

UNIVERSITY OF WESTMINSTER



## WestminsterResearch

<http://www.wmin.ac.uk/westminsterresearch>

### **Successive eigenvalue relaxation: a new method for the generalized eigenvalue problem and convergence estimates.**

**Evgueni Ovtchinnikov**  
**Leonidas Xanthis**

Harrow School of Computer Science

This is an electronic version of an article published in Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 457 (2006). pp. 441-451, February 2001. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences is available online at:

<http://www.journals.royalsoc.ac.uk/openurl.asp?genre=article&issn=1364-5021&volume=457&issue=2006&page=441>

---

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners. Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

---

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch.  
(<http://www.wmin.ac.uk/westminsterresearch>).

In case of abuse or copyright appearing without permission e-mail [wattsn@wmin.ac.uk](mailto:wattsn@wmin.ac.uk).

# Successive eigenvalue relaxation: a new method for the generalized eigenvalue problem and convergence estimates

BY EVGUENI E. OVTCHINNIKOV AND LEONIDAS S. XANTHIS

*Centre for Techno-Mathematics and Scientific Computing Laboratory,  
University of Westminster, London HA1 3TP, UK  
(eoo@wmin.ac.uk; lsx@wmin.ac.uk)*

*Received 24 July 2000; accepted 15 November 2000*

We present a new subspace iteration method for the efficient computation of several smallest eigenvalues of the generalized eigenvalue problem  $Au = \lambda Bu$  for symmetric positive definite operators  $A$  and  $B$ . We call this method *successive eigenvalue relaxation*, or the SER method (*homoechon* of the classical successive over-relaxation, or SOR method for linear systems). In particular, there are two significant features of SER which render it computationally attractive: (i) it can effectively deal with preconditioned large-scale eigenvalue problems, and (ii) its practical implementation does not require *any* information about the preconditioner used: it can routinely accommodate sophisticated preconditioners designed to meet more exacting requirements (e.g. three-dimensional elasticity problems with small thickness parameters). We endow SER with theoretical convergence estimates allowing for multiple and clusters of eigenvalues and illustrate their usefulness in a numerical example for a discretized partial differential equation exhibiting clusters of eigenvalues.

**Keywords:** large-scale eigenvalue problems; eigensolvers with preconditioning; subspace iteration; convergence rate; multiple and clustered eigenvalues

MNEMOSYNE AND THE CHILDREN OF MEMORY†

**Constantine Carathéodory** [1873–1950]

Mathematician towering and pre-eminent.

A chorus for his legacy:

*ὦ καλλίπυργον σοφίαν κλεινοτάτην ἐπασκῶν,  
ὡς ἡδύ σου τοῖσι λόγοις σῶφρον ἐπεστιν ἄνθος.*

—Aristophanes, *Clouds*

## Prolegomena

We are entering large, difficult ground, landmarked by such luminaries as Arnoldi, Lanczos and Wilkinson. This list could be extensive if we were to add the names of

† Title of an address delivered (12 November 1993) at the British Museum by His All Holiness The Ecumenical Patriarch Bartholomew I (it invokes the Muses with their inspired knowledge to weave a song of praise, and Mnemosyne—archetypal image of cultural and intellectual memory—to record our homage to the Mathematical Muse).

the several eminent mathematicians and computer scientists who, alongside younger talented guild members, constantly chart the field of eigenvalue computation (see, for example, Golub & van der Vorst 2000; van der Vorst & Golub 1997). Thus, we feel the need at the outset to delineate our niche in this crowded, and in spots foggy, landscape.

The present work (on eigenvalue computation) is but a clade branching off the main trunk of our research project aiming at the development of robust—with respect to certain parameters—numerical (iterative) methods for three-dimensional large-scale problems (see, for example, Ovtchinnikov & Xanthis 1998, 2000*a*). What prompted us to explore new pathways for computing large-scale eigenvalue problems was the result of an *aporia*, an impasse we reached when we tried to employ existing methods: no single method would satisfy our requirements for solid theoretical foundations (convergence estimates)—a *sine qua non* for computational reliability—and efficient implementation for the important cases of *multiple and clustered* eigenvalues.

