Constr. Approx. (1995) 11: 165–180

CONSTRUCTIVE APPROXIMATION

Multivariate Polynomials: A Spanning Question

A. Pinkus and B. Wajnryb

Abstract. The main result of this paper is the following. If g is any given polynomial of two variables, then

$$\mathrm{span}\{(g(\cdot - a, \cdot - b))^k : (a, b) \in \mathbf{R}^2, k \in \mathbf{Z}_+\}$$

contains all polynomials if and only if

$$\mathrm{span}\{g(\cdot - a, \cdot - b) : (a, b) \in \mathbf{R}^2\}$$

separates points. This result is not valid in $\mathbf{R}^d$ for $d \geq 4$.

1. Introduction and Motivation

There are many different methods of approximating multivariate functions. We have, for example, the classic methods of polynomials, Fourier series, or tensor products, and more modern methods using wavelets, radial basis functions, multivariate splines, or ridge functions. Many of these are natural generalizations of methods developed for approximating univariate functions. However, functions of many variables are fundamentally different from functions of one variable, and approximation techniques for such problems are much less developed and understood. We discuss in these pages one approach to this problem which is truly multivariate in character.

To motivate the approach considered, we recall Hilbert's 13th problem. Although not formulated in the following terms, Hilbert's 13th problem was interpreted by some as conjecturing that not all functions of three variables could be represented as *superpositions* (compositions) and *sums* of functions of two variables. Surprisingly it transpired that all functions could be so represented, and even more was true. Kolmogorov and his student Arnold proved in a series of papers in the late 1950s that fixed continuous one-variable functions h_{ij} exist such that every continuous function f of d variables on $[0, 1]^d$ could be represented in the form

$$f(x_1, \ldots, x_d) = \sum_{i=1}^{2d+1} l_i\left(\sum_{j=1}^{d} h_{ij}(x_j)\right),$$

where the continuous one-variable functions l_i are chosen.

Date received: June 17, 1993. Date revised: January 10, 1994. Communicated by Tim N. T. Goodman.
AMS classification: 41A10, 41A63.
Key words and phrases: Multivariate polynomials, Spanning set.

As with any such surprising and deep result, numerous questions and results were spawned. One question asked had to do with smoothness properties of the h_{ij}. Assuming f is in some class of smooth functions, can smooth h_{ij} be chosen?

Different answers were given in different frameworks. One answer due to Vitushkin and Henkin (separately and together) was given in the mid 1960s (see [2] for a survey of their results). It says that one cannot demand that the h_{ij} be C^1 functions. (Since the answer is in the negative, we only formulate it for functions of two variables x and y in $[0, 1]^2$.)

Theorem 1. *For any m fixed continuous functions $\psi_i(x, y)$, $i = 1, \ldots, m$, and continuously differentiable functions $\varphi_i(x, y)$, $i = 1, \ldots, m$, the set of functions*

$$\left\{ \sum_{i=1}^{m} \psi_i(x, y) l_i(\varphi_i(x, y)) : l_i \text{ continuous} \right\}$$

is nowhere dense in the space of all functions continuous in $[0, 1]^2$ with the topology of uniform convergence.

There is, however, a fundamental difference between "representing" and "approximating." While the above implies that representation is not possible in this setting, this does not mean that one cannot approximate arbitrarily well. Perhaps it might be possible to find a family Φ of smooth useful functions $\varphi: \mathbf{R}^d \to \mathbf{R}$ (an infinite number of such functions) such that

$$\operatorname{span}\{l(\varphi(\cdot)) : l \in C(\mathbf{R}),\ \varphi \in \Phi\}$$

is dense in our appropriate space.

One simple choice of a family Φ is to take a fixed function h of d variables, and to consider all its shifts (translates). That is, we consider the set

$$\mathscr{H} = \operatorname{span}\{l(h(\cdot - \mathbf{a})) : l \in C(\mathbf{R}),\ \mathbf{a} \in \mathbf{R}^d\},$$

for some fixed function h, and ask when such a space is dense in, say, the space of continuous real-valued functions defined on $\mathbf{R}^d$, endowed with the topology of uniform convergence on compact sets. (This space is taken for convenience only.)

One simple necessary condition is that

$$\operatorname{span}\{h(\cdot - \mathbf{a}) : \mathbf{a} \in \mathbf{R}^d\}$$

must separate points. In other words, for $\mathbf{x}, \mathbf{y} \in \mathbf{R}^d$, $\mathbf{x} \neq \mathbf{y}$, a point $\mathbf{a} \in \mathbf{R}^d$ must exist such that $h(\mathbf{x} - \mathbf{a}) \neq h(\mathbf{y} - \mathbf{a})$. If this condition does not hold, then all functions in $\mathscr{H}$ take exactly the same values at $\mathbf{x}$ and at $\mathbf{y}$, and so cannot possibly be dense in our space.

A simple sufficient condition is the following. If h is a C^1 function, and

$$\{\mathbf{x} : c \leq h(\mathbf{x}) \leq d\}$$

is a bounded nonempty set for some choice of $c < d$, then $\mathscr{H}$ is dense in $C(\mathbf{R}^d)$.

This can be proved using Fourier-transform-type arguments. This condition is not necessary.

