HAL open science

Generalized Perron Roots and Solvability of the Absolute Value Equation

Manuel Radons, Josué Tonelli-Cueto

To cite this version:

Manuel Radons, Josué Tonelli-Cueto. Generalized Perron Roots and Solvability of the Absolute Value Equation. SIAM Journal on Matrix Analysis and Applications, 2023, 44 (4), pp.1645-1666. 10.1137/22M1517184 . hal-03738197

HAL Id: hal-03738197

https://inria.hal.science/hal-03738197

Submitted on 25 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

arXiv:1912.08157v4 [math.OC] 22 Jul 2022

Generalized Perron roots and solvability of the absolute value equation

Manuel Radons
Technische Universität Berlin
Chair of Discrete Math./Geometry
Berlin, GERMANY
radons@math.tu-berlin.de

Josué Tonelli-Cueto
Inria Paris & IMJ-PRG
OURAGAN team
Paris, FRANCE
josue.tonelli.cueto@bizkaia.eu

Abstract

Let A be a $n \times n$ real matrix. The piecewise linear equation system $z - A|z| = b$ is called an absolute value equation (AVE). It is well-known to be equivalent to the linear complementarity problem. Unique solvability of the AVE is known to be characterized in terms of a generalized Perron root called the sign-real spectral radius of A. For mere, possibly non-unique, solvability no such characterization exists. We close this gap in the theory. That is, we define the concept of the aligned spectrum of A and prove, under some mild genericity assumptions on A, that the mapping degree of the piecewise linear function $F_A : \mathbb{R}^n \to \mathbb{R}^n$, $z \mapsto z - A|z|$ is congruent to $(k+1) \mod 2$, where k is the number of aligned values of A which are larger than 1. We also derive an exact—but more technical—formula for the degree of F_A in terms of the aligned spectrum. Finally, we derive the analogous quantities and results for the LCP.

1 Introduction

The *linear complementarity problem* $\mathrm{LCP}(q, M)$, where $q \in \mathbb{R}^n$, and $M \in \mathrm{M}_n(\mathbb{R})$, the space of $n \times n$ real matrices, is to determine $v, w \in \mathbb{R}^n_{\geq 0}$ with $v^T w = 0$ so that

$$v \;=\; Mw + q\,. \tag{1}$$

It provides a common framework for numerous optimization tasks in economics, engineering and computer science. Classical problems that can be reduced to solving an LCP include bimatrix games, linear and quadratic programs [1]. Recent applications are the correct formulation of numerical models for free-surface hydrodynamics [2], L1 regularization in reinforcement learning [3], and the massively parallel implementation of collision detection on CUDA GPUs [4, Chap. 33].

It is well known that an $\mathrm{LCP}(q, M)$ is uniquely solvable for arbitrary q if and only if M is a P-matrix, that is, a matrix whose principal minors are all positive. If $\mathrm{LCP}(q, M)$ is solvable—possibly non-uniquely—for arbitrary q, then M is called a *Q-matrix*. There exists no comprehensive characterization of Q-matrices. We will study this question by investigating the following equivalent problem [5]. Let $b \in \mathbb{R}^n$ and $A \in \mathrm{M}_n(\mathbb{R})$. Then the *absolute value equation* (AVE) poses the problem to find a vector $z \in \mathbb{R}^n$ so that

$$z - A|z| \;=\; b\,, \tag{2}$$

where $|\cdot|$ denotes the componentwise absolute value. The AVE is an interesting problem in its own right. For example, a result by Rump [6, Thm. 2.8] relates the number of solutions of (2) to the condition of the matrix A, which is noteworthy in light of recent developments in real algebraic geometry that deal with precisely such connections of complexity and condition [7, 8]. However, the main focus of theoretical investigations of the AVE is to obtain statements about the LCP. A recent success of this approach is the development of condition numbers for the AVE that lead to new error bounds for the LCP [9].

We will study solvability of the AVE, but with our eyes on the Q-matrix problem. To this end we investigate the piecewise linear function

$$\begin{aligned} F_A : \mathbb{R}^n &\to \mathbb{R}^n \\ z &\mapsto z - A|z| \end{aligned} \tag{3}$$

associated to the AVE (2) and determine its degree and its degree modulo 2 (see Section 3) in terms of the *aligned spectrum* of A (see Section 2):

$$\mathrm{Spec}^{\mathrm{a}}(A) := \{\lambda \geq 0 \mid \exists x \neq 0 \,:\, |Ax| = \lambda x\}. \tag{4}$$

The first main result of this article relates the degree of F_A to what we call the *aligned count* of A:

$$\mathrm{c}^{\mathrm{a}}(A) := \#\{\lambda \in \mathrm{Spec}^{\mathrm{a}}(A) \mid \lambda > 1\} \tag{5}$$

where the count on the right-hand side is with multiplicities.

The term *generic* in the theorem and the corollaries means that the statement holds for all matrices that have a specific property (see Definition 4), which is satisfied for all matrices except those in a given homogeneous hypersurface (see Section 4). This condition, akin to the general position condition in the polyhedral world, guarantees that a random matrix (with respect to a continuous distribution) is generic with probability 1.

Theorem 1. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic such that $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. Then the degree of F_A is well-defined and it satisfies that*

$$\deg F_A \;\equiv\; 1 + \mathrm{c}^{\mathrm{a}}(A) \mod 2 \tag{6}$$

Moreover, $\deg F_A$ equals 1 if all aligned values are smaller than 1, and it equals 0 if all aligned values are larger than 1.

Corollary 1. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be a generic matrix. Then the number of aligned values of A, counted with multiplicity, is odd.*

Corollary 2. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be a generic matrix such that $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. If $\mathrm{c}^{\mathrm{a}}(A)$ is even, then the AVE (2) has a solution for every $b \in \mathbb{R}^n$.*

After stating Theorem 1, we might wonder if there is an exact formula for the degree of F_A when A is generic (in the sense of Definition 4). Indeed, there is such formula and Theorem 1 is a direct consequence of the following more general—but more technical—theorem.

Theorem 2. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic and such that $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. Then the degree of F_A is well-defined and it satisfies that*

$$\deg F_A = 1 - \sum \{\mathrm{sign}(\chi'_{SA}(\lambda)) \mid \lambda > 1, S \in \mathcal{S}, \exists x \in \mathbb{R}_{>0} \,:\, SAx = \lambda x\}$$

where χ_{SA} is the characteristic polynomial of SA, and $\mathcal{S} \subseteq \mathrm{M}_n(\mathbb{R})$ is the set of sign matrices, i.e., diagonal matrices with ± 1 in the diagonal entries.

Observe that the right-hand side sum runs over all aligned values greater than 1, since we have that $|Ax| = \lambda x$ for some $x \neq 0$ if and only if $SAx = x$ for some $S \in \mathcal{S}_n$ and $x \in \mathbb{R}^n_{\geq 0}$. Now, for a generic matrix (see Definition 4), all aligned values correspond to simple eigenvalues of some SA, and so the right-hand side sum is nothing more than a "signed aligned count", i.e., a signed variation of the aligned count $\mathrm{c}^{\mathrm{a}}(A)$. In this way, Theorem 1 is just Theorem 2 reduced modulo 2.

As we mentioned above, a key reason to study the AVE is to gain insights into the equivalent LCP. To this end we derive LCP-analogues for the concept of the aligned spectrum and all statements about the degree of F_A listed in this introduction. Concerning our afore-stated interest in Q-matrices, we note that Corollary 2 directly translates into a statement about the latter, i.e., the coefficient matrix of an LCP is a Q-matrix if the LCP-equivalent of the aligned count is even (Corollary 4).

Organization In Section 2, we recall the sign-real spectrum and we introduce the aligned spectrum which naturally emerges during the study of the eigenproblem of F_A; in Section 3, we recall the topological notion of degree, its formula in terms of signed counts of preimages of a regular value, and some of its properties in the case of interest; in Section 4, we introduce the notion of genericity of a matrix relevant to our context and show some perturbation results. In Section 5, we show how the above results for AVEs can be transferred to LCPs. We conclude in Section 6 proving Theorems 1 and its corollaries, and also Theorem 2.

Acknowledgements M.R. would like to thank Jiri Rohn and Siegfried M. Rump for helpful discussions. Further, he would like to thank his advisor, Michael Joswig, for general support.

J.T.-C. was supported by a postdoctoral fellowship of the 2020 "Interaction" program of the *Fondation Sciences Mathématiques de Paris*, and was partially supported by ANR JCJC GALOP (ANR-17-CE40-0009), the PGMO grant ALMA, and the PHC GRAPE. He is also grateful to Evgenia Lagoda for moral support and Gato Suchen for the mathematical discussions regarding the proof of Theorem 2.

2 Sign-Real and Aligned Spectra

The sign-real and aligned spectra of A emerge naturally when we study the map F_A (3). Note that studying this map in order to understand (2) is a similar to the strategy in linear algebra to study $z \mapsto Az$ in order to understand the solvability of $Az = b$

We note that F_A is a positively homogeneous map, i.e., for every $z \in \mathbb{R}^n$ and $\lambda > 0$, $F_A(\lambda z) = \lambda F_A(z)$; and that F_A is piecewise linear, with linear parts of the form

$$\mathbb{I} - AS \tag{7}$$

for sign matrices $S \in \mathcal{S}_n := \{\mathrm{diag}(s_1, \ldots, s_n) \mid s_i \in \{-1, +1\}\}$.

