On the Numerical Solution of a Three-Dimensional Inverse Medium Scattering Problem

Thorsten Hohage†

† Industrial Mathematics Institute, Univ. Linz, Altenberger Str. 69, A-4040 Linz
after July 1,2000: Konrad-Zuse Zentrum, Takustr.7, D-14195 Berlin

Abstract. We examine the scattering of time-harmonic acoustic waves in inhomogeneous media. The problem is to recover a spatially varying refractive index in a three-dimensional medium from far-field measurements of scattered waves corresponding to incoming waves from all directions. This problem is exponentially ill-posed and of large scale since a solution of the direct problem corresponds to solving a PDE in $\mathbb{R}^3$ for each incident wave. We construct a preconditioner for the conjugate gradient method applied to the normal equation (CGNE) to solve the regularized linearized operator equation in each Newton step. This reduces the number of operator evaluations dramatically compared to standard regularized Newton methods. Our method can also be applied effectively to other exponentially ill-posed problems, e.g. in impedance tomography, heat conduction and obstacle scattering. To solve the direct problems, we use an improved fast solver for the Lippmann-Schwinger equation suggested by Vainikko.

AMS classification scheme numbers: 65J20, 65J15, 81U40

1. Introduction

A classical problem in inverse scattering is to determine the refractive index of a medium from measurements of far field patterns of scattered time-harmonic acoustic waves in this medium. The direct scattering problem for an incident field u_i (e.g. $u_\mathrm{i}(x) = \exp(-\mathrm{i}\kappa x \cdot d)$ with $d \in S^2 := \{x \in \mathbb{R}^3 : |x| = 1\}$) is described by the differential equation

$$\Delta u + \kappa^2 n u = 0, \qquad x \in \mathbb{R}^3, \tag{1a}$$

$$u = u_\mathrm{i} + u_\mathrm{s}, \tag{1b}$$

$$r\left(\frac{\partial u_\mathrm{s}}{\partial r} - i\kappa u_\mathrm{s}\right) = \mathrm{o}\,(1)\,, \qquad r = |x| \to \infty. \tag{1c}$$

Here $\kappa > 0$ denotes the wave number (we only work with one fixed κ!), n is the refractive index of the medium, u_s is the scattered field, and u is the total field. We assume that $\mathrm{Re}\, n \geq 0$, $\mathrm{Im}\, n \geq 0$ and that $n = 1 - a$ with $\mathrm{supp}\, a \subset B_\rho := \{x \in \mathbb{R}^3 : |x| \leq \rho\}$, $\rho > 0$. For a derivation of this model we refer to [2, 17]. Eq. $(1c)$, which is called

Sommerfeld radiation condition, implies that the scattered field behaves asymptotically like an outgoing spherical wave:

$$u_{\mathrm{s}}(x) = \frac{\mathrm{e}^{\mathrm{i}\kappa|x|}}{|x|}\left\{u_\infty\left(\frac{x}{|x|}\right) + \mathrm{O}\left(\frac{1}{|x|}\right)\right\}, \qquad |x| \to \infty.$$

The function $u_\infty : S^2 \to \mathbb{C}$ is called *far-field pattern* or *scattering amplitude* of u_{s}. It is always an analytic function. To indicate the dependence on the direction $d \in S^2$ of the incident field $u_{\mathrm{i}}(x, d) := \exp(-\mathrm{i}\kappa x \cdot d)$, we write $u_\infty(\hat{x}, d)$, $u_{\mathrm{s}}(x, d)$, and $u(x, d)$.

The inverse problem we consider is to determine $a : B_\rho \to \mathbb{C}$ from measurements of $u_\infty : S^2 \times S^2 \to \mathbb{C}$. Introducing the operator $F : H_0^s(B_\rho) \to L^2(S^2 \times S^2)$, $a \mapsto u_\infty$, $s > \frac{3}{2}$, this problem can be formulated as an operator equation

$$F(a) = u_\infty. \tag{2}$$

It can be shown (cf. [2, 10, 17, 20]) that a solution to (2) is unique if it exists. However, as indicated by the analyticity of far field patterns, eq. (2) is severely ill-posed.

In the literature a number of methods have been suggested to solve the inverse acoustic medium scattering problem numerically. Langenberg ([19]) considers a linearized version of (2). Gutman and Klibanov ([8, 9]) confine themselves to reconstructing a fixed number of Fourier coefficients of a. This makes the problem well-posed, and it allows an efficient implementation of a quasi-Newton method provided the directions of the incident fields and the measurement points are chosen in a special way. The number of Fourier coefficients which can be reconstructed by this method depends on κ. Kleinman and van den Berg propose a modified gradient method for an output-least-squares formulation of the problem, which iterates on both a and u ([18]). Finally, Colton and Monk have developed a so-called *dual space method* (cf. [2, 3]) which does not directly use (2), but leads to a least-squares minimization problem involving the Lippmann-Schwinger equation.

Here we propose a Newton-type method tailored to large-scale exponentially ill-posed problems. Newton-type methods use the linearized operator equation

$$F'(a_n)h_n = -F(a_n) + u_\infty \tag{3}$$

to find an update $h_n = a_{n+1} - a_n$ in each iteration step. For ill-posed problems regularization methods have to be employed. A straightforward implementation of Newton's method would involve setting up the whole matrix for the derivate operator $F'(a_n)$, i.e. evaluating $F'(a_n)\varphi_j$ for all elements φ_j of a basis of $H_0^s(B_\rho)$. A formal differentiation of (1a) shows that $F'(a_n)\varphi_j$ for fixed φ_j can be characterized with help of a differential equation similar to (1) (cf. Section 2). Hence the numerical evaluation of $F'(a_n)\varphi_j$ involves solving a 3D PDE for each incident wave u_{i}. It is clear that a straightforward implementation of a regularized Newton method for eq. (2) would be far too time consuming. It turns out that standard iterative solvers need far too many iterations to solve eq. (3). Here we take advantage of the exponentially ill-posed nature of eq. (3) (which, of course, is a very undesirable property otherwise) to construct an efficient preconditioner for the conjugate gradient iteration applied to the normal equation corresponding to (3). The construction uses the close connection of the Lanczos and the CG method. With this method the number of direct problem solutions is reduced remarkably.

Before we summarize the contents of this paper, let us make some remarks concerning convergence properties of regularized Newton methods for (2), which will not be discussed here. A convergence analysis under logarithmic source conditions, which are usually appropriate for exponentially ill-posed problems, has been carried out in [14] (cf. also [4, 13, 15]). Unfortunately all convergence results need some condition restricting the degree of nonlinearity of F, and none of these conditions has been verified for the inverse acoustic medium problem, yet. We also do not know how logarithmic source conditions can be interpreted for our problem. However, in a forthcoming paper ([11]) we derive new conditional stability estimates which make it possible to replace the logarithmic source conditions by simple smoothness conditions in terms of Sobolev spaces.

The plan of this paper is as follows:

In Section 2 we study properties of the Fréchet derivative $F'(a)$ occurring in (3). We show that $F'(a)$ is compact and one-to-one, and that the singular values satisfy $\sigma_j(F'(a)) = \mathrm{O}(\exp(-cj^{1/4}))$.

For the numerical solution of the direct problems we have chosen a method which has recently been developed by Vainikko ([23]). This method is very attractive since it converges of arbitrarily high order depending on the smoothness of a, it has time complexity $\mathrm{O}(N^3 \log N)$ where N is the number of grid points in one direction in an equidistant grid, and it is comparatively easy to implement. In Section 3 we describe and analyze an improved version of Vainikko's method and present numerical test examples.

In Section 4 we introduce the Newton method for exponentially ill-posed problems mentioned above. The material is presented in the framework of abstract operator equations in Hilbert spaces since our method is also applicable to other exponentially ill-posed problems, e.g. in impedance tomography, heat conduction and inverse obstacle scattering.

In Section 5 we discuss some practical issues concerning the implementation of this method both in general and for our particular problem.

Finally, a number of numerical examples presented in Section 6 show the practicality and efficiency of the proposed method.

2. On the Fréchet derivative of the solution operator

It is well known that the scattering problem (1) has an equivalent formulation as an integral equation of the second kind called the *Lippmann-Schwinger equation*:

$$u(x) + \kappa^2 \int_{B_\rho} \Phi(x-y)a(y)u(y)\,\mathrm{d}y = u_{\mathrm{i}}(x), \qquad x \in \mathbb{R}^3. \tag{4}$$

Here $\Phi(x) := \frac{1}{4\pi}\frac{e^{i\kappa|x|}}{|x|}$ is the fundamental solution to the Helmholtz equation in free space. The integral operator with kernel $\kappa^2\Phi(x-y)$ will be denoted by V. It can be shown that (4) has a unique solution for all κ and all right hand sides. We refer to [2, 17] for an analysis of the direct scattering problem via (4).

We will also make frequent use of a different form of (4) where both sides are multiplied by a:

$$(I + aV)v = au_{\mathrm{i}} \tag{5}$$

Here a denotes a multiplication operator, and we have introduced the new unknown $v := au$. Note that u can be recovered from v by the formula $u = u_{\mathrm{i}} - Vv$, and that u defined this way satisfies (4). Moreover the far field pattern is given by

$$u_\infty = -\frac{\kappa^2}{4\pi} Ev \qquad \text{with} \qquad (Ev)(\hat{x}) := \int_{B_\rho} \mathrm{e}^{-\mathrm{i}\kappa\hat{x}\cdot y} v(y)\,\mathrm{d}y, \quad \hat{x} \in S^2.$$

In the following we will need some basic properties of Hilbert-Schmidt operators (cf., e.g., [22]). Let X, Y be Hilbert spaces and $\mathcal{L}(X, Y)$ the space of bounded linear operators from X to Y. Given a Hilbert basis $\{\varphi_j : j \in J\}$ of X, a scalar product can be defined by

$$\langle K_1, K_2 \rangle_{\mathcal{HS}} := \sum_{j\in J} \langle K_1\varphi_j, K_2\varphi_j \rangle$$

for all $K_1, K_2 \in \mathcal{L}(X, Y)$ with $\|K_l\|^2_{\mathcal{HS}} = \langle K_l, K_l \rangle_{\mathcal{HS}} < \infty$ $(l = 1, 2)$, and this definition does not depend on the choice of the φ_j. The space $\mathcal{HS}(X, Y) := \{K \in \mathcal{L}(X, Y) : \|K\|_{\mathcal{HS}} < \infty\}$ of Hilbert-Schmidt operators equipped with this inner product is a Hilbert space. If U and Z are further Hilbert spaces and $S \in \mathcal{L}(U, X)$, $T \in \mathcal{L}(Y, Z)$, then TKS is again a Hilbert-Schmidt operator, and $\|TKS\|_{\mathcal{HS}} \leq \|T\|\|K\|_{\mathcal{HS}}\|S\|$. Let G_1 and G_2 be submanifolds of $\mathbb{R}^{d_1}$, $\mathbb{R}^{d_2}$, respectively. Then the mapping

$$\begin{aligned} J_{G_1,G_2} : L^2(G_1 \times G_2) &\to \mathcal{HS}(L^2(G_1), L^2(G_2)) \\ k \mapsto K, \quad (K\varphi)(x) &:= \int_{G_1} k(x, y)\varphi(y)\,\mathrm{d}y \end{aligned}$$

is unitary, i.e. the spaces $L^2(G_1 \times G_2)$ and $\mathcal{HS}(L^2(G_1), L^2(G_2))$ are isometrically isomorphic. If $G_1 = G_2$ we simply write J_{G_1} for J_{G_1,G_1}, and we also write $\mathcal{HS}(X)$ for $\mathcal{HS}(X, X)$, and $\mathcal{L}(X)$ for $\mathcal{L}(X, X)$.

