A JACOBI–DAVIDSON TYPE METHOD FOR A RIGHT DEFINITE TWO-PARAMETER EIGENVALUE PROBLEM*

MICHIEL HOCHSTENBACH† BOR PLESTENJAK‡§

Abstract. We present a new numerical iterative method for computing selected eigenpairs of a right definite two-parameter eigenvalue problem. The method works even without good initial approximations and is able to tackle large problems that are too expensive for existing methods. The new method is similar to the Jacobi–Davidson method for the eigenvalue problem. In each step we first compute Ritz pairs of a small projected right definite two-parameter eigenvalue problem and then expand the search spaces using approximate solutions of appropriate correction equations. We present two alternatives for the correction equations, introduce a selection technique that makes it possible to compute more than one eigenpair, and give some numerical results.

Key words. Right definite two-parameter eigenvalue problem, subspace method, Jacobi–Davidson method, correction equation, Ritz pair, inexact Newton's method.

AMS subject classifications. 65F15, 15A18, 15A69.

1. Introduction. We are interested in computing one or more eigenpairs of a right definite two-parameter eigenvalue problem

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu C_1 x, \\ A_2 y &= \lambda B_2 y + \mu C_2 y, \end{aligned} \tag{1.1}$$

where A_i, B_i, and C_i are given real symmetric $n_i \times n_i$ matrices for $i = 1, 2$ and $\lambda, \mu \in \mathbb{R}$, $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$. A pair (λ, μ) is called an eigenvalue if it satisfies (1.1) for nonzero vectors x, y. The tensor product $x \otimes y$ is the corresponding eigenvector. The condition for right definiteness is that the determinant

$$\begin{vmatrix} x^T B_1 x & x^T C_1 x \\ y^T B_2 y & y^T C_2 y \end{vmatrix} \tag{1.2}$$

is strictly positive for all nonzero vectors $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$. Right definiteness and symmetry of matrices A_i, B_i, and C_i imply that there exist $n_1 n_2$ linearly independent eigenvectors for the problem (1.1) [2].

Multiparameter eigenvalue problems of this kind arise in a variety of applications [1], particularly in mathematical physics when the method of separation of variables is used to solve boundary value problems [22].

Two-parameter problems can be expressed as two coupled generalized eigenvalue problems. On the tensor product space $S := \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$ of the dimension $N := n_1 n_2$ we define matrices

$$\begin{aligned} \Delta_0 &= B_1 \otimes C_2 - C_1 \otimes B_2, \\ \Delta_1 &= A_1 \otimes C_2 - C_1 \otimes A_2, \\ \Delta_2 &= B_1 \otimes A_2 - A_1 \otimes B_2 \end{aligned} \tag{1.3}$$

*Version: January 29, 2002

†Mathematical Institute, Utrecht University, PO Box 80 010, 3508 TA Utrecht, The Netherlands.
E-MAIL: *hochstenbach@math.uu.nl*

‡IMFM/TCS, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia.
E-MAIL: *bor.plestenjak@fmf.uni-lj.si*

§Supported in part by the Ministry of Education, Science, and Sport of Slovenia (Research Project Z1-3136).

(for details on the tensor product see for example [2]). Since the tensor product of symmetric matrices is symmetric, Δ_i is a symmetric matrix for $i = 0, 1, 2$. Atkinson [2, Theorem 7.8.2] proves that right definiteness of (1.1) is equivalent to the condition that Δ_0 is positive definite. He also shows that matrices $\Delta_0^{-1}\Delta_1$ and $\Delta_0^{-1}\Delta_2$ commute and that the problem (1.1) is equivalent to the associated problem

$$\begin{aligned} \Delta_1 z &= \lambda \Delta_0 z, \\ \Delta_2 z &= \mu \Delta_0 z, \end{aligned} \tag{1.4}$$

for decomposable tensors $z \in S$, $z = x \otimes y$. The eigenvectors of (1.1) are Δ_0-orthogonal, i.e. if $x_1 \otimes y_1$ and $x_2 \otimes y_2$ are eigenvectors of (1.1) corresponding to different eigenvalues, then

$$(x_1 \otimes y_1)^T \Delta_0 (x_2 \otimes y_2) = \begin{vmatrix} x_1^T B_1 x_2 & x_1^T C_1 x_2 \\ y_1^T B_2 y_2 & y_1^T C_2 y_2 \end{vmatrix} = 0. \tag{1.5}$$

Decomposable tensors $x_i \otimes y_i$ for $i = 1, \ldots, N$ form a complete basis for S.

There exist numerical methods for right definite two-parameter eigenvalue problems. First of all, the associated problem (1.4) can be transformed in such a way that it can be solved by numerical methods for simultaneous diagonalization of commutative symmetric matrices [14, 21]. This is only feasible for problems of low dimension as the size of the matrices of the associated problem is $N \times N$. Among other methods we mention those based on Newton's method [7], the gradient method [5, 6, 8], and the Minimal Residual Quotient Iteration [4]. A deficiency of these methods is that they require initial approximations close enough to the solution in order to avoid misconvergence.

The continuation method [16, 17] overcomes problems with initial approximations but since the ordering of the eigenvalues is not necessarily preserved in a continuation step we have to compute all eigenvalues, even if we are interested only in a small portion. In this paper we introduce a new numerical method which is similar to the Jacobi–Davidson method for the one-parameter eigenvalue problem [20]. The method can be used to compute selected eigenpairs and does not need good initial approximations.

Our method computes the exterior eigenvalue (λ, μ) of (1.1) which has the maximum value of $\lambda \cos\alpha + \mu \sin\alpha$ for a given α. We also present a version that computes the interior eigenpair closest to a given pair (λ_0, μ_0), i.e. the one with minimum $(\lambda - \lambda_0)^2 + (\mu - \mu_0)^2$.

The outline of the paper is as follows. We generalize the Rayleigh–Ritz approach to right definite two-parameter eigenvalue problems in §2. In §3 we present a Jacobi–Davidson type method for right definite two-parameter eigenvalue problems and introduce two alternatives for the correction equations. We discuss how the method can be used for exterior and interior eigenvalues in §4. In §5 we present a selection technique that allows to compute more than one eigenpair. The time complexity is given in §6 and some numerical examples are presented in §7. Conclusions are summarized in §8.

2. Subspace methods and Ritz pairs. The Jacobi–Davidson method [20] is one of the subspace methods that may be used for the numerical solution of one-parameter eigenvalue problems (for an overview of subspace methods see for example [3]). The common principle of subspace methods is to compute accurate eigenpairs from low dimensional subspaces. This approach reduces computational time and memory usage and thus enables us to tackle larger problems that are too expensive for methods that work in the entire space.

A subspace method works as follows. We start with a given search subspace from which approximations for eigenpairs are computed (*extraction*). In the extraction we usually have

to solve the same type of eigenvalue problem as the original one, but of a smaller dimension. After each step we expand the subspace by a new direction (*expansion*). The idea is that as the search subspace grows, the eigenpair approximations will converge to an eigenpair of the original problem. In order to keep computation costs low, we usually do not expand the search space to the whole space. If the process does not converge in a certain number of iterations then the method is restarted with a few selected approximations as the basis of a new search space. In this section we discuss the extraction, in the next section the algorithm and the expansion.

The Rayleigh–Ritz approach defines approximations for the eigenpairs that can be extracted from the given subspace (see for instance [15]). We generalize the Rayleigh–Ritz approach for the two-parameter eigenvalue problem as follows. Suppose that the k-dimensional search subspaces $\mathcal{U}_k$ of $\mathbb{R}^{n_1}$ and $\mathcal{V}_k$ of $\mathbb{R}^{n_2}$ are represented by matrices $U_k \in \mathbb{R}^{n_1 \times k}$ and $V_k \in \mathbb{R}^{n_2 \times k}$ with orthonormal columns, respectively. The Ritz–Galerkin conditions

$$\begin{aligned}(A_1 - \sigma B_1 - \tau C_1)u &\perp \mathcal{U}_k,\\ (A_2 - \sigma B_2 - \tau C_2)v &\perp \mathcal{V}_k,\end{aligned}$$

where $u \in \mathcal{U}_k \backslash \{0\}$ and $v \in \mathcal{V}_k \backslash \{0\}$, lead to the smaller projected right definite two-parameter problem

$$\begin{aligned}U_k^T A_1 U_k c &= \sigma U_k^T B_1 U_k c + \tau U_k^T C_1 U_k c,\\ V_k^T A_2 V_k d &= \sigma V_k^T B_2 V_k d + \tau V_k^T C_2 V_k d,\end{aligned} \tag{2.1}$$

where $u = U_k c \neq 0$, $v = V_k d \neq 0$, $c, d \in \mathbb{R}^k$, and $\sigma, \tau \in \mathbb{R}$.

