Journal of Computational and Applied Mathematics 235 (2011) 2626–2639

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

The Rayleigh–Ritz method, refinement and Arnoldi process for periodic matrix pairs

Eric King-Wah Chu [a,*], Hung-Yuan Fan [b], Zhongxiao Jia [c], Tiexiang Li [d], Wen-Wei Lin [e]

[a] School of Mathematical Sciences, Building 28, Monash University, VIC 3800, Australia
[b] Department of Mathematics, National Taiwan Normal University, Taipei 116, Taiwan
[c] Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China
[d] Department of Mathematics, Southeast University, Nanjing 211189, China
[e] CMMSC and NCTS, Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

Article history:
Received 2 December 2009

Keywords:
Arnoldi process
Periodic eigenvalues
Periodic matrix pairs
Rayleigh–Ritz method
Refinement
Ritz values

ABSTRACT

We extend the Rayleigh–Ritz method to the eigen-problem of periodic matrix pairs. Assuming that the deviations of the desired periodic eigenvectors from the corresponding periodic subspaces tend to zero, we show that there exist periodic Ritz values that converge to the desired periodic eigenvalues unconditionally, yet the periodic Ritz vectors may fail to converge. To overcome this potential problem, we minimize residuals formed with periodic Ritz values to produce the refined periodic Ritz vectors, which converge under the same assumption. These results generalize the corresponding well-known ones for Rayleigh–Ritz approximations and their refinement for non-periodic eigen-problems. In addition, we consider a periodic Arnoldi process which is particularly efficient when coupled with the Rayleigh–Ritz method with refinement. The numerical results illustrate that the refinement procedure produces excellent approximations to the original periodic eigenvectors.

1. Introduction

Let $E_j, A_j \in \mathbb{C}^{n\times n}$ $(j = 1, \ldots, p)$, where $E_{j+p} = E_j$ and $A_{j+p} = A_j$ for all j. We denote the periodic matrix pairs of periodicity p by $\{(A_j, E_j)\}_{j=1}^{p}$. In this paper, the indices j for all periodic coefficient matrices are chosen in $\{1, \ldots, p\}$ modulo p. The equations

$$\beta_j A_j x_{j-1} = \alpha_j E_j x_j \quad (j = 1, 2, \ldots, p) \tag{1}$$

with $x_0 = x_p$ define the nonzero periodic right eigenvectors $\{x_j\}_{j=1}^{p}$ for complex ordered pairs $\{(\alpha_j, \beta_j)\}_{j=1}^{p}$. Similarly, the equations

$$\beta_{j-1} y_j^H A_j = \alpha_j y_{j-1}^H E_{j-1} \quad (j = 1, 2, \ldots, p) \tag{2}$$

with $y_0 = y_p$ define the nonzero periodic left eigenvectors $\{y_j\}_{j=1}^{p}$. The ordered pairs $(\pi_\alpha, \pi_\beta) \equiv \left(\prod_{j=1}^{p} \alpha_j, \prod_{j=1}^{p} \beta_j\right)$ then constitute the spectrum, with the traditional eigenvalues being the quotients π_α/π_β. Because of the possibility of infinite eigenvalues, we shall deal with spectra in their ordered pair representation, with equality interpreted in the sense of the corresponding equivalent relationship for quotients. Using the notation $\mathrm{col}[x_j]_{j=1}^{p} \equiv [x_1^\top, \ldots, x_p^\top]^\top$ and

* Corresponding author. Tel.: +61 412 596430; fax: +61 3 99054403.
E-mail addresses: eric.chu@sci.monash.edu.au, eric.chu@monash.edu (E.K.-W. Chu), hyfan@ntnu.edu.tw (H.-Y. Fan), jiazx@tsinghua.edu.cn (Z. Jia), txli@seu.edu.cn (T. Li), wwlin@am.nctu.edu.tw (W.-W. Lin).

0377-0427/$ – see front matter
doi:10.1016/j.cam.2010.11.014

$$C\begin{pmatrix}\alpha_1,\ldots,\alpha_p\\ \beta_1,\ldots,\beta_p\end{pmatrix}\equiv\begin{bmatrix}\alpha_1E_1 & & & -\beta_1A_1\\ -\beta_2A_2 & \alpha_2E_2 & & \\ & \ddots & \ddots & \\ & & -\beta_pA_p & \alpha_pE_p\end{bmatrix},\tag{3}$$

the eigen-equations (1) and (2) can also be written as the multivariate eigen-problems, respectively,

$$C\begin{pmatrix}\alpha_1,\ldots,\alpha_p\\ \beta_1,\ldots,\beta_p\end{pmatrix}\mathrm{col}[x_j]_{j=1}^p=0\tag{4}$$

and

$$\left\{\mathrm{col}[y_j]_{j=1}^p\right\}^H C\begin{pmatrix}\alpha_2,\ldots,\alpha_p;\alpha_1\\ \beta_p;\beta_1\ldots,\beta_{p-1}\end{pmatrix}=0^\top.\tag{5}$$

In this paper, we consider only regular periodic matrix pairs for which

$$\det C\begin{pmatrix}\alpha_1,\ldots,\alpha_p\\ \beta_1,\ldots,\beta_p\end{pmatrix}=\sum_{k=0}^n c_k\pi_\alpha^k\pi_\beta^{n-k}\not\equiv 0,\tag{6}$$

and consequently all eigenvalues $(\pi_\alpha,\pi_\beta)\not\equiv(0,0)$. For regular periodic matrix pairs, at least one of the coefficients $c_k\neq 0$ and there are exactly n eigenvalues for $\{(A_j,E_j)\}_{j=1}^p$, counting multiplicities. The spectrum, or the set of all eigenvalue pairs, of $\{(A_j,E_j)\}_{j=1}^p$ is denoted by $\lambda(\{(A_j,E_j)\}_{j=1}^p)$.

For the periodic matrix pairs $\{(A_j,E_j)\}_{j=1}^p$, we have the periodic Schur decomposition of $\{(A_j,E_j)\}_{j=1}^p$ [1–3].

Theorem 1.1 (*Periodic Schur Decomposition*). *Let $\{(A_j,\ E_j)\}_{j=1}^p$ be regular matrix pairs. There exist unitary matrices Q_j, Z_j ($j=1,2,\ldots,p$) such that*

$$Q_j^HA_jZ_{j-1}=\hat{A}_j,\qquad Q_j^HE_jZ_j=\hat{E}_j\quad(j=1,2,\ldots,p)$$

are all upper triangular, with $Z_0=Z_p$. Moreover, the diagonal parts

$$\{[\mathrm{diag}(\alpha_{j1},\ldots,\alpha_{jn}),\mathrm{diag}(\beta_{j1},\ldots,\beta_{jn})]\}_{j=1}^p$$

of $\{(\hat{A}_j,\ \hat{E}_j)\}_{j=1}^p$ determine all the eigenvalues $\left\{\left(\prod_{j=1}^p\alpha_{jk},\ \prod_{j=1}^p\beta_{jk}\right)\right\}_{k=1}^n$ of $\{(A_j,\ E_j)\}_{j=1}^p$, which can be arranged in any order.

We can also generalize the concept of deflating subspaces as follows [4,3].

Definition. Let $\mathcal{X}_j,\mathcal{Y}_j$ ($j=1,2,\ldots,p$) be subspaces in $\mathbb{C}^n$ of equal dimension. The pairs $\{(\mathcal{X}_j,\ \mathcal{Y}_j)\}_{j=1}^p$ are called the periodic deflating subspaces of $\{(A_j,E_j)\}_{j=1}^p$ if

$$A_j\mathcal{X}_{j-1}\subset\mathcal{Y}_j,\qquad E_j\mathcal{X}_j\subset\mathcal{Y}_j\quad(j=1,2,\ldots,p)$$

with $\mathcal{X}_0=\mathcal{X}_p$. Furthermore, the subspaces $\{\mathcal{X}_j\}_{j=1}^p$ are called the periodic invariant subspaces of $\{(A_j,\ E_j)\}_{j=1}^p$.

We list some further results and definitions from [3].

(i) Theorem 1.1 implies that $\lambda(\{(A_j,\ E_j)\}_{j=1}^p)=\lambda(\{(A_j^\top,\ E_j^\top)\}_{j=1}^p)$.
(ii) An eigenvalue is said to be simple if it appears in a linear factor of the characteristic polynomial.
(iii) Let $Z_1^{(j)},Q_1^{(j)}\in\mathbb{C}^{n\times k}$ satisfy $(Z_1^{(j)})^HZ_1^{(j)}=(Q_1^{(j)})^HQ_1^{(j)}=I_k$, and let $\mathcal{X}_j=\mathrm{span}(Z_1^{(j)})$, $\mathcal{Y}_j=\mathrm{span}(Q_1^{(j)})$ for all j. It can be verified [3] that $\{(\mathcal{X}_j,\mathcal{Y}_j)\}_{j=1}^p$ are periodic deflating subspaces of the regular matrix pairs $\{(A_j,\ E_j)\}_{j=1}^p$ if and only if there exist unitary matrices $Z_j=[Z_1^{(j)},\ Z_2^{(j)}]$, $Q_j=[Q_1^{(j)},\ Q_2^{(j)}]\in\mathbb{C}^{n\times n}$ such that

$$Q_j^HA_jZ_{j-1}=\begin{bmatrix}A_{11}^{(j)} & A_{12}^{(j)}\\ 0 & A_{22}^{(j)}\end{bmatrix},\qquad Q_j^HE_jZ_j=\begin{bmatrix}E_{11}^{(j)} & E_{12}^{(j)}\\ 0 & E_{22}^{(j)}\end{bmatrix},\tag{7}$$

where $A_{11}^{(j)},E_{11}^{(j)}\in\mathbb{C}^{k\times k}$, and both $\{(A_{11}^{(j)},\ E_{11}^{(j)})\}_{j=1}^p$ and $\{(A_{22}^{(j)},\ E_{22}^{(j)})\}_{j=1}^p$ are regular for all j. Furthermore, if the intersection of the spectra of the two sub-matrix pairs is empty, the periodic deflation subspaces $\{(\mathcal{X}_j,\ \mathcal{Y}_j)\}_{j=1}^p$ are called simple periodic deflating subspaces, and $\{\mathcal{X}_j\}_{j=1}^p$ simple periodic invariant subspaces.

