On the efficient discretization of integral equations of the third kind

Sergei V.Pereverzev, Eberhard Schock, and Sergei G.Solodky

Abstract

We propose a new discretization scheme for solving ill-posed integral equations of the third kind. Combining this scheme with Morozov's discrepancy principle for Landweber iteration we show that for some classes of equations in such method a number of arithmetic operations of smaller order than in collocation method is required to approximately solve an equation with the same accuracy.

1 Introduction

In his fundamental papers on integral equations D.Hilbert [5] introduced the notion of integral equations of the first, second, and of the third kind. A linear integral equation

$$rx + Kx \equiv r(t)x(t) + \int_0^1 k(t,\tau)x(\tau)d\tau = y(t)) \tag{1}$$

is said to be of the first kind if $r \equiv 0$, of the second kind if r is a non-zero constant, and of the third kind if r is a function with zeros in its domain, otherwise the equation is equivalent to an equation of the second kind. If the function r is continuous and has a finite number of zeros, then the equation (1) is a special case of non-elliptic singular integral equations investigated by S.Prössdorf [11]. For above-mentioned function r with known zeros the approximate methods for solving integral equations (1) were proposed by N.S.Gabbasov (see, for example, [3]). But these methods are completely unusable if r is, for example, a characteristic function of a proper subset of positive measure. Moreover, as indicated in [12], if for each neighbourhood V of zero the inverse $r^{-1}(V)$ of V has positive measure, then the problem of solving the equation (1) is not well posed in the sense of J.Hadamard and regularization techniques are required for solving (1). In our opinion it makes sense to apply the regularization methods, even though the function r has a finite number of zeros but we do not know theirs location.

Usually, the application of some regularization method is preceded by the discretization of the problem and there is a close connection between an amount of discrete information and the choice of the regularization parameter. The aim of this paper is to discuss this connection for the approximate solution of ill-posed equations (1). Moreover, some estimate for the number of arithmetic operations required in order to reach fixed accuracy ε will be obtained too.

2 The discretization scheme

Throughout this paper we shall consider the integral equations (1) with operators K acting continuously from L_2 to the Sobolev space W_2^1 and with $y \in W_2^1$, where L_2 is the Hilbert space of square-summable functions on $[0,1]$ with the usual norm $\|\cdot\|$ and the usual inner product $(\cdot,\cdot)$, and W_2^1 is the normed space of functions $f(t)$ having square-summable derivatives $f' \in L_2$. Therewith

$$\|f\|_{W_2^1} = \|f\| + \|\frac{d}{dt}f\|.$$

Moreover, it will be assumed that the operators K have some additional smoothness. Namely,

$$K \in \mathcal{K}_\gamma^1 := \Big\{K : \|K\|_{L_2\to W_2^1} \le \gamma_1, \quad \|K^*\|_{L_2\to W_2^1} \le \gamma_2,$$
$$\|(\tfrac{d}{dt}K)^*\|_{L_2\to W_2^1} \le \gamma_3\Big\}, \quad \gamma = (\gamma_1,\gamma_2,\gamma_3),$$

where $\|\cdot\|_{X\to Y}$ is the usual norm in the space of all linear bounded operators from X into Y; B^* denotes the adjoint operator of $B: L_2 \to L_2$. If the kernel $k(t,\tau)$ of the integral operator K has mixed partial derivatives and

$$\int_0^1\int_0^1 \Big[\frac{\partial^{\imath+\jmath}k(t,\tau)}{\partial t^\imath \partial \tau^\jmath}\Big]^2 dt d\tau < \infty, \quad \imath,\jmath = 0,1,$$

then it is easy to see that $K \in \mathcal{K}_\gamma^1$ for some γ.

Let us consider the Haar orthonormal basis $\chi_1, \chi_2, \ldots, \chi_m, \ldots$ of piecewise constant functions, where $\chi_1(t) \equiv 1$, and for $m = 2^{k-1}+j$, $k = 1,2,\ldots; j = 1,2,\ldots,2^{k-1}$

$$\chi_m(t) = \begin{cases} 2^{(k-1)/2}, & t \in [(j-1)/2^{k-1}, (j-1/2)/2^{k-1}) \\ -2^{(k-1)/2}, & t \in [(j-1/2)/2^{k-1}, j/2^{k-1}) \\ 0, & t \notin [(j-1)/2^{k-1}, j/2^{k-1}] \end{cases},$$

and let P_m be the orthogonal projector on $span\{\chi_1, \chi_2, \ldots, \chi_m\}$, that is,

$$P_m f(t) = \sum_{i=1}^{m} (f, \chi_\imath)\chi_\imath(t).$$

It is well known that [6],p.81,82

$$\|I - P_m\|_{W_2^1 \to L_2} \le cm^{-1}, \tag{2}$$

where I is the identity operator and c is some absolute constant. Moreover, if $|r'(t)| \le d$ then for any $t \in [0,1]$

$$|r(t) - P_m r(t)| \le 3dm^{-1}. \tag{3}$$

To construct an efficient method for discretizing ill-posed equations (1) we shall use a specific "hyperbolic cross" approximation of the kernel function $k(t,\tau)$. This means that instead of (1) we consider now the equation

$$xP_{2^n} r + K_n x = P_{2^n} y, \tag{4}$$

where

$$K_n = \sum_{k=1}^{n} (P_{2^k} - P_{2^{k-1}})KP_{2^{n-k}} + P_1 K P_{2^n} = \sum_{(\imath,\jmath)\in\Gamma_n} \chi_\imath(\chi_\imath, K\chi_\jmath)(\chi_\jmath, \cdot),$$

$$\Gamma_n = \{1\} \times [1, 2^n] \bigcup_{k=1}^{n} (2^{k-1}, 2^k] \times [1, 2^{n-k}].$$

