Electronic Transactions on Numerical Analysis.
Volume 7, 1998, pp. 75-89.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

EFFICIENT EXPANSION OF SUBSPACES IN THE JACOBI-DAVIDSON METHOD FOR STANDARD AND GENERALIZED EIGENPROBLEMS*

GERARD L.G. SLEIJPEN*, HENK A. VAN DER VORST*, AND ELLEN MEIJERINK†

Abstract. We discuss approaches for an efficient handling of the correction equation in the Jacobi-Davidson method. The correction equation is effective in a subspace orthogonal to the current eigenvector approximation. The operator in the correction equation is a dense matrix, but it is composed from three factors that allow for a sparse representation. If the given matrix eigenproblem is sparse then one often aims for the construction of a preconditioner for that matrix. We discuss how to restrict this preconditioner effectively to the subspace orthogonal to the current eigenvector. The correction equation itself is formulated in terms of approximations for an eigenpair. In order to avoid misconvergence one has to make the right selection for the approximations, and this aspect will be discussed as well.

Key words. linear eigenproblems, generalized eigenproblems, Jacobi-Davidson, harmonic Ritz values, preconditioning.

AMS subject classification. 65F15.

1. Introduction. We will assume that A is an n by n matrix and we are interested in some of the eigenvalues and eigenvectors of A. The Jacobi-Davidson method [14] is based on the following two principles. The first one is to apply a Ritz-Galerkin approach for the eigenproblem $Ax = \lambda x$, with respect to some given subspace spanned by $v_1, \ldots, v_k$. The usage of other than Krylov subspaces was suggested by Davidson [1], who also suggested specific choices for the construction of orthonormal basis vectors v_j. If we define V_k as the matrix with columns v_1 up to v_k, then any vector x in the subspace can be written as $x = V_k s$, where s is a vector of length k. Approximate eigenpairs (θ, x) follow from the Ritz-Galerkin condition:

$$AV_k s - \theta V_k s \perp \{v_1, \ldots, v_k\},$$

and this leads to the reduced system

$$V_k^* A V_k s - \theta s = 0. \tag{1.1}$$

Equation (1.1) has k solutions $(\theta_j^{(k)}, s_j^{(k)})$. The k pairs $(\theta_j^{(k)}, u_j^{(k)} \equiv V_k s_j^{(k)})$ are called the Ritz values and Ritz vectors, respectively, of A with respect to the subspace spanned by the v_j. For certain choices of the v_j these Ritz pairs form suitable approximations for eigenpairs of A.

The other principle behind the Jacobi-Davidson approach goes back to Jacobi [6]. Suppose that we have an eigenvector approximation $u_j^{(k)}$ for a given eigenvalue λ. Then Jacobi suggested (in the original paper for strongly diagonally dominant symmetric matrices) to compute the orthogonal correction t for $u_j^{(k)}$ so that

$$A(u_j^{(k)} + t) = \lambda(u_j^{(k)} + t).$$

Since $t \perp u_j^{(k)}$, we can restrict ourselves to the subspace orthogonal to $u_j^{(k)}$. The operator A restricted to that subspace is given by

$$(I - u_j^{(k)} u_j^{(k)^*}) A (I - u_j^{(k)} u_j^{(k)^*}),$$

*Received February 6, 1998. Accepted for publication June 26, 1998. Recommended by R. Lehoucq.
†Mathematical Institute, Utrecht University, P.O. Box 80010, NL-3508 TA Utrecht, the Netherlands (`sleijpen@math.uu.nl, vorst@math.uu.nl, meijerin@math.uu.nl`).

and, with $r_j^{(k)} \equiv (A - \theta_j^{(k)} I) u_j^{(k)}$, we find that t satisfies the equation

$$(I - u_j^{(k)} {u_j^{(k)}}^*)(A - \lambda I)(I - u_j^{(k)} {u_j^{(k)}}^*) t = -r_j^{(k)}.$$

In practical situations we do not know λ and the obvious solution to this is to replace it by its approximation $\theta_j^{(k)}$, which leads to the *Jacobi-Davidson correction equation* for the update $t^{(k)}$:

$$(I - u_j^{(k)} {u_j^{(k)}}^*)(A - \theta_j^{(k)} I)(I - u_j^{(k)} {u_j^{(k)}}^*) t^{(k)} = -r_j^{(k)}. \tag{1.2}$$

This correction equation is often solved only approximately and its approximate solution $\tilde{t}^{(k)}$ is taken for the expansion of the subspace. This is the fundamental difference with the Krylov subspace methods; instead of selecting a subspace as powers of an operator acting on a given starting vector, we select some subspace without Krylov structure and we project the given eigenproblem onto that subspace. Any approximation technique for $\widetilde{t}^{(k)}$ is allowed in the Jacobi-Davidson framework, provided that the projectors $(I - u_j^{(k)} {u_j^{(k)}}^*)$ are taken into account.

It can be shown that the selection of the exact solution $t^{(k)}$ of (1.2) leads to quadratic convergence of the largest $\theta_j^{(k)}$ towards $\lambda_{\max}(A)$, for increasing k (similar statements can be made for the convergence towards other eigenvalues of A, provided that the Ritz values are selected appropriately in each step). The convergence is even cubic if A is symmetric.

The Jacobi-Davidson correction equation (1.2) is the key ingredient of the method and it is important to focus on how this equation can be handled. We will discuss the following aspects of solving this equation:

1. A frequently occurring situation is that one has some sparse approximation for the matrix $(A - \theta_j^{(k)} I)$. If one wishes to use that approximation as a preconditioner for the iterative solution of (1.2), then one has to restrict the preconditioner also to the subspace orthogonal to $u_j^{(k)}$. It may be not so obvious how to handle all the projections involved, but as we will see in §3, preconditioning can be implemented quite efficiently. In our discussions we will also include the generalized eigenproblem $Ax = \lambda Bx$.

2. The proper choice of the pair $(\theta_j^{(k)}, u_j^{(k)})$ is not a trivial one, especially not if one is heading for an interior eigenvalue λ. A nearby Ritz value does not necessarily represent the best choice. We will discuss this aspect in §4.

3. Finally, we will collect some reflections on the choice of preconditioners in §5. In particular, we will discuss stability effects, associated with ill-conditioning of $(A - \theta_j^{(k)} I)$ when $\theta_j^{(k)}$ is close to λ.

Before we start our discussions on these aspects, we will first describe the Jacobi-Davidson process for the generalized eigenproblem $Ax = \lambda Bx$. It turns out that much of our discussion for the standard eigenproblem carries over to the generalized eigenproblem.

