Math. Z. 220, 75–97 (1995)

Mathematische
Zeitschrift

Uniform denominators in Hilbert's Seventeenth Problem

Bruce Reznick*

University of Illinois, 1409 W. Green Street, Urbana, IL 61801, USA (e-mail: reznick@math.uiuc.edu)

Received 11 August 1993; in final version from 20 February 1994

1 Introduction and overview

In 1888, Hilbert [H4] proved that if $n \geq 3$, then there exist real positive semidefinite forms $p = p(x_1, \dots, x_n)$ which cannot be written as a sum of squares of forms. He also proved that for $n = 3$ such a form is always a sum of squares of rational functions. Hilbert's Seventeenth Problem asked whether a positive semidefinite form in any number of variables must be a sum of squares of rational functions. In the 1920s, Artin solved Hilbert's Seventeenth Problem in the affirmative by using the Artin-Schreier theory of real fields. This proof was not constructive.

A few years later, Pólya [P1] presented a concrete proof in one special case: if p is both positive definite and even, then for sufficiently large r, $p \cdot (\sum x_i^2)^r$ has positive coefficients, and so is *per se* a sum of squares of monomials. This fact implies that p is a sum of squares of rational functions with common denominator $(\sum x_i^2)^{r/2}$. In 1940, Habicht [H1] used Pólya's result to write an arbitrary positive definite form p as a quotient of two sums of squares of monomials. It follows from his proof that p is a sum of squares of rational functions with positive definite denominator. In each case, if p has rational coefficients, then so do the monomials. (See [H2, pp.57–59, 300–304].) These results are stronger than Artin's in specifying the nature of the denominators, but weaker in that they only apply to real positive definite forms.

The restriction to positive definite forms is necessary. There exist positive semidefinite forms p which have the remarkable property that, in any representation $p = \sum_k \phi_k^2$, where $\phi_k = f_k / g_k$ is a rational function, each form g_k must have a specified non-trivial zero. The existence of these so-called "bad points" insures that $p \cdot (\sum x_i^2)^r$ can never be a sum of squares of forms for *any* r. Habicht's theorem implies that no positive definite form can have a bad point. According

* The author was supported in part by NSF Grant DMS 92-04126

to Delzell's thesis [D3], bad points were first noted by E. G. Straus in an unpublished 1956 letter to G. Kreisel. Delzell's thesis contains an extensive history of the subject (see chapter V); this contains much material unpublished elsewhere. Another subject with an extensive history is that of "continuous" solutions to Hilbert's 17th problem over a real closed field R. These are formulas of the shape $p(x) = \sum_j \lambda_j(p)\phi_{j,p}^2$, where $\phi_{j,p} = f_{j,p}/g_{j,p}$, which hold for every positive semidefinite form p of fixed degree, under the assumptions that $0 \le \lambda_j(p) \in R$, $f_{j,p}(x)$ and $g_{j,p}(x)$ belong to $R[x_1, \dots, x_n]$ and $\lambda_j(p)$, $f_{j,p}$ and $g_{j,p}$ depend continuously on p. Such formulas exist in which the coefficients of the parameters have the shape $\sup_k \inf_\ell h_{k,\ell}$ for polynomials $h_{k,\ell}$ in the coefficients of p. See [D4, D5, D6, D7] for recent results on this subject.

In 1981, Becker ([B1, B2]) extended the Artin theory, giving necessary and sufficient conditions for a rational function p over a formally real field to be a sum of $2k$-th powers of rational functions. Roughly speaking, p must be psd, its degree must be a multiple of $2k$ and all "zeros" must have "$2k$-th order". A concrete application [B2, p.144] is that for all $k \ge 1$, there exist $0 < \lambda_{j,k} \in \mathbf{Q}$ and polynomials $f_{j,k}$ and $g_{j,k}$ in $\mathbf{Q}[t]$ such that

$$B(t) = \frac{1+t^2}{2+t^2} = \sum_j \lambda_{j,k} \left(\frac{f_{j,k}(t)}{g_{j,k}(t)}\right)^{2k}. \tag{1.1}$$

Explicit versions of (1.1) are known for $k = 1, 2$, and there has been interest in finding them for all k. Formula (1.3) is such an expression, but one in which $f_{j,k}, g_{j,k} \in \mathbf{R}[t]$.

One can deduce from recent work of Becker and Powers [B3] that there is an identity (1.1) in which each $g_{j,k}$ is positive definite. Schmid has also recently shown [S1, Cor. 4.1] that if f and g are real positive definite polynomials in one variable with the same degree, then an expression like (1.1) holds (with $(f/g)(t)$ replacing $B(t)$), in which $f_{j,k}$ and $g_{j,k}$ are definite polynomials of the *same* degree.

Here is an overview of the paper.

The author's memoir [R1] developed a set of notations for forms in several variables over a field K of characteristic 0; in this paper, $K \subseteq \mathbf{R}$. A brief review is given in section two. Let $H_d(K^n)$ denote the vector space of forms in $K[x_1, \cdots, x_n]$ of degree d. This usage of n and d is fixed and m will always denote an even integer. There are three closed convex cones in $H_m(\mathbf{R}^n)$ of particular interest: the cone of positive semidefinite forms, denoted $P_{n,m}$, the cone of sums of squares of forms, denoted $\Sigma_{n,m}$, and the cone of sums of m-th powers of linear forms, denoted $Q_{n,m}$. The solution to Hilbert's Seventeenth Problem amounts to the assertion that, if $p \in P_{n,m}$, then there exists $h \in H_d(\mathbf{R}^n)$ so that $h^2 p \in \Sigma_{n,m+2d}$. The main result of this paper is that, if $p \in P_{n,m}$ is positive definite, then for sufficiently large d, we can take $h = (\sum_i x_i^2)^r$ and $h^2 p \in Q_{n,m+4r} \subset P_{n,m+4r}$.

Let $G_n(x_1, \dots, x_n) = x_1^2 + \cdots + x_n^2$. As part of his solution of Waring's Problem, Hilbert [H5] established that $G_n^r \in Q_{n,2r}$ for all (n, r); in fact, G_n^r can be written as a sum of $2r$-th powers of linear forms with rational coefficients (see Proposition

2.6 and Theorems 6.9 and 6.12). Families of such "rational" representations are known explicitly for $r = 1, 2$, and for a few other (n, r). No general formula is known which applies to all (n, r), although it is not difficult to find specific examples; see for example Dickson [D9, pp.717–724], or [R1, §8,9]. (Hausdorff [H3] gave an explicit representation of G_n^r as a sum of $2r$-th powers of linear forms with real coefficients involving the roots of the Hermite polynomials.) Our analysis of (1.1) will require a trigonometric representation of $(x^2 + y^2)^r$ as a sum of $2r$-th powers (see (5.9)), which is not rational (except accidentally, for $r = 1, 3$).

Associated to a form $g \in H_d(\mathbf{R}^n)$ is the d-th order differential operator $g(D)$, defined by replacing each occurrence of x_j with $\frac{\partial}{\partial x_j}$. (For example, $G_n(D)$ is the Laplacian $\Delta = \sum \frac{\partial^2}{\partial x_j^2}$.) In the 19th century, Sylvester developed (and Clifford exploited [C2, p.119]) the method of "contravariant differentiation". We use this to give a simple representation for $g(D)$ applied to a sum of r-th powers (Proposition 2.8):

$$g(D)\sum_k(\alpha_{k1}x_1+\cdots+\alpha_{kn}x_n)^r = \frac{r!}{(r-d)!}\sum_k g(\alpha_{k1},\cdots,\alpha_{kn})(\alpha_{k1}x_1+\cdots+\alpha_{kn}x_n)^{r-d}.$$

An immediate consequence of these ideas is the first main result.

First Theorem (see Theorem 3.1) *If $h \in Q_{n,m+m'}$ and $f \in P_{n,m'}$, then $f(D)h \in Q_{n,m} \subset \Sigma_{n,m}$.*

We apply the First Theorem to $h = G_n^s \in Q_{n,2s}$. It is easy to see that if $f \in H_d(\mathbf{R}^n)$ and $s \ge d$, then G_n^{s-d} must be a factor of $f(D)G_n^s$, and the quotient must also have degree d. Define $\Phi_s(f) \in H_d(\mathbf{R}^n)$ by $f(D)G_n^s = \Phi_s(f)G_n^{s-d}$. If $f \in H_m(K^n)$ is psd, then by the First Theorem, $\Phi_s(f)G_n^{s-m} \in Q_{n,2s-m}$; in fact, it is a non-negative K-linear combination of the $(2s-m)$-th powers of linear forms whose $2s$-th powers sum to G_n^s.

We give explicit formulas for $\Phi_s(p)$ and $\Phi_s^{-1}(p)$ for $p \in H_d(\mathbf{R}^n)$ (see Theorems 3.7 and 3.9), written as linear combinations of $\{\Delta^k(p)G_n^k\}$, where Δ^k denotes the k-th iterated Laplacian and $\Delta^0(p) = p$. (These are finite sums, because $\Delta^k(p) = 0$ for $k > d/2$.) It turns out that a suitably normalized version of Φ_s converges to the identity. If p is positive definite, then for sufficiently large s, $\Phi_s^{-1}(p)$ is also positive definite by Theorem 3.11. The proofs of Theorems 3.7, 3.9 and 3.11 are deferred until the fourth section. For $0 \ne p \in P_{n,m}$, we need a measure of how close p is to having a non-trivial zero. Let

$$\epsilon(p) = \frac{\inf\{p(u) : u \in S^{n-1}\}}{\sup\{p(u) : u \in S^{n-1}\}}. \tag{1.2}$$

Second Theorem (see Theorem 3.12) *Suppose $p \in H_m(K^n)$ is positive definite. If $r \ge \frac{nm(m-1)}{(4\log 2)\epsilon(p)} - \frac{n+m}{2}$, then pG_n^r is a non-negative K-linear combination of a set of $(m+2r)$-th powers of linear forms in $\mathbf{Q}[x_1,\ldots,x_n]$. This set depends only on n, m and r.*

As noted earlier, Pólya proved a similar but weaker result: if p is an even positive definite form, then pG_n^r has positive coefficients for r sufficiently large. The condition on K also carries over; see [H2, p.304] for $K = \mathbf{Q}$. Pólya did not provide estimates for r. This has been done very recently by de Loera and Santos [D2], who also describe some new algorithms for implementing the representations. Interestingly, their bound on r also depends on $\epsilon(p)^{-1}$.

Let $P_{n,m}^{(\epsilon)}$ be the set of $p \in P_{n,m}$ so that $\epsilon(p) \geq \epsilon$; $P_{n,m} = \bigcup_{\epsilon \geq 0} P_{n,m}^{(\epsilon)}$. For each $\epsilon > 0$, the Second Theorem implies that if $p \in P_{n,m}^{(\epsilon)}$, $r \geq \frac{nm(m-1)}{(4 \log 2)\epsilon} - \frac{n+m}{2}$ and $G_n^{m+r} = \sum (\alpha_k \cdot)^{2m+2r}$, then $pG_n^r = \sum \lambda_k(p)(\alpha_k \cdot)^{m+2r}$, where $\lambda_k(p) \geq 0$ is linear in p. Such a representation clearly cannot hold over all of $P_{n,m}$; see Corollary 3.18 below.

The Second Theorem gives new, concrete information about representations as a sum of $2k$-th powers of rational functions. The following result (without the specification of the denominators) can be given an abstract proof using Becker's theory.

