János Karátson

Mesh independent superlinear convergence estimates of the conjugate gradient method for some equivalent self-adjoint operators

Applications of Mathematics, Vol. 50 (2005), No. 3, 277--290

Persistent URL: `http://dml.cz/dmlcz/134606`

MESH INDEPENDENT SUPERLINEAR CONVERGENCE ESTIMATES OF THE CONJUGATE GRADIENT METHOD FOR SOME EQUIVALENT SELF-ADJOINT OPERATORS*

János Karátson, Budapest

Abstract. A mesh independent bound is given for the superlinear convergence of the CGM for preconditioned self-adjoint linear elliptic problems using suitable equivalent operators. The results rely on K-condition numbers and related estimates for compact Hilbert-Schmidt operators in Hilbert space.

Keywords: conjugate gradient method, superlinear convergence, mesh independence, preconditioning operator

MSC 2000: : 65F10, 65N22

1. Introduction

The conjugate gradient method has become the most widespread way of solving large symmetric positive definite (SPD) linear algebraic systems

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1}$$

in the past fifty years. One of its the most important features is superlinear convergence, first proved in [15] (see also [7], [13], [27], [28]). A comprehensive summary on the convergence of the CGM, with much of the results covering nonsymmetric systems as well, is given in the book [1].

Since its first presentation in [16], the convergence properties of the CGM have been deeply understood, sometimes involving Hilbert space theory. The survey given in [1] includes a characterization of the three typical phases of convergence of the CGM (namely, sublinear, linear and superlinear), see also [2]. While the

* This research was supported by the Hungarian National Research Fund OTKA under grant No. T043765.

well-known linear estimate uses the standard condition number, the superlinear estimate of [2] is based on the K-condition number of $\mathbf{A}$. The superlinear convergence of the CGM can be further understood from the related Hilbert space result on compact perturbations of the identity operator (i.e. for operators of the form $A = I + C$ with C compact) proved in [28]. Here the convergence is estimated using the sequence of eigenvalues of the compact operator (see also [27]) and the author's paper [3] where the K-condition numbers are related to the eigenvalues of the compact operator).

However, when the CGM is applied to the discretizations of elliptic problems, the convergence estimates deteriorate under refinement, i.e., the number of iterations for prescribed accuracy tends to ∞ when the mesh parameter h tends to 0. In the case of the linear estimate the standard condition numbers κ_h satisfy $\kappa_h = O(h^{-2})$ (see [1]). This suggests that mesh independent convergence can only be expected if suitable preconditioning is applied as a remedy. For the linear convergence estimate, the problem of mesh independence has been rigorously described in the framework of equivalent operators even for nonsymmetric problems [11]. Using the discretizations of suitable linear preconditioning operators as preconditioning matrices, the linear convergence estimate is mesh independent. (Moreover, the equivalence of the two operators is essentially necessary for such a result.) We note that this idea has wide applications for both linear and nonlinear problems, see e.g. [19], [20], [22] and was considered by the author in [12], [18]. Moreover, it has been shown in [11] that the PCG using equivalent operators is competitive with multigrid methods. A central area of the application of equivalent operators is the case when an original elliptic operator with variable (i.e. nonconstant) coefficients is preconditioned by an elliptic operator with constant coefficients, since the latter is cheaper to solve [6], [8], [10], [21], [26].

The above-mentioned mesh independence results have been extended to the case of superlinear convergence for some nonsymmetric elliptic problems in [4]. Namely, mesh independent superlinear convergence of the CGM has been proved for the symmetric part preconditioner $S = \frac{1}{2}(L + L^*)$ for general elliptic operators L, and for constant coefficient preconditioners when the original operator L has also (other) constant coefficients. These cases leave a gap, i.e., the result is still unproved in general for an elliptic operator with variable coefficients and a preconditioner with constant coefficients.

Our purpose is to take a step towards filling this gap: in this paper the mesh independence result is proved when the original (variable coefficient) operator is also symmetric (self-adjoint). For nonsymmetric operators, this result can be used in the framework of outer-inner iterations: that is, if the symmetric part preconditioner is used in an outer iteration cycle and in each step a constant coefficient preconditioning operator is used in an inner iteration cycle, then (by joining the results of [4] and the present paper) the superlinear convergence is mesh independent in both cycles.

Some further additions to the results of [4] are given in this paper by exploiting the self-adjoint setting. Namely, on a general bounded domain the superlinear convergence estimate now has the explicit form $\|e_n\|^{1/n} \leqslant c/\sqrt{n}$, moreover, the (mesh independent) constant $c > 0$ has a computable bound. (For the nonsymmetric case this was proved for the special case when the original operator also has constant coefficients and the domain is the unit square.) We also note that, owing to the symmetry, we can now rely on the theory of Hilbert-Schmidt operators, and the corresponding estimates for the discretized operators involve K-condition numbers and their relation to Frobenius norms.

