Proceedings of Inverse Problems in Engineering: Theory and Practice
3rd Int. Conference on Inverse Problems in Engineering
June 13-18, 1999, Port Ludlow, WA, USA

IM04

ITERATIVE REGULARIZATION METHODS IN INVERSE SCATTERING

Thorsten Hohage
Institute for Industrial Mathematics, Univ. Linz
Altenberger Str. 69
A-4040 Linz
Austria
Email: hohage@indmath.uni-linz.ac.at

ABSTRACT

The numerical performances of Landweber iteration, the Newton-CG method, the Levenberg-Marquardt algorithm, and the iteratively Regularized Gauß-Newton method are compared for a nonlinear, severely ill-posed inverse scattering problem in two space dimensions. A modification of the Gauß-Newton method is suggested, which compares favorably with the above methods. A convergence proof is presented including the effects of the numerical approximation of the solution operator.

Keywords: inverse obstacle scattering, iterative regularization methods, operator approximations, convergence rates.

1 Introduction

We consider the following problem. Let $D \subset \mathbb{R}^2$ be a star-shaped, smooth domain describing the cross section of a long, cylindrical scattering obstacle. For an incident plane time-harmonic wave $u_i(x) := e^{ik\langle x,d\rangle}$, $|d| = 1, k > 0$ the scattered field u_s and the total field $u := u_i + u_s$ satisfy

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^2 \backslash \bar{D} \tag{1}$$

$$\sqrt{r}\left(\frac{\partial u}{\partial r} - iku\right) \to 0, \qquad r = |x| \to \infty \tag{2}$$

$$u = 0 \qquad \text{on } \partial D. \tag{3}$$

In acoustics, this describes scattering from a sound-soft obstacle, and in electromagnetics, it describes scattering of a polarized wave from a perfect conductor. The Sommerfeld radiation condition (2) implies the asymptotic behavior

$$u(x) = \frac{1}{|x|}\left(u_\infty\left(\frac{x}{|x|}\right) + O\left(\frac{1}{|x|}\right)\right), \qquad |x| \to \infty$$

(Colton, Kreß 97). The function $u_\infty : \{x : |x| = 1\} \to \mathbb{C}$, called the far-field pattern of u_s, is analytic. We want to solve the inverse problem to identify the shape of D from measurements of u_∞ for one fixed incident wave u_i. This problem is difficult to solve since it is nonlinear and severely ill-posed. To formulate it as a nonlinear operator equation

$$F(q) = u_\infty,$$

we describe the boundary ∂D by a radial function $q : [0, 2\pi] \to \mathbb{R}$:

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

The characterization of the Fréchet derivatives of F, which has been accomplished some years ago by Kirsch and others, has paved the way for the application of iterative regularization methods to solve this and related problems numerically. These methods have been intensively investigated recently, and convergence results have been obtained under some conditions on the nonlinearity of the operator. Examples include Landweber iteration (Hanke, Neubauer, Scherzer 95), the iteratively regularized Gauß-Newton method (IRGNM) (Bakushinskii 92, Blaschke/Kaltenbacher et al 97, Hohage 97) and inexact Newton methods such as the Levenberg-Marquardt and the Newton-CG algorithm (Hanke 97) or most recently a second degree method (Hettlich, Rundell 99). We compare the numerical performance of these methods applied to the above inverse scattering problem and suggest a new, more efficient modification of the IRGNM.

While the numerical solution of the model problem (1) - (3) is quite fast on modern computers, computation time becomes a central issue in large scale and 3-dimensional problems. Hence we address the problem of minimizing computation time. Roughly speaking, it does not pay off to pay too much effort in an accurate evaluation of the operator as long as one is still far away from the solution. We show how various discretization parameters have to be increased during the iteration such that the order of convergence established in the infinite dimensional setting is maintained.

2 Iterative Regularization methods

We consider the following abstract setting: Let X and Y be Hilbert spaces, and $F : X \supset D(F) \to Y$ a nonlinear operator that is continuously Fréchet differentiable on its domain $D(F)$. We want to solve the operator equation

$$F(x) = y, \tag{4}$$

and call the exact solution $x^\dagger$. Moreover, we assume that only noisy data y^δ are available satisfying

$$\|y^\delta - y\| \leq \delta \tag{5}$$

with some known noise level δ.

Definition 1. *An iterative method $x^\delta_{n+1} := \Phi(x^\delta_n, \ldots, x_0, F, y^\delta)$ together with a stopping rule $N(\delta, y^\delta)$ is called an* iterative regularization method *for F if for all $x^\dagger \in D(F)$, $y := F(x^\dagger)$, all y^δ satisfying (5) and all initial guesses x_0 sufficiently close to $x^\dagger$ the following conditions hold:*

- *x^δ_n is well defined for $n = 1, \ldots, N(\delta, y^\delta)$, and $N(\delta, y^\delta) < \infty$ for $\delta > 0$.*
- *For exact data ($\delta = 0$) either $N = N(\delta, y^\delta) < \infty$ and $x^\delta_N = x^\dagger$ or $N = \infty$ and $\|x_n - x^\dagger\| \to 0$ for $n \to \infty$.*
- *The following regularization property holds:*

$$\sup_{\|y^\delta - y\| \leq \delta} \|x^\delta_{N(\delta,y^\delta)} - x^\dagger\| \to 0, \qquad \delta \to 0. \tag{6}$$

The choice of the stopping index is a very important issue for iterative regularization methods since typically, for noisy data, the approximations deteriorate quite rapidly if one iterates too often. The most well-known stopping rule is the discrepancy principle which consists in stopping the iteration at the first index $N = N(\delta, y^\delta)$ for which

$$\|F(x_N) - y^\delta\| \leq \tau\delta \tag{7}$$

with some fixed constant $\tau > 1$.

