H01

Proceedings of the 5th International Conference on Inverse Problems in Engineering: Theory and Practice, Cambridge, UK, 11-15th July 2005

ON THE CHOICE OF THE REGULARIZATION PARAMETER IN THE CASE OF THE APPROXIMATELY GIVEN NOISE LEVEL OF DATA

U. HÄMARIK and T. RAUS
Institute of Applied Mathematics, University of Tartu, Estonia
e-mails: uno.hamarik@ut.ee, toomas.raus@mtk.ut.ee

Abstract - We consider linear ill-posed problems $Au = f$ in Hilbert spaces, mostly in case $A = A^* \geq 0$. Regularized approximations u_r to solutions u_* of problem $Au = f$ are obtained by a general regularization scheme, including the Lavrentiev method, iterative and other methods. We assume that instead of $f \in \mathcal{R}(A)$ noisy data $\tilde{f}$ are available with the approximately given noise level δ: it holds $\|\tilde{f} - f\|/\delta \leq c$ for $\delta \to 0$, but $c = \text{const}$ is unknown. We propose a new a-posteriori rule for the choice of the regularization parameter $r = r(\delta)$ guaranteeing $u_{r(\delta)} \to u_*$ for $\delta \to 0$. Note that such convergence is not guaranteed for the parameter choice given by the L-curve rule, by the GCV-rule and also for discrepancy principle $\|Au_r - \tilde{f}\| = b\delta$ with $b \leq c$. We give error estimates which in case $\|\tilde{f} - f\| \leq \delta$ are quasioptimal and order-optimal.

1. INTRODUCTION

We consider an operator equation

$$Au = f\,, \quad f \in R(A)\,, \tag{1}$$

where $A \in L(H, H)$, $A = A^* \geq 0$ is the linear continuous self-adjoint and non-negative operator; u and f are elements of the real Hilbert space H. In general our problem is ill-posed: the range $\mathcal{R}(A)$ may be non-closed, the kernel $N(A)$ may be non-trivial. As usual, in the treatment of ill-posed problems, we suppose that instead of the exact right-hand side f we have only an approximation $\tilde{f} \in H$ with noise $\tilde{f} - f$.

The approximate solution u_r of the ill-posed problem $Au = f$ is found by some regularization method and depends on the regularization parameter r. The important problem is how to choose the proper regularization parameter r. If there is some information about the noise level of the data, this information should be used for the choice of r. Consider now the choice of r in situations with a different amount of information about $\|\tilde{f} - f\|$.

Situation 1. Full information about the noise level is known: the exact noise level δ with $\|\tilde{f} - f\| \leq \delta$ is given. Then the proper parameter choice $r = r(\delta)$ guarantees $u_{r(\delta)} \to u_*$ for $\delta \to 0$, where u_* is the solution of $Au = f$, the nearest to the initial approximation u_0 (see Section 2; often $u_0 = 0$). In this situation proper rules for the choice of r are the discrepancy principle [9, 15, 16] and its modification [10].

Situation 2. Nothing is known about the noise level. If there is no information about noise level δ, parameter r may be chosen by the quasioptimality criterion [14, 15], by the GCV-rule [3, 17], by the L-curve rule [2, 8] or by rule of [7]. The serious drawback of these rules is that convergence $u_{r(\delta)} \to u_*$ for $\delta \to 0$ is not guaranteed (see [1]).

In applied inverse and ill-posed problems the situation is often between extreme Situations 1, 2: some approximate δ is known, but it is unknown, if the inequality $\|\tilde{f} - f\| \leq \delta$ holds or not. In this paper we are interested in the case of approximately given noise level δ: instead of the inequality $\|\tilde{f} - f\| \leq \delta$ we assume that $\|\tilde{f} - f\|/\delta \leq c$ for $\delta \to 0$, where c is an unknown constant. We give a rule for the parameter choice $r = r(\delta)$ guaranteeing $u_{r(\delta)} \to u_*$ for $\delta \to 0$. This rule was lately proposed in [5, 6], where convergence is also proven. In this paper we prove error estimates.

2. REGULARIZATION METHODS

We consider the regularization methods in the general form (see [15, 16])

$$u_r = (I - Ag_r(A))u_0 + g_r(A)\tilde{f}\,, \tag{2}$$

where u_r is the approximate solution, u_0 – initial approximation, r – regularization parameter, I – the identity operator and the function $g_r(\lambda)$ satisfies the conditions (3) and (4):

$$\sup_{0 \leq \lambda \leq a} |g_r(\lambda)| \leq \gamma r\,, \quad r \geq 0\,, \tag{3}$$

$$\sup_{0\le\lambda\le a} \lambda^p\left|1-\lambda g_r(\lambda)\right| \le \gamma_p r^{-p}\,, \quad r\ge 0\,,0\le p\le p_0\,. \tag{4}$$

Here p_0, γ and γ_p are positive constants, $a \ge \|A\|$, $\gamma_0 \le 1$ and the greatest value of p_0, for which the inequality (4) holds is called the qualification of method.

The following regularization methods are special cases of the general method (2).

M1 The Lavrentiev method $u_\alpha = (\alpha I + A)^{-1}\tilde{f}$. Here $u_0 = 0$, $r = \alpha^{-1}$, $g_r(\lambda) = (\lambda + r^{-1})^{-1}$, $p_0 = 1$, $\gamma = 1$, $\gamma_p = p^p(1-p)^{1-p}$.

M2 The iterative variant of the Lavrentiev method. Let $m \in \mathbf{N}$, $m \ge 1$, $u_0 = u_{0,\alpha} \in H$ – initial approximation and $u_{m,\alpha} = (\alpha I + A)^{-1}(\alpha u_{m-1,\alpha} + \tilde{f})$. Here $r = \alpha^{-1}$, $g_r(\lambda) = \frac{1}{\lambda}\left(1 - \left(\frac{1}{1+r\lambda}\right)^m\right)$, $p_0 = m$, $\gamma = m$, $\gamma_p = (p/m)^p(1-p/m)^{m-p}$.

