Chapter 7

Spectrahedral Approximations of Convex Hulls of Algebraic Sets

João Gouveia*[†] *and Rekha R. Thomas

This chapter describes a method for finding spectrahedral approximations of the convex hull of a real algebraic variety (the set of real solutions to a finite system of polynomial equations). The procedure creates a nested sequence of convex approximations of the convex hull of the variety. Computations can be done modulo the ideal generated by the polynomials which has several advantages. We examine conditions under which the sequence of approximations converges to the closure of the convex hull of the real variety, either asymptotically or in finitely many steps, with special attention to the case in which the very first approximation yields a semidefinite representation of the convex hull. These methods allow optimization, or approximation of the optimal value, of a linear function over a real algebraic variety via semidefinite programming.

7.1 Introduction

A central problem in optimization is to find the maximum (or minimum) value of a linear function over a set S in $\mathbb{R}^n$. For example, in a *linear program*

$$\text{maximize}\,\{\langle c,x\rangle : Ax \le b\}$$

with $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}$, and $b \in \mathbb{R}^m$, the set $S = \{x \in \mathbb{R}^n \ : \ Ax \le b\}$ is a polyhedron, while in a *semidefinite program*,

$$\text{maximize}\left\{\langle c,x\rangle \ : \ A_0 + \sum_{i=1}^{n} A_i x_i \succeq 0\right\}$$

[†]João Gouveia was partially supported by NSF grant DMS-0757371 and by Fundação para a Ciência e Technologia.

with $c \in \mathbb{R}^n$ and symmetric matrices $A_0, A_1, \ldots, A_n$, the feasible region is the set $S = \{x \in \mathbb{R}^n : A_0 + \sum_{i=1}^n A_i x_i \succeq 0\}$ which is a *spectrahedron*. In both cases, S is a convex *semialgebraic set* as it is convex and can be defined by a finite list of polynomial inequalities. A *real algebraic variety*, which is the set of all real solutions to a finite list of polynomial equations, is a special case of a semialgebraic set. Optimizing a linear function over any set $S \subset \mathbb{R}^n$, in particular, a real algebraic variety, is equivalent to optimizing the linear function over the closure of $\text{conv}(S)$, the convex hull of S. In this chapter we describe a method to construct semidefinite approximations of the closure of the convex hull of a real algebraic variety.

Representing the convex hull of a real algebraic variety is a multifaceted problem that arises in many contexts in both theory and practice. In Chapter 5 we saw a method using dual projective varieties for explicitly finding the polynomials that describe the boundary of the convex hull of a real variety. These bounding polynomials use the same variables as those describing the variety and can be highly complicated. Their computation boils down to eliminating variables from a larger polynomial system and can be challenging in practice, although they can be computed using existing computer algebra packages in examples with a small number of variables. If one is allowed to use more variables than those describing the variety, then there is more freedom in finding representations and approximations and the key idea then is to express the convex hull implicitly as the projection of a higher-dimensional object. This approach is more flexible than the former and has the potential to yield a representation of a complicated set as the projection of a simple set in higher dimensions. The method we will describe adopts this philosophy for finding approximations and representations of the convex hull of a real algebraic variety.

We present a procedure for finding a sequence of approximations of the convex hull of a real algebraic variety (sometimes just called an *algebraic set*) in the form of *projected spectrahedra*. While the convex hull of a real algebraic variety is a convex semialgebraic set, recall from Chapter 6 that it is not known which convex semialgebraic sets are projected spectrahedra. Regardless, we will develop an automatic method that finds semidefinite representations (as projected spectrahedra) for a sequence of outer approximations of $\text{conv}(S)$, when S is an algebraic set. In many cases, these approximations will converge to $\text{conv}(S)$. If our procedure yields an exact representation of $\text{conv}(S)$ as a projected spectrahedron, then as a by product we can optimize a linear function over S by solving a semidefinite program. In the nice cases where the representation uses spectrahedra of small size (relative to the size of S), semidefinite programming becomes an efficient method for optimizing a linear function over S. In fact, there are several families of algebraic sets where this spectrahedral approach yields polynomial time algorithms for linear optimization. Similarly, the spectrahedral approach can, in some cases, yield efficient algorithms for finding good approximations of the optimal value of a linear function over S.

While we will see many examples of real algebraic varieties (and their defining ideals) for which our method yields an exact representation of its convex hull in a few iterations of our procedure, many open questions remain. For instance, there is no complete understanding of when the method is guaranteed to converge to the convex hull of the variety in finitely many steps of the procedure. Even in the cases where finite convergence is guaranteed, good upper bounds on the number of iterations required by the procedure are lacking. The work presented in this chapter was inspired by a question posed by Lovász in [19] that asked for a characterization of ideals for which the first approximation in our hierarchy will yield a semidefinite representation of the convex hull of the variety of the ideal. In Section 7.3 we answer this question for finite varieties. The case of infinite varieties is far less understood. We identify conditions that prevent finite convergence of these approximations to the closure of the convex hull of the variety. However, again a full characterization is missing. Thus, the material in this chapter offers both advances in spectrahedral representations of algebraic sets as well as many avenues for further research.

This chapter is organized as follows. In Section 7.2 we explain the procedure for finding spectrahedral approximations of the convex hull of an algebraic set. These techniques were developed in [8], coauthored with Parrilo. One of the key theorems needed in this section (Theorem 7.6) was strengthened in this presentation with the help of Greg Blekherman. We illustrate the method with various examples and explain the underlying computations. In Section 7.3 we discuss the situations in which this method converges, either asymptotically or finitely, to an exact semidefinite representation of the convex hull of the variety. The most useful scenario is when the first approximation yields an exact semidefinite representation of the convex hull of the variety. We characterize all finite varieties for which this happens. We conclude in Section 7.4 with examples from combinatorial optimization where the underlying varieties are all finite. The methods we describe have algorithmic impact on certain classes of combinatorial optimization problems and the algebra becomes endowed with rich combinatorics in these cases.

7.2 The Method

Let $f_1, \ldots, f_m \in \mathbb{R}[x_1, \ldots, x_n] =: \mathbb{R}[x]$ be polynomials and

$$V_{\mathbb{R}}(f_1, \ldots, f_m) := \{x \in \mathbb{R}^n \,:\, f_1(x) = f_2(x) = \cdots = f_m(x) = 0\}$$

be their set of real zeros. We are interested in representing $\text{conv}(V_{\mathbb{R}}(f_1, \ldots, f_m))$, the convex hull of $V_{\mathbb{R}}(f_1, \ldots, f_m)$ in $\mathbb{R}^n$ as projected spectrahedra.

Recall that the *ideal* generated by $f_1, \ldots, f_m$ in $\mathbb{R}[x]$ is the set

$$I = \langle f_1, \ldots, f_m \rangle = \left\{ \sum_{i=1}^{m} g_i f_i \,:\, g_i \in \mathbb{R}[x],\ m \in \mathbb{N} \right\} \subset \mathbb{R}[x].$$

The *real variety* of I is the set $V_{\mathbb{R}}(I) := \{x \in \mathbb{R}^n \,:\, h(x) = 0 \text{ for all } h \in I\}$ of real zeros of all polynomials in I. Note that if $s \in V_{\mathbb{R}}(f_1, \ldots, f_m)$, then $s \in V_{\mathbb{R}}(I)$ since $f_i(s) = 0$ implies that $h(s) = \sum_{i=1}^m g_i(s) f_i(s) = 0$ for all $h \in I$. Conversely, if $s \in V_{\mathbb{R}}(I)$, then for all $i = 1, \ldots, m$, $f_i(s) = 0$ since $f_i \in I$. Therefore, $V_{\mathbb{R}}(f_1, \ldots, f_m) = V_{\mathbb{R}}(I)$, and our goal can be viewed more generally as wanting to find semidefinite representations of the convex hull of the real variety of an ideal in $\mathbb{R}[x]$, or approximations of it.

For any set $S \subseteq \mathbb{R}^n$, the closure of $\text{conv}(S)$ is exactly the intersection of all closed half spaces $\{x \in \mathbb{R}^n \,:\, l(x) \geq 0\}$ as l varies over all *linear* polynomials that are nonnegative on S. Throughout this chapter, linear polynomials include *affine* linear polynomials (those with a constant term). In particular, given an ideal I,

$$\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I))) = \bigcap_{l \text{ linear},\ l|_{V_{\mathbb{R}}(I)} \geq 0} \{x : l(x) \geq 0\}.$$

It is not so clear how to work with this description. Even for a single linear polynomial l, checking whether $l(x)$ is nonnegative on $V_{\mathbb{R}}(I)$ is a difficult task. A natural idea is to relax the condition $l|_{V_{\mathbb{R}}(I)} \geq 0$ to something easier to check, at the risk of losing some of the $l(x)$ in the above intersection, and obtaining a superset of $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$. As seen already in Chapters 3 and 4, the classical method to certify the nonnegativity of a polynomial on all of $\mathbb{R}^n$ is to write it as a *sum of squares* (sos) of other polynomials. In our case, we just need to certify that $l(x)$ is nonnegative on $V_{\mathbb{R}}(I)$, a subset of $\mathbb{R}^n$.

Let Σ denote the set of all sos polynomials in $\mathbb{R}[x]$, $\mathbb{R}[x]_k$ the set of all polynomials in $\mathbb{R}[x]$ of degree at most k, and Σ_{2k} the set of all sos polynomials $\sum h_j^2$, where $h_j \in \mathbb{R}[x]_k$. Nonnegativity of $l(x)$ on $V_{\mathbb{R}}(I)$ is guaranteed if

$$l(x) = \sigma(x) + \sum_{i=1}^{m} g_i(x) f_i(x) \tag{7.1}$$

for $\sigma(x) \in \Sigma$ and $g_i \in \mathbb{R}[x]$, since then for all $s \in V_{\mathbb{R}}(I)$, $l(s) = \sigma(s) \geq 0$. In Chapter 3 we saw that semidefinite programming can be used to check whether a polynomial is sos. In (7.1) we need to find both $\sigma(x)$ and the polynomials g_i to write $l(x)$ as *sos mod* I. Therefore, to check (7.1) in practice, we impose degree restrictions and proceed in one of two possible ways.

(i) In the first method, we ask that $\sigma \in \Sigma_{2k}$ and $g_i f_i \in \mathbb{R}[x]_{2k}$ for a fixed positive integer k and, if so, say that $l(x)$ is k-*sos mod* $\{f_1, \ldots, f_m\}$. This is the basic idea that underlies Lasserre's moment method for approximating the convex hull of a semialgebraic set described in Chapter 6.

(ii) In the second method, we ask only that $\sigma \in \Sigma_{2k}$ for a fixed positive integer k which reduces (7.1) to $l(x) = \sigma(x) + h(x)$ where $h(x) \in I$. If this is the case, we say that $l(x)$ is k-*sos mod* I. This method is more natural if one is interested in the geometry of $V_{\mathbb{R}}(I)$ and $\mathrm{conv}(V_{\mathbb{R}}(I))$ as it removes the dependence of the method on the choice of a particular generating set of I. The only issue is if the computation can be done in practice at the level of the ideal I and not the input $f_1, \ldots, f_m$.

Both methods yield a hierarchy of convex relaxations of $\mathrm{conv}(V_{\mathbb{R}}(I))$ obtained as the intersection of all half spaces $\{x : l(x) \geq 0\}$ as $l(x)$ ranges over the linear polynomials that are k-sos in the sense of the method. Since if $l(x)$ is k-sos mod $\{f_1, \ldots, f_m\}$ then it is also k-sos mod I, method (ii) yields a relaxation that is no worse than that from method (i) for each value of k. On the other hand, method (ii) requires the knowledge of a basis of $\mathbb{R}[x]/I$ as we will see below, which for some problems may be hard to compute in practice. To see the computational differences that can occur between the two methods, consult Remark 7.14.

In this chapter we focus on method (ii). The kth iteration of (ii) yields a closed convex set, called the kth *theta body* of I, defined as

$$\mathrm{TH}_k(I) := \{x \in \mathbb{R}^n : l(x) \geq 0 \text{ for all } l \text{ linear and } k\text{-sos mod } I\}.$$

Clearly $V_{\mathbb{R}}(I)$, and hence $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$, is contained in $\mathrm{TH}_k(I)$ for all k. Thus the theta bodies of I form a hierarchy of closed convex approximations of $\mathrm{conv}(V_{\mathbb{R}}(I))$ as follows:

$$\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \mathrm{TH}_k(I) \supseteq \mathrm{TH}_{k+1}(I) \supseteq \cdots \supseteq \mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I))).$$

An immediate question is when this hierarchy converges to $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$ either finitely or asymptotically. Finite convergence allows an exact representation of $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$ as a theta body which would be extremely useful if we can represent and optimize over a theta body efficiently. We will show in Section 7.2.2 that each $\mathrm{TH}_k(I)$ is the closure of a projected spectrahedron. This enables optimization over a real variety using semidefinite programming. In Section 7.4, we will learn the motivation for the name "theta bodies." We begin with some background on working modulo a polynomial ideal.

7.2.1 Sum of Squares Modulo an Ideal

Let $I \subseteq \mathbb{R}[x]$ be an ideal and $V_{\mathbb{R}}(I)$ be its real variety. For two polynomials $f, g \in \mathbb{R}[x]$, if $f - g \in I$, then $f(s) = g(s)$ for all $s \in V_{\mathbb{R}}(I)$. If $f - g \in I$, then f and g are said to be *congruent mod* I, written as $f \equiv g \bmod I$. Congruence mod I is an equivalence relation on $\mathbb{R}[x]$. The equivalence class of f is denoted as $f + I$, and the set of equivalence classes is denoted as $\mathbb{R}[x]/I$. The set $\mathbb{R}[x]/I$ is both an $\mathbb{R}$-vector space and a ring over $\mathbb{R}$ where addition, scalar multiplication, and multiplication are defined as follows. Given $f, g \in \mathbb{R}[x]$ and $\lambda \in \mathbb{R}$, $(f + I) + (g + I) = (f + g) + I$, $\lambda(f + I) = \lambda f + I$, and $(f + I)(g + I) = fg + I$. We will denote vector space bases of $\mathbb{R}[x]/I$ by $\mathcal{B}$ in this chapter. By the *degree* of an equivalence class $f + I$, we mean the smallest degree of an element in the class. With this definition, we may assume that the elements of $\mathcal{B}$ are listed in order of increasing degree. Further, for each $k \in \mathbb{N}$, the set $\mathcal{B}_k$ of all elements in $\mathcal{B}$ of degree at most k is then well-defined.

Computations in $\mathbb{R}[x]/I$ can be done via *Gröbner bases* of I. Recall that if G is any reduced Gröbner basis of I, then a polynomial h lies in I if and only if the *normal form* of h with respect to G is zero. Therefore, $f \equiv g \bmod I$ if and only if the normal form of $f - g$ with respect to G is zero, or equivalently, f and g have the same normal form with respect to G. This provides an algorithm to check whether two polynomials are congruent mod I. The unique normal form of all polynomials in the same equivalence class serves as a canonical representative for this class given G. If M is the *initial ideal* of I corresponding to the reduced Gröbner basis G, then recall that the *standard monomials* of M form an $\mathbb{R}$-vector space basis for $\mathbb{R}[x]/I$. Therefore, the normal form of a polynomial with respect to G can be written as an $\mathbb{R}$-linear combination of the standard monomials of the initial ideal M. The vector space $\mathbb{R}[x]/I$ has many other bases, some of which may be better suited for computations than the standard monomial bases coming from an initial ideal of I. See Chapter 3 for a discussion of alternative bases of $\mathbb{R}[x]$ and hence $\mathbb{R}[x]/I$. In this chapter we will use only a standard monomial basis of $\mathbb{R}[x]/I$. A quick tour of the algebraic notions needed in this chapter can be found in the appendix. For a thorough introduction to the theory of Gröbner bases and related notions, we refer the reader to [6].

We now come to sum of squares polynomials modulo an ideal I, and the question of how to check whether a polynomial $f \in \mathbb{R}[x]$ is k-sos mod I. A polynomial $f \in \mathbb{R}[x]$ is *sos mod I* if $f \equiv \sum h_j^2$ mod I for some $h_j \in \mathbb{R}[x]$, and *k-sos mod I* if $h_j \in \mathbb{R}[x]_k$ for all j. Hence, the equivalence classes of polynomials that are sos mod I (respectively, k-sos mod I) are precisely those in

$$\Sigma/I := \{\sigma + I \,:\, \sigma \in \Sigma\}$$

(respectively, Σ_{2k}/I). It is worthwhile to note that many polynomials that are not sos in $\mathbb{R}[x]$ can become sos mod an ideal I. For instance, the univariate linear polynomial x is congruent to x^2 mod the ideal $\langle x - x^2 \rangle \subset \mathbb{R}[x]$.

Let $[x]_k$ denote the vector of all monomials in $\mathbb{R}[x]_k$ in a fixed order, say degree lexicographic. Recall from Chapter 3 that a polynomial $f \in \Sigma_{2k}$ if and only if there exists a positive semidefinite matrix A, denoted $A \succeq 0$, such that $f = [x]_k^T A [x]_k$. The matrix A can be solved for using semidefinite programming and a Cholesky factorization of it as $A = V^T V$ yields an sos expression $\sum h_j^2$ for f, where $h_j(x)$ is the inner product of the jth row of V and the vector of monomials $[x]_k$. This method can be adapted to check whether f is k-sos mod I as follows. The vector $[x]_k$ can be replaced by the vector of monomials from $\mathcal{B}_k$, denoted as $[x]_{\mathcal{B}_k}$, since $\mathbb{R}[x]_k/I$ is spanned by $\mathcal{B}_k$. Since the size of $\mathcal{B}_k$ is no larger than the size of a basis of $\mathbb{R}[x]_k$, this can decrease the size of the unknown matrix A considerably, making the final SDP much smaller than before. Setting up A as a symmetric matrix of indeterminates A_{ij} and multiplying out $[x]_{\mathcal{B}_k}^T A [x]_{\mathcal{B}_k}$, we get a polynomial $g \in \mathbb{R}[x]_{2k}$. Let the normal forms of f and g with respect to a reduced Gröbner basis G of I be f' and g', respectively. Then since $f \equiv f'$ and $g \equiv g'$ mod I and f' and g' are fully reduced with respect to G, we have that $f \equiv g$ mod I if and only if $f' = g'$. Therefore, to check if f is k-sos mod I, we equate the coefficients of f' and g' for like monomials and check whether the resulting linear system in the A_{ij}'s has a solution with $A \succeq 0$.

Example 7.1. Consider the polynomial $f(x, y) = x^4 + y^4 + 2x^2y^2 - x^2 + y^2$ and the principal ideal $I = \langle f \rangle \subset \mathbb{R}[x, y]$. The real variety $V_{\mathbb{R}}(I)$, which is the set of real zeros of f, is a *Bernoulli lemniscate* (shown in Figure 7.1) with foci $(\pm\frac{1}{\sqrt{2}}, 0)$.

It is easy to check that the horizontal line $y = \frac{1}{\sqrt{8}}$ is a bitangent to $V_{\mathbb{R}}(I)$ and that $l(x, y) := -y + \frac{1}{\sqrt{8}}$ is nonnegative on $V_{\mathbb{R}}(I)$. Since f has degree 4 and l has degree 1, l cannot be 1-sos mod I but has a chance to be 2-sos mod I. We apply the method described above to verify this.