Unfortunately, convergence in (6) can be *arbitrarily slow* if (4) is ill-posed (see, e.g., Proposition 3.11 in (Engl et al. 96)). We would like to have an estimate of the form

$$\sup_{\|y^\delta - y\| \leq \delta} \|x^\delta_{N(\delta,y^\delta)} - x^\dagger\| = O(g(\delta)) \tag{8}$$

with some function g converging to 0 as $\delta \to 0$. Such convergence rates results can only be obtained if a so-called *source condition* is satisfied which for nonlinear problems usually has the form

$$x_0 - x^\dagger = f(F'[x^\dagger]^* F'[x^\dagger])w, \qquad \|w\| \leq \bar{w}. \tag{9}$$

Since $F'[x^\dagger]$ is typically smoothing, this can be seen as an abstract smoothness and closeness condition on the initial error. The most common choice of f is $f(\lambda) = \lambda^\mu$ with some $\mu > 0$. It can be shown that the best-possible g in (8) for these so-called *Hölder-type source conditions* is $g(\lambda) = \lambda^{\frac{2\mu}{2\mu+1}}$. However, for severely ill-posed problems such source conditions are usually far too restrictive. For our inverse scattering problem, they imply that the exact solution is known up to an analytic function! As discussed below, appropriate functions f for severely ill-posed problems are often of the form

$$f_p(\lambda) = (-\ln\lambda)^{-p}, \qquad \lambda > 0, \qquad f_p(0) = 0$$

with some parameter $p > 0$. In this case, the best possible g in (8) is $g = f_p$ (Mair 94), which is of course much worse than $g(\lambda) = \lambda^{\frac{2\mu}{2\mu+1}}$, reflecting the severe ill-posedness. In order to avoid the singularity of f_p at $\lambda = 0$, w.r.o.g. we will always assume that F is scaled such that $\|F'[x]^* F'[x]\| \leq \exp(-1)$ holds for x in a neighborhood of $x^\dagger$.

Let us relate these abstract results to our inverse scattering problem. First of all, we need the derivative of the domain-to-far-field mapping F. We consider F in the function spaces

$$F : \{q \in H^s[0, 2\pi] : q > 0\} \to L^2(S^1).$$

Here $H^s[0,2\pi]$ is the Sobolev space of index s of periodic functions on $[0,2\pi]$. It can be shown (see Kirsch 93) that F' exists and has the form $F'[q]h = L_q M_q h$ where $M_q h := \frac{q}{q^2+q'^2}\frac{\partial u}{\partial \nu}h$ and L_q maps a (parametrized) function f on ∂D_q to the far field u_∞ of the radiating solution v of the exterior Helmholtz equation with boundary values $v|_{\partial D} = f$. In other words, the derivative $u_s'[q]h = v$ of the scattered field u_s for the perturbation h has the boundary values $u_s'[q]h = M_q h$ on ∂D_q. Whereas M_q has a simple and stable inverse, all the ill-posedness of $F'[q]$ is in the operator L_q. If D is a disk, i.e. $q = const$, then

$$H^{s+p+\varepsilon}[0,2\pi] \subset R(f_p(L_q^* L_q)) \subset H^s[0,2\pi]$$

for all $\varepsilon > 0$ (Hohage 97). Hence, (9) with $f = f_p$ means, roughly speaking, that the initial error $x_0 - x^\dagger$ is small in some higher Sobolev norm, which is a quite natural condition. If, e.g., the unknown domain has a corner the location of which is not known, then convergence is very slow. These results, including the dependence on the smoothness parameter p came out quite neatly in numerical experiments. Similar (and often simpler) interpretations of logarithmic source conditions in terms of Sobolev spaces can be given for many other severely ill-posed problems, e.g. the backwards and sideways heat equation, an inverse problem in satellite gradiometry, an inverse potential problem, and other scattering problems (see Mair 94, Hohage 97,98,99, and Hohage, Schormann 98).

3 Comparison

Here we review some iterative regularization methods that have been suggested in the literature and compare the numerical results for the application to our inverse scattering problem. All convergence theorems need some condition restricting the degree of nonlinearity of the nonlinear operator F. Unfortunately none of these conditions could be verified for our inverse scattering problem, yet. Nevertheless, the convergence results have been confirmed in numerical experiments.

