SIAM J. MATRIX ANAL. APPL.
Vol. 42, No. 1, pp. 224–242

LOCATING CONICAL DEGENERACIES IN THE SPECTRA OF PARAMETRIC SELF-ADJOINT MATRICES*

GREGORY BERKOLAIKO† AND ADVAIT PARULEKAR‡

Abstract. A simple iterative scheme is proposed for locating the parameter values for which a two-parameter family of real symmetric matrices has a double eigenvalue. The convergence is proved to be quadratic. An extension of the scheme to complex Hermitian matrices (with three parameters) and to the location of triple eigenvalues (five parameters for real symmetric matrices) is also described. Algorithm convergence is illustrated in several examples: a real symmetric family, a complex Hermitian family, a family of matrices with an "avoided crossing" (no convergence), and a five-parameter family of real symmetric matrices with a triple eigenvalue.

Key words. conical points, eigenvalue multiplicity, Dirac points, dispersion relation, parametric families of self-adjoint matrices

AMS subject classifications. 15A18, 65F15, 81U30

DOI. 10.1137/20M134174X

1. Introduction.

A theorem of von Neumann and Wigner states that, generically, a two-parameter family of real symmetric matrices has multiple eigenvalues at isolated points [29]. In other words, the matrices with multiple eigenvalues have co-dimension 2 in the manifold of real symmetric matrices [1, Appendix 10]. In this paper, we would like to address the problem of locating these isolated points of eigenvalue multiplicity in the two-dimensional parameter space. To be more precise, we consider the following problem.

PROBLEM. *Given a smooth real symmetric matrix valued function* $A : \mathbb{R}^2 \mapsto \mathbb{R}^{n\times n}$*, locate the values of the parameters* (x, y) *which yield a matrix* $A(x, y)$ *with degenerate eigenvalues.*

To give a simple example, the function

$$A(x,y) = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}$$

has a double eigenvalue at the unique point $(x, y) = (0, 0)$. Its eigenvalues λ satisfy the equation $\lambda^2 = x^2 + y^2$, and the eigenvalue surface is a circular double cone in the space (x, y, λ). In contrast, the nonlinear function

$$A(x,y) = \begin{pmatrix} \cos(y)\sin(x) & 2 - 3\sin(y-x) \\ 2 - 3\sin(y-x) & 2\cos(y) - \sin(x) \end{pmatrix} \tag{1}$$

has multiple points of eigenvalue multiplicity; see Figure 1. Each point is isolated, and locally around each point the eigenvalue surface also looks like a cone.

*Received by the editors May 29, 2020; accepted for publication (in revised form) by P.-A. Absil December 10, 2020; published electronically March 1, 2021.
https://doi.org/10.1137/20M134174X
Funding: The work of the authors was supported by the National Science Foundation grant DMS-1815075 and the Binational Science Foundation grant 2016281.
†Department of Mathematics, Texas A&M University, College Station, TX 77843-3368 USA (gberkolaiko@gmail.com).
‡Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (adutheparulekar@gmail.com).

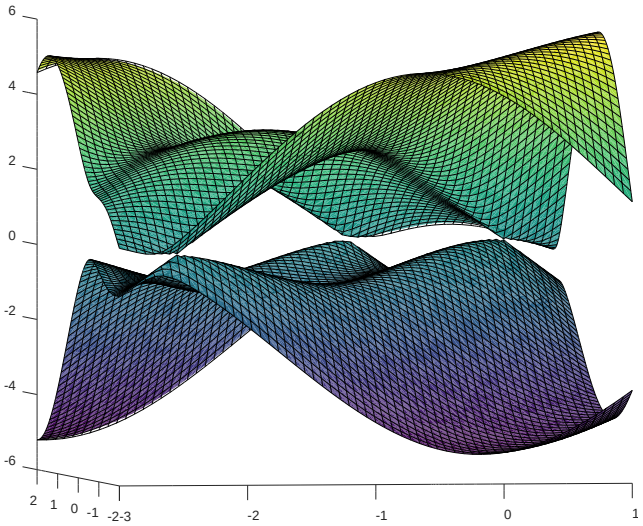

FIG. 1. *Eigenvalue surfaces corresponding to $A(x, y)$ from* (1). *There are three conical points; the surfaces appear to not touch at the middle point due to insufficient grid precision.*

For a family of complex Hermitian matrices, the co-dimension of the matrices with multiple eigenvalues is 3. Therefore, the analogous question can be posed about locating multiple eigenvalues of a Hermitian $A(x, y, z)$. We will formulate an extension of our results to complex Hermitian matrices but will concentrate on the real symmetric case in our proofs.

The problem of locating the points of eigenvalue multiplicity is of practical importance. In condensed matter physics [2] the wave propagation through a periodic medium is studied via Floquet–Bloch transform [19, 20] which results in a parametric family of self-adjoint operators (or matrices) with discrete spectrum. The eigenvalue surfaces (sheets of the "dispersion relation") may touch (see Figure 1), which has a profound effect on wave propagation and its sensitivity to a small perturbation of the medium. This touching corresponds precisely to a multiplicity in the eigenvalue spectrum. To give a well-studied example, the unusual electron properties of graphene occur due to the presence of eigenvalue multiplicity [6, 23]. It is also of practical relevance to be able to distinguish touching from "almost touching" (also known as "avoided crossing" in one-parameter problems).

The question of locating eigenvalue multiplicity in a family of 2×2 real symmetric matrices A has a straightforward solution (which also illustrates why the co-dimension is 2). The discriminant of $A \in \mathbb{R}^{2\times 2}$ can be written as a sum of two squares,

$$\operatorname{disc}(A) := (\lambda_1 - \lambda_2)^2 = (A_{11} - A_{22})^2 + 4A_{12}^2. \tag{2}$$

By definition, the discriminant is 0 if and only if two eigenvalues coincide. Therefore, we have two conditions that must simultaneously be met for the multiplicity to occur:

$$F(x, y) = \mathbf{0}, \qquad \text{where} \quad F : \mathbb{R}^2 \to \mathbb{R}^2, \qquad F(x, y) := \begin{pmatrix} A_{11}(x, y) - A_{22}(x, y) \\ A_{12}(x, y) \end{pmatrix}. \tag{3}$$

Unfortunately, for larger matrices the discriminant quickly becomes unwieldy and cannot be used in practical computations. The discriminant can still be written as a

sum of squares [17, 21, 25, 7], but the number of terms grows fast with the size of the matrix.

Thus, for an $n \times n$ real symmetric matrix $A(x, y)$ depending on two parameters x and y, there is only one easily computable function $\lambda_2(x, y) - \lambda_1(x, y)$ whose root, in variables x and y, we are seeking.[1] However, to apply a standard method with quadratic convergence, such as the Newton–Raphson algorithm, one needs two functions for two variables. One can search for the minimum of the square eigenvalue difference, $\big(\lambda_2(x, y) - \lambda_1(x, y)\big)^2$, which is smooth. But such a search would converge equally well to a point of "avoided crossing," a pitfall our proposed method manages to avoid; see sections 5.3 and 5.4.

One can change the basis to make $A(x, y)$ block-diagonal, with a 2×2 block corresponding to eigenvalues λ_1 and λ_2. The existence of this change in a neighborhood of the multiplicity point is assured (using Riesz projector) if $\lambda_{1,2}$ remain bounded away from the rest of the spectrum. However, the new basis will depend on the parameters (x, y) and is not directly accessible for numerical computations. Despite this obstacle, we will show that a "naive" approach produces equivalently good convergence: one can use a *constant* eigenvector basis which is recomputed[2] at each point of the Newton–Raphson iteration. More precisely, we establish the following theorem.

THEOREM 1.1. *Let $A(\mathbf{r}) : \mathbb{R}^2 \mapsto \mathbb{R}^{n\times n}$ be a real symmetric matrix valued function which is continuously twice differentiable in each entry, with a nondegenerate conical point (defined below) between λ_1 and λ_2 at parameter point α. For any $\mathbf{r}_i$, define $\mathbf{r}_{i+1}$ by*

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \begin{pmatrix} \left\langle v_1, \frac{\partial A}{\partial x} v_1 \right\rangle - \left\langle v_2, \frac{\partial A}{\partial x} v_2 \right\rangle & \left\langle v_1, \frac{\partial A}{\partial y} v_1 \right\rangle - \left\langle v_2, \frac{\partial A}{\partial y} v_2 \right\rangle \\ 2\left\langle v_1, \frac{\partial A}{\partial x} v_2 \right\rangle & 2\left\langle v_1, \frac{\partial A}{\partial y} v_2 \right\rangle \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 - \lambda_2 \\ 0 \end{pmatrix}, \tag{4}$$

where $\lambda_{1,2} = \lambda_{1,2}(\mathbf{r}_i)$ denote the eigenvalues of A at the point $\mathbf{r}_i$ and $v_{1,2} = v_{1,2}(\mathbf{r}_i)$ denote the corresponding eigenvectors.

