A Second Degree Method for
Nonlinear Inverse Problems

F. Hettlich [*]   W. Rundell [†]

November 17, 1998

Abstract: The paper is concerned with the solution of nonlinear ill-posed problems by methods that utilise the second derivative. A general predictor–corrector approach is developed; one which avoids solving quadratic equations during the iteration process. Combining regularisation of each iteration step with an adequate stopping condition leads to a general regularisation scheme for nonlinear equations. Possible implementations and discussion of the performance of this method are illustrated by applications to some well–known inverse problems.

Keywords: Nonlinear ill–posed problems, iterative regularisation methods, domain derivative

AMS subject classification: 35R30, 65J15, 65J20

1   Introduction

In inverse problems we are often concerned with solving the nonlinear equation

$$F(x) = g, \tag{1.1}$$

where $F : U \subseteq X \to Y$ is a differentiable operator between Hilbert spaces; those relating the unknown x (in our viewpoint, a coefficient in a partial differential operator) and the data g (typically some functional of the solution of the differential operator measured as output of the system).

A classical numerical approach would be the Newton method, or if (1.1) is formulated as a least squares problem to minimise the functional ϕ defined by $\phi(x) = \frac{1}{2}\|r(x)\|_2^2$,

$r(x) := F(x) - g$, the Gauß–Newton method or modifications [2]. In essence these schemes rely on a first order model; F is approximated by

$$F(x + h) \approx F(x) + F'[x]h \tag{1.2}$$

and the present approximation x is corrected to $x + h$ by setting $F(x + h) = g$. If the problem is ill–posed, as almost all inverse coefficient problems are, then certain difficulties occur. The ill–posedness stems from the fact that the map F is compact, a property inherited by the derivative F'. Thus the inversion of the map $F'[x]h$ will require regularisation and this is typically accomplished by finding an invertible approximant of F' or suitably restricting the class of allowable solutions x, and, possibly, combining this with a stopping rule governing the iteration process.

A second degree model would seek to determine the correction h to the current value x from the Taylor polynomial

$$F(x + h) \approx F(x) + F'[x]h + \frac{1}{2}F''[x](h, h). \tag{1.3}$$

In the least squares setting, the second derivative of ϕ involves the residual multiplied by the individual hessians. This can lead to diminishing returns if the residual is small.

In the optimisation literature there have been many schemes and implementations that utilise the second derivative in approximating solutions of (1.1). However, these have tended to be less popular than schemes that require at most a single derivative, see for example [2]. In particular, it is rare in the inverse problems literature to see the use of derivatives higher than the first. However, some results in the case of ill–conditioned problems in $\mathbb{R}^n$ are known from approaches using an approximation of the second derivative [23].

In the general theory, there are two main reasons cited for avoiding the use of schemes requiring the second derivative. The first is that if the residual is small, then the second derivative is also small; thus near convergence the contributions of the terms involving the higher order derivatives is negligible. This is particularly pronounced when the nonlinearity in the function F is mild. The second is based on the fact that in many applications the increased convergence rate of the higher order scheme does not repay the computational effort of computing the Hessian matrices.

As we will show in this paper, these objections are to some extent unfounded for many classical inverse problems in partial differential equations. As noted, such problems are frequently highly ill–posed and require some regularisation procedure, typically by constraining some aspect of the solution being sought. In consequence, the data, even if accurate, may arise from a function not in the designated constraint class and so the residual will never be zero. In the presence of noise, more regularisation will be required further restricting the degree to which the residual can be reduced. Indeed, it has been observed over a wide variety of problems that

iterative solution algorithms should usually terminate the iteration procedure before the residual reaches the minimum possible by the method [5].

For many undetermined coefficient problems in partial differential equations where one is trying to invert the map F from the unknown coefficient c to the data g, it is possible to represent the derivative F' as the solution of a differential equation with similar form to that specifying the map F. In this case advantage can be taken of the effort already performed in the computation of the direct map F to compute F' relatively cheaply. A similar situation also holds for the second derivative of the map. In the examples we use to illustrate the scheme, both of the first and second derivatives can be computed at less than the cost of another direct solve.

Even in those situations where derivatives cannot be evaluated without considerable additional computation or where either the nonlinearity is relatively weak or the data error small, there is potential advantage in using our second degree scheme. The advantage of using a "frozen" Newton method, where the derivative is held at an initial guess throughout the iteration process has been demonstrated in a wide variety of inverse coefficient problems. By choosing a sufficiently simple coefficient, for example a constant, it is sometimes possible to explicitly compute the relevant derivatives. In this situation one is also able to better estimate the degree of regularisation required for stable inversion. Even if this explicit calculation is not feasible, the method can provide computational advantages. One disadvantage of this approach is that inevitably more iterations are required. However, this can be completely offset by using a second degree scheme that freezes both the first and second derivative at a known solution. Indeed, in this case one can often get effective numerical convergence with less computational effort from the second degree scheme using frozen derivatives than one can by using the full Newton method.

There is a further difficulty in using a second degree model; to update the correction h in (1.3) requires the solution of a quadratic equation to be (regularised and) solved at each iteration step. To avoid this, we use the following predictor–corrector procedure. Let x_0 denote a starting guess and let $\tilde{h}$ be computed by a Newton step,

$$F'[x_n]\,\tilde{h} = g - F(x_n). \tag{1.4}$$

Then the next iteration $x_{n+1} = x_n + h$ is defined from the second degree Taylor remainder by the solution of the linear equation

$$T[x_n]h := F'[x_n]h + \frac{1}{2}F''[x_n](\tilde{h}, h) = g - F(x_n). \tag{1.5}$$

Of course, in each of (1.4) and (1.5) we must address the regularisation issue.

Some advantages are seen at once in one dimension. We can combine both steps and obtain the iteration scheme

$$x_{n+1} = H[f](x_n) := x_n - \frac{f(x_n)f'(x_n)}{(f'(x_n))^2 - \frac{1}{2}f''(x_n)f(x_n)}. \tag{1.6}$$

3

to approximate a root of $f : \mathbb{R} \to \mathbb{R}$. This scheme can handle several functions like $f(x) = \exp(ax) - 1$ with large $a > 0$ and a small but negative starting guess, where the classical Newton method slows down or fails. In fact, (1.6) has been known for some considerable time; its essence was used by Halley to extract roots of polynomials to high accuracy [7]. The connection with derivatives was only made some decades later and it was full century and a half before the iteration function formulation of (1.6) was adopted. For a historical survey of the method in one dimension see [22]. Even in one dimension the two stage predictor-corrector means of solving the second degree Taylor approximation of f (1.4)–(1.5), seems not to have been considered, although other relations to Newton's scheme have been noted. For example, it has been observed that $N[f/\sqrt{f'}] = H[f]$ where $N[f]$ is the usual Newton iteration function applied to f.

We are not claiming that the second degree method is a more stable algorithm or indeed that it offers superior final reconstructions. The lack of compactness in F that is inherited by F' is also inherited by the second derivative and so the additional term cannot be expected to significantly change the ill–conditioning of the problem. In some cases the second degree method with frozen derivatives involved a slightly better–conditioned inversion in the corrector step than in the predictor (which is the only one that would be used in a first degree model). However, we did find advantages in the final reconstruction and often a significant improvement in the number of iterations (and in the computational effort) required before the stopping condition terminated the iteration process.

In the next section of the paper we shall show that the scheme (1.4), (1.5) gives a general solver for well–posed nonlinear equations as long as the second derivative is available. Its application to ill–posed problems requires regularisation of both the predictor and the corrector step and of course we also need a stopping criterion to ensure a reliable approximation. The main part of section 2 will show that using Tikhonov's regularisation for the linear equations and a certain stopping rule will lead to convergence of the scheme under similar conditions as the regularised Newton method [8] or the Landweber iteration [10].

Retaining the predictor-corrector approach to the second degree method leaves many possible variations to be considered. For example, we will hold the derivative operators constant at the value obtained from the starting guess and refer to such schemes and their components as "frozen Newton", "frozen predictor", "frozen corrector" etc. We focus on the "full" and the "frozen" schemes. In a last short section some other possible modifications are discussed.

The subsequent sections are devoted to the analysis of our second degree method as applied to four standard undetermined coefficient problems. By this means we hope to bring out what we see as the main features of the method and using this collection of problems we will address the issues raised earlier in the introduction. In section 3 we apply the scheme to a problem where a parameter has to be determined

from multiple input sources and show that the convergence condition of section 2 holds. The excellent performance of the second degree method with frozen derivatives compared to full Newton is demonstrated. In section 4 we consider the classical inverse Sturm–Liouville problem and we will derive an explicit formula based on our predictor–corrector scheme where the derivatives are taken about the initial constant potential. We will show that this gives an excellent approximation to even relatively large potentials from the initial guess, that is, without recourse to computing even the direct map $F(q)$. In section 5 we consider the problem of recovering the support of an unknown source in Poisson's equation from Cauchy boundary data. In this case we see that the very mild nonlinearity involved makes the use of the second degree scheme less compelling. Finally, in section 6 inverse obstacle scattering is investigated. We show that both the first and second derivatives can be computed here from a scattering problem of identical nature to that for the computation of F. Thus the additional overhead in implementing the scheme (1.4), (1.5) is relatively low and in consequence there is considerable computational advantage in using the second degree method.

2   A second degree method

For ill–posed problems a regularisation strategy of the scheme (1.4), (1.5) is required. We will suggest such a method and in fact show convergence with respect to noise level, but we first consider well–posed problems; presenting a proof of convergence analogous to the established procedure for the classical Newton method [25].

Theorem 2.1 *Let X, Y be Banach spaces and $\hat{x} \in U \subseteq X$ denote a solution of $F(x) = g$. Assume $F'[\hat{x}]$ admits a bounded inverse and F' and F'' are uniformly bounded in U. Then there exists $\delta > 0$ such that the iteration (1.4), (1.5) with starting guess $x_0 \in B(\hat{x}, \delta) = \{x \in X : \|\hat{x} - x\| < \delta\}$ converges quadraticly to $\hat{x}$. If additionally the second derivative is Lipschitz continuous, i.e.*

$$\|F''[x](h, \tilde{h}) - F''[y](h, \tilde{h})\| \le L\|x - y\| \, \|h\| \, \|\tilde{h}\|$$

for all $x, y \in U$ with $h, \tilde{h} \in X$ and a constant $L > 0$, then

$$\|x_{n+1} - \hat{x}\| \le c\|x_n - \hat{x}\|^3$$

holds for $n = 0, 1, 2, \ldots$ with a constant $c > 0$.

Proof: The proof consists of three parts. First we prove that an iteration step is well defined in a neighbourhood of $\hat{x}$. Then we continue in showing that the iterates remain in such a neighbourhood. The third part establishes the estimates which are required for the convergence result.

From Newton's method we know, if $\delta' > 0$ is sufficiently small, that $(F'[x])^{-1}$ exists for $x \in B(\hat{x}, \delta')$, is uniformly bounded in this neighbourhood, and depends continuously on $x \in B(\hat{x}, \delta')$. Thus $T[x]$ defined in (1.5) depends continuously on x. Additionally we observe that $T[\hat{x}] = F'[\hat{x}]$. Again by a perturbation argument we can reduce δ' such that $(T[x])^{-1}$ exists in $B(\hat{x}, \delta')$ and is uniformly bounded (replace x_n by x and x_{n-1} by $\hat{x}$ in (2.1)). This shows that an iteration step (1.4), (1.5) is well defined if x_n is sufficiently close to $\hat{x}$.

Next we show that the iteration remains in a neighbourhood of $\hat{x}$. Let $\beta > 0$ denote an upper bound of $\|T(x)^{-1}\|$ and $\|F(x)^{-1}\|$ in $B(\hat{x}, \delta')$. We choose $0 < \alpha \leq \delta'/3$ such that $\|T[x] - T[\hat{x}]\| \leq \frac{1}{2\beta}$ for $\|x - \hat{x}\| \leq 3\alpha$ and $4\beta C''\alpha < 1$, where C'' denotes an upper bound of $\|F''[x]\| = sup_{\|h\|=\|\tilde{h}\|=1}\{\|F''[x](h,\tilde{h})\|\}$ for $x \in B(\hat{x}, \delta')$.

Now we define $0 < \delta < \alpha$ with $\|g - F(x)\| \leq \frac{\alpha}{\beta}$ for all $x \in B(\hat{x}, \delta)$. Let $\tilde{x}_n = x_{n-1} + (F'[x_{n-1}])^{-1}(g - F(x_{n-1}))$ be the predictor. Then $x_0 \in B(\hat{x}, \delta)$ yields $\|\tilde{x}_1 - x_0\| \leq \alpha$ and $\|x_1 - x_0\| \leq \alpha$. Considering the Neumann series in

$$T[x_n] = T[x_{n-1}]\Big(I + (T[x_{n-1}])^{-1}(T[x_n] - T[x_{n-1}])\Big) \tag{2.1}$$

leads to

$$\|(T[x_n])^{-1}\| \leq \frac{\|(T[x_{n-1}])^{-1}\|}{1 - \|(T[x_{n-1}])^{-1}\|/(2\beta)} \leq 2\beta.$$

From the inequalities

$$
\begin{aligned}
\|x_{n+1} - x_n\| &= \|(T[x_n])^{-1}(g - F(x_n))\| \\
&\leq \|(T[x_n])^{-1}\|\,\|F(x_n) - F(x_{n-1}) - T[x_{n-1}](x_n - x_{n-1})\| \\
&\leq C''\beta\left(\|x_n - x_{n-1}\|^2 + \|(F'[x_{n-1}])^{-1}(g - F(x_{n-1}))\|\,\|x_n - x_{n-1}\|\right) \\
&\leq \frac{1}{2}\|x_n - x_{n-1}\|
\end{aligned}
$$

induction shows that $\|x_{n+1} - x_n\| \leq \alpha$. This also holds for the predictor step $\tilde{x}_{n+1}$ (replace $T[x_n]^{-1}$ by $F'[x_n]^{-1}$). Moreover, we conclude

$$\|x_{n+1} - \hat{x}\| \leq \sum_{k=0}^{n}\|x_{k+1} - x_k\| + \|x_0 - \hat{x}\| = \alpha\sum_{k=0}^{n}\frac{1}{2^k} + \alpha \leq 3\alpha.$$

Quadratic convergence of the scheme is obvious by

$$\|\hat{x} - x_{n+1}\| = \|\hat{x} - x_n - (T[x_n])^{-1}(g - F(x_n))\| \tag{2.2}$$

$$\leq \beta\Big(\|F(\hat{x}) - F(x_n) - F'(x_n)(\hat{x} - x_n) - \frac{1}{2}F''(x_n)(\hat{x} - x_n, \hat{x} - x_n)\|$$

$$+ \frac{1}{2}\|F''(x_n)(\hat{x} - x_n - (F'[x_n])^{-1}(g - F(x_n)), \hat{x} - x_n)\|\Big)$$

and the quadratic convergence of the Newton method, i.e. $\|\hat{x} - \tilde{x}_{n+1}\| \leq c\|\hat{x} - x_n\|^2$ with $c > 0$.

