SIAM J. Matrix Anal. Appl.
Vol. 40, No. 2, pp. 672–695

SENSITIVITY ANALYSIS OF NONLINEAR EIGENPROBLEMS*

RAFIKUL ALAM† AND SK. SAFIQUE AHMAD‡

Abstract. Let $P : \Omega \subset \mathbb{C} \to \mathbb{C}^{n\times n}$ be given by $P(\lambda) := \sum_{j=0}^{m} A_j\phi_j(\lambda)$, where $\phi_j : \Omega \to \mathbb{C}$ for $j = 0, 1, \ldots, m$ are suitable functions. We present an eigenvector-free framework for the sensitivity analysis of eigenvalues of P. We analyze the Fréchet differentiability of a simple eigenvalue of P as a function of P and derive two equivalent representations of the Fréchet derivative and the gradient of the eigenvalue. Further, we derive three equivalent representations of the condition number $\mathrm{cond}(\lambda, P)$ of a simple eigenvalue λ of P. Specially, we present an eigenvector-free representation of $\mathrm{cond}(\lambda, P)$ which generalizes a result due to Smith [*Numer. Math.*, 10 (1967), pp. 232–240] for a standard eigenvalue problem to the case of a nonlinear eigenvalue problem and provides an alternative viewpoint of the sensitivity of eigenvalues. In the second part, we consider a homogeneous matrix-valued function $H : \mathbb{C}^2 \to \mathbb{C}^{n\times n}$ of the form $H(c, s) := \sum_{j=0}^{m} A_j\psi_j(c, s)$, where $\psi_j : \mathbb{C}^2 \to \mathbb{C}$ for $j = 0, 1, \ldots, m$ are homogeneous functions of degree ℓ. We present a simple and concise eigenvector-free framework for the sensitivity analysis of eigenvalues of H that avoids the apparatus of projective spaces. We analyze Fréchet differentiability of a simple eigenvalue of H as a function of H and derive two equivalent representations of the Fréchet derivative and the gradient of the eigenvalue. Furthermore, we derive three equivalent representations of the condition number $\mathrm{cond}((\lambda, \mu), H)$ of a simple eigenvalue (λ, μ) of H. Our eigenvector-free representation of $\mathrm{cond}((\lambda, \mu), H)$ generalizes Smith's eigenvector-free representation of the condition number of a simple eigenvalue of a matrix to the case of a homogeneous nonlinear eigenproblem.

Key words. sensitivity analysis, eigenvalue, condition number, matrix polynomial, matrix-valued function, singular value, eigenvector

AMS subject classifications. 65F15, 15A12, 15A18, 65F35

DOI. 10.1137/17M1153236

1. Introduction.

Let $P : \Omega \subset \mathbb{C} \to \mathbb{C}^{n\times n}$ be given by $P(\lambda) := \sum_{j=0}^{m} A_j\phi_j(\lambda)$, where $\phi_j : \Omega \to \mathbb{C}$ for $j = 0, 1, \ldots, m$ are suitable functions. We assume that $P(\lambda)$ is regular, that is, $\mathrm{rank}(P(\lambda)) = n$ for some $\lambda \in \Omega$, and consider the eigenvalue problem $P(\lambda)v = 0$. This setting encompasses a wide variety of nonlinear eigenvalue problems such as polynomial eigenproblems, rational eigenproblems, and holomorphic and meromorphic eigenproblems when $\phi_0, \ldots, \phi_m$ are polynomials, rational functions, holomorphic, and meromorphic functions, respectively. We also consider a homogeneous matrix-valued function $H : \mathbb{C}^2 \to \mathbb{C}^{n\times n}$ of the form $H(c, s) := \sum_{j=0}^{m} A_j\psi_j(c, s)$, where $\psi_j : \mathbb{C}^2 \to \mathbb{C}$ for $j = 0, 1, \ldots, m$ are homogeneous functions of degree ℓ. This encompasses homogeneous polynomial eigenproblems of degree m as a special case when $\psi_j(c, s) := c^{m-j}s^j$ for $j = 0, 1, \ldots, m$.

We assume that $P(\lambda)$ is analytic except possibly for poles in Ω. Then $\lambda \in \Omega$ is said to be an eigenvalue of P if $\mathrm{rank}(P(\lambda)) < n$. We denote by $\mathrm{eig}(P)$ the set of eigenvalues of P, that is, $\mathrm{eig}(P) := \{\lambda \in \Omega : \mathrm{rank}(P(\lambda)) < n\}$. We also assume that H is regular and consider the spectrum $\mathrm{eig}(H) := \{(c, s) \in \mathbb{C}^2\setminus\{0\} : \mathrm{rank}(H(c, s)) < n\}$. Note that $(\lambda, \mu) \in \mathrm{eig}(H) \Rightarrow \alpha(\lambda, \mu) \in \mathrm{eig}(H)$ for all $\alpha \neq 0$. Thus an eigenvalue (λ, μ) is uniquely represented by the line $[\lambda : \mu] := \mathrm{span}\{(\lambda, \mu)\}$ in $\mathbb{C}^2$. Let $\mathbb{CP}^1$ be the projective space

*Received by the editors October 23, 2017; accepted for publication (in revised form) by Z. Drmac March 27, 2019; published electronically May 30, 2019.
http://www.siam.org/journals/simax/40-2/M115323.html
†Corresponding author. Department of Mathematics, IIT Guwahati, Guwahati 781039, India (rafik@iitg.ernet.in, rafikul68@gmail.com).
‡School of Basic Sciences, Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Khandwa Road, Indore 453552, India (safique@iiti.ac.in).

of one dimensional subspaces of $\mathbb{C}^2$, that is, $\mathbb{CP}^1 := \{[c : s] : (c, s) \in \mathbb{C}^2 \setminus \{0\}\}$. Then $(\lambda, \mu) \in \text{eig}(H) \Rightarrow [\lambda : \mu] \in \mathbb{CP}^1$. We also refer to $[\lambda : \mu]$ as an eigenvalue of H.

Sensitivity analysis of eigenvalues plays an important role in analyzing the accuracy of computed eigenvalues and has been studied extensively over the years [34, 35, 36, 23, 10, 25, 12, 13, 29, 30, 31]. For example, if λ is a simple eigenvalue of a matrix A with left and right eigenvectors y and x, respectively, then it is shown by Wilkinson [34, 35, 36] that the condition number $\text{cond}(\lambda, A)$ measures the sensitivity of λ and, for the spectral norm, $\text{cond}(\lambda, A)$ is given by

$$\text{cond}(\lambda, A) := \limsup_{\|\Delta A\|_2 \to 0} \frac{\text{dist}(\lambda,\ \text{eig}(A + \Delta A))}{\|\Delta A\|_2} = \frac{\|x\|_2 \|y\|_2}{|y^* x|}, \tag{1.1}$$

where $\text{dist}(\lambda, \text{eig}(A+\Delta A)) := \min\{|\lambda - \mu| : \mu \in \text{eig}(A+\Delta A)\}$. On the other hand, for the spectral norm, Smith [23] derived an eigenvector-free representation of $\text{cond}(\lambda, A)$ given by

$$\text{cond}(\lambda, A) = \frac{\|\text{adj}(A - \lambda I)\|_2}{|\partial_s \det(A - \lambda I)|} = \frac{\|\text{adj}(A - \lambda I)\|_2}{\prod_{\mu \neq \lambda} |\lambda - \mu|}, \tag{1.2}$$

where $\text{adj}(A - \lambda I)$ is the "adjugate" of $A - \lambda I$, $\partial_s \det(A - \lambda I)$ is the derivative of the characteristic polynomial $\det(A - sI)$ with respect to s at λ, and the product $\prod_{\mu \neq \lambda} |\lambda - \mu|$ is taken over all the eigenvalues of A except for λ. The eigenvector-free representation of $\text{cond}(\lambda, A)$ in (1.2) reveals an interesting fact that when λ belongs to a cluster of eigenvalues of A, then λ is expected to be highly ill-conditioned and behave like a multiple eigenvalue. The representation in (1.2) is generalized to the case of arbitrary matrix norm in [1].

Motivated by various applications of nonlinear eigenproblems [2, 7, 8, 15, 17, 18, 19, 20, 24, 26, 32, 33], we analyze the sensitivity of eigenvalues of P and H. We present a unified framework for the sensitivity analysis of homogeneous as well as nonhomogeneous nonlinear eigenproblems from which various condition numbers of simple eigenvalues considered in the literature [34, 23, 25, 10, 4, 11, 12, 31, 14, 13] would follow as special cases. The main contributions of this paper are the following.

First, we present a simple and concise framework for the sensitivity analysis of eigenvalues of $P(\lambda)$ with respect to small perturbations in the coefficient matrices $A_0, \ldots, A_m$. A salient feature of our framework is that our analysis is eigenvector-free and is based on the analysis of the function $(\lambda, P) \longmapsto \det(P(\lambda))$. By invoking the implicit function theorem, we show that a simple eigenvalue $\lambda(P)$ of P evolves as a smooth function of P in a neighborhood of P. We derive two equivalent representations of the Fréchet derivative $\mathrm{D}\lambda(P)$, of which one is eigenvector-free. As a byproduct, we deduce that a simple eigenvalue $\lambda(t)$ of a parameter dependent nonlinear eigenproblem $P(t, \lambda)u = 0$, where $t \in \mathbb{C}^k$ is a parameter and $t \longmapsto P(t, .)$ is holomorphic, evolves holomorphically in a small neighborhood of t and determine the derivative $\mathrm{D}\lambda(t)$ and the gradient $\nabla\lambda(t)$.

Next, we define a norm $\|P\|_V := \|(\|A_0\|, \ldots, \|A_m\|)\|_V$, where $\|\cdot\|_V$ is a monotone norm on $\mathbb{C}^{m+1}$, and consider the condition number $\text{cond}(\lambda, P)$ of $\lambda \in \text{eig}(P)$ given by

$$\text{cond}(\lambda, P) := \limsup_{\|\Delta P\|_V \to 0} \frac{\text{dist}(\lambda,\ \text{eig}(P + \Delta P))}{\|\Delta P\|_V}, \tag{1.3}$$

where $\Delta P(\lambda) := \sum_{j=0}^{m} \Delta A_j \phi_j(\lambda)$. Let λ be a simple eigenvalue of P with left and right eigenvectors u and v, that is, $u^* P(\lambda) = 0$ and $P(\lambda)v = 0$. We show that $\text{cond}(\lambda, P) =$

$\|D\lambda\|$, where $D\lambda$ is the Fréchet derivative of λ at P, and derive three equivalent representations of $\mathrm{cond}(\lambda, P)$, of which two are eigenvector-free. Indeed, by setting $\Phi(\lambda) := [\phi_0(\lambda), \ldots, \phi_m(\lambda)]^\top$, we show that

$$\mathrm{cond}(\lambda, P) = \frac{\|\Phi(\lambda)\|_{V*}\|(\mathrm{adj}(P(\lambda)))^*\|_*}{|\partial_s \det(P(\lambda))|} = \frac{\|\Phi(\lambda)\|_{V*}\|uv^*\|_*}{|u^*\partial_s P(\lambda)v|}, \tag{1.4}$$

where $\mathrm{adj}(P(\lambda))$ is the adjugate of the matrix $P(\lambda)$, $\partial_s \det(P(\lambda))$ is the derivative of the characteristic polynomial $\det(P(s))$ at λ, and $\|\cdot\|_{V*}$ and $\|\cdot\|_*$ are the dual norms of $\|\cdot\|_V$ and the matrix norm $\|\cdot\|$, respectively. Observe that the first equality in (1.4) generalizes (1.2) to the case of nonlinear eigenproblem $P(\lambda)v = 0$.

Further, let $\sigma_1(P(\lambda)) \geq \cdots \geq \sigma_n(P(\lambda))$ be the singular values of the matrix $P(\lambda)$. Then, for the spectral norm on $\mathbb{C}^{n\times n}$, we show that

$$\mathrm{cond}(\lambda, P) = \frac{\|\Phi(\lambda)\|_{V*}\prod_{j=1}^{n-1}\sigma_j(P(\lambda))}{|\partial_s \det(P(\lambda))|} = \frac{\|\Phi(\lambda)\|_{V*}\|u\|_2\|v\|_2}{|u^*\partial_s P(\lambda)v|}.$$

Second, we extend the framework developed for the sensitivity analysis of nonlinear nonhomogeneous eigenproblems to the case of nonlinear homogeneous eigenproblems. Indeed, we present a concise framework for the sensitivity analysis of eigenvalues of $H(c, s)$ with respect to small perturbations in the coefficient matrices $A_0, \ldots, A_m$. The most important feature of our framework is that our analysis is eigenvector-free and is based on the analysis of the function

$$(c, s, H) \longmapsto (\det(H(c, s)),\ (c-\lambda)\bar{\lambda} + (s-\mu)\bar{\mu}), \ \text{ where } (\lambda, \mu) \in \mathrm{eig}(H),$$

which enables us to carry out the sensitivity analysis of (λ, μ) in the inner product space $\mathbb{C}^2$ and avoid the apparatus of projective spaces, especially the calculus on projective spaces. Again, by invoking the implicit function theorem, we show that a simple eigenvalue $(\lambda(H), \mu(H))$ of H evolves as a smooth function of H in a small neighborhood of H. We derive two equivalent representations of the Fréchet derivative $(D\lambda(H), D\mu(H))$, of which one is eigenvector-free. As a byproduct, we deduce that a simple eigenvalue $(\lambda(t), \mu(t))$ of a parameter dependent nonlinear homogenous eigenproblem $H(t, \lambda, \mu)u = 0$, where $t \in \mathbb{C}^k$ is a parameter and $t \longmapsto H(t, .)$ is holomorphic, evolves holomorphically in a small neighborhood of t and determine the derivative $(D\lambda(t), D\mu(t))$ and the gradient $(\nabla\lambda(t), \nabla\mu(t))$.

The homogeneous framework is convenient for dealing with infinite eigenvalues as it treats finite and infinite eigenvalues on equal footing. For $(\lambda, \mu) \in \mathrm{eig}(H)$, we define the condition number $\mathrm{cond}((\lambda, \mu), H)$ by

$$\mathrm{cond}((\lambda, \mu), H) := \limsup_{\|\Delta H\|_V \to 0} \frac{\mathrm{dist}((\lambda, \mu),\ \mathrm{eig}(H + \Delta H))}{\|(\lambda, \mu)\|_2\ \|\Delta H\|_V}, \tag{1.5}$$

where $\mathrm{dist}((\lambda, \mu), \mathrm{eig}(H+\Delta H)) := \min\{\|(\lambda, \mu)-(c, s)\|_2 : (c, s) \in \mathrm{eig}(H+\Delta H)$ and $(c-\lambda, s-\mu) \perp (\lambda, \mu)\}$ and $\Delta H(c, s) := \sum_{j=0}^{m} \Delta A_j \psi_j(c, s)$. Let $(\lambda, \mu) \in \mathrm{eig}(H)$ be simple with left and right eigenvectors u and v, respectively. We show that $\mathrm{cond}((\lambda, \mu), H) = \frac{\|(\|D\lambda\|, \|D\mu\|)\|_2}{\|(\lambda,\mu)\|_2}$, where $(D\lambda,\ D\mu)$ is the Fréchet derivative of (λ, μ) at H, and determine three equivalent representations of $\mathrm{cond}((\lambda, \mu), H)$, of which two are eigenvector-free. Indeed, by setting $\Psi(\lambda, \mu) := [\psi_0(\lambda, \mu), \ldots, \psi_m(\lambda, \mu)]^\top$, we show that

$$\begin{aligned}\mathrm{cond}((\lambda, \mu), H) &= \frac{\|\Psi(\lambda, \mu)\|_{V*}\|(\mathrm{adj}(H(\lambda, \mu)))^*\|_*}{|\bar{\mu}\partial_c \det(H(\lambda, \mu)) - \bar{\lambda}\partial_s \det(H(\lambda, \mu))|} \\ &= \frac{\|\Psi(\lambda, \mu)\|_{V*}\|uv^*\|_*}{|u^*(\bar{\mu}\partial_c H(\lambda, \mu) - \bar{\lambda}\partial_s H(\lambda, \mu))v|},\end{aligned}$$

where $\partial_c \det(H(\lambda,\mu))$ and $\partial_s \det(H(\lambda,\mu))$ (resp., $\partial_c H(\lambda,\mu)$ and $\partial_s H(\lambda,\mu)$) denote the partial derivatives of $\det(H(c,s)$ (resp., $H(c,s)$) with respect to c and s, respectively, evaluated at (λ,μ). Note that the first equality in $\mathrm{cond}((\lambda,\mu),H)$ generalizes (1.2) to the case of nonlinear homogeneous eigenproblem $H(\lambda,\mu)v = 0$. Also note that $\mathrm{cond}((\lambda,\mu),H)$ is invariant under the scaling of (λ,μ), that is, $\mathrm{cond}((\lambda,\mu),H) = \mathrm{cond}(\alpha(\lambda,\mu),H)$ for all nonzero $\alpha \in \mathbb{C}$.

