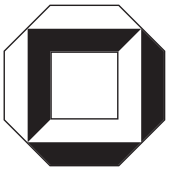


UNIVERSITÄT KARLSRUHE

A Convergenze Analysis Of The Newton-Type Regularization CG-Reginn With Application To Impedance Tomography

A. Lechleiter
A. Rieder

Preprint Nr. 07/01

Institut für Wissenschaftliches Rechnen
und Mathematische Modellbildung

76128 Karlsruhe

Anschriften der Verfasser:

Dipl.-Math. Armin Lechleiter
Graduiertenkolleg 1294: Analysis, Simulation und Design nanotechnologischer Prozesse
Universität Karlsruhe (TH)
D-76128 Karlsruhe

Prof. Dr. Andreas Rieder
Institut für Angewandte und Numerische Mathematik und
Institut für Wissenschaftliches Rechnen und Mathematische Modellbildung
Universität Karlsruhe
D-76128 Karlsruhe

A CONVERGENCE ANALYSIS OF THE NEWTON-TYPE REGULARIZATION CG-REGINN WITH APPLICATION TO IMPEDANCE TOMOGRAPHY

ARMIN LECHLEITER[†] AND ANDREAS RIEDER[‡]

Abstract. The Newton-type regularization `CG-REGINN` is an efficient tool for stably solving nonlinear ill-posed problems. In this paper a new convergence analysis for a slightly modified version of `CG-REGINN` is given, extending previous results by Hanke [*Numer. Funct. Anal. Optimiz.* 18, 971-993, 1997] and the second author [*SIAM Numer. Anal.* 43, 604-622, 2005]. Some numerical experiments from electrical impedance tomography illustrate the algorithm.

Key words. Nonlinear ill-posed problems, inexact Newton iteration, conjugate gradients, electrical impedance tomography.

AMS subject classifications. 65J20, 65J22.

1. Introduction. During the last 15 years a broad variety of Newton-like methods for regularizing nonlinear ill-posed problems have been suggested and analyzed, see, e.g., [1, 8, 12] for an overview. So different the methods are, they almost all rely on needful assumptions restricting the nonlinearity, especially when a posteriori stopping rules are employed. The weakest of these assumptions currently is the *tangential cone condition* which was presumably introduced by Scherzer in [14]. An operator $F\colon \mathsf{D}(F) \subset X \to Y$ satisfies the tangential cone condition if

$$\|F(v) - F(w) - F'(w)(v-w)\|_Y \le \omega \|F(w) - F(v)\|_Y \quad \text{for one } \omega < 1 \tag{1.1}$$

locally about a point in $\mathsf{D}(F)$, the domain of definition of F. Here, F' is the Fréchet derivative of F and X as well as Y always denote Hilbert spaces.

In this paper we consider the Newton-type regularization `CG-REGINN` and give a novel convergence analysis under (1.1). We complement and improve on previous results of Hanke [4] who investigated a slightly different version of `CG-REGINN`, also assuming the tangential cone condition. `CG-REGINN`'s ability for tackling severely ill-posed problems arising in electrical impedance tomography was demonstrated in [9].

Now we will set the stage for introducing `CG-REGINN` later in Section 3. We want to find a stable approximate solution of the nonlinear ill-posed

[†]Graduiertenkolleg 1294 "Analysis, Simulation and Design of Nanotechnological Processes", Universität Karlsruhe, 76128 Karlsruhe, Germany, `lechleiter@math.uni-karlsruhe.de`

[‡]Institut für Angewandte und Numerische Mathematik and Institut für Wissenschaftliches Rechnen und Mathematische Modellbildung, Universität Karlsruhe, 76128 Karlsruhe, Germany, `andreas.rieder@math.uni-karlsruhe.de`

equation

$$F(x) = y^\delta \tag{1.2}$$

where the right hand side y^δ is a noisy version of the exact but unknown data $y = F(x^+)$ satisfying

$$\|y - y^\delta\|_Y \leq \delta. \tag{1.3}$$

The non-negative *noise level* δ is assumed to be known. All Newton-type algorithms for solving (1.2) update the actual iterate x_n by adding a correction step s_n^{N} obtained from solving a linearization of (1.2):

$$x_{n+1} = x_n + s_n^{\mathrm{N}}, \quad n \in \mathbb{N}_0, \tag{1.4}$$

with an initial guess x_0. For obvious reasons we like to have s_n^{N} as close as possible to the *exact Newton step*

$$s_n^{\mathrm{e}} = x^+ - x_n.$$

Assuming F to be continuously Fréchet differentiable with derivative $F' : \mathsf{D}(F) \to \mathfrak{L}(X, Y)$ the exact Newton step satisfies the linear equation

$$F'(x_n)s_n^{\mathrm{e}} = y - F(x_n) - E(x^+, x_n) =: b_n \tag{1.5}$$

where $E(v, w) := F(v) - F(w) - F'(w)(v - w)$ is the linearization error. In the sequel we will use the notation

$$A_n = F'(x_n) \quad \text{and} \quad A = F'(x^+).$$

Of course, the above right hand side b_n is not available, however, a perturbed version is known:

$$b_n^\varepsilon := y^\delta - F(x_n) \quad \text{with } \|b_n - b_n^\varepsilon\|_Y \leq \delta + \|E(x^+, x_n)\|_Y. \tag{1.6}$$

Therefore, the correction step s_n^{N} is determined as a stable approximate solution of

$$A_n s = b_n^\varepsilon. \tag{1.7}$$

The various Newton-type regularizations for (1.2) differ in how they solve (1.7) and how they stop the Newton iteration (1.4).

Algorithm `CG-REGINN` applies the method of conjugate gradients (CG) to (1.7) which is terminated as soon as the relative residual falls below a certain tolerance. This termination criterion is typical for inexact Newton methods giving `CG-REGINN` its name: `REG`ularization by `IN`exact `N`ewton methods. The corresponding outer Newton iteration (1.4) is stopped by a discrepany principle. See Figure 3.1 below for algorithmic details.

We proceed our paper in Section 2 by recalling and proving well-known and less well-known properties of the CG-method which we will rely on later

in our convergence analysis of `CG-REGINN`. In Section 3 we explain all details of `CG-REGINN` and give a formulation in pseudo code. Our new convergence results are presented in Section 4 followed by a discussion of some consequences. Finally, Section 5 completes the paper with numerical illustrations from electrical impedance tomography where (1.1) is satisfied. In two appendices we collect results which are needed but do not fit comfortably in the body of the text.

2. Preliminaries about the CG-method. Let $T \in \mathcal{L}(X,Y)$ and $0 \neq g \in Y$. The method of conjugate gradients (CG) is an iteration for solving the normal equation $T^*Tf = T^*g$. Starting with $f_0 \in X$ the CG-method produces a sequence $\{f_m\}_{m\in\mathbb{N}_0}$ with the following minimization property

$$\|g - Tf_m\|_Y = \min\big\{\|g - Tf\|_Y \,\big|\, f \in X,\ f - f_0 \in U_m\big\}, \quad m \geq 1,$$

where U_m is the m-th Krylov space,

$$U_m := \operatorname{span}\big\{T^*r^0, (T^*T)T^*r^0, (T^*T)^2T^*r^0, \ldots, (T^*T)^{m-1}T^*r^0\big\} \subset \mathsf{N}(T)^\perp$$

with $r^0 := g - Tf_0$. Here, $\mathsf{N}(T)^\perp$ denotes the orthogonal complement of the null space $\mathsf{N}(T)$ of T. Since

$$\langle g - Tf_m, Tu\rangle_Y = 0 \quad \text{for all } u \in U_m, \tag{2.1}$$

see formula (5.19) in [12], we have that

$$\begin{aligned} &\langle g - Tf_m, Tf_m\rangle_Y = 0 \quad \text{for all } m \in \mathbb{N}_0 \\ &\text{provided } f_0 = 0 \text{ which we assume throughout.} \end{aligned} \tag{2.2}$$

Therefore,

$$\|g - Tf_m\|_Y^2 = \|g\|_Y^2 - \|Tf_m\|_Y^2 < \|g\|_Y^2 \ \text{ as } f_m \in \mathsf{N}(T)^\perp \setminus \{0\}. \tag{2.3}$$

