SIAM Review
Vol. 62, No. 2, pp. 463–482

First-Order Perturbation Theory for Eigenvalues and Eigenvectors*

Anne Greenbaum†
Ren-Cang Li‡
Michael L. Overton§

Dedicated to Peter Lancaster and G.W. (Pete) Stewart, Masters of Analytic Perturbation Theory and Numerical Linear Algebra on the Occasion of Their 90th and 80th Birthdays

Abstract. We present first-order perturbation analysis of a simple eigenvalue and the corresponding right and left eigenvectors of a general square matrix, not assumed to be Hermitian or normal. The eigenvalue result is well known to a broad scientific community. The treatment of eigenvectors is more complicated, with a perturbation theory that is not so well known outside a community of specialists. We give two different proofs of the main eigenvector perturbation theorem. The first, a block-diagonalization technique inspired by the numerical linear algebra research community and based on the implicit function theorem, has apparently not appeared in the literature in this form. The second, based on complex function theory and on eigenprojectors, as is standard in analytic perturbation theory, is a simplified version of well-known results in the literature. The second derivation uses a convenient normalization of the right and left eigenvectors defined in terms of the associated eigenprojector, but although this dates back to the 1950s, it is rarely discussed in the literature. We then show how the eigenvector perturbation theory is easily extended to handle other normalizations that are often used in practice. We also explain how to verify the perturbation results computationally. We conclude with some remarks about difficulties introduced by multiple eigenvalues and give references to work on perturbation of invariant subspaces corresponding to multiple or clustered eigenvalues. Throughout the paper we give extensive bibliographic commentary and references for further reading.

Key words. analytic perturbation theory, numerical linear algebra, eigenvalues, eigenvectors

AMS subject classifications. 47A55, 65F15

DOI. 10.1137/19M124784X

1. Introduction.

Eigenvalue perturbation theory is an old topic dating originally to the work of Rayleigh in the 19th century. Broadly speaking, there are two main streams of research. The most classical is analytic perturbation theory (APT), where

*Received by the editors March 4, 2019; accepted for publication (in revised form) June 11, 2019; published electronically May 7, 2020.
https://doi.org/10.1137/19M124784X
Funding: The work of the second author was partially supported by National Science Foundation grants CCF-1527104 and DMS-1719620. The work of the third author was partially supported by National Science Foundation grant DMS-1620083.

†Department of Applied Mathematics, University of Washington, Seattle, WA 98195 (greenbau@uw.edu).
‡Department of Mathematics, University of Texas, Arlington, TX 76019 (rcli@uta.edu).
§Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 (overton@cs.nyu.edu).

one considers the behavior of eigenvalues of a matrix or linear operator that is an analytic function of one or more parameters. Authors of well-known books describing this body of work include Kato [Kat66, Kat76, Kat82, Kat95],[1] Rellich [Rel69], Chatelin [Cha11], Baumgärtel [Bau85], and, in textbook form, Lancaster and Tismenetsky [LT85]. Kato [Kat82, p. XII]) and Baumgärtel [Bau85, p. 21] explain that it was Rellich who first established in the 1930s that when a Hermitian matrix or self-adjoint linear operator with an isolated eigenvalue λ of multiplicity m is subjected to a real analytic perturbation, that is, a convergent power series in a real parameter κ, then (1) it has exactly m eigenvalues converging to λ as $\kappa \to 0$, (2) these eigenvalues can also be expanded in convergent power series in κ, and (3) the corresponding eigenvectors can be chosen to be mutually orthogonal and may also be written as convergent power series. As Kato notes, these results are exactly what were anticipated by Rayleigh, Schrödinger, and others, but to prove them is by no means trivial, even in the finite-dimensional case.

The second stream of research is largely due to the numerical linear algebra (NLA) community. It is mostly restricted to matrices and generally concerns perturbation *bounds* rather than expansions, describing how to bound the change in the eigenvalues and associated eigenvectors or invariant subspaces when a given matrix is subjected to a perturbation with a given norm and structure. Here there is a wide variety of well-known results due to many of the founders of matrix analysis and NLA: Gerschgorin, Hoffman and Wielandt, Mirksy, Lidskii, Ostrowski, Bauer and Fike, Henrici, Davis and Kahan, Varah, Ruhe, Stewart, Elsner, Demmel, and others. These are discussed in many books, of which the most comprehensive include those by Wilkinson [Wil65], Stewart and Sun [SS90], Bhatia [Bha97, Bha07], and Stewart [Ste01], as well as Chatelin [Cha12], which actually covers both the APT and the NLA streams of research in some detail. See also the survey by Li [Li14]. An important branch of the NLA stream concerns the *pseudospectra* of a matrix; see the book by Trefethen and Embree [TE05] and the Pseudospectra Gateway website [ET].

This paper is inspired by both the APT and the NLA streams of research, and its scope is limited to an important special case: first-order perturbation analysis of a simple eigenvalue and the corresponding right and left eigenvectors of a general square matrix, not assumed to be Hermitian or normal. The eigenvalue result is well known to a broad scientific community. The treatment of eigenvectors is more complicated, with a perturbation theory that is not so well known outside a community of specialists. We give two different proofs of the main eigenvector perturbation theorem. The first, inspired by the NLA research stream and based on the implicit function theorem, has apparently not appeared in the literature in this form. The second, based on complex function theory and on eigenprojectors, as is standard in APT, is largely a simplified version of results in the literature that are well known. The second derivation uses a convenient normalization of the right and left eigenvectors that depends on the perturbation parameter, but although this dates back to the 1950s, it is rarely discussed in the literature. We then show how the eigenvector perturbation theory is easily extended to handle other normalizations that are often used in

[1] The first edition of Kato's masterpiece *Perturbation Theory for Linear Operators* was published in 1966 and a revised second edition appeared in 1976. The most recent edition is the 1995 reprinting of the second edition with minor corrections. Most of this book is concerned with linear operators, but the first two chapters treat the finite-dimensional case of matrices, and these appeared as a standalone short version in 1982. Since we are only concerned with matrices in this article, our references to Kato's book are to the 1982 edition, although in any case the equation numbering is consistent across all editions.

practice. We also explain how to verify the perturbation results computationally. In the final section, we illustrate the difficulties introduced by multiple eigenvalues with two illuminating examples and give references to work on perturbation of invariant subspaces corresponding to multiple or clustered eigenvalues.

2. First-Order Perturbation Theory for a Simple Eigenvalue. Throughout the paper we use $\|\cdot\|$ to denote the vector or matrix 2-norm, I_n to denote the identity matrix of order n, the superscript T to denote transpose, and $*$ to denote complex conjugate transpose. Greek lowercase letters denote complex scalars. Latin lowercase letters denote complex vectors, with the exception of i for the imaginary unit and j, k, ℓ, m, n for integers. Uppercase letters denote complex matrices or, in some cases, sets in the complex plane. We begin with an assumption that also serves to establish our notation.

Assumption 1. Let $A_0 \in \mathbb{C}^{n\times n}$ have a simple eigenvalue λ_0 corresponding to right eigenvector x_0 (so $Ax_0 = \lambda_0 x_0$ with $x_0 \neq 0$) and left eigenvector y_0 (so $y_0^* A = \lambda_0 y_0^*$ with $y_0 \neq 0$), normalized so that $y_0^* x_0 = 1$. Let $\tau_0 \in \mathbb{C}$ and let $A(\tau)$ be a complex-valued matrix function of a complex parameter τ that is analytic in a neighborhood of τ_0, satisfying $A(\tau_0) = A_0$.

Remark 1. The normalization $y_0^* x_0 = 1$ is always possible since the right and left eigenvectors corresponding to a simple eigenvalue cannot be orthogonal. Note that since x_0 and y_0 are unique only up to scalings, we may multiply x_0 by any nonzero complex scalar ω provided we also scale y_0 by the reciprocal of the conjugate of ω so that $y_0^* x_0$ remains equal to one. The use of the complex conjugate transpose in y_0^* instead of an ordinary transpose is purely a convention that is often, but not universally, followed. The statement that the matrix $A(\tau)$ is analytic means that each entry of $A(\tau)$ is analytic (equivalently, complex differentiable or holomorphic) in τ in a neighborhood of τ_0.

The most basic result in eigenvalue perturbation theory follows.

THEOREM 1 (eigenvalue perturbation theorem). *Under Assumption* 1*, $A(\tau)$ has a unique eigenvalue $\lambda(\tau)$ that is analytic in a neighborhood of τ_0, with $\lambda(\tau_0) = \lambda_0$ and with*

$$\lambda'(\tau_0) = y_0^* A'(\tau_0) x_0, \tag{1}$$

where $\lambda'(\tau_0)$ and $A'(\tau_0)$ are, respectively, the derivatives of $\lambda(\tau)$ and $A(\tau)$ at $\tau = \tau_0$.

The proof appears in the next section.

The quantity

$$\chi = \|x_0\|\|y_0\| \geq |y_0^* x_0| = 1,$$

introduced by [Wil65], is called the *eigenvalue condition number* for λ_0. We have $|\lambda'(\tau_0)| \leq \chi \|A'(\tau_0)\|$. In the real case, χ is the reciprocal of the cosine of the angle between x_0 and y_0. In the special case that A_0 is Hermitian, its right and left eigenvectors coincide so that $\chi = 1$, but in this article we are concerned with general square matrices.