It is obvious that K_n is the integral operator with degenerate kernel

$$k_n(t,\tau) = \sum_{(\imath,\jmath)\in\Gamma_n} \hat{k}(\imath,\jmath)\chi_\imath(t)\chi_\jmath(\tau),$$

where $\hat{k}(\imath,\jmath)$ denotes the Fourier coefficients of function $k(t,\tau)$ with respect to Haar system, i.e.

$$\hat{k}(\imath,\jmath) = \int_0^1\int_0^1 k(t,\tau)\chi_\imath(t)\chi_\jmath(\tau)dtd\tau.$$

Let now $card(\Gamma_n)$ be the number of Fourier coefficients $\hat{k}(\imath,\jmath)$ required to construct $k_n(t,\tau)$. It is easily verified that

$$card(\Gamma_n) \asymp n2^n.$$

As usual, we write $T(u) \asymp S(u)$ if there are constants c, c_1 such that for all u belonging to the domain of definition $T(u), S(u)$

$$cT(u) \leq S(u) \leq c_1 T(u).$$

Moreover, for simplicity we often use the same symbol c for possibly different constants.

If we denote by N_{disc} the number of all Fourier coefficients

$$\hat{k}(\imath,\jmath) = (\chi_\imath, K\chi_\jmath), \quad \hat{r}(\imath) = (r,\chi_\imath), \quad \hat{y}(\imath) = (y,\chi_\imath) \tag{5}$$

taking part in the definition of the equation (4), then

$$N_{disc} \asymp 2^{n+1} + card(\Gamma_n) \asymp n2^n. \tag{6}$$

The direct solution of (4) by means of some exact solution method for system of 2^n linear algebraic equations would take too many arithmetic operations, even when we assume for the moment that the solution of (4) exists and is unique. A more favorable way is the use of regularization methods which are degenerated by iteration procedures. In this paper we will consider the Landweber iteration

$$x_{m,n} = x_{m-1,n} - \mu B_n^*(B_n x_{m-1,n} - P_{2^n}y), \quad m = 1,2,\ldots, x_{0,n} = 0, \tag{7}$$

where

$$B_n f = fP_{2^n}r + K_n f, \quad B_n^* f = fP_{2^n}r + K_n^* f,$$
$$0 < \mu < 2/\|B_n\|^2_{L_2\to L_2}.$$

Further examples of iterative methods are discussed in [13].

The number of iteration steps m acts as a regularization parameter and the usual discussion of rates of convergence of iterative methods for ill-posed equations is done under the assumption, that the exact solution $\bar{x}$ of (1) belongs to the range of operatpr $|B|^p$ for some $p \geq p_0$, where $|B|^p = (B^*B)^{p/2}$ and

$$Bf(t) = r(t)f(t) + Kf(t).$$

Therefore from now on we assume that the exact solution of (1) fulfills the smoothness property

$$\bar{x} = |B|^p v, \quad \|v\| \leq \rho \tag{8}$$

for some $p \geq p_0 \geq 1$, and $K \in \mathcal{K}^1_\gamma, \quad |r'(t)| \leq d$,

$$y \in W^1_{2,1} = \Big\{ f : f \in W^1_2, \quad \|f\|_{W^1_2} \leq 1 \Big\}.$$

In the following for class of such equations (1) we use the notation $\Phi^{p_0}_{\gamma,\rho}$ (of course, ρ depends on γ and d).

In what follows we need

Lemma 2.1 *Let $K \in \mathcal{K}^1_\gamma$ and $|r'(t)| \leq d$. Then*

$$\|B - B_n\|_{L_2 \to L_2} \leq cn2^{-n},$$

where constant c depends on γ and d.

Proof: From the definion of operator K_n we find

$$P_{2^n}K - K_n = \sum_{k=1}^{n} (P_{2^k} - P_{2^{k-1}})K(I - P_{2^{n-k}}) + P_1 K(I - P_{2^n}).$$

With an argument like that in the proof of Lemma 3.2 of [8] for $K \in \mathcal{K}^1_\gamma$ we get the estimate

$$\|(P_{2^k} - P_{2^{k-1}})K(I - P_{2^{n-k}})\|_{L_2 \to L_2} \leq c2^{-n}.$$

Then by virtue of (2) we have

$$\|K - K_n\|_{L_2 \to L_2} \leq \|(I - P_{2^n})K\|_{L_2 \to L_2} + \|P_{2^n}K - K_n\|_{L_2 \to L_2} \leq$$

$$\leq c2^{-n}\|K\|_{L_2 \to W^1_2} + \sum_{k=1}^{n} \|(P_{2^k} - P_{2^{k-1}})K(I - P_{2^{n-k}})\|_{L_2 \to L_2} +$$

$$+ \|P_1 K(I - P_{2^n})\|_{L_2 \to L_2} \leq c2^{-n}\gamma_1 + cn2^{-n} + \|(I - P_{2^n})K^*\|_{L_2 \to L_2} \leq$$

$$\leq c2^{-n}(\gamma_1 + \gamma_2) + cn2^{-n} \leq cn2^{-n}.$$

Using this bound and (3) we obtain the estimate

$$\|B - B_n\|_{L_2 \to L_2} \le \max_{0 \le t \le 1} |r(t) - P_{2^n} r(t)| + \|K - K_n\|_{L_2 \to L_2} \le$$

$$\le 3d2^{-n} + cn2^{-n} \le cn2^{-n},$$

as claimed.