From the periodic Schur decomposition in Theorem 1.1, we also obtain the periodic Kronecker canonical form [5–7] of $\{(A_j,E_j)\}_{j=1}^p$.

Theorem 1.2 (*Periodic Kronecker Canonical Form*). *Suppose that the periodic matrix pairs $\{(A_j,\ E_j)\}_{j=1}^p$ are regular. Then there exist nonsingular matrices X_j and Y_j ($j=1,2,\ldots,p$) such that*

$$Y_j^HE_jX_j=\begin{bmatrix}I & 0\\ 0 & E_j^0\end{bmatrix},\qquad Y_j^HA_jX_{j-1}=\begin{bmatrix}A_j^f & 0\\ 0 & I\end{bmatrix},\tag{8}$$

where A_j^f and E_j^0 are all upper triangular,

$$J^{(j)} \equiv A_{j+p-1}^f A_{j+p-2}^f \dots A_j^f \quad (j = 1, 2, \dots, p) \tag{9}$$

are Jordan canonical forms corresponding to the finite eigenvalues of $\{(A_j,\ E_j)\}_{j=1}^p$*, and*

$$N^{(j)} \equiv E_j^0 E_{j+1}^0 \dots E_{j+p-1}^0 \quad (j = 1, 2, \dots, p) \tag{10}$$

are nilpotent Jordan canonical forms corresponding to the infinite eigenvalues.

Remarks.

(i) From [4], the matrices A_j^f and E_j^0 in (8) can be further reduced to block-upper triangular. Each individual block in A_j^f or E_j^0 relates to the corresponding Jordan block of a multiple eigenvalue of $\{(A_j,\ E_j)\}_{j=1}^p$.

(ii) For different values of j, the Jordan canonical forms $J^{(j)}$ and $N^{(j)}$ in (9) and (10) may have different structures. Thus, an eigenvalue with a certain algebraic multiplicity may have different geometric multiplicities dependent on j.

The eigen-problem of the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$ reflects the behavior of the linear discrete-time periodic systems

$$E_j x_{j+1} = A_j x_j \quad (j = 1, 2, \dots, p) \tag{11}$$

with respect to solvability and stability [8–12]. There has been much recent interest in periodic systems. It arises in a large variety of applications, including queueing network [13,14], analysis of bifurcations and computation of multipliers [15,16], multirate sampled-data systems, chemical processes, periodic time-varying filters and networks and seasonal phenomena; see [8,9] and the references therein for further information. Note that the periodic matrix eigen-problem is mathematically equivalent to the product matrix eigen-problem and the cyclic matrix eigen-problem [17,18]. Recently, some reliable numerical algorithms have been designed for the computation of the periodic stable invariant subspaces [1,19]. Perturbation analysis of eigenvalues and periodic deflating subspaces of periodic matrix pairs have been extensively studied in [20,5,4,21]. For the large product matrix eigen-problems and the periodic matrix eigen-problems with $E_j = I$ $(j = 1, 2, \dots, p)$, Kressner [17] presents a periodic Arnoldi process that generates orthonormal bases of certain periodic Krylov subspaces. Based on it, he proposes a periodic Arnoldi method for the product matrix eigen-problem and develops a periodic Arnoldi algorithm and a periodic Krylov–Schur algorithm.

The Rayleigh–Ritz method is widely used for the computation of approximations to an eigen-space $\mathcal{X}$ of an ordinary large matrix eigen-problem $Ax = \lambda x$, from an approximating subspace $\tilde{\mathcal{X}}$. The harmonic Rayleigh–Ritz method is an alternative for solving the interior eigen-problem (see, e.g., [22, Chapter 4]). Furthermore, when one is concerned with eigenvalues and eigenvectors, one can compute certain refined (harmonic) Ritz vectors whose convergence is guaranteed [23–27]; see also [22].

The purpose of this paper is to generalize the concept of the Rayleigh–Ritz approximation for the periodic matrix pairs, leading to the periodic Rayleigh–Ritz approximation. We study the convergence of the periodic Ritz values and the corresponding periodic Ritz vectors and extend some of the results in [26,27,22] to the periodic Rayleigh–Ritz approximation. Similar to the ordinary eigen-problem case (when $p = 1$) in [26,27,22], periodic Ritz vectors may fail to converge even if the corresponding periodic projection subspaces contain sufficiently accurate approximations to the desired periodic eigenvectors. It is thus necessary to refine the periodic Ritz vectors, as described in Section 5. We shall prove the convergence of the refined periodic Ritz vectors and propose an algorithm for their computation. All the convergence results are nontrivial generalizations of some of the known ones for Rayleigh–Ritz approximations and their refinement for the ordinary eigenvalue problem in [26,27]; see also [22]. As an important special case when the periodic Arnoldi process [17] is employed to generate the periodic orthonormal bases of the periodic Krylov subspaces, the refinement can be realized much more efficiently.

In the rest of the paper, $\|\cdot\|$ denotes both the Euclidean vector norm and the subordinate spectral matrix norm, unless otherwise stated. The conjugate of a complex number α is denoted by $\bar{\alpha}$ and the unit imaginary number is denoted by $\iota = \sqrt{-1}$.

The paper is organized as follows. We first consider the Rayleigh–Ritz procedure for the periodic eigen-problem (1) in Section 2. The convergence of the Ritz value pairs and their corresponding periodic Ritz vectors will be treated in Sections 3 and 4, respectively. In Section 5, we shall establish the convergence of the refined periodic Ritz vectors and propose a numerical method to compute them. In Section 6, we consider the special case when the periodic Krylov subspaces are generated by the periodic Arnoldi process. In Section 7, some numerical examples are given to illustrate the accuracy of the refined periodic Ritz vectors and the sharpness of their convergence bounds. The paper concludes with a brief summary in Section 8.

2. The periodic Rayleigh–Ritz approximation

As is known [17,18], the eigen-problem of the periodic matrices $\{A_j\}_{j=1}^p$ is very closely related to the product matrix eigen-problem and the cyclic matrix eigen-problem. Recently, based on the periodic Arnoldi process, a periodic Arnoldi algorithm and its Krylov–Schur version have been developed for solving eigenvalue problems associated with products of large and sparse matrices [17]. One of the central problems in this method is how to extract approximations to the desired eigenvalues and periodic eigenvectors from the given periodic subspaces $\{\tilde{\mathcal{X}}_j\}_{j=1}^p$. The algorithm is based on a variant of the Rayleigh–Ritz procedure applied to the eigen-problems (1) for the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$. It performs restarts and deflations via reordered periodic Schur decompositions and generates an approximate sequence of periodic subspaces $\{\tilde{\mathcal{X}}_j\}_{j=1}^p$ containing increasingly accurate approximations to the desired periodic eigenvectors.

For the periodic subspaces $\{\tilde{\mathcal{X}}_j\}_{j=1}^p$, suppose that they are spanned by the periodic orthonormal bases $\{U_j\}_{j=1}^p$ with $\dim(\tilde{\mathcal{X}}_j) = k$ $(j = 1, 2, \ldots, p)$. Compute the (thin or compact) QR-decompositions

$$E_jU_j = V_jN_j \quad (j = 1, 2, \ldots, p) \tag{12}$$

where $V_j^HV_j = I_k$ and N_j is upper triangular. Let, for all j,

$$V_j^HA_jU_{j-1} = M_j. \tag{13}$$

Then (12) and (13) define the periodic Rayleigh–Ritz pairs $\{(M_j, N_j)\}_{j=1}^p$ with respect to $\{U_j\}_{j=1}^p$. The following theorem shows that for any periodic orthonormal bases $\{U_j\}_{j=1}^p$, the periodic Rayleigh–Ritz pairs yield minimal residuals.

Theorem 2.1. *Let $\{(M_j, N_j)\}_{j=1}^p$ be the periodic Rayleigh–Ritz pairs with respect to the periodic bases $\{U_j\}_{j=1}^p$. Suppose that N_j is nonsingular for all j. Then the residuals*

$$R_j \equiv A_jU_{j-1} - E_jU_j(N_j^{-1}M_j) \quad (j = 1, 2, \ldots, p) \tag{14}$$

are minimal in the matrix 2-norm:

$$\min_{C_j\in\mathbb{C}^{k\times k}} \|A_jU_{j-1} - E_jU_jC_j\| = \|R_j\|. \tag{15}$$

Proof. Let $P_j \equiv N_j^{-1}M_j$ $(j = 1, 2, \ldots, p)$. For any $C_j \in \mathbb{C}^{k\times k}$, denote $\Delta_j \equiv P_j - C_j$. From $(I - V_jV_j^H)E_jU_j = 0$ and $C_j^HU_j^HE_j^HE_jU_jP_j = C_j^HU_j^HE_j^HA_jU_{j-1}$, we have

$$\begin{aligned}\|A_jU_{j-1} - E_jU_jC_j\|^2 &= \rho(U_{j-1}^HA_j^HA_jU_{j-1} - C_j^HU_j^HE_j^HA_jU_{j-1} - U_{j-1}^HA_j^HE_jU_jC_j + C_j^HU_j^HE_j^HE_jU_jC_j)\\ &= \rho(U_{j-1}^HA_j^HA_jU_{j-1} + \Delta_j^HU_j^HE_j^HE_jU_j\Delta_j - P_j^HU_j^HE_j^HE_jU_jP_j)\\ &= \rho[(U_{j-1}^HA_j^H - P_j^HU_j^HE_j^H)(A_jU_{j-1} - E_jU_jP_j) + \Delta_j^HU_j^HE_j^HE_jU_j\Delta_j]\\ &\geq \rho(R_j^HR_j) = \|R_j\|^2,\end{aligned}$$

where $\rho(\cdot)$ denotes the spectral radius. □

Remark. The residuals R_j's in (14) possess the following geometric meaning

$$\begin{aligned}R_j &= A_jU_{j-1} - E_jU_j(V_j^HE_jU_j)^{-1}(V_j^HA_jU_{j-1})\\ &= [I - E_jU_j(V_j^HE_jU_j)^{-1}V_j^H]A_jU_{j-1} = (I - P_{E_jU_j})A_jU_{j-1},\end{aligned}$$

where $P_{E_jU_j}$ is the orthogonal projector onto the subspace span(E_jU_j). Furthermore, if N_j is nonsingular, it is easily verified that $P_{E_jU_j} = V_jV_j^H$. So $\|R_j\|$ is the distance of A_jU_{j-1} from span(V_j) and should be minimal over all projections of A_jU_{j-1} onto span$(V_j) = (E_jU_j)$.

