper side $\{z = 7\} \cap \overline{\Omega_f}$ of Ω for times $t \in (0, 6.5)$. However, it reaches the upper side $\{z = 1\} \cap \overline{\Omega}$ of Ω at $t = 6$. This wave reaches the lower side $\{z = -6\} \cap \overline{\Omega_f}$ of Ω_f at $t = 1$, which means that the zero Dirichlet boundary condition on the lower side is incorrect. Still, for $t \in (0, 6.5)$, the wave reflected from



(a) (b)

Figure 8.1: (a) A schematic diagram of domains Ω, source, and detectors. (b) This figure explains how do we approximately choose the boundary condition $\tau(\mathbf{x})|_{\partial\Omega}$. We have chosen here a selected point $\mathbf{x} \in \Gamma_0$.

the lower side of Ω_f does not reach the upper side $\{z = 1\} \cap \overline{\Omega}$ of Ω, where the data for our CIP are given. Hence, this reflected wave does not affect our data, see Figure 8.1.

We use the explicit scheme. The grid step size in each spatial direction is $\Delta x = 1/32$ and in time direction $\Delta t = 0.002$. To avoid a substantial increase of the computational time, we do not decrease these step sizes. When solving the forward problem, we model the $\delta(\mathbf{x} - \mathbf{x}_0)$-function in (8.12) as

$$\tilde{\delta}(\mathbf{x} - \mathbf{x}_0) = \begin{cases} \frac{1}{\varepsilon} \exp\left(-\frac{1}{1-|\mathbf{x}-\mathbf{x}_0|^2/\varepsilon}\right), & \text{if } |\mathbf{x} - \mathbf{x}_0|^2 < \varepsilon = 0.01, \\ 0, & \text{otherwise.} \end{cases}$$

In computations of the inverse problem, for each test we use, we choose in the data (8.14) $T = \max_{\overline{\Omega}} \tau(\mathbf{x}) + 0.1$. We have observed that $T < T_0$ in all our tests. Next, we set $T_1 = T - \max_{\overline{\Omega}} \tau(\mathbf{x}) = 0.1$. An important question is on how do we figure out boundary conditions at $\partial\Omega$ for the function $\tau(\mathbf{x})$, that is, $\tau(\mathbf{x})|_{\partial\Omega}$ and also $\partial_z \tau(\mathbf{x})|_{\Gamma_0}$. In principle, for $\mathbf{x} \in \partial\Omega$, one should choose such a number $\tau_0(\mathbf{x})$ that $\tau_0(\mathbf{x}) = \min_t \{t : u(\mathbf{x}, t) > 0\}$. However, it is hard to choose in practice the number $\tau_0(\mathbf{x})$ satisfying this criterion. Therefore, we calculate such a number $\tilde{t}(\mathbf{x})$ at which the first wave with the largest maximal value arrives at the point $\mathbf{x} \in \partial\Omega$; see Figure 8.2. Next, we set $\tau_0(\mathbf{x})|_{\partial\Omega} := \tilde{t}(\mathbf{x})|_{\partial\Omega}$. To calculate the derivative $\partial_z \tau_0(\mathbf{x})|_{\Gamma_0}$, we compute the discrete normal derivative of $\tau_0(\mathbf{x})$ over the mesh in the forward problem.

To minimize the weighted cost functional $J_\lambda(W)$ in (8.98), we act similarly with the previous above cited works about numerical studies of the convexification for a number of other CIPs. More precisely, we write the differential operators involved in $J_\lambda(W)$ via finite differences and minimize with respect to the values of the discrete analog of the vector function W at grid points. As to the choice of the parameter λ, even though the above theory works only for sufficiently large values of λ, we have established in our computational experiments that the choice $\lambda = 1$ is an optimal one for all tests we have performed. This again repeats an observation of all above cited works on numerical studies of the convexification, in which the optimal choice was $\lambda \in [1,3]$. We have also tested two different values of the number N terms in the series (8.37): $N = 1$ and $N = 3$. Our computational results indicate that $N = 3$ provides results of a good quality. In all tests below, the starting point of GD is the vector function $W(\mathbf{x})$, which is generated by the coefficient $c(\mathbf{x}) \equiv 1$ in equation (8.11), $W_{c\equiv 1}(\mathbf{x})$.

8.6.2 A multilevel minimization method of the functional $J_\lambda(W)$

We have found in our computational experiments that the gradient descent method for our weighted cost functional $J_{\lambda,0}(W)$ converges rapidly on a coarse mesh. This provides us with a rough image. Hence, we have implemented a multi-level method [189]. Let $M_{h_1} \subset M_{h_2} \subset \cdots \subset M_{h_K}$ be nested finite difference meshes, that is, M_{h_k} is a refinement of $M_{h_{k-1}}$ for $k \le K$. Let P_{h_k} be the corresponding finite difference functional space. One the first level M_{h_1}, we solve the discrete optimization problem. In other words, let $W_{h_1,\min}$ be the minimizer on the finite difference analog of the set (8.98) of the following functional:

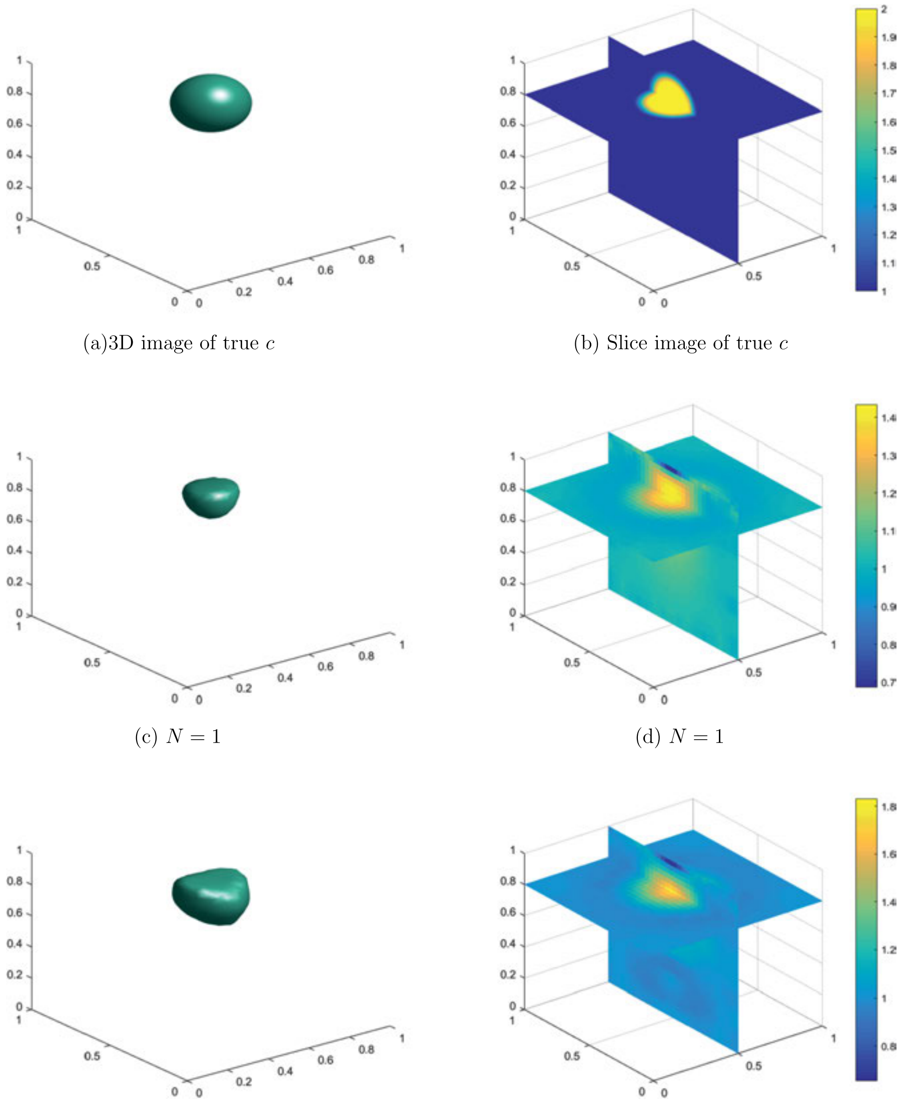
$$J_\lambda(W_{h_1}) = \int_\Omega \left(\Delta W_{h_1} + F(\nabla W_{h_1}, W_{h_1})\right)^2 e^{2\lambda z^2} d\mathbf{x}. \tag{8.100}$$

In (8.100) the integral and the derivatives are understood in the discrete sense, and $W_{h_1,\min}$ is found via the GD. Then we interpolate the minimizer $W_{h_1,\min}$ on the finer mesh M_{h_2} and use the resulting vector function $W_{h_2,\mathrm{int}}$ as the starting point of the gradient descent method of the optimization of the direct analog of functional (8.100) in which h_1 is replaced with h_2 and W_{h_1} is replaced with W_{h_2}. This process is repeated until we obtain the minimizer $W_{h_K,\min}$ on the K_{th} level on the mesh M_{h_K}.

Since $(x,y,z) \in (-1/2,1/2) \times (-1/2,1/2) \times (0,1)$, then our first level M_{h_1} is set to be the uniform mesh with the grid step $h_1 = 1/8$. For each mesh refinement, we will refine the mesh via setting the new grid step of the refined mesh in all directions to be 1/2 of the previous grid step. On each level M_{h_k}, we stop iterations as soon as we see that $\|\nabla J_\lambda^{(h_k)}(W_{h_k})\| \le 2 \times 10^{-2}$. Next, we refine the mesh and compute the solution on the next level $M_{h_{k+1}}$. In the end, we compute our approximation for the target coefficient $c(\mathbf{x})$ using the final minimizer $W_{h_K,\min}$.

8.6.3 Numerical testing

In the tests of this section, we demonstrate the efficiency of our numerical method for imaging of small inclusions as well as for imaging of a smoothly varying function $c(\mathbf{x})$.



(a) 3D image of true c (b) Slice image of true c

(c) $N = 1$ (d) $N = 1$

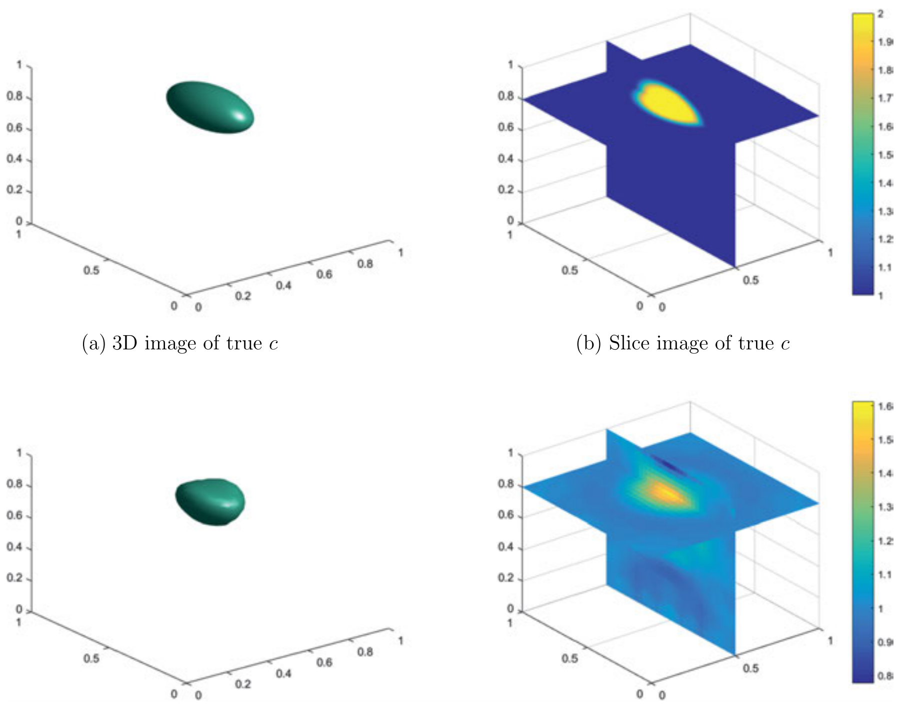
(e) $N = 3$ (f) $N = 3$

Figure 8.2: Results of Test 1. Imaging of one ball-shaped inclusion with $c = 2$ in it and $c = 1$ outside. Hence, the inclusion/background contrast is 2:1. We have stopped at the 3rd mesh refinement for all three values of N. (a) and (b) Correct images. (c) and (d) Computed images for $N = 1$. (e) and (f) Computed images for $N = 3$. The maximal value of the computed coefficient $c(\mathbf{x})$ is approximately 1.8.

In all tests, the background value of $c_{\text{bkgr}} = 1$. Note that a postprocessing of images was not applied in our numerical tests. As to the derivatives, we refer to Section 7.6.3. In Figures 7.2–7.8, slices are depicted to demonstrate the values of the computed function $c(\mathbf{x})$.

Test 1. First, we test the reconstruction by our method of a single ball shaped inclusion depicted in Figure 8.2(a). Here, $c = 2$ inside of this inclusion and $c = 1$ outside. Hence, the inclusion/background contrast is 2:1. We show the 3D image and slices for $N = 1, 3$; see Figures 8.2.

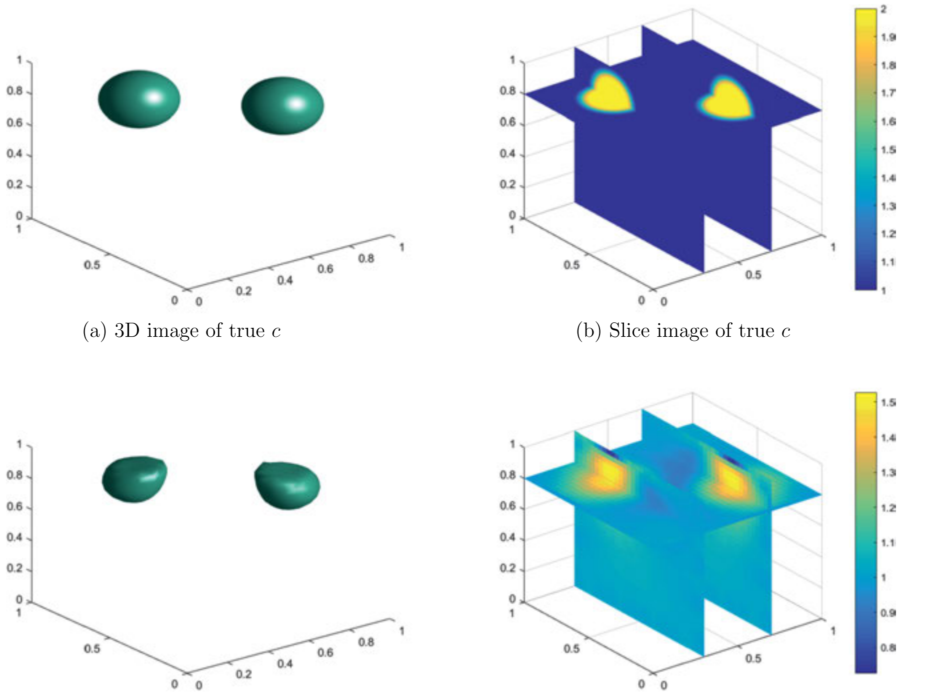
Test 2. Second, we test the reconstruction by our method of a single elliptically shaped inclusion depicted on Figure 8.3(a). Here, $c = 2$ inside of this inclusion and $c = 1$ outside. Hence, the inclusion/background contrast is 2:1. We show the 3D image and slices for $N = 3$; see Figures 8.3.



(a) 3D image of true c (b) Slice image of true c

(c) $N = 3$ (d) $N = 3$

Figure 8.3: Results of Test 2. Imaging of one elliptically shaped inclusion with $c = 2$ in it and $c = 1$ outside. Hence, the inclusion/background contrast is 2:1. We have stopped at the 3rd mesh refinement for all three values of N. (a) and (b) Correct images. (c) and (d) Computed images for $N = 3$. The maximal value of the computed coefficient $c(\mathbf{x})$ is approximately 1.6.



(a) 3D image of true c (b) Slice image of true c

(c) $N = 3$ (d) $N = 3$

Figure 8.4: Results of Test 3. Imaging of two ball-shaped inclusions with $c = 2$ in each of them and $c = 1$ outside. We have stopped on the 3rd mesh refinement for all three values of N. (a) and (b) Correct images. (c) and (d) Computed images for $N = 3$. In each imaged inclusion, the maximal value of the computed coefficient $c(\mathbf{x})$ is approximately 1.9.

Test 3. We now test the performance of our method for imaging of two ball-shaped inclusions depicted on Figure 8.4(a). Here, $c = 2$ inside of each inclusion and $c = 1$ outside of these inclusions. Figures 8.4 display results.

Test 4. We now test our method for the case when the function $c(\mathbf{x})$ is smoothly varying within an abnormality and with a wide range of variations approximately between 0.6 and 1.7. The results are shown in Figure 8.5. Thus, our method can accurately image not only "sharp" inclusions as in Tests 1–3, but abnormalities with smoothly varying functions $c(\mathbf{x})$ in them as well.

Test 5. In this example, we test the reconstruction by our method of a single ball-shaped inclusion with a high inclusion/background contrast; see Figure 8.6(a). Here, $c = 5$ inside of this inclusion and $c = 1$ outside. Hence, the inclusion/background contrast is 5:1. See Figures 8.6 for results.

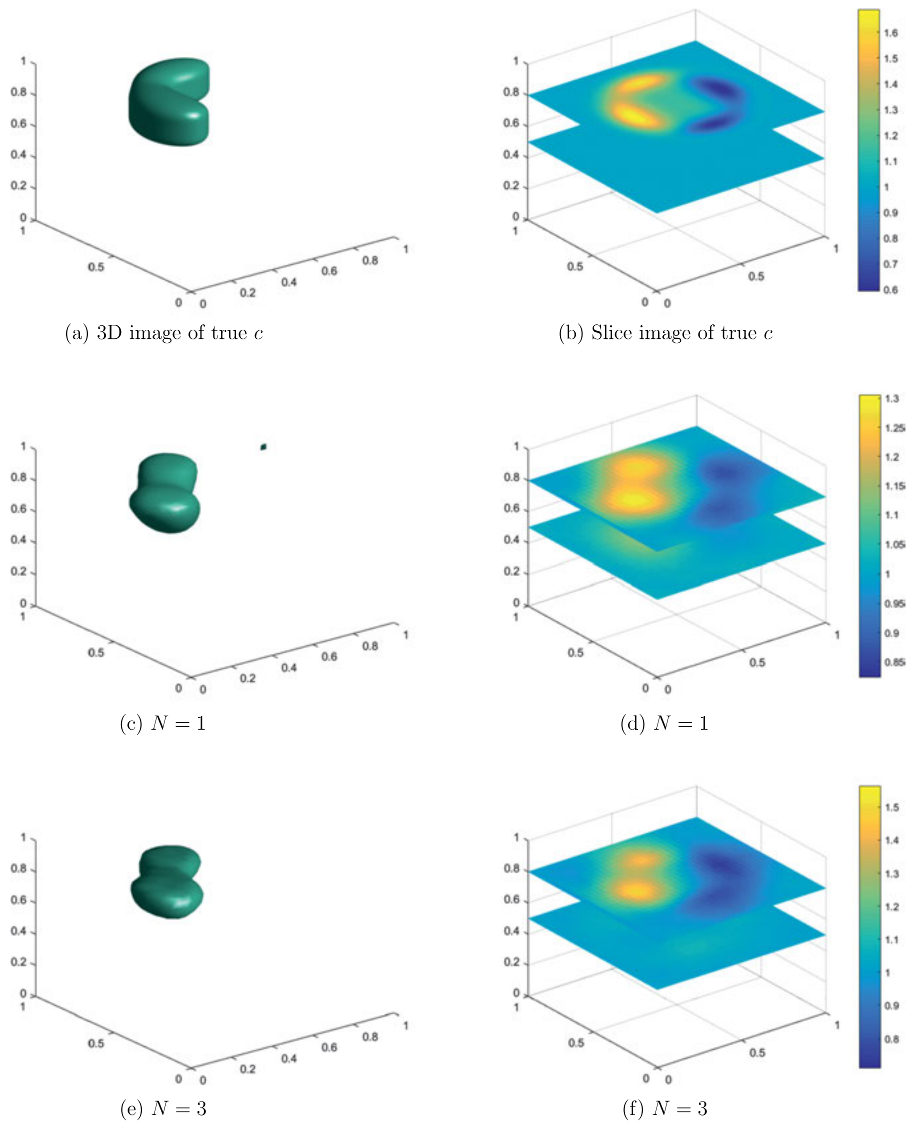
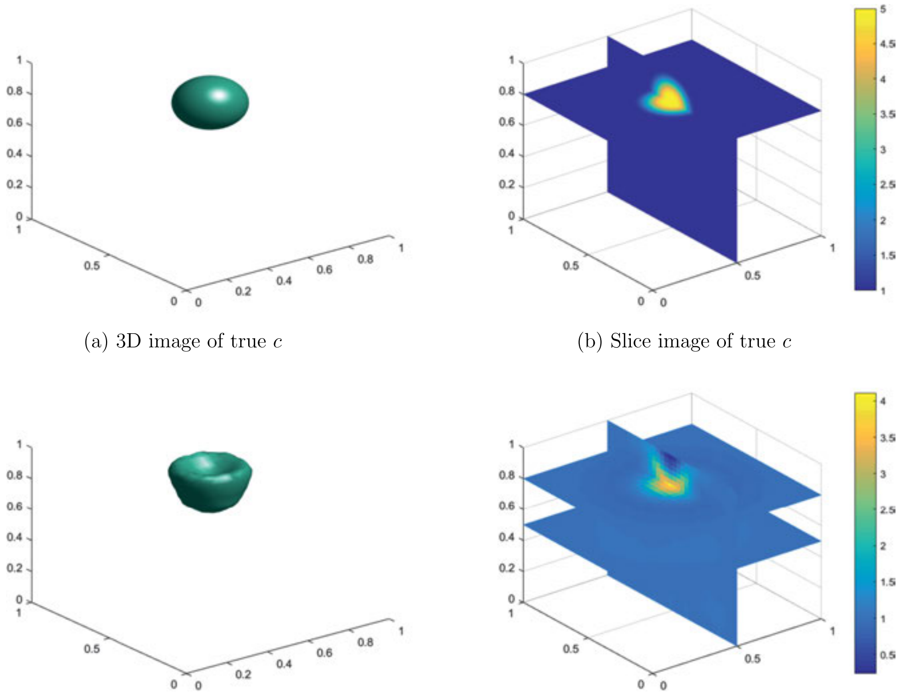


Figure 8.5: Results of Test 4. Imaging of a smoothly varying coefficient. The function $c(\mathbf{x})$ in the inclusion varies between 0.6 and 1.7. (a) and (b) Correct images. (c) and (d) Computed images for $N = 1$. (e) and (f) Computed images for $N = 3$. The computed function $c(\mathbf{x})$ in the inclusion varies approximately between 0.7 and 1.6.



(a) 3D image of true c (b) Slice image of true c

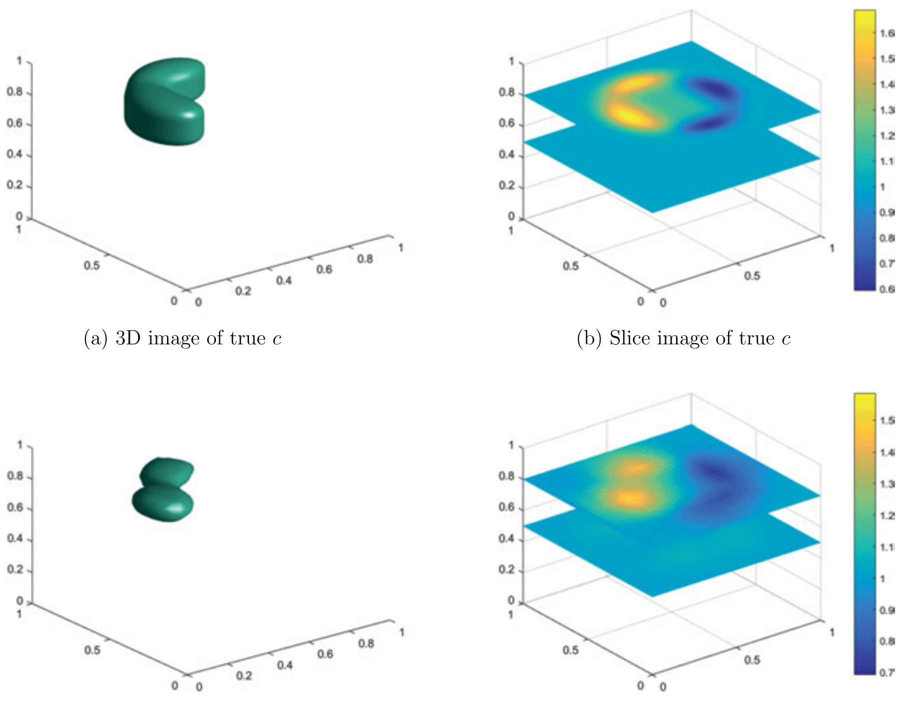
(e) $N = 3$ (f) $N = 3$

Figure 8.6: Results of Test 5. Imaging of one ball-shaped inclusion with $c = 5$ in it and $c = 1$ outside. Hence, the inclusion/background contrast is 5:1. We have stopped at the 3rd mesh refinement. (a) and (b) Correct images. (c) and (d) Computed images for $N = 3$. The maximal value of the computed coefficient $c(\mathbf{x})$ is approximately 4.

Test 6. In this example, we test the stability of our algorithm with respect to the random noise in the data. We test the stability for the case of the function $c(\mathbf{x})$ described in test 4. The noise is added for $\mathbf{x} \in \Gamma_0$ (see (8.96)) as

$$g_{0,\text{noise}}(\mathbf{x}, t) = g_0(\mathbf{x}, t)(1 + \epsilon\xi_t) \quad \text{and} \quad g_{1,\text{noise}}(\mathbf{x}, t) = g_1(\mathbf{x}, t)(1 + \epsilon\xi_t), \tag{8.101}$$

where functions $g_0(\mathbf{x}, t)$, $g_1(\mathbf{x}, t)$ are defined in (8.14), ϵ is the noise level, and ξ_t is a random variable depending only on the time t and uniformly distributed on $[-1, 1]$. We took $\epsilon = 5\,\%$ which is $5\,\%$ noise, see Figure 8.7.



(a) 3D image of true c (b) Slice image of true c

(c) $N = 3$ (d) $N = 3$

Figure 8.7: Results of Test 6. We test the reconstruction of the same function $c(\mathbf{x})$ as in Test 4 (Figures 8.6) but with the noise in the data. The level of noise in (6.5) is $\epsilon = 5\%$. We have stopped at the 3rd mesh refinement for $N = 3$. (a) and (b) Correct images. (c) and (d) Computed images for $N = 3$. The computed function $c(\mathbf{x})$ in the inclusion varies between 0.7 and 1.6.

9 Convexification for an inverse parabolic problem

In this chapter, we follow our publication [151]. Permission for republication is obtained from the publisher.

9.1 Introduction

In this chapter, we construct the convexification globally convergent numerical method for a Coefficient Inverse Problem (CIP) for a parabolic PDE. This CIP has applications in heat conduction [4, 5] and in medical optical imaging using the diffuse infrared light [70]. The first step toward the goal of this chapter was made in [141] for a similar CIP. However, there are some problems in [141], which prevent one from a numerical implementation of the idea of [141]. Indeed, although a weighted globally strictly convex Tikhonov-like functional is constructed in [141], the CWF in it is the same as the one in Section 2.3. However, this CIP is too complicated since it depends on two large parameters rather than on a single one. This means that the CWF of [141] changes too rapidly. The latter does not allow a numerical implementation. In addition, since [141] was published before [11], then uniqueness and existence of the minimizer as well as the global convergence of the gradient projection method to the correct solution of that CIP are not proven in [141]. Besides, numerical studies were not conducted in [141].

Thus, in this chapter we establish a new Carleman estimate with a simpler CWF, which can be used for computations. Our *central* result is the global strict convexity of our weighted Tikhonov-like functional. Next, we establish the existence and uniqueness of its minimizer, estimate the distance between that minimizer and the exact solution and prove the global convergence of the gradient projection method to the exact solution. Finally, we present results of our numerical experiments.

9.2 Statement of the coefficient inverse problem

Below $\mathbf{x} = (x, \overline{x}) \in \mathbb{R}^n$, where $\overline{x} = (x_2, \dots, x_n)$ and $x = x_1$. Let the numbers $A, B > 0$, and $A < B$. We introduce the cube $\Omega \subset \mathbb{R}^n$ and a part Γ of its boundary $\partial\Omega$ as

$$\Omega = \{\mathbf{x} : A < x, x_2, \dots, x_n < B\}, \quad \Gamma = \{x = B, A < x_2, \dots, x_n < B\}. \tag{9.1}$$

Let the number $T > 0$. Denote

$$Q_T^{\pm} = \Omega \times (-T, T), \quad S_T^{\pm} = \partial\Omega \times (-T, T), \quad \Gamma_T^{\pm} = \Gamma \times (-T, T).$$

Below $\alpha \in (0, 1)$, $m \geq 1$ is an integer and $C^{m+\alpha}(\overline{\Omega})$, $C^{2m+\alpha, m+\alpha/2}(\overline{Q_T^{\pm}})$ are Hölder spaces [173]. Let

$$b_j(\mathbf{x}), c(\mathbf{x}) \in C^{2+\alpha}(\overline{\Omega}); \quad j = 1, \dots, n.$$

https://doi.org/10.1515/9783110745481-009

We consider the elliptic operator L in the following form:

$$Lu = \Delta u + \sum_{j=1}^{n} b_j(\mathbf{x})u_{x_j} - c(\mathbf{x})u, \quad \mathbf{x} \in \Omega. \tag{9.2}$$

We assume that

$$c(\mathbf{x}) \geq 0 \quad \text{in } \overline{\Omega}. \tag{9.3}$$

The forward parabolic initial boundary value problem is stated as [173, 174].

Forward problem. Let the initial condition $f(\mathbf{x}) \in C^{4+\alpha}(\overline{\Omega})$. Find a function $u(\mathbf{x}, t) \in C^{4+\alpha,2+\alpha/2}(\overline{Q_T^{\pm}})$ satisfying the following conditions:

$$u_t = Lu \quad \text{in } Q_T^{\pm}, \tag{9.4}$$

$$u(\mathbf{x}, -T) = f(\mathbf{x}), \tag{9.5}$$

$$u|_{S_T^{\pm}} = g_0(\mathbf{x}, t). \tag{9.6}$$

If $n = 3$ and functions $b_j(\mathbf{x}) \equiv 0$ for $j = 1, \ldots, n$, then $c(\mathbf{x})$ is the absorption coefficient in the case of medical optical imaging using the diffuse infrared light [70].

If the domain Ω would have its boundary $\partial\Omega \in C^{4+\alpha}$ and if the Dirichlet condition $g_0(\mathbf{x}, t)$ would belong to $C^{4+\alpha,2+\alpha/2}(\overline{S_T^{\pm}})$ and also corresponding compatibility conditions would be satisfied [174], then the existence and uniqueness of the solution $u \in C^{4+\alpha,2+\alpha/2}(\overline{Q_T^{\pm}})$ of problems (9.2)–(9.6) would be ensured [174]. However, for the convenience of our derivations for the inverse problem, we have chosen the case of a piecewise smooth boundary $\partial\Omega$. Hence, we can only assume the existence of the solution $u \in C^{4+\alpha,2+\alpha/2}(\overline{Q_T^{\pm}})$ of problems (9.4)–(9.6). As to its uniqueness, it follows immediately from (9.3) and the maximum principle for parabolic PDEs.

Coefficient Inverse Problem (CIP). Let the number $t_0 \in (-T, T)$. Suppose that the function $g_0(\mathbf{x}, t)$ in (9.6) is known. Also, assume that the function $f(\mathbf{x})$ in (9.5) as well as the coefficient $c(\mathbf{x})$ are unknown. In addition, let the following two functions $g_1(\mathbf{x}, t)$ and $f_0(\mathbf{x})$ be known:

$$u_x|_{\Gamma_T^{\pm}} = g_1(\mathbf{x}, t), \tag{9.7}$$

$$u(\mathbf{x}, t_0) = f_0(\mathbf{x}). \tag{9.8}$$

Find the unknown coefficient $c(\mathbf{x})$.

Remarks 9.2.1.

1. Thus, the data for this CIP is the single pair of functions (g_1, f_0). These data are nonredundant, so as for all CIPs for which the convexification method works. In other words, the number m of free variables in the data equals the number n of free variables in the unknown coefficient, $m = n$. As to the globally convergent numerical methods for CIPs with redundant data with $m > n$, see, for example, [27, 71, 109–112].

2. Assuming that $f_0(\mathbf{x}) \neq 0$ in $\overline{\Omega}$, uniqueness of this CIP follows immediately from Theorem 3.4.2 even for the case if both functions $g_0(\mathbf{x}, t)$ and $g_1(\mathbf{x}, t)$ are given for $(\mathbf{x}, t) \in \Gamma' \times (-T', T')$, where $\Gamma' \subset \partial\Omega$ is an arbitrary part of $\partial\Omega$ and $T' \in (0, T]$ is an arbitrary number. Prior the publication of this book uniqueness of this and similar CIPs for any value of $T > 0$ was proven by the BK method [51] in, for example, Theorem 1.10.7 in [22], Theorem 2 in [122], Theorem 3.10 in [126], and Theorem 3.4 in [132]. We also refer to [96, 252] for the Lipschitz stability estimates for this CIP.
3. As soon as the coefficient $c(\mathbf{x})$ is found via the solution of the above CIP, one can uniquely determine the function $f(\mathbf{x})$ in (9.5) using functions g_0 and f_0 [184]: the knowledge of g_1 is not necessary then, as long as $c(\mathbf{x})$ is known. This is the so-called "parabolic problem with the reversed time." Since this problem is outside of the scope of the current publication, we now only provide some short comments. We refer to [253] for an early publication where both the coefficient $c(\mathbf{x})$ and the initial condition $f(\mathbf{x})$ where simultaneously reconstructed numerically for a similar CIP. As to the reconstruction of the function $f(\mathbf{x})$, results of [130] indicate that even if both Dirichlet and Neumann boundary conditions are given at the entire lateral boundary $S_T^{\pm}$, in addition to the function $f_0(\mathbf{x})$ in (9.8), only a logarithmic stability estimate can be obtained for the problem of finding $f(\mathbf{x})$. This means that one cannot anticipate a good stability of the latter problem. The same conclusion was drawn in [253]. A recent convergent numerical method for the parabolic problem with the reversed time can be found in [166].

9.3 Weighted globally strictly convex Tikhonov-like functional

We assume below that there exists a number $\mu > 0$ such that

$$f(\mathbf{x}) \geq \mu, \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{9.9}$$

$$g_0(\mathbf{x}, t) \geq \mu, \quad \forall (\mathbf{x}, t) \in \overline{S_T^{\pm}}. \tag{9.10}$$

Then (9.3), (9.9), (9.10), and the maximum principle for parabolic PDEs [79, 174] imply that

$$u(\mathbf{x}, t) \geq \mu \quad \text{in } \overline{Q_T^{\pm}}. \tag{9.11}$$

9.3.1 Nonlinear integral differential equation

Using (9.11), we introduce a new function $v(\mathbf{x}, t)$,

$$v(\mathbf{x}, t) = \ln u(\mathbf{x}, t) \to u = e^v. \tag{9.12}$$

Substituting (9.12) in (9.4)–(9.8), we obtain in $Q_T^{\pm}$:

$$v_t - \Delta v - (\nabla v)^2 - \sum_{k=1}^{n} b_j(\mathbf{x}) v_{x_j} = c(\mathbf{x}), \tag{9.13}$$

$$v|_{S_T^{\pm}} = \ln g_0(\mathbf{x}, t), \quad v_x|_{\Gamma_T^{\pm}} = (g_1/g_0)(\overline{x}, t), \tag{9.14}$$

$$v(\mathbf{x}, t_0) = \ln f_0(\mathbf{x}) := \widetilde{f}_0(\mathbf{x}). \tag{9.15}$$

For brevity, we set below $t_0 := 0$. The case $t_0 \neq 0$ can be considered along the same lines. Differentiate both sides of the nonlinear equation (9.13) with respect to t and denote $w(\mathbf{x}, t) = v_t(\mathbf{x}, t)$. Since the function $c(\mathbf{x})$ is independent on t, then the right-hand side of the resulting equation will be zero. By (9.15),

$$v(\mathbf{x}, t) = \int_0^t w(\mathbf{x}, \tau) d\tau + \widetilde{f}_0(\mathbf{x}), \quad (\mathbf{x}, t) \in Q_T^{\pm}. \tag{9.16}$$

Thus, (9.14)–(9.16) lead to a nonlinear integral differential PDE with Volterra integrals, supplied by the lateral Cauchy data,

$$\begin{aligned} N(w) = w_t &- Mw \\ &- 2\nabla w \int_0^t \nabla w(\mathbf{x}, \tau) d\tau = 0, \quad (\mathbf{x}, t) \in Q_T^{\pm}, \end{aligned} \tag{9.17}$$

$$w|_{S_T^{\pm}} = p_0(\mathbf{x}, t), \quad w_x|_{\Gamma_T^{\pm}} = p_1(\mathbf{x}, t), \tag{9.18}$$

where $p_0(\mathbf{x}, t) = (g_{0t}/g_0)(\mathbf{x}, t)$ and $p_1(\mathbf{x}, t) = \partial_t (g_1/g_0)(\overline{x}, t)$, where

$$Mw = Lw + 2\nabla w \nabla \widetilde{f}_0 = \Delta w + \sum_{j=1}^{n} b_j(\mathbf{x}) w_{x_j} + 2\nabla w \nabla \widetilde{f}_0. \tag{9.19}$$

9.3.2 The functional

First, we choose such a CWF which would work well computationally. The CWF of Section 2.3 changes too rapidly since it depends on two large parameters λ and ν. This is inconvenient for a numerical implementation. Thus, we choose the CWF $\varphi_\lambda(x, t)$ as

$$\varphi_\lambda(x, t) = \exp(2\lambda(x^2 - t^2)), \tag{9.20}$$

where $\lambda \geq 1$ is a parameter. This means that we need to prove the Carleman estimate with this CWF; see Theorem 9.4.1 in Section 9.4.

Since the function $\varphi_\lambda(x, t)$ depends only on one component x of the n-D vector $x = (x, \overline{x})$, then it follows from (9.26) and (9.29) that this Carleman estimate is valid only because the Dirichlet boundary condition $g_0(\mathbf{x},t)$ in (9.6) is given on the entire lateral boundary $S_T^{\pm}$ of the cylinder $Q_T^{\pm}$, whereas the Neumann boundary condition $g_1(\mathbf{x},t)$ in (9.7) is given only on $\Gamma_T^{\pm} \subset S_T^{\pm}$.

Let $[(n+1)/2]$ be the maximal integer, which does not exceed $(n+1)/2$. Denote $k_n = [(n+1)/2] + 2$. For example, we have for the most popular cases of $n = 1, 2, 3$:

$$k_n = \begin{cases} 3 & \text{if } n = 1, 2, \\ 4 & \text{if } n = 3. \end{cases}$$

We have chosen the number k_n in such a way that

$$H^{k_n}(Q_T^{\pm}) \subseteq H^3(Q_T^{\pm}), \tag{9.21}$$

$$H^{k_n}(Q_T^{\pm}) \subset C^1(\overline{Q_T^{\pm}}), \quad \|q\|_{C^1(\overline{Q_T^{\pm}})} \le C_0 \|q\|_{H^{k_n}(Q_T^{\pm})}, \quad \forall q \in H^{k_n}(Q_T^{\pm}), \tag{9.22}$$

where the number $C_0 = C_0(Q_T^{\pm}) > 0$ depends only on the domain $Q_T^{\pm}$. Relations (9.22) follow from (9.21) and the embedding theorem.

Let $R > 0$ be an arbitrary number. We define the bounded set of functions $B(R, p_0, p_1)$ as follows:

$$\begin{aligned} &B(R, p_0, p_1) \\ &\quad = \{w \in H^{k_n}(Q_T^{\pm}) : \|w\|_{H^{k_n}(Q_T^{\pm})} < R, w|_{S_T^{\pm}} = p_0, w_x|_{\Gamma_T^{\pm}} = p_1\}, \end{aligned} \tag{9.23}$$

see (9.18) for functions p_0, p_1.

Let $\beta > 0$ be a small regularization parameter and $N(w)$ be the nonlinear integral differential operator defined in (9.17). We construct our weighted Tikhonov-like functional with the CWF (9.20) in it as

$$J_{\lambda,\beta}(w) = e^{-2\lambda B^2} \int_{Q_T^{\pm}} (N(w))^2 \varphi_\lambda d\mathbf{x}dt + \beta \|w\|^2_{H^{k_n}(Q_T^{\pm})}. \tag{9.24}$$

Since $\max_{\overline{Q_T^{\pm}}} \varphi_\lambda = e^{2\lambda B^2}$, then the multiplier $e^{-2\lambda B^2}$ is introduced in (9.24) to balance two terms in the right hand side of (9.24).

Minimization problem. Minimize the functional $J_{\lambda,\beta}(w)$ on the set $\overline{B(R)}$ defined in (9.23).

Assume for a moment that a minimizer $w_{\min,\lambda,\beta}(\mathbf{x},t)$ of functional (9.24) exists and is computed. Then we substitute $w_{\min,\lambda,\beta}(\mathbf{x},t)$ in the integral of (9.16). Let the function $v_{\text{comp}}(\mathbf{x},t)$ be the resulting left-hand side of (9.16). Next, substituting $v_{\text{comp}}(\mathbf{x},t)$ in equation (9.13), we calculate an approximation for the target unknown coefficient $c(\mathbf{x})$. However, due to the inevitable computational errors as well as the noise in the data, the resulting left-hand side of (9.13) would depend on t. Hence, to calculate an approximation $c_{\mathrm{comp}}(\mathbf{x})$ for $c(\mathbf{x})$, we set

$$c_{\mathrm{comp}}(\mathbf{x}) = \frac{1}{2\gamma T}\int_{-\gamma T}^{\gamma T}\left(\partial_t v_{\mathrm{comp}} - \Delta v_{\mathrm{comp}} - (\nabla v_{\mathrm{comp}})^2 - \sum_{k=1}^{n} b_j(\mathbf{x})\partial_{x_j} v_{\mathrm{comp}}\right)dt, \tag{9.25}$$

where the number $\gamma \in (0, 1/\sqrt{3})$ is chosen in the formulation of Theorem 9.4.5 in Section 9.4. Thus, we focus below on the minimization problem.

9.4 Theorems

Introduce the subspaces $H_0^{2,1}(Q_T^{\pm}) \subset H^{2,1}(Q_T^{\pm})$ and $H_0^{k_n}(Q_T^{\pm}) \subset H^{k_n}(Q_T^{\pm})$ as

$$\begin{aligned} H_0^{2,1}(Q_T^{\pm}) &= \{u \in H^{2,1}(Q_T^{\pm}) : u|_{S_T^{\pm}} = 0, u_x|_{\Gamma_T^{\pm}} = 0\}, \\ H_0^{k_n}(Q_T^{\pm}) &= \{u \in H^{k_n}(Q_T^{\pm}) : u|_{S_T^{\pm}} = 0, u_x|_{\Gamma_T^{\pm}} = 0\}. \end{aligned} \tag{9.26}$$

Since it is well known that any Carleman estimate depends only on the principal part of the operator (see, e. g., Lemma 2.1.1), then we consider in Theorem 9.4.1 only the principal part $\partial_t - \Delta$ of the parabolic operator $\partial_t - L$. We prove Theorem 9.4.1 in Section 9.9.

Theorem 9.4.1 (Carleman estimate). *Suppose that the domain Ω and the CWF $\varphi_\lambda(x,t)$ are the same as in* (9.1) *and* (9.20), *respectively. Then there exist numbers λ_0, C,*

$$\lambda_0 = \lambda_0(\Omega) \geq 1, \quad C = C(\Omega, T) > 0 \tag{9.27}$$

depending only on listed parameters such that the following Carleman estimate holds:

$$\begin{aligned} \int_{Q_T^{\pm}} (u_t - \Delta u)^2 \varphi_\lambda d\mathbf{x}dt \geq{}& \frac{C}{\lambda}\int_{Q_T^{\pm}}\left(u_t^2 + \sum_{i,j=1}^{n} u_{x_i x_j}^2\right)\varphi_\lambda d\mathbf{x}dt \\ &+ C\lambda \int_{Q_T^{\pm}} [(\nabla u)^2 + \lambda^2 u^2]\varphi_\lambda d\mathbf{x}dt \qquad (9.28)\\ &- C\exp(2\lambda(B^2 - T^2)) \int_{\Omega} \lambda^2 u^2(\mathbf{x}, T)d\mathbf{x} \\ &- C\exp(2\lambda(B^2 - T^2))((\nabla u)^2 + \lambda^2 u^2)(\mathbf{x}, -T), \\ &\forall \lambda \geq \lambda_0, \forall u \in H_0^{2,1}(Q_T^{\pm}). \qquad (9.29) \end{aligned}$$

Remarks 9.4.1.
1. Since the normal derivative of the function $u \in H_0^{2,1}(Q_T^{\pm})$ equals zero only on the part $\Gamma_T^{\pm}$ of the lateral boundary $S_T^{\pm}$ of the time cylinder $Q_T^{\pm}$ rather than on the whole $S_T^{\pm}$, then one should carefully analyze integrals over $S_T^{\pm}$, which occur in the pointwise Carleman estimate: to make sure that these integrals equal zero.
2. We assume in all theorems below that the set $B(R, p_0, p_1) \neq \varnothing$ and that it has infinitely many elements. A simple sufficient condition of this is presented in this section below.

For brevity, let K denote below the following vector of numbers:

$$K = \left(\max_j \|b_j\|_{C(\overline{\Omega})}, \mu\right). \tag{9.30}$$

Theorem 9.4.2 (the central theorem of this chapter). *Let $\mu, R > 0$ be two numbers. Assume that condition* (9.11) *holds. Then the functional $J_{\lambda,\beta}(w)$ has the Fréchet derivative $J'_{\lambda,\beta}(w) \in H_0^{k_n}(Q_T^{\pm})$ for all $\lambda, \beta > 0$ and for all $w \in H^{k_n}(Q_T^{\pm})$ satisfying boundary conditions* (9.18). *Let $\lambda_0 \geq 1$ be the constant of Theorem* 9.4.1. *Then there exist constants*

$$\lambda_1 = \lambda_1(R, T, \Omega, K) \geq \lambda_0, \tag{9.31}$$

$$C_1 = C_1(R, T, \Omega, K) > 0 \tag{9.32}$$

depending only on listed parameters such that if $\lambda \geq \lambda_1$ and the regularization parameter $\beta \in [2e^{-\lambda T^2}, 1)$, then the functional $J_{\lambda,\beta}(w)$ is strictly convex on the set $\overline{B(R, p_0, p_1)}$ for all $\lambda \geq \lambda_1$, i. e. for all $w_1, w_2 \in \overline{B(R, p_0, p_1)}$ and for all $\lambda \geq \lambda_1$ the following estimate is valid:

$$\begin{aligned} &J_{\lambda,\beta}(w_2) - J_{\lambda,\beta}(w_1) - J'_{\lambda,\beta}(w_1)(w_2 - w_1) \\ &\quad \geq \frac{C_1}{\lambda} \exp(-2\lambda(T^2 + B^2 - A^2)) \|w_2 - w_1\|^2_{H^{2,1}(Q_T^{\pm})} + \frac{\beta}{2} \|w_2 - w_1\|^2_{H^{k_n}(Q_T^{\pm})}. \end{aligned} \tag{9.33}$$

Everywhere below $C > 0$ and $C_1 > 0$ denote different constants depending only on parameters listed in (9.27) and (9.32), respectively; also see (9.30) for K.

Theorem 9.4.3. *Let $\mu, R > 0$ be two numbers. Assume that condition* (9.11) *holds. Let parameters $\lambda_1, \lambda \geq \lambda_1$, and β be the same as the ones in Theorem* 9.4.2. *Then there exists the unique minimizer $w_{\min,\lambda,\beta} \in \overline{B(R, p_0, p_1)}$ of the functional $J_{\lambda,\beta}(w)$ on the set $\overline{B(R, p_0, p_1)}$. Furthermore, the following inequality holds:*

$$J'_{\lambda,\beta}(w_{\min,\lambda,\beta})(w - w_{\min,\lambda,\beta}) \geq 0, \quad \forall w \in \overline{B(R, p_0, p_1)}. \tag{9.34}$$

By one of main concepts of the regularization theory [244], we assume now that there exists an ideal, the so-called "exact" solution $c^*(\mathbf{x}) \in C^{2+\alpha}(\overline{\Omega})$ of the CIP (9.3), (9.4)–(9.8), where the data (9.6)–(9.8) are noiseless. Having the function $c^*(\mathbf{x})$, one can consider the noise-free solution $w^* \in H^{k_n}(Q_T^{\pm})$ of equation (9.17) with the noiseless boundary data p_0^*, p_1^* in (9.18) and the noiseless function $\widetilde{f}_0^*(\mathbf{x})$ in (9.17).

Let $c_{\text{comp}}(\mathbf{x})$ be the coefficient $c(\mathbf{x})$ reconstructed from the minimizer $w_{\min,\lambda,\beta}(\mathbf{x},t)$ via backwards calculations, as outlined in the last paragraph of Section 9.3 and, in particular, in (9.25). We now want to estimate how the distances between the minimizer $w_{\min,\lambda,\beta}$ and the function w^* as well as between coefficients $c^*(\mathbf{x})$ and $c_{\text{comp}}(\mathbf{x})$ depends on the level of noise $\delta \in (0, \min(1, \mu/2, R))$ in the data. Naturally, we assume that the number δ is sufficiently small.

The regularization theory [22, 76, 244] says that one can assume that the exact solution of an ill-posed problem belongs to a given bounded set in an appropriate Banach space. Hence, we assume that

$$w^* \in B(R-\delta, p_0^*, p_1^*). \tag{9.35}$$

By (9.8) $w^*(\mathbf{x}, t_0) = f_0^*(\mathbf{x})$. Hence, (9.9), (9.22), and (9.35) imply that

$$\|f_0^*\|_{C^1(\overline{\Omega})} < C_0 R, \quad \min_{\overline{\Omega}} f_0^* \geq \mu > 0, \tag{9.36}$$

where the number μ is the same as in (9.9), (9.11) and is independent on δ.

To obtain the desired estimate of the distance between $w_{\min,\lambda,\beta}$ and w^*, we arrange zero boundary conditions in an analog of (9.18). We assume that there exists an extension $G \in H^{k_n}(Q_T^{\pm})$ inside of the cylinder $Q_T^{\pm}$ the boundary conditions (9.18),

$$G|_{S_T^{\pm}} = p_0(\mathbf{x}, t), \quad G_x|_{\Gamma_T^{\pm}} = p_1(\mathbf{x}, t). \tag{9.37}$$

Since

$$w^*|_{S_T^{\pm}} = p_0^*(\mathbf{x}, t), \quad w_x^*|_{\Gamma_T^{\pm}} = p_1^*(\mathbf{x}, t), \tag{9.38}$$

then there also exists a function $G^* \in H^{k_n}(Q_T^{\pm})$ satisfying the latter boundary conditions. Since $R > 0$ is an arbitrary number, then taking into account (9.35), it is reasonable to assume that

$$\|G\|_{H^{k_n}(Q_T^{\pm})} < R, \quad \|G^*\|_{H^{k_n}(Q_T^{\pm})} < R. \tag{9.39}$$

Thus, (9.37)–(9.39) lead to

$$G \in B(R, p_0, p_1), \quad G^* \in B(R, p_0^*, p_1^*). \tag{9.40}$$

It follows from (9.40) that assumptions (9.35), (9.37)–(9.39), which are quite natural ones, imply that $B(R, p_0, p_1) \neq \varnothing$ and $B(R, p_0^*, p_1^*) \neq \varnothing$.

It follows from (9.37) and (9.38) that functions $G(\mathbf{x}, t)$ and $G^*(\mathbf{x}, t)$ can be treated as the first part of noisy and noiseless data respectively for our CIP. The second part of such data are functions f_0 and f_0^*. Thus, we assume below the presence of the noise of the level δ in functions G and f_0,

$$\|G - G^*\|_{H^{k_n}(Q_T^{\pm})} < \delta, \tag{9.41}$$

$$\|f_0 - f_0^*\|_{C^1(\overline{\Omega})} < \delta. \tag{9.42}$$

In particular, since $\delta < \min(\mu/2, R)$, then (9.36) and (9.42) imply that

$$\min_{\overline{\Omega}} f_0(\mathbf{x}) \geq \frac{\mu}{2}, \quad \|f_0\|_{C^1(\overline{\Omega})} < C_0 R. \tag{9.43}$$

We now show that if the above function $G \in B(R, p_0, p_1)$ exists, then the set $B(R, p_0, p_1)$ contains infinitely many functions. Indeed, let a and b be two numbers such that $0 < a < b$ and the set

$$\{\mathbf{x} \in \Omega : \operatorname{dist}(\mathbf{x}, \partial\Omega) > b\} \neq \varnothing,$$

where $\operatorname{dist}(\mathbf{x}, \partial\Omega)$ is the Hausdorff distance between the point x and the boundary $\partial\Omega$ of the domain Ω. Consider an arbitrary number $\xi \in ((a+b)/2, b)$. By (9.1), the domain Ω is a rectangular prism. Hence, it is well known from the Real Analysis course that there exists such a function $\chi_{(a,\xi)}(\mathbf{x}) \in C^{\infty}(\overline{\Omega})$ that

$$\chi_{(a,\xi)}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega \text{ and } \operatorname{dist}(\mathbf{x}, \partial\Omega) < a, \\ 0 & \text{if } \mathbf{x} \in \Omega \text{ and } \operatorname{dist}(\mathbf{x}, \partial\Omega) \in (\xi, b), \\ \text{between 0 and 1} & \text{for all other points } \mathbf{x} \in \Omega. \end{cases}$$

Then the function $G_{(a,\xi)}(\mathbf{x}, t) = \chi_{(a,\xi)}(\mathbf{x}) G(\mathbf{x}, t) \in H^{k_n}(Q_T^{\pm})$ and satisfies boundary conditions (9.37). Varying the parameter ξ between $(a+b)/2$ and b, we obtain infinitely many such functions $G_{(a,\xi)}(\mathbf{x}, t)$. Since $R > 0$ is an arbitrary number, then we can ensure that all these functions belong to the set $B(R, p_0, p_1)$. We took $\xi \in ((a+b)/2, b)$ rather than $\xi \in (a, b)$ in order to make sure that the function $\chi_{(a,\xi)}(\mathbf{x})$ does not change too rapidly for those values of ξ which are close to a. Indeed, such a rapid change would increase the number R.

Denote

$$W = w - G, \quad W^* = w^* - G^*. \tag{9.44}$$

Similarly, with (9.23) denote

$$B_0(2R) = \{W \in H^{k_n}(Q_T^{\pm}) : \|W\|_{H^{k_n}(Q_T^{\pm})} < 2R, W|_{S_T^{\pm}} = W_x|_{\Gamma_T^{\pm}} = 0\}. \tag{9.45}$$

Then (9.35)–(9.44) imply that

$$W \in B_0(2R), \quad \forall w \in B(R, p_0, p_1) \quad \text{and also} \quad W^* \in B_0(2R - \delta), \tag{9.46}$$

$$W + G \in B(3R, p_0, p_1), \quad \forall W \in B_0(2R). \tag{9.47}$$

Due to (9.47), it is convenient to denote below $\lambda_1(3R)$, $\lambda(3R)$, which means that the values of the parameters λ_1 and $\lambda \geq \lambda_1$ correspond to $B(3R, p_0, p_1)$ in Theorem 9.4.2 and, in particular, R is replaced with $3R$ in (9.31) and (9.32).