In the APT research stream, instead of eigenvectors, the focus is mostly on the eigenprojector[2] corresponding to λ_0, which can be defined as

$$\Pi_0 = x_0 y_0^* \tag{2}$$

[2]In APT, the standard term is "eigenprojection," while in NLA, "spectral projector" is often used. The somewhat nonstandard term "eigenprojector" is a compromise.

and which satisfies

$$A_0\Pi_0 = \lambda_0\Pi_0 = \Pi_0 A_0 \text{ and } \Pi_0^2 = \Pi_0.$$

Note that the eigenprojector does not depend on the normalization used for the eigenvectors x_0 and y_0 (assuming $y_0^* x_0 = 1$), which simplifies the associated perturbation theory, and note also that $\chi = \|\Pi_0\|$. Let tr denote trace and recall the property $\mathrm{tr}(XY) = \mathrm{tr}(YX)$. Clearly, equation (1) is equivalent to

$$\lambda'(\tau_0) = \mathrm{tr}\left(\Pi_0 A'(\tau_0)\right). \tag{3}$$

Kato [Kat82, p. XIII] explains that the results of Rellich for analytic perturbations of self-adjoint linear operators were extended (by Sz-Nagy, Kato, and others) to non-self-adjoint linear operators and therefore non-Hermitian matrices in the early 1950s using complex function theory, so (3), equivalently (1), was known at that time. However, it seems that these results did not become well known until the publication of the first edition of Kato's book in 1966 (although Kato did present a summary of these results for the linear case at a conference on matrix computations [Giv58, p. 104] in 1958). Equation (1) was independently obtained for the analytic case by Lancaster [Lan64], and for the linear case by Wilkinson [Wil65, pp. 68–69] and Lidskii [Lid66]. They all used the theory of algebraic functions to obtain their results, exploiting the property that eigenvalues are roots of the characteristic polynomial. A different technique is used by Stewart and Sun [SS90, p. 185], who show that the eigenvalue is differentiable w.r.t. its matrix argument using a proof depending on Gerschgorin circles; the result for a differentiable family $A(\tau)$ then follows from the ordinary chain rule.

We close this section with a brief discussion of multiple eigenvalues. The algebraic multiplicity of λ_0 is the multiplicity of the factor $\lambda - \lambda_0$ in the characteristic polynomial $\det(A_0 - \lambda I_n)$, while the geometric multiplicity (which is always less than or equal to the algebraic multiplicity) is the number of associated linearly independent right (equivalently, left) eigenvectors. A simple eigenvalue has both algebraic and geometric multiplicity equal to one. More generally, if the algebraic and geometric multiplicity are equal, the eigenvalue is said to be semisimple or nondefective. An eigenvalue whose geometric multiplicity is one is called nonderogatory.

3. First-Order Perturbation Theory for an Eigenvector Corresponding to a Simple Eigenvalue. We begin this section with a basic result from linear algebra; see [Ste01, Theorem 1.18 and eq. (3.10)] for a proof.

LEMMA 1. *Suppose Assumption* 1 *holds. There exist matrices* $X_1 \in \mathbb{C}^{n\times(n-1)}$, $Y_1 \in \mathbb{C}^{n\times(n-1)}$, *and* $B_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ *satisfying*

$$X = [x_0, X_1], \quad Y = [y_0, Y_1], \quad Y^*X = I_n, \quad \textit{and} \quad Y^*A_0X = \begin{bmatrix} \lambda_0 & 0 \\ 0 & B_1 \end{bmatrix}. \tag{4}$$

Note that, from $Y^*X = I_n$, it is immediate that the columns of X_1 and Y_1, respectively, span the null spaces of y_0^* and x_0^*. Furthermore, we have

$$I_n = XY^* = x_0y_0^* + X_1Y_1^*.$$

We also have $A_0X_1 = X_1B_1$ and $Y_1^*A_0 = B_1Y_1^*$, so the columns of X_1 and Y_1 are, respectively, bases for right and left $(n-1)$-dimensional invariant subspaces of A_0, and $I_n = \Pi_0 + \Pi_1$, where $\Pi_1 = X_1Y_1^*$ is the complementary projector to Π_0. If we assume that A_0 is diagonalizable, i.e., with n linearly independent eigenvectors, then

we can take the columns of X_1 and of Y_1 to be, respectively, right and left eigenvectors corresponding to the eigenvalues of A_0 that differ from λ_0, which we may denote by $\lambda_1, \ldots, \lambda_{n-1}$ (some of which could coincide, as diagonalizability implies only that the eigenvalues are semisimple, not that they are simple). In this case, we can take B_1 to be the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_{n-1})$. More generally, however, X_1 and Y_1 may be *any* matrices satisfying (4), ignoring the multiplicities and Jordan structure of the other eigenvalues.

Now let

$$S = X_1\big(B_1 - \lambda_0 I_{n-1}\big)^{-1} Y_1^*. \tag{5}$$

It then follows that

$$S\Pi_0 = \Pi_0 S = 0 \quad \text{and} \quad (A_0 - \lambda_0 I_n)S = S(A_0 - \lambda_0 I_n) = \Pi_1.$$

In the NLA stream of research, S is called the *group inverse* of $A_0 - \lambda_0 I_n$ [SS90, pp. 240–241], [MS88], [CM91], [GO11, Theorem 5.2.]. In the APT research stream, it is called the *reduced resolvent matrix* of A_0 w.r.t. the eigenvalue λ_0 (see [Kat82, eqs. I.5.28 and II.2.11]).[3]

We now give a first-order perturbation theorem for right and left eigenvectors corresponding to a simple eigenvalue.

THEOREM 2 (eigenvector perturbation theorem). *Suppose Assumption* 1 *holds and define* X_1, Y_1, *and* B_1 *as in Lemma* 1 *and* S *as in* (5). *Then there exist vector-valued functions* $x(\tau)$ *and* $y(\tau)^*$ *that are analytic in a neighborhood of* τ_0 *with* $x(\tau_0) = x_0$, $y(\tau_0) = y_0$, *and* $y(\tau)^* x(\tau) = 1$, *satisfying the right and left eigenvector equations*

$$A(\tau)x(\tau) = \lambda(\tau)x(\tau), \quad y(\tau)^* A(\tau) = \lambda(\tau)y(\tau)^*, \tag{6}$$

where $\lambda(\tau)$ *is the analytic function from Theorem* 1. *Furthermore, these can be chosen so that their derivatives,* $x'(\tau)$ *and* $(y^*)'(\tau)$, *satisfy* $y_0^* x'(\tau_0) = 0$ *and* $(y^*)'(\tau_0)x_0 = 0$, *with*[4]

$$x'(\tau_0) = -SA'(\tau_0)x_0, \tag{7}$$

$$(y^*)'(\tau_0) = -y_0^* A'(\tau_0)S. \tag{8}$$

Note that it is $y(\tau)^*$, not $y(\tau)$, that is analytic with respect to the complex parameter τ. However, $y(\tau)$ is differentiable w.r.t. the real and imaginary parts of τ. Note also that we do not claim that $x(\tau)$ and $y(\tau)$ are unique, even when they are chosen to satisfy (7) and (8). Sometimes, other normalizations of the eigenvectors, not necessarily satisfying (7) and (8), are preferred, as we shall discuss in section 3.4.

It follows from Theorem 2 that

$$\frac{\|x'(\tau_0)\|}{\|x_0\|} \le \kappa(X)\|(\lambda_0 I_{n-1} - B_1)^{-1}\|\|A'(\tau_0)\|,$$

where $\kappa(X) = \kappa(Y) = \|X\|\|Y\|$, the ordinary matrix condition number of X, equivalently of Y (as $Y^* = X^{-1}$), with the same bound also holding for $\|(y^*)'(\tau_0)\|/\|y_0\|$.

[3]The notion of group inverse or reduced resolvent extends beyond the simple eigenvalue context to multiple eigenvalues. If λ_0 is a defective eigenvalue with a nontrivial Jordan structure, the reduced resolvent matrix of A with respect to λ_0 must take account of "eigennilpotents." It is the same as the Drazin inverse of $A_0 - \lambda_0 I_n$, a generalization of the group inverse (see [Cha12, p. 98], [CM91], and, for a method to compute the Drazin inverse, [GOS15]).

[4]We use the notation $(y^*)'(\tau_0)$ to mean $\frac{d}{d\tau}(y(\tau)^*)|_{\tau=\tau_0}$.

In the diagonalizable case, as already noted above, we can take $B_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n-1})$, so

$$\frac{\|x'(\tau_0)\|}{\|x_0\|} \leq \frac{\kappa(X)\|A'(\tau_0)\|}{\min_{j=1,\ldots,n-1}\{|\lambda_0 - \lambda_j|\}},$$

with the same bound also holding for $\|(y^*)'(\tau_0)\|/\|y_0\|$. In this case, the formula (7) for the eigenvector derivative was given by Wilkinson [Wil65, p. 70]. He remarked (p. 109) that although his derivation is essentially classical perturbation theory, a simple but rigorous treatment did not seem to be readily available in the literature. Lancaster [Lan64] and Lidskii [Lid66] both showed that the perturbed eigenvector corresponding to a simple eigenvalue λ_0 may be defined to be differentiable at λ_0, but they did not give the first-order perturbation term. The books by Stewart and Sun [SS90, section V.2] and Stewart [Ste01, sections 1.3 and 4.2] contain excellent discussions of the issues summarized above as well as many additional related results. The eigenvector derivative formula (7) in Theorem 2 above is succinctly stated just below [Ste01, eq. (3.14), p. 46], where on the same page Theorem 3.11 stating it more rigorously, and providing additional bounds, is also given; see also [Ste01, line 4, p. 48]. The reader is referred to [Ste71] and [Ste73] for a proof. Stewart [Ste71] introduced the idea of establishing the existence of a solution to an algebraic Riccati equation by a fixed point iteration, a technique that was followed up in [Ste73, eq. (1.5), p. 730] and [Dem86, eq. (7.2), p. 187]. Alternatively, proofs of Theorem 2 may be derived by various approaches based on the implicit function theorem; see [Mag85, Sun85] and [Sun02, section 2.1]. A related argument appears in [Lax07, Theorem 9.8]. These approaches generally focus on obtaining results for the right eigenvector subject to some normalization; they can also be applied to obtain results for the left eigenvector, and these can be normalized further to obtain the condition $y(\tau)^* x(\tau) = 1$. The proof that we give in section 3.2 is also based on the implicit function theorem, using a block-diagonalization approach that obtains the perturbation results for the right and left eigenvectors simultaneously, ensuring that $y(\tau)^* x(\tau) = 1$. Note, however, that a fundamental difficulty with eigenvectors is their lack of uniqueness. In contrast, the eigenprojector is uniquely defined and satisfies the following perturbation theorem.