In this generality, it seems to be rather difficult to say much more. However, in case h is a polynomial, we were led to conjecture that the simple necessary condition of separation of points mentioned above might be sufficient. This conjecture is valid for $d = 2$, and in fact an even stronger statement holds in this case. This is the main result of this paper (Theorem 2). We prove that

$$\mathscr{G} = \operatorname{span}\{(g(\cdot - \mathbf{a}))^k \colon k \in \mathbf{Z}_+,\ \mathbf{a} \in \mathbf{R}^2\}$$

contains all algebraic polynomials in $\mathbf{R}^2$ if g is any algebraic polynomial such that

$$\operatorname{span}\{g(\cdot - \mathbf{a}) \colon \mathbf{a} \in \mathbf{R}^2\}$$

separates points. (This result is not valid in $\mathbf{R}^d$, $d \geq 4$.)

In Section 2 we state and prove Theorem 2. Our proof is more general than necessary, as in Section 3 we use this proof to state various results in connection with $d \geq 3$. In Section 3 we also discuss some related problems. For example, we prove that it suffices to consider only a much more restricted set of shifts.

It remains an open and interesting problem to determine, for $d \geq 3$, necessary and sufficient conditions on an algebraic polynomial g so that

$$\mathscr{G} = \operatorname{span}\{(g(\cdot - \mathbf{a}))^k \colon k \in \mathbf{Z}_+,\ \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials in $\mathbf{R}^d$ (or is dense in $C(\mathbf{R}^d)$).

2. Theorem 2 and Its Proof

In this section we prove the following result

Theorem 2. *Let g be an algebraic polynomial on $\mathbf{R}^2$. Then*

$$\mathscr{G} = \operatorname{span}\{(g\cdot - a, \cdot - b))^k \colon k \in \mathbf{Z}_+,\ (a, b) \in \mathbf{R}^2\}$$

contains all algebraic polynomials of x and y if and only if

$$\operatorname{span}\{g(\cdot - a, \cdot - b) \colon (a, b) \in \mathbf{R}^2\}$$

separates points.

One direction is obvious, and we restate it. If

$$\operatorname{span}\{g(\cdot - a, \cdot - b) \colon (a, b) \in \mathbf{R}^2\}$$

does not separate points, then $\mathscr{G}$ cannot possibly contain all algebraic polynomials. As such we concern ourselves only with the proof of the converse direction. It is more convenient to deal with different but equivalent statements of the fact that the above space separates points, and that we are dealing with all shifts. This we present in the next two lemmas. When possible, proofs are given in $\mathbf{R}^d$ rather than in $\mathbf{R}^2$. Let us also note some standard multivariate notation. For $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbf{R}^d$ and $\mathbf{j} = (j_1, \ldots, j_d) \in \mathbf{Z}_+^d$, we set

$$|\mathbf{j}| = j_1 + \cdots + j_d, \qquad \mathbf{j}! = j_1! \cdots j_d!$$

and

$$\mathbf{x}^{\mathbf{j}} = x_1^{j_1} \cdots x_d^{j_d}, \qquad D^{\mathbf{j}} = \frac{\partial^{|\mathbf{j}|}}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}.$$

If q is a polynomial, $q(\mathbf{x}) = \sum_{\mathbf{j}} a_{\mathbf{j}} \mathbf{x}^{\mathbf{j}}$, then by $q(D)$ we mean the differential operator

$$\sum_{\mathbf{j}} a_{\mathbf{j}} D^{\mathbf{j}}.$$

Lemma 3. *The following are equivalent for any algebraic polynomial* $h\colon \mathbf{R}^d \to \mathbf{R}$.

(1) $\operatorname{span}\{h(\cdot - \mathbf{a})\colon \mathbf{a} \in \mathbf{R}^d\}$ *does not separate points.*
(2) *The functions* $\{\partial h/\partial x_i\}_{i=1}^d$ *are linearly dependent.*

Proof. Assume (1) holds. Thus $\mathbf{x}, \mathbf{y} \in \mathbf{R}^d$, $\mathbf{x} \neq \mathbf{y}$, exist such that $h(\mathbf{x} - \mathbf{a}) = h(\mathbf{y} - \mathbf{a})$ for all $\mathbf{a} \in \mathbf{R}^d$. Letting $\mathbf{u} = \mathbf{x} - \mathbf{y}$ and considering $\mathbf{a}$ as the variable, it follows that

$$h(\cdot) = h(\cdot + \mathbf{u}).$$

Thus

$$h(\cdot) = h(\cdot + n\mathbf{u})$$

for all $n \in \mathbf{Z}$. For each $\mathbf{x} \in \mathbf{R}^d$, set

$$r(t) = h(\mathbf{x} + t\mathbf{u}) - h(\mathbf{x}).$$

Then r is a polynomial in one variable of some finite degree which vanishes at all integers. Thus $r = 0$. Differentiating with respect to t and then setting $t = 0$, we obtain

$$\sum_{i=1}^d u_i \frac{\partial h}{\partial x_i} = 0,$$

and (2) holds.

