A CONTINUATION METHOD FOR A RIGHT DEFINITE TWO-PARAMETER EIGENVALUE PROBLEM

BOR PLESTENJAK†

Abstract. The continuation method has been successfully applied to the classical $Ax = \lambda x$ and to the generalized $Ax = \lambda Bx$ eigenvalue problem. Shimasaki applied the continuation method to the right definite two-parameter problem which results as a discretization of a two-parameter Sturm–Liouville problem. We show that the continuation method can be used for a general right definite two-parameter problem and we give a sketch of the algorithm. For a local convergent method we use the Tensor Rayleigh Quotient Iteration (TRQI), which is a generalization of the Rayleigh iterative method to two-parameter problems. We show its convergence and compare it with Newton's method and with the Generalized Rayleigh Quotient Iteration (GRQI), studied by Ji, Jiang and Lee.

Key words. right definite two-parameter problem, continuation method, Newton's method, Rayleigh quotient iteration

AMS subject classifications. 65F15, 65H20, 15A69, 15A18

1. Introduction. We consider a two-parameter eigenvalue problem

$$\begin{aligned} A_1x &= \lambda B_1x + \mu C_1x, \\ A_2y &= \lambda B_2y + \mu C_2y, \end{aligned} \tag{1.1}$$

where A_i, B_i, C_i are symmetric $n_i \times n_i$ matrices over $\mathbb{R}$, $i = 1, 2$. We also require a definiteness condition

$$D(x,y) := \begin{vmatrix} x^TB_1x & x^TC_1x \\ y^TB_2y & y^TC_2y \end{vmatrix} \geq \delta_1 > 0 \tag{1.2}$$

for all vectors $\|x\| = \|y\| = 1$. We call the problem (1.1) *right definite* [27] if the condition (1.2) holds. If in addition the matrices B_1 and C_2 are positive definite, then we call the problem *diagonal right definite*.

We say that (λ, μ) is an *eigenvalue* of the problem (1.1) if

$$\ker(A_i - \lambda B_i - \mu C_i) \neq \{0\}, \quad i = 1, 2.$$

If $\dim\ker(A_i - \lambda B_i - \mu C_i) = 1$ for $i = 1, 2$ then (λ, μ) is a *simple eigenvalue* of (1.1).

On the tensor product space $V := \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$ of the dimension $N := n_1n_2$ we define operator determinants

$$\Delta_0 = \begin{vmatrix} B_1^\dagger & C_1^\dagger \\ B_2^\dagger & C_2^\dagger \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} A_1^\dagger & C_1^\dagger \\ A_2^\dagger & C_2^\dagger \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} B_1^\dagger & A_1^\dagger \\ B_2^\dagger & A_2^\dagger \end{vmatrix},$$

where $A_1^\dagger, A_2^\dagger, B_1^\dagger, B_2^\dagger, C_1^\dagger, C_2^\dagger$ are the induced linear transformations on V. For instance, $A_1^\dagger$ is defined on a decomposable tensor $x \otimes y$ by $A_1^\dagger(x \otimes y) = A_1x \otimes y$ and this definition is extended to all of V by linearity and continuity. Similarly, $A_2^\dagger(x \otimes y) = x \otimes A_2y$.

†IMFM/TCS, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia (bor.plestenjak@fmf.uni-lj.si)

A two-parameter system (1.1) is called nonsingular if the corresponding operator determinant Δ_0 is invertible. In the case of a nonsingular two-parameter system, the problem (1.1) is equivalent to the simultaneous problem in V

$$\begin{aligned} \Delta_1 z &= \lambda \Delta_0 z, \\ \Delta_2 z &= \mu \Delta_0 z, \end{aligned} \tag{1.3}$$

for decomposable tensors z (see Atkinson [3] for details). Atkinson [3, Theorem 7.8.2] also showed that right definiteness is equivalent to the condition that the operator determinant Δ_0 is positive definite.

Several numerical algorithms can be applied to the right definite two-parameter eigenvalue problem (1.1). Bohte [9] uses Newton's method to find an eigenvalue pair. The simultaneous problem (1.3) of the right definite problem (1.1) can be treated by standard numerical methods [25, 16]. Müller [21] uses the continuation method to compute eigenvalue curves starting from a given eigenvalue and Browne and Sleeman [11] use the gradient method [7, 8]. Blum and Chang [6] derived the Minimum Residual Quotient Iteration (MRQI) for the problem $Ax = \lambda Bx + \mu Cx$ subject to $\|x\| = 1$ and $f(x) = 0$, where f is a real functional. Ji, Jiang and Lee [18] generalized their approach to the right definite two-parameter problem (1.1) and derived the Generalized Rayleigh Quotient Iteration (GRQI).

Multiparameter eigenvalue problems arise in a variety of applications [2], particularly in mathematical physics when the method of separation of variables is used to solve boundary value problems [27]. When the separation constants cannot be decoupled, two-parameter Sturm–Liouville problems of the form

$$-\Big(p_i(x_i)y_i'(x_i)\Big)' + q_i(x_i)y_i(x_i) = \Big(\lambda a_{i1}(x_i) + \mu a_{i2}(x_i)\Big)y_i(x_i),$$

where $x_i \in [a_i, b_i]$, with boundary conditions

$$y_i(a_i)\cos\alpha_i - y_i'(a_i)\sin\alpha_i = 0, \quad 0 \le \alpha_i \le \pi,$$

$$y_i(b_i)\cos\beta_i - y_i'(b_i)\sin\beta_i = 0, \quad 0 \le \beta_i \le \pi,$$

will arise, where $\alpha_i \in [0, \pi)$, $\beta_i \in (0, \pi]$ and $p_i', q_i, a_{i1}, a_{i2}$ are real valued and continuous, for $i = 1, 2$. Many numerical methods for two-parameter Sturm–Liouville problems have been proposed, see for example [4, 5, 14, 15].

By using finite differences the problem (1) can be converted into a right definite problem (1.1), where matrices A_i, B_i, C_i satisfy the following conditions: A_i is an irreducible tridiagonal matrix and B_i, C_i are diagonal matrices, such that diagonal elements are all strictly positive or all strictly negative, for $i = 1, 2$. We will denote these conditions by TBC as in [17]. Some numerical methods have been developed specially for problems which satisfy TBC: Ji [17] uses the two-dimensional bisection method and Shimasaki [23] uses the continuation method with quadratic convergence rate [24] (cf. [1] for details about the continuation method).

The continuation method has been successfully applied to the one-parameter eigenvalue problems, see for example [12, 13, 19, 20]. We show that the continuation method with a homotopy similar to the one used in [23] can be applied to a general two-parameter problem (1.1), which does not necessarily satisfy TBC.

In §2 we show that every right definite problem can be transformed to a diagonal right definite problem by only a simple linear substitution of parameters λ and μ.

In §3 we construct a homotopy, derive bounds for the minimum distance between eigenvectors and give a sketch of the algorithm based on the continuation method. For a local convergent method in the continuation method we use Tensor Rayleigh Quotient Iteration (TRQI). We derive TRQI in §4, show its convergence and compare it with another generalization of the Rayleigh quotient, suggested by Ji, Jiang and Lee [18]. In §5 we give a numerical example which reflects the behaviour of our continuation method.

2. Transformation to a diagonal right definite problem.

LEMMA 2.1. *If (1.1) is a right definite problem, then*

(i) *at least one of the matrices B_1 and B_2 is definite (positive or negative)*

(ii) *at least one of the matrices C_1 and C_2 is definite (positive or negative)*

Proof. First we show that at least one of the matrices B_1, B_2 is definite.

If not, there exist vectors $x_0, y_0 \neq 0$ such that $x_0^T B_1 x_0 = 0$ and $y_0^T B_2 y_0 = 0$, which gives the counterexample $D(x_0, y_0) = 0$. It follows that at least one of the matrices B_1,B_2 is definite. In the same manner we show that at least one of the matrices C_1, C_2 is definite. $\square$

Since the problem (1.1) is right definite, there exist $i, j \in \{1, 2\}$ such that B_i and C_j are definite matrices (positive or negative). This is enough to transform the right definite problem (1.1) to a diagonal right definite problem.

LEMMA 2.2. *If the problem (1.1) is right definite, then it is possible to transform it to a diagonal right definite problem by a linear substitution of parameters λ and μ.*

Proof. Let us first consider all possibilities with the positive definite matrix B_1.

(i) The matrix C_2 is positive definite. This is already the situation we are looking for, so no transformation is needed.

