NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2002; **9**:379–397 (DOI: 10.1002/nla.280)

Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method

Xiao-Chuan Cai[1,*,†] and Jun Zou[2,‡]

[1]*Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309-0430, U.S.A.*
[2]*Department of Mathematics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong*

SUMMARY

Additive Schwarz preconditioned GMRES is a powerful method for solving large sparse linear systems of equations on parallel computers. The algorithm is often implemented in the Euclidean norm, or the discrete l^2 norm, however, the optimal convergence result is available only in the energy norm (or the equivalent Sobolev H^1 norm). Very little progress has been made in the theoretical understanding of the l^2 behaviour of this very successful algorithm. To add to the difficulty in developing a full l^2 theory, in this note, we construct explicit examples and show that the optimal convergence of additive Schwarz preconditioned GMRES in l^2 cannot be obtained using the existing GMRES theory. More precisely speaking, we show that the symmetric part of the preconditioned matrix, which plays a role in the Eisenstat–Elman–Schultz theory, has at least one negative eigenvalue, and we show that the condition number of the best possible eigenmatrix that diagonalizes the preconditioned matrix, key to the Saad–Schultz theory, is bounded from both above and below by constants multiplied by $h^{-1/2}$. Here h is the finite element mesh size. The results presented in this paper are mostly negative, but we believe that the techniques used in our proofs may have wide applications in the further development of the l^2 convergence theory and in other areas of domain decomposition methods. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: domain decomposition; additive Schwarz preconditioner; Krylov subspace iterative method; finite elements; eigenvalue; eigenmatrix

1. INTRODUCTION

Additive Schwarz (AS) preconditioned generalized minimum residual (GMRES) method is a powerful method for solving large sparse non-symmetric linear system of equations arising

*Correspondence to: Xiao-Chuan Cai, Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309-0430, U.S.A.
†E-mail: cai@cs.colorado.edu
‡E-mail: zou@math.cuhk.edu.hk

Contract/grant sponsor: The work of XCC was supported in part by the NSF grants ECS-9725504 and ACI-0072089
Contract/grant sponsor: The work of JZ was supported in part by the Hong Kong RGC; contract/grant no. CUHK4292/00P

Received 27 April 2001
Revised 10 January 2002

from discretizations of boundary value problems of partial differential equations, especially on parallel computers with large number of processors. AS/GMRES has been implemented in several large software packages for solving partial differential equations or general sparse linear systems, such as PETSc [1] and PSPARSLIB [2]. The Euclidean norm (l^2) has been used in all the implementations as far as we know. However, the convergence result is available only in the energy norm (or the equivalent Sobolev H^1 norm). For example, the theory for symmetric positive definite problems can be found in References [3–7] and for non-symmetric and indefinite problems in References [8, 9]. Very little progress has been made toward the theoretical understanding of the *optimal* convergence of AS/GMRES in l^2. As customary in the domain decomposition literature [3, 7], '*optimal*' refers to the fact that the convergence rate is independent of the finite element mesh size and the number of subdomains. In the past ten years, many numerical experiments have been carried out, using AS/GMRES, for solving linear systems obtained from the discretization of scalar and systems of linear and non-linear partial differential equations, see for examples References [10–15] and all seem to indicate that the method is optimal in l^2; i.e. the convergence rate has no or very little dependence on the mesh size, however, none of the existing AS/GMRES theory confirms this even for the one dimensional case. In this paper, we show constructively that the l^2 optimal convergence theory cannot be obtained by using any of the two existing GMRES theories.

We now briefly summarize the main findings of the paper. First, the Eisenstat–Elman–Schultz theory [16] for the convergence of GMRES requires that the symmetric part of the preconditioned system to be positive definite. Using a simple example, we show that this condition cannot be satisfied with AS/GMRES. At least one of the eigenvalues of the symmetric part is negative. Second, the Saad–Schultz theory [17] assumes that the preconditioned system is diagonalizable by a certain eigenmatrix X, and the convergence rate is bounded by the condition number of X in l^2 and the distribution of the eigenvalues. The theory is not very useful in practice because the estimate of the condition number of X is often too hard to obtain. Using the same example, we are able to estimate the condition number of X and show that it has a $h^{-1/2}$ dependence on the mesh size.

In the rest of this section, we recall some of the key components of the Eisenstat–Elman–Schultz theory and the Saad–Schultz theory. We consider a non-singular linear system of equations of size $n \times n$

$$Au = f \tag{1}$$

and let M^{-1} be a preconditioner for A. The solution of (1) is often obtained by solving iteratively the preconditioned system

$$Tu = g \tag{2}$$

where $T = M^{-1}A$ and $g = M^{-1}f$. For generality, we do not assume that T is symmetric and we use GMRES [18] as the iterative solver. We shall use $(\cdot, \cdot)$ and $\|\cdot\|_2$ to denote the usual Euclidean inner product and norm in $\mathbb{R}^n$, respectively. The main result of the paper concerns the optimal l^2 convergence rate of GMRES for solving (2) with additive Schwarz as the preconditioner.

Two types of convergence theory are available for the convergence of GMRES. One theory, due to Eisenstat *et al.* [16], is for the case when the symmetric part of the matrix T is positive definite. More precisely, let us assume that there exist two positive constants c_0 and C_0, such that

$$(Tx,x)\geqslant c_0\|x\|_2^2 \quad \text{and} \quad \|Tx\|_2\leqslant C_0\|x\|_2$$

for any $x\in\mathbb{R}^n$. Then the residuals satisfy

$$\|r_k\|_2\leqslant\left(1-\frac{c_0^2}{C_0^2}\right)^{k/2}\|r_0\|_2$$

Unfortunately, with the additive Schwarz preconditioner, the symmetric part of T with respect to the l^2 inner product is generally not positive definite, i.e. $c_0<0$. An example will be given later to illustrate the case. Another theory due to Saad and Schultz [17] assumes that the matrix T is diagonalizable, i.e. there exists a matrix X such that $T=X\Lambda X^{-1}$, where $\Lambda=\{\lambda_1,\dots,\lambda_n\}$ is a diagonal matrix of eigenvalues. Let Π_k be the space of polynomials of degree less than or equal to k, and

$$\varepsilon^{(k)}=\min_{p\in\Pi_k,\,p(0)=1}\ \max_{i=1,\dots,n}|p(\lambda_i)| \tag{3}$$

Then the residuals of GMRES satisfy

$$\|r_k\|_2\leqslant\kappa_2(X)\varepsilon^{(k)}\|r_0\|_2 \tag{4}$$

Here $\kappa_2(X)=\|X\|_2\|X^{-1}\|_2$ is the condition number of the eigenmatrix. Later in this paper, we prove that the additive Schwarz preconditioned linear operator is indeed diagonalizable, and also estimate $\kappa_2(X)$. The choice of X is not unique, however, we show, using an interesting result of Demmel [20], that even with the best possible eigenmatrix X, $\kappa_2(X)$ has a bound depending on the finite element mesh size from both above and below regardless the size of the overlap. Therefore we cannot claim that the method is optimal.

Other techniques have also been used to study the convergence of the preconditioned GMRES method, such as the method based on the field-of-values analysis in References [20, 21]. The rest of the paper is organized as follows. In Section 2, we review the classical additive Schwarz method. Section 3 is devoted to the Eisenstat–Elman–Schultz theory, and Section 4 to the Saad–Schultz theory.

