An Improved Version of Marti's Method
for Solving Ill-Posed Linear Integral Equations

By Heinz W. Engl and Andreas Neubauer*

Abstract. We propose an algorithm for solving linear integral equations of the first kind that can be viewed as a variant of Marti's method; as opposed to that method, our algorithm leads to optimal convergence rates (also with noisy data).

1. Introduction. Throughout this paper, let X and Y be real Hilbert spaces, $T: X \to Y$ a bounded linear operator with nonclosed range. Then the problem of determining the "best-approximate solution" of

$$(1.1) \qquad\qquad Tx = y$$

is ill-posed: The best-approximate solution exists only for $y \in D(T^\dagger) = R(T) \dotplus R(T)^\perp$ (which we assume from now on) and depends discontinuously on the right-hand side. Here $T^\dagger$ is the Moore-Penrose inverse of T (see [12]); the best-approximate solution is defined as the element of minimal norm that minimizes the residual $\|Tx - y\|$ and can be written as $T^\dagger y$. A prominent example for the ill-posed case of (1.1) is a Fredholm integral equation of the first kind. Ill-posed problems have to be solved by regularization methods, e.g., Tikhonov regularization. See [7] and [14] for more background.

An algorithm that has been used successfully in recent years is "Marti's method" (see [9], [10], [11]). In this method, a sequence of finite-dimensional subspaces $V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots$ of X with $\overline{\bigcup_{n \in \mathbf{N}} V_n} = X$ is used to compute approximate solutions of (1.1) as follows:

Let, for all $n \in \mathbf{N}$,

$$(1.2) \qquad\qquad a_n := \inf\{\|Tx - y\|/x \in V_n\},$$

P_n be the orthogonal projector of X onto V_n, and $b_n > 0$ be chosen such that

$$(1.3) \qquad \lim_{n \to \infty} \frac{\|P_n T^\dagger y - T^\dagger y\|}{b_n} = 0, \qquad \lim_{n \to \infty} b_n = 0$$

holds. Then x_n is defined by

$$(1.4) \qquad\qquad x_n \in V_n,$$

$$(1.5) \qquad\qquad \|Tx_n - y\|^2 \leqslant a_n^2 + b_n^2,$$

(1.6) $\|x_n\| = \inf\{\|x\|/x \in V_n \text{ and } x \text{ fulfills } (1.5)\}.$

This is Marti's algorithm as given in [11]; originally ([9], [10]), the right-hand side of (1.5) was $(a_n + b_n)^2$ and the algorithm was formulated only for $y \in R(T)$. The remarks of this section apply to both versions of the algorithm.

Marti proved that $\{x_n\}$ converges to $T^\dagger y$ as $n \to \infty$ and that

(1.7) $\|x_n - T^\dagger y\| = O\left(\sqrt{b_n}\right)$

if

(1.8) $T^\dagger y \in R(T^*)$

holds, which can be interpreted as an a priori smoothness assumption about the (unknown) exact solution. By T^* we denote the adjoint of T.

Incidentally, in [8] it is claimed that Marti's results are wrong as stated; this statement is formally correct if applied to [9], since there the condition (1.3) is missing. However, this condition appears in [10], and the results in that paper are correct. It has to be pointed out that the authors of [8] quote [10], so that their claim that Marti's result is in error and that they corrected it is not justified, since Marti himself had corrected his error before them in [10]. Moreover, it is easy to see where the error in [9] is: There, (5) is wrong, since (in the notation of that paper) $P_m f_0$ need not be in V_n. From there, one can immediately deduce the condition that is added in the corrected version of Marti's result ([8, Theorem 2.5]).

In [7, Section 4.3], C. W. Groetsch has given the following alternate formulation of Marti's algorithm:

x_n is determined by

(1.9) $\alpha_n x_n + T_n^* T_n x_n = T_n^* y,$

(1.10) $\|T x_n - y\|^2 = a_n^2 + b_n^2,$

where

(1.11) $T_n := T_{|V_n}.$

Note that (1.9) is just Tikhonov regularization, where the regularization parameter α_n is determined from (1.10), which can be interpreted as a "discrepancy principle" (cf. [13], [7, Section 3.3]). In view of [5] it cannot be expected that the convergence rate in (1.7) can be improved, even under stronger smoothness assumptions. However, our results in [2] and [3] can be used to modify (1.9), (1.10) (and thus Marti's method) in such a way that the convergence rate in (1.7) is improved to the best rate one can reasonably expect.