2. The generalized eigenproblem. The Jacobi-Davidson approach can also be followed for computing a few selected eigenpairs of generalized eigenproblems [16, 3] of the form

$$Ax - \lambda Bx = 0. \tag{2.1}$$

Here we suggest to follow a Petrov-Galerkin method for the construction of approximate solutions. An approximate solution in a search subspace spanned by $v_1, \ldots, v_k$ is tested against a test subspace spanned by vectors $w_1, \ldots, w_k$:

$$AV_k s - \theta BV_k s \perp \{w_1, \ldots, w_k\}. \tag{2.2}$$

As before, V_k is the matrix with the vectors v_j as its columns. Likewise W_k is the matrix with columns w_j. The k-vector $s_j^{(k)}$ and the scalar $\theta_j^{(k)}$ denote a solution of the k-dimensional generalized eigenvalue problem

$$W_k^* A V_k s_j^{(k)} - \theta_j^{(k)} W_k^* B V_k s_j^{(k)} = 0. \tag{2.3}$$

The pair $(\theta_j^{(k)}, u_j^{(k)} \equiv V_k s_j^{(k)})$ will be considered as approximation for an eigenpair of A. The value $\theta_j^{(k)}$ will be called a Petrov value and $u_j^{(k)}$ is a Petrov vector.
Of course, the test subspace could have been chosen to be equal to the search subspace, but linear combinations of AV_k and BV_k, seem to be more effective; see §4 and [16, 3],

We define the residual $r_j^{(k)}$ as $r_j^{(k)} \equiv -(A - \theta_j^{(k)} B) u_j^{(k)}$. The search subspace is expanded by the solution $t^{(k)} \perp u_j^{(k)}$ of the Jacobi-Davidson correction equation

$$(I - q_j^{(k)} {q_j^{(k)}}^*)(A - \theta_j^{(k)} B)(I - u_j^{(k)} {u_j^{(k)}}^*) t^{(k)} = -r_j^{(k)}. \tag{2.4}$$

The selected Petrov vector $u_j^{(k)}$ and the "test" vector $q_j^{(k)}$ are assumed to be normalized. For fast, asymptotically quadratic convergence, this test vector should be a linear combination of $Au_j^{(k)}$ and $Bu_j^{(k)}$, and orthogonal to the residual $r_j^{(k)}$ [16, 3]. In practical situations, we solve equation (2.4) only approximately under these conditions, which leads to an approximation $\widetilde{t}^{(k)}$. Of course, we loose the asymptotical quadratic convergence in that case.

3. Preconditioning for Jacobi-Davidson.

3.1. Preconditioning for the standard eigenproblem.

If we solve the correction equation (1.2) approximately by replacing the operator $A - \theta_j^{(k)} I$ by some approximation P_k, then we obtain the following simple formula for $\widetilde{t}^{(k)}$:

$$\widetilde{t}^{(k)} = -P_k^{-1} r_k + \alpha P_k^{-1} u_j^{(k)}.$$

The value for α follows from the orthogonality constraint $\widetilde{t}^{(k)} \perp u_j^{(k)}$. The resulting formula was proposed by Olsen et al [8]. Note that we need two actions with the preconditioner for the computation of $\widetilde{t}^{(k)}$.

We will now discuss how to use preconditioning for an iterative solver for the approximate solution of equation (1.2). We can then derive 'Olsen'-like formulas for the vectors occurring in the iteration process, but, as we will see, we will need only one action with the preconditioner per iteration step. Of course, the preconditioner P_k has to be restricted to the subspace orthogonal to $u_j^{(k)}$ as well, which means that we have to work effectively with

$$\widetilde{P}_k = (I - u_j^{(k)} {u_j^{(k)}}^*) P_k (I - u_j^{(k)} {u_j^{(k)}}^*).$$

This may look quite complicated, but it is not that difficult to work with $\widetilde{P}_k$, as we will show now. We will first discuss the usage of $\widetilde{P}_k$ as a left-preconditioner.

3.1.1. Left-preconditioning.

We will assume that we apply a Krylov solver like MINRES [10], or GMRES, etc., with starting vector $t_0^{(k)} = 0$. With left-preconditioning we apply the iterative solver with the operator $\widetilde{P}_k^{-1} \widetilde{A}$, with $\widetilde{A} = (I - u_j^{(k)} {u_j^{(k)}}^*)(A - \theta_j^{(k)} I)(I - u_j^{(k)} {u_j^{(k)}}^*)$. It is easy to see that, because of the factors $(I - u_j^{(k)} {u_j^{(k)}}^*)$, all vectors in the iterative solution process will be orthogonal to $u_j^{(k)}$.

ETNA
Kent State University
etna@mcs.kent.edu

In the solution process we will have to compute the vector $z = \widetilde{P}_k^{-1}\widetilde{A}y$, for vectors y generated in the Krylov solver. This can be done in two stages. Note that $y \perp u_j^{(k)}$, which simplifies the computation in the first stage. If we define $\widetilde{y} \equiv (A - \theta_j^{(k)} I)y$, then we have that:

$$\begin{aligned}\widetilde{A}y &= (I - u_j^{(k)} {u_j^{(k)}}^*)(A - \theta_j^{(k)} I)(I - u_j^{(k)} {u_j^{(k)}}^*)y \\ &= (I - u_j^{(k)} {u_j^{(k)}}^*)(A - \theta_j^{(k)} I)y \\ &= (I - u_j^{(k)} {u_j^{(k)}}^*)\widetilde{y}.\end{aligned}$$

Then we solve z from $\widetilde{P}_k z = (I - u_j^{(k)} {u_j^{(k)}}^*)\widetilde{y}$, or, expressed in terms of P_k:

$$(I - u_j^{(k)} {u_j^{(k)}}^*)P_k(I - u_j^{(k)} {u_j^{(k)}}^*)z = (I - u_j^{(k)} {u_j^{(k)}}^*)\widetilde{y}.$$

Since z has to be orthogonal to $u_j^{(k)}$, it is not difficult to see that the above expression can be rewritten as

$$P_k z = \widetilde{y} - \alpha u_j^{(k)}. \tag{3.1}$$

Let $\bar{y}$ be the solution of $P_k \bar{y} = \widetilde{y}$, and $\bar{u}$ the solution of $P_k \bar{u} = u_j^{(k)}$. Then it follows from (3.1) that

$$z = \bar{y} - \alpha \bar{u}. \tag{3.2}$$

The orthogonality constraint on z gives the formula for α:

$$\alpha = \frac{{u_j^{(k)}}^* \bar{y}}{{u_j^{(k)}}^* \bar{u}}. \tag{3.3}$$

Since $\bar{u}$ and ${u_j^{(k)}}^* \bar{u}$ have to be computed only once at the start of the iterative solution process for (1.2), we see that each action with the preconditioned operator involves only one matrix-vector product, one action with the preconditioner, one inner product, and one vector update. This is quite surprising in view of the four projections that play a role in the formal definitions of the involved operators.

3.2. Preconditioning for the generalized eigenproblem. Although the operator in the correction equation (2.4) can be viewed as acting on the subspace $u_j^\perp$, it is defined on the whole n-dimensional space. The projection factors take care for the proper action of the operators in the $(n-1)$-dimensional subspace. The fact that the image subspace $q_j^\perp$ may differ from the original subspace $u_j^\perp$ raises another problem. Iterative linear solvers of Krylov subspace type require that the operator is defined on its image subspace as well. Krylov subspace methods, as GMRES [12] and Bi-CGSTAB methods [19, 13], subject to appropriate preconditioning, can cope with this difficulty: the preconditioner can be designed to map the image subspace to the original subspace, while the Krylov subspace solver keeps the approximate solutions of the linear system in the original subspace.