M3 Explicit iteration scheme (the Landweber's method). Let $0 < \mu < 1/\|A\|$ be a constant and $u_n = u_{n-1} - \mu(Au_{n-1} - \tilde{f})$, $n = 1, 2, \ldots$. Here $r = n$, $g_r(\lambda) = \frac{1}{\lambda}(1-(1-\mu\lambda)^r)$, $p_0 = \infty$, $\gamma = \mu$, $\gamma_p = (p/(\mu e))^p$.

M4 Implicit iteration scheme. Let $\alpha > 0$ be a constant and $\alpha u_n + Au_n = \alpha u_{n-1} + \tilde{f}$, $n = 1, 2, \ldots$. Here $r = n$, $g_r(\lambda) = \frac{1}{\lambda}\left(1 - \left(\frac{\alpha}{\alpha+\lambda}\right)^r\right)$, $p_0 = \infty$, $\gamma = 1/\alpha$, $\gamma_p = (\alpha p)^p$.

M5 The method of the Cauchy problem: approximation u_r solves the Cauchy problem $u'(r) + Au(r) = \tilde{f}$, $u(0) = u_0$. Here $g_r(\lambda) = \frac{1}{\lambda}(1 - e^{-r\lambda})$, $p_0 = \infty$, $\gamma = 1$, $\gamma_p = (p/e)^p$.

Note that some other regularization methods are studied in recent book [4].

3. PARAMETER CHOICE IN THE CASE OF THE KNOWN NOISE LEVEL OF DATA

In regularization methods (2) the error $u_r - u_*$ depends crucially on the choice of a regularization parameter r. In the case of small r the approximation error of the exact solution is large and in the case of big r the error $u_r - u_*$ is large due to noise in data.

At first we consider the choice of r in the case when the noise level δ with $\|\tilde{f} - f\| \le \delta$ is known. Then the most prominent rule for methods M2–M5 is the discrepancy principle [9, 15, 16]. In this rule the regularization parameter $r = r_D$ is chosen as the solution of the equation $\left\|Au_r - \tilde{f}\right\| \approx b\delta$ with $b = \text{const} > 1$. The second rule in the case of known δ is the modification of the discrepancy principle (the MD rule) [10]. In this rule the regularization parameter $r = r_{MD}$ is chosen as the solution of the equation

$$\left\|B_r(Au_r - \tilde{f})\right\| \approx b\delta \quad \text{with} \quad b = \text{const} > 1\,, \quad B_r = \begin{cases} I & \text{for} \quad p_0 = \infty\,, \\ (I - Ag_r(A))^{1/p_0} & \text{for} \quad p_0 \ne \infty \end{cases}$$

where the operator B_r depends on the qualification p_0 of the method.

The discrepancy principle and its modification coincide for regularization methods M3–M5 where $p_0 = \infty$, but differ for methods M1, M2, where $\|B_r(Au_{\alpha,m} - \tilde{f})\| = \|Au_{\alpha,m+1} - \tilde{f}\|$.

Both rules guarantee convergence $\|u_r - u_*\| \to 0$ for $\delta \to 0$ and order-optimality: if $u_0 - u_* = A^p v$, $v \in H$, $\|v\| \le \varrho$, $p > 0$, then $\|u_r - u_*\| \le C_p \varrho^{\frac{1}{\varrho+1}} \delta^{\frac{p}{p+1}}$, where $p \in (0, p_0 - 1)$ for $r = r_D$ and $p \in (0, p_0)$ for $r = r_{\text{MD}}$. In contrast to the discrepancy principle the MD rule has also the quasioptimality property: there exists a constant c such that

$$\|u_{r_{MD}} - u_*\| \le c \inf_{r\ge 0}\{\|(I - Ag_r(A))(u_0 - u_*)\| + \gamma r\delta\}\,. \tag{5}$$

Note that in the case $\|\tilde{f} - f\| \le \delta$ relations (2)–(4) yield estimate $\|u_r - u_*\| \le \|(I - Ag_r(A))(u_0 - u_*)\| + \gamma r\delta$ which explains why rules with property (5) are called quasioptimal. It is obvious that if a rule is quasioptimal, then this rule is order-optimal for all $p \in (0, p_0)$.

The discrepancy principle and the MD-rule are unstable in the sense that if the actual error of the right-hand side is only slightly larger than $b\delta$, then the error of the approximate solution may be arbitrarily large, irrespective of the value of the ratio of the actual and supposed noise level. For example, if $b = 2$ and the actual noise level is three times larger than the noise level δ, which we use in the rule, then the error of the approximate solution may be arbitrarily large.

There are also heuristic parameter choice rules which do not use the noise level δ: the quasioptimality criterion [13, 14]. the Wahba's generalized cross-validation rule [17, 3], the Hansen's L-curve rule [8, 2] and the rules of [7]. These rules are formulated for non-selfadjoint problems, reformulations for selfadjoint problems are possible. For example, the selfadjoint variant of the quasioptimality criterion choses the parameter for which the function $k(r) = r\|B_r(Au_r - \tilde{f})\|$ has the global minimum.

Heuristic rules often work well, but as shown by Bakushinskii [1], one cannot prove the convergence of the approximate solution.

4. PARAMETER CHOICE IF THE NOISE LEVEL IS GIVEN APPROXIMATELY

In applied ill-posed problems the exact noise level is often unknown. Therefore in the following we assume that only some guesses about this level can be made. It means that the supposed error level $\delta > 0$ is given, but we do not know exactly, if $\|\tilde{f} - f\| \leq \delta$ holds or not. We give the rule for the stable parameter choice which guarantees the convergence of the approximate solution to the exact solution if only the ratio $\|\tilde{f} - f\|/\delta$ is bounded in the process $\delta \to 0$.