Third Theorem (see Theorems 3.15 and 3.16) *If $p \in K[x_1, \cdots, x_n]$ is a positive definite form of degree $m = 2kt$, then p is a non-negative K-linear combination of $2k$-th powers of rational functions in $\mathbf{Q}[x_1, \cdots, x_n]$ whose denominators are powers of G_n. If p and q are positive definite forms in $K[x_1, \cdots, x_n]$ and the degree of the rational function p/q is a multiple of $2k$, then p/q is a non-negative K-linear combination of $2k$-th powers of rational functions whose numerators are in $\mathbf{Q}[x_1, \cdots, x_n]$ and whose denominators are products of powers of G_n and q.*

The Second Theorem also contains the hard part of a curious equivalence (Corollary 3.18): $p \in P_{n,m}$ is positive definite if and only if $pG_n^r \in Q_{n,m+2r}$ for some $r \geq 1$. This is not a trivial consequence of the existence of bad points: there exist positive semidefinite forms p which are not definite and for which pG_n is a sum of squares of forms. (See for example [R3, p.273] and [C1, p.579].) For that matter, we show in section five that there exist positive definite forms p so that $\Phi_{r+m}^{-1}(p)$ is not psd, but $pG_n^r \in Q_{n,m+2r}$.

The fourth section contains proofs of the formulas needed in section three. The formula for $\Phi_s(f)$ follows from an old theorem used by Hobson in his studies of spherical harmonics. The formula for $\Phi_s^{-1}(f)$ requires some tricky but elementary computations. In order to give the estimate on s for which $\Phi_s^{-1}(f)$ is psd, we need a simple, but apparently new, bound on the $L_\infty(S^{n-1})$ norm of $\Delta : H_d(\mathbf{R}^n) \to H_{d-2}(\mathbf{R}^n)$.

Section five gives applications of the Second and Third Theorems to quadratic forms. We generalize $B(t)$ by considering quadratic forms in n variables and their dehomogenizations. We show (Corollary 5.3) that if $a > 0$, $b > 0$ and $2k + 1 \geq \frac{a}{b} \geq \frac{n}{n+2k}$, then $\frac{a+t_1^2+\ldots+t_n^2}{b+t_1^2+\ldots+t_n^2}$ is a sum of $\binom{n+4k}{n}$ $2k$-th powers of rational functions with denominator $b + t_1^2 + \ldots + t_n^2$.

Fix $k \geq 1$ and let $L_j(x,y) = (\cos\frac{j\pi}{k+2})x + (\sin\frac{j\pi}{k+2})y$ and $\lambda_j = 3k - (k+1)\cos(\frac{2j\pi}{k+2})$ for $0 \leq j \leq k+1$. Using the known trigonometric identity (5.9) for $(x^2+y^2)^s$ and the proof of Corollary 5.3, we obtain:

$$(1.3) \qquad B(t) = \frac{1+t^2}{2+t^2} = \frac{2^{4k-2}}{k(k+2)^2\binom{2k}{k}^2} \sum_{i=0}^{k+1}\sum_{j=0}^{k+1} \lambda_j \left(\frac{L_i(\sqrt{2},t)L_j(\sqrt{2},t)}{2+t^2} \right)^{2k}.$$

Although (1.3) gives $B(t)$ as a sum of $2k$-th powers in $\mathbf{R}(t)$, the summands are not in $\mathbf{Q}(t)$. Such a representation cannot yet be found by our methods, because there is no known analogue to (5.9) with coefficients in $\mathbf{Q}$.

Section six presents a self-contained proof of Hilbert's theorem about G_n^r and more, mostly using the ideas of [R1]. We give an integral representation for $[f, G_n^r]$ which implies that G_n^r is interior to $Q_{n,2r}$ (a result essentially due to Hilbert). In fact, $\Phi_s(f)$ can be given an integral interpretation, and if $0 \not\equiv f \in P_{n,m}$ and $r \geq 0$, then $\Phi_{m+r}(f)G_n^r$ is interior to $Q_{n,m+2r}$ (see Theorem 6.9). Finally, we show that any $h \in H_{2r}(K^n)$ in the interior of $Q_{n,2r}$ is a non-negative K-linear combination of $\binom{n+2r-1}{n-1}$ $2r$-th powers of linear forms with coefficients in $\mathbf{Q}$ (Theorem 6.12). Theorem 6.12 may fail for h on the boundary of $Q_{n,m}$: we have shown [R1, p.137] that for $r \geq 2$, $(x+\sqrt{2}y)^{2r} + (x-\sqrt{2}y)^{2r} \in \mathbf{Q}[x,y]$ is not a non-negative $\mathbf{Q}$-linear combination of $2r$-th powers of linear forms in $\mathbf{Q}[x,y]$.

I am happy to acknowledge much advice and many suggestions. Vicki Powers kept asking me about $B(t)$ until I finally realized how interesting it is; the methods in this paper were originally developed for the study of (1.1). Chip Delzell patiently answered my questions about the continuous versions of Hilbert's Seventeenth Problem, and reminded me about bad points, both grammatical and algebraic. Alex Prestel saved me from some rashly incorrect historical assertions. Eberhard Becker helped sharpen the focus of the paper on its main results and simplified some of the proofs. These colleagues also suffered through an earlier (and much more oblique) version of Theorem 3.11. Finally, Norm Levenberg and Doron Zeilberger gave me the valuable references [D8] and [S2], respectively.

2 Preliminary material

In this section we present some basic notations and facts about real forms in several variables. For the most part, this material can also be found in [R1, §1,2,3].

Let K be a field of characteristic 0, and let $H_d(K^n)$ denote the set of homogeneous polynomials in $K[x_1, \cdots, x_n]$ with degree d. (In this paper, K will always be a subfield of $\mathbf{R}$.) The index set for monomials in $H_d(K^n)$ consists of n-tuples of non-negative integers:

$$\mathcal{I}(n,d) = \left\{ i = (i_1, \ldots, i_n) : \sum_{k=1}^{n} i_k = d \right\}.$$

Write $N(n,d) = \binom{n+d-1}{n-1} = |\mathcal{I}(n,d)|$ and for $i \in \mathcal{I}(n,d)$, let $c(i) = \frac{d!}{i_1! \cdots i_n!}$ be the associated multinomial coefficient. The multinomial abbreviation u^i means $u_1^{i_1} \ldots u_n^{i_n}$, where u may be an n-tuple of constants or variables. Every $f \in H_d(K^n)$ can be written as

$$f(x_1,\dots,x_n)=\sum_{i\in\mathcal{I}(n,d)} c(i)a(f;i)x^i. \tag{2.1}$$

The identification of f with the $N(n,d)$-tuple $(a(f;i))$ shows that $H_d(K^n)\approx K^{N(n,d)}$ as a vector space. For $\alpha\in K^n$, we define $(\alpha\cdot)^d\in H_d(K^n)$ by

$$(\alpha\cdot)^d(x)=\left(\sum_{k=1}^n \alpha_k x_k\right)^d=\sum_{i\in\mathcal{I}(n,d)} c(i)\alpha^i x^i. \tag{2.2}$$

We define a fundamental symmetric bilinear form. For p and q in $H_d(\mathbf{R}^n)$, let

$$[p,q]=\sum_{i\in\mathcal{I}(n,d)} c(i)a(p;i)a(q;i). \tag{2.3}$$

For $i\in\mathcal{I}(n,d)$, let $D^i=\prod(\frac{\partial}{\partial x_k})^{i_k}$ and define $f(D)=\sum c(i)a(f;i)D^i$ to be the d-th order differential operator associated to $f\in H_d(\mathbf{R}^n)$. For any forms f and g (possibly of different degree), $(fg)(D)=f(D)g(D)=g(D)f(D)$ since the $\frac{\partial}{\partial x_k}$'s commute.

Lemma 2.4 *If $p,\ q\in H_d(\mathbf{R}^n)$, $g\in H_e(\mathbf{R}^n)$, $h\in H_{d+e}(\mathbf{R}^n)$, $\alpha\in\mathbf{R}^n$ and $i,j\in\mathcal{I}(n,d)$, then: (i) $[p,(\alpha\cdot)^d]=p(\alpha)$; (ii) $[p,x^i]=a(p;i)$; (iii) $d![p,q]=p(D)q$; (iv) $d![p,g(D)h]=(d+e)![gp,h]$.*

Proof. The first two formulas follow from (2.1), (2.2) and (2.3), since $a((\alpha\cdot)^d;i)=\alpha^i$ and $a(x^i;j)=\frac{1}{c(i)}$ if $i=j$, and 0 otherwise. Since both sides of (iii) are bilinear in p and q, it suffices to prove (iii) for monomials $p=c(i)x^i$, $q=x^j$. In this case, $[p,q]=0$ if $i\neq j$ and $[c(i)x^i,x^i]=1$ by (i). On the other hand,

$$\begin{aligned} c(i)D^ix^j &= c(i)\prod_{k=1}^n \frac{\partial^{i_k}(x_k^{j_k})}{\partial x_k^{i_k}} \\ &= \frac{d!}{i_1!\cdots i_n!}\prod_{k=1}^n \big(j_k\cdot(j_k-1)\cdots(j_k-(i_k-1))\big)x_k^{j_k-i_k}. \end{aligned} \tag{2.5}$$

It is easy to see from (2.5) that $c(i)D^ix^i=d!$. If $i\neq j$, then $\sum i_k=\sum j_k=d$ implies that $i_\ell>j_\ell$ for some ℓ, and so $c(i)D^ix^j=0$. Thus (iii) is true. Finally, two applications of (iii) give $(d+e)![gp,h]=((gp)(D))h=p(D)(g(D)h)=d![p,g(D)h]$, proving (iv). □

Note that Lemma 2.4(ii) implies that the bilinear product is non-degenerate: if $[p_1,q]=[p_2,q]$ for $p_1,p_2\in H_d(\mathbf{R}^n)$ and all $q\in H_d(\mathbf{R}^n)$, then $p_1=p_2$.

A set $C\in\mathbf{R}^N$ is a *closed convex cone* if it is closed under addition, multiplication by positive reals and is also closed topologically. Similarly, a set of forms $\mathcal{B}\subset H_d(\mathbf{R}^n)$ is a *closed convex cone in* $H_d(\mathbf{R}^n)$ if it is closed under addition, multiplication by positive reals and is also closed topologically. (Here, $f^{(j)}\to f$ if $a(f^{(j)};i)\to a(f;i)$ for all $i\in\mathcal{I}(n,d)$.) If $\mathcal{B}$ is a closed convex cone in $H_d(\mathbf{R}^n)$, then $\{(a(f;i)):f\in\mathcal{B}\}$ is a closed convex cone in $\mathbf{R}^{N(n,d)}$ and vice versa.

A form $p \in H_d(\mathbf{R}^n)$ is called *positive semidefinite* (or *psd*) if $p(x) \ge 0$ for every $x \in \mathbf{R}^n$. A nonzero psd form must have even degree m. The set of all psd forms in $H_m(\mathbf{R}^n)$ is denoted by $P_{n,m}$. It is easy to see that $P_{n,m}$ is closed under addition and multiplication by positive reals, and that if $\{p^{(j)}\}$ is a sequence in $P_{n,m}$ which converges to p, then $p \in P_{n,m}$. Thus, $P_{n,m}$ is a closed convex cone in $H_m(\mathbf{R}^n)$. A psd form p is *positive definite* (or *pd*) if $p(x) = 0$ only for $x = 0$. Equivalently, $p \in H_m(\mathbf{R}^n)$ is pd if there exists $\nu > 0$ so that $p(u) \ge \nu$ for all u on the unit sphere S^{n-1}. Suppose $p^{(j)} \to p$ and p is pd. Then $p^{(j)}$ is pd for j sufficiently large. (Look at $\inf(p^{(j)})$ on the (compact) set S^{n-1}.) Thus, the pd forms in $H_m(\mathbf{R}^n)$ comprise the interior of the cone $P_{n,m}$; the boundary of $P_{n,m}$ consists of the psd forms which are not pd; that is, those psd forms with non-trivial zeros.