2. A BRIEF REVIEW OF THE ADDITIVE SCHWARZ METHOD

Although the method we study is mainly for non-symmetric linear systems, we shall focus on a simple symmetric elliptic Dirichlet boundary value problem: Find $u\in H_0^1(\Omega)$ such that

$$a(u,v)=(f,v)_{L^2(\Omega)} \quad \forall v\in H_0^1(\Omega) \tag{5}$$

where $(\cdot,\cdot)_{L^2(\Omega)}$ is the continuous L^2 inner product, the bilinear form $a(u,v)$ is defined by $a(u,v)=\int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}\Omega$ and $f(x)\in L^2(\Omega)$ is given. Here Ω is an open bounded polygon with boundary $\partial\Omega$. We assume that the diameter of Ω is of order 1. To introduce the finite element discretization and the finite element space V_h, we let $\mathscr{T}^h=\{\tau_i\}$ be a standard quasi-uniform finite element triangulation of Ω with interior nodal points denoted as $W=\{x_1,x_2,\dots,x_n\}$ and the standard basis functions as $\{\phi_{x_i}(x)\}$, i.e. $\phi_{x_i}(x_j)=\delta_{ij}$. The finite element problem reads as follows: Find $u^*\in V_h\subset H_0^1(\Omega)$ such that

$$a(u^*,v)=(f,v)_{L^2(\Omega)}\quad \forall v\in V_h \tag{6}$$

To discuss the overlapping additive Schwarz methods, we introduce the partition of Ω into $\{\Omega_i\}$, such that no $\partial\Omega_i$ cuts through any elements τ_i, and $\bar{\Omega}=\bigcup_{i=1}^N\bar{\Omega}_i$. We assume Ω_i is an open domain. We extend each Ω_i to a larger subdomain $\Omega_i'\supset\Omega_i$, which is also assumed not to cut any fine mesh triangles. For each Ω_i', we define a finite element space $V_i\equiv V_h\cap H_0^1(\Omega_i')$ extended by zero outside Ω_i'. Let n_i be the dimension of V_i, that equals the number of interior nodes in Ω_i'. We now define the subdomain mapping operator $T_i: V_h\rightarrow V_i$ as

$$a(T_iu,v)=a(u,v)\quad \forall u\in V_h \text{ and } \forall v\in V_i \tag{7}$$

and

$$T=T_1+\cdots+T_N \tag{8}$$

Since $g=Tu^*$ can be pre-calculated without knowing u^*, we define the additive Schwarz preconditioned GMRES as follows:

Algorithm 2.1 (*AS/GMRES*)
Solve

$$Tu^*=g \tag{9}$$

by GMRES with any initial guess.

The focus of this paper is to understand the L^2 convergence of AS/GMRES. We first make a simple observation about the eigenvalues and eigenvectors of the operator T. We define W_Γ as a subset of W consisting of all the mesh points on the internal boundaries, i.e.

$$W_\Gamma\equiv\left\{x_i \mid \in W, x_i\in\left(\bigcup\partial\Omega_i'\right)\cap\Omega\right\}$$

For each $x_i\in W\backslash W_\Gamma$, using the definition of T_i, we have

$$T\phi_{x_i}(x)=\sigma_i\phi_{x_i}(x) \tag{10}$$

where $\sigma_i\geqslant 1$ is an integer which equals the number of subdomains that x_i belongs to. This implies that $\phi_{x_i}(x)$ $(x_i\in W\backslash W_\Gamma)$ are the eigenfunctions of T and the corresponding eigenvalues are integers. Note that this is not true for the nodal points on the internal boundaries. Much of the rest of the paper is devoted to the explicit calculation of the eigenvalues and eigenvectors associated with the nodes on the internal boundaries.

To simplify the notations, we shall mix up the notions of operators and matrices, finite element functions and vectors. The mix-up is always understood in the sense of the standard basis functions. For example in the vector sense, we have

$$(\phi_i,\phi_j)=\delta_{ij}$$

This may not be true in the continuous inner product $(\cdot,\cdot)_{L^2(\Omega)}$.

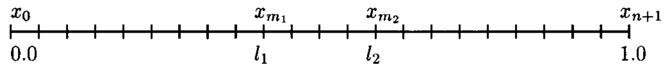

Figure 1. $\Omega_1' = (0, l_2)$ and $\Omega_2' = (l_1, 1.0)$.

3. ANALYSIS USING THE EISENSTAT–ELMAN–SCHULTZ THEORY

In this section, we study the optimal convergence of AS/GMRES using the Eisenstat–Elman–Schultz theory. In other words, we calculate the smallest eigenvalue of the symmetric part of T. We work on a simple one-dimensional problem, and show, sadly, that the smallest eigenvalue is negative. Therefore, we conclude that the Eisenstat–Elman–Schultz theory is generally not applicable for studying the L^2 optimal convergence of the AS/GMRES method.

Consider a one-dimensional Poisson problem defined on the unit interval $(0, 1)$ divided into two overlapping subdomains. As in Figure 1, we divide the unit interval into $n+1$ subintervals with length $h = 1/(n+1)$ and mesh points $x_i = ih$, $i = 0, \ldots, n+1$. We define the overlapping subdomains $\Omega_1' = (0, l_2)$ and $\Omega_2' = (l_1, 1.0)$. n is the total number of interior mesh points. For simplicity, we assume that

$$l_1 = 1 - l_2 \tag{11}$$

Let n_1^0 be the number of mesh points in $(0, l_1)$, n_2^0 the number of mesh points in $(l_2, 1.0)$, and n_{12} the number of mesh points in the overlapping region (l_1, l_2). Note that $n_1^0 + n_2^0 + n_{12} = n - 2$.

Let $V_h \subset H_0^1(0, 1)$ be the piecewise linear continuous finite element space and $\{\phi_{x_i}, i = 1, \ldots, n\}$ the collection of the usual basis functions; i.e. $\phi_{x_i} \in V_h$, $\phi_{x_i}(x_j) = 1$ if $i = j$ and $\phi_{x_i}(x_j) = 0$ if $i \neq j$. We define the subspaces $V_i = V_h \cap H_0^1(\Omega_i')$ whose dimension is n_i. We denote the two special mesh points l_1 and l_2 as x_{m_1} and x_{m_2} with m_1 and m_2 being two positive integers, and the corresponding basis functions as ϕ_{m_1} and ϕ_{m_2}.

To define the matrix form of T_i, we need to introduce an interpolation matrix I_i from V_i to V_h and a restriction operator from V_h to V_i. For any $v_h \in V_i$, let v_i and v be the coefficient vectors of v_h in terms of the basis of V_i and V_h, respectively, i.e. $v_i = (v_h(x_j))_{x_j \in \Omega_i'}$, $v = (v_h(x_j))_{x_j \in \Omega}$, then I_i is defined by $I_i v_i = v$, and I_i is a $n \times n_i$ matrix with all entries being either 0 or 1. The restriction matrix from V_h to V_i is defined to be the transpose of I_i, i.e. I_i^{t}. Let A and A_i be the stiffness matrices corresponding to the discretizations of the Poisson's problem on Ω and Ω_i' with zero Dirichlet boundary condition, respectively, then we have

$$T_i = I_i A_i^{-1} I_i^{\mathrm{t}} A \tag{12}$$

Obviously, the transpose of T_i is $T_i^{\mathrm{t}} = A I_i A_i^{-1} I_i^{\mathrm{t}}$. The Eisenstat–Elman–Schultz theory depends on the smallest eigenvalue of the symmetric part of T. We consider the following eigenvalue problem:

$$(T + T^{\mathrm{t}})u = \lambda u \tag{13}$$

We start the analysis with several lemmas.