By superposition, $(J_{S^2}F(a))g = \int_{S^2} u_\infty(\cdot, d)g(d)\,\mathrm{d}s(d)$ is the far field pattern corresponding to the incident field $u_{\mathrm{i}}^g := \int_{S^2} u_{\mathrm{i}}(\cdot, d)g(d)\,\mathrm{d}s(d)$, i.e. $(J_{S^2}F(a))g = -\frac{\kappa^2}{4\pi}E(I + aV)^{-1}au_{\mathrm{i}}^g$. (Multiply (5) by $g(d)$, integrate with respect to d, and change the order of integration!). For a bounded linear operator $K : L^2(G_1) \to L^2(G_2)$, the adjoint of K with respect to the *bilinear* products $\langle \varphi_1, \varphi_2 \rangle_{\mathrm{bl}} := \int_{A,B} \varphi_1\varphi_2\,\mathrm{d}x$ is denoted by K^{t}. In particular, we have $(E^{\mathrm{t}}g)(y) = \int_{S^2} \mathrm{e}^{-\mathrm{i}y\cdot d}\,\mathrm{d}s(d) = u_{\mathrm{i}}^g$, so

$$J_{S^2}F(a) = -\frac{\kappa^2}{4\pi}E(I + aV)^{-1}aE^{\mathrm{t}}.$$

We first consider F as an operator between $L^\infty(B_\rho) \to L^2(S^2 \times S^2)$. Later when considering singular values of $F'(a)$ and for the Newton method, we have to replace $L^\infty(B_\rho)$ by some Hilbert space with an implementable norm. The domain of F is defined as $\mathcal{D}(F) := \{a \in L^\infty(B_\rho) : (I + aV)\text{is boundedly invertible}\}$. $\mathcal{D}(F)$ contains all

a satisfying the physical restrictions $\mathrm{Re}(1-a), \mathrm{Im}(1-a) \geq 0$. Moreover, a Neumann series argument shows that $\mathcal{D}(F)$ is open. This justifies to ignore the physical restrictions $\mathrm{Re}(1-a), \mathrm{Im}(1-a) \geq 0$ in practice since Newton's method is only locally convergent, anyways.

Proposition 2.1 *The operator $F : L^\infty(B_\rho) \to L^2(S^2 \times S^2)$ is Fréchet differentiable on $\mathcal{D}(F)$.*

Proof: It follows from elementary rules about differentiation in Banach spaces that the mapping $L^\infty(B_\rho) \to \mathcal{L}(L^2(B_\rho), L^2(S^2))$, $a \mapsto -\frac{\kappa^2}{4\pi} E(I+aV)^{-1}a$ is Fréchet differentiable. The operator E^t is Hilbert-Schmidt since its kernel $\mathrm{e}^{-\mathrm{i}\kappa\hat{x}\cdot y}$ is in $L^2(S^2 \times B_\rho)$. Moreover, the linear mapping $\mathcal{L}(L^2(B_\rho), L^2(S^2)) \to \mathcal{HS}(L^2(S^2))$, $K \mapsto KE^\mathrm{t}$ is bounded. Hence $a \mapsto J_{S^2}F(a)$ is Fréchet differentiable as the composition of these mappings, and so is $a \mapsto F(a) = J_{S^2}^* J_{S^2} F(a)$. ■

It can be shown analogously that $a \mapsto v = J^*_{S^2,B_\rho}((I+aV)^{-1}aE^\mathrm{t})$ is Fréchet differentiable. Differentiating (5) with respect to a in direction h for some fixed u_i yields

$$(I+aV)v'_h = hu$$

since $u = u_\mathrm{i} - Vv$. We will show in the next section how to solve (5) numerically. Then we can compute both v and v'_h by solving the same equation twice for different right hand sides. This also yields $u'_{\infty,h} = -\frac{\kappa^2}{4\pi} E v'_h$.

It can be shown that the operator F is one-to-one (cf. [2, 10, 17, 20]). We can use results from this uniqueness proof to show that $F'(a)$ is also one-to-one:

Proposition 2.2 *The operator $F'(a)$ is one-to-one for any $a \in \mathcal{D}(F)$.*

Proof: Assume that $F'(a)h = 0$ for $a \in L^\infty(B_\rho)$ and $h \in L^\infty(B_\rho)$. We want to show that $h = 0$. From the preceding remarks we know that

$$-\frac{\kappa^2}{4\pi}\int_{B_\rho} u_\mathrm{i}(x,\tilde{d})v'_h(x,d)\,\mathrm{d}x = u'_{\infty,h}(\tilde{d},d) = (F'(a)h)(\tilde{d},d) = 0$$

for all $d, \tilde{d} \in S^2$. Obviously, the operator V and the multiplication operator a are both self-adjoint with respect to the bilinear product $\langle\cdot,\cdot\rangle_\mathrm{bl}$ on $L^2(B_\rho)$. Therefore,

$$\begin{aligned}
0 &= \left\langle u_\mathrm{i}(\cdot,\tilde{d}), v'_h(\cdot,d)\right\rangle_\mathrm{bl} \\
&= \left\langle u_\mathrm{i}(\cdot,\tilde{d}), (I+aV)^{-1}(hu(\cdot,d))\right\rangle_\mathrm{bl} \\
&= \left\langle (I+Va)^{-1}u_\mathrm{i}(\cdot,\tilde{d}), hu(\cdot,d)\right\rangle_\mathrm{bl} \\
&= \left\langle u(\cdot,\tilde{d}), hu(\cdot,d)\right\rangle_\mathrm{bl}
\end{aligned}$$

for all $d, \tilde{d} \in S^2$. A result of the uniqueness proof for F states that $\overline{\mathrm{span}\{u(\cdot,d) : d \in S^2\}}$ is dense in $X := \overline{\{v \in C^2(B_{2\rho}) : \Delta v + \kappa^2(1-a)v = 0\}}$ where both denseness and closure are understood with respect to the L^2-norm. Hence, $\int_{B_\rho} hv_1v_2\,\mathrm{d}x = 0$ for all $v_1, v_2 \in X$. According to the main idea of the uniqueness proof for F this implies $h = 0$. ■

Proposition 2.3 *$F'(a) : H_0^s(B_\rho) \to L^2(S^2 \times S^2)$, $s > \frac{3}{2}$ is compact, and the singular values of $F'(a)$ satisfy $\sigma_j(F'(a)) = \mathrm{O}(\exp(-cj^{1/4}))$, $j \to \infty$ for some $c > 0$.*

Proof: Let us first consider point sources $\tilde{u}_\mathrm{i}(x, \eta) := \Phi(x - \eta)$, $\eta \in \partial B_{2\rho}$ as incident fields and corresponding near-fields $\tilde{u}_\mathrm{s}(\xi, \eta)$, $\xi, \eta \in \partial B_{2\rho}$ as measured data. We introduce the operator $\tilde{E} \in \mathcal{HS}(L^2(B_\rho), L^2(\partial B_{2\rho}))$ by

$$(\tilde{E}\varphi)(\xi) := \int_{B_\rho} \Phi(\xi - y)\varphi(y)\,\mathrm{d}y, \xi \in \partial B_{2\rho},$$

and $\tilde{F}(a) := -\kappa^2 J^*_{\partial B_{2\rho}}(\tilde{E}(I + aV)^{-1} a\tilde{E}^\mathrm{t})$. Then, as in Proposition 2.1 , $\tilde{F}$ is Fréchet differentiable in $L^\infty(B_\rho) \to L^2(\partial B_{2\rho} \times \partial B_{2\rho})$. By Sobolev's Embedding Theorem the $\|\cdot\|_{H^s}$ norm is stronger than the $\|\cdot\|_\infty$ norm for $s > \frac{3}{2}$, so $\tilde{F}$ is also differentiable from $H_0^s(B_\rho) \to L^2(S^2 \times S^2)$. Let $K : L^2(\partial B_{2\rho}) \to L^2(S^2)$ be the operator which maps a function $f \in L^2(\partial B_{2\rho})$ to the far field pattern v_∞ of the radiating solution v to the homogeneous Helmholtz equation with boundary data f on $\partial B_{2\rho}$. Obviously $E = K\tilde{E}$ and $E^\mathrm{t} = \tilde{E}^\mathrm{t} K^\mathrm{t}$. Hence $J_{S^2}F(a) = K(J_{\partial B_{2\rho}}\tilde{F}(a))K^\mathrm{t}$. Let $\Omega \in \mathcal{L}(\mathcal{HS}(L^2(\partial B_{2\rho}), \mathcal{HS}(L^2(S^2))))$ be defined by $\Omega A := KAK^\mathrm{t}$. Then $J_{S^2}F(a) = \Omega J_{\partial B_{2\rho}}\tilde{F}(a)$ and hence $J_{S^2}F'(a) = \Omega J_{\partial B_{2\rho}}\tilde{F}'(a)$. It follows from the Courant-Fischer characterization of singular values that

$$\sigma_j(F'(a)) = \sigma_j(\Omega J_{\partial B_{2\rho}}\tilde{F}'(a)) \le \sigma_j(\Omega)\|J_{\partial B_{2\rho}}\tilde{F}'(a)\| = \sigma_j(\Omega)\|\tilde{F}'(a)\|.$$

This means, we have to estimate the singular values of Ω. We first compute the exact singular values of Ω. Let Y_l^m, $l \in \mathbb{N}_0$, $|m| \le l$ denote the sperical harmonics, which are a Hilbert basis of $L^2(S^2)$. Correspondingly, $\tilde{Y}_l^m(x) := (2\rho)^{-1}Y_l^m(x/|x|)$ is a Hilbert basis of $L^2(\partial B_{2\rho})$. It can be shown (cf. [2]) that $K\tilde{Y}_l^m = \gamma_l^m Y_l^m$ where $\gamma_l^m := \left(\kappa(-\mathrm{i})^{l+1}(2\rho\kappa)^2 h_l^{(1)}(2\rho\kappa)\right)^{-1}$ and $h_l^{(1)}$ denotes the spherical Hankel function of the first kind of order l. A Hilbert basis $\mathcal{HS}(L^2(S^2))$ is given by the operators $Z_{l_1m_1l_2m_2}$ defined by $Z_{l_1m_1l_2m_2}Y_l^m := \delta_{ll_1}\delta_{mm_1}Y_{l_2}^{m_2}$. Here $(l_1, m_1, l_2, m_2) \in \mathcal{I}$ with

$$\mathcal{I} := \{(l_1, m_1, l_2, m_2) \in \mathbb{N}_0 \times \mathbb{Z} \times \mathbb{N}_0 \times \mathbb{Z} : |m_1| \le l_1 \wedge |m_2| \le l_2\}.$$

A corresponding Hilbert basis of $\mathcal{HS}(L^2(\partial B_{2\rho}))$ by $\{\tilde{Z}_{l_1m_1l_2m_2} : (l_1, m_1, l_2, m_2) \in \mathcal{I}\}$ with $\tilde{Z}_{l_1m_1l_2m_2}\tilde{Y}_l^m := \delta_{ll_1}\delta_{mm_1}\tilde{Y}_{l_2}^{m_2}$. Since $\Omega\tilde{Z}_{l_1m_1l_2m_2} = \gamma_{l_1}^{m_1}\gamma_{l_2}^{m_2}Z_{l_1m_1l_2m_2}$, the singular values of Ω are $\sigma_{l_1m_1l_2m_2} := |\gamma_{l_1}^{m_1}||\gamma_{l_2}^{m_2}|$ with $(l_1, m_1, l_2, m_2) \in \mathcal{I}$.