Consider the functional $I_{\lambda,\beta}(W)$,

$$I_{\lambda,\beta}(W) = J_{\lambda,\beta}(W + G) \quad \text{for } W \in B_0(2R). \tag{9.48}$$

Theorem 9.4.4. *Let $\mu, R > 0$ be two numbers. Assume that condition* (9.11) *holds. Let parameters λ_1 and β be the same as in Theorem* 9.4.2, *except that R is replaced with $3R$ in* (9.31). *Then the functional $I_{\lambda,\beta}(W)$ is strictly convex on the ball $\overline{B_0(2R)}$ for all $\lambda \geq \lambda_1(3R, T, \Omega, K)$. Here, $\lambda_1(3R, T, \Omega, K)$ means* (9.31), *where R is replaced with $3R$. In other words, the following analog of* (9.33) *holds for all $\lambda \geq \lambda_1(3R, T, \Omega, K)$ and for all $W_1, W_2 \in \overline{B_0(2R)}$:*

$$\begin{aligned} I_{\lambda,\beta}(W_2) - I_{\lambda,\beta}(W_1) - I'_{\lambda,\beta}(W_1)(W_2 - W_1) \\ \geq \frac{C_1}{\lambda} \exp(-2\lambda(T^2 + B^2 - A^2)) \|W_2 - W_1\|^2_{H^{2,1}(Q_T^\pm)} + \frac{\beta}{2} \|W_2 - W_1\|^2_{H^{k_n}(Q_T^\pm)}, \end{aligned} \tag{9.49}$$

where $I'_{\lambda,\beta}(W) \in H_0^{k_n}(Q_T^\pm)$ is the Fréchet derivative of the functional $I_{\lambda,\beta}(W)$ at the point W, which exists due to Theorem 9.4.2 *and* (9.48). *Furthermore, there exists unique minimizer $W_{\min,\lambda,\beta} \in \overline{B_0(2R)}$ of the functional $I_{\lambda,\beta}(W)$ and the following inequality holds:*

$$I'_{\lambda,\beta}(W_{\min,\lambda,\beta})(W - W_{\min,\lambda,\beta}) \geq 0, \quad \forall W \in \overline{B_0(2R)}. \tag{9.50}$$

Theorem 9.4.5 (accuracy estimates). *Let $\mu, R > 0$ be two numbers. Assume that condition* (9.11) *holds. Suppose that conditions* (9.35), (9.36), (9.40)–(9.42), *and* (9.44) *hold and also let $T > \sqrt{3(B^2 - A^2)}$. Choose a number $\gamma \in (0, 1/\sqrt{3})$ such that $T^2(1-3\gamma^2)-3(B^2-A^2) > 0$. Denote*

$$\eta_1 = \gamma^2 T^2 + B^2 - A^2, \quad \eta_2 = (1-3\gamma^2)T^2 - 3(B^2 - A^2), \quad \rho = \frac{1}{2}\min\left(1, \frac{\eta_2}{3\eta_1}\right) \in \left(0, \frac{1}{2}\right]. \tag{9.51}$$

Let $\lambda_1 = \lambda_1(3R, T, \Omega, K)$ be the number of Theorem 9.4.4. *Choose a number $\lambda_2 = \lambda_2(3R, \gamma T, \Omega, K)$ such that $\lambda_2 \geq \lambda_1$ and*

$$\frac{1}{\lambda} \geq \exp[-\lambda(\gamma^2 T^2 + B^2 - A^2)], \quad \forall \lambda \geq \lambda_2. \tag{9.52}$$

Choose a sufficiently small number $\delta_0 > 0$ satisfying the following inequality:

$$\ln(\delta_0^{-1/(3\eta_1)}) \geq \lambda_2. \tag{9.53}$$

For each $\delta \in (0, \delta_0)$, choose $\lambda = \lambda(\delta)$ as

$$\lambda = \lambda(\delta) = \ln(\delta^{-1/(3\eta_1)}) > \ln(\delta_0^{-1/(3\eta_1)}).$$

Let the regularization parameter $\beta = \beta(\delta, T) = 2e^{-\lambda(\delta)T^2}$ *(see Theorem* 9.4.2*). Let*

$$w_{\min,\lambda(\delta),\beta(\delta,T)} = W_{\min,\lambda(\delta),\beta(\delta,T)} + G. \tag{9.54}$$

(*Theorem* 9.4.4) *and let* $c_{\text{comp}}(\mathbf{x})$ *be the function* $c(\mathbf{x})$ *computed from the function*

$$w_{\min,\lambda(\delta),\beta(\delta,T)}(\mathbf{x}, t)$$

by the procedure described in the last paragraph of Section 9.3*. Then the following accuracy estimates are valid:*

$$\|w^* - w_{\min,\lambda(\delta),\beta(\delta,T)}\|_{H^2(Q^{\pm}_{\gamma T})} \le C_1 \delta^{\rho}, \tag{9.55}$$

$$\|c^* - c_{\min,\lambda(\delta),\beta(\delta,T)}\|_{L_2(\Omega)} \le C_1 \delta^{\rho}. \tag{9.56}$$

We now construct the gradient projection method of the minimization of the functional $I_{\lambda,\beta}(W)$ defined in (9.48) on the set $\overline{B_0(2R)}$ defined in (9.45). Let $P_B : H^{k_n}(Q^{\pm}_T) \to \overline{B_0(2R)}$ be the orthogonal projection operator. Let $W_0 \in B_0(2R)$ be an arbitrary point of the ball $B_0(2R)$. Let the number $\omega \in (0, 1)$. We arrange the gradient projection method of the minimization of the functional $I_{\lambda,\beta}(W)$ as

$$W_n = P_B(W_{n-1} - \omega I'_{\lambda,\beta}(W_{n-1})), \quad n = 1, 2, \ldots. \tag{9.57}$$

Note that since $W_{n-1}, I'_{\lambda,\beta}(W_{n-1}) \in H_0^{k_n}(Q^{\pm}_T)$, then the function $W_{n-1} - \omega I'_{\lambda,\beta}(W_{n-1})$ has zero boundary conditions (9.18). The latter is important in the computational practice.

Theorem 9.4.6. *Let* $\mu, R > 0$ *be two numbers. Assume that condition* (9.11) *holds. Let the parameter* $\lambda_1(3R, T, \Omega, K)$ *be the same as in Theorem* 9.4.2*, except that* R *is replaced with* $3R$ *in* (9.31)*. Let* $\lambda \ge \lambda_1(3R, T, \Omega, K)$*. Let* $\beta \in [2e^{-\lambda T^2}, 1)$ *be the same as in Theorem* 9.4.2*. Then there exists a sufficiently small number* $\omega_0 = \omega_0(\Omega, R, T, \lambda)$ *such that for any* $\omega \in (0, \omega_0)$ *there exists a number* $\theta = \theta(\omega) \in (0, 1)$ *such that the sequence* (9.57) *converges to the unique minimizer* $W_{\min,\lambda(3R),\beta} \in \overline{B_0(2R)}$ (*Theorem* 9.4.4*) in the norm of the space* $H^{k_n}(Q^{\pm}_T)$*. More precisely,*

$$\|W_{\min,\lambda,\beta} - W_n\|_{H^{k_n}(Q^{\pm}_T)} \le \theta^n \|W_{\min,\lambda,\beta} - W_0\|_{H^{k_n}(Q^{\pm}_T)}. \tag{9.58}$$

Theorem 9.4.7 (global convergence to the exact solution of the gradient projection method). *Let the parameters* $\lambda_2(3R, \gamma T, \Omega, K)$ *and* δ_0 *be the same as in Theorem* 9.4.5*. For* $\delta \in (0, \delta_0)$*, let parameters* $\lambda = \lambda(\delta) > \lambda_2$ *and* $\beta = \beta(\delta, T) = 2e^{-\lambda(\delta)T^2}$ *be the same as in that theorem. Suppose that assumptions of Theorem* 9.4.5 *about numbers* T, γ, A, B *also hold. Let* $\rho \in (0, 1/2]$ *be the number defined in* (9.51)*. Let* $w_n = W_n + G$*,* $n = 0, 1, \ldots$*. Let* $c_{n,\text{comp}}(\mathbf{x})$ *be the function* $c(\mathbf{x})$ *obtained from the function* $w_n(\mathbf{x}, t)$ *by the procedure outlined in the end of Section* 9.3*. Then there exists a sufficiently small number* $\omega_1 =$

$\omega_1(\Omega, R, T, \lambda, \gamma) \in (0, \omega_0]$ *such that for any* $\omega \in (0, \omega_1)$ *there exists a number* $\theta = \theta(\omega) \in (0,1)$ *such that the following convergence estimates are valid for* $n = 1, 2, \ldots$:

$$\|w^* - w_n\|_{H^{2,1}(Q_{\gamma T}^{\pm})} \leq C_1\delta^{\rho} + \theta^n \|w_{\min,\lambda(\delta),\beta(\delta,T)} - w_0\|_{H^{k_n}(Q_T^{\pm})}, \tag{9.59}$$

$$\|c^* - c_{n,\mathrm{comp}}\|_{L_2(\Omega)} \leq C_1\delta^{\rho} + \theta^n \|w_{\min,\lambda(\delta),\beta(\delta,T)} - w_0\|_{H^{k_n}(Q_T^{\pm})}. \tag{9.60}$$

Remarks 9.4.2.
1. Since the starting point $W_0 \in B_0(2R)$ of the iterative process (9.57) is an arbitrary point of the ball $B_0(2R)$ and since $R > 0$ is an arbitrary number, then Theorem 9.4.7 ensures the *global convergence* of the gradient projection method (9.57) to the correct solution as long as the noise level δ tends to zero; see Definition 1.4.2.
2. We omit the proof of Theorem 9.4.3 since it is similar with the proof of Lemma 5.2.1. We omit the proof of Theorem 9.4.6 since it is similar with the proof of Theorem 5.2.1.

9.5 Proofs of Theorems 9.4.2 and 9.4.4

Lemma 9.5.1 follows immediately either from Lemma 3.1.1.

Lemma 9.5.1. *The following estimate holds for every function* $q \in L_2(Q_T^{\pm})$ *and for every* $\lambda \geq 1$:

$$\int\limits_{Q_T^{\pm}} \left(\int\limits_0^t q(\mathbf{x}, \tau) d\tau \right)^2 \varphi_\lambda(x,t) d\mathbf{x} dt \leq \frac{1}{4\lambda} \int\limits_{Q_T^{\pm}} q^2(\mathbf{x}, t) \varphi_\lambda(x,t) d\mathbf{x} dt.$$

9.5.1 Proof of Theorem 9.4.2

Let $w_1, w_2 \in \overline{B(R, p_0, p_1)}$ be two arbitrary functions. Denote $h = w_2 - w_1$. Then $w_2 = w_1 + h$ and also

$$h \in \overline{B_0(2R)}. \tag{9.61}$$

First, we evaluate the expression $(N(w_1+h))^2 - (N(w_1))^2$, where the nonlinear operator N is given in (9.17) and (9.19). Using (9.17) and (9.19), we obtain

$$\begin{aligned} &(N(w_1 + h))^2 \\ &\quad = \left[h_t - Mh - 2\nabla h \int\limits_0^t \nabla w_1(\mathbf{x}, \tau) d\tau - 2\nabla w_1 \int\limits_0^t \nabla h(\mathbf{x}, \tau) d\tau - 2\nabla h \int\limits_0^t \nabla h(\mathbf{x}, \tau) d\tau + N(w_1) \right]^2 \\ &\quad = \left(h_t - Mh - 2\nabla h \int\limits_0^t \nabla w_1(\mathbf{x}, \tau) d\tau - 2\nabla w_1 \int\limits_0^t \nabla h(\mathbf{x}, \tau) d\tau - 2\nabla h \int\limits_0^t \nabla h(\mathbf{x}, \tau) d\tau \right)^2 \end{aligned}$$

$$- 4N(w_1)\nabla h \int_0^t \nabla h(\mathbf{x},\tau)d\tau + \left(N(w_1)\right)^2$$

$$+ 2N(w_1)\left(h_t - Mh - 2\nabla h \int_0^t \nabla w_1(\mathbf{x},\tau)d\tau - 2\nabla w_1 \int_0^t \nabla h(\mathbf{x},\tau)d\tau \right).$$

Let Lin(h) be the linear, with respect to h, part of the above expression,

$$\mathrm{Lin}(h) = 2N(w_1)\left(h_t - Mh - 2\nabla h \int_0^t \nabla w_1(\mathbf{x},\tau)d\tau - 2\nabla w_1 \int_0^t \nabla h(\mathbf{x},\tau)d\tau \right). \tag{9.62}$$

Then

$$\begin{aligned}
&\left(N(w_1 + h)\right)^2 - \left(N(w_1)\right)^2 \\
&\quad = \mathrm{Lin}(h) \qquad\qquad (9.63) \\
&\quad + \left(h_t - Mh - 2\nabla h \int_0^t \nabla w_1(\mathbf{x},\tau)d\tau - 2\nabla w_1 \int_0^t \nabla h(\mathbf{x},\tau)d\tau - 2\nabla h \int_0^t \nabla h(\mathbf{x},\tau)d\tau \right)^2 \\
&\quad - 4N(w_1)\nabla h \int_0^t \nabla h(\mathbf{x},\tau)d\tau.
\end{aligned}$$

Using (9.22), (9.23), (9.61), and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&\left(h_t - Mh - 2\nabla h \int_0^t \nabla w_1(\mathbf{x},\tau)d\tau - 2\nabla w_1 \int_0^t \nabla h(\mathbf{x},\tau)d\tau - 2\nabla h \int_0^t \nabla h(\mathbf{x},\tau)d\tau \right)^2 \\
&\qquad - 4N(w_1)\nabla h \int_0^t \nabla h(\mathbf{x},\tau)d\tau \qquad\qquad (9.64) \\
&\qquad \geq \frac{1}{2}(h_t - Mh)^2 - C_1(\nabla h)^2 - C_1\left(\int_0^t \nabla h(\mathbf{x},\tau)d\tau \right)^2.
\end{aligned}$$

For the sake of clarity, we state here that the constant $C_1 > 0$ in (9.64) depends on

$$\sup_{w_1 \in B(R,p_0,p_1)} \left(\sup_{(\mathbf{x},t)\in Q_T^{\pm}} \left|N(w_1)(\mathbf{x},t)\right|, \sup_{(\mathbf{x},t)\in Q_T^{\pm}} \left|\nabla w_1(\mathbf{x},t)\right| \right), \qquad \sup_{\Omega} |\nabla \widetilde{f}_0|.$$

Thus, it follows from (9.15), (9.17), (9.19), (9.30), (9.36), and (9.43) that C_1 in (9.64) depends on R, T, Ω, K, which is exactly as in (9.32).

Using (9.24) and (9.62)–(9.64), we obtain

$$\begin{aligned} &J_{\lambda,\beta}(w_1+h)-J_{\lambda,\beta}(w_1)-e^{-2\lambda B^2}\int\limits_{Q_T^\pm}\mathrm{Lin}(h)\varphi_\lambda d\mathbf{x}dt-2\beta\{w,h\} \\ &\geq \frac{1}{2}e^{-2\lambda B^2}\int\limits_{Q_T^\pm}(h_t-Mh)^2\varphi_\lambda d\mathbf{x}dt \\ &\quad - C_1 e^{-2\lambda B^2}\int\limits_{Q_T^\pm}\left[(\nabla h)^2+\left(\int\limits_0^t\nabla h(\mathbf{x},\tau)d\tau\right)^2\right]\varphi_\lambda d\mathbf{x}dt+\beta\|h\|^2_{H^{k_n}(Q_T^\pm)}. \end{aligned} \tag{9.65}$$

Here and below, $\{\,,\,\}$ is the scalar product in $H^{k_n}(Q_T^\pm)$.

Consider now the functional $S(h): H_0^{k_n}(Q_T^\pm)\to\mathbb{R}$ defined as

$$S(h)=e^{-2\lambda B^2}\int\limits_{Q_T^\pm}\mathrm{Lin}(h)\varphi_\lambda d\mathbf{x}dt+2\beta\{w,h\}. \tag{9.66}$$

It is clear from (9.62) that $S(h)$ is a bounded linear functional. Hence, by the Riesz theorem there exists a function $Z\in H_0^{k_n}(Q_T^\pm)$ such that $S(h)=\{Z,h\}$, $\forall h\in H_0^{k_n}(Q_T^\pm)$. Furthermore, it follows from (9.63) that

$$\lim_{\|h\|_{H^{k_n}(Q_T^\pm)}\to 0}\left|J_{\lambda,\beta}(w_1+h)-J_{\lambda,\beta}(w_1)-S(h)\right|=0.$$

Hence, $S(h)$ is the Fréchet derivative of the functional $J_{\lambda,\beta}(w)$ at the point w_1,

$$S(h)=J'_{\lambda,\beta}(w_1)(h)=\{Z,h\}=\{J'_{\lambda,\beta}(w_1),h\},\quad \forall h\in H_0^{k_n}(Q_T^\pm), \tag{9.67}$$

that is, we can set $Z=J'_{\lambda,\beta}(w_1)$. Thus, we have proven the existence of the Fréchet derivative of the functional $J_{\lambda,\beta}(w)$ for all for all $\lambda,\beta>0$ and for all $w\in B(R,p_0,p_1)$. Obviously, the same proof is valid for all $\lambda,\beta>0$ and for all functions $w\in H^{k_n}(Q_T^\pm)$ satisfying boundary conditions (9.18).

Thus, (9.65)–(9.67) imply that

$$\begin{aligned} &J_{\lambda,\beta}(w_1+h)-J_{\lambda,\beta}(w_1)-J'_{\lambda,\beta}(w_1)(h) \\ &\geq \frac{1}{2}e^{-2\lambda B^2}\int\limits_{Q_T^\pm}(h_t-Mh)^2\varphi_\lambda d\mathbf{x}dt \\ &\quad - C_1 e^{-2\lambda B^2}\int\limits_{Q_T^\pm}\left[(\nabla h)^2+\left(\int\limits_0^t\nabla h(\mathbf{x},\tau)d\tau\right)^2\right]\varphi_\lambda d\mathbf{x}dt+\beta\|h\|^2_{H^{k_n}(Q_T^\pm)}. \end{aligned} \tag{9.68}$$

It follows from (9.19), (9.30), (9.32), and the Cauchy–Schwarz inequality that

$$(h_t-Mh)^2\geq\frac{1}{2}(h_t-\Delta h)^2-C_1(\nabla h)^2. \tag{9.69}$$

Hence, using Theorem 9.4.1, Lemma 9.5.1, (9.68), and (9.69), we obtain for all $\lambda \geq \lambda_0 \geq 1$,

$$\begin{aligned}
&\frac{1}{2}e^{-2\lambda B^2}\int_{Q_T^\pm}(h_t - Mh)^2\varphi_\lambda d\mathbf{x}dt \\
&\quad - C_1 e^{-2\lambda B^2}\int_{Q_T^\pm}\left[(\nabla h)^2 + \left(\int_0^t \nabla h(\mathbf{x},\tau)d\tau\right)^2\right]\varphi_\lambda d\mathbf{x}dt + \beta\|h\|^2_{H^{k_n}(Q_T^\pm)} \\
&\geq \frac{1}{4}e^{-2\lambda B^2}\int_{Q_T^\pm}(h_t - \Delta h)^2\varphi_\lambda d\mathbf{x}dt \\
&\quad - C_1 e^{-2\lambda B^2}\int_{Q_T^\pm}\left[(\nabla h)^2 + \left(\int_0^t \nabla h(\mathbf{x},\tau)d\tau\right)^2\right]\varphi_\lambda d\mathbf{x}dt + \beta\|h\|^2_{H^{k_n}(Q_T^\pm)} \qquad (9.70) \\
&\geq \frac{C}{\lambda}e^{-2\lambda B^2}\int_{Q_T^\pm}\left(h_t^2 + \sum_{i,j=1}^n h_{x_i x_j}^2\right)\varphi_\lambda d\mathbf{x}dt \\
&\quad + C\lambda e^{-2\lambda B^2}\int_{Q_T^\pm}[(\nabla h)^2 + \lambda^2 h^2]\varphi_\lambda d\mathbf{x}dt - C_1 e^{-2\lambda B^2}\int_{Q_T^\pm}(\nabla h)^2\varphi_\lambda d\mathbf{x}dt \\
&\quad - Ce^{-2\lambda T^2}\int_\Omega((\nabla h)^2(\mathbf{x},-T) + \lambda^2 h^2(\mathbf{x},T) + \lambda^2 h^2(\mathbf{x},-T))d\mathbf{x} + \beta\|h\|^2_{H^{k_n}(Q_T^\pm)}.
\end{aligned}$$

Choose $\lambda_1 = \lambda_1(R,T,\Omega,K) \geq \lambda_0$ so large that $C\lambda \geq 2C_1$ as well as $2e^{-\lambda T^2} \geq C\lambda^2 e^{-2\lambda T^2}$, for all $\lambda \geq \lambda_1$. Also, we keep in mind that by the trace theorem

$$\|u(\mathbf{x},\pm T)\|_{H^1(\Omega)} \leq C\|u\|_{H^{k_n}(Q_T^\pm)}, \quad \forall u \in H^{k_n}(Q_T^\pm).$$

Hence, choosing $\beta \in [2e^{-\lambda T^2}, 1)$ and using (9.68) and (9.70), we obtain

$$\begin{aligned}
&J_{\lambda,\beta}(w_1 + h) - J_{\lambda,\beta}(w_1) - J'_{\lambda,\beta}(w_1)(h) \qquad (9.71) \\
&\geq \frac{C_1}{\lambda}e^{-2\lambda B^2}\int_{Q_T^\pm}\left(h_t^2 + \sum_{i,j=1}^n h_{x_i x_j}^2\right)\varphi_\lambda d\mathbf{x}dt + C\lambda e^{-2\lambda B^2}\int_{Q_T^\pm}[(\nabla h)^2 + \lambda^2 h^2]\varphi_\lambda d\mathbf{x}dt \\
&\quad + \frac{\beta}{2}\|h\|^2_{H^{k_n}(Q_T^\pm)}, \quad \forall h \in \overline{B_0(2R)}, \forall \lambda \geq \lambda_1;
\end{aligned}$$

also, see (9.61). Finally, since $\varphi_\lambda(x,t) \geq \exp(-2\lambda(T^2 - A^2))$ for $(x,t) \in (A,B) \times (-T,T)$, then the target estimate (9.33) follows immediately from (9.71).

9.5.2 Proof of Theorem 9.4.4

Since by (9.48) $I_{\lambda,\beta}(W) = J_{\lambda,\beta}(W + G)$, $W \in \overline{B_0(2R)}$ and also since $W + G \in \overline{B(3R,p_0,p_1)}$, $\forall W \in \overline{B_0(2R)}$, then we take in Theorems 9.4.2 and 9.4.3 $\lambda \geq \lambda_1 = \lambda_1(3R,T,\Omega,K)$. This means that we replace in (9.31) R with $3R$. Denote $w_1 = W_1 + G$, $w_2 = W_2 + G$. Then $w_1, w_2 \in \overline{B(3R, p_0, p_1)}$. The rest of the proof follows immediately from Theorems 9.4.2 and 9.4.3.

9.6 Proof of Theorem 9.4.5

As to the parameters λ and β, we only initially assume in this proof that $\lambda \geq \lambda_2(3R, \gamma T, \Omega, K) \geq \lambda_1$ and

$$\beta = 2e^{-\lambda T^2}. \tag{9.72}$$

We will choose the dependencies of parameters λ and β on δ later in this proof. Recall that by (9.46) $w^* - G^* = W^* \in B_0(2R - \delta)$. Hence, by (9.23), (9.41), and (9.44),

$$W^* + G \in B(3R, p_0, p_1). \tag{9.73}$$

Recall that the operator $N(w)$ was defined in (9.17), (9.19). Also, the functional $I_{\lambda,\beta}(W)$ was defined in (9.48). In (9.17), replace w with $W + G$. We temporally denote the resulting expression as $N(W + G, f_0)$. Also, we temporally set that $I_{\lambda,\beta}(W, f_0)$ is the functional $I_{\lambda,\beta}(W)$. We are doing these in order to emphasize the presence of the vector function $\widetilde{f}_0 = \nabla \ln f_0$ in the corresponding formulas. Denote

$$I^0_{\lambda,\beta}(W, f_0) = e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} [N(W + G, f_0)]^2 \varphi_\lambda d\mathbf{x}dt. \tag{9.74}$$

By (9.17) $N(W^* + G^*, f_0^*) = 0$. Hence,

$$I^0_{\lambda,\beta}(W^*, f_0^*) = e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} [N(W^* + G^*, f_0^*)]^2 \varphi_\lambda d\mathbf{x}dt = 0. \tag{9.75}$$

Hence, it follows from (9.17), (9.36), (9.41), (9.42), (9.43), (9.44), (9.74), and (9.75) that

$$\begin{aligned} I^0_{\lambda,\beta}(W^*, f_0) &= e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} \left(N(W^* + G^* + (G - G^*), f_0^* + (f_0 - f_0^*))\right)^2 \varphi_\lambda d\mathbf{x}dt \\ &= e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} \left(N(W^* + G^*, f_0^*)\right)^2 \varphi_\lambda d\mathbf{x}dt + P_{\lambda,\delta} = P_{\lambda,\delta}. \end{aligned} \tag{9.76}$$

We now estimate $|P_{\lambda,\delta}|$ from the above. By (9.17) and (9.19),

$$\begin{aligned} &N(W^* + G^* + (G - G^*), f_0^* + (f_0 - f_0^*))(\mathbf{x}, t) \\ &\quad = N(W^* + G^*, f_0^*)(\mathbf{x}, t) + Q(W^* + G^*, f_0^*, G - G^*, f_0 - f_0^*)(\mathbf{x}, t) \\ &\quad = Q(W^* + G^*, f_0^*, G - G^*, f_0 - f_0^*)(\mathbf{x}, t). \end{aligned} \tag{9.77}$$

An obvious explicit formula can be written for $Q(W^* + G^*, f_0^*, G - G^*, f_0 - f_0^*)(\mathbf{x}, t)$, although we avoid doing so for brevity. Consider, for example, the following term in that formula:

$$Y(\mathbf{x}, t) = -2\left[\nabla(W^* + G^*)(\mathbf{x}, t)\int_0^t \nabla(G - G^*)(\mathbf{x}, \tau)d\tau\right].$$

Since $e^{-2\lambda B^2}\varphi_\lambda(x, t) \le 1$ for $(x, t) \in (A, B) \times (-T, T)$, then it follows from (9.40), (9.41), and (9.46) that

$$e^{-2\lambda B^2}\int_{Q_T^\pm} Y^2\varphi_\lambda d\mathbf{x}dt \le C_1\delta^2.$$

Similarly, using (9.17), (9.19), (9.36), (9.40)–(9.43), (9.46), and the Cauchy–Schwarz inequality, we obtain

$$e^{-2\lambda B^2}\int_{Q_T^\pm} [Q(W^* + G^*, f_0^*, G - G^*, f_0 - f_0^*)]^2\varphi_\lambda d\mathbf{x}dt \le C_1\delta^2.$$

Hence, using (9.76) and (9.77), we obtain

$$|P_{\lambda,\delta}| \le C_1\delta^2. \tag{9.78}$$

Since $I_{\lambda,\beta}(W^*, f_0) = I^0_{\lambda,\beta}(W^*, f_0) + \beta\|W^* + G\|^2_{H^{k_n}(Q_T^\pm)}$, then (9.40), (9.46), (9.76), and (9.78) imply that

$$I_{\lambda,\beta}(W^*, f_0) \le C_1(\delta^2 + \beta). \tag{9.79}$$

By (9.35) and (9.41),

$$\|W^* + G\|_{H^{k_n}(Q_T^\pm)} = \|(W^* + G^*) + (G - G^*)\|_{H^{k_n}(Q_T^\pm)} \le \|w^*\|_{H^{k_n}(Q_T^\pm)} + \delta < R.$$

Hence,

$$W^* + G \in B(R, p_0, p_1). \tag{9.80}$$

Let $W_{\min,\lambda,\beta} \in \overline{B_0(2R)}$ be the minimizer of the functional $I_{\lambda,\beta}(W, f_0)$, the existence and uniqueness of which on the set $\overline{B_0(2R)}$ is guaranteed by Theorem 9.4.3. For brevity, introduce the function $z(\mathbf{x}, t)$,

$$z(\mathbf{x}, t) = W^* - W_{\min,\lambda,\beta}(\mathbf{x}, t). \tag{9.81}$$

It follows from (9.48) that an obvious analog of (9.71) holds for the functional $I_{\lambda,\beta}$. Thus, we use that analog now. In addition, we use the fact that $Q_{\gamma T}^{\pm} \subset Q_T^{\pm}$. We take β as in (9.72). We obtain

$$\begin{aligned} & I_{\lambda,\beta}(W^*,f_0) - I_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0) - I'_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0)(W^* - W_{\min,\lambda,\beta}) \\ & \geq \frac{C_1}{\lambda} e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} \left(z_t^2 + \sum_{i,j=1}^{n} z_{x_i x_j}^2 \right) \varphi_\lambda d\mathbf{x}dt \\ & + C\lambda e^{-2\lambda B^2} \int\limits_{Q_T^{\pm}} [(\nabla z)^2 + \lambda^2 z^2] \varphi_\lambda d\mathbf{x}dt + e^{-\lambda T^2} \|z\|^2_{H^{k_n}(Q_T^{\pm})} \\ & \geq \frac{C_1}{\lambda} e^{-2\lambda B^2} \int\limits_{Q_{\gamma T}^{\pm}} \left(z_t^2 + \sum_{i,j=1}^{n} z_{x_i x_j}^2 \right) \varphi_\lambda d\mathbf{x}dt \\ & + C\lambda e^{-2\lambda B^2} \int\limits_{Q_{\gamma T}^{\pm}} [(\nabla z)^2 + \lambda^2 z^2] \varphi_\lambda d\mathbf{x}dt. \end{aligned} \tag{9.82}$$

By dropping the term $e^{-\lambda T^2} \|z\|^2_{H^{k_n}(Q_T^{\pm})}$ in the last line of (9.82), we have made that estimate stronger. Note that

$$\frac{1}{\lambda} e^{-2\lambda B^2} \varphi_\lambda(\mathbf{x},t) \geq \frac{1}{\lambda} \exp[-2\lambda(\gamma^2 T^2 + B^2 - A^2)] \quad \text{for } (\mathbf{x},t) \in Q_{\gamma T}^{\pm}. \tag{9.83}$$

Since (9.52) holds, then we replace (9.83) with a weaker estimate. The advantage of the latter estimate is that it allows us to obtain an explicit formula (9.90) for the dependence of the parameter $\lambda = \lambda(\delta)$ on δ. That estimate is

$$\frac{1}{\lambda} e^{-2\lambda B^2} \varphi_\lambda(\mathbf{x},t) \geq \exp[-3\lambda(\gamma^2 T^2 + B^2 - A^2)] \quad \text{for } (\mathbf{x},t) \in Q_{\gamma T}^{\pm}. \tag{9.84}$$

Hence, using (9.81), (9.82), and (9.84), we obtain

$$\begin{aligned} & I_{\lambda,\beta}(W^*,f_0) - I_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0) - I'_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0)(W^* - W_{\min,\lambda,\beta}) \\ & \geq C_1 \exp[-3\lambda(\gamma^2 T^2 + B^2 - A^2)] \|W^* - W_{\min,\lambda,\beta}\|^2_{H^{2,1}(Q_{\gamma T}^{\pm})}. \end{aligned} \tag{9.85}$$

By (9.50),

$$-I'_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0)(W^* - W_{\min,\lambda,\beta}) \leq 0. \tag{9.86}$$

Using (9.86), as well as the fact that $-I_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0) \leq 0$, we obtain

$$I_{\lambda,\beta}(W^*,f_0) - I_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0) - I'_{\lambda,\beta}(W_{\min,\lambda,\beta},f_0)(W^* - W_{\min,\lambda,\beta}) \leq I_{\lambda,\beta}(W^*,f_0).$$

Hence, (9.85) implies that

$$\exp[-3\lambda(\gamma^2 T^2 + B^2 - A^2)] \|W^* - W_{\min,\lambda,\beta}\|^2_{H^{2,1}(Q_{\gamma T}^{\pm})} \leq I_{\lambda,\beta}(W^*,f_0). \tag{9.87}$$

Recall that $\eta_1 = \gamma^2 T^2 + B^2 - A^2$. Hence, (9.79) and (9.87) imply that

$$\|W^* - W_{\min,\lambda,\beta}\|^2_{H^{2,1}(Q^{\pm}_{\gamma T})} \leq C_1(\delta^2 + \beta)\exp(3\lambda\eta_1). \tag{9.88}$$

We now specify the dependencies of λ and β on the noise level δ. Choose $\lambda = \lambda(\delta)$ such that

$$\exp(3\lambda(\delta)\eta_1) = \frac{1}{\delta}. \tag{9.89}$$

Hence,

$$\lambda(\delta) = \ln(\delta^{-1/(3\eta_1)}). \tag{9.90}$$

Since $\delta \in (0, \delta_0)$ and since (9.53) holds, then $\lambda(\delta) \geq \lambda_2$. Next, choose $\beta = \beta(\delta, T) = 2e^{-\lambda(\delta)T^2}$. Hence, in (9.88),

$$\beta\exp(3\lambda\eta_1) = 2\delta^{\eta_2/(3\eta_1)}. \tag{9.91}$$

Recalling that $2\rho = \min(1, \eta_2/(3\eta_1))$ and using (9.88)–(9.91), we obtain

$$\|W^* - W_{\min,\lambda(\delta),\beta(\delta,T)}\|_{H^{2,1}(Q^{\pm}_{\gamma T})} \leq C_1\delta^{\rho}. \tag{9.92}$$

Hence, the triangle inequality, (9.41), (9.54), (9.92), and the fact that $\rho \in (0, 1/2]$ lead to

$$\begin{aligned}\|w^* - w_{\min,\lambda(\delta),\beta(\delta,T)}\|_{H^{2,1}(Q^{\pm}_{\gamma T})} &\leq \|W^* - W_{\min,\lambda(\delta),\beta(\delta,T)}\|_{H^{2,1}(Q^{\pm}_{\gamma T})} \\ &\quad + \|G^* - G\|_{H^{2,1}(Q^{\pm}_{\gamma T})} \\ &\leq C_1\delta^{\rho} + \delta \leq (C_1 + 1)\delta^{\rho},\end{aligned} \tag{9.93}$$

which proves (9.55). Finally, since (9.55) holds, then (9.56) follows immediately from (9.25).

9.7 Proof of Theorem 9.4.7

Recall that Theorem 9.4.6 is valid; see item 2 in Remarks 9.4.2 (Section 9.4). By the triangle inequality, (9.41), (9.55), and (9.58),

$$\begin{aligned}\|w^* - w_n\|_{H^{2,1}(Q^{\pm}_{\gamma T})} &= \|w^* - w_{\min,\lambda(\delta),\beta(\delta,T)} + (w_{\min,\lambda(\delta),\beta(\delta,T)} - w_n)\|_{H^{2,1}(Q^{\pm}_{\gamma T})} \\ &\leq C_1\delta^{\rho} + \|w_{\min,\lambda(\delta),\beta(\delta,T)} - w_n\|_{H^{2,1}(Q^{\pm}_{\gamma T})} \\ &\leq C_1\delta^{\rho} + \|w_{\min,\lambda(\delta),\beta(\delta,T)} - w_n\|_{H^{2,1}(Q^{\pm}_{T})} \\ &= C_1\delta^{\rho} + \|W_{\min,\lambda(\delta),\beta(\delta,T)} - W_n\|_{H^{k_n}(Q^{\pm}_{T})} \\ &\leq C_1\delta^{\rho} + \theta^n\|W_{\min,\lambda(\delta),\beta(\delta,T)} - W_0\|_{H^{k_n}(Q^{\pm}_{T})} \\ &= C_1\delta^{\rho} + \theta^n\|w_{\min,\lambda(\delta),\beta(\delta,T)} - w_0\|_{H^{k_n}(Q^{\pm}_{T})},\end{aligned}$$

which proves (9.59). Estimate (9.60) follows immediately from (9.59) and the discussion in the last paragraph of Section 9.3.

9.8 Numerical testing

In the following tests, we set the domain $\Omega = (1,2) \times (1,2)$ and also

$$Lu = \Delta u - c(\mathbf{x})u.$$

To solve the inverse problem, we should first computationally simulate the data (9.7), (9.8) via the numerical solution of the forward problem (9.4). To solve problem (9.4), computationally, we have used the standard finite difference method. The spatial mesh size is $1/640 \times 1/640$ while the temporal one $T/512$. For the forward problem, we use the implicit scheme to compute the data needed for the inverse problem.

In computations of the inverse problem, the spatial mesh size is $1/16 \times 1/16$ and the temporal one $T/16$. When minimizing the functional $J_{\lambda,\beta}(w)$ in the discrete sense, we formulate the right-hand side of (9.24) via finite differences and minimize with respect to the values of the function w at grid points. To minimize the discretized functional, we use Matlab's built-in function *fminunc* with its option of *quasi-Newton algorithm*. This procedure calculates the gradient $\nabla J_{\lambda,\beta}(w)$ automatically and iterations stop when the condition $|\nabla J_{\lambda,\beta}(w)| < 10^{-2}$ holds. Note that even though our theory requires the application of the gradient projection method, our numerical observation is that we can avoid the use of the projection operator P_B and to work with just the conjugate gradient method. In fact, the use of the operator P_B would complicate the numerical implementation. The same observation took place in Chapters 7, 8, and 10 as well as in all our works on the convexification, which contain numerical studies [9, 115–117, 142–145, 145, 146, 150, 151, 164]. Also, we have minimized the functional $J_{\lambda,\beta}(w)$ rather than $I_{\lambda,\beta}(W)$ and it worked quite well.

As to (9.25), we have numerically discovered that rather than taking an average over $t \in [-\gamma T, \gamma T]$, better to use (9.13) at $\{t = t_0\}$. In numerical tests below, we took

$$\lambda = 1, \quad k_n = 3, \quad \beta = 0.01. \tag{9.94}$$

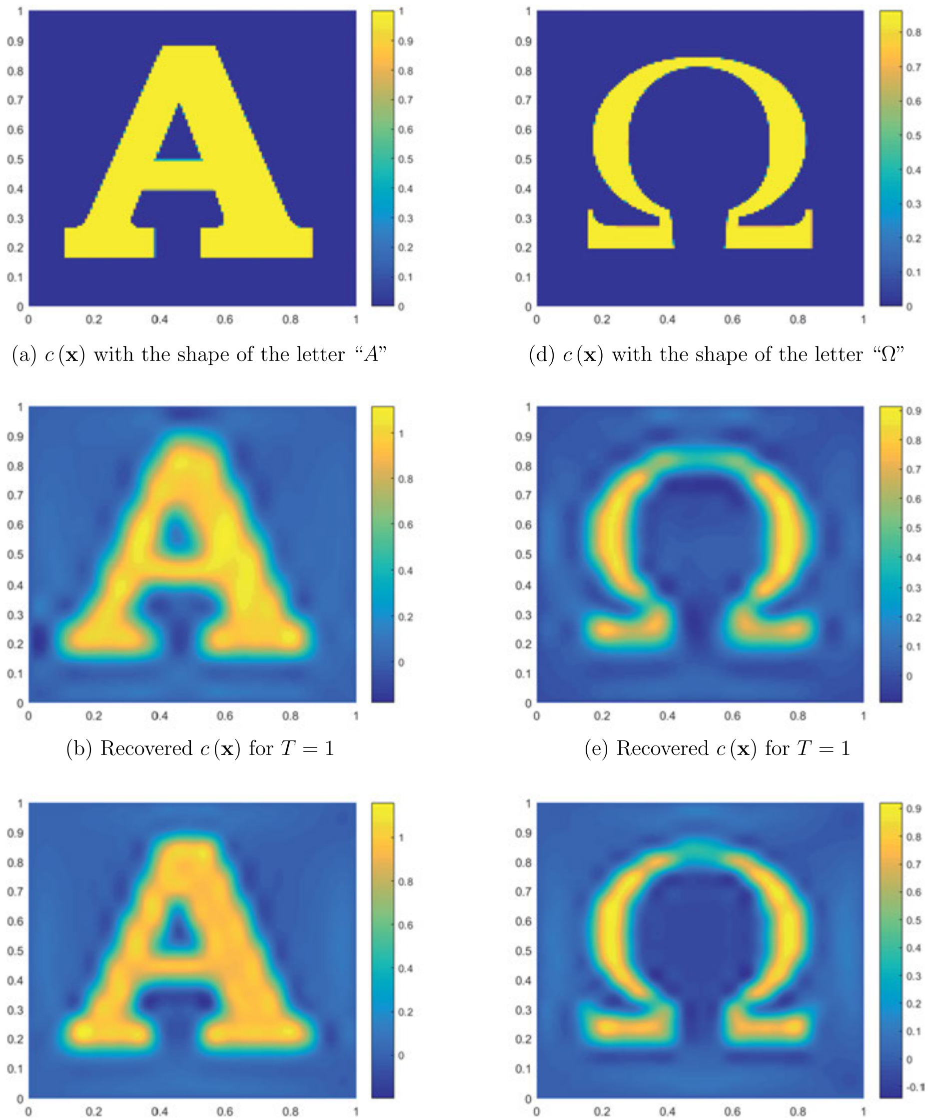
In the process of the minimization of the functional $J_{\lambda,\beta}(w)$, the starting point of iterations is always chosen to be the null function of value zero everywhere.

In the following three tests, we show the results of the recovery of the coefficients $c(\mathbf{x})$ with sophisticated structures. We choose the tested coefficients $c(\mathbf{x})$ having the shapes of the letters "A" and "Ω." We measure $g_1(x_1, x_2, t)$ on 16×32 detectors uniformly distributed on the rectangle $\Gamma_T^{\pm}$ and "measure" the function $f_0(x_1, x_2, t_0)$ on 16×16 detectors uniformly distributed on the square $(1,2)\times(1,2)\times\{t = t_0\}$. As initial and Dirichlet boundary conditions for the data simulations in (9.5), (9.6), we took

$$u(\mathbf{x}, -T) = 1 + \sin(\pi(x_1 - 1))\sin(\pi(x_2 - 1)) \quad \text{and} \quad u|_{S_T^{\pm}} = 1.$$

We allow in our tests the function $c(\mathbf{x})$ to be both positive and negative. Indeed, we have imposed the positivity condition (9.3) only to ensure that the function $u(\mathbf{x}, t) \neq 0$ in $\overline{Q_T^{\pm}}$. However, we have not observed any zeros of this function in our numerical studies.

Test 1. First, we test the reconstruction by our method of the coefficients $c(\mathbf{x})$ with the shapes of letters "A" and "Ω." In this test, we measure the data at time $\{t_0 = 0\}$ for the cases $T = 1$ and $T = 0.1$. The numerical results are shown in Figure 9.1.



(a) $c(\mathbf{x})$ with the shape of the letter "A"

(d) $c(\mathbf{x})$ with the shape of the letter "Ω"

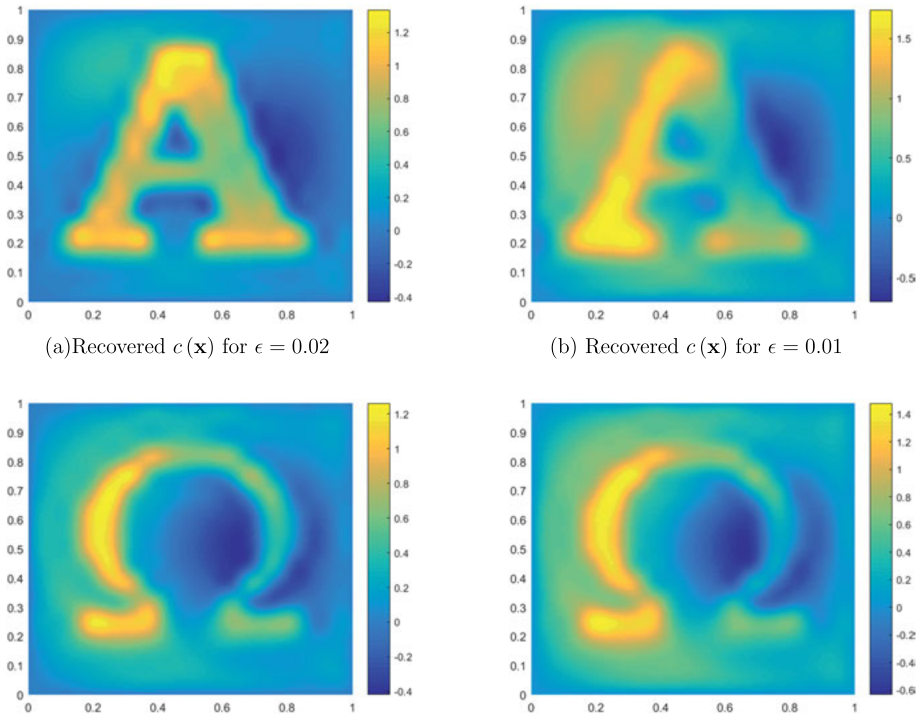
(b) Recovered $c(\mathbf{x})$ for $T = 1$

(e) Recovered $c(\mathbf{x})$ for $T = 1$

(c) Recovered $c(\mathbf{x})$ for $T = 0.1$

(f) Recovered $c(\mathbf{x})$ for $T = 0.1$

Figure 9.1: Results of Test 1. Here, $t_0 = 0$ in (9.8). (a) The coefficient $c(\mathbf{x})$ with the shape of the letter "A." (d) The coefficient $c(\mathbf{x})$ with the shape of the letter "Ω." (b) and (c) are the recovered $c(\mathbf{x})$ for $T = 1$ and $T = 0.1$ respectively for coefficient with the shape of the letter "A." (e) and (f) are the recovered $c(\mathbf{x})$ for $T = 1$ and $T = 0.1$, respectively, for coefficient with the shape of the letter "Ω."



(a)Recovered $c(\mathbf{x})$ for $\epsilon = 0.02$ (b) Recovered $c(\mathbf{x})$ for $\epsilon = 0.01$

(c) Recovered $c(\mathbf{x})$ for $\epsilon = 0.02$ (d) Recovered $c(\mathbf{x})$ for $\epsilon = 0.01$

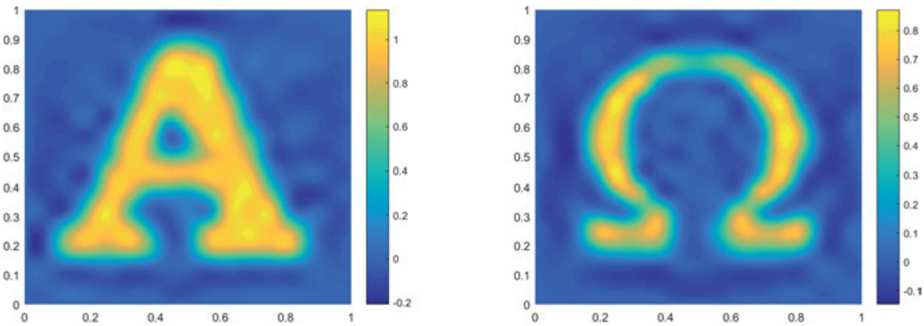
Figure 9.2: Results of Test 2. Here, $T = 0.1$ and in (9.8) $t_0 = 0$. We show the results in the case when the data are measured at a time t_0, which is close to the initial time moment $\{t = -T = -0.1\}$. (a) and (b) are the recovered $c(\mathbf{x})$ for $\epsilon = 0.02$ and $\epsilon = 0.01$, respectively, for coefficient with the shape of the letter "A." (c) and (d) are the recovered $c(\mathbf{x})$ for $\epsilon = 0.02$ and $\epsilon = 0.01$, respectively, for coefficient with the shape of the letter "Ω." Comparison with Figure 9.1 shows that the quality of images is better if t_0 is not too close to the initial time moment $\{t = -T\}$.

Test 2. In this test, we set $T = 0.1$. We show the results in the case when the data are measured at a time $\{t_0\}$, which is close to the initial time $\{t = -T = -0.1\}$. We take $t_0 = -T + \epsilon$ with $\epsilon = 0.02$ and $\epsilon = 0.01$. We test the reconstruction by our method of the coefficients $c(\mathbf{x})$ with the shapes of the letters "A" and "Ω." The numerical results are shown in Figure 9.2. In this test, we demonstrate the results when one measures the data at some time close to the initial time. It is numerically shown that the closer t_0 is to the initial time $t = -T$, the worse the result is.

Test 3. We now want to see how the random noise in the data influences our reconstruction. We add 5 % relative random noise to each detector on $\Gamma_T^{\pm}$ as well as on $(1,2) \times (1,2) \times \{t = 0\}$, that is, we work now with the noisy data,

$$u_x^{\text{noise}}|_{\Gamma_T^{\pm}} = g_1(\mathbf{x}, t) + \sigma \xi_{\mathbf{x},t} g_1(\mathbf{x}, t), \tag{9.95}$$

$$u^{\text{noise}}(\mathbf{x}, t_0) = f_0(\mathbf{x}) + \sigma \xi_{\mathbf{x}} f_0(\mathbf{x}). \tag{9.96}$$



(a)Recovered $c(\mathbf{x})$ with the shape "A" (b) Recovered $c(\mathbf{x})$ with the shape "Ω"

Figure 9.3: Results of Test 3. Here, $T = 1$ and $t_0 = 0$ in (9.8). (a) and (b) are the recovered coefficients $c(\mathbf{x})$ with the shapes of the letters "A" and "Ω," respectively. The measured data contain 5% relative random noise.

Here, $\sigma = 5\,\%$ is the noise level, $\xi_{\mathbf{x},t}$ and $\xi_{\mathbf{x}}$ are independent normally distributed random variables. To preprocess the noisy data, we use the thin plate spline smoother developed in [61]. The algorithm proposed in [61] provides a good approximation to the true function without knowing neither the noise level nor any other a priori information of the true function to be approximated. Then the cubic B-splines are employed to approximate the first- and second-order derivatives of the noisy data. In this test, we "measure" $g_1(x_1, x_2, t)$ on 16×32 detectors uniformly distributed on the plane $\Gamma_T^{\pm}$ and also "measure" $f_0(x_1, x_2, t_0)$ on 160×160 detectors uniformly distributed on the plane $(1, 2) \times (1, 2) \times \{t = 0\}$. We now set $T = 1$, in (9.8) $t_0 = 0$, and the noise is added to the data as in (9.95), (9.96). We test the reconstruction by our method of the coefficients with the shape of the letters "A" and "Ω." The numerical results are shown in Figure 9.3. We see that our method works still very well in the mild noisy case.

9.9 Proof of Theorem 9.4.1

We prove this theorem only for functions $u(\mathbf{x}, t)$ such that

$$u \in C^3(\overline{Q_T^{\pm}}), \quad u|_{S_T^{\pm}} = u_x|_{\Gamma_T^{\pm}} = 0. \tag{9.97}$$

The case $u \in H_0^{2,1}(Q_T^{\pm})$ follows immediately from this proof via density arguments. Below in this proof, $O(1/\lambda^k)$, $k \geq 1$ denotes different smooth functions, which are independent on u and for which the following estimate is valid $\|O(1/\lambda^k)\|_{C^1(\overline{Q_T^{\pm}})} \leq C/\lambda^k$, $\forall \lambda \geq 1$.

Recall that by (9.20) $\varphi_\lambda(x,t) = \exp(2\lambda(x^2 - t^2))$. Introduce a new function $v(\mathbf{x},t) = u(\mathbf{x},t)\exp(\lambda(x^2 - t^2))$. Then $u = v\exp(-\lambda(x^2 - t^2))$. Hence,

$$
\begin{aligned}
(u_t - \Delta u) &= \left(v_t - \Delta v + 4\lambda x v_x - 4\lambda^2 x^2 (1 - 1/(2\lambda x^2) + t/(2\lambda x^2))v\right)\exp\left(-\lambda(x^2 - t^2)\right) \\
&= \left[\left(-\Delta v - 4\lambda^2 x^2 (1 + O(1/\lambda))v\right) + (v_t + 4\lambda x v_x)\right]\exp\left(-\lambda(x^2 - t^2)\right).
\end{aligned}
$$

Hence,

$$
(u_t - \Delta u)^2 \varphi_\lambda \geq (2v_t + 8\lambda x v_x)\left(-\Delta v - 4\lambda^2 x^2 (1 + O(1/\lambda))v\right). \tag{9.98}
$$

Step 1. Estimate from the below the following term in (9.98):
$2v_t(-\Delta v - 4\lambda x^2(1 + O(1/\lambda))v + 2\lambda t v)$. We have

$$
\begin{aligned}
&2v_t\left(-\Delta v - 4\lambda x^2(1 + O(1/\lambda))v + 2\lambda t v\right) \\
&\quad = -2\sum_{i=1}^{n} v_{x_i x_i} v_t + \left(-4\lambda^2 x^2 (1 + O(1/\lambda))v^2 + 2\lambda t v^2\right)_t - 2\lambda v^2 \\
&\quad = \sum_{i=1}^{n}(-2v_{x_i} v_t)_{x_i} + 2\sum_{i=1}^{n} v_{x_i} v_{x_i t} + \left(-4\lambda^2 x^2 (1 + O(1/\lambda))v^2 + 2\lambda t v^2\right)_t - 2\lambda v^2 \\
&\quad = -2\lambda v^2 + \sum_{i=1}^{n}(-2v_{x_i} v_t)_{x_i} + \left((\nabla v)^2 - 4\lambda^2 x^2 (1 + O(1/\lambda))v^2 + 2\lambda t v^2\right)_t.
\end{aligned}
$$

Thus,

$$
2v_t\left(-\Delta v - 4\lambda x^2(1 + O(1/\lambda))v + 2\lambda t v\right) = -2\lambda v^2 + \operatorname{div} U_1 + V_{1t}, \tag{9.99}
$$

$$
\operatorname{div} U_1 = \sum_{i=1}^{n}(-2v_{x_i} v_t)_{x_i}, \tag{9.100}
$$

$$
V_1(\mathbf{x},t) = (\nabla v)^2 - 4\lambda^2 x^2 \left(1 + O(1/\lambda)\right)v^2. \tag{9.101}
$$

Step 2. Estimate from the below the following term in (9.98):
$8\lambda x v_x(-\Delta v - 4\lambda^2 x^2(1 + O(1/\lambda))v + 2\lambda t v)$. We have

$$
\begin{aligned}
&8\lambda x v_x\left(-\Delta v - 4\lambda^2 x^2 (1 + O(1/\lambda))v + 2\lambda t v\right) \\
&\quad = -8\lambda x v_x v_{xx} + \sum_{i=2}^{n}(-8\lambda x v_x v_{x_i x_i}) \\
&\qquad + \left(-16\lambda^3 x^3 (1 + O(1/\lambda))v^2 + 8\lambda^2 x t v^2\right)_x + 48\lambda^3 x^2 (1 + O(1/\lambda))v^2 \\
&\quad = \left(-4\lambda x v_x^2\right)_x + 4\lambda v_x^2 + \sum_{i=2}^{n}(-8\lambda x v_x v_{x_i})_{x_i} + \sum_{i=2}^{n}(8\lambda x v_{x x_i} v_{x_i}) \\
&\qquad + 48\lambda^3 x^2 (1 + O(1/\lambda))v^2 + \left(-16\lambda^3 x^3 (1 + O(1/\lambda))v^2 + 8\lambda^2 x t v^2\right)_x \\
&\quad = 4\lambda\left(v_x^2 - \sum_{i=2}^{n} v_{x_i}^2\right) + 48\lambda^3 x^2 (1 + O(1/\lambda))v^2
\end{aligned}
$$

$$+\left(-4\lambda x v_x^2 + 4\lambda x \sum_{i=2}^{n} v_{x_i}^2 - 16\lambda^3 x^3 (1 + O(1/\lambda)) v^2 + 8\lambda^2 x t v^2\right)_x$$
$$+ \sum_{i=2}^{n} (-8\lambda x v_x v_{x_i})_{x_i}.$$

Thus, we end up with the following estimate of Step 2:

$$8\lambda x v_x(-\Delta v - 4\lambda^2 x^2 (1 + O(1/\lambda)) v + 2\lambda t v)$$
$$= 4\lambda\left(v_x^2 - \sum_{i=2}^{n} v_{x_i}^2\right) + 48\lambda^3 x^2 (1 + O(1/\lambda)) v^2 + \operatorname{div} U_2, \tag{9.102}$$
$$\operatorname{div} U_2 = \left(-4\lambda x v_x^2 + 4\lambda x \sum_{i=2}^{n} v_{x_i}^2 - 16\lambda^3 x^3 (1 + O(1/\lambda)) v^2 + 8\lambda^2 x t v^2\right)_x \tag{9.103}$$
$$+ \sum_{i=2}^{n} (-8\lambda x v_x v_{x_i})_{x_i}.$$

Step 3. Analysis of boundary integrals over $S_T^{\pm}$.

