Available online at www.sciencedirect.com
ScienceDirect
Advances in Mathematics 228 (2011) 2606–2622
ADVANCES IN Mathematics
www.elsevier.com/locate/aim

Convex hulls of curves of genus one

Claus Scheiderer [1]

Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany

Received 27 April 2010; accepted 19 July 2011

Available online 6 August 2011

Communicated by Ezra Miller

Abstract

Let C be a real nonsingular affine curve of genus one, embedded in affine n-space, whose set of real points is compact. For any polynomial f which is nonnegative on $C(\mathbb{R})$, we prove that there exist polynomials f_i with $f \equiv \sum_i f_i^2 \pmod{\mathcal{I}_C}$ and such that the degrees $\deg(f_i)$ are bounded in terms of $\deg(f)$ only. Using Lasserre's relaxation method, we deduce an explicit representation of the convex hull of $C(\mathbb{R})$ in $\mathbb{R}^n$ by a lifted linear matrix inequality. This is the first instance in the literature where such a representation is given for the convex hull of a nonrational variety. The same works for convex hulls of (singular) curves whose normalization is C. We then make a detailed study of the associated degree bounds. These bounds are directly related to size and dimension of the projected matrix pencils. In particular, we prove that these bounds tend to infinity when the curve C degenerates suitably into a singular curve, and we provide explicit lower bounds as well.

MSC: primary 14P05; secondary 14H52, 90C22

Keywords: Real algebraic curves; Convex hulls; Elliptic curves; Linear matrix inequalities; Spectrahedra; Lasserre relaxation

E-mail address: claus.scheiderer@uni.constanz.de.

URL: http://www.math.uni-konstanz/~scheider.

[1] The author is indebted to Rekha Thomas for helpful conversations around the subject of this article. He profited from recent invitations to workshops at AIM and BIRS and would like to thank both institutions.

0001-8708/$ – see front matter
doi:10.1016/j.aim.2011.07.014

0. Introduction

Let $V \subset \mathbb{A}^n$ be an affine algebraic variety over $\mathbb{R}$ whose set $V(\mathbb{R})$ of real points is compact. The convex hull of $V(\mathbb{R})$ in $\mathbb{R}^n$ is a compact semi-algebraic set. Recently there has been a growing interest in describing this set, or its boundary, from different perspectives, see [16,8,4,22,20]. Part of the motivation comes from potential applications in semidefinite programming. If A_i $(i = 0, \ldots, n)$ are symmetric real matrices of some fixed size, an inequality

$$A_0 + x_1 A_1 + \cdots + x_n A_n \succcurlyeq 0$$

is called a linear matrix inequality (LMI) in the variables $x_1, \ldots, x_n$. (Here $\succcurlyeq$ denotes positive semidefiniteness of the matrix.) The set K of $x \in \mathbb{R}^n$ which satisfy the LMI is a basic closed and convex semi-algebraic subset of $\mathbb{R}^n$. From the view point of convex optimization, such a description is very useful since it allows quick and efficient optimization of linear functions on K, see e.g. [13,2,12].

Convex sets which allow an LMI representation are also called spectrahedra. Being a spectrahedron is a restrictive property, since these sets are not only basic closed but also rigidly convex, a property that is much stronger than just convexity [7]. In dimension $\leqslant 2$, rigid convexity characterizes spectrahedra [7, Thm. 2.2]. In higher dimensions it is currently unknown whether such a converse holds. For optimization purposes, however, a linear projection of a spectrahedron is just as good as a spectrahedron. That $K \subset \mathbb{R}^n$ is a projected spectrahedron means that there exist symmetric real matrices A_i $(0 \leqslant i \leqslant n)$ and B_j $(1 \leqslant j \leqslant m)$ such that K is the set of $x \in \mathbb{R}^n$ for which there exists $y \in \mathbb{R}^m$ with

$$A_0 + \sum_{i=1}^{n} x_i A_i + \sum_{j=1}^{m} y_j B_j \succcurlyeq 0.$$

One speaks of a lifted LMI representation of K, or of a semidefinite (SDP) representation. Projected spectrahedra form a much wider class than spectrahedra, and much research effort is currently spent on understanding their properties, e.g. [13,12,9,5,6,15,14,3]. In fact, Helton and Nie [5] have conjectured that every convex semi-algebraic set allows a lifted LMI representation.

Obtaining explicit lifted LMI representations for concretely given convex sets is a different matter. A general construction, called the relaxation method, is due to Lasserre [9] and applies in many cases. We will recall it (in specialized form) in Section 1 below. Other constructions are due to Helton and Nie [5,6], who proved the existence of lifted LMI representations for several large classes of convex sets.

Here we are interested in applying Lasserre's construction to the convex hull of a (compact) real algebraic curve C in the affine plane or in some higher-dimensional space. The key properties that are needed to make the relaxation method work are a partial stability property and a partial saturation property, each for the cone of sums of squares in the coordinate ring $\mathbb{R}[C]$ (see Section 1). Namely, every linear polynomial that is nonnegative on the curve has to be a sum of squares in $\mathbb{R}[C]$ with uniformly bounded degrees.

Our results apply when the curve C is nonsingular of genus one and its real part $C(\mathbb{R})$ is compact. It has been known for some time that every psd element in $\mathbb{R}[C]$ is a sum of squares. We prove that the sums of squares cone in $\mathbb{R}[C]$ is stable, which is our main result (Theorems 2.1, 2.13). The proof uses algebraic–geometric methods, and unfortunately it seems to be restricted to

genus one. No similar result is known for any curve of genus > 1 (with compact real points). On the other hand, our result gives the first construction of a lifted LMI representation for the convex hull of a nonrational real algebraic variety. We illustrate the application to such representations by means of some concrete examples (Section 3).

Since the sizes of the lifted LMI representations depend directly on the stability (degree) bounds, there exist good reasons to study these bounds in more detail. This is mainly done in Section 4. We succeed in making the bounds fairly explicit, and in a sense we arrive at the best possible bounds. As a result, we can make the lifted LMI representations completely explicit for many curves. We also study how the bounds change under variation of the curve, and we prove that they tend to infinity when the curve gets degenerated to a singular (rational) curve.

1. Convex hulls of algebraic sets and Lasserre relaxation

We give a brief review here of Lasserre's relaxation method for the construction of lifted LMI representations, however only in the special case which will be used later, to keep the exposition less technical.

1.1. For the following discussion, A can be any finitely generated $\mathbb{R}$-algebra. Let $V = \operatorname{Spec}(A)$ be the associated affine $\mathbb{R}$-variety. The set $V(\mathbb{R}) = \operatorname{Hom}_{\mathbb{R}}(A, \mathbb{R})$ of $\mathbb{R}$-algebra homomorphisms has a natural euclidean topology, namely the topology induced by the inclusion $V(\mathbb{R}) \hookrightarrow \mathbb{R}^n$, $p \mapsto (x_1(p), \ldots, x_n(p))$, where $x_1, \ldots, x_n$ is any system of generators of A. This embedding identifies $V(\mathbb{R})$ with a (closed) real algebraic subset of $\mathbb{R}^n$. As usual, we think of the elements $p \in V(\mathbb{R})$ as points and denote the pairing between $f \in A$ and $p \in V(\mathbb{R})$ by $f(p)$.

1.2. Let ΣA^2 denote the cone of sums of squares in A. By

$$A_+ = \big\{ f \in A\colon\ \forall p \in V(\mathbb{R})\ f(p) \geqslant 0 \big\}$$

we denote the cone of all *positive semidefinite (psd)* elements of A. Given any finite-dimensional linear subspace L of A, one can ask two questions:

(1) Is $L \cap A_+$ contained in ΣA^2 (and hence equal to $L \cap \Sigma A^2$)?
(2) Does there exist a finite-dimensional linear subspace W of A such that every $f \in L \cap \Sigma A^2$ can be written as $f = \sum_{i=1}^r a_i^2$ with $r \in \mathbb{N}$ and $a_1, \ldots, a_r \in W$?

