Identification of a Temperature Dependent Heat Conductivity by Tikhonov Regularization

Philipp Kügler [1] and **Heinz W. Engl** [2]

Abstract

If one applies the standard theory of Tikhonov regularization to the nonlinear ill-posed problem of identifying a temperature-dependent heat conductivity from temperature measurements, one needs rather strong conditions for obtaining convergence rates. By using a problem-adapted adjoint, we can obtain convergence rates under substantially weaker and more realistic conditions.

1 Introduction

In many applications modeled by the heat equation, for example in the context of steel production (see [10], [11], [12]), the heat conductivity q does not vary spatially but rather depends on the temperature u itself, which makes the problem nonlinear. We consider the stationary case, where the heat distribution is described by the nonlinear elliptic equation

$$-\nabla \cdot (q(u)\nabla u) = f(x) \text{ in } \Omega \tag{1.1}$$

with, e.g., the boundary condition

$$u(x) = 0 \text{ on } \partial\Omega. \tag{1.2}$$

Here, Ω is an open bounded connected domain in $\mathbb{R}^n$, $n \geq 1$, with a piecewise smooth boundary $\partial\Omega$, and f is a given heat source density.

Measuring a temperature-dependent heat conductivity can be difficult, sometimes (e.g. at high temperatures) impossible. In such cases, one has to resort to using indirect measurements, i.e., to identifying the conductivity from temperature or heat flux measurements. Such parameter identification problems are usually ill-posed, i.e. small perturbations in the observed data can strongly influence the determination of the parameter in a negative way. Hence, for their stable numerical solution some type of regularization is required. See [1] for a general background on inverse problems and regularization. The probably most frequently used approach is Tikhonov regularization, which we use in this paper.

[1]Institut für Industriemathematik, Johannes Kepler Universität, A–4040 Linz, Austria. E-Mail: kuegler@indmath.uni-linz.ac.at. Supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung, projects F013-08 and P13478-INF

[2]Institut für Industriemathematik, Johannes Kepler Universität, A–4040 Linz, Austria. E-Mail: engl@indmath.uni-linz.ac.at.

Numerical experience shows that Tikhonov regularization performs very well for various parameter identification problems. In order to provide some useful guidance under which conditions Tikhonov regularization will actually perform well and to give practitioners confidence in their numerical results, a mathematical analysis of the numerical behaviour achieved by Tikhonov regularization is necessary. Note that for ill-posed problems, convergence of any numerical algorithm can be arbitrarily slow. Hence, conditions for convergence rates are not only of theoretical interest but also of practical relevance as they tell engineers for which problems fast convergence of numerical algorithms can be expected.

Starting from [13], [14] and [15], there exists a large literature on the convergence theory for Tikhonov regularization for nonlinear ill-posed problems. Below we briefly summarize the existing general theory and indicate in which respect it has to be modified in order to yield good results for our problem.

Consider a nonlinear operator $F : K \subset Q \to U$ acting between two Hilbert-spaces Q and U - K is some admissible parameter set - and satisfying

- F is continuous
- F is weakly closed, i.e. $q_n \rightharpoonup q \wedge F(q_n) \rightharpoonup u \Rightarrow F(q) = u$.

We are interested in the case where the inversion (in some generalized sense, if necessary) of F is unstable, i.e., the equation $F(q) = u$ is ill-posed. For a fixed right-hand side $u \in U$, we want to find the q^*-minimum-norm solution $q^\dagger$ of $F(q) = u$ defined by

$$F(q^\dagger) = u$$

and

$$\|q^\dagger - q^*\|_Q = \min \{\|q - q^*\|_Q \mid F(q) = u\}.$$

We assume that there exists at least one q^*-minimum-norm solution, i.e., we suppose the data u to be attainable. In the case of more than one solution, q^* is some kind of selection criterion. Its choice should depend on a-priori information on the solution of $F(q) = u$, if available.

In the situation that instead of the exact data u only an approximation u^δ is given with

$$\|u - u^\delta\|_U \leq \delta,$$

we are looking for an approximation, say q_β^δ, of $q^\dagger$. Because of the ill-posedness, q_β^δ should be the solution of a well-posed auxiliary problem - the regularized problem - i.e. q_β^δ should

depend continuously on the perturbed data u^δ, so that it can be computed in a stable way. On the other hand we want q_β^δ to tend to $q^\dagger$ as δ decreases to zero and the regularization parameter β is chosen in an appropriate way.

In the context of Tikhonov regularization, q_β^δ is determined as the minimizer of the Tikhonov functional

$$\|F(q) - u^\delta\|_U^2 + \beta\|q - q^*\|_Q^2, \tag{1.3}$$

where $\beta > 0$. The assumptions made on F guarantee the existence of a minimizer, which of course will not be unique in general. Furthermore, we get from these assumptions

- stability for a fixed $\beta > 0$:
  For the sequences $\{u_k\}$ and $\{q_k\}$, where $u_k \to u^\delta$ and q_k is a minimizer of (1.3) with u^δ replaced by u_k, there exists a convergent subsequence of $\{q_k\}$, and every limit is a minimizer of (1.3).

- convergence:
  For $\beta(\delta) \to 0$ with $\delta^2/\beta(\delta) \to 0$ as $\delta \to 0$, every sequence $\{q_{\beta_k}^{\delta_k}\}$, where $\delta_k \to 0$, $\beta_k = \beta(\delta_k)$ and $\{q_{\beta_k}^{\delta_k}\}$ is a solution of (1.3), has a convergent subsequence, and every limit is a q^*-minimum-norm solution. If the latter is unique, the whole sequence converges.

Talking about the rate of convergence of a regularization method, we have that rate in mind at which

$$\|q_\beta^\delta - q^\dagger\|_Q \to 0 \ \text{ as } \ \delta \to 0.$$

In general the convergence rate for ill-posed problems is arbitrarily slow (see [9]). Convergence rates can only be given on a subset of K, i.e., under a-priori assumptions on the exact solution $q^\dagger$, or equivalently, on the exact data u. This a priori-assumption is usually formulated in terms of a "source condition": According to [1], a convergence rate

$$\|q_\beta^\delta - q^\dagger\|_Q = O(\sqrt{\delta})$$

of regularized solutions is obtained, for the choice $\beta \sim \delta$, under the additional assumptions that

- F is Fréchet-differentiable,
- there exists $\gamma \geq 0$ such that $\|F'(q^\dagger) - F'(q)\|_U \leq \gamma\|q^\dagger - q\|_Q$,
- there exists $w \in U$ satisfying the *source-condition*

$$q^\dagger - q^* = F'(q^\dagger)^* w, \tag{1.4}$$

- and the smallness condition

$$\gamma \|w\|_U < 1. \tag{1.5}$$

Here, $F'(q^\dagger)^*$ denotes the adjoint of the Fréchet-derivative.

This general theory has been applied to various inverse problems including parameter identification, see [16], [1] for elliptic problems and integral equations, and [17] for a parabolic equation. All these applications are for one-dimensional problems, since only there, the source condition (1.4) has a rather immediate explicit interpretation (usually requiring some additional smoothness and prescribed boundary behaviour for $q^\dagger - q^*$). For the nonlinear heat conduction problem discussed in this paper, as we assume the parameter only to be of H^1-regularity, the conditions of the general convergence theory, e.g. the Fréchet-differentiability and the Lipschitz-continuity of F', are not satisfied. Even suppose the general theory applies, an interpretation of the source condition (1.4) would be hard to find. The main problem is that there is no clear understanding of the derivative $F'(q)$ of the parameter-to-solution map and its adjoint $F'(q)^*$, since they often have no physical meaning, unlike $F(q)$ itself. And, finally, the smallness condition (1.5) on the source function is quite restrictive.

In this paper, we circumvent the restrictive requirements of the general convergence theory, i.e. the Fréchet differentiability of $F(q)$ and the uniform Lipschitz continuity of the Fréchet derivative $F'(q)$. We use a simpler source condition which is intimately related to the weak form of (1.1) and can be nicely interpreted. And, finally, we also get rid of the smallness condition. Altogether, we preserve the usual convergence rate $O(\sqrt{\delta})$ under much weaker and more realistic assumptions. Our new approach is based on the recent ideas of [2] for identifying a heat conductivity distribution $q = q(x)$, depending on the spatial variable. In [2], the parabolic case was considered, while we discuss the stationary case here. Our results can easily be reformulated for the parabolic case (see [3]).

