On the monotone error rule for choosing the regularization parameter in ill-posed problems

U. HÄMARIK* and U. TAUTENHAHN†

Published in: Inverse and Ill-Posed Problems Series: Ill-Posed and Non-Classical Problems of Math. Physics and Analysis, Proc. of the Int. Conf. Samarkand, Uzbekistan, eds. M. M. Lavrentiev et al. (Utrecht-Boston: VSP, 2003), pp. 27 – 55.

Abstract — We consider linear ill-posed problems $Ax = y$ in Hilbert spaces with minimum-norm solution $x^\dagger$ and suppose that instead of y noisy data y^δ are given satisfying $\|y - y^\delta\| \le \delta$ with known noise level δ. For the stable numerical solution regularization methods are considered including continuous regularization methods such as ordinary Tikhonov regularization $x_r^\delta = (A^*A + r^{-1}I)^{-1}A^*y^\delta$ and iterative regularization methods. For the proper choice of the regularization parameter r (which is the stopping index in iterative methods) we study the *monotone error rule* (ME rule): Choose $r = r_{ME}$ as the largest r-value for which it can be guaranteed that the error $\|x_r^\delta - x^\dagger\|$ is monotonically decreasing for $r \in (0, r_{ME}]$. We compare this rule with other *a posteriori* rules and give conditions for which convergence and order optimal convergence rate results can be guaranteed. For the computation of r_{ME} in Tikhonov methods some nonlinear equation has to be solved. Newton's iteration for this equation appears to be globally and monotonically convergent. Numerical experiments are provided that verify some of the theoretical results.

1. INTRODUCTION

In this paper we consider linear ill-posed problems

$$Ax = y \tag{1.1}$$

where $A \in \mathcal{L}(X, Y)$ is a bounded operator with non-closed range $R(A)$ and X, Y are infinite dimensional real Hilbert spaces with inner products $(\cdot,\cdot)$ and norms $\|\cdot\|$, respectively. We are interested in the minimum-norm solution $x^\dagger$ of problem (1.1) and assume that instead of *exact* data y there are given *noisy* data $y^\delta \in Y$ with $\|y - y^\delta\| \le \delta$ and known noise level δ.

Ill-posed problems (1.1) arise in a wide variety of problems in applied sciences. For their stable numerical solution regularization methods are necessary, see [4, 19, 26, 39].

*University of Tartu, Department of Mathematics, Liivi 2, 50409 Tartu, Estonia. E-mail: Uno.Hamarik@math.ut.ee

†University of Applied Sciences Zittau/Görlitz, Department of Mathematics, P.O. Box 1455, 02755 Zittau, Germany. E-mail: u.tautenhahn@hs-zigr.de

Regularization methods can be divided into continuous regularization methods that include ordinary Tikhonov regularization $x_r^\delta := R_r y^\delta = (A^*A + r^{-1}I)^{-1}A^*y^\delta$, and iterative regularization methods where the stopping index plays the role of the regularization parameter. The element $x_r^\delta = R_r y^\delta$ is called regularized approximation for the minimum-norm solution $x^\dagger$ of problem (1.1) provided

1. for any $r > 0$, $R_r : Y \to X$ is continuous and
2. for arbitrary $y \in Y$ with $Qy \in R(A)$, $\lim_{r\to\infty} \|R_r y - x^\dagger\| = 0$, where Q is the projection operator onto $\overline{R(A)}$

(see [36]). Traditional regularization methods possess the property that in the case of exact data the error $\|x_r^0 - x^\dagger\|$ as a function of r is monotonically decreasing for $r \to \infty$. This property is no longer true for the error $\|x_r^\delta - x^\dagger\|$. In the case of noisy data the situation is as follows:

(i) If r becomes too large, then the error $\|x_r^\delta - x^\dagger\|$ increases due to the fact of ill-posedness of the operator equation (1.1).

(ii) If r becomes too small, then the error $\|x_r^\delta - x^\dagger\|$ increases due to the fact that the error $\|x_r^0 - x^\dagger\|$ is monotonically increasing for decreasing r-values.

The monotone decrease of the error $\|x_r^\delta - x^\dagger\|$ for growing r-values can only be guaranteed for small r. Typically $\|x_r^\delta - x^\dagger\|$ diverges for $r \to \infty$. Therefore a rule for the proper choice of the regularization parameter r is necessary.

In the monotone error rule for choosing a proper regularization parameter the idea consists in *searching for a largest computable regularization parameter* $r = r_{ME}$ *for which we can guarantee that the error* $\|x_r^\delta - x^\dagger\|$ *is monotonically decreasing for* $r \in (0, r_{ME}]$. For continuous regularization methods this means that

$$\frac{\mathrm{d}}{\mathrm{d}r}\|x_r^\delta - x^\dagger\|^2 \le 0 \quad \text{for all} \quad r \in (0, r_{ME}]\,, \tag{1.2}$$

for iteration methods this means that

$$\|x_r^\delta - x^\dagger\| \le \|x_{r-1}^\delta - x^\dagger\| \quad \text{for all} \quad r = 1, 2, \ldots, r_{ME}\,. \tag{1.3}$$

Similar rules for the choice of the regularization parameter which are based on monotonicity properties of the error were proposed and studied for some iterative methods in [1, 2, 13, 15, 16], for the method of ordinary Tikhonov regularization in [34] and for the method of iterated Tikhonov regularization and some other continuous regularization methods in [14, 15, 35].

In this paper we study the ME rule for continuous regularization methods including Tikhonov methods and asymptotical regularization, and for iterative regularization methods including gradient type methods (Landweber's method, steepest descent method, minimal error method) and implicit iteration methods. We compare the ME rule with other *a posteriori* rules including Morozov's discrepancy principle and study questions concerning convergence $x_{r_{ME}}^\delta \to x^\dagger$ for $\delta \to 0$ and concerning order optimal error bounds under certain source conditions. For the computation of r_{ME} in Tikhonov methods some nonlinear equation $\tilde{d}_{ME}(r) = \delta$ has to be solved. The function $\tilde{d}_{ME}$ appears to be strictly monotonically decreasing and strictly convex. These properties guarantee global and monotone convergence for Newton's iteration. In the final section numerical examples are provided which verify some of the theoretical results.

2. CONTINUOUS REGULARIZATION METHODS

2.1. Continuous regularization methods and the ME rule

In continuous regularization methods which include the method of ordinary Tikhonov regularization $x_r^\delta = (A^*A + r^{-1}I)^{-1}A^*y^\delta$ we use in this section for the regularization parameter the traditional notation $\alpha = 1/r$ instead of r. We consider continuous regularization methods of the general form

$$x_\alpha^\delta = g_\alpha(A^*A)A^*y^\delta\,. \tag{2.1}$$

Here $g_\alpha(\lambda) : [0, a] \to \mathbb{R}$ with $a = \|A\|^2$ is a family of piecewise continuous functions depending on a positive regularization parameter $\alpha > 0$ and the operator function g_α is defined according to

$$g_\alpha(A^*A) = \int_0^a g_\alpha(\lambda)\,\mathrm{d}E_\lambda$$

where $A^*A = \int_0^a \lambda\,\mathrm{d}E_\lambda$ is the spectral decomposition of the operator A^*A. For the functions $g_\alpha(\lambda)$ we assume as in [4, 26, 38, 39] that there exist constants γ, γ_p and p_0 such that for $\alpha > 0$

$$\sup_{0\le\lambda\le a} |g_\alpha(\lambda)| \le \gamma\alpha^{-1} \tag{2.2}$$

and

$$\sup_{0\le\lambda\le a} \lambda^p|1 - \lambda g_\alpha(\lambda)| \le \gamma_p\alpha^p \quad \text{for} \quad 0 \le p \le p_0\,. \tag{2.3}$$

The largest constant p_0 in assumption (2.3) is called *qualification* of the regularization method (2.1) (see [39]). Three well known *a posteriori* rules for choosing the regularization parameter α in continuous regularization methods (2.1) are:

1. *Morozov's discrepancy principle* [4, 27, 38, 39]. In this principle (*D principle*) the parameter $\alpha = \alpha_D$ is chosen as the solution of the equation

$$d_D(\alpha) := \|y^\delta - Ax_\alpha^\delta\| = C\delta \quad \text{with} \quad C \ge 1\,.$$

2. *Rule of Raus* [30]. In this rule, which we call *R rule*, the regularization parameter $\alpha = \alpha_R$ is chosen as the solution of the equation

$$d_R(\alpha) := \left\|(I - g_\alpha(AA^*)AA^*)^{1/(2p_0)}(y^\delta - Ax_\alpha^\delta)\right\| = C\delta \quad \text{with} \quad C \ge 1\,.$$

   Here p_0 is the (largest) constant from assumption (2.3).

3. *Rule of Engl and Gfrerer* [3]. In this rule, which we call *EG rule*, the regularization parameter $\alpha = \alpha_{EG}$ is chosen as the solution of the equation

$$d_{EG}(\alpha) := \gamma^{-1/2}\alpha\Big(Ax_\alpha^\delta - y^\delta, \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big)^{1/2} = C\delta \quad \text{with} \quad C \ge 1\,.$$

   Here γ is the (smallest) constant from assumption (2.2).

For ordinary and iterated Tikhonov methods the R– and EG rules coincide. The resulting rule was also proposed in [8] and we call this rule *Raus-Gfrerer rule* (*RG rule*).