Introducing

$$\eta_m^2 := 1 - \frac{\|Tf_m\|_Y^2}{\|g\|_Y^2} < 1 \tag{2.4}$$

the above equation reads

$$\|g - Tf_m\|_Y = \eta_m \|g\|_Y.$$

LEMMA 2.1. *We have that $\eta_m \leq \eta_{m-1} \leq \eta_1 < 1$ for all $m \geq 2$.*

Proof. The assertion follows immediately from

$$\|Tf_{m-1}\|_Y \leq \|Tf_m\|_Y \leq \|g\|_Y \tag{2.5}$$

which we will prove now. By

$$\|g - Tf_m\|_Y^2 = \|g\|_Y^2 - \|Tf_m\|_Y^2 \text{ and } \|g - Tf_{m-1}\|_Y^2 = \|g\|_Y^2 - \|Tf_{m-1}\|_Y^2$$

and

$$\|g - Tf_m\|_Y^2 \leq \|g - Tf_{m-1}\|_Y^2$$

we directly find (2.5). ☐

LEMMA 2.2. *We have that*

$$\lim_{m\to\infty} \eta_m^2 = 1 - \frac{\|P_{\overline{\mathsf{R}(T)}} g\|_Y^2}{\|g\|_Y^2} = \frac{\|P_{\mathsf{R}(T)^\perp} g\|_Y^2}{\|g\|_Y^2}$$

where $P_M : Y \to Y$ is the orthogonal projection onto the subspace $M \subset Y$.

Proof. It is a well-known property of the CG-iteration that

$$\lim_{m\to\infty} Tf_m = P_{\overline{\mathsf{R}(A)}} g$$

whenever $f_0 \in \mathsf{N}(T)^\perp$, see, e.g., page 135 ff. in [12]. ☐

To formulate and to prove the following results we need to recall further facts of the CG-iteration. The m-th iterate f_m can be computed from the previous one by

$$f_m = f_{m-1} + \alpha_m p^m \quad \text{where } \alpha_m = \frac{\|T^*(g - Tf_m)\|_X^2}{\|Tp^m\|_Y^2}.$$

Also the search direction p^m obeys a recursion

$$p^{m+1} = T^*(g - Tf_m) + \beta_m p^m \quad \text{with } \beta_m = \frac{\|T^*(g - Tf_m)\|_X^2}{\|T^*(g - Tf_{m-1})\|_X^2}. \tag{2.6}$$

As soon as $T^*(g - Tf_m) = 0$ holds true the CG-sequence is finite: $f_m = f_k$ for all $k \geq m$. Accordingly,

$$\boldsymbol{m}_{\mathrm{T}} := \sup\{m \in \mathbb{N} \,|\, T^*(g - Tf_{m-1}) \neq 0\}$$

is called the ultimate termination index of the CG-method ($\boldsymbol{m}_{\mathrm{T}} = \infty$ is allowed and the supremum of the empty set is understood as zero).

REMARK 2.3. *The ultimate termination index is finite if and only if the right hand side g can be decomposed into $g = g_0 + g_1$ where $g_0 \in \mathsf{R}(T)^\perp$ and g_1 is a finite superposition of eigenvectors of TT^*, see, e.g., Hanke* [4, Sec. 2] *or* [12, Satz 5.3.4].

COROLLARY 2.4. *For $\mu \in \left]\|P_{\mathsf{R}(T)^\perp} g\|_Y / \|g\|_Y, 1\right]$ there is an $m^* \in \mathbb{N}$ such that*

$$\|g - Tf_m\|_Y \leq \mu \|g\|_Y \quad \textit{for all } m \geq m^*.$$

Moreover, $m^* \leq \boldsymbol{m}_{\mathrm{T}}$.

Proof. In view of Lemma 2.2 we only need to verify the second assertion. For $\boldsymbol{m}_{\mathrm{T}} = \infty$ nothing has to be shown. Now, let $\boldsymbol{m}_{\mathrm{T}} < \infty$. Then, $f_{\boldsymbol{m}_{\mathrm{T}}} =$

T^+g where T^+ is the Moore-Penrose inverse of T. Since $\|Tf_{\boldsymbol{m}_\mathrm{T}} - g\|_Y = \|P_{\mathsf{R}(T)^\perp} g\|_Y < \mu\|g\|_Y$ we have $m^* \le \boldsymbol{m}_\mathrm{T}$. □

The step length α_m, $1 \le m \le \boldsymbol{m}_\mathrm{T}$, is determined to minimize the strictly convex parabola

$$\begin{aligned}\Psi_m(\lambda) &:= \|g - T(f_{m-1} + \lambda\, p^m)\|_Y^2 \\ &= \|g - Tf_{m-1}\|_Y^2 + 2\lambda\langle g - Tf_{m-1}, Tp^m\rangle_Y + \lambda^2\|Tp^m\|_Y^2,\end{aligned} \tag{2.7}$$

that is,

$$\Psi_m'(\alpha_m) = 0 \quad \text{and} \quad \Psi_m''(\lambda) > 0,\ \lambda \in \mathbb{R}. \tag{2.8}$$

The search directions are T^*T-orthogonal and span the Krylov space:

$$\langle T^*Tp^k, p^m\rangle_X = \delta_{k,m}\,\|Tp^k\|_Y^2 \quad \text{and} \quad U_m = \operatorname{span}\{p^1, \dots, p^m\} \tag{2.9}$$

for k, $m \in \{1, \dots, \boldsymbol{m}_\mathrm{T}\}$.

LEMMA 2.5. *Let $m \in \{1, \dots, \boldsymbol{m}_\mathrm{T}\}$ and $\lambda \in [0, \alpha_m]$. Then,*

$$\|T(f_{m-1} + \lambda\, p^m)\|_Y \le \|g\|_Y.$$

Moreover, equality can only hold for $m = \boldsymbol{m}_\mathrm{T}$ and $\lambda = \alpha_{\boldsymbol{m}_\mathrm{T}}$.

Proof. We consider the function

$$\begin{aligned}\varphi(\lambda) &:= \|T(f_{m-1} + \lambda\, p^m)\|_Y^2 \\ &= \|Tf_{m-1}\|_Y^2 + 2\lambda\langle Tf_{m-1}, Tp^m\rangle_Y + \lambda^2\|Tp^m\|_Y^2.\end{aligned}$$

Since $f_{m-1} = \sum_{i=1}^{m-1} \alpha_i p^i$ we find $\langle Tf_{m-1}, Tp^m\rangle_Y = 0$ by the orthogonality (2.9). Hence, φ is strictly increasing on $[0, \infty[$ and the claimed estimate follows from $\varphi(\alpha_m) = \|Tf_m\|_Y^2 \le \|g\|_Y^2$, see (2.5). Now assume that $\varphi(\lambda) = \|g\|_Y^2$. The strict monotonicity of φ implies that $\lambda = \alpha_m$. Therefore, $\|Tf_m\|_Y^2 = \|g\|_Y^2$ and $\|g - Tf_m\|_Y^2 = 0$ by (2.3). □

The following theorem is essentially a re-formulation to our needs of Theorem 3.1 of Hanke [6]. Its rather technical proof is presented in Appendix A.

THEOREM 2.6. *For $g \in Y$ and $f \in X$ define $\gamma(f) := \|g - Tf\|_Y / \|g\|_Y$. Let $\mu \ge \max\{\sqrt{\gamma(f)}, \|P_{\mathsf{R}(T)^\perp} g\|_Y / \|g\|_Y\}$ and set $f_{m-1,\lambda} := f_{m-1} + \lambda p^m$ for $\lambda \in]0, \alpha_m]$.*[*]

If $\|g - Tf_{m-1,\lambda}\|_Y \ge \mu\|g\|_Y$[†] *then*

$$\|f_{m-1,\lambda} - f\|_X < \|f_{m-1} - f\|_X, \quad m \in \{1, \dots, \boldsymbol{m}_\mathrm{T}\}.$$

[*] Observe: $f_{m-1,\alpha_m} = f_m$.

[†] In particular, $m - 1 < \boldsymbol{m}_\mathrm{T}$, see Corollary 2.4.