2 The Direct Problem

Before formulating the inverse parameter identification problem we briefly discuss the underlying nonlinear direct problem (1.1) - (1.2). We set

$$V = H_0^1(\Omega) \text{ and } H = L^2(\Omega),$$

and denote the corresponding norms by $\|\cdot\|_V$ and $\|\cdot\|_H$.

Except in Section 5, we assume throughout the paper that

$$f \in V^*,$$

where V^* is the dual space to V. Then, by integration by parts we can immediately derive the variational formulation for (1.1)-(1.2):

Find $u \in V$ such that

$$\int_\Omega q(u)\nabla u \cdot \nabla v \, dx = \int_\Omega f(x)v \, dx \quad \text{for all} \quad v \in V \tag{2.1}$$

holds. Under the assumption

$$q \in H^1(\mathbb{R}) \ \text{ and } \ 0 < q_{\min} \leq q \leq q_{\max},$$

there exists a unique solution to the variational problem (2.1) in V. The existence proof - based on the continuous embedding $H^1(\mathbb{R}) \subseteq C(\mathbb{R})$ and the theory of quasi-monotone operators - can for example be found in [8] (SectionII.6). The uniqueness of the solution follows from an argument similar to that in [3] for the parabolic problem. Furthermore, there is an a-priori estimate for the solution, i.e. there is a constant $C > 0$ depending only on $q_{\max}$ and f such that

$$\|u_q\|_V \leq C,$$

where, in order to emphasize the fact that the solution u depends on the parameter q, we use the notation u_q, or $u_q(x)$. We quote [3] for the standard derivation of this result.

Already with respect to Section 5 we cite [5] and [7] for the classical analysis of (1.1)-(1.2): If we require more regularity for f, namely

$$f \in C^\alpha(\bar{\Omega}) \tag{2.2}$$

for an $\alpha \in (0,1)$ and for q, namely

$$q \in C^{1,\alpha}(\mathbb{R}) \ \text{ and } \ 0 < q_{\min} \leq q \leq q_{\max},$$

the unique classical solution is twice Hölder continuously differentiable:

$$u_q \in C^{2,\alpha}(\bar{\Omega}). \tag{2.3}$$

See Theorem 12.8. of [5], p. 276, for the existence result and Theorem 31 of [7], p. 151, for the uniqueness argument based on a maximum principle.

3 Output Least Squares Formulation with Tikhonov Regularization

In this section we formulate the inverse problem for the elliptic problem discussed above. Given a distributed observation of the solution of the direct problem, the aim is to recover

the physical parameter q on an appropriate interval using the observation data. In order to overcome the ill-posedness of this identification problem, we choose Tikhonov regularization for its stabilization.

By z(x) we denote the measurement of the solution u over Ω, i.e.

$$z(x) = u_{\text{measured}}(x)$$

and we want to identify the true thermal conductivity $q^\dagger$ satisfying

$$u_{q^\dagger}(x) = z(x) \ \text{ on } \ \Omega. \tag{3.1}$$

We assume the exact data to have the same regularity as the solution u, i.e. $z \in V$. Note that $q^\dagger$ may not exist, and even if it exists it may not be unique. However, we will always assume the existence of a true parameter $q^\dagger$, i.e., that the exact data z are attainable. Of course, the measured (noisy) data need not be attainable.

We carry out the recovery process by using the output least squares method: The output u_q corresponding to a parameter q out of a set K of admissible parameters should match z in the L^2-norm, i.e., we want to solve the optimization problem

$$J(q) = \int_\Omega |u_q - z|^2 dx \to \min_{q \in K}.$$

Note that the thermal conductivity is a function of a one dimensional variable τ, which varies in the range of possible temperatures. We cannot expect $q^\dagger$ to be identifiable on that whole range, but only on the range of $u_{q^\dagger}$, i.e., the set where the data are available. Unless we have a priori information on this range, the best we can do is to construct a finite interval - "containing mainly" the range of $u_{q^\dagger}$ - using the given data z. An additional difficulty is that if we approximate $u_{q^\dagger}$, the range of that approximation will be different. We return to this very crucial point when discussing a numerical example in Section 7.

For our theoretical investigations it is sufficient to choose a real interval

$$I = [U_{\min}, U_{\max}], \ \ U_{\min}, U_{\max} \in \mathbb{R} \tag{3.2}$$

as domain of the parameter, such that

$$U_{\min} \le u_q(x) \le U_{\max} \tag{3.3}$$

holds for the solution u_q of the direct problem for any parameter $q \in H^1(\mathbb{R})$ with $q_{\min} \le q \le q_{\max}$. Note that the assumption of the existence of such an interval I is not very restrictive, as we can choose $U_{\min}$, $U_{\max}$ arbitrarily small or large, respectively. Of course, we can only identify q on a subdomain where we have information about the system from the data z and from the solutions of the direct problem calculated during the

recovery process. Outside this domain, we have no information, so that an identification is impossible. In this sense we should look in (3.2) for an interval I of minimal length. Again, we refer to Section 7.

We define the set of admissible parameters as

$$K = \left\{ q \in H^1(I) \mid q_{\min} \leq q(\tau) \leq q_{\max} \text{ for } \tau \in I \right\}, \tag{3.4}$$

where $q_{\min}$ and $q_{\max}$ are fixed positive constants. As the embedding of $H^1(I)$ into $C(I)$ is compact, any $q \in K$ is continuous.

For later use, we introduce the set of the indefinite integrals of the parameters $q \in K$

$$S = \left\{ p \in H^2(I) \mid \frac{dp}{d\tau} \in K \text{ and } p(0) = 0 \right\}. \tag{3.5}$$

Because of (3.4), we have a common Lipschitz constant, namely $q_{\max}$, for the functions $p \in S$:

$$|p(\tau_1) - p(\tau_2)| \leq q_{\max} |\tau_1 - \tau_2|, \ \tau_1, \ \tau_2 \in I \tag{3.6}$$

for all $p \in S$.

So far, we only considered the output-least-squares term. As the inverse problem is ill-posed, we have to add some type of regularization, i.e. stabilization, into account. If we choose the $H^1(I)$-norm as regularizing term, which is appropriate since we assume $q \in H^1(I)$, we obtain the following regularized output-least-squares problem:

$$J_\beta(q) = \int_\Omega |u_q - z|^2 dx + \beta \|q - q^*\|_I^2 \to \min_{q \in K}$$

with

$$\|q - q^*\|_I^2 = \int_I (q(\tau) - q^*(\tau))^2 \, d\tau + \int_I \left(\frac{dq}{d\tau}(\tau) - \frac{dq^*}{d\tau}(\tau) \right)^2 d\tau.$$

Later on, it will turn out that this output-least-squares problem is really stabilized by the penalty term $\beta \|q - q^*\|_I^2$. We mention that q^* is not required to be an element of the constraint set K. The selection of q^* is crucial for the results about the convergence rate in Section 4. Available a-priori information about the true parameter $q^\dagger$ should be used for the choice of q^*, i.e., q^* should be interpreted as some kind of a-priori guess for $q^\dagger$. In the case of multiple solutions, q^* plays the role of a selection criterion, the choice of q^* influences which of the possible parameters is approximated. Recalling the lines subsequent to (3.3), q^* also determines the "identified" parameter $q^\dagger$ outside the domain of information in the case of I chosen too large.

In applications, the exact data $z(x)$ is not known precisely due to measurement errors. Hence, the actual data are available in the form

$$z^\delta(x) = z(x) + \text{noise},$$

where one needs some information

$$\|z - z^\delta\|_H \leq \delta \tag{3.7}$$

about the noise level.

Summarizing, we formulate the regularized problem, which will be the issue of our theoretical analysis:

Let the set K of admissible parameters and noisy data z^δ be given as in (3.4) and (3.7). Assume that the exact data z is attainable from a parameter $q^\dagger \in K$. Then, for $\beta > 0$, find a parameter $q_\beta^\delta \in K$ that minimizes

$$J_\beta(q) = \int_\Omega |u_q - z^\delta|^2 dx + \beta \|q - q^*\|_I^2 \tag{3.8}$$

over K for an appropriate choice of β and q^*.