We say that an eigenvalue (σ, τ) of (2.1) is a *Ritz value* for the two-parameter eigenvalue problem (1.1) and subspaces $\mathcal{U}_k, \mathcal{V}_k$. If (σ, τ) is an eigenvalue of (2.1) and $c \otimes d$ is the corresponding eigenvector, then $u \otimes v$ is a *Ritz vector*, where $u = U_k c$ and $v = V_k d$. Altogether we obtain k^2 *Ritz pairs* that are approximations to the eigenpairs of (1.1). It is easy to check that if $u \otimes v$ is a Ritz vector corresponding to the Ritz value (σ, τ) then σ and τ are equal to the tensor Rayleigh quotients [16]

$$\begin{aligned}\sigma = \rho_1(u,v) &= \frac{(u\otimes v)^T \Delta_1 (u\otimes v)}{(u\otimes v)^T \Delta_0 (u\otimes v)} = \frac{(u^T A_1 u)(v^T C_2 v) - (u^T C_1 u)(v^T A_2 v)}{(u^T B_1 u)(v^T C_2 v) - (u^T C_1 u)(v^T B_2 v)},\\ \tau = \rho_2(u,v) &= \frac{(u\otimes v)^T \Delta_2 (u\otimes v)}{(u\otimes v)^T \Delta_0 (u\otimes v)} = \frac{(u^T B_1 u)(v^T A_2 v) - (u^T A_1 u)(v^T B_2 v)}{(u^T B_1 u)(v^T C_2 v) - (u^T C_1 u)(v^T B_2 v)}.\end{aligned}$$

In order to obtain Ritz values we have to solve small right definite two-parameter eigenvalue problems. For this purpose one of the available numerical methods that computes all eigenpairs of a small right definite two-parameter eigenvalue problem can be used. For instance, the associated problem (1.4) can be solved using methods for simultaneous diagonalization of two commutative symmetric matrices [14, 21].

3. Jacobi–Davidson method. The Jacobi–Davidson method [20] is a subspace method where approximate solutions of certain correction equations are used to expand the search space. Jacobi–Davidson type methods restrict the search for a new direction to the subspace that is orthogonal or oblique to the last chosen Ritz vector.

Jacobi–Davidson type methods have been successfully applied to the eigenvalue problem [20, 13], to the generalized eigenvalue problem [18], and to the singular value problem [12]. In this paper we show that a Jacobi–Davidson type method can be applied to the right definite two-parameter problem as well.

A brief sketch of the Jacobi–Davidson type method for the right definite two-parameter problem is presented in Algorithm 1. In Step 2b we have to decide which Ritz pair to select. We give details of this step in §4 where we discuss how to deal with exterior and interior eigenvalues. In Step 2e we have to find new search directions in order to expand the search subspaces. We will discuss two possible correction equations for Step 2e later in this section.

Algorithm 1

1. **Start.** Choose initial nontrivial vectors u and v.
   a) Compute $u_1 = u/\|u\|$, $v_1 = v/\|v\|$ and set $U_1 = [u_1]$, $V_1 = [v_1]$.
   b) Set $k = 1$.
2. **Iterate.** Until convergence or $k > k_{\max}$ do:
   a) Solve the projected right definite two-parameter eigenvalue problem

$$\begin{aligned} U_k^T A_1 U_k c &= \sigma U_k^T B_1 U_k c + \tau U_k^T C_1 U_k c, \\ V_k^T A_2 V_k d &= \sigma V_k^T B_2 V_k d + \tau V_k^T C_2 V_k d. \end{aligned} \tag{3.1}$$

   b) Select an appropriate Ritz value (σ, τ) and the corresponding Ritz vector $u \otimes v$, where $u = U_k c$, $v = V_k d$.
   c) Compute the residuals

$$\begin{aligned} r_1 &= (A_1 - \sigma B_1 - \tau C_1)u, \\ r_2 &= (A_2 - \sigma B_2 - \tau C_2)v. \end{aligned} \tag{3.2}$$

   d) Stop if $\rho_k \leq \epsilon$ where

$$\rho_k = (\|r_1\|^2 + \|r_2\|^2)^{1/2}. \tag{3.3}$$

   e) Compute new search directions s and t.
   f) Expand the search subspaces. Set

$$\begin{aligned} U_{k+1} &= \mathrm{RGS}(U_k, s), \\ V_{k+1} &= \mathrm{RGS}(V_k, t), \end{aligned}$$

   where RGS denotes the repeated Gram–Schmidt orthonormalization.
   g) Set $k = k + 1$.
   h) Restart. If the dimension of U_k and V_k exceeds $l_{\max}$ then replace U_k, V_k with new orthonormal bases of dimension $l_{\min}$.

To apply this algorithm we need to specify a tolerance ϵ, a maximum number of steps $k_{\max}$, a maximum dimension of the search subspaces $l_{\max}$, and a number $l_{\min} < l_{\max}$ that specifies the dimension of the search subspaces after a restart.

A larger search space involves a larger projected problem (2.1). The existing methods are able to solve only low-dimensional two-parameter problems in a reasonable time. Therefore, we expand search spaces up to the preselected dimension $l_{\max}$ and then restart the algorithm. For a restart we take the most promising $l_{\min}$ eigenvector approximations as a basis for the initial search space.

Suppose that we have computed new directions s and t for the search spaces $\mathcal{U}_{k+1}$ and $\mathcal{V}_{k+1}$, respectively. We expand the search spaces simply by adding new columns to the matrices U_k and V_k. For reasons of efficiency and stability we want orthonormal columns and therefore we

orthonormalize s against U_k and t against V_k by a stable form of the Gram-Schmidt orthonormalization.

The next theorem expresses that if the residuals (3.2) are small then the Ritz value (σ, τ) is a good approximation to an eigenvalue of (1.1). This justifies the criterion in Step 2d.

THEOREM 3.1. *If (σ, τ) is a Ritz value and r_1, r_2 are the residuals (3.2), then there exists an eigenvalue (λ, μ) of the right definite two-parameter problem (1.1) such that*

$$(\lambda - \sigma)^2 + (\mu - \tau)^2 \leq \|\Delta_0^{-1}\| \Big[(\|B_1\|\|r_2\| + \|B_2\|\|r_1\|)^2 + (\|C_1\|\|r_2\| + \|C_2\|\|r_1\|)^2 \Big]. \tag{3.4}$$

Proof. In order to prove (3.4) we consider the associated problem (1.4). First we derive a relation between the residuals (3.2) and the residuals of the associated problem. We denote

$$\begin{aligned} p_1 &= \Delta_1(u \otimes v) - \sigma \Delta_0(u \otimes v), \\ p_2 &= \Delta_2(u \otimes v) - \tau \Delta_0(u \otimes v), \end{aligned} \tag{3.5}$$

where u, v are the normalized Ritz vectors from Step 2b. From (1.3) and (3.2) it follows that

$$\begin{aligned} p_1 &= -C_1 u \otimes r_2 + r_1 \otimes C_2 v, \\ p_2 &= B_1 u \otimes r_2 - r_1 \otimes B_2 v \end{aligned}$$

and we have the bounds

$$\begin{aligned} \|p_1\| &\leq \|C_1\|\|r_2\| + \|C_2\|\|r_1\|, \\ \|p_2\| &\leq \|B_1\|\|r_2\| + \|B_2\|\|r_1\|. \end{aligned} \tag{3.6}$$

Now we return to the residuals (3.5). As Δ_0 is a symmetric positive definite matrix we can transform (3.5) into

$$\begin{aligned} \Delta_0^{-1/2} p_1 &= G_1 w - \sigma w, \\ \Delta_0^{-1/2} p_2 &= G_2 w - \tau w, \end{aligned} \tag{3.7}$$

where $w = \Delta_0^{1/2}(u \otimes v)$ and $G_i = \Delta_0^{-1/2} \Delta_i \Delta_0^{-1/2}$ for $i = 1, 2$. The matrices G_1 and G_2 are symmetric and commute because the matrices $\Delta_0^{-1}\Delta_1$ and $\Delta_0^{-1}\Delta_2$ commute. As a result there exists a common orthonormal basis of eigenvectors $w_1, \ldots, w_N$ such that

$$\begin{aligned} G_1 w_i &= \lambda_i w_i, \\ G_2 w_i &= \mu_i w_i, \end{aligned} \tag{3.8}$$

where (λ_i, μ_i), $i = 1, \ldots, N$, are the eigenvalues of (1.1). In the eigenvector basis we can decompose w as $w = \sum_{j=1}^{N} \alpha_j w_j$. From (3.7) and (3.8) we get

$$\begin{aligned} \Delta_0^{-1/2} p_1 &= \sum_{j=1}^{N} \alpha_j (\lambda_j - \sigma) w_j, \\ \Delta_0^{-1/2} p_2 &= \sum_{j=1}^{N} \alpha_j (\mu_j - \mu) w_j \end{aligned} \tag{3.9}$$

and

$$\|\Delta_0^{-1/2}p_1\|^2 + \|\Delta_0^{-1/2}p_2\|^2 = \sum_{j=1}^{N} \alpha_j^2 \Big((\lambda_j - \sigma)^2 + (\mu_j - \tau)^2 \Big).$$