Further, let $\sigma_1(H(\lambda,\mu)) \geq \cdots \geq \sigma_n(H(\lambda,\mu))$ be the singular values of the matrix $H(\lambda,\mu)$. Then, for the spectral norm on $\mathbb{C}^{n\times n}$, we show that

$$\begin{aligned}\mathrm{cond}((\lambda,\mu),H) &= \frac{\|\Psi(\lambda,\mu)\|_{V*}\prod_{j=1}^{n-1}\sigma_j(H(\lambda,\mu))}{|\bar{\mu}\partial_c \det(H(\lambda,\mu)) - \bar{\lambda}\partial_s \det(H(\lambda,\mu))|}\\ &= \frac{\|\Psi(\lambda,\mu)\|_{V*}\|u\|_2\|v\|_2}{|u^*(\bar{\mu}\partial_c H(\lambda,\mu) - \bar{\lambda}\partial_s H(\lambda,\mu))v|}.\end{aligned}$$

The eigenvector-free representation of $\mathrm{cond}(\lambda,P)$ (resp., $\mathrm{cond}((\lambda,\mu),H)$) provides an alternative viewpoint of the sensitivity of λ (resp., (λ,μ)) by bringing into focus the influence of the $n-1$ largest singular values of the matrix $P(\lambda)$ (resp., $H(\lambda,\mu)$).

The rest of the paper is organized as follows. In section 2, we collect some basic results. In section 3, we present a concise framework for the sensitivity analysis of nonhomogeneous nonlinear eigenproblems and derive three equivalent explicit expressions of the condition number of a simple eigenvalue. In section 4, we analyze simple eigenvalues of a parameter dependent nonlinear eigenproblem in which coefficient matrices depend on a parameter $t \in \mathbb{C}^k$ and determine the derivatives of the eigenvalues. In section 5, we present a framework for the sensitivity analysis of homogeneous nonlinear eigenproblems and derive three equivalent explicit expressions of the condition number of a simple eigenvalue. Finally, in section 6 we analyze simple eigenvalues of a parameter dependent nonlinear homogeneous eigenproblem in which coefficient matrices dependent on a parameter $t \in \mathbb{C}^k$ and determine the derivatives of the eigenvalues.

2. Preliminaries. Let V be a finite dimensional Hilbert space equipped with an inner product $\langle , \rangle$. If $\|\cdot\| : V \to \mathbb{R}$ is a norm, then $\|\cdot\|_* : V \to \mathbb{R}$ defined by $\|x\|_* := \sup_{\|y\|=1}\{|\langle x, y\rangle| : y \in V\}$ is a norm on V and is called the dual norm of $\|\cdot\|$. It follows that $|\langle x, y\rangle| \leq \|x\|\,\|y\|_*$ for $x \in V$ and $y \in V$.

We consider the standard inner product $\langle x, y\rangle := y^*x$ on $\mathbb{C}^n$, where y^* is the conjugate transpose of y. The spectral norm on $\mathbb{C}^{n\times n}$ is given by $\|A\|_2 := \sup\{\|Ax\|_2 : \|x\|_2 = 1\}$, where $\|x\|_2 := \sqrt{\langle x, x\rangle}$. We consider the standard inner product $\langle X, Y\rangle := \mathrm{Tr}(Y^*X)$ on $\mathbb{C}^{m\times n}$, where $\mathrm{Tr}(X)$ is the trace of X. Then $\|X\|_F := \sqrt{\langle X, X\rangle}$ is the Frobenius norm on $\mathbb{C}^{m\times n}$.

Let $U \subset V$ be open and $\lambda : U \to \mathbb{C}$. The (Fréchet) derivative $\mathrm{D}\lambda(A)$ of λ at $A \in U$ is a linear map $\mathrm{D}\lambda(A) : V \to \mathbb{C}$ such that

$$\lim_{\|H\|\to 0}\frac{|\lambda(A+H) - \lambda(A) - \mathrm{D}\lambda(A)H|}{\|H\|} = 0.$$

If $\mathrm{D}\lambda(A)$ exists, then there is a unique vector $\nabla\lambda(A) \in V$, called the gradient of λ at A, such that $\mathrm{D}\lambda(A)H = \langle H, \nabla\lambda(A)\rangle$ for all $H \in V$. Consequently, we have

$$\|\mathrm{D}\lambda(A)\| = \sup_{\|H\|=1}|\mathrm{D}\lambda(A)H| = \|\nabla\lambda(A)\|_*.$$

The subdifferential of a norm will play an important role in constructing an optimal perturbation. Let $x_0 \in V$ be nonzero. Then it is well known [37, 38] that

$$\partial\|x_0\| := \{y \in V : \langle y,\, x_0\rangle = \|x_0\| \text{ and } \|y\|_* = 1\} \tag{2.1}$$

is the subdifferential (subgradient) of the map $x \mapsto \|x\|$ at x_0. If $\|\cdot\|$ is differentiable in a neighborhood of x_0, then $\partial\|x_0\| = \{\nabla\|x_0\|\}$. In such a case, we have $\|(\nabla\|x_0\|)\|_* = 1$ and $\langle x_0,\, \nabla\|x_0\|\rangle = \|x_0\|$. For example, if $\|\cdot\|$ is strictly convex, then it is differentiable on $V \setminus \{0\}$. For the special case of the Hölder p-norm on $\mathbb{C}^m$, it is easy to determine the subdifferential $\partial\|x_0\|$ (see [3, 37, 38]).

The adjugate of a matrix $A \in \mathbb{C}^{n\times n}$, denoted by $\mathrm{adj}(A)$, is defined by

$$(\mathrm{adj}(A))_{ji} := (-1)^{i+j}\det(A(i,j)),$$

where $A(i,j)$ is the matrix obtained from A by deleting the ith row and the jth column of A. Let $\det : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}, X \longmapsto \det(X)$, be the determinant map. It is well known that $A\,\mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)I$. Also det is a differentiable function and the derivative $\mathrm{D}\det(A)$ at A is given by the Jacobi formula

$$\mathrm{D}\det(A)X = \mathrm{Tr}(\mathrm{adj}(A)X) = \langle X,\, \mathrm{adj}(A)^*\rangle \ \text{ for } X \in \mathbb{C}^{n\times n}. \tag{2.2}$$

Note that $\nabla\det(A) := \mathrm{adj}(A)^*$ is the gradient of det at A. Also note that $\mathrm{adj}(A) = 0$ when $\mathrm{rank}(A) < n-1$ and $\mathrm{adj}(A) \neq 0$ when $\mathrm{rank}(A) \geq n-1$. If $\mathrm{rank}(A) = n-1$, then

$$\mathrm{adj}(A) = vu^* \tag{2.3}$$

for some nonzero vectors u and v such that $Av = 0$ and $u^*A = 0$ (see [1]).

Let $U \subset \mathbb{C}$ be open and $A : U \longrightarrow \mathbb{C}^{n\times n}$ be *regular*, that is, $\det(A(z)) \neq 0$ for some $z \in \mathbb{C}$. Let $\lambda \in U$. If $\mathrm{rank}(A(\lambda)) < n$, then λ is said to be an *eigenvalue* of $A(z)$. We denote the spectrum (set of eigenvalues) of $A(z)$ by $\mathrm{eig}(A)$. Let $\lambda \in \mathrm{eig}(A)$. Then there exist nonzero vectors x and y such that $A(\lambda)x = 0$ and $y^*A(\lambda) = 0$. The vectors x and y are called right and left eigenvectors of $A(z)$ corresponding to λ, respectively. Further, λ is said to be a multiple (resp., simple) eigenvalue of $A(z)$ if λ is a multiple (resp., simple) zero of $\mathbf{p}(z) := \det(A(z))$. By the Jacobi formula, we have $\partial_z\mathbf{p}(\lambda) = \mathrm{Tr}(\mathrm{adj}(A(\lambda))\partial_z A(\lambda))$, where $\partial_z\mathbf{p}(\lambda)$ and $\partial_z A(\lambda)$ are the derivatives of $\mathbf{p}(z)$ and $A(z)$ at λ. We say that (λ, x, y) is a *simple eigentriple* of $A(z)$ if λ is a simple eigenvalue of $A(z)$ and y and x are corresponding left and right eigenvectors.

THEOREM 2.1 (see [1]). *Let $A : U \longrightarrow \mathbb{C}^{n\times n}$ be regular and differentiable at $\lambda \in U$. Set $\mathbf{p}(z) := \det(A(z))$. Then λ is a multiple eigenvalue of $A(z)$ if and only if there exist left and right eigenvectors u and v, respectively, of $A(z)$ corresponding to λ such that $u^*\partial_z A(\lambda)v = 0$. Further, if (λ, x, y) is a simple eigentriple of $A(z)$, then*

$$\mathrm{adj}(A(\lambda)) = \frac{\partial_z\mathbf{p}(\lambda)xy^*}{y^*\partial_z A(\lambda)x} = \frac{\mathrm{Tr}(\mathrm{adj}(A(\lambda))\partial_z A(\lambda))xy^*}{y^*\partial_z A(\lambda)x}.$$

We mention that a proof of the first result in Theorem 2.1 can be found in [5].

3. Sensitivity analysis of nonhomogeneous eigenproblems.

We now present a framework for the sensitivity analysis of a nonlinear eigenproblem $P(\lambda)v = 0$ and determine the Fréchet derivative $\mathrm{D}\lambda(P)$ and the gradient $\nabla\lambda(P)$ of a simple eigenvalue $\lambda(P)$ of P as a function of P. Also, we derive three equivalent explicit expressions of the condition number of $\lambda(P)$. We proceed as follows.

Let $\Omega \subset \mathbb{C}$ be open and let $\phi_j : \Omega \to \mathbb{C}$ for $j = 0, 1, \ldots, m$ be suitable functions. For the rest of this section, we set $\Phi(z) := [\phi_0(z), \ldots, \phi_m(z)]^\top$ for $z \in \Omega$ and assume that $\phi_0, \ldots, \phi_m$ are analytic on Ω except possibly for poles, if any. We consider the vector space (over $\mathbb{C}$)

$$\mathbb{H}(\Phi, \Omega) := \{\phi_0 A_0 + \cdots + \phi_m A_m : A_j \in \mathbb{C}^{n\times n} \text{ for } j = 0, 1, \ldots, m\}.$$

For $x \in \mathbb{C}^{m+1}$ and $A \in \mathbb{C}^{n\times n}$, define $x^\top \otimes A \in \mathbb{H}(\Phi, \Omega)$ by $(x^\top \otimes A)(\lambda) := \sum_{j=0}^m x_j \phi_j(\lambda) A$.

Let $X, Y \in \mathbb{H}(\Phi, \Omega)$ be given by $X := \sum_{j=0}^m \phi_j X_j$ and $Y := \sum_{j=0}^m \phi_j Y_j$. Then

$$\langle X,\, Y\rangle := \mathrm{Tr}(Y_0^* X_0) + \cdots + \mathrm{Tr}(Y_m^* X_m)$$

defines an inner product on $\mathbb{H}(\Phi, \Omega)$ and $|\!|\!| X |\!|\!|_F := \sqrt{\langle X,\, X\rangle} = (\sum_{j=0}^m \|X_j\|_F^2)^{1/2}$ is a norm on $\mathbb{H}(\Phi, \Omega)$. For a monotone norm $\|\cdot\|_V$ on $\mathbb{C}^{m+1}$, we consider the norm

$$|\!|\!| X |\!|\!|_V := \|(\|X_0\|, \ldots, \|X_m\|)\|_V \tag{3.1}$$

on $\mathbb{H}(\Phi, \Omega)$. Then $|\!|\!| Y |\!|\!|_{V*} := \sup_{|\!|\!| X |\!|\!|_V = 1} |\langle X,\, Y\rangle|$ is the dual norm of $|\!|\!|\cdot|\!|\!|_V$. Note that $|\langle X,\, Y\rangle| \leq |\!|\!| X |\!|\!|_V |\!|\!| Y |\!|\!|_{V*}$ for $X, Y \in \mathbb{H}(\Phi, \Omega)$. It is easily seen that

$$|\!|\!| Y |\!|\!|_{V*} = \|(\|Y_0\|_*, \ldots, \|Y_m\|_*)\|_{V*}, \tag{3.2}$$

where $\|\cdot\|_{V*}$ and $\|\cdot\|_*$ are the dual norms of $\|\cdot\|_V$ and the matrix norm $\|\cdot\|$, respectively.

For example, the monotone norm $\|\cdot\|_V$ can be chosen to be the Hölder p-norm $\|\cdot\|_p$ for $1 \leq p \leq \infty$. Then we have $|\!|\!| X |\!|\!|_p = \|(\|X_0\|, \ldots, \|X_m\|)\|_p$ and its dual norm $|\!|\!| X |\!|\!|_q = \|(\|X_0\|_*, \ldots, \|X_m\|_*)\|_q$, where $p^{-1} + q^{-1} = 1$.

Now consider the map $F : \mathbb{H}(\Phi, \Omega) \times \mathbb{C} \to \mathbb{C}$ given by $F(X, s) := \det(X(s))$. If λ is an eigenvalue of $P \in \mathbb{H}(\Phi, \Omega)$, then obviously (P, λ) is a solution of $F(X, s) = 0$. Consider the variety $\mathbb{V}(F) := \{(X, s) \in \mathbb{H}(\Phi, \Omega) \times \mathbb{C} : F(X, s) = 0\}$. We say that (P, λ) is a *simple point* of $\mathbb{V}(F)$ if P is regular and λ is a simple eigenvalue of P. We now show that if (P, λ_P) is a simple point of $\mathbb{V}(F)$, then there is an open set $U \subset \mathbb{H}(\Phi, \Omega)$ containing P such that $\mathbb{V}(F) \cap (U \times \mathbb{C})$ is the graph of a smooth function $U \to \mathbb{C}, X \mapsto \lambda(X)$, such that $\lambda(P) = \lambda_P$ and $\lambda(X) \in \mathrm{eig}(X)$ is simple for $X \in U$.

THEOREM 3.1. *Let $P \in \mathbb{H}(\Phi, \Omega)$ be regular and $\lambda_P \in \mathrm{eig}(P)$ be simple. Then there is an open set $U \subset \mathbb{H}(\Phi, \Omega)$ containing P and a smooth function $\lambda : U \to \mathbb{C}$ such that $\lambda(P) = \lambda_P$ and $\lambda(X)$ is a simple eigenvalue of X for all $X \in U$. Moreover, for $X \in U$, the derivative $\mathrm{D}\lambda(X) : \mathbb{H}(\Phi, \Omega) \longrightarrow \mathbb{C}$ is given by*

$$\mathrm{D}\lambda(X)H = -\frac{\mathrm{Tr}(\mathrm{adj}(X(\lambda(X))H(\lambda(X)))}{\partial_s \det(X(\lambda(X)))} = -\frac{\mathrm{Tr}(\mathrm{adj}(X(\lambda(X))H(\lambda(X)))}{\mathrm{Tr}(\mathrm{adj}(X(\lambda(X)))\partial_s X(\lambda(X)))}$$

for $H \in \mathbb{H}(\Phi, \Omega)$, where $\partial_s \det(X(\lambda(X)))$ and $\partial_s X(\lambda(X))$ are the partial derivatives of $(X, s) \longmapsto \det(X(s))$ and $(X, s) \longmapsto X(s)$ with respect to s evaluated at $(X, \lambda(X))$. Then the gradient $\nabla\lambda(X)$ is given by

$$\nabla\lambda(X) = -\frac{\Phi(\lambda(X))^* \otimes (\mathrm{adj}(X(\lambda(X))))^*}{\overline{\partial_s \det(X(\lambda(X)))}},$$

where $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$.