2.1 Sign-real spectrum and bijectivity

The sign-real spectrum was used independently by Rump [6] and Rohn [10, Chap. 6] to determine when the function F_A is bijective, or equivalently, when the absolute value equation (2) is uniquely solvable for arbitrary b.

Definition 1. Let $A \in \mathrm{M}_n(\mathbb{R})$. The *sign-real spectrum* of A, $\mathrm{Spec}^\rho(A)$, is the set

$$\mathrm{Spec}^\rho(A) := \mathbb{R}_{\geq 0} \cap \bigcup_{S \in \mathcal{S}_n} \mathrm{Spec}(SA).$$

Theorem 3. *Let $A \in \mathrm{M}_n(\mathbb{R})$. Then the following are equivalent:*

(a) $\max \mathrm{Spec}^\rho(A) < 1$,

(b) F_A *is bijective,*

(c) The AVE (2) *has a unique solution for every* $b \in \mathbb{R}^n$. □

The largest element of the sign-real spectrum, $\max \mathrm{Spec}^\rho(A)$, is called the *sign-real spectral radius*. Due to Theorem 3 it can be considered as a generalization of contractivity conditions for linear operators. Rump [6] also showed that the sign-real spectral radius generalizes the Perron root to matrices that are not positive. Another generalization of Perron Frobenius theory, to homogeneous monotone functions on the positive cone, was introduced in [11]. It is not clear how these two generalized theories are related apart from their common origin.

2.2 Aligned spectrum and the eigenproblem

The aligned spectrum arises when studying the eigenproblem for F_A: Determine for which $\lambda \geq 0$ and $v \in \mathbb{R}^n \setminus 0$, we have

$$F_A(v) = \lambda v.$$

Proposition 1. *Let $A \in \mathrm{M}_n(\mathbb{R})$, $\lambda \geq 0$ and $v \in \mathbb{R}^n \setminus 0$. Then $F_A(v) = \lambda v$ if and only if there is some sign matrix $S \in \mathcal{S}_n$ such that $Sv \geq 0$ and $(1-\lambda, |v|)$ is an eigenpair of SA.*

Proof. If $F_A(v) = \lambda v$, then we have that $(1-\lambda)v = A|v|$. Now, let $S \in \mathcal{S}_n$ such that $Sv \geq 0$, by taking as the diagonal elements of S the signs of the components of v. Then $(1-\lambda)|v| = (1-\lambda)Sv = SA|v|$. Hence there is $S \in \mathcal{S}_n$ such that $Sv \geq 0$ and $(1-\lambda, |v|)$ is an eigenpair of SA.

Conversely, if such an S exists, then

$$F_A(v) = v - A|v| = v - S(SA)|v| = v - S(1-\lambda)|v| = \lambda v,$$

where we have used that $S^2 = \mathbb{I}$. Consequently, $(1-\lambda, |v|)$ is an eigenpair of SA, and $v = S|v|$. □

In view of the above proposition, the aligned spectrum is introduced. We note that the definition below is equivalent to that given in (4).

Definition 2. Let $A \in \mathrm{M}_n(\mathbb{R})$. An *aligned trio* of A is a triplet $(\lambda, S, v) \in \mathbb{R}_{\geq 0} \times \mathcal{S}_n \times (\mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0})$ such that

$$SAv = \lambda v.$$

Given such a trio, we call (S, v) an *aligned vector* and λ an *aligned value* of A. The *aligned spectrum* of A, which we denote $\mathrm{Spec}^{\mathrm{a}}(A)$, is the set of aligned values of A, i.e.,

$$\mathrm{Spec}^{\mathrm{a}}(A) := \{\lambda \geq 0 \mid \exists S \in \mathcal{S}_n,\ v \in \mathbb{R}^n_{\geq 0} \setminus 0 \,:\, SAv = \lambda v\}.$$

The following proposition shows how the aligned spectrum is related to the solution set of $F_A(z) = 0$. In analogy to linear maps, we call F_A *nondegenrate* if $F_A(z) = 0$ has only the trivial solution $z = 0$.

Proposition 2. *Let $A \in \mathrm{M}_n(\mathbb{R})$. Then $F_A(z) = 0$ has non-trivial solutions if and only if $1 \in \mathrm{Spec}^{\mathrm{a}}(A)$.*

Proof. By Proposition 1, non-trivial solutions of $F_A(z) = 0$ correspond to aligned trios of the form $(1, S, v)$. Hence, the claim follows. □

From the definitions it follows that

$$\mathrm{Spec}^{\mathrm{a}}(A) \subseteq \mathrm{Spec}^{\rho}(A). \tag{8}$$

However, this is not an equality in general.

Example 1. Let

$$A := \begin{pmatrix} 1 & 0 \\ -1/2 & 1/2 \end{pmatrix}. \tag{9}$$

One readily checks that

$$\mathrm{Spec}^{\mathrm{a}}(A) = \{1/2\} \subsetneq \mathrm{Spec}^{\rho}(A) = \{1/2, 1\}.$$

Moreover, by Theorem 3 and Proposition 2, this example shows that F_A might not be bijective, even though it is nondegenrate. Furthermore, it shows that the largest aligned value and the sign-real spectral radius do not necessarily coincide. In light of Theorem 1, this demonstrates that we may have $\deg F_A = 1$ without bijectivity.

We finish with the following example which shows that $A \mapsto \max \mathrm{Spec}^{\mathrm{a}}(A)$ is not continuous unlike $A \mapsto \max \mathrm{Spec}^{\rho}(A)$ which is [6, Corollary 2.5]. However, in the generic case (see Definition 4), we can recover continuity, since simple real eigenvalues cannot become complex and strictly positive vectors cannot become nonpositive under an arbitrarily small perturbation.

Example 2. Let t lie in a sufficiently small neighborhood of 0 and consider the following family of matrices:

$$A_t := \begin{pmatrix} 1 & -0.5-t \\ 0.5 & 0 \end{pmatrix} . \tag{10}$$

A straightforward calculation shows that $\mathrm{Spec}^{\rho}(A_t)$ is equal to

$$\left\{ \frac{1+\sqrt{-2t}}{2}, \frac{1-\sqrt{-2t}}{2}, \frac{\sqrt{2}\sqrt{1+t}-1}{2}, \frac{1+\sqrt{2}\sqrt{1+t}}{2} \right\},$$

if $t \leq 0$; and

$$\left\{ \frac{\sqrt{2}\sqrt{1+t}-1}{2}, \frac{1+\sqrt{2}\sqrt{1+t}}{2} \right\}$$

if $t > 0$. Similarly, we can see that $\mathrm{Spec}^{\mathrm{a}}(A_t)$ is equal to

$$\left\{ \frac{1+\sqrt{-2t}}{2}, \frac{1-\sqrt{-2t}}{2}, \frac{\sqrt{2}\sqrt{1+t}-1}{2} \right\},$$

if $t \leq 0$; and

$$\left\{ \frac{\sqrt{2}\sqrt{1+t}-1}{2} \right\}$$

if $t > 0$.

Hence, we have that

$$\max \mathrm{Spec}^{\rho}(A_t) = \frac{1+\sqrt{2}\sqrt{1+t}}{2} \geq \max \mathrm{Spec}^{\mathrm{a}}(A_t) = \begin{cases} \frac{1+\sqrt{-2t}}{2}, & \text{if } t \leq 0 , \\ \frac{\sqrt{2}\sqrt{1+t}-1}{2}, & \text{if } t > 0 . \end{cases}$$

This shows that the maximum of the aligned spectrum is not continuous.

3 Degree of a map

The degree of a continuous map $G : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ is a fundamental topological invariant that is preserved under homotopy. Intutively, we only have to think of the degree as the number of times that the map wraps $\mathbb{S}^{n-1}$ around itself. Its formal definition is as follows.

Definition 3. Let $G : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ be a continuous map. The *degree* of G, $\deg G$, is the unique integer d such that the induced map $H_{n-1}(f) : H_{n-1}(\mathbb{S}^{n-1}) \to H_{n-1}(\mathbb{S}^{n-1})$ of homology groups is given by

$$x \mapsto dx$$

under the choice of any fixed isomorphism $H_{n-1}(\mathbb{S}^{n-1}) \simeq \mathbb{Z}$.

Among the main properties of the degree, we have the following, cf. [12, p. 98 ff].

Proposition 3. *Let $G_0, G_1 : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ be continuous maps. Then the following holds:*

(1) $\deg \mathrm{id}_{\mathbb{S}^{n-1}} = 1$ *and* $\deg(-\mathrm{id}_{\mathbb{S}^{n-1}}) = (-1)^n$.

(2) $\deg G_1 \circ G_0 = \deg G_1 \deg G_0$.

(3) If there is a homotopy *between G_0 and G_1, that is, a continuous map $H : [0,1] \times \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ such that for all $x \in \mathbb{S}^{n-1}$, $H(0,x) = G_0(x)$ and $H(1,x) = G_1(x)$, then*

$$\deg G_0 = \deg G_1.$$

Moreover, the converse statement is also true.