Then there exists an open neighborhood $\Omega \subset \mathbb{R}^2$ of α and a constant $C > 0$ such that for all $\mathbf{r}_i \in \Omega$, the corresponding $\mathbf{r}_{i+1}$ satisfies the estimate

$$|\mathbf{r}_{i+1} - \alpha| < C|\mathbf{r}_i - \alpha|^2. \tag{5}$$

Before we prove this theorem in section 4, we explain in section 2 the geometrical picture behind the iterative procedure (4) and also point out the main differences between (4) and the Newton–Raphson method in a conventional setting. We also review related literature in section 2.1 once we introduce relevant notions. The precise definition and properties of "nondegenerate conical point" are given in section 3. Section 5 contains some computational examples.

1.1. Notation. We let $C^2(\mathbb{R}^2, \mathbb{R}^{n\times n})$ denote the set of matrix valued functions mapping $\mathbb{R}^2$ to $\mathbb{R}^{n\times n}$ with each element being continuously twice differentiable. The eigenvalues of the matrix-function $A \in C^2(\mathbb{R}^2, \mathbb{R}^{n\times n})$ are numbered in the increasing order $\lambda_1 \le \lambda_2 \le \lambda_3 \le \cdots \le \lambda_n$, and without loss of generality, we will look for $\mathbf{r} = (x, y) \in \mathbb{R}^2$ such that $\lambda_1(\mathbf{r}) = \lambda_2(\mathbf{r})$. Naturally, all results apply equally well to

[1] Here, without loss of generality, we have assumed that one is interested in the degeneracy $\lambda_1 = \lambda_2 < \lambda_3 < \cdots$.

[2] We are mostly motivated by the applications to tight-binding models of condensed matter physics [2] where the matrix dimension n is often of order 10 and computation of eigenvectors is relatively fast and precise. Another area of application is pointed out at the end of section 5.4.

any pair of consecutive eigenvalues. We remark that functions $\lambda_k(\mathbf{r})$ are continuous but not necessarily smooth: the points of eigenvalue multiplicity are typically the points where the eigenvalues involved are not differentiable; see Figure 1.

For any real symmetric matrix valued function A and any point $\mathbf{p} \in \mathbb{R}^2$, we let $A^{\mathbf{p}} = V^* A(\mathbf{r}) V$ denote the representation of A in the eigenvector basis computed at point $\mathbf{p}$. That is, V is a fixed orthogonal matrix whose columns are the eigenvectors of $A(\mathbf{p})$. The eigenvectors are assumed to be numbered according to the eigenvalue ordering. This means that $A^{\mathbf{p}} \in C^2(\mathbb{R}^2, \mathbb{R}^{n\times n})$ is a diagonal matrix at the point $\mathbf{p}$ but not necessarily anywhere else.

We let

$$\widetilde{A}^{\mathbf{p}}(\mathbf{r}) = \begin{pmatrix} A^{\mathbf{p}}_{11} & A^{\mathbf{p}}_{12} \\ A^{\mathbf{p}}_{21} & A^{\mathbf{p}}_{22} \end{pmatrix} := \begin{pmatrix} \langle v_1, A(\mathbf{r}) v_1 \rangle & \langle v_1, A(\mathbf{r}) v_2 \rangle \\ \langle v_2, A(\mathbf{r}) v_1 \rangle & \langle v_2, A(\mathbf{r}) v_2 \rangle \end{pmatrix} \tag{6}$$

denote the submatrix of $A^{\mathbf{p}}$ corresponding to the eigenvectors of the coalescing eigenvalues. We stress again that the eigenvectors $v_1 = v_1(\mathbf{p})$ and $v_2 = v_2(\mathbf{p})$ are computed at the point $\mathbf{p}$ and do not vary with $\mathbf{r}$. By the definition of $A^{\mathbf{p}}$, we have

$$\widetilde{A}^{\mathbf{p}}(\mathbf{p}) = \begin{pmatrix} \lambda_1(\mathbf{p}) & 0 \\ 0 & \lambda_2(\mathbf{p}) \end{pmatrix}. \tag{7}$$

We let

$$F\big(A^{\mathbf{p}}(\mathbf{r})\big) := \begin{pmatrix} A^{\mathbf{p}}_{11}(\mathbf{r}) - A^{\mathbf{p}}_{22}(\mathbf{r}) \\ 2A^{\mathbf{p}}_{12}(\mathbf{r}) \end{pmatrix} \tag{8}$$

denote the target function similar to (3). We stress that F is a function of $\mathbf{r}$.

Throughout this paper, D will denote the row vector of derivatives taken with respect to parameters $\mathbf{r} = (x, y)$,

$$\mathrm{D}f = \left(\frac{\partial f}{\partial x}, \ \frac{\partial f}{\partial y} \right).$$

If f is a vector-function, $\mathrm{D}f$ is a matrix with two columns. We use the notation $\mathrm{D}_{\mathbf{r}_0} f$ to denote the derivative evaluated at the point $\mathbf{r} = \mathbf{r}_0$, i.e.,

$$\mathrm{D}_{\mathbf{r}_0} f = \left(\frac{\partial f}{\partial x}(\mathbf{r}_0), \ \frac{\partial f}{\partial y}(\mathbf{r}_0) \right).$$

We use notation $J_{\mathbf{r}}(A^{\mathbf{p}})$ to denote the Jacobian of $F(A^{\mathbf{p}})$,

$$J_{\mathbf{r}}(A^{\mathbf{p}}) := \mathrm{D}_{\mathbf{r}} F(A^{\mathbf{p}}) = \begin{pmatrix} \big\langle v_1, \frac{\partial A}{\partial x} v_1 \big\rangle - \big\langle v_2, \frac{\partial A}{\partial x} v_2 \big\rangle & \big\langle v_1, \frac{\partial A}{\partial y} v_1 \big\rangle - \big\langle v_2, \frac{\partial A}{\partial y} v_2 \big\rangle \\ 2\big\langle v_1, \frac{\partial A}{\partial x} v_2 \big\rangle & 2\big\langle v_1, \frac{\partial A}{\partial y} v_2 \big\rangle \end{pmatrix}, \tag{9}$$

where v_1, v_2 are the eigenvectors of $A(\mathbf{p})$ and the derivatives $\frac{\partial A}{\partial x}$ and $\frac{\partial A}{\partial y}$ have been evaluated at point $\mathbf{r}$. This is the matrix appearing in Theorem 1.1. The factor 2 in the definition of $J_{\mathbf{r}}(A^{\mathbf{p}})$ arises naturally in calculations; it can also be used to put the second row terms in the more symmetric form,

$$2\Big\langle v_1, \frac{\partial A}{\partial x} v_2 \Big\rangle = \Big\langle v_1, \frac{\partial A}{\partial x} v_2 \Big\rangle + \Big\langle v_2, \frac{\partial A}{\partial x} v_1 \Big\rangle.$$

Finally, we remark that by our definitions $F(A) = F\big(\widetilde{A}\big)$ and $J_{\mathbf{r}}(A) = J_{\mathbf{r}}\big(\widetilde{A}\big)$. Therefore, the tilde (defined in (6)) will usually be omitted once we invoke functions F and J.

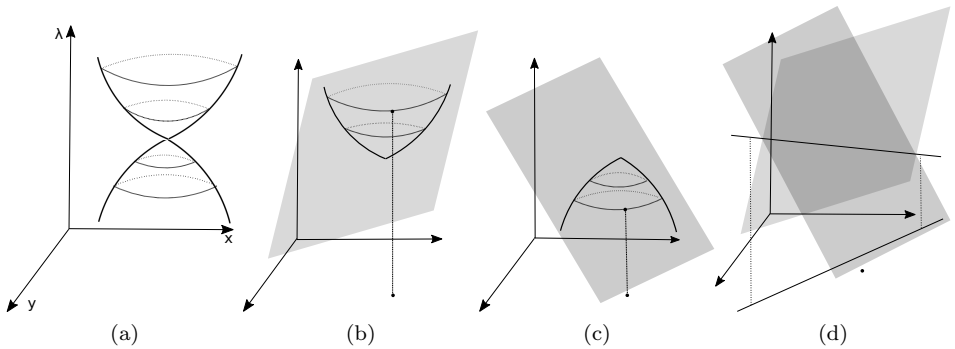

FIG. 2. (a) *Conical degeneracy of eigenvalues.* (b) *Linear approximation of the top eigenvalue about the initial guess.* (c) *Linear approximation of the bottom eigenvalue about the initial guess.* (d) *The intersection of the two linear approximations is a line, not a point. We need to use the conical nature of the intersection to determine a unique point to choose as our next guess.*

2. Discussion.

2.1. Geometric interpretation. What is described in this paper is a variation of the Newton–Raphson method searching for a zero of the objective function $\lambda_1(\mathbf{r}) - \lambda_2(\mathbf{r})$. This is only one condition on two parameters (in the real symmetric case), and leads to an underdetermined Newton–Raphson iteration. In particular, given an initial guess $\mathbf{r}_0$, we would like to update our guess to $\mathbf{r}_1$ such that

$$\mathrm{D}_{\mathbf{r}_0}\big(\lambda_1(\mathbf{r}) - \lambda_2(\mathbf{r})\big)\,(\mathbf{r}_1 - \mathbf{r}_0) = -\big(\lambda_1(\mathbf{r}_0) - \lambda_2(\mathbf{r}_0)\big). \tag{10}$$

However, there is a whole line of points $\mathbf{r}_1$ that satisfy this condition, as illustrated in Figure 2.