Let $\nu = \nu(\mathbf{x})$ be the unit outward looking normal vector to $\partial\Omega$ at the point $x \in \partial\Omega$. By Gauss' formula, (9.100) and (9.103),

$$\int_{Q_T^{\pm}} (\operatorname{div} U_1 + \operatorname{div} U_2) d\mathbf{x} dt = \int_{-T}^{T} \int_{\partial\Omega} \sum_{i=1}^{n} (U_{1,i} + U_{2,i}) \cos(\nu(\mathbf{x}), x_i) dS dt, \tag{9.104}$$

where $U_k = (U_{k,1}, \ldots, U_{k,n})$, $k = 1, 2$. Obviously, $U_{1i} = -2 v_{x_i} v_t$. Since (9.97) holds and since

$$v_t(\mathbf{x}, t) = (u_t - 2\lambda t u)(\mathbf{x}, t) \exp(\lambda(x^2 - t^2)),$$

then $v_t(\mathbf{x}, t) = 0$ for $\mathbf{x} \in \partial\Omega$. Hence, in (9.104),

$$\int_{-T}^{T} \int_{\partial\Omega} \sum_{i=1}^{n} U_{1,i} \cos(\nu(\mathbf{x}), x_i) dS dt = 0. \tag{9.105}$$

We now analyze the first term in the right-hand side of (9.103). We have

$$v_x(\mathbf{x}, t) = (u_x + 2\lambda x u)(\mathbf{x}, t) \exp(\lambda(x^2 - t^2)), \tag{9.106}$$
$$v_{x_i}(\mathbf{x}, t) = u_{x_i}(\mathbf{x}, t) \exp(\lambda(x^2 - t^2)), \quad i = 2, \ldots, n. \tag{9.107}$$

By (9.97), (9.103), (9.106), and (9.107),

$$\int_{-T}^{T} U_{2,1} \cos(\nu(\mathbf{x}), x_1) dS dt = 4\lambda A \int_{-T}^{T} \int_{\Gamma'} u_x^2(A, \overline{x}) e^{2\lambda(A^2 - t^2)} dS \geq 0, \tag{9.108}$$

where $\Gamma' = \{x = A\} \cap \partial\Omega$. Similarly,

$$\int_{-T}^{T}\int_{\partial\Omega} U_{2,i}\cos(\nu(\mathbf{x}), x_i)dSdt = 0, \quad i \geq 2. \tag{9.109}$$

Using (9.104), (9.105), (9.108), and (9.109), we obtain

$$\int_{Q_T^{\pm}} (\operatorname{div} U_1 + \operatorname{div} U_2)d\mathbf{x}dt \geq 0. \tag{9.110}$$

Step 4. Sum up (9.99) with (9.102), replace the right-hand side of (9.98) by the resulting inequality and then integrate the obtained inequality over $Q_T^{\pm}$. Use in that integral (9.97), Gauss' formula, (9.101) and (9.105)–(9.109). We obtain for all $\lambda \geq \lambda_0$ and for all $u \in C^2(\overline{Q_T^{\pm}}) \cap H_0^{2,1}(Q_T^{\pm})$,

$$\begin{aligned}\int_{Q_T^{\pm}} (u_t - \Delta u)^2\varphi_\lambda d\mathbf{x}dt \geq -4\lambda \int_{Q_T^{\pm}} (\nabla u)^2\varphi_\lambda d\mathbf{x}dt + 47\lambda^3 \int_{Q_T^{\pm}} u^2x^2\varphi_\lambda d\mathbf{x}dt \\ + \int_{\Omega} (V_1(\mathbf{x}, T) - V_1(\mathbf{x}, -T))d\mathbf{x}.\end{aligned} \tag{9.111}$$

The inconvenient point of (9.111) is the presence of the negative term in the first line of (9.111). Therefore, we continue.

Step 5. Estimate from the below $(u_t - \Delta u)u\varphi_\lambda$, and then estimate the corresponding integral over $Q_T^{\pm}$,

$$\begin{aligned}(u_t - \Delta u)u\varphi_\lambda &= \left(\frac{u^2}{2}\varphi_\lambda\right)_t + 2\lambda t u^2\varphi_\lambda + (-u_x u\varphi_\lambda)_x + u_x^2\varphi_\lambda + 4\lambda x u_x u\varphi_\lambda \\ &\quad + \sum_{i=2}^{n}(-u_{x_i}u\varphi_\lambda)_{x_i} + \sum_{i=2}^{n} u_{x_i}^2\varphi_\lambda \\ &= (\nabla u)^2\varphi_\lambda + \sum_{i=1}^{n}(-u_{x_i}u\varphi_\lambda)_{x_i} + \left(\frac{u^2}{2}\varphi_\lambda\right)_t + (2\lambda x u^2\varphi_\lambda)_x \\ &\quad - 8\lambda^2x^2(1 + O(1/\lambda))u^2\varphi_\lambda.\end{aligned}$$

Hence,

$$(u_t - \Delta u)u\varphi_\lambda \geq (\nabla u)^2\varphi_\lambda - 9\lambda^2x^2u^2\varphi_\lambda + \operatorname{div} U_3 + V_{2t}, \tag{9.112}$$

$$\operatorname{div} U_3 = \sum_{i=1}^{n}(-u_{x_i}u\varphi_\lambda)_{x_i} + (2\lambda x u^2\varphi_\lambda)_x, \tag{9.113}$$

$$V_2(\mathbf{x}, t) = \frac{u^2}{2}\varphi_\lambda. \tag{9.114}$$

Hence, by (9.97), (9.113), and Gauss' formula,

$$\int_{Q_T^\pm} \operatorname{div} U_3 d\mathbf{x}dt = 0. \tag{9.115}$$

Integrate (9.112) over $Q_T^\pm$ using (9.114) and (9.115). Then multiply the resulting inequality by 5λ and sum up with (9.111). We obtain

$$\begin{aligned} &5\lambda \int_{Q_T^\pm} (u_t - \Delta u) u\varphi_\lambda dxdt + \int_{Q_T^\pm} (u_t - \Delta u)^2 \varphi_\lambda d\mathbf{x}dt \\ &\geq \lambda \int_{Q_T^\pm} (\nabla u)^2 \varphi_\lambda d\mathbf{x}dt + 2\lambda^3 \int_{Q_T^\pm} u^2 x^2 \varphi_\lambda d\mathbf{x}dt \\ &+ \int_\Omega [(V_1 + 5\lambda V_2)(\mathbf{x}, T) - (V_1 + 5\lambda V_2)(\mathbf{x}, -T)] d\mathbf{x}. \end{aligned} \tag{9.116}$$

Next, by the Cauchy–Schwarz inequality,

$$\begin{aligned} &5\lambda \int_{Q_T^\pm} (u_t - \Delta u) u\varphi_\lambda dxdt + \int_{Q_T^\pm} (u_t - \Delta u)^2 \varphi_\lambda d\mathbf{x}dt \\ &\leq \frac{7}{2} \int_{Q_T^\pm} (u_t - \Delta u)^2 \varphi_\lambda d\mathbf{x}dt + \frac{5}{2}\lambda^2 \int_{Q_T^\pm} u^2 \varphi_\lambda d\mathbf{x}dt. \end{aligned} \tag{9.117}$$

Since for sufficiently large $\lambda_0 > 1$ and for $\lambda \geq \lambda_0$,

$$2\lambda^3 \int_{Q_T^\pm} u^2 x^2 \varphi_\lambda d\mathbf{x}dt - \frac{5}{2}\lambda^2 \int_{Q_T^\pm} u^2 \varphi_\lambda d\mathbf{x}dt \geq \lambda^3 \int_{Q_T^\pm} u^2 x^2 \varphi_\lambda d\mathbf{x}dt, \tag{9.118}$$

then (9.116)–(9.118) imply that for all $u \in H_0^{2,1}(Q_T^\pm)$ and for all $\lambda \geq \lambda_0$,

$$\begin{aligned} \int_{Q_T^\pm} (u_t - \Delta u)^2 \varphi_\lambda d\mathbf{x}dt &\geq C\lambda \int_{Q_T^\pm} [(\nabla u)^2 + \lambda^2 u^2] \varphi_\lambda d\mathbf{x}dt \\ &+ \int_\Omega [(V_1 + 5\lambda V_2)(\mathbf{x}, T) - (V_1 + 5\lambda V_2)(\mathbf{x}, -T)] d\mathbf{x}, \end{aligned} \tag{9.119}$$

which is a part of estimate (9.28).

Step 6. In this step, we incorporate terms with $u_t^2, u_{x_i x_j}^2$.

We have

$$(u_t - \Delta u)^2 \varphi_\lambda = u_t^2 \varphi_\lambda - 2u_t u_{xx} \varphi_\lambda - \sum_{i=2}^n 2u_t u_{x_i x_i} \varphi_\lambda + (\Delta u)^2 \varphi_\lambda. \tag{9.120}$$

Denote

$$z_1 = -2u_t u_{xx}\varphi_\lambda, \quad z_2 = -\sum_{i=2}^{n} 2u_t u_{x_i x_i}\varphi_\lambda, \quad z_3 = (\Delta u)^2\varphi_\lambda \tag{9.121}$$

and estimate each of terms in (9.121). First, we have

$$\begin{aligned} z_1 &= -2u_t u_{xx}\varphi_\lambda = (-2u_t u_x\varphi_\lambda)_x + 2u_{tx}u_x\varphi_\lambda + 8\lambda x u_t u_x\varphi_\lambda \\ &= (-2u_t u_x\varphi_\lambda)_x + \left(u_x^2\varphi_\lambda\right)_t + 2\lambda t u_x^2\varphi_\lambda + 8\lambda x u_t u_x\varphi_\lambda \\ &\geq -\frac{1}{2}u_t^2\varphi_\lambda - C\lambda^2 u_x^2\varphi_\lambda + (-2u_t u_x\varphi_\lambda)_x + \left(u_x^2\varphi_\lambda\right)_t. \end{aligned}$$

Thus,

$$z_1 \geq -\frac{1}{2}u_t^2\varphi_\lambda - C\lambda^2 u_x^2\varphi_\lambda + (-2u_t u_x\varphi_\lambda)_x + \left(u_x^2\varphi_\lambda\right)_t. \tag{9.122}$$

We now estimate z_2,

$$\begin{aligned} z_2 &= \sum_{i=2}^{n}(-2u_t u_{x_i}\varphi_\lambda)_{x_i} + \sum_{i=2}^{n} 2u_{tx_i}u_{x_i}\varphi_\lambda \\ &= \sum_{i=2}^{n}\left(u_{x_i}^2\varphi_\lambda\right)_t + 4\lambda t\sum_{i=2}^{n} u_{x_i}^2\varphi_\lambda + \sum_{i=2}^{n}(-2u_t u_{x_i}\varphi_\lambda)_{x_i} \\ &\geq -C\lambda\sum_{i=2}^{n} u_{x_i}^2\varphi_\lambda + \sum_{i=2}^{n}\left(u_{x_i}^2\varphi_\lambda\right)_t + \sum_{i=2}^{n}(-2u_t u_{x_i}\varphi_\lambda)_{x_i}. \end{aligned}$$

Thus,

$$z_2 \geq -C\lambda\sum_{i=2}^{n} u_{x_i}^2\varphi_\lambda + \sum_{i=2}^{n}\left(u_{x_i}^2\varphi_\lambda\right)_t + \sum_{i=2}^{n}(-2u_t u_{x_i}\varphi_\lambda)_{x_i}. \tag{9.123}$$

Now we estimate z_3,

$$\begin{aligned} z_3 &= (\Delta u)^2\varphi_\lambda = \left(u_{xx} + \sum_{i=2}^{n} u_{x_i x_i}\right)^2\varphi_\lambda \\ &= \sum_{i=1}^{n} u_{x_i x_i}^2\varphi_\lambda + 2\sum_{i=2}^{n} u_{xx}u_{x_i x_i}\varphi_\lambda + \sum_{i,j=2,i\neq j}^{n} u_{x_i x_i}u_{x_j x_j}\varphi_\lambda \\ &= \sum_{i=1}^{n} u_{x_i x_i}^2\varphi_\lambda + \left(2\sum_{i=2}^{n} u_x u_{x_i x_i}\varphi_\lambda\right)_x - 8\lambda x\sum_{i=2}^{n} u_x u_{x_i x_i}\varphi_\lambda \\ &\quad - 2\sum_{i=2}^{n} u_x u_{x x_i x_i}\varphi_\lambda + \left(\sum_{i,j=2,i\neq j}^{n} u_{x_i}u_{x_j x_j}\varphi_\lambda\right)_{x_i} - \sum_{i,j=2,i\neq j}^{n} u_{x_i}u_{x_i x_j x_j}\varphi_\lambda \\ &= \sum_{i=1}^{n} u_{x_i x_i}^2\varphi_\lambda + \left(2\sum_{i=2}^{n} u_x u_{x_i x_i}\varphi_\lambda\right)_x - 8\lambda x\sum_{i=2}^{n} u_x u_{x_i x_i}\varphi_\lambda \end{aligned} \tag{9.124}$$

$$+\left(-2\sum_{i=2}^{n}u_x u_{xx_i}\varphi_\lambda\right)_{x_i}+2\sum_{i=2}^{n}u_{xx_i}^2\varphi_\lambda+\left(\sum_{i,j=2,i\neq j}^{n}u_{x_i}u_{x_jx_j}\varphi_\lambda\right)_{x_i}$$
$$+\left(-\sum_{i,j=2,i\neq j}^{n}u_{x_i}u_{x_ix_j}\varphi_\lambda\right)_{x_j}+\sum_{i,j=2,i\neq j}^{n}u_{x_ix_j}^2\varphi_\lambda.$$

Since by the Cauchy–Schwarz inequality,

$$-8\lambda x\sum_{i=2}^{n}u_x u_{x_ix_i}\varphi_\lambda \geq -C\lambda^2(\nabla u)^2\varphi_\lambda-\frac{1}{2}\sum_{i=2}^{n}u_{x_ix_i}^2\varphi_\lambda,$$

then (9.124) implies that

$$\begin{aligned} z_3 \geq{} & \frac{1}{2}\sum_{i,j=1}^{n}u_{x_ix_j}^2\varphi_\lambda - C\lambda^2(\nabla u)^2\varphi_\lambda \\ & +\left(2\sum_{i=2}^{n}u_x u_{x_ix_i}\varphi_\lambda\right)_x+\left(-2\sum_{i=2}^{n}u_x u_{xx_i}\varphi_\lambda\right)_{x_i} \\ & +\left(\sum_{i,j=2,i\neq j}^{n}u_{x_i}u_{x_jx_j}\varphi_\lambda\right)_{x_i}+\left(-\sum_{i,j=2,i\neq j}^{n}u_{x_i}u_{x_ix_j}\varphi_\lambda\right)_{x_j}. \end{aligned} \tag{9.125}$$

Combining (9.120)–(9.125), we obtain

$$(u_t-\Delta u)^2\varphi_\lambda \geq \left(u_t^2+\sum_{i,j=1}^{n}u_{x_ix_j}^2\right)\varphi_\lambda-\widetilde{C}\lambda^2(\nabla u)^2\varphi_\lambda$$
$$+\operatorname{div}U_4+V_{3t}, \tag{9.126}$$

$$\int_{Q_T^{\pm}}\operatorname{div}U_4 d\mathbf{x}dt=0, \tag{9.127}$$

$$V_3=(\nabla u)^2\varphi_\lambda. \tag{9.128}$$

In (9.126), the number $\widetilde{C}=\widetilde{C}(\Omega,T)>0$ depends on the same parameters as the above number C. We introduce $\widetilde{C}$ here with the goal to obtain a better combination of (9.126) with (9.119).

Using (9.127) and (9.128), integrate (9.126) over $Q_T^{\pm}$. Next, multiply the resulting inequality by $C/(2\widetilde{C}\lambda)$. Next, sum up the resulting estimate with (9.119). In doing so, keep in mind that (9.101), (9.106), (9.114), and (9.128) imply for $\lambda \geq \lambda_0$,

$$\left(V_1+5\lambda V_2+\frac{C}{2\widetilde{C}\lambda}V_3\right)(\mathbf{x},T)\geq -C\lambda^2\exp(2\lambda(B^2-T^2))u^2(\mathbf{x},T),$$
$$-\left(V_1+5\lambda V_2+\frac{C}{2\widetilde{C}\lambda}V_3\right)(\mathbf{x},-T)\geq -C\exp(2\lambda(B^2-T^2))((\nabla u)^2+\lambda^2u^2)(\mathbf{x},-T).$$

Then, using two latter inequalities and taking into account the fact that C denotes different positive constants depending only on Ω and T, we obtain the target estimate (9.28) of Theorem 9.4.1.

10 Experimental data and convexification for the recovery of the dielectric constants of buried targets using the Helmholtz equation

In this chapter, we follow publications [115, 117]. Permissions for republishing are obtained from the publishers.

10.1 Introduction

While in Chapters 7–9, we have developed the convexification method for electrical impedance tomography, a CIP for a hyperbolic PDE and a CIP for a parabolic PDE, we develop here the globally convergent convexification method for a 3D CIP for the Helmholtz equation. In our CIP, the wavenumber (i. e., frequency) is fixed and the backscattering boundary data for the inversion are generated by the point source moving along an interval of a straight line. First, we develop the convexification analytically. Next, we test it on microwave experimental data collected by the research group of Klibanov: He is in charge of two experimental devices located in Grigg Building of the University of North Carolina at Charlotte. More precisely, we calculate the dielectric constants of targets buried in a sandbox. It is well known that dielectric constants of buried targets are very hard to calculate using standoff measurements. Values of dielectric constants, in turn give us locations and shapes of targets. We point out that two (2) out of five (5) of our tests are for *the most challenging case of blind experimental data*.

Targets of our primary interest are those which mimic antipersonnel land mines and improvised explosive devices (IEDs). It was stated in [170] (p. 33) that even though the knowledge of the dielectric constant alone is insufficient to identify an explosive, one can still hope that this knowledge might serve as an important piece of information, additional to the conventional ones, to help better identify explosives, and thus, to decrease the false alarm rate.

The coefficient of the Helmholtz equation, that is, the spatially distributed dielectric constant, is the subject of the solution of our CIP. In principle, we probably would need to work with the full Maxwell's system rather than with the single Helmholtz equation. However, it was demonstrated numerically in [155] that if the incident electric wave field has only a single nonzero component, then this component dominates two other components while propagating through a medium. Furthermore, the propagation of this dominated component is well described by the Helmholtz equation. A similar result was obtained in [20] for the time dependent case. Besides, we will see below that the description by the Helmholtz equation provides accurate results for experimental data, which is an ultimate justification of this mathematical model; the

https://doi.org/10.1515/9783110745481-010

same was observed in many previous publications about experimental data [22, 24, 25, 115, 116, 139, 143, 168, 170, 204, 205, 242, 243].

In principle, the case when the source is moving and the frequency is fixed enables one to consider a physically realistic problem when the dielectric constant depends not only on spatial variables but on the frequency as well. Indeed, if we repeat those measurements for an interval of frequencies, then we can find the dependence of the dielectric constant on both spatial variables and the frequency. But we assume here that the dielectric constant depends only on spatial variables and not on the frequency. The analytical part of this paper is devoted to the derivation of the method and its convergence analysis.

Unlike this chapter, previously the convexification method was constructed for some CIPs for the Helmholtz equation only for the case of a single direction of the incident plane wave with the wavenumber running over a certain interval [142, 143, 145]. We demonstrate in our numerical studies below that, in the moving source case, the convexification method accurately images all three components of targets of interest: locations, shapes, and the target/background contrasts in the dielectric constant. This is unlike the above mentioned previously studied case of a single direction of the incident plane wave, which ensured only first and third components, while shapes were not accurately imaged.

One of strengths of the convexification is that it works only with the non-overdetermined data. This means that the number m of independent variables in the data equals the number n of independent variables in the unknown coefficient, $m = n$. In particular, in our CIP $m = n = 3$.

As to the CIPs with the fixed wavenumber, we refer to numerical procedures developed during a long standing effort by the group of R. G. Novikov since about 1988 [211]; also, see, for example, [1, 3, 213, 215]. See Section 1.4 for further comments on this issue.

Just as in Chapters 6, 7, 11, and 12, we use a special orthonormal basis in $L_2(a, b)$, which was first introduced in [136]. This basis is described in Section 6.2.3. And just as in those chapters, we truncate the Fourier series of a certain function with respect to this basis. Then we work with the resulting approximate mathematical model and do not prove convergence of our method at $N \to \infty$, where N is the number of terms in that truncated series. We refer to Remark 7.3 for a discussion of this issue.

We also note that we consider here the problem of finding the coefficient of the Helmholtz equation from scattering data in the case when the both the magnitude and the phase of the scattered wave are known. But there are also cases when the phase is unknown. Corresponding CIPs were considered in works of Klibanov and Romanov; see, for example, [159, 160, 226] for some references. Finally, we refer to some works on inverse scattering problems in the case when one is recovering locations and shapes of inclusions but not the values of the unknown coefficients inside [104, 186–188].

10.2 Statement of the coefficient inverse problem

We model the propagation of the electric wave field by the Helmholtz equation instead of the Maxwell's equations. This modeling was numerically justified in the Appendix of the paper [155]. Such a mathematical model is true at least for rather simple medium consisting of a homogeneous background and a few embedded inclusions. Besides, good accuracies of reconstructions obtained by our research group from experimental data in publications [115, 116, 143, 204, 205], where the Helmholtz equation was used to model the wave propagation process, speak in favor of this modeling.

Denote $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. Let the number $R > 0$. We define the cube $\Omega \subset \mathbb{R}^3$ as

$$\Omega = \{\mathbf{x} : |x|, |y| < R, z \in (-b, b)\}. \tag{10.1}$$

Let $\Gamma \subset \partial\Omega$ be the lower part of the boundary of Ω where measurements of the backscatter data are conducted,

$$\Gamma := \{\mathbf{x} : |x|, |y| < R, z = -b\}. \tag{10.2}$$

Let $c := c(\mathbf{x}) \in [1, \infty)$ be a sufficiently smooth function that represents the dielectric function of the medium. We assume that

$$\begin{cases} c(\mathbf{x}) \geq 1 & \text{in } \mathbb{R}^3, \\ c(\mathbf{x}) = 1 & \text{in } \mathbb{R}^3 \backslash \Omega. \end{cases} \tag{10.3}$$

Here, $k > 0$ is the wavenumber. The function c is the spatially distributed and k dependent dielectric constant. The second assumption in (10.3) means that we have vacuum outside of the domain of interest Ω.

Let a_1, a_2 and d be three numbers such that $d > b$ and $a_1 < a_2$. We define the line of sources as

$$L_{\text{src}} := \{(\alpha, 0, -d) : a_1 \leq \alpha \leq a_2\}. \tag{10.4}$$

Obviously, this line is parallel to the x-axis. The distance from L_{src} to Γ is d, and the length of our the line of sources is $a_2 - a_1$. Since $d > b$, then $L_{\text{src}} \cap \overline{\Omega} = \varnothing$. Thus, for each $\alpha \in [-a, a]$ the corresponding point source is $\mathbf{x}_\alpha := (\alpha, 0, -d) \in L_{\text{src}}$.

First, we formulate the forward problem. Let $k = \text{const.} > 0$ and assume that the function c is known. For each source position $\mathbf{x}_\alpha \in L_{\text{src}}$ the forward problem is

$$\Delta u + k^2 c(\mathbf{x}) u = -\delta(\mathbf{x} - \mathbf{x}_\alpha) \quad \text{in } \mathbb{R}^3, \tag{10.5}$$

$$\lim_{r \to \infty} r(\partial_r u - \mathrm{i}ku) = 0 \quad \text{for } r = |\mathbf{x} - \mathbf{x}_\alpha|, \mathrm{i} = \sqrt{-1}. \tag{10.6}$$

Conditions (10.5)–(10.6) form the Helmholtz equation with the Sommerfeld radiation condition at the infinity. Let $u_0(\mathbf{x}, \alpha)$ be the solution of (10.5)–(10.6) with $c \equiv 1$,

$$u_0(\mathbf{x}, \alpha) = \frac{\exp(\mathrm{i}k|\mathbf{x} - \mathbf{x}_\alpha|)}{4\pi|\mathbf{x} - \mathbf{x}_\alpha|}. \tag{10.7}$$

Using the Helmholtz equation for $u_{0,\alpha} = u_0(\mathbf{x}, \alpha)$, we obtain from (10.5)–(10.6),

$$\Delta(u - u_{0,\alpha}) + k^2(u - u_{0,\alpha}) = -k^2(c(\mathbf{x}, k) - 1)u \quad \text{in } \mathbb{R}^3,$$
$$\lim_{r\to\infty} r[\partial_r(u - u_{0,\alpha}) - \mathrm{i}k(u - u_{0,\alpha})] = 0 \quad \text{for } r = |\mathbf{x} - \mathbf{x}_\alpha|.$$

In view of the fact that $c(\mathbf{x}) = 1$ in $\mathbb{R}^3 \setminus \Omega$, we thus find that the solution u to the system (10.5)–(10.6) satisfies the so-called Lippmann–Schwinger equation (see, e. g., [66, Section 8.2]), which reads for all $\mathbf{x} \in \mathbb{R}^3$ as

$$u(\mathbf{x}, \alpha) = u_0(\mathbf{x}, \alpha) + k^2 \int_\Omega \frac{\exp(\mathrm{i}k|\mathbf{x} - \mathbf{x}'|)}{4\pi|\mathbf{x} - \mathbf{x}'|}(c(\mathbf{x}') - 1)u(\mathbf{x}', \alpha)d\mathbf{x}'. \tag{10.8}$$

We now pose the CIP which we solve in this paper.

Coefficient Inverse Problem (CIP). Given a fixed wavenumber $k > 0$, determine the coefficient $c(\mathbf{x})$ for $\mathbf{x} \in \Omega$ in the system (10.5)–(10.6), assuming that the following function $F(\mathbf{x}, \mathbf{x}_\alpha)$ is given

$$F(\mathbf{x}, \mathbf{x}_\alpha) = u(\mathbf{x}, \alpha) \quad \text{for } \mathbf{x} \in \Gamma, \mathbf{x}_\alpha \in L_{\text{src}}, \tag{10.9}$$

where $u(\mathbf{x}, \alpha)$ is the solution to (10.5)–(10.6).

Physically, to reconstruct the dielectric function c of objects in Ω, one sends the incident wave field from the source $\mathbf{x}_\alpha$. This wave scatters when hitting the objects. Then one measures the backscattering wave on the square Γ at a single frequency. And the data (10.9) are used to reconstruct the unknown dielectric constant inside the cube Ω.

Uniqueness of this CIP is a long standing open problem. Currently uniqueness for multidimensional CIPs with non overdetermined data can be proven only by the BK method and only if the right-hand side of equation (10.5) is not vanishing in $\overline{\Omega}$; see Chapter 3 for a variety of uniqueness results. Nevertheless, uniqueness within the framework of our approximate mathematical model (Remarks 10.3.1) follows immediately from Theorem 10.3.2.

Remarks 10.2.1.

1. We are not interested here in a specification of smoothness condition imposed on the function $c(\mathbf{x})$. Thus, $c(\mathbf{x})$ is supposed to be sufficiently smooth with respect to $\mathbf{x}$; also see Remark 8.2.1. Some particular discussions concerning this matter can be found in, for example, [142] and references therein, where the smoothness of $c(\mathbf{x})$ is essential for the asymptotic behavior of the solution u to the forward problem (10.5)–(10.6). We also note that in studies of CIPs the smoothness conditions are usually not of a considerable concern; see, for example, [224, Theorem 4.1].
2. We solve the forward problem (10.5)–(10.6) using the integral equation (10.8) for all $x \in \Omega$. In doing so, we rely on numerical methods commenced in [246]. This way enables us to extract information of $u(\mathbf{x}, \alpha)|_\Gamma$, and by repeating this process for each $\alpha \in [-a, a]$ we obtain computationally the simulated data (10.9).

10.3 An auxiliary system of coupled quasilinear elliptic equations

10.3.1 An equation without the unknown coefficient

Observe that since L_{src} is located outside of $\overline{\Omega}$, then the point source $\mathbf{x}_\alpha = (\alpha, 0, -d)$ is not in $\overline{\Omega}$. Hence, (10.5)–(10.6) imply that for each $\alpha \in [a_1, a_2]$,

$$\Delta u + k^2 c(\mathbf{x}) u = 0 \quad \text{in } \Omega. \tag{10.10}$$

We now define the function $\log u(\mathbf{x}, \alpha)$. The conformal Riemannian metric generated by the function $c(\mathbf{x})$ is

$$d\tau = \sqrt{c(\mathbf{x})}|d\mathbf{x}|, \quad |d\mathbf{x}| = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}.$$

Following [159, 160], we assume that geodesic lines generated by this metric and originated at sources $\mathbf{x}_\alpha \in L_{\mathrm{src}}$ are regular. In other words, for each point $\mathbf{x} \in \mathbb{R}^3$ and for each point $\mathbf{x}_\alpha \in L_{\mathrm{src}}$ there exists unique geodesic line $\Gamma(\mathbf{x}, \mathbf{x}_\alpha)$ connecting them. See (8.10) and [225] for a sufficient condition of the regularity of geodesic lines. The travel time along $\Gamma(\mathbf{x}, \mathbf{x}_\alpha)$ is

$$\tau(\mathbf{x}, \mathbf{x}_\alpha) = \int\limits_{\Gamma(\mathbf{x}, \mathbf{x}_\alpha)} \sqrt{c(\xi)}\, ds.$$

Tentatively, we denote $u = u(\mathbf{x}, k, \alpha)$. It was established in [159] that, under certain conditions imposed on $c(\mathbf{x})$, which we do not discuss here (Remark 10.2.1), the asymptotic behavior of the function $u(\mathbf{x}, k, \alpha)$ at $k \to \infty$ is

$$u(\mathbf{x}, k, \alpha) = A(\mathbf{x}, \alpha) \exp\big(ik\tau(\mathbf{x}, \mathbf{x}_\alpha)\big)\big[1 + \mathcal{O}(1/k)\big], \quad \forall (\mathbf{x}, \alpha) \in \overline{\Omega} \times [a_1, a_2], \tag{10.11}$$

where the function $A(\mathbf{x}, \alpha) > 0$. Let $\overline{k} > 1$ be a number. Assuming that $\overline{k}$ is sufficiently large, and that $k \geq \overline{k}$, we obtain from (10.11) that $u(\mathbf{x}, k, \alpha) \neq 0$ for all $(\mathbf{x}, k, \alpha) \in \overline{\Omega} \times [\overline{k}, \infty) \times [a_1, a_2]$. Denoting the term $\mathcal{O}(1/k)$ in (10.11) as $\mathcal{O}(1/k) = s(\mathbf{x}, k, \alpha)$, we naturally assume that $|s(\mathbf{x}, k, \alpha)| < 1$. Hence, using (10.11), we uniquely define the function $\log u(\mathbf{x}, k, \alpha)$ as

$$\log u(\mathbf{x}, k, \alpha) = ik\tau(\mathbf{x}, \mathbf{x}_\alpha) + \ln A(\mathbf{x}, \alpha) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} s^n(\mathbf{x}, k, \alpha), \tag{10.12}$$

for all $(\mathbf{x}, k, \alpha) \in \overline{\Omega} \times [\overline{k}, \infty) \times [-a, a]$. In (10.12), the series is obviously taken from the power series expansion of the function $\log(1 + s(\mathbf{x}, k, \alpha))$.

Now, suppose that k is not so large, but still $k \geq \underline{k} > 0$ where the number $\underline{k}$ is such that

$$u(\mathbf{x}, k, \alpha) \neq 0 \quad \text{for all } (\mathbf{x}, k, \alpha) \in \overline{\Omega} \times [\underline{k}, \infty) \times [a_1, a_2]. \tag{10.13}$$

Then, using an idea of [157], we define the function $\varphi(\mathbf{x},k,\alpha)$ as

$$\varphi(\mathbf{x},k,\alpha) = -\int_k^{\overline{k}} \frac{\partial_k u(\mathbf{x},\eta,\alpha)}{u(\mathbf{x},\eta,\alpha)} d\eta + \log u(\mathbf{x},\overline{k},\alpha), \quad \forall(\mathbf{x},k,\alpha) \in \overline{\Omega}\times[\underline{k},\infty)\times[a_1,a_2], \tag{10.14}$$

where $\log u(\mathbf{x},\overline{k},\alpha)$ is defined in (10.12). Hence, $(\partial_k u - u\partial_k\varphi)(\mathbf{x},k,\alpha) = 0$. Multiplying both sides of the latter by $e^{-\varphi}$, we obtain $\partial_k(ue^{-\varphi})(\mathbf{x},k,\alpha) = 0$. Since $\varphi(\mathbf{x},\overline{k},\alpha) = u(\mathbf{x},\overline{k},\alpha)$, then $u(\mathbf{x},k,\alpha) = \exp(\varphi(\mathbf{x},k,\alpha))$. This uniquely defines the function $\log u(\mathbf{x},k,\alpha) = \varphi(\mathbf{x},k,\alpha)$ as long as (10.13) holds. Finally, we note that we use below only derivatives of the function $u(\mathbf{x},k,\alpha)$, which means that we do not use $\log u$ "directly."

In all of our above cited previous publications about numerical methods for CIPs for the Helmholtz equation, we have not observed numerically such values of the function $|u(\mathbf{x},k,\alpha)|$, which would be close to zero. The same is true for the current paper. Thus, we assume below that the fixed number k we work with is such $k \in [\underline{k},\infty)$. Hence, by (10.13)–(10.14), the function $\log u(\mathbf{x},k,\alpha) = \varphi(\mathbf{x},k,\alpha)$ is uniquely defined. Thus, we assume below that

$$u(\mathbf{x},\alpha) \neq 0 \quad \text{for all } (\mathbf{x},\alpha) \in \overline{\Omega}\times[a_1,a_2].$$

We set

$$\log u_0(\mathbf{x},\alpha) = \mathrm{i}k|\mathbf{x}-\mathbf{x}_\alpha| - \ln(4\pi|\mathbf{x}-\mathbf{x}_\alpha|). \tag{10.15}$$

Denote $v_0(\mathbf{x},\alpha) = u(\mathbf{x},\alpha)/u_0(\mathbf{x},\alpha)$ and define the function $v(\mathbf{x},\alpha)$ as

$$v(\mathbf{x},\alpha) = \log v_0(\mathbf{x},\alpha) = \log u(\mathbf{x},\alpha) - \log u_0(\mathbf{x},\alpha) \quad \text{for } \mathbf{x}\in\Omega, \alpha\in[a_1,a_2]. \tag{10.16}$$

Obviously,

$$\nabla v(\mathbf{x},\alpha) = \frac{\nabla v_0(\mathbf{x},\alpha)}{v_0(\mathbf{x},\alpha)}, \quad \Delta v(\mathbf{x},\alpha) = \frac{\Delta v_0(\mathbf{x},\alpha)}{v_0(\mathbf{x},\alpha)} - \left(\frac{\nabla v_0(\mathbf{x},\alpha)}{v_0(\mathbf{x},\alpha)}\right)^2. \tag{10.17}$$

Using (10.17), we obtain the equation for v:

$$\Delta v + (\nabla v)^2 + 2\nabla v\cdot\nabla(\log u_0(\mathbf{x},\alpha)) = -k^2(c(\mathbf{x},k)-1), \quad \mathbf{x}\in\Omega. \tag{10.18}$$

Differentiating (10.18) with respect to α, we obtain

$$\Delta\partial_\alpha v + 2\nabla v\cdot\nabla\partial_\alpha v + 2\nabla\partial_\alpha v\cdot\tilde{\mathbf{x}}_\alpha + 2\widehat{\mathbf{x}}_\alpha\cdot\nabla v = 0 \quad \text{for all } \mathbf{x}\in\Omega. \tag{10.19}$$

Recall that $\mathbf{x}-\mathbf{x}_\alpha = (x-\alpha, y, z+d)$. We have the following notation in (10.19):

$$\begin{aligned} \tilde{\mathbf{x}}_\alpha &= \frac{\mathrm{i}k(\mathbf{x}-\mathbf{x}_\alpha)}{|\mathbf{x}-\mathbf{x}_\alpha|} - \frac{\mathbf{x}-\mathbf{x}_\alpha}{|\mathbf{x}-\mathbf{x}_\alpha|^2}, \\ \widehat{\mathbf{x}}_\alpha &= \frac{\mathrm{i}k}{|\mathbf{x}-\mathbf{x}_\alpha|^3}\left(-y^2-(z+d)^2, (x-\alpha)y, (x-\alpha)z\right) \\ &\quad - \frac{1}{|\mathbf{x}-\mathbf{x}_\alpha|^4}\left((x-\alpha)^2-y^2-(z+d)^2, 2(x-\alpha)y, 2(x-\alpha)z\right). \end{aligned} \tag{10.20}$$

The notion behind this differentiation is to get rid of the α-independent dielectric function c in (10.18), and thus, the auxiliary equation depends only on v and $\partial_\alpha v$ is presented in (10.19). This approach is actually very similar with the first step of the BK method. To deal with the variable α in (10.19), we rely below on the orthonormal basis of Section 6.2.3 to reduce (10.19) to a system of coupled elliptic quasilinear PDEs.

Remark 10.3.1. Suppose that we have approximately computed the function $v(\mathbf{x}, \alpha)$. Then substituting this approximation in equation (10.18) and taking the average of the left-hand side with respect to α, we obtain an approximation for the unknown coefficient $c(\mathbf{x})$.

10.3.2 Truncated Fourier series

To approximately solve the auxiliary problem (10.19), we use a truncated Fourier series. To do this, we use the orthonormal basis in $L_2(a_1, a_2)$ of Section 6.2.3 denoted by $\{\psi_n(\alpha)\}_{n=0}^{\infty}$, $\alpha \in (a_1, a_2)$.

Consider the auxiliary function $v(\mathbf{x}, \alpha)$ that we have defined in (10.16). Given $N \geq 1$, our truncated Fourier series for v is written as

$$v(\mathbf{x}, \alpha) = \sum_{n=0}^{N-1} \langle v(\mathbf{x}, \cdot), \psi_n(\cdot) \rangle \psi_n(\alpha) \quad \text{for } \mathbf{x} \in \Omega, \alpha \in [a_1, a_2]. \tag{10.21}$$

Actually the sign "≈" should be used in (10.21). However, we use "=" for the further convenience of our work with our approximate mathematical model; see Remarks 10.3.1 about this model.

Remarks 10.3.1.

1. Just as in Chapters 6–8, the representation (10.21) is an approximation of the function $v(\mathbf{x}, \alpha)$ since the rest of the Fourier series is not counted here. Furthermore, we assume that the α-derivative $\partial_\alpha v(\mathbf{x}, \alpha)$ can be obtained via the term-by-term differentiation of the right-hand side of (10.21) with respect to α. Next, we suppose that the substitution of (10.21) and its α-derivative in the left-hand side of equation (10.19) give us zero in its right-hand side. In addition, we assume that the substitution of (10.21) in the left-hand side of (10.18) provides us with the exact coefficient $c(\mathbf{x})$ in its right-hand side. Finally, we impose in Section 10.3.3 the boundary condition (10.28) on $\partial\Omega \setminus \Gamma$.
2. The assumptions of item 1 form our *approximate mathematical model*. We cannot prove convergence as $N \to \infty$. Indeed, such a result is very hard to prove due to both the nonlinearity and the ill posedness of our CIP. Therefore, our goal below is to find spatially dependent Fourier coefficients $v_n(\mathbf{x}) = \langle v(\mathbf{x}, \cdot), \psi_n(\cdot) \rangle$. The number N should be chosen numerically, see Remarks 7.3 for a discussion of this issue.

3. Everywhere below we work only within the framework of this approximate mathematical model. As it was pointed out in Introduction, the fundamental underlying reason why we are accepting this model is that the original CIP is an extremely challenging one; also see Remarks 7.3 for a similar conclusion.

We now substitute (10.21) into (10.19) to get

$$\Delta\left(\sum_{n=0}^{N-1} v_n(\mathbf{x})\psi_n'(\alpha)\right) + 2\nabla\left(\sum_{n=0}^{N-1} v_n(\mathbf{x})\psi_n(\alpha)\right)\cdot\nabla\left(\sum_{n=0}^{N-1} v_n(\mathbf{x})\psi_n'(\alpha)\right)$$
$$+ 2\nabla\left(\sum_{n=0}^{N-1} v_n(\mathbf{x})\psi_n'(\alpha)\right)\cdot\tilde{\mathbf{x}}_\alpha + 2\bar{\mathbf{x}}_\alpha\cdot\nabla\left(\sum_{n=0}^{N-1} v_n(\mathbf{x})\psi_n(\alpha)\right) = 0.$$

This equation is equivalent with

$$\sum_{n=0}^{N-1}\psi_n'(\alpha)\Delta v_n(\mathbf{x}) + 2\sum_{n=0}^{N-1}\sum_{l=0}^{N-1}\psi_n(\alpha)\psi_l'(\alpha)\nabla v_n(\mathbf{x})\cdot\nabla v_l(\mathbf{x}) \tag{10.22}$$
$$+ 2\psi_n'(\alpha)\sum_{n=0}^{N-1}\nabla v_n(\mathbf{x})\cdot\tilde{\mathbf{x}}_\alpha + 2\psi_n(\alpha)\sum_{n=0}^{N-1}\bar{\mathbf{x}}_\alpha\cdot\nabla v_n(\mathbf{x}) = 0.$$

Multiply both sides of (10.22) by the function $\psi_m(\alpha)$ for $0 \le m \le N-1$ and then integrate the resulting equation with respect to α. We arrive at the following system of coupled quasi-linear elliptic equations:

$$\Delta V(\mathbf{x}) + K(\nabla V(\mathbf{x})) = 0 \quad \text{for } \mathbf{x} \in \Omega, \tag{10.23}$$
$$V(\mathbf{x}) = \varphi_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial\Omega, \tag{10.24}$$
$$V_z(\mathbf{x}) = \varphi_1(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Gamma, \tag{10.25}$$
$$K(\nabla V(\mathbf{x})) = M_N^{-1} f(\nabla V(\mathbf{x})), \tag{10.26}$$

where M_N is the matrix of Theorem 6.2.1. Recall that this matrix is invertible. Here, $\varphi_0(\mathbf{x})$ and $\varphi_1(\mathbf{x})$ are known boundary data and we explain in Section 10.7.4 how to obtain them. Above, the unknown vector function $V(\mathbf{x}) \in \mathbb{R}^N$ is given by

$$V(\mathbf{x}) = \begin{pmatrix} v_0(\mathbf{x}) & v_1(\mathbf{x}) & \cdots & \cdots & v_{N-1}(\mathbf{x}) \end{pmatrix}^T.$$

The nonlinearity $f = ((f_m)_{m=0}^{N-1})^T \in \mathbb{R}^N$ is quadratic with respect to the first derivatives of components of $V(\mathbf{x})$,

$$f_m(\nabla V(\mathbf{x})) = 2\sum_{n,l=0}^{N-1}\nabla v_n(\mathbf{x})\cdot\nabla v_l(\mathbf{x})\int_{-a}^{a}\psi_m(\alpha)\psi_n(\alpha)\psi_l'(\alpha)d\alpha$$
$$+ 2\sum_{n=0}^{N-1}\int_{-a}^{a}\psi_m(\alpha)\psi_n'(\alpha)\nabla v_n(\mathbf{x})\cdot\tilde{\mathbf{x}}_\alpha d\alpha \tag{10.27}$$
$$+ 2\sum_{n=0}^{N-1}\int_{-a}^{a}\psi_m(\alpha)\psi_n(\alpha)\bar{\mathbf{x}}_\alpha\cdot\nabla v_n(\mathbf{x})d\alpha.$$

It follows from (10.26) and (10.27) that the vector function $K(\nabla V)$ is quadratic with respect to components of ∇V.

The problems (10.23)–(10.25) are overdetermined since we have two boundary conditions (10.24), (10.25) instead of just one. Also, this is not a regular Cauchy problem for the system (10.23) since the Dirichlet data in (10.24) are given at the entire boundary $\partial\Omega$. If solving problems (10.23)–(10.25), then we would find the dielectric constant $c(\mathbf{x})$ via backwards calculations. Therefore, we focus below on the solution of problems (10.23)–(10.25).

10.3.3 Boundary data (10.24), (10.25)

We now explain how to find the boundary data for the vector function $V(\mathbf{x})$ in (10.24), (10.25). It follows from (10.9) and (10.21) that the Dirichlet data at $\mathbf{x} \in \Gamma$ for $V(\mathbf{x})$ are known. As it is known, several data completion methods are heuristically applied in inverse problems with incomplete data; see, for example, [204]. To complement the lack of the boundary data information on $\partial\Omega\setminus\Gamma$, we use the data completion for (10.10). More precisely, we set for each α:

$$u(\mathbf{x},\alpha)|_{\partial\Omega} = \begin{cases} F(\mathbf{x},\mathbf{x}_\alpha), & \text{if } \mathbf{x} \in \Gamma, \\ u_0(\mathbf{x},\alpha), & \text{if } \mathbf{x} \in \partial\Omega\setminus\Gamma, \end{cases} \tag{10.28}$$

where the $u_0(\mathbf{x},\alpha)$ is given in (10.7) and it is the solution of problem (10.5)–(10.6) for the case of the uniform background.

As to the Neumann data (10.25), usually measurements are performed far from the domain of interest, that is, on the plane $\{z = -D\}$, where $D > b$. It is time consuming to solve a CIP in a large domain. Besides, the data at the measurement plane are hard to use for an inversion algorithm since they do not look “nice,” see Figures 10.2(a)–10.6(a). To “move” the data closer to the target’s side, the so-called “data propagation” procedure can be applied to the measured data; see [204] for a detailed description of this procedure as well as Section 10.7.4. By this procedure, one obtains “propagated data,” for example, an approximation of the data at our desired rectangle $\Gamma \subset \{z = -R\}$. Besides, the propagated data look much better than the original data, for example, compare Figure 10.2(a) with Figure 10.2(b)) below. In addition, it is clear from the data propagation procedure that one of its outcomes is an approximation of the z-derivative of the function $u(\mathbf{x},k)$ at Γ. Thus, we assume that, in addition to the Dirichlet data at Γ, we know the Neumann boundary condition $u_z(\mathbf{x},\alpha) = G(\mathbf{x},\alpha)$ for $\mathbf{x} \in \Gamma$, $\mathbf{x}_\alpha \in L_{\text{src}}$. Having the function $G(\mathbf{x},\alpha)$ and using (10.21), one can easily find the Neumann boundary condition $\varphi_1(\mathbf{x})$ at $\mathbf{x} \in \Gamma$ in (10.25).

10.3.4 Lipschitz stability of the boundary value problem (10.23)–(10.25)

For any Banach space B considered below and any integer $X > 1$ we consider the Banach space $B_X = \underbrace{B \times B \times \cdots \times B}_{X \text{ times}}$ with the norm

$$\|g\|_{B_X} = \left(\sum_{j=1}^{X} \|g_j\|_B^2 \right)^{1/2} \quad \text{for all } g = (g_1, \dots, g_X) \in B_X.$$

Let the number $r > R$ and the number $\lambda > 0$. As we have stated above, a rather simple CWF is better to use in computations than a complicated CWF of (2.66). Thus, we choose the CWF as

$$\mu_\lambda(z) = \exp\left[2\lambda(z-\theta)^2\right], \quad z \in [-b, b]. \tag{10.29}$$

We choose $\theta > b$ since one of conditions imposed on the CWF in any Carleman estimate is that its gradient should not vanish in the closed domain of ones interest. Obviously, the function $\mu_\lambda(z)$ is decreasing for $z \in (-b, b)$ and

$$\max_{\overline{\Omega}} \mu_\lambda(z) = \exp\left[2\lambda(b+\theta)^2\right], \quad \min_{\overline{\Omega}} \mu_\lambda(z) = \exp\left[2\lambda(b-\theta)^2\right]. \tag{10.30}$$

In other words, by (10.1) and (10.2), the CWF (10.29) attains its maximal value in $\overline{\Omega}$ on the part Γ of the boundary where measurements are conducted, and it attains its minimal value on the opposite side.

Define the subspace $H_0^2(\Omega)$ of the space $H^2(\Omega)$ as

$$H_0^2(\Omega) := \left\{ v \in H^2(\Omega) : v|_{\partial\Omega} = 0, \partial_z v|_\Gamma = 0 \right\}. \tag{10.31}$$

Theorem 10.3.1 easily follows from Theorem 8.4.1.

Theorem 10.3.1 (Carleman estimate). *Let* $\mu_\lambda(z)$ *be the function defined in* (10.29). *Then there exist constants* $\lambda_0 = \lambda_0(\Omega, r) \geq 1$ *and* $C = C(\Omega, r) > 0$ *depending only on the domain* Ω *such that for every function* $u \in H_0^2(\Omega)$ *and for all* $\lambda \geq \lambda_0$ *the following Carleman estimate holds:*

$$\int_\Omega |\Delta u|^2 \mu_\lambda(z) d\mathbf{x} \geq \frac{C}{\lambda} \sum_{i,j=1}^{3} \int_\Omega |u_{x_i x_j}|^2 \mu_\lambda(z) d\mathbf{x} + C\lambda \int_\Omega \left[|\nabla u|^2 + \lambda^2 |u|^2\right] \mu_\lambda(z) d\mathbf{x}. \tag{10.32}$$

Suppose that there exist two vector functions $V^{(1)}(\mathbf{x})$ and $V^{(2)}(\mathbf{x})$ satisfying equation (10.23) with boundary conditions as in (10.24), (10.25),

$$V^{(1)}(\mathbf{x}) = \varphi_0^{(1)}(\mathbf{x}), \quad V^{(2)}(\mathbf{x}) = \varphi_0^{(2)}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial\Omega, \tag{10.33}$$

$$V_z^{(1)}(\mathbf{x}) = \varphi_1^{(1)}(\mathbf{x}), \quad V_z^{(2)}(\mathbf{x}) = \varphi_1^{(2)}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Gamma. \tag{10.34}$$

Suppose that there exist two vector functions $F_1, F_2 \in H_N^3(\Omega)$ satisfying boundary conditions (10.33), (10.34), that is,

$$F_1(\mathbf{x}) = \varphi_0^{(1)}(\mathbf{x}), \quad F_2(\mathbf{x}) = \varphi_0^{(2)}(\mathbf{x}), \quad \text{for } \mathbf{x} \in \partial\Omega, \tag{10.35}$$

$$\partial_z F_1(\mathbf{x}) = \varphi_1^{(1)}(\mathbf{x}), \quad \partial_z F_2(\mathbf{x}) = \varphi_1^{(2)}(\mathbf{x}), \quad \text{for } \mathbf{x} \in \Gamma. \tag{10.36}$$

Let $M > 0$ be a number. We assume that

$$V^{(1)}, V^{(2)}, F_1, F_2 \in G(M) = \{W \in H_N^3(\Omega) : \|W\|_{H_N^3(\Omega)} < M\}. \tag{10.37}$$

Note that by the embedding theorem

$$G(M) \subset C_N^1(\overline{\Omega}) \quad \text{and} \quad \|W\|_{C_N^1(\overline{\Omega})} \leq C_1 \quad \text{for all } W \in G(M). \tag{10.38}$$

Here and below, $C_1 = C_1(\Omega, N, M) > 0$ denotes different constants depending only on listed parameters.

Theorem 10.3.2 (Lipschitz stability estimate for problem (10.23)–(10.25)). *Let* $V^{(1)}(\mathbf{x})$ *and* $V^{(2)}(\mathbf{x})$ *be two solutions of equation* (10.23) *with boundary conditions* (10.33), (10.34). *Suppose that there exist two vector functions* $F_1, F_2 \in H_N^3(\Omega)$ *satisfying* (10.35), (10.36). *Also, let* (10.37) *holds. Then the following Lipschitz stability estimate is valid:*

$$\|V^{(1)} - V^{(2)}\|_{H_N^2(\Omega)} \leq C_1 \|F_1 - F_2\|_{H_N^2(\Omega)}. \tag{10.39}$$

Proof. Denote

$$Q_1(\mathbf{x}) = V^{(1)}(\mathbf{x}) - F_1(\mathbf{x}), \quad Q_2(\mathbf{x}) = V^{(2)}(\mathbf{x}) - F_2(\mathbf{x}), \tag{10.40}$$

$$\widetilde{Q}(\mathbf{x}) = Q_1(\mathbf{x}) - Q_2(\mathbf{x}), \quad \widetilde{F}(\mathbf{x}) = F_1(\mathbf{x}) - F_2(\mathbf{x}). \tag{10.41}$$

Then (10.27) and (10.33)–(10.41) imply that

$$\Delta\widetilde{Q}(\mathbf{x}) = T_1(\mathbf{x}) \cdot \nabla\widetilde{Q}(\mathbf{x}) + T_2(\mathbf{x}) \cdot \nabla\widetilde{F}(\mathbf{x}) - \Delta\widetilde{F}(\mathbf{x}), \tag{10.42}$$

$$\widetilde{Q}|_{\partial\Omega} = 0, \quad \widetilde{Q}_z|_{\Gamma} = 0, \tag{10.43}$$

$$T_1, T_2 \in C_N(\overline{\Omega}), \quad \|T_1\|_{C_N(\overline{\Omega})}, \|T_2\|_{C_N(\overline{\Omega})} \leq C_1. \tag{10.44}$$

Square absolute values of both sides of equation (10.42). Next, multiply the resulting equation by the CWF (10.29) and integrate over the domain Ω. Using (10.44), we obtain

$$\int_\Omega |\Delta\widetilde{Q}|^2 \mu_\lambda(z) d\mathbf{x} \leq C_1 \int_\Omega |\nabla\widetilde{Q}|^2 \mu_\lambda(z) d\mathbf{x} + C_1 \int_\Omega (|\Delta\widetilde{F}|^2 + |\nabla\widetilde{F}|^2) \mu_\lambda(z) d\mathbf{x}. \tag{10.45}$$

Taking into account (10.31) and (10.43) and also applying (10.32) to (10.45), we obtain for all $\lambda \geq \lambda_0 > 1$,

$$C_1 \int_\Omega (|\Delta\widetilde{F}|^2 + |\nabla\widetilde{F}|^2) \mu_\lambda(z) d\mathbf{x} + C_1 \int_\Omega |\nabla\widetilde{Q}|^2 \mu_\lambda(z) d\mathbf{x} \tag{10.46}$$
$$\geq \frac{1}{\lambda} \sum_{i,j=1}^{3} \int_\Omega |\widetilde{Q}_{x_i x_j}|^2 \mu_\lambda(z) d\mathbf{x} + \lambda \int_\Omega [|\nabla\widetilde{Q}|^2 + \lambda^2 |\widetilde{Q}|^2] \mu_\lambda(z) d\mathbf{x}.$$

Choose a number $\lambda_1 \geq \lambda_0$ such that $\lambda_1 \geq 2C_1$. Then (10.46) implies that

$$C_1 \int_\Omega (|\Delta\widetilde{F}|^2 + |\nabla\widetilde{F}|^2)\mu_{\lambda_1}(z)d\mathbf{x}$$
$$\geq \frac{1}{\lambda_1}\sum_{i,j=1}^{3}\int_\Omega |\widetilde{Q}_{x_ix_j}|^2\mu_{\lambda_1}(z)d\mathbf{x} + \frac{\lambda_1}{2}\int_\Omega [|\nabla\widetilde{Q}|^2 + |\widetilde{Q}|^2]\mu_{\lambda_1}(z)d\mathbf{x}.$$

This inequality and (10.30) lead to

$$C_1 \exp(4Rr\lambda_1)\int_\Omega (|\Delta\widetilde{F}|^2 + |\nabla\widetilde{F}|^2)d\mathbf{x}$$
$$\geq \frac{1}{\lambda_1}\sum_{i,j=1}^{3}\int_\Omega |\widetilde{Q}_{x_ix_j}|^2 d\mathbf{x} + \frac{\lambda_1}{2}\int_\Omega [|\nabla\widetilde{Q}|^2 + |\widetilde{Q}|^2]d\mathbf{x}.$$

Hence, with a new constant C_1 we have

$$\|\widetilde{Q}\|_{H^2_N(\Omega)} \leq C_1\|\widetilde{F}\|_{H^2_N(\Omega)}. \tag{10.47}$$

Next, by (10.40), (10.41), and triangle inequality,

$$\|\widetilde{Q}\|_{H^2_N(\Omega)} = \left\|(V^{(1)} - F_1) - (V^{(2)} - F_2)\right\|_{H^2_N(\Omega)}$$
$$\geq \left\|V^{(1)} - V^{(2)}\right\|_{H^2_N(\Omega)} - \|F_1 - F_2\|_{H^2_N(\Omega)}.$$

Combining this with (10.47), we obtain the target estimate (10.39) of this theorem. □

10.4 Weighted Tikhonov-like functional

For the convenience of the presentation, each N-D complex valued vector function $W = \operatorname{Re} W + i\operatorname{Im} W$ is considered below as the $2N$-D vector function with real valued components $(\operatorname{Re} W, \operatorname{Im} W) := (W_1, W_2) := W \in \mathbb{R}^{2N}$. All results and proofs below are for these $2N$-D vector functions. For any number $s \in \mathbb{C}$, its complex conjugate is denoted as $\bar{s}$.

We find an approximate solution of the problem (10.23)–(10.26) via the minimization of an appropriate weighted Tikhonov-like functional with the CWF (10.29) involved in it. Due to (10.23), denote

$$L(V)(\mathbf{x}) = \Delta V(\mathbf{x}) + K(\nabla V(\mathbf{x})). \tag{10.48}$$

Let $\gamma \in (0,1)$ be the regularization parameter. We now consider the following weighted Tikhonov-like functional $J_{\lambda,\gamma} : H^3_{2N}(\Omega) \to \mathbb{R}_+$:

$$J_{\lambda,\gamma}(V) = \exp[-2\lambda(b+\theta)^2]\int_\Omega \left|L(V)\right|^2\mu_\lambda(z)d\mathbf{x} + \gamma\|V\|^2_{H^3_N(\Omega)}. \tag{10.49}$$

Here, $\exp[-2\lambda(b+\theta)^2]$ is the balancing multiplier: to balance first and second terms in the right-hand side of (10.49); see (10.30). We use the $H_N^3(\Omega)$-norm in the regularization term here since $H_N^3(\Omega) \subset C_N^1(\overline{\Omega})$ and an obvious analog of (10.38) holds.