A preconditioner for the generalized problem can be implemented in a way similar to the standard eigenproblem (§3.1).

Suppose that P_k is an approximation of $A - \theta_k B$ that leads to systems $P_k x = y$ that can be solved efficiently. Since the preconditioner has to be restricted to subspaces of dimension $n-1$, we have to work effectively with (cf. §3.1)

$$\tilde{P}_k \equiv (I - q_j^{(k)}\, q_j^{(k)^*}) P_k (I - u_j^{(k)}\, u_j^{(k)^*}). \tag{3.4}$$

The solution z of the system

$$z \perp u_j^{(k)}, \qquad \tilde{P}_k z = y, \tag{3.5}$$

for $y \perp q_j^{(k)}$, can be written as

$$z = \tilde{y} - \tilde{q}\, \left(\frac{1}{\mu}\, u_j^{(k)^*} \tilde{y} \right), \tag{3.6}$$

where $\tilde{y} \equiv P_k^{-1} y$, $\tilde{q} \equiv P_k^{-1} q_j^{(k)}$, and $\mu \equiv u_j^{(k)^*} \tilde{q}$. Obviously, the same vector $\tilde{q}$ and the scalar μ can be used for different right-hand side vectors y. Note also that, for an arbitrary vector y, the vector z computed as in (3.6) satisfies the equation

$$z \perp q_j^{(k)}, \qquad \tilde{P}_k z = (I - q_j^{(k)} q_j^{(k)^*}) y. \tag{3.7}$$

This observation leads to a simplification in the computations with left-preconditioning within the Krylov solver.

An action of a preconditioned operator in a Krylov method on, say, a vector v, consists, in the case of left preconditioning, of a multiplication by the projected matrix

$$y = (I - q_j^{(k)}\, q_j^{(k)^*})(A - \theta_j^{(k)} B)(I - u_j^{(k)}\, u_j^{(k)^*}) v, \tag{3.8}$$

followed by solving the preconditioning equation (3.5). When v is orthogonal to $u_j^{(k)}$, equations (3.7) and (3.8) show that this is equivalent to a multiplication by $A - \theta_j B$, the computation of $y = (A - \theta_j B) v$, followed by the computation of z as in (3.6). Here, we used that $(I - q_j^{(k)}\, q_j^{(k)^*})(I - q_j^{(k)}\, q_j^{(k)^*}) = (I - q_j^{(k)}\, q_j^{(k)^*})$. Since the result vector z is orthogonal to $u_j^{(k)}$, we see that the Krylov subspace for the linear solver and the preconditioned operator can be built with multiplications by the matrix $A - \theta_j B$ itself (no projections!) and preconditioning steps as in (3.6), provided that the initial guess is orthogonal to $u_j^{(k)}$. The preconditioning step (3.6) can be coded as a standard preconditioning "solve $P_k \tilde{y} = y$ ", followed by a skew projection " $z = (I - \tilde{q} \frac{1}{\mu} u_j^{(k)^*}) \tilde{y}$ ".

Note that the choice $P_k = I$ provides on operator that maps the image subspace of the operator in (2.4) back to the original subspace and this is necessary if one has no other preconditioner available for an iterative solution process.

3.2.1. Right preconditioning. In practice, it may be inefficient to solve the correction equation accurately. Often a moderate accuracy already leads to efficient performance of the Jacobi-Davidson method. If the solution of the correction equation is more accurate, then it may be expected to form a better expansion vector for the search subspace and to lead to accurate eigenvector approximations for smaller k. It is usually unknown what degree of accuracy leads to the most efficient overall performance.

The Jacobi-Davidson method can be viewed as an accelerated inexact Newton method (see [15]). For inexact Newton methods there are some guidelines in the literature (cf., e.g.,

ETNA
Kent State University
etna@mcs.kent.edu

[2]) for effective reduction factors for the residuals in the Newton steps. The suggestion to solve the Jacobian system $DF(x_k)h = r_k$ in step k of the Newton method with a residual reduction by a factor 2^{-k} (that is, $\|r_k - DF(x_k)\tilde{h}\|_2/\|r_k\|_2 \le 2^{-k}$, $x_{k+1} = x_k + \tilde{h}$) seems to work well also for the Jacobi-Davidson method [3][1] Since residuals for a a right preconditioned system are also residuals for the unpreconditioned system, the right preconditioning seems to be more suitable for the "2^{-k}-criterion".

If $\tilde{t}$ solves

$$(A - \theta B)\tilde{P}_k^\dagger \tilde{t} = -r_k, \quad \text{where} \quad \tilde{P}_k^\dagger \equiv \left(I - \tilde{q}\frac{1}{\mu}u_j^{(k)^*}\right) P_k^{-1} \tag{3.9}$$

(cf. (3.7) and (3.6)), then orthogonality of r_k with respect to $q_j^{(k)}$ implies that $t = \tilde{P}_k^\dagger \tilde{t}$ solves (2.4). Also in this case of right preconditioning, the preconditioned operator can be written as the product of a standard preconditioning step, a skew projection, and a multiplication by the matrix $A - \theta_j B$. However, since the image of the operator in (3.9) needs not to be orthogonal to $q_j^{(k)}$, (3.9) may be inconsistent. Therefore, an iterative solution of (2.4) with right instead of left preconditioning requires an additional post projection with $I - q_j^{(k)} q_j^{(k)^*}$ in each step of the iterative method.

3.2.2. Preconditioning for a deflated eigenproblem. If more than one eigenvector is wanted, then the target can be re-selected after detection of an eigenpair. Especially when nearby eigenvalues are computed, deflation with the detected eigenvectors is recommended, since this improves the performance of the method considerably [3]. For stability and efficiency reasons, eigenpairs for (2.1) are computed via partial Schur decompositions. Deflation is also based on these decompositions.

A *partial Schur decomposition for the pencil* A, B is of the form

$$AQ_\ell = Z_\ell S, \qquad BQ_\ell = Z_\ell T, \tag{3.10}$$

with Q_ℓ and Z_ℓ orthonormal n by ℓ matrices and S and T upper triangular ℓ by ℓ matrices. Eigenpairs for the ℓ-dimensional generalized eigenproblem $Sy - \lambda Ty = 0$ yield eigenpairs for the original n-dimensional eigenproblem (2.1).

Suppose that a partial Schur decomposition (3.10) for a value of ℓ has been computed already. Then the next Schur pair u, q satisfies

$$A\left[\, Q_\ell \; u \,\right] = \left[\, Z_\ell \; q \,\right] \begin{bmatrix} S & a \\ 0^* & \theta_s \end{bmatrix}, \qquad B\left[\, Q_\ell \; u \,\right] = \left[\, Z_\ell \; q \,\right] \begin{bmatrix} T & b \\ 0^* & \theta_c \end{bmatrix}, \tag{3.11}$$

and (see [3]) u is precisely the solution of

$$u \perp Q_\ell, \quad \theta_c(I - Z_\ell Z_\ell^*)A(I - Q_\ell Q_\ell^*)u - \theta_s(I - Z_\ell Z_\ell^*)B(I - Q_\ell Q_\ell^*)u = 0. \tag{3.12}$$

In other words, $(\theta_s/\theta_c, u)$ is an eigenpair of the deflated system.