The above class of problems is of high theoretical and practical interest and currently pursued by several researchers, employing a variety of methods, but mostly through numerical calculations, often referred to in the literature as numerical *experiments*. But, is there an epistemology of experiment in general, and numerical experiment in particular, to provide confidence in experimental results/numerical calculations? On this important issue we are on rather uncharted ground. However, to begin reflecting on the epistemic role of experiment, we draw from two of the leading scientists of 20th century physics and mathematics: the late Richard Feynman, Nobel laureate in physics, and Peter Lax, member of the American Academy of Sciences. ‘*The test of all knowledge is experiment. Experiment is the sole judge of scientific “truth”*’ Feynman (Feynman *et al.* 1963) asserts (surely referring to experiments in the physical sciences). In the context of numerical methods we take ‘truth’ to refer to the *quality* of the method used in a calculation, and experiment to refer to the numerical testing of the method. But can such an experiment be the ‘sole judge’ of (mathematical) ‘truth’? Or, invoking Peter Lax’s resonant words ‘*How much confidence can one place in a calculation?*’ (our italics) (see Lax 2000, §3). One of the salient points conveyed by Lax is that full confidence in a calculation requires that the ‘theory is well understood’ (Lax 2000). With reference to iterative methods, which underlie the present work, the Lax message echoes the need for theoretical convergence estimates, established through *rigorous* mathematical *proof*. But, for large-scale *eigenvalue* computations, such estimates (and proofs) are not easy to come by, which explains their scarcity.

In the present paper we take up this challenge. We present a new method fully supported by theoretical convergence estimates (illustrated by numerical experiments). These include the more demanding cases of multiple and clustered eigenvalues. Thus, our position in the eigenvalue computation landscape is engraved on the epistemological rock of *apodeictics* (Ovtchinnikov & Xanthis 2000*b, c*), where the (mathematical) proof reigns supreme over the extensive land of numerical experiments and *heuristics*.

## 1. Introduction

The subject of eigenvalue analysis and computation has a long and rich history. It began with the work of Carl Gustav Jacobi in 1846 and slowly progressed, until the advent of the computer accelerated its pace and fertilized its blossoms. Today

we can reap its fruits in various areas of mathematical engineering, physics, chemistry, economics and finance. Several monographs and recent review papers consummately expound the eigenvalue-computation story† (see, for example, Wilkinson 1965; Parlett 1998; Chatelin 1993; Golub & van Loan 1997; Saad 1992; D'yakonov 1996; Bertolini 1998; Golub & van der Vorst 2000; van der Vorst & Golub 1997; Bai *et al.* 2000; Knyazev 1998). These stimulating xenagogies skilfully guide the inquisitive *xenos* (stranger, visitor and guest) to the 'eigenvalue palace' through the labyrinthine corridors of algorithms and countless alleys of their modifications built over the years by a host of *daedalic*, inventive researchers. When this tour of exploration, discovery and knowledge is complete, the enlightened *xenos* will reflect that the algorithms leading to the challenging area of *large-scale* eigenvalue computations invariably pass, at some stage, through the door of *preconditioning*. Further, they will notice that many of these algorithms largely rely on numerical experiments rather than on solid theoretical foundations.

Below we will endeavour to thread our way through the various concepts mentioned above and then proceed to present a new method, with convergence estimates, for the solution of the generalized eigenvalue problem

$$Au = \lambda Bu, \quad (1.1)$$

where  $A$  and  $B$  are symmetric positive definite operators.

Large-scale eigenvalue problems arise in a plethora of important applications in various fields (e.g. quantum chemistry, vibration analysis of geomechanical and aerospace structures). As it is known, such problems cannot be efficiently addressed by the standard 'matrix transformation' methods usually employed for problems of 'small' size. Thus, to face the challenge of large-scale computations, various alternative iterative methods had to be invented: operating on vectors rather than on matrices. Typically, such methods require on each iteration the solution of a system of the form  $(L - \mu)v = g$ , where  $L$  is a linear operator whose eigenvalues and associated eigenvectors are sought. For large-scale problems the solution of such systems by *direct* methods is prohibitively expensive in terms of arithmetic operations and memory storage requirements. At the same time, standard *iterative* methods may fail to converge due to the poor conditioning of  $L$ . This is often the case in applications related to partial differential equations. Here is where *preconditioning* enters to improve convergence by using iterations of the form

$$v^{n+1} = v^n - \tau_n L_0^{-1}((L - \mu)v^n - g) \quad (1.2)$$

with a 'proper' choice of the *preconditioner*  $L_0$ .