In particular, if x and y, respectively, are right and left eigenvectors of P corresponding to λ_P, then we have

$$\mathrm{D}\lambda(P)H = -\frac{\mathrm{Tr}(\mathrm{adj}(P(\lambda_P))H(\lambda_P))}{\partial_s \det(P(\lambda_P))} = -\frac{y^* H(\lambda_P)x}{y^*\partial_s P(\lambda_P)x} \quad \text{for } H \in \mathbb{H}(\Phi, \Omega),$$
$$\nabla\lambda(P) = -\frac{\Phi(\lambda_P)^* \otimes (\mathrm{adj}(P(\lambda_P)))^*}{\overline{\partial_s \det(P(\lambda_P))}} = -\frac{\Phi(\lambda_P)^* \otimes yx^*}{\overline{y^*\partial_s P(\lambda_P)x}}.$$

Proof. Consider the map $F : \mathbb{H}(\Phi, \Omega) \times \mathbb{C} \to \mathbb{C}$ given by $F(X, s) := \det(X(s))$ and the variety $\mathbb{V}(F) := \{(X, s) \in \mathbb{H}(\Phi, \Omega) \times \mathbb{C} : F(X, s) = 0\}$. Then $(P, \lambda_P) \in \mathbb{V}(F)$ is a simple point. Now $\partial_s F(P, \lambda_P) = \partial_s \det(P(\lambda_P)) \neq 0$. Hence by the implicit function theorem there is an open set $U \subset \mathbb{H}(\Phi, \Omega)$ containing P and a smooth function $\lambda : U \to \mathbb{C}$ such that $\lambda(P) = \lambda_P$ and $\mathbb{V}(F) \cap (U \times \mathbb{C}) = \{(X, \lambda(X)) : X \in U\}$. Hence $\lambda(P) = \lambda_P$ and $F(X, \lambda(X)) = 0$ for $X \in U$. Since $\mathbb{V}(F) \cap (U \times \mathbb{C}) = \{(X, \lambda(X)) : X \in U\}$ is the graph of $X \mapsto \lambda(X)$ for $X \in U$, it follows that $\lambda(X)$ is a simple eigenvalue of X for all $X \in U$. Indeed, if $\lambda(X_0)$ is a multiple eigenvalue of X_0 for some $X_0 \in U$, then the eigenvalue λ_P and an eigenvalue $\mu_P \neq \lambda_P$ of P must move and coalesce at $\lambda(X_0)$ when X varies from P to X_0. However, the intersection of two eigenvalue paths $(X, \lambda(X))$ and $(X, \mu(X))$ with $\mu(P) = \mu_P$ at $(X_0, \lambda(X_0))$ would contradict the fact that $\mathbb{V}(F) \cap (U \times \mathbb{C}) = \{(X, \lambda(X)) : X \in U\}$ is the graph of the smooth map $X \mapsto \lambda(X)$ for $X \in U$.

Now differentiating $F(X, \lambda(X)) = 0$ for $X \in U$ with respect to X, we have $\mathrm{D}\lambda(X)H = -\partial_X F(X, \lambda(X))H/\partial_s F(X, \lambda(X))$ for $H \in \mathbb{H}(\Phi, \Omega)$, where $\partial_X F(X, \lambda(X))$ and $\partial_s F(X, \lambda(X))$ are derivatives of $F(X, s)$ with respect to X and s, respectively, evaluated at $(X, \lambda(X))$. By the Jacobi formula, we have

$$\partial_s F(X, \lambda(X)) = \partial_s \det(X(\lambda(X))) = \mathrm{Tr}(\mathrm{adj}(X(\lambda(X)))\partial_s X(\lambda(X))),$$
$$\partial_X F(X, \lambda(X))H = \mathrm{Tr}(\mathrm{adj}(X(\lambda(X)))H(\lambda(X))) = \langle H, \nabla\lambda(X)\rangle\, \partial_s \det(X(\lambda(X))),$$

which yield $\mathrm{D}\lambda(X)$ and $\nabla\lambda(X)$.

Finally, considering $X = P$ and $\lambda(P) = \lambda_P$, the desired results follow from Theorem 2.1 and the fact that $\partial_s \det(P(\lambda_P)) = \mathrm{Tr}(\mathrm{adj}(P(\lambda_P))\partial_s P(\lambda_P))$. ∎

For a simple eigenvalue λ_P of $P(s)$, we have the following results for $\mathrm{cond}(\lambda_P, P)$.

THEOREM 3.2. *Let $P \in \mathbb{H}(\Phi, \Omega)$ be regular and let $\lambda_P \in \mathrm{eig}(P)$ be simple with left and right eigenvectors y and x, respectively. Then, for any matrix norm $\|\cdot\|$, we have*

$$\mathrm{cond}(\lambda_P, P) = \frac{\|\Phi(\lambda_P)\|_{V*}\|(\mathrm{adj}(P(\lambda_P)))^*\|_*}{|\partial_s \det(P(\lambda_P))|} = \frac{\|\Phi(\lambda_P)\|_{V*}\|yx^*\|_*}{|y^*\partial_s P(\lambda_P)x|}, \tag{3.3}$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Let $\sigma_1(P(\lambda_P)) \geq \cdots \geq \sigma_n(P(\lambda_P))$ be the singular values of $P(\lambda_P)$. Then, for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have

$$\mathrm{cond}(\lambda_P, P) = \frac{\|\Phi(\lambda_P)\|_{V*}\prod_{j=1}^{n-1}\sigma_j(P(\lambda_P))}{|\partial_s \det(P(\lambda_P))|} = \frac{\|\Phi(\lambda_P)\|_{V*}\|x\|_2\|y\|_2}{|y^*\partial_s P(\lambda_P)x|}. \tag{3.4}$$

Proof. In view of Theorem 3.1, it is easy to see that the condition number in (1.3) is given by $\mathrm{cond}(\lambda_P, P) = \|\mathrm{D}\lambda(P)\| = \|\nabla\lambda(P)\|_{V*}$, where $\|\mathrm{D}\lambda(P)\| := \sup\{|\mathrm{D}\lambda(P)H| : \|H\|_V = 1 \text{ and } H \in \mathbb{H}(\Phi, \Omega)\}$. Again, by Theorem 3.1, we have

$$\nabla\lambda(P) = -\frac{\Phi(\lambda_P)^* \otimes (\mathrm{adj}(P(\lambda_P)))^*}{\overline{\partial_s \det(P(\lambda_P))}} = -\frac{\Phi(\lambda_P)^* \otimes yx^*}{\overline{y^*\partial_s P(\lambda_P)x}}.$$

Note that $\Phi(\lambda_P)^* \otimes (\mathrm{adj}(P(\lambda_P)))^* \in \mathbb{H}(\Phi, \Omega)$. Hence by (3.2) we have $\|\!|\nabla\lambda(P)|\!\|_{V*} = \|\!|\Phi(\lambda_P)|\!\|_{V*}\|(\mathrm{adj}(P(\lambda_P)))^*\|_*/|\partial_s \det(P(\lambda_P))| = \|\!|\Phi(\lambda_P)|\!\|_{V*}\|yx^*\|_*/|y^*\partial_s P(\lambda_P)x|$, which proves (3.3).

For the spectral and the Frobenius norms, we have $\|yx^*\|_2 = \|x\|_2\|y\|_2 = \|yx^*\|_F$. Since $\mathrm{adj}(P(\lambda_P))$ is a rank-1 matrix, we have $\|\mathrm{adj}(P(\lambda)))\|_2 = \|\mathrm{adj}(P(\lambda)))\|_F$. By [23, Theorem 4], we have $\|\mathrm{adj}(P(\lambda_P))\|_2 = \prod_{j=1}^{n-1}\sigma_j(P(\lambda_P))$. Hence (3.4) follows from (3.3). ∎

Note that the first equality in (3.3) for $\mathrm{cond}(\lambda_P, P)$ generalizes the condition number (1.2) derived by Smith [23] for standard eigenproblems to the case of nonlinear eigenproblems. Since $\partial_s \det(P(\lambda_P)) = 0 \iff \lambda_P$ is a multiple eigenvalue, the eigenvector-free representation of the condition number in (3.4) provides an alternative viewpoint of the sensitivity of λ_P by bringing into focus the influence of the $n-1$ largest singular values of the matrix $P(\lambda_P)$ on the sensitivity of λ_P.

For polynomial eigenproblems, that is, when $\phi_j(s) = s^{j-1}, j = 0, 1, \ldots, m$, the second equality in (3.3) for a subordinate matrix norm yields

$$\mathrm{cond}(\lambda_P, P) = \frac{\|\!|[1, \lambda_P, \ldots, \lambda_P^m]^\top|\!\|_{V*}\|x\|\,\|y\|_*}{|y^*\partial_s P(\lambda_P)x|}. \tag{3.5}$$

The condition number $\mathrm{cond}(\lambda_P, P)$ in (3.5) is derived in [4] by employing a different method.

Let (λ_P, x, y) be a simple eigentriple of $P(s)$. Then by Theorem 3.1, we have

$$\lambda(P + \Delta P) = \lambda_P + \mathrm{D}\lambda(P)\Delta P + \mathcal{O}(\|\!|\Delta P|\!\|_V^2) = \lambda_P + \langle \Delta P,\, \nabla\lambda(P)\rangle + \mathcal{O}(\|\!|\Delta P|\!\|_V^2)$$

for sufficiently small $\Delta P \in \mathbb{H}(\Phi, \Omega)$. Hence by Theorem 3.1, we have the first order expansion

$$\begin{aligned}\lambda(P+\Delta P) &= \lambda_P - \frac{y^*\Delta P(\lambda_P)x}{y^*\partial_s P(\lambda_P)x} + \mathcal{O}(\|\!|\Delta P|\!\|_V^2)\\ &= \lambda_P - \frac{\mathrm{Tr}(\mathrm{adj}(P(\lambda_P))\Delta P(\lambda_P))}{\partial_s \det(P(\lambda_P))} + \mathcal{O}(\|\!|\Delta P|\!\|_V^2)\\ &= \lambda_P - \frac{\langle \Delta P,\, \Phi(\lambda_P)^* \otimes (\mathrm{adj}(P(\lambda_P)))^*\rangle}{\partial_s \det(P(\lambda_P))} + \mathcal{O}(\|\!|\Delta P|\!\|_V^2)\\ &= \lambda_P - \frac{\langle \Delta P,\, \Phi(\lambda_P)^* \otimes yx^*\rangle}{y^*\partial_s P(\lambda_P)x} + \mathcal{O}(\|\!|\Delta P|\!\|_V^2)\end{aligned}$$

and the first order bound $|\lambda(P+\Delta P) - \lambda_P| \lesssim \mathrm{cond}(\lambda_P, P)\|\!|\Delta P|\!\|_V$ for sufficiently small $\Delta P \in \mathbb{H}(\Phi, \Omega)$.

We say that $\Delta P \in \mathbb{H}(\Phi, \Omega)$ is a *fast perturbation* for $\lambda(P) = \lambda_P$ if $\|\!|\Delta P|\!\|_V = 1$ and $\mathrm{D}\lambda(P)\Delta P = \langle \Delta P,\, \nabla\lambda(P)\rangle = \|\!|\nabla\lambda(P)|\!\|_{V*}$. Note that if ΔP is a fast perturbation for λ_P, then the first order bound $|\lambda(P+t\Delta P)-\lambda_P| = \mathrm{cond}(\lambda_P, P)\|\!|t\Delta P|\!\|_V + \mathcal{O}(|t|^2)$ holds for sufficiently small $|t|$, where $t \in \mathbb{C}$. A fast perturbation for λ_P can be constructed as follows.

Let $X \in \partial\|yx^*\|_*$ and $\mu := [\mu_0, \ldots, \mu_m]^\top \in \partial\|\!|\Phi(\lambda_P)|\!\|_{V*}$. Then by (2.1), $\|X\| = 1$ and $\langle X,\, yx^*\rangle = \mathrm{Tr}(xy^*X) = \|yx^*\|_*$. Also $\|\!|\mu|\!\|_V = 1$ and $\langle \mu,\, \Phi(\lambda_P)\rangle = \|\!|\Phi(\lambda_P)|\!\|_{V*}$. Set $\omega := y^*\partial_s P(\lambda_P)x$. Define $\Delta P := \mu^* \otimes X/\mathrm{sign}(\omega) \in \mathbb{H}(\Phi, \Omega)$, where $\mathrm{sign}(\omega) := \bar{\omega}/|\omega|$. Then $\Delta P(z) = (\mu^* \otimes X)(z)/\mathrm{sign}(\omega) = \langle \Phi(z),\, \mu\rangle X/\mathrm{sign}(\omega)$ for $z \in \Omega$. Further, we have $\|\!|\Delta P|\!\|_V = \|\!|\mu|\!\|_V\|X\| = 1$ and $\langle \Delta P,\, \nabla\lambda(P)\rangle = \langle \mu^* \otimes X,\, \Phi(\lambda_P)^* \otimes yx^*\rangle/|\omega| = \langle \Phi(\lambda_P),\, \mu\rangle\langle X,\, yx^*\rangle/|\omega| = \|\!|\Phi(\lambda_P)|\!\|_{V*}\|yx^*\|_*/|\omega| = \|\!|\nabla\lambda(P)|\!\|_{V*}$ showing that ΔP is a fast perturbation for λ_P.

For a subordinate matrix norm or for a matrix norm satisfying $\|yx^*\|_* = \|x\|\,\|y\|_*$, the matrix $X \in \partial\|yx^*\|_*$ can be constructed as follows. Let $u \in \partial\|y\|_*$ and $v \in \partial\|x\|$. Define $X := uv^*$. Then $\langle X, yx^*\rangle = \mathrm{Tr}(xy^*uv^*) = \|x\|\,\|y\|_* = \|yx^*\|_*$ and $\|X\| = \|uv^*\| = \|u\|\,\|v\|_* = 1$, which shows that $X \in \partial\|yx^*\|_*$. This proves the following result.

PROPOSITION 3.3. *Let $P \in \mathbb{H}(\Phi, \Omega)$ be regular and let (λ_P, x, y) be a simple eigentriple of $P(s)$. Let $X \in \partial\|yx^*\|_*$ and $\mu := [\mu_0, \dots, \mu_m]^\top \in \partial\|\Phi(\lambda_P)\|_{V*}$. Define $\Delta P := \mu^* \otimes X/\mathrm{sign}(\omega)$, where $\omega := y^*\partial_s P(\lambda_P)x$. Then $\Delta P \in \mathbb{H}(\Phi, \Omega)$ is a fast perturbation for λ_P, that is, $\|\!|\Delta P|\!\|_V = \|\mu\|_V\|X\| = 1$ and $\langle \Delta P, \nabla\lambda(P)\rangle = \|\Phi(\lambda_P)\|_{V*}\|yx^*\|_*/|\omega| = \|\!|\nabla\lambda(P)|\!\|_{V*}$.*

In particular, for a subordinate matrix norm or for a matrix norm satisfying $\|yx^*\|_* = \|x\|\,\|y\|_*$, define $X := uv^*$, where $u \in \partial\|y\|_*$ and $v \in \partial\|x\|$. Then we have $X \in \partial\|yx^*\|_*$.

Weighted perturbation. Let $w \in \mathbb{R}^{m+1}$ be a nonnegative weight vector, that is, $w := [w_0, \dots, w_m]^\top$ and $w_j \geq 0$ for $j = 0, 1, \dots, m$. For $X := \sum_{j=0}^m X_j\phi_j \in \mathbb{H}(\Phi, \Omega)$, we define $w \odot X := \sum_{j=0}^m w_j X_j \phi_j$. Next, define $w^{-1} := [w_0^{-1}, \dots, w_m^{-1}]^\top$ with the convention that $w_j^{-1} = 0$ when $w_j = 0$ for $j = 0, 1, \dots, m$. We say that $Y \in \mathbb{H}(\Phi, \Omega)$ is w-admissible if $w^{-1} \odot w \odot Y = Y$. If $Y := \sum_{j=0}^m Y_j\phi_j \in \mathbb{H}(\Phi, \Omega)$ is w-admissible and if $w_j = 0$, then $Y_j = 0$. This means that if X is perturbed to $X + Y$, then X_j remains unperturbed. Let $\mathbb{H}_w(\Phi, \Omega)$ denote the subspace of w-admissible elements in $\mathbb{H}(\Phi, \Omega)$. Then

$$\|\!|X|\!\|_{w,V} := \|\!|w \odot X|\!\|_V \text{ for } X \in \mathbb{H}_w(\Phi, \Omega) \tag{3.6}$$

defines a norm on $\mathbb{H}_w(\Phi, \Omega)$. Indeed, if X is w-admissible, then $\|\!|X|\!\|_{w,V} = 0 \iff X = 0$. Hence $\|\!|\cdot|\!\|_{w,V}$ defines a norm on $\mathbb{H}_w(\Phi, \Omega)$. For $X, Y \in \mathbb{H}_w(\Phi, \Omega)$, we have

$$\langle X, Y\rangle = \langle w \odot X, w^{-1} \odot Y\rangle,$$

which shows that $\|\!|\cdot|\!\|_{w^{-1},V*}$ is the dual norm of the norm $\|\!|\cdot|\!\|_{w,V}$ on $\mathbb{H}_w(\Phi, \Omega)$ and that $|\langle X, Y\rangle| \leq \|\!|X|\!\|_{w,V}\|\!|Y|\!\|_{w^{-1},V*}$.

