THE METHOD OF CONJUGATE RESIDUALS FOR SOLVING THE GALERKIN EQUATIONS ASSOCIATED WITH SYMMETRIC POSITIVE SEMIDEFINITE ILL-POSED PROBLEMS

R. PLATO*

Abstract. For the numerical solution of the Galerkin equations associated with linear ill-posed problems that are symmetric and positive semidefinite, the method of conjugate residuals is considered. An a posteriori stopping rule is introduced, and associated estimates for the approximations are provided which are order-optimal with respect to noise in the right-hand side and with respect to the discretization error.

Key words. Ill-posed problems, first kind integral equations, conjugate gradient type methods, Galerkin method, regularization schemes, discrepancy principle, Symm's integral equation.

AMS subject classifications. 65J20, 65R30, 45E99, 65N38

1. Introduction. In this paper we consider equations of the form

$$Au = f_*, \tag{1.1}$$

where

$$A \in \mathcal{L}(\mathcal{H}), \qquad A = A^* \geq 0, \tag{1.2}$$

$$f_* \in \mathcal{R}(A). \tag{1.3}$$

Here $\mathcal{H}$ denotes a real Hilbert space with inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ and corresponding norm $\| \cdot \| : \mathcal{H} \to \mathbb{R}$, $\mathcal{L}(\mathcal{H})$ denotes the space of operators $T : \mathcal{H} \to \mathcal{H}$ that are bounded and linear, and $A^* \in \mathcal{L}(\mathcal{H})$ in (1.2) denotes the adjoint operator of A. Moreover, $\mathcal{R}(A)$ denotes the range of A which in general is non-closed, and then equation (1.1) is ill-posed. In the sequel we assume that only an approximation f^δ for f_* is available, more specifically,

$$f^\delta \in \mathcal{H}, \qquad \|f_* - f^\delta\| \leq \delta, \tag{1.4}$$

where $\delta \geq 0$ is a small and known error bound.

For practical reasons equation (1.1) has to be discretized, and in the sequel we shall consider Galerkin equations associated with (1.1). For this purpose let

$$\mathcal{V}_h \subset \mathcal{H}, \qquad 0 < h \leq h_0,$$

be given finite-dimensional linear subspaces. For fixed h, the corresponding Galerkin equations for (1.1) are given by

$$A_h u_h = \mathcal{P}_h f_*, \tag{1.5}$$

where

$$A_h : \mathcal{H} \to \mathcal{H}, \qquad u \mapsto \mathcal{P}_h A u, \tag{1.6}$$

* Fachbereich Mathematik, Technische Universität Berlin, Straße des 17. Juni 135, D - 10623 Berlin, Germany

and $\mathcal{P}_h$ denotes the orthogonal projection onto $\mathcal{V}_h$, i.e.,

$$\mathcal{P}_h \in \mathcal{L}(\mathcal{H}), \qquad \mathcal{P}_h = \mathcal{P}_h^2, \qquad \mathcal{P}_h = \mathcal{P}_h^*, \qquad \mathcal{R}(\mathcal{P}_h) = \mathcal{V}_h. \tag{1.7}$$

In order to determine a solution $u_* \in \mathcal{H}$ of equation (1.1) with noisy data as in (1.4) we shall apply the method of conjugate residuals to the Galerkin equations (1.5) (see the next section) associated with a discrepancy principle as stopping rule (cf. Section 3 for its introduction), and this shall be done for the following reasons:

1. Since the underlying equation (1.1) is ill-posed, usually also the finite-dimensional version (1.5) has to be regularized. It is shown in this paper that the method of conjugate residuals associated with the mentioned discrepancy principle as a stopping rule has this desired regularizing property.
2. Typically only a small number of iteration steps is needed until the method of conjugate residuals terminates according to the discrepancy principle, and thus this yields a numerically efficient solver of equation (1.1).

2. The method of conjugate residuals for the Galerkin equations.

2.1. Description of the method. We next introduce the method of conjugate residuals for the Galerkin equations (1.5), and for this purpose in the sequel we assume that the noise level $\delta > 0$ and the discretization level $h > 0$ in conditions (1.2)-(1.4) and in (1.6)-(1.7) are fixed. Moreover we introduce the Krylov subspaces with respect to A_h and a vector $r \in \mathcal{V}_h$,

$$\mathcal{K}_n(A_h, r) \;=\; \operatorname{span}\Big\{ r, A_h r, \ldots, A_h^{n-1} r \Big\} \;\subset\; \mathcal{V}_h, \qquad n = 0, 1, \ldots \;.$$

DEFINITION 2.1. *Let conditions (1.2)–(1.4) be fulfilled, and let $\mathcal{P}_h,\ A_h \in \mathcal{L}(\mathcal{H})$ be as in (1.7), (1.6). The method of conjugate residuals, applied to (1.5) with noisy data as in (1.4), iteratively generates a (terminating) sequence $u_0 = 0, u_1, u_2, \ldots,$ with $u_n \in \mathcal{V}_h$ and*

$$\left\{ \begin{array}{rcl} u_n & \in & \mathcal{K}_n(A_h, \mathcal{P}_h f^\delta), \\ \|A_h u_n - \mathcal{P}_h f^\delta\| & = & \inf_{u \in \mathcal{K}_n(A_h, \mathcal{P}_h f^\delta)} \|A_h u - \mathcal{P}_h f^\delta\|, \end{array} \right. \tag{2.1}$$

and the algorithm terminates, by definition, at step $n_* := n \le \dim \mathcal{V}_h$ if $A_h r_n = 0$. Here, r_n denotes the residual, i.e.,

$$r_n \;=\; A_h u_n - \mathcal{P}_h f^\delta \in \mathcal{V}_h, \qquad n = 0, 1, \ldots, n_*. \tag{2.2}$$

Remarks. 1. For notational convenience, $u_0 = 0$ is taken as starting vector for the method of conjugate residuals, and the dependence of δ and h on vectors and scalars that arise in the course of iteration is not stated explicitly.

2. The stopping index $\overline{n}$ to be defined in Definition 3.2 below usually is much smaller than the final index n_*.

3. For technical reasons we shall not consider here the classical method of conjugate gradients where the iterates minimize the energy functional $J(u) = \frac{1}{2}\langle A_h u, u\rangle - \langle u, \mathcal{P}_h f^\delta\rangle$ over $\mathcal{K}_n(A_h, \mathcal{P}_h f^\delta),\ n = 0, 1, \ldots\,.$ △

The basic algorithm for computing u_n given by Definition 2.1 is:

ALGORITHM 2.2. *(Method of conjugate residuals for (1.5) with noisy given right-hand side) Let conditions (1.2)–(1.4) be fulfilled, and let $\mathcal{P}_h,\ A_h \in \mathcal{L}(\mathcal{H})$ be as in (1.7), (1.6). <u>Step 0:</u> Let $u_0 := 0,\ r_0 = -\mathcal{P}_h f^\delta$.*

For $n = 0, 1, \dots$:

1) If $A_h r_n = 0$ *then terminate,* $n_* := n$*;*

2) If otherwise $A_h r_n \neq 0$*, then proceed with* <u>*step n+1*</u>*: compute from* $u_n,\ d_{n-1}$

$$d_n \;=\; -r_n \;+\; \beta_{n-1} d_{n-1}, \qquad \beta_{n-1} \;=\; \frac{\langle A_h r_n, r_n\rangle}{\langle A_h r_{n-1}, r_{n-1}\rangle}, \tag{2.3}$$

$$u_{n+1} \;=\; u_n \;+\; \alpha_n d_n, \qquad \alpha_n \;=\; \frac{\langle A_h r_n, r_n\rangle}{\|A_h d_n\|^2}. \tag{2.4}$$

Here we assume $d_{-1} = 0,\ \beta_{-1} = 0$.

It follows from (2.2)–(2.4) that for $0 \le n \le n_* - 1$ we have

$$A_h d_n \;=\; -A_h r_n \;+\; \beta_{n-1} A_h d_{n-1}, \qquad r_{n+1} \;=\; r_n \;+\; \alpha_n A_h d_n, \tag{2.5}$$

and in fact in any step for computational reasons $A_h d_n$ and r_{n+1} are computed as in (2.5) so that only one operator-vector multiplication (to obtain $A_h r_n$) has to be performed in each step.

2.2. Matrix formulation of the method of conjugate residuals for the Galerkin equations. A matrix formulation of the method of conjugate residuals for positive definite linear systems of equations is presented, e.g., in Stoer [27], and for other surveys on conjugate gradient type methods we refer to Ashby, Manteuffel & Saylor [1] and Freund, Golub & Nachtigal [6].

For completeness we present a matrix formulation of the method of conjugate residuals for our specific situation (1.5), and for this purpose we denote by N the dimension of $\mathcal{V}_h$. Then let $\Psi_1, \Psi_2, \dots, \Psi_N \in \mathcal{V}_h$ be a basis of $\mathcal{V}_h$, and let

$$\left\{\begin{array}{rcl} \mathbf{G} &=& \big(\langle \Psi_j, \Psi_i\rangle\big) \;\in\; \mathbb{R}^{(N,N)} \\ \mathbf{B} &=& \big(\langle A\Psi_j, \Psi_i\rangle\big) \;\in\; \mathbb{R}^{(N,N)} \\ \mathbf{f} &=& \big(\langle f^\delta, \Psi_i\rangle\big) \;\in\; \mathbb{R}^{N} \end{array}\right. \tag{2.6}$$

The approximations $u_n \in \mathcal{V}_h,\ n = 0, 1, \dots, n_*$, defined by Algorithm 2.2 then can be represented as follows,

$$u_n \;=\; \sum_{j=1}^{N} \mathbf{u}_{n,j} \Psi_j,$$

where $\mathbf{u}_n = (\mathbf{u}_{n,j}) \in \mathbb{R}^N$ is determined by the following algorithm:

ALGORITHM 2.3. *(Method of conjugate residuals for (1.5) with noisy given right-hand side, matrix formulation) Let* $\mathbf{G},\ \mathbf{B}$ *and* $\mathbf{f}$ *as in (2.6).* <u>*Step 0:*</u> *Let* $\mathbf{u}_0 := \mathbf{0} \in \mathbb{R}^N,\ \mathbf{r}_0 = -\mathbf{G}^{-1}\mathbf{f} \in \mathbb{R}^N$.

For $n = 0, 1, \dots$:

1) If $\mathbf{B}\mathbf{r}_n = 0$ *then terminate,* $n_* = n$*;*

2) If otherwise $\mathbf{B}\mathbf{r}_n \neq 0$*, then proceed with* <u>*step n+1*</u>*: compute from* $\mathbf{u}_n,\ \mathbf{d}_{n-1}$*:*

$$\mathbf{d}_n \;=\; -\mathbf{r}_n \;+\; \beta_{n-1} \mathbf{d}_{n-1}, \qquad \beta_{n-1} \;=\; \frac{\mathbf{r}_n^T \mathbf{B} \mathbf{r}_n}{\mathbf{r}_{n-1}^T \mathbf{B} \mathbf{r}_{n-1}},$$

$$\mathbf{u}_{n+1} \;=\; \mathbf{u}_n \;+\; \alpha_n \mathbf{d}_n, \qquad \alpha_n \;=\; \frac{\mathbf{r}_n^T \mathbf{B} \mathbf{r}_n}{\mathbf{e}_n^T \mathbf{G} \mathbf{e}_n},$$

$$\textit{where}\quad \mathbf{e}_n := \mathbf{G}^{-1}\mathbf{B}\mathbf{d}_n.$$

Here $\mathbf{d}_{-1} = \mathbf{0}$, $\beta_{-1} = 0$.