Now let us turn over to the ME rule. The general idea of this rule (see Chapter 1) consists in searching for a largest computable $r = r_{ME}$ for which we can guarantee that the error $\|x_r^\delta - x^\dagger\|$ is monotonically decreasing for $r \in (0, r_{ME}]$. The reformulation of this idea in terms of $\alpha = 1/r$ means: *Search for the smallest computable regularization parameter* $\alpha = \alpha_{ME}$ *for which we can guarantee that*

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\|x_\alpha^\delta - x^\dagger\|^2 \geq 0 \quad \text{for all} \quad \alpha \in [\alpha_{ME}, \infty)\,.$$

In order to guarantee this property we use the identity $g_\alpha(A^*A)A^* = A^*g_\alpha(AA^*)$ and estimate the derivative of the squared error $\|x_\alpha^\delta - x^\dagger\|^2$ with respect to α as follows:

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\alpha}\|x_\alpha^\delta - x^\dagger\|^2 &= \Big(x_\alpha^\delta - x^\dagger, A^*\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big) \\
&= \Big(Ax_\alpha^\delta - y^\delta + (y^\delta - y), \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big) \\
&\geq \Big\|\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big\| \left\{\frac{\Big(Ax_\alpha^\delta - y^\delta, \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big)}{\Big\|\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big\|} - \delta\right\}.
\end{aligned}$$

This estimate leads us to following ME rule for continuous regularization method (2.1):

ME rule. For regularization methods with monotonically increasing functions $d_{ME}(\alpha)$ in (2.4), choose $\alpha = \alpha_{ME}$ as the solution of the equation

$$d_{ME}(\alpha) := \frac{\Big(Ax_\alpha^\delta - y^\delta, \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big)}{\Big\|\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big\|} = \delta\,. \tag{2.4}$$

2.2. The ME rule for ordinary and iterated Tikhonov regularization

In these methods we start with $x_{\alpha,0}^\delta = 0$ and compute the regularized solution $x_\alpha^\delta := x_{\alpha,m}^\delta$ recursively by solving the m operator equations

$$(A^*A + \alpha I)x_{\alpha,k}^\delta = A^*y^\delta + \alpha x_{\alpha,k-1}^\delta\,, \quad k = 1, 2, ..., m\,. \tag{2.5}$$

For $m = 1$ this method is the *method of ordinary Tikhonov regularization.* In these methods we have $g_\alpha(\lambda) = [1 - (1 + \lambda/\alpha)^{-m}]/\lambda$ with some fixed positive integer $m \geq 1$. Let $r_{\alpha,m}$ denote the discrepancy of the regularized solution $x_\alpha^\delta = x_{\alpha,m}^\delta$, i.e.

$$r_{\alpha,m} = y^\delta - Ax_{\alpha,m}^\delta\,.$$

Using the identities

$$1 - \lambda g_\alpha(\lambda) = \left(\frac{\alpha}{\lambda + \alpha}\right)^m \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(\lambda) = -\frac{m}{\alpha^2}\left(\frac{\alpha}{\lambda + \alpha}\right)^{m+1}$$

we obtain

$$r_{\alpha,m} = [I - AA^*g_\alpha(AA^*)]y^\delta = [\alpha(AA^* + \alpha I)^{-1}]^m y^\delta$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta = -\frac{m}{\alpha^2}r_{\alpha,m+1}\,.$$

From these representations we conclude that the functions $d_{RG}(\alpha)$ and $d_{ME}(\alpha)$ for the RG– and ME rules have the form

$$d_{RG}(\alpha) = (r_{\alpha,m}, r_{\alpha,m+1})^{1/2} \quad \text{and} \quad d_{ME}(\alpha) = \frac{(r_{\alpha,m}, r_{\alpha,m+1})}{\|r_{\alpha,m+1}\|}\,. \tag{2.6}$$

The function $d_{ME}(\alpha)$ given in (2.6) possesses following properties, see [34] for $m = 1$ and [14, 35] for $m \geq 1$:

Theorem 2.1. *Let P denote the orthoprojection of Y onto $N(A^*) = \overline{R(A)}^{\perp}$ and let $A^*y^\delta \neq 0$. Then:*

(i) *$d_{ME}(\alpha)$ is strictly monotonically increasing and obeys*

$$d_{ME}(0) = \|Py^\delta\| \quad \textit{and} \quad \lim_{\alpha\to\infty} d_{ME}(\alpha) = \|y^\delta\|\,.$$

The equation $d_{ME}(\alpha) = \delta$ has a unique solution α_{ME} provided $\|Py^\delta\| < \delta < \|y^\delta\|$.

(ii) *For all $\alpha \in (\alpha_{ME}, \infty)$ there holds $\frac{\mathrm{d}}{\mathrm{d}\alpha}\|x^\delta_{\alpha,m} - x^\dagger\|^2 > 0$.*

(iii) *There holds*

$$d_{RG}(\alpha) < d_{ME}(\alpha) < d_D(\alpha)\,.$$

If $C = 1$ in the D principle and in the RG rule, then $\alpha_D < \alpha_{ME} < \alpha_{RG}$.

From parts (ii) and (iii) of the theorem there follows

$$\|x^\delta_{\alpha_{ME}} - x^\dagger\| < \|x^\delta_{\alpha_{RG}} - x^\dagger\|\,. \tag{2.7}$$

Hence, the ME rule provides always a smaller error than the RG rule. Exploiting the monotonicity property (ii) we obtain for the parameter choice $\alpha = \alpha_{ME}$ order optimal error bounds (see [35]):

Theorem 2.2. *Assume $x^\dagger = (A^*A)^{p/2}w$ with $\|w\| \leq E$. Then for $p \in (0, 2m]$*

$$\|x^\delta_{\alpha_{ME}} - x^\dagger\| \leq \left\{2^{\frac{p}{p+1}} + 2^{\frac{-1}{p+1}}\gamma_*(\gamma_{p/2})^{\frac{1}{p}}\right\} E^{\frac{1}{p+1}}\, \delta^{\frac{p}{p+1}} \tag{2.8}$$

with $\gamma_* = \sqrt{m}$ and $\gamma_{p/2} = [p/(2m)]^{p/2}[1 - p/(2m)]^{m-p/2} \leq 1$.

2.3. The ME rule for asymptotical regularization

In this regularization method the regularized solution is given by $x^\delta_\alpha = x^\delta(r)$ where $x^\delta(r)$ is the solution of the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}x^\delta(t) + A^*Ax^\delta(t) = A^*y^\delta \quad \text{for} \quad 0 < t \leq r\,, \quad x^\delta(0) = 0$$

with $r = 1/\alpha$. In this method we have $g_\alpha(\lambda) = (1 - e^{-\lambda/\alpha})/\lambda$. Using the identities

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(\lambda) = -\frac{1}{\alpha^2}e^{-\lambda/\alpha} = \frac{1}{\alpha^2}(1 - \lambda g_\alpha(\lambda))$$

we obtain that for the method of asymptotical regularization there holds

$$d_R(\alpha) = d_{EG}(\alpha) = d_{ME}(\alpha) = d_D(\alpha)\,.$$

From the identity $d_D(\alpha) = d_{ME}(\alpha)$ and the monotonicity property of our ME rule we conclude that for the method of asymptotical regularization the best constant C in the D principle is $C = 1$:

Theorem 2.3. *Let $A^*y^\delta \neq 0$, let x_α^δ the regularized solution obtained by the method of asymptotical regularization and let α_D the regularization parameter of the D principle with $C = 1$. Then*

$$\|x_{\alpha_D}^\delta - x^\dagger\| < \|x_\alpha^\delta - x^\dagger\| \quad \textit{for all} \quad \alpha > \alpha_D\,.$$

From the identity $d_R(\alpha) = d_{EG}(\alpha) = d_{ME}(\alpha) = d_D(\alpha)$ we conclude that all results known for the D principle (see, e.g., [4, 38, 39]) are also true for the R rule, the EG rule and the ME rule, respectively. Exploiting the monotonicity property of our ME rule we obtain order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$ with α chosen by the ME rule, or equivalently, the R rule, the EG rule or the D principle, respectively.

Theorem 2.4. *Assume $x^\dagger = (A^*A)^{p/2}w$ with $\|w\| \leq E$. Let x_α^δ the regularized solution obtained by the method of asymptotical regularization and let α be chosen by the ME rule. Then for all $p \in (0,\infty)$ the order optimal error estimate (2.8) holds true with $\gamma_* \approx 0.6382$ and $\gamma_{p/2} = (p/(2e))^p$.*

Proof. Let us introduce the two operators

$$K_\alpha = I - A^*Ag_\alpha(A^*A)\,, \quad \tilde{K}_\alpha = I - AA^*g_\alpha(AA^*)\,.$$

Then we obtain from (2.1)

$$x_\alpha^\delta - x^\dagger = -K_\alpha x^\dagger + g_\alpha(A^*A)A^*(y^\delta - Ax^\dagger)\,, \tag{2.9}$$

$$Ax_\alpha^\delta - y^\delta = -\tilde{K}_\alpha Ax^\dagger + \tilde{K}_\alpha(Ax^\dagger - y^\delta)\,. \tag{2.10}$$

The method of asymptotical regularization is characterized by $g_\alpha(\lambda) = (1 - e^{-\lambda/\alpha})/\lambda$. For this function there holds for arbitrary $\alpha > 0$ the estimate

$$\sup_{0\leq\lambda\leq\alpha} \sqrt{\lambda}|g_\alpha(\lambda)| \leq \gamma_*/\sqrt{\alpha} \tag{2.11}$$

with $\gamma_* \approx 0.6382$ (see [38, 39]). From (2.9) and (2.11) we obtain

$$\|x_\alpha^\delta - x^\dagger\| \leq \|K_\alpha x^\dagger\| + \gamma_*\delta/\sqrt{\alpha}\,. \tag{2.12}$$

Now we exploit the assumption $x^\dagger = (A^*A)^{p/2}w$ with $\|w\| \leq E$ and obtain from (2.3), which holds with $p_0 = \infty$ and $\gamma_p = (p/e)^{2p}$, the estimate

$$\|K_\alpha x^\dagger\| \leq \gamma_{p/2}\alpha^{p/2}E\,. \tag{2.13}$$

For arbitrary $D = D^* \geq 0$ and $0 \leq s \leq t$ there holds the moment inequality

$$\|D^s v\| \leq \|D^t v\|^{s/t}\|v\|^{1-s/t} \tag{2.14}$$

(see, e.g., [22]). We apply this inequality with $D = (A^*A)^{1/2}$, $v = K_\alpha w$, $s = p$, $t = p+1$ and obtain due to the relations $D^{p+1}K_\alpha w = \tilde{K}_\alpha Ax^\dagger$ and $\|K_\alpha\| \leq 1$ that

$$\begin{aligned} \|K_\alpha x^\dagger\| &= \|D^p K_\alpha w\| \\ &\leq \|D^{p+1}K_\alpha w\|^{p/(p+1)}\|K_\alpha w\|^{1/(p+1)} \\ &\leq \|\tilde{K}_\alpha Ax^\dagger\|^{p/(p+1)}E^{1/(p+1)}\,. \end{aligned} \tag{2.15}$$

For $\alpha = \alpha_{ME} = \alpha_D$ we obtain due to the relations $\|Ax^\delta_{\alpha_{ME}} - y^\delta\| = \delta$, $\|Ax^\dagger - y^\delta\| \le \delta$, $\|\tilde{K}_\alpha\| \le 1$ and (2.10) that

$$\|\tilde{K}_{\alpha_{ME}} Ax^\dagger\| \le \|Ax^\delta_{\alpha_{ME}} - y^\delta\| + \|\tilde{K}_{\alpha_{ME}}(Ax^\dagger - y^\delta)\| \le 2\delta\,.$$