3. Definition of `CG-REGINN`. Algorithm `CG-REGINN` was briefly explained in the introduction. Now we give its details in Figure 3.1.

A few comments are in order. The `repeat`-loop implements the CG-iteration applied to (1.7). It terminates when the relative linear residual $\|b_n^\varepsilon - A_n s_{n,m}\|_Y / \|b_n^\varepsilon\|_Y$ is smaller or equal to the user-chosen tolerance $\mu_n < 1$. If μ_n is not too small we indeed have termination.

LEMMA 3.1. *The* `repeat`*-loop of* `CG-REGINN` *terminates for any tolerance*

$$\mu_n \in \Big]\frac{\|P_{\mathsf{R}(F'(x_n))^\perp}(y^\delta - F(x_n))\|_Y}{\|y^\delta - F(x_n)\|_Y}, 1\Big].$$

Furthermore, $m_n \le \boldsymbol{m}_{\mathrm{T}}$.

Proof. By Corollary 2.4 we know that the `repeat`-loop terminates for

$$\mu_n \in \Big]\frac{\|P_{\mathsf{R}(A_n)^\perp} b_n^\varepsilon\|_Y}{\|b_n^\varepsilon\|_Y}, 1\Big]$$

and that $m_n \le \boldsymbol{m}_{\mathrm{T}}$. ∎

The `while`-loop realizes the Newton iteration. It stops as soon as the discrepancy principle is satisfied, that is, when the nonlinear residual $\|b_n^\varepsilon\|_Y$ is less than $R\delta$ where $R > 1$. Under reasonable assumptions the `while`-loop terminates as we will prove later (Theorem 4.2).

Our version of `CG-REGINN` differs a little from its original definition [4, 13] in an additional *backtracking step*: In the `else`-branch of the `if`-statement the full CG-step s_{n,m_n} is reduced to meet $\|A_n(s_{n,m_n-1} + \lambda^\star p^m) - b_n^\varepsilon\|_Y = \mu_n \|b_n^\varepsilon\|_Y$ where $\lambda^\star < \alpha_m$. Since all tolerances $\{\mu_n\}$ are bounded from below by $\mu_{\min}$ backtracking guarantees that the residual of the Newton step s_n^{N} satisfies $\|A_n s_n^{\mathrm{N}} - b_n^\varepsilon\|_Y \ge \mu_{\min}\|b_n^\varepsilon\|_Y$ which, in turn, makes Theorem 2.6 applicable yielding a monotone error reduction (Theorem 4.2).

The following lemma gives first evidence why `CG-REGINN` might work.

LEMMA 3.2. *Any direction* $s_{n,m}$ *computed by* `CG-REGINN` *in its* `repeat`*-loop is a descent direction for* $\varphi(\cdot) = \frac{1}{2}\|y^\delta - F(\cdot)\|_Y^2$. *More precisely,*

$$\langle \nabla\varphi(x_n), s_{n,m}\rangle_X = -2\theta_{n,m}^2\,\varphi(x_n) < 0, \quad 1 \le m \le \boldsymbol{m}_{\mathrm{T}},$$

where $\theta_{n,m} = \|A_n s_{n,m}\|_Y / \|b_n^\varepsilon\|_Y \in]0, 1]$. *Further,*

$$\langle \nabla\varphi(x_n), s_{n,m}\rangle_X \ge \langle \nabla\varphi(x_n), s_{n,m+1}\rangle_X \ge -\big\|P_{\overline{\mathsf{R}(A_n)}} b_n^\varepsilon\big\|_Y^2.$$

Proof. By $\nabla\varphi(\cdot) = -F'(\cdot)^*\big(y^\delta - F(\cdot)\big)$ we find that

$$\langle \nabla\varphi(x_n), s_{n,m}\rangle_X = -\langle b_n^\varepsilon, A_n s_{n,m}\rangle_Y \overset{(2.2)}{=} -\|A_n s_{n,m}\|_Y^2 = -\theta_{n,m}^2 \|b_n^\varepsilon\|_Y^2.$$

Due to Lemma 2.5 $\theta_{n,m}$ is in $]0, 1]$. The monotonicity result follows from (2.5). ∎

CG-REGINN$(x, R, \{\mu_n\}, \mu_{\min})$
%$\mu_{\min}$ is a lower bound for $\{\mu_n\}$
$n := 0$; $x_0 := x$;
while $\|b_n^\varepsilon\|_Y \geq R\delta$ do
{ $m := 0$; $s_{n,0} := 0$; $r^0 = b_n^\varepsilon$; $p^1 := d^0 := A_n^* b_n^\varepsilon$;
repeat
$m := m + 1$;
$q^m := A_n p^m$; $\alpha_m := \|d^{m-1}\|_X^2 / \|q^m\|_Y^2$;
$s_{n,m} := s_{n,m-1} + \alpha_m p^m$;
$r^m := r^{m-1} - \alpha_m q^m$;
$d^m := A_n^* r^m$; $\beta_m := \|d^m\|_X^2 / \|d^{m-1}\|_X^2$;
$p^{m+1} := d^m + \beta_m p^m$;
until $\|r^m\|_Y \leq \mu_n \|b_n^\varepsilon\|_Y$
$m_n := m$;
if $\mu_{\min}\|b_n^\varepsilon\|_Y \leq \|r^{m_n}\|_Y$
$s_n^{\mathrm{N}} := s_{n,m_n}$;
else
$\lambda^\star := \alpha_{m_n}\left(1 - \sqrt{1 + \dfrac{\mu_n^2\|b_n^\varepsilon\|_Y^2 - \|r^{m_n-1}\|_Y^2}{\alpha_{m_n}\|d^{m_n-1}\|_X^2}}\right)$;
%$\lambda^\star$ is determined from $\|A_n(s_{n,m_n-1} + \lambda^\star p^{m_n}) - b_n^\varepsilon\|_Y = \mu_n\|b_n^\varepsilon\|_Y$
%note that $\lambda^\star < \alpha_{m_n}$
$s_n^{\mathrm{N}} := s_{n,m_n-1} + \lambda^\star p^{m_n}$;
endif
$x_{n+1} := x_n + s_n^{\mathrm{N}}$;
$n := n + 1$;
}
$x := x_n$;

Figure 3.1: CG-REGINN with backtracking.

4. A local convergence analysis for CG-REGINN. Throughout we work with the following bound for the linearization error:

$$\begin{aligned}\|E(v,w)\|_Y \leq L\,\|F'(w)(v-w)\|_Y \quad &\text{for one } L < 1\\ &\text{and for all } v, w \in B_\rho(x^+) \subset \mathsf{D}(F).\end{aligned} \tag{4.1}$$

Here, $B_\rho(x^+)$ is the ball of radius $\rho > 0$ centered about x^+. From (4.1) we derive that

$$\|E(v,w)\|_Y \leq \omega\,\|F(w) - F(v)\|_Y \quad \text{where } \omega = \frac{L}{1-L} > L \tag{4.2}$$

which is identical to the tangential cone condition (1.1). Note that $\omega < 1$ whenever $L < 1/2$.

REMARK 4.1. *Actually,* (4.1) *and* (4.2) *are equivalent in the following sense:* (4.2) *for one* $\omega < 1$ *implies* (4.1) *with* $L = \frac{\omega}{1-\omega}$. *Here,* $L < 1$ *whenever* $\omega < 1/2$.