(ii) The matrix C_2 is negative definite. Since for all vectors $x, y \neq 0$ we have $(x^T B_1 x)(y^T C_2 y) < 0$ and $D(x, y) > 0$, it follows that both matrices C_1 and B_2 are definite. If not, there would exist a pair $x_0, y_0 \neq 0$ such that $(x_0^T C_1 x_0)(y_0^T B_2 y_0) = 0$ and $D(x_0, y_0) < 0$, but this is not possible. Since B_2 is a definite matrix, there exists $k \in \mathbb{R}$, such that $C_2 + kB_2$ is a positive definite matrix.
It is easy to see that the substitution $\lambda \mapsto \lambda + k\mu$ gives

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu(C_1 + kB_1)x \\ A_2 y &= \lambda B_2 y + \mu(C_2 + kB_2)y, \end{aligned}$$

which is a desired situation.

(iii) The matrix C_1 is definite (positive or negative). From $D(x, y) > 0$ and $(x^T B_1 x) > 0$ for all vectors $x, y \neq 0$ it follows

$$(y^T C_2 y) - \frac{(x^T C_1 x)}{(x^T B_1 x)}(y^T B_2 y) > 0.$$

As a consequence for an arbitrarily chosen vector $x_0 \neq 0$ the matrix

$$C_2 - \frac{(x_0^T C_1 x_0)}{(x_0^T B_1 x_0)} B_2$$

is positive definite and the substitution $\lambda \mapsto \lambda - k\mu$ gives

$$A_1 x = \lambda B_1 x + \mu(C_1 - kB_1)x$$

$$A_2y = \lambda B_2 y + \mu(C_2 - kB_2)y,$$

where

$$k = \frac{(x_0^T C_1 x_0)}{(x_0^T B_1 x_0)}.$$

If B_1 is a negative definite matrix, we multiply both equations in (1.1) by -1. This turns B_1 into a positive definite matrix while the determinant $D(x,y)$ remains the same. Once B_1 is a positive definite matrix, we can use (i), (ii) or (iii). Similarly, if B_2 is definite, we exchange equations in (1.1) and multiply one of the equations by -1 so that $D(x,y)$ remains unchanged. This shows that it is enough to prove the lemma only for the case when the matrix B_1 is definite. □

3. A continuation method. As shown in the previous section, we can assume that the matrices B_1 and C_2 in (1.1) are positive definite. In this case we construct the following homotopy:

$$H : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} \times \mathbb{R} \times [0,1] \longrightarrow \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} \times \mathbb{R},$$

$$H(x,y,\lambda,\mu,t) := \begin{bmatrix} (1-t)W_1x + tA_1x - \lambda B_1 x - t\mu C_1 x \\ (1-t)W_2y + tA_2y - t\lambda B_2 y - \mu C_2 y \\ \frac{1}{2}(1 - x^Tx) \\ \frac{1}{2}(1 - y^Ty) \end{bmatrix}, \tag{3.1}$$

where W_i are symmetric $n_i \times n_i$ matrices, such that eigenproblems

$$W_1 x = \lambda B_1 x \tag{3.2}$$

and

$$W_2 y = \mu C_2 y \tag{3.3}$$

have n_1 and n_2 distinct eigenvalues, respectively.

It is obvious that a solution of $H(x,y,\lambda,\mu,t) = 0$ is a solution of the two-parameter problem

$$\begin{aligned} (1-t)W_1x + tA_1x &= \lambda B_1x + \mu t C_1 x, \\ (1-t)W_2y + tA_2y &= \lambda t B_2 y + \mu C_2 y, \end{aligned} \tag{3.4}$$

which is equal to (1.1) for $t = 1$ and equal to (3.2), (3.3) for $t = 0$. Associated with the problem (3.4) are operator determinants

$$\Delta_0(t) = \begin{vmatrix} B_1^\dagger & tC_1^\dagger \\ tB_2^\dagger & C_2^\dagger \end{vmatrix},$$

$$\Delta_1(t) = \begin{vmatrix} (1-t)W_1^\dagger + tA_1^\dagger & tC_1^\dagger \\ (1-t)W_2^\dagger + tA_2^\dagger & C_2^\dagger \end{vmatrix},$$

$$\Delta_2(t) = \begin{vmatrix} B_1^\dagger & (1-t)W_1^\dagger + tA_1^\dagger \\ tB_2^\dagger & (1-t)W_2^\dagger + tA_2^\dagger \end{vmatrix},$$

where $\dagger$ denotes the induced linear transformation on V.

For a start we will show that the problem (3.4) is right definite for all $t \in [0,1]$. We denote the determinant

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

by $D(x,y,t)$. It is easy to see that for arbitrary vectors x, y and $t \in [0,1]$ we have

$$\min\{D(x,y,0), D(x,y,1)\} \le D(x,y,t). \tag{3.5}$$

LEMMA 3.1. *The problem (3.4) is right definite for all $t \in [0,1]$.*

Proof. As B_1 and C_2 are positive definite matrices and the problem (1.1) is right definite, it follows

$$D(x,y,0), D(x,y,1) > 0$$

for all vectors $x, y \neq 0$. Now (3.5) yields

$$D(x,y,t) > 0$$

for all $x, y \neq 0$. $\square$

A simple consequence of Lemma 3.1 is that there exists a constant $\delta_2 > 0$ such that

$$D(x,y,t) \ge \delta_2 \tag{3.6}$$

for all vectors $\|x\| = \|y\| = 1$ and $t \in [0,1]$. It is easy to see from (3.5) that

$$\delta_2 := \min\left\{ \min_{\|x\|=\|y\|=1} D(x,y,0), \min_{\|x\|=\|y\|=1} D(x,y,1) \right\} \tag{3.7}$$

satisfies (3.6).

Let

$$\Gamma := \{(x,y,\lambda,\mu,t) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} \times \mathbb{R} \times [0,1] \mid H(x,y,\lambda,\mu,t) = 0\}$$

denote a solution set for the homotopy (3.1).

If we denote

$$\begin{aligned} S_1(t) &= (1-t)W_1 + tA_1 - \lambda B_1 - t\mu C_1, \\ S_2(t) &= (1-t)W_2 + tA_2 - t\lambda B_2 - \mu C_2, \end{aligned}$$

then we can write the Jacobian of the homotopy (3.1) as

$$J(x,y,\lambda,\mu,t) = \begin{bmatrix} S_1(t) & 0 & -B_1 x & -tC_1 x \\ 0 & S_2(t) & -tB_2 y & -C_2 y \\ -x^T & 0 & 0 & 0 \\ 0 & -y^T & 0 & 0 \end{bmatrix}. \tag{3.8}$$

DEFINITION 3.2. *We say that $(x,y,\lambda,\mu,t) \in \Gamma$ is a singular point of the homotopy (3.1) if the Jacobian $J(x,y,\lambda,\mu,t)$ is singular.*

LEMMA 3.3. *Singular points of the homotopy (3.1) are precisely those points* $(x, y, \lambda, \mu, t) \in \Gamma$ *where* (λ, μ) *is a multiple eigenvalue of (3.4).*

Proof. First we prove that if $(x, y, \lambda, \mu, t) \in \Gamma$ and (λ, μ) is a simple eigenvalue of (3.4), then the Jacobian (3.8) is nonsingular. Let $(z, w, a, b) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} \times \mathbb{R}$ be such that

$$J(x, y, \lambda, \mu, t) \begin{bmatrix} z \\ w \\ a \\ b \end{bmatrix} = 0.$$

It follows that

$$S_1(t)z - aB_1x - btC_1x = 0, \tag{3.9}$$
$$S_2(t)w - atB_2y - bC_2y = 0, \tag{3.10}$$
$$x^T z = 0, \tag{3.11}$$
$$y^T w = 0. \tag{3.12}$$

Since $H(x, y, \lambda, \mu, t) = 0$,

$$S_1(t)x = 0, \tag{3.13}$$
$$x^T x = 1, \tag{3.14}$$
$$S_2(t)y = 0, \tag{3.15}$$
$$y^T y = 1. \tag{3.16}$$

From (3.13), (3.9) and $S_1(t) = S_1^T(t)$ it follows that

$$0 = z^T S_1(t)x = x^T S_1^T(t)z = x^T S_1(t)z = ax^T B_1 x + btx^T C_1 x. \tag{3.17}$$

Similarly it follows from (3.15) and (3.10) that

$$0 = w^T S_2(t)y = y^T S_2^T(t)w = y^T S_2(t)w = aty^T B_2 y + by^T C_2 y. \tag{3.18}$$

Equations (3.17) and (3.18) form a 2×2 homogeneous linear system for a and b. Lemma 3.1 implies that this linear system is nonsingular, thus the only solution is

$$a = b = 0.$$

From (3.9) and (3.10) it now follows $S_1(t)z = 0$ and $S_2(t)w = 0$. Since (λ, μ) is a simple eigenvalue for (3.4) we have $\mathrm{rank}(S_i(t)) = n_i - 1$ and there exist $\alpha, \beta \in \mathbb{R}$ such that

$$z = \alpha x$$

and

$$w = \beta y.$$

It follows from (3.11), (3.12), (3.14) and (3.16) that $\alpha = \beta = 0$.