We first quote a result from [3] about the existence of solutions to the minimization problem (3.8):

Theorem 3.1. *There exists at least one minimizer q_β^δ to the optimization problem* (3.8).

Next, we present an auxiliary result, once again from [3], which will take over the part of the weak closedness used in the general convergence theory:

Theorem 3.2. *For any sequence $\{q_n\}$ in K which converges weakly to some $q \in K$ in $H^1(I)$ as $n \to \infty$ we have, for a subsequence $\{q_{n_k}\}$,*

$$u_{q_{n_k}} \rightharpoonup u_q \quad in \ V$$

and

$$\lim_{k\to\infty} \int_\Omega |(u_{q_{n_k}} - z)|^2 dx = \int_\Omega |(u_q - z)|^2 dx.$$

We remark that though all proofs in [3] are given for the corresponding parabolic problem, they can be easily carried over to the elliptic case. The proofs are similar to those in [4] for the linear case. Furthermore, in [3] an additional weight function $q(u_q)$ is used in the output least squares term (like in the linear case as discussed in [2]), which

we omit for the sake of simplicity in this paper.

After these preliminary considerations, we now turn to the analysis of the regularized parameter identification problem. First, we consider its stability, where by stability we mean that - for a fixed regularization parameter β - the minimizers of Problem (3.8) depend continuously on the data z^δ. Since these minimizers need not be unique, stability is to be understood in the following multivalued sense:

Theorem 3.3. *For any $\beta > 0$ let $\{z_n\}$ be a sequence such that $z_n \to z^\delta$ in H, and $\{q_n\}$ be minimizers of (3.8) with z^δ replaced by z_n. Then there exists a subsequence of $\{q_n\}$ that converges in $H^1(I)$, and the limit of every such convergent subsequence is a minimizer of*

$$J_\beta(q) = \int_\Omega |u_q - z^\delta|^2 dx + \beta \|q - q^*\|_I^2. \tag{3.9}$$

Proof. By definition of $\{q_n\}$ we have

$$\begin{aligned} &\int_\Omega |u_{q_n} - z_n|^2 dx + \beta \|q_n - q^*\|_I^2 \\ &\leq \int_\Omega |u_q - z_n|^2 dx + \beta \|q - q^*\|_I^2 \end{aligned} \tag{3.10}$$

for any $q \in K$. Hence, $\{\|q_n\|_I\}$ and $\{\|u_{q_n}\|_H\}$ are bounded. Therefore a subsequence (again denoted by $\{q_n\}$) and $\bar{q}$ exist such that $q_n \to \bar{q}$ in $L^2(I)$ and

$$q_n \rightharpoonup \bar{q} \text{ in } H^1(I) \text{ and } u_{q_n} \rightharpoonup \bar{u} \text{ in } V$$

Now, Theorem 3.2 yields $\bar{u} = u_{\bar{q}}$ and

$$\lim_{n\to\infty} \int_\Omega |u_{q_n} - z_n|^2 dx = \int_\Omega |u_{\bar{q}} - z^\delta|^2 dx, \tag{3.11}$$

as $z_n \to z^\delta$. Using this, the weak lower semicontinuity of $\|\cdot\|_I$ and (3.10), we get

$$\begin{aligned} J_\beta(\bar{q}) &= \int_\Omega |u_{\bar{q}} - z^\delta|^2 dx + \beta \|\bar{q} - q^*\|_I^2 \\ &\leq \liminf_{n\to\infty} \left\{ \int_\Omega |u_{q_n} - z_n|^2 dx + \beta \|q_n - q^*\|_I^2 \right\} \\ &\leq \limsup_{n\to\infty} \left\{ \int_\Omega |u_{q_n} - z_n|^2 dx + \beta \|q_n - q^*\|_I^2 \right\} \\ &\leq \lim_{n\to\infty} \left\{ \int_\Omega |u_q - z_n|^2 dx + \beta \|q - q^*\|_I^2 \right\} \\ &= J_\beta(q) \end{aligned} \tag{3.12}$$

for all $q \in K$. Hence, $\bar{q}$ is a minimizer of (3.9).

Next, we show that $q_n \to \bar{q}$ in $H^1(I)$ by contradiction: If $q_n \not\to \bar{q}$, the weak lower semicontinuity of the norm gives

$$\mu := \limsup_{n\to\infty} \|q_n - q^*\|_I > \|\bar{q} - q^*\|_I \tag{3.13}$$

which implies that there exists a subsequence $\{q_m\}$, with $q_m \rightharpoonup \bar{q}$ in $H^1(I)$, such that

$$\lim_{m\to\infty} \|q_m - q^*\|_I = \mu$$

Choosing $q = \bar{q}$ in (3.12) we see that

$$\begin{aligned}\lim_{m\to\infty} \left\{ \int_\Omega |u_{q_m} - z_m|^2 dx + \beta \|q_m - q^*\|_I^2 \right\} \\ = \int_\Omega |u_{\bar{q}} - z^\delta|^2 dx + \beta \|\bar{q} - q^*\|_I^2\end{aligned}$$

Combining this with (3.13) we get

$$\begin{aligned}&\lim_{m\to\infty} \int_\Omega |u_{q_m} - z_m|^2 dx \\ = \;& \int_\Omega |u_{\bar{q}} - z^\delta|^2 dx + \beta \left(\|\bar{q} - q^*\|_I^2 - \mu^2 \right) \\ < \;& \int_\Omega |u_{\bar{q}} - z^\delta|^2 dx\end{aligned}$$

which is a contradiction to (3.11). Therefore $q_n \to \bar{q}$ in $H^1(I)$. □

After this stability result, we now investigate the convergence of the Tikhonov regularization method as the noise level tends to zero, with the regularization parameter chosen by an a priori rule. Again, convergence is to be understood in a multivalued sense:

Theorem 3.4. *For any positive sequence $\{\delta_k\}$, let $\beta_k := \beta(\delta_k)$ be such that $\beta_k \to 0$, $\delta_k^2/\beta_k \to 0$ as $k \to \infty$, and $z^{\delta_k} \in H$ satisfy (3.7). Then the sequence $\{q_{\beta_k}^{\delta_k}\}$ given by solutions of problem (3.8) with z^δ and β replaced by z^{δ_k} and β_k, has a subsequence which converges in $H^1(I)$. The limit of every convergent subsequence, say $\bar{q}$, is a q^*-minimum-norm-solution of (3.1), i.e., $u_{\bar{q}} = z$ and*

$$\|\bar{q} - q^*\|_I = \min_{q\in K} \{ \|q - q^*\|_I \mid u_q = z \}.$$

Proof. Let $q^\dagger$ be a q^*-minimum-norm-solution of (3.1), which we assumed to exist. As $q_{\beta_k}^{\delta_k}$ is a minimizer we get $J_{\beta_k}(q_{\beta_k}^{\delta_k}) \le J_{\beta_k}(q^\dagger)$, which implies

$$\|u_{q_{\beta_k}^{\delta_k}} - z^{\delta_k}\|_H^2 + \beta_k \|q_{\beta_k}^{\delta_k} - q^*\|_I^2 \le \delta_k^2 + \beta_k \|q^\dagger - q^*\|_I^2.$$

Therefore

$$\|u_{q_{\beta_k}^{\delta_k}} - z^{\delta_k}\|_H^2 \to 0 \ \text{ as } \ k \to \infty,$$

which by (3.7) and the triangle inequality leads to

$$u_{q_{\beta_k}^{\delta_k}} \to z \ \text{ in } \ H. \tag{3.14}$$

We also see that

$$\limsup_{k\to\infty} \|q_{\beta_k}^{\delta_k} - q^*\|_I \le \|q^\dagger - q^*\|_I. \tag{3.15}$$

Hence, $\{q_{\beta_k}^{\delta_k}\}$ is bounded in $H^1(I)$, so that there exists a subsequence, again denoted by $\{q_{\beta_k}^{\delta_k}\}$, such that

$$q_{\beta_k}^{\delta_k} \rightharpoonup \bar{q} \text{ in } H^1(I) \ \text{ and } \ q_{\beta_k}^{\delta_k} \to \bar{q} \text{ in } L^2(I) \ \text{ as } \ k \to \infty. \tag{3.16}$$

According to Theorem 3.2, we have

$$u_{q_{\beta_k}^{\delta_k}} \rightharpoonup u_{\bar{q}} \ \text{ in } \ V$$

which with (3.14) shows that $u_{\bar{q}} = z$, so that $\bar{q}$ is a minimizer of

$$\int_\Omega |u_q - z|^2 dx.$$

Now, by the weak lower semicontinuity of the norm, (3.15), (3.16) and the definition of $q^\dagger$ we have

$$\|q^\dagger - q^*\|_I \le \|\bar{q} - q^*\|_I \le \limsup_{k\to\infty} \|q_{\beta_k}^{\delta_k} - q^*\|_I \le \|q^\dagger - q^*\|_I,$$

which implies that $\bar{q}$ is a q^*-minimum-norm solution.