(4) If G_0 is not surjective, then $\deg G_0 = 0$. □

Our investigation will be centered around F_A which are nondegenerate. In this case, we can consider the spherical map

$$\begin{aligned} \bar{F}_A : \mathbb{S}^{n-1} &\to \mathbb{S}^{n-1} \\ x &\mapsto F_A(x)/\|F_A(x)\|_2 \end{aligned} \tag{11}$$

and define the *degree of* F_A as

$$\deg F_A := \deg \bar{F}_A. \tag{12}$$

This definition agrees with a more traditional count used for maps $\mathbb{R}^n \mapsto \mathbb{R}^n$. Recall that the *set of regular values of* F_A is the set given by

$$\mathrm{Reg}F_A := \{y \in \mathbb{R}^n \mid \forall x \in F_A^{-1}(y),\ \partial F_A(x) \text{ is well-defined and invertible}\},$$

where ∂F_A is the Jacobian of F_A.

Proposition 4. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be such that $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. Then the set of regular values of F_A is dense and for all $y \in \mathrm{Reg}F_A$ we have*

$$\deg F_A = \sum_{x \in F_A^{-1}(y)} \mathrm{sign}(\det(\partial F_A(x))). \tag{13}$$

Proof. By [1, p. 509 ff] the oriented preimage counts of a nondegenerate positively homogeneous function and its restriction to the sphere coincide. □

Moreover, the following proposition is helpful for our degree computations.

Proposition 5. *Let $A, A_0, A_1 \in \mathrm{M}_n(\mathbb{R})$. Then:*

(1) If $A : [0,1] \to \mathrm{M}_n(\mathbb{R})$ is a continuous path between A_0 and A_1 such that for all $t \in [0,1]$, $1 \notin \mathrm{Spec}^{\mathrm{a}}(A(t))$; then

$$\deg F_{A_0} = \deg F_{A_1}.$$

(2) If $\mathrm{Spec}^{\mathrm{a}}(A) \subseteq [0,1)$, then $\deg F_A = 1$.

(3) If $\mathrm{Spec}^{\mathrm{a}}(A) \subset (1,\infty)$, then $\deg F_A = 0$.

Proof. (1) Consider the following homotopy between $\bar{F}_{A_0}$ and $\bar{F}_{A_1}$:

$$[0,1] \times \mathbb{S}^{n-1} \ni (t,x) \mapsto H(t,x) = \bar{F}_{A(t)}(x) = \frac{F_{A(t)}(x)}{\|F_{A(t)}(x)\|_2}.$$

This homotopy is well-defined because, by assumption and Proposition 2, $F_{A(t)}(x)$ does not vanish at any (t,x). Hence $\bar{F}_{A_0}$ and $\bar{F}_{A_1}$ are homotopic and so they have the same degree.

(2) Consider the path

$$[0,1] \ni t \mapsto A(t) := (1-t)A.$$

This path joins A with $\mathbb{O}$ and it satisfies the condition of (1). Hence $\deg F_A = \deg F_{\mathbb{O}} = \deg \mathrm{id}_{\mathbb{S}^{n-1}} = 1$.

(3) Consider the following homotopy

$$[0,1] \times \mathbb{S}^{n-1} \ni (t,x) \mapsto H(t,x) = \frac{(1-t)x - A|x|}{\|(1-t)x - A|x|\|_2}.$$

Note that the map $(t,x) \mapsto (1-t)x - A|x|$ is continuous, so if it is not vanishing, then the above homotopy is well-defined and continuous. If $t < 1$, then

$$\mathbb{S}^{n-1} \ni x \mapsto (1-t)x - A|x| = (1-t)F_{A/(1-t)}(x)$$

cannot vanish by Proposition 2, since $1 \notin \operatorname{Spec}^{\mathrm{a}}(A/(1-t)) = \operatorname{Spec}^{\mathrm{a}}(A)(1-t) \subset (1,\infty)$. If $t = 1$, then

$$\mathbb{S}^{n-1} \ni x \mapsto A|x|$$

does not vanish, because otherwise $0 \notin \operatorname{Spec}^{\mathrm{a}}(A)$. Thus the desired map does not vanish and we obtain a homotopy between $\bar{F}_A$ and

$$\mathbb{S}^{n-1} \ni x \mapsto \frac{A|x|}{\|A|x|\|_2}.$$

If we precompose this map with $x \mapsto |x|$, it does not change. Now, since $x \mapsto |x|$ is not surjective,

$$\deg\left(x \mapsto \frac{A|x|}{\|A|x|\|_2}\right) = \deg\left(x \mapsto \frac{A|x|}{\|A|x|\|_2}\right)\deg(x \mapsto |x|) = 0,$$

as we wanted to show. □

We conclude with an example which shows that the relationship between the aligned spectrum and the degree in Theorem 1 holds only modulo 2

Example 3. Let ε be in a sufficiently small neigborhood of zero. Consider the family of matrices:

$$B_\varepsilon := \begin{pmatrix} 2.5 & -1.25-\varepsilon \\ 1.25 & 0 \end{pmatrix}. \tag{14}$$

We can see that

$$\operatorname{Spec}^{\mathrm{a}}(B_\varepsilon) = \left\{1.25\left(1+\sqrt{-0.8\varepsilon}\right), 1.25\left(1-\sqrt{-0.8\varepsilon}\right), 1.25\left(\sqrt{2}\sqrt{1+0.4\varepsilon}-1\right)\right\},$$

if $\varepsilon \leq 0$; and that

$$\operatorname{Spec}^{\mathrm{a}}(B_\varepsilon) = \left\{1.25\left(\sqrt{2}\sqrt{1+0.4\varepsilon}-1\right)\right\},$$

if $\varepsilon > 0$. And so we see that

$$\#\{\lambda \in \operatorname{Spec}^{\mathrm{a}}(B_\varepsilon) \mid \lambda > 1\} = \begin{cases} 2, & \text{if } \varepsilon < 0 \\ 0, & \text{if } \varepsilon > 0 \end{cases},$$

and that

$$\deg F_{B_\varepsilon} = 1$$

by Proposition 5, cf. Figure 1.

On the one hand, this shows that the degree is more stable than the number of aligned values greater than one. On the other hand, note that the change in the number of aligned values greater than one happens because for $\varepsilon = 0$, B_ε is not generic—it has a double aligned value.

Moreover, for $\varepsilon < 0$, B_ε is generic (see Definition 4) and it satisfies

$$\#\{\lambda \in \operatorname{Spec}^{\mathrm{a}}(B_\varepsilon) \mid \lambda > 1\} - 1 = \deg F_{B_\varepsilon};$$

and for $\varepsilon > 0$, B_ε is still generic, but

$$\#\{\lambda \in \operatorname{Spec}^{\mathrm{a}}(B_\varepsilon) \mid \lambda > 1\} + 1 = \deg F_{B_\varepsilon}.$$

This shows that the equality modulo 2 in Theorem 1 cannot be corrected in an easy way to obtain an equality between the degree and the number of aligned values greater than one. We need the more technical expression in Theorem 2 for this.

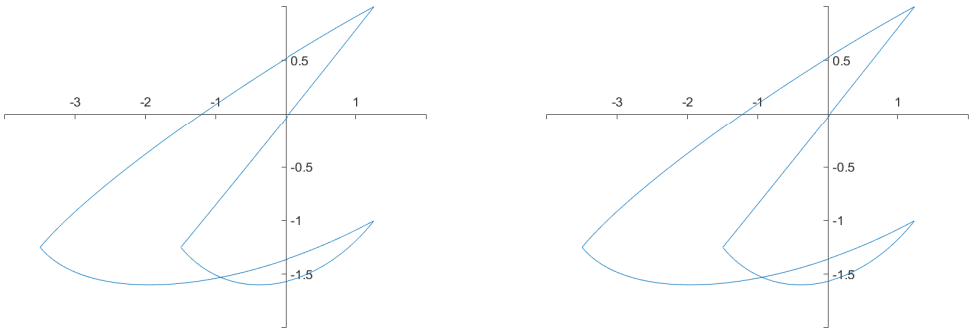

Figure 1: Image of the unit circle under F_{B_ε}. For $\varepsilon = -0.01$ (left) and $\varepsilon = 0.01$ (right) it winds around the origin.

4 Generic matrices

To make precise the statement of Theorem 1, we introduce a notion of genericity adapted to our setting.

Definition 4. A *generic* matrix is a matrix $A \in \mathrm{M}_n(\mathbb{R})$ such that for every aligned trio (λ, S, v) of A, a) λ is a simple eigenvalue of SA, and b) v is strictly positive.

Since genericity usually means a class of entities whose complement is contained in a proper algebraic hypersurface, we need to show that the above notion is indeed a generic one according to the common use.

Proposition 6. *The set of matrices that is not generic in* $\mathrm{M}_n(\mathbb{R})$ *is contained in a proper algebraic hypersurface.*

In particular, for any random matrix $\mathfrak{A} \in \mathrm{M}_n(\mathbb{R})$ *with an absolutely continuous distribution,* $\mathfrak{A}$ *is generic almost surely.*

Proof. We only have to prove that the set of matrices with double aligned values or aligned vectors in the boundary of the positive orthant is contained in a hypersurface.