Remark 10.4.1 (Underlying reasons of the convexification idea). Assuming for a moment that the nonlinear term $K(\nabla V(\mathbf{x}))$ is absent in (10.48), we remark that since the Laplace operator is linear, then one can also find an approximate solution of the problem (10.23)–(10.26) by the regular quasi-reversibility method with $\lambda = 0$ in (10.49); see Section 4.3. However, if $K(\nabla V(\mathbf{x})) \neq 0$, then the presence of the CWF serves three purposes: first, it controls this nonlinear term; second, it "maximizes" the influence of the important boundary data at $z = -R$; and third, it "convexifies" the cost functional globally. These are the underlying reasons of the convexification idea.

Below $(\cdot,\cdot)$ is the scalar product in the space $H_{2N}^3(\Omega)$. Let $M > 0$ be an arbitrary number. We define the set $B(M) \subset H_{2N}^3(\Omega)$ as

$$B(M) = \{V \in H_{2N}^3(\Omega) : \|V\|_{H_{2N}^3(\Omega)} < M, V|_{\partial\Omega} = \varphi_0, V_z|_\Gamma = \varphi_1\}. \tag{10.50}$$

By (10.38), we know that

$$B(M) \subset C_{2N}^1(\overline{\Omega}) \quad \text{and} \quad \|V\|_{C_{2N}^1(\overline{\Omega})} \leq C_1 \quad \text{for all } V \in B(M). \tag{10.51}$$

Minimization problem (MP). Minimize the cost functional $J_{\lambda,\gamma}(V)$ on the set $\overline{B(M)}$.

10.5 Analysis of the functional $J_{\lambda,\gamma}(V)$

10.5.1 Strict convexity on $\overline{B(M)}$

Theorem 10.5.1 is the central analytical result of this chapter. Note that in the proof of this theorem we do not "subtract" boundary conditions from the vector function V, which means that we do not arrange zero boundary conditions for the difference. Hence, we do not require here that our boundary conditions should be extended in the entire domain Ω. This is a new element compared with our proofs of related theorems in Chapters 7 and 8. Still, we use that subtraction in Theorems 10.5.4 and 10.6.1.

Theorem 10.5.1. *The functional $J_{\lambda,\gamma}(V)$ has its Fréchet derivative $J'_{\lambda,\gamma}(V)$ at any point $V \in \overline{B(M)}$. Let $\lambda_0 > 1$ be the number of Theorem* 10.3.1. *There exists a sufficiently large number $\lambda_2 = \lambda_2(M, N, r, \Omega) \geq \lambda_0$ such that the functional $J_{\lambda,\gamma}(V)$ is strictly convex on $\overline{B(M)}$ for all $\lambda \geq \lambda_2$. More precisely, for all $\lambda \geq \lambda_2$ the following inequality holds:*

$$\begin{aligned} &J_{\lambda,\gamma}(V^{(2)}) - J_{\lambda,\gamma}(V^{(1)}) - J'_{\lambda,\gamma}(V^{(1)})(V^{(2)} - V^{(1)}) \\ &\quad \geq C_1\|V^{(2)} - V^{(1)}\|^2_{H_{2N}^2(\Omega)} + \gamma\|V^{(2)} - V^{(1)}\|^2_{H_{2N}^3(\Omega)} \quad \textit{for all } V^{(1)}, V^{(2)} \in \overline{B(M)}. \end{aligned} \tag{10.52}$$

Proof. Let $V^{(1)}, V^{(2)} \in \overline{B(M)}$ be two arbitrary points. Define $H^3_{0,2N}(\Omega) = H^3_{2N}(\Omega) \cap H^2_{0,2N}(\Omega)$; see (10.31). Denote $h = (h_1, h_2) = V^{(2)} - V^{(1)}$. Then

$$h \in \overline{B(2M)} \quad \text{and} \quad h \in H^3_{0,2N}(\Omega). \tag{10.53}$$

Obviously, $|L(V^{(2)})|^2 = |L(V^{(1)} + h)|^2$. Observe that it follows from (10.23), (10.26), and (10.27) that the vector function $K(\nabla V)$ is the sum of linear and quadratic parts with respect to the gradients $\nabla v_n(\mathbf{x})$ of the components $v_n(\mathbf{x})$ of the vector function V. Using this as well as (10.48), we obtain

$$L(V^{(1)} + h) = L(V^{(1)}) + \Delta h + K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h). \tag{10.54}$$

Here, the vector functions K_1, K_2 are continuous with respect to $\mathbf{x}$ in $\overline{\Omega}$. Also, $K_1(\mathbf{x})$ is independent on h. vector function $K_2(\mathbf{x}, \nabla h)$, it is quadratic with respect to the gradients $\nabla h_n(\mathbf{x})$ of the components of the vector function h. The latter, (10.50) and (10.51) imply that

$$\big|K_2(\mathbf{x}, \nabla h)\big| \leq C_1 |\nabla h|^2 \quad \text{for all } \mathbf{x} \in \overline{\Omega}. \tag{10.55}$$

Squaring absolute values of both sides of (10.54), we obtain

$$\begin{aligned} \big|L(V^{(1)} + h)\big|^2 &= \big|L(V^{(1)})\big|^2 + 2\,\mathrm{Re}\{\overline{L(V^{(1)})}[\Delta h + K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)]\} \\ &\quad + \big|\Delta h + K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)\big|^2. \end{aligned} \tag{10.56}$$

In (10.56), we single out the linear, with respect to h, term as well as the term $|\Delta h|^2$. We obtain

$$\begin{aligned} &\big|L(V^{(1)} + h)\big|^2 - \big|L(V^{(1)})\big|^2 \\ &\quad = 2\,\mathrm{Re}\{\overline{L(V^{(1)})}[\Delta h + K_1(\mathbf{x})\nabla h]\} + |\Delta h|^2 + 2\,\mathrm{Re}[\overline{L(V^{(1)})}K_2(\mathbf{x}, \nabla h)] \\ &\qquad + 2\,\mathrm{Re}\{\overline{\Delta h}[K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)]\} + \big|K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)\big|^2. \end{aligned} \tag{10.57}$$

In (10.57), the term $2\,\mathrm{Re}\{\overline{L(V^{(1)})}[\Delta h + K_1(\mathbf{x})\nabla h]\}$ is linear with respect to h. Thus, we obtain

$$\begin{aligned} &J_{\lambda,\gamma}(V^{(1)} + h) - J_{\lambda,\gamma}(V^{(1)}) \\ &\quad = \mathrm{Lin}(h) + \gamma \|h\|^2_{H^3_{2N}(\Omega)} \\ &\qquad + e^{-2\lambda(R+r)^2} \int_\Omega \{|\Delta h|^2 + 2\,\mathrm{Re}[\overline{\Delta h} \cdot (K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h))]\}\mu_\lambda(z)\,d\mathbf{x} \\ &\qquad + e^{-2\lambda(R+r)^2} \int_\Omega [2\,\mathrm{Re}[\overline{L(V^{(1)})}K_2(\mathbf{x}, \nabla h)] + \big|K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)\big|^2]\mu_\lambda(z)\,d\mathbf{x}, \end{aligned} \tag{10.58}$$

where the functional $\mathrm{Lin}(h) : H^3_{0,2N} \to \mathbb{R}$ is linear with respect to $h = (h_1, h_2)$. It follows from (10.57) that it is generated by the term $2\,\mathrm{Re}\{\overline{L(V^{(1)})}[\Delta h + K_1(\mathbf{x})\nabla h]\}$,

$$\begin{aligned}\mathrm{Lin}(h) &= 2\gamma(V^{(1)}, h) \\ &\quad + 2e^{-2\lambda(R+r)^2} \int_\Omega \mathrm{Re}[\Delta h + (K_1(\mathbf{x})\nabla h)(\overline{\Delta V^{(1)} + K(\nabla V^{(1)})})]\mu_\lambda(z)d\mathbf{x}.\end{aligned} \tag{10.59}$$

Besides, it follows from (10.58) and (10.59) that

$$\lim_{\|h\|_{H^3_{2N}(\Omega)} \to 0^+} \left\{ \frac{1}{\|h\|_{H^3_{2N}(\Omega)}} [J_{\lambda,\gamma}(V^{(1)} + h) - J_{\lambda,\gamma}(V^{(1)}) - \mathrm{Lin}(h)] \right\} = 0.$$

Hence, the functional $\mathrm{Lin}(h)$ is the Frechét derivative of the functional $J_{\lambda,\gamma}$ at the point $V^{(1)} \in \overline{B(M)}$. By the Riesz theorem, there exists a unique point $J'_{\lambda,\gamma}(V^{(1)})$ such that

$$J'_{\lambda,\gamma}(V^{(1)}) \in H^3_{0,2N}(\Omega) \quad \text{and} \quad \mathrm{Lin}(h) = (J'_{\lambda,\gamma}(V^{(1)}), h) \quad \text{for all } h \in H^3_{0,2N}(\Omega). \tag{10.60}$$

Thus, we can rewrite (10.58) as

$$\begin{aligned}&J_{\lambda,\gamma}(V^{(1)} + h) - J_{\lambda,\gamma}(V^{(1)}) - (J'_{\lambda,\gamma}(V^{(1)}), h) \\ &\quad = \gamma\|h\|^2_{H^3_{2N}(\Omega)} \\ &\qquad + e^{-2\lambda(R+r)^2} \int_\Omega \{|\Delta h|^2 + 2\,\mathrm{Re}[\overline{\Delta h} \cdot (K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h))]\}\mu_\lambda(z)d\mathbf{x} \\ &\qquad + e^{-2\lambda(R+r)^2} \int_\Omega [2\,\mathrm{Re}[\overline{L(V^{(1)})}K_2(\mathbf{x}, \nabla h)] + |K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h)|^2]\mu_\lambda(z)d\mathbf{x}.\end{aligned} \tag{10.61}$$

We now estimate from the below the term in the second line of (10.61). By the Cauchy–Schwarz inequality, (10.51) and (10.55) we find that

$$2|\overline{\Delta h} \cdot (K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h))| \le \frac{1}{2}|\Delta h|^2 + C_1|\nabla h|^2.$$

Therefore,

$$\begin{aligned}&\int_\Omega \{|\Delta h|^2 + 2\,\mathrm{Re}[\overline{\Delta h} \cdot (K_1(\mathbf{x})\nabla h + K_2(\mathbf{x}, \nabla h))]\}\mu_\lambda(z)d\mathbf{x} \\ &\quad \ge \int_\Omega |\Delta h|^2\mu_\lambda(z)d\mathbf{x} - \frac{1}{2}\int_\Omega |\Delta h|^2\mu_\lambda(z)d\mathbf{x} - C_1\int_\Omega |\nabla h|^2\mu_\lambda(z)d\mathbf{x} \\ &\quad = \frac{1}{2}\int_\Omega |\Delta h|^2\mu_\lambda(z)d\mathbf{x} - C_1\int_\Omega |\nabla h|^2\mu_\lambda(z)d\mathbf{x}.\end{aligned} \tag{10.62}$$

Next, using (10.55), we estimate from the below the term in the third line of (10.61),

$$e^{-2\lambda(R+r)^2}\int\limits_{\Omega}[2\,\mathrm{Re}[\overline{L(V^{(1)})}K_2(\mathbf{x},\nabla h)] + |K_1(\mathbf{x})\nabla h + K_2(\mathbf{x},\nabla h)|^2]\mu_\lambda(z)d\mathbf{x}$$
$$\geq -C_1 e^{-2\lambda(R+r)^2}\int\limits_{\Omega}|\nabla h|^2\mu_\lambda(z)d\mathbf{x}. \tag{10.63}$$

Thus, (10.61)–(10.63) imply

$$J_{\lambda,\gamma}(V^{(1)}+h) - J_{\lambda,\gamma}(V^{(1)}) - (J'_{\lambda,\gamma}(V^{(1)}),h) \tag{10.64}$$
$$\geq \frac{e^{-2\lambda(R+r)^2}}{2}\Big[\int\limits_{\Omega}|\Delta h|^2\mu_\lambda(z)d\mathbf{x} - C_1\int\limits_{\Omega}|\nabla h|^2\mu_\lambda(z)d\mathbf{x}\Big] + \gamma\|h\|^2_{H^3_{2N}(\Omega)}.$$

Now we apply the Carleman estimate (10.32) to the second line of (10.64). This use is possible due to (10.53). For brevity, we do not count the multiplier $\exp[-2\lambda(R+r)^2]$ for a while. With a constant $\widetilde{C} = \widetilde{C}(\Omega, r, N) > 0$ and a number $\widetilde{\lambda}_0 = \widetilde{\lambda}_0(\Omega, r, N) \geq \lambda_0 > 1$ depending only on listed parameters, we obtain for all $\lambda \geq \widetilde{\lambda}_0$,

$$\frac{1}{2}\int\limits_{\Omega}|\Delta h|^2\mu_\lambda(z)d\mathbf{x} - C_1\int\limits_{\Omega}|\nabla h|^2\mu_\lambda(z)d\mathbf{x} \tag{10.65}$$
$$\geq \frac{\widetilde{C}}{\lambda}\sum_{i,j=1}^{3}\int\limits_{\Omega}|h_{x_ix_j}|^2\mu_\lambda(z)d\mathbf{x}$$
$$+\widetilde{C}\lambda\int\limits_{\Omega}[|\nabla h|^2 + \lambda^2|h|^2]\mu_\lambda(z)d\mathbf{x} - C_1\int\limits_{\Omega}|\nabla h|^2\mu_\lambda(z)d\mathbf{x}.$$

Choose the number $\lambda_2 = \lambda_2(M,\Omega,r,N) \geq \widetilde{\lambda}_0 > 1$ depending only on listed parameters such that $\widetilde{C}\lambda_2 > 2C_1$. Then we obtain from (10.65),

$$\frac{1}{2}\int\limits_{\Omega}|\Delta h|^2\mu_{\lambda_2}(z)d\mathbf{x} - C_1\int\limits_{\Omega}|\nabla h|^2\mu_{\lambda_2}(z)d\mathbf{x}$$
$$\geq \frac{\widetilde{C}}{\lambda_2}\sum_{i,j=1}^{3}\int\limits_{\Omega}|h_{x_ix_j}|^2\mu_{\lambda_2}(z)d\mathbf{x} + \frac{1}{2}\widetilde{C}\lambda_2\int\limits_{\Omega}[|\nabla h|^2 + \lambda_2^2|h|^2]\mu_{\lambda_2}(z)d\mathbf{x} \tag{10.66}$$
$$\geq C_1 e^{2\lambda_2(R-r)^2}\|h\|^2_{H^2_{2N}}.$$

Hence, combining (10.64)–(10.66), we arrive at

$$J_{\lambda,\gamma}(V^{(1)}+h) - J_{\lambda,\gamma}(V^{(1)}) - (J'_{\lambda,\gamma}(V^{(1)}),h) \geq C_1\|h\|^2_{H^2_{2N}} + \gamma\|h\|^2_{H^3_{2N}(\Omega)},$$

which is equivalent to our target estimate (10.52). □

10.5.2 The minimizer of $J_{\lambda,\gamma}(V)$ on $\overline{B(M)}$

We omit proofs of Theorems 10.5.2 and 10.5.3 since these proofs are similar with proofs of Theorem 5.3.1 and Lemma 5.2.1, respectively. In Theorem 10.5.2 below, we state the Lipschitz continuity of the Fréchet derivative $J'_{\lambda,\gamma}(V)$ on $\overline{B(M)}$.

Theorem 10.5.2. *For any $\lambda > 0$ the Fréchet derivative $J'_{\lambda,\gamma}(V)$ of the functional $J_{\lambda,\gamma}(V)$ is Lipschitz continuous on the set $\overline{B(M)}$. In other words, there exists a number $D = D(\Omega, r, N, M, \lambda, \gamma) > 0$ depending only on listed parameters such that for any $V^{(1)}, V^{(2)} \in \overline{B(M)}$ the following estimate holds:*

$$\|J'_{\lambda,\gamma}(V^{(2)}) - J'_{\lambda,\gamma}(V^{(1)})\|_{H^3_{2N}(\Omega)} \leq D\|V^{(2)} - V^{(1)}\|_{H^3_{2N}(\Omega)}.$$

Theorem 10.5.3. *Let the number $\lambda_2 = \lambda_2(M, N, r, \Omega) > 1$ be the one in Theorem* 10.5.1. *Then for any $\lambda \geq \lambda_2$ and for any $\gamma > 0$ the functional $J_{\lambda,\gamma}(V)$ has a unique minimizer $V_{\min,\lambda,\gamma} \in \overline{B(M)}$ on $\overline{B(M)}$. Furthermore, the following inequality holds:*

$$(J'_{\lambda,\gamma}(V_{\min,\lambda,\gamma}), V_{\min,\lambda,\gamma} - Q) \leq 0 \quad \text{for all } Q \in \overline{B(M)}. \tag{10.67}$$

10.5.3 The distance between the minimizer and the "ideal" solution

In accordance with the concept of Tikhonov for ill-posed problems [22, 244], assume now that there exists the "ideal" solution V^* of problem (10.23)–(10.26) with the "ideal" noiseless data φ_0^*, φ_1^*. It makes sense to obtain an estimate of the distance between V^* and the minimizer $V_{\min,\lambda,\gamma}$ of the functional $J_{\lambda,\gamma}(V)$ for the case of noisy data with the noise level $\delta \in (0, 1)$. This is what is done in the current subsection.

To obtain this estimate, we need to "extend" the boundary data φ_0, φ_1 in (10.24), (10.25) inside Ω. Recall that, unlike all previous works on the convexification, we have not done this extension in the proof of our central Theorem 10.5.1. Thus, we assume that there exists a vector function $G(\mathbf{x}) \in H^3_{2N}(\Omega)$ satisfying boundary conditions (10.24), (10.25),

$$G|_{\partial\Omega} = \varphi_0(\mathbf{x}), \quad G_z|_\Gamma = \varphi_1(\mathbf{x}). \tag{10.68}$$

On the other hand, the existence of the corresponding vector function $G^*(\mathbf{x}) \in H^3_{2N}(\Omega)$ satisfying boundary conditions with the "ideal" data,

$$G^*|_{\partial\Omega} = \varphi_0^*(\mathbf{x}), \quad G_z^*|_\Gamma = \varphi_1^*(\mathbf{x}) \tag{10.69}$$

follows from the existence of the ideal solution V^*. We assume that

$$\|G - G^*\|_{H^3_{2N}(\Omega)} < \delta. \tag{10.70}$$

In addition, we suppose that

$$\|V^*\|_{H^3_{2N}(\Omega)}, \|G^*\|_{H^3_{2N}(\Omega)} < M - \delta. \tag{10.71}$$

Using (10.70), (10.71), and the triangle inequality, we easily see that

$$\|G\|_{H^3_{2N}(\Omega)} < M. \tag{10.72}$$

Our goal now is to estimate $\|V_{\min,\lambda,\gamma} - V^*\|_{H^3_{2N}(\Omega)}$ via the noise parameter δ.

Theorem 10.5.4 (accuracy and stability of minimizers).
Suppose that conditions (10.68)–(10.71) *hold. Let* $\lambda_2 = \lambda_2(M, N, r, \Omega) > 1$ *be the number in Theorems* 10.5.1, 10.5.3. *Choose the number* $\lambda_3 = \lambda_2(3M, N, \Omega) > \lambda_2 > 1$. *Let* $\lambda = \lambda_3$ *and* $\gamma = \delta^2$. *Then the following accuracy estimate holds:*

$$\|V_{\min,\lambda,\gamma} - V^*\|_{H^2_{2N}(\Omega)} \le C_1\delta. \tag{10.73}$$

Proof of Theorem 10.5.4. We note first that since the boundary conditions for vector functions $V_{\min,\lambda,\gamma}$ and V^* are different, then we cannot apply directly the strict convexity inequality (10.52) here, setting, for example, that $V^{(2)} = V^*$ and $V^{(1)} = V_{\min,\lambda,\gamma}$.

For every vector function $V \in B(M)$, consider the vector function $W = V - G$. Then by (10.72) and the triangle inequality

$$W \in B_0(2M) = \{W : \|W\|_{H^3_{2N}(\Omega)} < 2M, W|_{\partial\Omega} = W_z|_\Gamma = 0\}. \tag{10.74}$$

On the other hand, (10.72) and (10.74) imply that

$$W + G \in B(3M) \quad \text{for all } W \in B_0(2M). \tag{10.75}$$

Now, for any $W \in B_0(2M)$ we have

$$\begin{aligned} &J_{\lambda,\gamma}(W^* + G) - J_{\lambda,\gamma}(W + G) - J'_{\lambda,\gamma}(W + G)(W^* - W) \\ &\quad = J_{\lambda,\gamma}(\widetilde{V}^*) - J_{\lambda,\gamma}(V) - J'_{\lambda,\gamma}(V)(\widetilde{V}^* - V), \end{aligned} \tag{10.76}$$

where $\widetilde{V}^* = W^* + G$ and $V = W + G$. Notice that by (10.74) and (10.75) both vector functions $\widetilde{V}^*, V \in B(3M)$. Hence, by Theorem 10.5.1 we can apply the estimate (10.52) to the second line of (10.76) with $\lambda = \lambda_3 = \lambda_2(3M, N, r, \Omega) > 1$. Thus,

$$\begin{aligned} &J_{\lambda_3,\gamma}(W^* + G) - J_{\lambda_3,\gamma}(W + G) - J'_{\lambda_3,\gamma}(W + G)(W^* - W) \\ &\quad \ge C_1\|W^* - W\|^2_{H^2_{2N}(\Omega)} + \gamma\|W^* - W\|^2_{H^3_{2N}(\Omega)} \quad \text{for all } W \in B_0(2M). \end{aligned} \tag{10.77}$$

Consider now the minimizer $V_{\min,\lambda_3,\gamma} \in \overline{B(M)}$ which is claimed by Theorem 10.5.3. Let $W_{\min,\lambda_3,\gamma} = V_{\min,\lambda_3,\gamma} - G \in B(2M)$. Then (10.77) implies that

$$\begin{aligned} &J_{\lambda_3,\gamma}(W^* + G) - J_{\lambda_3,\gamma}(V_{\min,\lambda_3,\gamma}) - J'_{\lambda_3,\gamma}(V_{\min,\lambda_3,\gamma})((W^* + G) - V_{\min,\lambda_3,\gamma}) \\ &\quad \ge C_1\|W^* - W_{\min,\lambda_3,\gamma}\|^2_{H^2_{2N}(\Omega)} + \gamma\|W^* - W_{\min,\lambda_3,\gamma}\|^2_{H^3_{2N}(\Omega)}. \end{aligned} \tag{10.78}$$

Using the triangle inequality, (10.70) and (10.71), we obtain

$$\begin{aligned}\|W^* + G\|_{H^3_{2N}(\Omega)} &= \|W^* + G^* + (G - G^*)\|_{H^3_{2N}(\Omega)}\\ &\le \|W^* + G^*\|_{H^3_{2N}(\Omega)} + \|G - G^*\|_{H^3_{2N}(\Omega)}\\ &= \|V^*\|_{H^3_{2N}(\Omega)} + \|G - G^*\|_{H^3_{2N}(\Omega)} < (M - \delta) + \delta = M.\end{aligned}$$

This means that $(W^* + G) \in B(M)$. Therefore, we use (10.67) to get

$$-J'_{\lambda_3,\gamma}(V_{\min,\lambda_3,\gamma})\big((W^* + G) - V_{\min,\lambda_3,\gamma}\big) \le 0.$$

Hence,

$$\begin{aligned}&J_{\lambda_3,\gamma}(W^* + G) - J_{\lambda_3,\gamma}(V_{\min,\lambda_3,\gamma}) - J'_{\lambda_3,\gamma}(V_{\min,\lambda_3,\gamma})\big((W^* + G) - V_{\min,\lambda_3,\gamma}\big)\\ &\quad\le J_{\lambda_3,\gamma}(W^* + G).\end{aligned}$$

Moreover, substituting this inequality in (10.78), we obtain

$$J_{\lambda_3,\gamma}(W^* + G) \ge C_1\|W^* - W_{\min,\lambda_3,\gamma}\|^2_{H^2_{2N}(\Omega)}. \tag{10.79}$$

We now estimate the left-hand side of (10.79). Note that the functional $J_{\lambda_3,\gamma}(V)$ can be represented as

$$J_{\lambda_3,\gamma}(V) = J^0_{\lambda_3,\gamma}(V) + \gamma\|V\|^2_{H^3_N(\Omega)}, \tag{10.80}$$

$$J^0_{\lambda_3,\gamma}(V) = \exp[-2\lambda(R + r)^2]\int_\Omega |L(V)|^2\mu_\lambda(z)d\mathbf{x}. \tag{10.81}$$

Since $W^* + G^* = V^*$ is the ideal solution, then $L(V^*)(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega$.

Next, using the finite increment formula and (10.48), we obtain

$$\begin{aligned}|L(W^* + G)|^2(\mathbf{x}) &= |L(W^* + G^* + G - G^*)|^2(\mathbf{x})\\ &= |L(V^*) + S(G - G^*)|^2(\mathbf{x}) = |S(G - G^*)|^2(\mathbf{x}),\end{aligned}$$

where by (10.30) and (10.70) the following estimate is valid:

$$\exp[-2\lambda(R + r)^2]\int_\Omega |S(G - G^*)|^2(\mathbf{x})\mu_\lambda(z)d\mathbf{x} \le C_1\delta^2.$$

This and (10.81) imply that

$$J^0_{\lambda_3,\gamma}(W^* + G) \le C_1\delta^2. \tag{10.82}$$

Next, using (10.72), (10.74), (10.80), and (10.82), we obtain

$$J_{\lambda_3,\gamma}(W^* + G) \le C_1(\delta^2 + \gamma). \tag{10.83}$$

Therefore, using (10.79), (10.83), and recalling that $\gamma = \delta^2$, we obtain

$$\|W^* - W_{\min,\lambda_3,\gamma}\|_{H^2_{2N}(\Omega)} \le C_1\delta. \tag{10.84}$$

Finally, using (10.70) and the triangle inequality, we obtain the following lower bound for the left-hand side of (10.84):

$$\begin{aligned}\|W^* - W_{\min,\lambda_3,\gamma}\|_{H^2_{2N}(\Omega)} &= \|(W^* + G^*) - (W_{\min,\lambda_3,\gamma} + G) + (G - G^*)\|_{H^2_{2N}(\Omega)} \\ &= \|(V^* - V_{\min,\lambda_3,\gamma}) + (G - G^*)\|_{H^2_{2N}(\Omega)} \\ &\ge \|V^* - V_{\min,\lambda_3,\gamma}\|_{H^2_{2N}(\Omega)} - \|G - G^*\|_{H^2_{2N}(\Omega)} \\ &\ge \|V^* - V_{\min,\lambda_3,\gamma}\|_{H^2_{2N}(\Omega)} - \delta.\end{aligned}$$

Substituting this in (10.84), we obtain the target estimate (10.73). □

Corollary 10.5.1. *The functional $I_{\lambda,\gamma}(W) := J_{\lambda,\gamma}(W + G)$ is strictly convex on $\overline{B_0(2M)}$ for all $\lambda \ge \lambda_3$, where λ_3 is the number defined in Theorem* 10.5.4.

Proof. It follows from the proof of Theorem 10.5.4 and (10.52) that the following analog of (10.77) holds for all $\lambda \ge \lambda_3$ and for all $W^{(1)}, W^{(2)} \in B_0(2M)$:

$$\begin{aligned}&I_{\lambda,\gamma}(W^{(2)}) - I_{\lambda,\gamma}(W^{(1)}) - I'_{\lambda,\gamma}(W^{(1)})(W^{(2)} - W^{(1)}) \\ &\quad \ge C_1\|W^{(2)} - W^{(1)}\|^2_{H^2_{2N}(\Omega)} + \gamma\|W^{(2)} - W^{(1)}\|^2_{H^3_{2N}(\Omega)}.\end{aligned}$$

□

10.6 The globally convergent gradient projection method

Now we construct an approximation for the vector function $W^* = V^* - G^*$ for $W^* \in B_0(2M)$. It follows from (10.74) that $B_0(2M) \subset H^3_{0,2N}(\Omega)$. Let $P_{\overline{B}} : H^3_{0,2N}(\Omega) \to \overline{B_0(2M)}$ be the orthogonal projection operator of the space $H^3_{0,2N}(\Omega)$ on the closed ball $\overline{B_0(2M)}$. Let $W^{(0)} \in B_0(2M)$ be an arbitrary point of the ball $B_0(2M)$ and let $\eta > 0$ be a number. The gradient projection method constructs the following sequence:

$$W^{(n)} = P_{\overline{B}}(W^{(n-1)} - \eta J'_{\lambda,\gamma}(W^{(n-1)} + G)), \quad n = 1, 2, \ldots. \tag{10.85}$$

It is important for computations that $(W^{(n-1)} - \eta J'_{\lambda,\gamma}(W^{(n-1)} + G)) =: Y_{n-1}(\mathbf{x}) \in H^3_{0,2N}(\Omega)$. Indeed, $W^{(n-1)} \in H^3_{0,2N}(\Omega)$; also, (10.60) holds. In other words, the vector function $Y_{n-1}(\mathbf{x})$ satisfies the boundary conditions $Y_{n-1}|_{\partial\Omega} = \partial_z Y_{n-1}|_{\Gamma} = 0$.

Theorem 10.6.1. *Let $\lambda \ge \lambda'_2 = \lambda_2(2M, N, r, \Omega) > 1$, where λ_2 was defined in Theorem* 10.5.1. *Let $W_{\min,\lambda,\gamma}$ be the minimizer of the functional $J_{\lambda,\gamma}(W + G)$ on the set $\overline{B_0(2M)}$, the existence and uniqueness of which follow from Theorem* 10.5.3 *and Corollary* 10.5.1. *Then there exists a sufficiently small number $\eta_0 = \eta_0(2M, N, r, \Omega, \lambda) \in (0, 1)$*

depending only on listed parameters such that for any $\eta \in (0, \eta_0)$ *we can find a number* $\theta = \theta(\eta) \in (0, 1)$ *such that the sequence* $\{W^{(n)}\}_{n=0}^{\infty}$ *converges to* $W_{\min,\lambda,\gamma}$ *in the* $H^3_{2N}(\Omega)$*-norm and the following convergence estimate holds:*

$$\|W_{\min,\lambda,\gamma} - W^{(n)}\|_{H^3_{2N}(\Omega)} \leq \theta^n \|W_{\min,\lambda,\gamma} - W^{(0)}\|_{H^3_{2N}(\Omega)}, \quad n = 1, 2, \ldots. \tag{10.86}$$

Theorem 10.6.1 follows immediately from the combination of Theorems 10.5.1–10.5.3 with Theorem 5.3.1.

Theorem 10.6.2. *Let* $\lambda = \lambda_4 = \lambda_2(2M, N, r, \Omega) \geq \lambda_2'$. *Suppose that conditions imposed in Theorems* 10.5.4 *and* 10.6.1 *hold. Then the following convergence estimates are valid for* $n = 1, 2, \ldots$:

$$\|W^* - W^{(n)}\|_{H^2_{2N}(\Omega)} \leq C_1\delta + \theta^n \|W_{\min,\lambda,\gamma} - W^{(0)}\|_{H^3_{2N}(\Omega)}, \tag{10.87}$$

$$\|c^*(\mathbf{x}) - c_n(\mathbf{x})\|_{L^2(\Omega)} \leq C_1\delta + \theta^n \|W_{\min,\lambda,\gamma} - W^{(0)}\|_{H^3_{2N}(\Omega)}, \tag{10.88}$$

where $c^*(\mathbf{x})$ *stands in the right-hand side of equation* (10.17) *in the case when* $W^*(\mathbf{x})$ *is replaced with* $V^*(\mathbf{x}) = W^*(\mathbf{x}) + G^*(\mathbf{x})$. *The function* $v^*(\mathbf{x}, \alpha)$ *is obtained via components of the vector function* $V^*(\mathbf{x})$ *and* (10.21) *and then this function is substituted in the left-hand side of* (10.18)*; see the first item of Remarks* 10.3.1. *The function* $c_n(\mathbf{x})$ *is obtained in the same way with the only replacement of* $V^*(\mathbf{x})$ *with* $V^{(n)}(\mathbf{x}) = W^{(n)}(\mathbf{x}) + G(\mathbf{x})$.

Proof. Combining (10.84) with (10.86), we obtain

$$\begin{aligned}\|W^* - W^{(n)}\|_{H^2_{2N}(\Omega)} &= \|(W^* - W_{\min,\lambda,\gamma}) + (W_{\min,\lambda,\gamma} - W^{(n)})\|_{H^2_{2N}(\Omega)} \\ &\leq \|W^* - W_{\min,\lambda,\gamma}\|_{H^2_{2N}(\Omega)} + \|W_{\min,\lambda,\gamma} - W^{(n)}\|_{H^2_{2N}(\Omega)} \\ &\leq C_1\delta + \theta^n \|W_{\min,\lambda,\gamma} - W^{(0)}\|_{H^3_{2N}(\Omega)},\end{aligned}$$

which proves (10.87). As to (10.88), it follows from (10.87) and the construction of functions $c^*(\mathbf{x})$, $c_n(\mathbf{x})$ described in the formulation of Theorem 10.6.2. □

Remarks 10.6.1. Since the starting point $W^{(0)}$ of the gradient projection method (10.85) is an arbitrary point of the ball $B_0(2M)$ and since smallness conditions are not imposed on M, then convergence estimates (10.87) and (10.88) mean the global convergence of the gradient projection method (10.85) to the correct solution; see Definition 1.4.2. In other words, a good first guess about the ideal solution is no longer required. We note that in the case of a nonconvex functional, the global convergence of a gradient-like method cannot be guaranteed.

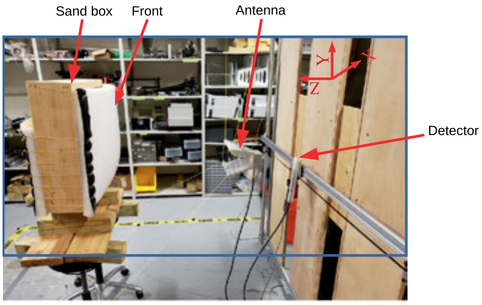
10.7 Work with experimental data

10.7.1 Experimental setup

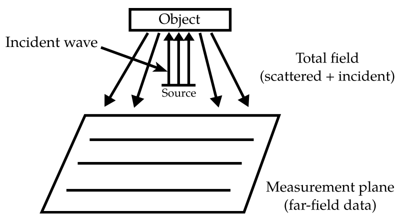
We now explain our experimental setup and data acquisition at the microwave facility of University of North Carolina at Charlotte (UNCC). Keeping in mind our target application mentioned in the first paragraph of Introduction to imaging of explosive-like devices, We have collected experimental for objects buried in a sandbox. More precisely, we have placed the targets of interest inside of a wooden framed box filled with the moisture-free sand. Besides, we cover the front and back sides of the sandbox by Styrofoam whose dielectric constant is close to 1, that is, to the dielectric constant of air. Hence, styrofoam should not affect neither the incident nor the scattered electric waves. Here, the front surface is physically defined as the foam layer closer to the transmitter fixed at a given position. On the other hand, the burial depths of objects do not exceed 10 cm, which really mimics a scanning and detecting action for shallow mine-like targets. Typically, the sizes of antipersonnel land mines and improvised explosive devices are between 5 and 15 cm; see, for example, [204]. The transmitter is a standard horn antenna, whose length is about 20 cm, and the detector is essentially a point probe. To get a better insight into the description we have detailed, the reader can take a look at Figures 10.1.

It is worth mentioning that there are several challenges we confront in this configuration, which actually reflect the difficulties met in the realistic detection of land mines. We now name some central challenges:

– *Distractions*. See Figure 10.1(a): we deliberately keep many other devices and items (made of different materials) on the desks outside the yellow caution bands. In other words, we do not use any isolations of our device from the outside world. This is reasonable since no isolation conditions can be created on a battlefield.



(a) A photograph of our experimental setup



(b) A schematic diagram of sources/detectors locations in our experimental setup

Figure 10.1: Our experimental setup (left) and a schematic diagram of our measurements (right).

Obviously, such unwanted obstacles and furniture can affect the quality of the raw backscatter signal. The presence of the Wi-Fi signal is also unavoidable in the room where we conduct the experiments. Moreover, it is technically very hard to place the antenna behind the measurement site. Therefore, the backscatter wave hits the antenna first and only then comes to detectors, which is another complicating factor.

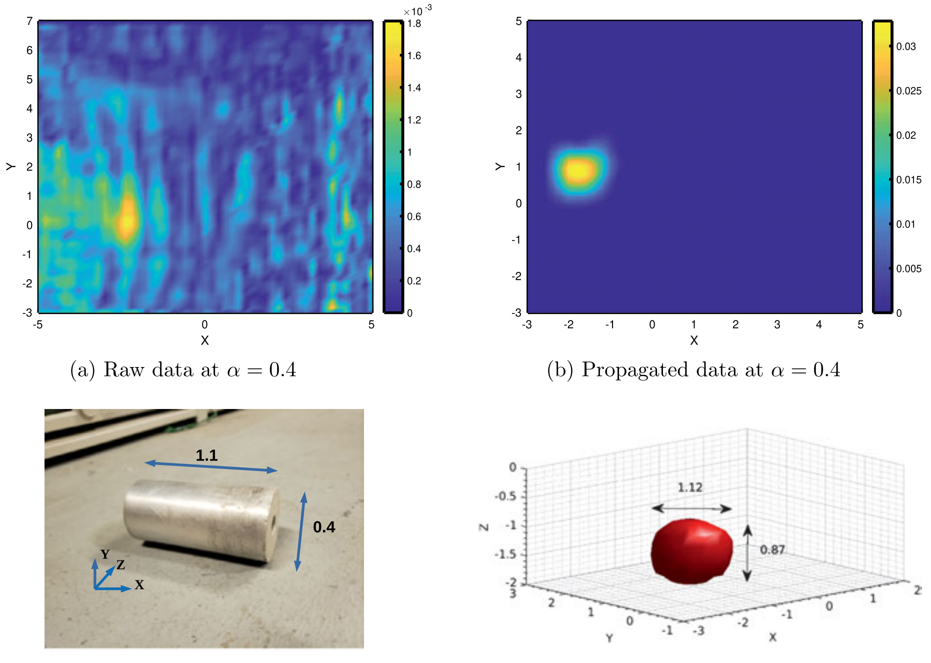
– *Random noise factor.* When facing real experiments, one cannot rarely estimate the noise level as well as its frequency dependent dynamics since they depends on hundreds of factors such as measurement process, unknown true data, distracting signals, etc.

10.7.2 Buried targets to be imaged

We present here five examples of computational reconstructions of buried objects mimicking typical metallic and nonmetallic land mines. The tested objects we use in the experiments are basic in-store items that one can easily purchase. The burial depth of any target is not of an interest here since all depths are just a few centimeters. The most valuable information for the engineering part is in estimating the values of dielectric constants of targets as well as their shapes.

Our five examples are:

– *Example 1*: An aluminum cylinder (see Figure 10.2(c)). As metallic mines usually caught in military services, this object can be shaped as the NO-MZ 2B, a Vietnamese antipersonnel fragmentation mine; cf., for example, [12]. It is known that metallic objects can be characterized by large values of dielectric constants [170]. Hence, we suppose that the true values of dielectric constants of metallic objects are large and are not fixed.
– *Example 2*: A glass bottle filled with the clear water (see Figure 10.3(c)). This object is more complicated than the one of Example 2 due to the presence of the cap on the top of the bottle. Example 2 is a good fit of the usual Glassmine 43 (cf. [217]), a nonmetallic antipersonnel land mine largely with a glass body that the Germans used to make detection harder in the World War II era. The true value of the dielectric constant in this case was measured to be 23.8 [242].
– *Example 3*: An U-shaped piece of a dry wood (see Figure 10.4(c)). This example is our next attempt to deal with a non-metallic object. Note that the shape is non convex now. In the spirit of Example 2, this wood-based object is well suited (in terms of the material) to the case of Schu-mine 42, an antipersonnel blast mine that the Germans developed during the World War II. The augmented complexity of the geometry of the object is just our purpose of this work since we wish to see how the reconstruction works with different front shapes. Given this, the maximal achievable value of the dielectric constant, which we see in [65], should be 6.



(a) Raw data at $\alpha = 0.4$

(b) Propagated data at $\alpha = 0.4$

(c) Aluminum cylinder

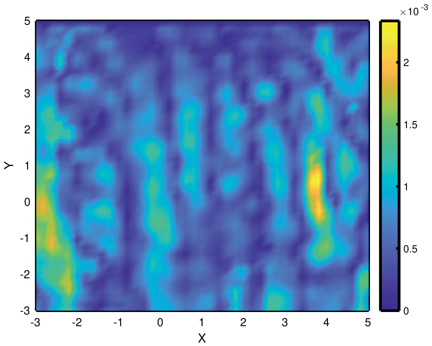
(d) Computed inclusion

Figure 10.2: Reconstruction results of Example 1 (aluminum tube). (a) Illustration of the absolute value of the raw far-field data; (b) Illustration of the absolute value of the near-field data after the data propagation procedure; (c) Photo of the experimental object; (d) The computed image of (c). All images are in the dimensionless variables.

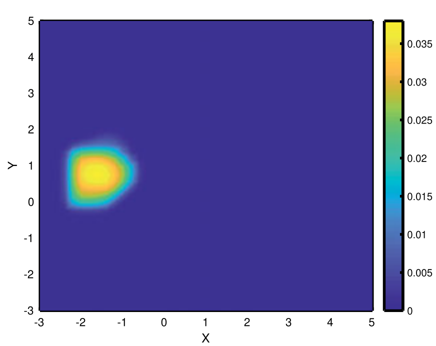
– *Blind Examples 4 and 5*: Metallic letters “A” and “O” (see Figures 10.5(c) and 10.6(c)). Shapes are nonconvex. These two tests are different from the above examples because they were *blind* tests. This means that we did not know any other information except of the measured data and the fact that these objects were buried close to the sand surface. Since they are metallic, the true contrast should be large as in Example 1.

10.7.3 The necessity of data propagation

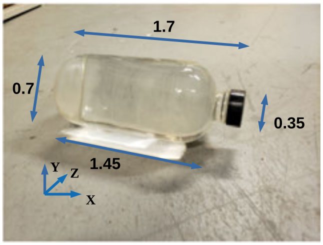
In the experimental setup, our observed and measured data are the source dependent backscattering data of the electric field. Although our experimental device measures the backscattering data with varied frequencies for each location of the point source, we use only a single frequency for each experiment when solving our CIP. Basically, these are varied far-field data; see Figures 10.1. However, these data are deficient, that is, it is unlikely that these data can be reasonably inverted; see Figures 10.2(a)–10.6(a)).



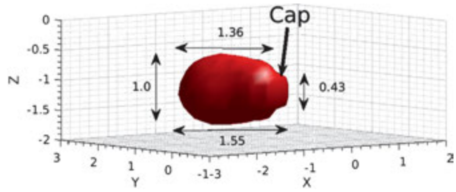
(a) Raw data at $\alpha = 0.5$



(b) Propagated data at $\alpha = 0.5$



(c) Glass bottle

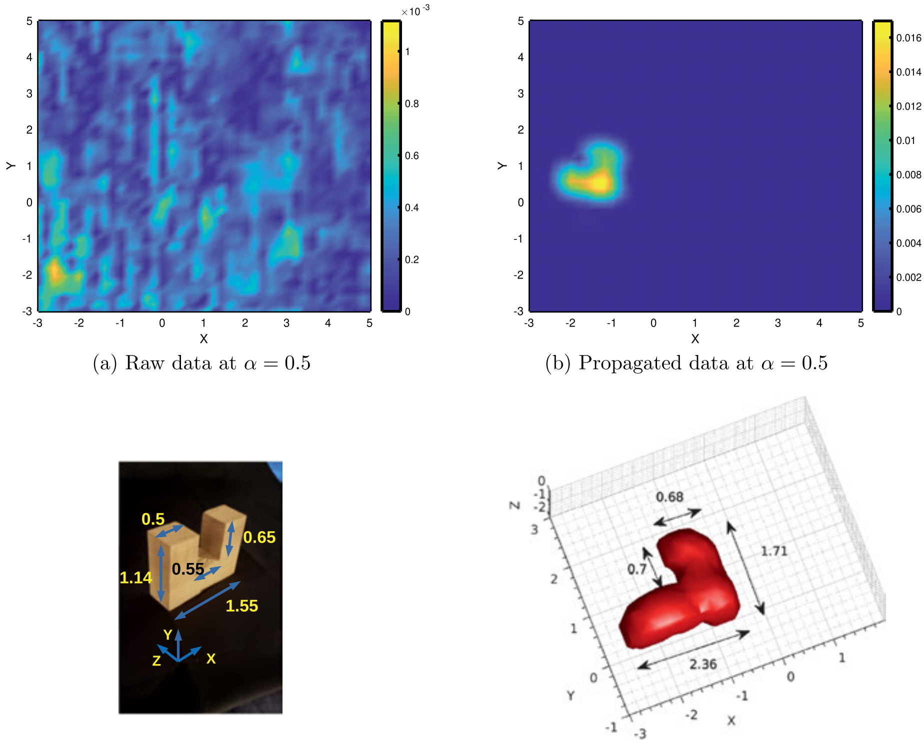


(d) Computed inclusion. Note that the cap of the bottle is clearly seen.

Figure 10.3: Reconstruction results of Example 2 (a glass bottle filled with clear water). (a) Illustration of the absolute value of the raw far-field data; (b) Illustration of the absolute value of the near-field data after the data propagation procedure; (c) Photo of the experimental object; (d) The computed image of (c). All images are in the dimensionless variables. An interesting point here is that we can even see the cap of the bottle in (d), which is challenging to image.

In fact, the same observation was made in previous publications of our research group on experimental data [143, 204, 242]. Hence, to make our data feasible for inversion, we apply the well-known data propagation procedure, which approximates the near field data. These approximate data form actual inputs of our minimization process. A rigorous justification of the data propagation procedure can be found in [204].

It is our experience that the good quality near-field data are not always obtained well enough from any far field data after the propagation. This requires a substantial workload in choosing proper data among a large amount of frequency dependent data sets. In other words, we have no choice but to select an acceptable frequency for each particular target we work with. So, for each considered target, we select its own frequency. Then we use this frequency for all positions of the source we work with. Depending on the specific target, frequencies varied between 3.16 GHz and 4.19 GHz, which corresponds to the variation of the dimensionless wavenumber k between 6.62 and 8.79 [115].



(a) Raw data at $\alpha = 0.5$

(b) Propagated data at $\alpha = 0.5$

(c) U-shaped piece of dry wood

(d) Computed inclusion

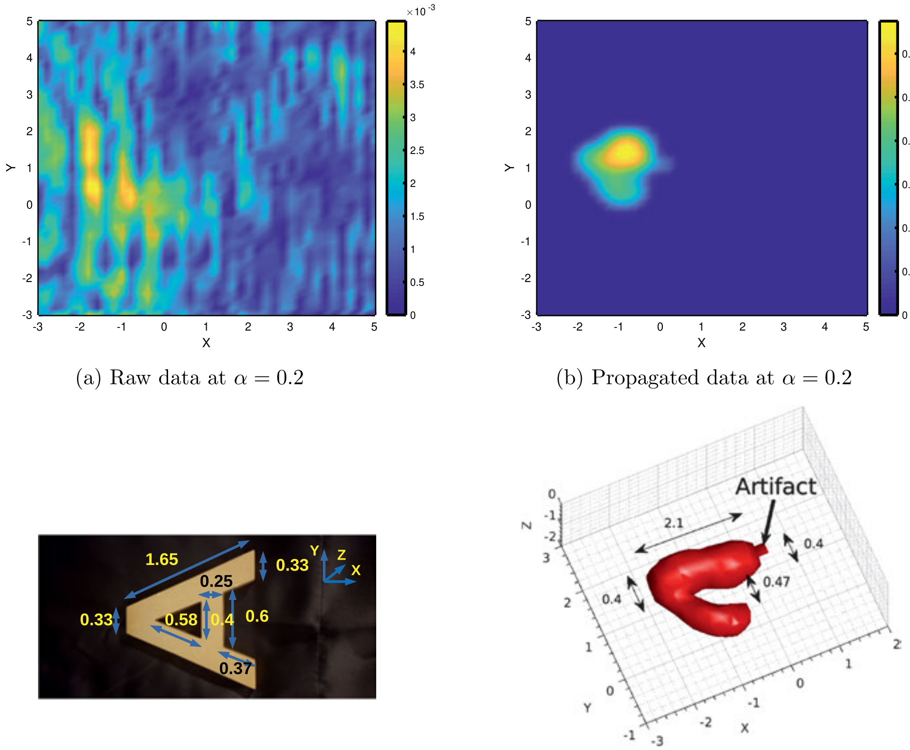
Figure 10.4: Reconstruction results of Example 3 (U-shaped piece of dry wood). Note that the shape is nonconvex, which is difficult to image. (a) Illustration of the absolute value of the raw far-field data; (b) Illustration of the absolute value of the near-field data after the data propagation procedure; (c) Photo of the experimental object; (d) The computed image of (c). Note that the void is clearly seen which is difficult to image. Our axes on (d) are oriented differently from ones on (c) due to some technical problem of the imaging software. These axes are comparable. All images are in the dimensionless variables.

10.7.4 Data propagation revisited

We know in advance that the half space $\{z < -b\} \subset \mathbb{R}^3$ is homogeneous, that is, $c(\mathbf{x}) = 1$ in this half-space. Therefore, the function u_s is a backscatter wave in $\{z < -b\}$ and it satisfies the following conditions:

$$\begin{cases} \Delta u_s + k^2 u_s = 0 & \text{for } \mathbf{x} \in \{z < -b\}, \\ \partial_r u_s - \mathrm{i} k u_s = \mathcal{O}(r^{-1}) & \text{for } r = |\mathbf{x} - \mathbf{x}_\alpha|, \mathrm{i} = \sqrt{-1}. \end{cases} \tag{10.89}$$

As was mentioned in Section 10.3.3, we actually measure the far field data, that is, we measure the function $u_s(x, y, -D, \mathbf{x}_\alpha)$, where the number $D > b$. Having the function $u_s(x, y, -D, \mathbf{x}_\alpha)$, we want to approximate the function $u_s(x, y, -b, \mathbf{x}_\alpha)$, that is, we want



(a) Raw data at $\alpha = 0.2$

(b) Propagated data at $\alpha = 0.2$

(c) Metallic letter "A". Blind test. The shape is non-convex, which is difficult to image.

(d) Computed inclusion

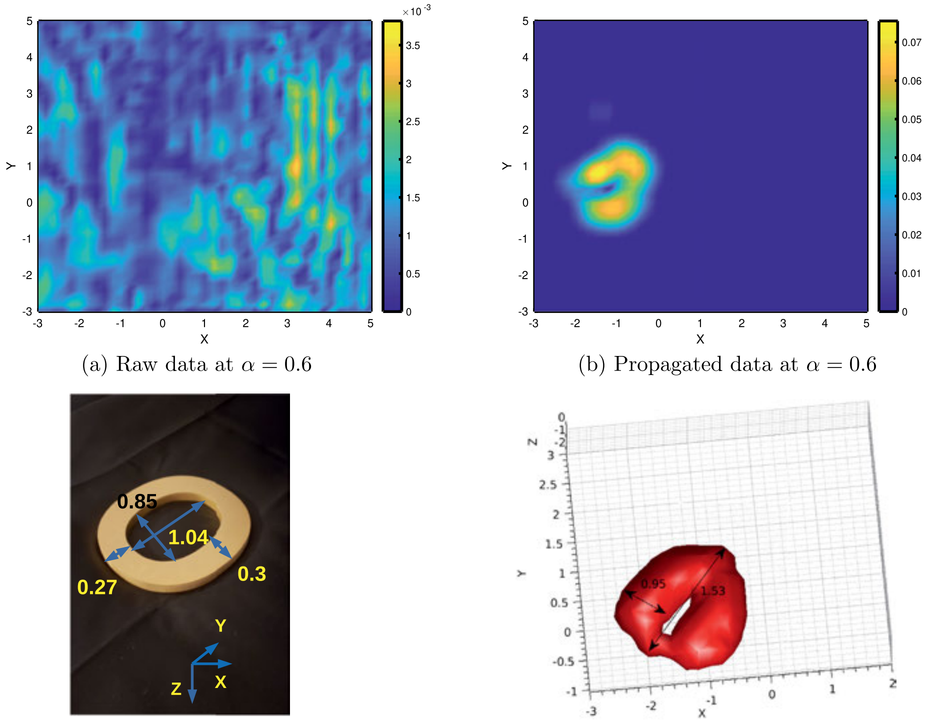
Figure 10.5: Reconstruction results of Example 4 (metallic letter "A"). This is a blind test. (a) Illustration of the absolute value of the raw far-field data; (b) Illustration of the absolute value of the near-field data after the data propagation procedure; (d) The computed image of (c). Note that the void is clearly seen, which is challenging to image. Also, sizes of the imaged target are close to the true ones. The strip of "A" is not seen since its width is 2.5 cm, which is less than the wavelength of 10.4 cm we have used with $k = 9.55$. All images are in the dimensionless variables.

to approximate the wave field in the near field zone. The data propagation procedure does exactly this. Denote

$$u_s(x,y,-b,\mathbf{x}_\alpha) = \mathbf{U}(x,y,\mathbf{x}_\alpha) \quad \text{and} \quad u_s(x,y,-D,\mathbf{x}_\alpha) = \mathbf{V}(x,y,\mathbf{x}_\alpha). \tag{10.90}$$

In this work, we rely on the data propagation procedure to unveil this difficulty as it has been successfully exploited in [204]. First, we apply the Fourier transform of the scattered field with respect to x, y, assuming that the corresponding integral converges:

$$\hat{u}_s(\rho_1,\rho_2,z,\mathbf{x}_\alpha) = \frac{1}{2\pi}\int\limits_{\mathbb{R}^2} u_s(x,y,z,\mathbf{x}_\alpha)e^{-\mathrm{i}(x\rho_1+y\rho_2)}dxdy \quad \text{for } \rho_1,\rho_2 \in \mathbb{R}. \tag{10.91}$$



(a) Raw data at $\alpha = 0.6$

(b) Propagated data at $\alpha = 0.6$

(c) Metallic letter "O". Blind test. The shape is non-convex, which is challenging to image.

(d) Computed inclusion

Figure 10.6: Reconstruction results of Example 5 (metallic letter "O"). This is a blind test. (a) Illustration of the absolute value of the raw far-field data; (b) Illustration of the absolute value of the near-field data after the data propagation procedure; (c) Photo of the experimental object; (d) The computed image of (c). Note that the void is clearly seen, which is not easy to image. All images are in the dimensionless variables.

Next, we apply this Fourier transform to the PDE in (10.89) and arrive at a second order ODE with respect to z:

$$\partial_{zz}^2 \hat{u}_s + (k^2 - \rho_1^2 - \rho_2^2)\hat{u}_s = 0 \quad \text{for } z < -b. \tag{10.92}$$

By (10.90), we also have

$$\hat{u}_s(\rho_1, \rho_2, -b, \mathbf{x}_\alpha) = \hat{\mathbf{U}}(\rho_1, \rho_2, \mathbf{x}_\alpha) \quad \text{and} \quad \hat{u}_s(\rho_1, \rho_2, -D, \mathbf{x}_\alpha) = \hat{\mathbf{V}}(\rho_1, \rho_2, \mathbf{x}_\alpha).$$

It follows from (10.92) that

$$\hat{u}_s(\rho_1, \rho_2, z, \mathbf{x}_\alpha) = \begin{cases} \hat{\mathbf{U}}(\rho_1, \rho_2, \mathbf{x}_\alpha) e^{\sqrt{\rho_1^2 + \rho_2^2 - k^2}(z+b)} & \text{if } \rho_1^2 + \rho_2^2 > k^2, \\ C_1 e^{-\mathrm{i}\sqrt{k^2 - \rho_1^2 - \rho_2^2}(z+b)} + C_2 e^{\mathrm{i}\sqrt{k^2 - \rho_1^2 - \rho_2^2}(z+b)} & \text{if } \rho_1^2 + \rho_2^2 < k^2, \end{cases} \tag{10.93}$$

where $z < -b$. It is not immediately clear which of two terms in the second line of the last formula should be taken. However, it was proven in Theorem 4.1 of [204] that one should set $C_2 := 0$. Thus, for $z < -b$,

$$\hat{u}_s(\rho_1, \rho_2, z, \mathbf{x}_\alpha) = \begin{cases} \hat{\mathbf{U}}(\rho_1, \rho_2, \mathbf{x}_\alpha) e^{\sqrt{\rho_1^2+\rho_2^2-k^2}(z+b)} & \text{if } \rho_1^2 + \rho_2^2 > k^2, \\ \hat{\mathbf{U}}(\rho_1, \rho_2, \mathbf{x}_\alpha) e^{-\mathrm{i}\sqrt{k^2-\rho_1^2-\rho_2^2}(z+b)} & \text{otherwise.} \end{cases} \tag{10.94}$$

Observe that if the Fourier frequency satisfies $\rho_1^2 + \rho_2^2 > k^2$, the function $\hat{u}_s(\rho_1, \rho_2, z, \mathbf{x}_\alpha)$ decays exponentially with respect to $z \to -\infty$. Therefore, if the measurement surface is far away from the domain of interest, that is, D is large, then we can neglect the term in the first line of (10.94). In other words, we can neglect high spatial frequencies in (10.94). Thus, we take $z = -D$ in (10.94) the case $\rho_1^2 + \rho_2^2 < k^2$ and then use the inverse Fourier transform to get

$$\begin{aligned} \mathbf{U}(x, y, \mathbf{x}_\alpha) &= u_s(x, y, -b, \mathbf{x}_\alpha) && (10.95) \\ &= \frac{1}{(2\pi)^2} \int\limits_{\rho_1^2+\rho_2^2<k^2} \left[\int\limits_{\mathbb{R}^2} u_s(\tilde{x}, \tilde{y}, -D, \mathbf{x}_\alpha) e^{-\mathrm{i}(\tilde{x}\rho_1+\tilde{y}\rho_2)} d\tilde{x} d\tilde{y} \right] \\ &\quad \times e^{\mathrm{i}\sqrt{k^2-\rho_1^2-\rho_2^2}(-D+b)} e^{\mathrm{i}(x\rho_1+y\rho_2)} d\rho_1 d\rho_2. && (10.96) \end{aligned}$$

The last formula of (10.95) is the actual data propagation procedure we will use in this work.

