ABSTRACT

PROFILI, DANIEL ALBERT. Conditions for Eigenvalue Configurations of Two Real Symmetric Matrices. (Under the direction of Hoon Hong).

For two real symmetric matrices, their eigenvalue configuration is the relative arrangement of their eigenvalues on the real line. We consider the following problem: given two parametric real symmetric matrices and an eigenvalue configuration, find a simple condition on the parameters such that the two matrices have the given eigenvalue configuration.

The eigenvalue configuration problem generalizes Descartes' rule of signs, which is a fundamental and widely used tool in computational real algebraic geometry, with numerous applications in science and engineering. Hence we expect that the eigenvalue configuration problem may have even more applications (for example: rank updates and constrained optimization).

In this thesis, we develop two distinct approaches for this problem. In the first approach, we solve the problem by producing a set of matrices whose entries are constructed from the entries of the input matrices. The signatures of these new matrices determine the eigenvalue configuration via a linear map which depends only on the sizes of the two matrices and not on their entries. In the second approach, we solve the problem by producing a set of symmetric polynomials whose roots are constructed from the eigenvalues of the input matrices. The numbers of positive roots of these new polynomials determine the eigenvalue configuration via another linear map which depends only on the sizes of the two matrices and not on their entries.

© Copyright 2026 by Daniel Albert Profili

All Rights Reserved

Conditions for Eigenvalue Configurations of Two Real Symmetric Matrices

by
Daniel Albert Profili

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina
2026

APPROVED BY:

Yairon Cid-Ruiz	Irina Kogan
Seth Sullivant	Juan Rafael Sendra External Member

Hoon Hong
Chair of Advisory Committee

BIOGRAPHY

Daniel Profili was born in Marietta, Georgia, a suburb of metro Atlanta. He earned Bachelor's degrees in Mathematics and Computer Science at Georgia Tech in 2021. Later that year, he began a PhD in Mathematics at North Carolina State University under the supervision of Professor Hoon Hong.

ACKNOWLEDGEMENTS

First, thank you to my advisor, Dr. Hoon Hong, for your guidance and support over the past five years. Your patience and wisdom was invaluable on innumerable occasions. Thank you to my committee for your help and support, particularly Dr. Juan Rafael Sendra, for editing my writing (many times).

Thanks to my fellow graduate students and dear friends Francisco Ponce, Jack Reever, and Elijah Rutter, and especially thanks to Francisco for putting up with me as a roommate for five years. Thank you to Justin Htay for being my friend for such a long time. Thanks also to my parents and my brother, and finally, thanks to Demeter.

TABLE OF CONTENTS

Authorship Statement v

Chapter 1 Introduction 1
1.1 Problem and Motivation 1
1.2 Contributions 3
1.3 Background 4

Chapter 2 Generic Pairs of Symmetric Matrices 14
2.1 Problem Statement 14
2.2 Signature Approach 19
2.2.1 Main Theorem 19
2.2.2 Proof 27
2.3 Symmetric Polynomials Approach 58
2.3.1 Main Theorem 58
2.3.2 Proof 64

Chapter 3 Arbitrary Pairs of Symmetric Matrices 80
3.1 Problem Statement 80
3.2 Signature Approach 83
3.2.1 Main Theorem 83
3.2.2 Proof 83
3.3 Symmetric Polynomials Approach 94
3.3.1 Main Theorem 94
3.3.2 Proof 97

References 109

Authorship Statement

Contributions of Daniel Profili and coauthors are listed below for each chapter.

Chapter 1

1. Daniel Profili: writing
2. Professor Hoon Hong: guidance on writing
3. Professor J. Rafael Sendra: guidance on writing

Hong, Hoon, Daniel Profili, and J. Rafael Sendra. "Conditions for eigenvalue configurations of two real symmetric matrices (signature approach)." SIAM Journal on Matrix Analysis and Applications (pending). (2026).

Chapter 2

1. Daniel Profili: conceptualizing, proving, writing
2. Professor Hoon Hong: guidance on conceptualizing, proving, writing
3. Professor J. Rafael Sendra: guidance on conceptualizing, proving, writing

Hong, Hoon, Daniel Profili, and J. Rafael Sendra. "Conditions for eigenvalue configurations of two real symmetric matrices (signature approach)." SIAM Journal on Matrix Analysis and Applications (pending). (2026).

Hong, Hoon, Daniel Profili, and J. Rafael Sendra. "Conditions for eigenvalue configurations of two real symmetric matrices (symmetric polynomial approach)." arXiv preprint arXiv:2401.00089 (2026).

Chapter 3

1. Daniel Profili: conceptualizing, proving, writing
2. Professor Hoon Hong: guidance on conceptualizing, proving, writing
3. Professor J. Rafael Sendra: guidance on conceptualizing, proving, writing

Hong, Hoon, Daniel Profili, and J. Rafael Sendra. "Conditions for eigenvalue configurations of two real symmetric matrices (signature approach)." SIAM Journal on Matrix Analysis and Applications (pending). (2026).

Hong, Hoon, Daniel Profili, and J. Rafael Sendra. "Conditions for eigenvalue configurations of two real symmetric matrices (symmetric polynomial approach)." arXiv preprint arXiv:2401.00089 (2026).

Use of generative artificial intelligence: No generative artificial intelligence was used to write this dissertation.

Chapter 1

Introduction

In this chapter, we give an informal overview of the problem, state the contributions of this thesis, and provide some relevant mathematical background.

1.1 Problem and Motivation

For two real symmetric matrices F and G, their eigenvalue configuration is the relative arrangement of their eigenvalues on the real line. In this thesis, we tackle the following problem: Given two parametric real symmetric matrices and an eigenvalue configuration, find a quantifier-free necessary and sufficient condition on the parameters so that the eigenvalue configuration of the matrices is the given one.

For some historical context, a fundamental problem in computational algebra and geometry is to determine quantifier-free necessary and sufficient conditions on the coefficients of a polynomial such that its roots lie in a given subset of the complex plane. This problem appears in a wide breadth of fields, including the theory of stable polynomials [1], combinatorics [2], graph theory [3], and functional analysis [4], among others. The eigenvalue configuration problem is a specific case of this problem.

The eigenvalue configuration problem generalizes Descartes' rule of signs, which is a fundamental and widely used (see e.g. [5], [6], [7]) tool in computational real algebraic geometry. Recall that Descartes' rule of signs states that, for a real polynomial g, the number of positive real roots of g (counted with multiplicity) is bounded above by the sign variation count of the coefficients of g; i.e., the number of times consecutive coefficients change sign, ignoring zeros. It is especially useful when the polynomial g has only real roots since the number of positive real roots of g is exactly the sign variation count of the coefficients. Viewed as an eigenvalue configuration problem, Descartes' rule of signs can be thought of as determining the eigenvalue

configuration of the 1×1 matrix $F = [0]$ and a symmetric matrix G whose characteristic polynomial is g. The eigenvalue configuration problem therefore extends Descartes' rule of signs by allowing two polynomials of arbitrary degrees; in addition, we reframe the problem slightly by considering characteristic polynomials of real symmetric matrices, as these have only real roots and occur naturally in many areas. Given the wide applicability of Descartes' rule of signs, we expect that a generalization may have even more applications. Below, we list a couple of potential applications.

- *Rank update*: Let $F \in \mathbb{R}^{n\times n}$ be symmetric and let $r \leq n$. A rank-r update of F is

$$G = F + UCU^T,$$

where

$$\begin{aligned} &U \in \mathbb{R}^{n\times r} \text{ is orthogonal and satisfies } \operatorname{rank}(UU^T) = r, \\ &C = \operatorname{diag}(C_1, \ldots, C_r) \in \mathbb{R}^{r\times r}_{>0} \text{ is diagonal.} \end{aligned}$$

A natural question is how the eigenvalues of F are impacted under a rank update; that is, where are the relative locations of the eigenvalues of the updated matrix G.

When $r = 1$, there is a well-known and very useful result that each eigenvalue β_i of G can move some distance closer to α_{i+1}, but cannot reach it; i.e., the eigenvalues interlace. A natural followup question is what happens to eigenvalues of F under rank-r updates for arbitrary r. Any progress in the eigenvalue configuration problem could provide a systematic algebraic tool for studying rank updates.

- *Constrained optimization*: Consider a system defined parametrically by the real symmetric matrices $F = [a_{ij}]$ and $G = [b_{ij}]$. In many such systems, some desired characteristic or constraint (e.g. stability) may be dependent on a certain configuration, c, of the eigenvalues of F and G. It may also be desirable to optimize some objective function $p(a_{ij}, b_{ij})$ of the system which depends on the parameters (e.g. to minimize cost or maximize profit), giving the constrained optimization problem

$$\begin{aligned} \underset{a_{ij}, b_{ij}}{\text{optimize}} \quad & p(a_{ij}, b_{ij}) \\ \text{subject to} \quad & \text{eigenvalues of } F \text{ and } G \text{ are arranged according to } c. \end{aligned}$$

However, there are two difficulties. First, the constraint refers to the eigenvalues of F and G and therefore involves quantifiers (see Section 2.1 for more details) and there is

no standard optimization theory for such constraints. Second, the objective function depends on the parameters of F and G, while the constraint depends on the eigenvalues. This is an issue because the eigenvalues of matrices larger than 4×4 cannot, in general, be expressed in terms of field operations and radicals on the parameters. Hence, any answer to this challenge would provide a method to obtain an equivalent constraint to "eigenvalues of F and G are arranged according to c" which has no quantifiers and is in terms of the parameters a_{ij} and b_{ij}.

The eigenvalue configuration problem is essentially a special case of the real quantifier elimination problem. Thus, it could in principle be solved using real quantifier elimination algorithms and software tools (see e.g. [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]). However, there are two major difficulties.

- The computation quickly becomes very time-consuming as m and n grow, making this approach practically infeasible.
- The output formula looks chaotic, making it practically incomprehensible.

As a consequence of these limitations, we need to develop a specialized approach which exploits the nature of the eigenvalue configuration problem.

1.2 Contributions

In this section, we summarize the original contributions in this thesis.

- Chapter 2: We tackle the eigenvalue configuration problem for *generic* pairs of matrices; that is, pairs of matrices that do not share any eigenvalues.
  - Section 2.1: We first develop a notion of eigenvalue configuration for generic pairs of matrices F and G.
  - Section 2.2: We give an approach based on the *signature of matrices*. We solve the problem by producing a set of matrices whose entries are constructed from the entries of the input matrices. The signatures of these new matrices determine the eigenvalue configuration via a linear map which depends only on the sizes of the two matrices and not on their entries.
  - Section 2.3: We give an entirely distinct approach based on the theory of *symmetric polynomials*. We solve the problem by producing a set of symmetric polynomials whose roots are constructed from the eigenvalues of the input matrices. The

numbers of positive roots of these new polynomials determine the eigenvalue configuration via another linear map which depends only on the sizes of the two matrices and not on their entries.

- Chapter 3: We tackle the eigenvalue configuration problem for arbitrary pairs of matrices by extending the above results for the generic case.
  - Section 3.1: We generalize the definition of eigenvalue configuration to arbitrary pairs of matrices. For an arbitrary pair of matrices F and G, its eigenvalue configuration is defined as the average eigenvalue configuration over all sufficiently small perturbations of the eigenvalues of F.
  - Section 3.2: We prove that the result from Section 2.2 still holds for arbitrary pairs of matrices.
  - Section 3.3: We prove that the result from Section 2.3 holds, with a slight extension, for arbitrary pairs of matrices.

1.3 Background

In this section, we will discuss previous work in this area, as well as some mathematical preliminaries. To our knowledge, there has not been any published work directly concerning the eigenvalue configuration problem, or any other generalization of Descartes' rule of signs to multiple univariate polynomials. Hence, in this section, we will focus on the fundamental tools we will use in this thesis from the areas of real algebraic geometry, combinatorics, and real polynomials.

Real symmetric matrices

Throughout this thesis, most matrices will be real symmetric. We rely heavily on the fact that all eigenvalues of real symmetric matrices are real.

Theorem 1. *Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Then every eigenvalue of A is real.*

Proof. Let λ be an arbitrary (not necessarily real) eigenvalue of A. Then there exists a nonzero vector $v \in \mathbb{R}^n$ so that $Av = \lambda v$. Let $*$ denote the complex conjugate of a matrix or vector. Taking the complex conjugate of both sides, we have that

$$(Av)^* = (\lambda v)^* = A^* v^* = \lambda^* v^*.$$

Since A is real, we have $A^* = A$. Thus $Av^* = \lambda^* v^*$. Further, since A is symmetric, we have $A^T = A$. Thus we have

$$(v^*)^T Av = (v^*)^T \lambda v = \lambda(v^* \cdot v).$$

On the other hand,

$$(v^*)^T Av = (Av^*)^T v = (\lambda^* v^*)^T v = \lambda^*(v^* \cdot v).$$

Hence $\lambda(v^* \cdot v) = \lambda^*(v^* \cdot v)$. Since v is nonzero, we have $v^* \cdot v \neq 0$, which implies $\lambda = \lambda^*$. Therefore λ is real. □

Real quantifier elimination

While developing the eigenvalue configuration problem in Chapters 2 and 3, we refer to the fact that *quantifier elimination* is feasible over the real numbers. Quantifier elimination (over the real numbers) refers to the following problem.

Problem 2 (Quantifier elimination over the real numbers)**.**

In: P*, a quantified Boolean combination of polynomial equalities and inequalities involving real variables*

Out: Q*, an equivalent quantifier-free Boolean combination of polynomial equalities and inequalities in the free variables of* P

An algorithm solving this problem was first given by the logician Alfred Tarski in [30] as the proof of the following theorem.

Theorem 3 (Tarski)**.** *For every quantified Boolean combination of polynomial equalities and inequalities involving real variables, there exists an equivalent quantifier-free Boolean combination of polynomial equalities and inequalities in the free variables.*

Descartes' rule of signs

The fundamental real root counting tool we will use heavily throughout the thesis is Descartes' rule of signs [31], which allows us to count the number of positive roots of a polynomial via the signs of its coefficients.

Definition 4 (Sign variation count)**.** *Let $f = a_m x^m + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$. The **sign variation count** of f, denoted $v(f)$, is the number of times adjacent elements in the sequence $(a_m, a_{m-1}, \ldots, a_1, a_0)$ have opposite sign, ignoring zeros.*

At times, we will overload the symbol v to refer to the sign variation count of a sign sequence instead, which is defined similarly.

Example 5.

1. $v(x^2+x+1)=0$
2. $v(x^4-1)=1$
3. $v(x^7-x^6+x-1)=3$
4. $v(+,-,+,-)=3$.

△

Theorem 6 (Descartes' rule of signs)**.** *Let $f \in \mathbb{R}[x]$. Let $p(f)$ denote the number of positive real roots of f, counting with multiplicity. Then $p(f) \leq v(f)$.*

Proof. We proceed by induction on the degree of f. For the base case, if $\deg f = 0$, then clearly $p(f)=0$ and $v(f)=0$; thus the claim holds for $\deg f = 0$.

Now suppose the claim holds for $\deg f < n$. It suffices to show the claim holds for $\deg f = n$. Let $f = a_k x^{e_k} + \cdots + a_0 x^{e_0}$, where

- $a_k, \ldots, a_0 \neq 0$
- $e_k > e_{k-1} > \cdots > e_0 \geq 0$.

We will now split the proof into cases.

$e_0 > 0$: Let $\overline{f} = f/x^{e_0} = a_k x^{e_k - e_0} + \cdots + a_0 x^0$. Note that $\deg \overline{f} < \deg f = n$. Further, note that $p(f) = p(\overline{f}) \leq v(\overline{f}) = v(f)$. Hence, the claim holds.

$e_0 = 0$: Let $\overline{f} = f'/x^{e_1 - 1} = a_k e_k x^{e_k - e_1} + \cdots + a_1 e_1$. We further split this case into two sub-cases.

$a_1 a_0 < 0$: Note that $v(f) = v(\overline{f}) + 1$. Further, by Rolle's theorem, we have $p(f) \leq p(\overline{f}) + 1$. Therefore $p(f) \leq p(\overline{f}) + 1 \leq v(\overline{f}) + 1 = v(f)$.

$a_1 a_0 > 0$: This case is treated similarly to the previous case.

□

Since in this work we deal exclusively deal with polynomials having only real roots, we will rely heavily on the following corollary to Descartes' rule of signs.

Corollary 7. *If all roots of f are real, then $p(f) = v(f)$.*

Proof. From Theorem 6, we have

1. $p(f) + d_1 = v(f)$, for some $d_1 \geq 0$
2. $n(f) + d_2 = v(f(-x))$, for some $d_2 \geq 0$, where $n(f)$ denotes the number of negative real roots of f, counting multiplicity.

Further, we have that $z(f) = \text{taildeg}(f)$, where $z(f)$ denotes the multiplicity of zero as a root of f. Adding those together, we have

$$p(f) + n(f) + z(f) + d_1 + d_2 = v(f) + v(f(-x)) + \text{taildeg}(f).$$

Since all roots of f are real, we have

$$\deg f + d_1 + d_2 = v(f) + v(f(-x)) + \text{taildeg}(f).$$

By induction on the number of terms of f, one can show that $v(f)+v(f(-x))+\text{taildeg}(f) \leq \deg f$. Putting these together, we get

$$\deg f + d_1 + d_2 \leq \deg f.$$

Since $d_1 + d_2 \leq 0$, this means that $d_1 = d_2 = 0$. Therefore

$$p(f) + d_1 = p(f) = v(f).$$

□

Example 8. *Let $f = (x-1)(x+2)(x+3) = x^3 + 4x^2 + x - 6$. Clearly f has only real roots and has exactly one positive root. Note*

$$\begin{aligned} v(f) &= v(x^3 + 4x^2 + x - 6) = 1 \\ &= p(f). \end{aligned}$$

△

Generalized real root counting

In the development of the signature approach in Chapters 2 and 3, we will leverage an algorithm which solves the *univariate generalized real root counting problem*, which refers to the following problem.

Problem 9.

In: $f, g_1, \ldots, g_r \in \mathbb{R}[x]$

Out: $N \in \mathbb{N}^{3^r}$ *whose entries are indexed by* $s \in \{-, 0, +\}^r$ *where*

$$N_s = \{x \in \mathbb{R} : f(x) = 0 \ \wedge \ \text{sign}(g_1(x), \ldots, g_r(x)) = s\}.$$

The solution to this problem relies on the notion of the signature of a matrix.

Definition 10 (Signature of a matrix)**.** *The signature of a square matrix* M*, denoted* $\sigma(M)$*, is*

$$\sigma(M) = \#\text{ positive eigenvalues of } M \ - \ \#\text{ negative eigenvalues of } M.$$

Theorem 11 (Ben-Or, Kozen, Reif [32])**.** *Let*

- $S \in \mathbb{Z}^{3^r}$ *be a column vector whose entries are indexed by* $e \in \{0,1,2\}^r$ *where* $S_e = \sigma(g_1^{e_1} \cdots g_r^{e_r})$
- $\overline{H} \in \mathbb{Z}^{3^r \times 3^r}$ *be the matrix given by* $\overline{H} = \underbrace{\overline{H}_1 \otimes \cdots \otimes \overline{H}_1}_{r}$ *where* $\overline{H}_1 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

Then $N = \overline{H}^{-1} S$.

Proof. We will prove a stronger claim since it is easier for an inductive argument. For this, we will use two shorthand notations.

$$\overline{H}_r \ = \ \underbrace{\overline{H}_1 \otimes \cdots \otimes \overline{H}_1}_{r} \quad \text{and } \overline{H}_0 = [1], \text{ the } 1 \times 1 \text{ matrix.}$$

$$\begin{aligned} \langle e_1 \cdots e_i \mid s_{i+1} \cdots s_r \rangle \ = \ & \#\left\{x \in \mathbb{R} : f(x) = 0, \ g_1^{e_1} \cdots g_i^{e_i} > 0, \ \text{sign}\left(g_{i+1}\right) = s_{i+1}, \ldots, \text{sign}\left(g_r\right) = s_r\right\} - \\ & \#\left\{x \in \mathbb{R} : f(x) = 0, \ g_1^{e_1} \cdots g_i^{e_i} < 0, \ \text{sign}\left(g_{i+1}\right) = s_{i+1}, \ldots, \text{sign}\left(g_r\right) = s_r\right\} \end{aligned}$$

If e_i or s_i are replaced with $*_i$ then they are expanded lexicographically with $0,1,2$ or $-,0,+$

respectively. For example

$$\langle e_1 *_2 \mid s_3 *_4 \rangle = \begin{bmatrix} \langle e_1 0 \mid s_3 - \rangle \\ \langle e_1 0 \mid s_3 0 \rangle \\ \langle e_1 0 \mid s_3 + \rangle \\ \langle e_1 1 \mid s_3 - \rangle \\ \langle e_1 1 \mid s_3 0 \rangle \\ \langle e_1 1 \mid s_3 + \rangle \\ \langle e_1 2 \mid s_3 - \rangle \\ \langle e_1 2 \mid s_3 0 \rangle \\ \langle e_1 2 \mid s_3 + \rangle \end{bmatrix}$$

Now we are ready to state a stronger claim.

Claim: $\underset{0\le k\le r}{\forall} \underset{e_1,\ldots,e_k\in\{0,1,2\}}{\forall} \langle e_1 \cdots e_k *_{k+1} \cdots *_r \mid\rangle = \overline{H}_{r-k} \langle e_1 \cdots e_k \mid *_{k+1} \cdots *_r \rangle$.

Note that the theorem is an instance ($k = 0$) of the above claim, since

$$S = \langle *_1 \cdots *_r \mid\rangle = \overline{H}_r \langle \mid *_1 \cdots *_r \rangle = \overline{H}_r N.$$

We will prove the claim by induction on k, from $k = r$ down to $k = 0$. The claim obviously holds for $k = r$ since $\overline{H}_0 = [1]$. Assume that the claim holds for $k+1$. We will to show that the claim holds for k. Let $e_1, \ldots, e_k \in \{0,1,2\}$ be arbitrary but fixed. We need to show

$$\langle e_1 \cdots e_k *_{k+1} \cdots *_r \mid\rangle = \overline{H}_{r-k} \langle e_1 \cdots e_k \mid *_{k+1} \cdots *_r \rangle .$$

Note

$$\begin{aligned}
&\langle e_1 \cdots e_k *_{k+1} \cdots *_r \mid\rangle \\
&= \begin{bmatrix} \langle e_1 \cdots e_k 0 *_{k+2} \cdots *_r \mid\rangle \\ \langle e_1 \cdots e_k 1 *_{k+2} \cdots *_r \mid\rangle \\ \langle e_1 \cdots e_k 2 *_{k+2} \cdots *_r \mid\rangle \end{bmatrix} \\
&= \begin{bmatrix} \overline{H}_{r-(k+1)} \langle e_1 \cdots e_k 0 \mid *_{k+2} \cdots *_r \rangle \\ \overline{H}_{r-(k+1)} \langle e_1 \cdots e_k 1 \mid *_{k+2} \cdots *_r \rangle \\ \overline{H}_{r-(k+1)} \langle e_1 \cdots e_k 2 \mid *_{k+2} \cdots *_r \rangle \end{bmatrix} \text{ by applying the induction hypothesis to each block} \\
&= \begin{bmatrix} \overline{H}_{r-(k+1)} (\langle e_1 \cdots e_k \mid - *_{k+2} \cdots *_r \rangle + \langle e_1 \cdots e_k \mid 0 *_{k+2} \cdots *_r \rangle + \langle e_1 \cdots e_k \mid + *_{k+2} \cdots *_r \rangle) \\ \overline{H}_{r-(k+1)} (\langle e_1 \cdots e_k \mid + *_{k+2} \cdots *_r \rangle - \langle e_1 \cdots e_k \mid - *_{k+2} \cdots *_r \rangle) \\ \overline{H}_{r-(k+1)} (\langle e_1 \cdots e_k \mid - *_{k+2} \cdots *_r \rangle + \langle e_1 \cdots e_k \mid + *_{k+2} \cdots *_r \rangle) \end{bmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} \overline{H}_{r-(k+1)} & \overline{H}_{r-(k+1)} & \overline{H}_{r-(k+1)} \\ -\overline{H}_{r-(k+1)} & & \overline{H}_{r-(k+1)} \\ \overline{H}_{r-(k+1)} & & \overline{H}_{r-(k+1)} \end{bmatrix} \begin{bmatrix} \langle e_1 \cdots e_k \mid -*_{k+2} \cdots *_r \rangle \\ \langle e_1 \cdots e_k \mid 0 *_{k+2} \cdots *_r \rangle \\ \langle e_1 \cdots e_k \mid +*_{k+2} \cdots *_r \rangle \end{bmatrix} \\
&= \left(\overline{H}_1 \otimes \overline{H}_{r-(k+1)}\right) \langle e_1 \cdots e_k \mid *_{k+1} \cdots *_r \rangle \\
&= \overline{H}_{r-k} \langle e_1 \cdots e_k \mid *_{k+1} \cdots *_r \rangle
\end{aligned}
$$

□

Rolle's theorem

During our development of the signature approach in Chapter 2, we will use Rolle's theorem to establish some properties of polynomials having only real roots.

Theorem 12 (Rolle, 1691)**.** *Let $f \in \mathbb{R}[x]$ and let $a < b \in \mathbb{R}$ be such that $f(a) = f(b) = 0$. Then there exists $c \in (a, b)$ so that $f'(c) = 0$.*

The proof can be found in most introductory real analysis texts; for example, [33].

Symmetric polynomials

During our development of the symmetric polynomials approach in Chapter 2, we will use the theory of symmetric polynomials. A *symmetric polynomial* is a multivariate polynomial which is invariant under every permutation of the indices of the variables.

Example 13. *Let*

$$h(x_1, x_2) = x_1 x_2 + x_1 + x_2 \in \mathbb{R}[x_1, x_2]$$

and consider a permutation τ of two elements. Clearly, if τ is the identity permutation, then we have $h(x_{\tau(1)}, x_{\tau(2)}) = h(x_1, x_2)$. If τ is the permutation which swaps 1 and 2, then

$$
\begin{aligned}
h(x_{\tau(1)}, x_{\tau(2)}) &= h(x_2, x_1) \\
&= x_2 x_1 + x_2 + x_1 \\
&= x_2 x_1 + x_2 + x_1 \\
&= h(x_1, x_2).
\end{aligned}
$$

Thus, $h(x_1, x_2) = h(x_{\tau(1)}, x_{\tau(2)})$ for every permutation τ of two elements. Therefore h is a symmetric polynomial. △

Definition 14. *Let $x = (x_1, \ldots, x_n)$ and let*

$$\begin{aligned}
e_1(x) &= \sum_{1 \le i_1 \le n} x_{i_1} \\
e_2(x) &= \sum_{1 \le i_1 < i_2 \le n} x_{i_1} x_{i_2} \\
&\vdots \\
e_k(x) &= \sum_{1 \le i_1 < \cdots < i_k \le n} x_{i_1} \cdots x_{i_k} \\
&\vdots \\
e_n(x) &== \sum_{1 \le i_1 < \cdots < i_n \le n} x_{i_1} \cdots x_{i_n}.
\end{aligned}$$

*The polynomial $e_k(x)$, or just e_k when the list of variables is clear, is called the k-th **elementary symmetric polynomial** on the variables $x = (x_1, \ldots, x_n)$.*

Example 15. *For $x = (x_1, x_2)$, we have*

$$\begin{aligned}
e_1(x) &= \textstyle\sum_{1 \le i_1 \le 2} x_{i_1} &&= x_1 + x_2 \\
e_2(x) &= \textstyle\sum_{1 \le i_1 < i_2 \le 2} x_{i_1} x_{i_2} &&= x_1 x_2.
\end{aligned}$$

△

Example 16. *For $x = (x_1, x_2, x_3, x_4)$, we have*

$$\begin{aligned}
e_1(x) &= \textstyle\sum_{1 \le i_1 \le 4} x_{i_1} &&= x_1 + x_2 + x_3 + x_4 \\
e_2(x) &= \textstyle\sum_{1 \le i_1 < i_2 \le 4} x_{i_1} x_{i_2} &&= x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4 \\
e_3(x) &= \textstyle\sum_{1 \le i_1 < i_2 < i_3 \le 4} x_{i_1} x_{i_2} x_{i_3} x_{i_4} &&= x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 \\
e_4(x) &= \textstyle\sum_{1 \le i_1 < i_2 < i_3 < i_4 \le 4} x_{i_1} x_{i_2} x_{i_3} x_{i_4} &&= x_1 x_2 x_3 x_4
\end{aligned}$$

△

Our symmetric polynomials approach relies on the Fundamental Theorem of Symmetric Polynomials, which is the fact that any symmetric polynomial can be written in terms of the elementary symmetric polynomials.

Theorem 17 (Fundamental Theorem of Symmetric Polynomials [34])**.**

Let $h \in \mathbb{R}[x_1, \ldots, x_n]$ be a symmetric polynomial and let $x = (x_1, \ldots, x_n)$. Then there exists a

polynomial $u \in \mathbb{R}[y_1, \ldots, y_n]$ *such that*

$$h = u(e_1(x), \ldots, e_n(x)).$$

Proof. We prove constructively by giving an algorithm adapted from [35].