Lemma 3.1

For two basis functions ϕ_{m_1} and ϕ_{m_2} at the nodes x_{m_1} and x_{m_2}, we have

$$T_1 \phi_{m_2} = -\frac{l_2 - h}{l_2} \tilde{\phi}_{m_2 - 1} \quad \text{and} \quad T_2 \phi_{m_1} = -\frac{1 - l_1 - h}{1 - l_1} \tilde{\phi}_{m_1 + 1} \tag{14}$$

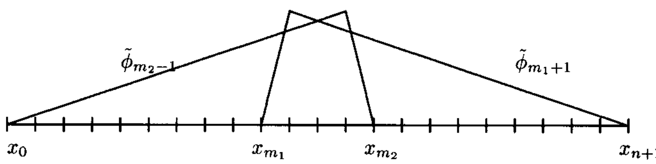

Figure 2. Two special piecewise linear functions $\tilde{\phi}_{m_1+1}$ and $\tilde{\phi}_{m_2-1}$.

where the piecewise linear continuous functions $\tilde{\phi}_{m_2-1}$ and $\tilde{\phi}_{m_1+1}$ are given below (Figure 2)

$$\tilde{\phi}_{m_1+1}(x)=\begin{cases} 0 & x\leqslant l_1 \\ \text{linear} & x\in(l_1, l_1+h) \\ 1 & x=l_1+h \\ \text{linear} & x\in(l_1+h, 1.0) \\ 0 & x=1.0 \end{cases} \qquad \tilde{\phi}_{m_2-1}(x)=\begin{cases} 0 & x=0.0 \\ \text{linear} & x\in(0, l_2-h) \\ 1 & x=l_2-h \\ \text{linear} & x\in(l_2-h, l_2) \\ 0 & x\geqslant l_2 \end{cases}$$

Proof
Using the definition of T_1, we have

$$a(T_1\phi_{m_2}, v_1)=a(\phi_{m_2}, v_1) \quad \forall v_1\in V_1 \tag{15}$$

We first observe that $a(\phi_{m_2}, v_1)=0$ if $v_1(x)$ is any of the interior nodal basis functions in $(0, l_2-h)$. Therefore, we claim that $T_1\phi_{m_2}$ is a discrete harmonic function in the interval $(0, l_2-h)$. This implies that $T_1\phi_{m_2}$ is a linear function in this interval. Since $T_1\phi_{m_2}$ must be linear in (l_2-h, l_2), we hence have $T_1\phi_{m_2}=\alpha\tilde{\phi}_{m_2-1}$ for some constant α. And α can be calculated by taking $v_1=\phi_{m_2-1}$ in (15), i.e.

$$\alpha=\frac{a(\phi_{m_2}, \phi_{m_2-1})}{a(\tilde{\phi}_{m_2-1}, \phi_{m_2-1})}=-\frac{l_2-h}{l_2}$$

The result for T_2 can be proved similarly. □

Lemma 3.2
For the adjoint operators T_i^{t}, $i=1,2$, we have

$$T_1^{\mathrm{t}}\phi_k=\phi_k-\frac{x_k}{l_2}\phi_{m_2} \quad \forall x_k\in(0, l_2); \qquad T_1^{\mathrm{t}}\phi_k=0 \quad \forall x_k\in[l_2, 1)$$
$$T_2^{\mathrm{t}}\phi_k=\phi_k-\frac{1-x_k}{1-l_1}\phi_{m_1} \quad \forall x_k\in(l_1, 1); \qquad T_2^{\mathrm{t}}\phi_k=0 \quad \forall x_k\in(0, l_1]$$

Proof
By definition, for any nodal point $x_k\in(0, l_2)$, we have

$$(T_1^{\mathrm{t}}\phi_k, v)=(\phi_k, T_1 v) \quad \forall v\in V \tag{16}$$

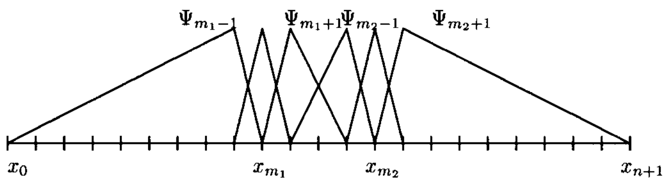

Figure 3. Four special piecewise linear functions and two standard basis functions at x_{m_1} and x_{m_2}.

Since $T_1\phi_j=\phi_j$ if $x_j<l_2$ and $T_1\phi_j=0$ if $x_j>l_2$, and also $(\phi_k,\phi_j)=\delta_{kj}$, we have

$$(T_1^{\mathrm{t}}\phi_k,\phi_j)=\begin{cases}\delta_{kj} & x_j<l_2\\ 0 & x_j>l_2\end{cases}$$

Hence $T_1^{\mathrm{t}}\phi_k$ must have the following expression

$$T_1^{\mathrm{t}}\phi_k=\phi_k+\alpha\phi_{m_2} \tag{17}$$

for some constant α to be determined below. To calculate α, we substitute (17) into (16) and take $v=\phi_{m_2}$,

$$\alpha(\phi_{m_2},\phi_{m_2})+(\phi_k,\phi_{m_2})=(\phi_k,T_1\phi_{m_2})$$

combining with (14) we obtain $\alpha=-x_k/l_2$. The fact that $T_1^{\mathrm{t}}\phi_k=0$ for $x_k\in[l_2,1)$ follows immediately from (16) by taking $v=\phi_j$ for all $x_j\in(0,1)$. The result for T_2^{t} can be proved in a similar way. □

We next show the main result of this section:

Theorem 3.1

For the one-dimensional Poisson problem defined on $(0,1)$, when the mesh size h is sufficiently small, the symmetric part of the additive Schwarz preconditioned operator T has at least two negative eigenvalues, one corresponding to a symmetric eigenfunction and the other corresponding to an anti-symmetric eigenfunction.

Proof

We first introduce four special piecewise linear continuous functions as shown in Figure 3:

$$\Psi_{m_1-1}(x)=\begin{cases}0 & x=0\\ \text{linear} & x\in(0,l_1-h)\\ \phi_{m_1-1} & x\in[l_1-h,l_1]\end{cases}$$

$$\Psi_{m_1+1}(x)=\begin{cases}\phi_{m_1+1} & x\in[l_1,l_1+h)\\ \text{linear} & x\in[l_1+h,l_2-h)\\ 0 & x=l_2-h\end{cases}$$

$$]\Psi_{m_2-1}(x)=\begin{cases} 0 & x=l_1+h \\ \text{linear} & x\in(l_1+h,l_2-h) \\ \phi_{m_2-1} & x\in[l_2-h,l_2] \end{cases}$$

$$\Psi_{m_2+1}(x)=\begin{cases} \phi_{m_2+1} & x\in[l_2,l_2+h) \\ \text{linear} & x\in[l_2+h,1) \\ 0 & x=1 \end{cases}$$

Let λ be the possible negative eigenvalue that we are looking for and $\Psi(x)$ be the corresponding eigenfunction of the following form

$$\Psi=\alpha_1\Psi_{m_1-1}+\alpha_2\phi_{m_1}+\alpha_3\Psi_{m_1+1}+\alpha_4\Psi_{m_2-1}+\alpha_5\phi_{m_2}+\alpha_6\Psi_{m_2+1}$$