Remark 10.7.1. Since we ignore first lines in (10.93) and in (10.94), then (10.95) means that the data propagation procedure actually cuts off high spatial frequencies.

10.7.5 Computational setup

We introduce dimensionless variables as $\mathbf{x}' = \mathbf{x}/(10\text{ cm})$ and keep the same notation as before, for brevity. This means that the dimensions we use in computations are 10 times less than the real ones in centimeters. We illustrate the choice of the coordinate system on Figures 10.2(c), 10.3(c), 10.4(c), 10.5(c), and 10.6(c): the x- and y-axis are horizontal and vertical sides, respectively, and z-axis is orthogonal to the measurement plane. The far-field data are measured on a rectangular surface of dimensions 100 cm × 100 cm, that is, 10 × 10 in dimensionless regime. Compare Figure 10.1(b) as to our mesh grid of the measurement plane, each step is 2 cm (0.2) over 100 cm (10) length row. The total number of steps in a row is 50, and the total number of steps in a column is also 50. The distance between the measurement plane and the sandbox with the foam layer, whose thickness is 5 cm, is about 110.5 cm (11.05). The length in the z direction of the sandbox without the foam is approximately 44 cm, but due to the bending foam layer, we reduce 10 % of this length. Henceforth, the domain Ω should be taken as $\Omega = \{\mathbf{x} \in \mathbb{R}^3 : |x|, |y| < 5, |z| < 2\}$, which implies that $R = 5$ and $b = 2$. The near-field or propagated measurement site is then assigned as $\Gamma := \{\mathbf{x} \in \mathbb{R}^3 : |x|, |y| < 5, z = -2\}$. Also, we take $D = 14$ for the far-field measurement site as we estimate the distance between this site and the zero point. Meanwhile, for all objects, for the line of sources L_{src} defined in (10.4), we have $d = 9$, $a_1 = 0.1$, and $a_2 = 0.6$. Besides, we take $\theta = 4$ in the CWF (10.29) and in the series (10.21) we take $N = 6$.

It remains to obtain the wavenumber k corresponding to the dimensionless spatial variables we are working with. It is well known that the relation between the wavelength ($\tilde{\lambda}$) and the wavenumber is expressed by $k = 2\pi/\tilde{\lambda}$. Basically, the wavelength can be computed via the formulation $\tilde{\lambda} = \tilde{v}/\tilde{f}$, where $\tilde{v} = 299{,}792{,}458$ (m/s) is the speed of light in vacuum and $\tilde{f}$ is the frequency in Hertz (Hz or s^{-1}). Hence, in the computational setting we compute (in cm^{-1})

$$k = \frac{2\pi}{2{,}997{,}924{,}580}\tilde{f}.$$

The choice of k relies on the performance of the data after preprocessing. More precisely, our criterion is heuristically based upon the best visualization of the propagated data that we obtain using the data propagation. For each example below, we then use its own frequency, which we specify in Table 10.1. Note that for each location of the detector we measure the backscatter data for 300 frequency points uniformly distributed between 1 GHz and 10 GHz.

Now, we summarize the crucial steps of the data preprocessing to obtain fine data for our inversion method from the raw ones.

– *Step 1.* For every frequency and for every location of the source, we subtract the reference data from the far-field measured data. A similar procedure was implemented in [143, 204, 242]. The reference data are the background ones measured when the sandbox is without a target. This subtraction helps to extract the pure signals from buried objects from the whole signal. Therefore, we reduce the noise this way.
– *Step 2.* We apply the data propagation procedure to obtain an approximation of the near-field data. This procedure provides a significantly better estimation for x, y coordinates of buried objects, and reduces the size of the computational domain in the z−direction; see Figures 10.2(a),(b)–10.6(a),(b).
– *Step 3.* We truncate the so obtained near-field data to get rid of random oscillations. The oscillations appear randomly during the data propagation and may cause unnecessary issues during our inversion procedure. This data truncation

Table 10.1: Wave numbers and frequencies for examples 1–5, see [258].

Example	1	2	3	4	5
k	8.51	6.62	11.43	9.55	8.79
Frequency (GHz)	4.06	3.16	5.45	4.55	4.19

was developed in [204] and now we improve it using the following two steps, given a function $g(x,y,\alpha)$ to be truncated:

– For each point source, we replace the function $g(x,y,\alpha)$ with a function $\tilde{g}(x,y,\alpha)$ defined as

$$\tilde{g}(x,y,\alpha) = \begin{cases} g(x,y,\alpha) & \text{if } |g(x,y,\alpha)| \geq \kappa_1 \max_{|x|,|y|\leq R} |g(x,y,\alpha)|, \\ 0 & \text{otherwise.} \end{cases}$$

  Here, we call $\kappa_1 > 0$ the truncation number. Even though this number should be dependent of the source position α and should be different from every single choice of the frequency point, we apply the same truncation number to all the examples below. By the trial and error procedure, we have chosen $\kappa_1 = 0.4$, which means that we only preserve those propagated near-field data whose values are least 40 % of the global maximum value.

– The next step would be smoothing the function $\tilde{g}$ using the Gaussian filter. However, we notice that when doing so, the maximum value of $\tilde{g}$ will be smaller than that of g. In order to preserve this important "peak" of g after truncation, we add back some percents of $\tilde{g}$ in the following manner:

$$\tilde{g}_{\text{new}}(x,y,\alpha) = \kappa_2 \tilde{g}_{\text{old}}(x,y,\alpha). \tag{10.97}$$

  Here, we call $\kappa_2 > 0$ the retrieval number. This number is computed by $\kappa_2 = \max(|\tilde{g}|)/\tilde{m}$, where $\tilde{m}$ is the maximal absolute value of the smoothed $\tilde{g}_{\text{old}}$.

Fully discrete setting

We now present our numerical approach of the approximation of the right-hand side of formula (10.95) in order to use it for our experimental data. First, we adapt the conventional Riemannian sum approximation to compute the Fourier transform of the function $\mathbf{V}$. Using (10.91) and the samples $\{u_s(\tilde{x}_i,\tilde{y}_j,-D,\mathbf{x}_\alpha)\}_{i,j=0}^{\tilde{N}-1}$ over a 2D finite domain, where we are experimentally measuring the far-field data, we find that

$$\hat{\mathbf{V}}(\rho_1,\rho_2,\mathbf{x}_\alpha) \approx \omega^2 \sum_{i,j=0}^{\tilde{N}-1} u_s(\tilde{x}_i,\tilde{y}_j,-D,\mathbf{x}_\alpha)\exp\big(-\mathrm{i}(\tilde{x}_i\rho_1 + \tilde{y}_j\rho_2)\big).$$

Here, a uniform sampling rate, that is, $\tilde{x}_i = i\Delta\tilde{x}_i$, $\tilde{y}_j = j\Delta\tilde{y}_j$, is used with $\Delta\tilde{x}_i = \Delta\tilde{y}_j = \omega \in (0,1)$. Next, we define the following truncated Fourier domain in 2D as $\Theta_k := \{(\rho_1,\rho_2) \in \mathbb{R}^2 : \rho_1^2 + \rho_2^2 < k^2\}$. We sample this truncated Fourier domain at uniformly discrete points $\rho_{1m_1} = m_1\omega_\rho$, $\rho_{2m_2} = m_2\omega_\rho$ for a number $\omega_\rho \in (0,1)$ and $0 \leq m_1, m_2 \leq \tilde{M}-1$, provided that these points are in the set Θ_k. Thus, we conclude that

$$\begin{aligned} &\mathbf{U}(x_p,y_q,\alpha_l) \\ &\quad\approx \frac{1}{(2\pi)^2}\omega_\rho^2 \sum_{m_1,m_2=0}^{\tilde{M}-1} \hat{\mathbf{V}}(\rho_{1m_1},\rho_{2m_2},\alpha_l)\exp\Big(\mathrm{i}\sqrt{k^2-\rho_{1m_1}^2-\rho_{2m_2}^2}(-D+b)\Big) \\ &\quad\times\exp\big(\mathrm{i}(x_p\rho_{1m_1}+y_q\rho_{2m_2})\big). \end{aligned} \tag{10.98}$$

In our experimental data, we have $N_p = N_q = 51$, where N_p and N_q are the number of discrete points in x and y directions, respectively. Therefore, we take $\tilde{N} = \tilde{M} = 51$, which gives $\omega = \omega_\rho = 1/50$. Thus, (10.98) gives us the approximate Dirichlet boundary condition $V(\mathbf{x}^h) = \varphi_0^h(\mathbf{x}^h)$ at $\{z = -b\}$ in (10.24) in the discrete form. Here, h is the grid step size and $\mathbf{x}^h$ is the grid point at $\{z = -b\}$. Since we also need the function $V_z(\mathbf{x}^h) = \varphi_1^h(\mathbf{x}^h)$ at $\{z = -b\}$ in (10.25), then to obtain it, we formally replace in (10.98) b with z, differentiate the right-hand side of the obtained equality with respect to z, then set again $z := -b$ and calculate the resulting sum. The result is $\varphi_1^h(\mathbf{x}^h)$ at $\{z = -b\}$ in (10.25). Hence,

$$\varphi_0^h(\mathbf{x}^h) = (\varphi_{01}^h, \dots, \varphi_{0N-1}^h)^T(\mathbf{x}^h), \quad \varphi_1^h(\mathbf{x}^h) = (\varphi_{11}^h, \dots, \varphi_{1N-1}^h)^T(\mathbf{x}^h). \tag{10.99}$$

Hence, the Cauchy boundary data in (10.24), (10.25) are in the fully discrete form now. Then we write the functional $J_{h,\lambda}(V^h)$ defined in (10.49) in the fully discrete form. In this fully discrete setting, we take into account the grid points in x, y, z directions, $\{(x_p, y_q, z_s)\}_{p,q,s=0}^{Z_h}$. For brevity, we do not bring in here this fully discrete form of $J_{h,\lambda}(V^h)$.

After the global minimum $V_{p,q,s}$ of the functional $J_{h,\lambda}(V^h)$ (in its discrete form) is obtained, we compute an approximation of the unknown dielectric constant $c_{p,q,s}$ using the following formula:

$$c_{p,q,s} = \operatorname{mean}_{\alpha_l}\left|\operatorname{Re}\left\{-\frac{\Delta^h v_{p,q,s,\alpha_l} + (\nabla^h v_{p,q,s,\alpha_l})^2 + 2\nabla^h v_{p,q,s,\alpha_l}\cdot\tilde{\mathbf{x}}_{p,q,s,\alpha_l}}{k^2}\right\}\right| + 1,$$

which resulted from (10.18); see Remark 10.3.1. Here, $v_{p,q,s,\alpha_l} = v(x_p, y_q, z_s, \alpha_j)$, also Δ^h and ∇^h are finite difference analogs of operators Δ and ∇, respectively. Recall that $\tilde{\mathbf{x}}_{p,q,s,\alpha_l}$ denote vectors $\tilde{\mathbf{x}}_\alpha$ at (x_p, y_q, z_s) for every α_l; see Section 10.3.1. Since the number of point sources is small, we apply the Gauss–Legendre quadrature method to compute the measured data in the Fourier mode.

Since this work focuses on the detection and identification of antipersonnel land mines and IEDs, we know that the sizes of these targets are between 5 and 15 cm; cf., for example, [204]. Therefore, we search for targets in a subdomain of Ω with only 20 cm in depth in the z-direction. Denote this subdomain by $\Omega_1 = \{-b \le z \le -b + 2\}$.

We consider the following vector $V_0^h = V_0(x_p, y_q, z_s)$ as the starting point of iterations in the minimization of the functional $J_{\lambda,h}(V^h)$:

$$V_0^h = \begin{pmatrix} v_{00}^h & v_{01}^h & \cdots & v_{0(N-1)}^h \end{pmatrix}^T, \quad v_{0n}^h = (\varphi_{0n}^h + \varphi_{1n}^h(z+b))\chi(z), \tag{10.100}$$

where φ_{0n}^h and φ_{1n}^h are Fourier coefficients in (10.99). Here, $\chi : [-b, b] \to \mathbb{R}$ is the smooth function given by

$$\chi(z) = \begin{cases} \exp\left(\frac{2(z+b)^2}{(z+b)^2 - b^2}\right) & \text{if } z < 0, \\ 0 & \text{otherwise.} \end{cases}$$

This function attains the maximum value 1 at $z = -b$ where the propagated data are given. Then it is easy to see that $v^h_{0n}|_{z=-b} = \varphi^h_{0n}$, $\partial_z v^h_{0n}|_{z=-b} = \varphi^h_{1n}$. On the other hand, χ tends to 0 as $z \to 0^+$, which, in particular, means that $v^h_{0n}|_{z=b} = \partial_z v^h_{0n}|_{z=b} = 0$. Thus, this starting point (10.100) of iterations satisfies the discrete form of boundary conditions (10.24), (10.25).

Although Theorems 10.6.1 and 10.6.2 claim the global convergence of the gradient projection method, we have successfully used the gradient descent method for the minimization of the target functional $J_{h,\lambda}(V^h)$ in (10.49). Clearly, the gradient descent method is easier to implement than the gradient projection method. Our success in working with the gradient descent method is similar with the success in Chapters 7–9 as well as in all previous publications discussing the numerical studies of the convexification [115–117, 142–145, 145, 146, 150, 151, 164]. As to the value of the parameter λ in $J_{h,\lambda}(V^h)$, even though the above analysis requires large values of λ, our numerical experience tells us that we can choose a moderate value $\lambda = 1.1$. Again similar values of $\lambda \in [1, 3]$ were chosen in all above cited publications on the convexification.

As to the step size γ of the gradient descent method, we start from $\gamma_1 = 10^{-1}$. For each iteration step $m \geq 1$, the following step size γ_m is reduced by the factor of 2 if the value of the functional on the step m exceeds its value of the previous step. Otherwise, $\gamma_{m+1} = \gamma_m$. The minimization process is stopped when either $\gamma_m < 10^{-10}$ or $|J_{h,\lambda}(V^h_m) - J_{h,\lambda}(V^h_{m-1})| < 10^{-10}$. As to the gradient $J'_{h,\lambda}$ of the discrete functional $J_{h,\lambda}$, we apply the technique of Kronecker deltas (cf., e. g., [172]) to derive its explicit formula, which significantly reduces the computational time. For brevity, we do not provide this formula here.

After the minimization procedure is stopped, we obtain numerically the coefficient of $c_{p,q,s}$, denoted by $\tilde{c}$. Our reconstructed solution, denoted by c_{comp}, is concluded after we smooth $\tilde{c}$ by the standard filtering via the `smooth3` built-in function in MATLAB. In fact, we find c_{comp} by using $c_{\mathrm{comp}} = \hat{\varrho}\,\mathrm{smooth}(|\tilde{c}|)$, for some $\hat{\varrho} > 0$ depending on every single example. This step is definitely similar to the smoothing procedure discussed in (10.97) and we do not repeat how to find $\hat{\varrho}$ here. We use this step to get better images.

10.7.6 Reconstruction results

Values of $\max(c_{\mathrm{true}})$ and $\max(c_{\mathrm{comp}})$ for all five tests are tabulated in Table 10.2. Values of $\max(c_{\mathrm{true}})$ for all tests, which were used, are published in [65, 170, 242]. More precisely, as to the metallic targets of Example 1 (aluminum cylinder), Example 4 (metallic letter “A”), and Example 5 (metallic letter “O”), it was numerically established that one can treat metals as materials with large values of the dielectric constant in the interval $c \in [10, 30]$; see the formula (7.2) of [170]. As to Example 2, the dielectric constant of the clear water for our frequency range was directly measured in [242], and it was 23.8: see the first line of Table 1 of [242]. As to the Example 3 (an U-shaped piece of a

Table 10.2: True c_{true} and computed $\max(c_{comp})$ dielectric constants of Examples 1–5 of experimental data. True values for Examples 1, 4, 5 were taken from formula (7.2) of [170]. True value for Example 2 (clear water) was taken from [242]. True value for Example 3 (wood) was taken from [65].

Example number	**1**	**2**	**3**	**4**	**5**
Object	Met. cyl.	Water	Wood	Met. let. 'A'	Met. let 'O'
c_{true}	$\in [10, 30]$	23.8	$\in [2, 6]$	$\in [10, 30]$	$\in [10, 30]$
$\max(c_{comp})$	18.72	23.29	6.56	15.01	16.25

dry wood), the table of dielectric constants [65] tells one that the dielectric constant of a dry wood is $c \in [2, 6]$.

Figures 10.2(a), 10.3(a), 10.4(a), 10.5(a), and 10.6(a) show how "bad" the far-field data look like. It is clear from these figures that the data need a preprocess to have a proper inversion. One can see the good shapes of the corresponding images after the data propagation procedure; see Figures 10.2(b), 10.3(b), 10.4(b), 10.5(b), and 10.6(b). For every test, we deliberately show the 2D illustrations (raw and propagated) of the data at a specific point source, where the images of the propagated data and the computed inclusion are congruent with each other.

3D images of computed inclusions are depicted by using the `isosurface` function in MATLAB with the associated `isovalue` being 10 % of the maximal value; see 10.2(d), 10.3(d), 10.4(d), 10.5(d), and 10.6(d). The most challenging targets to image were: (1) The U-shaped piece of dry wood, see Figure 10.4, (2) The metallic letter "A," see Figure 10.5, and (3) the metallic letter "O," see Figure 10.6. This is because these targets have the most complicated geometries. Nevertheless, we are still able to see their characteristic shapes in the images of computed inclusion.

Most notably, one can see voids in imaged letters "A" and "O," The latter is usually difficult to achieve. The "strip" of the letter "A" is not imaged since its width was 2.5 cm, which is less than the used wavelength of 10.4 cm with $k = 9.55$. Another interesting observation is that we can even see the cap on the bottle of water on Figure 10.3(d).

We also find that the lengths of parts of true and computed inclusions are quite compatible with each others. Note that even though the computed inclusions here are slightly larger (just a few centimeters) than the true ones, it is still useful in detection and identification of land mines and further in the mine-clearing operations. In fact, having information of smaller sizes is rather dangerous. Hence, we conclude that the dimensions of the computed inclusions are acceptable. Finally, we can accurately obtain approximations of the dielectric constants. Aside from the dielectric constant of metallic targets, we notice from Table 10.2 that the relative errors obtained for the bottle with water and for the wooden target are 2.14 % and 9.33 %, respectively.

11 Travel time tomography with formally determined incomplete data in 3D

In this chapter, we closely follow [138]. Permission for republication is obtained from the publisher.

11.1 Introduction

In this chapter, we construct the convexification globally convergent numerical method for the challenging 3D travel time tomography problem (TTTP) with formally determined incomplete data. “Formally determined” means that the number n of free variables in the unknown coefficient equals the number m free variables in the data. “Incomplete” means that the data are measured only at a part of the boundary of the domain of interest rather than on the whole boundary. For example, the latter was the case of the backscattering data of the CIP considered in Chapter 10. A similar version of the convexification method for the TTTP, although using slightly different assumptions, was constructed by Klibanov in [137].

The TTTP was first considered by Herglotz [88] in 1905 and then by Wiechert and Zoeppritz [250] in 1907 in the 1D case due to a geophysical application; also, see a detailed description of the 1D case in [224]. However, globally convergent numerical methods for the 3D TTTP with formally determined data were not developed so far.

In this paragraph, we indicate those ideas for the TTTP, which are presented here *for the first time*. Since we develop a numerical method, we are allowed to work here with an *approximate mathematical model,* as we did in Chapters 6–8, 10; see Remarks 7.3. The TTTP is considered in the semidiscrete form, that is, we consider the practically important case of finite differences with respect to two out of three variables. The Lipschitz stability estimate is obtained, which implies uniqueness. Our method does not use sophisticated geometrical properties to construct the above sequence. Rather, we straightforwardly minimize the above mentioned globally strictly convex Tikhonov-like functional.

The TTTP is the problem of the recovery of the spatially distributed speed of acoustic waves from first times of arrival of those waves. Another well-known term for the TTTP is the “inverse kinematic problem of seismic” [224]. Waves are originated by some point sources located on the boundary of the domain of interest. First times of arrival are recorded on a number of detectors located on that boundary. It is well known that the TTTP is actually a nonlinear problem of the integral geometry; see, e. g., [224]. The TTTP has important applications in geophysics [88, 224, 249, 250]. In addition, it was established in [159] that the TTTP arises in the phaseless inverse problem of scattering of electromagnetic waves at high frequencies. The specific TTTP con-

https://doi.org/10.1515/9783110745481-011

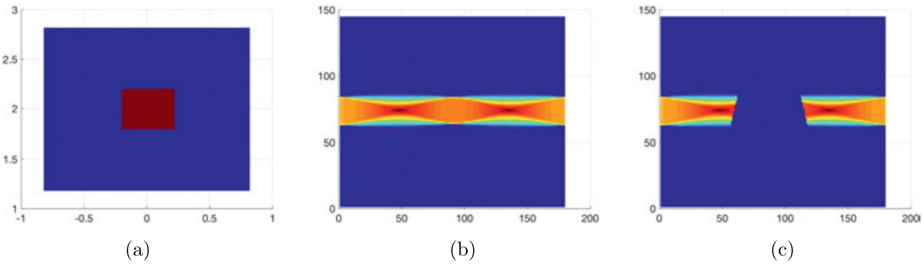


Figure 11.1: An illustration for complete and incomplete data in the 2D case; see details in [224]. To simplify, we assume in this figure that the geodesics are straight lines. Thus, we deal in this figure with the data of Radon transform, generated by the function "radon" of MATLAB. (a) The true function $m(\mathbf{x})$ to be imaged. (b) The complete data of the Radon transform of the function of (a). (c) The incomplete data of the Radon transform of (a) in the case when the source runs along an interval of a straight line, as in this paper below.

sidered here has potential applications in geophysics, checking the bulky baggage in airports, search for possible defects inside the walls, etc.

The sole purpose of Figure 11.1 is to illustrate this for a simple case when geodesics are straight lines.

Our *approximate mathematical model* consists of two items; see Remarks 11.6.1 in the Section 11.6 for more details. First, we consider a semidiscrete model. This means that partial derivatives with respect to two out of three variables are written in finite differences. Second, we assume that a certain function generated by the solution of the eikonal equation can be represented as a truncated Fourier series with respect to the orthonormal basis in the $L_2(0,1)$ space, which was constructed in Chapter 5. Functions of that basis depend only on the position of the source. The number $N \geq 1$ of terms of this series is not allowed to tend to the infinity.

As to the numerical methods for the TTTP in the n-D case, $n = 2, 3$, such a method for the 2D TTTP was published in [230]. Another numerical approach in 3D was published in [257]. Both publications [230, 257] use, at a certain stage, the minimization of a least squares cost functional. Since the convexity of those functionals of [230, 257] is not proven, then the problem of local minima is not addressed there. In both publications [230, 257], complete data are used, and these data are overdetermined ones in the 3D case of [257].

The first global Lipschitz stability and uniqueness result for the TTTP was obtained by Mukhometov in the 2D case [201]. Next, this result was extended in [36, 202, 224] to the n-D case, $n \geq 3$. We also refer to the related work [219] for the 2D case. In all of these references, the data are complete and the assumption of the regularity of geodesic lines is used. In addition, more recently the question of uniqueness in the 3D case when geodesic lines are not necessarily regular ones was considered in [239]. In the 2D case of [201, 219, 230], the data are formally determined. However, they are overdetermined in the n-D case with $n \geq 3$ [36, 202, 224, 239, 257]. Contrary to this, we work with non-overdetermined data, both in this and next chapters.

11.2 Statement of the problem

Below $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. Let numbers $A, \sigma = \text{const.} > 0$. Define the rectangular prism $\Omega \subset \mathbb{R}^3$ as

$$\Omega = \{\mathbf{x} = (x, y, z) : x, y \in (0,1), z \in (A, A+\sigma)\}. \tag{11.1}$$

Denote parts of the boundary $\partial\Omega$ as

$$B_A = \{\mathbf{x} = (x, y, z) : x, y \in (0,1), z = A\}, \tag{11.2}$$

$$B_{A+\sigma} = \{\mathbf{x} = (x, y, z) : x, y \in (0,1), z = A+\sigma\}, \tag{11.3}$$

$$\Gamma = \partial\Omega \setminus (B_A \cup B_{A+\sigma}). \tag{11.4}$$

Let $n(\mathbf{x})$ be the refractive index of the medium at the point $\mathbf{x}$. Hence, $c(\mathbf{x}) = 1/n(\mathbf{x})$ is the sound speed. Denote $m(\mathbf{x}) = n^2(\mathbf{x})$. Let the number $m_0 > 0$ be given. We impose the following assumptions on the function $m(\mathbf{x})$:

$$m(\mathbf{x}) \geq m_0, \quad \mathbf{x} \in \mathbb{R}^3, \tag{11.5}$$

$$m(\mathbf{x}) = 1, \quad \mathbf{x} \in \{z < A\}, \tag{11.6}$$

$$m \in C^2(\mathbb{R}^3), \tag{11.7}$$

$$m_z(\mathbf{x}) \geq 0, \quad \mathbf{x} \in \overline{\Omega}. \tag{11.8}$$

Remark 11.2.1. We note that the monotonicity condition (11.8) is not an overly restrictive one. Indeed, it can also be found in the end of Section 11.2 of Chapter 11.3 of the book [224]; see formulas (3.24) and (3.24′) there. Also, an analogous monotonicity condition was actually imposed in the 1D case in the originating classical works of Herglotz and Wiechert and Zoeppritz [88, 250]; see Section 3 of Chapter 3 of [224] for a description of their method. We also refer to Figures 5 and 10 in [249] for some geophysical information.

The function $m(\mathbf{x})$ generates the Riemannian metric

$$d\tau = \sqrt{m(\mathbf{x})}|d\mathbf{x}|, \quad |d\mathbf{x}| = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}. \tag{11.9}$$

The travel time from the point $\mathbf{x}_0 \in \mathbb{R}^3$ (source) to the point $\mathbf{x} \in \mathbb{R}^3$ (receiver) is [224]

$$\tau(\mathbf{x}, \mathbf{x}_0) = \int\limits_{\Gamma(\mathbf{x}, \mathbf{x}_0)} \sqrt{m(\mathbf{x}(s))}\, ds, \tag{11.10}$$

where $\Gamma(\mathbf{x}, \mathbf{x_0})$ is the geodesic line connecting points $\mathbf{x}$ and $\mathbf{x}_0$ and ds is the Euclidean arc length. We assume that the source $\mathbf{x}_0$ runs along an interval L of a straight line located in the plane $\{z = 0\}$,

$$L = \{\mathbf{x} = (x, y, z) : x = \alpha \in (0, 1), y = 1/2, z = 0\}. \tag{11.11}$$

Hence, $\mathbf{x}_0 = \mathbf{x}_\alpha = (\alpha, 1/2, 0)$, $\alpha \in (0, 1)$. Let $\tau(\mathbf{x},\alpha)$ be the travel time between points $\mathbf{x}$ and $\mathbf{x}_\alpha = (\alpha, 1/2, 0)$. Thus, we denote $\tau(\mathbf{x},\alpha) = \tau(\mathbf{x}, \mathbf{x}_\alpha)$. We denote $\Gamma(\mathbf{x}, \alpha)$ the geodesic line connecting points $\mathbf{x}$ and $\mathbf{x}_\alpha$. It is well known [224] that the function $\tau(\mathbf{x}, \alpha)$ satisfies eikonal equation as the function of $\mathbf{x}$,

$$\begin{gathered} \tau_x^2 + \tau_y^2 + \tau_z^2 = m(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3, \\ \tau(\mathbf{x}, \alpha) = O(|\mathbf{x} - \mathbf{x}_\alpha|), \quad \text{as } |\mathbf{x} - \mathbf{x}_\alpha| \to 0. \end{gathered} \tag{11.12}$$

Everywhere below we assume without further mentioning that the following property holds.

Regularity of geodesic lines. The function $\tau(\mathbf{x}, \alpha) \in C^2(\mathbb{R}^3 \times [0, 1])$. For any pair of points $(\mathbf{x}, \mathbf{x}_\alpha) \in \overline{\Omega} \times L$ there exists unique geodesic line $\Gamma(\mathbf{x}, \alpha)$ connecting these two points and $\Gamma(\mathbf{x}, \alpha) \cap B_A \neq \varnothing$. In addition, if any geodesic line, which starts at a point $\mathbf{x}_\alpha \in L$, intersects B_A, then it intersects it at a single point. Also, it does not go "downwards" in the z-direction, but instead intersects $\partial\Omega \setminus B_A$ at another single point; see (11.1)–(11.4). In addition, after intersecting $\partial\Omega \setminus B_A$, this line does not "come back" in the domain Ω but rather goes away from this domain. In other words, this line is not reflected back from any point of its intersection with $\partial\Omega$.

The following sufficient condition of the regularity of geodesic lines was derived in [225]:

$$\sum_{i,j=1}^{3} \frac{\partial^2 \ln n(\mathbf{x})}{\partial x_i \partial x_j} \xi_i \xi_j \geq 0, \quad \forall \xi \in \mathbb{R}^3, \forall \mathbf{x} = (x_1, x_2, x_3) \in \overline{\Omega}.$$

Travel Time Tomography Problem (TTTP). Suppose that the function $m(\mathbf{x})$ satisfies conditions (11.5)–(11.8). Assume that the following function $f(\mathbf{x}, a)$ is given:

$$\tau(\mathbf{x},\alpha) = f(\mathbf{x}, \alpha), \quad \forall \mathbf{x} \in \Gamma \cup B_{A+\sigma}, \forall \alpha \in (0, 1). \tag{11.13}$$

Determine the function $m(\mathbf{x})$ for $\mathbf{x} \in \Omega$.

In other words, by (11.1)–(11.4) and (11.13) the data for the travel time are given for all sources running along the line interval L defined in (11.11) and for the part $\Gamma \cup B_{A+\sigma}$ of the boundary $\partial\Omega$. Hence, the data (11.13) are both formally determined and incomplete.

11.3 A special orthonormal basis

We now describe the orthonormal basis in $L_2(0,1)$, which is a little bit different from the one in Section 6.2.3. Consider the set of functions $\{\xi_n(\alpha)\}_{n=0}^{\infty} = \{(\alpha + a)^n e^{\alpha}\}_{n=0}^{\infty}$, where $a = \text{const.} > 0$. This is a set of linearly independent functions. Besides, functions $\Psi_n(\alpha)$ are polynomials orthonormal in $L_2(0,1)$ with the weight $e^{2\alpha}$. The matrix $M_N = \{(\Psi_n', \Psi_m)\}_{m,n=1}^{N}$ is invertible since its elements $a_{mn} = (\Psi_n', \Psi_m)$ are such that $a_{mn} = 1$ if $m = n$ and $a_{mn} = 0$ if $n < m$.

Consider the function $q(\alpha)$ in the following form:

$$q(\alpha) = \sum_{n=1}^{N} q_n \Psi_n(\alpha), \quad q_n = \int_0^1 q(\alpha)\Psi_n(\alpha) d\alpha. \tag{11.14}$$

Below we need to impose such a sufficient condition on the vector of coefficients $q^N = (q_0, \ldots, q_{N-1})^T$ in the Fourier expansion (11.14), which would guarantee that the function $q(\alpha)$ is positive for all $\alpha \in [0,1]$. Consider vector functions $\xi^N(\alpha) = (\xi_1, \ldots, \xi_N)^T(\alpha)$ and $\Psi^N(\alpha) = (\Psi_1, \ldots, \Psi_{N-1})^T(\alpha)$. The desired condition is given in Lemma 11.3.1.

Lemma 11.3.1. *Let the $N \times N$ matrix X_N transforms the vector function $\xi^N(\alpha)$ in the vector function $\Psi^N(\alpha)$ via the Gram–Schmidt orthonormalization procedure, that is, $X_N \cdot \xi^N(\alpha) = \Psi^N(\alpha)$. Let the matrix X_N^T be the transpose of X_N. Let $X_N^T q^N = \tilde{q}^N = (\tilde{q}_0, \ldots, \tilde{q}_{N-1})^T$. Suppose that all numbers $\tilde{q}_0, \ldots, \tilde{q}_{N-1} > 0$. Then in* (11.14) *the function $q(\alpha) > 0$ for all $\alpha \in [0,1]$.*

Proof. It follows from the Gram–Schmidt procedure that elements of the matrix X_N are independent on α. Let the raw number n of the matrix X_N be $(x_{n1}, x_{n2}, \ldots, x_{nN})$, $n = 1, \ldots, N$. Then

$$\Psi_n(\alpha) = \sum_{j=1}^{N} x_{nj} \xi_j(\alpha).$$

Hence, by (11.14)

$$q(\alpha) = \sum_{n=1}^{N} q_n \sum_{j=1}^{N} x_{nj} \xi_j(\alpha) = \sum_{j=1}^{N} \left(\sum_{n=1}^{N} x_{nj} q_n \right) \xi_j(\alpha) = \sum_{j=1}^{N} \tilde{q}_j \xi_j(\alpha). \tag{11.15}$$

Since $\xi_j(\alpha) = (\alpha + a)^j e^{\alpha} > 0$, then (11.15) implies that $q(\alpha) > 0$ for $\alpha \in [0,1]$. □

11.4 Estimate of $\tau_z^2(\mathbf{x}, \alpha)$ from the below

Lemma 11.4.1. *Assume that conditions* (11.5)–(11.8) *hold. Then*

$$\tau_z^2(\mathbf{x}, \alpha) \geq \frac{A^2}{A^2 + 2}, \quad \forall \mathbf{x} \in \overline{\Omega}, \forall \alpha \in [0,1]. \tag{11.16}$$

Also,

$$\tau_z(\mathbf{x},\alpha) > 0, \quad \forall \mathbf{x} \in \overline{\Omega} \cup \{z \in (0,A]\}, \forall \alpha \in [0,1]. \tag{11.17}$$

Thus,

$$\tau_z(\mathbf{x},\alpha) \geq \frac{A}{\sqrt{A^2+2}}, \quad \forall \mathbf{x} \in \overline{\Omega}, \forall \alpha \in [0,1]. \tag{11.18}$$

Proof. Note that having proven (11.16) is not enough for our technique: we need to know the sign of the function $\tau_z(\mathbf{x},\alpha)$, that is (11.17), in Section 11.5 (more precisely, in (11.31)) where we consider the square root of $\tau_z^2(\mathbf{x},\alpha)$. Denote

$$p = \tau_x(x,y,z,\alpha), \quad q = \tau_x(x,y,z,\alpha), \quad r = \tau_z(x,y,z,\alpha). \tag{11.19}$$

The following equations for geodesic lines can be found on page 66 of [224]:

$$\frac{dx}{ds} = \frac{p}{m}, \quad \frac{dy}{ds} = \frac{q}{m}, \quad \frac{dz}{ds} = \frac{r}{m}, \tag{11.20}$$

$$\frac{dp}{ds} = \frac{m_x}{2m}, \quad \frac{dq}{ds} = \frac{m_y}{2m}, \quad \frac{dr}{ds} = \frac{m_z}{2m}, \tag{11.21}$$

where s is a parameter. Using (11.19), we obtain for $\tau = \tau(x(s),y(s),z(s),\alpha)$ along a geodesic line

$$\frac{d\tau}{ds} = \frac{\partial\tau}{\partial x}\frac{dx}{ds} + \frac{\partial\tau}{\partial y}\frac{dy}{ds} + \frac{\partial\tau}{\partial z}\frac{dz}{ds} = p\frac{p}{m} + q\frac{q}{m} + r\frac{r}{m} = 1. \tag{11.22}$$

Set

$$\tau\big(x(0),y(0),z(0),\alpha\big) = 0 \quad \text{for } s = 0. \tag{11.23}$$

Then (11.22) implies: $\tau(x(s),y(s),z(s),\alpha) = s$. Hence, the parameter s coincides with the travel time. In particular, we now rewrite equations (11.20), (11.21) in a different form: to have derivatives with respect to z rather than with respect to s. Hence, we obtain from (11.20) and (11.21):

$$\frac{dx}{dz} = \frac{p}{r}, \quad \frac{dy}{dz} = \frac{q}{r}, \quad \frac{dp}{dz} = \frac{m_x}{2r}, \quad \frac{dq}{dz} = \frac{m_y}{2r}, \quad \frac{dr^2}{dz} = m_z, \quad \frac{d\tau}{dz} = \frac{m}{r}, \tag{11.24}$$

$$x|_{z=0} = \alpha, \quad y|_{z=0} = 1/2, \quad p|_{z=0} = p_0, \quad q|_{z=0} = q_0, \quad r|_{z=0} = r_0, \quad \tau|_{z=0} = 0, \tag{11.25}$$

where $\alpha \in (0,1)$ and p_0, q_0 are some given numbers such that $p_0^2 + q_0^2 \leq 1$. The latter inequality follows from (11.12) and the fact that by (11.6) $m(x,y,0) = 1$. Also, by (11.12) $r_0 = \pm\sqrt{1-p_0^2-q_0^2}$. To prove that we should take "+" sign in the latter formula, we note that

$$\tau(x,y,z,\alpha) = \sqrt{(x-\alpha)^2 + (y-1/2)^2 + z^2} \quad \text{for } (x,y,z) \in (0,1)\times(0,1)\times(0,A). \tag{11.26}$$

Hence, $\tau_z = r = z/\tau > 0$ for $z \in (0, A)$. Hence,

$$r_0 = \sqrt{1 - p_0^2 - q_0^2} \geq 0, \tag{11.27}$$

$$\tau_z^2(x, y, A, \alpha) \geq \frac{A^2}{A^2 + 2} \quad \text{for } (x, y) \in (0,1) \times (0,1), \alpha \in (0,1). \tag{11.28}$$

Suppose that the geodesic line defined by (11.24) and (11.25) intersects the part B_A of the boundary $\partial\Omega$. Then the condition of the regularity of geodesic lines implies that there exists a single number $z_0 = z_0(p_0, q_0, \alpha) \in (A, A + \sigma]$ such that the point $(x(z_0), y(z_0), z_0) \in \partial\Omega \setminus B_A$ and for all numbers $z \in (A, z_0)$ all points $(x(z), y(z), z)$ of that geodesic line belong to Ω. Since by (11.24) $dr^2/dz = m_z$, then, using (11.24) and (11.28), we obtain

$$\begin{aligned} r^2(x(z), y(z), z, \alpha) &= \int_A^z m_z(x(t), y(t), t, \alpha)dt + r^2(x(A), y(A), A) \\ &\geq r^2(x(A), y(A), A) \geq \frac{A^2}{A^2 + 2}, \quad z \in (A, z_0), \end{aligned}$$

which establishes (11.16). To prove (11.17), we notice that it follows from (11.5), (11.8), (11.19), the last equation (11.21) and (11.23) that

$$\tau_z((x(s), y(s), z(s), \alpha)) = \int_0^s \left(\frac{m_z}{2m}\right)(x(t), y(t), z(t), \alpha)dt \geq 0. \tag{11.29}$$

Estimates (11.17) and (11.18) follow immediately from (11.16), (11.26), and (11.29). □

11.5 A boundary value problem for a system of nonlinear coupled integro-differential equations

11.5.1 A nonlinear integro-differential equation

Denote

$$u(\mathbf{x}, \alpha) = \tau_z^2(\mathbf{x}, \alpha), \quad \mathbf{x} \in \Omega, \alpha \in (0, 1). \tag{11.30}$$

By (11.17)

$$\tau_z(\mathbf{x}, \alpha) = \sqrt{u(\mathbf{x}, \alpha)}, \quad \mathbf{x} \in \Omega. \tag{11.31}$$

Hence, (11.13) and (11.31) imply that for all $\alpha \in (0, 1)$

$$\tau(x, y, z, \alpha) = -\int_z^{A+\sigma} \sqrt{u(x, y, t, \alpha)}dt + f(x, y, A + \sigma, \alpha), \quad (x, y, z) \in \Omega, \tag{11.32}$$

$$\tau_x(x,y,z,\alpha) = -\int_z^{A+\sigma} \left(\frac{u_x}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_x(x,y,A+\sigma,\alpha), \quad (x,y,z) \in \Omega, \tag{11.33}$$

$$\tau_y(x,y,z,\alpha) = -\int_z^{A+\sigma} \left(\frac{u_y}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_y(x,y,A+\sigma,\alpha), \quad (x,y,z) \in \Omega. \tag{11.34}$$

Substituting (11.30)–(11.34) in the eikonal equation (11.12), we obtain the following equation for $(x,y,z) \in \Omega$, $\alpha \in (0,1)$:

$$\begin{aligned} u(x,y,z,\alpha) &+ \left[-\int_z^{A+\sigma} \left(\frac{u_x}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_x(x,y,A+\sigma,\alpha)\right]^2 \\ &+ \left[-\int_z^{A_1+\sigma} \left(\frac{u_y}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_y(x,y,A+\sigma,\alpha)\right]^2 = m(x,y,z). \end{aligned} \tag{11.35}$$

Differentiating (11.35) with respect to α and using $\partial_\alpha m(x,y,z) \equiv 0$, we obtain for $(x,y,z) \in \Omega$, $\alpha \in (0,1)$,

$$\begin{aligned} u_\alpha(x,y,z,\alpha) &+ \frac{\partial}{\partial\alpha}\left[-\int_z^{A+\sigma} \left(\frac{u_x}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_x(x,y,A+\sigma,\alpha)\right]^2 \\ &+ \frac{\partial}{\partial\alpha}\left[-\int_z^{A+\sigma} \left(\frac{u_y}{2\sqrt{u}}\right)(x,y,t,\alpha)dt + f_x(x,y,A+\sigma,\alpha)\right]^2. \end{aligned} \tag{11.36}$$

11.5.2 Boundary value problem for a system of coupled integro- differential equations

Using (11.26) and (11.30), denote for $(x,y,\alpha) \in [0,1]^3$,

$$u_0(\mathbf{x},\alpha) = u(x,y,A,\alpha) = \frac{A^2}{(x-\alpha)^2 + (y-1/2)^2 + A^2} \geq \frac{A^2}{A^2+2}. \tag{11.37}$$

We seek the function $u(\mathbf{x},\alpha)$ in the form

$$u(\mathbf{x},\alpha) = u_0(\mathbf{x},\alpha) + v(\mathbf{x},\alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0,1], \tag{11.38}$$

$$v(x,y,A,\alpha) = 0, \quad (x,y,\alpha) \in [0,1]^3, \tag{11.39}$$

where the function $v(\mathbf{x},\alpha)$ is unknown for $\mathbf{x} \in \Omega$, $\alpha \in (0,1)$. Recall that the part Γ of the boundary of the domain Ω is defined in (11.4). We need to obtain zero boundary condition at Γ for a function associated with the function v. To do this, we assume first that there exists a function $g(\mathbf{x},\alpha) \in H^1(\Omega)$ for every $\alpha \in [0,1]$ such that

$$g(\mathbf{x},\alpha) = (f_z)^2(\mathbf{x},\alpha) - u_0(\mathbf{x},\alpha), \quad \forall \mathbf{x} \in \Gamma, \forall \alpha \in [0,1], \tag{11.40}$$

$$g(x,y,A,\alpha) = 0. \tag{11.41}$$

Introduce the function $w(\mathbf{x}, \alpha)$,

$$w(\mathbf{x}, \alpha) = v(\mathbf{x}, \alpha) - g(\mathbf{x}, \alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1]. \tag{11.42}$$

Then (11.38)–(11.42) imply that

$$u(\mathbf{x}, \alpha) = u_0(\mathbf{x}, \alpha) + w(\mathbf{x}, \alpha) + g(\mathbf{x}, \alpha), \mathbf{x} \in \overline{\Omega}, \quad \alpha \in [0, 1], \tag{11.43}$$

$$w(\mathbf{x}, \alpha) = 0, \quad \forall \mathbf{x} \in \Gamma, \forall \alpha \in [0, 1], \tag{11.44}$$

$$w(x, y, A, \alpha) = 0, \quad (x, y, \alpha) \in [0, 1]^3. \tag{11.45}$$

We assume that both functions $w(\mathbf{x}, \alpha)$ and $g(\mathbf{x}, \alpha)$ have the form of the truncated Fourier series with respect to the orthonormal basis $\{\Psi_n(\alpha)\}$,

$$w(\mathbf{x}, \alpha) = \sum_{n=1}^{N} w_n(\mathbf{x}) \Psi_n(\alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1], \tag{11.46}$$

$$g(\mathbf{x}, \alpha) = \sum_{n=1}^{N} g_n(\mathbf{x}) \Psi_n(\alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1]. \tag{11.47}$$

Here, coefficients $w_n(\mathbf{x})$ are unknown and coefficients $g_n(\mathbf{x})$ are known. And similarly for w_x, w_y, $w_{x\alpha}$, $w_{y\alpha}$, and the same derivatives of the function g. Furthermore, we assume that these functions, being substituted in equation (11.36), give us zero in its right-hand side for $\mathbf{x} \in \Omega$, $\alpha \in (0, 1)$. By (11.44)–(11.46),

$$w_n(\mathbf{x})|_\Gamma = 0, \quad w_n(x, y, A) = 0. \tag{11.48}$$

Denote

$$W(\mathbf{x}) = (w_1, \dots, w_N)^T(\mathbf{x}), \tag{11.49}$$

$$G(\mathbf{x}) = (g_1, \dots, g_N)^T(\mathbf{x}). \tag{11.50}$$

Let

$$f(x, y, A + \sigma, \alpha) = \sum_{n=1}^{N} f_n(x, y, A + \sigma) \Psi_n(\alpha), \quad (x, y, \alpha) \in (0, 1)^3, \tag{11.51}$$

$$F(x, y, A + \sigma) = (f_1, f_2, \dots, f_N)^T(x, y, A + \sigma), \quad (x, y) \in (0, 1)^2. \tag{11.52}$$

Keeping in mind (11.48) and (11.49), we define the spaces $C_N^1(\overline{\Omega})$, $C_{N,0}^1(\overline{\Omega})$ of N-D vector functions $W(\mathbf{x})$ defined in (11.49) as

$$C_{x,y,N}^1(\overline{\Omega}) = \left\{ \begin{aligned} & W(\mathbf{x}) : \|W\|_{C_{x,y,N}^1(\overline{\Omega})} \\ & \quad = \max_{n \in [1,N]} \left(\|w_n\|_{C(\overline{\Omega})} + \|w_{nx}\|_{C(\overline{\Omega})} + \|w_{ny}\|_{C(\overline{\Omega})} \right) < \infty \end{aligned} \right\},$$

$$C_{x,y,N,0}^1(\overline{\Omega}) = \left\{ W \in C_{x,y,N}^1(\overline{\Omega}) : W|_\Gamma = W(x, y, A) = 0 \right\}.$$

Keeping in mind (11.37)–(11.52), substitute functions w, g and their corresponding first derivatives with respect to x, y, α in equation (11.36). Next, multiply the resulting equation sequentially by functions $\Psi_n(\alpha)$, $n = 1, \ldots, N$ and integrate with respect to $\alpha \in (0, 1)$. Then multiply both sides of obtained system of nonlinear integral differential equations by the matrix M_N^{-1}, where the matrix M_N was introduced in Section 11.3. We obtain

$$W(\mathbf{x}) = M_N^{-1} P(W, W_x, W_y, G, G_x, G_y, F_x, F_y, \mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{11.53}$$

$$W|_\Gamma = W(x, y, A) = 0, \tag{11.54}$$

where P is the N-D vector function,

$$P = (P_1, \ldots, P_N)^T (W, W_x, W_y, G_x, G_y, F_x, F_y, \mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{11.55}$$

$$\begin{aligned} &P_n(W, W_x, W_y, G_x, G_y, F_x, F_y, \mathbf{x}) \\ &= \int_0^1 \Psi_n(\alpha) \frac{\partial}{\partial \alpha} \left(- \int_z^{A+\sigma} \frac{u_{0x} + w_x + g_x}{2\sqrt{u_0 + w + g}} dt + f_x(x, y, A, \alpha) \right)^2 d\alpha \\ &+ \int_0^1 \Psi_n(\alpha) \frac{\partial}{\partial \alpha} \left(- \int_z^{A+\sigma} \frac{u_{0y} + w_y + g_y}{2\sqrt{u_0 + w + g}} dt + f_y(x, y, A, \alpha) \right)^2 d\alpha. \end{aligned} \tag{11.56}$$

Thus, we have obtained the desired boundary value problem (11.53)–(11.56) for the system of nonlinear coupled integral differential equations. Below we focus on this problem.

11.5.3 The positivity of the function $(u_0 + w + g)(\mathbf{x}, \alpha)$

It follows from (11.36) and (11.43) that we need the function $(u_0 + w + g)(\mathbf{x}, \alpha)$ to be positive. We discuss this issue in the current section.

Using (11.16), (11.30), (11.37), (11.43), and (11.49), define the set of functions K as

$$K = \{w(\mathbf{x}, \alpha) := (w + g)(\mathbf{x}, \alpha) > 0, \text{ (11.46) holds}, \ W \in C^1_{x,y,N,0}(\overline{\Omega})\}, \tag{11.57}$$

where $(\mathbf{x}, \alpha) \in \overline{\Omega} \times [0, 1]$. Then by (11.37) and (11.43),

$$(u_0 + w + g)(\mathbf{x}, \alpha) \geq \frac{A^2}{A^2 + 2}, \quad \forall w \in K, (\mathbf{x}, \alpha) \in \overline{\Omega} \times [0, 1]. \tag{11.58}$$

To obtain a sufficient condition guaranteeing (11.57) in terms of the vector function W, we formulate Lemma 11.5.1.

Lemma 11.5.1. *Let* (11.46) *and* (11.47) *hold. Consider the vector function* $v^N(\mathbf{x}) = (w_1 + g_1, w_2 + g_2, \ldots, w_N + g_N)^T(\mathbf{x})$. *Let* X_N *be the* $N \times N$ *matrix of Lemma* 11.3.1. *Consider the* *vector function* $\widetilde{v}^N(\mathbf{x}) = X_N^T \cdot v^N(\mathbf{x})$. *Let* $\widetilde{v}^N(\mathbf{x}) = (v_1, \dots, v_N)^T(\mathbf{x})$. *Suppose that all functions* $v_n(\mathbf{x}) > 0$ *in* $\overline{\Omega}$. *Then the function* $w \in K$ *and, therefore,* (11.58) *holds.Also, the set* K *is convex.*

Proof. The first part of this lemma follows immediately from Lemma 11.3.1. We now prove the convexity of the set K. Suppose that two functions $w^{(1)}, w^{(2)} \in K$. Let the number $\theta \in (0, 1)$. Then by (11.57),

$$\theta w^{(1)}(\mathbf{x}, \alpha) > -\theta g(\mathbf{x}, \alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1],$$
$$(1-\theta) w^{(2)}(\mathbf{x}, \alpha) > -(1-\theta) g(\mathbf{x}, \alpha), \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1].$$

Summing up these two inequalities, we obtain

$$\theta w^{(1)}(\mathbf{x}, \alpha) + (1-\theta) w^{(2)}(\mathbf{x}, \alpha) + g(\mathbf{x}, \alpha) > 0, \quad \mathbf{x} \in \overline{\Omega}, \alpha \in [0, 1]. \qquad \square$$

11.5.4 Applying the multidimensional analog of Taylor formula

We specify in this section how the classical multidimensional analog of Taylor formula [247] can be applied to the right-hand side of equation (11.53). Let $R > 0$ be an arbitrary number. Denote

$$K(R) = \left\{ W : w \in K, \|W\|_{C^1_{x,y,N}(\overline{\Omega})} < R \right\}. \tag{11.59}$$

It follows from Lemma 11.5.1 and (11.59) that $K(R)$ is a convex set.

Lemma 11.5.2. *Let* $W^{(1)}, W^{(2)} \in K(R)$, *let* $G^{(1)}$, $G^{(2)}$ *be the vector functions* (11.50) *and* $F^{(1)}, F^{(2)}$ *be the vector functions in* (11.52). *Based on* (11.46) *and* (11.49), *denote* $\widetilde{w}(\mathbf{x}, \alpha) = w^{(1)}(\mathbf{x}, \alpha) - w^{(2)}(\mathbf{x}, \alpha)$, $\widetilde{w}_n(\mathbf{x}) = w_n^{(1)}(\mathbf{x}) - w_n^{(2)}(\mathbf{x})$. *Similarly, denote* $\widetilde{g}(\mathbf{x}, \alpha) = g^{(1)}(\mathbf{x}, \alpha) - g^{(2)}(\mathbf{x}, \alpha)$, *where* $g^{(k)}(\mathbf{x}, \alpha)$ *corresponds to the vector function* $G^{(k)}(\mathbf{x})$, $k = 1, 2$ *via* (11.47), (11.50). *Also, denote*

$$\widetilde{W} = W^{(2)} - W^{(1)} = (\widetilde{w}_1, \dots, \widetilde{w}_N), \quad \widetilde{G} = G^{(1)} - G^{(2)}, \quad \widetilde{F} = F^{(2)} - F^{(1)}.$$

And let functions $f^{(1)} = f^{(1)}(x, y, A+\sigma)$ *and* $\widetilde{f} = \widetilde{f}(x, y, A+\sigma)$ *correspond to the vector functions* $F^{(1)}(\mathbf{x})$ *and* $\widetilde{F}(\mathbf{x})$, *respectively, via* (11.51), (11.52). *Then the following form of the multidimensional Taylor formula is valid:*

$$\begin{aligned}
& M_N^{-1} P(W^{(2)}, W_x^{(2)}, W_y^{(2)}, G_x^{(2)}, G_y^{(2)}, F_x^{(2)}, F_y^{(2)}, \mathbf{x}) \\
& \quad - M_N^{-1} P(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)} \mathbf{x}) \\
& = \int_z^{A+\sigma} T_0(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, x, y, t) \widetilde{W}(x, y, t) dt
\end{aligned}$$

$$
\begin{aligned}
&+ \int_z^{A+\sigma} T_1(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, x, y, t)\widetilde{W}_x(x,y,t)dt \\
&+ \int_z^{A+\sigma} T_2(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, x, y, t)\widetilde{W}_y(x,y,t)dt \\
&+ \int_z^{A+\sigma} T_3(W^{(i)}, W_x^{(i)}, W_y^{(i)}, G_x^{(i)}, G_y^{(i)}, F_x^{(i)}, F_y^{(i)}, \widetilde{W}, \widetilde{W}_x, \widetilde{W}_y, x, y, t)dt \qquad (11.60) \\
&+ \int_z^{A+\sigma} S_1(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, x, y, t)\widetilde{G}_x(x,y,t)dt \\
&+ \int_z^{A+\sigma} S_2(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, x, y, t)\widetilde{G}_y(x,y,t)dt \\
&+ S_3(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, \mathbf{x})\widetilde{F}_x(x,y) \\
&+ S_4(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}\mathbf{x})\widetilde{F}_y(x,y),
\end{aligned}
$$

where $\mathbf{x} \in \Omega$*,* $i = 1,2$ *in* T_3 *means that this matrix depends on both vector functions* $(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)})$ *and* $(W^{(2)}, W_x^{(2)}, W_y^{(2)}, G_x^{(2)}, G_y^{(2)}, F_x^{(2)}, F_y^{(2)})$*. Also, all elements of matrices* T_j*,* S_k*,* $j = 0,1,2,3$*;* $k = 1,2,3,4$ *are continuous functions of their variables. Furthermore, the following estimates hold:*

$$\left|T_j(W^{(i)}, W_x^{(i)}, W_y^{(i)}, G_x^{(i)}, G_y^{(i)}, F_x^{(i)}, F_y^{(i)}, x, y, t)\right| \le C_1; \quad j = 0,1,2; (x,y,t) \in \overline{\Omega}, \qquad (11.61)$$

$$\begin{aligned}&\left|T_3(W^{(i)}, W_x^{(i)}, W_y^{(i)}, G_x^{(i)}, G_y^{(i)}, F_x^{(i)}, F_y^{(i)}, \widetilde{W}, \widetilde{W}_x, \widetilde{W}_y, x, y, t)\right| \qquad (11.62)\\ &\quad \le C_1(\widetilde{W}^2 + \widetilde{W}_x^2 + \widetilde{W}_y^2), \quad (x,y,t) \in \overline{\Omega},\end{aligned}$$

$$\left|S_k(W^{(1)}, W_x^{(1)}, W_y^{(1)}, G_x^{(1)}, G_y^{(1)}, F_x^{(1)}, F_y^{(1)}, \mathbf{x})\right| \le C_1; \quad k = 1,\dots,4; \mathbf{x} \in \overline{\Omega}, \qquad (11.63)$$

where the number

$$C_1 = C_1\Big(N, R, \max_{i=1,2}\|G^{(i)}\|_{C^1_{x,y,N}(\overline{\Omega})}, \max_{i=1,2}\|F^{(i)}\|_{C^1_{x,y,N}(\overline{\Omega})}\Big) > 0 \qquad (11.64)$$

depends only on listed parameters. Estimates (11.61)–(11.64) *mean estimates for each element of corresponding matrices. In terms of the integration with respect to α, matrices* T_j*,* S_j*,* $j = 0,1,2,3$ *depend on the integrals of the form*

$$\int_0^1 \Psi_n(\alpha)\mu_{jnkm}(\mathbf{x}, (u_0, u_{0x}, u_{0y})(x,y,t,\alpha), \Psi_k(\alpha), \Psi_m(\alpha), \Psi_k'(\alpha), \Psi_m'(\alpha))d\alpha,$$

where $n,k,m = 1,\dots,N$*, and* μ_{jnkm} *are continuous functions of their variables.*

Proof. Below $C_1 > 0$ denotes different constants depending only on parameters listed in (11.64). For $\mathbf{x} \in \Omega$ and $\alpha \in (0,1)$, consider the functions $\xi(\mathbf{x},\alpha)$ and $\eta(\mathbf{x},\alpha)$ defined as

$$\xi(w,\mathbf{x},\alpha) = \frac{u_{0x} + w_x + g_x}{2\sqrt{u_0 + w + g}}(\mathbf{x},\alpha), \quad \eta(w,\mathbf{x},\alpha) = \frac{u_{0y} + w_y + g_y}{2\sqrt{u_0 + w + g}}(\mathbf{x},\alpha), \tag{11.65}$$

where functions w, g have the forms (11.46), (11.47). Then $\xi(w^{(2)},\mathbf{x},\alpha) = \xi(w^{(1)} + \widetilde{w},\mathbf{x},\alpha)$ and $\eta(w^{(2)},\mathbf{x},\alpha) = \eta(w^{(1)} + \widetilde{w},\mathbf{x},\alpha)$.