For the second part, let (λ, μ) be a multiple eigenvalue of (3.4). Without any loss of the generality we can assume that $\mathrm{rank}(S_1(t)) < n_1 - 1$. There exists a vector $z \in \mathbb{R}^{n_1}$ such that $z \neq 0$, $S_1(t)z = 0$ and $x^T z = 0$. It follows that

$$J(x, y, \lambda, \mu, t) \begin{bmatrix} z \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0$$

and the Jacobian $J(x, y, \lambda, \mu, t)$ is singular. □

If the problem (1.1) satisfies the TBC condition, then it is easy to see that the homotopy (3.1) has no singular points. The reason for this is the well known fact that irreducible tridiagonal matrices can not have multiple eigenvalues [22]. Shimasaki [23] makes use of this in his algorithm.

Since the problem (3.4) is right definite it has N real eigenvalues for every $t \in [0,1]$. Therefore there exist N homotopy curves

$$\vartheta_i(t) := (x_i(t), y_i(t), \lambda_i(t), \mu_i(t), t), \quad i = 1, \ldots, N,$$

parametrizable with respect to t, such that $\vartheta_i(t) \in \Gamma$. The following theorem, which is a result of Browne and Sleeman, is of major importance as it shows that $\vartheta_i(t)$ is a holomorphic function of t.

THEOREM 3.4 (Browne and Sleeman, [10]). *Eigenvalues $\lambda_i(t), \mu_i(t)$ and vectors $x_i(t), y_i(t)$, for $i = 1, \ldots, N$, are holomorphic functions of t on the interval $[0,1]$ and $x_i(t) \otimes y_i(t)$ form a $\Delta_0(t)$-orthogonal basis of eigenvectors for (3.4), i.e. if $i \neq j$ then*

$$(x_i(t) \otimes y_i(t))^T \Delta_0(t)(x_j(t) \otimes y_j(t)) = 0.$$

Theorem 3.4 implies that we have N disjoint homotopy curves $\vartheta_i(t)$. This allows easy tracking of each curve from $t = 0$ to $t = 1$.

Eigenvalues of the problem (3.4) are not necessarily distinct (except for $t = 0$), so the eigenvalue curves $\gamma_i(t) := (\lambda_i(t), \mu_i(t))$, $i = 1, \ldots, N$, can intersect. It follows from Lemma 3.3 that whenever at least two eigenvalue curves $\gamma_i(t)$ intersect, we have a singular point. When we trace an eigenvalue curve from $t = 0$ to $t = 1$ numerically, we must be aware that there is a danger of switching to another curve whenever we hit a singular point. A numerical method finds one eigenvector from the eigensubspace and since the eigensubspace at a singular point is at least two-dimensional, the method can accidentally pick an eigenvector from another homotopy curve.

Since we choose the matrices W_1 and W_2 in such a way that the homotopy has no singular points at $t = 0$, the homotopy has a singular point only at finitely many values $t \in (0,1]$. The continuity of the eigenvalue curves assures that we can jump over the singular point at t using $\vartheta_i(t-h)$ as an initial approximation for $\vartheta_i(t+h)$ if h is small enough.

In order to prevent switching from one homotopy curve to another, we will derive a bound for the constant $\eta > 0$, independent of t, such that if $x_i(t) \otimes y_i(t)$ and $x_j(t) \otimes y_j(t)$ are eigenvectors for (3.4) and unit vectors $x_i(t), y_i(t)$ and $x_j(t), y_j(t)$ are parts of homotopy curves $\vartheta_i(t)$ and $\vartheta_j(t)$, respectively, then

$$\max(\|x_i(t) - x_j(t)\|^2, \|y_i(t) - y_j(t)\|^2) \geq \eta$$

for all $i \neq j$ and $t \in [0,1]$. Then we will write the sketch of an algorithm for calculating all eigentuples of the problem (1.1) using the continuation method.

Before we can state the theorem about the existence of such an η, we have to prove some auxiliary results.

LEMMA 3.5. *Let $A \in \mathbb{R}^{n \times n}$ be a positive definite symmetric matrix and let x and y be unit vectors, such that $x^T A y = 0$. It follows that*

$$\|x - y\|^2 \geq \frac{4\lambda_n}{\lambda_1 + \lambda_n},$$

where λ_1 and λ_n are the greatest and the smallest eigenvalue of A, respectively.

Proof. Since matrix A is symmetric, there exists an orthonormal basis of eigenvectors x_i, such that $Ax_i = \lambda_i x_i$, $i = 1, \ldots, n$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ and we can write vectors x and y as:

$$x = \sum_{i=1}^{n} \alpha_i x_i,$$
$$y = \sum_{i=1}^{n} \beta_i x_i.$$

The condition $x^T A y = 0$ is now equivalent to

$$\sum_{i=1}^{n} \lambda_i \alpha_i \beta_i = 0. \tag{3.19}$$

We are searching for the minimum of $\|x-y\|^2$. Under conditions $\sum_{i=1}^{n} \alpha_i^2 = 1$ and $\sum_{i=1}^{n} \beta_i^2 = 1$ we have

$$\|x-y\|^2 = \sum_{i=1}^{n} (\alpha_i - \beta_i)^2 = 2 - 2\sum_{i=1}^{n} \alpha_i \beta_i. \tag{3.20}$$

It follows from (3.19) and from the fact that matrix A is positive definite that in (3.20) not all terms in the sum on the right-hand side can be of equal sign. It is easy to see that candidates for which the minimum is attained have the form

$$x = \alpha_1 x_1 + \alpha_n x_n$$
$$y = -\alpha_1 x_1 + \alpha_n y_n.$$

The equation (3.19) together with $\|x\| = \|y\| = 1$ gives

$$\alpha_1^2 = \frac{\lambda_n}{\lambda_1 + \lambda_n}, \quad \alpha_n^2 = \frac{\lambda_1}{\lambda_1 + \lambda_n}$$

and

$$\|x-y\|^2 = \frac{4\lambda_n}{\lambda_1 + \lambda_n}. \qquad \square$$

LEMMA 3.6. *If x_1, x_2, y_1, y_2 are unit vectors such that*

$$\|(x_1 \otimes y_1) - (x_2 \otimes y_2)\|^2 \geq \epsilon^2,$$

then

$$\max(\|x_1 - x_2\|^2, \|y_1 - y_2\|^2) \geq 2(1 - \sqrt{1 - \frac{\epsilon^2}{2}}).$$

Proof. Since vectors x_1, x_2, y_1, y_2 are all normalized, we have

$$\|(x_1 \otimes y_1) - (x_2 \otimes y_2)\|^2 = 2 - 2\langle x_1, x_2 \rangle \langle y_1, y_2 \rangle \geq \epsilon^2.$$

It follows that

$$\langle x_1, x_2\rangle\langle y_1, y_2\rangle \le 1 - \frac{\epsilon^2}{2}$$

and

$$\min(\langle x_1, x_2\rangle, \langle y_1, y_2\rangle) \le \sqrt{1 - \frac{\epsilon^2}{2}}.$$

Finally, from

$$\|x_1 - x_2\|^2 = 2 - 2\langle x_1, x_2\rangle,$$

$$\|y_1 - y_2\|^2 = 2 - 2\langle y_1, y_2\rangle,$$

we get

$$\max(\|x_1 - x_2\|^2, \|y_1 - y_2\|^2) \ge 2(1 - \sqrt{1 - \frac{\epsilon^2}{2}}). \qquad \square$$

Now we are able to prove a theorem on the minimum distance between eigenvectors of the homotopy curves $\vartheta_i(t)$.

THEOREM 3.7. *Let $x_i(t) \otimes y_i(t)$ and $x_j(t) \otimes y_j(t)$ be eigenvectors of (3.4), $i \ne j$ and $t \in [0,1]$, where unit vectors $x_i(t), y_i(t)$ and $x_j(t), y_j(t)$ are parts of homotopy curves $\vartheta_i(t)$ and $\vartheta_j(t)$, respectively. Then there exists a constant $\eta > 0$, independent of t, such that*

$$\max(\|x_i(t) - x_j(t)\|^2, \|y_i(t) - y_j(t)\|^2) \ge \eta$$

for all $i \ne j$ and $t \in [0,1]$.