The paper is organized as follows. In Section 2 some background is given on K-condition numbers. In Section 3 we consider preconditioned operator equations in Hilbert space in which the operator is a compact Hilbert-Schmidt perturbation of the identity, and prove an independence result with respect to the discretization subspaces. Using the latter, in Section 4 the mesh independent estimate is derived for elliptic boundary value problems and their finite discretizations. The result is verified when the operator is preconditioned by its principal part, and a computable estimate of the arising constant is given here. Then examples illustrate that the scope of the main result includes some known efficient preconditioning methods.

2. Preliminaries

The superlinear estimate of the CGM in [1], [2] is based on the K-condition number

$$K(\mathbf{A}) = \Big(\frac{1}{k}\,\mathrm{trace}\,(\mathbf{A})\Big)^k \big/ \det(\mathbf{A}) = \Big(\frac{1}{k}\sum_{i=1}^{k}\lambda_i(\mathbf{A})\Big)^k \Big(\prod_{i=1}^{k}\lambda_i(\mathbf{A})\Big)^{-1}, \tag{2}$$

where $\lambda_i(\mathbf{A})$ are the eigenvalues of the $k \times k$ SPD matrix $\mathbf{A}$. Namely:

Proposition 1 ([1], [2]). *If $n \in \mathbb{N}$ is even and $n \geqslant 3 \ln K(\mathbf{A})$, then the residuals satisfy*

$$\frac{\|e_n\|}{\|e_0\|} \leqslant \Big(\frac{3\ln K(\mathbf{A})}{n}\Big)^{n/2}. \tag{3}$$

K-condition numbers can be related to Frobenius norms if the matrix is considered as a perturbation of the identity matrix. The Frobenius norm of a symmetric matrix $\mathbf{B}$ is defined via

$$\|\mathbf{B}\|_F^2 := \sum_{i=1}^{k} \lambda_i(\mathbf{B})^2. \tag{4}$$

Then the following relation holds, which follows from the proof of the assertion (1) of Theorem 8 in [3]:

Proposition 2. *If* $\mathbf{B}$ *is positive semidefinite then*

$$\ln K(\mathbf{I}+\mathbf{B}) \leqslant \frac{1}{2}\|\mathbf{B}\|_F^2 \tag{5}$$

where $\mathbf{I}$ *is the identity matrix.*

Corollary 1. *If* $\mathbf{A}=\mathbf{I}+\mathbf{B}$ *for some positive semidefinite matrix* $\mathbf{B}$, *then the* CGM *for* (1) *satisfies*

$$\frac{\|e_n\|}{\|e_0\|} \leqslant \left(\frac{3\|\mathbf{B}\|_F^2}{2n}\right)^{n/2} \tag{6}$$

($n \in \mathbb{N}$ *is even and* $n \geqslant \frac{3}{2}\|\mathbf{B}\|_F^2$*).*

Our study of mesh independence will be based accordingly on Frobenius norms.

Remark 2.1. (i) Estimates (5)–(6) can be extended to indefinite $\mathbf{B}$ if $\mathbf{A}=\mathbf{I}+\mathbf{B}$ is still positive definite: then in both estimates the factors 2 are replaced by $2\lambda_{\min}(\mathbf{A})$. This also follows from the quoted proof in [3].

(ii) As mentioned before, the superlinear convergence result extends to Hilbert space for perturbations of the identity [9], [15], [28]. In [3] an analogue of (6) is proved, also for the case allowing B to be indefinite. The superlinear convergence does not hold for general bounded self-adjoint positive operators [13], here only the linear estimate is valid [9]. That is, the assumption that A is a compact perturbation of the identity is in some sense necessary for superlinear convergence. Accordingly, for such operators one can expect that the superlinear convergence factor of the underlying operator is an upper bound for those of the discretized problems.

3. Mesh independence of the CGM for discretized operator equations in Hilbert space

Let H be a separable Hilbert space and let us consider a linear operator equation

$$Bu = g \tag{7}$$

with some $g \in H$, under the following

Assumptions 1.

(i) The operator B is decomposed as $B = S + Q$ where S is a self-adjoint operator in H with dense domain D and Q is a self-adjoint operator defined on the domain H.

(ii) There exists $p > 0$ such that $\langle Su, u\rangle \geqslant p\|u\|^2$ $(u \in D)$.

(iii) $\langle Qu, u\rangle \geqslant 0$ $(u \in H)$.

