Determinantal Representations and the Image of the Principal Minor Map

Abeer Al Ahmadieh

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:
Cynthia Vinzant, Chair
Farbod Shokrieh
Rekha Thomas

Program Authorized to Offer Degree:
Department of Mathematics

University of Washington

Abstract

Determinantal Representations and
the Image of the Principal Minor Map

Abeer Al Ahmadieh

Chair of the Supervisory Committee:
Cynthia Vinzant
Department of Mathematics

Research in algebraic geometry has interfaces with other fields, such as matrix theory, combinatorics, and convex geometry. It is a branch of mathematics that studies solution to systems of polynomial equations and inequalities. This dissertation consists of three projects, all of which use techniques from matrix theory, convex geometry and symbolic computation to approach problems in algebraic geometry.

In the first chapter we introduce some of the necessary background in classical, convex and real algebraic geometry. We also introduce the principal minor problem and its applications.

In the second chapter we study the image of the principal minor map of symmetric matrices. The *principal minor map* is the map that assigns to each $n \times n$ matrix the 2^n-vector of its principal minors. By exploiting a connection with symmetric determinantal representations, we characterize the image of the subspace of symmetric matrices through the condition that certain polynomials coming from the so-called Rayleigh differences are squares in the polynomial ring over any unique factorization domain R. In almost all cases, one can characterize this image using the orbit of Cayley's hyperdeterminant under the action of $(\mathrm{SL}_2(R))^n \rtimes S_n$. Over $\mathbb{C}$, this recovers a characterization of Oeding from 2011, and over $\mathbb{R}$, the orbit of a single additional quadratic inequality suffices to cut out the image.

In the third chapter we explore determinantal representations of multiaffine polynomials and consequences for the image of various spaces of matrices under the principal minor map. We show that a real multiaffine polynomial has a definite Hermitian determinantal representation if and only if all of its Rayleigh differences factor as Hermitian squares and use this characterization to conclude that the image of the space of Hermitian matrices under the principal minor map is cut out by the orbit of finitely many equations and inequalities under the action of $(\mathrm{SL}_2(\mathbb{R}))^n \rtimes S_n$. We also study such representations over more general fields with quadratic extensions. Factorizations of Rayleigh differences prove an effective tool for capturing subtle behavior of the principal minor map. In contrast to the Hermitian case, we give examples to show over any field $\mathbb{F}$, there is no finite set of equations whose orbit under $(\mathrm{SL}_2(\mathbb{F}))^n \rtimes S_n$ cuts out the image of $n \times n$ matrices under the principal minor map for every n.

In the fourth chapter we study the variety of the space of complete quadrics. It is the space of nondegenerate quadrics, representing nonsingular quadrics, in addition to the so-called degenerate quadrics. We aim at generalizing the space of complete quadrics associated to any hyperbolic polynomial. To any homogeneous polynomial h we naturally associate a variety Ω_h which maps birationally onto the graph of the gradient map ∇h and which agrees with the space of complete quadrics when h is the determinant of a generic symmetric matrix. We give a sufficient criterion for Ω_h being smooth which applies, for example, when h is an elementary symmetric polynomial. In this case Ω_h is a smooth toric variety associated to a certain generalized permutohedron. We also give examples when Ω_h is not smooth.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

This is the end of a long journey through which I have met many people who helped in shaping and supporting me, and at last I have a moment to reflect and express my deep gratitude.

First and foremost, I would like to thank my advisor Cynthia Vinzant for her endless guidance, patience, support and care. She gave me the skills to overcome many challenges and I gained a lot of invaluable knowledge in the process. Thank you, Cynthia, for the inspiring discussions and for always being there.

I would like to thank my committee members, Shayan Oveis Gharan, Farbod Shokrieh, and Rekha Thomas for their insightful feedback on my dissertation.

I would like to thank Bernd Sturmfels for his encouragement and support and for putting together a great team of researchers during the LSSM project. Thanks to Mario Kummer for his guidance during this project and for the valuable discussions we had with Miruna-Stefana Sorea.

I would like to thank Tye Lidman, for guiding me and putting me on the correct track at NCSU. I was lucky to have him as my academic advisor. Thanks to Hoon Hong for teaching me computer algebra at NCSU, one of the best classes that I attended during my graduate studies.

I would like to thank the professors at AUB who were the main reason behind me reaching this stage. Thanks to my masters thesis advisor Tamer Tlas, who was always there during my Ph.D studies to support me and to provide insightful advice whenever needed. I can't express my gratitude enough to Wissam Raji, whose support, enthusiasm and belief in me were the main reasons behind this journey in the U.S.

I would like to acknowledge the University of Washington graduate school for their financial support through the Top Scholar Award, and the mathematics department at the University of Washington for their financial support through the MCFellowship.

Last but not the least, I am extremely grateful to my family for their sincere encouragement and inspiration. Without their assistance, love, and care, I would have never reached where I am today. Special thanks to my niece Leah, for offering me the chance to spend such nice breaks with her and for supporting me with the purest positive energy in this world.

DEDICATION

to my family

Chapter 1

INTRODUCTION

In this chapter we introduce some of the necessary background on classical and convex algebraic geometry that will be used in this thesis.

1.1 Classical Algebraic Geometry

Algebraic geometry is the study of algebraic equations and their sets of solutions or *varieties*. It combines the visualization of geometry and the accurateness of algebra. Oscar Zariski considered it as

> *the best things that has happened to commutative algebra in a long time.*

In this section we give a brief introduction to the basic concepts of classical algebraic geometry with focus on the theorems that will be used later in this thesis. The reader is referred to [9, 20, 37, 64] for a more thorough introduction. Through out this chapter we will denote by $\mathbb{F}$ any field, $\mathbb{K}$ any algebraically closed field and by $\mathbb{K}[x_1, \dots, x_n]$ the polynomial ring over $\mathbb{K}$ in the variables $x_1, \dots, x_n$.

1.1.1 Affine Varieties and Regular Maps

We will start with some basic definitions.

Definition 1.1.1. Let I be an ideal of $\mathbb{K}[x_1, \dots, x_n]$. The *algebraic variety* $\mathcal{V}(I)$ is the set of common zeros of the polynomials in I:

$$\mathcal{V}(I) = \{a \in \mathbb{K}^n : f(a) = 0 \text{ for all } f \in I\}.$$

Any variety V can be defined as the vanishing set of a finite set of polynomials, by the Hilbert basis theorem.

Theorem 1.1.2. *(***Hilbert Basis Theorem***) Any ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ is finitely generated. That is, there exists $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ such that*

$$I = \langle f_1, \ldots, f_m \rangle = \left\{ \sum_{i=1}^{m} h_i f_i : h_i \in \mathbb{K}[x_1, \ldots, x_n] \right\}.$$

The algebraic varieties define a topology on $\mathbb{K}^n$. Unlike the Euclidean topology, all nonempty open sets in this topology are dense and thus it is far from being Hausdorff.

Definition 1.1.3. The **Zariski topology** is a topology on $\mathbb{K}^n$ where the closed sets are the algebraic varieties and the open sets are their complements.

To each subset $X \subset \mathbb{K}^n$, we associate an ideal $\mathcal{I}(X) \subset \mathbb{K}[x_1, \ldots, x_n]$ defined by

$$\mathcal{I}(X) = \{ f \in \mathbb{K}[x_1, \ldots, x_n] : f(a) = 0 \text{ for all } a \in X \}.$$

In particular, if J is an ideal and $\mathcal{V}(J)$ is its corresponding variety, then $\mathcal{I}(\mathcal{V}(J))$ equals the *radical* of J. This is one of the fundamental theorems of algebraic geometry, called Hilbert's Nullstellensatz.

Definition 1.1.4. The **radical** of an ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$, denoted by $\sqrt{I}$, is defined as

$$\sqrt{I} = \{ f \in \mathbb{K}[x_1, \ldots, x_n] : f^m \in I \text{ for some } m \in \mathbb{N} \}.$$

Theorem 1.1.5. *(Hilbert's Nullstellensatz) For any ideal $J \subseteq \mathbb{K}[x_1, \ldots, x_n]$,*

$$\mathcal{I}(\mathcal{V}(J)) = \sqrt{J}.$$

We associate to each variety V a coordinate ring $\mathbb{K}[V]$, which is the ring of polynomial functions of $\mathbb{K}^n$ restricted to V.

Definition 1.1.6. The **coordinate ring** of an algebraic variety V is defined by

$$\mathbb{K}[V] = \mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}(V).$$

We denote by $\mathbb{K}(V)$ the field of fractions of $\mathbb{K}[V]$.

Definition 1.1.7. The **field of rational functions** on an algebraic variety V is defined by

$$\mathbb{K}(V) = \left\{ \frac{f}{g} : f, g \in \mathbb{K}[V], g \not\equiv 0 \text{ and } \frac{f}{g} = \frac{h}{\ell} \iff f\ell - hg \in \mathcal{I}(V) \right\}.$$

An element f in $\mathbb{K}(V)$ is called a **rational function**.

Definition 1.1.8. A rational function $f \in \mathbb{K}(V)$ is **regular at a point** p in $\mathbb{K}^n$, if there exists an open neighborhood U of p and $g, h \in \mathbb{K}[V]$ such that $f = g/h$ on U and $h(q) \neq 0$ for all $q \in U$. We say that f is **regular on** U if f is regular at each point $p \in U$.

Definition 1.1.9. Let $X \subset \mathbb{K}^n$ and $Y \subset \mathbb{K}^m$. A **rational map** $\varphi : X \dashrightarrow Y$ is a map

$$\varphi : X \dashrightarrow \mathbb{K}^m \text{ given by } \varphi(x) = (f_1(x), \ldots, f_m(x)),$$

where each f_i is a rational function and such that $\varphi(X) \subseteq Y$. The map φ is **dominant** if $\varphi(X)$ is Zariski-dense in Y.

Definition 1.1.10. Let X and Y be two varieties. A **birational map** $\varphi : X \dashrightarrow Y$ is a dominant rational map such that there exists another dominant rational map $\psi : Y \dashrightarrow X$ such that $\varphi \circ \psi = \mathrm{id}_Y$ and $\psi \circ \varphi = \mathrm{id}_X$. The varieties X and Y are called **birationally equivalent** in this case.

Definition 1.1.11. Let X and Y be affine varieties and let $U \subseteq X$ be an open subset. A **morphism** $\varphi : U \longrightarrow Y$ is a rational map $\varphi : X \dashrightarrow Y$ such that φ is regular on U. If there exists W an open subset of Y such that $\varphi(U) \subseteq W$ and a morphism $\psi : W \longrightarrow X$ such that $\psi(W) \subseteq U$, then φ is called an **isomorphism**.

An example of a morphism that will be discussed in-depth in Chapter 2 is the *principal minor map.*

Example 1.1.12. Consider the map $\varphi : \mathbb{C}^{\binom{n+1}{2}} \longrightarrow \mathbb{C}^{2^n}$ which takes an $n \times n$ complex symmetric matrix to the length-2^n vector of its principal minors. We call this **the principal**

minor map. The map φ is a morphism, since it is defined by polynomial functions. The image is closed [41] and hence it defines a variety whose defining equations are the main object of study of Chapter 2.

1.1.2 Projective Varieties

In this section, we will extend the affine space $\mathbb{K}^n$ and add the "points at infinity".

Definition 1.1.13. Projective n-space over $\mathbb{K}$, denoted by $\mathbb{P}^n$, is the set of lines in $\mathbb{K}^{n+1}$ that passes through the origin. In other words

$$\mathbb{P}^n = \left(\mathbb{K}^{n+1} \setminus \{0\}\right) / \sim$$

where $(a_0, \dots, a_n) \sim (b_0, \dots, b_n)$ if $(a_0, \dots, a_n) = \lambda(b_0, \dots, b_n)$ for some $\lambda \in \mathbb{K}^*$. The equivalence class of an element $(a_0, \dots, a_n)$ will be denoted by $(a_0 : \dots : a_n)$ and will denote the **homogeneous coordinates** of a point in $\mathbb{P}^n$.

To build an algebro-geometric dictionary, similar to the affine case, we will start by defining the homogeneous polynomials and then homogeneous ideals. These ideals will be the key ingredients in defining projective varieties.

Definition 1.1.14. A polynomial $f \in \mathbb{K}[x_1, \dots, x_n]$ is called **homogeneous** of degree d if all the monomials in this polynomial have degree d. In other words

$$f(\lambda x_1, \dots, \lambda x_n) = \lambda^d f(x_1, \dots, x_n) \text{ for all } \lambda \in \mathbb{K}.$$

Definition 1.1.15. A polynomial $f \in \mathbb{K}[x_1, \dots, x_n]$ has **multidegree** $d = (d_1, \dots, d_n)$ if $\deg_{x_i}(f) = d_i$ for each $i = 1, \dots, n$.

Definition 1.1.16. A polynomial $f \in \mathbb{K}[x_1, \dots, x_n, y_1, \dots, y_n]$ is called **multi-homogeneous in** $(\mathbf{x_1}, \mathbf{y_1}), \dots, (\mathbf{x_n}, \mathbf{y_n})$ if f is homogeneous in each pair of variables x_i, y_i.

Example 1.1.17. Consider the polynomial $f \in \mathbb{R}[x_1, x_2, y_1, y_2]$ defined by

$$f(x_1, x_2, y_1, y_2) = x_1x_2 - x_1y_2 + 2x_2y_1 - 3y_1y_2.$$

f has total degree 2 and multi-degree $\mathbf{d} = (1, 1, 1, 1)$. It is a homogeneous polynomial and it is multi-homogeneous in the variables (x_1, y_1) and (x_2, y_2). The polynomial f is an example of a special class of polynomials called *multiaffine polynomials* that will be used heavily in this thesis.

Definition 1.1.18. A polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ is called **multiaffine** if it has multi-degree $d = (d_1, \ldots, d_n)$ where each $d_i \leq 1$.

Definition 1.1.19. A **homogeneous ideal** $I \subset \mathbb{K}[x_1, \ldots, x_n]$ is an ideal that is generated by homogeneous polynomials.

Definition 1.1.20. The **projective variety** $\mathcal{V}(J)$ of a homogeneous ideal J in $\mathbb{K}[x_1, \ldots, x_n]$ is the zero set of J, that is

$$\mathcal{V}(J) = \{(a_0 : \ldots : a_n) \in \mathbb{P}^n : f(a_0, \ldots, a_n) = 0 \text{ for all } f \in J\}.$$

As in the affine case, we define the Zariski topology on $\mathbb{P}^n$ by declaring that the projective varieties are closed. We also define the **homogeneous coordinate ring** of a projective variety V as

$$\mathbb{K}[V] = \mathbb{K}[x_0, \ldots, x_n]/\mathcal{I}(V),$$

where $\mathcal{I}(V)$ is the ideal of homogeneous polynomials that vanishes on V. The **field of rational functions** $\mathbb{K}(V)$ is defined as

$$\mathbb{K}(V) = \left\{ \frac{f}{g} : f, g \in \mathbb{K}[V] \text{ such that } f \text{ and } g \text{ are homogeneous of same degree} \right\} \Big/ \sim$$

where $f_1/g_1 \sim f_2/g_2$ if and only if $f_1 g_2 - f_2 g_1 \in \mathcal{I}(V)$. For two projective varieties $X \subset \mathbb{P}^n$ and $Y \subset \mathbb{P}^m$, we define a **rational map** $\varphi : X \dashrightarrow Y$ as

$$\varphi : X \dashrightarrow \mathbb{P}^m \text{ given by } x \longrightarrow (f_0(x) : \ldots : f_m(x))$$

where the image of φ is a subset of Y. We say φ is **regular** at a point $P \in \mathbb{P}^n$, if there exist rational functions $f_i \in \mathbb{K}(X)$ that are regular at P and $f_i(P) \neq 0$ for some i. We define dominant, rational, and birational maps as in the affine case. An example of a morphism map that will be used in Chapter 4 is the *gradient map* of a homogeneous polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]$.

Example 1.1.21. For a homogeneous polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]$, consider the map

$$\nabla h : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1} \quad \text{defined by } p \dashrightarrow \left(\frac{\partial h}{\partial x_1}(p) : \cdots : \frac{\partial h}{\partial x_n}(p) \right).$$

The map ∇h is regular at any point $p \in \mathbb{P}^{n-1}$ such that $p \notin V(h)$.

Definition 1.1.22. A projective variety X is called **rational** if X is birational to $\mathbb{P}^{\dim(X)}$.

An example of a rational variety that will be used in Chapter 2 is the variety of the special linear group defined over $\mathbb{C}$ and denoted by $\mathrm{SL}_2(\mathbb{C})$, this is the group of 2×2 matrices over $\mathbb{C}$ with determinant one.

Example 1.1.23. Consider the rational map

$$\varphi : \mathbb{C}^3 \dashrightarrow \mathrm{SL}_2(\mathbb{C}) \subset \mathbb{C}^4 \quad \text{given by } (a, b, c) \dashrightarrow \left(a, b, c, \frac{1 + bc}{a} \right).$$

Then φ defines a birational map between $\mathrm{SL}_2(\mathbb{C})$ and $\mathbb{C}^3$ and so $\mathrm{SL}_2(\mathbb{C})$ is a rational variety.

Next we give one of the fundamental theorems in computational algebraic geometry that will be used later in Chapter 3. For more in-depth introduction about computational algebraic geometry, the reader is referred to [20, 39].

Theorem 1.1.24. (The Projective Elimination Theorem) *Let $V = \mathcal{V}(I) \subset \mathbb{K}^{n+m}$ for some ideal $I \subset \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Let $I_n = I \cap \mathbb{K}[y_1, \ldots, y_m]$, then*

$$\mathcal{V}(I_n) = \pi\left(\overline{V}\right)$$

where $\pi : \mathbb{P}^n \times \mathbb{K}^m \longrightarrow \mathbb{K}^m$ is the projection and $\overline{V} \subset \mathbb{P}^n \times \mathbb{K}^m$ is the Zariski closure of the image of V under the embedding $\mathbb{K}^{n+m} \hookrightarrow \mathbb{P}^n \times \mathbb{K}^m$.

1.2 Convexity

1.2.1 Convex Sets

In this section we introduce the basic concepts from the theory of convexity with focus on the theorems and definitions needed for Chapter 4. We refer the reader to [28, 65] for more details.

Definition 1.2.1. A set $C \subset \mathbb{R}^n$ is called **convex**, if for all $x, y \in C$ and $0 \le \lambda \le 1$

$$\lambda x + (1 - \lambda) y \in C.$$

The intersection of an arbitrary collection of convex sets is convex.

Definition 1.2.2. The **convex hull** of a set X is the intersection of all convex sets containing X. In other words, it is the convex combinations of points of X

$$\operatorname{conv}(X) = \left\{ \sum_{i=1}^{m} \lambda_i x_i : \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \in \mathbb{R}_{\ge 0} \text{ and } x_i \in X \right\}.$$

Definition 1.2.3. A **hyperplane** of $\mathbb{R}^n$ is a subset $H \subset \mathbb{R}^n$ such that

$$H = \{x \in \mathbb{R}^n : \langle a, x \rangle = c\}$$

for some $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

A fundamental example of convex sets is a **convex polyhedron** which is a set of the form

$$\{x \in \mathbb{R}^n : \langle a_i, x \rangle \le c_i \text{ for } 1 \le i \le m\}$$

where $a_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$. A bounded polyhedron is called a **polytope**.

Definition 1.2.4. A **face** F of a convex set $C \subset \mathbb{R}^n$ is a convex subset of C such that for every pair $x, y \in C$, if any convex combination of x and y, that is not x or y, belongs to F, then x and y belong to F. A face is called **exposed** if $F = C \cap H$ for some hyperplane H. A **vertex** is a face of dimension zero. An **edge** is a face of dimension one.

Definition 1.2.5. For any two sets S and T in $\mathbb{R}^n$, the **Minkowski sum** of S and T is

$$S + T = \{x + y : x \in S, y \in T\}.$$

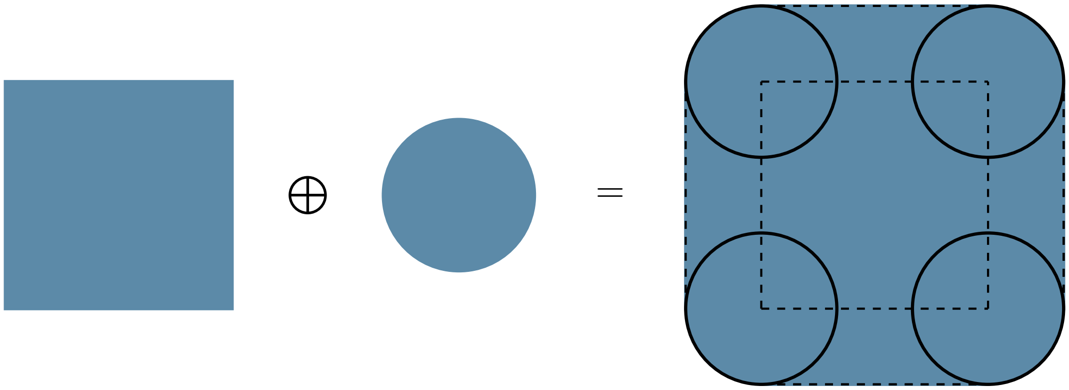

Figure 1.1: The Minkowski sum of a square and a circle

Definition 1.2.6. For $k \geq 0$, the **k-multiple** of a polytope $P = \text{conv}(S)$ is defined as $kP = \text{conv}(kS)$, where we multiply the coordinates of each $s \in S$ by k.

Theorem 1.2.7. *A convex polytope P has a finite number of vertices and it is the convex hull of these vertices. Conversely, the convex hull of a finite set of vertices is a convex polytope.*

A polytope whose vertices all have integer coordinates is called a **lattice polytope**. We will discuss such polytopes in-depth in Chapter 4. One of the fundamental examples of lattice polytopes is the **Newton polytope** of a polynomial f.

Example 1.2.8. Let $f = \sum_{i=1}^{m} c_i \mathbf{x}^{\mathbf{a}_i} \in \mathbb{K}[\mathbf{x}]$ where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{a}_i \in \mathbb{Z}^n_{\geq 0}$. The Newton polytope of f is defined by

$$\text{Newt}(f) = \text{conv}\{\mathbf{a}_i : c_i \neq 0\}.$$

Two special types of polytopes play a dominant role in the construction of toric varieties.

Definition 1.2.9. A lattice polytope P is called **very ample** if, for sufficiently large k, all lattice points in kP are sums of k lattice points of P.

Definition 1.2.10. An n-dimensional lattice polytope P is **smooth** if each vertex v of P is connected to exactly n edges with lattice points closest to v denoted by $v_1, \ldots, v_n$ and the vectors $\{v - v_i : i = 1, \ldots, n\}$ form a basis for the lattice spanned by $P \cap \mathbb{Z}^n$.

Proposition 1.2.11. *Every smooth full-dimensional polytope is very ample.*

1.2.2 Convex Cones

Definition 1.2.12. A set $C \subset \mathbb{R}^n$ is called a **convex cone** if for all $x, y \in C$ and $\lambda, \mu \in \mathbb{R}_{\geq 0}$

$$\lambda x + \mu y \in C.$$

Definition 1.2.13. The **conical hull** of a set X is the intersection of all convex cones containing X. In other words, it is the conical combination of the points of X

$$\mathrm{Cone}(X) = \left\{ \sum_{i=1}^{m} \lambda_i x_i : \lambda_i \geq 0 \text{ and } x_i \in X \right\}.$$

An example of a convex cone that will be used in Chapter 4 is the cone of positive semi-definite matrices. A real symmetric matrix is called **positive semi-definite** if all its eigenvalues are nonnegative. Such symmetric matrices form the positive semi-definite cone.

A **lattice cone**, in some books referred to as **rational cone**, is the conical hull of a finite subset of $\mathbb{Z}^n$. The theory of toric varieties is based on the concept of lattice cones. To each convex cone C we associate a dual cone.

Definition 1.2.14. The **dual cone** of a convex cone C is denoted by C^* and is defined by

$$C^* = \{c \in \mathbb{R}^n : \langle c, x \rangle \geq 0 \text{ for all } x \in C\}.$$

The dual of the positive semi-definite cone for instance, is the positive semi-definite cone itself and we say that the positive semi-definite cone is **self-dual**.

1.3 Convex Algebraic Geometry

This part is about the interplay of algebraic geometry and convex geometry. Many algebraic geometric theorems are proved using combinatorial geometric facts. In the theory of toric varieties for instance, the geometry of the variety is completely determined by the combinatorics of the associated polyhedron.

1.3.1 Toric Varieties

Here we give a very brief introduction of toric varieties, following the survey [51]. For more detailed references the reader is referred to [23, 29, 65]. We will start by defining an algebraic torus which is an essential ingredient in constructing toric varieties.

Definition 1.3.1. An **algebraic torus** denoted by T is an affine variety that is isomorphic to $(\mathbb{C}^*)^n$, where T inherits the group structure defined on $(\mathbb{C}^*)^n$ by the component-wise multiplication.

We will refer to an algebraic torus simply by a torus. To each torus, we associate two lattices M and N that are dual, that is $M = \mathrm{Hom}_{\mathbb{Z}}(N, \mathbb{Z})$.

Definition 1.3.2. A **character** of a torus T is an algebraic group homomophism

$$\chi : T \longrightarrow \mathbb{C}^*.$$

The characters of a torus T form a lattice $M \cong \mathbb{Z}^n$, with the addition operation

$$\chi_1 + \chi_2 : T \longrightarrow \mathbb{C}^* \quad \text{defined by } t \longrightarrow \chi_1(t)\chi_2(t).$$

All characters of a torus have a unique form.

Proposition 1.3.3. *All the characters of a torus T are of the form:*

$$T \longrightarrow \mathbb{C}^* \quad \textit{given by } (t_1, \ldots, t_n) \longrightarrow t_1^{a_1} \cdots t_n^{a_n}$$

for some $(a_1, \ldots, a_n) \in \mathbb{Z}^n$.

Definition 1.3.4. A **one-parameter subgroup** of a torus T is an algebraic group homomorphism

$$\psi : \mathbb{C}^* \longrightarrow T.$$

The set of one-parameter subgroups of $T \cong (\mathbb{C}^*)^n$ form a lattice $N \cong \mathbb{Z}^n$.

Proposition 1.3.5. *All one-parameter subgroups of a torus T are of the form*

$$\mathbb{C}^* \longrightarrow T \quad \textit{given by} \quad t \longrightarrow (t^{b_1}, \dots, t^{b_n})$$

for some $(b_1, \dots, b_n) \in \mathbb{Z}^n$.

Proposition 1.3.6. *Given a group homomorphism*

$$\varphi : T_1 \longrightarrow T_2 \quad \textit{with} \quad \varphi(t) = (t^{a_1}, \dots, t^{a_m}),$$

the image of φ is a torus in T_2 with a character lattice generated by $\{a_1, \dots, a_m\}$.

Now we are ready to define the affine toric variety

Definition 1.3.7. An **affine toric variety** denoted by Y_A, is the Zariski closure of the image of the following map

$$T \longrightarrow \mathbb{C}^m \quad \text{given by } t \longrightarrow (t^{a_1}, \dots, t^{a_m})$$

where $A = \{a_1, \dots, a_m\} \subset M$ is the character lattice associated to T.

Example 1.3.8. The varieties $\mathbb{C}^n$ and $(\mathbb{C}^*)^n$ are toric varieties.

All the toric varieties defined in this thesis are projective and constructed from smooth polytopes, therefore we will focus on such varieties.

Definition 1.3.9. Let T be a torus with two lattices M and N as defined above. A **projective toric variety** is the Zariski closure of the image of the map

$$\varphi : T \longrightarrow (\mathbb{C}^*)^n \subset \mathbb{P}^{n-1} \quad \text{given by} \quad t \longrightarrow (t^{a_1} : \cdots : t^{a_n})$$

for some $A = \{a_1, \dots, a_n\} \subset M$.

We note from the definition that each lattice polytope P gives a projective toric variety X_P. The very ample polytopes give special types of projective varieties.

Theorem 1.3.10. *Let P be a lattice polytope. P is very ample in the lattice it spans if and only if the set of lattice points in P define a normal projective toric variety X_P.*

The smoothness of a toric variety is determined by the smoothness of its polytope.

Theorem 1.3.11. *Let P be a full-dimensional polytope. The variety X_P is smooth if and only if P is smooth.*

1.3.2 Determinantal Representations and Hyperbolicity

Determinantal polynomials are polynomials that can be expressed as determinants of a matrix with linear entries. The problem of characterizing such polynomials is a classical problem of study [8, 24].

Definition 1.3.12. A polynomial p is **determinantal** if it can be written as

$$p = \det(M_0 + M_1x_1 + \cdots + M_nx_n) \text{ for some matrices } M_0, \ldots, M_n \text{ with entries in } \mathbb{C}.$$

We say p is **symmetric (Hermitian) determinantal** if M_i is symmetric (Hermitian) for each i.

Example 1.3.13. Consider the symmetric polynomial $f \in \mathbb{R}[x_1, x_2, x_3, x_4]$

$$f \;=\; x_1x_2x_3x_4 - (x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4) + 2(x_1 + x_2 + x_3 + x_4) - 3.$$

As we will see in Chapter 2, f is symmetric determinantal since f can be written as

$$f = \det \begin{pmatrix} x_1 & 1 & 1 & 1 \\ 1 & x_2 & 1 & 1 \\ 1 & 1 & x_3 & 1 \\ 1 & 1 & 1 & x_4 \end{pmatrix}.$$

Determinantal polynomials play fundamental role in many mathematical areas. See for instance [36, 40, 70]. In this thesis, we study their application to the principal minor assignment problem 1.5. They also form important examples of *hyperbolic polynomials.* These are an interesting class of real multivariate polynomials that have many applications to different fields of mathematics. See for instance [36, 48, 61].

Definition 1.3.14. A homogeneous polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]$ of degree d is called **hyperbolic** with respect to a point $e \in \mathbb{R}^n$ if $h(e) \neq 0$ and all the roots of the univariate polynomial $h(te + a) \in \mathbb{R}[t]$ are real for all $a \in \mathbb{R}^n$.

Helton and Vinnikov [40] prove that hyperbolicity is a necessary and sufficient condition for a plane curve to be definite determinantal. However, this fails for polynomials in more variables [15]. The connected component of $\mathbb{R}^n \setminus \mathcal{V}(h)$ that contains e is a cone and is called **the hyperbolicity cone** of h. For more details, see [14]. We will discuss these polynomials in-depth in Chapter 4. A polynomial that is hyperbolic with respect to any point in the positive orthant $\mathbb{R}^n_{\geq 0}$ is called **stable**. See [71] for an in-depth survey about stable polynomials and their applications.

Definition 1.3.15. Let $\mathcal{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. A polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$ is called **stable** if $f \equiv 0$ or $f(z) \neq 0$ for all $z = (z_1, \ldots, z_n) \in \mathcal{H}^n$.

Example 1.3.16. The polynomial f defined by

$$f(x_1, x_2, x_3) = x_1x_2x_3 - x_1 - x_2 - x_3 + 2$$

is definite determinantal. It can be written as

$$f(x_1, x_2, x_3) = \det \begin{pmatrix} x_1 & 1 & 1 \\ 1 & x_2 & 1 \\ 1 & 1 & x_3 \end{pmatrix}.$$

It is also stable.

Brändén [13] uses *Rayleigh differences* to characterize multiaffine stable polynomials.

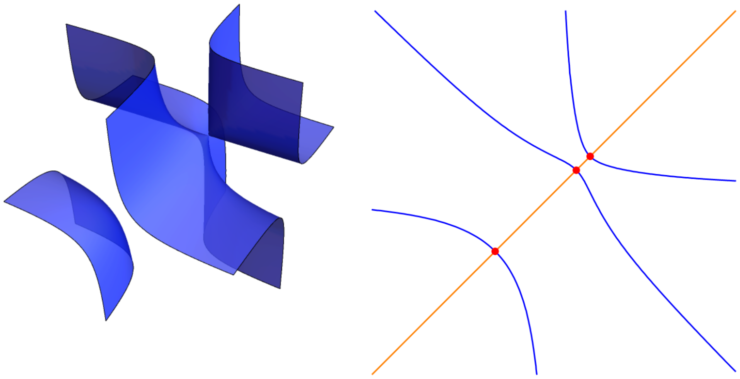

Figure 1.2: The real variety of the stable polynomial f in Example 1.3.16

Definition 1.3.17. Let $f \in \mathbb{R}[x_1, \ldots, x_n]$. The **Rayleigh difference of** f with respect to the variables x_i and x_j is defined as

$$\Delta_{ij}(f) = \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_i} - \frac{\partial^2 f}{\partial x_i \partial x_j} f.$$

Theorem 1.3.18 (Theorem 5.6,[13])**.** *Let* $f \in \mathbb{R}[x_1, \ldots, x_n]$ *be multiaffine. Then* f *is stable if and only if* $\Delta_{ij}(f)(a)$ *is nonnegative for all* $a \in \mathbb{R}^n$.

Kummer, Plaumann and Vinzant [48] use Rayleigh differences to characterize real multiaffine symmetric determinantal polynomials. We will use these polynomials extensively in Chapters 2 and 3 to characterize multiaffine symmetric and Hermitian determinantal polynomials.

1.4 Real Algebraic Geometry

The main objects of study of real algebraic geometry are called *semialgebraic sets*. These are the solutions sets of polynomial equations and inequalities. Real algebraic geometry has many applications to optimization and convex algebraic geometry. See [56] for a survey in this field and [10] for an in-depth introduction.

Definition 1.4.1. A **basic closed semialgebraic set** S is a subset of $\mathbb{R}^n$ that can be written in the form

$$S = \{p \in \mathbb{R}^n : f_1(p) \geq 0, \dots, f_m(p) \geq 0\}$$

for some polynomials $f_1, \dots, f_m \in \mathbb{R}[x_1, \dots, x_n]$.

Definition 1.4.2. A **semialgebraic** set is obtained from a finite number of unions, intersections and complements of basic closed semialgebraic sets.

Example 1.4.3. A fundamental example of a basic semialgebraic set in the field of optimization is the **spectrahedron**. A spectrahedron S is a subset of $\mathbb{R}^n$ defined by

$$S = \{x \in \mathbb{R}^n : A_0 + A_1x_1 + \cdots + A_nx_n \succeq 0\}$$

where A_i is a real symmetric matrix for each $i = 0, \dots, n$. The spectrahedron S is a basic closed semialgebraic set since it can be defined as the solution set of the inequalities formed by the principal minors being nonnegative. A spectrahedron of $n \times n$ positive semidefinite matrices with diagonal entries equal one is known as the *elliptope*. An example of the case $n = 3$ is displayed in Figure 1.3

$$\left\{(x, y, z) \in \mathbb{R}^3 : \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0\right\}.$$

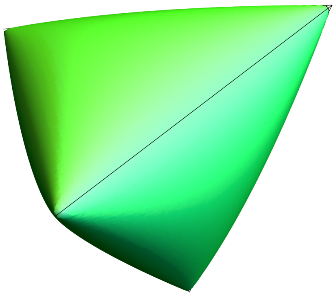

Figure 1.3: The 3×3 Elliptope

1.5 The Principal Minors Problem

Through out this thesis we will denote by R any commutative ring and $[n]$ the set $\{1, \ldots, n\}$. Let $S \subseteq [n]$ and $A \in R^{n\times n}$. The principal minor of A indexed by S and denoted by A_S is the determinant of the submatrix of A indexed on the columns and rows by S. The *principal minor map* is the map that assigns to each matrix A a 2^n-length vector of its principal minor, where we assume that $A_\emptyset = 1$, namely

$$\varphi : R^{n\times n} \longrightarrow R^{2^n} \text{ given by } \varphi(A) = \left(A_\emptyset, A_1, \ldots, A_{[n]}\right).$$

1.5.1 Historical Background

The principal minor map problem has been studied and considered in different versions over time. One version of this problem is to study the relations among the principal minors of a matrix. This dates back to the 19th century, when Nanson [55] and Muir [53] studied the principal minors of 4 × 4 symmetric and general matrices, respectively, using the *cycle sums*. More than hundred years later, Holtz and Sturmfels [41] consider this problem where they give a partial characterization of the image of the principal minor map over $\mathbb{R}$ using *Hadamard-inequalities* and conjectured that the image is cut out by the orbit of one equation called the *Hyperdeterminant* under the action of $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$. This conjecture was resolved set-theoretically by Oeding [57] in 2011 using techniques from representation theory and algebraic geometry. In Chapter 2 we reprove and generalize this result using tools from symbolic computation and classical algebraic geometry.

The general case however remains mysterious. Lin and Sturmfels completed Muir's work [53]. They prove that the image is closed, give an explicit list of generators for the ideal of polynomials vanishing on this image for 4 × 4 complex matrices, and conjecture that for any $n \times n$ matrix, the image is cut out by equations of degree 12. Huang and Oeding [42] solve this conjecture in the special case where all principal minors of same size are equal (*the symmetrized principal minor assignment problem*) where they use the cycle sums in their approach. They provide a minimal parametrization of the respective varieties in the cases of symmetric, skew symmetric and square complex matrices. Kenyon and Pemantle [44] adjust the principal minor map by adding *the almost principal minors* to the vector in the image and show that the ideal of the variety in this case is generated by translations of a single relation using rhombus tilings. We will discuss the general case in-depth in Chapter 3 .

A related problem is to understand the fiber of this map. More explicitly: what is the relationship between two matrices with entries in a field $\mathbb{F}$ and having equal corresponding principal minors? This remains also open in the general case. In the symmetric case, it was solved by Engel and Schneider [27]. In 1984 Loewy and Hartfiel [38] and then Loewy [50] gave sufficient conditions for two general matrices A and B to be *diagonally similar* and hence to belong to the same fiber, and they provided examples to prove that the converse is not true. Boussaıria and Cherguia [1] give a characterization of the fiber in the skew symmetric case in the special case where all non-diagonal entries are nonzero.

Another interesting problem is the problem of finding an efficient algorithm for reconstructing the matrix given the vector of its principal minors. Griffin and Tsatsomeros [34, 35] give a numerical algorithm that reconstructs a preimage, if it exists, over $\mathbb{C}$. Rising, Kulesza and Taskarc [62] provide an efficient algorithm for reconstruction in the symmetric case. Brunel, Moitra, Rigollet and Urschel [67] use the *graph cycle* to reconstruct symmetric matrices with increased efficiency.

1.5.2 Applications of the Principal Minor Problem

The principal minor map problem has many applications and in different fields, including statistical models, machine learning, combinatorics and matrix theory. One of its fundamental applications is to Determinantal Point Processes (DPP). These are probabilistic models that arise naturally in the study of random matrix theory [43] and machine learning [19, 31]. One way to define a DPP (or what is known as symmetric DPP) is through an $n \times n$ positive semi-definite matrix whose eigenvalues lie in the interval $[0, 1]$ and which is known as the *the kernel* of the model. The DPP associated to such matrix K is given by the distribution on

$Y \subseteq [n]$ such that

$$\mathbb{P}[J \subseteq Y] = \det(K_J)$$

for any $J \subseteq [n]$. Symmetric DPPs have attracted a lot of attention as they reflect repulsive behavior in modeling. For more in-depth details the reader is referred to [2, 12, 26, 47, 67]. Recently the case of non-symmetric kernel has started to get more attention in order to model both repulsive and attractive interactions, which can significantly improve modeling power [6, 17, 30].

Chapter 2

CHARACTERIZING PRINCIPAL MINORS OF SYMMETRIC MATRICES VIA DETERMINANTAL MULTIAFFINE POLYNOMIALS

The work presented in this chapter is based on a joint project with Cynthia Vinzant [4]

2.1 Introduction

In this chapter we will restrict the map of pricipal minors to the space of symmetric matrices $\mathrm{Sym}_n(R)$:

$$\varphi : \mathrm{Sym}_n(R) \to R^{2^n} \quad \text{given by} \quad \varphi(A) = (A_S)_{S\subseteq[n]},$$

where we take $A_\emptyset = 1$. Here we seek to characterize the image of the principal minor map over arbitrary unique factorization domains R, and in particular, arbitrary fields. For a background of the problem, the reader is referred to 1.5. In this chapter we generalize Oedings result [57] to hold over arbitrary unique factorization domains, except those with exactly three elements.

We will study this problem by associating to the matrix A the multiaffine polynomial

$$f_A = \det\left(\mathrm{diag}(x_1, \dots, x_n) + A\right) = \sum_{S\subseteq[n]} A_S \prod_{i\notin S} x_i.$$

This translates the problem of characterizing the image of the principal minor map to the problem of characterizing multiaffine polynomials in $R[x_1, \dots, x_n]$ with symmetric determinantal representations. Key to this characterization will be Rayleigh differences.