Algorithm 18 (Fundamental Theorem of Symmetric Polynomials (FTSP))**.**

In: $h \in \mathbb{R}[X]$ *symmetric polynomial*

$x = (x_1, \ldots, x_n)$*, a list of variables*

Out: $u \in \mathbb{R}[y_1, \ldots, y_n]$ *such that* $h = u(e_1(x), \ldots, e_n(x))$

1. *If* $n = 1$ *or* $\text{totaldeg}(h) = 0$ *then*
   (a) $u \leftarrow h(y_1, \ldots, y_n)$
   (b) *Return* u
2. $h_n \leftarrow h(x_1, \ldots, x_{n-1}, 0)$
3. $u_n \leftarrow \text{FTSP}(h_n, (x_1, \ldots, x_{n-1}))$
4. $g \leftarrow u_n(e_1, \ldots, e_{n-1})$
5. $d \leftarrow (h - g)/e_n$
6. $w \leftarrow FTSP(d, (x_1, \ldots, x_n))$
7. $u \leftarrow w\, y_n + u_n$
8. *Return* u

□

Example 19. *Let* $h = x_1^2 x_2^2 - 2x_1 x_2 + x_1 + x_2 \in \mathbb{R}[x_1, x_2]$*. By inspection, one can see that* h *is symmetric. The elementary symmetric polynomials in* $\mathbb{R}[x_1, x_2]$ *are*

$$e_1(x) = x_1 + x_2$$
$$e_2(x) = x_1 x_2.$$

Now observe that

$$h = x_1^2 x_2^2 - 2x_1 x_2 + x_1 + x_2$$

$$= (x_1 x_2)^2 - 2(x_1 x_2) + (x_1 + x_2)$$
$$= (e_2)^2 - 2e_2 + e_1.$$

Hence, if $u = y_2^2 - 2y_2 + y_1$, *then*

$$\begin{aligned} u(e_1(x), e_2(x)) &= e_2^2 - 2e_2 + e_1 \\ &= (x_1 x_2)^2 - 2(x_1 x_2) + (x_1 + x_2) \\ &= h(x_1, x_2). \end{aligned}$$

△

Chapter 2

Generic Pairs of Symmetric Matrices

In this chapter, we give two different approaches for the eigenvalue configuration problem for "generic" pairs of input matrices (i.e., the matrices do not share eigenvalues). In Chapter 3, we will generalize this to arbitrary pairs.

2.1 Problem Statement

In this section, we will state the problem precisely.

Defining eigenvalue configuration

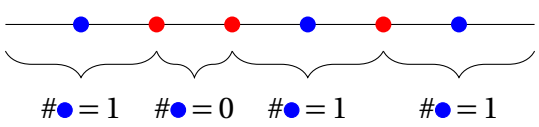
First, we will construct a suitable definition of eigenvalue configuration of generic pairs of matrices. Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be real symmetric matrices. First, let us consider the simplest case when F and G have only simple eigenvalues and do not share any eigenvalues. Since F and G are real symmetric, all their eigenvalues are real. This means we can meaningfully consider the ordering of both matrices' eigenvalues on the real line.

Consider the following example. Suppose F and G are 3×3 real symmetric matrices such that their eigenvalues are arranged as in the following diagram, where the red points denote eigenvalues of F and blue points denote eigenvalues of G.



Our goal is to define a notion of eigenvalue configuration which encapsulates the information depicted by diagram. One natural way to do this is to count the number of blue points between each pair of red points. In this example, there are 3 red points, and so they partition the real line into 4 open intervals. Then, we count the number of blue points in each interval,

as in the diagram below.



The number of blue points in each of these intervals is, respectively, $(1,0,1,1)$. Note that the sum of the entries of this vector is exactly the number of eigenvalues of G. Since all roots of G are real, we can save some space by omitting one of the counts; by arbitrary choice, we can omit the first count (i.e. the number of blue points which lie to the left of the first red point), which could later be reobtained by adding up the remaining entries and subtracting from the size of G. Thus, in this case, we will encode the eigenvalue configuration as $(0,1,1)$.

Note that this definition is inherently biased. We could just as easily have instead counted the number of red points between each pair of blue points. This choice is mostly arbitrary; however, in many potential applications (some of which we discuss in the next section), we are interested in the eigenvalue configuration of a matrix F and another matrix derived by F; for example, one might consider a perturbation of the entries of F to study how the eigenvalues change. In these contexts, it is natural to bias our definition in favor of F as the "base" matrix.

Now, we will make this definition precise. First, we introduce some notation.

Notation 20.

1. Let $F = [a_{ij}] \in \mathbb{R}^{m\times m}$ and $G = [b_{ij}] \in \mathbb{R}^{n\times n}$ be real symmetric matrices.
2. Let $\alpha = (\alpha_1, \ldots, \alpha_m)$ be the eigenvalues of F.

   Let $\beta = (\beta_1, \ldots, \beta_n)$ be the eigenvalues of G.

   Since F and G are real symmetric, all their eigenvalues are real. Thus without losing generality, let us index the eigenvalues so that $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_m$ and $\beta_1 \le \beta_2 \le \cdots \le \beta_n$.
3. Let A_t denote the set $\{x \in \mathbb{R} : \alpha_t < x < \alpha_{t+1}\}$ for $t = 1, \ldots, m$, where $\alpha_{m+1} = \infty$.

We have already addressed the case where F and G have simple and distinct eigenvalues; in that case, the eigenvalue configuration is the vector $c = (c_1, \ldots, c_m)$ where $c_t = \#\{i : \beta_i \in A_t\}$.

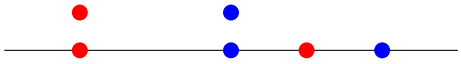
Next, let us generalize slightly to the case where F and G still do not share eigenvalues, but multiple eigenvalues are allowed. We will refer to this case often in the rest of the thesis, so we will give a special name.

Definition 21 (Generic)**.** *We say that the pair of real symmetric matrices F and G is **generic** if F and G do not share eigenvalues.*

Note that this definition describes a property of pairs *of matrices. We may sometimes be more casual with this phrasing and simply say that "F and G are generic matrices."*

The motivation for this terminology is the fact that if F and G are generic, then sufficiently small perturbation of the eigenvalues of F and G preserves the relative arrangement of their eigenvalues.

In the case of generic matrices which **do** contain multiple (repeated) eigenvalues, we can use the previous definition with one minor caveat. For instance, suppose F and G are such that their eigenvalues are arranged as in the following diagram.



In this case, the only extra consideration we need to make is that the interval between the first two red points is the empty set, so the corresponding entry in the eigenvalue configuration vector will be zero. This leads to the following definition for the eigenvalue configuration of generic matrices.

Definition 22 (Eigenvalue configuration for generic matrices)**.** *The eigenvalue configuration of generic pairs of matrices F and G, written as* $\mathrm{GEC}(F,G)$*, is the tuple*

$$c = (c_1, \ldots, c_m)$$

where

$$c_t = \#\{i : \beta_i \in A_t\}.$$

Example 23 (Eigenvalue configuration of generic matrices)**.** *Let $F \in \mathbb{R}^{5\times 5}$ and $G \in \mathbb{R}^{7\times 7}$ be symmetric matrices such that their corresponding eigenvalues are*

$$\alpha = (0,0,3,9,12), \qquad \beta = (-1,2,4,4,8,10,10).$$

Note that F and G are generic since they do not share any eigenvalues. The eigenvalues are arranged on the real line as in the following diagram.



Then

$$\begin{aligned} A_1 &= (\alpha_1, \alpha_2) \\ A_2 &= (\alpha_2, \alpha_3) \quad \ni \beta_2 \\ A_3 &= (\alpha_3, \alpha_4) \quad \ni \beta_3, \beta_4, \beta_5 \\ A_4 &= (\alpha_4, \alpha_5) \quad \ni \beta_6, \beta_7 \\ A_5 &= (\alpha_5, \infty) \end{aligned}$$

Therefore

$$\mathrm{GEC}(F, G) = (0, 1, 3, 2, 0).$$

Note that we counted eigenvalues of G with multiplicity. △

Stating the problem

Now, we will state the problem precisely. The goal of this thesis is to develop two algorithms for the following problem.

Problem 24 (Generic eigenvalue configuration problem)**.**

In: $F \in \mathbb{R}[p]^{m\times m}$ *and* $G \in \mathbb{R}[p]^{n\times n}$*, generic pair of symmetric matrices where* p *is a finite set of parameters.*

$c \in \mathbb{N}^m$*, an eigenvalue configuration.*

Out: a "simple condition" on p *such that* $c = \mathrm{GEC}(F, G)$.

Let us now clarify precisely what the problem is and what the challenge is. This amounts to clarifying what "simple condition" means. To illustrate, consider the following example. Suppose that F and G are generic. We will write down an equivalent condition for $c = \mathrm{GEC}(F, G)$

by explicitly writing all quantifiers and all implicit assumptions. We have

$$c = \mathrm{GEC}(F,G) \iff \exists_{\substack{\alpha_1,\ldots,\alpha_m \\ \beta_1,\ldots,\beta_n \\ \alpha_1\le\cdots\le\alpha_m \\ \beta_1\le\cdots\le\beta_n}} \bigwedge \left(\begin{array}{l} \bigwedge_{1\le i\le m} |\alpha_i I_m - F| = 0 \\ \forall_{\underline{x}} |x I_m - F| = 0 \Longrightarrow (x=\alpha_1 \vee \cdots \vee x = \alpha_m) \\ \hdashline \bigwedge_{1\le i\le n} \left|\beta_i I_n - F\right| = 0 \\ \forall_{\underline{x}} |x I_n - G| = 0 \Longrightarrow \left(x=\beta_1 \vee \cdots \vee x = \beta_n\right) \\ \hdashline \bigwedge_{1\le j\le c_0} \beta_j < \alpha_1 \\ \bigwedge_{c_0+1\le j\le c_0+c_1} \alpha_1 < \beta_j < \alpha_2 \\ \vdots \\ \bigwedge_{c_0+\cdots+c_{m-1}+1\le j\le c_0+\cdots+c_{m-1}+c_m} \alpha_m < \beta_j \end{array} \right) \tag{2.1}$$

where c_0 is a short hand for $n-c_1-\cdots-c_m$. The top block states that all the eigenvalues of F are exactly $\alpha_1,\ldots,\alpha_m$. The middle block states that all the eigenvalues of G are exactly $\beta_1,\ldots,\beta_n$. The bottom block states that the eigenvalue configuration is given by c, and that F and G are generic.

Obviously, the above is a condition on the parameters of F and G so that their eigenvalues are arranged in the given way. However, it is not "simple" in the sense that the expression involves quantified variables α, β, and x.

Note that the whole expression (2.1) is a quantified Boolean combination of polynomial equations and inequalities; in other words, it is a well-formed formula in the first-order theory of real closed fields. Alfred Tarski's celebrated theorem (Theorem 3) on real closed fields states that every first-order formula in the theory of real closed fields is equivalent to a quantifier-*free* formula [30]. Thus, by "simple condition", we mean a quantifier-*free* formula in the parameters of F and G; concretely, this means an expression which can be readily rewritten as a Boolean combination of equalities and inequalities of polynomials in the parameters of F and G.

Another important consideration in what constitutes a simple condition is the size and complexity of the output. In particular, Tarski's theorem guarantees that any quantified formula over the real numbers can be written as a quantifier-free formula in the language of the reals; precisely, this means a logical condition involving only real variables, polynomial ring operations, equality and inequality symbols, and logical conjunction and disjunction. However, such expressions can be unwieldy and difficult to understand, particularly when produced by general quantifier elimination algorithms. To mitigate this, we can introduce certain higher-level language constructs to improve the readability and reduce the size of the output condition. In this thesis, we will provide output conditions which allow the higher-level matrix

operations of multiplication, inverse, signature, and sign variation count. These high-level operations significantly reduce the complexity of the output, and we allow them because they can be trivially expanded into a quantifier-free Boolean combination of polynomial equalities and inequalities in the entries of the matrices.

Thus, the challenge is to produce a quantifier-free formula for all m and n efficiently and in a structured form using only real numbers, equality, matrix arithmetic, sign variation count, and matrix signature.

Remark 25. *Note that in Problem 24 there is an implicit condition on p so that F and G are a generic pair. As a result, the problem can be written in as finding a quantifier-free condition P such that*

$$\bigforall_{p:F,G \text{ generic}} \big(c = \mathrm{GEC}(F,G) \iff P(c,p)\big).$$

2.2 Signature Approach

In this section, we will solve the generic eigenvalue configuration problem using an approach based on the signature of matrices.

2.2.1 Main Theorem

In this subsection, we will state the main theorem for the signature approach. For this, we introduce the two central notions used in the main result. The first is a *combinatorial* object which depends only on the size of F, and the second is an *algebraic* object constructed from the parameters of F and G.

Definition 26 (Combinatorial part, signature approach)**.** *The matrix $C_{\text{sig}} \in \mathbb{Q}^{m \times 2^m}$ is defined as*

$$C_{\text{sig}} = V_m H_m^{-1}$$

where

1. *H_m is the $2^m \times 2^m$ Hadamard matrix whose rows are indexed by $e \in \{0,1\}^m$ and whose columns are indexed by $s \in \{-,+\}^m$ where*

   $$(H_m)_{e,s} = s_1^{e_1} \cdots s_m^{e_m},$$

   where $-$ is treated as -1 and $+$ treated as 1 in the calculation.

   We will drop the subscript and call this matrix H when the context is clear.

Equivalently, one could define H_m as

$$H_m = \underbrace{H_1 \otimes \cdots \otimes H_1}_{r},$$

where $H_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Note that this is a special case of the $\overline{H}$ matrix which appears in Theorem 11, where the rows and columns corresponding to the zero entries are removed.

2. *$V_m \in \mathbb{Z}^{m \times 2^m}$ is the matrix whose rows are indexed by $t \in [1, \ldots, m]$ and whose columns are lexicographically indexed by $s \in \{-,+\}^m$ where*

$$(V_m)_{t,s} = \delta_{v(s,+),m-t}$$

where δ denotes the Kronecker delta function, and $v(s,+)$ is the sign variation count of the sequence obtained by appending a $+$ to s; that is, the number of times two consecutive elements of $(s,+)$ are opposite sign, ignoring zeros.

We will drop the subscript and call this matrix V when the context is clear.

Example 27. *Let $m = 2$. We will produce the matrix C_{sig}.*

1. *First, we construct H_2.*

$$H_2 = \left[\begin{array}{c|cccc} e \backslash s & -- & -+ & +- & ++ \\ \hline 00 & 1 & 1 & 1 & 1 \\ 01 & -1 & 1 & -1 & 1 \\ 10 & -1 & -1 & 1 & 1 \\ 11 & 1 & -1 & -1 & 1 \end{array}\right]$$

Explanations on a few entries:

$$(H_2)_{01,--} = (-1)^0(-1)^1 = -1$$
$$(H_2)_{00,-+} = (-1)^0(1)^0 = 1$$
$$(H_2)_{11,+-} = (1)^1(-1)^1 = -1.$$

2. *Then, we construct V_2.*

$$V_2 = \left[\begin{array}{c|cccc} t \backslash s & -- & -+ & +- & ++ \\ \hline 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

Explanations on a few entries:

$$V_{1,--} = \delta_{\nu(--+),2-1} = \delta_{1,1} = 1$$
$$V_{2,--} = \delta_{\nu(--+),2-2} = \delta_{1,0} = 0$$
$$V_{2,++} = \delta_{\nu(+++),2-2} = \delta_{0,0} = 1.$$

Putting these together, we get that

$$\begin{aligned} C_{\text{sig}} &= V_2 H_2^{-1} \\ &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}. \end{aligned}$$

△

Definition 28 (Algebraic part, signature approach)**.** *A_{sig}, or $A_{\text{sig}}(F,G)$ to emphasize the dependence on the entries of F and G, is the column vector of length 2^m whose entries are indexed by $e \in \{0,1\}^m$ where*

$$\left(A_{\text{sig}}\right)_e = \sigma(f_e(G))$$

where

- *σ denotes the signature of a matrix (i.e., the difference between the number of positive eigenvalues and number of negative eigenvalues)*
- *f is the characteristic polynomial of F*
- *f_e is the polynomial*

$$f_e = \left(f^{(0)}\right)^{e_0} \cdots \left(f^{(m-1)}\right)^{e_{m-1}},$$

with $f^{(k)}$ denoting the k-th derivative of f.

Example 29. *Let $m = 2$ and $n = 3$. Then let F and G be matrices where each entry is an*

independent parameter; that is,

$$F = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{1,2} & a_{2,2} \end{bmatrix}, \qquad G = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{1,2} & b_{2,2} & b_{2,3} \\ b_{1,3} & b_{2,3} & b_{3,3} \end{bmatrix}.$$

We will compute one of the entries of $A_{\text{sig}}(F,G)$. By Definition 28 we have that

$$A_{\text{sig}} = \begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)). \end{bmatrix}$$

As an example, consider $e = (0,1)$. We will compute $f_{01}(G)$. We have

$$\begin{aligned} f &= \det(x I_2 - F) = x^2 + (-a_{2,2} - a_{1,1})x + a_{2,2}a_{1,1} - a_{1,2}^2 \\ f_{(0,1)} &= f^{(0)^0} f^{(1)^1} \\ &= 2x - a_{2,2} - a_{1,1}. \end{aligned}$$

Next, we compute

$$\begin{aligned} f_{(0,1)}(G) &= 2G + (-a_{2,2} - a_{1,1}) I_3 \\ &= 2 \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{1,2} & b_{2,2} & b_{2,3} \\ b_{1,3} & b_{2,3} & b_{3,3} \end{bmatrix} + (-a_{2,2} - a_{1,1}) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -a_{2,2} - a_{1,1} + 2b_{1,1} & 2b_{1,2} & 2b_{1,3} \\ 2b_{1,2} & -a_{2,2} - a_{1,1} + 2b_{2,2} & 2b_{2,3} \\ 2b_{1,3} & 2b_{2,3} & -a_{2,2} - a_{1,1} + 2b_{3,3} \end{bmatrix} \end{aligned}$$

Finally, we have that

$$\sigma(f_{01}(G)) = \#\textit{positive eigenvalues of } f_{01}(G) - \#\textit{negative eigenvalues of } f_{01}(G),$$

which clearly depends on the particular values of the parameters of F and G. △ △

Now, we are ready to state the main theorem which relates these two notions.

Theorem 30 (Signature approach for generic pairs of matrices)**.** *Let $F \in \mathbb{R}^{m \times m}$ and $G \in \mathbb{R}^{n \times n}$ be*

a pair of generic real symmetric matrices. We have

$$\mathrm{GEC}(F,G) \;=\; C_{\mathrm{sig}}\; A_{\mathrm{sig}}(F,G).$$

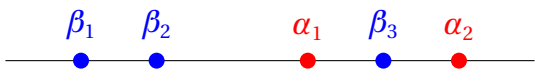
Remark 31. *Note that the entries of the symmetric matrices $f_e(G)$, whose signatures make up the vector $A_{\mathrm{sig}}(F,G)$, are polynomials in the parameters a_{ij} of F and b_{ij} of G. Hence, their signatures can easily be written as conditions on polynomials in the parameters of F and G since Descartes' rule of signs on the characteristic polynomial gives exact counts of the numbers of positive and negative eigenvalues. Thus the condition is quantifier-free.*

Remark 32. *The reader may notice that the main theorem in this section is stated almost identically to Theorem 66 which appears in the next section. However, they are completely different results, as C and A are defined as completely different concepts for each approach. The main theorem in this section is based on the signature of matrices while the main theorem in the next section is based on symmetric polynomials. However, in both theorems, there is a combinatorial part and an algebraic part; hence, we structure the theorem statements similarly to emphasize the super-structural similarities and the conceptual differences.*

Example 33 (General case with parametric matrices)**.** *Let $m=2$ and $n=3$ and let F and G be as in Example 29; that is,*

$$F=\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \qquad G=\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{bmatrix},$$

where each a_{ij} and b_{ij} is a fully independent parameter. We will now use Theorem 30 to write a condition on the parameters of F and G which is equivalent to the eigenvalues of F and G being arranged as in the diagram below.



That is, we will find a condition for $\mathrm{GEC}(F,G)=\begin{bmatrix}1\\0\end{bmatrix}$. *By Theorem 30, we have that*

$$\mathrm{GEC}(F,G)=\begin{bmatrix}1\\0\end{bmatrix} \iff \begin{bmatrix}1\\0\end{bmatrix}= C_{\mathrm{sig}}\; A_{\mathrm{sig}}(F,G).$$

In Example 27, we found that

$$C_{\text{sig}} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix},$$

and from Example 29 we have

$$A_{\text{sig}}(F,G) = \begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)) \end{bmatrix}.$$

Thus by Theorem 30 we have

$$\text{GEC}(F,G) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)) \end{bmatrix}.$$

With the fact that $-m \leq \sigma(f_e(G)) \leq m$, one can brute force the solutions to the above linear system and find that

$$\text{GEC}(F,G) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\Updownarrow$$

$$\begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)) \end{bmatrix} \in \left\{ \begin{bmatrix} 0 \\ 0 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\}.$$

Note that all the work done in this example used only m and did not depend on n; thus, the above set of possible signature vectors could be pre-computed for fixed m and used to construct a condition for any *$m \times m$ matrix F and* any *$n \times n$ matrix G (for arbitrary n) having the same eigenvalue configuration.* △

Example 34 (Matrices with only one parameter)**.** *In Example 33, we studied F and G which were symmetric matrices, each of whose elements was an independent parameter. However, it need not be the case that each entry is an independent parameter. We would, for instance, have the entries of both matrices be polynomials in some smaller list of parameters. Consider, for*

example, the matrices

$$F = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$$

and

$$G = \begin{bmatrix} 1 & p^2+p & p+1 \\ p^2+p & -p & p-1 \\ p+1 & p-1 & p \end{bmatrix},$$

where p is a (single) parameter. Suppose we wish to know for which values of p the eigenvalues of the matrix G interlace with that of F; that is, when $\mathrm{GEC}(F,G) = (1,1)$. *By Theorem 30 and reusing the computation from Example 27, we have that*

$$\begin{aligned} \mathrm{GEC}(F,G) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} &\iff \begin{bmatrix} 1 \\ 1 \end{bmatrix} = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G) \\ &\iff \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)) \end{bmatrix}, \end{aligned} \tag{2.2}$$

where

$$\sigma(f_{00}(G)) = \sigma\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

$$\sigma(f_{01}(G)) = \sigma\left(\begin{bmatrix} 0 & 2p^2+2p & 2p+2 \\ 2p^2+2p & -2-2p & 2p-2 \\ 2p+2 & 2p-2 & 2p-2 \end{bmatrix}\right)$$

$$\sigma(f_{10}(G)) = \sigma\left(\begin{bmatrix} p^4+2p^3+2p^2+2p-3 & -p^3-p^2-p-1 & (p-1)(p+1)^2 \\ -p^3-p^2-p-1 & p^4+2p^3+3p^2-2 & p^3+2p^2-p+2 \\ (p-1)(p+1)^2 & p^3+2p^2-p+2 & 3p^2-2p-1 \end{bmatrix}\right)$$

$$\sigma(f_{11}(G)) = \sigma\left(\begin{bmatrix} -2p^5-2p^4-4p^2-6p-2 & 2p^6+6p^5+12p^4+12p^3-2p^2-2p+4 & 2(p^3+p^2+p-1)(p+1)^2 \\ 2p^6+6p^5+12p^4+12p^3-2p^2-2p+4 & -4p^5-8p^4-12p^3-16p^2+8p & 2p^5+2p^4-16p^2-2p-2 \\ 2(p^3+p^2+p-1)(p+1)^2 & 2p^5+2p^4-16p^2-2p-2 & 4(p-1)(p^3+4p^2+1) \end{bmatrix}\right).$$

Taking into account the fact that $-m \le \sigma(f_e(G)) \le m$, one can find integer vector solutions to

the linear system by brute force and find that (2.2) is equivalent to the condition

$$\begin{bmatrix}\sigma(f_{00}(G))\\ \sigma(f_{01}(G))\\ \sigma(f_{10}(G))\\ \sigma(f_{11}(G))\end{bmatrix} \in \left\{ \begin{bmatrix}1\\2\\-1\\2\end{bmatrix}, \begin{bmatrix}2\\0\\0\\2\end{bmatrix}, \begin{bmatrix}2\\1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\2\\0\\0\end{bmatrix} \right\}.$$

Note that the above is a condition which depends only on the parameter p and not on the eigenvalues of F and G. As stated previously, the signatures of each (real symmetric) matrix $f_e(G)$ can be expanded using Descartes' rule of signs to express the positive and negative eigenvalues of $f_e(G)$ in terms of the sign variation count of the respective characteristic polynomials.

For example, consider

$$f_{01}(G) = \begin{bmatrix} 0 & 2p^2+2p & 2p+2 \\ 2p^2+2p & -2-2p & 2p-2 \\ 2p+2 & 2p-2 & 2p-2 \end{bmatrix}.$$

We have

$$\det(xI_2 - f_{01}(G)) = x^3 + 4x^2 - (4p^4+8p^3+16p^2+4)x + 8p^5 - 8p^4 - 32p^3 - 16p^2 - 8p - 8.$$

Since $f_e(G)$ is real symmetric, its characteristic polynomial has all real roots; thus by Descartes' rule of signs, the number of positive eigenvalues of $f_{01}(G)$ equals

$$v(1, 4, -(4p^4+8p^3+16p^2+4), 8p^5 - 8p^4 - 32p^3 - 16p^2 - 8p - 8)$$

and the number of negative eigenvalues equals

$$v(-1, 4, (4p^4+8p^3+16p^2+4), 8p^5 - 8p^4 - 32p^3 - 16p^2 - 8p - 8).$$

Therefore

$$\begin{aligned}\sigma(f_{01}(G)) = \;& v(1, 4, -(4p^4+8p^3+16p^2+4), 8p^5 - 8p^4 - 32p^3 - 16p^2 - 8p - 8) \\ & - v(-1, 4, (4p^4+8p^3+16p^2+4), 8p^5 - 8p^4 - 32p^3 - 16p^2 - 8p - 8).\end{aligned}$$

This expression can then be expanded to explicitly enumerate the possible signs for each of the polynomials, and the process can be repeated for $\sigma(f_{00}(G)), \sigma(f_{10}(G))$, and $\sigma(f_{11}(G))$ to obtain the full condition involving only p. △

In algorithm form, Theorem 30 can be stated as follows.

Algorithm 35 (Condition for EC: signature approach)**.**

In: $F \in \mathbb{R}[p]^{m\times m}$ *and* $G \in \mathbb{R}[p]^{n\times n}$*, symmetric, where* p *is a finite list of parameters*
$c \in \mathbb{N}^m$*, eigenvalue configuration*

Out: *P, quantifier-free condition on* p *such that* $c = \mathrm{GEC}(F,G)$

1. $f \longleftarrow \det(x I_m - F)$
2. *For* $e \in \{0,1\}^m$ *do*
$$f_e \leftarrow \left(f^{(0)}\right)^{e_0} \cdots \left(f^{(m-1)}\right)^{e_{m-1}}$$
3. $A_{\mathrm{sig}} \leftarrow$ *vector such that* $(A_{\mathrm{sig}})_e = \sigma(f_e(G))$ *for* $e \in \{0,1\}^m$
4. $V \leftarrow$ *matrix such that* $V_{t,s} = \delta_{v(s,+)=m-t}$ *for* $t \in [m]$ *and* $s \in \{-,+\}^m$
5. $H \leftarrow$ *matrix such that* $H_{e,s} = s_1^{e_1} \cdots s_m^{e_m}$ *for* $e \in \{0,1\}^m$ *and* $s \in \{-,+\}^m$
6. $C_{\mathrm{sig}} \leftarrow V H^{-1}$
7. $P \longleftarrow \bigvee_{S : c = C_{\mathrm{sig}}\, S} A_{\mathrm{sig}} = S$
8. *Return P*

2.2.2 Proof

In this section, we will provide a proof of Theorem 30 (Signature approach for generic matrices). We will do this in a series of lemmas that explain the motivation and reasoning behind each step. We divide the proof into two stages.

1. First, we prove a special case of Theorem 30 in the case where F and G are "strongly generic;" that is, the pair satisfies an additional condition (see below) in addition to being generic.

2. Next, we prove a slightly more general case of Theorem 30 in the case where the pair F and G is fully generic.

Definition 36 (Strongly generic)**.** *We say that the pair of real symmetric matrices* F *and* G *is* ***strongly generic*** *if both of the following hold.*

- *The pair F and G is generic, i.e. F and G do not share eigenvalues.*
- *Let $f(x)$ denote the characteristic polynomial of F. For each eigenvalue β of G and for all $k \in \{1, \ldots, m-1\}$, we have*

$$f^{(k)}(\beta) \neq 0,$$

where $f^{(k)}$ denotes the k-th derivative of f.

We will illustrate the above definition using two examples: one for strongly generic and the other for generic but not strongly generic. These will be used as running examples in the following subsections.