Next, we show that $q_{\beta_k}^{\delta_k} \to \bar{q}$ in $H^1(I)$. Similarly to (3.15) we derive, by using $J_{\beta_k}(q_{\beta_k}^{\delta_k}) \le J_{\beta_k}(\bar{q})$, that

$$\limsup_{k\to\infty} \|q_{\beta_k}^{\delta_k} - q^*\|_I \le \|\bar{q} - q^*\|_I.$$

This together with the identity

$$\|q_{\beta_k}^{\delta_k} - \bar{q}\|_I^2 = \|q_{\beta_k}^{\delta_k} - q^*\|_I^2 + \|q^* - \bar{q}\|_I^2 + 2\left(q_{\beta_k}^{\delta_k} - q^*, q^* - \bar{q}\right)_I$$

yields that

$$\limsup_{k\to\infty} \|q_{\beta_k}^{\delta_k} - \bar{q}\|_I^2 \le 2\|q^* - \bar{q}\|_I^2 + 2\limsup_{k\to\infty} \left(q_{\beta_k}^{\delta_k} - q^*, q^* - \bar{q}\right)_I = 0$$

and hence

$$\lim_{k\to\infty} q_{\beta_k}^{\delta_k} = \bar{q},$$

which completes the proof. □

4 Convergence Rates

Since in general the convergence rate of the regularized solutions to the exact solution may be arbitrarily slow, it is no surprise that additional assumptions are needed for a convergence rate. They have the form of a regularity assumption on the difference between the true parameter $q^\dagger$ and the a-priori guess q^*. The following theorem together with its interpretation in Section 5 is the main result of this paper:

Theorem 4.1. *Assume that there exists a function*

$$w \in V \tag{4.1}$$

such that

$$\left(q^\dagger - q^*, \psi\right)_I = -\int_\Omega \psi(u_{q^\dagger}) \nabla u_{q^\dagger} \nabla w dx \quad \forall \psi \in H^1(I) \tag{4.2}$$

holds. Furthermore, assume that

$$\Delta w \in H. \tag{4.3}$$

Then, with $\beta \sim \delta$, we have

$$\|u_{q_\beta^\delta} - z\|_H = O(\delta)$$

and

$$\|q_\beta^\delta - q^\dagger\|_I = O(\sqrt{\delta}).$$

Proof. As q_β^δ is a minimizer of (3.8), we get $J_\beta(q_\beta^\delta) \leq J_\beta(q^\dagger)$. This implies

$$\|u_{q_\beta^\delta} - z^\delta\|_H^2 + \beta \|q_\beta^\delta - q^*\|_I^2 \leq \delta^2 + \beta \|q^\dagger - q^*\|_I^2,$$

from which we obtain

$$\begin{aligned} & \|u_{q_\beta^\delta} - z^\delta\|_H^2 + \beta \|q^+ - q_\beta^\delta\|_I^2 \\ \leq \; & \delta^2 + \beta \|q^\dagger - q^*\|_I^2 + \beta \left\{ \|q^\dagger - q_\beta^\delta\|_I^2 - \|q_\beta^\delta - q^*\|_I^2 \right\} \\ = \; & \delta^2 + 2\beta \left(q^\dagger - q^*, q^\dagger - q_\beta^\delta\right)_I . \end{aligned} \tag{4.4}$$

Choosing $\psi = q^\dagger - q_\beta^\delta$ in the source condition (4.2) leads to

$$\left(q^\dagger - q^*, q^\dagger - q_\beta^\delta\right)_I = -\int_\Omega \left(q^\dagger(u_{q^\dagger}) - q_\beta^\delta(u_{q^\dagger})\right) \nabla u_{q^\dagger} \nabla w dx. \tag{4.5}$$

Using the direct problem formulation (see (2.1)) for $u_{q_\beta^\delta}$ and $u_{q^\dagger}$, we see by taking the difference that

$$\int_\Omega q_\beta^\delta(u_{q_\beta^\delta})\nabla u_{q_\beta^\delta}\cdot\nabla w dx - \int_\Omega q^\dagger(u_{q^\dagger})\nabla u_{q^\dagger}\cdot\nabla w dx = 0 \tag{4.6}$$

holds. Multiplying (4.5) by β and adding zero in the form of (4.6), it follows that

$$\begin{aligned}\beta\left(q^\dagger - q^*, q^\dagger - q_\beta^\delta\right)_I = \quad & \beta\int_\Omega \left(q_\beta^\delta(u_{q^\dagger}) - q^\dagger(u_{q^\dagger})\right)\nabla u_{q^\dagger}\nabla w dx \\ - \quad & \beta\int_\Omega q_\beta^\delta(u_{q_\beta^\delta})\nabla u_{q_\beta^\delta}\cdot\nabla w dx \\ + \quad & \beta\int_\Omega q^\dagger(u_{q^\dagger})\nabla u_{q^\dagger}\cdot\nabla w dx. \end{aligned}\tag{4.7}$$

Denoting the right hand side in (4.7) by I_2, we see that

$$I_2 = \beta\int_\Omega \left(q_\beta^\delta(u_{q^\dagger})\nabla u_{q^\dagger} - q_\beta^\delta(u_{q_\beta^\delta})\nabla u_{q_\beta^\delta}\right)\cdot\nabla w dx.$$

Using the antiderivative $p_\beta^\delta \in S$ (see (3.5)) of q_β^δ, we obtain

$$I_2 = \beta\int_\Omega \left(\nabla p_\beta^\delta(u_{q^+}) - \nabla p_\beta^\delta(u_{q_\beta^\delta})\right)\cdot\nabla w dx.$$

As $p_\beta^\delta(0) = 0$, integration by parts leads to

$$I_2 = -\beta\int_\Omega \left(p_\beta^\delta(u_{q^+}) - p_\beta^\delta(u_{q_\beta^\delta})\right)\Delta w dx.$$

As $q_{\max}$ is a common Lipschitz-constant for all p ∈ S (see (3.6)) we get

$$\begin{aligned}|I_2| &\le \beta\|p_\beta^\delta(u_{q^\dagger}) - p_\beta^\delta(u_{q_\beta^\delta})\|_H\|\Delta w\|_H \\ &\le q_{\max}\beta\|u_{q^\dagger} - u_{q_\beta^\delta}\|_H\|\Delta w\|_H.\end{aligned}$$

Applying the triangle inequality and Young's inequality

$$a\cdot b \le \varepsilon a^2 + \frac{b^2}{4\varepsilon},$$

for any $\varepsilon > 0$, we obtain

$$\begin{aligned}|I_2| &\le q_{\max}\beta\|u_{q_\beta^\delta} - z^\delta\|_H\|\Delta w\|_H \\ &\quad q_{\max}\beta\|z^\delta - u_{q^\dagger}\|_H\|\Delta w\|_H \\ &\le \varepsilon q_{\max}^2\delta^2 + \frac{\beta^2}{2\varepsilon}\|\Delta w\|_H^2 \\ &\quad +\varepsilon q_{\max}^2\|u_{q_\beta^\delta} - z^\delta\|_H^2.\end{aligned}$$