2. Main Results. As in Marti's method, let $V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots$ be a sequence of finite-dimensional subspaces of X with $\overline{\bigcup_{n \in \mathbf{N}} V_n} = X$. For each $n \in \mathbf{N}$, let

(2.1) $\gamma_n := \|(I - P_n)T^*\|,$

where P_n is the orthogonal projector of X onto V_n, and let $y_n \in Y$ with

(2.2) $\|y - y_n\| \leqslant \delta_n,$

where the sequence $\{\delta_n\}$ is assumed to be known. Let $\{b_n\}$ be a sequence in $\mathbf{R}^+$; we assume

(2.3) $\lim_{n \to \infty} b_n = 0, \qquad \lim_{n \to \infty} \delta_n = 0.$

Further properties of $\{b_n\}$ will be fixed below. For each $\alpha > 0$ and $n \in \mathbf{N}$ we denote by $x_{n,\alpha}^{\delta_n}$ the unique solution of

$$(2.4) \qquad\qquad \alpha x + T_n^* T_n x = T_n^* y_n$$

in V_n, where T_n is as in (1.11); with this definition, let

$$(2.5) \qquad\qquad \rho_n(\alpha) := \left\| T_n^* T_n x_{n,\alpha}^{\delta_n} - T_n^* y_n \right\|^2.$$

Note that ρ_n depends also on y_n and hence on δ_n, which is the norm of the data error. In our variant of Marti's method, the data error will be included from the beginning, since despite its different origin, it can be treated the same way as the approximation error symbolized by b_n.

Let D_1, D_2 be positive constants, and $p, q > 0$ be parameters that are fixed later. In our algorithm, $x_n \in V_n$ is defined as the unique solution of (2.4), where α is the solution of

$$(2.6) \qquad\qquad \rho_n(\alpha) = (D_1 b_n + D_2 \delta_n)^p \alpha^{-q}.$$

This can be seen as a variant of Marti's method as formulated by (1.9), (1.10) in the following sense:

The equations (1.9) and (2.4) are identical except that in (2.4) we only use the approximation y_n for y; as can be seen from [7, (4.3.5)], (1.10) can also be written as

$$(2.7) \qquad\qquad \left\| T_n x_n - Q_n y \right\|^2 = b_n^2,$$

where Q_n is the orthogonal projector of Y onto $T(V_n)$. If we replace the residual in (2.7) by the residual of the corresponding finite-dimensional normal equation, we obtain (2.6) with $\delta_n = 0$, $D_1 = 1$, $p = 2$, and $q = 0$.

From now on we denote by $x_n^{\delta_n}$ and x_n always the unique solution in V_n of (2.4) and (2.4) with y_n replaced by y, respectively, where α is determined by (2.6); this α will also be denoted by α_n. Of course, these quantities depend on p and q. We will determine p and q in such a way that the convergence rate $\| x_n^{\delta_n} - T^\dagger y \|$ is best possible.

To exclude trivial cases, we will always assume that

$$(2.8) \qquad\qquad T^* y \neq 0 \quad \text{and} \quad T_n^* y_n \neq 0,$$

where the latter assumption follows from the first when n is sufficiently large. The first result shows that our algorithm makes sense:

PROPOSITION 2.1. *For any $p, q > 0$, (2.6) is uniquely solvable.*

Proof. It follows from [4, Lemma 3.1] (cf. also [7, Theorem 3.3.1]), applied with T_n, V_n and y_n instead of T, X and b^δ, respectively, that ρ_n is continuous, strictly increasing and that

$$\lim_{\alpha \to 0} \rho_n(\alpha) = 0 \quad \text{and} \quad \lim_{\alpha \to +\infty} \rho_n(\alpha) = \left\| T_n^* y_n \right\|^2 > 0$$

because of (2.8). Thus, the assertion follows from the Intermediate Value Theorem. $\square$

Since $x_n^{\delta_n}$ and x_n minimize the Tikhonov functionals

$$x \to \| Tx - y_n \|^2 + \alpha_n \| x \|^2 \quad \text{and} \quad x \to \| Tx - y \|^2 + \alpha_n \| x \|^2,$$

respectively, over the subspace V_n, we obtain from [7, Lemma 4.2.7] that

$$(2.9) \qquad\qquad \left\| x_n^{\delta_n} - x_n \right\| \leqslant \delta_n \cdot \alpha_n^{-1/2}.$$

For the same reason, we can use [7, Lemma 4.2.3] to obtain

$$(2.10) \qquad\qquad \left\| x_n - x^{\alpha_n} \right\| \leqslant \left(1 + \frac{\gamma_n^2}{\alpha_n} \right)^{1/2} \cdot \left\| (I - P_n) x^{\alpha_n} \right\|,$$

where γ_n is defined in (2.1), and where

$$(2.11) \qquad\qquad x^\alpha := \left(T^*T + \alpha I \right)^{-1} T^*y$$

is the approximate solution obtained from infinite-dimensional Tikhonov regularization; of course, x^α is not used in our algorithm, which is strictly finite-dimensional.

LEMMA 2.2. $\lim_{n \to \infty} \alpha_n = 0$.

Proof. One shows, as in the proof of [2, Lemma 2.2], that if $\{\alpha_n\}$ would have a subsequence (again denoted by $\{\alpha_n\}$) with

$$(2.12) \qquad\qquad \lim_{n \to \infty} \alpha_n = +\infty,$$

then this would imply

$$(2.13) \qquad\qquad \lim_{n \to \infty} x_n^{\delta_n} = 0,$$

and hence

$$(2.14) \qquad\qquad \lim_{n \to \infty} \rho_n(\alpha_n) = \left\| T^*y \right\|^2$$

because of (2.2), (2.3) and the facts that $T_n^* = P_n T^*$ and $P_n \to I$ pointwise. Hence

$$0 = \lim_{n \to \infty} \left(D_1 b_n + D_2 \delta_n \right)^p = \lim_{n \to \infty} \left(\alpha_n^q \cdot \rho_n(\alpha_n) \right) = +\infty$$

because of (2.12), (2.14), and (2.8), which is a contradiction. Thus, (2.12) cannot hold for any subsequence, which implies the boundedness of $\{\alpha_n\}$.

