HOMOGENEOUS POLYNOMIAL SOLUTIONS TO CONSTANT COEFFICIENT PDE'S

Bruce Reznick

University of Illinois
1409 W. Green St.
Urbana, IL 61801
reznick@math.uiuc.edu

October 21, 1994

1. Introduction

Suppose p is a homogeneous polynomial over a field K. Let $p(D)$ be the differential operator defined by replacing each occurence of x_j in p by $\frac{\partial}{\partial x_j}$, defined formally in case K is not a subset of $\mathbf{C}$. (The classical example is $p(x_1, \cdots, x_n) = \sum_j x_j^2$, for which $p(D)$ is the Laplacian ∇^2.) In this paper we solve the equation $p(D)q = 0$ for homogeneous polynomials q over K, under the restriction that K be an algebraically closed field of characteristic 0. (This restriction is mild, considering that the "natural" field is $\mathbf{C}$.) Our Main Theorem and its proof are the natural extensions of work by Sylvester, Clifford, Rosanes, Gundelfinger, Cartan, Maass and Helgason. The novelties in the presentation are the generalization of the base field and the removal of some restrictions on p. The proofs require only undergraduate mathematics and Hilbert's Nullstellensatz. A fringe benefit of the proof of the Main Theorem is an Expansion Theorem for homogeneous polynomials over $\mathbf{R}$ and $\mathbf{C}$. This theorem is trivial for linear forms, goes back at least to Gauss for $p = x_1^2 + x_2^2 + x_3^2$, and has also been studied by Maass and Helgason.

Main Theorem (See Theorem 4.1). *Let K be an algebraically closed field of characteristic 0. Suppose $p \in K[x_1, \cdots, x_n]$ is homogeneous and $p = \prod_j p_j^{m_j}$ for distinct non-constant irreducible factors $p_j \in K[x_1, \cdots, x_n]$. Suppose further that $q \in K[x_1, \cdots, x_n]$ is homogeneous of degree d and let $m'_j = \min(m_j, d+1)$. Then $p(D)q = 0$ if and only if there exist $h_{jk} \in K[x_1, \cdots, x_n]$, homogeneous of degree $m'_j - 1$, and $(\alpha_{jk1}, \cdots, \alpha_{jkn}) \in K^n$, satisfying $p_j(\alpha_{jk1}, \cdots, \alpha_{jkn}) = 0$, so that*

$$q(x_1, \cdots, x_n) = \sum_j \sum_k h_{jk}(x_1, \cdots, x_n)(\alpha_{jk1}x_1 + \cdots + \alpha_{jkn}x_n)^{d-m'_j+1}. \qquad (1.1)$$

If p is itself irreducible, then (1.1) simplifies: $p(D)q = 0$ if and only if q is a sum of d-th powers of linear forms whose coefficients belong to the variety determined

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

by p. One direction of this theorem is easy to make plausible. If p is homogeneous of degree r, then a formal application of the chain rule to a function $F(t)$ gives:

$$p(D)F(\alpha_1 x_1 + \cdots + \alpha_n x_n) = p(\alpha_1, \cdots, \alpha_n)F^{(r)}(\alpha_1 x_1 + \cdots + \alpha_n x_n).$$

In particular, if $p(\alpha_1, \cdots, \alpha_n) = 0$, then $p(D)$ annihilates $(\alpha_1 x_1 + \cdots + \alpha_n x_n)^d$. (This was a commonplace 19-th century observation.) In the special case $n = 2$ and $p(x, y) = x^2 + y^2$, $p(\alpha_1, \alpha_2) = 0$ implies $\alpha_2 = \pm i\alpha_1$, and we recover from the Main Theorem the familiar fact that a homogeneous d-th degree polynomial $q(x, y)$ is harmonic if and only if it is a linear combination of $(x + iy)^d$ and $(x - iy)^d$. The Main Theorem has already been proved over $\mathbf{C}$ for $n = 2$, for $p = \sum_k x_k^2$, and for p without repeated factors.

Expansion Theorem (See Theorem 4.7). *Suppose* $K = \mathbf{R}$ *or* $\mathbf{C}$, $p \in K[x_1, \cdots, x_n]$ *is homogeneous of degree* r *and* $f \in K[x_1, \cdots, x_n]$ *is homogeneous of degree* d, *and let* $m = \lfloor \frac{d}{r} \rfloor$. *Then* f *has a unique representation*

$$f = f_0 + pf_1 + \cdots + p^m f_m, \tag{1.2}$$

where $f_j \in K[x_1, \cdots, x_n]$ *is homogeneous of degree* $d - jr$, $p(D)f_j = 0$ *if* $K = \mathbf{R}$ *and* $\bar{p}(D)f_j = 0$ *if* $K = \mathbf{C}$.

The Expansion Theorem is trivial if p is linear; for example, if $p = x_1$, (1.2) amounts to the assertion $K[x_1, \cdots, x_n] = K[x_2, \cdots, x_n][x_1]$. It is also trivial if $r > d$. For $p = x_1^2 + x_2^2 + x_3^2$, the expansion (1.2) is familiar in the classical study of spherical harmonics.