The fact that most modern methods for solving large-scale eigenvalue problems ultimately employ iterations with preconditioning of the type (1.2) suggests the idea of applying such iterations directly to the computation of eigenvectors of  $L$  rather than of some auxiliary vectors  $v$ . This idea may be seen behind several new—or newly rediscovered—methods for solving eigenvalue problems. They are known in the literature as eigensolvers with preconditioning (or, as Bramble *et al.* (1996) and Knyazev

† In the beginning was the word *eigenvalue*, the linguistic hybrid, alluding to another (biological) hybrid: the *Minotaur*! (Was such a formidable creature in the mind of Davidson (1993) when he wrote about monster matrices and their eigenvalues?!). The intricacies involved in the computation of eigenvalues are evocative of the Cretan *Labyrinth*: the archetype of complexity. Will *preconditioning* emulate the role of the long awaited hero *Theseus*? Can modern convergence theories for preconditioned eigensolvers serve as *Ariadne's* thread?

(1998) have it, ‘preconditioned eigensolvers’). However, preconditioning in eigenvalue computations is a theme scarcely investigated and certainly less understood in comparison with the more mature area of preconditioning in linear systems (Knyazev 1998). This is evident from the fact that the preconditioned eigensolvers proposed in the literature are almost entirely supported by numerical experiments rather than justified by theoretical convergence estimates (see, for example, Davidson’s (1993) method and its variants in, for example, Bai *et al.* (2000), Golub & van der Vorst (2000), van der Vorst & Golub (1997), and references therein). A laudable achievement in the realm of convergence analysis for eigensolvers with preconditioning is the work of D’yakonov & Knyazev (1992), D’yakonov (1996), Knyazev (1987), Bramble *et al.* (1996), Bergamaschi *et al.* (1997), Neymeyr (1999) (for a more thorough xenagogy visit Knyazev 1998). Nevertheless, the theory of eigensolvers with preconditioning is still in its early stages, and most convergence results available to date address some particular cases. Prime examples are: the case when  $L_0 = L$  (Bergamaschi *et al.* 1997) and the case when  $L$  has simple eigenvalues (Bergamaschi *et al.* 1997; Bramble *et al.* 1996). To date the most comprehensive theoretical results for a preconditioned method which includes convergence estimates valid for multiple and clusters of eigenvalues (as in the present paper) are given in D’yakonov & Knyazev (1992) and D’yakonov (1996). However, the practical use of the method proposed in D’yakonov & Knyazev (1992) (see also D’yakonov 1996, ch. 9, § 6) requires knowledge of an upper estimate for the eigenvalues of  $L_0^{-1}L$  which, for certain preconditioners, may *not* be available (see, for example, the preconditioner in Ovtchinnikov & Xanthis (1998) and Ovtchinnikov & Xanthis (2000*a*) dealing with three-dimensional elasticity problems for thin structures featuring a small parameter). The same difficulty is inherent in the method of Bramble *et al.* (1996). The modifications of the methods of Bramble *et al.* (1996), D’yakonov & Knyazev (1992) and D’yakonov (1996) suggested in Knyazev (1987) and Leinen *et al.* (1997) to avoid this difficulty have not been thoroughly studied. The above limitations prompted us to explore new avenues, leading to the work reported here.

Until now we have argued the case of eigensolvers with preconditioning for tackling large-scale eigenvalue problems. At this point we wish to emphasize that the results of the present paper are not confined to the preconditioning case alone but encompass the broader area of the *generalized* eigenvalue problem (1.1).

Several preconditioned methods for finding eigenvalues of a symmetric positive definite operator  $L$  can be interpreted as iterative methods for the solution of (1.1), where  $A = L_0^{-1}L$  and  $B = L_0^{-1}$  (preconditioned eigenproblem). We note that in many applications the operators  $L$  and  $L_0$  are derived from discretizations of some differential operators, in which case  $A$  can be associated with a discretization of some bounded operator, and  $B$  with some compact operator. Further, we note that problems of such a kind also appear in various other areas unrelated to preconditioning (cf. eigenvalue problems for integral equations, and projection methods for eigenvalue problems; see, for example, the Rayleigh–Ritz method used in § 5). It is for this reason that in the method we propose here we address the generalized eigenvalue problem (1.1) rather than the (standard) eigenvalue problem for  $L$ .

In this paper we introduce a new method of the subspace iteration type for the efficient computation of several smallest eigenvalues of the generalized symmetric eigenvalue problem (1.1). We call this method *successive eigenvalue relaxation*, or the SER method (*homoechon* of the classical successive over-relaxation, or SOR

method for linear systems). We endow SER with theoretical convergence estimates which cover the case of multiple eigenvalues and eigenvalue clusters and identify the important parameters essentially affecting the convergence. Finally, we illustrate the convergence of SER numerically by solving an eigenvalue problem for an elliptic partial differential equation that features multiple eigenvalues and eigenvalue clusters.