By (1.6) and (4.2) we can further bound the noise in b_n^ε for $x_n \in B_\rho(x^+)$:

$$\|b_n - b_n^\varepsilon\|_Y \le \delta + \omega\|y - F(x_n)\|_Y \le (1+\omega)\delta + \omega\|b_n^\varepsilon\|_Y.$$

As long as $\|b_n^\varepsilon\|_Y \ge R\delta$ we accordingly have

$$\frac{\|b_n - b_n^\varepsilon\|_Y}{\|b_n^\varepsilon\|_Y} \le \frac{1+\omega}{R} + \omega. \tag{4.3}$$

THEOREM 4.2. *Assume* (4.1) *to hold true for* $L < 1$ *so small that there exists a* $\Lambda \in]L, 1]$ *with*

$$(\Lambda - L)^2 > \frac{L}{1-L} = \omega.^\ddagger \tag{4.4}$$

Further, choose

$$R > \frac{1+\omega}{(\Lambda - L)^2 - \omega} > \frac{1+\omega}{1-\omega} \tag{4.5}$$

yielding

$$\mu_{\min} := \sqrt{\omega + \frac{1+\omega}{R}} < 1. \tag{4.6}$$

Restrict all tolerances $\{\mu_n\}$ *to* $[\mu_{\min}, \Lambda - L]$ *and start* `CG-REGINN` *with* $x_0 \in B_\rho(x^+)$.
Then, there exist an $N(\delta)$ *such that all iterates* $\{x_1, \dots, x_{N(\delta)}\}$ *of* `CG-REGINN` *are well defined and stay in* $B_\rho(x^+)$. *We even have a strictly monotone error reduction:*

$$\|x^+ - x_n\|_X < \|x^+ - x_{n-1}\|_X, \quad n = 1, \dots, N(\delta). \tag{4.7}$$

Moreover, only *the final iterate satisfies the discrepancy principle, that is,*

$$\|y^\delta - F(x_{N(\delta)})\|_Y < R\delta \le \|y^\delta - F(x_n)\|_Y, \quad n = 0, \dots, N(\delta) - 1, \tag{4.8}$$

and the nonlinear residuals decrease linearly at an estimated rate

$$\frac{\|y^\delta - F(x_{n+1})\|_Y}{\|y^\delta - F(x_n)\|_Y} \le \mu_n + \theta_n L \le \Lambda, \quad n = 0, \dots, N(\delta) - 1, \tag{4.9}$$

where $\theta_n = \|A_n s_n^{\mathrm{N}}\|_Y / \|b_n^\varepsilon\|_Y \le 1$.

‡In particular, $L < 1/2$.

Proof. Before we start with the proof let us discuss our assumptions on L, ω, Λ and R. Condition (4.4) guarantees that the denominator of the lower bound on R is positive. The lower bound on R and the right condition in (4.4) are needed to make $[\mu_{\min}, \Lambda - L]$ a well-defined non-empty interval.

We will argue inductively. Therefore assume the iterates $x_1, \dots, x_n$ to be well defined in $B_\rho(x^+)$. If $\|b_n^\varepsilon\|_Y < R\delta$ `CG-REGINN` will be stopped with $N(\delta) = n$. Otherwise, $\|b_n^\varepsilon\|_Y \geq R\delta$ and $\mu_n \in [\mu_{\min}, \Lambda - L]$ will provide a new Newton step:

$$\frac{\|P_{\mathsf{R}(A_n)^\perp} b_n^\varepsilon\|_Y}{\|b_n^\varepsilon\|_Y} = \frac{\|P_{\mathsf{R}(A_n)^\perp}(b_n^\varepsilon - b_n)\|_Y}{\|b_n^\varepsilon\|_Y} \overset{(4.3)}{\leq} \frac{1+\omega}{R} + \omega = \mu_{\min}^2 < \mu_{\min}.$$

By Lemma 3.1 the Newton step s_n^{N} and hence $x_{n+1} = x_n + s_n^{\mathrm{N}} \in X$ are well defined. It remains to verify the strictly monotone error reduction. Here we will benefit from Theorem 2.6. Please note that we do *not* need to distinguish the cases $s_n^{\mathrm{N}} = s_{n,m_n}$ and $s_n^{\mathrm{N}} = s_{n,m_n-1} + \lambda^\star p^{m_n}$ because Theorem 2.6 covers both cases.

We apply Theorem 2.6 with $g = b_n^\varepsilon$, $T = A_n$, $f = s_n^{\mathrm{e}} = x^+ - x_n$, and $f_{m-1} = s_{n,m-1}$, $m = 1, \dots, m_n$ and $f_{m_n-1,\lambda} = s_{n,m_n-1} + \lambda\, p^{m_n}$ where λ is either α_{m_n} or $\lambda^\star$. Hence, $\|f_{m-1} - f\|_X = \|x^+ - (x_n + s_{n,m-1})\|_X$, $m = 1, \dots, m_n$, especially, $\|f_0 - f\|_X = \|x^+ - x_n\|_X$, and $\|f_{m_n-1,\lambda} - f\|_X = \|x^+ - x_{n+1}\|_X$. Further,

$$\gamma(f) = \gamma(s_n^{\mathrm{e}}) = \frac{\|b_n^\varepsilon - A_n s_n^{\mathrm{e}}\|_Y}{\|b_n^\varepsilon\|_Y} = \frac{\|b_n^\varepsilon - b_n\|_Y}{\|b_n^\varepsilon\|_Y} \overset{(4.3)}{\leq} \mu_{\min}^2.$$

Since $\|g - Tf_{m-1}\|_Y = \|b_n^\varepsilon - A_n s_{n,m-1}\|_Y \geq \mu_{\min}\|b_n^\varepsilon\|_Y$, $m = 1, \dots, m_n$, and $\|g - Tf_{m_n-1,\lambda}\|_Y = \|b_n^\varepsilon - A_n(s_{n,m_n-1} + \lambda p^{m_n})\|_Y \geq \mu_{\min}\|b_n^\varepsilon\|_Y$ with $\mu_{\min}^2 \geq \gamma(f)$ we have by repeatedly applying Theorem 2.6 that

$$\begin{aligned}\|x^+ - x_{n+1}\|_X &= \|f_{m_n-1,\lambda} - f\|_X \\ &< \|f_{m_n-1} - f\|_X < \|f_{m_n-2} - f\|_X \\ &< \cdots < \|f_0 - f\|_X = \|x^+ - x_n\|_X.\end{aligned}$$

Finally, we prove the reduction rate. Since x_{n+1}, $x_n \in B_\rho(x^+)$ the bound (4.1) can be used to estimate

$$\big\|b_{n+1}^\varepsilon\big\|_Y = \big\|b_n^\varepsilon - A_n s_n^{\mathrm{N}} + E(x_{n+1}, x_n)\big\|_Y \leq \mu_n \big\|b_n^\varepsilon\big\|_Y + L\big\|A_n s_n^{\mathrm{N}}\big\|_Y.$$

The rate (4.9) follows from Lemma 2.5 which implies that $\|A_n s_n^{\mathrm{N}}\|_Y = \theta_n \|b_n^\varepsilon\|_Y$ with $\theta_n \leq 1$ (Again, Lemma 2.5 covers $s_n^{\mathrm{N}} = s_{n,m_n}$ as well as $s_n^{\mathrm{N}} = s_{n,m_n-1} + \lambda^\star p^{m_n}$). Now, the inductive step is complete. ∎

Corollary 4.3. *Adopt all assumptions and notations of Theorem* 4.2. *Additionally let F be weakly sequentially closed and let $\{\delta_j\}_{j\in\mathbb{N}}$ be a positive*

zero sequence.

Then, any subsequence of $\{x_{N(\delta_j)}\}_{j\in\mathbb{N}}$ *contains a subsequence which converges weakly to a solution of* $F(x) = y$.

Proof. Any subsequence of the bounded family $\{x_{N(\delta_j)}\}_{j\in\mathbb{N}} \subset B_\rho(x^+)$ is bounded and, therefore, has a weakly convergent subsequence. Let ξ be its weak limit. Since

$$\|y - F(x_{N(\delta_j)})\|_Y \overset{(4.8)}{<} (R+1)\delta_j \tag{4.10}$$

the images under F of the convergent subsequence converge (weakly) to y. Due to the weak closedness of F we have that $y = F(\xi)$. ∎

The whole family $\{x_{N(\delta_j)}\}_{j\in\mathbb{N}}$ converges weakly to x^+ if x^+ is the unique solution of $F(x) = y$ in $B_\rho(x^+)$. This follows, for instance, from Proposition 10.13 (2) in [17]. However, under the assumptions of Theorem 4.2, the latter can only happen if $\mathsf{N}(A)$, the null space of $A = F'(x^+)$, is trivial. In fact, if $0 \neq v \in \mathsf{N}(A)$ then

$$\|F(x^+ + tv) - y\|_Y = \|F(x^+ + tv) - F(x^+)\|_Y \overset{(4.1)}{\leq} (L+1)\,|t|\,\|Av\|_Y = 0$$

for any $t \in [0, \rho/\|v\|_X]$.