Now, to measure the sensitivity of $\lambda \in \mathrm{eig}(P)$ as P is perturbed to $P + \Delta P$ for small $\Delta P \in \mathbb{H}_w(\Phi, \Omega)$, we define the weighted condition number of λ by

$$\mathrm{cond}_w(\lambda, P) := \limsup_{\|\!|\Delta P|\!\|_{w,V} \to 0} \frac{\mathrm{dist}(\lambda,\ \mathrm{eig}(P + \Delta P))}{\|\!|\Delta P|\!\|_{w,V}}, \quad \text{where } \Delta P \in \mathbb{H}_w(\Phi, \Omega). \tag{3.7}$$

THEOREM 3.4. *Let $P \in \mathbb{H}(\Phi, \Omega)$ be regular and let (λ_P, x, y) be a simple eigentriple of $P(s)$. Then, for any matrix norm, we have*

$$\mathrm{cond}_w(\lambda_P, P) = \frac{\|w^{-1} \odot \Phi(\lambda_P)\|_{V*}\|(\mathrm{adj}(P(\lambda_P)))^*\|_*}{|\partial_s \det(P(\lambda_P))|} = \frac{\|w^{-1} \odot \Phi(\lambda_P)\|_{V*}\|yx^*\|_*}{|y^*\partial_s P(\lambda_P)x|},$$

where $w^{-1} \odot \Phi(\lambda_P)$ is the Hadamard (componentwise) product of w^{-1} and $\Phi(\lambda_P)$. Let $\sigma_1(P(\lambda_P)) \geq \cdots \geq \sigma_n(P(\lambda_P))$ be the singular values of $P(\lambda_P)$. Then, for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have

$$\mathrm{cond}_w(\lambda_P, P) = \frac{\|w^{-1} \odot \Phi(\lambda_P)\|_{V*}\prod_{j=1}^{n-1}\sigma_j(P(\lambda_P))}{|\partial_s \det(P(\lambda_P))|} = \frac{\|w^{-1} \odot \Phi(\lambda_P)\|_{V*}\|x\|_2\|y\|_2}{|y^*\partial_s P(\lambda_P)x|}.$$

Proof. By Theorem 3.1, we have

$$\nabla\lambda(P) = -\frac{\Phi(\lambda_P)^* \otimes (\mathrm{adj}(P(\lambda_P)))^*}{\overline{\partial_s \det(P(\lambda_P))}} = -\frac{\Phi(\lambda_P)^* \otimes yx^*}{\overline{y^*\partial_s P(\lambda_P)x}}$$

and $\mathrm{D}\lambda(P)H = \langle H, \nabla\lambda(P)\rangle = \langle w \odot H, w^{-1} \odot \nabla\lambda(P)\rangle$ for $H \in \mathbb{H}_w(\Phi, \Omega)$. Hence

$$\begin{aligned}\mathrm{cond}_w(\lambda_P, P) &= \sup\{|\mathrm{D}\lambda(P)H| : \|H\|_{w,V} = 1 \text{ and } H \in \mathbb{H}_w(\Phi,\Omega)\}\\ &= \sup\{|\langle w\odot H, w^{-1}\odot\nabla\lambda(P)\rangle| : \|w\odot H\|_V = 1 \text{ and } H \in \mathbb{H}_w(\Phi,\Omega)\}\\ &= \|w^{-1}\odot\nabla\lambda(P)\|_{V*},\end{aligned}$$

which yields the desired results. ∎

For a matrix polynomial $P(s) := \sum_{j=0}^m A_j s^j$, we have $\phi_j(s) = s^{j-1}$ for $j = 0, 1, \ldots, m$. Let (λ, x, y) be a simple eigentriple of $P(s)$. Then the condition number

$$\mathrm{cond}_w(\lambda, P) := \frac{\|x\|_2\|y\|_2 \sum_{j=0}^m \alpha_j|\lambda|^j}{|y^*\partial_s P(\lambda)x|}, \tag{3.8}$$

where $\alpha_0, \ldots, \alpha_m$ are nonnegative weights, is derived in [31]. Notice that $\mathrm{cond}_w(\lambda, P)$ in (3.8) follows from Theorem 3.4 with $w = [\alpha_0^{-1}, \ldots, \alpha_m^{-1}]^\top$, the spectral norm $\|\cdot\|_2$ on $\mathbb{C}^{n\times n}$, and $\|\cdot\|_V = \|\cdot\|_\infty$, the ∞-norm on $\mathbb{C}^{m+1}$. More generally, for these choices of norms and weight vector and for $P := \sum_{j=0}^m A_j\phi_j$, Theorem 3.4 yields

$$\mathrm{cond}_w(\lambda, P) = \frac{\|x\|_2\|y\|_2 \sum_{j=0}^m \alpha_j|\phi_j(\lambda)|}{|y^*\partial_s P(\lambda)x|}. \tag{3.9}$$

The condition number $\mathrm{cond}_w(\lambda, P)$ in (3.9) is derived in [13]. Finally, considering $P(\lambda) := A - \lambda I$ with $\Phi(\lambda) := [1, \lambda]^\top$ and $w := [1, 0]^\top$, we obtain from Theorem 3.4 the condition number $\mathrm{cond}(\lambda, A)$ in (1.1) and (1.2) derived in [34, 23]; see also [1].

Scaling of eigenproblems. The theory of condition is a classical subject that deals with a canonical way of defining the condition number of a problem [34, 21, 10, 27]. The condition number measures the sensitivity of the problem and provides a first order variation of the solution. The canonical approach to conditioning in [21] leads to the condition numbers in (1.1), (1.3), and (1.5) for standard, nonlinear nonhomogeneous, and homogeneous eigenproblems, respectively. Now consider the eigenproblem $P(\lambda)u = 0$ and observe that the eigenvalues are neutral to the scaling $P \mapsto \alpha P$ for $\alpha > 0$, that is, the scaling does not affect the eigenvalues. Note, however, that $\mathrm{cond}(\lambda_P, P)$ as well as $\mathrm{cond}_w(\lambda_P, P)$ are not neutral to the scaling $P \mapsto \alpha P$. Indeed, we have $\mathrm{cond}(\lambda_P, \alpha P) = \mathrm{cond}(\lambda_P, P)/\alpha$ and $\mathrm{cond}_w(\lambda_P, \alpha P) = \mathrm{cond}_w(\lambda_P, P)/\alpha$.

The condition number faithfully conveys the fact that scaling affects the behavior of eigenvalues under absolute perturbations $P \mapsto P + \Delta P$. Indeed, the unscaled perturbed problem $(P(\lambda) + \Delta P(\lambda))v = 0$ and the scaled perturbed problem $(\alpha P(\lambda) + \Delta P(\lambda))v = 0$ are very different eigenproblems. The first order expansions of simple eigenvalues of the unscaled and scaled eigenproblems are given by

$$\begin{aligned}\lambda(P + \Delta P) &= \lambda_P + \langle \Delta P, \nabla\lambda(P)\rangle + \mathcal{O}(\|\Delta P\|_V^2),\\ \lambda(\alpha P + \Delta P) &= \lambda_P + \langle \Delta P/\alpha, \nabla\lambda(P)\rangle + \mathcal{O}((\|\Delta P\|_V/\alpha)^2),\end{aligned} \tag{3.10}$$

which in turn yield the first order variation of the eigenvalue λ_P given by

$$\begin{aligned}|\lambda(\alpha P + \Delta P) - \lambda_P| &\le \mathrm{cond}(\lambda_P, P)\|\Delta P\|_V/\alpha + \mathcal{O}((\|\Delta P\|_V/\alpha)^2)\\ &= \mathrm{cond}(\lambda_P, \alpha P)\|\Delta P\|_V + \mathcal{O}((\|\Delta P\|_V/\alpha)^2).\end{aligned}$$

The expansion in (3.10) and the first order variation of eigenvalues show that the scaling $P \mapsto \alpha P$ affects the sensitivity of the eigenproblem and the condition number $\mathrm{cond}(\lambda_P, \alpha P)$ detects and faithfully reports the impact. This shows that a badly scaled problem will have an adverse impact on the computed eigenvalues. Similar results hold for weighted perturbations and $\mathrm{cond}_w(\lambda_P, P)$.

We mention that if the absolute perturbation $|\!|\!|\Delta P|\!|\!|_V$ (resp., $|\!|\!|\Delta P|\!|\!|_{w,V}$) is replaced with relative perturbation $|\!|\!|\Delta P|\!|\!|_V/|\!|\!|P|\!|\!|_V$ (resp., $|\!|\!|\Delta P|\!|\!|_{w,V}/|\!|\!|P|\!|\!|_{w,V}$) in (1.3) (resp., (3.7)), then the resulting *relative condition numbers*, which we denote by $\kappa(\lambda_P, P)$ and $\kappa_w(\lambda_P, P)$, will be neutral to the scaling of P. Indeed, we have

$$\kappa(\lambda_P, P) = \mathrm{cond}(\lambda_P, P)|\!|\!|P|\!|\!|_V \text{ and } \kappa_w(\lambda_P, P) = \mathrm{cond}_w(\lambda_P, P)|\!|\!|P|\!|\!|_{w,V},$$

which are easily seen to be neutral to the scaling $P \mapsto \alpha P$. Notice, however, that (3.10) still holds and the first order variation of eigenvalues is still affected by the scaling. Indeed, we have

$$|\lambda(\alpha P + \Delta P) - \lambda_P| \leq \kappa(\lambda_P, P)\frac{|\!|\!|\Delta P|\!|\!|_V}{|\!|\!|\alpha P|\!|\!|_V} + \mathcal{O}((|\!|\!|\Delta P|\!|\!|_V/\alpha)^2).$$

The inability of the relative condition number $\kappa(\lambda_P, P)$ to detect the (adverse) effect of the scaling of P on the perturbed eigenvalues renders the scale neutral relative condition number $\kappa(\lambda_P, P)$ to be an ineffective and impractical measure of sensitivity of eigenvalues. Thus, for the sensitivity analysis of eigenvalues under the perturbation $P \mapsto P + \Delta P$, the scale neutrality of the condition number $\kappa(\lambda_P, P)$ with respect to the scaling $P \mapsto \alpha P$ is not a desirable property. The same is true for $\kappa_w(\lambda_P, P)$.

4. Parameter dependent eigenproblem. The analysis of eigenvalues of a matrix $A(t) \in \mathbb{C}^{n\times n}$ depending on a parameter $t \in \mathbb{C}^k$ is a classical topic and has been studied extensively over the years; see [6, 16, 9, 22, 28] and the references therein. A parameter dependent nonlinear eigenvalue problem has been analyzed, for example, in [5]. We now show that the evolution of a simple eigenvalue of a parameter dependent nonlinear eigenproblem can be deduced from the framework in section 3 and Theorem 3.1.

Let $U \subset \mathbb{C}^k$ be open and let $P : U \times \Omega \longrightarrow \mathbb{C}^{n\times n}$ be such that $t \mapsto P(t, \cdot)$ is holomorphic on U and $z \mapsto P(\cdot, z)$ is homomorphic on Ω except possibly for poles. We say that $P(t, z)$ is regular at $\hat{t} \in U$ if $\det(P(\hat{t}, z)) \neq 0$ for some $z \in \Omega$.

THEOREM 4.1. *Let $\hat{t} \in U$ and $P(\hat{t}, z)$ be regular. Let $(\hat{\lambda}, x_0, y_0)$ be a simple eigentriple of $P(\hat{t}, z)$. Then there is an open set $\mathrm{Nbd}(\hat{t}) \subset U$ containing $\hat{t}$ and a holomorphic function $\lambda : \mathrm{Nbd}(\hat{t}) \longrightarrow \mathbb{C}$ such that $\lambda(\hat{t}) = \hat{\lambda}$ and $\lambda(t)$ is a simple eigenvalue of $P(t, z)$ for all $t \in \mathrm{Nbd}(\hat{t})$. Further, for $t \in \mathrm{Nbd}(\hat{t})$ we have*

$$\lambda(t + h) = \lambda(t) + \mathrm{D}\lambda(t)h + \mathcal{O}(\|h\|_2^2) = \lambda(t) + \langle h, \nabla\lambda(t)\rangle + \mathcal{O}(\|h\|_2^2)$$

for sufficiently small $\|h\|_2$. The derivative $\mathrm{D}\lambda(t)$ and the gradient $\nabla\lambda(t)$ are given by

$$\mathrm{D}\lambda(t)h = -\frac{\sum_{j=1}^k \mathrm{Tr}(\mathrm{adj}(P(t,\lambda(t)))\partial_{t_j} P(t,\lambda(t)))h_j}{\partial_z \det(P(t,\lambda(t)))},$$
$$\nabla\lambda(t) = -\left(\frac{[\mathrm{Tr}(\mathrm{adj}(P(t,\lambda(t)))\partial_{t_1} P(t,\lambda(t))), \ldots, \mathrm{Tr}(\mathrm{adj}(P(t,\lambda(t)))\partial_{t_k} P(t,\lambda(t)))]}{\partial_z \det(P(t,\lambda(t)))}\right)^*$$

for all $h \in \mathbb{C}^k$, where $\partial_{t_j} P(t, \lambda(t))$ is the partial derivative of $P(t, z)$ with respect to t_j and $\partial_z \det(P(t, \lambda(t)))$ is the partial derivative of $\det(P(t, z))$ with respect to z

evaluated at $(t, \lambda(t))$. *In particular, we have*

$$\frac{\partial \lambda(\hat{t})}{\partial t_j} = -\frac{\operatorname{Tr}\left(\operatorname{adj}(P(\hat{t}, \hat{\lambda}))\partial_{t_j} P(\hat{t}, \hat{\lambda})\right)}{\partial_z \det(P(\hat{t}, \hat{\lambda}))} = -\frac{y_0^* \partial_{t_j} P(\hat{t}, \hat{\lambda}) x_0}{y_0^* \partial_z P(\hat{t}, \hat{\lambda}) x_0},\ j = 1, 2, \dots, k.$$

Proof. Consider $p(t, z) := \det(P(t, z))$ for $(t, z) \in U \times \Omega$ and the variety $\mathbb{V}(p) := \{(t, z) \in U \times \Omega : p(t, z) = 0\}$. Then $(\hat{t}, \hat{\lambda}) \in \mathbb{V}(p)$ and $\partial_z p(\hat{t}, \hat{\lambda}) \neq 0$. Hence by the implicit function theorem and by similar arguments as those in the proof of Theorem 3.1, we obtain a holomorphic function $\lambda : \operatorname{Nbd}(\hat{t}) \to \mathbb{C}$ such that $\lambda(\hat{t}) = \hat{\lambda}$ and $\lambda(t)$ is a simple eigenvalue of $P(t, z)$ for $t \in \operatorname{Nbd}(\hat{t})$. Again by Theorem 3.1 and the chain rule, we have $\mathrm{D}\lambda(t)h = -\frac{\operatorname{Tr}(\operatorname{adj}(P(t,\lambda(t)))\mathrm{D}P(t,\lambda(t))h)}{\partial_z p(t,\lambda(t))}$ for all $h \in \mathbb{C}^k$, where $\mathrm{D}P(t, \lambda(t))$ is the derivative of $P(t, z)$ with respect to t evaluated at $(t, \lambda(t))$. The derivative $\mathrm{D}P(t, \lambda(t)) : \mathbb{C}^k \to \mathbb{C}^{n \times n}$ is a linear map and is given by $\mathrm{D}P(t, \lambda(t))h = \sum_{j=1}^k \partial_{t_j} P(t, \lambda(t)) h_j$ for all $h \in \mathbb{C}^k$. Hence for $h \in \mathbb{C}^k$, we have

$$\mathrm{D}\lambda(t)h = -\frac{\sum_{j=1}^k \operatorname{Tr}\left(\operatorname{adj}(P(t, \lambda(t)))\partial_{t_j} P(t, \lambda(t))\right) h_j}{\partial_z p(t, \lambda(t))} = \langle h,\, \nabla\lambda(t)\rangle,$$

which yields the desired results. $\square$

5. Sensitivity analysis of homogeneous eigenproblems. Sensitivity analysis of homogeneous polynomial eigenvalue problems has been considered in [11, 12]. In this section, we present a simple and concise framework for the sensitivity analysis of nonlinear homogeneous eigenvalue problems. The main feature of our framework is that our sensitivity analysis is eigenvector-free and is undertaken in the inner product space $\mathbb{C}^2$, which avoids the apparatus of projective spaces, and is akin to the framework developed for the sensitivity analysis of nonlinear nonhomogeneous eigenvalue problems.