### 2. Notation

In this paper we are concerned with the efficient numerical solution of the generalized eigenvalue problem (1.1) where  $A$  and  $B$  are linear operators in  $N$ -dimensional Euclidean space  $\mathbb{R}^N$ . We assume that  $A$  and  $B$  are symmetric positive definite in the scalar product  $[\cdot, \cdot] = (C\cdot, \cdot)$ , where  $C$  is a symmetric positive definite operator and  $(\cdot, \cdot)$  is the scalar product in  $\mathbb{R}^N$ . We denote by  $\alpha_0$  and  $\alpha^0$  two positive constants such that

$$\alpha_0[u, u] \leq [Au, u] \leq \alpha^0[u, u]. \tag{2.1}$$

We enumerate the eigenvalues of (1.1) in ascending order (starting from 0) and we assume that the first  $m \ll N$  eigenpairs (i.e. eigenvalues and corresponding eigenvectors) are of interest. The corresponding eigenvectors  $u_i$  are normalized by  $[Bu_i, u_i] = 1$ .

We denote the Rayleigh quotient for a vector  $v$  by  $\lambda(v)$  and the residual by  $r(v)$ , i.e.

$$\lambda(v) = \frac{[Av, v]}{[Bv, v]}, \quad r(v) = Av - \lambda(v)Bv.$$

The Ritz eigenvectors and eigenvalues of the problem (1.1) in a subspace  $H \subset \mathbb{R}^N$  are denoted by  $u_i(H)$  and  $\lambda_i(H)$  respectively, i.e.

$$[Au_i(H) - \lambda_i(H)Bu_i(H), v] = 0 \quad \forall v \in H. \tag{2.2}$$

Again, we enumerate the eigenvalues  $\lambda_i(H)$  in ascending order and normalize the eigenvectors  $u_i(H)$  by  $[Bu_i(H), u_i(H)] = 1$ . We also denote  $r_k(H) = r(u_k(H))$ .

### 3. The SER method

Consider the following iterative procedure. Given an approximation  $I^n$  to the invariant subspace  $I$  corresponding to the first  $m$  eigenvalues of (1.1) we choose a number  $k_n$ ,  $0 \leq k_n < m$ , and define the new approximation as

$$I^{n+1} = \text{span}\{u_i(\tilde{I}^n)\}_{i=0, m-1}, \quad \tilde{I}^n = I^n + \text{span}\{r_{k_n}(I^n)\}. \tag{3.1}$$

In other words, we add the residual on the Ritz vector  $u_{k_n}(I^n)$  to the set of Ritz vectors in  $I^n$ , then we apply the Rayleigh–Ritz projection onto the subspace  $\tilde{I}^n$  spanned by these  $m + 1$  vectors and we define the new approximation  $I^{n+1}$  to the invariant subspace  $I$  as the span of the first  $m$  Ritz eigenvectors in  $\tilde{I}^n$  (the last one is discarded so that the dimension of  $I^{n+1}$  is the same as  $I^n$ ). It is known that the residual vector  $r_{k_n}(I^n)$  is collinear to the gradient (in the scalar product  $[\cdot, \cdot]$ )

of the functional  $\lambda(u)$  at  $u = u_{k_n}(I^n)$ . Furthermore, from the minimax principle for eigenvalues (see, for example, Parlett 1998) it follows that

$$\lambda_i \leq \lambda_i(I^{n+1}) = \min_{\substack{H \subset \mathbb{R}^n \\ \dim H = m}} \lambda_i(H) \leq \lambda_i(I^n). \tag{3.2}$$

In the iterative procedure introduced here, this suggests the following interpretation of (3.1): we use the gradients of  $\lambda(u)$  for the Ritz vectors *successively* (i.e. one at a time) in order to minimize (or *relax*) the Ritz eigenvalues. In view of this property, we call the iterative procedure (3.1) the *successive eigenvalue relaxation*, or SER method.

The following lemma is pivotal to the analysis of the convergence of SER (for the proof see Ovtchinnikov & Xanthis (2000b)).