Let us introduce the function

$$\varphi(r) = \sqrt{r}\|A^{1/2}B_r^{3/2}(Au_r - \tilde{f})\| = \sqrt{r}\langle B_r(Au_r - \tilde{f}), AB_r^2(Au_r - \tilde{f})\rangle^{1/2}.$$

Note that for methods M1, M2 $B_r = (I + rA)^{-1}$ and $\varphi(r) = \varphi(\alpha^{-1}) = \frac{1}{\sqrt{\alpha}}\langle Au_{m+1,\alpha} - \tilde{f}, A(Au_{m+2,\alpha} - \tilde{f})\rangle^{1/2}$; for iterative methods M3, M4 $\varphi(r) = \varphi(n) = \sqrt{n}\langle Au_n - \tilde{f}, Au_n - \tilde{f})\rangle^{1/2}$.

Rule P. Let $0 \leq s \leq 1$ and b_1, b_2 be such constants that $b_2 \geq b_1 > C_m$, where the value of the constant C_m is $C_m = 1/2$, $C_m = 1/\sqrt{2m+3}$, $C_m = 1/\sqrt{2\mu e}$, $C_m = \sqrt{\alpha/2}$ and $C_m = 1/\sqrt{2e}$ for methods M1–M5 respectively. If $\varphi(1) \leq b_2\delta$ then choose $r(\delta) = 1$. In the contrary case we find at first $r_2(\delta) > 1$ such that

$$\varphi(r_2(\delta)) \leq b_2\delta\,, \qquad (6)$$

$$\varphi(r) \geq b_1\delta \qquad \forall\, r \in [1, r_2(\delta)]\,. \qquad (7)$$

For the regularization parameter $r(\delta)$ we choose the parameter r, for which the function $t(r) = r^s\|B_r(Au_r - \tilde{f})\|$ has the global minimum on the interval $[1, r_2(\delta)]$.

Let us reformulate the rule P for the choice of the stopping index $n(\delta)$ as the parameter r in iterative methods. For this rule P' the analogous results hold as for the rule P.

Rule P'. Let $0 \leq s \leq 1$ and b be the constant such that $b > C_m$. Find $n_2(\delta)$ as the first $n = 1, 2, \ldots$, for which $\varphi(n) \leq b\delta$. For the regularization parameter $n(\delta)$ we choose $n \in \mathbf{N}$, for which the function $t(n) = n^s\|Au_n - \tilde{f}\|$ has the global minimum on the interval $[1, n_2(\delta)]$.

Rules P and P' are similar to the rules in [11, 13, 14]. In [11] for the regularization parameter the parameter $r_2(\delta)$ was taken. We can consider the rule P as the generalization of this rule, since in case $s = 0$ these rules coincide while the function $\|B_r(Au_r - \tilde{f})\|$ is monotonically decreasing with respect to r. On the other hand, in case $s = 1$ the rule P is similar to the selfadjoint analogue of the quasioptimality criterion. In both rules the regularization parameter is chosen as the minimizer of the function $r\|B_r(Au_r - \tilde{f})\|$; only intervals for minimization are different: the intervals are $[1, r_2(\delta)]$ and $[1, \infty)$ respectively.

In [11] for methods M1–M5 the following results are proven:

(i) for each $\tilde{f} \in H$ we have $\lim_{r \to \infty} \varphi(r) = 0$;

(ii) if $\|\tilde{f} - f\| \leq \delta$, $\|u_0 - u_*\| \leq M$, $b \geq \overline{C}_m$, then for each r, $r \geq R_{M,\delta} = \overline{C}_m M/(b - C_m)\delta$ we have $\varphi(r) \leq b\delta$; here the constant $\overline{C}_m$ in methods M1–M5 has values $12\sqrt{15}/125$, $(3/2)^{(3/2)}m^m/(m+3/2)^{m+3/2}$, $(3/(2\mu e))^{3/2}$, $(3\alpha/2)^{(3/2)}$, $(3/(2e))^{3/2}$ respectively;

(iii) if $\frac{\|\tilde{f} - f\|}{\delta} \leq \text{const}$ for $\delta \to 0$ then $\|u_{r_2(\delta)} - u_*\| \to 0$ for $\delta \to 0$.

The property (i) and the continuity of the function $\varphi(r)$ guarantee that the choice of finite parameters $r_2(\delta)$ and $r(\delta) \leq r_2(\delta)$ according to Rule P is possible. From the property (ii) it follows that if we know a constant $M > 0$ such that $\|u_0 - u_*\| \leq M$, then it is sufficient to search for the parameter $r_2(\delta)$ in the finite interval $[1, R_{M,\delta}]$. Note that the function $\varphi(r)$ may be non-monotone and therefore in Rule P we must use the conditions (6)–(7) instead of inequalities $b_1\delta \leq \varphi(r) \leq b_2\delta$.

Note that the analogues of the results of the paper [11] for non-selfadjoint problems are presented in [12]. In [5, 6] the following convergence result is proven.

Theorem 1. *Let $A \in L(H, H)$, $A = A^* \geq 0$, $f \in R(A)$. Let the parameter $r(\delta)$ be chosen according to Rule P. If $\frac{\|\tilde{f} - f\|}{\delta} \leq$ const in the process $\delta \to 0$, then in methods M1–M5*

$$\left\|u_{r(\delta)} - u_*\right\| \to 0 \quad \textit{for} \quad \delta \to 0\,.$$

In the following theorem we give error estimates, using notation

$$\psi(r) := \|G_r(u_0 - u_*)\| + \gamma r \max\{\delta, \|\tilde{f} - f\|\}, \quad G_r := I - Ag_r(A).$$

Theorem 2. *Let $A \in L(H,H)$, $A = A^* \geq 0$, $f \in R(A)$. Let the parameter $r(\delta)$ be chosen according to Rule P with $s \in (0,1)$. Let $\delta_0 := \frac{1}{2}\|B_{r(\delta)}(Au_{r(\delta)} - \tilde{f})\|$. Then for methods M1–M5 the following error estimates hold*

1. If $\|\tilde{f} - f\| \leq \max\{\delta, \delta_0\}$, then

$$\|u_{r(\delta)} - u_*\| \leq C(b_1, b_*, d_*)\frac{1}{1-s}\inf_{r\geq 0}\psi(r), \quad d_* = \max_{r,r',r(\delta)\leq r\leq r'\leq r_2(\delta)+1}\left(\frac{r^s\|B_r(Au_r - \tilde{f})\|}{(\varrho r')^s\|B_{\varrho r'}(Au_{\varrho r'} - \tilde{f})\|}\right) \quad (8)$$

Here $\varrho := 1 + \gamma\gamma_1$, $b_* = \begin{cases} b_2, \text{ if } r(\delta) \geq R(\delta), \\ \max_{r(\delta)\leq r\leq R(\delta)} \varphi(r)/\delta, \text{ if } r(\delta) < R(\delta) \end{cases}$ and $R(\delta)$ is the greatest parameter for which $\varphi(r) = b_2\delta$.