The proofs in this paper primarily involve the application of simple linear algebra to vector spaces of polynomials. In section two, we present the machinery for discussing the vector spaces of homogeneous polynomials over fields of characteristic 0. This allows the succinct expression of some very old ideas and results. In section three, Hilbert's Nullstellensatz is given an unsurprising modification. Suppose φ, $f \in K[x_1, \cdots, x_n]$ are homogeneous, K is an algebraically closed field, and φ is irreducible. Then f vanishes to k-th order on the variety $\{\varphi = 0\} \subset K^n$ if and only if $\varphi^{k+1} \mid f$. In section four, we prove the Theorems and some corollaries. The proof of the Main Theorem is that the subspace of solutions to $p(D)q = 0$ and the subspace of forms of shape (1.1) have the same "perpendicular" subspaces. The proof of the Expansion Theorem follows from this perpendicularity. In section five, we summarize the literature for some circumstances in which the Main Theorem and Expansion Theorem have already been proven.

The Main Theorem was first proved for $n = 2$ in 1851 by Sylvester in the case of no multiple factors, and fully in 1886 by Gundelfinger. Gundelfinger's Theorem was studied afresh in the 1980's by Kung and Rota. The Main Theorem for $p(D) = \nabla^2$, was proved in 1871 by Clifford for $n = 3$ and for arbitrary n by Cartan in 1931. In 1981, Helgason proved Cartan's result with essentially the same perpendicularity argument of this paper. Finally, for $n = 3$, a general homogeneous form is irreducible, and the Main Theorem is the "Fundamental Theorem of Apolarity", which was given by Rosanes in 1873, and discussed by Coolidge in 1931, in the style of older algebraic geometry; i.e., without proof. A rigorous version of the Main Theorem for real irreducible forms in any number of variables was proved by Maass in 1959. Diaconis and Shahshahani proved a generalization of the Main Theorem in

1984 for $H_d(\mathbf{C}^2)$ when p is a product of distinct linear factors. Finally, in 1993, Pedersen solved the equation $p(D)q = 0$ over $\mathbf{R}$ for p with no multiple factors. The Expansion Theorem is trivial when p is a monomial. For $p = x_1^2 + x_2^2 + x_3^2$, it appears in an undated "Nachlass" in Gauss' *Werke*; for irreducible real p, this result was proved by Maass in 1959.

We are grateful to Bernard Beauzamy, Jerome Dégot, Igor Dolgachev and Paul S. Pedersen for helpful discussions and correspondence.

2. Vector spaces of forms

We begin with a discussion of the vector spaces of homogeneous polynomials. Versions of this discussion can be found in the author's work, with base field $\mathbf{R}$ in [25] and with base field $\mathbf{C}$ in [26,27]. It must be emphasized that the underlying ideas in this section are all very old. A parallel development using the "hypercube representation" can be found in the recent work of Beauzamy and his collaborators (see e.g. [1,2]).

Let $H_d(K^n)$ denote the set of homogeneous polynomials (forms) in n variables with degree d and coefficients in a field K of characteristic 0. (We have previously written $F_{n,d}$ for $H_d(\mathbf{R}^n)$ and $\mathcal{F}_{n,d}$ for $H_d(\mathbf{C}^n)$.) Suppose $n \geq 1$ and $d \geq 0$. The index set for monomials in $H_d(K^n)$ consists of n-tuples of non-negative integers:

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

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 $u_1^{i_1} \dots u_n^{i_n}$, 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 characteristic conclusion ensures that $c(i) \neq 0$ in K, and so can be factored in (2.1) from the coefficient of x^i in f.) 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 $\alpha \in K^n$, define $(\alpha \cdot)^d \in H_d(K^n)$ by

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

These d-th powers are of central importance, especially when combined with a natural symmetric bilinear form on $H_d(K^n)$. For p and q in $H_d(K^n)$, let

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

Note that $[p, x^j] = a(p;j)$, hence $[p,q] = 0$ for all $q \in H_d(K^n)$ if and only if $p = 0$, and so this bilinear form is non-degenerate. If $K = R$ is a formally real field, then char $R = 0$, so this discussion applies, and $[f,f] = 0$ implies $\sum_i c(i) a(f;i)^2 = 0$, so $a(f;i) = 0$, or $f = 0$. Thus, the bilinear form gives an inner product on $H_d(R^n)$.

This is false if K is not formally real. However, if $K = C = R[i]$, then $(f,g) = [f,\bar{g}]$ turns $H_d(C^n)$ into an inner product space over C. In particular, if $R = \mathbf{R}$ and $C = \mathbf{C}$, then $||f|| = [f,\bar{f}]$ is known as the Bombieri norm; see [1,2].

Let U be a vector subspace of $H_d(K^n)$, and by analogy to the real case, let

$$U^{\perp} = \{v \in H_d(K^n) : [u,v] = 0 \quad \text{for all} \quad u \in U\}. \tag{2.4}$$

It is clear that $U^{\perp}$ is also a vector subspace of $H_d(K^n)$. We give an explicit proof of an elementary observation, because it will be central in the proofs of the theorems.

Lemma 2.5. $(U^{\perp})^{\perp} = U$.

Proof. First observe that $U \subseteq (U^{\perp})^{\perp}$ from (2.4). Also, $\{0\}^{\perp} = H_d(K^n)$ and, as noted above, $p \in H_d(K^n)^{\perp}$ implies $p = 0$. Now suppose U is a proper vector subspace of $H_d(K^n)$ and $\dim(U) = m$. Then U is spanned by m linear independent elements $u^{(j)}$, $1 \le j \le m$, and so $v \in U^{\perp}$ if and only if $[u^{(j)}, v] \equiv 0$. This occurs if and only if the $N(n,d)$-tuple $\{X_i = a(v;i)\}$ solves the linear system

$$\sum_{i \in \mathcal{I}(n,d)} c(i) a(u^{(j)}; i) X_i = 0, \qquad 1 \le j \le m. \tag{2.6}$$

Since the $u^{(j)}$'s are linearly independent, the matrix in (2.6) has rank m and so the solution space to this system has dimension $N(n,d) - m$. That is, $\dim(U) + \dim(U^{\perp}) = N(n,d)$. It follows that $\dim((U^{\perp})^{\perp}) = m = \dim(U)$, and so $(U^{\perp})^{\perp} = U$. $\square$

It follows from (2.2) and (2.3) that

$$[p, (\alpha\cdot)^d] = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) a((\alpha\cdot)^d; i) = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) \alpha^i = p(\alpha). \tag{2.7}$$

This identification leads immediately to a folk-theorem which has often appeared in the literature, see [25,p.30]; for a stronger version, see Corollary 4.5.