The Jacobi-Davidson method can be applied to solve (3.12). If $(\theta_j^{(k)}, u_j^{(k)})$, with $u_j^{(k)} \perp Q_\ell$ and $u_j^{(k)} \perp V_k$, is an approximate eigenpair for (3.12), then, for some appropriate $q_j^{(k)} \perp Z$, the correction equation for (3.12) reads as

$$(I - [Z_\ell, q_j^{(k)}]\,[Z_\ell, q_j^{(k)}]^*)(A - \theta_j^{(k)} B)(I - [Q_\ell, u_j^{(k)}]\,[Q_\ell, u_j^{(k)}]^*)t = -r_k, \tag{3.13}$$

[1] This was also mentioned by A. Stathopoulos (personal communication).

where r_k is the residual for the deflated system,

$$-r_k \equiv -(I - Z_\ell Z_\ell^*)(A - \theta_j^{(k)} B)(I - Q_\ell Q_\ell^*) u_j^{(k)}, \tag{3.14}$$

and $t \perp [Q_\ell, u_j(k)]$ is the wanted expansion vector for the search subspace V_k.

The correction equation (3.13) can be obtained directly from correction equation (2.4), simply by replacing $q_j^{(k)}$ with $[Z_\ell, q_j^{(k)}]$ and $u_j^{(k)}$ with $[Q_\ell, u_j^{(k)}]$. Incorporating a preconditioner for the deflated equation can be done in a similar way: if P_k is a preconditioner for $A - \theta_j^{(k)} B$, then the approaches in the preceding subsections 3.2 and 3.2.1, with $q_j^{(k)}$ replaced with $[Z_\ell, q_j^{(k)}]$ and $u_j^{(k)}$ with $[Q_\ell, u_j^{(k)}]$, give a correct way of handling of the projected preconditioner for (3.13). Then μ transforms into the $\ell + 1$ by $\ell + 1$ matrix

$$[Q_\ell, u_j^{(k)}] P_k^{-1} [Z_\ell, q_j^{(k)}] = \begin{bmatrix} Q_\ell^* P_k^{-1} Z_\ell & Q_\ell^* \tilde{q} \\ {u_j^{(k)}}^* P_k^{-1} Z_\ell & {u_j^{(k)}}^* \tilde{q} \end{bmatrix}, \quad \text{with} \quad \tilde{q} \equiv P_k^{-1} q_j^{(k)}$$

and $\frac{1}{\mu}$ should be interpreted as matrix inversion.

4. The selection of approximate eigenpairs. The pair $(\theta, u \equiv V_k s)$ may be considered as the best approximate eigenpair in the search subspace if it solves the minimization problem

$$\min_{\theta, u} \{ \|Au - \theta Bu\| \mid u = V_k s, \|u\|_2 = 1 \}.$$

However, solving this problem is computationally more expensive than solving the Petrov-Galerkin system (2.3), and there may be Petrov-Galerkin solutions that are of comparable quality if the test subspace is appropriately chosen.

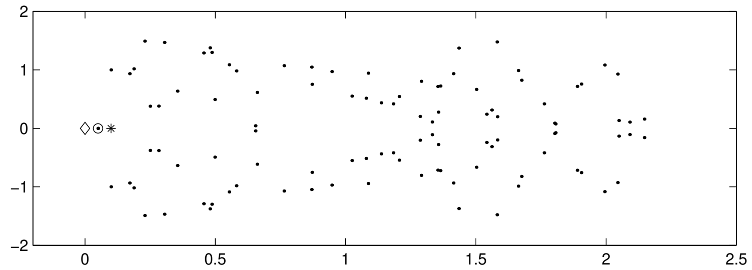
First, in §4.1, we will argue that test subspaces for so-called harmonic-Petrov values are appropriate and then, in §4.2, we will discuss an approach for finding the best approximate eigenvectors among the set of k Petrov vectors.

4.1. Harmonic Ritz values for standard eigenproblems. In order to simplify our discussions, we will assume that A is normal and $B = I$ (a standard eigenproblem). Then, there is an orthonormal basis of eigenvectors.

For a Ritz pair $(\theta, u \equiv V_k s)$, the search subspace is also the test space and (θ, u) satisfies $Au - \theta u \perp V_k$. Hence $\theta = (Au, u)/(u, u)$, from which it can be deduced that θ is a convex combination $\sum_i \gamma_i \lambda_i$ of the eigenvalues λ_i. If θ is close to an extremal eigenvalue λ_{i_0} then this eigenvalue will give a major contribution to θ, that is, $\gamma_i \ll \gamma_{i_0}$ $(i \neq i_0)$ and, consequently, the Ritz vector u will have a small angle with the eigenvector associated with λ_{i_0}: γ_{i_0} is the square of the cosine of this angle. If we are interested in this eigenvector, then we can safely select this Ritz pair for the Jacobi-Davidson correction equation. If θ is close to an interior eigenvalue λ_{i_0} then γ_{i_0} can very well be negligible: θ may be a combination of, for instance, eigenvalues far to the left and far to the right of θ.

Conclusion: if we are interested in extremal eigenvalues, say the one with smallest real part, then the Ritz pair, with Ritz value of smallest real part, forms a safe choice, at least for the Jacobi-Davidson correction equation, when we have some degree of convergence. For interior eigenvalues, the approach with Ritz pairs may fail.

Suppose that we are interested in eigenvalues close to some target value τ in the interior of the spectrum of A. Since $(A - \tau I)^{-1} x = x/\delta$ implies $Ax = (\tau + \delta) x$, we see that eigenvalues close to τ correspond to large, extremal, eigenvalues of $(A - \tau I)^{-1}$. Therefore, it is tempting to use the Ritz approach for $(A - \tau I)^{-1}$. With a test space of the form $(A - \tau I) V_k$,