Recall that the preordering ΣA^2 is called *saturated* if $A_+ = \Sigma A^2$ [23,26]. Therefore, a positive answer to (1) can be regarded as a partial saturatedness property of ΣA^2. On the other hand, ΣA^2 is called *stable* if (2) has a positive answer for any finite-dimensional L [19,25]. Therefore, a positive answer to (2) means a partial stability property of ΣA^2.

Remark 1.3. Assume we are fixing a system of generators of A, so that $A = \mathbb{R}[\mathrm{x}]/I$ for some ideal I of $\mathbb{R}[\mathrm{x}]$, where $\mathrm{x} = (x_1, \ldots, x_n)$ is a tuple of variables. For $d \geqslant 0$ let $\mathbb{R}[\mathrm{x}]_d$ be the space of polynomials of total degree $\leqslant d$ in $\mathbb{R}[\mathrm{x}]$, and put $A_d = (\mathbb{R}[\mathrm{x}]_d + I)/I$. Given integers d, $k \geqslant 0$, the ideal I is said to be (d, k)-*sos* in [4] if (1) and (2) hold for $L = A_d$ and $W = A_k$. The problem of characterizing the $(1, k)$-sos ideals in $\mathbb{R}[\mathrm{x}]$, and in particular the $(1, 1)$-sos ideals, was raised by Lovász [10], who showed that this question for certain 0-dimensional ideals is closely related to the stable set problem for graphs.

1.4. We now recall Lasserre's important relaxation construction [9]. Assume $A = \mathbb{R}[\mathrm{x}]/I$ for some ideal I of $\mathbb{R}[\mathrm{x}]$, where $\mathrm{x} = (x_1, \ldots, x_n)$. We denote the zero set of I in $\mathbb{R}^n$ by $V_{\mathbb{R}}(I)$. For convenience of exposition let us assume that I does not contain any nonzero polynomial of degree $\leqslant 1$.

Let $L = A_1 = \{f + I\colon\ f \in \mathbb{R}[\mathrm{x}],\ \deg(f) \leqslant 1\} \subset A$, and let W be some finite-dimensional linear subspace of A containing L. Let U be the linear subspace of A generated by all squares a^2 with $a \in W$; clearly $L \subset U$ and $\dim(U) < \infty$. Let $\rho : U^\vee \to L^\vee$ be the restriction map between the dual linear spaces induced by the inclusion $L \subset U$. Moreover, let $U_1^\vee$ (resp. $L_1^\vee$) denote the set of all linear forms λ in $U^\vee$ (resp. in $L^\vee$) with $\lambda(1) = 1$. Then $L_1^\vee$ is canonically identified with $\mathbb{R}^n$ via $\lambda \leftrightarrow (\lambda(\bar{x}_1), \ldots, \lambda(\bar{x}_n))$ where $\bar{x}_i := x_i + I$, and we always consider $V_{\mathbb{R}}(I)$ as a real algebraic subset of $L_1^\vee = \mathbb{R}^n$ in the natural way.

Let $M_W = \{\sum_{i=1}^r a_i^2\colon\ r \in \mathbb{N},\ a_1, \ldots, a_r \in W\} \subset A$ denote the set of sums of squares of elements of W. This is a convex cone in U, which is closed in U if I is a real radical ideal [19, Prop. 2.6]. We'll denote the dual of a convex cone C by C^*, so M_W^* is the dual cone of M_W in $U^\vee$. Then M_W^* is a spectrahedron in $U^\vee$, which means that M_W^* can be defined in $U^\vee$ by a (homogeneous) linear matrix inequality. Indeed, for $\mu \in U^\vee$ the symmetric bilinear form $\beta(\mu) : W \times W \to \mathbb{R}$, $(a, a') \mapsto \mu(aa')$ depends linearly on μ, and by definition it is psd if and only if $\mu \in M_W^*$.

The subset $M_W^* \cap U_1^\vee$ of M_W^* is an affine-linear section of M_W^*, and is therefore a spectrahedron as well. Its image

$$K_W := \rho\big(M_W^* \cap U_1^\vee\big) = L_1^\vee \cap \rho\big(M_W^*\big)$$

under the restriction map $\rho : U_1^\vee \to L_1^\vee = \mathbb{R}^n$ is a convex semi-algebraic subset of $L_1^\vee = \mathbb{R}^n$. For any point $p \in V_{\mathbb{R}}(I)$, evaluation in p is a linear form $\lambda_p \in M_W^* \cap U_1^\vee$ for which tautologically $\rho(\lambda_p) = p$ holds; this shows that K_W contains $V_{\mathbb{R}}(I)$. By construction, K_W is a linear projection of a spectrahedron. Increasing W results in decreasing K_W, so by making W larger and larger (of finite dimension) one gets a shrinking family of convex sets K_W which all contain $V_{\mathbb{R}}(I)$. For ease of exposition let us assume that the ideal I is real radical. Then the closure $\overline{K_W}$ is equal to $L_1^\vee \cap (L \cap M_W)^*$. Moreover, $\overline{K_W} = \overline{\operatorname{conv} V_{\mathbb{R}}(I)}$ holds if, and only if, L and W satisfy conditions (1) and (2) of 1.2. (See [9, Thm. 2] and [15, Prop. 3.1].) If these conditions are fulfilled, and if the convex hull of $V_{\mathbb{R}}(I)$ is closed, we have obtained an explicit representation of $\operatorname{conv} V_{\mathbb{R}}(I) = K_W$ by a lifted LMI. Note that $\operatorname{conv} V_{\mathbb{R}}(I)$ will be automatically closed if the real algebraic set $V_{\mathbb{R}}(I)$ is compact.

1.5. We keep the assumptions and notations of 1.4. Let us take $W = A_k = (\mathbb{R}[\mathrm{x}]_k + I)/I$ for some $k \geqslant 1$, and form the associated projected spectrahedron K_{A_k} as before. By 1.4, the ideal I is $(1, k)$-sos (see 1.3) if, and only if, $\overline{\operatorname{conv} V_{\mathbb{R}}(I)} = \overline{K_{A_k}}$. The *k-th theta body* of the ideal I, defined in [4] as

$$\mathrm{TH}_k(I) = \big\{x \in \mathbb{R}^n\colon\ \forall f \in L \cap M_{A_k}\ f(x) \geqslant 0\big\},$$

is (by definition) equal to $L_1^\vee \cap (L \cap M_{A_k})^*$, and is therefore equal to $\overline{K_{A_k}}$. The ideal I is said to be TH_k*-exact* in [4] if $\mathrm{TH}_k(I)$ is the closure of $\operatorname{conv} V_{\mathbb{R}}(I)$. We see that this is the case if and only if I is $(1, k)$-sos (assuming I real radical), see [4, Prop. 2.8].

1.6. Conditions (1) and (2) of 1.2 are also of interest for finite-dimensional subspaces L of A other than $L = A_1$. Given an arbitrary such subspace L (containing 1), let A' be the $\mathbb{R}$-subalgebra generated by L, and let $V' = \mathrm{Spec}(A')$. If $1 = u_0, u_1, \ldots, u_m$ is a vector space basis of L, then $x \mapsto (u_1(x), \ldots, u_m(x))$ is a closed embedding of V' into affine m-space. If there exists a finite-dimensional subspace W of A satisfying (1) and (2), and if $V'(\mathbb{R})$ is compact, we get a representation of the convex hull of $V'(\mathbb{R})$ in $\mathbb{R}^m$ as a projected spectrahedron.

Remark 1.7. In the above discussion we only considered sums of squares, corresponding on the geometric side to convex hulls of real algebraic sets in $\mathbb{R}^n$. We did so to simplify the exposition, and since the main results of this paper only concern this case. Note however that both the setup and the results of Lasserre relaxation generalize well to arbitrary finitely generated quadratic modules. On the geometric side, this corresponds to convex hulls of basic closed semi-algebraic sets. See also [15] and [3] for more details.