To incorporate our knowledge that the degeneracy occurs at an isolated point, we use a heuristic derived from Berry phase [14, 4, 27], a phenomenon which underlies the inability to find a smooth diagonalization around a degeneracy: on a loop in the parameter space around a nondegenerate conical point, a continuous choice of eigenvectors must rotate by π (as opposed to 0 mod 2π).

But if smoothly going in a loop around the degeneracy rotates the eigenvectors, the direction of minimal rotation is a direction *towards the point of degeneracy.* Let $\{v_1(\mathbf{r}), v_2(\mathbf{r})\}$ be a smooth choice of normalized eigenvectors around the point $\mathbf{r}_0$ (this is possible because $\mathbf{r}_0$ is not a point of eigenvalue multiplicity). Then we are looking for the direction in the parameter space in which the eigenvector v_1 as a function of $\mathbf{r}$ does not rotate in the plane spanned by $\{v_1(\mathbf{r}_0), v_2(\mathbf{r}_0)\}$ (it may still rotate "out of the plane"). This condition can be written as

$$\mathrm{D}_{\mathbf{r}_0}\big\langle v_1(\mathbf{r}), v_2(\mathbf{r}_0)\big\rangle\,(\mathbf{r}_1 - \mathbf{r}_0) = 0. \tag{11}$$

Conditions (10) and (11) together generically[3] define a unique point $\mathbf{r}$ which can be taken as the next step in the iteration. We can solve for it explicitly using the well-known perturbation formulas [5, 18],

$$\mathrm{D}_{\mathbf{r}_0}\lambda_1 = \mathrm{D}_{\mathbf{r}_0}A^{\mathbf{r}_0}_{11}, \qquad \mathrm{D}_{\mathbf{r}_0}\lambda_2 = \mathrm{D}_{\mathbf{r}_0}A^{\mathbf{r}_0}_{22}, \tag{12}$$

$$\mathrm{D}_{\mathbf{r}_0}\big\langle v_1(\mathbf{r}), v_2(\mathbf{r}_0)\big\rangle = \frac{\mathrm{D}_{\mathbf{r}_0}A^{\mathbf{r}_0}_{12}}{\lambda_1 - \lambda_2}, \tag{13}$$

[3] See sections 3 and 4 for a precise formulation.

where

$$A_{ij}^{\mathbf{r}_0} = A_{ij}^{\mathbf{r}_0}(\mathbf{r}) = \big\langle v_i(\mathbf{r}_0), A^{\mathbf{r}_0}(\mathbf{r}) v_j(\mathbf{r}_0) \big\rangle. \tag{14}$$

We stress that in (14) the eigenvectors v_1, v_2 are evaluated at the point $\mathbf{r}_0$ and do not depend on $\mathbf{r}$.

The tangent planes condition (10) and the nonrotation condition (11) can now be written succinctly as

$$\left[\mathrm{D}_{\mathbf{r}_0} \begin{pmatrix} A_{11}^{\mathbf{r}_0} - A_{22}^{\mathbf{r}_0} \\ 2A_{12}^{\mathbf{r}_0} \end{pmatrix}\right] (\mathbf{r}_1 - \mathbf{r}_0) = \big[\mathrm{D}_{\mathbf{r}_0} F\big(A^{\mathbf{r}_0}(\mathbf{r})\big)\big] (\mathbf{r}_1 - \mathbf{r}_0) = \begin{pmatrix} \lambda_2 - \lambda_1 \\ 0 \end{pmatrix}, \tag{15}$$

or, less succinctly, as

$$\begin{pmatrix} \big\langle v_1, \frac{\partial A}{\partial x} v_1 \big\rangle - \big\langle v_2, \frac{\partial A}{\partial x} v_2 \big\rangle & \big\langle v_1, \frac{\partial A}{\partial y} v_1 \big\rangle - \big\langle v_2, \frac{\partial A}{\partial y} v_2 \big\rangle \\ 2\big\langle v_1, \frac{\partial A}{\partial x} v_2 \big\rangle & 2\big\langle v_1, \frac{\partial A}{\partial y} v_2 \big\rangle \end{pmatrix} (\mathbf{r}_1 - \mathbf{r}_0) = \begin{pmatrix} \lambda_2 - \lambda_1 \\ 0 \end{pmatrix},$$

which immediately leads to (4).

Berry phase also lies at the heart of another set of works devoted to locating points of eigenvalue multiplicity. Pugliese, Dieci, and co-authors [26, 9, 10, 11, 8] developed a procedure which uses Berry phase to grid-search available space and identify regions with conical points. For the final convergence they used the standard Newton–Raphson method to locate the critical point of $(\lambda_2 - \lambda_1)^2$. The convergence rate of this final step is quadratic, as in Theorem 1.1; we refer to section 5.4 for a comparison of actual convergence in an example.

In terms of ease of application, coding (4) is straightforward and lack of convergence of the method also carries information (see sections 5.3 and 5.4). To perform a thorough search of all available space and to locate all conical points, it is preferable to use the methods of [26, 11, 8].

2.2. Relation to Newton–Raphson method. Recalling the definition of $\widetilde{A}^{\mathbf{r}_0}$ and, in particular, (7), we have

$$\begin{pmatrix} \lambda_2 - \lambda_1 \\ 0 \end{pmatrix} = -F\big(A^{\mathbf{r}_0}(\mathbf{r}_0)\big).$$

This allows us to rewrite (15) as

$$\big[\mathrm{D}_{\mathbf{r}_0} F\big(A^{\mathbf{r}_0}(\mathbf{r})\big)\big] (\mathbf{r}_1 - \mathbf{r}_0) = -F\big(A^{\mathbf{r}_0}(\mathbf{r}_0)\big),$$

which is the same as a single step of Newton–Raphson iteration applied to $F(\widetilde{A}^{\mathbf{r}_0})$. In other words, $\mathbf{r}_1 = (x_1, y_1)$ is chosen to be a solution to

$$\widetilde{A}^{\mathbf{r}_0}(\mathbf{r}_0) + (x_1 - x_0) \frac{\partial \widetilde{A}^{\mathbf{r}_0}}{\partial x}(\mathbf{r}_0) + (y_1 - y_0) \frac{\partial \widetilde{A}^{\mathbf{r}_0}}{\partial y}(\mathbf{r}_0) = \lambda I_2 \tag{16}$$

for some $\lambda \in \mathbb{R}$. Equivalently, $\mathbf{r}_1$ is the point where the linear approximation to $\widetilde{A}^{\mathbf{r}_0}(\mathbf{r})$ has a double eigenvalue.

To understand the difference of our algorithm from a seemingly conventional Newton–Raphson method, we need to revisit the computation of $\widetilde{A}$. It can be viewed as first expressing $A(\mathbf{r})$ in the eigenvector basis computed *at the point* $\mathbf{r}_0$ and then extracting the $\{1, 2\}$-subblock of the resulting matrix.

In this notation, the problem of finding the degeneracy is equivalent to finding a point $\mathbf{r}'$ such that

$$\widetilde{A}^{\mathbf{r}'}(\mathbf{r}') = \lambda I_2 \qquad \text{for some } \lambda \in \mathbb{R}. \tag{17}$$

In contrast, solving (16) is a first step in finding a point $\mathbf{r}'$ such that

$$\widetilde{A}^{\mathbf{r}_0}(\mathbf{r}') = \lambda I_2 \qquad \text{for some } \lambda \in \mathbb{R}. \tag{18}$$

Going all the way to find the solution $\mathbf{r}'$ to (18) is pointless; this is not the equation we need to solve. Instead, we go one step, computing the first Newton–Raphson approximation $\mathbf{r}_1$, and then update our target equation to

$$\widetilde{A}^{\mathbf{r}_1}(\mathbf{r}') = \lambda I_2 \qquad \text{for some } \lambda \in \mathbb{R},$$

compute the first Newton–Raphson approximation $\mathbf{r}_2$ to *that* equation, and so on.

2.3. Complex Hermitian matrices. Let us now consider a complex Hermitian matrix valued function $A \in C^2(\mathbb{R}^3, \mathbb{C}^{n\times n})$. To find a point of eigenvalue multiplicity, we typically need three real parameters (the off-diagonal terms can be complex, and that introduces an additional degree of freedom), which we still denote by $\mathbf{r} = (x, y, z)$.

