Math. Ann. 283, 431–464 (1989)

Mathematische Annalen

Forms Derived from the Arithmetic-Geometric Inequality

Bruce Reznick*

Department of Mathematics, University of Illinois, 1409 West Green Street, Urbana, IL 61801, USA

1. Introduction and Overview

A real homogeneous polynomial (a *form*) p in n variables is *positive semidefinite* or (*psd*) if $p(\underset{\sim}{x}) \geqq 0$ for all $\underset{\sim}{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. It is *sos* if it is a *sum of squares* of forms: $p(\underset{\sim}{x}) = \sum h_k^2(\underset{\sim}{x})$ for suitable h_k. Every sos form is psd. In 1888, Hilbert [13] proved that not every psd form is sos, but no explicit example was given for nearly eighty years. The set of psd forms $p(x_1, \ldots, x_n)$ of fixed degree comprises a closed convex cone. A psd form p is called *extremal* if it is extremal as an element of this cone: p is extremal if $p = h_1 + h_2$, where h_i is a psd form, implies that $h_i = \alpha_i p$ for some $\alpha_i \geqq 0$. Every psd form is a sum of finitely many extremal forms.

In general, it is difficult to determine whether a particular psd form is sos or extremal. Many examples from the literature arise from monomial substitution into the arithmetic-geometric inequality (AGI); we shall call these agiforms. In this paper, we determine a necessary condition for an agiform to be sos. If the monomials are algebraically independent, this condition is sufficient, and we obtain an explicit representation of the agiform as a sum of squares of binomials. We also determine a necessary and sufficient condition for an agiform to be extremal. These expand the pools of known extremal forms and of psd forms which are not sos.

The relevant conditions are always geometric. Associated to each agiform is a polytope with lattice point vertices and a distinguished interior lattice point: the convex hull of the set of exponents used in the substitution, and the exponent of the resulting weighted geometric mean. We shall study the set of lattice points contained in this polytope. For sums of squares, the condition involves writing lattice points as averages of even lattice points. For extremality, the condition involves the parity (mod 2) of the lattice points contained in the polytope.

Hilbert did not carry out his construction in detail, and the first explicit example of a psd form which is not sos was found by Motzkin in 1967. The AGI

* Author supported in part by the Alfred P. Sloan Foundation and the National Science Foundation

(see [12, p. 17]) states:

$$\lambda_1 t_1 + \ldots + \lambda_m t_m - t_1^{\lambda_1} \ldots t_m^{\lambda_m} \geqq 0 \tag{1.1}$$

if $t_i \geqq 0$, $\lambda_i \geqq 0$ and $\sum \lambda_i = 1$. Equality holds in (1.1) if and only if, for some $c \geqq 0$, $\lambda_i > 0$ implies $t_i = c$. Motzkin [16, p. 217] presented the form

$$M(x,y,z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2, \tag{1.2}$$

which is psd by (1.1), under the substitution $t_1 = x^4y^2$, $t_2 = x^2y^4$, $t_3 = z^6$, and $\lambda_i = \frac{1}{3}$ (and multiplication by 3); he showed that M is not sos. The polytope associated to M is the triangle with vertices $(4,2,0)$, $(2,4,0)$ and $(0,0,6)$ and the distinguished point is $(2,2,2)$.

Choi, Lam, and the author [3, 4, 17] have derived other psd forms which are not sos from monomial substitutions into the AGI. Two such examples are given in [3, p. 388]:

$$S(x,y,z) = x^4y^2 + y^4z^2 + z^4x^2 - 3x^2y^2z^2, \tag{1.3}$$

$$Q(x,y,z,w) = w^4 + x^2y^2 + x^2z^2 + y^2z^2 - 4xyzw. \tag{1.4}$$

(Not all examples of psd forms which are not sos come from the AGI. Others have been found by Robinson [21], Lax and Lax [15], Schmüdgen [22], and Choi, Lam, and the author [4–7].)

Hurwitz [14] proved in 1891 that

$$G_{2d}(y_1, \ldots, y_{2d}) := y_1^{2d} + \ldots + y_{2d}^{2d} - 2d y_1 \ldots y_{2d} = \sum g_i^2(y_1, \ldots, y_{2d}) \tag{1.5}$$

for appropriate forms g_i. Hurwitz explicitly alluded to Hilbert's recent work as indicating the non-triviality of this representation. The form G_{2d} arises from the substitution $m = 2d$, $\lambda_i = \dfrac{1}{2d}$, $t_i = y_i^{2d}$ into (1.1). In fact, Hurwitz used the sum of squares representation (1.5) to prove the AGI. Let c_i be non-negative integers summing to $2d$. Upon setting c_i of the variables equal to x_i for $i = 1, \ldots, n$, (1.5) becomes:

$$G(\underset{\sim}{c})(\underset{\sim}{x}) := c_1 x_1^{2d} + \ldots + c_n x_n^{2d} - 2d x_1^{c_1} \ldots x_n^{c_n} = \sum g_i^2(x_1, \ldots, x_1, x_2, \ldots, x_n).$$

Thus, $G(\underset{\sim}{c})$ is sos, and (1.1) is valid when $\lambda_i = \dfrac{1}{2d} c_i$; that is, for all rational $\underset{\sim}{\lambda}$ with $\lambda_i \geqq 0$ and $\sum \lambda_i = 1$. By continuity, (1.1) holds for real $\underset{\sim}{\lambda}$. The polytope associated to $G(\underset{\sim}{c})$ is the simplex with vertices $2d\underset{\sim}{e}_i$, where $\underset{\sim}{e}_i$ is the i-th unit vector, and the distinguished point is $\underset{\sim}{c}$. We shall repeatedly contrast the Motzkin form M and the Hurwitz form

$$H(x,y,z) = x^6 + y^6 + z^6 - 3x^2y^2z^2 \tag{1.6}$$

as prototypes of the not-sos and sos agiforms; $2H = G(\underset{\sim}{c})$ for $\underset{\sim}{c} = (2,2,2)$.

Following Choi and Lam, [4, p. 1], we let $P_{n,m}$ (resp. $\Sigma_{n,m}$) denote the convex cone of psd (resp. sos) forms in n variables with even degree m and let $\Delta_{n,m} = P_{n,m} \setminus \Sigma_{n,m}$. Hilbert proved that $\Delta_{n,m} = \phi$ if and only if $m = 2$ or $n = 2$ or $(n,m) = (3,4)$. By identifying an n-ary m-ic form with the M-tuple of its coefficients,

$M=\binom{m+n-1}{m}$, $P_{n,m}$ can be viewed as a cone lying in $\mathbb{R}^M$. From elementary convexity theory, it follows that every psd form is a sum of M extremal forms. [If p is extremal and sos, then clearly p is a perfect square, but not every perfect square is extremal: $(x^2+y^2)^2=(x^2-y^2)^2+(2xy)^2$.] There are few constructions of extremal psd forms in the literature; one general result is found in [2, p. 287]. A product of distinct irreducible indefinite factors is called "purely indefinite". If h is purely indefinite, then h^2 is extremal; if h is purely indefinite and p is extremal, then ph^2 is also extremal. The product of two extremal forms need not be extremal (see [3, p. 402].)

Choi and Lam proved that the agiforms M, S, and Q are extremal, as well as not sos [4, pp. 8–9]. The results inspired [17], in which the author derived the set of extremal psd forms with four or fewer terms. Such a form either is a monomial square, $c(\underline{x}^{\underline{v}})^2$ $(c>0)$ or, after a dilation, arises from a special class of monomial substitutions into (1.1) with $\lambda_i=\frac{1}{3}$.

We introduce some notation. An n-tuple $\underline{u}=(u_1,\ldots,u_n)$ is a *lattice point* if $\underline{u}\in\mathbb{Z}^n$; $\underline{u}$ is an *even* lattice point if $u_j\in 2\mathbb{Z}$, or $\underline{u}=2\underline{v}$, where $\underline{v}$ is a lattice point. For a lattice point $\underline{u}$ with $u_j\geqq 0$, and $\underline{x}\in\mathbb{R}^n$, $\underline{x}^{\underline{u}}$ is the monomial $x_1^{u_1}\ldots x_n^{u_n}$. (When n is small, we name the variables $x, y, z, w, v, u, \ldots$). If $\underline{u}$ is even then $\underline{x}^{\underline{u}}=(\underline{x}^{\underline{v}})^2\geqq 0$ for all $\underline{x}\in\mathbb{R}^n$. We use the term *framework* (and a capital gothic letter) to denote a set $\mathfrak{U}=\{\underline{u}_1,\ldots,\underline{u}_m\}$ of even lattice points in $\mathbb{R}^n$ for which $u_{ij}\geqq 0$ and $\sum_{j=1}^{n} u_{ij}=2d$ for all i and some d. (This last condition ensures that each monomial $\underline{x}^{\underline{u}_i}$ has degree $2d$.) A *trellis* is a framework in which $\underline{u}_1,\ldots,\underline{u}_m$ comprise the vertices of a simplex. (The name is suggested by the horticultural trellis).

We collect some conditions under which a framework is a trellis. Suppose $\sum_{i=1}^{m} c_i\underline{u}_i=\underline{0}$. Then, by summing the coordinates on both sides, $0=\sum_{j=1}^{n}\left(\sum_{i=1}^{m} c_iu_{ij}\right)=2d\left(\sum_{i=1}^{m} c_i\right)$, hence $\sum_{i=1}^{m} c_i=0$ and so $\sum_{i=2}^{m} c_i(\underline{u}_i-\underline{u}_1)=\underline{0}$. Since a polytope is a simplex if and only if $\{\underline{u}_i-\underline{u}_1: i\geqq 2\}$, the edges emanating from $\underline{u}_1$, are linearly independent, it follows that $\mathfrak{U}$ is a trellis if and only if $\mathfrak{U}$ is a linearly independent set in $\mathbb{R}^n$; $m\leqq n$ in a trellis. The monomials $\underline{x}^{\underline{u}_i}$ are algebraically independent when $\prod_{i=1}^{m}(\underline{x}^{\underline{u}_i})^{t_i}=1$ implies $t_i=0$ for all i, hence $\mathfrak{U}$ is a trellis if and only if the $\underline{x}^{\underline{u}_i}$'s are algebraically independent. Finally, if $m=n$, then $\mathfrak{U}$ is a trellis if and only if $\det[u_{ij}]\neq 0$.

Suppose $\mathfrak{U}$ is a framework. We let $C(\mathfrak{U})=cvx(\mathfrak{U})\cap\mathbb{Z}^n$ and $E(\mathfrak{U})=cvx(\mathfrak{U})\cap(2\mathbb{Z})^n$ denote the lattice points (and the even lattice points) contained in the convex hull of $\mathfrak{U}$. We are interested in two sets of averages of sets of lattice points. If $\mathfrak{B}\subset\mathbb{Z}^n$, let

$$A(\mathfrak{B})=\{\tfrac{1}{2}(\underline{s}+\underline{t}):\ \underline{s},\ \underline{t}\in(\mathfrak{B}\cap(2\mathbb{Z})^n)\}$$

denote the set of averages of even points from $\mathfrak{B}$ and

$$\bar{A}(\mathfrak{B})=\{\tfrac{1}{2}(\underline{s}+\underline{t}):\ \underline{s}\neq\underline{t},\ \underline{s},\underline{t}\in(\mathfrak{B}\cap(2\mathbb{Z})^n)\}$$

denote the set of averages of distinct even points from $\mathfrak{B}$, so $A(\mathfrak{B})=\bar{A}(\mathfrak{B})\cup(\mathfrak{B}\cap(2\mathbb{Z})^n)$. Observe that $C(\mathfrak{U})$ contains $A(E(\mathfrak{U}))$; we show in [20] that $n\leqq 3$ implies $A(E(\mathfrak{U}))=C(\mathfrak{U})$, but this is false for $n\geqq 4$ (see Theorem 6.18).

If $\underline{w}\in\mathbb{Z}^n$ and $\underline{w}=\sum\lambda_i\underline{u}_i$ with $\lambda_i\geqq 0$ and $\sum\lambda_i=1$, then $\underline{w}\in C(\mathfrak{U})$; conversely, if $\underline{w}\in C(\mathfrak{U})$ then at least one such $\underline{\lambda}$ exists. Let

$$\Lambda(\underline{w})=\{\underline{\lambda}:\lambda_i\geqq 0,\ \textstyle\sum\lambda_i=1 \text{ and } \underline{w}=\sum\lambda_i\underline{u}_i\}.$$

If $\mathfrak{U}$ is a trellis, then the linear independence of the $\underline{u}_i$'s implies that $\Lambda(\underline{w})=\{\underline{\lambda}\}$ is a singleton, and $(\lambda_1,\ldots,\lambda_m)$ are called the *barycentric coordinates* of $\underline{w}$ with respect to $\mathfrak{U}$. If $\lambda_i>0$ for all i, then $\underline{w}$ is *interior* to $\mathfrak{U}$.

Fix a framework $\mathfrak{U}$, and select $\underline{w}\in C(\mathfrak{U})$ and $\underline{\lambda}\in\Lambda(\underline{w})$. Under the substitution $t_i=\underline{x}^{\underline{u}_i}(\geqq 0)$ the AGI becomes

$$\lambda_1\underline{x}^{\underline{u}_1}+\ldots+\lambda_m\underline{x}^{\underline{u}_m}-(\underline{x}^{\underline{u}_1})^{\lambda_1}\ldots(\underline{x}^{\underline{u}_m})^{\lambda_m}\geqq 0. \tag{1.7}$$

Since the $\underline{u}_i$'s are even and $w_j=\sum\lambda_i\underline{u}_{ij}$,

$$(\underline{x}^{\underline{u}_1})^{\lambda_1}\ldots(\underline{x}^{\underline{u}_m})^{\lambda_m}=\prod_{i=1}^{m}\left(\prod_{j=1}^{n}|x_j|^{u_{ij}}\right)^{\lambda_i}=\prod_{j=1}^{n}|x_j|^{w_j}=|\underline{x}^{\underline{w}}|,$$

and since $\sum_{j=1}^{n}w_j=\sum_{j=1}^{n}\sum_{i=1}^{m}\lambda_i u_{ij}=2d$, it follows from (1.7) that

$$f(\mathfrak{U},\underline{\lambda},\underline{w})(\underline{x}):=\lambda_1\underline{x}^{\underline{u}_1}+\ldots+\lambda_m\underline{x}^{\underline{u}_m}-\underline{x}^{\underline{w}} \tag{1.8}$$

is a psd form. Any positive multiple of $f(\mathfrak{U},\underline{\lambda},\underline{w})$ is called an *agiform on* $\mathfrak{U}$. (Multiples are usually taken to clear the denominators of the coefficients.) If $f(\mathfrak{U},\underline{\lambda},\underline{w})(\underline{x})=0$ and $\lambda_i>0$ for $i\in I\subseteq\{1,\ldots,m\}$, then there exists $c\geqq 0$ so that $\underline{x}^{\underline{u}_i}=c$ when $i\in I$ and $\underline{x}^{\underline{w}}\geqq 0$; in particular, $f(\mathfrak{U},\underline{\lambda},\underline{w})(1,\ldots,1)=0$.

If $\mathfrak{U}$ is a trellis, f is called a *simplical agiform on* $\mathfrak{U}$. In this case, $\underline{\lambda}$ is redundant, so it is convenient to write $f=f(\mathfrak{U},\underline{w})$. The simplicial agiforms on a fixed trellis are indexed by the elements of $C(\mathfrak{U})$. If $\underline{w}\in A(\mathfrak{U})$, then either $\underline{w}=\underline{u}_i$ or $\underline{w}=\frac{1}{2}(\underline{u}_i+\underline{u}_j)$ and the agiform $f(\mathfrak{U},\underline{w})$ is simple. In the first case $\underline{\lambda}=\underline{e}_i$ and $f(\mathfrak{U},\underline{w})(\underline{x})=\underline{x}^{\underline{u}_i}-\underline{x}^{\underline{u}_i}=0$. In the second case, $\underline{\lambda}=\frac{1}{2}(\underline{e}_i+\underline{e}_j)$ and $f(\mathfrak{U},\underline{w})(\underline{x})=\frac{1}{2}\underline{x}^{\underline{u}_i}+\frac{1}{2}\underline{x}^{\underline{u}_j}-\underline{x}^{\underline{w}}=\frac{1}{2}(\underline{x}^{\underline{u}_i/2}-\underline{x}^{\underline{u}_j/2})^2$ is a binomial square. It turns out that every agiform is a convex combination of simplical agiforms (see Theorem 7.1).

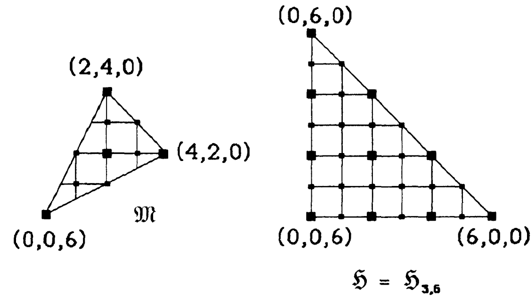
If f is an agiform on a framework $\mathfrak{U}$ and $\mathfrak{U}$ is a subset of a framework $\mathfrak{B}$, then, by taking the additional monomials with coefficient 0, f is an agiform on $\mathfrak{B}$. This creates a possible ambiguity of notation if $\mathfrak{U}$ is a trellis and $\mathfrak{B}$ is not. Accordingly, we say that the form f is a *simplicial agiform* if there exists a trellis $\mathfrak{U}$ so that f is a simplicial agiform on $\mathfrak{U}$. We may always choose $\mathfrak{U}$ so that $\underline{w}$ is an interior point; if $\lambda_j=0$, then $\underline{x}^{\underline{u}_j}$ does not occur in f, and $\underline{u}_j$ may be deleted from $\mathfrak{U}$.

We return in detail to the prototypical agiforms, M and H. Each trellis lies in the plane $t_1+t_2+t_3=6$ and no information is lost in Fig. 1 by projecting onto the first two coordinates. The elements of each trellis are labeled, the even lattice points are large squares, and the other lattice points are smaller squares.

(1.9) *Example.* We define the Motzkin trellis:

$$\mathfrak{M}=\{(4,2,0),\ (2,4,0),\ (0,0,6)\}.$$

Fig. 1



[Since the points in $\mathfrak{M}$ are linearly independent, $\mathfrak{M}$ is a trellis; alternately, $cvx(\mathfrak{M})$ is a simplex, viz. a triangle.] Let $\underline{w}=(2,2,2)$ so $\Lambda(\underline{w})=\{(\frac{1}{3},\frac{1}{3},\frac{1}{3})\}$; $M=3f(\mathfrak{M},\underline{w})$ [see (1.2)] is not sos. It is easy to see from Fig. 1 that $E(\mathfrak{M})=\mathfrak{M}\cup\{\underline{w}\}$ and, since each of the ten lattice points in $cvx(\mathfrak{M})$ is an average of two even points, $C(\mathfrak{M})=A(E(\mathfrak{M}))$. Considering the distinct averages, we see that $\bar{A}(E(\mathfrak{M}))=C(\mathfrak{M})\backslash E(\mathfrak{M})$. It is particularly significant that $\underline{w}\notin\bar{A}(E(\mathfrak{M}))$; this will imply that M is not sos. By Corollary 3.4, $f(\mathfrak{M},\underline{v})$ is not sos if $\underline{v}$ is any of the four points in $C(\mathfrak{M})\backslash\bar{A}(E(\mathfrak{M}))$.

(1.10) *Example.* We define a special case of the Hurwitz trellis (see Example 1.12):

$$\mathfrak{H}=\{(6,0,0),\ (0,6,0),\ (0,0,6)\}.$$

Again, if $\underline{w}=(2,2,2)$, then $\Lambda(\underline{w})=\{(\frac{1}{3},\frac{1}{3},\frac{1}{3})\}$; and $H=3f(\mathfrak{H},\underline{w})$, as in (1.6). By Hurwitz' Theorem, H is sos; the construction of [14] gives a representation of H as a sum of nine squares which reduces to (see [12, p. 55]):

$$2H(x,y,z)=(x^2+y^2+z^2)\left((x^2-y^2)^2+(x^2-z^2)^2+(y^2-z^2)^2\right). \tag{1.11}$$

The representations from Hurwitz' proof are not efficient with respect to the number of squares (see [19]); we write H as a sum of five squares of binomials in (5.2). Again, $|E(\mathfrak{H})|=10$ and $C(\mathfrak{H})=A(E(\mathfrak{H}))$. In contrast to $\mathfrak{M}$, $\bar{A}(E(\mathfrak{H}))=C(\mathfrak{H})\backslash\mathfrak{H}$; there are enough even points in $C(\mathfrak{H})$ so that every non-vertex even point is an average of two distinct even points. This will imply, independently of Hurwitz, that H is sos (see Theorem 4.4).

(1.12) *Example.* The Hurwitz trellis $\mathfrak{H}_{n,2d}$ is $\{2d\underline{e}_i\}$, so $\mathfrak{H}_{3,6}=\mathfrak{H}$. Again, it is easy to see that $\mathfrak{H}_{n,2d}$ is a trellis and,

$$C(\mathfrak{H}_{n,2d})=\{\underline{c}=(c_1,\ldots,c_n):0\leqq c_i\in\mathbb{Z}\text{ and }\textstyle\sum c_i=2d\}. \tag{1.13}$$

For $\underline{c}\in C(\mathfrak{H}_{n,2d})$, we have $\Lambda(\underline{c})=\left\{\frac{1}{2d}\underline{c}\right\}$ and, by Hurwitz' Theorem, $G(\underline{c})=2df(\mathfrak{H}_{n,2d},\underline{c})$ is also an sos simplicial agiform. In Example 2.4, we show that $\bar{A}(E(\mathfrak{H}_{n,2d}))=C(\mathfrak{H}_{n,2d})\backslash\mathfrak{H}_{n,2d}$; together with Theorem 4.4, this implies that $G(\underline{c})$ is sos.