Landweber iteration

Landweber iteration is defined by the formula

$$x_{n+1}^\delta := x_n^\delta + \mu F'[x_n^\delta]^*(y^\delta - F(x_n^\delta)). \tag{10}$$

μ is a scaling parameter that has to be chosen such that $\|F'[x]\| \leq 1/\eta$ for all x in a neighborhood of $x^\dagger$. Hanke, Neubauer, and Scherzer (95) have proved the following result:

Theorem 1. *If the nonlinearity condition*

$$\|F(x) - F(\bar{x}) - F'[x](x-\bar{x})\| \leq \eta \|F(x) - F(\bar{x})\| \tag{11}$$

holds for all $x, \bar{x}$ in a neighborhood of $x^\dagger$ and some $\eta < \frac{1}{2}$, then Landweber iteration together with the discrepancy principle with $\tau > 2\frac{1+\eta}{1-2\eta}$ is a regularization method in the sense of Definition 1.

In (Deuflhard et al. 98) an estimate

$$\|x_n^\delta - x^\dagger\| \leq C(\ln n)^{-p} \tag{12}$$

and optimal order convergence rates with $N = O((-\ln\delta)^{-2p}/\delta^2)$ have been shown for logarithmic source conditions, but under a nonlinearity condition that does not hold for scattering problems, as the case of concentric circles shows.

Landweber iteration has been applied to the sound-soft scattering problem by Hanke, Hettlich, and Scherzer (95), and a characterization of the adjoint $F'[x_n^\delta]^*$ has been given.

Inexact Newton methods

In inexact Newton methods, we want to solve the linearized equation

$$F'[x_n^\delta]h_n + F(x_n^\delta) = y^\delta \tag{13}$$

for the update $h_n = x_{n+1}^\delta - x_n^\delta$. As (13) inherits the ill-posedness from the nonlinear problem, it has to be regularized. This leads to formulae of the form

$$h_n := g_n(A_n^* A_n)A_n^*(y^\delta - F(x_n^\delta)). \tag{14}$$

where $A_n := F'[x_n^\delta]$ and $g_n(\lambda) \approx 1/\lambda$.

If Tikhonov regularization is applied to solve (13), we have $g_n(\lambda) = 1/(\lambda + \alpha_n)$. Then the updates $h_n \in X$ solve the minimization problems

$$\|A_n h + F(x_n^\delta) - y^\delta\|^2 + \alpha_n \|h\|^2 = \min! \tag{15}$$

This is the *Levenberg-Marquardt algorithm*. A convergence analysis was given by Hanke (97) under the assumption that α_n is chosen such that

$$\|A_n h_n + F(x_n^\delta) - y^\delta\| = \rho \|F(x_n^\delta) - y^\delta\| \tag{16}$$

with some $\rho < 1$ and that the discrepancy principle (7) is used with $\tau > 1/\rho$.

Another possibility is to use the CGNE (conjugate gradient method for the normal equation) in order to regularize (13), leading to the *Newton-CG method*. Here g_n is a polynomial depending on the right hand side, making the CGNE a nonlinear method. This method has also been investigated by Hanke (97) . He suggested to stop the inner iteration by a condition similar to (16) and use (7) with $\tau > 2/\rho^2$.

Theorem 2. *Let F satisfy the nonlinearity condition*

$$\|F(x) - F(\bar{x}) - F'[x](x - \bar{x})\| \leq c\|x - \bar{x}\|\|F(x) - F(\bar{x})\| \quad (17)$$

for all $x, \bar{x}$ in a neighborhood of $x^\dagger$. Then the versions of the Levenberg-Marquardt algorithm and the Newton-CG method described above are regularization methods in the sense of Definition 1.

If the adjoint $F'[x_n^\delta]^*$ can be computed directly, this has the advantage that a costly computation and inversion of the matrix for A_n can be avoided.

Second Degree methods

Recently, Hettlich and Rundell (99) have suggested a class of methods that use the second derivative of the operator F. A predictor-corrector procedure is used to avoid solving quadratic equations. The predictor $\tilde{h}_n$ is computed by a formula similar to (14). Then, the corrector h_n is obtained as a regularized solution of the linear equation

$$A_n h_n + \frac{1}{2} F''[x_n^\delta](h_n, \tilde{h}_n) = y^\delta - F(x_n^\delta).$$

The authors used Tikhonov regularization with constant regularization parameter in both the predictor and the corrector step. The following convergence result was shown:

Theorem 3. *The second degree method described above with the stopping rule (7) is a regularization method in the sense of Definition 1 if (17) holds, $\|F'\|$, and $\|F''\|$ are locally bounded, and if τ and the regularization parameter in the corrector step are chosen sufficiently large.*

In order to apply this method to the scattering problem in Section 1, boundary values of the second derivative $u_s''(h, \tilde{h})$ of the scattered field have been computed. A simpler proof, showing analyticity of F and giving boundary values of derivatives of arbitrary order for Dirichlet and Neumann boundary conditions, is given in (Hohage 99).

Numerical experiments have shown that better reconstructions can be obtained in the first iteration steps, but the asymptotic behavior is the same as for the Levenberg-Marquardt algorithm.