Proposition 2.8. *$H_d(K^n)$ is spanned by $\{(\alpha\cdot)^d : \alpha \in K^n\}$.*

Proof. Let U be the subspace of $H_d(K^n)$ spanned by $\{(\alpha\cdot)^d : \alpha \in K^n\}$ and suppose $q \in U^{\perp}$. Then by (2.7), $0 = [q, (\alpha\cdot)^d] = q(\alpha)$ for all $\alpha \in K^n$. Since $\mathrm{char}(K) = 0$, this implies $q = 0$. Thus, $U^{\perp} = \{0\}$, so $U = (U^{\perp})^{\perp} = \{0\}^{\perp} = H_d(K^n)$. $\square$

For $i \in \mathcal{I}(n,d)$, let $D^i = \prod (\frac{\partial}{\partial x_k})^{i_k}$ and let $f(D) = \sum c(i) a(f;i) D^i$ be the d-th order differential operator associated to $f \in H_d(K^n)$. (These operators are formal and will be applied only to polynomials.) Since $\frac{\partial}{\partial x_k}$ and $\frac{\partial}{\partial x_\ell}$ commute, it follows that $D^i D^j = D^{i+j} = D^j D^i$ for $i \in \mathcal{I}(n,d)$ and $j \in \mathcal{I}(n,e)$, and so by multilinearity, $(fg)(D) = f(D)g(D) = g(D)f(D)$ for any forms f and g (possibly of different degree).

Lemma 2.9. *Suppose $i \in \mathcal{I}(n,d)$ and $j \in \mathcal{I}(n,e)$. If $i_k > j_k$ for some k, then $D^i(x^j) = 0$. If $i_k \le j_k$ for all k and $j - i = \ell \in \mathcal{I}(n, e-d)$, then*

$$D^i(x^j) = D^i(x^{i+\ell}) = \frac{e!}{(e-d)!} \frac{c(\ell)}{c(i+\ell)} x^{\ell}. \tag{2.10}$$

Proof. We have

$$D^i(x^j) = \prod_{k=1}^n \left(\frac{\partial^{i_k}}{\partial x_k^{i_k}}\right) \prod_{k=1}^n x_k^{j_k} = \prod_{k=1}^n \frac{\partial^{i_k} x_k^{j_k}}{\partial x_k^{i_k}}.$$

If $i_r > j_r$, then the r-th factor above is zero. Otherwise, writing $j_k = i_k + \ell_k$, we have

$$D^i(x^{i+\ell}) = \prod_{k=1}^n \frac{\partial^{i_k} x_k^{i_k+\ell_k}}{\partial x_k^{i_k}} = \prod_{k=1}^n \frac{(i_k+\ell_k)!}{\ell_k!} x_k^{\ell_k}$$
$$= \prod_{k=1}^n (i_k+\ell_k)! \prod_{k=1}^n \frac{1}{\ell_k!} \prod_{k=1}^n x_k^{\ell_k} = \frac{e!}{c(i+\ell)} \frac{c(\ell)}{(e-d)!} x^\ell. \quad \square$$

If $d = e$ but $i \neq j$, then $i_k > j_k$ for some k, so $D^i(x^j) = 0$; (2.10) implies that $D^i(x^i) = \prod_k i_k! = d!/c(i)$. An easy consequence of this fact (see Theorem 2.11(i) below) is the essence of the connection between differential operators and the bilinear product.

Theorem 2.11. *(i) If $p, q \in H_d(K^n)$, then $p(D)q = q(D)p = d![p,q]$.*

(ii) If $f, hp \in H_d(K^n)$, where $p \in H_r(K^n)$ and $h \in H_{d-r}(K^n)$, then

$$[hp, f] = \frac{(d-r)!}{d!}\ [h, p(D)f]. \tag{2.12}$$

(In particular, (2.12) implies that $[hp, f] = 0$ if and only if $[h, p(D)f] = 0$.)

(iii) If $p \in H_d(K^n)$ and $q \in H_e(K^n)$ with $d \le e$, then

$$p(D)q = \frac{e!}{(e-d)!} \sum_{\ell \in \mathcal{I}(n,e-d)} c(\ell) \Bigg(\sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) a(q;i+\ell) \Bigg) x^\ell. \tag{2.13}$$

Proof. (i) We have

$$\begin{aligned} p(D)q &= \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) D^i \Bigg(\sum_{j \in \mathcal{I}(n,d)} c(j) a(q;j) x^j \Bigg) \\ &= \sum_{i \in \mathcal{I}(n,d)} \sum_{j \in \mathcal{I}(n,d)} c(i) c(j) a(p;i) a(q;j) D^i x^j. \end{aligned} \tag{2.14}$$