Consequently, (2.15) attains for $\alpha = \alpha_{ME}$ the form

$$\|K_{\alpha_{ME}} x^\dagger\| \le (2\delta)^{p/(p+1)} E^{1/(p+1)}\,. \tag{2.16}$$

Let $\alpha = \alpha_*$ be chosen such that the right hand sides of the estimates (2.13) and (2.16) coincide, i.e., let $\alpha_* = (\gamma_{p/2})^{-2/p} E^{-2/(p+1)} (2\delta)^{2/(p+1)}$. If $\alpha_{ME} \le \alpha_*$, then the desired estimate (2.8) follows from the monotonicity property $\|x_{\alpha_{ME}} - x^\dagger\| \le \|x_{\alpha_*} - x^\dagger\|$ (see Theorem 2.3) and from the estimates (2.12), (2.13) with $\alpha = \alpha_*$. If $\alpha_{ME} > \alpha_*$, then the desired estimate (2.8) follows from (2.12) with $\alpha = \alpha_{ME}$ and from the estimates (2.16) and $\gamma_*\delta/\sqrt{\alpha_{ME}} < \gamma_*\delta/\sqrt{\alpha_*}$. Hence, the proof is complete. $\square$

2.4. Newton's iteration for the ME rule in Tikhonov methods

Let us consider the methods of ordinary and iterated Tikhonov regularization (2.5). For choosing the regularization parameter α according to the D principle, the RG rule and the ME rule, respectively, the following nonlinear equations

$$d_D(\alpha) = C\delta\,, \quad d_{RG}(\alpha) = C\delta \quad \text{and} \quad d_{ME}(\alpha) = \delta$$

have to be solved numerically. For the iterative solution of these nonlinear equations the change of the variable α by $r = 1/\alpha$ is reasonable since the functions

$$\tilde{d}_D(r) := d_D(1/r)\,, \quad \tilde{d}_{RG}(r) := d_{RG}(1/r)\,, \quad \tilde{d}_{ME}(r) := d_{ME}(1/r)$$

are *monotonically decreasing and convex* for all $r > 0$. These two properties guarantee global monotone convergence, e.g., for Newton's iteration. For the function $\tilde{d}_{ME}(r)$ we prove the properties of monotonicity and convexity in Theorem 2.5. In order to prove these properties for the functions $\tilde{d}_D(r)$ and $\tilde{d}_{RG}(r)$ we use the representations

$$\tilde{d}_D(r) = \|R_r^m y^\delta\|\,, \quad \tilde{d}_{RG}(r) = \|R_r^{m+1/2} y^\delta\|$$

with $R_r = (I + rAA^*)^{-1}$, exploit the two identities

$$\frac{\mathrm{d}}{\mathrm{d}r}\|R_r^k y^\delta\|^2 = -2k\|A^* R_r^{k+1/2} y^\delta\|^2\,, \tag{2.17}$$

$$\frac{\mathrm{d}}{\mathrm{d}r}\|A^* R_r^{k+1/2} y^\delta\|^2 = -(2k+1)\|AA^* R_r^{k+1} y^\delta\|^2 \tag{2.18}$$

that hold true for arbitrary $r > 0$ and $k \ge 0$ (compare [35]) and obtain

$$\tilde{d}'_D(r) < 0\,, \quad \tilde{d}'_{RG}(r) < 0\,, \quad \tilde{d}''_D(r) > 0 \quad \text{and} \quad \tilde{d}''_{RG}(r) > 0\,. \tag{2.19}$$

For the proof of (2.19) see also [12] and [4]. In our next theorem we prove that the function

$$\tilde{d}_{ME}(r) = \frac{\|R_r^{m+1/2} y^\delta\|^2}{\|R_r^{m+1} y^\delta\|}$$

has analogous properties.

Theorem 2.5. *The function $\tilde{d}_{ME}(r)$ is analytic, strictly monotonically decreasing and strictly convex. For all $r > 0$ there hold the estimates*

$$\tilde{d}'_{ME}(r) \leq -\frac{m\|A^* R_r^{m+1} y^\delta\|^2}{\|R_r^{m+1} y^\delta\|}, \tag{2.20}$$

$$\tilde{d}''_{ME}(r) \geq \frac{m(m+1)\|R_r^{m+1/2} y^\delta\|^2 \|AA^* R_r^{m+2} y^\delta\|^2}{\|R_r^{m+1} y^\delta\|^3}. \tag{2.21}$$

Proof. From [35] we have that (2.20) holds true and that $\tilde{d}'_{ME}$ is given by

$$\begin{aligned}\tilde{d}'_{ME}(r) \;=\; & \frac{(m+1)\|R_r^{m+1/2} y^\delta\|^2 \|A^* R_r^{m+3/2} y^\delta\|^2}{\|R_r^{m+1} y^\delta\|^3} \\ & - \frac{(2m+1)\|R_r^{m+1} y^\delta\|^2 \|A^* R_r^{m+1} y^\delta\|^2}{\|R_r^{m+1} y^\delta\|^3}. \end{aligned} \tag{2.22}$$

In order to prove (2.21) we differentiate (2.22) as follows. We use (2.17), (2.18) with $k = m + 1/2$ and $k = m + 1$, respectively, apply the identity

$$\frac{\mathrm{d}}{\mathrm{d}r}\|R_r^{m+1} y^\delta\|^3 = \frac{\mathrm{d}}{\mathrm{d}r}(\|R_r^{m+1} y^\delta\|^2)^{3/2} = \frac{3}{2}\|R_r^{m+1} y^\delta\|(-2m-2)\|A^* R^{m+3/2} y^\delta\|^2,$$

use the notations

$$\begin{aligned} t_1 &:= \|R_r^{m+1/2} y^\delta\|^2 \|A^* R_r^{m+3/2} y^\delta\|^4 \\ t_2 &:= \|A^* R_r^{m+1} y^\delta\|^2 \|A^* R_r^{m+3/2} y^\delta\|^2 \|R_r^{m+1} y^\delta\|^2 \\ t_3 &:= \|R_r^{m+1} y^\delta\|^4 \|AA^* R_r^{m+3/2} y^\delta\|^2 \\ t_4 &:= \|R_r^{m+1/2} y^\delta\|^2 \|R_r^{m+1} y^\delta\|^2 \|AA^* R_r^{m+2} y^\delta\|^2 \end{aligned}$$

and obtain

$$\begin{aligned} \tilde{d}''_{ME}(r)\frac{\|R_r^{m+1} y^\delta\|^5}{m+1} &= 3(m+1)t_1 - 2(2m+1)t_2 + 2(2m+1)t_3 - (2m+3)t_4 \\ &= 3(m+1)(t_1 - t_2 + t_3 - t_4) + (m-1)(t_3 - t_2) + m t_4. \end{aligned} \tag{2.23}$$

We use the abbreviation $D := (AA^*)^{1/2}$, apply the identity

$$\|z\|^2 = \|R_r^{1/2} z\|^2 + r\|D R_r^{1/2} z\|^2 \tag{2.24}$$

with $z = R_r^{m+1} y^\delta$ and $z = D R_r^{m+1} y^\delta$, respectively, and obtain

$$t_3 - t_2 = \|R_r^{m+1} y^\delta\|^2 \left(\|R_r^{m+3/2} y^\delta\|^2 \|AA^* R^{m+3/2} y^\delta\|^2 - \|A^* R_r^{m+3/2} y^\delta\|^4\right) \geq 0. \tag{2.25}$$

Repeated use of (2.24) yields

$$\|D^i R_r^{m+2-k/2} y^\delta\|^2 = \sum_{j=0}^{k} \binom{k}{j} r^j a_{i+j} \qquad \text{with} \quad a_j = \|D^j R_r^{m+2} y^\delta\|^2.$$

We apply this formula for $k = 1, 2, 3$ and $i = 0, ..., 3 - k$ and obtain

$$\begin{aligned} t_1 - t_2 + t_3 - t_4 &= (a_0 + 3ra_1 + 3r^2a_2 + r^3a_3)(a_1 + ra_2)^2 \\ &\quad -(a_0 + 2ra_1 + r^2a_2)(a_1 + 2ra_2 + r^2a_3)(a_1 + ra_2) \\ &\quad +(a_0 + 2ra_1 + r^2a_2)^2(a_2 + ra_3) \\ &\quad -a_2(a_0 + 2ra_1 + r^2a_2)(a_0 + 3ra_1 + 3r^2a_2 + r^3a_3) \\ &= (r^4a_2 + 3r^3a_1 + 3r^2a_0)(a_3a_1 - a_2^2) \\ &\quad +r(a_1^{3/2} - a_3^{1/2}a_0)^2 + 2ra_0a_1[(a_3a_1)^{1/2} - a_2]\,. \end{aligned} \tag{2.26}$$

From the Cauchy-Schwarz inequality we have

$$a_3a_1 - a_2^2 = \|D^3R_r^{m+2}y^\delta\|^2\|DR_r^{m+2}y^\delta\|^2 - (D^3R_r^{m+2}y^\delta, DR_r^{m+2}y^\delta)^2 \geq 0\,,$$

and due to (2.26) we obtain $t_1 - t_2 + t_3 - t_4 \geq 0$. From this inequality, (2.25) and (2.23) we conclude that $\tilde{d}''_{ME}(r)\|R_r^{m+1}y^\delta\|^5 \geq m(m+1)t_4$ which is equivalent to (2.21). □

Let us rewrite the derivative (2.22) into some equivalent form which is suitable for numerical computations. We apply the identity $r_{\alpha,k} = y^\delta - Ax^\delta_{\alpha,k} = K_\alpha^k y^\delta = R_r^k y^\delta$ with $K_\alpha = \alpha(AA^* + \alpha I)^{-1}$ and obtain from (2.22)

$$\tilde{d}'_{ME}(r) = \frac{(m+1)(r_{\alpha,m}, r_{\alpha,m+1})(A^*r_{\alpha,m+1}, A^*r_{\alpha,m+2}) - (2m+1)\|A^*r_{\alpha,m+1}\|^2\|r_{\alpha,m+1}\|^2}{\|r_{\alpha,m+1}\|^3}$$

with $\alpha = 1/r$. This representation shows that the evaluation of the derivative $\tilde{d}'_{ME}(r)$ requires the computation of $r_{\alpha,m}$, $r_{\alpha,m+1}$ and $r_{\alpha,m+2}$. Note that an efficient evaluation of the derivatives $\tilde{d}'_D(r)$ and $\tilde{d}'_{RG}(r)$ requires only the computation of $r_{\alpha,m}$ and $r_{\alpha,m+1}$ (see [4]). As an alternative to Newton's method applied to the equation $\tilde{d}_{ME}(r) - \delta = 0$ one could use the secant method in which the expensive evaluation of $\tilde{d}'_{ME}(r)$ is not necessary.