Note that the numbers α_n and β_{n-1} in Algorithms 2.2 and 2.3 coincide, and the vectors $\mathbf{d}_n \in \mathbb{R}^N$ and $\mathbf{r}_n \in \mathbb{R}^N$ are the coordinates of $d_n \in \mathcal{V}_h$ and $r_n \in \mathcal{V}_h$, respectively, i.e.,

$$d_n \ = \ \sum_{j=1}^N \mathbf{d}_{n,j}\Psi_j, \qquad\qquad r_n \ = \ \sum_{j=1}^N \mathbf{r}_{n,j}\Psi_j.$$

Note also that the vectors $\mathbf{B}\mathbf{d}_n,\ \mathbf{r}_{n+1} \in \mathbb{R}^N$ can be computed efficiently for $0 \le n \le n_* - 1$,

$$\mathbf{B}\mathbf{d}_n \ = \ -\mathbf{B}\mathbf{r}_n \ + \ \beta_{n-1}\mathbf{B}d_{n-1}, \qquad\qquad \mathbf{r}_{n+1} \ = \ \mathbf{r}_n \ + \ \alpha_n\mathbf{G}^{-1}\mathbf{B}\mathbf{d}_n.$$

3. The main section.

3.1. Approximation properties of the subspaces $\mathcal{V}_h$. For a symmetric and positive semidefinite operator $A \in \mathcal{L}(\mathcal{H})$ and arbitrary real $\nu > 0$ we next define ν-norms on $\mathcal{R}(A^\nu)$,

$$\|u\|_\nu := \min\Big\{ \ \|z\| : \quad z \in \mathcal{H},\ A^\nu z = u \ \Big\}, \qquad u \in \mathcal{R}(A^\nu). \tag{3.1}$$

Remark. Fractional powers $A^\nu \in \mathcal{L}(\mathcal{H}),\ \nu > 0$, as well as their elementary properties are presented e.g., in a more general framework, in Fattorini [5], Chapter 6.3. For symmetric and positive semidefinite operators $A \in \mathcal{L}(\mathcal{H})$ that are compact and have an infinite-dimensional range $\mathcal{R}(A)$, the fractional powers $A^\nu \in \mathcal{L}(\mathcal{H})$ can be introduced in a simplified manner. In fact, the following spectral representation of A is valid then,

$$A \ = \ \sum_{j=1}^\infty \lambda_j \mathcal{Q}_j,$$

where $\lambda_1 > \lambda_2 > \ldots > 0$ denote the pairwise distinct, non-vanishing eigenvalues of $A \in \mathcal{L}(\mathcal{H})$, and $\mathcal{Q}_j \in \mathcal{L}(\mathcal{H})$ denote the associated orthogonal projections onto $\mathcal{N}(A - \lambda_j I)$ for $j \in \mathbb{N}$; the fractional powers $A^\nu \in \mathcal{L}(\mathcal{H})$ then are given by

$$A^\nu \ = \ \sum_{j=1}^\infty \lambda_j^\nu \mathcal{Q}_j. \qquad\qquad \triangle$$

In the sequel we shall assume that the subspaces $\{\mathcal{V}_h\}$ fulfill an *approximation property* with respect to A, this is, for some integer

$$\nu_1 \ge 1$$

and some known

$$0 < \xi_h \le 1, \quad 0 < h \le h_0, \qquad \text{with} \quad \xi_h \to 0 \quad \text{as } h \to 0, \tag{3.2}$$

we suppose that for certain constants $a_\nu > 0$

$$\forall\ 0 < \nu \le \nu_1 : \quad \inf_{v_h \in \mathcal{V}_h} \|u - v_h\| \ \le \ a_\nu \xi_h^\nu \|u\|_\nu, \quad u \in \mathcal{R}(A^\nu), \quad 0 < h \le h_0, \tag{3.3}$$

holds. Note that the approximation property (3.3) is equivalent to

$$\forall\, 0 < \nu \le \nu_1 : \qquad \|(I - \mathcal{P}_h)A^\nu\| \;\le\; a_\nu \xi_h^\nu, \qquad 0 < h \le h_0, \tag{3.4}$$

with $\mathcal{P}_h \in \mathcal{L}(\mathcal{H})$ as in (1.7). Note moreover that if conditions (3.2), (3.4) are valid then A necessarily is a compact operator. Finally we observe that the approximation property (3.4) implies

$$\forall\, 0 < \nu \le \nu_1 : \qquad \|A^\nu(I - \mathcal{P}_h)\| \;\le\; a_\nu \xi_h^\nu, \qquad 0 < h \le h_0, \tag{3.5}$$

which follows from the fact that A^ν and $\mathcal{P}_h$ are symmetric operators. The preceding notations are summarized in the following basic assumption:

Assumption 3.1. *1. $\mathcal{H}$ denotes a real Hilbert space, and $A \in \mathcal{L}(\mathcal{H})$ is symmetric and positive semidefinite.*

2. For $0 < h \le h_0$ let $\mathcal{V}_h \subset \mathcal{H}$ be a finite-dimensional linear subspace and let $\mathcal{P}_h,\ A_h \in \mathcal{L}(\mathcal{H})$ be as in (1.7), (1.6). We suppose that the approximation property (3.3) is fulfilled for some $\nu_1 \ge 1$, with ξ_h as in (3.2). In the sequel h is supposed to be fixed.

3. Let $u_* \in \mathcal{R}(A^\nu)$ with some $\nu > 0$, and let $\delta \ge 0$ and $f^\delta \in \mathcal{H}$ with

$$\|Au_* - f^\delta\| \;\le\; \delta.$$

4. Finally we introduce the numbers

$$\begin{aligned} \varrho &:= \|u_*\|_\nu, \\ b_\nu &:= a_{\min\{\nu,\nu_1\}}\|A\|^{\max\{0,\nu-\nu_1\}}, \\ \eta(h,\delta) &:= \delta \;+\; a_1 b_\nu \varrho \xi_h^{\min\{\nu,\nu_1\}+1}, \end{aligned} \tag{3.6}$$

cf. (3.1) for the definition of $\|\cdot\|_\nu$.

From Assumption 3.1 it follows immediately that

$$\|(I - \mathcal{P}_h)u_*\| \;\le\; b_\nu \varrho \xi_h^{\min\{\nu,\nu_1\}}, \tag{3.7}$$

$$\|A\mathcal{P}_h u_* - f^\delta\| \;\le\; \eta(h,\delta), \tag{3.8}$$

which are estimates that shall be used at several occasions.

3.2. The discrepancy principle as stopping rule for the method of conjugate residuals. Assume that the iteration process, described by Algorithm 2.2, generates iterates $0 = u_0,\ u_1,\ u_2, \ldots, u_{n_*} \in \mathcal{V}_h$, where $0 \le n_* < \infty$ denotes the final iteration step. Then for any $0 \le n \le n_*$ there exists a unique polynomial (depending on A_h and $\mathcal{P}_h f^\delta$)

$$q_n \in \Pi_{n-1} \tag{3.9}$$

such that

$$u_n \;=\; q_n(A_h)\mathcal{P}_h f^\delta. \tag{3.10}$$

Here, $\Pi_{-1} := \{0\}$, and $\Pi_{n-1} = \big\{\ q : q \text{ is a polynomial of degree } \le n-1\ \big\},\ n = 1, 2, \ldots$. We next introduce

$$\Delta_n := \|A_h u_n - \mathcal{P}_h f^\delta\|, \qquad 0 \le n \le n_*, \tag{3.11}$$

the norm of the residual. From (2.1) we obtain

$$\Delta_n \;\le\; \Delta_{n-1}, \qquad 1 \le n \le n_*, \tag{3.12}$$

and typically Δ_n decays fast and becomes small after a small number of iterations n. In the infinite-dimensional setting ($\mathcal{P}_h = I$), a regularization method can be obtained by stopping the iteration when $\Delta_n \approx \delta$ and taking $u_n \in \mathcal{H}$ then as approximation for the desired solution, for references see Subsection 3.3.2. In our situation ($\mathcal{P}_h \neq I$) several other cases, however, has to be taken into account to define a stopping rule as well as a corresponding approximation that guarantees best possible convergence rates, and in fact the precise stopping criterion is presented next.

DEFINITION 3.2. *(A discrepancy principle as stopping rule) Suppose that Assumption 3.1 holds. Let $u_n \in \mathcal{V}_h,\ n = 0, 1, \ldots,$ be generated by the method of conjugate residuals, and let $b > 1$. Stop iteration at step $\overline{n} := n$, if*

$$\Delta_n \;\le\; b\delta \qquad \textit{or} \qquad q_n(0) \;\ge\; \xi_h^{-1} \qquad \textit{or} \qquad n = n_*.$$

Then define

$$\overline{u}(h,\delta) := \begin{cases} u_{\overline{n}-1}, & \textit{if } q_{\overline{n}}(0) \;\ge\; \xi_h^{-1}, \\ u_{\overline{n}}, & \textit{if } q_{\overline{n}}(0) \;<\; \xi_h^{-1}. \end{cases} \tag{3.13}$$

We remark that the numbers $q_n(0),\ 0 \le n \le n_*$, increase as n increases, cf. also Lemma 5.1 below. Moreover, $q_n(0)$ can be computed easily from the three-term recurrence

$$\left\{ \begin{aligned} q_0(0) \;&=\; 0, \qquad q_1(0) = \alpha_0, \\ q_{n+1}(0) \;&=\; \Big(1 + \tfrac{\alpha_n\beta_{n-1}}{\alpha_{n-1}}\Big) q_n(0) \;-\; \tfrac{\alpha_n\beta_{n-1}}{\alpha_{n-1}} q_{n-1}(0) \;+\; \alpha_n, \\ & 1 \le n \le n_* - 1, \end{aligned} \right. \tag{3.14}$$

which follows immediately from the first equalities in (2.3) and (2.4) and from the uniqueness of $q_n(t)$ in (3.9), (3.10). Note that (3.14) in particular means that the situation $q_{\overline{n}}(0) \ge \xi_h^{-1}$ may arise only for $\overline{n} \ge 1$, i.e., the definition (3.13) makes sense.

3.3. Statement of the main result. We next present the main result of this paper.

THEOREM 3.3. *Suppose that Assumption 3.1 holds. Moreover, let $u_n \in \mathcal{V}_h,\ n = 0, 1, \ldots$ be defined by Algorithm 2.2, and let the iteration be stopped according to the stopping rule presented in Definition 3.2, with corresponding approximation $\overline{u}(h,\delta) \in \mathcal{V}_h$. Then*

$$\| u_* - \overline{u}(h,\delta) \| \;\le\; e_\nu \Big(\big(\varrho\delta^\nu\big)^{1/(\nu+1)} \;+\; \varrho\xi_h^{\min\{\nu,\nu_1\}} \Big). \tag{3.15}$$

The constant e_ν is independent of δ, h and ϱ (and depends on ν and b).

The proof of Theorem 3.3 shall be given in Section 5.