The above is essentially reversible. If (2) holds, i.e.,

$$\sum_{i=1}^d u_i \frac{\partial h}{\partial x_i} = 0$$

for some $\mathbf{u} = (u_1, \ldots, u_d) \neq \mathbf{0}$, then it follows that

$$h(\cdot) = h(\cdot + t\mathbf{u})$$

for all $t \in \mathbf{R}$ and thus (1) holds. ∎

Remark 4. If h is a polynomial in (x, y) (i.e., in $\mathbf{R}^2$), then the linear dependence of $\partial h/\partial x$ and $\partial h/\partial y$ is easily seen to be equivalent to the fact that h is of the form

$$h(x, y) = \sum_{i=0}^{n} a_i(cx + dy)^i$$

for some choice of $(c, d) \in \mathbf{R}^2$, $n \in \mathbf{Z}_+$, and $(a_0, \ldots, a_n) \in \mathbf{R}^{n+1}$.

We also wish to note that varying over all possible shifts (translates) is equivalent to considering all derivatives. This latter form proves more amenable.

Lemma 5. *For any algebraic polynomial* $h\colon \mathbf{R}^d \to \mathbf{R}$,

$$\operatorname{span}\{h(\cdot - \mathbf{a})\colon \mathbf{a} \in \mathbf{R}^d\} = \operatorname{span}\{D^{\mathbf{j}}h\colon \text{all } \mathbf{j} \in \mathbf{Z}_+^d\}.$$

Proof. The result is well known. Each of the above spaces is closed, and finite-dimensional. One inclusion follows from the fact that every derivative may be obtained as the limit of a differencing operation. The other inclusion is an immediate consequence of Taylor's formula. ∎

Based on the above two lemmas and the remark, we can now restate what we will prove.

Assume the polynomial g is not a polynomial in $cx + dy$ for any $(c, d) \in \mathbf{R}^2$. Then

$$\operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\, \partial y^j}\,(g)^k\colon k, i, j \in \mathbf{Z}_+\right\}$$

contains all algebraic polynomials.

Let $\mathbf{p} = (p_1, \ldots, p_d) \in \mathbf{N}^d$. We say that a polynomial h is $\mathbf{p}$-*homogeneous of degree* n if

$$h(\mathbf{x}) = \sum_{\mathbf{p}\cdot\mathbf{j}=n} a_{\mathbf{j}}\mathbf{x}^{\mathbf{j}},$$

where $\mathbf{p}\cdot\mathbf{j} = \sum_{i=1}^{d} p_i j_i$. For each $r \in \mathbf{N}$,

$$\begin{aligned}(h(\mathbf{x}))^r &= \left(\sum_{\mathbf{p}\cdot\mathbf{j}=n} a_{\mathbf{j}}\mathbf{x}^{\mathbf{j}}\right)^r \\ &= \sum_{\substack{\mathbf{p}\cdot\mathbf{j}^l=n \\ l=1,\ldots,r}} a_{\mathbf{j}^1}\cdots a_{\mathbf{j}^r}\mathbf{x}^{\mathbf{j}^1+\cdots+\mathbf{j}^r} \\ &= \sum_{\mathbf{p}\cdot\mathbf{j}=nr} c_{\mathbf{j}}^{(r)}\,\frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}\end{aligned}$$

due to the linearity of $\mathbf{p}\cdot\mathbf{j}$. Thus the rth power of a $\mathbf{p}$-homogeneous polynomial of degree n is a $\mathbf{p}$-homogeneous polynomial of degree nr.

Proposition 6. *Let h be a* $\mathbf{p}$*-homogeneous polynomial of degree n. The set*

$$\operatorname{span}\{D^{\mathbf{j}}(h)^r: \text{all } \mathbf{j}\in \mathbf{Z}_+^d,\ r\in \mathbf{Z}_+\}$$

does not contain all $\mathbf{p}$*-homogeneous polynomials of degree t if and only if a nontrivial* $\mathbf{p}$*-homogeneous polynomial q of degree t exists such that*

$$q(D)(h)^r = 0, \qquad r = 0, 1, \ldots .$$

Proof. One direction is obvious. Since $q(D)(h)^r = 0$ for all $r\in \mathbf{Z}_+$, we have that $q(D)D^{\mathbf{j}}(h)^r = D^{\mathbf{j}}q(D)(h)^r = 0$ for all $r\in\mathbf{Z}_+$ and all $\mathbf{j}\in\mathbf{Z}_+^d$. Thus $q(D)$ annihilates every polynomial in the above span. However, it is easily seen that $q(D)q \neq 0$.

For the other direction, let $\mathbf{m}\in\mathbf{Z}_+^d$ be such that $\mathbf{p}\cdot\mathbf{m}\le nr$. Then

$$\begin{aligned} D^{\mathbf{m}}(h(\mathbf{x}))^r &= D^{\mathbf{m}}\left(\sum_{\mathbf{p}\cdot\mathbf{j}=nr} c_{\mathbf{j}}^{(r)}\frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}\right)\\ &= \sum_{\substack{\mathbf{p}\cdot\mathbf{j}=nr\\ \mathbf{j}\ge\mathbf{m}}} c_{\mathbf{j}}^{(r)}\frac{\mathbf{x}^{\mathbf{j}-\mathbf{m}}}{(\mathbf{j}-\mathbf{m})!} = \sum_{\mathbf{p}\cdot\mathbf{j}=nr-\mathbf{p}\cdot\mathbf{m}} c_{\mathbf{j}+\mathbf{m}}^{(r)}\frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}. \end{aligned}$$