It remains to derive an estimate of these singular values ordered by decreasing size with multiplicity. It follows from the asymptotic behaviour of the Hankel functions for large indices (cf. [2]) that there are constants c_1 and c_2 such that $|\gamma_l^m| < c_1 \exp(-c_2 l)$ for all $l \in \mathbb{N}_0$. This implies that $\sigma_{l_1m_1l_2m_2} < c_1^2 \exp(-c_2 J)$ if $l_1 + l_2 \le J$. Hence

$$\begin{aligned}&\#\{(l_1, m_1, l_2, m_2) \in \mathcal{I} : \sigma_{l_1m_1l_2m_2} \ge c_1 \exp(-c_2 J)\}\\ &\le \#\{(l_1, m_1, l_2, m_2) \in \mathcal{I} : l_1 + l_2 \le J\} =: N_J\end{aligned}$$

To compute N_J, note that $\#\{(l_2, m_2) \in \mathbb{N}_0 \times \mathbb{Z} : |m_2| \le l_2 \wedge l_2 \le K\} = \sum_{l_2=0}^K (2l_2 + 1) = (K + 1)^2$. Therefore,

$$N_J = \sum_{l_1=0}^J (2l_1 + 1) \cdot \#\{(l_2, m_2) \in \mathbb{N}_0 \times \mathbb{Z} : |m_2| \le l_2 \wedge l_2 \le J - l_1\}$$

$$= \sum_{l_1=0}^{J} (2l_1+1)(J-l_1+1)^2 \le \sum_{l_1=0}^{J} 2(J+1)^3 = 2(J+1)^4.$$

This is a coarse estimate, but it is optimal w.r.t. the power of $J+1$. It follows from the definition of N_J that $\sigma_{2(J+1)^4}(\Omega) \le \sigma_{N_J}(\Omega) < c_1^2 \exp(-c_2 J)$ (the largest singular value is $\sigma_0(\Omega)$ by definition!) For arbitrary $j \ge 2$ we choose $J \in \mathbb{N}_0$ such that $2(J+1)^4 \le j < 2(J+2)^4$. Then

$$\sigma_j(\Omega) \le \sigma_{2(J+1)^4}(\Omega) < c_1^2 \exp(-c_2 J) \le c_1^2 \exp(2c_2) \exp\left(-c_2 \left(\frac{j}{2}\right)^{1/4}\right).$$

This estimates also holds true for $j = 0, 1$ after possibly increasing the value of c_1. This proves the assertion. ∎

Remark 2.4 *A corresponding analysis in the two dimensional case yields the estimate* $\sigma_j(F'(a)) = \mathrm{O}(\exp(-cj^{1/2}))$, $j \to \infty$.

3. The implementation of the direct problem

In [23] Vainikko has suggested a fast solution method for eq. (5) which we describe in the following.

Periodization. Define $k(x) := \kappa^2 \Phi(x)$ for $|x_i| \le 2\rho$ $(i = 1, 2, 3)$ and $k(x) = 0$ else. We extend k, a, v, and $f := au_{\mathrm{i}}$ as 4ρ-periodic functions with respect to x_1, x_2, and x_3 outside of $G_{2\rho} := \{x \in \mathbb{R}^3 : |x_i| \le 2\rho, i = 1, 2, 3\}$. This gives the multiperiodic integral equation

$$v(x) + a(x) \int_{G_{2\rho}} k(x-y)v(y)\,\mathrm{d}y = f(x), \qquad x \in G_{2\rho} \tag{6}$$

which is equivalent to (5).

Discretization. Let

$$\varphi_j(x) := (4\rho)^{-3/2} \exp\left(\frac{\mathrm{i}\pi}{2\rho} j \cdot x\right), \qquad j \in \mathbb{Z}^3$$

denote the trigonometric orthonormal basis of $L^2(G_{2\rho})$ and let $\mathcal{T}_N := \operatorname{span}\{\varphi_j : j \in \mathbb{Z}_N^3\}$ where $\mathbb{Z}_N^3 := \{j \in \mathbb{Z}^3 : -\frac{N}{2} \le j_1, j_2, j_3 < \frac{N}{2}\}$. The orthogonal projection P_N of L^2 onto $\mathcal{T}_N$ is given by

$$P_N v := \sum_{j \in \mathbb{Z}_N^3} \hat{v}(j)\varphi_j, \qquad \hat{v}(j) := \int_{G_{2\rho}} v\overline{\varphi_j}\,\mathrm{d}x.$$

The trigonometric interpolation operator $Q_N : C(G_{2\rho}) \to \mathcal{T}_N$ (in this paper all function spaces on $G_{2\rho}$ consist of multi-periodic functions by definition!) is well defined by

$$Q_N v \in \mathcal{T}_N, \qquad (Q_N v)(hj) = v(hj), \quad j \in \mathbb{Z}_N^3, h = \frac{4\rho}{N}.$$

We approximate eq. (6) by

$$v_N + Q_N a K P_N v_N = Q_N f. \tag{7}$$

Here $(Kv)(x) := \int_{G_{2\rho}} k(x-y)v(y)\,\mathrm{d}y$. Deviating from [23] we have introduced the operator P_N in (7). Note that this does not change the solution v_N since any solution v_N to (7) (with or without P_N) belongs to $\mathcal{T}_N$. The advantages of this modification occur in the application of the two-grid method and the solution of the adjoint equation and will be discussed there.

The operator on the left hand side of (7) can be implemented efficiently by FFT. For $w \in C(G_{2\rho})$ let $\underline{w}_N$ denote the vector of nodal values $w(hj)$, $j \in \mathbb{Z}_N^3$, $h = \frac{4\rho}{N}$, and $\underline{\hat{w}}_N$ the vector of Fourier coefficients $\hat{w}(j)$, $j \in \mathbb{Z}_N^3$ of w. Moreover, define the matrix $\mathcal{F}_N := (\exp(-\frac{2\pi \mathrm{i}}{N} k \cdot j))_{j,k \in \mathbb{Z}_N^3}$. Then $\underline{\hat{v}}_N = (4\rho)^{3/2} N^{-3} \mathcal{F}_N \underline{v}_N$ and $\underline{v}_N = (4\rho)^{-3/2} \mathcal{F}_N^* \underline{\hat{v}}_N$ for $v \in \mathcal{T}_N$. The application of the matrices $\mathcal{F}_N$ and $\mathcal{F}_N^{-1}$ can be implemented by FFT in $\mathrm{O}(N^3 \log N)$ arithmetical operations. With $\underline{\hat{K}}_N := (4\rho)^{3/2} N^{-3} \underline{\hat{k}}_N$ eq. (7) can be written in matrix form as follows:

$$\underline{v}_N + \underline{a}_N \mathcal{F}_N^{-1} \underline{K}_N \mathcal{F}_N \underline{v}_N = \underline{f}_N \tag{8}$$

The second term in this formula is an abbreviation for $\mathrm{diag}(\underline{a}_N)\mathcal{F}_N^{-1}\,\mathrm{diag}(\underline{K}_N)\mathcal{F}_N \underline{v}_N$ where $\mathrm{diag}(\underline{w})$ denotes the diagonal matrix with diagonal $\underline{w}$. (eq. (8) corrects eq. (19) in [23]. The error occurs in the formula $K\varphi_j = \hat{k}(j)\varphi_j$ which should be $K\varphi_j = (4\rho)^{3/2}\hat{k}(j)\varphi_j$.) The equivalent formulation using Fourier coefficients is

$$\underline{\hat{v}}_N + \mathcal{F}_N(\underline{a}_N \mathcal{F}_N^{-1} \underline{K}_N \underline{\hat{v}}_N) = \underline{\hat{f}}_N.$$

The vector $\underline{K}_N$ can be computed explicitly. If $\pi|j| \neq R$, $R := 2\rho\kappa$, then

$$(\underline{K}_N)_j = \frac{R^2 N^{-3}}{\pi^2 |j|^2 - R^2} \left\{ 1 - \mathrm{e}^{R\mathrm{i}} \left(\cos(\pi|j|) - \mathrm{i}\frac{R}{\pi|j|} \sin(\pi|j|) \right) \right\}, \quad j \neq 0$$

$$(\underline{K}_N)_0 = N^{-3} \left(\mathrm{e}^{R\mathrm{i}}(1 - R\mathrm{i}) - 1 \right),$$

otherwise $(\underline{K}_N)_j = -\mathrm{i}2^{-1/2} R N^{-3} (1 - \mathrm{e}^{R\mathrm{i}} R^{-1} \sin(R))$.

Two-Grid Method. With the notation $T_N := Q_N a K P_N$ eq. (7) can be written as $v_N + T_N v_N = f_N$. After multiplying both sides of this equation by $(I + T_M)^{-1}$ with some $M < N$, a simple manipulation leads to the fixed point iteration

$$v_N^0 := (I + T_M)^{-1} f_N \tag{9a}$$

$$v_N^l := (I + T_M)^{-1} \left\{ (T_M - T_N) v_N^{l-1} + f_N \right\} \qquad l = 1, 2, \ldots \tag{9b}$$

At this point it is convenient to have the operator P_N in T_N since then $\mathcal{T}_N$ and $\mathcal{T}_N^\perp$ are invariant subspaces of T_N. Hence, in order to apply $(I + T_M)^{-1}$ to v using Fourier coefficients, we can leave the large Fourier coefficients $\underline{\hat{v}}(j)$, $j \in \mathbb{Z}_N^3 \backslash \mathbb{Z}_M^3$ unchanged and apply $(I + T_M)^{-1}$ to the small Fourier coefficients $\underline{\hat{v}}(j)$, $j \in \mathbb{Z}_M^3$. Another consequence is that $\mathcal{T}_N$ and $\mathcal{T}_N^\perp$ are also invariant under T_N^* which will be convenient for evaluating $F'(a)^* g$.