The set $\{f\}$ is a reduced Gröbner basis of I with respect to every term order. The initial ideal of I under the total degree order with ties broken lexicographically with $x > y$, is generated by x^4. Hence a basis $\mathcal{B}$ for $\mathbb{R}[x, y]/I$ is given by the infinite set of standard monomials of $\langle x^4 \rangle \subseteq \mathbb{R}[x, y]$ which are all the monomials in x and y

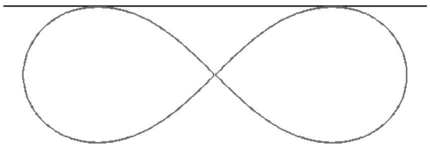

Figure 7.1. *The lemniscate* $x^4 + y^4 + 2x^2y^2 - x^2 + y^2 = 0$ *with a bitangent.*

that are not divisible by x^4. In particular, $\mathcal{B}_1 = \{1, x, y\}$, $\mathcal{B}_2 = \{1, x, y, x^2, xy, y^2\}$, and $[x]_{\mathcal{B}_2} = (1\ x\ y\ x^2\ xy\ y^2)$.

The general 2-sos polynomial mod I is therefore of the form

$$g = \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{12} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{13} & a_{23} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} & a_{46} \\ a_{15} & a_{25} & a_{35} & a_{45} & a_{55} & a_{56} \\ a_{16} & a_{26} & a_{36} & a_{46} & a_{56} & a_{66} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix},$$

where $A = (a_{ij}) \succeq 0$. Multiplying out the above expression we get that

$g := a_{11} + 2a_{12}x + 2a_{13}y + (2a_{14} + a_{22})x^2 + (2a_{23} + 2a_{15})xy + (2a_{16} + a_{33})y^2 + 2a_{24}x^3 + (2a_{34} + 2a_{25})x^2y + (2a_{26} + 2a_{35})xy^2 + 2a_{36}y^3 + a_{44}x^4 + 2a_{45}x^3y + (a_{55} + 2a_{46})x^2y^2 + 2a_{56}xy^3 + a_{66}y^4.$

We now reduce g by the Gröbner basis $\{f\}$, which means replacing every occurrence of x^4 with

$$-y^4 - 2x^2y^2 + x^2 - y^2,$$

and obtain the normal form of g, which is

$g' := a_{11} + 2a_{12}x + 2a_{13}y + (2a_{14} + a_{22} + a_{44})x^2 + (2a_{23} + 2a_{15})xy + (2a_{16} + a_{33} - a_{44})y^2 + 2a_{24}x^3 + (2a_{34} + 2a_{25})x^2y + (2a_{26} + 2a_{35})xy^2 + 2a_{36}y^3 + 2a_{45}x^3y + (a_{55} + 2a_{46} - 2a_{44})x^2y^2 + 2a_{56}xy^3 + (a_{66} - a_{44})y^4.$

Since $l(x, y) = -y + \frac{1}{\sqrt{8}}$ is already reduced with respect to $\{f\}$, if l is 2-sos mod I, then $l = g'$, and hence to verify this, we need to check whether there exists $A \succeq 0$ such that $a_{11} = \frac{1}{\sqrt{8}}, 2a_{13} = -1$, and all other coefficients of g' equal zero. Writing out all the linear conditions, we need to check whether there exists a positive semidefinite matrix of the form

$$\begin{bmatrix} \frac{1}{\sqrt{8}} & 0 & -\frac{1}{2} & a_{14} & a_{15} & a_{16} \\ 0 & a_{22} & -a_{15} & 0 & a_{25} & a_{26} \\ -\frac{1}{2} & -a_{15} & a_{33} & -a_{25} & -a_{26} & 0 \\ a_{14} & 0 & -a_{25} & a_{44} & 0 & a_{46} \\ a_{15} & a_{25} & -a_{26} & 0 & a_{55} & 0 \\ a_{16} & a_{26} & 0 & a_{46} & 0 & a_{44} \end{bmatrix}$$

that satisfies the conditions

$$2a_{14} + a_{22} + a_{44} = 0, \quad 2a_{16} + a_{33} - a_{44} = 0, \quad a_{55} + 2a_{46} - 2a_{44} = 0.$$

Check that the matrix

$$A = \begin{bmatrix} 2^{-3/2} & 0 & -1/2 & -2^{-3/2} & 0 & -2^{-3/2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1/2 & 0 & 2^{1/2} & 0 & 0 & 0 \\ -2^{-3/2} & 0 & 0 & 2^{-1/2} & 0 & 2^{-1/2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2^{-3/2} & 0 & 0 & 2^{-1/2} & 0 & 2^{-1/2} \end{bmatrix}$$

is positive semidefinite and satisfies the conditions given above. This matrix A factors as $A = V^T V$ with

$$V = \begin{bmatrix} -2^{-5/4} & 0 & 0 & 2^{-1/4} & 0 & 2^{-1/4} \\ -2^{-5/4} & 0 & 2^{1/4} & 0 & 0 & 0 \end{bmatrix},$$

and hence,

$$\left(\frac{1}{\sqrt{8}} - y\right) \equiv \frac{1}{4\sqrt{2}}\left(2x^2 + 2y^2 - 1\right)^2 + \sqrt{2}\left(y - \frac{1}{\sqrt{8}}\right)^2 \mod I.$$

In general, finding exact sos expressions, as above, is difficult. This particular sos decomposition was found by Bruce Reznick using a series of tricks. He showed that

$$\begin{aligned} &(\tfrac{1}{\sqrt{8}} - y) + \tfrac{1}{\sqrt{2}}((x^2+y^2)^2 - (x^2-y^2)) \\ &= \tfrac{1}{4\sqrt{2}}\left(2x^2 + 2y^2 - 1\right)^2 + \sqrt{2}\left(y - \tfrac{1}{\sqrt{8}}\right)^2. \end{aligned}$$

In practice, one can use an SDP solver to find A. Using MATLAB, to do this computation in YALMIP [17] we input the following code:

```
sdpvar a14 a15 a16 a22 a25 a26 a33 a44 a46 a55
A=[ 1/sqrt(8) 0   -1/2  a14  a15 a16;
    0         a22 -a15  0    a25 a26;
   -1/2      -a15  a33 -a25 -a26 0  ;
    a14       0   -a25  a44  0   a46;
    a15       a25 -a26  0    a55 0  ;
    a16       a26  0    a46  0   a44];
```

```
l1=2*a14 + a22 + a44;
l2=2*a16 + a33 - a44;
l3=a55 + 2*a46 -2*a44;
solvesdp([A>0,l1==0,l2==0,l3==0],0);
```

We ran this code with SeDuMi 1.1 as the underlying SDP solver in YALMIP. The matrix can now be recovered by simply typing `double(A)` and we obtain

$$A = \begin{bmatrix} 0.3536 & 0.0000 & -0.5000 & -0.4052 & 0.0000 & -0.1985 \\ 0.0000 & 0.1034 & 0.0000 & 0.0000 & -0.2924 & 0.0000 \\ -0.5000 & 0.0000 & 1.1041 & 0.2924 & 0.0000 & 0.0000 \\ -0.4052 & 0.0000 & 0.2924 & 0.7071 & 0.0000 & 0.2936 \\ 0.0000 & -0.2924 & 0.0000 & 0.0000 & 0.8270 & 0.0000 \\ -0.1985 & 0.0000 & 0.0000 & 0.2936 & 0.0000 & 0.7071 \end{bmatrix},$$

in which the entries are shown up to four digits of precision. After factorizing A as $V^T V$ we obtain the sos decomposition:

$$\begin{array}{ll} & \left(0.5946427499 - 0.8408409925\,y - 0.6814175403\,x^2 - 0.3338138740\,y^2\right)^2 \\ + & (0.3215587038\,x - 0.9093207446\,xy)^2 \\ + & \left(0.6301479392\,y - 0.4452348146\,x^2 - 0.4454261796\,y^2\right)^2 \\ + & \left(0.2110357686\,x^2 - 0.6263671431\,y^2\right)^2 \\ + & 0.0001357833655\,x^2y^2 \\ + & 0.004928018144\,y^4, \end{array}$$

which simplifies to

$$\begin{array}{ll} & 0.3536000000 - \ y \\ + & 0.707(x^4 + 2x^2y^2 + y^4 - x^2 + y^2) \\ + & 10^{-11}(8.089965190\,x^2y - 3.247827064\,y^3). \end{array}$$

This provides fairly strong computational evidence that $l = \frac{1}{\sqrt{8}} - y$ is 2-sos mod I even though it is not an exact 2-sos representation of l mod I.

The above approach becomes cumbersome as we search for higher and higher degree sums of squares modulo an ideal. Luckily there are ways of using the existing software to simplify our input. In our example, checking whether l is 2-sos modulo I is the same as checking if there exists some $\lambda \in \mathbb{R}$ such that $l(x, y) + \lambda f(x, y)$ is sos, which can be done via YALMIP with the following commands:

```
sdpvar x y lambda
f=x^4+y^4+2*x^2*y^2-x^2+y^2;
l=1/sqrt(8)-y;
F=sos(l+lambda*f);
solvesos(F,0,[],lambda);
sdisplay(sosd(F))
```

The last command will actually display a list of polynomials whose squares sum up to (approximately) $l(x, y) + \lambda f(x, y)$. In our example, the following output is obtained

```
'-0.5919274724+0.8880*y+0.6222*x^2+0.3571*y^2'
'-0.03240303655-0.5699*y+0.4037*x^2+0.6602*y^2'
'-0.3036*x+0.8587*x*y'
'-0.0461010126-0.1559*y+0.3963*x^2-0.3792*y^2'
'9.2958e-05*x+3.2868e-05*x*y'
'3.789017278e-05+1.3396e-05*y+1.4209e-05*x^2+4.7355e-06*y^2'
```

which should be interpreted as saying that $l(x, y)$ is the sum of squares of the polynomials shown on each line. Note that the last two polynomials in the list above again point to the fact that the software only provided reasonable evidence that $l(x, y)$ is 2-sos mod I. ■

The above computations also give a glimpse into the intertwining of algebraic and numerical methods that is prevalent in convex algebraic geometry. The question of whether a polynomial is a sum of squares modulo an ideal is purely algebraic. However, the search for an sos expression is done via semidefinite programming which is solved using numerical methods. The answer provided by these numerical solvers is often not exact. Massaging the numerical information into a certifiable answer can sometimes be an art.

Example 7.2. Consider the polynomial $g(x, y) := y^2(1 - x^2) - (x^2 + 2y - 1)^2$ and the ideal $I = \langle g(x, y)\rangle$ defining the *bicorn curve* shown in Figure 7.2. It is clear that $y \geq 0$ over the curve. Instead of checking if y is k-sos mod I for some k (which is never the case as we will see in the next section), it is in general more useful to search for the smallest μ such that $y + \mu$ is k-sos mod I. That way, if y is not sos mod I, we will at least obtain a valid inequality $y + \mu \geq 0$ on $V_{\mathbb{R}}(I)$ which will then be valid for $\mathrm{TH}_k(I)$. In general, $y + \mu$ is k-sos mod I if there exists some polynomial $h(x, y)$ of degree $2k - 4$ such that $(y + \mu) + h(x, y)g(x, y)$ is sos. As before, this can be checked easily using YALMIP.

```
k=2;
sdpvar x y mu
[h,c]=polynomial([x y],2*k-4);
g=y^2*(1-x^2)-(x^2+2*y-1)^2;
F=sos(y+mu-h*g);
solvesos(F,mu,[],[mu;c]);
```

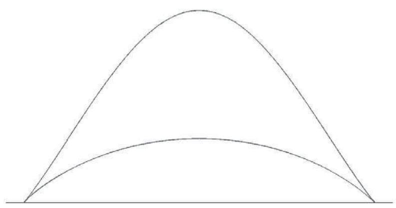

Figure 7.2. *A bicorn curve.*

By successively setting k to be 2, 3, and 4, we get that the minimum value of μ (recovered using `double(mu)`) is 0.1776, 0.0370, and 0.0161, respectively. So while μ is approaching 0, it seems that y is at least not 4-sos mod I. ■

7.2.2 Theta Bodies

We now come back to theta bodies of the ideal I and their representations. Recall that the kth theta body of I is

$$\mathrm{TH}_k(I) := \{x \in \mathbb{R}^n : l(x) \geq 0 \text{ for all } l \text{ linear and } k\text{-sos mod } I\}.$$

Given any polynomial, it is possible to check whether it is k-sos mod I using Gröbner bases and semidefinite programming as seen in Section 7.2.1. The bottleneck in using the definition of $\mathrm{TH}_k(I)$ in practice is that it requires knowledge of all the linear polynomials (infinitely many) that are k-sos mod I. To overcome this difficulty we will now derive an alternative description of $\mathrm{TH}_k(I)$ as a projected spectrahedron (up to closure) which enables computations via semidefinite programming.

We may assume that there are no linear polynomials in the ideal I since otherwise, some variable x_i is congruent to a linear combination of other variables mod I, and we may work in a smaller polynomial ring. Therefore, $\mathbb{R}[x]_1/I \cong \mathbb{R}[x]_1$ and $\{1 + I, x_1 + I, \ldots, x_n + I\}$ can be completed to a basis $\mathcal{B}$ of $\mathbb{R}[x]/I$. Recall the definition of degree of $f + I$. We will assume that each element in a basis $\mathcal{B} = \{f_i + I\}$ of $\mathbb{R}[x]/I$ is represented by a polynomial whose degree equals the degree of its equivalence class, and that $\mathcal{B}$ is ordered so that $\deg(f_i + I) \leq \deg(f_{i+1} + I)$. Further, $\mathcal{B}_k$ denotes the ordered subset of $\mathcal{B}$ of degree at most k.

Definition 7.3. *Let $I \subseteq \mathbb{R}[x]$ be an ideal. A basis $\mathcal{B} = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[x]/I$ is a θ-basis if it has the following properties:*

1. $\mathcal{B}_1 = \{1 + I, x_1 + I, \ldots, x_n + I\}$.
2. *If $\deg(f_i + I), \deg(f_j + I) \leq k$, then $f_i f_j + I$ is in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$.*

Our goal will be to first express the kth theta body $\mathrm{TH}_k(I)$ as the closure of a certain set of linear functionals on the k-sos polynomials mod I. This will be achieved in Theorem 7.6. In the case where I contains the polynomials $x_i^2 - x_i$ for all $i = 1, \ldots, n$, the closure can be removed (Theorem 7.8). Such ideals appear in combinatorial optimization and hence this result will have an important role in Section 7.4. After this, we use a θ-basis of the quotient ring $\mathbb{R}[x]/I$ to turn the description of $\mathrm{TH}_k(I)$ in Theorem 7.6 to an explicit semidefinite representation. This allows concrete computations and examples. We proceed toward Theorem 7.6.

In what follows, we identify a linear polynomial $\alpha + \langle a, x\rangle \in \mathbb{R}[x]_1$ with the vector $(\alpha, a) \in \mathbb{R}^{n+1}$. Let $\Sigma_1^k(I) := \{f + I : f \in \mathbb{R}[x]_1, f \text{ } k\text{-sos mod } I\}$. Then $\Sigma_1^k(I)$ is a cone in the vector space $\mathbb{R}[x]_1/I \cong \mathbb{R}[x]_1$, and its dual cone $\Sigma_1^k(I)^*$ lives in $(\mathbb{R}[x]_1/I)^* \cong \mathbb{R}[x]_1^* \cong \mathbb{R}^{n+1}$. Thus,

$$\Sigma_1^k(I)^* = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : \alpha t + \langle a, x\rangle \geq 0 \text{ for all } (\alpha, a) \in \Sigma_1^k(I)\}.$$

Consider the hyperplane $H := \{(1, x) : x \in \mathbb{R}^n\}$ in $\mathbb{R}^{n+1}$. We may think of H also as $H = \{L \in (\mathbb{R}[x]_1/I)^* : L(1 + I) = 1\}$. It then follows immediately that

$$\{1\} \times \mathrm{TH}_k(I) = \Sigma_1^k(I)^* \cap H. \tag{7.2}$$

Lemma 7.4. *The hyperplane H intersects the relative interior of $\Sigma_1^k(I)^*$.*

Proof. A sufficient condition for a hyperplane L to intersect the relative interior of a closed convex cone P is that $\mathrm{cl}(\mathrm{cone}(\mathrm{relint}(P \cap L))) = P$. If L does not intersect the relative interior of P, then $P \cap L$ is contained in some proper face F of P (possibly the empty face). Therefore, $\mathrm{cl}(\mathrm{cone}(\mathrm{relint}(P \cap L)))$ is also contained in this face which is a proper subset of P.

By (7.2), $C := \{(\lambda, \lambda x) : \lambda \geq 0, x \in \mathrm{relint}(\mathrm{TH}_k(I))\}$ is the cone over the relative interior of $\Sigma_1^k(I)^* \cap H$. We will show that $\mathrm{cl}(C) = \Sigma_1^k(I)^*$. Let $(\alpha, a) \in \Sigma_1^k(I)$ and $x \in \mathrm{relint}(\mathrm{TH}_k(I))$. Then since $x \in \mathrm{TH}_k(I)$, $0 \leq \alpha + \langle a, x\rangle = \langle(\alpha, a), (1, x)\rangle$ which implies that $0 \leq \langle(\alpha, a), (\lambda, \lambda x)\rangle$ for all $\lambda \geq 0$. Hence $C \subseteq \Sigma_1^k(I)^*$, and since $\Sigma_1^k(I)^*$ is closed, $\mathrm{cl}(C) \subseteq \Sigma_1^k(I)^*$.

Suppose $\Sigma_1^k(I)^* \not\subseteq \mathrm{cl}(C)$. Then there exists $(t, x) \in \Sigma_1^k(I)^* \backslash \mathrm{cl}(C)$. Since the constant polynomial 1 lies in $\Sigma_1^k(I)$ and $(t, x) \in \Sigma_1^k(I)^*$, $t \geq 0$. Also, since $\mathrm{cl}(C)$ is closed and there exists $(s, y) \in C$ with $s > 0$, we can find a small enough $\epsilon > 0$ such that $(t, x) + \epsilon(s, y) \in \Sigma_1^k(I)^* \backslash \mathrm{cl}(C)$, and the first coordinate of $(t, x) + \epsilon(s, y)$ is positive. Scaling this element, we may assume that there is an element $(1, x) \in \Sigma_1^k(I)^* \backslash \mathrm{cl}(C)$. Since $(1, x) \in \Sigma_1^k(I)^*$, $\alpha + \langle a, x\rangle \geq 0$ for all $(\alpha, a) \in \Sigma_1^k(I)$, which implies that $x \in \mathrm{TH}_k(I)$ and hence $(1, x) \in \mathrm{cl}(C)$, which is a contradiction. □

We will also need the following lemma which can be proved using standard tools of convex geometry.