Now assume that

$$(2.15) \qquad\qquad \lim_{n \to \infty} \alpha_n = \alpha > 0$$

holds for some subsequence (again denoted by $\{\alpha_n\}$). Then

$$\lim_{n \to \infty} \left\| \left(T_n^*T_n + \alpha_n I \right)^{-1} T_n^* - \left(T_n^*T_n + \alpha I \right)^{-1} T_n^* \right\| = 0,$$

so that

$$(2.16) \qquad\qquad \lim_{n \to \infty} \left\| x_n^{\delta_n} - \left(T_n^*T_n + \alpha I \right)^{-1} T_n^*y \right\| = 0.$$

Because of [7, Lemma 4.2.3],

$$(2.17) \qquad \left\| \left(T_n^*T_n + \alpha I \right)^{-1} T_n^*y - x^\alpha \right\|^2 \leqslant \left(1 + \frac{\gamma_n^2}{\alpha} \right) \cdot \left\| (I - P_n) x^\alpha \right\|^2,$$

where x^α is as in (2.11) with α as in (2.15). Since the right-hand side of (2.17) tends to 0 as $n \to \infty$, (2.16) and (2.17) imply

$$(2.18) \qquad\qquad \lim_{n \to \infty} \left\| x_n^{\delta_n} - x^\alpha \right\| = 0.$$

Since, as easy calculations show,

$$(2.19) \qquad \rho_n(\alpha) = \alpha^2 \cdot \left\| x_{n,\alpha}^{\delta_n} \right\|^2$$

holds for all $\alpha > 0$, (2.18) implies together with the definition of α_n that

$$0 = \lim_{n \to \infty} (D_1 b_n + D_2 \delta_n)^p = \lim_{n \to \infty} \left(\alpha_n^{q+2} \left\| x_n^{\delta_n} \right\|^2 \right) = \alpha^{q+2} \left\| x^\alpha \right\|^2,$$

which implies $x^\alpha = 0$ in contradiction to (2.8). Hence (2.15) cannot hold for any subsequence, which proves the lemma. $\square$

LEMMA 2.3. *If* $0 < p \leqslant 2q$, *then* $\lim_{n \to \infty} (\delta_n^2 \cdot \alpha_n^{-1}) = 0$.

Proof. Since $x_n^{\delta_n}$ minimizes the functional $x \to \| Tx - Q_n y_n \|^2 + \alpha_n \| x \|^2$ over V_n, where Q_n is the orthogonal projector of Y onto $T(V_n)$ (and hence $T_n^* Q_n = T_n^*$), we have

$$0 \leqslant \rho_n(\alpha_n) \leqslant \left\| T_n^* \right\|^2 \cdot \left[\left\| T_n x_n^{\delta_n} - Q_n y_n \right\|^2 + \alpha_n \left\| x_n^{\delta_n} \right\|^2 \right]$$

$$\leqslant \| T \|^2 \cdot \left[\left\| T P_n T^\dagger y - Q_n y_n \right\|^2 + \alpha_n \left\| T^\dagger y \right\|^2 \right].$$

Since $P_n T^\dagger y \to T^\dagger y$ and Q_n converges pointwise to the projector of Y onto $\overline{R(T)}$, i.e., to the extension of $TT^\dagger$ onto Y, $\| T P_n T^\dagger y - Q_n y_n \| \to 0$. This implies, together with the last inequality and with Lemma 2.2, that $\lim_{n \to \infty} \rho_n(\alpha_n) = 0$. Hence,

$$0 \leqslant \delta_n^p \cdot \alpha_n^{-p/2} \leqslant D_2^{-p} \cdot (D_1 b_n + D_2 \delta_n)^p \cdot \alpha_n^{-p/2} = D_2^{-p} \cdot \alpha_n^{q-p/2} \cdot \rho_n(\alpha_n);$$

the last expression tends to 0 as $n \to \infty$ because of the first part of this proof, Lemma 2.2 and the assumptions on p and q. Hence $\lim_{n \to \infty} \delta_n^p \cdot \alpha_n^{-p/2} = 0$, which implies the assertion. $\square$

The next result, where a condition on $\{ b_n \}$—the analogue of (1.3) in Marti's method—is introduced, shows that for a wide range of p and q, our method converges.

THEOREM 2.4. *Let* $0 < p \leqslant 2q$ *and assume that*

$$(2.20) \qquad \gamma_n = O \left(b_n^{p/2q} \right),$$

where γ_n *is defined in* (2.1). *Then* $\lim_{n \to \infty} x_n^{\delta_n} = T^\dagger y$.