Example 37 (Strongly generic)**.** *We will be using this example as a running example throughout the section.*

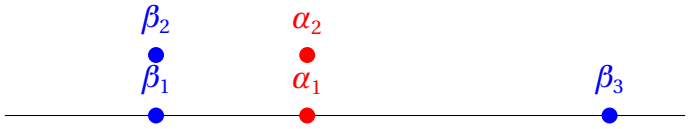
Let

$$F = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 2} \qquad G = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Then their respective eigenvalues are

$$\alpha = (4,4) \qquad \beta = (2,2,8),$$

as shown in the diagram below.



So,

$$\mathrm{GEC}(F,G) = (c_1, c_2) = (0,1).$$

Further,

$$f = \det(xI - F) = (x-4)^2$$
$$f^{(1)} = 2(x-4).$$

Note that F and G are generic because they do not share any eigenvalues. Furthermore, the derivative $f^{(1)}$ is nonzero on all eigenvalues of G. Thus F and G are strongly generic. △

Example 38 (Not strongly generic)**.** *We will be using this example as a running example throughout the section.*

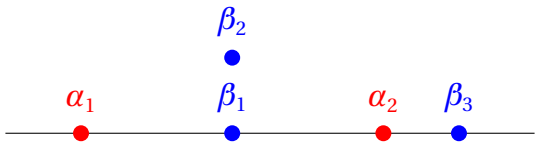
Let

$$F = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad G = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \in \mathbb{R}^{3\times 3},$$

and so

$$\alpha = (1,5), \qquad \beta = (3,3,6).$$

Hence their eigenvalue configuration is shown in the figure below.



and so

$$\mathrm{GEC}(F,G) = (2,1).$$

Note that F and G are generic because they do not share any eigenvalues. However, note that

$$f = \det(xI - F) = (x-1)(x-5)$$
$$f^{(1)} = 2(x-3)$$

Then we have

$$f^{(1)}(\beta_1) = f^{(1)}(\beta_2) = 0,$$

and so F and G are not strongly generic. △

Proof for strongly generic pair of matrices

We first prove a special case of Theorem 30 for strongly generic F and G.

Lemma 39. *Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be a real symmetric **strongly generic** pair of matrices. Then*

$$\mathrm{GEC}(F,G) \;=\; C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G).$$

Our process of proving Lemma 39 will follow the diagram below. In the lemmas in this

section, we repeatedly rewrite the statement $c = \mathrm{GEC}(F,G)$ with the goal of eliminating all references to α and β, so that we end with an expression involving only the entries of F and G.

$c = \mathrm{GEC}(F,G)$	
$\Vert$ by Definition 22	
$c_t = \#\{j : \beta_j \in A_t\}$	(Expression involving α and β)
$\Vert$ Lemma 40	
$c_t = \#\{j : \overline{\#}\{x : f(x) = 0 \wedge x > \beta_j\} = m - t\}$	(Expression involving β)
$\Vert$ Lemmas 43, 45, 47	
$c = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G)$	(Expression involving entries and signature)

As the first step, we will now find an expression for $\mathrm{GEC}(F,G)$ which does not refer to α.

Lemma 40 (Eliminate α)**.** *Let $c = \mathrm{GEC}(F,G)$. Then for all $t = 1,\ldots,m$ we have*

$$c_t = \#\left\{j : \overline{\#}\left\{x : f(x) = 0 \ \wedge\ x > \beta_j\right\} = m - t\right\}$$

where $f(x) = \det(x I_m - F)$ is the characteristic polynomial of F, and the symbol $\overline{\#}$ means to count with multiplicity.

Example 41 (Running)**.** *Recall that in the running example (Example 37) we have $(c_1, c_2) = (0,1)$. Recall the eigenvalue configuration depicted in Example 37. On the other hand,*

$$\begin{aligned} \#\{j : \overline{\#}\{x : f(x) = 0 \wedge x > \beta_j\} = 2-1\} &= \#\emptyset = 0 = c_1 \\ \#\{j : \overline{\#}\{x : f(x) = 0 \wedge x > \beta_j\} = 2-2\} &= \#\{3\} = 1 = c_2. \end{aligned}$$

△

Proof of Lemma 40. Recall the eigenvalue configuration condition (Definition 22):

$$c_t = \#\{j : \beta_j \in A_t\}. \tag{2.3}$$

We will rewrite the expression on the right-hand side, with the goal of eventually reducing the problem to a set of real root counting problems. First, we will repeatedly rewrite the condition $\beta_j \in A_t$ with the goal of replacing the *membership* checking of β_j with the *counting* of α_i.

$$
\begin{aligned}
&\beta_j \in A_t \\
&\quad \Updownarrow \text{ from the definition of } A_t \\
&\alpha_t < \beta_j < \alpha_{t+1} \\
&\quad \Updownarrow \text{ from the assumption that } \alpha_1 \le \alpha_2 \le \cdots \le \alpha_m \\
&\alpha_1 \le \cdots \le \alpha_t < \beta_j < \alpha_{t+1} \le \cdots \le \alpha_m \\
&\quad \Updownarrow \text{ from counting } i \text{ such that } \alpha_i \text{ is greater than } \beta_j \\
&\#\left\{ i : \alpha_i > \beta_j \right\} = m - t .
\end{aligned}
$$

Note that the counting of *indices* of the eigenvalues α effectively counts the eigenvalues with multiplicity.

By replacing $\beta_j \in A_t$ with $\#\left\{ i : \alpha_i > \beta_j \right\} = m - t$ in the right-hand side of (2.3), we arrive at

$$
c_t = \#\left\{ j : \#\left\{ i : \alpha_i > \beta_j \right\} = m - t \right\}. \tag{2.4}
$$

We will now rewrite the right-hand side of the above with the goal of eliminating α. Let $f(x)$ be the characteristic polynomial of F, that is,

$$
f(x) = \det(x I_m - F).
$$

Thus we have

$$
\#\left\{ i : \alpha_i > \beta_j \right\} = \overline{\#}\left\{ x : f(x) = 0 \ \wedge \ x > \beta_j \right\},
$$

where $\overline{\#}$ counts with multiplicity.

Finally, the claim of the lemma follows immediately by replacing $\#\left\{ i : \alpha_i > \beta_j \right\}$ with

$$
\overline{\#}\left\{ x : f(x) = 0 \ \wedge \ x > \beta_j \right\}
$$

in the right-hand side of (2.4). □

At this point, we now have an expression for each component of GEC(F,G) which does not contain any α's. The next step is to eliminate the β's. This is somewhat more complicated than eliminating the α's, so we split the next stage into several steps.

1. In Lemma 43, we will first eliminate the counting of roots of f; i.e., we find another way to express the quantity $\overline{\#}\{x : f(x) = 0 \wedge x > \beta_j\}$. This will be achieved by Descartes' rule of signs. With this, we will have rewritten each of the m entries of GEC(F,G) as a root counting problem which counts eigenvalues of G subject to a sign condition on the derivatives of f.

2. In Lemma 45, we will then rewrite in matrix form. This is a necessary prerequisite for the next step, in which we solve all m root counting problems together.

3. Finally, in Lemma 47, we will solve the root counting problems by rewriting them in terms of the signature (i.e., the difference between the number of positive and negative eigenvalues) of matrices derived from F and G.

We begin by eliminating the reference to counting the roots of f. First, we need a notation.

Notation 42. *The sign sequence of* $f = \det(xI_m - F)$ *evaluated at* $x \in \mathbb{R}$*, denoted* sseq(x)*, is*

$$\text{sseq}(x) = \text{sign}\left(f^{(0)}(x), \ldots, f^{(m-1)}(x)\right).$$

Note that we only go to the $(m-1)$*th derivative, since by construction the* m*-th derivative is always a positive constant; hence we can omit it for simpler presentation.*

Lemma 43 (Eliminate counting roots of f)**.** *Let* $c = \text{GEC}(F,G)$*. For all* $t = 1, \ldots, m$*, we have*

$$c_t = \#\left\{j : v\left(\text{sseq}(\beta_j), +\right) = m - t\right\}.$$

Example 44 (Running)**.** *Recall that in the running example (Example 37) we have*

$$\begin{array}{lcl} (c_1, c_2) & = & (0,1) \\ \alpha & = & (4,4) \\ \beta & = & (2,2,8) \\ f & = & \det(xI - F) = (x-4)^2 \\ f^{(1)} & = & 2(x-4). \end{array}$$

On the other hand,

$$v(\text{sseq}(\beta_1), +) = v(\text{sseq}(2), +) = v\left(\text{sign}\left(f(2), f^{(1)}(2)\right), +\right) = v(+,-,+) = 2$$

$$\begin{aligned} v(\text{sseq}(\beta_2),+) &= v(\text{sseq}(2),+) = v\left(\text{sign}\left(f(3), f^{(1)}(3)\right),+\right) = v(+,-,+) = 2 \\ v(\text{sseq}(\beta_3),+) &= v(\text{sseq}(8),+) = v\left(\text{sign}\left(f(6), f^{(1)}(6)\right),+\right) = v(+,+,+) = 0. \end{aligned}$$

In summary, we have:

j	$v(\text{sseq}(\beta_j),+)$
1	2
2	2
3	0

So then

$$\begin{aligned} \#\{j : v(\text{sseq}(\beta_j)) = 2-1\} &= \#\emptyset = 0 = c_1 \\ \#\{j : v(\text{sseq}(\beta_j)) = 2-2\} &= \#\{3\} = 1 = c_2 \end{aligned}$$

△

Proof of Lemma 43. Recall Lemma 40:

$$c_t = \#\left\{j : \overline{\#}\left\{x : f(x) = 0 \ \wedge \ x > \beta_j\right\} = m - t\right\}. \tag{2.5}$$

We will eliminate the inner count symbol $\overline{\#}$ from the right-hand side. For this, we will crucially use the corollary to Descartes' rule of signs (Corollary 7), which allows us to exactly count the positive roots. Note

$$\begin{array}{l} \overline{\#}\{x \mid f(x) = 0 \wedge x > \beta_j\} \\ \quad \Big\| \ \text{by introducing } r = x - \beta_j \\ \overline{\#}\{r \mid f(r+\beta_j) = 0 \wedge r > 0\} \\ \quad \Big\| \ \text{by introducing } p(r) = f\left(r+\beta_j\right) \\ \overline{\#}\{r \mid p(r) = 0 \wedge r > 0\} \\ \quad \Big\| \ \text{from Descartes' rule of signs (see a remark below for a detailed reasoning)} \\ v\left(\text{sign}\left(p_0, p_1, \ldots, p_m\right)\right) \qquad \text{where } p = p_m r^m + \cdots + p_0 r^0 \end{array}$$

$$
\begin{aligned}
&\Big\| \quad \text{since } p_j = \frac{p^{(j)}(0)}{j!} \\
&v\left(\operatorname{sign}\left(p^{(0)}(0), p^{(1)}(0), \ldots, p^{(m)}(0)\right)\right) \\
&\Big\| \quad \text{since } p^{(k)}(0) = f^{(k)}(\beta_j) \\
&v\left(\operatorname{sign}\left(f^{(0)}(\beta_j), f^{(1)}(\beta_j), \ldots, f^{(m)}(\beta_j)\right)\right) \\
&\Big\| \quad \text{since } f^{(m)}(\beta_j) > 0 \\
&v\left(\operatorname{sign}\left(f^{(0)}(\beta_j), f^{(1)}(\beta_j), \ldots, f^{(m-1)}(\beta_j)\right), +\right) \\
&\Big\| \quad \text{from definition of sseq} \\
&v\left(\operatorname{sseq}(\beta_j), +\right).
\end{aligned}
$$

Two remarks on the above rewriting steps:

- The third rewriting is based on the following detailed reasoning: (1) Since the matrix F is real-symmetric, all the roots of its characteristic polynomial f are real. (2) In turn, all the roots of the related polynomial p are also real. (3) Descartes' rule of signs is exact when all the roots are real. (4) Thus the number of positive real roots of p, counting multiplicity, is exactly the number of sign variation in the coefficients of p.
- If one uses the Budan-Fourier extension [36, 37] of Descartes' rule of signs, then one could skip a few steps in the above rewriting.

Finally the claim of the lemma follows immediately by replacing the expression

$$\overline{\#}\left\{x : f(x) = 0 \ \wedge \ x > \beta_j\right\}$$

with the expression $v\left(\operatorname{sseq}(\beta_j), +\right)$, in the right-hand side of (2.5). □

At this point, we have written each component of GEC(F, G) as a root counting problem counting eigenvalues of G:

$$c_t = \#\left\{j : v\left(\operatorname{sseq}(\beta_j), +\right) = m - t\right\}.$$

Next, we will rewrite these m root counting problems into matrix form.

Lemma 45 (Rewrite in matrix form)**.** *We have*

$$\mathrm{GEC}(F,G) = V\,q$$

where

- *V is the matrix whose rows are indexed by $t \in [1,\dots,m]$ and the columns are indexed lexicographically by $s \in \{-,+\}^m$ with $V_{t,s} = \delta_{v(s,+),m-t}$.*
- *q is the column vector whose rows are indexed lexicographically by $s \in \{-,+\}^m$ with*

$$q_s = \#\left\{ j : \mathrm{sseq}(\beta_j) = s \right\}$$

Example 46 (Running)**.** *Recall the running example (Example 37), where*

$$c = (0,1).$$

From the definition of V in the statement of Lemma 45, we have

$$V = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

and q is the vector

$$q = \begin{bmatrix} q_{--} \\ q_{-+} \\ q_{+-} \\ q_{++} \end{bmatrix} = \begin{bmatrix} \#\{j : \mathrm{sseq}(\beta_j) = (--)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (-+)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (+-)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (++)\} \end{bmatrix}.$$

From the previous installment in the running example (Example 44), we calculated that

$$\begin{aligned} \mathrm{sseq}(\beta_1) &= (+-) \\ \mathrm{sseq}(\beta_2) &= (+-) \\ \mathrm{sseq}(\beta_3) &= (++). \end{aligned}$$

Hence

$$q = \begin{bmatrix} \#\{j : \mathrm{sseq}(\beta_j) = (--)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (-+)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (+-)\} \\ \#\{j : \mathrm{sseq}(\beta_j) = (++)\} \end{bmatrix} = \begin{bmatrix} \#\emptyset \\ \#\emptyset \\ \#\{1,2\} \\ \#\{3\} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}.$$

Putting it together, we get

$$V q = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

△

Proof of Lemma 45. Recall Lemma 43:

$$c_t = \#\left\{ j : v\left(\operatorname{sseq}(\beta_j), +\right) = m - t \right\}. \tag{2.6}$$

We will eliminate the sign variation count symbol from v from the right-hand side. Let J_t denote the set whose entries are counted on the right-hand side of (2.6); that is,

$$J_t = \left\{ j : \ v\left(\operatorname{sseq}(\beta_j), +\right) = m - t \right\}.$$

We partition the set J_t according to the sign vector, obtaining

$$J_t = \biguplus_{\substack{s \in \{-,+\}^m \\ v(s,+) = m-t}} J_s \quad \text{where } J_s = \left\{ j : \ \operatorname{sseq}(\beta_j) = s \right\}.$$

Note that this is indeed a disjoint union, since by the generic and strongly generic conditions we have $f^{(k)}(\beta_j) \neq 0$ for all $k = 0, \ldots, m-1$.

Then we immediately have

$$c_t = \#J_t = \sum_{\substack{s \in \{-,+\}^m \\ v(s,+) = m-t}} \#J_s$$

In matrix form, we can write this as

$$c = V q$$

where

- V is the matrix where rows are indexed by $t \in [1, \ldots, m]$ and the columns are indexed lexicographically by $s \in \{-,+\}^m$ with $V_{t,s} = \delta_{v(s+),m-t}$.
- q is the column vector where rows are indexed lexicographically by $s \in \{-,+\}^m$ with $q_s = \#J_s$.

Finally the claim of the lemma follows immediately by replacing the expression in the right-hand side of (2.6) with the expression $c = Vq$. □

Lemma 47 (Rewrite q in terms of signature)**.** *We have*

$$q = H^{-1} A_{\mathrm{sig}}(F, G)$$

where $A_{\mathrm{sig}}(F,G)$ is the vector defined in Definition 28; that is, the vector indexed by $e \in \{0,1\}^m$ where

$$(A_{\mathrm{sig}}(F,G))_e = \sigma(f_e(G))$$

where

$$f_e = \left(f^{(0)}\right)^{e_0} \cdots \left(f^{(m-1)}\right)^{e_{m-1}}.$$

Example 48 (Running)**.** *In Example 46, we computed*

$$q = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}.$$

From the definition of H, we have that

$$H^{-1} = \left[\begin{array}{c|cccc} e \backslash s & -- & -+ & +- & ++ \\ \hline 00 & 1 & 1 & 1 & 1 \\ 01 & -1 & 1 & -1 & 1 \\ 10 & -1 & -1 & 1 & 1 \\ 11 & 1 & -1 & -1 & 1 \end{array}\right]^{-1}$$

$$= \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

We then compute

$$f_{00} = \left(f^{(0)}\right)^0 \left(f^{(1)}\right)^0 = 1$$
$$f_{01} = \left(f^{(0)}\right)^0 \left(f^{(1)}\right)^1 = 2(x-4)$$
$$f_{10} = \left(f^{(0)}\right)^1 \left(f^{(1)}\right)^0 = (x-1)^2$$

$$f_{11} = \left(f^{(0)}\right)^1 \left(f^{(1)}\right)^1 = 2(x-4)^3.$$

Then

$$f_{00}(G) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$f_{01}(G) = \begin{bmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$
$$f_{10}(G) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 16 \end{bmatrix}$$
$$f_{11}(G) = \begin{bmatrix} -16 & 0 & 0 \\ 0 & -16 & 0 \\ 0 & 0 & 128 \end{bmatrix}.$$

So we have

e	Eigenvalues of $f_e(G)$	# positive eigenvalues	# negative eigenvalues
00	$(1,1,1)$	3	0
01	$(-4,-4,8)$	1	2
10	$(4,4,16)$	3	0
11	$(-16,-16,128)$	1	2

Therefore

$$A_{\text{sig}}(F,G) = \begin{bmatrix} \sigma(f_{00}(G)) \\ \sigma(f_{01}(G)) \\ \sigma(f_{10}(G)) \\ \sigma(f_{11}(G)) \end{bmatrix} = \begin{bmatrix} 3-0 \\ 1-2 \\ 3-0 \\ 1-2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 3 \\ -1 \end{bmatrix}.$$

Finally, we have that

$$H^{-1}\sigma = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} = q.$$

△

Proof of Lemma 47. Recall q from Lemma 45: q is the column vector where rows are indexed lexicographically by $s \in \{-,+\}^m$ with

$$q_s = \#\left\{ j : \operatorname{sseq}(\beta_j) = s \right\}$$

We will rewrite it in terms of the signatures of certain symmetric matrices constructed from F and G.

From applying a special case of Theorem 11, we have

$$Hq = A_{\text{sig}}(F,G)$$

where

- H is the $2^m \times 2^m$ Hadamard matrix whose rows are indexed by $e \in \{0,1\}^m$ and whose columns are indexed by $s \in \{-,+\}^m$ where

$$(H_m)_{e,s} = s_1^{e_1} \cdots s_m^{e_m}$$

- A_{sig} is defined as in Definition 28; that is, A_{sig} is the column vector whose rows are indexed lexicographically by $e = (e_0, \ldots, e_{m-1}) \in \{0,1\}^m$ with

$$(A_{\text{sig}})_e = \sigma(f_e(G))$$

where

$$f_e = f^{(0)e_0} \cdots f^{(m-1)e_{m-1}},$$

and again $\sigma(f_e(G))$ is the signature of the matrix $f_e(G)$.

Since H is a Hadamard matrix, it is invertible, so we therefore have

$$q = H^{-1} A_{\text{sig}}(F,G)$$

and we are done.

□

Remark 49. *Note that the key step (the application of Theorem 11) in the previous proof relied on the fact that $f_e(\beta_j) \neq 0$. This is always true for strongly generic F and G, but does not hold for F and G which are only generic. This is the main obstacle that we will overcome when we generalize to pairs of matrices which are not strongly generic.*

Finally, we have arrived at an expression for $\text{GEC}(F,G)$ which does not contain any references to the eigenvalues α or β. Hence, we are ready to prove the main result (Theorem 30) for strongly generic F and G, namely Lemma 39.

Proof of Lemma 39. Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be real symmetric **strongly generic** matrices. We need to prove the following:

$$\text{GEC}(F,G) \;=\; C_{\text{sig}}\, A_{\text{sig}}(F,G).$$

For this, we begin by recalling the following lemmas from above.

1. From Lemma 45, we have
$$\text{GEC}(F,G) = V\,q$$
where V is the matrix where rows are indexed by $t \in [1,\ldots,m]$ and the columns are indexed lexicographically by $s \in \{-,+\}^m$ with
$$V_{t,s} = \delta_{v(s),m-t}\,.$$

2. From Lemma 47, we have
$$q = H^{-1} A_{\text{sig}}(F,G)$$
where A_{sig} is a column vector whose rows are indexed lexicographically by $e \in \{0,1\}^m$.

Together, we finally have

$$\text{GEC}(F,G) \underset{1}{=} V\,q \underset{2}{=} V\left(H^{-1} A_{\text{sig}}(F,G)\right) \underset{\text{Def. 26}}{=} C_{\text{sig}}\, A_{\text{sig}}(F,G).$$

We have proved Lemma 39. □

Proof for generic pair of matrices

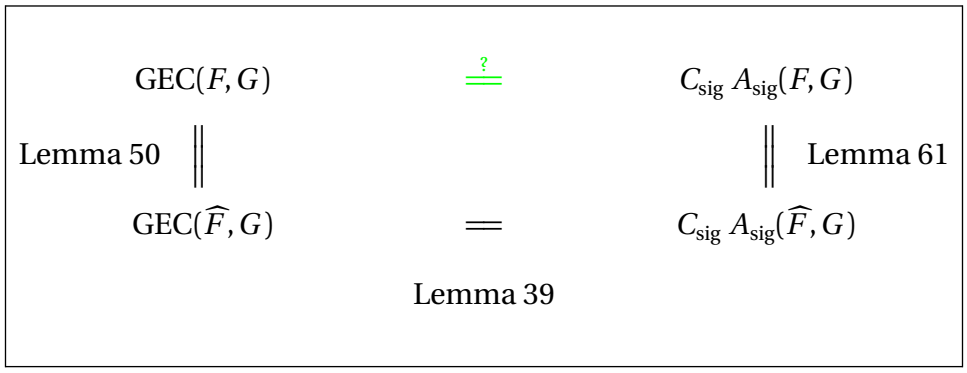
Now, we will extend the result on strongly generic matrices (Lemma 39) to prove Theorem 30 (Signature approach for generic pairs of matrices). Recall that Theorem 30 states that if F and G are a generic pair of real symmetric matrices, then $\mathrm{GEC}(F,G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G)$. The major difficulty extending Lemma 39 is that the version of the root counting method [32] used in Lemma 47 requires that $\mathrm{sseq}(\beta_j)$ contains no zeros; i.e., that F and G are strongly generic. To circumvent this difficulty, we take the following approach.

1. In Lemma 50, we will show that, given generic F and G, we can "safely" perturb F into a new real symmetric matrix $\widehat{F}$, so that $\widehat{F}$ and G are strongly generic and the eigenvalue configuration of F and G equals that of $\widehat{F}$ and G. Since $\widehat{F}$ and G are strongly generic, we can apply Lemma 39.

2. In Lemma 61, we will show that $C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F},G)$. However, before we can do that, we need to prove an intermediate result in Lemma 55, which generalizes Lemma 47 from the previous subsection.

   Once this is done, the proof will be complete.

Our strategy is illustrated by the following diagram.



First, we begin by showing that, given generic F and G, it is always possible to safely (i.e. without disturbing the eigenvalue configuration) shift the eigenvalues of F slightly to form a new symmetric matrix $\widehat{F}$ so that $\widehat{F}$ and G are strongly generic.

Lemma 50 (Safely perturb generic matrices to strongly generic matrices)**.** *Let F, G be generic matrices. Then there exists a real symmetric matrix $\widehat{F}$ so that*

- *$\widehat{F}$ and G are strongly generic and*

- $\mathrm{GEC}(F,G) = \mathrm{GEC}(\widehat{F},G)$.

Example 51 (Running)**.** *Recall the generic, but not strongly generic, matrices F and G from Example 38:*

$$F = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad G = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \in \mathbb{R}^{3\times 3},$$

and so

$$\mathrm{GEC}(F,G) = (2,1).$$

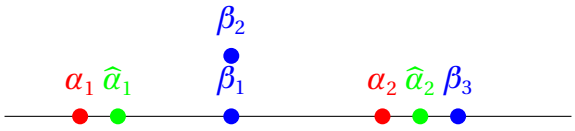
Now, let $\varepsilon = 0.5$ and set

$$\widehat{F} = F + \varepsilon I = \begin{bmatrix} 1+\varepsilon & 0 \\ 0 & 5+\varepsilon \end{bmatrix} = \begin{bmatrix} 1.5 & 0 \\ 0 & 5.5 \end{bmatrix}.$$

The derivatives of the characteristic polynomial $\widehat{f}$ of the new matrix $\widehat{F}$ are

$$\widehat{f}^{(0)}(x) = x^2 - 7x + 8.25$$
$$\widehat{f}^{(1)}(x) = 2x - 7.$$

The arrangement of the eigenvalues of F, G, and $\widehat{F}$ is shown in the diagram below.



Note that now $\widehat{F}$ and G are strongly generic because $\widehat{f}^{(0)}$ and $\widehat{f}^{(1)}$ are nonzero at β_1, β_2, and β_3. Further, because ε was chosen small enough, no eigenvalue of $\widehat{F}$ "crossed over" any β and so the eigenvalue configuration of $\widehat{F}$ and G remains the same as that of F and G. △

Proof of Lemma 50. If F and G are a pair of strongly generic matrices, then clearly $\widehat{F} = F$ satisfies the conclusion of the lemma; thus, for the remainder of this proof, assume without loss of generality that the pair F and G is generic but not strongly generic.

We will construct $\widehat{F}$ from F by shifting the eigenvalues of F by a positive real number ε. Let

$$\widehat{F} := F + \varepsilon I.$$

Obviously $\widehat{F}$ is symmetric and its eigenvalues are simply the eigenvalues of F plus ε. Now we

need to choose ε so that the claims of the lemma hold. Let us choose ε as follows.

$$\varepsilon := \frac{1}{2}\min(A, B, C)$$

where

$$\begin{array}{rcll} A & = & \min\{|\alpha_i - \alpha_{i+1}| & : \ i = 1, \ldots, m-1 \text{ and } \alpha_i \neq \alpha_{i+1}\} \\ B & = & \min\{|\alpha_i - \beta_j| & : \ i = 1, \ldots, m \text{ and } j = 1, \ldots, n\} \\ C & = & \min\{|\beta_j - \gamma| & : \ j = 1, \ldots, n \text{ and } f^{(k)}(\gamma) = 0 \text{ for some } 1 \leq k \leq m \text{ and } \gamma \neq \beta_j\} \end{array}$$

By construction, A, B, C, and therefore ε are positive real numbers. Let us now check that $\widehat{F}$ and G satisfy the two claims of the lemma, one by one.

- $\widehat{F}$ and G are strongly generic.

    1. $\widehat{F}$ and G do not share any eigenvalues, because by construction each eigenvalue of $\widehat{F}$ is, at minimum, a distance of $\frac{1}{2}\min(B) > 0$ from any eigenvalue of G.
    2. Each eigenvalue of G is, at minimum, a distance of $\varepsilon > 0$ from any root of any derivative of $\widehat{f} = \det(x I_m - \widehat{F})$.

- $\mathrm{GEC}(F, G) = \mathrm{GEC}(\widehat{F}, G)$.

    1. Since $\varepsilon < \min(A)$ and $\varepsilon < \min(B)$, no eigenvalue of $\widehat{F}$ crosses over an eigenvalue of F or G.
    2. Hence, the eigenvalue configuration of $\widehat{F}$ and G remains the same as the eigenvalue configuration of F and G.

□

Now that we have established that we can perturb generic F and G to get strongly generic $\widehat{F}$ and G without disturbing the eigenvalue configuration, we need to show that

$$C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F, G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F}, G).$$

To do this, we first need to extend some notions from the previous section. First, to aid in the computation, we need a key lemma which generalizes Lemma 47 from the previous section by finding a simple expression for the quantity $H^{-1} A_{\mathrm{sig}}(F, G)$. We first need to define the notion of the boundary and closure of a sign sequence.

Definition 52.

- *The boundary of $s \in \{-,+\}^m$, written as $\partial(s)$, is defined by*

$$\partial(s) = \{s' \in \{-,0,+\}^m : s' \text{ and } s \text{ differ by one or more zeros}\}.$$

- *The closure of $s \in \{-,+\}^m$, written as $\mathrm{cl}(s)$, is defined by*

$$\mathrm{cl}(s) = \{s\} \cup \partial s.$$

Example 53. *Consider the sign sequence $+-$. We have*

1. $\partial(+-) = \{+0, 0-, 00\}$
2. $\mathrm{cl}(+-) = \{+-, +0, 0-, 00\}$.

△

Remark 54.