Using this estimate for $|I_2|$, we get from (4.4) that

$$\begin{aligned}\|u_{q_\beta^\delta} - z^\delta\|_H^2 + \beta\|q_\beta^\delta - q^+\|_I^2 \leq \\ \leq \delta^2 + 2\varepsilon q_{\max}^2\|u_{q_\beta^\delta} - z^\delta\|_H^2 + \\ + 2\varepsilon q_{\max}^2\delta^2 + \frac{\beta^2}{\varepsilon}\|\Delta w\|_H^2,\end{aligned}$$

which is equivalent to

$$\begin{aligned}\left\{1 - 2\varepsilon q_{\max}^2\right\}\|u_{q_\beta^\delta} - z^\delta\|_H^2 + \\ + \beta\|q_\beta^\delta - q^\dagger\|_I^2 \leq \delta^2 + 2\varepsilon q_{\max}^2\delta^2 + \\ + \frac{\beta^2}{\varepsilon}\|\Delta w\|_H^2.\end{aligned}$$

With $\varepsilon < \frac{1}{2q_{\max}^2}$ we finally obtain that

$$\|u_{q_\beta^\delta} - z^\delta\|_H = O(\delta) \tag{4.8}$$

(hence, by (3.7), also $\|u_{q_\beta^\delta} - z\|_H = O(\delta)$) and

$$\|q_\beta^\delta - q^+\|_I = O(\sqrt{\delta}) \tag{4.9}$$

hold. The constants in the O-terms in (4.8) and (4.9) can be derived from the proof. □

Now, the theoretical analysis of our identification problem is complete. We have shown stability, convergence and given conditions on the rate of convergence. In the next section, we give sufficient conditions for the construction of a source function w that satisfies (4.1)-(4.3), which allows the interpretation of the source condition (4.2).

5 Discussion of the Source Condition

The crucial assumption in our result on the convergence rate is the existence of a source function

$$w \in V \tag{5.1}$$

with

$$\Delta w \in H \tag{5.2}$$

such that

$$\left(q^\dagger - q^*, \psi\right)_I = -\int_\Omega \psi(u_{q^\dagger})\nabla u_{q^\dagger}\nabla w dx \tag{5.3}$$

holds for all $\psi \in H^1(I)$. We recall that $(\cdot,\cdot)_I$ denotes the H^1-scalar product. We first note that the formulation of the source condition itself does not impose any more regularity on q^* and on $q^\dagger$ than they being in $H^1(I)$. This is opposed to the general theory where Fréchet-differentiability of the forward operator F and Lipschitz continuity of $F'(q)$ already require a more regular parameter. Furthermore, in the general theory the adjoint of $F'(q^\dagger)$ is needed which makes the source condition usually very difficult to comprehend. In our new approach only the parameter-to-solution map $u_{q^\dagger}$ itself is used, which has a direct physical meaning, not its linearization. This is also a marked difference to the approach of [2] for the case of a space-dependent parameter.

So far, source conditions could at best be interpreted in the sense that they gave some information on smoothness and boundary behaviour of the exact solutions needed for convergence rates (see [1]). Very recently, in [2] for the first time a source function could be constructed - at least for the one-dimensional case - under quite natural conditions for the problem of identifying a space dependent parameter in a linear parabolic equation. Though in our case things get even much more complicated because already the direct problem is nonlinear, we even succeeded in the construction of a source function in higher dimensions. This will be explained in detail now.

First, we look again at the interval I introduced in Section 3. If we denote the range of the true temperature $u_{q^\dagger}$ on Ω by the interval $I^\dagger = \left[I^\dagger_{\min}, I^\dagger_{\max}\right]$, i.e.

$$I^\dagger_{\min} = \min_{x\in\Omega} u_{q^\dagger} \text{ and } I^\dagger_{\max} = \max_{x\in\Omega} u_{q^\dagger},$$

we clearly have

$$I^\dagger \subseteq I.$$

If we can assume to know the true parameter $q^\dagger$ outside the true range $I^\dagger$ already from q^*, i.e., for $\rho := q^\dagger - q^*$ we have

$$\rho = 0 \text{ for } \tau \in I\backslash I^\dagger, \tag{5.4}$$

the source condition (5.3) turns to

$$(\rho,\psi)_{I^\dagger} = -\int_\Omega \psi(u_{q^\dagger})\nabla u_{q^\dagger}\nabla w dx \quad \forall \psi \in H^1(I).$$

This is a very natural assumption as outside of $I^\dagger$ the parameter is of no use for the physical system. Also, from the inverse point of view we only can expect to identify the parameter on the range of the true temperature as outside of $I^\dagger$ no information is available. Hence, on $I\backslash I^\dagger$ $q^\dagger$ is already determined by the choice of q^*.

For our further considerations it is necessary to recall the *co-area formula* , see e.g. [6]:

Theorem 5.1 (The Coarea Formula). *Let a function $t : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous. By L^n, we denote the Lebesque measure on $\mathbb{R}^n$, by H^{n-1}, we denote the (n-1)-dimensional Hausdorff measure. Then for each L^n-summable function $g : \mathbb{R}^n \to \mathbb{R}$,*

$$g|_{t^{-1}\{\tau\}} \text{ is } H^{n-1}\text{-summable for } L\text{-almost every } \tau \in \mathbb{R}$$

and

$$\int_{\mathbb{R}^n} g(x)Jt(x)dx = \int_{\mathbb{R}} \left[\int_{t^{-1}\{\tau\}} g dH^{n-1} \right] d\tau. \tag{5.5}$$

This co-area formula, which we give here for the case of $t : \mathbb{R}^n \to \mathbb{R}$, is a special case of a more general analogous formula for $t : \mathbb{R}^n \to \mathbb{R}^k$, which contains, as special cases, the change-of-variables formula (for $k = n$) and Fubini's Theorem (for t being a projection onto a subspace). The basic concept is that the integral of the (n-1)-dimensional Hausdorff measure ($\int_{t^{-1}\{\tau\}} dH^{n-1}$) of the level sets of t, integrated over all the levels, gives the L^n-integral of the Jacobian Jt. We will use this by integrating level-sets of a given temperature level, i.e., isotherms, over these temperature levels.

In order to use the coarea formula we have to assume more regularity (as in (2.2)) on f in the direct problem and to require that

$$q^\dagger - q^* \in H^4(I). \tag{5.6}$$

Then - because of the compact embedding $H^4(I^\dagger) \subset C^3(I^\dagger)$ - our assumptions on the a priori knowledge about $\rho = q^\dagger - q^*$ (see (5.4)) result in

$$\rho^{(i)}(I^\dagger_{\min}) = \rho^{(i)}(I^\dagger_{\max}) = 0 \quad \text{for} \quad i = 0, 1, 2, 3. \tag{5.7}$$

Furthermore, (5.6) implies (see (2.3)) that $u_{q^\dagger}$ belongs to $\in C^{2,\alpha}(\bar{\Omega})$, from which we immediately deduce that the transformation

$$u_{q^\dagger} : \Omega \to I^\dagger$$

is Lipschitz. Hence the coarea formula (5.5) is valid for the transformation $t = u_{q^\dagger}$, whose level-sets are isotherms in Ω. The next step is to find a suitable function g in (5.5) for our purpose. First, we define m to be the (n-1)-dimensional Hausdorff-measure of the level sets of $u_{q^\dagger}$, i.e.

$$m(\tau) = \int_{u^{-1}_{q^\dagger}\{\tau\}} dH^{n-1} \quad \text{for} \quad \tau \in I^\dagger,$$

and assume $u_{q^\dagger}$ to behave such that

$$(\rho'''(u_{q^\dagger}) - \rho'(u_{q^\dagger})) \cdot \frac{1}{m(u_{q^\dagger})} \cdot Ju_{q^\dagger} \in L^2(\Omega). \tag{5.8}$$

The only term that might cause a violation of (5.8) is $\frac{1}{m(u_{q^\dagger})}$. If $m(u_{q^\dagger})$ is bounded away from 0, i.e., if all temperatures in $I^\dagger$ are assumed on sets of non-vanishing H^{n-1}-measure, and if these measures depend in a reasonable way on the temperatures, then (5.8) is certainly fulfilled. This is a (weak) regularity condition on the measures of the isotherms, and it is reasonable, since one cannot expect identifiability of $q^\dagger$ for temperatures which are assumed only on a "small" set (of H^{n-1}-measure zero). But even such "small isotherms" are not excluded by (5.8): The only temperature values possibly attained by $u_{q^\dagger}$ at a set of H^{n-1}-measure zero are $I^\dagger_{\min}$ or $I^\dagger_{\max}$. This is a consequence of the continuity of $u_{q^\dagger}$ and the intermediate value theorem. But both the first term in the product (see (5.7)) and $Ju_{q^\dagger}$ (necessary condition for an extremum) vanish in the respective critical situation. Now if the product of these two expressions tends faster to zero than $m(u_{q^\dagger})$, the L^2-boundedness in (5.8) is maintained, even then.