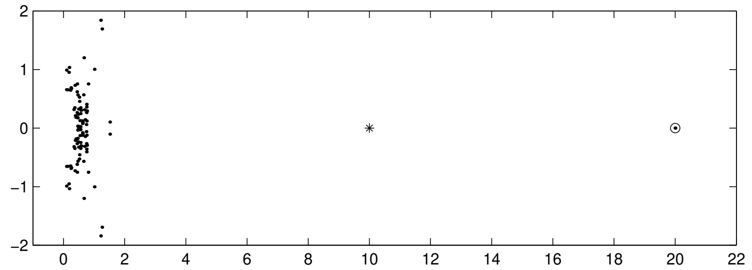


FIG. 4.1. *The figures show the set of eigenvalues (the $\cdot$'s) of a normal matrix A (top figure), and of $(A-\tau I)^{-1}$ (bottom figure). The diamond ($\diamond$) is the target value τ; the dotted circle ($\odot$) is the target eigenvalue, that is the eigenvalue closest to τ in the upper figure, the absolute largest eigenvalue in the bottom figure. The asterisk ($*$) is the (harmonic) Ritz value.*

the matrix inversion can be avoided. If we write $y \equiv (A - \tau I)V_k s$ then

$$\begin{aligned} V_k s - \frac{1}{\delta}(A - \tau I)V_k s &= (A - \tau I)^{-1} y - \frac{1}{\delta} y \\ &\perp (A - \tau I)V_k, \end{aligned}$$

or, equivalently, with $u \equiv V_k s$:

$$Au - (\tau + \delta)u \perp W_k \equiv (A - \tau I)V_k.$$

The Ritz approach for $(A-\tau I)^{-1}$ corresponds to the so-called *harmonic Ritz* space approach for A, where V_k spans the search subspace and $W_k = (A - \tau I)V_k$ spans the test subspace. The pair $(\theta \equiv \tau + \delta, u)$ is a harmonic Ritz pair for $A - \tau I$ with respect to V_k [9]. Since multiplication by $A - \tau I$ diminishes the eigenvector component which is of interest in u, u is taken as eigenvector approximation, rather than $y = (A - \tau I)u$. Arguments for preferring harmonic Ritz values for interior eigenvalues can be found in, e.g., [7, 14].

Apparently, misselection can be avoided, at least asymptotically, if we use harmonic Ritz pairs in the Jacobi-Davidson correction equation for the computation of interior eigenvalues. But in the non-Hermitian case, the harmonic Ritz value can be attractive also for extremal eigenvalues. If, for a small positive ϵ, A has eigenvalues, say, $\lambda_1 = \epsilon$, $\lambda_{2,3} = \pm i + 2\epsilon$ and all other eigenvalues have real part much larger than ϵ, then a Ritz value $\theta = 2\epsilon$ may be a combination of λ_2 and λ_3 with zero λ_1-component (see the top figure in Fig. 4.1). On the other hand, since a harmonic Ritz value 2ϵ for A corresponds to a Ritz value $\frac{1}{2\epsilon}$ for A^{-1}, and since the eigenvalues $\frac{1}{\lambda_2}$ and $\frac{1}{\lambda_3}$ of A^{-1} are still close to the imaginary axis, the Ritz vector

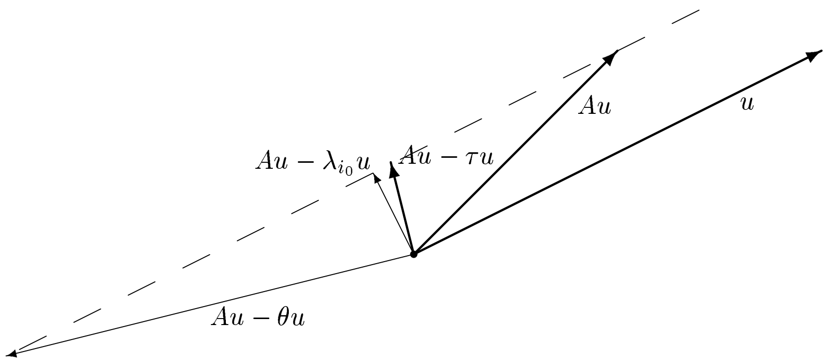


FIG. 4.2. *The effect of testing with harmonic Ritz values when (τ, u) approximates an eigenpair (λ_{i_0}, x_{i_0}) moderately well. The residual $r = Au - \theta u$ is orthogonal with respect to the test subspace spanned by $Au - \tau u$. Note that $|\lambda_{i_0} - \tau| \ll |\lambda_{i_0} - \theta|$.*

associated with $\frac{1}{2\epsilon}$ will have a large component in the eigenvector direction of the eigenvalue $\frac{1}{\lambda_1}$ (see the bottom figure in Fig. 4.1).

4.2. Locally minimal residuals. In this subsection, we will consider again the standard normal eigenproblem.

In the discussion in §4.1, we considered approximate eigenvalues that are already close to some target eigenvalue. In the non-asymptotic case, extremal Ritz values are not necessarily the best candidates for approximating extremal eigenvalues. If, for instance, A is symmetric with eigenvalues λ_i in increasing order, and the smallest Ritz value θ_1 is larger than λ_2, then θ_1 may be a combination of mainly λ_2 and λ_3 with almost zero λ_1-component, while θ_2 may have a very significant λ_1-component (and a smaller λ_n-component). Since harmonic Ritz values close to a target τ correspond to extremal Ritz values for the inverse of $A - \tau I$, a similar observation also holds for harmonic Ritz values with respect to interior eigenvalues close to τ.

It would be ideal to have a moderately accurate initial guess for the initial search space; a τ that is moderately close to some λ_{i_0}, while the eigenvector x_{i_0} associated with λ_{i_0} plus some other eigenvectors with nearby eigenvalues are moderately close to the subspace spanned by V_ℓ, with, say, angles less than 10°. Such an ideal situation is likely to occur after restarts (provided that the target is updated) and when, after detection of an eigenpair, deflation is applied for the computation for a next nearby eigenpair: in the preceding steps the search subspace will have gathered already components in the direction of the desired eigenvector. However, it is precisely this situation were, when using harmonic Ritz values, the confusion as sketched in the previous paragraph occurs. For a harmonic Ritz pair $(\theta, u \equiv V_\ell s)$ we have that $Au - \theta u \perp Au - \tau u$. If u is directionally close to x_{i_0} and λ_{i_0} is close to τ, then θ is far from τ (see Fig. 4.2). If u is directionally close to an eigenvector with eigenvalue not close to τ, then θ will approximate that eigenvalue well: either the angle between the harmonic Ritz vector and the target eigenvector is large or the harmonic Ritz value is far from the target eigenvalue. It can be shown that there is a harmonic Ritz vector, in this almost converged situation, that is directionally close to the desired eigenvector. The question is how to detect that harmonic Ritz vector and how to avoid working with the associated harmonic Ritz value, which seems to be inappropriate.

Efficient detection of the appropriate harmonic Ritz vector

If we have a good approximate eigenvector u_{old} from, for instance, a previous Jacobi-Davidson step, then we propose to scan all harmonic Ritz vectors by computing the angle with this approximate eigenvector. If the angle with a harmonic Ritz vector is less than, say, 45° then we may select that harmonic Ritz vector and stop further scanning. Since the harmonic Ritz vectors converge to an orthogonal system, we may expect correct selection already after moderate convergence. In practice, the search matrix V_ℓ has orthonormal columns which facilitates efficient computation of the angles, since, if $u_{\text{old}} = V_{\ell-1}s_{\text{old}}$ and $u = V_\ell s$ then $(u_{\text{old}}, u) = (s_{\text{old}}, \tilde{s})$, where $\tilde{s}$ is the vector composed from the first $\ell - 1$ components of s. If a QZ-factorization of the interaction matrices $W_\ell^* A V_\ell$ and $W_\ell^* V_\ell$ is used for computing the ℓ-dimensional vectors s, then some additional savings are possible by exploiting the structure of the triangular matrices in the factorization.