We now describe the periodic Rayleigh–Ritz (pRR) procedure with respect to $\{U_j\}_{j=1}^p$ to approximate an eigen-pair $((\pi_\alpha, \pi_\beta); \{x_j\}_{j=1}^p)$ of the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$.

(i) Construct the periodic orthonormal bases $\{U_j\}_{j=1}^p$, where $U_j \in \mathbb{C}^{n\times k}$.
(ii) Compute the QR-decompositions $E_jU_j = V_jN_j$ with $V_j^HV_j = I_k$ $(j = 1, 2, \ldots, p)$.
(iii) Compute $M_j = V_j^HA_jU_{j-1}$ $(j = 1, 2, \ldots, p)$.
(iv) Compute a desired eigenvalue pair $(\pi_\mu, \pi_\nu) \equiv \left(\prod_{j=1}^p \mu_j, \prod_{j=1}^p \nu_j\right)$ of the periodic Rayleigh–Ritz matrix pairs $\{(M_j, N_j)\}_{j=1}^p$ and the corresponding periodic right eigenvectors $\{z_j\}_{j=1}^p$ with $\|z_j\| = 1$, by using the periodic QZ algorithm with eigenvalue reordering techniques [1,2], such that

$$\nu_jM_jz_{j-1} = \mu_jN_jz_j \quad (j = 1, 2, \ldots, p).$$

(v) With $\tilde{x}_j \equiv U_jz_j$, take the Ritz value pair and periodic Ritz vectors $((\pi_\mu, \pi_\nu); \{\tilde{x}_j\}_{j=1}^p)$ as an approximate eigenvalue pair and periodic eigenvectors.

3. Convergence of Ritz value pairs

Let $(\pi_\alpha, \pi_\beta) \equiv \left(\prod_{j=1}^p \alpha_j, \prod_{j=1}^p \beta_j\right)$ be a simple eigenvalue pair of $\{(A_j, E_j)\}_{j=1}^p$ and $\{x_j\}_{j=1}^p$ be the corresponding periodic right eigenvectors with $\|x_j\| = 1$ $(j = 1, 2, \dots, p)$. That is, we have

$$\begin{cases} \beta_j A_j x_{j-1} = \alpha_j E_j x_j, \\ \|x_j\| = 1 \quad (j = 1, 2, \dots, p). \end{cases} \tag{16}$$

We assume that the periodic subspaces $\{\tilde{\mathcal{X}}_j\}_{j=1}^p$ contain accurate approximations to the periodic eigenvectors $\{x_j\}_{j=1}^p$. For given periodic orthonormal bases $\{U_j\}_{j=1}^p$ with $[U_j, U_j^\perp]$ being unitary, we define, for all j,

$$\theta_j = \angle(x_j, \tilde{\mathcal{X}}_j) \tag{17}$$

$$v_j = U_j^H x_j, \qquad v_j^\perp = (U_j^\perp)^H x_j. \tag{18}$$

Then it holds for all j that

$$\|v_j^\perp\| = \sin\theta_j, \qquad \|v_j\| = \sqrt{1 - \sin^2\theta_j} = \cos\theta_j, \tag{19}$$

assuming without loss of generality that all θ_j are in the first quadrant. We now show that the spectrum of the periodic Rayleigh–Ritz matrix pairs

$$(M_j, N_j) = (V_j^H A_j U_{j-1}, V_j^H E_j U_j) \quad (j = 1, 2, \dots, p) \tag{20}$$

obtained by (iv) in the pRR approximation contains a Ritz value pair $(\pi_\mu, \pi_\nu) \equiv \left(\prod_{j=1}^p \mu_j, \prod_{j=1}^p \nu_j\right)$ that converges to (π_α, π_β) when $\sin\theta_j \to 0$ for all j.

Theorem 3.1. *Let $\{(M_j, N_j)\}_{j=1}^p$ be the periodic Rayleigh–Ritz matrix pairs defined by (20). Then for all j, there exist matrices $\mathcal{E}_{M_j}$ and $\mathcal{E}_{N_j}$ which satisfy*

$$\|\mathcal{E}_{M_j}\| \le \frac{|\beta_j|}{|\alpha_j|^2 + |\beta_j|^2} \min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\} \tag{21}$$

and

$$\|\mathcal{E}_{N_j}\| \le \frac{|\alpha_j|}{|\alpha_j|^2 + |\beta_j|^2} \min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\} \tag{22}$$

with

$$\epsilon_j^{(1)} = |\alpha_j|\|E_j\| \left|1 - \frac{\cos\theta_j}{\cos\theta_{j-1}}\right| + |\alpha_j|\|E_j\| \frac{\sin\theta_j}{\cos\theta_{j-1}} + |\beta_j|\|A_j\| \tan\theta_{j-1} \tag{23}$$

and

$$\epsilon_j^{(2)} = |\beta_j|\|E_j\| \left|1 - \frac{\cos\theta_{j-1}}{\cos\theta_j}\right| + |\alpha_j|\|E_j\| \tan\theta_j + |\beta_j|\|A_j\| \frac{\sin\theta_{j-1}}{\cos\theta_j} \tag{24}$$

such that (π_α, π_β) is an eigenvalue pair of the periodic matrix pairs $\{(M_j + \mathcal{E}_{M_j}, N_j + \mathcal{E}_{N_j})\}_{j=1}^p$.

Proof. As $[U_j, U_j^\perp]$ $(j = 1, 2, \dots, p)$ are unitary, pre-multiplying the equations in (16) by V_j^H produces

$$\beta_j V_j^H A_j [U_{j-1}, U_{j-1}^\perp] \begin{bmatrix} U_{j-1}^H \\ (U_{j-1}^\perp)^H \end{bmatrix} x_{j-1} - \alpha_j V_j^H E_j [U_j, U_j^\perp] \begin{bmatrix} U_j^H \\ (U_j^\perp)^H \end{bmatrix} x_j = 0.$$

From (18) and (20), it follows that

$$\beta_j (M_j v_{j-1} + V_j^H A_j U_{j-1}^\perp v_{j-1}^\perp) - \alpha_j (N_j v_j + V_j^H E_j U_j^\perp v_j^\perp) = 0. \tag{25}$$

Let $\hat{v}_j \equiv v_j / \|v_j\|$ $(j = 1, 2, \dots, p)$. Dividing (25) by $\|v_{j-1}\|$, we obtain, for all j,

$$\beta_j M_j \hat{v}_{j-1} - \alpha_j N_j \hat{v}_j = \alpha_j N_j \hat{v}_j \frac{\|v_j\|}{\|v_{j-1}\|} - \alpha_j N_j \hat{v}_j + \alpha_j V_j^H E_j U_j^\perp \frac{v_j^\perp}{\|v_{j-1}\|} - \beta_j V_j^H A_j U_{j-1}^\perp \frac{v_{j-1}^\perp}{\|v_{j-1}\|}. \tag{26}$$

If we define the residuals

$$r_j \equiv \beta_j M_j \hat{v}_{j-1} - \alpha_j N_j \hat{v}_j \quad (j = 1, 2, \ldots, p), \tag{27}$$

then (26), (19) and (23) imply $\|r_j\| \le \epsilon_j^{(1)}$. Similarly, dividing (25) by $\|v_j\|$ yields $\|r_j\| \le \epsilon_j^{(2)}$, and consequently $\|r_j\| \le \min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\}$. Next we define, for all j,

$$\mathcal{E}_{M_j} \equiv \frac{-\bar{\beta}_j}{|\alpha_j|^2 + |\beta_j|^2} r_j \hat{v}_{j-1}^H, \qquad \mathcal{E}_{N_j} \equiv \frac{\bar{\alpha}_j}{|\alpha_j|^2 + |\beta_j|^2} r_j \hat{v}_j^H. \tag{28}$$

It then follows from (27) and (28) that

$$\alpha_j (N_j + \mathcal{E}_{N_j}) \hat{v}_j = \beta_j (M_j + \mathcal{E}_{M_j}) \hat{v}_{j-1} \quad (j = 1, 2, \ldots, p) \tag{29}$$

with $\mathcal{E}_{M_j}$ and $\mathcal{E}_{N_j}$ satisfying (21) and (22) by construction. □

Remark.

(i) Though $\epsilon_j^{(1)}, \epsilon_j^{(2)} \to 0$ as $\theta_j \to 0$ for $j = 1, 2, \ldots, p$, they are somehow complex and less clear. We shall simplify them, first by defining $\epsilon = \max_{j=1,2,\ldots,p} \sin\theta_j$. Applying Taylor expansions and observing that

$$\left|1 - \frac{\cos\theta_j}{\cos\theta_{j-1}}\right|, \qquad \left|1 - \frac{\cos\theta_{j-1}}{\cos\theta_j}\right| = O(\epsilon^2),$$

we obtain, by ignoring higher order small terms,

$$\epsilon_j^{(1)}, \epsilon_j^{(2)} \le (|\alpha_j| \, \|E_j\| + |\beta_j| \, \|A_j\|)\epsilon. \tag{30}$$

From Theorem 3.1 and the continuity of the eigenvalues of $\{(M_j, N_j)\}_{j=1}^p$, we immediately have the following corollary.

Corollary 3.2. *There exists a Ritz value pair (π_μ, π_ν) that converges to the simple eigenvalue pair (π_α, π_β) when $\sin\theta_j \to 0$ for all j.*

4. Convergence of periodic Ritz vectors

From Theorem 1.1, there are unitary matrices $[x_j, X_j]$ and $[y_j, Y_j]$, with $X_j, Y_j \in \mathbb{C}^{n\times(n-1)}$, such that

$$\begin{bmatrix} y_j^H \\ Y_j^H \end{bmatrix} A_j [x_{j-1}, X_{j-1}] = \begin{bmatrix} \alpha_j & l_j^H \\ 0 & L_j \end{bmatrix}, \qquad \begin{bmatrix} y_j^H \\ Y_j^H \end{bmatrix} E_j [x_j, X_j] = \begin{bmatrix} \beta_j & k_j^H \\ 0 & K_j \end{bmatrix}, \tag{31}$$

where the matrices L_j and K_j are $(n-1) \times (n-1)$ for $j = 1, 2, \ldots, p$. The periodic eigenvalue pairs of the periodic matrix pairs $\{(L_j, K_j)\}_{j=1}^p$ are the periodic eigenvalue pairs of $\{(A_j, E_j)\}_{j=1}^p$ other than (π_α, π_β). Also, (31) implies the spectral decompositions

$$A_j = \alpha_j y_j x_{j-1}^H + y_j l_j^H X_{j-1}^H + Y_j L_j X_{j-1}^H \tag{32}$$

and

$$E_j = \beta_j y_j x_j^H + y_j k_j^H X_j^H + Y_j K_j X_j^H. \tag{33}$$

For any approximate eigen-pair, we have the following residual bound for the approximate eigenvectors.