The above set is contained in the union of the sets

$$\{A \in \mathrm{M}_n(\mathbb{R}) \mid SA \text{ has a double eigenvalue}\} \tag{15}$$

and

$$\{A \in \mathrm{M}_n(\mathbb{R}) \mid SA \text{ has a non-zero eigenvector in } H\} \tag{16}$$

where $S \in \mathcal{S}_n$ runs over all sign matrices and H over all coordinate hyperplanes—of the form $X_i = 0$. Thus, if we show that each one of these sets is contained in a hypersurface, then we are done, since a finite union of hypersurfaces is a hypersurface.

On the one hand, the set 15 is given by the discriminant of the characteristic polynomial of SA, which is well-known to define a proper algebraic hypersurface in $\mathrm{M}_n(\mathbb{R})$. On the other hand, the set 16 is a proper algebraic hypersurface by [13, Propoposition 1.2]. Hence, all the sets are proper algebraic hypersurfaces, and the proof is complete. □

4.1 Interpretation of count for generic matrices

When we introduced the aligned count, see (5),

$$\mathrm{c}^{\mathrm{a}}(A) \;=\; \#\{\lambda \in \mathrm{Spec}^{\mathrm{a}}(A) \mid \lambda > 1\},$$

we said that the right-hand side counts multiplicities. Note that for generic A, an aligned value λ will always be a simple eigenvalue of the corresponding SA, where $S \in \mathcal{S}_n$. However, there might be more than one such SA. Because of this we have to include the "counted with multiplicity". The following proposition gives an alternative interpretation of the central quantity $\mathrm{c}^{\mathrm{a}}(A)$ for Theorem 1 in terms of $\bar{F}_A$.

Proposition 7. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic. Then $\mathrm{c}^{\mathrm{a}}(A)$ is equal to the number of fixed points of $\bar{F}_A$ such that its antipodal point is mapped to them. In other words,*

$$\mathrm{c}^{\mathrm{a}}(A) \;=\; \#\{x \in \mathbb{S}^{n-1} \mid x = \bar{F}_A(x) = \bar{F}_A(-x)\}.$$

For proving this proposition, the following proposition will be useful.

Proposition 8. *Let $A \in \mathrm{M}_n(\mathbb{R})$ and $x \in \mathbb{S}^{n-1}$. If F_A is nondegenerate, then x is a fixed point of $\bar{F}_A$ if and only if there is an aligned trio (λ, S, v) such that either $x = -Sv$ or $\lambda < 1$ and $x = Sv$. Moreover, when x is a fixed point of $\bar{F}_A$, the following are equivalent:*

- *$\bar{F}_A(-x) = x$.*
- *There is an aligned trio (λ, S, v) such that $\lambda > 1$ and $x = -Sv$.*

Proof of Proposition 7. By Proposition 8, the fixed points $x \in \mathbb{S}^{n-1}$ of $\bar{F}_A$ such that $\bar{F}_A(-x) = x$ are in one-to-one correspondence with the aligned trios (λ, S, v) such that $\lambda > 1$. Since A is generic, this means precisely the number of aligned values greater than one counted with multiplicity. □

Proof of Proposition 8. If (λ, S, v) is an aligned trio, then

$$F_A(Sv) = (1 - \lambda)Sv$$

and

$$F_A(-Sv) = (1 + \lambda)(-Sv).$$

In this way, $-Sv$ is always a fixed point of $\bar{F}_A$ and Sv is so if and only if $\lambda < 1$. This shows one direction.

If $x \in \mathbb{S}^{n-1}$ is a fixed point of $\bar{F}_A$, then for some $\mu > 0$,

$$F_A(x) = \mu x.$$

Let $S \in \mathcal{S}_n$ be such that $Sx \geq 0$, so that $v = Sx$. Then we have that

$$SAv = (1 - \mu)v.$$

If $1 - \mu \geq 0$, then $(1 - \mu, S, v)$ is an aligned trio such that $1 - \mu < 1$ and $x = Sv$. Otherwise, $1 - \mu < 0$, and then $(\mu - 1, -S, v)$ is an aligned trio and $x = -(-S)v$. Hence there is an aligned trio (λ, S, v) such that either $x = -Sv$ or $\lambda < 1$ and $x = Sv$.

We show the second equivalence. Let $x \in \mathbb{S}^{n-1}$ be a fixed point of $\bar{F}_A$. Then, by the first part, there is an aligned trio (λ, S, v) such that either $x = -Sv$ or $\lambda < 1$ and $x = Sv$. In the second case, we have that

$$\bar{F}_A(-x) = -x.$$

Thus we must have the first case. But then $\bar{F}_A(-x) = x$ if and only if $\lambda > 1$, because otherwise $-x$ is also a fixed point.

If $x = -Sv$ for some aligned trio (λ, S, v) such that $\lambda > 1$. Then, by the first equivalence, $x = -Sv$ is a fixed point of $\bar{F}_A$, and, by direct computation, $\bar{F}_A(-x) = x$. □

4.2 Perturbation of matrices to make them generic

The following proposition shows that matrices corresponding to non-degenerate maps can be slightly perturbed to obtain a generic matrix with the same corresponding degree.

Proposition 9. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be such that $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. Then:*

(a) The quantity

$$\kappa^{\mathrm{a}}(A) := \sup_{x \neq 0} \frac{\|x\|_2}{\|F_A(x)\|_2} \tag{17}$$

is finite.

(b) For every $\varepsilon \in (0, 1/\kappa^{\mathrm{a}}(A))$ and

$$\tilde{A} \in B_F(A, \varepsilon) := \{X \in \mathrm{M}_n(\mathbb{R}) \mid \|X - A\|_F < \varepsilon\},$$

$F_{\tilde{A}}$ is nondegenrate, and $\deg F_{\tilde{A}} = \deg F_A$.

(c) Let $\tilde{A} \in B_F(A, \varepsilon)$ be a random matrix with the uniform distribution on $B_F(A, \varepsilon)$, then $\tilde{A}$ is generic with probability one.

Proof. (a) We have that

$$\min_{x \neq 0} \frac{\|F_A(x)\|_2}{\|x\|_2}$$

is zero if and only if $1 \in \mathrm{Spec}^{\mathrm{a}}(A)$ by Proposition 2. Hence, it is a positive number if $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$ and its inverse, the quantity $\kappa^{\mathrm{a}}(A)$, must be finite.

(b) By the inequalities between matrix norms, we can show that

$$\frac{1}{\kappa^{\mathrm{a}}(\tilde{A})} \geq \frac{1}{\kappa^{\mathrm{a}}(A)} - \|\tilde{A} - A\|\,.$$

Hence, if $\tilde{A} \in B(A, \varepsilon)$, with the given choice of ε, then no matrix in the segment $[A, \tilde{A}]$ can have 1 as an aligned value. Consequently, we have a path between A and $\tilde{A}$, given by $t \mapsto (1-t)A + t\tilde{A}$, such that no matrix in the path has 1 as an aligned value, and so $\deg F_{\tilde{A}} = \deg F_A$ by Proposition 5 (1).

(c) This is a direct consequence of Proposition 6. □

We observe that the above proposition can only be applied when $1 \notin \mathrm{Spec}^{\mathrm{a}}(A)$. In that case, it allows us to produce a generic matrix $\tilde{A}$ such that $F_{\tilde{A}}$ has the same topological structure—degree—as F_A. Now, if $1 \in \mathrm{Spec}^{\mathrm{a}}(A)$, this perturbation trick will not produce $F_{\tilde{A}}$ with the same topological structure as the following examples show.

Example 4. Let

$$A := \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix} \tag{18}$$

for which $\mathrm{Spec}^{\mathrm{a}}(A) = \{1, \sqrt{2} - 1\}$. Now consider the following perturbation

$$A_\varepsilon := \begin{pmatrix} 2 & -1-\varepsilon \\ -1 & 0 \end{pmatrix}. \tag{19}$$

We can see that for $\varepsilon > 0$, $\deg F_{A_\varepsilon} = 1$, since all aligned values are smaller than one; and that for $\varepsilon < 0$, $\deg F_{A_\varepsilon} = 0$, after a straightforward computation. Moreover, Figure 2 indicates—and it can be checked by computation—that

$$F_{A_\varepsilon}(x) = \begin{pmatrix} r \\ 0 \end{pmatrix},$$

where $r > 0$, does not have a solution for $\varepsilon < 0$, cf. Figure 2.

Hence, when we perturb $A \in \mathrm{M}_n(\mathbb{R})$ with $1 \in \mathrm{Spec}^{\mathrm{a}}(A)$, neither do we get consistent topological information about $F_{\tilde{A}}$ via the perturbed matrix $\tilde{A}$, nor do we obtain consistent information about the general solvability of the AVE (2).

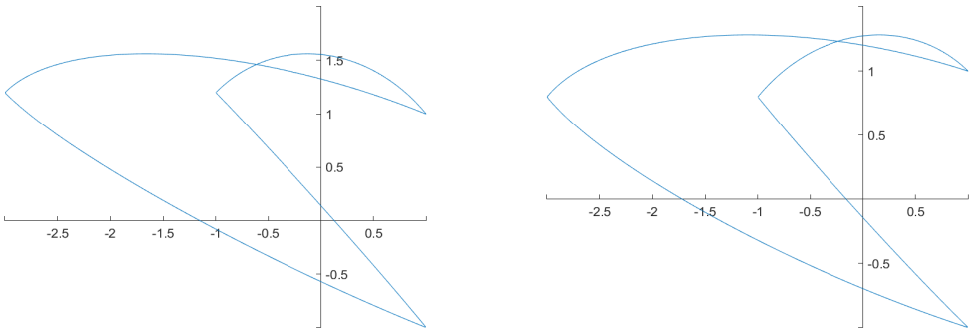

Figure 2: Image of the unit circle under F_{A_ε}. For $\varepsilon = 0.2$ it winds around the origin (left). For $\varepsilon = -0.2$ it does not (right).