2. Stability of sums of squares

From now on we will consider affine algebraic curves (mostly nonsingular) of genus one. By the genus of a real curve which is irreducible over $\mathbb{C}$, we mean the (geometric) genus of its nonsingular projective model. In this section we will prove:

Theorem 2.1. *Let C be an irreducible nonsingular affine curve of genus one over $\mathbb{R}$ which has at least one pair of conjugate nonreal points at infinity. Then the preordering of sums of squares in $\mathbb{R}[C]$ is stable and saturated.*

Remarks 2.2.

1. See 1.2 for the meaning of stable or saturated. Theorem 2.1 says that questions (1) and (2) in 1.2 have a positive answer for any finite-dimensional linear subspace L of $\mathbb{R}[C]$. Therefore, if $C(\mathbb{R})$ is compact, the relaxation construction 1.4 applies and gives lifted LMI representations of the convex hull of $C(\mathbb{R})$ for any closed embedding of C into affine space. We will discuss these applications in more detail in Section 3 below.
2. That $\Sigma\mathbb{R}[C]^2$ is saturated, i.e. that psd = sos holds on C, was already proved in [23] (and again, in much greater generality, in [24]). A special case of the stability part of Theorem 2.1 was mentioned in [19] (Example 2.17) without proof. The key argument was sketched in [18] (unpublished).

2.3. We always denote the nonsingular projective completion of C by $\overline{C}$. The (geometric) points of C at infinity are by definition the points in $\overline{C}(\mathbb{C}) \setminus C(\mathbb{C})$. The condition that C has at least one nonreal point at infinity says that at least one among these points is not real.

By assumption on C, there exists a pair $\infty \neq \overline{\infty}$ of conjugate nonreal points in $\overline{C}(\mathbb{C}) \setminus C(\mathbb{C})$. Let $C_0 = \overline{C} \setminus \{\infty, \overline{\infty}\}$, then C_0 is an affine curve over $\mathbb{R}$ that contains C as a Zariski open subset. Every psd element of $\mathbb{R}[C_0]$ is a sum of squares in $\mathbb{R}[C_0]$ (see [24]). By the following lemma, it suffices to prove Theorem 2.1 for C_0 instead of C:

Lemma 2.4. *Let C_0 be an affine curve over $\mathbb{R}$, and let C be a Zariski open subset of C_0. Assume that every psd element of $\mathbb{R}[C_0]$ is a sum of squares in $\mathbb{R}[C_0]$. If the preordering of sums of squares in $\mathbb{R}[C_0]$ is stable, then the preordering of sums of squares in $\mathbb{R}[C]$ is stable as well.*

Proof. There exists $s \in \mathbb{R}[C_0]$ such that $\mathbb{R}[C] = \mathbb{R}[C_0]_s$, the ring of fractions $\frac{f}{s^n}$ with $f \in \mathbb{R}[C_0]$ and $n \geqslant 0$. Let L be a finite-dimensional subspace of $\mathbb{R}[C]$, and choose $n \geqslant 0$ such that $s^{2n}L =: L_0$ is contained in $\mathbb{R}[C_0]$. Since $\Sigma\mathbb{R}[C_0]^2$ is stable in $\mathbb{R}[C_0]$, there is a finite-dimensional subspace W_0 of $\mathbb{R}[C_0]$ such that every element of $L_0 \cap \Sigma\mathbb{R}[C_0]^2$ is a sum of squares of elements of W_0. Now let $f \in L \cap \Sigma\mathbb{R}[C]^2$. Then $s^{2n}f$ lies in L_0, and it is psd on $C_0(\mathbb{R})$ since C_0 has no isolated real points (the latter by [24, Thm. 4.18]). By assumption, therefore, $s^{2n}f \in \Sigma\mathbb{R}[C_0]^2$, hence $s^{2n}f$ is a sum of squares of elements of W_0. So if we put $W := s^{-n}W_0$, every element of $L \cap \Sigma\mathbb{R}[C]^2$ is a sum of squares of elements of W. □

2.5. So it suffices to consider a nonsingular affine curve C of genus one over $\mathbb{R}$ with precisely one pair $\infty \neq \overline{\infty}$ of complex conjugate points at infinity. Note that this implies that $C(\mathbb{R})$ is compact. It follows from Riemann–Roch that C is isomorphic to a plane affine curve with equation $y^2 + q(x) = 0$, where (x, y) are plane affine coordinates and $q(x) \in \mathbb{R}[x]$ is a monic polynomial of degree 4 without multiple roots. We can also assume that $q(x)$ is indefinite, i.e. has (2 or 4) real roots, since otherwise $C(\mathbb{R})$ is empty (in which case the theorem is both true and uninteresting). Note that $\overline{C}$, the nonsingular projective model of C, is the normalization of the Zariski closure of C in $\mathbb{P}^2$, and is an elliptic curve over $\mathbb{R}$.

Conversely, every plane affine curve over $\mathbb{R}$ with equation $y^2 + q(x) = 0$ with q monic and separable of degree four is nonsingular of genus one and has precisely two complex conjugate points at infinity.

2.6. From now on C will always be a curve as in 2.5. Usually we shall not distinguish in our notation between a polynomial $f \in \mathbb{R}[x, y]$ and its restriction to C (i.e. the image under the canonical map $\mathbb{R}[x, y] \to \mathbb{R}[C]$). Instead of working with the ordinary (total) degree of polynomials we will use a variant which is better adapted to the curve C:

Let $\mathbb{R}(C)$ be the (real) function field of C. Given any point $p \in C(\mathbb{C})$ we let $v_p : \mathbb{R}(C)^* \to \mathbb{Z}$ be the associated discrete valuation of $\mathbb{R}(C)$. Given $f \in \mathbb{R}(C)$, we'll write

$$\delta(f) := -v_\infty(f) = -v_{\overline{\infty}}(f)$$

(putting $\delta(0) := -\infty$). So δ is the negative of a discrete valuation on $\mathbb{R}(C)$. For any $n \geqslant 1$, the elements x^i $(0 \leqslant i \leqslant n)$ and $x^j y$ $(0 \leqslant j \leqslant n-2)$ form a linear basis of the subspace $\{f \in \mathbb{R}[C]\colon \delta(f) \leqslant n\}$ of $\mathbb{R}[C]$.

In [23, Section 4], it was proved that every psd element of $\mathbb{R}[C]$ is a sum of squares in $\mathbb{R}[C]$. (At that time general results like [24, Thm. 4.18] were not yet available.) In order to prove the stability result of Theorem 2.1, we first need to review a part of the proof from [23] and analyze the involved δ-degrees. This is done in the next lemma:

Lemma 2.7. *Let $0 \neq f \in \mathbb{R}[C]$ be psd on $C(\mathbb{R})$, and assume that f has at least one nonreal zero in $C(\mathbb{C})$. Then there exist $g_1, g_2 \in \mathbb{R}[C]$ with*

(a) $\delta(f - g_1^2 - g_2^2) \leqslant \delta(f)$;
(b) $f - g_1^2 - g_2^2$ *has strictly less nonreal zeros than* f *in* $C(\mathbb{C})$*, or is identically zero;*
(c) $\delta(g_1), \delta(g_2) \leqslant \lceil \frac{1}{2}\delta(f) \rceil$.

Here the zeros of $0 \neq f \in \mathbb{R}[C]$ in $C(\mathbb{C})$ are counted with multiplicities. As usual we write $\lceil x \rceil = \min\{n \in \mathbb{Z}\colon\ x \leqslant n\}$ for $x \in \mathbb{R}$.