2. If $\max\{\delta, \delta_0\} < \|\tilde{f} - f\| \leq \delta_1 = \frac{1}{2}\|B_1(Au_1 - \tilde{f})\|$, then

$$\|u_{r(\delta)} - u_*\| \leq C\left(\frac{\|\tilde{f} - f\|}{\delta_0}\right)^{1/s}\inf_{r\geq 0}\psi(r). \quad (9)$$

To prove Theorem 2, we need the following lemmas.

Lemma 1. *Let $\|\tilde{f} - f\| \leq \delta$ and $\psi_1(r) := \|G_r(u_0 - u_*)\| + \gamma r\delta$. Then for methods M1–M5 the following assertions hold:*

a) if $\|B_r(Au_r - \tilde{f})\| \geq b_1\delta$, $b_1 > 1$, then $\psi_1(r) \leq c_1(b_1)\psi_1(r')$ for each $r' \leq r$;

b) if $\|B_r(Au_r - \tilde{f})\| \leq b_2\delta$, then $\psi_1(r) \leq c_2(b_2)\psi_1(r')$ for each $r' \geq r$;

c) if $\varphi(r) \geq b_1\delta$, $b_1 > C_m$, then $\psi_1(r) \leq c_3(b_1)\psi_1(r')$ for each $r' \leq r$;

d) if $\varphi(s') \leq b_2\delta$ for each $s' > r$, then $\psi_1(r) \leq c_4(b_2)\psi_1(r')$ for each $r' \geq r$.

Proof of assertions a) and b) can be found in [10], proof of assertions c) and d) can be found in [11].

Lemma 2. *Let f_0 be an element such that*

$$\|P(\lambda)(f_0 - f)\| = \|P(\lambda)(\tilde{f} - f)\|, \quad \langle dP(\lambda)(u_0 - u_*), f_0 - f\rangle \geq 0, \quad 0 \leq \lambda \leq a,$$

where $P(\lambda)$ is the spectral family of the projectors of the operator A. Then for methods M1–M5 in case $r \geq 1$ it holds

$$\gamma r\|B_{\varrho r}(Au_{\varrho r} - \tilde{f})\| \leq \|u_r^0 - u_*\|,$$

where $u_r^0 = G_r u_0 + g_r(A)f_0$.

Proof. We have

$$u_r - u_* = G_r(u_0 - u_*) + g_r(A)(f - \tilde{f}), \quad (10)$$

$$B_r(Au_r - \tilde{f}) = B_rG_rA(u_0 - u_*) - B_rG_r(\tilde{f} - f),$$

$$(\gamma r)^2\|B_{\varrho r}(Au_{\varrho r} - \tilde{f})\|^2 = (\gamma r)^2\|B_{\varrho r}G_{\varrho r}A(u_0 - u_*)\|^2$$
$$- 2(\gamma r)^2\langle B_{\varrho r}G_{\varrho r}A(u_0 - u_*), B_{\varrho r}G_{\varrho r}(\tilde{f} - f)\rangle + (\gamma r)^2\|B_{\varrho r}G_{\varrho r}(\tilde{f} - f)\|^2. \quad (11)$$

It is easy to show that for methods M1–M5 the function $g_r(\lambda)$ satisfies the following conditions:

a) $0 \leq 1 - \lambda g_r(\lambda) \leq 1$, $0 \leq \lambda \leq a$, $r \geq 0$;

b) the function $r \to 1 - \lambda g_r(\lambda)$ is monotonically decreasing for $r \geq 0$;

c) $\gamma r\lambda\beta_{\varrho r}(\lambda)(1-\lambda g_{\varrho r}(\lambda)) \le 1-\lambda g_r(\lambda)$, $0\le\lambda\le a$, $r\ge 0$, where

$$\beta_r(\lambda)=\begin{cases}1, & \text{if } p_0=\infty,\\ (1-\lambda g_r(\lambda))^{1/p_0}, & \text{if } p_0<\infty.\end{cases}$$

d) $\gamma r\beta_{r+q_0}(\lambda)(1-\lambda g_{r+q_0}(\lambda)) \le g_r(\lambda)$, $0\le\lambda\le a$, $r\ge 0$, where $q_0=1$ for method M4 and $q_0=0$ for other methods.

Using these properties of function $g_r(\lambda)$, we separately estimate the right-hand terms of the equality (11). If $r\ge 0$, then due to the property c)

$$(\gamma r)^2\|B_{\varrho r}G_{\varrho r}A(u_0-u_*)\|^2=(\gamma r)^2\int_0^a \beta^2_{\varrho r}(\lambda)(1-\lambda g_{\varrho r}(\lambda))^2\lambda^2 d\langle P(\lambda)(u_0-u_*),u_0-u_*\rangle$$
$$\le\int_0^a(1-\lambda g_r(\lambda))^2 d\langle P(\lambda)(u_0-u_*),u_0-u_*\rangle=\|G_r(u_0-u_*)\|^2.$$

If $r\ge 1$, then $r+q_0\le\varrho r$ and properties b), d) give

$$(\gamma r)^2\|B_{\varrho r}G_{\varrho r}(\tilde f-f)\|^2=(\gamma r)^2\int_0^a \beta^2_{\varrho r}(\lambda)(1-\lambda g_{\varrho r}(\lambda))^2 d\langle P(\lambda)(\tilde f-f),\tilde f-f\rangle$$
$$\le\int_0^a g_r^2(\lambda)d\langle P(\lambda)(\tilde f-f),\tilde f-f\rangle=\|g_r(A)(\tilde f-f)\|^2= g_r(A)(f_0-f)\|^2.$$