We consider a homogeneous matrix-valued function $P : \mathbb{C}^2 \to \mathbb{C}^{n \times n}$ of the form $P(c, s) := \sum_{j=0}^m A_j \psi_j(c, s)$, where $\psi_j : \mathbb{C}^2 \to \mathbb{C}$ for $j = 0, 1, \dots, m$ are homogeneous functions of degree ℓ. This encompasses homogeneous polynomial eigenproblems of degree m as a special case when $\psi_j(c, s) := c^{m-j} s^j$ for $j = 0, 1, \dots, m$. For the rest of this section, set $\Psi(c, s) := [\psi_0(c, s), \dots, \psi_m(c, s)]^\top$ for $(c, s) \in \mathbb{C}^2$ and consider the vector space (over $\mathbb{C}$)

$$\mathbb{H}(\Psi) := \{\psi_0 A_0 + \dots + \psi_m A_m : A_j \in \mathbb{C}^{n \times n} \text{ for } j = 0, 1, \dots, m\}.$$

For $x \in \mathbb{C}^{m+1}$ and $A \in \mathbb{C}^{n \times n}$, define $x^\top \otimes A \in \mathbb{H}(\Psi)$ by

$$(x^T \otimes A)(c, s) := \sum_{j=0}^m x_j \psi_j(c, s) A.$$

The norms and inner product on $\mathbb{H}(\Psi)$ are defined as in the case of $\mathbb{H}(\Phi, \Omega)$.

Let $P \in \mathbb{H}(\Psi)$ be regular, that is, $\operatorname{rank}(P(c, s)) = n$ for some $(c, s) \in \mathbb{C}^2$. Recall that a nonzero $(\lambda, \mu) \in \mathbb{C}^2$ is an eigenvalue of $P(c, s)$ if $\operatorname{rank}(P(\lambda, \mu)) < n$. Also $\operatorname{eig}(P) := \{(\lambda, \mu) \in \mathbb{C}^2 : (\lambda, \mu) \neq 0 \text{ and } \operatorname{rank}(P(\lambda, \mu)) < n\}$ is the spectrum of P. Let $(\lambda, \mu) \in \operatorname{eig}(P)$. Then the nonzero vectors x and y in $\mathbb{C}^n$ are right and left eigenvectors of $P(c, s)$ corresponding to (λ, μ) if $P(\lambda, \mu)x = 0$ and $y^* P(\lambda, \mu) = 0$, respectively. We refer to $((\lambda, \mu), x, y)$ as an eigentriple of $P(c, s)$.

Define $\mathbf{p}(c, s) := \det(P(c, s))$. Then $(\lambda, \mu) \in \operatorname{eig}(P) \iff \mathbf{p}(\lambda, \mu) = 0$ and $(\lambda, \mu) \neq 0$. We say that $(\lambda, \mu) \in \operatorname{eig}(P)$ is a simple eigenvalue of $P(c, s)$ if $(\partial_c \mathbf{p}(\lambda, \mu)),\ \partial_s \mathbf{p}(\lambda, \mu))$

$\neq 0$, where $\partial_c \mathbf{p}(\lambda, \mu)$ and $\partial_s \mathbf{p}(\lambda, \mu)$ are the partial derivatives of $\mathbf{p}(c, s)$ with respect to c and s, respectively, at (λ, μ). An eigentriple $((\lambda, \mu), x, y)$ is said to be simple when (λ, μ) is a simple eigenvalue of $P(c, s)$.

We denote the partial derivatives of $P(c, s)$ with respect to c and s at (λ, μ) by $\partial_c P(\lambda, \mu)$ and $\partial_s P(\lambda, \mu)$, respectively. The next result provides a characterization of multiple eigenvalues of $P(c, s)$.

THEOREM 5.1. *Let $P \in \mathbb{H}(\Psi)$ be regular and $(\lambda, \mu) \in \mathrm{eig}(P)$. Then the following conditions are equivalent.*

(a) *(λ, μ) is a multiple eigenvalue of $P(c, s)$.*
(b) $\bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu) = 0$.
(c) *There exist right and left eigenvectors x and y of $P(c, s)$ corresponding to (λ, μ) such that $y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x = 0$.*

Proof. Obviously (a) implies (b). Now suppose that (b) holds. By Jacobi formula, we have $\bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu) = \mathrm{Tr}(\mathrm{adj}(P(\lambda, \mu))(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu)))$. If $\mathrm{rank}(P(\lambda, \mu)) = n - 1$, then by (2.3) we have $\mathrm{adj}(P(\lambda, \mu)) = xy^*$ for some right and left eigenvectors x and y of $P(c, s)$ corresponding to (λ, μ). This shows that $y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x = \bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu) = 0$. On the other hand, if $\mathrm{rank}(P(\lambda, \mu)) < n - 1$, then the linear map

$$f : \mathcal{N}(P(\lambda, \mu)) \longrightarrow \mathbb{C}, u \longmapsto y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))u,$$

has a nontrivial null space and hence there is a nonzero $x \in \mathcal{N}(P(\lambda, \mu))$ such that $f(x) = 0$, where y is a left eigenvector of $P(c, s)$ corresponding to (λ, μ) and $\mathcal{N}(P(\lambda, \mu))$ is the null space of $P(\lambda, \mu)$. This shows that (b) implies (c).

Finally, suppose that (c) holds. By Jacobi formula, $(\partial_c \mathbf{p}(\lambda, \mu),\ \partial_s \mathbf{p}(\lambda, \mu)) = (\mathrm{Tr}(\mathrm{adj}(P(\lambda, \mu))\partial_c P(\lambda, \mu)),\ \mathrm{Tr}(\mathrm{adj}(P(\lambda, \mu))\partial_s P(\lambda, \mu))$. Now if $\mathrm{rank}(P(\lambda, \mu)) < n-1$, then $\mathrm{adj}(P(\lambda, \mu)) = 0 \implies (\partial_c \mathbf{p}(\lambda, \mu),\ \partial_s \mathbf{p}(\lambda, \mu)) = 0$, which shows that (λ, μ) is a multiple eigenvalue. On the other hand, if $\mathrm{rank}(P(\lambda, \mu)) = n - 1$, then by (2.3) we have $\mathrm{adj}(P(\lambda, \mu)) = vu^*$ for some left and right eigenvectors u and v of $P(c, s)$ corresponding to (λ, μ). Since $\mathrm{rank}(P(\lambda, \mu)) = n - 1$, we have $u = \alpha y$ and $v = \beta x$ for some nonzero scalars α and β. Consequently, $(\bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu)) = \bar{\alpha}\beta y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x = 0$. Next, note that $\mathbf{p}(c, s)$ is homogeneous. Hence we have $\lambda\partial_c \mathbf{p}(\lambda, \mu) + \mu\partial_s \mathbf{p}(\lambda, \mu) = \deg(\mathbf{p})\mathbf{p}(\lambda, \mu) = 0$. Since (λ, μ) and $(\bar{\mu}, -\bar{\lambda})$ are orthogonal and hence a basis of $\mathbb{C}^2$, we conclude that $(\partial_c \mathbf{p}(\lambda, \mu),\ \partial_s \mathbf{p}(\lambda, \mu)) = 0$, which shows that (λ, μ) is a multiple eigenvalue. This proves (a). ∎

Let $(\lambda, \mu) \in \mathrm{eig}(P)$ be fixed. Let $X \in \mathbb{H}(\Psi)$ be regular and $(\alpha, \beta) \in \mathrm{eig}(X)$. Since a scalar multiple of (α, β) is an eigenvalue of X, we normalize (α, β) such that $(\alpha - \lambda, \beta - \mu)$ is orthogonal to (λ, μ), which we write as $(\alpha - \lambda,\ \beta - \mu) \perp (\lambda, \mu)$. Such a normalization always holds unless (α, β) is a scalar multiple of (λ, μ).

PROPOSITION 5.2. *Let $P \in \mathbb{H}(\Psi)$ be regular and $(\lambda, \mu) \in \mathrm{eig}(P)$. Define*

$$G : \mathbb{C}^2 \longrightarrow \mathbb{C}^2, (c, s) \longmapsto (\det(P(c, s)),\ (c - \lambda)\bar{\lambda} + (s - \mu)\bar{\mu}).$$

Then (λ, μ) is simple if and only if (λ, μ) is a simple zero of G.

Let $((\lambda, \mu), x, y)$ be a simple eigentriple of $P(c, s)$. Set $\mathbf{p}(c, s) := \det(P(c, s))$. Then

$$\mathrm{adj}(P(\lambda, \mu)) = \frac{(\bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu))xy^*}{y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x}.$$

Proof. The Jacobian matrix of G at (λ, μ) is given by

$$J_G(\lambda, \mu) = \begin{bmatrix} \partial_c \mathbf{p}(\lambda, \mu) & \partial_s \mathbf{p}(\lambda, \mu) \\ \bar{\lambda} & \bar{\mu} \end{bmatrix}.$$

Since $\det(J_G(\lambda, \mu)) = \bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu)$, by Theorem 5.1 it follows that (λ, μ) is simple if and only if $J_G(\lambda, \mu)$ is nonsingular.

Finally, by (2.3) we have $\operatorname{adj}(P(\lambda, \mu)) = vu^*$ for some left and right eigenvectors u and v of $P(c, s)$ corresponding to (λ, μ). Since (λ, μ) is simple, we have $u = \alpha y$ and $v = \beta x$ for some nonzero scalars α and β. Hence $\operatorname{adj}(P(\lambda, \mu) = \bar{\alpha}\beta x y^*$. Now by the Jacobi formula we have

$$\begin{aligned}\det(J_G(\lambda, \mu)) &= \bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu) = \operatorname{Tr}(\operatorname{adj}(P(\lambda, \mu))(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))) \\ &= \bar{\alpha}\beta y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x,\end{aligned}$$

which gives $\bar{\alpha}\beta = (\bar{\mu}\partial_c \mathbf{p}(\lambda, \mu) - \bar{\lambda}\partial_s \mathbf{p}(\lambda, \mu))/y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x$. Hence the desired result follows from $\operatorname{adj}(P(\lambda, \mu)) = \bar{\alpha}\beta x y^*$. ∎

Let $P \in \mathbb{H}(\Psi)$ be regular and $(\lambda_P, \mu_P) \in \operatorname{eig}(P)$. Define

$$F : \mathbb{H}(\Psi) \times \mathbb{C}^2 \longrightarrow \mathbb{C}^2, (X, (c, s)) \longmapsto (\det(X(c, s)),\ (c - \lambda_P)\bar{\lambda}_P + (s - \mu_P)\bar{\mu}_P)$$

and consider the variety $\mathbb{V}(F) := \{(X, (c, s)) \in \mathbb{H}(\Psi) \times \mathbb{C}^2 : F(X, c, s) = 0\}$. If (λ, μ) is an eigenvalue of $X \in \mathbb{H}(\Psi)$ and $(\lambda - \lambda_P, \mu - \mu_P) \perp (\lambda_P, \mu_P)$, then $(X, (\lambda, \mu)) \in \mathbb{V}(F)$. We say that $(X, (\lambda, \mu)) \in \mathbb{V}(F)$ is a *simple point* of $\mathbb{V}(F)$ if X is regular and (λ, μ) is a simple eigenvalue of X. We now show that if $(P, (\lambda_P, \mu_P))$ is a simple point of $\mathbb{V}(F)$, then there is an open set $U \subset \mathbb{H}(\Psi)$ containing P such that $\mathbb{V}(F) \cap (U \times \mathbb{C}^2)$ is the graph of a smooth function $U \to \mathbb{C}^2, X \longmapsto (\lambda(X), \mu(X))$, such that $(\lambda(P), \mu(P) = (\lambda_P, \mu_P)$, $(\lambda(X) - \lambda_P, \mu(X) - \mu_P) \perp (\lambda_P, \mu_P)$ and $(\lambda(X), \mu(X)) \in \operatorname{eig}(X)$ is simple for $X \in U$.

THEOREM 5.3. *Let $P \in \mathbb{H}(\Psi)$ be regular and $(\lambda_P, \mu_P) \in \operatorname{eig}(P)$ be simple. Then there exists an open set $U \subset \mathbb{H}(\Psi)$ containing P and a smooth function $U \to \mathbb{C}^2, X \longmapsto (\lambda(X), \mu(X))$, such that $(\lambda(P), \mu(P)) = (\lambda_P, \mu_P)$, $(\lambda(X) - \lambda_P, \mu(X) - \mu_P) \perp (\lambda_P, \mu_P)$ and $(\lambda(X), \mu(X))$ is a simple eigenvalue of X for all $X \in U$. Moreover, for $X \in U$, the derivative $(\mathrm{D}\lambda(X), \mathrm{D}\mu(X)) : \mathbb{H}(\Psi) \longrightarrow \mathbb{C}^2$ is given by*

(5.1)
$$(\mathrm{D}\lambda(X)H,\ \mathrm{D}\mu(X)H) = -\frac{\operatorname{Tr}(\operatorname{adj}(X(\lambda(X), \mu(X)))H(\lambda(X), \mu(X)))\ (\bar{\mu}_P,\ -\bar{\lambda}_P)}{\bar{\mu}_P \partial_c \det(X(\lambda(X), \mu(X))) - \bar{\lambda}_P \partial_s \det(X(\lambda(X), \mu(X)))}$$

for $H \in \mathbb{H}(\Psi)$, where $\partial_z \det(X(\lambda(X), \mu(X)))$ is the partial derivative of $(X, (c, s)) \longmapsto \det(X(c, s))$ with respect to z evaluated at $(X, \lambda(X), \mu(X))$ for $z = c$ and $z = s$. Then the gradient is given by

(5.2)
$$(\nabla\lambda(X), \nabla\mu(X)) = -\frac{\Psi(\lambda(X), \mu(X))^* \otimes (\operatorname{adj}(X(\lambda(X), \mu(X))))^*\ (\mu_P,\ -\lambda_P)}{\overline{(\bar{\mu}_P \partial_c \det(X(\lambda(X), \mu(X))) - \bar{\lambda}_P \partial_s \det(X(\lambda(X), \mu(X))))}},$$

where $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$.

Set $\mathbf{p}(c, s) := \det(P(c, s))$. If x and y, respectively, are right and left eigenvectors of P corresponding to (λ_P, μ_P), then we have

$$\begin{aligned}(\mathrm{D}\lambda(P)H,\ \mathrm{D}\mu(P)H) &= -\frac{\operatorname{Tr}(\operatorname{adj}(P(\lambda_P, \mu_P))H(\lambda_P, \mu_P))}{\bar{\mu}_P \partial_c \mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s \mathbf{p}(\lambda_P, \mu_P)}(\bar{\mu}_P,\ -\bar{\lambda}_P) \\ &= -\frac{y^* H(\lambda_P, \mu_P)x}{y^*(\bar{\mu}_P \partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s P(\lambda_P, \mu_P))x}(\bar{\mu}_P,\ -\bar{\lambda}_P)\end{aligned}$$

for $H \in \mathbb{H}(\Psi)$, *and*

$$(\nabla\lambda(P),\ \nabla\mu(P)) = -\frac{\Psi(\lambda_P,\mu_P)^* \otimes (\mathrm{adj}(P(\lambda_P,\mu_P)))^*}{\overline{(\bar{\mu}_P\partial_c\mathbf{p}(\lambda_P,\mu_P) - \bar{\lambda}_P\partial_s\mathbf{p}(\lambda_P,\mu_P))}}(\mu_P,-\lambda_P)$$
$$= -\frac{\Psi(\lambda_P,\mu_P)^* \otimes yx^*}{\overline{(y^*(\bar{\mu}_P\partial_c P(\lambda_P,\mu_P) - \bar{\lambda}_P\partial_s P(\lambda_P,\mu_P))x)}}(\mu_P,-\lambda_P).$$

Proof. Consider $F : \mathbb{H}(\Psi)\times\mathbb{C}^2 \to \mathbb{C}^2, (X,(c,s)) \longmapsto (\det(X(c,s)),\ (c-\lambda_P)\bar{\lambda}_P + (s-\mu_P)\bar{\mu}_P)$ and the variety $\mathbb{V}(F) := \{(X,(c,s)) \in \mathbb{H}(\Psi)\times\mathbb{C}^2 : F(X,c,s) = 0\}$. Then note that $(P,(\lambda_P,\mu_P)) \in \mathbb{V}(F)$ is a simple point. Now for a fixed X, the Jacobian matrix of the map $\mathbb{C}^2 \to \mathbb{C}^2, (c,s) \longmapsto F(X,c,s)$ at (P,λ_P,μ_P) is given by

$$J(P,\lambda_P,\mu_P) = \begin{bmatrix} \partial_c\mathbf{p}(\lambda_P,\mu_P) & \partial_s\mathbf{p}(\lambda_P,\mu_P) \\ \bar{\lambda}_P & \bar{\mu}_P \end{bmatrix}.$$

By Proposition 5.2, $J(P,\lambda_P,\mu_P)$ is nonsingular. Hence by the implicit function theorem there is an open set $U \subset \mathbb{H}(\Psi)$ containing P and a smooth function $U \to \mathbb{C}^2, X \longmapsto (\lambda(X),\mu(X))$ such that $(\lambda(P),\mu(P)) = (\lambda_P,\mu_P)$ and $\mathbb{V}(F)\cap(U\times\mathbb{C}^2) = \{(X,(\lambda(X),\mu(X))) : X \in U\}$. Hence $F(X,\lambda(X),\mu(X)) = 0$ for $X \in U$. This shows that $(\lambda(X),\mu(X)) \in \mathrm{eig}(X)$ and $(\lambda(X)-\lambda_P,\ \mu(X)-\mu_P) \perp (\lambda_P,\mu_P)$ for $X \in U$ with $(\lambda(P),\mu(P)) = (\lambda_P,\mu_P)$.