Proof of Lemma 2.7. Let $m = \delta(f) \geqslant 1$. All divisors are calculated on the complexified curve $\overline{C}_{\mathbb{C}}$, i.e. they are finite integral linear combinations of the points in $\overline{C}(\mathbb{C})$. We write

$$\operatorname{div}(f) = 2D + \Theta - m(\infty + \overline{\infty})$$

where D, Θ are conjugation-invariant effective divisors such that the support of D contains only real points and the support of Θ contains no real point. By hypothesis $\Theta \neq 0$. Let $p = q = \frac{m}{2}$ if m is even, and put $p = \frac{m+1}{2}$, $q = \frac{m-1}{2}$ if m is odd. Then $p + q = m$, and the divisor $E := -D + p\infty + q\overline{\infty}$ satisfies $\deg(E) = \frac{1}{2}\deg(\Theta) \geqslant 1$. By Riemann–Roch there exists $g \in \mathbb{C}(C)^*$ with $\operatorname{div}(g) + E \geqslant 0$, hence with

$$\operatorname{div}(g\overline{g}) \geqslant 2D - m(\infty + \overline{\infty}).$$

It follows that $g\overline{g} \in \mathbb{R}[C]$, and the rational function $\varphi := g\overline{g}/f$ on C has no poles in $C(\mathbb{R})$. Let $c > 0$ be the maximum value that φ takes on the compact set $C(\mathbb{R})$, say $\varphi(p) = c$ with $p \in C(\mathbb{R})$. The regular function $h := f - \frac{1}{c}\, g\overline{g}$ on C is psd on $C(\mathbb{R})$ and vanishes at p. From $\delta(g\overline{g}) \leqslant m$ we see $\delta(h) \leqslant m$. Writing $\frac{1}{\sqrt{c}} g = g_1 + i g_2$ with $g_1, g_2 \in \mathbb{R}[C]$ we have $\frac{1}{c}\, g\overline{g} = g_1^2 + g_2^2$, and we see

$$\delta(g_1), \delta(g_2) \leqslant \max\{p, q\} = \left\lceil \frac{m}{2} \right\rceil.$$

For every point $q \in C(\mathbb{R})$ we have $v_q(h) \geqslant v_q(f)$, and even $v_p(h) \geqslant 2 + v_p(f)$ if $q = p$. Counting with multiplicity, h has therefore strictly more real zeros on C than f. Since $\delta(h) \leqslant \delta(f)$, we see that h has strictly less nonreal zeros than f (or else $h = 0$). □

By applying Lemma 2.7 inductively, we obtain the following reduction to psd regular functions with only real zeros:

Proposition 2.8. *Let* $0 \neq f \in \mathbb{R}[C]$ *be psd. There are finitely many regular functions* $0 \neq g_1, \ldots, g_r \in \mathbb{R}[C]$ *(*$r \geqslant 0$*) such that*

(a) $h := f - (g_1^2 + \cdots + g_r^2)$ *is psd on* $C(\mathbb{R})$*;*
(b) $h = 0$*, or all zeros of* h *on* C *are real;*
(c) $\delta(g_i) \leqslant \lceil \frac{1}{2}\delta(f) \rceil$ *for* $i = 1, \ldots, r$*, and* $\delta(h) \leqslant \delta(f)$.

Remark 2.9. From Lemma 2.7 we see that the number r of squares in Proposition 2.8 can be bounded by the number of nonreal zeros of f in $C(\mathbb{C})$, counted with multiplicities. In other words, $r \leqslant 2(m - k)$ where $\delta(f) = m$ and $2k$ is the number of real zeros of f, counted with multiplicities. On the other hand, it is well known that the Pythagoras number of $\mathbb{R}[C]$ is $\leqslant 4$.

The second step consists in studying the nonnegative regular functions on C with only real zeros. We will see that part of the conclusions made in Example 2.19 for linear psd polynomials generalizes to psd polynomials of any degree on C. Recall that C has the affine equation $y^2 + q(x) = 0$ where the monic quartic polynomial $q(x) \in \mathbb{R}[x]$ is square-free and indefinite. Let $\alpha < \beta$ denote the smallest resp. the largest real zero of $q(x)$. Let $\mathbb{R}(C)$ be the function field of C.

Proposition 2.10. *Let G be the subgroup of $\mathbb{R}(C)^*/\mathbb{R}(C)^{*2}$ which is generated by the cosets $f\mathbb{R}(C)^{*2}$ of all psd $0 \neq f \in \mathbb{R}[C]$ which have only real zeros on C. Then G has order four and is generated by the cosets of $x - \alpha$ and of $\beta - x$.*

Proof. Since the square classes of $x - \alpha$ and $\beta - x$ lie in G and are independent, it is enough to show $|G| = 4$. This was done in [23, Prop. 4.3], where $|G|$ was calculated in a more general setting. □

Definition 2.11. Let $0 \neq f \in \mathbb{R}[C]$. By $\theta(f)$ we denote the least integer $d \geqslant 0$ for which there exists a sums of squares representation $f = f_1^2 + \cdots + f_r^2$ with $r \in \mathbb{N}$ and $f_i \in \mathbb{R}[C]$ such that $\delta(f_i) \leqslant d$ for $i = 1, \ldots, r$. We put $\theta(f) = \infty$ if f is not a sum of squares in $\mathbb{R}[C]$.

Note that one obviously has $\theta(f + g) \leqslant \max\{\theta(f), \theta(g)\}$ and $\theta(fg) \leqslant \theta(f) + \theta(g)$.

Lemma 2.12. *Let $0 \neq f, g \in \mathbb{R}[C]$ be psd. Assume that g has only real zeros on C and that f/g is a square in $\mathbb{R}(C)$. If $g = b_1^2 + \cdots + b_r^2$ with $b_1, \ldots, b_r \in \mathbb{R}[C]$, then there exist $a_1, \ldots, a_r \in \mathbb{R}[C]$ with $f = a_1^2 + \cdots + a_r^2$ and with*

$$\delta(a_i) = \delta(b_i) + \frac{1}{2}\big(\delta(f) - \delta(g)\big)$$

($i = 1, \ldots, r$). In particular we have $2\theta(f) - \delta(f) \leqslant 2\theta(g) - \delta(g)$.

Proof. Let $h \in \mathbb{R}(C)^*$ with $\frac{f}{g} = h^2$. We have $f = \sum_i (b_i h)^2$, so it suffices to show that $a_i := b_i h$ lies in $\mathbb{R}[C]$ and $\delta(a_i)$ satisfies the identity of the lemma ($i = 1, \ldots, r$). Every pole of a_i on C is a zero of g, so it is real by the assumption. On the other hand, for $i = 1, \ldots, r$ and for every point $p \in C(\mathbb{R})$ we have $v_p(g) \leqslant 2v_p(b_i)$, hence $v_p(h) \geqslant \frac{1}{2}v_p(f) - v_p(b_i)$ and $v_p(a_i) \geqslant \frac{1}{2}v_p(f) \geqslant 0$. This proves $a_i \in \mathbb{R}[C]$. Clearly $\delta(a_i) = \delta(b_i h) = \delta(b_i) + \frac{1}{2}(\delta(f) - \delta(g))$, and this implies $\theta(f) \leqslant \theta(g) + \frac{1}{2}(\delta(f) - \delta(g))$. □

In Lemma 2.12, note that we have in fact $\theta(f) - \theta(g) = \frac{1}{2}(\delta(f) - \delta(g))$ if both f and g have only real zeros.

This discussion leads to the following result. It completes the proof of Theorem 2.1:

Theorem 2.13. *Let q be a quartic monic polynomial which is indefinite and has no multiple roots, and let C be the affine curve $y^2 + q(x) = 0$ over $\mathbb{R}$. There is an integer $N \geqslant 1$ such that*

$$\theta(f) \leqslant N + \left\lceil \frac{1}{2}\delta(f) \right\rceil$$

holds for every psd regular function f in $\mathbb{R}[C]$.

Proof. Let α (resp. β) be the smallest (resp. largest) real zero of $q(x)$, and put $l_1 = x - \alpha$, $l_2 = \beta - x$. Each of l_1, l_2 and $l_1 l_2$ has only real zeros on C and is a sum of squares in $\mathbb{R}[C]$ [23, Thm. 4.10(a)]. We claim that the theorem holds with $N = \max\{\theta(l_1), \theta(l_2), \theta(l_1 l_2)\}$.