The convexity of the set $K(R)$ allows us to use the multidimensional analog of Taylor formula in the following form:

$$\begin{aligned}\xi(w^{(2)},\mathbf{x},\alpha) = \xi(w^{(1)},\mathbf{x},\alpha) &+ s_1(w^{(1)},\mathbf{x},\alpha)\widetilde{w} + s_2(w^{(1)},\mathbf{x},\alpha)\widetilde{w}_x \\ &+ s_3(w^{(1)},\mathbf{x},\alpha)\widetilde{g}_x(\mathbf{x},\alpha) + s_4(w^{(1)},w^{(2)},\mathbf{x},\alpha)\widetilde{w}^2 \\ &+ s_5(w^{(1)},w^{(2)},\mathbf{x},\alpha)\widetilde{w}_x^2 + s_6(w^{(1)},w^{(2)},\mathbf{x},\alpha)\widetilde{w}_x\widetilde{w},\end{aligned} \tag{11.66}$$

where s_j are continuous functions of their listed variables. Furthermore,

$$|s_i(w^{(1)},\mathbf{x},\alpha)|, |s_j(w^{(1)},w^{(2)},\mathbf{x},\alpha)| \le C_1, \quad \text{where } i = 1,2,3 \text{ and } j = 4,5,6. \tag{11.67}$$

Next, substituting (11.46), (11.47), and (11.50) in (11.66), we obtain

$$\begin{aligned}&\xi(w^{(2)},\mathbf{x},\alpha) \\ &\quad = \xi_1 + \sum_{k=1}^{N}[s_1\widetilde{w}_k(\mathbf{x}) + s_2\widetilde{w}_{kx}(\mathbf{x}) + s_3\widetilde{g}_{kx}(\mathbf{x})]\Psi_k(\alpha) \\ &\qquad + \sum_{k,m=1}^{N}[s_4(\widetilde{w}_k\widetilde{w}_m)(\mathbf{x}) + s_5(\widetilde{w}_{kx}\widetilde{w}_{mx})(\mathbf{x}) + s_6(\widetilde{w}_{kx}\widetilde{w}_m)(\mathbf{x})]\Psi_k(\alpha)\Psi_m(\alpha),\end{aligned} \tag{11.68}$$

where for brevity $\xi_1 = \xi(w^{(1)},\mathbf{x},\alpha)$. Hence, it follows from (11.65)–(11.68) that the second line of (11.56) can be rewritten as

$$\begin{aligned}&\int_0^1 \Psi_n(\alpha)\frac{\partial}{\partial\alpha}\left(-\int_z^{A+\sigma}\frac{u_{0x} + w_x + g_x}{2\sqrt{u_0 + w + g}}dt + f_x(x,y,A,\alpha)\right)^2 d\alpha \\ &\quad = \int_0^1 \Psi_n(\alpha) \\ &\qquad \times \frac{\partial}{\partial\alpha}\left(-\int_z^{A+\sigma}\left(\xi_1 + \sum_{k=1}^{N}\left(V_k + \sum_{m=1}^{N}V_{km}\Psi_m(\alpha)\right)\Psi_k(\alpha)\right)dt + (f_x^{(1)} + \widetilde{f}_x)\right)^2 d\alpha,\end{aligned} \tag{11.69}$$

$$V_k(w^{(1)},\mathbf{x},\alpha,\widetilde{w}_n) = s_1\widetilde{w}_k(\mathbf{x}) + s_2\widetilde{w}_{kx}(\mathbf{x}) + s_3\widetilde{g}_{kx}(\mathbf{x}), \tag{11.70}$$

$$V_{km}(w^{(1)},w^{(2)},\mathbf{x},\alpha) = s_4(\widetilde{w}_k\widetilde{w}_m)(\mathbf{x}) + s_5(\widetilde{w}_{kx}\widetilde{w}_{mx})(\mathbf{x}) + s_6(\widetilde{w}_{kx}\widetilde{w}_m)(\mathbf{x}). \tag{11.71}$$

Similar formulas are obviously valid for the third line of (11.56). Thus, formulas (11.56), (11.65)–(11.71) imply (11.60)–(11.64). □

In Lemma 11.5.2, we have used second order terms in the Taylor formula. In addition, we will also need the formula which uses only linear terms. The proof of Lemma 11.5.3 is completely similar to the proof of Lemma 11.5.2.

Lemma 11.5.3. *Assume that conditions of Lemma* 11.5.2 *hold. Then the following analog of the final increment formula is valid:*

$$\begin{aligned}
&M_N^{-1}P(W^{(2)},W_x^{(2)},W_y^{(2)},G_x^{(2)},G_y^{(2)},F_x^{(2)},F_y^{(2)},\mathbf{x})\\
&\quad - M_N^{-1}P(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)}\mathbf{x})\\
&= \int_z^{A+\sigma} Y_0(W^{(i)},W_x^{(i)},W_y^{(i)},G_x^{(i)},G_y^{(i)},F_x^{(i)},F_y^{(i)},x,y,t)\widetilde{W}(x,y,t)dt\\
&\quad + \int_z^{A+\sigma} Y_1(W^{(i)},W_x^{(i)},W_y^{(i)},G_x^{(i)},G_y^{(i)},F_x^{(i)},F_y^{(i)},x,y,t)\widetilde{W}_x(x,y,t)dt\\
&\quad + \int_z^{A+\sigma} Y_2(W^{(i)},W_x^{(i)},W_y^{(i)},G_x^{(i)},G_y^{(i)},F_x^{(i)},F_y^{(i)},x,y,t)\widetilde{W}_y(x,y,t)dt\\
&\quad + \int_z^{A+\sigma} \widehat{S}_1(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},x,y,t)\widetilde{G}_x(x,y,t)dt\\
&\quad + \int_z^{A+\sigma} \widehat{S}_2(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},x,y,t)\widetilde{G}_y(x,y,t)dt\\
&\quad + \widehat{S}_3(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},\mathbf{x})\widetilde{F}_x(x,y)\\
&\quad + \widehat{S}_4(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},\mathbf{x})\widetilde{F}_y(x,y), \quad \mathbf{x}\in\Omega.
\end{aligned} \tag{11.72}$$

where $\mathbf{x}\in\Omega$*,* $i=1,2$*, all elements of matrices* Y_j*,* $j=0,1,2$ *are continuous functions of their variables and the following estimates are valid for* $t\in[z,A+\sigma]$*,* (x,y,t)*,* $\mathbf{x}\in\overline{\Omega}$*:*

$$\begin{aligned}
&\left|\sum_{j=0}^{2} Y_j(W^{(i)},W_x^{(i)},W_y^{(i)},G_x^{(i)},G_y^{(i)},F_x^{(i)},F_y^{(i)},x,y,t)\right|\\
&\quad + \sum_{k=1}^{2}\left|\widehat{S}_k(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},x,y,t)\right|\\
&\quad + \sum_{k=3}^{4}\left|\widehat{S}_k(W^{(1)},W_x^{(1)},W_y^{(1)},G_x^{(1)},G_y^{(1)},F_x^{(1)},F_y^{(1)},\mathbf{x})\right| \le C_1.
\end{aligned} \tag{11.73}$$

11.6 Problem (11.53), (11.54) in the semidiscrete form

We now rewrite equation (11.53) in the form of finite differences with respect to variables x, y while keeping the continuous derivative with respect to z. For brevity, we keep the same grid step size $h > 0$ in both directions x, y. Consider partitions of the intervals $x \in (0,1)$, $y \in (0,1)$ in small subintervals of the same length h with $B = 1/h$ and corresponding semidiscrete subdomains of the domains Ω and $\overline{\Omega}$,

$$0 = x_0 < x_1 < \cdots < x_{B-1} < x_B = 1, \quad x_i - x_{i-1} = h,$$
$$0 = y_0 < y_1 < \cdots < y_{B-1} < y_B = 1, \quad y_i - y_{i-1} = h,$$
$$\Omega^h = \{\mathbf{x}^h(z) = \{(x_i, y_j, z)\}_{i,j=1}^{B-1}, z \in (A, A+\sigma)\}. \tag{11.74}$$
$$\Omega_1^h = \{\mathbf{x}^h(z) = \{(x_i, y_j, z)\}_{i,j=0}^{B}, z \in (A, A+\sigma)\}. \tag{11.75}$$

Hence, $\mathbf{x}^h(z)$ is a z-dependent vector. Its dimension is $(B-2)^2$ in the case of Ω^h and $(B+1)^2$ in the case of Ω_1^h. By (11.74), only those points $(x_i, y_j, z) \in \Omega^h$, which are corresponding interior points of the domain Ω. On the other hand, in addition to points of Ω^h, the semidiscrete domain Ω_1^h contains boundary points which belong to the part Γ of the boundary $\partial\Omega$. We assume below that

$$h \in [h_0, 1), \quad h_0 = \text{const.} \in (0,1). \tag{11.76}$$

Remark 11.6.1. Unlike classical forward problems for PDEs, we do not let the grid step size h tend to zero. This is typical for numerical methods for many inverse problems: due to their ill-posed nature, see, for example, [156, 236]. In other words, the grid step size is often used as the regularization parameter.

Consider the N-D vector function $Q(\mathbf{x}) = (Q_1, \ldots, Q_N)^T(\mathbf{x})$ with $Q_n \in C(\overline{\Omega})$. We denote $Q^h(\mathbf{x}^h(z))$ the trace of this vector function on the set $\overline{\Omega}_1^h$. Thus,

$$Q^h(\mathbf{x}^h(z)) = (Q_1^h, \ldots, Q_N^h)^T(\mathbf{x}^h(z)) = ((Q_1^{i,j}(z))_{i,j=0}^{B}, \ldots, (Q_N^{i,j}(z))_{i,j=0}^{B})^T \tag{11.77}$$

is the matrix depending on the variable $z \in [A, A+\sigma]$. Here, $Q_k^{i,j}(z) = Q_k(x_i, y_j, z)$, where $k = 1, \ldots, N$. Hence, by (11.46), (11.47), (11.51), and (11.77) the finite difference analogs of functions $w(\mathbf{x}, \alpha)$, $g(\mathbf{x}, \alpha)$, $f(x, y, A+\sigma)$ are

$$w^h(\mathbf{x}^h(z), \alpha) = \sum_{n=1}^{N} w_n^h(\mathbf{x}^h(z))\Psi_n(\alpha), \quad \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \alpha \in [0,1], \tag{11.78}$$
$$g^h(\mathbf{x}^h(z), \alpha) = \sum_{n=1}^{N} g_n^h(\mathbf{x}^h(z))\Psi_n(\alpha), \quad \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \alpha \in [0,1], \tag{11.79}$$
$$f^h(\mathbf{x}^h(A+\sigma), \alpha) = \sum_{n=1}^{N} f_n^h(\mathbf{x}^h(A+\sigma))\Psi_n(\alpha), \quad \mathbf{x}^h(A+\sigma) \in \overline{\Omega}_1^h, \alpha \in [0,1]. \tag{11.80}$$

Next, by (11.49), (11.50), (11.52), and (11.78)–(11.80) the finite difference analogs of matrices W, G, and F are

$$W^h(\mathbf{x}^h(z)) = (w_1^h, \ldots, w_N^h)^T(\mathbf{x}^h(z)), \quad \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \tag{11.81}$$

$$G^h(\mathbf{x}^h(z)) = (g_1^h,\dots,g_N^h)^T(\mathbf{x}^h(z)), \quad \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \tag{11.82}$$

$$F^h(\mathbf{x}^h(z)) = (f_1^h,f_2^h,\dots,f_N^h)^T(\mathbf{x}^h(z)), \quad z = A+\sigma, \mathbf{x}^h(A+\sigma) \in \overline{\Omega}_1^h. \tag{11.83}$$

For an arbitrary number $z \in [A, A+\sigma]$, denote $\Omega_z^h = \{\mathbf{x}^h(z) = (x_i, y_j, z)_{i,j=1}^{B-1}, z \text{ is fixed}\}$. We introduce semidiscrete functional spaces for matrices like $Q^h(\mathbf{x}(z))$,

$$C_N^h(\overline{\Omega}_z^h) = \Big\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{C_N^h(\overline{\Omega}_z^h)}(z) = \max_{k\in[1,N]} \max_{i,j=1,\dots,B-1} |Q_k^{i,j}(z)| < \infty\Big\},$$

$$C_N^h(\overline{\Omega}^h) = \Big\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{C_N^h(\overline{\Omega}^h)} = \max_{z\in[A,A+\sigma]} \|Q^h\|_{C_N^h(\overline{\Omega}_z^h)}(z) < \infty\Big\},$$

$$C_N^h(\overline{\Omega}_{1,z}^h) = \Big\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{C^h(\overline{\Omega}_{1,z}^h)}(z) = \max_{k\in[1,N]} \max_{i,j=0,\dots B} |Q_k^{i,j}(z)| < \infty\Big\},$$

$$C_N^h(\overline{\Omega}_1^h) = \Big\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{C_N^h(\overline{\Omega}_1^h)} = \max_{z\in[A,A+\sigma]} \|Q^h\|_{C_N^h(\overline{\Omega}_{1z}^h)}(z) < \infty\Big\},$$

$$C_{N,0}^h(\overline{\Omega}_1^h) = \{Q^h(\mathbf{x}^h(z)) \in C_N^h(\overline{\Omega}_1^h) : Q^h(\mathbf{x}^h(z)) = 0 \text{ for } \mathbf{x}^h(z) \in \Gamma \cup \{z = A\}\},$$

$$L_{2,N}^h(\Omega_z^h) = \left\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{L_2^h(\Omega_z^h)}^2(z) = h^2 \sum_{k=1}^{N} \sum_{i,j=1}^{B-1} (Q_k^{i,j}(z))^2 < \infty\right\},$$

$$L_{2,N}^h(\Omega^h) = \left\{Q^h(\mathbf{x}^h(z)) : \|Q^h\|_{L_{2,N}^h(\Omega^h)}^2 = \int_A^{A+\sigma} \|Q^h\|_{L_{2,N}^h(\Omega_\theta^h)}^2(z)dz < \infty\right\},$$

$$H_{0,N}^{1,h}(\Omega_1^h) = \left\{\begin{array}{l} Q^h(\mathbf{x}^h(z)) : Q^h(\mathbf{x}^h(z))|_{\Gamma\cup B_A} = 0, \\ \|Q^h\|_{H_{0,N}^{1,h}(\Omega_1^h)}^2 \\ \quad = \int_A^{A+\sigma} [\|Q^h\|_{L_{2,N}^h(\Omega_z^h)}^2(z) + \|\partial_z Q^h\|_{L_{2,N}^h(\Omega_z^h)}^2(z)]dz < \infty \end{array}\right\}.$$

We approximate x, y derivatives of the vector function $W(\mathbf{x})$ via central finite differences [228]. It is convenient to write this in the equivalent form for the matrix function $W^h(\mathbf{x})$ as

$$W_{k,x}^{i,j}(z) = \frac{W_k^{i+1,j}(z) - W_k^{i-1,j}(z)}{2h}; \quad i,j = 1,\dots,B-1, \tag{11.84}$$

$$W_{k,y}^{i,j}(z) = \frac{W_k^{i,j+1}(z) - W_k^{i,j-1}(z)}{2h}, \quad i,j = 1,\dots,B-1, \tag{11.85}$$

$$W_x^h(\mathbf{x}^h(z)) = ((W_{0,x}^{i,j}(z))_{i,j=0}^B,\dots,(W_{0,x}^{i,j}(z))_{i,j=0}^B)^T, \tag{11.86}$$

$$W_y^h(\mathbf{x}^h(z)) = ((W_{0,y}^{i,j}(z))_{i,j=0}^B,\dots,(W_{0,y}^{i,j}(z))_{i,j=0}^B)^T. \tag{11.87}$$

See (11.77) for notation (11.86) and (11.87). Using (11.76) and (11.84)–(11.87), we obtain that there exists a constant $C_2 = C_2(h_0, N, \Omega^h) > 0$ depending only on listed parameters such that

$$\|Q_x^h\|_{C_N^h(\overline{\Omega}^h)}, \|Q_y^h\|_{C_N^h(\overline{\Omega}^h)} \le C_2 \|Q^h\|_{C_N^h(\overline{\Omega}_1^h)}, \quad \forall Q^h \in C_N^h(\overline{\Omega}_1^h). \tag{11.88}$$

In addition, by embedding theorem $H^{1,h}_{0,N}(\Omega^h_1) \subset C^h_{N,0}(\overline{\Omega}^h_1)$ and there exists a constant $C = C(A, \sigma, h_0, N) > 0$ such that

$$\|Q^h\|_{C^h_N(\overline{\Omega}^h_1)} \leq C\|Q^h\|_{H^{1,h}_{0,N}(\Omega^h_1)}, \quad \forall Q^h \in H^{1,h}_{0,N}(\Omega^h_1). \tag{11.89}$$

Thus, using (11.77)–(11.83), we obtain the following finite difference analog of problem (11.53), (11.54):

$$W^h(\mathbf{x}^h(z)) = M_N^{-1} P(W^h, W^h_x, W^h_y, G^h, G^h_x, G^h_y, F^h, F^h_x, F^h_y, \mathbf{x}^h(z)), \quad \mathbf{x}^h(z) \in \Omega^h, \tag{11.90}$$

$$W^h(\mathbf{x}^h(z)) \in H^{1,h}_{0,N}(\Omega^h_1). \tag{11.91}$$

Also, we assume that the vector functions $G^h(\mathbf{x}^h(z))$ and $F^h(\mathbf{x}^h(z))$ in (11.82) and (11.83) are such that

$$G^h(\mathbf{x}^h(z)), F^h(\mathbf{x}^h(z)) \in C^h_N(\overline{\Omega}^h_1). \tag{11.92}$$

As above, let $R > 0$ be an arbitrary number. We now introduce the finite difference analogs of sets (11.57) and (11.59). Assume that (11.79) holds. Then

$$K^h = \left\{ \begin{array}{l} w^h = w^h(\mathbf{x}^h(z), \alpha) := (w^h + g^h)(\mathbf{x}^h(z), \alpha) > 0, \\ \forall (\mathbf{x}^h(z), \alpha) \in \overline{\Omega}^h_1 \times [0,1], \text{ (11.78) holds for } w^h(\mathbf{x}^h(z), \alpha), \\ W^h(\mathbf{x}^h(z)) \in H^{1,h}_{0,N}(\Omega^h_1) \end{array} \right\}, \tag{11.93}$$

$$K^h(R) = \{W^h(\mathbf{x}^h(z)) : w^h \in K^h, \|W^h\|_{H^{1,h}_{0,N}(\Omega^h_1)} < R\}. \tag{11.94}$$

It follows from (11.89), (11.93), and (11.94) that $K^h(R) \subset K^h \subset C^h_{N,0}(\overline{\Omega}^h_1)$.

Similarly, with Lemmata 11.3.1 and 11.5.1, Lemma 11.6.1 provides a sufficient condition imposed on the components of the matrix $w^h(\mathbf{x}^h(z), \alpha)$, which guarantees that $w^h(\mathbf{x}^h(z), \alpha) \in K^h$.

Lemma 11.6.1. *Let the matrix $w^h = w^h(\mathbf{x}^h(z), \alpha) \in K^h$. Select a triple (i, j, z) with $i, j = 0, \dots, B$, $z \in [A, A + \sigma]$ and consider the vector $v(i, j, z) = (w_1 + g_1, w_2 + g_2, \dots, w_N + g_N)^T(i, j, z)$. Let X_N be the $N \times N$ matrix of Lemma* 11.3.1. *Consider the vector $\widetilde{v}(i, j, z) = X_N^T \cdot v(i, j, z)$. Let $\widetilde{v}(i, j, z) = (\widetilde{v}_1, \dots, \widetilde{v}_N)^T(i, j, z)$. Suppose that all numbers $\widetilde{v}_n(i, j, z) > 0$ for all $i, j = 0, \dots, B$, $z \in [A, A + \sigma]$. Then the function $w^h \in K$ and, therefore, by* (11.37) *and* (11.43) *the following analog of* (11.58) *holds for $\mathbf{x}^h(A + \sigma), \mathbf{x}^h(z) \in \overline{\Omega}^h_1$:*

$$u^h_0(\mathbf{x}(A + \sigma)) + w^h(\mathbf{x}^h(z), \alpha) + g^h(\mathbf{x}^h(z), \alpha) > \frac{A^2}{A^2 + \_2}.$$

Also, the set $K^h(R)$ is convex.

Proof. The first part of this lemma, the one about the positivity, follows immediately from Lemma 11.3.1. Consider now two matrices $w^{1,h}(\mathbf{x}^h(z),\alpha), w^{2,h}(\mathbf{x}^h(z),\alpha) \in K^h(R)$. Let the number $\theta \in [0,1]$. Then one can prove completely similarly with the proof of Lemma 11.5.1 that $\theta w^{1,h}(\mathbf{x}^h(z),\alpha) + (1-\theta)w^{2,h}(\mathbf{x}^h(z),\alpha) \in K^h$. Let $W^{1,h}(\mathbf{x}^h(z))$ and $W^{2,h}(\mathbf{x}^h(z))$ be two matrices corresponding to matrices $w^{1,h}(\mathbf{x}^h(z),\alpha)$ and $w^{2,h}(\mathbf{x}^h(z),\alpha)$, respectively, via (11.81). The triangle inequality and (11.94) imply that

$$\begin{aligned}&\|\theta W^{1,h}(\mathbf{x}^h(z)) + (1-\theta)W^{2,h}(\mathbf{x}^h(z),\alpha)\|_{H^{1,h}_{0,N}(\Omega_1^h)}\\ &\quad\le \theta\|W^{1,h}(\mathbf{x}^h(z))\|_{H^{1,h}_{0,N}(\Omega_1^h)} + (1-\theta)\|W^{2,h}(\mathbf{x}^h(z))\|_{H^{1,h}_{0,N}(\Omega_1^h)}\\ &\quad< \theta R + (1-\theta)R = R.\end{aligned}$$

□

Lemma 11.6.2 is a finite difference analog of Lemmata 11.5.2 and 11.5.3. The proof is completely similar and is therefore omitted.

Lemma 11.6.2. *Assume that* (11.92) *holds. Then the direct analogs of Lemmata* 11.5.2 *and* 11.5.3*, being applied to the right-hand side of* (11.90)*, are true, provided that all functions involved in these lemmata are replaced with their above semidiscrete analogs. The constant C_1 in* (11.64) *and* (11.73) *should be replaced with the constant $\widetilde{C}_1$ depending only on listed parameters, where*

$$\widetilde{C}_1 = \widetilde{C}_1\Big(h_0, N, R, \max_{i=1,2}\|G^{(i)}\|_{C_{x,y,N}(\overline{\Omega})}, \max_{i=1,2}\|F^{(i)}\|_{C^1_{x,y,N}(\overline{\Omega})}\Big) > 0.$$

Suppose that we have found such a matrix $W^h(\mathbf{x}^h(z)) \in K^h(R)$, that it solves problem (11.90), (11.91). Then, using (11.38), (11.42), (11.43), and (11.78)–(11.80), we set

$$w^{i,j}(z,\alpha) = \sum_{n=1}^{N} W_n^{i,j}(z)\Psi_n(\alpha); \quad i,j = 0,\dots,B, z \in [A, A+\sigma], \alpha \in (0,1), \tag{11.95}$$

$$w^h(\mathbf{x}^h(z),\alpha) = (w^{i,j}(z,\alpha))_{i,j=0}^{B}; \quad z \in [A, A+\sigma], \alpha \in (0,1), \tag{11.96}$$

$$v^h(\mathbf{x}^h(z),\alpha) = w^h(\mathbf{x}^h(z),\alpha) + g^h(\mathbf{x}^h(z),\alpha), \quad \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \alpha \in (0,1) \tag{11.97}$$

$$u^h(\mathbf{x}^h(z),\alpha) = u_0^h(\mathbf{x}^h(z),\alpha) + v^h(\mathbf{x}^h(z),\alpha). \tag{11.98}$$

Hence, by (11.37), (11.93), and (11.94),

$$u^h(\mathbf{x}^h(z),\alpha) > \frac{A^2}{A^2+2} > 0, \quad \forall \mathbf{x}^h(z) \in \overline{\Omega}_1^h, \forall \alpha \in [0,1]. \tag{11.99}$$

Using (11.30), (11.31), and (11.99), we set

$$\tau_z^h(\mathbf{x}^h(z),\alpha) = \sqrt{u^h(\mathbf{x}^h(z),\alpha)}. \tag{11.100}$$

The semidiscrete analogs of formulas (11.33) and (11.34) are

$$\tau_x^h(\mathbf{x}^h(z),\alpha) = -\int_z^{A+\sigma}\left(\frac{u_x^h}{2\sqrt{u^h}}\right)(\mathbf{x}^h(t),\alpha)dt + f_x^h(x,y,A+\sigma,\alpha), \quad \mathbf{x}^h(z)\in\Omega, \tag{11.101}$$

$$\tau_y^h(\mathbf{x}^h(z),\alpha) = -\int_z^{A+\sigma}\left(\frac{u_y^h}{2\sqrt{u^h}}\right)(\mathbf{x}^h(t),\alpha)dt + f_y^h(x,y,A+\sigma,\alpha), \quad \mathbf{x}^h(z)\in\Omega. \tag{11.102}$$

Next, using the original eikonal equation (11.12), we set its semidiscrete form as

$$[(\tau_x^h)^2 + (\tau_y^h)^2 + (\tau_z^h)^2](\mathbf{x}^h(z)) = m^h(\mathbf{x}^h(z)), \quad \mathbf{x}^h(z)\in\overline{\Omega}^h. \tag{11.103}$$

Remarks 11.6.1.

1. Equation (11.90) with condition (11.91) as well conditions (11.76), (11.78)–(11.80), and the assumption that the right-hand side of (11.103) is independent on the parameter α form our approximate mathematical model for the TTTP formulated in Section 11.2.
2. It is well known that the problems like proving the convergence of the numerical methods as ours when in (11.76), (11.78)–(11.80) actual regularization parameters $h_0 \to 0$ and $N \to \infty$ are, generally, extremely challenging ones in the field of inverse problems. The fundamental underlying reason of these challenges is the ill-posed nature of inverse problems. Therefore, we do not analyze this type of convergence here.

11.7 Lipschitz stability and uniqueness

Based on (11.56) as well as on (11.88) and (11.89), we consider everywhere below the matrix equation (11.90), as a system of coupled nonlinear Volterra integral equations whose solution satisfies (11.91). Denote

$$\Phi^h(\mathbf{x}^h(z)) = (G^h, G_x^h, G_y^h, F^h, F_x^h, F_y^h)(\mathbf{x}^h(z)); \quad \mathbf{x}^h(z)\in\overline{\Omega}^h,\ G^h, F^h \in C_N^h(\overline{\Omega}_1^h). \tag{11.104}$$

Theorem 11.7.1 (Lipschitz stability and uniqueness). *Let $R > 0$ be an arbitrary number. Consider two matrices $W^{(1),h}(\mathbf{x}^h(z)), W^{(2),h}(\mathbf{x}^h(z)) \in K^h(R)$. Suppose that these matrices generate two pairs of matrices in* (11.82), (11.83) *$G^{(i),h}(\mathbf{x}^h(z))$ and $F^{(i),h}(\mathbf{x}^h(A+\sigma))$, $i = 1,2$, which satisfy conditions* (11.92). *Let $\Phi^{(i),h}(\mathbf{x}^h(z))$, $i = 1,2$ be the corresponding matrices as in* (11.104). *Let $m^{(1),h}(\mathbf{x}(z))$ and $m^{(2),h}(\mathbf{x}(z))$ be the corresponding right-hand sides of equality* (11.103). *Assume that functions $m^{(1),h}(\mathbf{x}(z))$ and $m^{(2),h}(\mathbf{x}(z))$ are independent on α (Remarks* 11.6.1). *Then there exists a constant*

$$C_3 = C_3(h_0, A, \sigma, R, N, \Omega^h, \|\Phi^{(1),h}\|_{C_{2N}^h(\overline{\Omega}_1^h)}, \|\Phi^{(2),h}\|_{C_{2N}^h(\overline{\Omega}_1^h)}) > 0, \tag{11.105}$$

depending only on listed parameters such that the following Lipschitz stability estimate holds:

$$\|m^{(1),h} - m^{(2),h}\|_{C_1^h(\overline{\Omega}^h)} \leq C_3 \|\Phi^{(1),h} - \Phi^{(2),h}\|_{C_{2N}^h(\overline{\Omega}_1^h)}. \tag{11.106}$$

In particular, if $\Phi^{(1),h}(\mathbf{x}^h(z)) \equiv \Phi^{(2),h}(\mathbf{x}^h(z))$*, then* $m^{(1),h}(\mathbf{x}^h(z)) \equiv m^{(2),h}(\mathbf{x}^h(z))$*, that is, uniqueness holds.*

Proof. In this proof, $C_3 > 0$ denotes different constants depending on parameters listed in (11.105). Denote

$$\widetilde{W}^h = W^{(2),h} - W^{(1),h}, \quad \widetilde{\Phi}^h = \Phi^{(2),h} - \Phi^{(1),h}. \tag{11.107}$$

It follows from Lemma 11.6.2, equality (11.72), estimate (11.73) of Lemma 11.5.3, (11.88), (11.90), (11.91), (11.95)–(11.99), and (11.104) that the following inequality with the Volterra integral holds true:

$$|\widetilde{W}^h(\mathbf{x}^h(z))| \leq C_3 \int_z^{A+\sigma} \|\widetilde{W}^h(\mathbf{x}^h(t))\|_{C_N^h(\overline{\Omega}_t^h)} dt + C_3 \|\widetilde{\Phi}^h\|_{C_{2N}^h(\overline{\Omega}_1^h)},$$

where $\mathbf{x}(z) \in \overline{\Omega}_1^h$ and $z \in [A, A+\sigma]$. Hence, Gronwall's inequality leads to

$$\|\widetilde{W}^h\|_{C_N^h(\overline{\Omega}_1^h)} \leq C_3 \|\widetilde{\Phi}^h\|_{C_{2N}^h(\overline{\Omega}_1^h)}.$$

This is the key estimate of this proof. Having this estimate, the target estimate (11.106) follows immediately from (11.81)–(11.83), (11.88), (11.95)–(11.103), (11.104), and (11.107). Uniqueness follows from (11.106). □

11.8 Weighted globally strictly convex Tikhonov-like functional

11.8.1 Estimating an integral

Let $\lambda > 0$ be the parameter to be chosen later. We choose the "integral analog" of the CWF as

$$\varphi_\lambda(z) = e^{2\lambda z}. \tag{11.108}$$

Lemma 11.8.1. *The following estimate holds for all* $\lambda > 0$ *and for every function* $p \in L_1(A, A+\sigma)$:

$$\int_A^{A+\sigma} e^{2\lambda z} \left(\int_z^{A+\sigma} |p(y)| dy \right) dz \leq \frac{1}{2\lambda} \int_A^{A+\sigma} |p(z)| e^{2\lambda z} dz.$$

Proof. Interchanging the integrals, we obtain

$$I = \int_A^{A+\sigma} e^{2\lambda z}\left(\int_z^{A+\sigma} |p(y)|dy\right)dz = \int_A^{A+\sigma} |p(y)|\left(\int_A^{y} e^{2\lambda z}dz\right)dy$$
$$= \frac{1}{2\lambda}\int_A^{A+\sigma} |p(y)|(e^{2\lambda y} - e^{2\lambda A})dy \le \frac{1}{2\lambda}\int_A^{A+\sigma} |p(y)|e^{2\lambda y}dy. \qquad \square$$

11.8.2 The functional

To solve problem (11.90), (11.91) numerically, we consider the following minimization problem.

Minimization problem. Fix an arbitrary number $R > 0$ as well as the gird step size $h \in [h_0, 1)$. Let $\gamma > 0$ be the regularization parameter. Minimize the functional $J_{\lambda,\gamma}(W^h)$ on the closed set $\overline{K^h(R)}$, where

$$J_{\lambda,\gamma}(W^h) = e^{-2\lambda A}\|[W^h(\mathbf{x}^h(z)) - M_N^{-1}P(W^h, W_x^h, W_y^h, \Phi^h, \mathbf{x}^h(z))]e^{\lambda z}\|^2_{L_2^h(\Omega^h)} + \gamma\|W^h\|^2_{H_0^{1,h}(\Omega_1^h)}. \tag{11.109}$$

Here, we took into account (11.108). We use the multiplier $e^{-2\lambda A}$ in order to balance two terms in the right-hand side of (11.109). Note that since $R > 0$ is an arbitrary number, then we do not impose a smallness condition on the set $\overline{K^h(R)}$ where we search for the solution of problem (11.90), (11.91). This is why we are talking below about the global strict convexity and the globally convergent numerical method.

Theorem 11.8.1 (global strict convexity). *Let h_0 be the number defined in* (11.76) *and let $h \in [h_0, 1)$. At every point $W^h \in K^h(2R)$ and for all $\lambda > 0, \gamma \in (0,1)$ the functional $J_{\lambda,\gamma}(W^h)$ has the Fréchet derivative $J'_{\lambda,\gamma}(W^h) \in H_0^{1,h}(\Omega_1^h)$. Furthermore, this derivative is Lipschitz continuous on $K^h(2R)$, that is, there exists a constant*

$$C_4 = C_4(h_0, A, \sigma, R, N, \Omega^h, \|G^h\|_{C_N^h(\overline{\Omega}_1^h)}, \|F^h\|_{C_N^h(\overline{\Omega}_1^h)}) > 0 \tag{11.110}$$

depending only on parameters listed in (11.110) *such that for all $W^{(1),h}, W^{(2),h} \in K^h(2R)$*

$$\|J'_{\lambda,\gamma}(W^{(2),h}) - J'_{\lambda,\gamma}(W^{(1),h})\|_{H_0^{1,h}(\Omega_1^h)} \le \overline{C}\|W^{(2),h} - W^{(1),h}\|_{H_0^{1,h}(\Omega_1^h)}, \tag{11.111}$$

where the constant $\overline{C} > 0$ depends on the same parameters as ones listed in (11.110) *as well as on λ. In addition, there exists a sufficiently large number $\lambda_0 > 1$ depending on* *the same parameters as those listed in* (11.110) *such that for every* $\lambda \geq \lambda_0$ *and for every* $\gamma \in (0,1)$ *the functional* $J_{\lambda,\gamma}(W^h)$ *is strictly convex on the closed set* $\overline{K^h(R)}$. *More precisely, the following estimate holds for all* $W^{(1),h}, W^{(2),h} \in \overline{K^h(R)}$:

$$\begin{aligned} J_{\lambda,\gamma}(W^{(2),h}) - J_{\lambda,\gamma}(W^{(1),h}) - J'_{\lambda,\gamma}(W^{(1),h})(W^{(2),h} - W^{(1),h}) \\ \geq \frac{1}{8}\|W^{(2),h} - W^{(1),h}\|^2_{L_2^h(\Omega^h)} + \gamma\|W^{(2),h} - W^{(1),h}\|^2_{H_0^{1,h}(\Omega_1^h)}. \end{aligned} \tag{11.112}$$

Below C_4 denotes different constants depending on parameters listed in (11.110). Theorem 11.8.2 follows immediately from (11.111), (11.112), and Lemma 5.2.1.

Theorem 11.8.2. *Suppose that conditions of Theorem* 11.8.1 *are in place. Then for every* $\lambda \geq \lambda_0$ *and for every* $\gamma \in (0,1)$ *there exists a single minimizer* $W^h_{\min} \in \overline{K^h(R)}$ *of the functional* $J_{\lambda,\gamma}(W^h)$ *on the set* $\overline{K^h(R)}$. *Furthermore,*

$$J'_{\lambda,\gamma}(W^h_{\min})(W^h - W^h_{\min}) \geq 0, \quad \forall W^h \in \overline{K^h(R)}. \tag{11.113}$$

Let $P_{\overline{K^h(R)}} : H_0^{1,h}(\Omega_1^h) \to \overline{K^h(R)}$ be the orthogonal projection operator of the space $H_0^{1,h}(\Omega_1^h)$ on the closed set $\overline{K^h(R)} \subset H_0^{1,h}(\Omega_1^h)$. Consider an arbitrary point $W^{0,h} \in K^h(R)$. And minimize the functional $J_{\lambda,\gamma}(W^h)$ by the gradient projection method, which starts its iterations at the point $W^{0,h}$,

$$W^h_n = P_{\overline{K^h(R)}}(W^h_{n-1} - \rho J'_{\lambda,\alpha}(W^h_{n-1})), \quad n = 1, 2, \ldots. \tag{11.114}$$

Here, the number $\rho \in (0,1)$. Theorem 11.8.3 follows from a combination of Theorems 11.8.1 and 11.8.2 with Theorem 5.2.1.

Theorem 11.8.3. *Suppose that conditions of Theorem* 11.8.1 *are in place,* $\lambda \geq \lambda_0$ *and* $\gamma \in (0,1)$. *Let* $W^h_{\min} \in \overline{K^h(R)}$ *be the minimizer listed in Theorem* 11.8.2. *Then there exists a sufficiently small number* $\rho_0 \in (0,1)$ *depending on the same parameters as ones listed in* (11.110) *such that for every* $\rho \in (0,\rho_0)$ *there exists a number* $\eta = \eta(\rho) \in (0,1)$ *such that the sequence* $W^{0,h} \in K^h(R)$ *converges to* $W^h_{\min}$,

$$\|W^h_{\min} - W^h_n\|_{H^{1,h}_{0,N}(\Omega_1^h)} \leq \eta^n \|W^h_{\min} - W^h_0\|_{H^{1,h}_{0,N}(\Omega_1^h)}. \tag{11.115}$$

Recall that in the regularization theory, the minimizer $W^h_{\min}$ of functional (11.109) is called "regularized solution"; see, for example, [22, 76, 244]. We now need to show that regularized solutions converge to the exact one when the noise in the data tends to zero. Following the regularization theory, we assume that there exists an exact, that is, idealized, solution $W^{*,h} \in K^h(R)$ of problem (11.90), (11.91) with the noiseless data

$\Phi^{*,h}(\mathbf{x}(z))$,

$$\Phi^{*,h}(\mathbf{x}(z)) = (G^{*,h}, G_x^{*,h}, G_y^{*,h}, F^{*,h}, F_x^{*,h}, F_y^{*,h})(\mathbf{x}^h(z)); \quad G^{*,h}, F^{*,h} \in C_N^h(\overline{\Omega}_1^h), \tag{11.116}$$

where $\mathbf{x}^h(z) \in \overline{\Omega}^h$, see (11.104). We assume that there exists the exact, that is, idealized function $m^{*,h}(\mathbf{x}^h(z))$, $\mathbf{x}^h(z) \in \overline{\Omega}^h$ in (11.103), which produces the data (11.116).

Let the number $\delta \in (0,1)$ be the level of the error in the data G^h, F^h, that is,

$$\|G^{*,h} - G^h\|_{C_N^h(\overline{\Omega}_1^h)}, \|F^{*,h} - F^h\|_{C_N^h(\overline{\Omega}_1^h)} < \delta. \tag{11.117}$$

Denote $\widetilde{G}^h = G^{*,h} - G^h$, $\widetilde{F}^h = F^{*,h} - F^h$. Then (11.88) and (11.117) imply that with a constant $C_2 = C_2(h_0, N, \Omega^h) > 0$ depending only on listed parameters the following inequalities hold:

$$\|\widetilde{G}^h\|_{C_N^h(\overline{\Omega}_1^h)}, \|\widetilde{G}_x^h\|_{C_N^h(\overline{\Omega}^h)}, \|\widetilde{G}_y^h\|_{C_N^h(\overline{\Omega}^h)} < C_2\delta, \tag{11.118}$$

$$\|\widetilde{F}^h\|_{C_N^h(\overline{\Omega}_1^h)}, \|\widetilde{F}_x^h\|_{C_N^h(\overline{\Omega}^h)}, \|\widetilde{F}_y^h\|_{C_N^h(\overline{\Omega}^h)} < C_2\delta. \tag{11.119}$$

Since $\delta \in (0,1)$ and (11.117)–(11.119) hold, then, using the triangle inequality, we replace below the dependence of the constant $C_4 > 0$ on $\|\Phi^h\|_{C_{2N}^h(\overline{\Omega}_1^h)}$ in (11.110) with the dependence of C_4 on $\|\Phi^{*,h}\|_{C_{2N}^h(\overline{\Omega}_1^h)}$. Thus, everywhere below $C_4 > 0$ denotes different constants depending on the same parameters as ones listed in (11.110), except that $\|\Phi^h\|_{C_{2N}^h(\overline{\Omega}_1^h)}$ is replaced with $\|\Phi^{*,h}\|_{C_{2N}^h(\overline{\Omega}_1^h)}$.

Consider now the right-hand side of equation (11.90) for the case when W^h is replaced with $W^{*,h}$, whereas other arguments remain the same. By Lemma 11.6.2, we can use finite difference analogs of (11.72) and (11.73). In addition, we use now (11.116)–(11.119). Thus, we obtain

$$M_N^{-1} P(W^{*,h}, W_x^{*,h}, W_y^{*h}, \Phi^h, \mathbf{x}^h(z))$$
$$= M_N^{-1} P(W^{*,h}, W_x^{*,h}, W_y^{*h}, \Phi^{*,h}, \mathbf{x}^h(z)) + \widehat{P}(\mathbf{x}^h(z)), \mathbf{x}^h(z) \in \overline{\Omega}^h, \tag{11.120}$$
$$\|\widehat{P}\|_{C_N^h(\overline{\Omega}^h)} \leq C_4\delta. \tag{11.121}$$

Theorem 11.8.4. *Suppose that conditions of Theorem* 11.8.1 *are in place and also that there exists an exact solution* $W^{*,h} \in K^h(R)$ *of problem* (11.90), (11.91) *with the noiseless data* $\Phi^{*,h}$ *as in* (11.116). *Let the* $\lambda_0 > 1$ *be the number chosen in Theorem* 11.8.1. *Fix an arbitrary number* $\lambda = \lambda_1 \geq \lambda_0$ *in the functional* $J_{\lambda,\gamma}(W^h) = J_{\lambda_1,\gamma}(W^h)$. *Just as in the regularization theory, set* $\gamma = \gamma(\delta) = \delta^2$. *Let the numbers* $\rho \in (0, \rho_0)$, $\rho_0, \eta = \eta(\rho) \in (0,1)$ *be the same as in Theorem* 11.8.3. *Then the following estimates are valid:*

$$\|W^{h,*} - W_{\min}^h\|_{L_{2,N}^h(\Omega^h)} \leq C_4\delta, \tag{11.122}$$

$$\|W^{*,h} - W_n^h\|_{L_{2,N}^h(\Omega^h)} \le C_4\delta + \eta^n \|W_{\min}^h - W_0^h\|_{H_{0,N}^{1,h}(\Omega_1^h)}, \tag{11.123}$$

$$\|m^{*,h} - m_n^h\|_{L_{2,1}^h(\Omega^h)} \le C_4\delta + \eta^n \|W_{\min}^h - W_0^h\|_{H_{0,N}^{1,h}(\Omega_1^h)}, \tag{11.124}$$

where the functions $m_n^h(\mathbf{x}^h(z))$ are constructed from matrices W_n^h via the procedure described in Section 11.6 *with the final formula* (11.103).

Remark 11.8.1. Since $W^{0,h} \in K^h(R)$ is an arbitrary point and $R > 0$ is an arbitrary number, then Theorem 11.8.4 implies the global convergence of the gradient projection method (11.114); see Definition 1.4.2.

It follows from the above that we need to prove only Theorems 11.8.1 and 11.8.4.

11.9 Proofs of Theorems 11.8.1 and 11.8.4

11.9.1 Proof of Theorem 11.8.1

Denote (,) and [,] the scalar products in the spaces $L_{2,N}^h(\Omega^h)$ and $H_{0,N}^{1,h}(\Omega_1^h)$, respectively. Let $W^{(1),h}$ and $W^{(2),h}$ be two arbitrary points of the set $K^h(R)$. As above, denote $\widetilde{W}^h = W^{(2),h} - W^{(1),h}$. Hence, $W^{(2),h} = W^{(1),h} + \widetilde{W}^h$. Also, by (11.94),

$$\|\widetilde{W}^h\|_{H_{0,N}^{1,h}(\Omega_1^h)} < R. \tag{11.125}$$

Hence, by (11.89) and (11.125),

$$\|\widetilde{W}^h\|_{C_N^h(\overline{\Omega}^h)} \le 2CR. \tag{11.126}$$

It follows from Lemma 11.6.2 and (11.60) that

$$\begin{aligned}
M_N^{-1}P&\big(W^{(2),h}, W_x^{(2),h}, W_y^{(2),h}, \Phi^h, \mathbf{x}^h(z)\big)\\
&= M_N^{-1}P\big(W^{(1),h}, W_x^{(1),h}, W_y^{(1),h}, \Phi^h, \mathbf{x}^h(z)\big)\\
&\quad + \int_z^{A+\sigma} T_0\big(W^{(1),h}, W_x^{(1),h}, W_y^{(1),h}, \Phi^h, x, y, t\big)\widetilde{W}^h(\mathbf{x}^h(t))\,dt\\
&\quad + \int_z^{A+\sigma} T_1\big(W^{(1),h}, W_x^{(1),h}, W_y^{(1),h}, \Phi^h, x, y, t\big)\widetilde{W}_x^h(\mathbf{x}^h(t))\,dt\\
&\quad + \int_z^{A+\sigma} T_2\big(W^{(1),h}, W_x^{(1),h}, W_y^{(1),h}, \Phi^h, x, y, t\big)\widetilde{W}_y^h(\mathbf{x}^h(t))\,dt\\
&\quad + \int_z^{A+\sigma} T_3\big(W^{(i),h}, W_x^{(i),h}, W_y^{(i),h}, \Phi^h, \widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h, \mathbf{x}^h(t)\big)\,dt, \quad \mathbf{x}^h(z) \in \overline{\Omega}^h,
\end{aligned} \tag{11.127}$$

where $i = 1, 2$ and T_j, $j = 0, 1, 2, 3$ are continuous functions of their variables for $W^{(i),h} \in \overline{D^h(R)}$. In addition, by (11.62) and (11.126),

$$\int_z^{A+\sigma} \left|T_3(W^{(i),h}, W_x^{(i),h}, W_y^{(i),h}, \Phi^h, \widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h, \mathbf{x}^h(t))\right| \leq C_4 \int_z^{A+\sigma} ((\widetilde{W}^h)^2 + (\widetilde{W}_x^h)^2 + (\widetilde{W}_y^h)^2)(\mathbf{x}^h(t))dt, \quad \mathbf{x}^h(z) \in \overline{\Omega}^h. \tag{11.128}$$

Denote

$$D_1^h(\mathbf{x}^h(z)) = W^{(1),h}(\mathbf{x}^h(z)) - M_N^{-1}P(W^{(1),h}, W_x^{(1),h}, W_y^{(1),h}, \Phi^h, \mathbf{x}^h(z)), \tag{11.129}$$

$$D_2^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h, \mathbf{x}^h(z)) = \text{the sum of lines number 3, 4, 5 of (11.127)}, \tag{11.130}$$

$$D_3^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h, \mathbf{x}^h(z)) = \text{the line number 6 of (11.127)}. \tag{11.131}$$

Then it follows from (11.109) and (11.127)–(11.131) that

$$\begin{aligned} &J_{\lambda,\gamma}(W^{(2),h}) - J_{\lambda,\gamma}(W^{(1),h}) \\ &\quad = 2e^{-2\lambda(A+\sigma)}(D_1^h, \widetilde{W}^h - D_2^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)e^{2\lambda z}) + 2\gamma[\widetilde{W}^h, W^{(1),h}] \\ &\qquad - 2e^{-2\lambda A}(D_1^h, D_3^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)e^{2\lambda z}) \\ &\qquad + e^{-2\lambda A}\|[\widetilde{W}^h - D_2^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h) - D_3^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)]e^{\lambda z}\|^2_{L^h_{2,N}(\Omega^h)} \\ &\qquad + \gamma\|\widetilde{W}^h\|^2_{H^{1,h}_{0,N}(\Omega_1^h)}. \end{aligned} \tag{11.132}$$

In addition, by Lemma 11.6.2, (11.60), (11.61), (11.89), Lemma 11.8.1, (11.126), and (11.128)–(11.131) the following estimates hold:

$$\begin{aligned} &\left|-2(D_1^h, D_3^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)e^{2\lambda z})\right|e^{-2\lambda A} \\ &\quad \leq C_4 e^{-2\lambda A} \int_A^{A+\sigma} \left(\int_z^{A+\sigma} \|\widetilde{W}^h\|^2_{L_{2,N}(\Omega_t^h)} dt \right) e^{2\lambda z} dz \leq \frac{C_4}{\lambda} e^{-2\lambda A} \|\widetilde{W}^h e^{\lambda z}\|^2_{L_{2,N}(\Omega^h)}. \end{aligned} \tag{11.133}$$

Similarly, using the Cauchy–Schwarz inequality, Lemma 11.8.1, and assuming that the parameter $\lambda \geq \lambda_0$, where λ_0 is a sufficiently large number depending on the same parameters as ones listed in (11.110), we estimate from the below the sum of the fourth and fifth lines of (11.132) as

$$\begin{aligned} &e^{-2\lambda A}\|[\widetilde{W}^h - D_2^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h) - V_3^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)]e^{\lambda z}\|^2_{L^h_{2,N}(\Omega^h)} \\ &\qquad + \gamma\|\widetilde{W}^h\|^2_{H^{1,h}_{0,N}(\Omega_1^h)} \\ &\quad \geq \frac{1}{2}e^{-2\lambda A}\|\widetilde{W}^h e^{\lambda z}\|^2_{L_{2,N}(\Omega^h)} - \frac{C_4}{\lambda}e^{-2\lambda A}\|\widetilde{W}^h e^{\lambda z}\|^2_{L_{2,N}(\Omega^h)} + \gamma\|\widetilde{W}^h\|^2_{H^{1,h}_{0,N}(\Omega_1^h)} \\ &\quad \geq \frac{1}{4}e^{-2\lambda A}\|\widetilde{W}^h e^{\lambda z}\|^2_{L_{2,N}(\Omega^h)} + \gamma\|\widetilde{W}^h\|^2_{H^{1,h}_{0,N}(\Omega_1^h)}. \end{aligned} \tag{11.134}$$

It follows from (11.127), (11.129), and (11.130) that the expression in the second line of (11.132) is linear with respect to $\widetilde{W}^h$,

$$\mathrm{Lin}(\widetilde{W}^h) = 2e^{-2\lambda A}(D_1^h, \widetilde{W}^h - D_2^h(\widetilde{W}^h, \widetilde{W}_x^h, \widetilde{W}_y^h)e^{2\lambda z}) + 2\gamma[\widetilde{W}^h, W^{(1),h}]. \tag{11.135}$$

Obviously $|\mathrm{Lin}(\widetilde{W}^h)| \le C_4 \|\widetilde{W}^h\|_{H^{1,h}_{0,N}(\Omega_1^h)}$. Hence, $\mathrm{Lin}(\widetilde{W}^h) : H^{1,h}_{0,N}(\Omega_1^h) \to \mathbb{R}$ is a bounded linear functional. Hence, by Riesz theorem there exists a matrix $\Theta \in H^{1,h}_{0,N}(\Omega_1^h)$ such that

$$\mathrm{Lin}(\widetilde{W}^h) = [\Theta, \widetilde{W}^h], \quad \forall \widetilde{W}^h \in H^{1,h}_{0,N}(\Omega_1^h). \tag{11.136}$$

Besides, it follows from the above that

$$J_{\lambda,\gamma}(W^{(1),h} + \widetilde{W}^h) - J_{\lambda,\gamma}(W^{(1),h}) - [\Theta, \widetilde{W}^h] = o(\|\widetilde{W}^h\|_{H^{1,h}_{0,N}(\Omega_1^h)}),$$

as $\|\widetilde{W}^h\|_{H^{1,h}_{0,N}(\Omega_1^h)} \to 0$. Hence, $\Theta = J'_{\lambda,\gamma}(W^{(1),h}) \in H^{1,h}_{0,N}(\Omega_1^h)$ is the Frechét derivative of the functional $J_{\lambda,\gamma}$ at the point $W^{(1),h} \in K^h(R)$. The existence of the Frechét derivative in the larger set $K^h(2R)$ can be proven completely similarly. The Lipschitz continuity property (11.111) of the Frechét derivative $J'_{\lambda,\gamma}$ can be proven similarly with the proof of Theorem 3.1 of [9]. Therefore, we omit this proof for brevity.

Furthermore, using (11.132)–(11.136) and recalling that $\widetilde{W}^h = W^{(2),h} - W^{(1),h}$, we obtain for sufficiently large $\lambda_0 > 1$ and for $\lambda \ge \lambda_0$,

$$\begin{aligned} &J_{\lambda,\gamma}(W^{(2),h}) - J_{\lambda,\gamma}(W^{(1),h}) - J'_{\lambda,\gamma}(W^{(1),h})(W^{(2),h} - W^{(1),h}) \\ &\quad \ge \frac{1}{8}\|\widetilde{W}^h\|^2_{L_{2,N}(\Omega^h)} + \gamma\|\widetilde{W}^h\|^2_{H^{1,h}_{0,N}(\Omega_1^h)}, \end{aligned}$$

which is the same as the target estimate (11.112) of this theorem.

11.9.2 Proof of Theorem 11.8.4

Recall that in this theorem we fix and arbitrary number $\lambda = \lambda_1 \ge \lambda_0$, where λ_0 is the number of Theorem 11.8.1. To indicate the dependence of the functional $J_{\lambda_1,\gamma}$ on Φ^h, we write in this proof $J_{\lambda_1,\gamma}(W^h, \Phi^h)$.