To show a cubic convergence rate we proceed as in the case of the standard Newton method. For $x, y \in B(\hat{x}, \delta')$, by the Hahn–Banach theorem we define a bounded linear operator $A \in Y^*$ with $\|A\| = 1$ and

$$A\Big(F(y) - F(x) - F'[x](y - x) - \frac{1}{2}F''[x](y - x, y - x)\Big)$$

$$= \ \Big\|F(y) - F(x) - F'[x](y - x) - \frac{1}{2}F''[x](y - x, y - x)\Big\|.$$

Defining $g(t) := A(F(x+t(y-x))$ we observe that $g : [0, 1] \to \mathbb{C}$ is twice differentiable and $|g''(t) - g''(s)| \leq L|t - s|\,\|x - y\|^3$. Since g' and g'' are continuous, we obtain from

$$g(t) = g(0) + g'(0)t - \int_0^t g''(s)(s - t)\,ds$$

that

$$\|F(y) - F(x) - F'[x](y - x) - \frac{1}{2}F''[x](y - x, y - x)\|$$

$$= \ |g(1) - g(0) - g'(0) - \frac{1}{2}AF''[x](y - x, y - x)|$$

$$= \ \left|\int_0^1 \Big(g''(s) - AF''[x](y - x, y - x)\Big)(1 - s)\,ds\right|$$

$$\leq \ \int_0^1 L\|y - x\|^3 s(1 - s)\,ds \ = \ \frac{L}{6}\|y - x\|^3.$$

From equation (2.2) follows

$$\|\hat{x} - x_{n+1}\| \leq \beta\left(\frac{L}{6} + \frac{c\,C''}{2}\right)\|\hat{x} - x_n\|^3. \qquad \square$$

Our goal is to apply the method in the case of noisy data g^δ with $\|g^\delta - g\| \leq \delta$ for a given noise level $\delta > 0$ and an ill–posed problem, i.e. roughly speaking F and $F'[\hat{x}]$ do not allow a continuous inverse [5]. There are several ways to regularise equation (1.4) and (1.5). From now on we assume X and Y to be Hilbert spaces. Then one possibility is to apply Tikhonov regularisation to each iteration step

$$\big((T[x_n^\delta])^*T[x_n^\delta] + \alpha_2 I\big)\,(x_{n+1}^\delta - x_n^\delta) = g^\delta - F(x_n^\delta) \qquad (2.3)$$

where $T[x_n^\delta]$ is now defined by

$$T[x_n^\delta]h = F'[x_n^\delta]h + \frac{1}{2}F''[x_n^\delta](\tilde{h}, h)$$

and the regularised predictor step

$$\tilde{h} = ((F'[x_n^\delta])^* F'[x_n^\delta] + \alpha_1 I)^{-1} (F'[x_n^\delta])^* (g^\delta - F(x_n^\delta)).$$

Here x_n^δ denotes the n-th iteration arising from noisy data g^δ.

For simplicity, the regularisation parameters α_1 and $\alpha_2 > 0$ are chosen to be constant during the iteration since we only present numerical results using constant parameters. More sophisticated parameter strategies are certainly possible (see [8] in case of the Newton method). In common with other iterative schemes, the regularisation of each iteration step is not in itself sufficient. In a generalisation of the discrepancy principle (see [8, 10]) we require a stopping rule, terminating the iteration if

$$\|g^\delta - F(x_n^\delta)\| \leq \tau\delta \leq \|g^\delta - F(x_j^\delta)\| \tag{2.4}$$

is satisfied for all $j \in \{0, \ldots, n-1\}$, where $\tau > 1$ denotes an additional regularisation parameter.

Our proof of convergence of the above scheme requires an assumption on the non-linearity of the operator F. As for the Landweber iteration [10] and the Newton method [8] we assume that

$$\|F(y) - F(x) - F'[x](y - x)\| \leq C\|y - x\| \, \|F(y) - F(x)\| \tag{2.5}$$

with $C > 0$ is satisfied in a neighbourhood $U \subseteq X$ of a solution $\hat{x}$ of $F(x) = g$. Under this condition the following convergence result shows the regularising property of the scheme.

Theorem 2.2 *Let X, Y be Hilbert spaces and assume F satisfies (2.5) and $\|F'\| \leq C'$, $\|F''\| \leq C''$ are uniformly bounded on U. Let $\alpha_1 > 0$ and $x_0 \in U$ denotes a starting guess such that $R_0 = \|x_0 - \hat{x}\| < \rho/(C + \frac{C'C''}{2\alpha_1})$ with $0 < \rho < 1$ and $\{x \in X : \|x - \hat{x}\| \leq R_0\} \subseteq U$. Moreover, we assume*

$$\alpha_2 > \frac{\rho}{1 - \rho} \left(C' + \frac{C'C''}{2\alpha_1} \left(\sup_{x \in B(\hat{x}, R_0)} \|g - F(x)\| + \delta \right) \right)^2.$$

Then,

a) *if $g^\delta = g$, the sequence x_n computed by (2.3) is well defined and converges to a solution x^* of $F(x) = g$.*

b) *if $\|g^\delta - g\| \leq \delta$ and*

$$\tau > \frac{1 + CR_0}{\rho - R_0(C + \frac{C'C''}{2\alpha_1})},$$

there exists $n(\delta) \in \mathbb{N}_0$ such that condition (2.4) is satisfied. Moreover, the approximations $x_{n(\delta)}^\delta$ converge to the solution x^* if the noise level δ tends to zero.

8

The proof of this theorem is closely related to that for the Levenberg–Marquardt scheme presented in [8]. We first show the following monotonicity result.

Lemma 2.3 *Let $0 < \rho < 1 < \gamma$ with $\alpha_2 > \frac{\rho}{1-\rho}\|T[x_n]\|^2$ and $\hat{x} \in U$ be a solution of $F(x) = g$. We assume*

$$\|g^\delta - F(x_n) - T[x_n]\hat{h}_n\| \leq \frac{\rho}{\gamma}\|g^\delta - F(x_n)\| \tag{2.6}$$

with $\hat{h}_n = \hat{x} - x_n$. Then

$$\|\hat{x} - x_{n+1}\|^2 \leq \|\hat{x} - x_n\|^2 - \frac{2(\gamma - 1)\rho}{\gamma}\|g^\delta - F(x_n)\| \, \|v_n\| \tag{2.7}$$

and

$$\|\hat{x} - x_{n+1}\|^2 \leq \|\hat{x} - x_n\|^2 - \frac{2(\gamma - 1)\rho^2}{\alpha_2\gamma}\|g^\delta - F(x_n)\|^2 \tag{2.8}$$

are satisfied, where v_n is defined by $v_n = (T[x_n](T[x_n])^* + \alpha_2 I)^{-1}(g^\delta - F(x_n))$.

Proof: We use the notation $T = T[x_n]$, $\tilde{g} = g^\delta - F(x_n)$ and $h = x_{n+1} - x_n$. From the Tikhonov equation (2.3) we obtain

$$\tilde{g} - Th = \alpha_2(TT^* + \alpha_2 I)^{-1}\tilde{g}, \quad \text{and} \quad h = T^*(TT^* + \alpha_2 I)^{-1}\tilde{g}.$$

With the abbreviation $A = (TT^* + \alpha_2 I)$ we compute

$$
\begin{aligned}
\|h - \hat{h}_n\|^2 &= \|h\|^2 - 2(A^{-1}\tilde{g}, T\hat{h}_n) + \|\hat{h}_n\|^2 \\
&= \|h\|^2 - 2(A^{-1}\tilde{g}, \tilde{g}) + 2(A^{-1}\tilde{g}, \tilde{g} - T\hat{h}_n) + \|\hat{h}_n\|^2 \\
&= \|h\|^2 - 2(T^*A^{-1}\tilde{g}, T^*A^{-1}\tilde{g}) - 2\alpha_2(A^{-1}\tilde{g}, A^{-1}\tilde{g}) \\
&\quad + 2(A^{-1}\tilde{g}, \tilde{g} - T\hat{h}_n) + \|\hat{h}_n\|^2 \\
&= -\|h\|^2 - 2\alpha_2\|A^{-1}\tilde{g}\|^2 + 2(A^{-1}\tilde{g}, \tilde{g} - T\hat{h}_n) + \|\hat{h}_n\|^2.
\end{aligned}
$$

From $\alpha_2 \geq \frac{\rho}{1-\rho}\|T\|^2$ we obtain

$$\rho\|\tilde{g}\| \leq \rho\|A\|\|A^{-1}\tilde{g}\| \leq \rho\left(\frac{1-\rho}{\rho}\alpha_2 + \alpha_2\right)\|A^{-1}\tilde{g}\| = \alpha_2\|A^{-1}\tilde{g}\|$$

and using (2.6) this implies

$$
\begin{aligned}
\|\hat{h}_n\|^2 - \|h - \hat{h}_n\|^2 &= \|h\|^2 + 2\alpha_2\|A^{-1}\tilde{g}\|^2 - 2(A^{-1}\tilde{g}, \tilde{g} - T\hat{h}_n)) \\
&\geq 2\alpha_2\|A^{-1}\tilde{g}\|^2 - 2(A^{-1}\tilde{g}, \tilde{g} - T\hat{h}_n)) \\
&\geq 2\left(\rho - \frac{\rho}{\gamma}\right)\|\tilde{g}\| \, \|A^{-1}\tilde{g}\|.
\end{aligned}
$$

9

Inequality (2.7) follows and from $\|\tilde{g}\| \leq \|A\| \|v_n\| \leq \frac{\alpha_2}{\rho}\|v_n\|$ we conclude (2.8). $\quad\square$

Applying this lemma using noise–free data implies that $\hat{h}_n$ constitutes a Cauchy sequence and the limit x^* is a solution of the problem. For polluted data the lemma is used to show that the iterates are close to the iterates from the noise–free case provided the stopping condition does not apply. This is the basic idea of the following proof of the regularisation properties.

Proof of Theorem 2.2. a) Assume $\delta = 0$ and define $\gamma = \dfrac{\rho}{\left(C + \frac{C'C''}{2\alpha_1}\right)R_0} > 1$. Then the assumption 2.5 implies

$$\|g - F(x_0) - T(x_0)\hat{h}_0\| \leq \left(C + \frac{C'C''}{2\alpha_1}\right)\|\hat{h}_0\| \, \|g - F(x_0)\| = \frac{\rho}{\gamma}\|g - F(x_0)\| \quad (2.9)$$

by $\|((F'[x_0])^*F'[x_0] + \alpha_1 I)^{-1}\| \leq 1/\alpha_1$ (where $\hat{h}_n = \hat{x} - x_n$ is defined as in the previous Lemma). Lemma 2.3 yields $\|x_1 - \hat{x}\| \leq \|x_0 - \hat{x}\|$ and by induction it follows that

$$\|x_{n+1} - \hat{x}\| \leq \|x_n - \hat{x}\| \quad \text{for all } n \in \mathbb{N}.$$

Now we continue by showing that $\hat{h}_n$ defines a Cauchy sequence in X. Let $m, n \in \mathbb{N}$ with $m > n$ and define $l \in \{n, \ldots, m\}$ by

$$\|g - F(x_l)\| \leq \|g - F(x_j)\|, \quad \text{for } j \in \{n, \ldots, m\}.$$

By the definition of $T[x_j]$ we obtain

$$\begin{aligned}
\|T[x_j]\hat{h}_l\| &= \|T[x_j]\hat{h}_j - T[x_j](x_l - x_j)\| \\[4pt]
&\leq \|g - F(x_j) - T(x_j)\hat{h}_j\| + \|F(x_l) - F(x_j) - T[x_j](x_l - x_j)\| + \|g - F(x_l)\| \\[4pt]
&\leq \left(C + \frac{C'C''}{2\alpha_1}\right)\left(\|\hat{x} - x_j\| + \|x_l - x_j\|\right)\|g - F(x_j)\| + \|g - F(x_l)\| \quad (2.10) \\[4pt]
&\leq (3\rho + 1)\|g - F(x_j)\|. \quad\quad (2.11)
\end{aligned}$$

Defining v_j as in the previous lemma implies $x_{j+1} = x_j + (T[x_j])^*v_j$. From (2.11) and the lemma we obtain

$$\begin{aligned}
|(\hat{h}_l - \hat{h}_n, \hat{h}_l)| = \left|\sum_{j=n}^{l-1}((T[x_j])^*v_j, \hat{h}_l)\right| &\leq \sum_{j=n}^{l-1}\|T[x_j]\hat{h}_l\| \, \|v_j\| \\[6pt]
&\leq (3\rho + 1)\sum_{j=n}^{l-1}\|g - F(x_j)\| \, \|v_j\| \\[6pt]
&\leq \frac{(3\rho + 1)\gamma}{2(\gamma - 1)\rho}\left(\|\hat{x} - x_n\|^2 - \|\hat{x} - x_l\|^2\right)
\end{aligned}$$

10

which yields

$$
\begin{aligned}
\|\hat{h}_l - \hat{h}_n\|^2 &= 2(\hat{h}_l - \hat{h}_n, \hat{h}_l) + \|\hat{h}_n\|^2 - \|\hat{h}_l\|^2 \\
&\leq \left(\frac{(3\rho + 1)\gamma}{2(\gamma - 1)\rho} + 1 \right) \left(\|\hat{x} - x_n\|^2 - \|\hat{x} - x_l\|^2 \right).
\end{aligned}
$$

Since the analogue result holds for $\|\hat{h}_m - \hat{h}_l\|$, we conclude

$$
\begin{aligned}
\|x_m - x_n\|^2 &= \|\hat{h}_m - \hat{h}_n\|^2 \\
&\leq 2 \left(\frac{(3\rho + 1)\gamma}{2(\gamma - 1)\rho} + 1 \right) \left(\|\hat{x} - x_n\|^2 - \|\hat{x} - x_m\|^2 \right).
\end{aligned}
$$

By monotonicity this shows convergence for $n \to \infty$.

It remains to show that the limit x^* of x_n is a solution. Lemma 2.3 yields

$$
\sum_{n=0}^{N} \|g - F(x_n)\|^2 \leq \frac{\alpha_2 \gamma}{2(\gamma - 1)\rho^2} \left(\|\hat{x} - x_0\|^2 - \|\hat{x} - x_{N+1}\|^2 \right).
$$

Thus, the sum is bounded for $N \to \infty$ and we obtain $\lim\limits_{n \to \infty} \|g - F(x_n)\| = 0$. Continuity implies $F(x^*) = g$.

b) Now assume $\delta > 0$ and define x_n to be the computed solution using exact data and x_n^δ the solution computed from noisy data g^δ. If $n(\delta) > 0$, i.e. $\|g^\delta - F(x_0)\| > \tau\delta$, we obtain from the assumption on the Taylor remainder

$$
\begin{aligned}
&\|g^\delta - F(x_0) - T(x_0)(\hat{x} - x_0)\| \\
&\leq (1 + C\|\hat{x} - x_0\|)\delta + \left(C + \frac{C'C''}{2\alpha_1} \right) \|\hat{x} - x_0\| \, \|g^\delta - F(x_0)\| \\
&\leq \left(\frac{1}{\tau} + \frac{C}{\tau} R_0 + C R_0 + \frac{1}{2} \frac{C'C''}{\alpha_1} R_0 \right) \|g^\delta - F(x_0)\|.
\end{aligned}
$$

The choice of τ shows that $\gamma := \dfrac{\rho\tau}{1 + \left((1+\tau)C + \frac{C'C''\tau}{2\alpha_1} \right) \|\hat{x} - x_0\|} > 1$ and thus Lemma 2.3 implies

$$
\|\hat{x} - x_n^\delta\|^2 \leq \|\hat{x} - x_{n-1}^\delta\|^2 - \frac{2(\gamma - 1)\rho^2}{\alpha_2 \gamma} \|g^\delta - F(x_{n-1}^\delta)\|^2. \tag{2.12}
$$

The definition of $n(\delta)$ leads to

$$
n(\delta)\tau^2\delta^2 \leq \sum_{j=0}^{n(\delta)-1} \|g^\delta - F(x_j^\delta)\|^2 \leq \frac{\alpha_2 \gamma}{2(\gamma - 1)\rho^2} \|\hat{x} - x_0\|^2,
$$

11

which shows that $n(\delta)$ is finite.

We can easily show by induction the continuous dependence $x_n^\delta \to x_n$ for fixed $n \in \mathbb{N}$ and $\delta \to 0$. Now two cases have to be considered.

First, we assume that the sequence $n(\delta)$ has a finite accumulation point, i.e. without loss of generality $n(\delta) = N$ for $\delta \leq \delta_N$. Then follows

$$\lim_{\delta \to 0} x_{n(\delta)}^\delta = \lim_{\delta \to 0} x_N^\delta = x_N.$$

Moreover, by the definition of $n(\delta)$ we know $\|g^\delta - F(x_{n(\delta)}^\delta)\| \leq \tau\delta$. Taking the limit $\delta \to 0$ implies $F(x_N) = g$, and we conclude $\lim_{\delta \to 0} x_{n(\delta)}^\delta = x_N = x^*$.

