ON THE ANALOG OF THE MONOTONE ERROR RULE
FOR PARAMETER CHOICE IN THE (ITERATED)
LAVRENTIEV REGULARIZATION

U. HÄMARIK\textsuperscript{1}, T. RAUS\textsuperscript{1}, AND R. PALM\textsuperscript{2}

Abstract — We consider linear ill-posed problems in Hilbert spaces with a noisy right hand side and a given noise level. To solve non-self-adjoint problems by the (iterated) Tikhonov method, one effective rule for choosing the regularization parameter is the monotone error rule (Tautenhahn\&Hämarik, Inverse Problems, 1999, 15, 1487–1505). In this paper we consider the solution of self-adjoint problems by the (iterated) Lavrentiev method and propose for parameter choice an analog of the monotone error rule. We prove under certain mild assumptions the quasi-optimality of the proposed rule guaranteeing convergence and order optimal error estimates. Numerical examples show for the proposed rule and its modifications much better performance than for the modified discrepancy principle.

2000 Mathematics Subject Classification: 65J20, 47A52.

Keywords: ill-posed problem, regularization, iterated Lavrentiev method, regularization parameter choice, monotone error rule, quasi-optimality.

1. Introduction

We consider the operator equation

\[ Au = f, \quad f \in \mathcal{R}(A), \]  

where \( A \in \mathcal{L}(H, F) \) is a linear bounded operator, and \( H, F \) are Hilbert spaces with corresponding inner products \((\cdot, \cdot)\) and norms \( \| \cdot \| \). We do not suppose that \( \mathcal{R}(A) \) is closed and so in general our problem is ill-posed. The kernel \( \mathcal{N}(A) \) may be nontrivial. As usual in the treatment of ill-posed problems, we suppose that instead of the exact right hand side \( f \in F \) we have only an approximation \( f_\delta \in F \) with a given noise level \( \delta: \|f_\delta - f\| \leq \delta \) holds.

Standard regularization method for solving problem (1.1) is the Tikhonov method

\[ u_\alpha = (\alpha I + A^*A)^{-1}A^*f_\delta. \]  

The accuracy of this approximation may be increased by iteration. Let \( m \in \mathbb{N} \) be fixed and \( u_0 = u_{\alpha,0} \) be the initial approximation. The \( m \)-iterated Tikhonov approximation \( u_\alpha = u_{\alpha,m} \) is obtained by iterative computations

\[ u_{\alpha,i} = (\alpha I + A^*A)^{-1}(A^*f_\delta + \alpha u_{\alpha,i-1}), \quad (i = 1, \ldots, m). \]  

\textsuperscript{1}Institute of Mathematics, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia. E-mail: uno.hamarik@ut.ee and toomas.raus@ut.ee
\textsuperscript{2}Institute of Computer Science, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia. E-mail: reimo.palm@ut.ee
In the case that $m = 1$ and $u_0 = 0$, we return to the Tikhonov approximation. In this paper we consider the self-adjoint problem (1.1), assuming that $H = F$ and $A = A^* \geq 0$. In this case usually instead of the Tikhonov method (1.2) and the $m$-iterated Tikhonov method (1.3) the Lavrentiev method

$$u_\alpha = (\alpha I + A)^{-1} f_\delta$$

and the $m$-iterated Lavrentiev method $u_\alpha = u_{\alpha,m}$,

$$u_{\alpha,i} = (\alpha I + A)^{-1}(f_\delta + \alpha u_{\alpha,i-1}) \quad (i = 1, \ldots, m)$$

are used. An important problem is the choice of the regularization parameter $\alpha$, for which various a posteriori rules have been proposed. One effective rule for choosing $\alpha$ in methods (1.2), (1.3) is the monotone error rule (ME-rule) [6, 24, 26]. The name of this rule is justified by the property

$$\frac{d}{d\alpha} \|u_\alpha - u_*\| > 0 \text{ for all } \alpha \in (\alpha_{ME}, \infty),$$

where $u_*$ is the solution of (1.1), nearest to the initial approximation $u_0$. In this paper we propose and analyze a self-adjoint analog of the ME-rule. The paper is organized as follows. In Section 2, we give a short overview of the a posteriori rules in methods (1.2)–(1.5). In Section 3, we establish the monotonicity properties of the function $d_{ME}(\alpha)$ used in the proposed rule. In Section 4, we state the quasi-optimality property of our rule. In Section 5, we consider the extrapolation of the (iterated) Lavrentiev approximations which is used in later sections. Actually the main rule proposed does not satisfy condition (1.6) and in Section 6 we discuss the possibilities for better approximations of the genuine ME-rule for methods (1.4), (1.5); this permits some decrease in the error of approximation. The paper is concluded by numerical experiments in Section 7.

2. Rules for choosing the regularization parameter

In the prominent discrepancy principle [15, 27, 28] $\alpha_D = \alpha(\delta)$ is chosen from the equation $\|Au_\alpha - f_\delta\| \approx b\delta$, $b > 1$. For methods (1.2), (1.3) it guarantees convergence $\|u_{\alpha(\delta)} - u_*\| \to 0$ ($\delta \to 0$) and in the case of source-like solutions order-optimal error estimates

$$u_* - u_0 \in \mathcal{R}((A^*A)^{p/2}) \Rightarrow \|u_{\alpha(\delta)} - u_*\| \leq \text{const} \delta^s$$

with $p \leq 2m - 1$. The discrepancy principle leads to a divergence for method (1.4) and a convergence for method (1.5) with the error estimate (2.1) for source-like solutions with $p \leq m - 1$. Note that the discrepancy principle for the modifications of the Lavrentiev method in Hilbert scales was investigated in [16] and the rule $\|Au_\alpha - f_\delta\| = $ with $s \in (0, 1)$ in [14].

To consider the following rules for the parameter choice, we introduce the operator

$$B_\alpha = \begin{cases} \alpha^{1/2}(\alpha I + AA^*)^{-1/2} & \text{for methods (1.2), (1.3)}, \\ \alpha(\alpha I + A)^{-1} & \text{for methods (1.4), (1.5)}. \end{cases}$$

In the modified discrepancy principle $\alpha_{MD} = \alpha(\delta)$ is chosen for methods (1.2), (1.3) from equation [20] (see also [11, 25])

$$\|B_\alpha(Au_\alpha - f_\delta)\| \equiv (Au_\alpha - f_\delta, Au_{\alpha,m+1} - f_\delta)^{1/2} = b\delta, \quad b \geq 1,$$
and for methods (1.4), (1.5) from equation [19]
\[ \|B_{\alpha}(Au_{\alpha} - f_{\delta})\| \equiv \|Au_{\alpha,m+1} - f_{\delta}\| = b\delta, \quad b > 1. \] (2.2)

This rule guarantees convergence and order-optimal error estimates (2.1) for methods (1.2)–(1.5) for all values of \( p \leq 2m \) for methods (1.2), (1.3) and \( p \leq m \) for methods (1.4), (1.5).

For methods (1.2), (1.3) the modified discrepancy principle is outperformed by the monotone error rule (ME-rule) \([6, 24, 26]\) choosing \( \alpha_{ME} = \alpha(\delta) \) from the equation
\[ \frac{\|B_{\alpha}(Au_{\alpha} - f_{\delta})\|^2}{\|B_{\alpha}^2(Au_{\alpha} - f_{\delta})\|} \equiv \frac{(Au_{\alpha} - f_{\delta}, Au_{\alpha,m+1} - f_{\delta})}{\|Au_{\alpha,m+1} - f_{\delta}\|} = b\delta, \quad b \geq 1 \] (2.3)
and guaranteeing \( \|u_{\alpha,ME} - u_*\| \leq \|u_{\alpha,MD} - u_*\| \). In many other methods for the regularization of non-self-adjoint problems the ME-rule for parameter choice was studied in \([8, 9]\).

In this paper, we propose for methods (1.4), (1.5) a self-adjoint analog of rule (2.3), choosing \( \alpha_{ME} = \alpha(\delta) \) from the equation
\[ d_{ME}(\alpha) = \frac{\|B_{\alpha}(Au_{\alpha} - f_{\delta})\|^2}{\|B_{\alpha}^2(Au_{\alpha} - f_{\delta})\|} \equiv \frac{\|Au_{\alpha,m+1} - f_{\delta}\|^2}{\|Au_{\alpha,m+2} - f_{\delta}\|} = b\delta, \quad b > 1. \] (2.4)

As is well known (see, e.g., \([27, 28]\)), approximations (1.4), (1.5) have the form
\[ u_{\alpha} = u_0 + g_{\alpha,m}(A)(f_\delta - Au_0) \]
with the generating function
\[ g_{\alpha,m}(\lambda) = \sum_{j=1}^{m} \frac{\alpha^{j-1}}{(\alpha + \lambda)^j} = \frac{1}{\lambda} \left[ 1 - \left( \frac{\alpha}{\lambda + \alpha} \right)^m \right]. \] (2.5)