To write down the matrix form of (9*b*), we introduce the canonical restriction $R_{M,N} : \mathbb{C}^{\mathbb{Z}_N^3} \to \mathbb{C}^{\mathbb{Z}_M^3}$ which maps a vector $\underline{v} : \mathbb{Z}_N^3 \to \mathbb{C}$ to $R_{M,N}\underline{v} := \underline{v}|_{\mathbb{Z}_M^3}$. Note that the adjoint $R_{M,N}^* \mathbb{C}^{\mathbb{Z}_M^3} \to \mathbb{C}^{\mathbb{Z}_N^3}$ is a prolongation operator which extends a vector

$\underline{v} : \mathbb{Z}^3_M \to \mathbb{C}$ by 0's on $\mathbb{Z}^3_N \backslash \mathbb{Z}^3_M$, and that $I_N - R^*_{M,N} R_{M,N}$ is the orthogonal projection onto the subspace $\mathbb{C}^{\mathbb{Z}^3_N \backslash \mathbb{Z}^3_M}$ of large Fourier coefficients. We have

$$\begin{aligned}\underline{\hat{v}}^l_N := \Big((I_N - R^*_{M,N} R_{M,N}) + R^*_{M,N} (I_M + \mathcal{F}_M \underline{a}_M \mathcal{F}^{-1}_M \underline{K}_M)^{-1} R_{M,N} \Big) \\ \times \left\{ \left(\frac{N}{M} \right)^3 R^*_{M,N} \mathcal{F}_M \underline{a}_M \mathcal{F}^{-1}_M \underline{K}_M R_{M,N} \underline{\hat{v}}^{l-1}_N - \mathcal{F}_N \underline{a}_N \mathcal{F}^{-1}_N \underline{K}_N \underline{\hat{v}}^{l-1}_N + \underline{\hat{f}}_N \right\}\end{aligned}$$

Error Analysis. The error analysis is almost the same as for the case without P_N which has been treated in [23]. Let $H^s = H^s(G_{2\rho})$ denote the Sobolev space of multi-periodic functions on $G_{2\rho}$ of index $s \in \mathbb{R}$, and assume that $a \in H^s$, $s > \frac{3}{2}$. Then we have the estimate

$$\|aK - Q_N aK P_N\|_{\mathcal{L}(H^t, H^t)} \leq c \begin{cases} N^{t-s}, & s - \frac{3}{2} \leq t \leq s, \\ N^{-\frac{3}{2}}, & 0 \leq t \leq s - \frac{3}{2}. \end{cases} \tag{10}$$

This follows from results in [23] using

$$\|aKv - Q_N aK P_N v\|_t \leq \|aK\|_{\mathcal{L}(H^{t-3/2}, H^t)} \|v - P_N v\|_{t-3/2} + \|aK P_N v - Q_N aK P_N v\|_t.$$

From (10) the following theorem can be derived as in [23, Theorem 2]:

Theorem 3.1 *Assume that $a, f \in H^s(G_{2\rho})$, $s > \frac{3}{2}$ and $\operatorname{supp} a \subset B_\rho$. Then eq. (7) has a unique solution v_N for sufficiently large N and*

$$\|v - v_N\|_t \leq c \|v\|_t N^{t-s} \qquad \textit{for } \; 0 \leq t \leq s. \tag{11}$$

Note that (10) implies uniform boundedness of $\|(I + T_M)^{-1}\|$ for sufficiently large M. Hence we get the following estimate on the iteration operator in (9*b*) for $0 \leq t \leq s - \frac{3}{2}$:

$$\|(I + T_M)^{-1} (T_M - T_N)\|_{\mathcal{L}(H^t, H^t)} \leq c M^{-\frac{3}{2}}$$

This implies convergence of the two grid iteration for $cM^{-\frac{3}{2}} < 1$. Moreover, it can be shown in [23] that for $t = 0$ the order optimal accuracy $\mathrm{O}(N^{-s})$ can be achieved in just one iteration provided $M \sim N^\theta$ with some $\frac{s}{s+3/2} < \theta < 1$.

We solve the coarse-grid equations $(I + T_M)\tilde{g}_M = \tilde{f}_M$ by the conjugate gradient method applied to the normal equation (CGNE). It follows from the analysis in [23] that the total cost of solving the coarse grid equation to the accuracy of the machine error or the discretization error is $\mathrm{O}(M^3 \log M)$ or $\mathrm{O}(M^3 \log^2 M)$ arithmetical operations, respectively. We always solve the coarse grid equation up to machine accuracy for reasons explained below under "Implementation of the adjoint $F'(a)^*$". In both cases the total cost of the two-grid method is dominated by the operators $\mathcal{F}_N$ and $\mathcal{F}^*_N$ which require $\mathrm{O}(N^3 \log N)$ arithmetical operations. If we perform one two-grid iteration step, we need one fine-grid FFT transformation to compute $\underline{\hat{f}}_N$, two to compute $\underline{\hat{v}}^1_N$, and one to compute $\underline{v}^1_N$. This makes a total of 4 fine-grid FFT transformations as opposed to 6 fine-grid FFT transformations required by the algorithm in [23]. Moreover, we are more flexible in the choice of M and N since we do not need the requirement $N/M \in \mathbb{N}$.

Implementation of the adjoint $F'(a)^*$. The evaluation of $F'(a)^*g$ involves solving the adjoint equation to (7):

$$(I + T_N^*)w_N = f_N \tag{12}$$

Even though it is sufficient to solve (7) and (12) up to an accuracy of the order of the discretization error, it is advantageous in the context of CG and Lanczos methods that both approximations are adjoint to each other to a higher order (e.g. the machine error). In the L-step two-grid method, we use the following approximation to $(I + T_N)^{-1}$:

$$(I + T_N)^{-1} \approx \left\{ \sum_{l=0}^{L} \left((I + T_M)^{-1}(T_M - T_N) \right)^l \right\} (I + T_M)^{-1}$$

With a proper choice of L and M the approximation is of the order of the discretization error with respect to the matrix norm. In order to get an approximation to $(I + T_N^*)^{-1}$ we take the exact adjoint of the operator on the right hand side:

$$\begin{aligned}(I + T_N^*)^{-1} &\approx (I + T_M^*)^{-1} \left\{ \sum_{l=0}^{L} \left((T_M^* - T_N^*)(I + T_M^*)^{-1} \right)^l \right\} \\ &= \left\{ \sum_{l=0}^{L} \left((I + T_M^*)^{-1}(T_M^* - T_N^*) \right)^l \right\} (I + T_M^*)^{-1}\end{aligned}$$

This leads to the iteration scheme

$$\begin{aligned} w_N^l &:= (I + T_M^*)^{-1} \left\{ (T_M^* - T_N^*) w_N^{l-1} + f_N \right\} \qquad l = 1, \ldots, L \\ w_N^0 &:= (I + T_M^*)^{-1} f_N. \end{aligned}$$

Numerical results. As a test example we used the function

$$a(x) = \exp\left(\frac{-1}{1 - |x - \xi_1|^2}\right) - \frac{1}{2}\exp\left(\frac{-1}{1 - 2|x - \xi_2|^2}\right), \qquad \begin{aligned}\xi_1 &= (0.1, 0.2, 0.1)\\ \xi_2 &= (-0.4, -0.2, -0.1)\end{aligned}$$

with $\rho = 1$, $\kappa = 1$ and $N = 16, 32, 64, 128, 256$. We compare a one-step two-grid iteration to the CGNE iteration on the fine grid. Since it is necessary for the efficiency of FFT that the coarse grid constant M be a product of small prime factors, we choose M to be the smallest integer $>= N^{s/(s+3/2)}$, $s = 5$ such that M is a product of prime factors 2,3,5 and 7. For the FFT algorithm we used the implementation in [6].

Table 1 shows the computed values of u_∞ at the points $\hat{x} = d = (1, 0, 0)$. Fast convergence is clearly exhibited. The accuracy of the two-grid and the CGNE iteration are almost the same, but the two-grid iteration is much faster.

4. A Preconditioned Regularized Newton-CG method for exponentially ill-posed problems

In this section we consider a general operator equation

$$F(x) = y, \tag{13}$$

where $\mathcal{X} \supset D(F) \to \mathcal{Y}$ is an operator, which is Fréchet-differentiable on its domain $D(F)$, and $\mathcal{X}, \mathcal{Y}$ are Hilbert spaces. We assume that F is one-to-one. Let $x^\dagger$ denote the

		Two-Grid Iteration		CGNE Iteration	
N	M	$u_\infty(\hat{x}, d)$	time	$u_\infty(\hat{x}, d)$	time
16	9	0.1541043642553	0.04 s	0.1541073064510	0.09 s
32	15	0.1542513420299	0.33 s	0.1542514229206	0.86 s
64	25	0.1542512347807	3.58 s	0.1542512358846	14.9 s
128	42	0.1542535468730	26.2 s	0.1542535468878	135 s
256	72	0.1542535418092	203 s	0.1542535418093	1330 s

Table 1. Convergence of the direct solution method

exact solution to (13). We assume that the measured data y^δ are perturbed by noise such that

$$\|y - y^\delta\| \leq \delta. \tag{14}$$

Over the last years several iterative methods for solving ill-posed problems of the form (13) have been investigated intensively. Two well-known examples of such methods are Landweber iteration and the Levenberg-Marquardt algorithm. *Landweber iteration* is defined by

$$x_{n+1} := x_n + \mu F'(x_n)^*(y^\delta - F(x_n)) \tag{15}$$

where μ is a scaling parameter such that $\mu\|F'(x_n)\| \leq 1$ for all n. Applying Tikhonov regularization to the linearized operator equation (3) leads to the *Levenberg-Marquardt algorithm* given by $x_{n+1} := x_n + h_n$ with

$$h_n := \operatorname{argmin}_{h\in\mathcal{X}}(\|F'(x_n)h + F(x_n) - y^\delta\|^2 + \gamma_n\|h\|^2). \tag{16}$$

Here γ_n are regularization parameters which can be chosen by different rules, e.g. $\gamma_n = \gamma_0\overline{\gamma}^{-n}$ with some $\overline{\gamma} > 1$. An alternative is the iteratively regularized Gauss-Newton method (IRGNM, cf. [1, 13]) defined by

$$h_n := \operatorname{argmin}_{h\in\mathcal{X}}(\|F'(x_n)h + F(x_n) - y^\delta\|^2 + \gamma_n\|h + x_n - x_0\|^2). \tag{17}$$

The term $x_n - x_0$ has an additional regularizing effect and facilitates the convergence analysis.

Whereas one Landweber step is very cheap as it involves only one evaluation of $F(x_n)$ and one evaluation of $F'(x_n)^*g$, convergence is very slow. It turns out that the number of Landweber steps which is necessary to get the same accuracy as with N Newton steps grows exponentially (cf. [4])!

On the other hand, a straightforward implementation of (16) involves setting up the whole derivative matrix corresponding to $F'(x_n)$, i.e. evaluation of $F'(x_n)\varphi_j$ for all basis functions φ_j spanning the approximating subspace. For small-scale problems where a decomposition of the system matrix (i.e. the matrix representing the differential or integral equation) is an appropriate way of solving the direct problem, it has often been pointed out that the effort of computing the matrix for $F'(x_n)$ is acceptable. However, for the problem at hand the system matrix in (8) would not fit into the storage of a workstation, and even if we had a decomposition of that matrix, it would be inefficient to use it. Therefore, our only choice is to solve the direct problem by an iterative method, and therefore the computation of $F'(x_n)$ is much more expensive than the evaluation of $F(x_n)$. This is the typical situation for large-scale problems.

An obvious way to reduce the computation time is to use a matrix $F'(x_*)$ for some x_* for several Newton steps once it has been computed. A better alternative is to avoid setting up a derivative matrix altogether and solve the linearized equations

$$F'(x_*)h_n = y^\delta - F(x_n) \tag{18}$$

iteratively. As mentioned before, we have to use regularization methods since (18) is ill-posed. One possibility is to use the regularizing properties of Landweber iteration ([16]), ν-methods ([21]) or CGNE ([12]). Among these methods, Newton-CG is most attractive since for Landweber iteration and ν-methods the number of inner iterations grows exponentially. But the Newton-CG method also has some drawbacks: It is difficult to analyze due to the nonlinear character of CGNE, and numerical experience shows that convergence is typically much slower than for Levenberg-Marquardt in the long run (cf. [14]).