Here α_i $(i=1,\dots,6)$ are real parameters to be determined. By the definition of T_i, we have immediately

$$\begin{aligned} T_1\tilde{\Psi} &= \tilde{\Psi} \text{ for } \tilde{\Psi}=\Psi_{m_1-1},\phi_{m_1},\Psi_{m_1+1},\Psi_{m_2-1} \\ T_1\Psi_{m_2+1} &= 0 \\ T_2\tilde{\Psi} &= \tilde{\Psi} \text{ for } \tilde{\Psi}=\Psi_{m_1+1},\Psi_{m_2-1},\phi_{m_2},\Psi_{m_2+1} \\ T_2\Psi_{m_1-1} &= 0 \end{aligned}$$

Following Lemma 3.2, we also have that $T_1^{\mathrm{t}}\phi_{m_2}=T_1^{\mathrm{t}}\Psi_{m_2+1}=0$, $T_2^{\mathrm{t}}\Psi_{m_1-1}=T_2^{\mathrm{t}}\phi_{m_1}=0$,

$$T_1^{\mathrm{t}}\tilde{\Psi}=\tilde{\Psi}-\frac{1}{l_2}(\tilde{\Psi},x)\phi_{m_2} \quad \text{for } \tilde{\Psi}=\Psi_{m_1-1},\phi_{m_1},\Psi_{m_1+1},\Psi_{m_2-1} \tag{18}$$

and

$$T_2^{\mathrm{t}}\tilde{\Psi}=\tilde{\Psi}-\frac{1}{(1-l_1)}(\tilde{\Psi},1-x)\phi_{m_1} \quad \text{for } \tilde{\Psi}=\Psi_{m_1+1},\Psi_{m_2-1},\phi_{m_2},\Psi_{m_2+1} \tag{19}$$

To simplify the notation, we denote the four coefficients (without the minus sign) of ϕ_{m_2} in (18) and of ϕ_{m_1} in (19) by μ_1,μ_2,μ_3,μ_4 and $\tilde{\mu}_3,\tilde{\mu}_4,\tilde{\mu}_5,\tilde{\mu}_6$, respectively. Substituting these relations into

$$(T+T^{\mathrm{t}})\Psi=\lambda\Psi$$

then equating the coefficients in both sides and using the fact that

$$\begin{aligned} \tilde{\phi}_{m_2-1} &= \frac{l_1-h}{l_2-h}\Psi_{m_1-1}+\frac{l_1}{l_2-h}\phi_{m_1}+\frac{l_1+h}{l_2-h}\Psi_{m_1+1}+\Psi_{m_2-1} \\ \tilde{\phi}_{m_1+1} &= \Psi_{m_1+1}+\frac{1-l_2+h}{1-l_1-h}\Psi_{m_2-1}+\frac{1-l_2}{1-l_1-h}\phi_{m_2}+\frac{1-l_2-h}{1-l_1-h}\Psi_{m_2+1} \end{aligned}$$

we obtain a reduced eigenvalue problem with six variables

$$\lambda\alpha_1=2\alpha_1-\frac{l_1-h}{l_2}\alpha_5 \tag{20}$$

$$\lambda\alpha_2 = 2\alpha_2 - \frac{l_1}{l_2}\alpha_5 - \tilde{\mu}_3\alpha_3 - \tilde{\mu}_4\alpha_4 - \tilde{\mu}_5\alpha_5 - \tilde{\mu}_6\alpha_6 \tag{21}$$

$$\lambda\alpha_3 = 4\alpha_3 - \frac{l_1+h}{l_2}\alpha_5 - \frac{1-l_1-h}{1-l_1}\alpha_2 \tag{22}$$

$$\lambda\alpha_4 = 4\alpha_4 - \frac{l_2-h}{l_2}\alpha_5 - \frac{1-l_2+h}{1-l_1}\alpha_2 \tag{23}$$

$$\lambda\alpha_5 = 2\alpha_5 - \frac{1-l_2}{1-l_1}\alpha_2 - \mu_1\alpha_1 - \mu_2\alpha_2 - \mu_3\alpha_3 - \mu_4\alpha_4 \tag{24}$$

$$\lambda\alpha_6 = 2\alpha_6 - \frac{1-l_2-h}{1-l_1}\alpha_2 \tag{25}$$

Although there are only six variables, it seems not easy to solve it explicitly by hand. Since $l_1 = 1 - l_2$, the subdomain partition is symmetric. We tend to believe that the eigenfunctions are either symmetric or anti-symmetric with respect to the center point of the domain. Let us first work on the symmetric eigenfunctions case, namely we assume

$$\alpha_1 = \alpha_6, \quad \alpha_2 = \alpha_5, \quad \alpha_3 = \alpha_4$$

Then the following holds from (20), (22) and (23) by using the above assumption:

$$\lambda\alpha_1 = 2\alpha_1 - \frac{l_1-h}{l_2}\alpha_5 \tag{26}$$

$$\lambda\alpha_3 = 4\alpha_3 - \frac{l_1+h}{l_2}\alpha_5 - \frac{1-l_1-h}{1-l_1}\alpha_5 \tag{27}$$

$$\lambda\alpha_5 = 2\alpha_5 - \frac{1-l_2}{1-l_1}\alpha_5 - \mu_1\alpha_1 - \mu_2\alpha_5 - (\mu_3+\mu_4)\alpha_3 \tag{28}$$

As $(T+T^{\rm t})$ is symmetric, its eigenvalues must be real, and three of them are given by the eigenvalues of the 3×3 system (26)–(28). With these three eigenvalues, we can immediately get α_1, α_3 and α_5 from (26) to (28), and so find three symmetric eigenfunctions for $(T+T^{\rm t})$.

To prove that $(T+T^{\rm t})$ has at least one negative eigenvalue is equivalent to show that the determinant of the 3×3 coefficient matrix on the right-hand side of (26)–(28) is negative as the product of the three eigenvalues is just the determinant. Using the fact that $\mu_2 = l_1/l_2$, one can easily check that the determinant is given by

$$C_1 = 16 - 16\frac{l_1}{l_2} - 4\frac{l_1-h}{l_2}\mu_1 - \frac{2}{l_2}(\mu_3+\mu_4) \tag{29}$$

We next calculate μ_1 and $\mu_3+\mu_4$ explicitly. Recall that

$$\mu_1 = \frac{1}{l_2}(\Psi_{m_1-1}, x) = \frac{1}{l_2}\left(\sum_{j=1}^{m_1-1}\frac{x_j^2}{l_1-h}\right) = \frac{l_1(2l_1-h)}{6l_1 h} \tag{30}$$

and

$$\mu_3+\mu_4 = \frac{1}{l_2}(\Psi_{m_1+1}(x) + \Psi_{m_2-1}(x), x)$$

Noting that $\Psi_{m_1+1}(x) + \Psi_{m_2-1}(x) \equiv 1$ for $x \in [l_1 + h, l_2 - h]$, so

$$\mu_3 + \mu_4 = \frac{1}{l_2}\left(\sum_{j=m_1+1}^{m_2-1} x_j\right) = \frac{l_2 - l_1 - 2h}{2l_2 h} \tag{31}$$

Substituting (30) and (31) into (29) gives

$$C_1 = -\left\{\frac{16(l_1 - l_2)}{l_2} + \frac{2l_1(l_1 - h)(2l_1 - h)}{3l_2^2 h} + \frac{(l_2 - l_1 - h)(l_1 + l_2)}{2l_2^2 h}\right\}$$

Clearly C_1 is indeed always negative when h is sufficiently small. For example, let $\bar{x} = \frac{1}{2}$, $m = 1/(2h)$ or $mh = \bar{x}$. And $l_1 = \bar{x} - h$, $l_2 = \bar{x} + h$, then $C_1 < 0$ for all $h \leqslant 1/15$.

