COMPLEXITY OF PATH–FOLLOWING METHODS FOR THE EIGENVALUE PROBLEM

DIEGO ARMENTANO

ABSTRACT. A unitarily invariant projective framework is introduced to analyze the complexity of path–following methods for the eigenvalue problem. A condition number, and its relation to the distance to ill–posedness, is given. A Newton map appropriate for this context is defined, and a version of Smale's γ-Theorem is proven. The main result of this paper bounds the complexity of path–following methods in terms of the length of the path in the condition metric.

To the memory of Jean-Pierre Dedieu and Mario Wschebor

CONTENTS

1. Introduction and main results

1.1. **Introduction and background.**

In this paper we study the complexity of path–following methods to solve the eigenvalue problem:

$$(\lambda I_n - A)v = 0, \quad v \neq 0,$$

where $A \in \mathbb{K}^{n \times n}$, $v \in \mathbb{K}^n$, and $\lambda \in \mathbb{K}$. (The set $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$, and $n \geq 2$.) Here, the complexity of an algorithm should be understood as the study of the number of arithmetic operations required to pass from the input to the output.

Many algorithms have been used to solve the eigenvalue problem. A naive approach to solve this problem would be to compute the characteristic polynomial $\chi_A(z) = \det(zI_n - A)$ of A and then compute (i.e., find approximations of) its zeros. From these zeros, a correspondent eigenvector can be computed. Unfortunately, in some cases, polynomials $\chi_A(z)$ arising in this way may be ill–conditioned even when the original matrix A isn't, and therefore, the numerical stability may be destroyed under this process. In practice, algorithms for solving the eigenvalue problem have avoided this naive approach, and moreover, the tradition in numerical analysis in the last 20 years seems to go in the opposite direction. That is, in order to solve a polynomial in one variable the standard procedure is to search for the eigenvalues of the associated companion matrix; see Trefethen–Bau [34] for example.

Most of the algorithms used in practice for solving the eigenvalue problem may be divided into two classes: QR methods (including Hessenberg reduction, single or double shift strategy, deflation), and Krylov subspace methods; see Wilkinson [36], Golub–Van Loan [19], Stewart [32], or Watkins [35] for details. These algorithms are known to be stable but, surprisingly, the complexity in not well–understood; indeed, for each of these methods one of the following fundamental questions is still open:

(1) For which class of matrices there is guaranteed convergence?
(2) Is there a small bound on the average of number of steps, in a given probabilistic model on the set of inputs, to obtain a given accuracy on the output?

The two following examples show that such questions are particularly difficult:

- Rayleigh quotient iteration fails for a non–empty open set of matrices; see Batterson-Smillie [3, 4].
- The unshifted QR algorithm is convergent for almost every complex matrix. However, even for the simple choice of *Gaussian Orthogonal Ensemble* as a

probabilistic model, question (2) remains unanswered; see Deift [16]. (See also Pfrang et al. [26] for some statistics about QR and Toda algorithms.)

While in practice many numerical methods are available for computing the eigenvalues and eigenvectors, until now, a numerically stable algorithm that provides satisfactory answers to questions (1) and (2) is not available. On the one hand we have algorithms for which we can prove low complexity bounds but appears to be unstable in practice; see Pan [25] or Renegar [27]. On the other hand, we have algorithms which are stable (and even efficient) in practice but for which we cannot prove satisfactory complexity bounds.

The main theoretical open question for the complexity of the eigenvalue problem is to provide an algorithm to solve this problem which is numerically stable and works in average polynomial–time. The present paper may be considered as a step forward to achieve this goal; the underlying algorithmic approach are the so called homotopy methods.

In the last three decades path–following methods, or homotopy methods, have been applied to solve the eigenvalue problem. The advantage of using homotopy methods for the eigenvalue problem lies in the following facts.

(1) Path–following methods are numerically stable almost by definition. Being more precise, in order to follow a path of (problems,solutions) it is enough to compute a sequence of pairs such that each pair is a mere approximation of the path with some prescribed error; ; see, for instance, the recent analysis in Briquel et al. [9] for the case of complex polynomial systems.

(2) The recent success of homotopy methods for attacking Smale's 17th problem (see Beltrán–Pardo [6], Bürgisser–Cucker [10], and more recently Armentano–Shub [1]) brings some hope to reach the main goal for the complexity of the eigenvalue problem mentioned above.

(3) Given a matrix, the output of the homotopy method is a good approximation of some eigenvalue (and its corresponding eigenvector) of the given matrix. In this way we are avoiding to deal with the problem of computing an eigenvector from some approximation of the eigenvalue of the matrix (which will carry an extra cost and accuracy issues; cf. Remark 1.6).

The homotopy method for the eigenvalue problem was first studied by Chu [12], when A is a real symmetric matrix. In Li–Sauer–York [22] and Li–Sauer [21] homotopy methods were given for *deficient* polynomial systems, and in particular, the general eigenvalue problem is considered; see Li [20] for a general discussion. Since then, a substantial amount of papers dealing with homotopy methods for solving the eigenvalue problem have been written; see Lui–Keller–Kwok [23] and references therein. Even though these methods can achieve spectacular results in practice (even faster than QR in some cases) the complexity is still an open problem.

In this paper we consider the eigenvalue problem as a bilinear polynomial system of equations and we study the complexity of homotopy methods to solve it. Briefly, homotopy methods can be described as follows. The system $(\lambda I_n - A)v = 0$, $v \neq 0$, is the endpoint of a path of problems

$$(\lambda(t)I_n - A(t))v(t) = 0, \ v(t) \neq 0, \ 0 \leq t \leq 1,$$

with $(A(1), \lambda(1), v(1)) = (A, \lambda, v)$. Starting from a known triple $(A(0), \lambda(0), v(0))$ we "follow" this path to reach the target system $(\lambda I_n - A)v = 0$. The algorithmic

way to do so is to construct a finite number of triples

$$(A_k, \lambda_k, v_k), \quad 0 \le k \le K,$$

with $A_k = A(t_k)$, and $0 = t_0 < t_1 < \cdots < t_K = 1$, and where (λ_k, v_k) are approximations of $(\lambda(t_k), v(t_k))$. The complexity of the algorithm just described (defined more precisely below) is reduced to the number K of steps sufficient to validate the approximation, since the arithmetic cost of each iteration is linear in n.

The main result of this paper is to relate K with a geometric invariant, namely, the **condition length** of the path. In the next paragraphs of this section we give a succinct description of the main definitions in order to state the main results, pointing on our way to the location in this paper when notions are dealt at greater length.

1.2. **Solution Variety.** We begin with the geometric framework of our problem. Since the eigenvalue problem is homogeneous in $v \in \mathbb{K}^n$ and in $(A, \lambda) \in \mathbb{K}^{n \times n} \times \mathbb{K}$, we define the *solution variety* as

$$\mathcal{V} := \left\{ (A, \lambda, v) \in \mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}\left(\mathbb{K}^n\right) : \ (\lambda I_n - A)v = 0 \right\},$$

where $\mathbb{P}(\mathbb{E})$ denotes the projective space associated with the vector space $\mathbb{E}$. We speak interchangeably of a nonzero vector and its corresponding class in the projective space.

The solution variety $\mathcal{V}$ plays a crucial role in this paper. It is a connected smooth manifold of the same dimension as $\mathbb{P}(\mathbb{K}^{n \times n})$. It is possible to define a natural projection

$$\pi : \mathcal{V} \to \mathbb{P}(\mathbb{K}^{n \times n}) \quad \text{given by} \quad \pi(A, \lambda, v) = A.$$

This projection, for almost every $(A, \lambda, v) \in \mathcal{V}$, has a branch of the inverse image of π taking $A \in \mathbb{P}(\mathbb{K}^{n \times n})$ to $(A, \lambda, v) \in \mathcal{V}$. This branch of π^{-1} is usually called the *input–output map*. In this fashion, we may think of $\mathbb{P}(\mathbb{K}^{n \times n})$ as the space of inputs and $\mathcal{V}$ as the space of outputs. (Section 2.2 provides a detailed exposition of these facts.)

1.3. **Newton's method.** Given a nonzero matrix $A \in \mathbb{K}^{n \times n}$, we define the evaluation map $F_A : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$, by

$$F_A(\lambda, v) := (\lambda I_n - A)v.$$

The *Newton map* associated to F_A, is the map N_A on $\mathbb{K} \times (\mathbb{K}^n \setminus \{0\})$ given by

$$N_A(\lambda, v) := (\lambda, v) - \left(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}\right)^{-1} F_A(\lambda, v),$$

defined for every (λ, v) such that $DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}$ is invertible, where $DF_A(\lambda, v)$ denotes the derivative of F_A at (λ, v). Here $v^\perp$ is the Hermitian complement of v in $\mathbb{K}^n$.

In Section 4 we show that the map N_A is well–defined provided that $\Pi_{v^\perp}(\lambda I_n - A)\big|_{v^\perp}$ is invertible (where $\Pi_{v^\perp}$ denotes the orthogonal projection of $\mathbb{K}^n$ onto $v^\perp$). If this is the case, then the map N_A is given by $N_A(\lambda, v) = (\lambda - \dot{\lambda}, v - \dot{v})$, where

$$(1.1) \quad \dot{v} = \left(\Pi_{v^\perp}(\lambda I_n - A)\big|_{v^\perp}\right)^{-1} \Pi_{v^\perp}(\lambda I_n - A)v, \quad \dot{\lambda} = \frac{\langle (\lambda I_n - A)(v - \dot{v}), v \rangle}{\langle v, v \rangle}.$$

(Here $\langle \cdot, \cdot \rangle$ denotes the canonical Hermitian product on $\mathbb{K}^n$.)

Let $A \in \mathbb{K}^{n \times n}$ be a nonzero matrix, and let $(\lambda_0, v_0) \in \mathbb{K} \times \mathbb{K}^n$, $v_0 \neq 0$. We say that the triple (A, λ_0, v_0) is an *approximate solution* of the eigenvalue problem $(A, \lambda, v) \in \mathcal{V}$, when the sequence $(A, N_A^k(\lambda_0, v_0))$, $k = 0, 1, 2, \ldots$ converges immediately quadratically to the eigentriple $(A, \lambda, v) \in \mathcal{V}$, that is, if the given sequence satisfies

$$d_{\mathbb{P}^2}\left((A, N_A^k(\lambda_0, v_0)), (A, \lambda, v)\right) \leq \left(\frac{1}{2}\right)^{2^k - 1} d_{\mathbb{P}^2}\left((A, \lambda_0, v_0), (A, \lambda, v)\right),$$

for all positive integers k. Here $d_{\mathbb{P}^2}(\cdot, \cdot)$ is the induced Riemannian distance on $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$; see Section 2.1.1.

Remark 1.1. From the expression (1.1), it is easily seen that, if $N_A(\lambda, v) = (\lambda', v')$ then $N_{\alpha A}(\alpha \lambda, \beta v) = (\alpha \lambda', \beta v')$, for every nonzero scalars α and β. Hence, the sequence defined above with starting point (A, λ_0, v_0) and that starting at $(\alpha A, \alpha \lambda_0, \beta v_0)$, define the same sequence on $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$. Thus the property of being an approximate solution is well–defined on $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$.

Remark 1.2. From the last remark we conclude that the Newton map N_A induces a map from $\mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$ into itself (defined almost everywhere).

1.4. **The predictor–corrector algorithm.** Let $\Gamma(t) = (A(t), \lambda(t), v(t))$, $0 \leq t \leq 1$, be a path of eigentriples in $(\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$, i.e., $\Gamma \subset \mathcal{V}$. To approximate Γ by a finite sequence we use the following predictor–corrector strategy: given a mesh $0 = t_0 < t_1 < \cdots < t_K = 1$ and a pair $(\lambda_0, v_0) \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$, we define

$$(\lambda_{k+1}, v_{k+1}) := N_{A(t_{k+1})}(\lambda_k, v_k), \quad 0 \leq k \leq K - 1,$$

(in case it is defined). We say that the sequence $(A(t_k), \lambda_k, v_k)$, $0 \leq k \leq K$, *approximates* the path $\Gamma(t)$, $0 \leq t \leq 1$, when for any $k = 0, \ldots, K$, the triple $(A(t_k), \lambda_k, v_k)$ is an approximate solution of the eigentriple $\Gamma(t_k) \in \mathcal{V}$. In that case we define the *complexity of the sequence* by K.

1.5. **Condition of a triple and condition length.** Let $\mathcal{W} \subset \mathcal{V}$ be the set of *well–posed* problems, that is, the set of triples $(A, \lambda, v) \in \mathcal{V}$ such that the input–output map mentioned in Section 1.2, taking $A \in \mathbb{P}(\mathbb{K}^{n \times n})$ to $(A, \lambda, v) \in \mathcal{V}$, is locally defined. The set $\mathcal{W}$ is the open set of triples $(A, \lambda, v) \in \mathcal{V}$ such that λ is a simple eigenvalue; see Section 2 for further details. Let $\Sigma' := \mathcal{V} \setminus \mathcal{W}$ be the *ill–posed variety*.

When $(A, \lambda, v) \in \mathcal{W}$, the operator $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is invertible; see Section 2. The *condition number* of $(A, \lambda, v) \in \mathcal{W}$ is defined by

$$\mu(A, \lambda, v) := \max\left\{1, \|A\|_F \left\|(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\right\|\right\},$$

where $\|\cdot\|_F$ and $\|\cdot\|$ are the Frobenius and operator norms in the space of matrices; see Section 3.

Let $\Gamma(t)$, $0 \leq t \leq 1$, be an absolutely continuous path in $\mathcal{W}$. We define its *condition–length* as

$$\ell_\mu(\Gamma) := \int_0^1 \left\|\dot{\Gamma}(t)\right\|_{\Gamma(t)} \mu\left(\Gamma(t)\right) \, dt,$$

where $\left\|\dot{\Gamma}(t)\right\|_{\Gamma(t)}$ is the norm of $\dot{\Gamma}(t)$ in the unitarily invariant Riemannian structure on $\mathcal{V}$; see Section 2.1.1 and Section 5.1.

1.6. **Main results.** Recall that $\mathcal{V} \subset \mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) \times \mathbb{P}(\mathbb{K}^n)$. Let $\pi_2 : \mathcal{V} \to \mathbb{P}(\mathbb{K}^n)$ be the restriction to $\mathcal{V}$ of the canonical projection $(A, \lambda, v) \mapsto v$. Let $\mathcal{V}_v \subset \mathcal{V}$ be the inverse image of v under π_2.

Theorem 1 (Condition Number Theorem). *For $(A, \lambda, v) \in \mathcal{W}$, we get*

$$\mu(A, \lambda, v) \leq \max \left\{ 1, \frac{1}{\left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right)^{1/2}} \frac{1}{\sin(d_{\mathbb{P}^2}\left((A, \lambda, v), \Sigma' \cap \mathcal{V}_v\right))} \right\}.$$

This theorem is a version of the Condition Number Theorem which relates the condition number to the distance to ill–posed problems. Its proof is given in Section 3.6.

The main theorem concerning the convergence of Newton's iteration is the following.

Theorem 2 (Approximate Solution Theorem). *There is a universal constant $c_0 > 0$ with the following property. Let $A \in \mathbb{K}^{n \times n}$ be a nonzero matrix, and let (λ, v), (λ_0, v_0) in $\mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. If $(A, \lambda, v) \in \mathcal{W}$ and*

$$d_{\mathbb{P}^2}\left((A, \lambda_0, v_0), (A, \lambda, v)\right) < \frac{c_0}{\mu(A, \lambda, v)},$$

then, (A, λ_0, v_0) is an approximate solution of (A, λ, v). (One may choose $c_0 = 0.0739$.)

Theorem 2 is a version of the so called *Smale's γ-theorem* (see Blum et al. [8]), which gives the size of the basin of attraction of Newton's method. Different versions of Smale's γ-theorem for the symmetric eigenvalue problem and for the generalized eigenvalue problem are given in Dedieu [13] and Dedieu–Shub [15] respectively.

Theorem 2 is the main ingredient to prove complexity results for path–following methods. The proof of this theorem is included in Section 4.

Following these lines our main result is the following.

Theorem 3 (Main Theorem). *There is a universal constant $C > 0$ such that for any absolutely continuous path $\Gamma(t) = (A(t), \lambda(t), v(t))$ in $\mathcal{W}$, $0 \leq t \leq 1$, (with $\ell_\mu(\Gamma) < \infty$), there exists a sequence $(A(t_0), \lambda_0, v_0), \ldots, (A(t_K), \lambda_K, v_K)$ such that, $t_0 = 0$, $t_K = 1$, the triple $(A(t_k), \lambda_k, v_k)$ is an approximation of $\Gamma(t_k)$, $0 \leq k \leq K$, and*

$$K \leq C \, \ell_\mu(\Gamma) + 1.$$

(One may choose $C = 100$.)

The proof of Theorem 3 is given in Section 5.

Remark 1.3. The selection of a good starting triple $(A(0), \lambda_0, v_0)$ is an important issue that is beyond the scope of this paper. Nevertheless, we would like to conclude this section suggesting some simple candidates.

  (i) *Rank one matrices*: If we are thinking in a fixed family of starting points, a natural requirement would be that the condition number on this set is small. The family of triples $(vv^*, \|v\|^2, v) \in \mathcal{V}$, for $v \in \mathbb{K}^n \setminus \{0\}$, satisfy this requirement. More precisely, it is easily check that this family is a subset of the set of triples where the condition number reaches its minimum value. Note that 0 is a multiple eigenvalue of vv^*. (This example is a version of

an example conjectured by Shub–Smale [31] to be a good starting point for linear homotopy in the polynomial system case.)

(ii) *Roots of unity*: Another example is to consider $D = \mathrm{Diag}(1, \zeta, \ldots, \zeta^{n-1})$, where ζ is the nth primitive root of the unity. Then all the associated eigentriples $(D, 1, e_1), (D, \zeta, e_2), \ldots (D, \zeta^{n-1}, e_n)$ are well–posed, where $e_1, \ldots, e_n$ denotes the canonical basis of $\mathbb{K}^n$. Their condition numbers are constant equal to $\sqrt{n}/(2\sin(\pi/n))$. (This candidate is a version of the system of polynomials considered in Bürgisser–Cucker [10] as starting point.)

(iii) *Projection on a subspace*: A different approach is to consider the starting point as a function of the input. There are many different strategies to pursue. For instance, for a given matrix $A = (a_{ij})_{i,j=1,\ldots,n} \in \mathbb{K}^{n \times n}$ let

$$
A(0) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.
$$

Then one can consider $(A(0), a_{11}, e_1) \in \mathcal{V}$ as starting point. Note that $A(0)$ is the orthogonal projection of the matrix A into the subspace of matrices such that e_1 is an eigenvector. This kind of procedures have many advantages compared with the preceding examples. However, the main drawback is that in some cases we cannot assure that $(A(0), a_{11}, e_1)$ is well–posed. (This particular selection of starting point is a version of a starting point considered in Armentano–Shub [1] for the polynomial system case.)

1.7. **Comments.** In their seminal paper [29], Shub and Smale relate, in the context of polynomial system solving, the complexity K to three ingredients: the degree of the considered system, the length of the path $\Gamma(t)$, $0 \leq t \leq 1$, and the condition number of the path. Precisely, they obtain the complexity

$$(1.2) \qquad\qquad K \leq CD^{3/2}\ell(\Gamma)\mu(\Gamma)^2,$$

where C is a universal constant, D is the maximum of the degrees of the underlying system of polynomials, $\ell(\Gamma)$ is the length of Γ in the associated Riemannian structure, and $\mu(\Gamma) = \sup_{a \leq t \leq b} \mu(\Gamma(t))$.

In Shub [28] the complexity K of path–following methods for the polynomial system solving problem is analyzed in terms of the condition length of the path.

In the context of polynomial system solving, the eigenvalue problem $Av = \lambda v$ for a $n \times n$ matrix A, with unknowns λ and v, may be considered as a quadratic system of equations. When $\mathbb{K} = \mathbb{C}$, by Bézout's theorem, after homogenization, one expects 2^n roots. However, this system has at most n isolated roots. Therefore the eigenvalue problem as a quadratic system belongs to the subset of ill–posed problems, and hence [29] and [28] do not apply. For this reason, in order to analyze the complexity of the eigenvalue problem, a different framework is required.

In Shub–Smale [30] a unitarily invariant geometric framework is introduced to study the eigenvalue problem, where the input space is the space of matrices $\mathbb{C}^{n \times n}$, and the space of outputs is $\mathbb{C} \times \mathbb{P}(\mathbb{C}^n)$. In Dedieu–Shub [15] a complexity bound of type (1.2) is obtained for general multi–homogeneous analytic functions, which applies in particular for the generalized eigenvalue problem.