An appropriate discretization (4) and the number of iteration steps m in dependence of the predetermined order of accuracy $O(\varepsilon)$ for $\|\bar{x} - x_{m,n}\|$ have to be chosen. One of the most widely used strategies for choosing regularization parameter m which are also called "stopping rules" in literature is Morozov's discrepancy principle. We shall consider this discrepancy principle in the form tailored to the discretized version of Landweber iteration (7) for equations (1) from $\Phi^{p_0}_{\gamma,\rho}$: Let $d_1 > 1$. A stopping rule for (7) is given by choosing the first integer m such that $m \le m_{\max} \asymp \varepsilon^{-2/p_0}$ and

$$\|P_{2^n} f - B_n x_{m,n}\| \le d_1 \varepsilon^{\frac{p_0+1}{p_0}}. \tag{9}$$

If there is no $m \le m_{\max}$, such that (9) holds, then choose $m = [m_{\max}] + 1$, $[m_{\max}]$ denoting the largest integer which is not greater than $m_{\max} \asymp \varepsilon^{-2/p_0}$.

Now we can state the main result.

Theorem 2.1 *Let $n2^{-n} \asymp \varepsilon^{\frac{p_0+1}{p_0}}$, and let the number of iteration steps m in (7) be chosen according to the discrepancy principle (9). If equation (1) belongs to the class $\Phi^{p_0}_{\gamma,\rho}$, $p_0 \ge 1$, then*

$$\|\bar{x} - x_{m,n}\| = O(\varepsilon).$$

Proof. The regularization method (7) is generated by function

$$g_m(\lambda) = \lambda^{-1}[1 - (1 - \mu\lambda)^m], \quad \lambda > 0.$$

Namely, $x_{m,n} = R_{m,n} P_{2^n} y$, where

$$R_{m,n} = g_m(B_n^* B_n) B_n^*.$$

We put $S_{m,n} = I - R_{m,n} B_n$. From [10] one sees that

$$\|R_{m,n}\|_{L_2 \to L_2} \le c_1 m^{1/2}, \quad \|S_{m,n}\|_{L_2 \to L_2} \le c_2,$$

$$\|I - B_n R_{m,n}\|_{L_2 \to L_2} \le 1, \quad \|S_{m,n} |B_n|^p\|_{L_2 \to L_2} \le c_{1,p} m^{-p/2}, \tag{10}$$

$$\|B_n S_{m,n} |B_n|^p\|_{L_2 \to L_2} \le c_{2,p} m^{-\frac{(p+1)}{2}}.$$

Using (2), (10) and Lemma, from the definition $x_{m,n}$ we find

$$\begin{gathered}\|\bar{x}-x_{m,n}\| = \|S_{m,n}\bar{x} + R_{m,n}[(I-P_{2^n})y - (B-B_n)\bar{x}]\| \le \\ \le \|S_{m,n}\bar{x}\| + cm^{1/2}(\|(I-P_{2^n})y\| + \|B-B_n\|_{L_2\to L_2}) \le \\ \le \|S_{m,n}\bar{x}\| + cm_{\max}^{1/2}(2^{-n} + n2^{-n}) \le \\ \le \|S_{m,n}\bar{x}\| + c\varepsilon^{-1/p_0}\varepsilon^{\frac{p_0+1}{p_0}} \asymp \|S_{m,n}\bar{x}\| + \varepsilon.\end{gathered} \tag{11}$$

Let us estimate $\|S_{m,n}\bar{x}\|$. Using the inequality

$$\||B|^p - |B_n|^p\|_{L_2\to L_2} \le c\|B-B_n\|_{L_2\to L_2}^{\min\{1,p\}} |\ln(\|B-B_n\|_{L_2\to L_2})|$$

(see [14], p.93), (8), (10) and Lemma 2.1 for $p \ge p_0 \ge 1$ we have

$$\begin{gathered}\|S_{m,n}\bar{x}\| \le \|S_{m,n}|B_n|^p v\| + \|S_{m,n}(|B|^p - |B_n|^p)v\| \le \\ \le c_{1,p}\rho m^{-p/2} + c_2\rho n^2 2^{-n} \le c(m^{-p/2} + \varepsilon^{\frac{p_0+1}{p_0}} \ln\tfrac{1}{\varepsilon}).\end{gathered} \tag{12}$$

If $m > \varepsilon^{-2/p}$, the assertion of the theorem follows from (11), (12).

Assume now that $m < m_1 = [\varepsilon^{-2/p}] + 1$. With an argument like that in the proof of Theorem 3.3 of [10] we get the estimate

$$\|S_{m,n}\bar{x}\|^2 \le c\Big(\|S_{m_1,n}\bar{x}\|^2 + m_1\|B_nS_{m,n}\bar{x}\|^2\Big). \tag{13}$$

On the other hand, from (2), (9), (10) and Lemma 2.1 we know that

$$\begin{gathered}\|B_nS_{m,n}\bar{x}\| \le \|P_{2^n}y - B_nx_{m,n}\| + \|I - B_nR_{m,n}\|_{L_2\to L_2}\times \\ \times\Big(\|(B_n-B)\bar{x}\| + \|y - P_{2^n}y\|\Big) \le d_1\varepsilon^{\frac{p_0+1}{p_0}} + c(n2^{-n} + 2^{-n}) \asymp \\ \asymp \varepsilon^{\frac{p_0+1}{p_0}}.\end{gathered}$$

Moreover, using the inequality (12) we obtain

$$\|S_{m_1,n}\bar{x}\|^2 \le c\Big(m_1^{-p/2} + \varepsilon^{\frac{p_0+1}{p_0}} \ln\frac{1}{\varepsilon}\Big)^2 \le c\varepsilon^2.$$

Thus, from (13) one sees that for $m < m_1$ and $p \ge p_0$

$$\|S_{m,n}\bar{x}\|^2 \le c\varepsilon^2 + m_1\varepsilon^{\frac{2(p_0+1)}{p_0}} \le c(\varepsilon^2 + \varepsilon^{2-\frac{2}{p}+\frac{2}{p_0}}) \asymp \varepsilon^2.$$

Combining this estimate and (11) for $m < m_1$ we have

$$\|\bar{x} - x_{m,n}\| = O(\varepsilon).$$

The theorem is proved.