Using locally minimal residuals
Now, suppose we have detected a harmonic Ritz vector, say u_{hR}, that is moderately close to x_{i_0}, and suppose that the associated harmonic Ritz value θ_{hR} is far from λ_{i_0}. Then $\|Au_{\text{hR}} - \theta_{\text{hR}} u_{\text{hR}}\|_2$ is large (in comparison with $\|Au_{\text{hR}} - \tau u_{\text{hR}}\|_2$). To find a better residual and a better approximate eigenvalue, we can simply take the value $\theta = \theta_{\text{locR}}$ for which $\|Au_{\text{hR}} - \theta u_{\text{hR}}\|_2$ is minimal. This is precisely the Rayleigh quotient $(Au_{\text{hR}}, u_{\text{hR}})/(u_{\text{hR}}, u_{\text{hR}})$, which is a Ritz value with respect to the 1-dimensional space generated by u_{hR}. Note that the associated residual $r_{\text{locR}} = (A - \theta_{\text{locR}} I)u_{\text{hR}}$, the *locally minimal residual*, is orthogonal to u_{hR}. As for the Ritz approach, this mixed approach with harmonic Ritz vectors and local Ritz values leads to cubic convergence for standard normal eigenproblems, if the search subspace is expanded with exact solutions of the associated Jacobi-Davidson correction equation

$$t \perp u_{\text{hR}}, \qquad (I - u_{\text{hR}}\, u_{\text{hR}}^*)(A - \theta_{\text{locR}} I)(I - u_{\text{hR}}\, u_{\text{hR}}^*)t = -r_{\text{locR}}. \tag{4.1}$$

Summary
The approach with harmonic Ritz values avoids misselection when we have near convergence; selecting the approximate eigenvector from the set of harmonic Ritz vectors by inspection of the angles with the approximate eigenvector from the preceding step avoids a poor selection of $(\theta_j, u_j^{(k)})$ in case of moderate convergence and does not lead to a significant increase in the computational costs per step; for a selected approximate eigenvector, the norm of the residual can be minimized by selecting a suitable approximate θ_j, which avoids irregular convergence.

4.3. Harmonic Petrov values for generalized eigenproblems. The approaches sketched in §§4.1 and 4.2 can be generalized for $Ax = \lambda Bx$. Although, for this type of problem, a theoretical justification can generally not be given, the approach appears to be practical. We will sketch its main ingredients.

Suppose we are interested in eigenvalues near a target τ. Then *harmonic Petrov pairs* [3] $(\theta, u = V_k s)$ are formed from the solutions of the k-dimensional eigenproblem (2.3), with $W_k \equiv (A - \tau B)V_k$ (or with W_k an orthonormal matrix over $(A - \tau B)V_k$). Again, harmonic Petrov values closest to the target τ can be related to extremal Ritz values of a matrix involving the inverse of $A - \tau B$. However, in general the matrices involved are non-normal and then the Ritz values can not be expressed as a convex combination of appropriate eigenvalues. It cannot be proved that a Ritz vector is directionally close to an eigenvector if the associated eigenvalue is extremal and close to a Ritz value. Nevertheless the approach with harmonic Petrov vectors appears to work well.

As before, inspection of the angles, between the selected approximate eigenvector of Jacobi-Davidson step j versus the harmonic Petrov vectors of step $j + 1$, can help to avoid misselection in case of moderate convergence (and accurate target τ).

Minimizing the residual norm $\|Au_{\text{hR}} - \theta B u_{\text{hR}}\|_2$, with $\|u_{\text{hR}}\|_2 = 1$, can introduce instabilities if $\|Bu_{\text{hR}}\|_2$ is small relative to $\|B\|_2$, while $\|Au_{\text{hR}}\|_2$ is relatively not small: then the computed Bu_{hR} may be affected significantly by rounding errors, in contrast to the computed Au_{hR}. For this reason eigenvalues and approximate eigenvalues are "balanced" over both components [11]: the residual norm $\|\theta_c Au_{\text{hR}} - \theta_s Bu_{\text{hR}}\|_2$ is minimized over all complex pairs (θ_c, θ_s), with $\theta_c \in [0,1]$ and $\theta_c^2 + |\theta_s|^2 = 1$. Then the locally minimizing residual $r_{\text{locR}} = \theta_{c,\text{locR}} Au_{\text{hR}} - \theta_{s,\text{locR}} Bu_{\text{hR}}$ is proportional to the smallest singular vector of the system $E \equiv [Au_{\text{hR}}, Bu_{\text{hR}}]$. To be more precise, $r_{\text{locR}} = \sigma_1 p_1$, where $E[b_1, b_2] = [p_1, p_2]\text{diag}(\sigma_1, \sigma_2)$ is the singular value decomposition of E with $\sigma_1 \leq \sigma_2$, and $b_1 = (\theta_{c,\text{locR}}, -\theta_{s,\text{locR}})^*$. The largest singular vector p_2 is orthogonal to r_{locR} and is a linear combination of Au_{hR} and Bu_{hR}. The correction equation, as in (2.4) with $q_j^{(k)}$ replaced by p_2 and r_k by r_{locR}, leads to good expansion vectors $\tilde{t}$ for the search subspace V_k, and, with $A\tilde{t} - \tau B\tilde{t}$, to suitable expansion vectors for the test subspace W_k.

5. Some ideas for preconditioners. Unless A has some very special structure so that (the major part of) A can be inverted efficiently, we have to consider some kind of preconditioning, for instance a preconditioner based on incomplete LU decompositions. A major problem is that we would like to have an efficient preconditioner for the operator $(I - u_j^{(k)} u_j^{(k)^*})(A - \theta_j^{(k)} I)(I - u_j^{(k)} u_j^{(k)^*})$, but since this operator is represented by a dense matrix, we focus on preconditioners K for $A - \theta I$ and we use the projected preconditioner

$$(I - u_j^{(k)} u_j^{(k)^*})K(I - u_j^{(k)} u_j^{(k)^*}).$$

Note that this does not necessarily lead to a good preconditioner, because we approximate before we project. The projected preconditioner should behave like the projected $A - \theta I$ for eigenvectors associated with eigenvalues close to θ. In general this is a difficult goal to achieve with incomplete decompositions, especially if θ approximates some interior eigenvalue. This is a fundamental complication. In general it is difficult, for given A, to identify a successful preconditioner, but for eigenproblems the situation is even more complicated because we also have the shift θ, which varies with the eigenvalue that we are looking for. A preconditioning technique that leads to efficient preconditioners for given A for some values of θ may be impractical for the same A with a different set of θ's. In particular, for values of θ located in the interior of the spectrum of A, the matrix $A - \theta I$ will be highly indefinite. There is no guarantee that a standard incomplete decomposition will lead to errors that do not affect the eigenvectors of the preconditioner for eigenvalues close to the wanted one. In our experience we found that one has to admit a great deal of fill-in in order to get efficient preconditioning matrices for interior eigenvalues, sometimes so much that we could have used a full decomposition at virtually the same costs.