In this paper we pursue a different approach, considering the eigenvalue problem as a bilinear problem; see Section 2.5. The main difference of this projective framework compared to the frameworks mentioned above is that the complexity of the eigenvalue problem is not only unitarily invariant but invariant under the scaling of the matrix as well, and hence the natural space for the input is the projective space $\mathbb{P}(\mathbb{K}^{n \times n})$. This approach was greatly inspired by Michael Shub.

Remark 1.4. As it was mentioned before, there is a natural connection between the eigenvalue problem and the problem of finding a root of a polynomial in one variable. Given a $n \times n$ matrix A, the roots of the characteristic polynomial $\chi_A(z)$ are exactly the eigenvalues of A. Therefore in the case $\mathbb{K} = \mathbb{C}$, one may consider this approach to analyze the complexity of the eigenvalue problem, where proven average polynomial–time complexity –for homotopy methods– are given with respect to some natural Gaussian measure on the space of polynomials; see Armentano–Shub [1]. However, the push–forward measure (induced by the map $A \mapsto \chi_A$) of the natural Gaussian measure on $\mathbb{C}^{n \times n}$ is different from the restriction of the Gaussian measure, mentioned above, to the space of monic polynomials. Hence, it is not clear how the complexity of the eigenvalue problem is related to the complexity of solving polynomials in one variable. In addition, this approach has some important drawbacks:

- The complexity analysis of solving the characteristic polynomial $\chi_A(z)$ is not invariant under scaling of A. More precisely, the complexity theory for polynomial systems mentioned above, applied to non–homogeneous polynomials in one variable, is not invariant under scaling of the roots. This is in contrast to the natural scaling of the eigenvalue and the matrix.
- The complexity of finding eigenvectors is not considered under this approach.
- The map $A \mapsto \chi_A$ magnifies the condition number in some cases and hence the complexity may growth.

On the other hand, the magnifying effect of the condition number mentioned above is not true in general; see example (e) in Section 3.4. In particular, if one is interested only in the complexity of finding eigenvalues (and not eigenvectors), there exist a possibility that this approach may improve the complexity in some cases (in contrast to the belief prevailing among numerical analysts; cf. Trefethen-Bau [34]). It is hoped that the present paper may help to analyze this issue rigorously and may provide the elements to give theoretical proofs of the experimental tendencies of numerical analysts, in particular, the connection of polynomials in one variable and the companion matrices mentioned at the beginning of this paper. (See Beltrán–Shub [7] for a similar discussion and some interesting questions.)

Remark 1.5. A drawback of homotopy methods is that it is not always possible to lift a path of matrices $A(t)$, $0 \leq t \leq 1$, to the solution variety $\mathcal{V}$. However, this is the only requirement to analyze the complexity. One is able to lift this path when the "continued" eigenvalue of the homotopy $A(t)$ remains a simple eigenvalue of $A(t)$ for all $t \in [0, 1]$. (A completely different situation can occur when we restrict ourselves to the case $\mathbb{K} = \mathbb{R}$, since the projection $\pi : \mathcal{V} \to \mathbb{P}(\mathbb{R}^{n \times n})$ is not even surjective; see Section 2.4 for discussion on the connectivity of $\mathcal{W}$ for this case.) The *discriminant variety* $\Sigma := \pi(\Sigma')$ is an algebraic variety of $\mathbb{P}(\mathbb{K}^{n \times n})$ (see Remark 2.6), hence when

$\mathbb{K} = \mathbb{C}$ it has complex codimension one, and thus almost all straight line paths (at least) can be lifted.

Remark 1.6. Another possible strength of this paper is that one allows A to be non–normal. Moreover, one of the important arguments for pseudo–spectral techniques in numerical linear algebra has been that when applying the QR algorithm one will only get an approximation to the Schur form of the matrix, and hence one is solving a perturbed problem. This in turn suggests that the best error bound one can get is from bounding the perturbation and then one has to resort to pseudo–spectral theory in order to get a precise and reliable bound. As the pseudo–spectrum can be rather wild one may end up with very crude and rather poor error bounds. This is not the case in our Theorem 2. Thus, if $\ell_\mu(\Gamma)$ can be estimated (and it must not be too large) one may have a good alternative to the QR method for non–normal problems, and the method would come with nice error bounds.

Remark 1.7. Armentano [2] addressed, for the case $\mathbb{K} = \mathbb{C}$, the problem of finding short paths for the condition length. It is proved that for every problem $(A, \lambda, v) \in \mathcal{W}$ there exist a path Γ in $\mathcal{W}$ joining (A, λ, v) with $(e_1 e_1^*, 1, e_1)$ (where e_1 is the first element of the canonical basis of $\mathbb{C}^n$) such that

$$\ell_\mu(\Gamma) \leq C \sqrt{n} \left(C' + \log(\mu(A, \lambda, v)) \right),$$

for some universal constants C and C'. This type of results shed some light on the contribution of our paper. More precisely, this result combined with Theorem 3 means that it may be possible to solve the eigenvalue problem with a small complexity, precisely, logarithmic in the condition number of the ending triple. This motivates the study of short paths or geodesics in the condition metric. Any result on this matter is encouraging and a real challenge.

Remark 1.8. Theorem 3 states the existence of a sequence which approximates $\Gamma \subset \mathcal{W}$ with the given complexity. The sequence is described in the proof of this theorem but is not constructive. Our next objective is to transfer these theoretical considerations into a practical algorithm. This issue will be considered in another paper, and the construction of the path–following algorithm should be analogous to the constructions given by Beltrán [5], Bürgisser–Cucker [10], or Dedieu–Malajovich–Shub [14], for the polynomial system case.

Remark 1.9. For the purpose of this paper, we only require that the derivative of the path Γ is almost everywhere defined, and also that the length and the condition length of Γ are defined. For this reason we ask Γ to be an absolutely continuous path. This is in contrast to Shub [28] where the requirement for the path Γ is to be a C^1 path. The C^1 hypothesis seems to be more natural for the implementation of the algorithm.

Note: Throughout this paper we work with $\mathbb{K} = \mathbb{C}$. However most definitions and results can be extended immediately to the case $\mathbb{K} = \mathbb{R}$. Whenever it is necessary we shall state the difference.

algebraic properties of the solution variety. I gratefully acknowledge a large number of comments and remarks by the anonymous referees which lead me to enhance the presentation of this paper. (In particular Remark 1.6 and Remark 2.6 are due to them.) Part of this work was elaborated during the special semester on Foundations of Computational Mathematics in 2009, in the Fields Institute at Toronto. Thanks to the Fields Institute for its hospitality and financial support.

This paper is part of my PhD thesis, written under the supervision of Jean–Pierre Dedieu (Institut de Mathèmatiques de Toulouse, France) and Mario Wschebor (Universidad de la República, Uruguay).

Dedications. This paper is in the memory of my beloved friends and advisors Jean-Pierre Dedieu and Mario Wschebor for their inspiration and dedication. Jean-Pierre and Mario passed away, before the defense of my PhD thesis and during the submission of this paper, after a long battle with cancer. To them all my gratitude.

2. Solution Variety

2.1. **Introduction.** We start this section defining the canonical metric structures. Following this we define the solution variety $\mathcal{V}$ and the varieties Σ' and Σ, and we study some basic geometrical, topological and algebraic properties of these varieties.

2.1.1. *Canonical metric structures.* The space $\mathbb{K}^n$ is equipped with the canonical Hermitian inner product $\langle \cdot, \cdot \rangle$. The space $\mathbb{K}^{n \times n}$ is equipped with the Frobenius Hermitian inner product

$$\langle A, B \rangle_F := \mathrm{trace}\,(B^* A),$$

where B^* denotes the adjoint of B.

In general, if $\mathbb{E}$ is a finite dimensional vector space over $\mathbb{K}$ with the Hermitian inner product $\langle \cdot, \cdot \rangle$, we can define an Hermitian structure on $\mathbb{P}(\mathbb{E})$ in the following way: for $x \in \mathbb{E}$,

$$\langle w, w' \rangle_x := \frac{\langle w, w' \rangle}{\|x\|^2},$$

for all w, w' in the Hermitian complement $x^\perp$ of x in $\mathbb{E}$, which is a natural representation of the tangent space $T_x \mathbb{P}(\mathbb{E})$. Let $d_{\mathbb{P}}(x, y)$ be the angle between the vectors x and y.

In this way, the space $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) \times \mathbb{P}(\mathbb{K}^n)$ inherits the Hermitian product structure

$$(2.1) \qquad \langle (\dot{A}, \dot{\lambda}, \dot{v}), (\dot{B}, \dot{\eta}, \dot{w}) \rangle_{(A, \lambda, v)} := \langle (\dot{A}, \dot{\lambda}), (\dot{B}, \dot{\eta}) \rangle_{(A, \lambda)} + \langle \dot{v}, \dot{w} \rangle_v,$$

for all $(\dot{A}, \dot{\lambda}, \dot{v})$, $(\dot{B}, \dot{\eta}, \dot{w}) \in (A, \lambda)^\perp \times v^\perp$.

We denote by $d_{\mathbb{P}^2}(\cdot, \cdot)$ the induced Riemannian distance on $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) \times \mathbb{P}(\mathbb{K}^n)$.

Throughout this paper we denote by the same symbol $d_{\mathbb{P}}$ distances on $\mathbb{P}(\mathbb{K}^n)$, $\mathbb{P}(\mathbb{K}^{n \times n})$ and $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K})$.

2.2. The Varieties $\mathcal{V}$, Σ' and Σ.

Definition 2.1. We define the *solution variety* as
$$\mathcal{V} := \left\{ (A, \lambda, v) \in \mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}\left(\mathbb{K}^n\right) : \ (\lambda I_n - A)v = 0 \right\}.$$

The solution variety $\mathcal{V}$ is the set of equivalence classes of the set of solutions of $F = 0$, where F is the bilinear system given by

$$(2.2) \quad F : (\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times (\mathbb{K}^n \setminus \{0\}) \to \mathbb{K}^n, \quad F(A, \lambda, v) = (\lambda I_n - A)v.$$

Note that $F(\alpha A, \alpha \lambda, \beta v) = \alpha \beta F(A, \lambda, v)$, for all nonzero scalars α and β. Therefore $\mathcal{V}$ is an algebraic subvariety of the product $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$. Moreover, since 0 is a regular value of F we conclude that $\mathcal{V}$ is also a smooth submanifold of $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$. Its dimension over $\mathbb{K}$ is given by

$$\dim \mathcal{V} = \dim(\mathbb{K}^{n \times n} \times \mathbb{K} \times \mathbb{K}^n) - n - 2 = n^2 - 1.$$

Thus we have concluded the following result.

Proposition 2.2. *The solution variety $\mathcal{V}$ is a smooth submanifold of $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$ with the same dimension as $\mathbb{P}(\mathbb{K}^{n \times n})$. The tangent space $T_{(A, \lambda, v)}\mathcal{V}$ to $\mathcal{V}$ at (A, λ, v) is the set of triples $(\dot{A}, \dot{\lambda}, \dot{v}) \in \mathbb{K}^{n \times n} \times \mathbb{K} \times \mathbb{K}^n$, satisfying*

$$(2.3) \quad (\dot{\lambda} I_n - \dot{A})v + (\lambda I_n - A)\dot{v} = 0; \quad \langle \dot{A}, A \rangle_F + \dot{\lambda}\overline{\lambda} = 0; \quad \langle \dot{v}, v \rangle = 0. \quad \square$$

Remark 2.3. The solution variety $\mathcal{V}$ inherits the Hermitian structure from $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \times \mathbb{P}(\mathbb{K}^n)$ defined in (2.1).
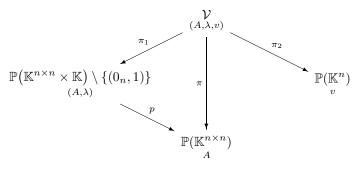
We denote by π_1 and π_2 the restriction to $\mathcal{V}$ of the canonical projections onto $\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right)$ and $\mathbb{P}(\mathbb{K}^n)$ respectively.

Note that $\pi_1(\mathcal{V}) \subset \mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right)$ does not include the pair $(0_n, 1)$. Therefore we can define the map

$$\pi : \mathcal{V} \to \mathbb{P}(\mathbb{K}^{n \times n}), \quad \pi := p \circ \pi_1,$$

where p is the canonical projection

$$(2.4) \quad p : (\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \setminus \{(0_n, 1)\}) \to \mathbb{P}(\mathbb{K}^{n \times n}), \quad p(A, \lambda) = A;$$

see the following diagram.

$$
\begin{array}{ccc}
& \mathcal{V} & \\
& (A, \lambda, v) & \\
{}^{\pi_1}\swarrow & \downarrow {\scriptstyle \pi} & \searrow {}^{\pi_2} \\
\mathbb{P}\left(\mathbb{K}^{n \times n} \times \mathbb{K}\right) \setminus \{(0_n, 1)\} & & \mathbb{P}(\mathbb{K}^n) \\
(A, \lambda) & & v \\
& \searrow {}^{p} & \\
& \mathbb{P}(\mathbb{K}^{n \times n}) & \\
& A &
\end{array}
$$

The derivative

$$(2.5) \qquad\qquad D\pi(A, \lambda, v) : T_{(A, \lambda, v)}\mathcal{V} \to T_A \mathbb{P}(\mathbb{K}^{n \times n}),$$

is a linear operator between spaces of equal dimension.

Definition 2.4. We say that the triple $(A, \lambda, v) \in \mathcal{V}$ is *well–posed* when $D\pi(A, \lambda, v)$ is an isomorphism. Let $\mathcal{W}$ be the set of well–posed triples, and $\Sigma' := \mathcal{V} \setminus \mathcal{W}$ be the *ill–posed variety*. Let $\Sigma = \pi(\Sigma') \subset \mathbb{P}(\mathbb{K}^{n \times n})$ be the *discriminant variety*, i.e., the subset of ill–posed inputs.

Lemma 2.5. *The ill–posed variety Σ' is the set of triples $(A, \lambda, v) \in \mathcal{V}$ such that λ is an eigenvalue of A of algebraic multiplicity ≥ 2.*

Proof. The linear operator (2.5) is given by

$$(2.6) \qquad D\pi(A, \lambda, v)(\dot{A}, \dot{\lambda}, \dot{v}) = \dot{A} + \frac{\dot{\lambda}\,\overline{\lambda}}{\|A\|_F^2} A, \qquad (\dot{A}, \dot{\lambda}, \dot{v}) \in T_{(A,\lambda,v)}\mathcal{V}.$$

According to (2.3), a non–trivial triple in the kernel of $D\pi(A, \lambda, v)$ has the form $(\frac{-\dot{\lambda}\,\overline{\lambda}}{\|A\|_F^2} A, \dot{\lambda}, \dot{v})$, where $\langle \dot{v}, v \rangle = 0$, $\dot{v} \neq 0$, and

$$\dot{\lambda}\left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right) v + (\lambda I_n - A)\dot{v} = 0.$$

Then, $\mathrm{rank}[(\lambda I_n - A)^2] < n - 1$, namely, λ is not a simple eigenvalue of A. Reciprocally, if the algebraic multiplicity of λ is ≥ 2, then there exists $0 \neq w \in v^\perp$ such that $(\lambda I_n - A)w = \alpha v$, for some $\alpha \in \mathbb{K}$. Then, $(\dot{A}, \dot{\lambda}, \dot{v})$ given by $\dot{A} = \frac{\alpha\overline{\lambda}}{\|A\|_F^2} A$, $\dot{\lambda} = -\alpha$ and $\dot{v} = (1 + \frac{|\lambda|^2}{\|A\|_F^2})w$, is a non–trivial triple belonging to $\ker D\pi(A, \lambda, v)$, and therefore $(A, \lambda, v) \in \Sigma'$. $\qquad\square$

Remark 2.6. From Lemma 2.5 we conclude that Σ' is an algebraic subvariety of $\mathcal{V}$. Since Σ is the set of matrices A such that the resultant of $\chi_A(x)$ and $\chi_A(x)'$ is zero, thus it is an algebraic variety of $\mathbb{P}(\mathbb{K}^{n \times n})$; see for example Blum et al. [8].

Lemma 2.7. *One has,*

$$\mathcal{W} = \{(A, \lambda, v) \in \mathcal{V} : \Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp} \text{ is invertible}\}.$$

Proof. Let $(A, \lambda, v) \in \mathcal{W}$. Then, from Lemma 2.5, $(\lambda I_n - A)v = 0$ and the rank of $(\lambda I_n - A)^2$ is $n - 1$. That is, $(\lambda I_n - A)v = 0$ and the range of the linear operator $(\lambda I_n - A)|_{v^\perp}$, which is a $(n-1)$-dimensional subspace of $\mathbb{K}^n$, does not contain the vector v. Thus $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is an invertible operator from $v^\perp$ into itself. Reciprocally, by contradiction, if we assume that $(A, \lambda, v) \in \mathcal{V}$ and λ is not a simple eigenvalue, then there exists $w \in v^\perp$, $w \neq 0$, such that $(\lambda I_n - A)w = \alpha v$ (for some $\alpha \in \mathbb{K}$). Then, the linear operator $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ has non–trivial kernel. $\qquad\square$

2.3. Unitarily invariance. Let $\mathbb{U}_n(\mathbb{K})$ stands for the unitary group when $\mathbb{K} = \mathbb{C}$ or the orthogonal group when $\mathbb{K} = \mathbb{R}$. The group $\mathbb{U}_n(\mathbb{K})$ acts on $\mathbb{P}(\mathbb{K}^n)$ in the natural way. In addition, $\mathbb{U}_n(\mathbb{K})$ acts on $\mathbb{P}(\mathbb{K}^{n \times n})$ by conjugation (i.e., $U \cdot A = UAU^{-1}$), and acts on $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K})$ by $U \cdot (A, \lambda) = (UAU^{-1}, \lambda)$. These actions define an action on the product space $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) \times \mathbb{P}(\mathbb{K}^n)$, namely,

$$(2.7) \qquad\qquad U \cdot (A, \lambda, v) \mapsto (UAU^{-1}, \lambda, Uv), \quad U \in \mathbb{U}_n(\mathbb{K}).$$

If $(A, \lambda, v) \in \mathcal{V}$, then $(UAU^{-1}, \lambda, Uv) \in \mathcal{V}$. Thus $\mathcal{V}$ is invariant under the product action (2.7). Furthermore, if $(A, \lambda, v) \in \mathcal{W}$, the condition of λ being a simple eigenvalue of A is invariant under the action of the group $\mathbb{U}_n(\mathbb{K})$ on (A, λ), hence the variety $\mathcal{W}$ is invariant under the action of $\mathbb{U}_n(\mathbb{K})$ as well. We have thus proved the following result.

Lemma 2.8. *The solution variety $\mathcal{V} \subset \mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big) \times \mathbb{P}(\mathbb{K}^n)$, and the subvariety $\mathcal{W} \subset \mathcal{V}$, are invariant under the action of $\mathbb{U}_n(\mathbb{K})$.* $\square$

Remark 2.9. The action of the group $\mathbb{U}_n(\mathbb{K})$ preserves the canonical structures defined on $\mathbb{P}(\mathbb{K}^n)$, $\mathbb{P}(\mathbb{K}^{n \times n})$, and $\mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big)$. Thus $\mathbb{U}_n(\mathbb{K})$ acts by isometries on these spaces. In particular, $\mathbb{U}_n(\mathbb{K})$ acts by isometries on $\mathcal{V}$. In addition, the projections π_1, π_2, and π are $\mathbb{U}_n(\mathbb{K})$–equivariant, i.e., they commute with the action of $\mathbb{U}_n(\mathbb{K})$.

2.4. Connectivity.
In this section we study the connectivity of the varieties $\mathcal{V}$ and $\mathcal{W}$.

Proposition 2.10. *The solution variety $\mathcal{V}$ is connected.*

Proof. Let $\widehat{\mathcal{V}} \subset (\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times (\mathbb{K}^n \setminus \{0\})$ be the inverse image of $\mathcal{V}$ under he canonical quotient projection $(\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times (\mathbb{K}^n \setminus \{0\}) \to \mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big) \times \mathbb{P}(\mathbb{K}^n)$, that is,

$$(2.8) \qquad \widehat{\mathcal{V}} := \{(A, \lambda, v) \in (\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times (\mathbb{K}^n \setminus \{0\}) : (\lambda I_n - A)v = 0\}.$$

It suffices to prove that $\widehat{\mathcal{V}}$ is connected.

The proof consists in the construction of a path connecting $(A, \lambda, v) \in \widehat{\mathcal{V}}$ with the triple $(e_1 e_1^* - I_n, 0, e_1) \in \widehat{\mathcal{V}}$. (Here e_1 is the first element of the canonical basis of $\mathbb{K}^n$, and e_1^* denotes the transpose of the column vector e_1.)