The conditions can now be written as

$$\left[\mathrm{D}_{\mathbf{r}_0} G\big(A^{\mathbf{r}_0}(\mathbf{r})\big)\right](\mathbf{r}_1 - \mathbf{r}_0) = \begin{pmatrix} \lambda_2 - \lambda_1 \\ 0 \\ 0 \end{pmatrix}, \tag{19}$$

where

$$G\big(A^{\mathbf{r}_0}\big) = \begin{pmatrix} A^{\mathbf{r}_0}_{11} - A^{\mathbf{r}_0}_{22} \\ 2A^{\mathbf{r}_0}_{12} \\ 2A^{\mathbf{r}_0}_{21} \end{pmatrix}. \tag{20}$$

One can equivalently use the objective function

$$G\big(A^{\mathbf{r}_0}\big) = \begin{pmatrix} A^{\mathbf{r}_0}_{11} - A^{\mathbf{r}_0}_{22} \\ 2\,\mathrm{Re}(A^{\mathbf{r}_0}_{12}) \\ 2\,\mathrm{Im}(A^{\mathbf{r}_0}_{21}) \end{pmatrix}. \tag{21}$$

3. Conical intersection. Let α be a point in the parameter space such that $A(\alpha)$ has a double eigenvalue $\lambda_1 = \lambda_2$. The existence of eigenvalue multiplicity precludes a smooth diagonalization in a region containing the degeneracy. However, a smooth block-diagonalization exists. The standard construction (see, for example, [18, II.4.2 and Remark 4.4 therein]) uses a Riesz projector.

We can choose a contour $\gamma : [0, 1] \mapsto \mathbb{C}$ with $\gamma(0) = \gamma(1)$ enclosing λ_1, λ_2 and no other point in the spectrum of $A(\alpha)$. This property of γ must persist for $A(\mathbf{r})$ when $\mathbf{r}$ is in a small neighborhood of α. The Riesz projector

$$P(\mathbf{r}) = \int_\gamma (A(\mathbf{r}) - \lambda I_n)^{-1} d\lambda \tag{22}$$

projects onto the space spanned by the eigenvectors of $\lambda_1(\mathbf{r})$ and $\lambda_2(\mathbf{r})$ [15]. The projector itself is smooth, as the points on the contour are all in the resolvent set of A (and so $A - \lambda I_n$ has a bounded inverse for all $\lambda \in \Gamma$). Starting with an arbitrary

eigenvector basis $\{v_1, v_2\}$ at α, we can obtain a basis at a nearby $\mathbf{r}$ by applying the Gram–Schmidt procedure to the set $\{P(\mathbf{r})v_1, P(\mathbf{r})v_2\}$, which preserves smoothness. We can do the same with the orthogonal complement $I - P(\mathbf{r})$ and a complementary basis to $\{v_1, v_2\}$. To summarize, for some region $\Omega \in \mathbb{R}^2$ with $\alpha \in \Omega$, we find a change of basis $M(\cdot) \in C^2(\Omega, R^{n\times n})$ such that

$$M(\mathbf{r})^* A(\mathbf{r}) M(\mathbf{r}) = B(\mathbf{r}) \oplus \Lambda(\mathbf{r}), \tag{23}$$

where $B \in C^2(\Omega, \mathbb{R}^{2\times 2})$ and $\Lambda \in C^2(\Omega, \mathbb{R}^{(n-2)\times(n-2)})$. We can further diagonalize both B and A at any point $\mathbf{r}_0$ to obtain

$$\Gamma(\mathbf{r})^* A^{\mathbf{r}_0}(\mathbf{r}) \Gamma(\mathbf{r}) = B^{\mathbf{r}_0}(\mathbf{r}) \oplus \Lambda(\mathbf{r}), \tag{24}$$

where $\Gamma(\mathbf{r}) = V^T M(\mathbf{r})(W \oplus I_{n-2}) \in C^2(\Omega, R^{n\times n})$, and both

$$A^{\mathbf{r}_0}(\cdot) := V^T A(\cdot) V \qquad \text{and} \qquad B^{\mathbf{r}_0}(\cdot) := W^T B(\cdot) W$$

are diagonal at $\mathbf{r}_0$. A stronger result from Hsieh and Sibuya [16] and Gingold [13] states that such block-diagonalization exists even for matrices that are not necessarily Hermitian, and for any closed rectangular region that contains an isolated degeneracy.

Note that since B is a 2×2 matrix which has an eigenvalue multiplicity at the point α, $B(\alpha)$ is a multiple of the identity. The eigenvalue multiplicity is detected by the *discriminant* of B which in the 2×2 case is defined as

$$\operatorname{disc}(B) := (\lambda_1 - \lambda_2)^2 = (B_{11} - B_{22})^2 + 4B_{12}^2. \tag{25}$$

The discriminant achieves its minimum value 0 at the point α. It is also a C^2 function of $\mathbf{r}$ and its Hessian is well-defined.

DEFINITION 3.1. *A point of eigenvalue multiplicity α is* a nondegenerate conical point *if* $\operatorname{disc}(B(\mathbf{r}))$ *has a nondegenerate critical point at* $\mathbf{r} = \alpha$.

In other words, there is a positive definite matrix H (the "Hessian") such that

$$\operatorname{disc}(B(\mathbf{r})) = \big\langle (\mathbf{r} - \alpha), H(\mathbf{r} - \alpha) \big\rangle + o\big(|\mathbf{r} - \alpha|^2\big)\,,$$

and therefore, along any ray originating at α, the eigenvalues are separating at a nonzero linear rate. This picture justifies the use of the term "conical."

Unfortunately, while existence of $B(\mathbf{r})$ is assured, it is not easily accessible. The following theorem provides a more practical method of checking if α is nondegenerate.

THEOREM 3.2. *The Hessian of* $\operatorname{disc}(B)$ *at α is given by*

$$\operatorname{Hess}_\alpha(\operatorname{disc}(B)) = 2J_\alpha(B)^T J_\alpha(B) = 2J_\alpha(A^\alpha)^T J_\alpha(A^\alpha). \tag{26}$$

Consequently, α is a nondegenerate conical point if and only if $\det J_\alpha(A^\alpha) \neq 0$.

The condition $\det J_\alpha(A^\alpha) \neq 0$ has a nice geometric meaning: it is precisely the condition that the manifold $\widetilde{A}^\alpha$ of 2×2 real symmetric matrices is transversal to the line of 2×2 symmetric matrices with repeated eigenvalues (cf. [24, Def. 1]).

The choice of basis in the definition of $\widetilde{A}^\alpha$ is assumed to align with the choice of basis used to compute $B(\mathbf{r})$, i.e., the first two columns of $M(\alpha)$ are the eigenvectors used to compute $\widetilde{A}^\alpha$. This choice does not affect the definition of the nondegenerate point because of the following lemma.

LEMMA 3.3. *Let $A \in C^2(\mathbb{R}^2, \mathbb{R}^{2\times 2})$ be a 2×2 matrix valued function of $\mathbf{r} \in \mathbb{R}^2$. Then for any orthogonal matrix $U \in \mathbb{R}^{2\times 2}$ there is an orthogonal matrix $W \in \mathbb{R}^{2\times 2}$ such that for all $\mathbf{r}$ we have*

$$F(U^T AU) = WF(A), \qquad J_{\mathbf{r}}(U^T AU) = WJ_{\mathbf{r}}(A), \tag{27}$$

and therefore,

$$|\det(J_{\mathbf{r}}(A))| = \left|\det(J_{\mathbf{r}}(U^T AU))\right|. \tag{28}$$

Proof. This identity for 2×2 matrix-functions can be checked by direct computation, but the details are excessively tedious. Instead, we use a more generalizable approach.

We fix an orthogonal U, and let $\mathcal{S}^2$ denote the linear space of 2×2 real symmetric matrices. The map F (see (8)), acts as a linear transformation from $\mathcal{S}^2$ to $\mathbb{R}^2$. It is obviously onto and has the kernel $\mathrm{Ker}(F)$ consisting of multiples of the identity. On the other hand, conjugation by U (namely the map $A \mapsto U^T AU$) is a linear transformation of $\mathcal{S}^2$ to itself. It maps multiples of the identity to themselves and therefore induces a linear transformation from the quotient space $\mathcal{S}^2/\mathrm{Ker}(F)$ to itself. This linear transformation, via the isomorphism F between $\mathcal{S}^2/\mathrm{Ker}(F)$ and $\mathbb{R}^2$, induces a linear transformation on $\mathbb{R}^2$ mapping $F(A)$ to $F(U^T AU)$.

We summarize the above in the commutative diagram

$$\begin{array}{ccc} \mathcal{S}^2 & \xrightarrow{A\mapsto U^T AU} & \mathcal{S}^2 \\ {\scriptstyle F}\downarrow & & {\scriptstyle F}\downarrow \\ \mathbb{R}^2 & \xrightarrow{\quad W \quad} & \mathbb{R}^2 \end{array}$$

In other words, for a given orthogonal U, there exists a constant 2×2 matrix W such that

$$F(U^T AU) = WF(A).$$

From the identity (see (25) for the definition of discriminant)

$$|F(A)|^2 = \mathrm{disc}(A) = \mathrm{disc}(U^T AU) = \left|F(U^T AU)\right|^2,$$

we conclude that W is orthogonal. Finally, taking derivatives we get

$$J(U^T AU) = WJ(A) \quad \Longrightarrow \quad \det(J(U^T AU)) = \det(WJ(A)) = \pm\det(J(A)),$$

since the determinant of an orthogonal matrix is either 1 or -1. ∎

The following identity will be helpful in the proof of Theorem 3.2 and also in section 4.