Lemma 7.5. *Let P be a closed convex cone and Q be a convex subcone of P such that $\mathrm{cl}(Q) = P$. Then $\mathrm{relint}(P) \subseteq Q$, and for any affine hyperplane H passing through the relative interior of P, $P \cap H = \mathrm{cl}(Q \cap H)$.*

We now examine the cone $\Sigma_1^k(I)^*$ more closely. Let $\Sigma^k(I)$ denote the set of all $f + I$ such that f is k-sos mod I. Then $\Sigma^k(I) = \Sigma_{2k}/I$ is a cone in $\mathbb{R}[x]_{2k}/I$, and $\Sigma_1^k(I) = \Sigma^k(I) \cap \mathbb{R}[x]_1/I$. Therefore, the dual cone of $\Sigma_1^k(I)$ in $(\mathbb{R}[x]/I)^*$ is the closure of the projection of $\Sigma^k(I)^*$ into $(\mathbb{R}[x]_1/I)^*$ as explained in Section 2.1 of Chapter 5. Hence we may identify $\Sigma_1^k(I)^*$ with the closure of the set

$$S_k(I) := \{(L(1 + I), L(x_1 + I), \ldots, L(x_n + I)) : L \in \Sigma^k(I)^*\}.$$

Further, define $Q_k(I) := \big\{(L(x_1 + I), \ldots, L(x_n + I)) : L \in \Sigma^k(I)^*, L(1 + I) = 1\big\}$. We will see shortly that $Q_k(I)$ is a projected spectrahedron, but first we establish the connection between $\mathrm{TH}_k(I)$ and $Q_k(I)$.

Theorem 7.6. $\mathrm{TH}_k(I) = \mathrm{cl}(Q_k(I))$.

Proof. Since $\{1\} \times Q_k(I) = S_k(I) \cap H$, we have $\{1\} \times \mathrm{cl}(Q_k(I)) = \mathrm{cl}(S_k(I) \cap H)$. Since $\mathrm{cl}(S_k(I)) = \Sigma_1^k(I)^*$, it follows from (7.2) that $\{1\} \times \mathrm{TH}_k(I) = \mathrm{cl}(S_k(I)) \cap H$. Therefore, the theorem will follow if we can show that

$$\mathrm{cl}(S_k(I)) \cap H = \mathrm{cl}(S_k(I) \cap H).$$

By Lemma 7.5, this equality holds if H intersects $S_k(I)$ in its relative interior. Again, by Lemma 7.5, $\mathrm{relint}(\Sigma_1^k(I)^*) \subseteq S_k(I)$. Lemma 7.4 showed that H intersects the relative interior of $\Sigma_1^k(I)^*$ and hence the relative interior of $S_k(I)$. ◻

We now focus on an important situation where the closure is not needed in Theorem 7.6. In many cases in practice, we are interested in finding the convex hull of a set $S \subseteq \mathbb{R}^n$ that may not be presented as the real variety of an ideal. However, the approximation $\mathrm{TH}_k(I)$ of $\mathrm{conv}(S)$ is defined with respect to an ideal I whose real variety is S. In this case, the canonical choice for such an ideal is the *vanishing ideal* of S, denoted as $I(S)$, which consists of all polynomials in $\mathbb{R}[x]$ that vanish on S. The *real radical* of an ideal $I \subseteq \mathbb{R}[x]$ is the ideal

$$\sqrt[\mathbb{R}]{I} = \left\{ f \in \mathbb{R}[x] : f^{2m} + \sum g_i^2 \in I, m \in \mathbb{N}, g_i \in \mathbb{R}[x] \right\},$$

and the ideal I is said to be real radical if $I = \sqrt[\mathbb{R}]{I}$. The *real Nullstellensatz* [21] states that I is real radical if and only if $I = I(V_{\mathbb{R}}(I))$. This is the analogue of Hilbert's Nullstellensatz for real algebraic varieties. Computing any ideal I such that $V_{\mathbb{R}}(I) = S$ might be hard, and in general, computing $I(S)$, given S, might also be hard. However, in many cases of practical interest, $I(S)$ is available. A large source of such examples is combinatorial optimization, where S is usually a finite set of 0/1 points for which a generating set for $I(S)$ can be computed using combinatorial arguments. We will see several such examples in Section 7.4. If S is a subset of $\{0,1\}^n$ and $I = I(S)$, then Theorem 7.6 can be improved to Theorem 7.8. We first prove a lemma.

Lemma 7.7. *Let J be any ideal that contains $x_i^2 - x_i$ for all $i = 1, \ldots, n$. Then $1 + J$ is in the relative interior of* $\Sigma^k(J) = \{f + J \,:\, f \text{ is } k\text{-sos mod } J\}$.

Proof. Let $\mathcal{I} := \langle x_i^2 - x_i \text{ for all } i = 1, \ldots, n\rangle$. We will first show that $1 + \mathcal{I}$ is in the relative interior of $\Sigma^k(\mathcal{I}) \subseteq \mathbb{R}[x]_{2k}/\mathcal{I}$. The cone $\Sigma^k(J)$ is a projection of $\Sigma^k(\mathcal{I})$ since $\mathcal{I} \subseteq J$, and hence, if $1 + \mathcal{I} \in \mathrm{relint}(\Sigma^k(\mathcal{I}))$, then $1 + J \in \mathrm{relint}(\Sigma^k(J))$. $1 + \mathcal{I}$ is in the relative interior of $\Sigma^k(\mathcal{I})$, which is a cone in the vector space $\mathbb{R}[x]_{2k}/\mathcal{I}$.

We will show that for any polynomial $p \in \mathbb{R}[x]_{2k}$, $(1 + \epsilon p) + \mathcal{I} \in \Sigma^k(\mathcal{I})$ for some $\epsilon > 0$. Since we are working modulo $\mathcal{I}$, we may assume that every monomial in p is square-free. Further, since every monomial is a square modulo $\mathcal{I}$, it suffices to show that $(1 - \epsilon q) + \mathcal{I} \in \Sigma^k(\mathcal{I})$ for any square-free monomial q of degree at most $2k$ and some $\epsilon > 0$. Write $q = q_1 q_2$ for some square-free monomials q_1, q_2 of degree at most k. Now note that

$$(1 - q_2)^2 = 1 - 2q_2 + q_2^2 \equiv 1 - q_2 \bmod \mathcal{I}, \text{ and}$$
$$(1 - q_1 + q_2)^2 = 1 + q_1^2 + q_2^2 - 2q_1 + 2q_2 - 2q_1 q_2 \equiv 1 - q_1 + 3q_2 - 2q_1 q_2 \bmod \mathcal{I}.$$

Therefore, $(1-q_1+q_2)^2+3(1-q_2)^2+q_1^2 \equiv 4-2q_1q_2 = 4-2q \mod \mathcal{I}$. Since $q_1, q_2 \in \mathbb{R}[x]_k$, it follows that $(4-2q)+\mathcal{I} \in \Sigma^k(\mathcal{I})$, which implies that $(1-\frac{q}{2})+\mathcal{I} \in \Sigma^k(\mathcal{I})$. ∎

Theorem 7.8. *If $S \subseteq \{0,1\}^n$ and $I = I(S)$, then* $\mathrm{TH}_k(I) = Q_k(I)$.

Proof. Since $S \subseteq \{0,1\}^n$, its vanishing ideal $I = I(S)$ contains $x_i^2 - x_i$ for all $i = 1, \dots, n$, and so by Lemma 7.7, $1 + I$ is in the relative interior of $\Sigma^k(I)$. Hence, $\Sigma_1^k(I)^* = S_k(I)$. (No closure operation is needed by [24, Corollary 16.4.2].) Therefore,

$$\{1\} \times \mathrm{TH}_k(I) = \Sigma_1^k(I)^* \cap H = S_k(I) \cap H = \{1\} \times Q_k(I),$$

and the result follows. ∎

We have thus far seen that the kth theta body $\mathrm{TH}_k(I)$ is the closure of $Q_k(I)$. However, this description is still abstract and in order to work with theta bodies in practice, we now give an explicit (coordinate based) description of $Q_k(I)$ using a basis of $\mathbb{R}[x]/I$ which will make it transparent that $Q_k(I)$ is the projection of a spectrahedron. This involves the theory of *moments* and *moment matrices* as explained below.

Fix a θ-basis $\mathcal{B} = \{f_i + I\}$ of $\mathbb{R}[x]/I$ and define $[x]_{\mathcal{B}_k}$ to be the column vector formed by all the elements of $\mathcal{B}_k$ in order. Then $[x]_{\mathcal{B}_k}[x]_{\mathcal{B}_k}^T$ is a square matrix indexed by $\mathcal{B}_k$ and its (i,j)-entry is equal to $f_if_j + I$. By hypothesis, the entries of $[x]_{\mathcal{B}_k}[x]_{\mathcal{B}_k}^T$ lie in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$. Let $\{\ \lambda_{i,j}^l\ \}$ be the unique set of real numbers such that $f_if_j + I = \sum_{f_l+I\in\mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$.

Definition 7.9. *Let I, $\mathcal{B}$, and $\{\ \lambda_{i,j}^l\ \}$ be as above. Let y be a real vector indexed by $\mathcal{B}_{2k}$ with $y_0 = 1$, where y_0 is the first entry of y, indexed by the basis element $1 + I$. The k*th reduced moment matrix *$M_{\mathcal{B}_k}(y)$ of I is the real matrix indexed by $\mathcal{B}_k$ whose (i,j)-entry is $[M_{\mathcal{B}_k}(y)]_{i,j} = \sum_{f_l+I\in\mathcal{B}_{2k}} \lambda_{i,j}^l y_l$.*

We now give examples of reduced moment matrices. For simplicity, we often write f for $f + I$. Also, in this chapter we consider only monomial bases of $\mathbb{R}[x]/I$ (i.e., f_i is a monomial for all $f_i + I \in \mathcal{B}$) which we can obtain via Gröbner basis theory. In this case, $[x]_{\mathcal{B}_k}$ is a vector of monomials and we identify the vector $[x]_{\mathcal{B}_k}$ with the vector of monomials that represent the elements of $\mathcal{B}_k$. The method is to compute a reduced Gröbner basis of I and take $\mathcal{B}$ to be the equivalence classes of the standard monomials of the corresponding initial ideal. If the reduced Gröbner basis is with respect to a *total degree ordering*, then the second condition in the definition of a θ-basis is satisfied by $\mathcal{B}$.

Example 7.10. Consider the ideal I generated by $f := (x+1)x(x-1)^2$. Clearly, $V_{\mathbb{R}}(I) = \{-1, 0, 1\}$ with a double root at 1, and $\mathrm{conv}(V_{\mathbb{R}}(I)) = [-1, 1]$. The polynomial $f = x^4 - x^3 - x^2 + x$ is the unique element in every reduced Gröbner basis of I with $\langle x^4 \rangle$ as initial ideal. The standard monomials of this initial ideal are

$1, x, x^2, x^3$, and hence $\mathcal{B} = \{1 + I, x + I, x^2 + I, x^3 + I\}$ is a θ-basis for $\mathbb{R}[x]/I$. The biggest reduced moment matrix we could construct is $M_{\mathcal{B}_3}(y)$, whose rows and columns are indexed by $\mathcal{B}_3 = \mathcal{B}$.

We have $[x]_{\mathcal{B}_3} = (1\ x\ x^2\ x^3)$ and

$$[x]_{\mathcal{B}_3}[x]_{\mathcal{B}_3}^T = \begin{bmatrix} 1 & x & x^2 & x^3 \\ x & x^2 & x^3 & x^4 \\ x^2 & x^3 & x^4 & x^5 \\ x^3 & x^4 & x^5 & x^6 \end{bmatrix},$$

which is entrywise equivalent mod I to

$$\begin{bmatrix} 1 & x & x^2 & x^3 \\ x & x^2 & x^3 & x^3 + x^2 - x \\ x^2 & x^3 & x^3 + x^2 - x & 2x^3 - x \\ x^3 & x^3 + x^2 - x & 2x^3 - x & 2x^3 + x^2 - 2x \end{bmatrix}.$$

We now linearize using $y = (1, y_1, y_2, y_3)$ and obtain

$$M_{\mathcal{B}_3}(y) = \begin{bmatrix} 1 & y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 & y_3 + y_2 - y_1 \\ y_2 & y_3 & y_3 + y_2 - y_1 & 2y_3 - y_1 \\ y_3 & y_3 + y_2 - y_1 & 2y_3 - y_1 & 2y_3 + y_2 - 2y_1 \end{bmatrix}.$$

The reduced moment matrices $M_{\mathcal{B}_1}(y)$ and $M_{\mathcal{B}_2}(y)$ are the upper left 2×2 and 3×3 principal submatrices of $M_{\mathcal{B}_3}(y)$. ∎

Example 7.11. Consider the ideal $I = \langle x^4 - y^2 - z^2, x^4 + x^2 + y^2 - 1 \rangle$. Using a computer algebra package such as Macaulay2 [10] one can calculate a total degree reduced Gröbner basis of I as follows:

```
Macaulay2, version 1.3

i1 : R = QQ[x,y,z,Weights => {1,1,1}];
i2 : I = ideal(x^4-y^2-z^2, x^4+x^2+y^2-1);
i3 : G = gens gb I
o3 = | x2+2y2+z2-1 4y4+4y2z2+z4-5y2-3z2+1 |
```

which says that this Gröbner basis consists of the two polynomials

$$x^2 + 2y^2 + z^2 - 1 \quad \text{and} \quad 4y^4 + 4y^2z^2 + z^4 - 5y^2 - 3z^2 + 1.$$

A basis for the quotient ring $\mathbb{R}[x, y, z]/I$ is given by the standard monomials of the initial ideal $\langle x^2, y^4 \rangle$, which gives the following partial bases:

$$\begin{aligned} \mathcal{B}_1 &= \{1, x, y, z\}, \\ \mathcal{B}_2 &= \mathcal{B}_1 \cup \{xy, y^2, xz, yz, z^2\}, \\ \mathcal{B}_3 &= \mathcal{B}_2 \cup \{xy^2, y^3, xyz, y^2z, xz^2, yz^2, z^3\}, \\ \mathcal{B}_4 &= \mathcal{B}_3 \cup \{xy^3, xy^2z, y^3z, xyz^2, y^2z^2, xz^3, yz^3, z^4\}. \end{aligned}$$

Linearizing the elements of $\mathcal{B}_4$, we get the following table:

1	x	y	z	xy	y^2	xz	yz	z^2	xy^2	y^3	xyz	y^2z	xz^2	yz^2	z^3
1	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}

xy^3	xy^2z	y^3z	xyz^2	y^2z^2	xz^3	yz^3	z^4
y_{16}	y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}	y_{23}.

We can now calculate various reduced moment matrices. For instance,

$$M_{\mathcal{B}_2}(y) = \begin{bmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \\ & T_1 & y_4 & y_6 & T_2 & y_9 & T_3 & y_{11} & y_{13} \\ & & y_5 & y_7 & y_9 & y_{10} & y_{11} & y_{12} & y_{14} \\ & & & y_8 & y_{11} & y_{12} & y_{13} & y_{14} & y_{15} \\ & & & & T_4 & y_{16} & T_5 & y_{17} & y_{19} \\ & & & & & T_6 & y_{17} & y_{18} & y_{20} \\ & & & & & & T_7 & y_{19} & y_{21} \\ & & & & & & & y_{20} & y_{22} \\ & & & & & & & & y_{23} \end{bmatrix},$$

where we have filled in only the upper triangular region. The unknowns $T_1, T_2, \ldots$ stand for the following expressions:

$$\begin{aligned} T_1 &= -2y_5 - y_8 + 1, \\ T_2 &= -2y_{10} - y_{14} + y_2, \\ T_3 &= -2y_{12} - y_{15} + y_3, \\ T_4 &= y_{20} + \tfrac{y_{23}}{2} - \tfrac{3y_5}{2} - \tfrac{3y_8}{2} + \tfrac{1}{2}, \\ T_5 &= -2y_{18} - y_{22} + 1, \\ T_6 &= -y_{20} - \tfrac{y_{23}}{4} + \tfrac{5y_5}{4} + \tfrac{3y_8}{4} - \tfrac{1}{4}, \\ T_7 &= -2y_{20} - y_{23} + y_8. \end{aligned}$$

The T_i's can be calculated using Macaulay2 by first finding the normal form of the needed monomial with respect to the Gröbner basis that was calculated and then linearizing using the y_i's. For instance, T_2 is the linearization of the normal form of x^2y, which by the calculation below, is $-2y^3 - yz^2 + y$.

```
i6 : x^2*y%G
          3      2
o6 = - 2y  - y*z  + y     █
```

The reduced moment matrix $M_{\mathcal{B}_k}(y)$ can also be defined in terms of linear functionals on $\mathbb{R}[x]_{2k}/I$. For a vector $y = (y_b) \in \mathbb{R}^{\mathcal{B}_{2k}}$, define $L_y \in (\mathbb{R}[x]_{2k}/I)^*$ as $L_y(b) := y_b$ for all $b \in \mathcal{B}_{2k}$. Then every $L \in (\mathbb{R}[x]_{2k}/I)^*$ is equal to L_y for $y = (L(b) : b \in \mathcal{B}_{2k}) \in \mathbb{R}^{\mathcal{B}_{2k}}$. If $y \in \mathbb{R}^{\mathcal{B}_{2k}}$, let $y_0 := y_{1+I}$, $y_i := y_{x_i+I}$ for $i = 1, \ldots, n$. Further, let $\pi_{\mathbb{R}^n}$ be the projection map that sends $y \in \mathbb{R}^{\mathcal{B}_{2k}}$ to $(y_1, \ldots, y_n) \in \mathbb{R}^n$.

Lemma 7.12.

1. *For a vector* $y \in \mathbb{R}^{\mathcal{B}_{2k}}$ *with* $y_0 = 1$, *the entry of* $M_{\mathcal{B}_k}(y)$ *indexed by* $b_i, b_j \in \mathcal{B}_k$ *is* $L_y(b_i b_j)$.
2. $M_{\mathcal{B}_k}(y) \succeq 0 \Leftrightarrow L_y(f^2 + I) \geq 0$ *for all* $f + I \in \mathbb{R}[x]_k/I$.

Proof. The first part follows from the definition of $M_{\mathcal{B}_k}(y)$ and L_y. For $f + I \in \mathbb{R}[x]_k/I$, let $\hat{f}$ be the unique vector in $\mathbb{R}^{\mathcal{B}_k}$ such that $f + I = \sum_{b_i \in \mathcal{B}_k} \hat{f}_i b_i$. Therefore, $f^2 + I = \sum_{b_i, b_j \in \mathcal{B}_k} \hat{f}_i \hat{f}_j (b_i b_j)$ which implies that

$$L_y(f^2 + I) = \sum_{b_i, b_j \in \mathcal{B}_k} \hat{f}_i \hat{f}_j L_y(b_i b_j) = \hat{f}^T M_{\mathcal{B}_k}(y) \hat{f}.$$

Therefore, $M_{\mathcal{B}_k}(y) \succeq 0 \Leftrightarrow L_y(f^2 + I) \geq 0$ for all $f + I \in \mathbb{R}[x]_k/I$. ☐

Putting all this together, we obtain the following specific semidefinite representation of $Q_k(I)$, and hence $\mathrm{TH}_k(I)$ up to closure. We will use this explicit coordinate based description of $\mathrm{TH}_k(I)$ in the the calculations below.