Example 5. Consider a generic matrix $A \in \mathrm{M}_n(\mathbb{R})$ so that all aligned values have odd multiplicity. Then for every $t > 0$, tA is generic as long as $1/t \notin \mathrm{Spec}^{\mathrm{a}}(A)$. As t increases from zero to infinity, we have that $\deg F_{tA}$ alternates parity as t crosses the inverses of the aligned values of A, by Theorem 1. This shows again that perturbing matrices with 1 as an aligned value will not produce consistent topological information.

5 Transfer of results to LCPs

Let us reformulate the linear complementary problem $\mathrm{LCP}(q, M)$ (see (1)) in terms of AVEs. Although our reduction is not general, it is almost general.

Proposition 10. *Let $M \in \mathrm{M}_n(\mathbb{R})$ and $q \in \mathbb{R}^n$. The map*

$$(v, w) \mapsto v - w$$

is a bijectivite correspondence between the solution $(v, w) \in \mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$ of $\mathrm{LCP}(q, M)$ and the solutions of

$$(\mathbb{I} + M)x + (\mathbb{I} - M)|x| = 2q. \tag{20}$$

In particular, if $\mathbb{I} + M$ is invertible, then $\mathrm{LCP}(q, M)$ is equivalent to solve the AVE given by

$$z - (\mathbb{I} + M)^{-1}(M - \mathbb{I})|z| = 2(\mathbb{I} + M)^{-1}q \tag{21}$$

Proof. We note that $(v, w) \mapsto v - w$ is a bijection between $\{(v, w) \in \mathbb{R}^n_{\geq 0} \mid v^t w = 0\}$ and $\mathbb{R}^n$. For seeing this, note that the condition $v^T w = 0$ is equivalent to v and w having disjoint supports. In this way, for $x = v - w$, we have that $v = \frac{x+|x|}{2}$ and $w = \frac{|x|-x}{2}$, and so the map has an inverse.

By substituting $v = \frac{x+|x|}{2}$ and $w = \frac{|x|-x}{2}$ into (1), we get (20). And viceversa, by substituting $x = v - w$ into (20), we recover (1), since $|v - w| = v + w$. Hence we have a bijective correspondence of solutions, as desired.

The last claim follows trivially. □

Because of the above proposition, we will consider the following piecewise linear map.

$$\begin{aligned} G_M : \mathbb{R}^n &\to \mathbb{R}^n \\ z &\mapsto (\mathbb{I} + M)z + (M - \mathbb{I})|z|. \end{aligned} \tag{22}$$

As it happened with AVEs and F_A, we have now that the topological degree G_M determines the solvability of $\mathrm{LCP}(q, M)$ for arbitrary $q \in \mathbb{R}^n$. To analyze the degree we introduce the LCP-variant of aligned trios of Definition 2 and of generic matrices of Definition 4—for which we can prove claims analogous to those of Section 4[1].

[1]For this, we only have to use that $\{M \in \mathrm{M}_n(\mathbb{R}) \mid \det(\mathbb{I} + M) = 0\}$ is an algebraic hypersurface.

Definition 5. Let $M \in \mathrm{M}_n(\mathbb{R})$. An *LCP-aligned trio* of M is a triplet $(\lambda, S, v) \in \mathbb{R}_{\geq 0} \times \mathcal{S}_n \times (\mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0})$ such that

$$(M - \mathbb{I})v \;=\; \lambda(\mathbb{I} + M)Sv\,.$$

Given such a trio, we call (S, v) an *LCP-aligned vector* and λ an *LCP-aligned value* of M. The *LCP-aligned spectrum* of M, which we denote $\mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$, is the set of LCP-aligned values of M.

Definition 6. An LCP-generic matrix $M \in \mathrm{M}_n(\mathbb{R})$ is a matrix such that a) $\mathbb{I} + M$ is invertible, and b) $(\mathbb{I} + M)^{-1}(M - \mathbb{I})$ is generic.

With these definitions, we can state the main theorems (Theorems 3 and 2) and their corollaries in the context of LCPs. Recall that $\deg G_M$ is the degree of the spherical map $\bar{G}_M : x \mapsto G_M(x)/\|G_M(x)\|_2$.

Theorem 4. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be LCP-generic such that $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$. Then the degree of G_M is well-defined and it satisfies that*

$$\deg G_M \;\equiv\; 1 + \mathrm{c}^{\mathrm{a}}_{\mathrm{L}}(M) \mod 2$$

where $\mathrm{c}^{\mathrm{a}}_{\mathrm{L}}(M)$ is the number of LCP-aligned values greater than one, counted with multiplicity. Moreover, we have $\deg G_M = \mathrm{sign}(\det(\mathbb{I} + M))$ if $\mathrm{c}^{\mathrm{a}}_{\mathrm{L}}(M) = 0$ and $\deg G_M = 0$ if all LCP-aligned values of M are larger than 1.

Theorem 5. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be LCP-generic such that $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$ Then the degree of G_M is well-defined and it satisfies that*

$$\deg G_M \;=\; \mathrm{sign}(\det(\mathbb{I} + M)) - \sum_{\substack{(\lambda,S,v)\ \textit{LCP-aligned trio of } M \\ \lambda > 1}} \mathrm{sign}\left(\chi'_{\mathbb{I}+M,(M-\mathbb{I})S}(\lambda)\right)\,.$$

where $\chi_{U,V} := \det(TU - V)$ is the generalized characteristic polynomial of $(U, V) \in \mathrm{M}_n(\mathbb{R})^2$.

Corollary 3. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be a LCP-generic matrix. Then the number of LCP-aligned values of M, counted with multiplicity, is odd.* □

Corollary 4. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be LCP-generic such that $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$. If $\mathrm{c}^{\mathrm{a}}_{\mathrm{L}}(M)$ is even, then M is a Q-matrix, i.e., for all $q \in \mathbb{R}^n$, $\mathrm{LCP}(q, M)$ has a solution.*

To show these theorems and their corollaries, we only need to show how the notions in the LCP-setting crrespond to the notions in the AVE-setting. The following three propositions allow us to do these trasnlations. The first one gives how aligned trios correspond to LCP-aligned trios; the second one gives sufficient condition for the degree of G_M to be well-defined; and the third one relates the degree of G_M to that F_A for an appropriate A.

Proposition 11. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be such that $\mathbb{I} + M$ is invertible. Then $(\lambda, S, v) \in \mathbb{R}_{\geq 0} \times \mathcal{S}_n \times (\mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0})$ is an LCP-aligned trio of M if and only if it is an aligned trio of $(\mathbb{I} + M)^{-1}(M - \mathbb{I})$.*

Proposition 12. *Let $M \in \mathrm{M}_n(\mathbb{R})$. Then $G_M(z) = 0$ has non-trivial solutions if and only if $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$. In particular, if $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$, the degree of G_M is well-defined, the set of regular values of G_M is dense and for all $y \in \mathrm{Reg}G_M$ we have*

$$\deg G_M = \sum_{x \in G_M^{-1}(y)} \mathrm{sign}(\det(\partial G_M(x))). \tag{23}$$

Proposition 13. *Let $M \in \mathrm{M}_n(\mathbb{R})$ be such that $1 \notin \mathrm{Spec}^{\mathrm{a}}_{\mathrm{L}}(M)$ and such that $\mathbb{I} + M$ is invertible. Then*

$$\deg G_M \;=\; \mathrm{sign}(\det(\mathbb{I} + M)) \cdot \deg F_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})}.$$

Proof of Proposition 11. Note that, when $\mathbb{I} + M$ is invertible, (λ, S, v) is a LCP-aligned trio of M if and only if $\lambda Sv = (\mathbb{I} + M)^{-1}(M - \mathbb{I})v$. The latter is equivalent to $\lambda v = S(\mathbb{I} + M)^{-1}(M - \mathbb{I})v$, which means exactly that (λ, S, v) is an aligned trio of $(\mathbb{I} + M)^{-1}(M - \mathbb{I})$. □

Proof of Proposition 12. The first claim is proven as in Proposition 2. The second one follow from [1, p. 509 ff], since G_M is a nondegenrate positively homogeneous map. □

Proof of Proposition 13. This follows from the multiplicative property of the degree and the fact that

$$L_M \circ \bar{F}_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})} = \bar{G}_M$$

where $L_M : x \mapsto (\mathbb{I}+M)x/\|(\mathbb{I}+M)x\|_2$. Recall that we have that $\deg L_M = \operatorname{sign}(\det(\mathbb{I}+M))$. □

We now give the proof of Theorems 4 and 5.