In the second case we assume that a subsequence of $n(\delta)$ is unbounded and obtain from (2.12) that $\|x^* - x_{n(\delta)}^\delta\| \leq \|x^* - x_m^\delta\|$ for all $m \leq n(\delta)$. For $\epsilon > 0$, choose m sufficiently large such that $\|x^* - x_m\| \leq \epsilon$. Without loss of generality we choose $\delta_m > 0$ with $m \leq n(\delta)$ and $\|x_m - x_m^\delta\| \leq \epsilon$ for all $\delta < \delta_m$ and we conclude

$$\|x^* - x_{n(\delta)}^\delta\| \leq \|x^* - x_m\| + \|x_m - x_m^\delta\| \leq 2\epsilon.$$

Combining both cases proves the assertion. $\square$

3   A parameter identification problem

Consider the problem of determining the unknown conductivity $a(x)$ from

$$-(au')' = f_j, \quad 0 < x < 1, \qquad u(0) = u(1) = 0. \tag{3.1}$$

We assume the input source functions are of the form $f_j(x) = \sin j\pi x$, $j \in \mathbb{N}$, and for a given value of a we denote by $u_j(x; a)$ the solution of the direct problem solving (3.1). The additional data needed to recover a is $g_j = u_j'(0; a)$.

Let $F : U \to \tilde{l}^2$ be the map $a \mapsto \{u_j'(0; a)\}_{j=1}^\infty$, where $U \subseteq \{a \in H^1[0, 1] : a(t) \geq \epsilon > 0, \ a(0) = 1 = a(1)\}$ and $\tilde{l}^2$ denotes the Hilbert space of sequences $\{c_j\}_{j=0}^\infty$ with $\sum_{j=1}^\infty \frac{|c_j|^2}{j^2} < \infty$. As in example 11.1 of [5], we observe that the condition (2.5) is satisfied which ensures convergence of the second degree method by Theorem 2.2.

The function $u_j(x; 1)$ can be explicitly computed, $u_j(x, 1) = \sin(j\pi x)/j^2\pi^2$. This allows us to compute the required derivatives for the first and second degree schemes utilising derivative operators held fixed at $a = 1$.

The Fréchet derivative $(F'[a]h)_j$ is represented by the value of $\{v_j(0; a; h)\}$ where v_j solves

$$-(av_j')' = (hu_j')', \quad 0 < x < 1, \qquad v_j(0) = v_j(1) = 0, \tag{3.2}$$

(see [3]). Integrating this equation from 0 to x gives

$$-a(x)v_j'(x) + a(0)v_j'(0) = h(x)\cos(j\pi x)/j\pi - h(0)/j\pi$$

and so, in the case $a = 1$ and $h = \sin(k\pi x)$ by once again integrating between $x = 0$ and $x = 1$ and using the boundary conditions on v_j we see that

$$(F'[1]h)_j = v_j'(0; 1; h) = \frac{1}{j\pi} \int_0^1 h(x) \cos(j\pi x)\, dx$$

$$= \frac{1}{j\pi} \int_0^1 \sin(k\pi x) \cos(j\pi x)\, dx =: c_{jk}$$

The values c_{jk} are computed to be

$$c_{jk} = \frac{k(1 - (-1)^{k+j})}{j\pi^2(k^2 - j^2)}, \quad j \neq k \quad \text{and} \quad c_{jj} = 0.$$

Now the predictor $\tilde{h}$ is taken to be the solution of the equations

$$(F'[1]\tilde{h})_j = g_j - u_j'(0; a)$$

where $\tilde{h}(x)$ has the series expansion $\tilde{h}(x) = \sum_{n=1}^N \tilde{h}_n^s \sin(n\pi x)$. This can be solved uniquely for the sequence $\{\tilde{h}_n^s\}$ to obtain

$$v_j'(x; 1; \tilde{h}) = -\frac{1}{j\pi}\tilde{h}(x)\cos(j\pi x) + g_j - u_j'(0; a). \tag{3.3}$$

A further differentiation of (3.1) with respect to a at $a = 1$ gives

$$-w_j''(x) = (\tilde{h}v_j'(x; 1; h) + hv_j'(x; 1; \tilde{h}))' \quad 0 < x < 1, \qquad w_j(0) = w_j(1) = 0 \tag{3.4}$$

(see [5]) and the value of $w_j'(0; 1; \tilde{h}, h)$ represent the second derivative of u. Integrating (3.4) in x leads to

$$-w_j'(x) + w_j'(0) = \tilde{h}(x)v_j'(x; 1; h) + h(x)v_j'(x; 1; \tilde{h}).$$

If we take $h(x) = \sin k\pi x$, then by a further integration from $x = 0$ to $x = 1$ using the boundary conditions on w_j, (3.3) and the series expansion for $\tilde{h}$ it follows that

$$w_j'(0; 1; \tilde{h}, h) = \frac{1}{j\pi}\Big(\int_0^1 \tilde{h}(x)\, dx \int_0^1 h(x) \cos j\pi x\, dx + \int_0^1 h(x)\, dx \int_0^1 \tilde{h}(x) \cos j\pi x\, dx$$

$$-2 \int_0^1 \tilde{h}(x)h(x) \cos j\pi x\, dx \Big)$$

$$= \sum_{m=1}^N \tilde{h}_m^s \Big(\frac{1 - (-1)^m}{m\pi} c_{jk} - \frac{2}{j\pi} \int_0^1 \sin(m\pi x) \sin(k\pi x) \cos(j\pi x)\, dx \Big)$$

$$+\frac{1 - (-1)^k}{k\pi}(g_j - u_j'(0; 1))$$

13

$$= -\frac{1}{2j\pi}\left[\tilde{h}^s_{k+j} + \text{sgn}(k-j)\tilde{h}^s_{|k-j|}\right] + c_{jk}\sum_{m=1}^{N}\frac{1-(-1)^m}{m\pi}\tilde{h}^s_m$$

$$+\frac{1-(-1)^k}{k\pi}(g_j - u'_j(0;a)).$$

The first degree model with frozen derivatives at $a = 1$ now requires at each iteration step the solution of the equation

$$J\tilde{h}^s = (g_j - u'_j(0,a_n))_{j=1...N} \tag{3.5}$$

with the update $a_{n+1} = a_n + \tilde{h}$, where $J_{jk} = v'_j(0;1;\sin k\pi x)$. The second degree method additionally seeks the solution $h^s = (h_k)_{k=1...N}$ of

$$\left(J + \frac{1}{2}H(\tilde{h})\right)h = (g_j - u'_j(0,a_n))_{j=1...N} \tag{3.6}$$

where $H_{jk}(\tilde{h}^s) = w'_j(0;1;\tilde{h};\sin k\pi x)$. The next iteration is then defined by $a_{n+1}(x) = a_n(x) + \sum_{k=1}^{N} h^s_k \sin k\pi x$.

Table 1: Comparison of three Iteration Schemes.

n	Frozen Newton $\|a_n - a\|$	R_n	Full Newton $\|a_n - a\|$	R_n	Second degree $\|a_n - a\|$	R_n
0	1.2484	0.08042	1.2484	0.08042	1.2484	0.08042
1	0.8409	0.01693	0.8414	0.01687	0.2873	0.00432
2	0.6649	0.01702	0.3916	0.00524	0.2568	0.00067
3	0.5686	0.00864	0.2220	0.00053	0.2304	0.00074
4	0.4807	0.00934	0.2109	0.00001	0.2222	0.00028
5	0.4289	0.00488	0.2110	0.00000	0.2167	0.00018
6	0.3757	0.00568	0.2110	0.00000	0.2142	0.00009
7	0.3442	0.00290	0.2110	0.00000	0.2127	0.00005
8	0.3097	0.00359	0.2110	0.00000	0.2120	0.00002
9	0.2903	0.00175	0.2110	0.00000	0.2115	0.00001
10	0.2672	0.00233	0.2110	0.00000	0.2113	0.00000

We implemented frozen versions of both the Newton and the second degree method, regularising equation (3.5) and (3.6) only by restricting the number of basis elements to $N = 10$, (although we could also have used Tikhonov's method). Taking the actual solution $a(x) = (1 + x\sin(2\pi x))^{-1}$, a run is presented in Table 1, which shows the fast convergence of the frozen second degree method in the beginning even compared to the full Newton scheme. In this example we used the starting guess $a = 1$ and computed numerical data using a Runge-Kutta scheme on a course mesh, estimating that resulting output was accurate to within about 1%. The table records the L^2 norm error $\|a_n - a\|_2$ and the residual $R_n := \|u'_j(0;a_n) - d\|_2$ for each of these methods. Figure 1 shows the final reconstruction obtained, where the dotted line corresponds to the exact solution. This was essentially identical in all three methods.
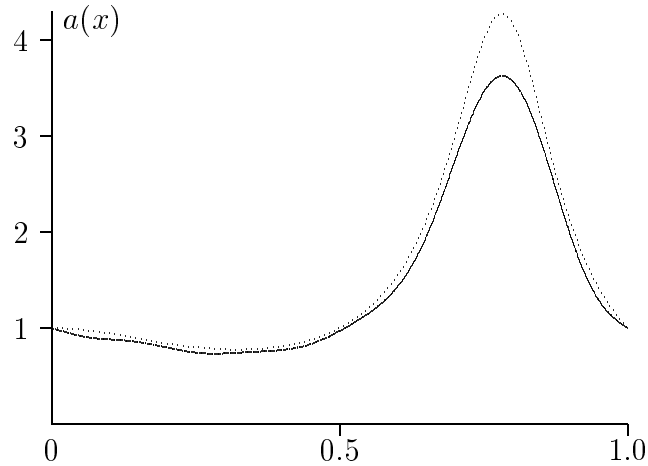
Figure 1: Identification of $a(x) = (1 + x \sin(2\pi x))^{-1}$

4 The inverse Sturm-Liouville problem

A classical inverse problem is the reconstruction of the potential $q(x)$ from the equation $-y'' + qy = \lambda_n y$ given the eigenvalues with respect to certain boundary conditions [16]. The celebrated result of Borg showed that if, say, $y(0) = 0$ then the complete sequence of eigenvalues $\{\lambda_n\}$ and $\{\mu_n\}$ corresponding to the two boundary conditions $y(1) = 0$ and $y'(1) = 0$ suffice to determine the potential q. Further, in the symmetric case $q(x) = q(1 - x)$ the single sequence of Dirichlet eigenvalues $\{\lambda_n\}$ suffices. We shall consider this latter problem here using an extension of the methods of [21] to incorporate the second degree method.

From the results of [21] a reformulation of the inverse Sturm-Liouville problem is to identify the potential q as the solution of an overposed boundary value problem for the Goursat problem

$$u_{tt} - u_{xx} + q(x)u = 0 \quad \text{for } 0 < t \le x < 1 \tag{4.1}$$

with

$$u(x, 0) = 0, \quad u(x, x) = \frac{1}{2} \int_0^x q(s) \, ds \tag{4.2}$$

and subject to the data on $x = 1$, $u(1, t) = g(t)$. We shall denote the solution of (4.1) and (4.2) for a given q by $u(x, t; q)$. As shown in [21], the function g can be constructed from the eigenvalue data and is the unique solution of the integral equation $\int_0^1 g(t) \sin \sqrt{\lambda_n} t \, dt = - \sin \sqrt{\lambda_n}$ for each $n = 1, 2, \ldots$. The asymptotics of the λ_n guarantees that this is easily and stably solved for $g(t)$. Some preliminary analysis shows also that one can reduce the problem to the situation of a potential

with zero mean, that is $\int_0^1 q = 0$ and thus the function $g(t)$ can be taken to be odd about the origin and $g(1) = 0$.

We now must consider the mapping taking potentials onto the data. In fact it is more convenient to choose the map such that $F : L^2[0,1] \to L^2[0,1]$ is defined by $F[q](t) = u_t(1,t) = g'(t)$.

One can compute the derivative of the map F, and in particular the derivative held at the zero potential $F'[0]h$ gives valuable insight to the problem. In fact, as shown in [21], the frozen Newton method gives remarkably good results. An easy calculation shows that $F'[q]h$ must be given as the solution $u'(x,t;q;h)$ of the Goursat problem

$$u'_{tt} - u'_{xx} + q(x)u' = -hu \quad \text{for } 0 < t \le x < 1$$

with

$$u'(x,0) = 0, \quad u'(x,x) = \frac{1}{2} \int_0^x h(s) \, ds$$

evaluated on the line $x = 1$. Now if $q = 0$ then it follows from uniqueness of solutions to (4.1) and (4.2) that $u(x,t;0)$ must be identically zero. Thus $u'(x,t;0;h)$ must satisfy the equation

$$u'_{tt} - u'_{xx} = 0 \quad \text{for } 0 < t \le x < 1$$

and this is easily computed in closed form as

$$u'(x,t;0;h) = \frac{1}{2} \int_{\frac{x-t}{2}}^{\frac{x+t}{2}} h(s) \, ds. \tag{4.3}$$

Now it follows that $u'_t(1,t;0;h) = \frac{1}{2}(h(\frac{1+t}{2}) + h(\frac{1-t}{2}))$ and the symmetry assumption on our potentials allows us to conclude that $u'_t(1,t;0;h) = h(\frac{1+t}{2})$.

The frozen Newton scheme now becomes

$$q_{n+1}(s) - q_n(s) = g'(2s-1) - u_t(1,2s-1;q_n) \tag{4.4}$$

for $s \in [0,1]$ and where $u(1,t;q)$ is extended as an odd function.

In a similar fashion we can compute the second derivative $F''[0](h_1, h_2)$ as the solution of the Goursat problem

$$u''_{tt} - u''_{xx} = f(x,t) \quad \text{for } 0 < t \le x < 1 \tag{4.5}$$

with $u''(x,0) = 0$ and $u''(x,x) = 0$ and the function f equal to

$$f(x,t) = -h_1(x)u'(x,t;0;h_2) - h_2(x)u'(x,t;0;h_1). \tag{4.6}$$

We must compute $u''_t(1,t;0;h_1,h_2)$. In view of our representation (4.3) for u' we can write (4.6) in the form

$$2f(x,t) = -h_1(x) \int_{\frac{x-t}{2}}^{\frac{x+t}{2}} h_2(s) \, ds - h_2(x) \int_{\frac{x-t}{2}}^{\frac{x+t}{2}} h_1(s) \, ds. \tag{4.7}$$

16

and in this case the solution of (4.5) at $x = 1$ becomes

$$2u''(1, t; 0; h_1, h_2) = \int_{\frac{1-t}{2}}^{1-t} \int_{1-t-s}^{s} f(s, r) \, ds dr + \int_{1-t}^{\frac{1+t}{2}} \int_{s+t-1}^{s} f(s, r) \, ds dr$$

$$+ \int_{\frac{1+t}{2}}^{1} \int_{s+t-1}^{1+t-s} f(s, r) \, ds dr.$$

This gives

$$2u_t''(1, t; 0; h_1, h_2) = -\int_0^{1-t} f(1-t, r) \, dr - \int_{\frac{1-t}{2}}^{1-t} f(s, s+t-1) \, ds$$

$$+ \frac{1}{2} \int_{\frac{3t-1}{2}}^{\frac{1+t}{2}} f(\tfrac{1+t}{2}, r) \, dr + \int_0^{1-t} f(1-t, r) \, dr$$

$$- \int_{1-t}^{\frac{1+t}{2}} f(s, s+t-1) \, ds - \frac{1}{2} \int_{\frac{3t-1}{2}}^{\frac{1+t}{2}} f(\tfrac{1+t}{2}, r) \, dr$$

$$+ \int_{\frac{1+t}{2}}^{1} f(s, 1+t-s) \, ds - \int_{\frac{1+t}{2}}^{1} f(s, s+t-1) \, ds.$$

