Local Ill–Posedness and Source Conditions of Operator Equations in Hilbert Spaces †

Bernd Hofmann
Fakultät für Mathematik
Technische Universität Chemnitz
D–09107 Chemnitz, Germany

and

Otmar Scherzer
Institut für Mathematik
Johannes–Kepler–Universität Linz
A–4040 Linz, Austria

1991 Mathematics Subject Classifications:
65 J15, 65 J20, 47 H15, 47 H17
Keywords: Nonlinear operator equations, local ill–posedness, Fréchet derivative, source conditions, stability estimates, ill–posedness measures, a posteriori estimates

Abstract

The characterization of the local ill–posedness and the local degree of nonlinearity are of particular importance for the stable solution of nonlinear ill–posed problems. We present assertions concerning the interdependence between the ill-posedness of the nonlinear problem and its linearization. Moreover, we show that the concept of the degree of nonlinearity combined with source conditions can be used to characterize the local ill–posedness and to derive a posteriori estimates for nonlinear ill–posed problems. A posteriori estimates are widely used in finite element and multigrid methods for the solution of nonlinear partial differential equations, but these techniques are in general not applicable to inverse and ill–posed problems. Additionally we show for the well–known Landweber method and the iteratively regularized Gauß–Newton method that they satisfy a posteriori estimates under source conditions; this can be used to prove convergence rates results.

†The work of B.H. is supported in part by the Alexander von Humboldt Foundation Bonn (Germany) and by the Johannes–Kepler–University Linz (Austria), the work of O.S. is supported in part by the Christian Doppler Society (Austria) and by the Fonds zur Förderung der Wissenschaftlichen Forschung (Austria) , SFB F1310

1 Introduction

In this paper we consider some local ill–posedness of inverse problems, which can be written in form of an operator equation

$$F(x) \;=\; y_0\,, \tag{1.1}$$

where the (possibly nonlinear) operator

$$F : D(F) \subseteq H_1 \to H_2,$$

with domain $D(F)$, is defined between separable infinite dimensional Hilbert spaces H_1 and H_2. We denote by $\|\cdot\|$ and $\langle\cdot\,,\,\cdot\rangle$ norms and inner products, respectively. We focus our attention to the local behaviour of equation (1.1) in a neighbourhood

$$B_r(x^\dagger) := \{x \in H_1 : \|x - x^\dagger\| \le r\} \quad (r > 0)$$

of a solution $x^\dagger \in D(F)$ satisfying $F(x^\dagger) = y_0$.

In the sequel of this paper indirect data y_0 of $x^\dagger$ are observed, but these data are in general superposed by some perturbations (e.g. measurement errors).

In the setting of this paper well–posedness of the problem of solving (1.1) means that $x^\dagger$ can be recovered accurately whenever data are available that approximate y_0. Otherwise we have local ill–posedness of (1.1) at $x^\dagger$. This local ill–posedness complicating the practical reconstruction process of $x^\dagger$ may have two reasons:

- If $x^\dagger$ fails to be an isolated solution point of (1.1), i.e., we have points $x_1^\dagger \in D(F)$ with $F(x_1^\dagger) = y_0$ and $x_1^\dagger \ne x^\dagger$ in all balls $B_r(x^\dagger)$ for arbitrarily small $r > 0$, then $x^\dagger$ cannot be recovered from the data uniquely even if y_0 is exactly known.
- The (possibly multi-valued) inverse mapping F^{-1} is not continuous in y_0.

The above discussion motivates the following definition of local ill–posedness and local well–posedness of (1.1) in a solution point $x^\dagger$.

Definition 1.1 *We call the operator equation (1.1) locally ill–posed at $x^\dagger \in D(F)$ if there exist, for all $r > 0$, sequences $\{x_n\} \subset B_r(x^\dagger) \cap D(F)$ satisfying the condition*

$$\|F(x_n) - F(x^\dagger)\| \to 0\,, \quad \textit{but} \quad \|x_n - x^\dagger\| \not\to 0\,, \quad \textit{as} \quad n \to \infty\,.$$

Otherwise, we call the equation (1.1) locally well–posed at $x^\dagger$, i.e., there exists an $r > 0$ such that, for all sequences $\{x_n\} \in B_r(x^\dagger) \cap D(F)$, convergence of $\|F(x_n) - F(x^\dagger)\|$ to zero implies convergence of $\|x_n - x^\dagger\|$ to zero.

If we apply Definition 1.1 to a linear operator equation

$$(1.2) \qquad A\,x \;=\; y\,,$$

where $A \in \mathcal{L}(H_1, H_1)$ is a bounded linear operator with domain $D(A) = H_1$, then the linear equation (1.2) is either well–posed or ill–posed for all $x \in H_1$.

Definition 1.2 *We call the linear operator equation (1.2) intrinsically well–posed (intrinsically ill–posed) if (1.2) is locally well–posed (locally ill–posed) at all points $x \in H_1$ in the sense of Definition 1.1.*

From the theory of bounded linear operators in Hilbert spaces (see e.g. [9]) we immediately obtain the following lemma characterizing the intrinsic well–posedness and ill–posedness of equations (1.2)

Lemma 1.3 *The linear operator equation (1.2) is intrinsically well–posed if and only if A is injective, i.e., the null-space $N(A) = \{0\}$ is trivial, and the range $R(A) = \overline{R(A)}$ is a closed subspace in H_2. Intrinsic well–posedness of (1.2) is characterized by a bounded Moore-Penrose inverse $A^\dagger \in \mathcal{L}(H_1, H_1)$, where $A^\dagger A = I$ is the identity operator in H_1.*

If we think on iterative methods for solving a nonlinear operator equation (1.1), then linear equations (1.2) occur from the linearization of the norm of the residual $\|F(x) - y^\delta\|^2$ in any iteration step (cf. e.g. the Landweber iteration [6, 10] and the iteratively regularized Gauß-Newton method [5, 3, 4]).

In particular the Fréchet derivatives $F'(x)$ are useful to characterize the strength (or degree) of local ill–posedness for a nonlinear ill–posed inverse problem whenever the Fréchet derivatives characterize F in a neighbourhood of $x^\dagger$ sufficiently well. In the papers [8] and [11] there are given some ideas and motivations for defining the local degree of ill–posedness of nonlinear inverse problems at a solution point $x^\dagger$ using ill–posedness measures of the Fréchet derivative $F'(x^\dagger)$, which can realistically assumed to be a compact linear operator. Namely, for linear ill–posed problems (1.2) with compact operators A ill–posedness measures (in particular, decay rates to zero of the singular values of A) have been studied comprehensively (cf. e.g. [14]). This approach assumes that the operator $F'(x^\dagger)$ is able to express the essential local ill–posedness properties of F in the point $x^\dagger$.

Therefore it makes sense to confront the nonlinear equation (1.1) with its linearization

$$(1.3) \qquad F'(x^\dagger)x = F'(x^\dagger)x^\dagger$$

at the point $x^\dagger$. In order to exploit the properties of $F'(x^\dagger)$ for evaluating the strength of local ill–posedness of F at $x^\dagger$ it would be desirable that the operator equations (1.1) and (1.3) show a similar behaviour with respect to well–posedness and ill–posedness.