(iv) The operator $S^{-1}Q$, defined on the energy space H_S, is a compact Hilbert-Schmidt operator, i.e.,

$$\|S^{-1}Q\|^2 \equiv \sum_{i=1}^{\infty} \lambda_i(S^{-1}Q)^2 < \infty \tag{8}$$

where $\lambda_i(S^{-1}Q)$ $(i \in \mathbb{N})$ are the eigenvalues of $S^{-1}Q$. (Here $\|S^{-1}Q\|$ is called the *Hilbert-Schmidt norm* of $S^{-1}Q$.)

We recall that the energy space H_S is the completion of D under the energy inner product $\langle u, v\rangle_S = \langle Su, v\rangle$, and the corresponding norm has the obvious notation $\|\cdot\|_S$. Assumption (ii) implies $H_S \subset H$. We also note that assumptions (i)–(ii) on S imply that $R(S) = H$ (see e.g. [24]), hence $S^{-1}Q$ makes sense indeed.

We replace equation (7) by its preconditioned form $(I + S^{-1}Q)u = S^{-1}g$. This is equivalent to the weak formulation

$$\langle u, v\rangle_S + \langle Qu, v\rangle = \langle g, v\rangle \quad (\forall\, v \in H_S), \tag{9}$$

which has a unique solution $u \in H_S$ since by assumption (iii) the bilinear form on the left is coercive on H_S.

Now equation (9) is solved numerically using a Galerkin discretization.

Construction of the discretization. Let $V = \operatorname{span}\{\varphi_1, \dots, \varphi_k\} \subset H_S$ be a given finite-dimensional subspace,

$$\mathbf{S} = \{\langle \varphi_i, \varphi_j\rangle_S\}_{i,j=1}^k \quad \text{and} \quad \mathbf{Q} = \{\langle Q\varphi_i, \varphi_j\rangle\}_{i,j=1}^k$$

the Gram matrices corresponding to S and Q. We look for the numerical solution $u_V \in V$ of equation (9) in V, i.e., for which

$$\langle u_V, v\rangle_S + \langle Qu_V, v\rangle = \langle g, v\rangle \quad (\forall\, v \in V). \tag{10}$$

Then $u_V = \sum_{j=1}^{k} c_j\varphi_j$, where $\mathbf{c} = (c_1, \dots, c_k) \in \mathbb{R}^k$ is the solution of the system

$$(\mathbf{S} + \mathbf{Q})\mathbf{c} = \mathbf{b} \tag{11}$$

with $\mathbf{b} = \{\langle g, \varphi_j \rangle\}_{j=1}^k$. (For simplicity the dependence on V in (11) is not indicated.) The matrix $\mathbf{S} + \mathbf{Q}$ is SPD.

As a discrete counterpart of the above used operator preconditioning with S, we introduce the matrix $\mathbf{S}$ as the preconditioner for the system (11) and replace this system by its preconditioned form

$$(\mathbf{I} + \mathbf{S}^{-1}\mathbf{Q})\mathbf{c} = \tilde{\mathbf{b}} \tag{12}$$

where $\tilde{\mathbf{b}} = \mathbf{S}^{-1}\mathbf{b}$ and $\mathbf{I}$ is the identity matrix in $\mathbb{R}^k$. Then we apply the CGM for the solution of the system (12).

Our main theorem below states that the Frobenius norm of $\mathbf{S}^{-1}\mathbf{Q}$ is bounded by the Hilbert-Schmidt norm of the operator $S^{-1}Q$. This bound is natural but not at all trivial (except for the case when the eigenvalues of $\mathbf{S}^{-1}\mathbf{Q}$ equal some of those of the operator $S^{-1}Q$, which does not hold for any practical discretization).

Theorem 1. *Let assumptions* 1 *hold. Then*

$$\|\mathbf{S}^{-1}\mathbf{Q}\|_F^2 \leqslant |\!|\!| S^{-1}Q |\!|\!|^2. \tag{13}$$

P r o o f. For simplicity let us only denote by λ_m the eigenvalues of $\mathbf{S}^{-1}\mathbf{Q}$. Let $\mathbf{c}^m = (c_1^m, \ldots, c_k^m) \in \mathbb{R}^k$ be the corresponding eigenvectors. Then

$$\mathbf{Q}\mathbf{c}^m = \lambda_m \mathbf{S}\mathbf{c}^m \tag{14}$$

for all m. Since $\mathbf{S}^{-1}\mathbf{Q}$ is self-adjoint with respect to the $\mathbf{S}$-inner product, therefore the eigenvalues are $\lambda_1, \ldots, \lambda_k$ (with multiplicity), and the vectors $\mathbf{c}^m$ $(m = 1, \ldots, k)$ are orthogonal in $\mathbb{R}^k$ with respect to the $\mathbf{S}$-inner product. Let us choose them such that they are also orthonormal:

$$\mathbf{S}c^m \cdot c^l = \delta_{ml} \quad (m, l = 1, \ldots, k), \tag{15}$$

where δ_{ml} is the Kronecker symbol.