On the other hand, if $\mathsf{N}(A)$ is trivial we even have a norm convergence.

Corollary 4.4. *Under the assumptions of Theorem* 4.2 *we have that*

$$\|x^+ - x_{N(\delta)}\|_A < \frac{1+R}{1-L}\,\delta$$

where $\|\cdot\|_A = \|A\cdot\|_Y$ *is a semi-norm in general.*

Proof. From (4.1) we obtain that

$$\|x^+ - x_{N(\delta)}\|_A \leq \frac{1}{1-L}\,\|y - F(x_{N(\delta)})\|_Y$$

which, in view of (4.10), implies the assertion. ∎

The above corollary yields norm convergence whenever $\mathsf{N}(A) = \{0\}$. In general, this norm is weaker than the standard norm in X. Under an assumption a little bit stronger than (4.1) convergence in X of $\{x_{N(\delta)}\}_{0<\delta\leq\delta_{\max}}$ to x^+ has been shown in [13]. Further, convergence rates have been given.

We close this chapter with further comments.

Remark 4.5. *We emphasize that the factors* θ_n *are virtually smaller than* 1*, so that the rate* (4.9) *is smaller than* $\mu_n + L$*. Only in rare situations, which practically do not occur, some of the factors* θ_n *will be equal to* 1 *(Lemma* 2.5 *and Remark* 2.3*).*

REMARK 4.6. *Some nonlinear ill-posed problems from real-life applications satisfy slightly stronger versions of* (4.1) *or* (4.2) *where* L *or* ω *are replaced by* $C\|v-w\|_X$*, see, e.g., Hanke* [5, Sec. 3] *for an example from groundwater filtration or* [10] *for an example from impedance tomography. Also nonlinear mappings between finite dimensional spaces satisfy the stronger version of* (4.2) *locally about a point where the derivative is injective and Lipschitz continuous, see Appendix* B.

In view of (4.7) *we expect in this situation the reduction rate* (4.9) *to approach* μ_n *as the Newton iteration progresses.*

REMARK 4.7. *Hanke* [6] *investigated a Newton-CG algorithm similar to* `CG-REGINN`*. Two features differ:* `CG-REGINN` *allows an adaptive strategy for the selection of the tolerances* $\{\mu_n\}$ *(see* [11, Sec. 6] *for more details) and the backtracking step (not needed by Hanke at the price of more restrictive assumptions. For instance,* R *and the fixed tolerance* μ *are coupled via* $R\mu^2 > 2$*). While we have been able to prove reduction rates for the nonlinear residuals under* (4.2)*, Hanke only obtained convergence without rates. Hanke's proof of convergence in the noise-free situation applies, with minor changes, to* `CG-REGINN` *with backtracking as well. Observe that* (4.9) *even holds for* $\delta = 0$ *where* $\mu_{\min} = \sqrt{\omega}$ *is allowed by letting* R *approach infinity.*

REMARK 4.8. *A stronger assumption than* (4.1) *is*

$$\begin{aligned} \|E(v,w)\|_Y \le \widetilde{L}\,\|F'(w)(v-w)\|_Y^{1+\kappa} \quad \textit{for one } \kappa > 0 \\ \textit{and for all } v, w \in B_\rho(x^+) \subset \mathsf{D}(F). \end{aligned} \tag{4.11}$$

Here, $\widetilde{L}$ *is allowed to be arbitrarily large. If* ρ *is sufficiently small we have* (4.1) *with*

$$L := 2^\kappa \rho^\kappa \widetilde{L} \max_{u \in B_\rho(x^+)} \|F'(u)\|^\kappa < 1.$$

Now, let ρ *be so small that all assumptions of Theorem* 4.2 *apply with* L *as above. Additionally, choose* $x_0 \in B_\rho(x^+)$ *satisfying* $\|y^\delta - F(x_0)\|_Y^\kappa \le L/\widetilde{L}$.[§] *Then, all assertions of Theorem* 4.2 *remain valid with the stronger rate*

$$\frac{\|y^\delta - F(x_{n+1})\|_Y}{\|y^\delta - F(x_n)\|_Y} \le \mu_n + \theta_n^{1+\kappa}\,\Lambda^{\kappa n}\,L \le \Lambda, \quad n = 0, \dots, N(\delta) - 1. \tag{4.12}$$

We only need to verify the rate. We have

$$\begin{aligned} \big\|b_{n+1}^\varepsilon\big\|_Y = \big\|b_n^\varepsilon - A_n s_n^{\mathrm{N}} + E(x_{n+1}, x_n)\big\|_Y &\overset{(4.11)}{\le} \mu_n \big\|b_n^\varepsilon\big\|_Y + \widetilde{L}\big\|A_n s_n^{\mathrm{N}}\big\|_Y^{1+\kappa} \\ &\le \Big(\mu_n + \widetilde{L}\theta_n^{1+\kappa}\big\|b_n^\varepsilon\big\|_Y^\kappa\Big)\,\big\|b_n^\varepsilon\big\|_Y \end{aligned}$$

[§]This bound implicitly forces $\|F(x^+) - y^\delta\|_Y^\kappa < L/\widetilde{L}$.

which inductively implies (4.12).

REMARK 4.9. *Both bounds* (4.1) *and* (4.11) *for the linearization error may be derived from the following* affine contravariant *Lipschitz condition:*

$$\begin{gathered}\big\|\big(F'(v)-F'(w)\big)(v-w)\big\|_Y \le L_\kappa \,\|F'(w)(v-w)\|_Y^{1+\kappa} \\ \text{for one } \kappa\in[0,1] \text{ and for all } v,\, w\in B_\rho(x^+)\end{gathered} \tag{4.13}$$

where $L_\kappa > 0$ *is the Lipschitz constant and in case* $\kappa = 0$ *we require* $L_0 < 1$. *Indeed,*

$$\begin{aligned}\|E(v,w)\|_Y &= \Big\| \int_0^1 \big(F'\big(w+t(v-w)\big) - F'(w)\big)(v-w)\,\mathrm{d}t\Big\|_Y \\ &\le \frac{L_\kappa}{1+\kappa}\,\|F'(w)(v-w)\|_Y^{1+\kappa}.\end{aligned}$$

For a general discussion of the importance of affine contravariance for Newton-like algorithms we refer to Section 1.2.2 *in the book* [3] *by Deuflhard. In particular, Section* 4.2 *of that book treats Gauß-Newton methods for (well-posed) finite dimensional least squares problems under* (4.13) *with* $\kappa = 1$ *and restricted to* $v, w \in \mathsf{D}(F)$ *where* $v - w \in \mathsf{N}\big(F'(w)\big)^\perp$.

REMARK 4.10. *For linear problems* `CG-REGINN` *does* not *reduce to standard CG but to CG with restart. Theorem* 4.2 *holds under* $L = \omega = 0$, *so that* $R > 1/\Lambda^2$ *and* $\mu_{\min} = \sqrt{1/R}$. *The restart guarantees strictly monotone error decrease until the discrepancy principle is satisfied. This feature is* not *met by standard CG. In contrary, standard CG is known to diverge rapidly if not stopped very close to its optimal stopping point. One likely misses the optimal stopping point when underestimating the noise.*
Please note that CG with restart and stopped by the discrepancy principle is a regularization scheme of optimal order. This fact follows by results from [13].
With Λ *close to* 1 *we may choose* R *close to* 1 *yielding* $\mu_{\min}$ *close to* 1. *In this situation* `CG-REGINN` *is likely to perform only one CG-step per outer iteration and therefore agrees with the nonlinear gradient descent considered by Scherzer* [15].
Conclusion: *For linear problems* `CG-REGINN` *with backtracking is a regularization scheme blending standard CG and steepest descent, thus combining the numerical efficiency of the former method with the strictly monotone error decrease of the latter one.*

5. A numerical example from electrical impedance tomography. In the final section we illustrate the performance of `CG-REGINN` by solving the inverse problem of electrical impedance tomography (EIT). In EIT one tries to recover the electric conductivity of an object by applying

electric currents at the boundary of the object and measuring the resulting voltages also on the boundary. In formulating the corresponding inverse problem mathematically we will be brief and refer to [9] for the mathematics as well as implementation details and for further original references.