Here is an overview of the rest of the paper.

In Sect. 2, we formally introduce an essential definition. If $\mathfrak{U}$ is a framework and $\mathfrak{L}$ is a set of lattice points containing $\mathfrak{U}$, then $\mathfrak{L}$ is "$\mathfrak{U}$-mediated" if every $\underline{w}$ in $\mathfrak{L}\backslash\mathfrak{U}$ is an average of two distinct even points in $\mathfrak{L}$. We give an algorithm for the construction of a maximal $\mathfrak{U}$-mediated set $\mathfrak{U}^*$ for which $A(\mathfrak{U}) \subseteqq \mathfrak{U}^* \subseteqq C(\mathfrak{U})$, and show that $\mathfrak{M}^* = A(\mathfrak{M})$ and $\mathfrak{H}^*_{n,2d} = C(\mathfrak{H}_{n,2d})$. Trellises such as these, for which $\mathfrak{U}^*$ is extreme, are called "M-trellises" and "H-trellises" respectively. We give some sufficient conditions for $\mathfrak{U}$ to be an M-trellis, which combine with previous results of the author on lattice point simplices to give a recipe for the construction of a large number of M-trellises.

In Sect. 3 we present part of the general theory of sos forms. If the agiform $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is sos, we prove that $\underline{w} \in \mathfrak{U}^*$, by constructing a $\mathfrak{U}$-mediated set from the exponents of the monomials involved in the squares. Thus, a non-zero simplicial agiform on an M-trellis is either a binomial square or is not sos. Combined with the construction of the last section, this gives an efficient mechanism for producing large numbers of psd forms which are not sos. The essential arguments used are rephrasings and generalizations of those used by Motzkin, Choi, and Lam to show that M, S, and Q are not sos. We review some results on the Newton polytope of a psd form $p(\underline{x}) = \sum c(\underline{v}) \underline{x}^{\underline{v}} : N(p) = cvx\{\underline{v} : c(\underline{v}) \neq 0\}$. [For example, $p \geqq q \geqq 0$ implies $N(p) \supseteqq N(q)$.] This is useful in studying extremality.

In Sect. 4 we show that the mediation relation $\underline{w} = \frac{1}{2}(\underline{s} + \underline{t})$, $\underline{s}$ and $\underline{t}$ even, implies an identity for the simplicial agiform $f(\mathfrak{U}, \underline{w})$ as a linear combination of $f(\mathfrak{U}, \underline{s})$, $f(\mathfrak{U}, \underline{t})$ and $(\underline{x}^{\underline{s}} - \underline{x}^{\underline{t}})^2$. If $\mathfrak{L}$ is a $\mathfrak{U}$-mediated set, we use this identity to derive a system of linear equations involving the simplicial agiforms on $\mathfrak{U}$. The solution to this system gives $f(\mathfrak{U}, \underline{w})$ as a sum of at most $|\mathfrak{L} \backslash \mathfrak{U}|$ squares of binomials. The results of sections two, three, and four combine to give one main result (Corollary 4.9):

Theorem. *The simplicial agiform $f(\mathfrak{U}, \underline{w})$ is* sos *if and only if $\underline{w} \in \mathfrak{U}^*$.*

It follows that every agiform on an H-trellis is sos (generalizing Hurwitz' Theorem) and an sos simplicial agiform is a sum of squares of binomials.

In Sect. 5, we apply the algorithm of Sect. 4 to write H and $M(x^k, y^k, z^k)$ (for $k = 2, 3$) explicitly as sums of squares in several inequivalent ways. In particular, we obtain H as a sum of five squares. We also discuss non-simplicial agiforms. Difficulties arise from the fact that $\Lambda(\underline{w})$ is not, in general, a singleton, so geometric information on $\underline{w}$ in $C(\mathfrak{U})$ need not translate into information about the agiform $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$.

In Sect. 6, we examine four families of agiforms introduced by Motzkin, Choi, and Lam, generalizing M (twice), S and Q to more variables, and we introduce alternate generalizations of S and Q. Five of these six families of agiforms are not sos; four of them are simplicial. We compute $C(\mathfrak{U})$ and $E(\mathfrak{U})$ for suitable trellises for later use.

In Sect. 7 we give a sufficient condition for extremality. We show that every agiform is a convex combination of simplicial agiforms. An agiform f is "primitive" if it cannot be written as a non-trivial sum of other agiforms; this is weaker than extremality. We show that $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is primitive if and only if it is simplicial and $E(\mathfrak{U}) \subseteqq (\mathfrak{U} \cup \{\underline{w}\})$. A study of the zero-sets of agiforms leads to the following equivalence relation on $\mathbb{Z}^n$: $\underline{v} \sim \underline{v}'$ if $\underline{\varepsilon}^{\underline{w}} = 1$ for $\underline{\varepsilon} \in \{-1, 1\}^n$ implies $\underline{\varepsilon}^{\underline{v}} = \underline{\varepsilon}^{\underline{v}'}$. This relation decomposes $C(\mathfrak{U})$ into equivalence classes $Z_1, \ldots, Z_r$, where $Z_1 \supseteqq \mathfrak{U} \cup \{\underline{w}\}$. We say that $\mathfrak{U}$ is "$\underline{w}$-thin" if $Z_1 = \mathfrak{U} \cup \{\underline{w}\}$ and Z_k is linearly independent for $k \geqq 2$. We show that if $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is extremal, then it is simplicial and $\mathfrak{U}$ is $\underline{w}$-thin.

In Sect. 8, we show that if $\mathfrak{U}$ is not $\underset{\sim}{w}$-thin, then one can construct h so that $f_\alpha = f + \alpha h$ is psd for small $|\alpha|$, so $f = \frac{1}{2}(f_\alpha + f_{-\alpha})$ is not extremal. Thus, we obtain our other main result (Corollary 8.11):

Theorem. *Let f be an agiform. Then f is extremal if and only if f is simplicial and $\mathfrak{U}$ is $\underset{\sim}{w}$-thin.*

We use this to verify the extremality of M, S, and Q. The section ends with a derivation, following [17], of the simplicial agiform as the simplest extremal form which is not a monomial square. This may be viewed as an independent motivation for the study of agiforms.

In Sect. 9, we show that M, S, and Q are each generalized by a family of extremal forms; here are the next forms for each:

(1.14) $$M_4(x,y,z,w) = x^4y^2z^2 + x^2y^4z^2 + x^2y^2z^4 + w^8 - 4x^2y^2z^2w^2,$$

(1.15) $$\overline{S_4}(x,y,z,w) = x^4y^2z^2 + y^4z^2w^2 + z^4w^2x^2 + w^4x^2y^2 - 4x^2y^2z^2w^2,$$

(1.16) $$\overline{Q_6}(x,y,z,w,v,u) = u^6 + x^2y^2z^2 + y^2z^2w^2 + z^2w^2v^2 + w^2v^2x^2 + v^2x^2y^2 - 6xyzwvu.$$

We discuss S_4 and another two non-extremal primitive agiforms in some detail, and begin an analysis of "almost-agiforms."

We conclude in Sect. 10 with some open questions and areas for further research.

Finally, we should observe that our discussion centers on forms rather than inhomogeneous polynomials. This is largely a matter of taste, and we follow Choi and Lam in this decision. It is easy to change from one to the other by homogenizing a polynomial or dehomogenizing a form. The properties of being psd and sos are preserved with the obvious modifications, which we omit. It follows that the definitions and theorems of this paper have straightforward translations from forms to polynomials. The only significant changes are the deletion of the condition $\Sigma_j u_{ij} = 2d$ in the definition of framework, an accompanying adjustment in the criteria for $\mathfrak{U}$ to be a trellis, and a reduction by one in the number of variables.

Motzkin's example was $h(x,y) = x^2y^4 + x^4y^2 + 1 - 3x^2y^2$, which homogenizes to $M: M(x,y,1) = h(x,y)$. One might also homogenize each variable separately and create the biform ([5, p. 20]) $L(x,y,z,w) = x^2y^4z^2 + x^4y^2w^2 + z^4w^4 - 3x^2y^2z^2w^2$. Note that $L(x,y,1,1) = h(x,y)$ and L is a simplicial agiform on $\mathfrak{L} = \{(2,4,2,0), (4,2,0,2), (0,0,4,4)\}$.

Suppose $\mathfrak{S}$ is a polytope in $\mathbb{R}^n$ with vertices $\{z_1, \ldots, z_m\} \subset \mathbb{Z}^n_+$ and let $d = \max_i \left(\sum_{j=1}^{n} z_{ij} \right)$. We embed $\mathfrak{S}$ in the hyperplane $\left\{ \sum_{j=1}^{n+1} z_j = d \right\} \subset \mathbb{R}^{n+1}$ by setting $z_{i,n+1} = d - \sum_{j=1}^{n} z_{ij} \geqq 0$, and call the resulting simplex $\mathfrak{S}'$. Then $\mathfrak{U} = 2\mathfrak{S}'$ is a framework, and $E(\mathfrak{U}) = 2(\mathfrak{S}' \cap \mathbb{Z}^{n+1})$, so $|E(\mathfrak{U})| = |\mathfrak{S} \cap \mathbb{Z}^n|$. We return to this construction in Example 2.6.

This paper is, to a great extent, a rewriting and generalization of the author's [17], which was written very early in his study of psd and sos forms. All his subsequent papers in this subject have been written in collaboration with Man-Duen Choi and Tsit-Yuen Lam (and others). This is thus his first opportunity in many years to thank Professors Choi and Lam in print for their friendship, guidance and assistance.

2. Mediated Sets

Let $\mathfrak{U}$ be a framework. A set $\mathfrak{L} \subset \mathbb{Z}^n$ is called $\mathfrak{U}$-*mediated* if

$$\text{(2.1)} \qquad \mathfrak{U} \subseteqq \mathfrak{L} \subseteqq (\bar{A}(\mathfrak{L}) \cup \mathfrak{U}).$$

That is, $\mathfrak{L}$ is $\mathfrak{U}$-mediated if it contains $\mathfrak{U}$, and every $\underline{v} \in \mathfrak{L} \backslash \mathfrak{U}$ is an average of two distinct even points in $\mathfrak{L}$; $\mathfrak{L}$ need not contain all of $\bar{A}(\mathfrak{L})$. (We will want to find the smallest $\mathfrak{U}$-mediated set containing a given lattice point.) By (2.1), every subset of $A(\mathfrak{U})$ containing $\mathfrak{U}$ is $\mathfrak{U}$-mediated (see also Theorem 2.8). We now give an algorithm for constructing the maximal $\mathfrak{U}$-mediated set, $\mathfrak{U}^*$.

(2.2) **Theorem.** *If $\mathfrak{U}$ is a framework, then there is a $\mathfrak{U}$-mediated set $\mathfrak{U}^*$ satisfying $A(\mathfrak{U}) \subseteqq \mathfrak{U}^* \subseteqq C(\mathfrak{U})$ which contains every $\mathfrak{U}$-mediated set.*

Proof. Define the sequence $\{\mathfrak{U}^k\}$ by $\mathfrak{U}^0 = C(\mathfrak{U})$ and $\mathfrak{U}^{k+1} = \bar{A}(\mathfrak{U}^k) \cup \mathfrak{U}$ for $k \geqq 0$. Then $\mathfrak{U}^1 = (\bar{A}(C(\mathfrak{U})) \cup \mathfrak{U}) \subseteqq C(\mathfrak{U}) = \mathfrak{U}^0$. If $\mathfrak{U}^k \subseteqq \mathfrak{U}^{k-1}$, then $\mathfrak{U}^{k+1} = (\bar{A}(\mathfrak{U}^k) \cup \mathfrak{U}) \subseteqq (\bar{A}(\mathfrak{U}^{k-1}) \cup \mathfrak{U}^k = \mathfrak{U}^k$, hence $\{\mathfrak{U}^k\}$ is a decreasing sequence of finite sets, which must stabilize. Let $\mathfrak{U}^r = \mathfrak{U}^{r+1} = \mathfrak{U}^*$; since $\mathfrak{U}^* = \mathfrak{U}^{r+1} = \bar{A}(\mathfrak{U}^r) \cup \mathfrak{U} = \bar{A}(\mathfrak{U}^*) \cup \mathfrak{U}$, $\mathfrak{U}^*$ is $\mathfrak{U}$-mediated. Further, for $k \geqq 0$, $A(\mathfrak{U}) = (\bar{A}(\mathfrak{U}) \cup \mathfrak{U}) \subseteqq (\bar{A}(\mathfrak{U}^k) \cup \mathfrak{U}) = \mathfrak{U}^{k+1}$, so $A(\mathfrak{U}) \subseteqq \mathfrak{U}^*$, and $\mathfrak{U}^* \subseteqq \mathfrak{U}^0 = C(\mathfrak{U})$.

Let $\mathfrak{L}$ be any $\mathfrak{U}$-mediated set and suppose $\underline{v} \in \mathfrak{L}$ is an extreme point of $cvx(\mathfrak{L})$. Then $\underline{v}$ cannot be an average of two distinct points in $\mathfrak{L}$; $\underline{v} \notin \bar{A}(\mathfrak{L})$. By (2.1), $\underline{v} \in \mathfrak{U}$. Thus, $cvx(\mathfrak{L}) \subseteqq cvx(\mathfrak{U})$ and $\mathfrak{L} \subseteqq C(\mathfrak{U}) = \mathfrak{U}^0$. Since $\mathfrak{L}$ is $\mathfrak{U}$-mediated, $\mathfrak{L} \subseteqq \mathfrak{U}^k$ implies $\mathfrak{L} \subseteqq (\bar{A}(\mathfrak{L}) \cup \mathfrak{U}) \subseteqq (\bar{A}(\mathfrak{U}^k) \cup \mathfrak{U}) = \mathfrak{U}^{k+1}$. It follows by induction that $\mathfrak{L} \subseteqq \mathfrak{U}^*$. □

Note that $\mathfrak{U}^* = C(\mathfrak{U})$ if and only if $C(\mathfrak{U})$ is $\mathfrak{U}$-mediated, that is, if and only if $\bar{A}(E(\mathfrak{U})) = C(\mathfrak{U}) \backslash \mathfrak{U}$.

(2.3) *Example.* (Continuing Example 1.9). Let $\mathfrak{M}$ be the Motzkin trellis. Referring to Fig. 1, we apply the algorithm of Theorem 2.2 to compute $\mathfrak{M}^*$: $\mathfrak{M}^0 = C(\mathfrak{M})$, $\mathfrak{M}^1 = \bar{A}(\mathfrak{M}^0) \cup \mathfrak{M}$. As $\underline{w} \notin \mathfrak{M}^1$, $\mathfrak{M}^1 \cap (2\mathbb{Z})^n = \mathfrak{M}$, so $\bar{A}(\mathfrak{M}^1) = \bar{A}(\mathfrak{M})$ and $\mathfrak{M}^2 = \bar{A}(\mathfrak{M}) \cup \mathfrak{M} = A(\mathfrak{M})$. Hence $A(\mathfrak{M}) \subseteqq \mathfrak{M}^* \subseteqq \mathfrak{M}^2 = A(\mathfrak{M})$, so $\mathfrak{M}^* = A(\mathfrak{M})$.

(2.4) *Example.* (Continuing Example 1.12). Let $\mathfrak{H}_{n,2d} = \{2d\underline{e}_i\}$ be the generic Hurwitz trellis. We show that $\bar{A}(E(\mathfrak{H}_{n,2d})) = C(\mathfrak{H}_{n,2d}) \backslash \mathfrak{H}_{n,2d}$, hence $\mathfrak{H}^*_{n,2d} = C(\mathfrak{H}_{n,2d})$. It suffices to write $\underline{c} \in C(\mathfrak{H}_{n,2d})$, $\underline{c} \neq 2d\underline{e}_i$, in the form $\underline{c} = \frac{1}{2}(\underline{s} + \underline{t})$, with $\underline{s} \neq \underline{t} \in E(\mathfrak{H}_{n,2d})$. Let $b_r = \sum_{i=1}^{r} c_i$, and choose k so that $b_{k-1} < d \leqq b_k$. If

$$\underline{s} = (2c_1, \ldots, 2c_{k-1}, 2d - 2b_{k-1}, 0, \ldots, 0)$$

and

$$\underline{t} = (0, \ldots, 0, 2b_k - 2d, 2c_{k+1}, \ldots, 2c_n),$$

then $\underline{c}=\frac{1}{2}(\underline{s}+\underline{t})$ with $\underline{s}, \underline{t} \in E(\mathfrak{H}_{n,2d})$. (If $\underline{s}=\underline{t}$, then $c_i=0$ for $i \neq k$; that is, $\underline{c}=2d\underline{e}_k$, a case we have excluded.)

We say that a trellis $\mathfrak{U}$ is an *M-trellis* if $\mathfrak{U}^*=A(\mathfrak{U})$; $\mathfrak{U}$ is an *H-trellis* if $\mathfrak{U}^*=C(\mathfrak{U})$. Every trellis in $\mathbb{Z}^3$ is either an H-trellis or an M-trellis (see [20]), but this is false in higher dimensions. If $A(\mathfrak{U})=C(\mathfrak{U})$, for example if $\mathfrak{U}=\mathfrak{H}_{n,2}$, then $\mathfrak{U}$ may be both an H-trellis and an M-trellis. The argument in Example 2.3 generalizes into a useful criterion.

(2.5) **Theorem.** *If $\mathfrak{U}$ is a trellis, and either $E(\mathfrak{U})=\mathfrak{U}$ or $E(\mathfrak{U})=\mathfrak{U}\cup\{\underline{w}\}$ (and $\underline{w}\notin A(\mathfrak{U})$), then $\mathfrak{U}$ is an M-trellis.*

Proof. First, suppose $\mathfrak{U}^k\cap(2\mathbb{Z})^n=\mathfrak{U}$ for some k. Then $\bar{A}(\mathfrak{U}^k)=\bar{A}(\mathfrak{U})$ and $A(\mathfrak{U})\subseteqq\mathfrak{U}^*\subseteqq\mathfrak{U}^{k+1}=(\bar{A}(\mathfrak{U}^k)\cup\mathfrak{U})=(\bar{A}(\mathfrak{U})\cup\mathfrak{U})=A(\mathfrak{U})$, so $\mathfrak{U}^*=A(\mathfrak{U})$. It thus suffices to show that the hypothesis imply $\mathfrak{U}^k\cap(2\mathbb{Z})^n=\mathfrak{U}$ for $k=0$ or 1. If $E(\mathfrak{U})=\mathfrak{U}$, then $\mathfrak{U}^0\cap(2\mathbb{Z})^n=C(\mathfrak{U})\cap(2\mathbb{Z})^n=E(\mathfrak{U})=\mathfrak{U}$. If $E(\mathfrak{U})=\mathfrak{U}\cup\{\underline{w}\}$, then $\mathfrak{U}^0\cap(2\mathbb{Z})^n=\mathfrak{U}\cup\{\underline{w}\}$ and

$$\mathfrak{U}^1=\mathfrak{U}\cup\bar{A}(\mathfrak{U}\cup\{\underline{w}\})=A(\mathfrak{U})\cup\{\tfrac{1}{2}(\underline{w}+\underline{u}_i)\}.$$

Since $\underline{w}\notin A(\mathfrak{U})$ by hypothesis, $\underline{w}\notin\mathfrak{U}^1$, so $\mathfrak{U}^1\cap(2\mathbb{Z})^n=\mathfrak{U}$. □

If $\mathfrak{U}$ is a trellis, $E(\mathfrak{U})=\mathfrak{U}\cup\{\underline{w}\}$ and $\underline{w}\in A(\mathfrak{U})$, then $\mathfrak{U}$ is not an M-trellis. It does not seem useful to elaborate on the conditions under which $\mathfrak{U}^k\cap(2\mathbb{Z})^n=\mathfrak{U}$ for $k\geqq 2$. The number of steps in the computation of $\mathfrak{U}^*$ is bounded above by $|E(\mathfrak{U})\backslash\mathfrak{U}|$, since $\mathfrak{U}^k\cap(2\mathbb{Z})^n=\mathfrak{U}^{k+1}\cap(2\mathbb{Z})^n$ implies $\mathfrak{U}^{k+1}=\mathfrak{U}^{k+2}$ $(=\mathfrak{U}^*)$. This bound is achieved by the trellis

$$\mathfrak{U}_p=\{(0,0,2p),(2,2p-2,0),(4,2,2p-6)\}$$

for $p\geqq 3$, which is discussed more fully in [20]. (Note that $\mathfrak{U}_3=\mathfrak{M}$.) It is not hard to see that

$$E(\mathfrak{U}_p)=\mathfrak{U}_p\cup\{(2,2,2p-4),(2,4,2p-6),\ldots,(2,2p-4,2)\}$$

and that $E(\mathfrak{U}_p^1)\backslash E(\mathfrak{U}_p)=(2,2,2p-4)$, $E(\mathfrak{U}_p^2)\backslash E(\mathfrak{U}_p^1)=(2,4,2p-6)$, etc.

We now show how to transform certain lattice point simplices into M-trellises via homogenization.