From (2.5) and the relations
\[ Au_{\alpha,i} - f_\delta = B_{\alpha}(Au_0 - f_\delta) \] (2.6)
it follows that
\[ \frac{du_{\alpha}}{d(\alpha^{-1})} = -mB_{\alpha}(Au_{\alpha} - f_{\delta}) \]
therefore
\[ \frac{du_{\alpha,m+1}}{d(\alpha^{-1})} = -(m + 1)B_{\alpha}(Au_{\alpha,m+1} - f_{\delta}) = -(m + 1)B_{\alpha}^2(Au_{\alpha} - f_{\delta}). \]

The last equality allows us to represent rule (2.4) in the form
\[ d_{ME}(\alpha) = \frac{m + 1}{m^2} \left\| \frac{du_{\alpha}}{d(\alpha^{-1})} \right\|^2 \left/ \left\| \frac{du_{\alpha,m+1}}{d(\alpha^{-1})} \right\| \right\| = b\delta, \quad b > 1. \] (2.7)

Let some sequence \( \alpha_i = \alpha_{i-1}q \) with \( q > 1 \) be given. Then we can approximate the derivatives as follows:
\[ \left\| \frac{du_{\alpha}}{d(\alpha^{-1})} \right\| \approx \|u_{\alpha+1} - u_{\alpha}\| \alpha_i \frac{1}{q - 1}, \quad \left\| \frac{du_{\alpha,m+1}}{d(\alpha^{-1})} \right\| \approx \|u_{\alpha+1,m+1} - u_{\alpha,m+1}\| \alpha_i \frac{1}{q - 1}. \]
Using the last formulas in (2.7), this rule takes the form

\[
\tilde{d}_{ME}(\alpha) = \frac{\|u_{\alpha,i} - u_\alpha\|^2}{\|u_{\alpha,i+1,m+1} - u_{\alpha,m+1}\|} \frac{\alpha_i \cdot m + 1}{q - 1} = b\delta, \quad b > 1. \tag{2.8}
\]

For the regularization parameter we take \( \alpha(\delta) = \alpha_i \).

Let us consider other choices of \( \alpha \) in methods (1.4), (1.5). In \([21,22]\) \( \alpha \) is chosen from the equation

\[
\alpha \|A^k(Au_{\alpha,m+k+1} - f_\delta)\| = b\delta, \quad b > 1, \quad 2k \in \mathbb{N}. \tag{2.9}
\]

Rule (2.9) guarantees convergence \( \|u_{\alpha} - u_*\| \to 0 \) (\( \delta \to 0 \)) also in the case where (instead of \( \|f_\delta - f\| \leq \delta \)) \( \|f_\delta - f\| \leq c\delta \) with unknown \( c = \text{const} \) holds. In the case \( \|f_\delta - f\| \leq \delta \) order-optimal error estimates (2.1) are stated with \( p \leq m \). Modifications of rule (2.9) were studied in [7]. In \([3,18,22]\) \( \alpha \) is chosen from the discrepancy of the extrapolated approximation

\[
\|A[(1 - q^{-m})^{-1}u_\alpha + (1 - q^m)^{-1}u_{\alpha,m}] - f_\delta\| = b\delta, \quad b > 1, \quad q \neq 1, \tag{2.10}
\]

and the order-optimal error estimates (2.1) for source-like solutions are stated with \( p \leq m \).

In \([22,23]\) for choosing \( \alpha \) differences of \( u_\alpha \) and \( u_{\alpha'} \) with some \( \alpha' \) are used. Namely, it is easy to check that in methods (1.4), (1.5)

\[
\|u_{\alpha/(1+\alpha)} - u_\alpha\|/m \leq \|B_\alpha(Au_\alpha - f_\delta)\| \leq \|u_\alpha - u_{\alpha/(1-\alpha)}\|/m \quad (\alpha < 1).
\]

Therefore the same results as for modified discrepancy principle hold for the analog of this principle, where \( \alpha = \alpha(\delta) \) is a solution of the equation

\[
\|u_\alpha - u_{\alpha/(1-\alpha)}\| = bm\delta, \quad b > 1. \tag{2.11}
\]

The parameter \( \alpha(\delta) \) from this rule satisfies \( \alpha(\delta) \in (\alpha_{MD}/(1 + \alpha_{MD}), \alpha_{MD}] \), where \( \alpha_{MD} \) is a solution of (2.2) with the same \( b \) as in (2.11).

Another analog [23] of the modified discrepancy principle uses the sequence \( \alpha_i = \alpha_{i-1}q \) with \( q > 1 \) as the popular balancing principle [12, 13, 17] and chooses \( \alpha(\delta) = \alpha_i \), where \( i \) is the first index, for which \( \|u_{\alpha_i} - u_{\alpha_{i-1}}\| > (q - 1)^{-1}m\alpha_i\delta \). Again, the same results as for the modified discrepancy principle hold.

Note also that in [1, 2] the Lavrentiev method \( u_\alpha = (\alpha L^s + A)^{-1}f_\delta \) with a strictly positive operator \( L : D(L) \to H \) and \( s \geq 0 \) in Hilbert scales was considered. Assuming \( c_1\|L^au\| \leq \|Au\| \leq c_2\|L^au\| \) for some positive \( a, c_1, c_2 \) and choosing \( \alpha \) from the equation

\[
\frac{\|B_\alpha^{3/2}D_s f_\delta\|^2}{\|B_\alpha^{3/2}D_s f_\delta\|^2} = \text{const} \delta, \quad D_s = (L^{-s/2}AL^{-s/2})^{-s/(2s+2a)}
\]

the error estimates (2.1) were stated. In the case \( s > 0 \), the operator \( D_s \) uses a negative power of \( A \), but the case \( s = 0 \) gives due to \( D_0 = I \) the practical rule (rule (2.3) with \( m = 1 \))

\[
\frac{(Au_{\alpha,1} - f_\delta, Au_{\alpha,2} - f_\delta)}{\|Au_{\alpha,2} - f_\delta\|} = \text{const} \delta. \tag{2.12}
\]
3. Monotonicity properties of the function $d_{ME}(\alpha)$

Theorem 3.1. The function $d_{ME}(\alpha)$ in (2.4) has the following properties: $d_{ME}(0) = \|Qf_\delta\|$, where $Q$ denotes the orthoprojection of $Y$ onto $\mathcal{N}(A) = \overline{\mathcal{R}(A)}^\perp$; $d_{ME}(\alpha)$ is monotonically increasing with $\lim_{\alpha \to \infty} d_{ME}(\alpha) = \|f_\delta\|$ and for $\delta \in (\|Qf_\delta\|, \|f_\delta\|)$ the equation $d_{ME}(\alpha) = \delta$ has a unique solution.

Proof. Due to the relation $Au_{\alpha,i} - f_\delta = -B^{i}_{\alpha} \hat{f}_\delta$ (see (2.6)) with $\hat{f}_\delta = f_\delta - Au_0$ ($i = 0, 1, \ldots$) the function $d_{ME}(\alpha)$ can be rewritten as

$$d_{ME}(\alpha) = \frac{\|B^{m+1}_{\alpha}\hat{f}_\delta\|^2}{\|B^{m+2}_{\alpha}\hat{f}_\delta\|}. \tag{3.1}$$

It is clear that $d_{ME}(\alpha)$ is continuous. From the inequality $\|B_{\alpha}\| \leq 1$ follows

$$\|B^{m+1}_{\alpha}\hat{f}_\delta\| \|B^{m+2}_{\alpha}\hat{f}_\delta\| \leq \|B^{m+1}_{\alpha}\hat{f}_\delta\|^2 = (B^{m+1}_{\alpha}\hat{f}_\delta, B^{m+2}_{\alpha}\hat{f}_\delta) \leq \|B^{m}_{\alpha}\hat{f}_\delta\| \|B^{m+2}_{\alpha}\hat{f}_\delta\|,$$

therefore,

$$\|B^{m+1}_{\alpha}\hat{f}_\delta\| \leq d_{ME}(\alpha) \leq \|B^{m}_{\alpha}\hat{f}_\delta\|. \tag{3.2}$$

Due to the relations $Au_{\alpha,i} - f_\delta = -B^{i}_{\alpha} \hat{f}_\delta$ and the limits $\lim_{\alpha \to 0} \|Au_{\alpha,i} - f_\delta\| = \|f_\delta\|$, $\lim_{\alpha \to \infty} \|Au_{\alpha,i} - f_\delta\| = \|Qf_\delta\|$ for all $i \in \mathbb{N}$ (cf. [27, 28]) the function $d_{ME}(\alpha)$ has the same limits.