**Lemma 3.1.** *Let  $H \subset \mathbb{R}^N$  be a subspace of dimension  $k + 1$  and let  $\tilde{H} = H + \text{span}\{r_k\}$ , where  $r_k = r(u_k(H))$ . Then*

$$\sum_{i=0}^k (\lambda_i(H) - \lambda_i(\tilde{H})) \geq \frac{1}{s(H)[Ar_k, r_k]} \sum_{i=0}^k [r_i, r_k]^2, \tag{3.3}$$

where  $r_i = r(u_i(H))$  and

$$s(H) = (1 + 2s_1)(1 + s_2), \quad s_i = \frac{1}{\lambda_0^{i/2}} \sum_{j=0}^k [A^{-1}r_j, r_j]^{i/2}.$$

**Corollary 3.2.**

$$[A^{-1}r_k(H), r_k(H)] \leq s(H) \sum_{i=0}^k (\lambda_i(H) - \lambda_i).$$

**Corollary 3.3.**

$$\sum_{n=0}^{\infty} [A^{-1}r_{k_n}^n, r_{k_n}^n] \leq s_0 \frac{\alpha_0}{\alpha_0} \sum_{i=0}^{m-1} (\lambda_i(I^0) - \lambda_i), \tag{3.4}$$

where

$$s_0 = (1 + 2s_1^0)(1 + s_2^0), \quad s_i^0 = \sum_{j=0}^{m-1} \left( \frac{\lambda_j(I^0)}{\lambda_0} \right)^{3/i}.$$

Based on lemma 3.1 we immediately obtain the following convergence result.

**Theorem 3.4.** *Let  $0 \leq k < m$  and suppose that for any number  $n$  there exists a number  $\tilde{n} > n$  such that  $k_{\tilde{n}} = k$ . Then*

$$\lambda_k(I^n) \rightarrow \lambda_{\tilde{k}}, \quad \text{as } n \rightarrow \infty, \tag{3.5}$$

where  $\tilde{k} \geq k$ .

*Proof.* As we see from equation (3.2), the sequence  $\{\lambda_k(I^n)\}_{n=0,\infty}$  is decreasing and bounded from below by  $\lambda_k$ . Hence it converges to some  $\lambda \geq \lambda_k$ . From corollary 3.3 we have:

$$\sum_{n: k_n=k} [A^{-1}r_k^n, r_k^n] < \infty,$$

where the sum has infinite number of terms. Hence  $[A^{-1}r_k^n, r_k^n] \rightarrow 0$ , which implies that  $\lambda$  is an eigenvalue of (1.1). ■

**Corollary 3.5.** *If  $\lambda_{m-1}(I^0) < \lambda_m$  and the condition of theorem 3.4 holds for all  $k < m$  then in (3.5) we have  $k = k$ , and the angles between the Ritz eigenvectors  $u_k(I^n)$  and the corresponding invariant subspace converge to 0.*

### 4. Convergence estimate

The convergence rate of SER iterations depends on the choice of  $k_n$ . In this paper we present the convergence estimates for the following selection procedure:

- (i) for  $n = lm, l = 0, 1, \dots$ , we choose  $k_n = 0$  and we set  $H_0 = \text{span}\{u_0(I^n)\}$ ; and
- (ii) for  $n = lm + k, 0 < k < m$ , we choose  $k_n \leq k$  for which the angle between  $u_{k_n}(I^n)$  and  $H_{k-1}$  measured in the scalar product  $[A \cdot, \cdot]$  is maximal and we set  $H_k = H_{k-1} + \text{span}\{u_{k_n}(I^n)\}$ .

The purpose of introducing the auxiliary subspaces  $H_k$  is to ensure that on each step we ‘relax’ the Ritz eigenvalues corresponding to Ritz eigenvectors which approximate different eigenvectors of (1.1). We note that in practice we do not need to compute (and store) explicitly the subspaces  $H_k$ : it is enough to calculate approximations to the angles between  $u_i(I^n)$  and  $H_k$  which would guarantee uniform positiveness of these angles. Here we do not elaborate on these rather technical calculations (for details see Ovtchinnikov & Xanthis (2000b)).

Theorem 4.1 below gives the convergence estimate for SER iterations with the above selection procedure for  $k_n$  (for the proof see Ovtchinnikov & Xanthis (2000b)). In order to allow for multiple eigenvalues and, moreover, to treat efficiently eigenvalues which are very close to each other this estimate is formulated in terms of groups (clusters) of eigenvalues, i.e. it is independent of the distances between individual eigenvalues in a cluster. We divide the set of  $m$  smallest eigenvalues into subsets  $\sigma_j = \{\lambda_i\}_{s_j \leq i < s_{j+1}}, 1 \leq j \leq J, \lambda_{s_j} > \lambda_{s_{j-1}}, s_{J+1} = m$ , and denote by  $m_j$  the number of eigenvalues in  $\sigma_j$ , i.e.  $m_j = s_{j+1} - s_j$ , and by  $\delta_j$  (for  $j > 1$ ) the relative distance between  $\sigma_{j-1}$  and  $\sigma_j$ , i.e.  $\delta_j = (\lambda_{s_j} - \lambda_{s_{j-1}})/\lambda_{s_j}$ .