Bakushinskii methods

Another class of iterative regularization methods is given by the recursion formula

$$x_{n+1} := x_0 + g_n(A_n^* A_n) A_n^* \Big(y^\delta - F(x_n^\delta) + A_n(x_n^\delta - x_0) \Big). \quad (18)$$

For the choice $g_n(\lambda) := \frac{1}{\alpha_n + \lambda}$, suggested originally by Bakushinskii (92), this is the *iteratively regularized Gauß-Newton method* (IRGNM). The regularization parameters α_n are chosen such that

$$1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq R \quad \text{and} \quad \lim_{n \to \infty} \alpha_n = 0 \quad (19)$$

with some $R > 1$, e.g.

$$\alpha_n = \alpha_0 R^{-n}. \quad (20)$$

From the theory of linear Tikhonov regularization, it easily follows that the updates $h_n = x_{n+1}^\delta - x_n^\delta$ solve the minimization problems

$$\|A_n h + F(x_n^\delta) - y^\delta\|^2 + \alpha_n \|h + x_n^\delta - x_0\|^2 = \min! \quad (21)$$

The additional term $x_n^\delta - x_0$ as compared to (15) has an additional regularizing effect and facilitates the convergence analysis. In particular, it allows to obtain convergence rates which are not yet available for inexact Newton methods.

Here, we suggest to use (18) with

$$g_n(\lambda) := \frac{(\lambda + \alpha_n)^l - \alpha_n^l}{\lambda(\lambda + \alpha_n)^l}, \quad (22)$$

$l \in \mathbb{N}$. This corresponds to iterated Tikhonov regularization and contains the IRGNM as the special case $l = 1$. The implementation of one iteration step is similar to linear iterated Tikhonov regularization (Engl et al. 96).

$$
\begin{aligned}
& h_1 := \text{solution to (21)} \\
& \text{for} \, (j = 2, \ldots, l) \\
& \quad h_j := \text{argmin}_{h \in X} (\|A_n h + F(x_n^\delta) - y^\delta\|^2 + \alpha_n \|h - h_{j-1}\|^2) \\
& x_{n+1}^\delta := x_n^\delta + h_l
\end{aligned}
$$

Note that the computation of $h_2, \ldots, h_l$ is very cheap since the same matrix has to be inverted as in the first inner step, and there is no further operator evaluation. Roughly speaking, one gets the

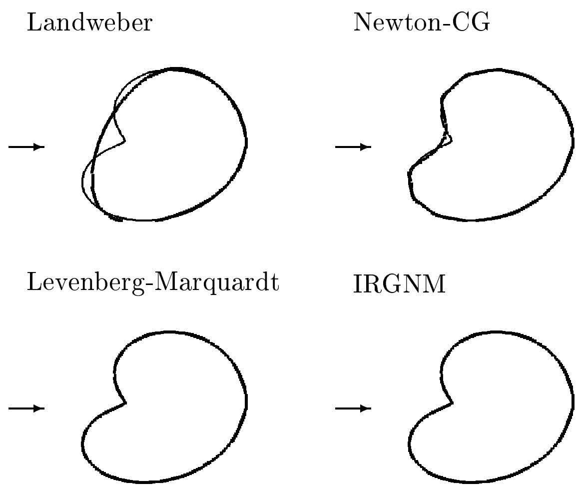

Figure 1. Reconstructions for exact data after 100 its

effect of almost l Newton steps for the cost of somewhat more than one.

It was shown in (Blaschke/Kaltenbacher et al. 97) that the IRGNM with the discrepancy principle is a regularization method under the nonlinearity condition

$$\begin{aligned} &F'[\bar{x}] = R(\bar{x},x)F'[x] + Q(\bar{x},x) \\ &\|I - R(\bar{x},x)\| \le C_R, \quad \|Q(\bar{x},x)\| \le C_Q \|F'[x^\dagger](\bar{x}-x)\| \end{aligned} \tag{23}$$

for $x,\bar{x}$ in a neighborhood of $x^\dagger$. Moreover, optimal order convergence rates have been shown for Hölder-type source conditions with $0 < \mu \le 1/2$. For logarithmic source conditions, it was shown in (Hohage 97) that optimal order convergence rates hold and that

$$\|x_n^\delta - x^\dagger\| \le C f_p(\alpha_n) \tag{24}$$

for $n \le N(\delta, y^\delta)$ with $N(\delta, y^\delta) = O(-\ln\delta)$. A convergence analysis for the method (18), (22) with $l > 1$ will be given below as a special case of Theorem 4.

Numerical Results

Figures 1 and 2 show reconstruction and corresponding error plots for regularization methods described above. We always chose the unit circle as initial guess, and $k = 1$ in (1).