In order to find the three remaining eigenvalues corresponding to some anti-symmetric eigenfunctions, we assume

$$\alpha_1 = -\alpha_6, \quad \alpha_2 = -\alpha_5, \quad \alpha_3 = -\alpha_4$$

Then the following holds from (20), (22) and (24):

$$\lambda\alpha_1 = 2\alpha_1 - \frac{l_1 - h}{l_2}\alpha_5 \tag{32}$$

$$\lambda\alpha_3 = 4\alpha_3 - \frac{l_1 + h}{l_2}\alpha_5 + \frac{1 - l_1 - h}{1 - l_1}\alpha_5 \tag{33}$$

$$\lambda\alpha_5 = 2\alpha_5 + \frac{1 - l_2}{1 - l_1}\alpha_5 - \mu_1\alpha_1 + \mu_2\alpha_5 + (\mu_4 - \mu_3)\alpha_3 \tag{34}$$

The three eigenvalues of this 3×3 system are also three eigenvalues of $(T + T^{\mathrm{t}})$. With these eigenvalues, three anti-symmetric eigenfunctions can then be calculated from (32) to (34). But the eigenfunctions are of no interest to us. Of the three eigenvalues, we prove that there is at least one which is negative. As for the previous symmetric case, it suffices to show that the determinant of the 3×3 coefficient matrix on the right-hand side of (32)–(34) is negative. This determinant is given by

$$C_2 = \frac{16}{l_2} - 4\frac{l_1 - h}{l_2}\mu_1 - 2(\mu_4 - \mu_3)\frac{1 - 2l_1 - 2h}{l_2} \tag{35}$$

Our calculation shows that

$$\mu_4 - \mu_3 = \frac{1}{l_2 - l_1 - 2h}\left\{\frac{(l_2 - h)l_2(2l_2 - h)}{3} - \frac{l_1(l_1 + h)(2l_1 + h)}{3} - \frac{(l_2 - l_1 - h)}{2}\right\}$$

Together with μ_1 given in (30), we have that

$$C_2 = \frac{16}{l_2} - \frac{2l_1(l_1 - h)(2l_1 - h)}{3l_2^2 h} - \frac{2(1 - 2l_1 - 2h)}{l_2(l_2 - l_1 - 2h)} \times\left\{\frac{l_2(l_2 - h)(2l_2 - h)}{3} - \frac{l_1(l_1 + h)(2l_1 + h)}{3} - \frac{(l_2 - l_1 - h)}{2}\right\} \tag{36}$$

Table I. A list of all the negative eigenvalue(s) of the symmetric part of the additive Schwarz preconditioned matrix. h is the mesh size.

h^{-1}	16	32	64	128	256	512	1024
λ_1	−0.1872	−0.6687	−1.3561	−2.3386	−0.4738	−1.2216	−2.2828
λ_2					−3.7390	−5.7292	−8.5515

Therefore, when h is small enough, C_2 is negative. We then conclude that at least one of the three eigenvalues is negative. □

Remark 3.1.
To make C_2 in (35) negative h has to be much smaller than what is required for making C_1 in (29) negative . This implies that as the mesh gets finer, the first negative eigenvalue shows up much earlier than the second negative eigenvalue.

Remark 3.2.
The theorem requires that the mesh size is sufficiently small. This does not mean h is smaller than any of the practically useful sizes. For example, in the one-dimensional space, a uniform mesh with 16 nodes would have one negative eigenvalue. The second negative eigenvalue shows up when the mesh is as fine as having 256 nodes. To be more clear, we present a numerical calculation. Here the domain is $(0,1)$ divided into two subdomains $(0,0.625)$ and $(0.375,1)$. The overlap is fixed at 25% of the size of the un-extended subdomains. All the negative eigenvalue(s) are listed in Table I. The corresponding eigenvectors are given in Figure 4. It is interesting to note that the two 'spiky' regions are all around the subdomain boundary points, and the eigenfunctions are linear in each of intervals if we remove the subdomain boundary points.

4. ANALYSIS USING THE SAAD–SCHULTZ THEORY

In this section, we investigate the convergence of AS/GMRES using the Saad–Schultz theory. We shall use the same one-dimensional example as in Section 3. We need to define two functions $\psi_{l_i}(x) \in V_h,\ i=1,2$, as follows:

$$\psi_{l_1}(x)=\begin{cases} 0 & x=0 \\ \text{linear} & x\in(0,l_1) \\ 1 & x=l_1 \\ \text{linear} & x\in(l_1,l_2) \\ 0 & x\in[l_2,1.0) \end{cases} \qquad \psi_{l_2}(x)=\begin{cases} 0 & x\in[0,l_1] \\ \text{linear} & x\in(l_1,l_2) \\ 1 & x=l_2 \\ \text{linear} & x\in(l_2,1.0) \\ 0 & x=1.0 \end{cases}$$

It is easy to see that the functions $\psi_{l_i}(x)$ have the following properties that are useful later,

$$\int_0^1 \psi_{l_1}^2(x)\,\mathrm{d}x = l_2/3, \quad \int_0^1 \psi_{l_2}^2(x)\,\mathrm{d}x = (1-l_1)/3, \quad \int_0^1 \psi_{l_1}(x)\psi_{l_2}(x)\,\mathrm{d}x = (l_2-l_1)/3$$

We first prove the following lemma

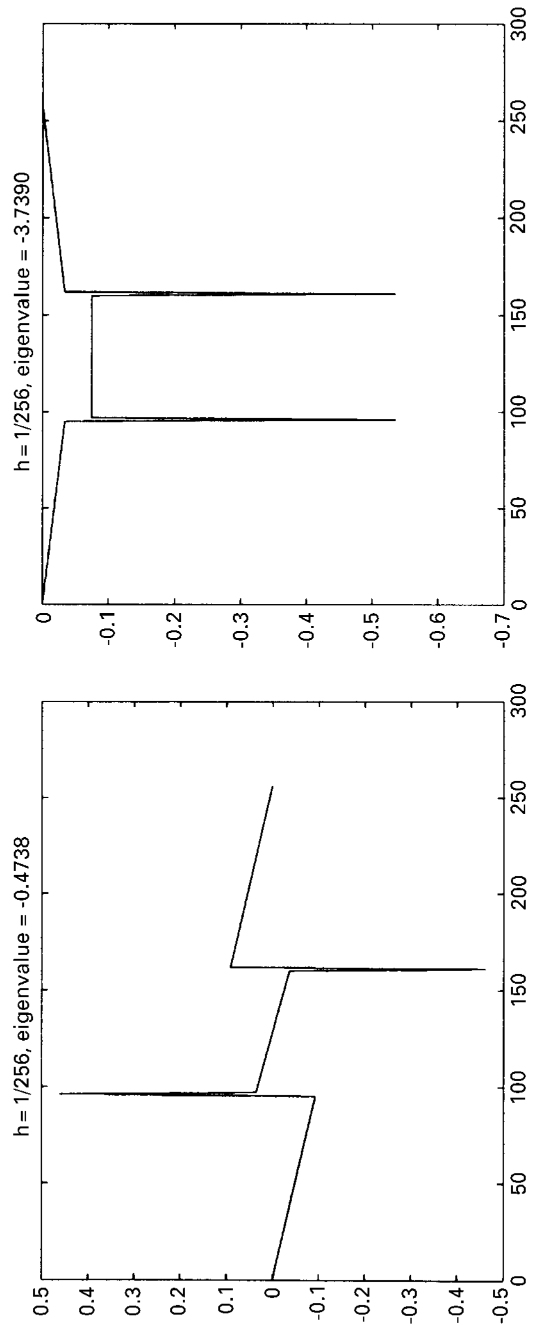

Figure 4. The eigenfunctions corresponding to the negative eigenvalues.