A form $p \in H_m(\mathbf{R}^n)$ is called a *sum of m-th powers* if there exist $\alpha_k \in \mathbf{R}^n$, $1 \le k \le T$, so that $p = \sum_{k=1}^{T} (\alpha_k \cdot)^m$; that is, if p is a sum of m-th powers of linear forms. There is no *a priori* upper bound on T, but Proposition 6.1 implies that $T \le N(n, m)$. The set of all sums of m-th powers in $H_m(\mathbf{R}^n)$ is denoted by $Q_{n,m}$. It follows from Proposition 6.2 that $Q_{n,m}$ is a closed convex cone. If $\lambda_k \ge 0$ and $\beta_k = \lambda_k^{1/m} \alpha_k \in \mathbf{R}^n$, then $\sum_k \lambda_k (\alpha_k \cdot)^m = \sum_k (\beta_k \cdot)^m$: every non-negative $\mathbf{R}$-linear combination of a sum of m-th powers is in $Q_{n,m}$.

The following proposition is central to the paper, and is a special case of a more general result proved at the end of the paper; see Theorems 6.9 and 6.12.

Proposition 2.6 (Hilbert) *For all (n, r), $G_n^r \in Q_{n,2r}$. Furthermore, there exist $\lambda_k = \lambda_k(n, r) \in \mathbf{Q}$, $0 \le \lambda_k(n, r)$, and $\alpha_{k\ell} = \alpha_{k\ell}(n, r) \in \mathbf{Z}$ so that*

$$G_n^r(x_1, \cdots, x_n) = (x_1^2 + \ldots + x_n^2)^r = \sum_{k=1}^{N(n,2r)} \lambda_k (\alpha_{k1} x_1 + \ldots + \alpha_{kn} x_n)^{2r}. \tag{2.7}$$

Contravariant differentiation gives a simple formula for result of a differential operator applied to a sum of powers of linear forms. It is convenient to use the falling factorial notation: $(t)_0 = 1$, $(t)_k = t(t-1)\cdots(t-(k-1))$ for a positive integer k. If n is a positive integer, then $(n)_k = n!/(n-k)! = k!\binom{n}{k}$ if $n \ge k$ and $(n)_k = 0$ if $n < k$.

Proposition 2.8 *If $g \in H_e(\mathbf{R}^n)$ and $h = \sum_k \lambda_k (\alpha_k \cdot)^{d+e} \in H_{d+e}(\mathbf{R}^n)$, then*

$$g(D)h = (d+e)_e \sum_k \lambda_k g(\alpha_k)(\alpha_k \cdot)^d. \tag{2.9}$$

Proof. Since (2.9) is bilinear in g and h, it suffices to prove the formula when $g(x) = x^i$ and $h = (\beta\cdot)^{d+e}$ is a single $(d+e)$-th power; that is, $D^i(\beta\cdot)^{d+e} = (d+e)_e \beta^i (\beta\cdot)^d$. But,

$$\left(\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{i_n} \left(\sum_{j=1}^{n} \beta_j x_j\right)^{d+e} = (d+e)\cdots(d+1)\beta_1^{i_1}\cdots\beta_n^{i_n}\left(\sum_{j=1}^{n} \beta_j x_j\right)^d. \quad \square$$

3 The main theorems

We begin this section with an application of Proposition 2.8.

Theorem 3.1 *Suppose $h \in Q_{n,2s}$ and $q \in P_{n,m}$. Then $q(D)h \in Q_{n,2s-m}$.*

Proof. If $h = \sum_k (\alpha_k \cdot)^{2s}$, then $q(D)h = (2s)_m \sum_k q(\alpha_k)(\alpha_k \cdot)^{2s-m}$ by Proposition 2.8. Since q is psd by hypothesis, $q(\alpha_k) \geq 0$ for all k, and we are done. □

Theorem 3.1 suggests an inverse problem: given a form p, can we find suitable q and h so that $p = q(D)h$? Fortunately, $G_n^s \in Q_{n,2s}$ for all (n, s) by Proposition 2.6, and it turns out that $q(D)G_n^s$ can always be explicitly computed. Before we state the general formulas, we consider low degrees. Note that

$$\text{(3.2)} \qquad \frac{\partial}{\partial x_k}(G_n^s) = \frac{\partial}{\partial x_k}(x_1^2 + \ldots + x_n^2)^s = 2sx_k(x_1^2 + \ldots + x_n^2)^{s-1}.$$

It follows easily from (3.2) that

$$\text{(3.3)} \qquad \begin{aligned} \frac{\partial^2}{\partial x_j \partial x_k}(G_n^s) &= 2s(2s-2)x_j x_k (x_1^2 + \ldots + x_n^2)^{s-2}, \qquad j \neq k; \\ \frac{\partial^2}{\partial x_k^2}(G_n^s) &= 2s(2s-2)x_k^2(x_1^2 + \ldots + x_n^2)^{s-2} + 2s(x_1^2 + \ldots + x_n^2)^{s-1}. \end{aligned}$$

More generally, if $f \in H_d(\mathbf{R}^n)$ and $s \geq d$, then every monomial term in $f(D)$ involves d differentiations, hence G_n^{s-d} divides $f(D)G_n^s$, with quotient of degree $2s - d - 2(s-d) = d$. Accordingly, for $f \in H_d(\mathbf{R}^n)$ and $s \geq d$, we define $\Phi_s(f) \in H_d(\mathbf{R}^n)$ by

$$\text{(3.4)} \qquad f(D)G_n^s = \Phi_s(f)G_n^{s-d}.$$

It is clear from (3.4) that $\Phi_s(= \Phi_{s,n,d})$ is linear. By (3.2) and (3.3), $\Phi_s(x_k) = 2sx_k$, $\Phi_s(x_j x_k) = 4s(s-1)x_j x_k$ if $j \neq k$ and $\Phi_s(x_k^2) = 4s(s-1)x_k^2 + 2sG_n$. Recall that $\Delta = G_n(D) = \sum_i \frac{\partial^2}{\partial x_i^2}$. Since $\Delta(x_k) = \Delta(x_j x_k) = 0$ and $\Delta(x_k^2) = 2$, we have by linearity:

$$\text{(3.5)(i)} \qquad \Phi_s(f) = 2sf, \quad \text{if } f \in H_1(\mathbf{R}^n),$$

$$\text{(3.5)(ii)} \qquad \Phi_s(f) = 4s(s-1)f + s(\Delta f)G_n, \quad \text{if } f \in H_2(\mathbf{R}^n).$$

Since $\Delta(G_n) = 2n$ and $\Phi_s(G_n) = (4s(s-1) + 2ns)G_n = 4s(\frac{n}{2} + s - 1)G_n$, we can invert:

$$\text{(3.6)(i)} \quad \Phi_s^{-1}(f) = \frac{1}{2s}f, \quad \text{if } f \in H_1(\mathbf{R}^n),$$

$$\text{(3.6)(ii)} \quad \Phi_s^{-1}(f) = \frac{1}{4s(s-1)}\left(f - \frac{1}{4(\frac{n}{2} + s - 1)}(\Delta f)G_n\right), \quad \text{if } f \in H_2(\mathbf{R}^n).$$

It turns out that, if p is psd, then $\Phi_s^{-1}(p)$ is psd for s sufficiently large. This result follows from three theorems whose proofs are deferred to the next section. We first obtain closed forms for Φ_s and Φ_s^{-1}; compare (3.8) to (3.5) and (3.10) to (3.6) for $d = 1, 2$. Also observe from (3.7) that $\frac{1}{(s)_d 2^d}\Phi_s(p) \to p$ as $s \to \infty$, since $\frac{(s)_{d-k}}{(s)_d} = \frac{1}{(s-d+k)_k} \to 0$ for $k \geq 1$.

Theorem 3.7 *If $f \in H_d(\mathbf{R}^n)$ and $s \geq d$, then*

$$\Phi_s(p) = \sum_{k \geq 0} \frac{(s)_{d-k} 2^d}{2^{2k} k!} \Delta^k(p) G_n^k. \tag{3.8}$$

Theorem 3.9 *If $f \in H_d(\mathbf{R}^n)$ and $s \geq d$, then*

$$\Phi_s^{-1}(p) = \frac{1}{(s)_d 2^d} \sum_{\ell \geq 0} \frac{(-1)^\ell}{2^{2\ell} \ell! (\frac{n}{2} + s - 1)_\ell} \Delta^\ell(p) G_n^\ell. \tag{3.10}$$

Theorem 3.11 *Suppose $p \in P_{n,m}$ is positive definite and recall the definition of $\epsilon(p)$ from (1.2). If $s \geq \frac{nm(m-1)}{(4 \log 2)\epsilon(p)} - \frac{n-m}{2}$, then $\Phi_s^{-1}(p) \in P_{n,m}$.*

Theorem 3.12 *If $p \in H_m(K^n)$ is positive definite and $r \geq \frac{nm(m-1)}{(4 \log 2)\epsilon(p)} - \frac{n+m}{2}$, then $pG_n^r \in Q_{n,m+2r}$; in fact, pG_n^r is a a non-negative K-linear combination of $(m+2r)$-th powers of linear forms in $\mathbf{Q}[x_1, \cdots, x_n]$.*

Proof. Let $h = G_n^{r+m}$ and $q = \Phi_{r+m}^{-1}(p)$. Taking $s = r+m$ in Theorem 3.11, we see that q is psd, and by Theorem 3.1, $q(D)h = q(D)G_n^{r+m} = \Phi_{r+m}(q)G_n^r = pG_n^r \in Q_{n,m+2r}$. In fact, Proposition 2.6 states that for suitable $0 \leq \lambda_k \in \mathbf{Q}$ and $\alpha_{k,\ell} \in \mathbf{Z}$,

$$G_n^{r+m} = \sum_{k=1}^{N(n,r+m)} \lambda_k (\alpha_{k1}x_1 + \ldots + \alpha_{kn}x_n)^{2r+2m}. \tag{3.13}$$

Apply $q(D)$ to both sides of (3.13) and use Proposition 2.8:

$$pG_n^r = (2m+2r)_m \sum_{k=1}^{N(n,r+m)} \lambda_k q(\alpha_k)(\alpha_{k1}x_1 + \ldots + \alpha_{kn}x_n)^{2r+m}. \tag{3.14}$$

Since the coefficients in (3.10) are rational, $p \in H_m(K^n)$ implies $q = \Phi_{r+m}^{-1}(p) \in H_m(K^n)$, so $q(\alpha_k) \in K$, and (3.14) gives the desired representation. □

Theorem 3.15 *If $p \in H_m(K^n)$ is positive definite and $2k|m$, then $p = \sum_j \lambda_j \left(f_j G_n^{-u}\right)^{2k}$, where $0 \leq \lambda_j \in K$ and $f_j \in \mathbf{Q}[x_1, \ldots, x_n]$.*

Proof. Let $m = 2kt$. By Theorem 3.12, for sufficiently large u, there exist linear forms L_j over $\mathbf{Q}$ and $0 \leq \lambda_j \in K$ so that $pG_n^{2ku} = \sum_j \lambda_j L_j^{2kt+4ku}$, or $p = \sum_j \lambda_j (L_j^{t+2u} G_n^{-u})^{2k}$. □

Theorem 3.16 *Suppose $p \in H_{m_1}(K^n)$ and $q \in H_{m_2}(K^n)$ are both positive definite, and $2k|m_1 - m_2$. Then for sufficiently large v,*

$$\frac{p}{q} = \sum_j \lambda_j \left(\frac{f_j}{qG_n^v} \right)^{2k}$$

for suitable $0 \le \lambda_j \in K$ and $f_j \in \mathbf{Q}[x_1, \ldots, x_n]$.