Now

$$\operatorname{span}\{D^{\mathbf{j}}(h)^r: \text{all } \mathbf{j}\in \mathbf{Z}_+^d,\ r\in \mathbf{Z}_+\}$$

will contain all $\mathbf{p}$-homogeneous polynomials of degree t if and only if

$$\operatorname{span}\{D^{\mathbf{m}}(h)^r: \mathbf{p}\cdot\mathbf{m} = nr - t,\ r\in\mathbf{Z}_+\}$$

contains all $\mathbf{p}$-homogeneous polynomials of degree t. For $\mathbf{m}$ such that $\mathbf{p}\cdot\mathbf{m} = nr - t$,

$$D^{\mathbf{m}}(h(\mathbf{x}))^r = \sum_{\mathbf{p}\cdot\mathbf{j}=t} c_{\mathbf{j}+\mathbf{m}}^{(r)}\frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}.$$

The number of linearly independent $\mathbf{p}$-homogeneous polynomials of degree t is finite. The above polynomials do not span this set if and only if constants $d_{\mathbf{j}}$, $\mathbf{p}\cdot\mathbf{j} = t$, not all zero, exist such that

$$\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}c_{\mathbf{j}+\mathbf{m}}^{(r)} = 0$$

for all $\mathbf{m}$ and r such that $\mathbf{p}\cdot\mathbf{m} = nr - t$. Set

$$q(\mathbf{x}) = \sum_{\mathbf{p}\cdot\mathbf{i}=t} d_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}.$$

Then

$$q(D)(h(\mathbf{x}))^r = \sum_{\mathbf{p}\cdot\mathbf{i}=t} d_{\mathbf{i}} D^{\mathbf{i}}\left(\sum_{\mathbf{p}\cdot\mathbf{j}=nr} c_{\mathbf{j}}^{(r)} \frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}\right)$$

$$= \sum_{\mathbf{p}\cdot\mathbf{i}=t} d_{\mathbf{i}} \sum_{\mathbf{p}\cdot\mathbf{j}=nr-t} c_{\mathbf{j}+\mathbf{i}}^{(r)} \frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}$$

$$= \sum_{\mathbf{p}\cdot\mathbf{j}=nr-t} \left(\sum_{\mathbf{p}\cdot\mathbf{i}=t} d_{\mathbf{i}} c_{\mathbf{i}+\mathbf{j}}^{(r)}\right) \frac{\mathbf{x}^{\mathbf{j}}}{\mathbf{j}!}.$$

The monomials $\mathbf{x}^{\mathbf{j}}/\mathbf{j}!$ are linearly independent. Thus

$$q(D)(h)^r = 0$$

if and only if

$$\sum_{\mathbf{p}\cdot\mathbf{i}=t} d_{\mathbf{i}} c_{\mathbf{i}+\mathbf{j}}^{(r)} = 0$$

for each $r \in \mathbf{Z}_+$ and all $\mathbf{j}$ with $\mathbf{p}\cdot\mathbf{j} = nr - t$. This proves the proposition. ∎

We use an equivalent form of the fact that

$$q(D)(h)^r = 0, \qquad r = 0, 1, \ldots.$$

To find this form, we first prove this next lemma.

Lemma 7. *Let h be a $\mathbf{p}$-homogeneous polynomial of degree n. For $\mathbf{j} \in \mathbf{Z}_+^d \backslash \{\mathbf{0}\}$,*

$$D^{\mathbf{j}}(h)^r = \sum_{i=1}^{|\mathbf{j}|} \frac{r!}{(r-i)!} h^{r-i} \varphi_i(h; \mathbf{j}; \cdot),$$

where $\varphi_i(h; \mathbf{j}; \cdot)$ is a polynomial which depends on h, $\mathbf{j}$, and i, but is independent of r. Furthermore,

$$\varphi_{|\mathbf{j}|}(h; \mathbf{j}; \cdot) = \left(\frac{\partial h}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial h}{\partial x_d}\right)^{j_d}.$$

Proof. Our proof is via induction on $|\mathbf{j}|$. If $|\mathbf{j}| = 1$, then $\mathbf{j} = \mathbf{e}_l$ for some $l \in \{1, \ldots, d\}$ and

$$\frac{\partial}{\partial x_l}(h)^r = r(h)^{r-1} \frac{\partial h}{\partial x_l}$$

is of the desired form.