(2.6) *Example.* A *k-point n-simplex* (see [18]) is a simplex in $\mathbb{R}^n$ with vertex-set $\mathfrak{S}=\{\underline{z}_0,\ldots,\underline{z}_n\}$, such that $cvx(\mathfrak{S})\cap\mathbb{Z}^n=\mathfrak{S}\cup\{\underline{v}_1,\ldots,\underline{v}_k\}$ and $\underline{v}_i=\sum\lambda_{ij}\underline{z}_j$ with $\sum\lambda_{ij}=1$ and $\lambda_{ij}>0$ [so the $\underline{v}_i$'s are strictly interior to $cvx(\mathfrak{S})$.] Since $\underline{\lambda}_i$ is the unique solution to a linear system with integer coefficients, $\lambda_{ij}\in\mathbb{Q}$. The term "0-point" is equivalent to "fundamental"; fundamental triangles are familiar, and fundamental tetrahedra have been extensively studied.

If $\mathfrak{S}$ is a 1-point n-simplex and $\underline{v}_1=\sum\lambda_i\underline{z}_i$, write $\lambda_i=a_i/D$. Then $\sum\{k\lambda_i\}>1$ for $2\leqq k\leqq D-1$ by [18, p. 228], where $\{x\}=x-[x]$ denotes the fractional part of x. Conversely [p. 230], suppose $n+1$ positive rationals $\lambda_i=a_i/D$ are given, so that $\sum\lambda_i=1$, $\sum\{k\lambda_i\}>1$ for $2\leqq k\leqq D-1$ and a_j and D are relatively prime for some j. (For example, $a_i=1$, $D=n+1$.) Then there exists a canonical 1-point n-simplex $\mathfrak{S}$ with interior point $\underline{v}=\sum\lambda_i\underline{z}_i$. The only such set of rationals satisfying these conditions for $n=2$ is $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$; when $n=3$, there are seven such sets, up to permutation.

Let $\mathfrak{S}=\{\underline{z}_i\}$ be a 0-point or 1-point simplex in $\mathbb{R}^n$, translated so that the components are all non-negative. As described in the introduction, $\mathfrak{S}$ may be homogenized so that we obtain the framework $2\mathfrak{S}'=\mathfrak{U}$ in $\mathbb{R}^{n+1}$. Since $\mathfrak{S}$ is a simplex, $\mathfrak{U}$ is a trellis; either $E(\mathfrak{U})=\mathfrak{U}$ or $E(\mathfrak{U})=\mathfrak{U}\cup\{2\underline{v}\}$, depending on whether $\mathfrak{S}$ is 0-point or 1-point. In the 1-point case, since $n\geqq 2$ and $\underline{v}$ is strictly interior to $\mathfrak{S}$, it follows that $2\underline{v}\notin A(\mathfrak{U})$. In either case, it follows from Theorem 2.5 that $\mathfrak{U}$ is an M-trellis.

As a concrete illustration of this construction, take $n=3$ and $\underline{\lambda}=(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4})$. Then [p. 230] $\mathfrak{S}$ is the tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(-1,-1,4)$, and $\underline{z}=(0,0,1)$ has barycentric coordinates equal to $\underline{\lambda}$. After translation by $(1,1,0)$, we have $d=4$ and

$$\mathfrak{U}=2\mathfrak{S}'=\{(2,2,0,4),\ (4,2,0,2),\ (2,4,0,2),\ (0,0,8,0)\}.$$

Upon permuting the third and fourth coordinates, $\mathfrak{U}=\mathfrak{M}_4$ (see Theorem 6.9). This construction, applied to $\underline{\lambda}=\underline{\varphi}_n=\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$, leads to $\mathfrak{M}_{n+1}$.

We prove the following result in [20], which complements the previous discussion by giving a general way to construct H-trellises.

(2.7) **Proposition.** If $\mathfrak{U}=\{\underline{u}_1,\ldots,\underline{u}_m\}$ is a trellis and $k\geqq\max(2,m-2)$, then $k\mathfrak{U}$ is an H-trellis.

The final result is really a remark used to prove Corollary 4.11.

(2.8) **Theorem.** *If $\mathfrak{U}$ is a trellis and $\underline{w}\in\mathfrak{U}^*$, then there is a $\mathfrak{U}$-mediated set $\mathfrak{F}$ which contains $\underline{w}$ and has at most $|E(\mathfrak{U})\cup\{\underline{w}\}|$ elements.*

Proof. Let $\mathfrak{F}$ consist of the even points in $\mathfrak{U}^*$ plus $\underline{w}$ (if it isn't even.) Then $\bar{A}(\mathfrak{F})=\bar{A}(\mathfrak{U}^*)$, so $\mathfrak{F}$ is $\mathfrak{U}$-mediated. □

When $\mathfrak{U}=\mathfrak{H}_{n,2d}$, we can do better than this. Given $\underline{w}$, use the construction in Example 2.4 to write $\underline{w}=\frac{1}{2}(\underline{s}+\underline{t})$. Since $\underline{s}$ and $\underline{t}$ are even and lie in $\mathfrak{H}_{n,2d}$, we may repeat the argument, and write $\underline{s}$ and $\underline{t}$ as averages, etc. Unless w_1 or w_n is greater than d, $\underline{s}$ and $\underline{t}$ each have fewer than n positive components. It can be shown that this process leads to a $\mathfrak{U}$-mediated set containing $\underline{w}$ which has considerably fewer than $|E(\mathfrak{U})|$ elements. (This is the algorithm of [19], at least for $n\geqq 3$.)

The concept of an H-trellis is related to Handelman's property of two-convexity (see [10, 11]), which arises in the study of the integral closure of certain commutative algebras associated to lattice point polytopes. A lattice point polytope $\mathfrak{S}$ in $\mathbb{R}^n$ is called *two-convex* if every lattice point in $2\mathfrak{S}$ is a sum of two lattice points in $\mathfrak{S}$ [10, p. 150]. After homogenization, $\mathfrak{U}=2\mathfrak{S}'$ is a framework in $\mathbb{R}^{n+1}$, and $\mathfrak{S}$ is two-convex if and only if $C(\mathfrak{U})=A(E(\mathfrak{U}))$. If $\mathfrak{S}$ is a simplex, then two-convexity is a weaker condition than being an H-trellis ($C(\mathfrak{U})=\bar{A}(E(\mathfrak{U}))\cup\mathfrak{U}$.) Handelman [11, p. 33] has proved that, among other properties, $k\mathfrak{U}$ is two-convex if $k\geqq n$. This part of his result is implied by Proposition 2.7.

3. Sos Agiforms: The Necessary Condition

We begin this section with some general remarks about sos forms. Suppose $p=\sum_{k=1}^{r} h_k^2$, where $p(\underline{x})=\sum_{\underline{u}} a(\underline{u})\underline{x}^{\underline{u}}$ and $h_k(\underline{x})=\sum_{\underline{v}} b_k(\underline{v})\underline{x}^{\underline{v}}$. Then

$$\sum_{\underline{u}} a(\underline{u})\underline{x}^{\underline{u}}=\sum_{k}\Big(\sum_{\underline{v}} b_k(\underline{v})\underline{x}^{\underline{v}}\Big)^2 . \tag{3.1}$$

Let $\underline{B}(\underline{v})$ denote the r-tuple whose k-th component is $b_k(\underline{v})$ and let $G(\underline{v},\underline{v}')=\underline{B}(\underline{v})\cdot\underline{B}(\underline{v}')=\sum_k b_k(\underline{v})b_k(\underline{v}')$. By comparing the coefficient of $\underline{x}^{\underline{u}}$ on both sides of (3.1), we obtain the equation:

$$a(\underline{u})=\sum_{\underline{v}+\underline{v}'=\underline{u}} G(\underline{v},\underline{v}')=\sum_{\underline{v}} G(\underline{v},\underline{u}-\underline{v}). \tag{3.2}$$

These observations were made, in effect, by Motzkin [16] and generalized by Choi and Lam [3, 4] into the "term-inspection" method. Refinements on the term-inspection method were made in [17], and results needed in our discussion are presented in Theorem 3.6. The general theory, to be developed in [7], proceeds from the fact that $[\underline{B}(\underline{v})\cdot\underline{B}(\underline{v}')]$ is a psd matrix.

(3.3) **Theorem.** *If $\mathfrak{U}$ is a framework and $f=f(\mathfrak{U},\underline{\lambda},\underline{w})$ is an* sos *agiform, then $\underline{w}\in\mathfrak{U}^*$.*

Proof. Using the notation of (3.1), suppose $f=\sum h_k^2$ and let

$$\mathfrak{N}=\{\underline{v}\colon b_k(\underline{v})\neq 0 \text{ for some } k\}.$$

Let $\mathfrak{L}=2\mathfrak{N}\cup\mathfrak{U}\cup\{\underline{w}\}$. We show that $\mathfrak{L}$ is $\mathfrak{U}$-mediated (and so $\underline{w}\in\mathfrak{L}\subseteq\mathfrak{U}^*$) by writing each $\underline{u}\in\mathfrak{L}\setminus\mathfrak{U}$ as a sum of two distinct points in $\mathfrak{N}$; this implies that $\underline{u}$ is an average of two distinct even points in $2\mathfrak{N}\subseteq\mathfrak{L}$.

If $G(\underline{v},\underline{v}')<0$, then $b_k(\underline{v})b_k(\underline{v}')<0$ for some k, hence $\underline{v}\neq\underline{v}'$ and $\underline{v}$ and $\underline{v}'$ belong to $\mathfrak{N}$. It thus suffices to show that, for $\underline{u}\in\mathfrak{L}\setminus\mathfrak{U}$, there exists $\underline{v}$ with $G(\underline{v},\underline{u}-\underline{v})<0$. Note that $a(\underline{u}_i)=\lambda_i$, $a(\underline{w})=-1$, and $a(\underline{u})=0$ otherwise. By (3.2), we have $-1=a(\underline{w})=\sum G(\underline{v},\underline{w}-\underline{v})$, so $G(\underline{v}_0,\underline{w}-\underline{v}_0)<0$ for some $\underline{v}_0$. If $\underline{u}\neq\underline{w}$, then $\underline{u}\in\mathfrak{L}\setminus(\mathfrak{U}\cup\{\underline{w}\})$ so $a(\underline{u})=0=\sum G(\underline{v},\underline{u}-\underline{v})$. But $\underline{u}\in 2\mathfrak{N}$, so $G(\frac{1}{2}\underline{u},\frac{1}{2}\underline{u})>0$ and there must exist $\underline{v}$ with $G(\underline{v},\underline{u}-\underline{v})<0$ to make the sum vanish. □

(3.4) **Corollary.** *If $\mathfrak{U}$ is an M-trellis and $\underline{w}\notin A(\mathfrak{U})$, then the simplicial agiform $f(\mathfrak{U},\underline{w})$ is not* sos. *Any non-zero* sos *agiform on a M-trellis is a perfect square.*

Proof. If $\mathfrak{U}$ is an M-trellis and $f(\mathfrak{U},\underline{w})$ is sos, then $\underline{w}\in\mathfrak{U}^*=A(\mathfrak{U})$, so either $\underline{w}=\underline{u}_i$ (and $f=0$) or $\underline{w}=\frac{1}{2}(\underline{u}_i+\underline{u}_j)$ [and $f=\frac{1}{2}(\underline{x}^{\underline{u}_i/2}-\underline{x}^{\underline{u}_j/2})^2$]. □

Using the construction of Example 2.6 and Corollary 3.4, one can produce many psd forms which are not sos. The argument used by Motzkin, Choi, and Lam to prove that M, S, and Q are not sos was basically this: if $\underline{w}\notin\mathfrak{U}^1$, then $f(\mathfrak{U},\underline{w})$ is not sos. Since $\mathfrak{U}^1\supseteq\mathfrak{U}^*$, this is a weaker version of Theorem 3.3.

The term-inspection method can be more fully developed. Suppose $p(\underline{x})=\sum a(\underline{u})\underline{x}^{\underline{u}}$ is a (not necessarily psd) form; let $N(p)=cvx\{\underline{v}\colon a(\underline{v})\neq 0\}$ denote the *Newton polytope* or *cage* of p. If f is a simplicial agiform, then $N(f)$ is the simplex whose vertices are $\{\underline{u}_i\colon\lambda_i>0\}$. We let "$\alpha\cdot\leq s$" denote the set $\{\underline{y}\in\mathbb{R}^n\colon\underline{\alpha}\cdot\underline{y}\leq s\}$ and say that it is a *supporting half-space* for $N(p)$ if $\underset{\sim}{\alpha}\cdot \underset{\sim}{z}\leqq s$ for all $\underset{\sim}{z}\in N(p)$ and $\underset{\sim}{\alpha}\cdot \underset{\sim}{z}_0=s$ for some $\underset{\sim}{z}_0\in N(p)$. [By the definition of $N(p)$, this implies that $\underset{\sim}{\alpha}\cdot \underset{\sim}{v}=s$ for some $\underset{\sim}{v}$ with $a(\underset{\sim}{v})\neq 0$.] If $\underset{\sim}{\alpha}\cdot \leqq s$ is a supporting half-space, then $\underset{\sim}{\alpha}\cdot = s$ is called a *supporting hyperplane*. The supporting hyperplanes of p have another interpretation. For p as above, with $a(\underset{\sim}{u})\neq 0$, $\underset{\sim}{\alpha}\in\mathbb{R}^n$ and $t>0$, let

$$p(\underset{\sim}{x},\underset{\sim}{\alpha})(t)=p(t^{\alpha_1}x_1,\ldots,t^{\alpha_n}x_n)=\sum(a(\underset{\sim}{u})\underset{\sim}{x}^{u})t^{\Sigma\alpha_i u_i}=\sum(a(\underset{\sim}{u})\underset{\sim}{x}^{u})t^{\underset{\sim}{\alpha}\cdot \underset{\sim}{u}},$$

and let $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)=\lim_{t\to\infty} t^{-s}p(\underset{\sim}{x},\underset{\sim}{\alpha})(t)$. If $s>\max\underset{\sim}{\alpha}\cdot\underset{\sim}{u}$, then $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)=0$. If $s=\max\underset{\sim}{\alpha}\cdot\underset{\sim}{u}$, then $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)=\{\sum a(\underset{\sim}{u})\underset{\sim}{x}^{u}:\underset{\sim}{\alpha}\cdot\underset{\sim}{u}=s\}<\infty$, and since the summation is over a non-empty set, it is not identically zero. If $0<L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)$ and $r<s$, then $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},r)=\infty$. Thus, if $s<\max\underset{\sim}{\alpha}\cdot\underset{\sim}{u}$, then $L(p,\underset{\sim}{y},\underset{\sim}{\alpha},s)=\infty$ for some $\underset{\sim}{y}$. We have thus proved the following lemma.

(3.5) **Lemma.** *The half-space $\underset{\sim}{\alpha}\cdot\leqq s$ $(=s(\underset{\sim}{\alpha},p))$ is a supporting half-space of $N(p)$ if and only if $|L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)|<\infty$ for all $\underset{\sim}{x}\in\mathbb{R}^n$ and $L(p,\underset{\sim}{y},\underset{\sim}{\alpha},s)\neq 0$ for some $\underset{\sim}{y}$.*

(3.6) **Theorem.** *Let $p(\underset{\sim}{x})=\sum a(\underset{\sim}{v})\underset{\sim}{x}^{v}$ be a* psd *form.*
(i) *If $p\geqq q\geqq 0$, then $N(p)\supseteq N(q)$.*
(ii) *If $p=\sum h_k^2$, then $N(h_j)\subseteq\frac{1}{2}N(p)$ for each h_j.*
(iii) *If $\underset{\sim}{v}_0$ is an extreme point of $N(p)$, then $\underset{\sim}{v}_0$ is an even lattice point and $a(\underset{\sim}{v}_0)>0$.*
(iv) *If F is a face of the polytope $N(p)$, then $p^{(F)}(\underset{\sim}{x})=\sum_{\underset{\sim}{v}\in F} a(\underset{\sim}{v})\underset{\sim}{x}^{v}$ is* psd.

Proof. (i) Since $p(\underset{\sim}{x},\underset{\sim}{\alpha})(t)\geqq q(\underset{\sim}{x},\underset{\sim}{\alpha})(t)$, it follows from multiplication by t^{-s} that $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)\geqq L(q,\underset{\sim}{x},\underset{\sim}{\alpha},s)$. By Lemma 3.5, it follows that $s(\underset{\sim}{\alpha},p)\geqq s(\underset{\sim}{\alpha},q)$. Thus every supporting half-space of $N(p)$ contains $N(q)$. But $N(p)$, as a convex body, is the intersection of its supporting half-spaces, so $N(p)\supseteq N(q)$.

(ii) If $p=\sum h_k^2$, then $p\geqq h_j^2$, so $N(p)\supseteq N(h_j^2)$. Since $L(h^2,\underset{\sim}{x},\underset{\sim}{\alpha},2s)=(L(h,\underset{\sim}{x},\underset{\sim}{\alpha},s))^2$, Lemma 3.5 implies that $N(h_j^2)=2N(h_j)$.

(iii) If $\underset{\sim}{v}_0$ is an extreme point of $N(p)$, then $a(\underset{\sim}{v}_0)\neq 0$ by the definition of $N(p)$, and there exists a supporting hyperplane for $N(p)$ containing only $\underset{\sim}{v}_0$. That is, there exists $\underset{\sim}{\alpha}$ so that $\{\underset{\sim}{v}:\underset{\sim}{\alpha}\cdot\underset{\sim}{v}=s(\underset{\sim}{\alpha},p)\}=\{\underset{\sim}{v}_0\}$ and $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)=a(\underset{\sim}{v}_0)\underset{\sim}{x}^{v_0}$. Since p is psd, $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)\geqq 0$, so $a(\underset{\sim}{v}_0)\underset{\sim}{x}^{v_0}$ is a psd form. Since $a(\underset{\sim}{v}_0)\underset{\sim}{\varepsilon}^{v_0}\geqq 0$ for $\underset{\sim}{\varepsilon}\in\{-1,1\}^n$, it follows that $a(\underset{\sim}{v}_0)\geqq 0$ and v_{0i} is even.

(iv) The argument of (iii) may be applied, choosing $\underset{\sim}{\alpha}\cdot\leqq s$ to be a supporting hyperplane containing F. Once again, $L(p,\underset{\sim}{x},\underset{\sim}{\alpha},s)=p^{(F)}(\underset{\sim}{x})\geqq 0$. □

Both (iii) and (iv) are generalizations of the fact that a non-negative polynomial has even degree and positive leading coefficient. Parts (i), (ii), and (iii) are contained in [17, pp. 365–366], where they are proved as above, but less carefully; parts (i) and (iv) were proved by Handelman [9, p. 53]. In [11, p. 69], Handelman proved that $N(fg)=N(f)+N(g)$, with the usual convex set-sum; this implies the argument used in (ii) that $N(h^2)=2N(h)$.

4. Sos Agiforms: The Sufficient Condition

In this section, we prove the converse of Theorem 3.3 when $\mathfrak{U}$ is a trellis: if $\underset{\sim}{w}\in\mathfrak{U}^*$, then the simplicial agiform $f(\mathfrak{U},\underset{\sim}{w})$ is sos. We need two lemmas: a useful identity and a technical matrix result.

(4.1) Lemma. *Let $\mathfrak{U}$ be a trellis and suppose $\underline{w}=\frac{1}{2}(\underline{s}+\underline{t})$, where $\underline{s}$, $\underline{t}\in E(\mathfrak{U})$. Then*

$$2f(\mathfrak{U},\underline{w})(\underline{x})=f(\mathfrak{U},\underline{s})(\underline{x})+f(\mathfrak{U},\underline{t})(\underline{x})+(\underline{x}^{s/2}-\underline{x}^{t/2})^2\,. \tag{4.2}$$

Proof. Write $\underline{w}$, $\underline{s}$, and $\underline{t}$ in terms of the $\underline{u}_i$'s: $\underline{w}=\sum\lambda_i\underline{u}_i$, $\underline{s}=\sum\sigma_i\underline{u}_i$ and $\underline{t}=\sum\tau_i\underline{u}_i$. By the uniqueness of barycentric coordinates, $\lambda_i=\frac{1}{2}(\sigma_i+\tau_i)$. It follows that the right-hand side of (4.2),

$$(\textstyle\sum\sigma_i\underline{x}^{u_i}-\underline{x}^{s})+(\sum\tau_i\underline{x}^{u_i}-\underline{x}^{t})+(\underline{x}^{s}-2\underline{x}^{w}+\underline{x}^{t})\,,$$

equals the left-hand side. □

If $\underline{s}=\underline{t}=\underline{w}$, (4.2) is vacuous. If $\underline{s}=\underline{u}_k\in\mathfrak{U}$, then $f(\mathfrak{U},\underline{u}_k)(\underline{x})=\underline{x}^{u_k}-\underline{x}^{u_k}$ vanishes identically. Thus one or both of the simplicial agiforms on the right-hand side of (4.2) may disappear.

When $\mathfrak{L}$ is a $\mathfrak{U}$-mediated set, each $\underline{w}\in\mathfrak{L}\backslash\mathfrak{U}$ can be written as $\frac{1}{2}(\underline{s}+\underline{t})$, and so each $f(\mathfrak{U},\underline{w})$ satisfies a non-trivial equation of shape (4.2). This leads to an inhomogeneous linear system involving the simplicial agiforms. The next lemma will be applied to the matrix of that system.

(4.3) Lemma. *Let $A=[a_{ij}]$ be a finite matrix such that $a_{ii}=2$ and $a_{ij}\in\{0,-1\}$ if $i\neq j$. Suppose each row of A has at most two -1's and there is no principal submatrix of A in which each row has exactly two -1's. Then A is invertible, and the entries of A^{-1} are non-negative.*

Proof. Write $A=2(I-P)$ and note that the entries of P are 0 and $\frac{1}{2}$. We shall show that all the eigenvalues of P have modulus <1, so $P^r\to 0$. In this case, $A(I+P+\ldots+P^{r-1})=2(I-P^r)$, and $A^{-1}=\frac{1}{2}(I+P+P^2+\ldots)$ exists and has non-negative entries.