In [17] the linearization is used to specify subsets $U_{\text{a-p}}(x^\dagger)$ where, although the problem of solving (1.1) on $D(F)$ is locally ill–posed at $x^\dagger$ in the sense of Definition 1.1, the problem is locally well–posed on the set $U_{\text{a-p}}(x^\dagger)$. In [17], even more, we were able to construct sets $U_{\text{a-p}}(x^\dagger)$ (depending on the linearization) where generalized a posteriori estimates hold. A posteriori estimates of general type allow to estimate the error of the solution by its residual error, i.e.,

$$\|x - x^\dagger\| \le f(\|F(x) - F(x^\dagger)\|) \,. \tag{1.4}$$

A posteriori error estimates play an important role in the theory of numerical solutions of partial differential equations with multi–grid and multi–level methods. As we will see below a general view of a posteriori estimates is particularly useful in the analysis of *iterative regularization techniques* for the solution of *inverse problems.*

2 Aspects of interdepence between the nonlinear and the linearized problem

Assumption 2.1 *Throughout this section we pose the following requirements:*

- *There exists a ball $B_r(x^\dagger)$ with radius $r > 0$ such that*

$$B_r(x^\dagger) \subseteq D(F). \tag{2.1}$$

- *The operator F is Fréchet differentiable from $D(F) \subseteq H_1$ into H_2 and the Fréchet derivative $F'(x)$ is continuous and locally Lipschitz continuous, i.e., there exists a constant L such that*

$$\|F'(x) - F'(x^\dagger)\| \;\le\; L\,\|x - x^\dagger\| \tag{2.2}$$

for all $x \in B_r(x^\dagger)$.

In [13] we have defined the *degree of nonlinearity* of a Fréchet differentiable mapping F at a point $x^\dagger$ as a vector of real numbers (c_1, c_2, c_3) satisfying $0 \le c_1, c_2 \le 1$ and $0 \le c_3 \le 2$, such that for all $x \in B_r(x^\dagger)$

$$\left\{ \begin{array}{ll} & \|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \\ \le & K\,\|F'(x^\dagger)(x - x^\dagger)\|^{c_1}\,\|F(x) - F(x^\dagger)\|^{c_2}\,\|x - x^\dagger\|^{c_3} \,. \end{array} \right. \tag{2.3}$$

In that paper we also presented some arguments that reveal some connections between the nonlinear operator F in a neighbourhood of $x^\dagger$ and the linear operator $F'(x^\dagger)$ if the operator F is of degree (c_1, c_2, c_3) with $c_2 = 1$. This situation occurs in particular if there exists a constant $\eta < 1$ such that for all $x \in B_r(x^\dagger)$ the following condition is satisfied

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \le \eta\,\|F(x) - F(x^\dagger)\| \,. \tag{2.4}$$

In the remaining of this paper we will refer to this condition as the "η–condition".

By straightforward calculations it can be seen that an operator F which satisfies the η–condition also satisfies

$$(2.5)\qquad (1-\eta)\,\|F'(x^\dagger)(x-x^\dagger)\| \;\le\; \|F(x)-F(x^\dagger)\| \;\le\; (1+\eta)\,\|F'(x^\dagger)(x-x^\dagger)\|$$

for all $x \in B_r(x^\dagger)$. Therefore the η condition guarantees a correlation between the nonlinear operator and its linearization.

In general Assumption 2.1 guarantees only a very weak correlation between the nonlinear and the linearized problem. In the very general situation that F is nonlinear of degree $(0,0,2)$ the norm of the remainder in the Taylor–series expansion, i.e., the norm of $F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger)$, has locally no majorants of the form $\tilde{K}\,\|F(x)-F(x^\dagger)\|^{c_2}$ $(c_2>0)$, and therefore there exists (in general) no correlation between the nonlinear operator and its linearization.

One aspect of finding correlations between the nonlinear operator F and its derivative $F'(x^\dagger)$ is to deduce local ill–posedness of the equation (1.1) from the intrinsic ill–posedness of (1.3) and vice versa. The following results show that such correlations can be found if the operator is nonlinear of degree $(0,1,0)$ (for the proofs see [12]):

Theorem 2.2 *Let the operator F of equation (1.1) satisfy an inequality*

$$(2.6)\qquad \|F(x)-F(x^\dagger)\| \;\le\; C_1\,\|F'(x^\dagger)(x-x^\dagger)\| \qquad \textit{for all}\quad x\in B_r(x^\dagger)$$

with a positive constant C_1. If then (1.1) is locally well–posed at $x^\dagger$, the associated linearized problem (1.3) is intrinsically well–posed. On the other hand, if (1.3) is intrinsically ill–posed, then (1.1) is locally ill–posed at $x^\dagger$.

Theorem 2.3 *Let the operator F of equation (1.1) satisfy an inequality*

$$(2.7)\qquad \|F'(x^\dagger)(x-x^\dagger)\| \;\le\; C_2\,\|F(x)-F(x^\dagger)\| \qquad \textit{for all}\quad x\in B_r(x^\dagger)$$

with a positive constant C_2. If then (1.1) is locally ill–posed at $x^\dagger$, the associated linearized problem (1.3) is intrinsically ill–posed. On the other hand, if (1.3) is intrinsically well–posed, then (1.1) is locally well–posed at $x^\dagger$.

As a consequence of both theorems we obtain:

Corollary 2.4 *Let the operator F of equation (1.1) satisfy the inequalities*

$$(2.8)\qquad \begin{gathered}\underline{C}\,\|F'(x^\dagger)(x-x^\dagger)\| \;\le\; \|F(x)-F(x^\dagger)\| \;\le\; \overline{C}\,\|F'(x^\dagger)(x-x^\dagger)\| \\ \textit{for all}\quad x\in B_r(x^\dagger)\,,\end{gathered}$$

where $0<\underline{C}\le\overline{C}<\infty$. Then (1.1) is locally well–posed at $x^\dagger$ if and only if (1.3) is intrinsically well–posed. Consequently, (1.1) is locally ill–posed at $x^\dagger$ if and only if (1.3) is intrinsically ill–posed.

Note that Corollary 2.4 applies in particular if an η-condition (2.4) with $0<\eta<1$ holds. Based on the equivalence of well–posedness and ill–posedness between nonlinear and corresponding linearized problems provided that (2.8) is valid, we can even show that under the Assumption 2.1 an equation (1.1) locally ill–posed at $x^\dagger$ can never correspond to an intrinsically well–posed linearized equation (1.3).