Theorem 4.1. *Let $\{(A_j, E_j)\}_{j=1}^p$ have the spectral representations (32) and (33) with $[x_j, X_j]$ and $[y_j, Y_j]$ being unitary for all j, $((\pi_{\tilde{\alpha}}, \pi_{\tilde{\beta}}); \{\tilde{x}_j\}_{j=1}^p)$ be an approximation to the simple eigen-pair $((\pi_\alpha, \pi_\beta); \{x_j\}_{j=1}^p)$,*

$$\tau_j \equiv \tilde{\alpha}_j E_j \tilde{x}_j - \tilde{\beta}_j A_j \tilde{x}_{j-1} \quad (j = 1, 2, \ldots, p), \tag{34}$$

and

$$\operatorname{sep}((\pi_{\tilde{\alpha}}, \pi_{\tilde{\beta}}), \{(L_j, K_j)\}_{j=1}^p) \equiv \|\mathcal{C}^{-1}\|^{-1} \tag{35}$$

with

$$\mathcal{C} \equiv \begin{bmatrix} \tilde{\alpha}_1 K_1 & & & -\tilde{\beta}_1 L_1 \\ -\tilde{\beta}_2 L_2 & \tilde{\alpha}_2 K_2 & & \\ & \ddots & \ddots & \\ & & -\tilde{\beta}_p L_p & \tilde{\alpha}_p K_p \end{bmatrix}.$$

If $\mathrm{sep}((\pi_{\tilde{\alpha}}, \pi_{\tilde{\beta}}), \{(L_j, K_j)\}_{j=1}^p) > 0$, *then*

$$\sqrt{\sum_{j=1}^{p} \sin^2 \angle(x_j, \tilde{x}_j)} \le \frac{\sqrt{\sum_{j=1}^{p} \|\tau_j\|^2}}{\mathrm{sep}((\pi_{\tilde{\alpha}}, \pi_{\tilde{\beta}}), \{(L_j, K_j)\}_{j=1}^p)}. \tag{36}$$

Proof. Pre-multiplying (34) by Y_j^H, we get, with the help of (32) and (33),

$$Y_j^H \tau_j = \tilde{\alpha}_j Y_j^H E_j \tilde{x}_j - \tilde{\beta}_j Y_j^H A_j \tilde{x}_{j-1} = \tilde{\alpha}_j K_j X_j^H \tilde{x}_j - \tilde{\beta}_j L_j X_{j-1}^H \tilde{x}_{j-1}. \tag{37}$$

This implies

$$\mathcal{C} \begin{bmatrix} X_1^H \tilde{x}_1 \\ \vdots \\ X_p^H \tilde{x}_p \end{bmatrix} = \begin{bmatrix} Y_1^H \tau_1 \\ \vdots \\ Y_p^H \tau_p \end{bmatrix}. \tag{38}$$

Note that $\mathcal{C}$ is invertible in the neighborhood of (π_α, π_β) if and only if the eigenvalue pair (π_α, π_β) is simple. As $[x_j, X_j]$ is unitary, we have $\sin \angle(x_j, \tilde{x}_j) \equiv \|X_j^H \tilde{x}_j\|$ for all j. Note that $\|Y_j^H \tau_j\| \le \|\tau_j\|$ for all j. The theorem then follows from inverting $\mathcal{C}$ in (38) and taking norms. □

Theorem 4.1 leads easily to the following corollary.

Corollary 4.2. *For* $j = 1, 2, \ldots, p$, *we have*

$$\sin \angle(x_j, \tilde{x}_j) \le \frac{\sqrt{p} \max_{j=1,2,\ldots,p} \|\tau_j\|}{\mathrm{sep}((\pi_{\tilde{\alpha}}, \pi_{\tilde{\beta}}), \{(L_j, K_j)\}_{j=1}^p)}. \tag{39}$$

In Corollary 3.2, we see that there is a Ritz value pair (π_μ, π_ν) approaching the simple eigenvalue pair (π_α, π_β) when $\sin \theta_j \to 0$ for all j. If, in addition, the p residual norms $\|\tau_j\|$ $(j = 1, 2, \ldots, p)$ defined in (34) approach zero, the periodic Ritz vectors $\{\tilde{x}_j\}_{j=1}^p$ converge to the periodic right eigenvectors $\{x_j\}_{j=1}^p$. Thus, Theorem 4.1 and Corollary 4.2 show that a converging Ritz value pair and vanishing residuals imply the convergence of the periodic Ritz vectors since $\|\mathcal{C}^{-1}\|$ is uniformly bounded when (π_μ, π_ν) converges to the simple eigenvalue (π_α, π_β).

When $p = 1$, it has been proved that the Ritz vector may fail to converge for a (nearly) multiple Ritz value (see, e.g. [26, 27]). We now perform a convergence analysis of the periodic Ritz vectors and establish some a priori error bounds, showing why the periodic Ritz vectors can fail to converge. Let the periodic Ritz pair $((\pi_\mu, \pi_\nu); \{\tilde{x}_j\}_{j=1}^p)$ be used to approximate the simple periodic eigen-pair $((\pi_\alpha, \pi_\beta); \{x_j\}_{j=1}^p)$.

Again, from Theorem 1.1, there are unitary matrices $[z_j, Z_j]$ and $[w_j, W_j]$, with $Z_j, W_j \in \mathbb{C}^{r \times (r-1)}$, such that

$$\begin{bmatrix} w_j^H \\ W_j^H \end{bmatrix} M_j [z_{j-1}, Z_{j-1}] = \begin{bmatrix} \mu_j & d_j^H \\ 0 & D_j \end{bmatrix}, \qquad \begin{bmatrix} w_j^H \\ W_j^H \end{bmatrix} N_j [z_j, Z_j] = \begin{bmatrix} \nu_j & f_j^H \\ 0 & F_j \end{bmatrix}, \tag{40}$$

where the matrices D_j and F_j are $(k-1) \times (k-1)$ for $j = 1, 2, \ldots, p$.

Since the only assumption on $\{\mathcal{X}_j\}_{j=1}^p$ is that they contain accurate approximations to the periodic eigenvectors $\{\tilde{x}_j\}_{j=1}^p$, the eigenvalue pairs of $\{(D_j, F_j)\}_{j=1}^p$ are not necessarily near the eigenvalue pairs of $\{(A_j, E_j)\}_{j=1}^p$ rather than (π_μ, π_ν). Particularly, this means that an eigenvalue pair of $\{(D_j, F_j)\}_{j=1}^p$ could be arbitrarily near and even equal to the Ritz value pair (π_μ, π_ν). For a multiple (π_μ, π_ν), there are more than one $\{\tilde{x}_j\}_{j=1}^p$ to approximate the unique periodic eigenvectors $\{x_j\}_{j=1}^p$. It will be impossible for the periodic Rayleigh–Ritz method to tell which particular approximation is better. If (π_μ, π_ν) is near an eigenvalue of $\{(D_j, F_j)\}_{j=1}^p$, we will get a unique periodic $\{\tilde{x}\}_{j=1}^p$, but there is no guarantee that it converges to $\{x_j\}_{j=1}^p$.

The above analysis leads us to postulate that the periodic Ritz vectors $\{\tilde{x}\}_{j=1}^p$ will converge provided that (π_μ, π_ν) is uniformly away from those eigenvalues (other Ritz values) of $\{(D_j, F_j)\}_{j=1}^p$, independent of θ_j, $j = 1, 2, \ldots, p$. We next prove that this is indeed the case quantitatively.

Theorem 4.3. *Assume that the periodic Rayleigh–Ritz pairs* $\{(M_j, N_j)\}_{j=1}^p$ *have the spectral decompositions* (40) *and*

$$\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p) \equiv \|\hat{\mathcal{C}}^{-1}\|^{-1} > 0 \tag{41}$$

with

$$\hat{\mathcal{C}} \equiv \begin{bmatrix} \alpha_1 F_1 & & & -\beta_1 D_1 \\ -\beta_2 D_2 & \alpha_2 F_2 & & \\ & \ddots & \ddots & \\ & & -\beta_p D_p & \alpha_p K_p \end{bmatrix}.$$

Let $\epsilon = \max_{j=1,2,\ldots,p} \sin\theta_j$. *Then for* $j = 1, 2, \ldots, p$, *we have*

$$\sin\angle(x_j, \tilde{x}_j) \le \sin\theta_j + \frac{\sqrt{p}\max\{\min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\}\}}{\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p)} \tag{42}$$

$$\le \left(1 + \frac{\sqrt{p}(|\alpha_j|\,\|E_j\| + |\beta_j|\,\|A_j\|)}{\mathrm{sep}((\pi_\alpha, \pi_\beta),\ \{(D_j, F_j)\}_{j=1}^p)}\right)\epsilon \tag{43}$$

with $\epsilon_j^{(1)}, \epsilon_j^{(2)}$ *defined as in (23) and (24).*

Proof. Let the periodic Ritz pair $((\pi_\mu, \pi_\nu); \{\tilde{x}_j\}_{j=1}^p)$ be an approximation to the periodic eigen-pair $((\pi_\alpha, \pi_\beta); \{x_j\}_{j=1}^p)$. As in the proof of Theorem 3.1, let $\hat{v}_j = U_j^H x_j / \|U_j^H x_j\|$. Then we get

$$r_j \equiv \beta_j M_j \hat{v}_{j-1} - \alpha_j N_j \hat{v}_j \quad (j = 1, 2, \ldots, p),$$

which is defined by (26). It is seen from the proof of Theorem 3.1 that

$$\|r_j\| \le \min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\}$$

with $\epsilon_j^{(1)}, \epsilon_j^{(2)}$ in (23) and (24).