Let $\{F_\lambda\}$ denote the spectral family of the operator $A$. Then we obtain for arbitrary $i \in \mathbb{R}$

$$\frac{d}{d\alpha}\|B^{i}_{\alpha}\hat{f}_\delta\|^2 = \frac{d}{d\alpha}(\|\alpha^{-1}A + I\|^{-2}T^i\hat{f}_\delta, \hat{f}_\delta) = \frac{d}{d\alpha} \int_{0}^{\|A\|} (\alpha^{-1}\lambda + 1)^{-2i}d\|F_\lambda\hat{f}_\delta\|^2 =$$

$$\frac{2i}{\alpha^2} \int_{0}^{\|A\|} \lambda(\alpha^{-1}\lambda + 1)^{-2i-1}d\|F_\lambda\hat{f}_\delta\|^2 = \frac{2i}{\alpha^2}(A(\alpha^{-1}A + I)^{-2i-1}T^i\hat{f}_\delta, \hat{f}_\delta) = \frac{2i}{\alpha^2}\|A^{1/2}B^{i+1/2}_{\alpha}\hat{f}_\delta\|^2. \tag{3.3}$$

Using this result for $i = m + 1$ and $i = m + 2$, respectively, the equality

$$\frac{d}{d\alpha}\|B^{m+2}_{\alpha}\hat{f}_\delta\| = \frac{d}{d\alpha}\left[\|B^{m+2}_{\alpha}\hat{f}_\delta\|^2\right]^{1/2} = \left(\frac{1}{2} \frac{d}{d\alpha}\|B^{m+2}_{\alpha}\hat{f}_\delta\|^2\right) / \|B^{m+2}_{\alpha}\hat{f}_\delta\|$$

and the quotient rule, we obtain

$$d'_{ME}(\alpha) = \frac{2(m + 1)\|A^{1/2}B^{m+3/2}_{\alpha}\hat{f}_\delta\|^2\|B^{m+2}_{\alpha}\hat{f}_\delta\|^2 - (m + 2)\|A^{1/2}B^{m+5/2}_{\alpha}\hat{f}_\delta\|^2\|B^{m+1}_{\alpha}\hat{f}_\delta\|^2}{\alpha^2\|B^{m+2}_{\alpha}\hat{f}_\delta\|^3}.$$

We use the relation

$$\|A^{1/2}B^{i}_{\alpha}\hat{f}_\delta\|^2 = \alpha\|B^{i-1/2}_{\alpha}\hat{f}_\delta\|^2 - \alpha\|B^{i}_{\alpha}\hat{f}_\delta\|^2$$

with $i = m + 3/2$ and $i = m + 5/2$, respectively, the notation $z = B^{m+1}_{\alpha}\hat{f}_\delta$ and obtain

$$d'_{ME}(\alpha) = \frac{m\|A^{1/2}B^{1/2}_{\alpha}z\|^2\|B_{\alpha}z\|^2}{\alpha^2\|B_{\alpha}z\|^3} + (m + 2)\|z\|^2\|B^{3/2}_{\alpha}z\|^2 - \alpha\|B_{\alpha}z\|^2\|B^{1/2}_{\alpha}z\|^2.$$
Both summands here are nonnegative. The nonnegativity of the second term can be stated by multiplying the inequalities

\[ \|B_{\alpha} z\| \leq \|B_{\alpha}^{3/2} z\|^{2/3} \|z\|^{{1/3}}, \quad \|B_{\alpha}^{1/2} z\| \leq \|B_{\alpha}^{3/2} z\|^{1/3} \|z\|^{{2/3}}, \]

which can be obtained from the inequality of moments (see [28])

\[ \|W^{p} z\| \leq \|W^{q} z\|^{p/q} \|z\|^{-p/q} \quad (W \in \mathcal{L}(H, H), \ W = W^{*} \geq 0, \ z \in H, \ p \leq q) \quad (3.4) \]

with \( W = B_{\alpha}, \ q = 3/2 \) and \( p = 1 \) or \( p = 1/2 \), respectively. Thus, the function \( d_{\alpha}(\alpha) \) is monotonically increasing and the equation \( d_{\alpha}(\alpha) = \delta \) has a unique solution for \( \delta \in (\|Qf_{\delta}\|, \|f_{\delta}\|) \).

4. On the quasi-optimality of the ME-rule

To characterize the quality of the ME-rule, we use in the following the property of weak quasi-optimality (see [22]). We say that the rule \( R \) for the a posteriori choice of the regularization parameter \( \alpha = \alpha(R) \) is weakly quasi-optimal if there exists a constant \( C \) (which does not depend on \( A, u_{\ast}, f_{\delta} \)) such that for each \( f_{\delta}, \|f_{\delta} - f\| \leq \delta \) the error estimate

\[ \|u_{\alpha}(\delta) - u_{\ast}\| \leq C \inf_{\alpha > 0} \Psi(\alpha) + O(\delta) \]

holds. Here the function

\[ \Psi(\alpha) = \|u_{\alpha} - u_{\ast}^{0}\| + m\alpha^{-1}\delta = \|B_{\alpha}^{m}(u_{\alpha} - u_{\ast})\| + m\alpha^{-1}\delta \]

is the upper bound of the error \( \|u_{\alpha} - u_{\ast}\| \) and \( u_{\ast}^{0} \) is the approximation in the (iterated) Lavrentiev method with \( f \) instead of \( f_{\delta} \). It was shown in [22] that if the rule \( R \) is weakly quasi-optimal and the method \( P \) with the a priori parameter choice is order-optimal on some set of solutions \( M \), then the method \( P \) with a parameter choice by the Rule \( R \) is also order-optimal on the set \( M \). For example, if the rule \( R \) is weakly quasi-optimal, then the regularization method with a parameter choice by the rule \( R \) is also order-optimal on the set \( M_{\rho_{0}} = \{ u \in H : u - u_{0} = A^{p}v, \|v\| \leq \rho, p > 0 \} \) for all values of \( p \in (0, \rho_{0}] \) Here \( \rho_{0} \) is the qualification of the method. The qualification of the \( m \) times iterated Lavrentiev method is \( m \).

We introduce the notations

\[ x_{\alpha} := 1 + \frac{\alpha}{\|A\|}, \quad \bar{u} := u_{0} - u_{\ast}, \quad \bar{f} := f_{\delta} - f, \quad t_{c}(\alpha) := \|B_{\alpha}^{m}\bar{u}\|^{2} + c\|g_{\alpha,m}(A)\bar{f}\|^{2} \]

with constant \( c > 0 \). For investigating the quasi-optimality of rule (2.4), we need the monotonicity property of the function \( t_{c}(\alpha) \). For error \( u_{\alpha} - u_{\ast} = B_{\alpha}^{m}\bar{u} + g_{\alpha,m}(A)\bar{f} \) we have \( \|u_{\alpha} - u_{\ast}\|^{2} \leq 2t_{1}(\alpha) \), therefore the monotonicity properties of function \( t_{c}(\alpha) \) are of interest.

**Lemma 4.1.** From the inequality

\[ x_{\alpha}^{1/2}\|B_{\alpha}^{m+1} A\bar{u}\| \geq \sqrt{cm}\|B_{\alpha}^{1+m/2}\bar{f}\|, \quad \alpha \in [\alpha_{0}, \alpha_{1}] \quad (4.1) \]

it follows that the function \( t_{c}(\alpha) \) is monotonically increasing in the interval \([\alpha_{0}, \alpha_{1}]\).
Therefore in case (4.1) the function \( g_{\alpha,m}(\lambda) \leq \frac{m}{\alpha} \frac{\alpha}{\alpha + \lambda} \) (see (2.5)) give

\[
\frac{1}{2} \frac{d}{d\alpha} \| g_{\alpha,m}(A) \bar{f} \|^2 = \left( g_{\alpha,m}(A) \bar{f}, \frac{dg_{\alpha,m}(A) \bar{f}}{d\alpha} \right) = - \frac{m}{\alpha^2} (g_{\alpha,m}(A) \bar{f}, B_{\alpha}^{m+1} \bar{f}) = - \frac{m}{\alpha^2} \| g_{\alpha,m}(A) B_{\alpha}^{(m+1)/2} \bar{f} \|^2 \geq - \frac{m}{\alpha^2} \| B_{\alpha}^{1+m/2} \bar{f} \|^2.
\]

This inequality and equality (3.3) with \( i = m \) and with \( \bar{u} \) instead of \( f_\delta \) give

\[
\frac{1}{2} t'_c(\alpha) \geq \frac{m}{\alpha^2} \left[ \| A^{1/2} B_{\alpha}^{m+1/2} \bar{u} \|^2 - \frac{cm}{\alpha} \| B_{\alpha}^{1+m/2} \bar{f} \|^2 \right].
\]

Since \( \| \alpha^{-1} AB_{\alpha} \| \leq \kappa_{\alpha} \), we have

\[
t'_c(\alpha) \geq 2m \alpha^{-2} \left[ \kappa_{\alpha} \| (\alpha^{-1} AB_{\alpha})^{1/2} A^{1/2} B_{\alpha}^{m+1/2} \bar{u} \|^2 - cm \alpha^{-1} \| B_{\alpha}^{1+m/2} \bar{f} \|^2 \right] = 2m \alpha^{-3} \left[ \kappa_{\alpha} \| AB_{\alpha}^{m+1} \bar{u} \|^2 - cm \| B_{\alpha}^{m+1} \bar{f} \|^2 \right].
\]

Therefore in case (4.1) the function \( t_c(\alpha) \) monotonically increases in the interval \([\alpha_0, \alpha_1]\). \( \Box \)

In the following, we need the condition

\[
\| B_{\alpha}^2 (Au_{\alpha} - f_\delta) \| \equiv \| B_{\alpha}^{m+2} \bar{A} \bar{u} - B_{\alpha}^{m+2} \bar{f} \| \geq \| B_{\alpha}^{m+2} \bar{f} \| \quad \forall \alpha \geq \alpha(\delta). \tag{4.3}
\]

This condition is satisfied in special cases if the error of the right hand side \( \bar{f} = f_\delta - f \) is such that \( (F_{\lambda} \bar{u}, \bar{f}) \leq 0, \forall \lambda \geq 0 \), or if \( \bar{f} \in N(A) \). Numerical experiments show that for most severely ill-posed problems in the case of a uniform distribution of the error \( \bar{f} \) the condition

\[
\| B_{\alpha}^2 (Au_{\alpha} - f_\delta) \| \geq c_* \| B_{\alpha}^{m+1} \bar{f} \|
\]

is fulfilled with constant \( c_* \approx 1 \).