Using $f(x, -t) = -f(x, t)$ and both cancelling and combining terms yields

$$2u_t''(1, t; 0; h_1, h_2) = -\int_{\frac{1-t}{2}}^{1} f(s, s+t-1) \, ds + \int_{\frac{1+t}{2}}^{1} f(s, 1-s+t) \, ds$$

$$=: \quad \Phi(t) + \Phi(-t).$$

In implementing the predictor step as a frozen Newton update, we must solve an equation of the form $h_1(t) = \tilde{g}'(t)$ at each iteration where $\tilde{g}(t) = \frac{1}{2}(g(2t-1) - u(1, 2t-1; q_n))$, which implies $\tilde{g}(t) = -\tilde{g}(1-t)$. In the initial approximation we have simply $\tilde{g}(t) = g(2t-1)/2$. Using (4.7) and setting $h = h_2$ we find

$$\Phi(t) = \int_{\frac{1-t}{2}}^{1} h(s) \int_{\frac{1-t}{2}}^{s - \frac{1-t}{2}} \tilde{g}'(r) \, dr \, ds + \int_{\frac{1-t}{2}}^{1} \tilde{g}'(s) \int_{\frac{1-t}{2}}^{s - \frac{1-t}{2}} h(r) \, dr \, ds$$

$$= \int_{\frac{1-t}{2}}^{1} \left[\tilde{g}(s - \tfrac{1-t}{2}) - \tilde{g}(\tfrac{1-t}{2}) \right] h(s) \, ds$$

$$+ \left(\tilde{g}(s) \int_{\frac{1-t}{2}}^{s - \frac{1-t}{2}} h(r) \, dr \right) \Big|_{s = \frac{1-t}{2}}^{s=1} - \int_{\frac{1-t}{2}}^{1} \tilde{g}(s) h(s - \tfrac{1-t}{2}) \, ds$$

$$= \int_{\frac{1-t}{2}}^{1} [\tilde{g}(s - \tfrac{1-t}{2}) - \tilde{g}(\tfrac{1-t}{2})] h(s) \, ds$$

17

$$+ \tilde{g}(\tfrac{1-t}{2}) \int_0^{\frac{1-t}{2}} h(s)\,ds - \int_0^{\frac{1+t}{2}} \tilde{g}(s + \tfrac{1-t}{2}) h(s)\,ds$$

where we have used the fact that $\tilde{g}(1) = 0$. Now the assumed symmetry of h yields

$$\Phi(t) = \int_{\frac{1-t}{2}}^1 \tilde{g}(s - \tfrac{1-t}{2}) h(s)\,ds - \int_0^{\frac{1+t}{2}} \tilde{g}(s + \tfrac{1-t}{2}) h(s)\,ds$$

$$= \int_0^{\frac{1+t}{2}} \left(\tilde{g}(\tfrac{1+t}{2} - s) - \tilde{g}(s + \tfrac{1-t}{2}) \right) h(s)\,ds.$$

Then

$$2u_t''(1,t;0;h_1,h_2) = \int_0^{\frac{1-t}{2}} \left(\tilde{g}(\tfrac{1+t}{2} - s)\tilde{g}(\tfrac{1-t}{2} - s) - \tilde{g}(s + \tfrac{1-t}{2}) - \tilde{g}(s + \tfrac{1+t}{2}) \right) h(s)\,ds$$

$$+ \int_{\frac{1-t}{2}}^{\frac{1+t}{2}} \left(\tilde{g}(\tfrac{1+t}{2} - s) - \tilde{g}(s + \tfrac{1-t}{2}) \right) h(s)\,ds.$$

Using the symmetry of the function $\tilde{g}$ the last integral must be zero. Thus our corrector formula when the derivatives are evaluated at $q = 0$ is given by the Volterra equation

$$\tilde{g}'(t) = h(t) + \frac{1}{4} \int_0^t \left(\tilde{g}(1 - t - s) + \tilde{g}(t - s) - \tilde{g}(1 + s - t) - \tilde{g}(t + s) \right) h(s)\,ds \quad (4.8)$$

where $\tilde{g}(t) = \frac{1}{2}(g(2t - 1) - u(1, 2t - 1; q))$ and we use an odd extension of both g and $u(1, .; q)$.

Convergence of the Newton scheme based on (4.1) and (4.2) is quite rapid, even with the frozen derivative version (4.4) is used [21]. As an illustration of the ability of the second degree method to improve an initial approximation we show in figure 2 the results of using the predictor (a Newton step from $q = 0$) and then the corrected version (the second degree method).

The actual potential is the function $q(x) = 50\sin(3\pi x) * e^{-5x}$, $x \in [0, 1/2]$, and the reconstruction uses the first 10 Dirichlet eigenvalues. These eigenvalues were quite accurate and the information contained in them was sufficient to recover the actual potential (shown as a dotted curve) to within about one percent as measured in the L^2 norm. The figure on the left was obtained by using formula (4.4) with $q_0 = 0$. Note that since $u(1, t; 0) = 0$ this means we have simply $q_1(x) = g'(2x - 1)$. The figure on the right used (4.8) where, since $q_0 = 0$, $\tilde{g}(x) = \frac{1}{2}g(x)$. This Volterra integral equation of the second kind is easily inverted to recover h and hence the approximation q_1.

The initial updated approximation is remarkably good using only the purely predicted step – the ratio $(1 - \|q_1 - \tilde{q}\|_2)/\|\tilde{q}\|_2 = 0.83$. In other words we get about

18

<div align="center">Using the predictor only   Using the corrector</div>
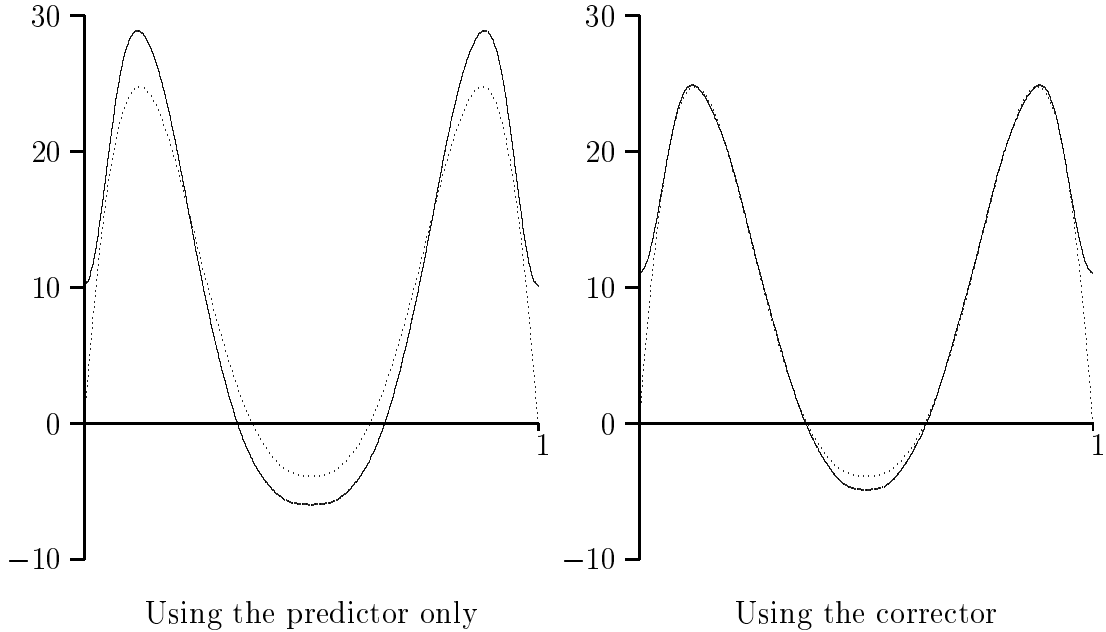
<div align="center">Figure 2: Reconstruction after a single step from $q = 0$.</div>

83% of the reconstruction in the first step without having even to compute a value of the direct map $u(1, x; q)$. However the superiority of the correcting term in (4.8) is quite apparent for now $(1 - \|q_1 - \tilde{q}\|_2)/\|\tilde{q}\|_2 = 0.97$ and this is quite close to the final reconstruction possible from the limited data.

The point to be made here is that for the very little additional cost of solving (4.8), as opposed to taking only the first term Neumann approximation of the Volterra equation, we gain a substantially improved reconstruction This extra effort is certainly much less than would be necessary to perform a direct computation to obtain $u(1, t; q)$. In fact in most cases, a single further iteration of the second degree method will suffice. That is, we can effectively reconstruct the potential q from eigenvalue data by solving the Volterra equation (4.8) twice and the Goursat problem (4.1), (4.2) once.

5   An inverse potential problem

Let $\Omega \subseteq \mathbb{R}^N$ be a smooth bounded domain with an included subdomain $D \subseteq \Omega$. Then the problem consist in the recovery of the shape of D, i.e. the boundary ∂D, from the knowledge of the Neumann boundary values

$$F(\partial D) = \frac{\partial u}{\partial \nu}\Big|_{\partial \Omega}$$

<div align="center">19</div>

of the solution u of the homogeneous Dirichlet problem

$$\triangle u = \chi_D \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega, \tag{5.1}$$

where χ_D denotes the characteristic function. This inverse problem is well known to be severely ill–posed. The existence of a unique solution $u \in H^2(\Omega)$ is well established (see e.g. [6]) and also the inverse problem is discussed by several authors (see e.g. [14], [12], [13]).

For our numerical purpose we assume Ω to be the unit disk in $\mathbb{R}^2$ and suppose D is starlike with respect to the origin, i.e. ∂D is a closed curve represented by

$$x(t) = q(t) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad t \in [0, 2\pi],$$

with $0 < q(t) < 1$ sufficiently smooth. In particular this set of admissible boundaries ensures uniqueness of the inverse problem. The proof is due to Novikov and can be found in [14].

An application of Green's second formula yields

$$\int_D v\, dx = \int_\Omega \triangle u\, v\, dx = \int_{\partial\Omega} \frac{\partial u}{\partial \nu} v\, ds \tag{5.2}$$

for all harmonic functions v in Ω. Choosing polar coordinates we compute from the identity (5.2) by $v \in \{1, r^n \cos nt, r^n \sin nt\}$ the Fourier coefficients of $\frac{\partial u}{\partial \nu}$ on $\partial\Omega$,

$$(F(q))_n^{c,s} = \frac{1}{\pi} \int_0^{2\pi} \frac{\partial u(x(t))}{\partial \nu} \left\{ \begin{array}{c} \cos nt \\ \sin nt \end{array} \right\} dt = \frac{1}{(n+2)\pi} \int_0^{2\pi} q^{n+2}(t) \left\{ \begin{array}{c} \cos nt \\ \sin nt \end{array} \right\} dt. \tag{5.3}$$

This equation also allows the determination of the Fourier coefficients of the Fréchet derivative of F with respect to variations of q by

$$(F'[q]h)_n^{c,s} = \frac{1}{\pi} \int_0^{2\pi} q^{n+1}(t) h(t) \left\{ \begin{array}{c} \cos nt \\ \sin nt \end{array} \right\} dt, \tag{5.4}$$

and we continue in computing the second derivative to

$$(F''(\tilde{h}, h))_n^{c,s} = \frac{(n+1)}{\pi} \int_0^{2\pi} q^n(t)\, h(t)\, \tilde{h}(t) \left\{ \begin{array}{c} \cos nt \\ \sin nt \end{array} \right\} dt. \tag{5.5}$$

Equation (5.4) suggests a simple and fast frozen Newton scheme for the numerical solution of the inverse problem can be obtained by fixing the derivative at a circle of radius r_0 and computing a new update $\tilde{q}_{j+1}$ from

$$F'[q_{r_0}](\tilde{q}_{j+1} - q_j) = g - F(q_j)$$

with measured data $g \approx \frac{\partial u}{\partial \nu}$ on $\partial\Omega$ (see also [12]).

We consider boundaries in a finite dimensional subspace represented by trigonometric polynomials, $h(t) = \frac{1}{2}h_0 + \sum_{n=1}^{N} h_n^c \cos nt + h_n^s \sin nt$. This leads to a finite dimensional linear equation

$$J\vec{\tilde{h}} = g - F(q_j) \qquad (5.6)$$

for the $2N+1$ Fourier coefficients of $\tilde{h} = \tilde{q}_{j+1} - q_j$ collected in $\vec{\tilde{h}}$. The Jacobian matrix J turns out to be diagonal with $J_{nn} = r_0^{n+1}$ if $0 \le n \le N$ and $J_{nn} = r_0^{n-N+1}$ if $N + 1 \le n \le 2N$.

We can also consider the second derivative evaluated at this circle q_{r_0}. We use $\tilde{h}$ computed by the frozen Newton step (5.6) as the predictor and obtain from equation (5.5) the Fourier coefficients

$$(F''(\tilde{h}, h))_n^{c,s} = \frac{(n+1)r_0^n}{2}\tilde{h}_0 \left\{ \begin{array}{c} h_0 \\ h_n^c \\ h_n^s \end{array} \right\}$$

$$+ \frac{(n+1)r_0^n}{2\pi} \sum_{j=1}^{N} \tilde{h}_j^c \int_0^{2\pi} h(t) \left\{ \begin{array}{c} \cos jt \\ \cos((n+j)t) + \cos((n-j)t) \\ \sin((n+j)t) + \sin((n-j)t) \end{array} \right\} dt$$

$$+ \tilde{h}_j^s \int_0^{2\pi} h(t) \left\{ \begin{array}{c} \sin jt \\ \sin((n+j)t) - \sin((n-j)t) \\ \cos((n-j)t) - \cos((n+j)t) \end{array} \right\} dt$$

$$= \left\{ \begin{array}{l} \dfrac{1}{2}\sum_{j=0}^{N} \tilde{h}_j^c h_j^c + \tilde{h}_j^s h_j^s \\[2ex] \dfrac{(n+1)r_0^n}{2}\left(\tilde{h}_n^c h_0 + \sum_{j=1}^{N}\left(\tilde{h}_{n+j}^c + \tilde{h}_{|n-j|}^c \right)h_j^c + \left(\tilde{h}_{n+j}^s - \mathrm{sgn}(n-j)\tilde{h}_{|n-j|}^s \right)h_j^s \right) \\[2ex] \dfrac{(n+1)r_0^n}{2}\left(\tilde{h}_n^s h_0 + \sum_{j=1}^{N}\left(\tilde{h}_{n+j}^s + \mathrm{sgn}(n-j)\tilde{h}_{|n-j|}^s \right)h_j^c + \left(\tilde{h}_{|n-j|}^c - \tilde{h}_{n+j}^c \right)h_j^s \right) \end{array} \right\}.$$

Thus, if we define the Hessian matrix by the following summation of blocks of Toeplitz and Hankel matrices

$$H = D \left(\begin{array}{c|c|c} \tilde{h}_0 & \tilde{h}_j^c & \tilde{h}_j^s \\ \hline \tilde{h}_n^c & \tilde{h}_{|n-j|}^c + \tilde{h}_{n+j}^c & \tilde{h}_{n+j}^s - \mathrm{sgn}(n-j)\tilde{h}_{|n-j|}^s \\ \hline \tilde{h}_n^s & \tilde{h}_{n+j}^s + \mathrm{sgn}(n-j)\tilde{h}_{|n-j|}^s & \tilde{h}_{|n-j|}^c - \tilde{h}_{n+j}^c \end{array} \right),$$
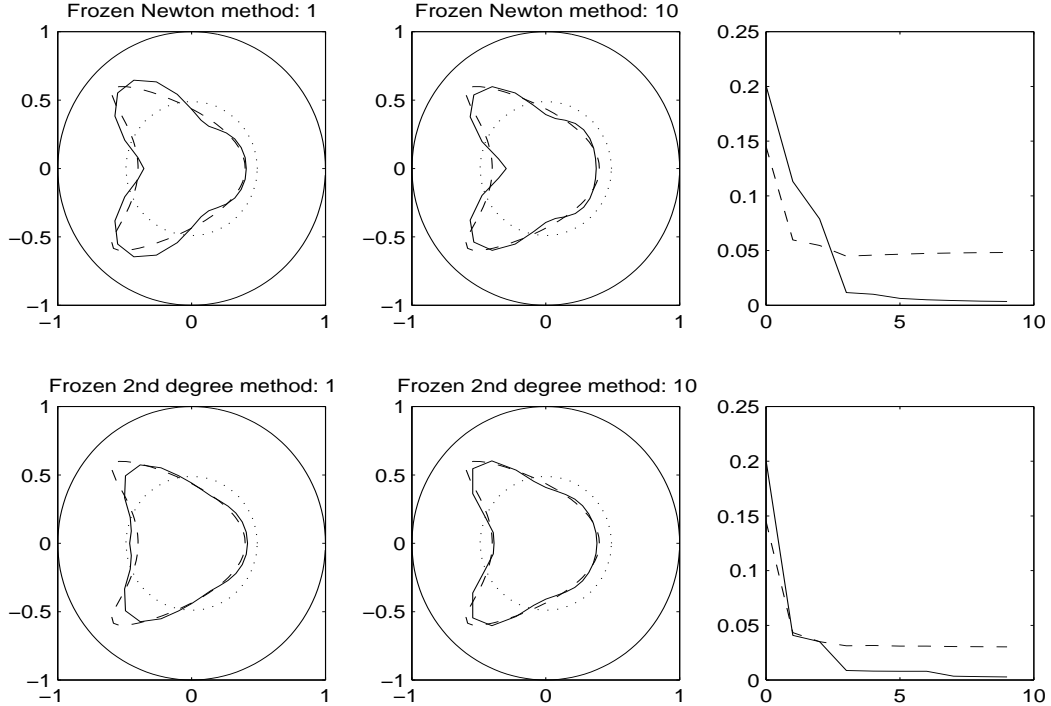
Figure 3: Reconstructions from noise–free data

where D denotes a diagonal matrix with $D_{nn} = \frac{n+1}{2}\,r_0^n$ if $0 \le n \le N$ and $D_{nn} = \frac{n-N+1}{2}\,r_0^{n-N}$ if $N+1 \le n \le 2N$, the frozen corrector step is given by the solution of the linear equation

$$\left(J + \frac{1}{2}H\right)\vec{h} = g - F(q_j). \tag{5.7}$$

A perturbation argument shows that this equation is uniquely solvable, if the predictor $\tilde{h}$ is sufficiently small. Since $\tilde{h}$ is defined by (5.6) we obtain the behaviour $H_{nj} = \mathcal{O}(r_0^{-j-1})$ with respect to the radius r_0. Thus, applying H on $\tilde{h}$ damps the j mode by this factor. The same asymptotic formula holds for the frozen Newton step in (5.6). Therefore we cannot expect a significant change in the condition number with respect to r_0 and N by adding the Hessian matrix in the second degree method.