Let $u_m = \sum\limits_{i=1}^k c_i^m \varphi_i \in V$ $(m = 1, \ldots, k)$. Then for all $m, l = 1, \ldots, k$

$$\langle u_m, u_l \rangle_S = \sum_{i,j=1}^k \langle \varphi_i, \varphi_j \rangle_S \, c_i^m c_j^l = \mathbf{S}\, c^m \cdot c^l, \tag{16}$$

hence (15) implies that $u_1, \ldots, u_k$ (as elements of H_S) form an orthonormal base in V with respect to the H_S-inner product. Then (14) and (15) yield

$$\mathbf{Q}\, c^m \cdot c^l = l_m \delta_{ml} \quad (m, l = 1, \ldots, k)$$

and, together with the analogue of (16) for Q, this implies

$$\langle Qu_m, u_l\rangle = \lambda_m \delta_{ml} \quad (m, l = 1, \ldots, k). \tag{17}$$

Let $u_{k+1}, u_{k+2}, \ldots$ be a complete orthonormal system in the orthocomplement of V in H_S. Then $u_1, u_2, \ldots$ form a complete orthonormal system in H_S. An invariance theorem on an arbitrary Hilbert-Schmidt operator L in a Hilbert space [14] asserts that

$$|||L|||^2 = \sum_{m,l=1}^{\infty} |\langle Lu_m, u_l\rangle|^2$$

for any complete orthonormal system (u_l). In our setting, we obtain for $S^{-1}Q$ in the space H_S that

$$|||S^{-1}Q|||^2 = \sum_{m,l=1}^{\infty} |\langle S^{-1}Qu_m, u_l\rangle_S|^2 = \sum_{m,l=1}^{\infty} |\langle Qu_m, u_l\rangle|^2. \tag{18}$$

Here, using (17), we have

$$\sum_{m=1}^{k} \lambda_m^2 = \sum_{m,l=1}^{k} |\langle Qu_m, u_l\rangle|^2 \tag{19}$$

with $\lambda_m = \lambda_m(\mathbf{S}^{-1}\mathbf{Q})$ $(m = 1, \ldots, k)$. Then (18) and (19) imply that (13) is satisfied.

□

Corollary 1 and Theorem 1 imply

Corollary 2. *The* CGM *applied to system* (12) *yields*

$$\frac{\|e_n\|}{\|e_0\|} \leqslant \left(\frac{3|||S^{-1}Q|||^2}{2n}\right)^{n/2} \tag{20}$$

(if $n \in \mathbb{N}$ is even and $n \geqslant \frac{3}{2}|||S^{-1}Q|||^2$). This estimate is independent of the subspace V.

Remark 3.2. (i) Theorem 1 can also be proved without the assumption $R(S) = H$, in this case the operator $S^{-1}Q$ is replaced by a suitable weak form.

(ii) Theorem 1 includes as a special case the non-preconditioned case when $S = I$ is the identity operator (and $D = H$). However, we are rather interested in the case when S is an unbounded operator including elliptic differential operators. The following section is devoted to such applications.

4. Mesh independence of the CGM for discretized linear elliptic problems

In this section we consider self-adjoint second order elliptic boundary value problems and their finite element discretizations. First we verify that by preconditioning an elliptic operator with its principal part, the superlinear convergence of the CGM becomes mesh independent. Then a computable estimate of the arising constant is given. Finally some examples illustrate that the arising preconditioned CG iteration has efficient realization—here we underline that our aim is not the construction of new numerical procedures but to demonstrate the mesh independence for known efficient methods.

4.1. The general mesh independence result

Let $N \leqslant 3$ and $\Omega \subset \mathbb{R}^N$ be a bounded domain. We consider the elliptic problem

$$\begin{cases} -\operatorname{div}(G\nabla u) + du = g, \\ u_{|\partial\Omega} = 0, \end{cases} \tag{21}$$

under the standard assumptions listed below. Our main interest is the case when the principal part has constant or separable coefficients, i.e.

$$G(x) \equiv G \in \mathbb{R}^{N\times N} \quad \text{or} \quad G(x) \equiv \operatorname{diag}\{G_i(x_i)\}_{i=1}^N,$$

whereas

$$d = d(x)$$

is a general variable (i.e. nonconstant) coefficient. Our goal is to use the principal part as preconditioning operator; such a setting has been considered e.g. in [6], [8], where the mesh independence of the rate of linear convergence has been observed.