Landweber iteration is clearly the slowest method. Although some good progress is made at the beginning in filtering out the low frequency components of the error, Landweber iteration is extremely slow at the high frequency components. From

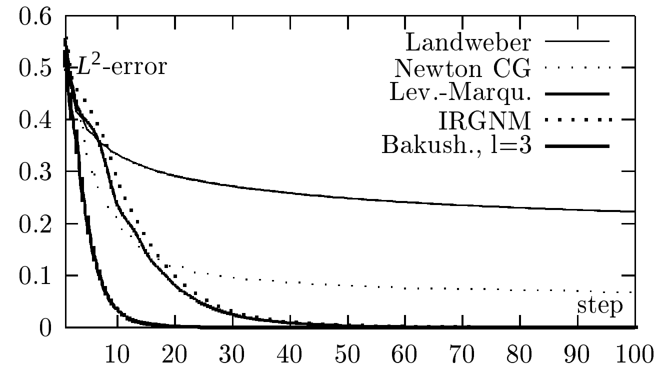

Figure 2. Convergence plots for different methods

(12), (20) and (24), we see that in order to get the effect of one IRGNM step one has to multiply the number of Landweber steps by R!

Although the Newton-CG method is significantly faster than Landweber iteration, and the low-frequency components of the error are eliminated remarkably fast, the asymptotic convergence is also very slow. This can (informally) be explained as follows: As for Landweber iteration, the updates h_n contain only very little high-frequency components since they are images of a reasonable-size functions under the highly smoothing operator $F'[x_n^\delta]$. We chose $\rho = 0.8$ and imposed a maximum number of 50 inner iterations which was rarely reached.

The performance of the *Levenberg-Marquardt algorithm* and the *IRGNM* is very similar if the same regularization parameters α_n are used. Levenberg-Marquardt is slightly faster, but somewhat less stable since high-frequency components may add up in course of the iteration. We have found in numerical experiments that the choice (16) of α_n almost leads to an asymptotic behavior of the form (20). Therefore, the numerical results with the a-priori choice (20) are similar to those with the a-posteriori choice (16). (16) has the advantage that only one free parameter ρ occurs, whereas (20) has two free parameters α_0 and R, and particularly α_0 is very problem dependent. Therefore, one may consider to chose α_0 by (16) when using (20). For an effective implementation of (16) we refer to Chapter 9 of (Engl et al. 96).

Let us now compare the methods (18), (22) for $l = 1$ and $l > 1$. As expected from the theoretical convergence result (24), for the same choice of α_n the order of convergence is the same, only the constant is somewhat smaller for $l > 1$. However, we may reduce α_n faster, i.e. increase R in (20), for $l > 1$, and then convergence becomes significantly faster. Of course, for ill-posed problems, fast convergence is to some extent always obtained at the cost of stability, but we claim that the method (18), (22) with $l > 1$ is a good trade-off. To support this, we have tested this method with noisy data on a kite-shaped do-

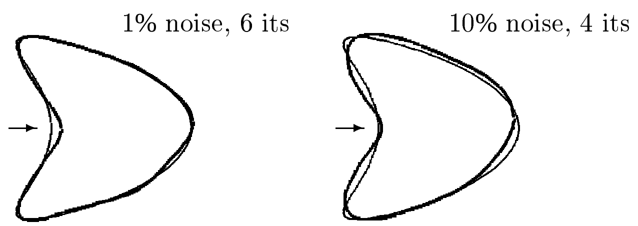

Figure 3. method (18), (22) with $l = 3$ for noisy data

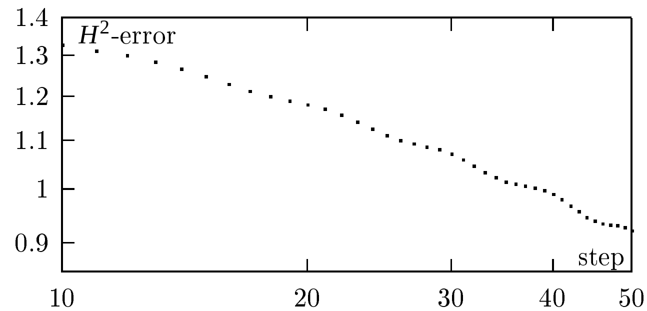

Figure 4. Speed of convergence for method (18), (22) with $l = 3$

main which is known as a difficult test example. It turned out that the method, with $l = 3$ and $R = 5$ in (20), is stable even with 10% white noise added to the data, and only very few iterations were needed to meet the stopping criterion (7) with $\tau = 1.1$. Moreover, to check our theoretical result on the speed of convergence, we have plotted $\|x_n^\delta - x^\dagger\|$ over n on a double logarithmic scale for the bean-shaped domain in Figure 1. From (24) we have $\ln \|x_n^\delta - x^\dagger\| \leq c - p \ln n$, so assuming this estimate to be asymptotically sharp, we expect the plot to be close to a straight line. This is confirmed in Figure 4.