Proof of Theorem 4. By Propositions 11, 12 and 13, we have that

$$\deg G_M \equiv \deg F_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})} \mod 2$$

and

$$\mathrm{c}^{\mathrm{a}}_{\mathrm{L}}(M) \equiv \mathrm{c}^{\mathrm{a}}((\mathbb{I}+M)^{-1}(M-\mathbb{I})) \mod 2.$$

Hence the theorem follows by Theorem 1. □

Proof of Theorem 5. By Propositions 11,

$$\deg G_M = \operatorname{sign}(\det(\mathbb{I}+M)) \deg F_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})}.$$

Hence, by Theorem 2 and Proposition 13,

$$\deg G_M = \operatorname{sign}(\det(\mathbb{I}+M)) - \sum_{\substack{(\lambda,S,v) \text{ LCP-aligned trio of } M \\ \lambda>1}} \operatorname{sign}\left(\det(\mathbb{I}+M)\chi'_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})S}(\lambda)\right).$$

Now,

$$\det(\mathbb{I}+M))\chi'_{(\mathbb{I}+M)^{-1}(M-\mathbb{I})S}(\lambda) = \chi'_{\mathbb{I}+M,(M-\mathbb{I})S}(\lambda),$$

so we obtain the desired result. □

6 Proof of Theorems 1 and 2 and Corollaries 1 and 2

We first show how Theorem 1 follows from Theorem 2. Then we prove the corollaries 1 and 2. We finish giving the proof of Theorem 2.

6.1 Proof of Theorem 1

Note that the formula of Theorem 1 can be rewritten as

$$\deg F_A = 1 - \sum_{\substack{(\lambda,S,v) \text{ aligned trio of } A \\ \lambda>1}} \operatorname{sign}(\chi'_{SA}(\lambda)).$$

Now, since A is generic, we have that for each aligned trio (λ, S, A), λ is a simple eigenvalue of SA and so a simple root of χ_{SA}. Hence $\operatorname{sign}(\chi'_{SA}(\lambda))$ is either $+1$ or -1 for each summand in the sum. Moreover, for a specific aligned value λ, the summand $\operatorname{sign}(\chi'_{SA}(\lambda))$ appears as many times as the size of

$$\{(S,v) \mid S \in \mathcal{S}_n,\ v \in \mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0}, (S,v) \text{ is the aligned vector of } A \text{ corresponding to } \lambda\}.$$

But this quantity, for generic A, is precisely the multiplicity of λ. Hence, we have that for all $\lambda > 1$,

$$\sum_{(\lambda,S,v) \text{ aligned trio of } A} \operatorname{sign}(\chi'_{SA}(\lambda))$$

is the multiplicity of λ mod 2. Hence

$$\sum_{\substack{(\lambda,S,v) \text{ aligned trio of } A \\ \lambda>1}} \operatorname{sign}(\chi'_{SA}(\lambda)) \equiv \mathrm{c}^{\mathrm{a}}(A) \mod 2,$$

and the first part of the theorem follows.

The second part is just Proposition 5(2)-(3).

6.2 Proof of Corollary 1

Since A is generic, the multiplicity of 0 as an aligned value is given by the size of the set

$$\{(S,v) \mid S \in \mathcal{S}_n,\ v \in \mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0}, (S,v) \text{ is the aligned vector of } A \text{ corresponding to } 0\}.$$

This set is invariant under the transformation $(S,v) \mapsto (-S,v)$. Hence, the multiplicity of 0 as an aligned value is always even.

By the proof of Lemma 1, we can consider perturbed matrix $\tilde{A}$ with the same number of positive aligned values as A, but such that 0 is not an aligned value of $\tilde{A}$. Since the multiplicity of 0 is even, this means that the parity of the number of aligned values of $\tilde{A}$ and A is the same. Therefore, without loss of generality, we can assume that A has only positive aligned values.

For $t > 0$, we have that

$$\operatorname{Spec}^{\mathrm{a}}(tA) = t \operatorname{Spec}^{\mathrm{a}}(A). \tag{24}$$

For some $t > 0$ sufficiently large, every aligned value of tA is larger than one, because, by assumption, A has only positive aligned values. Thus, by the second part of Theorem 1, $\deg F_{tA} = 0$, and so, by the first part of the same theorem, $\mathrm{c}^{\mathrm{a}}(tA)$ is odd. Hence tA has an odd number of aligned values, as we wanted to show.

6.3 Proof of Corollary 2

If $\mathrm{c}^{\mathrm{a}}(A)$ is even, then, by Theorem 1, $\deg F_A$ is odd, and so, in particular, non-zero. Hence, by Proposition 3(4), $\bar{F}_A$ is surjective, and so is F_A.

6.4 Proof of Theorem 2

By perturbing A randomly, we can assume without loss of generality that all aligned values of A are simple, i.e., that every aligned value of A appears only in one aligned trio (λ, S, v) of A. The following lemma allows us to do this.

Lemma 1. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic such that $1 \notin \operatorname{Spec}^{\mathrm{a}}(A)$. Then there is $\tilde{A}$ such that 1) $1 \notin \operatorname{Spec}^{\mathrm{a}}(A)$, 2) $\tilde{A}$ is generic, 3) F_A and $F_{\tilde{A}}$ have the same degree, 4) A and $\tilde{A}$ have the same aligned count, i.e., $\mathrm{c}^{\mathrm{a}}(A) = \mathrm{c}^{\mathrm{a}}(\tilde{A})$; and*

$$\sum_{\substack{(\tilde{\lambda},S,v) \text{ aligned trio of } \tilde{A} \\ \lambda>1}} \operatorname{sign}(\chi'_{S\tilde{A}}(\tilde{\lambda})) = \sum_{\substack{(\lambda,S,v) \text{ aligned trio of } A \\ \lambda>1}} \operatorname{sign}(\chi'_{SA}(\lambda));$$

and 5) for every aligned value $\tilde{\lambda}$ of $\tilde{A}$, there is a unique $S \in \mathcal{S}$ such that $\tilde{\lambda}$ is an eigenvalue of $S\tilde{A}$.

Proof. By Proposition 9, 1), 2), and 3) are guaranteed by taking $\tilde{A}$ in a sufficiently small neighborhood $B_F(A, \varepsilon)$ of A. Now, since A is generic, for each $S \in \mathcal{S}_n$, given an eigenvalue λ of SA, we proceed as follows:

(a) If $\lambda \notin [1, \infty) \subseteq \mathbb{C}$, then, by continuity of the eigenvalues, we have that for an arbitrarily small perturbation of A, $\tilde{A}$, the corresponding eigenvalue, $\tilde{\lambda}$, is still outside $[1, \infty)$.

(b) If $\lambda > 1$ is not an aligned value, then $(\lambda\mathbb{I} - SA)v$ does not vanish for $v \in \mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0}$. Therefore

$$\min_{v \in \mathbb{S}^{n-1} \cap \mathbb{R}^n_{\geq 0}} \|(\lambda\mathbb{I} - SA)v\|_2 > 0.$$

But this quantity is not only continuous in λ and A, but 1-Lipschitz in them. Hence, for an arbitrarily small perturbation $\tilde{A}$ of A, we can guarantee that the corresponding eigenvalue $\tilde{\lambda}$ does not become an aligned value of $\tilde{A}$.

(c) If $\lambda > 1$ is an aligned value, consider an aligned trio (λ, S, v) such that $\|v\|_2 = 1$. Since A is generic, λ is simple, and so we can apply the implicit function theorem to

$$\mathbb{R} \times \mathbb{S}^{n-1} \times \mathrm{M}_n(\mathbb{R}) \ni (\lambda, x, M) \mapsto (\det(\lambda\mathbb{I} - MS), (\lambda\mathbb{I} - MS)v)$$

at (λ, v, A). Hence, in a small neighborhood of A, we can write λ and v as smooth functions of A. Since v is strictly positive, this means that for an arbitrarily small perturbation $\tilde{A}$ the aligned value λ goes to an aligned value $\tilde{\lambda}$ which is still simple as an eigenvalue of $\tilde{A}S$. Moreover, if the perturbation is sufficiently small, $\tilde{\lambda}$ remains inside $(1, \infty)$, by continuity, and the signs of $\chi'_{SA}(\lambda)$ and of $\chi'_{S\tilde{A}}(\tilde{\lambda})$ coincide, by the continuity of $\chi'_{SM}(\mu)$ with respect to (M, μ).

Putting the above together, we have that taking $\tilde{A}$ in a sufficiently small neighborhood $B_F(A, \varepsilon)$ of A we can guarantee that it satisfies 1), 2), 3) and 4).

For 5), we only have to show that for almost all $\tilde{A} \in \mathrm{M}_n(\mathbb{R})$, the $S\tilde{A}$, with $S \in \mathcal{S}_n$, do not share any eigenvalue. Once this is done, we can guarantee, by the above, that we can choose a perturbation $\tilde{A}$ with the desired properties. Note that this is the only point where the perturbation is not arbitrary.

We will show that the set

$$M_S := \{X \in \mathrm{M}_n(\mathbb{R}) \mid X \text{ and } SX \text{ have an eigenvalue in common}\}$$

is a proper algebraic hypersurface. Then the set of X such that S_1X and S_2X share an eigenvalue for some $S_1, S_2 \in \mathcal{S}_n$ is given by

$$\bigcup_{S,T \in \mathcal{S}_n} SM_T.$$

Therefore it will be a proper algebraic hypersurface, since it is a finite union of proper algebraic hypersurfaces. Hence, we can choose $\tilde{A} \in B_F(A, \varepsilon)$ such that $\tilde{A}$ does not lie in it.