Proof. Let $m_1 - m_2 = 2kr$ and pick u large enough that $pG_n^{ku-m_1/2}$ and $qG_n^{ku-m_1/2}$ are in $Q_{n,2ku}$ and $Q_{n,2k(u-r)}$ respectively. Then by Theorem 3.12, there are linear forms $L_{j,1}$ and $L_{j,2}$ in $\mathbf{Q}[x_1, \ldots, x_n]$ and $\mu_j, \nu_j > 0$ in K so that:

(3.17)

$$\frac{p}{q} = \frac{pG_n^{ku-m_1/2}}{qG_n^{ku-m_1/2}} = \frac{\sum_j \mu_j (L_{j,1}^u)^{2k}}{\sum_j \nu_j (L_{j,2}^{u-r})^{2k}} = \frac{\left(\sum_j \mu_j (L_{j,1}^u)^{2k}\right) \left(\sum_j \nu_j (L_{j,2}^{u-r})^{2k}\right)^{2k-1}}{\left(\sum_j \nu_j (L_{j,2}^{u-r})^{2k}\right)^{2k}}.$$

The numerator in (3.17) is the product of $2k$ terms, each of which is a sum of $2k$-th powers, and so is a sum of $2k$-th powers. The (common) denominator in (3.17) is $\left(qG_n^{ku-m_1/2}\right)^{2k}$; p/q is a sum of $2k$-th powers of rational functions with denominator $qG_n^{ku-m_1/2}$. □

One might hope that Theorem 3.12 could be extended to psd forms, since Theorem 3.1 only requires that $\Phi_s^{-1}(p) \ge 0$ at any n-tuple $(\alpha_{k1}, \ldots, \alpha_{kn})$ used in (2.7). We shall see in the next section that it is possible for pG_n^{s-m} to be in $Q_{n,2s-m}$, even though $\Phi_s^{-1}(p)$ is not psd. However, the requirement that p be pd (not merely psd) in Theorem 3.12 is essential.

Corollary 3.18 *A form $p \in H_m(\mathbf{R}^n)$ is pd if and only if $pG_n^r \in Q_{n,m+2r}$ for some $r \ge 1$.*

Proof. One direction is Theorem 3.12. For the other, suppose $pG_n^r \in Q_{n,m+2r}$; clearly, p is psd; must it be pd? Suppose $p(\bar{u}) = 0$ for $\bar{u} \in S^{n-1}$; after an orthogonal change of variables (fixing G_n), assume $\bar{u} = (1, 0, \ldots, 0)$. Suppose there exist $\alpha_{k\ell} \in \mathbf{R}$ so that

$$p(x_1, \ldots, x_n)(x_1^2 + \ldots + x_n^2)^r = \sum_{k=1}^{T} (\alpha_{k1}x_1 + \ldots + \alpha_{kn}x_n)^{m+2r}. \tag{3.19}$$

Setting $x = \bar{u}$ in (3.19), we see that $0 = \sum_{k=1}^T \alpha_{k1}^{m+2r}$, hence $\alpha_{k1} = 0$ for all k. Thus x_1 does not appear at all on the right-hand side of (3.19). Let $x_1^a \bar{p}$ be the terms in p which contain x_1 to the maximum exponent; possibly, $a = 0$ and $\bar{p} = p$. Then $x_1^{a+2r}\bar{p}$ appears uncancelled in the left-hand side, a contradiction. □

4 Proofs of Theorems 3.7, 3.9 and 3.11

In this section we give combinatorial proofs of Theorems 3.7 and 3.9 and an analytic proof of Theorem 3.11. First, we prove Hobson's Theorem, which dates to 1892 [H6,H7] and can be found in [H8]. Another recent proof of Hobson's Theorem is given by Strasburger [S2]. We first need a lemma on the iterated Laplacian.

Lemma 4.1 *(i) If $q \in H_d(\mathbf{R}^n)$ and $k \geq 0$, then*

$$\Delta^k(x_\ell q) = 2k\frac{\partial(\Delta^{k-1}(q))}{\partial x_\ell} + x_\ell \Delta^k(q). \tag{4.2}$$

(ii) If $0 \leq k \leq s$, then

$$\Delta^k(G_n^s) = 2^{2k}(s)_k\left(\frac{n}{2} + s - 1\right)_k G_n^{s-k}. \tag{4.3}$$

Proof. If $2k > d + 1$, then degree considerations imply that both sides of (4.2) vanish. Otherwise, (4.2) is valid provided both sides have the same inner product with every $p \in H_{d+1-2k}(\mathbf{R}^n)$. But by the product rule and repeated applications of Lemma 2.4(iv),

$$\begin{aligned}
(d-2k+1)!\left[\Delta^k(x_\ell q),p\right] &= (d+1)!\left[x_\ell q, G_n^k p\right] = d!\left[q, \frac{\partial\left(G_n^k p\right)}{\partial x_\ell}\right] \\
&= d!\left[q, 2kx_\ell G_n^{k-1}p\right] + d!\left[q, G_n^k\frac{\partial p}{\partial x_\ell}\right] \\
&= (d-2k+2)!\left[\Delta^{k-1}(q), 2kx_\ell p\right] + (d-2k)!\left[\Delta^k(q), \frac{\partial p}{\partial x_\ell}\right] \\
&= (d-2k+1)!\left[2k\frac{\partial\left(\Delta^{k-1}(q)\right)}{\partial x_\ell}, p\right] + (d-2k+1)!\left[x_\ell\Delta^k(q), p\right].
\end{aligned}$$

Thus both sides of (4.2) have the same inner product with $(d-2k+1)!p$. For (ii), (3.3) implies that $\Delta(G_n^s) = 4s(\frac{n}{2}+s-1)G_n^{s-1}$, so (4.3) holds for $k = 1$. If (4.3) holds for k, then

$$\begin{aligned}
\Delta^{k+1}(G_n^s) &= \Delta^k(\Delta(G_n^s)) = 4s\Delta^k((\tfrac{n}{2}+s-1)G_n^{s-1}) \\
&= 4s(s-1)_k 2^{2k}(\tfrac{n}{2}+s-1)(\tfrac{n}{2}+s-2)_k G_n^{s-(k+1)} \\
&= 2^{2k+2}(s)_{k+1}(\tfrac{n}{2}+s-1)_{k+1}G_n^{s-(k+1)},
\end{aligned}$$

so (4.3) holds for $k+1$ as well, and the induction is complete. □

Proposition 4.4 (Hobson) *If $p \in H_d(\mathbf{R}^n)$, and F is sufficiently differentiable, then*

$$p(D)F(G_n) = \sum_{k=0}^{\lfloor d/2\rfloor}\frac{2^d}{2^{2k}k!}\Delta^k(p)F^{(d-k)}(G_n). \tag{4.5}$$

Proof. It suffices by linearity to prove (4.5) for monomials $p \in H_d(\mathbf{R}^n)$. It is trivial for $d = 0$; if $d = 1$, let $p = x_j$, amd $\frac{\partial}{\partial x_j} F(G_n) = F'(G_n)\frac{\partial G_n}{\partial x_j} = 2x_j F'(G_n) = \frac{2^1}{2^0 0!}\Delta^0(p)F^{(1-0)}(G_n)$, as desired. Now suppose (4.5) holds for monomials of degree d. If $x^i \in H_{d+1}(\mathbf{R}^n)$, then $x^i = x_\ell x^j$ for some ℓ, where $x^j \in H_d(\mathbf{R}^n)$. Then we have the following sequence of identities, explained after the fact:

$$\begin{aligned}
D^i F(G^n) &= \frac{\partial(D^j F(G_n))}{\partial x_\ell} = \sum_{k\geq 0} \frac{2^d}{2^{2k}k!}\frac{\partial(\Delta^k(x^j)F^{(d-k)}(G_n))}{\partial x_\ell} \\
&= \sum_{k\geq 0} \frac{2^d}{2^{2k}k!}\Delta^k(x^j)\frac{\partial(F^{(d-k)}(G_n))}{\partial x_\ell} + \sum_{k\geq 0}\frac{2^d}{2^{2k}k!}\frac{\partial(\Delta^k(x^j))}{\partial x_\ell}F^{(d-k)}(G_n) \\
&= \sum_{k\geq 0} \frac{2^d}{2^{2k}k!}\Delta^k(x^j)F^{(d+1-k)}(G_n)\frac{\partial G_n}{\partial x_\ell} \\
&\quad + \sum_{k\geq 0}\frac{2^2 2^d(k+1)}{2^{2k+2}(k+1)!}\frac{\partial(\Delta^k(x^j))}{\partial x_\ell}F^{(d+1-(k+1))}(G_n) \\
&= \sum_{k\geq 0} \frac{2^d}{2^{2k}k!}\Delta^k(x^j)F^{(d+1-k)}(G_n)2x_\ell \\
&\quad + \sum_{k\geq 0}\frac{2^{d+1}2k}{2^{2k}k!}\frac{\partial(\Delta^{k-1}(x^j))}{\partial x_\ell}F^{(d+1-k)}(G_n) \\
&= \sum_{k\geq 0}\frac{2^{d+1}}{2^{2k}k!}\left(x_\ell\Delta^k(x^j) + 2k\frac{\partial(\Delta^{k-1}(x^j))}{\partial x_\ell}\right)F^{(d+1-k)}(G_n) \\
&= \sum_{k\geq 0}\frac{2^{d+1}}{2^{2k}k!}\Delta^k(x_\ell x^j)F^{(d+1-k)}(G_n) = \sum_{k\geq 0}\frac{2^{d+1}}{2^{2k}k!}\Delta^k(x^i)F^{(d+1-k)}(G_n).
\end{aligned}$$

The induction hypothesis is used in the first line, and the product rule is used to get to the second line. In the third line, the chain rule is used in the first sum, and the second sum is written in terms of $k+1$, rather than k. In the fourth equation, the index of the second sum is shifted from $k+1$ to k. The summation should now be taken over $k \geq 1$, but the summand is 0 for $k = 0$, so it can be taken over $k \geq 0$. Terms are combined in the fifth equation, and Lemma 4.1(i) is used in the last line to complete the induction. □

Proof of Theorem 3.7 Take $F(z) = z^s$ above; $F^{(d-k)}(z) = (s)_{d-k}z^{s-(d-k)}$, so by (4.5),

$$\Phi_s(p)G_n^{s-d} = p(D)G_n^s = \sum_{k\geq 0}\frac{(s)_{d-k}2^d}{2^{2k}k!}\Delta^k(p)G_n^{s-d+k}. \tag{4.6}$$

After cancellation of G_n^{s-d}, (4.6) becomes (3.8). □

Proof of Theorem 3.9 By applying Φ_s to (3.10) and scaling, we find:

$$(s)_d 2^d p = \sum_{\ell \geq 0} \frac{(-1)^\ell}{2^{2\ell} \ell! (\frac{n}{2} + s - 1)_\ell} \Phi_s(\Delta^\ell(p) G_n^\ell). \tag{4.7}$$