To see this let $0 \neq f \in \mathbb{R}[C]$ be psd. By Proposition 2.8 there exists a psd element $h \in \mathbb{R}[C]$ which is either identically zero or has only real zeros on C, such that $\theta(f - h) \leqslant \lceil \frac{1}{2}\delta(f) \rceil$ and $\delta(h) \leqslant \delta(f)$. We can assume $h \neq 0$. By Proposition 2.10 there is $g \in \{1, l_1, l_2, l_1 l_2\}$ such that h/g is a square in $\mathbb{R}(C)^*$, and by Lemma 2.12 we have $\theta(h) \leqslant \theta(g) + \frac{1}{2}(\delta(h) - \delta(g))$. So we conclude $\theta(h) \leqslant N + \frac{1}{2}(\delta(f) + 1)$. Hence the same bound holds for $\theta(f)$ since $\theta(f) \leqslant \max\{\theta(h), \theta(f - h)\}$. □

Remark 2.14. Theorem 2.13 is a sharpening of the stability assertion of Theorem 2.1, as far as the plane curves $y^2 + q(x) = 0$ are concerned that are considered in 2.13. We would like to point out that 2.13 also yields a similar sharpening for the other curves discussed in 2.1. Indeed, the reduction Lemma 2.4 and its proof are explicit enough to permit a transfer of the assertion of 2.13 to Zariski open subcurves. Although we won't make this more explicit, it justifies to restrict the remaining discussions to plane curves as in 2.13.

Remark 2.15. A closer inspection of the last proof exhibits that Theorem 2.13 is true with

$$N = \theta\big((x - \alpha)(\beta - x)\big) - 1,$$

where α is the smallest and β is the largest real root of $q(x)$. Clearly, this is the smallest possible N, as we see by taking $f = (x - \alpha)(\beta - x)$ in Theorem 2.13.

Indeed, let us abbreviate $l_1 = x - \alpha$ and $l_2 = \beta - x$. From $l_1 + l_2 = \beta - \alpha$ we get $(\beta - \alpha) l_1 = l_1^2 + l_1 l_2$, which implies $\theta(l_1) \leqslant \theta(l_1 l_2)$. Similarly $\theta(l_2) \leqslant \theta(l_1 l_2)$. Let us distinguish the argument according to the parity of $\delta(h)$. If $\delta(h)$ is even then $g = 1$ or $g = l_1 l_2$, and $g = 1$ gives the bound $\theta(h) \leqslant \frac{1}{2}\delta(h) \leqslant \frac{1}{2}\delta(f)$, while $g = l_1 l_2$ gives the bound $\theta(h) \leqslant \theta(l_1 l_2) + \frac{1}{2}\delta(h) - 1 \leqslant \theta(l_1 l_2) + \frac{1}{2}\delta(f) - 1$. If $\delta(h)$ is odd then $g = l_j$ for $j \in \{1, 2\}$, and this gives the bound $\theta(h) \leqslant \theta(l_j) + \frac{1}{2}(\delta(h) - 1)$, which is at most $\theta(l_1 l_2) + \lceil \frac{1}{2}\delta(f) \rceil - 1$.

Definition 2.16. Let C have equation $y^2 + q(x) = 0$ with q monic, separable and indefinite of degree four, and let $\alpha < \beta$ be the smallest resp. largest real root of q. We write $N_C := \theta((x - \alpha)(\beta - x))$, and we call N_C the *stability constant* of the curve C.

Remark 2.15 has shown:

Corollary 2.17. *For every psd $f \in \mathbb{R}[C]$ we have $\theta(f) \leqslant N_C - 1 + \lceil \frac{1}{2}\delta(f) \rceil$.*

Corollary 2.18. *In the terminology of* [4]*, the ideal $\mathfrak{I}_C = (y^2 + q(x))$ of C in $\mathbb{R}[x, y]$ is $(d, N_C + d - 1)$-sos for every $d \geqslant 1$. In particular, this ideal is theta-exact of theta-rank N_C.*

Proof. Let $p \in \mathbb{R}[x, y]$ have degree d, let $\overline{p} = p + \mathfrak{I}_C \in \mathbb{R}[C]$. We have $\delta(\overline{p}) \leqslant 2d$, so if p is psd on $C(\mathbb{R})$, Corollary 2.17 shows that $p \equiv \sum_j p_j(x, y)^2 \pmod{\mathfrak{I}_C}$ where $\delta(\overline{p}_j) \leqslant N_C + d - 1$ for every j. Thus every p_j is congruent modulo $\mathfrak{I}_C$ to a polynomial of degree $\leqslant N_C + d - 1$, which proves the corollary. □

In the next section we shall study in more detail how N_C depends on the curve C, i.e. on the polynomial $q(x)$. In particular, we will see that N_C can become arbitrarily large.

Remark 2.19. Assume $f \in \mathbb{R}[x, y]$ is a *linear* polynomial that is nonnegative on $C(\mathbb{R})$, where $C\colon y^2 + q(x) = 0$ is a curve as in Theorem 2.13. In this case we can make the argument leading to the proof of the theorem entirely explicit. We assume that f has a real zero $p = (\xi, \eta)$ in $C(\mathbb{R})$. So $f = 0$ is the tangent line to the plane curve C at the point p.

Let us first assume that $\eta \neq 0$ (the tangent is not vertical), and that $f = 0$ is not a double tangent. Then f has a pair of complex conjugate nonreal zeros on C, and we can apply the construction from Lemma 2.7 with $g = x - \xi$. The rational function

$$\varphi(x, y) = \frac{(x-\xi)^2}{f(x, y)}$$

has no poles on $C(\mathbb{R})$; let $\gamma > 0$ be its maximum value. Then $h := f - \frac{1}{\gamma}(x-\xi)^2$ is psd on $C(\mathbb{R})$ and has only real zeros on C. If $q \in C(\mathbb{R})$ is the point where φ attains its maximum γ, then the conic $h(x, y) = 0$ is tangent to C in the points p and q. The psd function h lies in the square class of $(x-\alpha)(\beta - x)$ in $\mathbb{R}(C)^*/\mathbb{R}(C)^{*2}$. More explicitly, we have

$$h = \text{const} \cdot \frac{F^2}{(x-\alpha)(\beta - x)} \tag{2.1}$$

with a positive constant and with

$$F = \left(\xi^2 y - \eta x^2\right) + (\alpha + \beta)(\eta x - \xi y) + \alpha\beta(y - \eta). \tag{2.2}$$

Indeed, the above F is nonzero since $\eta \neq 0$ and has $\delta(F) \leqslant 2$, and F vanishes in $(\alpha, 0)$, $(\beta, 0)$ and $p = (\xi, \eta)$. If we call $\tilde{q}$ the fourth zero of F, then the rational function on the right of (2.1) has zero divisor $2(p + \tilde{q})$, while h has zero divisor $2(p + q)$. This implies $q = \tilde{q}$ unless q and $\tilde{q}$ are $(\alpha, 0)$ and $(\beta, 0)$, which is excluded by the assumption $\eta \neq 0$. Note that $\tilde{q} = q$ is the point where φ attains its maximum.

If $f = 0$ is a double tangent then $\operatorname{div}(f) = 2(p + q - \infty - \overline{\infty})$ with a real point q on C (possibly $q = p$), and the argument of the first case remains formally true (with $\gamma = \infty$, i.e. with $h = f$). So in this case

$$f = \text{const} \cdot \frac{F^2}{(x-\alpha)(\beta - x)}$$

with a positive constant and with F as in (2.2).

In summary, once we have an explicit representation $(x-\alpha)(\beta - x) = \sum_\nu g_\nu^2$ as a sum of squares in $\mathbb{R}[C]$, we immediately get an *explicit* sum of squares representation for every psd tangent line f to C, namely

$$f = \frac{1}{\gamma}(x-\xi)^2 + \text{const} \cdot \sum_\nu \left(\frac{F g_\nu}{(x-\alpha)(\beta - x)}\right)^2$$

with F as in (2.2). (This is correct when $\eta \neq 0$ and $f = 0$ is not a double tangent; when $f = 0$ is a double tangent it is true with $\gamma = \infty$; when $\eta = 0$ it is true with $F = (x - \alpha)(\beta - x)$.) All fractions on the right lie in $\mathbb{R}[C]$.