- *The notion of the boundary of a sign sequence s can be thought of as the topological boundary of the set in $\mathbb{R}^m$ of points whose coordinates have the signs given by s.*
- *Note that the number of elements in $\mathrm{cl}(s)$ where $s \in \{-,+\}^m$ is 2^m, because we can choose zero or more entries in s to be zero.*

Lemma 55. *Let F, G be an arbitrary (not necessarily generic) pair of real symmetric matrices. Then*

$$H^{-1} A_{\mathrm{sig}}(F,G) = \sum_{j=1}^{n} \gamma_j$$

where

$$\gamma_j = \frac{1}{\#S_j} \sum_{s' \in S_j} e_{s'}$$

$$S_j = \left\{ s \in \{-,+\}^m : \mathrm{sseq}(\beta_j) \in \mathrm{cl}(s) \right\}$$

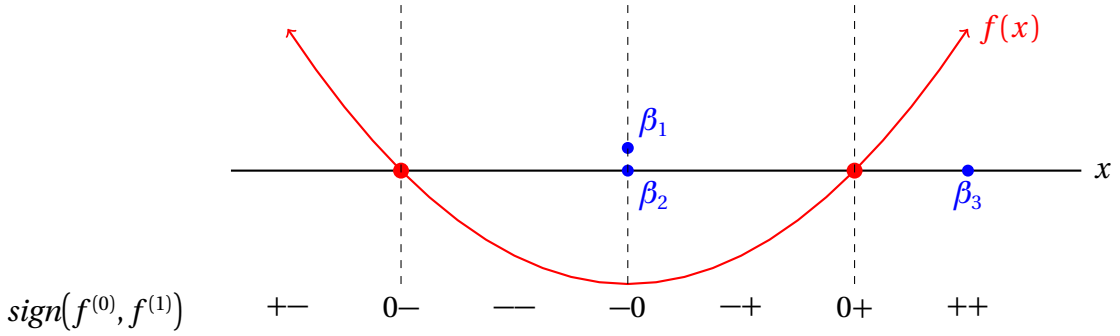
$e_{s'} =$ *standard unit vector with a 1 in the position of s' under lex order with $- < +$.*

Remark 56. *Note that the hypothesis of Lemma 55 does not require that F and G be generic. In this section, we assume that they are generic, but we will revisit this lemma later in the thesis when fully generalizing our result to arbitrary matrices.*

Example 57 (Running)**.** *Recall that in the running example we have*

$$\alpha = (1,5), \qquad \beta = (3,3,6).$$

Together with the characteristic polynomial of f, the eigenvalues of F and G are shown in the diagram below.



Lemma 55 should be seen as quantifying the contribution γ_j of each eigenvalue β_j to the eigenvalue configuration vector. To see this, we will now compute γ_j for each j in this example.

j	$\text{sseq}(\beta_j)$	S_j	γ_j
1	-0	$\{--,-+\}$	$\frac{1}{2}\left(\begin{bmatrix}1\\0\\0\\0\end{bmatrix}+\begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right)$
2	-0	$\{--,-+\}$	$\frac{1}{2}\left(\begin{bmatrix}1\\0\\0\\0\end{bmatrix}+\begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right)$
3	$++$	$\{++\}$	$\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

Observe that S_3 is a singleton set. This happens because $\text{sseq}(\beta_3)$ *does not contain zeros, because $f^{(k)}(\beta_3) \neq 0$ for all $k \in \{1,\ldots,m-1\}$. Hence, γ_3 is simply a unit vector with a 1 in the slot corresponding to the sign sequence of β_3.*

However, S_1 and S_2 are sets with two elements each. This is because $\text{sseq}(\beta_1)$ *and* $\text{sseq}(\beta_2)$ *each contain a zero, and so lie on the boundary between the region* $\{x \in \mathbb{R} : \text{sseq}(x) = --\}$ *and*

the region $\{x \in \mathbb{R} : \text{sseq}(x) = -+\}$. *As a result, the contributions of* β_1 *and* β_2 *to the eigenvalue configuration vector are split equally between those two regions.* $\triangle$

Remark 58. *In the following proof, we use bars over certain symbols (e.g.* $\overline{H}$*) to denote "augmented" versions of those objects used previously in this section. These are all defined analogously to their corresponding versions used earlier.*

Proof of Lemma 55. The proof is long, so we will divide the proof into a few stages.

1. First, we apply [32] as in Lemma 47, with the key difference that we now need to use a larger $3^m \times 3^m$ matrix $\overline{H}$ since the sign sequence of β_j's may now contain zeros. We will then split up the matrix $\overline{H}$ and find an expression for A_{sig}.

2. Next, we will find an explicit form for the term $B = H^T \widehat{H}$ appearing in the above expression.

3. Finally, we rewrite $H^{-1} A_{\text{sig}}$ in a form which allows us to quantify the contribution of each eigenvalue β_j to the overall eigenvalue configuration vector.

Now, we elaborate the details for each step.

1. First, we apply [32] as in Lemma 47, with the key difference that we now need to use a larger $3^m \times 3^m$ matrix $\overline{H}$ since the sign sequence of some β_j's may now contain zeros. We will then split up the matrix $\overline{H}$ and find an expression for A_{sig} in terms of these pieces of $\overline{H}$.

   Let

   $$\begin{aligned} (\overline{H})_{\overline{e},\overline{s}} &= \overline{s}_1^{\overline{e}_1} \cdots \overline{s}_m^{\overline{e}_m} && \text{where } \overline{e} \in \{0,1,2\}^m \text{ and } \overline{s} \in \{-,0,+\}^m \\ \overline{A_{\text{sig}_{\overline{e}}}} &= \sigma(f_{\overline{e}}(G)) && \text{where } \overline{e} \in \{0,1,2\}^m. \end{aligned}$$

   Note that H (as defined earlier) is a submatrix of $\overline{H}$, and $\overline{H}$ is the same matrix which appears in Theorem 11.

   Now, if we apply very similar reasoning as used in the proof of Lemma 47, we have

   $$\overline{A_{\text{sig}}} = \overline{H}\,\overline{q}. \tag{2.7}$$

   where

   $$\begin{aligned} \sigma_e &= \sigma(f_e(G)) && \text{for } e \in \{0,1\}^m \\ \widehat{\sigma} &= \overline{\sigma} \setminus \sigma \\ \overline{q}_s &= \#\left\{ j : \text{sseq}(\beta_j) = s \right\} && \text{for } s \in \{-,0,+\}^m \\ q_s &= \#\left\{ j : \text{sseq}(\beta_j) = s \right\} && \text{for } s \in \{-,+\}^m \\ \widehat{q} &= \overline{q} \setminus q. \end{aligned}$$

The notation $\overline{\sigma} \setminus \sigma$ (similarly for $\overline{q} \setminus q$) means that from the column vector $\overline{\sigma}$, we take away the elements from σ. Note that the key difference between (2.7) and the reasoning in Lemma 47 is that now we have $f_{\overline{e}}(\beta_j) = 0$ for some j and some $\overline{e}$, since F and G are no longer strongly generic. Hence, we must use the "augmented" matrix $\overline{H} \in \mathbb{N}^{3^m \times 3^m}$, rather than the previously used $H \in \mathbb{N}^{2^m \times 2^m}$. This hurts the space complexity (i.e., the number of matrices whose signatures need to be considered) by increasing the factor from 2^m to 3^m. We will now address this by rewriting (2.7) in such a way as to reduce the sizes of the matrices involved.

In block matrix form, after appropriately rearranging the rows and columns we have

$$\underbrace{\left[\begin{array}{c} \sigma \\ \hline \widehat{\sigma} \end{array}\right]}_{\overline{\sigma}} = \underbrace{\left[\begin{array}{c|c} H & \widehat{H} \\ \hline * & * \end{array}\right]}_{\overline{H}} \underbrace{\left[\begin{array}{c} q \\ \hline \widehat{q} \end{array}\right]}_{\overline{q}}.$$

Thus

$$\begin{aligned} \sigma &= Hq + \widehat{H}\widehat{q} \\ &= Hq + HH^{-1}\widehat{H}\widehat{q} \\ &= H(q + H^{-1}\widehat{H}\widehat{q}) \\ &= H\left(q + \frac{1}{2^m}\underbrace{H^T\widehat{H}}_{B}\widehat{q}\right). \end{aligned}$$

Note that the last equality follows because H, being a Hadamard matrix, has inverse $\frac{1}{2^m}H^T$.

Hence, we have found an expression for σ in terms of $H, q, \widehat{q}$, and $\widehat{H}$.

2. Next, we will find an explicit form for the term $B = H^T\widehat{H}$ appearing in the above expression. Note that in the following, the term $\widehat{s}$ denotes elements of $\{-,0,+\}^m \setminus \{-,+\}^m$, i.e., sign sequences of length m with at least one zero.

   Note that by construction, the matrix B can be indexed in its rows by $s \in \{-,+\}^m$ and in its columns by $\widehat{s} \in \{-,0,+\}^m \setminus \{-,+\}^m$. Thus we have

$$\begin{aligned} &B_{s,\widehat{s}} \\ &\quad\Big\| \quad \text{by expanding the matrix multiplication} \\ &\sum_{e \in \{0,1\}^m} (H^T)_{s,e}\widehat{H}_{e,\widehat{s}} \end{aligned}$$

$$\Big\| \quad \text{by the definition of } H$$

$$\sum_{e\in\{0,1\}^m} \left(s_1^{e_1}\cdots s_m^{e_m}\right)\left(\widehat{s}_1^{e_1}\cdots \widehat{s}_m^{e_m}\right)$$

$$\Big\| \quad \text{combining terms by exponents } e_i$$

$$\sum_{e\in\{0,1\}^m} \prod_{i=1}^{m} (s_i\widehat{s}_i)^{e_i}$$

$$\Big\| \quad \text{since } e_i \text{ only has values of 0 or 1}$$

$$\sum_{e\in\{0,1\}^m} \prod_{i=1}^{m} s_i\widehat{s}_i$$

$$\Big\| \quad \text{since the terms in the product are 1 if } e_i = 0$$

$$\sum_{e\in\{0,1\}^m} \prod_{\substack{i=1\\ e_i=1}}^{m} s_i\widehat{s}_i$$

$$\Big\| \quad \text{since the number of elements in } \{0,1\}^m \text{ equals the number of subsets } T\subseteq[m]$$

$$\sum_{T\subseteq[m]} \prod_{t\in T} s_t\widehat{s}_t$$

$$\Big\| \quad \text{since for nonzero } t\in T, \text{we have } s_t\widehat{s_t} = -1 \text{ if } s_t \text{ and } \widehat{s_t} \text{ differ or } -1 \text{ if they are the same}$$

$$\sum_{T\subseteq Z_{\widehat{s}}} \prod_{t\in T} \begin{cases} -1 & \text{if } s_t \neq \widehat{s}_t \\ 1 & \text{else} \end{cases} \quad \text{where } Z_{\widehat{s}} := \{t\in[m] : \widehat{s}_t \neq 0\}$$

$$\Big\| \quad \text{since only multiplications by } -1 \text{ affect the product}$$

$$\sum_{T\subseteq Z_{\widehat{s}}} (-1)^{\#\{t\in T : s_t\neq\widehat{s}_t\}}.$$

Let $U_T = \{t\in T : s_t \neq \widehat{s}_t\}$, and let $U = U_{Z_{\widehat{s}}}$.

Next, note that for each positive integer k, the number of subsets $T\subseteq Z_{\widehat{s}}$ such that $\#(T\cap U) = k$ is

$$\binom{\#U}{k} 2^{\#(Z_{\widehat{s}}\setminus U)}.$$

Further, note that the summand terms $(-1)^{\#U_T}$ depend only on the size of the sets $T\cap U$.

Hence, by partitioning the sum

$$\sum_{T \subseteq Z_{\widehat{s}}} (-1)^{\#U_T}$$

by the size of $\#(T \cap U)$, we get

$$\begin{aligned} B_{s,\widehat{s}} &= \sum_{T \subseteq Z_{\widehat{s}}} (-1)^{\#U_T} \\ &= \sum_{k=0}^{\#U} (-1)^k \binom{\#U}{k} 2^{\#(Z_{\widehat{s}} \setminus U)}. \end{aligned}$$

Rearranging, we get

$$\begin{aligned} B_{s,\widehat{s}} &= 2^{\#(Z_{\widehat{s}} \setminus U)} \sum_{k=0}^{\#U} (-1)^k \binom{\#U}{k} \\ &= \begin{cases} 2^{\#Z_{\widehat{s}}} & \text{if } U = \emptyset \\ 0 & \text{if } U \neq \emptyset. \end{cases} \end{aligned}$$

Note that by definition of ∂s, we have that $U = \emptyset$ if and only if $\widehat{s} \in \partial s$. Hence

$$B_{s,\widehat{s}} = \begin{cases} 2^{\#Z_{\widehat{s}}} & \text{if } \widehat{s} \in \partial s \\ 0 & \text{else .} \end{cases}$$

With that, we have found an explicit form for the B matrix.

3. Finally, we rewrite $H^{-1}A_{\text{sig}}$ in a form which allows us to quantify the "contribution" of each eigenvalue β_j to the overall eigenvalue configuration vector. From the above, for all $s \in \{-,+\}^m$ we have that the s-th component of $H^{-1}A_{\text{sig}}$ is

$$\begin{aligned} (H^{-1}A_{\text{sig}})_s &= \left(q + \frac{1}{2^m} B \widehat{q} \right)_s \\ &= q_s + \frac{1}{2^m} \sum_{\widehat{s} \in \partial s} 2^{\#Z_{\widehat{s}}} \widehat{q}_{\widehat{s}} \\ &= q_s + \sum_{\widehat{s} \in \partial s} \left(\frac{1}{2} \right)^{\#\{0 \in \widehat{s}\}} \widehat{q}_{\widehat{s}} \qquad \text{since } \#Z_{\widehat{s}} - m = \#\{0 \in \widehat{s}\}. \end{aligned}$$

Noting that $\text{cl}(s) = \partial s \cup s$, we can push the q_s term into the summation and get

$$(H^{-1}A_{\text{sig}})_s = \sum_{s' \in \text{cl}(s)} \left(\frac{1}{2} \right)^{\#\{0 \in s'\}} \overline{q}_{s'}.$$

Note that we switch to the symbol $\overline{q}_{s'}$ to denote the fact that s' could come from the set $\{-,0,+\}^m$ or $\{-,+\}^m$. Next, we rewrite $\overline{q}_{s'}$ as a summation of indicators.

$$(H^{-1}A_{\text{sig}})_s = \sum_{s' \in \text{cl}(s)} \left(\frac{1}{2}\right)^{\#\{0 \in s'\}} \sum_{j=1}^{n} \delta_{\text{sseq}(\beta_j), s'},$$

where here δ denotes the Kronecker delta function. Continuing, by rearranging the order of summation we have

$$(H^{-1}A_{\text{sig}})_s = \sum_{j=1}^{n} \sum_{s' \in \text{cl}(s)} \left(\frac{1}{2}\right)^{\#\{0 \in s'\}} \delta_{\text{sseq}(\beta_j), s'}.$$

Now, note that the summand is nonzero if and only if $\text{sseq}(\beta_j) = s'$. If we combine this with the innermost summation over elements of $\text{cl}(s)$, we can rewrite as

$$(H^{-1}A_{\text{sig}})_s = \sum_{j=1}^{n} \left(\frac{1}{2}\right)^{\#\{0 \in \text{sseq}(\beta_j)\}} \delta_{\text{sseq}(\beta_j) \in \text{cl}(s)}.$$

Now, fix $j \in [n]$ and recall that $S_j = \{s \in \{-,+\}^m : \text{sseq}(\beta_j) \in \text{cl}(s)\}$. If $\text{sseq}(\beta_j)$ has no zeroes, then the set S_j consists of exactly one element; namely, $\text{sseq}(\beta_j)$ itself. Now suppose that $\text{sseq}(\beta_j) = (\ldots, 0, \ldots)$; i.e., there is at least one zero. Then for each zero, there are two corresponding elements in S_j, which are $(\ldots, +, \ldots)$ and $(\ldots, -, \ldots)$. Hence

$$\#S_j = 2^{\#\{0 \in \text{sseq}(\beta_j)\}}.$$

Thus

$$\begin{aligned}
(H^{-1}A_{\text{sig}})_s &= \sum_{j=1}^{n} \left(\frac{1}{2}\right)^{\#\{0 \in \text{sseq}(\beta_j)\}} \delta_{\text{sseq}(\beta_j) \in \text{cl}(s)} \\
&= \sum_{j=1}^{n} \frac{1}{\#S_j} \delta_{\text{sseq}(\beta_j) \in \text{cl}(s)} \\
&= \sum_{j=1}^{n} \frac{1}{\#S_j} \sum_{s' \in S_j} \delta_{s', \text{sseq}(\beta_j)}.
\end{aligned}$$

Therefore, in full vector form, we have

$$H^{-1}A_{\text{sig}} = \sum_{j=1}^{n} \frac{1}{\#S_j} \sum_{s' \in S_j} e_{s'} = \sum_{j=1}^{n} \gamma_j.$$

□

Before we continue with the proof, we first need to take a short detour and establish a simple property of polynomials when they have only real roots.

Lemma 59 (Real roots property)**.** *Let $f \in \mathbb{R}[x]$ be a polynomial with only real roots and let $\beta \in \mathbb{R}$. Then we have*

$$\beta \text{ is a multiple root of } f' \qquad \Longrightarrow \qquad \beta \text{ is a multiple root of } f.$$

Proof. Let $f \in \mathbb{R}[x]$ have only real roots and let $\beta \in \mathbb{R}$. We proceed by contradiction. For this, we will assume that β is a multiple root of f' but that β is *not* a multiple root of f. Let m be the degree of f. Then, we can write f as

$$f(x) = (x-\alpha_1)^{\mu_1} \cdots (x-\alpha_t)^{\mu_t}$$

where

$$\begin{aligned}
&\alpha_1, \ldots, \alpha_t \text{ distinct real numbers} \\
&\mu_i \geq 1, \\
&\mu_1 + \cdots + \mu_t = m, \text{ since } f \text{ has only real roots.}
\end{aligned}$$

We will now proceed to count the roots of f' in two different ways. On the one hand, by the Gauss-Lucas theorem, we have that

$$\#\text{ real roots of } f' = m-1.$$

On the other hand, by Rolle's theorem, there exists at least one root of f' in the open interval (α_i, α_{i+1}) for $i \in \{1, \ldots, t-1\}$. Let ρ_i be the number of roots, counted with multiplicity, in the interval (α_i, α_{i+1}). Furthermore, the multiplicity of α_i as a root of f' is $\mu_i - 1$. Thus, counting the roots of f' in this way, we have

$$\begin{aligned}
\#\text{ real roots of } f' &= \underbrace{(\mu_1 - 1) + \cdots + (\mu_t - 1)}_{\text{roots coming from the } \alpha_i\text{'s}} + \underbrace{\sum_{i=1}^{t-1} \rho_i}_{\text{Rolle's}} \\
&= (\mu_1 + \cdots + \mu_t) - t \qquad + \sum_{i=1}^{t-1} \rho_i \\
&= m - t \qquad\qquad\qquad + \sum_{i=1}^{t-1} \rho_i.
\end{aligned}$$

Now, recall that we assumed that β is a multiple root of f'. Equivalently, we have that

$$f'(\beta) = 0 \ \wedge \ f''(\beta) = 0. \tag{2.8}$$

Further, we assumed that β is *not* a multiple root of f. This means that

$$\neg(f(\beta) = 0 \ \wedge \ f''(\beta) = 0)$$

equivalently

$$f(\beta) \neq 0 \ \vee \ f''(\beta) \neq 0. \tag{2.9}$$

The assumptions (2.8) and (2.9) together imply that $f(\beta) \neq 0$. Hence, we have that $\alpha_i \neq \beta$ for all i. Further, this implies that $\rho_i > 1$ for at least one $i \in \{1, \ldots, t-1\}$. This means that $\sum_{i=1}^{t-1} \rho_i \geq t$. Putting everything together, we have that

$$\begin{aligned} \#\text{ real roots of } f' &= m - t + \sum_{i=1}^{t-1} \rho_i \\ &\geq m - t + t = m. \end{aligned}$$

This is a contradiction, since earlier we saw that

$$\#\text{ real roots of } f' = m - 1.$$

Hence, we have that β must be a multiple root of f, and so the lemma is proved. □

Corollary 60 (Multiple roots of derivatives of polynomials with only real roots)**.** *Let $k \in \mathbb{N}$. Let $f \in \mathbb{R}[x]$ have only real roots and let $\beta \in \mathbb{R}$. Then*

$$\beta \textit{ is a multiple root of } f^{(k)} \qquad \Longrightarrow \qquad \beta \textit{ is a multiple root of } f.$$

Proof. Note

$$\begin{array}{lcll} \beta \text{ is a multiple root of } f^{(1)} & \Longrightarrow & \beta \text{ is a multiple root of } f^{(0)} & \text{by Lemma 59 to } f^{(0)} \\ \beta \text{ is a multiple root of } f^{(2)} & \Longrightarrow & \beta \text{ is a multiple root of } f^{(1)} & \text{by Lemma 59 to } f^{(1)} \\ & \vdots & & \\ \beta \text{ is a multiple root of } f^{(k-1)} & \Longrightarrow & \beta \text{ is a multiple root of } f^{(k)} & \text{by Lemma 59 to } f^{(k)}. \end{array}$$

Following the chain of implications gives

$$\beta \text{ is a multiple root of } f^{(k)} \qquad \Longrightarrow \qquad \beta \text{ is a multiple root of } f$$

and we are done. □

With that, we are ready to complete the final piece of the proof of Theorem 30.

Lemma 61. *Let F, G be generic but not strongly generic, and let $\widehat{F}$ be symmetric. If we have*

1. $\mathrm{GEC}(F,G) = \mathrm{GEC}(\widehat{F},G)$ *and*
2. $\widehat{F}$ *and* G *are strongly generic,*

then

$$C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F},G).$$

Proof. Assume that F, G are generic but not strongly generic, and let $\widehat{F}$ be symmetric. Suppose that

1. $\mathrm{GEC}(F,G) = \mathrm{GEC}(\widehat{F},G)$ and
2. $\widehat{F}$ and G are strongly generic.

From Lemma 55 , we have

$$H^{-1} A_{\mathrm{sig}}(F,G) = \sum_{j=1}^{n} \gamma_j$$

and

$$H^{-1} A_{\mathrm{sig}}(\widehat{F},G) = \sum_{j=1}^{n} \widehat{\gamma}_j$$

Thus, it suffices to show

$$V \sum_{j=1}^{n} \gamma_j = V \sum_{j=1}^{n} \widehat{\gamma}_j,$$

equivalently, to show

$$\sum_{j=1}^{n} V\gamma_j = \sum_{j=1}^{n} V\widehat{\gamma}_j.$$

Below, we will show a stronger result: for all $j = 1, \ldots, n$, we have

$$V\gamma_j = V\widehat{\gamma}_j.$$

Let $j \in \{1,\dots,n\}$ be arbitrary but fixed. Recall that

$$\gamma_j = \frac{1}{\#S_j} \sum_{s' \in S_j} e_{s'}$$

where $S_j = \{s \in \{-,+\}^m : \mathrm{sseq}(\beta_j) \in \mathrm{cl}(s)\}$, with $\widehat{\gamma}_j$ defined analogously with $\widehat{\beta}_j$.
Using the definition of V from the proof of Lemma 45, we then have the following (note that γ_j could be substituted with $\widehat{\gamma}_j$):

$$(V\gamma_j)_t = \sum_{\substack{s \in \{-,+\}^m \\ v(s,+)=m-t}} (\gamma_j)_s \qquad \text{(where subscript means "}s\text{-th" element of the vector in lex order)}$$

$$\Big\| \quad \text{by definition of } \gamma_j$$

$$\sum_{\substack{s \in \{-,+\}^m \\ v(s,+)=m-t}} \left(\frac{1}{\#S_j} \sum_{s' \in S_j} e_{s'} \right)_s$$

$$\Big\| \quad \text{by evaluating the innermost summation}$$

$$\sum_{\substack{s \in \{-,+\}^m \\ v(s,+)=m-t}} \frac{1}{\#S_j} \delta_{s \in S_j}$$

$$\Big\| \quad \text{by rearranging}$$

$$\frac{1}{\#S_j} \sum_{\substack{s \in \{-,+\}^m \\ v(s,+)=m-t}} \delta_{s \in S_j}$$

$$\Big\|$$

$$\frac{1}{\#S_j} \#\{s \in S_j : v(s,+) = m-t\}.$$

We now make the following claim.

Claim: We have

$$V\gamma_j = e_{m-v(\mathrm{sseq}_f(\beta_j),+)} \qquad \text{and} \qquad V\widehat{\gamma}_j = e_{m-v(\mathrm{sseq}_{\widehat{f}}(\beta_j),+)}. \tag{2.10}$$

We will show the proof for the former equality, because the proof for the latter is essentially

identical. To prove the claim, it suffices to show that for all $s \in S_j$, we have $v(s,+) = v(\text{sseq}(\beta_j), s)$.

Let $s \in S_j$. Then we have two cases: either $\text{sseq}(\beta_j) = s$ or not.

Case 1: $\text{sseq}(\beta_j) = s$. Then $S_j = \{s\}$, and so

$$\begin{aligned}(V\gamma_j)_t &= \frac{1}{\#S_j}\#\{s \in S_j : v(s,+) = m-t\} \\ &= \frac{1}{1}\#\{s \in \{s\} : v(s,+) = m-t\} \\ &= \delta_{v(\text{sseq}(\beta_j),+),m-t}\,.\end{aligned}$$

Thus, in matrix form, we have $V\gamma_j = e_{m-v(\text{sseq}(\beta_j),+)}$ and the claim follows.

Case 2: $\text{sseq}(\beta_j) \neq s$. Then $\text{sseq}(\beta_j)$ and s differ by one or more zeros. At first glance, it is tempting to use the fact that for all $y \in \mathbb{R}$, we have that

$$\#\{x : f(x) > 0 \wedge x > y\} = v(\text{sseq}(y),+),$$

which was proved as part of Lemma 43 using Descartes' rule of signs. However, it is not always true that there exists some $y \in \mathbb{R}$ so that $s = \text{sseq}(y)$. In fact, since S_j comprises the sign sequences of open sets in $\mathbb{R}^m$ which share a boundary with the (not open) set $\{x : \text{sseq}(x) = \text{sseq}(\beta_j)\}$, there are at most two sign sequences in S_j that are actually realizable. As a result, we need a different approach. We will instead study the possibilities for $(\text{sseq}(\beta_j),+)$ under our genericity assumptions for F.

First, recall that F has exactly m (counting with multiplicity) real eigenvalues. We then have the following:

1. By the Gauss-Lucas theorem, for $k \in \{0,\dots,m-1\}$, the polynomial $f^{(k)}$ has exactly $m-k$ real roots, counting with multiplicity.
2. By Corollary 60, there are no consecutive zeros in $\text{sseq}(\beta_j)$, since multiple zeros would imply that β_j is a multiple root of $f^{(k)}$ for some k, which by Corollary 60 implies that $f(\beta_j) = 0$, which we assumed is false due to the genericity of F and G.
3. By the assumption that F and G are generic (i.e. do not share eigenvalues), the first entry in $\text{sseq}(\beta_j)$ is nonzero.

Under these restrictions, the only subsequences of $(\text{sseq}(\beta_j),+)$ which contradict (2.10) are

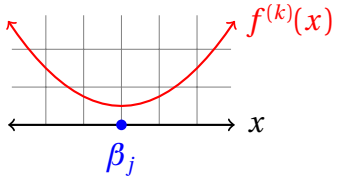
$$(\dots,+,0,+,\dots) \qquad \text{and} \qquad (\dots,-,0,-,\dots).$$

This is because changing the zero to either $+$ or $-$ in both of these subsequences can change the sign variation count of the subsequence, whereas doing so in the other possibilities $-,0,+$ and $+,0,-$ does not change the sign variation count.

Hence, without loss of generality, it suffices to show that $+,0,+$ cannot occur in $(\text{sseq}(\beta_j),+)$. (The argument for $-,0,-$ will be essentially identical.) Suppose that $(\text{sseq}(\beta_j),+)$ does contain $+,0,+$. Then there exists $k \in \{0,\dots,m\}$ so that

$$\begin{aligned} f^{(k)}(\beta_j) &> 0 \\ f^{(k+1)}(\beta_j) &= 0 \\ f^{(k+2)}(\beta_j) &> 0. \end{aligned}$$

Pictorially, we have the situation shown in the diagram below.



But this means that $f^{(k)}$ has fewer than $m-k$ real roots (counting multiplicity), which by Rolle's theorem contradicts the fact that f has all real roots. Therefore, $+,0,+$ (and, similarly, the subsequence $-,0,-$) cannot occur anywhere in $(\text{sseq}(\beta_j),+)$.