Since $\Delta^\ell(p) G_n^\ell \in H_d(\mathbf{R}^n)$ for each $\ell \leq d/2$, by (3.4) and Lemma 4.1(ii),

$$\begin{aligned}\Phi_s(\Delta^\ell(p)G_n^\ell)G_n^{s-d} &= (\Delta^\ell(p)G_n^\ell)(D)G_n^s = (\Delta^\ell(p))(D)(G_n^\ell)(D)(G_n^s)\\ &= (\Delta^\ell(p))(D)\Delta^\ell(G_n^s) = 2^{2\ell}(s)_\ell \left(\tfrac{n}{2} + s - 1\right)_\ell (\Delta^\ell(p))(D)(G_n^{s-\ell})\\ &= 2^{2\ell}(s)_\ell \left(\tfrac{n}{2} + s - 1\right)_\ell \Phi_{s-\ell}(\Delta^\ell(p))G_n^{(s-\ell)-(d-2\ell)}.\end{aligned} \tag{4.8}$$

Upon cancellation of G_n^{s-d} from both sides of (4.8), we obtain the formula

$$\Phi_s(\Delta^\ell(p)G_n^\ell) = 2^{2\ell}(s)_\ell(\tfrac{n}{2} + s - 1)_\ell \Phi_{s-\ell}(\Delta^\ell(p))G_n^\ell. \tag{4.9}$$

We now plug (4.9) into the right-hand side of (4.7) and then use (3.8), keeping in mind that $\Delta^\ell(p) \in H_{d-2\ell}(\mathbf{R}^n)$:

$$\begin{aligned}\sum_{\ell \geq 0} \frac{(-1)^\ell}{2^{2\ell}\ell!(\frac{n}{2} + s - 1)_\ell} \Phi_s(\Delta^\ell(p)G_n^\ell) &= \sum_{\ell \geq 0} \frac{(-1)^\ell (s)_\ell}{\ell!} \Phi_{s-\ell}(\Delta^\ell(p))G_n^\ell\\ &= \sum_{\ell \geq 0}\sum_{k \geq 0} \frac{(-1)^\ell (s)_\ell}{\ell!} \frac{(s-\ell)_{d-2\ell-k} 2^{d-2\ell}}{2^{2k} k!} \Delta^k(\Delta^\ell p) G_n^k G_n^\ell.\end{aligned} \tag{4.10}$$

Since $(s)_\ell(s - \ell)_{d-2\ell-k} = (s)_{d-\ell-k}$, we can stratify the final sum in (4.10) by $w = k + \ell$:

$$\sum_{w \geq 0} \left(\sum_{\ell=0}^{w} \frac{(-1)^\ell}{w!} \binom{w}{\ell} \right) (s)_{d-w} 2^{d-2w} \Delta^w(p) G_n^w. \tag{4.11}$$

The inner sum in (4.11) vanishes unless $w = 0$, in which case it equals 1, so the entire right-hand side of (4.7) reduces to $(s)_d 2^d p$, which was the assertion. $\square$

The proof of Theorem 3.11 requires a result which is interesting in its own right, but perhaps for a different paper! A key step requires a 1928 inequality due to Szegö which can be found in [D8, p.97], and is a sharpening of Bernstein's Inequality.

Proposition 4.12 (Szegö) *If $H(\theta) = \sum_{k=0}^d a_k \cos(k\theta) + \sum_{k=1}^d b_k \sin(k\theta)$ is a trigonometric polynomial of degree d and $|H(\theta)| \leq M$ for $\theta \in [0, 2\pi]$, then*

$$d^2(H(\theta))^2 + (H'(\theta))^2 \leq d^2 M^2 \quad \text{for } \theta \in [0, 2\pi]. \tag{4.13}$$

Theorem 4.14 *If $f \in H_d(\mathbf{R}^n)$ and $|f(u)| \leq M$ for $u \in S^{n-1}$, then $|\Delta^k f(u)| \leq n^k (d)_{2k} M$ for $u \in S^{n-1}$.*

Proof. If we could show that $|\frac{\partial f}{\partial x_j}(u)| \le dM$, then since $\frac{\partial f}{\partial x_j} \in H_{d-1}(\mathbf{R}^n)$, $|\frac{\partial^2 f}{\partial x_j^2}(u)| \le d(d-1)M = (d)_2 M$, hence $|\Delta f(u)| \le \sum_{j=1}^n |\frac{\partial^2 f}{\partial x_j^2}(u)| \le n(d)_2 M$. Then, $|(\Delta^{k+1} f)(u)| = |\Delta(\Delta^k(f))(u)| \le n(d-2k)_2 n^k (d)_{2k} M = n^{k+1}(d)_{2k+2} M$ by induction, and we would be done.

We first prove for $n = 2$; let $f \in H_d(\mathbf{R}^2)$ and suppose $|f(u_1, u_2)| \le M$ when $u_1^2 + u_2^2 = 1$. We define the d-th order trignometric polynomial $H(\theta) = f(\cos(\theta), \sin(\theta))$. The chain rule and Euler's PDE $df = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2}$, evaluated at $(\cos(\theta), \sin(\theta))$, give

$$\begin{aligned} H'(\theta) &= -\sin(\theta)\tfrac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta)) + \cos(\theta)\tfrac{\partial f}{\partial x_2}(\cos(\theta), \sin(\theta)), \\ dH(\theta) &= \cos(\theta)\tfrac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta)) + \sin(\theta)\tfrac{\partial f}{\partial x_2}(\cos(\theta), \sin(\theta)). \end{aligned} \tag{4.15}$$

We solve for $\frac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta))$ in (4.15):

$$\frac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta)) = d\cos(\theta)H(\theta) - \sin(\theta)H'(\theta). \tag{4.16}$$

Since $|A\cos(\theta) + B\sin(\theta)| \le \sqrt{A^2 + B^2}$, (4.16) and (4.13) imply that

$$\left|\frac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta))\right|^2 \le d^2(H(\theta))^2 + (H'(\theta))^2 \le d^2 M^2, \tag{4.17}$$

so that $|\frac{\partial f}{\partial x_1}(\cos(\theta), \sin(\theta))| \le dM$, and of course, similarly, $|\frac{\partial f}{\partial x_2}(\cos(\theta), \sin(\theta))| \le dM$.

Now suppose $f \in H_d(\mathbf{R}^n)$, $n \ge 3$. For fixed $\xi = (\xi_2, \cdots, \xi_n) \in S^{n-2}$, let $g_\xi(x_1, x_2) = f(x_1, \xi_2 x_2, \cdots, \xi_n x_2)$. Then $g_\xi \in H_d(\mathbf{R}^2)$ and, $u_1^2 + u_2^2 = u_1^2 + (\sum \xi_j^2)u_2^2$, hence $|g_\xi(u_1, u_2)| \le M$. By the last paragraph, $|\frac{\partial g_\xi}{\partial x_1}(u_1, u_2)| \le Md$. Since $\frac{\partial g_\xi}{\partial x_1} = \frac{\partial f}{\partial x_1}$, $|\frac{\partial f}{\partial x_1}(u_1, \xi_2 u_2, \cdots, \xi_n u_2)| \le Md$ for all $(u_1, \xi_2 u_2, \cdots, \xi_n u_2)$. But every point in $u \in S^{n-1}$ can be written in this way, hence $|\frac{\partial f}{\partial x_1}(u)| \le Md$ for $u \in S^{n-1}$, and, similarly, for each $\frac{\partial f}{\partial x_j}$. The argument of the first paragraph completes the proof. □

This bound is not far from the truth. There exist $p_{d,n} \in H_d(\mathbf{R}^n)$ such that $|p_{d,n}(u)| \le 1$ on S^{n-1}, but $\big(\Delta(p_{d,n})\big)(1, 0, \cdots, 0) = \max\{1, n-2\}(d)_2$. We shall discuss this elsewhere.

Proof of Theorem 3.11 Since $\epsilon(p) = \epsilon(\lambda p)$, scale p so that $1 \ge p(u) \ge \epsilon(p)$ for $u \in S^{n-1}$. Write $\epsilon = \epsilon(p)$ and let $\Psi_s(p) = (s)_m 2^m \Phi_s^{-1}(p)$; we show that $\big(\Psi_s(p)\big)(u) \ge 0$. By (3.10),

$$\begin{aligned} \big(\Psi_s(p)\big)(u) &= \big((s)_m 2^m \Phi_s^{-1}(p)\big)(u) \ge \\ & p(u) - \sum_{\ell \ge 1} \frac{1}{2^{2\ell}\ell!(\frac{n}{2} + s - 1)_\ell} |\Delta^\ell(p)(u)| G_n^\ell(u). \end{aligned} \tag{4.18}$$

Since $1 \ge p(u) \ge \epsilon$ and $G_n(u) \equiv 1$, Theorem 4.14 and (4.18) imply that

$$(4.19)\qquad \big(\Psi_s(p)\big)(u) \geq \epsilon - \sum_{\ell \geq 1} \frac{n^\ell (m)_{2\ell} \cdot 1}{2^{2\ell}\ell!(\frac{n}{2}+s-1)_\ell}.$$

The sum in (4.19) is finite because $\Delta^\ell(p) = 0$ for $\ell > \frac{m}{2}$. We have $(m)_{2\ell} \leq (m(m-1))^\ell$ and $(\frac{n}{2}+s-1)_\ell \geq (\frac{n}{2}+s-\ell)^\ell \geq (\frac{n-m}{2}+s)^\ell$. Thus, (4.19) implies

$$(4.20)\qquad \big(\Psi_s(p)\big)(u) \geq \epsilon - \sum_{\ell=1}^{m/2} \frac{1}{\ell!}\left(\frac{nm(m-1)}{2(n-m)+4s}\right)^\ell.$$

Since $s \geq \frac{nm(m-1)}{(4\log 2)\epsilon} - \frac{n-m}{2}$, we have $\frac{nm(m-1)}{2(n-m)+4s} \leq \epsilon \log 2$, and (4.20) implies

$$(4.21)\qquad \big(\Psi_s(p)\big)(u) > \epsilon - \sum_{\ell=1}^{\infty} \frac{(\epsilon\log 2)^\ell}{\ell!} = \epsilon - \big(e^{\epsilon\log 2} - 1\big) = 1 + \epsilon - 2^\epsilon.$$

Let $F(t) = 1 + t - 2^t$. Then $F'' < 0$ and $F(0) = F(1) = 0$, hence $F(t) \geq 0$ on $[0,1]$. In particular, $\epsilon \in [0,1]$, so $1 + \epsilon - 2^\epsilon \geq 0$ and (4.21) completes the proof. □

5 Quadratic examples

This section on quadratic polynomials does not use Theorems 3.7, 3.9 and 3.11.

Lemma 5.1 *If* $f(x_1,\dots,x_n) = \sum_{k=1}^{n} a_k x_k^2$, *then* $\Phi_s^{-1}(f)$ *is psd if and only if* $(2s+n-2)a_k \geq \sum_{j=1}^n a_j$ *for all* k, $1 \leq k \leq n$.

Proof. Observe that $\Delta f = 2\sum_{k=1}^n a_k$, hence by (3.6)(ii),

$$\begin{aligned} 4s(s-1)\Phi_s^{-1}(f) &= f - \frac{1}{4(\frac{n}{2}+s-1)}(\Delta f)G_n = \sum_{k=1}^{n}\left(a_k - \frac{2}{4(\frac{n}{2}+s-1)}\sum_{j=1}^{n} a_j\right)x_k^2 \\ &= \frac{1}{2(\frac{n}{2}+s-1)}\sum_{k=1}^{n}\left((2s+n-2)a_k - \sum_{j=1}^{n} a_j\right)x_j^2. \square \end{aligned}$$

Theorem 5.2 *If* $2r+1 \geq \lambda \geq \frac{n-1}{n+2r-1}$ *and* $p_\lambda(x) = \lambda x_1^2 + \sum_{k=2}^{n} x_k^2$, *then* $p_\lambda G_n^{r-1} \in Q_{n,2r}$.