LEMMA 3.4. *For any $A^{\mathbf{r}_0}$ and $B^{\mathbf{r}_0}$ as in* (24),

$$J_{\mathbf{r}_0}(B^{\mathbf{r}_0}) = J_{\mathbf{r}_0}(A^{\mathbf{r}_0}) + 2(\lambda_2 - \lambda_1)\begin{pmatrix} 0 & 0 \\ \left\langle \frac{\partial\gamma_1}{\partial x}, \gamma_2 \right\rangle & \left\langle \frac{\partial\gamma_1}{\partial y}, \gamma_2 \right\rangle \end{pmatrix}, \tag{29}$$

where $\gamma_{1,2} = \gamma_{1,2}(\mathbf{r}_0)$ are the first two columns of the matrix $\Gamma(\mathbf{r}_0)$.

Proof. We remark that identity (29) is only claimed for the Jacobian evaluated at the point where both $A^{\mathbf{r}_0}$ and $B^{\mathbf{r}_0}$ are diagonal. Therefore, $A^{\mathbf{r}_0}\gamma_j(\mathbf{r}_0) = \lambda_j(\mathbf{r}_0)\gamma_j(\mathbf{r}_0)$.

For all $\mathbf{r}$, $\gamma_j(\mathbf{r})$ are orthonormal, and differentiating $\langle\gamma_i, \gamma_j\rangle = \text{const}$, we get

$$\left\langle \frac{\partial\gamma_i}{\partial x}, \gamma_j \right\rangle = -\left\langle \gamma_i, \frac{\partial\gamma_j}{\partial x} \right\rangle. \tag{30}$$

We can now relate the derivatives of $A^{\mathbf{r}_0}$ to the derivatives of $B^{\mathbf{r}_0}$,

$$\begin{aligned}
\frac{\partial}{\partial x}(B_{ij}^{\mathbf{r}_0}) &= \frac{\partial}{\partial x}\langle\gamma_j, A^{\mathbf{r}_0}\gamma_i\rangle \\
&= \left\langle \gamma_i, \frac{\partial A^{\mathbf{r}_0}}{\partial x}\gamma_j \right\rangle + \left\langle \frac{\partial\gamma_i}{\partial x}, A^{\mathbf{r}_0}\gamma_j \right\rangle + \left\langle \gamma_i, A^{\mathbf{r}_0}\frac{\partial\gamma_j}{\partial x} \right\rangle \\
&= \frac{\partial A_{ij}^{\mathbf{r}_0}}{\partial x} + \lambda_j \left\langle \frac{\partial\gamma_i}{\partial x}, \gamma_j \right\rangle + \lambda_i \left\langle \gamma_i, \frac{\partial\gamma_j}{\partial x} \right\rangle \\
&= \frac{\partial A_{ij}^{\mathbf{r}_0}}{\partial x} + (\lambda_j - \lambda_i)\left\langle \frac{\partial\gamma_i}{\partial x}, \gamma_j \right\rangle, \qquad i, j \in \{1, 2\}.
\end{aligned}$$

The calculation is identical for y derivatives. ∎

Proof of Theorem 3.2. We write

$$\operatorname{disc}(B) = (B_{11} - B_{22})^2 + 4B_{12}^2 = \langle F(B), F(B)\rangle,$$

and note that $F(B(\alpha)) = \mathbf{0}$. The latter observation implies that the product rule for the second derivatives at the point α collapses to

$$\frac{\partial^2}{\partial x_i \partial x_j}\langle F(B), F(B)\rangle = 2\left\langle \frac{\partial F(B)}{\partial x_i}, \frac{\partial F(B)}{\partial x_j} \right\rangle, \qquad x_i, x_j \in \{x, y\}.$$

Therefore, the Hessian can be written as

$$\operatorname{Hess}_\alpha\langle F(B), F(B)\rangle = 2\begin{pmatrix} \frac{\partial F(B)^T}{\partial x} \\ \frac{\partial F(B)^T}{\partial y} \end{pmatrix}\begin{bmatrix} \frac{\partial F(B)}{\partial x} & \frac{\partial F(B)}{\partial y} \end{bmatrix} = 2J_\alpha(B)^T J_\alpha(B).$$

Finally, setting $\mathbf{r}_0 = \alpha$ in Lemma 3.4 yields

$$J_\alpha(B) = J_\alpha(A^\alpha), \tag{31}$$

and concludes the proof of (26). ∎

4. Proof of the main result. Here we restate the procedure used to locate the degeneracy in the notation that has been introduced.

THEOREM 4.1. *Let* $\sigma\colon C^2(\mathbb{R}^2, \mathbb{R}^{2\times 2}) \times \mathbb{R}^2 \to \mathbb{R}^2$ *be defined by*

$$\sigma(S, \mathbf{r}) = \mathbf{r} - J_{\mathbf{r}}(S)^{-1}F_{\mathbf{r}}(S). \tag{32}$$

Let $A \in C^2(\mathbb{R}^2, \mathbb{R}^{n\times n})$ *have a nondegenerate conical point at* α *between eigenvalues* λ_1 *and* λ_2. *Then there exists an open* $\Omega \subset \mathbb{R}^2$ *with* $\alpha \in \Omega$ *and* $\exists\, C \in \mathbb{R}$, *such that for all* $\mathbf{r} \in \Omega$,

$$|\sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r}) - \alpha| < C|\mathbf{r} - \alpha|^2, \tag{33}$$

where the 2×2 *matrix-function* $\widetilde{A}^{\mathbf{r}}(\cdot) \in C^2(\mathbb{R}^2, \mathbb{R}^{2\times 2})$ *is defined by*

$$\widetilde{A}^{\mathbf{r}}(\cdot) = V^T A(\cdot) V, \tag{34}$$

with the constant $n \times 2$ *matrix* $V = (v_1\ v_2)$ *whose columns are the eigenvectors of* $A(\mathbf{r})$.

We remark that nondegeneracy of the conical point is a generic property: any degenerate conical point can be made nondegenerate by a small perturbation of the function A.

We recall that the superscript in $\widetilde{A}^{\mathbf{r}}(\cdot)$ refers to the basis which is computed at the point $\mathbf{r}$ and in which the matrix $A(x, y)$ is represented. The derivatives of $\widetilde{A}^{\mathbf{r}}(\cdot)$ that are taken to compute $J_{\mathbf{r}}$ in (32) are also evaluated at the point $\mathbf{r}$. The result of evaluating $\sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r})$ is explicitly written out in (4).

Proof. Let B be the matrix defined in (23). We will see in Lemmas 4.2 and 4.4 that there is a neighborhood $\Omega \subset \mathbb{R}^2$ of the conical point α, and constants $C_1, C_2 > 0$ such that for all $\mathbf{r} \in \Omega$ we have

$$|\sigma(B, \mathbf{r}) - \alpha| < C_1 |\mathbf{r} - \alpha|^2$$

and

$$|\sigma(B, \mathbf{r}) - \sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r})| < C_2 |\mathbf{r} - \alpha|^2.$$

Together, these give us

$$|\sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r}) - \alpha| < (C_1 + C_2) |\mathbf{r} - \alpha|^2,$$

as desired. ∎

Now we establish the lemmas used in the proof of Theorem 4.1.

LEMMA 4.2. *There exists* $\Omega_1 \subset \mathbb{R}^2$ *with* $\alpha \in \Omega_1$ *and* $C_1 \in \mathbb{R}$ *such that*

$$|\sigma(B, \mathbf{r}) - \alpha| < C_1 |\mathbf{r} - \alpha|^2, \tag{35}$$

when $\mathbf{r} \in \Omega_1$.

Proof. This is the usual Newton–Raphson method applied to conical point search for the 2×2 matrix B. For completeness we provide the proof. For the function $F(\mathbf{r}) := F(B(\mathbf{r}))$, we have the Taylor expansion around the point $\mathbf{r}_0$ which is evaluated at the point α,

$$\mathbf{0} = F(\alpha) = F(\mathbf{r}_0) + \mathrm{D}_{\mathbf{r}_0} F \cdot (\alpha - \mathbf{r}_0) + O(|\alpha - \mathbf{r}_0|^2),$$

where the constant in $O(|\alpha - \mathbf{r}_0|^2)$ is *independent* of $\mathbf{r}_0$ as long as it is in a neighborhood $\widetilde{\Omega}_1$ of α. The dot denotes the matrix-by-vector multiplication (to distinguish it from the argument of the function F).