Let α be an eigenvalue of P and let $\underline{z}$ be a non-zero column vector with $A\underline{z}=\alpha\underline{z}$; let $\zeta=\max|z_k|$, $I=\{k:|z_k|=\zeta\}$, $T(i)=\{j:p_{ij}=\frac{1}{2}\}$, and $N(i)=|T(i)|$ [so $0\leqq N(i)\leqq 2$]. Then for $i\in I$,

$$|\alpha|\zeta=|\alpha z_i|=\left|\sum_{j\in T(i)}\frac{1}{2}\,z_j\right|\leqq\frac{1}{2}\,N(i)\zeta\leqq\zeta\,.$$

Since $\zeta>0$, $|\alpha|\leqq 1$. If $|\alpha|=1$, then $N(i)=2$ and for $j\in T(i)$, $|z_j|=\zeta$, so $T(i)\subseteq I$. That is, the principal submatrix of P with rows and columns taken from I has two $\frac{1}{2}$'s in each row. The corresponding principal submatrix in A has two -1's in each row, which violates the hypothesis. It follows that $|\alpha|<1$ for every eigenvalue α, completing the proof. □

(4.4) Theorem. *If $\mathfrak{U}$ is a trellis, $\mathfrak{L}$ is $\mathfrak{U}$-mediated and $\underline{w}\in\mathfrak{L}$, then the simplicial agiform $f=f(\mathfrak{U},\underline{w})$ is* sos. *To be specific, f is a sum of $|\mathfrak{L}\backslash\mathfrak{U}|$ squares of the form $c(\underline{x}^{s}-\underline{x}^{t})^2$, where $2\underline{s}$, $2\underline{t}\in\mathfrak{L}$ and $c\geqq 0$.*

Proof. Index the points of $\mathfrak{L}\backslash\mathfrak{U}$ as $\underline{w}_1,\ldots,\underline{w}_T$, with $\underline{w}=\underline{w}_1$. Since $\mathfrak{L}$ is $\mathfrak{U}$-mediated, at least one of the following three statements is true for each $\underline{w}_i$ and suitable distinct $\underline{u}_r$, $\underline{u}_s$, and $\underline{w}_j$, $\underline{w}_k\in\mathfrak{L}\cap(2\mathbb{Z})^n$:

(4.5) (i) $\quad \underline{w}_i=\frac{1}{2}(\underline{u}_r+\underline{u}_s)\,,$

(4.5) (ii) $\quad \underline{w}_i=\frac{1}{2}(\underline{u}_r+\underline{w}_k)\,,$

(4.5) (iii) $\quad \underline{w}_i=\frac{1}{2}(\underline{w}_j+\underline{w}_k)\qquad (j,k\neq i)\,.$

For the purposes of this proof, pick exactly one correct decomposition for each $\underline{w}_i$. By Lemma 4.1, the relationships in (4.5) have, respectively, the following implications [recall that $f(\mathfrak{U}, \underline{u}_i)=0$]:

(4.6) (i) $\quad 2f(\mathfrak{U}, \underline{w}_i)(\underline{x})=(\underline{x}^{\underline{u}_r/2}-\underline{x}^{\underline{u}_s/2})^2$,

(4.6) (ii) $\quad 2f(\mathfrak{U}, \underline{w}_i)(\underline{x})=f(\mathfrak{U}, \underline{w}_k)(\underline{x})+(\underline{x}^{\underline{u}_r/2}-\underline{x}^{\underline{w}_k/2})^2$,

(4.6) (iii) $\quad 2f(\mathfrak{U}, \underline{w}_i)(\underline{x})=f(\mathfrak{U}, \underline{w}_j)(\underline{x})+f(\mathfrak{U}, \underline{w}_k)(\underline{x})+(\underline{x}^{\underline{w}_j/2}-\underline{x}^{\underline{w}_k/2})^2$.

Define the $T \times T$ matrix $A=[a_{ij}]$ as follows. For all i, $a_{ii}=2$. If $\underline{w}_i$ satisfies (4.5) (i), then the other entries in the i-th row are 0. If $\underline{w}_i$ satisfies (4.5) (ii), then $a_{ik}=-1$ and the other entries in the i-th row are 0. If $\underline{w}_i$ satisfies (4.5) (iii), then $a_{ij}=a_{ik}=-1$ and the other entries in the i-th row are 0. Let h_i denote the binomial square appearing on the right-hand side of the appropriate expression for $2f(\mathfrak{U}, \underline{w}_i)$ in (4.6), so $h_i(\underline{x})=(\underline{x}^{\underline{s}_i}-\underline{x}^{\underline{t}_i})^2$ and $2\underline{s}_i, 2\underline{t}_i \in \mathfrak{L}$. Let $\underline{H}$ and $\underline{F}$ denote the column vectors whose i-th components are $h_i(\underline{x})$ and $f(\mathfrak{U}, \underline{w}_i)$, respectively. By the construction of A, (4.6) can be put into matrix form:

(4.7)
$$A\underline{F}=\underline{H}.$$

Lemma 4.3 was tailored to fit the matrix A; only one hypothesis is not obviously satisfied. Suppose A has a principal submatrix (with rows and columns from I) with two -1's in each row and let $\mathfrak{L}'=\{\underline{w}_i : i \in I\}$. Then every $\underline{w}_i \in \mathfrak{L}'$ satisfies (4.5) (iii) with j and k in I, so every point in $\mathfrak{L}'$ is an average of two other points in $\mathfrak{L}'$. This is clearly impossible for a finite set in $\mathbb{R}^n$; consider a point at maximum distance from the origin. It follows that A has no such principal submatrix, and Lemma 4.3 applies.

From (4.7),

(4.8)
$$\underline{F}=A^{-1}\underline{H},$$

and A^{-1} is a matrix with non-negative entries. The i-th component of $\underline{F}$ can thus be read off from (4.8) as a non-negative linear combination of the h_i's. Thus $f=f(\mathfrak{U}, \underline{w}_1)$ is a sum of $T=|\mathfrak{L}\setminus\mathfrak{U}|$ binomial squares. □

(4.9) **Corollary.** *A simplicial agiform* $f(\mathfrak{U}, \underline{w})$ *is* sos *if and only if* $\underline{w} \in \mathfrak{U}^*$.

Proof. Combine Theorems 3.3 and 4.4. □

(4.10) **Corollary.** *Every simplicial agiform on an H-trellis is* sos.

Example 2.4 and Corollary 4.10 generalize Hurwitz' Theorem.

(4.11) **Corollary.** *An* sos *simplicial agiform* $f(\mathfrak{U}, \underline{w})$ *is a sum of* $|E(\mathfrak{U}) \cup \{\underline{w}\}|-|\mathfrak{U}|$ *squares.*

Proof. Choose $\mathfrak{L}$ using Theorem 2.8. □

It is shown in [19, pp. 110, 111] that the Hurwitzian agiform $G(\underline{c})$ is a sum of $3n-4$ squares [where $\underline{c}=(c_1, \dots, c_n)$ and $\sum c_i=2d$]; this is much smaller than the bound in Corollary 4.11. Proposition 2.7 and Theorem 4.4 can be combined.

(4.12) **Proposition.** *If $f(\mathfrak{U}, \underline{w})$ is simplicial on $\mathfrak{U}=\{\underline{u}_1, \ldots, \underline{u}_m\}$ and $k \geqq \max(2, m-2)$, then the following form is* sos:

$$f(\mathfrak{U}, \underline{w})(x_1^k, \ldots, x_n^k)=\lambda_1 \underline{x}^{k\underline{u}_1}+\ldots+\lambda_m \underline{x}^{k\underline{u}_m}-\underline{x}^{k\underline{w}}=f(k\mathfrak{U}, k\underline{w})(\underline{x}).$$

Finally, one can define trellis isomorphism: $(\mathfrak{U}, \mathfrak{L}) \sim (\mathfrak{V}, \mathfrak{K})$ if, after relabeling, the relevant lattice points satisfy the same averages. Clearly, if $(\mathfrak{V}, \mathfrak{K})$ is the image of $(\mathfrak{U}, \mathfrak{L})$ under a linear map, then averages are preserved. It can be proved that this is the only possible instance of trellis isomorphism: the matrix A contains enough information to give the barycentric coordinates of $\mathfrak{L}$ with respect to $\mathfrak{U}$. Suppose $p(\underline{x})=\sum h_k^2(\underline{x})$, $h_k(\underline{x})=\sum b_k(\underline{v}) \underline{x}^{\underline{v}}$ and $[t_{ij}]$ is a matrix with rational entries so that $v_i'=\sum t_{ij} v_j$ is integral for every $\underline{v}$ occurring in any h_k. Under the formal substitution $x_i \rightarrow \underline{x}^{\underline{t}_i}$, we have $\underline{v} \rightarrow \underline{v}'$, and if p' and h_k' are the resulting forms, then

$$p'(\underline{x})=\sum_k h_k'^2(\underline{x})=\sum_k\Big(\sum_{\underline{v}} b_k(\underline{v}) \underline{x}^{\underline{v}'}\Big)^2.$$

In particular, if $\underline{t}_i=r\underline{e}_i$, then the substitution means replacing the variable x_i by x_i^r. It is not necessary for the t_{ij}'s to be integers; we can invert the previous example and replace x_i by $x_i^{1/r}$ in the form p'.

5. Examples of Sos Agiforms

In this section we implement the algorithm of Theorem 4.4, and write the forms $H(x, y, z)$ and $M(x^k, y^k, z^k)$ for $k=2,3$ as sums of squares. We briefly report some complications for non-simplicial agiforms.

(5.1) *Example* (Continuing Examples 1.10 and 2.4). We wish to write H as a sum of squares. Let $\mathfrak{U}=\mathfrak{H}$ and let $\mathfrak{L}$ consist of the following eight points: $\underline{u}_1=(6,0,0)$, $\underline{u}_2=(0,6,0)$, $\underline{u}_3=(0,0,6)$, $\underline{w}_1=(2,2,2)$, $\underline{w}_2=(2,4,0)$, $\underline{w}_3=(2,0,4)$, $\underline{w}_4=(4,2,0)$, and $\underline{w}_5=(4,0,2)$. We check that $\mathfrak{L}$ is $\mathfrak{U}$-mediated: $\underline{w}_1=\frac{1}{2}(\underline{w}_2+\underline{w}_3)$, $\underline{w}_2=\frac{1}{2}(\underline{w}_4+\underline{u}_2)$, $\underline{w}_3=\frac{1}{2}(\underline{w}_5+\underline{u}_3)$, $\underline{w}_4=\frac{1}{2}(\underline{w}_2+\underline{u}_1)$, and $\underline{w}_5=\frac{1}{2}(\underline{w}_3+\underline{u}_1)$. [By Theorem 2.8, we could have selected $E(\mathfrak{H})$ as the $\mathfrak{U}$-mediated set containing $(2,2,2)$; the advantage of $\mathfrak{L}$ is that it has fewer elements.] Using the terminology of the proof of Theorem 4.4, we have:

$$A=\begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix},$$

and hence

$$6A^{-1}=\begin{bmatrix} 3 & 2 & 2 & 1 & 1 \\ 0 & 4 & 0 & 2 & 0 \\ 0 & 0 & 4 & 0 & 2 \\ 0 & 2 & 0 & 4 & 0 \\ 0 & 0 & 2 & 0 & 4 \end{bmatrix}.$$

The 5-tuple of binomial squares is

$$((xy^2-xz^2)^2,\ (y^3-x^2y)^2,\ (z^3-x^2z)^2,\ (x^3-xy^2)^2,\ (x^3-xz^2)^2)^T.$$

Since $2H(\underline{x})=6f(\mathfrak{H}, \underline{w}_1)(\underline{x})$, we can read off $2H$ as a sum of five binomial squares from the first row of $6A^{-1}$, compare with (1.12) and [19, p. 111]:

$$\begin{aligned}(5.2)\qquad 2H(x,y,z)&=2x^6+2y^6+2z^6-6x^2y^2z^2\\ &=3(xy^2-xz^2)^2+2(y^3-x^2y)^2+2(z^3-x^2z)^2\\ &\quad+(x^3-xy^2)^2+(x^3-xz^2)^2.\end{aligned}$$

The variable x is distinguished in (5.2), and there are two similar expressions in which y and z are distinguished. These correspond to images of $\mathfrak{L}$ under a permutation of coordinates. It turns out that H is a sum of four squares [19, p. 112], but one is the square of a trinomial; H is not a sum of three squares of forms (see [7]).

(5.3) *Example* (Continuing Example 1.9). We turn to $M(x^2, y^2, z^2)$. By Proposition 2.7, $\mathfrak{U}=2\mathfrak{M}$ is an H-trellis; we choose a smaller set. It is easy to check that $\mathfrak{L}=2\mathfrak{M}\cup\{\underline{w}_1, \underline{w}_2, \underline{w}_3\}$ is $2\mathfrak{M}$-mediated, where $\underline{w}_1=(4,4,4)$, $\underline{w}_2=(6,6,0)$, $\underline{w}_3=(2,2,8)$. As before, there is only one way to write the $\underline{w}_i$'s as averages from $\mathfrak{L}$. After some simplification, we obtain the following representation:

$$\begin{aligned}M(x^2,y^2,z^2)&=x^8y^4+x^4y^8+z^{12}-3x^4y^4z^4\\ &=2(x^3y^3-xyz^4)^2+(x^4y^2-x^2y^4)^2+(z^6-x^2y^2z^2)^2.\end{aligned}$$

For $M(x^3, y^3, z^3)$, we use a shortcut: $M(x^3, y^3, z^3)=H(x^2y, xy^2, z^3)$. Thus we can take $x\to x^2y$, $y\to xy^2$, $z\to z^3$ in (5.2), and obtain

$$\begin{aligned}(5.4)\qquad 2M(x^3,y^3,z^3)&=2x^{12}y^6+2x^6y^{12}+2z^{18}-6x^6y^6z^6\\ &=3(x^4y^5-x^2yz^6)^2+2(x^3y^6-x^5y^4)^2\\ &\quad+2(z^9-x^4y^2z^3)^2+(x^6y^3-x^4y^5)^2+(x^6y^3-x^2yz^6)^2.\end{aligned}$$

We could, of course, invert the process and derive (5.2) from (5.4).

Here is a more leisurely representation of $M(x^3, y^3, z^3)$ as a sum of binomial squares; we have suppressed the implicit $3\mathfrak{M}$-mediated set.

$$\begin{aligned}10M(x^3,y^3,z^3)&=10x^{12}y^6+10x^6y^{12}+10z^{18}-30x^6y^6z^6\\ &=15(x^2y^3z^4-x^4y^3z^2)^2+12(xy^2z^6-x^3y^4z^2)^2\\ &\quad+9(x^2y^3z^4-x^6y^3)^2+8(z^9-x^2y^4z^3)^2\\ &\quad+6(x^3y^2z^4-x^3y^6)^2+4(xy^2z^6-x^3y^6)^2\\ &\quad+3(x^2yz^6-x^4y^3z^2)^2+2(z^9-x^4y^2z^3)^2\\ &\quad+(x^6y^3-x^2yz^6)^2.\end{aligned}$$

(In this case, $|\mathfrak{L}\backslash\mathfrak{U}|=9$; it is easier in practice to solve for M as a linear combination of nine specified binomial squares than it is to invert the full 9×9 matrix.)

The major obstacle in generalizing Theorem 4.4 to non-simplicial agiforms is Lemma 4.1. Suppose $\mathfrak{U}$ is not a trellis, and $\underline{w}=\frac{1}{2}(\underline{s}+\underline{t})$ with $\underline{s},\underline{t}\in E(\mathfrak{U})$. For a particular $\underline{\lambda}\in\Lambda(\underline{w})$, there may not exist $\underline{\sigma}\in\Lambda(\underline{s})$ and $\underline{\tau}\in\Lambda(\underline{t})$ with $\underline{\lambda}=\frac{1}{2}(\underline{\sigma}+\underline{\tau})$. In this case, no identity of the shape (4.2) is applicable to $f(\mathfrak{U},\underline{\lambda},\underline{w})$. For example, if

$$\mathfrak{T}=\mathfrak{H}\cup\mathfrak{M}=\{(6,0,0),\ (0,6,0),\ (0,0,6),\ (4,2,0),\ (2,4,0)\},$$

and $\underline{w}=(2,2,2)$, then M and H are both agiforms on $\mathfrak{T}$ with the same $\underline{w}$, but one is sos and the other is not. The Motzkin form corresponds to $\underline{\lambda}=(0,0,\frac{1}{3},\frac{1}{3},\frac{1}{3})\in\Lambda(\underline{w})$. There are several ways to write $\underline{w}=\frac{1}{2}(\underline{s}+\underline{t})$, but if $\underline{\lambda}=\frac{1}{2}(\underline{\sigma}+\underline{\tau})$ with $\underline{\sigma}\in\Lambda(\underline{s})$ and $\underline{\tau}\in\Lambda(\underline{t})$, then $\sigma_i=\tau_i=0$ for $i\leqq 2$, because $\sigma_i,\tau_i\geqq 0$. Thus $\sum\sigma_i\underline{u}_i$ and $\sum\tau_i\underline{u}_i$ belong to $E(\mathfrak{M})$; since $\underline{w}\notin\bar{A}(E(\mathfrak{M}))$, this is a contradiction. Of course, the failure of a lemma to generalize does not mean that the theorem also fails to generalize, and the points $(6,0,0)$ and $(0,6,0)$ in this case are basically irrelevant to the form $f(\mathfrak{U},\underline{\lambda},\underline{w})$.

Here is a less trivial example.

(5.5) *Example.* Let $\mathfrak{B}=\{(6,0,0),(4,2,0),(2,4,0),(0,0,6)\}$. It is not hard to see that

$$E(\mathfrak{B})=\mathfrak{B}\cup\{(2,2,2),\ (2,0,4),\ (4,0,2)\}.$$

For $\underline{w}=(2,2,2)$, a routine computation gives:

$$\Lambda(\underline{w})=\{\underline{\lambda}_t=(\tfrac{1}{3}t,\tfrac{1}{3}(1-2t),\tfrac{1}{3}(1+t),\tfrac{1}{3}):0\leqq t\leqq\tfrac{1}{2}\}.$$

We define the agiform $M_t(x,y,z)=3f(\mathfrak{B},\underline{\lambda}_t,\underline{w})$ for $0\leqq t\leqq\frac{1}{2}$:

$$\begin{aligned}(5.6)\qquad M_t(x,y,z)&=tx^6+(1-2t)x^4y^2+(1+t)x^2y^4+z^6-3x^2y^2z^2\\&=M(x,y,z)+t(x^3-xy^2)^2.\end{aligned}$$

Note that M_t is a simplicial agiform when $t=0$ and $t=\frac{1}{2}$, and is a convex combination of M_0 and $M_{1/2}$ (see also Theorem 7.1). The identity

$$M_{1/8}(x,y,z)=(x^2z-z^3)^2+\tfrac{1}{8}(3xy^2+x^3-4xz^2)^2$$

shows that $M_{1/8}$ is sos. A form is called *sbs* (see [6]) if it is a sum of squares of binomials. By Theorem 4.4, every sos simplicial agiform is also sbs. The following identity shows that $M_{1/2}$ is sbs:

$$M_{1/2}(x,y,z)=\tfrac{3}{2}(xy^2-xz^2)^2+\tfrac{1}{2}(x^3-xz^2)^2+(x^2z-z^3)^2.$$

We shall prove in [7] that M_t is sos if and only if $t\geqq 1/8$ and M_t is sbs if and only if $t=1/2$, so that the two conditions are not equivalent for non-simplicial agiforms.

6. Six Families of Agiforms

In this section we consider six families of agiforms, two generalizations for each of M, S, and Q. One family is due to Motzkin [16, p. 217], three are due to Choi and Lam [4, p. 5] and the other two are new. The subscript of a form always indicates the number of variables.

Motzkin defined a family of psd forms $\{M_n\}$ which are not sos:

$$(6.1)\qquad M_n(\underline{x})=x_1^2\dots x_{n-1}^2\left(\sum_{i=1}^{n-1}x_i^2\right)+x_n^{2n}-nx_1^2\dots x_n^2,$$

so $M_3 = M$ [cf. (1.14)]. [Motzkin actually defined $M_n(x_1, \ldots, x_{n-1}, 1)$; here as elsewhere in this section, we have renamed and renormalized the families.] Choi and Lam gave an alternative generalization of M:

(6.2) $$\bar{M}_n(\underline{x}) = (n-2)x_n^{2n} + \sum_{i \neq j} x_i^{2n-2} x_j^2 - n(n-2)x_1^2 \ldots x_n^2,$$

where the sum is taken over all pairs (i,j) with $1 \leqq i \neq j \leqq n-1$.

They also defined the family $\{S_n\}$:

(6.3) $$S_n(\underline{x}) = \sum_{i=1}^{n} x_i^{2n-2} x_{i+1}^2 - n x_1^2 \ldots x_n^2,$$

where $x_{n+1} = x_1$ and $S_3 = S$. We consider an alternative family:

(6.4) $$\overline{S_n}(\underline{x}) = x_1^2 x_2^2 \ldots x_n^2 \left(\sum_{i=1}^{n} x_i^{-2} x_{i+1}^2 - n \right),$$

where, again, $x_{n+1} = x_1$ [cf. (1.15)].