**Theorem 4.1.** *Assuming that  $k_n$  are chosen according to the above selection procedure and that  $\lambda_{m-1}(I^0) < \lambda_m$ , the following estimates are valid for the convergence of SER iterations:*

$$\sum_{i=0}^{m-1} (\lambda_i(I^{(l+1)m}) - \lambda_i) \leq \nu \sum_{i=0}^{m-1} (\lambda_i(I^{lm}) - \lambda_i), \tag{4.1}$$

$$\sum_{i=s_j}^{s_{j+1}-1} (\lambda_i(I^{(l+1)m}) - \lambda_i) \leq q_{j,l} \sum_{i=s_j}^{s_{j+1}-1} (\lambda_i(I^{lm}) - \lambda_i), \quad j = 1, \dots, J, \tag{4.2}$$



where  $q_{j,l} = \min\{1, q_j + \alpha\nu^l\}$ ,

$$q_j = 1 - \left( a_{m_j} \frac{\alpha^0}{\alpha_0} + b_{m_j} \frac{\lambda_m}{\lambda_m - \lambda_{s_j}} \right)^{-1} \frac{\lambda_m - \lambda_{s_{j+1}-1}}{\lambda_m - \lambda_{s_j}} \left( 1 - \frac{\lambda_{s_{j+1}-1}}{\lambda_m} \right). \quad (4.3)$$

$a_\mu$  and  $b_\mu$  are positive constants depending only on  $\mu$  (in particular,  $a_1 = 1$  and  $b_1 = 0$ ) and  $\alpha$  and  $\nu < 1$  are positive constants which depend on  $\alpha_0/\alpha^0$ ,  $\lambda_0$ ,  $\lambda_m$ ,  $m$ ,  $\max_j m_j$ ,  $\min_j \delta_j$  and  $1 - \lambda_{m-1}(I^0)/\lambda_m$  only.

**Remark 4.2.** To simplify our presentation of SER we omit explicit formulae for  $a_\mu$ ,  $b_\mu$ ,  $\alpha$ ,  $\rho$  and  $\nu$  from the above theorem. Instead, we explicitly identify the parameters of the problem (1.1) which essentially affect the convergence of SER. Below we demonstrate how this information can be used in practice.

## 5. Numerical example

In order to assess the benefits of the theoretical convergence estimates provided by theorem 4.1, we solve numerically an eigenvalue problem for a discretized partial differential equation exhibiting clusters of eigenvalues. In dealing with such problems it is important to ensure that the convergence of iterations is *robust*, i.e. it is not adversely affected by the discretization parameters. Below we demonstrate how theorem 4.1 can be used to verify that the convergence of SER iterations is robust.

Let  $V$  be the subspace of the Sobolev space  $H^1(\Omega)$ ,  $\Omega = (0, a) \times (0, b)$ , which contains functions vanishing at  $x = 0$  and  $x = a$ . Further, let  $w(x, y)$  be a function which is positive and bounded in  $\bar{\Omega}$ . We consider the following eigenvalue problem

$$u \in V : \mathcal{A}(u, v) = \lambda \mathcal{B}(u, v) \quad \forall v \in V \quad (5.1)$$

where

$$\mathcal{A}(u, v) = (w \nabla u, \nabla v)_\Omega, \quad \mathcal{B}(u, v) = (wu, v)_\Omega, \quad (v_1, v_2)_\Omega = \int_\Omega v_1 v_2 \, dx dy.$$

Problem (5.1) is the variational formulation of an eigenvalue problem for the second order partial differential operator  $L$  given by

$$L = -\frac{1}{w} \frac{\partial}{\partial x} w \frac{\partial}{\partial x} - \frac{1}{w} \frac{\partial}{\partial y} w \frac{\partial}{\partial y}. \quad (5.2)$$

We discretize problem (5.1) using the Rayleigh–Ritz method in the subspace  $V^{pq}$  of polynomials of degrees  $p$  and  $q$  in the variables  $x$  and  $y$ , respectively. This translates to simply replacing  $V$  by  $V^{pq}$  in (5.1). Introducing a basis  $\{v_i\}_{i=1, N_{pq}}$  in  $V^{pq}$  we eventually obtain the algebraic eigenvalue problem (1.1) where the operators  $A$  and  $B$  are represented by matrices with elements  $\mathcal{A}(v_i, v_j)$  and  $\mathcal{B}(v_i, v_j)$ , and  $C$  is the identity operator.