THEOREM 3 (eigenprojector perturbation theorem). *Suppose Assumption* 1 *holds and define* X_1, Y_1, *and* B_1 *as in Lemma* 1 *and* S *as in* (5). *Then there exists a matrix-valued function* $\Pi(\tau)$ *that is analytic in a neighborhood of* τ_0 *with* $\Pi(\tau_0) = \Pi_0$, *satisfying the eigenprojector equations*

$$A(\tau)\Pi(\tau) = \lambda(\tau)\Pi(\tau) = \Pi(\tau)A(\tau) \quad \textit{and} \quad \Pi(\tau)^2 = \Pi(\tau), \tag{9}$$

and with derivative given by

$$\Pi'(\tau_0) = -\Pi_0 A'(\tau_0) S - S A'(\tau_0)\Pi_0. \tag{10}$$

This result is well known in the APT research stream [Kat82, eq. (II.2.13)], and, like the eigenvalue perturbation result, goes back to the 1950s. Furthermore, while it's easy to see how Theorem 3 can be proved using Theorem 2, it is also the case that Theorem 2 can be proved using Theorem 3 by defining the eigenvectors appropriately in terms of the eigenprojector, as discussed below. This provides a convenient way to define eigenvectors uniquely.

We note that Theorems 1, 2, and 3 simplify significantly when $A(\tau)$ is a Hermitian function of a real parameter τ, because then the right and left eigenvectors coincide. The results for the Hermitian case lead naturally to perturbation theory for singular

values and singular vectors of a general rectangular matrix; see [Ste01, section 3.3.1] and [Sun02, section 3.1].

3.1. Nonrigorous Derivation of the Formulas in Theorems 1, 2, and 3. If we *assume* that for τ in some neighborhood of τ_0, the matrix $A(\tau)$ has an eigenvalue $\lambda(\tau)$ and corresponding right and left eigenvectors $x(\tau)$ and $y(\tau)$ with $y(\tau)^*x(\tau) = 1$ such that $\lambda(\tau)$, $x(\tau)$ and $y(\tau)^*$ are all analytic functions of τ satisfying $\lambda(\tau_0) = \lambda_0$, $x(\tau_0) = x_0$, and $y(\tau_0) = y_0$, then differentiating the equation $A(\tau)x(\tau) = \lambda(\tau)x(\tau)$ and setting $\tau = \tau_0$, we find

$$A'(\tau_0)x_0 + A_0x'(\tau_0) = \lambda'(\tau_0)x_0 + \lambda_0x'(\tau_0). \tag{11}$$

Multiplying on the left by y_0^* and using $y_0^*A_0 = \lambda_0 y_0^*$ and $y_0^*x_0 = 1$, we obtain the formula (1) for $\lambda'(\tau_0)$:

$$\lambda'(\tau_0) = y_0^*A'(\tau_0)x_0.$$

Equation (11) can be written in the form

$$(A'(\tau_0) - \lambda'(\tau_0)I_n)x_0 = -(A_0 - \lambda_0 I_n)x'(\tau_0). \tag{12}$$

Using Lemma 1, we can write

$$A_0 - \lambda_0 I_n = X \begin{bmatrix} 0 & 0 \\ 0 & B_1 - \lambda_0 I_{n-1} \end{bmatrix} Y^*,$$

and substituting this into (12) and multiplying on the left by Y^*, we find

$$Y^*(A'(\tau_0) - \lambda'(\tau_0)I_n)x_0 = -\begin{bmatrix} 0 & 0 \\ 0 & B_1 - \lambda_0 I_{n-1} \end{bmatrix} Y^*x'(\tau_0).$$

The first row equation here is $y_0^*(A'(\tau_0) - \lambda'(\tau_0)I_n)x_0 = 0$, which is simply the formula for $\lambda'(\tau_0)$. The remaining $n-1$ equations are

$$Y_1^*(A'(\tau_0) - \lambda'(\tau_0)I_n)x_0 = -(B_1 - \lambda_0 I_{n-1})Y_1^*x'(\tau_0),$$

and since $Y_1^*x_0 = 0$ and $B_1 - \lambda_0 I_{n-1}$ is invertible, we obtain the following formula for $Y_1^*x'(\tau_0)$:

$$Y_1^*x'(\tau_0) = -(B_1 - \lambda_0 I_{n-1})^{-1}Y_1^*A'(\tau_0)x_0.$$

Note that $x'(\tau_0)$ is not completely determined by this formula because each eigenvector $x(\tau)$ is determined only up to a multiplicative constant. *If* we can choose the scale factor in such a way that $y_0^*x'(\tau_0) = 0$, then, multiplying on the left by X_1 and recalling that $X_1Y_1^* = I_n - x_0y_0^*$, we obtain the formula in (7) for $x'(\tau_0)$:

$$x'(\tau_0) = -X_1(B_1 - \lambda_0 I_{n-1})^{-1}Y_1^*A'(\tau_0)x_0 = -SA'(\tau_0)x_0.$$

Similarly, the formula (8) for $(y^*)'(\tau_0)$ can be derived assuming that we can choose $y(\tau)$ so that $(y^*)'(\tau_0)x_0 = 0$.

Once formulas (7) and (8) are established, formula (10) follows immediately from

$$[x(\tau)y(\tau)^*]' = x'(\tau)y(\tau)^* + x(\tau)\left[y(\tau)^*\right]'.$$

Evaluating at $\tau = \tau_0$ and using formulas (7) and (8) for $x'(\tau_0)$ and $y'(\tau_0)$, we obtain formula (10) for $\Pi'(\tau_0)$.

In the following subsections, we establish the *assumptions* used here when λ_0 is a simple eigenvalue of A_0, and thus obtain *proofs* of Theorems 1, 2, and 3, in two different ways. The first involves finding equations that a similarity transformation must satisfy if it is to take $A(\tau)$ (or, more specifically, $Y^*A(\tau)X$) to a block-diagonal form like that in Lemma 1 for A_0. The implicit function theorem[5] is then invoked to show that these equations have a unique solution for τ in some neighborhood of τ_0, and that the solution is analytic in τ. The second proof uses the argument principle and the residue theorem from complex analysis to establish that, for τ in a neighborhood of τ_0, each matrix $A(\tau)$ has a simple eigenvalue $\lambda(\tau)$ that is analytic in τ and satisfies $\lambda(\tau_0) = \lambda_0$. It then follows from Lemma 1 that there is a similarity transformation taking $A(\tau)$ to block-diagonal form, but Lemma 1 says nothing about analyticity or even continuity of the associated matrices $X(\tau)$ and $Y(\tau)^*$. Instead, the similarity transformation is applied to the *resolvent* and integrated to obtain an expression for the eigenprojector $\Pi(\tau)$ that is shown to be analytic in τ. Finally, left and right eigenvectors satisfying the analyticity conditions along with the derivative formulas (7) and (8) are defined in terms of the eigenprojector.

Note that the assumptions used here do *not* generally hold when λ_0 is not a simple eigenvalue of A_0, as discussed in section 4.

3.2. First Proof of Theorems 1, 2, and 3, Using Techniques from the NLA Research Stream. The first proof that we give is inspired by the NLA research stream, but instead of Stewart's fixed point iteration technique mentioned previously, we rely on the implicit function theorem [Kra01, Theorem 1.4.11], which we now state in the form that we need.

THEOREM 4 (implicit function theorem). *Let $D \subset \mathbb{C} \times \mathbb{C}^\ell$ be an open set, $h = (h_1, \ldots, h_\ell) : D \to \mathbb{C}^\ell$ an analytic mapping, and $(\tau_0, z^0) \in D$ a point where $h(\tau_0, z^0) = 0$ and where the Jacobian matrix $\left(\frac{\partial h_j}{\partial z_k}\right)_{j,k=1}^{\ell}$ is nonsingular. Then the system of equations $h(\tau, z) = 0$ has a unique analytic solution $z = z(\tau)$ in a neighborhood of τ_0 that satisfies $z(\tau_0) = z^0$.*

We now exploit this result in our proof of Theorem 2. The setting of the stage before applying the implicit function theorem follows Demmel's variant of Stewart's derivation mentioned above. We obtain a proof of Theorem 1 along the way, and then give a proof of Theorem 3 as an easy consequence.

Using Lemma 1, define

$$\begin{bmatrix} \gamma_{11}(\tau) & c_{12}^*(\tau) \\ c_{21}(\tau) & C_{22}(\tau) \end{bmatrix} = Y^*A(\tau)X = \begin{bmatrix} \lambda_0 & 0 \\ 0 & B_1 \end{bmatrix} + Y^*(A(\tau) - A_0)X. \tag{13}$$

Here the scalar γ_{11}, the row and column vectors c_{12}^* and c_{21}, and the $(n-1) \times (n-1)$ matrix C_{22} are analytic functions of τ near τ_0, since $A(\tau)$ is. In what follows, we will transform this matrix into a block-diagonal matrix by a similarity transformation. We will choose $p(\tau)$, $q(\tau)$ so that

$$P(\tau) = \begin{bmatrix} 1 & -q(\tau)^* \\ p(\tau) & I_{n-1} \end{bmatrix}, \quad Q(\tau) = \begin{bmatrix} 1 & -p(\tau)^* \\ q(\tau) & I_{n-1} \end{bmatrix},$$

[5]Since the perturbation parameter τ and the matrix family $A(\tau)$ are complex, we need a version of the implicit function theorem from complex analysis, but in the special case that τ and $A(\tau)$ are real, we could use a more familiar version from real analysis. In that case, although some of the eigenvalues and eigenvectors of a real matrix may be complex, they occur in complex conjugate pairs and are easily represented using real quantities.

$$D(\tau) = Q(\tau)^* P(\tau) = \begin{bmatrix} 1 + q(\tau)^* p(\tau) & 0 \\ 0 & I_{n-1} + p(\tau) q(\tau)^* \end{bmatrix},$$

with $p(\tau)$ and $q(\tau)^*$, and consequently $P(\tau)$, $Q(\tau)^*$, and $D(\tau)$, analytic in a neighborhood of τ_0, with $p(\tau_0) = q(\tau_0) = 0$, and hence $P(\tau_0) = Q(\tau_0) = D(\tau_0) = I_n$. This transformation idea traces back to [Ste73, p. 730] who designed a P with $p = q$, but in the form given here it is due to [Dem86, p. 187].