Finally, Choi and Lam generalized Q for $n = 2m$:

$$Q_{2m}(\underline{x}) = (2m-2)!\, x_{2m}^{2m} + m!\,(m-1)! \sum_{i_1 < \ldots < i_m} x_{i_1}^2 \ldots x_{i_m}^2 - 2m(2m-2)!\, x_1 \ldots x_{2m},$$
(6.5)

where the sum is taken over all m-subsets of $\{1, \ldots, 2m-1\}$ and $Q_4 = 2Q$. We also consider an alternative family $\{\bar{Q}_{2m}\}$:

(6.6) $$\bar{Q}_{2m}(\underline{x}) = x_{2m}^{2m} + \sum_{i=1}^{2m-1} x_i^2 x_{i+1}^2 \ldots x_{i+(m-1)}^2 - 2m x_1 \ldots x_{2m}.$$

The summands are products of m consecutive elements of $\{x_1^2, x_2^2, \ldots, x_{2m-1}^2\}$, viewed cyclically, and $\overline{Q_4} = Q$ [c.f. (1.16)].

We shall show that M_n, S_n, $\bar{S}_n$, and $\bar{Q}_{2m}$ are simplicial agiforms on M-trellises, while $\bar{M}_n$ $(n \geqq 4)$ and Q_{2m} $(m \geqq 3)$ are non-simplicial agiforms. We use Theorem 3.3 to verify the assertions in [16] and [4] that M_n, S_n, and Q_{2m} are not sos; $\bar{S}_n$ and $\bar{Q}_{2m}$ are also not sos. We write $\bar{M}_4$ as a sum of squares, contradicting [4, p. 5]. We compute $E(\mathfrak{U})$ for the appropriate $\mathfrak{U}$; this is used in Theorem 9.1 to show that M_n, $\overline{S_n}$, and $\overline{Q_{2m}}$ are extremal.

Our arguments are simplified by a technical lemma and a proposition on the eigenvalues of a circulant matrix. For notational convenience, we let $\underline{\varphi}_k$ denote the k-tuple $\left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$.

(6.7) **Lemma.** *Suppose $\underline{\sigma} = (\sigma_1, \ldots, \sigma_n)$ with $\sigma_i \geqq 0$ and $\sum \sigma_i = 1$ and suppose that $2(\sigma_i - \sigma_j)$ is an integer for all i, j. Then $\underline{\sigma} = \underline{e}_j$, $\underline{\varphi}_n$, $\frac{1}{2}(\underline{e}_j + \underline{e}_k)$* or $\frac{1}{2}(\underline{e}_j + \underline{\varphi}_n)$.

Proof. Let $s = \min \sigma_i$, so $\underline{\sigma} = (s, \ldots, s) + \underline{\alpha}$, where each α_i is a multiple of $1/2$, $s \geqq 0$, and $ns + \sum \alpha_i = 1$. If $\sum \alpha_i = 1$, then $s = 0$ and $\underline{\sigma} = \underline{e}_j$ (if $\alpha_j = 1$) or $\underline{\sigma} = \frac{1}{2}(\underline{e}_j + \underline{e}_k)$ (if $\alpha_j = \alpha_k = 1/2$). If $\sum \alpha_i = 0$, then $\underline{\alpha} = 0$, $s = 1/n$ and $\underline{\sigma} = \underline{\varphi}_n$. If $\sum \alpha_i = 1/2$, then $\alpha_j = 1/2$, $s = 1/2n$ and $\underline{\sigma} = \frac{1}{2}(\underline{e}_j + \underline{\varphi}_n)$. $\square$

A *circulant* matrix is a square matrix whose rows are successive cyclic permutations of the first. To be specific, $C=[c_{ij}]=\operatorname{circ}(a_1,\dots,a_n)$ is the $n\times n$ matrix with $c_{ij}=a_{j-i+1}$, with the index k in a_k reduced mod n. Circulant matrices are studied in great detail in [8], from which we cite the following result (pp. 66–73).

(6.8) **Proposition.** *The eigenvalues of the matrix* $\operatorname{circ}(a_1,\dots,a_n)$ *are* $\lambda_k=\sum_{i=1}^{n} a_i\varepsilon^{k(i-1)}$, *for* $0\leqq k\leqq n-1$, *where* ε *is a primitive n-th root of unity. The eigenvalues of* $\operatorname{circ}(a_1+d,\dots,a_n+d)$ *are thus* $\lambda_0+nd,\ \lambda_1,\dots,\lambda_{n-1}$.

If $\mathfrak{U}=\{\underline{u}_1,\dots,\underline{u}_n\}\subset\mathbb{R}^n$ is a framework, then $\mathfrak{U}$ is a trellis if and only if $\det[u_{ij}]\neq 0$. For many examples in this section, the matrix $[u_{ij}]$ is circulant (or has a large circulant block), and we shall use Proposition 6.8 to show that the circulant block has non-zero determinant.

(6.9) **Theorem.** *For* $n\geqq 3$, *let*

(6.10)

$$\mathfrak{M}_n=\{(4,2,\dots,2,0),\ (2,4,\dots,2,0),\dots,(2,2,\dots,4,0),\ (0,0,\dots,0,2n)\}=\{\underline{u}_1,\dots,\underline{u}_n\}.$$

Then $\mathfrak{M}_n$ *is a trellis and* M_n *is a simplicial agiform on* $\mathfrak{M}_n$ *which is not* sos. *Further,* $E(\mathfrak{M}_n)=\mathfrak{M}_n\cup\{\underline{w}\}$, *where* $\underline{w}=(2,\dots,2,2)$, *and* $C(\mathfrak{M}_n)=A(E(\mathfrak{M}_n))$.

Proof. Note that $[u_{ij}]$ is a block matrix, consisting of the $n-1\times n-1$ block $\operatorname{circ}(4,2,\dots,2)$ and the 1×1 block $[2n]$. By applying Proposition 6.8 with $d=2$ to $\operatorname{circ}(2,0,\dots,0)=2I$, we find that $\lambda_0=2n$ and $\lambda_k=2$ for $k\neq 0$; thus $\operatorname{circ}(4,2,\dots,2)$ is non-singular and so $\mathfrak{U}$ is a trellis. It is easy to check via (6.1) that $\underline{w}=\sum\tau_i\underline{u}_i$, where $\underline{\tau}=\varphi_n$, so $M_n=nf(\mathfrak{M}_n,\underline{w})$. We shall show that $E(\mathfrak{M}_n)=\mathfrak{M}_n\cup\{\underline{w}\}$; as $\underline{w}\notin A(E(\mathfrak{M}_n))$ for $n\geqq 3$ ($v_n\in\{0,2n\}$ for $\underline{v}\in E(\mathfrak{M}_n)\setminus\{\underline{w}\}$), it follows from Corollary 3.4 that M_n is not sos.

Suppose $\underline{u}=(u_1,\dots,u_n)=\sum\lambda_i\underline{u}_i\in C(\mathfrak{M}_n)$. Then, from (6.10),

$$u_j=2\lambda_1+\dots+4\lambda_j+\dots+2\lambda_{n-1},\quad 1\leqq j\leqq n-1,\quad\text{and}\quad u_n=2n\lambda_n.$$

Since $\sum\lambda_i=1$, $u_j=2+2\lambda_j-2\lambda_n$ for $1\leqq j\leqq n-1$, so $2(\lambda_j-\lambda_n)$ is an integer; hence $2(\lambda_i-\lambda_j)$ is also an integer for all i,j. By Lemma 6.7, either $\underline{\lambda}=\underline{e}_j$ (and $\underline{u}=\underline{u}_j$), $\underline{\lambda}=\frac{1}{2}(\underline{e}_j+\underline{e}_k)$ [and $\underline{u}=\frac{1}{2}(\underline{u}_j+\underline{u}_k)$], $\underline{\lambda}=\varphi_n$ (and $\underline{u}=\underline{w}$), or $\underline{\lambda}=\frac{1}{2}(\underline{e}_j+\varphi_n)$ [and $\underline{u}=\frac{1}{2}(\underline{u}_j+\underline{w})$.] Since $\underline{w}$ and the $\underline{u}_i$'s are even, each average is a lattice point. Each of the distinct averages has at least one odd component, hence $E(\mathfrak{M}_n)=\mathfrak{M}_n\cup\{\underline{w}\}$ and $C(\mathfrak{M}_n)=A(E(\mathfrak{M}_n))$. □

(6.11) *Example.* Let $\bar{\mathfrak{M}}_n$ be the framework implicit in the definition of the family $\{\bar{M}_n\}$; $\bar{\mathfrak{M}}_n$ has $1+(n-1)(n-2)$ elements and so cannot be a trellis for $n\geqq 4$. When $n=3$, $\bar{M}_3=M$, which we know to be a non-sos simplicial agiform. It turns out that $\bar{M}_n=n(n-2)f(\bar{\mathfrak{M}}_n,\underline{\lambda},\underline{w})$ for $\underline{w}=(2,\dots,2)$ and suitable $\underline{\lambda}$. We have found representations for $\bar{M}_4$, $\bar{M}_5$, and $\bar{M}_6$ as a sum of squares and conjecture that $\bar{M}_n$ is sos for $n\geqq 4$. Here is $\bar{M}_4$:

$$\begin{aligned}
&\bar{M}_4(x,y,z,w)\\
&\quad=2w^8+x^6y^2+x^6z^2+y^6x^2+y^6z^2+z^6x^2+z^6y^2-8x^2y^2z^2w^2\\
&\quad=2(w^4-x^2y^2)^2+4(xyz^2-xyw^2)^2+2(x^2y^2-x^2z^2)^2\\
&\qquad+(x^3y-xy^3)^2+(y^3z-yz^3)^2+(z^3x-zx^3)^2.
\end{aligned}$$

(6.12) **Theorem.** *For $n \geqq 3$, let*

(6.13)

$$\mathfrak{S}_n = \{(2n-2, 2, 0, \ldots, 0), (0, 2n-2, 2, \ldots, 0), \ldots, (2, 0, 0, \ldots, 2n-2)\} = \{\underline{u}_1, \ldots, \underline{u}_n\}.$$

Then $\mathfrak{S}_n$ is a trellis and S_n is a simplicial agiform on $\mathfrak{S}_n$ which is not sos. *Further, $E(\mathfrak{S}_n) = \mathfrak{S}_n \cup \{\underline{w}\}$, where $\underline{w} = (2, \ldots, 2)$.*

Sketch of Proof. After showing that $[u_{ij}]$ is circulant and $\mathfrak{S}_n$ is a trellis, one shows that $S_n = nf(\mathfrak{S}_n, \underline{w})$ is a simplicial agiform with $\underline{\tau} = \varphi_n$. Again, $E(\mathfrak{S}_n) = \mathfrak{S}_n \cup \{\underline{w}\}$ will imply that $\underline{w}$ does not belong to $A(E(\mathfrak{S}_n))$, and so S_n is not sos.

If $\underline{u} = \sum \lambda_i \underline{u}_i \in E(\mathfrak{S}_n)$, then for all j, $u_j = 2\lambda_{j-1} + (2n-2)\lambda_j$ from (6.13), with $\lambda_0 = \lambda_n$. Since $\sum u_j = 2n$, if $u_j > 0$ for all j, then $u_j \geqq 2$, hence $\underline{u} = \underline{w}$. Otherwise $u_k = 0$ for some k, so $\lambda_{k-1} = \lambda_k = 0$ and $u_{k-1} = 2\lambda_{k-2}$. Thus λ_{k-2} is 0 or 1. If $\lambda_{k-2} = 1$, then $\underline{\lambda} = \underline{e}_{k-2}$ and $\underline{u} = \underline{u}_{k-2}$. If $\lambda_{k-2} = 0$, then $u_{k-1} = 0$, and the argument may be repeated. Since some u_j is positive, we eventually get $\underline{u} = \underline{u}_i \in \mathfrak{S}_n$ for some i. Thus $E(\mathfrak{S}_n) = \mathfrak{S}_n \cup \{\underline{w}\}$. The set $C(\mathfrak{S}_n)$ is larger than $A(E(\mathfrak{S}_n))$ for $n \geqq 4$ (see Theorem 9.3). □

(6.14) **Theorem.** *For $n \geqq 3$, let*

(6.15)

$$\bar{\mathfrak{S}}_n = \{(4, 2, 2, 2, \ldots, 0), (0, 4, 2, 2, \ldots, 2), (2, 0, 4, 2, \ldots, 2), \ldots\} = \{\underline{u}_1, \ldots, \underline{u}_n\},$$

so that $[u_{ij}]$ is circulant. Then $\bar{\mathfrak{S}}_n$ is a trellis and $\bar{S}_n$ is a simplicial agiform on $\bar{\mathfrak{S}}_n$ which is not sos. *Further, $E(\bar{\mathfrak{S}}_n) = \bar{\mathfrak{S}}_n \cup \{\underline{w}\}$, where $\underline{w} = (2, \ldots, 2)$, and $C(\bar{\mathfrak{S}}_n) = A(E(\bar{\mathfrak{S}}_n))$.*

The analysis of $\{\bar{\mathfrak{S}}_n\}$ is quite similar to that of $\{\mathfrak{M}_n\}$ and is omitted.

(6.16) **Theorem.** *For $m \geqq 2$, let $\mathfrak{Q}_{2m}$ denote the framework:*

(6.17)

$$\mathfrak{Q}_{2m} = \{(0, \ldots, 0, 2m)\} \cup \{(2a_1, \ldots, 2a_{2m-1}, 0) : a_j \in \{0, 1\}, \textstyle\sum a_j = m\} = \{\underline{u}_i\}.$$

Then $E(\mathfrak{Q}_{2m}) = \mathfrak{Q}_{2m}$ and Q_{2m} is an agiform on $\mathfrak{Q}_{2m}$ which is not sos.

Proof. As $\mathfrak{Q}_{2m}$ has $1 + \binom{2m-1}{m}$ elements, it is not a trellis for $2m \geqq 6$; it is easy to see that $\mathfrak{Q}_4$ is a trellis. Let $\underline{\lambda} = \left(\frac{1}{2m}, \gamma, \ldots, \gamma\right)$, where $\gamma = \frac{1}{2}\binom{2m-2}{m-1}^{-1}$ and let $\underline{w} = \sum \lambda_i \underline{u}_i \in C(\mathfrak{Q}_{2m})$. It is easy to check that $\sum \lambda_i = 1$ and $\underline{w} = (1, \ldots, 1, 1)$, so $Q_{2m} = 2m(2m-2)!\, f(\mathfrak{Q}_{2m}, \underline{\lambda}, \underline{w})$.

Since $\mathfrak{Q}_{2m}$ lies in the slab $0 \leqq x_j \leqq 2$ for $j = 1, \ldots, 2m-1$, if $\underline{u} \in E(\mathfrak{Q}_{2m})$, then $u_j = 0$ or 2 for $j < 2m$. If $u_k = 2$ for some k, then $\underline{u}$ is a convex combination of those $\underline{u}_i$'s for which $u_{i,k} = 2$, so $u_{2m} = 0$. Since $\sum u_j = 2m$, this implies that $\underline{u} \in \mathfrak{Q}_{2m}$. Otherwise, $u_j = 0$ for $1 \leqq j \leqq 2m-1$, so $\underline{u} = \underline{u}_1 \in \mathfrak{Q}_{2m}$. By looking at w_{2m}, we see that $\underline{w} \notin A(\mathfrak{Q}_{2m})$, so Q_{2m} is not sos. □

(6.18) **Theorem.** *For $m \geqq 2$ let $\bar{\mathfrak{Q}}_{2m} = \{\underline{u}_i\}$ be defined so that $[u_{ij}]$ has a $2m-1 \times 2m-1$ circulant block,* circ$(2, \ldots, 2, 0, \ldots, 0)$, *with m 2's and $m-1$ 0's, and the 1×1 block $2m$. Then $\bar{\mathfrak{Q}}_{2m}$ is a trellis, $E(\bar{\mathfrak{Q}}_{2m}) = \bar{\mathfrak{Q}}_{2m}$ and $C(\bar{\mathfrak{Q}}_{2m}) = A(\bar{\mathfrak{Q}}_{2m}) \cup \{\underline{w}\}$, where $\underline{w} = (1, \ldots, 1)$, and $\bar{Q}_{2m}$ is a non-*sos *simplicial agiform on $\bar{\mathfrak{Q}}_{2m}$.*

Proof. In this case, $\lambda_k = 2\sum_{i=1}^{m} \varepsilon^{k(i-1)}$, so $\lambda_0 = 2m$; as ε^m is a primitive $(2m-1)$-st root of unity, $\lambda_k = 2(1-\varepsilon^{km})/(1-\varepsilon^k) \neq 0$ for $k \geqq 1$. Thus $\bar{\mathfrak{Q}}_{2m}$ is a trellis; $\bar{Q}_{2m} = 2mf(\bar{\mathfrak{Q}}_{2m}, \underline{w})$. It is possible to argue as in Theorem 6.16 that $E(\bar{\mathfrak{Q}}_{2m}) = \bar{\mathfrak{Q}}_{2m}$; as $\underline{w} \notin A(E(\bar{\mathfrak{Q}}_{2m}))$, $\bar{Q}_{2m}$ is not sos. This may be proved more directly. For any permutation σ of $\{1, \ldots, 2m-1\}$, let

$$\sigma\underline{x} = (x_{\sigma(1)}, \ldots, x_{\sigma(2m-1)}, x_{2m}).$$

It is easily seen that $Q_{2m}(\underline{x})$ is the average (over σ) of $\bar{Q}_{2m}(\sigma\underline{x})$, up to a multiple. If $\bar{Q}_{2m}$ were sos, then Q_{2m} would also be sos, a contradiction.

Suppose $\underline{v} = \sum \lambda_i \underline{u}_i \in C(\bar{\mathfrak{Q}}_{2m})$, then

$$(6.19) \qquad v_i = 2(\lambda_i + \lambda_{i-1} + \ldots \lambda_{i-(m-1)}), \qquad 1 \leqq i \leqq m-1; \qquad v_{2m} = 2m\lambda_{2m},$$

where the indices are understood to cycle so that $\lambda_0 = \lambda_{2m-1}$, $\lambda_{-1} = \lambda_{2m-2}$, etc. With the same cyclic understanding, (6.19) implies that:

$$v_i + v_{i+(m-1)} = 4\lambda_i + 2\sum_{\substack{j=1 \\ j \neq i}}^{2m-1} \lambda_j = 2 + 2\lambda_i - 2\lambda_{2m} \quad \text{for} \quad 1 \leqq i \leqq 2m-1.$$

As before, $2\lambda_i - 2\lambda_{2m}$ is an integer for $1 \leqq i \leqq 2m-1$, and we may apply Lemma 6.7. In this case, $\frac{1}{2}(\underline{w} + \underline{u}_k)$ is not a lattice point because its $2m$-th component is not an integer. Thus $C(\bar{\mathfrak{Q}}_{2m}) = A(\bar{\mathfrak{Q}}_{2m}) \cup \{\underline{w}\}$ and it follows easily that $E(\bar{\mathfrak{Q}}_{2m}) = \bar{\mathfrak{Q}}_{2m}$. □

7. Extremal Agiforms: The Sufficient Condition

Recall that a psd form p is extremal if $p = h_1 + h_2$ with h_i psd implies that $h_i = \alpha_i p$, or equivalently, if $p(\underline{x}) \geqq h(\underline{x}) \geqq 0$ for all $\underline{x}$ (written $p \geqq h \geqq 0$) implies that $h = \alpha p$. Choi and Lam proved that the agiforms M, S, and Q are extremal, and it was shown in [17] that any extremal form with four or fewer terms is, up to a scaling of variables, either a monomial square or a simplicial agiform $f = f(\mathfrak{U}, \underline{w})$, where $|\mathfrak{U}| = 3$, $E(\mathfrak{U}) = \mathfrak{U} \cup \{\underline{w}\}$ and $\underline{w} = \frac{1}{3}(\underline{u}_1 + \underline{u}_2 + \underline{u}_3)$. In the next two sections, we generalize the methods of [17] to agiforms in more than three variables. In this section, we give a weaker version of extremality and discuss the zeros of agiforms. This gives a sufficient condition, which we later prove is necessary.

An agiform f is *primitive* if $f = \sum g_i$, where each g_i is an agiform, implies that $g_i = \alpha_i f$. (Keep in mind that agiforms do not comprise a cone.) An extremal agiform is primitive; it was shown in [17] that the converse is true when $n \leqq 3$, it is false when $n = 4$. It turns out that this is a geometric condition: f is primitive if and only if it is a simplicial agiform on a trellis $\mathfrak{U}$ with $E(\mathfrak{U}) \subseteq (\mathfrak{U} \cup \{\underline{w}\})$. (Necessity follows from Lemma 4.1.)

We begin with two decomposition theorems of independent interest.