Theorem 2.5 *Let equation (1.1) be locally ill–posed at $x^\dagger$. Then the linearized problem (1.3) is always intrinsically ill–posed.*

Proof: In order to find a contradiction, we assume that (1.1) is locally ill–posed at $x^\dagger$, but (1.3) is intrinsically well–posed. That means, the bounded linear operator $F'(x^\dagger)$ is injective, i.e. $N(F'(x^\dagger))=\{0\}$, has a closed range $R(F'(x^\dagger))=\overline{R(F'(x^\dagger))}$ and therefore there exists a bounded Moore–Penrose inverse $F'(x^\dagger)^\dagger \in \mathcal{L}(H_2,H_1)$. This implies $R(F'(x^\dagger)^*)=H_1$, since $[F'(x^\dagger)^*]^\dagger=[F'(x^\dagger)^\dagger]^* \in \mathcal{L}(H_1,H_2)$, $F'(x^\dagger)^*$ has a closed range and $R(F'(x^\dagger)^* \oplus N(F'(x^\dagger))=H_1$ (cf.[15]). Consequently, for all $x\in H_1$ there exists a uniquely determined element $w \in H_2 \ominus N(F'(x^\dagger)^*)$ satisfying the source condition

$$x-x^\dagger \;=\; F'(x^\dagger)^*\, w, \tag{2.9}$$

where $w=[F'(x^\dagger)^*]^\dagger(x-x^\dagger)$ and

$$\|w\| \;\le\; \|F'(x^\dagger)^\dagger\|\, \|x-x^\dagger\|. \tag{2.10}$$

If we choose a constant μ with $0<\mu<1$ and set $r:=\frac{2\mu}{L\,\|F'(x^\dagger)^\dagger\|}$, then we have $\|w\|\le\frac{2\mu}{L}$ for all $x\in B_r(x^\dagger)$. Now, we consider the remainder in the Taylor–series expansion

$$\begin{aligned} R \;&:=\; F(x)-F(x^\dagger)-F'(x^\dagger)(x-x^\dagger) \\ &=\; \int_0^1 \Big(F'(x^\dagger+t(x-x^\dagger))-F'(x^\dagger)\Big)\,(x-x^\dagger)\,dt. \end{aligned} \tag{2.11}$$

From Assumption 2.1 and (2.8) it follows that for all $x\in B_r(x^\dagger)$

$$\begin{aligned} \|R\| \;&\le\; \tfrac{L}{2}\|x-x^\dagger\|^2 \\ &=\; \tfrac{L}{2}\|F'(x^\dagger)^*\,w\|^2 \\ &=\; \tfrac{L}{2}\Big\langle F'(x^\dagger)F'(x^\dagger)^*w\,,\,w\Big\rangle \\ &\le\; \tfrac{L\|w\|}{2}\|F'(x^\dagger)F'(x^\dagger)^*w\| \\ &=\; \tfrac{L\|w\|}{2}\|F'(x^\dagger)(x-x^\dagger)\| \\ &\le\; \mu\|F'(x^\dagger)(x-x^\dagger)\|\;. \end{aligned}$$

By the triangle inequality we obtain

$$\|F'(x^\dagger)(x-x^\dagger)\| \le \|R\|+\|F(x)-F(x^\dagger)\| \le \mu\|F'(x^\dagger)(x-x^\dagger)\|+\|F(x)-F(x^\dagger)\|$$

and as a consequence

$$\|F'(x^\dagger)(x-x^\dagger)\| \le \frac{1}{1-\mu}\,\|F(x)-F(x^\dagger)\| \qquad \text{for all} \quad x \in B_r(x^\dagger).$$

By Theorem 2.3 we can conclude that (1.3) is intrinsically ill–posed. This contradicts the assumption that (1.3) is intrinsically well–posed and proves the theorem. □

As a consequence of Theorem 2.5 the condition (2.7) in Theorem 2.3 is superfluous. In spite of this result we must not conjecture that the equivalence of Corollary 2.4 is also valid in general even if condition (2.8) is violated. Namely, one can easily find examples, where the nonlinear problem (1.1) is locally well–posed at $x^\dagger$, but the Fréchet derivative $F'(x^\dagger)$ is non-injective. In such a case, the corresponding linearized problem (1.3) is intrinsically ill–posed.

Now, we are interested in studying situations, where the intrinsic ill–posedness of (1.3) carries over to the local ill–posedness of (1.1) at $x^\dagger$. From

$$\text{(2.12)} \quad \|F(x)-F(x^\dagger)\| = \left\| \int_0^1 F'(x^\dagger + t(x-x^\dagger))(x-x^\dagger)\,dt \right\| \qquad \text{for all} \quad x \in B_r(x^\dagger)$$

it becomes obvious that (2.6) cannot generally be derived from properties of $F'(x^\dagger)$ without additional assumptions on F in a neighbourhood of $x^\dagger$, since $\|F(x)-F(x^\dagger)\|$ depends on the whole family of operators $F'(x^\dagger+t(x-x^\dagger))$ for $0 \le t \le 1$. For an assertion on the local well–posedness or ill–posedness of (1.1) at $x^\dagger$ when (1.3) is intrinsically ill–posed, the coupling of the Fréchet derivatives in $B_r(x^\dagger)$ plays an important role.

There exist kinds of coupling between $F'(x^\dagger)$ and $F'(x)$ ($x \in B_r(x^\dagger)$) ensuring the local ill–posedness of (1.1) at $x^\dagger$ provided that the linearization (1.3) is intrinsically ill–posed.

Theorem 2.6 *For all $x \in B_r(x^\dagger)$ let exist bounded linear operators $G_x \in \mathcal{L}(H_2,H_2)$ depending on x and a uniform constant $\overline{K} > 0$ such that*

$$\text{(2.13)} \qquad F'(x) \,=\, G_x\,F'(x^\dagger), \qquad \|G_x\| \,\le\, \overline{K} < \infty.$$

If then (1.3) is intrinsically ill–posed, the equation (1.1) is locally ill–posed at $x^\dagger$. If in particular the null-space $N(F'(x^\dagger))$ of $F'(x^\dagger)$ is non-trivial, then $x^\dagger$ is no isolated solution of equation (1.1).

Proof: From formula (2.12) we obtain for all $x \in B_r(x^\dagger)$ with (2.13) the inequalities

$$\|F(x)-F(x^\dagger)\| \,\le\, \int_0^1 \|G_{x^\dagger+t(x-x^\dagger)}\|\,\|F'(x^\dagger)(x-x^\dagger)\|\,dt \le \overline{K}\,\|F'(x^\dagger)(x-x^\dagger)\|.$$

This guarantees an estimate of the form (2.6) and consequently Theorem 2.2 applies. If (2.6) holds and $F'(x^\dagger)$ has a non-trivial null-space, then there is an element $v \in B_r(x^\dagger)$, where $F'(x^\dagger)\,v = 0$ and $v \ne 0$, and all elements $x^\dagger + tv$ $(0 \le t \le 1) \in B_r(x^\dagger)$ are also solutions of equation (1.1). □

3 Stability and source conditions

In this section we restrict our attention to ellipsoids generated by source conditions. At the beginning we consider the condition

$$(3.1) \qquad x - x^\dagger \;=\; F'(x^\dagger)^*\, w \qquad (w \in H_2,\ \|w\| \le \tau),$$

where τ is a sufficiently small positive constant.