Let $(A, \lambda, v) \in \widehat{\mathcal{V}}$. With out loss of generality we can assume $\|v\| = 1$. Also note that we can connect v with e_1 by a rotation path $\{R_t\}_{t \in [0,1]} \subset \mathbb{U}_n(\mathbb{K})$. Since $\mathbb{U}_n(\mathbb{K})$ acts also on $\widehat{\mathcal{V}}$, from Lemma 2.8, we can assume $v = e_1$.

Let $A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$ be the matrix expression of A in the canonical basis $e_1, \ldots, e_n$ of $\mathbb{K}^n$, where $a \in \mathbb{K}^{1 \times (n-1)}$ and $\hat{A} \in \mathbb{K}^{(n-1) \times (n-1)}$. First assume that $A \neq \lambda I_n$. Then we can connect $(A, \lambda, e_1) \in \widehat{\mathcal{V}}$ with the triple $(A - \lambda I_n, 0, e_1) \in \widehat{\mathcal{V}}$ by the straight line path $\{(A - t\lambda I_n, (1-t)\lambda, e_1)\}_{0 \leq t \leq 1} \subset \widehat{\mathcal{V}}$. (Note that the condition $A \neq \lambda I_n$ is required.) In addition, $(A - \lambda I_n)|_{e_1^\perp} = \begin{pmatrix} a \\ \hat{A} - \lambda I_{n-1} \end{pmatrix}$ is a nonzero $n \times (n-1)$ matrix over $\mathbb{K}$. Since $n \geq 2$, the set $\mathbb{K}^{n \times (n-1)} \setminus \{0\}$ is connected. Then we can join the matrix $(A - \lambda I_n)|_{e_1^\perp}$ with the matrix $\begin{pmatrix} 0 \\ -I_{n-1} \end{pmatrix}$ by a path included in $\mathbb{K}^{n \times (n-1)} \setminus \{0\}$, and thus, $A - \lambda I_n$ and $e_1 e_1^* - I_n$ are connected by a path in $\mathbb{K}^{n \times n} \setminus \{0_n\}$ such that the first column remains invariant. Hence, we can join the triple $(A - \lambda I_n, 0, e_1)$ with the end point $(e_1 e_1^* - I_n, 0, e_1)$ by a path in $\widehat{\mathcal{V}}$.

If $A = \lambda I_n = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda I_{n-1} \end{pmatrix}$, then $\lambda \neq 0$, and therefore $(A, \lambda, e_1) \in \widehat{\mathcal{V}}$ is connected by a straight line path in $\widehat{\mathcal{V}}$ to the triple $(\lambda(I_n - e_1 e_1^*), 0, e_1) \in \widehat{\mathcal{V}}$. Now we are in the condition of the preceding argument. $\square$

Proposition 2.11.   (i) *When $\mathbb{K} = \mathbb{C}$, $\mathcal{W}$ is connected.*
   (ii) *When $\mathbb{K} = \mathbb{R}$ and n odd, $\mathcal{W}$ has two connected components.*
   (iii) *When $\mathbb{K} = \mathbb{R}$ and n even, $\mathcal{W}$ is connected.*

Since $\mathcal{V}$ is connected and Σ' is an algebraic subvariety of $\mathcal{V}$ (see Remark 2.6), the assertion (i) of this proposition follows from fact that a complex algebraic subvariety of $\mathcal{V}$ can not disconnect it; see for example Blum et al. [8, pp. 196].

For the proof of assertions (ii) and (iii) we need some definitions and a lemma.

Let $e_1, \ldots, e_n$ be the canonical basis of $\mathbb{R}^n$, and let $\det(\cdot)$ be the determinant function. Let $v \in \mathbb{R}^n$, $v \neq 0$. If $L : v^\perp \to v^\perp$ is a linear operator, then we define its determinant $\det_{v^\perp}(L)$ by

$$\det_{v^\perp}(L) = \det(v, Lv_2, \ldots, Lv_n),$$

where $v, v_2, \ldots, v_n$ is a positive orthonormal basis of $\mathbb{R}^n$, i.e., $\det(v, v_2, \ldots, v_n) = 1$.

Let $\widehat{\mathcal{V}}$ be the set defined in (2.8) and let

$$(2.9) \qquad \widehat{\mathcal{W}} := \{(A, \lambda, v) \in \widehat{\mathcal{V}} : \Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp} \text{ invertible}\}.$$

Let $D : \widehat{\mathcal{W}} \to \mathbb{R}$ be the function given by

$$D(A, \lambda, v) = \det_{v^\perp}(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}).$$

This function is the restriction, to $\widehat{\mathcal{W}}$, of a continuous function and thus continuous. Let $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$ be the inverse image, under D, of the rays $(0, +\infty)$ and $(-\infty, 0)$ respectively. (It is easily seen that $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$ are non–empty.) Then $\widehat{\mathcal{W}}$ is decomposed in the disjoint union of the open sets $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$.

Lemma 2.12. *The sets $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$ are connected.*

Proof. Let $S\mathcal{O}(n)$ be the special orthogonal group, that is, the subgroup of $\mathbb{U}_n(\mathbb{R})$ of matrices with determinant equal to one. The proof of this lemma is divided in several claims.

<u>Claim I:</u> The map $D : \widehat{\mathcal{W}} \to \mathbb{R}$ is invariant under the action of $S\mathcal{O}(n)$ on $\widehat{\mathcal{W}}$; hence the action of the special orthogonal group on $\widehat{\mathcal{W}}$ leave $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$ invariant: Let $(A, \lambda, v) \in \widehat{\mathcal{W}}$ and $U \in S\mathcal{O}(n)$. Let $v, v_2, \ldots, v_n$ be a positive orthonormal basis of $\mathbb{R}^n$. Since U is orthogonal and has determinant one, $Uv, Uv_2, \ldots, Uv_n$ is a positive orthonormal basis as well. Note that $Uv_i \in (Uv)^\perp$ and

$$\Pi_{(Uv)^\perp}(\lambda I_n - UAU^{-1})Uv_i = U(\lambda I_n - A)v_i - \alpha_i Uv,$$

where $\alpha_i = \langle U(\lambda I_n - A)v_i, Uv \rangle$. Then,

$$\begin{aligned}
D(UAU^{-1}, \lambda, Uv) &= \det_{(Uv)^\perp}(\Pi_{(Uv)^\perp}(\lambda I_n - UAU^{-1})|_{(Uv)^\perp}) \\
&= \det(Uv, U(\lambda I_n - A)v_2, \ldots, U(\lambda I_n - A)v_n) \\
&= \det_{v^\perp}(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}).
\end{aligned}$$

That is, $D(UAU^{-1}, \lambda, Uv) = D(A, \lambda, v)$, for every $U \in S\mathcal{O}(n)$, proving the claim.

Let $\hat{\pi}_2 : \widehat{\mathcal{V}} \to \mathbb{R}^n$ be the canonical projection $(A, \lambda, v) \in \widehat{\mathcal{V}} \mapsto v \in \mathbb{R}^n$. Note that $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}$ is the subset of triples $(A, \lambda, v) \in \widehat{\mathcal{W}}$ such that $v = e_1$.

<u>Claim II:</u> Each triple in $\widehat{\mathcal{W}}^+$ is connected, by path in $\widehat{\mathcal{W}}^+$, to a triple in $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$. Similarly, each triple in $\widehat{\mathcal{W}}^-$ is connected, by path in $\widehat{\mathcal{W}}^-$, to a triple in $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^-$:

Let $(A, \lambda, v) \in \widehat{\mathcal{W}}^+$. With out loss of generality we may assume $\|v\| = 1$. Let $U_t \in S\mathcal{O}(n)$, $0 \leq t \leq 1$, be a path in the special orthogonal group of $\mathbb{R}^n$ such that $U_0 = I_n$ and U_1 satisfying $U_1(v) = e_1$. Then by Lemma 2.8, the action $U_t \cdot (A, \lambda, v)$, $0 \leq t \leq 1$, is a path in $\widehat{\mathcal{W}}$ connecting the triple $(A, \lambda, v) \in \widehat{\mathcal{W}}$ to the triple $(U_1 A U_1^{-1}, \lambda, e_1) \in \hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}$. In addition, from Claim I the path is included in a level set of D, therefore, if the path started in a triple belonging to

$\widehat{\mathcal{W}}^+$ (respectively $\widehat{\mathcal{W}}^-$), then the path will remains on $\widehat{\mathcal{W}}^+$ (respectively on $\widehat{\mathcal{W}}^-$). This finishes the proof of the Claim II.

If $(A, \lambda, e_1) \in \hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}$, then we may write $A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$, where $a = \Pi_{e_1} A|_{e_1^\perp}$ belongs to $\mathbb{R}^{1 \times (n-1)}$ and $\hat{A} \in \mathbb{R}^{(n-1) \times (n-1)}$ invertible.

Let $\widehat{\mathcal{W}}_0$ be the subset of triples $(B, 0, e_1) \in \widehat{\mathcal{W}}^+$ such that e_1 is also a left eigenvector of B with eigenvalue 0, that is,

$$\widehat{\mathcal{W}}_0 = \{(B, 0, e_1) \in \widehat{\mathcal{W}} : \Pi_{e_1} B|_{e_1^\perp} = 0\},$$

and let $\widehat{\mathcal{W}}_0^+ = \widehat{\mathcal{W}}_0 \cap \widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}_0^- = \widehat{\mathcal{W}}_0 \cap \widehat{\mathcal{W}}^-$.

<u>Claim III:</u> $\widehat{\mathcal{W}}_0^+$ (respectively $\widehat{\mathcal{W}}_0^-$) is a deformation retract of $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$ (respectively $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^-$), and therefore the number of connected components is the same:

Let us prove that $\widehat{\mathcal{W}}_0^+$ is a deformation retract of $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$. (The proof for the set $\widehat{\mathcal{W}}_0^-$ is analogue.) Let $(A, \lambda, e_1) \in \hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$, and write $A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$.

Let $(A_t, \lambda_t, e_1) \in \widehat{\mathcal{V}}$, $0 \leq t \leq 1$, be the path of triples given by

$$A_t = \begin{pmatrix} (1-t)\lambda & (1-t)a \\ 0 & \hat{A} - t\lambda I_{n-1} \end{pmatrix}, \quad \lambda_t = (1-t)\lambda.$$

Since $A_t e_1 = (1-t)\lambda e_1$ and $\Pi_{e_1^\perp}(A_t - \lambda_t I_n)|_{e_1^\perp} = \hat{A} - \lambda I_{n-1}$, we conclude that $(A_t, \lambda_t, e_1) \in \hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$ for every $t \in [0, 1]$. Then, the continuous map

$$F : [0, 1] \times (\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+) \to \hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+, \quad F(t, (A, \lambda, v)) = (A_t, \lambda_t, e_1),$$

is a deformation retract of the space $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$ onto the subspace $\widehat{\mathcal{W}}_0^+$.

<u>Claim IV:</u> $\widehat{\mathcal{W}}_0^+$ and $\widehat{\mathcal{W}}_0^-$ are connected:

If $(B, 0, e_1) \in \widehat{\mathcal{W}}_0^+$ then we may write $B = \begin{pmatrix} 0 & 0 \\ 0 & \hat{B} \end{pmatrix}$, where $\hat{B} = \Pi_{e_1^\perp} B|_{e_1^\perp}$ belongs to the space of $(n-1) \times (n-1)$ invertible matrices with positive determinant, namely $\mathbb{G}l_{n-1}(\mathbb{R})^+$, which is a connected component of the linear group of $\mathbb{R}^{n-1}$. Then, under the identification

$$(B, 0, e_1) \in \widehat{\mathcal{W}}_0^+ \mapsto \Pi_{e_1^\perp} B|_{e_1^\perp} \in \mathbb{G}l_{n-1}(\mathbb{R})^+,$$

we have that $\widehat{\mathcal{W}}_0^+$ is homeomorphic to $\mathbb{G}l_{n-1}(\mathbb{R})^+$ and therefore connected. (The proof is analogue for the set $\widehat{\mathcal{W}}_0^-$.) This proves the claim.

Let us finish the proof of the lemma showing that $\widehat{\mathcal{W}}^+$ is connected. (The proof of the connectivity of $\widehat{\mathcal{W}}^-$ is analogue.) From Claim II each triple in $\widehat{\mathcal{W}}^+$ can be connected to a triple in $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$ by a path in $\widehat{\mathcal{W}}^+$. From Claim III the set $\hat{\pi}_2^{-1}(e_1) \cap \widehat{\mathcal{W}}^+$ has the same number of connected components as $\widehat{\mathcal{W}}_0^+$. Then Claim IV finishes the proof. $\square$

We have concluded from this lemma that $\widehat{\mathcal{W}}$ has two connected components, namely $\widehat{\mathcal{W}}^+$ and $\widehat{\mathcal{W}}^-$.

Proof of Proposition 2.11. The variety $\mathcal{W}$ is the quotient space of $\widehat{\mathcal{W}}$ under the equivalence relation of multiplying by nonzero real numbers the coordinates (A, λ)

and v respectively. Let $q : (\mathbb{K}^{n \times n} \setminus \{0_n\}) \times \mathbb{K} \times (\mathbb{K}^n \setminus \{0\}) \to \mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big) \times \mathbb{P}(\mathbb{K}^n)$ be the canonical quotient projection. In particular $q(\widehat{\mathcal{W}}) = \mathcal{W}$.

Let $(A, \lambda, v) \in \widehat{\mathcal{W}}$.

(ii) Let us assume that n is odd. Since

$$(2.10) \qquad D(\alpha A, \alpha \lambda, \beta v) = \det_{(\beta v)^\perp}(\Pi_{(\beta v)^\perp}(\alpha \lambda I_n - \alpha A)|_{(\beta v)^\perp})$$
$$= (\alpha)^{n-1} \det_{v^\perp}(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}),$$

for all α, $\beta \in \mathbb{R} \setminus \{0\}$, and $n - 1$ even, we conclude that the set of equivalence classes, of the quotient projection q, are include in one and only one of the connected components of $\widehat{\mathcal{W}}$. Hence $q(\widehat{\mathcal{W}}^+) \cap q(\widehat{\mathcal{W}}^-) = \emptyset$. Furthermore, since the quotient projection q is open and continuous we conclude, from Lemma 2.12, that $q(\widehat{\mathcal{W}}^+)$ and $q(\widehat{\mathcal{W}}^-)$ are (non–empty) open and connected sets. Thus $\mathcal{W}$ has two connected components, namely, $q(\widehat{\mathcal{W}}^+)$ and $q(\widehat{\mathcal{W}}^-)$.

(iii) When n is even, then $n-1$ is odd and therefore, from (2.10), the triples (A, λ, v) and $(-A, -\lambda, v)$ are in different components of $\widehat{\mathcal{W}}$, though they are equivalent triples in $\mathcal{W}$. Therefore every triple $(A, \lambda, v) \in \mathcal{W}$ has a representative in $\widehat{\mathcal{W}}^+$. Hence, from Lemma 2.12, we obtain that $\mathcal{W}$ is connected. $\qquad \square$

2.5. **Multidegree of $\mathcal{V}$.** The eigenvalue problem as a quadratic system belongs to the subset of ill–posed problems; see Section 1.7. The aim of this section is to prove that the bilinear approach considered in this paper gives the correct number of roots.

There are many different strategies to prove this with origins in algebraic geometry or algebraic topology. Yet we have not found a convenient proof in the literature to cite. One referee suggested that a possible scheme is to fix the matrix A and homogenize the eigenvalue λ in order to obtain a bi–homogeneous system in the variables (λ, v), and then apply some basic toric variety theory to obtain the result. Here we pursue a different strategy which make use in a more natural way our bilinear approach in (A, λ) and v.

For the sake of simplicity in the exposition we restrict ourself to the case $\mathbb{K} = \mathbb{C}$. This section follows closely D'Andrea–Krick–Sombra [11].

Since $\mathcal{V}$ is an algebraic subvariety of the product space $\mathbb{P}\big(\mathbb{C}^{n \times n} \times \mathbb{C}\big) \times \mathbb{P}(\mathbb{C}^n)$, there is a natural algebraic invariant associated to $\mathcal{V}$, namely, the *multidegree* of $\mathcal{V}$. This invariant is given by the numbers $\deg_{(n^2-1-i,i)}(\mathcal{V})$, $i = 0, \ldots, n-1$, where $\deg_{(n^2-1-i,i)}(\mathcal{V})$ is the number of points of intersection of $\mathcal{V}$ with the product $\Lambda \times \Lambda' \subset \mathbb{P}\big(\mathbb{C}^{n \times n} \times \mathbb{C}\big) \times \mathbb{P}(\mathbb{C}^n)$, where $\Lambda \subset \mathbb{P}\big(\mathbb{C}^{n \times n} \times \mathbb{C}\big)$ and $\Lambda' \subset \mathbb{P}(\mathbb{C}^n)$ are generic (n^2-1-i)-codimension plane and i-codimension plane respectively; see Fulton [18].

Lemma 2.13. *One has,*

$$\deg_{(n^2-1-i,i)}(\mathcal{V}) = \binom{n}{i+1}, \qquad for \quad i = 0, \ldots, n-1.$$

In order to give a proof of this lemma we recall some definitions from *intersection theory*; see Fulton [18]. (See also D'Andrea–Krick–Sombra [11].)

The Chow ring of $\mathbb{P}\big(\mathbb{C}^{n \times n} \times \mathbb{C}\big) \times \mathbb{P}(\mathbb{C}^n)$ is the graded ring

$$\mathcal{A}^*\left(\mathbb{P}\big(\mathbb{C}^{n \times n} \times \mathbb{C}\big) \times \mathbb{P}(\mathbb{C}^n)\right) = \mathbb{Z}[\omega_1, \omega_2]/(\omega_1^{n^2+1}, \omega_2^n),$$

where ω_1 and ω_2 denotes the rational equivalence classes of the inverse images of hyperplanes of $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$ and $\mathbb{P}(\mathbb{C}^n)$, under the projections $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)\to\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$ and $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)\to\mathbb{P}(\mathbb{C}^n)$ respectively.

Given a codimension n algebraic subvariety $\mathcal{X}\subset\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)$, the class of $\mathcal{X}$ in the Chow ring is

$$[\mathcal{X}] = \sum_{i=0}^{n-1} \deg_{(n^2-1-i,i)}(\mathcal{X})\,\omega_1^{i+1}\omega_2^{n-1-i} \in \mathcal{A}^*\left(\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)\right).$$

Proof of Lemma 2.13. Let F_i, $(i=1,\ldots,n)$, be the coordinate functions of F defined in (2.2). Since F_i is bilinear for each i, the class of $\{F_i=0\}$, as a subset of $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)$, is given by

$$[\{F_i=0\}] = \omega_1 + \omega_2 \in \mathcal{A}^*\left(\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})\times\mathbb{P}(\mathbb{C}^n)\right), \quad (i=1,\ldots,n).$$

Then, the class of $\mathcal{V}$ in the Chow ring is

$$[\mathcal{V}] = [\{F_1=0\}\cap\cdots\cap\{F_n=0\}] = \prod_{i=1}^{n}[\{F_i=0\}],$$

where the last equality follows from Bézout identity. Therefore one gets

$$[\mathcal{V}] = (\omega_1+\omega_2)^n \equiv \sum_{\ell=1}^{n}\binom{n}{\ell}\omega_1^\ell\omega_2^{n-\ell},$$

that is, $\deg_{(n^2-1-i,i)}(\mathcal{V}) = \binom{n}{i+1}$. $\qquad\square$

From Lemma 2.13 we obtain that the number of points of intersection of $\mathcal{V}$ with the product $\Lambda\times\mathbb{P}(\mathbb{C}^n)$ (for $\Lambda\subset\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$ a generic hyperplane of codimension n^2-1) is n. In particular the inverse image of $A\in\mathbb{P}(\mathbb{C}^{n\times n})\backslash\Sigma$, under the projection $\pi:\mathcal{V}\to\mathbb{P}(\mathbb{C}^{n\times n})$, is the intersection of $\mathcal{V}$ with $\Lambda_A\times\mathbb{P}(\mathbb{C}^n)$, where Λ_A is a particular hyperplane of codimension n^2-1 in $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$, namely, Λ_A is the projective line containing the pair of points $\{(A,0),(A,1)\}\in\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$. However, the family of all projective lines Λ_A, varying A on $\mathbb{P}(\mathbb{C}^{n\times n})$, is not a generic family on $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$. In the next proposition we prove that, actually, the family Λ_A for $A\in\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$ is included in the generic family of hyperplanes of codimension n^2-1 of $\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$ satisfying Lemma 2.13 for $i=0$. [1]

Let A be a finite set. We denote by $\#A$ the cardinal number of A.