(7.1) **Theorem.** *Every agiform is a convex combination of simplicial agiforms.*

Proof. Fix a framework $\mathfrak{U} = \{\underline{u}_1, \ldots, \underline{u}_m\}$, and let $f = f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ be an agiform on $\mathfrak{U}$. The set $A(\underline{w})$ is the intersection of an affine subspace of $\mathbb{R}^m$ (the solutions to $\sum \lambda_i \underline{u}_i = \underline{w}$) and a simplex ($\lambda_i \geqq 0$, $\sum \lambda_i = 1$), and so is a closed convex polytope in $\mathbb{R}^m$. Hence $\underline{\lambda}$ is a convex combination of extreme points, say $\underline{\lambda} = \sum \beta_j \underline{\sigma}_j$, where $\underline{\sigma}_j$ is extremal, $\beta_j > 0$ and $\sum \beta_j = 1$. Thus

$$f(\mathfrak{U}, \underline{\lambda}, \underline{w})(\underline{x}) = \sum_i \lambda_i \underline{x}^{\underline{u}_i} - \underline{x}^{\underline{w}} = \sum_i \left(\sum_j \beta_j \sigma_{ji} \underline{x}^{\underline{u}_i} \right) - \underline{x}^{\underline{w}}$$
$$= \sum_j \beta_j \left(\sum_i \sigma_{ji} \underline{x}^{\underline{u}_i} - \underline{x}^{\underline{w}} \right) = \sum_j \beta_j f(\mathfrak{U}, \underline{\sigma}_j, \underline{w})(\underline{x}).$$

It follows that f is not extremal unless $\underline{\lambda}$ is an extreme point in $\Lambda(\underline{w})$. We now show that if $\underline{\sigma}$ is an extreme point, then $f(\mathfrak{U}, \underline{\sigma}, \underline{w})$ is simplicial. After reindexing, we assume that $\sigma_i > 0$ precisely for $1 \leqq i \leqq r$; we now show that $\mathfrak{U}' = \{\underline{u}_i : 1 \leqq i \leqq r\}$ is a linearly independent set, so $f = f(\mathfrak{U}', \underline{w})$ is a simplicial agiform. Suppose $\sum_{i=1}^{r} \alpha_i \underline{u}_i = \underline{0}$; then $\sum_{j=1}^{n} \left(\sum_{i=1}^{r} \alpha_i u_{ij} \right) = 2d \left(\sum_{i=1}^{r} \alpha_i \right) = 0$. Let $\alpha_j = 0$ for $j > r$ and let $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$. Then $\sum_{i=1}^{m} (\sigma_i + c\alpha_i)\underline{u}_i = \underline{w}$ and $\sum_{i=1}^{m} (\sigma_i + c\alpha_i) = 1$ for all c. For $|c|$ sufficiently small, $\sigma_i + c\alpha_i \geqq 0$ for all i, so $\underline{\sigma} + c\underline{\alpha} \in \Lambda(\underline{w})$. This contradicts the extremality of $\underline{\sigma}$ unless $\underline{\alpha} = \underline{0}$. $\square$

(7.2) **Theorem.** *If $f = f(\mathfrak{U}, \underline{w})$ is a simplicial agiform, where $\underline{w}$ is interior to $\mathfrak{U}$, and $\underline{s} \in E(\mathfrak{U})$, $\underline{s} \notin (\mathfrak{U} \cup \{\underline{w}\})$, then $f(\mathfrak{U}, \underline{w}) = f(\mathfrak{B}, \underline{w}) + \beta f(\mathfrak{U}, \underline{s})$, where $\beta > 0$ and $\mathfrak{B}$ is a trellis contained in $E(\mathfrak{U})$.*

Proof. Suppose $\underline{s} = \sum \alpha_i \underline{u}_i \in E(\mathfrak{U})$, where $\underline{s} \notin \mathfrak{U} \cup \{\underline{w}\}$ and $\underline{\alpha} \in \Lambda(\underline{s})$. Since $\underline{\lambda}$ is interior, $\lambda_i > 0$; after reindexing, we may assume that $\alpha_1/\lambda_1 \geqq \alpha_j/\lambda_j$ for $2 \leqq j \leqq m$, so $\alpha_1 > 0$. Let $\mathfrak{B} = \{\underline{s}, \underline{u}_2, \ldots, \underline{u}_m\} (\neq \mathfrak{U})$. Define $\underline{\beta}$ by $\beta_1 = \lambda_1/\alpha_1$ and $\beta_j = \lambda_j - \alpha_j \beta_1$ for $2 \leqq j \leqq m$. Then $\beta_i > 0$ and a routine computation, which we omit, shows that the barycentric coordinates of $\underline{w}$ with respect to $\mathfrak{B}$ are given by $(\beta_1, \ldots, \beta_m)$. It follows that

$$f(\mathfrak{B}, \underline{\beta}, \underline{w})(\underline{x}) + \beta_1 f(\mathfrak{U}, \underline{\alpha}, \underline{s})(\underline{x})$$
$$= \beta_1 \underline{x}^{\underline{s}} + \sum_{j=2}^{m} (\lambda_j - \beta_1 \alpha_j) \underline{x}^{\underline{u}_j} - \underline{x}^{\underline{w}} + \beta_1 \left(\alpha_1 \underline{x}^{\underline{u}_1} + \sum_{j=2}^{m} \alpha_j \underline{x}^{\underline{u}_j} - \underline{x}^{\underline{s}} \right)$$
$$= \sum_{j=1}^{m} \lambda_j \underline{x}^{\underline{u}_j} - \underline{x}^{\underline{w}} = f(\mathfrak{U}, \underline{\lambda}, \underline{w})(\underline{x}). \quad \square$$

Geometrically speaking, the point $\underline{s}$ subdivides the simplex $cvx(\mathfrak{U})$ into sub-simplices, and $\underline{w}$ lies in the one which maximizes α_j/λ_j.

(7.3) **Theorem.** *The agiform $f = f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is primitive if and only if it is a simplicial agiform on a trellis $\mathfrak{U}$ for which $E(\mathfrak{U}) \subseteq (\mathfrak{U} \cup \{\underline{w}\})$.*

Proof. First suppose f is primitive; by Theorem 7.1, f is a simplicial agiform. Choose $\mathfrak{U}$ so that $f = f(\mathfrak{U}, \underline{w})$ and write $\underline{w} = \sum \lambda_i \underline{u}_i$ with $\lambda_i > 0$. If $\underline{s} \in E(\mathfrak{U})$, $\underline{s} \notin (\mathfrak{U} \cup \{\underline{w}\})$; by Theorem 7.2, $f(\mathfrak{U}, \underline{w})$ is a convex combination of $f(\mathfrak{U}', \underline{w})$ and $f(\mathfrak{U}, \underline{s})$. Since the term $\underline{x}^{\underline{s}}$ has coefficient -1 in $f(\mathfrak{U}, \underline{s})$ and does not appear in $f(\mathfrak{U}, \underline{w})$, $f(\mathfrak{U}, \underline{s}) \neq \alpha f(\mathfrak{U}, \underline{w})$; this contradicts the primitivity of f.

Suppose f is simplicial on $\mathfrak{U}$ with $E(\mathfrak{U}) \subseteq (\mathfrak{U} \cup \{\underline{w}\})$ and $f(\mathfrak{U}, \underline{w}) = \sum g_i$, where $g_i \neq \alpha_i f$ is an agiform. Since the coefficient of $\underline{x}^{\underline{w}}$ is -1 in f, it must be negative in some g_i, so $g_i = \alpha_i f(\mathfrak{B}, \underline{w})$ for some $\mathfrak{B}$. But $f \geqq g_i$, so by Theorem 3.6(i), $N(f) = \mathfrak{U} \supsetneq \mathfrak{B} = N(g_i)$. This is a contradiction, since $\underline{w}$ is not contained in the convex hull of a proper subset of $\mathfrak{U}$. $\square$

We remark that, if $f(\mathfrak{U}, \underline{w})$ is primitive, then either f is a binomial square or $\mathfrak{U}$ is an M-trellis by Theorem 2.5 and f is not sos. Since extremality implies primitivity, we have already shown that, if $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is extremal, then f is simplicial and $E(\mathfrak{U}) \subseteq \mathfrak{U} \cup \{\underline{w}\}$.

Let p be a (not necessarily psd) form. The *zero-set of* p, $\{\underline{z} : p(\underline{z}) = 0\}$, is denoted by $\mathfrak{z}(p)$. Since p is homogeneous, $\underline{z} \in \mathfrak{z}(p)$ if and only if $r\underline{z} \in \mathfrak{z}(p)$ for all real r. It will be convenient to consider $\underline{z}$ and $-\underline{z}$ separately, so we do not define $\mathfrak{z}(p)$ projectively. Let D_i denote the operator $\partial/\partial x_i$. If $h = h(x_1, \ldots, x_n)$ is a form and T is a subset of $\mathbb{R}^n$, then h is *second-order at* T if $D_i h(\underline{y}) = 0$ for all $\underline{y} \in T$ and $i = 1, \ldots, n$. If p is psd and $p(\underline{z}) = 0$, then $D_i p(\underline{z}) = 0$ for all i, so p is second-order at $\mathfrak{z}(p)$. If p is an *m-ic* form which is second-order at T, then $\sum x_i D_i p(\underline{x}) = mp(\underline{x})$ implies that $T \subseteq \mathfrak{z}(p)$.

(7.4) **Lemma.** *Suppose $z_i \neq 0$ for all i and all $\underline{z} \in T$ and $h(\underline{x}) = \sum_{k=1}^{s} c(\underline{v}_k)\underline{x}^{\underline{v}_k}$. Then h is second-order at T if and only if, for all $\underline{z} \in T$ and $i = 1, \ldots, n$,*

$$\sum_{k=1}^{s} \{v_{ki} \underline{z}^{\underline{v}_k}\} c(\underline{v}_k) = 0. \tag{7.5}$$

Proof. Since $x_i D_i(\underline{x}^{\underline{v}}) = v_i \underline{x}^{\underline{v}}$, the left-hand side of (7.5) is simply $z_i D_i h(\underline{z})$; $z_i D_i h(\underline{z}) = 0$ if and only if $D_i h(\underline{z}) = 0$ because $z_i \neq 0$. □

The relevance of second-order sets to extremal forms is shown by the next lemma, which follows from this discussion and Theorem 3.6.

(7.6) **Lemma.** *If p and q are forms and $p \geqq q \geqq 0$, then $N(p) \supseteq N(q)$, $\mathfrak{z}(p) \subseteq \mathfrak{z}(q)$ and q is second-order at $\mathfrak{z}(p)$.*

For $\underline{v} \in \mathbb{Z}^n$, let $\mathfrak{O}(\underline{v}) = \{j : v_j \text{ is odd}\}$. Let $G_n = \{-1, 1\}^n$; G_n forms a group under component-wise multiplication: $\underline{\varepsilon} \cdot \underline{\varepsilon}' = (\varepsilon_1 \varepsilon_1', \ldots, \varepsilon_n \varepsilon_n')$. For a set $I \subseteq \{1, \ldots, n\}$, let $G_n(I)$ denote the subgroup $\left\{\underline{\varepsilon} \in G_n : \prod_{i \in I} \varepsilon_i = 1\right\}$, so $G_n(\phi) = G_n$ and $|G_n(I)| = 2^{n-1}$ for $I \neq \phi$. If $k \in I \setminus J$, then $\Big(1, \ldots, \underset{k\text{-}th}{-1}, \ldots, 1\Big) \in G_n(J) \setminus G_n(I)$; $I \neq J$ implies $G_n(I) \neq G_n(J)$. As $\underline{\varepsilon}^{\underline{v}} = \{\prod \varepsilon_i : v_i \text{ is odd}\}$, $G_n(\mathfrak{O}(\underline{v})) = \{\underline{\varepsilon} : \underline{\varepsilon}^{\underline{v}} = 1\}$.

For $I \subseteq \{1, \ldots, n\}$, usually $I = \mathfrak{O}(\underline{w})$, we define I-congruence on $\mathbb{Z}^n$ by: $\underline{v} \equiv_I \underline{v}'$ if $\underline{\varepsilon}^{\underline{v}} = \underline{\varepsilon}^{\underline{v}'}$ for all $\underline{\varepsilon} \in G_n(I)$. [Equivalently, $\underline{v}$ and $\underline{v}'$ are I-congruent when $\underline{\varepsilon}^{\underline{w}} = 1$ implies $\underline{\varepsilon}^{\underline{v}} = \underline{\varepsilon}^{\underline{v}'}$, or when $\mathfrak{O}(\underline{v} - \underline{v}')$ is either ϕ or I.] We write $\equiv$ for $\equiv_\phi$; if $\underline{v} \equiv \underline{v}'$, then $\underline{v} \equiv_I \underline{v}'$ for all I. We shall be interested in the decomposition of $C(\mathfrak{U}) = Z_1 \cup \ldots \cup Z_r$ into I-equivalence classes. If $\mathfrak{U}$ is a trellis and $I = \mathfrak{O}(\underline{w})$, then $\underline{w} \equiv_I \underline{0}$; hence $\mathfrak{U} \cup \{\underline{w}\}$ always lies within one I-class. We shall always index the classes so that $(\mathfrak{U} \cup \{\underline{w}\}) \subseteq Z_1$.

(7.7) **Lemma.** *If $f = f(\mathfrak{U}, \underline{w})$ is a simplicial agiform, then $G_n(\mathfrak{O}(\underline{w})) \subset \mathfrak{z}(f)$.*

Proof. By (1.8), $f(\mathfrak{U}, \underline{w})(\underline{z}) = 0$ if $\underline{z}^{\underline{u}_1} = \ldots = \underline{z}^{\underline{u}_m} = \underline{z}^{\underline{w}} = 1$. For any $\underline{\varepsilon} \in G_n$, $\underline{\varepsilon}^{\underline{u}_i} = 1$; as noted earlier, $G_n(\mathfrak{O}(\underline{w})) = \{\underline{\varepsilon} \in G_n : \underline{\varepsilon}^{\underline{w}} = 1\}$. □

Simplicial agiforms have other zeros, but these are not useful to us. For example, suppose $\underline{z}^{\underline{u}_1} = \ldots = \underline{z}^{\underline{u}_m} = 0$ [as when $\mathfrak{U} = \mathfrak{M}$ and $\underline{z} = (1, 0, 0)$.] Then

$\underline{z}^{\underline{v}} = \prod (\underline{z}^{\underline{u}_i})^{\sigma_i} = 0$ for $\underline{v} = \sum \sigma_i \underline{u}_i \in C(\mathfrak{U})$ with $\sigma_i \geqq 0$, since at least one σ_i is positive. Thus, if $N(q) \subseteqq C(\mathfrak{U})$, then $\underline{z} \in \mathfrak{Z}(q)$.

(7.8) **Theorem.** *The form* $h(\underline{x}) = \sum_{j=1}^{s} c(\underline{v}_j)\underline{x}^{\underline{v}_j}$ *is second-order at* $G_n(I)$ *if and only if the vector equation*

(7.9)
$$\sum_{\underline{v}_j \in Z_k} c(\underline{v}_j)\underline{v}_j = \underline{0}$$

holds for each I-class Z_k *of the set* $\{\underline{v}_1, \ldots, \underline{v}_s\}$.

We require two technical lemmas (c.f. [2, pp. 295–296].)

(7.10) **Lemma.** *Suppose* $\underline{w}$ *is a lattice point and* $I \subseteqq \{1, \ldots, n\}$. *Then*

(7.11)
$$\sum_{\underline{\varepsilon} \in G_n(I)} \underline{\varepsilon}^{\underline{w}} = \begin{cases} |G_n(I)|, & \text{if } \mathfrak{O}(\underline{w}) = \phi \text{ or } I \\ 0, & \text{otherwise}. \end{cases}$$

Proof. Consider the subgroups $G_n(I)$ and $G_n(\mathfrak{O}(\underline{w})) = \{\underline{w} : \underline{\varepsilon}^{\underline{w}} = 1\}$. If $\mathfrak{O}(\underline{w}) = \phi$ or I, then $\underline{\varepsilon}^{\underline{w}}$ is always 1 on the left-hand side of (7.11), and the sum is $|G_n(I)|$. Otherwise, $G_n(I) \cap G_n(\mathfrak{O}(\underline{w}))$ is a proper subgroup of $G_n(I)$ which must have index two. Thus exactly half the summands in (7.11) are 1 and the other half are -1. [A slicker proof is this: $\chi_{\underline{w}}(\underline{\varepsilon}) = \underline{\varepsilon}^{\underline{w}}$ is a group character on $G = G_n(I)$, and either $\chi = 1_G$ or $\sum_{g \in G} \chi(g) = 0$.] $\square$

(7.12) **Lemma.** *Suppose A is a set of lattice points and* $I \subseteqq \{1, \ldots, n\}$. *Then the following two systems of linear equations are equivalent:*

(7.13)
$$\sum_{\underline{v} \in A} b(\underline{v})\underline{\varepsilon}^{\underline{v}} = 0 \quad \textit{for all } \underline{\varepsilon} \in G_n(I),$$

(7.14)
$$\sum_{\underline{v} \in Z} b(\underline{v}) = 0 \quad \textit{for every I-congruence class } Z \textit{ of } A.$$

Proof. Let $A = Z_1 \cup \ldots \cup Z_r$ be a decomposition of A into I-congruence classes, and choose $\underline{v}_k \in Z_k$. Then, multiplying (7.13) by $\underline{\varepsilon}^{\underline{v}_k}$ $(= \underline{\varepsilon}^{-\underline{v}_k})$ and summing over $\underline{\varepsilon} \in G_n(I)$ gives:

(7.15)
$$0 = \sum_{\underline{\varepsilon} \in G_n(I)} \left(\sum_{\underline{v} \in A} b(\underline{v})\underline{\varepsilon}^{\underline{v} - \underline{v}_k} \right) = \sum_{\underline{v} \in A} b(\underline{v}) \left(\sum_{\underline{\varepsilon} \in G_n(I)} \underline{\varepsilon}^{\underline{v} - \underline{v}_k} \right).$$

By Lemma 7.10, the inner sum is 0 in (7.15), unless $\mathfrak{O}(\underline{v} - \underline{v}_k) = \phi$ or I, in which case $\underline{v} \equiv_I \underline{v}_k$, and the sum is $|G_n(I)|$. Thus, (7.15) reduces to

$$0 = |G_n(I)| \cdot \sum_{\underline{v} \in Z_k} b(\underline{v}),$$

which implies (7.14). Conversely, suppose (7.14) holds for each Z_k. Fix $\underline{\varepsilon} \in G_n(I)$ and sum over the Z_i's, noting that $\underline{v} \in Z_k$ implies $\underline{\varepsilon}^{\underline{v}} = \underline{\varepsilon}^{\underline{v}_k}$. Then

$$\sum_{\underline{v} \in A} b(\underline{v})\underline{\varepsilon}^{\underline{v}} = \sum_{k=1}^{r} \left(\sum_{\underline{v} \in Z_k} b(\underline{v})\underline{\varepsilon}^{\underline{v}} \right) = \sum_{k=1}^{r} \underline{\varepsilon}^{\underline{v}_k} \left(\sum_{\underline{v} \in Z_k} b(\underline{v}) \right) = 0;$$

that is, (7.14) implies (7.13). $\square$

Proof of Theorem 7.8. Lemmas 7.6 and 7.12 combine to show that h is second-order at $G_n(I)$ if and only if, for all k,

$$\sum_{\underline{v}_j \in Z_k} c(\underline{v}_j) v_{ji} = 0, \quad \text{for } i = 1, \ldots, n. \tag{7.16}$$

But (7.16) is just the i-th component of the vector equation (7.9). □

Since $\underline{1} \in G_n(I)$, it follows that $\sum_{\underline{v}_j \in Z_k} c(\underline{v}_j) = 0$ for all classes Z_k.

Theorem 7.8 suggests the following definition. Suppose $\mathfrak{U}$ is a trellis, $\underline{w} \in C(\mathfrak{U})$ is an interior point, $I = \mathfrak{O}(\underline{w})$ and $C(\mathfrak{U})$ is partitioned into I-congruence classes $\{Z_k\}$, where $Z_1 \supseteq \mathfrak{U} \cup \{\underline{w}\}$. Then $\mathfrak{U}$ is *$\underline{w}$-thin* if

$$Z_1 = \mathfrak{U} \cup \{\underline{w}\}, \tag{7.17}$$

$$Z_k \text{ is a linearly independent set if } k \geqq 2. \tag{7.18}$$

If $\underline{w}$ is even, then $Z_1 = E(\mathfrak{U})$, so (7.17) is equivalent to primitivity. One can show using Theorem 2.5 that, if $\mathfrak{U}$ is $\underline{w}$-thin, then either $\mathfrak{U}$ is an M-trellis or $|\mathfrak{U}| = 2$ and $|E(\mathfrak{U})| = 3$. The same trellis may be $\underline{w}$-thin and not $\underline{w}'$-thin. Since $|E(\mathfrak{U}_p)| = p - 2$, the M-trellis $\mathfrak{U}_p$ cannot be $\underline{w}$-thin if $p \geqq 4$. Since $C(\mathfrak{U})$ is contained in $\{\sum x_i = m\}$, (7.18) is satisfied if $|Z_k| \leqq 2$.

(7.19) **Theorem.** *Suppose $f = f(\mathfrak{U}, \underline{w})$ is a simplicial agiform on a trellis $\mathfrak{U}$, where $\underline{w}$ is interior to $\mathfrak{U}$. If $\mathfrak{U}$ is $\underline{w}$-thin, then f is extremal.*

Proof. Write $f(\underline{x}) = \sum_{i=1}^{s} a(\underline{v}_i) \underline{x}^{\underline{v}_i}$, where $C(\mathfrak{U}) = \{\underline{v}_1, \ldots, \underline{v}_s\}$; $a(\underline{u}_i) = \lambda_i$ (>0 as $\underline{w}$ is interior to $\mathfrak{U}$), $a(\underline{w}) = -1$, and $a(\underline{v}) = 0$ otherwise. Suppose $f \geqq h \geqq 0$. As $N(f) \cap \mathbb{Z}^n = C(\mathfrak{U})$, Lemma 7.8 implies that $h(\underline{x}) = \sum_{i=1}^{s} c(\underline{v}_i) \underline{x}^{\underline{v}_i}$ is second-order at $\mathfrak{z}(f)$, which contains $G_n(I)$ by Lemma 7.7. By Theorem 7.8, the $c(\underline{v}_i)$'s satisfy (7.9), where the Z_k's are the I-congruence classes of $C(\mathfrak{U})$. Since $\mathfrak{U}$ is $\underline{w}$-thin, (7.18) implies that $c(\underline{v}) = 0$ if $\underline{v} \in Z_k$, $k \geqq 2$ and for Z_1, (7.9) and (7.17) imply:

$$\underline{0} = \sum_{i=1}^{m} c(\underline{u}_i) \underline{u}_i + c(\underline{w}) \underline{w}. \tag{7.20}$$

As $\mathfrak{U}$ is a trellis, the $\underline{u}_i$'s are linearly independent, and since $\underline{w} = \sum \lambda_i \underline{u}_i$, the complete solution to (7.20) is: $c(\underline{u}_i) = \alpha \lambda_i$, $c(\underline{w}) = -\alpha$ for some α. Thus $c(\underline{v}_i) = \alpha a(\underline{v}_i)$ for all $\underline{v}_i$; that is, $h = \alpha f$. Therefore, f is extremal. □

Parts of this argument were used by Choi and Lam to establish the extremality of M, S, and Q (see, for example, the remark after the proof of Theorem 3.8 in [4, p. 11].) These methods are not applicable to all extremal forms. For example, $p(x, y) = (x - y)^4$ is an extremal form in $P_{2,4}$, but if $q(x, y) = (x - y)^2 b(x, y)$, where b is any psd quadratic form, then $N(q) \subseteq N(p)$ and q is second-order at $\mathfrak{z}(p) = \{(r, r)\}$. Informally, our method is effective when p has only second-order zeros, except at the unit vectors, where $N(p)$ provides more information. [For example, the zeros of M are second-order at G_3, but sixth-order at $(0, 1, 0)$ in the direction of $(0, 0, 1)$, etc.]

