Existence of Positive Definite Noncoercive Sums of Squares

Gregory C. Verchota

Syracuse University

EXISTENCE OF POSITIVE DEFINITE NONCOERCIVE SUMS OF SQUARES IN $\mathbb{R}[x_1,\ldots,x_n]$

GREGORY C. VERCHOTA

ABSTRACT. Positive definite forms $f \in \mathbb{R}[x_1,\ldots,x_n]$ which are sums of squares of forms of $\mathbb{R}[x_1,\ldots,x_n]$ are constructed to have the additional property that the members of any collection of forms whose squares sum to f must share a nontrivial complex root in $\mathbb{C}^n$.

1. INTRODUCTION

Let $f \in \mathbb{R}[x_1,\ldots,x_n]$ be a *form*, i.e. homogeneous polynomial. Suppose f is a *sum of squares* (sos) of forms in $\mathbb{R}[x_1,\ldots,x_n]$ and is *positive definite* (pd), $f(\mathbf{a}) > 0$ for all $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Writing $f = \sum p_j^2$ this is equivalent to saying that the forms p_j share no common nontrivial real root from $\mathbb{R}^n$.

(1.1)
Suppose a positive definite *form f has at least one* sos *representation. Does f necessarily have a representation $f = \sum q_k^2$ with $q_k \in \mathbb{R}[x_1,\ldots,x_n]$ and the q_k sharing no common complex root from $\mathbb{C}^n \setminus \{\mathbf{0}\}$?*

For example,

(i) the *positive semi-definite* (psd) $x_1^2 = p^2 \in \mathbb{R}[x_1,x_2,x_3]$ is uniquely represented as an sos, and $p(0,1,i) = 0$;

(ii) $x_1^2 + x_2^2 \in \mathbb{R}[x_1,x_2]$ is pd with x_1 and x_2 sharing no common nontrivial complex root;

(iii) $f = (x_1^2 + x_2^2)^2 = p^2$ is pd with the quadratic form p having the root $(1,i) \in \mathbb{C}^2$. But also $f = (x_1^2)^2 + (\sqrt{2}x_1x_2)^2 + (x_2^2)^2$ or $(x_1^2 - x_2^2)^2 + (2x_1x_2)^2$ and in each case the quadratic forms now share no common nontrivial complex root.

Though not the subject of this article, the study of boundary value problems for elliptic partial differential equations (PDE) motivates question (1.1). Denote by $\partial = (\partial_1,\ldots,\partial_n) = (\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n})$ the vector of first partial derivatives for $\mathbb{R}^n$. Let $\alpha \in \mathbb{N}_0^n$ denote a multi-index. Define $|\alpha| = \alpha_1 + \cdots + \alpha_n$ and $\partial^\alpha = \partial_1^{\alpha_1}\cdots\partial_n^{\alpha_n}$.

A theorem of N. Aronszajn and K. T. Smith [Agm65] may be stated as

---

Date: February 2, 2008.

1991 *Mathematics Subject Classification.* 12D15,11E25,35J30,35J40.

The author gratefully acknowledges partial support provided by the National Science Foundation through award DMS-0401159.

Let $p_1, \ldots, p_r \in \mathbb{R}[x_1, \ldots, x_n]$ be forms of degree d. Let $\Omega \subset \mathbb{R}^n$ be a bounded open connected set with suitably regular boundary and let $\overline{\Omega}$ be its closure. Then the integro-differential quadratic form

$$\sum_j \int_\Omega |p_j(\partial)u|^2 dx \tag{1.2}$$

is coercive *over all functions u which have continuous partial derivatives of order d in Ω that extend continuously to $\overline{\Omega}$* if and only if *the system*

$$p_1 = p_2 = \cdots = p_r = 0$$

has no solution $\mathbf{a} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$.

For (1.2) to be *coercive* over the collection of functions u it is required, by definition, that there be constants $C > 0$ and $c_0 \in \mathbb{R}$ independent of the functions u so that

$$\sum_j \int_\Omega |p_j(\partial)u|^2 dx \geq C \int_\Omega \sum_{|\alpha| \leq d} |\partial^\alpha u|^2 dx - c_0 \int_\Omega |u|^2 dx \tag{1.3}$$

for all u in the collection. Once this estimate is obtained various elliptic boundary value problems can be solved.

The Aronszajn-Smith theorem gives a precise algebraic characterization of all integro-differential forms (1.2) for which the coercive estimate (1.3) can hold. The integro-differential forms (1.2) are termed *formally positive* because of their *sos* shape. S. Agmon [Agm58] improved this result by proving a necessary and sufficient (and more complicated) algebraic condition on all integro-differential forms

$$Re \sum_{|\alpha| \leq d} \sum_{|\beta| \leq d} \int_\Omega a_{\alpha\beta} \partial^\alpha u \overline{\partial^\beta u} dx \tag{1.4}$$

not only the formally positive, that give rise to self-adjoint linear properly elliptic differential operators

$$L(\partial) = \sum_{|\alpha| \leq d} \sum_{|\beta| \leq d} a_{\alpha\beta} \partial^{\alpha+\beta} \tag{1.5}$$

and their regular boundary value problems [Agm58][Agm60]. When $a_{\alpha\beta} \in \mathbb{R}$ and the integro-differential form is formally positive, L corresponds to a polynomial f of degree $2d$ that is a sum of squares.

With his algebraic characterization Agmon solved completely the *coerciveness problem for integro-differential forms* in the theory of linear PDE. However, the *coerciveness problem for linear differential operators* $L(\partial) = \sum_{|\alpha| \leq 2d} a_\alpha \partial^\alpha$ has not been solved. This problem can be stated in a way that leads back to the question about sums of squares in $\mathbb{R}[x_1, \ldots, x_n]$.

Instead of the integro-differential form one begins with the homogeneous constant coefficient operator in $\mathbb{R}^n$

$$L(\partial) = \sum_{|\alpha| = 2d} a_\alpha \partial^\alpha$$

$a_\alpha \in \mathbb{R}$. These will be self-adjoint. Suppose L is *elliptic* (equivalent to properly elliptic in this setting) $L(\xi) > 0$ for all $\xi \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. In general L can be rewritten an infinity of ways in the shape (1.5)

$$L(\partial) = \sum_{|\alpha| = |\beta| = d} a_{\alpha\beta} \partial^{\alpha+\beta} \tag{1.6}$$

and therefore admits an infinity of integro-differential forms (1.4). Is there any choice of rewriting (1.6) that yields a coercive estimate?

This fundamental question is broader than what can be answered here. Instead the question will be specialized to the setting of the Aronszajn-Smith theorem.

Suppose it is further known that the homogeneous differential operator is an *sos*, $L(\partial) = \sum p_j^2(\partial)$. Then the theorem provides the necessary and sufficient algebraic condition for the integro-differential form (1.2) to be coercive (1.3). If the form were to fail the algebraic condition and thus fail to be coercive is there another way to write the differential operator L *as a sum of squares* and thereby use the theorem again to obtain the coercive estimate for a new integro-differential form associated to L and thus solve boundary value problems for L? This is question (1.1).

All the results and proofs of this article are independent of these PDE considerations. Some more will be said about PDE in the last section.

Definition 1.1. $f \in \mathbb{R}[x_1,\dots,x_n]$ is called a *sum of squares* (an *sos*) if there exist polynomials $p_1,\dots,p_r \in \mathbb{R}[x_1,\dots,x_n]$ so that f has the representation $f = \sum_{j=1}^r p_j^2$

Definition 1.2. An sos $f \in \mathbb{R}[x_1,\dots,x_n]$ is called *coercive* or a *coercive sum of squares* if there exists a representation

$$f = \sum_{j=1}^{r} p_j^2 \tag{1.7}$$

with $p_1,\dots,p_r \in \mathbb{R}[x_1,\dots,x_n]$ such that there are *no* solutions $\mathbf{a} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ to the system

$$p_1 = \cdots = p_r = 0 \tag{1.8}$$

When such an f is homogeneous it is also called a *coercive form*.

To be clear

Definition 1.3. An sos $f \in \mathbb{R}[x_1,\dots,x_n]$ is called *noncoercive* or a *noncoercive sos* if there exists a representation (1.7) for f and if *every* such representation has a nontrivial solution in $\mathbb{C}^n$ to the *corresponding system* (1.8).

Question (1.1) asks if every positive definite sos is coercive. The aim of this article is to establish, by construction, the existence of positive definite noncoercive sums of squares. That this can be done is related to the well known fact that *not every positive definite polynomial is a sum of squares*.

If every pd polynomial were an sos the answer to question (1.1) would be *yes*. This follows because positive definiteness of f allows

$$f = [f - \epsilon(x_1^{2d} + \cdots + x_n^{2d})] + \epsilon(x_1^{2d} + \cdots) \tag{1.9}$$

with the bracketed term pd for $\epsilon > 0$ small enough. When the bracketed term is an sos, (1.9) is an sos representation for f that satisfies the definition of coercive sos.

We adopt standard notations for psd homogeneous polynomials [CL78][BCR98] p.111. $P_{n,d}$ denotes the set of $f \in \mathbb{R}[x_1,\dots,x_n]$ homogeneous of degree d that are *nonnegative* on $\mathbb{R}^n$. $\Sigma_{n,d}$ denotes the set of all $f \in P_{n,d}$ that are sos. These sets are nonempty only when d is an even number.

FOR THE REMAINDER OF THIS ARTICLE ALL POLYNOMIALS WILL BE HOMOGENEOUS POLYNOMIALS, OR FORMS.

(Homogenization can be used for other statements.)

The argument given above together with Hilbert's results on positive polynomials that are sos [Hil88], [Rez07] immediately yields the Theorem

(1.10)
If $n \leq 2$ and d is an even natural number, or if $d = 2$ and n is a natural number, or if $(n, d) = (3, 4)$, then every pd form of $P_{n,d}$ is a coercive sum of squares.

The result of Hilbert [Raj93], [Swa00], [Rud00], [Pfi04], [PR00] used here is that $P_{3,4} = \Sigma_{3,4}$, while $P_{2,2p} = \Sigma_{2,2p}$ and $P_{n,2} = \Sigma_{n,2}$ are elementary. See [BCR98] pp.111-112.

Hilbert further proved that in every other case $\Sigma_{n,2p}$ is a proper subset of $P_{n,2p}$, eliminating the argument based on (1.9). It was T. S. Motzkin [Mot67] who first published explicit examples of positive semi-definite polynomials that were not sos. There are now various examples of these, e.g. [Rob73],[CL78],[CL77],[LL78]; see [Rez00] for more. We found two of these to be very useful for the purpose here. Both are of Motzkin type and due to M. D. Choi and T. Y. Lam.

$$q(w, x, y, z) = w^4 + x^2y^2 + y^2z^2 + z^2x^2 - 4wxyz \tag{1.11}$$

and

$$s(x, y, z) = x^4y^2 + y^4z^2 + z^4x^2 - 3x^2y^2z^2$$

Both are nonnegative (psd) by the arithmetic-geometric mean inequality and neither is an sos. Thus $q \in P_{4,4} \setminus \Sigma_{4,4}$ and $s \in P_{3,6} \setminus \Sigma_{3,6}$.

For $\eta \geq 0$ define

$$q_\eta = q + \eta(x^4 + y^4 + z^4) \tag{1.12}$$

$$s_\eta = s + \eta(x^6 + y^6 + z^6)$$

For $\eta > 0$, q_η and s_η are pd. As long as η is small enough each is not an sos. This follows by an elementary topological argument first given by R. M. Robinson [Rob73] pp.267-268 which, moreover, shows the sets Σ to be topologically closed sets. It is also true that for all η large enough q_η and s_η are sos. See, for example, p.269 of [Rob73] (in the case of q_η it can be verified that the w^4 term in q obviates the need to add ηw^4). Consequently for each polynomial there is a smallest value of η, $\eta_0 > 0$, that makes q_η or s_η sos (cf. also the proof of Corollary 5.6 [CLR95] p.122). In Section 3 it is shown for the quartic q that the square root of this value is the smallest positive root of $X^3 - \frac{1}{2}X + \frac{1}{9} = 0$, and that

$$q_{\eta_0}(w, x, y, z) = (w^2 - \sqrt{\eta_0}(x^2 + y^2 + z^2))^2 + \frac{2}{9\sqrt{\eta_0}}[(3\sqrt{\eta_0}wx - yz)^2 + (3\sqrt{\eta_0}wy - zx)^2 + (3\sqrt{\eta_0}wz - xy)^2] \tag{1.13}$$

In addition, it is proved that there is exactly one Gram matrix (or Gramian [Gel89]) that represents the polynomial q_{η_0}. This means that every other sos representation for q_{η_0} is merely a sum of squares of quadratics that are linear combinations of the quadratics of (1.13). Thus any common complex roots must be the same among all representations.

The Gram matrix method of Choi, Lam and B. Reznick [CLR95], used for studying sos representations of polynomials, is put into a tensor setting in Section 2. Every form of degree $2p$ is nonuniquely represented by a symmetric matrix (rank-2 symmetric tensor) acting as a quadratic form on the vector space of rank-p symmetric tensors. These are termed *representation matrices* for the form. The Gram matrices are those representation matrices that are psd, necessary and sufficient for an sos representation.

The polynomial (1.13) provides an example of a positive definite quartic with a unique Gram matrix. A positive definite sextic with a unique Gram matrix has previously been identified by Reznick in [Pra06]. It is like the ones that will be constructed in Section 5 from the s_η.

However wonderful it is, q_{η_0} is coercive. It is proved in Section 4 that

$$(u^2 + v^2 + vw)^2 + q_{\eta_0}(w, x, y, z) \tag{1.14}$$

is *positive definite* and *noncoercive* in $\Sigma_{6,4}$. In effect the uniqueness of representation of (1.13) and the presence of the monomial vw forces a uniqueness of representation upon (1.14), while $(1, i, 0, 0, 0, 0)$ is a solution to the corresponding system of quadratic equations (1.8). It follows from the definition of coercive sos that any form $f \in \mathbb{R}[x_1, \ldots, x_n]$ of even degree d such that $f + x_{n+1}^d$ is a coercive sos must itself be a coercive sos. Consequently monomials $x_7^4, x_8^4, \ldots$ can be added to (1.14) preserving all required properties and the following theorem and partial answer to question (1.1) is obtained.