Now let $\mathbf{j} = \mathbf{k} + \mathbf{e}_l$ (i.e., $|\mathbf{j}| = |\mathbf{k}| + 1$). Assume the result holds for $\mathbf{k}$. Thus

$$D^{\mathbf{j}}(h)^r = \frac{\partial}{\partial x_l} D^{\mathbf{k}}(h)^r = \frac{\partial}{\partial x_l}\left(\sum_{i=1}^{|\mathbf{k}|} \frac{r!}{(r-i)!} h^{r-i} \varphi_i(h; \mathbf{k}; \cdot)\right),$$

where $\varphi_i(h; \mathbf{k}; \cdot)$ depends on h, $\mathbf{k}$, and i, but is independent of r, and

$$\varphi_{|\mathbf{k}|}(h; \mathbf{k}; \cdot) = \left(\frac{\partial h}{\partial x_1}\right)^{k_1} \cdots \left(\frac{\partial h}{\partial x_d}\right)^{k_d}.$$

Continuing, we have

$$\begin{aligned}
\frac{\partial^{|\mathbf{j}|}}{\partial \mathbf{x}^{\mathbf{j}}}(h)^r &= \sum_{i=1}^{|\mathbf{j}|-1} \frac{r!}{(r-i)!}(r-i)h^{r-i-1}\frac{\partial h}{\partial x_l}\varphi_i(h; \mathbf{k}; \cdot) + \sum_{i=1}^{|\mathbf{j}|-1} \frac{r!}{(r-i)!} h^{r-i}\frac{\partial \varphi_i}{\partial x_l}(h; \mathbf{k}; \cdot) \\
&= \sum_{i=2}^{|\mathbf{j}|} \frac{r!}{(r-i)!} h^{r-i}\frac{\partial h}{\partial x_l}\varphi_{i-1}(h; \mathbf{k}; \cdot) + \sum_{i=1}^{|\mathbf{j}|-1} \frac{r!}{(r-i)!} h^{r-i}\frac{\partial \varphi_i}{\partial x_l}(h; \mathbf{k}; \cdot) \\
&= \sum_{i=1}^{|\mathbf{j}|} \frac{r!}{(r-i)!} h^{r-i}\varphi_i(h; \mathbf{j}; \cdot),
\end{aligned}$$

where

$$\varphi_1(h; \mathbf{j}; \cdot) = \frac{\partial \varphi_1}{\partial x_l}(h; \mathbf{k}; \cdot),$$

$$\varphi_i(h; \mathbf{j}; \cdot) = \frac{\partial h}{\partial x_l}\varphi_{i-1}(h; \mathbf{k}; \cdot) + \frac{\partial \varphi_i}{\partial x_l}(h; \mathbf{k}; \cdot), \qquad i = 2, \ldots, |\mathbf{j}| - 1,$$

and

$$\varphi_{|\mathbf{j}|}(h; \mathbf{j}; \cdot) = \frac{\partial h}{\partial x_l}\varphi_{|\mathbf{j}|-1}(h; \mathbf{k}; \cdot) = \left(\frac{\partial h}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial h}{\partial x_d}\right)^{j_d}.$$

The result follows. ∎

In the above $h^i = 0$ for $i < 0$. Note that it is in fact easy to get formulae for each φ_i and thus show them to be independent of the "ordering."

For each $\mathbf{j} \in \mathbf{Z}_+^d \setminus \{\mathbf{0}\}$ and i, $1 \le i \le |\mathbf{j}|$, the function $\varphi_i(h; \mathbf{j}; \cdot)$ is well defined. In what follows we set $\varphi_i(h; \mathbf{j}; \cdot) = 0$ if $i > |\mathbf{j}|$, and also set $m = \max\{|\mathbf{j}|: \mathbf{p} \cdot \mathbf{j} = t\}$.

Proposition 8. *Let h be a* $\mathbf{p}$*-homogeneous polynomial of degree n and let*

$$q(\mathbf{x}) = \sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\mathbf{x}^{\mathbf{j}}.$$

Then

$$q(D)(h)^r = 0, \qquad r = 0, 1, 2, \ldots,$$

if and only if

$$\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\varphi_i(h; \mathbf{j}; \cdot) = 0, \qquad i = 1, 2, \ldots, m.$$

Proof. From Lemma 7,

$$\begin{aligned} q(D)(h)^r &\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}} D^{\mathbf{j}}(h)^r \\ &= \sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}} \sum_{i=1}^{|\mathbf{j}|} \frac{r!}{(r-i)!} h^{r-i}\varphi_i(h;\mathbf{j};\cdot) \\ &= \sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}} \sum_{i=1}^{m} \frac{r!}{(r-i)!} h^{r-i}\varphi_i(h;\mathbf{j};\cdot) \\ &= \sum_{i=1}^{m} \frac{r!}{(r-i)!} h^{r-i}\left(\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\varphi_i(h;\mathbf{j};\cdot)\right). \end{aligned}$$

If

$$\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\varphi_i(h;\mathbf{j};\cdot) = 0, \qquad i = 1, \ldots, m,$$

then $q(D)(h)^r = 0$ for all r. Assume $q(D)(h)^r = 0$ for all r. Set

$$g_i = h^{m-i}\left(\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\varphi_i(h;\mathbf{j};\cdot)\right).$$

Thus, for $r \geq m$,

$$0 = q(D)(h)^r = \sum_{i=1}^{m} \frac{r!}{(r-i)!} h^{r-m} g_i = h^{r-m}\left(\sum_{i=1}^{m} \frac{r!}{(r-i)!} g_i\right).$$

Since each of the factors of the right-hand side are polynomials, and h is nontrivial, it follows that

$$\sum_{i=1}^{m} \frac{r!}{(r-i)!} g_i = 0, \qquad r \geq m.$$

The $m \times m$ matrix

$$\left(\frac{r!}{(r-i)!}\right)_{i=1,r=m}^{m,2m+1}$$

is nonsingular. Thus $g_i = 0$ for $i = 1, \ldots, m$, which implies that

$$\sum_{\mathbf{p}\cdot\mathbf{j}=t} d_{\mathbf{j}}\varphi_i(h;\mathbf{j};\cdot) = 0, \qquad i = 1, \ldots, m. \qquad \blacksquare$$