Since $\sum_{j=1}^{N} \alpha_j^2 = 1$ it follows that

$$\begin{aligned} \min_{j=1,\ldots,N} \Big((\lambda_j - \sigma)^2 + (\mu_j - \tau)^2 \Big) &\leq \|\Delta_0^{-1/2}p_1\|^2 + \|\Delta_0^{-1/2}p_2\|^2 \\ &\leq \|\Delta_0^{-1}\|(\|p_1\|^2 + \|p_2\|^2). \end{aligned} \tag{3.10}$$

Finally, when we insert (3.6) into (3.10) we obtain (3.4). ∎

In the next theorem we show that if the Ritz vector $u \otimes v$ is close to an eigenvector $x \otimes y$ of problem (1.1), then the residuals r_1 and r_2 from (3.2) are of order $\mathcal{O}(\|u - x\|)$ and $\mathcal{O}(\|v - y\|)$, respectively. This shows that the criterion in Step 2d will be fulfilled if the Ritz vector $u \otimes v$ approximates an eigenvector of (1.1) well enough.

THEOREM 3.2. *Let (σ, τ) be a Ritz value of (1.1) with the corresponding Ritz vector $u \otimes v$, where u and v are normalized. If $(u+s)\otimes(v+t)$ is an eigenvector of (1.1) with the corresponding eigenvalue (λ, μ) then we can bound the error of (σ, τ) as*

$$\sqrt{(\lambda - \sigma)^2 + (\mu - \tau)^2} = \mathcal{O}(\|s\|^2 + \|t\|^2) \tag{3.11}$$

and the norm of the residuals r_1, r_2 from (3.2) as

$$\begin{aligned} \|r_1\| &\leq \|A_1 - \lambda B_1 - \mu C_1\|\|s\| + \mathcal{O}(\|s\|^2 + \|t\|^2), \\ \|r_2\| &\leq \|A_2 - \lambda B_2 - \mu C_2\|\|t\| + \mathcal{O}(\|s\|^2 + \|t\|^2). \end{aligned} \tag{3.12}$$

Proof. We write the residuals (3.2) as

$$\begin{aligned} r_1 &= -(A_1 - \lambda B_1 - \mu C_1)s + (\lambda - \sigma)B_1 u + (\mu - \tau)C_1 u, \\ r_2 &= -(A_2 - \lambda B_2 - \mu C_2)t + (\lambda - \sigma)B_2 v + (\mu - \tau)C_2 v. \end{aligned} \tag{3.13}$$

When we multiply equations (3.13) by u^T and v^T, respectively, and take into account that $u^T r_1 = v^T r_2 = 0$ then we obtain

$$\begin{bmatrix} u^T B_1 u & u^T C_1 u \\ v^T B_2 v & v^T C_2 v \end{bmatrix} \begin{bmatrix} \lambda - \sigma \\ \mu - \tau \end{bmatrix} = - \begin{bmatrix} s^T (A_1 - \lambda B_1 - \mu C_1)s \\ t^T (A_2 - \lambda B_2 - \mu C_2)t \end{bmatrix}. \tag{3.14}$$

The system (3.14) is nonsingular because of right definiteness. From (3.14) it follows that

$$\left\| \begin{bmatrix} \lambda - \sigma \\ \mu - \tau \end{bmatrix} \right\| = \left\| \begin{bmatrix} u^T B_1 u & u^T C_1 u \\ v^T B_2 v & v^T C_2 v \end{bmatrix}^{-1} \begin{bmatrix} s^T (A_1 - \lambda B_1 - \mu C_1)s \\ t^T (A_2 - \lambda B_2 - \mu C_2)t \end{bmatrix} \right\| = \mathcal{O}(\|s\|^2 + \|t\|^2)$$

and we get (3.11). The bound (3.12) is now a result of (3.13) and (3.11). ∎

In the following two subsections the expansion for our Jacobi–Davidson method is discussed. We present two alternatives for the correction equations for the right definite two-parameter eigenvalue problem. Let (σ, τ) be a Ritz value that approximates the eigenvalue (λ, μ) of (1.1) and let $u \otimes v$ be its corresponding Ritz vector. Let us assume that u and v are normalized.

3.1. Correction equations with orthogonal projections. The first alternative for the correction equations is a generalization of the approach used in [20] for the one-parameter eigenvalue problem. We are searching for orthogonal improvements of the vectors u and v of the form

$$A_1(u+s) = \lambda B_1(u+s) + \mu C_1(u+s), \tag{3.15}$$

$$A_2(v+t) = \lambda B_2(v+t) + \mu C_2(v+t), \tag{3.16}$$

where $s \perp u$ and $t \perp v$.

Let

$$r_1 = (A_1 - \sigma B_1 - \tau C_1)u,$$
$$r_2 = (A_2 - \sigma B_2 - \tau C_2)v$$

be the residuals of Ritz vector $u \otimes v$ and Ritz value (σ, τ). We can rewrite (3.15) and (3.16) as

$$(A_1 - \sigma B_1 - \tau C_1)s = -r_1 + (\lambda - \sigma)B_1 u + (\mu - \tau)C_1 u + (\lambda - \sigma)B_1 s + (\mu - \tau)C_1 s, \tag{3.17}$$

$$(A_2 - \sigma B_2 - \tau C_2)t = -r_2 + (\lambda - \sigma)B_2 v + (\mu - \tau)C_2 v + (\lambda - \sigma)B_2 t + (\mu - \tau)C_2 t. \tag{3.18}$$

In this subsection, we treat the equations (3.17) and (3.18) separately. From Theorem 3.2 it follows that $\|(\lambda - \sigma)B_1 u + (\mu - \tau)C_1 u\| = \mathcal{O}(\|s\|^2 + \|t\|^2)$. Asymptotically (i.e. when $u \otimes v$ is close to an eigenvector of (1.1)), s and t are first order corrections and $(\lambda - \sigma)B_1 u + (\mu - \tau)C_1 u$ represents some second order correction. In the same sense, the term $(\lambda - \sigma)B_1 s + (\mu - \tau)C_1 s$ can be interpreted as a third order correction.

If we ignore second and higher order terms in (3.17) then we obtain the equation

$$(A_1 - \sigma B_1 - \tau C_1)s = -r_1. \tag{3.19}$$

Because r_1 and s are orthogonal to u, we can multiply (3.19) with the orthogonal projection $(I - uu^T)$ and write $(I - uu^T)s$ instead of s. Thus we obtain the correction equation for the vector u

$$(I - uu^T)(A_1 - \sigma B_1 - \tau C_1)(I - uu^T)s = -r_1. \tag{3.20}$$

In a similar way we obtain from (3.18) the correction equation for the vector v

$$(I - vv^T)(A_2 - \sigma B_2 - \tau C_2)(I - vv^T)t = -r_2. \tag{3.21}$$

From (3.20) and (3.21) it is clear that the orthogonal projections preserve the symmetry of the matrices. Another advantage of orthogonal projections is that they are stable and easy to implement. The systems (3.20) and (3.21) for s and t are not of full rank but they are consistent. We solve them only approximatel with a Krylov subspace method with initial guess 0, for instance by a few steps of MINRES. If we do just one step of MINRES, then s and t are scalar multiples of r_1 and r_2, respectively, and then, in the sense that we expand the search spaces by the residuals, we have an Arnoldi type method, similar to the situation for the standard eigenproblem [20].

3.2. Correction equation with oblique projections. As in the correction equations with orthogonal projections we start with the equations (3.17) and (3.18). We neglect the third order correction terms $(\lambda-\sigma)B_1 s + (\mu-\tau)C_1 s$ and $(\lambda-\sigma)B_2 t + (\mu-\tau)C_2 t$, but rather then neglecting the second order terms $(\lambda-\sigma)B_1 u + (\mu-\tau)C_1 u$ and $(\lambda-\sigma)B_2 v + (\mu-\tau)C_2 v$ we project them to 0 using an oblique projection.