We would like to have, for τ sufficiently close to τ_0, the similarity transformation

$$Q(\tau)^* Y^* A(\tau) X P(\tau) D(\tau)^{-1} = \begin{bmatrix} \lambda(\tau) & 0 \\ 0 & B(\tau) \end{bmatrix}, \tag{14}$$

where $\lambda(\tau)$ and $B(\tau)$ are also analytic, with $\lambda(\tau_0) = \lambda_0$ and $B(\tau_0) = B_1$. Since $D(\tau)$ is block diagonal by definition, we need $Q(\tau)^* Y^* A(\tau) X P(\tau)$ to be block diagonal. Suppressing the dependence on τ, this last matrix is given by

$$\begin{bmatrix} \gamma_{11} + q^* c_{21} + c_{12}^* p + q^* C_{22} p & -q^*(\gamma_{11} I_{n-1} - C_{22}) + c_{12}^* - q^* c_{21} q^* \\ -(\gamma_{11} I_{n-1} - C_{22}) p + c_{21} - p c_{12}^* p & C_{22} - p c_{12}^* - c_{21} q^* + p \gamma_{11} q^* \end{bmatrix}. \tag{15}$$

For clarity, we introduce the notation $w(\tau)$ for the analytic row vector function $q(\tau)^*$. We then seek column and row vector analytic functions $p(\tau)$ and $w(\tau)$ making the off-diagonal blocks of (15) zero, i.e., satisfying

$$f(\tau, p(\tau)) := -\left[\gamma_{11}(\tau) I_{n-1} - C_{22}(\tau)\right] p(\tau) + c_{21}(\tau) - p(\tau) c_{12}(\tau)^* p(\tau) = 0, \tag{16}$$

$$g(\tau, w(\tau)) := -w(\tau)\left[\gamma_{11}(\tau) I_{n-1} - C_{22}(\tau)\right] + c_{12}(\tau)^* - w(\tau) c_{21}(\tau) w(\tau) = 0. \tag{17}$$

Taking $\ell = n - 1$, $z = 0$, and h equal to first f and then g with s equal to p and w, respectively, in Theorem 4, we note that since $f(\tau_0, 0) = 0$, $g(\tau_0, 0) = 0$, and the Jacobian matrices

$$\left.\frac{\partial f(\tau, p)}{\partial p}\right|_{(\tau,p)=(\tau_0,0)} = -(\lambda_0 I_{n-1} - B_1), \quad \left.\frac{\partial g(\tau, w)}{\partial w}\right|_{(\tau,w)=(\tau_0,0)} = -(\lambda_0 I_{n-1} - B_1)$$

are nonsingular, there are unique functions p and w, analytic in a neighborhood of τ_0, satisfying (16) and (17) with $p(\tau_0) = 0$, $w(\tau_0) = 0$, and

$$p'(\tau_0) = (\lambda_0 I_{n-1} - B_1)^{-1} c_{21}'(\tau_0), \quad w'(\tau_0) = (c_{12}^*)'(\tau_0)(\lambda_0 I_{n-1} - B_1)^{-1}, \tag{18}$$

where, using the definition (13), we have

$$c_{21}'(\tau_0) = Y_1^* A'(\tau_0) x_0 \quad \text{and} \quad (c_{12}^*)'(\tau_0) = y_0^* A'(\tau_0) X_1. \tag{19}$$

Thus, (15) is block diagonal and hence (14) holds, with $\lambda(\tau)$ an eigenvalue of $A(\tau)$ and with

$$\lambda(\tau) = [\gamma_{11} + w c_{21} + c_{12}^* p + w C_{22} p](1 + wp)^{-1},$$
$$B(\tau) = [C_{22} - p c_{12}^* - c_{21} w + p \gamma_{11} w](I_{n-1} + pw)^{-1},$$

again suppressing dependence on τ on the right-hand sides. These functions are analytic in a neighborhood of τ_0 satisfying $\lambda(\tau_0) = \lambda_0$ and $B(\tau_0) = B_1$, with

$$\lambda'(\tau_0) = \gamma_{11}'(\tau_0) = y_0^* A'(\tau_0) x_0,$$

proving Theorem 1.

Let e_1 be the first column of the identity matrix I_n. Multiplying (14) on the left by $XQ(\tau)^{-*}$ and on the right by e_1 we obtain, using $Y^*X = I$ and $D(\tau) = Q(\tau)^*P(\tau)$,

$$A(\tau)XP(\tau)D(\tau)^{-1}e_1 = \lambda(\tau)XP(\tau)D(\tau)^{-1}e_1,$$

so

$$x(\tau) := XP(\tau)D(\tau)^{-1}e_1 = \frac{1}{1 + q(\tau)^*p(\tau)}\,[x_0 + X_1p(\tau)]$$

is analytic with $x(\tau_0) = x_0$ and satisfies the desired right eigenvector equation in (6). Likewise, multiplying (14) on the right by $D(\tau)P(\tau)^{-1}Y^*$ and on the left by e_1^{T} gives

$$e_1^{\mathrm{T}}Q(\tau)^*Y^*A(\tau) = \lambda(\tau)e_1^{\mathrm{T}}Q(\tau)^*Y^*,$$

so

$$y(\tau)^* := e_1^{\mathrm{T}}Q(\tau)^*Y^* = y_0^* + q(\tau)^*Y_1^*$$

is analytic with $y(\tau_0) = y_0$ and satisfies the left eigenvector equation in (6). Furthermore, $y(\tau)^*x(\tau) = 1$, as claimed. Finally, differentiating $x(\tau)$ and $y(\tau)^*$ we have

$$x'(\tau_0) = X_1p'(\tau_0) \quad \text{and} \quad (y^*)'(\tau_0) = (q^*)'(\tau_0)Y_1^*,$$

so combining these with (18) and (19), recalling that $w(\tau) = q(\tau)^*$, and using the definition of S in (5), we obtain the eigenvector derivative formulas (7) and (8). The properties $y_0^*x'(\tau_0) = 0$ and $(y^*)'(\tau_0)x_0 = 0$ follow, so Theorem 2 is proved.

Finally, define

$$\Pi(\tau) = x(\tau)y(\tau)^*.$$

The eigenprojector equations (9) follow immediately. We have

$$\Pi'(\tau_0) = x'(\tau_0)y(\tau_0)^* + x(\tau_0)(y^*)'(\tau_0),$$

so Theorem 3 follows from (7) and (8). ∎

3.3. Second Proof of Theorems 1, 2, and 3, Using Techniques from the APT Research Stream. In this proof, in contrast to the previous one, we focus on proving Theorem 3 first, obtaining the proof of Theorem 1 along the way, and finally obtaining Theorem 2 as a consequence. This proof of Theorem 3 is based on complex function theory, as is standard in APT. However, our derivation is simpler than most given in the literature, which usually prove more general results such as giving complete analytic expansions for the eigenvalue and eigenprojector, while we are concerned only with the first-order term. The key to the last part of the proof, yielding Theorem 2, is to use an appropriate eigenvector normalization.

The main tool here is the residue theorem [MH06, pp. 293–294, Theorems 8.1 and 8.2].

THEOREM 5 (residue theorem). *Let D be a simply connected domain in $\mathbb{C}$ and let Γ be a simple closed positively oriented contour that lies in D. If f is analytic inside Γ and on Γ, except at the points $\zeta_1, \ldots, \zeta_m$ that lie inside Γ, then*

$$\int_\Gamma f(\zeta)\,d\zeta = 2\pi i \sum_{\ell=1}^{m} \mathrm{Res}[f, \zeta_\ell],$$

where if f has a simple pole at ζ_ℓ, then

$$\mathrm{Res}[f, \zeta_\ell] = \lim_{\zeta \to \zeta_\ell} (\zeta - \zeta_\ell) f(\zeta),$$

and if f has a pole of order k at ζ_ℓ, then

$$\mathrm{Res}[f,\zeta_\ell] = \frac{1}{(k-1)!}\lim_{\zeta\to\zeta_\ell}\frac{d^{k-1}}{d\zeta^{k-1}}\left[(\zeta-\zeta_\ell)^k f(\zeta)\right].$$

Let Γ be the boundary of an open set Δ in the complex plane containing the simple eigenvalue λ_0, with no other eigenvalues of A_0 in $\Delta\cup\Gamma$. First note that since $\det(A_0-\lambda I_n)$, the characteristic polynomial p_0 of A_0, does not vanish on Γ, the same will hold for all polynomials with coefficients sufficiently close to those of p_0; in particular, it will hold for $\det(A(\tau)-\lambda I_n)$, the characteristic polynomial p_τ of $A(\tau)$, if τ is sufficiently close to τ_0; say, $|\tau-\tau_0|\le\epsilon$. In what follows, we always assume that $|\tau-\tau_0|\le\epsilon$. By the argument principle [MH06, p. 328, Theorem 8.8], the number of zeros of p_τ inside Δ is

$$\frac{1}{2\pi i}\int_\Gamma \frac{\frac{d}{d\zeta}p_\tau(\zeta)}{p_\tau(\zeta)}\,d\zeta.$$

For $\tau=\tau_0$, this value is 1. Since, for each $\zeta\in\Gamma$, the integrand $\frac{d}{d\zeta}p_\tau(\zeta)/p_\tau(\zeta)$ is a continuous function of τ, the integral above is continuous as well. Since it is integer-valued, it must be the constant 1. So, let $\lambda(\tau)$ denote the unique root of p_τ in the region Δ, i.e., the unique eigenvalue of $A(\tau)$ in Δ. Note that this means that $p_\tau(\zeta)$ can be written in the form $(\zeta-\lambda(\tau))q(\zeta)$, where q has no roots in $\Delta\cup\Gamma$. It therefore follows from the residue theorem that

$$\frac{1}{2\pi i}\int_\Gamma \frac{\frac{d}{d\zeta}p_\tau(\zeta)}{p_\tau(\zeta)}\zeta\,d\zeta = \lambda(\tau), \tag{20}$$

since

$$\mathrm{Res}\left[\frac{\frac{d}{d\zeta}(p_\tau(\zeta))\,\zeta}{p_\tau(\zeta)},\lambda(\tau)\right] = \lim_{\zeta\to\lambda(\tau)}\frac{(\zeta-\lambda(\tau))\left[(\zeta-\lambda(\tau))\frac{d}{d\zeta}q(\zeta)+q(\zeta)\right]\zeta}{((\zeta-\lambda(\tau))\,q(\zeta)} = \lambda(\tau).$$

Since the left-hand side of (20) is an analytic function of τ, the right-hand side is as well. Thus $A(\tau)$ has a unique eigenvalue $\lambda(\tau)$ in Δ and $\lambda(\tau)$ is an analytic function of τ.