Proof of Theorem 2. We assume that

$$\operatorname{span}\{g(\cdot - a, \cdot - b)\colon (a, b) \in \mathbf{R}^2\}$$

separates points. By Remark 4 this implies that g is not a polynomial in $cx + dy$ for any $(c, d) \in \mathbf{R}^2$. In particular, g is of total degree n ($n \geq 2$):

Thus

$$g(x, y) = \sum_{i+j \le n} a_{ij} x^i y^j.$$

By a linear change of variables, which has no effect on the question of whether the span of our set contains all algebraic polynomials, we may assume that $a_{n0} \neq 0$. Furthermore, if

$$\sum_{i+j=n} a_{ij} x^i y^j = (cx + dy)^n$$

for some $(c, d) \in \mathbf{R}^2$, we are free to arrange this linear change of variable to have a_{n0} be the only nonzero coefficient a_{ij} with $i + j = n$. Thus it is no loss of generality, based on our assumption that g is not a polynomial in $cx + dy$ for any $(c, d) \in \mathbf{R}^2$, to assume that $a_{n0} \neq 0$, and either

$$\sum_{i+j=n} a_{ij} x^i y^j \neq (cx + dy)^n,$$

for any $(c, d) \in \mathbf{R}^2$, or $a_{ij} = 0$ for $i + j = n$, $(i, j) \neq (n, 0)$, and an $a_{ij} \neq 0$, where $i + j < n$ and $j \neq 0$, exists.

Let v/u be the irreducible fraction such that

$$\frac{v}{u} = \min\left\{\frac{n-i}{j} : a_{ij} \neq 0, j \neq 0\right\}.$$

Note that $ui + vj \le un$ for all (i, j) with $a_{ij} \neq 0$, and at least two distinct pairs of indices (i, j) exist such that $ui + vj = un$. One pair is simply $(n, 0)$. In addition, $v \ge u$, and $v = u$ if and only if

$$\sum_{i+j=n} a_{ij} x^i y^j \neq (cx + dy)^n,$$

for any $(c, d) \in \mathbf{R}^2$ (in which case $v = u = 1$).

We abuse notation by setting $\mathbf{p} = (u, v)$ and $\mathbf{k} = (i, j)$. Thus h is a $\mathbf{p}$-homogeneous polynomial of degree s if

$$h(x, y) = \sum_{\mathbf{p}\cdot\mathbf{k} = ui + vj = s} c_{ij} x^i y^j.$$

We write g in the form

$$g(x, y) = f(x, y) + r_1(x, y),$$

where

$$f(x, y) = \sum_{\mathbf{p}\cdot\mathbf{k} = un} a_{ij} x^i y^j,$$

$$r_1(x, y) = \sum_{\mathbf{p}\cdot\mathbf{k} < un} a_{ij} x^i y^j.$$

That is, f is a $\mathbf{p}$-homogeneous polynomial of degree un.

We claim that it suffices to prove that

$$\operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}(f)^k\colon k,\,i,\,j\in\mathbf{Z}_+\right\}$$

contains all algebraic polynomials. We verify this fact by induction. Assume that

$$\operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}(g)^k\colon k,\,i,\,j\in\mathbf{Z}_+\right\}$$

contains all polynomials of $\mathbf{p}$-degree $< t$, i.e., the span of all monomials $x^i y^j$ with $ui + vj < t$. Let h be any polynomial of $\mathbf{p}$-degree t. Then $h = h_1 + h_2$, where h_1 is $\mathbf{p}$-homogeneous of degree t, and h_2 is a polynomial of $\mathbf{p}$-degree $< t$. Now, by assumption,

$$h_1 \in \operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}(f)^k\colon k,\,i,\,j\in\mathbf{Z}_+\right\},$$

which implies, since f is the leading $\mathbf{p}$-homogeneous term of g, that an

$$\tilde{h} = h_1 + \tilde{h}_2 \in \operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}(g)^k\colon k,\,i,\,j\in\mathbf{Z}_+\right\},$$

where $\tilde{h}_2$ is of $\mathbf{p}$-degree $< t$, exists. By the induction hypothesis $h_2 - \tilde{h}_2$ is also in this span, and thus so is h. We remark that this same idea is valid in $\mathbf{R}^d$.

Assume that

$$\operatorname{span}\left\{\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}(f)^k\colon k,\,i,\,j\in\mathbf{Z}_+\right\}$$

does not contain all algebraic polynomials. Since f is a $\mathbf{p}$-homogeneous polynomial of degree un, it follows from Propositions 6 and 8 (and using the notation therein) that a nontrivial $\mathbf{p}$-homogeneous polynomial q of some degree t,

$$q(x, y) = \sum_{\mathbf{p}\cdot\mathbf{k}=t} d_{ij} x^i y^j,$$

exists such that

$$q(D)(f)^r = 0, \qquad r = 0, 1, \ldots,$$

and

$$\sum_{\mathbf{p}\cdot\mathbf{k}=t} d_{ij}\varphi_k(f; (i, j); \cdot) = 0, \qquad k = 1, \ldots, m,$$

where we can take $m = \max\{i + j\colon ui + vj = t,\ d_{ij} \neq 0\}$.