**Theorem 4.1.** If condition (4.3) holds and for \( \alpha = \alpha(\delta) \) the equality (2.4) holds with \( b > 2 \), then the function \( t_c(\alpha) \) with \( c = \frac{1}{m} \left( \frac{b}{2} - 1 \right) \) is monotonically increasing for \( \alpha \geq \alpha(\delta) \).

**Proof.** From the equality

\[
B_{\alpha} (Au_{\alpha} - f_\delta) = B_{\alpha}^{m+1} \bar{A} \bar{u} - B_{\alpha}^{m+1} \bar{f}
\]

follows

\[
\| B_{\alpha} (Au_{\alpha} - f_\delta) \|^2 \leq 2(\| B_{\alpha}^{m+1} \bar{A} \bar{u} \|^2 + \| B_{\alpha}^{m+1} \bar{f} \|^2)
\]

This inequality together with condition (4.3) gives

\[
d_{ME}(\alpha) = \frac{\| B_{\alpha} (Au_{\alpha} - f_\delta) \|^2}{\| B_{\alpha}^2 (Au_{\alpha} - f_\delta) \|} \leq \frac{2(\| B_{\alpha}^{m+1} \bar{A} \bar{u} \|^2 + \| B_{\alpha}^{m+1} \bar{f} \|^2)}{\| B_{\alpha}^{m+2} \bar{f} \|}. \tag{4.6}
\]

Assume now that \( \alpha \geq \alpha(\delta) \). Then according to Theorem 3.1 \( d_{ME}(\alpha) \geq b \delta \) and from (4.6) we conclude that

\[
\| B_{\alpha}^{m+1} \bar{A} \bar{u} \|^2 \geq \frac{b}{2} \delta \| B_{\alpha}^{m+2} \bar{f} \| - \| B_{\alpha}^{m+1} \bar{f} \|^2.
\]
To continue this estimation, we use the inequalities
\[ \|B\| \leq 1, \quad \|B^{m+1}\| \leq \|B^{m/2}\| \|B^{m/2+1}\| \leq \|B^{m/2+1}\|, \]
\[ \|B^{m/2+1}\| = (B^{m/2+1}, B^{m/2+1}) = (B^{m/2+1}, B^{m/2+1}) \leq \|B^{m/2+1}\| \delta, \]
and get
\[ \|B^{m/2+1}A\|^2 \geq b \|B^{m/2+1}\|^2, \quad \|B^{m/2+1}\|^2 = \left( \frac{b}{2} - 1 \right) \|B^{m/2+1}\|^2 \]
\[ \|B\|^2 \geq \|B\|^2 = \left( \frac{b}{2} - 1 \right) \|B\|^2 + \delta. \]

The monotonical increase in the function \( t_c(\alpha) \) with \( c = \frac{1}{m} \left( \frac{b}{2} - 1 \right) \) follows now from Lemma 4.1 and from the monotonical increase in function \( d_{ME}(\alpha) \). \( \square \)

**Remark 4.1.** The condition \( b > 2 \) in the formulation of Theorem 4.1 is due to the rough estimate (4.5). Typically \( \|B\| \approx \|B^{m+1}\| \approx \|B^{m/2}\| \|B^{m/2+1}\| \leq \|B^{m/2+1}\| \delta \) and for practice we suggest to use \( b > 1 \).

For further studying the quasi-optimality property of rule (2.4), we need the following auxiliary result, which is a self-adjoint analog of Lemma 3 in [22].

**Lemma 4.2.** Let \( \alpha(\delta) \) be such a parameter that for each \( \alpha \leq \alpha(\delta) \) the inequality
\[ \|B\| \leq \alpha \] holds. Then \( \alpha(\delta) \leq \alpha_0 \), where \( \alpha_0 \) is a parameter, for which the function
\[ \varrho(\alpha) = \|B\| + m \delta \]
has a global minimum.

*Proof.* Using equality (3.3) with \( i = m \) and with \( w \) instead of \( f_\delta \), we have
\[ \varrho(\alpha) = \frac{1}{2} d \left( \|B\|^2 \right) \|B\|^2 - m \frac{\alpha}{\alpha^2} c \delta \]
\[ \|A^2 \| \|B\|^2 \|B\|^2 \|B\|^2 - m \frac{\alpha}{\alpha^2} c \delta. \] (4.7)
If \( \lim_{\alpha \to \infty} \varrho(\alpha) = \varrho(\alpha') \) for all \( \alpha' < \infty \), the assertion of lemma holds. Otherwise since \( \varrho(\alpha) \to \infty \) as \( \alpha \to 0 \), there is a minimum point \( \alpha_0 \) in the interval \((0, \infty)\). Then \( \varrho'(\alpha_0) = 0 \) and for \( \alpha = \alpha_0 \) both terms in the difference of formula (4.7) are equal, therefore,
\[ c \delta = \|A^2 \|^2 \|B\|^2 \|B\|^2. \] (4.8)
Using the inequality of moments (3.4) with \( W = B_{\alpha 0} A, z = B_{\alpha 0}^m w, p = 1/2, q = 1 \), we have
\[ \|A^2 \| \|B\|^2 \|B\|^2 \|B\|^2 \leq \|A B_{\alpha 0}^m w\|^2 \|B\|^2 \|B\|^2. \] Using this estimate for the first term in (4.8), we have \( c \delta \leq \|A B_{\alpha 0}^m w\| \). Taking into account the equality \( B_{\alpha 0}^m A = A B_{\alpha 0}^m \) and comparing the last inequality with the assumption of Lemma 4.2, from monotonic increase in the function \( s(\alpha) = \|A B_{\alpha 0}^m w\| \) the assertion \( \alpha(\delta) \geq \alpha_0 \) follows. \( \square \)
Theorem 4.2. Let condition (4.3) hold and let the parameter $\alpha = \alpha(\delta)$ be chosen from equation (2.4) with $b > 2$. Then the error estimate (weak quasioptimality)

$$
\|u_{\alpha(\delta)} - u_*\| \leq \left[b + 1 + \sqrt{\max \left(\frac{m}{b/2 - 1}, \frac{b/2 - 1}{m}\right)}\right] \inf_{\alpha \geq 0} \Psi(\alpha)
$$

holds.

Proof. We denote by $\alpha_*$ the global minimum point of the function $t_c(\alpha)$ with $c = \frac{1}{m}(\frac{b}{2} - 1)$. According to Theorem 4.1 $\alpha(\delta) \geq \alpha_*$. From the representation $u_\alpha - u_* = B_{\alpha}^m \tilde{u} + g_{\alpha,m}(A) \bar{f}$ follows

$$
\|u_{\alpha(\delta)} - u_*\| \leq \|B_{\alpha(\delta)}^m \tilde{u}\| + \|g_{\alpha,m}(A) \bar{f}\|.
$$

(4.10)

Denoting $c = \min(c, 1)$ and $\bar{c} = \max(c, 1)$ we have

$$
\bar{c}^{-1} t_c(\alpha) \leq \|B_{\alpha}^m \tilde{u}\|^2 + \|g_{\alpha,m}(A) \bar{f}\|^2 \leq \left[\|B_{\alpha}^m \tilde{u}\| + \|g_{\alpha,m}(A) \bar{f}\|\right]^2 \leq \Psi(\alpha)^2,
$$

therefore,

$$
g_{\alpha(\delta),m}(A) \bar{f} \leq \|g_{\alpha,m}(A) \bar{f}\| \leq \sqrt{\frac{t_c(\alpha_*)}{c}} \leq \sqrt{\frac{\bar{c}}{c}} \inf_{\alpha \geq 0} \Psi(\alpha). \quad (4.11)
$$

To estimate the first summand in (4.10), we use the inequalities (see (3.2))

$$
\|B_{\alpha}(Au_{\alpha} - f_\delta)\| \leq d_{ME}(\alpha), \quad d_{ME}(\alpha(\delta)) \leq b\delta, \quad \|B_{\alpha}\| \leq 1, \quad \|\bar{f}\| \leq \delta
$$

and from (4.4) we get

$$
\|B_{\alpha(\delta)}^{m+1} A\tilde{u}\| \leq \|B_{\alpha(\delta)}(Au_{\alpha(\delta)} - f_\delta)\| + \|B_{\alpha(\delta)}^{m+1} \bar{f}\| \leq b\delta + \|B_{\alpha(\delta)}^{m+1}\| \|\bar{f}\| \leq (b + 1)\delta.
$$

The last relation allows us to use Lemma 4.2 with $c = b + 1$, $w = \tilde{u}$ and we get

$$
\|B_{\alpha(\delta)}^m \tilde{u}\| \leq \|B_{\alpha}^m \tilde{u}\| \leq \|B_{\alpha}^m \tilde{u}\| + m(b + 1)\frac{\delta}{\alpha_0} = \inf_{\alpha \geq 0} \left[\|B_{\alpha}^m \tilde{u}\| + m(b + 1)\frac{\delta}{\alpha}\right] \leq (b + 1)\inf_{\alpha \geq 0} \Psi(\alpha).
$$

Summation of last inequality and (4.11) gives asserted estimate (4.9).