We now verify that M, S, and Q are extremal. If $\mathfrak{U} = \mathfrak{M}$ and $\underline{w} = (2, 2, 2)$, then $\mathfrak{O}(\underline{w}) = \phi$ and the I-classes are congruence classes mod 2. Referring back to Fig. 1, we see that the I-classes for $C(\mathfrak{M})$ are:

$$Z_1 = \mathfrak{M} \cup \{\underline{w}\} = \{(4,2,0),(2,4,0),(0,0,6),(2,2,2)\},$$
$$Z_2 = \{(3,3,0),(1,1,4)\}, \qquad Z_3 = \{(2,1,3),(2,3,1)\} \quad \text{and}$$
$$Z_4 = \{(1,2,3),(3,2,1)\}.$$

These satisfy (7.17) and (7.18) so $\mathfrak{M}$ is $\underline{w}$-thin and M is extremal. If $\underline{w}' = (1,1,4)$, then $\mathfrak{O}(\underline{w}) = \{1,2\}$ and $\underline{v} \equiv_{I'} \underline{v}'$ if $v_1 - v_1'$ and $v_2 - v_2'$ are either both even or both odd. Thus, the I'-classes for $\mathfrak{M}$ are $Z_1' = Z_1 \cup Z_2$ and $Z_2' = Z_3 \cup Z_4$, which fail (7.17) and (7.18), so $\mathfrak{M}$ is not $\underline{w}'$-thin. Note that

$$6f(\mathfrak{U}, \underline{w}') = x^4y^2 + x^2y^4 + 4z^6 - 6xyz^4 = M(x,y,z) + 3(z^3 - xyz)^2$$

is not extremal. This is no accident.

If $\mathfrak{S} = \mathfrak{S}_3 = \bar{\mathfrak{S}}_3$ and $\underline{w} = (2,2,2)$, then the I-classes for $C(\mathfrak{S})$ are:

$$Z_1 = \mathfrak{S} \cup \{\underline{w}\} = \{(4,2,0),(0,4,2),(2,0,4),(2,2,2)\},$$
$$Z_2 = \{(1,3,2),(3,1,2)\}, \qquad Z_3 = \{(2,1,3),(2,3,1)\}, \quad \text{and}$$
$$Z_4 = \{(1,2,3),(3,2,1)\},$$

so $\mathfrak{S}$ is $\underline{w}$-thin and $S = 3f(\mathfrak{S}, \underline{w})$ is extremal.

If $\mathfrak{Q} = \mathfrak{Q}_4$ and $\underline{w} = (1,1,1,1)$, then $\mathfrak{O}(\underline{w}) = \{1,2,3,4\}$ and $\underline{v} \equiv_I \underline{v}'$ if $v_i - v_i'$ is either always even or always odd. Then the I-classes of $C(\mathfrak{Q})$ are:

$$Z_1 = \mathfrak{Q} \cup \{\underline{w}\} = \{(4,0,0,0),(0,2,2,0),(0,2,0,2),(0,0,2,2),(1,1,1,1)\},$$
$$Z_2 = \{(2,1,1,0),(0,1,1,2)\}, \qquad Z_3 = \{(2,1,0,1),(0,1,2,1)\} \quad \text{and}$$
$$Z_4 = \{(2,0,1,1),(0,2,1,1)\}.$$

Again, (7.17) and (7.18) are satisfied, so $Q = 4f(\mathfrak{Q}, \underline{w})$ is extremal.

8. Extremal Agiforms: The Necessary Condition

In this section we prove the converse to Theorem 7.19. We begin by showing that, if $\mathfrak{U}$ is not $\underline{w}$-thin, then there exists a form $h \neq \alpha f$ with $N(h) \subseteq N(f)$ which is second-order at $G_n(I)$. We study $f_\alpha = f + \alpha h$ on the orthants of $\mathbb{R}^n$ (so the signs of the variables are absorbed into the coefficients) by reversing the original substitution into the AGI. Although fractional exponents occur, the variables are non-negative. Finally, we show that f_α is psd for $|\alpha|$ small, so $f = \frac{1}{2}(f_\alpha + f_{-\alpha})$ is not extremal.

(8.1) **Lemma.** *If $\mathfrak{U}$ is not $\underline{w}$-thin, then there exists a non-zero $h \neq \alpha f$, involving monomials from a single I-class of $C(\mathfrak{U})$, which is second-order at $G_n(I)$.*

Proof. If (7.18) is violated, there is a non-zero solution to (7.9) for some $k \geq 2$; if (7.17) is violated, there is a solution to (7.9) for $k=1$ which is not a multiple of $\{a(\underline{v}_i)\}$. In either case, by Theorem 7.8, the resulting form $h(\underline{x}) = \sum c(\underline{v})\underline{x}^{\underline{v}}$ is second-order at $G_n(I)$ and $h \neq \alpha f$. The monomials in h come from the chosen Z_k. □

(8.2) **Theorem.** *If $\underline{w}$ is interior to $\mathfrak{U}$ and $\mathfrak{U}$ is not $\underline{w}$-thin, then $f(\mathfrak{U}, \underline{w})$ is not extremal.*

Proof. Construct h by Lemma 8.1 and let $f_\alpha(\underline{x}) = f(\underline{x}) + \alpha h(\underline{x})$; f_α is not a multiple of f if $\alpha \neq 0$. For $\underline{z} \in \mathbb{R}^n$, $f_\alpha(\underline{z}) \geqq 0$ if $f(\underline{z}) = h(\underline{z}) = 0$ or if $f(\underline{z}) > 0$ and $|h(\underline{z})/f(\underline{z})| \leqq |\alpha|^{-1}$. Write $G_n = \{\underline{\varepsilon}_1, \ldots, \underline{\varepsilon}_{2^n}\}$, and let

$$R_k = \{\underline{x} = (\varepsilon_{k1} z_1, \ldots, \varepsilon_{kn} z_n) : z_i \geqq 0\},$$

so $\mathbb{R}^n = \bigcup R_k$. We shall prove there exists $\delta_k > 0$ so that $f_\alpha(\underline{x}) \geqq 0$ for $\underline{x} \in R_k$ if $|\alpha| < \delta_k$. For $\gamma = \min \delta_k$, $f \geqq \frac{1}{2} f_\gamma \geqq 0$, so f is not extremal.

Fix k and write $\underline{\varepsilon}_k = \underline{\varepsilon}$, so $\underline{x}^{\underline{v}} = \underline{\varepsilon}^{\underline{v}} \underline{z}^{\underline{v}}$ with $z_i \geqq 0$. Then,

$$f(\underline{x}) = \lambda_1 \underline{z}^{\underline{u}_1} + \ldots + \lambda_n \underline{z}^{\underline{u}_n} - \underline{\varepsilon}^{\underline{w}} \underline{z}^{\underline{w}}, \tag{8.3}$$

$$h(\underline{x}) = \sum c(\underline{v}_j) \underline{\varepsilon}^{\underline{v}_j} \underline{z}^{\underline{v}_j}, \tag{8.4}$$

where, as in subsequent equations, the change of variables is implicit in the formulas. If $\underline{\varepsilon} \in G_n(I)$, then $\underline{\varepsilon}^{\underline{w}} = 1$ in (8.3) and the $\underline{\varepsilon}^{\underline{v}_j}$'s are all equal in (8.4) (since the $\underline{v}_j$'s all lie in the same I-class); thus $h(\underline{x}) = \pm \sum c(\underline{v}_j) \underline{z}^{\underline{v}_j}$. If $\underline{\varepsilon} \notin G_n(I)$, then $\underline{\varepsilon}^{\underline{w}} = -1$, and we can say nothing about the $\underline{\varepsilon}^{\underline{v}_j}$'s.

We reverse the substitution which turned (1.1) into f. Let $\underline{z}^{\underline{u}_i} = t_i \geqq 0$. If $\underline{v} = \sum \sigma_i \underline{u}_i$, then $\underline{z}^{\underline{v}} = \prod t_i^{\sigma_i}$; write $\underline{v}_j = \sum \sigma_{ji} \underline{u}_i$, where $\sigma_{ji} \geqq 0$ and $\sum \sigma_{ji} = 1$. Then (8.3) and (8.4) become:

$$f(\underline{x}) = \sum \lambda_i t_i - \underline{\varepsilon}^{\underline{w}} \left(\prod t_i^{\lambda_i}\right), \tag{8.5}$$

$$h(\underline{x}) = \sum (\underline{\varepsilon}^{\underline{v}_j}) c(\underline{v}_j) \left(\prod t_i^{\sigma_{ji}}\right). \tag{8.6}$$

These equations use the fact that $N(h) \subseteq C(\mathfrak{U})$ in a critical but implicit way: since $\sigma_{ji} \geqq 0$, the function $\prod t_i^{\sigma_{ji}}$ is continuous for $\underline{t} \in \mathbb{R}_+^m$. If $\underline{v} = \sum \sigma_i \underline{v}_i$ where $\sigma_k < 0$ for some k, then $\prod t_i^{\sigma_i}$ is unbounded near $t_k = 0$.

Using (8.5) and (8.6), let $F(\alpha, \underline{t}) = f_\alpha(\underline{x})$; $f_\alpha(\underline{x}) \geqq 0$ for $\underline{x} \in R_k$ if $F(\alpha, \underline{t}) \geqq 0$ for $\underline{t} \in \mathbb{R}_+^m$; $F(\alpha, \underline{t})$ is homogeneous in $\underline{t}$ of degree 1, but is not a form because of fractional exponents. By its homogeneity, $f_\alpha(\underline{x}) \geqq 0$ for $\underline{x} \in R_k$ if $F(\alpha, \underline{t}) \geqq 0$ on the compact set

$$K = \{\underline{t} = (t_1, \ldots, t_m) : t_i \geqq 0 \text{ and } \sum t_i = m\}.$$

If $\underline{\varepsilon}_k = \underline{\varepsilon} \notin G_n(I)$, then $\underline{\varepsilon}^{\underline{w}} = -1$ and, since $\lambda_i > 0$, $f(\underline{t})$ is strictly positive for $\underline{t} \in K$ from (8.5). Since f and h are continuous and $f > 0$, $\varphi = h/f$ is a continuous function on K and so is bounded; $|\varphi(\underline{x})| \leqq M$ for $\underline{t} \in K$. Then, as previously noted, $f_\alpha(\underline{x}) \geqq 0$ for $|\alpha| < 1/M$ and $\underline{x} \in R_k$.

If $\underline{\varepsilon}_k = \underline{\varepsilon} \in G_n(I)$, then

$$F(\alpha, \underline{t}) = \left\{\sum \lambda_i t_i - \prod t_i^{\lambda_i}\right\} \pm \alpha \left\{\sum c(\underline{v}_j) \left(\prod t_i^{\sigma_{ji}}\right)\right\}. \tag{8.7}$$

There is equality in the AGI on K only at $\underline{1} = (1, \ldots, 1)$: $f(\underline{1}) = 0$ and $f(\underline{t}) > 0$ for $\underline{t} \in K$, $\underline{t} \neq \underline{1}$. (As $\underline{w}$ is interior to $\mathfrak{U}$, $\lambda_i > 0$.) Since $\varphi = h/f$ is continuous on $K \setminus \{\underline{1}\}$, we need to show that φ is bounded near $\underline{1}$. Parameterize K by

$$\underline{t} = (1 + s_1, \ldots, 1 + s_m) \in K, \quad \sum s_i = 0, \tag{8.8}$$

and substitute into (8.7) to get the Taylor series for f and h at $\underline{1}$. First,

$$f(\underline{x}) = \sum \lambda_i (1 + s_i) - \prod \left(1 + \lambda_i s_i + \tfrac{1}{2} \lambda_i (\lambda_i - 1) s_i^2 + \text{higher-order terms}\right).$$

The constant and first-order terms cancel, leaving:

$$f(\underline{x})=\frac{1}{2}\sum\lambda_i(1-\lambda_i)s_i^2-\sum_{i<j}\lambda_i\lambda_j s_i s_j+\text{higher-order terms} \tag{8.9}$$
$$=\sum_{i<j}\frac{1}{2}\lambda_i\lambda_j(s_i-s_j)^2+\text{higher-order terms}.$$

Since $\lambda_i>0$, the leading term is a psd quadratic form which vanishes only at $(c,\dots,c)$. Since $s_m=-\sum_{i=1}^{m-1}s_i$, it is a strictly definite quadratic form in $s_1,\dots,s_{m-1}$, and so is bounded below by $\beta\sum_{i=1}^{m-1}s_i^2$ for some $\beta>0$.

Since $h(\underline{x})=\sum c(\underline{v}_j)\underline{x}^{\underline{v}_j}$ is second-order at $\underline{1}$, $\sum c(\underline{v}_j)=0$; as $\underline{v}_j=\sum\sigma_{ji}\underline{u}_i$, (7.9) implies that

$$\underline{0}=\sum_i\Big(\sum_j\sigma_{ji}c(\underline{v}_j)\Big)\underline{u}_i.$$

But the $\underline{u}_i$'s are linearly independent, so $\sum_j\sigma_{ji}c(\underline{v}_j)=0$ for each i. Thus,

$$h(\underline{x})=\sum_j c(\underline{v}_j)\Big(\prod_i(1+s_i)^{\sigma_{ij}}\Big)$$
$$=\sum_j c(\underline{v}_j)+\sum_i\Big(\sum_j\sigma_{ji}c(\underline{v}_j)\Big)s_i+\text{higher-order terms}.$$

It follows that leading term of h is second-order, say,

$$h(\underline{x})=\sum_{i,j}\gamma_{ij}s_i s_j+\text{higher-order terms} \tag{8.10}$$

for some γ_{ij}. Taking $s_m=-\sum_{i=1}^{m-1}s_i$, we see that the second-order term for h is bounded in absolute value by $\kappa\sum_{i=1}^{m-1}s_i^2$. Thus, (8.9) and (8.10) together imply that $\varphi=h/f$ is bounded near $\underline{t}=\underline{1}$. Since φ is continuous on $K\backslash\underline{1}$, it is bounded on K. As before, there exists $\delta_k>0$ so that $f_\alpha(\underline{x})\geqq 0$ for all $\underline{x}\in R_k$ and $|\alpha|\leqq\delta_k$. We conclude that f is not extremal. □

(8.11) **Corollary.** *The agiform f is extremal if and only if $f=f(\mathfrak{U},\underline{w})$ is simplicial, $\underline{w}$ is interior to $\mathfrak{U}$ and $\mathfrak{U}$ is $\underline{w}$-thin.*

Finally, we derive simplicial agiforms from "first principles"; this argument was given in [17] for $n=3$. Let $p(\underline{x})=\sum_{i=1}^{r}a(\underline{u}_i)\underline{x}^{\underline{u}_i}$ $(a(\underline{u}_i)\neq 0)$ be a form in n variables. We say that p has k *effective variables* if the $r\times n$ matrix $[u_{ij}]$ has rank k, or, equivalently, if $cvx(\{\underline{0},\underline{u}_1,\dots,\underline{u}_r\})$ lies in a k-dimensional subspace of $\mathbb{R}^n$. Since the $\underline{u}_i$'s are contained in the hyperplane $\underline{1}\cdot=2d$, it follows that $N(p)=cvx(\{\underline{u}_i\})$ is, geometrically, $(k-1)$-dimensional.

(8.12) **Theorem.** *Suppose $p(\underline{x})=\sum a(\underline{u}_i)\underline{x}^{\underline{u}_i}$ is a* psd *form in n variables with k effective variables and r terms, and $r\leqq k+1$. Then after a dilation of variables, p is a sum of monomial squares and (possibly) a simplicial agiform.*

Proof. As $N(p)$ is $(k-1)$-dimensional, it has at least k extreme points. By Theorem 3.6, if $\underline{v}$ is extreme, then it is even and $a(\underline{v})>0$, so $a(\underline{v})\underline{x}^{\underline{v}}$ is a monomial square. If p is not a sum of monomial squares, then $r=k+1$ and one of the $\underline{u}_i$'s is not an extreme point. Since $N(p)$ is $(k-1)$-dimensional and has k extreme points, it is a simplex. After relabeling, we denote the extreme points by $\underline{u}_1, \ldots, \underline{u}_k$, and the non-extreme point by $\underline{w}=\sum \lambda_i \underline{u}_i$, with $\lambda_i \geqq 0$ and $\sum \lambda_i = 1$. If $\underline{w}$ is even and $a(\underline{w})>0$, then p is still a sum of monomial squares. If $\underline{w}$ is not even (say w_j is odd) and $a(\underline{w})>0$, take $x_j \to -x_j$ so that $\underline{x}^{\underline{u}_i} \to \underline{x}^{\underline{u}_i}$ and $\underline{x}^{\underline{w}} \to -\underline{x}^{\underline{w}}$. We may thus assume that $p(\underline{x})=\sum a(\underline{u}_i)\underline{x}^{\underline{u}_i} - a(\underline{w})\underline{x}^{\underline{w}}$, where $a(\underline{u}_i)\geqq 0$ and $a(\underline{w})>0$. Suppose $\lambda_i>0$ for $i=1,\ldots,l$ and $\lambda_i=0$ for $i\geqq l+1$, and write $p(\underline{x})=p^*(\underline{x})+\sum_{i=l+1}^{r} a(\underline{u}_i)\underline{x}^{\underline{u}_i}$. In the notation of Theorem 3.6 (iv), $p^*=p^{(F)}$ is psd, where F is the facial simplex with vertices $\{\underline{u}_1, \ldots, \underline{u}_l\}$. Since a sum of monomial squares is still a sum of monomial squares after a dilation, it suffices to show that p^* decomposes in the desired way. For notational convenience, we assume that $l=k$.

By hypothesis, the $k \times n$ matrix $[u_{ij}]$ has full rank, thus, the inhomogeneous linear system of equations

$$\Sigma_j u_{ij} t_j = \log \lambda_i - \log a(\underline{u}_i), \qquad i=1,\ldots,l$$

has at least one solution: $\underline{t}=(s_1,\ldots,s_n)$ and $e^{\Sigma u_{ij}s_j}=\lambda_i(a(\underline{u}_i))^{-1}$. Under the scaling $x_j \to e^{s_j}x_j$, we have $\underline{x}^{\underline{u}_i} \to (e^{\Sigma u_{ij}s_j})\underline{x}^{\underline{u}_i} = \lambda_i(a(\underline{u}_i))^{-1}\underline{x}^{\underline{u}_i}$, and $p^*(\underline{x}) \to q(\underline{x}) = \sum \lambda_i \underline{x}^{\underline{u}_i} - c\underline{x}^{\underline{w}}$, where $c=(e^{\Sigma w_j s_j})a(\underline{w})>0$. Since p^* is psd, so is q and $0 \leqq q(1,\ldots,1)=1-c$. Thus $0<c\leqq 1$, and

$$q(\underline{x}) = c f(\mathfrak{U}, \underline{w})(\underline{x}) + (1-c)\left(\sum \lambda_i \underline{x}^{\underline{u}_i}\right)$$

as asserted. □

9. Examples of Extremal Forms

In Sect. 6, each of the forms M, S, and Q was generalized by two families of agiforms. Since $\{\bar{M}_n\}$ $(n>3)$ and $\{Q_{2m}\}$ $(m>2)$ are not simplicial, they cannot be extremal. In this section we use Theorem 7.19 to show that the families $\{M_n\}$, $\{\bar{S}_n\}$, and $\{\bar{Q}_{2m}\}$ consist of extremal forms, but S_n is not extremal for $n \geqq 4$. (We have already seen that M, S, and Q are extremal.) We also return to a primitive, but not extremal, agiform, first mentioned in [17].

(9.1) **Theorem.** *M_n, $\bar{S}_n$, and $\bar{Q}_{2m}$ are extremal for $n \geqq 3$ and $2m \geqq 4$.*

Proof. For all three cases, let $\underline{u}_{i,j}=\frac{1}{2}(\underline{u}_i+\underline{u}_j)$, $i<j$ and $\underline{w}_i=\frac{1}{2}(\underline{u}_i+\underline{w})$; $\underline{w}_i \in C(\mathfrak{U})$ if and only if $\underline{w}$ is even. If $\underline{v}$ and $\underline{v}'$ are in $C(\mathfrak{U})$ and $v_i \equiv_I v_i'$ for $i \leqq n-1$, then $\sum v_i = \sum v_i'$ implies $\underline{v} \equiv \underline{v}'$.