Theorem 1.4. *For $n \geq 6$, $\Sigma_{n,4}$ contains polynomials that are positive definite and noncoercive.*

Theorem 1.4 is really a statement about certain cones of polynomials. After a scaling (1.13) can be rewritten

$$a_1(x_1^2 - \gamma(x_2^2 + x_3^2 + x_4^2))^2 + a_2(x_1x_2 - x_3x_4)^2 + a_3(x_1x_3 - x_4x_2)^2 + a_4(x_1x_4 - x_2x_3)^2 \tag{1.15}$$

where it happens that for all values of γ, $0 < \gamma < \frac{1}{3}$ and all positive $a_1, \ldots, a_4$, the forms (1.15) are pd with a unique Gram matrices.

Corollary 1.5. *For $n \geq 6$ there exist nonempty collections of quadratic forms $\{p_1, \ldots, p_r\} \subset \mathbb{R}[x_1, \ldots, x_n]$ so that there exist no nontrivial solutions from $\mathbb{R}^n$ to the systems $p_1 = p_2 = \cdots = p_r = 0$, and so that every $f = \sum a_j p_j^2$, with positive coefficients $a_1, \ldots, a_r$, is a noncoercive sos.*

The Choi-Lam sextic form s (1.11) possesses more structure than its quartic counterpart q. First it is an *even* form. A form f is *even* if it is also a polynomial in $x_1^2, x_2^2, \ldots, x_n^2$. Second it is *symmetric*. A form f is *symmetric* if for every permutation σ on n objects $f(\mathbf{x}) = f(\sigma(\mathbf{x}))$. The construction (1.12) of the forms s_η preserves both of these properties. In Section 5, for $s_\eta(x, y, z)$ with a unique Gram matrix, it is proved that when x^2 is replaced with $w^2 + x^2$ the resulting form is pd and noncoercive.

Theorem 1.6. *For $n \geq 4$, $\Sigma_{n,6}$ contains polynomials that are positive definite and noncoercive.*

The additional structure provided by the non-sos s seems to be the reason Theorem 1.6 comes closer than Theorem 1.4 to being a complete result. As remarked on p.263 of [Rez00] and in [Har99], in any dimension every psd *even symmetric quartic* form is an sos. Further, the replacement of x^2 with $w^2 + x^2$ that works in the sextic construction seems to rely more on the *even* property than it does on symmetry. It turns out that every psd *even quartic* form in $n = 4$ or fewer variables is a sum of squares. This follows from results of P. H. Diananda [Dia62]. Thus constructing a quartic noncoercive sos for $n = 5$ from an even form in 4 variables in a way analogous to the sextic case is not possible. On the other hand the Horn form [HN63] pp. 334-335 [Dia62] p.25 [Rez00] p.260 provides a psd *even quartic* form for $n = 5$ that is not an sos. See [CL78] pp.394-396.

Between the coercive Theorem (1.10) and the noncoercive Theorems 1.4 and 1.6, dimensions 4 and 5 for the former and 3 for the latter remain obscure. This puzzle will be discussed further in Section 6.

2. A MULTILINEAR SETUP

At first let $\mathbf{e}^1, \dots, \mathbf{e}^n$ and $\mathbf{e}_1, \dots, \mathbf{e}_n$ be the standard (contravariant and covariant) basis vectors for $\mathbb{R}^n$. The scalar product of vector and covector is denoted $\mathbf{x} \cdot \mathbf{u} = \sum x_j u^j$ where $x_1, \dots, u^1, \dots$ are the standard coordinates of $\mathbf{x}$ and $\mathbf{u}$. The nonnegative integers are denoted $\mathbb{N}_0$. For a multi-index $\alpha \in \mathbb{N}_0$ its *order* is $|\alpha| = \alpha + \cdots + \alpha_n$, and $\alpha ! = \alpha_1 ! \cdots \alpha_n !$ For $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

The (contravariant) *tensors* $\mathbf{t}$ of *rank* p are multilinear (p-linear) forms mapping p vectors of $\mathbb{R}^n$ to $\mathbb{R}$ by

$$\mathbf{t} \cdot \mathbf{x}^1 \mathbf{x}^2 \cdots \mathbf{x}^p = \sum x^1_{j_1} x^2_{j_2} \cdots x^p_{j_p} t^{j_1 j_2 \cdots j_p}$$

The *coordinates* of $\mathbf{t}$ are $t^{j_1 \cdots j_p}$ and are obtained by $\mathbf{t} \cdot \mathbf{e}^{j_1} \cdots \mathbf{e}^{j_p} = t^{j_1 \cdots j_p}$. See [vdW70] pp.74-75, 80-81.

Given p (co)vectors $\mathbf{u}_1, \dots, \mathbf{u}_p$ a tensor $\mathbf{t}$ of rank p may be defined by the *tensor product*

$$\mathbf{t} = \mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_p$$

which acts multilinearly as

$$\mathbf{t} \cdot \mathbf{x}^1 \mathbf{x}^2 \cdots \mathbf{x}^p = (\mathbf{x}^1 \cdot \mathbf{u}_1)(\mathbf{x}^2 \cdot \mathbf{u}_2) \cdots (\mathbf{x}^p \cdot \mathbf{u}_p) \tag{2.1}$$

so that $t^{j_1 \cdots j_p} = u_1^{j_1} \cdots u_p^{j_p}$.

The collection of tensors of rank p, $T^p(\mathbb{R}^n)$, forms a vector space over $\mathbb{R}$ of dimension n^p with standard basis

$$\{\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p} : 1 \le j_\nu \le n\}$$

Let $\mathfrak{S}_p$ denote the symmetric group of all permutations of p objects. For each $\sigma \in \mathfrak{S}_p$ the map

$$P_\sigma(\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}) = \mathbf{e}_{j_{\sigma(1)}} \otimes \cdots \otimes \mathbf{e}_{j_{\sigma(p)}}$$

defines a permutation of the basis vectors of $T^p(\mathbb{R}^n)$ and thereby induces a (unique) linear isomorphism on $T^p(\mathbb{R}^n)$ [Yok92] p.43. If $P_\sigma(\mathbf{t}) = \mathbf{t}$ for all $\sigma \in \mathfrak{S}_p$, then $\mathbf{t}$ is called a *symmetric tensor*. The set of all symmetric tensors of rank p, $S^p(\mathbb{R}^n)$, also forms a vector space over $\mathbb{R}$. The linear operator

$$Sym = Sym_p = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} P_\sigma$$

is a projection from $T^p(\mathbb{R}^n)$ *onto* $S^p(\mathbb{R}^n)$ so that

$$\{Sym(\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}) : 1 \le j_1 \le \cdots \le j_p \le n\} \tag{2.2}$$

forms a basis for $S^p(\mathbb{R}^n)$. Further,

$$dim(S^p(\mathbb{R}^n)) = \begin{pmatrix} n+p-1 \\ p \end{pmatrix} \tag{2.3}$$

[Yok92] pp.47-48.

Given indices $j_1 \le \cdots \le j_p$ as in (2.2) let α_k equal the number of indices equal to k for $1 \le k \le n$. In this way the multi-indices $\alpha \in \mathbb{N}_0^n$ of order p are put in one-to-one correspondence with the basis elements of $S^p(\mathbb{R}^n)$. Denote

$$\mathbf{E}_\alpha = Sym(\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}) \tag{2.4}$$

for each basis element in (2.2) where α corresponds to $j_1 \leq \cdots \leq j_p$.

The coordinates of $\mathbf{E}_\alpha$ *(as a tensor in* $T^p(\mathbb{R}^n)$*)* $E_\alpha^{k_1 \cdots k_p}$ *are either* 0 *or* $\frac{\alpha!}{p!}$*, and sum to* 1.

Example 2.1. (i) For $p = 2$, $\mathbf{E}_{(2,0\ldots,0)} = \mathbf{e}_1 \otimes \mathbf{e}_1$ with $E^{11}_{(2,0\ldots,0)} = 1$ the only nonzero coordinate.
$\mathbf{E}_{(1,1,0\ldots,0)} = \frac{1}{2}(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1)$ with $E^{12}_{(1,1,0\ldots,0)} = E^{21}_{(1,1,0\ldots,0)} = \frac{1}{2}$ the only nonzero coordinates.

Thus $\{\mathbf{E}_\alpha : |\alpha| = 2\}$ is identified with an orthogonal basis for the $n \times n$ symmetric matrices under the Hilbert-Schmidt inner product.
(ii) For $p = 3$, $\mathbf{E}_{(3,0\ldots,0)} = \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1$.
$\mathbf{E}_{(2,1,0\ldots,0)} = \frac{1}{3}(\mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1)$.
$\mathbf{E}_{(1,1,1,0\ldots,0)} = \frac{1}{6}(\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_3 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_3 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_3 + \mathbf{e}_3 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1)$.

For a vector $\mathbf{x} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) each basis element $\mathbf{E}_\alpha \in S^p(\mathbb{R}^n)$ therefore acts multilinearly on $\mathbf{x}$ as

$$\mathbf{E}_\alpha \cdot \mathbf{x}\mathbf{x} \cdots \mathbf{x} = \mathbf{x}^\alpha \tag{2.5}$$

Therefore

The vector space $S^p(\mathbb{R}^n)$ *is isomorphic to the vector space of homogeneous polynomials of degree* p *from* $\mathbb{R}[x_1, \ldots, x_n]$.

See, for example, Theorem 2.5 p.67 of [Yok92].

In the same way the vector space of (covariant) tensors $T_p(\mathbb{R}^n)$ dual to $T^p(\mathbb{R}^n)$ ([Yok92], pp.53-54) is formed. Putting $\mathbf{s} = \mathbf{x}^1 \otimes \cdots \otimes \mathbf{x}^p \in T_p(\mathbb{R}^n)$, (2.1) can be rewritten as the dual pairing

$$\mathbf{t} \cdot \mathbf{s} = (\mathbf{x}^1 \cdot \mathbf{u}_1)(\mathbf{x}^2 \cdot \mathbf{u}_2) \cdots (\mathbf{x}^p \cdot \mathbf{u}_p) \tag{2.6}$$

A basis for the (covariant) symmetric tensors $S_p(\mathbb{R}^n)$ is defined similarly to (2.2), and basis elements $\mathbf{E}^\alpha$, $|\alpha| = p$, are defined as in (2.4). By the normalizations

$$\mathbf{N}_\alpha = \sqrt{\frac{p!}{\alpha!}}\mathbf{E}_\alpha \text{ and } \mathbf{N}^\alpha = \sqrt{\frac{p!}{\alpha!}}\mathbf{E}^\alpha \tag{2.7}$$

one obtains dual bases

$$\mathbf{N}_\alpha \cdot \mathbf{N}^\beta = \delta_\alpha^\beta \tag{2.8}$$

where the Dirac delta is equal to 0 when $\alpha \neq \beta$ and 1 otherwise.

Because these dual symmetric spaces are isomorphic, no longer will any distinction be made between them. Instead $S^p(\mathbb{R}^n)$ will be considered an inner product space with inner product formed as in (2.6). Bases will be written $\{\mathbf{E}_\alpha : |\alpha| = p\}$, $\{\mathbf{N}_\alpha : |\alpha| = p\}$ an orthogonal and an orthonormal basis respectively. Vectors of $\mathbb{R}^n$ will be enumerated $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{u}^1, \ldots$ with subscripts indicating coordinates $\mathbf{x} = (x_1, x_2, \ldots)$, $\mathbf{x}^1 = (x_1^1, x_2^1, \ldots), \ldots$

A convenient notation for the tensor product of p *identical* vectors is

$$\mathbf{x}^{\otimes p} = \mathbf{x} \otimes \cdots \otimes \mathbf{x} \in S^p(\mathbb{R}^n) \tag{2.9}$$

When $\mathbf{x} \neq \mathbf{0}$ the tensor $\mathbf{x}^{\otimes p}$ will be referred to as a *rank-one tensor* even though it is an element of $S^p(\mathbb{R}^n)$. For example, when $p = 2$ all $n \times n$ symmetric matrices that have rank 1 are given by $\mathbf{x}^{\otimes 2} = \mathbf{x} \otimes \mathbf{x}$. Now (2.5) becomes

$$\mathbf{E}_\alpha \cdot \mathbf{x}^{\otimes p} = \mathbf{x}^\alpha,\ |\alpha| = p.$$

Since $S^p(\mathbb{R}^n)$ is a real vector space, the foregoing can be done with it in place of $\mathbb{R}^n$. Of particular interest is the space $S^2(S^p(\mathbb{R}^n))$ isomorphic to the space of $\begin{pmatrix} n+p-1 \\ p \end{pmatrix} \times \begin{pmatrix} n+p-1 \\ p \end{pmatrix}$ real symmetric matrices. These matrices will be referred to below as the *representation matrices*.