If we define

$$M = \begin{bmatrix} A_1 - \sigma B_1 - \tau C_1 & 0 \\ 0 & A_2 - \sigma B_2 - \tau C_2 \end{bmatrix}$$

and

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

then we can reformulate (3.17) and (3.18) (without neglected third order correction terms) as

$$M \begin{bmatrix} s \\ t \end{bmatrix} = -r + (\lambda-\sigma) \begin{bmatrix} B_1 u \\ B_2 v \end{bmatrix} + (\mu-\tau) \begin{bmatrix} C_1 u \\ C_2 v \end{bmatrix}. \tag{3.22}$$

Let $V \in \mathbb{R}^{(n_1+n_2)\times 2}$ be a matrix with columns (for reasons of stability preferably orthonormal) such that

$$\operatorname{span}(V) = \operatorname{span}\left(\begin{bmatrix} B_1 u \\ B_2 v \end{bmatrix}, \begin{bmatrix} C_1 u \\ C_2 v \end{bmatrix} \right)$$

and let $W \in \mathbb{R}^{(n_1+n_2)\times 2}$ be

$$W = \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}.$$

With the oblique projection

$$P = I - V(W^T V)^{-1} W^T$$

onto $\operatorname{span}(V)^\perp$ along $\operatorname{span}(W)$, it follows that

$$Pr = r \quad \text{and} \quad P \begin{bmatrix} B_1 u \\ B_2 v \end{bmatrix} = P \begin{bmatrix} C_1 u \\ C_2 v \end{bmatrix} = 0. \tag{3.23}$$

Therefore, from multiplying (3.22) by P we obtain

$$PM \begin{bmatrix} s \\ t \end{bmatrix} = -r. \tag{3.24}$$

Furthermore, since $s \perp u$ and $t \perp v$ it follows that

$$P \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} s \\ t \end{bmatrix} \tag{3.25}$$

and the result is the correction equation

$$PMP \begin{bmatrix} s \\ t \end{bmatrix} = -r, \tag{3.26}$$

for $s \perp u$ and $t \perp v$.

The correction equation (3.26) is again not of full rank but consistent and it is often sufficient to solve it only approximately (e.g. by a few steps of GMRES). As before, if we do one step of GMRES then s and t are scalar multiples of r_1 and r_2, respectively.

The Jacobi-Davidson method for the one-parameter problem can be viewed as an accelerated inexact Newton scheme [19]. In a similar manner we now show that there is a connection between the Jacobi–Davidson correction equation (3.26) and Newton's method for the right definite two-parameter eigenvalue problem in [16].

Eigenpairs of the two-parameter problem (1.1) are solutions of the equation

$$G(x,y,\lambda,\mu) := \begin{bmatrix} A_1x - \lambda B_1x - \mu C_1x \\ A_2y - \lambda B_2y - \mu C_2y \\ \frac{1}{2}(x^Tx - 1) \\ \frac{1}{2}(y^Ty - 1) \end{bmatrix} = 0. \tag{3.27}$$

If we apply Newton's method to (3.27) and use u, v, σ, τ with $\|u\| = \|v\| = 1$ as an initial approximation, then in order to obtain the improved approximation $u+s, v+t, \lambda, \tau$ we have to solve the system

$$\begin{bmatrix} A_1 - \sigma B_1 - \tau C_1 & 0 & -B_1u & -C_1u \\ 0 & A_2 - \sigma B_2 - \tau C_2 & -B_2v & -C_2v \\ u^T & 0 & 0 & 0 \\ 0 & v^T & 0 & 0 \end{bmatrix} \begin{bmatrix} s \\ t \\ \lambda - \sigma \\ \mu - \tau \end{bmatrix} = \begin{bmatrix} -r_1 \\ -r_2 \\ 0 \\ 0 \end{bmatrix}. \tag{3.28}$$

LEMMA 3.3. *The Jacobi–Davidson correction equation (3.26), where $s \perp u$ and $t \perp v$, is equivalent to Newton's equation (3.28). That is, if (s,t) is a solution of (3.26), then there exist unique λ, μ such that $(s, t, \lambda - \sigma, \mu - \tau)$ is a solution of (3.28), and if $(s, t, \lambda - \sigma, \mu - \tau)$ is a solution of (3.28) then (s,t) is a solution of (3.26).*

Proof. We can rewrite the equation (3.28) as

$$M \begin{bmatrix} s \\ t \end{bmatrix} = -r + (\lambda - \sigma) \begin{bmatrix} B_1u \\ B_2v \end{bmatrix} + (\mu - \tau) \begin{bmatrix} C_1u \\ C_2v \end{bmatrix}$$

and $s \perp u$, $t \perp v$, which is exactly the equation (3.22) that appears in the derivation of the Jacobi–Davidson correction equation (3.26). The proof now follows from the relations (3.23), (3.25), and the fact that $\mathrm{Ker}(P) = \mathrm{span}(V)$. ∎

This shows that the Jacobi–Davidson type method with the correction equation (3.26) is a Newton scheme, accelerated by the projection of (1.1) onto the subspace of all previous approximations. Therefore, we expect locally at least quadratic convergence of the Jacobi–Davidson method when the correction equations are solved exactly.

4. Selection of Ritz values.

In this section we present different options for the selection of Ritz values in Step 2b of Algorithm 1.

4.1. Exterior eigenvalues.

First we discuss how to obtain the eigenvalue (λ, μ) of (1.1) with the maximum value of λ. We denote such an eigenvalue by $(\lambda_{\max}, \mu_{\max})$. We show that if we select the Ritz value (σ, τ) with the maximum value of σ in each Step 2b of Algorithm 1, then the Ritz pairs will converge monotonically to an eigenpair of (1.1).

LEMMA 4.1. *Let (σ, τ) be the Ritz value for problem (1.1) and subspaces $\mathcal{U}, \mathcal{V}$ with the maximum value of σ. Then*

$$\sigma = \max_{\substack{u \in \mathcal{U},\ v \in \mathcal{V} \\ u,v \neq 0}} \frac{(u \otimes v)^T \Delta_1 (u \otimes v)}{(u \otimes v)^T \Delta_0 (u \otimes v)}. \tag{4.1}$$

Proof. Let the columns of U and V be orthonormal bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. It follows from (1.1), (1.4) and (2.1) that if (σ, τ) is a Ritz pair then σ is an eigenvalue of a symmetric definite pencil

$$(U \otimes V)^T \Delta_1 (U \otimes V) - \sigma (U \otimes V)^T \Delta_0 (U \otimes V). \tag{4.2}$$

From the Minimax Theorem [11, p. 411] it follows that

$$\sigma = \max_{\substack{w \in \mathcal{U} \otimes \mathcal{V} \\ w \neq 0}} \frac{w^T \Delta_1 w}{w^T \Delta_0 w}.$$

Since pencil (4.2) is related to the two-parameter problem (2.1) we can restrict w to a decomposable tensor $w = u \otimes v$, where $u \in \mathcal{U}$ and $v \in \mathcal{V}$. From this (4.1) follows. □

If we select the Ritz value (σ_k, τ_k) in Step 2b of Algorithm 1 with the maximum σ_k, then it follows from Lemma 4.1 that

$$\sigma_k \leq \sigma_{k+1} \leq \lambda_{\max}.$$

We can not guarantee that the eigenvalue (λ, μ) of (1.1) to which (σ_k, τ_k) converges is equal to $(\lambda_{\max}, \mu_{\max})$, but convergence to a local optimum also may happen in the Jacobi–Davidson method for the symmetric eigenproblem and in all projection methods. Our numerical examples indicate that we usually do obtain the eigenvalue with the largest value of λ.

We can use the algorithm to obtain the eigenvalue (λ, μ) of (1.1) with the maximum value of $\lambda \cos\alpha + \mu \sin\alpha$ for a given parameter α if we apply the orthogonal linear substitution

$$\begin{aligned} \lambda &= \lambda' \cos\alpha - \mu' \sin\alpha, \\ \mu &= \lambda' \sin\alpha + \mu' \cos\alpha \end{aligned}$$

to the problem (1.1). The associated two-parameter eigenproblem with this substitution is now

$$\begin{aligned} A_1 x &= \lambda'(\cos\alpha B_1 + \sin\alpha C_1)x + \mu'(-\sin\alpha B_1 + \cos\alpha C_1)x, \\ A_2 y &= \lambda'(\cos\alpha B_2 + \sin\alpha C_2)y + \mu'(-\sin\alpha B_2 + \cos\alpha C_2)y. \end{aligned} \tag{4.3}$$

The operator determinant Δ_0 remains unchanged and the substituted problem (4.3) is right definite as well. Using orthogonal linear substitutions we can thus obtain exterior eigenvalues of (1.1) in chosen directions in the (λ, μ)-plane.