4 Operator Approximations

We now consider the situation where in the n-th step of the Bakushinskii-iteration (18) the operator F is approximated numerically by an operator $F^{(n)}$, and the derivative $F'[x]$ is approximated by an operators $A_x^{(n)}$. We will shortly write $A_n^{(n)}$ and $A_\dagger^{(n)}$ for $A_{x_n^\delta}^{(n)}$ and $A_{x^\dagger}^{(n)}$, rsp. Thus, (18) is replaced by the formula

$$x_{n+1}^\delta := x_0 + g_n(A_n^{(n)*} A_n^{(n)}) A_n^{(n)*} \left(y^\delta - F^{(n)}(x_n^\delta) + A_n^{(n)}(x_n^\delta - x_0) \right). \tag{25}$$

We will assume that the approximation errors satisfy

$$\max \left(\|F^{(n)}(x_n^\delta) - F(x_n^\delta)\|, \|A_\dagger^{(n)} - A_\dagger\| \right) \leq \eta_n, \tag{26}$$

and that $A_x^{(n)}$ is almost (if not exactly) the derivative of $F^{(n)}$, i.e.

$$\|A_x^{(n)} - F^{(n)\prime}[x]\| \leq \eta_n. \tag{27}$$

Finally, we will replace (23) by an analogous condition for the discrete operators:

$$\begin{aligned} &A_{\bar{x}}^{(n)} = R^{(n)}(\bar{x}, x) A_x^{(n)} + Q^{(n)}(\bar{x}, x) \\ &\|I - R^{(n)}(\bar{x}, x)\| \leq C_R, \quad \|Q^{(n)}(\bar{x}, x)\| \leq C_Q \|A_\dagger^{(n)}(\bar{x} - x)\| \end{aligned} \tag{28}$$

It may be advantageous to state the nonlinearity condition for the discrete operators since it may be easier to verify in this form, and a more careful analysis can also deal with a moderate dependence of the constants on n. However, we have not been able to provide such an analysis for the sound-soft scattering problem.

Theorem 4. *Assume that (4), (5), (7), (19), (22), and (25) - (28) hold with C_R, C_Q sufficiently small, τ and α_0 sufficiently large, and*

$$\eta_n \leq c_\eta \sqrt{\alpha_n} f_p(\alpha_n) \tag{29}$$

with c_η sufficiently small. Then, there exist $\rho, \bar{w} > 0$ such that for all x_0 satisfying $\|x_0 - x^\dagger\| \leq \rho$ and (9) with $f = f_p$, the iterates x_n^δ are well defined for $n \leq N(\delta, y^\delta)$ and satisfy the estimate (24) with a constant C independent of n, δ and y^δ. Moreover, $N(\delta, y^\delta) = O(-\ln \delta)$, and the order-optimal convergence rate (8) with $g = f_p$ holds.

Proof. Since parts of proof are similar to the proof for the IRGNM with exact operators (Hohage 97), we just outline the main ideas, and only describe the new features in some detail. A key observation, making Bakushinskii methods easier to analyze than other methods, is that the total error $e_n = x_n^\delta - x^\dagger$ can be decomposed into an approximation error $\tilde{e}_n^{\rm app}$, a propagated data noise error $e_n^{\rm noi}$, and a Taylor remainder $e_n^{\rm tay}$, i.e. $e_n = \tilde{e}_n^{\rm app} + e_n^{\rm tay} + e_n^{\rm noi}$ with

$$\begin{aligned} \tilde{e}_n^{\rm app} &:= r_n(S_n^{(n)}) e_0, \\ e_n^{\rm noi} &:= g_n(S_n^{(n)}) A_n^{(n)*}(y^\delta - y), \\ e_n^{\rm tay} &:= g_n(S_n^{(n)}) A_n^{(n)*}(y - F^{(n)}(x_n^\delta). + A_n^{(n)}(x_n^\delta - x^\dagger)) \end{aligned}$$

Here $r_n(\lambda) := 1 - \lambda g_n(\lambda)$, $S_n^{(n)} := A_n^{(n)*} A_n^{(n)}$, and correspondingly we will use $S_\dagger^{(n)} := A_\dagger^{(n)*} A_\dagger^{(n)}$ and $S_\dagger := A_\dagger^* A_\dagger$. The approximation error $\tilde{e}_n^{\rm app}$ is now further decomposed into an error $e_n^{\rm app}$ which corresponds to the approximation error for the linear equation $A_\dagger^{(n)} x = y$, an error $e_n^{\rm nl}$ describing the nonlinearity effect that

$A_n^{(n)} \neq A_\dagger^{(n)}$, and an error e_n^{discr} containing the effects of the discretization error in $A_\dagger^{(n)}$: $\tilde{e}_n^{\text{app}} = e_n^{\text{app}} + e_n^{\text{nl}} + e_n^{\text{discr}}$ with

$$\begin{aligned} e_n^{\text{app}} &:= r_n(S_\dagger^{(n)}) f_p(S_\dagger^{(n)}) w, \\ e_n^{\text{nl}} &:= \left(r_n(S_n^{(n)}) - r_n(S_\dagger^{(n)}) \right) f_p(S_\dagger^{(n)}) w, \\ e_n^{\text{discr}} &:= r_n(S_n^{(n)}) \left(f_p(S_\dagger) - f_p(S_\dagger^{(n)}) \right) w \end{aligned}$$

where we have used (9). Now all error terms are estimated separately to yield a recursive estimate of the total error of the form