In order to avoid these disadvantages, we first regularize (18) and apply CG to the regularized problem (16) instead of using the regularizing properties of CGNE. (16) is equivalent to finding the unique least-squares solution to

$$\begin{pmatrix} F'(x_*) \\ \sqrt{\gamma_n} I \end{pmatrix} h_n = \begin{pmatrix} y^\delta - F(x_n) \\ \sqrt{\gamma_n} b_n \end{pmatrix}. \tag{19}$$

with $b_n = 0$ for Levenberg-Marquardt and $b_n = x_0 - x_n$ for the IRGNM. Introducing the operator

$$T_n := \begin{pmatrix} F'(x_*) \\ \sqrt{\gamma_n} I \end{pmatrix} : \mathcal{X} \to \mathcal{Y} \times \mathcal{X}$$

and $g_n = (y^\delta - F(x_n), \sqrt{\gamma_n} b_n)^{\mathrm{t}}$, eq. (19) becomes $T_n h_n = g_n$. We will show below how to construct a preconditioner $M_n = C_n^2 \in \mathcal{L}(\mathcal{X})$ for the operator $T_n^* T_n = F'(x_*)^* F'(x_*) + \gamma_n I$. Here M_n and C_n are self-adjoint and positiv definite. The CG iteration is applied to the preconditioned normal equation

$$C_n^{-1} T_n^* T_n C_n^{-1} \tilde{h}_n = C_n^{-1} T_n^* g_n, \qquad \tilde{h}_n = C_n h_n. \tag{20}$$

This can be coded without explicit reference to C_n^{-1} and $\tilde{h}_n$ as follows:

$$
\begin{aligned}
&h_n^0 = 0;\ d^0 = g_n;\ r^0 = T_n^* d^0;\ p^1 = z^0 = M_n^{-1} r^0;\ k = 0;\\
&\text{while } \|r^k\| > \epsilon\gamma_n \|h_n^k\| \text{ do}\\
&\qquad k = k+1;\\
&\qquad q^k = T_n p^k;\\
&\qquad \alpha_k = (r^{k-1}\cdot z^{k-1})/\|q^k\|^2;\\
&\qquad h_n^k = h_n^{k-1} + \alpha_k p^k;\\
&\qquad d^k = d^{k-1} - \alpha_k q^k;\\
&\qquad r^k = T_n^* d^k;\\
&\qquad z^k = M_n^{-1} r^k;\\
&\qquad \beta_k = (r^k \cdot z^k)/(r^{k-1}\cdot z^{k-1});\\
&\qquad p^{k+1} = z^k + \beta_k p^k;
\end{aligned}
\tag{21}
$$

Of course, if $M_n = I$ then $z^k = r^k$, and we can dispense with the vectors z^k. Recall the following elementary properties of the CG iteration:

$$d^k = g_n - T_n h_n^k \tag{22a}$$

$$r^k = T_n^*(g_n - T_n h_n^k) \tag{22b}$$

$$\left\langle C_n^{-1} r_j, C_n^{-1} r_k \right\rangle = 0 \qquad \text{for } j \neq k \tag{22c}$$

$$\langle q_j, q_k \rangle = 0 \qquad \text{for } j \neq k \tag{22d}$$

We can show the following error estimates:

Proposition 4.1 *Assume that the preconditioner $M_n \in \mathcal{L}(\mathcal{X})$ is symmetric, positive definite and boundedly invertible. Then the iteration (21) terminates after a finite number K of steps, and we have the error estimate*

$$\|h_n^K - h_n\| \leq \epsilon \|h_n^K\| \leq \frac{\epsilon}{1-\epsilon}\|h_n\|. \tag{23}$$

Proof: It follows from standard convergence theory of CG (cf., e.g., [5]) that $\lim_{k\to\infty} h_n^k = h_n$ and $\lim_{k\to\infty} r^k = 0$ unless the iteration stops at the exact solution, i.e. $r_k = 0$ for some k. Therefore, the stopping criterion $\|r^k\| \leq \epsilon\gamma_n\|h_n^k\|$ is met after a finite number K of steps. The least-squares solution h_n to (19) satisfies the normal equation $T_n^* T_n h_n = T_n^* g_n$. It follows from (22$b$) and $\|r^K\| \leq \epsilon\gamma_n \|h_n^K\|$ that

$$\|T_n^* T_n (h_n - h_n^K)\| \leq \epsilon\gamma_n \|h_n^K\|.$$

Since $\langle T_n^* T_n x, x\rangle = \|F'(x_*)x\|^2 + \gamma_n\|x\|^2 \geq \gamma_n \|x\|^2$, we have $\|(T_n^* T_n)^{-1}\| \leq \gamma_n^{-1}$. Hence

$$\|h_n - h_n^K\| \leq \|(T_n^*T_n)^{-1}\| \|T_n^* T_n (h_n - h_n^K)\| \leq \epsilon \|h_n^K\|.$$

The second inequality follows from $\|h_n^K\| \leq \|h_n\| + \|h_n^K - h_n\| \leq \|h_n\| + \epsilon\|h_n^K\|$. ■

By choosing ϵ we can control the error that occurs from solving (19) approximately. This will be very useful for a convergence analysis of our method.

In the general case a lot of CG steps may be necessary to meet the stopping criterion since the condition number $\mathrm{cond}(T_n^* T_n) = (\gamma_n + \|F'(x_*)\|^2)\gamma_n^{-1}$ grows rapidly. However, if the singular values of $F'(x_*)$ decay like $\sigma_j(F'(x_*)) = \mathrm{O}(\exp(-cj^\sigma))$, then only $\mathrm{O}(n^{1/\sigma})$ singular values of T_n lie outside of the interval, say, $[\sqrt{\gamma_n}, \sqrt{2\gamma_n}]$ (at least for the choice $\gamma_n = \gamma_0\overline{\gamma}^{-n}$). Therefore, we expect that asymptotically only $\mathrm{O}(n^{1/\sigma})$ CG iterations are needed.

To construct a preconditioner $M_n = C_n^2$ for $T_n^*T_n$, we take advantage of the close connection between the CG iteration and the *Lanczos method.* Quantities arising in the iteration (21) with $M_n = I$ can be used to compute approximations to the largest eigenvalues and -vectors of $T_n^*T_n$ as follows: Since the vectors $\hat{r}^j := r^j/\|r^j\|$ are orthonormal due to (22c), the operator $R_k = [\hat{r}^0, \ldots, \hat{r}^{k-1}] : \mathbb{R}^k \to \mathcal{X}$ defined by $R_k\xi := \sum_{j=0}^{k-1}\xi_j\hat{r}^j$ is isometric. Analogously, the vectors $\hat{q}^j := q^j/\|q^j\|$ define an isometric operator $Q_k = [\hat{q}^1, \ldots, \hat{q}^k] : \mathbb{R}^k \to \mathcal{Y}$. Multiplying the identity $r^j = p^{j+1} - \beta_j p^j$ from the left by $\|r_j\|^{-1}T_n$ and using the relations $T_n p^j = q^j$, $\|q^{j+1}\|/\|r^j\| = \sqrt{1/\alpha_{j+1}}$ and $\|q^j\|/\|r^j\| = (\alpha_j\beta_j)^{-1/2}$ yields

$$T_n\hat{r}^j = \sqrt{\frac{1}{\alpha_{j+1}}}\hat{q}^{j+1} - \sqrt{\frac{\beta_j}{\alpha_j}}\hat{q}^j \qquad (j = 1, \ldots, k-1), \qquad T_n\hat{r}_0 = \sqrt{\frac{1}{\alpha_1}}\hat{q}_1$$

or

$$T_n R_k = Q_k B_k, \qquad B_k := \begin{bmatrix} \sqrt{\frac{1}{\alpha_1}} & -\sqrt{\frac{\beta_1}{\alpha_1}} & & \cdots & 0 \\ 0 & \sqrt{\frac{1}{\alpha_2}} & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & -\sqrt{\frac{\beta_{k-2}}{\alpha_{k-2}}} \\ 0 & \cdots & & 0 & \sqrt{\frac{1}{\alpha_{k-1}}} \end{bmatrix}.$$

Analogously, the identity $\alpha_j q^j = d^{j-1} - d^j$ multiplied from the left by $(\|q^j\|\alpha_j)^{-1}T_n^*$ together with the relations $T_n^* d^j = r^j$, $\|r^{j-1}\|/(\|q^j\|\alpha_j) = \sqrt{1/\alpha_j}$ and $\|r^j\|/(\|q^j\|\alpha_j) = \sqrt{\beta_j/\alpha_j}$ yields

$$T_n^*\hat{q}^j = \sqrt{\frac{1}{\alpha_j}}\hat{r}^{j-1} - \sqrt{\frac{\beta_j}{\alpha_j}}\hat{r}^j \qquad (j = 1, \ldots, k)$$

or $T_n^*Q_k = R_k B_k^{\mathrm{t}} - \sqrt{\frac{\beta_k}{\alpha_k}}\hat{r}^k e_k^{\mathrm{t}}$. Here $e_k^{\mathrm{t}} : \mathbb{R}^k \to \mathbb{R}$ is defined by $e_k^{\mathrm{t}}(\xi_j)_{j=1:k} := \xi_k$. Putting both together and using $e_k^{\mathrm{t}}B_k = \alpha_{k-1}^{-1/2}e_k^{\mathrm{t}}$ yields

$$T_n^*T_n R_k = T_n^*Q_k B_k = R_k B_k^{\mathrm{t}} B_k - \frac{1}{\|q^k\|\alpha_k^{3/2}} r^k e_k^{\mathrm{t}}.$$

Let us first consider the case $r^k = 0$. Let $B_k^{\mathrm{t}}B_k = V_k D_k V_k^{\mathrm{t}}$ be an eigenvalue decomposition of the tridiagonal matrix $B_k^{\mathrm{t}}B_k$, i.e. V_k orthogonal and D_k diagonal. Then $T_n^*T_n R_k V_k = (R_k V_k)D_k$, so $R_k V_k e_j$, $j = 0, \ldots, k-1$ are eigenvectors of $T_n^*T_n$ with eigenvalues $(D_k)_{jj}$. In the typical case $r^k \neq 0$ the vectors $R_k V_k e_j$ usually converge rapidly to the largest eigenvectors of $T_n^*T_n$. This is made precise in the Kaniel-Paige theory (cf., e.g., [7] and the references therein). Since $T_n^*T_n = F'(x_*)^*F'(x_*) + \gamma_n I$, eigenvectors of $T_n^*T_n$ are also eigenvectors of $F'(x_*)^*F'(x_*)$, and the eigenvalues are shifted by γ_n.

Assume that $\lambda_1 \geq \lambda_2 \geq \ldots$ are the eigenvalues of $F'(x_*)^* F'(x_*)$ with corresponding eigenvectors v_j. We define M_n as follows:

$$M_n x := \gamma_n x + \sum_{j=1}^{k} \lambda_j \langle x, v_j \rangle v_j \tag{24}$$

This implies that

$$M_n^{-1} x = \gamma_n^{-1} x + \sum_{j=1}^{k} \left(\frac{1}{\gamma_n + \lambda_j} - \frac{1}{\gamma_n} \right) \langle x, v_j \rangle v_j$$

M_n is designed such that $C_n^{-1} T_n^* T_n C_n^{-1} v_j = v_j$ for $j = 1, \ldots, k$, i.e. the large eigenvalues are shifted to 1. The spectrum and the condition number of $C_n^{-1} T_n^* T_n C_n^{-1}$ are $\sigma(C_n^{-1} T_n^* T_n C_n^{-1}) = \{1\} \cup \{1 + \frac{\lambda_j}{\gamma_n} : j > k\}$ and $\mathrm{cond}(C_n^{-1} T_n^* T_n C_n^{-1}) = 1 + \frac{\lambda_{k+1}}{\gamma_n}$. If λ_{k+1} is not much larger than γ_n, then the condition number of $C_n^{-1} T_n^* T_n C_n^{-1}$ is small and very few CG steps are necessary to solve (20).