Theorem 7.13. *The kth theta body of* I, $\mathrm{TH}_k(I)$, *is the closure of*

$$Q_k(I) = \pi_{\mathbb{R}^n} \left\{ y \in \mathbb{R}^{\mathcal{B}_{2k}} \,:\, M_{\mathcal{B}_k}(y) \succeq 0,\ y_0 = 1 \right\}.$$

Proof. Recall that $Q_k(I)$ is the set

$$\left\{ (L(x_1 + I), \ldots, L(x_n + I)) \,:\, \begin{array}{l} L(g + I) \geq 0 \text{ for all } g + I \in \Sigma_{2k}/I, \\ L(1 + I) = 1 \end{array} \right\}.$$

Equivalently, $Q_k(I)$ is the set

$$\left\{ (L(b) \,:\, b \in \mathcal{B}_1 \backslash \{1 + I\}) \,:\, \begin{array}{l} L(f^2 + I) \geq 0 \text{ for all } f + I \in \mathbb{R}[x]_k/I, \\ L(1 + I) = 1 \end{array} \right\}.$$

By Lemma 7.12 (2), it then follows that

$$Q_k(I) = \pi_{\mathbb{R}^n} \left\{ y \in \mathbb{R}^{\mathcal{B}_{2k}} \,:\, M_{\mathcal{B}_k}(y) \succeq 0,\ y_0 = 1 \right\} =: Q_{\mathcal{B}_k}(I). \quad ☐$$

When working with a specific basis $\mathcal{B}$, we use $Q_{\mathcal{B}_k}(I)$ instead of $Q_k(I)$ to make the choice of basis clear. In the examples that follow, please bear in mind that this abuse of notation is simply to keep track of which θ-basis of $\mathbb{R}[x]/I$ was used in the explicit semidefinite representation of $Q_k(I)$. The proof of Theorem 7.13 shows that any θ-basis of $\mathbb{R}[x]/I$ can be used to coordinatize $Q_k(I)$.

Example 7.10 continued. We write down $Q_{\mathcal{B}_k}(I)$ for $k = 1, 2, 3$ for the ideal $I = \langle (x+1)x(x-1)^2 \rangle$ from Example 7.10. Using the matrix $M_{\mathcal{B}_3}(y)$ (with $y_0 = 1$) that was already computed we see that

$$Q_{\mathcal{B}_1}(I) = \{ y_1 \,:\, \exists (y_1, y_2) \in \mathbb{R}^2 \text{ s.t. } y_2 \geq y_1^2 \},$$

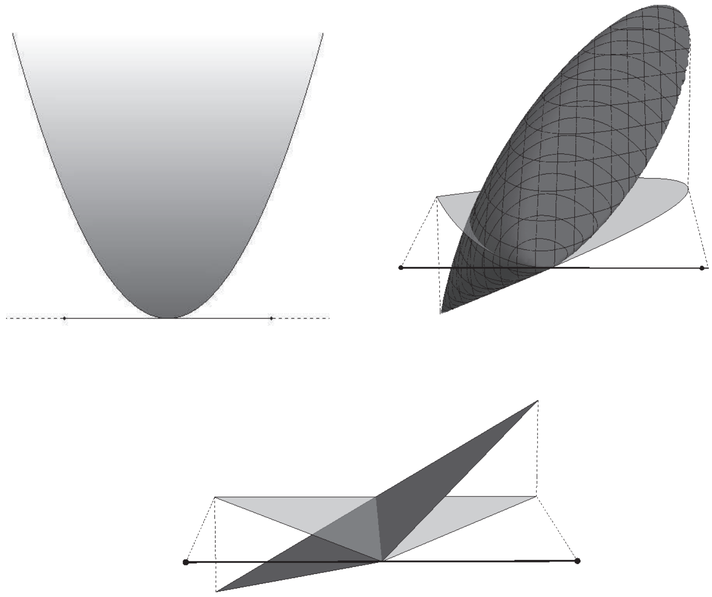

Figure 7.3. *The spectrahedra* $\{y \in \mathbb{R}^{\mathcal{B}_{2k}} : y_0 = 1, M_{\mathcal{B}_k}(y) \succeq 0\}$ *for* $k = 1, 2, 3$ *for* $I = \langle (x+1)x(x-1)^2 \rangle$ *and their projections to the* y_1*-axis.*

which is the projection onto the y_1-axis of the convex hull of the parabola $y_2 = y_1^2$. Therefore, $Q_{\mathcal{B}_1}(I) = \mathbb{R}$ and hence $\mathrm{TH}_1(I) = \mathbb{R}$, which is a trivial relaxation of $\mathrm{conv}(V_{\mathbb{R}}(I)) = [-1, 1]$.

The body $Q_{\mathcal{B}_2}(I) = \{y_1 : \exists y \in \mathbb{R}^3 \text{ s.t. } M_{\mathcal{B}_2}(y) \succeq 0\}$. We know the exact form of the moment matrices so we can use YALMIP to find $\mathrm{cl}(Q_{\mathcal{B}_2}(I))$, by minimizing x and $-x$ over that body.

```
sdpvar y1 y2 y3
M=[1  y1 y2;
   y1 y2 y3;
   y2 y3 y3+y2-y1];
solvesdp(M>0,y1);
double(y1)
solvesdp(M>0,-y1);
double(y1)
```

We then get $\mathrm{cl}(Q_{\mathcal{B}_2}(I)) \approx [-1.0000, 1.0417]$, and we will later see that it is actually exactly $[-1, \frac{25}{24}]$.

To finish, we compute $Q_{\mathcal{B}_3}(I) = \{y_1 : \exists y \in \mathbb{R}^3 \text{ s.t. } M_{\mathcal{B}_3}(y) \succeq 0\}$. This is the projection onto the y_1-coordinate of the spectrahedron in $\mathbb{R}^3$ described by all the

Figure 7.4. *The variety of Example* 7.11 *and its first theta body.*

Figure 7.5. *The second theta body from Example* 7.11.

inequalities obtained from the condition $M_{\mathcal{B}_3}(y) \succeq 0$. This body is the convex hull of the moment vectors (x, x^2, x^3) evaluated at $x = -1, 0, 1$, which is the triangle with vertices $(-1, 1, -1), (0, 0, 0), (1, 1, 1)$. Projecting onto the y_1-coordinate, we get $\text{cl}(Q_{\mathcal{B}_3}(I)) = [-1, 1]$. See Figure 7.3 for $Q_{\mathcal{B}_i}(I)$, $i = 1, 2, 3$, and their spectrahedral preimages.

Example 7.11 continued. We now draw a few theta bodies of the ideal

$$I = \langle x^4 - y^2 - z^2, x^4 + x^2 + y^2 - 1 \rangle$$

from Example 7.11, where we calculated the second reduced moment matrix $M_{\mathcal{B}_2}(y)$. This allows us to write down $Q_{\mathcal{B}_1}(I)$ and $Q_{\mathcal{B}_2}(I)$.

From the Gröbner basis of I that we computed, we see that the polynomial $x^2 + 2y^2 + z^2 - 1$ is in I. We will see in Example 7.36 that the first theta body of I is the ellipsoid $\{(x, y, z) \in \mathbb{R}^3 : x^2 + 2y^2 + z^2 \leq 1\}$. This ellipsoid along with $V_{\mathbb{R}}(I)$ (the two black rings) is shown in Figure 7.4. The second theta body is shown in Figure 7.5 and it appears to equal $\text{conv}(V_{\mathbb{R}}(I))$.

Remark 7.14. *This example shows the difference between Lasserre's method to convexify $V_{\mathbb{R}}(I)$ and the reduced moment method that underlies theta bodies. Recall that in step k of Lasserre's method, the relaxation of* $\text{conv}(V_{\mathbb{R}}(I))$ *that is computed is the common intersection of all half spaces $l(x) \geq 0$ containing $V_{\mathbb{R}}(I)$ and*

Figure 7.6. *The second Lasserre relaxation for Example* 7.11.

$l(x) = \sigma(x) + \sum_{i=1}^{m} g_i(x) f_i(x)$, where $\sigma(x)$ is a k-sos polynomial and $g_i(x)f_i(x) \in \mathbb{R}[x]_{2k}$. Using the software package Bermeja [25] *we can draw the second relaxation in Lasserre's method which is shown in Figure* 7.6.

Now that we have seen several examples of theta bodies of ideals, we give a few comments and examples to point out some of the subtleties involved. We start with an example to show that $Q_{\mathcal{B}_k}(I)$ may not be closed, which emphasizes the need to take its closure to get $\mathrm{TH}_k(I)$.

Example 7.15. Consider the principal ideal $I = \langle x_1^2 x_2 - 1 \rangle \subset \mathbb{R}[x_1, x_2]$. Then $\mathrm{conv}(V_{\mathbb{R}}(I)) = \{(s_1, s_2) \in \mathbb{R}^2 \,:\, s_2 > 0\}$, which is not a closed set. Any linear polynomial that is nonnegative over $V_{\mathbb{R}}(I)$ is of the form $\alpha x_2 + \beta$, where $\alpha, \beta \geq 0$. Since $\alpha x_2 + \beta \equiv (\sqrt{\alpha} x_1 x_2)^2 + (\sqrt{\beta})^2 \bmod I$, $\mathrm{TH}_2(I) = \mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$.

The set $\mathcal{B} = \bigcup_{k \in \mathbb{N}} \{x_1^k + I, x_2^k + I, x_1 x_2^k + I\}$ is a θ-basis for $\mathbb{R}[x_1, x_2]/I$ for which

$$\mathcal{B}_4 = \{1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1x_2^2, x_1^3, x_2^3, x_1x_2^3, x_1^4, x_2^4\} + I.$$

The reduced moment matrix $M_{\mathcal{B}_2}(y)$ for $y = (1, y_1, \ldots, y_{11}) \in \mathbb{R}^{\mathcal{B}_4}$ is

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{array}
\begin{array}{c} \begin{array}{cccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \end{array} \\
\left(\begin{array}{cccccc}
1 & y_1 & y_2 & y_3 & y_4 & y_5 \\
y_1 & y_3 & y_4 & y_6 & 1 & y_7 \\
y_2 & y_4 & y_5 & 1 & y_7 & y_8 \\
y_3 & y_6 & 1 & y_9 & y_1 & y_2 \\
y_4 & 1 & y_7 & y_1 & y_2 & y_{10} \\
y_5 & y_7 & y_8 & y_2 & y_{10} & y_{11}
\end{array} \right) \end{array} .$$

If $M_{\mathcal{B}_2}(y) \succeq 0$, then the principal minor indexed by x_1 and x_1x_2 implies that $y_2 y_3 \geq 1$, and so in particular, $y_2 \neq 0$ for all $y \in Q_{\mathcal{B}_2}(I)$. However, since $Q_{\mathcal{B}_2}(I) \supseteq \mathrm{conv}(V_{\mathbb{R}}(I)) = \{(s_1, s_2) \in \mathbb{R}^2 \,:\, s_2 > 0\}$, it must be that $Q_{\mathcal{B}_2}(I) = \mathrm{conv}(V_{\mathbb{R}}(I))$, which shows that $Q_{\mathcal{B}_2}(I)$ is not closed. ∎

We will see in the next section that when S is a finite set of points in $\mathbb{R}^n$, the ideal $I = I(S)$ of all polynomials that vanish on S, has the property that

$\mathrm{TH}_l(I) = \mathrm{conv}(V_{\mathbb{R}}(I)) = \mathrm{conv}(S)$ for a finite l that depends on I. However, since $\mathrm{conv}(S) \subseteq Q_{\mathcal{B}_l}(I) \subseteq \mathrm{TH}_l(I)$, we also get that $Q_{\mathcal{B}_l}(I)$ is closed. Even in this case, $Q_{\mathcal{B}_k}(I)$ may not be closed for some $k < l$.

Example 7.16. Consider the finite set of points $S = \{(\pm t, 1/t^2) : t = 1, \ldots, 7\}$ lying on the curve $x_1^2 x_2 = 1$. Then

$$I(S) = \langle x_1^2 x_2 - 1, (x_1^2 - 1)(x_1^2 - 4)(x_1^2 - 9)(x_1^2 - 16)(x_1^2 - 25)(x_1^2 - 36)(x_1^2 - 49)\rangle.$$

This is a zero-dimensional ideal, and a basis for $\mathbb{R}[x_1, x_2]/I(S)$ is given by

$$\mathcal{B} = \{1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1x_2^2, x_1^3, x_2^3, x_1x_2^3, x_1^4, x_2^4, x_1^5, x_1x_2^4\} + I.$$

In particular, $\mathcal{B}_4$ is the same as the $\mathcal{B}_4$ in Example 7.15 and the initial ideal of $I(S)$ whose standard monomials are the monomials in $\mathcal{B}$ is generated by $\{x_1^2 x_2, x_2^5, x_1^6\}$. Therefore, $M_{\mathcal{B}_2}(I(S))$ and $Q_{\mathcal{B}_2}(I(S))$ agree with those in Example 7.15, which implies that $Q_{\mathcal{B}_2}(I(S))$ is not closed. ∎

Another natural question is whether the theta bodies of different ideals with the same real variety can have drastically different behaviors, especially with respect to convergence to the convex hull of the variety. For instance, an ideal I and its real radical $\sqrt[\mathbb{R}]{I}$ have the same real variety and $I \subseteq \sqrt[\mathbb{R}]{I}$, $\mathrm{TH}_k(\sqrt[\mathbb{R}]{I}) \subseteq \mathrm{TH}_k(I)$ for all k.

Theorem 7.17. *Fix an ideal I. Then there exists a function $\Psi : \mathbb{N} \to \mathbb{N}$ such that $\mathrm{TH}_{\Psi(k)}(I) \subseteq \mathrm{TH}_k(\sqrt[\mathbb{R}]{I})$ for all k.*

We refer the reader to [9, Section 2.2] for a proof. The main message to take away from this result is that whether or not the theta body hierarchy of an ideal converges to $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$ is determined by the real variety of I. In particular, whether the theta body sequence of an ideal converges to $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$ in finitely many steps, or not, is determined by $\sqrt[\mathbb{R}]{I}$.

7.2.3 Possible Extensions

The focus of this chapter is on polynomial equations, and sums of squares relaxations. However, all this theory can potentially be adapted to work in some more complicated cases. In this section we give examples of some constructions that give a flavor of possible extensions. Similar constructions were also seen in Chapter 6, and we refer to [22] for a more systematic study of the types of techniques we will see below (in a slightly different setting).

Example 7.18. The theta body sequence can be modified to deal with polynomial inequalities, using Lasserre's ideas. Given an ideal I and some polynomials $g_1, \ldots, g_t$, we might want to find the convex hull of the semialgebraic set $S = \{x \in V_{\mathbb{R}}(I) : g_1(x) \geq 0, \ldots, g_t(x) \geq 0\}$. To do this we use *shifted reduced moment matrices* in addition to the reduced moment matrices of I.

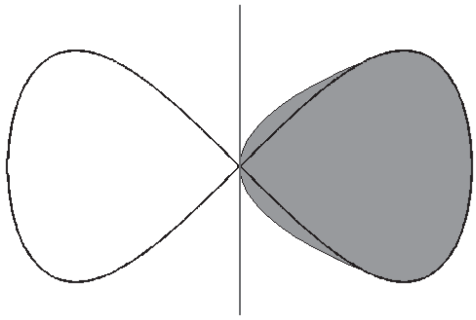

Figure 7.7. *Sum of squares approximation to the half-lemniscate of Gerono.*

Recall that to obtain the kth reduced moment matrix $M_{\mathcal{B}_k}(y)$ of I, we would take the matrix $[x]_{\mathcal{B}_k}[x]_{\mathcal{B}_k}^T$, write it in terms of a basis $\mathcal{B}$ of $\mathbb{R}[x]/I$, and linearize using the new variables y with $y_0 = 1$. To define the shifted reduced moment matrix $M_{\mathcal{B}_k}(g * y)$ (with respect to g), we take the matrix $g(x)[x]_{\mathcal{B}_k}[x]_{\mathcal{B}_k}^T$ and do precisely as before.

Consider for example the ideal $I = \langle x^4 - x^2 + y^2 \rangle$ of the *lemniscate of Gerono*, together with the inequality $x \geq 0$. The semialgebraic set S in this case is the right half-lemniscate shown in Figure 7.7. The second reduced moment matrix of I is given by

$$\begin{pmatrix} 1 & x & y & w_2^0 & w_1^1 & w_0^2 \\ x & w_2^0 & w_1^1 & w_3^0 & w_2^1 & w_1^2 \\ y & w_1^1 & w_0^2 & w_2^1 & w_1^2 & w_0^3 \\ w_2^0 & w_3^0 & w_2^1 & w_2^0 - w_0^2 & w_3^1 & w_2^2 \\ w_1^1 & w_2^1 & w_1^2 & w_3^1 & w_2^2 & w_1^3 \\ w_0^2 & w_1^2 & w_0^3 & w_2^2 & w_1^3 & w_0^4 \end{pmatrix},$$

where w_i^j is the linearization of $x^i y^j$. The combinatorial moment matrix shifted by x and truncated at $k = 1$ is

$$\begin{pmatrix} x & w_2^0 & w_1^1 \\ w_2^0 & w_3^0 & w_2^1 \\ w_1^1 & w_2^1 & w_1^2 \end{pmatrix}.$$

If we force both matrices to be positive semidefinite and project over the x, y coordinates, we get an approximation of the convex hull of the right half of the lemniscate, as shown in Figure 7.7. By increasing the truncation parameter of the reduced moment matrix and the shifted moment matrix we get better approximations to the convex hull.

Note that in this example we are essentially searching for certificates of nonnegativity of the form $l(x, y) \equiv \sigma_0(x, y) + x\sigma_1(x, y) \mod I$, where σ_0 and σ_1 are 2-sos and 1-sos, respectively. ∎

Example 7.19. Consider the *teardrop curve* given by $p(x, y) := x^4 - x^3 + y^2 = 0$. We will see in Corollary 7.45 that the singularity at the origin will prevent the theta bodies of $\langle p \rangle$ from converging in a finite number of steps to the convex hull of the curve. We can, however, get rid of that problem by strengthening the hierarchy in a simple way. Recall that the second theta body in this case will be obtained as the closure of the set of all points $(x, y) \in \mathbb{R}^2$ for which there exists a positive definite matrix of the form

$$\begin{pmatrix} 1 & x & y & w_2^0 & w_1^1 & w_0^2 \\ x & w_2^0 & w_1^1 & w_3^0 & w_2^1 & w_1^2 \\ y & w_1^1 & w_0^2 & w_2^1 & w_1^2 & w_0^3 \\ w_2^0 & w_3^0 & w_2^1 & w_3^0 - w_0^2 & w_3^1 & w_2^2 \\ w_1^1 & w_2^1 & w_1^2 & w_3^1 & w_2^2 & w_1^3 \\ w_0^2 & w_1^2 & w_0^3 & w_2^2 & w_1^3 & w_0^4 \end{pmatrix},$$

where w_i^j is a variable that linearizes the monomial $x^i y^j$, and so the rows and columns are indexed by $\{1, x, y, x^2, xy, y^2\}$. One can in this case strengthen the condition by adding a new row and column to the matrix, indexed not by a monomial but by the fraction $\frac{y}{x}$ that we linearize as w_{-1}^1. We then use the same strategy as before, of linearizing all resulting products modulo the relation $x^4 = x^3 - y^2$ (which allows us to get rid of $w_{4,0}$) and the relations $\frac{y^2}{x} = x^2 - x^3$ and $\frac{y^2}{x^2} = x - x^2$ (which eliminates two more variables). This new *pseudomoment matrix* is given by

$$M(x, y, w) = \begin{pmatrix} 1 & x & y & w_2^0 & w_1^1 & w_0^2 & w_{-1}^1 \\ x & w_2^0 & w_1^1 & w_3^0 & w_2^1 & w_1^2 & y \\ y & w_1^1 & w_0^2 & w_2^1 & w_1^2 & w_0^3 & w_2^0 - w_3^0 \\ w_2^0 & w_3^0 & w_2^1 & w_3^0 - w_0^2 & w_3^1 & w_2^2 & w_1^1 \\ w_1^1 & w_2^1 & w_1^2 & w_3^1 & w_2^2 & w_1^3 & w_0^2 \\ w_0^2 & w_1^2 & w_0^3 & w_2^2 & w_1^3 & w_0^4 & w_{-1}^3 \\ w_{-1}^1 & y & w_2^0 - w_3^0 & w_1^1 & w_0^2 & w_{-1}^3 & x - w_2^0 \end{pmatrix}.$$

Since the original moment matrix is a submatrix of $M(x, y, w)$, the body $Q = \{(x, y) \ : \ \exists w \text{ s.t. } M(x, y, w) \succeq 0\}$ must be contained in $\mathrm{TH}_2(\langle p \rangle)$, and a simple numeric computation seems to show that Q actually matches the convex hull of the real variety $V_{\mathbb{R}}(p)$, as we can see in Figure 7.8. In this figure we see a comparison of the second theta body and Q, drawn numerically using YALMIP. The fact that Q seems to be exact is related to the fact that we can now use the term $\frac{x}{y}$ to get sos certificates. For example, $x = x^2 + (\frac{x}{y})^2$ modulo the new identities that we introduced. ∎

Exercise 7.20. Let $I = \langle x^2 \rangle$.