The numerical experiments presented below used artificial data, which were computed by a boundary integral equation approach to the direct problem (see [17]). In modelling noise we added a random vector of 5% of the L^2-norm approximation of these highly accurate data. Fortunately, a good starting guess (represented by the dotted curve) is available using $v = 1$ in (5.2), which shows that the volume of D is given by the mean of the data. Therefore we used as a starting curve a circle bounding a disk of this volume centered at the origin.

Without any additional regularisation the simple frozen iteration schemes (5.6) and
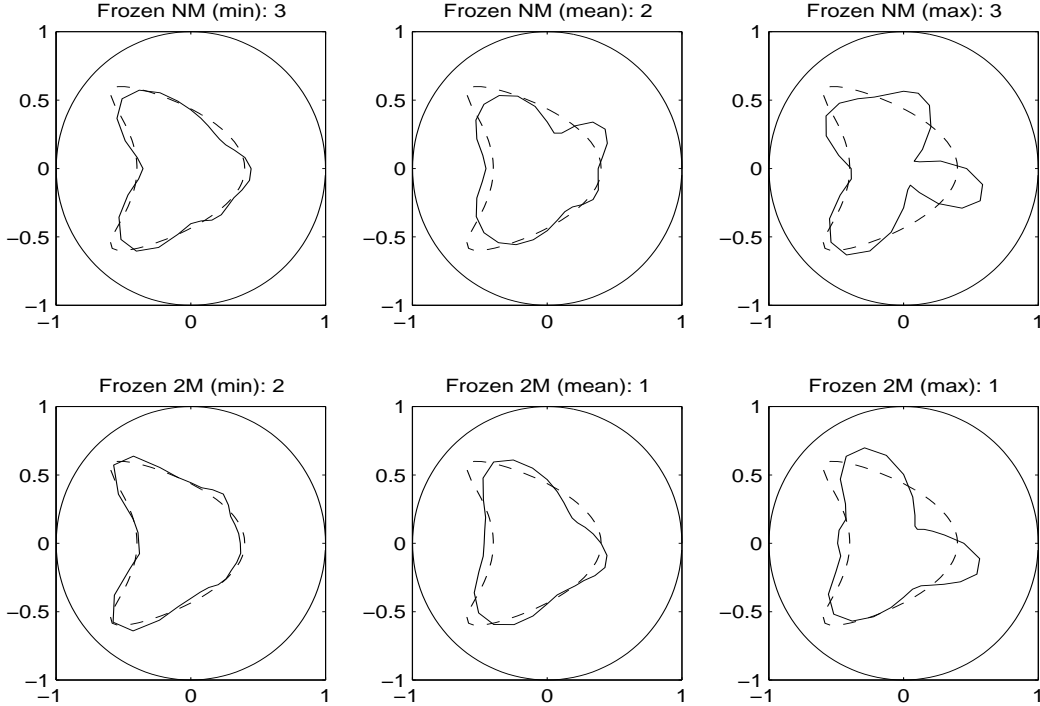
22

Figure 4: Reconstructions from noisy data

(5.7) allow the use of frequencies up to $N \leq 4$, which is only sufficient to represent objects D which do not contain much structure. Therefore we consider Tikhonov's regularisation of (5.6) and (5.7) as suggested in section 2 using $N = 10$ in all numerical experiments presented below.

Figure 3 shows the reconstruction of a kite shaped object from noise–free data using the regularisation parameters $\alpha = 10^{-4}$ for the frozen Newton scheme and $\alpha_1 = 10^{-5}, \alpha_2 = 10^{-4}$ for the second degree method. We observe a good reconstruction by the second degree method on the first iteration step. The figures show these reconstruction after the first and the 10th iteration. The residual error during iteration (the full line) and the error in the reconstruction with respect to the true boundary curve (dashed line) are given in the figure on the right.

The differences in the performance of the two schemes, using the predictor step only (frozen Newton method) and the frozen second degree approach, decreases if we consider much smaller objects. The reasons seems to be that the higher regularisation which is then required cancels the positive effect of considering the second derivative. This is also observed in case of data polluted by high levels of noise. In Figure 4 we show the best, the mean, and the worst reconstruction with respect to the error in the boundary curves from a hundred different runs with 5% random noise in the data and the same regularisation parameters as before. The figures show the

reconstructions, when the stopping condition applies (here $\|g - F(q_j)\|/\|g\| \leq 0.07$ was chosen). The performance in the worst case suggests more regularisation. Increasing the Tikhonov parameters lead to smaller differences in the two methods and to more stable but poorer results. Decreasing the Tikhonov parameters and increasing the stopping value can often lead to quite good reconstructions, but this highly dependent on the actual data set used and on average the results show more instability.

The same performance with respect to appropriate regularisation parameters occurs for the full iteration methods using the derivatives at q_j. A more significant difference in using the full second degree method is observed in an inverse scattering problem considered in the following section.

6   An inverse scattering problem

As a final example we consider the scattering of time–harmonic incident plane acoustic waves by a sound–soft obstacle D in two dimensions. Let $D \subseteq \mathbb{R}^2$ be a bounded domain with smooth boundary ∂D and connected complement $\mathbb{R}^2 \backslash \overline{D}$. Assuming a wave number $k > 0$ the scattering problem is modeled by a solution $u \in C^2(\mathbb{R}^2 \backslash \overline{D}) \cap C(\mathbb{R}^2 \backslash D)$ of the Helmholtz equation

$$\triangle u + k^2 u = 0 \quad \text{in } \mathbb{R}^2 \backslash \overline{D}. \tag{6.1}$$

In $\mathbb{R}^2 \backslash \overline{D}$ we have the decomposition $u = u^s + u^i$, where $u^i(x) = e^{ikx \cdot d}$ is the incident plane wave with direction $d \in S^1$, and u^s denotes the radiating field satisfying Sommerfeld's condition,

$$\lim_{r=|x|\to\infty} r^{-\frac{1}{2}} \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0 \tag{6.2}$$

uniformly in the directions $x/|x| \in S^1$. Supposing a sound–soft obstacle leads to the Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial D. \tag{6.3}$$

For a detailed analysis of this problem including existence of a unique solution of (6.1)–(6.3) by boundary integral equations we refer to [4].

The radiation condition implies the asymptotic behaviour

$$u^s(x) = \frac{e^{ik|x|}}{\sqrt{|x|}} \left\{ u_\infty(\hat{x}) + O\left(\frac{1}{|x|}\right) \right\}$$

uniformly in $\hat{x} = x/|x|$. The function u_∞ is the far field pattern or scattering amplitude of u^s. The inverse problem under consideration consists in determining the shape of D from measurements of the far field pattern. We define the far field

operator $F : X \rightarrow L^2(S^1)$ which maps a set X of admissible boundaries on the far field patterns of the scattering problem for a fixed wave number k and an incident direction $d \in S^1$. The inverse problem can be written as the nonlinear and ill–posed equation

$$F(\partial D) = u_\infty.$$

Recent investigations show that the linearisation of this operator using a representation of the derivative of F leads to efficient iterative regularisation schemes solving the inverse scattering problem (see e.g. [9, 13, 15, 18, 19]). All of these methods are based on a representation of the Fréchet derivative in the sense of a bounded linear operator $F'[\partial D]$ with

$$\lim_{\|h\|_{C^1} \rightarrow 0} \frac{1}{\|h\|_{C^1}} \|F(\partial D_h) - F(\partial D) - F'[\partial D]h\| = 0$$

for variations $h \in (C^1(\partial D))^2$ of the boundary. The value of the derivative is given by the far field pattern

$$F'[\partial D]\,h = u'_\infty$$

of the solution u' of the exterior Dirichlet boundary value problem,

$$\triangle u' + k^2 u' \;=\; 0 \quad \text{in } \mathbb{R}^2 \backslash \overline{D} \tag{6.4}$$

$$u' \;=\; -h_\nu \frac{\partial u}{\partial \nu} \quad \text{on } \partial D, \tag{6.5}$$

where ν denotes the unit outward normal at ∂D, u the solution of the scattering problem, and $h_\nu = h \cdot \nu$ the normal component of h [15].

The existence of suitably defined higher order derivatives can be established [20, 24]. Here we are interested in a representation of the second derivative for an efficient implementation of the second degree approach. The following theorem states this result and is proven in the appendix.

Theorem 6.1 *Let ∂D be in the class C^3. The operator F is two times differentiable at ∂D with second derivative $F''[\partial D](h_1, h_2) = u''_\infty$ The function u''_∞ is the far field pattern of the radiating solution $u'' \in H^1_{\text{loc}}(\mathbb{R}^2 \backslash \overline{D})$ of the exterior Dirichlet problem*

$$\triangle u'' + k^2 u'' \;=\; 0 \quad \text{in } \mathbb{R}^2 \backslash \overline{D}, \tag{6.6}$$

$$u'' \;=\; -h_{1,\nu} \frac{\partial u'_2}{\partial \nu} - h_{2,\nu} \frac{\partial u'_1}{\partial \nu}$$

$$+ (h_{1,\nu} h_{2,\nu} - h_{1,\tau} h_{2,\tau}) \kappa \frac{\partial u}{\partial \nu}$$

$$+ \left(h_{1,\tau} (\tau \cdot \nabla (h_{2,\nu})) + h_{2,\tau} (\tau \cdot \nabla (h_{1,\nu})) \right) \frac{\partial u}{\partial \nu} \quad \text{on } \partial D, \quad (6.7)$$

where u is the solution of the scattering problem, u'_j ($j = 1, 2$) is the solution of the boundary value problem (6.4),(6.5) with respect to the variation h_j and κ denotes the curvature of ∂D.

Here $H^1_{\mathrm{loc}}(\mathbb{R}^2 \backslash \overline{D})$ denotes the convex space of function, locally in H^1, i.e. in the Sobolev space $H^1(\Omega)$ for any bounded subset $\Omega \subseteq \mathbb{R}^2 \backslash \overline{D}$. We denote by $h_\tau = h \cdot \tau$ the tangential component of h and $\tau \cdot \nabla$ is the tangential gradient. We should remark that the regularity assumptions on ∂D and h_1, h_2 seem to be necessary to ensure the representation (6.6), (6.7). Under this assumption we obtain $u \in H^3(\Omega)$, and from the boundary value problem (6.4), (6.5) this implies that $u' \in H^2(\Omega)$ and finally $u'' \in H^1(\Omega)$ (see [6]).

Unfortunately, we were not able to verify condition (2.5) for a specification of the set of admissible boundaries in a Hilbert space which would imply convergence of the iterative schemes. However, we did implement the second degree method using the representation (6.6) and (6.7).

From the iteration (1.4), (1.5) and the above representation, we observe that we do need a solver for the exterior Dirichlet problem which allows the computation of the far field patterns u_∞, u'_∞, u''_∞ and the Neumann boundary values $\frac{\partial u}{\partial \nu}$, $\frac{\partial u'}{\partial \nu}$. We consider u^s represented by a mixed single– and double–layer potential

$$u^s(x) = \int_{\partial D} \left(\frac{\partial \Phi(x, y)}{\partial \nu_y} - ik\Phi(x, y) \right) \varphi(y) \, ds_y, \quad x \in \mathbb{R}^2 \backslash \overline{D}$$

with a density $\varphi \in C(\partial D)$ and the fundamental solution of the Helmholtz equation

$$\Phi(x, y) = \frac{i}{4} H_0^{(1)}(k|x - y|), \quad x \neq y$$

(see [4]). The jump conditions imply that u^s is a solution of the exterior boundary value problem provided the density φ solves the Fredholm equation

$$(I + K - ik\,S)\varphi = 2f \quad \text{on } \partial D, \tag{6.8}$$

where $f \in C(\partial D)$ denotes the boundary value on ∂D and the weakly singular integral operators are defined by

$$K\varphi(x) = 2 \int_{\partial D} \frac{\partial \Phi(x, y)}{\partial \nu_y} \varphi(y) \, ds_y \quad \text{and} \quad S\varphi(x) = 2 \int_{\partial D} \Phi(x, y)\varphi(y) \, ds_y \quad x \in \partial D.$$

Having solved (6.8) the far field pattern can be computed by evaluating

$$u_\infty(\hat{x}) = \sqrt{\frac{k}{8\pi}} \, e^{-i\frac{\pi}{4}} \int_{\partial D} (\nu_y \cdot \hat{x} + 1) \, e^{-ik\hat{x} \cdot y} \varphi(y) \, ds_y, \quad |\hat{x}| = 1, \tag{6.9}$$

and the Neumann boundary values by

$$\frac{\partial u^s}{\partial \nu} = \frac{1}{2}(ik\,I - ik\,K' + T)\varphi \tag{6.10}$$

26

with the weakly singular operator

$$K'\varphi(x) = 2 \int_{\partial D} \frac{\partial \Phi(x,y)}{\partial \nu_x} \varphi(y)\, ds_y, \quad x \in \partial D$$

and the singular operator

$$T\varphi(x) = 2\frac{\partial}{\partial \nu_x} \int_{\partial D} \frac{\partial \Phi(x,y)}{\partial \nu_y} \varphi(y)\, ds_y, \quad x \in \partial D.$$

For an implementation of the second degree method we have to specify the set of admissible boundaries. Here we assume starlike obstacles, i.e. the boundary can be represented by

$$\partial D = \left\{ x = r(t) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} : t \in [0, 2\pi) \right\}$$

with a 2π–periodic positive function $r \in C^3(\mathbb{R})$. By the n-th iteration r_n the next step is defined by $r_{n+1} = r_n + h$ from solving the linear equations (1.4) and (1.5).