Given any $\mathbf{t} \in S^p(\mathbb{R}^n)$ the notation of (2.9) will be applied as $\mathbf{t}^{\otimes 2} = \mathbf{t} \otimes \mathbf{t} \in S^2(S^p(\mathbb{R}^n))$. Given also s, we introduce the notation

$$\mathbf{s} \otimes_s \mathbf{t} = \mathbf{s} \otimes \mathbf{t} + \mathbf{t} \otimes \mathbf{s}$$

noting that

$$\mathbf{t} \otimes_s \mathbf{t} = 2\mathbf{t} \otimes \mathbf{t}$$

and

$$(\mathbf{s} + \mathbf{t})^{\otimes 2} = \mathbf{s}^{\otimes 2} + \mathbf{s} \otimes_s \mathbf{t} + \mathbf{t}^{\otimes 2}$$

A basis for the vector space $S^2(S^p(\mathbb{R}^n))$ is

$$\{\mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta : |\alpha| = |\beta| = p\} \tag{2.10}$$

It contains $\begin{pmatrix} \begin{pmatrix} n+p-1 \\ p \end{pmatrix} + 1 \\ \\ 2 \end{pmatrix}$ elements. More general elements of $S^2(S^p(\mathbb{R}^n))$ will be denoted in script as with $\mathcal{S}$ or $\mathcal{G}$. All act as symmetric bilinear (quadratic) forms on $S^p(\mathbb{R}^n)$

$$\mathcal{S} \cdot \mathbf{st} = \mathcal{S} \cdot \mathbf{ts}$$

For example

$$\frac{p!p!}{\alpha!\beta!} \mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta \cdot \mathbf{E}_\psi \mathbf{E}_\omega = \delta_\alpha^\psi \delta_\beta^\omega + \delta_\alpha^\omega \delta_\beta^\psi = 0, 1 \text{ or } 2$$

and in particular

$$\frac{1}{2} \mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta \cdot \mathbf{x}^{\otimes p} \mathbf{x}^{\otimes p} = \mathbf{x}^{\alpha+\beta} \tag{2.11}$$

By choosing a linear ordering for the multi-indices of order p, an isomorphism of $S^2(S^p(\mathbb{R}^n))$ and the $\begin{pmatrix} n+p-1 \\ p \end{pmatrix} \times \begin{pmatrix} n+p-1 \\ p \end{pmatrix}$ symmetric matrices can be made explicit. Given (2.11) the one that is apparently most computationally convenient is induced by the mapping

$$\mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta \mapsto \left(\delta_\alpha^\psi \delta_\beta^\omega + \delta_\alpha^\omega \delta_\beta^\psi\right)_{|\psi|=|\omega|=p} \tag{2.12}$$

In this way an element of $S^2(S^p(\mathbb{R}^n))$ is assigned a *representation matrix* and *vice versa*. For example, with linear order $\alpha \prec \beta \prec \cdots$, the tensor $(a_\alpha \mathbf{E}_\alpha + a_\beta \mathbf{E}_\beta + \cdots)^{\otimes 2} =$

$a_\alpha^2\mathbf{E}_\alpha^{\otimes 2} + a_\alpha a_\beta \mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta + \cdots$ is assigned the matrix $\begin{pmatrix} a_\alpha^2 & a_\alpha a_\beta & \cdots \\ a_\alpha a_\beta & a_\beta^2 & \cdots \\ \vdots & \vdots & \end{pmatrix}$, and by (2.11) represents the form $a_\alpha^2\mathbf{x}^{2\alpha} + 2a_\alpha a_\beta \mathbf{x}^{\alpha+\beta} + \cdots = (a_\alpha \mathbf{x}^\alpha + a_\beta \mathbf{x}^\beta + \cdots)^2$.

A tensor of $S^2(S^p(\mathbb{R}^n))$ and its representation matrix will be denoted by the same symbol.

In addition (2.11) shows that

Every element of $S^2(S^p(\mathbb{R}^n))$ represents *a homogeneous polynomial in $\mathbb{R}[x_1,\ldots,x_n]$ of degree $2p$, and every such homogeneous polynomial* can be represented *by an element of $S^2(S^p(\mathbb{R}^n))$.*

Such representations are not unique. $S^2(S^p(\mathbb{R}^n))$ is not isomorphic to $S^{2p}(\mathbb{R}^n)$. The respective dimensions are related by

$$\begin{pmatrix} \begin{pmatrix} n+p-1 \\ p \end{pmatrix} + 1 \\ \\ 2 \end{pmatrix} > \begin{pmatrix} n+2p-1 \\ 2p \end{pmatrix} \tag{2.13}$$

The following can be found on p.109 of [CLR95].

The subspace

$$A^{2,p,n} = \{\Delta \in S^2(S^p(\mathbb{R}^n)) : \Delta \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p} = 0 \text{ for every } \mathbf{x} \in \mathbb{R}^n\} \tag{2.14}$$

has as its dimension the difference of the two numbers in (2.13).

To see this, the basis (2.10) for $S^2(S^p(\mathbb{R}^n))$ can be partitioned into classes

$$\{\mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta : \alpha + \beta = \gamma\}$$

for each $|\gamma| = 2p$, with the number of classes equal to $dim(S^{2p}(\mathbb{R}^n))$. Beginning with a distinguished member of a class, the same span is obtained by the collection

$$\{\mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta, \mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta - \mathbf{E}_{\alpha'} \otimes_s \mathbf{E}_{\beta'}, \mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta - \mathbf{E}_{\alpha''} \otimes_s \mathbf{E}_{\beta''}, \ldots\} \tag{2.15}$$

where $\alpha + \beta = \alpha' + \beta' = \cdots = \gamma$. Every element after the first is in the subspace $A^{2,p,n}$.

By the definition of $A^{2,p,n}$,

Two representation matrices for the same homogeneous polynomial of degree $2p$ always differ by a member of $A^{2,p,n}$.

The members of the subspace $A^{2,p,n}$ when added to a representation matrix for a polynomial *change the representation* of the polynomial and *do not change the polynomial*. When a polynomial has an sos representation, adding what will be called a *change* Δ to that representation might or might not yield another sos representation. In the case it does yield another, it cannot alter the facts that the polynomials of degree p that are squared share or do not share a common *real* root. That they share or do not share a common complex root from $\mathbb{C}^n \setminus \mathbb{CR}^n$, however, possibly can be altered by adding a Δ. Here $\mathbb{CR}^n = \{a\mathbf{x} : a \in \mathbb{C} \text{ and } \mathbf{x} \in \mathbb{R}^n\}$.

Example 2.2. $\Delta = \mathbf{E}_{(2,0)} \otimes_s \mathbf{E}_{(0,2)} - 2\mathbf{E}_{(1,1)}^{\otimes 2}$ may be allowed to serve as the only basis element for $A^{2,2,2}$. Letting $a \in \mathbb{R}$

(2.16)
$$\mathcal{S}_a := (\mathbf{E}_{(2,0)} + \mathbf{E}_{(0,2)})^{\otimes 2} + a\Delta = \mathbf{E}_{(2,0)} + (1+a)\mathbf{E}_{(2,0)} \otimes_s \mathbf{E}_{(0,2)} + \mathbf{E}_{(0,2)} - 2a\mathbf{E}_{(1,1)}^{\otimes 2}$$

when applied to $\mathbf{x}^{\otimes 2}\mathbf{x}^{\otimes 2}$ always yield the pd polynomial $(x_1^2 + x_2^2)^2$. Choosing a linear order $(2,0) \prec (0,2) \prec (1,1)$ for the basis elements of $S^2(\mathbb{R}^2)$, the isomorphism (2.12), of $S^2(S^2(\mathbb{R}^2))$ with the symmetric 3×3 matrices, yields

$$\mathcal{S}_a = \begin{pmatrix} 1 & 1+a & 0 \\ 1+a & 1 & 0 \\ 0 & 0 & -2a \end{pmatrix}$$

The eigenvalues are $-a$, $2+a$ and $-2a$. Using these together with the corresponding unit eigenvectors suggests that (2.16) be written

$$\mathcal{S}_a = -\frac{a}{2}(\mathbf{E}_{(2,0)} - \mathbf{E}_{(0,2)})^{\otimes 2} + \frac{2+a}{2}(\mathbf{E}_{(2,0)} + \mathbf{E}_{(0,2)})^{\otimes 2} - 2a\mathbf{E}_{(1,1)}^{\otimes 2}$$

The representation matrix is psd if and only if $-2 \leq a \leq 0$ if and only if

$$\mathcal{S}_a \cdot \mathbf{x}^{\otimes 2}\mathbf{x}^{\otimes 2} = -\frac{a}{2}(x_1^2 - x_2^2)^2 + \frac{2+a}{2}(x_1^2 + x_2^2)^2 - 2a(x_1x_2)^2$$

is an sos representaion. Among these, each quadratic term has the complex root $\mathbf{x} = (1, i)$ when $a = 0$, while there are no common complex roots when $-2 \leq a < 0$.

This example used the fact that a real symmetric $m \times m$ matrix may be written as an element of $S^2(\mathbb{R}^m)$

$$\sum_{j=1}^{m} \lambda_j \mathbf{u}^j \otimes \mathbf{u}^j \tag{2.17}$$

where the λ_j are eigenvalues counted by multiplicity and $\mathbf{u}^j \in \mathbb{R}^m$ are the corresponding *unit* eigenvectors.

The following proposition can be found in [CLR95] p.106, Proposition 2.3. We include a proof in the multilinear language used here.

Proposition 2.3. *A form $f \in \mathbb{R}[x_1, \dots, x_n]$ of degree $2p$ is an sos* if and only if *there is a psd representation matrix $\mathcal{G}$ such that $f(x) = \mathcal{G} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$.*

Proof. When $\mathcal{G}$ is psd and a representation matrix for f, then $\mathcal{G}$ can be written as a matrix $\sum_{|\beta|=p} \lambda_\beta \mathbf{u}^\beta \otimes \mathbf{u}^\beta$ where the $\mathbf{u}^\beta$ are the unit eigenvectors with $\begin{pmatrix} n+p-1 \\ p \end{pmatrix}$ real components u_α^β for $|\alpha| = p$ and $\lambda_\beta \geq 0$ are the corresponding eigenvalues. By the isomorphism (2.12) it is a tensor $\mathcal{G} = \sum_{|\beta|=p} \lambda_\beta (\sum_{|\alpha|=p} u_\alpha^\beta \mathbf{E}_\alpha)^{\otimes 2}$ that acts as $\mathcal{G} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p} = \sum_{|\beta|=p} \lambda_\beta (\sum_{|\alpha|=p} u_\alpha^\beta \mathbf{x}^\alpha)^2$. Thus f is sos.

If f is sos, then it is a sum of forms

$$\left(\sum_{|\alpha|=p} a_\alpha \mathbf{x}^\alpha\right)^2 = \left(\sum_{|\alpha|=p} a_\alpha \mathbf{E}_\alpha \cdot \mathbf{x}^{\otimes p}\right)^2 = \left(\sum_{|\alpha|=p} a_\alpha \mathbf{E}_\alpha\right)^{\otimes 2} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$$

$a_\alpha \in \mathbb{R}$. $\mathcal{G}$ can be taken to be a sum of tensors $\left(\sum_{|\alpha|=p} a_\alpha \mathbf{E}_\alpha\right)^{\otimes 2}$ each with a psd representation matrix. $\square$

A psd *representation matrix* $\mathcal{G} \in S^2(S^p(\mathbb{R}^n))$ *is also called a* Gram matrix. *For a form f of degree $2p$ to be an sos it is necessary and sufficient that it have a representation* $f(x) = \mathcal{G} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$ *for* some *Gram matrix* $\mathcal{G}$.

An element of $S^2(S^p(\mathbb{R}^n))$ may also be viewed as a linear transformation $\mathbf{t} \mapsto \mathcal{S}\mathbf{t}$ on $S^p(\mathbb{R}^n)$ so that $\mathcal{S} \cdot \mathbf{st} = \mathbf{s} \cdot \mathcal{S}\mathbf{t}$.

Two more elementary but useful observations follow from the characterization of sums of squares given by Proposition 2.3 and elementary properties of psd matrices.

Suppose $\mathcal{G}$ is a Gram matrix. Then the form $\mathcal{G} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$ is positive definite if and only if *the tensor* $(\mathcal{G} + \Delta)\mathbf{x}^{\otimes p} \neq \mathbf{0}$ *for all nonzero* $\mathbf{x} \in \mathbb{R}^n$ *and for all changes* Δ.

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ put $\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathbb{C}^n$. Then formally using the binomial expansion

$$\mathbf{z}^{\otimes p} = \sum_{m=0}^{p} \begin{pmatrix} p \\ m \end{pmatrix} i^m Sym(\mathbf{x}^{\otimes(p-m)} \otimes \mathbf{y}^{\otimes m}) =$$

$$\mathbf{x}^{\otimes p} - \begin{pmatrix} p \\ 2 \end{pmatrix} Sym(\mathbf{x}^{\otimes(p-2)} \otimes \mathbf{y} \otimes \mathbf{y}) + \cdots + i(p\, Sym(\mathbf{x}^{\otimes(p-1)} \otimes \mathbf{y}) - \cdots)$$

$$:= Re\, \mathbf{z}^{\otimes p} + iIm\, \mathbf{z}^{\otimes p}$$

A linear transformation on $S^p(\mathbb{R}^n)$ is extended to complex valued tensors by $\mathcal{S}(\mathbf{s} + i\mathbf{t}) = \mathcal{S}\mathbf{s} + i\mathcal{S}\mathbf{t}$. It follows that $\Delta \cdot \mathbf{z}^{\otimes p}\mathbf{z}^{\otimes p} = 0$ for all changes Δ. This is because the coefficients on the powers of the *real* variable t in $\Delta \cdot (\mathbf{x} + t\mathbf{y})^{\otimes p}(\mathbf{x} + t\mathbf{y})^{\otimes p} = 0$ must all vanish. The same coefficients occur on the unreduced powers of i in $\Delta \cdot \mathbf{z}^{\otimes p}\mathbf{z}^{\otimes p}$. Or one can invoke the multi-index formalism. Similarly, by comparing coefficients between binomial expansions, (2.5) extends to complex rank-one tensors

$$\mathbf{E}_\alpha \cdot \mathbf{z}^{\otimes p} = \mathbf{z}^\alpha$$

(2.18)
Let $\mathcal{S}$ be a representation matrix. Then $\mathcal{S} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$ is a coercive sos if and only if *there exists a Δ such that $\mathcal{S} + \Delta$ is a Gram matrix, and for every nonzero $z \in \mathbb{C}^n$ the tensor* $(\mathcal{S} + \Delta)\mathbf{z}^{\otimes p} \neq \mathbf{0}$.