Corollary 2.1 *Let N_{disc} be an amount of discrete information (5) required to construct an approximate solution $x_{m,n}$. From the Theorem 2.1 and (6) it follows that within the framework of discretization scheme (4) we can guarantee on the class $\Phi^{p_0}_{\gamma,\rho}$ the order of accuracy ε in the case when*

$$N_{disc} \asymp \varepsilon^{-\frac{p_0+1}{p_0}} \ln^2 \frac{1}{\varepsilon}.$$

3 Complexity of the algorithm

Let us estimate the number N_{op} of arithmetic operations on the values of Fourier coefficients (5) required to construct an approximate solution $x_{m,n}$.

Proposition 3.1 *Let*

$$g(t) = \sum_{\imath=1}^{2^n} g_\imath \chi_\imath(t)$$

be an arbitrary element of subspace $span\{\chi_1, \chi_2, \ldots, \chi_{2^n}\}$. To represent an element

$$f(t) = g(t) P_{2^n} r(t) \in span\{\chi_1, \chi_2, \ldots, \chi_{2^n}\}$$

in the standard form

$$f(t) = \sum_{\imath=1}^{2^n} f_\imath \chi_\imath(t) \tag{14}$$

it suffices to perform no more than $c2^n$ arithmetic operations on the coefficients $g_\imath$ and $\hat{r}(\imath)$.

Proof. Note that $g(t), P_{2^n} r(t)$ and $f(t)$ are the constants on the dyadic intervals

$$\Delta_{n,\imath} = \Big(\frac{\imath - 1}{2^n}, \frac{\imath}{2^n}\Big), \quad \imath = 1, 2, \ldots, 2^n.$$

Keeping in mind that (see [6],p.78)

$$P_{2^n} \varphi(t) = 2^n \int_{\Delta_{n,\imath}} \varphi(\tau) d\tau, \quad t \in \Delta_{n,\imath},$$

for any $t \in \Delta_{n,\imath}, \imath = 1, 2, \ldots, 2^n$, we have

$$\begin{aligned} P_{2^n} f(t) &= 2^n \textstyle\int_{\Delta_{n,\imath}} g(\tau) P_{2^n} r(\tau) d\tau = \\ &= 2^n \textstyle\int_{\Delta_{n,\imath}} g(\tau) d\tau \cdot 2^n \int_{\Delta_{n,\imath}} P_{2^n} r(\tau) d\tau = \\ &= P_{2^n} g(t) P_{2^n} r(t) = g(t) P_{2^n} r(t) = f(t). \end{aligned} \tag{15}$$

Thus, $f(t) = P_{2^n} f(t) \in span\{\chi_1, \chi_2, \ldots, \chi_{2^n}\}$.

Let us denote by h_{kj}, $k = 1, 2, \ldots, n$, $j = 1, 2, \ldots, 2^{k-1}$, the Haar functions $\chi_2, \chi_3, \ldots, \chi_{2^n}$, labeled by two indices. Namely, for $m = 2^{k-1} + j$

$$\chi_m(t) = h_{k,j}(t).$$

Then for any $\varphi \in L_2$ we put $\varphi(k, j) = \hat{\varphi}(2^{k-1} + j)$. We also introduce the averages $\bar{\varphi}(k, j)$ of $\varphi(t)$ on $\Delta_{k,j}$ as

$$\bar{\varphi}(k, j) = |\Delta_{k,j}|^{-1} \int_{\Delta_{k,j}} \varphi(\tau) d\tau,$$

where $|\Delta_{k,j}|$ denotes the length of $\Delta_{k,j}$. It is well known (see, for example, [6], p.78) that

$$\begin{aligned} \bar{\varphi}(1,1) &= \hat{\varphi}(1) + \varphi(1,1), \\ \bar{\varphi}(1,2) &= \hat{\varphi}(1) - \varphi(1,1), \end{aligned} \tag{16}$$

and further

$$\begin{aligned} \bar{\varphi}(m, 2j-1) &= \bar{\varphi}(m-1, j) + 2^{(m-1)/2} \varphi(m, j), \\ \bar{\varphi}(m, 2j) &= \bar{\varphi}(m-1, j) - 2^{(m-1)/2} \varphi(m, j), \\ m = 2, 3, \ldots, n; &\quad j = 1, 2, \ldots, 2^{m-1}. \end{aligned} \tag{17}$$

It is easy to see that using (16), (17) we can compute the averages $\bar{g}(n, i), \bar{r}(n, i)$, $i = 1, 2, \ldots, 2^n$, of functions $g(t), r(t)$, and evaluating the whole set of these averages requires no more than $c2^n$ arithmetic operations on the coefficients $g_j, \hat{r}(j)$.