Proof. We first show that

$$(2.21) \qquad \left\{ \gamma_n^2 / \alpha_n \right\} \text{ is bounded.}$$

To see this, note that (cf. the proof of Proposition 2.1)

$$(D_1 b_n + D_2 \delta_n)^p \alpha_n^{-q} = \rho_n(\alpha_n) \leqslant \left\| T_n^* y_n \right\|^2 \leqslant \| T \|^2 \cdot \left(\| y \| + \delta_n \right)^2,$$

hence $\alpha_n \geqslant C (D_1 b_n + D_2 \delta_n)^{p/q}$ for all $n \in \mathbf{N}$ with a suitable $C > 0$. Together with (2.20), this implies (2.21). Now,

$$\left\| x_n^{\delta_n} - T^\dagger y \right\| \leqslant \left\| x_n^{\delta_n} - x_n \right\| + \left\| x_n - x^{\alpha_n} \right\| + \left\| x^{\alpha_n} - T^\dagger y \right\|,$$

where x^{α_n} is defined by (2.11) for $\alpha = \alpha_n$. Because of (2.9) and Lemma 2.3, the first term of this estimate tends to 0. Because of Lemma 2.2 and standard results about infinite-dimensional Tikhonov regularization (see, e.g., [7, Sections 2.1 and 3.1]), the third term tends to 0. Because of (2.10), the second term can be estimated as follows:

$$\left\| x_n - x^{\alpha_n} \right\| \leqslant \left(1 + \frac{\gamma_n^2}{\alpha_n} \right)^{1/2} \cdot \left[\left\| (I - P_n) \right\| \cdot \left\| x^{\alpha_n} - T^\dagger y \right\| + \left\| (I - P_n) T^\dagger y \right\| \right],$$

which tends to 0 because of (2.21), the convergence of x^{α_n} to $T^\dagger y$ and the fact that $\{P_n\} \to I$ pointwise. This implies that $\{x_n^{\delta_n}\} \to T^\dagger y$. $\square$

Now we proceed to give estimates for the rate of convergence for various choices of p and q. For this, we need the following estimate:

LEMMA 2.5. *Let the assumptions of Theorem 2.4 be fulfilled. Then there exist constants $C_1, C_2 > 0$ such that for all integers n, $C_1 \leqslant (D_1 b_n + D_2 \delta_n)^p \alpha_n^{-q-2} \leqslant C_2$ holds.*

Proof. Because of the definition of α_n and (2.19), $(D_1 b_n + D_2 \delta_n)^p \alpha_n^{-q-2} = \|x_n^{\delta_n}\|^2$, which converges to $\|T^\dagger y\|^2$. Together with (2.8), this implies the assertion. $\square$

As usual with ill-posed problems, convergence rates can only be obtained under a priori assumptions about the exact solution:

THEOREM 2.6. (a) *Assume that (1.8) holds, and let $p - 2 = q \geqslant 2$ and*

$$(2.22) \qquad\qquad \gamma_n = O\big(b_n^{1/2 + 1/q}\big).$$

Then $\|x_n^{\delta_n} - T^\dagger y\| = O((b_n + \delta_n)^{1/2})$.
(b) *Assume that*

$$(2.23) \qquad\qquad T^\dagger y \in R(T^*T)$$

and let $\frac{3}{2} p - 2 = q \geqslant 1$ and

$$(2.24) \qquad\qquad \gamma_n = \begin{cases} O\big(b_n^{1/3 + 2/3q}\big) & \text{if } q < 2, \\ O\big(b_n^{2/3}\big) & \text{if } q \geqslant 2. \end{cases}$$

Then $\|x_n^{\delta_n} - T^\dagger y\| = O((b_n + \delta_n)^{2/3})$.

Proof. Because of (2.9), $\|x_n^{\delta_n} - x_n\| \leqslant D_2^{-1} \cdot (D_1 b_n + D_2 \delta_n) \cdot \alpha_n^{-1/2}$. Since (2.22) as well as (2.24) imply (2.20), we can apply Lemma 2.5 and obtain with C_1 and C_2 as defined there:

$$\left\| x_n^{\delta_n} - x_n \right\| \leqslant C_3 \cdot \big(D_1 b_n + D_2 \delta_n \big)^{1 - p/2(q+2)}, \quad \text{where } C_3 := D_2^{-1} \cdot C_2^{1/2(q+2)}.$$

Hence

$$(2.25) \qquad \left\| x_n^{\delta_n} - T^\dagger y \right\| \leqslant \left\| x_n - T^\dagger y \right\| + C_3 \cdot \big(D_1 b_n + D_2 \delta_n \big)^{1 - p/2(q+2)}.$$

Now, assume $p - 2 = q \geqslant 2$ and (2.22) (and hence $p \leqslant 2q$ and (2.20)). Because of Lemma 2.5, $(D_1 b_n + D_2 \delta_n)^{1/2} = O(\alpha_n^{(q+2)/2p}) = O(\alpha_n^{1/2})$, and hence

$$\gamma_n = O\big(b_n^{1/2 + 1/q}\big) = O\big(b_n^{1/2}\big) = O\big((D_1 b_n + D_2 \delta_n)^{1/2}\big) = O\big(\alpha_n^{1/2}\big).$$

Thus, we can apply [7, Theorem 4.2.5] and obtain

$$(2.26) \qquad\qquad \|x_n - T^\dagger y\| = O\big(\alpha_n^{1/2}\big).$$

Because of Lemma 2.5,

$$\alpha_n^{1/2} = O\big((D_1 b_n + D_2 \delta_n)^{p/2(q+2)}\big) = O\big((D_1 b_n + D_2 \delta_n)^{1/2}\big),$$

so that (2.25) and (2.26) imply

$$\left\| x_n^{\delta_n} - T^\dagger y \right\| = O\left((D_1 b_n + D_2 \delta_n)^{1/2} \right) = O\left((b_n + \delta_n)^{1/2} \right).$$

This proves (a).