Let the problem (21) satisfy the following assumptions:

(i) The symmetric matrix-valued function $G \in C^1(\overline{\Omega}, \mathbb{R}^{N\times N})$ satisfies

$$G(x)\xi\cdot\xi \geqslant m|\xi|^2 \quad (\xi \in \mathbb{R}^N) \tag{22}$$

with some $m > 0$ independent of ξ.

(ii) $d \in C(\overline{\Omega})$ and $d \geqslant 0$.

(iii) $\partial\Omega$ is piecewise C^2 and Ω is locally convex at the corners.

(iv) $g \in L^2(\Omega)$.

Then the problem (21) has a unique weak solution in $H_0^1(\Omega)$.

Let $V_h \subset H_0^1(\Omega)$ be a given FEM subspace. We look for the numerical solution u_h of (21) in V_h:

$$\int_\Omega (G\nabla u_h \cdot \nabla v + du_h v) = \int_\Omega gv \qquad (v \in V_h). \tag{23}$$

The corresponding linear algebraic system has the form

$$(\mathbf{G}_h + \mathbf{D}_h)\,\mathbf{c} = \mathbf{g}_h, \tag{24}$$

where $\mathbf{G}_h$ and $\mathbf{D}_h$ are the corresponding stiffness and mass matrices, respectively. We apply the matrix $\mathbf{G}_h$ as preconditioner, thus the preconditioned form of (24) is

$$(\mathbf{I}_h + \mathbf{G}_h^{-1}\mathbf{D}_h)\,\mathbf{c} = \tilde{\mathbf{g}}_h \tag{25}$$

with $\tilde{\mathbf{g}}_h = \mathbf{G}_h^{-1}\mathbf{g}_h$.

Let us apply the CGM to the system (25).

Theorem 2. *There exists a constant $\sigma > 0$ independent of the subspace V_h such that*

$$\|\mathbf{G}^{-1}\mathbf{D_h}\|_F \leqslant \sigma, \tag{26}$$

and hence the CGM *applied to the system* (25) *yields*

$$\frac{\|e_n\|}{\|e_0\|} \leqslant \Big(\frac{3\sigma^2}{2n}\Big)^{n/2} \tag{27}$$

if $n \in \mathbb{N}$ is even and $n \geqslant \frac{3}{2}\sigma^2$.

Proof. Let us consider the Hilbert space $H = L^2(\Omega)$ endowed with the usual inner product. Let $D = H^2(\Omega) \cap H_0^1(\Omega)$. We define the operators

$$Su \equiv -\operatorname{div}(G\nabla u) \quad (u \in D) \qquad \text{and} \qquad Qu \equiv du \quad (u \in L^2(\Omega)).$$

Then

$$\langle Su, u\rangle \geqslant m \int_\Omega |\nabla u|^2 \geqslant m\nu \int_\Omega u^2 \quad (u \in D)$$

and

$$\langle Qu, u\rangle = \int_\Omega du^2 \geqslant 0 \quad (u \in L^2(\Omega))$$

where $\nu > 0$ comes from the Sobolev inequality. By assumption (iii) the symmetric operator S maps onto $L^2(\Omega)$ (see [17]), hence it is self-adjoint [23] as well as the bounded operator Q. Further, $H_S = H_0^1(\Omega)$ with $\langle u, v\rangle_S = \int_\Omega G\nabla u \cdot \nabla v$ and

$$\langle S^{-1}Qu, v\rangle_S = \langle Qu, v\rangle = \int_\Omega duv \quad (v \in H_0^1(\Omega)),$$

hence $S^{-1}Q$ is a Hilbert-Schmidt operator in H_S [3].

This means that Assumptions 1 are satisfied, and therefore Theorem 1 and Corollary 2 hold for the system (25) with $\mathbf{S} = \mathbf{G}_h$ and $\mathbf{Q} = \mathbf{D}_h$. These imply (26) and (27), respectively, with

$$\sigma = \|S^{-1}Q\|. \tag{28}$$

□

Remark 4.2. Theorem 2 holds with more general conditions on (21) as well:

(i) It suffices to assume $G, d \in L^\infty$ instead of $G \in C^1$ and $q \in C$, further, the assumption (iii) on Ω may be omitted. In this case the proof uses Remark 3.2 (i) instead of Theorem 1.

(ii) We may allow $d(x) \geqslant -d_0 > -m\mu_1$, where μ_1 is the first eigenvalue of $-\Delta$ on Ω. Then the operator $I + S^{-1}Q$ is still strictly positive and we may rely on Remark 2.1.