To summarize, we have established that $(\text{sseq}(\beta_j),+)$ has at least one zero, and that each zero must appear as in one of the subsequences $+,0,-$ or $-,0,+$. Note that in both of these, changing the zero to either a $+$ or a $-$ does not change the sign variation count. Hence, it follows that for all $s \in S_j$, we have $v(s,+) = v(\text{sseq}(\beta_j),+)$. Then

$$\begin{aligned} (V\gamma_j)_t &= \frac{1}{\#S_j}\#\{s \in S_j : v(s,+) = m-t\} \\ &= \frac{1}{\#S_j}\begin{cases} \#S_j & \text{if } v(\text{sseq}(\beta_j),+) = m-t \\ 0 & \text{otherwise.} \end{cases} \\ &= \delta_{v(\text{sseq}(\beta_j),+),m-t}\,. \end{aligned}$$

Thus, in matrix form, we have $V\gamma_j = e_{m-v(\text{sseq}(\beta_j),+)}$ and the claim follows.

From the two cases above, the claim (2.10) is now proved. Thus we have shown that

$$V\gamma_j = e_{m-\nu(\mathrm{sseq}_f(\beta_j),+)} \qquad \text{and} \qquad V\widehat{\gamma}_j = e_{m-\nu(\mathrm{sseq}_{\widehat{f}}(\beta_j),+)}.$$

Finally, note that

$$\begin{aligned}
&\nu(\mathrm{sseq}_f(\beta_j),+)\\
&\qquad \Big\| \quad \text{by the reasoning used in Lemma 43}\\
&\#\{x : f(x)=0 \wedge x>\beta_j\}\\
&\qquad \Big\| \quad \text{since } \mathrm{GEC}(F,G)=\mathrm{GEC}(\widehat{F},G)\\
&\#\{x : \widehat{f}(x)=0 \wedge x>\beta_j\}\\
&\qquad \Big\| \quad \text{by the reasoning used in Lemma 43}\\
&\nu(\mathrm{sseq}_{\widehat{f}}(\beta_j),+).
\end{aligned}$$

Thus $V\gamma_j = V\widehat{\gamma}_j$, and so $C_{\mathrm{sig}}\ A_{\mathrm{sig}}(F,G) = C_{\mathrm{sig}} A_{\mathrm{sig}}(\widehat{F},G)$. We have proved the lemma. □

Finally, we are now ready to prove Theorem 30 (Signature approach for generic pairs of matrices).

Proof of Theorem 30. Let F and G be a generic pair of real symmetric matrices. We need to prove that

$$\mathrm{GEC}(F,G) = C_{\mathrm{sig}}\ A_{\mathrm{sig}}(F,G).$$

We consider two cases.

Case 1: F and G are strongly generic.

From Lemma 39, we have

$$\mathrm{GEC}(F,G) = C_{\mathrm{sig}}\ A_{\mathrm{sig}}(F,G)$$

and so we are done.

Case 2: F and G are not strongly generic. Recall the following.

1. From Lemma 50, there exists some real symmetric matrix $\widehat{F}$ so that $\widehat{F}$ and G are strongly generic and $\mathrm{GEC}(F,G) = \mathrm{GEC}(\widehat{F},G)$.
2. From Lemma 39, we have $\mathrm{GEC}(\widehat{F},G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F},G)$.
3. From Lemma 61, we have $C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F},G) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G)$.

Putting these together, we therefore have

$$\mathrm{GEC}(F,G) \underset{1}{=} \mathrm{GEC}(\widehat{F},G) \underset{2}{=} C_{\mathrm{sig}}\, A_{\mathrm{sig}}(\widehat{F},G) \underset{3}{=} C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G).$$

□

2.3 Symmetric Polynomials Approach

In this section, we will solve the generic eigenvalue configuration problem using an approach based on the Fundamental Theorem of Symmetric Polynomials.

2.3.1 Main Theorem

In this subsection, we will state the main theorem for this section. For this, we first introduce two notions which are central to the main result: one is purely *combinatorial* (depending only on the size of F) and the other is *algebraic* (depending on the entries/parameters of both F and G).

Definition 62 (Combinatorial part, symmetric polynomials approach)**.** *The matrix $C_{\mathrm{sym}} \in \mathbb{Q}^{m\times m}$ is the matrix*

$$C_{\mathrm{sym}} = V_m W_m$$

where

$$(V_m)_{rt} = \frac{(-1)^{r-1}}{2^{t-1}}\binom{t}{r} \quad \textit{and} \quad (W_m)_{ts} = (-1)^{t+s}\binom{m-s}{m-t}.$$

Example 63. *Let $m = 4$. We will construct the matrix C_{sym}. Consider for example the entry at row 3, column 3. Then we have*

$$\begin{aligned}(C_{\mathrm{sym}})_{3,3} &= \sum_{t=1}^{m} (V_4)_{3,t}(W_4)_{t,3} \\ &= \sum_{t=1}^{4} \frac{1}{2^{t-1}}\binom{t}{3}(-1)^{t+3}\binom{1}{4-t} && \textit{by definition of } V \textit{ and } W\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2^0}\binom{1}{3}(-1)^4\binom{1}{3} + \frac{1}{2^1}\binom{2}{3}(-1)^5\binom{1}{2} \\
&\quad + \frac{1}{2^2}\binom{3}{3}(-1)^6\binom{1}{1} + \frac{1}{2^3}\binom{4}{3}(-1)^7\binom{1}{0} \quad \textit{by expanding the sum} \\
&= 0 + 0 - \frac{1}{4} + 0 \\
&= -\frac{1}{4}.
\end{aligned}
$$

Repeating this process for all entries of C_{sym}, we have

$$
C_{\text{sym}} = \begin{bmatrix} -\frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 & -\frac{3}{4} \\ \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{2} \\ \frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{1}{8} \end{bmatrix}.
$$

△

Definition 64 (Algebraic part, symmetric polynomials approach)**.** *The column vector A_{sym}, or $A_{\text{sym}}(F,G)$ to emphasize the dependence on the entries of F and G, is the column vector of length m whose r-th entry is*

$$(A_{\text{sym}})_r = v(D_r),$$

where v denotes the sign variation count of the coefficients of a polynomial and D_r is a polynomial in $\mathbb{R}[a,b][x]$ such that

$$D_r(a,b,x) = h_r(\alpha,\beta,x),$$

where h_r is the polynomial

$$h_r = \prod_{\substack{I\subset[m],\,\#I=r \\ j\in[n]}} \left(x + \prod_{i\in I}(\alpha_i - \beta_j) \right),$$

and where $a = (a_1,\ldots,a_m)$ and $b = (b_1,\ldots,b_n)$ are such that

$$
\begin{aligned}
f &= \det(xI_m - F) = x^m - a_1x^{m-1} + a_2x^{m-2} - \cdots + (-1)^m a_m x^0 \\
g &= \det(xI_n - G) = x^n - b_1x^{n-1} + b_2x^{n-2} - \cdots + (-1)^n b_n x^0.
\end{aligned}
$$

Note the alternating signs and reverse indexing from the usual indexing of polynomial coefficients. Equivalently, the coefficient a_i is the coefficient of x^{m-i} in the polynomial $\det(xI_m + F)$ *(and similarly with b_i and G).*

Example 65. *Let $m = n = 2$. We will construct the entry D_2 in A_{sym}. First, we have that*

$$h_2 = \prod_{\substack{I\subset\{1,2\},\,\#I=2\\ j\in\{1,2\}}} \left(x + \prod_{i\in I} (\alpha_i - \beta_j) \right)$$
$$(x + (\alpha_1-\beta_1)(\alpha_2-\beta_1))(x + (\alpha_1-\beta_2)(\alpha_2-\beta_2)).$$

Then, we set

$$D_2 = x^2 + (-a_1 b_1 + b_1^2 + 2a_2 - 2b_2)x + a_1^2 b_2 - a_1 a_2 b_1 - a_1 b_1 b_2 + a_2 b_1^2 + a_2^2 - 2a_2 b_2 + b_2^2,$$

where a_1, a_2, b_1, b_2 are the respective coefficients of the characteristic polynomials of F and G. One can then verify that $D_2(a, b, x) = h_2(\alpha, \beta, x)$ by expressing a_1, a_2, b_1, b_2 in terms of the eigenvalues α and β. △

Now we are ready to state our main result.

Theorem 66 (Symmetric polynomials approach for generic pairs of matrices)**.** *Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be a generic pair of real symmetric matrices. We have*

$$\text{GEC}(F, G) \;=\; C_{\text{sym}}\; A_{\text{sym}}(F, G).$$

Remark 67. *Note that the matrix C_{sym} is entirely numeric and depends only on m. In addition, the vector A_{sym} depends on the sign variation count of the polynomials D_r, whose coefficients are polynomials in the entries of F and G. Hence, the right-hand side of the above contains no references to the eigenvalues of F and G and is therefore quantifier-free.*

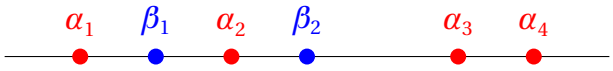
Remark 68. *The reader may notice that the main theorem in this section is stated almost identically to Theorem 30 which appeared in the previous section. However, they are completely different results, as C and A are defined as completely different concepts for each approach. The main theorem in this section is based on the theory of symmetric polynomials while the main theorem in the previous section is based on the signature of matrices. However, in both theorems, there is a combinatorial part and an algebraic part; hence, we structure the theorem statements similarly to emphasize the super-structural similarities and the conceptual differences.*

Example 69 (General case with parametric matrices)**.** *Let $m = 4$ and $n = 2$ and let F and G be*

parametric matrices with each entry being an independent parameter; that is, let

$$F = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{1,2} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{1,3} & a_{2,3} & a_{3,3} & a_{3,4} \\ a_{1,4} & a_{2,4} & a_{3,4} & a_{4,4} \end{bmatrix}, \qquad G = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{1,2} & b_{2,2} \end{bmatrix},$$

where each a_{ij} and b_{ij} is an independent parameter. We will now use Theorem 66 to write a condition on these parameters so that the eigenvalues of F and G are arranged as in the following picture.



That is, we will find a quantifier-free condition for $\mathrm{GEC}(F,G) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$. *By Theorem 66, we have*

$$\mathrm{GEC}(F,G) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = C_{\mathrm{sym}}\, A_{\mathrm{sym}}(F,G).$$

In Example 63 we found that

$$C_{\mathrm{sym}} = \begin{bmatrix} -\frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 & -\frac{3}{4} \\ \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{2} \\ \frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{1}{8} \end{bmatrix}.$$

From Definition 64 we have

$$A_{\mathrm{sym}}(F,G) = \begin{bmatrix} v(D_1) \\ v(D_2) \\ v(D_3) \\ v(D_4) \end{bmatrix},$$

where each D_r can be computed using Definition 64 as in Example 65. Hence we have that

$$\mathrm{GEC}(F,G)=\begin{bmatrix}1\\1\\0\\0\end{bmatrix} \quad\Longleftrightarrow\quad \begin{bmatrix}1\\1\\0\\0\end{bmatrix}=\begin{bmatrix}-\frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2}\\ 0 & \frac{1}{4} & 0 & -\frac{3}{4}\\ \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{2}\\ \frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{1}{8}\end{bmatrix}\begin{bmatrix}v(D_1)\\v(D_2)\\v(D_3)\\v(D_4)\end{bmatrix}.$$

Solving the linear system above gives the solution

$$\begin{bmatrix}v(D_1)\\v(D_2)\\v(D_3)\\v(D_4)\end{bmatrix}=\begin{bmatrix}3\\7\\5\\1\end{bmatrix}.$$

We therefore have that

$$\mathrm{GEC}(F,G)=\begin{bmatrix}1\\1\\0\\0\end{bmatrix} \quad\Longleftrightarrow\quad \begin{bmatrix}v(D_1)\\v(D_2)\\v(D_3)\\v(D_4)\end{bmatrix}=\begin{bmatrix}3\\7\\5\\1\end{bmatrix}.$$

Since each D_r is a polynomial with coefficients which are themselves coefficients in the parameters a_{ij} and b_{ij}, the right-hand side of the above is quantifier-free. △

Example 70 (Matrices with only one parameter)**.** *In Example 69, we studied F and G which were symmetric matrices, each of whose elements was an independent parameter. However, it need not be the case that each entry is an independent parameter. We would, for instance, have the entries of both matrices be polynomials in some smaller list of parameters. Consider, for example, the matrices*

$$F=\begin{bmatrix}1 & -2\\ -2 & 1\end{bmatrix}$$

and

$$G=\begin{bmatrix}1 & p^2+p & p+1\\ p^2+p & -p & p-1\\ p+1 & p-1 & p\end{bmatrix},$$

where p is a (single) parameter. Suppose we wish to know for which values of p the eigenvalues

of the matrix G interlace with that of F; that is, when $\mathrm{GEC}(F,G)=(1,1)$. *By Theorem 66 and Definition 62, we have that*

$$\mathrm{GEC}(F,G)=\begin{bmatrix}1\\1\end{bmatrix} \iff \begin{bmatrix}1\\1\end{bmatrix}=C_{\mathrm{sym}}\,A_{\mathrm{sym}}(F,G)$$
$$\iff \begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}0 & 1\\ \frac{1}{2} & -\frac{1}{2}\end{bmatrix}\begin{bmatrix} v(D_1)\\ v(D_2)\end{bmatrix},$$

where each D_r is the symmetric polynomial h_r written in terms of p using the Fundamental Theorem of Symmetric Polynomials; i.e.,

$$\begin{aligned}h_1(\alpha,\beta,x)&=(x+\alpha_1-\beta_1)(x+\alpha_2-\beta_1)(x+\alpha_1-\beta_2)(x+\alpha_2-\beta_2)(x+\alpha_1-\beta_3)(x+\alpha_2-\beta_3)\\ D_1(p,x)&=x^6+4x^5+(-2p^4-4p^3-8p^2-10)x^4+(2p^5-6p^4-16p^3-20p^2-2p-38)x^3\\ &\quad+(p^8+4p^7+12p^6+20p^5+14p^4-12p^3-4p+13)x^2\\ &\quad+(-2p^9-2p^8+4p^7+28p^6+64p^5+16p^4-44p^3+28p^2-22p+58)x\\ &\quad+p^{10}-2p^9-11p^8-12p^7-30p^6-32p^5-42p^4-4p^3+69p^2-14p+13\end{aligned}$$

$$\begin{aligned}h_2(\alpha,\beta,x)&=(x+(\alpha_1-\beta_1)(\alpha_2-\beta_1))(x+(\alpha_1-\beta_2)(\alpha_2-\beta_2))(x+(\alpha_1-\beta_3)(\alpha_2-\beta_3))\\ D_2(p,x)&=x^3+(2p^4+4p^3+8p^2-6)x^2\\ &\quad+(p^8+4p^7+12p^6+12p^5+6p^4-12p^3-48p^2+4p+5)x\\ &\quad+p^{10}-2p^9-11p^8-12p^7-30p^6-32p^5-42p^4-4p^3+69p^2-14p+13.\end{aligned}$$

Note that each $v(D_r)$ can be readily written in terms of p using Descartes' rule of signs. △

In algorithm form, Theorem 66 can be stated as follows.

Algorithm 71 (Condition for EC: symmetric polynomials approach)**.**

In: $F\in\mathbb{R}[p]^{m\times m}$ *and* $G\in\mathbb{R}[p]^{n\times n}$, *symmetric, where p is a finite list of parameters*
$c\in\mathbb{N}^m$, *eigenvalue configuration*

Out: *P, quantifier-free condition on p such that* $c=\mathrm{GEC}(F,G)$

1. *For* $r=1,\dots,m$ *do*

   (a) $Y_r \longleftarrow \left\{(i_1,\dots,i_r,j)\,:\,1\le i_1<\cdots<i_r\le m \ \wedge\ 1\le j\le n\right\}$

   (b) $h_r \longleftarrow \prod_{\substack{I\subset[m],\,\#I=r\\ j\in[n]}}\left(x+\prod_{i\in I}(\alpha_i-\beta_j)\right)$ $\qquad\in\mathbb{Z}[\alpha,\beta][x]$

(c) $u_r \longleftarrow \mathrm{FTSP}(h_r, [\alpha_1, \ldots, \alpha_m], [\beta_1, \ldots, \beta_n]) \quad \in \mathbb{Z}[\gamma, \delta][x]$

(d) $d_r \longleftarrow u_r\Big([\mathrm{coeff}_{z^0}(\det(zI_m + F)), \ldots, \mathrm{coeff}_{z^{m-1}}(\det(xI_m + F))],$
$[\mathrm{coeff}_{z^0}(\det(xI_n + G)), \ldots, \mathrm{coeff}_{z^{n-1}}(\det(xI_n + G))],\ x\Big) \quad \in \mathbb{Z}[p][x]$

2. $T \longleftarrow$ *matrix in* $\mathbb{Z}^{m\times m}$ *where* $T_{r,s} = \sum_{t=1}^{m} \binom{m-t}{r-t}(-2)^{t-1}\binom{s}{t}$

3. $y \longleftarrow Tc$

4. $P \longleftarrow \bigwedge_{1\le r\le m} v(d_r) = y_r$

5. *Return* P

Algorithm 72 (FTSP)**.**

In : $h \in \mathbb{R}[\alpha, \beta][x]$*, a polynomial symmetric in variables* $\alpha = (\alpha_1, \ldots, \alpha_m)$ *and* $\beta = (\beta_1, \ldots, \beta_n)$

Out: $u \in \mathbb{R}[\gamma, \delta][x]$ *such that* $h = u\big([e_1(\alpha), \ldots, e_m(\alpha)], \big[e_1(\beta), \ldots, e_n(\beta)\big], x\big)$

2.3.2 Proof

In this subsection, we will prove Theorem 66 (Symmetric polynomials approach for generic matrices). The proof is structured as follows.

1. First, in Lemmas 74 through 79, we establish a bijective combinatorial correspondence between the number of positive roots of the h polynomials from Definition 64 and the eigenvalue configuration vector.

2. Then, in Proposition 80, we use the fact that the polynomials h_r are symmetric in the *eigenvalues* α and β of F and G, respectively, to rewrite them in terms of the *coefficients* of the characteristic polynomials of F and G. This step eliminates all references to the eigenvalues and therefore concludes the proof.

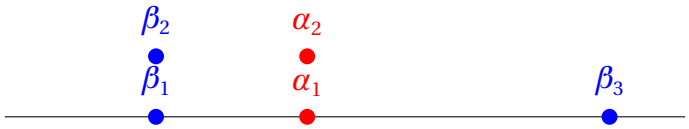
We will repeatedly revisit the following running example throughout the proof.

Example 73 (Running example)**.** *Let*

$$F = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 2} \qquad G = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Then their respective eigenvalues are

$$\alpha = (4,4) \qquad \beta = (2,2,8)$$



So,

$$\mathrm{GEC}(F,G) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

△

Lemma 74 (Transform)**.** *We have*

$$c = \mathrm{GEC}(F,G) \implies y = T_m c$$

where

$$\begin{aligned} y_r &= \#\text{ positive roots of } h_r\text{, counting multiplicity} \\ (T_m)_{rs} &= \#\{ I \subset [m] : \#I = r \ \wedge \ \#\{i \in I : i \le s\} \text{ is odd}\} \\ h_r &= \prod_{\substack{I \subset [m],\, \#I = r \\ j \in [n]}} \left(x + \prod_{i \in I} (\alpha_i - \beta_j) \right) \qquad \text{from Definition 64.} \end{aligned}$$

Example 75 (Running example)**.** *Recall Example 73, where we had F and G such that*

$$\mathrm{GEC}(F,G) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Using the fact that the respective eigenvalues of F and G are $\alpha = (4,4)$ and $\beta = (2,2,8)$, together with the definition of h_r from Definition 64, we compute

$$\begin{aligned} h_1 &= \prod_{\substack{I \subset [2],\, \#I = 1 \\ j \in [3]}} \left(x + \prod_{i \in I} (\alpha_i - \beta_j) \right) \\ &= (x + \alpha_1 - \beta_1)(x + \alpha_2 - \beta_1)(x + \alpha_1 - \beta_2)(x + \alpha_2 - \beta_2)(x + \alpha_1 - \beta_3)(x + \alpha_2 - \beta_3) \\ &= (x+2)^4 (x-4)^2 \\ h_2 &= \big(x + (-\beta_1 + \alpha_1)(-\beta_1 + \alpha_2)\big)\big(x + (-\beta_2 + \alpha_1)(-\beta_2 + \alpha_2)\big)\big(x + (\alpha_1 - \beta_3)(\alpha_2 - \beta_3)\big) \end{aligned}$$

$$= (x+4)^2(x+16).$$

We therefore have

$$y_1 = \#\ \textit{positive roots of } h_1\textit{, counting multiplicity} = 2$$
$$y_2 = \#\ \textit{positive roots of } h_2\textit{, counting multiplicity} = 0.$$

On the other hand, using the definition of T_m we construct

$$T_2 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

Therefore

$$T_2 c = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = y.$$

△

Proof of Lemma 74. Assume that $c = \mathrm{GEC}(F,G)$. It suffices to show $y = T_m c$. Recall that

$$y_r = \#\text{ positive roots of } h_r\text{, counting multiplicity.}$$

Equivalently, using the definition of h_r, we have

$$y_r = \#\left\{(i_1,\ldots,i_r,j) \in Y_r : \prod_{p=1}^{r}\left(\alpha_{i_p} - \beta_j\right) < 0\right\},$$

where

$$Y_r = \{(i_1,\ldots,i_r,j) : 1 \le i_1 < \cdots < i_r \le m \ \wedge\ j \in [n]\}.$$

We proceed by repeatedly rewriting the definition of y_r, with the goal of expressing it in terms of the eigenvalue configuration vector. We begin with the definition of y_r.

$$y_r = \#\left\{\left(i_1,\ldots,i_r,j\right) \in Y_r : \prod_{p=1}^{r}\left(\alpha_{i_p} - \beta_j\right) < 0\right\}$$

Note that for each $(i_1,\ldots,i_r)$ satisfying $1 \le i_1 < \cdots < i_r \le m$, we also have $(i_1,\ldots,i_r,j) \in Y_r$ simply by appending each $j = 1,\ldots,m$. Hence, we can rewrite this action of counting over the set Y_r as

a summation over all such tuples $(i_1, \ldots, i_r)$. Thus we obtain

$$y_r = \sum_{1 \le i_1 < \cdots < i_r \le m} \# \left\{ j : \prod_{p=1}^{r} (\alpha_{i_p} - \beta_j) < 0 \right\}$$

We can then introduce another summation by partitioning the set on the right-hand side depending on which interval A_s each β_j belongs to.

$$y_r = \sum_{1 \le s \le m} \sum_{1 \le i_1 < \cdots < i_r \le m} \# \left\{ j : \beta_j \in A_s \ \wedge \ \prod_{p=1}^{r} (\alpha_{i_p} - \beta_j) < 0 \right\}$$

Then, we eliminate the product symbol by observing that the product $\prod_{p=1}^{r} (\alpha_{i_p} - \beta_j)$ is negative if and only if there are an odd number of p's such that $\alpha_{i_p} - \beta_j < 0$, or equivalently $\alpha_{i_p} < \beta_j$.

$$y_r = \sum_{1 \le s \le m} \sum_{1 \le i_1 < \cdots < i_r \le m} \# \left\{ j : \beta_j \in A_s \ \wedge \ \# \left\{ p : \alpha_{i_p} < \beta_j \right\} \text{ is odd} \right\}$$

Next, we use the fact that the α's are indexed in ascending order, and we note that $\alpha_{i_p} < \beta_j$ if and only if $i_p \le s$, since $\beta_j \in A_s$. Thus

$$y_r = \sum_{1 \le s \le m} \sum_{1 \le i_1 < \cdots < i_s \le m} \# \left\{ j : \beta_j \in A_s \ \wedge \ \# \left\{ p : i_p \le s \right\} \text{ is odd} \right\}$$

Now, note that if $\#\{p : i_p \le s\}$ is even, then the size of the set in the summand is zero. If $\#\{p : i_p \le s\}$ is odd, then the size of the set in the summand equals $\#\{j : \beta_j \in A_s\}$, which is exactly c_s. Hence

$$y_s = \sum_{1 \le s \le m} \sum_{1 \le i_1 < \cdots < i_r \le m} \begin{cases} 0 & \text{if } \# \left\{ p : i_p \le s \right\} \text{ is even} \\ c_s & \text{if } \# \left\{ p : i_p \le s \right\} \text{ is odd} \end{cases}$$

Next, we factor out the c_s and move it outside the innermost summation.

$$\begin{aligned} y_r &= \sum_{1 \le s \le m} \sum_{1 \le i_1 < \cdots < i_r \le m} \left(\begin{cases} 0 & \text{if } \# \left\{ p : i_p \le s \right\} \text{ is even} \\ 1 & \text{if } \# \left\{ p : i_p \le s \right\} \text{ is odd} \end{cases} c_s \right) \\ &= \sum_{1 \le s \le m} \left(\sum_{1 \le i_1 < \cdots < i_r \le m} \begin{cases} 0 & \text{if } \# \left\{ p : i_p \le r \right\} \text{ is even} \\ 1 & \text{if } \# \left\{ p : i_p \le r \right\} \text{ is odd} \end{cases} \right) c_s \end{aligned}$$

Then, we fold the conditional into the summation.

$$y_r = \sum_{1 \le s \le m} \left(\sum_{\substack{1 \le i_1 < \cdots < i_r \le m \\ \#\{p : i_p \le s\} \text{ is odd}}} 1 \right) c_s$$

Since the innermost summand is just a summation of 1's, we can view it as counting the elements of the set $\{(i_1, \ldots, i_r) : 1 \le i_1 < \cdots < i_r \le m \wedge \#\{p : i_p \le s\} \text{ is odd}\}$. Rewriting this, we see that

$$\begin{aligned} \#\{(i_1, \ldots, i_r) : 1 \le i_1 < \cdots < i_r \le m \wedge \#\{p : i_p \le s\} \text{ is odd}\} &= \#\{I \subset [m] : \#I = r \ \wedge \ \#\{i \in I : i \le s\} \text{ is odd}\} \\ &= (T_m)_{rs}. \end{aligned}$$

Hence

$$y_r = \sum_{1 \le s \le m} (T_m)_{rs} c_s.$$

Note this is just the multiplication of the vector $c = (c_1, \ldots, c_m)^T$ by the matrix T_m whose entries are defined as above. Thus

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} (T_m)_{11} & \cdots & (T_m)_{1m} \\ \vdots & & \vdots \\ (T_m)_{m1} & \cdots & (T_m)_{mm} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

that is

$$y = T_m c.$$

Therefore

$$c = \mathrm{GEC}(F, G) \quad \Longrightarrow \quad y = T_m c.$$

□

Now, we will establish the other direction of the correspondence between c and y. To do this, we will give a closed-form expression for T_m, and show that it is invertible.

Lemma 76. *We have $T_m = L_m U_m$ where*

$$\begin{aligned} (L_m)_{rt} &= \binom{m-t}{r-t} \\ (U_m)_{ts} &= (-2)^{t-1} \binom{s}{t}. \end{aligned}$$

In particular, we have that T_m is an invertible matrix.

Example 77 (Running example)**.** *In Example 75, we found that*

$$T_2 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

Observe that

$$\det(T_2) = 0 - 2 = -2 \neq 0,$$

and so T_2 is invertible. △

Proof of Lemma 76. It suffices to prove that $\det(T_m) \neq 0$. We prove it in several stages. First, we establish a recurrence relation for the entries of T_m, and then apply that to decompose T_m into a product of triangular matrices to more easily compute the determinant.

Claim 1: We have

$$(T_m)_{rs} = \begin{cases} s & \text{if } (r,s) \in \{1\} \times \{1,\ldots,m\} \\ \begin{cases} 0 & \text{if } s \text{ is even} \\ 1 & \text{if } s \text{ is odd} \end{cases} & \text{if } (r,s) \in \{m\} \times \{1,\ldots,m\} \\ \begin{cases} \begin{pmatrix} m \\ r \end{pmatrix} & \text{if } r \text{ is odd} \\ 0 & \text{if } r \text{ is even} \end{cases} & \text{if } (r,s) \in \{2,\ldots,m-1\} \times \{m\} \\ (T_{m-1})_{r-1,s} + (T_{m-1})_{r,s} & \text{if } (r,s) \in \{2,\ldots,m-1\} \times \{1,\ldots,m-1\}. \end{cases}$$

Proof of the claim: There are four cases in the above. We will prove them one by one.

1. $(r,s) \in \{1\} \times \{1,\ldots,m\}$.

   By definition, we have

   $$(T_m)_{1s} = \#\{ I \subset [m] : \#I = 1 \wedge \#\{i \in I : i \leq s\} \text{ is odd} \}.$$

   Since we are only considering $\#I = 1$, then $\#\{i \in I : i \leq s\}$ can only be zero or one. If that quantity is zero, then $(T_m)_{1,s}$ is zero, since zero is even. If instead $\#\{i \in I : i \leq s\} = 1$, then the single element of I is between 1 and s, so there are s choices for that element. Thus,

in this case $(T_m)_{1,s} = s$. Together, we have that

$$(T_m)_{1s} = \begin{cases} 0 & \text{if } s \text{ is even.} \\ 1 & \text{if } s \text{ is odd.} \end{cases}$$

2. $(r,s) \in \{m\} \times \{1,\ldots,m\}$.