As noted above, $D^i x^j = 0$ if $i \neq j$ and so (2.14) reduces to

$$p(D)q = \sum_{i \in \mathcal{I}(n,d)} c(i) c(i) a(p;i) a(q;i) D^i x^i = \sum_{i \in \mathcal{I}(n,d)} c(i) c(i) a(p;i) a(q;i) \frac{d!}{c(i)}$$
$$= d! \sum c(i) a(p;i) a(q;i) = d![p,q] = d![q,p] = q(D)p.$$

(ii) By two applications of (i), we have

$$d![hp,f] = (hp)(D)f = h(D)p(D)f = h(D)(p(D)f) = (d-r)![h,p(D)f].$$

(iii) As in (i), we have

$$\begin{aligned} p(D)q &= \sum_{i\in\mathcal{I}(n,d)} c(i)a(p;i)D^i\Bigg(\sum_{j\in\mathcal{I}(n,e)} c(j)a(q;j)x^j\Bigg) \\ &= \sum_{i\in\mathcal{I}(n,d)}\sum_{j\in\mathcal{I}(n,e)} c(i)c(j)a(p;i)a(q;j)D^ix^j. \end{aligned} \tag{2.15}$$

By Lemma 2.9, the only non-zero terms in (2.15) have $j = i+\ell$ for $\ell \in \mathcal{I}(n, e-d)$, so

$$\begin{aligned} p(D)q &= \sum_{i\in\mathcal{I}(n,d)}\sum_{\ell\in\mathcal{I}(n,e-d)} c(i)c(i+\ell)a(p;i)a(q;i+\ell)D^ix^{i+\ell} \\ &= \sum_{i\in\mathcal{I}(n,d)}\sum_{\ell\in\mathcal{I}(n,e-d)} c(i)c(i+\ell)a(p;i)a(q;i+\ell)\frac{e!}{(e-d)!}\frac{c(\ell)}{c(i+\ell)}x^\ell \\ &= \frac{e!}{(e-d)!}\sum_{i\in\mathcal{I}(n,d)}\sum_{\ell\in\mathcal{I}(n,e-d)} c(i)c(\ell)a(p;i)a(q;i+\ell)x^\ell \\ &= \frac{e!}{(e-d)!}\sum_{\ell\in\mathcal{I}(n,e-d)} c(\ell)\Bigg(\sum_{i\in\mathcal{I}(n,d)} c(i)a(p;i)a(q;i+\ell)\Bigg)x^\ell. \quad \square \end{aligned}$$

A few remarks on this theorem. If $p \in H_d(K^n)$, $q \in H_e(K^n)$ and $d > e$, then $p(D)q$ is identically zero; if $d = e$, then $\mathcal{I}(n, d-d) = \{(0,\cdots,0)\}$ and $c((0,\cdots,0)) = 1$, so (iii) reduces to (i). Finally, if $d \le e$ but $p(D)q = 0$, then (2.13) implies that $\{a(q;j)\}$ satisfies a finite linear recurrence: for all $\ell \in \mathcal{I}(n, e-d)$,

$$\sum_{i\in\mathcal{I}(n,d)} c(i)a(p;i)a(q;i+\ell) = 0. \tag{2.16}$$

Most instances of $[p,q]$ in the literature have been in terms of $p(D)q$. In modern times, Stein and Weiss in 1971 [29,p.139] used the fact that $(p,q) = p(D)\bar{q}$ is an inner product on $H_d(\mathbf{C}^n)$. So did Helgason [11,12] in the early 1980s, in proving the Main Theorem for $p = \sum_k x_k^2$ over $\mathbf{C}$. Generalizations of (ii) over $\mathbf{C}$ using the hypercube and the Bombieri norm can be found in [2].

We conclude this section of introductory material with a result which combines many of its ideas. Suppose K is a field of characteristic 0, $p \in H_r(K^n)$ and d is fixed. Define

$$V_p(= V_{p,d}) = \{q \in H_d(K^n) : p(D)q = 0\}. \tag{2.17}$$

Clearly, V_p is a subspace of $H_d(K^n)$.

Theorem 2.18. $V_p^\perp = X_p(= X_{p,d}) = \{f \in H_d(K^n) : p \mid f\}$.

Proof. If $r > d$, then $V_p = H_d(K^n)$ and $X_p = \{0\}$, so the result is trivial. Otherwise, $r \le d$ and $f \in X_p$ if and only if $f = pg$ for some $g \in H_{d-r}(K^n)$. Thus, $q \in X_p^\perp$ if and only if $[pg,q] = 0$ for all $g \in H_{d-r}(K^n)$; that is, if and only if $[px^j,q] = 0$ for all $j \in \mathcal{I}(n,d-r)$. We conclude from Theorem 2.11(ii) that $q \in X_p^\perp$ if and only if $[x^j, p(D)q] = 0$ for all $j \in \mathcal{I}(n,d-r)$; that is, if and only if $p(D)q = 0$. Thus $X_p^\perp = V_p$ and the conclusion follows by Lemma 2.5. $\square$

3. A variation on the Nullstellensatz

For $\varphi \in H_r(K^n)$, let $\mathcal{Z}(\varphi) = \{\varphi = 0\} = \{(\alpha_1, \cdots, \alpha_n) \in K^n : \varphi(\alpha_1, \cdots, \alpha_n) = 0\}$. We recall Hilbert's Nullstellensatz, e.g. from [15,pp.409–413].

Hilbert's Nullstellensatz. *Suppose K is an algebraically closed extension field of a field F and suppose I is a proper ideal of $F[x_1, \cdots, x_n]$. Let*

$$V(I) = \{(a_1, \cdots, a_n) \in K^n : g(a_1, \cdots, a_n) = 0 \text{ for all } g \in I\} = \bigcap_{g \in I} \mathcal{Z}(g).$$

Then $f(a_1, \cdots, a_n) = 0$ for all $(a_1, \cdots, a_n) \in V(I)$ if and only if $f^m \in I$ for some $m \geq 1$.