Lemma 4.1
For the one-dimensional Poisson problem defined in Section 3, the operator T has four distinct eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 1 - \left(\frac{l_1}{l_2}\frac{1-l_2}{1-l_1}\right)^{1/2}, \quad \lambda_4 = 1 + \left(\frac{l_1}{l_2}\frac{1-l_2}{1-l_1}\right)^{1/2}$$

with a total multiplicity n $(n = n_1^0 + n_2^0 + n_{12} + 2)$ and we have

$$T\phi_{x_i} = \lambda_1 \phi_{x_i} \quad \text{for } x_i \in (0, l_1) \cup (l_2, 1), \qquad T\psi_3 = \lambda_3 \psi_3 \tag{37}$$

$$T\phi_{x_i} = \lambda_2 \phi_{x_i} \quad \text{for } x_i \in (l_1, l_2), \qquad T\psi_4 = \lambda_4 \psi_4 \tag{38}$$

where ψ_3 and ψ_4 are given by

$$\psi_3(x) = \alpha_3(\sqrt{l_1(1-l_1)}\psi_{l_1}(x) + \sqrt{l_2(1-l_2)}\psi_{l_2}(x))$$
$$\psi_4(x) = \alpha_4(\sqrt{l_1(1-l_1)}\psi_{l_1}(x) - \sqrt{l_2(1-l_2)}\psi_{l_2}(x))$$

Moreover, all the eigenfunctions listed in (37) and (38) are linearly independent and therefore form a complete basis of V_h. We choose the constants α_3 and α_4 so that $\|\psi_3\| = \|\psi_4\| = 1$.

Proof
The eigen-relations for λ_1 and λ_2 in (37) and (38) follow immediately from the definition of T_i, and their proofs are omitted here.

We next prove the eigen-relation for λ_3. The proof for λ_4 is similar. To derive the expressions of the eigenvalue and eigenfunction λ_3 and ψ_3, we first see that for each eigenfunction ϕ_{x_i} related to λ_1 and λ_2, there exists always a basis function ϕ_{x_j} with $x_j \neq l_1, l_2$ such that

$$a(\phi_{x_i}, \phi_{x_j}) \neq 0$$

Now assume that ψ_3 is another eigenfunction which violates this condition, namely

$$a(\psi_3, \phi_{x_j}) = 0 \quad \forall x_j \neq l_1, l_2$$

This implies ψ_3 is discrete harmonic in $(0, l_1)$, (l_1, l_2) and $(l_2, 1)$ respectively, hence it is linear in each of these subintervals and can then be expressed as

$$\psi_3(x) = \psi_3(l_1)\psi_{l_1}(x) + \psi_3(l_2)\psi_{l_2}(x) \tag{39}$$

It is easy to verify that

$$a(T_1\psi_{l_2}, \psi_{x_i}) = a(\psi_{l_2}, \psi_{x_i}) = 0 \quad \forall x_i \in (0, l_1) \cup (l_1, l_2)$$

so $T_1\psi_{l_2}$ is also discrete harmonic in $(0, l_1)$ and (l_1, l_2), respectively, this indicates $T_1\psi_{l_2} = \alpha\psi_{l_1}$ for some constant α. To obtain this α, we have by definition

$$a(T_1\psi_{l_2}, \psi_{l_1}) = a(\psi_{l_2}, \psi_{l_1})$$

or $\alpha a(\psi_{l_1}, \psi_{l_1}) = a(\psi_{l_2}, \psi_{l_1})$. Then by a simple computation we derive $\alpha = -l_1/l_2$.

Similarly we have $T_2\psi_{l_1} = \beta\psi_{l_2}$ for some constant β and $\beta = -(1-l_2)/(1-l_1)$. Substituting these relations along with $T_1\psi_{l_1} = \psi_{l_1}$ and $T_1\psi_{l_2} = \psi_{l_2}$ into $T\psi_3 = \lambda_3\psi_3$ gives

$$(\psi_3(l_1) + \alpha\psi_3(l_2))\psi_{l_1}(x) + (\beta\psi_3(l_1) + \psi_3(l_2))\psi_{l_2}(x) = \lambda_3(\psi_3(l_1)\psi_{l_1}(x) + \psi_3(l_2)\psi_{l_2}(x))$$

This implies

$$(\lambda_3 - 1)\psi_3(l_1) = \alpha\psi_3(l_2), \quad (\lambda_3 - 1)\psi_3(l_2) = \beta\psi_3(l_1) \tag{40}$$

We easily get $(\lambda_3 - 1)^2 = \alpha\beta$, or equivalently

$$\lambda_3 = 1 - \left(\frac{l_1}{l_2}\frac{1-l_2}{1-l_1}\right)^{1/2}$$

Note that we take only one root here, the other is for λ_4. Using (40), we have

$$\psi_3(l_1) = \frac{\alpha}{\lambda_3 - 1}\psi_3(l_2) = \left(\frac{l_1}{l_2}\frac{1-l_2}{1-l_1}\right)^{1/2}\psi_3(l_2)$$

The expression of ψ_3 follows now from this and (39). □

Remark 4.1.
The lemma says that T has only four different eigenvalues, and according to (3) and (4), we know immediately that there exists a polynomial of degree three such that

$$\varepsilon^{(3)} = 0$$

This implies that at most three GMRES iterations are needed regardless the mesh size, the overlapping size, the starting vector and the stopping condition. Therefore the condition number of the eigenmatrix does not affect the convergence at all. However, this type of eigen distribution rarely happens in practice, and the condition number of the eigenmatrix does play a role. Hence, we will spend the rest of the section on estimating $\kappa_2(X)$.

Remark 4.2.
If the points $x = l_1$ and $x = l_2$ are symmetric with respect to the centre of the interval $(0, 1)$, then the two eigenfunctions ψ_3 and ψ_4 are orthogonal in $(\cdot,\cdot)$., i.e. $(\psi_3, \psi_4) = 0$. Otherwise ψ_3 and ψ_4 are orthogonal only in the inner product $a(\cdot,\cdot)$.

We next consider the conditioning of the eigenmatrix X consisting all eigenvectors. Let e_i be the ith unit column vector of length n. Let $1 < m_i < n$ be the integer corresponding to the node l_i, see Figure 1. The eigenmatrix has the form

$$X = [e_1, \ldots, u, \ldots, w, \ldots, e_n]$$

Here

$$u \equiv (u_1 \ldots u_{m_1} \ldots u_{m_2} \ldots, u_n)^{\mathrm{T}} \equiv (\psi_3(x_1) \ldots \psi_3(x_{m_1}) \ldots \psi_3(x_{m_2}) \ldots \psi_3(x_n))^{\mathrm{T}}$$

and

$$w \equiv (w_1 \ldots w_{m_1} \ldots w_{m_2} \ldots, w_n)^{\mathrm{T}} \equiv (\psi_4(x_1) \ldots \psi_4(x_{m_1}) \ldots \psi_4(x_{m_2}) \ldots \psi_4(x_n))^{\mathrm{T}}$$