By assumption $\det(J_\alpha) \neq 0$, and, by smoothness, we know that $\mathrm{D}_{\mathbf{r}_0} F = J_{\mathbf{r}_0}$ is boundedly invertible in some region $\Omega_1 \subset \widetilde{\Omega}_1$ containing α. Therefore, for the point $\mathbf{r}_1 = \sigma(B, \mathbf{r}_0)$, or, equivalently,

$$J_{\mathbf{r}_0} \cdot (\mathbf{r}_1 - \mathbf{r}_0) = -F(\mathbf{r}_0),$$

we have

$$\mathbf{0} = J_{\mathbf{r}_0} \cdot (\alpha - \mathbf{r}_1) + O(|\alpha - \mathbf{r}_0|^2),$$

with the estimate (35) following by inverting $J_{\mathbf{r}_0}$. ∎

LEMMA 4.3. *For any $B \in C^2(\mathbb{R}^2, \mathbb{R}^{n\times n})$ and constant, orthogonal U, we have*

$$\sigma(B, \mathbf{r}) = \sigma(U^T B U, \mathbf{r}). \tag{36}$$

Proof. Equation (36) follows directly from the definition of the one-step iteration function σ and Lemma 3.3. ∎

LEMMA 4.4. *There exists $\Omega_2 \subset \mathbb{R}^2$ with $\alpha \in \Omega_2$ and $C_2 \in \mathbb{R}$ such that*

$$|\sigma(B, \mathbf{r}) - \sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r})| < C_2 |\mathbf{r} - \alpha|^2, \tag{37}$$

when $\mathbf{r} \in \Omega_2$.

Proof. By the assumption that α is a nondegenerate conical point and (26), we have that $J_{\mathbf{r}}(B)$. Therefore, $J_{\mathbf{r}}(B^{\mathbf{r}})$ has a bounded inverse in a region around α. By (29) we conclude that $J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}})$ also has a bounded inverse in some region Ω_2 around α where $\lambda_1 - \lambda_2$ is small. We can express the difference of the inverses as

$$\begin{aligned} J_{\mathbf{r}}(B^{\mathbf{r}})^{-1} - J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}})^{-1} &= J_{\mathbf{r}}(B^{\mathbf{r}})^{-1} \left(J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}}) - J_{\mathbf{r}}(B^{\mathbf{r}}) \right) J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}})^{-1} \\ &= (\lambda_1 - \lambda_2) J_{\mathbf{r}}(B^{\mathbf{r}})^{-1} \begin{pmatrix} 0 & 0 \\ \left\langle \frac{\partial \gamma_1}{\partial x}, \gamma_2 \right\rangle & \left\langle \frac{\partial \gamma_1}{\partial y}, \gamma_2 \right\rangle \end{pmatrix} J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}})^{-1}, \end{aligned}$$

and so, using boundedness of Γ and its derivatives, we get

$$\left\| J_{\mathbf{r}}(B^{\mathbf{r}})^{-1} - J_{\mathbf{r}}(\widetilde{A}^{\mathbf{r}})^{-1} \right\| = O(\lambda_1 - \lambda_2).$$

We also recall that by definition of $A^{\mathbf{r}}$ and $B^{\mathbf{r}}$,

$$F(B^{\mathbf{r}}) = F(\widetilde{A}^{\mathbf{r}}) = \begin{pmatrix} \lambda_1(\mathbf{r}) - \lambda_2(\mathbf{r}) \\ 0 \end{pmatrix}.$$

Finally, abbreviating $J = J_{\mathbf{r}}$, we estimate

$$\begin{aligned} \left| \sigma(B^{\mathbf{r}}, \mathbf{r}) - \sigma(\widetilde{A}^{\mathbf{r}}, \mathbf{r}) \right| &= \left| J(B^{\mathbf{r}})^{-1} F(B^{\mathbf{r}}) - J(\widetilde{A}^{\mathbf{r}})^{-1} F(\widetilde{A}^{\mathbf{r}}) \right| \\ &= \left| \left(J(B^{\mathbf{r}})^{-1} - J(\widetilde{A}^{\mathbf{r}})^{-1} \right) F(\widetilde{A}^{\mathbf{r}}) \right| \\ &\le \left\| J(B^{\mathbf{r}})^{-1} - J(\widetilde{A}^{\mathbf{r}})^{-1} \right\| \left| F(\widetilde{A}^{\mathbf{r}}) \right| \\ &= O\left((\lambda_2 - \lambda_1)^2 \right) = O\left(|\mathbf{r} - \alpha|^2 \right). \end{aligned}$$

Equation (37) now follows by applying Lemma 4.3 to get $\sigma(B^{\mathbf{r}}, \mathbf{r}) = \sigma(B, \mathbf{r})$. ∎

5. Examples.

5.1. Elements of A are linear in parameters. If A is linear in each parameter, we have $A = \Lambda I + xA_x + yA_y = \Lambda I + \alpha I + \beta\sigma_1 + \gamma\sigma_3$, where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

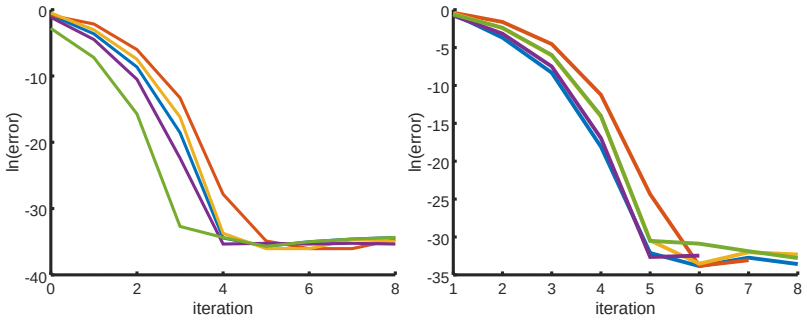

FIG. 3. (a) *Logarithm of the distance from the ith iteration* $\mathbf{r}_i$ *to the conical point* $(\frac{\pi}{3}, \frac{\pi}{3})$ *of* $A(x, y)$ *from* (38)*, plotted as a function of* i*; the algorithm saturates at the limit of numerical precision in* 3–5 *steps.* (b) *Logarithm of* $|\mathbf{r}_{i+1} - \mathbf{r}_i|$ *where* $\mathbf{r}_i$ *is the ith iteration of the algorithm applied to* $A(x, y, z)$ *given by* (39)*. Several independent runs are plotted, each beginning at a random point in* $[-\pi, \pi]$*.*

for some α, β that depend on x, y, and A. The eigenvalues of this matrix are values of λ where

$$\det(A - \lambda I) = \det(\Lambda I + \alpha I + \beta\sigma_1 + \gamma\sigma_3 - \lambda I) = 0,$$

$$(\Lambda + \alpha - \lambda)^2 = \beta^2 + \gamma^2,$$

$$\lambda = \Lambda + \alpha \pm \sqrt{\beta^2 + \gamma^2},$$

which is a cone in the new parameter space. In fact, a simple calculation shows that the degeneracy of the function

$$\hat{A}(\alpha, \beta) = \begin{pmatrix} \beta & \gamma \\ \gamma & -\beta \end{pmatrix},$$

which has the same eigenvectors and shifted eigenvalues, can be located using a single step of iteration (4).

5.2. Nonlinear examples. Consider the following matrix-function example:

$$A(x, y) = \begin{pmatrix} 2\cos(x) & 0 & 0 & 1 \\ 0 & 0.5 + \cos(y) & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{38}$$

Since $A(x, y)$ is a rank-one perturbation of a diagonal matrix, it can be shown that there is a double eigenvalue 1 at the point given by

$$2\cos(x) = 0.5 + \cos(y) = 1,$$

or $x = y = \pi/3$. The result of running the algorithm of Theorem 1.1 with random starting points in the rectangle $(\frac{\pi}{3}, \frac{\pi}{3}) \pm \frac{1}{2}$ is shown in Figure 3a.

The complex Hermitian case described in section 2.3 is demonstrated in Figure

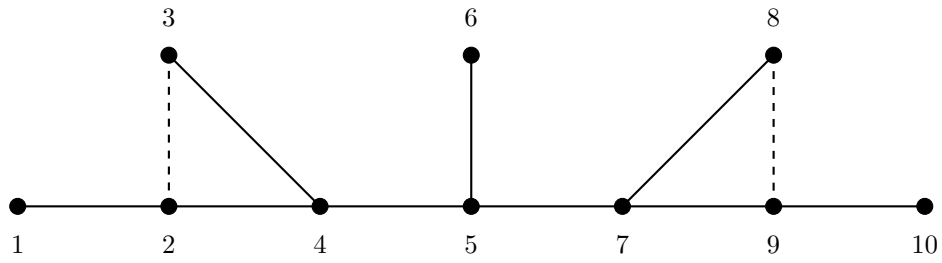

FIG. 4. *Graph corresponding to* (39).

3b. The matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & z \\ 1 & 3 & e^{ix} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-ix} & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & e^{iy} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & e^{-iy} & 3 & 1 \\ z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{39}$$

corresponds to the discrete Laplacian of the graph shown in Figure 4 with dashed edges carrying a magnetic potential (x and y correspondingly). The parameter z is introduced artificially, and the conical point found numerically has value $z = 0$. Since the location of the conical point is not known analytically, the error is estimated using the norms of updates $\|\mathbf{r}_i - \mathbf{r}_{i+1}\|$ instead of $\|\mathbf{r}_i - \alpha\|$. The result of several runs of the algorithm is shown in Figure 3b.