If the averages $\bar{g}(n, i), \bar{r}(n, i)$ are known then by virtue of (15)

$$\bar{f}(n, i) = \bar{g}(n, i)\bar{r}(n, i), \quad i = 1, 2, \ldots, 2^n,$$

and evaluating the whole set of $\bar{f}(n, i)$ requires 2^n multiplications. Now according to the method for calculating the Haar coefficients [1] the rest of the averages $\bar{f}(m, j)$ and the Fourier coefficients $f(m, j)$ can be computed from the formulas

$$\bar{f}(m-1, j) = \tfrac{1}{2}(\bar{f}(m, 2j-1) + \bar{f}(m, 2j)),$$

$$f(m, j) = 2^{-(m+1)/2}(\bar{f}(m, 2j-1) - \bar{f}(m, 2j)),$$

$$m = n, n-1, \ldots, 2; \quad j = 1, 2, \ldots, 2^{m-1},$$

$$f(1,1) = \tfrac{1}{2}(\bar{f}(1,1) - \bar{f}(1,2)),$$

$$\hat{f}(1) = \tfrac{1}{2}(\bar{f}(1,1) + \bar{f}(1,2)).$$

One can see that evaluating the whole set of averages and Fourier coefficients requires $2^{n+1}-2$ additions and 2^{n+1} multiplications. To complete the proof it only remains for us to note that in representation (14) $f_1=\hat{f}(1)$ and for $\imath=2^{k-1}+\jmath$

$$f_\imath=f(k,\jmath), \quad k=1,2,\ldots,n, \quad j=1,2,\ldots,2^{k-1}.$$

With an argument like that in the proof of Lemma 18.2 of [9],p.300 we get the following assertion.

Proposition 3.2 *Let $g(t)$ be an arbitrary element of $span\{\chi_1,\chi_2,\ldots,\chi_{2^n}\}$. To represent the elements $K_n g, \quad K_n^* g\in span\{\chi_1,\chi_2,\ldots,\chi_{2^n}\}$ in the standard form (14) it suffices to perform no more than $cn2^n$ arithmetic operations on the coefficients $\hat{g}(\imath)$ and $\hat{k}(\imath,\jmath)$.*

Theorem 3.1 *In conditions of Theorem 2.1 we can guarantee on the class $\Phi^{p_0}_{\gamma,\rho}$ the order of accuracy ε for*

$$N_{op}=O\Big(\varepsilon^{-\frac{p_0+3}{p_0}}\ln^2\frac{1}{\varepsilon}\Big). \tag{18}$$

Proof. By virtue of (7) for any $m=1,2,\ldots,$ we have

$$x_{m,n}=x_{m-1,n}-\mu\delta_{m-1}P_{2^n}r-\mu K_n^*\delta_{m-1},$$

$$\delta_{m-1}=x_{m-1,n}P_{2^n}r+K_n x_{m-1,n}-P_{2^n}y.$$

From the definition of operator K_n and (15) one sees that $x_{m,n}\in span\{\chi_1,\chi_2,\ldots,\chi_{2^n}\}$ for any m. Let $card(AO)$ be the number of arithmetic operations required for the passage from $x_{m-1,n}$ to $x_{m,n}$. From the Theorem 2.1 and Propositions 3.1, 3.2 it follows that

$$card(AO)\le cn2^n\asymp\varepsilon^{-\frac{p_0+1}{p_0}}\ln^2\frac{1}{\varepsilon}.$$

On the other hand, within the framework of stopping rule (9)

$$N_{op}\le m_{\max}card(AO)\asymp\varepsilon^{-\frac{2}{p_0}}card(AO)\asymp$$

$$\asymp\varepsilon^{-\frac{p_0+3}{p_0}}\ln^2\tfrac{1}{\varepsilon},$$

as claimed.

Remark. Let us assume that the equation (1) belongs to $\Phi^{p_0}_{\gamma,\rho}$ but the function $r(t)$ has a finite number of known zeros. In this case the collocation method proposed in

[3] can be applied. Within the framework of this method finding the approximate solution x_n of (1) reduces to solving a system of $O(n)$ linear algebraic equations. Moreover, from the Theorem 1 [3] it follows that

$$\|\bar{x} - x_n\| = O\Big(\frac{1}{\sqrt{n}}\Big).$$

Then for guaranteeing accuracy ε it is necessary to solve the system consisting of $n \asymp \varepsilon^{-2}$ algebraic equations. To solve this system, for example, by Gaussian elimination it is necessary to perform $N_1 \asymp n^3 \asymp \varepsilon^{-6}$ arithmetic operations. When N_1 is compared with estimation (18) it is apparent that for the class $\Phi^{p_0}_{\gamma,\rho}$ the scheme (4), (7) with stopping rule (9) is more economical than collocation method of [3].

4 Differential Equations and Integral Equations of the Third Kind

Integral equations of the third kind are closely related to some singular problems in differential equations.

4.1 Volterra Equations

Let A, B be (n, n)-matrices with entries a_{jk}, b_{jk} and c an n-vector with entries c_j, which are continuous resp. differentiable real or complex functions.

The system of linear ordinary differential equations

$$Ay' = By + c$$

is a system of differential-algebraic equations (see e.g. [4]), if the matrix A is singular. On the other hand, since

$$(Ay)' = A'y + Ay',$$

we have

$$(Ay)(t) = \int_0^t (A'(\tau) + B(\tau))\, y(\tau) d\tau + c(t).$$

This is a system of Volterra equations of the third kind.

In the special case $n = 1$, Kress ([7], p.34) has shown, that a Volterra integral equation of the first kind is equivalent to a Volterra integral equation of the second kind, if the kernel function does not vanish on the diagonal ($k(\tau, \tau) \neq 0$ for all τ). If the kernel function has zeros on the diagonal, then this equation is equivalent to a Volterra equation of the third kind.

4.2 Fredholm Equations

Let L be a linear differential operator with a continuous inverse T, let A, B, c be as above. Then the boundary problem

$$L(Ay) = By + c$$

is equivalent to the system of integral equations of the third kind

$$Ay = TBy + Tc.$$

In the case $n = 1$ and if a_{11} has zeros, then we have boundary value problems with "regular" and with "irregular" singularities (see e.g. [2], p.299).

References

[1] Beylkin, G., Coifman, R.,and Rokhlin, V. (1991), Fast wavelet transforms and numerical algorithms I, *Comm.Pure Appl.Math.* **44**, 141–183.