Proof. By Lemma 3.1, the matrix $\Delta_0(t)$ is symmetric and positive definite for all $t \in [0,1]$. From Theorem 3.4 we have

$$(x_i(t) \otimes y_i(t))^T \Delta_0(t)(x_j(t) \otimes y_j(t)) = 0$$

and we can apply Lemma 3.5. It follows that for every $t \in [0,1]$ there exists a constant $\epsilon(t) > 0$, such that

$$\|x_i(t) \otimes y_i(t) - x_j(t) \otimes y_j(t)\|^2 \ge \epsilon^2(t).$$

By Lemma 3.6 we have

$$\max(\|x_i(t) - x_j(t)\|^2, \|y_i(t) - y_j(t)\|^2) \ge 2(1 - \sqrt{1 - \frac{\epsilon^2}{2}}),$$

where

$$\epsilon := \max_{t\in[0,1]} \epsilon(t)$$

and the constant we are looking for is

$$\eta := 2(1 - \sqrt{1 - \frac{\epsilon^2}{2}}). \qquad \square$$

In order to use Theorem 3.7 to obtain a bound η on the minimum distance between the eigenvectors for different eigenvalues of (3.4) we need information about the extreme eigenvalues of $\Delta_0(t)$. Since $\Delta_0(t)$ is a matrix of order N, the exact computation would demand too much computational work. Therefore it is better to obtain an upper and a lower bound for the maximum and the minimum eigenvalue of $\Delta_0(t)$, respectively.

The upper bound is easily obtained from the fact that

$$\|\Delta_0(t)\| = \|B_1 \otimes C_2 - t^2 C_1 \otimes B_2\| \leq \|B_1\|\|C_2\| + \|C_1\|\|B_2\|.$$

It follows that $\lambda \leq \|B_1\|\|C_2\| + \|C_1\|\|B_2\|$ for an arbitrary eigenvalue λ of $\Delta_0(t)$, $t \in [0,1]$.

It is not so trivial to obtain a lower bound for the minimum eigenvalue of $\Delta_0(t)$. Let

$$\delta(t) := \min_{\|x\|=\|y\|=1} D(x,y,t).$$

Volkmer [26, Theorem 4.5] showed that

$$\delta(t) \cdot \min(\{n_1, n_2\})^{-1} \leq \lambda$$

for an arbitrary eigenvalue λ of $\Delta_0(t)$.

If follows from (3.6) that $\delta_2 \leq \delta(t)$ for every $t \in [0,1]$. We summarize these results in the following lemma.

LEMMA 3.8. *We consider the right definite problem (3.4) where the assumption (3.6) is valid. Let λ be an eigenvalue of $\Delta_0(t)$, $t \in [0,1]$. Then*

$$\delta_2 \cdot \min(\{n_1, n_2\})^{-1} \leq \lambda \leq \|B_1\|\|C_2\| + \|C_1\|\|B_2\|. \tag{3.21}$$

Since $x_i(t)$ and $y_i(t)$ are holomorphic functions of t, for each $\eta > 0$ there exists $h > 0$, such that

$$\|x_i(t) - x_i(t')\|, \|y_i(t) - y_i(t')\| \leq \eta$$

when $|t - t'| \leq h$ and $t, t' \in [0,1]$. If we take h small enough, then $\vartheta_i(t)$ will be a good approximation for $\vartheta_i(t+h)$. If we use a numerical method which converges from the initial approximation $\vartheta_i(t)$ to the nearest solution of (3.4) at $t+h$, then switching from one curve to another is possible only when passing through a singular point.

We propose the following continuation method for tracing a homotopy curve $\vartheta(t) = \vartheta_i(t)$ from $t = 0$ to $t = 1$.

ALGORITHM 1.

Initialization

(i) Solve (3.2) and (3.3) to obtain an initial eigenvalue $(\lambda(0), \mu(0))$ and initial eigenvectors $x(0)$, $y(0)$.

(ii) Set $t = 0$ and set the initial step size h.

Main loop

Repeat:

1) If $t + h > 1$ then set $h = 1 - t$.

2) Use a numerical method to calculate $\vartheta(t+h)$ using $\vartheta(t)$ as an initial approximation.

3) Calculate $M := \max(\|(x(t+h) - x(t)\|^2, \|y(t+h) - y(t)\|^2)$
4) If
   a) $M \geq c_1\eta$, where $c_1 \in (0,1]$ is a prescribed constant and η is the constant from Theorem 3.7, or
   b) $t + h = 1$ and the Jacobian $J(x(1), y(1), \lambda(1), \mu(1), 1)$ is singular

   then set t to the last good value, set $h = 2h/3$ and go to step 2).
5) If $M \leq c_2\eta$, where $0 < c_2 < c_1$ is a prescribed constant then set $h = 3h/2$.
6) Set $t = t + h$.

Until $t \geq 1 - \epsilon_1$, where $\epsilon_1 > 0$ is a prescribed constant.

The criterion in step 4a) detects if the point $\vartheta(t+h)$ still lies on the same homotopy curve. If $M \geq c_1\eta$, then it is possible that we have switched to another curve. The reason is that either the step size h is too large either we have passed a singular point. The solution is to decrease the step size h and start again from the last good point t. Instead of halving the step size we choose to multiply it with $2/3$ in order to avoid another calculation at the point where the problem was detected in step 4). By using this approach we jump over all singular points.

When checking for singular points we could use Lemma 3.3 and check the singularity of the Jacobian $J(x(t), y(t), \lambda(t), \mu(t), t)$. These checks are necessary only when $t = 1$ as we can not jump over $t = 1$. The jump is not possible since the problem (3.4) is not necessarily positive definite for $t = 1 + h$. The solution at $t = 1$ is to always check the singularity of the Jacobian in addition to the criterion $M \geq c_1\eta$. We do this in step 4b). In case of a singular point at $t = 1$ a recipe is to successfully decrease the stepsize h and take $\vartheta(t)$ for the solution at $t = 1$ when t is close enough to 1.

In step 1) we adjust the step size if $t + h$ is greater than 1 and in step 5) we increase the step size when the homotopy curve is flat. The constants c_1 and c_2 have to be suitably chosen to balance the computational time and the precision of the final results.

4. Numerical methods. The main principle of the continuation method is that as the parameter t changes from 0 to 1, $\vartheta(t)$ is always a good initial approximation for $\vartheta(t+h)$. In order to follow a homotopy curve from $t = 0$ to $t = 1$ numerically, we need a local convergent method, which will converge to the eigentuple of a right definite two-parameter eigenvalue problem (1.1) providing a good initial approximation is given. For such a method we can take Newton's method.

We view $u = (x, y, \lambda, \mu)$ as a $(K+2)$-dimensional vector where $K := n_1 + n_2$ and define

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

Newton's method applied to the equation $F(u) = 0$ gives

$$F'(u_k)(u_{k+1} - u_k) = -F(u_k), \tag{4.1}$$

where

$$F'(u) = \begin{bmatrix} A_1 - \lambda B_1 - \mu C_1 & 0 & -B_1x & -C_1x \\ 0 & A_2 - \lambda B_2 - \mu C_2 & -B_2y & -C_2y \\ x^T & 0 & 0 & 0 \\ 0 & y^T & 0 & 0 \end{bmatrix}.$$

To perform one step (4.1) of Newton's method, we have to solve the linear system

$$\begin{bmatrix} A_1-\lambda_kB_1-\mu_kC_1 & 0 & -B_1x_k & -C_1x_k \\ 0 & A_2-\lambda_kB_2-\mu_kC_2 & -B_2y_k & -C_2y_k \\ x_k^T & 0 & 0 & 0 \\ 0 & y_k^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta y_k \\ \Delta\lambda_k \\ \Delta\mu_k \end{bmatrix} = \begin{bmatrix} -(A_1-\lambda_kB_1-\mu_kC_1)x_k \\ -(A_2-\lambda_kB_2-\mu_kC_2)y_k \\ \frac{1}{2}(1-x_k^Tx_k) \\ \frac{1}{2}(1-y_k^Ty_k) \end{bmatrix} \tag{4.2}$$

and to calculate new approximations

$$\begin{aligned} x_{k+1} &= x_k+\Delta x_k, \\ y_{k+1} &= y_k+\Delta y_k, \\ \lambda_{k+1} &= \lambda_k+\Delta\lambda_k, \\ \mu_{k+1} &= \mu_k+\Delta\mu_k. \end{aligned}$$