4.2. Interior eigenvalues. Suppose that we are interested in the eigenvalue (λ, μ) of (1.1) closest to a specific target (λ_0, μ_0). Let us denote such an eigenvalue as $(\lambda_{\text{int}}, \mu_{\text{int}})$.

Similar to the algorithm for exterior eigenvalues we decide to select the Ritz value nearest to the target in each Step 2b of Algorithm 1. The convergence for interior Ritz values is not so nice as for the exterior ones. If a Ritz value (σ, τ) is close enough to $(\lambda_{\max}, \mu_{\max})$ then the Ritz vector corresponding to (σ, τ) is a good approximation to the eigenvector corresponding to $(\lambda_{\max}, \mu_{\max})$. On the contrary, if (σ, τ) is close to $(\lambda_{\text{int}}, \mu_{\text{int}})$ then the Ritz vector corresponding to (σ, τ) may be a poor approximation to the eigenvector corresponding to $(\lambda_{\text{int}}, \mu_{\text{int}})$, just as in the real symmetric eigenproblem.

Numerical examples in §7 show that although the convergence is very irregular, the method can still be used to compute the eigenvalue closest to the target. It turns out that for interior eigenvalues good approximations for new search directions are needed which may be obtained

with more GMRES steps for the correction equations. The number of GMRES steps is of large influence. The more steps of GMRES we take, the better updates for the approximate eigenvectors will be added to the search spaces. If we take too many steps then the method often converges to an eigenvalue $(\lambda, \mu) \neq (\lambda_{\text{int}}, \mu_{\text{int}})$. On the other hand, if we take too few GMRES steps then we need many outer iterations or we have no convergence at all.

If we are interested in interior eigenvalues of a symmetric eigenproblem $Ax = \lambda x$ then one of the possible tools are harmonic Ritz values. The question remains how to generalize harmonic Ritz values to a right definite two-parameter eigenvalue problem. We believe that any progress on this subject might lead to better methods for interior eigenvalues.

Remark. It is easy to see that Step 2b of Algorithm 1 can be modified in a similar manner if we are interested in the eigenvalue (λ, μ) of (1.1) with the maximum value of $\lambda^2 + \mu^2$.

5. Computing more eigenpairs. Suppose that we are interested in $p > 1$ eigenpairs of (1.1). In one-parameter problem various deflation techniques can be applied in order to compute more than one eigenpair. In this section we first show difficulties that are met when we try to translate standard deflation ideas from one-parameter problems to two-parameter problems. We then propose a selection method for Ritz vectors that makes it possible to obtain more than one eigenpair for two-parameter problems.

If (ξ, z) is an eigenpair of a symmetric matrix A then all other eigenpairs can be computed from the projection of A onto the subspace $z^{\perp}$. Similarly, if (λ, μ) is an eigenvalue of (1.1) and $x \otimes y$ is the corresponding eigenvector then all other eigenvectors lie in the subspace

$$(x \otimes y)^{\perp_{\Delta_0}} := \{z \in S : z^T \Delta_0 (x \otimes y) = 0\}$$

of the dimension $n_1 n_2 - 1$. By comparing the dimensions it is clear that the subspace $(x \otimes y)^{\perp_{\Delta_0}}$ can not be written as $\mathcal{U} \otimes \mathcal{V}$, where $\mathcal{U} \subset \mathbb{R}^{n_1}$ and $\mathcal{V} \subset \mathbb{R}^{n_2}$. Therefore, this kind of deflation can not be applied to Algorithm 1.

Another popular deflation of a symmetric matrix A is to use the matrix $A' = A - \xi z z^T$. Matrix A' has the same eigenvalues as matrix A except for ξ which is transformed into 0. A generalization of this approach would be to transform the two-parameter problem (1.1) into a two-parameter problem with the same eigenvalues as of (1.1) except for the eigenvalue (λ, μ) which should be transformed into $(0, 0)$. Since in a two-parameter problem there can exist eigenvalues (λ, μ) and (λ', μ') with eigenvectors $x \otimes y$ and $x' \otimes y'$, respectively, such that $(\lambda, \mu) \neq (\lambda', \mu')$ and $x = x'$, this approach would again work only if we apply the associated problem (1.4) in the tensor product space S. But, then we have to work with large Δ_i matrices and this is too expensive.

We propose the following approach. Suppose that we have already found p eigenvalues (λ_i, μ_i) and eigenvectors $x_i \otimes y_i$, $i = 1, \ldots, p$. Based on the fact that eigenvectors are Δ_0-orthogonal (see (1.5)) we adjust Algorithm 1 so that in Step 2b we consider only those Ritz vectors $u \otimes v$ which satisfy

$$|(u \otimes v)^T \Delta_0 (x_i \otimes y_i)| < \eta \text{ for } i = 1, \ldots, p \tag{5.1}$$

for an $\eta > 0$. Suppose that we are interested in eigenvalues with the maximum values of λ. Then in Step 2b we first order Ritz pairs $(\sigma_i, \tau_i), u_i \otimes v_i$ by their σ values so that $\sigma_i \geq \sigma_j$ for $i < j$ and then we select the Ritz pair that satisfies (5.1) and has the minimal index. In the case of interior eigenvalues a different ordering is used.

If none of the Ritz pairs meets (5.1) then we take the Ritz pair with index 1, but in this case the algorithm is not allowed to stop. This is achieved by a change of the stopping criterion

in Step 2d where in addition to a small residual norm (3.3) we now also require that the Ritz vector $u \otimes v$ satisfies (5.1). This guarantees that the method does not converge to the already computed eigenpairs.

The bound η should not be taken too small in order to avoid that none of the Ritz vectors is sufficiently Δ_0-orthogonal to the set of already computed eigenvectors. In numerical experiments in §7 we use

$$\eta = \frac{1}{2} \max_{i=1,\ldots,p} |(x_i \otimes y_i)^T \Delta_0 (x_i \otimes y_i)|$$

and that value successfully prevents the method from converging to the already computed eigenpairs.

All other steps of Algorithm 1 remain unchanged. Numerical results in §7 show that this approach enables us to compute more than one eigenpair.

6. Time complexity. We examine the time complexity of one outer iteration step of Algorithm 1. Let $n = n_1 = n_2$, let k be the dimension of the search spaces, and let m be the number of GMRES (MINRES) steps for a correction equation. The two steps that largely determine the time complexity are Step 2a and Step 2e. In Step 2a we first construct the smaller projected problem (3.1). We need to compute only the last row (and column) of matrices in (3.1). In the second part of Step 2a we solve (3.1) by solving its associated problem with matrices of size k^2 and thus we need $\mathcal{O}(k^6)$ [9].

First we assume that matrices A_i, B_i, and C_i are sparse. This is true in many applications, for instance when two-parameter Sturm-Liouville problems [10] are discretized. Because MINRES and GMRES are methods intended for sparse matrices the Jacobi–Davidson type method can in principle handle very large sparse problems. For such problems the time complexities of Steps 2a and 2e can be expressed as 6 MV + $\mathcal{O}(k^6)$ and $6m$ MV, respectively, where MV stands for a matrix-vector multiplication with an $n \times n$ matrix.

The analysis for dense matrices A_i, B_i, and C_i is as follows. In Step 2a we need $\mathcal{O}(n^2)$ for the construction of the smaller problem (3.1) and additional $\mathcal{O}(k^6)$ for the solution of (3.1). As in practice only very small values of k are used we can assume that $k = \mathcal{O}(n^{1/3})$ and thus the time complexity of Step 2a is $\mathcal{O}(n^2)$.

If we use correction equations (3.20), (3.21) with orthogonal projections and perform m steps of MINRES then the time complexity of Step 2e is $\mathcal{O}(mn^2)$ when we perform m matrix-vector multiplications. We obtain the same time complexity for Step 2e when we use the correction equation (3.26) with oblique projections and do m steps of GMRES. The only difference is that we are working with one matrix of size $2n$ while we are working with two matrices of size n if we use orthogonal projections.

Based on the above assumptions the time complexity of one outer step of Algorithm 1 for dense matrices is $\mathcal{O}(mn^2)$. Also important is the storage requirement. If an algorithm works with matrices A_i, B_i, and C_i as Algorithm 1 does then it requires $\mathcal{O}(n^2)$ memory. The methods that work with the associated system (1.4) need $\mathcal{O}(n^4)$ memory, which may exceed memory fast, even for modest values of n.