We take a basis in  $V^{pq}$  for which  $(\nabla v_i, \nabla v_j)_\Omega = \delta_{ij}$  (the construction details are rather technical and are omitted). Thus, if  $w_0 \leq w(x, y) \leq w^0$  on  $\bar{\Omega}$ , then one can take  $\alpha_0 = w_0$  and  $\alpha^0 = w^0$  (see § 2). From these facts, together with the fact that the eigenvalues of the discretized problem converge to those of (5.1), we conclude based on theorem 4.1 that in the present case the convergence of SER iterations is *uniform* in the discretization parameters  $p$  and  $q$ . The numerical results below confirm this conclusion.

Table 1. First 25 eigenvalues of the discretized problem (5.1) ( $p = q = 20$ )

1.0625	8.0626	13.0625	20.0626	26.0626
2.0625	9.0625	16.0625	20.0627	26.0630
4.0625	10.0625	17.0625	25.0627	29.0627
5.0625	10.0625	17.0626	25.0629	29.0631
5.0625	13.0625	18.0626	25.0629	32.0632

Table 2. Number of SER iterations versus discretization parameters

$p$	$q$				
	12	14	16	18	20
12	377	407	381	378	349
14	368	372	409	345	362
16	431	351	405	370	380
18	358	373	423	376	405
20	359	384	377	375	395

In our numerical experiments we take  $w(x, y) = e^{\alpha x}$  so that the eigenvalues of problem (5.1) are

$$\{\lambda_k\}_{k=0,\infty} = \{\lambda_{k_{ij}}\}_{i=1,\infty;j=0,\infty}, \quad \lambda_{k_{ij}} = \frac{1}{4}\alpha^2 + \left(\frac{\pi i}{a}\right)^2 + \left(\frac{\pi j}{b}\right)^2 \frac{pij}{b}. \quad (5.3)$$

Using this formula it is easy to arrange that the problem (5.1) contains clusters of eigenvalues. Specifically, we set  $a = b = \pi$  so that the problem (5.1) has double eigenvalues (e.g.  $\lambda_{k_{12}} = \lambda_{k_{21}}$ ), triple eigenvalues (e.g.  $\lambda_{k_{34}} = \lambda_{k_{43}} = \lambda_{k_{50}}$ ) and eigenvalues which are close to each other (e.g.  $\lambda_{k_{40}}, \lambda_{k_{41}}, \lambda_{k_{14}}$ ). Consequently, the discretized problem contains clusters of eigenvalues. Furthermore, the distances between eigenvalues in a cluster corresponding to a multiple eigenvalue of the differential problem generally depend on the discretization. However, the convergence estimate for SER is independent of the distances between individual eigenvalues in a cluster (see § 4). Thus, we can assert that the convergence of SER is robust even in the presence of clusters and multiple eigenvalues.

Table 1 shows the first (smallest) 25 eigenvalues of the discretized problem (5.1) with  $w(x, y) = e^{x/2}$  calculated using SER. Table 2 summarizes the convergence results for these eigenvalues. As is common in practical subspace iteration calculations, we take  $m$  slightly greater than the number of eigenpairs sought, namely  $m = 35$ , and we apply the transformation (3.1), with  $k_n$  chosen as described in § 4, only to  $u_{k_n}$  for which the relative residual

$$\sqrt{[Ar(u_{k_n}^n), r(u_{k_n}^n)]/[Au_{k_n}^n, u_{k_n}^n]}$$

is greater than 0.01 for  $k_n < 25$  and greater than 0.5 for  $k_n \geq 25$ . Table 2 shows the number of iterations (3.1) versus the discretization parameters  $p$  and  $q$ . We observe that the convergence of SER is robust, i.e. uniform in  $p$  and  $q$ .

## Epilegomena

Here we conclude with some reflections on SER. In a previous paper (Ovtchinnikov & Xanthis 2000a) we have introduced a methodology for computing the minimal eigenvalue (with theoretical convergence estimates) for three-dimensional large-scale problems for thin elastic structures (e.g. rods, plates and shells). Our efforts to extend this work to several eigenvalues (including multiple and clustered ones) based on existing eigensolvers revealed—after thorough scrutiny of the literature—the difficulties already discussed at the outset. With SER we radically overcome these difficulties, thus eloquently justifying its *raison d'être* and unique role amongst other eigensolvers (cf. Prolegomena and Introduction). Although SER is particularly attractive for preconditioned large-scale problems, our numerical example was chosen to accentuate and illustrate the fact that SER performs well also for other classes of problems unrelated to preconditioning. The full computational capabilities of SER will emerge when one undertakes extensive numerical (large-scale) computations to demonstrate its performance and relative merits compared with other methods (eigensolvers). Our current work aims at providing such a demonstration.