Theorem 4.3. Let condition (4.3) hold and let the parameter $\alpha = \alpha(\delta)$ be chosen from equation (2.4) with $b > 2$. Then $\|u_{\alpha(\delta)} - u_*\| \to 0$ as $\delta \to 0$ and in the case of a source-like solution, the order-optimal error estimate (2.1) holds with $p \leq m$.

Proof. The convergence $\alpha(\delta)$ of $\|u_{\alpha(\delta)} - u_*\| \to 0$ is guaranteed for such a priori parameter choice that $\alpha(\delta) \to 0$ (\(\delta \to 0\)) and $\frac{\delta}{\alpha(\delta)} \to 0$ (cf [27, 28]). In the case of a source-like solution, the a priori choice $\alpha(\delta) = \delta^{1/(p+1)}$ guarantees the order-optimal error estimate (2.1) with $p \leq m$ (cf [27, 28]). The quasi-optimality property of rule (2.4) in Theorem 4.2 guarantees for the (iterated) Lavrentiev approximation the same order of accuracy as for the best a priori parameter choice.

The quasi-optimality of the ME-rule (2.4) is proved under the special condition (4.3) for $\bar{f}$, but remains an open problem in the general case. Extensive numerical experiments have shown that with a high probability the quasi-optimality holds also in the general case.

In the following we will also use extrapolation which is useful for the parameter choice in the (iterated) Lavrentiev method or for the construction of alternative approximate solutions.
5. On the extrapolated Lavrentiev method

Let \( u_{\alpha_i}, i = 1, \ldots, n \) be approximations of Lavrentiev methods with different parameters \( \alpha_i = q_i \alpha \) (\( q_j \neq q_j \), if \( i \neq j \)). The corresponding extrapolated approximation \( v_{n,\alpha} \) has the form

\[
v_{n,\alpha} = \sum_{i=1}^{n} d_i u_{\alpha_i}, \quad d_i = \prod_{j=1, j \neq i}^{n} (1 - \alpha_i/\alpha_j)^{-1}.
\]  

(5.1)

As shown in [4], the extrapolated Lavrentiev approximation (5.1) coincides with the corresponding approximations

\[
u_n = (\alpha_n I + A)^{-1}(\alpha_n u_{n-1} + f_\delta), \quad (n = 1, 2, \ldots, u_0 = 0)
\]

of nonstationary implicit iterative methods. In the following we refer to some results from papers [4, 5].

In \( v_{n,\alpha} \) both indexes can be viewed as regularization parameters.

1) Let the sequence \( \alpha_1, \alpha_2, \ldots \) be given (\( \alpha \) is fixed) and consider the choice of \( n \) in the extrapolated Lavrentiev approximation \( v_{n,\alpha} \). Then we recommend to choose \( n \) by the discrepancy principle: \( n = n(\delta) \) is the first \( n \) with \( \|Av_{n,\alpha} - f_\delta\| \leq b\delta \), \( b \geq 1 \). It guarantees the convergence \( \|v_{n,\alpha} - u_\ast\| \to 0 \) (\( \delta \to 0 \)) and for source-like solutions the error estimate (2.1) holds for all \( p > 0 \).

2) Let \( n \geq 2 \) and \( q_1, \ldots, q_{n+1} \) be fixed. Consider the choice of \( \alpha \).

**Theorem 5.1** [4]. The function \( d_P(\alpha) = \|Av_{n,\alpha} - f_\delta\| \) is monotonically decreasing. If \( \alpha \) is chosen from the discrepancy principle \( d_P(n) = b\delta, \ b > 1, \) then \( \|v_{n,\alpha} - u_\ast\| \to 0 \) (\( \delta \to 0 \)) and for source-like solutions the error estimate (2.1) holds with \( p \leq n - 1 \). If \( \alpha \) is chosen from the modified discrepancy principle \( \|Av_{n+1,\alpha} - f_\delta\| = b\delta, \ b > 1, \) then \( \|v_{n,\alpha} - u_\ast\| \to 0 \) (\( \delta \to 0 \)) and for source-like solutions the error estimate (2.1) holds with \( p \leq n \).

Consider now the extrapolation of the Lavrentiev method iterated \( m \) times. For different \( \alpha_i = q_i \alpha \) (\( i = 1, \ldots, n \)) a different number of iterations \( m_1, \ldots, m_n \) may be used. We take for the approximate solution

\[
v_{n,\alpha} = \sum_{i=1}^{n} \sum_{k=1}^{m_i} d_{i,k} u_{k,\alpha_i},
\]

(5.2)

where the coefficients \( d_{i,k} \) can be uniquely determined from the relation

\[
\sum_{i=1}^{n} \sum_{k=1}^{m_i} d_{i,k}(1 + \lambda/q_i)^{-k} = \prod_{i=1}^{n} (1 + \lambda/q_i)^{-m_i} \quad (\forall \lambda \in \mathbb{R}).
\]

**Theorem 5.2** [4]. If \( n \) and \( q_1, \ldots, q_n \) are fixed and \( \alpha \) is chosen from the discrepancy principle \( d_P(n) = C\delta, \) then \( \|v_{n,\alpha} - u_\ast\| \to 0 \) (\( \delta \to 0 \)) and for source-like solutions the error estimate (2.1) holds with \( p \leq m_1 + m_2 + \cdots + m_n - 1 \).

6. On the approximation of the monotone error rule

As opposed to the ME-rule (2.3) for methods (1.2), (1.3), rule (2.4) for methods (1.4, 1.5) does not guarantee for the error the monotonicity property (1.6). Let us consider the question,
for what rule this property can be fulfilled at least approximately. For simplicity we take now \( u_0 = 0 \). From (4.2) with \( f_\delta \) instead of \( \hat{f} \) follows

\[
\frac{1}{2} \frac{d}{d\alpha} \|u_\alpha - u_*\|^2 = \left( u_\alpha - u_*, \frac{d}{d\alpha} u_\alpha \right) = (u_\alpha - u_*, -\frac{m}{\alpha^2} B^{m+1}_\alpha f_\delta). \tag{6.1}
\]

In the case of \( f_\delta \in \mathcal{R}(A) \), the equation \( Av = f_\delta \) has solution \( v \) and due to the relations \( f_\delta = Av, B^{m+1}_\alpha A = AB^{m+1}_\alpha, f_\delta - Au_\alpha = B^{m}_\alpha f_\delta, \|f_\delta - f\| \leq \delta \) relation (6.1) can be continued as

\[
\frac{\alpha^2}{2m} \frac{d}{d\alpha} \|u_\alpha - u_*\|^2 = (u_\alpha - u_*, B^{m+1}_\alpha A v) = (A(u_\alpha - u_\alpha), B^{m+1}_\alpha v) =
\]

\[
(f_\delta - Au_\alpha + f - f_\delta, B^{m+1}_\alpha v) \geq (B^{m}_\alpha f_\delta, B^{m+1}_\alpha v) - \delta \|B^{m+1}_\alpha v\|. \tag{6.2}
\]

The monotonicity property (1.6) for \( \alpha \geq \alpha_{ME} \) holds if \( \bar{\alpha}_{ME} \) is chosen from the equation

\[
\tilde{d}_{ME}(\alpha) \equiv \frac{(B^{m}_\alpha f_\delta, B^{m+1}_\alpha v)}{\|B^{m+1}_\alpha v\|} = b\delta, \quad b \geq 1,
\]

(6.3)
since one can show a monotonical increase in the function \( \tilde{d}_{ME}(\alpha) \) (proof is analogous to the proof of monotonicity of the function \( d_{ME}(\alpha) \) for rule (2.4) in Section 3). The best constant here and in rule (2.3) is \( b = 1 \). The last rule is not practical, while in the case of \( f_\delta \notin \mathcal{R}(A) \), \( v \) does not exist and if \( f_\delta \notin \mathcal{R}(A) \), then \( v \) is unknown. Substituting \( \nu \) in (6.3) by different approximate solutions of the equation \( Av = f_\delta \), we get different approximations of the monotone error rule (6.3). Substitution of \( \nu \) in (6.3) by the Lavrentiev approximation \( \alpha^{-1}B_\alpha f_\delta \) gives rule (2.4). Replacing \( \nu \) in (6.3) by a more accurate approximate solution of the equation \( Av = f_\delta \) we get better approximations of rule (6.3), increasing the probability of the monotonicity property (1.6).