1. Show that x is not k-sos mod I for any k.
2. Show that for any $\varepsilon > 0$, the polynomial $x + \varepsilon$ is 1-sos mod I.
3. Describe $\mathrm{TH}_1(I)$.

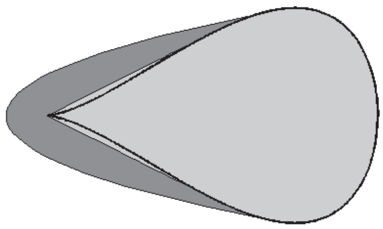

Figure 7.8. *In the darker color we see* $\mathrm{TH}_2(\langle p \rangle)$*, while in the lighter color we see the strengthening* Q *as defined in Example* 7.19*. In black we see the variety itself.*

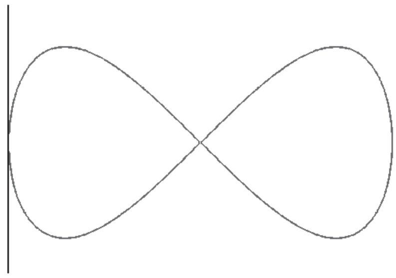

Figure 7.9. *Lemniscate of Gerono.*

Exercise 7.21. Using YALMIP or other software, find the smallest ϵ such that $x + \epsilon$ is 2-sos modulo the ideal $I = \langle x^4 - x^3 + y^2 \rangle$. What about 3-sos? What about 4-sos?

Exercise 7.22. The lemniscate of Gerono is given by the equation $x^4 - x^2 + y^2 = 0$ shown in Figure 7.9. Using YALMIP give an approximate 2-sos decomposition of $x + 1$ modulo the equation of the curve. Can you find an exact one?

Exercise 7.23. Using reduced moment matrices, give semidefinite descriptions of the following bodies:

1. $Q_{\mathcal{B}_2}(I)$ for the ideal of the lemniscate of Gerono.

2. $Q_{\mathcal{B}_1}(I)$ and $Q_{\mathcal{B}_2}(I)$ where $I = \langle y^2 - x - 1, x^2 - y - 1 \rangle$.

3. $Q_{\mathcal{B}_1}(I)$ where I is the vanishing ideal of the vertices of the 0/1 cube in $\mathbb{R}^3$.

Exercise 7.24. Let I be the vanishing ideal of a finite set of points in $\mathbb{R}^n$.

1. Prove that $p(x)$ is nonnegative over $V_{\mathbb{R}}(I)$ if and only if it is a sum of squares modulo the ideal I.

2. Using the above fact, prove that for $\mathcal{B}$, a θ-basis of $\mathbb{R}[x]/I$, the spectrahedron $\{y \in \mathbb{R}^{\mathcal{B}} : M_{\mathcal{B}}(y) \succeq 0, y_0 = 1\}$ is the simplex whose vertices are the vectors $(f_i(s) : f_i + I \in \mathcal{B})$ as s varies over the finitely many points in $V_{\mathbb{R}}(I)$.

7.3 Convergence of Theta Bodies

One of the main questions after defining a sequence of approximations to a convex set is if they actually approximate the set, and further, if some approximation in the sequence is guaranteed to coincide with the set. In this section we examine conditions under which the sequence of theta bodies of an ideal I converges, either finitely or asymptotically, to $\operatorname{conv}(V_{\mathbb{R}}(I))$.

Definition 7.25. *Let $I \subset \mathbb{R}[x]$ be an ideal.*

1. *The theta body sequence of I converges to* $\operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$ *if*

$$\bigcap_{k=1}^{\infty} \mathrm{TH}_k(I) = \operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I))).$$

2. *For a finite integer k, the ideal I is* TH_k-exact *if* $\mathrm{TH}_k(I) = \operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$.

3. *If I is TH_k-exact for a finite integer k, then we say that the theta body sequence of I converges to* $\operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$ *in finitely many steps. If the theta body sequence of I converges to* $\operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$ *but there is no finite k for which I is TH_k-exact, then we say that the theta body sequence of I converges asymptotically to* $\operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$.

We will see in Section 7.3.1 that if $V_{\mathbb{R}}(I)$ is finite, then there is always some finite k for which I is TH_k-exact. However, tight bounds on k for which I is TH_k-exact are not known in general. The best scenario is when I is TH_1-exact. We characterize finite varieties whose real radical ideal is TH_1-exact. Recall from the discussion following Theorem 7.17 that there is no loss of generality in passing to the real radical of I in discussing convergence.

When $V_{\mathbb{R}}(I)$ is infinite, much less is understood about the convergence of the theta body sequence of I. In Section 7.3.2 we explain what we know about this case. The best general result is that when $V_{\mathbb{R}}(I)$ is compact, the theta body sequence is guaranteed to converge to $\operatorname{cl}(\operatorname{conv}(V_{\mathbb{R}}(I)))$ asymptotically. However, finite convergence, and even convergence in the first step are sometimes possible for infinite varieties, although no characterization is known in either case. We show that certain singularities can prevent finite convergence when the variety is compact.

7.3.1 Finite Real Varieties

Theorem 7.26. *Let I be an ideal such that $V_{\mathbb{R}}(I)$ is finite; then there exists some k such that* $\mathrm{TH}_k(I) = \mathrm{conv}(V_{\mathbb{R}}(I))$.

Proof. First note that by Theorem 7.17 we just need to prove the existence of such a k for $J = \sqrt[\mathbb{R}]{I}$. Let $V_{\mathbb{R}}(I) := \{P_1, \dots, P_m\} \subset \mathbb{R}^n$ and, for each P_i, let q_i be a polynomial such that $q_i(P_i) = 1$ and $q_i(P_j) = 0$ for $j \neq i$. Then given any polynomial $f(x)$ that is nonnegative on $V_{\mathbb{R}}(I)$ we have that

$$f(x) - \sum_{j=1}^{m} \left(\sqrt{f(P_j)} q_j(x) \right)^2$$

vanishes at all P_i, and hence it belongs to J, and f is sos modulo J. So all nonnegative polynomials on $V_{\mathbb{R}}(J)$ are sos modulo J, which in particular implies that each of them is nonnegative over some $\mathrm{TH}_k(J)$. Since the convex hull of $V_{\mathbb{R}}(I)$ is a polytope, it is cut out by a finite number of linear inequalities. Pick k large enough for all these linear inequalities to be valid on $\mathrm{TH}_k(J)$ simultaneously. Then $\mathrm{conv}(V_{\mathbb{R}}(I)) = \mathrm{TH}_k(J)$. □

Clearly, Theorem 7.26 implies that when $V_{\mathbb{C}}(I)$ is finite, the ideal I is TH_k-exact for some finite k. When the ideal I is also radical, finite convergence of its theta body sequence to the convex hull of the variety was proved by Parrilo (see Theorem 2.4 in [16]). Having established finite convergence of the theta body sequence of I when $V_{\mathbb{R}}(I)$ is finite, one can ask the more ambitious question of when such an I is TH_1-exact. This is the most useful and computationally practical case of finite convergence. If the ideal defining a finite set of points is always assumed to be the vanishing ideal of the variety (and hence real radical), we can give a complete geometric characterization of when they are TH_1-exact. We will need the following fact about real radical ideals.

Lemma 7.27 ([8]). *If $I \subset \mathbb{R}[x]$ is a real radical ideal, then a linear inequality $l(x) \geq 0$ is valid for $\mathrm{TH}_k(I)$ if and only if $l(x)$ is k-sos modulo I.*

In order to characterize real radical ideals with finite real varieties, we need a new definition.

Definition 7.28. *Given a polytope P, we say that P is* 2-level *if for each facet F of P and its affine span H_F, all vertices of P are either in F or in a unique translate of H_F.*

Example 7.29. In $\mathbb{R}^3$, up to affine equivalence there are five three-dimensional 2-level polytopes, shown in the upper part of Figure 7.10. It is easy to see that a 2-level polytope must be affinely equivalent to a 0/1-polytope. In the bottom of Figure 7.10 we show the three remaining 0/1-polytopes (up to affine equivalence) with a face that fails to verify the 2-level condition highlighted. ■

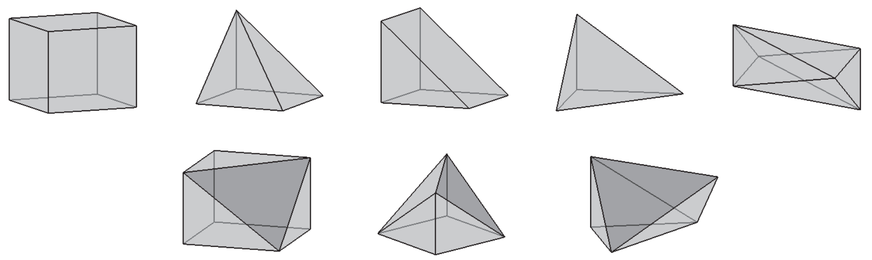

Figure 7.10. *The top row contains all* $0/1$ *three-dimensional* 2*-level polytopes (up to affine equivalence). The bottom row contains all* $0/1$ *three-dimensional polytopes (up to affine equivalence) that are not* 2*-level.*

Theorem 7.30. *Let* I *be real radical with* $S := V_{\mathbb{R}}(I)$ *finite. Then* I *is* TH_1*-exact if and only if* S *is the set of vertices of a* 2*-level polytope.*

Proof. Assume without loss of generality that S spans the entire space and let $f_1(x) \geq 0, \ldots, f_m(x) \geq 0$ be a minimal list of linear inequalities describing $P := \mathrm{conv}(S)$, i.e., each f_i corresponds to a facet F_i of P and is zero on that facet. By Lemma 7.27, I is TH_1-exact if and only if all f_i are 1-sos mod I, since every affine linear polynomial that is nonnegative on S is a nonnegative linear combination of the f_i's.

If I is TH_1-exact, for each $i = 1, \ldots, m$, we have $f_i(x) \equiv \sum (h_k(x))^2 \mod I$, where all h_k are linear. But since f_i vanishes on $S \cap F_i$ so must all h_k and therefore, since they are linear, they must vanish on the affine space generated by F_i. This means that they are actually just scalar multiples of f_i and we have $f_i(x) \equiv \lambda (f_i(x))^2 \mod I$, for some nonnegative λ. In particular, all points $P \in S$ must satisfy either $f_i(P) = 0$ or $f_i(P) = 1/\lambda$ proving the 2-level condition.

Suppose now that P is 2-level. Then for each f_i, all points $P \in S$ must satisfy $f_i(P) = 0$ or $f_i(P) = \lambda_i$, for some fixed $\lambda_i > 0$. But then $f_i(f_i - \lambda_i)$ vanishes on S, and therefore belongs to I. This implies $f_i \equiv (1/\lambda_i) f_i^2 \mod I$ and f_i is 1-sos modulo I. ∎

Theorem 7.30 will turn out to be very useful in the context of combinatorial optimization as we will see in the next section. Polytopes with integer vertices that are 2-level are called *compressed polytopes* in the literature [34, 35] and play an important role in other research areas. Being 2-level is a highly restrictive condition that immediately gives us much information on the polytope. Since all the vertices of a 2-level polytope in $\mathbb{R}^n$ can be assumed to be $0/1$ vectors, it is clear that they have at most 2^n vertices. It was shown in [8] that they also have at most 2^n facets which is not obvious. There are many infinite families of 2-level polytopes such as simplices, hypercubes, cross polytopes, and hypersimplices.

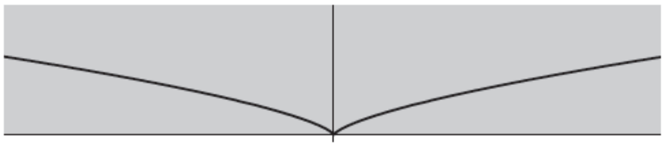

Figure 7.11. *Cusp and its convex hull.*

7.3.2 Infinite Real Varieties

We begin by showing that unlike for finite varieties, the theta body approximations can fail drastically when $V_{\mathbb{R}}(I)$ is infinite. The following simple example is adapted from Example 1.3.2 in [21].

Example 7.31. Consider the ideal $I = \langle x^2 - y^3 \rangle$ defining the cusp in Figure 7.11. The closure of the convex hull of this curve is the upper half-plane, so the only linear inequalities valid on the curve are of the form $l_\varepsilon(x, y) = y + \varepsilon$, where $\varepsilon \geq 0$. Suppose there exists some l_ε with an sos certificate modulo I, then $l_\varepsilon(x, y) \equiv \sum p_i(x, y)^2 \mod I$ for some polynomials p_i. Note that any polynomial p has a unique standard form of the type $a(y) + xb(y)$ modulo this ideal, which we can obtain by reducing all multiples of x^2, using the fact that $x^2 \equiv y^3 \mod I$. Two polynomials are the same modulo the ideal if they have the same standard form. Since $l_\varepsilon(x, y)$ is already in this form, we can simply reduce the right-hand side in the congruence relation to its standard form too. Suppose each $p_i = a_i(y) + xb_i(y)$. Then it is easy to check that

$$\sum p_i(x, y)^2 \equiv \sum (a_i(y)^2 + y^3 b_i(y)^2) + \sum (2x a_i(y) b_i(y)) \mod I.$$

Since the right-hand side is in standard form, to be congruent to l_ε it must be the same as l_ε. Looking at the maximum degree of y in the first sum on the right, we see that it is smaller than two only if the a_i's are all constants and the b_i's are all zero, since the highest degree terms cannot all cancel. In particular we get $y + \varepsilon$ is a constant, which is clearly a contradiction. This proves that $\mathrm{TH}_k(I) = \mathbb{R}^2$ for all k, and the theta bodies are completely ineffective in approximating $\mathrm{conv}(V_{\mathbb{R}}(I))$. In fact, the same proof would work for any curve of the form $x^2 - p(y)$ where p has odd degree. ∎

However, despite the existence of "badly behaved" varieties such as the one presented above, there is a large, very interesting class of infinite real varieties where such behavior never occurs, namely, compact varieties.

Theorem 7.32. *Let I be an ideal such that $V_{\mathbb{R}}(I)$ is compact. Then the theta body sequence of I converges to the convex hull of the variety $V_{\mathbb{R}}(I)$ in the sense that*

$$\bigcap_{k=1}^{\infty} \mathrm{TH}_k(I) = \mathrm{conv}(V_{\mathbb{R}}(I)).$$

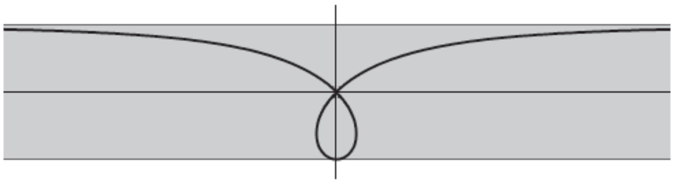

Figure 7.12. *Strophoid curve and its convex hull.*

This is an immediate consequence of Schmudgën's *Positivstellensatz* (see Chapter 3). To see the connection, just consider any set of generators $\{g_1, \ldots, g_t\}$ for I and the semialgebraic set $S = \{x \in \mathbb{R}^n : \pm g_1 \geq 0, \ldots \pm g_t \geq 0\} = V_{\mathbb{R}}(I)$. When applied to S, Schmudgën's Positivstellensatz guarantees that every linear polynomial that is strictly positive over $V_{\mathbb{R}}(I)$ is sos modulo I.

Example 7.33. The existence of varieties as in Example 7.31 does not imply that for all unbounded varieties we have problems with the theta body sequence. Consider the *strophoid* curve given by $p(x, y) := (1 - y)x^2 - (1 + y)y^2 = 0$, shown in Figure 7.12. The closure of the convex hull of this variety is the band B defined by $-1 \leq y \leq 1$. We claim that $\mathrm{TH}_2(I) = B$. To show this it is enough to prove that both $1 - y$ and $1 + y$ are 2-sos modulo I, which is true since

$$1 \pm y = \left(1 \pm \frac{1}{2}y - \frac{1}{2}y^2\right)^2 + \frac{1}{4}\left(\mp y - y^2\right)^2 + \frac{1}{2}\left(xy - x\right)^2 + \frac{1}{2}(y - 1)p(x, y). \qquad \blacksquare$$

In what follows we concentrate our efforts on the compact case, where asymptotic convergence of the theta body sequence is guaranteed. The next natural question when $V_{\mathbb{R}}(I)$ is infinite but compact is whether we can understand when the theta body sequence converges in finitely many steps to $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$. Finite convergence would prove that $\mathrm{conv}(V_{\mathbb{R}}(I))$ is the projection of a spectrahedron, which is an important feature of a convex semialgebraic set as seen in Chapter 6. There is no complete understanding of this situation, but in the remainder of this section, we discuss the known results.

$\mathbf{TH_1}$-exactness. We begin by discussing the strongest scenario within finite convergence, namely TH_1-exactness of an ideal. In spite of the strength of this property, there are surprisingly many interesting examples of such ideals with infinite real varieties. We begin by taking a general look at the notion of TH_1-exactness for all ideals. Roughly speaking, TH_1-exact ideals are those whose quadratic elements are enough to describe their convex geometry, a statement that will be made precise shortly. We start with a small lemma concerning convex quadrics.

Lemma 7.34. *If $p \in \mathbb{R}[x]$ is a convex quadric polynomial, then $\langle p \rangle$ is* TH_1*-exact.*

Proof. This result will follow from Proposition 7.41, where we will show that the first theta body of any quadric is simply the convex hull of its graph intersected with the x-plane. This intersection is precisely $\text{conv}(p)$ if p is convex. □

We now give an alternative characterization of $\text{TH}_1(I)$ for any ideal I.

Proposition 7.35. *For any ideal* $I \subseteq \mathbb{R}[x]$, $\text{TH}_1(I)$ *equals the intersection of* $\text{conv}(V_{\mathbb{R}}(p))$ *as* p *varies over all convex quadrics in* I.