Example 2.20. The previous remark allows us to write down parametric sums of squares representations of the positive tangents to the real curve, where the tangents and their representations are parametrized by the point of contact with the curve. Let us illustrate this using the curve C defined by $x^4 + y^2 = 1$. The set $C(\mathbb{R})$ of real points is just one convex oval, and with the notation of 2.19 we have $2(x - \alpha)(\beta - x) = 2(1 - x^2) = y^2 + (1 - x^2)^2$. If $(\xi, \eta) \in C(\mathbb{R})$ satisfies $\xi \neq 0$, and if $f = 0$ is the positive tangent to C in (ξ, η), the method of 2.19 gives the following representation as a sum of three squares:

$$f = \xi^2(x - \xi)^2 + \frac{((\xi^2 - 1)(x^2 + 1) + \eta y)^2 + (\eta(1 - x^2) + (\xi^2 - 1)y)^2}{4(1 - \xi^2)}.$$

Note that this representation passes to the limit cases $(\xi, \eta) = (\pm 1, 0)$, giving the representation $2(1 - \xi x) = (x - \xi)^2 + \frac{1}{2}y^2 + \frac{1}{2}(1 - x^2)^2$ for $\xi = \pm 1$.

Remarks 2.21.

1. It is not known whether Theorem 2.1 extends to curves of genus greater than one. For simplicity, let us restrict the discussion to irreducible affine and nonsingular curves C over $\mathbb{R}$ with $C(\mathbb{R}) \neq \varnothing$. When all points of C at infinity are real, then the sums of squares (sos) cone in $\mathbb{R}[C]$ is known to be stable [19]. However, as soon as the genus $g_C \geqslant 1$, this assumption implies that the sos cone in $\mathbb{R}[C]$ is (much) smaller than the psd cone [23]. On the other side, when C has nonreal points at infinity (for example, when $C(\mathbb{R})$ is compact), then the psd and the sos cone in $\mathbb{R}[C]$ coincide [24]. However, there is not a single such curve of genus $\geqslant 2$ for which it is known whether or not the sos cone is stable.
2. It is natural to weaken the question, and to ask only for partial stability, as in 1.2(2). For example, when C is a plane nonsingular curve of genus greater than one with $C(\mathbb{R})$ compact, can every linear polynomial nonnegative on $C(\mathbb{R})$ be written as a sum of squares in $\mathbb{R}[C]$, with the degrees of the summands bounded uniformly? Of course, this would be much weaker a property than full stability, and perhaps the answer is not so hard.

3. Application: Lifted LMI representations

Here we sketch how the main results of the previous section, combined possibly with further explicit results on degree bounds from the next, lead to very explicit lifted LMI representations of the convex hull of the curves considered.

First we record:

Corollary 3.1. *Let $C \subset \mathbb{A}^n$ be an irreducible real curve of genus one for which $C(\mathbb{R})$ is compact. Then the convex hull of $C(\mathbb{R})$ in $\mathbb{R}^n$ has a lifted LMI representation.*

Proof. Let $\tilde{C} \to C$ be the normalization of C. Since Theorem 2.1 applies to $\tilde{C}$, the Lasserre relaxation construction 1.4 becomes exact on every finite-dimensional linear subspace L of $\mathbb{R}[\tilde{C}]$.

Let $\mathbb{R}[x]_1 = \{f \in \mathbb{R}[x]\colon \deg(f) \leqslant 1\}$, and perform the relaxation construction to the image L of $\mathbb{R}[x]_1$ under $\mathbb{R}[x] \twoheadrightarrow \mathbb{R}[C] \hookrightarrow \mathbb{R}[\tilde{C}]$. □

We wish to demonstrate the explicitness of the construction by two examples. For this we restrict to discussing plane affine curves with equation $y^2 + q(x) = 0$ as in 2.13.

Remark 3.2. Let C be the curve $y^2 + q(x) = 0$, and let $L = \mathbb{R}[C]_1$ be the subspace of $\mathbb{R}[C]$ spanned by 1, x and y. Let $N := N_C$ be the stability constant of C (2.16). By Corollary 2.17, Lasserre's relaxation construction 1.4 works using the subspaces $W = \{f\colon \theta(f) \leqslant N\}$ and $U = \{f\colon \theta(f) \leqslant 2N\}$ of $\mathbb{R}[C]$. Since $\dim(W) = 2N$ and $\dim(U) = 4N$, this presents the convex hull of $C(\mathbb{R})$ in $\mathbb{R}^2$ in the form

$$\operatorname{conv} C(\mathbb{R}) = \left\{(x, y) \in \mathbb{R}^2\colon \exists z_1, \ldots, z_{2N-3} \text{ with } xA + yB + C_0 + \sum_{j=1}^{4N-3} z_j C_j \succcurlyeq 0\right\}$$

where A, B, C_j ($j = 0, \ldots, 4N-3$) are real symmetric matrices of size $2N \times 2N$ that are easy to make explicit. (Here $S \succcurlyeq 0$ means that the symmetric matrix S is positive semidefinite.)

Example 3.3. For an illustration, consider the simplest case, which is curves C with $N_C = 2$. Up to a linear coordinate change, these are precisely the curves with equation $y^2 + (x^2 - 1)(x^2 + b) = 0$ where $b \geqslant -1$, $b \neq 0$ (see 4.4 below). If this equation is written $y^2 + x^4 + Ax^2 + B = 0$, then $\operatorname{conv} C(\mathbb{R})$ is the set of $(x, y) \in \mathbb{R}^2$ for which there are $u_1, u_2, u_3, v_1, v_2 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & x & u_2 & y \\ x & u_2 & u_3 & v_1 \\ u_2 & u_3 & u_4 & v_2 \\ y & v_1 & v_2 & -B - Au_2 - u_4 \end{pmatrix} \succcurlyeq 0.$$

This matrix is obtained using the basis $1, x, x^2, y$ of W and the basis x^j, $x^k y$ ($0 \leqslant j \leqslant 4$, $0 \leqslant k \leqslant 2$) of U, resp. its dual basis of $U^\vee$.

Example 3.4. As pointed out in 1.6, we can expect interesting results as well from using construction 1.4 for subspaces L different from $\mathbb{R}[C]_1$. For example, we get concrete descriptions of the convex hulls of embeddings of C into higher-dimensional spaces, or of singular quotients of the curve, or of combinations of both. To present one more illustration, consider the curve $y^2 + x^4 = 1$, and perform construction 1.4 with the subspace L of $\mathbb{R}[C]$ spanned by 1, x and xy. This gives the "figure eight" curve $C' = \{w^2 = x^2(1 - x^4)\}$ and its convex hull in the (x, w)-plane. Since $N_C = 2$ (see 4.4 below), every psd element f of L satisfies $\theta(f) \leqslant 3$ by 2.17, so the construction works with $W = \{f\colon \delta(f) \leqslant 3\}$ and $U = \{f\colon \delta(f) \leqslant 6\}$. This yields a lifted LMI representation of $\operatorname{conv} C'(\mathbb{R})$ by symmetric 6×6 matrices with 9 free variables, namely as the set of $(x, w) \in \mathbb{R}^2$ for which there exist real numbers u_j ($2 \leqslant j \leqslant 6$) and v_j ($j \in \{0, 2, 3, 4\}$) which make the following matrix nonnegative:

$$\begin{pmatrix} 1 & x & u_2 & u_3 & v_0 & w \\ x & u_2 & u_3 & u_4 & w & v_2 \\ u_2 & u_3 & u_4 & u_5 & v_2 & v_3 \\ u_3 & u_4 & u_5 & u_6 & v_3 & v_4 \\ v_0 & w & v_2 & v_3 & 1-u_4 & x-u_5 \\ w & v_2 & v_3 & v_4 & x-u_5 & u_2-u_6 \end{pmatrix}.$$

Remark 3.5. As far as we are aware, this is the first example in the literature where explicit semidefinite representations are given for convex hulls of nonrational real algebraic varieties. An explicit representation of the convex hull of a genus one curve in 3-space as a spectrahedron was given recently by Rostalski and Sturmfels [21, Example 4.5]. Semidefinite representations of convex hulls of rational curves were given by Parrilo [16] (unpublished) and by Henrion [8], who also treats the quadratic Veronese surface. The arguments in these cases are elementary.