Sometimes the situation is more favourable. For the discretized Poisson operator A, it has been shown that an incomplete decomposition K has almost the same eigenvectors as A for a few of the smallest eigenvalues. This has been used as an argument to explain the effectiveness of ILU-preconditioning [18, 17]. If such a preconditioner is used for values of θ close to zero, then the ILU-process will yield a preconditioner that is also effective for a number of nearby small eigenvalues. This helps to explain the success of ILU preconditioning in the Jacobi-Davidson method for computing a number of the smallest eigenvalues of discretized elliptic operators and other eigenproblems that stem from discretized pde's. A procedure that has been reported to work well is to construct an incomplete LU-decomposition, for the initial target value, only once and to use this for various values of θ, even for different eigenvalues in combination with deflation [3]. In [3] it has been argued that deflation may help to improve the effectiveness of a given fixed preconditioner, since possible large errors in the incomplete decomposition in critical eigenvector directions may be removed by the deflation process. In [3, §4.7] an example is given where the usage of a fixed but deflated preconditioner is

ETNA
Kent State University
etna@mcs.kent.edu

still very efficient for the twentieth eigenvalue (when ordered algebraically), notwithstanding the increased costs per iteration involved with deflation against the previously determined 19 eigenvectors.

It is well-known that MILU [5] often leads to a significant reduction in iteration steps when used as a preconditioner for Krylov subspace iteration methods for the solution of discretized PDE's with a relatively smooth solution. For an incomplete decomposition of the 5-point discretized Poisson operator, over a rectangular grid, the diagonal elements of a MILU decomposition are typically smaller than those for standard ILU. Consequently, the off-diagonal decomposition errors with MILU are larger, but their effect is compensated by corrections to the diagonal. When we try to compute, with Jacobi-Davidson, the smallest eigenvalue of the Poisson matrix A, with corresponding smooth eigenvector, then MILU will still work well as a preconditioner for the correction equation as long as the involved θ's are small as well. If θ is located more in the interior of A's spectrum, then the diagonal elements of the MILU preconditioner will be smaller, making the decomposition errors bigger. Therefore, for larger values of θ the MILU preconditioner may be less effective or even become unstable.

Moreover, if we want to compute interior eigenvalues of A, then the corresponding eigenvectors will also be more oscillatory and hence the errors in the MILU approximation will not have a small effect in those eigenvector directions (remember that the error matrix in MILU is designed to have a small effect for almost constant vectors by compensating the decomposition errors with corrections to the diagonal). This may help to explain our observations that for interior eigenvalues, ILU-type preconditioning was more effective than MILU preconditioning (see Fig. 5.1). We have not systematically investigated whether the discussed effects can be diminished for MILU, for instance, by including appropriate fill-in, or by other correction mechanisms for the diagonal elements.

5.1. Ill-conditioned Preconditioners. If we attempt to approximate $A - \theta_j^{(k)} I$, then when $\theta_j^{(k)}$ is close to an eigenvalue an accurate approximation should necessarily be close to a singular matrix. This can be easily controlled by checking the diagonal elements of an incomplete factorization. The problem is that, since $A - \theta_j^{(k)} I$ is in general not an M-matrix, a flawless incomplete decomposition is not guaranteed to exist and a small diagonal element may occur as the result of approximation errors in the incomplete decomposition process, as well as because of the near-singularity. For the discretized Poisson operator we can prove that for θ close to the smallest eigenvalue of A, the incomplete decomposition process will lead to a small diagonal element only at the end of the decomposition. Anyway, with a near-singular preconditioner we run the risk of stability problems. If the preconditioner is denoted by P_k, then the preconditioned vector in the projected subspace is formally represented as (cf. formula (3.2))

$$z = P_k^{-1} \widetilde{y} - \alpha P_k^{-1} u_j^{(k)}. \tag{5.1}$$

The effect of cancellation is most easily seen when we take $P_k = A - \theta I$, and this has been discussed in [14]. There is an easy way to remove the cancellation problem. It is readily verified that the solution of the correction equation (1.2) remains unaltered if we replace $A - \theta I$ by $A - \theta I + \mu u_j^{(k)} {u_j^{(k)}}^*$.[2] For $\mu \neq 0$ we have removed the near-singularity, which kills

[2] In [20] it is suggested to use this this operator right away for the construction of a correction, that is, without the projections. In their case that would lead to a different approximation than our $\widetilde{t}^{(k)}$. In particular, the usage of the update in [20] does not cure the effects of ill-conditioning of $A - \theta_j^{(k)}$. Also, their update does not satisfy the constraint to be orthogonal to $u_j^{(k)}$.

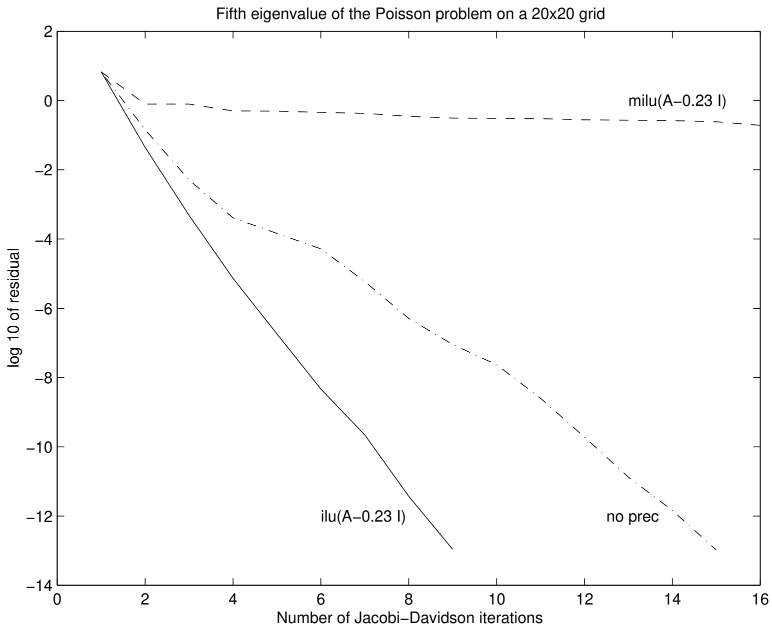


FIG. 5.1. *The* $\log_{10}$ *of the residual norm for iterative solves of the correction equations with 15 steps of* GMRES *without preconditioning* (*the dash-dotted curve*: $- \cdot$)*, with* ILU *preconditioner* (*the solid curve:* —)*, and with* MILU *preconditioner* (*the dashed curve*: $- -$)*. A is the* 5*-point discretized Poisson operator over a* 20 *by* 20 *grid. Both preconditioners are incomplete factorizations of* $A - \theta I$ *with* $\theta = 0.23$*. With this value of* θ *we are aiming for the 5th smallest eigenvalue.*

the large components in each of the two vectors in the right-hand side of (5.1) in the direction of $u_j^{(k)}$. Although this cures the problem, we note that this may not be very practical, since the correction $u_j^{(k)} u_j^{(k)^*}$ will be a dense matrix and destroys all sparsity in the given matrix. In other words, it may be difficult to identify a good preconditioner for the rank-one updated matrix. With the Sherman-Woodbury formula [4], for the inversion of a rank-one updated matrix, the problems with the ill-conditioning will not be avoided. The correction equation gives us more freedom with respect to corrections in the direction of $u_j^{(k)}$. For instance, we may correct very tiny diagonal elements, in the factors L and U of P_k, by rank-one updates with unit-vectors:

$$\widetilde{P_k} = (L + u_j^{(k)} e_{i_1}^*)(U + e_{i_2} u_j^{(k)^*}).$$

This, in exact arithmetic, does not change the effect of P_k in the subspace orthogonal to $u_j^{(k)}$. More of these updates may be included if necessary. Note that the factors of this new factorization can also be cheaply solved without cancellation problems, unless the i_1-th or i_2-th element of $u_j^{(k)}$ is zero.