Proof. The inclusion $\text{TH}_1(I) \subseteq \text{conv}(V_{\mathbb{R}}(p))$ for all convex quadrics $p \in I$ is easy, since a linear inequality is valid over the second set if and only if it is 1-sos modulo $\langle p \rangle$, which immediately implies that it is 1-sos modulo I and therefore valid on $\text{TH}_1(I)$. For the second inclusion note that if $l(x)$ is 1-sos mod I, then

$$l(x) = \sigma(x) + g(x),$$

where σ is a sum of squares and g is a quadric in I. But note that $-\nabla^2 g = \nabla^2 \sigma \succeq 0$ which implies $-g$ is a convex quadric in I, and $l(x)$ is 1-sos modulo $\langle -g \rangle$. Therefore, $l(x) \geq 0$ is valid on $\text{conv}(V_{\mathbb{R}}(-g))$ and hence also valid on the intersection of $\text{conv}(V_{\mathbb{R}}(p))$ as p varies over all convex quadrics in I. □

Example 7.36. Consider the ideal $I = \left\langle x^4 - y^2 - z^2, x^4 + x^2 + y^2 - 1 \right\rangle$ that we introduced in Example 7.11. This is the intersection of two quartic surfaces in $\mathbb{R}^3$. The Gröbner basis computation we did then shows that there exists a single quadric in this ideal (up to scalar multiplication), which is the polynomial $-1+x^2+2y^2+z^2$. Therefore, $\text{TH}_1(I)$ equals the ellipsoid $\{(x, y, z) \in \mathbb{R}^3 : x^2 + 2y^2 + z^2 \leq 1\}$, as seen in Figure 7.4. ■

Proposition 7.35 can sometimes be used to prove TH_1-exactness.

Example 7.37. Consider the ideal $I = \langle x^2 + y^2 + z^2 - 4, (x-1)^2 + y^2 - 1 \rangle$, from Example 7.47. Note that the quadratic polynomials $p_1 = (x-1)^2 + y^2 - 1$ and $p_2 = 2x + z^2 - 4$ belong to I. Write $I_1 = \langle p_1 \rangle$ and $I_2 = \langle p_2 \rangle$. Then we claim that

$$\text{conv}(V_{\mathbb{R}}(I)) = \text{conv}(V_{\mathbb{R}}(I_1)) \cap \text{conv}(V_{\mathbb{R}}(I_2)),$$

and therefore I is TH_1-exact. To see this note that the variety $V_{\mathbb{R}}(I)$ can be written as

$$\{(x, \pm\sqrt{1-(x-1)^2}, \pm\sqrt{4-2x}) : 0 \leq x \leq 2\}.$$

In particular for each fixed x we get four points, and the rectangle they form must be contained in the convex hull of $V_{\mathbb{R}}(I)$. This means

$$\{(x, y, z) \in \mathbb{R}^3 : |y| \leq \sqrt{1-(x-1)^2}, |z| \leq \sqrt{4-2x}, 0 \leq x \leq 2\} \subseteq \text{conv}(V_{\mathbb{R}}(I)),$$

but it is clear that this set can be rewritten as

$$\{(x, y, z) \in \mathbb{R}^3 : y^2 \leq 1-(x-1)^2, z^2 \leq 4-2x\} = \text{conv}(V_{\mathbb{R}}(I_1)) \cap \text{conv}(V_{\mathbb{R}}(I_2)),$$

which contains $\text{conv}(V_{\mathbb{R}}(I))$, so we get the intended equality. ■

An important open question concerning TH_1-exactness of varieties comes from oriented Grassmannians and illustrates that the TH_1 relaxation can be surprisingly powerful. For the purposes of this discussion, we define the oriented Grassmannian $G_{k,n}$ to be the set of all oriented k-subspaces of $\mathbb{R}^n$, embedded in $\mathbb{R}^{\binom{n}{k}}$ by taking Plücker coordinates, i.e., by picking an oriented basis of the space, writing the vectors as an $n \times k$ matrix, and taking all $k \times k$ minors and scaling them by a positive scalar to a point on the sphere $S^{\binom{n}{k}-1}$.

The ideal $I_{k,n}$, generated by all the quadratic relations among the $k \times k$ minors of an $n \times k$ matrix, is called the *Plücker* ideal. The Grassmann variety is then the compact real variety of the ideal $I = I_{n,k} + \langle 1 - \|x\|^2 \rangle$, so it makes sense to approximate it with theta bodies. It is unknown whether all Grassmann varieties are TH_1-exact, in fact even the $G_{3,6}$ case is unknown, but numerical simulations seem to say it is, at least for the relatively small examples for which numerical computations are doable. Unpublished work by Sanyal and Rostalski [26] makes connections between TH_1-exactness of these ideals and some classical open questions of Harvey and Lawson on calibrated geometries [12].

Exactness in one step for principal ideals. Principal ideals are the simplest ideals with infinite varieties. However, even in this case, TH_1-exactness is not to be expected. In fact, if p has degree d and $2k < d$, $\mathrm{TH}_k(p)$ is the full ambient space $\mathbb{R}^n$, since any k-sos linear inequality would verify $l(x) = \sigma(x) + g(x)$ with degree of the sums of squares σ less than or equal to $2k$. But the degree of $g \in I$ must be at least d so there would be no cancellation of the highest degree and the sum could never be a linear polynomial. An interesting question in this case is whether and when the first meaningful theta body would equal $\mathrm{conv}(V_{\mathbb{R}}(p))$ when $I = \langle p \rangle$. We will focus on the following problem: given a polynomial p of degree $2k$, decide if $\langle p \rangle$ is TH_k-exact. In this generality there is a simple necessary criterion, but we have to introduce a few definitions in order to state it.

Definition 7.38. *Consider a polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ and define $\tilde{p} = x_0 - p(x_1, \dots, x_n) \in \mathbb{R}[x_0, x_1, \dots, x_n]$. Consider the convex set $C = \mathrm{conv}(V_{\mathbb{R}}(\tilde{p}))$, which is simply the convex hull of the graph of p, and define the* shadow area *of p, denoted by $\mathrm{sh}(p)$, as the intersection of C with the plane $x_0 = 0$.*

This shadow area clearly contains $\mathrm{conv}(V_{\mathbb{R}}(p))$ since it is convex and contains the variety. However we can easily establish a more interesting inclusion.

Proposition 7.39. *For $p \in \mathbb{R}[x]$ of degree $2k$, $\mathrm{sh}(p) \subseteq \mathrm{TH}_k(\langle p \rangle)$. In particular if $\mathrm{sh}(p)$ strictly contains the closure of the convex hull of $V_{\mathbb{R}}(p)$, then $\langle p \rangle$ is not TH_k-exact.*

Proof. Let $l(x)$ be k-sos modulo $\langle p \rangle$, i.e., $l(x) = \sigma(x) + \lambda p(x)$ where σ is a sum of squares of degree at most $2k$ and $\lambda \in \mathbb{R}$. Then $l(x) - \lambda p(x) = \sigma(x)$ implies $l(x) - \lambda p(x) \geq 0$ everywhere and therefore $\tilde{l}(x_0, x) := l(x) - \lambda x_0$ is valid over $V_{\mathbb{R}}(\langle \tilde{p} \rangle)$ and hence over its convex hull too. But by intersecting with $x_0 = 0$ we

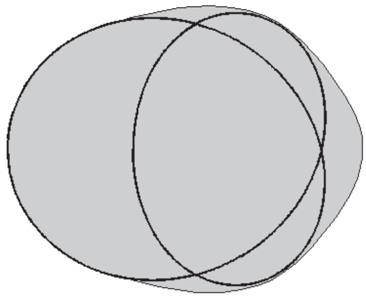

Figure 7.13. *Scarabaeus curve and its third theta body.*

get that $l(x) \geq 0$ must be valid on $\mathrm{sh}(p)$. From the definition of $\mathrm{TH}_k(I)$ it follows immediately that $\mathrm{sh}(p) \subseteq \mathrm{TH}_k(I)$ as intended. $\square$

Despite the simplicity of the criterion, it is a handy tool to prove that a principal ideal is not exact at the first step, without relying on numerical approximations.

Example 7.40. Consider the *scarabaeus* curve given by

$$p(x, y) := (x^2 + y^2)(x^2 + y^2 + 4x)^2 - (x^2 - y^2)^2 = 0.$$

A simple numerical computation with an SDP solver shows us that $\mathrm{TH}_3(\langle p \rangle)$ does not match the convex hull of the curve, as can be seen in Figure 7.13. To provide a short exact proof, one just has to point out that $p(-4, 0) = 256$ and $p(1, 0) = 24$, and since the point $(\frac{4}{7}, 0, 0)$ lies in the segment between $(-4, 0, 256)$ and $(1, 0, 24)$, the point $\xi = (\frac{4}{7}, 0)$ must be contained in $\mathrm{sh}(p)$ and therefore in $\mathrm{TH}_3(\langle p \rangle)$. It is, however, easy to calculate that the maximum value that x attains on the curve is $(-50 + 11\sqrt{22})/27 \approx 0.06$, which implies that the convex hull must not contain ξ. $\blacksquare$

In some very special cases we can actually say a bit more about the first meaningful theta body.

Proposition 7.41. *Let p be a polynomial in n variables and degree $2d$. Then*

1. *if $n = 1$, $\mathrm{sh}(p) = \mathrm{TH}_d(\langle p \rangle)$;*
2. *if $d = 1$, $\mathrm{sh}(p) = \mathrm{TH}_1(\langle p \rangle)$;*
3. *if $n = 2$ and $d = 2$, $\mathrm{sh}(p) = \mathrm{TH}_2(\langle p \rangle)$.*

Proof. We just have to prove that in these cases $\mathrm{sh}(p) \supseteq \mathrm{TH}_d(\langle p \rangle)$. To do this let $l(x) > 0$ be a valid linear inequality over $\mathrm{sh}(p)$. This means that the line $L = \{(x_0, x) : x_0 = 0, l(x) = 0\}$ does not intersect $C = \mathrm{conv}(V_{\mathbb{R}}(\langle \tilde{p} \rangle))$. By the

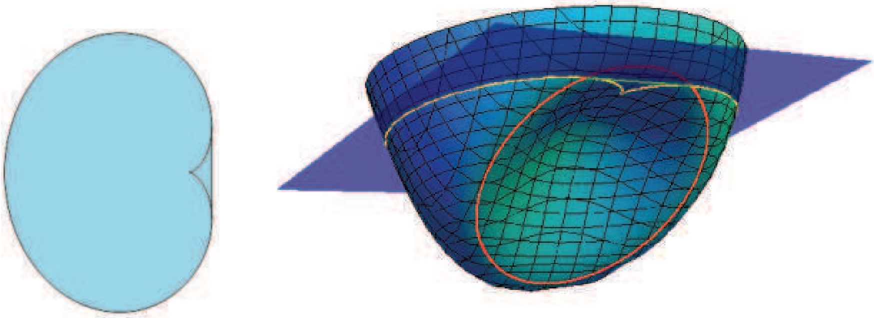

Figure 7.14. *On the left we see the cardioid $p(x) = 0$ and its convex hull. On the right we see the graph of p, its intersection with the plane $z = 0$ and the ellipsoidal region where the graph and the boundary of its convex hull differ.*

separation theorem for convex sets we can therefore take a hyperplane H that strictly separates L and C. Since H does not touch the graph of p, it depends on x_0, and since it does not touch L, it must be parallel to it. Therefore we have a hyperplane of the form $l'(x_0, x) := x_0 + \lambda(l(x) - \varepsilon) = 0$, with $\lambda \neq 0$, $\varepsilon > 0$. Since $\tilde{p}(x_0, x) = x_0 - p(x)$, this means that $\sigma(x) := p(x) + \lambda(l(x) - \varepsilon)$ is always nonnegative or always nonpositive. Without loss of generality assume it is always nonnegative (which implies $\lambda > 0$). Since the degree and number of variables of this polynomial fall under Hilbert's result (see Chapter 4), $\sigma(x)$ is a sum of squares. Hence, $l(x) = \sigma(x)/\lambda + \varepsilon - p(x)/\lambda$ is d-sos modulo the ideal, which implies that $l(x) \geq 0$ is valid over $\mathrm{TH}_d(\langle p \rangle)$, proving the inclusion. □

Example 7.42. We use the above result to prove TH_2-exactness of the following principal ideal. Consider

$$p(x, y) = (x^2 + y^2 + 2x)^2 - 4(x^2 + y^2)$$

defining a *cardioid*, and the function

$$q(x, y) = \begin{cases} p(x, y) & \text{if } (x+1)^2 + y^2 \geq 3, \\ 8x - 4 & \text{if } (x+1)^2 + y^2 < 3. \end{cases}$$

One can check that q is smooth and convex by noticing that $p(x, y) = ((x+1)^2 + y^2 - 3)^2 + 8x - 4$ and by looking at its Hessian. Furthermore, the convex hull of the graph of p is just the region above the graph of q. Therefore $\mathrm{sh}(p) = \{(x, y) : q(x, y) \leq 0\}$, and we can see in Figure 7.14 that $\mathrm{sh}(p)$ is the convex hull of the cardioid. ■

Even for one-variable polynomials this result is interesting.

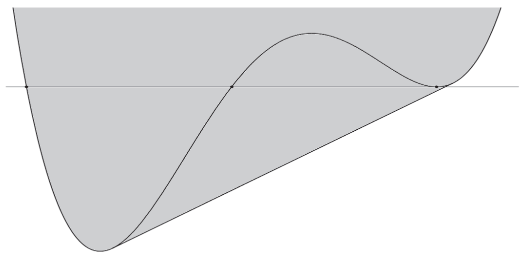

Figure 7.15. *Graph of the polynomial $x - x^2 - x^3 + x^4$, its convex hull, and intersection with the x-axis.*

Example 7.43. Consider the polynomial $p(x) = x - x^2 - x^3 + x^4$. In Figure 7.15 we can see that this polynomial is not TH_2-exact, and why that happens. The double root at $x = 1$ forces the convex hull of the graph to include some points to the right of $x = 1$. In fact one can compute precisely the double tangent that defines the boundary of the convex hull and show that $\mathrm{TH}_2(\langle p \rangle) = \left[-1, \frac{25}{24}\right]$. ∎

Singularities and convergence. We now return to the more general question of finite convergence of the theta body sequence for an ideal with an infinite real variety. There is no complete understanding of the obstructions to finite convergence, but we now show that if $V_{\mathbb{R}}(I)$ has certain types of *singularities*, then finite convergence is not possible.

Given an ideal I and a point P on the real variety of I, we define the *normal space* $N_P(I)$ to be the linear space $\{\nabla f(P) : f \in I\}$.

Proposition 7.44. *Let $l(x)$ be an affine polynomial such that $l(P) = 0$ for some P in $V_{\mathbb{R}}(I)$. If $\nabla l \notin N_P(I)$, then l is not a sum of squares modulo I.*

Proof. Suppose l is a sum of squares. Then

$$l(x) = \sigma(x) + g(x) \tag{7.3}$$

for some sum of squares σ and some polynomial $g \in I$. By evaluating at P we get that $\sigma(P) = 0$, which immediately implies $\nabla\sigma(P) = 0$. By differentiating (7.3) we get

$$\nabla l = \nabla\sigma(x) + \nabla g(x), \tag{7.4}$$

and by evaluating at P we get that $\nabla l = \nabla g(P) \in N_P(I)$. □

If I is real radical we can say even more.

Corollary 7.45. *If I is real radical and $l(x) \geq 0$ is a linear inequality valid on $V_{\mathbb{R}}(I)$ with $l(P) = 0$ at a point $P \in V_{\mathbb{R}}(I)$ such that $\nabla l \notin N_P(I)$, then I is not TH_k-exact for any k.*

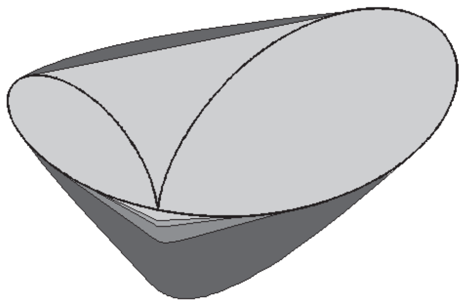

Figure 7.16. $\mathrm{TH}_2(I)$, $\mathrm{TH}_3(I)$, $\mathrm{TH}_4(I)$, *and* $\mathrm{TH}_5(I)$*: all contain the origin in their interior.*

Proof. This follows from the previous proposition and Lemma 7.27. ☐

Example 7.46. Let $p(x, y) = (x^2 + y^2)^2 - (x + 5y)x^2$ and $I = \langle p \rangle$. This ideal defines a *bifolium* with a singularity at the origin, which implies $N_{(0,0)}(I) = \{(0, 0)\}$. Furthermore the linear inequality $x + 5y \geq 0$ is valid on the variety and holds with equality at the origin. Since $(1, 5) \notin N_{(0,0)}(I)$ we immediately have that this inequality does not hold for any theta body relaxation of this ideal. In Figure 7.16 we can see $\mathrm{TH}_k(I)$ for $k = 2, 3, 4, 5$, and see that in fact the inequality does not hold for any of them. ∎

Corollary 7.45 essentially tells us that certain singularities of the ideal I that are in the boundary of the convex hull of $V_{\mathbb{R}}(I)$ affect the convergence of the theta bodies of I. For a point $P \in V_{\mathbb{R}}(I)$, the expected dimension of the normal space $N_P(I)$ is the codimension of $V_{\mathbb{R}}(I)$. A reasonable notion of a singularity of I is a point $P \in V_{\mathbb{R}}(I)$ for which $N_P(I)$ has smaller dimension than expected. The next example will show that just the existence of singularities of I on the boundary of $\mathrm{conv}(V_{\mathbb{R}}(I))$ is not enough for Corollary 7.45 to apply.

Example 7.47. Consider the variety $V_{\mathbb{R}}(I)$ in $\mathbb{R}^3$ defined by the ideal

$$I = \langle x^2 + y^2 + z^2 - 4, (x - 1)^2 + y^2 - 1 \rangle.$$

As seen in Figure 7.17, this variety looks like a curved figure-eight and has a singularity at the point $p = (2, 0, 0)$, which belongs to the boundary of $\mathrm{conv}(V_{\mathbb{R}}(I))$. This happens since $N_P(I) = \mathbb{R}\{(1, 0, 0)\}$ has dimension one, smaller than the codimension of the variety, which is two. However, $(2, 0, 0)$ does not cause problems for the convergence of theta bodies since the only linear polynomial that is zero at p and nonnegative on $V_{\mathbb{R}}(I)$ is the polynomial $2 - x$, whose gradient is in $N_P(I)$. Indeed, the first theta body of I already equals $\mathrm{conv}(V_{\mathbb{R}}(I))$, as we will see in Example 7.37. ∎

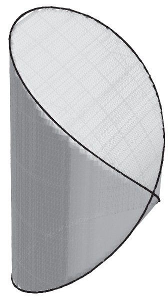

Figure 7.17. *The curved eight variety and its convex hull.*

A better, more refined, way of looking at singularities was introduced by Omar and Osserman in [23]. They introduce a stronger notion of nonnegativity over varieties that yields a stronger necessary condition for finite convergence of the theta body hierarchy. As a byproduct they prove the following result.

Theorem 7.48. *Let $f(x)$ be a polynomial such that there exists some positive integer n and an $\mathbb{R}$-algebra homomorphism $\varphi : \mathbb{R}[x]/I \to \mathbb{R}[\varepsilon]/\langle\varepsilon^n\rangle$ for which $\varphi(f) = a_0 + a_1\varepsilon + \cdots + a_{n-1}\varepsilon^{n-1}$. If the first nonzero (leading) coefficient a_i is negative, then f is not a sum of squares modulo I.*

Proof. Just note that homomorphisms send sums of squares to sums of squares, and sums of squares in $\mathbb{R}[\varepsilon]/\langle\varepsilon^n\rangle$ always have their leading coefficient nonnegative. □

Again this immediately gives us a new criterion.