Now, we look at the Poisson equation

$$\begin{aligned} \Delta w &= (\rho'''(u_{q^\dagger}) - \rho'(u_{q^\dagger})) \cdot \frac{1}{m(u_{q^\dagger})} \cdot Ju_{q^\dagger} \text{ in } \Omega \qquad (5.9) \\ w &= 0 \text{ on } \partial\Omega, \qquad (5.10) \end{aligned}$$

which has a unique (weak) solution $w \in V$ as the right-hand side belongs to H (see (5.8)). We now show that this function w satisfies the source condition (5.3). Note that we automatically fulfilled the conditions (5.1) and (5.2) of our convergence rate theorem by this approach.

Theorem 5.2. *Assume that* (5.4), (5.6) *and* (5.8) *hold. Then the weak solution of* (5.9), (5.10) *satisfies* (5.3).

Proof. We choose an arbitrary test function $\psi \in H^1(I^\dagger)$, denote its indefinite integral by Ψ and fix it by $\Psi(0) = 0$. First we get

$$\int_\Omega \psi(u_{q^\dagger}) \nabla u_{q^\dagger} \nabla w dx = \int_\Omega \nabla \Psi(u_{q^\dagger}) \nabla w dx.$$

Now, integration by parts yields

$$\begin{aligned} -\int_\Omega \nabla \Psi(u_{q^\dagger}) \nabla w dx &= -\int_{\partial\Omega} \Psi(u_{q^\dagger}) \frac{dw}{dn} dS \\ &\quad + \int_\Omega \Psi(u_{q^\dagger}) \Delta w dx \end{aligned}$$

As we have chosen the integration constant such that $\Psi(0) = 0$ and because $u_{q^\dagger}$ vanishes on the boundary $\partial\Omega$, we get

$$-\int_\Omega \psi(u_{q^\dagger}) \nabla u_{q^\dagger} \nabla w dx = \int_\Omega \Psi(u_{q^\dagger}) \Delta w dx.$$

Next, we replace Δw by the right hand side of (5.9) and obtain

$$-\int_\Omega \psi(u_{q^\dagger})\nabla u_{q^\dagger}\nabla w dx = \int_\Omega \Psi(u_{q^\dagger})(\rho'''(u_{q^\dagger}) - \rho'(u_{q^\dagger}))\frac{1}{m(u_{q^\dagger})} J u_{q^\dagger} dx$$

Using the coarea formula (5.5) with

$$g(x) = \Psi(u_{q^\dagger}(x))(\rho'''(u_{q^\dagger}(x)) - \rho'(u_{q^\dagger}(x)))\frac{1}{m(u_{q^\dagger}(x))},$$

which is L^n-summable because of (5.8) and $\psi \in H^1(I)$, results - as g is constant on any level set $u_{q^\dagger}^{-1}(\tau)$ - in

$$\begin{aligned} -\int_\Omega \psi(u_{q^\dagger})\nabla u_{q^\dagger}\nabla w dx &= \int_{I^\dagger} \Psi(\tau)(\rho'''(\tau) - \rho'(\tau))\frac{1}{m(\tau)} m(\tau) d\tau \\ &= \int_{I^\dagger} \Psi(\rho''' - \rho') d\tau. \end{aligned}$$

Finally, integration by parts with respect to τ yields

$$\begin{aligned} \int_{I^\dagger} \Psi(\rho''' - \rho') d\tau &= \Psi(\rho'' - \rho)|_{I^\dagger_{\min}}^{I^\dagger_{\max}} \\ &\quad - \int_{I^\dagger} \psi(\rho'' - \rho) d\tau \\ &= \left[\Psi(\rho'' - \rho) - \psi\rho'\right]|_{I^\dagger_{\min}}^{I^\dagger_{\max}} \\ &\quad + \int_{I^\dagger} (\psi'\rho' + \psi\rho) d\tau \end{aligned} \tag{5.11}$$

As the boundary terms in (5.11) vanish because of (5.7) we obtain

$$-\int_\Omega \psi(u_{q^\dagger})\nabla u_{q^\dagger}\nabla w dx = \left(q^\dagger - q^*, \psi\right)_{I^\dagger}.$$

□

Thus, under the assumptions of Theorem 5.2, the convergence rates from Theorem 4.1 are valid, since we could verify the source condition. The essential conditions needed for Theorem 5.2 were

- sufficient smoothness of $q^\dagger - q^*$ as required in (5.6)
- sufficient knowledge about $q^\dagger$ on the boundary of the temperature interval where measurements are available: (5.7)
- the assumption (5.8), which essentially says that the isotherms of $u_{q^\dagger}$ depend in a sufficiently regular way on the temperature level, where this regularity is rather weak.

6 Variants and Extensions

The results achieved in this paper are still in the infinite-dimensional setting, unrealistically assuming that we have distributed measurements of u. Of course, in reality only discrete measurements are available, which requires to extend the theory by adding discretization in solving the forward problem and in approximating the unknown parameter. This can be done in a rather standard way (cf. [1]).

In our optimization Problem (3.8), we use a H^1-regularization term. Instead, we could also consider L^2-regularization, i.e.,

$$\min_{q\in K} J(q) := \int_\Omega |u_q - z^\delta|^2 dx + \beta \int_I (q-q^*)^2 d\tau. \tag{6.1}$$

Results corresponding to Theorems 3.3, 3.4 and 4.1 can be proven in a completely analogous way, where the H^1-scalar product $(q^\dagger - q^*, \psi)_I$ in (4.2) is replaced by the L^2-scalar product $\int_I (q^\dagger - q^*)\psi d\tau$. A solution to the source condition can be constructed as in Section 5 under even weaker regularity assumptions on $q^\dagger$.

The existence of minimizers of (6.1) cannot be guaranteed by Theorem 3.1. This difficulty could either be resolved by using some weak-closedness argument or circumvented by incorporating a tolerance η into the minimization, i.e., replacing minimizers of (6.1) by elements $q^\delta_{\beta,\eta} \in K$ such that

$$J(q^\delta_{\beta,\eta}) \le \inf_{q\in K} J(q) + \eta.$$

As long as $\eta = O(\delta^2)$, all proofs carry over (see [13]). This can of course also be done for (3.8).

Instead of applying the general theory of Tikhonov regularization to our problem, we succeeded in proving results about convergence rates under much weaker conditions. We mention that this problem-adapted new approach also works for the corresponding nonlinear parabolic problem, where we identify the parameter from distributed measurements of the temperature u over a final time period $[T-\sigma, T]$. In this case the source condition changes from (4.2) to

$$\left(q^\dagger - q^*, \psi\right)_I = -\int_{T-\sigma}^T \int_\Omega \psi(u_{q^\dagger}) \nabla u_{q^\dagger} \nabla w dx dt \quad \forall \psi \in H^1(I),$$

where σ is a small positive constant. Now, the source function w also depends on the time variable which requires additional assumptions, i.e. $w(x,T) = 0$ and the L^2-boundedness of w_t. These can be interpreted as smoothness conditions on the behaviour of the temperature level sets with respect to the time.

Furthermore, the domain of measurements can be reduced to a subdomain of Ω both in the elliptic and the parabolic case. Identification from boundary measurements will be discussed in [18].

7 A Numerical Example

For the numerical test we consider the (one-dimensional) parabolic problem. We assume a temperature field

$$u_{q^\dagger} = (x - x^2) \cdot \sin(\pi t) \text{ on } [0,1] \times [0, 0.5]$$

corresponding to a true heat conductivity

$$q^\dagger(\tau) = 2 + \sin(8\pi\tau). \tag{7.1}$$

The source term f(x,t) is chosen such that it matches the right hand side of the direct problem, i.e.

$$f(x,t) = \frac{\partial u_{q^\dagger}}{\partial t} - \frac{\partial}{\partial x}(q^\dagger(u_{q^\dagger})\frac{\partial u_{q^\dagger}}{\partial x}).$$

Now, the goal is to recover $q^\dagger$ by a sequence of data

$$z_i(x,t) \text{ on } [0,1] \times [0.47436, 0.5]$$

given on a short final time period. Hence, our objective functional to be minimized during the identification process is

$$J_\beta(q) = \int_{0.47436}^{0.5} \int_0^1 |u_q - z|^2 dxdt + \beta \|q - q^*\|_I^2,$$

where u_q again denotes the solution of the direct problem for the input parameter q. For the numerical realization we used a BFGS-Quasi-Newton method. A detailed explanation can be found in [3].