4. Degree bounds: A detailed study

Since explicit bounds for the stability constant (see 2.16) are necessary to produce concrete lifted LMI presentations, see Remark 3.2, we think it worthwile to discuss this constant and its dependence on the individual curve in greater detail.

4.1. We keep the assumptions of 2.5. So $q \in \mathbb{R}[x]$ is a monic quartic polynomial which is indefinite and separable, and C is the affine real curve with equation $y^2 + q(x) = 0$. Let $\alpha < \beta$ be the smallest resp. the largest real root of q, write $f = (x-\alpha)(x-\beta)$, and let $h \in \mathbb{R}[x]$ be the monic quadratic polynomial with $q = fh$. We have seen that the stability constant $N_C = \theta(-f)$ governs all degree bounds for sums of squares decompositions in $\mathbb{R}[C]$ (2.15).

Lemma 4.2. *Let d be the smallest number for which there is an identity $th - sf = 1$ with sums of squares s, t in $\mathbb{R}[x]$ and with $\deg(s) = \deg(t) \leqslant d$. Then $N_C = \frac{d}{2} + 2$.*

Proof. Since $-f = \sum_i (a_i + b_i y)^2$ with $a_i, b_i \in \mathbb{R}[x]$ implies $-f = \sum_i a_i^2 - q \sum_i b_i^2$, we only need to consider identities $-f = s' - tq$ with $a_i, b_i \in \mathbb{R}[x]$ and $s' = \sum_i a_i^2$, $t = \sum_i b_i^2$. Clearly,

$$\max_i \{\theta(a_i), \theta(b_i y)\} = \frac{1}{2}\deg(s') = 2 + \frac{1}{2}\deg(t). \tag{4.1}$$

Since f has only real zeros, f necessarily divides every a_i, and so f^2 divides s'. Dividing by f and putting $s = s'/f^2$ (a sum of squares in $\mathbb{R}[x]$) we get $-1 = sf - th$. The lemma follows. □

4.3. By a linear change of variables we can normalize the equation of C so that it becomes

$$y^2 + (x^2 - 1)h(x) = 0, \quad h(x) = x^2 + ax + b, \tag{4.2}$$

where h is separable and $h(x) > 0$ for $|x| \geqslant 1$. So the smallest (resp. the largest) real root of $q(x) = (x^2-1)h(x)$ is -1 (resp. $+1$). For our study of how the stability constant N_C depends on the curve C, it will be convenient to assume that C has this normalized form. The conditions on h mean that (a, b) lies in the set

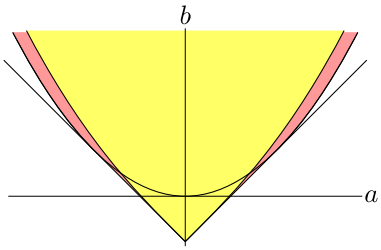
$$P := \{(a,b) \in \mathbb{R}^2 \colon a^2 - 4b < 0 \vee (a^2 - 4b > 0 \wedge |a| < \min\{2, |b|+1\})\}.$$

By $C_{a,b}$ we denote the affine curve with Eq. (4.2). We abbreviate its stability constant by $N(a,b) := N_{C_{a,b}}$, for $(a,b) \in P$.

4.4. By Lemma 4.2 we have $N(a,b) = 2 + \frac{1}{2}\deg(t)$, where $s, t \in \mathbb{R}[x]$ are psd polynomials with

$$1 = t(x)\big(x^2 + ax + b\big) - s(x)\big(x^2 - 1\big)$$

and $\deg(s) = \deg(t)$ is as small as possible. It is therefore clear that always $N(a,b) \geqslant 2$ holds, and that $N(a,b) = 2$ if and only if $a = 0$. Without proof we remark that $N(a,b) \leqslant 3$ if and only if $\frac{a^4}{16} + a^2 \leqslant (b+1)^2$. The following picture shows the parameter set P, the yellow part corresponding to $N \leqslant 3$ and the red part to $N \geqslant 4$ (for interpretation of the references to color in this picture, the reader is referred to the web version of this article):



For the next lemma let P' denote the boundary of the closure of P, so

$$P' = \big\{(a,b) \in \mathbb{R}^2 \colon\ a^2 = 4b \geqslant 4 \vee |a| = b + 1 \leqslant 2\big\}.$$

Lemma 4.5. *Let $(a_\nu, b_\nu)_{\nu \geqslant 1}$ be a sequence in P that converges to $(a,b) \in \mathbb{R}^2$ for $\nu \to \infty$. If the sequence $N(a_\nu, b_\nu)$ is bounded, and if $(a,b) \neq (0,-1)$, then $(a,b) \notin P'$. If in addition $(a,b) \in P$ then $N(a,b) \leqslant \sup_\nu N(a_\nu, b_\nu)$.*

(If $(a,b) \notin P \cup P'$ then $a^2 - 4b = 0$ and $|a| < 2$.)

Proof of Lemma 4.5. Assume that the sequence $N(a_\nu, b_\nu)$ is bounded. By 4.4 this means that there are $d \geqslant 0$ and sums of squares $s_\nu(x)$, $t_\nu(x)$ in $\mathbb{R}[x]$ with $\deg(s_\nu) = \deg(t_\nu) \leqslant 2d$ and

$$1 = \big(x^2 + a_\nu x + b_\nu\big)t_\nu(x) - \big(x^2 - 1\big)s_\nu(x) \tag{4.3}$$

for every ν. We first assume that the coefficients of the t_ν and s_ν are uniformly bounded for all ν. After passing to a suitable subsequence we can then assume that we have (coefficient-wise) convergences $s_\nu \to s$ and $t_\nu \to t$, where $s, t \in \mathbb{R}[x]$ are clearly sums of squares. Passing (4.3) to the limit $\nu \to \infty$ we see

$$1 = \big(x^2 + ax + b\big)t(x) - \big(x^2 - 1\big)s(x). \tag{4.4}$$

If $(a,b) \in P$, it follows that $N(a,b) \leqslant d$. Assume $(a,b) \in P'$ and $(a,b) \neq (0,-1)$. If $|a| > 2$ then $a^2 = 4b$, so $x^2 + ax + b = (x + \frac{a}{2})^2$ has a double zero at $-\frac{a}{2}$, which contradicts (4.4).

There remains the case where the coefficients or s_ν or t_ν are unbounded for $\nu \to \infty$. We scale (4.3) for each ν by the factor $\frac{1}{c_\nu}$ where $c_\nu > 0$ is the maximum absolute value of the coefficients of $s_\nu(x)$ and $t_\nu(x)$. After passing to a subsequence we have convergence $c_\nu^{-1} s_\nu(x) \to s(x)$ and $c_\nu^{-1} t_\nu(x) \to t(x)$, and again $s, t \in \mathbb{R}[x]$ are sums of squares. Both are nonzero since each has a coefficient ± 1. Taking (4.3) to the limit gives $(x^2 + ax + b)t(x) = (x^2 - 1)s(x)$. This implies that $(x^2 - 1)(x^2 + ax + b)$ is a psd polynomial, which only happens for $(a, b) = (0, -1)$. □

Corollary 4.6.

(a) *For each $N \geqslant 0$, the set $\{(a, b) \in P\colon\ N(a, b) \leqslant N\}$ is relatively closed in P.*
(b) *When (a, b) moves in P towards a boundary point $\neq (0, -1)$ in P', then $N(a, b)$ tends to infinity.*

Note that (b) is not necessarily true when $(a, b) \to (0, -1)$, for example since $N(0, b) = 2$ for all $b > -1$.

Remark 4.7. Degeneration of $(a, b) \in P$ towards a boundary point $(a_0, b_0) \in P'$, $(a_0, b_0) \neq (0, -1)$, corresponds to degenerating the curve $C_{a,b}$ into a nodal curve (for $|a_0| \neq 1$) or a cuspidal curve (for $a_0 = \pm 1$), rational in either case.