5.3. Avoided crossing. While a nondegenerate conical point is stable under small perturbations of the real symmetric matrix-function $A(x, y)$, the eigenvalue multiplicity may be lifted by an addition of a small complex perturbation. It is instructive to investigate the run results of our algorithm in this case.

Consider the matrix-function

$$A = \begin{pmatrix} x + 3\sin(y) & y + \epsilon i \\ y - \epsilon i & -x - x^2 \end{pmatrix}. \tag{40}$$

It has a conical point at $(0, 0)$ when $\epsilon = 0$ and no eigenvalue multiplicities when $\epsilon \neq 0$. We plot in Figure 5 the results of several runs with $\epsilon = 0$ (left) and with $\epsilon = 10^{-4}$ (right). For $\epsilon = 0$ the algorithm converges quadratically, as in the previous examples. For $\epsilon \neq 0$, the algorithm initially approaches the position of the former conical point, but gets repelled, resulting in oscillations. Conversely, such oscillations (within the limits of numerical precision) should be considered a tell-tale sign of eigenvalue surfaces nearly but not exactly touching.

We remark that for $\epsilon \neq 0$, the square eigenvalue difference $(\lambda_1 - \lambda_2)^2$ has the minimal value of order ϵ^2. If one is using optimization of $(\lambda_1 - \lambda_2)^2$ to find the multiplicity location, it would be difficult to tell apart genuine points of multiplicity from avoided crossings. This observation is investigated further in the next example.

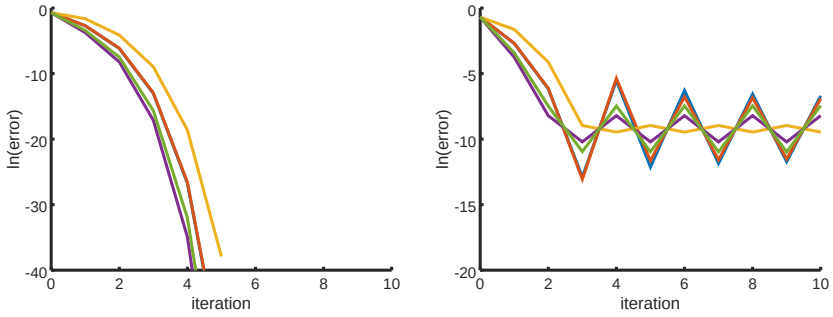

FIG. 5. *Logarithm of distance to* $(0,0)$ *as a function of the iteration step for several runs of the algorithm for* $A(x,y)$ *given by* (40) *with* $\epsilon = 0$ *(left) and with* $\epsilon = 10^{-4}$ *(right), i.e., an avoided crossing. Note the difference in vertical scales. Runs are initialized with random points on the circle of radius* $1/2$ *around* $(0,0)$.

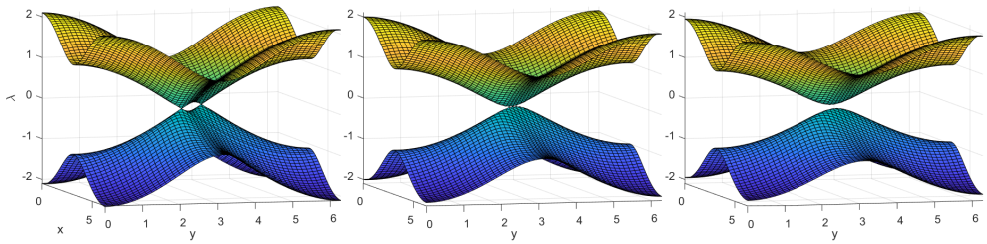

FIG. 6. *Two Dirac points (left) which are colliding (center) and disappearing (right) in the dispersion relation of* (41) *with parameter* p *values* 0.6, 0.5, *and* 0.45 *correspondingly.*

5.4. Merging Dirac points. In condensed matter physics literature, the conical points in the dispersion relation of a periodic structure are know as the "Dirac points," because the effective equation of the wave propagation at the corresponding energy is of Dirac type (see [12] for a mathematical formulation of this physics result). When the material parameters change, the Dirac point may undergo a fold bifurcation, where two points collide and annihilate. The physical consequences of this collision were studied, for example, in [22]; an experimental observation in a tunable honeycomb lattice was reported in [28]. In this section we use the basic model from [22],

$$A(x,y) := \begin{pmatrix} 0 & -1 - \frac{1}{2}e^{ix} - pe^{iy} \\ -1 - \frac{1}{2}e^{-ix} - pe^{-iy} & 0 \end{pmatrix}, \tag{41}$$

where the bifurcation occurs at $p = \frac{1}{2}$: for $p > \frac{1}{2}$ there are two Dirac points and for $p < \frac{1}{2}$ there are none; see Figure 6.

Despite $A(x,y)$ being a complex matrix, the problem of locating Dirac points in this setting is analogous to the real symmetric case due to the presence of the inversion symmetry $\overline{A(-x,-y)} = A(x,y)$. The correct target function F (cf. (8) and (21)) is

$$F(A^{\mathbf{p}}) := \begin{pmatrix} A^{\mathbf{p}}_{11} - A^{\mathbf{p}}_{22} \\ 2\,\mathrm{Im}(A^{\mathbf{p}}_{12}) \end{pmatrix}.$$

In Figure 7 we present a comparison between the convergence of iterations of Theorem 1.1 and a standard quasi-Newton search for the minimum of $g(x,y) = (\lambda_1 - \lambda_2)^2$.

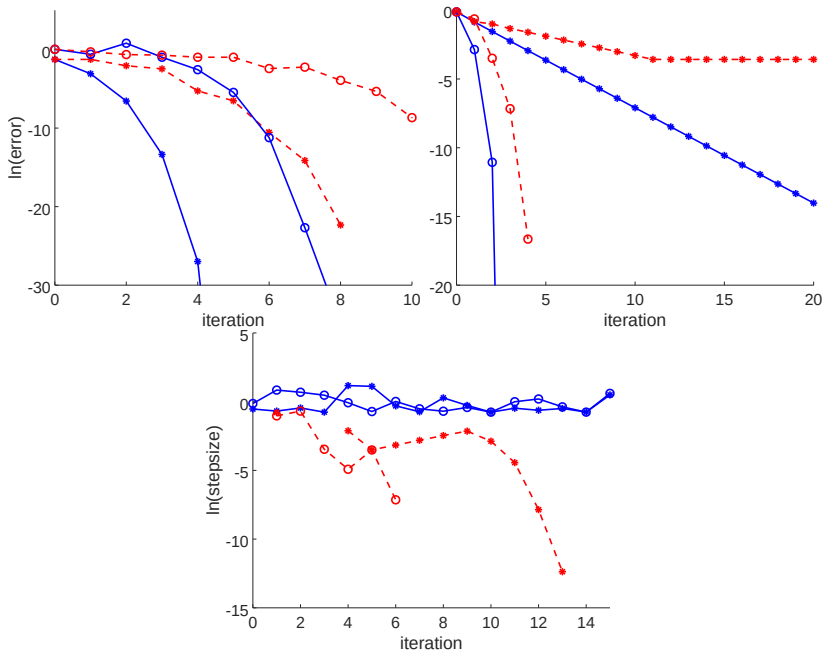

FIG. 7. *Convergence of iterations for matrix family* (41)*: Applying Theorem* 1.1 *(blue, solid) versus quasi-Newton minimization of* $(\lambda_1 - \lambda_2)^2$ *(red, dashed). The parameter* p *is* 0.6 *(left),* 0.5 *(center), and* 0.45 *(right). Two starting points used in each figure,* $(0.8\pi, 0.8\pi)$ *(empty circles) and* $(0.8\pi, 1.2\pi)$ *(stars). (Figure in color online.)*

Figure 7(left) is for $p = 0.6$, where the convergence of both methods is quadratic, although Theorem 1.1 is faster. Figure 7(center) is for $p = 0.5$, where the multiplicity point is *degenerate*. While Theorem 1.1 is no longer applicable, the iteration still converges when the matrix pseudoinverse is used in (4). The speed of iteration is highly dependent on the direction, presumably because the cross-section of the eigenvalue surface is parabolic in one direction and conical in the other. Again, the algorithm of Theorem 1.1 converges faster, while quasi-Newton iteration fails altogether for the second initial point.

Finally, in Figure 7(right), the Y-axis shows the logarithm of the last taken step, since the distance to the conical point is undefined: there is no conical point. While the quasi-Newton iteration converges, correctly, to the minimum of $(\lambda_1 - \lambda_2)^2$ located at (π, π), the algorithm of Theorem 1.1 is not converging, indicating the absence of the conical point in that area.

To interpret the results, recall that a quasi-Newton minimization is searching for the zero of the gradient of g using a numerical approximation of the Hessian of g. But according to Theorem 3.2, the matrix appearing in (4) is *equal* to the leading term of the Hessian (or its square root) around the conical point. It is, therefore, natural to expect a faster convergence.