For when $\mathcal{S} + \Delta$ is a Gram matrix it may be written $\sum \mathbf{g}_j \otimes \mathbf{g}_j$ with the collection of $\mathbf{g}_j \in S^p(\mathbb{R}^n)$ linearly independent; and $(\mathcal{S} + \Delta)\mathbf{z}^{\otimes p} = \sum(\mathbf{g}_j \cdot \mathbf{z}^{\otimes p})\mathbf{g}_j$.

The strategy, then, for showing that a positive definite sos is a coercive sos is to change the Gram matrix, preserving its psd property, in order to eliminate from the null space all 2-dimensional subspaces of the form $span\{\mathbf{s}, \mathbf{t}\}$ where $\mathbf{s} + i\mathbf{t} = \mathbf{z}^{\otimes p}$ for nonzero $\mathbf{z} \in \mathbb{C}$. In this way the point of view of this article is opposite that of some literature growing out of Hilbert's theorems on sums of squares. For example, the coercive result (1.10) is achieved by eliminating the nontrivial null space altogether, i.e. showing that pd Gram matrices exist for those cases. On the other hand, the most remarkable and difficult result of Hilbert's is that for the cone $P_{3,4}$, where the rank of a Gram matrix can be as large as 6, every polynomial can be written a sum of just 3 squares. Out of this came the general idea of the *length* or minimum number of squares required for an sos representation and out of this the *Pythagoras number*, the minimum number of squares needed over a collection of sos polynomials. See, for example, [BCR98], [CLR95], [Pfi95], [PD01] and others.

For coerciveness the length of an sos is often an undesirable number, and one naturally wishes to *maximize* the number of independent squares in a representation. That this is an interesting problem is shown here by demonstrating, in the case of a *positive definite* polynomial with psd representation (Gram) matrix, that the rank of its Gram matrices cannot in general be increased enough to achieve the desired end, vis. coerciveness.

We end this section by restating question (1.1) in multilinear language and by outlining the construction by which the answer is shown to be *no* in general.

Suppose $\mathcal{G} \in S^2(S^p(\mathbb{R}^n))$ *is a Gram matrix and* $\mathcal{G}\mathbf{x}^{\otimes p} \neq \mathbf{0}$ *for all rank-one tensors . Does there exist a change* Δ *such that* $\mathcal{G} + \Delta$ *is a Gram matrix and* $(\mathcal{G} + \Delta)\mathbf{z}^{\otimes p} \neq \mathbf{0}$ *for all nonzero* $z \in \mathbb{C}$?

Or less precisely, can a Gram matrix $\mathcal{G}$ that is pd on the rank-one tensors be *changed* to be a Gram matrix that is pd on all subspaces of the form $span\{\mathbf{s}, \mathbf{t}\}$ where $\mathbf{s} + i\mathbf{t} = \mathbf{z}^{\otimes p}$ for some nonzero $\mathbf{z} \in \mathbb{C}$?

The question is answered below in the negative, for the cases $n \geq 6$, $p = 2$ and $n \geq 4$, $p = 3$, by the construction

(2.19) Construct a Gram matrix $\mathcal{G}$ such that

(i) $\mathcal{G}$ is positive definite on the rank-one tensors.

(ii) there exists a nonzero $\mathbf{z} \in \mathbb{C}^n$ such that the tensor $\mathcal{G}\mathbf{z}^{\otimes p} = \mathbf{0}$.

(iii) $\mathcal{G} + \Delta$ is never a Gram matrix whenever $\Delta\mathbf{z}^{\otimes p} \neq \mathbf{0}$.

A uniqueness condition stronger than (iii) is

(iii)$'$ $\mathcal{G} + \Delta$ is never a Gram matrix whenever $\Delta \neq \mathbf{0}$.

3. A POSITIVE DEFINITE QUARTIC WITH A UNIQUE GRAM MATRIX

In this section an element of $\Sigma_{4,4}$ is constructed that satifies (i) and (iii)$'$ of the construction (2.19), but not (ii).

The vector space of representation matrices $S^2(S^p(\mathbb{R}^n))$ inherits a topology from the Euclidean space of the same dimension. The closed cone of Gram matrices will have as its interior the cone of positive definite Gram matrices. The boundary of this cone is the set of Gram matrices with rank less than $\begin{pmatrix} n+p-1 \\ p \end{pmatrix}$.

Part (ii) of the construction (2.19) cannot be realized if $\mathcal{G}$ is taken in the interior of the cone. Thus $\mathcal{G}$ must be on the boundary if one hopes to realize (ii) and one is led to consider pd polynomials of degree $2p$ that border those that are not sums of squares. Historically pd and psd polynomials that are not sos are difficult to locate. It is therefore sensible to begin with a known pd polynomial that is not sos, i.e. does not have a Gram matrix but is definite on the rank-one tensors, and perturb it in such a way so that one arrives at the boundary of the Gram matrices while maintaining the rank-one definiteness. Here we take $n = 4, p = 2$, let $\mathbf{x} \in \mathbb{R}^4$ correspond to (w, x, y, z) and begin with the Choi-Lam quartics q_η (1.11), (1.12), letting η increase until the quartic (1.13) is achieved.

Except for the uniqueness of representation claim, all other claims made for (1.13) in Section 1 can be quickly proved.

1. By expanding the right side of (1.13) and collecting terms the right side meets the definition of q_{η_0} (1.12) if the coefficients on the x^2y^2, y^2z^2 and z^2x^2 terms equal 1. This occurs when

2. $\sqrt{\eta_o}$ is a root of $X^3 - \frac{1}{2}X + \frac{1}{9} = 0$.
3. $\sqrt{\eta_o}$ must be chosen to be the smallest positive root, else η_0 would not be the smallest η that makes q_η an sos.

Since degree and dimension are low in this section, tensors $\mathbf{E}_\alpha$ will be denoted by using only the entries of each multi-index as subscripts, as in $\mathbf{E}_{ijkl}$ instead of $\mathbf{E}_{(i,j,k,l)}$. Thus $\mathbf{E}_{2000} \cdot \mathbf{x}^{\otimes 2} = x_1^2 = w^2$, etc.

4. That η_0, as described in Claims 2 and 3, is the *smallest* η for which q_η is an sos will follow once it is proved that

$$\mathcal{Q}_{\eta_0} = (\mathbf{E}_{2000} - \sqrt{\eta_0}(\mathbf{E}_{0200} + \mathbf{E}_{0020} + \mathbf{E}_{0002}))^{\otimes 2} + \frac{2}{9\sqrt{\eta_0}}\left[(3\sqrt{\eta_0}\mathbf{E}_{1100} - \mathbf{E}_{0011})^{\otimes 2} + (3\sqrt{\eta_0}\mathbf{E}_{1010} - \mathbf{E}_{0101})^{\otimes 2} + (3\sqrt{\eta_0}\mathbf{E}_{1001} - \mathbf{E}_{0110})^{\otimes 2}\right] \tag{3.1}$$

is the *unique* Gram matrix $\mathcal{G}$ for which $q_{\eta_0}(\mathbf{x}) = \mathcal{G} \cdot \mathbf{x}^{\otimes 2}\mathbf{x}^{\otimes 2}$. For if q_η were an sos for some $\eta < \eta_0$, then

$$q_{\eta_0} = q_\eta + (\eta_0 - \eta)(x^4 + y^4 + z^4) = q_\eta + (\eta_0 - \eta)((x^2 - y^2)^2 + (\sqrt{2}xy)^2 + z^4) \tag{3.2}$$

and the polynomial identity presents two different Gram matrices for q_{η_0}. Letting $\mathcal{Q}_\eta$ be, by Proposition 2.3, a Gram matrix for q_η, q_{η_0} now has both

$$\mathcal{Q}_\eta + (\eta_0 - \eta)(\mathbf{E}_{0200}^{\otimes 2} + \mathbf{E}_{0020}^{\otimes 2} + \mathbf{E}_{0002}^{\otimes 2}) \tag{3.3}$$

and

$$\mathcal{Q}_\eta + (\eta_0 - \eta)((\mathbf{E}_{0200} - \mathbf{E}_{0020})^{\otimes 2} + 2\mathbf{E}_{0110}^{\otimes 2} + \mathbf{E}_{0002}^{\otimes 2}) \tag{3.4}$$

as Gram matrices. They differ by $\Delta = (\eta_0 - \eta)(2\mathbf{E}_{0110}^{\otimes 2} - \mathbf{E}_{0200} \otimes_s \mathbf{E}_{0020})$ contradicting the uniqueness of $\mathcal{Q}_{\eta_0}$.

Remark 3.1. In contrast, the identity $2x^4 + 2y^4 = (x^2 - y^2)^2 + (x^2 + y^2)^2$ suggests $2\mathbf{E}_{0200}^{\otimes 2} + 2\mathbf{E}_{0020}^{\otimes 2}$ and $(\mathbf{E}_{0200} - \mathbf{E}_{0020})^{\otimes 2} + (\mathbf{E}_{0200} + \mathbf{E}_{0020})^{\otimes 2}$ which are identical Gram matrices. The two polynomial expressions are said to be obtained from one another by *orthogonal transformation*. See Proposition 2.10 of [CLR95], p.108. It is for this reason that by themselves it is not clear that each of (3.3) or (3.4) differs from $\mathcal{Q}_{\eta_0}$ since $\mathcal{Q}_\eta$ is unspecified.

5. That q_{η_0} is *coercive* is seen by showing that the corresponding homogeneous system of four quadratic equations has no solution in $\mathbb{C}^4 \setminus \{\mathbf{0}\}$. One starts with assuming a solution (w, x, y, z) has one of its coordinates equal to zero, cases that can be quickly eliminated. Then, assuming a solution has all nonzero coordinates, one has by using the last three quadratics of (1.13), $y^2z = 3\sqrt{\eta_0}wxy = zx^2$ etc., whence $x^2 = y^2 = z^2$, whence $3\sqrt{\eta_0}|w| = |x|$ by any of the last three quadratics. Then $|w|^2 = 3\sqrt{\eta_0}|x|^2$ by the first, whence $\sqrt{\eta_0} = \frac{1}{3}$ which is not true by Claim 2.

The only task remaining is to prove the uniqueness of the Gram matrix $\mathcal{Q}_{\eta_0}$. Before that is done a bit more will be said about finding (1.13).

An initial choice of representation matrices for the forms q_η is

$$\mathcal{S}_\eta = \mathbf{E}_{2000}^{\otimes 2} + \mathbf{E}_{0110}^{\otimes 2} + \mathbf{E}_{0011}^{\otimes 2} + \mathbf{E}_{0101}^{\otimes 2} - \frac{2}{3}(\mathbf{E}_{1100} \otimes_s \mathbf{E}_{0011} + \mathbf{E}_{1010} \otimes_s \mathbf{E}_{0101} + \mathbf{E}_{1001} \otimes_s \mathbf{E}_{0110}) + \eta(\mathbf{E}_{0200}^{\otimes 2} + \mathbf{E}_{0020}^{\otimes 2} + \mathbf{E}_{0002}^{\otimes 2}) \tag{3.5}$$

The q_η are symmetric in x, y and z. As η increases, if $\mathcal{G}$ becomes the first Gram matrix encountered so would be $\mathcal{G}'$ where $\mathcal{G}'$ is derived from $\mathcal{G}$ by permuting the indices for x, y and z. Averaging all such permutations would produce a first Gram matrix that was symmetric in x, y and z. Therefore the symmetry in the choice of $\mathcal{S}_\eta$ is no loss of generality, and we expect that if a Gram matrix uniquely represents a q_η, then it will be symmetric in x, y and z.

Arrange the basis elements $\mathbf{E}_{2000}, \ldots$ according to the linear order $w^2 \prec x^2 \prec y^2 \prec z^2 \prec wx \prec yz \prec wy \prec zx \prec wz \prec xy$. Then the matrix for $\mathcal{S}_\eta$ with respect to the basis (2.4) is

$$\tag{3.6} \begin{pmatrix} 1 & -b & -b & -b & & & & & & \\ -b & \eta & a & a & & & & & & \\ -b & a & \eta & a & & & & & & \\ -b & a & a & \eta & & & & & & \\ & & & & 2b & -\frac{2}{3} & & & & \\ & & & & -\frac{2}{3} & 1-2a & & & & \\ & & & & & & 2b & -\frac{2}{3} & & \\ & & & & & & -\frac{2}{3} & 1-2a & & \\ & & & & & & & & 2b & -\frac{2}{3} \\ & & & & & & & & -\frac{2}{3} & 1-2a \end{pmatrix}$$

when the *parameters* $a = b = 0$. The unmarked entries are $zero$.

The two parameters permit the addition of six *changes* in a way that also obey the symmetry considerations in x, y and z. The smallest value of η that allows a choice of a and b so that each of the four block matrices becomes rank-1 and psd is the η_0 defined above. The minimizing choices are $a = \eta_0$ and $b = \sqrt{\eta_0}$.