For ζ not an eigenvalue of $A(\tau)$, define the *resolvent* of $A(\tau)$ by

$$R(\zeta;A(\tau)) := (A(\tau)-\zeta I_n)^{-1}. \tag{21}$$

Lemma 1 states that there exist left and right eigenvectors $y_0(\tau)$ and $x_0(\tau)$ associated with $\lambda(\tau)$ and satisfying $y_0(\tau)^*x_0(\tau)=1$, along with matrices $X_1(\tau)\in\mathbb{C}^{n\times(n-1)}$, $Y_1(\tau)\in\mathbb{C}^{n\times(n-1)}$, and $B_1(\tau)\in\mathbb{C}^{(n-1)\times(n-1)}$ satisfying

$$X(\tau)=[x_0(\tau),X_1(\tau)],\quad Y(\tau)=[y_0(\tau),Y_1(\tau)],\quad Y(\tau)^*X(\tau)=I_n,\quad\text{and}$$

$$Y(\tau)^*A(\tau)X(\tau)=\begin{bmatrix}\lambda(\tau) & 0\\ 0 & B_1(\tau)\end{bmatrix}.$$

Note that we do not claim that $X(\tau)$ and $Y(\tau)^*$ are analytic or even continuous functions of τ. It follows that the resolvent of $A(\tau)$ satisfies

$$\begin{aligned} R(\zeta;A(\tau)) &= X(\tau)\begin{bmatrix}(\lambda(\tau)-\zeta)^{-1} & 0\\ 0 & (B_1(\tau)-\zeta I_{n-1})^{-1}\end{bmatrix}Y(\tau)^*\\ &= (\lambda(\tau)-\zeta)^{-1}x_0(\tau)y_0(\tau)^* + S(\zeta;\tau), \end{aligned} \tag{22}$$

where $S(\zeta;\tau) = X_1(\tau)(B_1(\tau) - \zeta I_{n-1})^{-1}Y_1(\tau)^*$. Now, $S(\zeta;\tau)$ is a matrix-valued function of ζ with no poles in Δ, so it follows from the residue theorem (applied to the functions associated with each entry of $S(\zeta;\tau)$) that $\int_\Gamma S(\zeta;\tau)\,d\zeta = 0$ and therefore from (22) that

$$\begin{aligned} -\frac{1}{2\pi i}\int_\Gamma R(\zeta;A(\tau))\,d\zeta &= -\frac{1}{2\pi i}\left(\int_\Gamma (\lambda(\tau)-\zeta)^{-1}\,d\zeta\right) x_0(\tau)y_0(\tau)^* \\ &= x_0(\tau)y_0(\tau)^* = \Pi(\tau). \end{aligned} \tag{23}$$

For $\tau = \tau_0$, this is $\Pi(\tau_0) = \Pi_0$.

From the definition of the resolvent (21), it follows that since $A(\tau)$ is an analytic function of τ, the resolvent $R(\zeta;A(\tau))$ is also analytic, provided that ζ is not an eigenvalue of $A(\tau)$. Differentiating the equation

$$(A(\tau) - \zeta I_n)R(\zeta;A(\tau)) = I_n$$

with respect to τ gives

$$\frac{\partial}{\partial\tau}R(\zeta;A(\tau)) = -R(\zeta;A(\tau))A'(\tau)R(\zeta;A(\tau)).$$

Considering expression (23) for $\Pi(\tau)$, it follows that $\Pi(\tau)$ is also analytic and its derivative at $\tau = \tau_0$ is

$$\Pi'(\tau_0) = \frac{1}{2\pi i}\int_\Gamma R(\zeta;A_0)A'(\tau_0)R(\zeta;A_0)\,d\zeta.$$

Using (22) and writing $S_0(\zeta)$ for $S(\zeta;\tau_0)$, we obtain

$$\begin{aligned} \Pi'(\tau_0) &= \frac{1}{2\pi i}\int_\Gamma \left[(\lambda_0-\zeta)^{-1}\Pi_0 + S_0(\zeta)\right] A'(\tau_0)\left[(\lambda_0-\zeta)^{-1}\Pi_0 + S_0(\zeta)\right]\,d\zeta \\ &= \frac{1}{2\pi i}\left(\int_\Gamma (\lambda_0-\zeta)^{-2}\,d\zeta\right)\Pi_0A'(\tau_0)\Pi_0 + \frac{1}{2\pi i}\int_\Gamma S_0(\zeta)A'(\tau_0)S_0(\zeta)\,d\zeta \\ &\quad + \frac{1}{2\pi i}\int_\Gamma (\lambda_0-\zeta)^{-1}\left[\Pi_0A'(\tau_0)S_0(\zeta) + S_0(\zeta)A'(\tau_0)\Pi_0\right]\,d\zeta. \end{aligned}$$

From the residue theorem, the first term is zero since the integrand has a pole of order 2 at λ_0, with $\mathrm{Res}[(\lambda_0-\zeta)^{-2},\lambda_0] = 0$. The second term is also zero because the integrand has no poles inside Γ. Since the integrand of the remaining term has a simple pole at λ_0 with $\mathrm{Res}[(\lambda_0-\zeta)^{-1},\lambda_0] = -1$, we have

$$\Pi'(\tau_0) = -\Pi_0A'(\tau_0)S_0(\lambda_0) - S_0(\lambda_0)A'(\tau_0)\Pi_0, \tag{24}$$

where $S_0(\lambda_0) = S(\lambda_0;\tau_0) = X_1(\tau_0)(B_1(\tau_0) - \lambda_0 I)^{-1}Y_1(\tau_0)$ is the same as S defined in (5). This proves Theorem 3.

Now we define the eigenvectors in terms of the eigenprojector, using a normalization that goes back to [SN51] (see [Bau85, eq. (7.1.12)]):

$$x(\tau) = (y_0^*\Pi(\tau)x_0)^{-1/2}\Pi(\tau)x_0, \quad y(\tau)^* = (y_0^*\Pi(\tau)x_0)^{-1/2}y_0^*\Pi(\tau), \tag{25}$$

where we use the principal branch of the square root function (and assume that ϵ is small enough that $|\tau - \tau_0| \le \epsilon$ implies that the quantities under the square roots,

which are 1 for $\tau = \tau_0$, are bounded away from the origin and the negative real axis).[6] Since $\Pi(\tau)$ is analytic, it follows that $x(\tau)$ and $y(\tau)^*$ are analytic as well. From (25) we have

$$A(\tau)x(\tau) = (y_0^*\Pi(\tau)x_0)^{-1/2}A(\tau)\Pi(\tau)x_0 = (y_0^*\Pi(\tau)x_0)^{-1/2}\lambda(\tau)\Pi(\tau)x_0 = \lambda(\tau)x(\tau)$$

and similarly $y(\tau)^*A(\tau) = \lambda(\tau)y(\tau)^*$, so the eigenvector equations (6) hold as required, and, since $\Pi(\tau)^2 = \Pi(\tau)$, we obtain

$$y(\tau)^*x(\tau) = (y_0^*\Pi(\tau)x_0)^{-1}y_0^*\Pi(\tau)^2x_0 = 1$$

as claimed in Theorem 2.

To obtain the eigenvalue derivative, we differentiate the equation

$$(A(\tau) - \lambda(\tau)I_n)x(\tau) = 0$$

and evaluate it at $\tau = \tau_0$:

$$(A'(\tau_0) - \lambda'(\tau_0)I_n)x_0 + (A_0 - \lambda_0 I_n)x'(\tau_0) = 0.$$

Multiplying by y_0^* on the left, this becomes

$$\lambda'(\tau_0) = y_0^*A'(\tau_0)x_0,$$

proving Theorem 1.

Finally, using (10), we have

$$\begin{aligned} x'(\tau_0) &= -\frac{1}{2}(y_0^*\Pi_0x_0)^{-3/2}(y_0^*\Pi'(\tau_0)x_0)\Pi_0x_0 + [y_0^*\Pi_0x_0]^{-1/2}\Pi'(\tau_0)x_0 \\ &= -\frac{1}{2}(y_0^*[-\Pi_0A'(\tau_0)S - SA'(\tau_0)\Pi_0]x_0)\Pi_0x_0 - [\Pi_0A'(\tau_0)S + SA'(\tau_0)\Pi_0]x_0. \end{aligned}$$

The first three terms are zero since $Sx_0 = -X_1(\lambda_0 I - B_1)^{-1}Y_1^*x_0 = 0$ and likewise $y_0^*S = 0$. So, as $\Pi_0x_0 = x_0$, we have

$$x'(\tau_0) = -SA'(\tau_0)x_0.$$

Similarly, $(y^*)'(\tau_0) = -y_0^*A'(\tau_0)S$. The properties $y_0^*x'(\tau_0) = 0$ and $(y^*)'(\tau_0)x_0 = 0$ follow, so the proof of Theorem 2 is complete. ∎

3.4. Eigenvector Normalizations. Theorem 2 does not state formulas for the analytic eigenvector functions $x(\tau)$ and $y(\tau)^*$, specifying only that they exist, satisfying $y(\tau)^*x(\tau) = 1$, with derivatives given by (7) and (8). Furthermore, the first proof given in section 3.2 does not provide formulas for $x(\tau)$, $y(\tau)$, showing only that they exist via the implicit function theorem. However, the second proof given in section 3.3 *does* provide formulas for $x(\tau)$ and $y(\tau)$ in terms of the eigenprojector $\Pi(\tau)$ (which *is* uniquely defined) and the eigenvectors x_0 and y_0. This formula (25) may be viewed as a *normalization* because it provides a way to define the eigenvectors uniquely and, furthermore, it has the property that $x(\tau)$ and $y(\tau)^*$ are analytic near τ_0 and satisfy $y(\tau)^*x(\tau) = 1$. Let us refer to (25) as "normalization 0."