7. Numerical examples. We present some numerical examples obtained with Matlab 5.3. If the dimension of the matrices is $n = n_1 = n_2 = 100$ then none of the existing methods that work in the tensor product space is able to compute all eigenpairs in a reasonable time [16]. Therefore, we construct right definite two-parameter examples where the exact eigenpairs are known, which enables us to check the obtained results.

We construct our right definite two-parameter examples in the following way. We take matrices

$$A_i = Q_i F_i Q_i^T, \quad B_i = Q_i G_i Q_i^T, \quad C_i = Q_i H_i Q_i^T, \tag{7.1}$$

where F_i, G_i, and H_i are diagonal matrices and Q_i is a random orthogonal matrix for $i = 1, 2$. We select diagonal elements of matrices F_1, F_2, G_2, and H_1 as uniformly distributed random numbers from the interval $(0, 1)$ and diagonal elements of G_1 and H_2 as uniformly distributed random numbers from the interval $(1, 2)$. The determinant (1.2) is clearly strictly positive for nonzero x, y and the obtained two-parameter problem is right definite. All matrices are of dimension $n \times n$.

Let us denote $F_i = \mathrm{diag}(f_{i1}, \ldots, f_{in})$, $G_i = \mathrm{diag}(g_{i1}, \ldots, g_{in})$, and $H_i = \mathrm{diag}(h_{i1}, \ldots, h_{in})$. It is easy to see that eigenvalues of the two-parameter problem (1.1) are solutions of linear systems

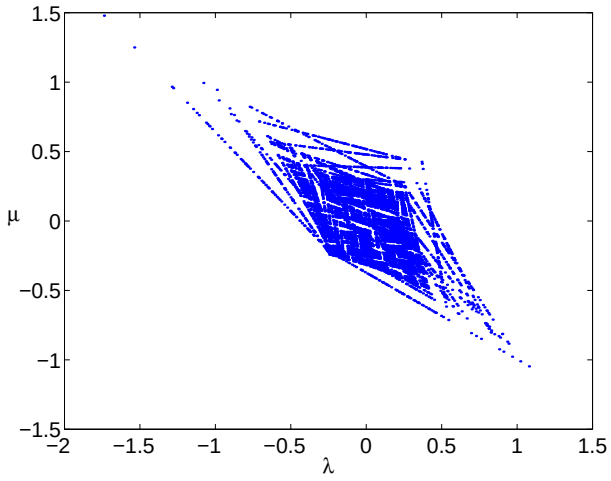
$$\begin{aligned} f_{1i} &= \lambda g_{1i} + \mu h_{1i}, \\ f_{2j} &= \lambda g_{2j} + \mu h_{2j} \end{aligned}$$

for $i, j = 1, \ldots, n$. This enables us to compute all the eigenvalues from the diagonal elements of F_i, G_i, H_i for $i = 1, 2$. In order to construct a two-parameter problem that has the point $(0, 0)$ in the interior of the convex hull of all the eigenvalues we take the shifted problem

$$\begin{aligned} (A_1 - \lambda_0 B_1 - \mu_0 C_1)x &= (\lambda - \lambda_0)B_1 x + (\mu - \mu_0)C_1 x, \\ (A_2 - \lambda_0 B_2 - \mu_0 C_2)y &= (\lambda - \lambda_0)B_2 y + (\mu - \mu_0)C_2 y, \end{aligned}$$

where the shift (λ_0, μ_0) is the arithmetic mean of all the eigenvalues. Figure 7.1 shows the distribution of eigenvalues obtained for $n = 100$.

FIG. 7.1. *Distribution of eigenvalues for a right definite two-parameter problem of size* $n = 100$.



For the following numerical examples we use GMRES instead of MINRES in the correction equation with orthogonal projections because MINRES is not standard available in Matlab 5.3.

Example 1. In the first example we use the Jacobi–Davidson type method for the exterior eigenvalues. Our goal is to compute the eigenvalue $(\lambda_{\max}, \mu_{\max})$ with the maximum value of λ.

We are interested in the number of iterations that the Jacobi–Davidson method needs for sufficiently accurate approximations and also in the percentage of the convergence to the eigenvalue $(\lambda_{\max}, \mu_{\max})$ for a test set of 250 different initial vectors.

We test both alternatives for the correction equations using various numbers of GMRES steps. Each combination is tested on the same set of 250 random initial vectors. The algorithm is restarted after every 10 iterations with the current eigenvector approximation, so $l_{\max} = 10$ and $l_{\min} = 1$. The value $\epsilon = 10^{-8}$ is used for the test of convergence and flops count in Matlab are used for a measure of time complexity.

TABLE 7.1

Statistics of the Jacobi–Davidson type method for the eigenvalue $(\lambda_{\max}, \mu_{\max})$ using different correction equations and number of GMRES steps for right definite two-parameter problems of size $n = 100$ and $n = 200$: average number of outer iterations, percentage of convergence to $(\lambda_{\max}, \mu_{\max})$, and average number of flops over 250 trials with different random initial vectors. Correction equations: JO(m) - orthogonal projections and m steps of GMRES, JS(m) - oblique projections and m steps of GMRES.

correction equation	$n = 100$			$n = 200$		
	iterations	percentage	flops	iterations	percentage	flops
JO(1)=JS(1)	105.4	100.0 %	$4.6 \cdot 10^8$	68.9	100.0 %	$3.4 \cdot 10^8$
JO(2)	50.0	100.0 %	$2.2 \cdot 10^8$	35.6	100.0 %	$2.0 \cdot 10^8$
JO(4)	26.7	100.0 %	$1.1 \cdot 10^8$	25.7	100.0 %	$1.6 \cdot 10^8$
JO(8)	23.3	99.2 %	$1.1 \cdot 10^8$	27.7	99.2 %	$2.1 \cdot 10^8$
JO(16)	25.4	30.0 %	$1.4 \cdot 10^8$	34.0	48.4 %	$3.6 \cdot 10^8$
JO(32)	29.8	38.0 %	$2.2 \cdot 10^8$	42.8	10.4 %	$7.2 \cdot 10^8$
JO(64)	33.1	28.0 %	$4.0 \cdot 10^8$	51.6	9.6 %	$16.0 \cdot 10^8$
JS(2)	96.4	100.0 %	$4.6 \cdot 10^8$	94.4	100.0 %	$6.1 \cdot 10^8$
JS(4)	99.9	100.0 %	$5.0 \cdot 10^8$	92.9	100.0 %	$6.6 \cdot 10^8$
JS(8)	63.9	100.0 %	$3.3 \cdot 10^8$	62.4	100.0 %	$5.2 \cdot 10^8$
JS(16)	45.2	94.0 %	$2.6 \cdot 10^8$	53.5	98.4 %	$6.0 \cdot 10^8$
JS(32)	41.9	82.4 %	$3.2 \cdot 10^8$	55.4	70.8 %	$9.6 \cdot 10^8$
JS(64)	39.7	66.0 %	$4.9 \cdot 10^8$	56.0	35.6 %	$17.6 \cdot 10^8$

Table 7.1 contains results obtained for $n = 100$ and $n = 200$. JO(m) and JS(m) denote that m steps of GMRES are used for the correction equation with orthogonal projections or with oblique projections, respectively. For each combination we list the average number of outer iterations for convergence, the percentage of eigenvalues that converged to the eigenvalue $(\lambda_{\max}, \mu_{\max})$, and the average number of flops in Matlab, all obtained on the same set of 250 different initial vectors.