The determinant of the Sylvester matrix of the characteristic polynomials of X and SX is zero if and only if X and SX share an eigenvalue (see [14, Ch. 3, Prop. 8]). Hence M_S is described by the zero set of a single polynomial. If it is not the full set $\mathrm{M}_n(\mathbb{R})$, then it is a proper algebraic hypersurface, as we wanted to show.

We show that M_S does not contain all matrices, by constructing a matrix not in it. Without loss of generality, we can assume that

$$S = \begin{pmatrix} \mathbb{I}_r & \\ & -\mathbb{I}_{n-r} \end{pmatrix}$$

with $r < n$. Consider the matrix

$$A = \begin{pmatrix} 0 & -\mathbb{I}_r & & & \\ 1 & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & n-r-1 \end{pmatrix}$$

whose eigenvalues are $1, \mathrm{e}^{\frac{2\pi i}{r+1}}, \mathrm{e}^{\frac{4\pi i}{r+1}} \dots, \mathrm{e}^{\frac{2r\pi i}{r+1}}, 1, \dots, n-r-1$. Now, for S as defined above, we have that

$$SA = \begin{pmatrix} 0 & -\mathbb{I}_r & & & \\ -c & 0 & & & \\ & & -a_1 & & \\ & & & \ddots & \\ & & & & -a_{n-r-1} \end{pmatrix}.$$

Hence the eigenvalues of SA are $\mathrm{e}^{\frac{\pi i}{r+1}}, \mathrm{e}^{\frac{3\pi i}{r+1}}, \mathrm{e}^{\frac{5\pi i}{r+1}} \dots, \mathrm{e}^{\frac{(2r+1)\pi i}{r+1}}, -1, \dots, -(n-r-1)$, and so A and SA do not have common eigenvalues for a sufficiently general choice of c and the a_i. Hence $M_S \neq \mathrm{M}_n(\mathbb{R})$ as we wanted to show. □

We will be considering the map F_{tA} as t goes from arbitrarily small values to 1. By (24), the map F_{tA} is non-degenerate as long as $1/t \notin \mathrm{Spec}^{\mathrm{a}}(A)$. Let

$$1 < \lambda_1 < \cdots < \lambda_c$$

be the aligned values greater than one of A. Further, let $S_1, \dots, S_c \in \mathcal{S}_n$ be the corresponding sign matrices of their unique aligned trios, and λ_0 be the largest aligned value smaller than 1 or zero if there are no such aligned values. By our initial assumption, they are not repeated.

When $t < 1/\lambda_c$, we have, by Proposition 5 (2), that F_{tA} has degree 1 and that $\mathrm{c}^{\mathrm{a}}(tA) = 0$. We aim to show that for $t \in (1/\lambda_{k+1}, 1/\lambda_k)$, we have

$$\deg F_{tA} = 1 - \sum_{i=k+1}^{c} \mathrm{sign}(\chi'_{AS_i}(\lambda_i)),$$

which gives the desired claim when $k = 0$. To do so, we only have to show that the degree of F_{tA} changes by $-\chi'_{AS_i}(\lambda_i)$ when t passes through $1/\lambda_i$. Note that, when t varies within $(1/\lambda_{k+1}, 1/\lambda_k)$, neither the aligned count nor the degree changes—the latter by Proposition 5(1). Hence, without loss of generality, it is enough to prove the following proposition.

Proposition 14. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic and $(1, \mathbb{I}, v)$ one of its aligned trios. Assume that $1 \in \mathrm{Spec}^{\mathrm{a}}(A)$ is a simple aligned value such that for all $S \in \mathcal{S} \setminus \{\mathbb{I}\}$, 1 is not an eigenvalue of SA. Then there is an $\varepsilon > 0$ such that for all $t, s \in (0, \varepsilon)$,*

$$\deg F_{(1+t)A} = \deg F_{(1-s)A} - \mathrm{sign}(\chi'_A(1)).$$

Once this proposition is shown, we obtain the following proposition by applying the above one to AS/λ, where (λ, S, v) is the considered aligned trio.

Proposition 15. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be generic and (λ, S, v) one of its aligned trios. Assume that $\lambda \in \mathrm{Spec}^{\mathrm{a}}(A)$ is a simple aligned value such that for all $T \in\in \mathcal{S} \setminus \{S\}$, λ is not an eigenvalue of SA. Then there is an $\varepsilon > 0$ such that for all $t, s \in (0, \varepsilon)$,*

$$\deg F_{(1/\lambda+t)A} = \deg F_{(1/\lambda-s)A} - \mathrm{sign}(\chi'_{AS}(\lambda)).$$ □

With this proposition, the desired claim follows, i.e., that F_{tA} changes the degree as wanted each time t passes through the inverse of an aligned value. Note that by our original perturbation, due to Lemma 1, the assumption is satisfied at each crossing.

Proof of Proposition 14

Since $\mathrm{Spec}^{\rho}(A)$ is discrete, there is some $\varepsilon_0 > 0$ such that $I_0 = (1-\varepsilon_0, 1+\varepsilon_0)$ does not contain other elements from $\mathrm{Spec}^{\rho}(A)$. For this interval, we have that for all $t \in I_0 \setminus \{1\}$, $\mathbb{I} - tA$ is invertible; and for all $t \in I_0$ and $S \in \mathcal{S} \setminus \{\mathbb{I}\}$, $\mathbb{I} - tAS$ is invertible. If $\mathbb{I} - tAS$ is not invertible, then $1/t \in \mathrm{Spec}^{\rho}(A)$ and so, by construction of I_0, $t = 1$ and, by assumption on the aligned value 1, $S = \mathbb{I}$.

Now, since 1 is an aligned value and A generic, let $v \in \mathbb{S}^{n-1} \cap \mathbb{R}^n_{>0}$ be the associated aligned vector with all positive entries. Now, we prove the following three lemmas in the context of Proposition 14.

Lemma 2. *Let $r > 0$. Then there exist $x \in B_2(-v, r)$ and $\delta > 0$ such that for all $t \in (1-\delta, 1+\delta) \setminus \{1\}$, x is a regular value of F_{tA} such that*

$$\sum_{\substack{z \in F_{tA}^{-1} \\ z \notin \mathbb{R}^n_{\geq 0}}} \operatorname{sign}(\det(\partial F_{tA}(z))) \tag{25}$$

does not depend on t.

Lemma 3. *There exist $r > 0$ and $\delta > 0$ such that for $t \in (1-\delta, 1)$,*

$$B(-v, r) \cap F_{tA}\left(\mathbb{R}^n_{\geq 0}\right) = \varnothing$$

and for $t \in (1, 1+\delta)$,

$$B(-v, r) \subset F_{tA}\left(\mathbb{R}^n_{\geq 0}\right).$$

Once these lemmas are proved, we only have to choose $r > 0$, $\delta > 0$ and $x \in B(-v, r)$ so that both Lemmas 3 and 2 apply. For this, we choose $r > 0$ and $\delta > 0$ as in Lemma 2, and then choose $x \in B(-v, r)$ and, if necessary, a smaller $\delta > 0$, following Lemma 3. In this way, for $s \in (1-\delta, 1)$, we obtain

$$\deg F_{sA} = \sum_{\substack{z \in F_{tA}^{-1} \\ z \notin \mathbb{R}^n_{\geq 0}}} \operatorname{sign}(\det(\partial F_{tA}(z)))$$

and for $t \in (1, 1+\delta)$,

$$\deg F_{tA} = \sum_{\substack{z \in F_{tA}^{-1} \\ z \notin \mathbb{R}^n_{\geq 0}}} \operatorname{sign}(\det(\partial F_{tA}(z))) + \operatorname{sign}(\det(\mathbb{I} - tA)),$$

since x has a preimage in the positive orthant when $t > 1$. Now, $\det(\mathbb{I} - tA) = \chi_{tA}(1)$. We have that $\chi_A(1) = 0$. Thus, for $t \in (1, 1+\delta)$ sufficiently small,

$$\operatorname{sign}(\det(\mathbb{I} - tA)) = \operatorname{sign}\left(\left.\frac{d}{dt}\right|_{t=1} \chi_{tA}(1)\right).$$

Now, $\chi_{tA}(1) = \sum_{k=0}^n c_k(A) t^{n-k}$, where $c_k(A)$ is the kth coefficient. Thus

$$\left.\frac{d}{dt}\right|_{t=1} \chi_{tA}(1) = \sum_{k=0}^{n} (n-k) c_k(A) = n \sum_{k=0}^{n-1} c_k(A) - \sum_{k=0}^{n-1} k c_k(A) = -\sum_{k=0}^{n} k c_k(A) = -\chi_A'(1),$$

where $\sum_{k=0}^{n-1} c_k(A) = -c_n(A)$ follows from $\chi_A(1) = 0$. Hence, the desired formula follows.

We finish the proof, by proving the lemmas above.