If $r\ge 1$, then properties a)–d) and assumptions of Lemma 2 yield

$$-(\gamma r)^2\langle B_{\varrho r}G_{\varrho r}A(u_0-u_*),B_{\varrho r}G_{\varrho r}(\tilde f-f)\rangle=-(\gamma r)^2\int_0^a \beta^2_{\varrho r}(\lambda)(1-\lambda g_{\varrho r}(\lambda))^2\lambda d\langle P(\lambda)(u_0-u_*),\tilde f-f\rangle$$
$$\le(\gamma r)^2\int_0^a \beta^2_{\varrho r}(\lambda)(1-\lambda g_{\varrho r}(\lambda))^2\lambda d\langle P(\lambda)(u_0-u_*),f_0-f\rangle\le\int_0^a(1-\lambda g_r(\lambda))g_r(\lambda)d\langle P(\lambda)(u_0-u_*),f_0-f\rangle$$
$$=\langle G_r(u_0-u_*),g_r(A)(f_0-f)\rangle$$

Using now (11), (10) and the last estimates, we can estimate in case $r\ge 1$

$$(\gamma r)^2\|B_{\varrho r}G_{\varrho r}A(u_0-u_*)\|^2\le\|G_r(u_0-u_*)\|^2+2\langle G_r(u_0-u_*),g_r(A)(f_0-f)\rangle+\|g_r(A)(f_0-f)\|^2=\|u_r^0-u_*\|^2,$$

which proves the lemma.

Lemma 3. *In methods M1–M5 for each $r\in[r(\delta),r_2(\delta)]$ it holds the inequality*

$$\frac{\|u_r-u_{r(\delta)}\|}{\gamma r\|B_{\varrho r}(Au_{\varrho r}-\tilde f)\|}\le\frac{cd_*}{1-s}.$$

Proof. At first we estimate the quantity $\|u_r-u_{r(\delta)}\|$, where $r\in[r(\delta),r_2(\delta)]$. It is easy to show that for methods M1–M5 the function $g_r(\lambda)$ satisfies the following condition

$$g_r(\lambda)-g_{r-1}(\lambda)\le\gamma\beta_{r-1}(\lambda)(1-\lambda g_{r-1}(\lambda)) \tag{12}$$

Denote

$$\overline{r}=\begin{cases} r, & \text{if } r-r(\delta) \text{ is integer}\\ r(\delta)+\text{int}(r-r(\delta))+1, & \text{otherwise.}\end{cases} \tag{13}$$

Now using (12) and the equality $B_r(Au_r-\tilde f)=B_rG_r(Au_0-\tilde f)$, we can estimate

$$\|u_r-u_{r(\delta)}\|=\|(g_r(A)-g_{r(\delta)}(A))(\tilde f-Au_0)\|\le\|(g_{\overline{r}}(A)-g_{r(\delta)}(A))(\tilde f-Au_0)\|=\|u_{\overline{r}}-u_{r(\delta)}\|=$$
$$\Big\|\sum_{j=r(\delta)+1}^{\overline{r}}(u_j-u_{j-1})\Big\|\le\sum_{j=r(\delta)+1}^{\overline{r}}\|u_j-u_{j-1}\|\le\gamma\sum_{j=r(\delta)+1}^{\overline{r}}\|B_{j-1}(Au_{j-1}-\tilde f)\|.$$

By assumption of lemma from definition of d_* follows

$$(j-1)^s\|B_{j-1}(Au_{j-1}-\tilde{f})\| \le d_*(\varrho\overline{r})^s\|B_{\varrho\overline{r}}(Au_{\varrho\overline{r}}-\tilde{f})\|\,, \tag{14}$$

where $r(\delta) \le j-1 \le \overline{r}-1$, $r(\delta) \le \overline{r} \le \overline{r}_2$ and $\overline{r}_2$ is defined as in (13) (replace r by r_2).

Using (12), (14) and monotone decrease of the function $r \to \|B_r(Au_r-\tilde{f})\|$ we get for $r \in [r(\delta), r_2(\delta)]$

$$\frac{\|u_r-u_{r(\delta)}\|}{\gamma r\|B_{\varrho r}(Au_{\varrho r}-\tilde{f})\|} \le \frac{\sum\limits_{j=r(\delta)+1}^{\overline{r}}\|B_{j-1}(Au_{j-1}-\tilde{f})\|}{r\|B_{\varrho\overline{r}}(Au_{\varrho r}-\tilde{f})\|} \le d_*\varrho^s\frac{\overline{r}^s}{r}\sum_{j=r(\delta)+1}^{\overline{r}}\frac{1}{(j-1)^s} \le$$

$$d_*\varrho^s\frac{\overline{r}^s}{r}\left[\frac{1}{(r(\delta))^s}+\int\limits_{r(\delta)}^{\overline{r}-1}x^{-s}dx\right] \le d_*\varrho^s\frac{\overline{r}^s}{r}\left[\frac{1}{(r(\delta))^s}+\frac{1}{1-s}\overline{r}^{1-s}\right] \le d_*\varrho^s\left[2+\frac{2}{1-s}\right] \le \frac{cd_*}{1-s}\,,$$

which proves Lemma 3.

Proof of Theorem 2. Part I. At first we consider the case $\|\tilde{f}-f\| \le \max\{\delta,\delta_0\}$. Let r_* be a parameter for which the function $\psi(r)$ has a global minimum. To prove the estimate (8), we separately consider three cases: 1) $r_* \le r(\delta)$, 2) $r_* \ge r_2(\delta)$ and 3) $r(\delta) \le r_* \le r_2(\delta)$.