Assume $v > u$. In this case there is a unique choice of (i, j), which we denote by (i^*, j^*), such that $d_{i^* j^*} \neq 0$,

$$i^* + j^* = m, \qquad ui^* + vj^* = t.$$

That is,

$$0 = \sum_{\mathbf{p}\cdot\mathbf{k}=t} d_{ij}\varphi_m(f; (i,j); \cdot) = d_{i^*j^*}\varphi_m(f; (i^*, j^*); \cdot).$$

Now, from Lemma 7,

$$\varphi_m(f; (i^*, j^*); \cdot) = \left(\frac{\partial f}{\partial x}\right)^{i^*}\left(\frac{\partial f}{\partial y}\right)^{j^*}.$$

Thus either

$$\frac{\partial f}{\partial x} = 0 \quad \text{or} \quad \frac{\partial f}{\partial y} = 0.$$

However, this contradicts our construction of f.

If $u = v$, then $\mathbf{p}\cdot\mathbf{k} = |(i,j)| = i + j$, and $m = t$. Thus

$$0 = \sum_{\mathbf{p}\cdot\mathbf{k}=t} d_{ij}\varphi_t(f; (i,j); \cdot) = \sum_{i+j=t} d_{ij}\left(\frac{\partial f}{\partial x}\right)^{i}\left(\frac{\partial f}{\partial y}\right)^{j}.$$

Every nontrivial homogeneous polynomial in two variables (x, y) can be factored into linear components. That is,

$$\sum_{i+j=t} d_{ij}x^i y^j = \prod_{k=1}^{t} (a_k x + b_k y),$$

where $(a_k, b_k) \neq (0, 0)$, $k = 1, \ldots, t$. Thus

$$0 = \sum_{i+j=t} d_{ij}\left(\frac{\partial f}{\partial x}\right)^{i}\left(\frac{\partial f}{\partial y}\right)^{j} = \prod_{k=1}^{t}\left(a_k\frac{\partial f}{\partial x} + b_k\frac{\partial f}{\partial y}\right).$$

However, this implies that

$$a_k\frac{\partial f}{\partial x} + b_k\frac{\partial f}{\partial y} = 0$$

for some $k \in \{1, \ldots, t\}$, again contradicting our construction of f. This completes the proof of Theorem 2. ■

3. Additional Remarks

In this section we gather together various facts connected with the problem under consideration.

In Theorem 2 we characterized all polynomials of two variables for which

$$\operatorname{span}\{(g(\cdot - \mathbf{a}))^k : k \in \mathbf{Z}_+, \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials. If $d = 3$, we do not know if the necessary condition of separation of points is also sufficient. For $d \geq 4$, it is insufficient as we now show.

Example 9. The polynomial

$$g(x_1, x_2, x_3, x_4) = x_1x_4^2 + x_2 + x_3x_4$$

is such that the linear span of its shifts separates points, but the linear span of its shifts and powers does not contain all algebraic polynomials in $\mathbf{R}^4$.

Proof. As noted in Lemma 3, the linear span of the shifts of a polynomial separate points if and only if the first partials of the polynomial are linearly independent. A simple calculation shows this to be the case here.

g is a $\mathbf{p}$-homogeneous polynomial of degree 3 with $\mathbf{p} = (1, 3, 2, 1)$. From Proposition 6, the linear span of its shifts and powers does not contain all algebraic polynomials in $\mathbf{R}^4$ if and only if a nontrivial $\mathbf{p}$-homogeneous polynomial q exists such that

$$q(D)(g)^r = 0, \qquad r = 0, 1, 2, \ldots.$$

Let

$$q(x_1, x_2, x_3, x_4) = x_1x_2 - x_3^2.$$

q is a $(1, 3, 2, 1)$-homogeneous polynomial of degree 4, and

$$q(D)(g)^r = \left(\frac{\partial^2}{\partial x_1\, \partial x_2} - \frac{\partial^2}{\partial x_3^2}\right)(x_1x_4^2 + x_2 + x_3x_4)^r = 0$$

for all r. ∎

For most polynomials

$$\operatorname{span}\{(g(\cdot - \mathbf{a}))^k \colon k \in \mathbf{Z}_+, \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials. This follows from the analysis of Section 2, as we now show.

Proposition 10. *For*

$$g(\mathbf{x}) = \sum_{|\mathbf{j}| \le n} a_{\mathbf{j}}\mathbf{x}^{\mathbf{j}},$$

set

$$f(\mathbf{x}) = \sum_{|\mathbf{j}| = n} a_{\mathbf{j}}\mathbf{x}^{\mathbf{j}},$$

If the first partials of f, $\{\partial f/\partial x_i\}_{i=1}^d$, *are algebraically independent, then*

$$\operatorname{span}\{(g(\cdot - \mathbf{a}))^k \colon k \in \mathbf{Z}_+, \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials.