Note that we can regard $((\pi_\alpha, \pi_\beta); \{\hat{v}_j\}_{j=1}^p)$ as an approximate periodic eigen-pair to the periodic eigen-pair $((\pi_\mu, \pi_\nu); \{z_j\}_{j=1}^p)$ of $\{(M_j, N_j)\}_{j=1}^p$. Then from Corollary 4.2, it follows for $j = 1, 2, \ldots, p$ that

$$\sin\angle(z_j, \hat{v}_j) \le \frac{\sqrt{p}\max\limits_{j=1,2,\ldots,p}\|r_j\|}{\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p)} \le \frac{\sqrt{p}\max\limits_{j=1,2,\ldots,p}\min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\}}{\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p)}.$$

Since U_j is orthonormal for $j = 1, 2, \ldots, p$, we have from the definitions of $\tilde{x}_j$, $\hat{v}_j$ and θ_j that

$$\sin\angle(z_j, \hat{v}_j) = \sin\angle(U_j z_j, U_j \hat{v}_j) = \sin\angle(\tilde{x}_j, U_j U_j^H x_j).$$

Note the triangle inequality

$$\angle(x_j, \tilde{x}_j) \le \angle(x_j, U_j U_j^H x_j) + \angle(U_j U_j^H x_j, \tilde{x}_j) = \angle(x_j, \tilde{\mathcal{X}}_j) + \angle(\tilde{x}_j, U_j U_j^H x_j)$$

with the equality holding when the vectors x_j, $\tilde{x}_j$ and $U_j U_j^H x_j$ are linearly dependent. Therefore, we get

$$\sin\angle(x_j, \tilde{x}_j) \le \sin\theta_j + \sin\angle(z_j, \hat{v}_j) \le \sin\theta_j + \frac{\sqrt{p}\max\limits_{j=1,2,\ldots,p}\{\min\{\epsilon_j^{(1)}, \epsilon_j^{(2)}\}\}}{\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p)},$$

which proves (42). Furthermore, from (30) we get (43). □

From Theorem 3.1, since the Ritz value pair (π_μ, π_ν) approaches the eigenvalue pair (π_α, π_β) as $\theta_j \to 0$ for $j = 1, 2, \ldots, p$, by the continuity argument we have

$$\mathrm{sep}((\pi_\alpha, \pi_\beta), \{(D_j, F_j)\}_{j=1}^p) \to \mathrm{sep}((\pi_\mu, \pi_\nu), \{(D_j, F_j)\}_{j=1}^p).$$

We see that a sufficient condition for the convergence of the periodic Ritz vectors $\{\tilde{x}_j\}_{j=1}^p$ is that $\mathrm{sep}((\pi_\mu, \pi_\nu), \{(D_j, F_j)\}_{j=1}^p)$ is uniformly bounded away from zero. This condition can be checked during the procedure. However, as we have argued above, $\mathrm{sep}((\pi_\mu, \pi_\nu),\ \{(D_j, F_j)\}_{j=1}^p)$ can be arbitrarily small (and even be exactly zero) when (π_μ, π_ν) is arbitrarily near other eigenvalue pairs (or is associated with a multiple eigenvalue pair) of $\{(M_j, N_j)\}_{j=1}^p$. Consequently, while the periodic Ritz value pair converges unconditionally once $\theta_j \to 0$ for $j = 1, 2, \ldots, p$, its corresponding periodic Ritz vectors may fail to converge or may converge very slowly or irregularly.

5. Refinement of periodic Ritz vectors

As we have seen, the periodic Ritz vectors may fail to converge. Since the Ritz value pair is known to converge to the simple eigenvalue pair (π_α, π_β) when $\sin\theta_j \to 0$ for all j, this suggests that we can deal with non-converging Ritz vectors by retaining the Ritz value pair while replacing the periodic Ritz vectors with a set of unit vectors $\hat{x}_j \in \tilde{\mathcal{X}}_j = \mathrm{span}(U_j)$ $(j = 1, 2, \ldots, p)$ with suitably small residuals. Thus, we construct $\hat{x}_j$ $(j = 1, 2, \ldots, p)$ from

$$\min_{\hat{x}_j} \sqrt{\sum_{j=1}^{p} \|\mu_j E_j \hat{x}_j - \nu_j A_j \hat{x}_{j-1}\|^2} \quad \text{subject to} \quad \hat{x}_j \in \mathrm{span}(U_j),\ \|\hat{x}_j\| = 1 \quad (j = 1, 2 \ldots, p). \tag{44}$$

We call the minimizer $\{\hat{x}_j\}_{j=1}^{p}$, the refined periodic Ritz vectors.

The following theorem shows that the refined periodic Ritz vectors converge when $\sin\theta_j \to 0$ for all j.

Theorem 5.1. *Let $\{(A_j, E_j)\}_{j=1}^{p}$ have spectral representations (32) and (33), where $\|x_j\| = 1$ for all j. Let $(\pi_\mu, \pi_\nu) \equiv (\prod_{j=1}^{p} \mu_j, \prod_{j=1}^{p} \nu_j)$ be a Ritz value pair with respect to the orthonormal bases $\{U_j\}_{j=1}^{p}$ and let $\{\hat{x}_j\}$ be the corresponding refined periodic Ritz vectors. If $\mathrm{sep}((\pi_\mu, \pi_\nu), \{(L_j, K_j)\}_{j=1}^{p}) > 0$, then*

$$\sin\angle(x_j, \hat{x}_j) \le \frac{\|\eta\|}{\mathrm{sep}((\pi_\mu, \pi_\nu), \{(L_j, K_j)\}_{j=1}^{p})} \quad (j = 1, 2, \ldots, p), \tag{45}$$

where $\eta = [\eta_1, \ldots, \eta_p]^\top$ and

$$\eta_j \equiv \frac{\rho_j + 2|\mu_j||\beta_j| \sin^2\frac{\theta_{j-1}}{2} + 2|\nu_j||\alpha_j| \sin^2\frac{\theta_j}{2}}{\cos\theta_{j-1}\cos\theta_j} + \frac{|\mu_j|\|E_j\| \sin\theta_j}{\cos\theta_j} + \frac{|\nu_j|\|A_j\| \sin\theta_{j-1}}{\cos\theta_{j-1}} \tag{46}$$

with $\rho_j \equiv |\mu_j\beta_j - \alpha_j\nu_j|$ for all j.

Proof. Let $x_j = q_j + q_j^\perp$, where $q_j = U_j U_j^H x_j$ and $q_j^\perp = (I - U_j U_j^H) x_j$ for all j. Then $\|q_j\| = \cos\theta_j$ and $\|q_j^\perp\| = \sin\theta_j$. Define the normalized vectors

$$\hat{q}_j \equiv \frac{q_j}{\|q_j\|} = \frac{q_j}{\cos\theta_j}, \quad j = 1, 2, \ldots, p. \tag{47}$$

By (47), the residuals $\hat{r}_j$ satisfy

$$\begin{aligned} \hat{r}_j &\equiv \mu_j E_j \hat{q}_j - \nu_j A_j \hat{q}_{j-1} \\ &= \frac{\mu_j E_j q_j}{\cos\theta_j} - \frac{\nu_j A_j q_{j-1}}{\cos\theta_{j-1}} = \frac{\mu_j E_j (x_j - q_j^\perp)}{\cos\theta_j} - \frac{\nu_j A_j (x_{j-1} - q_{j-1}^\perp)}{\cos\theta_{j-1}}. \end{aligned} \tag{48}$$

Denote the ith column of the identity matrix by e_i. Pre-multiplying (48) by the unitary matrix $\hat{Y}_j^H \equiv [y_j, Y_j]^H$ and using (32) and (33), we have, for all j,

$$\begin{aligned} \hat{Y}_j^H \hat{r}_j &= \frac{\mu_j\beta_j e_1}{\cos\theta_j} - \frac{\nu_j\alpha_j e_1}{\cos\theta_{j-1}} - \frac{\mu_j \hat{Y}_j^H E_j q_j^\perp}{\cos\theta_j} + \frac{\nu_j \hat{Y}_j^H A_j q_{j-1}^\perp}{\cos\theta_{j-1}} \\ &= \frac{\mu_j\beta_j\cos\theta_{j-1} - \nu_j\alpha_j\cos\theta_j}{\cos\theta_j\cos\theta_{j-1}} e_1 - \frac{\mu_j \hat{Y}_j^H E_j q_j^\perp}{\cos\theta_j} + \frac{\nu_j \hat{Y}_j^H A_j q_{j-1}^\perp}{\cos\theta_{j-1}}. \end{aligned} \tag{49}$$

Using the identity $\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$ and taking norm of (49), we obtain $\|\hat{r}_j\| \le \eta_j$ for all j. By the minimization in (44), we also have

$$\sum_{j=1}^{p} \|\mu_j E_j \hat{x}_j - \nu_j A_j \hat{x}_{j-1}\|^2 \le \sum_{j=1}^{p} \|\hat{r}_j\|^2 \le \|\eta\|^2. \tag{50}$$

The inequalities (45) then follow from Theorem 4.1 and (50). □

Since (π_μ, π_ν) converges to (π_α, π_β) as $\theta_j \to 0$ for $j = 1, 2, \ldots, p$, we have

$$\mathrm{sep}((\pi_\mu, \pi_\nu),\ \{(L_j, K_j)\}_{j=1}^{p}) \to \mathrm{sep}((\pi_\alpha, \pi_\beta), \{(L_j, K_j)\}_{j=1}^{p}),$$

which is a positive constant independent of the procedure, whenever (π_α, π_β) is a simple eigenvalue pair of $\{(A_j, E_j)\}_{j=1}^{p}$. So the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^{p}$ converge provided that $\|\eta\|$, i.e., ρ_j, $j = 1, 2, \ldots, p$ in (46), tends to zero.