1. Consider the case $r_* \le r(\delta)$. If $\delta \ge \delta_0$ then from (7) follows that $\varphi(r(\delta)) \ge b_1\delta = b_1\max\{\delta, \|\tilde{f}-f\|\}$ and due to assertion c) of Lemma 1 we get

$$\|u_{r(\delta)}-u_*\| \le \psi(r(\delta)) \le c_3(b_1)\psi(r_*)\,. \tag{15}$$

If $\delta < \delta_0$ then $\|B_{r(\delta)}(Au_{r(\delta)}-\tilde{f})\| = 2\delta_0 \ge 2\max\{\delta, \|\tilde{f}-f\|\}$ and from assertion a) of Lemma 1 we get

$$\|u_{r(\delta)}-u_*\| \le \psi(r(\delta)) \le c_1(2)\psi(r_*)\,. \tag{16}$$

2. Now we consider the case $r_* \ge r_2(\delta)$. Taking into account the inequalities $\varphi(s') \le b_*\delta \le b_*\max\{\delta, \|\tilde{f}-f\|\}$ for each s', $s' \ge r_2(\delta)$, Lemma 1 gives

$$\|u_{r_2(\delta)}-u_*\| \le \psi(r_2(\delta)) \le c_4(b_*)\psi(r_*)\,.$$

Using now (15) and Lemmas 2 and 3, we can estimate

$$\|u_{r(\delta)}-u_*\| \le \|u_{r(\delta)}-u_{r_2(\delta)}\|+\|u_{r_2(\delta)}-u_*\| \le \frac{\|u_{r(\delta)}-u_{r_2(\delta)}\|}{\gamma r_2(\delta)\|B_{\varrho r_2(\delta)}(Au_{\varrho r_2(\delta)}-\tilde{f})\|}\|u^0_{r_2(\delta)}-u_*\|+\|u_{r_2(\delta)}-u_*\|$$

$$\le (1+cd_*/(1-s))\psi(r_2(\delta)) \le c_4(b_*)(1+cd_*/(1-s))\psi(r_*)\,. \tag{17}$$

3. Case $r(\delta) \le r_* \le r_2(\delta)$. Analogously to the previous case we have

$$\|u_{r(\delta)}-u_*\| \le \|u_{r(\delta)}-u_{r_*}\|+\|u_{r_*}-u_*\| \le$$

$$\frac{\|u_{r(\delta)}-u_{r_*}\|}{\gamma r_*\|B_{\varrho r_*}(Au_{\varrho r_*}-\tilde{f})\|}\|u^0_{r_*}-u_*\|+\|u_{r_*}-u_*\| \le (1+cd_*/(1-s))\psi(r_*)\,. \tag{18}$$

Now the error estimate (8) easily follows from (15)–(18).

Part II. Consider the case $\max\{\delta,\delta_0\| < \|\tilde{f}-f\| \le \frac{1}{2}\|B_1(Au_1-\tilde{f})\|$. Let r_0 be a parameter for which $\|B_{r_0}(Au_{r_0}-\tilde{f})\| = 2\|\tilde{f}-f\|$. Then Lemma 1 yields

$$\|G_{r_0}(u_0-u_*)\|+\gamma r_0\|\tilde{f}-f\| \le \max\{c_1(2), c_2(2)\}\inf_{r\ge 0}\psi(r)\,. \tag{19}$$

From equality $\|B_{r(\delta)}(Au_{r(\delta)}-\tilde{f})\| = 2\delta_0$ and the fact that the function $r \to \|B_r(Au_r-\tilde{f})\|$ is monotonically decreasing, follows the inequality $1 \le r_0 \le r(\delta)$. Remember that $r(\delta)$ is the global minimum point of the function $t(r) = r^s\|B_r(Au_r-\tilde{f})\|$ on the interval $[1, r_2(\delta)]$. Therefore

$$r_0^s\|B_{r_0}(Au_{r_0}-\tilde{f})\| \ge (r(\delta))^s\|B_{r(\delta)}(Au_{r(\delta)}-\tilde{f})\|\,,$$

from which follows

$$r(\delta) \le r_0(\|\tilde{f}-f\|/\delta_0)^{1/s}\,. \tag{20}$$

Using now (19), (20) and the monotone decrease of the function $r \to 1 - \lambda g_r(\lambda)$, we can estimate

$$\|u_{r(\delta)} - u_*\| \le \|G_{r(\delta)}(u_0 - u_*)\| + \gamma r(\delta)\|\tilde{f} - f\| \le \|G_{r_0}(u_0 - u_*)\| + \gamma r_0(\|\tilde{f} - f\|/\delta_0)^{1/s}\|\tilde{f} - f\|$$

$$\le (\|\tilde{f} - f\|/\delta_0)^{1/s}(\|G_{r_0}(u_0 - u_*)\| + \gamma r_0\|\tilde{f} - f\|) \le C(\|\tilde{f} - f\|/\delta_0)^{1/s} \inf_{r \ge 0} \psi(r),$$

which proves the estimate (9).

Remark 1. If the function $t(r) = r^s\|B_r(Au_r - \tilde{f})\|$ is monotonously increasing on the interval $[r(\delta), \varrho r_2(\delta) + 1]$, then $d_* \le 1/\varrho^s$. In most of numerical examples (see Table 2) we had $d_* \le 1$.

Remark 2. One can show that in methods M1, M2, M3 and M5 coefficient $c(b_1, b_*, d_*) \le 2.5$, if $b_1 = b_2 = 1.5C_m$, $b_* = b_2$, $d_* \le 1/\varrho^s$.

Remark 3. If the problem (1) is non-selfadjoint, regularization methods have form $u_r = G_r u_0 + g_r(A^*A)A^*\tilde{f}$ with $G_r = (I - A^*Ag_r(A^*A))$ (compare (2)). For the choice of the regularization parameter r the analogue of rule P can be used, where $s \in (0, 1/2)$ and notations C_m, $\varphi(r)$, $t(r)$ are replaced by $C_m = [\gamma_{p_0/(2p_0+2)}]^{1+1/p_0}$, $\varphi(r) = \sqrt{r}\|A^*G_r^{1/p_0}(Au_r - \tilde{f})\|$, $t(r) = r^s\|G_r^{1/(2p_0)}(Au_r - \tilde{f})\|$ respectively. For the corresponding rule the analogues of Theorems 1, 2 hold.