The following equations can be derived from (4.2)

$$(A_1-\lambda_kB_1-\mu_kC_1)x_{k+1} = \Delta\lambda_kB_1x_k+\Delta\mu_kC_1x_k, \tag{4.3}$$

$$(A_2-\lambda_kB_2-\mu_kC_2)y_{k+1} = \Delta\lambda_kB_2y_k+\Delta\mu_kC_2y_k, \tag{4.4}$$

$$x_k^Tx_{k+1} = \tfrac{1}{2}(x_k^Tx_k+1),$$

$$y_k^Ty_{k+1} = \tfrac{1}{2}(y_k^Ty_k+1).$$

To solve for $\Delta\lambda_k, \Delta\mu_k$ we denote

$$\begin{aligned} v_k &= (A_1-\lambda_kB_1-\mu_kC_1)^{-1}B_1x_k, \\ w_k &= (A_1-\lambda_kB_1-\mu_kC_1)^{-1}C_1x_k, \\ p_k &= (A_2-\lambda_kB_2-\mu_kC_2)^{-1}B_2y_k, \\ q_k &= (A_2-\lambda_kB_2-\mu_kC_2)^{-1}C_2y_k. \end{aligned} \tag{4.5}$$

We multiply the above equalities with x_k^T and y_k^T and obtain the following linear system

$$\begin{bmatrix} x_k^Tv_k & x_k^Tw_k \\ y_k^Tp_k & y_k^Tq_k \end{bmatrix} \begin{bmatrix} \Delta\lambda_k \\ \Delta\mu_k \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(x_k^Tx_k+1) \\ \frac{1}{2}(y_k^Ty_k+1) \end{bmatrix} \tag{4.6}$$

for $\Delta\lambda_k$ and $\Delta\mu_k$. After we solve the system (4.6) for $\Delta\lambda_k, \Delta\mu_k$ we calculate x_{k+1} and y_{k+1} from (4.3) and (4.4).

To improve the convergence we use the following generalization of Rayleigh quotient to two-parameter problem (1.1). Remember that in the classical eigenvalue problem $Ax=\lambda x$ the formula for the Rayleigh quotient is

$$\rho(x) = \frac{x^TAx}{x^Tx}.$$

If instead of the problem (1.1) we consider the related system (1.3) than we define the Rayleigh quotient $\rho(x_k, y_k, A_1, A_2, B_1, B_2, C_1, C_2)$ as a pair (ρ_1, ρ_2), where

$$\begin{aligned} \rho_1 &= \frac{\langle \Delta_1 z, z\rangle}{\langle \Delta_0 z, z\rangle}, \\ \rho_2 &= \frac{\langle \Delta_2 z, z\rangle}{\langle \Delta_0 z, z\rangle}. \end{aligned} \tag{4.7}$$

We call (4.7) a *tensor Rayleigh quotient*. Since the two-parameter problem (1.1) is right definite, all tensor Rayleigh quotients are well defined.

Although it looks at first sight that a computation with $\Delta_0, \Delta_1, \Delta_2$, which are all of order N, will increase the computational work and make this method inefficient, it is possible to compute the tensor Rayleigh quotient quite efficiently. The reason is that we are always dealing only with the decomposable tensors $z = x \otimes y$. Then it is easy to see from the structure of Δ_i that

$$\begin{aligned} \rho_1 &= \frac{(x^T C_1 x)(y^T A_2 y) - (x^T A_1 x)(y^T C_2 y)}{(x^T B_1 x)(y^T C_2 y) - (x^T C_1 x)(y^T B_2 y)}, \\ \rho_2 &= \frac{(x^T A_1 x)(y^T B_2 y) - (x^T B_1 x)(y^T A_2 y)}{(x^T B_1 x)(y^T C_2 y) - (x^T C_1 x)(y^T B_2 y)}. \end{aligned} \tag{4.8}$$

This shows that we do not have to work explicitly with matrices Δ_i.

In a modification of Newton's method with the tensor Rayleigh quotient we respectively replace λ_k and μ_k with $\rho_{k,1}$ and $\rho_{k,2}$ in equations (4.3),(4.4) and (4.5). Let x_0 and y_0 be two initial approximations for the eigenvector of (1.1). An initial approximation for the eigenvalue is not needed, since we can calculate one using the tensor Rayleigh quotient. The following algorithm is called Tensor Rayleigh Quotient Iteration (TRQI):

Repeat for $k = 0, 1, \ldots$:

1) Calculate the tensor Rayleigh quotient
$$(\rho_{k,1}, \rho_{k,2}) = \rho(x_k, y_k, A_1, A_2, B_1, B_2, C_1, C_2).$$
2) If any of the matrices $A_1 - \rho_{k,1}B_1 - \rho_{k,2}C_1$ or $A_2 - \rho_{k,1}B_2 - \rho_{k,2}C_2$ is singular, give $(\rho_{k,1}, \rho_{k,2})$ a small perturbation.
3) Calculate the vectors
$$\begin{aligned} v_k &= (A_1 - \rho_{k,1}B_1 - \rho_{k,2}C_1)^{-1} B_1 x_k, \\ w_k &= (A_1 - \rho_{k,1}B_1 - \rho_{k,2}C_1)^{-1} C_1 x_k, \\ p_k &= (A_2 - \rho_{k,1}B_2 - \rho_{k,2}C_2)^{-1} B_2 y_k, \\ q_k &= (A_2 - \rho_{k,1}B_2 - \rho_{k,2}C_2)^{-1} C_2 y_k. \end{aligned}$$
4) Calculate $\Delta\lambda_k$ and $\Delta\mu_k$ from the linear system
$$\begin{bmatrix} x_k^T v_k & x_k^T w_k \\ y_k^T p_k & y_k^T q_k \end{bmatrix} \begin{bmatrix} \Delta\lambda_k \\ \Delta\mu_k \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(x_k^T x_k + 1) \\ \frac{1}{2}(y_k^T y_k + 1) \end{bmatrix}.$$
5) Calculate the vectors
$$\begin{aligned} \widetilde{x}_{k+1} &= \Delta\lambda_k v_k + \Delta\mu_k w_k, \\ \widetilde{y}_{k+1} &= \Delta\lambda_k p_k + \Delta\mu_k q_k, \\ x_{k+1} &= \widetilde{x}_{k+1}/\|\widetilde{x}_{k+1}\|, \\ y_{k+1} &= \widetilde{y}_{k+1}/\|\widetilde{y}_{k+1}\|. \end{aligned}$$

6) Set $k = k + 1$.

Until $\left(\|(A_1 - \rho_{k,1}B_1 - \rho_{k,2}C_1)x_{k+1}\|^2 + \|(A_2 - \rho_{k,1}B_2 - \rho_{k,2}C_2)y_{k+1}\|^2\right)^{\frac{1}{2}} \le \epsilon_2$.

To justify the terminating criterion we have to prove some results. Let us define the residual

$$r(x, y, \lambda, \mu) := \begin{bmatrix} (A_1 - \lambda B_1 - \mu C_1)x \\ (A_2 - \lambda B_2 - \mu C_2)y \end{bmatrix}. \tag{4.9}$$

LEMMA 4.1 (Ji, Jiang and Lee, [18]). *Let λ, μ be arbitrary scalars and x, y arbitrary unit vectors. Let $(x^*, y^*, \lambda^*, \mu^*)$ be an eigentuple of (1.1). If the matrix*

$$\begin{bmatrix} x^T B_1 x^* & x^T C_1 x^* \\ y^T B_2 y^* & y^T C_2 y^* \end{bmatrix}$$

is nonsingular, then

$$\sqrt{(\lambda^* - \lambda)^2 + (\mu^* - \mu)^2} \le \left\| \begin{bmatrix} x^T B_1 x^* & x^T C_1 x^* \\ y^T B_2 y^* & y^T C_2 y^* \end{bmatrix}^{-1} \right\| \|r\|_2, \tag{4.10}$$

where $r = r(x, y, \lambda, \mu)$ is the residual (4.9).

Proof. Since $(x^*, y^*, \lambda^*, \mu^*)$ is an eigentuple, we have

$$\begin{aligned} A_1 x^* - \lambda^* B_1 x^* - \mu^* C_1 x^* &= 0, \\ A_2 y^* - \lambda^* B_2 y^* - \mu^* C_2 y^* &= 0. \end{aligned} \tag{4.11}$$

From (4.9) it follows

$$\begin{aligned} x^T A_1 x^* - x^T \lambda B_1 x^* - x^T \mu C_1 x^* &= x^{*T} r_1, \\ y^T A_2 y^* - y^T \lambda B_2 y^* - y^T \mu C_2 y^* &= y^{*T} r_2, \end{aligned} \tag{4.12}$$

where $r = [r_1\ r_2]^T$. If we multiply relations (4.11) by x^T and y^T and subtract them from (4.12), we obtain the following linear system

$$\begin{bmatrix} x^T B_1 x^* & x^T C_1 x^* \\ y^T B_2 y^* & y^T C_2 y^* \end{bmatrix} \begin{bmatrix} \lambda^* - \lambda \\ \mu^* - \mu \end{bmatrix} = \begin{bmatrix} x^{*T} r_1 \\ y^{*T} r_2 \end{bmatrix}.$$

The bound (4.10) follows readily. □

COROLLARY 4.2 (Ji, Jiang and Lee, [18]). *If the sequence $(x_k, y_k, \rho_{k,1}, \rho_{k,2})$ converges to the eigentuple $(x^*, y^*, \lambda^*, \mu^*)$ of (1.1), then for a sufficiently large k we have*

$$\sqrt{(\lambda^* - \rho_{k,1})^2 + (\mu^* - \rho_{k,2})^2} \le c\, \delta_1^{-1} \|r_k\|_2,$$

where $r_k = r(x_k, y_k, \rho_{k,1}, \rho_{k,2})$ is the residual (4.9), c is a constant independent of r, k and δ_1 is the constant from (1.2).