Now we assume variations of r_n only in a finite dimensional setting represented by a trigonometric polynomial, $h(t) = \frac{1}{2}h_0 + \sum_{j=1}^{N} h_j^c \cos(jt) + h_j^s \sin(jt)$. Then an iteration step consists in solving the direct scattering problem with boundary represented by r_n by the integral equation (6.8) and computing the far field pattern $F(r_n)$ and the Neumann boundary values. Next we compute the $2N + 1$ derivatives u' in the direction of the basis function again using the integral equation (6.8) with $f = -h_\nu \frac{\partial u}{\partial \nu}$. This leads to the construction of the Jacobian matrix J and we solve for the predictor step

$$(J^*J + \alpha_1 R)\vec{h} = J^*(g - F(r_n)) \tag{6.11}$$

to obtain the Fourier coefficients $\vec{h}$ of $\tilde{h}$. Here the matrix R is diagonal with entries $R_{jj} = 1 + j^2$ for $j = 0, \dots N$ and $R_{jj} = 1 + (j - N)^2$ for $j = N + 1, \dots, 2N$ corresponding to the cosine and the sine modes in $\tilde{h}$. This approximates the Tikhonov regularisation with respect to the H^2 norm and shows a slightly better performance in suppressing oscillations in the reconstruction. Now using $\tilde{h}$ (suppressing the imaginary part) we can compute the corrected h by solving again $2N + 1$ direct problems for the second derivative in the direction of any basis function. This leads to a linear system

$$(T^*T + \alpha_2 R)\vec{h} = T^*(g - F(r_n)) \tag{6.12}$$

for the Fourier coefficients of h. Since the system is complex valued we again consider only the real part and use the update $r_{n+1}(t) = r_n(t) + \mathrm{Re}(h(t))$.

Note that the effort of computing the derivatives is small since we only have to change the right hand side of the integral equation (6.8). The Newton method of course only required using the prescribed predictor step. We also emphasize that we choose trigonometric polynomials for simplicity. The use of different basis functions can be of advantage in certain cases (see [4]).
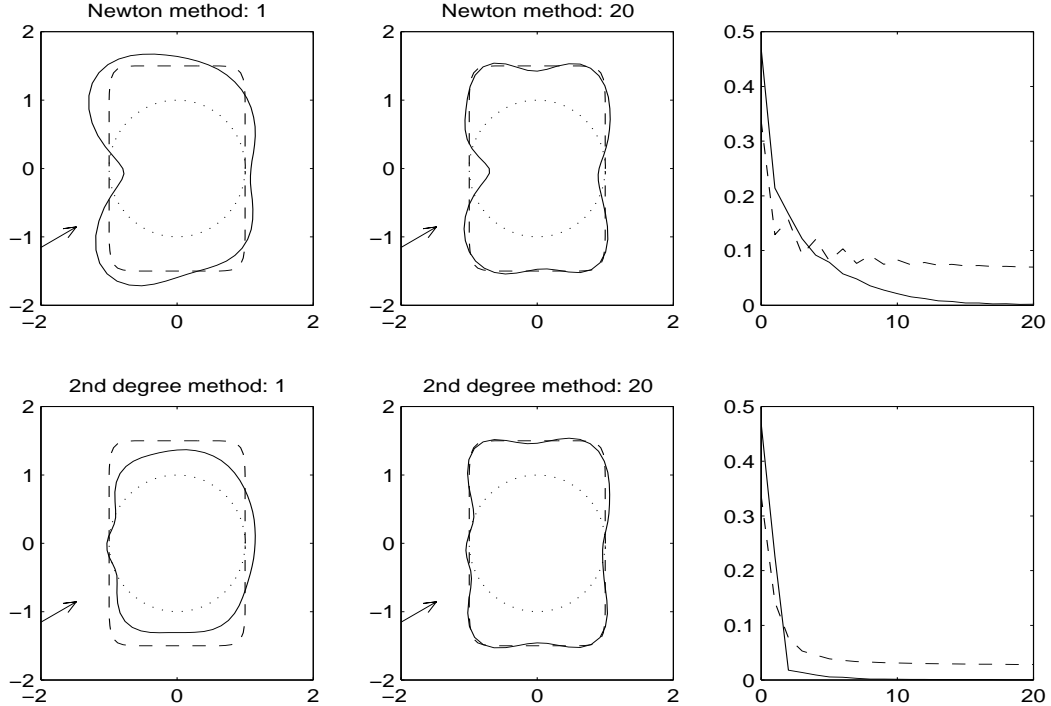
Figure 5: Reconstructions from noise–free data

In the examples used to illustrate the ideas the integral equation (6.8) is solved by Nyström's method at 64 discretisation points on ∂D using quadrature rules which take into account the weakly singular kernels (see [4]). K' is also computed by such a quadrature rule. The hypersingular operator T is evaluated by trigonometric interpolation as suggested in [17]. Using the trapezoidal rule for integrating the far field pattern the approach allows fast and accurate solutions of the direct problem. The dimension of the trigonometric polynomials for the reconstruction was chosen to be 21 ($N = 10$) in all examples.

To illustrate the performance of the second degree method we consider the wave number $k = 1$ and a rounded rectangle defined by

$$x_1^{10} + \left(\frac{2}{3}x_2\right)^{10} = 1$$

as the actual scatterer. This is one of the examples we tested, where a significant difference to the Newton method occurred. In some experiments the performance of both methods was very similar, but we never observed a case where the Newton method reached a significantly better reconstruction than the second degree approach.

Accurate artificial data computed by a different boundary integral equation based on the representation theorem were used. Figure 5 shows the result after the first
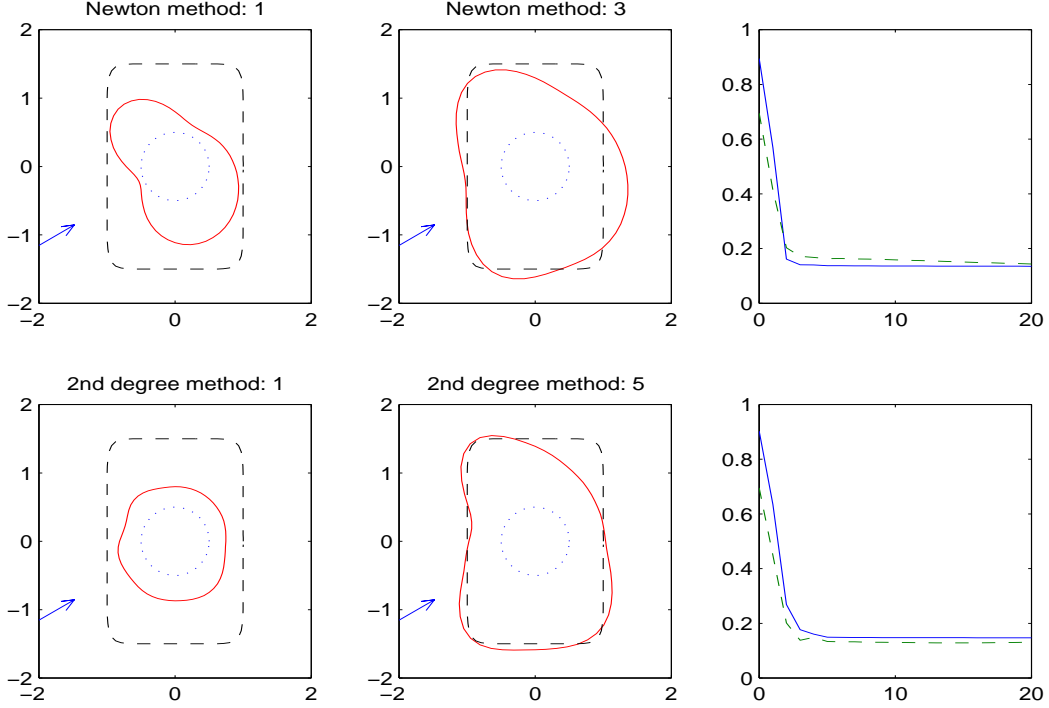
Figure 6: Reconstructions from noisy data

iteration and after 20 steps. The full line represents the reconstruction, the dashed line the actual scatterer and the dotted line the starting guess. Observe that the artificial indentation constructed by the Newton method in the initial iterations cannot be completely corrected. The second degree method avoids this artifact entirely. This difference cannot be seen in the relative residual error $\|u_\infty^\delta - F(r_n)\|/\|u_\infty^\delta\|$ plotted as the solid line curve on the bottom. The dashed error curve shows the approximation of the L^2–error of the difference in the reconstruction and the original boundary.

For such noise–free testing the regularisation parameters were fixed with $\alpha = 0.01$ for the Newton method and $\alpha_1 = 0.001$, $\alpha_2 = 0.01$ for the second degree scheme. Increasing the parameter in the Newton method can smooth the indentation but with the price of getting poorer reconstructions in the shadow region. The incident direction d of the plane wave is marked by an arrow in the figures. Of course, if we use more information by adding further incident waves we can improve the reconstructions in both methods.

As is usual in such problems, the choice of the regularisation parameter is more delicate in the case of noisy data. Figure 6 shows reconstructions from a more difficult initial guess when 10% random noise was added to the data. We model noisy data by adding a random vector of 10% of its approximated L^2 norm. The

29

results presented in figure 6 show the reconstruction closest to the mean value of the error from 100 runs with varying random noise. We have chosen $\alpha = 0.5$, $\alpha_1 = 0.05$ and $\alpha_2 = 0.5$ which was sufficient for all tests with this noise level. Note that the second degree iteration starts slower. This is compensated after a few iteration steps. The second picture in each method was stopped when the relative residual error was less than 15%. The number of iterations required is provided at the top of the pictures.

7 Some variations on the basic scheme

In the predictor–corrector scheme (1.4), (1.5) it is the second inversion that utilises the second degree model, the first step in (1.4) is required mainly to obtain the value of $\tilde{h}$ in order that (1.5) be a linear operator in h. This suggests that we might have flexibility in modifying the predictor step and indeed we considered several options. However, none of these appeared to improve on the scheme (1.4), (1.5).

First, at the nth iteration step we can simply take $\tilde{h} = h_{n-1}$, namely the value for h used at the previous iteration. This has the advantage of simplicity and not requiring an inversion of (1.4) along with attendant concerns about correct regularisation. In practice this approach worked poorly. The rapid initial convergence of the scheme guaranteed that the value of the perturbation h varied significantly from the previous iteration and so the predicted value $\tilde{h}$ was invariably of too large a norm.

Second, we can use a Landweber update for the predictor step. This would amount to replacing (1.4) by
$$\tilde{h} = \mu(F'[x_n])^*(g^\delta - F(x_n))$$
with a scaling parameter $\mu > 0$. If an inequality like (2.9) is still satisfied then convergence could also be obtained. The effect of a Landweber step is quite different from that of a Newton one. The former tends to preferentially improve the low frequencies of the solution whereas the latter tends to improve the high frequencies. This might seem an ideal choice since the predictor and corrector would then have properties that might be beneficially complementary. A drawback is the need to compute the adjoint operator of F'. This is sometimes easier to obtain than the derivative itself, but we are forced to compute F' anyway on the way to computing F'' as required by (1.5). From tests of this approach we usually could not observe any advantage since a better regularisation effect could be compensated by choosing larger values for α_1 and α_2. As a another suggestion along these lines, we can incorporate the second degree method into a Levenberg-Marquardt scheme applying it to one or both the predictor or corrector step.

We can also modify the corrector step in a variety of ways. One of the most obvious, and again borrowed ideas from the numerical solution of ordinary differential equations, is to iterate the corrector step. This means we use h computed again in

the same equation (1.5) as a predictor and determine a (hopefully) better update. Sometimes slightly better reconstructions can be obtained by this idea but we rarely found it worth the (slight) additional computational cost.

For Newton's method, Bakushinski [Ba] showed that adding an additional penalty term of the form $\beta_j(x_j - x_0)$ to the inversion step can act as a further regularisation. For a scheme using Tikhonov regularisation this is

$$x_{j+1} = x_j - ((F'[x_j])^* F'[x_j] + \alpha_j I)^{-1} \Big((F'[x_j])^* (F[x_j] - g) + \alpha_j (x_j - x_0) \Big).$$

He showed a convergence result for this method and established a stopping criterion for dealing with noisy data. The additional term penalises the solution for differing by too much from the initial guess. While this can stabilise an otherwise divergent sequence one can often achieve the same effect simply by increasing the Tikhonov parameter α_j. The Bakushinski-modified scheme usually requires substantially more iterations than the unmodified one. The idea can be applied to either the predictor or, we would expect with greater possible advantage, to the corrector. Of course we could implement this additional penalty in both steps.

Both (1.4) and (1.5) require regularisation for stable inversion, but we can use quite different regularisation schemes in the predictor and the corrector. We have noticed that good results are often obtained by using a smaller value of regularisation parameter in the predictor step than in the corrector, in fact a smaller value than would be necessary to stabilise the scheme using only the predictor (or Newton's method). This is only a statement that the second degree step is able to "correct" relatively minor levels of instability arising from the first stage. There is nothing to prevent the use of say spectral-cut off in one step and Tikhonov in the other, and indeed such a combination may be of value for certain types of problems.

Finally, we comment on the case where the derivatives in (1.4) and (1.5) evaluated at a general point x cannot be computed without considerable computational expense. Assuming that the computation is possible either analytically or numerically at some value $\hat{x}$ (possibly corresponding to a constant coefficient or simple geometry) then we can view the frozen predictor-corrector scheme described above as a quasi-Newton method with the second degree term "correcting" the "predicted" derivative. Applying a first order Taylor approximation to F' we obtain $F'[x]h = F'[\hat{x}]h + F''[\hat{x}](x - \hat{x}, h)$. If we use again a Newton step as the predictor then this would lead to an alternative quasi Newton scheme differing from the frozen version of (1.5) due to the absense of the factor of $\frac{1}{2}$. On the real line we arrive at the iteration scheme

$$x_{n+1} = x_n - \frac{f(x_n)f'(x_n)}{(f'(x_n))^2 - f''(x_n)f(x_n)}. \tag{7.1}$$

instead of (1.6). The scheme (7.1) is also known in the literature. It appears to be originally due to Schröder (see [ST]) and was designed to handle multiple roots. This suggests there may well be alternative correctors with quite different properties and

in fact their applicability may extend beyond the use as quasi-Newton approximants to the first derivative.

Appendix: Proof of Theorem 6.1

Let us start with some general remarks on the second derivative of an operator with respect to a domain. The problem in defining a second derivative lies in the definition of the first derivative in direction "h_1" on a perturbed boundary $\partial D_{h_2} = \{x + h_2(x) : x \in \partial D\}$. We consider two sufficiently small variations of the boundary ∂D described by vector fields $h_1, h_2 \in C^3(\partial D)$. Without changing the notation we shall use $h_1, h_2 \in C^3(\{|x| < R\})$ for extensions, which are assumed to have a support in a neighbourhood of ∂D. If we allow an arbitrary extension of h_1 (as in the proof of the frist derivative, [15]), we would obtain $(\partial D_{h_2})_{h_1} = \{x + h_2(x) + h_1(x + h_2(x)) : x \in \partial D\}$. But this is not symmetric and the variation of ∂D_{h_2} is depending on h_2.

If we consider a set of admissible domains fixed by a certain type of parametrisation (for instance starlike, see below), the variation of the perturbed domain is given by

$$(\partial D_{h_2})_{h_1} = \{x + h_2(x) + h_1(x) : x \in \partial D\},$$

which leads to a linear space of variations. Under such an assumption the existence of heigher order derivatives for boundary integral operators is established in [20]. A different approach is suggested in [24]. From the previous observation we define the second derivative by a bounded bilinear operator satisfying

$$\lim_{\|h_2\| \to 0} \frac{1}{\|h_2\|} \sup_{\|h_1\|=1} \left\| F'[\partial D_{h_2}]\tilde{h}_1 - F'[\partial D]h_1 - F''[\partial D](h_1, h_2) \right\| = 0, \qquad \text{(A.1)}$$

where $\tilde{h}_1 = h_1 \circ \psi_2$ with the inverse function ψ_2 of the diffeomorphism $\varphi_2(x) = x + h_2(x)$ using a sufficiently small extension of h_2 with compact support. Then we obtain from the first order approximation $\tilde{h}_1 = h_1 - J_{h_1}h_2 + \mathcal{O}(\|h_2\|^2)$ and continuity of F' with respect to the boundary that the second derivative satisfies

$$F''[\partial D](h_1, h_2) = (F'[\partial D]h_1)'h_2 - F'[\partial D](J_{h_1}h_2). \qquad \text{(A.2)}$$

We use the notation J_h for the Jacobian matrix of a vector field. The first term on the left means the domain derivative in direction h_2 of the first derivative u'_∞ with respect to the direction h_1. The second term is already known and is described by the far field pattern of a radiating solution v of the Helmholtz equation with boundary condition $v = -(J_{h_1}h_2 \cdot \nu)\frac{\partial u}{\partial \nu}$.