[2] Boyce, W.E., DiPrima, R.C. (1995), *Elementary Differential Equations*, Wiley&Sons, New York.

[3] Gabbasov, N.S. (1991), New variants of collocation method for integral equations of the third kind, *Mat.Zametki* **50**, 47–53.

[4] Griepentrog, E., März, R. (1986), *Differential-Algebraic Equations and Their Numerical Treatment*, Teubner, Leipzig.

[5] Hilbert, D. (1912), *Grundzüge einer allgemeinen Theorie der linear Integralgleichungen*, Leipzig.

[6] Kashin, B.S., and Saakian, A.A. (1984), *Orthogonal series*, Nauka, Moscow.

[7] Kress, R. (1989), *Linear Integral Equations*, Springer, Heidelberg.

[8] Pereverzev, S.V. (1995), Optimization of Projection Methods for solving Ill- Posed Problems, *Computing* **55**, 113–124.

[9] Pereverzev, S.V. (1996), *Optimization of Methods for Approximate Solution of Operator Equations*, Nova, New York.

[10] Plato, R., and Vainikko, G. (1990), On the Regularization of Projection Methods for solving Ill-posed Problems, *Numer.Math.* **57**, 63–70.

[11] Prössdorf, S. (1974), *Einige Klassen singulärer Gleichungen*, Akademie-Verlag, Berlin.

[12] Schock, E. (1985), Integral equations of the third kind, *Studia Mathematica* **81**, 1–11.

[13] Schock, E. (1988), Poinwise rational approximation and iterative methods for ill-posed problems, *Numer.Math.* **54**, 91-103.

[14] Vainikko, G.M., and Veretennikov, A.Yu. (1986), *Iteration procedures in ill-posed problems*, Nauka, Moscow.

A The finite-dimensional realization of the general residue principle for Tikhonov regularization

In this appendix the discretized version of the general residue principle is studied and within the framework of this principle the discretization scheme of the form (4) is used. One shows that this scheme is more economical for some classes of equations than standard Galerkin scheme.

A.1 Introduction

Let H be a real Hilbert space with the norm $\|\cdot\|$ and inner product $(\cdot,\cdot)$. In this appendix we study the schemes of discretization for solving ill-posed problems of the form

$$Tx = f, \tag{19}$$

where T is a linear compact operator from H into H and the free term f belongs to the $Range(T) := \{f : f = Tg, \ g \in H\}$, i.e. Eq. (19) is solvable. However, as a rule, instead of the free term f we have some approximation $f_\delta \in H$ such that $\|f - f_\delta\| \le \delta$, where δ is a small positive number which is usually known.

To get an approximation to a solution of (19) we have to discretize the problem. The traditional approach to such discretization lies in the following. We choose some finite-dimensional operator T_{disc} such that $rankT_{disc} = N$ and $\|T - T_{disc}\|_{H\to H} \le \varepsilon_N$, where ε_N depends on δ. Further, as the approximate solution of (19) we take the minimizer x_α^{disc} of the so-called Tikhonov functional

$$\Omega_\alpha(T_{disc}, x) = \|T_{disc}x - f_\delta\|^2 + \alpha\|x\|^2, \tag{20}$$

where α is the regularization parameter depending on δ. We may define x_α^{disc} from the Euler equation for (20)

$$\alpha x + T_{disc}^* T_{disc} x = T_{disc}^* f_\delta, \tag{21}$$

where the star denotes the adjoint operator. Note that the solution of (21) belongs to the $Range(T_{disc}^*)$, $\dim Range(T_{disc}^*) = rank(T_{disc}^*) = N$, and the finding an element x_α^{disc} reduces to solving a system of N linear algebraic equations. The problem is now to choose the regularization parameter α in dependence of δ in order to obtain the best possible order of accuracy with respect to δ as $\delta \to 0$.

The usual discussion of the order of accuracy of solution techniques for ill-posed problems (19) is done under the assumption that the minimum norm solution $x^\dagger$ lies in the range of $(T^*T)^\nu$, $\nu > 0$. From [7] it follows that under these assumptions for any solution technique connected with Tikhonov regularization (20), (21) the best possible order of accuracy in the power scale is $\delta^{2\nu/(2\nu+1)}$, $\nu \in (0,1]$.

One of the most widely used way to choose the regularization parameter α is by the residue principle. The general form of such principle was proposed in [8]. Within the framework of this general residue principle we choose $\alpha = \alpha(\delta)$ such that

$$\|Tx_\alpha - f_\delta\| \leq c_{p,q}\frac{\delta^p}{\alpha^q}, \tag{22}$$

where $p, q, c_{p,q}$ are some positive constants and x_α is the solution of non-discretized regularized equation

$$\alpha x + T^*Tx = T^*f_\delta.$$

A special case of general residue principle (22) is well-known Morozov's principle for $p = 1,\ q = 0$. The discretized version of this special case was studied in [6], [4]. But Morozov's principle leads to an optimal convergence rate $\delta^{2\nu/(2\nu+1)}$ of Tikhonov regularization only for $\nu \in (0, 1/2]$. For instance, within the framework of Morozov's principle we have not the possibility to obtain the best possible order of accuracy of Tikhonov regularization $\delta^{2/3}$. On the other hand, from [1], [3] it follows that for $p = \frac{2}{3}(q + 1)$ a non-discretized version of the general residue principle (22) leads to the order of accuracy $\delta^{2/3}$. The aim of this paper is to study the residue principle (22) for $p \leq q + 1$ in the form tailored to the discretized version of Tikhonov regularization (20), (21). Namely, we will choose $\alpha = \alpha(\delta)$ such that

$$c_1\frac{\delta^p}{\alpha^q} \leq \|T_{disc}x_\alpha^{disc} - f_\delta\| \leq c_2\frac{\delta^p}{\alpha^q}, \tag{23}$$

where, and below, c_1, c_2, etc. are positive generic constants which may take different values at different contexts.