3.3.1. Conclusions. (1) The estimate in (3.15) is order-optimal with respect to the noise in the right-hand side and with respect to the discretization, respectively. Moreover, no knowledge of the degree of smoothness of the solution u_* is required to obtain the approximations $\overline{u}(h,\delta)$. Finally, no *inverse property* associated with quasi-uniformity of the spaces $\mathcal{V}_h$ is needed and no *stability* of the Galerkin method for the solution of (1.1) is required (both properties usually are needed for the convergence analysis of Galerkin methods).

(2) Suppose that u_* is a solution of equation (1.1) that satisfies $u_* \in \mathcal{R}(A^\nu)$ for some $0 < \nu \le \nu_1$. If h is chosen sufficiently small, i.e., if $\xi_h^{\nu+1} \le \delta$ is satisfied, then Theorem 3.3 guarantees an order-optimal estimate

$$\|u_* - \overline{u}(h,\delta)\| \;=\; \mathcal{O}\big(\delta^{\nu/(\nu+1)}\big)$$

for the approximations.

(3) In Theorem 3.3, the noise-free case $\delta = 0$ itself is of interest. Note that the stopping rule given in Definition 3.2 as well as the corresponding error estimate in Theorem 3.3 makes sense also in this case.

(4) In [25] a similar approach is carried out for the classical conjugate gradient method of Hestenes and Stiefel applied to a normalized system of equations associated with arbitrary linear ill-posed problems in Hilbert spaces.

(5) We note that the operator A_h can be conceived as a specific perturbation of the operator $A \in \mathcal{L}(\mathcal{H})$. More generally as in our situation, Nemirovskiĭ [21] considers conjugate gradient type methods for linear ill-posed problems with *arbitrary* linear perturbations of the underlying operator; for example, in the situation (1.2)–(1.4) the method of conjugate residuals is applied to the perturbed equation $A_\eta u = f^\delta$ where $A_\eta \in \mathcal{L}(\mathcal{H})$ is an arbitrary positive semidefinite operator satisfying $\|A_\eta - A\| \le \eta$. It is shown in [21] for a specific stopping rule providing approximations $\overline{u}(\eta,\delta) \in \mathcal{H}$ that an estimate of the following kind is satisfied, $\|u_* - \overline{u}(\eta,\delta)\| \;=\; \mathcal{O}\big((\eta+\delta)^{\nu/(\nu+1)}\big)$. We thus can conclude that the result (3.15) improves the result obtained in [21], if one considers specific operator perturbations generated by projection methods.

3.3.2. Further bibliographical remarks on conjugate gradient type methods for linear ill-posed problems. In this subsection we refer to related results obtained for conjugate gradient type methods to solve linear ill-posed problems.

(1) Convergence results for precise data are presented in Kammerer & Nashed [14], Nemirovskiĭ & Polyak ([22], [23]), Brakhage [2], Louis [19] and Hanke [13].

(2) Results for noisy right-hand sides (and precisely given operators) are obtained e.g., in King [15], Lardy [18], in Eicke, Louis & Plato [3], Plato ([24], [26]), Hanke [12], and in Gilyazov [8]. For recent monographs containing associated results we refer to Gilyazov [7], Hanke [11], Engl, Hanke & Neubauer [4] and Kirsch [16].

4. Numerical Illustrations.

4.1. The spaces. In our numerical experiments, as underlying space we consider the real space of square-integrable real-valued functions on the interval $[0,1]$,

$$\mathcal{H} = L^2[0,1],$$

supplied with the inner product

$$\langle u, v\rangle \;=\; \int_0^1 u(t)v(t)\,dt, \qquad u,\ v \in L^2[0,1],$$

and the corresponding norm is

$$\|u\| \;=\; \langle u, u\rangle^{1/2}, \qquad u \in L^2[0,1].$$

For the Galerkin scheme (1.5) we use spaces of linear splines

$$\mathcal{V}_h \;=\; \Big\{\; u_h \in C[0,1] : \; u_h \text{ is linear on } [t_{j-1}, t_j], \quad j = 2,\ldots,N \;\Big\}, \tag{4.1}$$

where

$$\begin{aligned} h &= 1/(N-1), \\ t_j &= (j-1)h, \qquad j = 1, 2, \ldots, N, \end{aligned}$$

and $C[0,1]$ denotes the space of real-valued continuous functions defined on the interval $[0,1]$. The following approximation property is valid for those functions $\psi : [0,1] \to \mathbb{R}$ where ψ and ψ' are absolutely continuous functions and where $\psi'' \in L^2[0,1]$, cf. Hackbusch [9], Chapter 4.5.9:

$$\|(I - \mathcal{P}_h)\psi\| \le \frac{1}{\sqrt{90}} h^2 \|\psi''\|, \qquad h > 0. \tag{4.2}$$

Finally, as basis functions for $\mathcal{V}_h$ the standard hat functions $\Psi_j \in \mathcal{V}_h,\ j = 1, \ldots, N$, are taken, i.e., one has

$$\Psi_j(t_k) = \begin{cases} 1, & k = j \\ 0, & k \ne j \end{cases},$$

and the Gram matrix $\mathbf{G} = \big(\langle \Psi_j, \Psi_i \rangle\big)$ then has the form

$$\mathbf{G} = \frac{h}{6} \begin{bmatrix} 2 & 1 & 0 & \ldots & 0 \\ 1 & 4 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4 & 1 \\ 0 & \ldots & 0 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{(N,N)}.$$

4.2. Symm's integral equation for circles Γ.

4.2.1. Introduction. In our numerical experiments, for circles $\Gamma = \Gamma_\rho = \{ x \in \mathbb{R}^2 : |x| = \rho \}$ with radius ρ we shall consider Symm's weakly singular integral equation which for a standard parametrization of Γ_ρ looks as follows,

$$(Au)(t) := -\int_0^1 \log|2\rho \sin \pi(t-s)| u(s)\, ds = f(t), \qquad t \in [0,1], \tag{4.3}$$

see Vainikko [29] or Chapter 3.3 in Kirsch [16] for an introduction. It turns out that for $0 < \rho \le 1$, the operator $A : L^2[0,1] \to L^2[0,1]$ is compact, symmetric and positive semidefinite, and the following decomposition is valid (here presented in complex form),

$$(Au)(t) = -(\log \rho)\hat{u}(0) + \frac{1}{2} \sum_{0 \ne n \in \mathbb{Z}} \frac{1}{|n|} \hat{u}(n) e^{i2\pi nt}, \qquad t \in [0,1], \tag{4.4}$$

where $\hat{u}(n) = \int_0^1 u(t) e^{-i2\pi nt}\, dt$ denotes the n-th Fourier coefficient of $u \in L^2[0,1]$. It follows from the decomposition (4.4) that equation (4.3) is modestly ill-posed.

We next show that the approximation property (3.3) is valid for $\xi_h = h,\ \nu_1 = 2$ $\big($and for $\mathcal{H} = L^2[0,1]$, for the operator A as in (4.3) and for the subspaces $\mathcal{V}_h$ as in (4.1)$\big)$. For this purpose let $H^\lambda[0,1],\ \lambda > 0$, be the Sobolev space of (real-valued) functions $u \in L^2[0,1]$ with

$$\|u\|_{H^\lambda} := \Big(|\hat{u}(0)|^2 + \sum_{0 \ne n \in \mathbb{Z}} |n|^{2\lambda} |\hat{u}(n)|^2 \Big)^{1/2} < \infty.$$

From the eigenvalue decomposition (4.4) of A one easily derives

$$\mathcal{R}(A^\nu) \;=\; H^\nu[0,1], \qquad \nu > 0, \tag{4.5}$$

for $0 < \rho < 1$; for $\rho = 1$, "=" in (4.5) has to be replaced by "⊂". Moreover we see that $\sup_{0\neq u \in L^2[0,1]} \|A^2 u\|_{H^2}/\|u\| < \infty$, and then the estimate (4.2) yields that the estimate in (3.3) is valid for the specific case $\nu = 2$ (and $\xi_h = h$); the interpolation inequality (cf. Fattorini [5], Examples 6.3.6 and 6.3.7) then finally yields that the approximation property in (3.3) is valid in its general form.

4.2.2. Specific right-hand sides. In the numerical experiments with Symm's integral equation for circles the following radius is chosen,

$$\rho = \frac{1}{2},$$

and as right-hand side we consider

$$\begin{aligned} f_*(t) \;&=\; \begin{cases} (2\pi t)^2, & 0 \le t \le 0.5, \\ \big(2\pi(1-t)\big)^2, & 0.5 \le t \le 1, \end{cases} \\ &=\; \frac{\pi^2}{3} \;+\; 4\sum_{n=1}^\infty \frac{(-1)^n}{n^2}\cos(2\pi n t), \qquad 0 \le t \le 1. \end{aligned}$$

Then the function

$$u_*(s) \;=\; \frac{\pi^2}{3\log 2} \;+\; 8\sum_{n=1}^\infty \frac{(-1)^n}{n}\cos(2\pi n s), \qquad 0 \le s \le 1,$$

solves the equation $Au = f_*$, where A is as in (4.3), and we obviously have

$$u_* \in H^\lambda[0,1] \quad \text{for} \quad \lambda < \frac{1}{2}, \tag{4.6}$$

$$u_* \notin H^{1/2}[0,1]. \tag{4.7}$$

4.2.3. Specific choice of the parameters, and the experiments. In the numerical experiments we choose perturbed right-hand sides $f^\delta = f_* + \delta \cdot v$, where $v \in \mathcal{H}$ has uniformly distributed random values with $\|v\| \le 1$, and where

$$\delta \;=\; \|f_*\| \cdot \%/100,$$

with % noise $\in \{\; 0.11,\; 0.33,\; 1.0,\; 3.0,\; 9.0 \;\}$ in the implementations. The dimension of the underlying system of equations is chosen as follows,

$$N = 128.$$

Table 4.1 contains the results for the method of conjugate residuals, cf. Algorithm 2.2, which is terminated by the stopping rule described in Definition 3.2, with

$$b = 1.5.$$

All computations are performed in MATLAB on an IBM RISC/6000.

Due to (4.5), (4.7) one cannot derive from Theorem 3.3 that the entries in the third column stay bounded as % of noise decreases. On the other hand, however, due to (4.5), (4.6) it is no surprise that these entries in fact stay bounded in our experiments.