Proof. If $\Phi_{r+1}^{-1}(p_\lambda)$ is psd, then $\Phi_{r+1}^{-1}(p_\lambda)(D)G_n^{r+1} = p_\lambda G_n^{r-1} \in Q_{n,2r}$ by Theorem 3.1. Since $\sum_j a_j = \lambda + n - 1$ and a_k takes the values 1 and λ, Lemma 5.1 implies that $\Phi_{r+1}^{-1}(p_\lambda)$ is psd if and only if $(n+2r)\lambda \geq n-1+\lambda$ and $n+2r \geq n-1+\lambda$. □

Let's compare Theorem 5.2 with Theorem 3.11. If $0 \le \lambda \le 1$, then clearly $\epsilon(p_\lambda) = \lambda$. By Theorem 3.11, $\Phi_{r+1}^{-1}(p_\lambda)$ is psd provided $r + 1 \ge \frac{2n}{(4\log 2)\lambda} - \frac{n-2}{2}$. By Theorem 5.2, $\Phi_{r+1}^{-1}(p_\lambda)$ is psd precisely when $\lambda \ge \frac{n-1}{n+2r-1}$, or $r \ge \frac{n-1}{2\lambda} - \frac{n-1}{2}$. Theorem 3.11 is only off by a factor of $\log 2$. We show later that $p_\lambda G_2 \in Q_{2,4}$ does not imply $\Phi_2^{-1}(p_\lambda) \in P_{2,2}$.

Corollary 5.3 *Suppose $a > 0$, $b > 0$ and $2k + 1 \ge \frac{a}{b} \ge \frac{n}{n+2k}$. Then there exist quadratic polynomials $\{\xi_i(t_1, \cdots, t_n)\}$, $1 \le i \le N(n+1, 4k)$, so that*

$$\frac{a + t_1^2 + \ldots + t_n^2}{b + t_1^2 + \ldots + t_n^2} = \sum_{i=1}^{N(n+1,4k)} \left(\frac{\xi_i(t_1, \cdots, t_n)}{b + t_1^2 + \ldots + t_n^2} \right)^{2k}. \tag{5.4}$$

Proof. Suppose $2k + 1 \ge \lambda \ge \frac{n}{n+2k}$. Then by Proposition 2.6 and Theorem 5.2, there exist linear forms $\{\beta_j(x_1, \ldots, x_{n+1})\}$ and $\{\gamma_\ell(x_1, \ldots, x_{n+1})\}$ so that

$$(x_1^2 + \ldots + x_{n+1}^2)^k = \sum_{\ell=1}^{N(n+1,2k)} \gamma_\ell^{2k}(x_1, \ldots, x_{n+1}); \tag{5.5)(i}$$

$$\begin{aligned} &(\lambda x_1^2 + x_2^2 + \cdots + x_{n+1}^2)(x_1^2 + \ldots + x_{n+1}^2)^{k-1} \\ &= \sum_{j=1}^{N(n+1,2k+2)} \beta_j^{2k}(x_1, \ldots, x_{n+1}). \end{aligned} \tag{5.5)(ii}$$

Multiply (5.5)(i) and (5.5)(ii):

$$(\lambda x_1^2 + x_2^2 + \cdots + x_{n+1}^2)(x_1^2 + \ldots + x_{n+1}^2)^{2k-1} = \sum_{i=1}^{M} \psi_i^{2k}(x_1, \ldots, x_{n+1}), \tag{5.6}$$

where $M = N(n+1, 2k)N(n+1, 2k+2)$ and the ψ_i's are the (quadratic) pairwise products $\{\beta_j \gamma_\ell\}$. By Carathéodory's Theorem (see Proposition 6.1), a sum of M forms in $H_{4k}(\mathbf{R}^{n+1})$ is a non-negative linear combination of $N(n+1, 4k) < M$ of them. By absorbing constants into the powers, we can rewrite (5.6) as

$$(\lambda x_1^2 + x_2^2 + \cdots + x_{n+1}^2)(x_1^2 + \ldots + x_{n+1}^2)^{2k-1} = \sum_{i=1}^{N(n+1,4k)} \phi_i^{2k}(x_1, \ldots, x_{n+1}), \tag{5.7}$$

Substitute $x_1 = \sqrt{b}$, $x_2 = t_1, \ldots, x_{n+1} = t_n$, $\lambda = \frac{a}{b}$ into (5.7) and write $\phi_i(\sqrt{b}, t_1, \ldots, t_n) = \xi_i(t_1, \ldots, t_n)$:

$$(a + t_1^2 + \ldots + t_n^2)(b + t_1^2 + \ldots + t_n^2)^{2k-1} = \sum_{i=1}^{N(n+1,4k)} \xi_i^{2k}(t_1, \cdots, t_n). \tag{5.8}$$

Finally, divide both sides of (5.8) by $(b + t_1^2 + \ldots + t_n^2)^{2k}$ to obtain (5.4). $\square$

We now apply the argument of Corollary 5.3 in detail to $B(t)$, with the goal of proving (1.3). The first step is to find expressions for $(\frac{1}{2}x^2+y^2)(x^2+y^2)^{k-1}$ and $(x^2+y^2)^k$ as sums of $2k$-th powers of linear forms. This uses a venerable formula which is proved in [R1, Thm. 9.5]. Suppose s and v are positive integers and $v \geq s+1$. Then

$$(5.9) \qquad (x^2+y^2)^s = \frac{2^{2s}}{v\binom{2s}{s}} \sum_{j=0}^{v-1} \left(\cos\left(\frac{j\pi}{v}\right) x + \sin\left(\frac{j\pi}{v}\right) y \right)^{2s}.$$

A pleasant non-trivial illustration of (5.9) occurs for $s=3$ and $v=4$: $(x^2+y^2)^3 = \frac{4}{5}\left(x^6 + \left(\frac{x+y}{\sqrt{2}}\right)^6 + y^6 + \left(\frac{-x+y}{\sqrt{2}}\right)^6\right)$, an identity which can be verified by hand.

Proof of (1.3) Two instances of (5.9) (taking $v = k+2$ and $s = k$ and $k+1$) give:

$$(x^2+y^2)^k =$$

$$(5.10)(\mathrm{i}) \qquad \frac{2^{2k}}{(k+2)\binom{2k}{k}} \sum_{j=0}^{k+1} \left(\cos\left(\frac{j\pi}{k+2}\right) x + \sin\left(\frac{j\pi}{k+2}\right) y \right)^{2k},$$

$$(x^2+y^2)^{k+1}$$

$$(5.10)(\mathrm{ii}) \quad = \frac{2^{2k+2}}{(k+2)\binom{2k+2}{k+1}} \sum_{j=0}^{k+1} \left(\cos\left(\frac{j\pi}{k+2}\right) x + \sin\left(\frac{j\pi}{k+2}\right) y \right)^{2k+2}.$$

Let $L_j(x,y) = (\cos\frac{j\pi}{k+2})x + (\sin\frac{j\pi}{k+2})y$. The proof of Lemma 5.1 gives

$$(5.11) \quad \Phi_{k+1}^{-1}(\tfrac{1}{2}x^2+y^2) = \frac{\frac{1}{2}x^2+y^2}{4k(k+1)} - \frac{(1+2)(x^2+y^2)}{16k(k+1)^2} = \frac{(2k-1)x^2+(4k+1)y^2}{16k(k+1)^2}.$$

Then by (2.9), (5.10)(ii) and (5.11),

$$(5.12) \quad \left(\frac{1}{2}x^2+y^2\right)(x^2+y^2)^{k-1} = \frac{2^{2k+2}}{(k+2)\binom{2k+2}{k+1}}(2k+2)_2$$

$$\times \sum_{j=0}^{k+1} \left(\frac{2k-1}{16k(k+1)^2}\cos^2\left(\frac{j\pi}{k+2}\right) + \frac{4k+1}{16k(k+1)^2}\sin^2\left(\frac{j\pi}{k+2}\right)\right) L_j(x,y)^{2k}.$$

Now multiply (5.10)(i) and (5.12), divide both sides by $(x^2+y^2)^{2k}$ and set $x = \sqrt{2}$ and $y = t$. Algebraic simplification gives (1.3); key substitutions are $\binom{2k+2}{k+1} = \frac{4k+2}{k+1}\binom{2k}{k}$ and $(2k-1)\cos^2\theta + (4k+1)\sin^2\theta = 3k - (k+1)\cos 2\theta$. □

This construction is not designed to minimize the number of summands. We know from Carathéodory's Theorem that $(\frac{1}{2}x^2+y^2)(x^2+y^2)^{2k-1}$ is a non-negative linear combination of $N(2,4k) = 4k+1$ of the $(k+1)^2$ $2k$-th powers of quadratics given in (1.3). However, we don't know *which* ones to choose.

The identity (1.3) is also fundamentally defective, in that $B(t)$ is not given as a sum of $2k$-th powers in $\mathbf{Q}(t)$. The need for $\sqrt{2}$ could be eliminated by writing $(2x^2+y^2)^k$ as in (2.7); this is possible by Theorem 6.12. We can find $h(x,y) = ax^2+by^2$ so that $h(D)(2x^2+y^2)^{k+1} = (x^2+y^2)(2x^2+y^2)^{k-1}$ and proceed as above. Unfortunately, there are no known general explicit formulas for writing any $(\lambda x^2+y^2)^k$, $0<\lambda\in\mathbf{Q}$, as a $\mathbf{Q}$-linear combination of $2k$-th linear forms over $\mathbf{Q}$.

We can use the methods of [R1] (see (1.23), Corollary 3.17 and Theorem 5.1) to find necessary and sufficient conditions under which $(ax^2+by^2)(x^2+y^2)^{r-1}\in Q_{2,2r}$. In the interest of space, we shall summarize the results at the end of this section.

The case $r=2$ can be done from first principles. By scaling, we may assume that $(a,b)=(\lambda,1)$. If $(\lambda x^2+y^2)(x^2+y^2) = \lambda x^4+(\lambda+1)x^2y^2+y^4 = \sum(c_kx+d_ky)^4$, then $\lambda = \sum c_k^4$, $\frac{\lambda+1}{6} = \sum c_k^2d_k^2$ and $1=\sum d_k^4$, hence by the Cauchy-Schwarz inequality, $\lambda \ge (\frac{\lambda+1}{6})^2$, so $\lambda^2-34\lambda+1\le 0$, or $17-12\sqrt{2}\le\lambda\le 17+12\sqrt{2}$. On the other hand,

$$\frac{1}{2}\left(\left((\sqrt{2}\pm 1)x+y\right)^4+\left((\sqrt{2}\pm 1)x-y\right)^4\right) = \left((17\pm 12\sqrt{2})x^2+y^2\right)\left(x^2+y^2\right).$$

Thus, by convexity, $(\lambda x^2+y^2)(x^2+y^2)\in Q_{2,4}$ precisely when $17-12\sqrt{2}\le\lambda\le 17+12\sqrt{2}$. By Lemma 5.1, $\Phi_2^{-1}(\lambda x^2+y^2)$ is psd precisely for $\frac{1}{3}\le\lambda\le 3$. Thus, if λ or λ^{-1} are in $(3, 17+12\sqrt{2}\,]$, then $\Phi_2^{-1}(\lambda x^2+y^2)$ is not psd but $(\lambda x^2+y^2)(x^2+y^2)\in Q_{2,4}$.