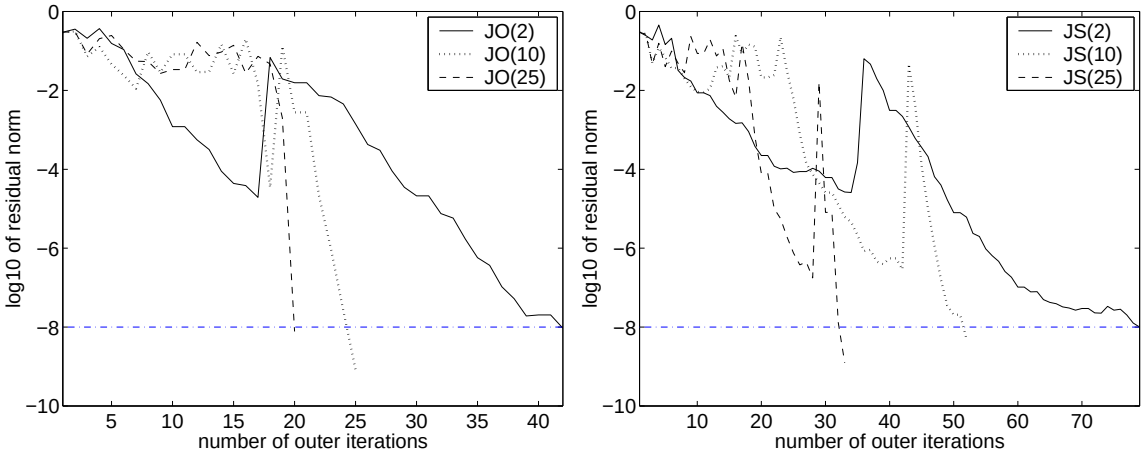
The results in Table 7.1 indicate that the method is likely to converge to an unwanted eigenvalue if we solve the correction equation too accurately, i.e. if too many GMRES steps are used to solve the correction equation. A comparison of the flops suggests that the best approach is to do a few steps of GMRES. We also see that for larger n the number of GMRES steps has more impact on the time complexity than the number of outer iterations. The reason is that for larger n the factor k^6 becomes relatively smaller compared to mn^2.

The correction equations with orthogonal projections behave similarly to the one with oblique projections but require less operations. The experiments suggest to use the correction equations with orthogonal projections in combination with a small number of GMRES steps in each outer iteration for $(\lambda_{\max}, \mu_{\max})$.

Example 2. In the second example the convergence to the exterior eigenvalue for the two-

parameter problem of dimension $n = 100$ and initial vectors $u = v = [1 \ \cdots \ 1]^T$ is examined. We compare the convergence for 2, 10, and 25 GMRES steps per iteration for the correction equation with orthogonal and the one with oblique projections, respectively. Figure 7.2 shows the $\log_{10}$ plot of residual norm ρ_k (3.3) versus the outer iteration number k. In all six cases the Ritz values converge to the eigenvalue $(\lambda_{\max}, \mu_{\max})$.

FIG. 7.2. *Convergence plot for the exterior eigenvalue* $(\lambda_{\max}, \mu_{\max})$ *for* $n = 100$ *and* $u = v = [1 \ \cdots \ 1]^T$. *The plots show the* $\log_{10}$ *of the residual norm* ρ_k *(3.3) versus the outer iteration number* k *for the Jacobi–Davidson type method for the eigenvalue* $(\lambda_{\max}, \mu_{\max})$ *using 2 (solid line), 10 (dotted line), and 25 (dashed line) GMRES steps to solve the correction equation with orthogonal projections (left plot) and oblique projections (right plot), respectively.*



It is clear from Figure 7.2 that convergence near the solution is faster if more GMRES steps are used. Experiments indicate that if only a few steps of GMRES are applied then the convergence near the solution is about linear, similar to the Jacobi–Davidson method for the standard eigenvalue problem [20, p. 419].

Example 3. In this example we examine the convergence of the Jacobi–Davidson type method for the interior eigenvalues. We look for the eigenvalue closest to $(0, 0)$. We use the same $n = 100$ two-parameter problem as in Example 1 and again test both correction equations with different number of GMRES steps on a set of 250 different initial vectors. The algorithm is restarted after every 10 iterations with the current eigenvector approximation. For the convergence test we take $\epsilon = 10^{-6}$. The reason for a more relaxed criterion is an irregular convergence of the interior eigenvalues (see the peaks in Figure 7.3).

The results, presented in Table 7.2, show that the method may also be used effectively for interior eigenvalues. In contrast to Example 1, more GMRES steps are required for one outer iteration step. If too many steps are applied then the process converges to an unwanted eigenvalue, similar to Example 1. On the other hand, if we do not take enough GMRES steps then we need many outer iteration steps and the results may be worse. This is different from Example 1 where the process converges in reasonable time even if only one GMRES step is applied per Jacobi–Davidson iteration step. The correction equation with oblique projections is more effective than the one with orthogonal projections. It is more expensive but the probability of coming close to the eigenvalue closest to $(0, 0)$ is higher.

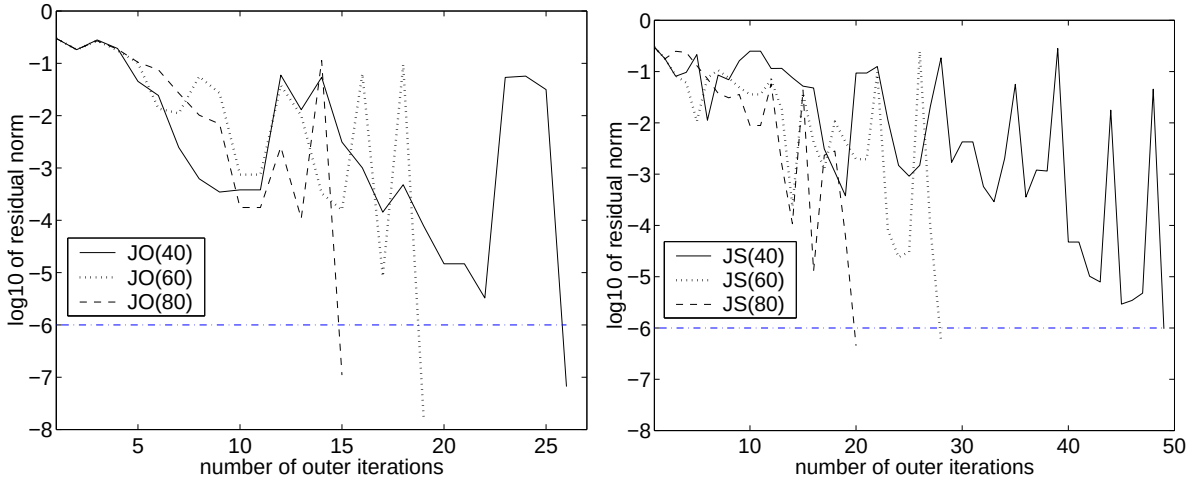
Example 4. We examine the convergence to the eigenvalue closest to $(0, 0)$ for the two-parameter problem of size $n = 100$ and initial vectors $u = v = [1 \ \cdots \ 1]^T$. Figure 7.3 shows

TABLE 7.2

Statistics of the Jacobi–Davidson type method for the eigenvalue closest to $(0,0)$ using different correction equations and different inner iteration processes for a right definite two-parameter problem of size $n = 100$: average number of iterations, percentage of convergence to the eigenvalue closest to $(0,0)$, and average number of flops over 250 trials with different random initial vectors. Correction equations: JO(m) - orthogonal projections and m steps of GMRES, JS(m) - oblique projections and m steps of GMRES.

correction equation	iterations	percentage	flops
JO(90)	15.2	80.8 %	$2.4 \cdot 10^8$
JO(80)	15.9	89.2 %	$2.2 \cdot 10^8$
JO(70)	18.9	90.0 %	$2.4 \cdot 10^8$
JO(60)	23.3	91.2 %	$2.5 \cdot 10^8$
JO(50)	32.8	79.6 %	$3.2 \cdot 10^8$
JO(40)	41.4	81.6 %	$3.5 \cdot 10^8$
JO(30)	76.5	72.8 %	$5.8 \cdot 10^8$
JO(20)	219.2	63.2 %	$14.4 \cdot 10^8$
JS(90)	20.2	92.4 %	$4.7 \cdot 10^8$
JS(80)	21.1	96.4 %	$4.3 \cdot 10^8$
JS(70)	24.2	95.6 %	$4.4 \cdot 10^8$
JS(60)	29.0	94.4 %	$4.7 \cdot 10^8$
JS(50)	38.1	93.2 %	$5.4 \cdot 10^8$
JS(40)	47.0	93.2 %	$5.7 \cdot 10^8$
JS(30)	82.9	94.0 %	$8.5 \cdot 10^8$
JS(20)	239.7	84.0 %	$20.5 \cdot 10^8$

FIG. 7.3. *Convergence plot for the eigenvalue closest to $(0,0)$ for $n = 100$ and $u = v = [1 \ \cdots \ 1]^T$. The plots show the $\log_{10}$ of the residual norm ρ_k (3.3) versus the outer iteration number k for the Jacobi–Davidson type method for the eigenvalue closest to $(0,0)$ using 40 (solid line), 60 (dotted line), and 80 (dashed line) GMRES steps to solve the correction equation with orthogonal projections (left plot) and oblique projections (right plot), respectively.*