In [24] it is shown that F'' defined by (A.2) ensures a second order expansion of F. Here we will prove a representation of F'' and that F'' is the second Fréchet derivative in the sense of (A.1), which confirms the existence result in [20].

We have to determine the first term in (A.2). Let us consider the weak formulation of the scattering problem (6.1)–(6.3). Green's formula yields

$$\int_\Omega \left(\nabla u \nabla \overline{v} - k^2 u \overline{v} \right) dx - \langle \Lambda u, v \rangle = \int_{\partial \Omega} \left(\frac{\partial u^i}{\partial \nu} - \Lambda u^i \right) \overline{v} \, ds \qquad \text{(A.3)}$$

for all $v \in H = \{w \in H^1(\Omega) : w|_{\partial D} = 0\}$, where $\Omega = \mathbb{R}^N \setminus \overline{D} \cap \{x \in \mathbb{R}^N : |x| < R\}$ ($N = 2, 3$) with $R > 0$ such that $D \subseteq \{|x| \leq R\}$. $H^s(\Omega)$ denotes the usual Sobolev spaces of order $s \in \mathbb{R}$. The radiation condition is incorporated in equation (A.3) by a non local boundary condition on $|x| = R$ described by the Dirichlet-to-Neumann operator $\Lambda : H^{1/2}(\{|x| = R\}) \to H^{-1/2}(\{|x| = R\})$ with $\Lambda u = \frac{\partial w}{\partial \nu}$, where w is the radiating solution of the exterior Dirichlet problem, $(\triangle + k^2)w = 0$ in $\mathbb{R}^N \setminus \{|x| \leq R\}$ and $w = u$ on the boundary. Thus, the scattered part u^s of a solution of (A.3) can be extended to a radiating solution $u^s \in H^1_{\text{loc}}(\mathbb{R}^N \setminus \overline{D})$ of the Helmholtz equation. The dual pairing $\langle ., . \rangle$ in $H^{-1/2}(\partial D) \times H^{1/2}(\partial D)$ is defined by the bounded extension of

$$\langle f, v \rangle = \int_{|x|=R} f v \, ds \quad \text{for } f \in L^2(\{|x| = R\}) \quad \text{and } v \in H^{1/2}(\{|x| = R\}).$$

With this notation we define the sesquilinear form

$$\mathcal{L}(u, v) = \int_\Omega \left[\nabla u \nabla \overline{v} - k^2 u \overline{v} \right] dx - \langle \Lambda u, v \rangle.$$

Then the scattering problem is equivalent to

$$\mathcal{L}(u, v) = (f, v), \quad \text{for all } v \in H$$

with

$$(f, v) = \int_{|x|=R} \left(\frac{\partial u^i}{\partial \nu} - \Lambda u^i \right) \overline{v} \, ds.$$

Moreover, this variational equation is uniquely solvable for any $f \in H^{-1}(\Omega)$, [15].

With an extension of h_1 and h_2 we can define the material derivative

$$w_i = u_i' + h_i \cdot \nabla u \in H,$$

where u_i' denotes the domain derivative defined by (6.4), (6.5). A slight modification of the proof presented in [15] (see also [11]) shows that $w_i \in H$ is the derivative of u in Ω in the sense of

$$\frac{1}{\|h_i\|_{C^1}} \|\tilde{u}_{h_i} - u - w_i\|_{H^1} \to 0, \quad \text{if } \|h_i\|_{C^1} \to 0. \tag{A.4}$$

Here we use the notation $\tilde{u}_{h_i} = u_{h_i}(x + h_i(x))$ with the solution u_{h_i} of the scattering problem with boundary $\partial D_{h_i} = \{x + h_i(x) : x \in \partial D\}$. The material derivative w_i is the unique solution of

$$\mathcal{L}(w_i, v) = \int_\Omega \left[\nabla u \, A_i \nabla \overline{v} + k^2 u \overline{v} \operatorname{div}(h_i) \right] dx \tag{A.5}$$

for all $v \in H$ with $A_i = J_{h_i} + J_{h_i}^\top - \operatorname{div}(h_i)I$ and the Jacobian matrix J_{h_i} of the extension of h_i.

Now we proceed as follows. For fixed h_1 we consider the variation of w_1 with respect to h_2. Therefore we introduce the notation w_{1,h_2} which is the material derivative in the direction h_1 replacing D by $D_2 = \{x + h_2(x) : x \in D\}$. From the sesquilinear form $\mathcal{L}$ acting on this function we obtain by certain estimates the existence of a derivative with respect to h_2. Finally we can split this derivative in $w_1' = (u_1')' + g$ with a radiating solution of the Helmholtz equation $(u_1')'$ and a locally supported function g. Computing the boundary values of g leads to the boundary value problem representing the domain derivative $(u_1')'$.

The definition (A.5) of w_1 leads to

$$\int_{\Omega_{h_2}} \left[\nabla w_{1,h_2} \cdot \nabla \overline{v} - k^2 w_{1,h_2} \overline{v} \right] dx - \langle \Lambda w_{1,h_2}, v \rangle$$

$$= \int_{\Omega_{h_2}} \left[\nabla u_{h_2} A_1 \nabla \overline{v} + k^2 u_{h_2} \overline{v} \operatorname{div}(h_1) \right] dx.$$

We define the diffeomorphism $\varphi_i(x) = x + h_i(x)$ mapping Ω onto Ω_{h_i} assuming h_i is sufficiently small ($i = 1, 2$). In order to avoid the dependence of the domain of integration on h_2 a change of variables leads to

$$\int_{\Omega} \left[\nabla \tilde{w}_{1,h_2} \left(J_{\psi_2} J_{\psi_2}^{\top} \det(J_{\varphi_2}) \right) \nabla \overline{v} - k^2 \tilde{w}_{1,h_2} \overline{v} \det(J_{\varphi_2}) \right] dx - \langle \Lambda \tilde{w}_{1,h_2}, v \rangle$$

$$= \int_{\Omega} \left[\nabla \tilde{u}_{h_2} \left(J_{\psi_2} \tilde{A}_1 J_{\psi_2}^{\top} \det(J_{\varphi_2}) \right) \nabla \overline{v} + k^2 \tilde{u}_{h_2} \overline{v} \widetilde{\operatorname{div}(h_1)} \det(J_{\varphi_2}) \right] dx,$$

for all $v \in H$, where $\tilde{w}(x) = w(\varphi_2(x))$, $\tilde{A}_1(x) = A_1(\varphi_2(x))$ and ψ_2 denotes the inverse function of φ_2. By equation (A.5) it follows that the difference to the material derivative w_1 is

$$\mathcal{L}(\tilde{w}_{1,h_2} - w_1, v) \tag{A.6}$$

$$= \int_{\Omega} \left[\nabla \tilde{w}_{1,h_2} \left(I - J_{\psi_2} J_{\psi_2}^{\top} \det(J_{\varphi_2}) \right) \nabla \overline{v} - k^2 \tilde{w}_{1,h_2} \overline{v}(1 - \det(J_{\varphi_2})) \right] dx$$

$$+ \int_{\Omega} \left[\nabla \tilde{u}_{h_2} \left(J_{\psi_2} \tilde{A}_1 J_{\psi_2}^{\top} \det J_{\varphi_2} - A_1 \right) \nabla \overline{v} + k^2 \tilde{u}_{h_2} \overline{v} \left(\widetilde{\operatorname{div}(h_1)} \det(J_{\varphi_2}) - \operatorname{div}(h_1) \right) \right] dx$$

$$+ \int_{\Omega} \left[\nabla (\tilde{u}_{h_2} - u) A_1 \nabla v + k^2 (\tilde{u}_{h_2} - u) \overline{v} \operatorname{div}(h_1) \right] dx.$$

Elementary calculations provide the estimates

$$\|\det(J_{\varphi_i}) - 1 - \operatorname{div} h_i\|_{\infty} = \mathcal{O}(\|h_i\|_{C^1}^2)$$

$$\|J_{\psi_j} J_{\psi_i}^{T} \det(J_{\varphi_i}) - I + A_i\|_{\infty} = \mathcal{O}(\|h_i\|_{C^1}^2)$$

$$\|\widetilde{\operatorname{div} h_1} \det(J_{\varphi_2}) - \operatorname{div}(h_1) - \operatorname{div}(h_1) \operatorname{div}(h_2) - \nabla(\operatorname{div} h_1) \cdot h_2\|_{\infty} = \mathcal{O}(\|h_2\|_{C^1}^2)$$

$$\|J_{\psi_2} \tilde{A}_1 J_{\psi_2}^{T} \det(J_{\varphi_2}) - A_1 - A_1 \operatorname{div} h_2 + J_{h_2} A_1 + A_1 J_{h_2}^{\top} - A_1'(h_2)\|_{\infty} = \mathcal{O}(\|h_2\|_{C^1}^2),$$

34

where the matrix $A'_1(h_2)$ is defined by its elements $(A'_1(h_2))_{k,l} = \nabla((A_1)_{k,l}) \cdot h_2$.

By the difference (A.6) a perturbation argument as in [15] yields that $w_1 \in H$ depends continuously on h_2. Moreover, this continuity implies through the above estimates that w_1 is locally differentiable with respect to h_2 in the sense of (A.4) and the derivative w'_1 satisfies

$$
\begin{aligned}
\mathcal{L}(w'_1, v) &= \int_\Omega \left[\nabla w_1 A_2 \nabla \overline{v} + k^2 w_1 \overline{v} \operatorname{div}(h_2) \right] dx \\
&\quad - \int_\Omega \left[\nabla u \left(J_{h_2} A_1 + A_1 J_{h_2}^\top - A_1 \operatorname{div}(h_2) - A'_1(h_2) \right) \nabla \overline{v} \right] dx \\
&\quad + k^2 \int_\Omega \left[\nabla(\operatorname{div} h_1) \cdot h_2 + \operatorname{div}(h_1) \operatorname{div}(h_2) \right] u \overline{v} \, dx \\
&\quad + \int_\Omega \left[\nabla w_2 A_1 \nabla \overline{v} + k^2 w_2 \overline{v} \operatorname{div}(h_1) \right] dx
\end{aligned}
\tag{A.7}
$$

for all $v \in H$. We observe that the right hand side of this equation has a support in the neighbourhood of ∂D. Therefore there exists a splitting $w'_1 = (u'_1)' + g$, with a radiating solution $(u'_1)'$ of the Helmholtz equation in $\mathbb{R}^N \setminus \overline{D}$ which is identical to w'_1 in the exterior of the support of h_1 and h_2. Since $w'_1 \in H$ vanishes on ∂D, the Dirichlet boundary values of $(u'_1)'$ are determined by g.

Consider the first and the last integral in the variational equation (A.7). By the identity

$$
\nabla w \left(J_h + J_h^T - \operatorname{div}(h) I \right) \nabla \overline{v}
\tag{A.8}
$$
$$
= \operatorname{div}\left[(h \cdot \nabla w) \nabla \overline{v} + (h \cdot \nabla \overline{v}) \nabla w - (\nabla w \cdot \nabla \overline{v}) h \right] - (h \cdot \nabla \overline{v}) \triangle w - (h \cdot \nabla w) \triangle \overline{v}
$$

and using the boundary conditions $w_1 = v = 0$ on ∂D and the divergence theorem we obtain

$$
\begin{aligned}
\int_\Omega & \left[\nabla w_j A_i \nabla \overline{v} + k^2 w_j \overline{v} \operatorname{div}(h_i) \right] dx \\
&= \int_\Omega \operatorname{div}\left[(h_i \cdot \nabla w_j) \cdot \nabla \overline{v} + (h_i \cdot \nabla \overline{v}) \nabla w_j - (\nabla w_j \cdot \nabla \overline{v}) h_i \right] dx \\
&\quad - \int_\Omega \left[(h_i \cdot \nabla w_j) \triangle \overline{v} + (h_i \cdot \nabla \overline{v}) \triangle w_j \right] dx + k^2 \int_\Omega w_j \overline{v} \operatorname{div}(h_i) \, dx \\
&= \int_\Omega \left[\nabla(h_i \cdot \nabla w_j) \nabla \overline{v} - k^2 (h_i \cdot \nabla w_j) \overline{v} \right] dx \\
&\quad - \int_\Omega (h_i \cdot \nabla \overline{v}) \triangle (u'_j + h_j \cdot \nabla u) \, dx - k^2 \int_\Omega (h_i \cdot \nabla \overline{v}) w_j \, dx \\
&= \mathcal{L}(h_i \cdot \nabla w_j, v) - \int_\Omega (h_i \cdot \nabla \overline{v}) \left(\triangle (h_j \cdot \nabla u) + k^2 (h_j \cdot \nabla u) \right) dx
\end{aligned}
$$

35

for $i \neq j$, since u'_j is a solution of the Helmholtz equation. Equation (A.7) yields

$$\mathcal{L}(w'_1 - h_1 \cdot \nabla w_2 - h_2 \cdot \nabla w_1, v) \tag{A.9}$$

$$= -\int_\Omega \left[\nabla u (J_{h_2} A_1 + A_1 J_{h_2}^\top - A_1 \mathrm{div} h_2 - A'_1(h_2)) \nabla \overline{v} \right] dx$$

$$+ k^2 \int_\Omega \left(\nabla(\mathrm{div} h_1) \cdot h_2 + \mathrm{div}(h_1) \mathrm{div}(h_2) \right) u \overline{v} \, dx$$

$$- \int_\Omega (h_1 \cdot \nabla \overline{v}) \left(\triangle (h_2 \cdot \nabla u) + k^2 (h_2 \cdot \nabla u) \right) dx$$

$$- \int_\Omega (h_2 \cdot \nabla \overline{v}) \left(\triangle (h_1 \cdot \nabla u) + k^2 (h_1 \cdot \nabla u) \right) dx.$$

In extension of equation (A.8) we compute the identity

$$\nabla u \left(J_h A + A J_h^T - (\mathrm{div} h) A - (A'(h)) \right) \nabla v$$

$$= \mathrm{div} \left[(h \cdot \nabla u) A \nabla v + (h \cdot \nabla v) A^\top \nabla u - ((A^\top \nabla u) \cdot \nabla v) h \right]$$

$$- (h \cdot \nabla u) \mathrm{div}(A \nabla v) - (h \cdot \nabla v) \mathrm{div}(A^\top \nabla u).$$

Applying this equation in the first integrand on the right hand side of (A.9) implies by the divergence theorem

$$\int_\Omega \left[\nabla u \left(J_{h_2} A_1 + A_1 J_{h_2}^\top - A_1 \mathrm{div}(h_2) - A'_1(h_2) \right) \nabla \overline{v} \right] dx \tag{A.10}$$

$$= \int_\Omega \mathrm{div} \left[(h_2 \cdot \nabla u) A_1 \nabla \overline{v} + (h_2 \cdot \nabla \overline{v}) A_1 \nabla u - ((A_1 \nabla u) \cdot \nabla \overline{v}) h_2 \right] dx$$

$$- \int_\Omega \left[(h_2 \cdot \nabla u) \mathrm{div}(A_1 \nabla \overline{v}) + (h_2 \cdot \nabla \overline{v}) \mathrm{div}(A_1 \nabla u) \right] dx$$

$$= \int_\Omega \left[\nabla(h_2 \cdot \nabla u) A_1 \nabla \overline{v} - (h_2 \cdot \nabla \overline{v}) \mathrm{div}(A_1 \nabla u) \right] dx$$