TABLE 4.1
Numerical results for Symm's integral equation (4.3) for the circle $\Gamma_{1/2}$

% noise	$\Vert\overline{u}(h,\delta)-u_*\Vert$	$\Vert\overline{u}(h,\delta)-u_*\Vert/\big(\delta^{1/3}+h^{1/2}\big)$	$\overline{n}$	♯ flops
9.0	3.01	3.65	2	0.93e+06
3.0	1.93	3.22	3	0.98e+06
1.0	1.15	2.60	4	1.01e+06
0.33	0.76	2.29	5	1.05e+06
0.11	0.37	1.42	7	1.12e+06

4.3. Harmonic continuation of a function.

4.3.1. Introduction. To illustrate the results numerically once more, we next consider the problem of harmonic continuation. For this purpose let

$$\mathbb{D} = \Big\{\ x \in \mathbb{R}^2 :\ x_1^2 + x_2^2 < 1\ \Big\}$$

be the open unit disk in the plane, let the function $v : \overline{\mathbb{D}} \to \mathbb{R}$ be continuous on $\overline{\mathbb{D}}$, the closure of $\mathbb{D}$, and let v be harmonic on $\mathbb{D}$, i.e.,

$$(\Delta v)(x) \ = \ \Big(\frac{\partial^2 v}{\partial x_1^2} \ + \ \frac{\partial^2 v}{\partial x_2^2}\Big)(x) \ = \ 0, \qquad x \in \mathbb{D}.$$

The problem then can be described as follows: we assume that v is known approximately on the boundary of a concentric disk of radius $0 < \rho < 1$, i.e., the function

$$f(t) \ = \ v(\rho, 2\pi t), \qquad 0 \le t \le 1,$$

is assumed to be known approximately; from these informations we wish to determine v on the boundary of $\mathbb{D}$, i.e.,

$$u(s) \ = \ v(1, 2\pi s), \qquad 0 \le s \le 1,$$

is the unknown function which has to be determined. The correspondence between the functions u and f can be stated in terms of the following integral equation of the first kind (cf. Kress [17], Problem 15.3, or Mikhlin [20], Chapter 13):

$$(4.8) \qquad (Au)(t) := \int_0^1 k(t-s)u(s)\ ds \ = \ f(t), \qquad 0 \le t \le 1,$$

$$(4.9) \qquad k(t) := \frac{1-\rho^2}{1+\rho^2-2\rho\cos(2\pi t)}.$$

It turns out that for $0 < \rho < 1$, the operator $A : L^2[0,1] \to L^2[0,1]$ in (4.8) is compact, symmetric and positive semidefinite, and the following decomposition is valid (again given in complex form),

$$(4.10) \qquad (Au)(t) \ = \ \sum_{n\in\mathbb{Z}} \rho^{|n|}\hat{u}(n)e^{i2\pi nt}, \qquad t \in [0,1];$$

from this decomposition (4.10) it follows that (4.8)–(4.9) is severely ill-posed. The representation (4.10) yields moreover that for arbitrarily small $\nu_1 > 0$ one has $\mathcal{R}(A^{\nu_1}) \subset H^2[0,1]$ and

$$\sup_{0\neq u\in L^2[0,1]} \Vert A^{\nu_1}u\Vert_{H^2}/\Vert u\Vert < \infty,$$

and then estimate (4.2) and the interpolation inequality yield that the approximation property in (3.3) is valid in our situation for $\nu_1 = 1$ and $\xi_h = h^2$.

4.3.2. Specific right-hand sides. In our numerical illustrations for the problem of harmonic continuation we consider

$$f_*(t) \;=\; 1 \;+\; 2\sum_{n=1}^{\infty} \rho^{2n} \cos(2\pi n t), \qquad 0 \le t \le 1, \tag{4.11}$$

as right-hand side in (4.8), and the representation (4.10) yields that

$$u_*(s) \;=\; 1 \;+\; 2\sum_{n=1}^{\infty} \rho^{n} \cos(2\pi n s), \qquad 0 \le s \le 1,$$

solves $Au = f_*$, where A is as in (4.8), (4.9); moreover,

$$u_* \in \mathcal{R}(A^\nu) \ \text{ for } 0 < \nu < 1, \qquad\qquad u_* \notin \mathcal{R}(A),$$

and thus it is no big surprise that the entries in the third column in the following Table 4.2 stay bounded as % of noise decreases.

4.3.3. Specific choice of the parameters, and the experiments. The following table contains the results for the method of conjugate residuals, cf. Algorithm 2.2, which again is stopped according to the stopping rule described in Definition 3.2, with $b = 1.5$. The perturbations of the right-hand side f_* in (4.11) are employed similar as for Symm's integral equation in Section 4.2. The dimension of the underlying system of equations again is $N = 128$. The constant ρ is chosen as follows, $\rho = 1/2$.

TABLE 4.2
Numerical results for the problem of harmonic continuation

% noise	$\Vert\overline{u}(h,\delta) - u_*\Vert$	$\Vert\overline{u}(h,\delta) - u_*\Vert/(\delta^{1/2} + h^2)$	$\overline{n}$	$\sharp$ flops
9.0	0.35	1.13	2	0.94e+06
3.0	0.24	1.33	3	0.98e+06
1.0	0.11	1.07	3	0.98e+06
0.33	0.07	1.25	4	1.02e+06
0.11	0.04	1.17	5	1.05e+06

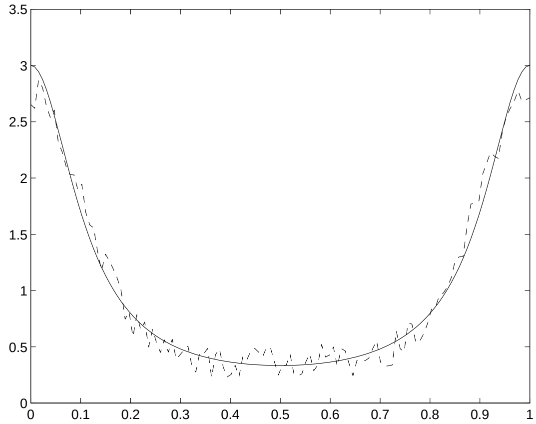
In Figure 4.1, the result for % noise = 1.0 is demonstrated.

5. Basic properties of the method of conjugate residuals. Throughout this section let $A_h \in \mathcal{L}(\mathcal{H})$ be as in (1.6). We start with a preliminary subsection on a spectral representation of A_h (cf. Section 5.1), and then (cf. Sections 5.2-5.4) we shall recall some well-known results for the method of conjugate residuals, and for the sake of convenience of the reader we frequently also provide the corresponding proofs.

5.1. Spectral decomposition of A_h. In the sequel, at several places we shall make use of the following (uniquely determined) spectral representation of A_h,

$$A_h \;=\; \sum_{j=1}^{m} \lambda_j \mathcal{Q}_j.$$

FIG. 4.1. *Exact solution u_* (solid line) and reconstruction $\overline{u}(h,\delta)$ (dashed line) for* 1.0% *noise in the right-hand side.*



Here, $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_m$ denote the pairwise distinct, non-vanishing eigenvalues of $A_h \in \mathcal{L}(\mathcal{H})$, and $\mathcal{Q}_j \in \mathcal{L}(\mathcal{H})$ denote the associated orthogonal projections onto $\mathcal{N}(A_h - \lambda_j I)$ for $j = 1, 2, \ldots, m$. For later notational convenience we also define

$$\lambda_0 = 0$$

and denote by $\mathcal{Q}_0 \in \mathcal{L}(\mathcal{H})$ the orthogonal projection onto $\mathcal{N}(A_h)$.

For any bounded function $\psi : \mathbb{R}_+ \to \mathbb{R}$, an operator $\psi(A_h) \in \mathcal{L}(\mathcal{H})$ is defined by

$$\psi(A_h) := \sum_{j=0}^{m} \psi(\lambda_j)\mathcal{Q}_j. \tag{5.1}$$

Note that for polynomials ψ, this definition (5.1) coincides with the usual meaning of $\psi(A_h)$. It is also useful to introduce the resolution of the identity,

$$\mathcal{F}_\tau := \sum_{j\geq 0:\ \lambda_j \leq \tau} \mathcal{Q}_j, \qquad \tau > 0, \tag{5.2}$$

and then we have

$$I - \mathcal{F}_\tau \ = \sum_{j>0:\ \lambda_j > \tau} \mathcal{Q}_j, \qquad \tau > 0. \tag{5.3}$$

The following properties will be useful: for any bounded function $\psi : \mathbb{R}_+ \to \mathbb{R}$ we have

$$\|\mathcal{F}_\tau \psi(A_h)\| \ \leq \ \sup_{0\leq t\leq \tau} |\psi(t)|, \qquad \tau > 0, \tag{5.4}$$

$$\mathcal{P}_h \psi(A_h) \ = \ \psi(A_h)\mathcal{P}_h, \tag{5.5}$$

where $\mathcal{P}_h \in \mathcal{L}(\mathcal{H})$ is as in (1.7). In fact, (5.4) follows immediately from the definitions, while (5.5) is obtained from the following facts: for any j there exists a polynomial p such that $p(A_h) = \mathcal{Q}_j$, and since $\mathcal{P}_h p(A_h) = p(A_h)\mathcal{P}_h$ is valid for each polynomial p, it is now obvious that $\mathcal{Q}_j \mathcal{P}_h \ = \ \mathcal{P}_h \mathcal{Q}_j$, and the definition (5.1) then finally yields (5.5).

5.2. First properties of the method of conjugate residuals. In the sequel we assume that Assumption 3.1 is fulfilled, and we assume that the iteration process described in Algorithm 2.2 generates iterates $u_0,\ u_1,\ldots,u_{n_*}\in\mathcal{V}_h$, where $n_*\geq 0$ denotes the final step. Then for the residuals $r_n = A_hu_n-\mathcal{P}_hf^\delta$, cf. (2.2), we have

$$r_n \;=\; -p_n(A_h)\mathcal{P}_hf^\delta,$$

for any $0\leq n\leq n_*$, with residual polynomials

$$p_n(t) \;=\; 1-tq_n(t), \tag{5.6}$$

cf. (3.9), (3.10) for the introduction of $q_n(t)$. We next state an important property of the method of conjugate residuals which is an immediate consequence of (2.1).

Minimum property. For $0\leq n\leq n_*$,

$$\|r_n\| \;=\; \|\big(I-A_hq_n(A_h)\big)\mathcal{P}_hf^\delta\| \;\leq\; \|\big(I-A_hq(A_h)\big)\mathcal{P}_hf^\delta\| \qquad \text{for all } \ q\in\Pi_{n-1},$$

or equivalently,

$$\|r_n\| \;\leq\; \|s(A_h)\mathcal{P}_hf^\delta\| \qquad \text{for any } s\in\Pi_n \ \text{ with } \ s(0)=1. \tag{5.7}$$

As an immediate consequence of (5.7) we obtain:

Conjugacy property. For $0\leq n\leq n_*$ we have $A_hr_n\in\mathcal{K}_n(A_h,\mathcal{P}_hf^\delta)^\perp$, the orthogonal complement of $\mathcal{K}_n(A_h,\mathcal{P}_hf^\delta)$, and thus

$$\begin{aligned} 0 \;&=\; \langle A_hp_n(A_h)\mathcal{P}_hf^\delta, s(A_h)\mathcal{P}_hf^\delta\rangle \\ &=\; \sum_{j=1}^{m}\lambda_jp_n(\lambda_j)s(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_hf^\delta\|^2 \end{aligned} \tag{5.8}$$

for all $s\in\Pi_{n-1}$.