To give an analogy, consider finding the root of $f(x) = x^2 - a$ via the Newton–Raphson scheme (thus computing f' as done in Theorem 1.1) or via minimization of $g(x) = f^2(x)$ (thus computing g'' in the course of finding the root of g'). Of course, close to the root, $g'' \approx (f')^2$, so the two schemes give equivalent rates of convergence, but having an analytical expression for f' naturally produces better results than

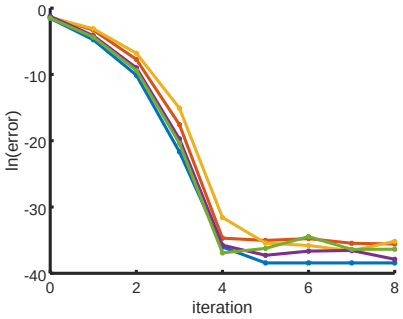

FIG. 8. *Logarithm of distance to* $\mathbf{0}$ *as a function of the iteration step for several runs of the algorithm for* $A(x, y, z, u, v)$ *given by* (43). *Several independent runs are plotted, each beginning at a random point in* $[-0.2, 0.2]^5$.

performing a numerical approximation of g''.

Theorem 1.1 would thus be beneficial in any situation where computing two eigenvectors is not significantly more expensive than sampling the eigenvalues several times.[4] One example of such circumstances is given by differential operators on metric graphs [3], where the eigenvalues are found by solving the "secular equation" of the form $\det\left(I - S(\sqrt{\lambda})\right) = 0$, and once an eigenvalue is identified, the corresponding eigenvector of $S(\sqrt{\lambda})$ gives the (Fourier coefficients of the) eigenvector on the graph. The latter operation is inexpensive relative to repeated evaluation of the determinant necessary for locating the root λ.

5.5. Locating points of higher multiplicity. We can apply a modification of the method to search for points of higher multiplicity in a family of matrices with a sufficient number of parameters. For example, for locating a triple eigenvalue of a five-parameter family A we use

$$F(A^{\mathbf{p}}) = \begin{pmatrix} A^{\mathbf{p}}_{11} - A^{\mathbf{p}}_{22} \\ A^{\mathbf{p}}_{22} - A^{\mathbf{p}}_{33} \\ 2A^{\mathbf{p}}_{12} \\ 2A^{\mathbf{p}}_{13} \\ 2A^{\mathbf{p}}_{23} \end{pmatrix}, \tag{42}$$

where $A^{\mathbf{p}}$ is the function $A(\cdot)$ expressed in the eigenbasis calculated at point $\mathbf{p}$; the first three eigenvectors are assumed to correspond to the consecutive eigenvalues whose point of coalescing we are seeking. As before, $J_{\mathbf{r}}(A^{\mathbf{p}}) = \mathrm{D}_{\mathbf{r}} F(A^{\mathbf{p}})$, and a point α of triple multiplicity is *nondegenerate* if $\det J_\alpha(A^\alpha) \neq 0$.

To demonstrate the performance of our method in locating a triple multiplicity, we consider the function

$$A = \begin{pmatrix} 1 + v + w + x - 3y - z & 2x + y + 2z & x + xz + y \\ 2x + y + 2z & 1 + x + yz & 2v - w + z \\ x + xz + y & 2v - w + z & 1 + vw \end{pmatrix} \tag{43}$$

with triple eigenvalue at $(0, 0, 0, 0, 0)$. The results of several runs are shown in Figure 8; the convergence is clearly quadratic until the limit of numerical precision is reached in about four steps.

[4] In the quasi-Newton experiment above, the eigenvalues were computed five times per iteration step in order to estimate the Hessian.

Acknowledgments. Numerous illuminating discussions with Igor Zelenko are gratefully acknowledged. The authors are particularly grateful to the referees for their deep reading of the manuscript and several suggestions which resulted in significant improvement of this presentation.

REFERENCES

[1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Grad. Texts Math. 60, Springer-Verlag, New York, 1978, translated from the Russian by K. Vogtmann and A. Weinstein.
[2] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Holt, Rinehart, and Winston, New York, London, 1976.
[3] G. Berkolaiko, *An elementary introduction to quantum graphs*, in Geometric and Computational Spectral Theory, Contemp. Math. 700, A. Girouard, D. Jakobson, M. Levitin, N. Nigam, I. Polterovich, and F. Rochon, eds., AMS, 2017, pp. 41–72.
[4] M. V. Berry, *Quantal phase factors accompanying adiabatic changes*, Proc. Roy. Soc. London Ser. A, 392 (1984), pp. 45–57.
[5] M. Born and V. Fock, *Beweis des Adiabatensatzes*, Z. Physik, 51 (1928), pp. 165–180.
[6] A. Castro Neto, F. Guinea, N. Peres, K. Novoselov, and A. Geim, *The electronic properties of graphene*, Rev. Mod. Phys., 81 (2009), pp. 109–162.
[7] M. Dana and K. D. Ikramov, *On the codimension of the variety of symmetric matrices with multiple eigenvalues*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 323 (2005), Chisl. Metody i Vopr. Organ. Vychisl., 18 (2005), pp. 34–46, 224.
[8] L. Dieci, A. Papini, and A. Pugliese, *Approximating coalescing points for eigenvalues of Hermitian matrices of three parameters*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 519–541, https://doi.org/10.1137/120898036.
[9] L. Dieci and A. Pugliese, *Singular values of two-parameter matrices: An algorithm to accurately find their intersections*, Math. Comput. Simulation, 79 (2008), pp. 1255–1269.
[10] L. Dieci and A. Pugliese, *Two-parameter SVD: Coalescing singular values and periodicity*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 375–403, https://doi.org/10.1137/07067982X.
[11] L. Dieci and A. Pugliese, *Hermitian matrices depending on three parameters: Coalescing eigenvalues*, Linear Algebra Appl., 436 (2012), pp. 4120–4142.
[12] C. L. Fefferman and M. I. Weinstein, *Wave packets in honeycomb structures and two-dimensional Dirac equations*, Comm. Math. Phys., 326 (2014), pp. 251–286.
[13] H. Gingold, *A method of global blockdiagonalization for matrix-valued functions*, SIAM J. Math. Anal., 9 (1978), pp. 1076–1082, https://doi.org/10.1137/0509086.
[14] G. Herzberg and H. C. Longuet-Higgins, *Intersection of potential energy surfaces in poluatomic molecules*, Discuss. Faraday Soc., 35 (1963), pp. 77–82.
[15] P. D. Hislop and I. M. Sigal, *Introduction to Spectral Theory. With applications to Schrödinger operators*, Appl. Math. Sci. 113, Springer-Verlag, New York, 1996.
[16] P.-F. Hsieh and Y. Sibuya, *A global analysis of matrices of functions of several variables*, J. Math. Anal. Appl., 14 (1966), pp. 332–340.
[17] N. V. Ilyushechkin, *The discriminant of the characteristic polynomial of a normal matrix*, Mat. Zametki, 51 (1992), pp.16–23.
[18] T. Kato, *Perturbation Theory for Linear Operators*, Classics Math., Springer-Verlag, Berlin, 1995, reprint of the 1980 edition.
[19] P. Kuchment, *Floquet Theory for Partial Differential Equations*, Oper. Theory Adv. Appl. 60, Birkhäuser Verlag, Basel, 1993.
[20] P. Kuchment, *An overview of periodic elliptic operators*, Bull. Amer. Math. Soc. (N.S.), 53 (2016), pp. 343–414.
[21] P. D. Lax, *On the discriminant of real symmetric matrices*, Comm. Pure Appl. Math., 51 (1998), pp. 1387–1396.
[22] G. Montambaux, F. Piéchon, J.-N. Fuchs, and M. O. Goerbig, *Merging of Dirac points in a two-dimensional crystal*, Phys. Rev. B, 80 (2009), 153412.
[23] K. Novoselov, *Nobel lecture: Graphene: Materials in the flatland*, Rev. Mod. Phys., 83 (2011), pp. 837–849.
[24] K. A. O'Neil, *Critical points of the singular value decomposition*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 459–473, https://doi.org/10.1137/040611719.
[25] B. N. Parlett, *The (matrix) discriminant as a determinant*, Linear Algebra Appl., 355 (2002), pp. 85–101.
[26] A. Pugliese, *Theoretical and Numerical Aspects of Coalescing of Eigenvalues and Singular Values of Parameter Dependent Matrices*, Ph.D. thesis, Georgia Institute of Technology,

Atlanta, GA, 2008.
[27] B. Simon, *Holonomy, the quantum adiabatic theorem, and Berry's phase*, Phys. Rev. Lett., 51 (1983), pp. 2167–2170.
[28] L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, and T. Esslinger, *Creating, moving and merging Dirac points with a Fermi gas in a tunable honeycomb lattice*, Nature, 483 (2012), pp. 302–305.
[29] J. von Neumann and E. Wigner, *Uber merkwürdige diskrete Eigenwerte. Uber das Verhalten von Eigenwerten bei adiabatischen Prozessen*, Phys. Z., 30 (1929), pp. 467–470.