[6] See also [Kat82, eq. (II.3.24)], which uses a related definition but without the square root, resulting in $y(\tau)^*x_0 = y_0^*x(\tau) = 1$ instead of $y(\tau)^*x(\tau) = 1$.

Although the beautifully simple normalization (25) dates back to the 1950s, it seems to be rarely used. In this subsection we discuss some other normalizations that are more commonly used in practice. Let us denote the resulting normalized eigenvectors by $\hat{x}(\tau)$ and $\hat{y}(\tau)$ and relate them to $x(\tau)$ and $y(\tau)$, as defined in (25), by

$$\hat{x}(\tau) = \alpha(\tau)\, x(\tau), \quad \hat{y}(\tau)^* = \beta(\tau)^*\, y(\tau)^*, \tag{26}$$

where $\alpha(\tau)$ and $\beta(\tau)$ are two nonzero complex-valued scalar functions of τ to be defined below. Here we use the complex conjugate of β in the definition to be consistent with the conjugated left eigenvector notation. The analyticity of $\hat{x}(\tau)$ and $\hat{y}(\tau)^*$ near τ_0 depends on that of $\alpha(\tau)$ and $\beta(\tau)^*$. We consider several possible normalizations, continuing to assume that $y_0^* x_0 = 1$ but not necessarily that $\hat{y}(\tau)^*\hat{x}(\tau) = 1$ for $\tau \neq 0$. In all cases, the formula

$$\hat{x}'(\tau_0) = \alpha(\tau_0)x'(\tau_0) + \alpha'(\tau_0)x_0, \quad (\hat{y}^*)'(\tau_0) = \beta^*(\tau_0)(y^*)'(\tau_0) + (\beta^*)'(\tau_0)y_0^* \tag{27}$$

follows immediately from (26), so to determine the derivatives $\hat{x}'(\tau_0)$ and $(\hat{y}^*)'(\tau_0)$, we need only determine $\alpha(\tau_0)$, $\alpha'(\tau_0)$, $\beta^*(\tau_0)$, and $(\beta^*)'(\tau_0)$, obtaining $x'(\tau_0)$ and $(y^*)'(\tau_0)$ from (7) and (8), as stated in Theorem 2. Note that the derivatives of the normalized eigenvectors, $\hat{x}'(\tau_0)$ and $(\hat{y}^*)'(\tau_0)$, do *not* necessarily satisfy $y_0^*\hat{x}'(\tau_0) = 0$ and $(\hat{y}^*)'(\tau_0)x_0 = 0$, unlike the derivatives $x'(\tau_0)$ and $(y^*)'(\tau_0)$. We now define several different normalizations:

1. $e_1^{\mathrm{T}}\hat{x}(\tau) = e_1^{\mathrm{T}}\hat{y}(\tau) = 1$ (i.e., the first entries of $\hat{x}(\tau)$ and $\hat{y}(\tau)$ are 1). This is possible for τ sufficiently close to τ_0 if $e_1^{\mathrm{T}}x_0 \neq 0$ and $e_1^{\mathrm{T}}y_0 \neq 0$. Suppose this is the case. Then

$$\alpha(\tau) = \frac{1}{e_1^{\mathrm{T}}x(\tau)}, \quad \text{so} \quad \alpha'(\tau_0) = -\frac{e_1^{\mathrm{T}}x'(\tau_0)}{(e_1^{\mathrm{T}}x_0)^2}, \tag{28}$$

and

$$\beta(\tau)^* = \frac{1}{y(\tau)^*e_1}, \quad \text{so} \quad (\beta^*)'(\tau_0) = -\frac{(y^*)'(\tau_0)e_1}{(y_0^*e_1)^2}. \tag{29}$$

Here $\alpha(\tau)$, $\beta(\tau)^*$ and hence $\hat{x}(\tau)$, $\hat{y}^*(\tau)$ are analytic near τ_0.

2. $e_1^{\mathrm{T}}\hat{x}(\tau) = 1$ and $\hat{y}(\tau)^*\hat{x}(\tau) = 1$. This normalization is defined near τ_0 under the assumption that $e_1^{\mathrm{T}}x_0 \neq 0$; no additional assumption is needed since $y_0^*x_0 = 1$. Clearly (28) holds as before. In addition, we have

$$\beta(\tau)^* = \frac{e_1^{\mathrm{T}}x(\tau)}{y(\tau)^*x(\tau)} = e_1^{\mathrm{T}}x(\tau), \quad (\beta^*)'(\tau_0) = e_1^{\mathrm{T}}x'(\tau_0). \tag{30}$$

Again, $\alpha(\tau)$ and $\beta(\tau)^*$ are analytic near τ_0.

3. $\hat{x}(\tau)^{\mathrm{T}}\hat{x}(\tau) = 1$ and $\hat{y}(\tau)^*\hat{x}(\tau) = 1$. This normalization is defined near τ_0 if $x_0^{\mathrm{T}}x_0 \neq 0$, which may not be the case when x_0 is complex. Suppose this does hold. We have

$$\alpha(\tau) = \pm\frac{1}{(x(\tau)^{\mathrm{T}}x(\tau))^{1/2}}, \quad \alpha'(\tau_0) = \mp\frac{x'(\tau_0)^{\mathrm{T}}x_0}{(x_0^{\mathrm{T}}x_0)^{3/2}}, \tag{31}$$

(32)
$$\beta(\tau)^* = \pm\frac{(x(\tau)^{\mathrm{T}}x(\tau))^{1/2}}{y(\tau)^*x(\tau)} = \pm(x(\tau)^{\mathrm{T}}x(\tau))^{1/2}, \quad (\beta^*)'(\tau_0) = \pm\frac{x'(\tau_0)^{\mathrm{T}}x_0}{(x_0^{\mathrm{T}}x_0)^{1/2}}.$$

Either sign may be used, resulting in analytic $\alpha(\tau)$ and $\beta^*(\tau)$ near $\tau = \tau_0$.

4. $\hat{x}(\tau)^*\hat{x}(\tau) = 1$ and $\hat{y}(\tau)^*\hat{x}(\tau) = 1$. In a way this is the most natural choice of normalization, because it is possible without any assumptions on x_0, y_0 beyond $y_0^*x_0 = 1$. The problem, however, is that it does not define the eigenvectors uniquely. Suppose we choose $\alpha(\tau) = 1/\|x(\tau)\|$, $\beta(\tau) = \|x(\tau)\|$. Then α, β^*, $\hat{x}$, and $\hat{y}^*$ are not analytic in τ, but they are differentiable w.r.t. to the real and imaginary parts of τ. However, we could equally well multiply α and β by any unimodular complex number $e^{i\theta}$, so there are infinitely many different choices of α, β that are smooth w.r.t. the real and imaginary parts of τ (though not analytic in τ). A variant on this normalization is known as real-positive (RP) compatibility [GO11]: it requires $y(\tau)^*x(\tau)$ to be real and positive with $\|x(\tau)\| = \|y(\tau)\| = 1$.

More generally we could define the normalization in terms of two functions $\psi(\hat{x}(\tau), \hat{y}(\tau)) = 1$ and $\omega(\hat{x}(\tau), \hat{y}(\tau))) = 1$. Depending on what choice is made, there are several possible outcomes: there could be unique $\alpha(\tau)$ and $\beta(\tau)$ that satisfy the two normalization equations, as in cases 1 and 2 above when $e_1^{\mathrm{T}}x_0 \neq 0$ and $e_1^{\mathrm{T}}y_0 \neq 0$; there could be two choices as in case 3 when $x_0^{\mathrm{T}}x_0 \neq 0$; there could be an infinite number of choices as in case 4; or there could be no $\alpha(\tau)$ and $\beta(\tau)$ that satisfy the normalization equations.

Perturbation theory for normalized eigenvectors, using several of the normalizations given above, was extensively studied by Meyer and Stewart [MS88] and by Bernasconi, Choirat, and Seri [BCS11].

3.5. Computation. Suppose we wish to verify the results of Theorems 1, 2, and 3 computationally—always a good idea! Let us consider how to do this in MATLAB, where eigenvalues as well as right and left eigenvectors can be conveniently computed by the function `eig`. Let Assumption 1 hold, assuming for convenience that $\tau_0 = 0$ and $A(\tau) = A_0 + \tau\Delta A$ for some given matrices A_0 and ΔA with $\|A_0\| = \|\Delta A\| = 1$. Take $|\tau|$ sufficiently small that only one computed eigenvalue $\tilde{\lambda}$ of $\tilde{A} \equiv A(\tau)$ is close to the eigenvalue λ_0 of A_0. Then, assuming the eigenvalue is not too badly conditioned, we can easily verify the eigenvalue perturbation formula (1) in Theorem 1 by computing the finite difference quotient $(\tilde{\lambda} - \lambda_0)/\tau$ and comparing it with (1). Since the exact eigenvalue $\lambda(\tau)$ of $A + \tau\Delta A$ is analytic in τ, it is clear that, mathematically, the difference between the difference quotient and the derivative should be $O(|\tau|)$ as $\tau \to 0$, but numerically, if $|\tau|$ is too small, rounding error dominates the computation instead [Ove01, Chapter 11].