Proof. For a sufficiently large k we have $x_k = x^* + o(1)$ and $y_k = y^* + o(1)$. It follows that

$$\begin{vmatrix} x_k^T B_1 x^* & x_k^T C_1 x^* \\ y_k^T B_2 y^* & y_k^T C_2 y^* \end{vmatrix} = D(x^*, y^*) + o(1).$$

Since $D(x^*, y^*) \geq \delta_1$ we can apply Lemma 4.1. It is easy to see that there exists a constant c such that

$$\left\| \begin{bmatrix} x_k^T B_1 x^* & x_k^T C_1 x^* \\ y_k^T B_2 y^* & y_k^T C_2 y^* \end{bmatrix}^{-1} \right\| \leq \delta_1^{-1} c. \qquad \square$$

We will show that when TRQI converges, its convergence rate is quadratic. We denote by $u_k = (x_k, y_k, \rho_{k,1}, \rho_{k,2})$ the sequence generated by TRQI with the initial vectors x_0 and y_0. Here $\|x_k\| = \|y_k\| = 1$, $k = 0, 1, \ldots$. Let $u^* = (x^*, y^*, \rho_1^*, \rho_2^*)$ stands for a true solution of the problem, ρ_k for $(\rho_{k,1}, \rho_{k,2})$ and ρ^* for (ρ_1^*, ρ_2^*).

We need the following lemma on the error bound for the Rayleigh quotient.

LEMMA 4.3. *Let A be a symmetric matrix with eigenpairs $Ax_i = \lambda_i x_i$, for $i = 1, \ldots, n$. Let a unit vector x be an approximation for the eigenvector x_1 and let $\mu = x^T A x$ be the Rayleigh quotient for x. Then*

$$|\mu - \lambda_1| \leq 2\|A\|\|x - x_1\|^2. \tag{4.13}$$

Proof. Since matrix A is symmetric, its eigenvectors form an orthogonal basis. In this basis we can write $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$, where $\alpha_1^2 + \cdots + \alpha_n^2 = 1$. It is easy to see that

$$\mu = x^T A x = \lambda_1 \alpha_1^2 + \lambda_2 \alpha_2^2 \cdots + \lambda_n \alpha_n^2$$

and

$$\|x - x_1\|^2 = (1 - \alpha_1)^2 + \alpha_2^2 + \cdots + \alpha_n^2 = 2(1 - \alpha_1).$$

It follows that

$$\mu - \lambda_1 = \sum_{i=2}^{n} (\lambda_i - \lambda_1) \alpha_i^2.$$

Finally,

$$\|\lambda_i - \lambda_1\| \leq 2\|A\|$$

yields

$$|\mu - \lambda_1| \leq 2\|A\|(1 - \alpha_1^2) \leq 4\|A\|(1 - \alpha_1) = 2\|A\|\|x - x_1\|^2. \qquad \square$$

THEOREM 4.4. *Suppose that $F'(u^*)$ is nonsingular. If the sequence $\{u_k\}$ converges to u^* as $k \to \infty$, then for k large enough there exists a constant c such that*

$$\|u_{k+1} - u^*\| \leq c\|u_k - u^*\|^2.$$

Proof. We notice that we obtain $u'_{k+1} := (\widetilde{x}_{k+1}, \widetilde{y}_{k+1}, \rho_{k,1} + \Delta\lambda_k, \rho_{k,2} + \Delta\mu_k)$ by Newton's method using the initial approximation u_k. Since $F'(u^*)$ is nonsingular, Newton's method has a quadratic convergence rate and for k large enough there exists a constant c_1, such that

$$\|\widetilde{x}_{k+1} - x^*\|, \|\widetilde{y}_{k+1} - y^*\| \leq c_1 \|u_k - u^*\|^2.$$

It is easy to see that there exists a constant c_2, such that

$$\|x_{k+1} - x^*\|, \|y_{k+1} - y^*\| \le c_2 \|u_k - u^*\|^2. \tag{4.14}$$

In order to complete the proof we have to find a suitable bound for $\|\rho_{k+1} - \rho^*\|$. We can write the system (1.3) in the form

$$\begin{aligned} G_1 w &= \lambda w, \\ G_2 w &= \mu w, \end{aligned}$$

where $G_i = \Delta_0^{-1/2}\Delta_i\Delta_0^{-1/2}$ for $i = 1, 2$ and $w = \Delta_0^{1/2} z$. Let us write

$$\begin{aligned} z_{k+1} &= x_{k+1} \otimes y_{k+1}, \\ w_{k+1} &= \Delta_0^{1/2} z_{k+1}, \\ z^* &= x^* \otimes y^*, \\ w^* &= \Delta_0^{1/2} z^*. \end{aligned}$$

The tensor Rayleigh quotient $(\rho_{k+1,1}, \rho_{k+1,2}) = \rho(x_{k+1}, y_{k+1}, A_1, A_2, B_1, B_2, C_1, C_2)$ is equal to

$$\begin{aligned} \rho_{k+1,1} &= \frac{\langle G_1 w_{k+1}, w_{k+1}\rangle}{\langle w_{k+1}, w_{k+1}\rangle}, \\ \rho_{k+1,2} &= \frac{\langle G_2 w_{k+1}, w_{k+1}\rangle}{\langle w_{k+1}, w_{k+1}\rangle}. \end{aligned}$$

The matrices G_1, G_2 are symmetric and we can apply Lemma 4.3. It follows that

$$\|\rho_{k+1} - \rho^*\| \le c_3 \|w_{k+1} - w^*\|^2 \le c_3 \|\Delta_0^{1/2}\|^2 \|z_{k+1} - z^*\|^2, \tag{4.15}$$

where $c_3 := 2\max(\|G_1\|, \|G_2\|)$. Together with (4.14) this gives a constant c_4, such that

$$\|\rho_{k+1} - \rho^*\| \le c_4 \|u_k - u^*\|^4. \tag{4.16}$$

The proof follows from the bounds (4.14) and (4.16). $\square$

Let us remark that the bound (4.16) shows that in TRQI the eigenvalue part converges faster than in Newton's method. This justifies the numerical results in the next section.

Another generalization of Rayleigh quotient was studied by Blum and Chang in [6] and by Ji, Jiang and Lee in [18]. In the classical eigenvalue problem $Ax = \lambda x$ the Rayleigh quotient

$$\rho(x) = \frac{x^T A x}{x^T x}$$

minimizes the residual $\|Ax - \mu x\|$ for a given $x \neq 0$ over all $\mu \in \mathbb{R}$. Based on this property we define the *generalized Rayleigh quotient* $\widetilde{\rho}(x, y, A_1, A_2, B_1, B_2, C_1, C_2)$ for the two-parameter problem (1.1) to be a pair $(\widetilde{\rho}_1, \widetilde{\rho}_2)$ which minimizes the norm of the residual

$$\left(\|(A_1 - s_1 B_1 - s_2 C_1)x\|^2 + \|(A_2 - s_1 B_2 - s_2 C_2)y\|^2 \right)^{1/2} \tag{4.17}$$

for a given $x, y \neq 0$ over all $s_1, s_2 \in \mathbb{R}$.