Let $B \subset \mathbb{R}^2$ be the domain of interest with electric conductivity σ. Assume that we use p (disjoint) electrodes $E_1, \dots, E_p \subset \partial B$ to apply currents and measure voltages. The so-called *complete electrode model* gives rise to the following forward problem of impedance tomography: Given a current vector $I \in \mathbb{R}^p_\diamond := \{v \in \mathbb{R}^p | \sum_{j=1}^p v_j = 0\}$ and positive contact impedances $z_1, \dots, z_p$ find a voltage potential $u \in H^1(B)$ and a set of electrode voltages $U \in \mathbb{R}^p_\diamond$ such that

$$-\nabla\big(\sigma \nabla u\big) = 0 \quad \text{in } B, \tag{5.1a}$$

$$u + z_j\, \sigma\, \nu \nabla u = U_j \quad \text{on } E_j, \tag{5.1b}$$

$$\int_{E_j} \sigma\, \nu \nabla u \, \mathrm{d}S = I_j \quad \text{for } j = 1, \dots p, \tag{5.1c}$$

$$\sigma\, \nu \nabla u = 0 \quad \text{on } \partial B \setminus \cup_{j=1}^p \overline{E}_j. \tag{5.1d}$$

The conditions $I \in \mathbb{R}^p_\diamond$ and $U \in \mathbb{R}^p_\diamond$ can be interpreted as conservation of charge and grounding the potential, respectively. Both restrictions are necessary to guarantee existence and uniqueness of a weak solution $(u, U) \in H^1(B) \oplus \mathbb{R}^p_\diamond$ of (5.1) for a positive conductivity $\sigma \in L^\infty(B)$ bounded away from zero, see Somersalo et al. [16].

Let us briefly explain the complete electrode model. The domain B is assumed to have no electric sources or drains. Hence, the electric flux $\sigma \nabla u$ is divergence free which yields (5.1a). In a medical application the conductivity between skin and electrodes may be increased by dermal moisture. This effect of contact impedance is taken into account by the Robin boundary condition (5.1b). The equations in (5.1c) model the electrodes as perfect conductors: the total electric flux over an electrode agrees with the electric current on that electrode. As there is no flux over the boundary of B in-between electrodes we have the Neumann boundary condition (5.1d).

The nonlinear forward operator F_p describing the EIT-problem is given by

$$F_p \colon D(F_p) \subset L^\infty(B) \to \mathbb{R}^p_\diamond, \quad \sigma \mapsto U,$$

with domain of definition $D(F_p) = \big\{\sigma \in L^\infty(B) \big| \sigma \geq c_0\big\}$ for some positive constant c_0. Note that F_p depends on I and $z_1, \dots, z_p$. The forward operator is Fréchet differentiable where $F_p'(\sigma)\eta$ is determined as a weak solution of an elliptic problem akin to (5.1), see Kaipio et al. [7, Theorem 2.3].

At present we do not know whether F_p satisfies the tangential cone condition (1.1), however, in the limit case of infinite many electrodes of in-

finitesimal width[¶] the 'limit operator' satisfies a slightly stronger version of (1.1) between suitable Hilbert spaces, see [10].

For our practical computations we restrict B to a regular polygonal domain with $2p$ corners where every second side of the boundary polygon serves as electrode. Further, we only search for conductivities $\sigma \in \mathcal{A} := S_\triangle \cap D(F_p)$ where $S_\triangle$ denotes the space of piecewise constant functions over the triangulation $\triangle$ of B. Finally, we discretize the weak formulation of (5.1) by conforming piecewise linear finite elements with respect to a regular triangulation of B with mesh size h. In applying ℓ different current patterns I^i, $i = 1, \dots, \ell$, we have the nonlinear mapping

$$F_{p,h} \colon \mathcal{A} \subset S_\triangle \to (\mathbb{R}^p_\diamond)^\ell, \quad \sigma \mapsto (U_h^1, \dots, U_h^\ell),$$

where U_h^i is the voltage potential induced by I^i.

If p is sufficiently large and h is sufficiently small then the Jacobian $F'_{p,h}(\sigma) \colon \mathbb{R}^{\dim S_\triangle} \to \mathbb{R}^{p\ell}$ is injective, see [10]. Note that the restriction

$$\dim S_\triangle \leq p\,\ell \tag{5.2}$$

is necessary for injectivity.

According to Appendix B the injectivity of $F'_{p,h}(\sigma^+)$ together with its Lipschitz continuity

$$\|F'_{p,h}(\sigma) - F'_{p,h}(\varsigma)\| \lesssim \|\sigma - \varsigma\|,$$

see again [10], implies the tangential cone condition

$$\|F_{p,h}(\sigma) - F_{p,h}(\varsigma) - F'_{p,h}(\varsigma)(\sigma - \varsigma)\| \lesssim \|\sigma - \varsigma\|\,\|F_{p,h}(\sigma) - F_{p,h}(\varsigma)\| \tag{5.3}$$

in a ball about σ^+ (As $F_{p,h}$ maps between finite-dimensional spaces we do not need to specify norms). Hence, our convergence result for `CG-REGINN` in Theorem 4.2 applies to $F_{p,h}$ in a sufficiently small ball about σ^+, provided the number of electrodes p is large enough and the mesh size h is small enough.

Figure 5.1 shows the `CG-REGINN` iterates for reconstructing a smooth conductivity distribution attaining values above and below the background conductivity 1 which serves as starting guess. We worked with $p = 32$ electrodes and applied the $\ell = p - 1$ linear independent current patterns $I^1 = (1, -1, 0, \dots, 0)^t$, $I^2 = (0, 1, -1, 0, \dots, 0)^t, \dots, I^{p-1} = (-1, 0, \dots, 0, 1)^t$. The number of independent degrees of freedom provided by these data is $p(p-1)/2$, see, e.g., Cheney et al. [2, Sec. 5]. In choosing $S_\triangle$ with $\dim S_\triangle = 446$ we do not try to recover more degrees of freedom than possible and also meet the bound (5.2).

[¶]With the phrase 'limit case of infinite many electrodes' we refer to the situation where the Neumann-to-Dirichlet mapping corresponding to σ is known.