Proposition 2.14. *For all $A\in\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$ we have $\#\pi^{-1}(A)=\deg_{(n^2-1,0)}(\mathcal{V})$.*

Proof. Recall that if $A\in\mathbb{P}(\mathbb{C}^{n\times n})\backslash\Sigma$ then the projection π is a local diffeomorphism between a neighbourhood of each inverse image of A in $\mathcal{V}$ and a neighbourhood of A in $\mathbb{P}(\mathbb{C}^{n\times n})\backslash\Sigma$. Then the number of inverse image is locally constant on $\mathbb{P}(\mathbb{C}^{n\times n})\backslash\Sigma$. Furthermore, since Σ is an algebraic subvariety of $\mathbb{P}(\mathbb{C}^{n\times n})$ (see Remark 2.6), then $\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$ is connected; cf. proof (i) of Proposition 2.11. Thus the number of inverse images under π is constant on $\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$. If $(A,\lambda,v)\in\mathcal{V}\setminus\Sigma'$ then, from Lemma 2.5, the number of inverse image of (A,λ), under $\pi_1:\mathcal{V}\to\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$, is one. Hence the restriction $\pi_1|_{(\mathcal{V}\setminus\Sigma')}:(\mathcal{V}\setminus\Sigma')\to\mathbb{P}(\mathbb{C}^{n\times n}\times\mathbb{C})$ is a bijective map onto its image $\pi_1(\mathcal{V}\setminus\Sigma')$. Therefore given $A\in\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$, we have

---

[1] If $A\in\mathbb{P}(\mathbb{C}^{n\times n})\setminus\Sigma$ then A has n distinct eigenvalues, thus the cardinal number of $\pi^{-1}(A)$ coincides with $\deg_{(n^2-1,0)}(\mathcal{V})$. In Proposition 2.14 we give an independent proof of this fact which we consider interesting *per se*.

$\#\pi^{-1}(A) = \#(p|_{\pi_1(\mathcal{V})})^{-1}(A)$, where p is the projection map given in (2.4). In addition, from [24, Corollary 5.6], we get that $\#(p|_{\pi_1(\mathcal{V})})^{-1}(A) = \deg \pi_1(\mathcal{V})$, where deg is the degree of the projective algebraic subvariety $\pi_1(\mathcal{V}) \subset \mathbb{P}(\mathbb{C}^{n\times n} \times \mathbb{C})$. Since $\dim \pi_1(\mathcal{V}) = \dim(\mathcal{V})$ and the fact that $\pi_1|_{(\mathcal{V}\setminus\Sigma')} : (\mathcal{V} \setminus \Sigma') \to \pi_1(\mathcal{V} \setminus \Sigma')$ is bijective, we get that $\#(\Lambda \times \mathbb{P}(\mathbb{C}^n)) \cap \mathcal{V} = \#\Lambda \cap \pi_1(\mathcal{V})$, for a generic $(n^2 - 1)$-codimension plane $\Lambda \subset \mathbb{P}(\mathbb{C}^{n\times n} \times \mathbb{C})$. Then we conclude that $\deg \pi_1(\mathcal{V}) = \deg_{(n^2-1,0)}(\mathcal{V})$. $\square$

Remark 2.15. From Proposition 2.14 and Lemma 2.13 we get that the restriction of the projection $\pi|_{(\mathcal{V}\setminus\pi^{-1}(\Sigma))} : (\mathcal{V} \setminus \pi^{-1}(\Sigma)) \to \mathbb{P}(\mathbb{C}^{n\times n}) \setminus \Sigma$ is an n-fold covering map.

3. Condition number

3.1. **Introduction.** In this section we introduce the eigenvalue and eigenvector condition numbers. We study some basic properties of these condition numbers and we show some examples. We define the condition number of the eigenvalue problem. We discuss the condition number theorem for this framework, which relates the condition number with the distance to ill–posed problems. In the last part of this section we study the rate of change of condition numbers.

3.2. **Eigenvalue and eigenvector condition numbers.** When (A, λ, v) belongs to $\mathcal{W}$, according to the implicit function theorem, π has an inverse defined in some neighbourhood $\mathcal{U}_A \subset \mathbb{P}(\mathbb{K}^{n\times n})$ of A such that $\pi^{-1}(A) = (A, \lambda, v)$. This map $\mathscr{S} = \pi^{-1}|_{\mathcal{U}_A} : \mathcal{U}_A \to \mathcal{V}$ is called the *solution map*. It associates to any matrix $B \in \mathcal{U}_A$ the eigentriple (B, λ_B, v_B) close to (A, λ, v). Its derivative

$$D\mathscr{S}(A, \lambda, v) : T_A\mathbb{P}(\mathbb{K}^{n\times n}) \to T_{(A,\lambda,v)}\mathcal{V},$$

is called the *condition operator* at (A, λ, v).

If $(A, \lambda, v) \in \mathcal{W}$, the derivative $D\mathscr{S}(A, \lambda, v)$ associates to each $\dot{B} \in T_A\mathbb{P}(\mathbb{K}^{n\times n})$ a triple $(\dot{A}, \dot{\lambda}, \dot{v}) \in T_{(A,\lambda,v)}\mathcal{V}$. This association defines two linear maps,

$$D\mathscr{S}_\lambda(A, \lambda, v)\dot{B} = (\dot{A}, \dot{\lambda}) \quad \text{and} \quad D\mathscr{S}_v(A, \lambda, v)\dot{B} = \dot{v},$$

namely, the condition operators of the eigenvalue and eigenvector respectively.

Recall that $\mathbb{P}(\mathbb{K}^{n\times n})$ is equipped with the canonical Hermitian structure induced by the Frobenius Hermitian product on $\mathbb{K}^{n\times n}$.

Definition 3.1. The condition numbers of the eigenvalue and eigenvector, at $(A, \lambda, v) \in \mathcal{W}$, are defined by

$$\mu_\lambda(A, \lambda, v) := \sup_{\substack{\dot{B} \in A^\perp \\ \|\dot{B}\|_F = \|A\|_F}} \|D\mathscr{S}_\lambda(A, \lambda, v)\dot{B}\|_{(A,\lambda)},$$

$$\mu_v(A, \lambda, v) := \sup_{\substack{\dot{B} \in A^\perp \\ \|\dot{B}\|_F = \|A\|_F}} \|D\mathscr{S}_v(A, \lambda, v)\dot{B}\|_v.$$

Proposition 3.2. *Let* $(A, \lambda, v) \in \mathcal{W}$. *Then,*

(i)

$$\mu_\lambda(A, \lambda, v) = \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} \left[1 + \frac{\|v\|^2 \|u\|^2}{|\langle v, u\rangle|^2} \right]^{1/2},$$

where $u \in \mathbb{K}^n$ *is any left eigenvector of* A *with eigenvalue* λ: *a nonzero vector satisfying* $(\lambda I_n - A)^* u = 0$.

(ii)
$$\mu_v(A, \lambda, v) = \|A\|_F \, \|(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\|,$$

where $\|\cdot\|$ is the operator norm.

Remark 3.3. $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is a linear map from the Hermitian complement of v in $\mathbb{K}^n$ into itself. Hence the operator norm of its inverse is independent of the representative of v in $\mathbb{P}(\mathbb{K}^n)$.

For the proof of Proposition 3.2 we need two lemmas.

Lemma 3.4. *Let* $(A, \lambda, v) \in \mathcal{W}$. *Then for* $\dot{B} \in T_A \mathbb{P}(\mathbb{K}^{n \times n})$, *one gets:*

(i)
$$D\mathscr{S}_\lambda(A, \lambda, v)\dot{B} = \left(\dot{B} - \dot{\lambda}\frac{\overline{\lambda}}{\|A\|_F^2}A, \dot{\lambda}\right), \quad \text{where } \dot{\lambda} = \frac{\langle \dot{B}v, u \rangle}{\left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right)\langle v, u \rangle},$$

where $u \in \mathbb{K}^n$ is any left eigenvector of A with eigenvalue λ;

(ii)
$$D\mathscr{S}_v(A, \lambda, v)\dot{B} = (\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\Pi_{v^\perp}(\dot{B}v).$$

Proof. (i): Let $\dot{B} \in A^\perp$, and let $(\dot{A}, \dot{\lambda}) \in (A, \lambda)^\perp$ such that $D\mathscr{S}_\lambda(A, \lambda, v)\dot{B} = (\dot{A}, \dot{\lambda})$. Then, by the definition of $D\mathscr{S}_\lambda(A, \lambda, v)$ and (2.6) we get

$$(3.1) \qquad\qquad \dot{B} = \dot{A} + \frac{\lambda \overline{\lambda}}{\|A\|_F^2}A.$$

Let $u \in \mathbb{K}^n$ be any left eigenvector of A with eigenvalue λ. Since u is in the Hermitian complement of the range of $(\lambda I_n - A)|_{v^\perp}$, then, from (2.3) we get $\langle \dot{A}v, u \rangle = \dot{\lambda}\langle v, u \rangle$. Furthermore, since $(A, \lambda, v) \in \mathcal{W}$, then $v \notin \text{Im}(A - \lambda I_n)$, and $\langle v, u \rangle \neq 0$. Thus

$$(3.2) \qquad\qquad \dot{\lambda} = \frac{\langle \dot{A}v, u \rangle}{\langle v, u \rangle}.$$

From (3.1) and (3.2) follows

$$\dot{\lambda} = \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}}\frac{\langle \dot{B}v, u \rangle}{\langle v, u \rangle}.$$

(ii): From (2.3) again one gets $\Pi_{v^\perp}(\lambda I_n - A)\dot{v} = \Pi_{v^\perp}\dot{A}v$. Furthermore, since $(A, \lambda, v) \in \mathcal{W}$, then

$$\dot{v} = (\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\Pi_{v^\perp}(\dot{A}v).$$

Since, from (3.1), one has $\Pi_{v^\perp}(\dot{B}v) = \Pi_{v^\perp}(\dot{A}v)$, the result follows. $\square$

Lemma 3.5. *Let* $(A, \lambda, v) \in \mathcal{W}$ *and let* $u \in \mathbb{K}^n$ *be any left eigenvector of* A *with eigenvalue* λ. *Then*

$$\sup_{\substack{\dot{B} \in A^\perp \\ \|\dot{B}\|_F = \|A\|_F}} \left|\langle \dot{B}v, u \rangle\right| = \|A\|_F \sqrt{\|v\|^2\|u\|^2 - \frac{|\lambda|^2}{\|A\|_F^2}|\langle v, u \rangle|^2}.$$

Proof. Note that $\langle Mv, u\rangle = \langle M, uv^*\rangle_F$ for every matrix $M \in \mathbb{K}^{n\times n}$. Write

$$uv^* = \langle uv^*, \frac{A}{\|A\|_F}\rangle_F \frac{A}{\|A\|_F} + \alpha C,$$

where $C \in A^\perp$ and $\|C\|_F = 1$. Then

$$\sup_{\substack{\dot{B}\in A^\perp \\ \|\dot{B}\|_F=\|A\|_F}} \left|\langle \dot{B}v, u\rangle\right| = \sup_{\substack{\dot{B}\in A^\perp \\ \|\dot{B}\|_F=\|A\|_F}} \left|\langle \dot{B}, \alpha C\rangle_F\right| = |\alpha|\, \|A\|_F.$$

Furthermore,

$$|\alpha|^2 = \|uv^*\|_F^2 - |\langle uv^*, \frac{A}{\|A\|_F}\rangle_F|^2,$$

where $\|uv^*\|_F = \|u\|\, \|v\|$. Since $Av = \lambda v$, then $|\langle uv^*, \frac{A}{\|A\|_F}\rangle_F| = \frac{|\lambda|}{\|A\|_F}|\langle u, v\rangle|$. $\square$

Proof of Proposition 3.2. (i): From Lemma 3.4, for any $\dot{B}$ such that $\langle \dot{B}, A\rangle_F = 0$, and $\|\dot{B}\|_A = 1$,

$$(3.3) \qquad \|D\mathscr{S}_\lambda(A, \lambda, v)\dot{B}\|_{(A,\lambda)}^2 = \frac{\|A\|_F^2 + |\dot{\lambda}|^2 \left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right)}{\|A\|_F^2 + |\lambda|^2}$$

$$= \frac{\|A\|_F^2 + \left|\frac{\langle \dot{B}v,u\rangle}{\langle v,u\rangle}\right|^2 \left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right)^{-1}}{\|A\|_F^2 + |\lambda|^2}.$$

Then, the proof of (i) can be deduced from Lemma 3.5.

(ii): Since $Av = \lambda v$, we have $\Pi_{v^\perp}(\dot{B}v) = \Pi_{v^\perp}((\dot{B} + \alpha A)v)$, for any $\alpha \in \mathbb{K}$ and $\dot{B} \in A^\perp$. Then, from Lemma 3.4 we get:

$$\mu_v(A, \lambda, v) = \sup_{\substack{\dot{B}\in A^\perp \\ \|\dot{B}\|_F=\|A\|_F}} \left\|(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\Pi_{v^\perp}(\dot{B}v)\right\|_v$$

$$= \sup_{\substack{\dot{B}\in \mathbb{K}^{n\times n} \\ \|\dot{B}\|_F=1}} \|A\|_F \left\|(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\Pi_{v^\perp}(\dot{B}v)\right\|_v.$$

Since $\{\Pi_{v^\perp}(\dot{B}v) : \dot{B} \in \mathbb{K}^{n\times n}, \|\dot{B}\|_F = 1\}$ fills the ball of radius $\|v\|$ in $v^\perp$, the result follows. $\square$

3.3. **Some basic properties.** In the next paragraphs we show some basic properties concerning the condition numbers μ_λ and μ_v.

Proposition 3.6. *The condition numbers μ_λ and μ_v are invariant under the action of the group $\mathbb{U}_n(\mathbb{K})$, that is,*

$$\mu_\lambda(UAU^{-1}, \lambda, Uv) = \mu_\lambda(A, \lambda, v),$$
$$\mu_v(UAU^{-1}, \lambda, Uv) = \mu_v(A, \lambda, v),$$

for every $U \in \mathbb{U}_n(\mathbb{K})$.

Remark 3.7. The proof of Proposition 3.6 can be deduced from the expressions of μ_λ and μ_v given in Proposition 3.2. However, we prefer to give a different proof which emphasize the fact that the property of the condition numbers of being unitarily invariant resides on the natural election of our Hermitian structures given in our geometric framework.

Proof of Proposition 3.6. The condition operators of the eigenvalue and eigenvector are given by the derivative of the (locally defined) maps $\mathscr{S}_\lambda = \pi_1 \circ \mathscr{S}$ and $\mathscr{S}_v = \pi_2 \circ \mathscr{S}$ respectively. From Remark 2.9 the projections π, π_1 and π_2 are $\mathbb{U}_n(\mathbb{K})$-equivariants, hence the action of $\mathbb{U}_n(\mathbb{K})$ commutes with $\mathscr{S}$, $\mathscr{S}_\lambda$ and $\mathscr{S}_v$. In addition, since $\mathbb{U}_n(\mathbb{K})$ acts by isometries on all the intervening spaces, the result follows. $\qquad\square$

Lemma 3.8. *The condition numbers μ_λ and μ_v are bounded below by $1/\sqrt{2}$.*

Proof. Let $(A, \lambda, v) \in \mathcal{W}$. Since $|\lambda| \le \|A\|_F$, the proof for μ_λ follows immediately by Proposition 3.2.

For the proof for μ_v, first fix a representative of $(A, \lambda, v) \in \mathcal{W}$ such that $\|A\|_F = 1$ and $\|v\| = 1$. In addition, since the action of $\mathbb{U}_n(\mathbb{K})$ on $\mathbb{P}(\mathbb{K}^n)$ is transitive, by Proposition 3.6 we may assume that v is the first element of the canonical basis. Then A has the form $\begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$, where $a \in \mathbb{K}^{1 \times (n-1)}$ and $\hat{A} \in \mathbb{K}^{(n-1) \times (n-1)}$. Under these assumptions $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp} = \lambda I_{n-1} - \hat{A}$, hence

$$\|\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\| \le \|\hat{A}\| + |\lambda| \le \|\hat{A}\|_F + |\lambda| \le \sqrt{2}\|A\|_F,$$

where last inequality follows from the inequality $x + y \le \sqrt{2}(x^2 + y^2)^{1/2}$, for x, y, in $\mathbb{R}$. Since we assume $\|A\|_F = 1$, we obtain, $\|\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\| \le \sqrt{2}$. Therefore from Proposition 3.2 we get

$$1 = \| \left(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\right)^{-1} \Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\| \le \sqrt{2}\mu_v(A, \lambda, v).$$

$\qquad\square$

Remark 3.9. Examples (a) and (b) of Section 3.4 show that the lower bound in Lemma 3.8 is sharp.

Remark 3.10. Let $(A, \lambda, v) \in \mathcal{W}$. Then $(A + \alpha I_n, \lambda + \alpha, v) \in \mathcal{W}$, for all $\alpha \in \mathbb{K}$, and $\mu_\lambda(A + \alpha I_n, \lambda + \alpha, v)$ is constant independent of α. On the other hand, this is not the case for the eigenvector condition number. More precisely,

$$\mu_v(A + \alpha I_n, \lambda + \alpha, v) = \frac{\|A + \alpha I_n\|_F}{\|A\|_F}\mu_v(A, \lambda, v).$$

In particular, it is an easy exercise to check that $\mu_v(A + \alpha I_n, \lambda + \alpha, v)$ is minimized, as a function of α, when the matrix $A + \alpha I_n$ has trace equal to zero, namely, $\alpha = -\mathrm{tr}(A)/n$. As we see in the next section, this procedure may improve drastically μ_v in some cases, and thus, it could be used as a natural pre–conditioning.

Remark 3.11. Let $(A, \lambda, v) \in \mathcal{W}$. If $(\lambda I_n - A)^* v = 0$, that is, if v is also a left eigenvector of A with eigenvalue λ, then, from Proposition 3.2, one has

$$\mu_\lambda(A, \lambda, v) = \frac{\sqrt{2}}{1 + \frac{|\lambda|^2}{\|A\|_F^2}}.$$

Thus $\mu_\lambda(A, \lambda, v) \le \sqrt{2}$.

From the previous remark we conclude that when A is normal, i.e., $A^*A = AA^*$, the eigenvalue condition number μ_λ is not related to the distance to the discriminant variety Σ. On the other hand, μ_v happens to be more interesting since, roughly speaking, $\mu_v(A, \lambda, v)$ measures how close to λ others eigenvalues of A are. More precisely, we have the following result.

Lemma 3.12. *Let A be a normal matrix. If $(A, \lambda, v) \in \mathcal{W}$ then*

$$\mu_\lambda(A, \lambda, v) \leq \sqrt{2}; \qquad \mu_v(A, \lambda, v) = \frac{\|A\|_F}{\min_i |\lambda - \lambda_i|},$$

where the minimum is taken for λ_i eigenvalue of A different from λ.

Proof. The inequality for μ_λ follows from Remark 3.11.
Since A is normal, by Proposition 3.6, we may assume that A is the diagonal matrix $\mathrm{Diag}(\lambda, \lambda_2, \ldots, \lambda_n)$, where λ, λ_i are the eigenvalues of A. Furthermore, since $(A, \lambda, v) \in \mathcal{W}$, then $\lambda \neq \lambda_i$ for $i = 2, \ldots n$. Thus the result follows from Proposition 3.2.   □

3.4. **Some examples.** In this paragraph we compute the eigenvalue and eigenvector condition numbers for some simple matrices. We denote by e_1 the first element of the canonical basis of the underlying $\mathbb{K}^n$.

(a) Let $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Then $(A_1, 1, e_1) \in \mathcal{W}$. Since A_1 is symmetric and has eigenvalues -1 and 1, we have $\mu_v(A_1, 1, e_1) = 1/\sqrt{2}$.

(b) Let $A_2 = e_1 e_1^* \in \mathbb{K}^{n \times n}$. Then $(A_2, 1, e_1) \in \mathcal{W}$, where $\mu_\lambda(A_2, 1, e_1) = 1/\sqrt{2}$ and $\mu_v(A_2, 1, e_1) = 1$. (Note that, when $n = 2$, A_2 and $A_1 + I_2$ are in the same equivalent class of $\mathbb{P}(\mathbb{K}^{2 \times 2})$.)

(c) Let $B_\epsilon = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \epsilon \end{pmatrix}$, where $\epsilon > 0$. Then $(B_\epsilon, 1, e_1) \in \mathcal{W}$. One has $\mu_\lambda(B_\epsilon, 1, e_1) \leq \sqrt{2}$, and $\mu_v(B_\epsilon, 1, e_1) = \sqrt{1 + |1 - \epsilon|^2}/\epsilon$. In particular, as $\epsilon \to 0$ we have $\mu_v(B_\epsilon, 1, e_1) \to +\infty$. Surprisingly, the behaviour of μ_v can be drastically changed by the pre–conditioning procedure described in Remark 3.10. More precisely, the matrix $B_\epsilon - (\mathrm{tr}(B_\epsilon)/2)I_2 = (\epsilon/2)A_1$ (where A_1 is given in the example (a) above), and hence μ_v attains its minimum value on the associated eigentriple.