There are, however, *twenty* independent changes Δ in $S^2(S^2(\mathbb{R}^4))$ altogether. Though the type of argument being given can be made rigorous and lead to a uniqueness proof for $\mathcal{Q}_{\eta_0}$, we will instead present another argument which will also be elementary, but also clearly decisive while computationally not too long if Maple^{TM}10 is used. It is based on the observation

Suppose $\mathcal{G}$ is a Gram matrix. Then a necessary (but not sufficient) condition for $\mathcal{G}+\Delta$ to be a Gram matrix is that Δ be psd *on $Null(\mathcal{G})$, the null space of $\mathcal{G} : S^p(\mathbb{R}^n) \to S^p(\mathbb{R}^n)$, i.e. for every $\mathbf{t} \in Null(\mathcal{G})$ it is necessary that $\Delta \cdot \mathbf{tt} \geq 0$.*

Let N be a nonempty subspace of $S^p(\mathbb{R}^n)$. When $\mathcal{S} \cdot \mathbf{tt} \geq 0$ fails to hold for some $\mathbf{t} \in N$ while $\mathcal{S} \cdot \mathbf{ss} > 0$ for an $\mathbf{s} \in N$, $\mathcal{S}$ is said to be *not definite* on N. Thus

(3.7)

If $f(x) = \mathcal{G} \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$ where $\mathcal{G}$ is a Gram matrix and if every *nonzero $\Delta \in A^{2,p,n}$ is* not definite *on $Null(\mathcal{G})$, then $\mathcal{G}$ is the* unique *Gram matrix for f.*

This is in fact a statement about subspaces of $S^p(\mathbb{R}^n)$ and the Gram matrices that can be supported on their orthogonal complements. Consequently

Let N be a subspace of $S^p(\mathbb{R}^n)$ and $\{\mathbf{t}_1, \ldots, \mathbf{t}_r\}$ a basis for its orthogonal complement M. Suppose every *nonzero $\Delta \in A^{2,p,n}$ is* not definite *on N. Let T be any linear* *transformation on M. Then $\mathcal{G}_T = (T(\mathbf{t}_1))^{\otimes 2}+\cdots+(T(\mathbf{t}_r))^{\otimes 2}$ is the* unique *Gram matrix for the sos $f_T(\mathbf{x}) = \mathcal{G}_T \cdot \mathbf{x}^{\otimes p}\mathbf{x}^{\otimes p}$. The collection of all such f_T is a convex cone of $\Sigma_{n,2p}$.*

The last statement follows because if $\mathcal{G}_T$ and $\mathcal{G}_U$ are psd on M so is their sum which will be given by some $\mathcal{G}_V$ with the linear transformation V on M derived, for example, by using (2.17).

Remark 3.2. If, for example, I is the identity on M and U is an orthogonal transformation on M, then $f_I = f_U$. This is Proposition 2.10 of [CLR95] again.

Given a subspace $N \subset S^p(\mathbb{R}^n)$ of dimension m the following steps will be carried out in order to prove that certain *sums of squares*, supported like the above f_T on the orthogonal complement of N, have unique Gram matrices.

1. Form a general linear combination $\mathbf{t} = a\mathbf{t}_1 + b\mathbf{t}_2 + \cdots$ of the m basis elements of N.

2. Apply each element Δ of a basis for $A^{2,p,n}$ (2.14) to the general linear combination, as $\Delta \cdot \mathbf{tt}$, yielding a set of homogeneous quadratic polynomials in the m *variables* $a, b, \ldots$

3. Thinking of each quadratic polynomial from Step 2 as a *linear* expression in the monomials $a^2, b^2, \ldots, ab, ac, \ldots, bc, bd, \ldots$, write the $\left(\begin{array}{c}\left(\begin{array}{c} n+p-1 \\ p \end{array}\right)+1 \\ 2\end{array}\right) - \left(\begin{array}{c} n+2p-1 \\ 2p \end{array}\right)$ by $\left(\begin{array}{c} m+1 \\ 2 \end{array}\right)$ coefficient matrix for these linear expressions.

4. Bring the coefficient matrix of Step 3 to reduced row echelon form thereby obtaining a set of quadratic polynomials that is equivalent to the set of Step 2, i.e. each set of quadratics consists of only linear combinations of quadratics from the other.

5. Show that no nontrivial linear combination of the quadratics from Step 4 yields a definite or semi-definite quadratic in the m variables.

Remark 3.3. Steps 1 through 4 can be thought of as supplying details for an algorithm designed to show a certain semi-algebraic set consists (here) of one point (the origin). See the second algorithmic step and the remark that follows on p. 101 of [PW98]. Here it is Step 5 that is uncertain.

In the case of interest here, there are $m = 6$ variables a, b, c, d, e, f and the coefficient matrix is 20×21, more quadratic monomials than quadratic polynomials.

To simplify calculation, $\mathbb{R}^4$ (and thus (3.1)) is scaled in the variable w, replaced with $\frac{w}{3\sqrt{\eta_0}}$. Define

$$\gamma_0 := 27\eta_0^{3/2}$$

Then (3.1) is a linear combination with *positive* coefficients of the tensors

$$(3.8) \quad (3\mathbf{E}_{2000} - \gamma(\mathbf{E}_{0200} + \mathbf{E}_{0020} + \mathbf{E}_{0002}))^{\otimes 2}, \\ (\mathbf{E}_{1100} - \mathbf{E}_{0011})^{\otimes 2}, (\mathbf{E}_{1010} - \mathbf{E}_{0101})^{\otimes 2}, \text{ and } (\mathbf{E}_{1001} - \mathbf{E}_{0110})^{\otimes 2}$$

when $\gamma = \gamma_0$. By Claims 2 and 3 at the beginning of this section the estimate $\sqrt{\eta_0} < 1/3$ holds, whence $0 < \gamma_0 < 1$. Thus all assertions about q_{η_0} (1.13) will hold once the following theorem is proved.

Theorem 3.4. *Given any γ, $0 < \gamma < 1$, and any choice of $a_j > 0,\ j = 1,2,3,4$, the quartic form of* $\mathbb{R}[w,x,y,z]$

$$a_1(3w^2 - \gamma(x^2+y^2+z^2))^2 + a_2(wx-yz)^2 + a_3(wy-zx)^2 + a_4(wz-xy)^2 \tag{3.9}$$

is coercive and has a unique Gram matrix.

Proof. Coerciveness follows as for q_{η_0} in Claim 5 at the beginning of this section.

Fix any $0 < \gamma < 1$ and denote by $\mathcal{G}_\gamma$ any linear combination, with *positive* coefficients, of the tensors (3.8). A basis for the null space of $\mathcal{G}_\gamma$ is supplied by

$$\mathbf{E}_{1100} + \mathbf{E}_{0011},\ \mathbf{E}_{1010} + \mathbf{E}_{0101},\ \mathbf{E}_{1001} + \mathbf{E}_{0110},\ \gamma\mathbf{E}_{2000} + \mathbf{E}_{0200} + \mathbf{E}_{0020} + \mathbf{E}_{0002},$$
$$\mathbf{E}_{0200} - \mathbf{E}_{0020},\ \text{and}\ \mathbf{E}_{0200} - \mathbf{E}_{0002}$$

as (2.3), (2.7) and (2.8) show. A general linear combination of these is $\mathbf{g} = 2a\mathbf{E}_{1100} + 2a\mathbf{E}_{0011} + 2b\mathbf{E}_{1010} + 2b\mathbf{E}_{0101} + 2c\mathbf{E}_{1001} + 2c\mathbf{E}_{0110} + \gamma d\mathbf{E}_{2000} + (d+e+f)\mathbf{E}_{0200} + (d-e)\mathbf{E}_{0020} + (d-f)\mathbf{E}_{0002}$

A basis for the changes $A^{2,2,4}$ divides into three sets depending on the number of multi-indices α with $\alpha! = 2$ that are used to express a Δ. The first type has two such α as in

$$\mathbf{E}_{0110}^{\otimes 2} - \frac{1}{2}\mathbf{E}_{0200} \otimes_s \mathbf{E}_{0020}$$

there are 6 of these altogether. The second type uses one as in

$$\frac{1}{2}\mathbf{E}_{2000} \otimes_s \mathbf{E}_{0110} - \frac{1}{2}\mathbf{E}_{1100} \otimes_s \mathbf{E}_{1010}$$

There are 12 of these. Finally there are only 2 independent changes that use no $\alpha! = 2$. We will use

$$\frac{1}{2}\mathbf{E}_{1100} \otimes_s \mathbf{E}_{0011} - \frac{1}{2}\mathbf{E}_{1010} \otimes_s \mathbf{E}_{0101}\ \text{and}\ \frac{1}{2}\mathbf{E}_{1100} \otimes_s \mathbf{E}_{0011} - \frac{1}{2}\mathbf{E}_{1001} \otimes_s \mathbf{E}_{0110}$$

The last type was used implicitly in the initial choice (3.5). The first type was introduced by the parameters in (3.6).

Keeping in mind that by (2.7) and (2.8) $\mathbf{E}_\alpha \cdot \mathbf{E}_\alpha = \frac{\alpha!}{2}$ and computing $\Delta \cdot \mathbf{gg}$ we obtain

$$\begin{aligned}
&a^2 - \gamma d(d+e+f)\\
&b^2 - \gamma d(d-e)\\
&c^2 - \gamma d(d-f)\\
&a^2 - (d-e)(d-f)\\
&b^2 - (d+e+f)(d-f)\\
&c^2 - (d+e+f)(d-e)
\end{aligned}$$

then

$$\begin{aligned}
&\gamma da - bc\\
&\gamma db - ac\\
&\gamma dc - ab\\
&(d+e+f)a - bc\\
&(d+e+f)b - ac\\
&(d+e+f)c - ab\\
&(d-e)a - bc\\
&(d-e)b - ac\\
&(d-e)c - ab\\
&(d-f)a - bc
\end{aligned}$$

$(d-f)b-ac$
$(d-f)c-ab$

and then

a^2-b^2
a^2-c^2

Linearly ordering the monomial squares in alphabetical order followed by the indefinite monomials in alphabetical order $a^2, b^2, \dots, f^2, ab, ac, \dots, af, bc, \dots, df, ef$ the 20×21 coefficient matrix of Step 3 above is obtained. Passing to reduced row echelon form, a matrix that consists of a 20×20 *identity* matrix together with a *21st column* with successive entries

$$\frac{\gamma}{1-\gamma}, \frac{\gamma}{1-\gamma}, \frac{\gamma}{1-\gamma}, \frac{1}{1-\gamma}, 2, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0$$

is obtained.

Thus an equivalent set of quadratic polynomials is

$$\begin{gathered} a^2+\frac{\gamma}{1-\gamma}ef \\ b^2+\frac{\gamma}{1-\gamma}ef \\ c^2+\frac{\gamma}{1-\gamma}ef \\ d^2+\frac{1}{1-\gamma}ef \\ e^2+2ef \\ f^2+2ef \end{gathered} \tag{3.10}$$

together with the collection of 14 indefinite monomials $ab, ac, \dots, df$ (ef not included). Precisely when $0<\gamma<1$ is there no nontrivial linear combination of these that yields a definite or semi-definite quadratic polynomial. Thus uniqueness follows from (3.7). $\square$

More generally, the quartics (3.9) are pd whenever $\gamma\neq 0$ and $\gamma\neq 1$. When $\gamma<0$, expanding the first square makes it transparent that the quartics (3.9) have *positive definite* Gram matrices and are thus coercive sos. When $\gamma>1$ it is not clear in this way, but it is clear from (3.10) that there is a Δ that is positive definite on the null space of the $\mathcal{G}_\gamma$ (from the proof) that represents a (3.9). By taking $\epsilon>0$ small enough $\mathcal{G}_\gamma+\epsilon\Delta$ will be pd by the proposition below.

In some cases there only exist nontrivial Δ that are positive *semi-definite* on the null space of a psd $\mathcal{G}$. In those cases the proposition below gives necessary and sufficient conditions for $\mathcal{G}+\epsilon\Delta$ to be psd, i.e. for the associated sos to *not* have a unique Gram matrix. When $Null(\mathcal{G})\cap Null(\Delta)\neq Null(\mathcal{G})$ the propsition gives necessary and sufficient conditions for $\mathcal{G}+\epsilon\Delta$ to be psd with greater rank than $\mathcal{G}$. It provides conditions to *build up* the ranks of Gram matrices associated to an sos in an attempt to prove coerciveness of the sos.

The *length* of a vector $\mathbf{x}\in\mathbb{R}^m$ is denoted $|\mathbf{x}|$ and the *operator norm* of an $m\times m$ matrix B, as a transformation on $\mathbb{R}^m$, is denoted $|B| = \max_{|\mathbf{x}|=1}|B\mathbf{x}|$.

Proposition 3.5. *Let A be real symmetric positive semi-definite $m \times m$ matrix. Let B be real symmetric $m \times m$ matrix that is psd on $Null(A) \subset \mathbb{R}^m$, i.e. $\mathbf{z} \cdot B\mathbf{z} \geq 0$ for all $\mathbf{z} \in Null(A)$.*

Then for all $\epsilon > 0$ small enough $A + \epsilon B$ is a positive semi-definite matrix if and only if *whenever $\mathbf{z}_1 \in Null(A)$ and $\mathbf{z}_1 \cdot B\mathbf{z}_1 = 0$ it follows that $B\mathbf{z}_1 = \mathbf{0}$.*

In the case $A + \epsilon B$ is psd $Null(A + \epsilon B) \subset Null(A)$ for all $\epsilon > 0$ small enough, with strict containment when $\mathbf{z} \cdot B\mathbf{z}$ does not vanish for every $\mathbf{z} \in Null(A)$.

If B is pd on $Null(A)$ then $A + \epsilon B$ is pd for all $\epsilon > 0$ small enough.

Proof. A and B are assumed nontrivial. The last statement is proved first.

Let $a > 0$ be the smallest nonzero eigenvalue of A. Let $b > 0$ be the smallest number satisfying $\mathbf{z} \cdot B\mathbf{z} \geq b|\mathbf{z}|^2$ for all $\mathbf{z} \in Null(A)$. Each $\mathbf{x} \in \mathbb{R}^m$ has a unique decomposition $\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{z} \in Null(A)$ and $\mathbf{y}$ is orthogonal to $Null(A)$, i.e. by the symmetry of A, each $\mathbf{y}$ is a sum of the eigenvectors of A that have *positive* eigenvalues. Thus

$$\begin{aligned} \mathbf{x} \cdot (A + \epsilon B)\mathbf{x} = \mathbf{y} \cdot A\mathbf{y} + \epsilon\mathbf{y} \cdot B\mathbf{y} + 2\epsilon\mathbf{y} \cdot B\mathbf{z} + \epsilon\mathbf{z} \cdot B\mathbf{z} \geq \\ a|\mathbf{y}|^2 - \epsilon|B||\mathbf{y}|^2 - 2\epsilon|B||\mathbf{y}||\mathbf{z}| + \epsilon b|\mathbf{z}|^2 \end{aligned} \tag{3.11}$$

For $\mathbf{x} \neq \mathbf{0}$ this last quantity will always be positive for any ϵ satisfying $0 < \epsilon < \frac{ab}{|B|^2 + b|B|}$, proving the positive definiteness of $A + \epsilon B$.