Now, let the assumptions of part (b) hold. Since also $p \leqslant 2q$ and (2.20) hold, we can apply Lemma 2.3 and obtain

$$(2.27) \qquad\qquad (D_1 b_n + D_2 \delta_n)^{2/3} = O(\alpha_n)$$

and

$$(2.28) \qquad\qquad \alpha_n = O\left((D_1 b_n + D_2 \delta_n)^{2/3} \right).$$

For $q \geqslant 2$, (2.24) and (2.27) imply $\gamma_n = O(\alpha_n)$. For $q < 2$, we obtain the same result since $b_n^{1/3 + 2/3q} = O(b_n^{2/3})$. Thus we can apply [7, Theorem 4.2.6] to obtain

$$(2.29) \qquad\qquad \left\| x_n - T^\dagger y \right\| = O(\alpha_n).$$

The result of part (b) now follows from (2.25), (2.27), (2.29), and (2.28). $\square$

Remark 2.7. Theorem 2.6 shows that our method improves Marti's method in two ways: First, data errors are included. It can be easily seen from the proofs, that in (2.20), (2.22) and (2.24), b_n could be replaced by $(b_n + \delta_n)$, which yields (formally) weaker conditions. Second, our convergence rates are better than those for Marti's method. For simplicity, we discuss this aspect for error-free data, i.e., $\delta_n = 0$. We first look at part (a) of Theorem 2.6. If we choose $q = 2$, we obtain the rate $O(\sqrt{b_n})$, which becomes $O(\sqrt{\gamma_n})$, if $b_n \sim \gamma_n$, i.e., if the correct asymptotic behavior of γ_n is known. If we choose higher values of q, we can get arbitrarily close to the optimal rate $O(\gamma_n)$; e.g., for $q = 10$ we can take $b_n \sim \gamma_n^{5/3}$ and thus obtain the rate $O(\gamma_n^{5/6})$. Of course, if we underestimate γ_n, i.e., if we replace O by o in (2.22), we get worse rates in terms of γ_n, which is not surprising. Note that γ_n represents the best possible convergence rate of elements in V_n toward solutions of (1.1) that fulfill (1.8). While in Marti's method, one can obtain at most the square root of the optimal rate (cf. (1.7)), we can come arbitrarily close to the optimal rate, even under the weak smoothness assumption (1.8). In Marti's method, a better convergence rate cannot be obtained under the stronger assumption (2.23) (cf. the concluding remarks of Section 1); Theorem 2.6(b) shows that we obtain the optimal convergence rate $O(\gamma_n)$ for $q \geqslant 2$ if we take $b_n \sim \gamma_n^{3/2}$. In the presence of data errors, the convergence rates are also best possible in terms of δ_n (cf., e.g., [1]).

If $\delta_n = 0$, one does not need the assumption $p \leqslant 2q$, which was needed in Lemma 2.3. In this case, one can show that if (1.8) and (2.20) hold, then $\| x_n - T^\dagger y \| = O(b_n^{p/2(q+2)})$; i.e., if one takes $b_n \sim \gamma_n^{2q/p}$, then one obtains the rate $O(\gamma_n^{q/(q+2)})$ for any choice of $p, q > 0$. In this case, (2.6) reads $\rho_n(\alpha) = D \cdot (\gamma_n^2/\alpha)^q$ with a suitable constant D.

Similarly, if $\delta_n = 0$ and (2.24) holds, and if either $0 < q < 2$ and $\gamma_n = O(b_n^{p/2q})$ or $q \geqslant 2$ and $\gamma_n = O(b_n^{p/(q+2)})$, then $\| x_n - T^\dagger y \| = O(b_n^{p/(q+2)})$ holds; i.e., if one takes $b_n \sim \gamma_n^{2q/p}$ for $q < 2$ or $b_n \sim \gamma_n^{(q+2)/p}$ for $q \geqslant 2$, one obtains the rates $O(\gamma_n^{2q/(q+2)})$ for $q < 2$ and the optimal rate $O(\gamma_n)$ for $q \geqslant 2$. The equation (2.6) then reads $\rho_n(\alpha) = D \cdot (\gamma_n^2/\alpha)^q$ and $\rho_n(\alpha) = D \cdot \gamma_n^{q+2} \cdot \alpha^{-q}$, respectively (with suitable constants D).

The proof of these statements follows along the lines of the proof of Theorem 2.6.