In particular, suppose $I = (\varphi)$ is the ideal generated by an irreducible form $\varphi \in H_r(K^n)$, $r \geq 1$. Then $V(I) = \mathcal{Z}(\varphi)$ and the Nullstellensatz states that $f \in H_d(K^n)$ vanishes on $\mathcal{Z}(\varphi)$ if and only if $f^m \in I$ for some m; that is, if and only if $\varphi \mid f^m$ for some m. Since φ is irreducible, $\varphi \mid f^m$ if and only if $\varphi \mid f$. Hence the only forms which vanish on the variety $\{\varphi = 0\}$ are multiples of φ. This is the simplest case of a variation on the Nullstellensatz which is needed for the proof of the Main Theorem.

Proposition 3.1. *Suppose $\varphi \in H_r(K^n)$, $r \geq 1$, is irreducible, $f \in H_d(K^n)$ and $0 \leq k \leq d$. Then $D^i f(\alpha) = 0$ for all $i \in \mathcal{I}(n,k)$ and all $\alpha \in \mathcal{Z}(\varphi)$ if and only if $\varphi^{k+1} \mid f$.*

Proof. We first show that $D^i f(\alpha) = 0$ for all $i \in \mathcal{I}(n,k)$ if and only if $D^i f(\alpha) = 0$ for all $i \in \mathcal{I}(n,\ell)$ for all $\ell \leq k$. One direction is obvious. For the other, fix $\alpha \in K^n$ and assume $D^i f \in H_{d-k}(K^n)$ for all $i \in \mathcal{I}(n,k)$. Theorem 2.11(i) implies that $D^i f(\alpha) = [D^i f, (\alpha\cdot)^{d-k}]$, hence by Theorem 2.11(ii), $D^i f(\alpha) = 0$ for all $i \in \mathcal{I}(n,k)$ if and only if $[f, (\alpha\cdot)^{d-k} x^i] = 0$ for all $i \in \mathcal{I}(n,k)$; that is, if and only if $[f, (\alpha\cdot)^{d-k}\psi] = 0$ for all $\psi \in H_k(K^n)$. In particular, if $\psi = (\alpha\cdot)^{k-\ell}\zeta$ for $\zeta \in H_\ell(K^n)$, then $[f, (\alpha\cdot)^{d-\ell}\zeta] = 0$. Now take $\zeta = x^j$ for $j \in \mathcal{I}(n,\ell)$. This completes the proof of the first assertion. (Note that if $k > d$, then $D^i f \equiv 0$ without restriction on f, and no non-trivial conclusion may be drawn about the lower order derivatives.)

We let $\mathcal{A}(f,k;\varphi)$ denote the assertion that $D^i f(\alpha) = 0$ for all $i \in \mathcal{I}(n,\ell)$ for all $\ell \leq k$ and all $\alpha \in \mathcal{Z}(\varphi)$, so the content of this Proposition is that $\mathcal{A}(f,k;\varphi)$ is true if and only if $\varphi^{k+1} \mid f$. If $f = 0$, this result is trivial, so assume $f \neq 0$. The proof is by induction on k.

In the base case, $\mathcal{A}(f,0;\varphi)$ states that $f(\alpha) = 0$ for all $\alpha \in \mathcal{Z}(\varphi)$, and as noted earlier, the Nullstellensatz implies that this is equivalent to $\varphi = \varphi^{0+1} \mid f$. Now assume the induction hypothesis holds for $k = s-1 \geq 0$ and suppose $\mathcal{A}(f,s;\varphi)$ holds. If $i \in \mathcal{I}(n,s)$, then $i_\ell \geq 1$ for at least one ℓ, so $D^i f = D^j \frac{\partial f}{\partial x_\ell}$ for some $j \in \mathcal{I}(n,s-1)$. Taking into account all $i \in \mathcal{I}(n,s)$, we may conclude that $\mathcal{A}(f,s;\varphi)$ is equivalent to $\mathcal{A}(\frac{\partial f}{\partial x_\ell}, s-1;\varphi)$ for $1 \leq \ell \leq n$. By the induction hypothesis, this is equivalent to $\varphi^s \mid \frac{\partial f}{\partial x_\ell}$ for $1 \leq \ell \leq n$. The proof is complete if we can show that this fact implies that $\varphi^{s+1} \mid f$.

Since $\mathcal{A}(f,s-1;\varphi)$ holds (by the first paragraph), the induction hypothesis implies that $f = \varphi^s\psi$, $s \geq 1$, $\psi \neq 0$, and since $\varphi^s \mid \frac{\partial f}{\partial x_\ell}$ for $1 \leq \ell \leq n$, we have

$$\varphi^s \;\Big|\; \frac{\partial f}{\partial x_\ell} = \frac{\partial(\varphi^s\psi)}{\partial x_\ell} = \varphi^s\frac{\partial\psi}{\partial x_\ell} + s\varphi^{s-1}\frac{\partial\varphi}{\partial x_\ell}\psi. \tag{3.2}$$

Upon dividing both sides of (3.2) by φ^{s-1}, we find that $\varphi \mid \varphi \frac{\partial \psi}{\partial x_\ell} + s\psi \frac{\partial \varphi}{\partial x_\ell}$, hence $\varphi \mid \psi \frac{\partial \varphi}{\partial x_\ell}$ for all $1 \le \ell \le n$. Since φ is non-constant, $\frac{\partial \varphi}{\partial x_\ell} \ne 0$ for some ℓ; since $\deg \frac{\partial \varphi}{\partial x_\ell} < \deg \varphi$, the irreducibility of φ implies that $\varphi \mid \psi$. Thus $\varphi^{s+1} \mid f$, as desired. $\square$

4. Proof of the Main Theorem

We first restate the Main Theorem using the notation of the last two sections.