In the first part of this section let hold the following assumption:

Assumption 3.1 *Let*

$$(3.2) \qquad D(F) \;:=\; \{x \in H_1 :\ x \ \textit{satisfies}\ (3.1)\}\,.$$

Moreover, let F be Fréchet differentiable with Lipschitz–continuous derivative, i.e., for all $x \in D(F)$, F satisfies (2.2).

In addition to the operator F and equation (1.1) we consider a reference operator $\tilde{F} : D(\tilde{F}) := D(F) \subseteq H_1 \to H_2$ and a corresponding reference equation

$$(3.3) \qquad \tilde{F}(x) \;=\; y\ .$$

From Definition 1.1 we immediately derive the following lemma:

Lemma 3.2 *Let F satisfy*

$$(3.4) \quad \underline{c}\,\|\tilde{F}(x) - \tilde{F}(x^\dagger)\| \;\le\; \|F(x) - F(x^\dagger)\| \;\le\; \overline{c}\,\|\tilde{F}(x) - \tilde{F}(x^\dagger)\| \ \textit{for all}\ x \in D(F),$$

where $0 < \underline{c} \le \overline{c} < \infty$ are constants. Then (1.1) is locally well–posed at $x^\dagger$ if and only if (3.3) is locally well–posed at $x^\dagger$. Consequently, (1.1) is locally ill–posed at $x^\dagger$ if and only if (3.3) is locally ill–posed at $x^\dagger$.

Now we consider a constrained version of the linearized problem (1.3) with a source condition domain (3.2) and $\tilde{F}(x) := F'(x^\dagger)x$. Then, we have the following lemma:

Lemma 3.3 *The constrained version of the linearized equation (1.3), where*

$$(3.5) \qquad F'(x^\dagger) : D(F) \subset H_1 \to H_2,$$

is always locally well–posed at $x^\dagger$. Namely, we have the stability estimate

$$(3.6) \qquad \|x - x^\dagger\| \;\le\; \sqrt{\tau}\,\sqrt{\|F'(x^\dagger)\,x - F'(x^\dagger)\,x^\dagger\|} \qquad \textit{for all} \quad x \in D(F).$$

Proof: From (3.1) it follows that

$$\begin{aligned}\|x-x^\dagger\| &= \sqrt{\langle F'(x^\dagger)^*w, F'(x^\dagger)^*w\rangle}\\ &= \sqrt{\langle F'(x^\dagger)F'(x^\dagger)^*w, w\rangle}\\ &= \sqrt{\langle F'(x^\dagger)(x-x^\dagger), w\rangle}\\ &\le \sqrt{\|F'(x^\dagger)(x-x^\dagger)\|}\,\sqrt{\|w\|}\\ &\le \sqrt{\tau}\,\sqrt{\|F'(x^\dagger)x - F'(x^\dagger)x^\dagger\|}\,.\end{aligned}$$

Hence, for a sequence $x_n \in D(F)$ which satisfies $\|F'(x^\dagger)x_n - F'(x^\dagger)x^\dagger\| \to 0$ we have also $\|x_n - x^\dagger\| \to 0$, and therefore the constrained version of (1.3) is locally well–posed at $x^\dagger$ □

Theorem 3.4 *Let $\tau < \frac{2}{L}$ for the constant τ in (3.1) with L from Assumption 2.1. Then the nonlinear operator equation (1.1) with the source condition domain (3.2) is always locally well–posed at $x^\dagger$.*

Proof: From (2.11) we obtain the estimate $\|R\| \le \frac{L\tau}{2}\|F'(x^\dagger)(x-x^\dagger)\|$ and the η–condition holds with $\eta := \frac{L\tau}{2} < 1$. Thus (2.5) holds for all $x \in D(F)$ ($D(F)$ as introduced in (3.2)). This is an estimation of the form (3.4) with $\tilde{F}(x) := F'(x^\dagger)x$. From the Lemmas 3.2 and 3.3 it follows the local well–posedness of (1.1) restricted to the source condition domain (3.2). □

The ellipsoid $D(F)$ forming the source condition domain (3.2) for sufficiently small τ excludes all ill–posed situations. Such an ellipsoid is given by the range of the adjoint of the Fréchet derivative $F'(x^\dagger)$. The more $F'(x^\dagger)$ is smoothing, the 'smaller' the range $R(F'(x^\dagger)^*)$ becomes. This yields a further argument for the fact that in all situations properties of the Fréchet derivative $F'(x^\dagger)$ are able to characterize the local degree of ill–posedness of (1.1) at $x^\dagger$. For another argument concerning convergence rates of Tikhonov regularized solutions we refer to [11].

In the literature on regularization methods source conditions of the form

$$x - x^\dagger = \left(F'(x^\dagger)^*F'(x^\dagger)\right)^{\frac{p}{2}} v \qquad (v \in H_1,\ \|v\| \le \nu)\,, \tag{3.7}$$

are frequently used. Here ν is a sufficiently small positive constant.

In the next part of this section we assume that the following assumptions hold:

Assumption 3.5 *Let*

$$D(F) := \{x \in H_1 :\ x \textit{ satisfies } (3.7)\}\,. \tag{3.8}$$

The operator F is Fréchet differentiable on $D(F)$ and satisfies (2.2) for all $x \in D(F)$.

We try to extend the stability results of Lemma 3.3 and Theorem 3.4 to this alternative form of source conditions. For the case $p \geq 1$ such an extension is given below.

Lemma 3.6 *For $D(F)$ from (3.8) with $p \geq 1$ the constrained version of the linearized equation (1.3), where*

$$F'(x^\dagger) : D(F) \subset H_1 \to H_2, \tag{3.9}$$

is always locally well–posed at $x^\dagger$. Namely, we have the stability estimate

$$\|x - x^\dagger\| \leq \sqrt{\nu}\, \|F'(x^\dagger)\|^{\frac{p-1}{2}} \sqrt{\|F'(x^\dagger)\, x - F'(x^\dagger)\, x^\dagger\|} \qquad \textit{for all} \quad x \in D(F). \tag{3.10}$$

Proof: By the source condition (3.7) we can write for all $x \in D(F)$ and fixed $p \geq 1$

$$\begin{aligned}
\|x - x^\dagger\| &= \sqrt{\left\langle (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}}\, v, (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}}\, v \right\rangle} \\
&= \sqrt{\left\langle (F'(x^\dagger)^* F'(x^\dagger))^{\frac{1}{2}}\, (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}}\, v, (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p-1}{2}}\, v \right\rangle} \\
&\leq \sqrt{\|F'(x^\dagger)\, (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}}\, v\|}\, \sqrt{\|v\|}\, \|F'(x^\dagger)\|^{\frac{p-1}{2}} \\
&\leq \sqrt{\nu}\, \|F'(x^\dagger)\|^{\frac{p-1}{2}} \sqrt{\|F'(x^\dagger)x - F'(x^\dagger)x^\dagger\|}\ .
\end{aligned}$$

Hence, for a sequence $x_n \in D(F)$, $\|F'(x^\dagger)x_n - F'(x^\dagger)x^\dagger\| \to 0$ implies $\|x_n - x^\dagger\| \to 0$, as $n \to \infty$, and the constrained version of (1.3) is locally well–posed at $x^\dagger$ over $D(F)$. $\square$