5.3. Some properties of the residual polynomials $p_n(t)$. The first equalities in (2.3) and (2.4) and the uniqueness of $q_n(t)$ in (3.9), (3.10) imply the three-term recurrence

$$\begin{aligned} p_0(t) \;&=\; 1, \qquad p_1(t)=1-\alpha_0t, \\ p_{n+1}(t) \;&=\; \Big(1+\frac{\alpha_n\beta_{n-1}}{\alpha_{n-1}}-\alpha_nt\Big)p_n(t) \;-\; \frac{\alpha_n\beta_{n-1}}{\alpha_{n-1}}p_{n-1}(t), \qquad 1\leq n\leq n_*-1, \end{aligned}$$

which means that the system $p_0(t),\ldots,p_{n_*}(t)$ forms, up to a normalization, a Sturm sequence, cf. Chapter 5.6 in Stoer & Bulirsch [28]. Hence for $n=1,\ldots,n_*$, the zeros $\{t_{j,n}\}_{j=1,\ldots,n}$ of $p_n(t)$ are simple and fulfill an intertwining property; more explicitly, if the zeros are ordered,

$$0<t_{1,n}<t_{2,n}<\ldots<t_{n,n}, \qquad n=1,\ldots,n_*, \tag{5.9}$$

then

$$t_{k,n}<t_{k,n-1}<t_{k+1,n}, \qquad k=1,\ldots,n-1, \qquad n=2,\ldots,n_*, \tag{5.10}$$

is satisfied. Due to $p_n(0)=1$ we then have the following representation,

$$p_n(t) \;=\; \prod_{k=1}^{n}\Big(1-\frac{t}{t_{k,n}}\Big), \tag{5.11}$$

which implies

$$0\leq p_n(t)\leq 1 \qquad \text{for all } \ 0\leq t\leq t_{1,n}. \tag{5.12}$$

5.4. Some properties of the polynomials $q_n(t)$. Property (5.6) yields

$$q_n(t) \;=\; \frac{1 - p_n(t)}{t}, \qquad t > 0 \qquad\qquad (0 \le n \le n_*), \tag{5.13}$$

and further properties of $q_n(t)$ are listed in the following lemma.

LEMMA 5.1. *Let $\{q_n(t)\}_{0 \le n \le n_*}$ and $\{p_n(t)\}_{1 \le n \le n_*}$ be (arbitrary) polynomials fulfilling (3.9) and (5.6), respectively, and let the roots $\{t_{k,n}\}_{1 \le k \le n}$ of $p_n(t)$ be ordered as in (5.9). For $1 \le n \le n_*$ we have:*

$$q_n(0) \;=\; -p_n'(0) \;=\; \sum_{k=1}^{n} t_{k,n}^{-1}, \tag{5.14}$$

$$q_n(t) \;\ge\; 0, \qquad 0 < t \le t_{1,n}, \tag{5.15}$$

$$q_n(0) \;=\; \sup_{0 \le t \le t_{1,n}} q_n(t). \tag{5.16}$$

If additionally the interlacing property (5.10) is satisfied, then

$$q_{n-1}(0) \;\le\; q_n(0), \tag{5.17}$$

$$q_n(0) \;\le\; t_{1,n}^{-1} \;+\; q_{n-1}(0). \tag{5.18}$$

Proof. The equalities in (5.14) follow from the representations (5.11) and (5.13). Moreover, (5.15) follows from (5.12) and (5.13). In order to prove (5.16) we observe that $p_n(t)$ is convex on $[0, t_{1,n}]$:

$$p_n'(t) \;=\; -\sum_{k=1}^{n} \frac{1}{t_{k,n}} \prod_{\substack{j=1 \\ j \ne k}}^{n} \Big(1 - \frac{t}{t_{j,n}}\Big),$$

$$p_n''(t) \;=\; \sum_{k} \frac{1}{t_{k,n}} \sum_{l \ne k} \frac{1}{t_{l,n}} \prod_{j \ne k,l} \Big(1 - \frac{t}{t_{j,n}}\Big).$$

Now (5.16) follows from (5.13). Moreover, (5.17) is fulfilled trivially for $n = 1$, and for $n \ge 2$ the intertwining property (5.10) yields

$$q_{n-1}(0) \;=\; \sum_{k=1}^{n-1} t_{k,n-1}^{-1} \;\le\; \sum_{k=1}^{n-1} t_{k,n}^{-1} \;\le\; \sum_{k=1}^{n} t_{k,n}^{-1} \;=\; q_n(0),$$

this is (5.17) for $n \ge 2$. Finally, the intertwining property (5.10) yields also (for $n \ge 2$; the case $n = 1$ in (5.18) is trivial)

$$q_n(0) \;=\; t_{1,n}^{-1} \;+\; \sum_{k=2}^{n} t_{k,n}^{-1} \;\le\; t_{1,n}^{-1} \;+\; \sum_{k=1}^{n-1} t_{k,n-1}^{-1} \;=\; t_{1,n}^{-1} \;+\; q_{n-1}(0),$$

this is (5.18). This completes the proof. □

We conclude this section with one more useful lemma.

LEMMA 5.2. *Let $\mu > 0$, and let n be an integer. For any polynomial $q_n(t) \in \Pi_{n-1}$ such that $p_n(t) = 1 - tq_n(t)$ has increasingly ordered positive roots $\{t_{k,n}\}_{1 \le k \le n}$, cf. (5.9), the following estimate is valid,*

$$\Phi(t) := p_n(t) t^{\mu} \;\le\; \Big(\mu q_n(0)^{-1}\Big)^{\mu}, \qquad 0 \le t \le t_{1,n}. \tag{5.19}$$

Proof. From the definition of Φ and the product representation (5.11) of $p_n(t)$ we get immediately $\Phi(0) = \Phi(t_{1,n}) = 0$, and $\Phi(t) > 0$ for $0 < t < t_{1,n}$. Moreover,

$$p_n'(t) \;=\; -p_n(t)\sum_{k=1}^{n}\frac{1}{t_{k,n}-t}.$$

Now let $0 < \bar{t} < t_{1,n}$ with

$$\Phi(\bar{t}) \;=\; \sup_{0\le t\le t_{1,n}} \Phi(t).$$

Hence $0 = \Phi'(\bar{t})$, and then (5.14) yields

$$\begin{aligned}\mu p_n(\bar{t})\bar{t}^{\mu-1} \;&=\; p_n(\bar{t})\bar{t}^{\mu}\sum_{k=1}^{n}\frac{1}{t_{k,n}-\bar{t}}\\ &\ge\; p_n(\bar{t})\bar{t}^{\mu} q_n(0),\end{aligned}$$

therefore $\bar{t} \le \mu q_n(0)^{-1}$, and thus

$$\sup_{0\le t\le t_{1,n}} \Phi(t) \;=\; \Phi(\bar{t}) \;=\; p_n(\bar{t})\bar{t}^{\mu} \;\le\; \bar{t}^{\mu} \;\le\; \Big(\mu q_n(0)^{-1}\Big)^{\mu},$$

and this completes the proof of (5.19). □

6. Subsidiary results and the proof of Theorem 3.3. The following diagram illustrates the relations between the several lemmas and corollaries presented in this section:

$$\text{Lemma 6.1} \Longrightarrow \left\{\begin{array}{lcc} \text{Lemma 6.2} & \Longrightarrow & \left\{\begin{array}{l}\text{Corollary 6.3}\\ \text{Corollary 6.4}\end{array}\right. \\ \left.\begin{array}{l}\text{Lemma 6.5}\\ \text{Lemma 6.6}\end{array}\right\} & \Longrightarrow & \text{Corollary 6.7} \end{array}\right\} \Longrightarrow \text{Proposition 6.8.}$$

Here, Corollaries 6.3 and 6.4 provide estimates for $\|u_* - u_n\|$ while Corollary 6.7 yields an estimate for the norm Δ_{n-1} of the residual. Finally, Corollary 6.7 and Proposition 6.8 then provide the tools for completing the proof of Theorem 3.3.

6.1. Preliminaries. In the sequel, all arising constants $c_1, c_2, \ldots$ are independent of δ, h and ϱ (introduced in Assumption 3.1), and they may depend on b (introduced in Definition 3.2), on a_ν, $0 \le \nu \le \nu_1$, (cf. (3.3)), and on $\|A\|$, if not further specified. We start with a lemma that can be applied also to other than conjugate gradient type methods; for a similar result we refer to Hämarik [10].

LEMMA 6.1. *Let parts 1. and 2. of Assumption 3.1 be fulfilled. Let $\psi : [0, a] \to \mathbb{R}$ be a bounded function, and let $\mu > 0$. Let $\mathcal{F}_\tau$, $\tau > 0$, be the resolution of the identity associated with A_h, cf. definition (5.2). Then*

$$\begin{aligned} &\|\mathcal{F}_\tau \psi(A_h)\mathcal{P}_h A^\mu\| \\ &\le \Big(\sup_{0\le t\le\tau}|\psi(t)|t^\mu\Big) + c_1\sum_{j=0}^{\lfloor\mu\rfloor}\Big(\sup_{0\le t\le\tau}|\psi(t)|t^j\Big)\xi_h^{\min\{\mu-j,\nu_1+1\}} && (6.1)\\ &\le c_2\Big(\sup_{0\le t\le\tau}|\psi(t)|\Big)\Big(\tau^\mu + \xi_h^{\min\{\mu,\nu_1+1\}}\Big), \qquad \tau > 0. && (6.2)\end{aligned}$$

Here, $\lfloor\mu\rfloor$ denotes the greatest integer $\le \mu$, and c_1 and c_2 are constants that are independent of h and ψ (and depend on μ).

Proof. We have the following decomposition (with $k = \lfloor\mu\rfloor$):

$$\begin{aligned}\mathcal{P}_h A^\mu &= \mathcal{P}_h\Big[A^\mu - A_h^k A^{\mu-k}\Big] + \mathcal{P}_h A_h^k (A^{\mu-k} - A_h^{\mu-k}) + \mathcal{P}_h A_h^\mu \\ &= \mathcal{P}_h\Big[\sum_{j=0}^{k-1} A_h^j A (I-\mathcal{P}_h) A^{\mu-j-1}\Big] + \mathcal{P}_h A_h^k (A^{\mu-k} - A_h^{\mu-k}) + \mathcal{P}_h A_h^\mu \qquad (6.3)\end{aligned}$$

(note that $\mathcal{P}_h A_h^j = \mathcal{P}_h A_h^j \mathcal{P}_h$ for $j \ge 0$). Then multiplying both sides of (6.3) with $\mathcal{F}_\tau \psi(A_h)$ yields (6.1), if we take into account that (5.4) as well as (3.4), (3.5) are valid, and if we moreover use the estimates

$$\|A^\gamma - A_h^\gamma\| \le 2\|A - A_h\|^\gamma \le 4 a_1^\gamma \xi_h^\gamma, \qquad 0 < \gamma \le 1, \qquad (6.4)$$

where the first of the estimates in (6.4) follows immediately from Lemma 1.1 in Chapter 4 of Vainikko & Veretennikov [30]. To obtain (6.2), we use the elementary estimates

$$\sup_{0\le t\le\tau} |\psi(t)| t^s \le \Big(\sup_{0\le t\le\tau} |\psi(t)|\Big)\tau^s, \qquad s > 0,$$

as well as (recall that we assume $\xi_h \le 1$)

$$\left\{\begin{array}{lcl} \tau^j \xi_h^{\mu-j} & \le & \xi_h^\mu + \tau^\mu, \\ \tau^j \xi_h^{\mu-j} & \le & \xi_h^{\nu_1+1} + \tau^\mu, \qquad \text{if } \mu \ge \nu_1 + 1, \\ \tau^j \xi_h^{\nu_1+1} & \le & \xi_h^{\nu_1+1} + \tau^\mu, \end{array}\right. \qquad (6.5)$$

(which are applied in (6.1) for the three situations (i) $\mu \le \nu_1+1$, (ii) $\mu \ge \nu_1+1$, $\mu - j \le \nu_1 + 1$ and (iii) $\mu \ge \nu_1 + 1$, $\mu - j \ge \nu_1 + 1$). This completes the proof. □

LEMMA 6.2. *Let Assumption 3.1 be valid. Let $\varphi : \mathbb{R}_+ \to \mathbb{R}$ be bounded, and let*

$$\Delta := \|\big(I - A_h\varphi(A_h)\big)\mathcal{P}_h f^\delta\|.$$

Then for any $\tau > 0$ we have

$$\begin{aligned}\|u_* - \varphi(A_h)\mathcal{P}_h f^\delta\| &\le \tau^{-1}\Big(\Delta + \eta(h,\delta)\Big) + \Big(\sup_{0\le t\le\tau} |\varphi(t)|\Big)\eta(h,\delta) \\ &\quad + c_3\varrho\Big(\sup_{0\le t\le\tau} |1 - t\varphi(t)|\Big)\Big(\tau^\nu + \xi_h^{\min\{\nu,\nu_1\}}\Big),\end{aligned}$$

where c_3 is a constant that is independent of δ, h, ϱ and φ (and depends on ν).