The Rayleigh difference of a polynomial f with respect to $i, j \in [n]$ is defined to be

$$\Delta_{ij}(f) = \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} - f\frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{2.1}$$

These polynomials play a prominent role in the theory of stable polynomials [13]. Using Dodgson condensation [25], one can see that for the determinantal polynomial f_A, all Rayleigh differences $\Delta_{ij}(f_A)$ are squares in the polynomial ring $R[x_1, \dots, x_n]$. In 2015, Kummer, Plaumann and the second author prove the converse over $\mathbb{R}$ [48] and here we prove it over an arbitrary unique factorization domain.

Formally, to $\mathbf{a} = (a_S)_{S \subseteq [n]}$ in R^{2^n} we associate the polynomial $f_{\mathbf{a}} = \sum_{S \subseteq [n]} a_S \prod_{i \notin S} x_i$.

Theorem 2.3.5. Let R be a unique factorization domain. An element $\mathbf{a} = (a_S)_{S \subseteq [n]}$ in R^{2^n} is in the image of $\mathrm{Sym}_n(R)$ under the principal minor map if and only if $a_\emptyset = 1$ and for every $i, j \in [n]$, $\Delta_{ij}(f_{\mathbf{a}})$ is a square in $R[x_1, \dots, x_n]$.

For $n = 3$, $\Delta_{12}(f_{\mathbf{a}})$ is a quadratic polynomial in the remaining variable x_3, namely

$$\Delta_{12}(f_{\mathbf{a}}) = (a_1a_2 - a_\emptyset a_{12})x_3^2 + (a_1a_{23} + a_2a_{13} - a_3a_{12} - a_\emptyset a_{123})x_3 + (a_{13}a_{23} - a_3a_{123}).$$

For this polynomial to be a square, its discriminant must vanish, giving us a necessary equation on the coefficients of $f_{\mathbf{a}}$. The discriminant of $\Delta_{12}(f_{\mathbf{a}})$ with respect to x_3 equals the well-known Cayley $2 \times 2 \times 2$ hyperdeterminant,

$$\begin{aligned}\mathrm{HypDet}(\mathbf{a}) &= (a_1a_{23} + a_2a_{13} - a_3a_{12} - a_\emptyset a_{123})^2 - 4(a_1a_2 - a_\emptyset a_{12})(a_{13}a_{23} - a_3a_{123}) \\ &= a_\emptyset^2 a_{123}^2 + a_1^2a_{23}^2 + a_2^2a_{13}^2 + a_3^2a_{12}^2 - 2a_\emptyset a_1a_{23}a_{123} - 2a_\emptyset a_2a_{13}a_{123} - 2a_\emptyset a_3a_{12}a_{123} \\ &\quad - 2a_1a_2a_{13}a_{23} - 2a_1a_3a_{12}a_{23} - 2a_2a_3a_{12}a_{13} + 4a_\emptyset a_{23}a_{13}a_{12} + 4a_{123}a_1a_2a_3.\end{aligned}$$

The coefficients of 1 and x_3^2 in $\Delta_{12}(f_{\mathbf{a}})$ are $a_{13}a_{23} - a_3a_{123}$ and $a_1a_2 - a_\emptyset a_{12}$, respectively. We see that $\Delta_{12}(f_{\mathbf{a}})$ is a square if and only if these two coefficients are squares in R and the discriminant, $\mathrm{HypDet}(\mathbf{a})$, is zero. One can check that $\mathrm{Discr}_{x_3}\Delta_{12}(f_{\mathbf{a}})$, $\mathrm{Discr}_{x_2}\Delta_{13}(f_{\mathbf{a}})$ and $\mathrm{Discr}_{x_1}\Delta_{23}(f_{\mathbf{a}})$ are all the same and equal to $\mathrm{HypDet}(\mathbf{a})$. Therefore a vector $\mathbf{a} \in R^{2^3}$ with $a_\emptyset = 1$ is in the image of the principal minor map if and only if $\mathrm{HypDet}(\mathbf{a}) = 0$ and for every $i, j \in [3]$ with $\{k\} = [3] \backslash \{i, j\}$, both $a_{ik}a_{jk} - a_ka_{ijk}$ and $a_ia_j - a_\emptyset a_{ij}$ are squares in R.

Our main result is that, under the action of $\mathrm{SL}_2(R)^n \rtimes S_n$, these conditions characterize the image of the principal minor map for general n.

Theorem 2.5.1. Let R be a unique factorization domain with $|R| \neq 3$ and $\mathbf{a} = (a_S)_{S \subseteq [n]} \in R^{2^n}$ with $a_\emptyset = 1$. There exists a symmetric matrix over R with principal minors $\mathbf{a}$ if and only if

(i) for every $i, j \in [n]$, $a_i a_j - a_{ij}$ is a square in R, and

(ii) for every $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$.

While the description in (ii) involves a potentially infinite set of quartic polynomials, we give an explicit set of $\binom{n}{3}5^{n-3}$ elements $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$ that are necessary and sufficient in this characterization (see Remark 2.5.2). As observed in [58, Observation III.15], when R is a field of characteristic zero, this is precisely the dimension of the linear space in $R[a_S : S \subseteq [n]]$ spanned by the polynomials $(\gamma \cdot \mathrm{HypDet})(\mathbf{a})$.

As a corollary of Theorem 2.5.1, we obtain another proof of Oeding's result over $\mathbb{C}$:

Corollary 2.5.3. Let $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{C}^{2^n}$ with $a_\emptyset = 1$. Then $\mathbf{a}$ belongs to the image of the principal minor map over $\mathbb{C}$ if and only if for every $\gamma \in \mathrm{SL}_2(\mathbb{C})^n \rtimes S_n$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$.

We also get a semialgebraic description of the image of the principal minor map over $\mathbb{R}$.

Corollary 2.5.4. Let $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{R}^{2^n}$ with $a_\emptyset = 1$. Then $\mathbf{a}$ belongs to the image of the principal minor map over $\mathbb{R}$ if and only if for every $i, j \in [n]$, $a_i a_j - a_{ij} \geq 0$ and for every $\gamma \in \mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$.

For real symmetric matrices, the inequalities $A_i A_j - A_{ij} \geq 0$ are a subset of the well-known Hadamard-Fischer inequalities $A_{S \cup i} A_{S \cup j} - A_S A_{S \cup ij} \geq 0$, which were also used by Holtz and Sturmfels in a partial characterization of the image of the principal minor map over $\mathbb{R}$, [41, Theorem 6]. Corollary 2.5.4 states that these inequalities and the equations given by the images of the $2 \times 2 \times 2$ hyperdeterminant under $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$ cut out the image of the principal minor map over $\mathbb{R}$. The image of the principal minor map over $\mathbb{R}$ is of special interest, as for positive semidefinite matrices A, the discrete probability measure on $2^{[n]}$ given by $\mathrm{Prob}(S) \propto A_S$ forms a determinantal point process. These distributions have several nice properties, such as negative association, and appear in a wide range of applications [11, 46].

Over fields of characteristic two, the discriminant of a univariate quadratic is a square. From Theorem 2.5.1, we then recover the results of van Geeman and Marrani [68] that the image is cut out by quadratic equations:

Corollary 2.5.5. Let $\mathbf{a} = (a_S)_{S\subseteq[n]} \in R^{2^n}$ with $a_\emptyset = 1$ where R has characteristic two. There exists a symmetric matrix over R with principal minors $\mathbf{a}$ if and only if

(i) for every $i, j \in [n]$, $a_i a_j - a_{ij}$ is a square in R, and

(ii) for every $\gamma \in \mathrm{SL}_2(\mathbb{F}_2)^n \rtimes S_n$, $\gamma \cdot (a_\emptyset a_{123} + a_1 a_{23} + a_2 a_{13} + a_3 a_{12}) = 0$.

In particular, for $R = \mathbb{F}_2$, (i) is always satisfied and the image of the principal minor map is cut out by the quadratic equations in (ii).

The chapter is organized as follows. In Section 2.2, we establish notation and introduce the action of $\mathrm{SL}_2(R)^n \rtimes S_n$. In Section 2.3, we establish the connection between square Rayleigh differences and determinantal representations and prove Theorem 2.3.5. In Section 2.4, we use the group action of $\mathrm{SL}_2(R)^n \rtimes S_n$ to characterize the set multiquadratic squares in $R[x_1, \ldots, x_n]$ and use this to characterize the image of the principal minor map in Section 2.5. To conclude, in Section 3.5, we discuss some consequences for other determinantal representations as well as connections to the Grassmannian $\mathrm{Gr}_{\mathbb{F}}(d, n)$ over arbitrary fields.

2.2 Background and notation

Throughout the chapter, we take R to be a unique factorization domain. Let $R[\mathbf{x}]$ denote the polynomial ring $R[x_1, \ldots, x_n]$. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $S \subseteq [n]$, we use the notation $\mathbf{x}^\alpha$ for $\prod_{i=1}^n x_i^{\alpha_i}$ and $\mathbf{x}^S$ for $\prod_{i\in S} x_i$. For $f \in R[\mathbf{x}]$, let $\deg_i(f)$ denote the degree of f in the variable x_i. Given $\mathbf{d} = (\mathbf{d}_1, \ldots, \mathbf{d}_n) \in \mathbb{Z}_{\geq 0}^n$, let $R[\mathbf{x}]_{\leq \mathbf{d}}$ denote the set of polynomials with degree at most $\mathbf{d}_i$ in x_i for each $i = 1, \ldots, n$. These form an R-module of rank $\prod_{i=1}^{n}(\mathbf{d}_i+1)$. When $\mathbf{d}_1 = \ldots = \mathbf{d}_n = m$, we abbreviate $R[\mathbf{x}]_{\leq(m,\ldots m)}$ by $R[\mathbf{x}]_{\leq \mathbf{m}}$. Of particular interest are *multiaffine polynomials*, $R[\mathbf{x}]_{\leq \mathbf{1}}$, with degree ≤ 1 in each variable, and *multiquadratic polynomials*, $R[\mathbf{x}]_{\leq \mathbf{2}}$, with degree ≤ 2 in each variable. We will often consider multi-homogenezations of these polynomials. Let $R[\mathbf{x},\mathbf{y}]_{\mathbf{d}}$ denote the set of polynomials in the variables $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ that are homogeneous of degree $\mathbf{d}_i$ in each pair of variables x_i, y_i. For $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha$, let $f^{\mathbf{d}-\mathrm{hom}}$ in $R[\mathbf{x},\mathbf{y}]_{\mathbf{d}}$ denote the polynomial

$$f^{\mathbf{d}-\mathrm{hom}} = \prod_{i=1}^{n} y_i^{\mathbf{d}_i} \cdot f\left(x_1/y_1,\ldots,x_n/y_n\right) = \sum_\alpha c_\alpha \mathbf{x}^\alpha \mathbf{y}^{\mathbf{d}-\alpha}.$$

To a polynomial $f \in R[\mathbf{x}]_{\leq \mathbf{2}}$, its discriminant with respect to any variable x_k, denoted $\mathrm{Discr}_{x_k}(f)$, equals b^2-4ac where $f = ax_k^2+bx_k+c$ and a, b, c do not involve the variable x_k. Similarly, for a multiquadratic polynomial $f \in R[\mathbf{x},\mathbf{y}]_{\mathbf{2}}$, we can write $f = ax_k^2+bx_ky_k+cy_k^2$ and define its discriminant with respect to (x_k,y_k) to be $\mathrm{Discr}_{(x_k,y_k)}(f) = b^2-4ac$.

The symmetric group acts on $R[\mathbf{x}]$ by permuting the variables. That is, for $\pi \in S_n$, $\pi \cdot f$ equals $f(x_{\pi(1)},\ldots,x_{\pi(n)})$. The action of $\mathrm{SL}_2(R)^n$ on $R[\mathbf{x}]_{\leq \mathbf{d}}$ is defined as follows. Let $\gamma = (\gamma_i)_{i\in[n]}$ in $\mathrm{SL}_2(R)^n$ where $\gamma_i = \left(\begin{smallmatrix} a_i & b_i \\ c_i & d_i \end{smallmatrix}\right)$. Then for $f \in R[\mathbf{x}]_{\leq \mathbf{d}}$,

$$\gamma \cdot f = \prod_{i=1}^{n}(c_ix_i+d_i)^{\mathbf{d}_i} \cdot f\left(\frac{a_1x_1+b_1}{c_1x_1+d_1},\ldots,\frac{a_nx_n+b_n}{c_nx_n+d_n}\right).$$

One way to interpret this action is via the multi-homogenezation of f. The induced action of γ on $f^{\mathbf{d}-\mathrm{hom}}$ is just an R-linear change of coordinates:

$$\gamma \cdot f^{\mathbf{d}-\mathrm{hom}} = f^{\mathbf{d}-\mathrm{hom}}\left(\gamma_1 \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix},\ldots,\gamma_n \cdot \begin{pmatrix} x_n \\ y_n \end{pmatrix}\right).$$

Restricting to $y_1 = \ldots = y_n = 1$ gives back $\gamma \cdot f$. Similarly, we can extend the action of S_n to $R[\mathbf{x},\mathbf{y}]_{\mathbf{d}}$ by simultaneous permutations of the x_i and y_i coordinates, i.e. $\pi \cdot f = f(x_{\pi(1)},\ldots,x_{\pi(n)},y_{\pi(1)},\ldots,y_{\pi(n)})$.

Note that $R[\mathbf{x}]_{\leq \mathbf{1}}$ and R^{2^n} are isomorphic R-modules, and so the action of $\mathrm{SL}_2(R)^n \rtimes S_n$ on $R[\mathbf{x}]_{\leq \mathbf{1}}$ also gives one on R^{2^n}. Specifically, to an element $\mathbf{a} = (a_S)_{S\subseteq[n]}$ in R^{2^n} we associate the multiaffine polynomial $f_{\mathbf{a}} = \sum_{S\subseteq[n]} a_S \mathbf{x}^{[n]\setminus S}$ and to any polynomial $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$ we associate the point $\mathbf{a} = (a_S)_{S \subseteq [n]}$ in R^{2^n} with $a_S = \mathrm{coeff}(f, \mathbf{x}^{[n]\setminus S})$. Note that if A is a symmetric matrix with $a_S = A_S$, then $f_{\mathbf{a}} = \det\left(\mathrm{diag}(x_1, \ldots, x_n) + A\right)$. For any $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$, we define $\gamma \cdot \mathbf{a}$ by the relation $f_{\gamma \cdot \mathbf{a}} = \gamma \cdot f_{\mathbf{a}}$. Similarly, we define the action of $\mathrm{SL}_2(R)^n \rtimes S_n$ on the polynomial ring $R[a_S : S \subseteq [n]]$ by $\gamma \cdot F(\mathbf{a}) = F(\gamma \cdot \mathbf{a})$.

Example 2.2.1. For $n = 3$, consider $\gamma = \left(\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), \mathrm{Id}_2, \mathrm{Id}_2\right)$ in $\mathrm{SL}_2(R)^3$. For any point $\mathbf{a} = (a_S)_{S \subseteq [3]} \in R^{2^3}$,

$$\gamma \cdot f_{\mathbf{a}} = x_1 f_{\mathbf{a}}(-x_1^{-1}, x_2, x_3) = \sum_{T \ni 1} a_T \mathbf{x}^{[3]\setminus(T\setminus 1)} - \sum_{T \not\ni 1} a_T \mathbf{x}^{[3]\setminus(T\cup 1)} = \sum_{S \not\ni 1} a_{S\cup 1}\mathbf{x}^{[3]\setminus S} - \sum_{S \ni 1} a_{S\setminus 1}\mathbf{x}^{[3]\setminus S}.$$

Taking coefficients of $\gamma \cdot f_{\mathbf{a}}$ shows that $(\gamma \cdot \mathbf{a})_S$ equals $a_{S\cup 1}$ if $1 \notin S$ and $-a_{S\setminus 1}$ if $1 \in S$. For $F(\mathbf{a}) = a_2 a_3 - a_\emptyset a_{23}$, we see that $\gamma \cdot F(\mathbf{a}) = F(\gamma \cdot \mathbf{a}) = a_{12}a_{13} - a_1 a_{123}$. From this we see that the image of F under the group $\mathrm{SL}_2(R)^3 \rtimes S_3$ includes all six polynomials of the form $a_i a_j - a_\emptyset a_{ij}$ and $a_{ik}a_{jk} - a_k a_{ijk}$ for $\{i, j, k\} = \{1, 2, 3\}$.

Proposition 2.2.2. *Consider an element $\gamma \in \mathrm{SL}_2(R)^n$ that acts by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in the k-th coordinate and the identity in all others. For any $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$,*

$$\Delta_{ij}(\gamma \cdot f) = \begin{cases} \Delta_{ij}(f) & \text{if } k = i, j \\ \gamma \cdot \Delta_{ij}(f) & \text{otherwise.} \end{cases}$$

Proof. For each $k \in [n]$, let f_k denote the derivative of f with respect to x_k and let f^k denote its specialization to $x_k = 0$. We can then write $f = x_k f_k + f^k$. Then

$$\gamma \cdot f \;=\; (a x_k + b) f_k + (c x_k + d) f^k \;=\; x_k(a f_k + c f^k) + (b f_k + d f^k).$$

In particular, $\frac{\partial}{\partial x_k}(\gamma \cdot f) = (a f_k + c f^k)$ and $(\gamma \cdot f)|_{x_k = 0} = b f_k + d f^k$.

To see how this action affects Rayleigh differences, we write the polynomial $\Delta_{ij}(f)$ as

$$\Delta_{ij}(f) = f_i f_j - f f_{ij} = f_i^j f_j^i - f^{ij} f_{ij},$$

where f_i^j for example denotes $\frac{\partial f}{\partial x_i}|_{x_j = 0}$. If $k = i$, applying γ then gives

$$\Delta_{ij}(\gamma \cdot f) = (a f_i^j + c f^{ij})(b f_{ij} + d f_j^i) - (b f_i^j + d f^{ij})(a f_{ij} + c f_j^i) = (ad - bc)(f_i^j f_j^i - f^{ij} f_{ij}),$$

showing that $\Delta_{ij}(f)$ is invariant. Another way to see this is to view $\Delta_{ij}(f)$ as the resultant of f_j and f^j with respect to x_i. Note that for $k \neq i, j$, γ commutes with taking the derivatives with respect to x_i, x_j and restricting x_i, x_j to zero. Therefore $\Delta_{ij}(\gamma \cdot f) = \gamma \cdot \Delta_{ij}(f)$. $\square$

Corollary 2.2.3. *The set of polynomials $f \in R[\mathbf{x}]_{\leq 1}$ such that $\Delta_{ij}(f)$ is a square for all $i, j \in [n]$ is invariant under the action of* $\mathrm{SL}_2(R)^n \rtimes S_n$.

Proof. Note that the set of squares in $R[\mathbf{x}]_{\leq \mathbf{2}}$ is invariant under the action of $\mathrm{SL}_2(R)^n \rtimes S_n$. If $g = h^2$ where $h \in R[\mathbf{x}]_{\leq \mathbf{1}}$ then for $\pi \in S_n$, $\pi \cdot g = (\pi \cdot h)^2$. Similarly for $\gamma \in \mathrm{SL}_2(R)^n$, $\gamma \cdot g = (\gamma \cdot h)^2$. Note here that γ acts on g as an element of $R[\mathbf{x}]_{\leq \mathbf{2}}$ and acts on h as an element of $R[\mathbf{x}]_{\leq \mathbf{1}}$, regardless of their degrees.

First we note invariance under the symmetric group. For any $\pi \in S_n$, $\Delta_{ij}(\pi \cdot f)$ equals $\pi \cdot \Delta_{\pi^{-1}(i)\pi^{-1}(j)}(f)$. Therefore if f has the property that $\Delta_{ij}(f)$ is a square for all i, j, then so does $\pi \cdot f$. Similarly, by Proposition 2.2.2, for any $\gamma \in \mathrm{SL}_2(R)^n$, if $\Delta_{ij}(f) \in R[\mathbf{x}]_{\leq \mathbf{2}}$ is a square, then so is $\Delta_{ij}(\gamma \cdot f)$. $\square$

We will also use the usual homogenization of a polynomial to some total degree d, using a single homogenizing variable y. That is, for $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in R[\mathbf{x}]$ of total degree $d = \deg(f)$, its homogenization is

$$\overline{f} = y^d f\left(x_1/y, \ldots, x_n/y\right) = \sum_\alpha c_\alpha \mathbf{x}^\alpha y^{d-|\alpha|} \ \in \ R[\mathbf{x}, y].$$

To end this section, we remark that the condition that $\Delta_{ij}(f)$ is a square is robust to various homogenizations.

Proposition 2.2.4. *Let $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$ and let $\overline{f}$ denote the homogenization of f in $R[\mathbf{x}, y]$. Then the following are equivalent*

(a) $\Delta_{ij}(f)$ *is a square in* $R[\mathbf{x}]$, (b) $\Delta_{ij}(\overline{f})$ *is a square in* $R[\mathbf{x}, y]$,

(c) $\overline{\Delta_{ij}(f)}$ *is a square in* $R[\mathbf{x}, y]$, (d) $(\Delta_{ij}(f))^{\mathbf{2}-\mathrm{hom}}$ *is a square in* $R[\mathbf{x}, \mathbf{y}]_{\mathbf{2}}$.

Proof. The implications (b)⇒(a), (c)⇒(a), and (d)⇒(a) follow from restricting to $y = 1$ or $y_1 = \ldots = y_n = 1$. For (a)⇒(c) and (a)⇒(d), we note that if $\Delta_{ij}(f) = g^2$ for some

$g \in R[\mathbf{x}]$, then $\overline{\Delta_{ij}(f)} = (\overline{g})^2$ and $(\Delta_{ij}(f))^{\mathbf{2}-\mathrm{hom}} = (g^{\mathbf{1}-\mathrm{hom}})^2$. For (a)⇒(b), let $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$ with total degree d and suppose that $\Delta_{ij} f = g^2$ for some $g \in R[\mathbf{x}]$. Let $m = \deg(g)$. By definition, $\Delta_{ij}(\overline{f}) \in R[\mathbf{x}, y]$ is homogeneous of degree $2d-2$. Its restriction to $y = 1$ equals $\Delta_{ij} f$. Therefore $\Delta_{ij}(\overline{f})$ equals $y^{2d-2-2m}\overline{\Delta_{ij}(f)}$, showing that $\Delta_{ij}(\overline{f})$ can be written as $(y^{d-1-m}\overline{g})^2$. □

2.3 Squares to Determinantal Representations

In this section we prove Theorem 2.3.5. This relies heavily on the structure of the polynomials $\Delta_{ij}(f)$ defined in (2.1). If f is multiaffine, then $\Delta_{ij}(f)$ does not involve the variables x_i and x_j and has degree ≤ 2 in the other variables. In particular, if $\Delta_{ij}(f) = (g_{ij})^2$ for some $g_{ij} \in R[x_1, \ldots, x_n]$, then g_{ij} does not involve the variables x_i and x_j and has degree ≤ 1 in the rest. We first work with more general determinantal representations in a larger ring $R[x_1, \ldots, x_m] = R[x_{n+1}, \ldots, x_m][\mathbf{x}]$ with $n < m$.

Theorem 2.3.1. *Let $f \in R[x_1, \ldots, x_m]$ be a homogeneous polynomial of degree $n < m$. Suppose f is multiaffine in the variables $x_1, \ldots, x_n$ and its coefficient of $x_1 \cdots x_n$ equals one. Then $f = \det(\mathrm{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_j M_j)$ for some $M_j \in \mathrm{Sym}_n(R)$ if and only if $\Delta_{ij} f$ is a square in $R[x_1 \ldots, x_m]$ for all $1 \leq i, j \leq n$.*

Proof of (⇒). This follows from a classical equality on the principal minors of an $n \times n$ matrix, used by Dodgson [25] as a method for computing determinants. For subsets $S, T \subset [n]$ of equal cardinality, let $M(S,T)$ denote the submatrix of M obtained by *dropping* rows S and columns T from M. Then for any $i \neq j \in [n]$,

$$\det(M(i,i)) \cdot \det(M(j,j)) - \det(M)\det(M(\{i,j\},\{i,j\})) = \det(M(i,j)) \cdot \det(M(j,i)). \quad (2.2)$$

Note that for $M = \mathrm{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_j M_j$ and any subset $S \subseteq [n]$, the principal minor $\det(M(S,S))$ equals the derivative of f with respect to the variables in S, $\left(\prod_{i \in S} \frac{\partial}{\partial x_i}\right) f$. The equation above then gives that $\Delta_{ij}(f)$ equals $\det(M(i,j)) \cdot \det(M(j,i))$. Since M is symmetric, this shows that $\Delta_{ij}(f) = (\det(M(i,j)))^2$. □

We prove the other direction of this theorem after the following lemma.

Lemma 2.3.2. *Let* $f \in R[x_1, \ldots, x_m]$ *be multiaffine in the variables* $x_1, \ldots, x_n$ *and its coefficient of* $x_1 \cdots x_n$ *equal one. If* $f = g \cdot h$ *for some* $g, h \in R[x_1, \ldots, x_m]$*, then* g *and* h *are multiaffine in disjoint subsets of the variables* $x_1, \ldots, x_n$ *and we can take their leading coefficients in these variables to be one. Moreover,* $\Delta_{ij} f$ *is a square if and only if* $\Delta_{ij} g$ *and* $\Delta_{ij} h$ *are squares.*

Proof. For any $i \in [n]$, the degree of f in x_i must be the sum of the degrees of g and h in x_i. Since this sum of nonnegative numbers is one for each $i \in [n]$, we see that for some subset $I \subseteq [n]$, g is multiaffine in $\{x_i : i \in I\}$, h is multiaffine in $\{x_j : j \notin I\}$, and $\deg_i(h) = \deg_j(g) = 0$ for any $i \in I$ and $j \notin I$.

The highest degree term in f with respect to the variables $x_1, \ldots, x_n$, $\prod_{i=1}^n x_i$, is the product of the highest degree terms in g and h. Therefore for some $r, s \in R$, these terms are $r \prod_{i \in I} x_i$ and $s \prod_{j \notin I} x_j$, respectively. Since $rs = 1$, we can replace g with sg and h with rh to obtain a factorization in which both have leading coefficient equal to 1.

For $i \in I$, $\partial(g \cdot h)/\partial x_i = h \cdot \partial g/\partial x_i$ and similarly, for $j \notin I$, $\partial(g \cdot h)/\partial x_j = g \cdot \partial h/\partial x_j$. From this, one can check that $\Delta_{ij}(gh)$ equals $h^2 \Delta_{ij}(g)$ for $i, j \in I$, $g^2 \Delta_{ij}(h)$ for $i, j \in [n] \backslash I$ and zero otherwise. In each case, we see that $\Delta_{ij}(gh)$ is a square in $R[x_1, \ldots, x_m]$ if and only if both $\Delta_{ij}(g)$ and $\Delta_{ij}(h)$ are squares. □

Proof of Theorem 3.3.1(⇐). Let S denote the ring $R[x_1, \ldots, x_m]$. Suppose that f is irreducible in S. For each $i \in [n]$, let $g_{ii} = \frac{\partial f}{\partial x_i}$ and for each $i < j$, suppose that $\Delta_{ij} f$ equals $(g_{ij})^2$ for some $g_{ij} \in S$. This implies that $\frac{\partial f}{\partial x_i} \cdot \frac{\partial f}{\partial x_j}$ is equivalent to $(g_{ij})^2$ modulo $\langle f \rangle$. For $1 < i < j$, the polynomials $(g_{11} g_{ij})^2$ and $(g_{1j} g_{i1})^2$ are both equivalent to $(\frac{\partial f}{\partial x_1})^2 \frac{\partial f}{\partial x_i} \cdot \frac{\partial f}{\partial x_j}$, showing that

$$(g_{11} g_{ij} - g_{1j} g_{i1})(g_{11} g_{ij} + g_{1j} g_{i1}) = (g_{11} g_{ij})^2 - (g_{1j} g_{i1})^2 \equiv 0 \mod \langle f \rangle.$$

Since f is irreducible, $S/\langle f \rangle$ is an integral domain. Therefore one of the two factors above must be zero in $S/\langle f \rangle$. After changing the sign of g_{ij} if necessary, we can assume that it is the first factor, giving that $g_{11}g_{ij} - g_{1j}g_{i1} \in \langle f \rangle$. Let $G \in \mathrm{Sym}_n(S)$ be the symmetric matrix with (i,j)th entry $g_{ij} = g_{ji}$. We claim that the 2×2 minors of G lie in $\langle f \rangle$. Note that by construction, for any $i, j, k, l \in [n]$,

$$g_{11}^2(g_{ij}g_{kl} - g_{il}g_{kj}) = (g_{11}g_{ij})(g_{11}g_{kl}) - (g_{11}g_{il})(g_{11}g_{kj}) \equiv g_{1i}g_{1j}g_{1k}g_{1l} - g_{1i}g_{1l}g_{1k}g_{1j} = 0 \mod \langle f \rangle.$$

Since f is irreducible and $g_{11} = \partial f / \partial x_1$ has smaller degree, g_{11} is not a zero-divisor in $S/\langle f \rangle$. Therefore the minor $g_{ij}g_{kl} - g_{il}g_{kj}$ belongs to $\langle f \rangle$.

From this it follows that f^{k-1} divides the $k \times k$ minors of G for every $2 \le k \le n$, see [59, Lemma 4.7]. In particular, f^{n-2} divides the entries of the adjugate matrix G^{adj}. Let

$$M = (1/f^{n-2}) \cdot G^{\mathrm{adj}}.$$

Also f^{n-1} divides $\det(G)$, and since these both have degree $n(n-1)$, there must be some constant $\lambda \in R$ for which $\det(G) = \lambda \cdot f^{n-1}$.

We can see that $\lambda = 1$ by specializing x_ℓ to 0 for all $\ell > n$. For any polynomial $h \in S$, let $h(\mathbf{x}, 0)$ denote the specialization of h with $x_\ell = 0$ for $\ell = n+1, \ldots, m$. Then $f(\mathbf{x}, 0)$ equals $x_1 \cdots x_n$ and $g_{ii}(\mathbf{x}, 0) = \prod_{j \neq i} x_j$. Recall that $g_{ij} \in R[x_k : k \neq i, j]$ has total degree $n-1$ and degree at most one in each variable x_k for $k \in [n] \backslash \{i, j\}$. Therefore every monomial appearing in g_{ij} with non-zero coefficient must involve a variable x_ℓ for $\ell > n$, giving that $g_{ij}(\mathbf{x}, 0) = 0$. Specializing all entries of G to $x_\ell = 0$ for $\ell > n$, gives the diagonal matrix

$$G(\mathbf{x}, 0) = \mathrm{diag}\left(\prod_{j \neq 1} x_j, \ldots, \prod_{j \neq n} x_j\right) = \prod_{j=1}^{n} x_j \cdot \mathrm{diag}\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right).$$

Its determinant is $\prod_{i=1}^{n} x_i^{n-1}$ which equals $f(\mathbf{x}, 0)^{n-1}$, showing that $\lambda = 1$. From this and the equation $G \cdot G^{\mathrm{adj}} = \det(G) \cdot \mathrm{Id}_n$, it follows that

$$\det(M) = \frac{1}{f^{n(n-2)}} \cdot \det(G^{\mathrm{adj}}) = \frac{1}{f^{n(n-2)}} \det(G)^{n-1} = \frac{1}{f^{n(n-2)}} f^{(n-1)^2} = f.$$

Note that the entries of M have degree $\le (n-1)^2 - n(n-2) = 1$, so we can write M as $\sum_{i=1}^{m} x_i M_i$ for some matrices $M_i \in \mathrm{Sym}_n(R)$. To finish the proof it suffices to show that

$\sum_{i=1}^n x_i M_i = \operatorname{diag}(x_1, \ldots, x_n)$. Indeed, using the previous formula for $G(\mathbf{x}, 0)$ we see that

$$\begin{aligned}
M(\mathbf{x}, 0) &= \frac{1}{\left(\prod_{j=1}^n x_j\right)^{n-2}} \cdot \left(\prod_{j=1}^n x_j \cdot \operatorname{diag}\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right)\right)^{\text{adj}} \\
&= \frac{\left(\prod_{j=1}^n x_j\right)^{n-1}}{\left(\prod_{j=1}^n x_j\right)^{n-2}} \cdot \left(\operatorname{diag}\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right)\right)^{\text{adj}} \\
&= \prod_{j=1}^n x_j \cdot \operatorname{diag}\left(\prod_{j\neq 1} \frac{1}{x_j}, \ldots, \prod_{j\neq n} \frac{1}{x_j}\right) \\
&= \operatorname{diag}\left(x_1, \ldots, x_n\right).
\end{aligned}$$

For general f, we take a factorization of f into irreducible polynomials $f = \prod_k f_k$. By Lemma 3.3.3, $\Delta_{ij}(f_k)$ is a square for each i, j, k and so by the arguments above, f_k has a determinantal representation of the correct form. Taking a block diagonal representation of these representations (and permuting the rows and columns if necessary to reorder $x_1, \ldots, x_n$) gives a determinantal representation for f. $\square$

Example 2.3.3. For $n = 4$ and $R = \mathbb{Z}$, we apply this algorithm to the symmetric quartic

$$f \;=\; x_1x_2x_3x_4 - (x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4) + 2(x_1 + x_2 + x_3 + x_4) - 3.$$

For each $i \in [4]$, we take g_{ii} to be $\frac{\partial f}{\partial x_i} = \prod_{j\neq i} x_j - \sum_{j\neq i} x_j + 2$. For every $i \neq j$, we find that $\Delta_{ij}(f) = \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} - f\frac{\partial^2 f}{\partial x_i \partial x_j}$ equals $(x_k - 1)^2(x_\ell - 1)^2$ where $\{k, \ell\} = [4]\backslash\{i, j\}$. For each $j = 2, 3, 4$, we can choose $g_{1j} = (1 - x_k)(x_\ell - 1)$ with $\{k, \ell\} = [4]\backslash\{1, j\}$. Then for $\{j, k, \ell\} = \{2, 3, 4\}$, we find that $g_{11}(1 - x_1)(x_j - 1) - g_{1k}g_{1\ell}$ equals $(1 - x_j)f$ and so we also take $g_{k\ell} = (1 - x_1)(x_j - 1)$. We then construct the 4×4 matrix $G = (g_{ij})_{1\leq i,j\leq 4} =$

$$\begin{pmatrix}
x_2x_3x_4 - x_2 - x_3 - x_4 + 2 & -x_3x_4 + x_3 + x_4 - 1 & -x_2x_4 + x_2 + x_4 - 1 & -x_2x_3 + x_2 + x_3 - 1 \\
-x_3x_4 + x_3 + x_4 - 1 & x_1x_3x_4 - x_1 - x_3 - x_4 + 2 & -x_1x_4 + x_1 + x_4 - 1 & -x_1x_3 + x_1 + x_3 - 1 \\
-x_2x_4 + x_2 + x_4 - 1 & -x_1x_4 + x_1 + x_4 - 1 & x_1x_2x_4 - x_1 - x_2 - x_4 + 2 & -x_1x_2 + x_1 + x_2 - 1 \\
-x_2x_3 + x_2 + x_3 - 1 & -x_1x_3 + x_1 + x_3 - 1 & -x_1x_2 + x_1 + x_2 - 1 & x_1x_2x_4 - x_1 - x_2 - x_4 + 2
\end{pmatrix}.$$

Note that only the diagonal entries of this matrix have degree three, so if we homogenize all entries to have degree three and then set the homogenizing variable equal to zero, the result is the diagonal matrix $G(\mathbf{x}, 0) = x_1x_2x_3x_4\text{diag}(x_1^{-1}, x_2^{-1}, x_3^{-1}, x_4^{-1})$ appearing in the proof of Theorem 3.3.1($\Leftarrow$). Moreover, the 2×2 minors of this matrix are divisible by f and so its 3×3 minors are divisible by f^2. Taking the adjugate of G and dividing by f^2, we find the desired symmetric matrix whose determinant gives the polynomial f:

$$M = \frac{1}{f^2}G^{\text{adj}} = \begin{pmatrix} x_1 & 1 & 1 & 1 \\ 1 & x_2 & 1 & 1 \\ 1 & 1 & x_3 & 1 \\ 1 & 1 & 1 & x_4 \end{pmatrix}.$$

Before moving on to applications to the principal minor problem, we remark that the rank of the matrices M_j can be recovered from f. Here we define the rank of a matrix $M \in \text{Sym}_n(R)$ to be the maximum $r \in \mathbb{N}$ so that there is a non-zero $r \times r$ minor of M. Note that the rank of M over R is the same as its rank over the field of fractions of R.

Lemma 2.3.4. *Let R be an integral domain. If $f = \det(\text{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_jM_j)$ where $M_j \in \text{Sym}_n(R)$, then the rank of M_j equals the degree f in the variable x_j.*

Proof. The bound $\deg_j(f) \leq \text{rank}(M_j)$ follows from the Laplace expansion of the determinant. To see equality, it suffices to take $j = m = n + 1$. Let $f = \det(\text{diag}(x_1, \ldots, x_n) + yA)$ where $A \in \text{Sym}_n(R)$. Then $f = \sum_{S\subseteq[n]} A_S\mathbf{x}^{[n]\setminus S}y^{|S|}$ equals the homogenization of f_A. From this we see that the degree of f in the variable y equals the size of the largest nonzero *principal* minor of A. By the so-called Principal Minor Theorem [45, Strong PMT 2.9], this coincides with the size of the largest nonzero minor of A, i.e. rank(A). $\square$

Recall that to an element $\mathbf{a} = (a_S)_{S\subseteq[n]}$ in R^{2^n} we associate the multiaffine polynomial

$$f_{\mathbf{a}} = \sum_{S\subseteq[n]} a_S\mathbf{x}^{[n]\setminus S}.$$

Theorem 2.3.5. *Let R be a unique factorization domain. An element $\mathbf{a} = (a_S)_{S\subseteq[n]}$ in R^{2^n} is in the image of $\text{Sym}_n(R)$ under the principal minor map if and only if $a_\emptyset = 1$ and for every $i, j \in [n]$, $\Delta_{ij}(f_{\mathbf{a}})$ is a square in $R[\mathbf{x}]$.*

Proof. By Corollary 2.2.3, $\Delta_{ij}(f_{\mathbf{a}})$ is a square in $R[\mathbf{x}]$ if and only if $\Delta_{ij}(\overline{f_{\mathbf{a}}})$ is a square in $R[\mathbf{x}, y]$. Furthermore, by Theorem 3.3.1, $\Delta_{ij}(\overline{f_{\mathbf{a}}})$ is a square in $R[\mathbf{x}, y]$ if and only if there exists a symmetric matrix $A \in \mathrm{Sym}_n(R)$ for which $\overline{f_{\mathbf{a}}} = \det(\mathrm{diag}(x_1, \ldots, x_n) + yA)$. $\square$

Corollary 2.3.6. *Let R be a unique factorization domain. Then the image of* $\mathrm{Sym}_n(R)$ *under the principal minor map is invariant under the action of $G^n \rtimes S_n$, where G is the subgroup of* $\mathrm{SL}_2(R)$ *defined by* $G = \left\{ \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \mid r \in R \right\}$.

Proof. Let $\mathbf{a} = (a_S)_{S \subseteq [n]}$ be an element in the image of the principal minor map and let $f_{\mathbf{a}}$ be the associated multiaffine polynomial. Consider $\mathbf{b} = (b_S)_{S \subseteq [n]}$ with $\mathbf{b} = \gamma \cdot \mathbf{a}$ where $\gamma \in G^n \rtimes S_n$. Theorem 3.3.1 implies that $a_\emptyset = 1$ and for every i, j, $\Delta_{ij}(f_{\mathbf{a}})$ is a square in $R[\mathbf{x}]$. Corollary 2.2.3 implies that for every i, j, $\Delta_{ij}(f_{\mathbf{b}})$ is a square in $R[\mathbf{x}]$. Using Theorem 3.3.1 again, it is enough to show that $b_\emptyset$, the coefficient of $x_1 \cdots x_n$ in $f_{\mathbf{b}}$, equals one. It is clear that this coefficient is invariant under the action of S_n, and so it suffices to take $\gamma = (\gamma_1, \ldots, \gamma_n) \in G^n$ where $\gamma_i = \left(\begin{smallmatrix} 1 & r_i \\ 0 & 1 \end{smallmatrix}\right)$, for which

$$f_{\mathbf{b}} = \gamma \cdot f_{\mathbf{a}} = f_{\mathbf{a}}(x_1 + r_1, \ldots, x_n + r_n).$$

From this, we see that coefficient of $x_1 \cdots x_n$ in $f_{\mathbf{b}}$ is equal to one. $\square$

The subgroup G is the maximal subgroup that preserves the leading coefficients of polynomials f_A. Consider an element of $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(R)$ acting on the first coordinate x_1 of f_A. Then

$$\mathrm{coeff}(\gamma \cdot f_A, x_1 \cdots x_n) = \mathrm{coeff}\left((ax_1 + b)\frac{\partial f_A}{\partial x_1} + (cx_1 + d)(f_A|_{x_1=0}), x_1 \cdots x_n \right) = aA_\emptyset + cA_1.$$

In order to preserve the coefficient of $x_1 \cdots x_n$, γ must satisfy $1 = a + cA_1$ for every value of A_1. This implies that $c = 0$ and $a = 1$. The condition $ad - bc = 1$ then implies that $d = 1$.