## References

- Bai, Z., Demmel, J., Dongarra, J., Ruhe, A. & van der Vorst, H. (eds) 2000 *Templates for the solution of algebraic eigenvalue problems: a practical guide*. Philadelphia, PA: SIAM.
- Bergamaschi, L., Gambolati, G. & Pini, G. 1997 Asymptotic convergence of conjugate gradient methods for the partial symmetric eigenproblem. *Numer. Linear Algebra Appl.* **4**, 69–84.
- Bertolini, A. F. 1998 Review of eigensolution procedures for linear dynamic finite element analysis. *Appl. Mech. Rev.* **51**, 155–171.
- Bramble, J. H., Pasciak, J. E. & Knyazev, A. V. 1996 A subspace preconditioning algorithm for eigenvector/eigenvalue computation. *Adv. Comput. Math.* **6**, 159–189.
- Chatelin, F. 1993 *Eigenvalues of matrices*. Wiley.
- Davidson, E. R. 1993 Monster matrices: their eigenvalues and eigenvectors. *Comput. Phys.* **7**, 519–522.
- D'yakonov, E. G. 1996 *Optimization in solving elliptic problems*. Boca Raton, FL: CRC Press.
- D'yakonov, E. G. & Knyazev, A. V. 1992 On an iterative method for finding lower eigenvalues. *Russ. J. Numer. Analysis Math. Model.* **7**, 473–486.
- Feynman, R. P., Leighton, R. B. & Sands, M. 1963 *The Feynman lectures on physics*, vol. 1, p. 1. Addison-Wesley.
- Golub, G. H. & van der Vorst, H. A. 2000 Eigenvalue computation in the 20th century. *J. Comput. Appl. Math.* **123**, 35–65.
- Golub, G. H. & van Loan, C. F. 1997 *Matrix computations*. Baltimore, MD: Johns Hopkins University Press.
- Knyazev, A. V. 1987 Convergence rate estimates for iterative methods for mesh symmetric eigenvalue problem. *Sov. J. Numer. Analysis Math. Model.* **2**, 371–396.
- Knyazev, A. V. 1998 Preconditioned eigensolvers—an oxymoron? *Electron. Trans. Numer. Analysis* **7**, 104–123.
- Lax, P. D. 2000 Mathematics and computing. In *Mathematics: frontiers and perspectives* (ed. V. Arnold, M. Atiyah, P. Lax & B. Mazur), pp. 417–432. International Mathematical Union, American Mathematical Society.
- Leinen, P., Lembach, W. & Neymeyr, K. 1997 An adaptive subspace method for elliptic eigenproblems with hierarchical basis preconditioning. SFB 382 report 68, University of Tübingen.

- Neymeyr, K. 1999 A geometric theory for preconditioned inverse iteration applied to a subspace. SFB 382 report 131, University of Tübingen.
- Ovtchinnikov, E. E. & Xanthis, L. S. 1998 Iterative subspace correction methods for thin elastic structures and Korn's type inequality in subspaces. *Proc. R. Soc. Lond. A* **454**, 2023–2039.
- Ovtchinnikov, E. E. & Xanthis, L. S. 2000a Effective dimensional reduction algorithm for eigenvalue problems for thin elastic structures—a paradigm in three dimensions. *Proc. Natn. Acad. Sci. USA* **97**, 967–971.
- Ovtchinnikov, E. E. & Xanthis, L. S. 2000b Successive eigenvalue relaxation. I. Theory. TM-SCLab Report 1-6-00, University of Westminster.
- Ovtchinnikov, E. E. & Xanthis, L. S. 2000c The apodeictics of SER—proofs and praxis. TM-SCLab Report 1-12-00, University of Westminster.
- Parlett, B. N. 1998 *The symmetric eigenvalue problem*. Classics In Applied Mathematics, no. 20. Philadelphia, PA: SIAM.
- Saad, Y. 1992 *Numerical methods for large eigenvalue problems*. Manchester: Manchester University Press.
- van der Vorst, H. A. & Golub, G. H. 1997 150 years old and still alive: eigenproblems. In *The state of the art in numerical analysis* (ed. I. S. Duff & G. A. Watson), pp. 93–119. Oxford: Clarendon.
- Wilkinson, J. H. 1965 *The algebraic eigenvalue problem*. Oxford University Press.