Let $\mathfrak{U}=\mathfrak{M}_n$. Since $\underline{w}=(2,\ldots,2)$ and $E(\mathfrak{M}_n)=\mathfrak{M}_n \cup \{\underline{w}\}$ by Theorem 6.9, (7.17) is satisfied, and $\equiv_I$ is congruence mod 2. For $i<j\leqq n-1$, $\underline{u}_{i,j}=(a_1,\ldots,a_{n-1},0)$ with 3's in the i-th and j-th place and 2's elsewhere. For $i \leqq n-1$, $\underline{u}_{i,n}=(1,\ldots,2,\ldots,1,n)$ and $\underline{w}_i=(2,\ldots,3,\ldots,2,1)$ with the i-th place distinguished, and $\underline{w}_n=(1,\ldots,1,\mathrm{n}+1)$. It is easy to see that no two of these are congruent, so each Z_k is a singleton for $k \geqq 2$, confirming (7.18).

A similar argument, whose details we omit, works for $\mathfrak{U}=\mathfrak{S}_n$.

For $\mathfrak{U}=\bar{\mathfrak{Q}}_{2m}$, $m\geqq 3$, $\underline{w}=(1,\dots,1)$ and $\underline{v}\equiv\underline{v}'$ provided v_i-v_i' is always even or always odd. Let $\underline{e}_i$ denote the i-th unit vector. Then, for $i<j\leqq 2m-1$, $\underline{u}_{ij}=\underline{e}_i+\dots+\underline{e}_{i+m-1}+\underline{e}_j+\dots+\underline{e}_{j+m-1}$, with the subscripts reduced $\bmod 2m-1$; $\underline{u}_{i,2m}=\underline{e}_i+\dots+\underline{e}_{i+m-1}+(0,\dots,0,m)$, and, by Theorem 6.18, $C(\bar{Q}_{2m})$ consists of these vectors and $\bar{Q}_{2m}\cup\{\underline{w}\}$. For suitable k, the first $2m-1$ components of the $\underline{u}_{ij}$'s have the following pattern mod 2, taken cyclically: $1^k0^{m-k}1^k0^{m-1-k}$. It follows that $Z_1=\bar{Q}_{2m}\cup\{\underline{w}\}$, so (7.17) is satisfied, and no two other elements of $C(\bar{Q}_{2m})$ are I-congruent, so (7.18) is satisfied and Q_{2m} is extremal. □

It can be shown that two conditions, satisfied above, are sufficient, but not necessary, for $\mathfrak{U}$ to be $\underline{w}$-thin:

$$\text{(9.2)(i)} \qquad E(\mathfrak{U})=(\mathfrak{U}\cup\{\underline{w}\}) \quad\text{and}\quad C(\mathfrak{U})=A(\mathfrak{U}\cup\{\underline{w}\}),$$

$$\text{(9.2)(ii)} \qquad E(\mathfrak{U})=\mathfrak{U} \quad\text{and}\quad C(\mathfrak{U})=A(\mathfrak{U})\cup\{\underline{w}\}.$$

(9.3) **Theorem.** *For $n\geqq 4$, S_n is primitive but not extremal.*

Proof. The primitivity of S_n follows from Theorems 6.12 and 7.3. We shall show that $(\alpha,\beta,\gamma,\delta,2,\dots,2)\in C(\mathfrak{S}_n)$ for each permutation $(\alpha,\beta,\gamma,\delta)$ of $(1,1,3,3)$. These six linearly dependent points are congruent mod 2, so they belong to the same Z_k, $\mathfrak{S}_n$ is not $\underline{w}$-thin and S_n is not extremal.

Suppose $\underline{v}=(v_1,\dots,v_n)\in C(\mathfrak{S}_n)$ with $v_i\geqq 0$ and $\sum v_i=2n$. By the proof of Theorem 6.12, the barycentric coordinates of $\underline{v}$ satisfy the equations $v_j=2\lambda_{j-1}+(2n-2)\lambda_j$ for $1\leqq j\leqq n$, with the indices viewed cyclically. It is easy to invert this system. Let $\varrho=-(n-1)^{-1}$; then

$$\text{(9.4)} \qquad v_j+\varrho v_{j-1}+\varrho^2v_{j-2}+\dots+\varrho^{n-1}v_{j+1}=2\{(n-1)+\varrho^{n-1}\}\lambda_j.$$

Let $t_j=v_j+\varrho v_{j-1}$. If n is even, (9.4) implies that λ_j is a non-negative linear combination of $t_j, t_{j-2},\dots,t_{j+2}$. If n is odd, there is the "leftover" term $\varrho^{n-1}v_{j+1}$. In any event, if $t_j\geqq 0$ for all j (and $v_j\geqq 0$ if n is odd), then $\lambda_j\geqq 0$ and $\underline{v}\in C(\mathfrak{S}_n)$. Since $(n-1)t_j=(n-1)v_j-v_{j-1}$, any $\underline{v}$ with $\sum v_j=2n$ and $v_j\in\{1,2,3\}$ belongs to $C(\mathfrak{S}_n)$ when $n\geqq 4$. □

(9.5) *Example.* We decompose S_4 into psd forms by carrying out in detail the proof of Theorem 8.2. The lattice points $(3,1,3,1)$, $(3,1,1,3)$, $(1,3,3,1)$, and $(1,3,1,3)$ are linearly dependent members of the same Z_k for $C(\mathfrak{S}_4)$, and $h(x,y,z,w)=xyzw(x^2-y^2)(z^2-w^2)$ is second-order at G_4. Let

$$\text{(9.6)} \qquad \begin{aligned} f_\alpha(x,y,z,w)&=x^6y^2+y^6z^2+z^6w^2+w^6x^2-4x^2y^2z^2w^2\\ &\quad+\alpha xyzw(x^2-y^2)(z^2-w^2).\end{aligned}$$

Since $f_\alpha(x,y,z,-w)=f_{-\alpha}(x,y,z,w)$, f_α is psd if and only if $f_{-\alpha}$ is psd. We show that f_2 is psd. The next two identities may be routinely verified:

$$\text{(9.7)} \qquad f_2(x,y,z,w)=(x^3y+z^3w)^2+(y^3z+w^3x)^2-2xyzw(yz+xw)^2,$$

$$\text{(9.8)} \qquad \begin{aligned} 10f_2(x,y,z,w)&=(7x^6y^2+y^6z^2+3z^6w^2+9w^6x^2-20x^3yzw^3)\\ &\quad+(3x^6y^2+9y^6z^2+7z^6w^2+w^6x^2-20xy^3z^3w)\\ &\quad+20xyzw(xz-yw)^2.\end{aligned}$$

If $xyzw \leqq 0$, then (9.7) shows that $f_2(x,y,z,w) \geqq 0$. If $xyzw \geqq 0$, then (9.8) gives $10f_2$ as a sum of two simplicial agiforms and a non-negative term, so $f_2(x,y,z,w) \geqq 0$. Thus f_2 is psd and $S_4 = \frac{1}{2}(f_2 + f_{-2})$ is not extremal. We have been unable to prove that f_2 is extremal, and suspect that it is not.

(9.9) *Example.* A version of this example was announced in [17, p. 373]. Let

$$D(x,y,z,w) = 2x^4y^2 + 2x^2y^4 + z^4w^2 + z^2w^4 - 6x^2y^2zw, \tag{9.10}$$

$D = 6f(\mathfrak{D}, \underline{w})$, where $\underline{w} = (2,2,1,1)$ and

$$\mathfrak{D} = \{(4,2,0,0), (2,4,0,0), (0,0,4,2), (0,0,2,4)\}.$$

It is easy to check that $C(\mathfrak{D}) = A(\mathfrak{D}) \cup \{\underline{w}, \underline{w}'\}$, where $\underline{w}' = (1,1,2,2)$, so the I-classes of $C(\mathfrak{D})$ with respect to $\underline{w}$ are:

$$\begin{aligned} (9.11)\quad & Z_1 = \{(4,2,0,0), (2,4,0,0), (0,0,4,2), (0,0,2,4), (0,0,3,3), (2,2,1,1)\}, \\ & Z_2 = \{(1,1,2,2), (3,3,0,0)\}, \qquad Z_3 = \{(1,2,1,2), (1,2,2,1)\}, \\ & Z_4 = \{(2,1,1,2), (2,1,2,1)\}. \end{aligned}$$

We see that Z_k is independent for $k \geqq 2$, but $\mathfrak{D}$ is not $\underline{w}$-thin, because Z_1 contains $(0,0,3,3)$. Let

$$E(x,y,z,w) = 4x^4y^2 + 4x^2y^4 + z^4w^2 + 2z^3w^3 + z^2w^4 - 12x^2y^2zw; \tag{9.12}$$

$2D(\underline{x}) = E(\underline{x}) + (z^2w - zw^2)^2$; E is not an agiform because $(0,0,3,3)$ is not even. We shall show that E is psd; since $D \geqq \frac{1}{2}E$, this verifies that D is not extremal. Observe that (9.12) arises from the substitution $t_1 = x^4y^2$, $t_2 = x^2y^4$, $t_3 = z^4w^2$, $t_4 = z^3w^3$, $t_5 = z^2w^4$, and $\underline{\lambda} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{12}, \frac{1}{6}, \frac{1}{12})$ into (1.1). If $zw \geqq 0$, then $t_3 \geqq 0$, so $E \geqq 0$. If $zw \leqq 0$, then

$$E(x,y,z,w) = 4x^4y^2 + 4x^2y^4 + (z^2w + zw^2)^2 + (-12zw)x^2y^2$$

shows that $E \geqq 0$. Thus E is psd, as asserted.

Finally, we show that E is extremal. Note that $N(E) = C(\mathfrak{D})$ and $\mathfrak{z}(E)$ contains $G_4(\{3,4\})$ and $(0,0,1,-1)$. If $E \geqq q \geqq 0$, then q is second-order at $\mathfrak{z}(E)$ and $N(q) \subseteq C(\mathfrak{D})$. Write $q(\underline{x}) = \sum c(\underline{v}_i)\underline{x}^{\underline{v}_i}$ and apply Theorem 7.8 to $G_4(\{3,4\})$. Since Z_2, Z_3, and Z_4 are linearly independent, $c(\underline{v}_i) = 0$ unless $\underline{v}_i \in Z_1$. After relabeling, we have

$$\begin{aligned} & q(x,y,z,w) = c_1x^4y^2 + c_2x^2y^4 + c_3z^4w^2 + c_4z^3w^3 + c_5z^2w^4 + c_6x^2y^2zw, \\ (9.13)\quad & 4c_1 + 2c_2 + 2c_6 = 2c_1 + 4c_2 + 2c_6 = 4c_3 + 3c_4 + 2c_5 + c_6 \\ & = 2c_3 + 3c_4 + 4c_5 + c_6 = 0. \end{aligned}$$

One parametric solution of the system in (9.13) is: $c_1 = c_2 = 4\alpha$, $c_3 = c_5 = \alpha + \beta$, $c_4 = 2\alpha - 2\beta$, $c_6 = -12\alpha$. Hence,

$$q(x,y,z,w) = \alpha E(x,y,z,w) + \beta(z^2w - zw^2)^2. \tag{9.14}$$

Taking $\alpha = \beta = \frac{1}{2}$ in (9.14) returns D; this is not surprising, as we have only used $\mathfrak{z}(D)$ so far. Finally, $q(0,0,1,-1) = 0$ implies $\beta = 0$, so $q = \alpha E$, and we conclude that E is extremal.

This example suggests that some classification may be possible for "almost agiforms" such as E.

10. Further Questions

We conclude with some open questions about agiforms. Of course, these are dwarfed in importance by the questions which motivated this paper. Given a psd form p, how can one tell whether it is sos? Given a psd form p, how can one tell whether it is extremal?

We begin with mediated sets. Given a framework $\mathfrak{U}$, is there an algorithm for computing $\mathfrak{U}^*$ which is more efficient than Theorem 2.2? Does the set of $\mathfrak{U}$-mediated sets for fixed $\mathfrak{U}$ have interesting properties? Given $\underline{w} \in \mathfrak{U}^*$, is there an easy way to compute a "small" $\mathfrak{U}$-mediated set containing $\underline{w}$? [One algorithm is to list all averages of distinct points in $E(\mathfrak{U})$, then build a $\mathfrak{U}$-mediated set containing $\underline{w}$ by starting with the average $\underline{w} = \frac{1}{2}(\underline{s} + \underline{t})$, then finding averages for $\underline{s}$ and $\underline{t}$ (unless they are in $\mathfrak{U}$) etc.; the branching looks horrendous.] What is the worst-case bound on size? If $\mathfrak{U} \subset \mathbb{R}^n$ lies in the hyperplane $\underline{1} \cdot = 2d$ and $\underline{w} \in \mathfrak{U}^*$, then by Theorem 2.8, $\underline{w}$ is contained a $\mathfrak{U}$-mediated set with at most $E(\mathfrak{H}_{n,2d}) + 1 = \binom{d+n-1}{d} + 1$ elements. One expects to do better. Handelman's publications contain many results and open questions related to the material in Proposition 2.7, as will [20].

After the generality of Theorem 3.3, it is painful that Theorem 4.4 is limited to simplicial agiforms; but the examples in Sect. 5 cast doubt on a general "yes-or-no" lattice point criterion for deciding whether $f(\mathfrak{U}, \underline{\lambda}, \underline{w})$ is sos in the non-simplicial case. For fixed $\underline{w}$, can anything be said about $\{\underline{\lambda} \in \Lambda(\underline{w}) : f(\mathfrak{U}, \underline{\lambda}, \underline{w}) \text{ is sos}\}$? What is the "probability" that a simplicial agiform is sos? (Is this question meaningful?) It is proved in [20] that every trellis for $m = 3$ is either an H-trellis, or the image of some $\mathfrak{U}_p$ with $p \geqq 3$, so "almost every" ternary simplicial agiform is sos.

Let R be a commutative ring, and suppose x is a sum of squares in R; x has *length* k if there exist $y_i \in R$ so that $x = y_1^2 + \ldots + y_k^2$, but x is not a sum of $k-1$ squares in R. As noted after Corollary 4.11, the Hurwitz agiform $G(\underline{c})$ has length $\leqq 3n-4$, when $R = \mathbb{R}[x_1, \ldots, x_n]$, see also [19]. The achievement of these small numbers relies on the fact that a psd binary form is a sum of two squares, which need not be binomials. By Theorem 4.4, the estimate on the size of a $\mathfrak{U}$-mediated set containing $\underline{w}$ gives an upper bound on the length. How good is this estimate in general? That is, how many fewer squares arise when we allow more terms? Can these "bigger" squares be incorporated into our algorithmic structure?

If p is a psd form, let $Y(p) = \{k \in \mathbb{Z} : p(x_1^k, \ldots, x_n^k) \text{ is sos}\}$. In [20] we study $Y(f)$ for simplicial agiforms; $Y(M_n) = \{k : k \geqq n/2\}$. If $\underline{w}$ is even and $f = f(\mathfrak{U}, \underline{w})$ is simplicial, then $Y(f) = \{k : k \geqq k_0\}$ for some k_0; for any agiform, $k \in Y(f)$ implies $k+2 \in Y(f)$. What can be said about $L(f, k)$, the length of $f(x_1^k, \ldots, x_n^k)$, beyond the obvious fact that $L(f, rk) \leqq L(f, k)$? The "Horn form" (see also [3, p. 396]) is an even quartic form $H(x_1, \ldots, x_5)$; we prove that $Y(H) = \phi$. Is $Y(p) = \phi$ possible for a psd form p in three or four variables?

In Sect. 6, circulant matrices make it easy to compute $C(\mathfrak{U})$ and $E(\mathfrak{U})$. Can we find other interesting trellises (and simplicial agiforms) using circulant matrices?

The condition of $\underline{w}$-thinness seems rather mysterious; does it have a more natural reformulation? Theorem 8.12 suggests the question of presenting the psd forms in k effective variables with $k+2$ terms and the determination of the resulting extremal forms. The form E, from Example 9.9, with $k=4$, might play the role of M for the "almost-agiforms".

There are several open algebraic questions about agiforms of a different nature. A celebrated theorem of Cassels, Ellison, and Pfister [1] states that $h(x, y) = M(x, y, 1)$ has length four when $R = \mathbb{R}(x, y)$. (By Artin's solution to Hilbert's Seventeenth Problem, every psd form is a sum of squares of *rational* functions.) What can be said about the length of the dehomogenizations of other simplicial agiforms in three variables? What about the extremal ones? These questions are probably very hard; [1] uses elliptic curves. It is known that the maximal length of an element in $\mathbb{R}(x_1, \ldots, x_n)$ is at most 2^n. Can this bound be achieved by the dehomogenization of another agiform?

Another question involves irreducibility. Binomial squares are reducible. It follows from the identity

$$a^3 + b^3 + c^3 - 3abc = (a+b+c)(a^2 + b^2 + c^2 - ab - ac - bc)$$

[c.f. (1.11)], that $f(\mathfrak{U}, \underline{w})$ is reducible when $\mathfrak{U} = \{3\underline{u}_1, 3\underline{u}_2, 3\underline{u}_3\}$ and $\underline{w} = \underline{u}_1 + \underline{u}_2 + \underline{u}_3$. Are these the only reducible simplicial agiforms?

What can be said about monomial substitutions into agiforms? The identities $z^2 M(x, y, z) = Q(xy, xz, yz, z^2)$ and $x^4 z^2 S(x, y, z) = M(x^2, yz, xz)$, and others, are used in [4] to simplify the proofs of extremality. Is there an efficient way to find formulas such as these?

References

1. Cassels, J.W.S., Ellison, W.J., Pfister, A.: On sums of squares and on elliptic curves over function fields. J. Number Theory **3**, 125–149 (1971), (MR 45 #1863)
2. Choi, M.D., Knebusch, M., Lam, T.Y., Reznick, B.: Transversal zeros and positive semidefinite forms. In: Colliot-Thélène, J.-L., Coste, M., Mahé, L., Roy, M.-F. (eds.) Géométrie algébrique réelle et formes quadratiques. Proceedings, Rennes 1981. (Lecture Notes Mathematics Vol. 959, pp. 273–298) Berlin Heidelberg New York: Springer 1982, (MR 84b.10027)
3. Choi, M.D., Lam, T.Y.: An old question of Hilbert. In: Orzech, G. (ed.) Proceedings Queens Univ. Quadratic Forms Conference 1976, Queens Pap. Pure Appl. Math. **46**, 385–405, (1977) (MR 58 #16503)
4. Choi, M.D., Lam, T.Y.: Extremal positive semidefinite forms. Math. Ann. **231**, 1–18 (1977), (MR 58 #16512)
5. Choi, M.D., Lam, T.Y., Reznick, B.: Real zeros of positive semidefinite forms, I. Math. Z. **171**, 1–26 (1980), (MR 81d.10012)
6. Choi, M.D., Lam, T.Y., Reznick, B.: Even symmetric sextics. Math. Z. **195**, 559–580 (1987)
7. Choi, M.D., Lam, T.Y., Reznick, B.: A combinatorial theory for sums of squares. In preparation
8. Davis, P.J.: Circulant matrices. New York: Wiley 1979 (MR 81a:15053)
9. Handelman, D.: Positive polynomials and product type actions of compact groups. Mem. Am. Math. Soc. **54** (1985), No. 320, (MR 86h:46091)
10. Handelman, D.: Integral body-building in $\mathbb{R}^3$. J. Geom. **27**, 140–152 (1986), (MR 87k:52032)
11. Handelman, D.: Positive polynomials, convex integral polytopes and a random walk problem. Lecture Notes in Mathematics Vol. 1282, Berlin Heidelberg New York: Springer 1987

12. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge: Cambridge University Press 1967
13. Hilbert, D.: Über die Darstellung definiter Formen als Summe von Formenquadraten, Math. Ann. **32**, 342–350 (1888) (=Ges. Abh., **2**, 154–161)
14. Hurwitz, A.: Über den Vergleich des arithmetischen und des geometrischen Mittels. J. Reine Angew. Math. **108**, 266–268 (1891) (=Werke, **2**, 505–507)
15. Lax, A., Lax, P.: On sums of squares. Linear Algebra Appl. **20**, 71–75 (1978), (MR 57 #3074)
16. Motzkin, T.S.: The arithmetic-geometric inequality. In: Shisha, O. (ed.) Inequalities, pp. 205–224, New York: Academic Press 1967, (=Selected Papers, pp. 203–222), (MR 36 #6569)
17. Reznick, B.: Extremal psd forms with few terms. Duke Math. J. **45**, 363–374 (1978), (MR 58 #511)
18. Reznick, B.: Lattice point simplices. Discr. Math. **60**, 219–242 (1986), (MR 87i.52022)
19. Reznick, B.: A quantitative version of Hurwitz' theorem on the arithmetic-geometric inequality. J. Reine Angew. Math. **377**, 108–112 (1987)
20. Reznick, B.: Midpoint polytopes and the map $x_i \to x_i^k$. In preparation
21. Robinson, R.M.: Some definite polynomials which are not sums of squares of real polynomials. In: Selected questions of algebra and logic. Izdat. "Nauka" Sibirsk. Otdel. Novosibirsk, pp. 264–282, 1973 [Abstract in Notices A.M.S. **16**, 554 (1969)], (MR 49 #2647)
22. Schmüdgen, K.: An example of a positive polynomial which is not a sum of squares of polynomials. A positive, but not strongly positive functional. Math. Nachr. **88**, 385–390 (1979), (MR 81b.12024)

Received October 1, 1987; in revised form May 30, 1988