Theorem 4.1. *Suppose K is an algebraically closed field of characteristic 0 and suppose $p \in H_r(K^n)$ has the factorization $p = \prod_{j=1}^m p_j^{m_j}$ into distinct non-constant irreducible forms p_j over K. Suppose $q \in H_d(K^n)$ and let $m_j' = \min(m_j, d+1)$. Then $p(D)q = 0$ if and only if there exist $h_{jk} \in H_{m_j'-1}(K^n)$ and $\alpha_{jk} \in \mathcal{Z}(p_j)$ so that*

$$q = \sum_{j=1}^{m} \sum_{k=1}^{N(n,d)} h_{jk}(\alpha_{jk}\cdot)^{d-m_j'+1}. \tag{4.2}$$

Proof. We recall (2.17) and define another subspace of $H_d(K^n)$:

$$W_p (= W_{p,d}) = \left\{ \sum_{j=1}^{m} \sum_{k=1}^{N(n,d)} h_{jk}(\alpha_{jk}\cdot)^{d-m_j+1} : h_{jk} \in H_{m_j'-1}(K^n), \alpha_{jk} \in \mathcal{Z}(p_j) \right\}. \tag{4.3}$$

We shall prove that $V_p^\perp = W_p^\perp$. By Lemma 2.5, this is equivalent to $V_p = W_p$. First suppose $m_j \ge d+1$ for some j. Then $\deg p \ge m_j \deg p_j \ge d+1 > \deg q$, and so $p(D)q = 0$ for all $q \in H_d(K^n)$. On the other hand, $m_j' = d+1$, so we can write $q = h_{j1}$ in (4.2) for all $q \in H_d(K^n)$. Having disposed of this annoying special case, we henceforth assume that $m_j \le d$ for all j and write m_j for m_j'.

By Theorem 2.18, it suffices to show that $W_p^\perp = X_p$. Observe that $f \in W_p^\perp$ if and only if $[f, g] = 0$ for each permissible summand $g = h(\alpha_j\cdot)^{d-m_j+1}$ in (4.3). Expand h as a linear combination of monomials x^i with $i \in \mathcal{I}(n, m_j - 1)$. Thus $f \in W_p^\perp$ if and only if

$$[f, x^i(\alpha_j\cdot)^{d-m_j+1}] = 0 \tag{4.4}$$

for all $i \in \mathcal{I}(n, m_j - 1)$ and all $\alpha_j \in \mathcal{Z}(p_j)$, $1 \le j \le m$. By Theorem 2.11(ii), (4.4) is equivalent to $D^i f(\alpha_j) = 0$, and by Proposition 3.1, this is equivalent to $p_j^{m_j} \mid f$. This is true for $1 \le j \le m$, and since the p_j's are distinct irreducibles, $f \in W_p^\perp$ if and only if $p \mid f$; that is, if and only if $f \in X_p$. $\square$

Suppose $m_j \equiv 1$, so p has no multiple factors. Then the auxiliary forms h_{jk} disappear from (4.2), and $p(D)q = 0$ if and only if q is a sum of d-th powers of linear forms taken from $\{p = 0\}$. In particular, if $r > d$, then $p(D)q = 0$ for all $q \in H_d(K^n)$, and we obtain a strengthening of Proposition 2.8:

Corollary 4.5. *Suppose $p \in H_r(K^n)$ has no multiple factors and $r > d$. Then $H_d(K^n)$ is spanned by $\{(\alpha\cdot)^d : \alpha \in \mathcal{Z}(p)\}$.*

It is easy to compute the dimension of V_p from Lemma 2.5 and Theorem 2.18.

Corollary 4.6. *If $d \geq r$, then $dim(V) = N(n,d) - N(n,d-r)$; if $d < r$ then $dim(V) = N(n,d)$.*

An equivalent version of Corollary 4.6 for $H_d(\mathbf{C}^2)$ was used by Sylvester to give the canonical representation of a binary d-ic as a sum of d-th powers of linear forms. For $n \geq 3$, however, the resulting bound is larger than the correct number; see [27].

A final corollary (of Theorem 2.18) is the Expansion Theorem.

Theorem 4.7. *Suppose $K = \mathbf{R}$ or $\mathbf{C}$, $p \in H_r(K^n)$ and $f \in H_d(K^n)$, and let $m = \lfloor \frac{d}{r} \rfloor$. Then there is a unique representation*

$$f = f_0 + pf_1 + \cdots + p^m f_m, \tag{4.8}$$

where $f_j \in H_{d-jr}(K^n)$, $p(D)f_j = 0$ if $K = \mathbf{R}$ and $\bar{p}(D)f_j = 0$ if $K = \mathbf{C}$.