   By definition, we have

$$\begin{aligned}(T_m)_{ms} &= \#\{\, I \subset [m] : \#I = m \ \wedge\ \#\{i \in I : i \leq s\} \text{ is odd}\,\} \\ &= \begin{cases} 1 & \text{if } s \text{ is odd} \\ 0 & \text{if } s \text{ is even.} \end{cases}\end{aligned}$$

3. $(r,s) \in \{2,\ldots,m-1\} \times \{m\}$.

   By definition of $(T_m)_{rs}$ we have

$$(T_m)_{rm} = \#\{\, I \subset [m] : \#I = r \ \wedge\ \#\{i \in I : i \leq m\} \text{ is odd}\,\}.$$

   But $i \leq m$ for all $i \in I$ trivially, so $\#\{i \in I : i \leq m\} = \#I = r$. Therefore

$$\begin{aligned}(T_m)_{rm} &= \#\{\, I \subset [m] : \#I = r \ \wedge\ \#\{i \in I : i \leq m\} \text{ is odd}\,\} \\ &= \#\{I \subset [m] : \#I = r \quad \wedge \quad r \text{ is odd}\} \\ &= \begin{cases} \#\{I \subset [m] : \#I = r\} & \text{if } r \text{ is odd} \\ \#\{\} & \text{if } r \text{ is even} \end{cases} \\ &= \begin{cases} \begin{pmatrix} m \\ r \end{pmatrix} & \text{if } r \text{ is odd} \\ 0 & \text{if } r \text{ is even.} \end{cases}\end{aligned}$$

4. $(r,s) \in \{2,\ldots,m-1\} \times \{1,\ldots,m-1\}$.

   By definition, we have

$$(T_m)_{rs} = \#\{\, I \subset [m] : \#I = r \ \wedge\ \#\{i \in I : i \leq s\} \text{ is odd}\,\}.$$

   We now partition the set we are counting into disjoint sets with $m \in I$ and $m \notin I$. Then we have

$$(T_m)_{rs} = \qquad \#\{\, I \subset [m] : \#I = r \ \wedge\ m \in I \ \wedge\ \#\{i \in I : i \leq s\} \text{ is odd}\,\}$$

$$+\#\{\, I \subset [m] : \#I = r \ \wedge\ m \notin I \ \wedge\ \#\{i \in I : i \le s\} \text{ is odd}\}.$$

If $m \in I$, then $I \setminus \{m\} \subset [m-1]$. In addition, we still have that $\#\{i \in I : i \le s\}$ is odd, because $s < m$. These are exactly the tuples that are counted by $(T_{m-1})_{r-1,s}$; thus

$$\#\{\, I \subset [m] : \#I = r \ \wedge\ m \in I \ \wedge\ \#\{i \in I : i \le s\} \text{ is odd}\} \ = \ (T_{m-1})_{r-1,s}.$$

On the other hand, if $m \notin I$, then $I \subset [m-1]$ and again $\#\{i \in I : i \le s\}$ remains odd. These tuples are counted by $(T_m)_{r,s}$, and so

$$\#\{\, I \subset [m] : \#I = r \ \wedge\ m \notin I \ \wedge\ \#\{i \in I : i \le s\} \text{ is odd}\} \ = \ (T_{m-1})_{r,s}.$$

Putting those together, we get that

$$(T_m)_{rs} = (T_{m-1})_{r-1,s} + (T_{m-1})_{r,s}.$$

We have proved Claim 1.

Claim 2: We have

$$T_m = L_m U_m$$

where

$$(L_m)_{rt} = \begin{pmatrix} m-t \\ r-t \end{pmatrix} \quad \text{and} \quad (U_m)_{ts} = (-2)^{t-1} \begin{pmatrix} s \\ t \end{pmatrix};$$

that is, for $1 \le r, s \le m$ we have

$$(T_m)_{rs} = \sum_{t=1}^{m} \binom{m-t}{r-t} (-2)^{t-1} \binom{s}{t}.$$

Proof of the claim. It suffices to show

$$(T_m)_{rs} = \sum_{t=1}^{m} (L_m)_{rt} (U_m)_{ts}$$

We proceed by induction on m. Note that when $m = 1$, we have

$$1 = (T_1)_{11} = \binom{1-1}{1-1} (-2)^{1-1} \binom{1}{1} = L_1 U_1.$$

For the induction hypothesis, suppose the claim is true for some $m-1$. We will prove the claim for m. There are four cases in the above. We will prove them one by one.

1. $(r,s) \in \{1\} \times \{1,\ldots,m\}$.

   We have

$$\begin{aligned}
&\sum_{t=1}^{m} (L_m)_{1t}(U_m)_{ts} \\
&= \sum_{t=1}^{m} \binom{m-t}{1-t} (-2)^{t-1} \binom{s}{t} \\
&= \binom{m-1}{1-1} (-2)^{1-1} \binom{s}{1} \qquad \text{since } \binom{m-t}{1-t} = 0 \text{ for } t > 1 \\
&= s \\
&= (T_m)_{1s}
\end{aligned}$$

2. $(r,s) \in \{m\} \times \{1,\ldots,m\}$.

   We have

$$\begin{aligned}
&\sum_{t=1}^{m} (L_m)_{mt}(U_m)_{ts} \\
&= \sum_{t=1}^{m} \begin{pmatrix} m-t \\ m-t \end{pmatrix} (-2)^{t-1} \begin{pmatrix} s \\ t \end{pmatrix} \\
&= \sum_{t=1}^{s} (-2)^{t-1} \begin{pmatrix} s \\ t \end{pmatrix} \qquad \text{since } \begin{pmatrix} s \\ t \end{pmatrix} = 0 \text{ for } t > s \\
&= \frac{1}{-2} \left(-1 + \sum_{t=0}^{s} (1)^{s-t} (-2)^{t} \begin{pmatrix} s \\ t \end{pmatrix} \right) \qquad \text{by the binomial theorem} \\
&= \frac{1}{-2} \left(-1 + (1-2)^{s} \right) \qquad \text{by the binomial theorem} \\
&= \begin{cases} 1 & \text{if } s \text{ is odd} \\ 0 & \text{if } s \text{ is even} \end{cases} \\
&= (T_m)_{ms}
\end{aligned}$$

3. $(r,s) \in \{2,\ldots,m-1\} \times \{m\}$.

   We have

$$\begin{aligned}
\sum_{t=1}^{m} (L_m)_{rt}(U_m)_{tm} &= \sum_{t=1}^{n} \binom{m-t}{r-t} (-2)^{t-1} \binom{m}{t} \\
&= \sum_{t=1}^{m} \frac{(m-t)!}{(m-r)!(r-t)!} (-2)^{t-1} \frac{m!}{(m-t)!t!} \qquad \text{expanding}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{t=1}^{m} \frac{m!}{(m-r)!} \frac{1}{(r-t)!\,t!} (-2)^{t-1} && \text{cancelling terms} \\
&= \sum_{t=1}^{m} \frac{m!}{(m-r)!\,r!} \frac{r!}{(r-t)!\,t!} (-2)^{t-1} && \text{multiplying by } r!/r! \\
&= \sum_{t=1}^{m} \binom{m}{r} \binom{r}{t} (-2)^{t-1} \\
&= \binom{m}{r} \sum_{t=1}^{m} \binom{r}{t} (-2)^{t-1} && \text{rearranging} \\
&= \binom{m}{r} \left(\frac{1}{-2} \left(-1 + \sum_{t=0}^{m} \binom{r}{t} (1)^{r-t} (-2)^{t} \right) \right) && \text{binomial thm.} \\
&= \binom{m}{r} \left(\frac{1}{-2} \left(-1 + (1-2)^{r} \right) \right) && \text{binomial thm.} \\
&= \begin{cases} \binom{m}{r} & \text{if } r \text{ is odd} \\ 0 & \text{if } r \text{ is even} \end{cases} \\
&= (T_m)_{rm}
\end{aligned}
$$

4. $(r,s) \in \{2,\ldots,m-1\} \times \{1,\ldots,m-1\}$.

   Recall that we have

$$(T_m)_{rs} = (T_{m-1})_{r-1,s} + (T_{m-1})_{r,s}.$$

   By the induction hypothesis, we have

$$
\begin{aligned}
(T_m)_{rs} &= \sum_{t=1}^{m-1} (L_{m-1})_{r-1,t} (U_{m-1})_{t,s} + \sum_{t=1}^{m-1} (L_{m-1})_{r,t} (U_{m-1})_{t,s} && \text{by def. of } L \text{ and } U \\
&= \sum_{t=1}^{m-1} (L_{m-1})_{r-1,t} (U_{m-1})_{t,s} + (L_{m-1})_{r,t} (U_{m-1})_{t,s} && \text{combining sums} \\
&= \sum_{t=1}^{m-1} ((L_{m-1})_{r-1,t} + (L_{m-1})_{r,t}) (U_{m-1})_{t,s} && \text{factoring} \\
&= \sum_{t=1}^{m-1} \left(\binom{m-1-t}{r-1-t} + \binom{m-1-t}{r-t} \right) (U_{m-1})_{t,s} && \text{by def. of } L \\
&= \sum_{t=1}^{m-1} \binom{m-t}{r-t} (U_{m-1})_{t,s} && \text{Pascal's rule} \\
&= \sum_{t=1}^{m-1} (L_m)_{r,t} (U_{m-1})_{t,s} && \text{by def. of } L.
\end{aligned}
$$

But $(U_{m-1})_{t,s}$ does not depend on m. Thus $(U_{m-1})_{t,s} = (U_m)_{t,s}$. Hence

$$\begin{aligned}(T_m)_{rs} &= \sum_{t=1}^{m-1} (L_m)_{r,t}(U_m)_{t,s} \\ &= \sum_{t=1}^{m-1} (L_m)_{r,t}(U_m)_{t,s} + \underbrace{\binom{m-m}{r-m}(-2)^{r-1}\binom{s}{m}}_{=0 \text{ since } s<m} \\ &= \sum_{t=1}^{m} (L_m)_{r,t}(U_m)_{t,s}.\end{aligned}$$

We have proved Claim 2.

Now note that L_m is lower triangular, since if $s > r$ then $r - s < 0$ and so $\binom{m-s}{r-s} = 0$. Similarly, U_m is upper triangular, since if $r > s$ then $\binom{s}{r} = 0$. Further, we have

$$\det(L_m) = \prod_{t=1}^{m} (L_m)_{tt} = \prod_{t=1}^{m} \binom{m-t}{t-t} = \prod_{t=1}^{m} \binom{m-t}{0} = \prod_{t=1}^{m} 1 = 1,$$

and

$$\det(U_m) = \prod_{t=1}^{m} (U_m)_{tt} = \prod_{t=1}^{m} (-2)^{t-1}\binom{t}{t} = \prod_{t=1}^{m} (-2)^{t-1} = (-2)^{0+1+\cdots+m-1} = (-2)^{\binom{m}{2}}.$$

Thus we have

$$\det(T_m) = \det(L_m)\det(U_m) = (-2)^{\binom{m}{2}} \neq 0.$$

Hence T_m is invertible. □

Next, we establish the relationship between C_{sym} and the matrix T_m; in particular, they are inverses of each other.

Proposition 78. *We have that* $C_{\text{sym}} = T_m^{-1}$.

Proof. Recall from Defintion 62 that $C_{\text{sym}} = V_m W_m$ where

$$(V_m)_{rt} = \frac{(-1)^{r-1}}{2^{t-1}}\binom{t}{r} \quad \text{and} \quad (W_m)_{ts} = (-1)^{t+s}\binom{m-s}{m-t}.$$

Recall again from Lemma 76 that $T_m = L_m U_m$ where

$$(L_m)_{rt} = \begin{pmatrix} m-t \\ r-t \end{pmatrix} \quad \text{and} \quad (U_m)_{ts} = (-2)^{t-1}\begin{pmatrix} s \\ t \end{pmatrix}.$$

Note that

$$T_m^{-1} = (L_m U_m)^{-1} = U_m^{-1} L_m^{-1};$$

hence, to show that $C_{\text{sym}} = T_m^{-1}$, it suffices to show that

1. $V_m U_m = I_m$
2. $L_m W_m = I_m$

where I_m denotes the $m \times m$ identity matrix.

1. $V_m U_m = I_m$: We have that the (r,s)-entry of $V_m U_m$ is

$$\begin{aligned}
(V_m U_m)_{rs} &= \sum_t V_{rt} U_{ts} \\
&= \sum_t \frac{(-1)^{r-1}}{2^{t-1}} \binom{t}{r} (-2)^{t-1} \binom{s}{t} \\
&= \sum_t (-1)^{r+t} \binom{t}{r} \binom{s}{t} \\
&= \sum_u (-1)^u \binom{u+r}{r} \binom{s}{u+r} && \text{reindexing with } t = u+r \\
&= \sum_u (-1)^u \frac{(u+r)!}{r!u!} \frac{s!}{(u+r)!(s-u-r)!} \\
&= \sum_u (-1)^u \frac{s!}{r!u!(s-u-r)!} \cdot \frac{(s-r)!}{(s-r)!} && \text{by multiplying by } \frac{(s-r)!}{(s-r)!} \\
&= \sum_u (-1)^u \frac{s!}{r!(s-r)!} \frac{(s-r)!}{u!(s-r-u)!} && \text{grouping terms} \\
&= \sum_u (-1)^u \binom{s}{r} \binom{s-r}{u} \\
&= \binom{s}{r} \sum_u (-1)^u \binom{s-r}{u} (1)^{s-r-u} \\
&= \binom{s}{r} (1-1)^{s-r} && \text{binomial thm.} \\
&= \begin{cases} 1 & \text{if } s = r \\ 0 & \text{if } s \neq r \end{cases} \\
&= \delta_{r,s}.
\end{aligned}$$

Since the (r,s)-entry of $V_m U_m$ is $\delta_{r,s}$, it follows that $V_m U_m = I_m$.

2. $L_m W_m = I_m$: We have that the (r,s)-entry of $L_m W_m$ is

$$\begin{aligned}
(L_m W_m)_{rs} &= \sum_t (L_m)_{rt}(W_m)_{ts} \\
&= \sum_t \binom{m-t}{r-t}(-1)^{t+s}\binom{m-s}{m-t} \\
&= \sum_u \binom{m-r+u}{u}(-1)^{r-u+s}\binom{m-s}{m-r+u} && \text{reindexing } t = r-u \\
&= \sum_u (-1)^{r-u+s}\frac{(m-r+u)!}{u!(m-r)!}\frac{(m-s)!}{(m-r+u)!(r-s-u)!} \\
&= \sum_u (-1)^{r-u+s}\left(\frac{(m-s)!}{u!(m-r)!(r-s-u)!}\right)\cdot\frac{(r-s)!}{(r-s)!} && \text{multiplying by } \frac{(r-s)!}{(r-s)!} \\
&= \sum_u (-1)^{r-u+s}\frac{(r-s)!}{u!(r-s-u)!}\frac{(m-s)!}{(m-r)!(r-s)!} && \text{grouping terms} \\
&= \sum_u (-1)^{r-u+s}\binom{r-s}{u}\binom{m-s}{m-r} \\
&= (-1)^{r+s}\binom{m-s}{m-r}\sum_u (-1)^u\binom{r-s}{u}(1)^{r-s-u} \\
&= (-1)^{r+s}\binom{m-s}{m-r}(1-1)^{r-s} && \text{binomial thm.} \\
&= \begin{cases} 1 & \text{if } r = s \\ 0 & \text{if } r \neq s \end{cases} \\
&= \delta_{r,s}.
\end{aligned}$$

Since the (r,s)-entry of $L_m W_m$ is $\delta_{r,s}$, it follows that $L_m W_m = I_m$.

Hence, we have shown that $V_m U_m = I_m$ and $L_m W_m = I_m$, and therefore

$$\begin{aligned}
C_{\text{sym}} &= V_m W_m \\
&= U_m^{-1} L_m^{-1} \\
&= (L_m U_m)^{-1} \\
&= T_m^{-1}.
\end{aligned}$$

□

Now, we can summarize Lemmas 74 and 76 and Proposition 78 into the following lemma.

Lemma 79. *We have*

$$\mathrm{GEC}(F,G) \;=\; C_{\mathrm{sym}}\; y.$$

Proof. Lemma 76 shows that T_m is invertible, hence T_m is a 1-1 linear map. Proposition 78 shows that $C_{\mathrm{sym}} = T_m^{-1}$. Therefore, we have that

$$T_m^{-1} y = C_{\mathrm{sym}}\; y = c \implies \mathrm{GEC}(F,G) = c\,.$$

Together with Lemma 74, this means that $\mathrm{GEC}(F,G) = T_m^{-1} y = C_{\mathrm{sym}}\; y$, and the lemma is proved. □

At this point, we have found an equivalent condition for $\mathrm{GEC}(F,G) = c$. However, this condition still contains quantifiers, as the vector y (via the h polynomials) is still computed using the eigenvalues of F and G. The remainder of the derivation will focus on solving this issue by providing a way to count the positive roots of the h polynomials without referring to the eigenvalues.

Recall that, for each $r \in [m]$, the polynomial h_r is symmetric with only real roots. This is the crucial fact which allows us to construct another polynomial D_r involving only the *parameters* of F and G rather than their *eigenvalues*. This will give us a quantifier-free condition.

Proposition 80. *Let $\alpha = (\alpha_1, \ldots, \alpha_m)$, and similarly for β, a, and b. For each $r \in [m]$, there exists a polynomial $D_r \in \mathbb{R}[a,b][x]$ such that*

$$D_r(a,b,x) \;=\; h_r(\alpha,\beta,x).$$

Proof. Let $r \in [m]$ be arbitrary but fixed. For the purposes of this proof, we view h_r as being a polynomial in the variables

$$\begin{aligned}
&\alpha = (\alpha_1, \ldots, \alpha_m),\\
&\beta = (\beta_1, \ldots, \beta_n),\\
&x,
\end{aligned}$$

with real number coefficients. Recall from Definition 64 that

$$h_r = \prod_{\substack{I \subset [m],\, \#I = r \\ j \in [n]}} \left(x + \prod_{i \in I} (\alpha_i - \beta_j) \right).$$

Note that by construction h_r is symmetric in both α and β. We will now apply the Fundamental Theorem of Symmetric Polynomials. For this, let $e_k(\alpha)$ denote the k-th elementary symmetric

polynomial in the variables $\alpha = (\alpha_1, \ldots, \alpha_m)$; that is,

$$e_k(\alpha) = \sum_{1 \le i_1 < \cdots < i_k \le m} \alpha_{i_1} \cdots \alpha_{i_k}.$$

Let $e(\alpha)$ denote the list of all elementary symmetric polynomials; i.e.,

$$e(\alpha) = (e_1(\alpha), \ldots, e_m(\alpha)),$$

and respectively for $e_k(\beta)$ and $e(\beta)$.

Recall that by the Fundamental Theorem of Symmetric Polynomials, since h_r is symmetric in α and β this means that there exists a polynomial

$$D_r \in \mathbb{Z}[y_1, \ldots, y_m,\ z_1, \ldots, z_n,\ x]$$

such that

$$D_r(e(\alpha),\ e(\beta),\ x) = h_r(\alpha, \beta, x).$$

Recall from Definition 64 that $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$ were the coefficients of the respective characteristic polynomials f and g of F and G labelled so that

$$\begin{aligned} f &= \det(x I_m - F) = x^m - a_1 x^{m-1} + a_2 x^{m-2} - \cdots + (-1)^m a_m x^0 \\ g &= \det(x I_n - G) = x^n - b_1 x^{n-1} + b_2 x^{n-2} - \cdots + (-1)^n b_n x^0. \end{aligned}$$

With this labelling of the coefficients of f and g, we have that

$$\begin{aligned} a_1 &= \sum_{1 \le i_1 \le m} \alpha_{i_1} = \alpha_1 + \cdots + \alpha_m = e_1(\alpha) \\ a_2 &= \sum_{1 \le i_1 < i_2 \le m} \alpha_{i_1} \alpha_{i_2} = e_2(\alpha) \\ &\vdots \\ a_m &= \sum_{1 \le i_1 < \cdots < i_m \le m} \alpha_{i_1} \cdots \alpha_{i_m} = \alpha_1 \cdots \alpha_m = e_m(\alpha). \end{aligned}$$

Similarly

$$\begin{aligned} b_1 &= e_1(\beta) \\ &\vdots \\ b_n &= e_n(\beta). \end{aligned}$$

Hence, we have that

$$\begin{aligned} h_r(\alpha,\beta,x) &= D_r(\underbrace{e_1(\alpha)}_{a_1},\ldots,\underbrace{e_m(\alpha)}_{a_m},\ \underbrace{e_1(\beta)}_{b_1},\ldots,\underbrace{e_n(\beta)}_{b_n},\ x) \\ &= D_r(a,b,x) \in \mathbb{R}[a,b][x]. \end{aligned}$$

□

With that, we are ready to prove Theorem 66 (Symmetric polynomials approach for generic pairs of matrices).

Proof of Theorem 66.

Let $y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = T_m c$. For each $r \in [m]$, by Proposition 80 there exists $D_r \in \mathbb{R}[a,b][x]$ such that

$$D_r(a,b,x) = h_r(\alpha,\beta,x).$$

Recall that, by definition of y, we have

$$\begin{aligned} y_r &= \#\text{ positive roots of } h_r(x), \text{ counting multiplicity} \\ &= \#\text{ positive roots of } D_r(x), \text{ counting multiplicity} && \text{by Proposition 80} \\ &= v(D_r(x)) && \text{by Corollary 7} \end{aligned}$$

By the above and Lemma 79, we have that

$$\begin{aligned} \mathrm{GEC}(F,G) &= T_m^{-1} y \\ &= T_m^{-1} \begin{bmatrix} v(D_1) \\ \vdots \\ v(D_m) \end{bmatrix} \\ &= C_{\mathrm{sym}}\, A_{\mathrm{sym}}(F,G). \end{aligned}$$

Thus Theorem 66 is proved. □

Chapter 3

Arbitrary Pairs of Symmetric Matrices

In this chapter, we will extend the results from the previous chapter on generic pairs of matrices to arbitrary pairs of matrices.

3.1 Problem Statement

In this section, we will state the problem precisely for arbitrary pairs of matrices.

Defining eigenvalue configuration for arbitrary pairs of matrices

In Section 2.1, we gave a definition (Definition 22) of eigenvalue configuration for generic pairs of matrices F and G. Next, the challenge is to extend this definition to pairs of matrices which are not generic (that is, matrices which share eigenvalues) in such a way that Theorems 30 and 66 still hold for input matrices which are not generic.

In the generic eigenvalue definition, we count eigenvalues of G which lie between the *open* intervals spanned by adjacent eigenvalues of F; thus, if F and G share eigenvalues, some eigenvalues of G will not be counted. Hence, we need to expand this definition for matrices which share eigenvalues.

The key observation we can make in this case is that if the pair F and G is not generic, then sufficiently small perturbation of their eigenvalues results in a new pair of matrices which is generic. One natural idea is then to define the eigenvalue configuration for arbitrary (i.e. non-generic) pairs of real symmetric matrices as the average of the eigenvalue configurations of the generic "neighbor" configurations, those being the configurations that can be obtained by slightly perturbing the eigenvalues.

Definition 81 (Eigenvalue configuration of arbitrary matrices)**.** *Let F and G be real symmetric matrices. The eigenvalue configuration of F and G, written as* $\mathrm{EC}(F,G)$*, is defined as*

$$\mathrm{EC}(F,G) = \frac{1}{2^m} \sum_{d \in \{-1,1\}^m} \mathrm{GEC}(F_d, G)$$

where

$$\begin{aligned} \varepsilon &= \text{“small enough” positive number (see Remark below)} \\ F_d &= \text{matrix with eigenvalues } \alpha_i + \varepsilon d_i \text{ for } i = 1, \ldots, m. \end{aligned}$$

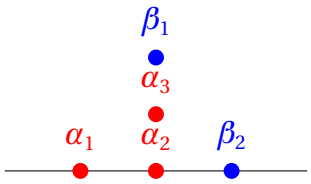
Remark 82. *Some notes on Definition 81:*

1. *The number ε is “small enough” if no eigenvalue of F is perturbed far enough to cross over the next or previous eigenvalue. We will explicitly construct a suitable value for ε in Lemma 86 in the next section.*
2. *Note that the pairs F_d and G are generic for all $d \in \{-1,1\}^m$.*
3. *If the pair F and G is generic, then* $\mathrm{GEC}(F_d, G) = \mathrm{GEC}(F,G)$*. This is because ε is chosen so that eigenvalues of F do not move far enough to cross any other eigenvalues; hence, the arrangement of the eigenvalues of generic matrices is not disturbed. Therefore* $\mathrm{EC}(F,G) = \mathrm{GEC}(F,G)$*, and so the generic eigenvalue configuration definition (Definition 22) is a special case of this one.*

Example 83 (Eigenvalue configuration of arbitrary matrices)**.** *Let $F \in \mathbb{R}^{3\times 3}$ and $G \in \mathbb{R}^{2\times 2}$ be real symmetric matrices such that their corresponding eigenvalues are*

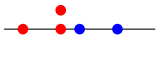
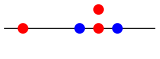
$$\alpha = (0,1,1), \qquad \beta = (1,2).$$

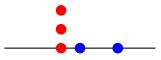
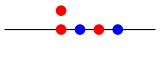
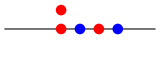
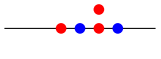
Pictorially, the eigenvalue arrangement is as shown in the diagram below.



Let $\varepsilon = 0.5$ (which can be verified to satisfy Definition 81). Now, for each $d \in \{-1,1\}^3$, we compute the eigenvalue configuration of F_d and G in the tables below. Note that each pair F_d

and G is generic.

d	*Picture*	$\mathrm{GEC}(F_d, G)$
$---$		$\begin{bmatrix}0\\0\\2\end{bmatrix}$
$--+$		$\begin{bmatrix}0\\1\\1\end{bmatrix}$
$-+-$		$\begin{bmatrix}0\\1\\1\end{bmatrix}$
$-++$		$\begin{bmatrix}1\\0\\1\end{bmatrix}$

d	*Picture*	$\mathrm{GEC}(F_d, G)$
$+--$		$\begin{bmatrix}0\\0\\2\end{bmatrix}$
$+-+$		$\begin{bmatrix}0\\1\\1\end{bmatrix}$
$++-$		$\begin{bmatrix}0\\1\\1\end{bmatrix}$
$+++$		$\begin{bmatrix}1\\0\\1\end{bmatrix}$

Then, applying Definition 81, we have

$$\mathrm{EC}(F,G) = \frac{1}{2^m}\sum_{d\in\{-1,1\}^m}\mathrm{GEC}(F_d,G) = \frac{1}{2^3}\left(2\begin{bmatrix}0\\0\\2\end{bmatrix}+4\begin{bmatrix}0\\1\\1\end{bmatrix}+2\begin{bmatrix}1\\0\\1\end{bmatrix}\right) = \frac{1}{8}\begin{bmatrix}2\\4\\10\end{bmatrix} = \begin{bmatrix}1/4\\1/2\\5/4.\end{bmatrix}.$$

△

Stating the problem

Now, we state the eigenvalue configuration problem for arbitrary pairs of matrices.

Problem 84 (Arbitrary eigenvalue configuration problem)**.**

In: $F \in \mathbb{R}[p]^{m\times m}$ *and* $G \in \mathbb{R}[p]^{n\times n}$*, arbitrary pair of symmetric matrices where* p *is a finite set of parameters.*

$c \in \mathbb{Q}^m$*, an eigenvalue configuration.*

Out: *a quantifier-free condition on* p *such that* $c = \mathrm{EC}(F,G)$.

All the considerations discussed in Section 2.1 still apply.

3.2 Signature Approach

In this section, we will generalize Theorem 30 (Signature approach for generic pairs of matrices) from the previous chapter to arbitrary pairs of matrices.

3.2.1 Main Theorem

In this subsection, we state the theorem for arbitrary pairs of matrices.

Theorem 85 (Signature approach for arbitrary pairs of matrices)**.** *Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be **arbitrary** real symmetric matrices. We have*

$$\mathrm{EC}(F,G) \;=\; C_{\mathrm{sig}}\; A_{\mathrm{sig}}(F,G),$$

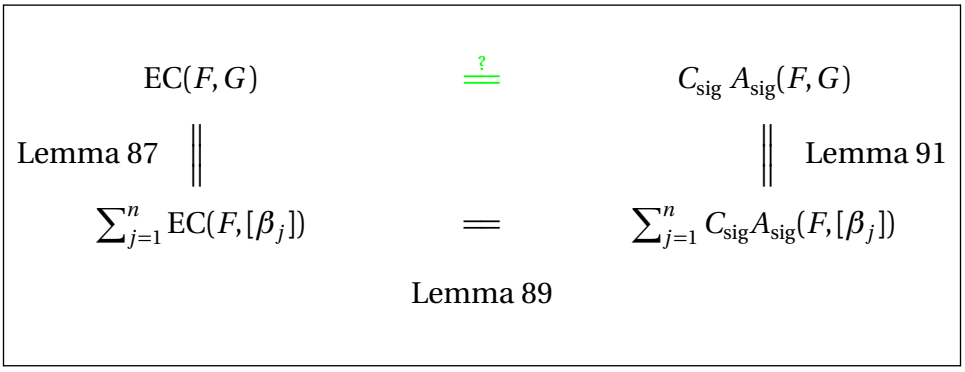
where C_{sig} and A_{sig} are defined as in Definitions 26 and 28, respectively.