Theorem 3.7 *Let $p \geq 1$ and $\nu < \frac{2}{L\|F'(x^\dagger)\|^{p-1}}$, where p and ν are as in (3.7) and L is as in Assumption 2.1. Then the nonlinear operator equation (1.1) with the source condition domain (3.8) is always locally well–posed at $x^\dagger$.*

Proof: As a consequence of the interpolation inequality we have

$$\left\|(F'(x^\dagger)^* F'(x^\dagger))^{\frac{p-1}{2}}\right\| \leq \left\|(F'(x^\dagger)^* F'(x^\dagger))^{\frac{1}{2}}\right\|^{p-1} = \left\|F'(x^\dagger)\right\|^{p-1}\ .$$

Therefore, from (2.11) it follows

$$\begin{aligned}
\|R\| &\leq \tfrac{L}{2}\|x - x^\dagger\|^2 \\
&= \tfrac{L}{2}\left\|\left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{p}{2}} v\right\|^2 \\
&= \tfrac{L}{2}\left\langle \left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{1}{2}} \left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{p}{2}} v, \left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{p-1}{2}} v \right\rangle \\
&\leq \tfrac{L\|v\|\|F'(x^\dagger)\|^{p-1}}{2} \|F'(x^\dagger)\left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{p}{2}} v\| \\
&\leq \tfrac{L\nu\|F'(x^\dagger)\|^{p-1}}{2} \|F'(x^\dagger)(x - x^\dagger)\|\ .
\end{aligned}$$

Therefore F satisfies the η–condition for $x \in D(F)$ with $\eta := \frac{L\,\nu\,\|F'(x^\dagger)\|^{p-1}}{2} < 1$, and consequently satisfies (2.5), which then immediately shows the assertion. $\square$

The following theorems have been proven in [17] and allow to show local well–posedness of constraint problems under the source conditions with $p \leq 1$. For proving these results a very similar technique has been used as we use in the proof of Lemma 3.6 and in the proof of Theorem 3.7.

Theorem 3.8 *Let $x^\dagger$ be a solution of (1.1) and let $x_h \in D(F)$ fulfil the equation*

$$x^\dagger - x_h = F'(x^\dagger)^* w_h + r_h \tag{3.11}$$

for some $w_h \in Y$ and $r_h \in X$. If additionally the operator F is Fréchet differentiable and satisfies (2.3) with $(c_1, c_2, c_3) = (0, 0, 2)$ and

$$\left\{ \begin{array}{rcl} 2K\|w_h\| & \leq & C_2 < 1\,, \\ \|r_h\|^2 & \leq & \max\left\{C_1\|F(x_h) - F(x^\dagger)\|, \hat{C}_1\|x_h - x^\dagger\|^2\right\}\,, \\ \hat{C}_1 & < & 4K\|w_h\|(1 - 2K\|w_h\|)\,, \end{array} \right.$$

then

$$\|x_h - x^\dagger\|^2 = O\left(\|F(x_h) - F(x^\dagger)\|\right)\,.$$

Theorem 3.9 *Let $x^\dagger$ be a solution of (1.1). Moreover, let F be Fréchet differentiable with*

$$\|F'(x^\dagger)\| \leq 1$$

and satisfy (2.3) with $(c_1, c_2, c_3) = (1, 0, 0)$. We assume that $x_h \in D(F)$ fulfils for some $v_h \in X, r_h \in X$ and $0 < p < 1$ the equation

$$x^\dagger - x_h = \left(F'(x^\dagger)^* F'(x^\dagger)\right)^{\frac{p}{2}} v_h + r_h\,,$$

with

$$\max\{\|r_h\|, \|v_h\|\} \leq C_1$$

and

$$\|r_h\|^p \leq \max\left\{C_2\|x_h - x^\dagger\|, C_3\|F(x_h) - F(x^\dagger)\|^{(1-p)p}\right\}\,,$$

with

$$1 - 2C_1^{1-p}C_2 > 0\,.$$

Then it follows that

$$\|x_h - x^\dagger\| = O\left(\|F(x_h) - F(x^\dagger)\|^{(1-p)p}\right)\,.$$

Moreover, If $\|r_h\| \leq C_4\|F(x_h) - F(x^\dagger)\|$, then

$$\|x_h - x^\dagger\| = O\left(\|F(x_h) - F(x^\dagger)\|^{\frac{p}{1+p}}\right)\,.$$

4 The degree of nonlinearity and a posteriori estimates

A posteriori estimates are widely used in finite element methods (e.g. in two–level methods) for the solution of nonlinear partial differential equations.

In order to explain exemplarily the use of a posteriori estimates in finite element methods for the solution of nonlinear differential equations we first outline the idea of two–level methods as discussed e.g. in papers by Axelsson and Kaporin [1], Axelsson and Layton [2], Xu [19], to name but a few. Let F be a differentiable operator defined on a Hilbertian Sobolev space X, which maps into the dual space Y. Moreover, let X_H and X_h be two subsets of X satisfying $X_H \subseteq X_h$. A two–level method consists of two steps:

1. Find a finite element approximation $x_H \in X_H$ of the least squares problem, to minimize the functional
   $$\|F(x) - y\|^2 \tag{4.1}$$
   by solving the nonlinear equation
   $$< F(x_H) - y, \phi >_{Y,X} = 0 \text{ for all } \phi \in X_H \ . \tag{4.2}$$
2. An update $e_h \in X_h$ is calculated by solving the *linear* equation
   $$< F(x_H) - y - F'(x_H)(x_H + e_h), \phi >_{Y,X} = 0 \text{ for all } \phi \in X_h \ , \tag{4.3}$$
   where $x_H + e_h$ is taken as the approximation to the minimizer over all $x \in D(F)$ of the least squares functional $\|F(x) - y\|^2$.

The basic idea of two–level methods is that the nonlinear equation (4.2), which is a finite element approximation on the coarse grid X_H, and the linearized equation (4.3) together can be solved cheaper than the nonlinear equation on the finer grid X_h. The accuracy of the solution obtained with this two–level algorithm very much requires the optimal balance of the discretization parameters h and H. A reasonable choice of these parameters can be found by balancing reasonable estimates for $\|x_h - x^\dagger\|$ with some reasonable estimates for $\|(x_H + e_h) - x_h\|$. Here we do not specify the relevant norms, since several norms might be of interest. Possibly interesting norms are $\|\cdot\|_X$ and $\|\cdot\|_{L^2}$, depending on the practical application considered. In this setting x_h denotes the least squares finite element solution of the nonlinear equation

$$< F(x_h) - y, \phi >_{Y,X} = 0 \text{ for all } \phi \in X_h \ . \tag{4.4}$$

To be more concrete: Let $C(h)$ an upper bound for $\|x_h - x^\dagger\|$ and let $D(H,h)$ be an upper bound for $\|(x_H + e_h) - x_h\|$. Then it follows from triangle inequality that

$$\|(x_H + e_h) - x^\dagger\| \leq \|x_h - x^\dagger\| + \|(x_H + e_h) - x_h\| \leq C(h) + D(H,h) \ .$$

In order to find a quasi–optimal discretization parameter h the right hand side $C(h) + D(H,h)$ is minimized with respect to h (for given H). In partial differential equations estimates for $C(h)$ can be derived from standard estimates for finite element approximations. Estimates for $D(H,h)$ can be derived from a posteriori estimates which usually are of the form

$$\|(x_H + e_h) - x_h\| \leq f\left(\|F(x_H) - y\|\right) , \tag{4.5}$$

where $f : [0, \infty[\to \mathbb{R}$ is an increasing, continuous, and positive function. In practice f is known and $\|F(x_H) - y\|_Y$ is computationally available, and consequently estimates for $\|(x_H + e_h) - x_h\|$ can be derived.