(d) Let $B_\epsilon = \begin{pmatrix} 1 & \epsilon \\ 1 & 1 \end{pmatrix}$, where $\epsilon > 0$. This matrix was studied by Wilkinson in [37] as an example of ill–conditioned matrix. One has $(B_\epsilon, \lambda_\epsilon, v_\epsilon) \in \mathcal{W}$, where $v_\epsilon = (\sqrt{\epsilon}, 1)^T$ and $\lambda_\epsilon = 1 + \sqrt{\epsilon}$. Then, we have $\mu_v(B_\epsilon, \lambda_\epsilon, v_\epsilon) = \sqrt{3 + \epsilon^2}/(2\sqrt{\epsilon})$ and $\mu_\lambda(B_\epsilon, \lambda_\epsilon, v_\epsilon) = \sqrt{1 + 6\epsilon + \epsilon^2}/(4\sqrt{\epsilon})$. So, as ϵ decrease to zero, both condition numbers growth to $+\infty$. We return to this example in Section 3.6.

(e) Let $B_\epsilon = \begin{pmatrix} 1 & 1/\epsilon \\ 0 & 2 \end{pmatrix}$, where $\epsilon > 0$. Then $(B_\epsilon, 1, e_1) \in \mathcal{V}$. It is easily to check that μ_λ and μ_v, at $(B_\epsilon, 1, e_1)$, are larger that $1/(2\epsilon)$. Therefore both condition numbers growth to infinity as ϵ decrease to zero, even though $\chi_{B_\epsilon}(z) = (z-1)(z-2)$ is a well–posed polynomial; cf. Remark 1.4.

3.5. **Condition number of the eigenvalue problem.** The condition number of a computational problem is usually defined as the operator norm of the map giving the first order variation of the output in terms of the first order variation of the input; c.f. Definition 3.1. In our case the condition number should be the operator norm of the condition operator $D\mathscr{S}(A, \lambda, v)$ given in Section 3.2, i.e.,

$$\|D\mathscr{S}(A, \lambda, v)\| := \sup_{\substack{\dot{B} \in A^\perp \\ \|\dot{B}\|_F = \|A\|_F}} \|D\mathscr{S}(A, \lambda, v)\dot{B}\|_{(A, \lambda, v)}.$$

However, instead of this definition, we define the condition number of the eigenvalue problem in the following way.

Definition 3.13 (Condition Number of the Eigenvalue Problem)**.** The condition number of the eigenvalue problem is defined by

$$\mu(A, \lambda, v) := \max\{1, \mu_v(A, \lambda, v)\}, \qquad (A, \lambda, v) \in \mathcal{W}.$$

In item (ii) of the next proposition we show that this definition and the usual one are essentially equivalent.

Proposition 3.14. *Let* $(A, \lambda, v) \in \mathcal{W}$*. Then,*

(i) $\mu_\lambda(A, \lambda, v) < (1 + \frac{|\lambda|^2}{\|A\|_F^2})^{-1} (2 + \mu_v(A, \lambda, v)^2)^{1/2}$;

(ii) $\mu(A, \lambda, v) < \|D\mathscr{S}(A, \lambda, v)\| < 2\,\mu(A, \lambda, v)$.

Proof. Fix a representative of $(A, \lambda, v) \in \mathcal{W}$ such that $\|A\|_F = 1$ and $\|v\| = 1$. Furthermore, by Proposition 3.6, without loss of generality we may assume that v is the first element of the canonical basis, and thus we may write $A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$, where $a \in \mathbb{K}^{1 \times (n-1)}$ and $\hat{A} \in \mathbb{K}^{(n-1) \times (n-1)}$.

(i) Since $A - \lambda I_n = \begin{pmatrix} 0 & a \\ 0 & \hat{A} - \lambda I_{n-1} \end{pmatrix}$, a straightforward computation shows that $u = (1, -[(\hat{A} - \lambda I_{n-1})^*]^{-1} a^*)^T$ is a solution of $(A - \lambda I_n)^* u = 0$, i.e., u is a left eigenvector associated to λ. Here $\cdot^T$ and $\cdot^*$ denote the transpose and conjugate transpose respectively. Then,

$$\frac{|\langle v, u \rangle|}{\|v\|\,\|u\|} = \frac{1}{\sqrt{1 + \|[(\hat{A} - \lambda I_{n-1})^*]^{-1} a^*\|^2}} \geq \frac{1}{\sqrt{1 + \|[(\hat{A} - \lambda I_{n-1})^*]^{-1}\|^2\,\|a\|^2}}$$

Since for every invertible matrix B, $\|(B^*)^{-1}\| = \|(B^{-1})^*\| = \|B^{-1}\|$, then

$$\frac{|\langle v, u \rangle|}{\|v\|\,\|u\|} \geq \frac{1}{\sqrt{1 + \|(\hat{A} - \lambda I_{n-1})^{-1}\|^2\,\|a\|^2}}.$$

Furthermore, since $(A, \lambda, v) \in \mathcal{W}$, then $|\lambda|$ and $\|\hat{A}\|_F$ cannot be zero at the same time (if this is the case then $\lambda = 0$ is a multiple eigenvalue). Then we have $1 = \|A\|_F = (|\lambda|^2 + \|a\|^2 + \|\hat{A}\|_F^2)^{1/2} > \|a\|$, and therefore from Proposition 3.2

$$\frac{|\langle v, u \rangle|}{\|v\|\,\|u\|} > \frac{1}{\sqrt{1 + \mu_v(A, \lambda, v)^2}}.$$

(ii): From the definition of the condition operator $D\mathscr{S}(A, \lambda, v)$ and equation (3.3), we obtain, for every $\dot{B} \in T_A \mathbb{P}(\mathbb{K}^{n \times n})$, with $\|\dot{B}\|_F = 1$, that

$$\|D\mathscr{S}(A, \lambda, v)(\dot{B})\|_{(A, \lambda, v)}^2 \geq \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} + \|D\mathscr{S}_v(A, \lambda, v)(\dot{B})\|_v^2.$$

Then, maximizing over $\dot{B} \in T_A \mathbb{P}(\mathbb{K}^{n \times n})$ such that $\|\dot{B}\|_A = 1$, we get the lower bound

$$\|D\mathscr{S}(A, \lambda, v)\|^2 \geq \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} + \mu_v(A, \lambda, v)^2.$$

Now, the lower bound in (ii) follows from the following claim.
<u>Claim:</u> Let $(A, \lambda, v) \in \mathcal{W}$, then $\mu(A, \lambda, v)^2 < \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} + \mu_v(A, \lambda, v)^2$:

If $\mu_v(A, \lambda, v) \geq 1$ then $\mu(A, \lambda, v)^2 = \mu_v(A, \lambda, v)^2 < \frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} + \mu_v(A, \lambda, v)^2$. There-

fore from Lemma 3.8, it suffices to prove the claim for the range $\frac{1}{\sqrt{2}} \leq \mu_v(A, \lambda, v) < 1$.
In this range $\mu(A, \lambda, v) = 1$. Furthermore, since $\frac{1}{1 + \frac{|\lambda|^2}{\|A\|_F^2}} \geq \frac{1}{2}$ and $\mu_v(A, \lambda, v)^2 \geq \frac{1}{2}$

we reduce our problem to prove that the last two inequalities cannot be equalities at
the same time. This assertion follows from the fact that the condition $\|A\|_F = |\lambda|$
implies $\mu_v(A, \lambda, v) = 1$ whenever $(A, \lambda, v) \in \mathcal{W}$.

Let us prove the second inequality in (ii). By the definition of the condition
operator we get

$$\|D\mathscr{S}(A, \lambda, v)\|^2 \leq \mu_v(A, \lambda, v)^2 + \mu_\lambda(A, \lambda, v)^2.$$

Then, from assertion (i) of this proposition we obtain that

$$\mu_v(A, \lambda, v)^2 + \mu_\lambda(A, \lambda, v)^2 < \mu(A, \lambda, v)^2 + (2 + \mu_v(A, \lambda, v)^2) \leq 4\mu(A, \lambda, v)^2,$$

proving the upper bound. $\square$

Remark 3.15. Example (c) in Section 3.4 shows that the inequality (i) in Proposi-
tion 3.14 is far from be sharp.

The next result follows immediately from Proposition 3.6.

Proposition 3.16. *The condition number μ is invariant under the action of $\mathbb{U}_n(\mathbb{K})$,*
i.e., for every $U \in \mathbb{U}_n(\mathbb{K})$, one has $\mu(UAU^{-1}, \lambda, Uv) = \mu(A, \lambda, v)$. $\square$

The next section is included for the sake of completeness but is not needed for
the proof of our main results.

3.6. **Condition Number Theorem.** In this section we study the relation of
$\mu(A, \lambda, v)$ with the distance of (A, λ, v) to Σ'. The main objective of this section is
to prove the following theorem.

Theorem 4. *For every $(A, \lambda, v) \in \mathcal{W}$,*

$$\mu_v(A, \lambda, v) \leq \frac{1}{\left(1 + \frac{|\lambda|^2}{\|A\|_F^2}\right)^{1/2}} \frac{1}{\sin(d_{\mathbb{P}^2}((A, \lambda, v), \Sigma_v'))},$$

where Σ_v' is the intersection of the fiber $\mathcal{V}_v = \pi_2^{-1}(v) \subset \mathcal{V}$ with the ill–posed variety
Σ'.

The proof of Theorem 1 follows immediately from Definition 3.13 and Theorem 4.

In general, if $(\mathbb{E}, \langle \cdot, \cdot \rangle)$ is a finite dimensional Hermitian vector space over $\mathbb{K}$, and
$q : \mathbb{E} \setminus \{0\} \to \mathbb{P}(\mathbb{E})$ is the canonical quotient projection defining the projective space
$\mathbb{P}(\mathbb{E})$, then, given $\Lambda \subset \mathbb{P}(\mathbb{E})$, we define

(3.4) $\widehat{\Lambda} := q^{-1}(\Lambda) \subset (\mathbb{E} \setminus \{0\}).$

Recall from the introduction that we write interchangeably a nonzero vector and
its corresponding class in the projective space.

With this notation the following result is elementary.

Lemma 3.17. *Given $x \in \mathbb{E}$, $x \neq 0$, and $\Lambda \subset \mathbb{P}(\mathbb{E})$ we have*

$$\sin(d_\mathbb{P}(x, \Lambda)) = \frac{d_\mathbb{E}(x, \widehat{\Lambda})}{\|x\|},$$

where $d_{\mathbb{E}}(x, \widehat{\Lambda}) = \inf\{\|x - y\| : y \in \widehat{\Lambda}\}$ and $d_{\mathbb{P}}(x, \Lambda) = \inf\{d_{\mathbb{P}}(x, z) : z \in \Lambda\}$. $\qquad \square$

The next proposition is a version, adapted to this context, of a known result given by Shub–Smale [30].

Recall that $\Sigma = \pi(\Sigma') \subset \mathbb{P}(\mathbb{K}^{n \times n})$.

Lemma 3.18. *Let $(A, \lambda, v) \in \mathcal{W}$. Then*

$$\mu_v(A, \lambda, v) = \frac{\|A\|_F}{d_F(A - \lambda I_n, \widehat{\Sigma_{0,v}})},$$

where $\Sigma_{0,v} = \{B \in \mathbb{P}(\mathbb{K}^{n \times n}) : Bv = 0, rank(B^2) < n - 1\} \subset \Sigma$.

Proof. In Shub–Smale [30] it is proved that, for a fixed triple $(A, \lambda, v) \in \widehat{\mathcal{W}}$,

$$d_F(\lambda I_n - A, \widehat{\Sigma_{0,v}}) = \frac{1}{\|(\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp})^{-1}\|}.$$

Then, the result follows from Proposition 3.2. $\qquad \square$

From Lemma 3.17 and 3.18 we conclude the following result.

Proposition 3.19. *Let $(A, \lambda, v) \in \mathcal{W}$. Then*

$$\mu_v(A, \lambda, v) = \frac{\|A\|_F}{\|A - \lambda I_n\|_F} \frac{1}{\sin(d_{\mathbb{P}}(A - \lambda I_n, \Sigma_{0,v}))},$$

where $\Sigma_{0,v} = \{B \in \mathbb{P}(\mathbb{K}^{n \times n}) : Bv = 0, rank(B^2) < n - 1\} \subset \Sigma$. $\qquad \square$

Remark 3.20. From Proposition 3.19 and the fact that $\sin(\cdot) \leq 1$, we conclude that, if $(A, \lambda, v) \in \mathcal{W}$, then $\mu(A - \lambda I_n, 0, v) = \sin(d_{\mathbb{P}}(A - \lambda I_n, \Sigma_{0,v}))^{-1}$.

Proof of Theorem 4. Since (A, λ, v) and Σ'_v are included in the fiber $\mathcal{V}_v$, the distance $d_{\mathbb{P}^2}((A, \lambda, v), \Sigma'_v)$ coincides with the projective distance of $\pi_1(A, \lambda, v)$ and $\pi_1(\Sigma'_v)$, where π_1 is the canonical projection $\pi_1 : \mathcal{V} \to \mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K})$, that is,

$$(3.5) \qquad\qquad d_{\mathbb{P}^2}((A, \lambda, v), \Sigma'_v) = d_{\mathbb{P}}((A, \lambda), \pi_1(\Sigma'_v)).$$

Note that

$$\pi_1(\Sigma'_v) = \{(B, \eta) \in \mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) : (B - \eta I_n)v = 0, \text{rank}((B - \eta I_n)^2) < n - 1\}.$$

Fix a representative of $(A, \lambda, v) \in \mathcal{W}$, i.e., we assume $(A, \lambda, v) \in \widehat{\mathcal{W}}$. Let $d_{\mathbb{K}^{n \times n} \times \mathbb{K}}$ be the canonical distance on $\mathbb{K}^{n \times n} \times \mathbb{K}$.

<u>Claim:</u>

$$d_{\mathbb{K}^{n \times n} \times \mathbb{K}}((A, \lambda), \widehat{\pi_1(\Sigma'_v)}) \leq d_F(A - \lambda I_n, \widehat{\Sigma_{0,v}}),$$

where $\Sigma_{0,v}$ is defined in Lemma 3.18:

Since $\mathbb{U}_n(\mathbb{K})$ acts by isometries on $\widehat{\mathcal{V}}$ (see Remark 2.9) we may assume that $v = e_1$. Write $A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix}$. Then we have

$$(3.6) \ \ d_{\mathbb{K}^{n \times n} \times \mathbb{K}}((A, \lambda), \widehat{\pi_1(\Sigma'_v)}) = \inf\{(\|A - B\|_F^2 + |\lambda - \eta|^2)^{1/2} : (B, \eta) \in \widehat{\pi_1(\Sigma'_v)}\}.$$

If we consider the subset of pairs $(B, \eta) \in \widehat{\pi_1(\Sigma'_v)}$ such that $\eta = \lambda$, we get

$$\|A - B\|_F^2 + |\lambda - \eta|^2 = \|(A - \lambda I_n) - (B - \lambda I_n)\|_F^2,$$

where $(B - \lambda I_n)v = 0$ and $\text{rank}((B - \lambda I_n)^2) < n - 1$. Then,

$$d_{\mathbb{K}^{n \times n} \times \mathbb{K}}((A, \lambda), \widehat{\pi_1(\Sigma'_v)}) \leq d_F(A - \lambda I_n, \widehat{\Sigma_{0,v}}),$$

and the claim follows.

Now, from this claim and Lemma 3.18, we get

$$d_{\mathbb{K}^{n\times n}\times\mathbb{K}}((A,\lambda),\widehat{\pi_1(\Sigma'_v)}) \leq \frac{\|A\|_F}{\mu_v(A,\lambda,v)}.$$

Then, from (3.5) and Lemma 3.17, we conclude

$$\sin(d_{\mathbb{P}^2}((A,\lambda,v),\Sigma'_v)) \leq \frac{\|A\|_F}{(\|A\|_F^2 + |\lambda|^2)^{1/2}}\frac{1}{\mu_v(A,\lambda,v)}$$

$\square$

Remark 3.21. When we fix a representative of $(A,\lambda,v) \in \mathcal{W}$, we obtain from Proposition 3.19 that the condition number $\mu_v(A,\lambda,v)$ is comparable to the inverse of the sine of the projective distance of A to the set of ill–posed matrices such that λ is not a simple eigenvalue with eigenvector v. However, if we remove the last condition, the distance of A to the discriminant variety Σ could be much smaller. This is the case of the example (d) in Section 3.4. In that case, when ϵ is small enough, $\mu_v(B_\epsilon,\lambda_\epsilon,v_\epsilon)$ has order $\epsilon^{-1/2}$ and hence $d_{\mathbb{P}}(B_\epsilon - \lambda_\epsilon I_n, \Sigma_{0,v})$ has order $\epsilon^{1/2}$, however, the order of $d_{\mathbb{P}}(B_\epsilon,\Sigma)$ is, at least, smaller than ϵ; cf. Wilkinson [36].

3.7. **Sensitivity.** For the proof of Theorem 3 we have to study the rate of change of the condition number μ defined in Definition 3.13.

The main result of this section is the following proposition.

Proposition 3.22. *Given $\varepsilon > 0$, there exist $C_\varepsilon > 0$ such that, if (A,λ,v), (A',λ',v') belong to $\mathcal{W}$ and*

$$d_{\mathbb{P}^2}\big((A,\lambda,v),(A',\lambda',v')\big) \leq \frac{C_\varepsilon}{\mu(A,\lambda,v)},$$

then

$$\frac{\mu(A,\lambda,v)}{1+\varepsilon} \leq \mu(A',\lambda',v') \leq (1+\varepsilon)\mu(A,\lambda,v).$$

(One may choose $C_\varepsilon = \frac{\arctan\left(\frac{\varepsilon}{\sqrt{2}+\alpha(1+\varepsilon)}\right)}{(1+\varepsilon)}$, where $\alpha := (1+\sqrt{5})2\sqrt{2}$.)

Before proving Proposition 3.22 we need some additional notation.

Given $w \in \mathbb{K}^n$, $w \neq 0$, we define the linear operator

$$(3.7) \qquad \hat{\Pi}_{w^\perp} : \mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}, \quad \text{given by} \quad \hat{\Pi}_{w^\perp} B := \tau_w \circ \Pi_{w^\perp} B,$$

for every $B \in \mathbb{K}^{n\times n}$, where $\tau_w : w^\perp \hookrightarrow \mathbb{K}^n$ is the inclusion map. That is,

$$\hat{\Pi}_{w^\perp} Bz = Bz - \langle Bz, \frac{w}{\|w\|}\rangle\frac{w}{\|w\|}.$$

When $\mathbb{E}$ is a finite dimensional vector space over $\mathbb{K}$ equipped with the Hermitian inner product $\langle\cdot,\cdot\rangle$, we define

$$(3.8) \qquad\qquad\qquad d_T(w,w') := \tan(d_{\mathbb{P}}(w,w')),$$

for all $w, w' \in \mathbb{P}(\mathbb{E})$. In particular, $d_T(w,w') = \|w - w'\|_w$, whenever w and w' satisfy $\langle w - w', w\rangle = 0$.

Note that $d_{\mathbb{P}}(\cdot,\cdot) \leq d_T(\cdot,\cdot)$. Moreover, from elementary facts we have the following result.

Lemma 3.23. *Let w, $w' \in \mathbb{P}(\mathbb{E})$ such that $d_{\mathbb{P}}(w, w') \leq \theta < \pi/2$. Then*

$$d_{\mathbb{P}}(w, w') \leq d_T(w, w') \leq \frac{\tan(\theta)}{\theta} \, d_{\mathbb{P}}(w, w'), \ \ for \ all \ w, \ w' \in \mathbb{P}(\mathbb{E}). \quad \square$$

With the notation given above we have the following result.