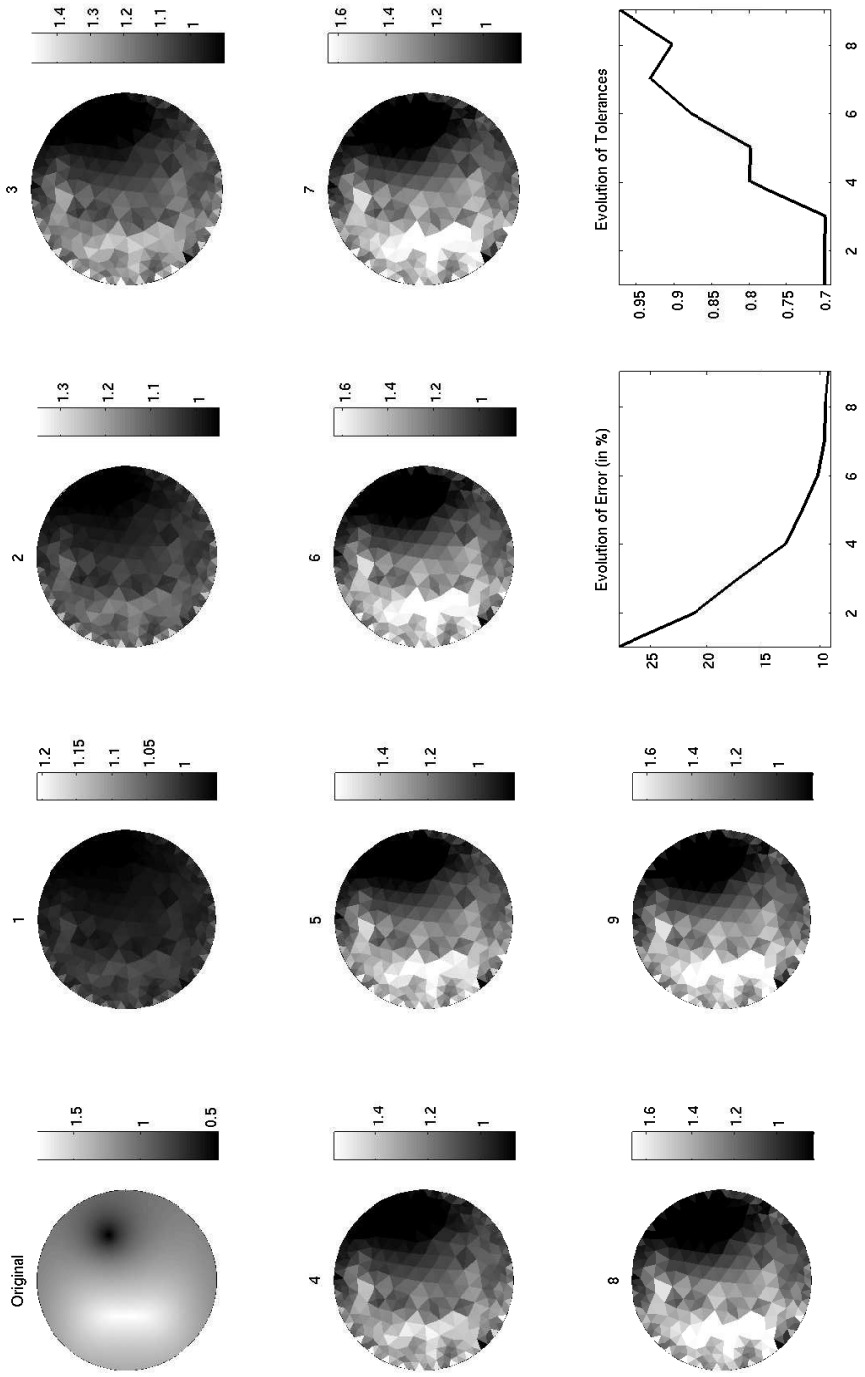


FIG. 5.1. *Reconstruction of a smooth conductivity distribution by* `CG-REGINN`. *The EIT-system has* 32 *electrodes with contact impedances* $z_j = 0.25$, $j = 1, \ldots, 32$, *see* (5.1b), *and* 31 *linear independent current patterns have been applied. The number on top of the reconstructions denotes the iteration number. Two graphs visualize the evolutions of the relative reconstruction errors and of the tolerances as the iteration progresses.*

For our numerical computations we perturbed the simulated voltages $(U_h^1, \dots, U_h^{p-1}) \in \mathbb{R}^{p \times p-1}$ artificially by noise of a relative magnitude of 1% measured in the Frobenius norm. The tolerances $\{\mu_n\}$ are adapted dynamically as explained in [11], see also [9]. In view of (5.3) and (4.6) a theoretically supported choice for the minimal tolerance $\mu_{\min}$ would be a number slightly larger than $\sqrt{1/R}$ (≈ 0.8165 for $R = 1.5$ in our experiment). However, we started with a $\mu_0 = 0.7 < \sqrt{1/R}$ and we did not implement backtracking. Nevertheless `CG-REGINN` worked just fine with a strongly monotone error decrease, see Figure 5.1. Hence, Theorem 4.2 seems not to tell the whole convergence story of `CG-REGINN`.

Appendix A. Proof of Theorem 2.6. We follow the arguments of Hanke [6, Proof of Theorem 3.1] with necessary changes.

First, we introduce an alternative representation of the search directions (2.6): There is a polynomial $w_m \in \Pi_m$ with $w_m(0) > 0$ such that

$$p^m = T^* w_m(TT^*) g,$$

see, e.g, Hanke [4, formula (2.7)].

A straightforward calculation reveals that

$$\begin{aligned}\|f - f_{m-1}\|_X^2 - \|f - f_{m-1,\lambda}\|_X^2 \\ &= \lambda \langle g - Tf_{m-1}, w_m(TT^*)g\rangle_Y + \lambda \langle g - Tf_{m-1,\lambda}, w_m(TT^*)g\rangle_Y \\ &\quad - 2\lambda \langle g - Tf, w_m(TT^*)g\rangle_Y.\end{aligned}$$

To proceed we rewrite w_m as $w_m(t) = w_m(0) + tq(t)$ where $q \in \Pi_{m-1}$. Hence, $w_m(TT^*)g = w_m(0)g + Tu$ with $u = T^*q(TT^*)g \in U_{m-1}$. Applying (2.1) and (2.2) we obtain

$$\begin{aligned}\langle g - Tf_{m-1}, w_m(TT^*)g\rangle_Y &= w_m(0)\langle g - Tf_{m-1}, g\rangle_Y + \langle g - Tf_{m-1}, Tu\rangle_Y \\ &= w_m(0)\|g - Tf_{m-1}\|_Y^2.\end{aligned}$$

Analogously,

$$\langle g - Tf_{m-1,\lambda}, w_m(TT^*)g\rangle_Y = w_m(0)\langle g - Tf_{m-1,\lambda}, g\rangle_Y + \langle g - Tf_{m-1,\lambda}, Tu\rangle_Y.$$

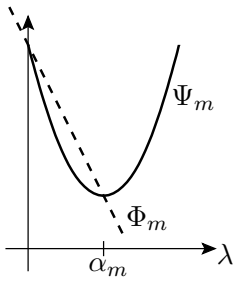
Here, things are little bit more involved. We first investigate the right-most term:

$$\langle g - Tf_{m-1,\lambda}, Tu\rangle_Y = \underbrace{\langle g - Tf_{m-1}, Tu\rangle_Y}_{=0 \text{ by (2.1)}} - \lambda \underbrace{\langle Tp^m, Tu\rangle_Y}_{=0 \text{ by (2.9)}} = 0.$$

Now we will verify that

$$\Phi_m(\lambda) := \langle g - Tf_{m-1,\lambda}, g\rangle_Y \geq \|g - Tf_{m-1,\lambda}\|_Y^2 = \Psi_m(\lambda)$$

for $\lambda \in [0, \alpha_m]$. The graph of Ψ_m is a strict convex parabola with vertex $(\alpha_m, \|g - Tf_m\|_Y^2)$, see (2.7) and (2.8). The graph of the linear function Φ_m connects $(0, \|g - Tf_{m-1}\|_Y^2)$ with the vertex $(\alpha_m, \|g - Tf_m\|_Y^2)$. Therefore, $\Phi_m(\lambda) \geq \Psi_m(\lambda)$, $\lambda \in [0, \alpha_m]$, see figure on the right. Finally we conclude that



$$\langle g - Tf_{m-1,\lambda}, w_m(TT^*)g\rangle_Y \geq w_m(0)\|g - Tf_{m-1,\lambda}\|_Y^2$$

and

$$\|f - f_{m-1}\|_X^2 - \|f - f_{m-1,\lambda}\|_X^2 \geq \lambda\, w_m(0)\Big(\|g - Tf_{m-1}\|_Y^2 + \|g - Tf_{m-1,\lambda}\|_Y^2 - 2\,\|g - Tf\|_Y\,\frac{\|w_m(TT^*)g\|_Y}{w_m(0)}\Big).$$

The normalized polynomial $w_m/w_m(0)$ is denoted $p_m^{[2]}$ by Hanke [4]. By his Theorem 3.2 we have

$$\frac{\|w_m(TT^*)g\|_Y}{w_m(0)} < \frac{\|w_0(TT^*)g\|_Y}{w_0(0)} = \|g\|_Y.$$

Our assumptions together with $\mu\|g\|_Y \leq \|g - Tf_{m-1,\lambda}\|_Y \leq \|g - Tf_{m-1}\|_Y$, $\lambda \in]0, \alpha_m]$, yield

$$\|f_{m-1} - f\|_X^2 - \|f_{m-1,\lambda} - f\|_X^2 > 2\lambda\, w_m(0)\|g\|_Y^2\big(\mu^2 - \gamma(f)\big) \geq 0$$

which completes the proof of Theorem 2.6.