The constants α_3 and α_4 in the expressions of Ψ_3 and Ψ_4 are so chosen that

$$\|u\|=1 \quad \text{and} \quad \|w\|=1 \tag{41}$$

We shall estimate $\|X\|$ and $\|X^{-1}\|$ separately. To bound $\|X\|$, we use the definition

$$\|X\|=\sup_{y\neq 0}\frac{\|Xy\|}{\|y\|}$$

Let $y=(y_1,\ldots,y_n)^{\mathrm{T}}\in\mathbb{R}^n$, we have

$$Xy=\sum_{j\neq m_1,m_2} y_j e_j + y_{m_1}u + y_{m_2}w$$

therefore

$$\|Xy\|^2\leqslant 3\left(\left\|\sum_{j\neq m_1,m_2} y_j e_j\right\|^2 + y_{m_1}^2\|u\|^2 + y_{m_2}^2\|w\|^2\right)=3\|y\|^2$$

That is

$$\sup_{y\neq 0}\frac{\|Xy\|}{\|y\|}\leqslant\sqrt{3}$$

On the other hand, by taking a vector $y=(y_1,\ldots,y_n)^{\mathrm{T}}\in\mathbb{R}^n$ such that $y_{m_1}=y_{m_2}=0$, we get

$$\|Xy\|^2=\sum_{j\neq m_1,m_2} y_j^2=\sum_{j=1}^{n} y_j^2=\|y\|^2$$

which implies that $\sup_{y\neq 0}\|Xy\|/\|y\|\geqslant 1$. Thus we have proved that

$$1\leqslant\|X\|\leqslant\sqrt{3} \tag{42}$$

We next turn to $\|X^{-1}\|$. It is important to have estimates of the values u_{m_1}, u_{m_2}, w_{m_1}, w_{m_2} and the corresponding determinant

$$D\equiv u_{m_1}w_{m_2}-u_{m_2}w_{m_1} \tag{43}$$

in terms of the mesh size h.

Lemma 4.2
There exist two positive constants C_0 and C_1 independent of h, such that

$$C_0 h\leqslant u_{m_1}^2, u_{m_2}^2, w_{m_1}^2, w_{m_2}^2, D\leqslant C_1 h \tag{44}$$

Proof
We need to make several observations below. From the expressions of ψ_3 and ψ_4, we know

$$u_{m_1}=\mu_0 u_{m_2}, \quad w_{m_1}=-\mu_0 w_{m_2}, \tag{45}$$

for some positive constant μ_0, independent of h. By a straightforward calculation, we obtain

$$\frac{h}{3}\leqslant\|\psi_i(x)\|_{L^2}^2\leqslant h \quad \text{for } i=3,4 \tag{46}$$

Using the linearity of ψ_3 and ψ_4 in the subintervals $(0, l_1)$, (l_1, l_2) and $(l_2, 1)$ and the identity $\|\psi_i\|_{L^2}^2 = \int_0^{l_1} \psi_i^2(x)\,\mathrm{d}x + \int_{l_1}^{l_2} \psi_i^2(x)\,\mathrm{d}x + \int_{l_2}^{1} \psi_i^2(x)\,\mathrm{d}x$, for $i = 3, 4$, we obtain

$$\beta_1(u_{m_1}^2 + u_{m_2}^2) \leqslant \|\psi_3\|_{L^2}^2 \leqslant \beta_2(u_{m_1}^2 + u_{m_2}^2) \tag{47}$$

$$\beta_1(w_{m_1}^2 + w_{m_2}^2) \leqslant \|\psi_4\|_{L^2}^2 \leqslant \beta_2(w_{m_1}^2 + w_{m_2}^2) \tag{48}$$

where β_1 and β_2 are two positive constants given by

$$\beta_1 = \tfrac{1}{12} \min\{l_1 + l_2, 2 - l_1 - l_2\} \quad \text{and} \quad \beta_2 = \tfrac{1}{4} \max\{l_1 + l_2, 2 - l_1 - l_2\}.$$

The desired proof follows immediately from (46), (45), (47), and (48). □

To estimate the norm $\|X^{-1}\| = \sup_{y \neq 0} \|X^{-1}y\|/\|y\|$, we need to form the inverse of X explicitly. We note that X has the form

$$X = I + (u - e_{m_1} \;\; w - e_{m_2}) \begin{pmatrix} e_{m_1}^{\mathrm{t}} \\ e_{m_2}^{\mathrm{t}} \end{pmatrix} \equiv I + UV$$

Using the Sherman–Morrison inverse formula (cf. Reference [22]), we have

$$X^{-1} = I - U(I + VU)^{-1}V$$

A simple calculation shows that

$$(I + VU)^{-1} = \begin{pmatrix} u_{m_1} & w_{m_1} \\ u_{m_2} & w_{m_2} \end{pmatrix}^{-1} = \frac{1}{D} \begin{pmatrix} w_{m_2} & -w_{m_1} \\ -u_{m_2} & u_{m_1} \end{pmatrix}$$

Using this relation we get

$$\begin{aligned} X^{-1} &= I - \frac{1}{D} U \begin{pmatrix} w_{m_2} e_{m-1}^{\mathrm{t}} - w_{m_1} e_{m_2}^{\mathrm{t}} \\ -u_{m_2} e_{m_1}^{\mathrm{t}} + u_{m_1} e_{m_2}^{\mathrm{t}} \end{pmatrix} \\ &= I - \frac{1}{D}((w_{m_2}(u - e_{m_1}) - u_{m_2}(w - e_{m_2}))e_{m_1}^{\mathrm{t}} \\ &\quad + (-w_{m_1}(u - e_{m_1}) + u_{m_1}(w - e_{m_2}))e_{m_2}^{\mathrm{t}}) \\ &\equiv I - \frac{1}{D}(z_{m_1} e_{m_1}^{\mathrm{t}} + z_{m_2} e_{m_2}^{\mathrm{t}}) \end{aligned}$$

Here the vectors z_{m_1} and z_{m_2} are defined in the second and third lines of the above formula. Then for any $y = (y_1 \ldots y_{m_1} \ldots y_{m_2} \ldots y_n)^{\mathrm{T}} \in \mathbb{R}^n$ and $y \neq 0$, we have

$$X^{-1}y = y - \frac{1}{D}(y_{m_1} z_{m_1} + y_{m_2} z_{m_2})$$

which implies that

$$\|X^{-1}y\|^2 \leqslant 3\left(\|y\|^2 + \frac{y_{m_1}^2}{D^2}\|z_{m_1}\|^2 + \frac{y_{m_2}^2}{D^2}\|z_{m_2}\|^2\right) \tag{49}$$

We easily see that

$$\|z_{m_1}\|^2 = \sum_{j\neq m_1,m_2} (w_{m_2}u_j - u_{m_2}w_j)^2 + (D - w_{m_2})^2 + u_{m_2}^2$$
$$\|z_{m_2}\|^2 = \sum_{j\neq m_1,m_2} (-w_{m_1}u_j + u_{m_1}w_j)^2 + w_{m_1}^2 + (D - u_{m_1})^2$$

Using the Cauchy–Schwarz inequality, we further obtain

$$\|z_{m_1}\|^2 \leqslant \sum_{j\neq m_1,m_2} (w_{m_2}^2 + u_{m_2}^2)(u_j^2 + w_j^2) + 2(D^2 + w_{m_2}^2) + u_{m_2}^2$$
$$\leqslant 4(w_{m_2}^2 + u_{m_2}^2) + 2D^2$$

Similarly, $\|z_{m_2}\|^2 \leqslant 4(u_{m_1}^2 + w_{m_1}^2) + 2D^2$. According to Lemma 4.2, we know that $\|z_{m_i}\|^2$ is of order h, i.e.