$$\|e_{n+1}\| \leq a_1 f_p(\alpha_n) + a_2 \|A_\dagger^{(n)} e_n\|. \tag{30}$$

The estimates on e_n^{noi} and e_n^{tay}, derived from (28), (26), (27), and $\|g_n(S_n^{(n)}) A_n^{(n)*}\| \leq \sqrt{\frac{l}{\alpha_n}}$, are very similar to those in (Hohage 97). We have

$$\begin{aligned} \|e_n^{\text{noi}}\| &\leq \sqrt{\frac{l}{\alpha_n}} \delta \leq \sqrt{\frac{l}{\alpha_n}} \frac{C_R + 1 + \frac{1}{2}\|e_n\| C_Q}{\tau - 1} \|A_\dagger^{(n)} e_n\|, \\ \|e_n^{\text{tay}}\| &\leq \sqrt{\frac{l}{\alpha_n}} \left((2C_R + \frac{3}{2}\|e_n\| C_Q) \|A_\dagger^{(n)} e_n\| + (1 + \|e_n\|) \eta_n \right). \end{aligned}$$

Moreover, from $r_n(\lambda) = (\alpha_n/(\alpha_n + \lambda))^l$, it can be shown that the estimates

$$\|e_n^{\text{app}}\| \leq c_1 f_p(\alpha_n), \qquad \|A_\dagger^{(n)} e_n^{\text{app}}\| \leq c_2 \sqrt{\alpha_n} f_p(\alpha_n). \tag{31}$$

derived for Tikhonov regularization, also hold for iterated Tikhonov regularization. To estimate $\|e_n^{\text{discr}}\|$, we use the inequality

$$\|f_p(S_\dagger) - f_p(S_\dagger^{(n)})\| \leq c\, f_p \left(\|S_\dagger - S_\dagger^{(n)}\| \right)$$

(Hohage 99). This, together with (26), (29), and $\|r_n(S_n^{(n)})\| \leq \|r_n\|_\infty \leq 1$ yields $\|e_n^{\text{discr}}\| = O(f_p(\eta_n)) = O(f_p(\alpha_n))$. The estimate on e_n^{nl} depends heavily on the form of r_n. For iterated Tikhonov regularization we use the formula

$$\begin{aligned} & r_n(S_n^{(n)}) - r_n(S_\dagger^{(n)}) \\ &= \alpha_n^l (S_n^{(n)} + \alpha_n)^{-l} \left((S_\dagger^{(n)} + \alpha_n)^l - (S_n^{(n)} + \alpha)^l \right) \\ &\quad \times (S_\dagger^{(n)} + \alpha_n)^{-l} \\ &= \alpha_n^l (S_n^{(n)} + \alpha_n)^{-l} \sum_{i=i}^{l} \binom{i}{l} \alpha_n^{l-i} \left((S_\dagger^{(n)})^i - (S_n^{(n)})^i \right) \\ &\quad \times (S_\dagger^{(n)} + \alpha_n)^{-l}, \end{aligned}$$

and treat each term in the sum as in (Hohage 97) using (28) and (31). This yields an estimate of the form

$$\|e_n^{\text{nl}}\| \leq c \left(1 + \frac{\|A_\dagger^{(n)} e_n\|}{\sqrt{\alpha_n}} \right) f_p(\alpha_n) \|w\|. \tag{32}$$

A recursive estimate of the form

$$\|A_\dagger^{(n)} e_{n+1}\| \leq \tilde{a}_3 \|A_\dagger^{(n)} e_n^{\text{app}}\| + a_4 \|A_\dagger^{(n)} e_n\| + a_5 \|A_\dagger^{(n)} e_n\|^2$$

can be derived analogously. It follows that

$$\|A_\dagger^{(n+1)} e_{n+1}\| \leq a_3 \sqrt{\alpha_n} f_p(\alpha_n) + a_4 \|A_\dagger^{(n)} e_n\| + a_5 \|A_\dagger^{(n)} e_n\|^2.$$

Now an induction argument shows that under certain smallness assumptions on $a_1, \ldots, a_5$ corresponding to smallness assumptions on $C_R, C_Q, 1/\tau, 1/\alpha_n, \eta_n, \rho$, and $\bar{w}$, we have

$$\|e_n\| \leq \rho, \qquad \frac{\|A_\dagger^{(n)} e_n\|}{\sqrt{\alpha_n} f_p(\alpha_n)} \leq C$$

for $n \leq N$ with C independent of n, δ, and y^δ. Together with (30), this yields (24).