Now assume B is psd on $Null(A)$. The first conclusion is proved next.

Assume for some $\epsilon > 0$ that $A + \epsilon B$ is psd. Let $\mathbf{z}_0 \in Null(A)$ and assume $\mathbf{z}_0 \cdot B\mathbf{z}_0 = 0$. Thus $\mathbf{z}_0 \cdot (A + \epsilon B)\mathbf{z}_0 = 0$. Since $A + \epsilon B$ has a psd square root it follows that $(A + \epsilon B)\mathbf{z}_0 = 0$ whence $B\mathbf{z}_0 = \mathbf{0}$.

For the other direction and for each $\mathbf{x} \in \mathbb{R}^m$, with $\mathbf{x} = \mathbf{y} + \mathbf{z}$ as before, the *equality* in (3.11) is again obtained. Each $\mathbf{z} \in Null(A)$ has a unique decomposition $\mathbf{z} = \mathbf{z}_0 + \mathbf{z}_1$ where $\mathbf{z}_0 \in Null(A) \cap Null(B)$ and $\mathbf{z}_1 \in Null(A)$ is *orthogonal* to $Null(A) \cap Null(B)$. In the event $Null(A) \cap Null(B) = Null(A)$ it follows that $\mathbf{z} = \mathbf{z}_0$ and (3.11) yields $\mathbf{x} \cdot (A + \epsilon B)\mathbf{x} \geq a|\mathbf{y}|^2 - \epsilon|B||\mathbf{y}|^2 \geq 0$ for every $\mathbf{x}$ if ϵ is small enough, with vanishing occurring only when $\mathbf{x} \in Null(A)$. Otherwise there is a smallest number $b_1 > 0$ such that $\mathbf{z}_1 \cdot B\mathbf{z}_1 \geq b_1|\mathbf{z}_1|^2$ for all $\mathbf{z}_1 \in Null(A)$ orthogonal to $Null(A) \cap Null(B)$. This follows by the hypothesis, $\mathbf{z}_1 \cdot B\mathbf{z}_1 = 0$ implies $B\mathbf{z}_1 = \mathbf{0}$, whence $\mathbf{z}_1 \in Null(A) \cap Null(B)$ whence $\mathbf{z}_1 = \mathbf{0}$. Consequently $\mathbf{z}$ may be replaced by $\mathbf{z}_1$ and b by b_1 in (3.11). For all $\mathbf{x} \notin Null(A) \cap Null(B)$ and $\epsilon > 0$ small enough (3.11) is then *positive*, completing the proof of the first conclusion.

It has been shown for $\epsilon > 0$ small enough that positivity of (3.11) fails only when $\mathbf{x} \in Null(A) \cap Null(B)$, proving the second conclusion.

$\square$

Example 3.6. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is psd and $B = \begin{pmatrix} 0 & b & 0 \\ b & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is psd on $Null(A)$, but whenever $b \neq 0$ and $\epsilon \neq 0$ $A + \epsilon B$ is not psd.

This phenomenon persists when the B are specialized to represent *changes* Δ. Consider the coercive sos in noncoercive representation $(x^2 + y^2)^2 + z^4 + y^2z^2 + x^2z^2$, i.e. with Gram matrix $\mathcal{A} = (\mathbf{E}_{200} + \mathbf{E}_{020})^{\otimes 2} + \mathbf{E}_{002}^{\otimes 2} + \mathbf{E}_{011}^{\otimes 2} + \mathbf{E}_{101}^{\otimes 2}$. Then $\Delta = \mathbf{E}_{002} \otimes_s \mathbf{E}_{110} - \mathbf{E}_{011} \otimes_s \mathbf{E}_{101}$ is trivially psd on $Null(\mathcal{A})$, but $\mathcal{A} + \epsilon\Delta$ is not psd unless $\epsilon = 0$. Here $\Delta \cdot \mathbf{E}_{110}\mathbf{E}_{110} = 0$ while $\Delta\mathbf{E}_{110} = \frac{1}{2}\mathbf{E}_{002}$.

4. Proof of Theorem 1.4

Theorem 1.4 follows from the next theorem.

Theorem 4.1. *Given γ, $0 < \gamma < 1/3$, the positive definite quartic form of $\mathbb{R}[u, v, w, x, y, z]$*

(4.1)
$$f = (u^2+v^2+vw)^2+(w^2-\gamma(x^2+y^2+z^2))^2+(wx-yz)^2+(wy-zx)^2+(wz-xy)^2$$

is a noncoercive sum of squares.

Proof. The last four terms sum to a pd form over $\mathbb{R}^4$ as shown in the last section. From this, positive definiteness over $\mathbb{R}^6$ follows. On the other hand $(1, i, 0, 0, 0, 0) \in \mathbb{C}^6$ is a root for each of the five squared quadratics, i.e. the real and imaginary parts of

$$(\mathbf{e}^1 + i\mathbf{e}^2)^{\otimes 2} = \mathbf{E}_{200000} - \mathbf{E}_{020000} + 2i\mathbf{E}_{110000} := \mathbf{r} + i\mathbf{q} \tag{4.2}$$

are in the null space of the Gram matrix $\mathcal{G}_0$ that gives representation (4.1) for f. Using (2.18), noncoerciveness of f will be proved by showing that *every* Gram matrix for f contains $\mathbf{r}$ and $\mathbf{q}$ (4.2) in its null space.

Denote $\Delta_1 = -\frac{1}{2}\mathbf{E}_{200000} \otimes_s \mathbf{E}_{020000} + \mathbf{E}_{110000}^{\otimes 2}$. Then

$$\Delta_1 \cdot \mathbf{rr} = \Delta_1 \cdot \mathbf{qq} = 1 \tag{4.3}$$

There is a basis

$$\{\Delta_1, \Delta_2, \ldots, \Delta_{105}\} \tag{4.4}$$

for $A^{2,2,6}$ with Δ_1 (4.3) as its first member so that

$$\Delta_j \cdot \mathbf{rr} = \Delta_j \cdot \mathbf{qq} = 0 \tag{4.5}$$

for all $j = 2, 3, \ldots, 105$. This follows because the basis elements of (2.15) $\mathbf{E}_\alpha \otimes_s \mathbf{E}_\beta - \mathbf{E}_{\alpha'} \otimes_s \mathbf{E}_{\beta'}$, $\alpha + \beta = \alpha' + \beta'$, permit one of the equalities in (4.5) *not* to hold only when either both $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$ are contained in $\{\mathbf{E}_{200000}, \mathbf{E}_{020000}, \mathbf{E}_{110000}\}$ or both $\mathbf{E}_{\alpha'}$ and $\mathbf{E}_{\beta'}$ are contained. The only basis element like this is $\pm\Delta_1$.

Remark 4.2. This relationship between a $\mathbf{z}^{\otimes 2}$, $\mathbf{z} \in \mathbb{C}^n$, and some basis for $A^{2,2,n}$ is general. The uniqueness does not quite hold in $A^{2,p,n}$, $p \geq 3$, however. For example, both $\mathbf{E}_{12}^{\otimes 2} - \frac{1}{2}\mathbf{E}_{21} \otimes_s \mathbf{E}_{03}$ and $\mathbf{E}_{21}^{\otimes 2} - \frac{1}{2}\mathbf{E}_{12} \otimes_s \mathbf{E}_{30}$ are nonzero as quadratic forms on the real and imaginary parts of $(\mathbf{e}^1 + i\mathbf{e}^2)^{\otimes 3}$.

If Δ_1 is removed from the basis (4.4) and Δ is taken in the subsequent span so that $\mathcal{G}_0 + \Delta$ is a Gram matrix, Proposition 3.5 and (4.5) then imply that $\mathbf{r}$ and $\mathbf{q}$ will also be in the null space of $\mathcal{G}_0 + \Delta$. Together with (4.3) this implies

(4.6)
Any linear combination Δ of basis elements (4.4)*, for which $\mathcal{G}_0 + \Delta$ is a Gram matrix and for which at least one of $\mathbf{r}$ or $\mathbf{q}$ is* not *in the null space of $\mathcal{G}_0 + \Delta$, must have a* positive *coefficient on Δ_1.*

Hence let $\delta > 0$ and consider the following principal submatrix of $\mathcal{G}_0 + 2\delta\Delta_1$ where the order $\mathbf{E}_{011000} \prec \mathbf{E}_{200000} \prec \mathbf{E}_{020000} \prec \mathbf{E}_{002000} \prec \mathbf{E}_{000200} \prec \mathbf{E}_{110000} \prec \mathbf{E}_{101000} \prec \mathbf{E}_{010100} \prec \mathbf{E}_{100100}$ (i.e. $vw \prec u^2 \prec v^2 \prec w^2 \prec x^2 \prec uv \prec uw \prec vx \prec ux$) has been chosen, and $a = b = c = d = e = 0$. Blank entries are *zero*.

$$(4.7)\qquad \begin{pmatrix} 1+2a & 1-c & 1 & & & & & & \\ 1-c & 1 & 1-\delta & -d & e & & & & \\ 1 & 1-\delta & 1 & -a & b & & & & \\ & -d & -a & 1 & -\gamma & & & & \\ & e & b & -\gamma & \gamma^2 & & & & \\ & & & & & 2\delta & c & & \\ & & & & & c & 2d & & \\ & & & & & & & -2b & \\ & & & & & & & & -2e \end{pmatrix}$$

The following notation for principal submatrices of (4.7) will be used. $[1\ 3]$ denotes the submatrix $\begin{pmatrix} 1+2a & 1 \\ 1 & 1 \end{pmatrix}$ formed from the 1st and 3rd rows and columns of (4.7), etc.

The parameters a, b, c, d, e correspond to the *changes* $2\mathbf{E}_{011000}^{\otimes 2} - \mathbf{E}_{020000} \otimes_s \mathbf{E}_{002000}$, $\mathbf{E}_{020000} \otimes_s \mathbf{E}_{000200} - 2\mathbf{E}_{010100}^{\otimes 2}$, $\mathbf{E}_{110000} \otimes_s \mathbf{E}_{101000} - \mathbf{E}_{200000} \otimes_s \mathbf{E}_{011000}$, $2\mathbf{E}_{101000}^{\otimes 2} - \mathbf{E}_{200000} \otimes_s \mathbf{E}_{002000}$, $\mathbf{E}_{200000} \otimes_s \mathbf{E}_{000200} - 2\mathbf{E}_{100100}^{\otimes 2}$ respectively.

No other *nonzero* entries may be altered: The four 1's in $[1\ 2\ 3]$ because there is no basis element of $A^{2,2,6}$ that is expressed using these positions. The three entries with δ because the only change possible has already been chosen. $[4\ 5]$ because the quartic form $f(0,0,w,x,y,z)$ has a unique Gram matrix by Theorem 3.4, and $[4\ 5]$ is a submatrix of that Gram matrix; if a $\mathcal{G}_0 + \Delta$ is a Gram matrix, then by deleting all rows and columns that involve the variables u and v one obtains a Gram matrix for $f(0,0,w,x,y,z)$.

When $a = c = 0$ it follows that $\det[1\ 2\ 3] = -\delta^2 < 0$. Since all principal minors of a psd matrix must be nonnegative, $a = c = 0$ cannot hold. It will first be shown that $a = 0$ is necessary and then that $c = \delta$ is necessary, leading to a contradiction that proves the theorem .

The determinant of $[1\ 3]$ forces $a \geq 0$. Introducing b, $\det[3\ 4\ 5] = -(b - a\gamma)^2$ whence $b = a\gamma \geq 0$. But submatrix $[8]$ implies $b \leq 0$ whence $a = 0$ also.

With $a = 0$ it follows that $\det[1\ 2\ 3] = -(\delta - c)^2$ whence $c = \delta$. Consequently $[6\ 7]$ requires $d > 0$. Now $\det[2\ 4\ 5] = -(e - d\gamma)^2$ whence $e > 0$ contradicting submatrix $[9]$. $\square$

5. A 6TH ORDER EXAMPLE

Consider the family of sextics

(5.1)
$$f_\rho(x,y,z) = x^2(\rho^2 x^2 + \rho y^2 - \frac{1}{2}z^2)^2 + y^2(\rho^2 y^2 + \rho z^2 - \frac{1}{2}x^2)^2 + z^2(\rho^2 z^2 + \rho x^2 - \frac{1}{2}y^2)^2$$

The three cubic polynomials that are squared have a common nontrivial root only when $\rho = 0, \rho^3 = -\frac{1}{2}, \rho^3 = -\frac{5+3\sqrt{3}}{4}$ or $\rho^3 = \frac{-5+3\sqrt{3}}{4}$. In each case the root can be taken in $\mathbb{R}^3$. Thus f_ρ is pd if and only if ρ^3 does not take the four listed values. In addition, every pd form f_ρ is coercive.

Put $\eta_0 = (1+\sqrt{5})^{-3}$. Then for the Choi-Lam sextics (1.12), $s_{\eta_0} = (1+\sqrt{5})f_\rho$ when $\rho = (1+\sqrt{5})^{-1}$. It will be shown that $(1+\sqrt{5})^{-1}$ belongs to an interval of ρ's for which the f_ρ have unique Gram matrices. This uniqueness implies, as in the quartic case, that η_0 is the smallest value of η for which s_η is an sos. The identity used in (3.2) may be replaced with $x^6+y^6=(x^3-2xy^2)^2+(y^3-2x^2y)^2$.