In the following we distinguish between parameter plots and temperature plots, e.g. in figure 1. Although we suppose the data to be available on $[0.47436, 0.5]$, a temperature plot always only shows the situation at the end of the time period, i.e. t = 0.5. Our sequence of (noisy) data actually contains four members

$$z_i(x,t) \in [0,1] \times [0.47436, 0.5] \quad \text{i = 1,...,4.}$$

The noise level

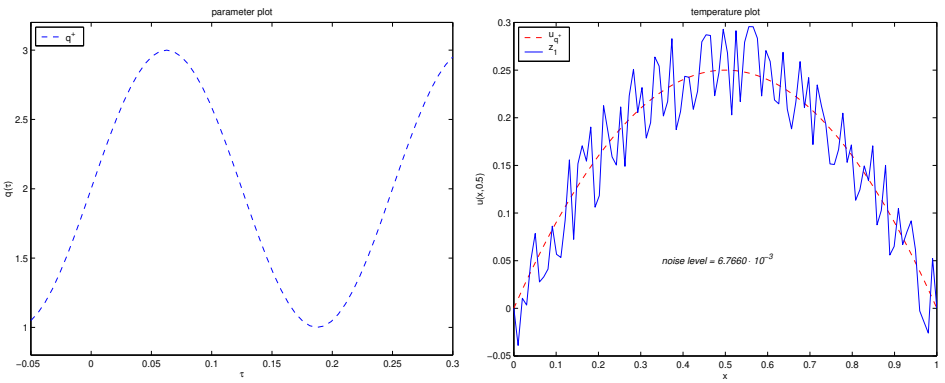

Figure 1: the true parameter $q^\dagger$ and the noisy data z_1 at t = 0.5

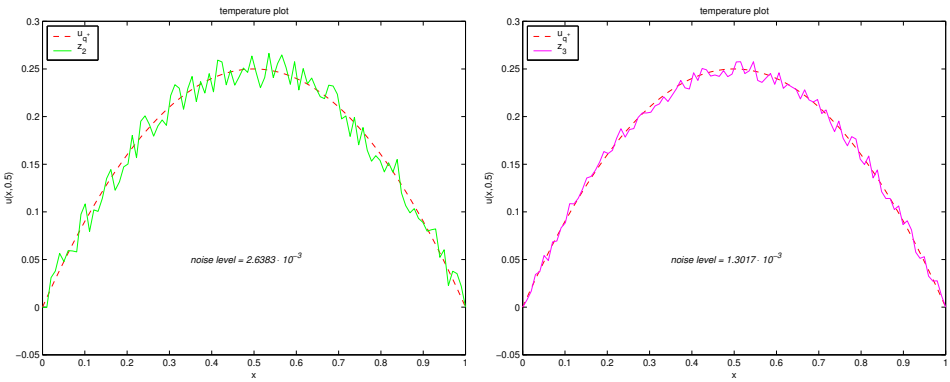

Figure 2: the noisy data z_2 and z_3 at t = 0.5

$$\delta_i = \sqrt{\int_{0.47436}^{0.5} \int_0^1 |u_{q^\dagger} - z_i|^2 dx dt}$$

decreases as the index i increases (see figures 1 and 2) until we reach

$$z_4(x,t) = u_{q^\dagger},$$

the case of an error free measurement of the temperature. Actually, the noise levels are

i	δ_i
1	$6.7660 \cdot 10^{-3}$
2	$2.6383 \cdot 10^{-3}$
3	$1.3017 \cdot 10^{-3}$
4	0

The parameters computed by the help of these data, i.e. the minimizers of

$$J_{\beta_i}(q) = \int_{0.47436}^{0.5} \int_0^1 |u_q - z_i|^2 dxdt + \beta_i \|q - q^*\|_I^2, \tag{7.2}$$

over the set K ($q_{\min} = 0.01$, $q_{\max} = 10$) are denoted by q_i. Given the data z_i we still cannot start our computations unless we specify the initial guess q^* and especially the interval I. Assuming that there is no a priori information on I available we can only ask the given data, say z_1, for help. From figure 1 we know that $z_{1\min} \approx -0.05$ and $z_{1\max} \approx 0.3$. Even at the best case this only gives a hint at the range of $u_{q^\dagger}$ for $t \in [0.47436, 0.5]$. Not knowing if $u_{q^\dagger}$ exceeds the range $[-0.05, 0.3]$ at some time $t \in [0, 0.47436]$, we first choose a sufficiently large (it is to be hoped) interval

$$\tilde{I} = [-0.25, 0.5].$$

These considerations are important as the time loop in the solver for the direct problem starts at $t = 0$ and not at $t = 0.47436$, i.e., the interval must cover the whole range of the temperatures. It turns out that capturing the "right" interval I during the numerical computations is the crucial point.

For a constant initial guess $q^* = 4$ we get, after 300 optimization steps, the left side of figure 3. Only looking at the solid line q_1, we learn that our initial guess q^* is not

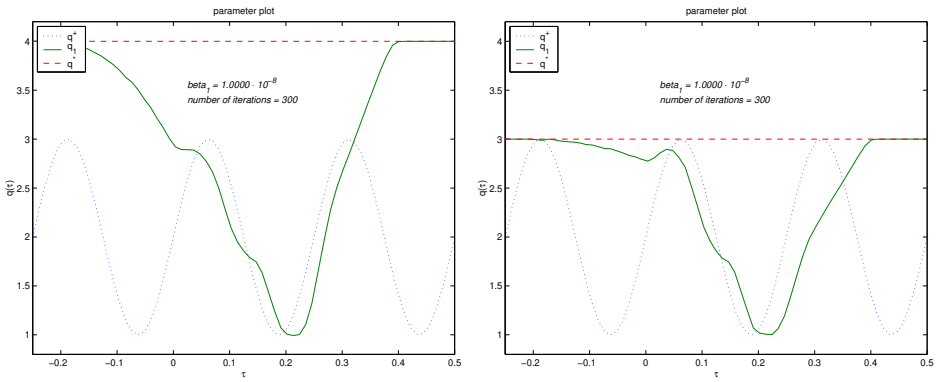

Figure 3: q_1 identified from z_1 with $q^* = 4$ and $q^* = 3$

that good. We maybe even have chosen the interval $\tilde{I}$ too large, as for example nothing changes in $[0.4, 0.5]$. But in order to make sure, we try $\tilde{I}$ again and use a lower initial guess $q^* = 3$. The result is shown on the right side of figure 3. The kink at $\tau \approx 0$ hints that q^* still is too large. We are also confirmed in our suspicion that changes of q^* outside of $[z_{1\min}, z_{1\max}]$ arise only because of the H^1-regularization term in (7.2). We try $\tilde{I}$ once more and choose $q^* = 2$ which yields figure 4. Figure 3 and the left part of figure 4 may

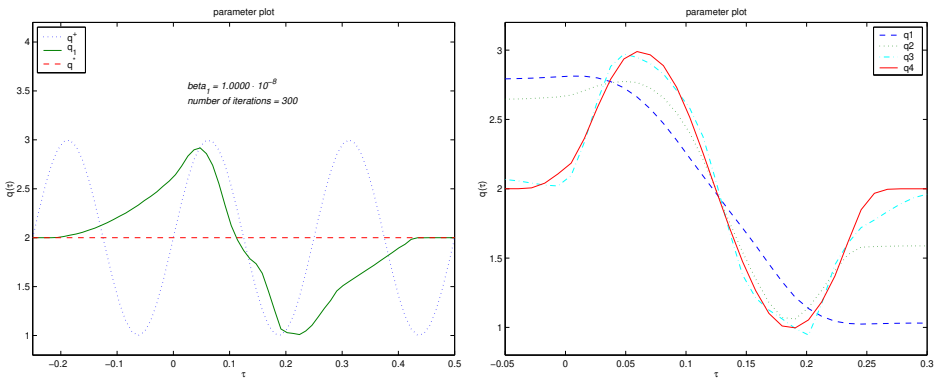

Figure 4: q_1 identified from z_1 with $q^* = 2$ and the sequence of solutions q_i

justify that from now on we set the interval I as

$$I = [-0.05, 0.3].$$

Of course, given only the data z_1 we still cannot exclude that the temperature over the whole time interval exceeds this range. Tests omitting the regularization term in order to avoid its influence on the changes are impossible because of the arising instabilities, especially in the case of noisy data. Naturally, we also tried a larger number of maximal iterations than 300 but that did not improve the results perceptibly. This example shows that even if no a-priori information on the interval $I^\dagger$ is available, a reliable approximation I can be found, which is crucial for the final aim of computing $q^\dagger$.