3. ITERATIVE REGULARIZATION METHODS

3.1. Iteration methods and the ME rule

For approximately solving linear ill-posed problems (1.1) with noisy data $y^\delta \in Y$ we consider iteration methods of the general form

$$x_n^\delta = x_{n-1}^\delta + A^*z_{n-1}\,, \quad n = 1, 2, ... \tag{3.1}$$

with $z_n \in X$ and initial guess $x_0^\delta = 0$. The elements z_n characterize the special iteration method. For example, $z_n = \beta(y^\delta - Ax_n^\delta)$ with $\beta \in (0, 2/\|A\|^2]$ leads to the well-known Landweber iteration.

Iteration methods for approximately solving ill-posed problems are especially attractive for *large scale problems* (cf. [19]). Such problems arise e.g. in the field of parameter identification in differential equations. Exploiting ideas from control theory, in such identification problems the elements A^*z_n can effectively be computed by solving one direct problem and one associated adjoint problem.

The iterates x_n^δ of the iteration process (3.1) generally diverge. Nevertheless these iterates allow a stable approximation of $x^\dagger$ provided the iteration is stopped after an appropriate number of iteration steps. Two well-known *a posteriori* rules of choosing the stopping index $n = n(\delta)$ are:

1. *Morozov's discrepancy principle* [27, 38, 39]. By this principle (*D principle*) the stopping index in (3.1) is chosen as the first index $n = n_D$ satisfying

$$d_D(n) := \|r_n\| \le C\delta \quad \text{with} \quad r_n = y^\delta - Ax_n^\delta \quad \text{and} \quad C \ge 1\,. \tag{3.2}$$

2. *Rule of Engl and Gfrerer* [3]. This rule may be applied to iteration methods (3.1) where z_n has the special form $z_n = h_n(\beta_n, AA^*)r_n$ with $r_n = y^\delta - Ax_n^\delta$. Here h_n is some operator function which depends on $\beta_n \in \mathrm{I\!R}$ and β_n is allowed to depend on the noisy data y^δ. In this rule (which we call *EG rule*) the stopping index in (3.1) is chosen as the first index $n = n_{EG}$ satisfying

$$d_{EG}(n) := \frac{(r_n + r_{n+1}, h_n(\beta_n, AA^*)r_n)^{1/2}}{(2\kappa_n)^{1/2}} \le C\delta \quad \text{with} \quad C \ge 1\,. \tag{3.3}$$

Here κ_n is a constant with

$$\kappa_n = \sup\{h_n(\beta_n, \lambda) \,|\, 0 \le \lambda \le \|A\|^2\}. \tag{3.4}$$

In the monotone error rule (ME rule) for choosing the stopping index we focus our attention on the monotonicity property (1.3). Our aim consists in searching for a *largest computable iteration number* n_{ME} for which (1.3) can be guaranteed. Exploiting (3.1) and using the notation $r_n = y^\delta - Ax_n^\delta$ we obtain

$$\begin{aligned}
\|x_n^\delta - x^\dagger\|^2 - \|x_{n-1}^\delta - x^\dagger\|^2 &= \|x_{n-1}^\delta + A^*z_{n-1} - x^\dagger\|^2 - \|x_{n-1}^\delta - x^\dagger\|^2 \\
&= 2(x_{n-1}^\delta - x^\dagger, A^*z_{n-1}) + \|A^*z_{n-1}\|^2 \\
&= (x_{n-1}^\delta + x_n^\delta - 2x^\dagger, A^*z_{n-1}) \\
&= (2(y^\delta - y) - (r_{n-1} + r_n), z_{n-1}) \\
&\le 2\|z_{n-1}\| \left\{ \delta - \frac{(r_{n-1} + r_n, z_{n-1})}{2\|z_{n-1}\|} \right\}.
\end{aligned} \tag{3.5}$$

This estimate leads us to the following ME rule for iteration methods (3.1):

ME rule. Choose n_{ME} as the first index n satisfying

$$d_{ME}(n) := \frac{(r_n + r_{n+1}, z_n)}{2\|z_n\|} \le \delta\,. \tag{3.6}$$

By this *a posteriori* choice of the stopping index in iteration methods (3.1) the monotonicity property (1.3) can be guaranteed:

Proposition 3.1. *Let $\|z_n\| \ne 0$ for $n = 0, 1, 2, \ldots$ and let n_{ME} be chosen by the ME rule (3.6). Then the monotonicity property (1.3) holds true.*

Proof. From the ME rule for iteration methods we have that $d_{ME}(n-1) > \delta$ for $n = 1, 2, \ldots, n_{ME}$. Consequently, from (3.5) there follows for all $n = 1, 2, \ldots, n_{ME}$ that

$$\|x_n^\delta - x^\dagger\|^2 \leq \|x_{n-1}^\delta - x^\dagger\|^2 + 2\|z_{n-1}\|\Big\{\delta - d_{ME}(n-1)\Big\} < \|x_{n-1}^\delta - x^\dagger\|^2 \qquad (3.7)$$

which completes the proof. $\square$

3.2. The ME rule in gradient type methods

Let us consider gradient type methods of the form (3.1) with $z_n = \beta_n r_n$ where $r_n = y^\delta - Ax_n^\delta$ is the discrepancy and $\beta_n > 0$ is some properly chosen stepsize. For such methods the iteration (3.1) attains the form

$$x_n^\delta = x_{n-1}^\delta + \beta_{n-1} A^*(y^\delta - Ax_{n-1}^\delta)\,, \quad n = 1, 2, \ldots. \qquad (3.8)$$

Since in the EG rule $h_n(\beta_n, \lambda) = \beta_n$, for the constant κ_n of (3.4) we have $\kappa_n = \beta_n$. Consequently, the functions $d_{EG}(n)$ and $d_{ME}(n)$ of the EG- and ME rules, respectively, attain the form

$$d_{EG}(n) = \frac{(r_n + r_{n+1}, r_n)^{1/2}}{\sqrt{2}} \quad \text{and} \quad d_{ME}(n) = \frac{(r_n + r_{n+1}, r_n)}{2\|r_n\|}\,. \qquad (3.9)$$

For gradient type methods (3.8) following properties are valid:

Theorem 3.2. *Let P denote the orthoprojection of Y onto $N(A^*) = \overline{R(A)}^\perp$, let $A^*y^\delta \neq 0$ and let β_n be chosen such that*

$$0 < \beta_n \leq \frac{\|A^* r_n\|^2}{\|AA^* r_n\|^2}\,. \qquad (3.10)$$

Then for the iterates of (3.8) following properties are valid:

(i) *The function $d_D(n) = \|r_n\|$ is strictly monotonically decreasing and obeys*

$$\|r_{n+1}\|^2 \leq (r_n, r_{n+1}) < \|r_n\|^2\,.$$

(ii) *The functions $d_{ME}(n)$ and $d_{EG}(n)$ are strictly monotonically decreasing and obey*

$$d_D(n+1) < d_{ME}(n) < d_{EG}(n) < d_D(n)\,.$$

(iii) *Let $\beta_n \geq c > 0$ with some positive constant c, then*

$$\lim_{n\to\infty} d_{ME}(n) = \lim_{n\to\infty} d_{EG}(n) = \lim_{n\to\infty} d_D(n) = \|Py^\delta\|\,. \qquad (3.11)$$

(iv) *If $\|Py^\delta\| < C\delta$, then the stopping indices n_D and n_{EG} are well defined. For $C = 1$ also n_{ME} is well defined and there holds*

$$n_D - 1 \leq n_{ME} \leq n_{EG} \leq n_D\,.$$

(v) *If $\|Ax_n^\delta - y^\delta\| \geq \delta$, then*

$$\|x_n^\delta - x^\dagger\| < \|x_{n-1}^\delta - x^\dagger\|\,.$$

Proof. From (3.8) we conclude that $r_{n+1} = (I - \beta_n AA^*)r_n$. Consequently,

$$\text{(a)} \quad (r_n, r_{n+1}) = \|r_n\|^2 - \beta_n \|A^* r_n\|^2 \,,$$
$$\text{(b)} \quad \|r_{n+1}\|^2 = \|r_n\|^2 - 2\beta_n \|A^* r_n\|^2 + \beta_n^2 \|AA^* r_n\|^2 \,.$$

From (a), $\beta_n > 0$ and $A^* y^\delta \neq 0$ we obtain the right inequality of (i). Combining (a) and (b) we have

$$\|r_{n+1}\|^2 = (r_n, r_{n+1}) - \beta_n \|A^* r_n\|^2 + \beta_n^2 \|AA^* r_n\|^2 \,.$$

From this equation and assumption (3.10) we obtain that the left inequality of (i) holds true. In order to prove the left inequality of (ii) we use the inequality $2ab < a^2 + b^2$ with $a \neq b$ and obtain with the help of the left inequality of (i) that

$$2\|r_n\| \|r_{n+1}\| < \|r_n\|^2 + \|r_{n+1}\|^2 \leq \|r_n\|^2 + (r_n, r_{n+1}) \,,$$

which is equivalent to $d_D(n+1) < d_{ME}(n)$. The middle and right inequalities of (ii) are equivalent to $(r_n, r_{n+1}) < \|r_n\|^2$ and follow from part (i). For proving (3.11) we proceed according to (3.5) and obtain for arbitrary $w \in X$ and arbitrary iteration methods (3.8) with $\beta_n > 0$ that

$$\begin{aligned} \|x_n^\delta - w\|^2 - \|x_{n-1}^\delta - w\|^2 &= \Big(2(y^\delta - Aw) - (r_{n-1} + r_n), \beta_{n-1} r_{n-1}\Big) \\ &\leq 2\beta_{n-1} \|r_{n-1}\| \Big\{\|y^\delta - Aw\| - d_{ME}(n-1)\Big\} \,. \end{aligned} \tag{3.12}$$