Corollary 7.49. *Let I be a real radical ideal and $l(x) \geq 0$ a linear inequality valid on $V_{\mathbb{R}}(I)$. If there exists an $\mathbb{R}$-algebra homomorphism $\varphi : \mathbb{R}[x]/I \to \mathbb{R}[\varepsilon]/\langle\varepsilon^n\rangle$ for which $\varphi(l)$ has negative leading coefficient, then I is not* TH_k*-exact for any k.*

This corollary is much stronger than Corollary 7.45, and examples showing the difference are presented in [23]. In our next example we just show that we can recover Corollary 7.45 from Corollary 7.49 for the variety in Example 7.46 but, in fact, we can do so for any variety just by considering maps to $\mathbb{R}[\varepsilon]/\langle\varepsilon^2\rangle$.

Example 7.50. Let $p(x, y) = (x^2+y^2)^2-(x+5y)x^2$ and $I = \langle p\rangle$ as in Example 7.46. Then the map $\varphi : \mathbb{R}[x, y]/I \to \mathbb{R}[\varepsilon]/\langle\varepsilon^2\rangle$ defined by $\varphi(x) = \varphi(y) = -\varepsilon$ is well defined, since $\varphi(p) = 0$. However, $\varphi(x+5y) = -6\varepsilon$ has a negative leading coefficient despite $x + 5y \geq 0$ being valid on the variety. Hence, $\langle p\rangle$ is not TH_k-exact for any k. ∎

One should keep in mind that singularities are not necessarily the only things that prevent finite convergence of the theta body sequence to $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$. For compact smooth curves and surfaces, Scheiderer proved that nonnegativity and

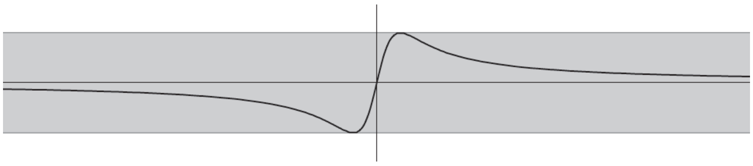

Figure 7.18. *Serpentine curve and the closure of its convex hull.*

sums of squares modulo the ideal are equivalent [28, 29]. However, even in these cases, it is an open question if one can bound the degree needed to represent every nonnegative affine polynomial as a sum of squares modulo the ideal. Thus there might be examples of smooth curves and surfaces with no finite convergence of the theta body hierarchy to $\mathrm{conv}(V_{\mathbb{R}}(I))$. The only cases where we know a little more is when the genus of the curve is one.

Proposition 7.51 (Theorem 2.1 [30]). *If $V_{\mathbb{R}}(I)$ is a smooth curve of genus 1 with at least one nonreal point at infinity, then I is TH_k-exact for some k.*

Genus zero curves can be rationally parametrized which allows semidefinite representations of their convex hulls by means of sums of squares, as seen in [13]. However such constructions do not automatically translate to finite convergence of the theta body sequence to the convex hull of the curve, even in the smooth case.

For varieties of dimension greater than two, there always exist nonnegative polynomials that are not sums of squares modulo any ideal that defines them, even in the smooth compact case, as seen in [27]. It is therefore very natural to expect examples of smooth compact varieties with no finite convergence of the theta body hierarchy, but we do not know a concrete example at this point.

Exercise 7.52. Consider the serpentine curve given by $p(x) := y(x^2+1) - x = 0$, depicted in Figure 7.18. The closure of its convex hull is the band cut out by the inequalities $-1/2 \leq y \leq 1/2$. Show that the ideal $I = \langle p \rangle$ is TH_2-exact by giving an exact expression of $1 - 2y$ and $1 + 2y$ as 2-sos polynomials modulo I.

Exercise 7.53. Using Proposition 7.35 show that the first theta body of the vanishing ideal of the points $\{(0,0), (1,0), (0,1), (2,2)\}$ is cut out by precisely two polynomial inequalities, and write them explicitly.

Exercise 7.54. Consider the ideal $I = \left\langle y^2 - x^5, z - x^3 \right\rangle$. The inequality $z \geq 0$ is valid on the variety $V_{\mathbb{R}}(I)$.

1. Can we use Proposition 7.44 to prove that z is not k-sos modulo I for any k?

2. Use Theorem 7.48 to prove that z is not k-sos modulo I for any k.

Exercise 7.55. Similarly to our definition of 2-level polytope, we can define a k-level polytope to be one where given a facet F, and the affine plane H_F that it spans, all vertices of the polytope are contained either in H_F or in one of $k-1$ parallel translates of H_F. Prove that if S is the set of vertices of a $(k+1)$-level polytope then the vanishing ideal of S, $I(S)$, is TH_k-exact.

Exercise 7.56. Consider the univariate quartic polynomial $p(x) = x^4 - 3x^3 + 3x^2 - 3x + 2$ which has two real roots, 1 and 2. Compute $\mathrm{TH}_2(\langle p \rangle)$ exactly. Is the ideal TH_2-exact?

Exercise 7.57. Consider the bifolium given by $p(x,y) := (x^2+y^2)^2 - yx^2 = 0$. This curve has a singularity at the origin, which is also on the boundary of its convex hull and satisfies the conditions of Corollary 7.45, and hence we know that its theta body hierarchy does not converge. Using the same ideas as in Example 7.19, add to the second moment matrix of $I = \langle p \rangle$ a row and a column indexed by $\frac{y^2}{x}$. Plot the resulting approximation and compare it with the convex hull of the curve.

7.4 Combinatorial Optimization

In this final section, we focus on combinatorial optimization where a typical problem involves optimizing a linear function over all combinatorial objects of a certain kind. Many of these problems are modeled using graphs and can sometimes be studied combinatorially. However, a more systematic approach is to model these problems as integer or linear programs, which puts an emphasis on the underlying geometry. These models work as follows. The combinatorial objects of interest are typically defined as subsets of the ground set $[n] := \{1, 2, \ldots, n\}$ and the object $T \subseteq [n]$ is recorded via its *characteristic vector* $\chi^T \in \{0,1\}^n$ defined as $\chi_i^T = 1$ if $i \in T$ and $\chi_i^T = 0$ otherwise. This creates a simple bijection between the objects and certain elements of $\{0,1\}^n$. Then, for a vector $c \in \mathbb{R}^n$, maximizing $\sum_{i \in T} c_i$ over all the objects $\{T\}$ is equivalent to maximizing $\sum c_i x_i$ over the characteristic vectors $\{\chi^T\}$ which in turn is equivalent to maximizing $\sum c_i x_i$ over $\mathrm{conv}(\{\chi^T\})$ which is a 0/1 polytope by construction. (Recall that a 0/1 polytope in $\mathbb{R}^n$ is the convex hull of vectors in $\{0,1\}^n$.) In principle this is a linear program but the difficulty is that no description of $\mathrm{conv}(\{\chi^T\})$ is usually known, and one resorts to relaxations of $\mathrm{conv}(\{\chi^T\})$ over which $\sum c_i x_i$ is maximized to obtain an upper bound on the value of $\max\{\langle c, x \rangle \,:\, x \in \mathrm{conv}(\{\chi^T\})\}$.

The theory of integer programming offers general methods to construct polyhedral relaxations of $\mathrm{conv}(\{\chi^T\})$ by first finding a polytope whose integer points are precisely $\{\chi^T\}$. See [31, Chapter 23] for linear programming–based methods. Polyhedral relaxations can sometimes be found using combinatorial arguments that depend explicitly on the structure of the problem. Automatic methods for constructing relaxations have also come about from *lift-and-project* methods that find a sequence of polyhedral or spectrahedral relaxations of $\mathrm{conv}(\{\chi^T\})$. Some examples of lift-and-project methods besides, the theta body method described in this chapter, can be found in [2, 14, 20, 33] (see also [15]). Theta bodies construct relaxations of $\text{conv}(V_{\mathbb{R}}(I))$ for an ideal I. In the special case of the combinatorial optimization model described above, the starting point is the finite set $\{\chi^T\}$ which is a finite algebraic variety, and we typically take its vanishing ideal as the ideal whose theta bodies are to be computed. As we saw in Section 7.3.1, these real radical ideals are always TH_k-exact for some finite k. We take a closer look at some combinatorial optimization problems whose theta bodies have been explored.

7.4.1 Stable Sets in a Graph

An example that is at the heart of the history of theta bodies is the *maximum stable set problem* in an undirected graph $G = ([n], E)$ with vertex set $[n]$ and edge set E. A *stable set* in G is a set $U \subseteq [n]$ such that for all $i, j \in U$, $\{i, j\} \notin E$. The maximum stable set problem seeks the stable set of largest cardinality in G, the size of which is the *stability number*, $\alpha(G)$, of G.

The maximum stable set problem can be modeled as follows. For each stable set $U \subseteq [n]$, let $\chi^U \in \{0,1\}^n$ be its characteristic vector defined as $\chi_i^U = 1$ if $i \in U$ and $\chi_i^U = 0$ otherwise. Let $S_G \subseteq \{0,1\}^n$ be the set of characteristic vectors of all stable sets in G. Then $\text{STAB}(G) := \text{conv}(S_G)$ is called the *stable set polytope* of G and the maximum stable set problem is, in theory, the linear program $\max\{\sum_{i=1}^n x_i \ : \ x \in \text{STAB}(G)\}$ with optimal value $\alpha(G)$. However, $\text{STAB}(G)$ is not known a priori, and so one resorts to relaxations of it over which to optimize $\sum_{i=1}^n x_i$.

Polyhedral relaxations of $\text{STAB}(G)$ can be constructed from combinatorial arguments. For instance, a well-known relaxation is the polytope

$$\text{FRAC}(G) := \{x \in \mathbb{R}^n \ : \ x_i + x_j \leq 1 \text{ for all } \{i,j\} \in E, \ \ x_i \geq 0 \text{ for all } i \in [n]\},$$

where the constraint $x_i + x_j \leq 1$ for $\{i, j\} \in E$ comes from the fact that both endpoints of an edge cannot be in a stable set. It can be checked that $\text{STAB}(G)$ is exactly the convex hull of the integer points in $\text{FRAC}(G)$. The polytope $\text{FRAC}(G)$ and several tighter polyhedral relaxations of $\text{STAB}(G)$ have been studied extensively in the literature; see [11, Chapter 9].

Since the set S_G is an algebraic variety, the theta bodies of its vanishing ideal offer convex relaxations of $\text{STAB}(G)$. This vanishing ideal is:

$$I_G := \langle x_i^2 - x_i \text{ for all } i \in [n], \ \ x_i x_j \text{ for all } \{i,j\} \in E\rangle \subset \mathbb{R}[x_1, \dots, x_n].$$

For $U \subseteq [n]$, let $x^U := \prod_{i \in U} x_i$. From the generators of I_G it follows that if $f \in \mathbb{R}[x]$, then $f \equiv g \bmod I_G$ where g is in the $\mathbb{R}$-span of the set of monomials $\{x^U \ : \ U \text{ is a stable set in } G\}$. In particular,

$$\mathcal{B} := \{x^U + I_G \ : \ U \text{ stable set in } G\}$$

is a θ-basis of $\mathbb{R}[x]/I_G$ (containing $1 + I_G, x_1 + I_G, \dots, x_n + I_G$). This implies that $\mathcal{B}_k = \{x^U + I_G \ : \ U \text{ stable set in } G, \ |U| \leq k\}$, and for $x^{U_i} + I_G, x^{U_j} + I_G \in \mathcal{B}_k$, their product is $x^{U_i \cup U_j} + I_G$, which is $0 + I_G$ if $U_i \cup U_j$ is not a stable set in G. This product formula allows us to compute $M_{\mathcal{B}_k}(y)$, where we index the element

$x^U + I_G \in \mathcal{B}_k$ by the set U. Since $S_G \subseteq \{0,1\}^n$ and $I(G)$ is the vanishing ideal of S_G, by Theorems 7.8, we have that

$$\mathrm{TH}_k(I_G) = \left\{ y \in \mathbb{R}^n \ : \ \begin{array}{l} \exists\, M \succeq 0,\ M \in \mathbb{R}^{|\mathcal{B}_k|\times|\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i \\ M_{UU'} = 0 \text{ if } U \cup U' \text{ is not stable in } G \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

In particular, indexing the one-element stable sets by the vertices of G,

$$\mathrm{TH}_1(I_G) = \left\{ y \in \mathbb{R}^n \ : \ \begin{array}{l} \exists\, M \succeq 0, M \in \mathbb{R}^{(n+1)\times(n+1)} \text{ such that} \\ M_{00} = 1, \\ M_{0i} = M_{i0} = M_{ii} = y_i \ \forall\, i \in [n] \\ M_{ij} = 0 \text{ for all } \{i,j\} \in E \end{array} \right\}.$$

Example 7.58. Let $G = ([5], \{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,5\}\})$ be a 5-cycle. The vanishing ideal of the characteristic vectors of stable sets in G is

$$I_G = \langle x_1x_2, x_2x_3, x_3x_4, x_4x_5, x_1x_5, x_i^2 - x_i \text{ for all } i = 1,\ldots,5\rangle,$$

and a θ-basis for $\mathbb{R}[x]/I_G$ is given by

$$\mathcal{B} = \{1, x_1, x_2, x_3, x_4, x_5, x_1x_3, x_1x_4, x_2x_4, x_2x_5, x_3x_5\} + I_G.$$

Let $y \in \mathbb{R}^{10}$ be the vector of variables whose coordinates are indexed by $\mathcal{B}$ in the given order and with $y_0 = 1$. Then

$$\mathrm{TH}_1(I_G) = \left\{ y \in \mathbb{R}^5 \ : \ \exists y_6,\ldots,y_{10} \text{ s.t. } M_{\mathcal{B}_1}(y) \succeq 0 \right\},$$

where

$$M_{\mathcal{B}_1}(y) = \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 \\ y_1 & y_1 & 0 & y_6 & y_7 & 0 \\ y_2 & 0 & y_2 & 0 & y_8 & y_9 \\ y_3 & y_6 & 0 & y_3 & 0 & y_{10} \\ y_4 & y_7 & y_8 & 0 & y_4 & 0 \\ y_5 & 0 & y_9 & y_{10} & 0 & y_5 \end{pmatrix}.$$

Note that $x_i \equiv x_i^2$ and $1 - x_i \equiv (1-x_i)^2$ mod I_G for any graph G, so $\mathrm{TH}_1(I_G)$ is always contained in the $[0,1]$ cube. ∎

The first example of an SDP relaxation of a combinatorial optimization problem was the *theta body* of a graph $G = ([n], E)$ constructed by Lovász in [18] while studying the Shannon capacity of graphs. The theta body of G, denoted as $\mathrm{TH}(G)$, is a relaxation of $\mathrm{STAB}(G)$ that was originally defined as the intersection of the infinitely many half spaces that arise from the *orthonormal representations* of G. Several equivalent definitions can be found in [18] and [11, Chapter 9]. However, none of them point to an obvious generalization of the construction to other discrete optimization problems. In [20], Lovász and Schrijver observe that $\mathrm{TH}(G)$ can be formulated via semidefinite programming exactly as the formulation for $\mathrm{TH}_1(I_G)$ shown above. This is still specialized to the stable set problem. Then in [19], Lovász observes that, in fact, $\mathrm{TH}(G)$ is cut out by all linear polynomials that are 1-sos mod the ideal I_G. For the stable set problem, this fact can be proven without all the machinery introduced in this paper. This connection leads naturally to the definition of $\mathrm{TH}_k(I_G)$ for any positive integer k and more generally $\mathrm{TH}_k(I)$ for any ideal $I \subseteq \mathbb{R}[x]$ and any k. Problem 8.3 in [19] (roughly) asks to characterize all ideals $I \subseteq \mathbb{R}[x]$ such that $\mathrm{cl}(\mathrm{conv}(V_{\mathbb{R}}(I)))$ equals $\mathrm{TH}_1(I)$ or more generally, $\mathrm{TH}_k(I)$. It was this problem that motivated us to study theta bodies in general and develop the methods in this chapter.

Example 7.59. Let us return to the example Example 7.58. When Lovász introduced the theta body of a graph G, he also introduced the concept of *theta number* of a graph, $\vartheta(G)$ (c.f. Chapter 2). This is just the number

$$\max\left\{\sum_{i=1}^{n} x_i : x \in \mathrm{TH}(G) = \mathrm{TH}_1(I_G)\right\},$$

which is an upper bound (and approximation) for the stability number $\alpha(G)$ of a graph. We can now easily compute $\vartheta(C_5)$, the theta number of the 5-cycle, numerically using YALMIP, since we have the precise structure of the reduced moment matrix.

```
y=sdpvar(1,10);
M=[1    y(1) y(2) y(3)  y(4) y(5) ;
   y(1) y(1) 0    y(6)  y(7) 0    ;
   y(2) 0    y(2) 0     y(8) y(9) ;
   y(3) y(6) 0    y(3)  0    y(10);
   y(4) y(7) y(8) 0     y(4) 0    ;
   y(5) 0    y(9) y(10) 0    y(5) ];
obj=y(1)+y(2)+y(3)+y(4)+y(5);
solvesdp(M>=0,-obj);
double(obj)
```

This will return the answer $\vartheta(C_5) \approx 2.361$. Note that $\alpha(C_5) = 2$, so we do get an upper approximation as expected, but it is clear that I_{C_5} is not TH_1-exact. ∎

A particular reason for Lovász's interest in [19, Problem 8.3] was due to the fact that $\mathrm{STAB}(G) = \mathrm{TH}(G)$ if and only if G is a *perfect graph* [11, Corollary 9.3.27]. Recall that a graph is perfect if and only if it has no induced odd cycle of length at least five or its complement [4]. Since $\mathrm{TH}(G) = \mathrm{TH}_1(I_G)$ for all graphs G, it follows that I_G is TH_1-exact if and only if G is perfect. The pentagon in Example 7.58 is not perfect, which justifies our observation that its ideal I_G is not TH_1-exact. Chvátal and Fulkerson had shown that $\mathrm{STAB}(G) = \mathrm{QSTAB}(G)$ if and only if G is a perfect graph where

$$\mathrm{QSTAB}(G) := \left\{x \in \mathbb{R}^n \,:\, x_i \geq 0 \text{ for all } i \in [n],\ \sum_{i\in K} x_i \leq 1 \text{ for all cliques } K \text{ in } G\right\}.$$

A clique in G is a complete subgraph in G. Since every edge in G is a clique, $\mathrm{FRAC}(G) \supseteq \mathrm{QSTAB}(G) \supseteq \mathrm{STAB}(G)$ in general. A hexagon is perfect, in which case, $\mathrm{FRAC}(G) = \mathrm{QSTAB}(G)$ since the only cliques in G are its edges. Therefore, for the hexagon, $\mathrm{STAB}(G) = \mathrm{TH}(G) = \mathrm{TH}_1(I_G) = \mathrm{QSTAB}(G) = \mathrm{FRAC}(G)$. Since I_G is TH_1-exact if and only if G is perfect, by Theorem 7.30, we also have that $\mathrm{STAB}(G)$ is 2-level if and only if G is perfect.