Hence, an apparent Gram matrix $\mathcal{G}_\rho$ for each f_ρ is

$$(\rho^2\mathbf{E}_{300}+\rho\mathbf{E}_{120}-\frac{1}{2}\mathbf{E}_{102})^{\otimes 2}+(\rho^2\mathbf{E}_{030}+\rho\mathbf{E}_{012}-\frac{1}{2}\mathbf{E}_{210})^{\otimes 2}+(\rho^2\mathbf{E}_{003}+\rho\mathbf{E}_{201}-\frac{1}{2}\mathbf{E}_{021})^{\otimes 2}$$

acting on the space $S^3(\mathbb{R}^3)$ which has 10 dimensions. Therefore using $\mathbf{E}_\alpha\cdot\mathbf{E}_\alpha=\frac{\alpha!}{6}$ the null space for $\mathcal{G}_\rho$ is spanned by the vectors $\mathbf{E}_{300}-3\rho\mathbf{E}_{120},\mathbf{E}_{120}+2\rho\mathbf{E}_{102},\mathbf{E}_{030}-3\rho\mathbf{E}_{012},\mathbf{E}_{012}+2\rho\mathbf{E}_{210},\mathbf{E}_{003}-3\rho\mathbf{E}_{201},\mathbf{E}_{201}+2\rho\mathbf{E}_{021}$ and $\mathbf{E}_{111}$. A general linear combination is

$$\begin{aligned}\mathbf{g}=a\mathbf{E}_{300}+3(b-\rho a)\mathbf{E}_{120}+6\rho b\mathbf{E}_{102}+c\mathbf{E}_{030}+3(d-\rho c)\mathbf{E}_{012}+6\rho d\mathbf{E}_{210}\\+e\mathbf{E}_{003}+3(f-\rho e)\mathbf{E}_{201}+6\rho f\mathbf{E}_{021}+6g\mathbf{E}_{111}\end{aligned}\tag{5.2}$$

The 27 dimensions of the subspace $A^{2,3,3}$ of *changes* may be briefly described as follows.

$$\frac{1}{2}\mathbf{E}_{300}\otimes_s\mathbf{E}_{120}-\frac{1}{2}\mathbf{E}_{210}\otimes_s\mathbf{E}_{210}$$

is representative of 6 changes.

$$\frac{1}{2}\mathbf{E}_{300}\otimes_s\mathbf{E}_{111}-\frac{1}{2}\mathbf{E}_{210}\otimes_s\mathbf{E}_{201}$$

is representative of 3.

$$\frac{1}{2}\mathbf{E}_{300}\otimes_s\mathbf{E}_{030}-\frac{1}{2}\mathbf{E}_{210}\otimes_s\mathbf{E}_{120}$$

representative of 3.

$$\frac{1}{2}\mathbf{E}_{300}\otimes_s\mathbf{E}_{021}-\frac{1}{2}\mathbf{E}_{201}\otimes_s\mathbf{E}_{120}$$

representative of 6.

$$\frac{1}{2}\mathbf{E}_{300}\otimes_s\mathbf{E}_{021}-\frac{1}{2}\mathbf{E}_{210}\otimes_s\mathbf{E}_{111}$$

representative of 6.

$$\frac{1}{2}\mathbf{E}_{210}\otimes_s\mathbf{E}_{012}-\frac{1}{2}\mathbf{E}_{111}\otimes_s\mathbf{E}_{111}$$

representative of 3. Keeping in mind the examples $\mathbf{E}_{300}\cdot\mathbf{E}_{300}=1,\mathbf{E}_{120}\cdot\mathbf{E}_{120}=1/3$ and $\mathbf{E}_{111}\cdot\mathbf{E}_{111}=1/6$, and computing $\Delta\cdot\mathbf{gg}$ for each change yields the quadratic polynomials

$-\rho a^2+ab-4\rho^2d^2$
$-\rho c^2+cd-4\rho^2f^2$
$-\rho e^2+ef-4\rho^2b^2$
$2\rho ab-\rho^2e^2-f^2+2\rho ef$
$2\rho cd-\rho^2a^2-b^2+2\rho ab$
$2\rho ef-\rho^2c^2-d^2+2\rho cd$

$ag-2\rho df+2\rho^2de$
$cg-2\rho bf+2\rho^2af$

$eg - 2\rho bd + 2\rho^2 bc$

$ac + 2\rho^2 ad - 2\rho bd$
$ce + 2\rho^2 cf - 2\rho df$
$ae + 2\rho^2 be - 2\rho bf$

$3\rho af - \rho^2 ae - bf + \rho be$
$3\rho bc - \rho^2 ac - bd + \rho ad$
$3\rho ed - \rho^2 ce - df + \rho cf$
$-\rho ac + ad - 4\rho^2 bd$
$-\rho ce + cf - 4\rho^2 df$
$-\rho ae + be - 4\rho^2 bf$

$af - dg$
$bc - fg$
$de - bg$
$-\rho ac + ad + \rho eg - fg$
$-\rho ce + cf + \rho ag - bg$
$-\rho ae + be + \rho cg - dg$

$-2\rho^2 cd + 2\rho d^2 - g^2$
$-2\rho^2 ef + 2\rho f^2 - g^2$
$-2\rho^2 ab + 2\rho b^2 - g^2$

Linearly order the 28 quadratic monomials $a^2, b^2, \ldots, g^2, ab, ac, \ldots, ag, bc, \ldots, eg, fg$ as before and put the resulting 27×28 coefficient matrix into reduced echelon form. When the 26th column (the ef column) is removed the result is the *identity* matrix. Putting $\sigma = \frac{1-16\rho^3}{\rho(1-4\rho^3)}, \tau = \frac{3\rho}{1-4\rho^3}$ and $\phi = \frac{4\rho^2(2\rho^3+1)}{1-4\rho^3}$, the 26th column has successive entries

$$-\sigma, -\tau, -\sigma, -\tau, -\sigma, -\tau, -\phi, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0$$

Thus an equivalent set of polynomials is

$$\begin{gathered} a^2 - \sigma ef \\ b^2 - \tau ef \\ c^2 - \sigma ef \\ d^2 - \tau ef \\ e^2 - \sigma ef \\ f^2 - \tau ef \\ g^2 - \phi ef \\ ab - ef \\ cd - ef \end{gathered} \tag{5.3}$$

together with the remaining 18 indefinite monomials none of which appear in the polynomials (5.3). By (3.7) a sufficient requirement for f_ρ to have a unique Gram matrix is that there exists no nontrivial linear combination of the polynomials (5.3) that is a definite or semi-definite quadratic polynomial in the variables $a, \ldots, g$. This requirement is equivalent to showing for a given ρ that every nontrivial choice of parameters $A, B, C, D, E, F, G, J, K$ in

(5.4)
$$\begin{pmatrix} 2A & J{-}A\sigma{-}B\tau &&&&& \\ J{-}A\sigma{-}B\tau & 2B &&&&& \\ && 2C & K{-}C\sigma{-}D\tau &&& \\ && K{-}C\sigma{-}D\tau & 2D &&& \\ &&&& 2E & -E\sigma{-}F\tau{-}G\phi{-}J{-}K & \\ &&&& -E\sigma{-}F\tau{-}G\phi{-}J{-}K & 2F & \\ &&&&&& 2G \end{pmatrix}$$

produces an indefinite matrix.

When σ, τ and ϕ are not all of the same sign there exist, by (5.3), choices of positive $A, \ldots, G$ that make (5.4) definite. Lack of a common sign holds for $-1/2 < \rho^3 < 0$ and $1/16 \le \rho^3$. When $\rho^3 < -1/2$ each of σ, τ and ϕ is negative while each is positive for $0 < \rho^3 < 1/16$.

Restricting to those nontrivial choices with $G = J = K = 0$, all produce indefinite matrices (5.4) if and only if $\sigma\tau > 1$. For example, the 2×2 minor $4EF - (E\sigma + F\tau)^2 < 0$ if and only if $\sigma\tau > 1$ if and only if $-\frac{5+3\sqrt{3}}{4} < \rho^3 < \frac{-5+3\sqrt{3}}{4}$. Therefore the remaining intervals for ρ^3 for which all nontrivial (5.4) are possibly not definite are the open intervals $(-\frac{5+3\sqrt{3}}{4}, \frac{-1}{2})$ and $(0, \frac{-5+3\sqrt{3}}{4})$. That ϕ shares the same sign with σ and τ in these intervals shows that choosing $G > 0$ does not restrict these intervals further. Neither can nonzero choices of J and K. The endpoints of the intervals yield f_ρ that are not pd.

The foregoing proves

Theorem 5.1. *The forms f_ρ are pd and have unique Gram matrices if and only if $-\frac{5+3\sqrt{3}}{4} < \rho^3 < -\frac{1}{2}$ or $0 < \rho^3 < \frac{-5+3\sqrt{3}}{4}$. All other pd f_ρ have Gram matrices of rank 10. Each f_ρ, for ρ^3 not equal to the endpoints of the above intervals, is coercive.*

$(\sqrt{5}+1)^{-3}$ is contained in the second interval.

To prove Theorem 1.6 we will be content with a single example. Take $\rho = -1$.

Theorem 5.2. *The positive definite sextic form of* $\mathbb{R}[w, x, y, z]$

$$g(w,x,y,z) := f_{-1}(\sqrt{w^2+x^2}, y, z) = (w^3 + wx^2 - wy^2 - \frac{1}{2}wz^2)^2 \tag{5.5}$$
$$+(xw^2+x^3-xy^2-\frac{1}{2}xz^2)^2+(y^3-yz^2-\frac{1}{2}yw^2-\frac{1}{2}yx^2)^2+(z^3-zw^2-zx^2-\frac{1}{2}zy^2)^2$$

is a noncoercive sum of squares.

Proof. Let $\mathcal{G}_0$ denote the apparent Gram matrix for g and let $\mathcal{F}_{-1}$ denote the unique Gram matrix for f_{-1}.

For $\mathbf{z} \in \mathbb{C}^4$ denote $z_1^2 + z_2^2 = \xi^2$. Then the precise relationship between common complex roots for sos representations of g and f_{-1} is $\mathcal{G}_0\mathbf{z}^{\otimes 3} = \mathbf{0}$ if and only if $\mathcal{F}_{-1}(\xi, z_3, z_4)^{\otimes 3} = \mathbf{0}$. Consequently by Theorem 5.1 and (2.18) $\xi = z_3 = z_4 = 0$ when $\mathbf{z}^{\otimes 3}$ is in the null space of $\mathcal{G}_0$. Thus $\mathbf{z} := (1, i, 0, 0)$ may be taken, up to scaling, as the only nontrivial common root in the sos representation (5.5) for g.

For g to be coercive there must exist a Δ such that $\mathcal{G}_0 + \Delta$ is a Gram matrix and $\Delta\mathbf{z}^{\otimes 3} \neq \mathbf{0}$ (2.18). Therefore, similarly to the quartic case, at least one of $\Delta_1 = \mathbf{E}_{1200}^{\otimes 2} - \frac{1}{2}\mathbf{E}_{2100} \otimes_s \mathbf{E}_{0300}$ or $\Delta_2 = \mathbf{E}_{2100}^{\otimes 2} - \frac{1}{2}\mathbf{E}_{1200} \otimes_s \mathbf{E}_{3000}$ (see Remark 4.2) must be included in Δ with a *positive* coefficient. However, if Δ' is obtained from Δ by permuting the 1st and 2nd components of each multi-index of the basis elements (2.4), then $\mathcal{G}_0 + \Delta'$ would also be a Gram matrix because of the symmetry in w and x of (5.5). Further, because of positive semi-definiteness, $\mathbf{z}^{\otimes 3}$ is not in the null space of $\mathcal{G}_0 + \frac{1}{2}\Delta + \frac{1}{2}\Delta'$ when it is not in the null space of $\mathcal{G}_0 + \Delta$. Consequently, for g to be coercive, values for the parameters a, b, c, d with $a + b + c + d = 0$ in

(5.6)
$$\begin{pmatrix}
1+2\delta & 1-\delta & -1+a & & & & & & & \\
1-\delta & 1 & -1 & & & & & & & \\
-1+a & -1 & 1 & & & & & & & \\
 & & & 1+2\delta & 1-\delta & -1+b & & & & \\
 & & & 1-\delta & 1 & -1 & & & & \\
 & & & -1+b & -1 & 1 & & & & \\
 & & & & & & 1 & -\frac{1}{2} & -\frac{1}{2} & \\
 & & & & & & -\frac{1}{2} & \frac{1}{4} & \frac{1}{4}+c & \\
 & & & & & & -\frac{1}{2} & \frac{1}{4}+c & \frac{1}{4} & \\
 & & & & & & & & & d
\end{pmatrix}$$

must be found that make (5.6) psd when $\delta > 0$. Here (5.6) is the principal submatrix of $\mathcal{G}_0 + 2\delta\Delta_1 + 2\delta\Delta_2$ corresponding to $\mathbf{E}_{1200} \prec \mathbf{E}_{3000} \prec \mathbf{E}_{1020} \prec \mathbf{E}_{2100} \prec \mathbf{E}_{0300} \prec \mathbf{E}_{0120} \prec \mathbf{E}_{0030} \prec \mathbf{E}_{2010} \prec \mathbf{E}_{0210} \prec \mathbf{E}_{1110}$ (i.e. $wx^2 \prec w^3 \prec wy^2 \prec w^2x \prec x^3 \prec xy^2 \prec y^3 \prec w^2y \prec x^2y \prec wxy$). The parameters represent the three *changes* $\mathbf{E}_{2100} \otimes_s \mathbf{E}_{0120} - \mathbf{E}_{1110} \otimes_s \mathbf{E}_{1110}, \mathbf{E}_{2010} \otimes_s \mathbf{E}_{0210} - \mathbf{E}_{1110} \otimes_s \mathbf{E}_{1110},$ and $\mathbf{E}_{1200} \otimes_s \mathbf{E}_{1020} - \mathbf{E}_{1110} \otimes_s \mathbf{E}_{1110}$.