5. NUMERICAL EXPERIMENTS

The following Fredholm integral equations of the first kind

$$\int_a^b \mathcal{K}(t, s)u(s)ds = f(t), \quad a \le t \le b, \tag{21}$$

with $\mathcal{K} \in L^2([a, b], [a, b])$, $u \in L^2[a, b]$ were solved by the Lavrentiev method using the choice of the parameter by the rule P (with the parameters $b_2 = 1.5C_m = 0.75$, $b_1 = 1.5C_m = 0.74$, $s = 0.75$) and by the MD rule (with the parameters $b_2 = 1.40$, $b_1 = 1.38$).

Example 1. Kernel $\mathcal{K}(t, s) = [t + s)/2 + ts + 1/3$, exact solution $u(s) = 1$, right-hand term $f(t) = t + 7/12$, $a = 0$, $b = 1$.

Example 2. $\mathcal{K}(t, s) = 1/(\pi((s - t)^2 + 1))$, $u(s) = (1 - s^2)^2$, $a = -1$, $b = 1$.

Example 3.

$$\mathcal{K}(t, s) = \begin{cases} \pi^2 t(1 - s), & \text{if} \quad t \le s, \\ \pi^2 s(1 - t), & \text{if} \quad t > s, \end{cases}$$

$u(s) = 1$, $a = 0$, $b = 1$.

Example 4. $\mathcal{K}(t, s) = \exp(ts)$, $u(s) = 1$, $f(t) = (\exp(t) - 1)/t$, $a = 0$, $b = 1$.

Example 5. $\mathcal{K}(t, s) = ts$, $u(s) = s/2$, $f(t) = t/6$, $a = 0$, $b = 1$.

Example 6.

$$\mathcal{K}(t, s) = \begin{cases} t(1 - s), & \text{if} \quad t \le s, \\ s(1 - t), & \text{if} \quad t > s, \end{cases}$$

$u(s) = s - 2s^3 + s^4$, $f(t) = (3t - 5t^3 + 3t^5 - t^6)/30$, $a = 0$, $b = 1$.

Example 7.

$$\mathcal{K}(t, s) = \begin{cases} t(1 - s)(2s - s^2 - t^2), & \text{if} \quad t \le s, \\ s(1 - t)(2t - s^2 - t^2), & \text{if} \quad t > s, \end{cases}$$

$u(s) = 1$, $f(t) = (t - 2t^3 + t^4)/24$, $a = 0$, $b = 1$.

After the discretization of equation (21) we get

$$h\sum_{j=1}^{n} K_{ij}u_j = f_i, \quad i = 1, 2, \ldots, n. \tag{22}$$

where $K_{ij} = \mathcal{K}(t_i, s_j)$, $u_j = u(s_j)$, $h = \dfrac{b - a}{n}$, $s_i = a + h/2 + (i - 1)h$, $i = 1, 2, \ldots, n$. For discretization we used $n = 50$. In the case of the given right-hand side of (21) we used $f_i = f(t_i)$, in the contrary case the numbers f_i, $i = 1, 2, \ldots, n$ were computed by formula (22). For obtaining the approximate right-hand side $\tilde{f} = \{\tilde{f}_i\}_i^n$ the vector $f = \{f_i\}_i^n$ was randomly perturbed with relative noise $\|\tilde{f} - f\|/\|f\| = 10^{-(k+1)/2}$, $k = 1, 2, \ldots, 13$. The following norm in the space R^n was used: $\|v\|_{R^n} = \left(h\sum_{i=1}^{n} v_i^2\right)^{1/2}$. The approximate solution $u_r = \{u_i\}_i^n$ was computed as the solution of linear system $(r^{-1}\mathbf{I} + \mathbf{K})u_r = \tilde{f}$, where matrix $\mathbf{K} =$

Table 1. Averages and maximums of terms V_1, V_2, b_2/b_*, d_*, δ_0/δ

	Averages					Maximums				
$\|\tilde{f}-f\|/\delta$	V_1	b_*/b_2	d_*	V_2	δ_0/δ	V_1	b_*/b_2	d_*	V_2	δ_0/δ
1	0.91	1.00	0.64	0.99	0.81	2.38	1.00	2.20	2.80	14.33
3	0.87	1.00	0.93	0.94	2.17	2.38	1.00	2.20	2.80	42.99
5	0.87	1.01	0.93	0.94	3.56	2.38	1.22	2.20	2.80	71.65
10	0.88	1.03	1.06	0.95	7.04	2.38	1.97	2.20	2.80	143.31
20	0.98	1.14	1.23	1.06	13.73	4.68	3.00	3.12	5.27	286.61
50	1.26	1.22	1.58	1.36	33.55	12.61	5.06	6.02	14.07	716.52
100	1.45	1.18	1.95	1.57	66.33	15.60	5.08	9.30	16.60	1433.05
1000	5.79	1.09	3.57	6.45	577.70	106.31	2.48	35.89	125.07	14330.48

Table 2. Percentual characterization of experiments

	% of experiments for which			
$D=\|\tilde{f}-f\|/\delta$	$V_2 \le 1$	$b_*/b_2 \le 1$	$b_*/b_2 \le 2$	$d_* \le 1$
1	90.11	100.00	100.00	92.31
3	85.71	100.00	100.00	95.60
5	82.42	92.31	100.00	86.81
10	82.42	89.01	100.00	74.73
20	79.12	81.32	93.41	76.92
50	74.73	84.62	91.21	79.12
100	73.63	92.31	93.41	81.32
1000	67.03	80.22	96.70	85.71
All	79.40	89.97	96.84	84.07