It can be shown that $q_\lambda(x,y) = (\lambda x^2+y^2)(x^2+y^2)^{r-1}\in Q_{2,2r}$ for $\lambda\in[M_r^{-1},M_r]$, where M_r is a zero of the catalecticant of q_λ. Theorem 5.2 with $n=2$ implies $M_r\ge 2r+1$. Calculations using Mathematica show that $M_3 = 13+6\sqrt{5}\approx 26.42$, $M_4 = \frac{139+80\sqrt{3}}{11}\approx 25.23$, $M_5\approx 25.64$, $M_6\approx 26.69$, $M_7\approx 28.05$, $M_8\approx 29.60$, $M_9\approx 31.26$, $M_{10}\approx 32.99$, $M_{11}\approx 34.78$ and $M_{12}\approx 36.61$. If $k\ge 13$, then $M_k\ge 2k+1\ge 27$ by Theorem 5.2, hence we can draw the following conclusion.

Proposition 5.13 *Let* $M = M_4 = \frac{139+80\sqrt{3}}{11}$. *If* $a,b>0$ *and* $1/M\le a/b\le M$, *then for all* $k\ge 1$, $\frac{a+t^2}{b+t^2}$ *is a sum of* $2k$*-th powers of rational functions with denominator* $b+t^2$.

6 A self-contained proof of Hilbert's theorem

In this section, we give a self-contained proof of a generalization of Proposition 2.6. This requires more background from [R1]. We begin with Carathéodory's Theorem, which can be traced back to Hilbert [H4] (see [R1, p.27]).

Proposition 6.1 (Carathéodory's Theorem) *Suppose* V *is a vector space over a field* $K\subseteq\mathbf{R}$ *and* $\dim V = N$. *Suppose there exist* $x_k\in V$ *and* $\gamma_k\in K$, $\gamma_k>0$, *so that* $y=\sum_{k=1}^L\gamma_kx_k$. *Then there exist* $\lambda_k\in K$, $\lambda_k\ge 0$, *so that* $y=\sum_{k=1}^N\lambda_kx_{s_k}$.

Proof. There is nothing to prove if $L \le N$, and it suffices by induction to consider $L = N + 1$. Since $\{x_1, \ldots, x_{N+1}\}$ is linearly dependent, there exist $c_k \in K$, not all 0, such that $\sum_k c_k x_k = 0$. Let $\beta_k = c_k/\gamma_k$. We may assume that $0 < |\beta_{N+1}| = \max_k |\beta_k|$. Then

$$y = \sum_{k=1}^{N+1} \gamma_k x_k - \frac{1}{\beta_{N+1}} \sum_{k=1}^{N+1} c_k x_k = \sum_{k=1}^{N+1} \left(1 - \frac{\beta_k}{\beta_{N+1}}\right) \gamma_k x_k = \sum_{k=1}^{N} \left(1 - \frac{\beta_k}{\beta_{N+1}}\right) \gamma_k x_k,$$

and $\lambda_k = (1 - \beta_k/\beta_{N+1})\gamma_k \ge 0$. □

Proposition 6.1, applied to $(K, V) = (\mathbf{R}, H_m(\mathbf{R}^n))$, retrieves Hilbert's result that every $p \in Q_{n,m}$ is a sum of at most $N(n,m)$ m-th powers of linear forms. Applied to $(K, V) = (K, H_m(K^n))$, we see that if $p = \sum \gamma_k (\alpha_k \cdot)^m$, $\alpha_k \in \mathbf{Q}^n$ and $\gamma_k \in K$, then p is a non-negative K-linear combination of $N(n,m)$ m-th powers of linear forms in $\mathbf{Q}[x_1, \ldots, x_n]$.

Proposition 6.2 *For all (n,m), $Q_{n,m}$ is a closed convex cone.*

Proof. We already know that $Q_{n,m}$ is a convex cone. Suppose $p^{(j)} \in Q_{n,m}$ and $p^{(j)} \to p$. By Proposition 6.1, we can write

$$p^{(j)}(x_1, \ldots, x_n) = \sum_{k=1}^{N(n,m)} (\alpha_{k1}^{(j)} x_1 + \cdots + \alpha_{kn}^{(j)} x_n)^m.$$

(The linear forms are over $\mathbf{R}$, so the coefficients can be absorbed into the linear forms.) Let $e_\ell = (0, \cdots, 1, \cdots, 0)$ denote the ℓ-th unit vector, so $p^{(j)}(e_\ell)$, the coefficient of x_ℓ^m, equals $\sum_k \left(\alpha_{k\ell}^{(j)}\right)^m$ and converges to $p(e_\ell)$. Thus, for fixed (k,ℓ) and j sufficiently large, $|\alpha_{k\ell}^{(j)}| \le (2p(e_\ell))^{1/m}$. Hence $|\alpha_{k\ell}^{(j)}|$ is uniformly bounded over all $\{(j,k,\ell)\}$, and there exists a sequence $\{j_r\} \to \infty$ such that $\{\alpha_{k\ell}^{(j_r)}\}$ converges for each (k,ℓ), say to $\alpha_{k\ell}$. Then

$$p(x_1, \ldots, x_n) = \sum_{k=1}^{N(n,m)} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^m \in Q_{n,m}. \quad \square$$

The following material can be found in [R1, p.26]. If $[\cdot,\cdot]$ is an inner product on $\mathbf{R}^N$, then the *dual cone* to a closed convex cone $C \subset \mathbf{R}^n$ is defined by $C^* = \{x \in \mathbf{R}^N : [x,y] \ge 0 \text{ for all } y \in C\}$; C^* is also a closed convex cone. Minkowski's Separation Theorem implies that $(C^*)^* = C$. We now show that $P_{n,m}$ and $Q_{n,m}$ are dual closed convex cones in $H_m(\mathbf{R}^n)$, using (2.3) as the underlying inner product; see [R1, p.38] for more discussion.

Proposition 6.3 *For all (n,m), $P_{n,m}$ and $Q_{n,m}$ are dual cones.*

Proof. By definition, $p \in Q_{n,m}^*$ if and only if $[p,q] \ge 0$ for every $q \in Q_{n,m}$. Using Lemma 2.4(i), we see that this means $[p,q] = [p, \sum (\alpha_k \cdot)^m] = \sum p(\alpha_k) \ge 0$ for any set $\{\alpha_k\} \subset \mathbf{R}^n$, hence $p(\alpha) \ge 0$ for all $\alpha \in \mathbf{R}^n$; that is, $p \in P_{n,m}$. □

There is a simple geometric criterion for $x \in C$ to be in the interior of C.

Proposition 6.4 *Suppose $C \subset \mathbf{R}^N$ is a closed convex cone with inner product $[\cdot,\cdot]$. Then $x \in C$ is in the interior of C if and only if $[x,y] > 0$ for all non-zero y in C^*.*

Proof. Suppose $x \in C$ and $[x,y] = 0$ for some non-zero $y \in C^*$. Then for every $\eta > 0$, $[x - \eta y, y] = -\eta[y,y] < 0$, hence $x - \eta y \notin (C^*)^* = C$ and so x is not in the interior of C. Conversely, if $[x,y] > 0$ for all non-zero y in C^*, then $[x,u] \geq \delta > 0$ on the compact set $\{u \in C^* : ||u||^2 = [u,u] = 1\}$. By linearity, $[x,y] \geq \delta||y||$ for all $y \in C^*$. Now suppose $z \in \mathbf{R}^N$ and $||x - z|| \leq \delta$. Then by the foregoing and the Cauchy-Schwartz inequality, we have $[z,y] = [x,y] - [x - z, y] \geq \delta||y|| - ||x - z|| \cdot ||y|| \geq 0$ for $y \in C^*$, hence $z \in C$. Thus C contains a δ-ball around x. □

It is easy to show that the d-th powers span $H_d(\mathbf{R}^n)$. This is known (see [R1, p.30] for references) and explains why $Q_{n,d}$ is not interesting for odd d: $(-\alpha\cdot)^d = -(\alpha\cdot)^d$, so every linear combination of d-th powers is a sum of d-th powers, and so $Q_{n,d} = H_d(\mathbf{R}^n)$.

Proposition 6.5 *For all (n,d), $H_d(\mathbf{R}^n)$ is spanned by $\{(\alpha\cdot)^d : \alpha \in \mathbf{Q}^n\}$.*

Proof. Let V be the subspace of $H_d(\mathbf{R}^n)$ spanned by $\{(\alpha\cdot)^d : \alpha \in \mathbf{Q}^n\}$ and suppose $q \in V^\perp$. Then $0 = [q, (\alpha\cdot)^d] = q(\alpha)$ for all $\alpha \in \mathbf{Q}^n$. By continuity, $q(\alpha) = 0$ for all $\alpha \in \mathbf{R}^n$, so $q = 0$. Since $V^\perp = \{0\}$, we must have $V = H_d(\mathbf{R}^n)$. □

We say that $\{\alpha_1, \ldots, \alpha_{N(n,d)}\} \in \mathbf{R}^n$ is a *basic set of nodes* for $H_d(\mathbf{R}^n)$ if $\{(\alpha_k\cdot)^d\}$ is a basis for $H_d(\mathbf{R}^n)$. (This is a classical term in numerical analysis.) Proposition 6.5 asserts the existence in $\mathbf{Q}^n$ of a basic set of nodes for every (n,d). In fact, Biermann's Theorem (see [R1, p.31]) states that $\mathcal{I}(n,d) \subset \mathbf{Z}^n$ is a basic set of nodes for $H_d(\mathbf{R}^n)$.

The following integral representation for G_n^r can be found in [R1, p.105], where several other proofs are given.

Proposition 6.6 *Suppose $p \in H_{2r}(\mathbf{R}^n)$ and let $d\mu$ be Lebesgue measure on $S^{n-1} \subset \mathbf{R}^n$, normalized so that $\int_{u\in S^{n-1}} \cdots \int u_1^{2r}\, d\mu = 1$. Then,*

$$[p, G_n^r] = \int_{u\in S^{n-1}} \cdots \int p(u)\, d\mu. \tag{6.7}$$

Proof. Since (6.7) is bilinear, it suffices to verify on a spanning set for $H_{2r}(\mathbf{R}^n)$; namely, $\{(\alpha\cdot)^{2r} : \alpha \in \mathbf{R}^n\}$. That is, we must show that for $p = (\alpha\cdot)^{2r}$,

$$[(\alpha\cdot)^{2r}, G_n^r] = \int_{u\in S^{n-1}} \cdots \int (\alpha_1 u_1 + \ldots + \alpha_n u_n)^{2r}\, d\mu. \tag{6.8}$$

But the left-hand side of (6.8) is $G_n^r(\alpha) = (\sum_j \alpha_j^2)^r = |\alpha|^{2r}$ by Lemma 2.4(i). On the right-hand side, make any orthogonal change of variables in which $v_1 =$

$|\alpha|^{-1}(\alpha_1 u_1 + \ldots + \alpha_n u_n)$. The integrand becomes $|\alpha|^{2r} v_1^{2r}$ and the right-hand side becomes:

$$|\alpha|^{2r} \int_{v \in S^{n-1}} \cdots \int v_1^{2r}\, d\mu = |\alpha|^{2r}. \square$$

A standard integral formula (see e.g. [D1, p.374]) gives

$$\int_{u \in S^{n-1}} \cdots \int u_1^{2r}\, du = 2 \frac{\Gamma(\frac{2r+1}{2}) \left(\Gamma(\frac{1}{2})\right)^{n-1}}{\Gamma(\frac{2r+n}{2})} = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} \prod_{j=0}^{r-1} \frac{2j+1}{2j+n}.$$

The following theorem gives a large class of forms in the interior of $Q_{n,m}$, as well as presenting another interpretation to Φ_s.