We illustrate the effect of rank-one updating to the operator, *in the context of the correction equation* (1.2), by a simple experiment. In Fig. 5.2 we have displayed the convergence history, that is, the norm of the residual vectors $Au_1^{(k)} - \theta_1^{(k)} u_1^{(k)}$, for the 5-point discretized Poisson operator over a 20 by 20 grid. One curve shows what happens if we solve by direct solution with $A - \theta_1^{(k)} I$, taking into account the projections in (1.2), but without reorthogonalization

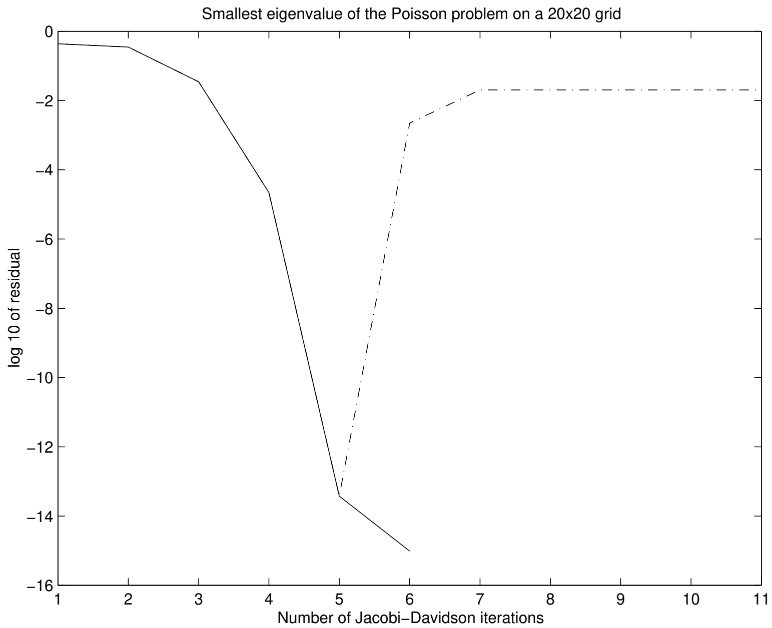


FIG. 5.2. *The $\log_{10}$ of the residual norm for direct solves of the correction equations with reorthogonalization* (*the solid curve*: —) *and without reorthogonalization* (*the dash-dotted curve*: $-\ \cdot$). *In the first part of the picture both curves coincide.*

of the basis vectors v_k. The other curve shows that reorthogonalization is necessary. If we update the matrix, inside the projections in (1.2), with $\mu u_1^{(k)} u_1^{(k)^*}$ then we get virtually the same curve as with reorthogonalization. That means that one has the choice between a single rank-one update or a reorthogonalization against all previous basis vectors and one may select the most convenient strategy depending on whether the rank-one update complicates the direct solution step or not.

REFERENCES

[1] E. R. DAVIDSON, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices*, J. Comput. Phys., 17 (1975), pp. 87–94.
[2] R. S. DEMBO, S. C. EISENSTAT AND T. STEIHAUG, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408.
[3] D. R. FOKKEMA, G. L. G. SLEIJPEN AND H. A. VAN DER VORST, *Jacobi-Davidson style QR and QZ algorithms for the partial reduction of matrix pencils*, SIAM J. Sci. Comput., 20 (1998), pp. 94–125.
[4] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996.
[5] I. GUSTAFSSON, *A Class of First Order Factorization Methods*, BIT, 18 (1978), pp. 142–156.
[6] C. G. J. JACOBI, *Ueber ein leichtes Verfahren, die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, Journal für die reine und angewandte Mathematik, (1846), pp. 51–94.
[7] R. B. MORGAN, *Computing interior eigenvalues of large matrices*, Linear Algebra Appl., 154/156 (1991), pp. 289–309.
[8] J. OLSEN, P. JØRGENSEN, AND J. SIMONS, *Passing the one-billion limit in full configuration-interaction (FCI) calculations*, Chemical Physics Letters, 169 (1990), pp. 463–472.
[9] C. C. PAIGE, B. N. PARLETT, AND H. A. VAN DER VORST, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Num. Lin. Algebra Appl., 2 (1995), pp. 115–133.

[10] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[11] A. RUHE, *Rational Krylov algorithms for nonsymmetric eigenvalue problems II. Matrix Pairs*, Lin. Algebra Appl., 197/198 (1994), pp. 283–295.

[12] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[13] G. L. G. SLEIJPEN AND D. R. FOKKEMA, *BiCGstab*(ℓ) *for linear equations involving matrices with complex spectrum*, Electron. Trans. Numer. Anal., 1 (1993), pp. 11–32.

[14] G. L. G. SLEIJPEN AND H. A. VAN DER VORST, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.

[15] ———, *The Jacobi-Davidson method for eigenvalue problems and its relation to accelerated inexact Newton schemes*, in Iterative Methods in Linear Algebra, II, S. D. Margenov and P. S. Vassilevski, eds., IMACS Ann. Comput. Appl. Math. 3, Baltzer, Basel, 1996, pp. 377-389.

[16] G. L. G. SLEIJPEN, J. G. L. BOOTEN, D. R. FOKKEMA, AND H. A. VAN DER VORST, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36:3 (1996), pp. 595–633.

[17] H. A. VAN DER VORST AND G. L. G. SLEIJPEN, *The effect of incomplete decomposition preconditioning on the convergence of Conjugate Gradients*, in Incomplete Decompositions (ILU) - Algorithms, Theory, and Applications, W. Hackbusch and G. Wittum, eds., Vieweg, Braunschweig, 1993, pp. 179–187.

[18] H. A. VAN DER VORST, *Preconditioning by Incomplete Decompositions*, Ph.D. thesis, Utrecht University, Utrecht, The Netherlands, 1982.

[19] ———, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.

[20] K. WU, Y. SAAD, AND A. STATHOPOULOS, *Inexact Newton preconditioning techniques*, Electron. Trans. Numer. Anal., 7 (1998), pp. 166–184.