Lemma 3.24. *Let v, $w \in \mathbb{P}(\mathbb{K}^n)$ and $B \in \mathbb{K}^{n \times n}$. Then*

$$\left\| \hat{\Pi}_{v^\perp} B - \hat{\Pi}_{w^\perp} B \right\| \leq 2\|B\| \, d_T(v, w).$$

Proof. Take representatives of v and w such that $\|v\| = 1$ and $\langle v - w, v \rangle = 0$. Let $u \in \mathbb{K}^n$, then

$$\left\| \left(\hat{\Pi}_{v^\perp} B - \hat{\Pi}_{w^\perp} B \right) u \right\| = \left\| Bu - \langle Bu, v \rangle v - \left(Bu - \langle Bu, \frac{w}{\|w\|} \rangle \frac{w}{\|w\|} \right) \right\|$$

$$= \left\| \langle Bu, \frac{w}{\|w\|} \rangle \frac{w}{\|w\|} - \langle Bu, v \rangle v \right\|$$

$$= \left\| \langle Bu, \frac{w}{\|w\|} - v \rangle \frac{w}{\|w\|} + \langle Bu, v \rangle \left(\frac{w}{\|w\|} - v \right) \right\|$$

$$\leq 2\|Bu\| \left\| \frac{w}{\|w\|} - v \right\| \leq 2\|Bu\| \, d_T(v, w).$$

$$\square$$

Notation 3.25. Given $(A, \lambda) \in \mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big)$, we denote $A_\lambda := (\lambda I_n - A)$.

Remark 3.26. Since $\left(\hat{\Pi}_{v^\perp} A_\lambda \right) v = 0$ for all $(A, \lambda, v) \in \mathcal{W}$, then $\left\| (\hat{\Pi}_{v^\perp} A_\lambda)^\dagger \right\|$ and $\left\| (\Pi_{v^\perp} A_\lambda|_{v^\perp})^{-1} \right\|$ are equal, where $\dagger$ denotes taking the Moore-Penrose inverse. Then Proposition 3.2 yields

$$(3.9) \qquad \mu_v(A, \lambda, v) = \|A\|_F \, \left\| (\hat{\Pi}_{v^\perp} A_\lambda)^\dagger \right\|.$$

Let d_{T^2} be the product function defined over $\mathbb{P}\big(\mathbb{K}^{n \times n} \times \mathbb{K}\big) \times \mathbb{P}(\mathbb{K}^n)$ by

$$d_{T^2}((A, \lambda, v), (A', \lambda', v')) := \big(d_T((A, \lambda), (A', \lambda'))^2 + d_T(v, v')^2 \big)^{1/2},$$

where d_T is given in (3.8).

Proposition 3.27. *Let $\alpha := (1 + \sqrt{5})2\sqrt{2}$. Let $(A, \lambda, v), (A', \lambda', v') \in \mathcal{W}$ such that*

$$d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big) < \frac{1}{\alpha \, \mu_v(A, \lambda, v)}.$$

Then, the following inequality holds:

$$\mu_v(A', \lambda', v') \leq \frac{\big(1 + \sqrt{2} d_{T^2}((A, \lambda, v), (A', \lambda', v'))\big) \, \mu_v(A, \lambda, v)}{1 - \alpha \, \mu_v(A, \lambda, v) \, d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big)}.$$

Proof. Consider representatives of (A, λ, v) and (A', λ', v') such that: $\|A\|_F = 1$, $\|v\| = 1$, $(A, \lambda) - (A', \lambda')$ perpendicular to (A, λ) in $\mathbb{K}^{n \times n} \times \mathbb{K}$, and $v - v'$ perpendicular to v in $\mathbb{K}^n$. From Notation 3.25 and (3.9), by Wedin's Theorem (see

Stewart–Sun [33, Theorem 3.9]) we have

$$\left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger - \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\| \le$$
$$\frac{1+\sqrt{5}}{2} \left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\| \left\| \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\| \left\| \hat{\Pi}_{v^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A'_{\lambda'} \right\|.$$

Since $\left\| \left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\| - \left\| \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\| \right\| \le \left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger - \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\|$, then,

$$\left\| \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\| \le \frac{\left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\|}{1 - \frac{1+\sqrt{5}}{2} \left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\| \left\| \hat{\Pi}_{v^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A'_{\lambda'} \right\|}.$$

Note that

$$\left\| \hat{\Pi}_{v^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A'_{\lambda'} \right\| \le \left\| \hat{\Pi}_{v^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A_\lambda \right\| + \left\| \hat{\Pi}_{v'^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A'_{\lambda'} \right\|$$
$$(3.10) \qquad\qquad \le 2 \left\| A_\lambda \right\| d_T(v, v') + \left\| A_\lambda - A'_{\lambda'} \right\|,$$

where the second inequality follows from Lemma 3.24 and the fact that the operator norm of $\hat{\Pi}_{v^\perp} : \mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}$, given in (3.7), is less or equal than one. In addition, taking into account that $(A, \lambda, v) \in \mathcal{W}$ and the choice of elected representatives, we get $\|A\|_F^2 + |\lambda|^2 \le 2$, and

$$\|A_\lambda - A'_{\lambda'}\| \le \|A - A'\| + |\lambda - \lambda'| \le \sqrt{2}(\|A - A'\|^2 + |\lambda - \lambda'|^2)^{1/2}$$
$$\le \sqrt{2}\, d_T((A, \lambda), (A', \lambda')) \sqrt{\|A\|_F^2 + |\lambda|^2}$$
$$\le 2\, d_T((A, \lambda), (A', \lambda')),$$

and hence from (3.10), and the fact $\|A_\lambda\| \le \|A\| + |\lambda|$, we get

$$\left\| \hat{\Pi}_{v^\perp} A_\lambda - \hat{\Pi}_{v'^\perp} A'_{\lambda'} \right\| \le 4\, d_T(v, v') + 2\, d_T((A, \lambda), (A', \lambda'))$$
$$\le 4(d_T(v, v') + d_T((A, \lambda), (A', \lambda')))$$
$$\le 4\sqrt{2}(d_T(v, v')^2 + d_T((A, \lambda), (A', \lambda'))^2)^{1/2}$$
$$= 4\sqrt{2}d_{T^2}((A, \lambda, v), (A', \lambda', v')).$$

Then we conclude
(3.11)

$$\left\| \left(\hat{\Pi}_{v'^\perp} A'_{\lambda'} \right)^\dagger \right\| \le \frac{\left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\|}{1 - (1 + \sqrt{5})2\sqrt{2} \left\| \left(\hat{\Pi}_{v^\perp} A_\lambda \right)^\dagger \right\| d_{T^2}((A, \lambda, v), (A', \lambda', v'))}.$$

In addition, by the triangle inequality we have $\|A'\|_F \le 1 + \|A - A'\|_F$. Then

$$\|A - A'\|_F \le \frac{\sqrt{2}}{(\|A\|_F^2 + |\lambda|^2)^{1/2}} \|A - A'\|_F \le \sqrt{2}d_T((A, \lambda), (A', \lambda')),$$

and hence $\|A'\|_F \le 1 + \sqrt{2}d_T((A, \lambda), (A', \lambda'))$. Then, the proof follows by multiplying both sides of equation (3.11) by $\|A'\|_F$. $\qquad\square$

Proposition 3.28. *Given $\varepsilon > 0$, there exist $c_\varepsilon > 0$ such that, if (A, λ, v), (A', λ', v') lie in $\mathcal{W}$ and*

$$d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big) \leq \frac{c_\varepsilon}{\mu(A, \lambda, v)},$$

then

$$\mu(A', \lambda', v') \leq (1 + \varepsilon)\mu(A, \lambda, v).$$

(One may choose $c_\varepsilon = \dfrac{\varepsilon}{\sqrt{2} + \alpha(1 + \varepsilon)}$, where $\alpha = (1 + \sqrt{5})2\sqrt{2}$.)

Proof. The condition

$$d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big) \leq \frac{c}{\mu(A, \lambda, v)},$$

implies

$$d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big) \leq \frac{c}{\mu_v(A, \lambda, v)}.$$

From Proposition 3.27 and the fact that $\mu \geq 1$, if $c < 1/\alpha$ and

$$\frac{1 + \sqrt{2}c}{1 - \alpha c} \leq 1 + \varepsilon,$$

we get

$$\mu_v(A', \lambda', v') \leq (1 + \varepsilon)\mu_v(A, \lambda, v).$$

Then

$$
\begin{aligned}
\mu(A', \lambda', v') &= \max\{1, \mu_v(A', \lambda', v')\} \\
&\leq \max\{1, (1 + \varepsilon)\mu_v(A, \lambda, v)\} \\
&\leq (1 + \varepsilon)\max\{1, \mu_v(A, \lambda, v)\} = (1 + \varepsilon)\mu(A, \lambda, v).
\end{aligned}
$$

One may choose $c_\varepsilon = \dfrac{\varepsilon}{\sqrt{2} + \alpha(1 + \varepsilon)}$. $\qquad\square$

Corollary 3.29. *Given $\varepsilon > 0$, there exist $c'_\varepsilon > 0$ such that, if (A, λ, v), (A', λ', v') lie in $\mathcal{W}$ and*

$$d_{\mathbb{P}^2}\big((A, \lambda, v), (A', \lambda', v')\big) \leq \frac{c'_\varepsilon}{\mu(A, \lambda, v)},$$

then

$$\mu(A', \lambda', v') \leq (1 + \varepsilon)\mu(A, \lambda, v).$$

(One may choose $c'_\varepsilon = \arctan\left(\frac{\varepsilon}{\sqrt{2}+\alpha(1+\varepsilon)}\right)$ where $\alpha := (1 + \sqrt{5})2\sqrt{2}$.)

Proof. By Lemma 3.23, if

$$d_{\mathbb{P}^2}\big((A, \lambda, v), (A', \lambda', v')\big) \leq \frac{c'}{\mu(A, \lambda, v)},$$

then

$$
\begin{aligned}
d_{T^2}\big((A, \lambda, v), (A', \lambda', v')\big) &\leq \frac{\tan(c')}{c'} d_{\mathbb{P}^2}\big((A, \lambda, v), (A', \lambda', v')\big) \\
&\leq \frac{\tan(c')}{\mu(A, \lambda, v)},
\end{aligned}
$$

so we just need to choose c' such that $\tan(c') \leq c_\varepsilon$ from Proposition 3.28. $\qquad\square$

Proof of Proposition 3.22. From Corollary 3.29, there exist $c' > 0$ such that, if (A, λ, v), $(A', \lambda', v') \in \mathcal{W}$ are such that

$$d_{\mathbb{P}^2}\big((A, \lambda, v), (A', \lambda', v')\big)\, \mu(A, \lambda, v) \leq c',$$

then

$$\mu(A', \lambda', v') \leq (1 + \varepsilon)\mu(A, \lambda, v).$$

It is enough to take c' such that $c' \leq \arctan\left(\frac{\varepsilon}{\sqrt{2} + \alpha(1+\varepsilon)}\right)$. In this case we have

$$d_{\mathbb{P}^2}\big((A, \lambda, v), (A', \lambda', v')\big)\, \mu(A', \lambda', v') \leq c'(1 + \varepsilon).$$

Then, by the same argument, if $c'(1 + \varepsilon) \leq \arctan\left(\frac{\varepsilon}{\sqrt{2} + \alpha(1+\varepsilon)}\right)$ we have the other inequality. $\qquad\square$

4. Newton's Method

4.1. **Introduction.** In this section we start describing the Newton method defined in Section 1.3. The main goal of this section is to prove Theorem 2.

Let us recall some definitions from the introduction.

Given a nonzero matrix $A \in \mathbb{K}^{n \times n}$, let $F_A : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$ be the evaluation map

$$F_A(\lambda, v) := (\lambda I_n - A)v.$$

This map is homogeneous of degree 1 in v. Its derivative $DF_A(\lambda, v) : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$ satisfies

$$(4.1) \qquad\qquad DF_A(\lambda, \alpha v)(\dot{\lambda}, \alpha\dot{v}) = \alpha DF_A(\lambda, v)(\dot{\lambda}, \dot{v}),$$

for all $(\dot{\lambda}, \dot{v}) \in \mathbb{K} \times \mathbb{K}^n$, and nonzero scalar α.

Definition 4.1. Given a nonzero matrix $A \in \mathbb{K}^{n \times n}$, we define the *Newton map associated to A* to be the map $N_A : \mathbb{K} \times (\mathbb{K}^n \setminus \{0\}) \to \mathbb{K} \times (\mathbb{K}^n \setminus \{0\})$ given by

$$N_A(\lambda, v) := (\lambda, v) - \big(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}\big)^{-1} F_A(\lambda, v),$$

defined for all (λ, v) such that $DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}$ is invertible.

Note that, from (4.1), the map N_A induces a map from $\mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$ into itself (defined almost everywhere); cf. Remark 1.2.

Lemma 4.2. *Let $A \in \mathbb{K}^{n \times n}$ be a nonzero matrix and $(\lambda, v) \in \mathbb{K} \times \mathbb{K}^n$, $v \neq 0$. The map N_A is well–defined at (λ, v) if and only if $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is invertible.*

Proof. The map N_A is well–defined provided that the linear operator $DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}$, from $\mathbb{K} \times v^\perp$ into $\mathbb{K}^n$, is invertible. Differentiating F_A with respect to λ and v yields

$$DF_A(\lambda, v)(\dot{\lambda}, \dot{v}) = \dot{\lambda}v + (\lambda I_n - A)\dot{v}, \quad (\dot{\lambda}, \dot{v}) \in \mathbb{K} \times \mathbb{K}^n.$$

Fix a basis of $\mathbb{K}^n$ and let $w \in \mathbb{K}^n$. Solving the linear equation $DF_A(\lambda, v)(\dot{\lambda}, \dot{v}) = w$, for $(\dot{\lambda}, \dot{v}) \in \mathbb{K} \times v^\perp$, is equivalent to solve the system of equations:

$$(4.2) \qquad \begin{pmatrix} v & \lambda I_n - A \\ 0 & v^* \end{pmatrix} \begin{pmatrix} \dot{\lambda} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} w \\ 0 \end{pmatrix}, \qquad \text{for} \quad (\dot{\lambda}, \dot{v}) \in \mathbb{K} \times \mathbb{K}^n.$$

Hence $N_A(\lambda, v)$ is well defined if and only if the matrix given in (4.2) is invertible.

Let $U \in \mathbb{U}_n(\mathbb{K})$ such that $Uv = \|v\|e_1$. Then,

$$\begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v & \lambda I_n - A \\ 0 & v^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U^* \end{pmatrix} = \begin{pmatrix} \|v\|e_1 & U(\lambda I_n - A)U^* \\ 0 & \|v\|e_1^* \end{pmatrix}.$$

Now, expanding the determinant of $\begin{pmatrix} \|v\|e_1 & U(\lambda I_n - A)U^* \\ 0 & \|v\|e_1^* \end{pmatrix}$ by the first column, and thereafter by the last row, we conclude that the matrix $\begin{pmatrix} v & \lambda I_n - A \\ 0 & v^* \end{pmatrix}$ is invertible if and only if the operator $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is invertible. $\qquad\square$

Remark 4.3. From the proof of Lemma 4.2 we obtain that N_A has the simple matrix expression

$$N_A \begin{pmatrix} \lambda \\ v \end{pmatrix} = \begin{pmatrix} \lambda \\ v \end{pmatrix} - \begin{pmatrix} v & \lambda I_n - A \\ 0 & v^* \end{pmatrix}^{-1} \begin{pmatrix} (\lambda I_n - A)v \\ 0 \end{pmatrix}.$$

Furthermore, solving the system (4.2) for $w = (\lambda I_n - A)v$, we conclude that if $\Pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}$ is invertible then the map N_A is given by $N_A(\lambda, v) = (\lambda - \dot\lambda, v - \dot v)$, where

$$\dot v = \left(\Pi_{v^\perp}(\lambda I_n - A)\big|_{v^\perp}\right)^{-1} \Pi_{v^\perp}(\lambda I_n - A)v; \qquad \dot\lambda = \frac{\langle(\lambda I_n - A)(v - \dot v), v\rangle}{\langle v, v\rangle}.$$

Definition 4.4. Let $A \in \mathbb{K}^{n \times n}$ be a nonzero matrix, and let (λ_0, v_0) in $\mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. We say that the triple (A, λ_0, v_0) is an *approximate solution* of the eigenvalue problem $(A, \lambda, v) \in \mathcal{V}$, if the sequence $(A, N_A^k(\lambda_0, v_0))$, $k = 0, 1, \ldots$ is defined and satisfies

$$d_{\mathbb{P}^2}\left((A, N_A^k(\lambda_0, v_0)), (A, \lambda, v)\right) \leq \left(\frac{1}{2}\right)^{2^k - 1} d_{\mathbb{P}^2}\left((A, \lambda_0, v_0), (A, \lambda, v)\right),$$

for all positive integers k.

Recall from Remark 1.1 that the notion of approximate solution, and the sequence $(A, N_A^k(\lambda_0, v_0))$, $k = 0, 1, \ldots$, are well–defined on $\mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K}) \times \mathbb{P}(\mathbb{K}^n)$.

4.2. **Approximate Solution Theorem.** The main tool to prove Theorem 2 is the following result.

Proposition 4.5. *Let* $0 < c \leq 1/(2\sqrt{2})$. *Let* $A \in \mathbb{K}^{n \times n}$ *such that* $\|A\|_F = 1$, *and let* $(\lambda, v), (\lambda_0, v_0) \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. *If* $(A, \lambda, v) \in \mathcal{W}$ *and*

$$(|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2} < \frac{c}{\mu(A, \lambda, v)},$$

then, the sequence $(\lambda_k, v_k) := N_A^k(\lambda_0, v_0)$ *satisfies*

$$(|\lambda_k - \lambda|^2 + d_{\mathbb{P}}(v_k, v)^2)^{1/2} \leq \left(\frac{2\tan(c)}{1 - \sqrt{2}\,c}\right) \left(\frac{1}{2}\right)^{2^k - 1} (|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2},$$

for all positive integers k.

(Since we do not find an appropriate version in the literature to cite, we include a proof of this proposition in the appendix of this paper.)

Remark 4.6. Some expressions given in Proposition 4.5 are not scale invariant in (A, λ) (and (A, λ_0)), and thus a restriction on $\|A\|$ is required.

Picking c in Proposition 4.5 such that $0 < c \leq 1/(2\sqrt{2})$ and $2\tan(c)/(1 - \sqrt{2}c) \leq 1$ we have the following result, which is interesting *per se*.

Theorem 5. *There is a universal constant $c_0 > 0$ with the following property. Let $A \in \mathbb{K}^{n \times n}$ such that $\|A\|_F = 1$, and let $(\lambda, v), (\lambda_0, v_0) \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. If $(A, \lambda, v) \in \mathcal{W}$ and*

$$(|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2} < \frac{c_0}{\mu(A, \lambda, v)},$$

then the sequence $(\lambda_k, v_k) := N_A^k(\lambda_0, v_0)$ satisfies

$$(|\lambda_k - \lambda|^2 + d_{\mathbb{P}}(v_k, v)^2)^{1/2} \leq \left(\frac{1}{2}\right)^{2^k - 1} (|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2},$$

for all positive integers k. (One may choose $c_0 = 0.288$.) $\qquad\square$

This theorem is a version, for the map $N_A : \mathbb{K} \times \mathbb{P}(\mathbb{K}^n) \to \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$, of a well–known theorem in the literature, namely, the *Smale γ-Theorem* (or *Approximate Solution Theorem*), which gives the size of the basin of attraction of Newton's method; see Blum et al. [8, Theorem 1, pp. 263].

4.3. **Proof of Theorem 2.** For the proof of Theorem 2 we need a technical lemma. Its proof is included in the appendix.

Lemma 4.7. *Let $A \in \mathbb{K}^{n \times n}$ such that $\|A\|_F = 1$, and let $(\lambda, v), (\lambda', v') \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$.*

    (1) *If $|\lambda - \lambda'| \leq c < \sqrt{2}$, then,*

$$d_{\mathbb{P}^2}\big((A, \lambda, v), (A, \lambda', v')\big) \leq \beta_c \left(|\lambda - \lambda'|^2 + d_{\mathbb{P}}(v, v')^2\right)^{1/2},$$

        where $\beta_c = (1 - c^2/2)^{-1/2}$.

    (2) *If $d_{\mathbb{P}^2}\big((A, \lambda, v), (A, \lambda', v')\big) < \theta < \pi/4$, then,*

$$(|\lambda - \lambda'|^2 + d_T(v, v')^2)^{1/2} \leq R_\theta \, d_{\mathbb{P}^2}((A, \lambda.v), (A, \lambda', v')),$$

        where $R_\theta = [\sqrt{2}/\cos(\theta + \pi/4)^3]^{1/2}$.

Let θ_0 such that $R_{\theta_0} \theta_0 = 1/(2\sqrt{2})$, where R_θ is given in Lemma 4.7 ($\theta_0 \approx 0.1389$).