First, we consider $J_{\lambda_1,\gamma}(W^{*,h}, \Phi^{*,h})$. Since

$$W^{*,h}(\mathbf{x}^h(z)) - M_N^{-1}P(W^{*,h}, W_x^{*,h}, W_y^{*h}, \Phi^{*,h}, \mathbf{x}^h(z)) = 0, \quad \mathbf{x}^h(z) \in \Omega^h, \tag{11.137}$$

then, using (11.109), we obtain

$$J_{\lambda_0,\gamma}(W^{*,h}, \Phi^{*,h}) = \gamma\|W^{*,h}\|^2_{H_0^{1,h}(\Omega_1^h)} \le \gamma R^2. \tag{11.138}$$

Next, by (11.109), (11.120), (11.121), (11.137), and (11.138),

$$J_{\lambda_1,\gamma}(W^{*,h},\Phi^h) = J_{\lambda_1,\gamma}(W^{*,h},\Phi^{*h}) + Z(W^{*,h},W^h,\Phi^{*h},\Phi^h), \tag{11.139}$$

where $Z(W^{*,h}, W^h, \Phi^{*h}, \Phi^h)$ satisfies the following estimate:

$$\left|Z(W^{*,h},W^h,\Phi^{*h},\Phi^h)\right| \leq C_4\delta^2 e^{2\lambda_1\sigma}. \tag{11.140}$$

Since λ_1 is a fixed arbitrary number such that $\lambda_1 \geq \lambda_0$ and since λ_0 depends on the same parameters as those listed in (11.110) for C_4, then recalling that $\gamma = \delta^2$, we obtain from (11.138)–(11.140)

$$J_{\lambda_1,\gamma}(W^{*,h},\Phi^h) \leq C_4\delta^2. \tag{11.141}$$

Next, (11.112) implies that

$$\begin{aligned} &J_{\lambda_1,\gamma}(W^{*,h},\Phi^h) - J_{\lambda_1,\gamma}(W^h_{\min},\Phi^h) - J'_{\lambda_1,\gamma}(W^h_{\min},\Phi^h)(W^{*,h} - W^h_{\min}) \\ &\quad\geq \frac{1}{8}\|W^{*,h} - W^h_{\min}\|^2_{L^h_2(\Omega^h)}. \end{aligned} \tag{11.142}$$

Since by (11.113) $-J'_{\lambda_0,\gamma}(W^h_{\min},\Phi^h)(W^{*,h} - W^h_{\min}) \leq 0$, then, using (11.141) and (11.142), we obtain

$$\|W^{*,h} - W^h_{\min}\|^2_{L^h_2(\Omega^h)} \leq C_4\delta^2. \tag{11.143}$$

The first target estimate (11.122) of this theorem follows from (11.143). Combining (11.115) and (11.122) with the procedure of Section 11.6, which led to (11.103), we obtain two other target estimates (11.123) and (11.124) of Theorem 11.8.4.

12 Numerical solution of the linearized travel time tomography problem with incomplete data

In this chapter, we follow [148]. Permission for republication is obtained from the publisher.

12.1 Introduction

At the moment of the submission of this book, the authors do not yet have numerical results for the convexification method of Chapter 11 for the full nonlinear Travel Time Tomography Problem (TTTP), or, by the terminology of V. G. Romanov [224, Chapter 3], Inverse Kinematic Problem of Seismic (IKPS). Nevertheless, even the simpler linearized TTTP/IKPS has a significant applied interest in geophysics; see, for example, the corresponding discussion in Section 4 of Chapter 3 of the book [224]. Probably the first numerical method for the latter problem was developed in Chapter 9 of an earlier book of Romanov [223]. Furthermore, it is intriguing that results of testing of that method on experimentally collected geophysical data are presented in [223]. However, certain specific shapes of geodesic lines are assumed in [223] and we do not impose such an assumption here. In addition to [148, 223], a linearized TTTP was solved numerically in [199].

In this chapter, we develop a new numerical method for the linearized TTTP for the d-D case, $d = 2, 3, \dots$. Our data are both nonredundant and incomplete. Using results of [236], we establish the convergence of our method. In addition, we provide results of numerical experiments in the 2D case. In particular, we demonstrate that our method provides good accuracy of images of complicated objects with 5 % noise in the data. Furthermore, a satisfactory accuracy of images is demonstrated even for very high levels of noise between 80 % and 170 %.

In fact, both the idea of our method and sources/detectors configuration are close to those of our recent works [156, 236] and it is the same as in Chapter 11, which, in turn reflects results of [138]. However, our case is substantially more difficult than ones in [156, 236] since the waves in our case propagate along geodesic lines, rather than a radiation propagating along straight lines in these references. Still, although we formulate here results related to the convergence of our method, we do not prove them. This is because proofs are very similar to those in [236]. In other words, surprisingly, the analytical apparatus of the convergence theory developed in [236] works well for the problem considered in this paper. This is basically because we work here with a version of the quasi-reversibility method (QRM). And the convergence theory for this method, which is presented in Chapter 4, can be easily extended to the case considered in the current chapter. Recall that the convergence theory of Chapter 4 is again based on Carleman estimates.

https://doi.org/10.1515/9783110745481-012

In the isotropic case of acoustic/seismic waves propagation, the TTTP/IKPS is the problem of the recovery of the spatially distributed speed of propagation of acoustic/seismic waves from the first times of arrival of those waves. In the electromagnetic case, this is the problem of the recovery of the spatially distributed dielectric constant from those times. Waves are originated by some sources located either at the boundary of a bounded domain of interest or outside of this domain. Times of first arrival from those sources are measured on a part of the boundary of that domain.

The pioneering papers about the solution of the 1D TTTP/IKPS were published by Herglotz [88] (1905) and then by Wiechert and Zoeppritz [250] (1907). Their method is described in the book of Romanov [224, Section 3 of Chapter 3]. It was recently discovered that, in addition to geophysics, the TTTP/IKPS has applications in the phaseless inverse scattering problem [159, 226]. We refer to Section 11.1 for other comments about the past publications on TTTP/IKPS.

Below $d \geq 2$ is the dimension of the space $\mathbb{R}^d$. Points of this space are denoted as $\mathbf{x} \in \mathbb{R}^d$. Let $\overline{c} \equiv \text{const.} > 0$ be the speed of sound in a certain reference medium in $\mathbb{R}^d$, which we do not specify, and $c(\mathbf{x}) > 0$ be the variable speed of sound. Then the refractive index is [224, Chapter 3]

$$\mathbf{n}(\mathbf{x}) = \overline{c}/c(\mathbf{x}). \tag{12.1}$$

To linearize, one should assume that $\mathbf{n}(\mathbf{x}) = \mathbf{n}_0(\mathbf{x}) + \mathbf{n}_1(\mathbf{x})$, where $\mathbf{n}_0(\mathbf{x})$ is the known background function and $\mathbf{n}_1(\mathbf{x})$ with $|\mathbf{n}_1(\mathbf{x})| \ll \mathbf{n}_0(\mathbf{x})$ is its unknown perturbation, which is the subject to the solution of the linearized TTTP. Thus, one assumes that the refractive index is basically known, whereas its small perturbation $\mathbf{n}_1$ is unknown. This problem is also called the *geodesic X-ray transform problem* [199].

In our derivation, we end up with an overdetermined boundary value problem for a system of coupled linear PDEs of the first order. It is well known that the QRM is an effective tool for numerical solutions of overdetermined boundary value problems for PDEs, see Chapter 4. Another important feature of this chapter is a special orthonormal basis in the space $L^2(-\overline{\alpha}, \overline{\alpha})$, which was presented in Section 6.2.3. Here, $\overline{\alpha} > 0$ is a certain number. The functions of this basis depend only on the position of the point source. We consider a truncated Fourier series with respect to this basis. This assumption forms the first element of that model. The second element is that we assume that the first derivatives with respect to all variables are written via finite differences, and the step size of these finite differences with respect to all variables, except of one, is bounded from below by a positive number $h_0 > 0$. We do not prove convergence for the case when $h_0 \to 0^+$ and the number N of terms in that truncated series tends to infinity. Thus, we come up with a finite dimensional approximate mathematical model; see Remarks 7.3.

12.2 The linearization

Consider numbers R, a, b such that $R > 1$ and $0 < a < b$. Set

$$\Omega = (-R, R)^{d-1} \times (a, b) \subset \mathbb{R}^d. \tag{12.2}$$

Recall that by (12.1) $\mathbf{n}(\mathbf{x}) = \overline{c}/c(\mathbf{x})$, where $c(\mathbf{x})$ is the speed of sound propagation and $n(\mathbf{x})$ is the refractive index. Let the function $\mathbf{n}_0(\mathbf{x})$ be the known refractive index of the background. We assume that

$$\mathbf{n}_0, \mathbf{n} \in C^2(\mathbb{R}^d); \quad \mathbf{n}_0^2(\mathbf{x}), \mathbf{n}^2(\mathbf{x}) \geq 1 \quad \text{in } \Omega, \tag{12.3}$$

$$\mathbf{n}_0^2(\mathbf{x}) = \mathbf{n}^2(\mathbf{x}) = 1 \quad \text{for } \mathbf{x} \in \mathbb{R}^d \setminus \Omega. \tag{12.4}$$

For any two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathbb{R}^d$, define the geodesic line generated by $\mathbf{n}_0$ connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ as

$$\Gamma_0(\mathbf{x}_1, \mathbf{x}_2) = \operatorname{argmin}\left\{ \int_\gamma \mathbf{n}_0(\boldsymbol{\xi}) d\sigma(\boldsymbol{\xi}) \text{ where } \gamma : [0,1] \to \mathbb{R}^d \right.$$

$$\left. \text{is a smooth map with } \gamma(0) = \mathbf{x}_1, \gamma(1) = \mathbf{x}_2 \right\}. \tag{12.5}$$

Here, $d\sigma(\xi)$ is the elementary arc length. Note that by (12.5) the geodesic line $\Gamma_0(\mathbf{x}_1, \mathbf{x}_2)$ connects points $\mathbf{x}_1$ and $\mathbf{x}_2$. Let

$$\mathbf{a}_0(\mathbf{x}) = \mathbf{n}_0^2(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^d. \tag{12.6}$$

The corresponding travel time between $\mathbf{x}_1$ and $\mathbf{x}_2$ is the integral

$$\int_{\Gamma_0(\mathbf{x}_1, \mathbf{x}_2)} \mathbf{n}_0(\boldsymbol{\xi}) d\sigma(\boldsymbol{\xi}) = \int_{\Gamma_0(\mathbf{x}_1, \mathbf{x}_2)} \sqrt{\mathbf{a}_0(\boldsymbol{\xi})} d\sigma(\boldsymbol{\xi}).$$

Introduce the line of sources L_s located on the x_1-axis as

$$L_s = [-\overline{\alpha}, \overline{\alpha}] \times \{(0, 0, \ldots, 0)\}, \tag{12.7}$$

where $\overline{\alpha}$ is a fixed positive number. It follows from (12.2) and (12.7) that

$$\overline{\Omega} \cap L_s = \varnothing. \tag{12.8}$$

For $\mathbf{x}_\alpha \in L_s$, the travel time along $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$ of the wave from $\mathbf{x}_\alpha$ to $\mathbf{x}$ is

$$u_0(\mathbf{x}, \mathbf{x}_\alpha) = \int_{\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)} \sqrt{\mathbf{a}_0(\boldsymbol{\xi})} d\sigma(\boldsymbol{\xi}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{12.9}$$

Assumption 12.2.1 (regularity of geodesic lines). We assume everywhere in this paper that the geodesic lines are regular in the following sense: for each point $\mathbf{x}$ of the closed domain $\overline{\Omega}$ and for each point $\mathbf{x}_\alpha$ of the line of sources L_s there exists a single geodesic line $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$ connecting them.

The inverse problem we consider arises from the highly nonlinear and severely ill-posed inverse kinematic problem. We now present the formal linearization arguments in exactly the same way as they are presented in the book of Romanov [224, Section 4 of Chapter 3]. Just as in that book, we avoid a setting via functional spaces, for brevity.

Assume that the function $\mathbf{a}(\mathbf{x}) = \mathbf{n}^2(\mathbf{x})$ contains a perturbation term of the background function $\mathbf{a}_0(\mathbf{x}) = \mathbf{n}_0^2(\mathbf{x})$. In other words,

$$\mathbf{a}(\mathbf{x}) = \mathbf{a}_0(\mathbf{x}) + 2\epsilon\sqrt{\mathbf{a}_0(\mathbf{x})}p(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{12.10}$$

where $\epsilon > 0$ is a sufficiently small number. Here, the function $p \in C(\mathbb{R}^d)$ and $p(\mathbf{x}) = 0$ for $\mathbf{x} \notin \overline{\Omega}$. Hence, by (12.8) $p(\mathbf{x}) = 0$ for points $\mathbf{x}$ in a small neighborhood of the line of sources L_s. Denote

$$u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha) = \int\limits_{\Gamma_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)} \mathbf{n}(\boldsymbol{\xi}) d\sigma(\boldsymbol{\xi})$$

the travel time from the point $\mathbf{x}_\alpha \in L_s$ to the point $\mathbf{x} \in \Omega$, where $\Gamma_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)$ is the geodesic line generated by the function $\mathbf{n}(\mathbf{x})$. It is well known that $u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)$ satisfies the Eikonal equation [224, Chapter 3]

$$\left|\nabla_{\mathbf{x}} u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)\right|^2 = \mathbf{a}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \mathbf{x}_\alpha \in L_s. \tag{12.11}$$

Let $u_0(\mathbf{x}, \mathbf{x}_\alpha)$ be the travel time function in (12.9) corresponding to the background $\mathbf{a}_0$. Then

$$\left|\nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha)\right|^2 = \mathbf{a}_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \mathbf{x}_\alpha \in L_s. \tag{12.12}$$

Due to (12.10), we represent $\nabla_{\mathbf{x}} u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)$ as

$$\nabla_{\mathbf{x}} u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha) = \nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha) + \epsilon \nabla_{\mathbf{x}} u^{(1)}(\mathbf{x}, \mathbf{x}_\alpha). \tag{12.13}$$

Hence, ignoring the term with ϵ^2, we obtain

$$\left|\nabla_{\mathbf{x}} u_{\mathbf{n}}(\mathbf{x}, \mathbf{x}_\alpha)\right|^2 \approx \left|\nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha)\right|^2 + 2\epsilon \nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha) \nabla_{\mathbf{x}} u^{(1)}(\mathbf{x}, \mathbf{x}_\alpha). \tag{12.14}$$

Denoting

$$u^{(1)} := u \tag{12.15}$$

and comparing (12.14) with (12.10) and (12.12), we obtain

$$\frac{\nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\sqrt{\mathbf{a}_0(\mathbf{x})}} \cdot \nabla_{\mathbf{x}} u(\mathbf{x}, \mathbf{x}_\alpha) = p(\mathbf{x}). \tag{12.16}$$

Thus, equation (12.16) is the "linearization" of the nonlinear equation (12.11). In fact, a similar PDE for the case when geodesic lines are straight lines was derived in [87, Chapter 3] and was used then in [156] to invert incomplete Radon transform data.

By (12.12) $|\nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha)|/\sqrt{\mathbf{a}_0(\mathbf{x})} \equiv 1$. Hence, this is a unit vector, which is tangent to the curve $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$ at the point $\mathbf{x}$ [224, Chapter 3]. Hence, the left-hand side of (12.16) is the derivative of the function $u(\mathbf{x}, \mathbf{x}_\alpha)$ along the curve $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$. Thus, integrating, we obtain [224, Chapter 3]

$$u(\mathbf{x}, \mathbf{x}_\alpha) = \int\limits_{\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)} p(\boldsymbol{\xi}) d\sigma(\boldsymbol{\xi}). \tag{12.17}$$

Let $\partial\Omega_{\mathrm{sm}}$ be the smooth part of the boundary $\partial\Omega$ of the domain Ω. For each $\alpha \in (-\overline{\alpha}, \overline{\alpha})$, define

$$\partial\Omega_\alpha^- = \{\mathbf{x} \in \partial\Omega_{\mathrm{sm}} : \nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha) \cdot \nu(\mathbf{x}) < 0\},$$
$$\partial\Omega_\alpha^+ = \{\mathbf{x} \in \partial\Omega_{\mathrm{sm}} : \nabla_{\mathbf{x}} u_0(\mathbf{x}, \mathbf{x}_\alpha) \cdot \nu(\mathbf{x}) > 0\},$$

where $\mathbf{x}_\alpha = (\alpha, 0, \dots, 0) \in L_{\mathrm{s}}$ and $\nu(\mathbf{x})$ is the outward looking unit normal vector at the point $\mathbf{x} \in \partial\Omega_{\mathrm{sm}}$. If $\mathbf{n}_0 \equiv 1$, then $\Gamma_0(\mathbf{x}_1, \mathbf{x}_2)$ is the line segment connecting these two points. Hence, it follows from (12.2), (12.7), (12.8), (12.13), and (12.15) that

$$u(\mathbf{x}, \mathbf{x}_\alpha) = 0, \quad \mathbf{x} \in \partial\Omega_\alpha^-. \tag{12.18}$$

The aim of this paper is to solve the following inverse problem.

Problem 12.2.1 (linearized travel time tomography problem). Let the function $u = u(\mathbf{x}, \mathbf{x}_\alpha) \in C^1(\overline{\Omega} \times [-\overline{\alpha}, \overline{\alpha}])$ be the solution of boundary value problem (12.16), (12.18). Given the data $f(\mathbf{x}, \mathbf{x}_\alpha)$,

$$f(\mathbf{x}, \mathbf{x}_\alpha) = \begin{cases} u(\mathbf{x}, \mathbf{x}_\alpha), & \mathbf{x} \in \partial\Omega_\alpha^+, \mathbf{x}_\alpha \in L_{\mathrm{s}}, \\ 0, & \mathbf{x} \in \partial\Omega_\alpha^-, \mathbf{x}_\alpha \in L_{\mathrm{s}}, \end{cases} \tag{12.19}$$

determine the function $p(\mathbf{x})$, $\mathbf{x} \in \Omega$.

Note that the data (12.19) are nonredundant ones. Indeed, the source $\mathbf{x}_\alpha \in L_{\mathrm{s}}$ depends on one variable and the point $\mathbf{x} \in \partial\Omega_\alpha^+$ depends on $d-1$ variables. Hence the function $f(\mathbf{x}, \mathbf{x}_\alpha)$ depends on d variables, so as the target unknown function $p(\mathbf{x})$.

From now on, to separate the coordinate number d of the point $\mathbf{x}$, we write $\mathbf{x} = (x_1, \dots, x_{d-1}, z)$. The transport equation in (12.16) is read as

$$\frac{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\sqrt{\mathbf{a}_0(\mathbf{x})}} \partial_z u(\mathbf{x}, \mathbf{x}_\alpha) + \sum_{i=1}^{d-1} \frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\sqrt{\mathbf{a}_0(\mathbf{x})}} \partial_{x_i} u(\mathbf{x}, \mathbf{x}_\alpha) = p(\mathbf{x}), \tag{12.20}$$

which is equivalent to

$$\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u(\mathbf{x}, \mathbf{x}_\alpha) + \sum_{i=1}^{d-1} \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_{x_i} u(\mathbf{x}, \mathbf{x}_\alpha) = \sqrt{\mathbf{a}_0(\mathbf{x})}p(\mathbf{x}) \tag{12.21}$$

for all $\mathbf{x} \in \Omega$, $\mathbf{x}_\alpha \in L_s$.

12.3 A boundary value problem for a system of coupled PDEs of the first order

This section aims to derive a system of partial differential equations, which can be stably solved by the quasi-reversibility method in the semi-finite difference scheme. The solution of this system yields the desired numerical solution to Problem 12.2.1. Recall that Problem 12.2.1 is the linearized travel time tomography problem, and it is labeled this way by its title.

We employ now the special orthonormal basis of $L_2(-\overline{\alpha}, \overline{\alpha})$ of Section 6.2.3, where $2\overline{\alpha}$ is the length of the line of source L_{sc}; see (12.7). For each $n = 1, 2, \ldots$, let $\phi_n(\alpha) = \alpha^{n-1}\exp(\alpha)$. The set $\{\phi_n\}_{n=1}^{\infty}$ is complete in $L^2(-\overline{\alpha}, \overline{\alpha})$. Applying the Gram–Schmidt orthonormalization process to this set, we obtain a basis of $L^2(-\overline{\alpha}, \overline{\alpha})$, named as $\{\Psi_n\}_{n=1}^{\infty}$.

Proposition 12.3.1 (see Theorem 6.2.1). *Any function Ψ_n is not identically zero. For all $m, n \geq 1$,*

$$s_{mn} = \int_{-\overline{\alpha}}^{\overline{\alpha}} \Psi_n'(\alpha)\Psi_m(\alpha)d\alpha = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } n < m. \end{cases}$$

Thus, the matrix $S_N = (s_{mn})_{m,n=1}^{N}$, is invertible for all integers $N \geq 1$.

We now derive a system of partial differential equations for the Fourier coefficients of the function

$$w(\mathbf{x}, \mathbf{x}_\alpha) = u(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha), \quad \mathbf{x} \in \Omega, \mathbf{x}_\alpha \in L_s \tag{12.22}$$

with respect to the basis $\{\Psi_n\}_{n=1}^{\infty}$. Differentiate (12.21) with respect to α. We obtain

$$\frac{\partial}{\partial \alpha}\left[\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u(\mathbf{x}, \mathbf{x}_\alpha) + \sum_{i=1}^{d-1} \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_{x_i} u(\mathbf{x}, \mathbf{x}_\alpha)\right] = 0 \tag{12.23}$$

for all $\mathbf{x} \in \Omega$, $\mathbf{x}_\alpha \in L_s$. From now on, we impose the following condition.

Assumption 12.3.1 (Monotonicity condition in the z-direction). The traveling time function u_0, defined in (12.9) with $\mathbf{n}$ replaced by $\mathbf{n}_0$, is strictly increasing with respect to z. In other words, $\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha) > 0$ for all $\mathbf{x} = (x_1, \ldots, x_{d-1}, z) \in \Omega$ and for all $\mathbf{x}_\alpha \in L_s$.

Assumption 12.3.1 means that the higher in the z-direction, the longer the traveling time is. A sufficient condition for Assumption 12.3.1 to be true is formulated in (12.24) of Lemma 11.3.1. A similar monotonicity condition can be found in formulas (3.24) and (3.24′) of Section 2 of Chapter 3 of the book [224]. Also, a similar condition was imposed in originating works for the 1D problem of Herglotz and Wiechert and Zoeppritz [88, 250]: see Section 3 of Chapter 3 of [224]. Besides, Figures 5 and 10 of [249] justify this condition from the geophysical standpoint. Recall that by (12.6) and (12.3) $\mathbf{a}_0 \in C^2(\mathbb{R}^d)$ and $\mathbf{a}_0(\mathbf{x}) \geq 1$ in $\mathbb{R}^d$. Therefore, we recall the full analog of Lemma 11.4.1.

Lemma 12.3.1 (a full analog of Lemma 11.4.1). *Let conditions* (12.3) *and* (12.6) *hold. Also, let*

$$\partial_z \mathbf{a}_0(\mathbf{x}) \geq 0 \quad \text{for all } \mathbf{x} \in \overline{\Omega}. \tag{12.24}$$

Then

$$\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha) \geq \frac{a}{\sqrt{a^2+2}} \quad \text{for all } \mathbf{x} \in \overline{\Omega}, \alpha \in [-\overline{\alpha}, \overline{\alpha}].$$

Although this lemma is proven in Chapter 11 only in the 3D case, the proof in the d-D case is very similar and is, therefore, avoided. Let $w(\mathbf{x}, \mathbf{x}_\alpha)$ be the function defined in (12.22). Then

$$\begin{aligned} \partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u(\mathbf{x}, \mathbf{x}_\alpha) &= \partial_z w(\mathbf{x}, \mathbf{x}_\alpha) - u(\mathbf{x}, \alpha)\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha) \\ &= \partial_z w(\mathbf{x}, \mathbf{x}_\alpha) - w(\mathbf{x}, \alpha)\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}. \end{aligned} \tag{12.25}$$

Also, for $i = 1, \dots, d-1$

$$\begin{aligned} \partial_{x_i} u(\mathbf{x}, \mathbf{x}_\alpha) &= \frac{\partial}{\partial x_i}\left(\frac{w(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\right) \\ &= \frac{\partial_{x_i} w(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha) - w(\mathbf{x}, \mathbf{x}_\alpha)\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2} \end{aligned} \tag{12.26}$$

for all $\mathbf{x} \in \Omega$, $\mathbf{x}_\alpha \in L_{\text{sc}}$. Combining (12.23), (12.25), and (12.26), we obtain

$$\begin{aligned} &\frac{\partial}{\partial \alpha}\Bigg[\partial_z w(\mathbf{x}, \mathbf{x}_\alpha) - w(\mathbf{x}, \mathbf{x}_\alpha)\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \\ &\quad + \sum_{i=1}^{d-1} \frac{\partial_{x_i} w(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha) - w(\mathbf{x}, \mathbf{x}_\alpha)\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2}\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)\Bigg] = 0. \end{aligned} \tag{12.27}$$

This is equivalent to

$$\begin{aligned} &\partial_{\alpha z} w(\mathbf{x}, \mathbf{x}_\alpha) - \frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\partial_\alpha w(\mathbf{x}, \mathbf{x}_\alpha) - \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\right)w(\mathbf{x}, \mathbf{x}_\alpha) \\ &\quad + \sum_{i=1}^{d-1}\left[\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\partial_{\alpha x_i} w(\mathbf{x}, \mathbf{x}_\alpha) + \frac{\partial}{\partial \alpha}\left(\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\right)\partial_{x_i} w(\mathbf{x}, \mathbf{x}_\alpha)\right. \end{aligned}$$

$$- \frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2} \partial_\alpha w(\mathbf{x}, \mathbf{x}_\alpha)$$
$$- \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2} \right) w(\mathbf{x}, \mathbf{x}_\alpha) \Bigg] = 0. \tag{12.28}$$

We recall now the orthonormal basis $\{\Psi_n\}_{n=1}^\infty$ constructed at the beginning of this section. For each $\mathbf{x} \in \Omega$ and for all $\mathbf{x}_\alpha \in L_{\mathrm{sc}}$, we write

$$w(\mathbf{x}, \mathbf{x}_\alpha) = \sum_{n=1}^\infty w_n(\mathbf{x}) \Psi_n(\alpha) \approx \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n(\alpha), \tag{12.29}$$

$$w_n(\mathbf{x}) = \int_{-\bar{\alpha}}^{\bar{\alpha}} w(\mathbf{x}, \mathbf{x}_\alpha) \Psi_n(\alpha) d\alpha. \tag{12.30}$$

The "cut-off" number N is chosen numerically. We discuss the choice of N in more details in Section 12.5. We assume that the approximation $\approx$ in (12.29) is an equality as well as

$$\partial_\alpha w(\mathbf{x}, \mathbf{x}_\alpha) = \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n'(\alpha). \tag{12.31}$$

Plugging (12.29) and (12.31) into (12.28) gives

$$\sum_{n=1}^N \partial_z w_n(\mathbf{x}) \Psi_n'(\alpha) - \frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n'(\alpha)$$
$$- \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \right) \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n(\alpha) + \sum_{i=1}^{d-1} \Bigg[\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \sum_{n=1}^N \partial_{x_i} w_n(\mathbf{x}) \Psi_n'(\alpha)$$
$$+ \frac{\partial}{\partial \alpha}\left(\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \right) \sum_{n=1}^N \partial_{x_i} w_n(\mathbf{x}) \Psi_n(\alpha) - \frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2} \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n'(\alpha)$$
$$- \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2} \right) \sum_{n=1}^N w_n(\mathbf{x}) \Psi_n(\alpha) \Bigg] = 0.$$

For each $m \in \{1, \dots, N\}$, multiply the latter equation by $\Psi_m(\alpha)$ and then integrate the resulting equation with respect to α. We get

$$\sum_{n=1}^N s_{mn} \partial_z w_n(\mathbf{x}) + \sum_{n=1}^N a_{mn}(\mathbf{x}) w_n(\mathbf{x}) + \sum_{n=1}^N \sum_{i=1}^{d-1} b_{mn,i}(\mathbf{x}) \partial_{x_i} w_n(\mathbf{x}) = 0 \tag{12.32}$$

for all $\mathbf{x} \in \Omega$ where s_{mn} is defined as in Proposition 12.3.1,

$$a_{mn}(\mathbf{x}) = \int_{-\bar{\alpha}}^{\bar{\alpha}} \Bigg[-\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \Psi_n'(\alpha) - \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zz} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)} \right) \Psi_n(\alpha)$$

$$- \sum_{i=1}^{d-1} \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2}\right)\Psi_n'(\alpha)$$
$$- \sum_{i=1}^{d-1} \frac{\partial}{\partial \alpha}\left(\frac{\partial_{zx_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{(\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha))^2}\right)\Psi_n(\alpha)\Bigg]\Psi_m(\alpha)d\alpha \tag{12.33}$$

and for $i = 1, \dots, d-1$,

$$b_{mn,i}(\mathbf{x}) = \int_{-\bar{\alpha}}^{\bar{\alpha}} \left[\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\Psi_n'(\alpha) + \frac{\partial}{\partial \alpha}\left(\frac{\partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha)}{\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)}\right)\Psi_n(\alpha)\right]\Psi_m(\alpha)d\alpha, \tag{12.34}$$

for all $\mathbf{x} \in \Omega$. For each $\mathbf{x} \in \Omega$, let $W(\mathbf{x}) = (w_1(\mathbf{x}), \dots, w_N(\mathbf{x}))^T$, $S = (s_{mn})_{m,n=1}^N$, $A(\mathbf{x}) = (a_{mn}(\mathbf{x}))_{m,n=1}^N$ and $B_i(\mathbf{x}) = (b_{mn,i}(\mathbf{x}))_{m,n=1}^N$ for $i = 1, \dots, d-1$. Since (3.8) holds true for every $m = 1, \dots, N$, it can be rewritten as

$$S_N \partial_z W(\mathbf{x}) + A(\mathbf{x})W(\mathbf{x}) + \sum_{i=1}^{d-1} B_i(\mathbf{x})\partial_{x_i} W(\mathbf{x}) = 0. \tag{12.35}$$

Since S is invertible, see Proposition 12.3.1, then (12.35) implies the following system of transport equations:

$$\partial_z W(\mathbf{x}) + S_N^{-1} A(\mathbf{x})W(\mathbf{x}) + \sum_{i=1}^{d-1} S_N^{-1} B_i(\mathbf{x})\partial_{x_i} W(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega. \tag{12.36}$$

The boundary data for W are

$$W|_{\partial\Omega} = F(\mathbf{x}) = (f_n)_{n=1}^N, \quad f_n(\mathbf{x}) = \int_{-\bar{\alpha}}^{\bar{\alpha}} f(\mathbf{x}, \mathbf{x}_\alpha)\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha)\Psi_n(\alpha)d\alpha \tag{12.37}$$

where f is the given data; see (12.19).

Remark 12.3.1. Problem 12.2.1 is reduced to the problem of finding the vector valued function W satisfying the system (12.36) and the boundary condition (12.37). Assume this vector function is computed and denote it as $W^{\text{comp}} = (w_1^{\text{comp}}, \dots, w_n^{\text{comp}})$. Then we can compute the function $w^{\text{comp}}(\mathbf{x}, \mathbf{x}_\alpha)$ and then the function $u^{\text{comp}}(\mathbf{x}, \mathbf{x}_\alpha)$ sequentially via (12.29) and (12.22). The computed target function $p^{\text{comp}}(\mathbf{x})$ is given by (12.20).

We find an approximate solution of the boundary value problem (12.36)–(12.37) by the QRM. This means that we minimize the functional

$$J_\epsilon(W) = \int_\Omega \left|\partial_z W(\mathbf{x}) + \sum_{i=1}^{d-1} S_N^{-1} B_i(\mathbf{x})\partial_{x_i} W(\mathbf{x}) + S_N^{-1} A(\mathbf{x})W(\mathbf{x})\right|^2 d\mathbf{x}$$
$$+ \epsilon\|W\|^2_{H^1(\Omega)^N} \tag{12.38}$$

on the set of vector functions $W \in H^1(\Omega)^N$ satisfying the boundary constraint (12.37). Here, the space $H^1(\Omega)^N = \underbrace{H^1(\Omega) \times \cdots \times H^1(\Omega)}_{N}$ with the commonly defined norm. Similar to [236], we analyze the functional $J_\epsilon(W)$ for the case when derivatives in (12.38) with respect to x_i, $i = 1, \ldots, d-1$ are written in finite differences.

12.4 The QRM in partial finite differences

For brevity, we describe and analyze here the QRM in the case when $d = 2$. The arguments for higher dimensions can be done in the same manner. In 2D, $\Omega = (-R, R) \times (a, b)$. We arrange an $M_x \times M_z$ grid of points on $\overline{\Omega}$,

$$\begin{aligned} \mathcal{G} = \{(x_i, z_j) : x_i = -R + (i-1)h_x, z_j = a + (j-1)h_z, \\ i = 1, \ldots, M_x, j = 1, \ldots, M_z\}, \end{aligned} \tag{12.39}$$

where $h_x \in [h_0, \beta_x)$ and $h_z \in (0, \beta_z)$ are grid step sizes in the x and z directions, respectively, and $h_0, \beta_x, \beta_z > 0$ are certain numbers. Here, N_x and N_z are two positive integers. Let $\mathbf{h} = (h_x, h_z)$. We define the discrete set $\Omega^{\mathbf{h}}$ as the set of those points of the set (12.39), which are interior points of the rectangle Ω and $\partial\Omega^{\mathbf{h}}$ is the set of those points of the set (12.39), which are located on the boundary of Ω,

$$\begin{aligned} \Omega^{\mathbf{h}} &= \{(x_i, z_j) : x_i = -R + (i-1)h_x, z_j = a + (j-1)h_z : \\ &\quad i = 2, \ldots, M_x - 1; j = 2, \ldots, N_z - 1\}, \\ \partial\Omega^{\mathbf{h}} &= \{(\pm R, z_j) : j = 1, \ldots, M_z\} \cup \{(x_i, z) : i = 1, \ldots, M_x, z \in \{a, b\}\}, \\ \overline{\Omega}^{\mathbf{h}} &= \Omega^{\mathbf{h}} \cup \partial\Omega^{\mathbf{h}}. \end{aligned}$$

For any continuous function v defined on Ω its finite difference version is $v^{\mathbf{h}} = v|_{\mathcal{G}}$. Here, $\mathbf{h}$ denotes the pair (h_x, h_z). The partial derivatives of the function v are given via forward finite differences as

$$\begin{aligned} \partial_x^{h_x} v^{\mathbf{h}}(x_i, z_j) &= \frac{v^{\mathbf{h}}(x_{i+1}, z_j) - v^{\mathbf{h}}(x_i, z_j)}{h_x}, \\ \partial_z^{h_z} v^{\mathbf{h}}(x_i, z_j) &= \frac{v(x_i, z_{j+1}) - v(x_i, z_j)}{h_z} \end{aligned} \tag{12.40}$$

for $i = 0, \ldots, N_x - 1$ and $j = 0, \ldots, N_z - 1$. We denote the finite difference analogs of the spaces $L^2(\Omega)$ and $H^1(\Omega)$ as $L^{2,\mathbf{h}}(\Omega)$ and $H^{1,\mathbf{h}}(\Omega)$. Norms in these spaces are defined as

$$\|v^{\mathbf{h}}\|_{L^{2,\mathbf{h}}(\Omega^{\mathbf{h}})} = \left[h_x h_z \sum_{i=1}^{M_x} \sum_{j=1}^{M_z} [v^{\mathbf{h}}(x_i, z_j)]^2 \right]^{1/2},$$

$$\|v^{\mathbf{h}}\|_{H^{1,\mathbf{h}}(\Omega^{\mathbf{h}})} = \left[\|v^{\mathbf{h}}\|^2_{L^{2,\mathbf{h}}(\Omega^{\mathbf{h}})} + h_x h_z \sum_{i=1}^{N_x-1} \sum_{j=1}^{N_z-1} [\partial_x^{h_x} v^{\mathbf{h}}(x_i, z_j)]^2 + [\partial_z^{h_z} v^{\mathbf{h}}(x_i, z_j)]^2 \right]^{1/2}.$$

Let $F^{\mathbf{h}} = F|_{\partial\Omega^{\mathbf{h}}}$. The problem (12.36)–(12.37) becomes

$$L^h(W^{\mathbf{h}}) = \partial_z^{h_z} W^{\mathbf{h}}(x_i, z_j) + S_N^{-1} B_1(\mathbf{x}_i, z_j) \partial_x^{h_x} W^{\mathbf{h}}(x_i, z_j) + S_N^{-1} A(x_i, z_j) W^{\mathbf{h}}(x_i, z_j) = 0 \tag{12.41}$$

for $i = 0, \dots, N_x - 1$; $j = 0, \dots, N_z - 1$ and

$$W^{\mathbf{h}}|_{\partial\Omega^{\mathbf{h}}} = F^{\mathbf{h}}. \tag{12.42}$$

To solve problem (12.41)–(12.42) numerically, we introduce the finite difference version of the functional J_ϵ, defined in (12.38),

$$J_\epsilon^{\mathbf{h}}(W^{\mathbf{h}}) = h_x h_z \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_z-1} \left|\partial_z^{h_z} W^{\mathbf{h}}(x_i, z_j) + S_N^{-1} B_1(\mathbf{x}_i, z_j) \partial_x^{h_x} W^{\mathbf{h}}(z_i, z_j) + S_N^{-1} A(x_i, z_j) W^{\mathbf{h}}(x_i, z_j)\right|^2 + \epsilon \|W^{\mathbf{h}}\|^2_{H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})}, \tag{12.43}$$

where $H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}}) = [H^{1,\mathbf{h}}(\Omega^{\mathbf{h}})]^N$ and similarly for $L_N^{2,\mathbf{h}}(\Omega^{\mathbf{h}})$. We consider the following problem.

Problem 12.4.1 (Minimization problem). Minimize the functional $J_\epsilon^{\mathbf{h}}(W^{\mathbf{h}})$ on the set of such vector functions $W^{\mathbf{h}} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ that satisfy boundary condition (12.42).

The convergence theory for this problem is formulated in Theorems 12.4.1 and 12.4.2. Proofs of these theorems follow closely the arguments of [236, Section 5] and are, therefore, not repeated in this paper. Theorem 12.4.1 guarantees the existence and uniqueness of the minimizer of $J_\epsilon^{\mathbf{h}}(W^{\mathbf{h}})$, and this result can be proven on the basis of Riesz theorem. The next natural although a more difficult question is about the convergence of regularized solutions (i. e., minimizers) to the exact one when the level of the noise in the data tends to zero, that is, Theorem 12.4.2. A close analog of Theorem 12.4.3 is proven in [236] via applying a new discrete Carleman estimate: recall that conventional Carleman estimates are in the continuous form. In other words, these two theorems confirm the effectiveness of our proposed numerical method for solving Problem 12.2.1.

Theorem 12.4.1 (existence and uniqueness of the minimizer). *For any $h = (h_x, h_z)$ with $h_x \in [h_0, \beta_x)$, $h_z \in (0, \beta_z)$, any $\epsilon > 0$ and for any matrix $F^{\mathbf{h}}$ of boundary conditions there exists unique minimizer $W^{\mathbf{h}}_{\min,\epsilon} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ of the functional satisfying boundary condition* (12.42).

As it is always the case in the regularization theory, assume now that there exists an “ideal” solution $W_*^{\mathbf{h}} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ of problem (12.41)–(12.42) satisfying the following boundary condition:

$$W_*^{\mathbf{h}}|_{\partial\Omega^{\mathbf{h}}} = F_*^{\mathbf{h}}, \tag{12.44}$$

where $F_*^{\mathbf{h}}$ is the “ideal” noiseless boundary data. Since $W_*^{\mathbf{h}}$ exists, then (12.44) implies that there exists an extension $G_*^{\mathbf{h}} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ with $G_*^{\mathbf{h}}|_{\partial\Omega^{\mathbf{h}}} = F_*^{\mathbf{h}}$ of the matrix $F_*^{\mathbf{h}}$ in $\Omega^{\mathbf{h}}$. As to the data $F^{\mathbf{h}}$ in (11.31), we assume now that there exists an extension $G^{\mathbf{h}} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ with $G^{\mathbf{h}}|_{\partial\Omega^{\mathbf{h}}} = F^{\mathbf{h}}$ of $F^{\mathbf{h}}$ in $\Omega^{\mathbf{h}}$. Let $\delta > 0$ be the level of the noise in $G^{\mathbf{h}}$. We assume that

$$\|G^{\mathbf{h}} - G_*^{\mathbf{h}}\|_{H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})} < B\delta, \tag{12.45}$$

where the constant $B > 0$ is independent on δ.

It is convenient to replace the above notation of the minimizer $W_{\min,\epsilon}^{\mathbf{h}}$ with $W_{\min,\epsilon,\delta}^{\mathbf{h}}$, thus, indicating its dependence on δ. In [236, Section 5], to prove a direct analog of Theorem 12.4.2 (formulated below), a new Carleman estimate for the finite difference operator $\partial_z^{h_z} v$ was proven first. The Carleman weight function of this estimate depends only on the discrete variable z. The value of this function at the point $z_j = a + (j-1)h_z$ is $e^{2\lambda(j-1)h_z}$, where $\lambda \geq 1$ is a parameter. This estimate is valid only if $\lambda h_z < 1$ (Lemma 4.7 of [236, Section 5]). The latter explains the condition of Theorem 12.4.2 imposed on the grid step size h_z in the z-direction.

We now explain why do we impose the condition that the grid step size h_x in the x-direction must be bounded from below as $h_x \geq h_0 = \text{const.} > 0$. Indeed, this bound guarantees that with a constant $C > 0$ independent on $\mathbf{h}$, we have $\|\partial_x^{h_x} W^{\mathbf{h}}\|_{L^{2,\mathbf{h}}(\Omega^{\mathbf{h}})} \leq C\|W^{\mathbf{h}}\|_{L^{2,\mathbf{h}}(\Omega^{\mathbf{h}})}$, which is exactly inequality (4.8) of [236, Section 4]. Note that proofs of convergence results in [236, Section 5] use the latter inequality quite essentially.

Theorem 12.4.2 (convergence of regularized solutions). *Let conditions* (12.44) *and* (12.45) *be valid. Let* $L^{\mathbf{h}}$ *be the operator in* (12.41). *Let* $W_{\min,\epsilon,\delta}^{\mathbf{h}} \in H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})$ *be the minimizer of the functional* $J_\epsilon^{\mathbf{h}}(W^{\mathbf{h}})$ *with boundary condition* (12.42). *Then there exists a sufficiently small number* $\overline{h}_z > 0$ *depending only on* h_0, a, b, R, N, $L^{\mathbf{h}}$ *such that the following estimate is valid for all* $(h_x, h_z) \in [h_0, \beta_x) \times (0, \overline{h}_z)$ *and all* $\epsilon, \delta > 0$ *with a constant* $C > 0$ *independent on* ϵ, δ:

$$\|W_{\min,\epsilon,\delta}^{\mathbf{h}} - W_*^{\mathbf{h}}\|_{L_N^{2,\mathbf{h}}(\Omega^{\mathbf{h}})} \leq C(\delta + \sqrt{\epsilon}\|W_*^{\mathbf{h}}\|_{H_N^{1,\mathbf{h}}(\Omega^{\mathbf{h}})}).$$

We also note that Lipschitz stability estimate for problem (12.41)–(12.42) is valid as a direct analog of Theorem 5.5 of [236, Section 5]. Therefore, uniqueness also takes place for problem (12.41)–(12.42).

12.5 Numerical implementation

In this section, we solve Problem 12.2.1 in the 2D case. The domain Ω is

$$\Omega = (-1,1) \times (1,3). \tag{12.46}$$

The line of sources L_s is set to be $(-\overline{\alpha}, \overline{\alpha})$ with $\overline{\alpha} = 3$.

We solve the forward problem to compute the simulated data as follows. Given the background function $\mathbf{n}_0$, instead of solving the nonlinear Eikonal equation (12.12), we find $u_0(\mathbf{x}, \mathbf{x}_\alpha)$ using (12.9). To do this, we first find the geodesic line $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$ in (12.9) connecting points $\mathbf{x} \in \Omega$ and $\mathbf{x}_\alpha \in L_s$.We do the latter by using the 2D Fast Marching toolbox which is built in Matlab. The Fast Marching is very similar to the Dijkstra algorithm to find the shortest paths on graphs. We refer the reader to [233] for more details about fast marching. Next, with this geodesic line $\Gamma_0(\mathbf{x}, \mathbf{x}_\alpha)$ in hands, we compute the function $u(\mathbf{x}, \mathbf{x}_\alpha)$ via (12.17). It is clear that this function u solves (12.16). The point $\mathbf{x}_\alpha$ above is chosen as $(\alpha_i, 0)$ where $\alpha_i = 2(i-1)\overline{\alpha}/N_\alpha$. We set in our computations $M_\alpha = 209$.

Remark 12.5.1. Denote by $f^*(\mathbf{x}, \mathbf{x}_\alpha)$ the noiseless data $u(\mathbf{x}, \mathbf{x}_\alpha)$, $\mathbf{x} \in \partial\Omega$, $\mathbf{x}_\alpha \in L_s$. The corresponding noisy data at the noise level $\delta > 0$ are set as

$$f^\delta(\mathbf{x}, \mathbf{x}_\alpha) = f^*(\mathbf{x}, \mathbf{x}_\alpha)(1 + \delta\,\mathrm{rand}(\mathbf{x}, \mathbf{x}_\alpha)), \quad \mathbf{x} \in \partial\Omega^+_\alpha, \mathbf{x}_\alpha \in L_s, \tag{12.47}$$

where rand is the uniformly distributed function of random numbers taking values in the range $[-1, 1]$. Recall that by (12.19) $f^*(\mathbf{x}, \mathbf{x}_\alpha) = 0$ for $\mathbf{x} \in \partial\Omega^-_\alpha$. This noise generates a noise in the boundary condition $F^{\mathbf{h}}$ in (12.42). Hence, using (12.44), we obtain $F^{\mathbf{h}} = F^{\mathbf{h}}_* + \sigma^{\mathbf{h}}$, where $\sigma^{\mathbf{h}}$ is generated by the noisy part of (12.47).

The choice of appropriate values of parameters is always a difficult task. We have selected an appropriate cut-off number N in (12.29) by a trial and error procedure. More precisely, we took Test 4 in subsection 12.5.1 with the noise level 5 % as a reference test and have selected such a value of N, which gave us the best reconstruction result. We have selected $N = 35$ this way. Then we have used the same $N = 35$ for all other tests.

We have numerically observed that the additional regularization term with the second derivatives in (12.43) is crucial. If this term is absent, then our numerical results do not meet our expectations; see Figure 12.1(g).

Remark 12.5.2. The above Theorems 12.4.1 and 12.4.2 are valid only for the case when the regularization term with the second derivatives is absent in (12.43). We also recall that proofs of those theorems are presented in [236, Section 5]. We are not sure that those theorems can be extended to the case when the second derivatives are present in (12.43). Thus, we have a discrepancy between the theory and computations. It is well known, however, that such discrepancies quite often occur in numerical studies of truly hard problems, such as, for example, the one of this publication is.

The procedure of computing $p(\mathbf{x})$ is summarized in Algorithm 1.

Algorithm 1 The procedure to solve Problem 12.2.1.

1: Choose the cut-off number $N = 35$. Find $\{\Psi_n\}_{n=1}^N$.
2: Compute the boundary data of the vector valued function $W(\mathbf{x})$.
3: Minimize the functional $I_\epsilon(W)$ subjected to the boundary condition (12.37) to obtain $W^{\text{comp}}(\mathbf{x})$, $\mathbf{x} \in \Omega$.
4: Set $w^{\text{comp}}(\mathbf{x}, \mathbf{x}_\alpha) = \sum_{n=1}^N w_n^{\text{comp}} \Psi_n(\alpha)$, $\mathbf{x} \in \Omega$, $\alpha \in [-\overline{\alpha}, \overline{\alpha}]$.
5: Set $u^{\text{comp}} = w^{\text{comp}}/\partial_z u_0$. Compute p^{comp} by the average of the left-hand side of (2.4), namely

$$p^{\text{comp}} = \frac{1}{2\overline{\alpha}\sqrt{\mathbf{a}_0(\mathbf{x})}} \int_{-\overline{\alpha}}^{\overline{\alpha}} \Big[\partial_z u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_z u^{\text{comp}}(\mathbf{x}, \mathbf{x}_\alpha) + \sum_{i=1}^{d-1} \partial_{x_i} u_0(\mathbf{x}, \mathbf{x}_\alpha) \partial_{x_i} u^{\text{comp}}(\mathbf{x}, \mathbf{x}_\alpha) \Big] d\alpha. \tag{12.48}$$

12.5.1 Numerical tests

We perform four (4) numerical tests in this paper. When indicating dependence of any function below on x, z, we assume that $(x, z) \in \Omega$, where the domain Ω is defined in (12.46). In all our tests, the noise level δ is as in (12.47). In all tests with all noise levels in the data, we use $\epsilon = 10^{-8}$.

Remark 12.5.3. In all our tests below, the function a_0 is far away from the constant background function. Therefore, Problem 12.2.1 is not considered as a small perturbation of the problem of the inverse Radon transform with incomplete data, see [156]. Some functions a_0 in our tests might not be smooth in $\mathbb{R}^2$. Still, $a_0 \in C^1(\overline{\Omega})$ in Tests 2, 3. Thus, the second derivatives of the corresponding function u_0 are well-defined in these two tests. Even though $a_0 \notin C^1(\overline{\Omega})$ in Test 1, numerically we have not experienced problems with second derivatives of the function u_0.

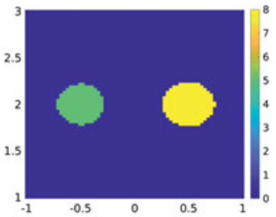
Test 1. The true source function p is given by

$$p^{\text{true}}(x, z) = \begin{cases} 8 & (x - 0.5)^2 + (z - 2)^2 < 0.24^2, \\ 5 & (x + 0.5)^2 + (z - 2)^2 < 0.22^2, \\ 0 & \text{otherwise.} \end{cases}$$

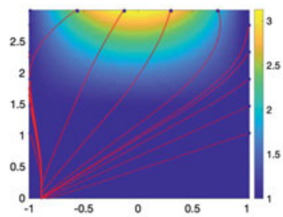
The background function $\mathbf{a}_0$ is

$$\mathbf{a}_0(x, z) = \begin{cases} 1 + 0.3(1 - x^2)(z^2 - 2) & \text{if } z^2 - 2 > 0, \\ 1 & \text{otherwise.} \end{cases}$$

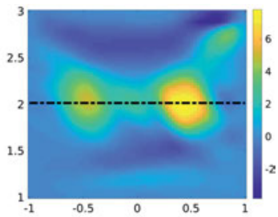
The numerical results of this test are displayed in Figure 12.1.



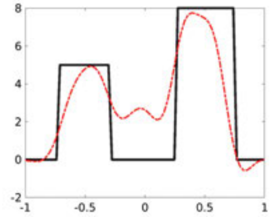
(a) The function p^{true}



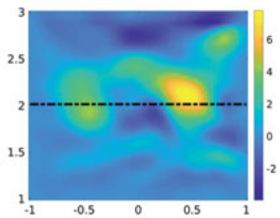
(b) The function $\mathbf{a}_0$ and some geodesic lines, generated by the Fast Marching package in Matlab



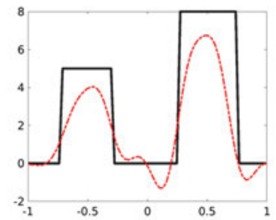
(c) The function p^{comp} computed by Algorithm 1 with 5% noise in the data



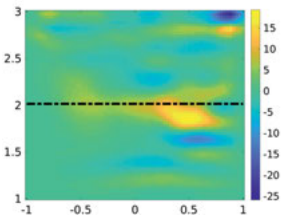
(d) The function p^{true} and p^{comp} with 5% noise in the data on the set $\{z = 2\}$, indicated by a dash-dot line in (c)



(e) The function p^{comp} computed by Algorithm 1 with 170% noise in the data



(f) The function p^{true} and p^{comp} with 170% noise in the data on the set $\{z = 2\}$, indicated by a dash-dot line in (e)



(g) The function p^{comp} in the case when second derivatives are absent in the regularization term of (12.43). 5% noise in the data. The image quality is significantly deteriorated, compared with (c).

Figure 12.1: Test 1. The true and reconstructed source functions using Algorithm 1 from noisy data.

The support of p^{true} in test 1 consists of two discs. The value of the function p in the right disc is higher than the value in the left disc. Our method detects both these inclusions very well; see Figures 12.1(c)–12.1(f). There are some unwanted artifacts near $\partial\Omega$ where we measure the noisy data. The higher level of noisy data, the more artifacts present. When the noise level $\delta = 5\,\%$, the computed maximal value of p^{comp} in the left inclusion is 4.97 (relative error 0.6 %) and the computed maximal value of p^{comp} in the right inclusion is 7.79 (relative error 2.62 %). When the noise level $\delta = 170\,\%$, the computed maximal value of p^{comp} in the left inclusion is 4.327 (relative error 13.46 %)

and the computed maximal value of p^{comp} in the right inclusion is 7.811 (relative error 2.36 %).

To verify the necessity of the presence of the second derivatives in the regularization term of (12.43), we also conduct computations for test 12.1 in the case when only the first derivatives are present in the regularization term of (12.43). The result for the case of 5 % noise in the data is depicted on Figure 12.1(g). Comparison with Figure 12.1(c) makes it evident that the presence of the second derivatives in the regularization term of (12.43) is important.

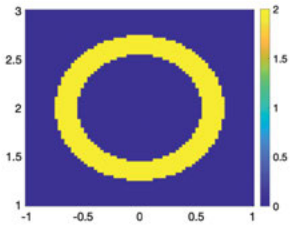
Test 2. We test a complicated case when the support of p_{true} looks like a ring. In this test,

$$p^{\text{true}}(x,z) = \begin{cases} 2 & 0.55^2 < r^2 = x^2 + (z-2)^2 < 0.75^2, \\ 0 & \text{otherwise.} \end{cases}$$

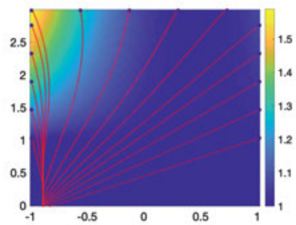
The background function $\mathbf{a}_0$ is given by

$$\mathbf{a}_0(x,z) = \begin{cases} 1 + 0.25(x-0.5)^2 \ln(z) & z > 1, \\ 1 & \text{otherwise.} \end{cases}$$

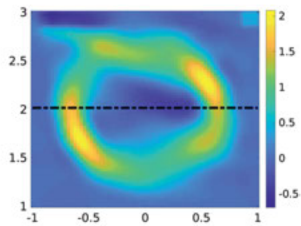
The numerical results of this test are displayed in Figure 12.2.



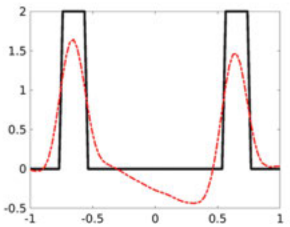
(a) The function p_{true}



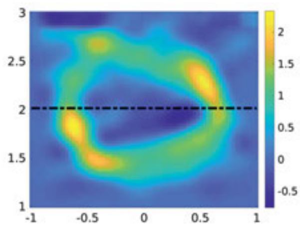
(b) The function $\mathbf{a}_0$ and some geodesic line, generated by the Fast Marching package in Matlab



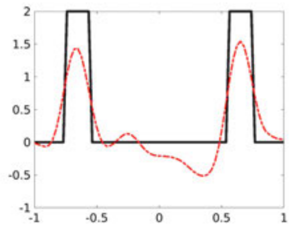
(c) The function p^{comp} computed by Algorithm 1 with 5% noise in the data



(d) The function p^{true} and p^{comp} with 5% noise in the data on the set $\{z = 2\}$, indicated by a dash-dot line in (c)



(e) The function p^{comp} computed by Algorithm 1 with 100% noise in the data



(f) The function p^{true} and p^{comp} with 100% noise in the data on the set $\{z = 2\}$, indicated by a dash-dot line in (e)

Figure 12.2: Test 2. The true and reconstructed source functions using Algorithm 1 from noisy data.

In this test, it is evident that the reconstructed "ring" is acceptable; see Figures 12.2(c) and 12.2(e). The position of the ring is detected quite well; see Figures 12.2(d) and 12.2(f). When the noise level is 5 %, the reconstructed maximal value of p^{comp} in the ring is 2.078 (relative error 3.9 %). When the noise level is 100 %, the reconstructed maximal value of p^{comp} in the ring is 2.329 (relative error 16.45 %).

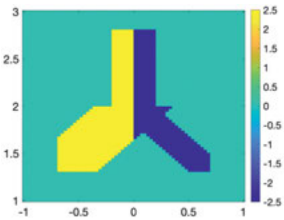
Test 3. We test an interesting and complicated case of the up-side-down letter Y having both positive and negative values. In this test, the function p_{true} is given by

$$p^{\text{true}}(x,z) = \begin{cases} 2.5 & |x-(z-2)| < 0.35, \max\{|x|,|z-2|\} < 0.7, z < 2, x < 0, \\ -2.5 & |x+(z-2)| < 0.2, \max\{|x|,|z-2|\} < 0.7, z < 2, x > 0, \\ 2.5 & |x| < 0.2, \max\{|x|,|z-2|\} < 0.8, z > 2, x < 0, \\ -2.5 & |x| < 0.2, \max\{|x|,|z-2|\} < 0.8, z > 2, x > 0. \end{cases}$$

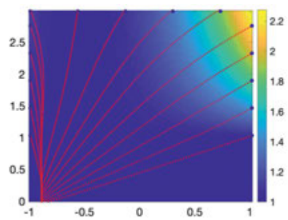
The background function $\mathbf{a}_0$ is given by

$$\mathbf{a}_0(x,z) = \begin{cases} 1 + 0.5(x+0.5)^2 \ln(z) & z > 1, \\ 1 & \text{otherwise.} \end{cases}$$

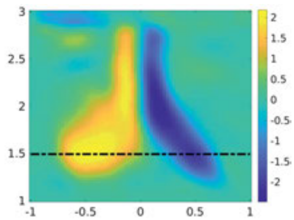
The numerical results of this test are displayed in Figure 12.3.