(note that A_j is symmetric). It follows from the Helmholtz equation applied to u that $\mathrm{div}(A_1^\top \nabla u - \nabla(h_1 \cdot \nabla u)) = k^2 \mathrm{div}(h_1) u + k^2 h_1 \cdot \nabla u$. Applying this identity to the second term of (A.10) and the equation (A.8) to the first term, we obtain by the divergence theorem and the boundary conditions that

$$\int_\Omega \left[\nabla u (J_{h_2} A_1 + A_1 J_{h_2}^\top - A_1 \mathrm{div} h_2 - A'_1(h_2)) \nabla \overline{v} \right] dx$$

$$= \int_\Omega \mathrm{div} \left[(h_1 \cdot \nabla(h_2 \cdot \nabla u)) \nabla \overline{v} + (h_1 \cdot \nabla \overline{v}) \nabla(h_2 \cdot \nabla u) - (\nabla(h_2 \cdot \nabla u) \cdot \nabla \overline{v}) h_1 \right] dx$$

36

$$-\int_\Omega (h_1 \cdot \nabla(h_2 \cdot \nabla u)) \bigtriangleup \overline{v}\, dx \;-\; \int_\Omega (h_1 \cdot \nabla \overline{v}) \bigtriangleup (h_2 \cdot \nabla u)\, dx$$

$$-\int_\Omega (h_2 \cdot \nabla \overline{v}) \Big(\bigtriangleup (h_1 \cdot \nabla u) + k^2 \mathrm{div}(h_1)u + k^2 h_1 \cdot \nabla u \Big)\, dx$$

$$= \int_\Omega \nabla(h_1 \cdot \nabla(h_2 \cdot \nabla u)) \cdot \nabla \overline{v}\, dx$$

$$- \int_\Omega \Big[(h_1 \cdot \nabla \overline{v}) \bigtriangleup (h_2 \cdot \nabla u) + (h_2 \cdot \nabla \overline{v}) \bigtriangleup (h_1 \cdot \nabla u) \Big]\, dx$$

$$-k^2 \int_\Omega (h_2 \cdot \nabla \overline{v}) u\, \mathrm{div}(h_1) + (h_1 \cdot \nabla u)(h_2 \cdot \nabla \overline{v})\, dx.$$

We insert this equation in (A.9) and obtain using the boundary condition $u = v = 0$ on ∂D

$$\mathcal{L}(w_1' - h_1 \cdot \nabla w_2 - h_2 \cdot \nabla w_1, v)$$

$$= k^2 \int_\Omega \Big(\nabla(\mathrm{div}\, h_1) \cdot h_2 + \mathrm{div}(h_1)\mathrm{div}(h_2) \Big) u\overline{v}\, dx$$

$$- \int_\Omega \nabla(h_1 \cdot \nabla(h_2 \cdot \nabla u)) \cdot \nabla \overline{v}\, dx$$

$$+ k^2 \int_\Omega (h_2 \cdot \nabla \overline{v}) u\, \mathrm{div}(h_1)\, dx - k^2 \int_\Omega (h_1 \cdot \nabla \overline{v})(h_2 \cdot \nabla u)\, dx.$$

$$= - \int_\Omega \nabla(h_1 \cdot \nabla(h_2 \cdot \nabla u)) \cdot \nabla \overline{v}\, dx$$

$$-k^2 \int_\Omega \Big[(h_2 \cdot \nabla u)\overline{v}\, \mathrm{div}(h_1) + (h_1 \cdot \nabla \overline{v})(h_2 \cdot \nabla u) \Big]\, dx$$

$$= -\mathcal{L}\left(h_1 \cdot \nabla(h_2 \cdot \nabla u), v \right).$$

The last identity shows by the invertibility of the sesquilinear form that

$$w_1' = (u_1')' + h_1 \cdot \nabla w_2 + h_2 \cdot \nabla w_1 - h_1 \cdot \nabla(h_2 \cdot \nabla u)$$

in $H^1(\Omega)$, where $(u_1')'$ can be extended to a radiating solution of the Helmholtz equation in $\mathbb{R}^N \setminus \overline{D}$. Thus, on the boundary ∂D we obtain

$$(u_1')' = -h_1 \cdot \nabla u_2' - h_{2,\nu} \frac{\partial w_1}{\partial \nu}. \tag{A.11}$$

Note that this domain derivative of u_1' depends only on the normal component of h_2.

In deriving the boundary condition for the second derivative u'' defined by (A.2) we have to subtract from (A.11) the boundary value $-\nu^\top J_{h_1} h_2 \frac{\partial u}{\partial \nu}$ of the domain derivative of u with respect to the direction $J_{h_1} h_2$. Let us consider the two dimensional case. By the boundary condition of u_2' we obtain

$$
\begin{aligned}
u'' \;=\; & -h_{1,\nu}\frac{\partial u_2'}{\partial \nu} - h_{2,\nu}\frac{\partial u_1'}{\partial \nu} + h_{1,\tau}\tau \cdot \nabla\big(h_{2,\nu}\frac{\partial u}{\partial \nu}\big) - h_{2,\nu}\frac{\partial(h_1 \cdot \nabla u)}{\partial \nu} \\
& + \Big(\nu^\top J_{h_1}\tau\, h_{2,\tau} + \nu^\top J_{h_1}\nu\, h_{2,\nu}\Big)\frac{\partial u}{\partial \nu}.
\end{aligned}
$$

The following Lemma and the identity $\tau \cdot \nabla\big(\frac{\partial u}{\partial \nu}\big) = \tau H_u \nu$, which follows from $u = 0$ on the boundary, we continue in computing

$$
\begin{aligned}
u'' \;=\; & -h_{1,\nu}\frac{\partial u_2'}{\partial \nu} - h_{2,\nu}\frac{\partial u_1'}{\partial \nu} + \kappa h_{1,\nu} h_{2,\nu}\frac{\partial u}{\partial \nu} \\
& + h_{1,\tau}\tau \cdot \nabla(h_{2,\nu})\frac{\partial u}{\partial \nu} + \nu^\top J_{h_1}\tau\, h_{2,\tau}\frac{\partial u}{\partial \nu}.
\end{aligned}
$$

By orthogonality holds $\nu^\top J_{h_1} h_{2,\tau} = h_{2,\tau} \cdot \nabla(h_{1,\nu}) - h_{2,\tau} h_{1,\tau}(\tau^\top J_\nu \tau)$ and we obtain

$$
\begin{aligned}
u'' \;=\; & -h_{1,\nu}\frac{\partial u_2'}{\partial \nu} - h_{2,\nu}\frac{\partial u_1'}{\partial \nu} + \kappa h_{1,\nu} h_{2,\nu}\frac{\partial u}{\partial \nu} \\
& + \Big[h_{1,\tau}\tau \cdot \nabla(h_{2,\nu}) + h_{2,\tau}\tau \cdot \nabla(h_{1,\nu}) - (\tau^\top J_\nu \tau)h_{1,\tau} h_{2,\tau}\Big]\frac{\partial u}{\partial \nu}.
\end{aligned}
$$

Since for the curvature holds $\kappa = \tau^\top J_\nu \tau$ we conclude the boundary condition (6.7).

Similar we can obtain from (A.11) the boundary condition in $\mathbb{R}^3$. Note that the above proof shows the second Fréchet derivative of u_∞ for instance for starlike domains. Finally we complete the proof by the following Lemma.

Lemma A.1 *Let $u \in C^2(\mathbb{R}^n \setminus D)$ $(n = 2, 3)$ be a solution of $\triangle u + k^2 u = 0$ with $u|_{\partial D} = 0$. Then*

$$
\nu^\top H_u \nu = -(n-1)\kappa \frac{\partial u}{\partial \nu} \quad \text{on } \partial D
$$

is satisfied, where H_u is the Hessian matrix of u and κ denotes the (mean) curvature of the boundary ∂D.

Proof: We start with the two dimensional case: Let Φ be a local parametrisation of the boundary in a neighbourhood of a point $x_0 = \Phi(0) \in \partial D$ with $|\dot\Phi(0)| = 1$. Then at x_0 we compute from $\nu(x_0) = (\dot\Phi_2(0), -\dot\Phi_1(0))^\top$

$$
\nu^\top H_u \nu = (\dot\Phi_2)^2 \frac{\partial^2 u}{\partial x_1^2} + (\dot\Phi_1)^2 \frac{\partial^2 u}{\partial x_2^2} - 2\dot\Phi_1 \dot\Phi_2 \frac{\partial^2 u}{\partial x_1 \partial x_2}.
$$

On the other hand, differentiating the boundary condition $u(\Phi(v)) = 0$ twice leads to

$$0 = (\dot{\Phi}_1)^2 \frac{\partial^2 u}{\partial x_1^2} + (\dot{\Phi}_2)^2 \frac{\partial^2 u}{\partial x_2^2} + 2\dot{\Phi}_1 \dot{\Phi}_2 \frac{\partial^2 u}{\partial x_1 \partial x_2} + \frac{\partial u}{\partial x_1} \ddot{\Phi}_1 + \frac{\partial u}{\partial x_2} \ddot{\Phi}_2.$$

The sum of both, the Helmholtz equation, and the boundary condition implies

$$\nu^\top H_u \nu = \nabla u \cdot \ddot{\Phi} = -\kappa \frac{\partial u}{\partial \nu}.$$

The case of three spatial dimensions is similar: Let $\Phi(v_1, v_2)$ denote a local parametrisation of ∂D in a neighbourhood of $\Phi(0,0) = x_0 \in \partial D$ with

$$\left| \frac{\partial \Phi}{\partial v_1}(0,0) \right| = \left| \frac{\partial \Phi}{\partial v_2}(0,0) \right| = 1, \quad \frac{\partial \Phi}{\partial v_1}(0,0) \cdot \frac{\partial \Phi}{\partial v_2}(0,0) = 0$$

and

$$\nu(x_0) = \frac{\partial \Phi}{\partial v_1}(0,0) \times \frac{\partial \Phi}{\partial v_2}(0,0).$$

A lengthy but straightforward calculation shows

$$\begin{aligned} \nu^\top H_u \nu &= \triangle u - \left(\frac{\partial \Phi}{\partial v_1} \right)^\top H_u \frac{\partial \Phi}{\partial v_1} - \left(\frac{\partial \Phi}{\partial v_2} \right)^\top H_u \frac{\partial \Phi}{\partial v_2} \\ &= -\left(\frac{\partial \Phi}{\partial v_1} \right)^\top H_u \frac{\partial \Phi}{\partial v_1} - \left(\frac{\partial \Phi}{\partial v_2} \right)^\top H_u \frac{\partial \Phi}{\partial v_2} \end{aligned}$$

at x_0 by the Helmholtz equation and the boundary condition $u|_{\partial D} = 0$.

From the boundary condition we also observe $\nabla u \cdot \frac{\partial \Phi}{\partial v_j} = 0$, $j = 1, 2$. Differentiating again with respect to v_j leads to

$$\left(\frac{\partial \Phi}{\partial v_j} \right)^\top H_u \frac{\partial \Phi}{\partial v_j} + \nabla u \cdot \frac{\partial^2 \Phi}{\partial v_j^2} = 0.$$

Now we consider the mean curvature at x_0 and obtain

$$2\kappa = \frac{\partial \nu}{\partial v_1} \cdot \frac{\partial \Phi}{\partial v_1} + \frac{\partial \nu}{\partial v_2} \cdot \frac{\partial \Phi}{\partial v_2} = -\nu \cdot \left(\frac{\partial^2 \Phi}{\partial v_1^2} + \frac{\partial^2 \Phi}{\partial v_2^2} \right)$$

Combining the last three equations yields

$$\nu^\top(x_0) H_u(x_0) \nu(x_0) = -2\kappa(x_0) \frac{\partial u}{\partial \nu}(x_0)$$

which finishes the proof.

References

[1] A.B. BAKUSHINSKII, *The problem of the convergence of the iteratively regularised Gauss-Newton method*, Comput. Maths. Math. Phys., **21** (1992), pp. 1353-1359.

[2] A. BJORK, *Numerical Methods for Least Squares Problems,* SIAM, Philadelphia, 1996, pp. 339ff.

[3] F. COLONIUS AND K. KUNISCH, *Output least squares stability in elliptic systems,* Appl. Math. Optim. **19** (1989), pp. 33–63.

[4] D.L. COLTON AND R. KRESS, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd edt., Springer-Verlag, Berlin, Heidelberg, New York, etc. 1998.

[5] H.W. ENGL, M. HANKE AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[6] D. GILBARG, N.S. TRUDINGER, *Elliptic Partial Differential Equation of Second Order*, Springer-Verlag, Berlin, Heidelberg, New York, etc. 1983.

[7] E. HALLEY, *A new, exact, and easy method of finding the roots of any equations generally, and that without any previous reduction,* [Latin], Philos. Trans. Roy. Soc. London, **18**, (1694), pp. 136–148. [English Translation: Philos. Trans. Roy. Soc. (1809)]

[8] M. HANKE, *A Regularizing Levenberg–Marquardt Scheme with Aplications to Inverse Groundwater Filtration Problems,* Inverse Problems **13** (1997), pp. 79–95.

[9] M. HANKE, F. HETTLICH, O. SCHERZER, *The Landweber Iteration for an Inverse Scattering Problem,* in Proc. of the 1995 design engineering technical conferences, Vol.3 Part C, K.-W. Wang et. al. (eds.), The American Society of Mechanical Engineers, New York, 1995, pp. 909-915.

[10] M. HANKE, A. NEUBAUER AND O. SCHERZER, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems,* Numer. Math. **72** (1995), pp. 21–37.

[11] F. HETTLICH, *Fréchet Derivatives in Inverse Obstacle Scattering,* Inverse Problems **11** (1995), pp. 371-382.

[12] F. HETTLICH AND W. RUNDELL, *Iterative Methods for the Reconstruction of an Inverse Potential Problem,* Inverse Problems **12** (1996), pp. 251-266.

[13] T. HOHAGE, *Logarithmic convergence rates of the iteratively regularized Gauss–Newton method for an inverse potential and an inverse scattering problem,* Inverse Problems **13** (1997), pp. 1279–1299.

[14] V. Isakov, *Inverse Source Problems*, Mathematical Surveys and Monographs, Vol 34, American Mathematical Society, Providence, Rhode Island, 1990.

[15] A. KIRSCH, *The Domain Derivative and Two Applications in Inverse Scattering Theory,* Inverse Problems **9** (1993), pp. 81-96.

[16] A. KIRSCH, *An Introduction to the Mathematical Theory of Inverse Problems,* Springer, New–York, 1996.

[17] R. KRESS, *Linear Integral Equations,* Springer, Berlin, Heidelberg, 1989.

[18] R. KRESS, *A Newton method in inverse obstacle scattering,* in Proc. Second International Symposium on Inverse Problems in Engineering Mechanics, ISIP, 1994.

[19] R. KRESS AND W. RUNDELL, *A quasi-Newton method in inverse obstacle scattering,* Inverse Problems **10** (1994), pp. 1145–1157.

[20] R. POTTHAST, *Fréchet differentiability of boundary integral operators,* Inverse Problems **10** (1994), pp. 431-447.

[21] W. RUNDELL AND P.E. SACKS, *Reconstruction Techniques for Classical Inverse Sturm–Liouville Problems,* Math. Comp. **58/197** (1992), pp. 161–183.

[22] T. R. SCAVO AND J. B. THOO, *On the Geometry of Halley's Method,* Amer. Math. Monthly, **102**, (1995), pp. 417–426.

[23] R.B. SCHNABEL AND P.D. FRANK, *Solving Systems of Nonlinear Equations by Tensor Methods,* in Proc. Inst. Math. Appl. Conf. Ser., **9** (1987), Oxford Univ. Press.

[24] J. SIMON, *Second Variations for Domain Optimization Problems,* in Proc. 4th Int. Conf. Control of Distributed Parameter Systems, Birkhäuser, Basel, Boston and Berlin, 1988, pp. 361–378.

[25] E. ZEIDLER, *Nonlinear Functional Analysis and its Applications I, Fixed Point Theorems,* Springer, New York, 1986.