Corollary 2.3.7. *Let R be a unique factorization domain and let $\mathbf{a}$ be the vector of principal minors of a matrix in* $\mathrm{Sym}_n(R)$. *Fix $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$ and let λ denote the coefficient of*

$\prod_{i=1}^n x_i$ in $\gamma \cdot f_{\mathbf{a}}$. If $\lambda \neq 0$, then for some $A \in \mathrm{Sym}_n(R)$,

$$\gamma \cdot f_{\mathbf{a}} = \lambda \cdot \det\left(\mathrm{diag}(x_1, \ldots, x_n) + \frac{1}{\lambda}A\right).$$

That is, $\frac{1}{\lambda}\gamma \cdot \mathbf{a}$ belongs the image of the principal minor map over $\frac{1}{\lambda}R$.

Proof. Let $\mathbf{b} = \frac{1}{\lambda}\gamma \cdot \mathbf{a}$ and let $f_{\mathbf{b}}$ be the multiaffine polynomial associated to $\mathbf{b}$. Then $\mathrm{coeff}\,(f_{\mathbf{b}}, \prod_i x_i) = \frac{1}{\lambda}\lambda = 1$ and by Corollary 2.2.3, $\Delta_{ij} f_{\mathbf{b}} = \frac{1}{\lambda^2}\Delta_{ij}(\gamma \cdot f_{\mathbf{a}})$ is a square for every i, j. Hence, by Theorem 3.3.1, there exists a symmetric matrix B with entries in $R(\frac{1}{\lambda})$ with

$$f_{\mathbf{b}} = \det\left(\mathrm{diag}(x_1, \ldots, x_n) + B\right).$$

We claim that λB has entries in R. To see this, note that $\mathbf{b} = \frac{1}{\lambda}\gamma \cdot \mathbf{a}$ is the vector of principal minors of B and that the entries of $\gamma \cdot \mathbf{a}$ belong to R. So all principal minors of B belong to $\lambda^{-1}R$. This immediately shows that the diagonal elements of λB belong to R.

For the off-diagonal elements, fix $i \neq j \in [n]$ and let z denote the (i, j)th entry of B. Then $b_i b_j - b_{ij} = z^2$, where $\mathbf{b} = (b_S)_{S \subseteq [n]}$, which implies that $\lambda^2 z^2 = (\lambda b_i)(\lambda b_j) - \lambda^2 b_{ij} \in R$. By construction, $z \in R(1/\lambda)$ so we can take the minimal $m \in \mathbb{N}$ for which $\lambda^m z \in R$. So $\lambda^m z = r \in R$ where either $m = 0$ or λ does not divide r. Then $\lambda^{2m-2}(\lambda^2 z^2) = (\lambda^m z)^2 = r^2$. Since $\lambda^2 z^2 \in R$, we see that λ^{2m-2} divides r^2. If $m > 1$, this contradicts the assumption that λ does not divide r. Therefore $A = \lambda B \in \mathrm{Sym}_n(R)$, as desired. □

Example 2.3.8. For $R = \mathbb{Z}$ and $n = 2$ consider the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and its vector of principal minors $\mathbf{a} = (1, 2, 1, 1)$, giving $f_{\mathbf{a}} = x_1x_2 + x_1 + 2x_2 + 1$. The image of $f_{\mathbf{a}}$ under the action of $\gamma = \left(\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \in \mathrm{SL}_2(R)^2$ is

$$\gamma \cdot f_{\mathbf{a}} = 4x_1x_2 + 3x_1 - x_2 - 1.$$

Since $4 \neq 1$, the vector $(4, -1, 3, -1)$ is not the vector of principal minors of any symmetric matrix over $\mathbb{Z}$. However, as promised by Corollary 2.3.7, the vector $\frac{1}{4}\gamma \cdot \mathbf{a} = \left(1, \frac{-1}{4}, \frac{3}{4}, \frac{-1}{4}\right)$ is the vector of principal minors of the matrix $B = \frac{1}{4}\begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix}$ over $\frac{1}{4}\mathbb{Z}$.

Corollary 2.3.9. *Let $R = \mathbb{F}$ be an infinite field. The Zariski closure of the image of the principal minor map in $\mathbb{P}^{2^n-1}(\mathbb{F})$ is invariant under the action of $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$.*

Proof. The image is immediately invariant under the action of S_n, so it suffices to show invariance under $\mathrm{SL}_2(\mathbb{F})^n$. Using the action of S_n, it suffices to show this by acting with $\mathrm{SL}_2(\mathbb{F})$ on the first coordinate. Let $A \in \mathrm{Sym}_n(\mathbb{F})$, giving a point $\varphi(A) = (A_S)_S$ in the image of the principal minor map. Consider the open subset

$$\begin{aligned} U &= \{\gamma \in \mathrm{SL}_2(\mathbb{F}) : \mathrm{coeff}(\gamma \cdot f_A, x_1 \cdots x_n) \neq 0\} \\ &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{F}^{2\times 2} : ad - bc = 1 \ \text{ and } \ aA_\emptyset + cA_1 \neq 0 \right\}. \end{aligned}$$

By Corollary 2.3.7, for every $\gamma \in U$, $\gamma \cdot \varphi(A)$ belongs to the image of the principal minor map, up to scaling. The parametrization $(a, b, c) \dashrightarrow (a, b, c, (1 + bc)/a)$ shows that $\mathrm{SL}_2(\mathbb{F})$ is a rational variety over $\mathbb{F}$. Since $\mathbb{F}$ is infinite, the set of $(a, b, c) \in \mathbb{F}^3$ such that $a \neq 0$ and $aA_\emptyset + cA_1 \neq 0$ is Zariski-dense in $\mathbb{F}^3$. It follows that U is Zariski-dense in $\mathrm{SL}_2(\mathbb{F})$. Since $\gamma \mapsto \gamma \cdot \varphi(A)$ defines a rational map $\mathrm{SL}_2(\mathbb{F}) \to \mathbb{P}^{2^n-1}(\mathbb{F})$, it follows that for any $\gamma \in \mathrm{SL}_2(\mathbb{F})$, $\gamma \cdot \varphi(A)$ belongs to the Zariski-closure of the image of φ in $\mathbb{P}^{2^n-1}(\mathbb{F})$. □

2.4 Defining the set of multiquadratic squares

The polynomials $\Delta_{ij}(f)$ appearing in Theorem 2.3.5 have degree $\leq$ two in each variable. In order to make use of this characterization, in this section we find algebraic conditions characterizing the set of squares in $R[\mathbf{x}]_{\leq \mathbf{2}}$. In fact, to simplify the notation used in the arguments below we consider the multihomogenezations, as in Proposition 3.2.1, and characterize the set of squares in $R[\mathbf{x}, \mathbf{y}]_{\mathbf{2}}$.

In a slight abuse of notation, we define $\mathbb{P}^1(R)$ to be the following subset of R^2:

$$\mathbb{P}^1(R) = \{(r, 1) : r \in R\} \cup \{(1, 0)\}.$$

Lemma 2.4.1. *Let $g \in R[\mathbf{x}, \mathbf{y}]_{\mathbf{d}}$ where $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{N}^n$. Let $P_i \subseteq \mathbb{P}^1(R)$ be a set of size $d_i + 1$. Then g is the zero polynomial if and only if $g(p) = 0$ for all $p \in P_1 \times \cdots \times P_n$.*

Proof. We prove this by induction on n. Note that for $n = 1$, a polynomial $g \in R[x, y]_d$ vanishes at $(a, b) \in \mathbb{P}^1(R)$ if and only if $bx - ay$ divides g. Therefore a bivariate form

$g \in R[x,y]_d$ cannot have more than d roots in $\mathbb{P}^1(R)$. Now suppose that $n \geq 1$. Fix $(a,b) \in P_{n+1}$. The polynomial $g(\mathbf{x}, a, \mathbf{y}, b)$ vanishes on $P_1 \times \cdots \times P_n$ and is therefore identically zero by induction. This means that, considered as a bivariate form in x_{n+1}, y_{n+1} over the ring $R[\mathbf{x}, \mathbf{y}]$, g vanishes at the $d_{n+1}+1$ points in P_{n+1} and it is therefore identically zero in $R[\mathbf{x}, \mathbf{y}][x_{n+1}, y_{n+1}] = R[\mathbf{x}, x_{n+1}, \mathbf{y}, y_{n+1}]$. $\square$

Lemma 2.4.2. *A polynomial* $g(\mathbf{x}, s, \mathbf{y}, t) = g_2(\mathbf{x}, \mathbf{y})s^2 + g_1(\mathbf{x}, \mathbf{y})st + g_0(\mathbf{x}, \mathbf{y})t^2 \in R[\mathbf{x}, s, \mathbf{y}, t]_{\mathbf{2}}$ *is a square in* $R[\mathbf{x}, s, \mathbf{y}, t]$ *if and only if* $g_0(\mathbf{x}, \mathbf{y})$ *and* $g_2(\mathbf{x}, \mathbf{y})$ *are squares in* $R[\mathbf{x}, \mathbf{y}]$ *and* $\mathrm{Discr}_{(s,t)}(g) = 0$*.*

Proof. Suppose that $g_2 = (h_2)^2$ and $g_0 = (h_0)^2$, where $h_2, h_0 \in R[\mathbf{x}, \mathbf{y}]$ and

$$\mathrm{Discr}_{(s,t)}(g) = g_1^2 - 4g_0g_2 = 0 \quad \text{in} \quad R[\mathbf{x}, \mathbf{y}].$$

Then $4g_0g_2 = (2h_0h_2)^2 = g_1^2$, giving that $(2h_0h_2)^2 - g_1^2 = (2h_0h_2 - g_1)(2h_0h_2 + g_1) = 0$ in $R[\mathbf{x}, \mathbf{y}]$. It follows that $g_1 = \pm 2h_0h_2$. Changing the sign of h_0 if necessary, we can assume that $g_1 = 2h_0h_2$. Then

$$g = (h_2)^2 s^2 + 2h_0h_2 st + (h_0)^2 t^2 = (sh_2 + th_0)^2$$

is a square, as desired. Conversely, if $g = (sh_2 + th_0)^2$, we see that $g_0 = (h_0)^2$ and $g_2 = (h_2)^2$ and $\mathrm{Discr}_{(s,t)}(g) = 0$. $\square$

Theorem 2.4.3. *Let R be a unique factorization domain with $|R| \neq 3$. A multiquadratic polynomial* $g = \sum_{\alpha \in \{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha \mathbf{y}^{\mathbf{2}-\alpha}$ *is a square, i.e.* $g = h^2$ *with* $h \in R[\mathbf{x}, \mathbf{y}]_{\mathbf{1}}$*, if and only if for every* $\beta \in \{0,1\}^n$*,* $c_{2\beta}$ *is a square in* R *and* $\mathbf{c} = (c_\alpha)_{\alpha \in \{0,1,2\}^n}$ *satisfies the images of*

$$(c_{(1,\mathbf{0})})^2 - 4c_{(0,\mathbf{0})}c_{(2,\mathbf{0})} = 0 \tag{2.3}$$

under the action of $\mathrm{SL}_2(R')^n \rtimes S_n$ *where* R' *is any nontrivial subring of* R *with* $1_R = 1_{R'}$ *and size* ≥ 4 *whenever* $\mathrm{char}(R) \neq 2$*.*

Proof. $(\Rightarrow)$ If $g = h^2$, then for every $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$, $\gamma \cdot g = (\gamma \cdot h)^2$. Note here that the action on g comes from the action on $R[\mathbf{x}, \mathbf{y}]_{\mathbf{2}}$ and the action on h comes from the action on $R[\mathbf{x},\mathbf{y}]_{\mathbf{1}}$. The specialization of $\gamma \cdot g$ to $(\mathbf{x},\mathbf{y}) = (x_1, \mathbf{0}, y_1, \mathbf{1})$ will be a square in $R[x_1, y_1]$. Moreover, the coefficient c_β of $\mathbf{x}^{2\beta}\mathbf{y}^{\mathbf{2}-2\beta}$ in g is the square of the coefficient $\mathbf{x}^{\beta}\mathbf{y}^{\mathbf{1}-\beta}$ in h, and in particular the square of an element in the ring R.

($\Leftarrow$) We prove this by induction on n. This holds immediately for $n = 1$ by Lemma 2.4.2. Now suppose $n \geq 1$ and take $g = g_2 s^2 + g_1 st + g_0 t^2 \in R[\mathbf{x}, s, \mathbf{y}, t]$. Fixing $n+1$ in S_{n+1} and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ for the $(n+1)$st coordinate in $(\mathrm{SL}_2(R))^{n+1}$, we see that both g_0 and g_2 satisfy the hypothesis of the theorem and so, by induction, are squares in $R[\mathbf{x},\mathbf{y}]$. Here 1 denotes the common multiplicative identity of R and R'.

If $\mathrm{char}(R) \neq 2$ then $|R'| \geq 4$ and we can take a set $P \subset \mathbb{P}^1(R')$ of size five. Define a map $\varphi : \mathbb{P}^1(R') \to \mathrm{SL}_2(R')$ by $\varphi((1,0)) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and for $r \in R'$, $\varphi((r,1)) = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$. Then to $(p_1, \ldots, p_n) \in P^n$, we associate the element $\gamma = (\varphi(p_1), \ldots, \varphi(p_n), \mathrm{Id}_2)$ of $\mathrm{SL}_2(R')^{n+1}$. Acting on g by γ and then specializing to $x_1 = \ldots = x_n = 0$ and $y_1 = \ldots = y_n = 1$ gives

$$\big(\gamma \cdot g\big)|_{\mathbf{x}=\mathbf{0},\mathbf{y}=\mathbf{1}} = g(\mathbf{a}, s, \mathbf{b}, t) \quad \text{and} \quad \big(\mathrm{Discr}_{(s,t)}(\gamma \cdot g)\big)|_{\mathbf{x}=\mathbf{0},\mathbf{y}=\mathbf{1}} = \big(\mathrm{Discr}_{(s,t)} g\big)|_{\mathbf{x}=\mathbf{a},\mathbf{y}=\mathbf{b}} \qquad (2.4)$$

where $p_i = (a_i, b_i)$ and $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$. Transposing (x_1, y_1) and (s,t) using the action of S_{n+1}, we see that by assumption, this evaluation of the discriminant of $\gamma \cdot g$ must be zero. Since the discriminant has degree ≤ 4 in each variable, Lemma 2.4.1 implies that it is identically zero. Then by Lemma 2.4.2, g is a square in $R[\mathbf{x}, s, \mathbf{y}, t]$.

If $\mathrm{char}(R) = 2$, the discriminant $\mathrm{Discr}_{(s,t)} g$ simplifies to a square, g_1^2, which by (2.4), must vanish at the points $\{(1,0),(1,1),(0,1)\}^n \subseteq (\mathbb{P}^1(R))^n$. Since g_1 has degree 2 in each variable and must vanish at these points, Lemma 2.4.1 implies that g_1 is identically zero. Therefore by Lemma 2.4.2, g is a square. □

Remark 2.4.4. The proof of Theorem 2.4.3 reveals that only a small subset of $\mathrm{SL}_2(R)$ is needed in each coordinate to characterize multiquadratic squares, specifically a set of size five.

Up to isomorphism, there is only one ring of size three, namely $\mathbb{F}_3$. The exclusion of $\mathbb{F}_3$ in the statement of Theorem 2.4.3 is a necessary one, as the following example demonstrates.

Example 2.4.5. For $R = \mathbb{F}_3$ and $n = 3$ consider the multiquadratic form

$$\begin{aligned} g &= x_1^2(x_2y_3 - x_3y_2)^2 - 2x_1y_1(x_2y_3 + x_3y_2)(x_2x_3 - y_2y_3) + y_1^2(x_2x_3 + y_2y_3)^2 \\ &= x_1^2x_2^2y_3^2 + x_1^2x_3^2y_2^2 + x_2^2x_3^2y_1^2 - 2x_1^2x_2x_3y_2y_3 - 2x_1x_2^2x_3y_1y_3 - 2x_1x_2x_3^2y_1y_2 \\ &\quad + 2x_1x_2y_1y_2y_3^2 + 2x_1x_3y_1y_2^2y_3 + 2x_2x_3y_1^2y_2y_3 + y_1^2y_2^2y_3^2. \end{aligned}$$

The discriminant of g with respect to (x_1, y_1) is given by

$$\mathrm{Discr}_{(x_1,y_1)}g = 16x_2y_2(x_2 + y_2)(x_2 - y_2)x_3y_3(x_3 + y_3)(x_3 - y_3).$$

This polynomial is non-zero in $\mathbb{F}_3[\mathbf{x}, \mathbf{y}]_{\mathbf{4}}$ but vanishes on all points in $(\mathbb{P}^1(\mathbb{F}_3))^2$. The polynomial g is invariant under permutations of indices, so the discriminants $\mathrm{Discr}_{(x_2,y_2)}g$ and $\mathrm{Discr}_{(x_3,y_3)}g$ have the same property.

Before using this result to characterize the image of the principal minor map, we record a few of the notable special cases of Theorem 2.4.3. In particular, over $\mathbb{C}$, the set of multiquadratic squares is defined by the orbit of a single polynomial and the only additional constraint over $\mathbb{R}$ is the nonnegativity of the orbit of one other polynomial.

Corollary 2.4.6. *Let $g = \sum_{\alpha\in\{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha \in \mathbb{C}[\mathbf{x}]$. The polynomial g is a square, $g = h^2$ with $h \in \mathbb{C}[\mathbf{x}]_{\leq\mathbf{1}}$ if and only if $\mathbf{c} = (c_\alpha)_{\alpha\in\{0,1,2\}^n}$ satisfies the images of*

$$(c_{(1,\mathbf{0})})^2 - 4c_{(0,\mathbf{0})}c_{(2,\mathbf{0})} = 0 \tag{2.5}$$

under the action of $\mathrm{SL}_2(\mathbb{C})^n \rtimes S_n$. Moreover, g is a square over $\mathbb{R}$, $g = h^2$ with $h \in \mathbb{R}[\mathbf{x}]_{\leq\mathbf{1}}$, if and only if its coefficients $\mathbf{c} = (c_\alpha)_{\alpha\in\{0,1,2\}^n}$ satisfy the images of

$$(c_{(1,\mathbf{0})})^2 - 4c_{(0,\mathbf{0})}c_{(2,\mathbf{0})} = 0 \quad \text{and} \quad c_{(0,\mathbf{0})} \geq 0. \tag{2.6}$$

under the action of $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$.

As seen in the proof of Theorem 2.4.3, in characteristic two, linear equations in the coefficients suffice to cut out of the set of squares in $R[\mathbf{x}]_{\leq\mathbf{2}}$.

Corollary 2.4.7. *Let* $g = \sum_{\alpha \in \{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha \in R[\mathbf{x}]$ *where* R *is a UFD of characteristic two. The polynomial* g *is a square,* $g = h^2$ *with* $h \in R[\mathbf{x}]_{\leq \mathbf{1}}$ *if and only if* $\mathbf{c} = (c_\alpha)_{\alpha \in \{0,1,2\}^n}$ *satisfy*

(i) for every $\beta \in \{0,1\}^n$*,* $c_{2\beta}$ *is a square in* R*, and*

(ii) for every $\gamma \in \mathrm{SL}_2(\mathbb{F}_2)^n \rtimes S_n$*,* $(\gamma \cdot \mathbf{c})_{(1,\mathbf{0})} = 0$*.*

2.5 Characterizing the image of the principal minor map

We can now combine the characterization of multiaffine determinantal polynomials from Section 2.3 and the characterization of multiquadratic squares in Section 2.4 to give a complete description of the principal minor map over any unique factorization domain of size $\neq 3$.

Theorem 2.5.1. *Let* R *be an UFD with* $|R| \neq 3$ *and let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in R^{2^n}$ *with* $a_\emptyset = 1$*. There exists a symmetric matrix over* R *with principal minors* $\mathbf{a}$ *if and only if*

(i) for every $i, j \in [n]$*,* $a_i a_j - a_{ij}$ *is a square in* R*, and*

(ii) for every $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$*,* $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$*.*

Proof. ($\Rightarrow$) If $\mathbf{a}$ belongs to the image of the principal minor map over R, then $a_i a_j - a_{ij}$ is the square of the (i,j)th entry of the representing matrix A, and so is a square in R. Also, by Theorem 3.3.1 $\Delta_{ij} f_{\mathbf{a}}$ is a square for all $i, j \in [n]$. Then, by Corollary 2.2.3, for any $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$, $\Delta_{ij}(\gamma \cdot f_{\mathbf{a}})$ is square. In particular,

$$\mathrm{HypDet}(\gamma \cdot \mathbf{a}) = \mathrm{Discr}_{x_3}(\Delta_{12}(\gamma \cdot f_{\mathbf{a}}))|_{x_4 = \ldots = x_n = 0} = 0.$$

($\Leftarrow$) First, we show that $\mathbf{a}$ belongs to the image of the principal minor map over the algebraic closure of the fraction field of R, $\mathbb{F} = \overline{(\mathrm{frac}(R))}^{alg}$. Then R is a subring of $\mathbb{F}$ and has size ≥ 4 if $\mathrm{char}(R) = \mathrm{char}(\mathbb{F}) \neq 2$. Let $f = f_{\mathbf{a}}$. Every element of R is a square over $\mathbb{F}$. Then by Theorem 2.4.3, $\Delta_{ij} f = \sum_{\alpha \in \{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha$ is a square in $\mathbb{F}[\mathbf{x}]$ if and only if $\mathbf{c} = (c_\alpha)_{\alpha \in \{0,1,2\}^n}$ and its images under $\mathrm{SL}_2(R)^n \rtimes S_n$ satisfy $(c_{(1,\mathbf{0})})^2 - 4c_{(0,\mathbf{0})}c_{(2,\mathbf{0})} = 0$. This condition for all $i, j \in [n]$ is equivalent to the condition that $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$ for all

$\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$. Therefore $\Delta_{ij} f$ is a square in $\mathbb{F}[\mathbf{x}]$ for all i, j. By Theorem 2.3.5, $\mathbf{a}$ is the vector of principal minors of some matrix A with entries in $\mathbb{F}$. The diagonal entries of A are entries a_i in this vector and therefore belong to R. Let z denote the (i, j)th entry of A for some $i \neq j$. By assumption $z^2 = a_i a_j - a_{ij} = r^2$ for some $r \in R$. Then $(z + r)(z - r) = 0$ implying that $z = \pm r \in R$. Therefore $A \in \mathrm{Sym}_n(R)$. □

Remark 2.5.2. Note that in part (ii) of the characterization in Theorem 2.5.1, it suffices to take $\binom{n}{3} \cdot 5^{n-3}$ elements $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$. Specifically, let $P \subseteq R$ be a set of size 5. For any subset $\{i, j, k\} \subseteq [n]$ and point $\mathbf{p} \in P^{n-3}$, we get an equation

$$\mathrm{Discr}_{x_k}(\Delta_{ij}(f_{\mathbf{a}}))|_{\mathbf{x}=\mathbf{p}} = 0.$$

This is enough to ensure that $\mathrm{Discr}_{x_k}(\Delta_{ij}(f_{\mathbf{a}}))$ is identically zero. More generally, we can take evaluations of $\mathrm{Discr}_{(x_k,y_k)}(\Delta_{ij}(f_{\mathbf{a}}))^{\mathbf{2}-\mathrm{hom}}$ at 5^{n-3} points in $\mathbb{P}^1(R)$, as in Section 2.4. While this expression appears to depend on the ordering of i, j, k, one can check that for any $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$, $\mathrm{Discr}_{x_k}(\Delta_{ij}(f)) = \mathrm{Discr}_{x_j}(\Delta_{ik}(f)) = \mathrm{Discr}_{x_i}(\Delta_{jk}(f))$. See, for example, the proof of Theorem 3.1 in [72]. When R is a ring of characteristic 2, it suffices to take P to have size 3, giving a total of $\binom{n}{3} \cdot 3^{n-3}$ equations.

Applying this to $R = \mathbb{C}$, $\mathbb{R}$, and $\mathbb{F}_2$, we find the following immediate consequences.

Corollary 2.5.3. *Let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{C}^{2^n}$ *with* $a_\emptyset = 1$. *There exists a symmetric matrix over* $\mathbb{C}$ *with principal minors* $\mathbf{a}$ *if and only if* $\mathbf{a}$ *and all its images under the action of* $\mathrm{SL}_2(\mathbb{C})^n \rtimes S_n$ *satisfy the* $2 \times 2 \times 2$ *hyperdeterminant* $\mathrm{HypDet}(\mathbf{a}) = 0$.

Corollary 2.5.4. *Let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{R}^{2^n}$ *with* $a_\emptyset = 1$. *There exists a symmetric matrix over* $\mathbb{R}$ *with principal minors* $\mathbf{a}$ *if and only if* $\mathbf{a}$ *and all its images under the action of* $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$ *satisfy*

$$\mathrm{HypDet}(\mathbf{a}) = 0 \textit{ and } a_1 a_2 - a_\emptyset a_{12} \geq 0.$$

Corollary 2.5.5. *Let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in R^{2^n}$ *with* $a_\emptyset = 1$ *where* R *has characteristic two. There exists a symmetric matrix over* R *with principal minors* $\mathbf{a}$ *if and only if* *(i) for every* $i, j \in [n]$, $a_i a_j - a_{ij}$ *is a square in* R, *and*

(ii) for every $\gamma \in \mathrm{SL}_2(\mathbb{F}_2)^n \rtimes S_n$, $\gamma \cdot (a_\emptyset a_{123} + a_1 a_{23} + a_2 a_{13} + a_3 a_{12}) = 0$.

In particular, for $R = \mathbb{F}_2$, (i) *is always satisfied and the image of the principal minor map is cut out by the quadratic equations in* (ii).

It is unclear whether or not Theorem 2.5.1 can be extended to $R = \mathbb{F}_3$. Example 2.4.5 shows that this would likely require a different proof technique. Interestingly, the polynomial g in this example is of the form $\Delta_{12}(f)$ for some $f \in \mathbb{F}_3[x_1, \ldots, x_5]_{\leq \mathbf{1}}$, but for all such f we have found, the discriminant of some other $\Delta_{ij}(f)$ fails to vanish on $(\mathbb{P}^1(\mathbb{F}_3))^2$.

Question 2.5.6. *Does the equivalence in Theorem 2.5.1 hold for* $R = \mathbb{F}_3$*?*

2.6 Other determinantal representations and connections to $\mathrm{Gr}_{\mathbb{F}}(d, n)$

2.6.1 Other multiaffine determinantal representations

In this section we restrict ourselves to fields and consider the set of multiaffine determinantal polynomials. Formally, let $\mathbb{F}$ be an arbitrary field. We call a polynomial $f \in \mathbb{F}[\mathbf{x}]_{\leq \mathbf{1}}$ *determinantal* if it can be written in the form

$$f(\mathbf{x}) = \lambda \det\left(V \mathrm{diag}(x_1, \ldots, x_n) V^T + W\right) = \lambda \det\left(\sum_{i=1}^{n} x_i v_i v_i^T + W\right) \tag{2.7}$$

for some $\lambda \in \mathbb{F}$, some matrix $V = (v_1, \ldots, v_n) \in \mathbb{F}^{m \times n}$ and some $W \in \mathrm{Sym}_m(\mathbb{F})$ for some m. Note that when we take V to be the $n \times n$ identity matrix, this is exactly the principal minor polynomial f_W. When $m < n$, the coefficient of $x_1 \cdots x_n$ in f is necessarily zero.

Theorem 2.6.1. *A polynomial* $f \in \mathbb{F}[\mathbf{x}]_{\leq \mathbf{1}}$ *has a determinantal representation* (3.5) *if and only if for all* $i, j \in [n]$, $\Delta_{ij} f$ *is a square in* $\mathbb{F}[\mathbf{x}]$. *Moreover, one can always take a representation of size* $m = \deg(f)$ *in* (3.5).

Proof. ($\Rightarrow$) Without loss of generality, we show that $\Delta_{12}(f)$ is a square. First suppose v_1 and v_2 are linearly dependent, i.e. let $v_1 = \alpha v_2$ for some $\alpha \in \mathbb{F}$. Then $v_1 v_1^T = \alpha^2 v_2 v_2^T$ and

$f(x_1,\dots,x_n) = f(0,\alpha^2x_1+x_2,x_3,\dots,x_n)$. Taking partial derivatives shows that $\frac{\partial f}{\partial x_1} = \alpha^2\frac{\partial f}{\partial x_2}$ and that $\frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$. Therefore $\Delta_{12}(f) = \alpha^2(\frac{\partial f}{\partial x_2})^2$ and so is a square.

If v_1 and v_2 are linearly independent, then there is an invertible matrix U with $Uv_1 = e_1$ and $Uv_2 = e_2$. Then

$$\det(U)^2 f = \lambda \det\left(U\left(\sum_{i=1}^n x_iv_iv_i^T + W\right)U^T\right) = \lambda\det\left(\operatorname{diag}(x_1,x_2,\mathbf{0}) + \sum_{i=3}^n x_i\widetilde{v_i}\widetilde{v_i}^T + \widetilde{W}\right).$$

where $\widetilde{v_i} = Uv_i$ and $\widetilde{W} = UWU^T$. These matrices are still symmetric and so by equation (3.2), $\Delta_{12}(f)$ is a square.

($\Leftarrow$) Let $d = \deg(f)$. We can assume, without loss of generality, that the coefficient of $x_1\cdots x_d$ in f is nonzero. Moreover since the set of polynomials of the form (3.5) is invariant under scaling, we can assume that this coefficient equals one. Let $\overline{f} \in \mathbb{F}[\mathbf{x}, y]$ denote the homogenezation of f to total degree d. By Theorem 3.3.1, there are matrices $M_{d+1},\dots,M_{n+1}$ in $\mathrm{Sym}_d(\mathbb{F})$ so that

$$\overline{f} = \det\left(\operatorname{diag}(x_1,\dots,x_d) + \sum_{i=d+1}^n x_iM_i + yM_{n+1}\right).$$

We take $W = M_{n+1}$. It remains to show that for each $i = d+1,\dots,n$, the matrix M_i has the form $v_iv_i^T$ for some $v_i \in \mathbb{F}^d$. Without loss of generality we do this for $i = d+1$. Let g denote the specialization of $\overline{f}$ to $x_k = 0$ for $k = d+2,\dots,n$ and $y = 0$. Note that

$$g = \sum_{S\subseteq[d]} (M_{d+1})_S (x_{d+1})^{|S|} \prod_{j\in[d]\setminus S} x_j.$$

However f has degree ≤ 1 in x_{d+1}, and thus so does g. Therefore by Lemma 3.5.7, M_{d+1} has rank $\leq$ one. To examine its diagonal entries $(M_{d+1})_i$, note that for every $i = 1,\dots,d$, $\Delta_{i(d+1)}g$ is a square. Moreover, the restriction of g to $x_i = x_{d+1} = 0$ is identically zero, showing that

$$\begin{aligned}\Delta_{i(d+1)}g &= g_i^{d+1}g_{d+1}^i - g^{i(d+1)}g_{i(d+1)} = g_i^{d+1}g_{d+1}^i \\ &= \Big((M_{d+1})_\emptyset \prod_{j\in[d]\setminus i} x_j\Big)\Big((M_{d+1})_i \prod_{j\in[d]\setminus i} x_j\Big) = (M_{d+1})_i\left(\frac{x_1\cdots x_d}{x_i}\right)^2,\end{aligned}$$

where we use the notation $g_j = \frac{\partial g}{\partial x_j}$ and $g^j = g|_{x_j=0}$. It follows that $(M_{d+1})_i$ is a square in $\mathbb{F}$. Since M_{d+1} has rank $\leq$ one, then for some choice of square root $v_i = \sqrt{(M_{d+1})_i} \in \mathbb{F}$, the matrix M_{d+1} equals vv^T with $v = (v_1, \ldots, v_d)^T$. □

Using Corollary 2.2.3, this immediately gives the invariance of the set of determinantal polynomials.

Corollary 2.6.2. *The set of polynomials in* $\mathbb{F}[\mathbf{x}]_{\leq \mathbf{1}}$ *with a determinantal representation* (3.5) *is invariant under the action of* $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$.

Together, Theorems 2.4.3 and 3.5.6 characterize the set of determinantal polynomials in $\mathbb{F}[\mathbf{x}]$.

Corollary 2.6.3. *A polynomial* $f = \sum_{S\subseteq[n]} a_S \mathbf{x}^{[n]\backslash S} \in \mathbb{F}[\mathbf{x}]_{\leq \mathbf{1}}$ *has a determinantal representation* (3.5) *if and only if*

(i) for every $i, j \in [n]$ *and* $S \subseteq [n]\backslash\{i,j\}$, $a_{S\cup i}a_{S\cup j} - a_S a_{S\cup ij}$ *is a square over* $\mathbb{F}$, *and*

(ii) for every $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$.

Proof. By Theorem 3.5.6, f has a determinantal representation (3.5) if and only if for all i, j $\Delta_{ij} f$ is a square in $\mathbb{F}[\mathbf{x}]$. Since $\Delta_{ij} f$ has degree ≤ 2 in each variable, Theorem 2.4.3 implies that $\Delta_{ij} f = \sum_{\alpha\in\{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha$ is a square if and only if for every $\beta \in \{0,1\}^n$, $c_{2\beta}$ is a square in R and $\mathbf{c} = (c_\alpha)_{\alpha\in\{0,1,2\}^n}$ satisfy the images of

$$(c_{(1,\mathbf{0})})^2 - 4c_{(0,\mathbf{0})}c_{(2,\mathbf{0})} = 0 \tag{2.8}$$

under the action $\mathrm{SL}_2(\mathbb{F}) \rtimes S_n$. This is in turn equivalent to the condition that for every γ in $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$ and for every $i, j \in [n]$ and $S \subseteq [n]\backslash\{i,j\}$, $a_{S\cup i}a_{S\cup j} - a_S a_{S\cup ij}$ is a square in $\mathbb{F}$.

To see this, consider $S \subseteq [n]\backslash\{i,j\}$ and let $\beta \in \{0,1\}^n$ denote the indicator vector of $[n]\backslash(S \cup ij)$. We claim that that the coefficient of $\mathbf{x}^{2\beta}$ in $\Delta_{ij}(f)$ equals $a_{S\cup i}a_{S\cup j} - a_S a_{S\cup ij}$.

Since f, $\frac{\partial f}{\partial x_i}$, $\frac{\partial f}{\partial x_j}$, and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ have degree ≤ 1 in each variable, only the $\mathbf{x}^\beta$ terms in each of these polynomials contribute to the $\mathbf{x}^{2\beta}$ term in $\Delta_{ij}(f) = \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} - f\frac{\partial^2 f}{\partial x_i \partial x_j}$. That is,

$$\operatorname{coeff}\left(\Delta_{ij}(f), \mathbf{x}^{2\beta}\right) = \operatorname{coeff}\left(\frac{\partial f}{\partial x_i}, \mathbf{x}^\beta\right) \cdot \operatorname{coeff}\left(\frac{\partial f}{\partial x_j}, \mathbf{x}^\beta\right) - \operatorname{coeff}\left(f, \mathbf{x}^\beta\right) \cdot \operatorname{coeff}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}, \mathbf{x}^\beta\right).$$

Note that the coefficient of $\mathbf{x}^\beta$ in f is $a_{S\cup ij}$. The coefficient of $\mathbf{x}^\beta$ in $\frac{\partial f}{\partial x_i}$ equals the coefficient of $x_i \cdot \mathbf{x}^\beta$ in f, which is $a_{S\cup j}$. Similarly, the coefficients of $\mathbf{x}^\beta$ in $\frac{\partial f}{\partial x_j}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ are $a_{S\cup i}$ and a_S, respectively. □

2.6.2 Connections with the Grassmannian

Given a $d \times n$ matrix V of full-rank d, consider the polynomial f from (3.5) with $W = 0$:

$$f(\mathbf{x}) = \lambda \det\left(V \operatorname{diag}(x_1, \ldots, x_n) V^T\right) = \lambda \det\left(\sum_{i=1}^{n} x_i v_i v_i^T\right) = \lambda \sum_{S \in \binom{[n]}{d}} \left(V_S\right)^2 \mathbf{x}^S. \tag{2.9}$$

Here $\binom{[n]}{d}$ denotes the collection of size-d subsets of $[n]$ and V_S denotes the $d \times d$ minor of V obtained by taking columns indexed by S. If V has full rank d, the coefficients of f are the squares of the Plücker coordinates given by the Plücker embedding of the rowspan of V into $\operatorname{Gr}_{\mathbb{F}}(d, n)$. Otherwise f is identically zero.

Formally, consider the Plücker embedding of $\operatorname{Gr}_{\mathbb{F}}(d, n)$ into $\mathbb{P}^{\binom{n}{d}-1}(\mathbb{F})$. Given a subspace $L \subseteq \mathbb{F}^n$ of dimension d, its image in $\mathbb{P}^{\binom{n}{d}-1}(\mathbb{F})$ is the length-$\binom{n}{d}$ vector of $d \times d$ minors of any $d \times n$ matrix V whose rowspan equals L. The map $[p_S]_S \mapsto [(p_S)^2]_S$ defines a morphism $\mathbb{P}^{\binom{n}{d}-1}(\mathbb{F}) \to \mathbb{P}^{\binom{n}{d}-1}(\mathbb{F})$. Let $\operatorname{Gr}^2_{\mathbb{F}}(d, n)$ denote the image of $\operatorname{Gr}_{\mathbb{F}}(d, n)$ under this morphism. Corollary 2.6.3 then gives an immediate characterization of $\operatorname{Gr}^2_{\mathbb{F}}(d, n)$ via the hyperdeterminantal equations in $\mathbb{F}^{2^n}$. In fact, setting $x_n = 1$ in (2.9), we can study this image via multiaffine determinantal representations in the variables $x_1, \ldots, x_{n-1}$ and use the hyperdeterminantal equations in $\mathbb{F}^{2^{n-1}}$.

Corollary 2.6.4. *Let* $\mathbf{q} = (q_S)_{S \in \binom{[n]}{d}} \in \mathbb{P}^{\binom{n}{d}-1}(\mathbb{F})$ *and let* $\mathbf{a} \in \mathbb{P}^{2^{n-1}-1}(\mathbb{F})$ *denote the vector* *given by*

$$a_S = \begin{cases} q_S & \textit{if } S \subseteq [n-1], |S| = d \\ q_{S\cup n} & \textit{if } S \subseteq [n-1], |S| = d-1 \\ 0 & \textit{otherwise .} \end{cases}$$

The vector $\mathbf{q}$ *belongs to* $\mathrm{Gr}^2_{\mathbb{F}}(d,n)$ *if and only if*

(i) for all $i,j \in [n]$ *and* $S \subseteq [n]\backslash\{i,j\}$, $|S| = d-1$, $q_{S\cup i}q_{S\cup j}$ *is a square over* $\mathbb{F}$, *and*

(ii) for every $\gamma \in \mathrm{SL}_2(\mathbb{F})^{n-1} \rtimes S_{n-1}$, $(\gamma \cdot \mathrm{HypDet})(\mathbf{a}) = 0$.

Proof. $(\Rightarrow)$ This follows from applying Corollary 2.6.3 to (2.9).

$(\Leftarrow)$ Let $f = \sum_{S\in\binom{[n]}{d}} q_S \mathbf{x}^S$. This equals $f = f_{\mathbf{b}}$ where $\mathbf{b} \in \mathbb{F}^{2^n}$ is given by $b_S = q_S$ for $|S| = d$ and $b_S = 0$ otherwise. If $\mathbf{a} = (a_S)_S$ is defined as above, then $f_{\mathbf{a}}$ equals the restriction of f to $x_n = 1$ and f is the homogenezation of $f_{\mathbf{a}}$ to degree d with homogenizing variable x_n.