Proof. We fix r and argue by induction on d. The assertion is trivial if $d < r$, so assume that $d \geq r$. Suppose $K = \mathbf{R}$. Then the bilinear product is an inner product and there is an orthogonal decomposition $H_d(\mathbf{R}^n) = V_p \oplus X_p$. Hence there is a unique representation $f = f_0 + \tilde{f}$, where $f_0 \in V_p$ and $\tilde{f} \in X_p$; that is $f = f_0 + p\hat{f}_1$, where $p(D)f_0 = 0$ and $\hat{f}_1 \in H_{d-r}(\mathbf{R}^n)$. By induction, $\hat{f}_1$ has an expansion as in (4.8), and the proof is complete. In case $K = \mathbf{C}$, the inner product involves the conjugate and X_p is orthogonal to $\overline{V}_p$. □

We remark that Theorem 4.7 also applies with $\mathbf{R}$ replaced by any formally real field R, and $\mathbf{C}$ replaced by any field $C = R[i]$, with R formally real. It also can be shown [3] that (4.8) is a true orthogonal decomposition (i.e., $[p^i f_i, p^j f_j] = 0$ if $i < j$) when p is linear or $p = \sum_k x_k^2$, but not in general.

5. Antecedents of the main theorem

The notation of section two is ungainly for binary forms over $\mathbf{C}$. Suppose $r \leq d$ and let

$$p(x,y) = \sum_{t=0}^{r} c_t x^t y^{r-t} = \prod_{j=1}^{m} (\beta_j x - \alpha_j y)^{m_j}, \tag{5.1}$$

$$q(x,y) = \sum_{u=0}^{d} \binom{d}{u} a_u x^j y^{u-j}. \tag{5.2}$$

In (5.1), we assume $\beta_j \alpha_k \neq \beta_k \alpha_j$ for $j \neq k$. On taking $x_1 = x$ and $x_2 = y$, $a(p;(t,r-t)) = c_t / \binom{t}{r}$ and $a(q;(u,d-u)) = a_u$. Then by (2.13), $p(D)q = 0$ if and only if

$$\sum_{t=0}^{r} c_t a_{t+\ell} = 0, \quad \ell = 0, \dots, d-r,$$

or in matrix form,

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5.3}$$

In his studies of the canonical forms for $H_d(\mathbf{C}^2)$, Sylvester [30,31] proved that if p and q are given by (5.1) (with $m_j \equiv 1$) and (5.2), then there exist $\lambda_j \in \mathbf{C}$ so that

$$q(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^d$$

if and only if (5.3) holds. Later, Gundelfinger [10] removed the restriction on m_j and showed that (5.3) holds if and only if

$$q(x,y) = \sum_{j=1}^{m} h_j(x,y)(\alpha_j x + \beta_j y)^{d-(m_j-1)}$$

for arbitrary forms h_j of degree $m_j - 1$. This is, of course, the Main Theorem for $n = 2$ and $K = \mathbf{C}$. The modern discussion of this topic by Kung [16,17,18], and Kung and Rota [19] (using different algebraic techniques) has considered a covariant expression $\{p, q\}$ which equals $[p, \tilde{q}]$ up to a multiple, where $\tilde{q}(x, y) = q(y, -x)$. Diaconis and Shashahani [7,p.176] studied the equation $p(D)q$ for a binary form p which is the product of r distinct linear factors and $q(x, y) \in C^n[0, 1]^2$. Their solution reduces to Sylvester's theorem when q is a form. Helmke [13] has recently studied the problem of representing $q \in H_d(K^2)$ as a sum of d-th powers of linear forms, where K is real closed or algebraically closed, using continued fractions and Padé approximations.

In case $p(x, y, z) = x^2 + y^2 + z^2$, $p(D)q = 0$ becomes $\nabla^2 q = 0$, in which case q is called a spherical harmonic. There is an enormous classical literature on this topic (see e.g. [14,21,32]) discussing all analytic solutions to this equation, not just polynomial ones. The Main Theorem asserts that $\nabla^2 q = 0$ for $q \in H_d(K^n)$ if and only if $q(x_1, \cdots, x_n) = \sum_k (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^d$, where $\sum_j \alpha_{kj}^2 = 0$. Clifford [5] proved this for $n = 3$, using Sylvester's theory of contravariants and dimension-counting. However, there are much more concrete and useful ways to represent spherical harmonics, and this direction was not pursued. Later, Cartan [4,p.285] proved the result for general n. A rigorous version for of the Main Theorem for p real and irreducible forms was given by Maass [20] in 1959. Helgason's proofs [11,p.29;12,p.17] of Cartan's result can be regarded as a direct antecedent of the proof of Theorem 4.1 given here. (However, $n = 2$ was treated separately, because $x_1^2 + x_2^2$ is reducible.) He also generalized the Cartan and Maass versions of the Main Theorem (e.g. [12,p.381]) in ways that are deeper than this paper.

The author presented a version of the Main Theorem (with $m_j \equiv 1$ and $K = \mathbf{C}$) at the 1990 Summer Meeting of the A.M.S. at Columbus, Ohio under the title "Whose Theorem Is This?" [24]. Prof. Igor Dolgachev [8] answered the question by saying that this special case of the Main Theorem (for $n = 3$) is known as "The Fundamental Theorem of Apolarity". He cited Coolidge [6,pp.410–416], which is based on Rosanes [28]. The proof in [6] is geometric, in the style of older algebraic geometry, which means that it is proved for general p and q. Singular cases are excluded, and not discussed explicitly. Very recently, Pedersen [22] has given a complete description of real solutions to $p(D)q = 0$ when p has no multiple factors. Since the Nullstellensatz is inoperative over $\mathbf{R}$, his discussion of the solutions is necessarily more complicated than the Main Theorem. Pedersen also proves the equivalent of Corollary 4.6 for the dimension of $\{q : p(D)q = 0\}$.