Motivated from the above discussion we study in this paper a posteriori estimates of general type

$$\|x - z\| \leq f(\|F(x) - F(z)\|) \text{ for } x \in U_{\text{a-p}}(z) . \tag{4.6}$$

We do not restrict ourselves to the assumption that x and z are finite element or finite difference approximations. As one realizes from the discussion in the previous sections, there exist sets $U_{\text{a-p}}(z)$ on which a posteriori estimates for ill–posed problems hold.

In the following we derive convergence rates results based on a posteriori estimates for iterative regularization techniques like Landweber's method and the iteratively regularized Gauß–Newton method. Convergence rates results based on a posteriori estimates for the Tikhonov regularization have been derived in [17].

4.1 Landweber iteration

In [10] we have generalized Landweber's well–known iterative scheme for the regularized solution of linear ill–posed problems to cope with nonlinear applications. Let F be Fréchet differentiable and (1.1) be scaled such that locally in a neighborhood of $x^\dagger$

$$\|F'(x)\| \leq 1 \tag{4.7}$$

holds. Then the nonlinear Landweber iteration, introduced in [10], is defined as follows:

$$x_{k+1} = x_k - F'(x_k)^*(F(x_k) - y) , \quad k = 0, 1, 2, ... \tag{4.8}$$

where the initial guess x_0 may be chosen according to a given a priori knowledge on a solution $x^\dagger$ of (1.1). If F is nonlinear of degree $(1, 0, 0)$ in a neighborhood of $x^\dagger$ (see (2.3)) with $K < \frac{1}{2}$ and $x_0 \in B_r(x^\dagger)$, then the iterates x_k converge to $x^\dagger$. Given arbitrarily small $\delta > 0$, let y^δ satisfy

$$\|y^\delta - y\| \leq \delta \tag{4.9}$$

and denote by x_k^δ the iterates of (4.8) when y is replaced by y^δ. Then, due to the ill–posedness of the problem, the iteration will typically diverge. In the course of the iteration, however, some iterations will be close to a solution of (1.1). Therefore, in [10], a stopping rule has been proposed on the basis of a generalized discrepancy principle to obtain reasonable approximations of $x^\dagger$. When the stopping index is determined in this way, Landweber iteration is a regularization method in the sense of Tikhonov and Arsenin [18]. To be concrete let F satisfy (2.4) with $\eta < \frac{1}{2}$ and let the Landweber iteration be terminated by the stopping criterion

$$\|F(x_{k_*}) - y^\delta\| \leq \tau\delta < \|F(x_k) - y^\delta\| \qquad \text{for } k = 0, .., k_* - 1 \tag{4.10}$$

where

$$\tau > 2\frac{1+\eta}{1-2\eta} > 2\,. \tag{4.11}$$

That means, the approximation is determined by stopping the Landweber iteration when for the first time the residual is about the magnitude of the data error. The termination index will be denoted by $k_* = k_*(\delta)$.

In the following we prove convergence rates results based on a posteriori estimates. Using the notation

$$A_x = F'(x), \quad A = F'(x^\dagger)\,, \tag{4.12}$$

we require that the operator F satisfies for any $x \in B_r(x^\dagger)$

$$A_x = R_x A\,, \tag{4.13}$$

where $R_x : Y \to Y$ is a bounded linear operator which satisfies

$$\|R_x - I\| \leq \max\{\hat{C}_0\|x - x^\dagger\|, \tilde{C}_0\} < 1\,. \tag{4.14}$$

From (4.13) and (4.14) it follows that (cf. [10]) that for $x \in B_r(x^\dagger)$

$$\begin{aligned} \|F(x) - F(x^\dagger) - F'(x)(x - x^\dagger)\| &\leq \tfrac{\hat{C}_0}{2}\|x - x^\dagger\|\|F(x) - F(x^\dagger)\| \\ &\leq \hat{C}_1\|F(x) - F(x^\dagger)\|\,, \end{aligned} \tag{4.15}$$

with $\hat{C}_1 < 1$ a generic constant. Therefore F is nonlinear of degree $(0, 0, 1)$ at the point $x^\dagger$.

We note that from the assumption

$$\hat{C}_1 < 1 \tag{4.16}$$

the operators R_x are invertible and satisfy

$$\|R_x^{-1}\| \leq \hat{C}_2\,. \tag{4.17}$$

The following lemma will be central for our further considerations:

Lemma 4.1 *Let $x_0 - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}} v_0$, for $0 < p \le 1$ and let F satisfy (4.13), (4.14), (4.7), and Assumption 2.1. Then for any $0 \le k \le k_*(\delta), k \in \mathbb{N}$,*

$$x_k - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^{\frac{p}{2}} v_k \,, \tag{4.18}$$

with $\|v_k\|$ uniformly bounded.

Proof: We sketch the proof; the missing details can be found in Theorem 3.2 in [10]. All along this proof C will denote a generic constant. Following the lines in Theorem 3.2 in [10] we see that with the notation $A = F'(x^\dagger)$

$$x_k - x^\dagger = (A^*A)^{\frac{p}{2}} \left(f_k + \sum_{j=0}^{k-1} (I - A^*A)^j (A^*A)^{\frac{1-p}{2}} (z - z^\delta) \right) \,,$$

where

$$z - z^\delta = (A^*A)^{-\frac{1}{2}} A^* (y - y^\delta) \,.$$

Note that

$$\|z - z^\delta\| = \|y - y^\delta\| \,.$$

From the proof of Theorem 3.2 in [10] it immediately follows that

$$\|f_k\| \le C \|v_0\| \,.$$

Moreover,

$$\begin{aligned}
\left\| \sum_{j=0}^{k-1} (I - A^*A)^j (A^*A)^{\frac{1-p}{2}} \right\| &\le \sum_{j=0}^{k-1} \left\| (I - A^*A)^j (A^*A)^{\frac{1-p}{2}} \right\| \\
&\le \sum_{j=0}^{k-1} (j+1)^{\frac{p-1}{2}} \\
&\le C k^{\frac{1+p}{2}} \,.
\end{aligned}$$

The last inequality in the above chain can easily be proven by noting that $\sum_{j=0}^{k-1} (j+1)^{\frac{p-1}{2}}$ is a Riemann sum of the integral $\int_0^{k-1} (x+1)^{\frac{p-1}{2}}\, dx$.