(hK_{ij}) and $\mathbf{I}$ is the identity matrix. In rule P the function $\varphi(r) = r^{1/2}\|\mathbf{K}^{1/2}(r^{-1}\mathbf{I}+\mathbf{K})^{-2}(\mathbf{K}u_r - \tilde{f})\|_{R^n}$ was used and for the supposable noise level $\delta = \|\tilde{f}-f\|/D$ was taken, where values of D were 1, 3, 5, 10, 20, 50, 100, 1000. In the MD rule the actual noise level $\delta = \|\tilde{f}-f\|$ was used. The ratios

$$V_1 = \frac{\|u_{r(\delta)} - u_*\|}{\Phi_2(\|\tilde{f}-f\|)}, \quad V_2 = \frac{\|u_{r(\delta)} - u_*\|}{\|u_{r_{\mathrm{MD}}} - u_*\|}$$

with $\Phi_2(\|\tilde{f}-f\|) = \inf_{r\ge 0}\{\|(I - Ag_r(A))(u_0 - u_*)\|^2 + (\gamma_r\|\tilde{f}-f\|)^2]^{1/2}$ were computed. Note that V_1 characterizes the coefficient of the quasioptimality (compare with formula (5)).

The results of numerical experiments are given in Tables 1 and 2. The results show that in the case of the exactly estimated noise level ($D \equiv \|\tilde{f}-f\|/\delta = 1$) and in case $D \le 10$ the rule P works nearly as well as the MD rule or even better (average of ratio V_2 was 0.95). As expected, in the case of the essentialy underestimated noise level ($D \ge 20$) the error of the approximate solution for the rule P is larger than for the MD rule which uses the exact noise level. But the ratio of errors of these approximate solutions is relatively small in comparison with the error made by estimating the noise level. For example, if the noise level was $D = 100$ times smaller than the real value, the average of the error of the approximate solution for the rule P was only 57% larger than for the MD rule and for 73% of problems the rule P gave better results than the MD rule.

Note that in problems with smooth kernel (examples 1, 2, 4, 5) the rule P gave similar results as MD rule independently from D (maximum of V_2 was 1.56), but in problems with nonsmooth kernel (examples 3, 6, 7) V_2 depends more on D.

Table 2 shows that in most experiments we had $b_* = b_2$ (90% of all experiments, 100% of experiments with $D \le 3$) and $D_* \le 1$ (independently from D), therefore coefficients $C(b_1, b_*, d_*)$ in Theorem 2 are small. Table 2 indicates also that δ_0 increases with increase of D. Numerical examples also show that in most cases $r_2(\delta) \ge R(\delta)$.

6. CONCLUSION

For the choice of the regularization parameter r it is recommendable to use the noise level, while heuristic rules such as the L-curve rule, the GCV-rule etc do not guarantee the convergence of the approximations.

If the noise level is given only approximately and inequality $\|\tilde{f}-f\| \leq \delta$ is not guaranteed, the discrepancy principle and its modification are unstable. If δ with $\|\tilde{f}-f\|/\delta \leq$ const for $\delta \rightarrow 0$ is given, we recommend to use our rules P and P', guaranteeing convergence and in case $\|\tilde{f}-f\| \leq \delta$ also quasioptimal error estimates.

Note that by increasing parameter $s \in (0,1)$ the error estimate (8) increases and estimate (9) decreases. Therefore, if we are almost sure in inequality $\|\tilde{f}-f\| \leq \delta$, smaller values of s are recommended. Numerical examples show that it is reasonable to take the parameter s in the interval $[0.6, 0.8]$.

Acknowledgement
This work was supported by the Estonian Science Foundation, Grant No. 5785.

REFERENCES

1. A. Bakushinskii, Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *U.S.S.R. Comput. Math. Math. Phys.* (1984) **24**, 181–182.
2. D. Calvetti, S. Morigi, L. Reichel and F. Sgallari, Tikhonov regularization and the L-curve for large discrete ill-posed problems. Numerical analysis 2000, Vol. III. Linear algebra. *J. Comput. Appl. Math.* (2000) **123** (1–2), 423–446.
3. G. H. Golub and Urs. von Matt, Generalized cross-validation for large-scale problems. *J. Comput. Graph. Statist.* (1997) **6** (1), 1–34.
4. A. I. Grebennikov, *Spline Approximation Method and its Applications*, MAX Press, Moscow, 2004.
5. U. Hämarik and T. Raus, On regularization parameter choice for solving self-adjoint ill-posed problems with approximately given noise level of data. *Proc. Estonian Acad. Sci. Phys. Math.* (2004) **53** (2), 77–83.
6. U. Hämarik and T. Raus, On the regularization parameter choice in case of approximately given error bounds of data, *ENUMATH2003 – Proceedings of the 5th European Conference on Numerical Mathematics and Advanced Applications*, (eds. M. Feistauer, V. Dolejši, P. Knobloch, K. Najzar), Springer, Berlin, 2004, pp. 400–409.
7. M. Hanke and T. Raus, A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM J. Sci. Comput.* (1996) **17**, 954–972.
8. P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* (1992) **34**, 561–580.
9. V. Morozov, On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.* (1966) **7**, 414–417.
10. T. Raus, Residual principle for ill-posed problems. *Acta Comment. Univ. Tartuensis* (1984) **672**, 16–26 (in Russian).
11. T. Raus, An a posteriori choice of the regularizartion parameter in case of approximately given error bound of data. *Acta Comment. Univ. Tartuensis* (1990) **913**, 73–87.
12. T. Raus, About regularization parameter choice in case of approximately given error bounds of data. *Acta Comment. Univ. Tartuensis* (1992) **937**, 77–89.
13. A. N. Tikhonov and V. B. Glasko, Use of the regularization method in non-linear problems. *USSR Comput. Math. Math. Phys.* (1965) **5** (3), 93–107.
14. A. Tikhonov and V. Arsenin, *Solution of Ill-Posed Problems.* Wiley, New York, 1977.
15. G. Vainikko, The discrepancy principle for a class of regularization methods. *U.S.S.R. Comput. Math. Math. Phys.* (1982) **22** (3), 1–19.
16. G. Vainikko and A. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Nauka, Moscow, 1986 (in Russian).
17. G. Wahba, Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* (1977) **14**, 651–657.