For any $i \neq j \in [n-1]$ and $S \subseteq [n-1]\backslash\{i,j\}$, the sizes of S and $S \cup \{i,j\}$ differ by two, implying that $a_S a_{S\cup ij}$ equals zero. Then $a_{S\cup i}a_{S\cup j} - a_S a_{S\cup ij} = a_{S\cup i}a_{S\cup j}$ is a square in $\mathbb{F}$ by assumption (i). By Theorem 3.5.6 and Corollary 2.6.3, $\Delta_{ij}(f_{\mathbf{a}})$ is a square in $\mathbb{F}[x_1, \ldots, x_{n-1}]$ for all $i,j \in [n-1]$, implying that $\Delta_{ij}(f)$ is a square in $\mathbb{F}[\mathbf{x}]$ for all $i,j \in [n-1]$. In particular, $\mathrm{Discr}_{x_n}(\Delta_{ij}(f))$ is identically zero. One can check that for any i,j,n, $\mathrm{Discr}_{x_n}(\Delta_{ij}(f))$ equals $\mathrm{Discr}_{x_j}(\Delta_{in}(f))$. This shows that $(\gamma \cdot \mathrm{HypDet})(\mathbf{b}) = 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$. Assumption (i) implies that $b_{S\cup i}b_{S\cup j} - b_S b_{S\cup ij}$ is a square in $\mathbb{F}$ for all $i,j \in [n]$ and $S \subseteq [n]\backslash\{i,j\}$, since this is either zero or of the form $q_{S\cup i}q_{S\cup j}$. Corollary 2.6.3 then gives a representation $f = \lambda \det(\sum_{i=1}^n x_i v_i v_i^T + W)$ where $v_i \in \mathbb{F}^d$ and $W \in \mathrm{Sym}_d(\mathbb{F})$. The polynomial $\lambda \det(\sum_{i=1}^n x_i v_i v_i^T + yW) \in \mathbb{F}[\mathbf{x}, y]$ equals the homogenization of f to degree d. Since f is already homogeneous of degree d, this equals f and belongs to $\mathbb{F}[\mathbf{x}]$. Specializing to $y = 0$ gives the desired representation $f = \lambda \det(\sum_{i=1}^n x_i v_i v_i^T)$. □

Example 2.6.5. ($d = 2$, $n = 4$) The Grassmannian $\mathrm{Gr}_{\mathbb{F}}(2,4)$ is cut out by one Plücker relation $p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23} = 0$ in $\mathbb{P}^5(\mathbb{F})$. Taking $q_{ij} = p_{ij}^2$ and eliminating the variables

p_{ij} gives the defining equation

$$q_{12}^2q_{34}^2 + q_{13}^2q_{24}^2 + q_{14}^2q_{23}^2 - 2q_{12}q_{13}q_{24}q_{34} - 2q_{12}q_{14}q_{23}q_{34} - 2q_{13}q_{14}q_{23}q_{24} = 0$$

for $\mathrm{Gr}_{\mathbb{F}}^2(2,4)$. This is exactly the hyperdeterminant $\mathrm{HypDet}(\mathbf{a})$ where $\mathbf{a} = (a_S)_{S\subseteq[3]} \in \mathbb{F}^{2^3}$ is given by $a_\emptyset = a_{123} = 0$, $a_i = q_{i4}$ and $a_{ij} = q_{ij}$ for all $i, j \in [3]$.

2.6.3 Determinantal representations in higher degrees

For any $r \in \mathbb{Z}_+$, let $\mathrm{Sym}_n^r(\mathbb{F})$ denote the set of symmetric matrices over $\mathbb{F}$ that can be written as a sum of r rank-one matrices over $\mathbb{F}$, i.e.

$$\mathrm{Sym}_n^r(\mathbb{F}) = \left\{ \sum_{i=1}^r v_i v_i^T \ : \ v_1, \ldots, v_r \in \mathbb{F}^n \right\}.$$

If $\mathbb{F}$ is algebraically closed with $\mathrm{char}(\mathbb{F}) \neq 2$, this is just the set of matrices of rank $\leq r$. For $\mathbb{F} = \mathbb{R}$ this is the set of positive semidefinite matrices of rank $\leq r$.

Theorem 2.6.6. *The set of polynomials* $\mathbb{F}[\mathbf{x}]_{\leq \mathbf{d}}$ *with a determinantal representation*

$$f = \lambda \det\left(\sum_{i=1}^n x_i A_i + B\right) \ \textit{with } A_i \in \mathrm{Sym}_m^{d_i}(\mathbb{F}) \textit{ and } B \in \mathrm{Sym}_m(\mathbb{F}) \tag{2.10}$$

for some $m \in \mathbb{N}$ *is invariant under the action of* $\mathrm{SL}_2(\mathbb{F})^n$.

Proof. The invariance under the action of S_n is immediate. It remains to check the invariance under $\mathrm{SL}_2(\mathbb{F})^n$. Suppose that $f = \det(\sum_{i=1}^n x_i A_i + B)$, where $A_i \in \mathrm{Sym}_m^{d_i}(\mathbb{F})$ and $B \in \mathrm{Sym}_m(\mathbb{F})$. First, suppose that $\mathbf{d} = \mathbf{1}$ and let $\gamma \in \mathrm{SL}_2(\mathbb{F})^n$. By Corollary 2.2.3, $\Delta_{ij}(\gamma \cdot f)$ is a square for all i, j. Then by Theorem 3.5.6, $\gamma \cdot f$ has a determinantal representation as in (3.5).

Now consider arbitrary $\mathbf{d} = (d_1, \ldots, d_n)$. By definition, we can write each matrix A_i as a sum of d_i matrices A_{ij} each of the form vv^T for some $v \in \mathbb{F}^m$. Then consider

$$F = \det\left(\sum_{i=1}^n \sum_{j=1}^{d_i} y_{ij} A_{ij} + B\right) \ \in \ \mathbb{F}\big[y_{ij} : i \in [n], j \in [d_i]\big].$$

Note that there is an inclusion $\phi : \mathrm{SL}_2(\mathbb{F})^n \to \mathrm{SL}_2(\mathbb{F})^{d_1+\dots+d_n}$, given by $(\phi(\gamma))_{ij} = \gamma_i$ for all $i \in [n]$ and $j \in [d_i]$. By construction, the restriction of $\phi(\gamma) \cdot F$ given by $y_{ij} = x_i$ for all i, j gives $\gamma \cdot f$. That is,

$$(\phi(\gamma) \cdot F)\,|_{y_{ij}=x_i} = \gamma \cdot f.$$

By the case $\mathbf{d} = \mathbf{1}$, $\phi(\gamma) \cdot F$ is determinantal. That is, there are some matrices $C_{11}, \dots, C_{nd_n}, D$ with $C_{ij} \in \mathrm{Sym}^1_m(\mathbb{F})$ so that $\phi(\gamma) \cdot F$ equals $\det(\sum_{i=1}^n \sum_{j=1}^{d_i} y_{ij} C_{ij} + D)$. Then $\gamma \cdot f$ equals $\det(\sum_{i=1}^n x_i C_i + D)$ where $C_i = \sum_{j=1}^{d_i} C_{ij} \in \mathrm{Sym}^{d_i}_m(\mathbb{F})$. $\square$

One motivation for studying such polynomials comes from definite determinantal representations over $\mathbb{R}$ and their connection with stable polynomials. A real polynomial $f \in \mathbb{R}[\mathbf{x}]$ is *stable* if it has no zeros with strictly positive imaginary parts, i.e. $f(\mathbf{z}) \neq 0$ for all $\mathbf{z} \in \mathbb{C}^n$ with $\mathrm{Im}(\mathbf{z}) \in \mathbb{R}^n_+$. Equivalently, f is stable if and only if the polynomial $f(t\mathbf{v} + \mathbf{w}) \in \mathbb{R}[t]$ is real-rooted for all $\mathbf{v} \in \mathbb{R}^n_+$ and $\mathbf{w} \in \mathbb{R}^n$. Over $\mathbb{F} = \mathbb{R}$, any polynomial with a determinantal representation of the form (2.10) is stable (see, e.g. [72, Prop. 2.1]), but not every stable polynomial has such a representation (see [15]).

The action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{C}$ given by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}$ preserves the upper half plane $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ and so the set of stable polynomials in $\mathbb{R}[\mathbf{x}]$ is invariant under the action of $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$. One consequence of Theorem 2.6.6 is that the set of polynomials with a semidefinite determinantal representation is also invariant under the action of this group.

Corollary 2.6.7. *The set of polynomials in* $\mathbb{R}[\mathbf{x}]_{\leq \mathbf{d}}$ *with a determinantal representation*

$$f = \lambda \det\left(\sum_{i=1}^{n} x_i A_i + B\right) \text{ with } A_1, \dots, A_n, B \in \mathrm{Sym}_m(\mathbb{R}) \text{ and } A_1, \dots, A_n \succeq 0$$

for some $m \in \mathbb{N}$ *and* $\lambda \in \mathbb{R}$ *is invariant under the action of* $\mathrm{SL}_2(\mathbb{R})^n$.

Chapter 3

DETERMINANTAL REPRESENTATIONS AND THE IMAGE OF PRINCIPAL MINOR MAP THE HERMITIAN AND GENERAL CASES

This chapter is based on a joint work with Cynthia Vinzant [5].

3.1 Introduction

In this chapter we consider the image of any $n \times n$ complex matrix under the principal minor map. As in the symmetric case, to each vector $\mathbf{a} = (a_S)_{S \subset [n]} \in \mathbb{F}^{2^n}$, we assign a multiaffine polynomial $f_{\mathbf{a}}$ where $f_{\mathbf{a}} = \sum_{S \subset [n]} a_S \prod_{i \in [n] \setminus S} x_i$. This transforms the problem of characterizing the image of the principal minor map to the problem of characterizing multiaffine polynomials with determinantal representation, these are polynomials that can be written in the following form: $f = \det(\operatorname{diag}(x_1, \ldots, x_n) + A)$ for some $n \times n$ matrix A. Symmetric (Hermitian) multiaffine determinantal polynomials are determinantal polynomials that corresponds to symmetric (Hermitian) matrices. In the previous chapter, we prove that the class of symmetric determinantal multiaffine polynomials is characterized by their Rayleigh differences being squares. In this chapter we also use them to characterize determinantal multiaffine polynomials.

This chapter is organized as follows. In Section 3.2, we introduce terminology and the basic properties of determinantal representations and the action of $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$. In Section 3.3, we give a characterization of multiaffine determinantal polynomials involving the factoring of Rayleigh differences. For Hermitian determinantal representations, this condition simplifies and we give an algorithm for constructing such representations from a factorization, as described in Section 3.4 and Section 3.5. In Section 3.6 we give a characterization of multiaffine stable determinantal polynomials and prove Theorem 3.6.4. In Section 3.7, we translate these conditions into explicit equations and inequalities whose orbit under $\mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$ cuts of the image of Hermitian matrices under the principal minor map. Finally, in Section 3.8, we conclude by presenting a family of examples that disproves the existing of such a finite description for the image of general $n \times n$ matrices under the principal minor map.

3.2 Background and notation

For a commutative ring R, we use $R[\mathbf{x}]$ to denote the polynomial ring $R[x_1, \ldots, x_n]$ and for $f \in R[\mathbf{x}]$, we use $\deg_i(f)$ to denote the degree of f in the variable x_i. For $\mathbf{d} = (\mathbf{d}_1, \ldots, \mathbf{d}_n)$ with $d_i \in \mathbb{Z}_{\geq 0}$, let $\mathbb{F}[\mathbf{x}]_{\leq \mathbf{d}}$ denote the set of polynomials with degree at most $\mathbf{d}_i$ in x_i for each $i = 1, \ldots, n$. These form an R-module of rank $\prod_{i=1}^n (\mathbf{d}_i + 1)$. When $\mathbf{d}_1 = \ldots = \mathbf{d}_n = m$, we abbreviate $R[\mathbf{x}]_{\leq (m,\ldots m)}$ by $R[\mathbf{x}]_{\leq \mathbf{m}}$. Of particular interest are *multiaffine polynomials*, with degree ≤ 1 in each variable, and *multiquadratic polynomials*, with degree ≤ 2 in each variable. These are denoted by $R[\mathbf{x}]_{\leq \mathbf{1}} = R[\mathbf{x}]_{\mathrm{MA}}$ and $R[\mathbf{x}]_{\leq \mathbf{2}} = R[\mathbf{x}]_{\mathrm{MQ}}$, respectively.

We use $\mathrm{Mat}_n(\mathbb{F})$ to denote the set of $n \times n$ matrices with entries in $\mathbb{F}$. When $\mathbb{K}$ is a field with an automorphic involution $a \mapsto \overline{a}$, we use $\mathrm{Her}_n(\mathbb{K})$ to denote the set of matrices $A \in \mathrm{Mat}_n(\mathbb{K})$ for which $\overline{A} = A^T$. Note that for $\mathbb{K} = \mathbb{C}$ and $a \mapsto \overline{a}$ given by complex conjugation, this is the usual set of $n \times n$ Hermitian matrices.

3.2.1 The action of $\mathrm{SL}_2(R)^n \rtimes S_n$ *and homogenezations*

The action of $\mathrm{SL}_2(R)^n$ on $R[\mathbf{x}]_{\leq \mathbf{d}}$ is defined as follows. Let $\gamma = (\gamma_i)_{i \in [n]}$ in $\mathrm{SL}_2(R)^n$ where $\gamma_i = \left(\begin{smallmatrix} a_i & b_i \\ c_i & d_i \end{smallmatrix}\right)$. Then for $f \in R[\mathbf{x}]_{\leq \mathbf{d}}$,

$$\gamma \cdot f = \prod_{i=1}^n (c_i x_i + d_i)^{\mathbf{d}_i} \cdot f\left(\frac{a_1 x_1 + b_1}{c_1 x_1 + d_1}, \ldots, \frac{a_n x_n + b_n}{c_n x_n + d_n}\right).$$

One way to interpret this action is with the multihomogenezation of f. Let $f^{\mathbf{d}-\mathrm{hom}}$ in $R[x_1, \ldots, x_n, y_1, \ldots, y_n]_{\mathbf{d}}$ denote the polynomial

$$f^{\mathbf{d}-\mathrm{hom}} = \prod_{i=1}^n y_i^{\mathbf{d}_i} \cdot f\left(x_1/y_1, \ldots, x_n/y_n\right).$$

The induced action of γ on $f^{\mathbf{d}-\mathrm{hom}}$ is just a linear change of coordinates:

$$\gamma \cdot f^{\mathbf{d}-\mathrm{hom}} = f^{\mathbf{d}-\mathrm{hom}}\left(\gamma_1 \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \ldots, \gamma_n \cdot \begin{pmatrix} x_n \\ y_n \end{pmatrix}\right).$$

Restricting to $y_1 = \ldots y_n = 1$ gives back $\gamma \cdot f$.

We will also use the usual homogenization of a polynomial to some total degree d, using a single homogenizing variable y. That is, for $f = \sum_\alpha c_\alpha \mathbf{x}^\alpha \in R[\mathbf{x}]$ of total degree $d = \deg(f)$, its homogenization is

$$f^{\mathrm{hom}} = y^d f\left(x_1/y, \ldots, x_n/y\right) = \sum_\alpha c_\alpha \mathbf{x}^\alpha y^{d-|\alpha|} \ \in \ R[\mathbf{x}, y].$$

Suppose that $\mathbb{K}$ is a field with an automorphic involution $a \mapsto \overline{a}$ with fixed field $\mathbb{F}$. This extends to an involution on $\mathbb{K}[\mathbf{x}]$ by acting on the coefficients. We will say that a polynomial $q \in \mathbb{F}[\mathbf{x}]$ is a *Hermitian square* if $q = g\overline{g}$ for some $g \in \mathbb{K}[\mathbf{x}]$. To end this section, we remark that for $f \in \mathbb{F}[\mathbf{x}]$, the condition that $\Delta_{ij}(f)$ is a Hermitian square is robust to homogenization.

Proposition 3.2.1. *Suppose that $\mathbb{K}$ is a field with an automorphic involution $a \mapsto \overline{a}$ with fixed field $\mathbb{F}$. Let $f \in \mathbb{F}[\mathbf{x}]$. For $i, j \in [n]$, the polynomial $\Delta_{ij}(f)$ is a Hermitian square if any only if $\Delta_{ij}(f^{\mathrm{hom}})$ is a Hermitian square.*

Proof. If $\Delta_{ij}(f^{\mathrm{hom}})$ is a Hermitian square, then specializing to $y = 1$ gives a representation of $\Delta_{ij}(f)$ as a Hermitian square. For the converse, let $f \in \mathbb{F}[\mathbf{x}]$ with total degree d and suppose that $\Delta_{ij} f = g\overline{g}$ for some $g \in \mathbb{K}[\mathbf{x}]$. Let $m = \deg(g) = \deg(\overline{g})$. By definition, $\Delta_{ij}(f^{\mathrm{hom}}) \in \mathbb{F}[\mathbf{x}, y]$ is homogeneous of degree $2d - 2$. Its restriction to $y = 1$ equals $\Delta_{ij} f$. Therefore $\Delta_{ij}(f^{\mathrm{hom}})$ equals $y^{2d-2-2m}(\Delta_{ij}(f))^{\mathrm{hom}}$, which is the Hermitian square $h\overline{h}$ where h is the homogenezation of g to total degree $2d - 1$. $\square$

3.2.2 The action of $\mathrm{SL}_2(\mathbb{F})$ *on matrices*

Given a matrix $A \in \mathrm{Mat}_n(\mathbb{F})$, consider the multiaffine polynomial $f = \det(\mathrm{diag}(x_1, \ldots, x_n) + A)$. For $\gamma = (\gamma_i)_{i\in[n]}$ in $\mathrm{SL}_2(\mathbb{F})^n$ with $\gamma_i = \left(\begin{smallmatrix} a_i & b_i \\ c_i & d_i \end{smallmatrix}\right)$, $\gamma \cdot f$ is defined by:

$$\gamma \cdot f = \prod_{i=1}^{n} (c_i x_i + d_i) \cdot \det\left(\mathrm{diag}\left(\frac{a_1 x_1 + b_1}{c_1 x_1 + d_1}, \ldots, \frac{a_n x_n + b_n}{c_n x_n + d_n}\right) + A\right).$$

Let A_i denote the ith column of A and e_i the vector whose ith entry is one and zero otherwise. By using the factor $(c_i x_i + d_i)$ to scale the ith column, we see that

$$\gamma \cdot f = \det\left(C\,\mathrm{diag}(x_1, \ldots, x_n) + B\right)$$

where C is the matrix with ith column $C_i = (a_i e_i + c_i A_i)$ and B is the matrix with ith column $B_i = b_i e_i + d_i A_i$. When the matrix C is invertible, this gives

$$\gamma \cdot f = \det(C) \det\left(\mathrm{diag}(x_1, \ldots, x_n) + C^{-1}B\right).$$

Up to the scalar multiple $\det(C)$, the coefficients of $\gamma \cdot f$ are the principal minors of the matrix $C^{-1}B$.

3.2.3 Resultants

For two univariate polynomials $a = \sum_{j=0}^{d} a_j t^j$ with $a_d \neq 0$ and $b = b_1 t + b_0$ with $b_1 \neq 0$ we define the resultant of a, b with respect to the variable t to be

$$\mathrm{Res}_t(a, b) = \sum_{j=0}^{d} a_j (-b_0)^j (b_1)^{d-j}.$$

Over an algebraically closed field, this polynomial vanishes if and only if the univariate polynomials a and b have a common root. See, for example, [22, §3.5]. We will focus on multiaffine polynomials and so focus on resultants in degree $d = 1$. For $k = 1, \ldots, n$, define

$$\mathrm{res}_{x_k}(g, h) = (g|_{x_k=0}) \cdot \frac{\partial}{\partial x_k} h - (h|_{x_k=0}) \cdot \frac{\partial}{\partial x_k} g.$$

In particular, if g and h both have degree one in x_k, this agrees with $\mathrm{Res}_{x_k}(g, h)$. The benefit of this degree-dependent definition is that it is invariant under the action of $\mathrm{SL}_2(R)$.

If $f \in R[\mathbf{x}]$ has degree ≤ 1 in both x_i and x_j, then

$$\Delta_{ij}(f) = \mathrm{res}_{x_i}\left(\frac{\partial f}{\partial x_j}, f|_{x_j=0}\right) = \mathrm{res}_{x_j}\left(\frac{\partial f}{\partial x_i}, f|_{x_i=0}\right). \tag{3.1}$$

Proposition 3.2.2. *If $f \in R[x_1, \dots, x_n]$ has degree one in each of x_i and x_j, then $\Delta_{ij}(f) = 0$ if and only if f factors into polynomial $g \cdot h$ with $g \in R[x_k : k \neq i]$ and $h \in R[x_k : k \neq j]$.*

Proof. By assumption we can write $f = ax_ix_j + bx_i + cx_j + d$ for $a, b, c, d \in R[x_k : k \neq i, j]$. Then $\Delta_{ij}(f) = bc - ad$. If $\Delta_{ij}(f) = 0$, then there is some factorization $b = b_1b_2$ and $c = c_1c_2$ for which $a = b_1c_1$ and $d = b_2c_2$. Then $f = (b_1x_i + c_2)(c_1x_j + b_2)$. Similarly, if $f = (b_1x_i+c_2)(c_1x_j+b_2)$ for some $b_1, b_2, c_1, c_2 \in R[x_k : k \neq i, j]$, then $\Delta_{ij}(f) = bc-ad = 0$. □

Proposition 3.2.3. *Let $g \in R[\mathbf{x}]_{\leq \mathbf{d}}$ and $h \in R[\mathbf{x}]_{\leq \mathbf{e}}$ with $\mathbf{d}_k = \mathbf{e}_k = 1$. For $\gamma \in \mathrm{SL}_2(R)^n$,*

$$\gamma \cdot \mathrm{res}_k(g, h) = \mathrm{res}_k(\gamma \cdot g, \gamma \cdot h),$$

where γ acts of on $\mathrm{res}_k(g, h)$ as polynomial of multidegree $\leq \mathbf{d} + \mathbf{e} - 2 \cdot \mathbf{1}_k$ with $\mathbf{1}_k$ is the vector with kth entry is 1 and zero otherwise.

Proof. Write $g = g_1x_k + g_0$ and $h = h_1x_k + h_0$ where g_1, g_0, h_1, h_0 are polynomials in the polynomial ring $R[x_j : j \neq k]$. The resultant $\mathrm{res}_k(g, h)$ is the determinant of the 2×2 matrix $\begin{pmatrix} h_1 & h_0 \\ g_1 & g_0 \end{pmatrix}$. Consider $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(R)$ acting on the jth coordinate. If $j = k$, then

$$\gamma \cdot g = g_1(ax_k + b) + g_0(cx_k + d), \quad \text{and} \quad \gamma \cdot h = h_1(ax_k + b) + h_0(cx_k + d).$$

Taking coefficients with respect to $\{1, x_k\}$, we see that the $\mathrm{res}_{x_k}(\gamma \cdot g, \gamma \cdot h)$ equals

$$\det\begin{pmatrix} ah_1 + ch_0 & bh_1 + dh_0 \\ ag_1 + cg_0 & bg_1 + dg_0 \end{pmatrix} = \det\left(\begin{pmatrix} h_1 & h_0 \\ g_1 & g_0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \det\begin{pmatrix} h_1 & h_0 \\ g_1 & g_0 \end{pmatrix} = \mathrm{res}_{x_k}(g, h).$$

Since γ acts on $R[\mathbf{x}]_{\leq \mathbf{d}+\mathbf{e}-2\cdot\mathbf{1}_k}$ as the identity, this equals $\gamma \cdot \mathrm{res}_{x_k}(g, h)$.

If $j \neq k$, then $\gamma \cdot g = (\gamma \cdot g_1)x_k + (\gamma \cdot g_0)$ and $\gamma \cdot h = (\gamma \cdot h_1)x_k + (\gamma \cdot h_0)$, where γ acts on g_1, g_0 and h_1, h_0 as elements of multidegree $\mathbf{d} - \mathbf{1}_k$ and $\mathbf{e} - \mathbf{1}_k$, respectively. It follows that

$$\operatorname{res}_{x_k}(\gamma \cdot g, \gamma \cdot h) = \det \begin{pmatrix} \gamma \cdot h_1 & \gamma \cdot h_0 \\ \gamma \cdot g_1 & \gamma \cdot g_0 \end{pmatrix} = \gamma \cdot \operatorname{res}_{x_k}(g, h).$$

□

From (3.1), this gives the following:

Corollary 3.2.4. *Consider an element* $\gamma \in \mathrm{SL}_2(R)^n$ *that acts by* $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *in the* k*-th coordinate and the identity in all others. For any* $f \in R[\mathbf{x}]_{\leq \mathbf{1}}$,

$$\Delta_{ij}(\gamma \cdot f) = \begin{cases} \Delta_{ij}(f) & \text{if } k = i, j \\ \gamma \cdot \Delta_{ij}(f) & \text{otherwise.} \end{cases}$$

3.3 Determinantal Representations and Rayleigh Differences

Let R be a unique factorization domain and denote by $\mathrm{Mat}_n(R)$ the set of $n \times n$ matrices with entries in R.

Theorem 3.3.1. *Let* $f \in R[x_1, \ldots, x_n]$ *be multiaffine in the variables* $x_1, \ldots, x_n$ *with its coefficient of* $x_1 \cdots x_n$ *equals one. Then* $f = \det(\mathrm{diag}(x_1, \ldots, x_n) + A)$ *for some* $A \in \mathrm{Mat}_n(R)$ *if and only if for every* $i \neq j \in [n]$*, the polynomials* $\Delta_{ij}(f)$ *factor as the product* $g_{ij} \cdot g_{ji}$ *where*

(a) $g_{ij} \in R[x_k : k \neq i, j]$ *is multiaffine in* $x_1, \ldots, x_n$ *and*

(b) for every $k \in [n] \backslash \{i, j\}$, $\operatorname{res}_{x_k}(g_{ij}, f) = g_{ik} g_{kj}$.

In this case, we can take g_{ij} *to be the* (i, j)*th entry of* $(\mathrm{diag}(x_1, \ldots, x_n) + A)^{\mathrm{adj}}$*, with* M^{adj} *represents the adjugate matrix of* M*.*

Proof of (⇒). This follows from a classical equality on the principal minors of an $n \times n$ matrix, used by Dodgson [25] as a method for computing determinants. This is also known as the Desnanot-Jacobi identity or more generally as Sylvester's determinantal identity. For subsets $S, T \subset [n]$ of equal cardinality, let $M(S,T)$ denote the submatrix of M obtained by *dropping* rows S and columns T from M. Then for any $i \neq j \in [n]$,

$$\det(M(i,k)) \cdot \det(M(j,\ell)) - \det(M) \cdot \det(M(\{i,j\},\{k,\ell\})) = \det(M(i,\ell)) \cdot \det(M(j,k)). \tag{3.2}$$

Note that for $M = \operatorname{diag}(x_1, \ldots, x_n) + A$ and any subset $S \subseteq [n]$, the principal minor $\det(M(S,S))$ equals the derivative of f with respect to the variables in S, $\left(\prod_{i \in S} \frac{\partial}{\partial x_i}\right) f$. The equation above with $k = i$ and $\ell = j$ then gives that $\Delta_{ij}(f)$ equals $\det(M(i,j)) \cdot \det(M(j,i))$.

For every $i, j \in [n]$, let g_{ij} denote $\det(M(i,j))$. Then $g_{ij} \in R[x_k : k \neq i,j]$ is multiaffine in $x_1, \ldots, x_n$ and $\Delta_{ij}(f) = g_{ij} g_{ji}$. Under an appropriate choice of indices, (3.2) gives

$$g_{kk} \cdot g_{ij} - f \cdot q = g_{ik} \cdot g_{kj} \quad \text{where} \quad q = \det(M(\{i,k\},\{j,k\})).$$

Note that $g_{kk} = \frac{\partial f}{\partial x_k}$ is the coefficient of x_k in f and q is the coefficient of x_k in g_{ij}. Therefore $g_{kk} \cdot g_{ij} - f \cdot q$ is the resultant of g_{ij} and f with respect to x_k. □

Example 3.3.2. For $n \geq 5$, one cannot remove condition (b) from Theorem 3.3.1. Consider

$$\begin{aligned} f = {} & x_1x_2x_3x_4x_5 + x_1x_2x_3x_4 + x_1x_2x_3x_5 + x_1x_2x_4x_5 + x_1x_3x_4x_5 + x_2x_3x_4x_5 \\ & + x_1x_2x_4 + x_1x_2x_5 + x_1x_3x_4 + x_2x_3x_5 + x_3x_4x_5. \end{aligned}$$

One can check that for every $i, j \in [5]$, $\Delta_{ij}(f)$ factors as the product of two multiaffine polynomials in $\mathbb{Q}[x_1, \ldots, x_5]$. For example, $\Delta_{12}(f) = -x_3x_4x_5(x_4x_5 - x_3 + x_4 + x_5)$. Since there in an irreducible factor involving all three variables, there is only one possible factorization of $\Delta_{12}(f)$ as the product of two multiaffine polynomials $g_{12} \cdot g_{21}$, up to scalar multiples and switching the factors, namely $g_{12} = -x_3x_4x_5$ and $g_{21} = x_4x_5 - x_3 + x_4 + x_5$. Taking the resultant of g_{21} and f with respect to x_3 gives

$$\operatorname{Res}_{x_3}(g_{21}, f) = (x_1x_5 + x_1 + x_5)(x_2x_4 + x_2 + x_4)(x_4x_5 + x_4 + x_5).$$

Each of the three quadratic factors are irreducible and so there is no way of writing this resultant as the product of *two* multiaffine polynomials. Therefore there is no choice of polynomials g_{23} and g_{31} satisfying the conditions in Theorem 3.3.1.

Lemma 3.3.3. *Let $f \in R[x_1, \dots, x_n]$ be multiaffine in the variables $x_1, \dots, x_n$ and its coefficient of $x_1 \cdots x_n$ equals one. If $f = g \cdot h$ for some $g, h \in R[x_1, \dots, x_n]$, then g and h are multiaffine in disjoint subsets of the variables $x_1, \dots, x_n$ and we can take their leading coefficients in these variables to be one. Moreover, if the polynomials $\Delta_{ij}(f)$ have factorizations satisfying conditions (a) and (b) in Theorem 3.3.1, then so do $\Delta_{ij}(g)$ and $\Delta_{ij}(h)$.*

Proof. For any $i \in [n]$, the degree of f in x_i must be the sum of the degrees of g and h in x_i. Since this sum of nonnegative numbers is one for each $i \in [n]$, we see that for some subset $I \subseteq [n]$, g is multiaffine in $\{x_i : i \in I\}$, h is multiaffine in $\{x_j : j \in [n] \backslash I\}$, and $\deg_i(h) = \deg_j(g) = 0$ for any $i \in I$ and $j \notin I$.

The highest degree term in f with respect to the variables $x_1, \dots, x_n$, $\prod_{i=1}^n x_i$, is the product of the highest degree terms in g and h. Therefore after rescaling, we can assume that both g and h have leading coefficient in these variables equal to 1. For $i \in I$ and $j \notin I$, $\partial(g \cdot h)/\partial x_i = h \cdot \partial g/\partial x_i$ and $\partial(g \cdot h)/\partial x_j = g \cdot \partial h/\partial x_j$. From this, one can check that $\Delta_{ij}(gh)$ equals $h^2 \Delta_{ij}(g)$ for $i, j \in I$, $g^2 \Delta_{ij}(h)$ for $i, j \in [n] \backslash I$ and zero otherwise.

Suppose that for $i, j \in [n]$, $\Delta_{ij}(f) = m_{ij} m_{ji}$ with m_{ij} multiaffine in $x_1, \dots, x_n$ and $\mathrm{res}_{x_k}(m_{ij}, f) = m_{ik} m_{kj}$ for every i, j, k. For $i, j \in I$, we see that $m_{ij} m_{ji} = h^2 \Delta_{ij}(g)$. Since m_{ij}, m_{ji} are multiaffine, they both must be divisible by h, leaving $\tilde{m}_{ij} \tilde{m}_{ji} = \Delta_{ij}(g)$, where $\tilde{m}_{ij}, \tilde{m}_{ji}$ are multiaffine in x_i for $i \in I$. Moreover, for k also in I,

$$h^2 \tilde{m}_{ik} \tilde{m}_{kj} = m_{ik} m_{kj} = \mathrm{Res}_{x_k}(m_{ij}, f) = \mathrm{res}_{x_k}(\tilde{m}_{ij} h, gh) = h^2 \mathrm{res}_{x_k}(\tilde{m}_{ij}, g)$$

showing that $\tilde{m}_{ik} \tilde{m}_{kj} = \mathrm{res}_{x_k}(\tilde{m}_{ij}, g)$. The desired factorization for $\Delta_{ij}(h)$ with $i, j \in [n] \backslash I$ follows similarly. $\square$

Proof of ($\Leftarrow$). Suppose that f is irreducible and homogeneous of degree n. Let G denote the $n \times n$ matrix with (i, j)th entry g_{ij} for $i \neq j$ and $g_{ii} := \frac{\partial f}{\partial x_i}$ for $i = j$.

We claim that all of the 2×2 minors of G lie in $\langle f \rangle$. This is immediate for the symmetric minors, as $g_{ii} g_{jj} - g_{ij} g_{ji} = f \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}$. Moreover, since $\frac{\partial f}{\partial x_1}$ is the coefficient of x_1 in f, the resultant $\mathrm{res}_{x_1}(g_{ij}, f)$ has the form $\frac{\partial f}{\partial x_1} g_{ij} - qf$ for some q. This gives $g_{11} g_{ij} - g_{i1} g_{1j} = qf$.

Finally, suppose that i, j, k, ℓ are all distinct. Then

$$g_{11}^2(g_{ij}g_{kl} - g_{il}g_{kj}) = (g_{11}g_{ij})(g_{11}g_{kl}) - (g_{11}g_{il})(g_{11}g_{kj}) \equiv g_{1i}g_{1j}g_{1k}g_{1l} - g_{1i}g_{1l}g_{1k}g_{1j} = 0 \mod \langle f\rangle.$$

Since f is irreducible and $g_{11} = \partial f/\partial x_1$ has smaller degree, g_{11} is not a zero-divisor in $R[x_1, \dots, x_n]/\langle f\rangle$. Therefore the minor $g_{ij}g_{kl} - g_{il}g_{kj}$ belongs to $\langle f\rangle$.

From this it follows that f^{k-1} divides the $k \times k$ minors of G for every $2 \leq k \leq n$, see [59, Lemma 4.7]. In particular, f^{n-2} divides the entries of the adjugate matrix G^{adj}. Let

$$M = (1/f^{n-2}) \cdot G^{\text{adj}}. \tag{3.3}$$

Also f^{n-1} divides $\det(G)$, and since these both have degree $n(n-1)$, there must be some constant $\lambda \in R$ for which $\det(G) = \lambda \cdot f^{n-1}$.

We can see that $\lambda = 1$ by taking top degree terms. Since $\deg(f_i) = n - 1$ and $\deg(g_{ij}) \leq n - 2$ for all $i \neq j$, the leading degree term of $\det(G)$ comes uniquely from the product of the diagonals $f_1 \cdots f_n$ and is therefore $(\prod_{i=1}^n x_i)^{n-1}$. On the righthand side, the leading degree term of f^{n-1} is also $(\prod_{i=1}^n x_i)^{n-1}$, showing that $\lambda = 1$. Then

$$\det(M) = \frac{1}{f^{n(n-2)}} \cdot \det(G^{\text{adj}}) = \frac{1}{f^{n(n-2)}} \det(G)^{n-1} = \frac{1}{f^{n(n-2)}} f^{(n-1)^2} = f.$$

Note that the entries of M have degree $\leq (n-1)^2 - n(n-2) = 1$, so we can write M as $M_0 + \sum_{i=1}^n x_i M_i$ for some matrices $M_i \in R^{n\times n}$. We claim that $\sum_{i=1}^n x_i M_i = \text{diag}(x_1, \dots, x_n)$.

To see this, first note that a non-principal $(n-1)\times(n-1)$ minor of G involves at most $n-2$ elements from the diagonal of G and therefore has degree $\leq (n-2)(n-1)+(n-2) = n(n-2)$, since the off-diagonal entries of G have degree $\leq n - 2$. Therefore the off diagonal entries of M have degree $\leq n(n-2) - n(n-2) = 0$.

Moreover in the expansion of any principal minor of G, there is a *unique* term of degree $(n-1)^2$, namely the product of the leading terms of the diagonal elements, $\prod_{j\neq i} \text{LT}(g_{jj})$.

We can therefore take the leading terms (3.3) to find that

$$\begin{aligned}\sum_{i=1}^{n} x_i M_i &= \frac{1}{(\mathrm{LT}(f))^{n-2}} \cdot \left(\mathrm{diag}\left(\mathrm{LT}(g_{11}), \ldots, \mathrm{LT}(g_{nn})\right)\right)^{\mathrm{adj}} \\ &= \frac{1}{\left(\prod_{j=1}^{n} x_j\right)^{n-2}} \cdot \left(\prod_{j=1}^{n} x_j \cdot \mathrm{diag}\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right)\right)^{\mathrm{adj}} \\ &= \mathrm{diag}\left(x_1, \ldots, x_n\right).\end{aligned}$$

Finally, for general f, we take a factorization of f into irreducible polynomials $f = \prod_\ell f_\ell$. By Lemma 3.3.3, for every i, j, $\Delta_{ij}(f_\ell)$ has a factorization $m_{ij}m_{ji}$ so into multiaffine polynomials m_{ij} with $\mathrm{Res}_{x_k}(m_{ij}, f) = m_{ik}m_{kj}$. By the arguments above, f_k has a determinantal representation of the correct form. Taking a block diagonal representation of these representations (and permuting the rows and columns if necessary to reorder $x_1, \ldots, x_n$) gives a determinantal representation for f. □

Remark 3.3.4. Theorem 3.3.1 the matrix $G = (g_{ij})_{ij}$ and corresponding determinantal representation $\mathrm{diag}(x_1, \ldots, x_n) + A$ of f satisfy

$$G = (\mathrm{diag}(x_1, \ldots, x_n) + A)^{\mathrm{adj}} \text{ and } (\mathrm{diag}(x_1, \ldots, x_n) + A) = f^{2-n} G^{\mathrm{adj}}.$$

Corollary 3.3.5. *Let* $f = \det(\mathrm{diag}(x_1, \ldots, x_n) + A)$ *with* $A \in \mathrm{Mat}_n(R)$ *and* $\gamma \in \mathrm{SL}_2(R)^n$. *If* $\beta = \mathrm{coeff}(\gamma \cdot f, \prod_{i=1}^{n} x_i)$ *is nonzero, then for some* $n \times n$ *matrix* B *with entries in* $\frac{1}{\beta}R$,

$$\gamma \cdot f = \beta \det(\mathrm{diag}(x_1, \ldots, x_n) + B).$$

Proof. Let $g_{ij} \in R[\mathbf{x}]$ denote the (i, j)th entry of $(\mathrm{diag}(x_1, \ldots, x_n) + A)^{\mathrm{adj}}$. We claim that $\frac{1}{\beta}\gamma \cdot f$ and $\frac{1}{\beta}\gamma \cdot g_{ij}$ in $R(\frac{1}{\beta})[\mathbf{x}]$ satisfy the conditions in Theorem 3.3.1. Here γ acts of f as a polynomial of multidegree $\mathbf{1}_{[n]}$ and on g_{ij} as a polynomial of multidegree $\mathbf{1}_{[n]\setminus\{i,j\}}$.

It is immediate that $\frac{1}{\beta}(\gamma \cdot f) \in R(\frac{1}{\beta})[\mathbf{x}]$ is multiaffine in $x_1, \ldots, x_n$ and has coefficient of $x_1 \cdots x_n$ equal to one. We first note that

$$\Delta_{ij}(\tfrac{1}{\beta}(\gamma \cdot f)) = \tfrac{1}{\beta^2}(\gamma \cdot \Delta_{ij}(f)) = (\tfrac{1}{\beta}\gamma \cdot g_{ij})(\tfrac{1}{\beta}\gamma \cdot g_{ji})$$

where γ acts on $\Delta_{ij}(f)$ as a polynomial of multidegree $2 \cdot \mathbf{1}_{[n]\setminus\{i,j\}}$. By Proposition 3.2.3,

$$\operatorname{res}_{x_k}\left(\tfrac{1}{\beta}(\gamma \cdot g_{ij}), \tfrac{1}{\beta}(\gamma \cdot f)\right) = \tfrac{1}{\beta^2}(\gamma \cdot \operatorname{res}_{x_k}(g_{ij}, f)) = \left(\tfrac{1}{\beta}\gamma \cdot g_{ik}\right)\left(\tfrac{1}{\beta}\gamma \cdot g_{kj}\right).$$

As the polynomials $\frac{1}{\beta}\gamma \cdot g_{ij}$ are multiaffine, this finishes the claim.

By Theorem 3.3.1, $\frac{1}{\beta}\gamma \cdot f$ equals $\det(\operatorname{diag}(x_1, \ldots, x_n) + B)$ for some $B \in \operatorname{Mat}_n(R(\frac{1}{\beta}))$.