A.2 Convergence Analysis

Investigating a posteriori strategies of the sort (23) always starts by proving that choosing α according to (23) is equivalent to a residue principle with exact data and exact operator.

Lemma A.1 *If α is chosen to satisfy (23) for $p \leq q + 1$, $\|T - T_{disc}\|_{H\to H} \leq \delta$ then there exist c_3, c_4 such that*

$$c_3\frac{\delta^p}{\alpha^q} \leq \|T\hat{x}_\alpha - f\| \leq c_4\frac{\delta^p}{\alpha^q},$$

where $\hat{x}_\alpha = (\alpha I + T^*T)^{-1}T^*f$.

Proof. From Lemma 7 [2] it follows that

$$
\begin{aligned}
&\|T_{disc}x_\alpha^{disc} - f_\delta\| - \tfrac{9}{4}\|T - T_{disc}\|_{H\to H}\|x^\dagger\| - 2\|f - f_\delta\| \le \\
&\le \|T\hat{x}_\alpha - f\| \le \|T_{disc}x_\alpha^{disc} - f_\delta\| + \tfrac{9}{4}\|T - T_{disc}\|_{H\to H}\|x^\dagger\| + \qquad (24)\\
&+2\|f - f_\delta\|.
\end{aligned}
$$

Moreover, from Lemma 1 [8] we know that if α is chosen according to (23) then

$$\alpha \le c_6 \delta^{\frac{p}{q+1}} \tag{25}$$

and therefore for $p \le q+1$

$$\frac{\delta^p}{\alpha^q} = \delta \frac{\delta^{p-1}}{\alpha^q} \ge \delta c_6^{-q} \delta^{p-1-\frac{qp}{q+1}} \ge c_7 \delta. \tag{26}$$

Combining (23)–(26) we obtain the assertion of the lemma.

Now we are in a position to obtain a discretized version of the main results connected with general residue principle (22) for Tikhonov regularization (see [8], [1], [3]).

Theorem A.1 *If $x^\dagger \in Range((T^*T)^\nu)$, $\nu \in (0,1]$ and α is chosen according to (23) for $p \le q+1$, $\|T - T_{disc}\|_{H\to H} \le \delta$ then*

$$\|x^\dagger - x_\alpha^{disc}\| \le c\delta^s,$$

where $s = \min\left\{\frac{p\nu}{q+1}, 1 - \frac{p}{2q} + \frac{p}{4q(q+1)}\right\}$.

Proof. Using Lemma 6 [2] we have

$$
\begin{aligned}
\|x^\dagger - x_\alpha^{disc}\| &\le \frac{\delta}{2\sqrt{\alpha}} + \frac{\|T - T_{disc}\|_{H\to H}\|x^\dagger\|}{\sqrt{\alpha}} + c_{12}\alpha^\nu \le \\
&\le \frac{c_{13}\delta}{\sqrt{\alpha}} + c_{12}\alpha^\nu. \qquad (27)
\end{aligned}
$$

From Lemma 2 [8] it follows that if α is such that

$$\frac{c_{14}\delta^p}{\alpha^q} \le \|Tx_\alpha - f_\delta\| \le \frac{c_{15}\delta^p}{\alpha^q} \tag{28}$$

then

$$\frac{\delta^p}{\sqrt{\alpha}} \le c_{16}\delta^{1-\frac{p}{2q}+\frac{p}{4q(q+1)}}. \tag{29}$$

Let us verify the condition (28) for α chosen to satisfy (23). It is easy to see that

$$\|T\hat{x}_\alpha - f\| - \|T(\hat{x}_\alpha - x_\alpha)\| - \|f - f_\delta\| \leq \|Tx_\alpha - f_\delta\| \leq$$
$$\leq \|T\hat{x}_\alpha - f\| + \|T(\hat{x}_\alpha - x_\alpha)\| + \|f - f_\delta\|. \tag{30}$$

Moreover,

$$\|T(\hat{x}_\alpha - x_\alpha)\| = \|T(\alpha I + T^*T)^{-1}T^*(f - f_\delta)\| =$$
$$= \|(\alpha I + TT^*)^{-1}TT^*(f - f_\delta)\| \leq \delta\|(\alpha I + TT^*)^{-1}TT^*\|_{H\to H} \leq \delta.$$

Keeping in mind this estimate, (26), (30) and Lemma A.1 we can see that α chosen according to (23) satisfies the condition (28). Now the assertion of the theorem follows from (27), (25) and (29).

Corollary A.1 *If α is chosen according to (23) for $\|T - T_{disc}\|_{H\to H} \leq \delta$ and*

$$p \leq \min\left\{\frac{2(q+1)}{2\nu+1}, \frac{2(q+1)}{1+\frac{1-\nu}{2q}}\right\} \tag{31}$$

then

$$\|x^\dagger - x_\alpha^{disc}\| \leq c\delta^{s_1} \tag{32}$$

where

$$s_1 = \min\left\{\frac{p\nu}{q+1}, 1 - \frac{p}{2q} + \frac{(\nu+1)p}{4q(q+1)}\right\}.$$

Proof. From Lemma 3 [3] it follows that if p and α satisfy the conditions (31) and (28) respectively then

$$\frac{\delta}{\sqrt{\alpha}} \leq c\delta^{1-\frac{p}{2q}+\frac{(\nu+1)p}{4q(q+1)}}.$$

The assertion of the corollary follows from this estimate and (27), (25) .