Our algorithm can be summarized as follows:

```
Compute F(x_0);
n = 0;
while (||F(x_n) − y^δ|| ≥ τδ)
    if (√(n+1) ∈ ℕ)
        x_* = x_n;
        Solve T_n h_n = g_n by (21) with M_n = I;
        Compute the largest eigenvalues and -vectors of F'(x_*)^* F'(x_*);
    else
        Solve T_n h_n = g_n by (21) with M_n given by (24);
    x_{n+1} = x_n + h_n;
    n = n + 1;
    Compute F(x_n);
```

Note that we used the discrepancy principle (with $\tau = 2$) to stop the outer iteration. A new preconditioner is constructed whenever $n + 1$ is a square number. This is, of course, just a first approach, and it may be worth looking for other alternatives. We used the value $\epsilon = \frac{1}{3}$ in (21) for the preconditioned iteration, and $\epsilon = 10^{-9}$ when computing eigenvectors and -values. To avoid loss of orthogonality of the vectors r^k, which is a well-known problem of Lanczos iteration, we used a complete reorthogonalization as described in [7, Section 9.2] in each Lanczos/CGNE step.

5. Discretization of $\mathcal{X}$ and $\mathcal{Y}$

Let $\mathcal{X}_k \subset \mathcal{X}$ be a finite dimensional subspace with basis $\mathcal{X}_k = \mathrm{span}\{\varphi_1, \ldots, \varphi_{d_k}\}$, and let $Q_m : \mathcal{Y} \to \mathbb{R}^{d_m}$ be a sequence of linear mappings such that $|Q_m y|_2 \to \|y\|$ for all $y \in R(F)$. E.g., $Q_m y$ may be the coefficient vector $(\langle y, \psi_i \rangle)_{i=1,\ldots,d_m}$ with respect to some orthonormal system $\{\psi_1, \ldots, \psi_{d_m}\} \subset Y$. Alternatively, Q_m may be a collocation operator. We replace (13) by the finite dimensional equation

$$Q_m F(x) = Q_m y, \qquad x \in \mathcal{X}_k \tag{25}$$

and briefly discuss how to apply the method of the previous section to (25). Let $x \in \mathcal{X}_k$ be represented by its coordinate vector $\underline{x}$ with respect to the basis $\{\varphi_j\}$, i.e. $x = \sum_{j=1}^{d_k} \underline{x}_j \varphi_j$, and let $A_z \in \mathbb{R}^{d_m \times d_k}$ $(z \in \mathcal{X}_k)$ be defined by $Q_m F'[z]x = A_z \underline{x}$. To compute the norm of x and the matrix of the adjoint operator $(Q_m F'[z])^*$ we need the Gramian matrix $G := [\langle \varphi_i, \varphi_j \rangle_{\mathcal{X}}]$. A simple computation yields $\|x\|^2 = \underline{x}^{\mathrm{t}} G \underline{x}$. Let $G = B^{\mathrm{t}} B$ be a decomposition of G, e.g. the Cholesky decomposition. Then B defines a coordinate transformation $\tilde{\underline{x}} := B\underline{x}$ such that the discrete norm and the continuous norm coincide: $\|x\| = |\tilde{\underline{x}}|_2$. $\tilde{x}$ is the coefficient vector of x with respect to the orthonormal basis $\{\tilde{\varphi}_1, \ldots, \tilde{\varphi}_{d_k}\}$, $\tilde{\varphi}_i := \sum_j (B^{-1})_{ij} \varphi_j$. With respect to this basis, $Q_m F'(z)$ is represented by $A_z B^{-1}$, i.e. $Q_m F'(z)x = A_z B^{-1} \tilde{\underline{x}}$. Then $(Q_m F'[z])^*$ is represented by $B^{-\mathrm{t}} A_z^{\mathrm{t}}$, i.e. $(Q_m F'[z])^* g = \sum_{j=1}^{d_k} (B^{-\mathrm{t}} A_z^{\mathrm{t}} g)_j \tilde{\varphi}_j$. Therefore, we compute the coefficient vector $\tilde{\underline{h}}_n$ of the update $h_n = \sum_{j=1}^{d_k} (\tilde{\underline{h}}_n)_j \tilde{\varphi}_j$ in (21) using the substitutions $T_n \to \begin{pmatrix} A_z B^{-1} \\ \sqrt{\gamma_n} I \end{pmatrix}$ and $T_n^* \to \begin{pmatrix} B^{-\mathrm{t}} A_z^{\mathrm{t}} \\ \sqrt{\gamma_n} I \end{pmatrix}$. Since $\{\varphi_j\}$ may be a better basis to represent x_n than $\{\tilde{\varphi}_j\}$, we have to transform back at the end by computing $\underline{x}_{n+1} = \underline{x}_n + B^{-1} \tilde{\underline{h}}_n$.

In [2, Section 3.6] a quadrature rule

$$\int_{S^2} f(x)\, \mathrm{d}s(x) \approx \sum_{j=1}^{2m^2} \alpha_j f(x_j)$$

with $2m^2$ quadrature points $x_j \in S^2$ and quadrature weights $\alpha_j \in \mathbb{R}$ is constructed based on the Gauss-Legendre and the trapezoidal rule. It converges exponentially for analytic functions. Since farfield patterns $v_\infty(\hat{x}, d)$ are known to be analytic in both variables, $\int_{S^2} \int_{S^2} |v_\infty(\hat{x}, d)|^2 \, \mathrm{d}s(\hat{x})\, \mathrm{d}s(d) \approx \sum_{i=1}^{2m^2} \sum_{j=1}^{2m^2} \alpha_i \alpha_j |v_\infty(x_i, x_j)|^2$ is an excellent approximation. Therefore, the definition $(Q_m v_\infty)_{ij} := \sqrt{\alpha_i \alpha_j}\, v_\infty(x_i, x_j)$, $1 \le i, j \le m$ yields $\|v_\infty\|_{L^2} \approx |Q_m v_\infty|_2$.

When choosing the subspace $\mathcal{X}_k$ and basis functions $\{\varphi_1, \ldots, \varphi_{d_k}\} \subset \mathcal{X}_k$, we want to meet the following requirements:

(i) The basis functions should be smooth in order to guarantee fast convergence of the direct solver.

(ii) A factorization $G = BB^{\mathrm{t}}$ of the Gram matrix should be available, and the application of B, B^{-1} and $B^{-\mathrm{t}}$ should be cheap.

(iii) The basis functions φ_j should have good approximation properties.

For finite elements requirements (i) are (iii) are usually not satisfied. Therefore, it is not advisible to use finite elements in connection with the method discussed in Section 3. Since this method is based on periodization, one may think of using trigonometric interpolation. This works fine for $s = 0$, i.e. $X = L^2(B_\rho)$. However, the case $s = 0$ (which is not covered by our theoretical investigations) yields only poor reconstructions (cf. Fig. 4). For the better choices $s = 1, 2, \ldots$ the Lagrange functions of trigonometric interpolation are not in $H_0^s(B_\rho)$, and it is not clear how to compute the Gram matrix G of (any smoothly cut-off version of) these Lagrange functions. Hence requirement (ii) is violated.

One possibility to meet all the requirements above is to use tensor products with respect to spherical coordinates. We expand $a \in L^2(B_\rho)$ into spherical harmonics using the notations $r = |x|$ and $\hat{x} = x/|x|$:

$$a(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_{lm}(r) Y_l^m(\hat{x}), \qquad R_{lm}(r) := \int_{S^2} a(r\hat{x}) \overline{Y_l^m(\hat{x})} \, \mathrm{d}s(\hat{x})$$

Note that the functions R_{lm}, extended to $r < 0$ by the definition above, are even (odd) for l even (odd) since the Y_l^m's are even (odd). Moreover, $a \in C^L(B_\rho)$ implies $R_{ml} \in C^L((0,\rho))$. Near the origin the following result holds true:

Lemma 5.1 *Let a be L times continuously differentiable at 0. Then there are coefficients $c_{lm}^k \in \mathbb{C}$ such that for all $l \leq L$ and $m = -l, \ldots, l$*

$$R_{lm}(r) = \sum_{0 \leq k \leq (L-l)/2} c_{lm}^k r^{l+2k} + \mathrm{o}\left(r^L\right), \quad r \to 0. \tag{26}$$

Proof: The Taylor formula is $a(x) = \sum_{|\alpha| \leq L} \frac{(D^\alpha a)(0)}{\alpha!} x^\alpha + \mathrm{o}\left(|x|^L\right)$ for $|x| \to 0$. To prove the assertion, it suffices to establish the existence of coefficients $c_{lm}^k \in \mathbb{C}$ such that

$$\sum_{|\alpha| \leq L} \frac{(D^\alpha a)(0)}{\alpha!} x^\alpha = \sum_{0 \leq k \leq (L-l)/2} c_{lm}^k r^{l+2k} Y_l^m(\hat{x}).$$

Note that the functions $r^{l+2k} Y_l^m(\hat{x})$ are linearly independent polynomials since both r^{2k} and $r^l Y_l^m(\hat{x})$ are polynomials and the Y_l^m's are linearly independent. For $K \in \mathbb{N}_0$ let $B_K := \{r^{l+2k} Y^m(\hat{x}) : l + 2k = K; l, k \geq 0; |m| \leq l\}$. It is easy to show that $\#(B_K) = (K+1)(K+2)/2$, and that $(K+1)(K+2)/2$ is also the dimension of the space of homogeneous polynomials of degree K. Therefore, B_K is a basis of this space, and this implies the existence of the coefficients c_{lm}^k. ■

This lemma motivates the definition of the following spline spaces: Given $t_0 = 0 < t_1 < \ldots < t_N = \rho$, $\sigma \in \mathbb{N}$, and $l \in \{0, 1, \ldots\}$, let

$$S_l^\sigma(\{t_j\}) := \left\{ f \in C^{\sigma-1}([0,\infty)) : \begin{array}{l} \forall j \; f|_{[t_j,t_{j+1}]} \in \Pi^\sigma; \; f|_{[\rho,\infty)} \equiv 0; \\ (r^{-l} f(r))|_{(t_0,t_1]} \text{ is bounded and even} \end{array} \right\}.$$

Here Π^σ denotes the spaces of univariate polynomials of degree $\leq \sigma$. Note that for $l > \sigma$ we have $f|_{[t_0,t_1]} \equiv 0$ for all $f \in S_l^\sigma(\{t_j\})$. Given $\sigma, \{t_j\}$ as above and $L \in \{0, 1, \ldots\}$, we define an approximating subspace of $H_0^{\sigma-1}(B_\rho)$ by

$$W^\sigma(L, \{t_j\}) := \left\{ \sum_{l=0}^{L} \sum_{m=-l}^{l} R_{ml} \otimes Y_l^m : R_{ml} \in S_l^\sigma(\{t_j\}) \right\}$$

Here $R_{ml} \otimes Y_l^m$ denotes the function $x \mapsto R_{ml}(r) Y_l^m(\hat{x})$ where $r = |x|$ and $\hat{x} = x/|x|$ as above. Obviously $W^\sigma(L, \{t_j\})$ inherits the good approximation properties of splines and spherical harmonics, and $W^\sigma(L,\{t_j\}) \subset C_0^{\sigma-1}(B_\rho) \cap H_0^\sigma(B_\rho)$. Therefore, the requirements (i) and (iii) are satisfied.