We claim that βB has entries in R. By construction we have

$$\operatorname{diag}(x_1, \ldots, x_n) + B = (\tfrac{1}{\beta}\gamma \cdot f)^{2-n}(\tfrac{1}{\beta}\gamma \cdot G)^{\mathrm{adj}} = \tfrac{1}{\beta}(\gamma \cdot f)^{2-n}(\gamma \cdot G)^{\mathrm{adj}}.$$

For the last equality, we use that $(\frac{1}{\beta})^{2-n}(\frac{1}{\beta})^{n-1} = \frac{1}{\beta}$. Multiplying by β then gives

$$\beta\left(\operatorname{diag}(x_1, \ldots, x_n) + B\right) = (\gamma \cdot f)^{2-n}(\gamma \cdot G)^{\mathrm{adj}} \in \operatorname{Mat}_n(R[\mathbf{x}]), \tag{3.4}$$

showing that the entries of βB belong to R. □

From this, we see that $\mathrm{SL}_2(\mathbb{F})^n$ acts rationally on the set of matrices $A \in \operatorname{Mat}_n(\mathbb{F})$ for $\mathbb{F} = \operatorname{frac}(R)$. Namely, if $f = \operatorname{diag}(x_1, \ldots, x_n) + A$ and $\gamma \in \mathrm{SL}_2(\mathbb{F})^n$ with $\operatorname{coeff}(\gamma \cdot f, \prod_{i=1}^n x_i) = \beta \neq 0$, then as (3.4) in the proof of Corollary 3.3.5, $\beta\left(\operatorname{diag}(x_1, \ldots, x_n) + B\right) = (\gamma \cdot f)^{2-n}(\gamma \cdot G)^{\mathrm{adj}}$ for some $B \in \operatorname{Mat}_n(\mathbb{F})$. We can then define $\gamma \cdot A = B$.

Similarly, for a field $\mathbb{F}$, the mutiplicative group $(\mathbb{F}^*)^n$ acts on $n \times n$ matrices by diagonal conjugation. Namely, for $\lambda = (\lambda_1, \ldots, \lambda_n)$ we define

$$\lambda \cdot A := D^{-1}AD,$$

where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

Proposition 3.3.6. *The action of* $\mathrm{SL}_2(\mathbb{F})^n$ *on* $\operatorname{Mat}_n(\mathbb{F})$ *commutes with diagonal conjugation.*

Proof. Let $A \in \operatorname{Mat}_n(\mathbb{F})$ with $f = \det(\operatorname{diag}(x_1, \ldots, x_n) + A)$ and $\gamma \in \mathrm{SL}_2(\mathbb{F})^n$ for which $\operatorname{coeff}(\gamma \cdot f, \prod_{i=1}^n x_i) = \beta \neq 0$. Let $G = (g_{ij})_{ij} = (\operatorname{diag}(x_1, \ldots, x_n) + A)^{\mathrm{adj}}$.

For $\lambda \in (\mathbb{F}^*)^n$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, we see that $\frac{\lambda_j}{\lambda_i}(\gamma \cdot g_{ij}) = \gamma \cdot (\frac{\lambda_j}{\lambda_i} g_{ij})$ and so $\gamma \cdot (D^{-1}GD) = D^{-1}(\gamma \cdot G)D$. Then

$$\begin{aligned}\operatorname{diag}(x_1, \ldots, x_n) + D^{-1}(\gamma \cdot A)D &= \alpha(\gamma \cdot f)^{2-n} D^{-1}(\gamma \cdot G)^{\mathrm{adj}} D \\ &= \alpha(\gamma \cdot f)^{2-n}(\gamma \cdot (D^{-1}GD))^{\mathrm{adj}} \\ &= \operatorname{diag}(x_1, \ldots, x_n) + \gamma \cdot (D^{-1}AD).\end{aligned}$$

□

3.4 *Multiaffine algebra for constructing Hermitian factorizations*

In this section, we develop an algorithm for constructing factorizations that satisfy the conditions in Theorem 3.3.1. To do this, we find it most convenient to work in the following level of generality throughout this section. Let S be a unique factorization domain with an automorphic involution $a \mapsto \overline{a}$. We use 0 and 1 to denote the additive and multiplicative identities of S. The map $S \to S$ given by $a \mapsto \overline{a}$ then must satisfy

$$\overline{(\overline{a})} = a, \ \overline{0} = 0, \ \overline{1} = 1, \ \overline{a+b} = \overline{a} + \overline{b} \text{ and } \overline{a \cdot b} = \overline{a} \cdot \overline{b}.$$

for all $a, b \in S$. Let R be the subring of elements fixed by this automorphism, that is $R = \{a \in S : \overline{a} = a\}$.

The example of interest is the ring $S = \mathbb{C}[x_{n+1}, \ldots, x_m]$ of polynomials with complex coefficients with the involution given by complex conjugation. In this case the fixed ring is the subring of polynomials whose coefficients are real, $R = \mathbb{R}[x_{n+1}, \ldots, x_m]$.

Assumptions 3.4.1. *: Let $f \in R[x_1, \ldots, x_n]$ satisfy the following:*

1. *f is irreducible in $R[x_1, \ldots, x_n]$,*

2. *f has degree ≤ 1 in each variable $x_1, \ldots, x_n$,*

3. *the coefficient $\prod_{i=1}^{n} x_i$ in f is nonzero,*

4. *for every* $1 \leq i < j \leq n$, $\Delta_{ij}(f)$ *factors as* $g_{ij}\overline{g_{ij}}$ *in* $S[x_1, \ldots, x_n]$, *and*

5. *for every* $1 \leq i \leq n$, *the partial derivative* $\frac{\partial f}{\partial x_i}$ *is irreducible in* $R[x_1, \ldots, x_n]$ *up to a constant. That is, for any factorization* $\frac{\partial f}{\partial x_i} = g \cdot h$ *in* $R[x_1, \ldots, x_n]$, $g \in R$ *or* $h \in R$.

In what follows, we will build up tools to show that under these assumptions Algorithm 1 produces the desired representation of f. We first exploit some properties of multiaffine polynomials. For any disjoint subsets $S, T \subset [n]$, let

$$f_S^T = \prod_{i \in S} \partial_i \cdot f|_{\{x_j = 0 \,:\, j \in T\}}.$$

Note that if f is multiaffine in $x_1, \ldots, x_n$, then for any $1 \leq i < j \leq n$, we have

$$f = x_i f_i + f^i, \quad f_i = x_j f_{ij} + f_i^j, \quad \text{and} \quad f^i = x_j f_j^i + f^{ij}.$$

From this, one can check that the formula for $\Delta_{ij} f$ can be written without x_i, x_j:

$$\Delta_{ij} f = f_i^j \cdot f_j^i - f^{ij} \cdot f_{ij}$$

If in addition we assume that f and all its partial derivatives are irreducible, then $\Delta_{ij}(f)$ will have degree exactly 2 in each variable, as the following lemma shows.

Lemma 3.4.2. *If* f *satisfies Assumptions 3.4.1, then for all* $1 \leq i, j \leq n$, $\Delta_{ij}(f)$ *is quadratic in each variable* x_k *for* $k \in [n] \setminus \{i, j\}$.

Proof. For $1 \leq i < j \leq n$, we write $\Delta_{ij}(f)$ as a quadratic polynomial in the variable x_k:

$$\Delta_{ij}(f) = f_i f_j - f_{ij} f = (f_{ik} x_k + f_i^k)(f_{jk} x_k + f_j^k) - (f_{ijk} x_k + f_{ij}^k)(f_k x_k + f^k),$$

which gives

$$\operatorname{coeff}(\Delta_{ij}(f), x_k^2) = f_{ik} f_{jk} - f_{ijk} f_k = \Delta_{ij}(f_k).$$

If $\Delta_{ij}(f_k) = 0$, then by Proposition 3.2.2, f_k is reducible, contradicting Assumptions 3.4.1(5).

□

We next use ring maps given by taking resultants with f. For any $i = 1, \ldots, n$, define

$$\varphi_i : S[x_1, \ldots, x_m] \to S[x_k : k \neq i] \text{ by } \varphi_i(g) = \operatorname{Res}_{x_i}(g, f).$$

For instance if we restrict to polynomials $g = g_j x_j + g^j$ with degree one in x_j, then

$$\varphi_j(g, f) = -g_j f^j + g^j f_j.$$

First we will start by listing some of the properties of these maps.

Lemma 3.4.3. *If f satisfies Assumptions 3.4.1, then, for all $g \in S[\mathbf{x}]$, the maps $\varphi_1, \ldots, \varphi_n$ satisfy the following:*

1. $\varphi_j(f_i) = \Delta_{ij}(f)$ *for all* $1 \leq i < j \leq n$,
2. $\varphi_j(\Delta_{ik}(f)) = \Delta_{ij}(f)\Delta_{jk}(f)$ *for all distinct* $1 \leq i, j, k \leq n$,
3. *if* $\deg_j(g) = 0$, *then* $\varphi_j(g \cdot h) = g \cdot \varphi_j(h)$ *for all* $h \in S[\mathbf{x}]$,
4. *if* $\deg_j(g) > 0$ *and* $\deg_j(h) > 0$, *then* $\varphi_j(g \cdot h) = \varphi_j(g) \cdot \varphi_j(h)$ *for all* $1 \leq j \leq n$,
5. *If* $\deg_i(g) = \deg_j(g) = 1$ *and* $sg_j \notin \langle f_j \rangle$ *for all* $s \in S$, *then* $\varphi_j \circ \varphi_i(g) = \Delta_{ij} f \cdot \varphi_j(g)$,
6. *If* $\deg_j(g) = 1$, $\varphi_j(g) \equiv f_j \cdot g$ *modulo* $\langle f \rangle$.

Proof. We will prove (5) and (6) and all the other properties follow similarly by direct computations. To prove property (5), we write $g = g_{ij} x_i x_j + g_i^j x_i + g_j^i x_j + g^{ij}$, then

$$\begin{aligned}
\varphi_j \circ \varphi_i(g) &= \varphi_j(-g_{ij} x_j f^i - g_i^j f^i + g_j^i f_i x_j + g^{ij} f_i) \\
&= \varphi_j \left((-g_{ij} f_j^i + g_j^i f_{ij}) x_j^2 + (-g_{ij} f^{ij} + g_j^i f_i^j + g^{ij} f_{ij} - g_i^j f_j^i) x_j + (g^{ij} f_i^j - g_i^j f^{ij}) \right).
\end{aligned}$$

Since for all $s \in S$, $sg_j \notin \langle f_j \rangle$, we see that $\operatorname{Coeff}_{x_j^2}(\varphi_i(g)) \neq 0$. Otherwise $g_j^i f_{ij} = g_{ij} f_j^i$, and since $f_j = f_{ij} x_i + f_j^i$ is irreducible up to a constant, then f_{ij} and f_j^i are relatively prime up to a constant $s \in S$. This implies that f_j^i and f_{ij} divide sg_j^i and sg_{ij} respectively and this implies that $sg_j \in \langle f_j \rangle$.

Applying the map φ_j and simplifying then gives

$$\varphi_j \circ \varphi_i(g) = \Delta_{ij}(f)(-g_j f^j + g^j f_j) = \Delta_{ij}(f)(\varphi_j(g))$$

To prove (6), we write g as $g = g_j x_j + g^j$ and we use $f^j = f - f_j x_j$

$$\varphi_j(g) = -g_j f^j + g^j f_j = -g_j(f - f_j x_j) + g^j f_j = -g_j f + f_j \ g$$

Therefore $\varphi_j(g) \equiv f_j \ g$ modulo $\langle f \rangle$. □

Lemma 3.4.4. *If f satisfies Assumptions 3.4.1 and $\Delta_{ij}(f) = p\overline{p}$ for some $1 \leq i < j \leq n$, then for every $k \in [n] \setminus \{i, j\}$, there is a factorization of each $\Delta_{ik}(f)$ and $\Delta_{jk}(f)$ into $q\overline{q}$ and $r\overline{r}$, respectively, such that $\varphi_k(p) = qr$.*

Proof. Since $\Delta_{ik}(f)$ and $\Delta_{jk}(f)$ factor into two conjugates, we can write

$$\Delta_{ik}(f) = a_1 \cdots a_s \cdot \overline{a_1} \cdots \overline{a_s} \qquad \text{and} \qquad \Delta_{jk}(f) = b_1 \cdots b_t \cdot \overline{b_1} \cdots \overline{b_t}$$

where $a_1, \ldots, a_s, b_1, \ldots, b_t$ are irreducible in $S[x_1, \ldots, x_m]$ that are multiaffine in $x_1, \ldots, x_n$. Then

$$\varphi_k(p)\varphi_k(\overline{p}) = \varphi_k(\Delta_{ij}(f)) = \Delta_{ik}(f)\Delta_{jk}(f) = a_1 \cdots a_s \cdot \overline{a_1} \cdots \overline{a_s} \cdot b_1 \cdots b_t \cdot \overline{b_1} \cdots \overline{b_t}.$$

After switching a_i with $\overline{a_i}$ and b_i with $\overline{b_i}$ if necessary, we get

$$\varphi_k(p) = a_1 \cdots a_s \cdot b_1 \cdots b_t = q \cdot r$$

where $q = a_1 \cdots a_s$ and $r = b_1 \cdots b_t$ are multiaffine polynomials such that $\Delta_{ik}(f) = q\overline{q}$ and $\Delta_{jk}(f) = r\overline{r}$ as desired. □

Lemma 3.4.5. *If f satisfies Assumptions 3.4.1 and for some distinct $1 \leq i, j, k \leq n$, the polynomials $\Delta_{ij}(f) = p\overline{p}$, $\Delta_{ik}(f) = q\overline{q}$ and $\Delta_{jk}(f) = r\overline{r}$ such that $\varphi_k(p) = qr$, then*

$$\varphi_j(q) = p\overline{r} \quad \textit{and} \quad \varphi_i(r) = p\overline{q}.$$

Proof. We will prove the first equality and the second holds similarly. First notice that since $\deg_i(p) = \deg_j(p) = 0$, $sp \notin \langle f_j \rangle$ for all $s \in S$. Also, $\deg_j(r) = \deg_k(r) = 0$. Then using the properties in Lemma 3.4.3 we get

$$\varphi_j(q)r = \varphi_j(qr) = \varphi_j \circ \varphi_k(p) = \Delta_{jk}(f) \cdot \varphi_j(p) = \Delta_{jk}(f) \cdot p.$$

Since $\Delta_{jk}(f) = r\overline{r}$, dividing the above equation by r gives the desired result. □

The following algorithm gives the desired factorizations of $\Delta_{ij}(f)$ into $g_{ij}\overline{g_{ij}}$ that satisfy the hypothesis of Theorem 3.3.1, which will in turn give the desired Hermitian determinantal representation in Theorem 3.5.1.

Algorithm 1 Compatible Hermitian factorizations of Rayleigh differences

Input: $f \in R[x_1, \ldots, x_n]$ satisfying Assumptions 3.4.1

Output: Polynomials $\{g_{jk} : 1 \leq j < k \leq n\}$ in $S[x_1, \ldots, x_n]$

```
Take $g_{12} \in \mathbb{K}[x_1, \ldots, x_n]$ so that $\Delta_{12}f = g_{12} \cdot \overline{g_{12}}$
for $k = 3, k \leq n, k{+}{+}$ do
    $Q_0 := \gcd\{\Delta_{1k}(f), \varphi_k(g_{12}), \ldots, \varphi_k(g_{1(k-1)})\}$
    Factor $\Delta_{1k}(f) = p_{k,1} \cdots p_{k,m_k} \cdot \overline{p_{k,1}} \cdots \overline{p_{k,m_k}}$ with $p_{k,j}$ irreducible for all $j$
    for $j = 1, j \leq m_k, j{+}{+}$ do
        if $p_{k,j}\overline{p_{k,j}}$ divides $Q_{j-1}$ then $Q_j := Q_{j-1}/\overline{p_{k,j}}$
        else $Q_j := Q_{j-1}$
    $g_{1k} := Q_{m_k}$
    for $j = 2, j \leq k-1, j{+}{+}$ do
        $g_{jk} := \overline{\varphi_k(g_{1j})/g_{1k}}$
```

Proposition 3.4.6. *The polynomials* $\{g_{ik}\}_{1 \leq i < k \leq n}$ *constructed in Algorithm 1 satisfy*

(a) g_{1k} *is multiaffine in* $x_1, \ldots, x_n$ *for all* $k > 1$*,*

(b) $\varphi_k(g_{1i}) = g_{1k}\overline{g_{ik}}$ *for all* $1 < i < k$*, and*

(c) $\Delta_{ik}(f) = g_{ik}\overline{g_{ik}}$ *for all* $1 \leq i < k$*.*

Proof. (a) This is immediate for $k = 2$. For $2 < k \leq n$, notice that $\Delta_{1k}(f)$ has degree two in $x_1, \ldots, x_n$. Let $\ell \in [n]\backslash\{1, k\}$ and let $p_{k,j}, \overline{p_{k,j}}$ be the unique irreducible factors of $\Delta_{1k}(f)$ with degree one in x_ℓ. By construction, g_{1k} divides Q_j, which in turn divides $\Delta_{1k}(f)/\overline{p_{k,j}}$. Since this quotient only has degree one in x_ℓ, g_{1k} must have degree ≤ 1 in x_ℓ.

(b) follows directly from construction.

(c) We proceed by induction on k. It is trivially true for $k = 2$. For the inductive step, we will prove the claim for $\Delta_{1k}(f)$ and the other cases follow. By construction, $g_{1k}\overline{g_{1k}}$ divides $\Delta_{1k}(f)$. To see this, note that for each $j = 1, \ldots, m_k$ in Algorithm 1, we can take $q_j = p_{k,j}$ if $p_{k,j}$ divides g_{1k} and $q_j = \overline{p_{k,j}}$ otherwise. Then, by construction, g_{1k} divides $q = \prod_{j=1}^{m_k} q_j$ and $q \cdot \overline{q} = \Delta_{1k}(f)$, showing that $g_{1k} \cdot \overline{g_{1k}}$ divides $\Delta_{1k}(f)$.

Suppose for the sake of contradiction that $\Delta_{1k}(f) \neq g_{1k}\overline{g_{1k}}$. Then there is some irreducible factor p of $\Delta_{1k}(f)$ such that $p\overline{p}$ does not divide $g_{1k}\overline{g_{1k}}$. We claim that for every $1 < i < k$, either p or $\overline{p}$ divides $\varphi_k(g_{1i})$. By induction, for $1 < i < k$, $g_{1i}\overline{g_{1i}} = \Delta_{1i}(f)$. Applying φ_k gives

$$\varphi_k(g_{1i}) \cdot \varphi_k(\overline{g_{1i}}) = \varphi_k(\Delta_{1i}(f)) = \Delta_{1k}(f) \cdot \Delta_{ik}(f).$$

Since p is irreducible and divides $\Delta_{1k}(f)$, it must divide either $\varphi_k(g_{1i})$ or $\varphi_k(\overline{g_{1i}}) = \overline{\varphi_k(g_{1i})}$. In the latter case, $\overline{p}$ divides $\varphi_k(g_{1i})$. Since neither p nor its conjugate divide g_{1k}, it follows from the construction that neither p nor $\overline{p}$ divide $Q_0 = \gcd\{\Delta_{1k}(f), \varphi_k(g_{12}), \ldots, \varphi_k(g_{1(k-1)})\}$. Hence there exists distinct $2 \leq i, j < k$ such that neither p divide $\varphi_k(g_{1i})$ nor $\overline{p}$ divide $\varphi_k(g_{1j})$. By switching p and $\overline{p}$ if needed, we can assume $i < j$.

By induction (on k), we know that $\Delta_{1i}(f) = g_{1i}\overline{g_{1i}}$, $\Delta_{1j}(f) = g_{1j}\overline{g_{1j}}$ and $\Delta_{ij}(f) = g_{ij}\overline{g_{ij}}$. Moreover, by (b), $\varphi_j(g_{1i}) = g_{1j}\overline{g_{ij}}$. Lemma 3.4.5 then implies that $\varphi_1(g_{ij}) = g_{1i}\overline{g_{1j}}$ and

$$\Delta_{1k}(f)\varphi_k(g_{ij}) = \varphi_k(\varphi_1(g_{ij})) = \varphi_k(g_{1i}\overline{g_{1j}}) = \varphi_k(g_{1i})\varphi_k(\overline{g_{1j}}).$$

Now neither $\varphi_k(g_{1i})$ nor $\varphi_k(\overline{g_{1j}}) = \overline{\varphi_k(g_{1j})}$ is divisible by p while p divides $\Delta_{1k}(f)$ and this gives the desired contradiction. Therefore $\Delta_{1k}(f) = g_{1k}\overline{g_{1k}}$.

For $1 < i < k$, we calculate that

$$g_{ik} \cdot \overline{g_{ik}} = \frac{\varphi_k(\overline{g_{1i}})}{\overline{g_{1k}}} \cdot \frac{\varphi_k(g_{1i})}{g_{1k}} = \frac{\varphi_k(\overline{g_{1i}}g_{1i})}{\Delta_{1k}(f)} = \frac{\Delta_{1k}(f)\Delta_{ik}(f)}{\Delta_{1k}(f)} = \Delta_{ik}(f).$$

□

Corollary 3.4.7. *If $f \in R[x_1, \dots, x_n]$ satisfies Assumptions 3.4.1, then there exists a factorization of $\Delta_{ij}(f)$ into $g_{ij}g_{ji}$ such that $g_{ij} \in S[x_1, \dots, x_n]$, $g_{ji} = \overline{g_{ij}}$, and $\varphi_k(g_{ij}) = g_{ik}g_{kj}$ for all distinct $1 \leq i, j, k \leq n$.*

Proof. Let $\{g_{ij} : 1 \leq i < j \leq n\}$ be the polynomials given by Algorithm 1 and for $i < j$ let $g_{ji} = \overline{g_{ij}}$. By Proposition 3.4.6, $\Delta_{ij}(f) = g_{ij}\overline{g_{ij}} = g_{ji}g_{ij}$. Since $\Delta_{ij}(f)$ is quadratic in each variable $x_1, \dots, x_n$, then g_{ij} is multiaffine in $x_1, \dots, x_n$. We will show that $\varphi_k(g_{ij}) = g_{ik}g_{kj}$ for all distinct i, j, k. Assuming that $i < j < k$ and using Proposition 3.4.6 we get

$$\begin{aligned} \operatorname{res}_{x_k}(g_{1i}, f) &= \varphi_k(g_{1i}) = g_{1k}\overline{g_{ki}} = g_{1k}g_{ki}, \text{ and} \\ \operatorname{res}_{x_k}(g_{j1}, f) &= \varphi_k(g_{j1}) = \overline{\varphi_k(g_{1j})} = \overline{g_{1k}\overline{g_{kj}}} = g_{jk}g_{k1}. \end{aligned}$$

Multiplying the above two equations and using Properties 3.4.3 we get

$$\varphi_k(g_{1i}g_{j1}) = \Delta_{1k}(f)g_{ki}g_{jk}.$$

Using Proposition 3.4.6 again, we know that $\varphi_j(g_{1i}) = g_{1j}g_{ji}$ and Lemma 3.4.5 implies that $\varphi_1(g_{ji}) = g_{j1}g_{1i}$. Again using Properties 3.4.3 we find that

$$\Delta_{1k}(f)\varphi_k(g_{ji}) = \Delta_{1k}(f)g_{ki}g_{jk}.$$

Since f is irreducible, $\Delta_{1k}(f)$ is nonzero and we conclude that $\varphi_k(g_{ij}) = \overline{\varphi_k(g_{ji})} = \overline{g_{ki}g_{jk}} = g_{ik}g_{kj}$. Using Lemma 3.4.5, we get that $\varphi_j(g_{ik}) = g_{ij}g_{jk}$ and $\varphi_i(g_{jk}) = g_{ji}g_{ik}$ as desired. □

Example 3.4.8. $(n = 4)$. Consider $f \in \mathbb{R}[x_1, x_2, x_3, x_4]$ given by

$$f(x_1, x_2, x_3, x_4) = x_1x_2x_3x_4 - x_1x_2 - x_1x_3 - x_1x_4 - x_2x_3 - x_2x_4 - x_3x_4 + 1$$

For any distinct $i, j, k, \ell \in [4]$, the Raleigh differences of f with respect to x_i and x_j is

$$\Delta_{ij}(f) = (x_k^2+1)(x_\ell^2+1) = (x_k - \mathbb{i})(x_k + \mathbb{i})(x_\ell + \mathbb{i})(x_\ell - \mathbb{i}).$$

Using Algorithm 1, we can choose g_{12} as any multiaffine factor of $\Delta_{12}(f)$ of degree two. There are two possibilities, namely $g_{12} = (x_3 - \mathbb{i})(x_4 - \mathbb{i})$ or $g_{12} = (x_3 - \mathbb{i})(x_4 + \mathbb{i})$ and one can check that either choice works. We will start with the first option and compute

$$\varphi_3(g_{12}) = -\mathbb{i}(x_1 + \mathbb{i})(x_2 + \mathbb{i})(x_4 - \mathbb{i})(x_4 + \mathbb{i}).$$

To choose g_{13} we compute $\gcd(\Delta_{13}(f), \varphi_3(g_{12})) = -\mathbb{i}(x_2 + \mathbb{i})(x_4 - \mathbb{i})(x_4 + \mathbb{i})$. Thus, up to a constant, we have two choices for g_{13}, namely $-\mathbb{i}(x_2 + \mathbb{i})(x_4 - \mathbb{i})$ or $-\mathbb{i}(x_2 + \mathbb{i})(x_4 + \mathbb{i})$. We will choose the first option, giving

$$g_{23} = \overline{(\varphi_3(g_{12})/g_{13})} = (x_1 + \mathbb{i})(x_4 + \mathbb{i}).$$

To find g_{14}, we compute the $\gcd(\Delta_{14}(f), \varphi_4(g_{12}), \varphi_4(g_{13})) = -\mathbb{i}(x_2 + \mathbb{i})(x_3 + \mathbb{i})$ and we get g_{24} and g_{34} similarly. The final matrix is $M =$

$$\begin{pmatrix} -x_2x_3x_4 - x_2 - x_3 - x_4 & (x_3 - \mathbb{i})(x_4 - \mathbb{i}) & -\mathbb{i}(x_2 + \mathbb{i})(x_4 - \mathbb{i}) & -\mathbb{i}(x_2 + \mathbb{i})(x_3 + \mathbb{i}) \\ (x_3 + \mathbb{i})(x_4 + \mathbb{i}) & x_1x_3x_4 - x_1 - x_2 - x_3 & (x_1 + \mathbb{i})(x_4 - \mathbb{i}) & (x_1 - \mathbb{i})(x_3 + \mathbb{i}) \\ \mathbb{i}(x_2 - \mathbb{i})(x_4 + \mathbb{i}) & (x_1 - \mathbb{i})(x_4 + \mathbb{i}) & x_1x_2x_4 - x_1 - x_2 - x_4 & \mathbb{i}(x_1 - \mathbb{i})(x_2 - \mathbb{i}) \\ \mathbb{i}(x_2 - \mathbb{i})(x_3 - \mathbb{i}) & (x_1 + \mathbb{i})(x_3 - \mathbb{i}) & -\mathbb{i}(x_1 + \mathbb{i})(x_2 + \mathbb{i}) & x_1x_2x_3 - x_1 - x_2 - x_3 \end{pmatrix}.$$

Now we compute the adjugate matrix of M and divide its entries by f^2 we get

$$A = \frac{1}{f^2}M^{\mathrm{adj}} = \begin{pmatrix} x_1 & -1 & \mathbb{i} & \mathbb{i} \\ -1 & x_2 & -1 & -1 \\ -\mathbb{i} & -1 & x_3 & -\mathbb{i} \\ -\mathbb{i} & -1 & \mathbb{i} & x_4 \end{pmatrix}$$

and one can check that $\det(A) = f$. The algorithm gives all the possible representations of f, up to diagonal equivalence, namely

$$\left\{ \begin{pmatrix} x_1 & -1 & \mathbb{i} & \mathbb{i} \\ -1 & x_2 & -1 & -1 \\ -\mathbb{i} & -1 & x_3 & -\mathbb{i} \\ -\mathbb{i} & -1 & \mathbb{i} & x_4 \end{pmatrix}, \begin{pmatrix} x_1 & -1 & \mathbb{i} & \mathbb{i} \\ -1 & x_2 & -1 & -1 \\ -\mathbb{i} & -1 & x_3 & \mathbb{i} \\ -\mathbb{i} & -1 & -\mathbb{i} & x_4 \end{pmatrix}, \begin{pmatrix} x_1 & -1 & \mathbb{i} & -\mathbb{i} \\ -1 & x_2 & -1 & -1 \\ -\mathbb{i} & -1 & x_3 & -\mathbb{i} \\ \mathbb{i} & -1 & \mathbb{i} & x_4 \end{pmatrix} \right\}.$$

3.5 Hermitian determinantal representations

Let $\mathbb{K}$ be a field with an automorphism $a \mapsto \overline{a}$ of order two. Let $\mathbb{F}$ be the fixed field of this automorphism. We call a matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ Hermitian if $A = \overline{A}^T$.

3.5.1 Consequences of Algorithm 1

Theorem 3.5.1. *Let $f \in \mathbb{F}[x_1, \ldots, x_m]$ be a polynomial of total degree $n \leq m$ that is multiaffine in $x_1, \ldots, x_n$ and coefficient of $x_1 \cdots x_n$ equals to one. There exist Hermitian matrices $A_{n+1}, \ldots, A_m, A_0$ so that*

$$f = \det\left(\mathrm{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_j A_j + A_0\right)$$

if and only if for all $i, j \in [n]$, $\Delta_{ij}(f)$ is a Hermitian square in $\mathbb{K}[x_1, \ldots, x_m]$.

Lemma 3.5.2. *Let $\mathbb{F}$ be an infinite field and $f \in \mathbb{F}[x_1, \ldots, x_m]$ be irreducible, multiaffine in the variables $x_1, \ldots, x_n$ and have coefficient of $x_1 \cdots x_n$ equals to 1. Let $R = \mathbb{F}[x_{n+1}, \ldots, x_m]$. For a generic element $\gamma \in \mathrm{SL}_2(\mathbb{F})^n$, the derivatives $\frac{\partial}{\partial x_j}(\gamma \cdot f)$ are irreducible in $R[x_1, \ldots, x_n]$ for $j = 1, \ldots, n$, up to a constant and the coefficient of $\prod_{i=1}^{n} x_i$ is nonzero.*

Proof. Consider $\gamma_j = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{F})$ acting on x_j. Then $\partial_j(\gamma \cdot f) = af_j + cf^j$ where $f = f_j x_j + f^j$. Consider the set

$$\mathcal{X} = \{(a, c) \in \mathbb{F}^2 : af_j + cf^j \text{ is reducible in } R[x_1, \ldots, x_n] \text{ up to constants}\},$$

Suppose that the multidegree of f in $x_1, \ldots, x_m$ is given by $\mathbf{d} \in \mathbb{N}^m$. By assumption, $\mathbf{d}_i = 1$ for $i = 1, \ldots, n$. Note that $\mathcal{X}$ is contained in the union $\bigcup_{\mathbf{e}} \mathcal{X}_{\mathbf{e}}$ where

$$\mathcal{X}_{\mathbf{e}} = \{(a, c) \in \mathbb{F}^2 : af_j + cf^j \in \mathbb{F}^{\mathrm{alg}}[x_1, \ldots, x_m]_{\leq \mathbf{e}} \cdot \mathbb{F}^{\mathrm{alg}}[x_1, \ldots, x_m]_{\leq \mathbf{d}-\mathbf{e}}\}$$

and the union is taken over all vectors $\mathbf{e} \in \mathbb{N}^m$ that are coordinate-wise $\leq \mathbf{d}$ with the property that $\mathbf{e}_i = 1$ and $\mathbf{e}_k = 0$ for some $i, k \in [n] \backslash \{j\}$. Here $\mathbb{F}^{\mathrm{alg}}$ denotes the algebraic closure of $\mathbb{F}$. To see this, suppose $(a, c) \in \mathcal{X}$, meaning $af_j + cf^j = g \cdot h$ where $g, h \in R[x_1, \ldots, x_n]$ are not constants (i.e. elements of R). In particular, for some $i, k \in [n]\backslash\{j\}$, $\deg_i(g) > 0$ and $\deg_k(h) > 0$. Since $af_j + cf^j$ has degree at most one in each of x_i and x_k, it follows that $\deg_i(g) = \deg_k(h) = 1$ and $\deg_k(g) = \deg_i(h) = 0$. Taking $\mathbf{e} \in \mathbb{N}^m$ to be the multidegree of g gives $g \in \mathbb{F}[x_1, \ldots, x_m]_{\leq \mathbf{e}}$ and $h \in \mathbb{F}[x_1, \ldots, x_m]_{\leq \mathbf{d}-\mathbf{e}}$. Then $(a, c) \in \mathcal{X}_{\mathbf{e}}$.

By the projective elimination theorem, the image $\mathbb{F}^{\text{alg}}[x_1, \ldots, x_m]_{\leq \mathbf{e}} \cdot \mathbb{F}^{\text{alg}}[x_1, \ldots, x_m]_{\leq \mathbf{d}-\mathbf{e}}$ is Zariski-closed in the vectorspace $\mathbb{F}^{\text{alg}}[x_1, \ldots, x_m]_{\leq \mathbf{d}}$. Intersecting with the $\mathbb{F}$-subspace spanned by $\{f_j, f^j\}$ shows that $\mathcal{X}_{\mathbf{e}}$ and hence $\cup_{\mathbf{e}} \mathcal{X}_{\mathbf{e}}$ is Zariski-closed in $(\mathbb{F})^2$. Therefore this union is either all of $\mathbb{F}^2$ or is an algebraic set of codimension ≥ 1. Suppose, for the sake of contradiction, that it is all of $\mathbb{F}^2$. Then there exists some $\mathbf{e}$ for which $\mathcal{X}_{\mathbf{e}} = \mathbb{F}^2$. By assumption, there are $i, k \in [n]$ so that $\mathbf{e}_i = 1$ and $(\mathbf{d} - \mathbf{e})_k = 1$. By Proposition 3.2.2, it follows that for all $a, c \in \mathbb{F}$, $\Delta_{ik}(af_j + cf^j)$ is identically zero. For $c = 1$, this corresponds to the evaluation of $\Delta_{ik}(f)$ at $x_j = a$. It follows that the polynomial $\Delta_{ik}(f)$ is identically zero (see e.g. [4]). Proposition 3.2.2 then implies that f factors as the product of two nonconstant elements in $R[x_1, \ldots, x_n]$. $\square$

Proof of Theorem 3.5.1. First assume that $\mathbb{F}$ is an infinite field and that f is irreducible in $\mathbb{F}[x_1, \ldots, x_m]$. Let $R = \mathbb{F}[x_{n+1}, \ldots, x_m]$. By Lemma 3.5.2, there exists a generic $\gamma \in \text{SL}_2(\mathbb{F})^n$ such that the partial derivatives of $\gamma \cdot f$ are all irreducible in $R[x_1, \ldots, x_n]$ and the coefficient of $x_1 \cdots x_n$ in $\gamma \cdot f$ is nonzero. Then by Corollary 3.4.7, there exists $\{g_{ij}\}_{1 \leq i \neq j \leq n}$ with g_{ij} in $\mathbb{K}[x_\ell : \ell \neq i, j]$ satisfying $\Delta_{ij}(\gamma \cdot f) = g_{ij} g_{ji}$, $g_{ji} = \overline{g_{ij}}$, and $\text{res}_{x_k}(g_{ij}, \gamma \cdot f) = g_{ik} g_{kj}$ for all distinct $i, j, k \in [n]$. Acting by γ^{-1} on $\gamma \cdot f$ and using Proposition 3.2.3 we get $h_{ij} = \gamma^{-1} \cdot g_{ij}$ such that $\Delta_{ij}(f) = h_{ij} \overline{h_{ij}}$ and $\text{res}_{x_k}(h_{ij}, f) = h_{ik} h_{kj}$ and thus using Theorem 3.3.1 we get a determinantal representation of f over $\mathbb{K}$ and since $h_{ji} = \overline{h_{ij}}$, the matrix will be Hermitian.

Now suppose that f is reducible. Let g be an irreducible factor of f with $\deg_i(g) = 1$ for $i \in I \subseteq [n]$ and $\deg_j(g) = 0$ for $j \in [n]\backslash I$ where the coefficient of $\prod_{i \in I} x_i$ equals one. By Lemma 3.3.3, for every $i, j \in I$, $\Delta_{ij}(g)$ is a Hermitian square. Therefore g has a determinantal representation of the correct form, $g = \det(\text{diag}(x_i : i \in I) + \sum_{j=n+1}^m x_j A_{ij} + A_{i0})$. Taking a block diagonal representation of these representations (and permuting the rows and columns to reorder $x_1, \ldots, x_n$) gives a determinantal representation for f.

Now suppose that $\mathbb{F}$ is a finite field. Consider the transcendental extension of $\mathbb{F}$ to $\mathbb{F}(t)$ and of $\mathbb{K}$ to $\mathbb{K}(t)$. Then by the arguments above, $f = \det(\operatorname{diag} x_1, \ldots, x_n + A(t))$ for some Hermitian matrix $A(t) \in \mathrm{Mat}_n(\mathbb{K}(t))$. The (i,j)th entry of $A(t)$ can be written as $a_{ij} = \frac{p_{ij}}{q_{ij}}$ where $p_{ij}, q_{ij} \in \mathbb{K}[t]$ are relatively prime and the polynomial q_{ij} is nonzero. Specializing to $t = 0$ will give a determinantal representation of f over $\mathbb{K}$. To do this, we need to check that $q_{ij}(0)$ is nonzero for all i, j. If $a_{ij} = 0$, then we can take $p_{ij} = 0$ and $q_{ij} = 1$. Suppose that for some $i, j \in [n]$, p_{ij} is nonzero and $q_{ij}(0) = 0$. Then t divides q_{ij} and so also divides $\overline{q_{ij}}$. Notice that

$$a_{ii} = \operatorname{coeff}(f, \prod_{k \neq i} x_k) \quad \text{and} \quad a_{ij}a_{ji} = a_{ii}a_{jj} - \operatorname{coeff}\left(f, \prod_{k \neq i,j} x_k\right)$$

are both in $\mathbb{F}$ and hence $p_{ij}\overline{p_{ij}} = r q_{ij}\overline{q_{ij}}$ for some $r \in \mathbb{F}^*$. We get the desired contradiction by noticing that t^2 divides the left-hand side of the equation, while it does not divide the right-hand side since p_{ij} and q_{ij} are relatively prime. Therefore we can specialize both sides of the equation $f = \det(\operatorname{diag} x_1, \ldots, x_n + A(t))$ to $t = 0$, which gives a Hermitian determinantal representation of f. □

Example 3.5.3. Consider the polynomial $f = x_1x_2x_3 + x_1 + x_2 + x_3 + 1$ over the field $\mathbb{F} = \mathbb{F}_2$. The Rayleigh difference $\Delta_{12}(f) = x_3^2 + x_3 + 1$ does not factor in $\mathbb{F}_2[x_3]$, showing that the coefficient vector of f is not in the image of $\mathrm{Mat}_3(\mathbb{F}_2)$ under the principal minor map.

Consider the quadratic extension $\mathbb{K} = \mathbb{F}_2[\alpha]/\langle \alpha^2 + \alpha + 1 \rangle$. The map $\alpha \mapsto 1 + \alpha$ extends to an automorphic involution on $\mathbb{K}$ that fixes $\mathbb{F}_2$. Over $\mathbb{K}$, the Rayleigh differences factor into multiaffine polynomials, namely $\Delta_{ij}(f) = (x_k + \alpha)(x_k + 1 + \alpha)$, for distinct i, j, k. As then guaranteed by Theorem 3.5.1, f has a Hermitian determinantal representation over $\mathbb{K}$:

$$f = \det \begin{pmatrix} x_1 & 1+\alpha & 1+\alpha \\ \alpha & x_2 & 1+\alpha \\ \alpha & \alpha & x_3 \end{pmatrix}.$$

Corollary 3.5.4. *Let* $f \in \mathbb{R}[x_1, \dots, x_m]$ *be a polynomial of total degree* $n \leq m$ *that is multiaffine in* $x_1, \dots, x_n$ *and coefficient of* $x_1 \cdots x_n$ *equals to one. There exist Hermitian matrices* $A_{n+1}, \dots, A_m, A_0$ *so that*

$$f = \det\left(\operatorname{diag}(x_1, \dots, x_n) + \sum_{j=n+1}^{m} x_j A_j + A_0\right)$$

if and only if for all $i, j \in [n]$, $\Delta_{ij}(f)$ *factors as* $g_{ij}\overline{g_{ij}}$ *for* $g_{ij} \in \mathbb{C}[x_1, \dots, x_m]$.