Theorem 6.9 *Suppose $r, s \geq 0$ and $0 \neq f \in P_{n,2s}$. Then $\Phi_{2s+r}(f) G_n^r$ is interior to $Q_{n,2r+2s}$. In particular, taking $s = 0$ and $f = 1$, G_n^r is interior to $Q_{n,2r}$.*

Proof. If $p \in P_{n,2r+2s}$, then $pf \in P_{n,2r+4s}$, and by Proposition 6.6,

$$0 \leq \int_{u \in S^{n-1}} \cdots \int p(u) f(u)\, d\mu = [pf, G_n^{r+2s}], \tag{6.10}$$

with equality if and only if $p(u)f(u) = 0$ almost everywhere on S^{n-1}. By continuity, this implies $pf = 0$, and $f \neq 0$, so $p = 0$. But then by Lemma 2.4 and (6.10),

$$\begin{aligned} 0 &\leq (2r+4s)![pf, G_n^{r+2s}] \\ &= (2r+2s)![p, f(D) G_n^{r+2s}] = (2r+2s)![p, \Phi_{r+2s}(f) G_n^r], \end{aligned} \tag{6.11}$$

with equality if and only if $p = 0$. This is true for all by $p \in P_{n,2r+2s}$, hence (6.11) and Propositions 6.3 and 6.4 imply that $\Phi_{r+2s}(f) G_n^r$ is interior to $Q_{n,2r+2s}$. If $f = 1$, then this simplifies and $\Phi_{r+2s}(f) G_n^r = G_n^r$ is interior to $Q_{n,2r}$. $\square$

If $0 \neq \lambda_i$ and $f \in H_d(\mathbf{R}^n)$, define $f_\lambda \in H_d(\mathbf{R}^n)$ by $f_\lambda(x_1, \cdots, x_n) = f(\lambda_1 x_1, \cdots, \lambda_n x_n)$. Then it is easy to see that $[p, q] = [p_\lambda, q_{1/\lambda}]$. It follows by Proposition 6.4 that, if $0 \neq \lambda_i \in K$ and $h \in H_m(K^n)$ is interior to $Q_{n,m}$, then so is every h_λ. It is not hard to prove a substantial generalization of this remark: If T is any invertible linear transformation in $\mathbf{R}^n$ and p is interior to $Q_{n,m}$, then so is $p \circ T$. See also [R1, p.40] and [R2, p.1065].

Our last major theorem, which is implicit in [H5], combines with Theorem 6.9 to prove Proposition 2.6. A sketchier proof of this result (applying only to $G_n^{m/2}$ and $K = \mathbf{Q}$) can be found in Ellison [E1].

Theorem 6.12 *Suppose $h \in K[x_1, \ldots, x_n]$ and h is in the interior of $Q_{n,m}$. Then there exist $0 \leq \lambda_k \in K$ and $\alpha_k \in \mathbf{Q}^n$ so that*

$$h(x_1, \ldots, x_n) = \sum_{k=1}^{N(n,m)} \lambda_k (\alpha_{k1} x_1 + \ldots + \alpha_{kn} x_n)^m.$$

Proof. Let $\{\alpha_k : 1 \le k \le N(n,m)\} \subset \mathbf{Q}^n$ be a basic set of nodes for $H_m(\mathbf{R}^n)$. Since $\{(\alpha_k\cdot)^m\}$ spans $H_m(\mathbf{R}^n)$ and is contained in $H_m(K^n)$, it spans that space as well. Thus, if $g \in H_m(K^n)$, then there exist unique $\lambda_k \in K$ so that $g = \sum_\ell \lambda_k(\alpha_k\cdot)^m$.

Since h is interior to $Q_{n,m}$, there exists $\epsilon > 0$ so that $h - \epsilon\sum_k(\alpha_k\cdot)^m \in Q_{n,m}$. Thus there exist $\beta_\ell \in \mathbf{R}^n$ such that

$$h - \epsilon\sum_{k=1}^{N(n,m)}(\alpha_k\cdot)^m = \sum_{\ell=1}^{N(n,m)}(\beta_\ell\cdot)^m. \tag{6.13}$$

Now take $\bar\beta_\ell \in \mathbf{Q}^n$ close to β_ℓ and express $(\beta_\ell\cdot)^m - (\bar\beta_\ell\cdot)^m$ in terms of the basis $\{(\alpha_k\cdot)^m\}$:

$$(\beta_\ell\cdot)^m - (\bar\beta_\ell\cdot)^m = \sum_{k=1}^{N(n,m)}\delta_{k,\ell}(\alpha_k\cdot)^m. \tag{6.14}$$

We may choose $\bar\beta_\ell$ so close to β_ℓ that $|\delta_{k,\ell}| < N(n,m)^{-1}\epsilon$. Sum (6.14) over ℓ:

$$\sum_{\ell=1}^{N(n,m)}(\beta_\ell\cdot)^m = \sum_{\ell=1}^{N(n,m)}(\bar\beta_\ell\cdot)^m + \sum_{k=1}^{N(n,m)}\delta_k(\alpha_k\cdot)^m, \tag{6.15}$$

where $\delta_k = \sum_\ell\delta_{k,\ell}$ and $|\delta_k| < \epsilon$. Combine (6.13) and (6.15) to get

$$h = \sum_{k=1}^{N(n,m)}(\epsilon+\delta_k)(\alpha_k\cdot)^m + \sum_{\ell=1}^{N(n,m)}(\bar\beta_\ell\cdot)^m. \tag{6.16}$$

By construction, $\epsilon + \delta_k > 0$; we must check that $\epsilon + \delta_k \in K$. But, $\bar\beta_\ell \in \mathbf{Q}$, so $\bar h = h - \sum_\ell(\bar\beta_\ell\cdot)^m \in H_m(K^n)$, and so by the remark of the first paragraph, $\epsilon + \delta_k \in K$. Thus, (6.16) gives h as a non-negative K-linear combination of $2N(n,m)$ m-th powers of linear forms, each of which is in $\mathbf{Q}[x_1,\dots,x_n]$. It follows from Proposition 6.1 that h is a non-negative K-linear combination of $N(n,m)$ of these m-th powers. □

References

[B1] E. Becker, *Valuations and real places in the theory of formally real fields*, Géométrie algébrique réelle et formes quadratiques, Proceedings, Rennes 1981 (J.-L.Colliot Thélène, M. Coste, L. Mahé et M.-F. Roy, eds.), Lecture Notes in Math. No. 959, Springer-Verlag Berlin, Heidelberg, New York 1982, pp. 1–40

[B2] E. Becker, *The real holomorphy ring and sums of 2n-th powers*, Géométrie algébrique réelle et formes quadratiques, Proceedings, Rennes 1981 (J.-L.Colliot Thélène, M. Coste, L. Mahé et M.-F. Roy, eds.), Lecture Notes in Math. No. 959, Springer-Verlag Berlin, Heidelberg, New York 1982, pp. 139–181

[B3] E. Becker and V. Powers, *Sums of powers in rings and the real holomorphy ring*, preprint

[C1] M. D. Choi, T. Y. Lam and B. Reznick, *Even symmetric sextics* Math. Z. **195** (1987), 559–580

[C2] W. K. Clifford, *On syzygetic relations among the powers of linear quantics* originally in Proc. Lond. Math. Soc., vol. 3, 1869, Paper XIV in Mathematical papers, Chelsea, New York, 1968

[D1] P. J. Davis and P. Rabinowitz, *Methods of numerical integration*, 2nd ed., Academic Press, Orlando, 1984

[D2] J. A. de Loera and F. Santos, *An effective version of Pólya's thoerem on positive definite forms*, preprint

[D3] C. N. Delzell, *A constructive, continuous solution to Hilbert's 17th problem, and other results in semi-algebraic geometry*, Ph.D. thesis, Stanford University, 1980

[D4] C. N. Delzell, *Case distinctions are necessary for representing polynomials as sums of squares*, Proceedings of the Herbrand Symposium Logic Colloquium '81 (J. Stern, ed.), North-Holland, 1982, pp. 87–103

[D5] C. N. Delzell, *Continuous, piecewise-polynomial functions which solve Hilbert's 17th problem*, J. Reine Angew. Math. **440** (1993), 157–173

[D6] C. N. Delzell, *Nonexistence of analytically varying solutions to Hilbert's 17th problem*, Recent Advances in Real Algebraic Geometry and Quadratic Forms, Proc. RAGSQUAD Year, Berkeley 1990–1991 (W.B. Jacob, et al., eds.), Contemp. Math. **155**, Amer. Math. Soc. 1994, pp. 107–117

[D7] C. N. Delzell, L. González-Bega, H. Lombardi, *A continuous and rational solution to Hilbert's 17th problem and several cases of the Positivstellensatz*, Computational algebraic geometry (F. Eyssette, A. Galligo, eds.), Prog. Math. **109**, Birkhäuser (1993), 61–75

[D8] R. A. Devore and G. G. Lorentz, *Constructive approximation*, Springer-Verlag, 1993

[D9] L. E. Dickson, *History of the theory of numbers*, Vol. II, Chelsea, New York, 1966, Originally published by Carnegie Institute of Washington, 1920

[E1] W. J. Ellison, *Waring's problem*, Amer. Math. Monthly **78** (1971), 10–36

[H1] W. Habicht, *Über die Zerlegung strikte definiter Formen in Quadrate* Comment. Math. Helv. **12** (1940), 317–322

[H2] G. H. Hardy, J. Littlewood and G. Pólya, *Inequalities*, 2nd ed., Cambridge University Press, 1967

[H3] F. Hausdorff, *Zur Hilbertschen Lösung des Waringschen Problems* Math. Ann. **67** (1909), 301–305

[H4] D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten* Math. Ann. **32** (1888), 342–350; see Ges. Abh. 2, 154–161, Springer, Berlin, 1933, reprinted by Chelsea, New York, 1981

[H5] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n-ter Potenzen (Waringsches Problem)* Math. Ann. **67** (1909), 281–300; see Ges. Abh. 1, 510–527, Springer, Berlin, 1932, reprinted by Chelsea, New York, 1981

[H6] E. W. Hobson, *On a theorem in differentiation, and its application to spherical harmonics*, Proc. of the London Math. Soc. **24** (1892–1893), 55–67

[H7] E. W. Hobson, *On a theorem in the differential calculus* Mess. of Math. **23** (1893–1894), 115–119

[H8] E. W. Hobson, *The theory of spherical and ellipsoidal harmonics*, Cambridge University Press, 1931

[P1] G. Pólya, *Über positive Darstellung von Polynomen* Vierteljschr. Naturforsch. Ges. Zürich **73** (1928), 141–145; see Collected Papers, Vol. 2, pp. 309–313, MIT Press, 1974

[R1] B. Reznick, *Sums of even powers of real linear forms* Mem. Amer. Math. Soc. **96** (1992), no. 463

[R2] B. Reznick, *An inequality for products of polynomials* Proc. Amer. Math. Soc. **117** (1993), 1063–1073

[R3] R. M. Robinson, *Some definite polynomials which are not sums of squares in polynomials*, Selected questions of algebra and logic (a collection dedicated to the memory of A. I. Mal'cev), Indat. "Nauka" Sibirsk. Otdel. Novosibirsk (1973), 264–282, Abstract in Notices Amer. Math. Soc. 16 (1969), p. 554

[S1] J. Schmid, *On totally positive units of real holomorphy rings* Israel J. Math. **85** (1994), 339–350

[S2] A. Strasburger, *A generalization of the Bochner identity* Exposition. Math. **11** (1993), 153–157