For actual computations, the choice of D_1 and D_2 is of course important. Since we want to give convergence rates only, we do not enter into theoretical discussions on the choice of these constants (cf. [3] in a similar situation). In our computations, we chose $D_1 = D_2 = 10^{-6q/p}$, which turned out to be effective.

3. Numerical Aspects and Examples. For computing the solution $x_n^{\delta_n}$ of (2.4) and (2.6), one chooses a basis $\{v_1, \ldots, v_{d(n)}\}$ of V_n, computes the $d(n) \times d(n)$-matrices $B_n := (\langle Tv_i, Tv_j \rangle)$ and $M_n := (\langle v_i, v_j \rangle)$ and the vector $w_n := (\langle Tv_i, y_n \rangle)$. It is easy to see that if $(\alpha, \lambda) \in \mathbf{R}^{d(n)+1}$ solves

$$(3.1) \qquad\qquad (B_n + \alpha M_n)\lambda = w_n,$$

$$(3.2) \qquad\qquad \alpha^{q+2}\lambda^T M_n \lambda = (D_1 b_n + D_2 \delta_n)^p,$$

then $x_n^{\delta_n} = \sum_{i=1}^{d(n)} \lambda_i v_i$. Note that the system (3.1), (3.2) is very similar to the computational form of Marti's method (see, e.g., equations (3a), (3b) of [10]). Note that for obtaining (3.2), we used (2.19).

Let for any $\alpha > 0$, $\lambda(\alpha)$ be the unique solution of (3.1), and

$$(3.3) \qquad f(\alpha) := \alpha^{q+2}\lambda(\alpha)^T M_n \lambda(\alpha) - (D_1 b_n + D_2 \delta_n)^p.$$

One shows as in the proof of [3, Proposition 2.2]: f is differentiable, for all $\alpha > 0$,

$$(3.4) \qquad f'(\alpha) = (q+2)\alpha^{q+1}\lambda(\alpha)^T M_n \lambda(\alpha)$$
$$- 2\alpha^{q+2}\lambda(\alpha)^T M_n (B_n + \alpha M_n)^{-1} M_n \lambda(\alpha) > 0;$$

for all $q \geqslant 1$, Newton's method

$$(3.5) \qquad\qquad \alpha^{k+1} := \alpha^k - \frac{f(\alpha^k)}{f'(\alpha^k)}$$

converges to the unique zero α_n of f (defined in (3.3)). The convergence is global, for $k \geqslant 2$ the iterates decrease monotonically to α_n. Obviously, the vector $(\alpha_n, \lambda(\alpha_n))$ solves (3.1), (3.2). The solution of (3.1) that is needed in each iteration of (3.5) is done by Cholesky decompositions. For finding a suitable sequence b_n, one needs information about γ_n (as defined in (2.1)); estimates for γ_n for spline spaces can be found, e.g., in [6]. In an analogous way, numerous estimates for the approximating power of finite-element spaces could be used to estimate γ_n.

In our example, we choose V_n as a space of linear splines on a uniform grid of $(n + 1)$ points in $[0, 1]$. As basis functions, we take $v_1, \ldots, v_{n+1}$ with the property that $v_i((i - 1)/n) = 1$ and v_i vanishes at all other nodes. The elements of M_n are computed explicitly, the elements Tv_i are computed by Gaussian quadrature with two nodes on each subinterval $[(i - 1)/n, i/n]$. Finally, the scalar products needed for computing the elements of B_n are approximated by the trapezoidal rule. This is (nearly) identical to the procedure chosen in [9] and [10], so that it is fair to compare the results. All examples are Fredholm integral equations of the first kind on $[0, 1]$ with kernel k:

$$\int_0^1 k(t, s)x(s)\,ds = y(s).$$

Example 3.1. Here the kernel is always given by

$$k(t, s) := \begin{cases} 0, & t \leqslant s, \\ t - s, & t > s \end{cases}$$

(cf. the example in [10]). It follows from [6] that $\gamma_n = O(n^{-2})$.

(a) $y(t) = \frac{1}{24}(6t^2 - 4t^3 + t^4)$. The exact solution is $(T^\dagger y)(s) = \frac{1}{2}(s - 1)^2 \in R(T^*)$, since $T^\dagger y = T^*1$. We use the variant of Theorem 2.6 given in Remark 2.7 under the assumption (1.8) and choose $p = 1$, $q = 2$, $b_n := n^{-8}$. According to the theory, we should obtain the convergence rate $O(n^{-1})$. The results are as follows:

n	α_n	e_n	$e_n \cdot n \cdot 10^2$
4	$1.4 * 10^{-4}$	$3.4 * 10^{-3}$	1.4
8	$3.4 * 10^{-5}$	$1.1 * 10^{-3}$	0.90
16	$8.3 * 10^{-6}$	$4.3 * 10^{-4}$	0.68
32	$2.1 * 10^{-6}$	$1.7 * 10^{-4}$	0.55
64	$5.2 * 10^{-7}$	$7.1 * 10^{-5}$	0.45

Here, $e_n := \|x_n - T^\dagger y\|$. The last column shows that the convergence seems to be slightly faster than $O(n^{-1})$.