Proof. For $\tau > 0$ we have, with $\mathcal{F}_\tau$ as in (5.2):

$$\begin{aligned}&\|u_* - \varphi(A_h)\mathcal{P}_h f^\delta\| \\ &\overset{(5.5)}{\le} \|(I - \mathcal{F}_\tau)\Big(u_* - \varphi(A_h)\mathcal{P}_h f^\delta\Big)\| + \|\mathcal{F}_\tau\mathcal{P}_h\Big(u_* - \varphi(A_h)\mathcal{P}_h f^\delta\Big)\| \\ &\qquad + \|(I - \mathcal{P}_h)u_*\| \\ &\overset{(3.7)}{\le} \|(I - \mathcal{F}_\tau)\Big(u_* - \varphi(A_h)\mathcal{P}_h f^\delta\Big)\| + \|\mathcal{F}_\tau\mathcal{P}_h\Big(u_* - \varphi(A_h)\mathcal{P}_h f^\delta\Big)\| \qquad (6.6) \\ &\qquad + b_\nu \varrho \xi_h^{\min\{\nu,\nu_1\}}.\end{aligned}$$

In the sequel we shall estimate the first two terms on the right-hand side of (6.6). First, with $\mathcal{Q}_j$ as in Section 5.1 we have, cf. (5.3),

$$\begin{aligned}
\|(I-\mathcal{F}_\tau)\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|^2 &= \sum_{j:\ \lambda_j>\tau}\|\mathcal{Q}_j\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|^2\\
&\le \tau^{-2}\sum_{j:\ \lambda_j>\tau}\lambda_j^2\|\mathcal{Q}_j\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|^2\\
&\le \tau^{-2}\sum_{j=1}^{m}\lambda_j^2\|\mathcal{Q}_j\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|^2\\
&= \tau^{-2}\|A_h\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|^2\\
&\le \tau^{-2}\Big(\|A_hu_*-\mathcal{P}_hf^\delta\| \ + \ \Delta\Big)^2\\
&\overset{(3.8)}{\le} \tau^{-2}\Big(\eta(h,\delta)+\Delta\Big)^2.
\end{aligned}$$

Second, by assumption we have the representation, $u_*=A^\nu z,\ \|z\|=\varrho$, cf. (3.6), thus

$$u_*-\varphi(A_h)\mathcal{P}_hf^\delta \ = \ \Big(I-\varphi(A_h)A_h\Big)A^\nu z \ + \ \varphi(A_h)\Big(A_hu_*-\mathcal{P}_hf^\delta\Big), \tag{6.7}$$

and then Lemma 6.1 for $\mu=\nu$ and the properties (5.4), (5.5) yield

$$\begin{aligned}
&\|\mathcal{F}_\tau\mathcal{P}_h\Big(u_*-\varphi(A_h)\mathcal{P}_hf^\delta\Big)\|\\
&\le \|\mathcal{F}_\tau\Big(I-\varphi(A_h)A_h\Big)\mathcal{P}_hA^\nu z\| \ + \ \|\mathcal{F}_\tau\varphi(A_h)\Big(A_hu_*-\mathcal{P}_hf^\delta\Big)\|\\
&\le \|\mathcal{F}_\tau\Big(I-\varphi(A_h)A_h\Big)\mathcal{P}_hA^\nu\|\varrho \ + \ \|\mathcal{F}_\tau\varphi(A_h)\|\eta(h,\delta)\\
&\le c_2\varrho\sup_{0\le t\le\tau}\Big(|1-t\varphi(t)|\Big)\big(\tau^\nu \ + \ \xi_h^{\min\{\nu,\nu_1\}}\big) \ + \ \Big(\sup_{0\le t\le\tau}|\varphi(t)|\Big)\eta(h,\delta)
\end{aligned}$$

(recall that we assume $\xi_h\le 1$). This completes the proof. ∎

In a first corollary we state a result which is useful if the stopping rule given in Definition 3.2 leads to an immediate termination of the method of conjugate residuals, i.e., if we have $\overline{n}=0$.

COROLLARY 6.3. *Let Assumption 3.1 be valid, and suppose that*

$$\|\mathcal{P}_hf^\delta\| \ \le \ C\Big(\delta \ + \ \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\Big) \tag{6.8}$$

holds for some constant $C>0$. *Then*

$$\|u_*\| \ \le \ c_4\Big((\varrho\delta^\nu)^{1/(\nu+1)} \ + \ \varrho\xi_h^{\min\{\nu,\nu_1\}}\Big),$$

where c_4 *is a constant that is independent of* δ, h *and* ϱ *(and depends on* C *and* ν*).*

Proof. It follows immediately from Lemma 6.2, with $\varphi=0$, as well as from (6.8) that for $\tau>0$ we have the following estimate,

$$\|u_*\| \ \le \ \Big(C+\max\{1,a_1b_\nu\}\Big)\tau^{-1}\Big(\delta+\varrho\xi_h^{\min\{\nu,\nu_1\}+1}\Big) \ + \ c_3\varrho\Big(\tau^\nu+\xi_h^{\min\{\nu,\nu_1\}}\Big). \tag{6.9}$$

Without loss of generality we may suppose that $\varrho \neq 0$, and then we shall estimate the right-hand side of (6.9) for the specific choice

$$\tau \;=\; \Big(\frac{\delta}{\varrho}\Big)^{1/(\nu+1)} \;+\; \xi_h.$$

In fact,

$$\begin{aligned} \tau^{-1}\delta \;&\le\; \Big(\frac{\delta}{\varrho}\Big)^{-1/(\nu+1)}\delta \;=\; \big(\varrho\delta^\nu\big)^{1/(\nu+1)}, \\ \tau^{-1}\xi_h \;&\le\; 1, \\ \varrho\tau^\nu \;&\le\; \varrho 2^\nu\Big[\Big(\frac{\delta}{\varrho}\Big)^{\nu/(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}}\Big] \;=\; 2^\nu\Big(\big(\varrho\delta^\nu\big)^{1/(\nu+1)} \;+\; \varrho\xi_h^{\min\{\nu,\nu_1\}}\Big) \end{aligned}$$

(recall that we assume $\xi_h \le 1$). This completes the proof. □

6.2. Subsidiary results for the method of conjugate residuals. Now we return to the method of conjugate residuals. From Lemma 6.2 we get the following corollary.

Corollary 6.4. *Let Assumption 3.1 be valid. Let $\{u_n\}_{0\le n\le n_*}$ be as in Definition 2.1, and let $\{q_n(t)\}_n$ and $\{\Delta_n\}_n$ be as in (3.9)–(3.10) and (3.11), respectively. For any $1 \le n \le n_*$ we have*

$$\|u_* - u_n\| \;\le\; \tau^{-1}\Big(\Delta_n + \eta(h,\delta)\Big) + q_n(0)\eta(h,\delta) + c_3\varrho\Big(\tau^\nu + \xi_h^{\min\{\nu,\nu_1\}}\Big), \quad 0 < \tau \le t_{1,n},$$

with constant c_3 as in Lemma 6.2.

Proof. We shall apply Lemma 6.2 with $\varphi(t) = q_n(t)$. In fact, we have

$$0 \;\le\; p_n(t) \;=\; 1 - tq_n(t) \;\le\; 1, \qquad 0 < t \le t_{1,n},$$

cf. (5.6), (5.12), and from

$$0 \;\le\; q_n(t) \;\le\; q_n(0), \qquad 0 < t \le t_{1,n},$$

cf. (5.15), (5.16), the assertion then follows immediately. □

We have obtained a first estimate for the error $\|u_* - u_n\|$. The next two lemmas provide reasonable estimates for the norm Δ_n of the residual.

Lemma 6.5. *Let Assumption 3.1 be valid. Let $\{u_n\}_{0\le n\le n_*}$ be as in Definition 2.1, and let $\{q_n(t)\}_n$ and $\{\Delta_n\}_n$ be as in (3.9)–(3.10) and (3.11), respectively. Then for any $1 \le n \le n_*$ we have*

$$\Delta_n \;\le\; \delta \;+\; c_5\varrho\Big(q_n(0)^{-(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}+1}\Big), \tag{6.10}$$

where c_5 denotes a constant that is independent of δ, h, ϱ and n $\big($and depends on $\nu\big)$.

Proof. We decompose and estimate as follows,

$$\begin{aligned} \Delta_n^2 \;&=\; \sum_{j=0}^m p_n^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \\ &=\; \sum_{j:\ \lambda_j\le t_{1,n}} p_n^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \;+\; \sum_{j:\ \lambda_j > t_{1,n}} p_n^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \end{aligned}$$

$$\leq \sum_{j:\ \lambda_j \leq t_{1,n}} p_n^2(\lambda_j) \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2$$

$$+ \ t_{1,n}^{-1} \sum_{j:\ \lambda_j > t_{1,n}} \lambda_j p_n^2(\lambda_j) \Big(\frac{\lambda_j}{t_{1,n}} - 1 \Big)^{-1} \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2, \tag{6.11}$$

where the latter estimate follows from $\lambda(\frac{\lambda}{t_{1,n}} - 1)^{-1} \geq t_{1,n}$ for $\lambda > t_{1,n}$. In order to estimate (6.11), we observe that the conjugacy property (5.8) yields

$$\sum_{j:\lambda_j > t_{1,n}} \lambda_j p_n^2(\lambda_j) \Big(\frac{\lambda_j}{t_{1,n}} - 1 \Big)^{-1} \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2 = \sum_{j:\lambda_j \leq t_{1,n}} \lambda_j p_n^2(\lambda_j) \Big(1 - \frac{\lambda_j}{t_{1,n}} \Big)^{-1} \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2,$$

and this applied in (6.11) yields, cf. also (5.2),

$$\begin{aligned}
\Delta_n^2 \ &\leq \sum_{j:\ \lambda_j \leq t_{1,n}} p_n^2(\lambda_j) \Big[1 + \frac{\lambda_j}{t_{1,n}} \Big(1 - \frac{\lambda_j}{t_{1,n}} \Big)^{-1} \Big] \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2 \\
&= \sum_{j:\ \lambda_j \leq t_{1,n}} p_n^2(\lambda_j) \Big(1 - \frac{\lambda_j}{t_{1,n}} \Big)^{-1} \|\mathcal{Q}_j \mathcal{P}_h f^\delta\|^2 \\
&= \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h f^\delta\|^2,
\end{aligned} \tag{6.12}$$

with (cf. (5.11))