Similarly, we can compute right and left eigenvectors $\tilde{x}$ and $\tilde{y}$ corresponding to $\tilde{\lambda}$, normalize them so that $\tilde{y}^*\tilde{x} = 1$, compute the eigenprojector $\widetilde{\Pi} = \tilde{x}\tilde{y}^*$, and verify that the matrix difference quotient $(\widetilde{\Pi} - \Pi_0)/\tau$ approximates formula (10) for the eigenprojector derivative given by Theorem 3.

What about the eigenvector derivative formulas? Implementing "normalization 0" defined in (25), we can compute normalizations of $\tilde{x}$ and $\tilde{y}$ by

$$\tilde{\tilde{x}} = (y_0^*\widetilde{\Pi}x_0)^{-1/2}\widetilde{\Pi}x_0 = \left(\frac{\tilde{y}^*x_0}{y_0^*\tilde{x}}\right)^{1/2}\tilde{x}, \tag{33}$$

$$\tilde{\tilde{y}} = (x_0^*\widetilde{\Pi}^*y_0)^{-1/2}\widetilde{\Pi}^*y_0 = \left(\frac{\tilde{x}^*y_0}{x_0^*\tilde{y}}\right)^{1/2}\tilde{y}. \tag{34}$$

Note that the formulas on the right-hand side of (33) and (34) avoid computing the eigenprojector, which can be advantageous if n is large. Then we can verify that the

difference quotients $(\tilde{\tilde{x}}-x_0)/\tau$ and $(\tilde{\tilde{y}}^*-y_0^*)/\tau$ approximate the eigenvector derivatives (7) and (8) given by Theorem 2.

Now consider normalizations 1 to 4 as defined in section 3.4. For normalization 1 (respectively, normalization 2), the formulas for the normalized eigenvectors and their derivatives given by (26) and (27) together with (28) and (29) (respectively, (28) and (30)) can be verified easily provided the first components of x_0 and y_0 are not zero. Of course, the index 1 is arbitrary. A better choice is to use $e_j^{\mathrm{T}}\hat{x}(\tau) = e_k^{\mathrm{T}}\hat{y}(\tau) = 1$, where j (respectively, k) is the index of an entry of x_0 (respectively, y_0) with maximum modulus, but this requires access to x_0 and y_0. In the case of normalization 3, the formulas for the eigenvectors and their derivatives given by (26) and (27) together with (31) and (32) can also be verified easily, with a caveat due to the freedom in the choice of sign. Specifically, when $\hat{x}(\tau)$ and $\hat{y}(\tau)$ are obtained from the computed vectors $\tilde{x}$ and $\tilde{y}$, it is important to ensure that the signs of $\hat{x}(\tau)$ and $\hat{x}(\tau_0)$ (and therefore $\hat{y}(\tau)$ and $\hat{y}(\tau_0)$) are consistent; this can be done by choosing the signs of the real parts (or the imaginary parts) of $e_j^{\mathrm{T}}\hat{x}(\tau)$ and $e_j^{\mathrm{T}}\hat{x}(\tau_0)$ to be the same, where j is the index of an entry of $\hat{x}(\tau_0)$ with maximum modulus.

As for normalization 4, although we could arbitrarily choose α and β^* to be smooth functions w.r.t. the real and imaginary parts of τ, there is no way to know how to obtain smoothly varying computed eigenvectors from the unnormalized computed eigenvectors $\tilde{x}$ and $\tilde{y}$.

Summarizing, the formulas for the derivatives of the eigenvalue and the eigenprojector given in Theorems 1 and 3 are easily verified numerically, while for eigenvector normalization 0 given by (25) and normalizations 1, 2, and 3 defined in section 3.4, the formulas for the eigenvector derivatives can also usually be verified computationally. However, perhaps surprisingly, there is no panacea when it comes to the eigenvectors. Normalization 0 always requires access to the eigenvectors x_0 and y_0, while the only way to ensure that normalizations 1 and 2 are well defined is by providing access to x_0 and y_0 so that indices j and k can be used instead of index 1 if necessary. As for normalization 3, it may not be well defined, and even if it is, care must be taken to avoid inconsistent sign choices. Finally, normalization 4 is simply not well defined.

Verification of the eigenvalue, eigenprojector, and eigenvector formulas is illustrated by MATLAB programs in the supplemental materials [local/web 13KB].[7]

4. Multiple Eigenvalues. Theorems 1, 2, and 3 do not generally hold when λ_0 is a multiple eigenvalue. In this section we consider two illuminating examples where multiple eigenvalues enter the picture. Recall that algebraic and geometric multiplicity, along with the terms semisimple (nondefective) and nonderogatory, were defined at the end of section 1. In what follows, we use the principal branch of the square root function for definiteness, but any branch would suffice.

Example 1. Let

$$A(\tau) = \begin{bmatrix} 0 & 1 \\ \tau & 0 \end{bmatrix}$$

with $\tau \to \tau_0 = 0$. The limit matrix $A_0 = A(0)$ is a Jordan block, with 0 a defective, nonderogatory eigenvalue with algebraic multiplicity 2 and geometric multiplicity 1. The corresponding right and left eigenvectors are $x_0 = [1,\ 0]^{\mathrm{T}}$ and $y_0 = [0,\ 1]^{\mathrm{T}}$, but these are mutually orthogonal and cannot be scaled so that $y_0^* x_0 = 1$. For $\tau \neq 0$, $A(\tau)$ has two simple eigenvalues $\lambda_{1,2}(\tau) = \pm\tau^{1/2}$, which are not analytic in

[7] To run these programs, download them from the website and see the main routine `eigValProjVecVaryDemo.m`.

any neighborhood of 0. The corresponding right and left eigenvectors are uniquely defined, up to scalings, by

$$x_{1,2}(\tau) = \begin{bmatrix} 1 \\ \pm\tau^{1/2} \end{bmatrix} \quad \text{and} \quad y_{1,2}(\tau)^* = \begin{bmatrix} \pm\tau^{1/2} & 1 \end{bmatrix}. \tag{35}$$

The right eigenvectors $x_{1,2}(\tau)$ are not analytic near 0, and they *both* converge to the unique right eigenvector x_0 of A_0 as $\tau \to 0$. Likewise, $y_{1,2}(\tau)^*$ are not analytic, and they both converge to y_0^* as $\tau \to 0$. For $\tau \neq 0$, we can scale the eigenvectors so that $y_1(\tau)^* x_1(\tau) = y_2(\tau)^* x_2(\tau) = 1$, but then either $x_j(\tau)$ or $y_j(\tau)$ (or both) must diverge as $\tau \to 0$ for $j = 1, 2$.

This example is easily extended to the $n \times n$ case where $A(\tau)$ is zero except for a single superdiagonal of 1's and a bottom left entry τ, so that A_0 is a single Jordan block with 0 a nonderogatory eigenvalue with algebraic multiplicity n. The eigenvalues of $A(\tau)$ are then the nth roots of unity times $\tau^{1/n}$. In fact, Lidskii [Lid66] gave a remarkable general perturbation theory for eigenvalues of a linear family $A(\tau)$ for which an eigenvalue of A_0 may have *any* algebraic and geometric multiplicities and indeed any Jordan block structure; see also [Bau85, section 7.4] and [MBO97]. These results are not described in Kato's books. However, Kato does treat eigenvalue perturbation in detail in the case that the eigenvalues of A_0 are semisimple [Kat82, section II.2.3]; see also [LMZ03]. Even in this case, the behavior can be unexpectedly complex, as the following example shows.

Example 2 [LT85, p. 394]. Let

$$A(\tau) = \begin{bmatrix} 0 & \tau \\ \tau^2 & 0 \end{bmatrix}$$

with $\tau \to 0$. This time the limit matrix $A_0 = A(0)$ is the zero matrix, with 0 a semisimple eigenvalue with algebraic and geometric multiplicity 2, so we can take *any* vectors in $\mathbb{C}^2$ as right or left eigenvectors of A_0. The eigenvalues of $A(\tau)$ are $\lambda_{1,2}(\tau) = \pm\tau^{3/2}$, which are not analytic. The corresponding right and left eigenvectors are uniquely defined, up to scaling, by the same formulas (35) as in the previous example. So again $x_{1,2}(\tau)$ (respectively, $y_{1,2}(\tau)^*$) are not analytic, and both converge to the same vector x_0 (respectively, y_0^*) as in the previous example when $\tau \to 0$, although in this case A_0 has two linearly independent right (and left) eigenvectors. There is no right eigenvector of $A(\tau)$ that converges to a vector that is linearly independent of x_0 as $\tau \to 0$, and likewise no left eigenvector that converges to a vector linearly independent of y_0.

In the examples above, a multiple eigenvalue splits apart under perturbation. To avoid dealing with this complexity, one may study the average behavior of a cluster of eigenvalues, and the corresponding invariant subspace, under perturbation. There is a large body of work on this topic: see the books by Kato [Kat82, section II.2.1], Gohberg, Lancaster and Rodman [GLR06], Stewart [Ste01, Chapter 4] and Stewart and Sun [SS90, Chapter 5], the unpublished technical report by Sun [Sun02, section 2.3], two surveys on Stewart's many contributions by Ipsen [Ips10] and Demmel [Dem10], and papers such as [BDF08, Dem87, DF01, KK14].

In the case of a Hermitian family $A(\tau)$, the perturbation theory for multiple eigenvalues simplifies greatly; the pioneering results of Rellich were mentioned earlier in section 1. See [LT85, section 11.7] and [GLR85] for more details.

5. Concluding Remarks. In this paper we have presented two detailed yet accessible proofs of first-order perturbation results for a simple eigenvalue of a matrix and its associated right and left eigenvectors and eigenprojector. We hope this will facilitate the dissemination of these important results to a much broader community of researchers and students than has hitherto been the case. We have also tried to convey the breadth and depth of work in the two principal relevant research streams, analytic perturbation theory and numerical linear algebra. There are, of course, many generalizations that we have not even begun to explore in this article. As just one example, nonlinear eigenvalue problems, where one replaces $A - \lambda I_n$ by a matrix function $F(A, \lambda)$ with polynomial or more general nonlinear dependence on λ, arise in many important applications. Perturbation theory for nonlinear eigenvalue problems, representing the APT and NLA communities, respectively, may be found in [ACL93] and [BH13]. Finally, we remark that our bibliography, while fairly extensive, is in no way intended to be comprehensive.