LEMMA 4.5. *The generalized Rayleigh quotient* $\widetilde{\rho}(x, y, A_1, A_2, B_1, B_2, C_1, C_2)$ *is equal to*

$$\begin{aligned} \widetilde{\rho}_1 &= \frac{\langle Az, Bz\rangle\langle Cz, Cz\rangle - \langle Az, Cz\rangle\langle Bz, Cz\rangle}{\langle Bz, Bz\rangle\langle Cz, Cz\rangle - \langle Bz, Cz\rangle^2}, \\ \widetilde{\rho}_2 &= \frac{\langle Az, Cz\rangle\langle Bz, Bz\rangle - \langle Az, Bz\rangle\langle Bz, Cz\rangle}{\langle Bz, Bz\rangle\langle Cz, Cz\rangle - \langle Bz, Cz\rangle^2}. \end{aligned} \tag{4.18}$$

where

$$z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}.$$

Proof. We consider the least squares problem

$$\min_{(\widetilde{\rho}_1, \widetilde{\rho}_2) \in \mathbb{R}^2} \|Az - \widetilde{\rho}_1 Bz - \widetilde{\rho}_2 Cz\| \tag{4.19}$$

and the corresponding overdetermined linear system

$$\begin{bmatrix} B_1 x & C_1 x \\ B_2 y & C_2 y \end{bmatrix} \begin{bmatrix} \widetilde{\rho}_1 \\ \widetilde{\rho}_2 \end{bmatrix} = \begin{bmatrix} A_1 x \\ A_2 y \end{bmatrix}. \tag{4.20}$$

It is obvious that the norm (4.17) is equal to the norm of the residual (4.19). Let us show that the overdetermined system (4.20) has rank 2. Namely, if the $K \times 2$ matrix

$$\begin{bmatrix} B_1 x & C_1 x \\ B_2 y & C_2 y \end{bmatrix} \tag{4.21}$$

has rank smaller than 2, then if follows that

$$\begin{vmatrix} x^T B_1 x & x^T C_1 x \\ y^T B_2 y & y^T C_2 y \end{vmatrix} = D(x, y) = 0$$

which contradicts the fact that the problem (1.1) is right definite.

Since the matrix (4.21) has rank 2, the least squares solution of (4.20) can be expressed as the solution of the normal equations

$$\begin{bmatrix} (B_1 x)^T & (B_2 y)^T \\ (C_1 y)^T & (C_2 y)^T \end{bmatrix} \begin{bmatrix} B_1 x & C_1 x \\ B_2 y & C_2 y \end{bmatrix} \begin{bmatrix} \widetilde{\rho}_1 \\ \widetilde{\rho}_2 \end{bmatrix} = \begin{bmatrix} (B_1 x)^T & (B_2 y)^T \\ (C_1 y)^T & (C_2 y)^T \end{bmatrix} \begin{bmatrix} A_1 x \\ A_2 y \end{bmatrix}. \tag{4.22}$$

Using Cramer's rule it is easy to see that the solution of the system (4.22) is identical to (4.18). □

If we compare the formulas (4.8) and (4.18) then we see that the tensor Rayleigh quotient requires less operations to compute than the generalized Rayleigh quotient. If we use the generalized Rayleigh quotient instead of the tensor Rayleigh quotient in Step 1 of TRQI algorithm, then we call this method the Generalized Rayleigh Quotient Iteration algorithm (GRQI).

5. Numerical example. For a numerical example we take the following two-parameter system

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x - \lambda \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & -0.5 \\ 1 & -0.5 & 2 \end{bmatrix} x - \mu \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} x = 0,$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} y - \lambda \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} y - \mu \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 4 \end{bmatrix} y = 0.$$

Matrices are chosen in such a way that $(0,0)$ is a quadruple eigenvalue. For the construction of the homotopy (3.1) we take matrices

$$W_1 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad W_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

It is easy to see that $(0,0)$ is a quadruple eigenvalue for the homotopy (3.1) at $t = 1/2$. Thus the homotopy has a singular point at $t = 1/2$ which some of the homotopy curves $\vartheta_i(t)$ have to pass.

Using the gradient method we calculate the value

$$\delta_2 = \min\{ \min_{\|x\|=\|y\|=1} D(x,y,0), \min_{\|x\|=\|y\|=1} D(x,y,1)\}$$

and obtain $\delta_2 = 0.2188$. It follows from Lemma 3.8 that the smallest eigenvalue of Δ_0 is greater or equal to $\delta_2/3 = 0.0729$ (for a comparison, the exact smallest eigenvalue of Δ_0 is 0.2112). The upper bound $\|B_1\|\|C_2\| + \|C_1\|\|B_2\|$ from (3.21) is equal to 22.6292 (for a comparison, the exact largest eigenvalue of Δ_0 is 18.8374). Now we use Theorem 3.7 to obtain $\eta = 0.0064$. This is the value we use in the continuation method.

For the initial eigentuple $\vartheta(0)$ we take

$$\begin{aligned} \lambda(0) &= -1.0445, \\ \mu(0) &= -0.4529, \\ x(0) &= (0.8067, -0.3435, -0.4808)^T, \\ y(0) &= (0.4326, -0.8326, 0.3460)^T. \end{aligned}$$

We choose $h = 0.02$ as the initial step size and set $c_1 = 0.5$ and $c_2 = 0.1$. We set $\epsilon_1 = \epsilon_2 = 10^{-8}$ for the constants in the continuation method and TRQI (GRQI) algorithm, respectively. As a result, the continuation method using these initial values takes 94 steps to reach $t = 1$ from $t = 0$. In each step maximum 3 steps of TRQI algorithm and maximum 4 steps of GRQI algorithm are needed to reach the prescribed precision.

Next figures present the behaviour of the continuation method for these initial values. In Figure 5.1 we see the graphs of $\lambda(t)$ and $\mu(t)$. We see that the homotopy curve passes the singular point $(0,0)$ at $t = 1/2$. As it can be seen from Figures 5.2 and 5.3, where we see graphs of $x(t)$ and $y(t)$, no curve switching occurs in the singular point. Figure 5.4 depicts how the step size h changes from $t = 0$ to $t = 1$. If we compare Figure 5.4 with the previous ones, then it is clearly seen how the adaptive algorithm appropriately reduces or increases the step size h. We also notice that the step size h reduces near $t = 1$, where we have another singular point. It is clearly seen that the algorithm slows down in sensitive areas and speeds up when the homotopy curve is flat.

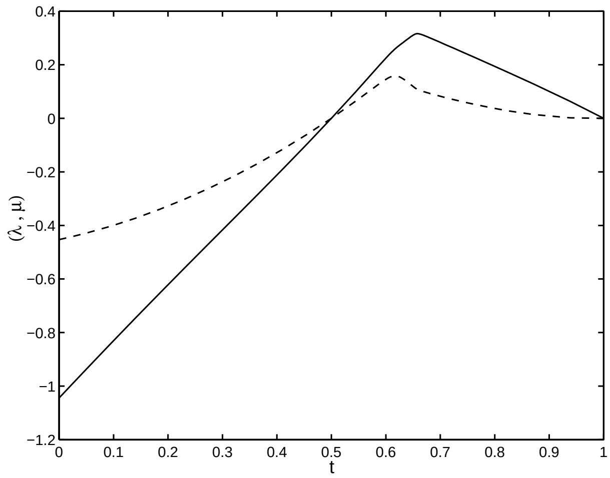

FIG. 5.1. *Eigenvalue pair* (λ, μ) *versus* t*. Solid curve:* $\lambda(t)$*; dashed curve:* $\mu(t)$*.*

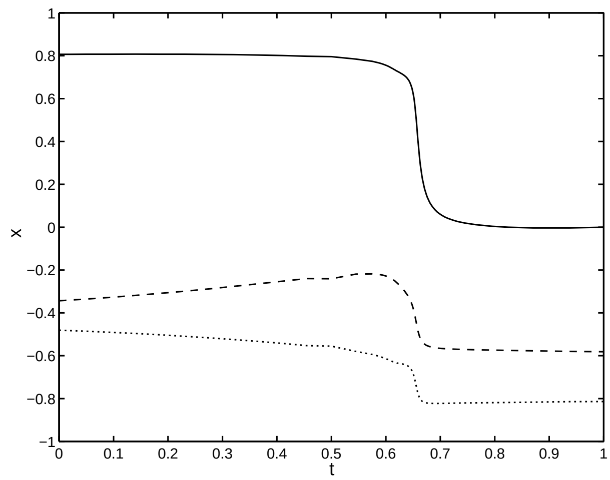

FIG. 5.2. *Eigenvector* x *versus* t*. Solid curve:* $x_1(t)$*; dashed curve:* $x_2(t)$*; dotted curve:* $x_3(t)$*.*

Similar results can be shown for the remaining 8 homotopy curves. Table 5.1 presents the number of steps it takes the continuation method to calculate the eigentuples from $t = 0$ to $t = 1$. In all homotopy curves the majority of steps is calculated on the interval $[0.6, 0.7]$. The only exception is $i = 6$ where the number of steps is so small because $\vartheta_6(t)$ is a constant. As it can be seen from Table 5.1 and also from Figure 5.4, the large number of steps is a result of a steep homotopy curve and not a result of passing a singular point.