This provides a partial converse to [48, Corollary 4.3], which states that if some power of a polynomial f has a definite determinantal representation, then for all i, j, the Rayleigh difference $\Delta_{ij}(f)$ is a sum of squares. In particular, Hermitian representations of f give real symmetric determinantal representations of f^2. We might hope for the following.

Conjecture 3.5.5. *If* $f \in \mathbb{R}[x_1, \dots, x_m]$ *is multiaffine in* $x_1, \dots, x_n$ *and coefficient of* $x_1 \cdots x_n$ *is nonzero, then some power of* f *has a definite real symmetric determinantal representation if and only if for all* i, j, $\Delta_{ij}(f)$ *is a sum of squares in* $\mathbb{R}[x_1, \dots, x_m]$.

3.5.2 Other multiaffine determinantal representations

In this section we restrict ourselves to fields and consider the set of multiaffine determinantal polynomials of the form

$$f(\mathbf{x}) = \lambda \det\left(V \operatorname{diag}(x_1, \dots, x_m) V^* + W\right) = \lambda \det\left(\sum_{i=1}^{m} x_i v_i v_i^* + W\right) \tag{3.5}$$

for some $\lambda \in \mathbb{F}$, some matrix $V = (v_1, \dots, v_m) \in \mathbb{K}^{n \times m}$ and some $n \times n$ Hermitian matrix W. Note that when we take V to be the $n \times n$ identity matrix and $\lambda = 1$, this is the principal minor polynomial f_W. When $n < m$, the coefficient of $x_1 \cdots x_m$ in f is necessarily zero.

Theorem 3.5.6. *A polynomial* $f \in \mathbb{F}[\mathbf{x}]_{\mathrm{MA}}$ *has a determinantal representation* (3.5) *if and only if for all* $i, j \in [n]$, $\Delta_{ij} f$ *is Hermitian square in* $\mathbb{K}[\mathbf{x}]$. *Moreover, one can always take a representation of size* $n = \deg(f)$ *in* (3.5).

Proof. ($\Rightarrow$) Without loss of generality, we show that $\Delta_{12}(f)$ is a Hermitian square. First suppose v_1 and v_2 are linearly dependent, i.e. let $v_1 = \alpha v_2$ for some $\alpha \in \mathbb{K}$. Then $v_1 v_1^* = \alpha\overline{\alpha} v_2 v_2^*$ and $f(x_1, \ldots, x_m) = f(0, \alpha\overline{\alpha}x_1 + x_2, x_3, \ldots, x_m)$. Taking partial derivatives shows that $\frac{\partial f}{\partial x_1} = \alpha\overline{\alpha}\frac{\partial f}{\partial x_2}$ and that $\frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$. Then $\Delta_{12}(f) = (\alpha\frac{\partial f}{\partial x_2})(\overline{\alpha}\frac{\partial f}{\partial x_2})$ is a Hermitian square.

If v_1 and v_2 are linearly independent, then there is an invertible matrix $U \in \mathbb{K}^{n\times n}$ with $Uv_1 = e_1$ and $Uv_2 = e_2$. Then

$$|\det(U)|^2 f = \lambda \det\left(U\left(\sum_{i=1}^m x_i v_i v_i^* + W\right)U^*\right) = \lambda \det\left(\operatorname{diag}(x_1, x_2, \mathbf{0}) + \sum_{i=3}^m x_i \widetilde{v_i}\widetilde{v_i}^* + \widetilde{W}\right).$$

where $\widetilde{v_i} = Uv_i$ and $\widetilde{W} = UWU^*$. These matrices are still Hermitian and so by equation (3.2), $\Delta_{12}(f)$ is Hermitian square.

($\Leftarrow$) Let $d = \deg(f)$. We can assume, without loss of generality, that the coefficient of $x_1 \cdots x_d$ in f is nonzero. Moreover since the set of polynomials of the form (3.5) is invariant under scaling, we can assume that this coefficient equals one. By Theorem 3.5.1, there are Hermitian matrices $A_0, A_{d+1}, \ldots, A_m$ so that

$$f = \det\left(\operatorname{diag}(x_1, \ldots, x_d) + \sum_{j=d+1}^m x_j A_j + A_0\right).$$

We take $W = A_0$. By Lemma 3.5.7 below, for every $k = n+1, \ldots, m$, the rank of A_k equals the degree of f in x_k, which is one. It remains to show that the matrix A_k has the form $v_k v_k^*$ for some $v_k \in \mathbb{K}^d$.

By homogenizing and specializing variables to zero, it suffices to consider polynomials of the form $f = \det(\operatorname{diag}(x_1, \ldots, x_d) + x_{d+1}A)$ where $A \in \mathbb{K}^{n\times n}$ is Hermitian and rank-one. Then $f = \prod_{i=1}^d x_i + \sum_{j=1}^d A_{jj} \prod_{i\in[d]\backslash\{j\}} x_i$, where A_{jj} is the jth entry of A. Then for $j = 1, \ldots, d$,

$$\Delta_{j(d+1)}(f) = f_j^{d+1} f_{d+1}^j - f_{j(d+1)} f^{j(d+1)} = \left(\prod_{i=1}^d x_i\right)\left(A_{jj} \prod_{i\in[d]\backslash\{j\}} x_i\right) = A_{jj}\left(\prod_{i\in[d]\backslash\{j\}} x_i\right)^2.$$

By assumption, $\Delta_{j(d+1)}(f)$ is a Hermitian square, and so we see that $A_{jj} = \alpha_j\overline{\alpha_j}$ for some $\alpha_j \in \mathbb{K}$. Since A has rank one, we can write it as λuu^* for some $\lambda \in \mathbb{F}^*$ and $u \in \mathbb{K}^n$. If

$u_j \neq 0$, then $\lambda u_j \overline{u_j} = \alpha_j \overline{\alpha_j}$, meaning that $\lambda = \beta\overline{\beta}$ for $\beta = \alpha_j / u_j$. It follows that $A = vv^*$ for $v = \beta u$. ☐

Lemma 3.5.7. *If* $f = \det(\operatorname{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_j A_j + A_0)$ *where* $A_j \in \mathbb{K}^{n\times n}$ *are Hermitian. Then the rank of* A_j *equals the degree* f *in the variable* x_j.

Proof. The bound $\deg_j(f) \leq \operatorname{rank}(A_j)$ follows from the Laplace expansion of the determinant. To see equality, it suffices to take $j = m = n+1$ and $A_0 = 0$. Let f_A be the polynomial $f_A = \det(\operatorname{diag}(x_1, \ldots, x_n) + A)$ where $A \in \mathbb{K}^{n\times n}$ is Hermitian. Then $f = \sum_{S\subseteq[n]} A_S \mathbf{x}^{[n]\setminus S} y^{|S|}$ equals the homogenization of f_A. From this we see that the degree of f in the variable y equals the size of the largest nonzero *principal* minor of A. By the so-called Principal Minor Theorem [45, Strong PMT 2.9], this coincides with the size of the largest nonzero minor of A, i.e. $\operatorname{rank}(A)$. Therefore for a general polynomial $f = \det(\operatorname{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} x_j A_j + x_0 A_0)$, the restriction to $x_k = 0$ for $k \in \{n+1, \ldots, m\}\setminus\{j\}$ and $x_0 = 0$ has degree $\operatorname{rank}(A_j)$ in x_j, showing that $\deg_j(f) \geq \operatorname{rank}(A_j)$ ☐

This immediately gives the invariance of the set of determinantal polynomials.

Corollary 3.5.8. *The set of polynomials in* $\mathbb{F}[\mathbf{x}]_{\mathrm{MA}}$ *with a determinantal representation* (3.5) *is invariant under the action of* $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$.

Proof. By Corollary 3.2.4, for any $\gamma \in \mathrm{SL}_2(\mathbb{F})^n$, $\Delta_{ij}(\gamma \cdot f) = \gamma \cdot \Delta_{ij}(f)$. If $\Delta_{ij}(f)$ is a Hermitian square $g\overline{g}$ with $g \in \mathbb{K}[\mathbf{x}]$ then so is $\Delta_{ij}(\gamma \cdot f) = (\gamma \cdot g)\overline{(\gamma \cdot g)}$. ☐

3.6 Determinantal Stable Polynomials

In this section we consider polynomials over $\mathbb{R}$ and $\mathbb{C}$ and show that any real stable multiaffine polynomial with a complex linear determinantal representation has a definite Hermitian determinantal representation (Theorem 3.6.4). Moreover, if the original polynomial is irreducible, then the matrix is diagonally similar to a Hermitian one (Theorem 3.6.6).

We build up to the proofs of these statements with a series of useful lemmas.

Lemma 3.6.1. *Let $f \in \mathbb{R}[x_1, \ldots, x_m]$ be multiaffine in the variables $x_1, \ldots, x_n$ for some $n \leq m$ with coefficient of $x_1 \cdots x_n$ equals to one. If f is irreducible, then for a generic element $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$, $\partial^S(\gamma \cdot f)$ is irreducible for every $S \subset [n]$.*

Proof. For each $S \subset [n]$, the set of $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$ for which $\partial^S(\gamma \cdot f)$ is irreducible is Zariski-open. Therefore it suffices to show that this set is nonempty for each $S \subset [n]$. Then the intersection of these nonempty, Zariski-open sets will be nonempty and Zariski open.

We will proceed by induction on $|S|$. For $|S| = 0$, this is immediate, so suppose that $|S| \geq 1$ and let $i \in S$. Note that $\partial^S(f) = \partial_i\left(\partial^{S\backslash\{i\}} f\right)$. By induction, for generic $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$, $\partial^{S\backslash\{i\}}(\gamma \cdot f)$ is irreducible. Moreover, its coefficient of $\prod_{j \in ([n]\backslash S) \cup \{i\}} x_j$ is nonzero. Therefore, up to a scalar multiple, $\partial^{S\backslash\{i\}}(\gamma \cdot f)$ satisfies the hypothesis of Lemma 3.5.2, and hence for generic $\widetilde{\gamma} \in \mathrm{SL}_2(\mathbb{R})$ acting on the ith coordinate,

$$\partial_i\left(\widetilde{\gamma} \cdot \partial^{S\backslash\{i\}}(\gamma \cdot f)\right) = \partial^S\left(\widetilde{\gamma} \cdot \gamma \cdot f\right)$$

is irreducible. Here we use that $\widetilde{\gamma}$ commutes with the differential operator $\partial^{S\backslash\{i\}}$, since $\widetilde{\gamma}$ acts as the identity in the coordinates indexed by elements of $S\backslash\{i\}$. It follows that for a generic element $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$, $\partial^S\left(\gamma \cdot f\right)$ is irreducible. $\square$

Lemma 3.6.2. *If $g = ax_1^2 + bx_1 + c$ is nonnegative on $\mathbb{R}^m$ where $a, b, c \in \mathbb{R}[x_2, \ldots, x_m]$, then the polynomial a is nonnegative on $\mathbb{R}^{m-1}$.*

Proof. Fix $\mathbf{p} \in \mathbb{R}^{m-1}$ and consider the specialization $g(x_1, \mathbf{p}) = a(\mathbf{p})x_1^2 + b(\mathbf{p})x_1 + c(\mathbf{p})$ in $\mathbb{R}[x_1]$. Since g is globally nonnegative on $\mathbb{R}^m$, $g(x_1, \mathbf{p})$ is nonnegative on $\mathbb{R}$ and so its leading coefficient $a(\mathbf{p})$ must be nonnegative. $\square$

Lemma 3.6.3. *Suppose $g, h \in \mathbb{C}[x_1, \ldots, x_m]$ are multiaffine in $x_1, \ldots, x_n$ and $\partial^{[n]}g$ and $\partial^{[n]}h$ are nonzero polynomials in $x_{n+1}, \ldots, x_m$ of total degree at most one. If the product $g \cdot h$ has real coefficients and is nonnegative as a function on $\mathbb{R}^m$, then h is a positive scalar multiple of $\overline{g}$, i.e. $h = \lambda \overline{g}$ for some $\lambda \in \mathbb{R}_{>0}$.*

Proof. ($n = 0$) Let $g = a + \mathbb{i}b$ and $h = c + \mathbb{i}d$ for some $a, b, c, d \in \mathbb{R}[x_1, \ldots, x_m]$. Since $g \cdot h \in \mathbb{R}[x_1, \ldots, x_m]$, we see that $ad = -bc$. Note that if $b = 0$, then $d = 0$ and so both g

and h are real. In order for $g \cdot h$ to be nonnegative on $\mathbb{R}^n$, we must have $h = \lambda \cdot g$ for some $\lambda \in \mathbb{R}_{>0}$. The case $d = 0$ follows similarly.

Otherwise, since g and h are linear and thus irreducible, either $a = \lambda b$ and $c = -\lambda d$ or $a = \lambda c$ and $b = -\lambda d$ for some nonzero $\lambda \in \mathbb{R}$. In the first case, $g = (\lambda + \mathbb{i})b$ and $h = (-\lambda + \mathbb{i})d = (\lambda - \mathbb{i})(-d)$ and thus $g \cdot h = (\lambda^2 + 1)(-b \cdot d) \geq 0$ on $\mathbb{R}^n$. Thus $-d = \mu b$ for some $\mu \in \mathbb{R}_{>0}$. It follows that $h = (\lambda - \mathbb{i})(\mu b) = \mu \overline{g}$. The second case gives $g = \lambda \overline{h}$. Since $g \cdot h = \lambda h \cdot \overline{g}$ is nonnegative on $\mathbb{R}^n_{\geq 0}$, we conclude $\lambda > 0$, as desired.

($n \geq 1$) Now suppose $n \geq 1$ and write $g = g_n x_n + g^n$ and $h = h_n x_n + h^n$. Since $g \cdot h$ is real and nonnegative, so is its coefficient of x_n^2, $g_n \cdot h_n$. In particular, g_n, h_n satisfy the hypothesis of the theorem and so by induction, $h_n = \lambda \overline{g_n}$ for some $\lambda \in \mathbb{R}_{>0}$. Moreover, for every $\mathbf{a} \in \mathbb{R}^{m-1}$ with $g_n(\mathbf{a}) \neq 0$, the roots (in x_n) of the specialization of $g \cdot h$ at $\mathbf{x} = \mathbf{a}$ come in complex conjugate pairs. It follows that $-h^n/h_n = -\overline{g^n}/\overline{g_n}$ as rational functions in $\mathbb{C}(x_k : k \neq n)$. Together with $h_n = \lambda \overline{g_n}$, this gives that $h = \lambda \overline{g}$. Moreover, since $g \cdot h = \lambda \cdot g \cdot \overline{g}$ is nonnegative on $\mathbb{R}^n$, we see that $\lambda > 0$. □

Theorem 3.6.4. *Let $f \in \mathbb{R}[x_1, \ldots, x_m]$ be stable and complex determinantal, i.e.*

$$f = \det\left(\operatorname{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} A_j x_j + A_0\right)$$

for some $n \times n$ complex matrices A_j. Then there exists Hermitian matrices $B_0, B_{n+1}, \ldots, B_m$ for which $f = \det\left(\operatorname{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} B_j x_j + B_0\right)$.

Proof. First suppose f is irreducible. By Lemma 3.6.1, there is $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$, such that $\partial^S(\gamma \cdot f)$ is irreducible for all $S \subset [n]$. By Corollary 3.3.5, we can replace f by $\gamma \cdot f$, and thereby assume that all the coefficients of $\prod_{k \in [n] \setminus \{i,j\}} x_k^2$ in the polynomials $\Delta_{ij}(\gamma \cdot f)$ are non-zero. To see this, notice that by induction on n, we can prove that

$$\operatorname{coeff}\left(\Delta_{ij}(f), \prod_{k \in [n] \setminus \{i,j\}} x_k^2\right) = \Delta_{i,j}\left(\partial^{[n] \setminus \{i,j\}}(f)\right).$$

If this coefficient is zero, then Lemma 3.3.3 implies that $\partial^{[n] \setminus \{i,j\}}(f)$ is reducible.

Let $i < j \in [n]$. Since f is determinantal, by Theorem 3.3.1, the polynomial $\Delta_{ij}(f)$ factors as $g_{ij} \cdot g_{ji}$ where g_{ij}, g_{ji} are multiaffine in $\{x_k : k \in [n]\backslash\{i,j\}\}$ and has total degree $\leq n-1$. In particular, the coefficient of $\prod_{k\in[n]\backslash\{i,j\}} x_k$ in both g_{ij} and g_{ji} has degree ≤ 1 in $x_{n+1}, \ldots, x_m$. By the arguments above we can assume this coefficient is nonzero. Since f is real stable, $\Delta_{ij}(f)$ is also globally nonnegative on $\mathbb{R}^n$ [13]. Therefore by Lemma 3.6.3, $g_{ji} = \lambda\overline{g_{ij}}$ for some g_{ij}. It follows that $\Delta_{ij}(f)$ factors as a Hermitian square $h_{ij} \cdot \overline{h_{ij}}$ where $h_{ij} = \sqrt{\lambda}g_{ij}$. Theorem 3.5.1 then gives the desired Hermitian determinantal representation.

Now suppose f is reducible, say $f = f_1 \cdots f_r$ where each factor f_k is irreducible and multiaffine in the variables x_i for $i \in I_k \subset [n]$. Each factor is stable. Moreover, by Lemma 3.3.3, $\Delta_{ij}(f_k)$ is either zero or factors as a product of two polynomials that are multiaffine in $\{x_\ell : \ell \in I_k\}$ and with total degree $\leq |I_k| - 1$. Since f_k is irreducible, the arguments above show that for every $i, j \in I_k$, $\Delta_{ij}(f_k)$ is a Hermitian square, from which it follows that $\Delta_{ij}(f) = \Delta_{ij}(f_k) \cdot \prod_{\ell\neq k} f_\ell^2$ is a Hermitian square. Theorem 3.5.1 then gives the desired Hermitian determinantal representation. $\square$

Remark 3.6.5. Theorem 3.6.4 cannot hold for arbitrary real stable polynomials. For example, consider f to be the basis generating polynomial of the Vámos matriod, defined in [15]. It was shown by Wagner and Wei [73] that f is stable. By the theory of matrix factorizations, some power f^r of f has a complex linear determinantal representation (see [69, §3.3]). This power is necessarily stable, but as shown by Brändén [15], f^r does not have a definite Hermitian determinantal representation.

When f is reducible, one can easily construct determinantal representations of f that are not Hermitian by taking block upper triangular representations. For example, x_1x_2 equals $\det \begin{pmatrix} x_1 & 1 \\ 0 & x_2 \end{pmatrix}$. However, when f is irreducible and real stable, we see that all complex linear determinantal representations are Hermitian, up to conjugation by diagonal matrices.

Theorem 3.6.6. *Let $f \in \mathbb{R}[x_1, \ldots, x_m]$ be stable, irreducible, and complex determinantal,* *i.e.*

$$f = \det\left(\operatorname{diag}(x_1, \ldots, x_n) + \sum_{j=n+1}^{m} A_j x_j + A_0\right)$$

for some $n \times n$ complex matrices A_j. Then there exists a real diagonal matrix $D \in \mathbb{R}^{n\times n}$ such that $D^{-1}A_jD$ is Hermitian for all j.

Proof. By Lemma 3.6.1, there is $\gamma \in \mathrm{SL}_2(\mathbb{R})^n$, such that $\partial^S(\gamma \cdot f)$ is irreducible for all $S \subset [n]$. By Corollary 3.3.5, we can replace f by $\gamma \cdot f$, and thereby assume that all the coefficients of $\prod_{k\in[n]\setminus\{i,j\}} x_k^2$ in the polynomials $\Delta_{ij}(\gamma \cdot f)$ are non-zero, as in the proof of Theorem 3.6.4.

Let $A(x) = \sum_{k=n+1}^{m} A_k x_k + A_0$ and let $a_{ij} \in \mathbb{C}[x_{n+1}, \ldots, x_m]$ denote the (i, j)th entry of $A(x)$. Then the coefficient of $\prod_{k\in[n]\setminus\{i,j\}} x_k^2$ in $\Delta_{ij} f$ is $a_{ij}a_{ji}$. Since f is stable, the polynomial $\Delta_{ij}(f)$ is nonnegative on $\mathbb{R}^m$. Then by Lemma 3.6.2, it follows that the coefficient $a_{ij}a_{ji}$ of $\prod_{k\in[n]\setminus\{i,j\}} x_k^2$ in $\Delta_{ij}(f)$ is nonnegative on $\mathbb{R}^{n-m}$. By Lemma 3.6.3, we can conclude that for each $1 \le i < j \le n$, there is some $\lambda_{ij} \in \mathbb{R}_{>0}$ such that $a_{ij} = \lambda_{ij}\overline{a_{ji}}$.

We claim that the scalars λ_{ij} satisfy $\lambda_{ij} = \lambda_{ik}\lambda_{kj}$ for all $1 \le i < k < j \le n$. For simplicity, we show this for $i = 1$, $k = 2$, $j = 3$ and the proof in general is identical. By the arguments above, the starting determinantal representation of f has the form

$$\operatorname{diag}(x_1, \ldots, x_n) + A(x) = \begin{pmatrix} x_1 + a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \lambda_{12}\overline{a_{12}} & x_2 + a_{22} & a_{23} & \cdots & \\ \lambda_{13}\overline{a_{13}} & \lambda_{23}\overline{a_{23}} & x_3 + a_{33} & & \\ \vdots & \vdots & & \ddots & \\ \lambda_{1n}\overline{a_{1n}} & & & & x_n + a_{nn} \end{pmatrix}.$$

Recall that by Dodgson condensation, the polynomial $\Delta_{ij}(f)$ factors as $\det(M[i, j])\cdot\det(M[j, i])$ where $M[i, j]$ is the matrix obtained from $M = \operatorname{diag}(x_1, \ldots, x_n) + A(x)$ by removing the ith row and jth column. These polynomials are affine in x_k for $k \in [n]\setminus\{i, j\}$. In particular,

$$g := \partial^{[n]\setminus\{1,2,3\}} \det(M[3, 1]) = a_{12}a_{23} - a_{13}(x_2 + a_{22}), \text{ and}$$
$$h := \partial^{[n]\setminus\{1,2,3\}} \det(M[1, 3]) = \lambda_{12}\lambda_{23}\overline{a_{12}a_{23}} - \lambda_{13}\overline{a_{13}}(x_2 + a_{22}).$$

These polynomials satisfy the hypotheses of Lemma 3.6.3, and so there is some $\mu \in \mathbb{R}_{>0}$ for which $h = \mu\overline{g}$. Since a_{ij} is nonzero for all i, j and a_{22} is invariant under conjugation, we see that $\lambda_{12}\lambda_{23} = \mu = \lambda_{13}$. More generally $\lambda_{ij} = \lambda_{ik}\lambda_{kj}$ for any $i < k < j$.

Now define $D = \operatorname{diag}(1, \sqrt{\lambda_{12}}, \ldots, \sqrt{\lambda_{1n}})$. For $i < j$, $\lambda_{1j} = \lambda_{1i}\lambda_{ij}$ we calculate the (i, j)th and (j, i)th entries of $D^{-1}A(x)D$ as

$$(D^{-1}A(x)D)_{ij} = \frac{\sqrt{\lambda_{1j}}}{\sqrt{\lambda_{1i}}} a_{ij} = \sqrt{\lambda_{ij}} a_{ij} \quad \text{and} \quad (D^{-1}A(x)D)_{ji} = \frac{\sqrt{\lambda_{1i}}}{\sqrt{\lambda_{1j}}} \lambda_{ij}\overline{a_{ij}} = \sqrt{\lambda_{ij}}\overline{a_{ij}}.$$

□

3.7 *Defining the set of factoring multiquadratic polynomials and the image of the principal minor map*

In this section we give a complete characterization of the image of the principal minor map of Hermitian matrices using the characterization of Hermitian multiaffine determinantal polynomials from Section 3.5 and the characterization of multiquadratic polynomials that are Hermitian squares. This set is invariant under the action of $\mathrm{SL}_2(R)^n \rtimes S_n$ and we derive the defining equations and numerical conditions as the orbit of a finite set under the action of this group, where R is a unique factorization domain. In this section, we will restrict to rings and fields of characteristic $\neq 2$.

Lemma 3.7.1. *Let $g = ax^2 + bx + c \in R[x]$. The polynomial g factors in to two linear factors in $R[x]$ if and only if its discriminant $\mathrm{Discr}_x(g)$ is a square in R.*

Proof. $(\Rightarrow)$ If g factors, then it has a root in the fraction field of R. By the quadratic formula, this implies that the discriminant is a square in $\mathrm{frac}(R)$, and hence in R.

$(\Leftarrow)$ Suppose that $b^2 - 4ac = q^2$ for some $q \in R$. We can rewrite this as $(b-q)(b+q) = 4ac$. Since R is a unique factorization domain, there is some choice of factorization of $a = a_1a_2$ and $c = c_1c_2$ so that $b - q = 2a_1c_1$ and $b + q = 2a_2c_2$. If $a = 0$, then g factors as $1 \cdot (bx + c)$, so we can assume $a \neq 0$. We can then write g as

$$g = a\left(x - \frac{-b+q}{2a}\right)\left(x - \frac{-b-q}{2a}\right) = a_1a_2\left(x + \frac{c_1}{a_2}\right)\left(x + \frac{c_2}{a_1}\right) = (a_2x + c_1)(a_1x + c_2).$$

□

This lemma does not hold over rings of characteristic two. See [21, Section 2.4, Exercise 6] for further discussion. Note that for $g \in R[x,y]_{\mathrm{MQ}}$, $\mathrm{Discr}_x(g)$ is a polynomial of degree 4 in y whose coefficients are quadratic in the coefficients of g.

Lemma 3.7.2. *Let* $h(x) = \sum_{i=0}^{4} b_i x^i \in R[x]_4$ *a univariate quartic. Then* h *is a square in* $R[x]$ *if and only if* b_0, b_4 *and* $h(1) = \sum_j b_j$ *are squares in* R *and the point* $(b_0, b_1, b_2, b_3, b_4)$ *satisfies*

$$b_4 b_1^2 - b_3^2 b_0 = 0, \qquad b_3^3 - 4b_4 b_3 b_2 + 8b_4^2 b_1 = 0, \qquad b_1^3 - 4b_0 b_1 b_2 + 8b_0^2 b_3 = 0 \tag{3.6}$$

$$b_2 b_3^2 - 4b_2^2 b_4 + 2b_1 b_3 b_4 + 16 b_0 b_4^2 = 0, \quad \textit{and} \quad b_1^2 b_2 - 4b_0 b_2^2 + 2b_0 b_1 b_3 + 16 b_0^2 b_4 = 0.$$

Proof. ($\Rightarrow$) If $h(x)$ is a square in $R[x]$, then $h(x) = \sum_{i=0}^{4} b_i x^i = (\alpha x^2 + \beta x + \delta)^2$ for some $\alpha, \beta, \delta \in R$. We see that $b_4 = \alpha^2$, $b_0 = \delta^2$, and $\sum_{i=0}^{4} b_i = (\alpha + \beta + \delta)^2$ are all squares in R. Each of the coefficients b_i is a polynomial in α, β, δ and one can quickly check that all the cubics in (3.6) vanish identically on this parametrization.

($\Leftarrow$) Let $b_4 = \alpha^2$, $b_0 = \delta^2$, and $\sum_j b_j = \lambda^2$ for some α, δ, $\lambda \in R$. From $b_0 b_3^2 = b_1^2 b_4$, we see that $\delta b_3 = \pm b_1 \alpha$, and replacing α with $-\alpha$ if necessary, we can take $\delta b_3 = b_1 \alpha$.

If b_3 is nonzero, we see from the second equation that b_4, and hence α, must also be nonzero. Define $\beta = b_3/(2\alpha) \in \mathrm{frac}(R)$. It follows immediately that $b_0 = \delta^2$, $b_1 = 2\delta\beta$, $b_3 = 2\beta\alpha$, and $b_4 = \alpha^2$. If $b_3 \neq 0$, the second equation implies that

$$b_2 = \frac{1}{4b_3 b_4}(b_3^3 + 8b_1 b_4^2) = \frac{1}{8\beta\alpha^3}(8\beta^3\alpha^3 + 16\delta\beta\alpha^4) = \beta^2 + 2\delta\alpha,$$

from which we conclude that $(\alpha x^2 + \beta x + \delta)^2 = h(x)$. Similarly, if b_1 is nonzero then so are b_0 and δ. We can define $\beta = b_1/(2\delta)$ and use $4b_0 b_1 b_2 = b_1^3 + 8b_0^2 b_3$ to conclude that $(\alpha x^2 + \beta x + \delta)^2 = h(x)$. In either case, evaluating at $x = 1$ gives that $\alpha + \beta + \delta = \pm\lambda$, and $\beta = \pm\lambda - \alpha - \delta \in R$.

If $b_1 = b_3 = 0$, the equations simplify to $4b_4(b_2^2 - 4b_0 b_2) = 0$ and $4b_0(b_2^2 - 4b_0 b_2) = 0$. If b_0 or b_4 is nonzero, then $b_2 = \pm\delta\alpha$ and $h(x)$ is $(\alpha x^2 \pm \delta)^2$. Otherwise $b_0 = b_1 = b_3 = b_4 = 0$, in which case $b_2 = \lambda^2$ and $h(x) = (\lambda x)^2$. □

Corollary 3.7.3. *A quartic* $h(x) = \sum_{j=0}^{4} b_j x^j$ *is a square in* $R[x]$ *if and only if for all* γ *in* $\mathrm{SL}_2(\{0, \pm 1\})$*,* $(\gamma \cdot h)_{x=0}$ *is a square in* R *and* $B_x(\gamma \cdot h) = C_x(\gamma \cdot h) = D_x(\gamma \cdot h) = 0$*, where*

$$B_x(h) = b_4 b_1^2 - b_3^2 b_0, \;\; C_x(h) = b_1^3 - 4b_0 b_1 b_2 + 8b_0^2 b_3, \;\; \text{and } D_x(h) = b_1^2 b_2 - 4b_0 b_2^2 + 2b_0 b_1 b_3 + 16 b_0^2 b_4.$$

Proof. It suffices to show that we can recover the conditions in Lemma 3.7.2, which we can do this with three elements of $\mathrm{SL}_2(\{0, \pm 1\})$: the identity, $\gamma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\gamma_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, representing the fractional linear transformations $x \mapsto -1/x$ and $x \mapsto x + 1$, respectively. Note that $(\gamma_1 \cdot h)(0) = b_4$ and $(\gamma_2 \cdot h)(0) = h(1) = \sum_j b_j$, so from (i), we recover that all of these are squares in R. The element γ_1 induces the transposition $b_k \mapsto (-1)^k b_{4-k}$ for each k. One can quickly check that we recover the two missing cubics from this action of γ_1. □

Remark 3.7.4. The ideal generated by the five cubics in Lemma 3.7.2 is not saturated with respect to the ideal $\langle b_0, \ldots, b_4 \rangle$. Its saturation is minimally generated by these five cubics together with $-b_1 b_3^2 + 4b_1 b_2 b_4 - 8b_0 b_3 b_4$ and $-b_1^2 b_3 + 4b_0 b_2 b_3 - 8b_0 b_1 b_4$.

Note that the coefficients of $\mathrm{Discr}_y(\gamma \cdot g)$ have degree two in the coefficients c_α of g, and so the polynomials listed in (ii) above have degree six. For example,

$$\begin{aligned} B_x(\mathrm{Discr}_y(\gamma \cdot g)) = {} & 4(c_{01}c_{11} - 2(c_{10}c_{02} + c_{00}c_{12}))^2 (c_{21}^2 - 4c_{20}c_{22}) \\ & - 4(c_{01}^2 - 4c_{00}c_{02})(c_{11}c_{21} - 2(c_{20}c_{12} + c_{10}c_{22}))^2. \end{aligned}$$

Theorem 3.7.5. *Let* R *be a unique factorization domain with* $\mathrm{char}(R) \neq 2$ *and* $|R| \geq 13$*. A polynomial* $g = \sum_{\alpha \in \{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha \in R[\mathbf{x}]$ *is the product of multiaffine polynomials if and only if for all* $\gamma \in \mathrm{SL}_2(R)^n \rtimes S_n$*,*

(i) $\mathrm{Discr}_{x_1}(\gamma \cdot g)|_{x_2 = \ldots = x_n = 0} = \gamma \cdot (c_{\mathbf{10}}^2 - 4c_{\mathbf{00}}c_{\mathbf{20}})$ *is a square in* R*,*

(ii) the sextic polynomials in $\mathbf{c}$ *given by specializing* $B_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$*,* $C_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ *and* $D_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ *to* $x_3 = \ldots = x_n = 0$ *are all zero.*

Proof. We can express $g = \sum_{\beta \in \{0,1,2\}^2} g_\beta x_1^{\beta_1} x_2^{\beta_2}$ where $g_\beta \in R[x_3, \dots, x_n]_{\leq \mathbf{2}}$. The polynomial $B_{x_2}(\mathrm{Discr}_{x_1}(g))$ has degree six in the coefficients g_β and so degree ≤ 12 in each variable x_j.

Consider $I \subset R$ with $|I| = 13$. For $\lambda_1 = \lambda_2 = 0$ and $\lambda_3, \dots, \lambda_n \in I$, consider the element $\gamma = \left(\left(\begin{smallmatrix} 1 & \lambda_j \\ 0 & 1 \end{smallmatrix}\right)\right)_j$ in $\mathrm{SL}_2(R)^n$. For any polynomial $F \in R[\mathbf{x}]$ the evaluation of $\gamma \cdot F$ at $\mathbf{x} = 0$ equals the evaluation of F at $\mathbf{x} = (\lambda_1, \dots, \lambda_n)$. In particular, (ii) implies that the polynomials $B_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$, $C_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ and $D_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ vanish at the point $\mathbf{x} = (\lambda_1, \dots, \lambda_n)$ for every choice of $\lambda_j \in I$. Since these polynomials have degree ≤ 12 in each variable x_j and $|I| \geq 13$, it follows that each of these polynomials is identically zero, using [**?**, Lemma 4.1].

We can now proceed by induction on n. The $n = 1$ case is the content of Lemma 3.7.1 together with the observation that the discriminant is invariant under the action of $\mathrm{SL}_2(R)$, so we suppose $n \geq 2$. Let $h = \mathrm{Discr}_{x_1}(g) \in S[x_2]$ where $S = R[x_3, \dots, x_n]$. By induction, for every $\gamma \in \mathrm{SL}_2(R)$ acting on the variable x_2, $(\gamma \cdot g)|_{x_2=0}$ factors into multiaffine polynomials and so $(\gamma \cdot h)|_{x_2=0} = \mathrm{Discr}_{x_1}((\gamma \cdot g)|_{x_2=0})$ is a square in $S = R[x_3, \dots, x_n]$.

By Corollary 3.7.3, it follows that $\mathrm{Discr}_{x_1}(g)$ is a square in $S[x_2]$. Then by Lemma 3.7.1, g factors into linear factors in x_1 in the ring $S[x_1, x_2] = R[\mathbf{x}]$. Using the action of S_n, we see that every irreducible factor of g must have degree ≤ 1 in each variable. $\square$

Remark 3.7.6. For every choice of $i \neq j \in [n]$ and $\lambda \in I^{n-3}$ we obtain three equations by evaluating $B_{x_j}(\mathrm{Discr}_{x_i}(g))$, $C_{x_j}(\mathrm{Discr}_{x_i}(g))$ and $D_{x_j}(\mathrm{Discr}_{x_i}(g))$ at the point λ, along with additional two polynomials from the two missing analogous polynomials in Lemma 3.7.2, which can be recovered from the SL_2-action on x_j. This gives a total of $5n(n-1)13^{n-3}$ sextic equations in the coefficients of g.

Lemma 3.7.7. *Let S be a unique factorization domain with $\mathrm{char}(S) \neq 2$ and an automorphic involution $a \to \overline{a}$ and let R be the fixed ring under this involution. The polynomial $g = ax^2 + bx + c \in R[x]$ is a Hermitian square in $S[x]$ if and only if a and c are Hermitian squares in S and the discriminant $\mathrm{Discr}_x(g) = q^2$ with $q \in S[x]$ and $\overline{q} = -q$.*

Proof. ($\Rightarrow$) If g factors into two conjugates $(sx+t)(\overline{s}x+\overline{t})$, then $a=s\overline{s}$ and $c=t\overline{t}$ and

$$\mathrm{Disc}_x(g)=b^2-4ac=(s\overline{t}+t\overline{s})^2-4s\overline{s}t\overline{t}=(s\overline{t}-t\overline{s})^2$$

which satisfies the desired property.

($\Leftarrow$) Assume that $b^2-4ac=q^2$ such that $\overline{q}=-q$. If $a=0$, then $b=\pm q$ and thus $\overline{b}=-b$. Since $b\in R$, then $b=0$ and $g=c$ is a Hermitian square as desired. If $a\neq 0$, then $(b-q)(b+q)=(b-q)(b-\overline{q})=4ac=4s\overline{s}t\overline{t}$, where $a=s\overline{s}$ and $c=t\overline{t}$. Thus, after relabeling if needed, we may assume that $b-\overline{q}=2st$. Thus, we can write g as

$$g=a\left(x-\frac{-b+\overline{q}}{2a}\right)\left(x-\frac{-b+q}{2a}\right)=s\overline{s}\left(x+\frac{t}{\overline{s}}\right)\left(x+\frac{\overline{t}}{s}\right)=(sx+t)(\overline{s}x+\overline{t}).$$

□

Theorem 3.7.8. *Let S be a unique factorization domain with* $\mathrm{char}(S)\neq 2$ *and an automorphic involution $a\mapsto\overline{a}$. Let R be the fixed ring of this automorphism with $|R|\geq 13$. The polynomial $g=\sum_{\alpha\in\{0,1,2\}^n}c_\alpha\mathbf{x}^\alpha$ in $R[\mathbf{x}]_{\mathrm{MQ}}$ is a Hermitian square if and only if $(\gamma\cdot g)|_{x_3=\ldots=x_n=0}$ is a Hermitian square in $S[x_1,x_2]$ for all $\gamma\in\mathrm{SL}_2(R)^n\rtimes S_n$.*

Proof. If for all $\gamma\in\mathrm{SL}_2(R)^n\rtimes S_n$, the polynomial $(\gamma\cdot g)|_{x_3=\ldots=x_n=0}$ is a Hermitian square in $S[x_1,x_2]$, then by Lemma 3.7.1, $\mathrm{Discr}_{x_1}(\gamma\cdot g)|_{x_3=\ldots=x_n=0}$ is a square in $S[x_2]$. Using Corollary 3.7.3 we see that the two conditions of Theorem 3.7.5 are satisfied and hence we deduce that g is a product of multiaffine polynomials in $S[\mathbf{x}]$. To prove that g is a Hermitian square, we will proceed by induction on n. The case $n=2$ is trivially satisfied. For the inductive step, write g as $g=p_2x_1^2+p_1x_1+p_0$ for some $p_2,p_1,p_0\in\tilde{R}=R[x_2,\ldots,x_n]$. By induction we see that p_2 and p_0 are both Hermitian squares and as g is a product of multiaffine polynomials, then by Lemma 3.7.1, we see that $\mathrm{Disc}_{x_1}(g)=p_1^2-4p_2p_0=q^2$ for some $q\in S[x_2,\ldots,x_n]$. Since $p_1^2-4p_2p_0\in\tilde{R}$, then $q^2\in\tilde{R}$ and so $q=-\overline{q}$ or $q=\overline{q}$. In the former case, Lemma 3.7.7 implies that g is a Hermitian square and we are done. Otherwise we get $(\gamma\cdot q)|_{\mathbf{x=0}}=\overline{(\gamma\cdot q)|_{\mathbf{x=0}}}$ for all $\gamma\in\mathrm{SL}_2(R)^{n-1}$. Notice that by induction on the other hand, $(\gamma\cdot g)|_{\mathbf{x=0}}$ is a Hermitian square and hence

$$\mathrm{Disc}_{x_1}((\gamma\cdot g)_{\mathbf{x=0}})=\big(\gamma\cdot(p_1^2-4p_2p_0)\big)_{\mathbf{x=0}}=(\gamma\cdot q)^2_{\mathbf{x=0}}\text{ with }(\gamma\cdot q)|_{\mathbf{x=0}}=-\overline{(\gamma\cdot q)|_{\mathbf{x=0}}}.$$

Thus we conclude that $(\gamma \cdot q)|_{\mathbf{x}=\mathbf{0}} = 0$ for all $\gamma \in \mathrm{SL}_2(R)^{n-1}$. Consider $\gamma = (\gamma_i)_{2 \leq i \leq n}$ where $\gamma_i = \begin{pmatrix} 1 & \lambda_i \\ 0 & 1 \end{pmatrix}$ for $\lambda_i \in R$. Notice that $\gamma \cdot q|_{(x_2 = \cdots = x_n = 0)} = q|_{(x_2 = \lambda_2, \ldots, x_n = \lambda_n)} = 0$. Since $|R| \geq 3$, [**?**, Lemma 4.1] implies that $q \equiv 0$ and thus $q = -\overline{q}$ and we apply Lemma 3.7.7 again to deduce that g is a Hermitian square. □

Let $\mathbb{F}$ be a field of $\mathrm{char}(\mathbb{F}) \neq 2$ with $|\mathbb{F}| \geq 13$ and $\mathbb{K}$ be a degree two extension field. Let δ denote the square root of the discriminant of the minimal polynomial of this field extension. Then $\mathbb{K} = \mathbb{F}(\delta)$ and the involution $\delta \longrightarrow \overline{\delta} = -\delta$ extends to an automorphism of $\mathbb{K}$ with fixed field $\mathbb{F}$.