From the proof of Theorem 3.2 in [10] it follows that for any $0 \le k \le k_*$

$$k^{\frac{1+p}{2}} \le \frac{C \|v_0\|}{\delta} \,,$$

which shows that

$$\left\| \sum_{j=0}^{k-1} (I - A^*A)^j (A^*A)^{\frac{1-p}{2}} (z - z^\delta) \right\| \le C \|v_0\| \,.$$

Consequently

$$x_k - x^\dagger = (A^*A)^{\frac{p}{2}} v_k \,,$$

with

$$v_k = f_k + \sum_{j=0}^{k-1} (I - A^*A)^j (A^*A)^{\frac{1-p}{2}} (z - z^\delta)\,,$$

and

$$\|v_k\| \le C\|v_0\|\,.$$

□

Corollary 4.2 *Under the assumptions of Lemma 4.1 let $0 < p \le 1$ and $\|v_0\|$ be sufficiently small. Moreover, let x_{k_*} satisfy*

$$\|F(x_{k_*}) - y^\delta\| \le \tau\delta\,,$$

then

$$\|x_{k_*} - x^\dagger\| \le C\delta^{\frac{p}{1+p}}\,,$$

with a generic constant C.

Proof: From Lemma 4.1 it follows that x_k satisfies the assumptions of Theorem 3.8 and Theorem 3.9, which then immediately yields the assertion. □

If instead of the assumption $0 < p \le 1$ the assumption $p > 1$ holds, then we obtain from Lemma 3.6

$$\|x_{k_*} - x^\dagger\| \le C\sqrt{\delta}\,.$$

This once again shows the saturation property of the discrepancy principle (see Morozov [16])!

4.2 An iteratively regularized Gauß–Newton technique

The iteratively regularized Gauß–Newton technique (first considered by Bakushinskii [5]) is defined as follows

$$x_{k+1} = x_k - (\alpha_k I + F'(x_k)^* F'(x_k))^{-1} \left(F'(x_k)^*(F(x_k) - y^\delta) + \alpha_k(x_k - x_0) \right)\,,$$

where x_0 is an initial guess which may incorporate a priori knowledge of an exact solution $x^\dagger$ and α_k is a sequence of positive parameters. In [4] we have proved convergence rates results under the source condition

$$x_0 - x^\dagger = \left(F'(x^\dagger)^* F'(x^\dagger) \right)^{\frac{p}{2}} v_0\,, \quad 0 < p \le 1\,. \tag{4.19}$$

In this section we prove analogous convergence results as in [4]. The technique of a posteriori estimates helps to simplify the proof of convergence rates of the iteratively regularized Gauß–Newton technique.

Theorem 4.3 *Let the operator F satisfy (4.13), (4.14) for all $x \in B_r(x^\dagger)$. Moreover, let (4.19) hold, with $\|v_0\|$ sufficiently small. Let the parameters in the iteratively regularized Gauß–Newton technique satisfy*

$$\alpha_k > 0, \quad 1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq \rho \quad \lim_{k\to\infty} \alpha_k = 0 \,. \tag{4.20}$$

Then for any $k \leq k_*(\delta)$ which satisfy (4.10) with τ sufficiently large we have

$$\|x_k - x^\dagger\| \leq C \|F(x_k) - y\|^{\frac{p}{1+p}} \,.$$

Proof: All along this section we assume that $k \leq k_*(\delta)$, and we use the notation $A = F'(x^\dagger)$, $A_k = F'(x_k)$.

From the definition of the iteratively regularized Gauß–Newton technique it follows that

$$x_{k+1} - x^\dagger = -\alpha_k(\alpha_k I + A^*A)^{-1}(A^*A)^{\frac{p}{2}} v_0 + r_k \,, \tag{4.21}$$

with

$$r_k = r_k^1 + r_k^2 + r_k^3$$

and

$$\begin{aligned}
r_k^1 &= -\alpha_k(\alpha_k I + A_k^* A_k)^{-1} A_k^* (A_k - A)(\alpha_k I + A^*A)^{-1}(A^*A)^{\frac{p}{2}} v_0 \\
r_k^2 &= -\alpha_k(\alpha_k I + A_k^* A_k)^{-1} (A_k^* - A^*) A (\alpha_k I + A^*A)^{-1}(A^*A)^{\frac{p}{2}} v_0 \\
r_k^3 &= -\alpha_k(\alpha_k I + A_k^* A_k)^{-1} A_k^* (F(x_k) - y - A_k(x_k - x^\dagger)) \,.
\end{aligned}$$

In the sequel we will use the following notation

$$v_k = -\alpha_k(\alpha_k + A^*A)^{-1} v_0$$

and Now, we estimate $\|r_k\|$: It follows from standard spectral estimates that

$$\begin{aligned}
\|r_k^1\| &\leq \tfrac{C}{\sqrt{\alpha_k}} \|x_k - x^\dagger\| \|(A^*A)^{\frac{p+1}{2}} v_k\| \\
\|r_k^2\| &\leq \tfrac{C}{\sqrt{\alpha_k}} \|x_k - x^\dagger\| \|(A^*A)^{\frac{p+1}{2}} v_k\| \\
\|r_k^3\| &\leq \tfrac{C}{\sqrt{\alpha_k}} \|x_k - x^\dagger\| \|A(x_k - x^\dagger)\| \,,
\end{aligned}$$

where C is a generic constant. Following the lines of the proof of Lemma 2.2 in [4] it can be seen that

$$\left\{ \begin{aligned}
\|A(x_{k+1} - x^\dagger)\| &\leq C \|(A^*A)^{\frac{p+1}{2}} v_k\| \\
\\
\|x_{k+1} - x^\dagger\| &\leq \|(A^*A)^{\frac{p}{2}} v_k\| \\
&\quad + \tfrac{C}{\sqrt{\alpha_k}} \|(A^*A)^{\frac{p+1}{2}} v_k\| \|x^\dagger - x_k\| \\
&\quad + \tfrac{C}{\sqrt{\alpha_k}} \|A(x_k - x^\dagger)\| \|x_k - x^\dagger\| \\
&\leq \|(A^*A)^{\frac{p}{2}} v_k\| + \tfrac{C}{\sqrt{\alpha_k}} \|(A^*A)^{\frac{p+1}{2}} v_k\| \|x^\dagger - x_k\| \,,
\end{aligned} \right. \tag{4.22}$$

where C always denotes a generic constant. Now, we follow the lines of the proof of Lemma 2.2 in [4], by taking into account that the iteratively regularized Gauß–Newton method is terminated by the a posteriori stopping (4.10) with τ sufficiently large and noting that under the assumptions of this theorem $\|F(x)-y\| \leq C\|A(x-x^\dagger)\|$. Since $\|(A^*A)^{\frac{p+1}{2}}v_k\|$ is small if $\|v_0\|$ (introduced in (4.19)) is small, we can deduce with an inductive argument that

$$\|x_{k+1}-x^\dagger\| \leq C\|(A^*A)^{\frac{p}{2}}v_k\| \,. \tag{4.23}$$

Therefore, it follows that $\|x_{k+1}-x^\dagger\| \leq r$ as long as $\|v_0\|$ is sufficiently small.