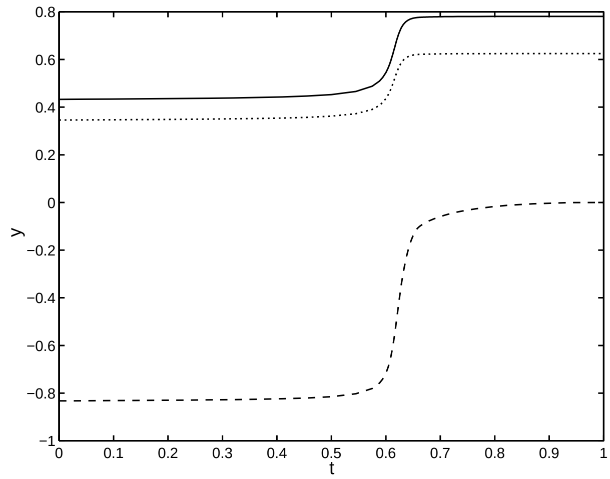

FIG. 5.3. *Eigenvector y versus t. Solid curve: $y_1(t)$; dashed curve: $y_2(t)$; dotted curve: $y_3(t)$.*

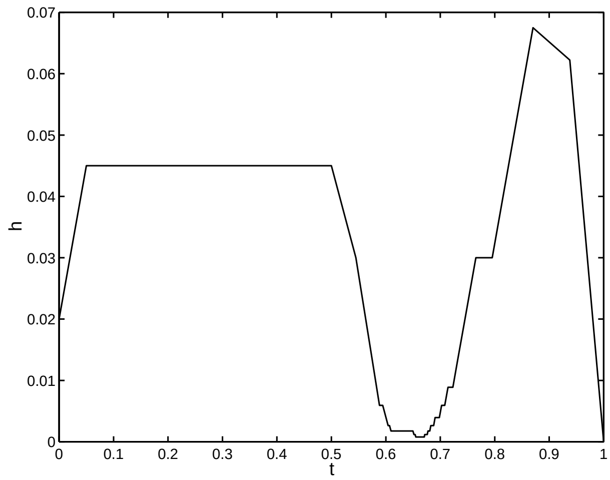

FIG. 5.4. *Step size h versus t.*

We also tested Newton's method without tensor or generalized Rayleigh quotient. The numbers in the fourth, the fifth and the sixth column of Table 5.1. represent the number of calculated Newton iterations, tensor and generalized Rayleigh quotients, respectively. As predicted in Theorem 4.4 TRQI converges faster than Newton's method. On the other hand, GRQI is most expensive to compute and has the slowest convergence among these three methods. Usually two steps of TRQI and three steps of GRQI are needed for one step of the continuation method. Based on these results and based on the theory from the previous section we recommend the use of TRQI algorithm instead of Newton's method or GRQI.

The number of steps could be further reduced by using better bounds for the norm of the difference between two distinct eigenvectors of the two-parameter right definite problem (1.1).

TABLE 5.1

Statistics for all 9 eigenvalues: starting $\lambda(0)$, $\mu(0)$ and obtained $\lambda(1)$, $\mu(1)$; number of steps of the continuation method; number of evaluations of Newton's method, TRQI and GRQI.

			No. of	Number of evaluations					passes
i	$\lambda(0)$	$\mu(0)$	steps	Newton	TRQI	GRQI	$\lambda(1)$	$\mu(1)$	$(0,0)$
1	-1.0445	0.8279	159	360	323	493	-0.2154	0.3659	no
2	1.0445	0.8279	174	399	346	534	1.8297	1.2842	no
3	0	0.8279	216	501	430	654	0.0090	0.4623	no
4	-1.0445	0	129	289	254	395	0	0	yes
5	1.0445	0	54	145	106	179	1.0445	0	no
6	0	0	10	21	9	9	0	0	yes
7	-1.0445	-0.4529	94	270	189	316	0	0	yes
8	1.0445	-0.4529	161	405	321	538	4.8629	-4.0229	no
9	0	-0.4529	104	253	208	311	0	0	yes

Acknowledgements. The author would like to thank Dr. Tomaž Košir for his helpful comments and suggestions.

REFERENCES

[1] E. L. ALLGOWER, K. GEORG, *Numerical Continuation Methods*, Springer–Verlag, Berlin, 1990.

[2] F. V. ATKINSON, *Multiparameter Spectral Theory*, Bull. Amer. Math. Soc., 74 (1968), pp. 1–27.

[3] ———, *Multiparameter Eigenvalue Problems*, Academic Press, New York, 1972.

[4] P. B. BAILEY, *The Automatic Solution of Two-Parameter Sturm–Liouville Eigenvalue Problems in Ordinary Differential Equations*, Appl. Math. Comp., 8 (1981), pp. 251–259.

[5] P. A. BINDING, P. J. BROWNE, X. JI, *A Numerical Method Using the Prüfer Transformation for the Calculation of Eigenpairs of Two-Parameter Sturm–Liouville Problems*, IMA J. Num. Anal., 13 (1993), pp. 559–569.

[6] E. K. BLUM, A. F. CHANG, *A Numerical Method for the Solution of the Double Eigenvalue Problem*, J. Inst. Math. Appl., 22 (1978), pp. 29–41.

[7] E. K. BLUM, A. R. CURTIS, *A Convergent Gradient Method for Matrix Eigenvector-Eigentuple Problems*, Numer. Math., 31 (1978), pp. 247–263.

[8] E. K. BLUM, P. B. GELTNER, *Numerical Solution of Eigentuple-Eigenvector Problems in Hilbert Spaces by a Gradient Method*, Numer. Math., 31: (1978), pp. 231–246.

[9] Z. BOHTE, *Numerical Solution of Some Two-Parameter Eigenvalue Problems*, Anton Kuhelj Memorial Volume, Slov. Acad. Sci. Art., Ljubljana (1982), pp. 17–28.

[10] P. J. BROWNE, B. D. SLEEMAN, *Analytic Perturbations of Multiparameter Eigenvalue Problems*, Quart. J. Math. Oxford (2), 30 (1979), pp. 257–263.

[11] ———, *A Numerical Technique for Multiparameter Eigenvalue Problems*, IMA J. of Num. Anal., 2 (1982), pp. 451–457.

[12] M. T. CHU, *A Simple Application of the Homotopy Method to Symmetric Eigenvalue Problems*, Lin. Alg. Appl., 59 (1984), pp. 85–90.

[13] M. T. CHU, T. Y. LI, T. SAUER, *Homotopy Method for General λ-matrix Problems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 528–536.

[14] L. FOX, L. HAYES, D. F. MAYERS, *The Double Eigenvalue Problem*, Topics in Numerical Analysis, Proc. Royal Irish Academy Conf. on Numerical Analysis, J. Miller, ed., Academic Press, New York, 1973, pp. 93–112.

[15] B. A. HARGRAVE, B. D. SLEEMAN, *The Numerical Solution of Two-Parameter Eigenvalue Problems in Ordinary Differential Equations with an Application to the Problem of Diffraction by a Plane Angular Sector*, J. Inst. Math. Appl., 14 (1974), pp. 9–22.

[16] X. JI, *Numerical Solution of Joint Eigenpairs of a Family of Commutative Matrices*, Appl. Math. Lett., 4 (1991), pp. 57–60.

[17] ———, *A Two-Dimensional Bisection Method for Solving Two-Parameter Eigenvalue Problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1085–1093.

[18] X. R. JI, H. JIANG, B. H. K. LEE, *A Generalized Rayleigh Quotient Iteration For Coupled Eigenvalue Problems*, unpublished.

[19] T. Y. LI, N. H. RHEE, *Homotopy Algorithm for Symmetric Eigenvalue Problems*, Numer. Math., 55 (1989), pp. 265–280.

[20] T. Y. LI, Z. ZENG, L. CONG, *Solving Eigenvalue Problems of Real Nonsymmetric Matrices*

with Real Homotopies, SIAM J. Numer. Anal., 29 (1992), pp. 229–248.

[21] R. E. Müller, *Numerical Solution of Multiparameter Eigenvalue Problems*, ZAMM 62 (1982), pp. 681–686.

[22] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice–Hall, Englewood Cliffs, NJ, 1980.

[23] M. Shimasaki, *Homotopy Algorithm for Two-Parameter Eigenvalue Problems* (Japanese), Trans. Japan SIAM, 5 (1995), pp. 121–129.

[24] ———, *Numerical Method Based on Homotopy Algorithm for Two-parameter Problems* (Japanese), Trans. Japan SIAM, 6 (1996), pp. 205-218.

[25] T. Slivnik, G. Tomšič, *A Numerical Method for the Solution of Two-Parameter Eigenvalue Problem*, J. Comp. Appl. Math., 15 (1986), pp. 109–115.

[26] H. Volkmer, *On the Minimal Eigenvalue of a Positive Definite Operator Determinant*, Proc. Roy. Soc. Edinburgh, 103A (1986), pp. 201–208.

[27] ———, *Multiparameter Problems and Expansion Theorems*, vol. 1356 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, New York, 1988.