Remark 3.7.9. In the field $\mathbb{K}$, $\overline{q} = -q$ is equivalent to requiring $q = \delta r$ for some $r \in \mathbb{F}$.

Lemma 3.7.10. *Let* $g = \sum_{\alpha \in \{0,1,2\}^2} c_\alpha \mathbf{x}^\alpha \in \mathbb{F}[x_1, x_2]_{MQ}$. *The polynomial* g *is a Hermitian square in* $\mathbb{K}[x_1, x_2]$ *if and only if for all* $\gamma \in \mathrm{SL}_2(\{0, \pm 1\})^2 \rtimes S_2$*:*

(i) $\gamma \cdot c_{(0,0)}$ *is a Hermitian square in* $\mathbb{K}$*.*

(ii) $\frac{1}{\delta^2}\mathrm{Discr}_{x_1}(\gamma \cdot g)$ *is a square in* $\mathbb{F}[x_2]$*.*

Proof. Write g as $g = p_2 x_1^2 + p_1 x_1 + p_0$ where p_2, p_1 and p_0 are quadratics in $\mathbb{F}[x_2]$. Using Lemma 3.7.7, we see that g is a product of two conjugate factors if and only if p_2 and p_0 are product of two conjugates in $\mathbb{K}[x_2]$ and $\mathrm{Disc}_{x_1} g = q^2$ where $\overline{q} = -q$ for some $q \in \mathbb{K}[x_2]$. Notice that by Remark 3.7.9, this condition is equivalent to $q = \delta r$ where $r \in \mathbb{F}[x_2]$ and thus requiring that $\frac{1}{\delta^2}\mathrm{Disc}_{x_1} g$ is a square in $\mathbb{F}[x_2]$. Using Lemma 3.7.7, p_2 and p_0 are conjugates if and only if $c_{(i,j)}$ is a product of two conjugates for $i, j \in \{0, 2\}$ and $\frac{1}{\delta^2}\mathrm{Discr}_{x_2}(\gamma \cdot g)|_{x_1 = 0}$ is a square for $\gamma \in \mathrm{SL}_2(\mathbb{F})$ and this gives the desired equivalence. □

Theorem 3.7.11. *A polynomial* $g = \sum_{\alpha \in \{0,1,2\}^n} c_\alpha \mathbf{x}^\alpha \in \mathbb{F}[\mathbf{x}]$ *is a Hermitian square in* $\mathbb{K}[\mathbf{x}]$ *if and only if for all* $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$*,*

(i) $(\gamma \cdot c_{\mathbf{0}})$ *is a Hermitian square in* $\mathbb{K}$*,*

(ii) $\frac{1}{\delta^2}\mathrm{Discr}_{x_1}(\gamma \cdot g)|_{x_2 = \ldots = x_n = 0} = \gamma \cdot \left(\frac{1}{\delta^2}(c_{\mathbf{10}}^2 - 4 c_{\mathbf{00}} c_{\mathbf{20}})\right)$ *is a square in* $\mathbb{F}$*,*

(iii) the sextic polynomials in $\mathbf{c}$ *given by specializing* $B_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$, $C_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ *and* $D_{x_2}(\mathrm{Discr}_{x_1}(\gamma \cdot g))$ *to* $x_3 = \ldots = x_n = 0$ *are all zero.*

Proof. Using Lemma 3.7.8, g is a Hermitian square in $\mathbb{K}[\mathbf{x}]$ if and only if for all $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$, $(\gamma \cdot g)|_{x_3=\ldots=x_n=0}$ is a Hermitian square in $\mathbb{K}[x_1, x_2]$. Now Lemma 3.7.10, shows that this is equivalent to $\gamma \cdot c_{\mathbf{0}}$ is a product of two conjugates and $\frac{1}{\delta^2}\mathrm{Discr}_{x_1}(\gamma \cdot g)$ is a square in $\mathbb{F}[x_2]$, which is equivalent to conditions (ii) and (iii) above using Corollary 3.7.3. □

Now we are ready to give a complete characterization of the image of the principal minor map of Hermitian matrices using the characterization of Hermitian multiaffine determinantal polynomials from Section 3.5 and the characterization of multiquadratic polynomials that are Hermitian squares.

Recall that to each element $\mathbf{a} = (a_S)_{S \subseteq [n]}$ in $\mathbb{F}^{2^n}$ we associate the multiaffine polynomial

$$f_{\mathbf{a}} = \sum_{S \subseteq [n]} a_S \mathbf{x}^{[n] \setminus S}.$$

For $n = 3$, the discriminant of the Rayleigh difference $\Delta_{12}(f)$ with respect to x_3 is Cayley's $2 \times 2 \times 2$ hyperdeterminant

$$\begin{aligned}\mathrm{HypDet}(\mathbf{a}) &= (a_1 a_{23} + a_2 a_{13} - a_3 a_{12} - a_\emptyset a_{123})^2 - 4(a_1 a_2 - a_\emptyset a_{12})(a_{13} a_{23} - a_3 a_{123}) \\ &= a_\emptyset^2 a_{123}^2 + a_1^2 a_{23}^2 + a_2^2 a_{13}^2 + a_3^2 a_{12}^2 - 2a_\emptyset a_1 a_{23} a_{123} - 2a_\emptyset a_2 a_{13} a_{123} - 2a_\emptyset a_3 a_{12} a_{123} \\ &\quad - 2a_1 a_2 a_{13} a_{23} - 2a_1 a_3 a_{12} a_{23} - 2a_2 a_3 a_{12} a_{13} + 4a_\emptyset a_{23} a_{13} a_{12} + 4a_{123} a_1 a_2 a_3.\end{aligned}$$

This quartic polynomial therefore appears in the arithmetic conditions on the image of the principal minor map.

Theorem 3.7.12. *Let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{F}^{2^n}$ *with* $a_\emptyset = 1$. *There exists a Hermitian matrix over* $\mathbb{K}$ *with principal minors* $\mathbf{a}$ *if and only if for every* $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$*:*

(i) $\gamma \cdot (a_1 a_2 - a_\emptyset a_{12})$ *is a Hermitian square in* $\mathbb{K}$*,*

(ii) $\frac{1}{\delta^2}\,\mathrm{HypDet}(\gamma \cdot \mathbf{a})$ *is a square in* $\mathbb{F}$*, and*

(iii) $\gamma \cdot \mathbf{a}$ *satisfies the degree-12 polynomials given by specializing* $B_{x_4}(\mathrm{Discr}_{x_3}(\gamma \cdot \Delta_{12} f_{\mathbf{a}}))$, $C_{x_4}(\mathrm{Discr}_{x_3}(\gamma \cdot \Delta_{12} f_{\mathbf{a}}))$ *and* $D_{x_4}(\mathrm{Discr}_{x_3}(\gamma \cdot \Delta_{12} f_{\mathbf{a}}))$ *to* $x_5 = \ldots = x_n = 0$.

Here the operators B_x, C_x, D_x are defined in Corollary 3.7.3.

Proof of Theorem 3.7.12. By Theorem 3.5.1 with $n = m$, $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{F}^{2^n}$ is in the image of the principal minor map if and only if $\Delta_{ij}(f_{\mathbf{a}})$ is a Hermitian square for all $i, j \in [n]$, which according to Theorem 3.7.8, is satisfied if and only if for all $\gamma \in \mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$, $\gamma \cdot \Delta_{34}(f_{\mathbf{a}})|_{x_5 = \ldots = x_n = 0}$ is a Hermitian square in $\mathbb{K}[x_1, x_2]$. This is equivalent to the three hypothesis of Theorem 3.7.11, which in turn is equivalent to the three hypotheses of the theorem. $\square$

Taking $\mathbb{K} = \mathbb{C}$ with the action complex conjugation then gives the following.

Corollary 3.7.13. *Let* $\mathbf{a} = (a_S)_{S \subseteq [n]} \in \mathbb{R}^{2^n}$ *with* $a_\emptyset = 1$. *There exists a Hermitian matrix over* $\mathbb{C}$ *with principal minors* $\mathbf{a}$ *if and only if for every* $\gamma \in \mathrm{SL}_2(\mathbb{R})^n \rtimes S_n$

(ii) $\gamma \cdot (a_1 a_2 - a_\emptyset a_{12}) \geq 0$,

(ii) $\mathrm{HypDet}(\gamma \cdot \mathbf{a}) \leq 0$, *and*

(ii) $\gamma \cdot \mathbf{a}$ *satisfies the three degree-12 equations given by restricting* $B_{x_4}(\mathrm{Discr}_{x_3}(\Delta_{12} f_{\gamma \cdot \mathbf{a}}))$, $C_{x_4}(\mathrm{Discr}_{x_3}(\Delta_{12} f_{\gamma \cdot \mathbf{a}}))$ *and* $D_{x_4}(\mathrm{Discr}_{x_3}(\Delta_{12} f_{\gamma \cdot \mathbf{a}}))$ *to* $x_5 = \ldots = x_n = 0$.

3.8 A family of counterexamples

Let $\mathbb{F}$ be a field and for $n \geq 2$, consider the multiaffine polynomial $f_{2n+1} \in \mathbb{F}[x_1, \ldots, x_{2n+1}]$ given by

$$f_{2n+1} = x_1 \cdot \prod_{j=1}^{n} (x_{2j+1} x_{2j+2} + 1) + \prod_{j=1}^{n} (x_{2j} x_{2j+1} + 1) \tag{3.7}$$

where we take $x_{2n+2} = x_2$. We show that this polynomial is not determinantal, i.e. its vector of coefficients do not belong to the image of the principal minor map, but is determinantal after specializing any one variable:

Theorem 3.8.1. *There is* no finite set of equations *whose orbit under* $\mathrm{SL}_2(\mathbb{F})^n \rtimes S_n$ *set-theoretically cuts out the image of the principal minor map for all* n.

Let $I_n \subset \mathbb{F}[a_S : S \subseteq [n]]$ be the homogeneous ideal of polynomials vanishing on the image of $n \times n$ matrices under the principal minor map in $\mathbb{P}^{2^n-1}(\mathbb{F})$. There is a natural inclusion of I_n into $\mathbb{F}[a_S : S \subseteq [n+1]]$.

Theorem 3.8.2. *The coefficient vector of the polynomial* f_{2n+1} *belongs to the variety of polynomials in the orbit* $(\mathrm{SL}_2(\mathbb{F})^{2n+1} \rtimes S_{2n+1}) \cdot I_{2n}$ *but not the variety of* I_{2n+1}.

The proof of this theorem relies on the fact that the coefficient of any generic specialization of f_{2n+1} lies in the image of the principal minor map, up to scaling. One key observation is that the Rayleigh differences of f_{2n+1} do not all factor as the product of *two* multiaffine polynomials, but do have such factorizations after specializing anyone variable. We show this explicitly by writing down the determinantal representations of these specializations.

Lemma 3.8.3. *The rational function* $\frac{1}{1+x_1} f_{2n+1}$ *can be written as* $\det(\mathrm{diag}(x_2, \ldots, x_{2n+1}) + A)$ *where for* $2 \leq i, j \leq 2n+1$,

$$A_{ij} = \begin{cases} 1/(1+x_1) & \textit{if } i \textit{ is odd, } j \textit{ is even, and } i > j, \\ -x_1/(1+x_1) & \textit{if } i \textit{ is odd, } j \textit{ is even, and } i < j, \\ -1 & \textit{if } i \textit{ is even, } j = i+1, \\ 1 & \textit{if } i \textit{ is even, } j = i-1, \\ -x_1 & \textit{if } i = 2,\ j = 2n+1, \textit{ and} \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Let D denote the determinant of the matrix $M = \det(\mathrm{diag}(x_2, \ldots, x_{2n+1}) + A)$. By definition, D is a polynomial in $\frac{1}{1+x_1}, x_1, x_2, \ldots, x_n$. Moreover the entries for which $x_1 + 1$ appears in the denominator form a square submatrix whose rows correspond to odd indices and whose columns correspond to even ones. It has the form

$$\frac{1}{1+x_1}\begin{pmatrix} 1 & -x_1 & -x_1 & \dots & -x_1 \\ 1 & 1 & -x_1 & \dots & -x_1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & 1 & \ddots & \ddots & -x_1 \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} = \frac{1}{1+x_1}J - U$$

where J is the all ones matrix and U is an upper triangular matrix with $U_{ij} = 1$ for $i < j$ and $U_{ij} = 0$ otherwise. Since J has rank one, the exponent of $1 + x_1$ appearing in the denominator of *any* minor of this matrix is at most one. This also shows that despite the many appearances of x_1 in numerator of this matrix, it does not appear in the numerator of any minor. There is only one other entry in M containing x_1, and so the determinant D can be written as $(x_1 + 1)^{-1}p_1 + p_2$ where p_1 and p_2 are multiaffine in $x_1, \dots, x_{2n+1}$. Moreover, the only term in the Laplace expansion of the determinant of M avoiding this submatrix is the product of the diagonal $\prod_{j=2}^{2n+1} x_j$. Therefore we can write D as $(x_1 + 1)^{-1}p$ where p is multiaffine in $x_1, \dots, x_{2n+1}$. Therefore to show that $p = f_{2n+1}$ it suffices to show that they have the same specialization at $x_1 = 0$ and the same coefficient of x_1.

When we specialize x_1 to zero, M becomes a block upper-triangular l matrix with diagonal blocks of the form $\begin{pmatrix} x_{2j} & -1 \\ 1 & x_{2j+1} \end{pmatrix}$. Its determinant agrees with the specialization of $\frac{1}{1+x_1}f_{2n+1}$ to $x_1 = 0$.

Consider the rational function g obtained by inverting x_1 in $\frac{1}{1+x_1}f_{2n+1}$, which is

$$\frac{x_1}{1+x_1}f_{2n+1}(x_1^{-1}, x_2, \dots, x_n) = \frac{1}{1+x_1} \cdot \left(\prod_{j=1}^{n}(x_{2j+1}x_{2j+2} + 1) + x_1 \cdot \prod_{j=1}^{n}(x_{2j}x_{2j+1} + 1) \right).$$

Let M' be the matrix obtained from M by replacing x_1 by x_1^{-1} and then multiplying the column indexed by 2 by x_1^{-1} and the row indexed by 2 by x_1. The entries are now rational functions in x_1 with only $1+x_1$ appearing in the denominator. After specializing M' to $x_1 = 0$ and cyclic shifting the rows and columns by one, we find another block upper triangular ma-

$$\begin{pmatrix} x_2 & -1 & 0 & 0 & 0 & -x_1 \\ \frac{1}{x_1+1} & x_3 & -\frac{x_1}{x_1+1} & 0 & -\frac{x_1}{x_1+1} & 0 \\ 0 & 1 & x_4 & -1 & 0 & 0 \\ \frac{1}{x_1+1} & 0 & \frac{1}{x_1+1} & x_5 & -\frac{x_1}{x_1+1} & 0 \\ 0 & 0 & 0 & 1 & x_6 & -1 \\ \frac{1}{x_1+1} & 0 & \frac{1}{x_1+1} & 0 & \frac{1}{x_1+1} & x_7 \end{pmatrix} \sim \begin{pmatrix} x_2 & 0 & 0 & -1 & 0 & -x_1 \\ 0 & x_4 & 0 & 1 & -1 & 0 \\ 0 & 0 & x_6 & 0 & 1 & -1 \\ \frac{1}{x_1+1} & -\frac{x_1}{x_1+1} & -\frac{x_1}{x_1+1} & x_3 & 0 & 0 \\ \frac{1}{x_1+1} & \frac{1}{x_1+1} & -\frac{x_1}{x_1+1} & 0 & x_5 & 0 \\ \frac{1}{x_1+1} & \frac{1}{x_1+1} & \frac{1}{x_1+1} & 0 & 0 & x_7 \end{pmatrix}$$

Figure 3.1: The matrix A in Lemma 3.8.3 for $2n + 1 = 7$.

trix with diagonal blocks of the form $\begin{pmatrix} x_{2j+1} & -1 \\ 1 & x_{2j+2} \end{pmatrix}$ for $j = 1, \ldots, n-1$ and $\begin{pmatrix} x_{2n+1} & 1 \\ -1 & x_2 \end{pmatrix}$. Therefore the determinant of M' restricted to $x_1 = 0$ is given by $\prod_{j=1}^{n}(x_{2j+1}x_{2j+2} + 1)$.

By definition, the determinant of M' equals $D(x_1^{-1}, x_2, \ldots, x_n) = \frac{x_1}{1+x_1} p(x_1^{-1}, \ldots, x_n)$. Restricting to $x_1 = 0$ gives the coefficient of x_1 in p, which must be $\prod_{j=1}^{n}(x_{2j+1}x_{2j+2} + 1)$. Therefore p agrees with the polynomial f_{2n+1}. $\square$

Lemma 3.8.4. *For every $m = 2, \ldots, 2n + 1$, the coefficients of $\frac{1}{x_m} f_{2n+1}$ are the principal minors of a $2n \times 2n$ matrix with entries in $\{0, \pm 1, x_m^{\pm 1}\}$. In particular, the rational function $\frac{1}{x_{2n+1}} f_{2n+1}$ can be written as $\det(\operatorname{diag}(x_1, \ldots, x_{2n}) + B)$ where for $1 \le i, j \le 2n$,*

$$B_{ij} = \begin{cases} 1 & \text{if } j = i+1 \text{ and } i > 1 \text{ or } (i,j) = (1,1) \text{ or } (i,j) = (2,1), \\ -1 & \text{if } i \text{ is even and } j = i-1 \text{ or } (i,j) = (2n,1), \\ x_{2n+1} & \text{if } i \text{ odd, } i \ge 3, \text{ and } j = 1, \\ 1/x_{2n+1} & \text{if } i \in \{1,2\} \text{ and } j \text{ is even, and} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $M = \det(\operatorname{diag}(x_1, \ldots, x_{2n}) + B)$ and let D denote its determinant. As in the proof of Lemma 3.8.3, the entries of M with x_{2n+1} appearing in the denominator appear in a submatrix of rank-one. The entries with x_{2n+1} appearing in the numerator are contained in the first column. Moreover, in the Laplace expansion of the determinant, the only terms avoiding the submatrix of entries x_{2n+1}^{-1} must include the $(1,1)$ and $(2,2)$ entries, and so will not involve any entries with x_{2n+1}. It follows that D can be written as $x_{2n+1}^{-1}p$ where p is multiaffine in $x_1, \ldots, x_{2n+1}$. Therefore it suffices to check that f_{2n+1} and p have the same restriction to $x_1 = 0$ and same coefficient of x_1.

We see that the coefficient of x_1 in D is the determinant of the matrix M after removing the first row and column. This minor is a block matrix with one block of the form $(x_2 + 1/x_{2n+1})$ and the rest of the form $\begin{pmatrix} x_{2j+1} & 1 \\ -1 & x_{j+2} \end{pmatrix}$. Therefore the coefficient of x_1 in p and f_{2n+1} agree.

The specialization of M to $x_1 = 0$ is a matrix has the form $\begin{pmatrix} 1 & b^T \\ c & A \end{pmatrix}$. Using Schur complements, we see that the determinant equals the determinant of $A - cb^T$. One can check that the matrix $A - cb^T$ is a block-lower triangular matrix with diagonal blocks $\begin{pmatrix} x_{2j} & 1 \\ -1 & x_{2j+1} \end{pmatrix}$ for $j = 1, \ldots, n-1$ and $x_{2n} + 1/x_{2n+1}$. This shows that the restriction of p to $x_1 = 0$ agrees with that of f_{2n+1}.

For the corresponding statement with arbitrary $m \neq 1$, we use the symmetries of f_{2n+1} under the action of a dihedral group of order n with the cyclic action $j \mapsto j+2$ (identifying $2n + j = j$ for $j \geq 2$ and reflection $n + 1 - j \leftrightarrow n + 2 - j$. There is some element of this group that moves m to $2n + 1$, and we can take the image of the representation above. □

To show that f_{2n+1} does not belong to I_{2n+1}, we will use the following:

Lemma 3.8.5. *The set of polynomials*

$$\mathcal{F}_n = \left\{ f \in \mathbb{F}[\mathbf{x}]_{\mathrm{MA}} : \text{ for all } i, j \in [n], \Delta_{ij}(f) = g_{ij} \cdot h_{ij} \text{ for some } g_{ij}, h_{ij} \in \mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MA}} \right\}$$

is Zariski closed in $\mathbb{F}[\mathbf{x}]_{\mathrm{MA}} \cong \mathbb{F}^{2^{[n]}}$, *where* $\mathbb{F}^{\mathrm{alg}}$ *denotes the algebraic closure of* $\mathbb{F}$.

Proof. The set of multiquadratic polynomials in $\mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MQ}}$ that factor as the product of two multiaffine polynomials is the image of $\mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MA}} \times \mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MA}}$ under $(g, h) \mapsto g \cdot h$.

$$\begin{pmatrix} x_1+1 & \frac{1}{x_7} & 0 & \frac{1}{x_7} & 0 & \frac{1}{x_7} \\ 1 & x_2+\frac{1}{x_7} & 1 & \frac{1}{x_7} & 0 & \frac{1}{x_7} \\ x_7 & 0 & x_3 & 1 & 0 & 1 \\ 0 & 0 & -1 & x_4 & 1 & 0 \\ x_7 & 0 & 0 & 0 & x_5 & 1 \\ -1 & 0 & 0 & 0 & -1 & x_6 \end{pmatrix} \qquad \begin{pmatrix} x_2 & 1 & 0 & 0 & 0 \\ -1 & x_3 & 0 & 0 & 0 \\ 0 & -1 & x_4 & 1 & 0 \\ -1 & 0 & -1 & x_5 & 0 \\ \frac{1}{x_7} & 0 & \frac{1}{x_7} & -1 & x_6+\frac{1}{x_7} \end{pmatrix}$$

Figure 3.2: The matrices B (left) and $A - cb^T$ (right) in Lemma 3.8.4 for $2n + 1 = 7$.

Since this map is bilinear, it follows from the projective elimination theorem that the set $\{q \in \mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MQ}} : q = g \cdot h \text{ for some } g, h \in \mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MA}}\}$ is Zariski-closed in $\mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MQ}}$.

Pulling back by the map Δ_{ij}, it follows that for each $i, j \in [n]$, the set of polynomials $f \in \mathbb{F}^{\mathrm{alg}}[\mathbf{x}]_{\mathrm{MA}}$ for which $\Delta_{ij}(f)$ factors as the product of two multiaffine polynomials is Zariski-closed, as is their intersection over all $i, j \in [n]$. It follows that its intersection with $\mathbb{F}[\mathbf{x}]_{\mathrm{MA}}$ is Zariski-closed $\mathbb{F}[\mathbf{x}]_{\mathrm{MA}}$. $\square$

Theorem 3.3.1 implies that the image of $\mathbb{F}^{n\times n}$ under the principal minor map is a subset of the variety $\mathcal{F}_n$, although as Example 3.3.2 shows, this containment can be strict. In order to show that f_{2n+1} does not belong to the variety of I_{2n+1}, it suffices to show that f_{2n+1} does not belong to $\mathcal{F}_{2n+1}$.

Recall that for $f = \sum_{S\subseteq[n]} a_S \mathbf{x}^{[n]\setminus S}$, the coefficient vector of f is defined to be

$$\mathrm{coeff}(f) = (a_S)_{S\subseteq[n]} \in \mathbb{F}^{2^{[n]}}.$$

Proof of Theorem 3.8.2. For convenience, let $f = f_{2n+1}$. Let $P \in I_{2n}$ be a homogenous polynomial vanishing on the image of $\mathbb{F}^{2n\times 2n}$ under the principal minor map. Let Q denote the image of P under inclusion into $\mathbb{F}[a_S : S \subseteq [2n+1]]$. Note that Q only involves a_S with $2n+1 \notin S$. Since our indexing of coefficients is inclusion reversing, we see that the evaluation of Q at the coefficient vector of f depends only on coefficients of monomials containing x_{2n+1}.

In particular, its evaluation at the coefficient vector of f equals the evaluation of P at the coefficient vector of derivative of f with respect to x_{2n+1}, i.e.

$$Q(\operatorname{coeff}(f)) = P(\operatorname{coeff}(\partial f/\partial x_{2n+1})). \tag{3.8}$$

If $\mathbb{F}$ is finite, it suffices to replace it with any infinite field extension, such as $\mathbb{F}(t)$ or $\mathbb{F}^{\text{alg}}$. Let $(\gamma, \pi) \in \mathrm{SL}_2(\mathbb{F})^{2n+1} \rtimes S_{2n+1}$, with γ generic, where $\mathbb{F}^{\text{alg}}$ denote the (necessarily infinite) algebraic closure of $\mathbb{F}$. We can write (γ, π) as the composition of elements $(\widehat{\gamma}, \widehat{\pi})$ in $\mathrm{SL}_2(\mathbb{F})^{2n} \rtimes S_{2n}$ and (γ_{2n+1}, σ), where $\gamma_{2n+1} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{F})$ acts on x_{2n+1} and σ is the transposition $\sigma = (m(2n+1)) \in S_{2n+1}$. Then

$$(\gamma_{2n+1}, \sigma) \cdot f = (cx_{2n+1} + d) f\left(x_1, \ldots, x_{m-1}, \frac{ax_{2n+1} + b}{cx_{2n+1} + d}, x_{m+1}, \ldots, x_{2n}, x_m\right).$$

By the genericity of γ, $c \neq 0$ and

$$\frac{\partial}{\partial x_{2n+1}}((\gamma_{2n+1}, \sigma) \cdot f) = cf\big|_{\{x_m = a/c,\, x_{2n+1} = x_m\}}.$$

Call this polynomial g. The coefficient λ of $\prod_{i=1}^{2n} x_i$ in g is $a + c$ for $m = 1$ and a for $m > 1$. In either case, we can assume it is nonzero by the genericity of γ.

By Lemma 3.8.3 for $m = 1$ and Lemma 3.8.4 for $m > 1$, the polynomial $\frac{1}{\lambda} g$ is determinantal and its coefficient vector belongs to the image of the principal minor map. Since the image of the principal minor map is invariant under the action of $(\mathrm{SL}_2(\mathbb{F})^{2n} \rtimes S_{2n})$, by (3.8),

$$0 = P(\operatorname{coeff}(g)) = P(\operatorname{coeff}((\widehat{\gamma}, \widehat{\pi}) \cdot g)) = Q(\operatorname{coeff}((\gamma, \pi) \cdot f)).$$

This shows that the coefficient vector of f belongs to the variety of $(\mathrm{SL}_2(\mathbb{F})^{2n+1} \rtimes S_{2n+1}) \cdot I_{2n}$.

On the other hand, we calculate that

$$\Delta_{12}(f) = (x_3 - x_{2n+1}) \prod_{i=3}^{2n} (x_i x_{i+1} + 1).$$

These form a cycle of length $2n - 1$ of irreducible bivariate factors, which cannot be factored as the product of two multiaffine polynomials. It follows that f does not belong to the variety $\mathcal{F}_{2n+1}$ from Lemma 3.8.5, which, by Theorem 3.3.1, contains the variety of I_{2n+1}. $\square$

The polynomial f_{2n+1} shows that the orbit of the ideal I_{2n} under $(\mathrm{SL}_2(\mathbb{F})^{2n+1} \rtimes S_{2n+1})$ is not enough to cut out the set of polynomials $f \in \mathbb{F}[x_1, \ldots, x_{2n+1}]$ all of whose Rayleigh differences factor as the product of two multiaffine polynomials. As Example 3.3.2 shows, even this is not enough to cut out the image of the principal minor map.

Chapter 4

A GENERALIZATION OF THE SPACE OF COMPLETE QUADRICS

This chapter is joint work with Mario Kummer and Miruna-Stefana Sorea appearing in Matematiche (Catania) [3].

4.1 Introduction

Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a homogeneous polynomial of degree d. We will always assume that there is no invertible linear change of coordinates T such that $h(Tx) \in \mathbb{R}[x_1, \dots, x_{k-1}]$. The *gradient map* of h is the rational map

$$\nabla h : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1},\ x \mapsto [\nabla h(x)] = [\frac{\partial}{\partial x_1} h(x) : \cdots : \frac{\partial}{\partial x_n} h(x)].$$

It is a regular map on the open subset $U \subset \mathbb{P}^{n-1}$ of all points where h does not vanish. Its graph Γ_h is the Zariski closure of all pairs $(x, \nabla h(x))$ in $\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$ with $x \in U$. In this chapter we will study resolutions of singularities of Γ_h for certain h and thereby address Question 43 in [66]:

Question 43. Can we define a generalization of the space of complete quadrics where the role of the symmetric determinant is played by an arbitrary hyperbolic polynomial h? Such a manifold could be a canonical resolution of the graph of the gradient map of h.

One motivation for this questions is the study of so-called *hyperbolic exponential families* in [52] where the gradient map plays a prominent role. While the construction we present in this article is motivated by the theory of hyperbolic polynomials, hyperbolicity plays a subordinate role in the rest of the paper. Instead, other properties such as M-convexity will matter.

In the case when $h = \det(X)$ is the determinant of the $n \times n$ generic symmetric matrix X, such a resolution of singularities is given by the *space of complete quadrics*. For any integer $0 < i < n$ and any symmetric matrix $A \in \mathbb{S}^n$ we denote by $\wedge^i A \in \mathbb{S}^{\binom{n}{i}}$ the representing matrix of the linear map $\wedge^i \mathbb{R}^n \to \wedge^i \mathbb{R}^n$ induced by A. Note that $\wedge^i A$ is nonzero if $\det(A) \neq 0$. Now the space of complete quadrics $\Omega_{\det X}$ is the Zariski closure of all tuples $([A], [\wedge^2 A], \ldots, [\wedge^{n-1} A])$ in $\mathbb{P}(\mathbb{S}^n) \times \mathbb{P}(\mathbb{S}^{\binom{n}{2}}) \times \cdots \times \mathbb{P}(\mathbb{S}^{\binom{n}{n-2}}) \times \mathbb{P}(\mathbb{S}^n)$ with A invertible. The projection of $\Omega_{\det X}$ onto the first and the last coordinate is a birational map onto $\Gamma_{\det(X)}$. Moreover it was shown for example in [49] that $\Omega_{\det X}$ is smooth.

Here we will define a variety Ω_h for an arbitrary homogeneous polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]$ together with a regular and birational map to Γ_h which agrees with the space of complete quadrics when $h = \det(X)$ is the determinant of the generic symmetric matrix. Before we give the definition of Ω_h, we recall the definition of a hyperbolic polynomial.

Definition 4.1.1. A homogeneous polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]$ is *hyperbolic* with respect to $e \in \mathbb{R}^n$ if the univariate polynomial $h(te - v) \in \mathbb{R}[t]$ has only real zeros for all $v \in \mathbb{R}^n$. The *hyperbolicity cone* of h at e is

$$\Lambda_e(h) = \{v \in \mathbb{R}^n : \ h(te - v) \text{ has only nonnegative roots}\}.$$

The prototype of a hyperbolic polynomial is the determinant of the generic symmetric matrix $\det(X)$. Indeed, since a real symmetric matrix has only real eigenvalues, the polynomial $\det(X)$ is hyperbolic with respect to the identity matrix I. The hyperbolicity cone of $\det(X)$ at I is the cone of positive semidefinite matrices.

The entries of $\wedge^{k+1} X$ cut out the variety of symmetric matrices with rank at most k. For a real symmetric matrix A the algebraic and geometric multiplicity of an eigenvalue agree. Thus the rank of A equals to the degree of the univariate polynomial $\det(tI + A)$. In fact the same holds true when we replace I by any positive definite matrix. This shows that we can express the degeneracy locus of the rational map $\mathbb{P}(\mathbb{S}^n) \dashrightarrow \mathbb{P}(\mathbb{S}^{\binom{n}{k+1}})$, $[A] \mapsto [\wedge^{k+1} A]$ in terms of the hyperbolic rank function of $\det(X)$:

Definition 4.1.2. Let $h \in \mathbb{R}[x_1, \ldots, x_n]$ be hyperbolic with respect to $e \in \mathbb{R}^n$. The *hyperbolic rank function* of h is defined as

$$\operatorname{rank}_{h,e} : \mathbb{R}^n \to \mathbb{N},\ v \mapsto \deg(h(e + tv)).$$

It was shown in [15, Lemma 4.4] that $\operatorname{rank}_{h,e} = \operatorname{rank}_{h,a}$ for any $a \in \operatorname{int}(\Lambda_e(h))$. We let $d = \deg(h)$, $0 \leq k < d$ and $v \in \mathbb{R}^n$. Then we have $\operatorname{rank}_{h,e}(v) \leq d - k - 1$ if and only if all kth order partial derivatives $\frac{\partial^k h}{\partial x_{i_1} \cdots \partial x_{i_k}}$ of h vanish in v. Lets denote by $D_1^k, \ldots, D_{m_k}^k$ a basis of the span of all kth order partial derivatives of h. We consider the rational map

$$\Delta h : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{m_1 - 1} \times \cdots \times \mathbb{P}^{m_{d-1} - 1},$$

$$[x] \mapsto ([D_1^1(x) : \cdots : D_{m_1}^1(x)], \ldots, [D_1^{d-1}(x) : \cdots : D_{m_{d-1}}^{d-1}(x)]).$$

We define the variety Ω_h to be the normalisation of the image of this rational map. The projection on the first and the last coordinate gives a birational morphism $\omega_h : \Omega_h \to \Gamma_h$. Moreover, when $h = \det(X)$ is the determinant of the generic symmetric matrix, then $\Omega_{\det(X)}$ is isomorphic to the space of complete quadrics as defined above and thus $\Omega_{\det(X)}$ is smooth in that case.

Another important example for hyperbolic polynomials are the elementary symmetric polynomials.

Theorem 4.1.3. *Let $\sigma_{d,n}$ be the elementary symmetric polynomial of degree d in n variables. Then $\Omega_{\sigma_{d,n}}$ is a smooth toric variety.*

It is well-known that $\sigma_{d,n}$ is hyperbolic with respect to every point in the positive orthant. Such polynomials are called *stable*. The theory of stable polynomials connects nicely to discrete convex analysis [54]. We denote by $\delta_k \in \mathbb{Z}^n$ the kth unit vector.

Definition 4.1.4. A nonempty set of integer points $B \subset \mathbb{Z}^n$ is called *M-convex* if for all $x, y \in B$ and every index i with $x_i > y_i$, there exists an index j with $x_j < y_j$ such that $x - \delta_i + \delta_j \in B$ and $y + \delta_i - \delta_j \in B$.

Theorem 4.1.5 (Theorem 3.2 in [13])**.** *Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a homogeneous stable polynomial. Then the support of h is M-convex.*

In Section 4.5 we will give a sufficient criterion for Ω_h being smooth when the support of h is M-convex. We will apply this criterion for proving Theorem 4.1.3. However, there are also stable (and thus hyperbolic) polynomials h for which Ω_h is not smooth.

Example 4.1.6. Consider the polynomial

$$h = w(2x + 4y + 7z)(4x + 2y + 7z)$$

$$+x^3 + 11x^2y + 11xy^2 + y^3 + 15x^2z + 46xyz + 15y^2z + 37xz^2 + 37yz^2 + 21z^3.$$

One can check that h is stable. Further, using the the computer algebra system `Macaulay2` [33], one checks that Ω_h is not smooth.

4.2 A simple polymatroid

In this section we prepare the proof of Theorem 4.1.3. Recall that a *polymatroid* on the ground set $[n] = \{1, \dots, n\}$ is a function $r : 2^{[n]} \to \mathbb{Z}_{\geq 0}$ such that for all $S, T \subset [n]$ we have:

1. $r(S) \leq r(T)$ if $S \subset T$,
2. $r(S \cup T) + r(S \cap T) \leq r(S) + r(T)$, and
3. $r(\emptyset) = 0$.

If further $r(\{i\}) \leq 1$ for all $i \in [n]$, then r is called a *matroid*. The second property is usually called *submodularity*. We call the number $d = r([n])$ the *rank* of r. See [74, Chapter 18] for a general reference on the theory of polymatroids.

Example 4.2.1. Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be hyperbolic with respect to $e \in \mathbb{R}^n$. The function that sends $S \subset [n]$ to $\operatorname{rank}_{h,e}(\sum_{i \in S} \delta_i)$ is a polymatroid [15, Proposition 3.2].

For all $0 \leq k \leq d$ the *kth truncation* r_k is the polymatroid defined by

$$r_k(S) = \min(d-k, r(S))$$

for all $S \subset [n]$. It follows directly from the definition that the sum of polymatroids is again a polymatroid. We define the following polymatroid

$$\overline{r} = r_0 + \ldots + r_d.$$

To every polymatroid r one associates the *independence polytope*

$$P(r) = \{x \in (\mathbb{R}_{\geq 0})^n : \sum_{i \in S} x_i \leq r(S) \text{ for all } S \subset [n]\}.$$

The goal of this section is to show that for every polymatroid r on $[n]$ the polytope $P(\overline{r})$ is simple. A characterization of polymatroids, whose independence polytope is simple, was given in [32, Theorem 2]. The following lemma will enable us to apply this criterion.

Definition 4.2.2. Let r be a polymatroid on $[n]$. We say that a subset $S \subset [n]$ is *r-inseparable* if for every two disjoint and nonempty subsets $S_1, S_2 \subset [n]$ with $S = S_1 \cup S_2$ we have $r(S) < r(S_1) + r(S_2)$.

Remark 4.2.3. If $|S| \leq 1$, then S is r-inseparable for every polymatroid r.

Lemma 4.2.4. *Let r, r' be polymatroids on $[n]$. If $S \subset [n]$ is r-inseparable, then S is $(r+r')$-inseparable.*

Proof. Assume that S is not $(r + r')$-inseparable. Let $\emptyset \neq S_1, S_2 \subset [n]$ such that S is the disjoint union of S_1 and S_2. If $r(S) + r'(S) \geq r(S_1) + r'(S_1) + r(S_2) + r'(S_2)$, then by submodularity of r' we get $r(S) \geq r(S_1) + r(S_2)$ which shows that S is not r-inseparable. $\square$

Remark 4.2.5. Let $|S| \geq 2$ and let $x \in [n]$ be a *loop* of r, i.e. $r(\{x\}) = 0$. If $x \in S$, then S is not r-inseparable: $r(S) = r(S \setminus \{x\}) + r(\{x\})$.

Lemma 4.2.6. *Let $S \subset [n]$ with $|S| \geq 2$ and r a polymatroid on $[n]$. Then S is $\overline{r}$-inseparable if and only if S does not contain a loop of r.*

Proof. We first observe that $x \in [n]$ is a loop of r if and only if x is a loop of all truncations r_k and thus of $\overline{r}$. Now the "only if" direction follows from Remark 4.2.5. For the "if" direction assume that S does not contain any loop of r. By Lemma 4.2.4 it suffices to show that S is r_{d-1}-inseparable. This is clear since

$$r_{d-1}(S) = 1 < 2 = r_{d-1}(S_1) + r_{d-1}(S_2)$$

for all nonempty subsets $S_1, S_2 \subset S$. □

Lemma 4.2.7. *Let r be a polymatroid on $[n]$ of rank d. Let $S, T \subset [n]$ such that*

1. $S \cap T \neq \emptyset$, $S \not\subset T$, $T \not\subset S$,
2. $\overline{r}(S \cap T) < \overline{r}(S)$, $\overline{r}(S \cap T) < \overline{r}(T)$, *and*
3. *the sets* $S, T, S \cup T$ *are* $\overline{r}$*-inseparable.*

Then $\overline{r}(S \cap T) + \overline{r}(S \cup T) < \overline{r}(S) + \overline{r}(T)$.

Proof. We proceed by induction on d. We first show that for $d \leq 1$ there are no subsets $S, T \subset [n]$ satisfying $(1), (2), (3)$. If $d = 0$, then r and $\overline{r}$ are both the zero function. Thus there are no subsets $S, T \subset [n]$ satisfying (2). If $d = 1$, we still have $r = \overline{r}$. Condition (1) implies that $|S| \geq 2$. Thus (3) and Lemma 4.2.6 imply that S contains no loop of r. Therefore, we have $r(S) = r(S \cap T) = 1$ contradicting (2).