For iteration methods (3.8) with stepsizes β_n satisfying (3.10) there holds

$$\|r_n\|^2 - \|r_{n+1}\|^2 = \beta_n \Big(2\|A^* r_n\|^2 - \beta_n \|AA^* r_n\|^2\Big) \geq \beta_n \|A^* r_n\|^2 \,.$$

Passing to the limit on both sides yields due to $\beta_n \geq c > 0$ that $\lim_{n\to\infty} \|A^* r_n\| = 0$. From $y^\delta = Ax_n^\delta + r_n$ and the Cauchy-Schwarz-inequality we have

$$\|r_n\|^2 - \|Aw - y^\delta\|^2 = 2\left(r_n, Aw - Ax_n^\delta\right) - \|Ax_n^\delta - Aw\|^2 \leq 2\|A^* r_n\| \|x_n^\delta - w\| \,. \tag{3.13}$$

Now we will use a contradicition argument and assume that $\lim_{n\to\infty} d_D(n) > \|Py^\delta\|$. Under this condition there exists some element $w \in X$ with the property that

$$\lim_{n\to\infty} \|r_n\| > \|Aw - y^\delta\| \,. \tag{3.14}$$

From (3.14) and (i) we have $\|r_n\| > \|Aw - y^\delta\|$ for all $n \in \mathbb{N}$. This property provides together with (3.12) and (ii) that

$$\|x_n^\delta - w\|^2 - \|x_{n-1}^\delta - w\|^2 < 2\beta_{n-1} \|r_{n-1}\| \{\|r_n\| - d_{ME}(n-1)\} \leq 0 \,.$$

Hence, $\|x_n^\delta - w\|$ is bounded by $\|w\|$ for all $n \in \mathbb{N}$. Passing to the limit on both sides of (3.13) provides due to $\|x_n^\delta - w\| \leq \|w\|$ and $\lim_{n\to\infty} \|A^* r_n\| = 0$ that $\lim_{n\to\infty} \|r_n\| \leq \|Aw - y^\delta\|$. This contradicts (3.14) and shows that our assumption $\lim_{n\to\infty} d_D(n) > \|Py^\delta\|$ cannot hold true. Hence, due to (ii) we obtain (3.11). Assertion (iv) follows from (3.11), the definition of the stopping rules and (ii). In order to prove (v) we use the first part of (3.7), property (ii) as well as the assumption $d_D(n) \geq \delta$ and obtain

$$\begin{aligned} \|x_n^\delta - x^\dagger\|^2 - \|x_{n-1}^\delta - x^\dagger\|^2 &\leq 2\|z_{n-1}\| \{\delta - d_{ME}(n-1)\} \\ &< 2\|z_{n-1}\| \{\delta - d_D(n)\} \\ &\leq 0 \,, \end{aligned}$$

which finishes the proof. □

Remark 3.1. The limit relation (3.11) can not only be guaranteed for strictly positive stepsizes $\beta_n \geq c > 0$, but also for stepsizes tending to zero not too fast and satisfying

$$0 \leq \beta_i \leq \frac{1}{\|A\|^2} \quad \text{and} \quad \lim_{n\to\infty} \sum_{i=0}^{n} \beta_i = \infty\,.$$

For the proof of this result we use the identity $r_{n+1} = (I - \beta_n AA^*)r_n$ which gives

$$r_{n+1} = r_n(AA^*)y^\delta \quad \text{with} \quad r_n(\lambda) = \prod_{i=0}^{n}(1-\beta_i\lambda)\,.$$

Hence, for $n \to \infty$ we have for all $\lambda \in (0, \|A\|^2]$

$$|r_n(\lambda)| \to 0 \quad \Leftrightarrow \quad \ln\prod_{i=0}^{n}|1-\beta_i\lambda| \to -\infty \quad \Leftrightarrow \quad \sum_{i=0}^{n}\ln|1-\beta_i\lambda| \to -\infty\,.$$

Using the estimate $\ln|1-\xi| \leq -\xi$ for $\xi \in [0,1]$ with $\xi = \beta_i\lambda$ we obtain the result that for $0 \leq \beta_i \leq 1/\|A\|^2$ there holds $|r_n(\lambda)| \to 0$ provided $\sum_{i=0}^n \beta_i \to \infty$. Finally we conclude that from $|r_n(\lambda)| \to 0$ for all $\lambda \in (0, \|A\|^2]$ there follows $\lim_{n\to\infty}\|r_n\| = \|Py^\delta\|$ and the proof is complete.

Remark 3.2. Part (v) of Theorem 3.2 shows that the iteration (3.8) should not be stopped as long as $\|Ax_n^\delta - y^\delta\| \geq \delta$ holds. Let us modify the D principle as follows: Choose $n = n_D$ as the first index n satisfying

$$\|Ax_{n+1}^\delta - y^\delta\| < C\delta \quad \text{with} \quad C \geq 1\,.$$

Let n_1 and n_2 the stopping indices of this modified D principle with $C = C_1$ and $C = C_2$, respectively. If $1 \leq C_1 \leq C_2$, then, since $d_D(\alpha)$ is monotonically decreasing there follows $n_1 \geq n_2$, and due to part (v) of Theorem 3.2 we obtain $\|x_{n_1}^\delta - x^\dagger\| \leq \|x_{n_2}^\delta - x^\dagger\|$. Hence, the best possible choice for the constant C in this modified D principle is $C = 1$.

Before we will study some special methods that fit into the framework of Theorem 3.2 let us derive some useful inequality that is helpful for checking condition (3.10).

Proposition 3.3. *Let $D \in \mathcal{L}(Y,Y)$, $D = D^* \geq 0$ and $\eta \geq 0$. Then for all $v \in Y$*

$$\|D^\eta v\|\|Dv\| \leq \|D^{\eta+1}v\|\|v\|\,. \tag{3.15}$$

Proof. We apply the moment inequality (2.14), first with $s = \eta$ and $t = \eta + 1$, second with $s = 1$ and $t = \eta + 1$ and obtain the inequalities

$$\|D^\eta v\| \leq \|D^{\eta+1}v\|^{\frac{\eta}{\eta+1}}\|v\|^{\frac{1}{\eta+1}} \quad \text{and} \quad \|Dv\| \leq \|D^{\eta+1}v\|^{\frac{1}{\eta+1}}\|v\|^{\frac{\eta}{\eta+1}}\,.$$

We multiply both inequalities and obtain (3.15). □

Now we are ready to study some special gradient type methods (3.8) that fit into the framework of Theorem 3.2. In the methods M3 – M5 below we assume that y^δ is not an eigenelement of the operator AA^* since in the opposite case there follows $r_1 = 0$ and the stepsizes β_n, $n = 1, 2, \ldots$ are not defined.

METHOD M1: Landweber's method with $\beta_n = \beta \in (0, 1/\|A\|^2]$. This method may actually be applied with stepsizes $\beta_n = \beta \in (0, 2/\|A\|^2)$, see [1, 4, 5, 13, 17, 23, 25, 28, 30, 38, 39]. However, due to the inequality $\|AA^* r_n\| \le \|A\| \|A^* r_n\|$ we realize that condition (3.10) and hence the results of Theorem 3.2 hold true for $\beta \in (0, 1/\|A\|^2]$.

METHOD M2: Nonstationary Landweber's method. This method is characterized by (3.8) with variable $\beta_n \in [c, 1/\|A\|^2]$ and $c > 0$, see [32]. As in method M1 we conclude that condition (3.10) and hence the results of Theorem 3.2 hold true for $\beta_n \in [c, 1/\|A\|^2]$.

METHOD M3: Steepest descent method (see [1, 2, 7, 9, 10, 11, 16, 20, 22, 31, 32]). This method is characterized by (3.8) with $\beta_n = \|A^* r_n\|^2 / \|AA^* r_n\|^2$. Since this stepsize satisfies (3.10), the results of Theorem 3.2 hold true for this method.

METHOD M4: α-processes with stepsizes $\beta_n = \|D^{\alpha+1} r_n\|^2 / \|D^{\alpha+2} r_n\|^2$, $D = (AA^*)^{1/2}$ and $\alpha \ge 0$. These methods may actually be applied with $\alpha \ge -1$, see [10, 16, 22, 31, 32]. However, assumption (3.10) holds true only for $\alpha \ge 0$. In order to check (3.10) we apply inequality (3.15) with $D = (AA^*)^{1/2}$, $v = (AA^*)^{1/2} r_n$ and $\eta = \alpha$ and obtain the inequality $\|D^{\alpha+1} r_n\| \|D^2 r_n\| \le \|D^{\alpha+2} r_n\| \|D r_n\|$ which gives $\beta_n \le \|D r_n\| / \|D^2 r_n\|$. Since furthermore $\beta_n \ge \|D^{\alpha+1} r_n\|^2 / (\|D\|^2 \|D^{\alpha+1} r_n\|^2) = 1/\|A\|^2$ we obtain

$$\frac{1}{\|A\|^2} \le \beta_n \le \frac{\|A^* r_n\|^2}{\|AA^* r_n\|^2} \tag{3.16}$$

and hence (3.10). Since the application of (3.15) requires $\eta \ge 0$ we realize that Theorem 3.2 holds true for α-processes with $\alpha \ge 0$.

METHOD M5: Method (3.8) with $\beta_n = \max\{\|D^{\alpha+2} r_n\|^2 / (\|D\|^4 \|D^{\alpha+1} r_n\|^2), \beta\}$ where $D = (AA^*)^{1/2}$, $\beta \in (0, 1/\|A\|^2]$ and $\alpha \ge -1$ (see [16, 32]). For this method there holds

$$\beta \le \beta_n \le \max\left\{\frac{\|D\|^2 \|D^{\alpha+1} r_n\|^2}{\|D\|^4 \|D^{\alpha+1} r_n\|^2}, \beta\right\} = \frac{1}{\|A\|^2}.$$

From this estimate it follows as in method M1 that condition (3.10) and therefore the results of Theorem 3.2 hold true.

Note that method M3 is a special case of α-processes with $\alpha = 0$. In this method the stepsize β_n minimizes the functional $g(\beta_n) := \|r_{n+1}\|^2 = \|(I - \beta_n AA^*) r_n\|^2$.

The results of Theorem 3.2 can be used to establish convergence- and convergence rate results for the above discussed *a posteriori* stopping rules in iteration methods M1 – M5.