Since $\mathbb{V}(F)\cap(U\times\mathbb{C}^2) = \{(X,(\lambda(X),\mu(X))) : X \in U\}$ is the graph of $X \longmapsto (\lambda(X),\mu(X))$ for $X \in U$, it follows that $(\lambda(X),\mu(X))$ is a simple eigenvalue of X for all $X \in U$. Indeed, if $(\lambda(X_0),\mu(X_0))$ is a multiple eigenvalue of X_0 for some $X_0 \in U$, then the eigenvalue (λ_P,μ_P) and an eigenvalue $(\alpha_P,\beta_P) \neq (\lambda_P,\mu_P)$ of P must move and coalesce at $(\lambda(X_0),\mu(X_0))$ when X varies from P to X_0. However, the intersection of two eigenvalue paths $(X,(\lambda(X)\mu(X)))$ and $(X,(\alpha(X),\beta(X)))$ with $(\alpha(P),\beta(P)) = (\alpha_P,\beta_P)$ at $(X_0,(\lambda(X_0),\mu(X_0)))$ would contradict the fact that $\mathbb{V}(F)\cap(U\times\mathbb{C}^2) = \{(X,(\lambda(X),\mu(X))) : X \in U\}$ is the graph of the smooth map $X \longmapsto (\lambda(X),\mu(X))$ for $X \in U$.

Now differentiating $F(X,\lambda(X),\mu(X)) = 0$ for $X \in U$ with respect to X, we have $(\mathrm{D}\lambda(X)H,\ \mathrm{D}\mu(X)H) = -(\partial_{(c,s)}F(X,\lambda(X),\mu(X)))^{-1}\partial_X F(X,\lambda(X),\mu(X))H$ for $H \in \mathbb{H}(\Psi)$, where $\partial_X F(X,\lambda(X),\mu(X))$ and $\partial_{(c,s)}F(X,\lambda(X),\mu(X))$ are derivatives of $F(X,c,s)$ with respect to X and (c,s), respectively, evaluated at $(X,(\lambda(X),\mu(X)))$. Now the matrix of $\partial_{(c,s)}F(X,\lambda(X),\mu(X))$ is the Jacobian matrix

$$J(X,\lambda(X),\mu(X)) = \begin{bmatrix} \partial_c\det(X(\lambda(X),\mu(X)) & \partial_s\det(X(\lambda(X),\mu(X)) \\ \bar{\lambda}_P & \bar{\mu}_P \end{bmatrix}.$$

By Jacobi formula, we have

$$\partial_X F(X,\lambda(X),\mu(X))H = (\mathrm{Tr}(\mathrm{adj}(X(\lambda(X),\mu(X)))H(\lambda(X),\mu(X))),\ 0)$$

for $H \in \mathbb{H}(\Psi)$. This in turn leads us to

$$[\mathrm{D}\lambda(X)H,\ \mathrm{D}\mu(X)H]^\top$$
$$= (J(X,\lambda(X),\mu(X)))^{-1}[\mathrm{Tr}(\mathrm{adj}(X(\lambda(X),\mu(X)))H(\lambda(X),\mu(X))),\ 0]^\top,$$

which yields (5.1).

Next note that for $H \in \mathbb{H}(\Psi)$ we have $\mathrm{Tr}(\mathrm{adj}(X(\lambda(X),\mu(X)))H(\lambda(X),\mu(X))) = \langle H,\ \Psi(\lambda(X),\mu(X))^* \otimes (\mathrm{adj}(X(\lambda(X),\mu(X))))^*\rangle$. Hence (5.2) follows from (5.1).

Finally, considering $X = P$ and the fact that $(\lambda(P), \mu(P)) = (\lambda_P, \mu_P)$, the desired expressions for $(\mathrm{D}\lambda(P)H,\ \mathrm{D}\mu(P)H)$ and $(\nabla\lambda(P), \nabla\mu(P))$ follow from (5.1), (5.2), and Proposition 5.2. ∎

Let (λ_P, μ_P) be a simple eigenvalue of $P(c, s)$. Then we have the following result for $\mathrm{cond}((\lambda_P, \mu_P), P)$.

THEOREM 5.4. *Let $P \in \mathbb{H}(\Psi)$ be regular and $(\lambda_P, \mu_P) \in \mathrm{eig}(P)$ be simple with left and right eigenvectors y and x, respectively. Set $\mathbf{p}(c, s) := \det(P(c, s))$ for $(c, s) \in \mathbb{C}^2$. Then, for any matrix norm $\|\cdot\|$, we have*

$$\begin{aligned}\mathrm{cond}((\lambda_P, \mu_P), P) &= \frac{\|\Psi(\lambda_P, \mu_P)\|_{V*}\,\|(\mathrm{adj}(P(\lambda_P, \mu_P)))^*\|_*}{\left|\left(\bar{\mu}_P\partial_c\mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P\partial_s\mathbf{p}(\lambda_P, \mu_P)\right)\right|}\\ &= \frac{\|\Psi(\lambda_P, \mu_P)\|_{V*}\|yx^*\|_*}{\left|\left(y^*(\bar{\mu}_P\partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P\partial_s P(\lambda_P, \mu_P))x\right)\right|},\end{aligned}$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Let $\sigma_1(P(\lambda_P, \mu_P)) \geq \cdots \geq \sigma_n(P(\lambda_P, \mu_P))$ be the singular values of $P(\lambda_P, \mu_P)$. Then, for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have

$$\begin{aligned}\mathrm{cond}((\lambda_P, \mu_P), P) &= \frac{\|\Psi(\lambda_P, \mu_P)\|_{V*}\prod_{j=1}^{n-1}\sigma_j(P(\lambda_P, \mu_P))}{\left|\left(\bar{\mu}_P\partial_c\mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P\partial_s\mathbf{p}(\lambda_P, \mu_P)\right)\right|}\\ &= \frac{\|\Psi(\lambda_P, \mu_P)\|_{V*}\|x\|_2\|y\|_2}{\left|\left(y^*(\bar{\mu}_P\partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P\partial_s P(\lambda_P, \mu_P))x\right)\right|}.\end{aligned}$$

Proof. Note that, in view of Theorem 5.3, it is easily seen that the condition number in (1.5) is given by

$$\mathrm{cond}((\lambda_P, \mu_P), P) = \frac{\|(\|\mathrm{D}\lambda(P)\|,\ \|\mathrm{D}\mu(P)\|)\|_2}{\|(\lambda_P, \mu_P)\|_2} = \frac{\|(\|\nabla\lambda(P)\|_{V*},\ \|\nabla\mu(P)\|_{V*})\|_2}{\|(\lambda_P, \mu_P)\|_2},$$

where $\|\mathrm{D}\lambda(P)\| = \sup\{|\mathrm{D}\lambda(P)H| : \|H\|_V = 1 \text{ and } H \in \mathbb{H}(\Psi)\}$. Hence the desired results follow from Theorem 5.3 and the fact that $\mathrm{adj}(P(\lambda_P, \mu_P))$ is a rank-1 matrix so that $\|\mathrm{adj}(P(\lambda_P, \mu_P))\|_F = \|\mathrm{adj}(P(\lambda_P, \mu_P))\|_2 = \prod_{j=1}^{n-1}\sigma_j(P(\lambda_P, \mu_P))$, where the last equality follows from [23, Theorem 4]. ∎

Note that the first eigenvector-free representation of $\mathrm{cond}((\lambda_P, \mu_P), P)$ in Theorem 5.4 generalizes the condition number (1.2) derived by Smith [23] for standard eigenproblems to the case of nonlinear homogeneous eigenproblems. The quantity $\mathrm{D}_N\mathbf{p}(\lambda_P, \mu_P) := \left(\bar{\mu}_P\partial_c\mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P\partial_s\mathbf{p}(\lambda_P, \mu_P)\right)$ appearing in the denominator of $\mathrm{cond}((\lambda_P, \mu_P), P)$ is the directional derivative of $\mathbf{p}(c, s)$ at (λ_P, μ_P) along the direction $(\bar{\mu}_P, -\bar{\lambda}_P)$, which is orthogonal to (λ_P, μ_P). Thus $\mathrm{D}_N\mathbf{p}(\lambda_P, \mu_P)$ is the normal derivative of $\mathbf{p}(c, s)$ at (λ_P, μ_P) along the normal $(\bar{\mu}_P, -\bar{\lambda}_P)$ to the line $\mathrm{span}\{(\lambda_P, \mu_P)\}$. By Theorem 5.1, $\mathrm{D}_N\mathbf{p}(\lambda_P, \mu_P) = 0 \iff (\lambda_P, \mu_P)$ is a multiple eigenvalue. Hence the smaller the value of $|\mathrm{D}_N\mathbf{p}(\lambda_P, \mu_P)|$ the more ill-conditioned the eigenvalue (λ_P, μ_P) is expected to be. Additionally, the eigenvector-free representation of $\mathrm{cond}((\lambda_P, \mu_P), P)$ brings into focus the influence of the $n-1$ largest singular values of the matrix $P(\lambda_P, \mu_P)$ on the sensitivity of (λ_P, μ_P).

Consider a regular homogeneous matrix polynomial $P(c, s) := \sum_{j=0}^{m} A_j c^{m-j}s^j$. Let $((\lambda, \mu), x, y)$ be a simple eigentriple of $P(c, s)$. Then for the spectral norm on $\mathbb{C}^{n\times n}$ the condition number

$$\mathrm{cond}((\lambda, \mu), P) = \frac{\|[\lambda^m, \lambda^{m-1}\mu, \ldots, \lambda\mu^{m-1}, \mu^m]^\top\|_2\|x\|_2\,\|y\|_2}{\left|\left(y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x\right)\right|} \tag{5.3}$$

is derived in [12] by employing an entirely different method; see also [14, Theorem 2.3]. Notice that the condition number in (5.3) follows from Theorem 5.4 as a special case when $\psi_j(c,s) = c^{m-j}s^j, j = 0,1,\ldots,m$, $\|\cdot\|_V = \|\cdot\|_2$ and the matrix norm is the spectral norm. Our derivation of the condition number $\mathrm{cond}((\lambda_P,\mu_P),P)$ leading to Theorem 5.4 avoids the apparatus of projective spaces and is concise and simple. Finally, considering $P(c,s) := cA - sB$ with $\Psi(c,s) := [c,s]^\top$, $\|\cdot\|_V = \|\cdot\|_2$, and the eigenvalue $(\lambda,\mu) := (y^*Bx, y^*Ax)$, we obtain from Theorem 5.4 the condition number

$$\mathrm{cond}((\lambda,\mu),P) = \frac{\|x\|_2\|y\|_2}{\sqrt{|y^*Ax|^2 + |y^*Bx|^2}}$$

derived by Stewart and Sun [25] when the matrix norm is the spectral norm.

Let $P \in \mathbb{H}(\Psi)$ be regular and let $((\lambda_P,\mu_P),x,y)$ be a simple eigentriple of $P(c,s)$. Then by Theorem 5.3, we have

$$\begin{aligned}
\lambda(P+\Delta P) - \lambda_P &= \mathrm{D}\lambda(P)\Delta P + \mathcal{O}(\|\Delta P\|_V^2) = \langle \Delta P,\ \nabla\lambda(P)\rangle + \mathcal{O}(\|\Delta P\|_V^2),\\
\mu(P+\Delta P) - \mu_P &= \mathrm{D}\mu(P)\Delta P + \mathcal{O}(\|\Delta P\|_V^2) = \langle \Delta P,\ \nabla\mu(P)\rangle + \mathcal{O}(\|\Delta P\|_V^2)
\end{aligned}$$

for sufficiently small $\Delta P \in \mathbb{H}(\Psi)$. Hence we have the first order expansion

$$\begin{aligned}
\lambda(P+\Delta P) &= \lambda_P - \bar\mu_P \frac{y^*\Delta P(\lambda_P,\mu_P)x}{\big(y^*(\bar\mu_P\partial_c P(\lambda_P,\mu_P) - \bar\lambda_P\partial_s P(\lambda_P,\mu_P))x\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \lambda_P - \bar\mu_P \frac{\mathrm{Tr}(\mathrm{adj}(P(\lambda_P,\mu_P))\Delta P(\lambda_P,\mu_P))}{\big(\bar\mu_P\partial_c \mathbf{p}(\lambda_P,\mu_P) - \bar\lambda_P\partial_s \mathbf{p}(\lambda_P,\mu_P)\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \lambda_P - \bar\mu_P \frac{\langle \Delta P,\ \Psi(\lambda_P,\mu_P)^* \otimes (\mathrm{adj}(P(\lambda_P,\mu_P)))^*\rangle}{\big(\bar\mu_P\partial_c \mathbf{p}(\lambda_P,\mu_P) - \bar\lambda_P\partial_s \mathbf{p}(\lambda_P,\mu_P)\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \lambda_P - \bar\mu_P \frac{\langle \Delta P,\ \Psi(\lambda_P,\mu_P)^* \otimes yx^*\rangle}{\big(y^*(\bar\mu_P\partial_c P(\lambda_P,\mu_P) - \bar\lambda_P\partial_s P(\lambda_P,\mu_P))x\big)} + \mathcal{O}(\|\Delta P\|_V^2),\\
\mu(P+\Delta P) &= \mu_P + \bar\lambda_P \frac{y^*\Delta P(\lambda_P,\mu_P)x}{\big(y^*(\bar\mu_P\partial_c P(\lambda_P,\mu_P) - \bar\lambda_P\partial_s P(\lambda_P,\mu_P))x\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \mu_P + \bar\lambda_P \frac{\mathrm{Tr}(\mathrm{adj}(P(\lambda_P,\mu_P))\Delta P(\lambda_P,\mu_P))}{\big(\bar\mu_P\partial_c \mathbf{p}(\lambda_P,\mu_P) - \bar\lambda_P\partial_s \mathbf{p}(\lambda_P,\mu_P)\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \mu_P + \bar\lambda_P \frac{\langle \Delta P,\ \Psi(\lambda_P,\mu_P)^* \otimes (\mathrm{adj}(P(\lambda_P,\mu_P)))^*\rangle}{\big(\bar\mu_P\partial_c \mathbf{p}(\lambda_P,\mu_P) - \bar\lambda_P\partial_s \mathbf{p}(\lambda_P,\mu_P)\big)} + \mathcal{O}(\|\Delta P\|_V^2)\\
&= \mu_P + \bar\lambda_P \frac{\langle \Delta P,\ \Psi(\lambda_P,\mu_P)^* \otimes yx^*\rangle}{\big(y^*(\bar\mu_P\partial_c P(\lambda_P,\mu_P) - \bar\lambda_P\partial_s P(\lambda_P,\mu_P))x\big)} + \mathcal{O}(\|\Delta P\|_V^2)
\end{aligned}$$

and the first order bounds

$$\begin{aligned}
|\lambda(P+\Delta P) - \lambda_P| &\lesssim |\mu_P|\,\mathrm{cond}((\lambda_P,\mu_P),P)\|\Delta P\|_V,\\
|\mu(P+\Delta P) - \mu_P| &\lesssim |\lambda_P|\mathrm{cond}((\lambda_P,\mu_P),P)\|\Delta P\|_V
\end{aligned}$$

for sufficiently small $\Delta P \in \mathbb{H}(\Psi)$.

Recall that $\mathbb{CP}^1 := \{[c:s] : (c,s) \in \mathbb{C}^2 \setminus \{0\}\}$, where $[c:s] := \mathrm{span}\{(c,s)\}$. The Riemannian metric on $\mathbb{CP}^1$ is given by

$$\mathbf{d}_{\mathrm{R}}([c:s],[\lambda:\mu]) := \arccos\left(\frac{|\langle (c,s),\ (\lambda,\mu)\rangle|}{\|(c,s)\|_2\|(\lambda,\mu)\|_2}\right).$$

Define $\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu]) := \tan\left(\mathbf{d}_{\mathrm{R}}([c:s],[\lambda:\mu])\right)$. Although, $\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu])$ is not a metric on $\mathbb{CP}^1$, we have $\mathbf{d}_{\mathrm{R}}([c:s],[\lambda:\mu]) \le \mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu])$ (see [11]).