$$v_n(t) \ = \ \Big(1 - \frac{t}{t_{1,n}} \Big) \Big[\prod_{k=2}^{n} \Big(1 - \frac{t}{t_{k,n}} \Big)^2 \Big], \qquad 0 \leq t \leq t_{1,n},$$

and $v_n(t) := 0$ for $t > t_{1,n}$. Note that $0 \leq v_n(t) \leq 1$ for all $0 \leq t \leq t_{1,n}$. We proceed with an estimation of (6.12): from Assumption 3.1 and estimate (5.4) we obtain

$$\begin{aligned}
\Delta_n \ &\leq \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h f^\delta\| \\
&\leq \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h A^{\nu+1} z\| \ + \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h (A u_* - f^\delta)\| \\
&\leq \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h A^{\nu+1}\| \varrho \ + \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h)\| \delta \\
&\leq \ \|\mathcal{F}_{t_{1,n}} v_n^{1/2}(A_h) \mathcal{P}_h A^{\nu+1}\| \varrho \ + \ \Big(\sup_{0 \leq t \leq t_{1,n}} v_n(t) \Big)^{1/2} \delta \\
&\leq \ \|\mathcal{F}_{t_{1,n}} \sqrt{|p_n|}(A_h) \mathcal{P}_h A^{\nu+1}\| \varrho \ + \ \delta,
\end{aligned} \tag{6.13}$$

where it has been taken into account that $0 \leq 1 - \frac{t}{t_{k,n}} \leq 1$ holds for $0 \leq t \leq t_{1,n},\ k = 2, \ldots, n$. In order to provide further estimations of the right-hand side in (6.13), we next shall apply Lemma 6.1 with $\mu = \nu + 1$; in fact, from the estimates (5.19) and (6.1) we get

$$\Delta_n \ \leq \ \Big[\Big(2(\nu+1) q_n(0)^{-1} \Big)^{\nu+1} \ + \ c_1 \sum_{j=0}^{\lfloor \nu+1 \rfloor} \big(2j q_n(0)^{-1} \big)^j \xi_h^{\min\{\nu-j, \nu_1\}+1} \Big] \varrho \ + \ \delta,$$

and proceeding then as in (6.5), with τ replaced by $q_n(0)^{-1}$, yields the assertion (6.10). This completes the proof. ☐

Lemma 6.6. *Let Assumption 3.1 be valid. Let $\{u_n\}_{0 \leq n \leq n_*}$ be as in Definition 2.1, and let $\{q_n(t)\}_n$ and $\{\Delta_n\}_n$ be as in (3.9)–(3.10) and (3.11), respectively. Fix $1 \leq n \leq n_*$, $\theta > 2$ and $2 < \kappa \leq 2(\theta - 1)$. If*

$$\theta q_{n-1}(0) \ \leq \ q_n(0), \tag{6.14}$$

then we have

$$\frac{\kappa-2}{\kappa-1}\Delta_{n-1} \;\le\; \delta \;+\; c_6\varrho\Big(q_n(0)^{-(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}+1}\Big),$$

where c_6 is a constant that is independent of δ, h, ϱ and n $\big($and depends on θ and $\kappa\big)$.

Proof. Let

$$s(t) := p_n(t)\Big(1-\frac{t}{t_{1,n}}\Big)^{-1} \;=\; \prod_{k=2}^{n}\Big(1-\frac{t}{t_{k,n}}\Big), \tag{6.15}$$

where $p_n(t)$ is as in (5.6), (5.11). We have $s\in\Pi_{n-1}$ and $s(0)=1$, and the minimum property (5.7) then yields

$$\begin{aligned}
\Delta_{n-1}^2 \;&\le\; \|s(A_h)\mathcal{P}_h f^\delta\|^2\\
&=\; \sum_{j:\ \lambda_j\le\kappa t_{1,n}} s^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \;+\; \sum_{j:\ \lambda_j>\kappa t_{1,n}} s^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2\\
&\le\; \sum_{j:\ \lambda_j\le\kappa t_{1,n}} s^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \;+\; (\kappa-1)^{-2}\sum_{j:\ \lambda_j>\kappa t_{1,n}} p_n^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \qquad (6.16)\\
&\le\; \sum_{j:\ \lambda_j\le\kappa t_{1,n}} s^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2 \;+\; (\kappa-1)^{-2}\sum_{j=0}^{m} p_n^2(\lambda_j)\|\mathcal{Q}_j\mathcal{P}_h f^\delta\|^2\\
&=\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h f^\delta\|^2 \;+\; (\kappa-1)^{-2}\Delta_n^2\\
&\le\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h f^\delta\|^2 \;+\; (\kappa-1)^{-2}\Delta_{n-1}^2, \qquad (6.17)
\end{aligned}$$

where (6.16) is valid since we have $(1-\frac{t}{t_{1,n}})^{-2}\le(\kappa-1)^{-2}$ for $t\ge\kappa t_{1,n}$; moreover, (6.17) follows from (3.12). A reformulation and further estimation of the right-hand side in (6.17) then yields

$$\begin{aligned}
\frac{\kappa-2}{\kappa-1}\Delta_{n-1} \;&\le\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h f^\delta\|\\
&\le\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h A^{\nu+1}z\| \;+\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h(Au_*-f^\delta)\|\\
&\le\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\mathcal{P}_h A^{\nu+1}\|\varrho \;+\; \|\mathcal{F}_{\kappa t_{1,n}} s(A_h)\|\delta. \qquad (6.18)
\end{aligned}$$

Below we show that the estimates

$$\kappa t_{1,n} \;\le\; 2t_{2,n}, \qquad \text{if } n\ge 2, \tag{6.19}$$

$$t_{1,n} \;\le\; (1-\theta^{-1})^{-1}q_n(0)^{-1}, \qquad \text{if } n\ge 1, \tag{6.20}$$

are valid. Then, due to

$$\begin{aligned}
|s(t)| \;&\le\; 1, \qquad 0\le t\le 2t_{2,n}, \qquad \text{if } n\ge 2, \qquad \big(\text{cf. (6.15)}\big)\\
s(t) \;&=\; 1, \qquad t\in\mathbb{R}, \qquad \text{if } n=1,
\end{aligned}$$

a further estimation of (6.18) yields, cf. (5.4), (6.2),

$$\begin{aligned}
\frac{\kappa-2}{\kappa-1}\Delta_{n-1} \;&\le\; \delta \;+\; c_2\varrho\Big(\big(\kappa t_{1,n}\big)^{\nu+1} \;+\; \xi_h^{\min\{\nu,\nu_1\}+1}\Big)\\
&\le\; \delta \;+\; c_2\varrho\Big(\big(\beta q_n(0)\big)^{-(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}+1}\Big),
\end{aligned}$$

with $\beta := (1-\theta^{-1})/\kappa$, and with constant c_2 as in Lemma 6.1.

It remains to show that the estimates (6.19), (6.20) are valid. For this purpose we recall that the assumption (6.14) and the estimate (5.18) yield

$$\theta q_{n-1}(0) \ \le \ q_n(0) \ \le \ t_{1,n}^{-1} \ + \ q_{n-1}(0), \tag{6.21}$$

and this implies immediately (6.20). We next show that (6.19) is valid: in fact, from the estimate (6.21) we obtain

$$(\theta-1)t_{1,n-1}^{-1} \overset{(5.14)}{\le} (\theta-1)q_{n-1}(0) \ \le \ t_{1,n}^{-1},$$

and thus

$$\kappa t_{1,n} \ \le \ \frac{\kappa t_{1,n-1}}{\theta-1} \ \le \ 2t_{1,n-1},$$

and finally $t_{1,n-1} \le t_{2,n}$, cf. (5.10), then yields (6.19). This completes the proof. □

Remarks. The proofs of Lemmas 6.5 and 6.6 follow the ideas of the paper by Nemirovskiĭ [21] until estimates (6.13) and (6.18), respectively. Beyond that we make use of Lemma 6.1 which is applicable in our specific situation and leads to better results than in [21]; cf. also part (5) in the conclusions in Section 3.3.1. △

From Lemmas 6.5 and 6.6 we get:

COROLLARY 6.7. *Let Assumption 3.1 be valid. Let $\{u_n\}_{0\le n\le n_*}$ be as in Definition 2.1, and let $\{q_n(t)\}_n$ and $\{\Delta_n\}_n$ be as in (3.9)–(3.10) and (3.11), respectively. For all $0<\gamma<1$ we have*

$$\gamma\Delta_{n-1} \ \le \ \delta \ + \ c_7\varrho\Big(q_n(0)^{-(\nu+1)} \ + \ \xi_h^{\min\{\nu,\nu_1\}+1}\Big), \qquad 1\le n\le n_*,$$

where c_7 is a constant that is independent of δ, h, ϱ and n (and depends on γ and ν).

Proof. Let $\kappa>2$ such that $\gamma=(\kappa-2)/(\kappa-1)$ is valid. Moreover let $\theta>2$ such that $\kappa=2(\theta-1)$ holds. From Lemmas 6.5 and 6.6 we obtain the desired result by considering the two different cases "$\theta q_{n-1}(0)\le q_n(0)$" and "$\theta q_{n-1}(0)>q_n(0)$". Note that in the latter case we necessarily have $n\ge 2$. □

Corollaries 6.4 and 6.7 enable us to prove the following lemma which is presented in a general form so that it may be applied for the cases $n=\overline{n}$ and $n=\overline{n}-1$, cf. the specific situations described in Subsection 6.3, respectively.

PROPOSITION 6.8. *Let Assumption 3.1 be valid, and let $b>1$. Let $\{u_n\}_{0\le n\le n_*}$ be as in Definition 2.1, and let $\{q_n(t)\}_n$ and $\{\Delta_n\}_n$ be as in (3.9)–(3.10) and (3.11), respectively. Suppose that for some fixed n with $1\le n\le n_*$ we have*

$$q_n(0) \ \le \ \xi_h^{-1}, \qquad b\delta \ \le \ \Delta_{n-1}, \qquad \Delta_n \ \le \ C\Big(\delta \ + \ \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\Big), \tag{6.22}$$

where C denotes some constant. Then

$$\|u_*-u_n\| \ \le \ c_8\Big(\big(\varrho\delta^\nu\big)^{1/(\nu+1)} \ + \ \varrho\xi_h^{\min\{\nu,\nu_1\}}\Big),$$

where c_8 is a constant that is independent of δ, h, ϱ and n (and depends on C and ν).

Proof. Corollary 6.4 and the assumption on Δ_n, cf. (6.22), yield for $0<\tau\le t_{1,n}$

$$\|u_*-u_n\| \ \le \ C_1\Big(\big[\tau^{-1}+q_n(0)\big]\Big(\delta+\varrho\xi_h^{\min\{\nu,\nu_1\}+1}\Big) \ + \ \varrho\Big(\tau^\nu+\xi_h^{\min\{\nu,\nu_1\}}\Big)\Big), \tag{6.23}$$

with a constant C_1, and we shall estimate the right-hand side of (6.23). For this purpose let

$$\begin{aligned} \varepsilon \;&=\; \Big(\frac{\delta}{\varrho}\Big)^{1/(\nu+1)} \;+\; \xi_h, \\ \tau \;&=\; \min\Big\{\varepsilon,\; q_n(0)^{-1}\Big\}. \end{aligned} \tag{6.24}$$

Then we have, cf. $\big(5.14\big)$, $\tau \le q_n(0)^{-1} \le t_{1,n}$, hence (6.23) is valid for this specific choice of τ. Now the different terms on the right-hand side of (6.23) remain to be estimated.