If we use the approximation of rule (6.3) for the parameter choice in the (iterated) Lavrentiev method for the basic equation \( Au = f \), then finding approximate \( v \) of the problem \( Av = f_\delta \) is a subproblem. This subproblem differs from the basic problem in that here \( f_\delta \) are exact data unlike the basic problem where \( f_\delta \) are noisy data. Therefore, to solve the problem \( Av = f_\delta \) by the Lavrentiev method, one can recommend to use a smaller regularization parameter \( \beta \) than \( \alpha \) used in the (iterated) Lavrentiev method for the basic equation \( Au = f \). If we take \( \beta = \nu \alpha \) with \( \nu \leq 1 \) and substitute \( \nu \) in (6.3) by the Lavrentiev approximation \( (\nu \alpha)^{-1}B_{\nu \alpha} f_\delta \), then this rule will take the form

\[
\tilde{d}_{ME}(\alpha) \equiv \frac{(B^{m}_\alpha f_\delta, B^{m}_{\nu \alpha} f_\delta)}{\|B^{m+1}_{\nu \alpha} f_\delta\|} = \frac{(B^{2m+1}_\alpha f_\delta, B^{m+1}_\nu f_\delta)}{\|B^{m+1}_{\nu} f_\delta\|} = b\delta, \quad b \geq 1, \quad \nu \leq 1. \tag{6.4}
\]

For \( \nu = 1 \) this rule coincides with rule (2.4). If \( f_\delta \in \mathcal{R}(A) \), then \( (\nu \alpha)^{-1}B_{\nu \alpha} f_\delta \rightarrow v \) as \( \nu \alpha \rightarrow 0 \) and rule (6.4) approximates rule (6.3) in the process \( \nu \rightarrow 0 \). However, very small values of \( \nu \) cause numerical unstability and the numerical experiments suggest to use \( \nu \geq 10^{-3} \). In numerical examples of the following section, we got the best results with \( \nu = 0.17 \) for the nonsmooth solutions and with \( \nu = 0.4 \) for the source-like solutions (2.1) with \( p = 1 \).

Considerations analogous to those used for the modification rule (2.4) via (6.3) to rule (6.4) may be applied for the modification of rule (2.8). Namely, the difference between rules (6.4) and (2.4) is that in rule (2.4) all iterations are performed with the same parameter \( \alpha \), and in rule (6.4) the last iteration in the numerator and denominator uses a smaller parameter \( \nu \alpha \). In rule (2.8) replacing \( \alpha_i \) by a smaller \( \alpha \) corresponds to replacing the index \( i \) by \( i + l \).
with some \( l \in \mathbb{N} \). Such replacement in the denominator and the replacement of one side of the scalar product \((u_{i+1} - u_i, u_{i+1} - u_i)\) in the numerator by \(u_{i+1} - u_i\) gives the rule

\[
\bar{d}^{(l)}_{\text{ME}}(\alpha) = \frac{(u_{\alpha,i} - u_{\alpha,i+1} - u_{\alpha,i+1} - u_{\alpha,i})}{\|u_{\alpha,i+1,m+1} - u_{\alpha,i+1,m+1}\|} \quad \alpha, \quad m + 1 \quad q - 1 \quad m^2 = b \delta, \quad b > 1, \quad l \in \mathbb{N}. \tag{6.5}
\]

We made numerical experiments with \( q = 1.2 \) and than as it appeared the optimal \( l \) was \( l = 5 \) for nonsmooth solutions and \( l = 4 \) for source-like solutions (2.1) with \( p = 1 \).

Alternatively, the element \( v \) in (6.3) can be approximated by the iterated Lavrentiev approximation (1.5) \( u_{\alpha,k} = g_{\alpha,k}(A) f_{\delta} = \alpha^{-1} \sum_{j=1}^{k} B^j_{\alpha} f_{\delta} \) (see (2.5)), then rule (6.3) takes the form

\[
\bar{d}^{(k)}_{\text{ME}}(\alpha) \equiv \frac{(B^m_{\alpha} f_{\delta}, \sum_{n=m+2}^{m+k+1} B^n_{\alpha} f_{\delta})}{\|\sum_{n=m+2}^{m+k+1} B^n_{\alpha} f_{\delta}\|} \equiv \frac{(\alpha u_{\alpha} - f_{\delta}, \sum_{n=m+2}^{m+k+1} (A u_{\alpha,n} - f_{\delta}))}{\|\sum_{n=m+2}^{m+k+1} (A u_{\alpha,n} - f_{\delta})\|} = b \delta, \quad b \geq 1, \quad k \in \mathbb{N}. \tag{6.6}
\]

For \( k = 1 \) this rule coincides with rule (2.4). If \( f_{\delta} \in \mathcal{R}(A) \), then from the limit \( u_{\alpha,k} \rightarrow v \) \((k \rightarrow \infty)\) follows \( \bar{d}^{(k)}_{\text{ME}}(\alpha) \rightarrow \bar{d}_{\text{ME}}(\alpha) \), therefore rules (6.6) can be considered as approximations of ME-rule (6.3). Substituting in (6.2) \( v \) by \( u_{\alpha,\infty} = \alpha^{-1} \sum_{j=1}^{\infty} B^j_{\alpha} f_{\delta} \) we get

\[
\frac{\alpha^2}{2m} \frac{d}{d\alpha} \|u_{\alpha} - u\|^2 = \left( B^m_{\alpha} f_{\delta} + (f - f_{\delta}), \alpha^{-1} \sum_{j=1}^{\infty} B^j_{\alpha} f_{\delta} \right) = \sum_{j=1}^{\infty} (B^m_{\alpha} f_{\delta} + (f - f_{\delta}), \alpha^{-1} \sum_{j=1}^{\infty} B^j_{\alpha} f_{\delta}) \geq \alpha^{-1} \sum_{j=1}^{\infty} [(B^m_{\alpha} f_{\delta}, B^j_{\alpha} f_{\delta}) - \delta \|B^j_{\alpha} f_{\delta}\|].
\]

Therefore, a sufficient (but quite strict) condition for monotonicity of the error is

\[
\bar{d}^{(j)}_{\text{ME}}(\alpha) = (B^m_{\alpha} f_{\delta}, B^j_{\alpha} f_{\delta})/\|B^j_{\alpha} f_{\delta}\| \geq \delta \text{ for all } j \geq 1. \quad \text{Although the function } \bar{d}^{(j)}_{\text{ME}}(\alpha) \text{ is monotonically decreasing in } j \text{ and increasing in } \alpha, \text{ the last relation does not give a practical rule because } j \text{ is not bounded.}
\]

As shown in the previous section, one useful tool for generating approximate solutions or for choosing the regularization parameter for the standard approximation method is extrapolation. Besides the rules for choosing \( \alpha \) in the (iterated) Lavrentiev method considered above one may recommend to choose \( \alpha \) from the rule analogous to (6.6)

\[
\bar{d}^{(k)}_{\text{ME}}(\alpha) \equiv \frac{(- B^m_{\alpha} f_{\delta}, \sum_{n=m+2}^{m+k+1} (A v_{\alpha,n} - f_{\delta}))}{\|\sum_{n=m+2}^{m+k+1} (A v_{\alpha,n} - f_{\delta})\|} = b \delta, \quad b \geq 1, \quad \tag{6.7}
\]

using extrapolated approximations \( v_{\alpha,n} \) instead of iterated approximations. The approximations \( v_{\alpha,n} \) can be computed by formula (5.1), for example, with \( q_i = q^{-i[(k+1)/2]} \), \( q = \text{const} > 1 \), where \([x]\) is the smallest integer greater than or equal to \( x \).

7. Numerical experiments

We have solved the self-adjoint test problems from [10]: deriv2, foxgood, gravity, phillips, shaw. We used the discretization parameter 100, the other parameters had standard values.
All problems were normalized in such a way that the norms of the operator and the right-hand side were 1.

Instead of exact data \( f \) randomly perturbed data \( f_\delta \) were used with error \( \|f - f_\delta\| = \delta \), where the values of \( \delta \) were 0.5 and \( 10^{-i}, i = 1, \ldots, 7 \). The problems were regularized by the Lavrentiev method and by the Lavrentiev method 2, 3 times iterated (using \( u_0 = 0 \)) and by the extrapolated Lavrentiev method

\[
v_{n,\alpha}^{(k)} = \sum_{i=k-n+1}^{k} d_i u_{\alpha_i}, \quad d_i = \prod_{j=1,j \neq i}^{n} \frac{\alpha_j}{\alpha_j - \alpha_i} = \prod_{j=1,j \neq i}^{n} \frac{1}{1 - q^{i-j}}.
\]

with \( n = 2 \) or \( n = 3 \), using \( \alpha \)-values \( \alpha_i = q\alpha_{i-1}, q = 1.2 \).

In the model equations, the exact solutions are known. We found \( \alpha_{\text{opt}} \) as \( \alpha \) with the smallest error: \( \|u_{\alpha_{\text{opt}}} - u_*\| = \min \{ \|u_{\alpha} - u_*\|, \alpha \geq 0 \} \). We solved these problems 10 times using parameter choice rules with constant \( b = 1 \) and with the optimal constant found by optimization over all problems, all \( \delta \) and 10 runs. Though for some parameter choices the maxima of errors for \( b = 1 \) were smaller than the maxima for optimal \( b \), in general the convergence of the approximate solutions at \( b = 1 \) is not guaranteed. The results are given in Tables 7.1, 7.2: the first table corresponds to the original problems of [10] with the solution \( u_* \) and for the second table smoothed solutions \( Au_* \) with the right-hand side \( A^2u_* \) were used (then \( p = 1 \) in (2.1)).