PROPOSITION 5.5. *Let $(c,s) \in \mathbb{C}^2 \setminus \{0\}$ and $(\lambda,\mu) \in \mathbb{C}^2 \setminus \{0\}$ be such that $(c-\lambda, s-\mu) \perp (\lambda,\mu)$. Then we have*

$$\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu]) = \frac{\|(c,s)-(\lambda,\mu)\|_2}{\|(\lambda,\mu)\|_2}.$$

Proof. Since $(c-\lambda, s-\mu) \perp (\lambda,\mu)$, we have $\|(\lambda,\mu)\|_2^2 = \langle (c,s),\,(\lambda,\mu)\rangle$. Consequently, we have

$$\begin{aligned}\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu]) &= \sqrt{\frac{\|(c,s)\|_2^2\|(\lambda,\mu)\|_2^2}{|\langle (c,s),\,(\lambda,\mu)\rangle|^2}-1} = \sqrt{\frac{\|(c,s)\|_2^2}{\|(\lambda,\mu)\|_2^2}-1}\\ &= \sqrt{\frac{\|(c,s)-(\lambda,\mu)\|_2^2+\|(\lambda,\mu)\|_2^2}{\|(\lambda,\mu)\|_2^2}-1} = \frac{\|(c,s)-(\lambda,\mu)\|_2}{\|(\lambda,\mu)\|_2}.\end{aligned}$$

This completes the proof. ∎

By Theorems 5.3 and 5.4 and Proposition 5.5, we have the following result.

COROLLARY 5.6. *Let $(\lambda,\mu) \in \mathrm{eig}(P)$ be simple and $\Delta P \in \mathbb{H}(\Psi)$. Then for sufficiently small $\|\Delta P\|_V$, there exists a simple eigenvalue $(c,s) \in \mathrm{eig}(P+\Delta P)$ such that $(c-\lambda, s-\mu) \perp (\lambda,\mu)$ and*

$$\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu]) \le \mathrm{cond}((\lambda,\mu),P)\|\Delta P\|_V + \mathcal{O}(\|\Delta P\|_V^2).$$

Remark 5.7. Notice that if $(\lambda,\mu) \in \mathrm{eig}(P)$ is simple, then $\mathrm{cond}((\lambda,\mu),P)$ is invariant with respect to the scaling of (λ,μ), that is, $\mathrm{cond}((\lambda,\mu),P) = \mathrm{cond}(\alpha(\lambda,\mu),P)$ for all nonzero $\alpha \in \mathbb{C}$. Hence it makes sense to define $\mathrm{cond}([\lambda:\mu],P) := \mathrm{cond}((\lambda,\mu),P)$. Alternatively, $\mathrm{cond}([\lambda:\mu],P)$ can be defined by

$$\mathrm{cond}([\lambda:\mu],P) := \limsup_{\|\Delta P\|_V \to 0} \frac{\min\{\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu]) :\ (c,s) \in \mathrm{eig}(P+\Delta P)\}}{\|\Delta P\|_V}. \tag{5.4}$$

Then by Proposition 5.5 and (1.5), we have $\mathrm{cond}([\lambda:\mu],P) = \mathrm{cond}((\lambda,\mu),P)$. This shows that $\mathrm{cond}((\lambda,\mu),P)$ can be defined via $\mathbf{d}_{\mathrm{T}}([c:s],[\lambda:\mu])$ and (5.4) can be taken as a definition of $\mathrm{cond}((\lambda,\mu),P)$.

We say that $\Delta P \in \mathbb{H}(\Psi)$ is a *fast perturbation* for $(\lambda(P),\mu(P)) = (\lambda_P,\mu_P)$ if $\|\Delta P\|_V = 1$ and $|\mathrm{D}\lambda(P)\Delta P| = |\langle \Delta P,\,\nabla\lambda(P)\rangle| = \|\nabla\lambda(P)\|_{V*}$, $|\mathrm{D}\mu(P)\Delta P| = |\langle \Delta P,\,\nabla\mu(P)\rangle| = \|\nabla\mu(P)\|_{V*}$. Note that if ΔP is a fast perturbation for (λ_P,μ_P), then the first order bounds $|\lambda(P+t\Delta P)-\lambda_P| = |\mu_P|\mathrm{cond}(\lambda_P,\mu_P,P)\|t\Delta P\|_V + \mathcal{O}(|t|^2)$ and $|\mu(P+t\Delta P)-\mu_P| = |\lambda_P|\mathrm{cond}(\lambda_P,\mu_P,P)\|t\Delta P\|_V + \mathcal{O}(|t|^2)$ hold for sufficiently small $|t|$, where $t \in \mathbb{C}$. A fast perturbation for (λ_P,μ_P) can be constructed as follows.

Let $X \in \partial\|yx^*\|_*$ and $\xi := [\xi_0,\ldots,\xi_m]^\top \in \partial\|\Psi(\lambda_P,\mu_P)\|_{V*}$. Then $\|X\| = 1$ and $\langle X,\, yx^*\rangle = \mathrm{Tr}(xy^*X) = \|yx^*\|_*$. Also $\|\xi\|_V = 1$ and $\langle \xi,\,\Psi(\lambda_P,\mu_P)\rangle = \|\Psi(\lambda_P,\mu_P)\|_{V*}$. Set $\omega := y^*(\bar{\mu}_P\partial_c P(\lambda_P,\mu_P) - \bar{\lambda}_P\partial_s P(\lambda_P,\mu_P))x$. Define $\Delta P := \xi^* \otimes X/\mathrm{sign}(\omega) \in \mathbb{H}(\Psi)$, where $\mathrm{sign}(\omega) := \bar{\omega}/|\omega|$. Then

$$\Delta P(c,s) = (\xi^* \otimes X)(c,s)/\mathrm{sign}(\omega) = \langle \Psi(c,s),\,\xi\rangle X/\mathrm{sign}(\omega) \ \text{ for } \ (c,s) \in \mathbb{C}^2.$$

Now $\|\Delta P\|_V = \|\xi\|_V\|X\| = 1$ and $|\langle \Delta P,\,\nabla\lambda(P)\rangle| = |\mu_P|\,|\langle \Psi(\lambda_P,\mu_P),\,\xi\rangle\langle X,\,yx^*\rangle|/|\omega| = |\mu_P|\,\|\Psi(\lambda_P,\mu_P)\|_{V*}\|yx^*\|_*/|\omega| = \|\nabla\lambda(P)\|_{V*}$. Similarly, we have $|\langle \Delta P,\,\nabla\mu(P)\rangle| = |\lambda_P|\,\|\Psi(\lambda_P,\mu_P)\|_{V*}\|yx^*\|_*/|\omega| = \|\nabla\mu(P)\|_{V*}$. This shows that ΔP is a fast perturbation for (λ_P,μ_P).

As shown in section 3, for a subordinate matrix norm or for a matrix norm satisfying $\|yx^*\|_* = \|x\|\,\|y\|_*$, the matrix $X \in \partial\|yx^*\|_*$ can be constructed as follows. Let $u \in \partial\|y\|_*$ and $v \in \partial\|x\|$. Define $X := uv^*$. Then $X \in \partial\|yx^*\|_*$.

Thus we have the following result.

Proposition 5.8. *Let $P \in \mathbb{H}(\Psi)$ be regular and let $((\lambda_P, \mu_P), x, y)$ be a simple eigentriple of $P(c, s)$. Set $\omega := y^*(\bar{\mu}_P \partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s P(\lambda_P, \mu_P))x$. Let $X \in \partial\|yx^*\|_*$ and $\xi := [\xi_0, \dots, \xi_m]^\top \in \partial\|\Psi(\lambda_P, \mu_P)\|_{V*}$. Define $\Delta P := \xi^* \otimes X/\mathrm{sign}(\omega)$. Then $\Delta P \in \mathbb{H}(\Psi)$ is a fast perturbation for (λ_P, μ_P), that is, $\|\!|\Delta P|\!\|_V = \|\xi\|_V \|X\| = 1$ and*

$$|\langle \Delta P,\, \nabla\lambda(P)\rangle| = |\mu_P|\, \|\Psi(\lambda_P, \mu_P)\|_{V*}\|yx^*\|_*/|\omega| = \|\!|\nabla\lambda(P)|\!\|_{V*},$$
$$|\langle \Delta P,\, \nabla\mu(P)\rangle| = |\lambda_P|\, \|\Psi(\lambda_P, \mu_P)\|_{V*}\|yx^*\|_*/|\omega| = \|\!|\nabla\mu(P)|\!\|_{V*}.$$

In particular, for a subordinate matrix norm or for a matrix norm satisfying $\|yx^*\|_* = \|x\|\,\|y\|_*$, define $X := uv^*$, where $u \in \partial\|y\|_*$ and $v \in \partial\|x\|$. Then $X \in \partial\|yx^*\|_*$.

Weighted perturbations. Let $\mathbb{H}_w(\Psi)$ denote the subspace of w-admissible elements in $\mathbb{H}(\Psi)$. Let $P \in \mathbb{H}(\Psi)$ be regular. Now, in view of Remark 5.7, to measure the sensitivity of $(\lambda_P, \mu_P) \in \mathrm{eig}(P)$ as P is perturbed to $P + \Delta P$ for small $\Delta P \in \mathbb{H}_w(\Psi)$, we consider the weighted condition number

$$\mathrm{cond}_w((\lambda_P, \mu_P), P) := \limsup_{\|\!|\Delta P|\!\|_{w,V} \to 0} \frac{\min\{\mathbf{d}_{\mathrm{T}}([c : s], [\lambda_P : \mu_P]) :\ (c, s) \in \mathrm{eig}(P + \Delta P)\}}{\|\!|\Delta P|\!\|_{w,V}},$$

where $\Delta P \in \mathbb{H}_w(\Psi)$. We have the following result.

Theorem 5.9. *Let $P \in \mathbb{H}(\Psi)$ be regular and let $((\lambda_P, \mu_P), x, y)$ be a simple eigentriple of $P(c, s)$. Set $\mathbf{p}(c, s) := \det(P(c, s))$ for $(c, s) \in \mathbb{C}^2$. Then, for any matrix norm on $\mathbb{C}^{n\times n}$, we have*

$$\begin{aligned}\mathrm{cond}_w((\lambda_P, \mu_P), P) &= \frac{\|w^{-1} \odot \Psi(\lambda_P, \mu_P)\|_{V*}\, \|(\mathrm{adj}(P(\lambda_P, \mu_P)))^*\|_*}{|\left(\bar{\mu}_P \partial_c \mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s \mathbf{p}(\lambda_P, \mu_P)\right)|} \\ &= \frac{\|w^{-1} \odot \Psi(\lambda_P, \mu_P)\|_{V*}\|yx^*\|_*}{|\left(y^*(\bar{\mu}_P \partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s P(\lambda_P, \mu_P))x\right)|},\end{aligned}$$

where $w^{-1} \odot \Psi(\lambda_P)$ is the Hadamard (componentwise) product of w^{-1} and $\Psi(\lambda_P)$. Let $\sigma_1(P(\lambda_P, \mu_P)) \geq \cdots \geq \sigma_n(P(\lambda_P, \mu_P))$ be the singular values of $P(\lambda_P, \mu_P)$. Then, for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have

$$\begin{aligned}\mathrm{cond}_w((\lambda_P, \mu_P), P) &= \frac{\|w^{-1} \odot \Psi(\lambda_P, \mu_P)\|_{V*} \prod_{j=1}^{n-1} \sigma_j(P(\lambda_P, \mu_P))}{|\left(\bar{\mu}_P \partial_c \mathbf{p}(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s \mathbf{p}(\lambda_P, \mu_P)\right)|} \\ &= \frac{\|w^{-1} \odot \Psi(\lambda_P, \mu_P)\|_{V*}\|x\|_2\|y\|_2}{|\left(y^*(\bar{\mu}_P \partial_c P(\lambda_P, \mu_P) - \bar{\lambda}_P \partial_s P(\lambda_P, \mu_P))x\right)|}.\end{aligned}$$

Proof. Note that for $H \in \mathbb{H}_w(\Psi)$, we have

$$\mathrm{D}\lambda(P)H = \langle H,\, \nabla\lambda(P)\rangle = \langle w \odot H,\, w^{-1} \odot \nabla\lambda(P)\rangle,$$
$$\mathrm{D}\mu(P)H = \langle H,\, \nabla\mu(P)\rangle = \langle w \odot H,\, w^{-1} \odot \nabla\mu(P)\rangle.$$

This shows that

$$\begin{aligned}\|\mathrm{D}\lambda(P)\|_w &:= \sup\{|\mathrm{D}\lambda(P)H| : \|H\|_{w,V} = 1 \text{ and } H \in \mathbb{H}_w(\Psi)\}\\ &= \sup\{|\langle w \odot H,\, w^{-1} \odot \nabla\lambda(P)\rangle| : \|w \odot H\|_V = 1 \text{ and } H \in \mathbb{H}_w(\Psi)\}\\ &= \|w^{-1} \odot \nabla\lambda(P)\|_{V*}.\end{aligned}$$

Similarly, we have $\|\mathrm{D}\mu(P)\|_w = \|w^{-1} \odot \nabla\mu(P)\|_{V*}$. Now, in view of Theorem 5.3 and Proposition 5.5 , we have

$$\mathrm{cond}_w((\lambda_P, \mu_P), P) = \frac{\|(\|w^{-1} \odot \nabla\lambda(P)\|_{V*},\ \|w^{-1} \odot \nabla\mu(P)\|_{V*})\|_2}{\|(\lambda_P, \mu_P)\|_2}.$$

Hence the desired result follows from Theorem 5.3. ∎

For a matrix polynomial $P(c, s) := \sum_{j=0}^m A_j c^{m-j} s^j$, we have $\psi_j(c, s) = c^{m-j} s^j$ for $j = 0, 1, \ldots, m$. Let $((\lambda, \mu), x, y)$ be a simple eigentriple of $P(c, s)$. Then the condition number

$$\mathrm{cond}_w((\lambda, \mu), P) := \frac{\|x\|_2\|y\|_2 \sum_{j=0}^m \alpha_j |\lambda|^{m-j}|\mu|^j}{|y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x|},$$

where $\alpha_0, \ldots, \alpha_m$ are nonnegative weights, follows from Theorem 5.9 with $w = (\alpha_0^{-1}, \ldots, \alpha_m^{-1})$, the spectral norm $\|\cdot\|_2$ on $\mathbb{C}^{n\times n}$ and $\|\cdot\|_V = \|\cdot\|_\infty$, the ∞-norm on $\mathbb{C}^{m+1}$. On the other hand, considering $\|\cdot\|_V = \|\cdot\|_2$, the 2-norm on $\mathbb{C}^{m+1}$, we have

$$\mathrm{cond}_w((\lambda, \mu), P) := \frac{\|x\|_2\|y\|_2 \sqrt{\sum_{j=0}^m \alpha_j^2 |\lambda|^{2(m-j)}|\mu|^{2j}}}{|y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x|}.$$

More generally, for $P := \sum_{j=0}^m A_j \psi_j$ and the above-mentioned choices of norms and the weight vector, Theorem 5.9 yields

$$\mathrm{cond}_w((\lambda, \mu), P) := \frac{\|x\|_2\|y\|_2 \sum_{j=0}^m \alpha_j |\psi_j(\lambda, \mu)|}{|y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x|}$$

and

$$\mathrm{cond}_w((\lambda, \mu), P) := \frac{\|x\|_2\|y\|_2 \sqrt{\sum_{j=0}^m \alpha_j^2 |\psi_j(\lambda, \mu)|^2}}{|y^*(\bar{\mu}\partial_c P(\lambda, \mu) - \bar{\lambda}\partial_s P(\lambda, \mu))x|}.$$

Finally, we mention that an analogue of Corollary 5.6 also holds for weighted perturbations. Indeed, we have the following bound:

$$\mathbf{d}_{\mathrm{T}}([c : s], [\lambda : \mu]) \leq \mathrm{cond}_w((\lambda, \mu), P)\|\Delta P\|_{w,V} + \mathcal{O}(\|\Delta P\|_{w,V}^2).$$

A similar bound is presented in [14, bound (2.3)] when P is a homogeneous matrix polynomial.

Scaling. Note that a homogeneous eigenproblem $P(\lambda, \mu)v = 0$ can be scaled in two ways without affecting the eigenvalues, namely, the scaling of the eigenproblem $P \mapsto \alpha P$ and the scaling of the eigenvalues $(\lambda_P, \mu_P) \mapsto \alpha(\lambda_P, \mu_P)$. The condition number $\mathrm{cond}((\lambda_P, \mu_P), P)$ as defined in (1.5) takes into account the scaling of (λ_P, μ_P) by considering relative distance between eigenvalues. Consequently, $\mathrm{cond}((\lambda_P, \mu_P), P)$ is neutral to the scaling $(\lambda_P, \mu_P) \mapsto \alpha(\lambda_P, \mu_P)$, which we have already seen. The same is true for the weighted condition number $\mathrm{cond}_w((\lambda_P, \mu_P), P)$.