3.2.2 Proof

In this subsection, we will prove the above theorem (Theorem 85). We take the following approach.

1. In Lemma 87, we reduce to the case where $n = 1$; that is, where G is a 1×1 matrix.
2. In Lemma 89, we prove Theorem 85 for the case $n = 1$.
3. Finally, in Lemma 91, we build back up from the $n = 1$ case to arbitrary n. The proof is then complete.

Our strategy is illustrated by the following diagram.



First, let us revisit Definition 81. For real symmetric matrices F and G, we have

$$\mathrm{EC}(F,G) = \frac{1}{2^m} \sum_{d \in \{-1,1\}^m} \mathrm{GEC}(F_d, G)$$

where

$$\begin{aligned} \varepsilon &= \text{ “small enough” positive number} \\ F_d &= \text{matrix with eigenvalues } \alpha_i + \varepsilon d_i \text{ for } i = 1, \ldots, m. \end{aligned}$$

First, we show that Definition 81 can always be satisfied. To do this, we explicitly construct ε and F_d.

Lemma 86 (Explicit construction of Definition 81)**.** *Let F and G be real symmetric matrices. Let*

$$\begin{aligned} \varepsilon &= \frac{1}{2} \min_{i,j} \begin{cases} 1 & \text{if } \alpha_i = \beta_j \\ |\alpha_i - \beta_j| & \text{if } \alpha_i \neq \beta_j \end{cases} \\ F_d &= F + A\,\mathrm{diag}(\varepsilon d) A^T \qquad\qquad \text{for } d \in \{-1,1\}^m \\ A &= \text{ orthonormal matrix such that } F = A\,\mathrm{diag}(\alpha_1, \ldots, \alpha_m) A^T. \end{aligned}$$

Then Definition 81 is satisfied.

Proof. First, note that since F is real symmetric, it admits an orthogonal eigendecomposition. Hence there exists an orthogonal matrix A so that

$$F = A\,\mathrm{diag}(\alpha) A^T.$$

Then, for $d \in \{-1,1\}^m$, we have

$$\begin{aligned} F_d &= F + A\,\mathrm{diag}(d) A^T \\ &= A\,\mathrm{diag}(\alpha) A^T + A\,\mathrm{diag}(\varepsilon d) A^T \\ &= A\big(\mathrm{diag}(\alpha) + \mathrm{diag}(\varepsilon d)\big) A^T. \end{aligned}$$

Thus the eigenvalues of F_d are exactly $\alpha_i + \varepsilon d_i$, which satisfies Definition 81.

To conclude, let

$$\varepsilon = \frac{1}{2} \min_{i,j} \begin{cases} 1 & \text{if } \alpha_i = \beta_j \\ |\alpha_i - \beta_j| & \text{if } \alpha_i \neq \beta_j \end{cases}.$$

It remains to show that this ε satisfies Definition 81. By construction, this ε is strictly less than

the minimum distance between α_i and β_j over all i and j. Since the eigenvalues of F_d are exactly $\alpha_i + \varepsilon d_i = \alpha_i \pm \varepsilon$, this means that no eigenvalue of F travels far enough to cross over an eigenvalue of G. Hence, the condition of Definition 81 is satisfied, and we are done. □

Having established that Definition 81 can always be satisfied, we can now proceed.

Lemma 87 (Decompose EC into eigenvalues)**.** *For all real symmetric F and G we have*

$$\mathrm{EC}(F,G) = \sum_{j=1}^{n} \mathrm{EC}(F,[\beta_j]).$$

Example 88 (Running)**.** *Recall from the running example that the matrix F has the eigenvalues $(\alpha_1,\alpha_2,\alpha_3) = (0,1,1)$ and G has eigenvalues $(\beta_1,\beta_2) = (1,2)$. We computed in Example 83 that*

$$\mathrm{EC}(F,G) = \begin{bmatrix} 1/4 \\ 1/2 \\ 5/4 \end{bmatrix}.$$

On the other hand, by applying Definition 81 directly, we have that

$$\mathrm{EC}(F,[\beta_1]) = \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}$$

$$\mathrm{EC}(F,[\beta_2]) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

and

$$\begin{aligned} \mathrm{EC}(F,[\beta_1]) + \mathrm{EC}(F,[\beta_2]) &= \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/2 \\ 5/4 \end{bmatrix} \\ &= \mathrm{EC}(F,G). \end{aligned}$$

△

Proof of Lemma 87. We break the proof into two cases.

Case 1: F and G are generic. Recall Definition 22:

$$\mathrm{EC}(F,G) = \mathrm{GEC}(F,G) = (c_1,\ldots,c_m) \text{ where } c_t = \#\{j : \beta_j \in A_t\}.$$

Note that

$$\begin{aligned} c_t &= \#\{j : \beta_j \in A_t\} \\ &= \sum_{j=1}^{n} \begin{cases} 1 & \text{if } \beta_j \in A_t \\ 0 & \text{else,} \end{cases} \end{aligned}$$

and the summand is exactly the t-th component of $\mathrm{EC}(F,[\beta_j])$.

Case 2: F and G are not generic. Recall Definition 81:

$$\mathrm{EC}(F,G) = \frac{1}{2^m} \sum_{d\in\{-1,1\}^m} \mathrm{GEC}(F_d, G).$$

Note that the pair F_d and G is generic. Then we have

$$\begin{aligned} \mathrm{EC}(F,G) &= \frac{1}{2^m} \sum_{d\in\{-1,1\}^m} \mathrm{GEC}(F_d, G) && \text{by Definition 81} \\ &= \frac{1}{2^m} \sum_{d\in\{-1,1\}^m} \sum_{j=1}^{n} \mathrm{GEC}(F_d, [\beta_j]) && \text{by Case 1} \\ &= \sum_{j=1}^{n} \underbrace{\frac{1}{2^m} \sum_{d\in\{-1,1\}^m} \mathrm{GEC}(F_d, [\beta_j])}_{=\mathrm{EC}(F,[\beta_j]) \text{ by Definition 81}} \\ &= \frac{1}{2^m} \sum_{j=1}^{n} \mathrm{EC}(F,[\beta_j]). \end{aligned}$$

□

Lemma 89 ($n = 1$)**.** *Let F be an arbitrary real symmetric matrix and let $\beta \in \mathbb{R}$. Then*

$$\mathrm{EC}(F,[\beta]) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta]).$$

Example 90 (Running)**.** *Let $F = \mathrm{diag}(0,1,1)$ and let $\beta = 1$. In Example 88, we found via Definition 81 that*

$$\mathrm{EC}(F,[\beta]) = \begin{bmatrix} 1/2 \\ 1/4 \\ 1/2 \end{bmatrix}.$$

On the other hand, using Definitions 26 and 28 directly, we have

$$C_{\text{sig}}A_{\text{sig}}(F,[\beta]) = VH^{-1}\begin{bmatrix}\sigma(f_{000}(G))\\ \sigma(f_{001}(G))\\ \sigma(f_{010}(G))\\ \sigma(f_{011}(G))\\ \sigma(f_{100}(G))\\ \sigma(f_{101}(G))\\ \sigma(f_{110}(G))\\ \sigma(f_{111}(G))\end{bmatrix}$$

$$= \begin{bmatrix}0&0&0&0&1&1&1&0\\ 1&1&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}1&1&1&1&1&1&1&1\\ -1&1&-1&1&-1&1&-1&1\\ -1&-1&1&1&-1&-1&1&1\\ 1&-1&-1&1&1&-1&-1&1\\ -1&-1&-1&-1&1&1&1&1\\ 1&-1&1&-1&-1&1&-1&1\\ 1&1&-1&-1&-1&-1&1&1\\ -1&1&1&-1&1&-1&-1&1\end{bmatrix}^{-1}\begin{bmatrix}1\\1\\0\\0\\0\\0\\0\\0\end{bmatrix}$$

$$= \begin{bmatrix}1/4\\ 1/2\\ 1/4\end{bmatrix}.$$

△

Proof of Lemma 89. Note that if F and G are generic (i.e. β is not an eigenvalue of F), then Theorem 30 applies and this lemma immediately follows as a special case.

Thus, for the remainder of this proof, suppose that F and G are not generic. In other words, assume that the eigenvalues of F are

$$\alpha_1 \leq \cdots \leq \alpha_l < \alpha_{l+1} = \cdots = \alpha_{l+\mu} < \alpha_{l+\mu+1} \leq \cdots \leq \alpha_m,$$

where $\alpha_{l+1} = \cdots = \alpha_{l+\mu} = \beta$. That is, suppose that β is an eigenvalue of F of multiplicity μ, starting at index $l+1$ in the list of eigenvalues of F (repeated with multiplicity).

To prove the lemma, we will show that

$$\text{EC}(F,[\beta]) = C_{\text{sig}}\, A_{\text{sig}}(F,[\beta]) = \frac{1}{2^\mu}\sum_{k=0}^{\mu}\binom{\mu}{k}e_{l+\mu-k}. \tag{3.1}$$

We will first simplify the left-hand side; i.e., we will simplify

$$\mathrm{EC}(F,[\beta])$$

by opening up Definition 81. We have

$$\begin{aligned}
\mathrm{EC}(F,[\beta]) &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} \mathrm{GEC}(F_d,G) && \text{by Definition 81}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F_d,[\beta]) && \text{by Theorem 82, since } F_d \text{ and } [\beta] \text{ are generic}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} V\sum_{j=1}^{1}\gamma_{d,j} && \text{by Lemma 55}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} e_{m-v(\mathrm{sseq}_{f_d}(\beta),+)} && \text{by Claim 2.10 in Lemma 61, with } f_d=\det(xI_m-F_d).
\end{aligned}$$

Next, we will examine the term $v(\mathrm{sseq}_{f_d}(\beta),+)$. Recall that by construction, the eigenvalues of F_d are $\alpha_i+\varepsilon d_i$ for $i\in\{1,\ldots,m\}$. Recall from Lemma 43 that

$$v(\mathrm{sseq}_{f_d}(\beta),+) = \#\{i : \alpha_i+\varepsilon d_i > \beta\}.$$

Note further that $\alpha_i+\varepsilon d_i>\beta$ if and only if either

1. $\alpha_i>\beta$,
2. or $\alpha_i=\beta$ and $d_i>0$.

This means that

$$\begin{aligned}
\#\{i:\alpha_i+\varepsilon d_i>\beta\} &= \#\{i:\alpha_i>\beta\} + \#\{i:d_i>0\ \wedge\ \alpha_i=\beta\}\\
&= m-(l+\mu) \ + \#\{i:d_i>0\ \wedge\ \alpha_i=\beta\}\\
&= m-l-\mu \quad + \#\{i:d_i>0\ \wedge\ i\in\{l+1,\ldots,l+\mu\}\}.
\end{aligned}$$

Thus, continuing from above we have

$$\begin{aligned}
\mathrm{EC}(F,[\beta]) &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} e_{m-v(\mathrm{sseq}_{f_d}(\alpha),+)}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} e_{m-(m-l-\mu+\#\{i:d_i>0\wedge i\in\{l+1,\ldots,l+\mu\}\})} && \text{by substituting the above}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} e_{l+\mu-\#\{i:d_i>0\wedge i\in\{l+1,\ldots,l+\mu\}\}}.
\end{aligned}$$

Next, we will partition the sum by the values of $\#\{i : d_i > 0 \wedge i \in \{l+1,\dots,l+\mu\}\}$. The possible values for this term are 0 through μ inclusive. Further, for fixed $k \in \{0,\dots,k\}$, the number of $d \in \{-1,1\}^m$ so that $\#\{i : d_i > 0 \wedge i \in \{l+1,\dots,l+\mu\}\} = k$ is exactly $2^{m-\mu}\binom{\mu}{k}$ (i.e. by choosing the arbitrary $m-\mu$ entries, then independently choosing k of the μ remaining positive entries). Therefore

$$\begin{aligned}
\mathrm{EC}(F,[\beta]) &= \frac{1}{2^m} \sum_{d\in\{-1,1\}^m} e_{l+\mu-\#\{i:d_i>0 \wedge i\in\{l+1,\dots,l+\mu\}\}} \\
&= \frac{1}{2^m} \sum_{k=0}^{\mu} 2^{m-\mu} \binom{\mu}{k} e_{l+\mu-k} \\
&= \frac{1}{2^\mu} \sum_{k=0}^{\mu} \binom{\mu}{k} e_{l+\mu-k}.
\end{aligned}$$

This is the same as (3.1), so we are done with the left-hand side.

Now, we will simplify the right-hand side by rewriting the expression

$$C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta]).$$

Recall that by Lemma 55 we have

$$C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta]) = V\gamma$$

where

$$\gamma = \frac{1}{\#S} \sum_{s\in S} e_s$$

where $S = \left\{ s \in \{-,+\}^m : \mathrm{sseq}(\beta) \in \mathrm{cl}(s) \right\}$. Hence we have

$$\begin{aligned}
C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta]) = V\gamma &= V\left(\frac{1}{\#S} \sum_{s\in S} e_s \right) \\
&= \frac{1}{\#S} \begin{bmatrix} \#\{s\in S : v(s,+) = m-1\} \\ \#\{s\in S : v(s,+) = m-2\} \\ \vdots \\ \#\{s\in S : v(s,+) = m-m\} \end{bmatrix}.
\end{aligned}$$

Recall the proof of Lemma 55 for more detailed reasoning on the last step.

Next, fix $t \in \{1, \ldots, m\}$. We will consider the t-th element of the above vector; that is,

$$\frac{1}{\#S}\{s \in S : v(s,+) = m - t\}. \tag{3.2}$$

Recall the definition of S from Lemma 55:

$$S = \{s \in \{-,+\}^m : \text{sseq}_f(\beta) \in \text{cl}(s)\}.$$

Note that since β is a root of f with multiplicity μ, we have that

$$\text{sseq}_f(\beta) = (\underbrace{0, \ldots, 0}_{\mu \text{ zeros}}, \omega),$$

where $\omega \in \{-, 0, +\}^{m-\mu}$. Note that $\omega_1 \neq 0$, because by assumption β has multiplicity exactly μ. Let z denote the number of zeros in ω. With this, we have that

$$S = \{s \in \{-,+\}^m : s \text{ and } \text{sseq}_f(\beta) \text{ differ by one or more zeros}\}.$$

In other words, the set S is the set of sign sequences of length m which can be obtained from $\text{sseq}_f(\beta)$ by replacing each 0 with either $+$ or $-$. Since there are $\mu + z$ zeros in $\text{sseq}_f(\beta)$, this means that S has $2^{\mu+z}$ elements.

Further, by the same reasoning used in the proof of Lemma 61, we have the following properties about ω:

1. ω does not contain consecutive zeros; this is because any consecutive zeros in ω would indicate a multiple root which propagates all the way upward (see Lemma 60), which contradicts the fact that $\omega_1 \neq 0$.

2. Zeros only appear in ω as $(\ldots, -, 0, +, \ldots)$ or $(\ldots, +, 0, -, \ldots)$, because of the assumption that f has only real roots.

Now, let $s \in S$. We will now rewrite $v(s,+) = v(s_1, \ldots, s_m, +)$. Note

$$\begin{aligned}
&v(s_1, \ldots, s_m, +) \\
&\qquad \Big\| \quad \text{trivially, by splitting at } s_{\mu+1} \\
&v(s_1, \ldots, s_\mu, s_{\mu+1}) + v(s_{\mu+1}, \ldots, s_m, +) \\
&\qquad \Big\| \quad \text{since } \omega_1 \neq 0, \text{ so } s_{\mu+1} = \omega_1
\end{aligned}$$

$$
\begin{array}{l}
v(s_1,\ldots,s_\mu,\omega_1)+v(\omega_1,\ldots,s_m,+) \\
\qquad \Big\| \quad \text{by above points 1 and 2} \\
v(s_1,\ldots,s_\mu,\omega_1)+v(\omega,+) \\
\qquad \Big\| \quad \text{since } v(\omega,+)=v(\text{sseq}(\beta),+) \\
v(s_1,\ldots,s_\mu,\omega_1)+v(\text{sseq}(\beta),+) \\
\qquad \Big\| \quad \text{since } v(\text{sseq}(\beta),+)=\#\{i:\alpha_i>\beta\} \text{ by Lemma 43} \\
v(s_1,\ldots,s_\mu,\omega_1)+\#\{i:\alpha_i>\beta\} \\
\qquad \Big\| \quad \text{by the hypotheses of the current lemma} \\
v(s_1,\ldots,s_\mu,\omega_1)+m-l-\mu.
\end{array}
$$

Now, substituting the above into (3.2) gives

$$
\begin{aligned}
\frac{1}{\#S}\#\{s\in S:v(s,+)=m-t\} &= \frac{1}{2^{\mu+z}}\#\{s\in S:v(s,+)=m-t\} \\
&= \frac{1}{2^{\mu+z}}\#\{s\in S:v(s_1,\ldots,s_\mu,\omega_1)+m-l-\mu=m-t\} \\
&= \frac{1}{2^{\mu+z}}\#\{s\in S:v(s_1,\ldots,s_\mu,\omega_1)=l+\mu-t\}
\end{aligned}
$$

Note that in counting $\#\{s\in S:v(s_1,\ldots,s_\mu,\omega_1)=l+\mu-t\}$, we choose the first μ entries so that there are exactly $l+\mu-t$ sign variations, and we arbitrarily choose z zeros from the ω part. Hence

$$
\begin{aligned}
\frac{1}{\#S}\#\{s\in S:v(s,+)=m-t\} &= \frac{1}{2^{\mu+z}}2^z\binom{\mu}{l+\mu-t} \\
&= \frac{1}{2^\mu}\binom{\mu}{l+\mu-t}.
\end{aligned}
$$

Now, we return to the full vector form for $V\gamma$. Note that in the below, we use the convention that m-dimensional standard unit vectors e_i are zero when $i\notin\{1,\ldots,m\}$ and binomial coefficients are likewise zero when the denominator is out of range. With this, we have

$$
C_{\text{sig}}\,A_{\text{sig}}(F,[\beta])=V\gamma
$$

$$
\begin{aligned}
&= \frac{1}{2^\mu} \sum_t \binom{\mu}{l+\mu-t} e_t \\
&= \frac{1}{2^\mu} \sum_k \binom{\mu}{k} e_{l+\mu-k} \qquad \text{by reindexing with } k = l+\mu-t \\
&= \frac{1}{2^\mu} \sum_{k=0}^{\mu} \binom{\mu}{k} e_{l+\mu-k} \qquad \text{since the summand is zero when } k \notin \{1,\dots,m\}.
\end{aligned}
$$

This is exactly the same as (3.1); as a result, we have shown that

$$
\mathrm{EC}(F,[\beta]) = \frac{1}{2^\mu} \sum_{k=0}^{\mu} \binom{\mu}{k} e_{l+\mu-k} = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta])
$$

and we are done. □

In the next lemma, we show that $\sum_{j=1}^{n} C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta_j])$ equals $C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G)$.

Lemma 91. *Let F and G be arbitrary real symmetric matrices. Let $\beta_1,\dots,\beta_n$ be the eigenvalues of G, including multiple eigenvalues. Then*

$$
\sum_{j=1}^{n} C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,[\beta_j]) = C_{\mathrm{sig}}\, A_{\mathrm{sig}}(F,G).
$$

Proof of Lemma 91. Let F and G be arbitrary and let $\beta_1,\dots,\beta_n$ be the eigenvalues of G. Note

$$
\sum_{j=1}^{n} C_{\mathrm{sig}} A_{\mathrm{sig}}(F,[\beta_j]) = C_{\mathrm{sig}} \sum_{j=1}^{n} A_{\mathrm{sig}}(F,[\beta_j]).
$$

Hence it suffices to show that

$$
\sum_{j=1}^{n} A_{\mathrm{sig}}(F,[\beta_j]) = A_{\mathrm{sig}}(F,G).
$$

Recall that for any two matrices F and G, the column vector $A_{\mathrm{sig}}(F,G)$ is defined as the vector indexed by $e \in \{0,1\}^m$ where

$$
\begin{aligned}
(A_{\mathrm{sig}}(F,G))_e &= \sigma(f_e(G)) \\
&= \#\{j : f_e(\beta_j) > 0\} - \#\{j : f_e(\beta_j) < 0\}.
\end{aligned}
$$

Hence

$$(A_{\text{sig}}(F,[\beta_j]))_e = \begin{cases} 1 & \text{if } f_e(\beta_j) > 0 \\ -1 & \text{if } f_e(\beta_j) < 0. \end{cases}$$

Note further that

$$\begin{aligned}
(A_{\text{sig}}(F,G))_e &= \#\{j : f_e(\beta_j) > 0\} - \#\{j : f_e(\beta_j) < 0\} \\
&= \sum_{j=1}^{n} \begin{cases} 1 & \text{if } f_e(\beta_j) > 0 \\ 0 & \text{else} \end{cases} - \sum_{j=1}^{n} \begin{cases} 1 & \text{if } f_e(\beta_j) < 0 \\ 0 & \text{else} \end{cases} && \text{by splitting each \# into a summation} \\
&= \sum_{j=1}^{n} \begin{cases} 1 & \text{if } f_e(\beta_j) > 0 \\ -1 & \text{if } f_e(\beta_j) < 0 \end{cases} && \text{combining the summations} \\
&= \sum_{j=1}^{n} \sigma\left(f_e\left([\beta_j]\right)\right) && \text{by definition of } \sigma \\
&= \sum_{j=1}^{n} (A_{\text{sig}}(F,[\beta_j]))_e && \text{by Lemma 89.}
\end{aligned}$$

Since this holds for all $e \in \{0,1\}^m$, we have

$$\sum_{j=1}^{n} A_{\text{sig}}(F,[\beta_j]) = A_{\text{sig}}(F,G),$$

and we are done. □

With that, we are now ready to prove Theorem 85 (Signature approach for arbitrary real symmetric matrices).

Proof of Theorem 85. Let F and G be arbitrary real symmetric matrices. Then

$$\begin{aligned}
&\text{EC}(F,G) \\
&\qquad \Big\| \quad \text{by Lemma 87} \\
&\sum_{j=1}^{n} \text{EC}(F,[\beta_j]) \\
&\qquad \Big\| \quad \text{by Lemma 89}
\end{aligned}$$

$$\sum_{j=1}^{n} C_{\text{sig}}\ A_{\text{sig}}(F,[\beta_j])$$

$$\Big\| \quad \text{by Lemma 91}$$

$$C_{\text{sig}} A_{\text{sig}}(F,G).$$

□

3.3 Symmetric Polynomials Approach

In this section, we will generalize Theorem 66 (Symmetric polynomimals approach for generic pairs of matrices) from the previous chapter to arbitrary pairs of matrices.

3.3.1 Main Theorem

In this subsection, we state the theorem for arbitrary pairs of matrices.

Theorem 92 (Symmetric polynomials approach for arbitrary pairs of matrices)**.** *Let $F \in \mathbb{R}^{m\times m}$ and $G \in \mathbb{R}^{n\times n}$ be **arbitrary** real symmetric matrices. We have*

$$\text{EC}(F,G) \ = \ C_{\text{sym}}\ A_{\text{sym}}(F,G),$$

where

1. *C_{sym} is the $m \times m$ matrix defined as in Definition 62*
2. *A_{sym} is the column vector of length m whose r-th entry is*

$$(A_{\text{sym}})_r = \overline{v}(D_r),$$

where $\overline{v}$ is defined as

$$\overline{v}(D_r) = v(D_r) + \frac{\text{taildeg}(D_r)}{2},$$

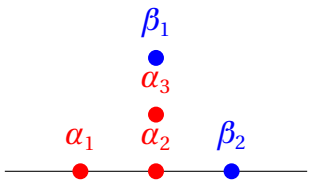
where $\text{taildeg}(D_r)$ *denotes the multiplicity of zero as a root of D_r.*

Remark 93. *Note that the only difference between the above theorem statement and that of Theorem 66 is the use of $\overline{v}$ instead of simply v. It is easily seen from the definition of D_r that, if F and G are a generic pair, then* $\text{taildeg}(D_r) = 0$ *for all r; therefore the above theorem reduces to Theorem 66 for generic pairs F and G.*

Example 94. *Let $F \in \mathbb{R}^{3\times 3}$ and $G \in \mathbb{R}^{2\times 2}$ be real symmetric matrices such that their corresponding eigenvalues are*

$$\alpha = (0,1,1), \qquad \beta = (1,2).$$

Pictorially, the eigenvalue arrangement is as shown below.



In Example 83, we computed directly from Definition 81 that

$$\mathrm{EC}(F,G) = \begin{bmatrix} 1/4 \\ 1/2 \\ 5/4. \end{bmatrix}$$

On the other hand, we compute

$$C_{\mathrm{sym}} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

$$A_{\mathrm{sym}}(F,G) = \begin{bmatrix} \overline{v}(D_1) \\ \overline{v}(D_2) \\ \overline{v}(D_3) \end{bmatrix} = \begin{bmatrix} \overline{v}\left(x^6 - 5x^5 + 9x^4 - 7x^3 + 2x^2\right) \\ \overline{v}\left(x^6 + 5x^5 + 8x^4 + 4x^3\right) \\ \overline{v}\left(x^2 - 2x\right) \end{bmatrix}.$$

For example, consider $\overline{v}\left(x^6 + 5x^5 + 8x^4 + 4x^3\right)$. We have

$$\begin{aligned} \overline{v}\left(x^6 + 5x^5 + 8x^4 + 4x^3\right) &= v(x^6 + 5x^5 + 8x^4 + 4x^3) + \frac{\mathrm{taildeg}(x^6 + 5x^5 + 8x^4 + 4x^3)}{2} \\ &= 0 + \frac{3}{2} \\ &= 3/2. \end{aligned}$$

Repeating this process for the other two entries of $A_{\mathrm{sym}}(F,G)$, we find that

$$A_{\mathrm{sym}}(F,G) = \begin{bmatrix} 5 \\ 3/2 \\ 3/2 \end{bmatrix}.$$

Observe now that

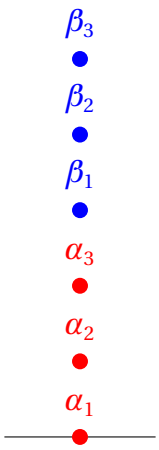
$$\begin{aligned} C_{\text{sym}} A_{\text{sym}}(F,G) &= \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 5 \\ 3/2 \\ 3/2 \end{bmatrix} \\ &= \begin{bmatrix} 1/4 \\ 1/2 \\ 5/4 \end{bmatrix} \\ &= \mathrm{EC}(F,G). \end{aligned}$$

△

Example 95. *Let $F \in \mathbb{R}^{3\times 3}$ and $G \in \mathbb{R}^{3\times 3}$ be real symmetric matrices such that their corresponding eigenvalues are*

$$\alpha = (0,0,0), \qquad \beta = (0,0,0).$$

Pictorially, the eigenvalue arrangement is as shown below.



Using Definition 81, one can compute that

$$\mathrm{EC}(F,G) = \begin{bmatrix} 9/8 \\ 9/8 \\ 3/8 \end{bmatrix}.$$

On the other hand, we compute

$$C_{\text{sym}} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$$
$$A_{\text{sym}} = \begin{bmatrix} \overline{v}\left(x^9\right) \\ \overline{v}\left(x^9\right) \\ \overline{v}\left(x^3\right) \end{bmatrix}.$$

For example, consider $\overline{v}\left(x^9\right)$*. We have*

$$\begin{aligned} \overline{v}\left(x^9\right) &= v(x^9) + \frac{\text{taildeg}(x^9)}{2} \\ &= 0 + 9/2 \\ &= 9/2. \end{aligned}$$

Repeating this process for the other two entries of A_{sym}*, we find that*

$$A_{\text{sym}} = \begin{bmatrix} 9/2 \\ 9/2 \\ 3/2 \end{bmatrix}.$$

Observe now that

$$\begin{aligned} C_{\text{sym}} A_{\text{sym}} &= \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 9/2 \\ 9/2 \\ 3/2 \end{bmatrix} \\ &= \begin{bmatrix} 9/8 \\ 9/8 \\ 3/8 \end{bmatrix} \\ &= \text{EC}(F, G). \end{aligned}$$

△

3.3.2 Proof

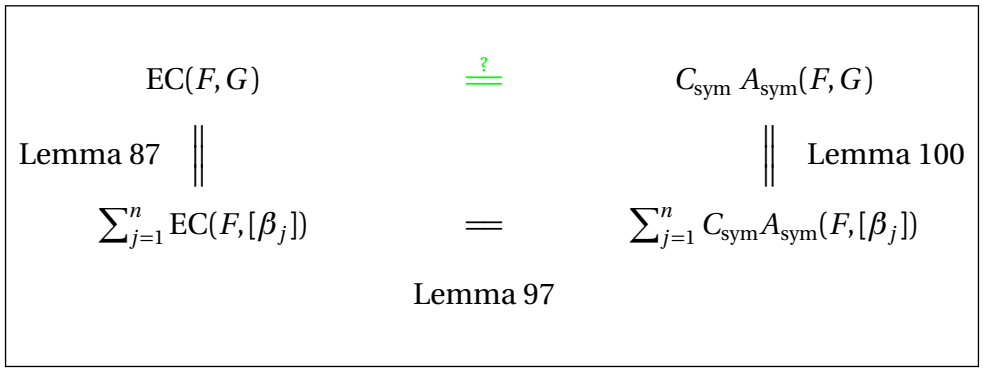
In this subsection, we will prove the above theorem (Theorem 92). We take the following approach.