(iii) The theorem also holds with mixed boundary conditions

$$u_{|\Gamma_D} = 0, \quad G\nabla u \cdot \nu_{|\Gamma_N} = \gamma$$

where $\partial\Omega$ is decomposed in measurable subparts Γ_D, Γ_N and $\gamma \in L^2(\Gamma_N)$. Then the FEM subspace is chosen as $V_h \subset H_D^1(\Omega) := \{u \in H^1(\Omega)\colon\ u_{|\Gamma_D} = 0\}$ and γ enters in (23) as

$$\int_\Omega (G\nabla u_h \cdot \nabla v + du_h v) = \int_\Omega gv + \int_{\Gamma_N} \gamma v \quad (v \in V_h).$$

Defining now the operator $Su = -\operatorname{div}(G\nabla u)$ with domain $D = H^2(\Omega) \cap H_D^1(\Omega)$, the proof of the theorem goes on in the same way using Remark 3.2 (i) again.

4.2. A computable estimate of the constant in the bound

The constant σ in Theorem 2 can be estimated as follows. We underline that (29) below not only gives an a priori bound for (26)–(27) but also helps to avoid a more costly calculation or estimation of the K-condition number of the matrix in (25).

Proposition 3. *Let* $d_\infty = \sup_\Omega d$ *and* m *be the lower spectral bound of* G *from* (22).

(1) *Let* $N = 2$ *and let* $R = [0,a] \times [0,b] \subset \mathbb{R}^2$ *be the smallest rectangle that contains a translate of* Ω. *Then*

$$\sigma^2 \leqslant \frac{d_\infty^2}{m^2\pi^4} \sum_{k,l=1}^{\infty} \left(\frac{k^2}{a^2} + \frac{l^2}{b^2}\right)^{-2}. \tag{29}$$

(2) *For* $N = 3$ *the obvious analogue holds, i.e,* $k^2/a^2 + l^2/b^2$ *is replaced by* $k^2/a^2 + l^2/b^2 + m^2/c^2$.

P r o o f. Let $\lambda_j = \lambda_j(S^{-1}Q)$ ($j \in \mathbb{N}^+$), then $\|S^{-1}Q\|^2 = \sum_{j=1}^{\infty} \lambda_j^2$. Since $S^{-1}Q$ is a compact operator in H_S, the variational characterization of the eigenvalues asserts that

$$\lambda_j = \inf_{v_1,\dots,v_{j-1}\in H_S} \sup_{u\perp \operatorname{span}\{v_1,\dots,v_{j-1}\}} \frac{\langle S^{-1}Qu, u\rangle_S}{\|u\|_S^2}.$$

Here

$$\frac{\langle S^{-1}Qu, u\rangle_S}{\|u\|_S^2} = \frac{\langle Qu, u\rangle}{\|u\|_S^2} = \frac{\int_\Omega du^2}{\int_\Omega G\nabla u\cdot\nabla u} \leqslant \frac{d_\infty \int_\Omega u^2}{m\int_\Omega |\nabla u|^2}.$$

Hence

$$\lambda_j \leqslant \frac{d_\infty}{m} \inf_{v_1,\dots,v_{j-1}\in H_S} \sup_{u\perp \operatorname{span}\{v_1,\dots,v_{j-1}\}} \frac{\int_\Omega u^2}{\int_\Omega |\nabla u|^2} = \frac{d_\infty}{m}\frac{1}{\mu_j(\Omega)},$$

where $\mu_j(\Omega)$ is the jth eigenvalue of $-\Delta$ on Ω with the Dirichlet boundary conditions. If $R \subset \mathbb{R}^2$ contains a translate of Ω then $\mu_j(R) \leqslant \mu_j(\Omega)$ (see [25]). Let $R = [0,a] \times [0,b]$ and let us re-index the eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}^+}$ into $\{\lambda_{kl}\}_{k,l\in\mathbb{N}^+}$ and similarly for μ_j. We thus obtain

$$\lambda_{kl} \leqslant \frac{d_\infty}{m}\frac{1}{\mu_{kl}(\Omega)} \leqslant \frac{d_\infty}{m}\frac{1}{\mu_{kl}(R)} = \frac{d_\infty}{m\pi^2}\left(\frac{k^2}{a^2} + \frac{l^2}{b^2}\right)^{-1} \quad (k, l \in \mathbb{N}^+).$$

This, together with (28), yields the required estimate.

The analogous statement for $N = 3$ is obvious. □

For an illustration of the magnitude of σ, let us consider the Dirichlet problem for the equation $-\Delta u + du = g$ on the unit square as a special case of (21). Then Proposition 3 yields that $\sigma \leqslant 0.0663\, d_\infty$.

(In general, the rough estimate $(k/a)^2 + (l/b)^2 \geqslant 2kl/ab$ yields the bound $\pi^4(\frac{1}{12}ab)^2$ for the infinite sum in (29), which implies $\sigma \leqslant \frac{1}{12}(ab/m)d_\infty$.)