Proof of Lemma 2. Denote by

$$\mathcal{H} := \{y \in \mathbb{R}^n \mid \text{there is some } i \text{ such that } x_i = 0\} \tag{26}$$

the union of the coordinate hyperplanes, and fix $t \in I_0 \setminus \{1\}$. Then for all $x \in \mathbb{R}^n$, x is a regular point of F_{tA} if and only if $x \notin F_{tA}(\mathcal{H})$. If a preimage z of x does not lie on $\mathcal{H}$, then F_{tA} is differentiable at z. Moreover, at that point, the Jacobian will be of the form $\mathbb{I} - tAS$, for some $S \in \mathcal{S}_n$, and it will be invertible since $t \in I_0 \setminus \{1\}$.

Now, for all $t \in I_0$,

$$\tilde{\mathcal{H}}_t := \bigcup_{S \in \mathcal{S}_n} (\mathbb{I} - tAS)\mathcal{H} \supseteq F_{tA}(\mathcal{H})$$

is a hyperplane arrangement. Therefore we can choose a strictly negative $x \in B_2(-v, r)$ such that $x \notin \tilde{\mathcal{H}}_1$. If we show that $I_0 \ni t \mapsto \operatorname{dist}(x, \tilde{\mathcal{H}}_t)$ is a continuous map, then, since $\operatorname{dist}(x, \tilde{\mathcal{H}}_1) > 0$, we

can choose $\delta > 0$ such that for all $t \in (1-\delta, 1+\delta)$, $\operatorname{dist}(x, \tilde{\mathcal{H}}_t) > 0$. Thus for all $t \in (1-\delta, 1+\delta)$, $x \notin F_{tA}(\mathcal{H})$ and, by the discussion in the above paragraph, x is a regular point of F_{tA}.

To show that $I_0 \ni t \mapsto \operatorname{dist}(x, \tilde{\mathcal{H}}_t)$ is continuous, we only have to observe that for all i, $S \in \mathcal{S}_n$ and $t \in I_0$,

$$(\mathbb{I} - tAS)\{x \mid x_i = 0\}$$

is a hyperplane, since $\{x \mid x_i\}$ does not contain any vector in the kernel of $\mathbb{I} - tAS$. Only $\mathbb{I} - A$ has a kernel, but it is a line spanned by the strictly positive vector v. Moreover, taking the generalized cross product of the $(\mathbb{I} - tAS)e_j$, $j \neq i$, and normalizing it, we can obtain a continuous map

$$I_0 \ni t \mapsto n(S, i, t) \in \mathbb{S}^{n-1}$$

that maps t to a normal vector of $(\mathbb{I} - tAS)\{x \mid x_i = 0\}$. Hence

$$\operatorname{dist}(x, \tilde{\mathcal{H}}_t) = \min\left\{|n(S, i, t)^t x| \mid i \in \{1, \dots, n\},\, S \in \mathcal{S}_n\right\}$$

is a continuous function as we wanted to show.

Now, for all $t \in (1-\delta, 1+\delta)$ and all $S \in \mathcal{S}_n \setminus \{1\}$, we have that

$$(\mathbb{I} - tAS)^{-1}x \notin \mathcal{H}.$$

Note that, by choosing $\delta > 0$ small enough if necessary, we can guarantee $t \in I_0$ and that $\mathbb{I} - tAS$ is invertible. Hence we have that the signs of $(\mathbb{I} - tAS)^{-1}x$ are constant. This means that, independently of $t \in (1-\delta, 1+\delta)$,

$$F_{tA}^{-1}(x) \cap S\mathbb{R}^n_{>0}$$

is either empty or has size one. Now, for $z \in S\mathbb{R}^n_{>0}$, $\operatorname{sign}(\det(\delta F_{tA}(z))) = \operatorname{sign}(\det(\mathbb{I} - tAS))$ is constant. Hence the sum (25) remains constant as desired. □

For the proof of Lemma 3, we will need the following auxiliary lemma.

Lemma 4. *Let $A \in \mathrm{M}_n(\mathbb{R})$. If λ is a simple eigenvalue of A, then*

$$\ker(\lambda\mathbb{I} - A) \cap \operatorname{im}(\lambda\mathbb{I} - A) = 0.$$

Proof of Lemma 4. If $v \in \ker(\lambda\mathbb{I} - A) \cap \operatorname{im}(\lambda\mathbb{I} - A)$ is non-zero, then $v \in \ker(\lambda\mathbb{I} - A)^2$, but $v \notin \ker(\lambda\mathbb{I} - A)$. Thus v is a nontrivial generalized eigenvector of rank 2 of A corresponding to λ. However, λ is a simple eigenvalue of A, so this is impossible. □

Proof of Lemma 3. By Lemma 4, we have that

$$-v \notin (\mathbb{I} - A)\mathcal{H},$$

where $\mathcal{H}$ is the union of the coordinate hyperplanes, as in the proof of Lemma 2. Arguing as in that proof, we have that $I_0 \ni t \mapsto \operatorname{dist}(-v, (\mathbb{I} - tA)\mathcal{H})$ is continuous, and so, since $\operatorname{dist}(-v, (\mathbb{I} - A)\mathcal{H})$ is positive, we have that there is $\delta > 0$ and $r > 0$ such that for all $t \in (1-\delta, 1+\delta) \subseteq I_0$,

$$\operatorname{dist}(-v, (\mathbb{I} - tA)\mathcal{H}) > r.$$

We show now that these are the desired r and δ.

Fix $t \in (1-\delta, 1)$. $F_{tA}(\mathbb{R}^n_{\geq 0})$ is a closed pointed cone, since it is image of a closed pointed cone under the invertible linear map $\mathbb{I} - tA$. This cone contains v, so it does not contain $-v$. Now,

$$\operatorname{dist}(-v, F_{tA}(\mathbb{R}^n_{\geq 0})) = \operatorname{dist}(-v, \operatorname{bd}F_{tA}(\mathbb{R}^n_{\geq 0})) \geq \operatorname{dist}(-v, (\mathbb{I} - A)\mathcal{H}),$$

since the nearest point in $F_{tA}(\mathbb{R}^n_{\geq 0})$ to $-v$ lies in the boundary, $\operatorname{bd}F_{tA}(\mathbb{R}^n_{\geq 0})$ of $F_{tA}(\mathbb{R}^n_{\geq 0})$, which is contained inside $(\mathbb{I} - A)\mathcal{H}$. Hence, by the first paragraph in the proof, $\operatorname{dist}(-v, F_{tA}(\mathbb{R}^n_{\geq 0})) > r$, and so $B(-v, r) \cap F_{tA}(\mathbb{R}^n_{\geq 0})) = \varnothing$, as desired.

Fix $t \in (1, 1+\delta)$. $F_{tA}(\mathbb{R}^n_{>0})$ is a full-dimensional cone, since it is the image of a full-dimensional cone under the invertible linear map $\mathbb{I} - tA$. Now, we have that

$$\operatorname{dist}(-v, \operatorname{bd}F_{tA}(\mathbb{R}^n_{\geq 0})) \geq \operatorname{dist}(-v, (\mathbb{I} - A)\mathcal{H}).$$

Therefore, by the first paragraph in the proof, $\operatorname{dist}(-v, \operatorname{bd}F_{tA}(\mathbb{R}^n_{\geq 0})) > r$, and so $B(-v, r) \subseteq F_{tA}(\mathbb{R}^n_{\geq 0}))$, as desired. □

References

[1] R.W. Cottle, J.-S. Pang, and R. Stone. *The linear complementarity problem.* Academic Press, 1992.

[2] L. Brugnano and V. Casulli. Iterative solution of piecewise linear systems. *SIAM Journal on Scientific Computing*, 30(1):463–472, 2008.

[3] J. Johns, C. Painter-wakefield, and R. Parr. Linear complementarity for regularized policy evaluation and improvement. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems*, volume 23. Curran Associates, Inc., 2010.

[4] H. Nguyen. *Gpu Gems 3.* Addison-Wesley Professional, first edition, 2007.

[5] O.L. Mangasarian and R.R. Meyer. Absolute value equations. *Linear Algebra and Its Applications*, 419:359–367, 2006.

[6] S. M Rump. Theorems of perron-frobenius type for matrices without sign restrictions. *Linear Algebra and Its Applications*, 266:1–42, 1997.

[7] D. Amelunxen and M. Lotz. Average-case complexity without the black swans. *Journal of Complexity*, 41:82–101, 2017.

[8] P. Bürgisser and F. Cucker. *Condition – The Geometry of Numerical Algorithms.* Springer, 2013.

[9] M. Zamani and M. Hladik. Error bounds and a condition number for the absolute value equations. *Math. Program.*, 2022.

[10] A. Neumaier. *Interval methods for systems of equations.* Cambridge University Press, 1990.

[11] S. Gaubert and J. Gunawardena. The perron-frobenius theorem for homogeneous, monotone functions. *Transactions of the American Mathematical Society*, 356(12):4931–4950, 2004.

[12] E. Outerelo and J.M. Ruiz. *Mapping Degree Theory.* American Mathematical Society, 2009.

[13] G. Ottaviani and B. Sturmfels. Matrices with eigenvectors in a given subspace. *Proc. Amer. Math. Soc.*, 141(4):1219–1232, 2013.

[14] D. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms.* Undergraduate Texts in Mathematics. Springer, New York, third edition, 2007. An introduction to computational algebraic geometry and commutative algebra.