To construct a basis of $S_l^\sigma(\{t_j\})$ we add points $t_{-\sigma} < \ldots < t_{-1} < t_0$ to the set $\{t_j\}$ and start with the usual B-Splines $B_{i,\sigma+1}$ supported on $[t_i, t_{i+\sigma+1}]$ $(i = -\sigma, \ldots, N-\sigma-1)$. Then we use the Marsden identity to define linear combinations of these B-splines satisfying $\tilde{B}_{-i,\sigma+1}(x) = x^i$, for $x \in [t_0, t_1]$ and $i = 0, \ldots, \sigma$:

$$\tilde{B}_{-i,\sigma+1} := \sum_{j=-\sigma}^{0} (-1)^{\sigma-i} \psi_{j,\sigma}^{(\sigma-i)}(0) \frac{(i)!}{\sigma!} B_{j,\sigma+1}, \qquad \psi_{i,\sigma}(\xi) := \prod_{j=1}^{\sigma} (t_{i+j} - \xi)$$

For $j = 1, \ldots, N-\sigma-1$ we define $\tilde{B}_{j,\sigma+1} := B_{j,\sigma+1}$. Then the set $\bigcup_{0 \le k \le \frac{\sigma-l}{2}} \{\tilde{B}_{-l-2k,\sigma+1}\} \cup \{\tilde{B}_{1,\sigma+1}, \tilde{B}_{2,\sigma+1}, \ldots, \tilde{B}_{N-\sigma-1,\sigma+1}\}$ is a basis of $S_l^\sigma(\{t_j\})$.

In order to compute the Gram matrix, we use the norm

$$\|v\|_{H_0^s(B_\rho)} := \|(I-\Delta)^{s/2} v\|_{L^2(B_\rho)}, \qquad s \in \mathbb{N}_0.$$

Since $\Delta_{S^2} Y_l^m = -l(l+1) Y_l^m$ (Δ_{S^2} denotes the Laplace-Beltrami operator on S^2), we have

$$\begin{aligned}(I-\Delta)\left(\tilde{B}_{i,\sigma+1} \otimes Y_l^m\right) &= \left(I - r^{-2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - r^{-2}\Delta_{S^2}\right)\left(\tilde{B}_{i,r+1} \otimes Y_l^m\right) \\ &= \left(D_l \tilde{B}_{i,r+1}\right) \otimes Y_l^m\end{aligned}$$

with $D_l := I - r^{-2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{l(l+1)}{r^2} I$. For s even, $s \le \sigma$ we get

$$\left\langle \tilde{B}_{i,\sigma+1} \otimes Y_l^m, \tilde{B}_{i',\sigma+1} \otimes Y_{l'}^{m'} \right\rangle_{H_0^s(B_\rho)} = \left\langle D_l^{s/2} \tilde{B}_{i,\sigma+1}, D_{l'}^{s/2} \tilde{B}_{i',\sigma+1} \right\rangle_{L^2(0,\rho)} \delta_{ll'} \delta_{mm'},$$

and for s odd, $s+1 \le \sigma$

$$\left\langle \tilde{B}_{i,\sigma+1} \otimes Y_l^m, \tilde{B}_{i',\sigma+1} \otimes Y_{l'}^{m'} \right\rangle_{H_0^s(B_\rho)} = \left\langle D_l^{(s+1)/2} \tilde{B}_{i,\sigma+1}, D_{l'}^{(s-1)/2} \tilde{B}_{i',\sigma+1} \right\rangle_{L^2(0,\rho)} \delta_{ll'} \delta_{mm'},$$

It follows that G is a block-diagonal matrix where each block corresponds to one couple (l,m). Moreover, due to the local character of the B-splines, each block is band-limited. Therefore, requirement (ii) is satisfied.

6. Numerical results

We first look at the effect of preconditioning on the number of inner CG iterations. In two space dimensions the preconditioned equations can be solved by a very small number of CG iterations (≤ 4) whereas the unpreconditioned equations require much more CG iterations (cf. Figure 1). The error is almost the same with and without preconditioning. In three space dimensions the number of inner CG iterations is larger since the number of relevant singular values of $F'[a_n]$ grows much more rapidly with the outer iteration number n (cf. Proposition 2.3 and Remark 2.4). Nevertheless, the total number of operator evaluations is reduced remarkably by using the preconditioner.

Figure 2 shows the reconstruction of a real-valued refractive index with almost exact data ($\delta = 10^{-7}$). We used the parameters $\kappa = 1$, $\rho = 1$, and $m = 4$, i.e. $2m^2 = 32$ collocation points and incident waves. The direct scattering problems were solved by a two-grid iteration with $N = 64$ and $M = 32$. For the approximating subspace $\mathcal{X}_k$ we used spherical harmonics up to the order $L = 15$ and splines of order $\sigma = 3$ on an equidistant mesh with $N + 1 = 21$ knots. To reduce the dimension of $\mathcal{X}_k$ we restricted the spline spaces $S_l^\sigma(\{t_j\})$ with $l > \sigma$ by the additional condition $f|_{[t_0,t_\lambda]} \equiv 0$, $\lambda = 1+\left[(l-\sigma-1)\frac{N-\sigma-1}{L-\sigma-1}\right]$, and obtained $\dim \mathcal{X}_k = 2060$. The regularization parameters in the preconditioned iteratively regularized Gauss-Newton method were chosen to be $\gamma_n = 0.1 \cdot 2^{-n}$. To avoid an inverse crime when creating synthetic data, we used a periodization interval of 2.2ρ instead of 2ρ and $N = 128$, $M = 48$. Moreover, the exact refractive index was not in the approximating subspace since a was identically 0 on large parts of B_ρ. The computation time for this very accurate reconstruction was 36 hours on an SGI R12000 machine.

If the noise level δ is larger or if less accurate reconstruction are required, the computation time is much shorter. Figure 3 shows a reconstruction with 1% data noise error which took only 20 minutes. This example shows that it is also possible to reconstruct large and complex valued refractive indices. In particular, the Born approximation is no longer valid for this example.

Acknowledgments

This work has been supported by the Austrian Science Fund (FWF) under the project SFB F013 "Numerical and Symbolic Scientific Computing".

References

[1] A. B. Bakushinskii. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Maths. Math. Phys.*, 32:1353–1359, 1992.

[2] D. Colton and R. Kreß. *Inverse Acoustic and Electromagnetic Scattering.* Springer Verlag, Berlin Heidelberg New York, second edition, 1997.

[3] D. Colton and P. Monk. The numerical solution of the three dimensional inverse scattering problem for time-harmonic acoustic waves. *SIAM J. Sci. Stat. Comp.*, 8:278–291, 1987.

[4] P. Deuflhard, H. Engl, and O. Scherzer. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions. *Inverse Problems*, 14:1081 – 1106, 1998.

[5] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publisher, Dordrecht Boston London, 1996.

[6] M. Frigo and S. G. Johnson. *FFTW User's Manual.* MIT, Boston, May 1999. available on `http://www.fftw.org/`.

[7] G. H. Golub and C. F. van Loan. *Matrix Computations.* The John Hopkins University Press, Baltimore, second edition, 1983.

[8] S. Gutman and M. Klibanov. Regularized quasi-Newton method for inverse scattering problems. *Math. Comput. Modelling*, 18:5–31, 1993.

[9] S. Gutman and M. Klibanov. Iterative methods for multi-dimensional inverse scattering problems at fixed frequencies. *Inverse Problems*, 10:5–31, 1994.

[10] P. Hähner. *On Acoustic, Electromagnetic, and Elastic Scattering Problems in Inhomogeneous Media.* Habilitation thesis, Göttingen, 1998.

[11] P. Hähner and T. Hohage. New stability estimates for the inverse acoustic inhomogeneous medium problem and applications. in preparation.

[12] M. Hanke. Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems. *Numer. Funct. Anal. Optim.*, 18:971–993, 1997.

[13] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauß-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13:1279–1299, 1997.

[14] T. Hohage. *Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially Ill-Posed Problems.* PhD thesis, University of Linz, 1999.

[15] T. Hohage. Regularization of exponentially ill-posed problems. *Numer. Funct. Anal. Optim.*, 21:439–464, 2000.

[16] B. Kaltenbacher. Some Newton-type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13:729–753, 1997.

[17] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Springer, New York, Berlin, Heidelberg, 1996.

[18] R. E. Kleinman and P. M. van den Berg. A modified gradient method for two-dimensional problems in tomography. *J. Comput. Appl. Math.*, 42:17–36, 1992.

[19] K. J. Langenberg. Applied inverse problems for acoustic, electromagnetic and elastic wave scattering. In Sabatier, editor, *Basic Methods of Tomography and Inverse Problems*, pages 127–467. Adam Hilger, Bristol and Philadelphia, 1987.

[20] A. Nachman. Reconstructions from boundary measurements. *Annals of Math.*, 128:531–576, 1988.

[21] A. Rieder. On the regularization of nonlinear ill-posed problems via inexact Newton iterations. *Inverse Problems*, 15:309–327, 1999.

[22] M. Taylor. *Partial Differential Equations*, volume 1. Springer Verlag, New York, 1996.

[23] G. M. Vainikko. Fast solvers of the Lippmann-Schwinger equation. In R. P. Gilbert, J. Kajiwara, and Y. S. Xu, editors, *Direct and Inverse Problems of Mathematical Physics.* Kluwer Acad. Publ., Dordrecht, 1999.

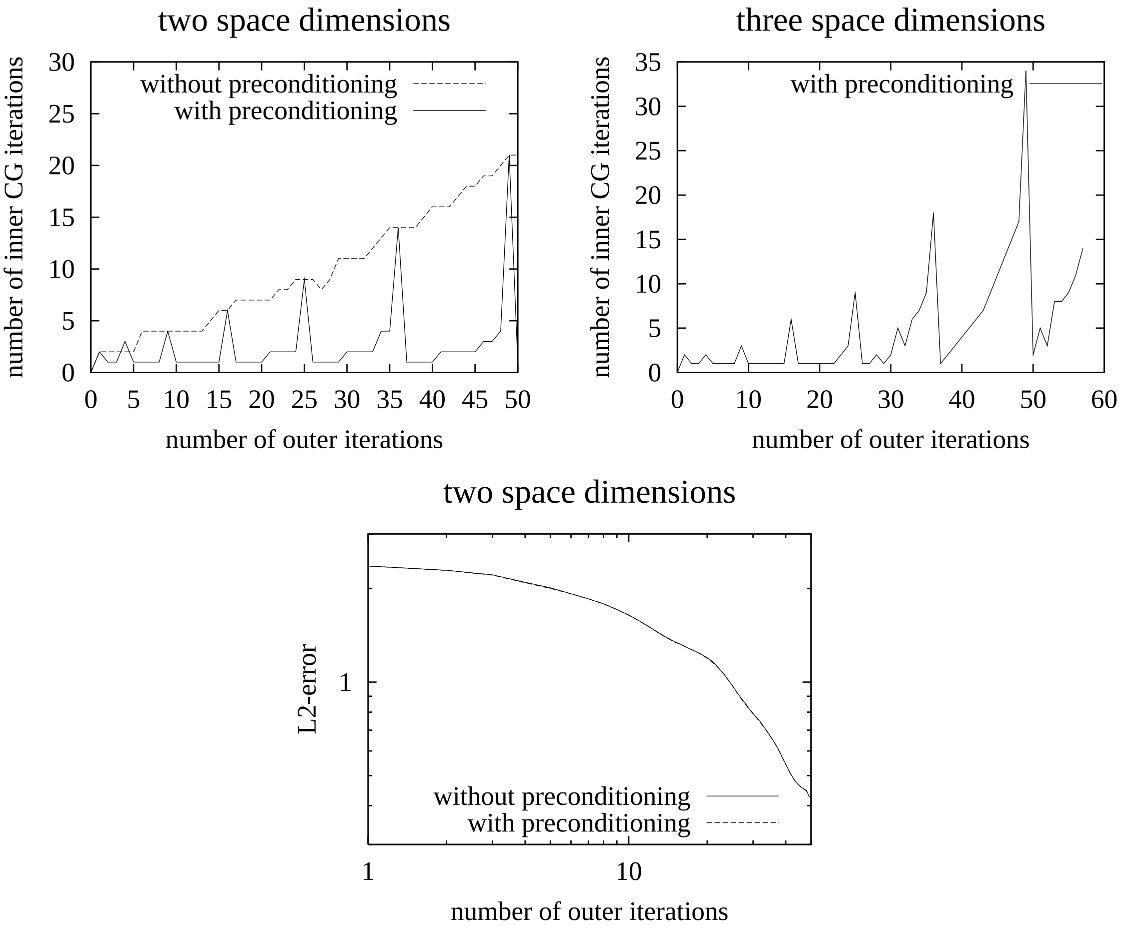

Figure 1. The effect of preconditioning on the number of inner CG iterations and the error.