4.8. We do not know how to express $N(a, b)$ for arbitrary $(a, b) \in P$. We conclude with proving an explicit lower bound for $N(a, b)$ in the $|a| > 2$ part. We keep the normalizations 4.3 and write $h = x^2 + ax + b$ and $f = x^2 - 1$.

Assume that one of $h'(-1) > 0$ or $h'(1) < 0$ holds. Either condition implies $h(x) > 0$ for all $x \in \mathbb{R}$. Let us assume $h'(1) < 0$, and let $s, t \in \mathbb{R}[x]$ be psd polynomials with $1 = th - sf$ (cf. 4.2). We conclude

$$t(x) \geqslant \frac{1}{h(x)} \quad \text{for } |x| \geqslant 1, \qquad 0 \leqslant t(x) \leqslant \frac{1}{h(x)} \quad \text{for } |x| \leqslant 1. \tag{4.5}$$

In particular $t(1) = \frac{1}{h(1)}$. Since h is quadratic, $h'(1) < 0$ implies $h'(x) < 0$ for all $x \leqslant 1$, and so $\frac{1}{h}$ is strictly increasing for $x \leqslant 1$. Hence $0 \leqslant t(x) \leqslant t(1) = \frac{1}{h(1)}$ for $|x| \leqslant 1$. On the other hand, (4.5) implies $t'(1) \geqslant (\frac{1}{h})'(1) = -\frac{h'(1)}{h(1)^2}$.

According to Markov's inequality [11,1], any polynomial $p \in \mathbb{R}[x]$ of degree $\leqslant n$ satisfies $\|p'\|_{[-1,1]} \leqslant n^2 \cdot \|p\|_{[-1,1]}$, where $\|p\|_{[-1,1]} = \max\{|p(x)|\colon\ |x| \leqslant 1\}$. Applying this to $p = t - \frac{1}{2h(1)}$ we conclude

$$\deg(t)^2 \geqslant 2 \cdot \frac{(1/h)'(1)}{(1/h)(1)} = -2\frac{h'(1)}{h(1)}. \tag{4.6}$$

Writing $h = x^2 + ax + b$, the assumption $h'(1) < 0$ means $a + 2 < 0$, and (4.6) becomes

$$\deg(t)^2 \geqslant -\frac{2(a + 2)}{1 + a + b}.$$

If instead of $h'(1) < 0$ we assume $h'(-1) > 0$, we get a symmetric estimate. Altogether we have shown:

Proposition 4.9. *Consider the affine curve* $y^2 + (x^2 - 1)(x^2 + ax + b) = 0$ *with* $(a, b) \in P$. *If* $|a| > 2$ *then*

$$N(a, b) \geqslant 2 + \sqrt{\frac{|a| - 2}{2(1 + b - |a|)}}.$$

Example 4.10. For $\gamma > 0$ consider the curve C_γ with equation $y^2 + (x^2 - 1)h_\gamma(x) = 0$ where

$$h_\gamma(x) = x^2 + \left(2 + \frac{2}{\gamma}\right)x + \left(1 + \frac{2}{\gamma} + \frac{4}{\gamma^2}\right) = \left(x + 1 + \frac{1}{\gamma}\right)^2 + \frac{3}{\gamma^2}.$$

Via Markov's inequality we get the lower bound $N_{C_\gamma} \geqslant 2 + \frac{\sqrt{\gamma}}{2}$ from Proposition 4.9, which tends to infinity for $\gamma \to \infty$. However, this bound does not seem close to being sharp. A small series of numerical experiments using Parrilo's `sostools` package [17] resulted in the following observations:

N	$4(N-2)^2$	$\gamma_{\max}(N)$
3	4	2.57
4	16	6.92
5	36	12.95
6	64	20.70
7	100	30.17
8	144	41.35
9	196	54.25

For given N let $\gamma_{\max}(N)$ be the maximal $\gamma > 0$ for which $N_{C_\gamma} \leqslant N$. The Markov estimate gives $\gamma_{\max}(N) \leqslant 4(N-2)^2$, which is the second column. The approximate true value of $\gamma_{\max}(N)$ is shown in the last column.

References

[1] P. Borwein, T. Erdélyi, Polynomials and Polynomial Inequalities, Grad. Texts in Math., vol. 161, Springer, New York, 1995.
[2] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge Univ. Press, Cambridge, 2004.
[3] J. Gouveia, T. Netzer, Positive polynomials and projections of spectrahedra, SIAM J. Optim., in press.
[4] J. Gouveia, P. Parrilo, R. Thomas, Theta bodies for polynomial ideals, SIAM J. Optim. 20 (2010) 2097–2118.
[5] W. Helton, J. Nie, Sufficient and necessary conditions for semidefinite representability of convex hulls and sets, SIAM J. Optim. 20 (2009) 759–791.
[6] W. Helton, J. Nie, Semidefinite representation of convex sets, Math. Program. 122 (Ser. A) (2010) 21–64.
[7] W. Helton, V. Vinnikov, Linear matrix inequality representation of sets, Comm. Pure Appl. Math. 60 (2007) 654–674.
[8] D. Henrion, Semidefinite representation of convex hulls of rational varieties, preprint, 2009, http://arxiv.org/abs/0901.1821.
[9] J.-B. Lasserre, Convex sets with semidefinite representation, Math. Program. 120 (Ser. A) (2009) 457–477.
[10] L. Lovász, Semidefinite programs and combinatorial optimization, in: Recent Advances in Algorithms and Combinatorics, in: CMS Books Math., vol. 11, Springer, New York, 2003, pp. 137–194.
[11] A.A. Markov, On a problem of D.I. Mendeleev, Acad. Sci. St. Petersburg 62 (1889) 1–24.
[12] A. Nemirovsky, Advances in convex optimization: Conic programming, in: Int. Cong. Math., vol. I, Eur. Math. Soc., Zürich, 2007, pp. 413–444.

[13] Yu. Nesterov, A. Nemirovsky, Interior-point Polynomial Algorithms in Convex Programming, SIAM Stud. Appl. Math., vol. 13, SIAM, Philadelphia, PA, 1994.
[14] T. Netzer, On semidefinite representations of non-closed sets, Linear Algebra Appl. 432 (2010) 3072–3078.
[15] T. Netzer, D. Plaumann, M. Schweighofer, Exposed faces of semidefinite representable sets, SIAM J. Optim. 20 (2010) 1944–1955.
[16] P. Parrilo, Exact semidefinite representations for genus zero curves, Talk at workshop "Positive Polynomials and Optimization", BIRS, Banff, 2006.
[17] P. Parrilo, sostools. Free software package, http://www.cds.caltech.edu/sostools.
[18] D. Plaumann, Stabilität von Quadratsummen auf reellen algebraischen Varietäten, Diplomarbeit, Univ. Duisburg, 2004.
[19] V. Powers, C. Scheiderer, The moment problem for non-compact semialgebraic sets, Adv. Geom. 1 (2001) 71–88.
[20] K. Ranestad, B. Sturmfels, On the convex hull of a space curve, preprint, 2009, http://arxiv.org/abs/0912.2986.
[21] Ph. Rostalski, B. Sturmfels, Dualities in convex algebraic geometry, Rend. Mat. (VII) 30 (2010) 249–291.
[22] R. Sanyal, F. Sottile, B. Sturmfels, Orbitopes, Mathematika, in press.
[23] C. Scheiderer, Sums of squares of regular functions on real algebraic varieties, Trans. Amer. Math. Soc. 352 (1999) 1039–1069.
[24] C. Scheiderer, Sums of squares on real algebraic curves, Math. Z. 245 (2003) 725–760.
[25] C. Scheiderer, Non-existence of degree bounds for weighted sums of squares representations, J. Complexity 21 (2005) 823–844.
[26] C. Scheiderer, Positivity and sums of squares: A guide to recent results, in: Emerging Applications of Algebraic Geometry, in: IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 271–324.