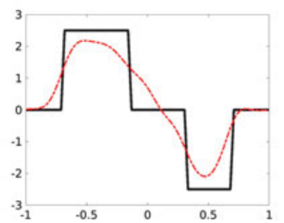
(a) The function p^{true}



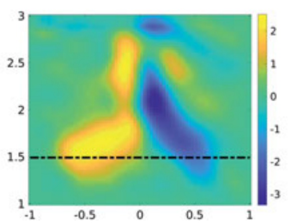
(b) The function $\mathbf{a}_0$ and some geodesic line, generated by the Fast Marching package in Matlab



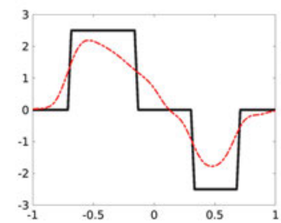
(c) The function p^{comp} computed by Algorithm 1 with 5% noise in the data



(d) The function p^{true} and p^{comp} with 5% noise in the data on the set $\{z = 1.5\}$, indicated by a dash-dot line in (c)



(e) The function p^{comp} computed by Algorithm 1 with 100% noise in the data



(f) The function p_{true} and p^{comp} with 100% noise in the data on the set $\{z = 1.5\}$, indicated by a dash-dot line in (e)

Figure 12.3: Test 3. The true and reconstructed source functions using Algorithm 1 from noisy data.

It is clear from Figure 12.3 that both positive and negative parts of the function $p(x,z)$ are successfully identified. When the noise level $\delta = 5\,\%$, the reconstructed maximal value of the positive part of p^{comp} is 2.186 (relative error 12.56 %) and the reconstructed minimal value of p^{comp} of the negative part is −2.482 (relative error 0.72 %.) When the noise level is $\delta = 100\,\%$, the reconstructed maximal value of p^{comp} of the positive part is 2.492 (relative error 0.32 %) and the reconstructed minimal value of p^{comp} of the negative part is −3.327 (relative error 33.08 %.)

Test 4. In this test, we reconstruct the letter λ. The function p^{true} is given by

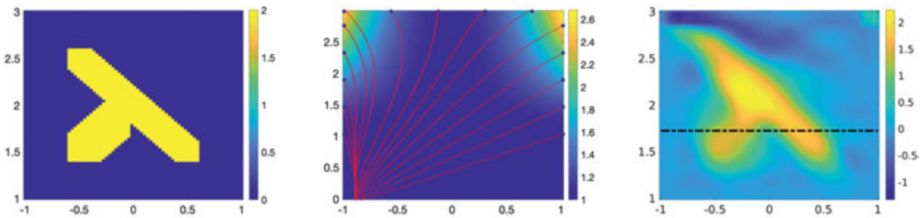
$$p^{\text{true}}(x,z) = \begin{cases} 2 & |x-(z-2)| < 0.325, \max\{|x|,|z-2|\} < 0.7 \text{ and } x < -0.03, \\ 2 & |x+(z-2)| < 0.2 \text{ and } \max\{|x|,|z-2|\} < 0.7, \\ 0 & \text{otherwise.} \end{cases}$$

In this test, we choose $\mathbf{a}_0$ as

$$\mathbf{a}_0(x,z) = \begin{cases} 1 + x^2 \ln(z) & z > 1, \\ 1 & \text{otherwise.} \end{cases}$$

The numerical results of this test are displayed in Figure 12.4.

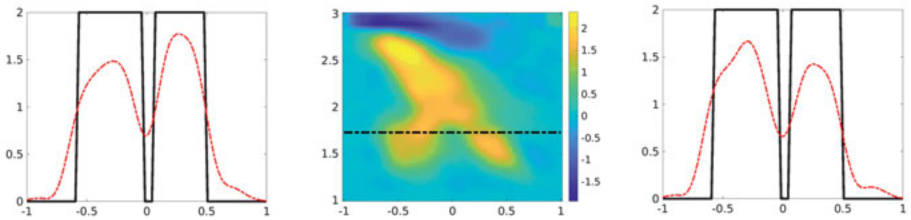
The letter λ and the values of the function p^{true} are successfully reconstructed. The computed position of λ is a quite accurate one; see Figures 12.4(d) and 12.4(f). When



(a) The function p^{true}

(b) The function $\mathbf{a}_0$ and some geodesic line, generated by the Fast Marching package in Matlab

(c) The function p^{comp} by Algorithm 1 with 5% noise in the data



(d) The function p^{true} and p^{comp} with 5% noise in the data on the set $\{z = 1.7\}$, indicated by a dash-dot line in (c)

(e) The function p^{comp} computed by Algorithm 1 with 80% noise in the data

(f) The function p_{true} and p^{comp} with 80% noise in the data on the set $\{z = 1.7\}$, indicated by a dash-dot line in (e)

Figure 12.4: Test 4. The true and reconstructed source functions using Algorithm 1 from noisy data.

the noise level $\delta = 5\,\%$, the computed maximal value of p^{comp} is 2.24 (relative error 12.0 %). When the noise level $\delta = 80\,\%$, the computed maximal value of p^{comp} is 2.375 (relative error 18.75 %).

Bibliography

[1] A. D. Agaltsov, T. Hohage and R. G. Novikov, An iterative approach to monochromatic phaseless inverse scattering, *Inverse Probl.*, 35, 24001, 2019.
[2] G. S. Alberti, H. Ammari, B. Jin, J.-K. Seo and W. Zhang, The linearized inverse problem in multifrequency electrical impedance tomography, *SIAM J. Imaging Sci.*, 9, 1525–1551, 2016.
[3] N. V. Alexeenko, V. A. Burov and O. D. Rumyanyseva, Solution of the three-dimensional acoustical inverse scattering problem. The modified Novikov algorithm, *Acoust. Phys.*, 54, 407–419, 2008.
[4] O. M. Alifanov, *Inverse Heat Conduction Problems*, Springer, New York, 1994.
[5] O. M. Alifanov, E. A. Artukhin and S. V. Rumyantcev, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begell House, New York, 1995.
[6] H. Ammari, G. S. Alberti, B. Jin, J. K. Seo and W. Zhang, The linearized inverse problem in multifrequency electrical impedance tomography, *SIAM J. Imaging Sci.*, 9, 1525–1551, 2016.
[7] H. Ammari, L. Qiu, F. Santosa and W. Zhang, Determining anisotropic conductivity using diffusion tensor imaging data in magneto-acoustic tomography with magnetic induction, *Inverse Probl.*, 33, 125006, 2017.
[8] S. Andrieux, T. N. Baranger and A. Ben Abda, Solving Cauchy problems by minimizing an energy-like functional, *Inverse Probl.*, 20, 115–133, 2006.
[9] A. B. Bakushinskii, M. V. Klibanov and N. A. Koshev, Carleman weight functions for a globally convergent numerical method for ill-posed Cauchy problems for some quasilinear PDEs, *Nonlinear Anal., Real World Appl.*, 34, 201–224, 2017.
[10] A. B. Bakushinskii and M. Yu. Kokurin, *Iterative Methods for Approximate Solutions of Inverse Problems*, Springer, New York, 2004.
[11] A. B. Bakushinsky and A. V. Goncharsky, *Ill-Posed Problems: Theory and Applications*, Kluwer Academic Publ., Boston, 1994.
[12] E. Banks, *Anti-Personnel Landmines: Recognising & Disarming (Brassey's Essential Guides)*, Pavilion, London, 1998.
[13] L. Baudouin, E. Cerpa, E. Crépeau and A. Mercado, Lipschitz stabilty in an inverse problem for the Kuramoto–Sivashinsky equation, *Appl. Anal.*, 92, 2084–2103, 2013.
[14] L. Baudouin, E. Crépeau and J. Valein, Global Carleman estimate on a network for the wave equation and application to an inverse problem, *Math. Control Relat. Fields*, 1, 1–24, 2011.
[15] L. Baudouin, M. de Buhan and S. Ervedoza, Global Carleman estimates for waves and applications, *Commun. Partial Differ. Equ.*, 38, 823–859, 2013.
[16] L. Baudouin, M. de Buhan and S. Ervedoza, Convergent algorithm based on Carleman estimates for the recovery of a potential in the wave equation, *SIAM J. Numer. Anal.*, 55, 1578–1613, 2017.
[17] L. Baudouin, M. de Buhan, S. Ervedoza and A. Osses, Carleman-based reconstruction algorithm for the waves, preprint, available at: https://hal.archives-ouvertes.fr/hal-02458787, 2020.
[18] L. Baudouin and A. Mercado, An inverse problem for Schrödinger equations with discontinuous main coefficient, *Appl. Anal.*, 87, 1145–1165, 2008.
[19] L. Baudouin and J.-P. Puel, Uniqueness and stability in an inverse problem for the Schrödinger equation, *Inverse Probl.*, 18, 1537–1554, 2002.
[20] L. Beilina, Energy estimates and numerical verification of the stabilized domain decomposition finite element/finite difference approach for the Maxwell's system in time domain, *Cent. Eur. J. Math.*, 11, 702–733, 2013.
[21] L. Beilina and M. V. Klibanov, A globally convergent numerical method for a coefficient inverse problem, *SIAM J. Sci. Comput.*, 31, 478–509, 2008.

https://doi.org/10.1515/9783110745481-013

[22] L. Beilina and M. V. Klibanov, *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*, Springer, New York, 2012.

[23] L. Beilina and M. V. Klibanov, Globally strongly convex cost functional for a coefficient inverse problem, *Nonlinear Anal., Real World Appl.*, 22, 272–288, 2015.

[24] L. Beilina, N. T. Thành, M. Klibanov and J. B. Malmberg, Reconstruction of shapes and refractive indices from blind backscattering experimental data using the adaptivity, *Inverse Probl.*, 105007, 2014.

[25] L. Beilina, N. T. Thánh, M. V. Klibanov and M. A. Fiddy, Reconstruction from blind experimental data for an inverse problem for a hyperbolic equation, *Inverse Probl.*, 30, 025002, 2014.

[26] M. I. Belishev, Recent progress in the boundary control method, *Inverse Probl.*, 23, R1–R67, 2007.

[27] M. I. Belishev, I. B. Ivanov, I. V. Kubyshkin and V. S. Semenov, Numerical testing in determination of sound speed from a part of boundary by the BC-method, *J. Inverse Ill-Posed Probl.*, 24, 159–180, 2016.

[28] M. Bellassoued, Uniqueness and stability in determining the speed of propagation of second-order hyperbolic equation with variable coefficients, *Appl. Anal.*, 83, 983–1014, 2004.

[29] M. Bellassoued, O. Yu. Imanuvilov and M. Yamamoto, Inverse problem of determining the density and two Lamé coefficients by boundary data, *SIAM J. Math. Anal.*, 40, 238–265, 2008.

[30] M. Bellassoued and M. Yamamoto, Logarithmic stability in determination of a coefficient in an acoustic equation by arbitrary boundary observation, *J. Math. Pures Appl.*, 85, 193–224, 2006.

[31] M. Bellassoued and M. Yamamoto, Inverse source problem for a transmission problem for a parabolic equation, *J. Inverse Ill-Posed Probl.*, 14, 47–56, 2006.

[32] M. Bellassoued and M. Yamamoto, Carleman estimates and an inverse heat source problem for the thermoelasticity system, *Inverse Probl.*, 27, 015006, 2011.

[33] M. Bellassoued and M. Yamamoto, Carleman estimate with second large parameter for second order hyperbolic operators in a Riemannian manifold and applications to thermoelasticity cases, *Appl. Anal.*, 91, 35–67, 2012.

[34] M. Bellassoued and M. Yamamoto, Carleman estimate and inverse source problem for Biot's equations describing wave propagation in porous media, *Inverse Probl.*, 29, 115002, 2013.

[35] M. Bellassoued and M. Yamamoto, *Carleman Estimates and Applications to Inverse Problems for Hyperbolic Systems*, Springer Japan KK, 2017.

[36] I. N. Bernstein and M. L. Gerver, On a problem of integral geometry for family of geodesics and the inverse kinematic problem of seismic, *Dokl. Akad. Nauk SSSR*, 243, 302–305, 1978.

[37] F. Berntsson, V. A. Kozlov, L. Mpinganzima and B. O. Turesson, An alternating iterative procedure for the Cauchy problem for the Helmholtz equation, *Inverse Probl. Sci. Eng.*, 22, 45–62, 2014.

[38] L. Borcea, Electrical impedance tomography, *Inverse Probl.*, 18, R99–R136, 2002.

[39] M. Boulakia, C. Grandmont and A. Osses, Some inverse stability results for the bi-stable reaction-diffusion equation using Carleman estimates, *C. R. Acad. Sci. Paris, Ser. I*, 347, 619–622, 2009.

[40] L. Bourgeois, A mixed formulation of quasi-reversibility to solve the Cauchy problem for Laplace's equation, *Inverse Probl.*, 21(3), 1087–1104, 2005.

[41] L. Bourgeois, Convergence rates for the quasi-reversibility method to solve the Cauchy problem for Laplace's equation, *Inverse Probl.*, 22, 413–430, 2006.

[42] L. Bourgeois and J. Dardé, A duality-based method of quasi-reversibility to solve the Cauchy problem in the presence of noisy data, *Inverse Probl.*, 26, 095016, 2010.

[43] L. Bourgeois and J. Dardé, A quasi-reversibility approach to solve the inverse obstacle problem, *Inverse Probl. Imaging*, 4, 351–377, 2010.

[44] L. Bourgeois and J. Dardé, About identification of defects in an elastic-plastic medium from boundary measurements in the antiplane case, *Appl. Anal.*, 90, 1481–1497, 2011.
[45] L. Bourgeois and J. Dardé, The “exterior approach” to solve the inverse obstacle problem for the Stokes system, *Inverse Probl. Imaging*, 8, 23–51, 2014.
[46] L. Bourgeois and E. Linéville, The method of quasi-reversibility to solve the Cauchy problems for elliptic partial differential equations, *Proc. Appl. Math. Mech.*, 7, 1042101–1042102, 2007.
[47] L. Bourgeois, D. Ponomarev and J. Dardé, And inverse obstacle problem for the wave equation in a finite domain, *Inverse Probl. Imaging*, 13, 377–400, 2019.
[48] A. L. Bukhgeim, Carleman estimates for Volterra operators and uniqueness of inverse problems, in: *Non-Classical Problems of Mathematical Physics*, Published by Computing Center of the Siberian Branch of USSR Academy of Science, Novosibirsk, pages 54–64, 1981 (in Russian).
[49] A. L. Bukhgeim, Carleman estimates for Volterra operators and uniqueness of inverse problems, *Sib. Math. J.*, 25, 43–50, 1984.
[50] A. L. Bukhgeim, *Introduction in the Theory of Inverse Problems*, VSP, Utrecht, The Netherlands, 2000.
[51] A. L. Bukhgeim and M. V. Klibanov, Global uniqueness of a class of multidimensional inverse problems, *Dokl. Akad. Nauk SSSR*, 260, 269–272, 1981. English translation: A. L. Bukhgeim and M. V. Klibanov, Global uniqueness of a class of multidimensional inverse problems, *Soviet Mathematics Doklady*, 24, 244–247, 1981.
[52] V. A. Burov, N. V. Alekseenko and O. D. Rumyantseva, Multifrequency generalization of the Novikov algorithm for the two-dimensional inverse scattering problem, *Acoust. Phys.*, 55, 843–856, 2009. 6.
[53] V. A. Burov, S. A. Morozov and O. D. Rumyantseva, Reconstruction of fine-scale structure of acoustical scatterers on large-scale contrast background, *Acoust. Imag.*, 26, 231–238, 2002.
[54] F. Cakoni and D. Colton, *Qualitative Methods in Inverse Scattering Theory. An Introduction*, Springer, Berlin, 2006.
[55] A. P. Calderon, Uniqueness in the Cauchy problem for partial differential equations, *Am. J. Math.*, 80, 16–36, 1958.
[56] A. P. Calderon, On an inverse boundary value problem, in: *Seminar on Numerical Analysis and Its Applications to Continuum Physics (Rio de Janeiro)*, Soc. Brasil. Mat., Rio de Janeiro, pages 65–73, 1980.
[57] H. Cao, M. V. Klibanov and S. V. Pereverzev, A Carleman estimate and the balancing principle in the quasi-reversibility method for solving the Cauchy problem for the Laplace equation, *Inverse Probl.*, 25, 35005, 2009.
[58] T. Carleman, Sur un probléme d’unicité pur les systeémes d’équations aux dériveés partielles à deux varibales indépendantes, *Ark. Mat. Astron. Fys.*, 26B(17), 1–9, 1939.
[59] C. Cavaterra, A. Lorenzi and M. Yamamoto, A stability result via Carleman estimates for an inverse source problem related to a hyperbolic integro-differential equation, *Comput. Appl. Math.*, 25, 229–250, 2006.
[60] G. Chavent, *Nonlinear Least Squares for Inverse Problems: Theoretical Foundations and Step-by-Step Guide for Applications*, Springer, New York, 2009.
[61] Z. Chen, R. Tuo and W. Zhang, Stochastic convergence of a nonconforming finite element method for the thin plate spline smoother for observational data, *SIAM J. Numer. Anal.*, 56, 635–659, 2018.
[62] M. Chouli, O. Yu. Imanuvilov, J.-P. Puel and M. Yamamoto, Inverse source problem for linearized Navier–Stokes equations with data in arbitrary sub-domain, *Appl. Anal.*, 92, 2127–2143, 2013.
[63] Y. T. Chow and J. Zou, A numerical method for reconstructing the coefficient in a wave equation, *Numer. Methods Partial Differ. Equ.*, 31, 289–307, 2015.

[64] C. Clason and M. V. Klibanov, The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medim, *SIAM J. Sci. Comput.*, 30, 1–23, 2007.
[65] Clipper Controls Inc. (n.d.) Dielectric Constants of various materials, http://www.clippercontrols.com/pages/Dielectric-Constant-Values.html#.
[66] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 3rd edn, Springer, New York, 2013.
[67] M. Cristofol, P. Gaitan and H. Ramoul, Inverse problems for a 2×2 reaction-diffusion system using a Carleman estimate with one observation, *Inverse Probl.*, 22, 1561–1573, 2006.
[68] J. Dardé, Iterated quasi-reversibility method applied to elliptic and parabolic data completion problems, *Inverse Probl. Imaging*, 10, 379–407, 2016.
[69] J. Dardé, A. Hannukainen and N. Hyvönen, An H_{div}-based mixed quasi-reversibility method for solving elliptic Cauchy problems, *SIAM J. Numer. Anal.*, 51, 2123–2148, 2013.
[70] B. Das, F. Liu and R. R. Alfano, Time-resolved fluorescence and photon migration studies in biomedical and model random media, *Rep. Prog. Phys.*, 60, 227–292, 1997.
[71] M. V. de Hoop, P. Kepley and L. Oksanen, Recovery of a smooth metric via wave field and coordinate transformation reconstruction, *SIAM J. Appl. Math.*, 78, 1931–1933, 2018.
[72] A. A. Duchkov and A. L. Karchevskii, Application of temperature monitoring to estimate the heat flux and thermophysical properties of bottom sediments, *Dokl. Earth Sci.*, 458(Part 2), 1285–1288, 2014.
[73] H. Egger, H. W. Engl and M. V. Klibanov, Global uniqueness and Hölder stability for recovering a nonlinear source term in a parabolic equation, *Inverse Probl.*, 21, 271–290, 2005.
[74] L. Eldén, Numerical solution of the sideways heat equation by difference approximation in time, *Inverse Probl.*, 11, 913–923, 1995.
[75] L. Eldén, F. Berntsson and T. Regińska, Wavelet and Fourier methods for solving the sideways heat equation, *SIAM J. Sci. Comput.*, 21, 2187–2205, 2000.
[76] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Ill-Posed Problems*, Kluwer Academic Publishers Group, Dordrecht, 2000.
[77] L. Fat Ho, Observabilité frontierè de l'equation des ondes, *C. R. Acad. Sci. Paris, Ser. I*, 302, 443–446, 1986.
[78] G. Floridia and H. Takase, Inverse problems for first-order hyperbolic equations with time dependent coefficients, arXiv:2009.12039, 2020.
[79] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice Hall, Inc., Englewood Cliffs, 1962.
[80] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer Verlag, Berlin, 1983.
[81] A. V. Goncharsky and S. Y. Romanov, Supercomputer technologies in inverse problems of ultrasound tomography, *Inverse Probl.*, 29, 075004, 2013.
[82] A. V. Goncharsky and S. Y. Romanov, Iterative methods for solving coefficient inverse problems of wave tomography in models with attenuation, *Inverse Probl.*, 33, 025003, 2017.
[83] A. V. Goncharsky and S. Y. Romanov, A method of solving the coefficient inverse problems of wave tomography, *Comput. Math. Appl.*, 77, 967–980, 2019.
[84] J. P. Guillement and R. G. Novikov, Inversion of weighted Radon transforms via finite Fourier series weight approximation, *Inverse Probl. Sci. Eng.*, 22, 787–802, 2013.
[85] B. Harrach, Uniqueness, stability and global convergence for a discrete inverse elliptic Robin transmission problem, *Numer. Math.*, published online https://doi.org/10.1007/s00211-020-01162-8, 2020.
[86] B. Harrach and M. N. Minh, Enhancing residual-based techniques with shape reconstruction features in electrical impedance tomography, *Inverse Probl.*, 32, 125002, 2016.
[87] A. H. Hasanoğlu and V. G. Romanov, *Introduction to Inverse Problems for Differential Equations*, Springer, 2017.

[88] G. Herglotz, Über die Elastizitaet der Erde bei Beruecksichtigung ihrer variablen Dichte, *Z. Angew. Math. Phys.*, 52, 275–299, 1905.

[89] N. Hyvonen, P. Piiroinen and O. Seiskari, Point measurements for a Neumann-to-Dirichlet map and the Calderon problem in the plane, *SIAM J. Math. Anal.*, 44, 3526–3536, 2012.

[90] D. N. Hào, A noncharacteristic Cauchy problem for linear parabolic equations II: a variational method, *Numer. Funct. Anal. Optim.*, 13, 541–564, 1992.

[91] D. N. Hào, N. V. Duc and N. V. Thang, Backward semi-linear parabolic equations with time-dependent coefficients and local Lipschitz source, *Inverse Probl.*, 34, 05510, 2018.

[92] D. N. Hào and D. Lesnic, The Cauchy problem for Laplace's equation via the conjugate gradient method, *IMA J. Appl. Math.*, 65, 199–217, 2000.

[93] L. Hörmander, *Linear Partial Differential Operators*, Springer, Berlin, 1963.

[94] L. Ignat, A. F. Pazoto and L. Rosier, Inverse problem for the heat equation and the Schrödinger equation on a tree, *Inverse Probl.*, 28, 015011, 2012.

[95] O. Yu. Imanuvilov, V. Isakov and M. Yamamoto, An inverse problem for the dynamical Lamé system with two sets of boundary data, *Commun. Pure Appl. Math.*, 56, 1–17, 2003.

[96] O. Yu. Imanuvilov and M. Yamamoto, Lipschitz stability in inverse parabolic problems by the Carleman estimate, *Inverse Probl.*, 14, 1229–1245, 1998.

[97] O. Yu. Imanuvilov and M. Yamamoto, Global uniqueness and stability in determining coefficients of wave equations, *Commun. Partial Differ. Equ.*, 26, 1409–1425, 2001.

[98] O. Yu. Imanuvilov and M. Yamamoto, Global Lipschitz stability in an inverse hyperbolic problem by interior observations, *Inverse Probl.*, 17, 717–728, 2001.

[99] O. Yu. Imanuvilov and M. Yamamoto, Determination of a coefficient in an acoustic equation with a single measurement, *Inverse Probl.*, 19, 157–171, 2003.

[100] V. Isakov, *Inverse Source Problems*, American Mathematical Society, Providence, RI, 1990.

[101] V. Isakov, Carleman estimates and applications to inverse problems, *Milan J. Math.*, 72, 249–271, 2004.

[102] V. Isakov, *Inverse Problems for Partial Differential Equations*, 2nd edn, Springer, New York, 2006.

[103] V. Isakov and M. Yamamoto, Carleman estimate with the Neumann boundary condition and its applications to the observability inequality and inverse hyperbolic problems, *Contemp. Math.*, 268, 191–225, 2000.

[104] K. Ito, B. Jin and J. Zou, A direct sampling method to an inverse medium scattering problem, *Inverse Probl.*, 28, 025003, 2012.

[105] B. Jin, T. Khan and P. Maass, A reconstruction algorithm for electrical impedance tomography based on sparsity regularization, *Int. J. Numer. Methods Eng.*, 89, 337–353, 2012.

[106] B. Jin, Y. Xu and J. Zou, A convergent adaptive finite element method for electrical impedance tomography, *IMA J. Numer. Anal.*, 37, 1520–1550, 2017.

[107] S. I. Kabanikhin and A. L. Karchevsky, Optimizational method for solving the Cauchy problem for an elliptic equation, *J. Inverse Ill-Posed Probl.*, 3, 21–46, 1995.

[108] S. I. Kabanikhin, N. S. Novikov, I. V. Osedelets and M. A. Shishlenin, Fast Toeplitz linear system inversion for solving two-dimensional acoustic inverse problem, *J. Inverse Ill-Posed Probl.*, 23, 687–700, 2015.

[109] S. I. Kabanikhin, K. K. Sabelfeld, N. S. Novikov and M. A. Shishlenin, Numerical solution of the multidimensional Gelfand-Levitan equation, *J. Inverse Ill-Posed Probl.*, 23, 439–450, 2015.

[110] S. I. Kabanikhin, A. D. Satybaev and M. A. Shishlenin, *Direct Methods for Solving Multidimensional Inverse Hyperbolic Problems*, VSP, Utrecht, The Netherlands, 2004.

[111] S. I. Kabanikhin, A. D. Satybaev and M. A. Shishlenin, *Direct Methods of Solving Inverse Hyperbolic Problems*, VSP, The Netherlands, 2005.

[112] S. I. Kabanikhin and M. A. Shishlenin, Numerical algorithm for two-dimensional inverse acoustic problem based on Gel'fand-Levitan–Krein equation, *J. Inverse Ill-Posed Probl.*, 18, 979–995, 2011.

[113] B. Kaltenbacher and M. V. Klibanov, An inverse problem for a nonlinear parabolic equation with applications in population dynamics and magnetics, *SIAM J. Math. Anal.*, 39, 1863–1889, 2008.

[114] M. Kazemi and M. V. Klibanov, Stability estimates for ill-posed Cauchy problem involving hyperbolic equation and inequalities, *Appl. Anal.*, 50, 93–102, 1993.

[115] V. A. Khoa, G. W. Bidney, M. V. Klibanov, L. H. Nguyen, A. J. Sullivan, L. Nguyen and V. N. Astratov, Convexification and experimental data for a 3D inverse scattering problem with the moving point source, *Inverse Probl.*, 36, 085007, 2020.

[116] V. A. Khoa, G. W. Bidney, M. V. Klibanov, L. H. Nguyen, A. J. Sullivan, L. Nguyen and V. N. Astratov, An inverse problem of a simultaneous reconstruction of the dielectric constant and conductivity from experimental backscattering data, *Inverse Probl. Sci. Eng.*, 29, 712–735, 2021.

[117] V. A. Khoa, M. V. Klibanov and L. H. Nguyen, Convexification for a three-dimensional inverse scattering problem with the moving point source, *SIAM J. Imaging Sci.*, 13, 871–904, 2020.

[118] Y. Kian, Q. S. Phan and E. Soccorsi, A Carleman estimate for infinite cylindrical quantum domains and the application to inverse problems, *Inverse Probl.*, 30, 055016, 2014.

[119] Y. Kian, Q. S. Phan and E. Soccorsi, Hölder stable determination of a quantum stable potential in unbounded cylindrical domains, *J. Math. Anal. Appl.*, 426, 194–210, 2015.

[120] M. V. Klibanov, Uniqueness of solutions in the 'large' of some multidimensional inverse problems, in: *Non-Classical Problems of Mathematical Physics*, Published by Computing Center of the Siberian Branch of the USSR Academy of Science, Novosibirsk, pages 101–114, 1981 (in Russian).

[121] M. V. Klibanov, On a class of inverse problems, *Sov. Math. Dokl.*, 26, 248–250, 1982.

[122] M. V. Klibanov, Inverse problems in the 'large' and Carleman bounds, *Differ. Equ.*, 20, 755–760, 1984.

[123] M. V. Klibanov, Uniqueness of inverse problems in the large for one class of differential equations, *Differ. Equ.*, 20, 1667–1671, 1984.

[124] M. V. Klibanov, Uniqueness of solutions of two inverse problems for the Maxwellian system, *USSR Comput. Math. Math. Phys.*, 26, 1063–1071, 1986.

[125] M. V. Klibanov, A class of inverse problems for nonlinear parabolic equations, *Sib. Math. J.*, 27, 698–707, 1987.

[126] M. V. Klibanov, Inverse problems and Carleman estimates, *Inverse Probl.*, 8, 575–596, 1992.

[127] M. V. Klibanov, Global convexity in a three-dimensional inverse acoustic problem, *SIAM J. Math. Anal.*, 28, 1371–1388, 1997.

[128] M. V. Klibanov, Carleman estimates and inverse problems in the last two decades, in: *Surveys on Solutions Methods for Inverse Problems*, Springer, Wien, pages 119–146, 2000.

[129] M. V. Klibanov, Global uniqueness of a multidimensional inverse problem for a nonlinear parabolic equation, *Inverse Probl.*, 20, 1003–1032, 2004.

[130] M. V. Klibanov, Estimates of initial conditions of parabolic equations and inequalities via lateral Cauchy data, *Inverse Probl.*, 22, 495–514, 2006.

[131] M. V. Klibanov, Uniqueness of an inverse problem with single measurement data generated by a plane wave in partial finite differences, *Inverse Probl.*, 27, 115005, 2011.

[132] M. V. Klibanov, Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems, *J. Inverse Ill-Posed Probl.*, 21, 477–510, 2013.

[133] M. V. Klibanov, Thermoacoustic tomography with an arbitrary elliptic operator, *Inverse Probl.*, 29, 025014, 2013.

[134] M. V. Klibanov, Carleman estimates for the regularization of ill-posed Cauchy problems, *Appl. Numer. Math.*, 94, 46–74, 2015.

[135] M. V. Klibanov, Carleman weight functions for solving ill-posed Cauchy problems for quasilinear PDEs, *Inverse Probl.*, 31, 125007, 2015.

[136] M. V. Klibanov, Convexification of restricted Dirichlet-to-Neumann map, *J. Inverse Ill-Posed Probl.*, 25, 669–685, 2017.
[137] M. V. Klibanov, On the travel time tomography problem in 3D, *J. Inverse Ill-Posed Probl.*, 27, 591–607, 2019.
[138] M. V. Klibanov, Travel time tomography with formally determined incomplete data in 3D, *Inverse Probl. Imaging*, 13, 1367–1393, 2019.
[139] M. V. Klibanov, M. A. Fiddy, L. Beilina, N. Pantong and J. Schenk, Picosecond scale experimental verification of a globally convergent numerical method for a coefficient inverse problem, *Inverse Probl.*, 26, 045003, 2010.
[140] M. V. Klibanov and O. V. Ioussoupova, Uniform strict convexity of a cost functional for three-dimensional inverse scattering problem, *SIAM J. Math. Anal.*, 26, 147–179, 1995.
[141] M. V. Klibanov and V. G. Kamburg, Globally strictly convex cost functional for an inverse parabolic problem, *Math. Methods Appl. Sci.*, 39, 930–940, 2016.
[142] M. V. Klibanov and A. E. Kolesov, Convexification of a 3-D coefficient inverse scattering problem, *Comput. Math. Appl.*, 77, 1681–1702, 2019.
[143] M. V. Klibanov, A. E. Kolesov and D.-L. Nguyen, Convexification method for a coefficient inverse problem and its performance for experimental backscatter data for buried targets, *SIAM J. Imaging Sci.*, 12, 576–603, 2019.
[144] M. V. Klibanov, A. E. Kolesov, L. Nguyen and A. Sullivan, Globally strictly convex cost functional for a 1-D inverse medium scattering problem with experimental data, *SIAM J. Appl. Math.*, 77, 1733–1755, 2017.
[145] M. V. Klibanov, A. E. Kolesov, L. Nguyen and A. Sullivan, A new version of the convexification method for a 1-D coefficient inverse problem with experimental data, *Inverse Probl.*, 34, 115014, 2018.
[146] M. V. Klibanov, A. N. Koshev, J. Li and A. G. Yagola, Numerical solution of an ill-posed Cauchy problem for a quasilinear parabolic equation using a Carleman weight function, *J. Inverse Ill-Posed Probl.*, 24, 761–776, 2016.
[147] M. V. Klibanov, A. V. Kuzhuget, S. I. Kabanikhin and D. V. Nechaev, A new version of the quasi-reversibility method for the thermoacoustic tomography and a coefficient inverse problem, *Appl. Anal.*, 87, 1227–1254, 2008.
[148] M. V. Klibanov, T. T. Le and L. H. Nguyen, Convergent numerical method for a linearized travel time tomography problem with incomplete data, *SIAM J. Sci. Comput.*, 45, B1173–B1192, 2020.
[149] M. V. Klibanov, J. Li and W. Zhang, Electrical impedance tomography with restricted Dirichlet-to-Neumann map data, *Inverse Probl.*, 35, 035005, 2019.
[150] M. V. Klibanov, J. Li and W. Zhang, Convexification for the inversion of a time dependent wave front in a heterogeneous medium, *SIAM J. Appl. Math.*, 79, 1722–1747, 2019.
[151] M. V. Klibanov, J. Li and W. Zhang, Convexification for an inverse parabolic problem, *Inverse Probl.*, 36, 085008, 2020.
[152] M. V. Klibanov, J. Li and W. Zhang, Linear Lavrent'ev integral equation for the numerical solution of a nonlinear coefficient inverse problem, arXiv:2010.14144, to appear in *SIAM J. Appl. Math.* in 2021.
[153] M. V. Klibanov and J. Malinsky, Newton–Kantorovich method for 3-dimensional potential inverse scattering problem and stability for the hyperbolic Cauchy problem with time dependent data, *Inverse Probl.*, 7, 577–596, 1991.
[154] M. V. Klibanov, D.-L. Nguyen, L. H. Nguyen and H. Liu, A globally convergent numerical method for a 3D coefficient inverse problem with a single measurement of multi-frequency data, *Inverse Probl. Imaging*, 12, 493–523, 2018.
[155] M. V. Klibanov, D. L. Nguyen and L. H. Nguyen, A coefficient inverse problem with a single measurement of phaseless data, *SIAM J. Appl. Math.*, 79, 1–27, 2019.

[156] M. V. Klibanov and L. H. Nguyen, PDE-based numerical method for a limited angle X-ray tomography, *Inverse Probl.*, 35, 045009, 2019.
[157] M. V. Klibanov, L. H. Nguyen, A. Sullivan and L. Nguyen, A globally convergent numerical method for a 1-D inverse problem with experimental data, *Inverse Probl. Imaging*, 10, 1057–1085, 2016.
[158] M. V. Klibanov and Rakesh, Numerical solution of a timelike Cauchy problem for the wave equation, *Math. Methods Appl. Sci.*, 15, 559–570, 1992.
[159] M. V. Klibanov and V. G. Romanov, Reconstruction procedures for two inverse scattering problems without the phase information, *SIAM J. Appl. Math.*, 76, 178–196, 2016.
[160] M. V. Klibanov and V. G. Romanov, Two reconstruction procedures for a 3-d phaseless inverse scattering problem for the generalized Helmholtz equation, *Inverse Probl.*, 32, 015005, 2016.
[161] M. V. Klibanov and F. Santosa, A computational quasi-reversibility method for Cauchy problems for Laplace's equation, *SIAM J. Appl. Math.*, 51, 1653–1675, 1991.
[162] M. V. Klibanov, A. V. Smirnov, K. A. Vo, A. J. Sullivan and L. H. Nguyen, Through-the-wall nonlinear SAR imaging, *IEEE Transactions on Geoscience and Remote Sensing*, published online by this journal https://doi.org/10.1109/TGRS.2021.3055805, 2021.
[163] M. V. Klibanov, J. Su, N. Pantong, H. Shan and H. Liu, A globally convergent numerical method for an inverse elliptic problem of optical tomography, *Appl. Anal.*, 89, 861–891, 2010.
[164] M. V. Klibanov and N. T. Thánh, Recovering of dielectric constants of explosives via a globally strictly convex cost functional, *SIAM J. Appl. Math.*, 75, 518–537, 2015.
[165] M. V. Klibanov and A. Timonov, *Carleman Estimates for Coefficient Inverse Problems and Numerical Applications*, VSP, 2004.
[166] M. V. Klibanov and A. G. Yagola, Convergent numerical methods for parabolic equations with reversed time, *Inverse Probl.*, 34, 115012, 2019.
[167] M. V. Klibanov and M. Yamamoto, Lipschitz stability for an inverse problem for an acoustic equation, *Appl. Anal.*, 85, 515–538, 2006.
[168] A. E. Kolesov, M. V. Klibanov, L. H. Nguyen, D.-L. Nguyen and N. T. Thanh, Single measurement experimental data for an inverse medium problem inverted by a multi-frequency globally convergent numerical method, *Appl. Numer. Math.*, 120, 176–196, 2017.
[169] V. A. Kozlov, V. G. Maz'ya and A. V. Fomin, An iterative method for solving the Cauchy problem for elliptic equations, *Ž. Vyčisl. Mat. Mat. Fiz.*, 31, 64–74, 1991.
[170] A. V. Kuzhuget, L. Beilina, M. V. Klibanov, A. Sullivan, L. Nguyen and M. A. Fiddy, Blind experimental data collected in the field and an approximately globally convergent inverse algorithm, *Inverse Probl.*, 28, 095007, 2012.
[171] A. V. Kuzhuget, L. Beilina, M. V. Klibanov, A. Sullivan, L. Nguyen and M. A. Fiddy, Quantitative image recovery from measured blind backscattered data using a globally convergent inverse method, *IEEE Trans. Geosci. Remote Sens.*, 51, 2937–2948, 2013.
[172] A. V. Kuzhuget and M. V. Klibanov, Global convergence for a 1-D inverse problem with application to imaging of land mines, *Appl. Anal.*, 89, 125–157, 2010.
[173] O. A. Ladyzhenskaya, *Boundary Value Problems of Mathematical Physics*, Springer Verlag, Berlin, 1985.
[174] O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Uralceva, *Linear and Quasilinear Equations of Parabolic Type*, AMS, Providence, R. I., 1968.
[175] A. Lakhal, A decoupling-based imaging method for inverse medium scattering for Maxwell's equations, *Inverse Probl.*, 26, 015007, 2010.
[176] I. Lasiecka, B. Priyasad and R. Triggiani, Finite dimensional boundary uniform stabilization of the Boussinesq system in Besov spaces by critical use of Carleman estimate-based inverse theory, preprint, 2020.
[177] I. Lasiecka, R. Triggiani and P. F. Yao, Inverse/observability estimates for second order hyperbolic equations with variable coefficients, *J. Math. Anal. Appl.*, 235, 13–57, 1999.

[178] I. Lasiecka, R. Triggiani and X. Zhang, Nonconservative wave equations with unobserved Neumann BC: global uniqueness and observability, *Contemp. Math.- Am. Math. Soc.*, 268, 227–326, 2000.
[179] I. Lasiecka, R. Triggiani and X. Zhang, Global uniqueness, observability and stabilization of non-conservative Schrödinger equations via pointwise Carleman estimates. Part I: $H^1(\Omega)$-estimates, *J. Inverse Ill-Posed Probl.*, 12, 1–81, 2004.
[180] I. Lasiecka, R. Triggiani and X. Zhang, Global uniqueness, observability and stabilization of non-conservative Schrödinger equations via pointwise Carleman estimates. Part II: $L_2(\Omega)$-estimates, *J. Inverse Ill-Posed Probl.*, 12, 182–231, 2004.
[181] J. Lassoued, M. Mahjoub and N. Zemzemi, Stability results for the parameter identification inverse problem in cardiac electrophysiology, *Inverse Probl.*, 32, 115002, 2016.
[182] R. Lattes and J.-L. Lions, *The Method of Quasireversibility: Applications to Partial Differential Equations*, Elsevier, New York, 1969.
[183] M. M. Lavrent'ev, On an inverse problem for the wave equation, *Sov. Math. Dokl.*, 5, 970–972, 1964.
[184] M. M. Lavrent'ev, V. G. Romanov and S. P. Shishatskii, *Ill-Posed Problems of Mathematical Physics and Analysis*, AMS, Providence, R. I., 1986.
[185] T. T. Le and L. H. Nguyen, A convergent numerical method to recover the initial condition of nonlinear parabolic equations from lateral Cauchy data, *J. Inverse Ill-Posed Probl.*, published online https://doi.org/10.1515/jiip-2020-0028, 2020.
[186] J. Li, P. Li, H. Liu and X. Liu, Recovering multiscale buried anomalies in a two-layered medium, *Inverse Probl.*, 31, 105006, 2015.
[187] J. Li, H. Liu, Z. Shang and H. Sun, Two single-shot methods for locating multiple electromagnetic scattereres, *SIAM J. Appl. Math.*, 73, 1721–1746, 2013.
[188] J. Li, H. Liu and Q. Wang, Enhanced multilevel linear sampling methods for inverse scattering problems, *J. Comput. Phys.*, 257, 554–571, 2014.
[189] J. Li and J. Zou, A multilevel model correction method for parameter identification, *Inverse Probl.*, 23, 1759, 2007.
[190] Q. Li and L. H. Nguyen, Recovering the initial condition of parabolic equations from lateral Cauchy data via the quasi-reversibility method, *Inverse Probl. Sci. Eng.*, 28, 580–598, 2020.
[191] S. Li, An inverse problem for Maxwell's equations in bi-isotropic media, *SIAM J. Math. Anal.*, 37, 1027–1043, 2005.
[192] S. Liu and R. Triggiani, Global uniqueness in determining electric potentials for a system of strongly coupled Schröödinger equations with magnetic potential terms, *J. Inverse Ill-Posed Probl.*, 19, 223–254, 2011.
[193] S. Liu and R. Triggiani, Global uniqueness and stability in determining the damping coefficient of an inverse hyperbolic problem with non homogeneous Neumann B. C. through an additional Dirichlet boundary trace, *SIAM J. Math. Anal.*, 43, 1631–1666, 2011.
[194] S. Liu and R. Triggiani, Global uniqueness and stability in determining the damping and potential coefficients of an inverse hyperbolic problem, *Nonlinear Anal., Real World Appl.*, 12, 1562–1590, 2011.
[195] S. Liu and R. Triggiani, Global uniqueness and stability in determining in determining the damping coefficient of an inverse hyperbolic problem with non-homogeneous Dirichlet BC through an additional localized Neumann boundary condition, *Appl. Anal.*, 91, 1551–1581, 2012.
[196] Y. Liu, J. Su, Z.-J. Lin, S. Teng, A. Rhoden, N. Pantong and H. Liu, Reconstructions for continuous-wave diffuse optical tomography by a globally convergent method, *J. Appl. Math. Phys.*, 2, 204–213, 2014.
[197] Q. Lü, Carleman estimate for stochastic parabolic equations and inverse stochastic parabolic problems, *Inverse Probl.*, 28, 045008, 2012.

[198] A. Mercado, A. Osses and L. Rosier, Inverse problems for the Schrödinger equation via Carleman inequalities with degenerate weights, *Inverse Probl.*, 24, 015017, 2008.

[199] F. Monard, Numerical implementation of geodesic X-ray transforms and their inversion, *SIAM J. Imaging Sci.*, 7, 1335–1357, 2014.

[200] J. L. Mueller and S. Siltanen, The D-bar method for electrical impedance tomography—demystified, *Inverse Probl.*, 36, 093001, 2020.

[201] R. G. Mukhometov, The reconstruction problem of a two-dimensional Riemannian metric and integral geometry, *Sov. Math. Dokl.*, 18, 32–35, 1977.

[202] R. G. Mukhometov and V. G. Romanov, On the problem of determining an isotropic Riemannian metric in the *n*-dimensional space, *Dokl. Acad. Sci. USSR*, 19, 1330–1333, 1978.

[203] A. Nachman, Global uniqueness for a two-dimensional inverse boundary value problem, *Ann. Math.*, 143, 71–96, 1996.

[204] D. L. Nguyen, M. V. Klibanov, L. H. Nguyen and M. A. Fiddy, Imaging of buried objects from multi-frequency experimental data using a globally convergent inversion method, *J. Inverse Ill-Posed Probl.*, 26, 501–522, 2018.

[205] D. L. Nguyen, M. V. Klibanov, L. H. Nguyen, A. E. Kolesov, M. A. Fiddy and H. Liu, Numerical solution of a coefficient inverse problem with multi-frequency experimental raw data by a globally convergent algorithm, *J. Comput. Phys.*, 345, 17–32, 2017.

[206] L. Nguyen, M. Ressler and J. Sichina, Sensing through the wall imaging using the Army Research Lab ultra-wideband synchronous impulse reconstruction (UWB SIRE) radar, in: *Proc. SPIE*, Vol. 6047, 6947/0B, 2008.

[207] L. H. Nguyen, An inverse space-dependent source problem for hyperbolic equations and the Lipschitz-like convergence of the quasi-reversibility method, *Inverse Probl.*, 35, 0037, 2019.

[208] L. H. Nguyen, A new algorithm to determine the creation or depletion term of parabolic equations from boundary measurements, *Comput. Math. Appl.*, 80, 2135–2149, 2020.

[209] L. H. Nguyen, Q. Li and M. V. Klibanov, A convergent numerical method for a multi-frequency inverse source problem in inhomogeneous media, *Inverse Probl. Imaging*, 13, 1067–1094, 2019.

[210] P. M. Nguyen and L. H. Nguyen, A numerical method for an inverse source problem for parabolic equations and its application to a coefficient inverse problem, *J. Inverse Ill-Posed Probl.*, published online https://doi.org/10.1515/jiip-2019-0026, 2020.

[211] R. G. Novikov, A multidimensional inverse spectral problem for the equation $-\Delta\psi + (v(x) - Eu(x))\psi = 0$, *Funct. Anal. Appl.*, 22, 263–272, 1988.

[212] R. G. Novikov, The inverse scattering problem on a fixed energy level for the two-dimensional Schrödinger operator, *J. Funct. Anal.*, 103, 409–463, 1992.

[213] R. G. Novikov, The ∂-bar approach to approximate inverse scattering at fixed energy in three dimensions, *Int. Math. Res. Reports*, 6, 287–349, 2005.

[214] R. G. Novikov, Formulas for phase recovering from phaseless scattering data at fixed frequency, *Bull. Sci. Math.*, 139, 923–936, 2015.

[215] R. G. Novikov, An iterative approach to non-overdetermined inverse scattering at fixed energy, *Sb. Math.*, 206, 120–134, 2015.

[216] R. G. Novikov, Explicit formulas and global uniqueness for phaseless inverse scattering in multidimensions, *J. Geom. Anal.*, 26, 346–359, 2016.

[217] Office of the Chief of Ordnance 1945 Catalog of enemy ordnance material World War II Operational Documents, N2228-E, OCLC 464601649, http://cgsc.contentdm.oclc.org/cdm/ref/collection/p4013coll8/id/2758.

[218] N. Pantong, J. Su, H. Shan, M. V. Klibanov and H. Liu, A globally accelerated reconstruction algorithm for diffusion tomography with continuous-wave source in arbitrary convex shape domain, *J. Opt. Soc. Am. A*, 26, 456–472, 2009.

[219] L. Pestov and G. Uhlmann, Two dimensional simple Riemannian manifolds are boundary distance rigid, *Ann. Math.*, 161, 1093–1110, 2005.
[220] O. Poisson, Carleman estimates for the heat equation with discontinuous diffusion coefficients, *Appl. Anal.*, 87, 1129–1144, 2008.
[221] K. G. Reznickaya, Connection between solutions of different types of Cauchy problems and inverse problems, in: *Mathematical Problems of Geophysics*, issue 5, part 1, Published by Computing Center of the Siberian Branch of the USSR Academy of Science, Novosibirsk, pages 55–62, 1974 (in Russian).
[222] G. Rizzuti and A. Gisolf, An iterative method for 2D inverse scattering problems by alternating reconstruction of medium properties and wavefields: theory and application to the inversion of elastic waveforms, *Inverse Probl.*, 33, 035003, 2017.
[223] V. G. Romanov, *Integral Geometry and Inverse Problems for Hyperbolic Equations*, Springer-Verlag, Berlin, 1974.
[224] V. G. Romanov, *Inverse Problems of Mathematical Physics*, VNU Science Press, Utrecht, 1987.
[225] V. G. Romanov, Inverse problems for differential equations with memory, *Eurasian J. Math. Comput. Appl.*, 2(4), 51–80, 2014.
[226] V. G. Romanov, Problem of determining the permittivity in the stationary system of Maxwell equations, *Dokl. Math.*, 95, 1–5, 2017.
[227] V. G. Romanov and M. Yamamoto, Recovering a Lamé kernel in a viscoelastic equation by a single boundary measurement, *Appl. Anal.*, 89, 377–390, 2010.
[228] A. A. Samarskii, *The Theory of Difference Schemes*, CRC Press, 2001.
[229] J. A. Scales, M. L. Smith and T. L. Fischer, Global optimization methods for multimodal inverse problems, *J. Comput. Phys.*, 103, 258–268, 1992.
[230] U. Schrőder and T. Schuster, An iterative method to reconstruct the refractive index of a medium from time-off-light measurements, *Inverse Probl.*, 32, 085009, 2016.
[231] J. K. Seo, J. Lee, S. W. Kim, H. Zribi and E. J. Woo, Frequency-difference electrical impedance tomography (fdEIT): algorithm development and feasibility study, *Physiol. Meas.*, 29, 929–944, 2008.
[232] J. K. Seo and E. J. Woo, Magnetic resonance electrical impedance tomography (MREIT), *SIAM Rev.*, 53, 40–68, 2011.
[233] J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, Cambridge Monogr. Appl. Comput. Math., Cambridge University Press, Cambridge, 1999.
[234] H. Shan, M. V. Klibanov, J. Su, N. Pantong and H. Liu, A globally accelerated numerical method for optical tomography with continuous wave source, *J. Inverse Ill-Posed Probl.*, 16, 765–792, 2008.
[235] S. Siltanen, J. Mueller and D. Isaacson, An implementation of the reconstruction algorithm of A. Nachman for the 2-D inverse conductivity problem, *Inverse Probl.*, 16, 681–699, 2000.
[236] A. V. Smirnov, M. V. Klibanov and L. H. Nguyen, On and inverse source problem for the full radiative transfer equation with incomplete data, *SIAM J. Sci. Comput.*, 41, B929–B952, 2019.
[237] A. V. Smirnov, M. V. Klibanov and L. H. Nguyen, Convexification for a 1D hyperbolic coefficient inverse problem with single measurement data, *Inverse Probl. Imaging*, 14, 913–938, 2020.
[238] A. V. Smirnov, M. V. Klibanov, A. J. Sullivan and L. H. Nguyen, Convexification for an inverse problem for a 1D wave equation with experimental data, *Inverse Probl.*, 36, 095008, 2020.
[239] P. Stefanov, G. Uhlmann and A. Vasy, Local and global boundary rigidity and the geodesic X-ray transform in the normal gauge, arXiv:1702.03638v2, 2017.
[240] J. Su, M. V. Klibanov, Y. Liu, Z. Lin, N. Pantong and H. Liu, Optical imaging of phantoms from real data by an approximately globally convergent inverse algorithm, *Inverse Probl. Sci. Eng.*, 21, 1125–1150, 2013.

[241] J. Sylvester and G. Uhlmann, Global uniqueness theorem for an inverse boundary problem, *Ann. Math.*, 39, 91–112, 1987.

[242] N. T. Thành, L. Beilina, M. V. Klibanov and M. A. Fiddy, Imaging of buried objects from experimental backscattering time dependent measurements using a globally convergent inverse algorithm, *SIAM J. Imaging Sci.*, 8, 757–786, 2015.

[243] N. T. Thánh, L. Beilina, M. V. Klibanov and M. A. Fiddy, Reconstruction of the refractive index from experimental backscattering data using a globally convergent inverse method, *SIAM J. Sci. Comput.*, 36, B273–B293, 2014.

[244] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer, London, 1995.

[245] R. Triggiani and P. F. Yao, Carleman estimates with no lower order terms for general Riemannian wave equations. Global uniqueness and observability in one shot, *Appl. Math. Optim.*, 46, 331–375, 2002.

[246] G. Vainikko, Fast solvers of the Lippmann–Schwinger equation, in: *Direct and Inverse Problems of Mathematical Physics*, Springer US, 2000, pages 423–440.

[247] M. M. Vajnberg, *Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations*, John Wiley & Sons, Washington, DC, 1973.

[248] F. P. Vasiliev, *Numerical Methods of Solutions of Extremal Problems*, Nauka, Moscow, 1989 (in Russian).

[249] L. Volgyesi and M. Moser, The inner structure of the Earth, *Period. Polytech., Chem. Eng.*, 26, 155–204, 1982.

[250] E. Wiechert and K. Zoeppritz, Uber Erdbebenwellen, *Nachr. Koenigl. Geselschaft Wiss. Gottingen*, 4, 415–549, 1907.

[251] M. Yamamoto, On an inverse problem of determining source terms in Maxwell's equations with a single measurement, in: *Inverse Problems, Tomography, and Image Processing*, A. Ramm, ed., Plenum Press, New York, pages 241–256, 1998.

[252] M. Yamamoto, Carleman estimates for parabolic equations. Topical Review, *Inverse Probl.*, 25, 123013, 2009.

[253] M. Yamamoto and J. Zou, Simultaneous reconstruction of the initial temperature and heat radiative coefficient, *Inverse Probl.*, 17, 1181–1202, 2001.

[254] G. Yuan and M. Yamamoto, Lipschitz stability in inverse problems for a Kirchhoff plate equation, *Asymptot. Anal.*, 53, 29–60, 2007.

[255] G. Yuan and M. Yamamoto, Carleman estimates for the Schrödinger equation and applications to an inverse problem and an observability inequality, *Chin. Ann. Math., Ser. B*, 31, 555–578, 2010.

[256] B. N. Zakhariev and A. A. Suzko, *Potentials in Quantum Scattering. Direct and Inverse Problems*, Energoatomizdat, Moskow, 1985.

[257] H. Zhao and Y. Zhong, A hybrid adaptive phase space method for reflection traveltime tomography, *SIAM J. Imaging Sci.*, 12, 28–53, 2019.

[258] https://www.translatorscafe.com/unit-converter/en-US/frequency-wavelength/5-29/gigahertz-wavelength%20in%20centimetres/.

Index

Inverse and Ill-Posed Problems Series

Volume 62
Vitalii P. Tanana, Anna I. Sidikova
Optimal Methods for Ill-Posed Problems. With Applications to Heat Conduction
ISBN 978-3-11-057573-6, e-ISBN 978-3-11-057721-1, e-ISBN (ePUB) 978-3-11-057583-5

Volume 61
Anatoly B. Bakushinsky, Mikhail M. Kokurin, Mikhail Yu. Kokurin
Regularization Algorithms for Ill-Posed Problems
ISBN 978-3-11-055630-8, e-ISBN 978-3-11-055735-0

Volume 60
Jin Cheng, Masahiro Yamamoto
Complex Methods for Inverse Problems
ISBN 978-3-11-040241-4, e-ISBN 978-3-11-040669-6

Volume 59
Alexander Andreevych Boichuk, Anatolii M. Samoilenko
Generalized Inverse Operators. And Fredholm Boundary-Value Problems
ISBN 978-3-11-037839-9, e-ISBN 978-3-11-037844-3

Volume 58
Shuai Lu, Sergei V. Pereverzev
Regularization Theory for Ill-posed Problems
ISBN 978-3-11-028646-5, e-ISBN 978-3-11-028649-6

Volume 56
Yanfei Wang, Anatoly G. Yagola, Changchun Yang
Computational Methods for Applied Inverse Problems
ISBN 978-3-11-025904-9, e-ISBN 978-3-11-025905-6

Volume 55
Sergey I. Kabanikhin
Inverse and Ill-posed Problems. Theory and Applications
ISBN 978-3-11-022400-9, e-ISBN 978-3-11-022401-6

Volume 54
Anatoly B. Bakushinsky, Mihail Yu. Kokurin, Alexandra Smirnova
Iterative Methods for Ill-Posed Problems. An Introduction
ISBN 978-3-11-025064-0, e-ISBN 978-3-11-025065-7

www.degruyter.com