Appendix B. Tangential cone condition in finite dimensional spaces. Let $F : D(F) \subset \mathbb{R}^n \to \mathbb{R}^m$, $n \leq m$, be a differentiable mapping. We will show that there is a $\rho > 0$ such that

$$\|F(v) - F(w) - F'(w)(v - w)\| \lesssim \|v - w\|^\alpha \|F(v) - F(w)\| \quad \text{for all } v, w \in B_\rho(x^+) \tag{B.1}$$

whenever $F'(x^+)$ has a trivial null space and F' is Hölder continuous of order α, that is,

$$\|F'(x) - F'(y)\| \leq L\|x - y\|^\alpha \quad \text{for all } v, w \in B_r(x^+)$$

for one $r > 0$ and one $L > 0$.

We prove (B.1) only for the Euclidean norm. Then, (B.1) holds true for any norm by norm equivalence. By the Hölder continuity of F' there is a $r_1 \in]0, r]$ such that

$$\|(F'(v)^*F'(v))^{-1}\| \leq 2\|(F'(x^+)^*F'(x^+))^{-1}\| =: M \quad \text{for all } v \in B_{r_1}(x^+).$$

For $E(v,w) := F(v) - F(w) - F'(w)(v-w)$ we have that

$$\begin{aligned}\|E(v,w)\| &= \Big\| \int_0^1 \big(F'(w+t(v-w)) - F'(w)\big)(v-w)\,\mathrm{d}t \Big\| \\ &\leq \frac{L}{1+\alpha}\|v-w\|^{1+\alpha} \quad \text{for all } v,w \in B_r(x^+).\end{aligned}$$

We proceed with

$$\begin{aligned}\|F(v)-F(w)\|^2 &= \|E(v,w) + F'(w)(v-w)\|^2 \\ &= \|F'(w)(v-w)\|^2 + \|E(v,w)\|^2 \\ &\qquad + 2\langle E(v,w), F'(w)(v-w)\rangle \\ &\geq M^{-1}\|v-w\|^2 + 2\langle E(v,w), F'(w)(v-w)\rangle \\ &\geq M^{-1}\|v-w\|^2 - \frac{2L}{1+\alpha}K\|v-w\|^{2+\alpha}\end{aligned}$$

where $K = \sup_{u\in B_{r_1}(x^+)} \|F'(u)\|$. There exists a $r_2 \in]0, r_1]$ such that

$$M^{-1} - \frac{2L}{1+\alpha}K\|v-w\|^{\alpha} \geq M^{-1}/2 \quad \text{for all } v,w \in B_{r_2}(x^+).$$

Hence,

$$\|F(v)-F(w)\| \geq \frac{1}{\sqrt{2M}}\,\|v-w\| \quad \text{for all } v,w \in B_{r_2}(x^+).$$

Finally,

$$\|E(v,w)\| \leq \frac{L}{1+\alpha}\|v-w\|^{1+\alpha} \leq \frac{L\sqrt{2M}}{1+\alpha}\|v-w\|^{\alpha}\,\|F(v)-F(w)\|$$

which is (B.1) with $\rho = r_2$.

REFERENCES

[1] A. B. Bakushinsky and M. Y. Kokurin, *Iterative Methods for Approximate Solution of Inverse Problems*, vol. 577 of Mathematics and Its Applications, Springer, Dordrecht, The Netherlands, 2004.

[2] M. Cheney, D. Isaacson, and J. Newell, *Electrical impedance tomography*, SIAM Review, 41 (1999), pp. 85–101.

[3] P. Deuflhard, *Newton Methods for Nonlinear Problems*, vol. 35 of Springer Series in Computational Mathematics, Springer, Berlin, 2004.

[4] M. Hanke, *Conjugate Gradient Type Methods for Ill-Posed Problems*, vol. 327 of Pitman Research Notes in Mathematics, Longman Scientific & Technical, Harlow, UK, 1995.

[5] ———, *A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems, 13 (1997), pp. 79–95.

[6] ———, *Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems*, Numer. Funct. Anal. Optim., 18 (1998), pp. 971–993.

[7] J. Kaipio, V. Kolehmainen, E. Somersalo, and M. Vaukhonen, *Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography*, Inverse Problems, 16 (2000), pp. 1487–1522.

[8] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, 2007. To appear.

[9] A. Lechleiter and A. Rieder, *Newton regularizations for impedance tomography: a numerical study*, Inverse Problems, 22 (2006), pp. 1967–1887.

[10] ———, *On the ill-posedness of impedance tomography: the tangential cone condition*, 2006. Work in progress.

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

[12] ———, *Keine Probleme mit Inversen Problemen*, Vieweg, Wiesbaden, Germany, 2003.

[13] ———, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal., 43 (2005), pp. 604–622.

[14] O. Scherzer, *The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems*, Computing, 51 (1993), pp. 45–60.

[15] ———, *A convergence analysis of a method of steepest descent and a two-step algorithm for nonlinear ill-posed problems*, Numer. Funct. Anal. and Optimiz., 17 (1996), pp. 197–214.

[16] E. Somersalo, M. Cheney, and D. Isaacson, *Existence and uniqueness for electrode models for electric current computed tomography*, SIAM J. Appl. Math., 52 (1992), pp. 1023–1040.

[17] E. Zeidler, *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*, Springer-Verlag, New York, 1993.

IWRMM-Preprints seit 2004

Nr. 04/01 Andreas Rieder: Inexact Newton Regularization Using Conjugate Gradients as Inner IteractionMichael

Nr. 04/02 Jan Mayer: The ILUCP preconditioner

Nr. 04/03 Andreas Rieder: Runge-Kutta Integrators Yield Optimal Regularization Schemes

Nr. 04/04 Vincent Heuveline: Adaptive Finite Elements for the Steady Free Fall of a Body in a Newtonian Fluid

Nr. 05/01 Götz Alefeld, Zhengyu Wang: Verification of Solutions for Almost Linear Complementarity Problems

Nr. 05/02 Vincent Heuveline, Friedhelm Schieweck: Constrained H^1-interpolation on quadrilateral and hexahedral meshes with hanging nodes

Nr. 05/03 Michael Plum, Christian Wieners: Enclosures for variational inequalities

Nr. 05/04 Jan Mayer: ILUCDP: A Crout ILU Preconditioner with Pivoting and Row Permutation

Nr. 05/05 Reinhard Kirchner, Ulrich Kulisch: Hardware Support for Interval Arithmetic

Nr. 05/06 Jan Mayer: ILUCDP: A Multilevel Crout ILU Preconditioner with Pivoting and Row Permutation

Nr. 06/01 Willy Dörfler, Vincent Heuveline: Convergence of an adaptive *hp* finite element strategy in one dimension

Nr. 06/02 Vincent Heuveline, Hoang Nam-Dung: On two Numerical Approaches for the Boundary Control Stabilization of Semi-linear Parabolic Systems: A Comparison

Nr. 06/03 Andreas Rieder, Armin Lechleiter: Newton Regularizations for Impedance Tomography: A Numerical Study

Nr. 06/04 Götz Alefeld, Xiaojun Chen: A Regularized Projection Method for Complementarity Problems with Non-Lipschitzian Functions

Nr. 06/05 Ulrich Kulisch: Letters to the IEEE Computer Arithmetic Standards Revision Group

Nr. 06/06 Frank Strauss, Vincent Heuveline, Ben Schweizer: Existence and approximation results for shape optimization problems in rotordynamics

Nr. 06/07 Kai Sandfort, Joachim Ohser: Labeling of n-dimensional images with choosable adjacency of the pixels

Nr. 06/08 Jan Mayer: Symmetric Permutations for I-matrices to Delay and Avoid Small Pivots During Factorization

Nr. 06/09 Andreas Rieder, Arne Schneck: Optimality of the fully discrete filtered Backprojection Algorithm for Tomographic Inversion

Nr. 06/10 Patrizio Neff, Krzysztof Chelminski, Wolfgang Müller, Christian Wieners: A numerical solution method for an infinitesimal elasto-plastic Cosserat model

Nr. 06/11 Christian Wieners: Nonlinear solution methods for infinitesimal perfect plasticity

Nr. 07/01 Armin Lechleiter, Andreas Rieder: A Convergenze Analysis of the Newton-Type Regularization CG-Reginn with Application to Impedance Tomography

Eine aktuelle Liste aller IWRMM-Preprints finden Sie auf:

www.mathematik.uni-karlsruhe.de/iwrmm/seite/preprints