4.3. Some efficient realizations and applications

As mentioned previously, our main interest is the case when the principal part of the elliptic operator has constant or separable coefficients, whereas $d = d(x)$ is a general variable (i.e. nonconstant) coefficient. Such a setting has been considered in various papers (e.g. [6], [8], [19], [22]) since there exist well-known fast solution methods for separable problems [5], [21], [26] that have made the separable principal part an efficient preconditioner. In these applications the mesh independence of the rate of linear convergence has been usually observed. Now our result of Subsection 4.1 provides mesh independence also for the superlinear convergence of CG in these preconditioning methods. We refer to some of these settings and related applications.

(a) *Separable principal part.* If the matrix G in (21) has the special form $G(x) \equiv \mathrm{diag}\{a_i(x_i)\}_{i=1}^N$ then the corresponding operator S is separable. Problems containing only the operator S can be solved efficiently by fast direct solvers [5], [21], [26], and the cost of this is of smaller order than for the original problem (21) whenever the function d is not separable.

(b) *Scaling of diffusion problems.* Let us consider the diffusion problem

$$\begin{cases} -\operatorname{div}(a\nabla u) = f, \\ u_{|\partial\Omega} = 0, \end{cases} \tag{30}$$

where $a \in C^2(\overline{\Omega})$ and $a(x) \geqslant m > 0$. Using the method of scaling [8], the problem (30) can be rewritten as

$$\begin{cases} -\Delta v + dv = g, \\ v_{|\partial\Omega} = 0, \end{cases} \tag{31}$$

with $d = a^{-1/2}\Delta(a^{1/2})$ and $g = a^{-1/2}f$. Hence we obtain a special case of (21) where now some fast Poisson solver [21], [26] can be applied for the efficient solution of the auxiliary problems.

(c) *Outer-inner iterations for nonlinear reaction-diffusion equations.* The Newton linearization of a nonlinear reaction-diffusion equation

$$\begin{cases} -\Delta u + f(x,u) = 0, \\ u_{|\partial\Omega} = 0 \end{cases}$$

involves auxiliary equations of the form

$$\begin{cases} -\Delta p_n + d_n p_n = g_n, \\ p_{n|\partial\Omega} = 0 \end{cases} \tag{32}$$

at step u_n, where $d_n = \partial_u f(x,u_n)$ and $g_n = \Delta u_n - f(x,u_n)$. For autocatalytic reactions, i.e. if $\partial_u f(x,u) \geqslant 0$, the problem (32) falls into the type (21). In view of our results, when the CGM is applied to solve the problems (32) using the principal part as preconditioner, then the overall outer-inner iteration converges superlinearly and at the same time it requires only Laplacian solvers.

(d) *Elliptic systems.* Theorem 2 can be extended to systems in a natural way. That is, for simplicity let us consider the system

$$\begin{cases} -\Delta u_i + d_{i1}u_1 + \ldots + d_{is}u_s = g_i, \\ u_{i|\partial\Omega} = 0 \quad (i = 1, \ldots, s), \end{cases} \tag{33}$$

with a symmetric positive semidefinite variable coefficient matrix $\{d_{ij}\}_{i,j=1}^{s}$. Then the mesh independent superlinear convergence can be proved in an analogous way using the operator n-tuple $S(u_1, \ldots, u_s) \equiv (-\Delta u_i)_{i=1}^{s}$. Since this S consists of independent Laplacians, the auxiliary problems are not only separable but also have smaller size than the original one. In the context of the previous paragraph (c), such systems arise in the Newton linearization of a nonlinear reaction-diffusion system which corresponds to a potential $\varphi(u_1, \ldots, u_s) = \int_\Omega \Big(\frac{1}{2} \sum_{i=1}^{s} |\nabla u_i|^2 + F(u_1, \ldots, u_s)\Big)$.

Remark 4.3. Most of the above fast solvers also extend to the case when a separable or constant coefficient lower order term is added to the separable principal part. Such an operator can also be proposed for the role of S (see [8], [19]) and, clearly, in this case the bounds in Proposition 3 are improved to some extent. The details are left to the interested reader.

References

[1] *O. Axelsson*: Iterative Solution Methods. Cambridge University Press, Cambridge, 1994.

[2] *O. Axelsson, I. Kaporin*: On the sublinear and superlinear rate of convergence of conjugate gradient methods. Mathematical journey through analysis, matrix theory and scientific computation (Kent, OH, 1999). Numer. Algorithms *25* (2000), 1–22.

[3] *O. Axelsson, J. Karátson*: On the rate of convergence of the conjugate gradient method for linear operators in Hilbert space. Numer. Funct. Anal. Optmization *23* (2002), 285–302.