Theorem 3.4. *Assume $A^* y^\delta \ne 0$ and $\|P y^\delta\| < \delta$. Let x_n^δ the regularized approximation obtained by one of the methods M1 – M5 and let n_D, n_{EG} and n_{ME} the stopping indices according to the discrepancy principle with $C \ge 1$, the EG rule (3.3) with $C \ge 1$ and the ME rule (3.6), respectively. Then for all $n \in \{n_D, n_{ME}, n_{EG}\}$ there holds:*

(i) $\|x_n^\delta - x^\dagger\| \to 0$ *for* $\delta \to 0$.

(ii) If $x^\dagger \in R((A^* A)^{p/2})$, *then*

$$\|x_n^\delta - x^\dagger\| = O(\delta^{p/(p+1)}) \quad \textit{for all} \quad p > 0\,.$$

Proof. First, let us discuss the case $C > 1$. In this case the proof for $n = n_D$ is known from the literature, see [38] for method M1, [31, 32] for methods M2–M5 and [9, 10] for method M3. Let $n_{ME,C}$ the stopping index of the ME rule (3.6) with δ replaced by $C\delta$. From the validity of Theorem 3.4 for $n = n_D$ and part (ii) of Theorem 3.2 we obtain that assertions (i), (ii) are valid for n_{EG} and $n_{ME,C}$ as well. Second, let us consider the case $C = 1$. From the monotonicity of $d_{ME}(n)$ we obtain that $n_{ME,C} \le n_{ME,1}$. Consequently, due to Proposition 3.1 there follows

$$\|x^\delta_{n_{ME,1}} - x^\dagger\| \le \|x^\delta_{n_{ME,C}} - x^\dagger\| .$$

Hence, the assertions (i) and (ii) of the theorem are true for $n = n_{ME,1}$. From this result and part (iv) of Theorem 3.2 we conclude that assertions (i) and (ii) of the theorem are also valid for $n = n_D$ and $n = n_{EG}$. □

Some weaker results compared with the results of Theorems 3.2 and 3.4 can be proved for method M4 with $\alpha \in [-1, 0)$. This class of methods contains for $\alpha = -1$ the *minimal error method* in which the stepsize β_n is given by $\beta_n = \|r_n\|^2/\|A^*r_n\|^2$, see [1, 2, 6, 9, 10, 11, 16, 22, 31, 32]. In the minimal error method the stepsize β_n minimizes the norm $\|x^\delta_{n+1} - A^{-1}y^\delta\|$ provided $A^{-1}y^\delta$ exists. For this method there holds $(r_n, r_{n+1}) = 0$ which shows that the left inequality of part (i) of Theorem 3.2 generally does not hold.

Theorem 3.5. *Let $A^*y^\delta \ne 0$. Then in method M4 with $\alpha \in [-1, 0)$ following properties are valid:*

(i) For all $n \in \mathbb{N}$ there holds

$$\frac{1}{2} d_D(n) \le \frac{1}{\sqrt{2}} d_{EG}(n) \le d_{ME}(n) < d_{EG}(n) < d_D(n) .$$

(ii) Denote by $n_{ME,C}$ the stopping index of the ME rule with δ replaced by $C\delta$. Then, $n_{ME,C}$ is well defined for $C > 1$. In addition, $n_{EG,C} := n_{EG}$ is well defined for $C > \sqrt{2}$ and $n_{D,C} := n_D$ is well defined for $C > 2$. For arbitrary $C > 1$ there holds

$$n_{D,2C} \le n_{EG,\sqrt{2}C} \le n_{ME,C} \le n_{EG,C} \le n_{D,C} . \tag{3.17}$$

(iii) For $n = n_{D,2C}$, $n = n_{EG,\sqrt{2}C}$ and $n = n_{ME,C}$ with $C > 1$ there holds

$$\|x^\delta_n - x^\dagger\| \to 0 \quad \textit{for} \quad \delta \to 0 .$$

Proof. By elementary computations it can be shown that the first two inequalities of assertion (i) are equivalent to $(r_n, r_{n+1}) \ge 0$. This inequality, however, is equivalent to

$$\beta_n \le \frac{\|r_n\|^2}{\|A^*r_n\|^2}$$

and follows from (3.15) with $\eta = \alpha + 1$ and $v = r_n$. The final two inequalities of assertion (i) are both equivalent to $(r_n, r_{n+1}) < \|r_n\|^2$. This inequality, however, holds for arbitrary stepsize $\beta_n > 0$ since due to $r_{n+1} = [I - \beta_n AA^*]r_n$ and $A^*y^\delta \ne 0$ we have

$$(r_n, r_{n+1}) = \|r_n\|^2 - \beta_n \|A^*r_n\|^2 < \|r_n\|^2 .$$

Now let us prove indirectly that for $C > 1$ there exists a finite stopping index $n_{ME,C}$. For this aim we assume that $d_{ME}(n) > C\delta$ for all $n \in \mathbb{N}$. Then, due to (3.5), $\|x_n^\delta - x^\dagger\| < \|x_{n-1}^\delta - x^\dagger\|$, and the limit $\lim_{n\to\infty} \|x_n^\delta - x^\dagger\|$ exists. We summize the inequalities (3.5) with $z_n = \beta_n r_n$ for all $n \geq 1$ and obtain due to $x_0^\delta = 0$ that

$$\|x^\dagger\|^2 - \lim_{n\to\infty} \|x_n^\delta - x^\dagger\|^2 \geq 2\sum_{n=0}^{\infty} \beta_n \|r_n\| \left\{d_{ME}(n) - \delta\right\} .$$

Since the right hand side is finite there follows that $\lim\limits_{n\to\infty} \beta_n \|r_n\| \left\{d_{ME}(n) - \delta\right\} = 0$. Due to $\beta_n \geq 1/\|A\|^2$, $\|r_n\| > d_{ME}(n)$ (compare (i)) and the assumption $d_{ME}(n) > C\delta$ there follows

$$\beta_n \|r_n\| \geq \frac{\|r_n\|}{\|A\|^2} > \frac{d_{ME}(n)}{\|A\|^2} > \frac{C\delta}{\|A\|^2} .$$

Consequently, $\lim_{n\to\infty} \left\{d_{ME}(n) - \delta\right\} = 0$. This contradicts our assumption $d_{ME}(n) > C\delta$ for all $n \in \mathbb{N}$ and proves that there has to exist some finite stopping index $n_{ME,C}$ for $C > 1$. Hence, due to the first two inequalities of (i) there follows that $n_{EG,C}$ is well defined for $C > \sqrt{2}$ and $n_{D,C}$ is well defined for $C > 2$. Now the proof of (3.17) follows from the definition of the corresponding stopping rules and (i). In order to prove (iii) we conclude from $n_{D,2C} \leq n_{EG,\sqrt{2}C} \leq n_{ME,C}$ and Proposition 3.1 that

$$\|x_{n_{ME,C}}^\delta - x^\dagger\| \leq \|x_{n_{EG,\sqrt{2}C}}^\delta - x^\dagger\| \leq \|x_{n_{D,2C}}^\delta - x^\dagger\| .$$

Since assertion (iii) holds true for $n_{D,2C}$ with $C > 1$ (see [10]), we conclude that (iii) holds also true for $n_{EG,\sqrt{2}C}$ and $n_{ME,C}$ with $C > 1$. $\square$

Remark 3.3. The estimate (3.5) which led us to the ME rule can also be exploited for finding stepsizes $\beta_n > 0$ in iteration methods (3.8) which guarantee that x_{n+1}^δ is a better approximation for $x^\dagger$ than x_n^δ. Here the stepsize β_n may not only depend on y^δ, but also on the noise level δ. Exploiting (3.7), (3.6) and $r_{n+1} = (I - \beta_n AA^*)r_n$ we obtain for $\beta_n > 0$

$$\|x_{n+1}^\delta - x^\dagger\|^2 - \|x_n^\delta - x^\dagger\|^2 < 2\beta_n \|r_n\| \left\{\delta - \|r_n\| + \beta_n \frac{\|A^* r_n\|^2}{2\|r_n\|}\right\} . \tag{3.18}$$

The right hand side of (3.18) is negative for $\|r_n\| > \delta$ and $0 < \beta_n < \frac{2\|r_n\|}{\|A^* r_n\|^2}(\|r_n\| - \delta)$ which shows that for such stepsizes the element x_{n+1}^δ is a better approximation for $x^\dagger$ than x_n^δ. Minimizing the right hand side of (3.18) with respect to β_n yields

$$\beta_n = \frac{\|r_n\|}{\|A^* r_n\|^2}(\|r_n\| - \delta) .$$

Substituting into (3.18) shows (see also [1], p. 69) that for this stepsize the improvement of the squared error can be estimated by

$$\|x_{n+1}^\delta - x^\dagger\|^2 - \|x_n^\delta - x^\dagger\|^2 < -\frac{\|r_n\|^2}{\|A^* r_n\|^2}(\|r_n\| - \delta)^2 .$$

Some gradient type methods that do not fit into the class of methods (3.8) are *conjugate gradient methods*. The conjugate gradient method for the normal equations $A^*Ax = A^*y^\delta$ (see [1, 2, 9, 10, 11]) is known as powerful method for the approximate solution of ill-posed problems. The advantage of this method over methods of type (3.8) consists in the fact that the stopping index is generally much smaller. In the standard variant of this method (see [1], p. 51) the regularized approximations x_n^δ have the form

$$x_{n+1}^\delta = x_n^\delta + \beta_n p_n\,, \qquad p_n = A^* r_n + \gamma_{n-1} p_{n-1}\,, \qquad r_n = y^\delta - A x_n^\delta\,,$$

$$\beta_n = \frac{(A^* r_n, p_n)}{\|A p_n\|^2}\,, \qquad \gamma_{n-1} = \frac{\|A^* r_n\|^2}{\|A^* r_{n-1}\|^2}\,, \qquad p_0 = A^* r_0\,.$$

It can easily be realized that this method fits into the framework of methods (3.1). Hence, the ME rule (3.6) for choosing an appropriate stopping index can be applied. In [1] it is shown that for this method the ME function (3.6) has the form

$$d_{ME}(n) = \frac{\|r_n\|^2 + \|r_{n+1}\|^2}{2} \cdot \frac{\sum\limits_{i=0}^{n} 1/\|A^* r_i\|^2}{\sum\limits_{i=0}^{n} \|r_i\|/\|A^* r_i\|^2}\,.$$

This representation for $d_{ME}(n)$ has been used in [1] to prove that the regularized approximation $x_{n_{ME}}^\delta$ converges to $x^\dagger$ for $\delta \to 0$. Unfortunately, we don't know results concerning convergence rates although such results are known for the D principle (see [4]).