Tables 7.1, 7.2 show the averages and maxima (over all problems, all \( \delta \) and 10 runs) of error ratios \( r = \|u_{\alpha_{\text{opt}}} - u_*\|/\|u_{\alpha_{\text{opt}}} - u_*\| \), where the numerator is error of the computed approximation under the given parameter choice rule and the denominator is the minimal error of the Lavrentiev approximation. In the additional column, the averages of the monotonicity indicators \( \max(\alpha_{\text{mon}} - \alpha(\delta), 0)/\alpha_{\text{mon}} \), where \( \alpha_{\text{mon}} = \min \{ \alpha' : \frac{d}{d\alpha}\|u_{\alpha} - u_*\| \geq 0, \alpha \geq \alpha' \} \) are given. Rules with smaller indicators better satisfy the monotonicity condition (1.6). Four zeros in monotonicity indicators for optimal \( b \) show that in the corresponding approximations the chosen parameter \( \alpha(\delta) \) was always larger than \( \alpha_{\text{mon}} \) in the Lavrentiev method. In the first table, rule (2.12) and the modified discrepancy principle gave the largest averages and maximums of \( r \), in the second table rule (2.12) gave a large averages and maxima at optimal \( b \), the modified discrepancy principle gave large maximum of \( r \) at optimal \( b \). As expected, the methods with a higher qualification (2 or 3 times iterated Lavrentiev methods and extrapolated Lavrentiev methods with 2 or 3 terms have qualification 2 or 3, respectively) had in the smooth case small averages of \( r \) at optimal \( b \).

### Table 7.1. Results for \( p = 0 \)

| Approx   | Parameter choice | Opt \( b \) | Avg \( r \) | Max \( r \) | Monot indic | \( b \) | Avg \( r \) | Max \( r \) | Monot indic |
|----------|------------------|-------------|-------------|-------------|-------------|--------|-------------|-------------|-------------|
| Lavrt(2.12) |                             | 1.059       | 1.594       | 20.459      | 0.085       | 1      | 1.722       | 19.877      | 0.268       |
| Lavrt mod diser |                          | 1.143       | 1.581       | 20.641      | 0.107       | 1      | 1.878       | 19.304      | 0.413       |
| Lavrt(2.4) |                          | 1.364       | 1.089       | 2.567       | 0.090       | 1      | 1.648       | 8.657       | 0.517       |
| Lavrt(6.4), \( \nu = 0.17 \) |                      | 1.096       | 1.066       | 2.136       | 0.038       | 1      | 1.115       | 3.117       | 0.076       |
| Lavrt(6.6), \( k = 2 \) |                       | 1.251       | 1.077       | 2.366       | 0.062       | 1      | 1.486       | 7.555       | 0.455       |
| Lavrt(6.7), \( k = 1 \) |                       | 1.421       | 1.078       | 2.278       | 0.093       | 1      | 1.700       | 9.181       | 0.537       |
| Lavrt(6.7), \( k = 2 \) |                       | 1.280       | 1.063       | 2.093       | 0.063       | 1      | 1.509       | 7.865       | 0.471       |
| 2-iter Lavrt mod diser |                             | 1.187       | 1.146       | 9.783       | 0.000       | 1      | 1.467       | 8.720       | 0.094       |
| 3-iter Lavrt discrepancy |                            | 1.530       | 0.954       | 3.053       | 0.000       | 1      | 1.714       | 9.709       | 0.094       |
| 2-extr Lavrt mod diser |                             | 1.236       | 1.119       | 8.742       | 0.000       | 1      | 1.513       | 7.565       | 0.092       |
| 3-extr Lavrt discrepancy |                             | 1.535       | 0.956       | 3.059       | 0.000       | 1      | 1.722       | 9.758       | 0.092       |
| Lavrt(2.8) |                             | 1.209       | 1.092       | 2.510       | 0.106       | 1      | 1.135       | 2.141       | 0.271       |
| Lavrt(6.5), \( l = 5 \) |                             | 1.004       | 1.019       | 1.315       | 0.003       | 1      | 1.019       | 1.313       | 0.004       |
Table 7.2. Results for $p = 1$

| Approx       | Parameter choice | Opt $b$ | Avg $r$ | Max $r$ | Monot indic | $b$ | Avg $r$ | Max $r$ | Monot indic |
|--------------|------------------|--------|---------|---------|-------------|----|---------|---------|-------------|
| Lavr         | (2.12)           | 1.021  | 1.350   | 3.291   | 0.113       | 1  | 1.397   | 4.343   | 0.168       |
| Lavr mod discr | (2.4)           | 1.263  | 1.019   | 1.437   | 0.040       | 1  | 1.437   | 6.048   | 0.428       |
| Lavr         | (2.4), $\nu = 0.4$ | 1.583  | 1.015   | 1.093   | 0.063       | 1  | 1.708   | 7.689   | 0.536       |
| Lavr         | (6.4), $k = 2$   | 1.365  | 1.006   | 1.083   | 0.030       | 1  | 1.498   | 6.616   | 0.459       |
| Lavr         | (6.7), $k = 1$   | 1.661  | 1.018   | 1.114   | 0.073       | 1  | 1.784   | 8.124   | 0.559       |
| Lavr         | (6.7), $k = 2$   | 1.416  | 1.006   | 1.055   | 0.035       | 1  | 1.548   | 6.919   | 0.482       |
| 2-iter Lavr mod discr |         | 1.352  | 0.523   | 1.007   | 0.000       | 1  | 0.847   | 5.214   | 0.063       |
| 3-iter Lavr discrepancy |     | 1.856  | 0.552   | 1.235   | 0.000       | 1  | 1.240   | 7.820   | 0.063       |
| 2-extr Lavr mod discr |         | 1.443  | 0.530   | 1.034   | 0.000       | 1  | 0.921   | 5.717   | 0.063       |
| 3-extr Lavr discrepancy |     | 1.862  | 0.555   | 1.240   | 0.000       | 1  | 1.249   | 7.882   | 0.062       |
| Lavr         | (2.8)           | 1.430  | 1.025   | 1.146   | 0.090       | 1  | 1.140   | 1.233   | 0.324       |
| Lavr         | (6.5), $l = 4$  | 1.041  | 1.003   | 1.066   | 0.017       | 1  | 1.005   | 1.057   | 0.054       |

In the Lavrentiev method, rules (2.4) and (2.8) were good, and their modifications were even somewhat better. Note that in the smooth case for the 3-extrapolated Lavrentiev approximation the modified discrepancy principle gave 20% smaller averages and maxima of $r$ than the discrepancy principle.

Table 7.3 presents the results for the nonsmooth case ($p = 0$) where the supposed noise level $\delta$ is $d$ times higher than $\|f - f_\delta\|$ and the optimal constant $b$ is used. Again, the first two rules gave the largest numbers. Setting aside the first two rules, the other rules gave smaller maxima of error ratios than the 2-iterated Lavrentiev method in the case of all $d$ and the 3-iterated Lavrentiev method in the case $d = 1, 2$. As in Tables 7.1, 7.2, also in Table 7.3 the smallest values of averages and maxima of $r$ in the Lavrentiev method were given by the last rule (6.5), (6.4) with $\nu = 0.17$ and (6.7) with $k = 2$; an especially good performance was shown by the last rule (6.5) in the nonsmooth case. However, the set of problems was not large enough for more far-reaching conclusions.

Table 7.3. Results for rough $\delta$

| Approx       | Parameter choice | Avg $r$ $d = 2$ | Max $r$ $d = 2$ | Avg $r$ $d = 3$ | Max $r$ $d = 3$ | Avg $r$ $d = 10$ | Max $r$ $d = 10$ |
|--------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Lavr         | (2.12)           | 2.117           | 28.998          | 2.467           | 35.544          | 3.914           | 64.736          |
| Lavr mod discr | (2.4)           | 2.007           | 29.216          | 2.342           | 35.772          | 3.754           | 64.793          |
| Lavr         | (6.4), $\nu = 0.17$ | 1.281           | 4.920           | 1.447           | 7.343           | 2.333           | 23.648          |
| Lavr         | (6.6), $k = 2$   | 1.361           | 3.979           | 1.554           | 5.936           | 2.436           | 19.465          |
| Lavr         | (6.7), $k = 1$   | 1.299           | 4.536           | 1.473           | 6.779           | 2.367           | 22.145          |
| Lavr         | (6.7), $k = 2$   | 1.251           | 4.297           | 1.404           | 6.420           | 2.241           | 21.034          |
| 2-iter Lavr mod discr |         | 1.266           | 3.889           | 1.425           | 5.807           | 2.262           | 19.219          |
| 3-iter Lavr discrepancy |     | 1.473           | 15.544          | 1.737           | 20.346          | 2.943           | 44.607          |
| 2-extr Lavr mod discr |         | 1.123           | 5.822           | 1.306           | 8.626           | 2.275           | 26.993          |
| 3-extr Lavr discrepancy |     | 1.415           | 14.040          | 1.663           | 18.512          | 2.825           | 41.625          |
| Lavr         | (2.8)           | 1.245           | 3.591           | 1.466           | 5.341           | 2.301           | 18.633          |