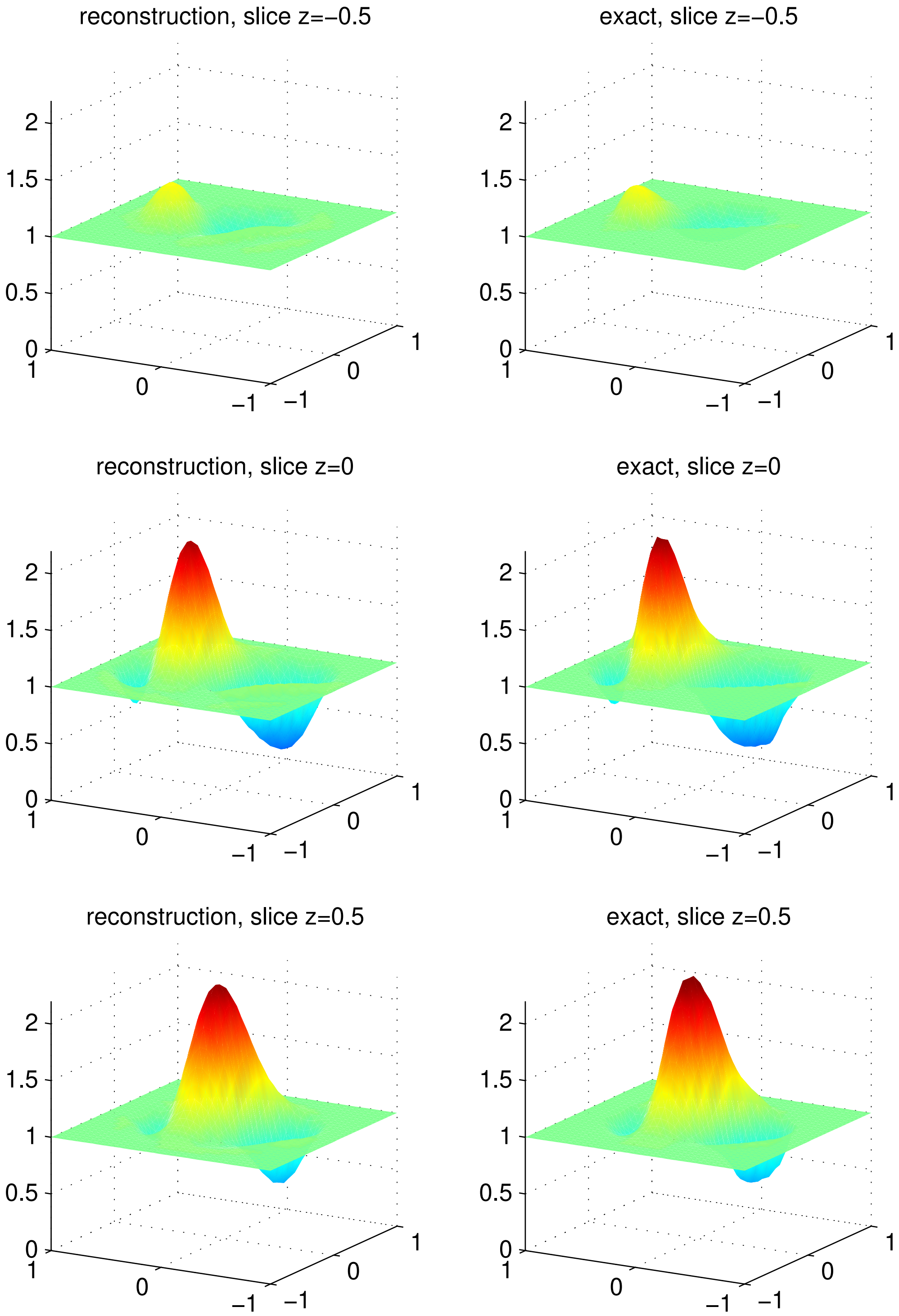

Figure 2. Slices of the reconstructed and the exact refractive index $n = 1 - a$.

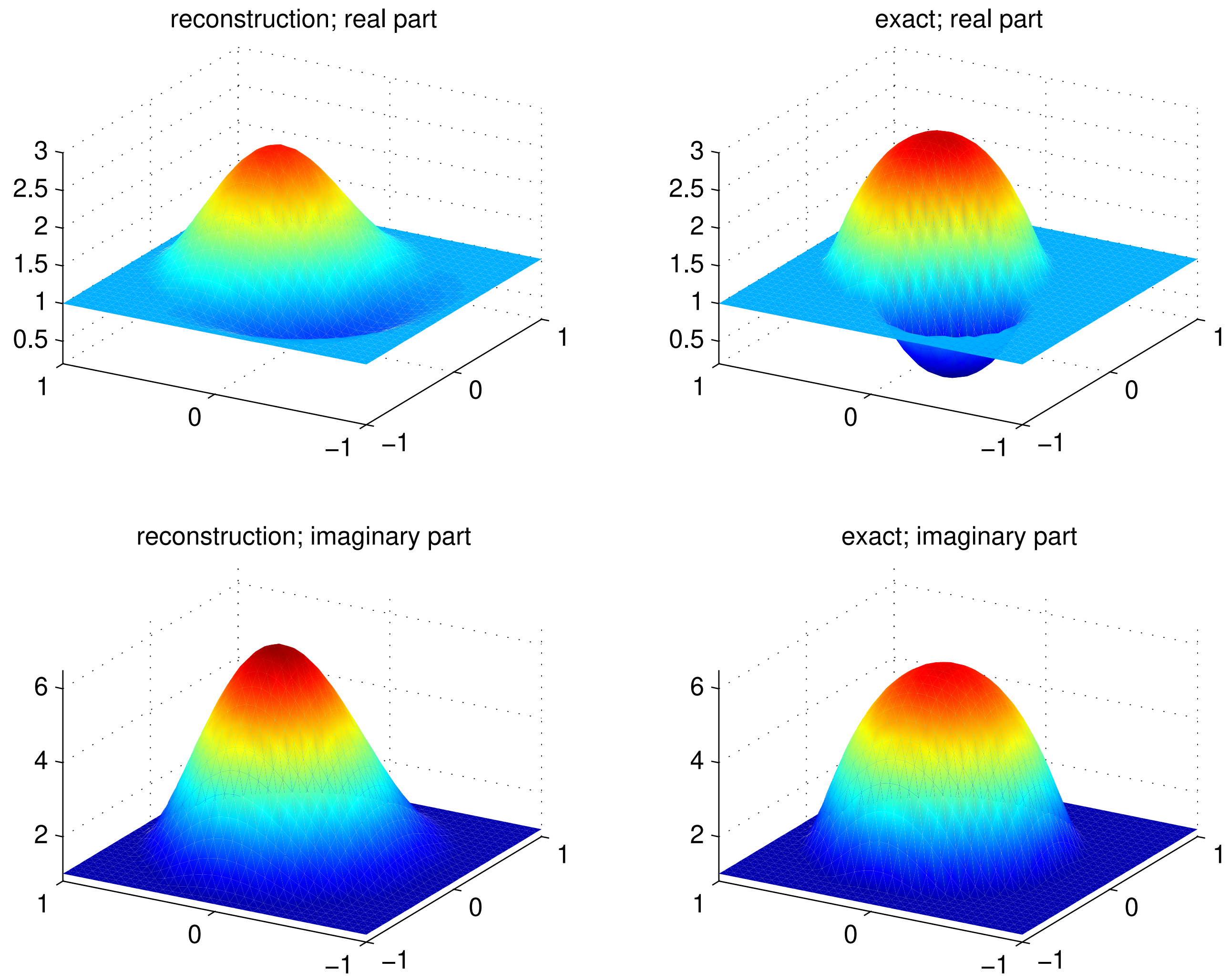

Figure 3. Reconstruction of a 3D complex-valued refractive index $n = 1 - a$ with 1% noise; slice $z = 0$.

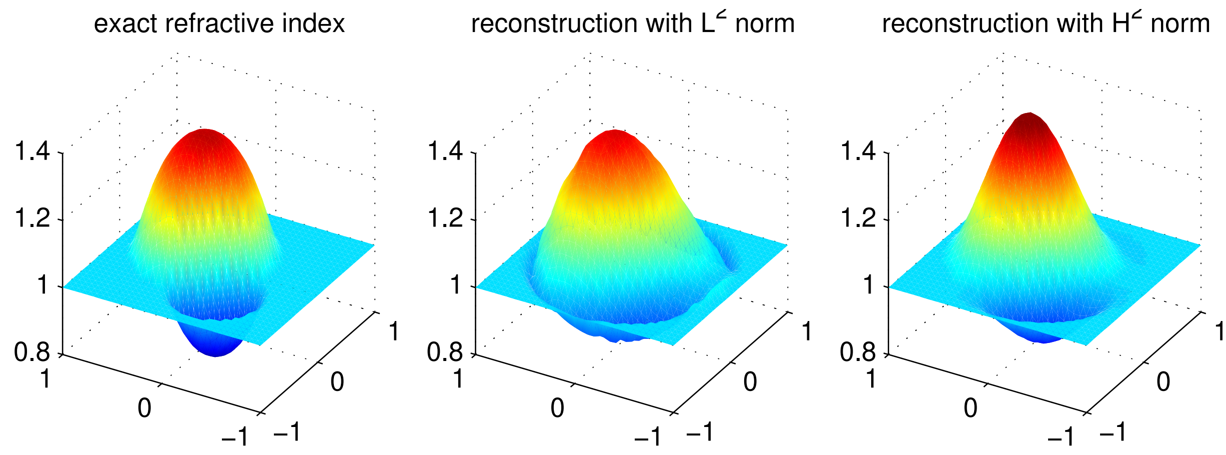

Figure 4. Influence of the choice of the Sobolev norm on the reconstruction; 0.1% noise.