Now we turn to the sequence of data $\{z_i\}$. Although we already know from the above results that the true parameter behaves somehow like a sine function on the interval $[-0.05, 0.3]$, we stay with the initial guess $q^* = 2$. As we want to compare the results q_i for different data z_i to each other and also look at the convergence rate, we have also to keep the interval

$$I = [-0.05, 0.3]$$

fixed, although the data z_2, z_3 and z_4 suggest to even try smaller intervals, respectively. Choosing in each situation the regularization parameter β_i as the smallest one admitting a stable result, we get the right side of figure 4 and see that our sequence converges to the true parameter (7.1). Note, that q_4 does not show the plot of (7.1) but the solution computed from the data z_4. On the range of the true parameter $I^\dagger = [0, 0.25]$ we obtain a remarkably good result for $q^\dagger$. Outside of course we get a smooth continuation due to the interval I chosen too large, since we have no data there.

Next, we summarize this result with respect to the rate of convergence. Denoting the

residuals corresponding to the temperatures u_{q_i} by

$$\|u_{q_i} - z_i\| = \sqrt{\int_{0.47436}^{0.5} \int_0^1 |u_{q_i} - z_i|^2 dx dt}$$

we have

i	δ_i	β_i	$\|u_{q_i} - z_i\|$	$\|q_i - q^\dagger\|_I$	$\frac{\delta_i^2}{\beta_i}$
1	$6.766 \cdot 10^{-3}$	$7.5 \cdot 10^{-11}$	$6.6816 \cdot 10^{-3}$	9.8028	$6.104 \cdot 10^5$
2	$2.6383 \cdot 10^{-3}$	$1 \cdot 10^{-9}$	$2.6592 \cdot 10^{-3}$	7.9609	$6.961 \cdot 10^3$
3	$1.3017 \cdot 10^{-3}$	$1 \cdot 10^{-8}$	$1.2922 \cdot 10^{-3}$	6.6671	$1.695 \cdot 10^2$
4	0	$7.5 \cdot 10^{-8}$	$1.5316 \cdot 10^{-5}$	5.6464	0

Finally, we call the attention to a discussion with respect to the initial guess q^* which not only determines the regularization term but also is the starting value in the optimization routine. We only consider error free measurements in order to obtain a result easier to survey and choose

$$q^* = 4.$$

First, we compute the temperature u_{q^*} corresponding to the input parameter q^* and get the left side of figure 5. Compared to the range $I^\dagger$ of the true temperature $u_{q^\dagger}$, u_{q^*} only

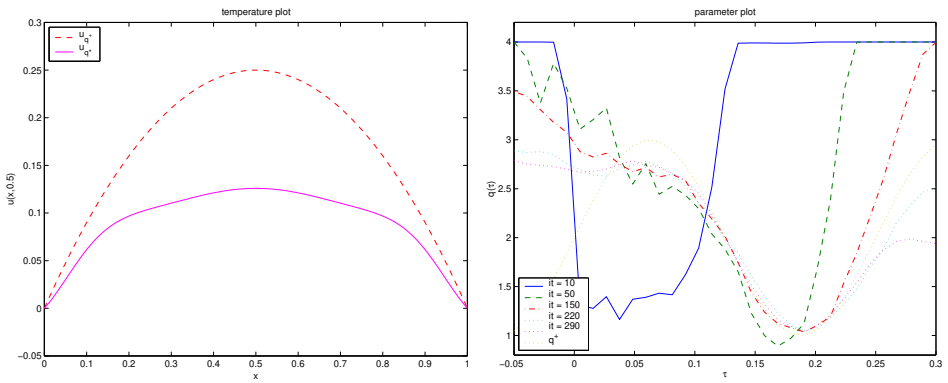

Figure 5: the solution for $q^* = 4$ and the process of conquering I

covers over $\approx [0, 0.125]$. This effect is the reason for another problem in the recovery process. From [3] we know that the components of the gradient of J_{β_4}, evaluated at q^*, vanish on the part of I which is not covered by the range of u_{q^*}. Outside, the information first must be gained during the optimization process, which explains the comparatively large numbers of iterations. The right side of figure 5 illustrates how the whole interval I - being a hint at the true range of the temperature - is conquered. Recall that in our situation we cannot expect the full conquest of I as it is chosen based on the noisy data z_1,

while here the calculations are based on the exact data z_4 covering a smaller range than z_1.

We have seen that identifying a nonlinearity entails a lot of numerical difficulties even in the one-dimensional case. Nevertheless, a numerical treatment based on a deep analysis of Tikhonov regularization yields satisfactory results.

Acknowledgement: We thank Prof. Paul Müller (Linz) for pointing out the coarea formula to us.

References

[1] H.W. Engl, M. Hanke and A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, 1996 (Paperback edition 2000)

[2] H.W. Engl and J. Zou, A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction, Inverse Problems 16 (2000), 1907-1923

[3] P. Kügler, Identification of a Temperature Dependent Heat Conductivity by Tikhonov Regularization, Diploma thesis, Johannes Kepler University, Linz, Austria, January 2000

[4] Y. Keung and J. Zou, Identifying parameters in parabolic systems by finite element methods, Research Report of The Chinese University of Hong Kong, 1997

[5] D. Gilbarg and N. Trudinger, Elliptic Partial Differential Equations of Second Order, Springer, 1977

[6] L. Evans and R. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, 1992

[7] M. Protter and H. Weinberger, Maximum Principles in Differential Equations, Prentice-Hall, 1967

[8] R.E. Showalter, Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, American Mathematical Society, 1997

[9] E. Schock, Approximate solution of ill-posed equations: arbitrarily slow convergence vs. superconvergence, in: G. Hämmerlin and K. Hoffmann, eds., Constructive Methods for the Practical Treatment of Integral Equations, Birkhäuser, 1985

[10] W. Grever, Optimal Cooling Strategies in Continuous Casting of Steel with Variable Casting Speed, Dissertation, Johannes Kepler Universität Linz, 1995

[11] H.W. Engl, T. Langthaler and P. Manselli, On an inverse problem for a nonlinear heat equation connected with continuous casting of steel, in: K.H. Hoffmann, W. Krabs (eds.), Optimal Control with Partial Differential Equations II, ISNM 78, Birkhäuser 1987, 67-89

[12] H.W. Engl and T. Langthaler, Control of the solidification front by secondary cooling in continuous casting of steel, in: H.W. Engl, H. Wacker and W. Zulehner (eds.), Case Studies in Industrial Mathematics, Teubner/Kluwer 1988, 51-77

[13] H.W. Engl, K. Kunisch and A. Neubauer, Convergence rates for Tikhonov regularization of nonlinear ill-posed problems, Inverse Problems 5 (1989), 523-540.

[14] T.I. Seidman and C.R. Vogel, Well–posedness and convergence of some regularization methods for nonlinear ill–posed problems, Inverse Problems 5 (1989), 227-238.

[15] A. Neubauer, Tikhonov regularization for nonlinear ill–posed problems: optimal convergence and finite–dimensional approximation, Inverse Problems 5 (1989), 541–557.

[16] K. Kunisch and W. Ring, Regularization of nonlinear illposed problems with closed operators, Numer. Funct. Anal. and Optimiz. 14 (1993), 389-404.

[17] O. Scherzer, H.W. Engl and R.S. Anderssen, Parameter identification from boundary measurements in a parabolic equation arising from geophysics, Nonlinear Anal. 20 (1993), 127–156.

[18] P. Kügler, Identification of a temperature dependent heat conductivity from boundary measurements, in preparation