$$\|z_{m_i}\|^2 \leqslant Ch \quad \text{for } i = 1,2$$

Substituting these bounds into (49) and using the bound for D from Lemma 4.2, yields

$$\|X^{-1}y\|^2 \leqslant C\left(\|y\|^2 + \frac{1}{D}y_{m_1}^2 + \frac{1}{D}y_{m_2}^2\right) \leqslant C\frac{1}{h}\|y\|^2$$

that is $\sup_{y\neq 0} \|X^{-1}y\|/\|y\| \leqslant Ch^{-1/2}$. On the other hand, to get the lower bound, we take a vector $\tilde{y}$ such that $\tilde{y}_i = 0$ for all $i \neq m_1$, then, we have $X^{-1}\tilde{y} = \tilde{y}_{m_1}z_{m_1}/D$, and

$$\|X^{-1}\tilde{y}\|^2 = \frac{\tilde{y}_{m_1}^2}{D^2}\|z_{m_1}\|^2 \geqslant \frac{\tilde{y}_{m_1}^2}{D^2}u_{m_2}^2$$
$$\geqslant \tilde{C}\frac{1}{h}\tilde{y}_{m_1}^2 = \tilde{C}\frac{1}{h}\|\tilde{y}\|^2$$

Therefore $\sup_{y\neq 0} \|X^{-1}y\| \geqslant \|X^{-1}\tilde{y}\|/\|\tilde{y}\| \geqslant Ch^{-1/2}$ and we have proved that

$$\tilde{C}h^{-1/2} \leqslant \|X^{-1}\| \leqslant Ch^{-1/2} \tag{50}$$

Before giving our main result of this section, we need the following beautiful Lemma due to Demmel [19].

Lemma 4.3 (*Demmel*)
Let S be a non-singular matrix and of the form

$$S = [S_1 \; S_2 \cdots \; S_b] \tag{51}$$

where each S_i consists of a certain number of columns of S and these columns are orthonormal to each other. Then

$$\kappa(S) \leqslant \sqrt{b}\,\kappa(S_{\text{best}})$$

where S_{best} stands for a matrix $\tilde{S}$ which is a scaled matrix of S (i.e. multiplying each column by a real number) and has the smallest condition number among all the scaled matrices of S.

By definition, the eigenmatrix X can be re-organized to have the form

$$X = [X_1 \ \ u \ \ w]$$

where X_1 consists of all the eigenvectors ϕ_{x_i} with $i \neq m_1, m_2$. Clearly, this X satisfies the condition of Lemma 4.3. Let X_{best} be the best possible eigenmatrix associated with the additive Schwarz preconditioned matrix T, we have, by combining (42) and (50), the following result

Theorem 4.1
Let X be the eigenmatrix of the one-dimensional Poisson problem defined in Section 3, and X_{best} be an optimally scaled version of X, then we have

$$\frac{1}{\sqrt{3}}\kappa(X) \leqslant \kappa(X_{\text{best}}) \leqslant \kappa(X)$$

and

$$\tilde{C}h^{-1/2} \leqslant \kappa(X) \leqslant Ch^{-1/2}$$

for some positive constants $\tilde{C}$ and C independent of the mesh size h.

Remark 4.3.
This theorem indicates that one of the factors in (4) that controls the convergence of GMRES grows at a rate like $h^{-1/2}$ as one refines the finite element mesh.

5. CONCLUDING REMARKS

In this paper, we study the eigenproperties of the additive Schwarz preconditioned elliptic operator T in the l^2 norm. Two interesting 'bad' properties of T are discovered. (1) For a one-dimensional elliptic finite element problem, we show theoretically that the symmetric part of the additive Schwarz preconditioned operator is not positive definite. (2) For the same finite element problem, if X is the optimally scaled eigenmatrix of T, then the condition number of X is proportional to $h^{-1/2}$ as the mesh size h goes to zero. The results are completely different from what were proved by Cai and Widlund in the energy norm (i.e. $a(\cdot,\cdot)$) in Reference [8]. On the other hand the operator T obtained from this particular example has a 'good' property, namely it has only four distinct eigenvalues, but we do not believe this property can be carried over to other finite element problems.

ACKNOWLEDGEMENTS

We thank the two anonymous referees for many constructive and useful suggestions.

REFERENCES

1. Balay S, Gropp W, McInnes L, Smith B. *The Portable, Extensible Toolkit for Scientific Computing, version 2.1.0*. `http://www.mcs.anl.gov/petsc`, code and documentation, 2001.

2. Saad Y, Malevsky A. P-SPARSLIB: a portable library of distributed memory sparse iterative solvers, (version 2.15, May 1997). *Technical Report UMSI 95-180*, Minnesota Supercomputer Institute, 1995.
3. Chan T, Mathew T. Domain decomposition algorithms. *Acta Numerica* 1994; 61–143.
4. Chan T, Smith B, Zou J. Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids. *Numerische Mathematik* 1996; **73**:149–167.
5. Dryja M, Widlund O. Towards a unified theory of domain decomposition algorithms for elliptic problems. in *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Chan T, Glowinski R, Périaux J, Widlund O (eds). SIAM: Philadelphia, PA, 1990.
6. Dryja M, Widlund O. Domain decomposition algorithms with small overlap. *SIAM Journal on Scientific Computing* 1994; **15**:604–620.
7. Smith B, Bjørstad P, Gropp W. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press: Cambridge, 1996.
8. Cai X-C, Widlund O. Domain decomposition algorithms for indefinite elliptic problems. *SIAM Journal on Scientific and Statistical Computing* 1992; **13**:243–258.
9. Chan T, Zou J. A convergence theory of multilevel additive Schwarz methods on unstructured meshes. *Numerical Algorithms* 1996; **13**:365–398.
10. Cai X-C, Farhat C, Sarkis M. Schwarz methods for the unsteady compressible Navier–Stokes equations on unstructured meshes. In *Domain Decomposition Methods in Sciences and Engineering*, Glowinski R, Periaux J, Shi Z, Widlund O (eds). Wiley: New York, 1997.
11. Cai X-C, Gropp W, Keyes D. A comparison of some domain decomposition and ILU preconditioned iterative methods for nonsymmetric elliptic problems. *Numerical Linear Algebra with Applications* 1994; **1**:477–504.
12. Cai X-C, Gropp W, Keyes D, Melvin R, Young D. Parallel Newton–Krylov–Schwarz algorithms for the transonic full potential equation. *SIAM Journal on Scientific Computing* 1998; **19**:246–265.
13. Gropp W, Kaushik D, Keyes D, Smith B. High performance parallel implicit CFD. *Parallel Computing* 2001; **27**:337–362.
14. Gropp W, Keyes D, McInnes L, Tidriri M. Parallel implicit PDE computations: algorithms and software. *Proceedings of Parallel CFD'97*, Ecer A, *et al.* (eds). Manchester, UK, 1997.
15. Keyes D, Kaushik D, Smith B. Prospects for CFD on petaflops systems. In *CFD Review 1997*, Hafez M, *et al.* (eds). 1997.
16. Eisenstat S, Elman H, Schultz M. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis* 1983; **20**:345–357.
17. Saad Y, Schultz M. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**:856–869.
18. Saad Y. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Company: Boston 1995.
19. Demmel J. The condition number of equivalence transformations that block diagonalize matrix pencils. *Matrix Pencils*, Lecture Notes in Mathematics, Vol. 973. Springer: Berlin, 1983; 1–15.
20. Klawonn A, Starke G. Block triangular preconditioners for nonsymmetric saddle point problems: field-of-values analysis. *Numerische Mathematik* 1999; **81**:577–594.
21. Starke G. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numerische Mathematik* 1997; **78**:103–117.
22. Henderson H, Searle S. On deriving the inverse of a sum of matrices. *SIAM Review* 1981; **23**:53–60.