Proposition 4.8. *Let $0 < c \leq \theta_0$. Let $A \in \mathbb{K}^{n \times n}$ be a nonzero matrix, and let $(\lambda, v), (\lambda_0, v_0) \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. If $(A, \lambda, v) \in \mathcal{W}$ and*

$$d_{\mathbb{P}^2}\big((A, \lambda_0, v_0), (A, \lambda, v)\big) < \frac{c}{\mu(A, \lambda, v)},$$

then the sequence $(A, N_A^k(\lambda_0, v_0))$, $k = 0, 1, \dots$ satisfies

$$d_{\mathbb{P}^2}\big((A, N_A^k(\lambda_0, v_0)), (A, \lambda, v)\big) \leq$$

$$\leq R_c \, \beta_{c\,R_c} \left(\frac{2\tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \left(\frac{1}{2}\right)^{2^k - 1} d_{\mathbb{P}^2}\big((A, \lambda_0, v_0), (A, \lambda, v)\big),$$

for all positive integers k, where $\delta(c) := c/(1 - c)$.

Proof. From Remark 1.2, we may assume $\|A\|_F = 1$.

Since $d_{\mathbb{P}}(\cdot) \leq d_T(\cdot)$, by Lemma 4.7–(2), one has

$$(4.3) \qquad\qquad (|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2} \leq \frac{cR_c}{\mu(A, \lambda, v)}.$$

Since $c \leq \theta_0$, we have $cR_c \leq 1/(2\sqrt{2})$, and then Proposition 4.5 yields

$$(4.4) \quad (|\lambda_k - \lambda|^2 + d_{\mathbb{P}}(v_k, v)^2)^{1/2} \leq$$

$$\left(\frac{2 \tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \left(\frac{1}{2}\right)^{2^k - 1} (|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2},$$

for all $k > 0$, where $(\lambda_k, v_k) := N_A^k(\lambda_0, v_0)$. Since $(|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2} < c\,R_c$, we deduce from Lemma 4.7 and (4.4) that

$$d_{\mathbb{P}^2}((A, \lambda_k, v_k), (A, \lambda, v)) \leq$$

$$\leq \beta_{c\,R_c} \left(\frac{2 \tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \left(\frac{1}{2}\right)^{2^k - 1} (|\lambda_0 - \lambda|^2 + d_{\mathbb{P}}(v_0, v)^2)^{1/2}$$

$$\leq R_c\,\beta_{c\,R_c} \left(\frac{2 \tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \left(\frac{1}{2}\right)^{2^k - 1} d_{\mathbb{P}^2}((A, \lambda_0, v_0), (A, \lambda, v)).$$

(Note that $c \leq \theta_0 < \frac{\pi}{4}$.) □

4.3.1. *Proof of Theorem 2.*

Proof of Theorem 2. From Proposition 4.8, proof of Theorem 2 follows picking $c_0 > 0$ such that $c_0 \leq \theta_0$ and $R_{c_0}\,\beta_{c_0\,R_{c_0}} \left(\frac{2 \tan(c_0\,R_{c_0})}{1 - \sqrt{2}\,c_0\,R_{c_0}}\right) \leq 1$. (One may choose $c_0 = 0.0739$.) □

5. Complexity bound

5.1. Condition length. Let us start recalling some basic definition.

Let $\mathbb{E}$ be a finite dimensional Hilbert space. A function $\alpha : [0, 1] \to \mathbb{E}$ is an *absolutely continuous path* if it is almost everywhere differentiable, its derivative $\dot{\alpha}(t)$ is an integrable function, and

$$\alpha(t) = \alpha(0) + \int_0^t \dot{\alpha}(s)\,ds.$$

We say that the (projective) path $\alpha : [0, 1] \to \mathbb{P}(\mathbb{E})$ is an absolutely continuous path if it is the projection, under the quotient canonical map $\mathbb{E} \setminus \{0\} \to \mathbb{P}(\mathbb{E})$, of an absolutely continuous path in $\mathbb{E} \setminus \{0\}$.

Let us recall some definition from the introduction.

Definition 5.1. The *condition length* of an absolutely continuous path $\Gamma : [0, 1] \to \mathcal{W}$ is defined by

$$\ell_\mu(\Gamma) := \int_a^b \|\dot{\Gamma}(t)\|_{\Gamma(t)}\,\mu(\Gamma(t))\,dt.$$

The next proposition is useful for the proof of our main theorem.

Proposition 5.2. *Given $\varepsilon > 0$, $C_\varepsilon > 0$ as in Proposition 3.22, and $\Gamma : [0, 1] \to \mathcal{W}$ an absolutely continuous path (with $\ell_\mu(\Gamma) < \infty$), define the sequence $t_0, t_1, \ldots$ in $[0, 1]$ such that:*
- *$t_0 = 0$;*
- *t_k such that $\mu(\Gamma(t_{k-1})) \int_{t_{k-1}}^{t_k} \|\dot{\Gamma}(s)\|_{\Gamma(s)} ds = C_\varepsilon$,*
whenever $\mu(\Gamma(t_{k-1})) \int_{t_{k-1}}^{1} \|\dot{\Gamma}(s)\|_{\Gamma(s)} ds > C_\varepsilon$;

• *else define $t_k = t_K = 1$.*
Then,

$$K \leq \frac{1+\varepsilon}{C_\varepsilon} \ell_\mu(\Gamma) + 1.$$

Proof. Whenever $k < K$ (where $K \in \mathbb{N} \cup \{\infty\}$), given $t \in [t_{k-1}, t_k]$,

$$d_{\mathbb{P}^2}(\Gamma(t_{k-1}), \Gamma(t)) \leq \int_{t_{k-1}}^{t_k} \|\dot{\Gamma}(s)\|_{\Gamma(s)} ds = \frac{C_\varepsilon}{\mu(\Gamma(t_{k-1}))}.$$

By the first inequality in Proposition 3.22, we get

$$\int_{t_{k-1}}^{t_k} \|\dot{\Gamma}(s)\|_{\Gamma(s)} \mu(\Gamma(s)) \, ds \geq \frac{\mu(\Gamma(t_{k-1}))}{1+\varepsilon} \int_{t_{k-1}}^{t_k} \|\dot{\Gamma}(s)\|_{\Gamma(s)} dt = \frac{C_\varepsilon}{1+\varepsilon}.$$

Since $\ell_\mu(\Gamma) < \infty$, then $K < \infty$, and adding, yields

$$\ell_\mu(\Gamma) \geq (K-1)\frac{C_\varepsilon}{1+\varepsilon}.$$

$\square$

5.2. **Proof of Theorem 3.**

Proof of Theorem 3: Since $\Gamma(t) = (A(t), \lambda(t), v(t))$, $0 \leq t \leq 1$, is an absolutely continuous path in $\mathcal{W}$, we may assume that the path $A(t)$, in $\mathbb{K}^{n \times n} \setminus \{0_n\}$, is absolutely continuous. In addition, without loss of generality we may assume that $\|A(t)\|_F = 1$, for every $t \in [0, 1]$.

The idea of the proof is to show that the mesh $0 = t_0 < t_1 < \cdots < t_K = 1$ given in Proposition 5.2, for some $\varepsilon > 0$ to be defined afterwards, guarantees that the predictor–corrector sequence $\hat{\Gamma}(t_k) := (A(t_k), \lambda_k, v_k)$, where

$$(\lambda_{k+1}, v_{k+1}) := N_{A(t_{k+1})}(\lambda_k, v_k), \quad 0 \leq k \leq K - 1,$$

approximates the path Γ provided that $(A(0), \lambda_0, v_0)$ is an approximate solution of $\Gamma(0)$.

The proof is by induction.

Let $\varepsilon > 0$, and let C_ε as in Proposition 3.22. Assume that $\Gamma(t_k)$, $\hat{\Gamma}(t_k)$, $\Gamma(t_{k+1})$ are such that,

$$d_{\mathbb{P}^2}(\Gamma(t_k), \Gamma(t_{k+1})) \leq \frac{C_\varepsilon}{\mu(\Gamma(t_k))}, \quad \text{and} \quad d_{\mathbb{P}^2}(\Gamma(t_k), \hat{\Gamma}(t_k)) \leq \frac{C_\varepsilon}{\mu(\Gamma(t_k))}.$$

Then,

$$d_{\mathbb{P}^2}(\Gamma(t_{k+1}), (A(t_{k+1}), \lambda_k, v_k)) \leq$$
$$\leq d_{\mathbb{P}^2}(\Gamma(t_{k+1}), \Gamma(t_k)) + d_{\mathbb{P}^2}(\Gamma(t_k), \hat{\Gamma}(t_k)) + d_{\mathbb{P}^2}(\hat{\Gamma}(t_k), (A(t_{k+1}), \lambda_k, v_k))$$
$$< \frac{2C_\varepsilon}{\mu(\Gamma(t_k))} + d_{\mathbb{P}^2}(\hat{\Gamma}(t_k), (A(t_{k+1}), \lambda_k, v_k)).$$

Note that

$$d_{\mathbb{P}^2}(\hat{\Gamma}(t_k), (A(t_{k+1}), \lambda_k, v_k)) = d_{\mathbb{P}}((A(t_k), \lambda_k), (A(t_{k+1}), \lambda_k)).$$

Claim: One has

$$d_{\mathbb{P}}((A(t_k), \lambda_k), (A(t_{k+1}), \lambda_k)) \leq d_{\mathbb{P}}(A(t_k), A(t_{k+1})) :$$

For the ease of notation let us denote $a := A(t_k)$, $a' := A(t_{k+1})$, $\lambda := \lambda_k$, $\theta_0 := d_{\mathbb{P}}(a, a')$, and $\theta_\lambda := d_{\mathbb{P}}((a, \lambda), (a', \lambda))$. Since $\|A(t)\|_F = 1$, $0 \leq t \leq 1$, we have $\|a\|_F = \|a'\|_F = 1$.

In addition, by the law of cosines we have that $\cos\theta_0 = 1 - \|a - a'\|_F^2/2$, and $\cos\theta_\lambda = 1 - \|a - a'\|_F^2/(2(1 + |\lambda|^2))$. Then $\cos\theta_0 \le \cos\theta_\lambda$. Since $\theta_0, \theta_\lambda \in [0, \pi]$, we conclude $\theta_\lambda \le \theta_0$.

Furthermore,

$$d_\mathbb{P}(A(t_k), A(t_{k+1})) \le \int_{t_k}^{t_{k+1}} \|\dot{A}(s)\|_{A(s)}\, ds$$

$$\le \sqrt{2} \int_{t_k}^{t_{k+1}} \|D\mathscr{S}_\lambda(\Gamma(t))\dot{A}(s)\|_{(A(s), \lambda(s))}\, ds$$

$$\le \sqrt{2} \int_{t_k}^{t_{k+1}} \|\dot{\Gamma}(s)\|_{\Gamma(s)}\, ds,$$

where the second inequality follows from the trivial lower bound which one may obtain from (3.3) and the assumption $\|A(s)\|_F = 1$ (and hence $\langle \dot{A}(s), A(s)\rangle_F = 0$).

Since, by construction, $\int_{t_k}^{t_{k+1}} \|\dot{\Gamma}(s)\|_{\Gamma(s)}\, ds \le C_\varepsilon/\mu(\Gamma(t_k))$ we conclude

$$d_{\mathbb{P}^2}(\Gamma(t_{k+1}), (A(t_{k+1}), \lambda_k, v_k)) < \frac{(2 + \sqrt{2})C_\varepsilon}{\mu(\Gamma(t_k))}.$$

Furthermore, since $d_{\mathbb{P}^2}(\Gamma(t_k), \Gamma(t_{k+1})) < C_\varepsilon\mu(\Gamma(t_k))^{-1}$, Proposition 3.22 yields

$$d_{\mathbb{P}^2}(\Gamma(t_{k+1}), (A(t_{k+1}), \lambda_k, v_k)) \le \frac{(1 + \varepsilon)(2 + \sqrt{2})C_\varepsilon}{\mu(\Gamma(t_{k+1}))}.$$

From Proposition 4.8, if $c := (1 + \varepsilon)C_\varepsilon(2 + \sqrt{2}) \le \theta_0$, then

$$d_{\mathbb{P}^2}(\hat{\Gamma}(t_{k+1}), \Gamma(t_{k+1})) \le$$

$$\le R_c\, \beta_{c\,R_c} \left(\frac{2\tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \frac{1}{2}\, d_{\mathbb{P}^2}((A(t_{k+1}), \lambda_k, v_k), \Gamma(t_{k+1}))$$

$$\le \frac{R_c\, \beta_{c\,R_c} \left(\frac{2\tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \frac{1}{2}\, c}{\mu(\Gamma(t_{k+1}))}.$$

Then, if ε is small enough such that $c \le \theta_0$ and $R_c\, \beta_{c\,R_c} \left(\frac{2\tan(c\,R_c)}{1 - \sqrt{2}\,c\,R_c}\right) \frac{1}{2}\, c < C_\varepsilon$, we get

$$d_{\mathbb{P}^2}(\Gamma(t_{k+1}), \hat{\Gamma}(t_{k+1})) \le \frac{C_\varepsilon}{\mu(\Gamma(t_k))}.$$

Moreover, if the ε picked above also satisfies $C_\varepsilon \le c_0$ (where c_0 is given in Theorem 2), then we have concluded that $\hat{\Gamma}(t_{k+1})$ is an approximate solution of $\Gamma(t_{k+1})$ provided that $\hat{\Gamma}(t_k)$ is an approximate solution of $\Gamma(t_k)$. This just finishes the induction step. (One can choose $\varepsilon = 0.1640$, $C_\varepsilon \approx 0.01167$, and $C = 100$.) $\square$

6. Appendix

This section is divided in two parts. In the first part we include a proof of Proposition 4.5. In the second, we prove Lemma 4.7.

6.1. **Proof of Proposition 4.5.** Given a nonzero matrix $A \in \mathbb{K}^{n\times n}$, recall that evaluation map $F_A : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$ is given by $F_A(\lambda, v) := (\lambda I_n - A)v$.

Throughout this section, we consider the canonical Hermitian structure on $\mathbb{K} \times \mathbb{K}^n$.

6.1.1. *Preliminaries and technical lemmas.* The next result follows by elementary computations.

Lemma 6.1. *Let v, $v' \in \mathbb{P}(\mathbb{K}^n)$ such that $d_{\mathbb{P}}(v, v') < \pi/2$. Let $\Pi_{v^\perp}|_{v'^\perp} : v'^\perp \to v^\perp$ be the restriction of the orthogonal projection $\Pi_{v^\perp}$ of $\mathbb{K}^n$ onto $v'^\perp$. Then,*

$$\| (\Pi_{v^\perp}|_{v'^\perp})^{-1} \| = \frac{1}{\cos(d_{\mathbb{P}}(v, v'))}. \quad \square$$

(In the preceding lemma, we consider the spaces $v^\perp$ and $v'^\perp$ as subspaces of $\mathbb{K}^n$ with the canonical Hermitian structure.)

Lemma 6.2. *Let $(A, \lambda, v) \in \mathcal{W}$ and $v' \in \mathbb{P}(\mathbb{K}^n)$ such that $d_{\mathbb{P}}(v, v') < \pi/2$.*

(i) *For every $(\dot{\lambda}, \dot{v}) \in \mathbb{K} \times v^\perp$ we have*

$$\left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}(\dot{\lambda}, \dot{v}) = (\dot{\lambda}, (\Pi_{v^\perp}|_{v'^\perp})^{-1}(\dot{v})).$$

(ii)

$$\| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp} \| = \frac{1}{\cos(d_{\mathbb{P}}(v, v'))};$$

(iii)

$$\| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} \| \leq \frac{\| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp} \right)^{-1} \|}{\cos(d_{\mathbb{P}}(v, v'))}.$$

Proof. (i): Given $(\dot{\lambda}, \dot{v}) \in \mathbb{K} \times v^\perp$, let $(\dot{\eta}, \dot{w}) \in \mathbb{K} \times v'^\perp$ such that

$$(\dot{\eta}, \dot{w}) = \left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}(\dot{\lambda}, \dot{v}).$$

Then,

$$\dot{\eta}v + (\lambda I_n - A)\dot{w} = \dot{\lambda}v + (\lambda I_n - A)\dot{v}.$$

Since $(A, \lambda, v) \in \mathcal{W}$, we deduce that $\dot{\eta} = \dot{\lambda}$ and $\Pi_{v^\perp}\dot{w} = \dot{v}$. Then,

$$\left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp}(\dot{\lambda}, \dot{v}) = \left(\dot{\lambda}, (\Pi_{v^\perp}|_{v'^\perp})^{-1}(\dot{v}) \right).$$

(ii): Taking the canonical norm of $\mathbb{K} \times \mathbb{K}^n$ in (i), and maximizing on the unit sphere in $\mathbb{K} \times v^\perp \subset \mathbb{K} \times \mathbb{K}^n$, the assertion (ii) follows from Lemma 6.1.

(iii): Note that

$$\| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} \| \leq$$
$$\| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v'^\perp} \right)^{-1} DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp} \| \, \| \left(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp} \right)^{-1} \|,$$

then the result follows from (ii). $\qquad \square$

Lemma 6.3. *Let $A \in \mathbb{K}^{n \times n}$, and $(\lambda, v) \in \mathbb{K} \times \mathbb{K}^n$. Then $\|D^2 F_A(\lambda, v)\| \leq 1$.*

Proof. Differentiating F_A twice, we get

$$D^2 F_A(\lambda, v)(\dot{\lambda}, \dot{v})(\dot{\eta}, \dot{u}) = \dot{\lambda}\dot{u} + \dot{\eta}\dot{v}, \quad \text{for all} \quad (\dot{\lambda}, \dot{v}), (\dot{\eta}, \dot{u}) \in \mathbb{K} \times \mathbb{K}^n.$$

Then,

$$\|D^2 F_A(\lambda, v)(\dot{\lambda}, \dot{v})(\dot{\eta}, \dot{u})\| \leq |\dot{\lambda}| \, \|\dot{u}\| + \|\dot{v}\| \, |\dot{\eta}|$$
$$\leq (|\dot{\lambda}|^2 + \|\dot{v}\|^2)^{1/2} \, (|\dot{\eta}|^2 + \|\dot{u}\|^2)^{1/2},$$

where the second inequality follows from Cauchy-Schwarz inequality. $\qquad \square$

Recall Neumann's series result (see for example Stewart–Sun [33]):

Lemma 6.4. *Let $\mathbb{E}$ be a Hermitian space, and $A, I_{\mathbb{E}} : \mathbb{E} \to \mathbb{E}$ be linear operators where $I_{\mathbb{E}}$ is the identity. If $\|A - I_{\mathbb{E}}\| < 1$, then A is invertible and*

$$\|A^{-1}\| \leq \frac{1}{1 - \|A - I_{\mathbb{E}}\|}. \quad \square$$

Proposition 6.5. *Let $0 < c \leq 1/(2\sqrt{2})$.*
Let $A \in \mathbb{K}^{n \times n}$ and $(\lambda, v) \in \mathbb{K} \times \mathbb{K}^n$, such that $\|v\| = 1$ and $(A, \lambda, v) \in \mathcal{W}$. Let $(\lambda_0, v_0) \in \mathbb{K} \times \mathbb{P}(\mathbb{K}^n)$. If

$$(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2} < \frac{c}{\|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\|},$$

then the sequence $(\lambda_k, v_k) := N_A^k(\lambda_0, v_0)$ satisfies

$$(|\lambda_k - \lambda|^2 + d_T(v_k, v)^2)^{1/2} \leq \sqrt{2}\,\delta(\sqrt{2}\,c)\left(\frac{1}{2}\right)^{2^k - 1}(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2},$$

for all positive integers k, where $\delta(c) := c/(1 - c)$.

Proof. Take a representative of v_0 such that $\langle v - v_0, v_0 \rangle = 0$. Thus we have $\|v_0\|\, d_T(v, v_0) = \|v - v_0\|$ and $\|v_0\| \leq 1$.