The criterion for numerical comparison of accuracy of these methods and parameter choice rules is the amount of computational work. Namely, for choosing the parameter $\alpha$ in the (iterated) Lavrentiev method $u_{\alpha,m} = \alpha^{-1} \sum_{j=1}^{m} B_{\alpha} f_{\delta}$ many rules require additional iterations for computing iterated Lavrentiev approximations $u_{\alpha,n}$, where $n > m$, and then $u_{\alpha,n}$ can be considered as an approximate solution of (1.1). Our rule (2.4) requires computa-
tion of \( u_{\alpha,m+2} \). But once \( u_{\alpha,m+2} \) has been computed, one can take for approximate solution \( u_{\alpha,m+1} \) with choice of \( \alpha \) from the modified discrepancy principle
\[
\| B_\alpha (Au_{\alpha,m+1} - f_\delta) \| \equiv \| Au_{\alpha,m+2} - f_\delta \| = b \delta, \quad b \geq 1,
\]
or take for the approximate solution \( u_{\alpha,m+2} \) with the choice of \( \alpha \) from the discrepancy principle
\[
\| Au_{\alpha,m+2} - f_\delta \| = b \delta,
\]
giving the same \( \alpha \) as (7.1). If the solution has the source-like representation (2.1), then for \( u_{\alpha,m+1} \) and for \( u_{\alpha,m+2} \) with \( \alpha \) from (7.1) the error estimate (2.1) holds with \( p \leq m+1 \), for \( u_{\alpha,m} \) with proper \( \alpha \) the error estimate (2.1) holds with \( p \leq m \) (and on the other hand, \( u_{\alpha,m+1} \) and \( u_{\alpha,m+2} \) have large maximums in Table 7.3).

Note that rule (6.6) requires computation of \( u_{\alpha,m+k+1} \), but rules (2.8), (6.5) only use \( u_{\alpha,m+1} \) with different \( \alpha \) and rule (6.7) uses \( u_{\alpha,m} \) with different \( \alpha \). Actually, if \( u_{\alpha,m} \) with different \( \alpha \) are available, in the case of the smooth solution we strongly recommend to use the extrapolation of the iterated Lavrentiev approximation (5.2) instead of \( u_{\alpha,m} \) because of the much higher qualification (qualification is \( nm \), where \( n \) is the number of terms in \( v_{n,\alpha} \), see [4]). Also Tikhonov approximations can be extrapolated analogously, the corresponding numerical results can be found in [4,5].

Note also that if \( \delta \) is large, the operator is not normalized and the norm of the operator is less than 1, then we recommend to multiply the left hand side of rule (2.4) by \( \kappa_\alpha = 1 + \alpha / \| A \| \) and to choose \( \alpha \) from the equation
\[
d_{ME}(\alpha) = \frac{\kappa_\alpha \| B_\alpha (Au - f_\delta) \|^2}{\| B_\alpha^2 (Au - f_\delta) \|} = \frac{\kappa_\alpha \| Au_{\alpha,m+1} - f_\delta \|^2}{\| Au_{\alpha,m+2} - f_\delta \|} = b \delta, \quad b = \text{const} \geq 1.
\]

It is clear that the monotonicity properties \( d_{ME}(\alpha) \) (see Theorem 3.1) are preserved when multiplied by the monotonic function \( \kappa_\alpha \). The corresponding modification can also be useful for other rules.

Acknowledgement. This work was supported by the Estonian Science Foundation, grant No. 7489.

References

1. S. George and M.T. Nair, An optimal order yielding discrepancy for simplified regularization of ill-posed problems in Hilbert scales, International J. Math. and Math. Sciences, (2003), no. 39, pp. 2487–2499.
2. S. George and M.T. Nair, An optimal order yielding discrepancy principle for simplified regularization in Hilbert scales: finite dimensional realizations, International J. Math. and Math. Sciences, (2004), no. 37, pp. 1973–1996.
3. U. Hämärk, On the parameter choice in the regularized Ritz-Galerkin method, Proc. Estonian Acad. Sci. Phys. Math., 42 (1993), no. 2, pp. 133–143.
4. U. Hämärk, R. Palm, and T. Raus, Use of extrapolation in regularization methods, J. Inverse and Ill-Posed Problems, 15 (2007), no. 3, pp. 277–294.
5. U. Hämärk, R. Palm, and T. Raus, Extrapolation of Tikhonov and Lavrentiev regularization methods, in: ICIPE2008, International Conference on Inverse Problems in Engineering: Theory and Practice, Dourdan (Paris), France, June 15–19, 2008, 51, 2008, pp. 1–8.
6. U. Hämärk and T. Raus, On the a posteriori parameter choice in regularization methods, Proc. Estonian Acad. Sci. Phys. Math., 48 (1999), no. 2, pp. 133–145.
7. U. Hämärk and T. Raus, On the choice of the regularization parameter in ill-posed problems with approximately given noise level of data, J. Inv. and Ill-Posed Problems, 14 (2006), no. 3, pp. 251–266.
8. U. Hämärk and U. Tautenhahn, On the monotone error rule for parameter choice in iterative and continuous regularization methods, BIT Numerical Mathematics, 41 (2001), no. 5, pp. 1029–1038.
9. U. Hämärk and U. Tautenhahn, On the monotone error rule for choosing the regularization parameter in ill-posed problems, in: Proc. of the International Conference, Samarkand, Uzbekistan 2000, Ill-Posed
and Non-Classical Problems of Mathematical Physics and Analysis (M. M. Lavrentiev and S. I. Kabanikhin, eds.), VSP, Utrecht-Boston, 2003, pp. 27–55.

10. P. C. Hansen, Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems, Numerical Algorithms, 6 (1994), no. 1, pp. 1–35.

11. J. Janno and U. Tautenhahn, On Lavrentiev regularization for ill-posed problems in Hilbert scales, Numer. Funct. Anal. Optim., 24 (2003), no. 5–6, pp. 531–555.

12. R. D. Lazarov, S. Lu, and S. V. Pereverzyev, On the balancing principle for some problems of numerical analysis, Numerische Mathematik, 106 (2007), pp. 659–689.

13. P. Mathé and S. V. Pereverzyev, Geometry of linear ill-posed problems in variable Hilbert scales, Inverse Problems, 19 (2003), no. 3, pp. 789–803.

14. J. Janno and U. Tautenhahn, On Lavrentiev regularization for ill-posed problems in Hilbert scales, Numer. Funct. Anal. Optim., 24 (2003), no. 5–6, pp. 531–555.

15. R. D. Lazarov, S. Lu, and S. V. Pereverzyev, On the balancing principle for some problems of numerical analysis, Numerische Mathematik, 106 (2007), pp. 659–689.

16. P. Mathé and S. V. Pereverzyev, Geometry of linear ill-posed problems in variable Hilbert scales, Inverse Problems, 19 (2003), no. 3, pp. 789–803.

17. S. Morigi, L. Reichel, and F. Sgallari, An iterative Lavrentiev regularization method, BIT Numerical Mathematics, 46 (2006), pp. 589–606.

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

19. M. T. Nair and U. Tautenhahn, Convergence rates for Lavrentiev-type regularization in Hilbert scales, Computational Methods in Applied Mathematics, 8 (2008), no. 3, pp 279–293.

20. S. V. Pereverzyev and E. Schock, On the adaptive selection of the parameter in the regularization of ill-posed problems, SIAM J. Numerical Analysis, 43 (2005), pp. 2060–2076.

21. R. Plato and U. Hämairik, On pseudo-optimal parameter choices and stopping rules for regularization methods in Banach spaces, Numer. Funct. Anal. Optim., 17 (1996), pp. 181–195.

22. T. Raus, The principle of the residual in the solution of ill-posed problems, Acta et comment. Univ. Tartuensis, 672 (1984), pp. 16–26. (in Russian)

23. T. Raus, The principle of the residual in the solution of ill-posed problems with non-selfadjoint operator, Acta et comment. Univ. Tartuensis, 715 (1985), pp. 12–20. (in Russian)

24. T. Raus, An a posteriori choice of the regularization parameter in case of approximately given error bound of data, Acta et comment. Univ. Tartuensis, 913 (1990), pp. 73–87.

25. T. Raus and U. Hämairik, On the quasioptimal regularization parameter choices for solving ill-posed problems, J. Inverse Ill-Posed Problems, 15 (2007), no. 4, pp. 419–439.

26. T. Raus and U. Hämairik, About the Lepskii-type balancing principle for choice of the regularization parameter, in: ICIPE2008, International Conference on Inverse Problems in Engineering: Theory and Practice, Dourdan (Paris), France, June 15–19, 2008, 100, 2008, pp.1–8.

27. 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, 1998, pp. 118–125.

28. U. Tautenhahn, Lavrentiev regularization of nonlinear ill-posed problems, Vietnam J. Math., 32 (2004), no. Special Issue, pp. 29–41.

29. U. Tautenhahn and U. Hämairik, The use of monotonicity for choosing the regularization parameter in ill-posed problems, Inverse Problems, 15 (1999), no. 6, pp. 1487–1505.

30. G. M. Vainikko, The principle of the residual for a class of regularization methods, USSR Comput. Math. Math. Phys., 22 (1982), no. 3, pp. 1–19.

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