[4] *O. Axelsson, J. Karátson*: Superlinearly convergent CG methods via equivalent preconditioning for nonsymmetric elliptic operators. Numer. Math. *99* (2004), 197–223. SpringerLink DOI: 10.1007/s00211-004-0557-2 (electronic).

[5] *R. E. Bank, D. J. Rose*: Marching algorithms for elliptic boundary value problems. I. The constant coefficient case. SIAM J. Numer. Anal. *14* (1977), 792–829.

[6] *R. E. Bank*: Marching algorithms for elliptic boundary value problems. II. The variable coefficient case. SIAM J. Numer. Anal. *14* (1977), 950–970.

[7] *B. Beckermann, A. B. J. Kuijlaars*: Superlinear convergence of conjugate gradients. SIAM J. Numer. Anal. *39* (2001), 300–329. Electronic.

[8] *P. Concus, G. H. Golub*: Use of fast direct methods for the efficient numerical solution of nonseparable elliptic equations. SIAM J. Numer. Anal. *10* (1973), 1103–1120.

[9] *J. W. Daniel*: The conjugate gradient method for linear and nonlinear operator equations. SIAM J. Numer. Anal. *4* (1967), 10–26.

[10] *H. C. Elman, M. H. Schultz*: Preconditioning by fast direct methods for nonself-adjoint nonseparable elliptic equations. SIAM J. Numer. Anal. *23* (1986), 44–57.
[11] *V. Faber, T. Manteuffel, and S. V. Parter*: On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations. Adv. Appl. Math. *11* (1990), 109–163.
[12] *I. Faragó, J. Karátson*: Numerical solution of nonlinear elliptic problems via preconditioning operators. Theory and applications. Advances in Computation, Vol. 11. NOVA Science Publishers, Huntington, 2002.
[13] *Z. Fortuna*: Some convergence properties of the conjugate gradient method in Hilbert space. SIAM J. Numer. Anal. *16* (1979), 380–384.
[14] *I. Gohberg, S. Goldberg, and M. A. Kaashoek*: Classes of linear operators, Vol. I. Operator Theory: Advances and Applications, Vol. 49. Birkhäuser-Verlag, Basel, 1990.
[15] *R. M. Hayes*: Iterative methods of solving linear problems in Hilbert space. Natl. Bur. Stand.; Appl. Math. Ser. *39* (1954), 71–103.
[16] *M. R. Hestenes, E. Stiefel*: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand., Sect. B *49* (1952), 409–436.
[17] *J. Kadlec*: On the regularity of the solution of the Poisson problem on a domain with boundary locally similar to the boundary of a convex open set. Czechoslovak Math. J. *14(89)* (1964), 386–393. (In Russian.)
[18] *J. Karátson, I. Faragó*: Variable preconditioning via quasi-Newton methods for nonlinear problems in Hilbert space. SIAM J. Numer. Anal. *41* (2003), 1242–1262.
[19] *T. Manteuffel, J. Otto*: Optimal equivalent preconditioners. SIAM J. Numer. Anal. *30* (1993), 790–812.
[20] *J. W. Neuberger*: Sobolev gradients and differential equations. Lecture Notes in Math., No. 1670. Springer-Verlag, Berlin, 1997.
[21] *T. Rossi, J. Toivanen*: A parallel fast direct solver for block tridiagonal systems with separable matrices of arbitrary dimension. SIAM J. Sci. Comput. *20* (1999), 1778–1793.
[22] *T. Rossi, J. Toivanen*: Parallel fictitious domain method for a non-linear elliptic Neumann boundary value problem. Czech-US Workshop in Iterative Methods and Parallel Computing, Part I (Milovy, 1997). Numer. Linear Algebra Appl. *6* (1999), 51–60.
[23] *W. Rudin*: Functional Analysis. McGraw-Hill, New York, 1991.
[24] *F. Riesz, B. Sz.-Nagy*: Vorlesungen über Funktionalanalysis. VEB Deutscher Verlag der Wissenschaften, Berlin, 1982.
[25] *L. Simon, E. Baderko*: Linear Partial Differential Equations of Second Order. Tankönyvkiadó, Budapest, 1983. (In Hungarian.)
[26] *P. N. Swarztrauber*: The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle. SIAM Rev. *19* (1977), 490–501.
[27] *Yu. V. Vorobyev*: Methods of Moments in Applied Mathematics. Gordon and Breach, New York, 1965.
[28] *R. Winter*: Some superlinear convergence results for the conjugate gradient method. SIAM J. Numer. Anal. *17* (1980), 14–17.

Author's address: *J. Karátson*, Department of Applied Analysis, ELTE University, H-1518 Budapest, Hungary, e-mail: karatson@cs.elte.hu.