In particular, the hypothesis implies that

$$\|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\|\,\|(\lambda_0 - \lambda, v_0 - v)\| < c.$$

Taylor's expansion of F_A and DF_A in a neighbourhood of (λ, v) are given by

$$(6.1) \qquad F_A(\lambda', v') = DF_A(\lambda, v)(\lambda' - \lambda, v' - v) + \frac{1}{2}D^2F_A(\lambda, v)(\lambda' - \lambda, v' - v)^2,$$

and

$$DF_A(\lambda', v') = DF_A(\lambda, v) + D^2F_A(\lambda, v)(\lambda' - \lambda, v' - v).$$

One has

$$\left(DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp}\right)^{-1} DF_A(\lambda_0, v_0)|_{\mathbb{K} \times v_0^\perp} - I_{\mathbb{K} \times v_0^\perp} =$$

$$= \left(DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp}\right)^{-1}\left(DF_A(\lambda_0, v_0)|_{\mathbb{K} \times v_0^\perp} - DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp}\right)$$

$$= \left(DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp}\right)^{-1} D^2F_A(\lambda, v))(\lambda_0 - \lambda, v_0 - v)|_{\mathbb{K} \times v_0^\perp}.$$

Then, taking norms, we get

$$\left\|\left(DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp}\right)^{-1} DF_A(\lambda_0, v_0)|_{\mathbb{K} \times v_0^\perp} - I_{\mathbb{K} \times v_0^\perp}\right\| \leq$$

$$\leq \left\|(DF_A(\lambda, v)|_{\mathbb{K} \times v_0^\perp})^{-1}\right\|\,\left\|D^2F_A(\lambda, v))(\lambda_0 - \lambda, v_0 - v)\right\|$$

$$\leq \frac{1}{\cos(d_{\mathbb{P}}(v, v_0))}\left\|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\right\|\,\left\|(\lambda_0 - \lambda, v_0 - v)\right\|,$$

where the last inequality follows from Lemma 6.2 and Lemma 6.3.
In the range of angles under consideration we have $\|v_0\| = \cos(d_{\mathbb{P}}(v, v_0)) \geq 1/\sqrt{2}$.
Then, by the condition $0 < c \leq 1/(2\sqrt{2})$, we can deduce from Lemma 6.4 that

$DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp}$ is invertible and

(6.2) $\left\| \left(DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp} \right)^{-1} DF_A(\lambda, v)\big|_{\mathbb{K} \times v_0^\perp} \right\| \le$

$$\le \frac{1}{1 - \frac{1}{\cos(d_{\mathbb{P}}(v, v_0))} \, \|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\| \, \|(\lambda_0 - \lambda, v_0 - v)\|}.$$

Furthermore,

$$N_A(\lambda_0, v_0) - (\lambda, v) =$$

$$= (\lambda_0 - \lambda, v_0 - v) - \left(DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp} \right)^{-1} F_A(\lambda_0, v_0)$$

$$= \left(DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp} \right)^{-1}$$

$$\left(DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp} (\lambda_0 - \lambda, v_0 - v) - F_A(\lambda_0, v_0) \right).$$

Then, from (6.1) we get

$$N_A(\lambda_0, v_0) - (\lambda, v) =$$

$$= \frac{1}{2} \left(DF_A(\lambda_0, v_0)\big|_{\mathbb{K} \times v_0^\perp} \right)^{-1} D^2 F_A(\lambda, v)(\lambda_0 - \lambda, v_0 - v)^2.$$

Taking the canonical norm in $\mathbb{K} \times \mathbb{K}^n$, we get

$$\|N_A(\lambda_0, v_0) - (\lambda, v)\| \le$$

$$\le \frac{1}{2} \left\| (DF_A(\lambda_0, v_0)|_{\mathbb{K} \times v_0^\perp})^{-1} \right\| \, \|D^2 F_A(\lambda, v)(\lambda_0 - \lambda, v_0 - v)^2\|.$$

Then, from (6.2) and Lemma 6.2,

(6.3) $\|N_A(\lambda_0, v_0) - (\lambda, v)\| \le$

$$\le \frac{\sqrt{2} \, \left\| (DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1} \right\| \, \frac{1}{2} \|D^2 F_A(\lambda, v)(\lambda_0 - \lambda, v_0 - v)^2\|}{1 - \sqrt{2} \, \|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\| \, \|(\lambda_0 - \lambda, v_0 - v)\|}.$$

Therefore Lemma 6.3 yields

$$\|N_A(\lambda_0, v_0) - (\lambda, v)\| \le$$

$$\le \frac{\sqrt{2} \, \left\| (DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1} \right\| \, \|(\lambda_0 - \lambda, v_0 - v)\|}{1 - \sqrt{2} \, \|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\| \, \|(\lambda_0 - \lambda, v_0 - v)\|} \, \frac{1}{2} \|(\lambda_0 - \lambda, v_0 - v)\|.$$

Then,

$$\|N_A(\lambda_0, v_0) - (\lambda, v)\| \le$$

$$\le \frac{\sqrt{2} \, \left\| (DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1} \right\| \, (|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}}{1 - \sqrt{2} \, \|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\| \, (|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}}$$

$$\frac{1}{2} (|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}.$$

Let $(\lambda_1, v_1) := N_A(\lambda_0, v_0)$.
From the proof of Lemma 6.3 we have $D^2 F_A(\lambda, v)(\lambda_0 - \lambda, v_0 - v)^2 = 2(\lambda_0 - \lambda)(v_0 - v)$, then, from (6.3) one can deduce that $\|v_1 - v\| < \delta(\sqrt{2}\,c)\|v_0 - v\|$,

where $\delta(c) = c/(1-c)$. Since $c \leq 1/(2\sqrt{2})$, we have $\delta(\sqrt{2}\,c) \leq 1$, then from Lemma 2–(4) of Blum et al. [8, pp. 264] we get

$$d_T(v_1, v) \leq \frac{\|v_1 - v\|}{\|v_0\|} \leq \sqrt{2}\,\|v_1 - v\|.$$

Hence

$$(6.4) \quad (|\lambda_1 - \lambda|^2 + d_T(v_1, v)^2)^{1/2} \leq$$

$$\leq \frac{2\,\left\|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\right\|\,(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}}{1 - \sqrt{2}\,\|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\|\,(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}}$$

$$\frac{1}{2}\,(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}.$$

Therefore

$$(6.5) \qquad (|\lambda_1 - \lambda|^2 + d_T(v_1, v)^2)^{1/2} \leq \sqrt{2}\,\delta(\sqrt{2}\,c)\,\frac{1}{2}\,(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2}.$$

From (6.5), (6.4), and the fact that $\delta(\sqrt{2}\,c) \leq 1$, working by induction we get

$$(|\lambda_k - \lambda|^2 + d_T(v_k, v)^2)^{1/2} \leq \sqrt{2}\,\delta(\sqrt{2}\,c)\,\left(\frac{1}{2}\right)^{2^k - 1}\,(|\lambda_0 - \lambda|^2 + d_T(v_0, v)^2)^{1/2},$$

for all $k > 0$, where $(\lambda_k, v_k) := N_A^k(\lambda_0, v_0)$. $\qquad\square$

Proposition 6.6. *Let $(A, \lambda, v) \in \mathcal{W}$, such that $\|A\|_F = 1$ and $\|v\| = 1$. Then,*

$$\mu(A, \lambda, v) \leq \|(DF_A(\lambda, v)|_{\mathbb{K} \times v^\perp})^{-1}\| \leq 2\,\mu(A, \lambda, v).$$

Proof. Since the action of $\mathbb{U}_n(\mathbb{K})$ on $\mathbb{P}(\mathbb{K}^n)$ is transitive, by Remark 2.9, we may assume that $v = e_1$, where $e_1, \ldots, e_n$ is the canonical basis of $\mathbb{K}^n$. Then in this basis we have.

$$A = \begin{pmatrix} \lambda & a \\ 0 & \hat{A} \end{pmatrix},$$

where $a \in \mathbb{K}^{1 \times (n-1)}$, and $\hat{A} = \Pi_{e_1^\perp} A|_{e_1^\perp} \in \mathbb{K}^{(n-1) \times (n-1)}$.

Recall that $DF_A(\lambda, e_1)(\dot{\lambda}, \dot{v}) = \dot{\lambda}e_1 + (\lambda I_n - A)\dot{v}$. Then in the basis $(1, 0)$, $(0, e_2), \ldots, (0, e_n)$ of $\mathbb{K} \times e_1^\perp$ we have

$$DF_A(\lambda, e_1)|_{\mathbb{K} \times e_1^\perp} = \begin{pmatrix} 1 & -a \\ 0 & \lambda I_{n-1} - \hat{A} \end{pmatrix}.$$

Note that $(DF_A(\lambda, e_1)|_{\mathbb{K} \times e_1^\perp})^{-1} = \begin{pmatrix} 1 & a(\lambda I_{n-1} - \hat{A})^{-1} \\ 0 & (\lambda I_{n-1} - \hat{A})^{-1} \end{pmatrix}$. Hence

$$\|(DF_A(\lambda, e_1)|_{\mathbb{K} \times e_1^\perp})^{-1}\| \geq \max\{1, \|(\lambda I_{n-1} - \hat{A})^{-1}\|\} = \mu(A, \lambda, e_1).$$

On the other hand,

$$\|(DF_A(\lambda, e_1)|_{\mathbb{K} \times e_1^\perp})^{-1}\| \leq$$

$$\leq \left\|\begin{pmatrix} 1 & a(\lambda I_{n-1} - \hat{A})^{-1} \\ 0 & 0 \end{pmatrix}\right\| + \left\|\begin{pmatrix} 0 & 0 \\ 0 & (\lambda I_{n-1} - \hat{A})^{-1} \end{pmatrix}\right\|$$

$$\leq \max\{1, \|a(\lambda I_{n-1} - \hat{A})^{-1}\|\} + \|(\lambda I_{n-1} - \hat{A})^{-1}\|$$

$$\leq 2\,\mu(A, \lambda, e_1),$$

where the last inequality follows from $\|a\| \leq \|A\|_F = 1$. $\qquad\square$

Remark 6.7. In the last proposition the result may be not longer true when $\|A\|_F = 1$ or $\|v\| = 1$ are not satisfied.

6.1.2. *Proof of Proposition 4.5.*

Proof of Proposition 4.5. Pick a representative of v and v_0 such that $\|v\| = 1$ and $\langle v - v_0, v_0 \rangle = 0$. Then the proof follows directly from Proposition 6.5, Proposition 6.6 and Lemma 3.23. □

6.2. **Proof of Lemma 4.7.**

Lemma 6.8. *Let $A \in \mathbb{K}^{n \times n}$ such that $\|A\|_F = 1$. Let λ, $\lambda' \in \mathbb{K}$ such that $|\lambda| \leq 1$.*

(1) *If $|\lambda' - \lambda| \leq c$ for some $0 \leq c < \sqrt{2}$, then there exists $\beta_c > 1$ such that*

$$d_{\mathbb{P}}((A, \lambda), (A, \lambda)) \leq \beta_c \, |\lambda' - \lambda|.$$

One may choose $\beta_c = (1 - c^2/2)^{-1/2}$.

(2) *If $d_{\mathbb{P}}((A, \lambda), (A, \lambda')) \leq \hat{\theta}$ for some $0 \leq \hat{\theta} < \pi/4$, then there exist $R_{\hat{\theta}} > 1$ such that*

$$|\lambda' - \lambda| \leq R_{\hat{\theta}} \, d_{\mathbb{P}}((A, \lambda), (A, \lambda')).$$

One may choose $R_{\hat{\theta}} = [\sqrt{2}/\cos(\hat{\theta} + \pi/4)^3]^{1/2}$.

Proof. Let $\theta := d_{\mathbb{P}}((A, \lambda), (A, \lambda'))$. (Thus $0 \leq \theta \leq \hat{\theta}$.) By the law of cosines we know that

$$|\lambda - \lambda'|^2 = 1 + |\lambda|^2 + 1 + |\lambda'|^2 - 2\sqrt{1 + |\lambda|^2}\sqrt{1 + |\lambda'|^2}\cos\theta.$$

Then,

$$
\begin{align}
(6.6) \qquad |\lambda - \lambda'|^2 &= \left(\sqrt{1 + |\lambda|^2} - \sqrt{1 + |\lambda'|^2}\right)^2 + \\
&\quad + 2\sqrt{1 + |\lambda|^2}\sqrt{1 + |\lambda|^2}\,(1 - \cos\theta).
\end{align}
$$

From (6.6) we get that

$$|\lambda - \lambda'|^2 \geq 2\sqrt{1 + |\lambda|^2}\sqrt{1 + |\lambda|^2}\,(1 - \cos\theta),$$

i.e.,

$$(6.7) \qquad 1 - \cos\theta \leq \frac{|\lambda - \lambda'|^2}{2\sqrt{1 + |\lambda|^2}\sqrt{1 + |\lambda|^2}} \leq \frac{|\lambda' - \lambda|^2}{2}.$$

Therefore $1 - \cos\theta \leq \frac{c^2}{2}$, and hence the angle θ is bounded above by $\arccos(1 - c^2/2)$. By the Taylor expansion of cosine near 0 we get the bound

$$\theta^2 \leq \frac{2}{1 - c^2/2}\,(1 - \cos\theta).$$

Then, from (6.7) we can deduce the upper bound in (1).

For the lower bound in (2), we rewrite the cosine law and get:

$$
\begin{align}
|\lambda - \lambda'|^2 &= \left(\frac{|\lambda|^2 - |\lambda'|^2}{\sqrt{1 + |\lambda|^2} + \sqrt{1 + |\lambda'|^2}}\right)^2 + \\
&\quad + 2\sqrt{1 + |\lambda|^2}\sqrt{1 + |\lambda'|^2}\,(1 - \cos\theta).
\end{align}
$$

Since $||\lambda| - |\lambda'|| \le |\lambda - \lambda'|$ and $1 - \cos\theta \le \theta^2/2$, then,

$$(6.8) \qquad |\lambda - \lambda'|^2 \le \left(\frac{|\lambda| + |\lambda'|}{\sqrt{1 + |\lambda|^2} + \sqrt{1 + |\lambda'|^2}} \right)^2 |\lambda - \lambda'|^2 +$$
$$+ \sqrt{1 + |\lambda|^2} \sqrt{1 + |\lambda'|^2} \, \theta^2$$

Since $0 \le |\lambda| \le 1$, it is easily seen that

$$\frac{|\lambda| + |\lambda'|}{\sqrt{1 + |\lambda|^2} + \sqrt{1 + |\lambda'|^2}} \le \frac{1 + |\lambda'|}{\sqrt{2} + \sqrt{1 + |\lambda'|^2}}.$$

Furthermore, by elementary arguments one can see that $|\lambda'| \le \tan(\hat\theta + \pi/4)$, and therefore

$$\frac{|\lambda| + |\lambda'|}{\sqrt{1 + |\lambda|^2} + \sqrt{1 + |\lambda'|^2}} \le \frac{1 + \tan(\hat\theta + \pi/4)}{\sqrt{2} + \sqrt{1 + \tan(\hat\theta + \pi/4)^2}}$$
$$\le \frac{\tan(\hat\theta + \pi/4)}{\sqrt{1 + \tan(\hat\theta + \pi/4)^2}} = \sin(\hat\theta + \pi/4),$$

where the second inequality holds since $\tan(\hat\theta + \pi/4) \ge 1$. Then, from (6.8),

$$|\lambda - \lambda'|^2 \le \frac{\sqrt{1 + |\lambda|^2} \sqrt{1 + |\lambda'|^2}}{\cos(\hat\theta + \pi/4)^2} \, \theta^2,$$

and hence

$$|\lambda - \lambda'|^2 \le \frac{\sqrt{2}}{\cos(\hat\theta + \pi/4)^3} \, \theta^2.$$

$\square$

Remark 6.9. Note that if $(A, \lambda) \in \pi_1(\mathcal{V}) \subset \mathbb{P}(\mathbb{K}^{n \times n} \times \mathbb{K})$ then $|\lambda| \le \|A\|_F$ is always satisfied.

6.2.1. *Proof of Lemma 4.7.*

Proof of Lemma 4.7. The proof of (1) and (2) follows directly from the definition of $d_{\mathbb{P}^2}$ and Lemma 6.8. $\square$

REFERENCES

[1] D. Armentano and M. Shub: Smale's Fundamental Theorem of Algebra reconsidered. To appear in *Foundation of Computational Mathematics*.

[2] D. Armentano: PH. D. Thesis. Universidad de la República, Uruguay, and Université Paul Sabatier, France.

[3] S. Batterson and J. Smillie: Rayleigh quotient iteration fails for nonsymmetric matrices, *Applied Math. Letters*, 2 (1989) 19-20.

[4] S. Batterson and J. Smillie: Rayleigh quotient iteration for nonsymmetric matrices, *Math. of Comp.*, 191 (1990) 169-178.

[5] C. Beltrán: A continuation method to solve polynomial systems, and its complexity, *Numerische Mathematik* 117, no. 1 (2011), 89113.

[6] C. Beltrán and L. M. Pardo, *Smale's 17th problem: average polynomial time to compute affine and projective solutions*, J. Amer. Math. Soc. 22 (2009), no. 2, 363–385. MR 2476778 (2009m:90147)

[7] C. Beltrán and M.Shub: The complexity and geometry of numerically solving polynomial systems. arXiv:1211.1528.

[8] L. Blum, F. Cucker, M. Shub and S. Smale: Complexity and real computation. *Springer-Verlag*, New York, 1998.

[9] I. Briquel, F. Cucker, J. Peña, and V. Roshchina: Fast Computation of Zeros of Polynomial Systems with Bounded Degree under Variable-precision. To appear at *Mathematics of Computation*

[10] P. Bürguisser and F. Cucker: On a problem posed by Steve Smale. To appear in *Annals of Mathematics.*

[11] C. D'Andrea, T. Krick and M. Sombra: Heights of varieties in multiprojective spaces and arithmetic Nullstellensätze. Preprint 2011.

[12] M.T. Chu: A simple application of the homotopy method to symmetric eigenvalue problems. *Linear Algebra Appl.*, 59 (1984), pp. 8590.

[13] J.P. Dedieu: Points Fixes, Zéros et la Méthode de Newton. Mathématiques & Applications *SMAI-Springer Verlag* , 54, 2006.

[14] J-P. Dedieu, G. Malajovich and M. Shub: Adaptative step size selection for homotopy methods to solve polynomial equations, to appear.

[15] J.P. Dedieu and M. Shub: Multihomogeneous Newton Methods. *Mathematics Of Computation* vol. 69, no. 231, (1999), 1071-1098.

[16] P. Deift: Some open problems in random matrix theory and the theory of integrable systems. *Contemporary Mathematics*, 458 (2008), 419-430.

[17] J. Demmel: The Probability that a numerical analysis problem is difficult. *Mathematics Of Computation* vol. 50,(1988), 449-480.

[18] W. Fulton: Intersection Theory. Springer Verlag, 1984.

[19] G. Golub and C. van Loan: Matrix Computations, third edition, *The Johns Hopkins University Press*, 1996.

[20] T.Y. Li: Numerical solution of multivariate polynomial systems by homotopy continuation methods. *Acta numerica*, 6, Cambridge Univ. Press, Cambridge,(1997), 399436.

[21] T.Y. Li, and T. Sauer: Homotopy method for generalized eigenvalue problems Ax=Bx. *Linear Algebra Appl.* 91 (1987), 6574.

[22] T.Y. Li, T. Sauer, and J. Yorke: Numerical solution of a class of deficient polynomial systems. *SIAM J. Numer. Anal.* 24 (1987), no. 2, 435451.

[23] S. H. Lui, H. B. Keller, and T. W. C. Kwok: Homotopy method for the large, sparse, real nonsymmetric eigenvalue problem, SIAM J. Matrix Anal. Appl. 18(2), volume 18, 312333, 1997

[24] D. Mumford: Algebraic Geometry I Complex Projective Varieties. Springer-Verlag, 1976.

[25] V.Y. Pan: Optimal and nearly optimal algorithms for approximating polynomial zeros, *Comput. Math. Appl.*, vol 31,(1996), 97-138.

[26] C. W. Pfrang, P. Deift, G. Menon: How long does it take to compute the eigenvalues of a random symmetric matrix? http://arxiv.org/abs/1203.4635

[27] J. Renegar: On the worst-case arithmetic complexity of approximating zeros of polynomials. *J. Complexity* 3 (1987), no. 2, 90113.

[28] M. Shub: Complexity of Bezout's theorem. VI. Geodesics in the Condition (Number) Metric. *Foundation of Computational Mathematics.* 9, (2009), 171-178.

[29] M. Shub and S. Smale: Complexity of Bezout's theorem. I. Geometrical Aspects. *J. Amer. Math. Soc.* vol. 6 , no. 1, (1993), 459-501.

[30] M. Shub and S. Smale: Complexity of Bezout's theorem. IV. Probability of success; extensions. *SIAM J. Numer. Anal.* 33 , no. 1, (1996),128-148.

[31] M. Shub and S. Smale: Complexity of Bezout's theorem. V. Polynomial time. *Theoret. Comput. Sci.* 133 , no. 1, (1994) ,141-164.

[32] G.W. Stewart:Matrix Algorithms: Eigensystems, *SIAM*, 2001.

[33] G.W. Stewart and J. Sun : Matrix perturbation theory. *Academic Press*, 1990.

[34] L.N. Trefethen and D. Bau: Numerical linear algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. xii+361.

[35] D. S. Watkins: The Matrix Eigenvalue Problem. GR and Krylov Subspaces Methods. *SIAM*, 2007.

[36] J.H. Wilkinson: The Algebraic Eigenvalue Problem, *Clarendon Press, Oxford*, 1965.

[37] J.H. Wilkinson: Note on matrices with a very ill–conditioned eigenproblem. *Numer. Math.*, vol 19, (1972), 176-178.

Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay

E-mail address: diego@cmat.edu.uy