From (4.12), (4.13) it follows that (using the representation $R_k^{-1}A_k = A$)

$$\begin{aligned}\|Ar_k^1\| &\leq \alpha_k\|R_k^{-1}\|\|A_k(\alpha_k I + A_k^*A_k)^{-1}A_k^*\|\|R_k - I\|\|A(\alpha_k I + A^*A)^{-1}(A^*A)^{\frac{p}{2}}v\|\\ &\leq C\|x_k - x^\dagger\|\|(A^*A)^{\frac{1+p}{2}}v_k\| \,.\end{aligned}$$

where here again C is a generic constant. Analogously, we obtain estimates

$$\|Ar_k^2\| \leq C\|x_k-x^\dagger\|\|(A^*A)^{\frac{1+p}{2}}v_k\| \,,$$

and

$$\|Ar_k^3\| \leq C\|x_k-x^\dagger\|\|F(x_k)-y\| \,.$$

Moreover, from (4.22) and (4.20) it follows that

$$\frac{\|A(x_{k+1}-x^\dagger)\|}{\sqrt{\alpha_{k+1}}} \leq C\rho\frac{\|(A^*A)^{\frac{1+p}{2}}v_k\|}{\sqrt{\alpha_k}} \,.$$

Since $\frac{\left\|(A^*A)^{\frac{1+p}{2}}v_k\right\|}{\sqrt{\alpha_k}}$ is bounded (and small if $\frac{\left\|(A^*A)^{\frac{1+p}{2}}v_0\right\|}{\sqrt{\alpha_0}}$ is sufficiently small) we can deduce again by an inductive argument that

$$\frac{\|A(x_{k+1}-x^\dagger)\|}{\sqrt{\alpha_{k+1}}} \leq C \,, \tag{4.24}$$

where C is again a generic constant. Therefore we have

$$\|Ar_k\| \leq C\|x_k-x^\dagger\|\|(A^*A)^{\frac{1+p}{2}}v_k\| \,, \tag{4.25}$$

and thus it follows from (4.25)

$$\begin{aligned}\|(A^*A)^{\frac{1+p}{2}}v_k\| &\leq \|(A^*A)^{\frac{1}{2}}(x_{k+1}-x^\dagger)\| + \|(A^*A)^{\frac{1}{2}}r_k\|\\ &\leq \|(A^*A)^{\frac{1}{2}}(x_{k+1}-x^\dagger)\| + C\|x_k-x^\dagger\|\|(A^*A)^{\frac{p+1}{2}}v_k\| \,.\end{aligned} \tag{4.26}$$

If $\|x_0 - x^\dagger\|$ is small, then we obtain from (4.23) that $\|x_k - x^\dagger\|$ is small, and consequently, we can assume without loss of generality that $C\|x_k - x^\dagger\| < 1$, which shows that

$$\|(A^*A)^{\frac{1+p}{2}} v_k\| \leq C\|F(x_{k+1}) - F(x^\dagger)\| \ .$$

From interpolation inequality it follows that

$$\begin{aligned} \|(A^*A)^{\frac{p}{2}} v_k\| &\leq \|(A^*A)^{\frac{1+p}{2}} v_k\|^{\frac{p}{1+p}} \|v_k\|^{\frac{1}{1+p}} \\ &\leq C\|F(x_{k+1}) - F(x^\dagger)\|^{\frac{p}{1+p}} \ . \end{aligned}$$

Therefore, from (4.22) we have

$$\|x_{k+1} - x^\dagger\| \leq C\|F(x_{k+1}) - y\|^{\frac{p}{1+p}} \ .$$

□

In this section we have shown that the iterates of the Landweber's method and iterates of the iteratively regularized Gauß–Newton method satisfy a posteriori estimates. These estimates were derived by proving that the iterates project onto a subset, which is characterized by the degree of nonlinearity and source conditions.

References

[1] Axelsson, O.;Kaporin I.E., Least squares finite element methods for systems of nonlinear partial differential equations of first order: a posteriori error estimates. *Preprint* (1996), submitted.

[2] Axelsson, O.; Layton B.:, A two–level method for the discretization of nonlinear boundary value problems. *SIAM J. Numer. Anal.* **17** (1996), 2359–2374.

[3] Bakushinsky, A.B.; Goncharsky, A.V.: *Ill–Posed Problems: Theory and Applications.* Kluwer Academic Publishers, Dordrecht, 1994.

[4] Blaschke, B.; Neubauer, A.; Scherzer O.: On convergence rates for the iteratively regularized Gauß-Newton method. *IMA J. Numer. Anal.* **17** (1997), 421-436.

[5] Bakushinskii, A.B.: The problem of the convergence of the iteratively regularized Gauß-Newton method *Comput. Maths. Math. Phys.* **32** (1992), 1353-1359.

[6] Binder, A.; Hanke, M.; Scherzer, O.: On the Landweber iteration for nonlinear ill–posed problems. *J. Inv. ill–posed Problems* **4** (1996), 381-389.

[7] Engl, H.W.; Kunisch, K.; Neubauer, A.: Convergence rates for Tikhonov regularization of non-linear ill–posed problems. *Inverse Problems* **5** (1989), 523-540.

[8] GORENFLO, R.; HOFMANN, B.: On autoconvolution and regularization. *Inverse Problems* **10** (1994), 353-373.

[9] GROETSCH, C.W.: *Inverse Problems in the Mathematical Sciences.* Vieweg, Braunschweig/Wiesbaden 1993.

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

[11] HOFMANN, B.: On the degree of ill–posedness for nonlinear problems. *Journal of Inverse and ill–posed Problems* **2** (1994), 61-76.

[12] HOFMANN, B.: On ill–posedness and local ill–posedness of operator equations in Hilbert spaces. Preprint 97-8, Techn. Univ. Chemnitz-Zwickau, Faculty of Mathematics, Chemnitz 1997.

[13] HOFMANN, B.; SCHERZER, O.: Factors influencing the ill–posedness of nonlinear problems. *Inverse Problems* **10** (1994), 1277-1297.

[14] HOFMANN, B.; TAUTENHAHN, U.: On ill–posedness measures and space change in Sobolev scales. *Z. Anal. Anw.* **16** (1997), 979-1000.

[15] NASHED, M.Z.: *Generalized Inverses and Applications.* Academic Press, New York, 1976.

[16] MOROZOV, V.A.: *Methods for Solving Incorrectly Posed Problems.* Springer Verlag, New York, Berlin, Heidelberg, 1984.

[17] SCHERZER O.: A posteriori estimates for nonlinear (ill–posed) operator equations, preprint.

[18] TIKHONOV, A. N.; ARSENIN, V. Y.: *Solutions of ill–posed problems*, John Wiley & Sons, Washington, D.C., 1977, Translation editor Frity John (English)

[19] XU, J.: A novel two–grid method for semilinear elliptic problems. SIAM J. SCI. COMPUT. **15** (1994) 231–237.