Remarks. In order to ensure the convergence of ρ_j, we should renormalize the complex ordered pairs $\{(\alpha_j, \beta_j)\}_{j=1}^p$ and $\{(\mu_j, \nu_j)\}_{j=1}^p$ in Theorem 5.1 by periodic complex numbers of modulo one so that

(i)

$$\begin{cases} \alpha_j := |\alpha_j|, \quad \beta_j := |\beta_j| \quad (j = 1, 2, \ldots, p-1), \\ \alpha_p := |\alpha_p| e^{\iota\left[\sum_{j=1}^{p} \arg(\alpha_j) - \arg(\beta_j)\right]}, \quad \beta_p := |\beta_p|, \end{cases}$$

whenever $\pi_\alpha \neq 0$ and $\pi_\beta \neq 0$;

(ii)

$$\alpha_j := |\alpha_j|, \quad \beta_j := |\beta_j| \quad (j = 1, 2, \ldots, p),$$

whenever $\pi_\alpha = 0$ or $\pi_\beta = 0$. A similar renormalization for $\{(\mu_j, \nu_j)\}_{j=1}^p$ can also be carried out. With these new normalized ordered pairs $\{(\alpha_j, \beta_j)\}_{j=1}^p$ and $\{(\mu_j, \nu_j)\}_{j=1}^p$, by Theorem 3.1 and the periodic Bauer–Fike theorem [5], we have $\rho_j \to 0$ when $\sin\theta_j \to 0$ for all j. It follows from Theorem 5.1 that $\angle(x_j, \hat{x}_j) \to 0$; i.e., unlike the periodic Ritz vectors, the refined Ritz vectors are guaranteed to converge.

(iii) Again, let $\epsilon = \max_{j=1,2,\ldots,p} \sin\theta_j$. Then using Taylor expansions and ignoring higher order terms, we have

$$\eta_j \leq \rho_j + (|\mu_j| \, \|E_j\| + |\gamma_j| \, \|A_j\|)\epsilon. \tag{51}$$

We now propose a numerical procedure to compute the refined periodic Ritz vectors efficiently and reliably.

From (44), the set of refined periodic Ritz vectors can be computed via the following constrained minimization problem

$$\begin{aligned} &\min_{\hat{z}} \quad f(\hat{z}) \equiv \sum_{j=1}^{p} \|\mu_j E_j U_j \hat{z}_j - \nu_j A_j U_{j-1} \hat{z}_{j-1}\|^2 \\ &\text{subject to} \quad c_j(\hat{z}) \equiv \hat{z}_j^H \hat{z}_j - 1 = 0 \quad (j = 1, 2 \ldots, p), \end{aligned} \tag{52}$$

where $\hat{z} \equiv [\,\hat{z}_1^\top, \ldots, \hat{z}_p^\top\,]^\top \in \mathbb{C}^{kp}$. Newton's method can be applied to the Lagrangian function of the constrained optimization problem (52), with the periodic Ritz vectors utilized as the feasible initial iterate. An approximate solution to (52) will be acceptable in the sense of Theorems 3.1, 4.1 and 5.1 if the associated residuals are reasonably small.

Remarks. (i) For periodicity $p = 1$, the minimization problem (44) can be solved via the singular value decomposition (SVD). Indeed, as mentioned in [26,27], it is easily seen that the refined Ritz vector $\hat{x}_1 = U_1\hat{z}_1$, where $\hat{z}_1$ is the right singular vector of $(\mu_1 E_1 - \nu_1 A_1)U_1$ corresponding to its smallest singular value. Unfortunately, the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^p$ with periodicity $p \geq 2$ cannot be computed via (52) by any SVD-like algorithm since, instead of $\|[\,\hat{z}_1^\top, \hat{z}_2^\top, \ldots, \hat{z}_p^\top\,]^\top\| = 1$, the constraints $\|\hat{z}_j\| = 1$ $(j = 1, 2, \ldots, p)$ have to be satisfied simultaneously.

(ii) The Newton optimization of (52) is straightforward, but can be expensive since E_jU_j and A_jU_{j-1} $(j = 1, 2, \ldots, p)$, though already available when forming the periodic Rayleigh–Ritz pairs $\{(M_j, N_j)\}_{j=1}^p$, are $n \times k$. However, it is possible to reduce the optimization problem to a much smaller one when the Rayleigh–Ritz method is applied to certain special periodic Krylov subspaces. In Section 6, we will consider a periodic Arnoldi process that generates periodic orthonormal bases of the periodic Krylov subspaces. Based on it, we propose the refined periodic Arnoldi method and show that Newton optimization is particularly efficient.

6. Refined periodic Ritz vectors from a periodic Arnoldi process

Recall the eigen-equations in (1):

$$\beta_j A_j x_{j-1} = \alpha_j E_j x_j \quad (j = 1, 2, \ldots, p).$$

Without loss of generality, E_j can be assumed to be nonsingular, as a shift can always be applied to the periodic eigenvalue problem. To apply the Arnoldi process for matrix products [17] to our periodic matrix pairs, we may consider two different products

$$\mathbb{P}_l \equiv (E_p^{-1}A_p)(E_{p-1}^{-1}A_{p-1}) \cdots (E_1^{-1}A_1), \qquad \mathbb{P}_r \equiv (A_pE_{p-1}^{-1})(A_{p-1}E_{p-2}^{-1}) \cdots (A_1E_p^{-1}).$$

First, construct $\mathbb{A} \in \mathbb{R}^{np\times np}$ from C in (3) by substituting $\alpha_j = 0, \beta_j = -1$ $(j = 1, 2, \ldots, p)$. Similarly, denote by $\mathbb{E}$ the matrix constructed with $\alpha_j = 1, \beta_j = 0$ $(j = 1, 2, \ldots, p)$ in (3). Denote C in (3) slightly differently as $C(\mathbb{A}, \mathbb{E};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$, with $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_p]^\top$ and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]^\top$. From the equivalence of (1) and (2) to (4) and (5), respectively, we can see clearly that $C(\mathbb{A}, \mathbb{E};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$ defines the periodic eigenvalue problem under consideration. An appropriate shift σ can be applied to $C(\mathbb{A}, \mathbb{E};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$, producing an equivalent periodic eigenvalue problem defined by $C(\mathbb{A}, \mathbb{E} - \sigma\mathbb{A};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$. Obviously, an eigenvalue $(\pi_\alpha, \pi_\beta) = (\prod_{j=1}^p \alpha_j, \prod_{j=1}^p \beta_j)$ for $C(\mathbb{A}, \mathbb{E};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$ is transformed to $(\tilde{\pi}_\alpha, \tilde{\pi}_\beta) = \left(\prod_{j=1}^p \alpha_j, \prod_{j=1}^p (\beta_j + \sigma\alpha_j)\right)$ for $C(\mathbb{A}, \mathbb{E} - \sigma\mathbb{A};\ \boldsymbol{\alpha}, \boldsymbol{\beta})$, with identical eigenvectors and $\beta_j + \sigma\alpha_j \neq 0$ for all j.

Utilizing $\mathbb{P}_l$, the Arnoldi process is applied to the equivalent eigenvalue equations after inverting E_j:

$$\beta_j E_j^{-1} A_j x_{j-1} = \alpha_j x_j \quad (j = 1, 2, \ldots, p),$$

resulting in the refinement of the corresponding periodic Ritz vectors as summarized in (52).

Alternatively with $\mathbb{P}_r$, we consider another set of equivalent eigenvalue equations:

$$\begin{aligned} &\beta_j A_j E_{j-1}^{-1}(E_{j-1}x_{j-1}) = \alpha_j(E_j x_j) \\ &\Rightarrow \beta_j \mathcal{A}_j(E_{j-1}x_{j-1}) = \alpha_j(E_j x_j), \end{aligned} \tag{53}$$

where $\mathcal{A}_j \equiv A_j E_{j-1}^{-1}$ $(j = 1, 2, \ldots, p)$ and $\mathbb{P}_r = \prod_{k=p}^1 \mathcal{A}_k$.

With the k-step periodic Arnoldi process for $\{\mathcal{A}_j\}_{j=1}^p$, we have

$$\begin{aligned} &\mathcal{A}_1 U_p = U_1 H_1, \ldots, \mathcal{A}_j U_{j-1} = U_j H_j, \ldots, \\ &\mathcal{A}_p U_{p-1} = U_p H_p + h_{k+1,k} u_{k+1}^p e_k^\top, \end{aligned} \tag{54}$$

where $H_1, \ldots, H_{p-1} \in \mathbb{C}^{k\times k}$ are upper triangular and $H_p \in \mathbb{C}^{k\times k}$ is upper Hessenberg. Denote $\widehat{H}_p = \begin{bmatrix} H_p \\ h_{k+1,k}e_k^\top \end{bmatrix}$, we have

$$\mathcal{A}_p U_{p-1} = [U_p | u_{k+1}^p]\widehat{H}_p.$$

It is easy to show that $H_j = M_j N_j^{-1}$ $(j = 1, 2, \ldots, p)$, with M_j and N_j as defined in (20).

Without loss of generality, assume $\nu_j = 1$ $(\forall j)$. We then select $\hat{x}_j$ $(j = 1, 2, \ldots, p)$ from

$$\begin{aligned} &\min_{\hat{x}_j} \quad \sum_{j=1}^p \|A_j \hat{x}_{j-1} - \mu_j E_j \hat{x}_j\|^2 \\ &\text{subject to} \quad \hat{x}_j \in \operatorname{span}(E_j^{-1}U_j),\ \|\hat{x}_j\| = 1 \quad (j = 1, 2 \ldots, p). \end{aligned} \tag{55}$$

It is easy to show that the refinement in (55) is equivalent to the one in (44) with the periodic Arnoldi process providing $\{U_j\}$, when $\{E_j^{-1}U_j\}$ are orthogonalized. There is no reason in doing so because of the saving in reusing computed quantities from the periodic Arnoldi process, as shown below.