REFERENCES

[ACL93] A. L. Andrew, K.-W. E. Chu, and P. Lancaster, *Derivatives of eigenvalues and eigenvectors of matrix functions*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 903–926, https://doi.org/10.1137/0614061. (Cited on p. 480)

[Bau85] H. Baumgärtel, *Analytic Perturbation Theory for Matrices and Operators*, Oper. Theory Adv. Appl. 15, Birkhäuser Verlag, Basel, 1985. (Cited on pp. 464, 474, 479)

[BCS11] M. Bernasconi, C. Choirat, and R. Seri, *Differentials of eigenvalues and eigenvectors in undamped discrete systems under alternative normalizations*, in Proceedings of the World Congress on Engineering, International Association of Engineers, 2011, pp. 285–287. (Cited on p. 477)

[BDF08] D. Bindel, J. Demmel, and M. Friedman, *Continuation of invariant subspaces in large bifurcation problems*, SIAM J. Sci. Comput., 30 (2008), pp. 637–656, https://doi.org/10.1137/060654219. (Cited on p. 479)

[BH13] D. Bindel and A. Hood, *Localization theorems for nonlinear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1728–1749, https://doi.org/10.1137/130913651. (Cited on p. 480)

[Bha97] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math. 169, Springer-Verlag, New York, 1997, https://doi.org/10.1007/978-1-4612-0653-8. (Cited on p. 464)

[Bha07] R. Bhatia, *Perturbation Bounds for Matrix Eigenvalues*, Classics Appl. Math. 53, SIAM, Philadelphia, 2007, https://doi.org/10.1137/1.9780898719079; reprint of the 1987 original. (Cited on p. 464)

[Cha11] F. Chatelin, *Spectral Approximation of Linear Operators*, Classics Appl. Math. 65, SIAM, Philadelphia, 2011, https://doi.org/10.1137/1.9781611970678; reprint of the 1983 original. (Cited on p. 464)

[Cha12] F. Chatelin, *Eigenvalues of Matrices*, Classics Appl. Math. 71, SIAM, Philadelphia, 2012, https://doi.org/10.1137/1.9781611972467; revised reprint of the 1993 edition. (Cited on pp. 464, 467)

[CM91] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Dover, New York, 1991; corrected reprint of the 1979 original. (Cited on p. 467)

[Dem86] J. W. Demmel, *Computing stable eigendecompositions of matrices*, Linear Algebra Appl., 79 (1986), pp. 163–193, https://doi.org/10.1016/0024-3795(86)90298-3. (Cited on pp. 468, 471)

[Dem87] J. W. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing, 38 (1987), pp. 43–57, https://doi.org/10.1007/BF02253743. (Cited on p. 479)

[Dem10] J. W. Demmel, *The SVD, eigenproblem, and invariant subspaces: Algorithms*, in G. W. Stewart: Selected Works with Commentaries, Contemp. Math., Birkhäuser Verlag, Basel, 2010, pp. 95–104, https://doi.org/10.1007/978-0-8176-4968-5_8. (Cited on p. 479)

[DF01] L. Dieci and M. J. Friedman, *Continuation of invariant subspaces*, Numer. Linear Algebra Appl., 8 (2001), pp. 317–327, https://doi.org/10.1002/nla.245. (Cited on p. 479)

[ET] M. Embree and L. N. Trefethen, *Pseudospectral Gateway*, https://www.cs.ox.ac.uk/pseudospectra/. (Cited on p. 464)

[Giv58] W. Givens, *Conference on matrix computations*, J. ACM, 5 (1958), pp. 100–115, https://doi.org/10.1145/320911.320923. (Cited on p. 466)

[GLR85] I. Gohberg, P. Lancaster, and L. Rodman, *Perturbation of analytic Hermitian matrix functions*, Appl. Anal., 20 (1985), pp. 23–48, https://doi.org/10.1080/00036818508839556. (Cited on p. 479)

[GLR06] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Classics Appl. Math. 51, SIAM, Philadelphia, 2006, https://doi.org/10.1137/1.9780898719093; reprint of the 1986 original. (Cited on p. 479)

[GO11] N. Guglielmi and M. L. Overton, *Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1166–1192, https://doi.org/10.1137/100817048. (Cited on pp. 467, 477)

[GOS15] N. Guglielmi, M. L. Overton, and G. W. Stewart, *An efficient algorithm for computing the generalized null space decomposition*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 38–54, https://doi.org/10.1137/140956737. (Cited on p. 467)

[Ips10] I. C. F. Ipsen, *The eigenproblem and invariant subspaces: Perturbation theory*, in G. W. Stewart: Selected Works with Commentaries, Contemp. Math., Birkhäuser Verlag, Basel, 2010, pp. 71–93, https://doi.org/10.1007/978-0-8176-4968-5_7. (Cited on p. 479)

[Kat66] T. Kato, *Perturbation Theory for Linear Operators*, Grundlehren Math. Wiss. 132, Springer-Verlag, New York, 1966. (Cited on p. 464)

[Kat76] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Grundlehren Math. Wiss. 132, Springer-Verlag, New York, 1976. (Cited on p. 464)

[Kat82] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1982. (Cited on pp. 464, 466, 467, 468, 475, 479)

[Kat95] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995; reprint of the 1980 edition. (Cited on p. 464)

[KK14] M. Karow and D. Kressner, *On a perturbation bound for invariant subspaces of matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 599–618, https://doi.org/10.1137/130912372. (Cited on p. 479)

[Kra01] S. G. Krantz, *Function Theory of Several Complex Variables*, AMS, Providence, RI, 2001, https://doi.org/10.1090/chel/340; reprint of the 1992 edition. (Cited on p. 470)

[Lan64] P. Lancaster, *On eigenvalues of matrices dependent on a parameter*, Numer. Math., 6 (1964), pp. 377–387, https://doi.org/10.1007/BF01386087. (Cited on pp. 466, 468)

[Lax07] P. D. Lax, *Linear Algebra and Its Applications*, 2nd ed., Pure Appl. Math. (Hoboken), Wiley-Interscience, Hoboken, NJ, 2007. (Cited on p. 468)

[Li14] R. Li, *Matrix perturbation theory*, in Handbook of Linear Algebra, 2nd ed., Discrete Math. Appl. (Boca Raton), L. Hogben, ed., CRC Press, Boca Raton, FL, 2014, pp. (21–1)–(21–20). (Cited on p. 464)

[Lid66] V. B. Lidskii, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Math. and Math. Phys., 6 (1966), pp. 73–85. (Cited on pp. 466, 468, 479)

[LMZ03] P. Lancaster, A. S. Markus, and F. Zhou, *Perturbation theory for analytic matrix functions: The semisimple case*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 606–626, https://doi.org/10.1137/S0895479803423792. (Cited on p. 479)

[LT85] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Comput. Sci. Appl. Math., Academic Press, Orlando, FL, 1985. (Cited on pp. 464, 479)

[Mag85] J. Magnus, *On differentiating eigenvalues and eigenvectors*, Econom. Theory, 1 (1985), pp. 179–191. (Cited on p. 468)

[MBO97] J. Moro, J. V. Burke, and M. L. Overton, *On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817, https://doi.org/10.1137/S0895479895294666. (Cited on p. 479)

[MH06] J. H. Mathews and R. W. Howell, *Complex Analysis for Mathematics and Engineering*, Jones and Bartlett, Sudbury, MA, 2006. (Cited on pp. 472, 473)

[MS88] C. D. Meyer and G. W. Stewart, *Derivatives and perturbations of eigenvectors*, SIAM J. Numer. Anal., 25 (1988), pp. 679–691, https://doi.org/10.1137/0725041. (Cited on pp. 467, 477)

[Ove01] M. L. Overton, *Numerical Computing with IEEE Floating Point Arithmetic*, SIAM, Philadelphia, 2001, https://doi.org/10.1137/1.9780898718072. (Cited on p. 477)

[Rel69] F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach Science, New York, London, Paris, 1969. (Cited on p. 464)

[SN51] B. Sz.-Nagy, *Perturbations des transformations linéaires fermées*, Acta Sci. Math. Szeged, 14 (1951), pp. 125–137. (Cited on p. 474)

[SS90] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Computer Science and Scientific Computing, Academic Press, Boston, MA, 1990. (Cited on pp. 464, 466, 467, 468, 479)

[Ste71] G. W. Stewart, *Error bounds for approximate invariant subspaces of closed linear operators*, SIAM J. Numer. Anal., 8 (1971), pp. 796–808, https://doi.org/10.1137/0708073. (Cited on p. 468)

[Ste73] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev., 15 (1973), pp. 727–764, https://doi.org/10.1137/1015095. (Cited on pp. 468, 471)

[Ste01] G. W. Stewart, *Matrix Algorithms. Vol.* II*: Eigensystems*, SIAM, Philadelphia, 2001, https://doi.org/10.1137/1.9780898718058. (Cited on pp. 464, 466, 468, 469, 479)

[Sun85] J. G. Sun, *Eigenvalues and eigenvectors of a matrix dependent on several parameters*, J. Comput. Math., 3 (1985), pp. 351–364. (Cited on p. 468)

[Sun02] J.-G. Sun, *Stability and Accuracy: Perturbation Analysis of Algebraic Eigenproblems*, Tech. Report UMINF 98.07, Umeå University, 1998; revised 2002, https://people.cs.umu.se/jisun/Jiguang-Sun-UMINF98-07-rev2002-02-20.pdf. (Cited on pp. 468, 469, 479)

[TE05] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra. The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005. (Cited on p. 464)

[Wil65] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. (Cited on pp. 464, 465, 466, 468)