To prove (8), it suffices to show that $\alpha_N^l = O(\delta)$. This can be done in a fashion similar to (Hohage 97). First, it is shown that $A_\dagger^{(n)} e_n$ is dominated by $A_\dagger^{(n)} e_n^{\text{app}}$, in the sense that

$$\|A_\dagger^{(n+1)} e_{n+1}\| \geq c \|A_\dagger^{(n)} e_n^{\text{app}}\|.$$

For the right hand side, with $n = N-1$, we have the estimate

$$\begin{aligned} \|A_\dagger^{(N-1)} e_{N-1}^{\text{app}}\| &= \|r_{N-1} \left(A_\dagger^{(N-1)} A_\dagger^{(N-1)*} \right) A_\dagger^{(N-1)} e_0\| \\ &\geq \frac{\alpha_{N-1}^l}{\left(\left\| A_\dagger^{(N-1)} \right\|^2 + \alpha_0 \right)^l} \|A_\dagger^{(N-1)} e_0\|. \end{aligned}$$

Together with the estimate $\delta \geq \frac{\tilde{c}}{\tau+1} \|A_\dagger^N e_N\|$, which can be derived from (7) and (28), the proof is complete. ■

Typically, the operators $F^{(n)}$ and $A_x^{(n)}$ are of the form $F^{(n)} = Q^{(q_n)}\tilde{F}^{(f_n)}$ and $A_x^{(n)} = Q^{(q_n)}\tilde{A}_x^{(a_n)}P^{(p_n)}$ where $P^{(m)}$ is an orthogonal projection on a finite dimensional subspace X_m of X, $Q^{(m)}$ is a projection on a finite dimensional subspace Y_m of Y, and $\tilde{F}^{(m)}$ and $\tilde{A}_x^{(m)}$ are numerical approximations of F and $F'[x]$, rsp. Q_n is not necessarily orthogonal, e.g., it can be a collocation operator.

If $g_n(\lambda) = (\lambda+\alpha_n)^{-1}$, then $h_n = x_{n+1}^\delta - x_n^\delta$ solves the minimization problem

$$\|A_n^{(n)}h + F_n^{(n)}(x_n^\delta) - y^\delta\|^2 + \alpha_n\|h + x_n^\delta - x_0\|^2 = \min! \quad h \in X. \tag{33}$$

If $X_0 \subset X_1 \subset X_2 \subset \ldots$, an induction argument shows that $h_n \in X_{p_n}$ and $x_n^\delta - x_0 \in X_{p_{n-1}}$. This means that (33) is actually a finite-dimensional minimization problem. We may replace y^δ in (33) by $\tilde{Q}^{(q_n)}y^\delta$ where $\tilde{Q}^{(q_n)}$ is the *orthogonal* projection on Y_{q_n}.

By the triangle inequality, we have the estimate

$$\begin{aligned}\|A_x^{(n)} - F'[x]\| &\leq \|Q^{(q_n)}\left(\tilde{A}_x^{(a_n)} - F'[x]\right)P^{(p_n)}\| \\ &+ \|\left(Q^{(q_n)} - I\right)F'[x]P^{(p_n)}\| + \|F'[x]\left(P^{(p_n)} - I\right)\|.\end{aligned}$$

If the sound-soft scattering problem is implemented as suggested in (Hohage 97, 98), then X_m and Y_m are spaces of trigonometric polynomials of order $\leq m$, and a_n, f_n correspond to the number of quadrature points for the integral operators. Using $\|F'[x](P^{(p_n)} - I)\| = \|(P^{(p_n)} - I)F'[x]^*\|$ and exponential convergence of the trigonometric interpolation error for analytic functions, it can be shown that (26)-(27) are satisfied if the discretization parameter are chosen of the form $p_n, q_n, a_n, f_n = d_0 + nd_1$ with sufficiently large d_0, d_1. We do not discuss this in detail here since for the two-dimensional sound-soft scattering problem computation time is actually not a real problem. However, Theorem 4 will be useful for large-scale problems.

ACKNOWLEDGMENT

This work has been supported by the Austrian Science Fund 'Fonds zur Förderung der wissenschaftlichen Forschung (FWF)' under project SFB F013 'Numerical and Symbolic Scientific Computing'.

REFERENCES

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

B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauß-Newton method. *IMA Journal of Numerical Analysis*, 17:421–436, 1997.

D. Colton and R. Kreß. *Inverse Acoustic and Electromagnetic Scattering*. Springer Verlag, Berlin,Heidelberg, second edition, 1997.

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.

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

M. Hanke. A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13:79–95, 1997.

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

M. Hanke, F. Hettlich, and O. Scherzer. The Landweber iteration for an inverse scattering problem. In K.W.Wang et al., editors, *Proceedings of the 1995 Design Engineering Technical Conferences Vol. 3 Part C*, pages 909–915, New York, 1995. ASME.

M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numer. Math.*, 72:21–37, 1995.

F. Hettlich and W. Rundell. A second degree method for nonlinear ill-posed problems. *SIAM J. Numer. Anal.*, to appear.

T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauß-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13:1279–99, 1997.

T. Hohage. Convergence rates of a regularized Newton method in sound-hard inverse scattering. *SIAM J.Numer.Anal.*, 36:125–42, 1998.

T. Hohage. *Convergence of Regularized Newton Methods in Inverse Obstacle Scattering*. PhD thesis, University of Linz, 1999.

T. Hohage and C. Schormann. A Newton-type method for a transmission problem in inverse scattering. *Inverse Problems*, 14:1207–1227, 1998.

A. Kirsch. The domain derivative and two applications in inverse scattering theory. *Inverse Problems*, 9:81–96, 1993.

B. A. Mair. Tikhonov regularization for finitely and infinitely smoothing operators. *SIAM J. Math. Anal.*, 25:135–47, 1994.