From (55), the set of refined periodic Ritz vectors can be computed via the following constrained minimization problem

$$\begin{aligned} &\min_{\hat{z}} \quad f(\hat{z}) \equiv \sum_{j=1}^p \|A_j(E_{j-1}^{-1}U_{j-1})\hat{z}_{j-1} - \mu_j E_j(E_j^{-1}U_j)\hat{z}_j\|^2 \\ &\text{subject to} \quad \|(E_j^{-1}U_j)\hat{z}_j\| = 1 \quad (j = 1, 2 \ldots, p), \end{aligned} \tag{56}$$

where $\hat{z} \equiv [\hat{z}_1^\top, \ldots, \hat{z}_p^\top]^\top \in \mathbb{C}^{kp}$. By (54), we have

$$\begin{aligned} &\min_{\hat{z}} \quad f(\hat{z}) \equiv \sum_{j=1}^p \|A_j(E_{j-1}^{-1}U_{j-1})\hat{z}_{j-1} - \mu_j E_j(E_j^{-1}U_j)\hat{z}_j\|^2 \\ &\Rightarrow \min_{\hat{z}} \quad f(\hat{z}) \equiv \sum_{j=1}^{p-1} \|U_j(H_j\hat{z}_{j-1} - \mu_j\hat{z}_j)\|^2 + \left\| [U_j|u_{k+1}] \left(\widehat{H}_p\hat{z}_{p-1} - \mu_p \begin{bmatrix} \hat{z}_p \\ 0 \end{bmatrix} \right) \right\|^2 \\ &\Rightarrow \min_{\hat{z}} \quad f(\hat{z}) \equiv \sum_{j=1}^{p-1} \|H_j\hat{z}_{j-1} - \mu_j\hat{z}_j\|^2 + \left\| \left(\widehat{H}_p\hat{z}_{p-1} - \mu_p \begin{bmatrix} \hat{z}_p \\ 0 \end{bmatrix} \right) \right\|^2. \end{aligned} \tag{57}$$

Then we consider the Lagrangian function of the constrained optimization problem (56):

$$L(\hat{z}, \lambda) = f(\hat{z}) + \sum_{j=1}^p \lambda_j(\hat{z}_j^H B_j \hat{z}_j - 1), \tag{58}$$

where $\lambda = [\lambda_1, \ldots, \lambda_p]^\top$, $B_j \equiv (E_j^{-1}U_j)^\top(E_j^{-1}U_j)$.

The derivatives of $L(\hat{z}, \lambda)$ are:

$$\begin{aligned} &f_1 \equiv \frac{\partial L}{\partial \hat{z}_1} = 2(\mu_1^2\hat{z}_1 - \mu_1 H_1\hat{z}_p + H_2^\top H_2\hat{z}_1 - \mu_2 H_2^\top\hat{z}_2 + \lambda_1 B_1\hat{z}_1), \\ &f_j \equiv \frac{\partial L}{\partial \hat{z}_j} = 2(\mu_j^2\hat{z}_j - \mu_j H_j\hat{z}_{j-1} + H_{j+1}^\top H_{j+1}\hat{z}_j - \mu_{j+1}H_{j+1}^\top\hat{z}_{j+1} + \lambda_j B_j\hat{z}_j) \quad (j = 2, \ldots, p-2), \end{aligned}$$

$$f_{p-1} \equiv \frac{\partial L}{\partial \hat{z}_{p-1}} = 2(\mu_{p-1}^2 \hat{z}_{p-1} - \mu_{p-1} H_{p-1} \hat{z}_{p-2} + \widehat{H}_p^\top \widehat{H}_p \hat{z}_{p-1} - \mu_p H_p^\top \hat{z}_p + \lambda_{p-1} B_{p-1} \hat{z}_{p-1}),$$
$$f_p \equiv \frac{\partial L}{\partial \hat{z}_p} = 2(\mu_p^2 \hat{z}_p - \mu_p H_p \hat{z}_{p-1} + H_1^\top H_1 \hat{z}_p - \mu_1 H_1^\top \hat{z}_1 + \lambda_p B_p \hat{z}_p);$$
$$f_{p+j} \equiv \frac{\partial L}{\partial \lambda_j} = \hat{z}_j^H B_j \hat{z}_j - 1, \quad (j = 1, 2, \ldots, p).$$

We then apply Newton's method to $f = [f_1^\top, \ldots, f_{2p}^\top]^\top = 0$, which can be formulated as

$$\begin{bmatrix} \hat{z} \\ \lambda \end{bmatrix}_{\text{new}} = \begin{bmatrix} \hat{z} \\ \lambda \end{bmatrix} - J_f^{-1} f, \qquad J_f = \begin{bmatrix} J_{11} & J_{12} \\ J_{12}^H & 0 \end{bmatrix},$$

where $J_{12} = 2(B_1 \widehat{z_1} \oplus \cdots \oplus B_p \widehat{z_p})$ and

$$J_{11} \equiv \begin{bmatrix} \Delta_1 & \widetilde{H}_2^\top & & & & & \widetilde{H}_1 \\ \widetilde{H}_2 & \Delta_2 & \widetilde{H}_3^\top & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \widetilde{H}_j & \Delta_j & \widetilde{H}_{j+1}^\top & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \widetilde{H}_{p-1} & \Delta_{p-1} & \widetilde{H}_p^\top \\ \widetilde{H}_1^\top & & & & & \widetilde{H}_p & \Delta_p \end{bmatrix},$$

with $\widetilde{H}_j \equiv -\mu_j H_j$ $(j = 1, 2, \ldots, p)$ and

$$\Delta_j = \mu_j^2 I_m + H_{j+1}^\top H_{j+1} + \lambda_j B_j \quad (j = 1, \ldots, p-2),$$
$$\Delta_{p-1} = \mu_{p-1}^2 I_m + \widehat{H}_p^\top \widehat{H}_p + \lambda_{p-1} B_{p-1},$$
$$\Delta_p = \mu_p^2 I_m + H_1^\top H_1 + \lambda_p B_p.$$

After convergence of the Newton optimization step, we obtain the refined $\{\hat{z}_j\}_{j=1}^p$, and in turn the refined Ritz vectors $\hat{x}_j = E_j^{-1} U_j \hat{z}_j$.

Note that the efficiency of the above Rayleigh–Ritz method with refinement comes from the fact that the matrices H_j, $B_j \equiv (E_j^{-1} U_j)^\top (E_j^{-1} U_j) \in \mathbb{R}^{k \times k}$ $(j = 1, 2, \ldots, p)$ and $\widehat{H}_p \in \mathbb{R}^{(k+1) \times k}$ are small in dimensions (relative to n). In addition, H_j and $E_j^{-1} U_j$ $(j = 1, 2, \ldots, p)$ are inherited from the periodic Arnoldi process, without further computations required.

7. Numerical examples

In the following numerical experiments, we shall illustrate the convergence of refined periodic Ritz vectors and the feasibility of our refinement strategy. All computations were performed in MATLAB/version 6.5 on a PC. The machine precision is approximately 2.22×10^{-16}.

Example 1. Consider the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$ with periodicity $p = 3$, with

$$A_j = \text{diag}(0, 1, -1), \qquad E_j = I_3, \quad j = 1, 2, 3.$$

It is easily seen that $x_j = [\,1, 0, 0\,]^\top \in \mathcal{R}^3$, $j = 1, 2, 3$, are the periodic right eigenvectors of $\{(A_j, E_j)\}_{j=1}^p$ corresponding to the simple eigenvalue $\lambda \equiv \pi_\alpha / \pi_\beta = 0$. Assume that by some method (e.g. the periodic Arnoldi algorithm or the periodic Krylov–Schur algorithm presented in [17]), we have come up with the orthonormal bases

$$U_j = \begin{bmatrix} 1 & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}, \quad j = 1, 2, 3.$$

Then the periodic Rayleigh–Ritz matrix pairs $\{(M_j, N_j)\}_{j=1}^p$ are given by

$$M_j = U_j^\top A_j U_{j-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad N_j = I_2, \quad j = 1, 2, 3.$$

Thus, any set of nonzero vectors, say the unnormalized $z_j = [\,1, \sqrt{2}\,]^\top$ $(j = 1, 2, 3)$, forms the periodic right eigenvectors of $\{(M_j, N_j)\}_{j=1}^p$ corresponding to the zero eigenvalue. Then we have $\tilde{x}_j \equiv U_j z_j = [\,1, 1, 1\,]^\top$ $(j = 1, 2, 3)$ as the approximate periodic eigenvectors, which are completely wrong. Therefore, the periodic Rayleigh–Ritz procedure can fail, even though the spaces $\{\tilde{\mathcal{X}}_j\}_{j=1}^p \equiv \{\text{span}(U_j)\}_{j=1}^p$ contain the desired eigenvectors. However, refinement finds the refined periodic Ritz vectors $\hat{x}_j = x_j,\ j = 1, 2, 3$ exactly.

In reality, we cannot expect M_j to be exactly zero. For example, if we perturb each U_j by a matrix of random normal variables with standard deviation of 10^{-4}, then we compute the periodic Rayleigh–Ritz matrix pairs $\{(M_j, N_j)\}_{j=1}^p$ again by

$$M_j \equiv V_j^\top A_j U_{j-1} = (U_j^\top U_j)^{-1} U_j^\top A_j U_{j-1}, \qquad N_j \equiv V_j^\top E_j U_j = I_2$$

for $j = 1, 2, 3$. Applying the Rayleigh–Ritz method, the approximate periodic Ritz vectors corresponding to the smallest eigenvalue of the periodic matrix pairs $\{(M_j, N_j)\}_{j=1}^p$ are

$$\tilde{x}_1 = \begin{bmatrix} 8.9825\text{e–}01 & 3.1098\text{e–}01 & 3.1092\text{e–}01 \end{bmatrix}^\top,$$
$$\tilde{x}_2 = \begin{bmatrix} -9.2454\text{e–}01 & 2.6936\text{e–}01 & 2.6930\text{e–}01 \end{bmatrix}^\top,$$
$$\tilde{x}_3 = \begin{bmatrix} 8.5145\text{e–}01 & 3.7075\text{e–}01 & 3.7074\text{e–}01 \end{bmatrix}^\top,$$

completely meaningless approximations to the original eigenvectors $x_j = [\,1, 0, 0\,]^\top$ $(j = 1, 2, 3)$.

On the other hand, the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^p$ are

$$\hat{x}_1 = \begin{bmatrix} 1.0001\text{e+}00 & 2.8944\text{e–}05 & -2.8946\text{e–}05 \end{bmatrix}^\top,$$
$$\hat{x}_2 = \begin{bmatrix} 9.9997\text{e–}01 & -4.8199\text{e–}05 & 4.8196\text{e–}05 \end{bmatrix}^\top,$$
$$\hat{x}_3 = \begin{bmatrix} 1.0000\text{e+}00 & 3.3310\text{e–}05 & -3.3306\text{e–}05 \end{bmatrix}^\top,$$

which are excellent approximations to the original eigenvectors $x_j = [\,1, 0, 0\,]^\top$ $(j = 1, 2, 3)$ of the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$, in view of perturbation matrices with norms of order 10^{-4}.