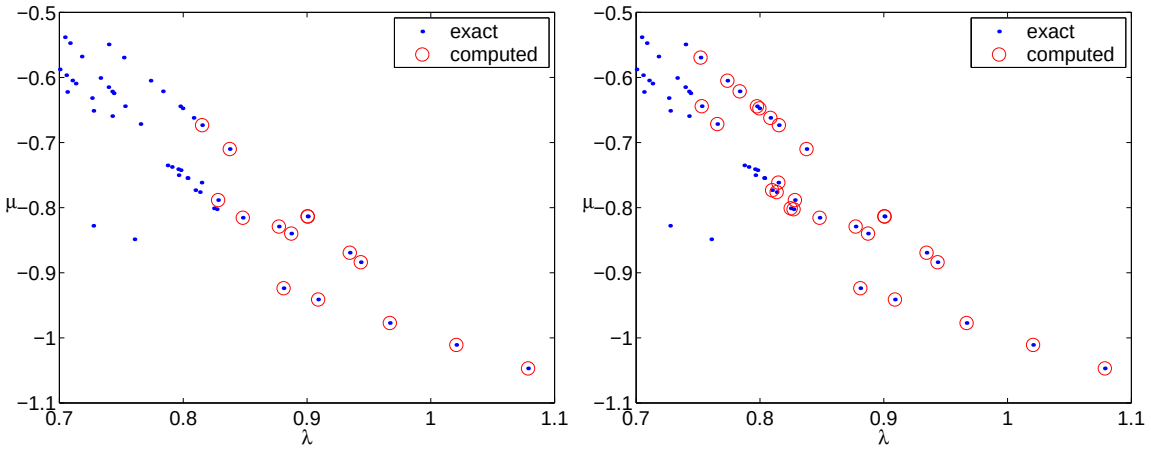


the $\log_{10}$ plot of residual norms ρ_k (3.3) versus the outer iteration number k. We compare 40, 60, and 80 GMRES steps for the correction equation with orthogonal and with oblique projections, respectively. In all six cases the Ritz values converge to the eigenvalue closest to $(0,0)$. We observe that the more GMRES steps are taken, the fewer iteration steps are needed.

The convergence is not as smooth as in Figure 7.2 for Example 2 but the algorithm is clearly useful for interior eigenvalues.

Example 5. In the last example we test the selection technique from §5 for computing more eigenpairs for the two-parameter problem of dimension $n = 100$. With 5 GMRES steps for the correction equation with orthogonal projections we try to compute 30 successive eigenvalues with the maximum value of λ. Figure 7.4 shows how well the first 15 and all 30 computed eigenvalues agree with the desired eigenvalues, respectively.

FIG. 7.4. *First 15 (left plot) and first 30 (right plot) computed eigenvalues with maximum value of λ for a two-parameter problem of size $n = 100$ computed using selection for Ritz vectors. The Jacobi–Davidson type method used 5 GMRES steps for the correction equation with orthogonal projections.*



The eigenvalues are not necessarily computed in the same order as their λ values. This explains the situation in Figure 7.4 where some eigenvalues that are in the top 30 by their λ values are not among the 30 computed eigenvalues. In order to obtain the top k eigenvalues with high probability it is therefore advisable to always compute more than k eigenvalues.

8. Conclusions. We have presented a new Jacobi–Davidson type method for a right definite two-parameter eigenvalue problem. It has several advantages over the existing methods. It can compute selected eigenpairs and it does not require good initial approximations. Probably the most important advantage is that it can tackle very large two-parameter problems, especially if the matrices A_i, B_i, and C_i are sparse.

We have proposed two correction equations. On one hand orthogonal projections are generally more stable than oblique projections and in addition, orthogonal projections preserve symmetry. On the other hand, the correction equation with oblique projections can be viewed as an inexact Newton scheme which guarantees asymptotically quadratic convergence. Numerical results indicate that the correction equation with oblique projections is more reliable but more expensive. It is therefore more suitable for the interior eigenvalues while the one with orthogonal projections may be used for the exterior eigenvalues.

Numerical results indicate that the probability of misconvergence is low when parameters are optimal. The number of GMRES steps is important. Experiments suggest to take up to 5 GMRES steps for exterior eigenvalues and more GMRES steps for interior eigenvalues. Restarts also impact the behaviour of the method. In our experiments we restart the method after every 10 iterations with the current eigenvector approximations, but a different setting may further improve the method.

Because standard deflation techniques for an one-parameter problem can not be applied to two-parameter problems, we came up with a new selection technique for Ritz vectors.

Acknowledgments. The authors are grateful to Gerard Sleijpen and Henk van der Vorst for suggestions that improved the paper.

REFERENCES

[1] F. V. Atkinson, Multiparameter spectral theory, *Bull. Amer. Math. Soc.* **74**, 1968, 1–27.
[2] ———, *Multiparameter eigenvalue problems*, Academic Press, New York, 1972.
[3] Z. Bai, J. Demmel, J .Dongarra, A. Ruhe, H. van der Vorst, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, Philadelphia, 2000.
[4] E. K. Blum, A. F. Chang, A numerical method for the solution of the double eigenvalue problem, *J. Inst. Math. Appl.*, **22**, 1978, 29–41.
[5] E. K. Blum, A. R. Curtis, A convergent gradient method for matrix eigenvector-eigentuple problems, *Numer. Math.*, **31**, 1978, 247–263.
[6] E. K. Blum, P. B. Geltner, Numerical solution of eigentuple-eigenvector problems in Hilbert spaces by a gradient method, *Numer. Math.*, **31**, 1978, 231–246.
[7] Z. Bohte, Numerical solution of some two-parameter eigenvalue problems, *Anton Kuhelj Memorial Volume*, Slov. Acad. Sci. Art., Ljubljana, 1982, 17–28.
[8] P. J. Browne, B. D. Sleeman, A numerical technique for multiparameter eigenvalue problems, *IMA J. of Num. Anal.*, **2**, 1982, 451–457.
[9] A. Bunse-Gerstner, R. Byers, V. Mehrmann, Numerical methods for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.*, **14**, 1993, 927–949.
[10] M. Faierman, *Two-Parameter Eigenvalue Problems in Ordinary Differential Equations*, volume 205 of *Pitman Research Notes in Mathematics*, Longman Scientific and Technical, Harlow, 1991.
[11] G. H. Golub, C. F. van Loan, *Matrix computations, 3rd edition*, The Johns Hopkins University Press, Baltimore, 1996.
[12] M. Hochstenbach, A Jacobi–Davidson type SVD method, *SIAM J. Sci. Comp.*, **23**, 2001, 606–628.
[13] M. Hochstenbach, G. L. G. Sleijpen, Two-sided and alternating Jacobi–Davidson, *Preprint 1196*, Dept. of Mathematics, Utrecht University, 2001. To appear in *Lin. Alg. Appl.*.
[14] X. Ji, Numerical solution of joint eigenpairs of a family of commutative matrices, *Appl. Math. Lett.*, **4**, 1991, 57–60.
[15] B. Parlett, *The symmetric eigenvalue problem*, Corrected reprint of the 1980 original, SIAM, Philadelphia, 1998.
[16] B. Plestenjak, A continuation method for a right definite two-parameter eigenvalue problem, *SIAM J. Matrix Anal. Appl.*, **21**, 2000, 1163–1184.
[17] M. Shimasaki, Homotopy algorithm for two-parameter eigenvalue problems (Japanese), *Trans. Japan SIAM*, **5**, 1995, 121–129.
[18] G. L. G. Sleijpen, A. G. L. Booten, D. R. Fokkema, H. A. van der Vorst, Jacobi–Davidson type method for generalized eigenproblems and polynomial eigenproblems, *BIT* **36**, 1996, 595-633.
[19] G. L. G. Sleijpen, H. A. van der Vorst, The Jacobi–Davidson method for eigenvalue problems and its relation with accelerated inexact Newton schemes, in *Iterative Methods in Linear Algebra, II.*, Proceedings of the Second IMACS International Symposium on Iterative Methods in Linear Algebra, June 17-20, 1995, Blagoevgrad, S. D. Margenov and P. S. Vassilevski, eds., Vol. 3 of IMACS series in Computational and Applied Mathematics, 1996, 377–389.
[20] ———, A Jacobi–Davidson iteration method for linear eigenvalue problems, *SIAM J. Matrix Anal. Appl.* **17**, 1996, 401-425.
[21] T. Slivnik, G. Tomšič, A numerical method for the solution of two-parameter eigenvalue problem, *J. Comp. Appl. Math.*, **15**, 1986, 109–115.
[22] H. Volkmer, *Multiparameter problems and expansion theorems*, vol. 1356 of *Lecture notes in mathematics*, Springer, New York, 1988.