(a) Similar as in the proof of Lemma 6.3 we obtain the following elementary estimates $\big($recall that $\xi_h \le 1\big)$,

$$\varrho\tau^\nu \;\le\; \varrho\varepsilon^\nu \;\le\; 2^\nu\Big(\big(\varrho\delta^\nu\big)^{1/(\nu+1)} \;+\; \varrho\xi_h^{\min\{\nu,\nu_1\}}\Big).$$

(b) Since by assumption $\xi_h \le \varepsilon$ and $\xi_h \le q_n(0)^{-1}$, cf. (6.22) and (6.24), respectively, we have $\xi_h \le \tau$ and thus $\tau^{-1}\xi_h \le 1$.

(c) We still have to estimate $\tau^{-1}\delta$ sufficiently good, and for this purpose we consider two different cases. First, if $q_n(0) \le \varepsilon^{-1}$, then

$$\tau \;=\; \varepsilon \;\ge\; \Big(\frac{\delta}{\varrho}\Big)^{1/(\nu+1)},$$

and thus

$$\tau^{-1}\delta \;\le\; \big(\varrho\delta^\nu\big)^{1/(\nu+1)},$$

which provides a sufficiently good estimate. Now suppose that

$$q_n(0) > \varepsilon^{-1}. \tag{6.25}$$

Then Corollary 6.7 and the assumption on Δ_{n-1} in (6.22) yields

$$b\gamma\delta \;<\; \gamma\Delta_{n-1} \;\le\; \delta \;+\; c_7\varrho\Big(q_n(0)^{-(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}+1}\Big).$$

Here, $0 < \gamma < 1$ is chosen such that $b\gamma > 1$, and then c_7 is chosen according to Corollary 6.7. Subtracting δ on both sides of the last inequality and multiplying then both sides with τ^{-1} yields

$$\begin{aligned} (b\gamma - 1)\tau^{-1}\delta \;&\le\; c_7\varrho\Big(q_n(0)^{-\nu} \;+\; \xi_h^{\min\{\nu,\nu_1\}}\Big) \\ &\le\; c_7\big(2^\nu + 1\big)\varrho\Big[\Big(\frac{\delta}{\varrho}\Big)^{1/(\nu+1)} \;+\; \xi_h^{\min\{\nu,\nu_1\}}\Big], \end{aligned}$$

where the assumption $\tau = q_n(0)^{-1}$, cf. (6.25), as well as the estimates in parts (a) and (b) of this proof have been applied. Thus $\tau^{-1}\delta$ is estimated sufficiently good in this case.

(d) Finally we shall estimate $q_n(0)\big(\delta + \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\big)$. In fact, $q_n(0) \le \xi_h^{-1}$ holds by assumption, and moreover we have $q_n(0) \le \tau^{-1}$, thus with the result (c) of this proof we can estimate $q_n(0)\delta$ sufficiently good. This completes the proof. ∎

6.3. The proof of Theorem 3.3. As a preparation we remark that we have by definition $A_h r_{n_*} = 0$, and thus

$$\Delta_{n_*} \;=\; \|r_{n_*}\| \;=\; \operatorname{dist}\big(\mathcal{P}_h f^\delta, \mathcal{R}(A_h)\big) \;\le\; \eta(h,\delta). \tag{6.26}$$

Now we are in a position to prove the main result.

(1) We first suppose that

$$\text{“}\overline{n} = 0\text{”} \qquad \text{or} \qquad \text{“}\overline{n} = 1, \quad q_1(0) \ge \xi_h^{-1}\text{”}, \tag{6.27}$$

any of these two cases leading to $\overline{u}(h,\delta) = 0$. We show below that in both cases we have

$$\|\mathcal{P}_h f^\delta\| \;\le\; C\Big(\delta \;+\; \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\Big) \tag{6.28}$$

for an appropriate constant $C > 0$, which together with Corollary 6.3 yields an adequate estimate of $\|u_* - \overline{u}(h,\delta)\| \;=\; \|u_*\|$ for these cases in (6.27). We start with the proof of (6.28): first, $\overline{n} = 0$ means $A_h r_0 = 0$ $\big($and (6.26) then yields $\Delta_0 = \|\mathcal{P}_h f^\delta\| \le \eta(h,\delta)\big)$ or $\Delta_0 = \|\mathcal{P}_h f^\delta\| \le b\delta$, and thus we have $\|\mathcal{P}_h f^\delta\| \le \max\{b, a_1 b_\nu\}\big(\delta + \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\big)$, if $\overline{n} = 0$. Moreover, if $\overline{n} = 1$ and $q_1(0) \ge \xi_h^{-1}$, then Corollary 6.7 yields $\big($recall that $\xi_h \le 1\big)$

$$\gamma\|\mathcal{P}_h f^\delta\| \;=\; \gamma\Delta_0 \;\le\; \delta \;+\; 2c_7\varrho\xi_h^{\min\{\nu,\nu_1\}+1},$$

where $0 < \gamma < 1$ is arbitrary, and c_7 is chosen as in Corollary 6.7. This completes the proof for the cases in (6.27).

(2) We next consider the remaining cases, i.e., $\overline{n} \ge 1$ and $q_1(0) < \xi_h^{-1}$.

(a) First, if $q_{\overline{n}}(0) < \xi_h^{-1}$, then we have $\Delta_{\overline{n}} \le b\delta$ or $A_h r_{\overline{n}} = 0$, in any case we have $\Delta_{\overline{n}} \le \max\{b, a_1 b_\nu\}\big(\delta + \varrho\xi_h^{\min\{\nu,\nu_1\}+1}\big)$, and then a sufficiently good estimate for $\|u_* - \overline{u}(h,\delta)\|$ follows from Proposition 6.8 for $n = \overline{n}$, since also $b\delta < \Delta_{\overline{n}-1}$ holds.

(b) Next suppose $q_{\overline{n}}(0) \ge \xi_h^{-1}$, and then $\overline{n} \ge 2$ and $\overline{u}(h,\delta) = u_{\overline{n}-1}$ holds. Corollary 6.7 then provides the estimate

$$\gamma\Delta_{\overline{n}-1} \;\le\; \delta \;+\; 2c_7\varrho\xi_h^{\min\{\nu,\nu_1\}+1},$$

where $0 < \gamma < 1$ is arbitrary, and c_7 is chosen as in Corollary 6.7. Since moreover the estimates $b\delta < \Delta_{\overline{n}-2}$ and $q_{\overline{n}-2}(0) < \xi_h^{-1}$ are valid, the assertion follows from Proposition 6.8 applied with $n = \overline{n} - 1$. This completes the proof of Theorem 3.3. □

REFERENCES

[1] S. Ashby, T. Manteuffel, and P. Saylor, *A taxomony for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[2] H. Brakhage, *On ill-posed problems and the method of conjugate gradients*, in Inverse and Ill-Posed Problems, Proc. St. Wolfgang 1986, H. Engl and C. Groetsch, eds., Boston, 1987, Academic Press, pp. 165–175.

[3] B. Eicke, A. Louis, and R. Plato, *The instability of some gradient methods for ill-posed problems*, Numer. Math., 58 (1990), pp. 129–134.

[4] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1st ed., 1996.

[5] H. Fattorini, *The Cauchy Problem*, Addison-Wesley, Reading, 1st ed., 1983.

[6] R. Freund, G. Golub, and N. Nachtigal, *Iterative solution of linear systems*, Acta Numerica, (1991), pp. 1–44.

[7] S. Gilyazov, *Approximate Solution of Ill-posed Problems (in Russian)*, Publisher Moscow University, Moscow, 1st ed., 1995.
[8] ———, *Regularizing conjugate-direction methods*, Comput. Maths. Math. Phys., 35 (1995), pp. 385–394.
[9] W. Hackbusch, *Integral Equations*, Birkhäuser, Basel, 1st ed., 1995.
[10] U. Hämarik, *On a parameter choice in the regularized Ritz-Galerkin method*, Proc. Estonian Acad. Sci. Phys. Math., 42 (1993), pp. 133–143.
[11] M. Hanke, *Conjugate Gradient Type Methods for Ill-Posed Problems*, Longman House, Harlow, 1995.
[12] ———, *The minimal error method is a regularization method*, Proc. Am. Math. Soc., 123 (1995), pp. 3487–3497.
[13] ———, *Asymptotics of orthogonal polynomials and the numerical solution of ill-posed problems*, Numer. Algor., 1-4 (1996), pp. 203–214.
[14] W. Kammerer and M. Nashed, *On the convergence of the conjugate gradient method for singular linear operator equations*, SIAM J. Numer. Anal., 9 (1972), pp. 165–181.
[15] J. King, *A minimal error conjugate gradient method for ill-posed problems*, J. Optim. Theory Appl., 60 (1989), pp. 297–304.
[16] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1st ed., 1996.
[17] R. Kress, *Linear Integral Equations*, Springer-Verlag, New York, 1st ed., 1989.
[18] L. Lardy, *A class of iterative methods of conjugate gradient type*, Numer. Funct. Anal. Optim., 11 (1990), pp. 283–302.
[19] A. Louis, *Convergence of the conjugate gradient method for compact operators*, in Inverse and Ill-Posed Problems, Proc. St. Wolfgang 1986, H. Engl and C. Groetsch, eds., Boston, 1987, Academic Press, pp. 177–183.
[20] S. Mikhlin, *Mathematical Physics, an advanced course*, North-Holland, Amsterdam, 1st ed., 1970.
[21] A. Nemirovskiĭ, *The regularizing properties of the adjoint gradient method in ill-posed problems*, U.S.S.R. Comput. Math. Math. Phys., 26 (1986), pp. 7–16.
[22] A. Nemirovskiĭ and B. Polyak, *Iterative methods for solving linear ill-posed problems under precise information I*, Moscow Univ. Comput. Math. Cybern., 22 (1984), pp. 1–11.
[23] ———, *Iterative methods for solving linear ill-posed problems under precise information II*, Moscow Univ. Comput. Math. Cybern., 22 (1984), pp. 50–56.
[24] R. Plato, *Optimal algorithms for linear ill-posed problems yield regularization methods*, Numer. Funct. Anal. Optim., 11 (1990), pp. 111–118.
[25] ———, *Über die Diskretisierung und Regularisierung schlecht gestellter Probleme*, PhD thesis, Technische Universität Berlin, Berlin, 1990.
[26] ———, *Die Effizienz des Diskrepanzprinzips für Verfahren vom Typ der konjugierten Gradienten*, in Festschrift H. Brakhage, E. Schock, ed., Aachen, 1994, Verlag Shaker, pp. 288–297.
[27] J. Stoer, *Solution of large systems of equations by conjugate gradient type methods*, in Mathematical Programming. The State of the Art. Bonn 1982, M. Grötschel, A. Bachem and B. Korte, eds., New York, 1983, Springer-Verlag, pp. 540–565.
[28] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1th ed., 1980.
[29] G. Vainikko, *Periodic Integral and Pseudo-Differential Operators*, Research Reports C13, Helsinki University of Technology, 1996.
[30] G. Vainikko and A. Veretennikov, *Iteration Procedures in Ill-Posed Problems (in Russian)*, Nauka, Moscow, 1st ed., 1986.