(b) $y(t) = \frac{1}{720}(t^6 - 20t^3 + 45t^2)$, $(T^\dagger y)(s) = \frac{1}{24}(s^4 - 4s + 3) \in R(T^*T)$, since $T^\dagger y = T^*T1$. Choosing the parameters as in (a), the results outlined in Remark 2.7 show that the convergence rate should be the optimal rate $O(n^{-2})$. This is in fact the case:

n	α_n	e_n	$e_n \cdot n^2 \cdot 10^2$
4	$2.5 * 10^{-4}$	$9.1 * 10^{-4}$	1.5
8	$6.3 * 10^{-5}$	$2.2 * 10^{-4}$	1.4
16	$1.6 * 10^{-5}$	$5.2 * 10^{-5}$	1.3
32	$3.9 * 10^{-6}$	$1.2 * 10^{-5}$	1.2
64	$9.7 * 10^{-7}$	$2.9 * 10^{-6}$	1.2

(c) y as in (a); y is randomly perturbed by $\delta_n = n^{-2}\|y\|$. We use Theorem 2.6(a) with $b_n = n^{-2}$, $p = 4$, $q = 2$. According to the theory, we should obtain the rate $\|x_n^{\delta_n} - T^\dagger y\| = O(n^{-1}) = O(\delta_n^{1/2})$.

n	α_n	$\tilde{e}_n$	$\tilde{e}_n \cdot n \cdot 10$
4	$1.4 * 10^{-4}$	$5.6 * 10^{-2}$	2.2
8	$3.4 * 10^{-5}$	$7.9 * 10^{-2}$	6.3
16	$8.8 * 10^{-6}$	$3.1 * 10^{-2}$	4.9
32	$2.2 * 10^{-6}$	$1.3 * 10^{-2}$	4.2
64	$5.5 * 10^{-7}$	$5.9 * 10^{-3}$	3.8

Here (and below), $\tilde{e}_n := \|x_n^{\delta_n} - T^\dagger y\|$.

(d) y as in (b); y is randomly perturbed by $\delta_n = n^{-3}\|y\|$. We use Theorem 2.6(b) with $b_n = n^{-3}$, $p = 8/3$, $q = 2$. According to the theory, we should obtain $\|x_n^{\delta_n} - T^\dagger y\| = O(n^{-2}) = O(\delta_n^{2/3})$.

n	α_n	$\tilde{e}_n$	$\tilde{e}_n \cdot n^2 \cdot 10^2$
4	$2.5 * 10^{-4}$	$3.6 * 10^{-3}$	5.8
8	$6.1 * 10^{-5}$	$2.1 * 10^{-3}$	13
16	$1.5 * 10^{-5}$	$3.6 * 10^{-4}$	9.3
32	$3.8 * 10^{-6}$	$8.4 * 10^{-5}$	8.6
64	$9.5 * 10^{-7}$	$2.1 * 10^{-5}$	8.6

Computations with larger values for δ_n give comparable results.

Example 3.2. Here the kernel is given by

$$k(t, s) := \begin{cases} t(1 - s), & t \leqslant s, \\ s(1 - t), & t > s \end{cases}$$

(cf. Example b) in [9]). Again, $\gamma_n = O(n^{-2})$.

In parts (a)–(d) of this example, the same choices of the relevant parameters are made and the same statements about convergence rates hold as in the corresponding parts of Example 3.1.

(a) $y(t) = \frac{1}{24}(t - 2t^3 + t^4)$, $(T^\dagger y)(s) = \frac{1}{2}(s - s^2)$.

n	α_n	e_n	$e_n \cdot n \cdot 10^2$
4	$2.1 * 10^{-4}$	$4.0 * 10^{-3}$	1.6
8	$5.2 * 10^{-5}$	$1.4 * 10^{-3}$	1.1
16	$1.3 * 10^{-5}$	$5.5 * 10^{-4}$	0.88
32	$3.3 * 10^{-6}$	$2.2 * 10^{-4}$	0.72
64	$8.1 * 10^{-7}$	$9.3 * 10^{-5}$	0.60

(b) $y(t) = \frac{1}{30}(3t - 5t^3 + 3t^5 + t^6)$, $(T^\dagger y)(s) = s - 2s^3 + s^4$.

n	α_n	e_n	$e_n \cdot n^2 \cdot 10^2$
4	$1.4 * 10^{-4}$	$6.2 * 10^{-3}$	10.0
8	$3.4 * 10^{-5}$	$1.5 * 10^{-3}$	9.6
16	$8.3 * 10^{-6}$	$3.7 * 10^{-4}$	9.5
32	$2.1 * 10^{-6}$	$9.3 * 10^{-5}$	9.5
64	$5.2 * 10^{-7}$	$2.3 * 10^{-5}$	9.6

(c) y as in (a); y is randomly perturbed by $\delta_n = n^{-2}\|y\|$.