Notice that the condition numbers $\mathrm{cond}((\lambda_P, \mu_P), P)$ and $\mathrm{cond}_w((\lambda_P, \mu_P), P)$ are not neutral to the scaling $P \mapsto \alpha P$ for $\alpha > 0$. In fact, we have

$$\begin{aligned}\mathrm{cond}((\lambda_P, \mu_P), \alpha P) &= \mathrm{cond}((\lambda_P, \mu_P), P)/\alpha \text{ and } \mathrm{cond}_w((\lambda_P, \mu_P), \alpha P)\\ &= \mathrm{cond}_w((\lambda_P, \mu_P), P)/\alpha,\end{aligned}$$

which show the effect of scaling of P on the sensitivity of eigenvalues. Recall that the first order variation of each component of the eigenvalue (λ_P, μ_P) is given by

$$\begin{aligned}|\lambda(\alpha P + \Delta P) - \lambda_P| &\leq |\mu_P|\, \mathrm{cond}((\lambda_P, \mu_P), \alpha P)\|\!|\Delta P|\!\|_V + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2),\\ |\mu(\alpha P + \Delta P) - \mu_P| &\leq |\lambda_P|\mathrm{cond}((\lambda_P, \mu_P), \alpha P)\|\!|\Delta P|\!\|_V + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2),\end{aligned}$$

and the first order variation of the eigenvalue (λ_P, μ_P) is given by

$$\mathbf{d}_{\mathrm{T}}([c:s], [\lambda_P:\mu_P]) \leq \mathrm{cond}((\lambda_P, \mu_P), \alpha P)\|\!|\Delta P|\!\|_V + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2),$$

which clearly shows the effect of the scaling $P \mapsto \alpha P$ on the perturbed eigenvalues, where $(c, s) := (\lambda(\alpha P + \Delta P),\ \mu(\alpha P + \Delta P))$. Similar results hold for weighted perturbations and the weighted condition number $\mathrm{cond}_w((\lambda_P, \mu_P), P)$.

Let $\kappa((\lambda_P, \mu_P), P)$ (resp., $\kappa_w((\lambda_P, \mu_P), P)$) denote the condition number obtained from (1.5) (resp., (5.4)) by replacing the absolute perturbation $\|\!|\Delta P|\!\|_V$ (resp., $\|\!|\Delta P|\!\|_{w,V}$) with the relative perturbation $\|\!|\Delta P|\!\|_V/\|\!|P|\!\|_V$ (resp., $\|\!|\Delta P|\!\|_{w,V}/\|\!|P|\!\|_{w,V}$). Then we have

$$\begin{aligned}\kappa((\lambda_P, \mu_P), P) &= \mathrm{cond}((\lambda_P, \mu_P), P)\|\!|P|\!\|_V \text{ and}\\ \kappa_w((\lambda_P, \mu_P), P) &= \mathrm{cond}_w((\lambda_P, \mu_P), P)\|\!|P|\!\|_{w,V},\end{aligned}$$

which shows that the relative condition numbers are neutral to the scaling of P. The first order variation of each component of the eigenvalue (λ_P, μ_P) can be rewritten as

$$\begin{aligned}|\lambda(\alpha P + \Delta P) - \lambda_P| &\leq |\mu_P|\, \kappa((\lambda_P, \mu_P), P)\frac{\|\!|\Delta P|\!\|_V}{\|\!|\alpha P|\!\|_V} + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2),\\ |\mu(\alpha P + \Delta P) - \mu_P| &\leq |\lambda_P|\kappa((\lambda_P, \mu_P), P)\frac{\|\!|\Delta P|\!\|_V}{\|\!|\alpha P|\!\|_V} + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2)\end{aligned}$$

and the first order variation of the eigenvalue (λ_P, μ_P) can be rewritten as

$$\mathbf{d}_{\mathrm{T}}([c:s], [\lambda_P:\mu_P]) \leq \kappa((\lambda_P, \mu_P), P)\frac{\|\!|\Delta P|\!\|_V}{\|\!|\alpha P|\!\|_V} + \mathcal{O}((\|\!|\Delta P|\!\|_V/\alpha)^2),$$

which shows that the relative condition number $\kappa((\lambda_P, \mu_P), P)$ fails to capture the influence of the scaling $P \mapsto \alpha P$ on the first order variation of the eigenvalue (λ_P, μ_P).

6. Parameter dependent homogeneous eigenproblem. We now show that evolutions of simple eigenvalues of a parameter dependent nonlinear homogeneous eigenproblem can be deduced from the framework in section 5 and Theorem 5.3.

Let $U \subset \mathbb{C}^k$ be open and let $P : U \times \mathbb{C}^2 \longrightarrow \mathbb{C}^{n\times n}$ be such $t \mapsto P(t, \cdot, \cdot)$ is holomorphic on U and $(c, s) \mapsto P(\cdot, c, s)$ is homogeneous on $\mathbb{C}^2$. We say that $P(t, c, s)$ is regular at $\hat{t} \in U$ if $\det(P(\hat{t}, c, s)) \neq 0$ for some $(c, s) \in \mathbb{C}^2$.

THEOREM 6.1. *Let* $\hat{t} \in U$ *and* $P(\hat{t}, c, s)$ *be regular. Let* $((\lambda_0, \mu_0), x_0, y_0)$ *be a simple eigentriple of* $P(\hat{t}, c, s)$. *Then there is an open set* $\mathrm{Nbd}(\hat{t}) \subset U$ *containing* $\hat{t}$ *and a holomorphic function* $(\lambda, \mu) : \mathrm{Nbd}(\hat{t}) \longrightarrow \mathbb{C}^2$ *such that* $(\lambda(\hat{t}), \mu(\hat{t})) = (\lambda_0, \mu_0)$,

$(\lambda(t) - \lambda_0,\ \mu(t) - \mu_0) \perp (\lambda_0, \mu_0)$ and $(\lambda(t), \mu(t))$ is a simple eigenvalue of $P(t, c, s)$ for all $t \in \mathrm{Nbd}(\hat{t})$. Further, for $t \in \mathrm{Nbd}(\hat{t})$ we have

$$\begin{aligned}
\lambda(t+h) &= \lambda(t) + \mathrm{D}\lambda(t)h + \mathcal{O}(\|h\|_2^2) = \lambda(t) + \langle h,\, \nabla\lambda(t)\rangle + \mathcal{O}(\|h\|_2^2),\\
\mu(t+h) &= \mu(t) + \mathrm{D}\mu(t)h + \mathcal{O}(\|h\|_2^2) = \mu(t) + \langle h,\, \nabla\mu(t)\rangle + \mathcal{O}(\|h\|_2^2)
\end{aligned}$$

for sufficiently small $\|h\|_2$. Set $\mathbf{p}(t, c, s) := \det(P(t, c, s))$ for $t \in U$ and $(c, s) \in \mathbb{C}^2$. Then the derivatives $\mathrm{D}\lambda(t)$ and $\mathrm{D}\mu(t)$ are given by

$$(\mathrm{D}\lambda(t)h, \mathrm{D}\mu(t)h) = (-\bar{\mu}_0, \bar{\lambda}_0)\, \frac{\sum_{j=1}^{k} \mathrm{Tr}(\mathrm{adj}(P(t, \lambda(t), \mu(t)))\partial_{t_j} P(t, \lambda(t), \mu(t)))h_j}{\big(\bar{\mu}_0 \partial_c \mathbf{p}(t, \lambda(t), \mu(t)) - \bar{\lambda}_0 \partial_s \mathbf{p}(t, \lambda(t), \mu(t))\big)}$$

for all $h \in \mathbb{C}^k$, where $\partial_{t_j} P(t, \lambda(t), \mu(t))$ is the partial derivative of $P(t, c, s)$ with respect to t_j and $\partial_c \mathbf{p}(t, \lambda(t), \mu(t))$ and $\partial_s \mathbf{p}(t, \lambda(t), \mu(t))$ are the partial derivatives of $\mathbf{p}(t, c, s)$ with respect to c and s, respectively, evaluated at $(t, \lambda(t), \mu(t))$. In particular,

$$\begin{aligned}
\left(\frac{\partial\lambda(\hat{t})}{\partial t_j}, \frac{\partial\mu(\hat{t})}{\partial t_j}\right) &= (-\bar{\mu}_0, \bar{\lambda}_0)\, \frac{\mathrm{Tr}(\mathrm{adj}(P(\hat{t}, \lambda_0, \mu_0))\partial_{t_j} P(\hat{t}, \lambda_0, \mu_0))}{\big(\bar{\mu}_0 \partial_c \mathbf{p}(\hat{t}, \lambda_0, \mu_0) - \bar{\lambda}_0 \partial_s \mathbf{p}(\hat{t}, \lambda_0, \mu_0)\big)}\\
&= (-\bar{\mu}_0, \bar{\lambda}_0)\, \frac{y_0^* \partial_{t_j} P(\hat{t}, \lambda_0, \mu_0) x_0}{\big(\bar{\mu}_0 \partial_c \mathbf{p}(\hat{t}, \lambda_0, \mu_0) - \bar{\lambda}_0 \partial_s \mathbf{p}(\hat{t}, \lambda_0, \mu_0)\big)},\ j = 1, 2, \ldots, k.
\end{aligned}$$

Proof. Consider $F : U \times \mathbb{C}^2 \to \mathbb{C}^2, (t, (c, s)) \longmapsto (\mathbf{p}(t, c, s)),\ (c-\lambda_0)\bar{\lambda}_0 + (s-\mu_0)\bar{\mu}_0)$ and the variety $\mathbb{V}(F) := \{(t, (c, s)) \in U \times \mathbb{C}^2 : F(t, c, s) = 0\}$. Then $(\hat{t}, (\lambda_0, \mu_0)) \in \mathbb{V}(F)$ and the Jacobian matrix of the map $\mathbb{C}^2 \to \mathbb{C}^2, (c, s) \longmapsto F(t, c, s)$ at $(\hat{t}, \lambda_0, \mu_0)$ given by

$$\begin{bmatrix} \partial_c \mathbf{p}(\hat{t}, \lambda_0, \mu_0) & \partial_s \mathbf{p}(\hat{t}, \lambda_0, \mu_0) \\ \bar{\lambda}_0 & \bar{\mu}_0 \end{bmatrix}$$

is nonsingular. Hence by the implicit function theorem there is an open set $\mathrm{Nbd}(\hat{t}) \subset U$ such that the desired function $\mathrm{Nbd}(\hat{t}) \longrightarrow \mathbb{C}^2$, $t \longmapsto (\lambda(t), \mu(t))$ exists. The rest of the proof follows from Theorem 5.3 and is analogous to Theorem 4.1 concerning the nonhomogeneous case. ∎

Conclusion. We have presented two simple and concise eigenvector-free frameworks for the sensitivity analysis of nonhomogeneous as well as homogeneous nonlinear eigenproblems. We have shown that the eigenvector-free frameworks simplify the analysis by avoiding the use of eigenvectors. For homogeneous eigenproblems, we have shown that our framework avoids the apparatus of projective spaces and enables us to carry out analysis in the inner product space $\mathbb{C}^2$. We have shown that a simple eigenvalue evolves as a smooth function of the matrix-valued function and have determined the Fréchet derivative of the eigenvalue. Further, we have derived three equivalent representations of the condition number of a simple eigenvalue of which two are eigenvector-free. Our eigenvector-free representation of the condition number generalizes a result due to Smith for a standard eigenproblem to the case of a nonlinear eigenproblem.

REFERENCES

[1] R. Alam, *On the sensitivity analysis of eigenvalues,* Electron. J. Linear Algebra, 29 (2015), pp. 223–235.

[2] R. Alam and N. Behera, *Linearizations for rational matrix functions and Rosenbrock system polynomials*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 354–380.

[3] S. Ahmad, R. Alam, and R. Byers, *Pseudospectra, critical points and multiple eigenvalues of matrix pencils*, SIAM J. Matrix Anal Appl., 31 (2010), pp. 1915–1933.

[4] B. Adhikari, R. Alam, and D. Kressner, *Structured eigenvalue condition numbers and linearizations for matrix polynomials,* Linear Algebra Appl., 435 (2011), pp. 2193–2221.

[5] 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.

[6] H. Baumgartel, *Analytic Perturbation Theory for Matrices and Operators*, Birkhäuser-Verlag, Basel, Switzerland, 1985.

[7] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems,* ACM Trans. Math. Software, 39 (2013), pp. 7–28.

[8] C. Conca, J. Planchard, and M. Vanninathan, *Existence and location of eigenvalues for fluid-solid structures*, Comput. Methods Appl. Mech. Engrg., 77 (1989), pp. 253–291.

[9] K.-W. E. Chu, *On multiple eigenvalues of matrices depending on several parameters,* SIAM J. Numer. Anal., 27 (1990), pp. 1368–1385.

[10] J. W. Demmel, *On condition numbers and the distance to the nearest ill-posed problem,* Numer. Math., 51 (1987), pp. 251–289.

[11] J. P. Dedieu, *Condition operators, condition numbers, and condition number theorem for the generalized eigenvalue problem,* Linear Algebra Appl., 263 (1997), pp. 1–24.

[12] J. P. Dedieu and F. Tisseur, *Perturbation theory for homogeneous polynomial eigenvalue problems,* Linear Algebra Appl., 358 (2003), pp. 71–94.

[13] S. Guettel and F. Tisseur, *The nonlinear eigenvalue problem,* Acta Numer., 26 (2017), pp. 1–94.

[14] H. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials,* SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1005–1028.

[15] T.-M. Hwang, W.-W. Lin, W.-C. Wang, and W. Wang, *Numerical simulation of three dimensional pyramid quantum dot,* J. Comput. Phys., 196 (2004), pp. 208–232.

[16] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1980.

[17] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Structured polynomial eigenvalue problems: Good vibrations from good linearizations,* SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1029–1051.

[18] V. Mehrmann and C. Schröder, *Nonlinear eigenvalue and frequency response problems in industrial practice,* J. Math. Ind., 1 (2011), https://doi.org/10.1186/2190-5983-1-7.

[19] V. Mehrmann and H. Voss, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods,* GAMM-Mitt., 27 (2004), pp. 121–152.

[20] J. Planchard, *Eigenfrequencies of a tube bundle placed in a confined fluid,* Comput. Methods Appl. Mech. Engrg., 30 (1982), pp. 75–93.

[21] J. R. Rice, *A theory of condition,* SIAM J. Numer. Anal., 3 (1966), pp. 287–310.

[22] F. Rellich, *Perturbation Theory of Eigenvalue Problems,* Gordon and Breach, New York, 1969.

[23] R. A. Smith, *The condition numbers of the matrix eigenvalue problem,* Numer. Math., 10 (1967), pp. 232–240.

[24] S. I. Solovëv, *Preconditioned iterative methods for a class of nonlinear eigenvalue problems,* Linear Algebra Appl., 415 (2006), pp. 210–229.

[25] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[26] Y. Su and Z. Bai, *Solving rational eigenvalue problems via linearization*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 201–216.

[27] M. Shub and S. Smale, *Complexity of Bezout's Theorem* III: *Condition number and packing,* J. Complexity, 9 (1993), pp. 4–14.

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

[29] J. G. Sun, *Multiple eigenvalue sensitivity anaylysis,* Linear Algebra Appl., 37/138 (1990), pp. 183–211.

[30] J. G. Sun, *On condition numbers of a nondefective multiple eigenvalue,* Numer. Math., 61 (1992), pp. 265–275.

[31] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems,* Linear Algebra Appl., 309 (2000), pp. 339–361.

[32] H. Voss, *A rational spectral problem in fluid-solid vibration,* Electron. Trans. Numer. Anal., 16 (2003), pp. 93–105.
[33] H. Voss, *Iterative projection methods for computing relevant energy states of a quantum dot,* J. Comput. Phys., 217 (2006), pp. 824–833.
[34] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.
[35] J. H. Wilkinson, *Notes on matrices with a very ill-conditioned eigenproblem,* Numer. Math., 19 (1972), pp. 175–178.
[36] J. H. Wilkinson, *Sensitivity of eigenvalues,* Util. Math., 25 (1984), pp. 5–76.
[37] G. A. Watson, *Characterization of the Subdifferential of Some Matrix Norms,* Linear Algebra Appl., 170 (1992), pp. 33–44.
[38] K. Zietak, *Subdifferentials, Faces, and Dual Matrices,* Linear Algebra Appl., 185 (1993), pp. 125–141.