3.3. The ME rule in implicit iteration methods

Let us consider implicit iteration methods of the form (3.1) with $z_n = h_n(\beta_n, AA^*) r_n = (\beta_n I + AA^*)^{-1} r_n$ where $r_n = y^\delta - A x_n^\delta$ is the discrepancy and $\beta_n > 0$ is some suitably chosen real number. For such methods the iteration (3.1) attains the form

$$\begin{aligned} x_n^\delta &= x_{n-1}^\delta + A^*(AA^* + \beta_{n-1} I)^{-1}(y^\delta - A x_{n-1}^\delta) \\ &= (A^*A + \beta_{n-1} I)^{-1}(\beta_{n-1} x_{n-1}^\delta + A^* y^\delta)\,, \qquad n = 1, 2, \ldots\,. \end{aligned} \tag{3.19}$$

For methods (3.19) the constant κ_n in (3.4) is given by $\kappa_n = \beta_n^{-1}$. From (3.19) we obtain that $r_n = \beta_{n-1}(AA^* + \beta_{n-1} I)^{-1} r_{n-1}$. Consequently, the element $z_n = (\beta_n I + AA^*)^{-1} r_n$ has the form $z_n = \beta_n^{-1} r_{n+1}$ and the functions $d_{EG}(n)$ and $d_{ME}(n)$ of the EG- and ME rules of Subsection 3.1 attain the form

$$d_{EG}(n) = \frac{(r_n + r_{n+1}, r_{n+1})^{1/2}}{\sqrt{2}} \quad \text{and} \quad d_{ME}(n) = \frac{(r_n + r_{n+1}, r_{n+1})}{2\|r_{n+1}\|}\,, \tag{3.20}$$

respectively. For implicit iteration methods with *arbitrary* positive parameters $\beta_n > 0$ following properties are valid:

Theorem 3.6. *Let $A^* y^\delta \neq 0$ and $\beta_n > 0$ arbitrary. Then for the iterates of (3.19) following properties are valid:*

(i) The function $d_D(n) = \|r_n\|$ is strictly monotonically decreasing and obeys

$$\|r_{n+1}\|^2 < (r_n, r_{n+1}) < \|r_n\|^2\,.$$

(ii) The functions $d_{ME}(n)$ and $d_{EG}(n)$ are strictly monotonically decreasing and obey

(a) $d_D(n+1) < d_{EG}(n) < d_{ME}(n) < d_D(n)$

(b) $\lim\limits_{n\to\infty} d_{ME}(n) = \lim\limits_{n\to\infty} d_{EG}(n) = \lim\limits_{n\to\infty} d_D(n)$

(iii) *Let $\lim_{n\to\infty} d_D(n) < C\delta$ and $C \geq 1$. Then the stopping indices n_D and n_{ME} are well defined. If $C = 1$, then also n_{ME} is well defined and there holds*

$$n_D - 1 \leq n_{EG} \leq n_{ME} \leq n_D \,.$$

(iv) *If $\|Ax_n^\delta - y^\delta\| \geq \delta$, then*

$$\|x_n^\delta - x^\dagger\| < \|x_{n-1}^\delta - x^\dagger\| \,.$$

Proof. From (3.19) we conclude that $r_n = (\beta_n^{-1}AA^* + I)r_{n+1}$. Consequently,

$$\begin{aligned} &\text{(a)} \quad (r_n, r_{n+1}) = \|r_{n+1}\|^2 + \beta_n^{-1}\|A^*r_{n+1}\|^2 \,, \\ &\text{(b)} \quad \|r_n\|^2 = \|r_{n+1}\|^2 + 2\beta_n^{-1}\|A^*r_{n+1}\|^2 + \beta_n^{-2}\|AA^*r_{n+1}\|^2 \,. \end{aligned}$$

From (a), (b) and $A^*y^\delta \neq 0$ we obtain assertion (i). From (3.20), the Cauchy-Schwarz inequality, the triangle inequality and inequality (i) we obtain

$$d_{ME}(n) := \frac{(r_n + r_{n+1}, r_{n+1})}{2\|r_{n+1}\|} \leq \frac{1}{2}\Big(\|r_n\| + \|r_{n+1}\|\Big) < \|r_n\| = d_D(n) \,.$$

Hence, the right inequality of (ii) holds true. The two other inequalities of (ii) are equivalent to $\|r_{n+1}\|^2 < (r_n, r_{n+1})$ and follow from part (i). Now part (b) of (ii) is a consequence of part (a) of (ii) and assertion (iii) of the theorem is a consequence of assertion (ii). The proof of (iv) can be done in analogy to the proof of part (iv) of Theorem 3.2. □

As in explicit iteration methods, property (iv) of the theorem shows that also in implicit iteration methods (3.19) with arbitrary $\beta_n > 0$ the iteration should not be stopped as long as $\|Ax_n^\delta - y^\delta\| \geq \delta$ holds. Here the same discussion as in Remark 3.2 can be made. Further note that the limit $\lim_{n\to\infty} d_D(n)$ always exists since $d_D(n)$ is monotonically decreasing and bounded by zero. However, under the additional condition

$$\lim_{n\to\infty} \sum_{j=0}^{n} \frac{1}{\beta_j} = \infty \,, \tag{3.21}$$

which is necessary for convergence $x_n \to x^\dagger$ for $n \to \infty$ in the case of exact data, it can be shown that $\lim_{n\to\infty} d_D(n) = \|Py^\delta\|$ where P denotes the orthoprojection of Y onto $N(A^*) = \overline{R(A)}^{\perp}$ (cf., e.g., [18]). Condition (3.21) excludes that the parameter sequence $\{\beta_n\}$ is growing faster than the sequence $\{n^\alpha\}$ with $\alpha > 1$.

Our next aim in this section is to establish convergence- and convergence rate results for implicit iteration methods (3.19) provided the stopping index is chosen from the D principle, the EG rule and the ME rule, respectively. For the D principle there is known that the proof of convergence rates can be done under the additional assumption

$$\frac{1}{\beta_n} \leq c \sum_{j=0}^{n-1} \frac{1}{\beta_j} \quad \text{for some} \quad c > 0 \tag{3.22}$$

(see [18]). These results and the results of Theorem 3.6 allow the proof of convergence- and convergence rate results for the EG- and ME rules.

Theorem 3.7. *Assume $A^*y^\delta \neq 0$, $\|Py^\delta\| < \delta$ and (3.21). Let x_n^δ the regularized approximation obtained by (3.19) and let n_D, n_{EG} and n_{ME} the stopping indices according to the discrepancy principle with $C \geq 1$, the EG rule (3.3) with $C \geq 1$ and the ME rule (3.6), respectively. Then for all $n \in \{n_D, n_{ME}, n_{EG}\}$ there holds:*

(i) $\|x_n^\delta - x^\dagger\| \to 0$ *for* $\delta \to 0$.

(ii) If $x^\dagger \in R((A^*A)^{p/2})$ *and (3.22) hold, then*

$$\|x_n^\delta - x^\dagger\| = O(\delta^{p/(p+1)}) \quad \text{for all} \quad p > 0\,.$$

Proof. The proof is along the lines of the proof of Theorem 3.4 and uses that for $C > 1$ the results of the theorem for $n = n_D$ are known from [18]. □

Now let us study some special implicit iteration methods (3.19) that fit into the framework of Theorem 3.7.

METHOD M6: Stationary implicit iteration method (see [13, 24, 28, 30, 38, 39]). This method is characterized by (3.19) with fixed $\beta_n := \beta > 0$. Obviously, condition (3.21) holds and condition (3.22) is satisfied with $c = 1$. Hence the results of Theorem 3.7 hold true for this method.

METHOD M7: Nonstationary implicit iteration method with $\beta_n = \beta q^n$, $\beta > 0$ and $q \in (0, 1)$. For this method condition (3.21) can easily be checked and condition (3.22) holds true with $c = 1/q$ (see [18]). Hence, Theorem 3.7 applies for this method.

METHOD M8: Nonstationary implicit iteration method with $\beta_n \in [c_1, c_2]$, $0 < c_1 < c_2$. For this method (see [32]) property (3.21) is valid due to $\beta_n \le c_2$. From

$$\frac{1}{\beta_n} \le \frac{1}{c_1} \le \frac{1}{c_1} \cdot \frac{c_2}{n} \sum_{j=0}^{n-1} \frac{1}{\beta_j}$$

we obtain (3.22) with $c = c_2/c_1$. Consequently, Theorem 3.7 applies for this method.

METHOD M9: Nonstationary implicit iteration method with

$$\beta_n = \max\left\{ \frac{\|D^{\alpha+2} r_n\|^2}{\|D^{\alpha+1} r_n\|^2}\,,\, \beta \right\}\,, \quad \beta > 0\,, \quad \alpha \ge -1$$

and $D = (AA^*)^{1/2}$. A modification of this method with $\alpha = -1$ and without lower bound β has been studied in [29, 33, 37]. Method M9 (see [16, 32, 37]) is a special case of method M8 with

$$\beta \le \beta_n \le \max\left\{ \|D\|^2\,,\, \beta \right\}\,. \tag{3.23}$$

The inequalities (3.23) follow from the definition of β_n. Hence, the results of Theorem 3.7 hold true for this method.

4. NUMERICAL EXPERIMENTS

In this section we summarize some of our numerical results for integral equations of the first kind

$$(Ax)(s) = \int_0^1 K(s,t)x(t)\,\mathrm{d}t = y(s)\,, \quad 0 \le s \le 1 \tag{4.1}$$

in a L^2-space setting with $X = Y = L^2(0, 1)$. We discretized problem (4.1) by the collocation method with 100 piecewise constant spline basis functions on a uniform mesh.

Table 1: Errors in the method of ordinary Tikhonov regularization

δ	e_D	e_{RG}	e_{ME}	$e_D/\delta^{1/2}$	$e_{RG}/\delta^{3/5}$	$e_{ME}/\delta^{3/5}$
10^{-1}	.05878	.05982	.05355	.1859	.2381	.2132
10^{-2}	.02332	.01905	.01710	.2332	.3019	.2710
10^{-3}	.00896	.00589	.00527	.2833	.3716	.3325
10^{-4}	.00338	.00183	.00163	.3380	.4597	.4094
10^{-5}	.00119	.00055	.00049	.3763	.5500	.4900

Instead of y randomly perturbed data y^δ with $\|y - y^\delta\| \le \delta$ have been used. Our test problem is taken from [26] and corresponds to equation (4.1) with

$$K(s,t) = \begin{cases} 0 & \text{for} \quad t \le s \\ 1 & \text{for} \quad s < t\,, \end{cases} \qquad y(s) = \frac{1-\cos(\pi s)}{\pi}\,, \qquad x^\dagger(t) = \sin(\pi t)\,.$$

For this test problem we have $x^\dagger \in R((A^*A)^{p/2})$ for all $p \in (0, 3/2)$.