With the same notation as in the quartic case, the submatrix $[2\ 3]$ of (5.6) is *fixed* because it is a submatrix of the unique Gram matrix $\mathcal{F}_{-1}$ for $g(w, 0, y, z) = f_{-1}(w, y, z)$. So is $[5\ 6]$ because $g(0, x, y, z) = f_{-1}(x, y, z)$. In the same way $[7\ 8]$ and $[7\ 9]$ are fixed. With no other choices and $\det[1\ 2\ 3] = -(a - \delta)^2$ it follows that $a = \delta$ is forced. In the same way $b = \delta$ and $c = 0$. Thus $d = -2\delta$, a contradiction, and g cannot be coercive.

□

Remark 5.3. By Nullstellensätze (see pp. 56-57 of [Pfi95]) every collection of homogeneous polynomials $p_1, \ldots, p_r \in \mathbb{C}[x_1, \ldots, x_n]$ with $1 \leq r < n$ has a common nontrivial zero $\mathbf{a} \in \mathbb{C}^n$ to the system of equations $p_1 = \cdots = p_r = 0$ while the corresponding statement, for the polynomial ring $\mathbb{R}[x_1, \ldots, x_n]$ and $\mathbb{R}^n$ in place of $\mathbb{C}^n$, holds only when all of the degrees $d_1, \ldots, d_r$ of the polynomials $p_1, \ldots, p_r$ are not *even*. Thus the sextic example here is required to be the sum of at least 4 squares in order to be pd while the quartic examples are pd with but 5 squares of quadratics in the 6 indeterminates. The 5 quadratics necessarily share a nontrivial complex root while the 4 cubics need not, though they do.

6. The game

Starting with the collection of $pd\ sos$ in $P_{3,4}$ one can obtain the coercive result (1.10) without using Hilbert's theorem on ternary quartics by considering several generic cases. One shows that the ranks of the Gram matrices arising in each case can be built up by adding changes Δ as delineated in Proposition 3.5. When attempting to show that the pd elements of $\Sigma_{4,4}$ are coercive the number of cases is significantly higher.

The vector space $S^2(\mathbb{R}^n)$ is isomorphic to the space of real symmetric $n \times n$ matrices by assigning $\mathbf{t} \in S^2(\mathbb{R}^n)$ to the matrix with the coordinates $\mathbf{t} \cdot \mathbf{e}^i\mathbf{e}^j$ as entries $1 \le i, j \le n$. Every change $\Delta \in A^{2,2,n}$ yields a quadratic form $\Delta \cdot \mathbf{tt}$ that is a linear combination of the 2×2 minors of the symmetric matrix $\mathbf{t}$. The argument of Section 3 that can show that a pd quartic with Gram matrix $\mathcal{G}$ has $\mathcal{G}$ as its unique Gram matrix amounts to showing that a general matrix in $Null(\mathcal{G}) \subset S^2(\mathbb{R}^n)$ has the property that every nontrivial linear combination of its 2×2 minors is indefinite. For example, the pd quartic

$$f = (x_1^2 + x_2^2 - x_3^2 - x_4^2)^2 + (2x_2x_3 - x_3^2 + x_4^2)^2 + (x_1x_3 - x_2x_4)^2 + (x_1 - x_3)^2x_4^2$$

has the basis $2\mathbf{E}_{1100}, \mathbf{E}_{2000} - \mathbf{E}_{0200}, \mathbf{E}_{2000} + \mathbf{E}_{0200} + \mathbf{E}_{0020} + \mathbf{E}_{0002}, \mathbf{E}_{0020} - \mathbf{E}_{0002} + 2\mathbf{E}_{0110}, 2\mathbf{E}_{1010} + 2\mathbf{E}_{0101}, 2\mathbf{E}_{1001} + 2\mathbf{E}_{0011}$ for the null space of its apparent Gram matrix. The first two basis elements are the imaginary and real parts of $\mathbf{z} \otimes \mathbf{z}$ for $\mathbf{z} = (1, i, 0, 0)$ the common complex root for the sos f. A general linear combination of the basis elements corresponds to the 4×4 matrix

$$\mathbf{t} = \begin{pmatrix} b+c & a & e & f \\ a & -b+c & d & e \\ e & d & c+d & f \\ f & e & f & c-d \end{pmatrix} \tag{6.1}$$

Here, however, there is a nontrivial linear combination of the 2×2 minors that is not indefinite. Otherwise f would provide a noncoercive example for $n = 4$. To prove that f is coercive it is necessary to produce a linear combination of minors of the form

$$a^2 + b^2 - c^2 + \Delta \cdot \mathbf{tt} \tag{6.2}$$

that is psd and where the last term does not include the principle 2×2 minor $\det[1\ 2]$. This is the same observation as (4.6). It might not be clear that the last term can be made up of minors that yield a positive coefficient on the monomial c^2 without introducing more indefiniteness. However, it can be done. To express c^2 itself as a linear combination of the remaining 19 independent 2×2 minors it is necessary to use 18 of them. In fact, (6.2) can be made pd and thus f possesses a Gram matrix of full rank by Proposition 3.5.

By a linear change of variables in $\mathbb{R}^n$ any nontrivial common complex root for a pd quartic sos may be taken to be $(1, i, 0, 0, \ldots)$. Therefore the precise setup of the principal submatrix $[1\ 2]$ of (6.1), together with the presence in some way of the variable c outside $[1\ 2]$, is a typical setup for the null spaces of Gram matrices when trying to answer question (1.1) in the quartic cases. When c does not occur outside $[1\ 2]$ real values may be assigned to the variables making $[1\ 2]$ and $\mathbf{t}$ rank-1 matrices, contradicting the positive definiteness of the form f. When only a and b occur in $[1\ 2]$ f can be written as a sos that includes the term $(x_1^2 + x_2^2)^2 = x_1^4 + 2x_1^2x_2^2 + x_2^4$ in the sum.

These observations lead to the following diversion.

1. Set up the principal submatrix $[1\ 2]$ of an $n \times n$ symmetric matrix $\mathbf{t}$ exactly as in (6.1).

2. Write linear combinations of c and a number of other real variables for the remaining entries. Variable c must be used while a and b may not.

3. The choices made in Step 2 are not allowed to result in a rank-1 matrix for any choice of real variable values. This can usually be checked by inspecting for zeros an sos quartic form, i.e. Gram matrix, which will have the $n \times n$ matrix as its null space.

4. Search for a linear combination of 2×2 minors (not including $\det[1\ 2]$) which when added to $a^2 + b^2 - c^2$ results in a psd quadratic form.

When $n = 4$ or 5 there are two or three ways to win this game. Find a setup for which the goal of Step 4 cannot be achieved. Or, when Step 4 does result in a psd quadratic but never a pd quadratic, show that the resulting change Δ always satisfies $\Delta \mathbf{t} \neq \mathbf{0}$ for some choice of real variable values. See Proposition 3.5. Or, prove that neither of these outcomes is ever possible for any $\mathbf{t}$ constructed according to Steps 1, 2 and 3, thus proving that every $pd\ sos$ is coercive.

7. Final remark on coercive integro-differential forms

The results of this article when combined with the Aronszajn-Smith Theorem show that there exist homogeneous constant coefficient elliptic operators L with formally positive integro-differential forms (1.2) for which a coercive estimate like (1.3) is never true. However, such an L could have an integro-differential form like (1.4) which is not formally positive but which satisfies the coercive estimate (1.3) when (1.4) is used on the left side in place of (1.2). The author claims this to be always true in the quartic, i.e. 4th order operator, cases. The proof necessarily uses Agmon's characterization of coerciveness and will appear elsewhere. Thus Agmon's characterization is needed in order to answer the coerciveness problem for differential operators even when those operators possess formally positive integro-differential forms.

References

[Agm58] Shmuel Agmon, *The coerciveness problem for integro-differential forms*, J. Analyse Math. **6** (1958), 183–223. MR 24 #A2748

[Agm60] ———, *Remarks on self-adjoint and semi-bounded elliptic boundary value problems*, Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960), Jerusalem Academic Press, Jerusalem, 1960, pp. 1–13. MR 24 #A3417

[Agm65] ———, *Lectures on elliptic boundary value problems*, Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Van Nostrand Mathematical Studies, No. 2, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965. MR MR0178246 (31 #2504)

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36, Springer-Verlag, Berlin, 1998, Translated from the 1987 French original, Revised by the authors. MR MR1659509 (2000a:14067)

[CL77] Man Duen Choi and Tsit Yuen Lam, *An old question of Hilbert*, Conference on Quadratic Forms—1976 (Proc. Conf., Queen's Univ., Kingston, Ont., 1976), Queen's Univ., Kingston, Ont., 1977, pp. 385–405. Queen's Papers in Pure and Appl. Math., No. 46. MR 58 #16503

[CL78] ———, *Extremal positive semidefinite forms*, Math. Ann. **231** (1977/78), no. 1, 1–18. MR MR0498384 (58 #16512)

[CLR95] M. D. Choi, T. Y. Lam, and B. Reznick, *Sums of squares of real polynomials*, K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992), Proc. Sympos. Pure Math., vol. 58, Amer. Math. Soc., Providence, RI, 1995, pp. 103–126. MR MR1327293 (96f:11058)

[Dia62] P. H. Diananda, *On non-negative forms in real variables some or all of which are non-negative*, Proc. Cambridge Philos. Soc. **58** (1962), 17–25. MR MR0137686 (25 #1136)

[Gel89] Bernard Gelbaum, *Linear algebra*, North-Holland Publishing Co., New York, 1989, Basics, practice, and theory. MR MR1087557 (92c:15001)

[Har99] William R. Harris, *Real even symmetric ternary forms*, J. Algebra **222** (1999), no. 1, 204–245. MR MR1728161 (2001a:11070)

[Hil88] David Hilbert, *Uber die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), 342–350.

[HN63] Marshall Hall, Jr. and Morris Newman, *Copositive and completely positive quadratic forms*, Proc. Cambridge Philos. Soc. **59** (1963), 329–339. MR MR0147484 (26 #5000)

[LL78] Anneli Lax and Peter D. Lax, *On sums of squares*, Linear Algebra and Appl. **20** (1978), no. 1, 71–75. MR MR0463112 (57 #3074)

[Mot67] T. S. Motzkin, *The arithmetic-geometric inequality*, Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965), Academic Press, New York, 1967, pp. 205–224. MR MR0223521 (36 #6569)

[PD01] Alexander Prestel and Charles N. Delzell, *Positive polynomials*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001, From Hilbert's 17th problem to real algebra. MR MR1829790 (2002k:13044)

[Pfi95] Albrecht Pfister, *Quadratic forms with applications to algebraic geometry and topology*, London Mathematical Society Lecture Note Series, vol. 217, Cambridge University Press, Cambridge, 1995. MR MR1366652 (97c:11046)

[Pfi04] ______, *On Hilbert's theorem about ternary quartics*, Algebraic and arithmetic theory of quadratic forms, Contemp. Math., vol. 344, Amer. Math. Soc., Providence, RI, 2004, pp. 295–301. MR MR2060205 (2005f:11058)

[PR00] Victoria Powers and Bruce Reznick, *Notes towards a constructive proof of Hilbert's theorem on ternary quartics*, Quadratic forms and their applications (Dublin, 1999), Contemp. Math., vol. 272, Amer. Math. Soc., Providence, RI, 2000, pp. 209–227. MR MR1803369 (2001h:11049)

[Pra06] S. Prajna, *Theory and algorithms of linear matrix inequalities*, The American Institute of Mathematics, http://www.aimath.org, March 12, 2006, Questions and Discussions of the Literature.

[PW98] Victoria Powers and Thorsten Wörmann, *An algorithm for sums of squares of real polynomials*, J. Pure Appl. Algebra **127** (1998), no. 1, 99–104. MR MR1609496 (99a:11047)

[Raj93] A. R. Rajwade, *Squares*, London Mathematical Society Lecture Note Series, vol. 171, Cambridge University Press, Cambridge, 1993. MR MR1253071 (94m:11047)

[Rez00] Bruce Reznick, *Some concrete aspects of Hilbert's 17th Problem*, Real algebraic geometry and ordered structures (Baton Rouge, LA, 1996), Contemp. Math., vol. 253, Amer. Math. Soc., Providence, RI, 2000, pp. 251–272. MR MR1747589 (2001i:11042)

[Rez07] ______, *On Hilbert's construction of positive polynomials*, preprint (2007), 28 pages.

[Rob73] Raphael M. Robinson, *Some definite polynomials which are not sums of squares of real polynomials*, Selected questions of algebra and logic (collection dedicated to the memory of A. I. Mal′cev) (Russian), Izdat. "Nauka" Sibirsk. Otdel., Novosibirsk, 1973, pp. 264–282. MR 49 #2647

[Rud00] Walter Rudin, *Sums of squares of polynomials*, Amer. Math. Monthly **107** (2000), no. 9, 813–821. MR MR1792413 (2002c:12003)

[Swa00] Richard G. Swan, *Hilbert's theorem on positive ternary quartics*, Quadratic forms and their applications (Dublin, 1999), Contemp. Math., vol. 272, Amer. Math. Soc., Providence, RI, 2000, pp. 287–292. MR MR1803372 (2001k:11065)

[vdW70] B. L. van der Waerden, *Algebra. Vol 1*, Translated by Fred Blum and John R. Schulenberger, Frederick Ungar Publishing Co., New York, 1970. MR MR0263582 (41 #8187a)

[Yok92] Takeo Yokonuma, *Tensor spaces and exterior algebra*, Translations of Mathematical Monographs, vol. 108, American Mathematical Society, Providence, RI, 1992, Translated from the 1977 Japanese edition by the author. MR MR1187759 (93j:15020)

215 CARNEGIE, SYRACUSE UNIVERSITY, SYRACUSE NY 13244
E-mail address: gverchot@syr.edu