Example 2. In this example we look more at the convergence of the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^p$ as summarized in Theorem 5.1. For $j = 1, 2, \ldots, p$, let

$$A_j = \begin{bmatrix} 1 & -\sin\phi_j & & 0 \\ & 2^{1/p} & \ddots & \\ & & \ddots & -\sin\phi_j \\ 0 & & & n^{1/p} \end{bmatrix}, \qquad E_j = \begin{bmatrix} n^{1/p} & \cos\phi_j & & 0 \\ & (n-1)^{1/p} & \ddots & \\ & & \ddots & \cos\phi_j \\ 0 & & & 1 \end{bmatrix},$$

where $\phi_j = 2\pi j/p$ for all j. We consider the periodic matrix pairs $\{(A_j, E_j)\}_{j=1}^p$ with $p = 8$, $n = 100, 500, 1000$.

For the Ritz value pair $(\pi_\mu, 1)$ and approximate periodic eigenvectors $\{\tilde{x}_j\}$ computed by the periodic Krylov–Schur algorithm in [17] and the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^p$ computed by Newton's method, the quantities "res_1", "res_2", "$\sin_1$" and "$\sin_2$" are defined as

$$\sin_1 = \max_{1\le j\le p} \sin\angle(\tilde{x}_j, x_j), \qquad \sin_2 = \max_{1\le j\le p} \sin\angle(\hat{x}_j, x_j),$$
$$\text{res}_1 = \max_{1\le j\le p} \|A_j\tilde{x}_{j-1} - \mu_j E_j \tilde{x}_j\|^2, \qquad \text{res}_2 = \max_{1\le j\le p} \|A_j\hat{x}_{j-1} - \mu_j E_j \hat{x}_j\|^2,$$

where the periodic eigenvectors $\{x_j\}_{j=1}^p$ are computed from (4) with $(\alpha_j, \beta_j) = (j^{1/p}, (n-j+1)^{1/p})$ $(j = 1, \ldots, n)$.

The numerical results for the largest eigenvalue $\lambda \equiv \pi_\alpha/\pi_\beta = n$ are shown in Table 1. Here k is the dimension of periodic Krylov subspaces. From Table 1, the refined periodic Ritz vectors $\{\hat{x}_j\}_{j=1}^p$ converge to the periodic eigenvectors $\{x_j\}_{j=1}^p$ corresponding to $\lambda = n$ when the dimensions of Krylov subspaces increase.

For the results in Table 1, we have the following comments.

(1) The quantities $\sin_2$ (the sine of the maximum angle between the refined periodic Ritz vectors and the eigenvectors) exceeds $\sin_1$ (the sine of the maximum angle between the periodic Ritz vectors and the eigenvectors) twice when $k = 10$ and $n = 500, 1000$, when the bases for U_j from the Arnoldi process do not contain enough information.
(2) The quantity $\sin_2$ is bounded by $\|\eta\|/\text{sep}$ as stated in Theorem 5.1. From our numerical experiments, the bound is sharp and over-estimates $\sin_2$ by 100 folds for $n = 100, 500$, to a sharper 10 folds for $n = 500, 1000$.
(3) Refinement always improves the accuracy of the Ritz vectors, with $\text{res}_1 > \text{res}_2$ in Table 1. However, the improvement is not drastic for this example, but could have been as suggested by Example 1 or Section 4.
(4) Increasing k from 10 to 15 in the Arnoldi process improves the accuracy of the Ritz vectors, but increasing it further from 15 to 25 worsen the accuracy slightly. Refinement is clearly necessary if higher accuracy is required.

Table 1
Numerical results for Example 2.

n	k	$\sin_1$	$\sin_2$	$\|\eta\|/\text{sep}$	res_1	res_2
100	10	2.461e−9	2.284e−9	3.890e−7	2.283e−9	8.820e−10
100	15	8.517e−14	5.370e−14	3.736e−12	1.950e−13	2.913e−14
100	25	6.606e−13	6.250e−13	8.318e−11	5.640e−13	4.799e−14
500	10	1.738e−8	2.477e−8	2.442e−6	1.845e−8	6.661e−9
500	15	2.050e−12	1.918e−12	2.110e−10	3.176e−12	6.682e−13
500	25	3.178e−12	2.295e−12	6.796e−11	4.750e−12	5.032e−13
1000	10	2.231e−8	2.608e−8	4.099e−7	2.412e−7	8.410e−8
1000	15	1.075e−12	4.919e−13	6.252e−12	1.358e−11	1.534e−12
1000	25	2.317e−12	7.646e−13	7.264e−12	2.353e−11	2.907e−12

8. Conclusions

In this paper, we first proposed the periodic Rayleigh–Ritz method for the eigen-problem of periodic matrix pairs and showed how to compute the periodic Ritz values and the periodic Ritz vectors. Then we established convergence theory of the Ritz values and the periodic Ritz vectors, revealing the possible non-convergence of the periodic Ritz vectors. To overcome this drawback, we introduced the refined periodic Ritz vectors, which, unlike ordinary periodic Ritz vectors, are guaranteed to converge whenever the angles between desired periodic vectors and the approximate periodic subspaces approach zero. These results generalized the corresponding ones of the standard Rayleigh–Ritz approximation and its refined version in [26,27,22]. Numerical examples demonstrated that the refined periodic Ritz vectors are excellent approximations to the desired periodic eigenvectors, and confirmed the sharpness of the upper bounds in (45) for the angles between the refined periodic Ritz vectors and the periodic eigenvectors. The computation of the refined Ritz vectors is especially efficient when coupled with the periodic Arnoldi process in Section 6.

Acknowledgements

Third author's work was supported by the National Basic Research Program of China 2011CB302400 and the National Science Foundation of China (Nos. 11071140 and 10771116). First, second and fourth author's work were partially supported by the National Centre of Theoretical Sciences, of ROC in Taiwan.

References

[1] A. Bojanczyk, G.H. Golub, P. Van Dooren, The periodic Schur decomposition, algorithms and applications, in: Proceedings of the SPIE Conference, San Diego, vol. 1770, 1992, pp. 31–42.
[2] J.J. Hench, A.J. Laub, Numerical solution of the discrete-time periodic Riccati equation, IEEE Trans. Automat. Control 39 (1994) 1197–1210.
[3] W.-W. Lin, P. Van Dooren, Q.-F. Xu, Periodic invariant subspaces in control, in: Workshop on Periodic Systems and Control, Como, Italy, 2001, pp. 21–25.
[4] W.-W. Lin, J.-G. Sun, Perturbation analysis for the eigenproblem of periodic matrix pairs, Linear Algebra Appl. 337 (2001) 157–187.
[5] E.K.-W. Chu, W.-W. Lin, Perturbation of eigenvalues for periodic matrix pairs via the Bauer–Fike theorems, Linear Algebra Appl. 378 (2004) 183–202.
[6] Y.-C. Kuo, W.-W. Lin, S.-F. Xu, Regularization of linear discrete-time periodic descriptor systems by derivative and proportional state feedback, SIAM J. Matrix Anal. Appl. 25 (2004) 1046–1073.
[7] V.V. Sergeichuk, Computation of canonical matrices for chains and cycles of linear mappings, Linear Algebra Appl. 376 (2004) 235–263.
[8] S. Bittati, P. Colaneri, G. De Nicolao, The difference periodic Riccati equation for the periodic prediction problem, IEEE Trans. Automat. Control 33 (1988) 706–712.
[9] B. Franci, T. Georgiou, Stability theory for linear time-invariant plants with periodic digital controllers, IEEE Trans. Automat. Control 33 (1988) 820–832.
[10] O. Grasselli, S. Longhi, The geometric approach for linear periodic discrete-time systems, Linear Algebra Appl. 158 (1991) 27–60.
[11] J. Sreedhar, P. Van Dooren, Periodic descriptor systems: solvability and conditionability, IEEE Trans. Automat. Control 44 (1999) 310–313.
[12] A. Varga, P. Van Dooren, Computational methods for periodic systems—an overview, in: Proceedings of IFAC Workshop on Periodic Control Systems, Como, Italy, 2001, pp. 171–176.
[13] F. Bonhoure, Y. Dalley, W.J. Stewart, On the use of periodicity properties for the efficient numerical solution of certain Markov chains, Numer. Linear Algebra Appl. 1 (1994) 265–286.
[14] W.J. Stewart, Introduction to the Numerical Solution of Markov Chains, Princeton University Press, Princeton, NJ, 1994.
[15] K. Lust, Numerical bifurcation analysis of periodic solution of partial differential equations, Ph.D. Thesis, Department of Computer Science, KU Leuven, Belgium, 1997.
[16] K. Lust, Improved numerical Floquet multipliers, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 11 (2001) 2389–2410.
[17] D. Kressner, A periodic Krylov–Schur algorithm for large matrix products, Numer. Math. 103 (2006) 461–483.
[18] D.S. Watkins, Product eigenvalue problems, SIAM Rev. 47 (2005) 3–40.
[19] E.K.-W. Chu, H.-Y. Fan, W.-W. Lin, C.-S. Wang, Structure-preserving algorithms for periodic discrete-time algebraic Riccati equations, Internat. J. Control 77 (2004) 767–788.
[20] P. Benner, V. Mehrmann, H. Xu, Perturbation analysis for the eigenvalue problem of a formal product of matrices, BIT 42 (2002) 1–43.
[21] J.-G. Sun, Perturbation bounds for subspaces associated with periodic eigenproblems, Taiwanese J. Math. 9 (2005) 17–38.
[22] G.W. Stewart, Matrix Algorithms, Volume II: Eigensystems, SIAM, Philadelphia, 2001.
[23] Z. Jia, Refined iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems, Linear Algebra Appl. 259 (1997) 1–23.
[24] Z. Jia, The refined harmonic Arnoldi method and an implicitly restarted refined algorithm for computing interior eigenpairs of large matrices, Appl. Numer. Math. 42 (2002) 489–512.
[25] Z. Jia, The convergence of harmonic Ritz values, harmonic Ritz vectors and refined harmonic Ritz vectors, Math. Comp. 74 (2005) 1441–1456.
[26] Z. Jia, G.W. Stewart, On the convergence of Ritz values, Ritz vectors, and refined Ritz vectors, TR-99-08, Institute for Advanced Computer Studies and TR-3986, Department of Computer Science, University of Maryland, College Park, 1999. Available on the web at: http://www.cs.umd.edu/stewart/.
[27] Z. Jia, G.W. Stewart, An analysis of the Rayleigh–Ritz method for approximating eigenspaces, Math. Comp. 70 (2001) 637–647.