Now let $d > 1$ and assume that the claim is true for the polymatroid r_1 of rank $d - 1$. We assume for the sake of a contradiction that $S, T \subset [n]$ satisfy $(1), (2), (3)$ but $\overline{r}(S \cap T) + \overline{r}(S \cup T) = \overline{r}(S) + \overline{r}(T)$. Again (1) implies that $|S| \geq 2$. So by (3) and Lemma 4.2.6 the set $S \cup T$ contains no loop of r. Since $d > 1$, this implies that $S \cup T$ contains no loop of r_1 as well. Thus again by Lemma 4.2.6 the sets $S, T, S \cup T$ are $\overline{r_1}$-inseparable. By submodularity and because $\overline{r} = r + \overline{r_1}$ we have

$$r(S) + r(T) = r(S \cap T) + r(S \cup T) \text{ and } \overline{r_1}(S) + \overline{r_1}(T) = \overline{r_1}(S \cap T) + \overline{r_1}(S \cup T).$$

So by induction hypothesis we have without loss of generality that $\overline{r_1}(S \cap T) = \overline{r_1}(S)$, which implies $r_1(S \cap T) = r_1(S)$, and $r(S \cap T) < r(S)$. Thus we must have $r(S) = d$ and the equation

$$d + r(T) = r(S) + r(T) = r(S \cap T) + r(S \cup T) = r(S \cap T) + d$$

implies that $r(T) = r(S \cap T)$. This in turn shows that $\overline{r}(T) = \overline{r}(S \cap T)$ contradicting (2). $\square$

Lemma 4.2.8. *Let r be a polymatroid on $[n]$ of rank d. Let $k \geq 2$ and $S_1, \ldots, S_k \subset [n]$ nonempty and pairwise disjoint. Let $S \subset [n]$ $\overline{r}$-inseparable with $\cup_{i=1}^k S_i \subset S$ and $\overline{r}(\cup_{i=1}^k S_i) = \overline{r}(S)$. Then $\overline{r}(\cup_{i=1}^k S_i) < \sum_{i=1}^k \overline{r}(S_i)$.*

Proof. We first observe that since $|S| \geq 2$ and S is $\overline{r}$-inseparable, Lemma 4.2.6 implies that S contains no loop of r. Thus each S_i also contains no loop of r.

We proceed again by induction on d. If $d = 0$, then there every element is a loop contradicting the assumptions. If $d = 1$, then we have

$$\overline{r}(\cup_{i=1}^k S_i) = 1 < 2 \leq k = \sum_{i=1}^k \overline{r}(S_i).$$

Now let $d > 1$. Then because S contains no loop of r, it also contains no loop of r_1 which shows that S is $\overline{r_1}$-inseparable. Further $\cup_{i=1}^k S_i \subset S$ and $\overline{r}(\cup_{i=1}^k S_i) = \overline{r}(S)$ imply that $\overline{r_1}(\cup_{i=1}^k S_i) = \overline{r_1}(S)$. By induction hypothesis we have $\overline{r_1}(\cup_{i=1}^k S_i) < \sum_{i=1}^k \overline{r_1}(S_i)$ which implies the claim because $r = r_0$ is submodular. $\square$

Theorem 4.2.9. *Let r be a polymatroid on $[n]$. Then the polytope $P(\overline{r})$ is simple.*

Proof. A characterization of simple independence polytopes of polymatroids was given in [32, Theorem 2]. It says that the polytope $P(\overline{r})$ is simple if and only if the conclusion of the two preceding Lemmas 4.2.7 and 4.2.8 holds. $\square$

We will be interested in the base polytope of a polymatroid rather than in its independent polytope. If $r : 2^{[n]} \to \mathbb{R}$ is a submodular function, then its *base polytope* $B(r)$ is defined as

$$B(r) = \{x \in (\mathbb{R}_{\geq 0})^n : \sum_{i \in S} x_i \leq r(S) \text{ for all } S \subset [n] \text{ and } \sum_{i=1}^n x_i = r([n])\}.$$

Corollary 4.2.10. *Let r be a polymatroid on $[n]$. Then the polytope $B(\overline{r})$ is simple.*

Proof. Clearly, the base polytope is a face of the independence polytope. Thus the claim follows from Theorem 4.2.9. $\square$

Remark 4.2.11. Taking the sum of submodular functions is compatible with taking the Minkowski sum of their base polytopes [54, Theorem 4.23(1)]. Thus if r is a polymatroid on $[n]$ of rank d. Then we have that $B(\overline{r}) = B(r_0) + \ldots + B(r_d)$.

We end this section with describing the polytope $B(\overline{r})$ explicitely when r is the rank function of a matroid. We start with the following easy lemma.

Lemma 4.2.12. *Let $r = r_{\mathcal{M}}$ be the rank function of a matroid $\mathcal{M}$ of rank d on $[n]$. There is a basis B of $\mathcal{M}$ such that for all $i \in [n]$ we have*

$$a_i := r([i]) - r([i-1]) = \begin{cases} 1 & \text{if } i \in B \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Since $\mathcal{M}$ is a matroid of rank d, we have $a_i \in \{0,1\}$ and $B = \{i \in [n] : a_i = 1\}$ has cardinality d. Let $k_1 < \cdots < k_d$ the elements of B. We show by induction on m that $I_m := \{k_1, \ldots, k_m\}$ is independent. Assume that I_{m-1} is independent. Since $r([k_m]) = m$ there is an independent subset I of $[k_m]$ of cardinality m. Thus there is an element $e \in I \backslash I_{m-1}$ such that $I_{m-1} \cup \{e\}$ is independent. Since $r([k_m - 1]) = m - 1$, we must have $e = k_m$. $\square$

Proposition 4.2.13. *Let $r = r_{\mathcal{M}}$ be the rank function of a matroid $\mathcal{M}$ of rank d on $[n]$. The vertices of the polytope $B(\overline{r})$ are exactly those points $v \in \mathbb{R}^n$ whose support is a basis of $\mathcal{M}$ and whose nonzero entries comprise the numbers $1, \ldots, d$.*

Proof. Let v be a vertex of $B(\overline{r})$. Then v is also a vertex of $P(\overline{r})$. Then by [74, §18.4, Theorem 1] there exists an integer $0 \leq k \leq n$ and a bijection $\pi : [n] \to [n]$ such that $v_{\pi(j)} = \overline{r}(\{\pi(1), \ldots, \pi(j)\}) - \overline{r}(\{\pi(1), \ldots, \pi(j-1)\})$ if $j \in [k]$ and $v_{\pi(j)} = 0$ otherwise. Since v is a vertex of $B(\overline{r})$, we can assume without loss of generality that $k = n$. Further, after relabeling, we can assume that π is the identity map. Now let $B = \{k_1, \dots, k_d\}$ with $k_1 < \cdots < k_d$ be the basis of $\mathcal{M}$ as in the preceding lemma. Then we have $v_j = d - m + 1$ if $j = k_m$ and zero if $j \notin B$. This shows that v is of the desired form.

Conversely, take $v \in \mathbb{R}^n$ whose support $\{i_1, \dots, i_d\}$ is a basis of $\mathcal{M}$ such that $v_{i_j} = d-j+1$ for $j = 1, \dots, d$. Since $v_{i_j} = \overline{r}(\{i_1, \dots, i_j\}) - \overline{r}(\{i_1, \dots, i_{j-1}\})$ for $j = 1, \dots, d$ and all other entries of v are zero, it is a vertex of $P(\overline{r})$ by [74, §18.4, Theorem 1]. One checks that $\sum_{i=1}^n v_i = \overline{r}([n])$, so v is a vertex of $B(\overline{r})$. □

Example 4.2.14. For instance when $\mathcal{M} = U(2,4)$ is the uniform matroid on 4 elements of rank 2, then $B(r_1)$ is the standard 3-simplex in $\mathbb{R}^4$ and $B(r_0)$ is the octahedron whose vertices are the permutations of $(1,1,0,0)$ (and thus is not simple). The Minkowski sum $B(\overline{r}) = B(r_0) + B(r_1)$ is simple by Corollary 4.2.10. It is the truncated tetrahedron whose vertices are the permutations of $(2,1,0,0)$.

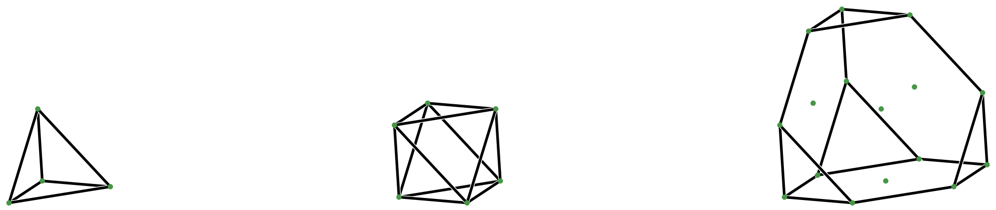

Figure 4.1: Lattice polytopes from Example 4.2.14 (left to right): $B(r_0)$, $B(r_1)$ and $B(\overline{r})$. The figures were created using `polymake` [7].

4.3 Polynomials with M-convex support

Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a homogeneous polynomial of degree d and assume that its support $\mathrm{supp}(h) \subset \mathbb{Z}^n$ is M-convex (see Definition 4.1.4). Recall that the *Newton polytope* $\mathrm{Newt}(h)$ of h is defined as the convex hull of $\mathrm{supp}(h)$ in $\mathbb{R}^n$. The statements in the following theorem are standard in the literature of polymatroids. Proofs can be found for example in [54, §4.4].

Theorem 4.3.1. *Consider the function* $\rho_h : 2^{[n]} \to \mathbb{Z}_{\geq 0}$ *defined by*

$$\rho_h(S) = \max\{\sum_{i\in S} \alpha_i : \alpha \in \operatorname{supp}(h)\}$$

for all $S \subset [n]$. *Then* ρ_h *is a polymatroid of rank* d, $\operatorname{Newt}(h) = B(\rho_h)$ *and* $\operatorname{supp}(h) = B(\rho_h) \cap \mathbb{Z}^n$.

Remark 4.3.2. If h is stable, then all coefficients of h have the same sign, see e.g. [15, Lemma 4.3]. This implies that for every $e \in (\mathbb{R}_{>0})^n$ we have that

$$\rho_h(S) = \operatorname{rank}_{h,e}(\sum_{i\in S} \delta_i)$$

as there can be no cancellation of terms.

An intriguing class of polynomials with M-convex support are Lorentzian polynomials.

Definition 4.3.3. Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a homogeneous polynomial of degree d whose support is M-convex and all of whose coefficients are nonnegative. Then h is *Lorentzian* if for every $i_1, \dots, i_{d-2} \in [n]$ the Hessian of the derivative

$$\frac{\partial^{d-2}}{\partial x_{i_1} \cdots \partial x_{i_{d-2}}} h$$

has at most one positive eigenvalue.

Remark 4.3.4. Let us clarify the relations between the different classes of polynomials that appeared so far. Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a homogeneous polynomial with nonnegative coefficients. Then we have the following implications:

$$\begin{aligned} & h \text{ is stable} \Leftrightarrow h \text{ is hyperbolic w.r.t. every } a \in (\mathbb{R}_{>0})^n \\ \Rightarrow \quad & h \text{ is Lorentzian} \\ \Rightarrow \quad & \operatorname{supp}(h) \text{ is } M\text{-convex} \end{aligned}$$

There are Lorentzian polynomials that are not stable [16, Example 2.3]. Furthermore, not every homogeneous polynomial with M-convex support and nonnegative coefficients is Lorentzian.

Theorem 4.3.5 (Theorem 3.10 in [16]). *A subset $B \subset (\mathbb{Z}_{\geq 0})^n$ is M-convex if and only if there is a Lorentzian polynomial $h \in \mathbb{R}[x_1, \dots, x_n]$ with $B = \operatorname{supp}(h)$.*

Lemma 4.3.6 (Corollary 2.11 in [16]). *Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be a Lorentzian polynomial and $e \in (\mathbb{R}_{\geq 0})^n$. The derivative*

$$D_e h = \sum_{i=1}^{n} e_i \frac{\partial h}{\partial x_i}$$

is Lorentzian as well. In particular, the support of $D_e h$ is M-convex.

Lemma 4.3.7. *If $h \in \mathbb{R}[x_1, \dots, x_n]$ is Lorentzian of degree d and $e \in (\mathbb{R}_{>0})^n$, then we have for all $0 \leq k \leq d$ that $(\rho_h)_k = \rho_{D_e^k h}$.*

Proof. It suffices to prove the claim in the case $k = 1$ because the general case follows from an iterative application of this case.

Let $S \subset [n]$. Since $\mathrm{D}_e h$ has degree $d-1$, we have $\rho_{\mathrm{D}_e h}(S) \leq d-1$. If $\rho_h(S) = d$, then there is an $\alpha \in \operatorname{supp}(h)$ such that $\sum_{i \in S} \alpha_i = d$. For any $j \in [n]$ with $\alpha_j > 0$ we have $\alpha' = \alpha - \delta_j \in \operatorname{supp}(\mathrm{D}_e h)$ and thus $\rho_{\mathrm{D}_e h}(S) \geq d-1$. Now let $\rho_h(S) < d$ and $\alpha \in \operatorname{supp}(h)$ such that $\sum_{i \in S} \alpha_i = \rho_h(S)$. Since the degree of h is d, there must be an index $j \in [n] \setminus S$ such that $\alpha_j > 0$. We have $\alpha' = \alpha - \delta_j \in \operatorname{supp}(\mathrm{D}_e h)$ and thus $\rho_{\mathrm{D}_e h}(S) \geq \rho_h(S)$. If $\beta \in \operatorname{supp}(\mathrm{D}_e h)$ satisfies $\rho_{\mathrm{D}_e h}(S) = \sum_{i \in S} \beta_i$, then there is a $j \in [n]$ such that $\beta + \delta_j \in \operatorname{supp}(h)$ so $\rho_{\mathrm{D}_e h}(S) \leq \rho_h(S)$. $\square$

The following lemma connects the polymatroid $\overline{\rho_h}$ with the variety Ω_h.

Proposition 4.3.8. *Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be homogeneous of degree d with $\operatorname{supp}(h)$ being M-convex. Consider the polymatroid $r = \rho_h$. For each $0 \leq k \leq d$ the set $B(r_k) \cap \mathbb{Z}^n$ agrees with the set B_k of all $\alpha \in \mathbb{Z}^n$ such that the monomial $\prod_{i=1}^n x_i^{\alpha_i}$ is in the support of a kth order partial derivative of h.*

Proof. Both r_k and B_k only depend on the support of h. Thus we can assume without loss of generality that h is Lorentzian by Theorem 4.3.5. Then for any $e \in (\mathbb{R}_{>0})^n$ we have that B_k is the support of $\mathrm{D}_e^k h$ because h has nonnegative coefficients. Thus B_k is M-convex by Lemma 4.3.6 and the result follows from Theorem 4.3.1 and the preceding lemma. $\square$

Remark 4.3.9. Let $h \in \mathbb{R}[x_1, \ldots, x_n]$ be homogeneous of degree d with $\mathrm{supp}(h)$ being M-convex and let B_k the set of all $\alpha \in \mathbb{Z}^n$ such that the monomial $\prod_{i=1}^{n} x_i^{\alpha_i}$ is in the support of a kth order partial derivative of h. Then it follows from the previous proposition and Corollary 4.4.2 that the Minkowski sum

$$B_1 + \ldots + B_{d-1}$$

is the set of lattice points in a smooth polytope. In general, if we drop the assumption of M-convexity, this is no longer true. Consider for example $h = a \cdot x_1 x_2^2 + b \cdot x_3^3$ with nonzero a, b. Then $B_1 + B_2$ is the set of lattice points in a simple polytope that is not smooth.

4.4 Preliminaries from algebraic and toric geometry

In this section we revisit some notions and results from algebraic geometric that will be used in the final section. Let A be a nonzero $(m+1) \times (n+1)$ matrix. Recall that a rational map $\pi : \mathbb{P}^n \dashrightarrow \mathbb{P}^m$ of the form $[x] \mapsto [Ax]$ is called a *linear projection* and that the *centre* of π is the linear subspace E of all $[x] \in \mathbb{P}^n$ such that $Ax = 0$. Clearly, the rational map π is regular on $\mathbb{P}^n \setminus E$. Thus if $X \subset \mathbb{P}^n$ is a projective variety with $X \cap E = \emptyset$, the restriction $f = \pi|_X : X \to \pi(X)$ is a morphism. This map f is *finite*, so in particular it has only finite fibers. See [63, §I.5.3] for the definition and proofs. We will use the following standard facts on finite morphisms which follow for example from [37, Lemma 14.8].

Lemma 4.4.1. *Let $f_i : X_i \to Y_i$, $i = 1, 2$, be finite morphisms of projective varieties. The product map $f_1 \times f_2$ is also finite. If $Y_1 = X_2$, then $f_2 \circ f_1$ is finite. If $Z \subset X_1$ is closed, then $f_1|_Z : Z \to f_1(Z)$ is finite.*

Given a projective variety X, a *normalisation* of X is a normal variety X^ν with a finite birational morphism $\nu : X^\nu \to X$. The normalisation is unique up to isomorphism. In particular, if $Y \to X$ is a finite birational morphism from a smooth variety Y, then Y is the normalisation of X because every smooth variety is normal. See [64, §II.5] for definitions and proofs.

We will especially consider *toric varieties*. The book [23] gives a comprehensive introduction to toric varieties and we will adopt their notation. For example, given a lattice polytope $P \subset \mathbb{R}^n$, we denote by X_P the associated toric variety [23, §2.3]. Similarly, for any finite set $A \subset \mathbb{Z}^n$ of lattice points we denote by $X_A \subset \mathbb{P}^{|A|-1}$ the image of the monomial map whose exponents are given by the elements of A. Note that in general $X_{P\cap\mathbb{Z}^n}$ is not necessarily isomorphic to X_P but when P is a smooth polytope this is the case by [23, Proposition 2.4.4]: A lattice polytope $P \subset \mathbb{R}^n$ is called *smooth* if its associated toric variety X_P is smooth [23, §2.4]. We further have:

Corollary 4.4.2. *Let r be a polymatroid on $[n]$. Then the polytope $B(\overline{r})$ is a smooth lattice polytope.*

Proof. By the Corollary to [74, §18.4, Theorem 1] the independence polytope $P(\overline{r})$ is a lattice polytope. Since $B(\overline{r})$ is a face of $P(\overline{r})$, it is a lattice polytope as well. Corollary 4.2.10 states that $B(\overline{r})$ is simple and in [18, Theorem 1.2] states that $B(\overline{r})$ is a so-called *generalized permutohedron*. Now the claim follows from [60, Corollary 3.10] which says that a simple lattice polytope, which is a generalized permutohedron, is smooth. □

4.5 A sufficient criterion for smoothness

Let $h \in \mathbb{R}[x_1, \ldots, x_n]$ be a homogeneous polynomial of degree d whose support is M-convex with nonnegative coefficients and $r = \rho_h$. Recall that we denote by $D_1^k, \ldots, D_{m_k}^k$ a basis of the span of all kth order partial derivatives of h. For all $1 \leq k < d$ consider the rational map

$$\Delta^k h : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{m_k-1},\ [x] \mapsto [D_1^k(x) : \cdots : D_{m_k}^k(x)].$$

By Proposition 4.3.8 we can decompose the map $\Delta^k h$ as $\pi_k \circ f_k$ where f_k is the monomial map associated to the polytope $B(r_k)$ (whose image is $X_{B(r_k)\cap\mathbb{Z}^n}$) and π_k the linear projection given by summing the monomials in each D_i^k.

Example 4.5.1. Let $h = x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_2x_3 + x_2^2x_3$. Then $\Delta^1 h(x)$ equals:

$$[2x_1x_2 + x_2^2 + 2x_1x_3 + x_2x_3 : x_1^2 + 2x_1x_2 + x_1x_3 + 2x_2x_3 : x_1^2 + x_1x_2 + x_2^2].$$

We further have $f_1 : \mathbb{P}^2 \dashrightarrow \mathbb{P}^4, [x_1 : x_2 : x_3] \mapsto [x_1^2 : x_1x_2 : x_1x_3 : x_2^2 : x_2x_3]$. We label the coordinates on $\mathbb{P}^4$ by z_{ij} where i and j keep track of the exponent of x_1 and x_2 respectively. The image X of f_1 is cut out by $z_{10}z_{02} - z_{11}z_{01}, z_{11}z_{10} - z_{20}z_{01}$ and $z_{11}^2 - z_{20}z_{02}$. Furthermore, the projection π_1 sends $[z_{20} : z_{11} : z_{10} : z_{02} : z_{01}]$ to

$$[2z_{11} + z_{02} + 2z_{10} + z_{01} : z_{20} + 2z_{11} + z_{10} + 2z_{01} : z_{20} + z_{11} + z_{02}].$$

The centre of π_1 is spanned by $[0 : -1 : 0 : 1 : 1]$ and $[1 : -1 : 1 : 0 : 0]$ and is disjoint from X. Thus π_1 restricts to a finite morphism $X \to \mathbb{P}^2$.

Proposition 4.5.2. *If the centre of the linear projection* π_k *is disjoint from* $X_{B(r_k)\cap\mathbb{Z}^n}$ *for each* $1 \leq k < d$, *then* Ω_h *is smooth. More precisely, it is isomorphic to the smooth toric variety* $X_{B(\overline{r_1})}$.

Proof. Let $P = B(\overline{r_1})$. By definition Ω_h is the normalisation of the image $Y \subset \prod_{i=1}^{d-1} \mathbb{P}^{m_i - 1}$ of the birational map $\Delta h(x) = (\Delta^1 h(x), \ldots, \Delta^{d-1}h(x))$. Consider the rational map $f : \mathbb{P}^{n-1} \dashrightarrow \prod_{i=1}^{d-1} \mathbb{P}^{|B(r_i)\cap\mathbb{Z}^n|-1}$ given by $f(x) = (f_1(x), \ldots, f_{d-1}(x))$ and let $\pi = \prod_{i=1}^{d-1} \pi_i$. By construction Δh factors as $\pi \circ f$. If we compose f with the Segre embedding of $\prod_{i=1}^{d-1} \mathbb{P}^{|B(r_i)\cap\mathbb{Z}^n|-1}$, we obtain the monomial map associated to the Minkowski sum $\sum_{i=1}^{d-1} B(r_i)$ which is P by Remark 4.2.11. Thus the image of f can be identified with $X_{P\cap\mathbb{Z}^n}$ and Y is the image of $X_{P\cap\mathbb{Z}^n}$ under the rational map $\pi : \prod_{i=1}^{d-1} \mathbb{P}^{|B(r_i)\cap\mathbb{Z}^n|-1} \dashrightarrow \prod_{i=1}^{d-1} \mathbb{P}^{m_i-1}$. Because P is smooth by Corollary 4.4.2 the variety $X_{P\cap\mathbb{Z}^n}$ is the smooth toric variety X_P. Letting E_i be the centre of π_i, this rational map π is regular on $U = \prod_{i=1}^{d-1}(\mathbb{P}^{|B(r_i)\cap\mathbb{Z}^n|-1} \setminus E_i)$. Since the projection of X_P on the ith factor $\mathbb{P}^{|B(r_i)\cap\mathbb{Z}^n|-1}$ is $X_{B(r_i)\cap\mathbb{Z}^n}$, which is disjoint from E_i by assumption, it follows that $X_P \subset U$. Thus restricting π gives a surjective morphism $p = \pi|_{X_P} : X_P \to Y$ which is finite since each $\pi_i|_{X_{B(r_i)\cap\mathbb{Z}^n}}$ is finite and by Lemma 4.4.1. Since $\Delta h = p \circ f$ is birational and f is birational, it follows that p is also birational. Thus X_P is the normalisation of Y. □

Example 4.5.3. Consider the polynomial

$$h = x^3 + 11x^2y + 11xy^2 + y^3 + 15x^2z + 46xyz + 15y^2z + 37xz^2 + 37yz^2 + 21z^3$$

$$+w(29x^2 + 90xy + 29y^2 + 150xz + 150yz + 137z^2).$$

One can check that h is stable. Note that h has the same support as the polynomial in Example 4.1.6 but different coefficients. Using the computer algebra system `Macaulay2` [33], one checks that conditions of Proposition 4.5.2 are fulfilled and thus Ω_h is smooth. It is the toric variety associated to the triangular frustum whose vertices are obtained by permuting the last three entries of $(0, 3, 0, 0)$ and $(2, 1, 0, 0)$. Here the coordinates correspond to the variables in alphabetic order.

Now let $h = \sigma_{d,n}$ be the elementary symmetric polynomial of degree d.

Lemma 4.5.4. *The centre of the linear projection π_k is disjoint from $X_{B(r_k)\cap\mathbb{Z}^n}$ for each $1 \leq k < d$.*

Proof. The lattice points of $B(r_k)$ are exactly the points $v \in \mathbb{R}^n$ with $d-k$ entries equal to 1 and all other entries 0. Thus $B(r_k)$ is the hypersimplex Δ_{d-k}. We denote $X = X_{\Delta_{d-k}\cap\mathbb{Z}^n} \subset \mathbb{P}^{\binom{n}{d-k}-1}$ and we label the coordinates on $\mathbb{P}^{\binom{n}{d-k}-1}$ by z_S for $S \subset [n]$ of size $d-k$. Every kth order derivative of $\sigma_{d,n}$ is an elementary symmetric polynomial of degree $d-k$ in the variables indexed by some subset $T \subset [n]$ of size $n-k$. Thus the centre E of π_k is the common zero set of all linear forms $L_T = \sum_{S\subset T,\,|S|=d-k} z_S$ for subsets $T \subset [n]$ of size $n-k$. The statement of [52, Lemma 6.4] is that X is disjoint from the common zero set E' of all linear forms $H_i = \sum_{S\subset[n]\setminus i,\,|S|=d-k} z_S$ for $i \in [n]$. We have for all $i \in [n]$:

$$\binom{n+k-1-d}{n-d} \cdot H_i = \sum_{T\subset[n]\setminus\{i\},\,|T|=n-k} L_T.$$

This implies $E \subset E'$ and thus X is also disjoint from E. □

Proof of Theorem 4.1.3. This follows from Lemma 4.5.4 and Proposition 4.5.2. □

Remark 4.5.5. By Proposition 4.2.13, we have that $\Omega_{\sigma_{d,n}}$ is the smooth toric variety X_P where P is the convex hull of all permutations of $(1, \ldots, d-1, 0, \ldots, 0) \in \mathbb{R}^n$.

Remark 4.5.6. Fix some M-convex $S \subset (\mathbb{Z}_{\geq 0})^n$ and let V be the vector space of polynomials h with $\operatorname{supp}(h) \subset S$. Note that there is an integer d such that any $h \in V$ is homogeneous of degree d. There are matrices A_h^k whose entries depend linearly on the coefficients of h such that, when $\operatorname{supp}(h) = S$, the centre of π_k is the set of all $[x]$ with $A_h^k x = 0$. Consider the incidence correspondence

$$\Sigma_k = \{([h], [x]) \in \mathbb{P}(V) \times X_{B(r_k) \cap \mathbb{Z}^n} : A_h^k x = 0\}.$$

This is a projective variety. Thus the projection of Σ_k onto the first factor is a closed subvariety Y_k of $\mathbb{P}(V)$. By construction the criterion from Proposition 4.5.2 applies to a polynomial $h \in V$ with $\operatorname{supp}(h) = S$ if and only if $[h]$ is not contained in any of the Y_k. Therefore, depending on S, either Proposition 4.5.2 applies to no polynomial with support S, or to a generic such polynomial. We say that S is *torically smoothable* if the latter is the case. The support of the elementary symmetric polynomial $\sigma_{d,n}$ is torically smoothable by Lemma 4.5.4. Based on experiments we conjecture that S is torically smoothable at least when S contains the support of $\sigma_{d,n}$. This condition is empty when $d > n$.

Remark 4.5.7. If h has nonnegative coefficients, then we can assume the same for each D_i^k. Then the linear projection π_k is at least regular on the nonnegative part of $X_{B(r_k)}$ as there can be no cancellation of terms. Thus we have at least a regular map on the nonnegative part of $X_{B(\overline{r_1})}$ that maps birationally onto the graph $\Gamma_{h,+}$ of ∇h restricted to the nonnegative orthant. In general, even when h is stable, we cannot expect π_k to be regular on all of $X_{B(r_k)}$. Take for instance the stable polynomial from Example 4.1.6. In this case $B(r_2)$ is the triangular frustum whose vertices are obtained by permuting the last three entries of $(0, 2, 0, 0)$ and $(1, 1, 0, 0)$. Using `Macaulay2` [33] one checks that the centre of π_2 intersects the toric variety $X_{B(r_2)}$ in a real point of the torus orbit corresponding to the face with vertices $(1, 1, 0, 0), (1, 0, 1, 0)$ and $(1, 0, 0, 1)$.

Remark 4.5.8. Let $h \in \mathbb{R}[x_1, \dots, x_n]$ be hyperbolic with respect to $\in \mathbb{R}^n$. In the spirit of the preceding remark one can speculate whether hyperbolicity of h guarantees smoothness of Ω_h at least at some distinguished subset. To make this more precise let $U \subset \mathbb{P}^{n-1}$ be the set of all $[p]$ such that p is in the interior of $\Lambda_e(h)$. Since h does not vanish on U, the gradient map ∇h is regular on U. We can thus consider the subset $C = \omega_h^{-1}(\nabla h(U))$ of Ω_h. We think it is reasonable to ask whether the Euclidean closure of C contains only smooth points of Ω_h.

BIBLIOGRAPHY

[1] Brahim Cherguia Abderrahim Boussaıria. Skew-symmetric matrices and their principal minors. *Linear Algebra and its applications*, 485(1):47–57, 2015.

[2] Raja Hafiz Affandi, Emily Fox, Ryan Adams, and Ben Taskar. Learning the parameters of determinantal point process kernels. In *International Conference on Machine Learning*, pages 1224–1232. PMLR, 2014.

[3] Abeer Al Ahmadieh, Mario Kummer, and Murina-Stefana Sorea. A generalization of the space of complete quadrics. *Le Matematiche*, 76(2), 2021.

[4] Abeer Al Ahmadieh and Cynthia Vinzant. Characterizing principal minors of symmetric matrices via determinantal multiaffine polynomials. *arXiv preprint arXiv:2105.13444*, 2021.

[5] Abeer Al Ahmadieh and Cynthia Vinzant. Determinantal representations and the image of the principal minor map. *arXiv preprint arxiv.2205.05267*, 2022.

[6] Yeganeh Alimohammadi, Nima Anari, Kirankumar Shiragur, and Thuy-Duong Vuong. Fractionally log-concave and sector-stable polynomials: counting planar matchings and more. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 433–446, 2021.

[7] Benjamin Assarf, Ewgenij Gawrilow, Katrin Herr, Michael Joswig, Benjamin Lorenz, Andreas Paffenholz, and Thomas Rehn. Computing convex hulls and counting integer points with polymake. *Mathematical Programming Computation*, 9(1):1–38, 2017.

[8] Arnaud Beauville. Determinantal hypersurfaces. *Michigan Mathematical Journal*, 48(1):39–64, 2000.

[9] Mauro Beltrametti, Ettore Carletti, Giacomo Monti Bragadin, and Dionisio Gallarati. *Lectures on curves, surfaces and projective varieties: a classical view of algebraic geometry*, volume 9. European Mathematical Society, 2009.

[10] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.

[11] Julius Borcea, Petter Brändén, and Thomas Liggett. Negative dependence and the geometry of polynomials. *Journal of the American Mathematical Society*, 22(2):521–567, 2009.

[12] Alexei Borodin. Determinantal point processes. *arXiv preprint arXiv:0911.1153*, 2009.

[13] Petter Brändén. Polynomials with the half-plane property and matroid theory. *Advances in Mathematics*, 216(1):302–320, 2007.

[14] Petter Brändén. Notes on hyperbolicity cones. *https://math. berkeley. edu/b̃ernd/branden. pdf*, 2010.

[15] Petter Brändén. Obstructions to determinantal representability. *Advances in Mathematics*, 226(2):1202–1212, 2011.

[16] Petter Brändén and June Huh. Lorentzian polynomials. *Annals of Mathematics*, 192(3):821–891, 2020.

[17] Victor-Emmanuel Brunel. Learning signed determinantal point processes through the principal minor assignment problem. *Advances in Neural Information Processing Systems*, 31, 2018.

[18] Federico Castillo and Fu Liu. Deformation cones of nested braid fans. *International Mathematics Research Notices*, 2022(3):1973–2026, 2022.

[19] Wei-Lun Chao, Boqing Gong, Kristen Grauman, and Fei Sha. Large-margin determinantal point processes. In *UAI*, 2015.

[20] David Cox, John Little, and Donal OShea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra.* Springer Science & Business Media, 2013.

[21] David A. Cox. *Galois theory.* Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2012.

[22] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms.* Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015. An introduction to computational algebraic geometry and commutative algebra.

[23] David A Cox, John B Little, and Henry K Schenck. *Toric varieties*, volume 124. American Mathematical Soc., 2011.

[24] Leonard Eugene Dickson. Determination of all general homogeneous polynomials expressible as determinants with linear elements. *Transactions of the American Mathematical Society*, 22(2):167–179, 1921.

[25] Charles Lutwidge Dodgson. Condensation of determinants, being a new and brief method for computing their arithmetical values. *Proc. R. Soc. Lond.*, 15:150–155, 1867.

[26] Christophe Dupuy and Francis Bach. Learning determinantal point processes in sublinear time. In *International Conference on Artificial Intelligence and Statistics*, pages 244–257. PMLR, 2018.

[27] Gernot M Engel and Hans Schneider. Matrices diagonally similar to a symmetric matrix. *Linear Algebra and its Applications*, 29:131–138, 1980.

[28] Günter Ewald. *Combinatorial convexity and algebraic geometry*, volume 168. Springer Science & Business Media, 1996.

[29] William Fulton. *Introduction to Toric Varieties.(AM-131), Volume 131.* Princeton university press, 2016.

[30] Mike Gartrell, Victor-Emmanuel Brunel, Elvis Dohmatob, and Syrine Krichene. Learning nonsymmetric determinantal point processes. *Advances in Neural Information Processing Systems*, 32, 2019.

[31] Mike Gartrell, Ulrich Paquet, and Noam Koenigstein. Bayesian low-rank determinantal point processes. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 349–356, 2016.

[32] Eberhard Girlich and Michail Kovalev. Classification of polyhedral matroids. *Mathematical Methods of Operations Research*, 43(2):143–159, 1996.

[33] Daniel R. Grayson and Michael E. Stillman. Macaulay2, a software system for research in algebraic geometry. *`https://faculty.math.illinois.edu/Macaulay2/`*.

[34] Kent Griffin and Michael J Tsatsomeros. Principal minors, part i: A method for computing all the principal minors of a matrix. *Linear Algebra and its applications*, 419(1):107–124, 2006.

[35] Kent Griffin and Michael J Tsatsomeros. Principal minors, part ii: The principal minor assignment problem. *Linear Algebra and its applications*, 419(1):125–171, 2006.

[36] Osman Güler. Hyperbolic polynomials and interior point methods for convex programming. *Mathematics of Operations Research*, 22(2):350–377, 1997.

[37] Joe Harris. *Algebraic geometry: a first course*, volume 133. Springer Science & Business Media, 2013.

[38] DJ Hartfiel and R Leowy. On matrices having equal corresponding principal minors. *Linear algebra and its applications*, 58:147–167, 1984.

[39] Brendan Hassett. *Introduction to algebraic geometry.* Cambridge University Press, 2007.

[40] J William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 60(5):654–674, 2007.

[41] Olga Holtz and Bernd Sturmfels. Hyperdeterminantal relations among symmetric principal minors. *Journal of Algebra*, 316(2):634–648, 2007.

[42] Huajun Huang and Luke Oeding. Symmetrization of principal minors and cycle-sums. *Linear and Multilinear Algebra*, 65(6):1194–1219, 2017.

[43] Kurt Johansson. Random matrices and determinantal processes. *arXiv preprint math-ph/0510038*, 2005.

[44] Richard Kenyon and Robin Pemantle. Principal minors and rhombus tilings. *Journal of Physics A: Mathematical and Theoretical*, 47(47):474010, 2014.

[45] Vijay Kodiyalam, T. Y. Lam, and R. G. Swan. Determinantal ideals, Pfaffian ideals, and the principal minor theorem. In *Noncommutative rings, group rings, diagram algebras and their applications*, volume 456 of *Contemp. Math.*, pages 35–60. Amer. Math. Soc., Providence, RI, 2008.

[46] Alex Kulesza and Ben Taskar. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning*, 5(2-3):123–286, 2012.

[47] John A Kulesza. *Learning with determinantal point processes.* University of Pennsylvania, 2012.

[48] Mario Kummer, Daniel Plaumann, and Cynthia Vinzant. Hyperbolic polynomials, interlacers, and sums of squares. *Mathematical Programming*, 153(1):223–245, 2015.

[49] Dan Laksov. Completed quadrics and linear maps. In *Proceedings of Symposia in Pure Mathematics*, volume 46, pages 371–387, 1987.

[50] Raphael Loewy. Principal minors and diagonal similarity of matrices. *Linear algebra and its applications*, 78:23–64, 1986.

[51] Mateusz Michałek. Selected topics on toric varieties. In *The 50th anniversary of Gröbner bases*, pages 207–252. Mathematical Society of Japan, 2018.

[52] Mateusz Michałek, Bernd Sturmfels, Caroline Uhler, and Piotr Zwiernik. Exponential varieties. *Proceedings of the London Mathematical Society*, 112(1):27–56, 2016.

[53] Thomas Muir. X.—the relations between the coaxial minors of a determinant of the fourth order. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, 39(2):323–339, 1900.

[54] Kazuo Murota. Discrete convex analysis. *Mathematical Programming*, 83(1):313–371, 1998.

[55] EJ Nanson. Xlvi. on the relations between the coaxial minors of a determinant. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):362–367, 1897.

[56] Tim Netzer. Real algebraic geometry and its applications. *Internat. Math. Nachrichten*, 232:1–18, 2016.

[57] Luke Oeding. Set-theoretic defining equations of the variety of principal minors of symmetric matrices. *Algebra & Number Theory*, 5(1):75–109, 2011.

[58] Luke Aaron Oeding. *G-varieties and the principal minors of symmetric matrices.* Texas A&M University, 2009.

[59] Daniel Plaumann and Cynthia Vinzant. Determinantal representations of hyperbolic plane curves: an elementary approach. *J. Symbolic Comput.*, 57:48–60, 2013.

[60] Alexander Postnikov, Victor Reiner, and Lauren Williams. Faces of generalized permutohedra. *Doc.Math*, 13:207–273, 2008.

[61] James Renegar. Hyperbolic programs, and their derivative relaxations. Technical report, Cornell University Operations Research and Industrial Engineering, 2004.

[62] Justin Rising, Alex Kulesza, and Ben Taskar. An efficient algorithm for the symmetric principal minor assignment problem. *Linear Algebra and its Applications*, 473:126–144, 2015.

[63] Igor Rostislavovich Shafarevich. *Basic algebraic geometry*. Springer-Verlag, Berlin-New York, study edition, 1977.

[64] Igor Rostislavovich Shafarevich and Miles Reid. *Basic algebraic geometry*, volume 2. Springer, 1994.

[65] Bernd Sturmfels. *Grobner bases and convex polytopes*, volume 8. American Mathematical Soc., 1996.

[66] Bernd Sturmfels. 3264 questions about symmetric matrices. pdf file on:. *https://www.orlandomarigliano.com/3264questions.pdf*, 2020.

[67] John Urschel, Victor-Emmanuel Brunel, Ankur Moitra, and Philippe Rigollet. Learning determinantal point processes with moments and cycles. In *International Conference on Machine Learning*, pages 3511–3520. PMLR, 2017.

[68] Bert van Geemen and Alessio Marrani. Lagrangian Grassmannians and spinor varieties in characteristic two. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 15:Paper No. 064, 22, 2019.

[69] Victor Vinnikov. Lmi representations of convex semialgebraic sets and determinantal representations of algebraic hypersurfaces: past, present, and future. In *Mathematical methods in systems, optimization, and control*, pages 325–349. Springer, 2012.

[70] Cynthia Vinzant. What is... a spectrahedron. *Notices Amer. Math. Soc*, 61(5):492–494, 2014.

[71] David Wagner. Multivariate stable polynomials: theory and applications. *Bulletin of the American Mathematical Society*, 48(1):53–84, 2011.

[72] David G. Wagner. Multivariate stable polynomials: theory and applications. *Bull. Amer. Math. Soc. (N.S.)*, 48(1):53–84, 2011.

[73] David G. Wagner and Yehua Wei. A criterion for the half-plane property. *Discrete Math.*, 309(6):1385–1390, 2009.

[74] Dominic JA Welsh. *Matroid theory*. Courier Corporation, 2010.