In the special case $p = x_1^2 + x_2^2 + x_3^2$, the Expansion Theorem is well-known in the classical study of spherical harmonics. The cited texts all note that every form f of degree d can be written $f = f_0 + pf_1 + p^2 f_2 + \cdots$, where f_j is harmonic of degree $d - 2j$. It is difficult to determine who first discovered this decomposition; Prasad [23,p.114] assigns credit to Gauss [9,pp.630–631]. Maass [20] proved the Expansion Theorem for real irreducible p in the same paper cited above. Helgason also generalized the Expansion Theorem in [12]; again, his approach is much deeper than ours.

References

[1] Bernard Beauzamy, Enrico Bombieri, Per Enflo and Hugh L. Montgomery, *Products of polynomials in many variables*, J. Number Theory **36** (1990), 219–245.

[2] Bernard Beauzamy and Jérome Dégot, *Differential identities*, to appear, Proc. Amer. Math. Soc..

[3] Bernard Beauzamy and Bruce Reznick, *in preparation*.

[4] E. Cartan, *Leçons sur la géométrie projective complexe*, Gauthier-Villars, Paris, 1931.

[5] William Kingdon Clifford, *On a canonical form of spherical harmonics*, Report of the forty-first meeting of the British Association for the Advancement of Science, held in Edinburgh, August 1871, p.10, Paper XXV in Mathematical papers, Chelsea, New York, 1968, Originally published by Macmillan in 1882.

[6] Julian Lowell Coolidge, *A treatise on algebraic plane curves*, Dover, New York, 1959, Originally published by Oxford University Press in 1931.

[7] Persi Diaconis and Mehrdad Shahshahani, *On nonlinear functions of linear combinations*, SIAM J. Sci. Statist. Comput. **5** (1984), 175–191.

[8] Igor Dolgachev, *personal communication*.

[9] Carl Freiderich Gauss, *Werke, Bd. V – Mathematische Physik*, Göttingen, 1867.

[10] S. Gundelfinger, *Zur theorie der binären formen*, J. Reine Angew. Math. **100** (1886), 413–424.

[11] Sigurdur Helgason, *Topics in harmonic analysis on homogeneous spaces*, Birkhäuser, Boston, 1981.

[12] Sigurdur Helgason, *Groups and geometric analysis*, Academic Press, Orlando, 1984.

[13] U. Helmke, *Waring's problem for binary forms*, J. Pure Appl. Alg. **80** (1992), 29–45.

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

[15] Thomas W. Hungerford, *Algebra*, Springer-Verlag, New York, 1974.

[16] Joseph P. S. Kung, *Gundelfinger's theorem on binary forms*, Stud. Appl. Math. **75** (1986), 163–170.

[17] Joseph P. S. Kung, *Canonical forms for binary forms of even degree*, Invariant theory (S. S. Koh, ed.), Lecture Notes in Mathematics, 1278, Springer-Verlag, New York, 1987, pp. 52–61.

[18] Joseph P. S. Kung, *Canonical forms of binary forms: Variations on a theme of Sylvester*, Invariant theory and tableaux, Proceedings of the Workshop on Invariant Theory and Tableaux held at the Institute of Mathematics and its Applications on March 21-25, 1988 (D. Stanton, ed.), Springer-Verlag, 1989, pp. 46–58.

[19] Joseph P. S. Kung and Gian-Carlo Rota, *The invariant theory of binary forms*, Bull. Amer. Math. Soc. (N. S.) **10** (1984), 27–85.

[20] Hans Maass, *Theorie der hamonischen Formen*, Math. Ann. **137** (1959), 142–149.

[21] T. M. MacRobert, *Spherical harmonics*, Third edition, revised with the assistance of I. N. Sneddon, Pergamon, Oxford, 1967.

[22] Paul S. Pedersen, *A function theory for finding a basis for all polynomial solutions to linear constant coefficient pde's of homogeneous order*, Complex Variables Theory Appl. **24** (1993), 79–87.

[23] Ganesh Prasad, *A treatise on spherical harmonics and the functions of Bessel and Lamé, Part 1 (Elementary)*, Benares Mathematical Society, 1930.

[24] Bruce Reznick, *Whose theorem is this? (859–14–98)*, Abstracts Amer. Math. Soc. **11** (1990), 311.

[25] Bruce Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992), no. 463.
[26] Bruce Reznick, *An inequality for products of polynomials*, Proc. Amer. Math. Soc. **117** (1993), 1063–1073.
[27] Bruce Reznick, *Sums of powers of linear forms*, in preparation.
[28] J. Rosanes, *Ueber ein Princip der Zuordnung algebraischer Formen*, J. Reine Angew. Math. **LXXVI** (1873), 312–330.
[29] Elias M. Stein and Guido Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, 1971.
[30] James Joseph Sylvester, *An essay on canonical forms, supplement to a sketch of a memoir on elimination, transformation, and canonical forms*, originally published separately by George Bell, Fleet Street 1851, Paper 34 in Mathematical Papers, vol. 1, Chelsea, New York, 1973, Originally published by Cambridge University Press in 1904.
[31] James Joseph Sylvester, *On a remarkable discovery in the theory of canonical forms and of hyperdeterminants*, originally in Philosophical Magazine, vol. 2, 1851, Paper 42 in Mathematical Papers, vol. 1, Chelsea, New York, 1973, Originally published by Cambridge University Press in 1904.
[32] Sir William Thomson (Lord Kelvin) and Peter Guthrie Tait, *Principles of Mechanics and Dynamics*, (Formerly titled Treatise on Natural Philosophy), Dover, New York, 1961, Originally published by Cambridge University Press in 1879–1912.