In a first experiment the discretized problem was regularized by the method of ordinary Tikhonov regularization (2.5) with $m = 1$. The regularization parameter α has been chosen according to the ME rule, the RG rule and the D principle, respectively. For C we have used the constant $C = 1$. In Table 1 the errors

$$e_D = \|x^\delta_{\alpha_D} - x^\dagger\|\,, \quad e_{RG} = \|x^\delta_{\alpha_{RG}} - x^\dagger\| \quad \text{and} \quad e_{ME} = \|x^\delta_{\alpha_{ME}} - x^\dagger\|$$

are given. The results in Table 1 verify the theoretical results of the estimate (2.8) of Theorem 2.2 which tells us that for $x^\dagger \in R((A^*A)^{p/2})$ with $p \le 2$ there holds $e_{ME} = O(\delta^{p/(p+1)})$. Table 1 also illustrates the well known results that for $1 \le p \le 2$ there holds $e_{RG} = O(\delta^{p/(p+1)})$ and $e_D = O(\delta^{1/2})$. In addition we observed that always $e_{ME} < e_{RG}$ holds true.

In a second experiment we solved the discretized problem by one explicit and one implicit iteration method, namely by Landweber's method with $\beta = 1/\|A\|^2$ and by the stationary implicit iteration method with $\beta = 1$. Both iterations were stopped with index $n_* + 10$, where n_* is the first n with $\|x^\delta_{n+1} - x^\dagger\| \ge \|x^\delta_n - x^\dagger\|$. In all experiments the final inequality appeared to be true for all $n = n_* + 1, \ldots, n_* + 10$. In Table 2 comparisons of the indices n_{ME}, n_* and the corresponding errors

$$e_{ME} = \|x_{n_{ME}} - x^\dagger\| \quad \text{and} \quad e_* = \|x_{n_*} - x^\dagger\|$$

are given. Note that in these and in all other numerical experiments we observed that $n_{ME} \le n_*$ and $n_{ME} = n_D$ or $n_{ME} = n_D - 1$, which is in agreement with part (iv) of Theorem 3.2 and part (iii) of Theorem 3.6. We made some further experiments in which instead of random perturbations some special perturbed data y^δ have been used. In these experiments we observed essentially smaller quotients $(n_* - n_{ME})/n_*$ compared with those which follow from Table 2. For some further numerical experiments for comparing the ME rule with the RG rule and the D principle in the method of ordinary Tikhonov regularization see [14, 16, 21, 35].

Table 2: Indices n_{ME}, n_* and errors e_{ME}, e_* in iteration methods

	Landweber's method				Implicit iteration method			
δ	n_{ME}	n_*	e_{ME}	e_*	n_{ME}	n_*	e_{ME}	e_*
10^{-1}	15	16	.0630	.0621	32	41	.1131	.1060
10^{-2}	38	67	.0453	.0400	103	166	.0406	.0356
10^{-3}	160	332	.0137	.0111	495	750	.0097	.0087
10^{-4}	947	1545	.0033	.0030	2422	3741	.0031	.0029
10^{-5}	4794	8643	.0009	.0008	12951	19667	.0008	.0008

Acknowledgement

The authors acknowledge assistance from Reimo Palm, University of Tartu, for providing a MATLAB-code for the numerical computations. Parts of this work have been done during a visit of the first author at the University of Applied Sciences Zittau/Görlitz (June – July 2000). The work was supported by DFG (Deutsche Forschungsgemeinschaft), grant 436 EST and by the Estonian Science Foundation, grant 4345.

REFERENCES

1. O. M. Alifanov, E. A. Artjukhin and S. V. Rumjancev, *Extremal methods for solving ill-posed problems and its applications to the inverse problems of the heat exchange*, Nauka, Moscow, 1988 (in Russian).

2. O. M. Alifanov and S. V. Rumjancev, *On the stability of iterative methods for the solution of linear ill-posed problems*, Soviet Math. Dokl. 20 (1979), 1133–1136.

3. H.W. Engl and H. Gfrerer, *A posteriori parameter choice for general regularization methods for solving linear ill-posed problems*, Appl. Numer. Math. 4 (1988), 395–417.

4. H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

5. V. M. Fridman, *The method of successive approximations for the Fredholm integral equation of the first kind*, Usp. Mat. Nauk 11 (1956), 233–234 (in Russian).

6. V. M. Fridman, *New methods of solving a linear operator equation*, Dokl. Akad. Nauk SSSR 128 (1959), 482–484 (in Russian).

7. V. M. Fridman, *On the convergence of methods of steepest descent type*, Usp. Mat. Nauk 17 (1962), 201–208 (in Russian).

8. H. Gfrerer, *An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates*, Math. Comp. 49 (1987), 507–522.

9. S. F. Gilyazov, *Methods for Solving Linear Ill-Posed Problems*, Moscow Univ., Moscow, 1987 (in Russian).

10. S. F. Gilyazov, *Approximate solution of ill-posed problems*, Moscow Univ., Moscow, 1995 (in Russian).

11. S. F. Gilyazov and N. L. Goldman, *Regularization of ill-posed problems by iteration methods*, Kluwer, Dordrecht, 2000.

12. V. I. Gordonova and V. A. Morozov, *Numerical parameter selection algorithms in the regularization methods*, USSR Comp. Math. Math. Phys. 13 (1974), 1–9.

13. U. Hämarik, *Monotonicity of error and choice of the stopping index in iterative regularization methods*. In: Differential and Integral Equations: Theory and Numerical Analysis, A. Pedas ed., Estonian Math. Society, Tartu, 1999, pp. 15–30.

14. U. Hämarik and T. Raus, *On the a posteriori parameter choice in regularization methods*, Proc. Estonian Acad. Sci. Phys. Math. 48 (1999), 133–145.

15. U. Hämarik and U. Tautenhahn, *The monotone error rule for parameter choice in regularization methods*. In: ENUMATH 99-Proc. of the 3rd European Conf. on Numer. Math. and Advanced Appl., P. Neittaanmäki, T. Tiihonen and P. Tarvainen eds., World Scientific, Singapore, 2000, pp. 518–525.

16. U. Hämarik and U. Tautenhahn, *On the monotone error rule for parameter choice in iterative and continuous regularization methods*, BIT 41 (2001), 1029–1038.

17. M. Hanke, Accelerated Landweber iterations for the solution of ill-posed equations, Numer. Math. 60 (1991), 341–373.

18. M. Hanke and C. W. Groetsch, Nonstationary iterated Tikhonov regularization, J. Optimization Theory Appl. 98 (1998), 37–53.

19. M. Hanke and P.C. Hansen, *Regularization methods for large-scale problems*, Surv. Math. Ind. 4 (1993), 253–315.

20. W. J. Kammerer and M. Z. Nashed, *Steepest descent for singular linear operators with nonclosed range*, Appl. Anal. 1 (1971), 143–159.

21. E. M. Kiss, *Ein Beitrag zur Regularisierung und Diskretisierung des inversen Problems der Identifikation des Memorykernes in der Viskoelastizität*, Dissertation Techn. Univ. Bergakademie Freiberg, Shaker, Aachen, 2000.

22. M. A. Krasnoselskij, G. M. Vainikko, P. P. Zabreiko, Ya. B. Rutitskij, V. Ya. Stetsenko, *Approximate solution of operator equations*, Wolters-Noordhoff, Groningen, 1972.

23. L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math. 73 (1951), 615–624.

24. L. J. Lardy, *A series representation for the generalized inverse of a closed linear operator*, Atti della Accademia Nazionale dei Lincei, Rendiconti della Classe di Science Fisiche Matematiche, e Naturali, serie VIII 58 (1975), 152–157.

25. M. Lavrentiev, *Some Improperly Posed Problems of Mathematical Physics*, Springer, New York, 1967.

26. A. K. L.Louis, *Inverse und schlecht gestellte Probleme*, Teubner, Stuttgart, 1989.

27. V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl. 7 (1966), 414–417.

28. V. A. Morozov, *On the regularizing families of operators*. In: Vychisl. Metody Program. 8, Moscow Univ., Moscow, 1967, pp. 63–95 (in Russian).

29. S. V. Pereverzev and E. Schock, *Brakhage's implicit iteration method and information complexity of equations with operators having closed range*, J. Complexity 15 (1999), 881–890.

30. T. Raus, *On the discrepancy principle for solution of ill-posed problems with non-selfadjoint operators*, Acta et Comment. Univ. Tartuensis 715 (1985), 12–20 (in Russian).

31. L. Sarv, *A certain family of nonlinear iterative methods for solving ill-posed problems*, Izv. Akad. Nauk Est. SSR, Fiz. Mat. 31 (1982), 261–269 (in Russian).

32. L. Sarv, *Iteration Methods for Linear Ill-Posed Problems*, Dissertation, Sverdlovsk, 1984 (in Russian).

33. V. I. Starostenko and S. M. Oganesjan, *Regularizing iterative processes.* In: Methods for solution of ill-posed problems and its applications, A. N. Tikhonov ed., Novosibirsk, 1982, pp. 119-128 (in Russian).

34. U. Tautenhahn, *On a new parameter choice rule for ill-posed inverse problems*, In: Hellenic-European Research on Computer Mathematics and its Applications, E.A. Lipitakis ed., Lea Publishers, Athens, 1999, pp. 118–125.

35. U. Tautenhahn and U. Hämarik, *The use of monotonicity for choosing the regularization parameter in ill-posed problems*, Inverse Problems 15 (1999), 1487–1505.

36. A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, Wiley, New York, 1977.

37. V. N. Trushnikov, *A non-linear regularizing algorithm and some of its applications*, USSR Comp. Math. Math. Phys. 19 (1979), 22–31.

38. G. M. Vainikko, *The discrepancy principle for a class of regularization methods*, USSR Comp. Math. Math. Phys. 22 (1982), 1–19.

39. G.M. Vainikko and A.Y. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Nauka, Moscow, 1986 (in Russian).