Remark. For $p = \frac{2}{3}(q+1)$ and $\nu = 1$ the estimate (32) leads to an optimal convergence rate $\delta^{2/3}$ of Tikhonov regularization. But if $x^\dagger \in Range((T^*T)^\nu)$, $\nu \in (0,1)$, then for $p = \frac{2}{3}(q+1)$ the residue principles (22), (23) lead to the convergence rate with order $\delta^{\frac{2}{3}\nu}$. For $\nu \in (0, 3/4)$ this convergence rate is worse than convergence rate of Tikhonov regularization with Morozov's residue principle ($p = 1$, $q = 0$). Moreover, for $\nu \in (0, 1/2]$ Morozov's principle leads to optimal convergence rate $\delta^{2\nu/(2\nu+1)}$ automaticly. There is a very interesting open problem here. Namely, is it possible to construct the residue principle for Tikhonov regularization leading automaticly to optimal convergence rate $\delta^{2\nu/(2\nu+1)}$ for all $\nu \in (0, 1]$?

A.3 Complexity of the discretization schemes.

Let $e_1, e_2, \dots, e_n, \dots$ be some orthonormal basis of Hilbert space H, and let P_n be the orthogonal projector on $span\{e_1, e_2, \dots, e_n\}$, that is

$$P_n f = \sum_{i=1}^{n} (f, e_i) e_i.$$

We denote by $H^r, \ \ r = 1, 2, \dots$, the linear subspace of H which is equipped with the norm

$$\|f\|_{H^r} := \|f\| + \|L^r f\|,$$

where L is some linear operator acting from H^r to H, and for $n = 1, 2, \dots$

$$\|I - P_n\|_{H^r \to H} = c_r n^{-r}.$$

It will be assumed now that the operators T have some "smoothness". Namely,

$$T \in \mathcal{H}_\gamma^r := \{T : \|T\|_{H \to H^r} \leq \gamma_1, \ \ \|T^*\|_{H \to H^r} \leq \gamma_2, \ \ \|(L^r T)^*\|_{H \to H^r} \leq \gamma_3\},$$

$$\gamma = (\gamma_1, \gamma_2, \gamma_3).$$

It is easy to see that the space H^r and the class $\mathcal{H}_\gamma^r$ are a generalization of the space of smooth functions W_2^1 and of the class $\mathcal{K}_\gamma^1$ of integral operators with kernels having square-summable mixed partial derivatives considered in Section 2.

The standard approach to the discretization of the problem (19) lies in the application of the Galerkin method. This means that

$$T_{disc} = P_n T P_n. \tag{33}$$

With an arguments like that in the proof of Theorem 5.5 [5] we get the estimate

$$\sup_{T \in \mathcal{H}_\gamma^r} \|T - P_n T P_n\|_{H \to H} \asymp \|I - P_n\|_{H^r \to H} \asymp n^{-r}.$$

Then from Theorem A.1 it follows that within the framework of the above mentioned standard approach (33 and with residue principle (23 we can guarantee the same order of accuracy of Tikhonov regularization as for non-discretized version (22 in the case when $n \asymp \delta^{-1/r}$.

Denote by $Card(IP)$ the number of inner products of the form (e_i, Te_j) required to construct T_{disc}. Then for standard approach (33)

$$Card(IP) = n^2 \asymp \delta^{-2/r}. \tag{34}$$

Now we combine the general residue principle (23) with "hyperbolic cross" approximation (4). This means that now $n = 2^m$ and

$$T_{disc} = \sum_{k=1}^{m} (P_{2^k} - P_{2^{k-1}}) T P_{2^{m-k}} + P_1 T P_{2^m}. \tag{35}$$

We remind that for this discretization scheme

$$Card(IP) \asymp m2^m.$$

Lemma A.2 *If for discretization scheme (35)* $Card(IP) \asymp \delta^{-1/r} \log^{1+1/r} \frac{1}{\delta}$ *and* $T \in \mathcal{H}_\gamma^r$ *then*

$$\|T - T_{disc}\|_{H \to H} \le \delta.$$

Proof. With an argument like that in the proof of Lemma 2.1 for $T \in \mathcal{H}_\gamma^r$ we have (for some constant c)

$$\|T - T_{disc}\|_{H \to H} \le cm2^{-mr} \asymp (Card(IP))^{-r} \, m^{r+1} \le \delta$$

as claimed.

When Lemma A.2 is compared with (34) it is apparent that within the framework of general residue principle (23) the discretization scheme (35) is more economical for $T \in \mathcal{H}_\gamma^r$ than standard Galerkin scheme (33).

References

[1] Guacaneme, J.E. (1988), An optimal parameter choice for regularized ill-posed problems, *Integral Equations and Operator Theory* **11**, 610–613.

[2] Maass, P., Rieder, A. (1997), Wavelet-accelerated Tikhonov-regularization with applications, in *Inverse Problems in Medical Imaging*, H.W. Engl, A.K. Louis and W. Rundell, eds., Springer.

[3] Nair, M.T. (1992), A generalization of Arcangeli's method for ill-posed problems leading to optimal rates, *Integral Equations and Operator Theory* **15**, 1042–1046.

[4] Pereverzev, S.V., Solodky, S.G. (1996), An efficient discretization for solving ill-posed problems, *Lectures in Applied Mathematics* **32**, 643–649.

[5] Pereverzev, S.V. (1996), *Optimization of Methods for Approximate Solution of Operator Equations*, Nova, New York.

[6] Plato, R., Vainikko, G. (1990), On the Regularization of Projection Methods for solving Ill-posed Problems, *Numer.Math.* **57**, 63–70.

[7] Schock, E. (1984), On the asymptotic order of accuracy of Tikhonov regularization, *J. Optim. Th. Appl.* **44**, 95–104.

[8] Schock, E. (1984), Parameter choice by discrepancy principle for the approximate solution of ill-posed problems, *Integral Equations and Operator Theory* **7**, 895–898.