1. First, we invoke Lemma 87 from the section on the signature algorithm (Section 3.2) to reduce to the case where $n = 1$ (i.e., where G is a 1×1 matrix).

2. Second, in Lemma 97 we prove the above theorem in the case where $n = 1$.

3. Finally, in Lemma 100, we build back up from the $n = 1$ case to arbitrary n. The proof is then complete.

Our proof strategy is illustrated by the following diagram.



Remark 96. *Note that the proof strategy in this section is nearly identical to the strategy taken in the signature approach for arbitrary pairs of matrices (Section 3.2).*

First, we reduce to the case where $n = 1$ (i.e., where G is a 1×1 matrix). For this, we invoke Lemma 87 from Section 3.2. Recall that Lemma 87 only used the definitions of EC and GEC and did not refer to the signature approach specifically; hence, the lemma applies here as well.

Next, we prove the above theorem in the case where $n = 1$.

Lemma 97. *Let $F \in \mathbb{R}^{m \times m}$ be arbitrary and let $\beta \in \mathbb{R}$. Then we have*

$$\mathrm{EC}(F, [\beta]) = C_{\mathrm{sym}} A_{\mathrm{sym}}(F, [\beta]).$$

Example 98. *Let $F = \mathrm{diag}(0, 1, 1)$ and let $\beta = 1$. One can use Definition 81 to compute*

$$\mathrm{EC}(F, [\beta]) = \begin{bmatrix} 1/2 \\ 1/4 \\ 1/2 \end{bmatrix}.$$

On the other hand, using the definitions of C_{sym} (from Definition 62) and A_{sym} (from the statement

of Theorem 92), we have

$$C_{\text{sym}}A_{\text{sym}}(F,[\beta]) = \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \overline{v}(D_1) \\ \overline{v}(D_2) \\ \overline{v}(D_3) \end{bmatrix}$$

Note that for the purposes of easier numeric computation and reasoning, we can replace D_r in the above with h_r since we know the eigenvalues of F and $G = [\beta]$. With this we can compute

$$\begin{aligned} \begin{bmatrix} \overline{v}(D_1) \\ \overline{v}(D_2) \\ \overline{v}(D_3) \end{bmatrix} &= \begin{bmatrix} \overline{v}(h_1) \\ \overline{v}(h_2) \\ \overline{v}(h_3) \end{bmatrix} \\ &= \begin{bmatrix} v(h_1) + \frac{\text{taildeg}(h_1)}{2} \\ v(h_2) + \frac{\text{taildeg}(h_2)}{2} \\ v(h_3) + \frac{\text{taildeg}(h_3)}{2} \end{bmatrix} \\ &= \begin{bmatrix} v(x^3 - x^2) + \frac{\text{taildeg}(x^3 - x^2)}{2} \\ v(x^3) + \frac{\text{taildeg}(x^3)}{2} \\ v(x) + \frac{\text{taildeg}(x)}{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 + 2/2 \\ 0 + 3/2 \\ 0 + 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 3/2 \\ 1/2 \end{bmatrix}, \end{aligned}$$

and therefore

$$\begin{aligned} C_{\text{sym}}A_{\text{sym}}(F,[\beta]) &= \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 \\ 3/2 \\ 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 \\ 1/4 \\ 1/2 \end{bmatrix} \\ &= \text{EC}(F,[\beta]). \end{aligned}$$

△

Proof of Lemma 97. If F does not have β as an eigenvalue (i.e. F and $[\beta]$ are a generic pair), then the lemma follows trivially by Theorem 66. Thus, for the remainder of the proof, suppose that β is an eigenvalue of F. By definition 81, the left-hand side of the lemma claim is

$$\begin{aligned}
\mathrm{EC}(F,[\beta]) &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} \mathrm{GEC}(F_d,[\beta]) \\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} C_{\mathrm{sym}} A_{\mathrm{sym}}(F_d,[\beta]) \qquad \text{since } F_d \text{ and } [\beta] \text{ are a generic pair} \\
&= C_{\mathrm{sym}}\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} A_{\mathrm{sym}}(F_d,[\beta]) \qquad \text{by linearity.}
\end{aligned}$$

Let $h_{r,d}$ (resp. $D_{r,d}$) be the polynomial constructed via Definition 64 for the perturbed matrix F_d and the matrix $G=[\beta]$; that is,

$$h_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r}} \left(x + \prod_{i\in I}\left((\alpha_i+\varepsilon d_i)-\beta\right)\right),$$

and respectively for $D_{r,d}$ (being those polynomials $h_{r,d}$ expressed in terms of the entries of F and G using the FTSP). Then for $r\in[m]$ we have

$$\begin{aligned}
(A_{\mathrm{sym}}(F_d,[\beta]))_r &= v(D_{r,d}) \\
&= v(h_{r,d}) \qquad \text{since by Def. 64 } D_{r,d}(p)=h_{r,d}(\alpha,\beta).
\end{aligned}$$

Now, we examine the right-hand side of the lemma claim. Recall that $A_{\mathrm{sym}}(F,G)$ is defined differently for generic pairs versus arbitrary pairs of matrices. Note that for $r\in[m]$ we have

$$\begin{aligned}
(A_{\mathrm{sym}}(F,[\beta]))_r &= \overline{v}(D_r) \\
&= v(D_r) + \frac{\mathrm{taildeg}(D_r)}{2} \qquad \text{by the definition in Theorem 92} \\
&= v(h_r) + \frac{\mathrm{taildeg}(h_r)}{2} \qquad \text{by the definition of } D_r.
\end{aligned}$$

Let $r\in[m]$ be arbitrary but fixed. To prove the lemma, it therefore suffices to show that

$$\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} v(h_{r,d}) = v(h_r) + \frac{\mathrm{taildeg}(h_r)}{2}.$$

We approach this proof by studying the left-hand side of the above.

Recall that

$$h_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r}} \left(x + \prod_{i\in I} \big((\alpha_i + \varepsilon d_i) - \beta\big) \right).$$

Let $M = \{i : \alpha_i = \beta\}$ and let

$$h^{=}_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}} \left(x + \prod_{i\in I} \big((\alpha_i + \varepsilon d_i) - \beta\big) \right)$$

$$h^{\neq}_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r\\ I\cap M=\emptyset}} \left(x + \prod_{i\in I} \big((\alpha_i + \varepsilon d_i) - \beta\big) \right).$$

Then, clearly we have $h_{r,d} = h^{\neq}_{r,d} h^{=}_{r,d}$. Informally speaking, the polynomial $h^{\neq}_{r,d}$ comprises the terms of $h_{r,d}$ which are invariant under perturbation, because it only involves α's which do not coincide with β. The polynomial $h^{=}_{r,d}$ collects all the terms of $h_{r,d}$ which contain at least one α which coincides with β, and therefore comprises all terms which are affected by perturbation and depend on d.

For each $I \subset [m]$ with $\#I = r$ and $I \cap M \neq \emptyset$, let

$$I_- = \{i \in I : \alpha_i < \beta\}$$
$$I_0 = \{i \in I : \alpha_i = \beta\}$$
$$I_+ = \{i \in I : \alpha_i > \beta\}.$$

Then, by construction $I_0 \neq \emptyset$ and $I = I_- \uplus I_0 \uplus I_+$ and so we have

$$h^{=}_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}} \left(x + \prod_{i\in I_-} \big((\alpha_i + \varepsilon d_i) - \beta\big) \prod_{i\in I_+} \big((\alpha_i + \varepsilon d_i) - \beta\big) \prod_{i\in I_0} \big((\alpha_i + \varepsilon d_i) - \beta\big) \right).$$

Let

$$p(h^{=}_{r,d}) = \#\text{ positive roots of } h^{=}_{r,d}.$$

Then, from the above we have

$$p(h^{=}_{r,d}) = \#\{I : -(-1)^{\#I_-}(1)^{\#I_+} \prod_{i\in I_0} d_i = 1\}.$$

Note that since $h^{=}_{r,d}$ is a polynomial with only real roots, by Descartes' rule of signs we have that $v(h^{=}_{r,d}) = p(h^{=}_{r,d})$. Hence it suffices to show that

$$\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h_{r,d}) = p(h_{r,d}) + \frac{\text{taildeg}(h_{r,d})}{2}.$$

We first prove the following claim.

Claim: We have

$$\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h^{=}_{r,d}) = \underset{d\in\{-1,1\}^m}{\text{average}}\, p(h^{=}_{r,d}) \;=\; \frac{\deg h^{=}_{r,d}}{2}.$$

Proof of claim: For a concise notation, let $S_{I,d} = -(-1)^{\#I_-}(1)^{\#I_+}\prod_{i\in I_0} d_i$. With this notation, we then have

$$p(h^{=}_{r,d}) = \#\{I : S_{I,d} = 1\}. \tag{3.3}$$

Note that since $\alpha_i + \varepsilon d_i \neq \beta$ for all i, we have that $h^{=}_{r,d}$ has only nonzero real roots. Hence, for each I and d, we have $S_{I,d} = \pm 1$. Thus, to prove the claim, it suffices to show that

$$\#\{(I,d) : S_{I,d} = 1\} \;=\; \#\{(I,d) : S_{I,d} = -1\}.$$

It suffices to show that for every fixed I and d, there exists a unique d' so that

$$S_{I,d} = -S_{I,d'}.$$

Now let

$$\begin{aligned} &I \subset [m] \text{ such that } \#I = r \text{ and } I \cap M \neq \emptyset \\ &d \in \{-1,1\}^m \end{aligned}$$

be fixed. Recall that $I_0 = \{i : \alpha_i = \beta\} = I \cap M \neq \emptyset$. Now let k be the smallest element of I_0. Then let

$$d' = (d_1, \ldots, d_{k-1}, -d_k, d_{k+1}, \ldots, d_m);$$

i.e., let d' be a copy of d but with the sign of the k-th entry flipped. Obviously, the map $d \mapsto d'$ is a 1-1 correspondence. Further, clearly $S_{I,d'} = -S_{I,d}$, since d and d' differ by exactly one sign in the kth slot, and $k \in I_0$. Hence, we have that

$$\#\{(I,d) : S_{I,d} = 1\} \;=\; \#\{(I,d) : S_{I,d} = -1\};$$

in other words, for fixed I, we have that exactly half of the d vectors give $S_{I,d} = 1$ while the other

half give -1. Since the size of the set $\{-1,1\}^m$ is 2^m, this means that for fixed I, we have

$$\#\{d \in \{-1,1\}^m : S_{I,d} = 1\} = \frac{1}{2}2^m = 2^{m-1}. \tag{3.4}$$

Now, note that

$$\begin{aligned}
\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h^{=}_{r,d}) &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} \#\{I : S_{I,d} = 1\} && && \text{by (3.3)}\\
&= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m}\sum_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}} && \delta_{S_{I,d},1} && \text{counting becomes sum}\\
&= \frac{1}{2^m}\sum_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}}\sum_{d\in\{-1,1\}^m} && \delta_{S_{I,d},1} && \text{switching the order of summation}\\
&= \frac{1}{2^m}\sum_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}} && 2^{m-1} && \text{by (3.4)}\\
&= \frac{1}{2}\sum_{\substack{I\subset[m]\\ \#I=r\\ I\cap M\neq\emptyset}} && 1\\
&= \frac{1}{2}\quad \deg h^{=}_{r,d}.
\end{aligned}$$

Note that the key step in the above was switching the order of summation and summing over d inside first. This is because it is very difficult to reason when the summation over I is inside (i.e. reasoning with fixed d); it is much easier to reason that, for fixed I, exactly half of the pairs (I,d) have $S_{I,d} = 1$. We explain this idea more in depth in Example 99. With this, the claim is proved.

To finish the proof of the lemma, recall that

$$h^{\neq}_{r,d} = \prod_{\substack{I\subset[m]\\ \#I=r\\ I\cap M=\emptyset}}\left(x + \prod_{i\in I}\big((\alpha_i + \varepsilon d_i) - \beta\big)\right).$$

Note that since $h^{\neq}_{r,d}$ involves only α's which do not coincide with β, by a small enough choice of ε (see Lemma 86 for details) we have that

$$\operatorname{sign}((\alpha_i + \varepsilon d_i) - \beta) = \operatorname{sign}(\alpha_i - \beta).$$

Therefore

$$\begin{aligned} p(h^{\neq}_{r,d}) &= v(h^{\neq}_{r,d}) && \text{by Descartes' rule of signs} \\ &= v(h_r) && \text{by the discussion above} \\ &= p(h_r) && \text{by Descartes' rule of signs, again.} \end{aligned}$$

Putting everything together, we therefore have

$$\begin{aligned} \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} v(h_{r,d}) &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h_{r,d}) && \text{by Descartes' rule of signs} \\ &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h^{\neq}_{r,d}) + p(h^{=}_{r,d}) && \text{since } h_{r,d} = h^{\neq}_{r,d}h^{=}_{r,d} \\ &= \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h_r) + p(h^{=}_{r,d}) && \text{since } p(h^{\neq}_{r,d}) = p(h_r), \text{ by the discussion above} \\ &= p(h_r) + \frac{1}{2^m}\sum_{d\in\{-1,1\}^m} p(h^{=}_{r,d}) && \text{since } p(h_r) \text{ does not depend on } d \\ &= p(h_r) + \frac{\deg h^{=}_{r,d}}{2} && \text{by the claim.} \end{aligned}$$

Finally, note that if we (by slight abuse of notation) substitute $d = (0,\dots,0)$ into the expression for $h_{r,d}$, we find that

$$\begin{aligned} \operatorname{taildeg}(h_r) &= \operatorname{taildeg}(h_{r,0}) \\ &= \operatorname{taildeg}\left(h^{\neq}_{r,0}\right) + \operatorname{taildeg}\left(h^{=}_{r,0}\right) && \text{since } h_{r,0} = h^{\neq}_{r,0}h^{=}_{r,0} \\ &= 0 + \operatorname{taildeg}\left(h^{=}_{r,0}\right) && \text{since all roots of } h^{\neq}_{r,0} \text{ are nonzero} \\ &= 0 + \deg\left(h^{=}_{r,d}\right) && \forall d, \text{ since plugging in } d \neq 0 \text{ preserves degree} \\ &= \deg(h^{=}_{r,d}). \end{aligned}$$

Finally, if we continue from above, we have

$$\begin{aligned} p(h_r) + \frac{\deg h^{=}_{r,d}}{2} &= p(h_r) + \frac{\operatorname{taildeg}(h_r)}{2} && \text{by the above} \\ &= v(h_r) + \frac{\operatorname{taildeg}(h_r)}{2} && \text{by Descartes' rule of signs} \\ &= \overline{v}(h_r). \end{aligned}$$

Thus we have shown that

$$\frac{1}{2^m}\sum_{d\in\{-1,1\}^m} v(h_{r,d}) = v(h_r) + \frac{\text{taildeg}(h_r)}{2},$$

which means that

$$\text{EC}(F,[\beta]) = C_{\text{sym}} A_{\text{sym}}(F,[\beta]).$$

Thus the lemma is proved. □

Example 99. *Since the proof of Lemma 97 is quite long and complicated, let us provide an example to better illustrate the key idea which led to the above proof. The crucial observation in the above proof was (3.4); that is, over all I and d, exactly half give $S_{I,d} = 1$.*

Let us consider an example where $F = \text{diag}(0,0,1)$ and $\beta = 0$. Using the notation established in the proof above, we have $m = 3$ and $M = \{i : \alpha_i = \beta\} = \{1,2\}$. We will consider the case where $r = 2$. Note

$$\begin{aligned} h^{=}_{2,d} &= \prod_{\substack{I\subset[3]\\ \#I=2\\ I\cap\{1,2\}\neq\emptyset}} \left(x + \prod_{i\in I}(\alpha_i + \varepsilon d_i - \beta)\right) \\ &= \underbrace{(x + (\alpha_1+\varepsilon d_1-\beta)(\alpha_2+\varepsilon d_2-\beta))}_{I=\{1,2\}} \\ &\quad \underbrace{(x + (\alpha_1+\varepsilon d_1-\beta)(\alpha_3+\varepsilon d_3-\beta))}_{I=\{1,3\}} \\ &\quad \underbrace{(x + (\alpha_2+\varepsilon d_2-\beta)(\alpha_3+\varepsilon d_3-\beta))}_{I=\{2,3\}}. \end{aligned}$$

Note that each of the three terms in $h^{=}_{2,d}$ is labelled with the corresponding I. For each I, we compute the corresponding sets I_-, I_0, and I_+.

I	I_-	I_0	I_+
$\{1,2\}$	$\emptyset$	$\{1,2\}$	$\emptyset$
$\{1,3\}$	$\emptyset$	$\{1\}$	$\{3\}$
$\{2,3\}$	$\emptyset$	$\{2\}$	$\{3\}$

Now, consider $d = (d_1,d_2,d_3) \in \{-1,1\}^3$. Recall from the above proof that $S_{I,d} = -(-1)^{\#I_-}(1)^{\#I_+}\prod_{i\in I_0} d_i$. We compute

$$\begin{aligned} S_{12,d} &= -(-1)^0\ (1)^0\ \ d_1 d_2\ = -d_1 d_2 \\ S_{13,d} &= -(-1)^0\,(1)^1\ \ d_1\qquad = -d_1 \end{aligned}$$

$$S_{23,d} = -(-1)^0\,(1)^1 \;\; d_2 \qquad = -d_2.$$

Finally, with this information, we now compute $S_{I,d}$ for each I enumerated above and for each $d \in \{-1,1\}^3$. The table below is constructed as follows. The rows are indexed by the I sets, where the leftmost column shows the decompositions I_-, I_0, and I_+, the disjoint union of which is I. The columns are indexed by $d \in \{-1,1\}^m$. Each interior cell contains the entries of the d vector which correspond to the indices in I_0.Immediately below in the same cell is the sign of $S_{I,d}$. For example, the top-left-most cell corresponds to $I_0\{1,2\}$ and $d = ---$, and so the cell contains $d_1, d_2 = -,-$, and beneath it is the sign of $S_{I,d} = -$ colored red. (Positive $S_{I,d}$ are colored green.) Each interior cell is also color-paired with the unique cell in the same row whose d vector is obtained by reversing the sign of d_k, where k is the smallest index in I_0, as in the proof. For example, in the top-left-most cell the first entry $d_1 = -$ is colored cyan and is paired with the corresponding cell in the same row which has $d_1 = +$.

$I \setminus d$	$---$	$--+$	$-+-$	$-++$	$+--$	$+-+$	$++-$	$+++$
$I_- = \emptyset$ $I_0 = 12$ $I_+ = \emptyset$	$d_1, d_2 = -,-$ $S_{I,d} = -$	$-,-$ $-$	$-,+$ $+$	$-,+$ $+$	$+,-$ $+$	$+,-$ $+$	$+,+$ $-$	$+,+$ $-$
$I_- = \emptyset$ $I_0 = 1$ $I_+ = 3$	$d_1 = -$ $S_{I,d} = +$	$-$ $+$	$-$ $+$	$-$ $+$	$+$ $-$	$+$ $-$	$+$ $-$	$+$ $-$
$I_- = \emptyset$ $I_0 = 2$ $I_+ = 3$	$d_2 = -$ $S_{I,d} = +$	$-$ $+$	$+$ $-$	$+$ $-$	$-$ $+$	$-$ $+$	$+$ $-$	$+$ $-$

The key takeaway of the above table is as follows. Observe that in each row, exactly half the cells have a green plus in the bottom entry (the $S_{I,d}$ value) and half have a red minus. This shows that $\#\{(I,d) : S_{I,d} = 1\} = \#\{(I,d) : S_{I,d} = -1\}$. Further, in each row, there is exactly one cyan minus and one cyan plus, and respectively for blue, red, and magenta. Each colored pair represents the map $d \mapsto d'$ discussed in the proof, which is obtained by reversing the sign of the entry of d corresponding to the smallest index in I_0.

Finally, observe that summing over all (I,d) corresponds to simply summing over the entire table. If we sum down the rows first (i.e., summing over I with d held constant first), it is unclear what the pattern is, if any. However, if we instead sum across the columns (i.e., summing over d with I held constant), it is easy to see that exactly half the cells have $S_{I,d} = 1$ and the other half have $S_{I,d} = -1$. This corresponds exactly with switching the order of summation in the previous proof. △

Finally, we build back up from the $n = 1$ case to arbitrary n.

Lemma 100. *Let F and G be arbitrary real symmetric matrices. Then*

$$\sum_{j=1}^{n} C_{\text{sym}} A_{\text{sym}}(F, [\beta_j]) = C_{\text{sym}} A_{\text{sym}}(F, G).$$

Proof. Recall that

$$(A_{\text{sym}}(F, [\beta_j]))_r = \overline{v}(D_{r,j})$$

where

$$D_{r,j} = \prod_{I \subset [m], \#I = r} \left(x + \prod_{i \in I} (\alpha_i - \beta_j) \right).$$

Note that we also have

$$\begin{aligned} D_r &= \prod_{\substack{I \subset [m], \#I = r \\ j \in [n]}} \left(x + \prod_{i \in I} (\alpha_i - \beta_j) \right) \\ &= \prod_{j \in [n]} D_{r,j}. \end{aligned}$$

Then

$$\begin{aligned} \sum_{j=1}^{n} (A_{\text{sym}}(F, [\beta_j]))_r &= \sum_{j=1}^{n} \overline{v}(D_{r,j}) \\ &= \sum_{j=1}^{n} v(D_{r,j}) + \frac{\text{taildeg}(D_{r,j})}{2} \\ &= \sum_{j=1}^{n} v(D_{r,j}) + \frac{1}{2} \sum_{j=1}^{n} \text{taildeg}\, D_{r,j} \\ &= v(D_r) + \frac{\text{taildeg}\, D_r}{2} \qquad \text{since } D_r = \prod_{j=1}^{n} D_{r,j} \\ &= \overline{v}(D_r), \end{aligned}$$

which is exactly the r-th element of $A_{\text{sym}}(F, G)$. Therefore, we have

$$\sum_{j=1}^{n} C_{\text{sym}} A_{\text{sym}}(F, [\beta_j]) = C_{\text{sym}} \sum_{j=1}^{n} A_{\text{sym}}(F, [\beta_j])$$

$$= C_{\text{sym}} A_{\text{sym}}(F, [\beta_j]).$$

Thus the proof is complete. □

With that, we are now ready to prove Theorem 92 (Symmetric polynomials approach for arbitrary real symmetric matrices).

Proof of Theorem 92. Let F and G be arbitrary real symmetric matrices. Then

$$\begin{aligned}
&\text{EC}(F, G) \\
&\qquad \Big\| \quad \text{by Lemma 87} \\
&\sum_{j=1}^{n} \text{EC}(F, [\beta_j]) \\
&\qquad \Big\| \quad \text{by Lemma 97} \\
&\sum_{j=1}^{n} C_{\text{sym}}\ A_{\text{sym}}(F, [\beta_j]) \\
&\qquad \Big\| \quad \text{by Lemma 100} \\
&C_{\text{sym}} A_{\text{sym}}(F, G).
\end{aligned}$$

□

REFERENCES

[1] M. Michelen and J. Sahasrabudhe. Central limit theorems and the geometry of polynomials. *Journal of the European Mathematical Society*, 2024.

[2] G. Codenotti, S. Gardoll, and T. Theobald. Combinatorics and preservation of conically stable polynomials, 2022.

[3] L. Gurvits, N. Klein, and J. Leake. From trees to polynomials and back again: New capacity bounds with applications to tsp, 2023.

[4] A. Aleman, K. Perfekt, S. Richter, C. Sundberg, and J. Sunkes. Cyclicity in the drury-arveson space and other weighted besov spaces, 2023.

[5] A. Chattopadhyay, T. Mandal, and A. Marrani. Near-extremal freudenthal duality. *Journal of High Energy Physics*, 2023(8):1–30, 2023.

[6] A. Abu Al Haija, M. Dong, B. Liang, and G. Boudreau. Resource allocation for massive mimo hetnets with quantize-forward relaying, 2021.

[7] J. Huang, T. Cao, and M. Zhang. Superradiant stability of five and six-dimensional extremal reissner–nordstrom black holes. *The European Physical Journal C*, 81(10), October 2021.

[8] S. McCallum. *An Improved Projection Operator for Cylindrical Algebraic Decomposition*. PhD thesis, University of Wisconsin-Madison, 1984.

[9] M. Ben-Or, D. Kozen, and J. H. Reif. The complexity of elementary algebra and geometry. *J. Comput. System Sci.*, 32(2):251–264, 1986.

[10] D. Yu. Grigor'ev. The complexity of deciding Tarski algebra. *Journal of Symbolic Computation*, 5(1,2):65–108, 1988.

[11] V. Weispfenning. A new approach to quantifier elimination for real algebra. In B. Caviness and J. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer Verlag, 1996. Texts and Monographs in Symbolic Computation.

[12] L. Gonzalez-Vega, H. Lombardi, T. Recio, and M.-F. Roy. Sturm-Habicht sequences. In *Proceedings of the ACM-SIGSAM 1989 International Symposium on Symbolic and Algebriaic Computation*, pages 136–146, July 1989.

[13] H. Hong. An improvement of the projection operator in cylindrical algebraic decomposition. In *International Symposium of Symbolic and Algebraic Computation (ISSAC-90)*, pages 261–264. ACM, 1990.

[14] H. Hong. *Improvements in CAD–based Quantifier Elimination*. PhD thesis, The Ohio State University, 1990.

[15] G. E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, sep 1991.

[16] H. Hong. Simple solution formula construction in cylindrical algebraic decomposition based quantifier elimination. In *International Conference on Symbolic and Algebraic Computation ISSAC-92*, pages 177–188, 1992.

[17] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. *Journal of Symbolic Computation*, 13(3):255–352, 1992.

[18] J. F. Canny. Improved algorithms for sign and existential quantifier elimination. *Computer Journal*, 36:409–418, 1993. In a special issue on *computational quantifier elimination*, edited by H. Hong.

[19] R. Loos and V. Weispfenning. Applying linear quantifier elimination. *Computer Journal*, 36(5):450–462, 1993. In a special issue on *computational quantifier elimination*, edited by H. Hong.

[20] V. Weispfenning. Quantifier elimination for real algebra – the cubic case. In *International Symposium on Symbolic and Algebraic Computation 94*, pages 258–263, 1994.

[21] L. Gonzalez-Vega. A combinatorial algorithm solving some quantifier elimination problems. In B. Caviness and J. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer Verlag, 1996. Texts and Monographs in Symbolic Computation.

[22] S. McCallum. Factors of iterated resultants and discriminants. *J. Symb. Comput.*, 27(4):367–385, 1999.

[23] S. McCallum and C. W. Brown. On delineability of varieties in cad-based quantifier elimination with two equational constraints. In *ISSAC*, pages 71–78, 2009.

[24] C. W. Brown. Improved projection for cylindrical algebraic decomposition. *J. Symb. Comput.*, 32(5):447–465, 2001.

[25] C. W. Brown. Simple cad construction and its applications. *J. Symb. Comput.*, 31(5):521–547, 2001.

[26] A. Strzebonski. Cylindrical algebraic decomposition using validated numerics. *J. Symb. Comput.*, 41(9):1021–1038, 2006.

[27] C. Chen, M. Moreno Maza, B. Xia, and L. Yang. Computing cylindrical algebraic decomposition via triangular decomposition. In *ISSAC*, pages 95–102, 2009.

[28] C. W. Brown. Fast simplifications for tarski formulas based on monomial inequalities. *J. Symb. Comput.*, 47(7):859–882, 2012.

[29] H. Hong and M. S. El Din. Variant quantifier elimination. *J. Symb. Comput.*, 47(7):883–901, 2012.

[30] A. Tarski. The completeness of elementary algebra and geometry. 1930. Reprinted in 1967.

[31] R. Descartes. *Géométrie*. 1636. In A source book in Mathematics, pages 90-131. Harvard University Press, 1969.

[32] M. Ben-Or, D. Kozen, and J. Reif. The complexity of elementary algebra and geometry. *Journal of Computer and System Sciences*, 32(2):251–264, 1986.

[33] M. Stoll. *Introduction to real analysis*. Chapman and Hall/CRC, 2021.

[34] R. P. Stanley. *Enumerative Combinatorics*. Cambridge University Press, 1999.

[35] P. Garrett. Symmetric polynomials. `https://www-users.cse.umn.edu/~garrett/m/algebra/notes_2023-24/15.pdf`. Accessed: 2026-02-17.

[36] F. Budan de Boislaurent. *Nouvelle méthode pour la résolution des équations*. 1807.

[37] J. Fourier. *Analyse des équations déterminées*. F. Didot, Paris, 1831.