n	α_n	$\tilde{e}_n$	$\tilde{e}_n \cdot n \cdot 10^2$
4	$2.0 * 10^{-4}$	$8.3 * 10^{-3}$	3.3
8	$5.3 * 10^{-5}$	$1.5 * 10^{-2}$	12.4
16	$1.3 * 10^{-5}$	$6.2 * 10^{-3}$	9.8
32	$3.3 * 10^{-6}$	$2.6 * 10^{-3}$	8.2
64	$8.3 * 10^{-7}$	$1.3 * 10^{-3}$	8.3

(d) y as in (b), y is randomly perturbed by $\delta_n = n^{-3}\|y\|$.

n	α_n	$\tilde{e}_n$	$\tilde{e}_n \cdot n^2 \cdot 10$
4	$1.3 * 10^{-4}$	$5.8 * 10^{-3}$	0.93
8	$3.3 * 10^{-5}$	$3.8 * 10^{-3}$	2.4
16	$8.2 * 10^{-6}$	$8.3 * 10^{-4}$	2.1
32	$2.0 * 10^{-6}$	$1.8 * 10^{-4}$	1.8
64	$5.1 * 10^{-7}$	$4.1 * 10^{-5}$	1.7

Again, computations with larger values for δ_n give comparable results.

(e) $y(t) = \frac{1}{6}(t - t^3)$, $(T^\dagger y)(s) = s$. Note that (1.8) is not fulfilled. Nevertheless, we use the same parameters (except for the values of D_1 and D_2) as in (a) and obtain:

n	α_n	e_n	$e_n \cdot n^{1/2} \cdot 10$
4	$2.8 * 10^{-6}$	$2.7 * 10^{-1}$	5.37
8	$6.8 * 10^{-7}$	$1.9 * 10^{-1}$	5.38
16	$1.7 * 10^{-7}$	$1.4 * 10^{-1}$	5.40
32	$4.1 * 10^{-8}$	$9.9 * 10^{-2}$	5.58

The last column shows that instead of the rate $O(n^{-1})$, which would be achieved if (1.8) were fulfilled, the rate seems to be $O(n^{-1/2})$ here. Incidentally, the errors are significantly larger than in (a) and a bit (but not much) smaller than in [9, Example c]. This confirms Marti's observation that for this example, Tikhonov regularization using the L^2-norm is not good. If one uses Tikhonov regularization using $\|x''\|^2$ instead of $\|x\|^2$, the results are better (cf. [3, Example 2]).

Institut für Mathematik
Johannes-Kepler-Universität
A-4040 Linz, Austria

1. H. W. ENGL, "On the convergence of regularization methods for ill-posed linear operator equations," in *Improperly Posed Problems and Their Numerical Treatment* (G. Hämmerlin and K. H. Hoffmann, eds.), Birkhäuser, Basel, 1983, pp. 81–95.

2. H. W. ENGL, "Discrepancy principles for Tikhonov regularization of ill-posed problems leading to optimal convergence rates," *J. Optim. Theory Appl.* (To appear.)

3. H. W. ENGL & A. NEUBAUER, "Optimal discrepancy principles for the Tikhonov-regularization of integral equations of the first kind," in *Constructive Methods for the Practical Treatment of Integral Equations* (G. Hämmerlin and K. H. Hoffmann, eds.), Birkhäuser, Basel, 1985, pp. 120–141.

4. C. W. GROETSCH, "The parameter choice problem in linear regularization," in *Ill-Posed Problems, Theory and Practise* (M. Z. Nashed, ed.). (To appear.)

5. C. W. GROETSCH, "Comments on Morozov's discrepancy principle," in *Improperly Posed Problems and Their Numerical Treatment* (G. Hämmerlin and K. H. Hoffmann, eds.), Birkhäuser, Basel, 1983, pp. 97–104.

6. C. W. GROETSCH, J. T. KING & D. MURIO, "Asymptotic analysis of a finite element method for Fredholm equations of the first kind," in *Treatment of Integral Equations by Numerical Methods* (C. T. H. Baker and G. F. Miller, eds.), Academic Press, London, 1982, pp. 1–11.

7. C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.

8. K. HICKEY & G. LUECKE, "Remarks on Marti's method for solving first kind equations," *SIAM J. Numer. Anal.*, v. 19, 1982, pp. 623–628.

9. J. T. MARTI, "An algorithm for computing minimum norm solutions of Fredholm integral equations of the first kind," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 1071–1076.

10. J. T. MARTI, "On the convergence of an algorithm computing minimum-norm solutions of ill-posed problems," *Math. Comp.*, v. 34, 1980, pp. 521–527.

11. J. T. MARTI, "On a regularization method for Fredholm equations of the first kind using Sobolev spaces," in *Treatment of Integral Equations by Numerical Methods* (C. T. H. Baker and G. F. Miller, eds.), Academic Press, London, 1982, pp. 59–66.

12. M. Z. NASHED (ed.), *Generalized Inverses and Applications*, Academic Press, New York, 1976.

13. A. MOROZOV, "On the solution of functional equations by the method of regularization," *Soviet Math. Dokl.*, v. 7, 1966, pp. 414–417.

14. A. N. TIKHONOV & V. Y. ARSENIN, *Solution of Ill-Posed Problems*, English transl., Wiley, New York, 1977.