The above discussion leads naturally to the question of which graphs G have the property that I_G is TH_2-exact, or more generally, TH_k-exact. These problems are open at the moment, although isolated examples of TH_k-exact ideals are known for specific values of $k > 1$. In practice it is quite difficult to find examples of graphs G for which I_G is not TH_2-exact although such graphs have to exist unless $P = NP$. Recent results of Au and Tunçel prove that if G is the line graph of the complete graph on $2n+1$ vertices, then the smallest k for which I_G is TH_k-exact grows linearly with n [1].

7.4.2 A General Framework

The stable set problem and many others in combinatorial optimization can be modeled as arising from a simplicial complex. A *simplicial complex* or *independence system*, Δ, with vertex set $[n]$, is a collection of subsets of $[n]$, called the *faces* of the Δ, such that whenever $S \in \Delta$ and $T \subset S$, then $T \in \Delta$. The *Stanley–Reisner* ideal of Δ is the ideal J_Δ generated by the square-free monomials $x_{i_1} x_{i_2} \cdots x_{i_k}$ such that $\{i_1, i_2, \ldots, i_k\} \subseteq [n]$ is not a face of Δ. If $I_\Delta := J_\Delta + \langle x_i^2 - x_i \,:\, i \in [n]\rangle$, then $V_{\mathbb{R}}(I_\Delta) = \{s \in \{0,1\}^n \,:\, \mathrm{support}(s) \in \Delta\}$. The support of a vector $v \in \mathbb{R}^n$ is the set $\{i \in [n] \,:\, v_i \neq 0\}$. Further, for $T \subseteq [n]$, if $x^T := \prod_{i \in T} x_i$, then $\mathcal{B} := \{x^T + I_\Delta \,:\, T \in \Delta\}$ is a θ-basis of $\mathbb{R}[x]/I_\Delta$. This implies that the kth theta body of I_Δ is

$$\mathrm{TH}_k(I_\Delta) = \pi_{\mathbb{R}^n}\{y \in \mathbb{R}^{\mathcal{B}_{2k}} \,:\, M_{\mathcal{B}_k}(y) \succeq 0,\, y_0 = 1\}.$$

Since $\mathcal{B}$ is in bijection with the faces of Δ and $x_i^2 - x_i \in I_\Delta$ for all $i \in [n]$, the theta body can be written explicitly as

$$\mathrm{TH}_k(I_\Delta) = \left\{ y \in \mathbb{R}^n \;:\; \begin{array}{l} \exists\, M \succeq 0,\, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i, \\ M_{UU'} = 0 \text{ if } U \cup U' \notin \Delta, \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

If the dimension of Δ is $d-1$ (i.e., the largest faces in Δ have size d), then I_Δ is TH_d-exact since all elements of $\mathcal{B}$ have degree at most d and hence the last possible theta body $\mathrm{TH}_d(I_\Delta)$ must coincide with $\mathrm{conv}(V_{\mathbb{R}}(I_\Delta))$ as $V_{\mathbb{R}}(I_\Delta)$ is finite. However, in many examples, I_Δ could be TH_k-exact for a k much smaller than d.

In the case of the stable set problem on $G = ([n], E)$, Δ is the set of all stable sets in G. This is a simplicial complex with vertex set $[n]$ whose nonfaces are the sets $T \subseteq [n]$ containing a pair $i, j \in [n]$ such that $\{i,j\} \in E$. Hence the minimal nonfaces (by set inclusion) are precisely the edges of G and so $J_\Delta = \langle x_i x_j \,:\, \{i,j\} \in E\rangle$.

Then $I_\Delta = J_\Delta + \langle x_i^2 - x_i \,:\, i \in [n]\rangle$, which is precisely the ideal I_G from Section 7.4.1, and the remaining facts about the θ-basis $\mathcal{B}$ used in Section 7.4.1 and the structure of the theta bodies of I_G follow from the general set up described above.

An example from combinatorial optimization that does not follow the simplicial complex framework is the *maximum cut problem* of finding the largest size *cut* in a graph. Recall that a cut in G is the collection of edges that go between the two parts of a partition of the vertices of G. Note that a subset of a cut is not necessarily a cut and hence the set of cuts in a graph do not form a simplicial complex. In [7] the theta body hierarchy for the maximum cut problem, and more generally for *binary matroids*, is studied. In this case, a θ-basis for the ideal in question is not obvious as in the simplicial complex model.

7.4.3 Triangle-free Subgraphs in a Graph

We finish the chapter with a second example from combinatorial optimization that fits the simplicial complex model. A subgraph H of a graph $G = ([n], E)$ is *triangle-free* if it does not contain a triangle (K_3, the complete graph on 3 vertices). Given weights on the edges of G, the *triangle-free subgraph problem* in G asks for a triangle-free subgraph of G of maximum weight. If all the edge weights are one, then the problem seeks a triangle-free subgraph in G with the most number of edges. The triangle-free subgraph problem is known to be NP-hard [36] and is relevant in various contexts within optimization.

The integer programming formulation of the triangle-free subgraph problem optimizes the linear function $\sum_{e \in E} w_e x_e$, where w_e is the weight on edge $e \in E$, over the characteristic vectors $\{\chi^H \,:\, H \text{ is triangle-free in } G\}$. This is equivalent to maximizing $\sum_{e \in E} w_e x_e$ over

$$P_{\text{tf}}(G) := \text{conv}\{\chi^H \,:\, H \text{ is triangle-free in } G\},$$

the *triangle-free subgraph polytope* of G. Note that $P_{\text{tf}}(G)$ is a full-dimensional $0/1$ polytope in $\mathbb{R}^E$. The triangle-free subgraph polytope of a graph has been studied by various authors (see, for instance, [3, 5]), and a number of facet defining inequalities of the polytope are known, although a full inequality description is not known or expected.

Taking Δ to be the simplicial complex on E consisting of all triangle-free subgraphs in G, and $I_{\text{tf}}(G) := I_\Delta$, we have that

$$V_{\mathbb{R}}(I_{\text{tf}}(G)) = \{\chi^H \,:\, H \text{ is triangle-free in } G\}.$$

Hence the theta bodies of $I_{\text{tf}}(G)$ provide convex relaxations of the triangle-free subgraph polytope $P_{\text{tf}}(G)$. From the general framework in Section 7.4.2, $\mathcal{B} = \{x^H + I_{\text{tf}}(G) \,:\, H \text{ triangle-free in } G\}$ is a θ-basis of $\mathbb{R}[x]/I_{\text{tf}}(G)$. Therefore, the rows and columns of $M_{\mathcal{B}_k}(y)$ are indexed by the triangle-free subgraphs in G with at most k edges. For ease of exposition, let us denote the entry of $M_{\mathcal{B}_k}(y)$ corresponding to row indexed by x^{H_1} and column indexed by x^{H_2} by $M_{\mathcal{B}_k}(y)_{H_1 H_2}$, let $H_1 \cup H_2$ denote the subgraph of G whose edge set is the union of the edge sets of H_1 and H_2, and y_H denote the entry of $y \in \mathbb{R}^{\mathcal{B}}$ corresponding to the basis element $x^H + I_{\text{tf}}(G)$. Then

$$\text{TH}_k(I_{\text{tf}}(G)) = \left\{ y \in \mathbb{R}^E \ : \ \begin{array}{l} \exists\, M \succeq 0,\ M \in \mathbb{R}^{|\mathcal{B}_k|\times|\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{H_1H_2} = \begin{cases} 0 \text{ if } & H_1 \cup H_2 \text{ has a triangle} \\ y_{H_1\cup H_2} & \text{otherwise} \end{cases} \end{array} \right\}.$$

Since all subgraphs of G with at most two edges are triangle-free, and $\mathcal{B}_1 = \{1 + I_{\text{tf}}(G)\} \cup \{x_e + I_{\text{tf}}(G) \ : \ e \in E\}$, $\text{TH}_1(I_{\text{tf}}(G))$ is exactly the same as the first theta body of the ideal $\langle x_e^2 - x_e \ : \ e \in E\rangle$ which is TH_1-exact by Theorem 7.30. Hence $\text{TH}_1(I_{\text{tf}}(G)) = [0,1]^E$, and $I_{\text{tf}}(G)$ is TH_1-exact if and only if every subgraph of G is triangle-free, or equivalently, G is triangle-free.

For graphs G that contain triangles, the second theta body of $I_{\text{tf}}(G)$ is more interesting as triples and quadruples of edges in G can contain triangles which forces some of the entries in $M_{\mathcal{B}_2}(y)$ to be zero.

Example 7.60. Suppose $G = K_3$ with edges labeled $1, 2, 3$. Then $P_{\text{tf}}(G)$ is the convex hull of all $0/1$ vectors in $\mathbb{R}^3$ except $(1, 1, 1)$ which is the first polytope shown in the second row of polytopes in Figure 7.10. This polytope is TH_2-exact since

$$\mathcal{B}_2 = \{1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3\} + I_{\text{tf}}(G) = \mathcal{B}.$$

Denoting $y \in \mathbb{R}^{\mathcal{B}_2}$, with first entry one, to be $y = (1, y_1, y_2, y_3, y_{12}, y_{13}, y_{23})$, we have that

$$M_{\mathcal{B}_2}(y) = \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} \\ y_1 & y_1 & y_{12} & y_{13} & y_{12} & y_{13} & 0 \\ y_2 & y_{12} & y_2 & y_{23} & y_{12} & 0 & y_{23} \\ y_3 & y_{13} & y_{23} & y_3 & 0 & y_{13} & y_{23} \\ y_{12} & y_{12} & y_{12} & 0 & y_{12} & 0 & 0 \\ y_{13} & y_{13} & 0 & y_{13} & 0 & y_{13} & 0 \\ y_{23} & 0 & y_{23} & y_{23} & 0 & 0 & y_{23} \end{pmatrix}.$$

Hence the triangle-free subgraph polytope of K_3 has the spectrahedral description $P_{\text{tf}}(G) = \{(y_1, y_2, y_3) \ : \ M_{\mathcal{B}_2}(y) \succeq 0\}$. ∎

Several families of facet inequalities for the triangle-free subgraph polytope of a graph can be found in the literature, and a complete facet description of $P_{\text{tf}}(G)$ for an arbitrary graph is unknown. An easy class of facets of $P_{\text{tf}}(G)$ come from the obvious fact that in any triangle in G at most two edges can be in a triangle-free subgraph. Mathematically, if $a, b, c \in E$ induce a triangle in G, then $2 - x_a - x_b - x_c \geq 0$ is a valid inequality for $P_{\text{tf}}(G)$. We now show that this inequality is valid for $\text{TH}_2(I_{\text{tf}}(G))$. First check that

$$(1 - x_c - x_ax_b) \equiv (1 - x_c - x_ax_b)^2 \bmod I_{\text{tf}}(G)$$

and also

$$(1 - x_a - x_b + x_ax_b) \equiv (1 - x_a - x_b + x_ax_b)^2 \bmod I_{\text{tf}}(G).$$

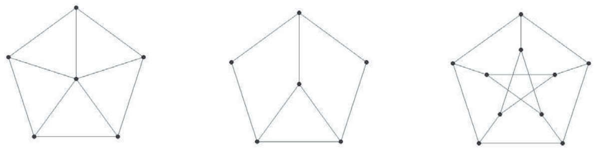

Figure 7.19. 5-*wheel, partial* 5-*wheel, and Petersen graph.*

This implies that $2 - x_a - x_b - x_c = (1 - x_a - x_b + x_a x_b) + (1 - x_c - x_a x_b)$ is 2-sos mod $I_{\mathrm{tf}}(G)$ and hence $2 - x_a - x_b - x_c \geq 0$ is valid for $\mathrm{TH}_2(I_{\mathrm{tf}}(G))$.

Exercise 7.61. We saw in Example 7.59 how to compute $\vartheta(G)$ numerically for a graph G. Find $\vartheta(G)$ for the graphs in Figure 7.19.

1. G a 5-wheel;

2. G the 5-wheel with two missing nonconsecutive rays;

3. G the Petersen graph.

Exercise 7.62. Compute the value of $\vartheta(G)$ for the 5-cycle exactly. (Hint: take advantage of the symmetries of the graph.)

Exercise 7.63. Prove that for any graph G, $\mathrm{TH}_1(I_G) \subseteq \mathrm{QSTAB}(G)$. Note that it is enough to prove that x_i and $1 - \sum_{i \in C} x_i$ are 1-sos mod I_G for all vertices i and all cliques C.

Exercise 7.64. It is known that the stable set polytope of C_{2k+1}, the odd cycle of $2k+1$ nodes, is defined by the inequalities $x_i \geq 0$ for all $i \in [2k+1]$, $x_i + x_j \leq 1$ for all $\{i, j\} \in E$, which by the previous exercise are 1-sos mod I_G, and the single *odd cycle inequality* $\sum x_i \leq k$ [32, Corollary 65.12a].

1. Show that C_5 is TH_2-exact.

2. Show that C_{2k+1} is TH_2-exact for all k.

Exercise 7.65. In Exercise 7.55 we have shown that the vanishing ideal of the set of vertices of a $(k+1)$-level polytope is TH_k-exact. We also have seen in Theorem 7.30 that the reverse implication is true for $k = 1$: if a real radical ideal is TH_1-exact, then its variety must be the set of vertices of a 2-level polytope. Using what we know of the theta body approximations to the stable set polytope, show that the reverse implication (TH_k-exact $\Rightarrow$ k-level) fails for $k \geq 2$.

Exercise 7.66. The triangle-free subgraph problem is closely related to another important problem in combinatorial optimization, the K_3-cover subgraph problem.

A subgraph of G is said to be a K_3-cover if it contains at least an edge of every triangle of G. What is the relation between a maximum triangle-free subgraph and a minimum K_3-cover? How is that reflected in the polytopes underlying those combinatorial problems?

Exercise 7.67. A $(2k+1)$-odd wheel is the graph on $2k+2$ vertices with $2k+1$ of the vertices forming a $2k+1$-cycle and the last vertex connected to each of the vertices of the cycle. Such a wheel yields the inequality $\sum_{e \in EW} x_e \leq 3k+1$ that is valid for the triangle-free subgraph polytope of G. For example, an induced 5-wheel in a graph gives the inequality

$$x_{12} + x_{23} + x_{34} + x_{45} + x_{15} + x_{16} + x_{26} + x_{36} + x_{46} + x_{56} \leq 7,$$

which is valid for the triangle-free subgraph polytope of the graph.

1. Use YALMIP to see that the 5-wheel and 7-wheel inequalities appear to be 2-sos mod $I_{\text{tf}}(G)$, where G is the corresponding wheel.
2. Can you express them exactly as 2-sos modulo the ideals?
3. Can you prove that all odd wheel inequalities are 2-sos modulo its ideal?

Exercise 7.68. Another version of the triangle-free subgraph problem is vertex-based. Given a subset of nodes of G we say it is triangle-free if its induced subgraph is triangle-free. This also falls into the simplicial complex model, so we know how to construct reduced moment matrices. Using the first theta body, compute an approximation for the maximum triangle-free subset of nodes of the 4-wheel.

Bibliography

[1] Y. H. Au and L. Tunçel. Complexity analyses of Bienstock-Zuckerberg and Lasserre relaxations on the matching and stable set polytopes. In *Integer Programming and Combinatorial Optimization*, Lecture Notes in Comput. Sci. 6655, Springer, Heidelberg, 2011, pp. 14–26.

[2] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Math. Program.*, 58:295–324, 1993.

[3] F. Bendali, A. R. Mahjoub, and J. Mailfert. Composition of graphs and the triangle-free subgraph polytope. *J. Comb. Optim.*, 6:359–381, 2002.

[4] M. Chudnovsky, N. Robertson, P. Seymour, and R. R. Thomas. The strong perfect graph theorem. *Ann. of Math.* (2), 164:51–229, 2006.

[5] M. Conforti, D. G. Corneil, and A. R. Mahjoub. K_i-covers. I. Complexity and polytopes. *Discrete Math.*, 58:121–142, 1986.

[6] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Springer-Verlag, New York, 1992.

[7] J. Gouveia, M. Laurent, P. A. Parrilo, and R. R. Thomas. A new hierarchy of semidefinite programming relaxations for cycles in binary matroids and cuts in graphs. *Math. Program., Ser. A*, 2010, to appear.

[8] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20:2097–2118, 2010.

[9] J. Gouveia and R. R. Thomas. Convex hulls of algebraic sets. In M. Anjos and J.-B. Lasserre, editors, *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, to appear.

[10] D. Grayson and M. Stillman. Macaulay 2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2.

[11] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, 2nd edition, Algorithms Combin. Springer-Verlag, Berlin, 1993.

[12] R. Harvey and H. B. Lawson, Jr. Calibrated geometries. *Acta Math.*, 148:47–157, 1982.

[13] D. Henrion. Semidefinite representation of convex hulls of rational varieties. LAAS-CNRS Research Report 09001, 2009.

[14] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11:796–817, 2001.

[15] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28:470–496, 2003.

[16] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, IMA Vol. Math. Appl. 149. Springer, Berlin, 2009.

[17] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[18] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25:1–7, 1979.

[19] L. Lovász. Semidefinite programs and combinatorial optimization. In *Recent Advances in Algorithms and Combinatorics*, CMS Books Math./Ouvrages Math. SMC 11. Springer, New York, 2003, pp. 137–194.

[20] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1:166–190, 1991.

[21] M. Marshall. *Positive Polynomials and Sums of Squares*, Math. Surveys Monogr. 146. American Mathematical Society, Providence, RI, 2008.

[22] J. Nie. First order conditions for semidefinite representations of convex sets defined by rational or singular polynomials. *Math. Program.*, 131:1–36, 2012.

[23] M. Omar and B. Osserman. Strong nonnegativity and sums of squares on real varieties. arXiv:1101.0826.

[24] R. T. Rockafellar. *Convex Analysis*, Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton, NJ, 1996.

[25] P. Rostalski. Bermeja, Software for Convex Algebraic Geometry. Available at http://math.berkeley.edu/∼philipp/Software/Software.

[26] R. Sanyal. Orbitopes and theta bodies. Talk at IPAM Workshop on Convex Optimization and Algebraic Geometry, slides available at http://math.berkeley.edu/∼bernd/raman.pdf, 2010.

[27] C. Scheiderer. Sums of squares of regular functions on real algebraic varieties. *Trans. Amer. Math. Soc.*, 352:1039–1069, 2000.

[28] C. Scheiderer. Sums of squares on real algebraic curves. *Math. Z.*, 245:725–760, 2003.

[29] C. Scheiderer. Sums of squares on real algebraic surfaces. *Manuscripta Math.*, 119:395–410, 2006.

[30] C. Scheiderer. Convex hulls of curves of genus one. *Adv. Math.*, 228:2606–2622, 2011.

[31] A. Schrijver. *Theory of Linear and Integer Programming*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 1986.

[32] A. Schrijver. *Combinatorial Optimization. Polyhedra and Efficiency. Vol. B*, Algorithms Combin. 24. Springer-Verlag, Berlin, 2003.

[33] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3:411–430, 1990.

[34] R. P. Stanley. Decompositions of rational convex polytopes. *Ann. Discrete Math.*, 6:333–342, 1980.

[35] S. Sullivant. Compressed polytopes and statistical disclosure limitation. *Tohoku Math. J.* (2), 58:433–445, 2006.

[36] M. Yannakakis. Edge-deletion problems. *SIAM J. Comput.*, 10:297–309, 1981.