Proof. As noted in the proof of Theorem 2,

$$\operatorname{span}\{(g(\cdot - \mathbf{a}))^k \colon k \in \mathbf{Z}_+, \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials if

$$\operatorname{span}\{(f(\cdot - \mathbf{a}))^k : k \in \mathbf{Z}_+, \mathbf{a} \in \mathbf{R}^d\}$$

contains all algebraic polynomials. The polynomial f is homogeneous. If this latter span does not contain all algebraic polynomials, then from Proposition 8

$$\sum_{|\mathbf{j}|=t} d_{\mathbf{j}} \varphi_i(f; \mathbf{j}; \cdot) = 0, \qquad i = 1, 2, \ldots, t,$$

for some t and nontrivial choice of $d_{\mathbf{j}}$. For $i = t$,

$$0 = \sum_{|\mathbf{j}|=t} d_{\mathbf{j}} \varphi_t(f; \mathbf{j}; \cdot) = \sum_{|\mathbf{j}|=t} d_{\mathbf{j}} \left(\frac{\partial f}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial f}{\partial x_d}\right)^{j_d}.$$

Thus the first partials of f are algebraically dependent. ∎

For a homogeneous polynomial f, a necessary condition for our set to contain all algebraic polynomials is that the first partials must be linearly independent. A sufficient condition is that they be algebraically independent. For $d = 2$ these are one and the same. It is of interest to try to determine necessary and sufficient conditions in general, and for homogeneous polynomials in particular.

Until now we have considered all possible shifts of our polynomial. However, it is possible to consider a much more restricted set of shifts and get exactly the same result.

Proposition 11. *For every polynomial h,*

$$\operatorname{span}\{h(\cdot - \mathbf{a}) : \mathbf{a} \in \mathbf{R}^d\} = \operatorname{span}\{h(\cdot - \mathbf{a}) : \mathbf{a} \in \mathscr{A}\}$$

if $\mathscr{A} \subseteq \mathbf{R}^d$ is not contained in any algebraic variety (i.e., no nontrivial polynomial vanishes on $\mathscr{A}$).

Proof. For any given h, both

$$\mathscr{C} = \operatorname{span}\{h(\cdot - \mathbf{a}) : \mathbf{a} \in \mathbf{R}^d\}$$

and

$$\mathscr{D} = \operatorname{span}\{h(\cdot - \mathbf{a}) : \mathbf{a} \in \mathscr{A}\}$$

are finite-dimensional linear subspaces. Let $k = \dim \mathscr{C}$ and $l = \dim \mathscr{D}$. Thus $\mathbf{a}^1, \ldots, \mathbf{a}^k \in \mathbf{R}^d$ exist such that

$$\mathscr{C} = \operatorname{span}\{h(\cdot - \mathbf{a}^i) : i = 1, \ldots, k\}.$$

The functions $\{h(\cdot - \mathbf{a}^i)\}_{i=1}^k$ are linearly independent. Thus distinct $\{\mathbf{x}^j\}_{j=1}^k$ in $\mathbf{R}^d$ exist such that

$$(*) \qquad \det(h(\mathbf{x}^j - \mathbf{a}^i))_{i,j=1}^k \neq 0.$$

If $\mathscr{C} \neq \mathscr{D}$, i.e., $k > l$, then a vector $\mathbf{c} = (c_1, \ldots, c_k) \neq \mathbf{0}$ exists such that

$$\sum_{j=1}^{k} c_j h(\mathbf{x}^j - \mathbf{a}) = 0$$

for all $\mathbf{a} \in \mathscr{A}$. Set

$$p(\mathbf{a}) = \sum_{j=1}^{k} c_j h(\mathbf{x}^j - \mathbf{a}).$$

Then p is a nontrivial polynomial ($\mathbf{c} \neq \mathbf{0}$ and $(*)$) which vanishes on $\mathscr{A}$. This is a contradiction to our assumption on $\mathscr{A}$. Thus we must have $\mathscr{C} = \mathscr{D}$. ∎

We were originally partially motivated by the question of when, for some fixed continuous function h,

$$\mathscr{H} = \operatorname{span}\{l(h(\cdot - \mathbf{a})): l \in C(\mathbf{R}), \mathbf{a} \in \mathbf{R}^d\}$$

is dense in the space of continuous real-valued functions defined on $\mathbf{R}^d$, endowed with the topology of uniform convergence on compact sets. In this topology

$$\bar{\mathscr{H}} = \overline{\operatorname{span}}\{(h(\cdot - \mathbf{a}))^k: \mathbf{a} \in \mathbf{R}^d, k \in \mathbf{Z}_+\}.$$

However, we certainly do not need to take powers (or polynomials) of the $h(\cdot - \mathbf{a})$. Other methods may be used.

It is known, see [1], that if $\sigma \in C(\mathbf{R})$, then

$$\operatorname{span}\{\sigma(c \cdot + d): c, d \in \mathbf{R}\}$$

is dense in the space of continuous real-valued functions defined on $\mathbf{R}$, endowed with the topology of uniform convergence on compact sets, if and only if σ is *not* a polynomial. As such we have the following.

Proposition 12. *If $h \in C(\mathbf{R}^d)$, then, for any fixed $\sigma \in C(\mathbf{R})$ which is not a polynomial,*

$$\overline{\operatorname{span}}\{l(h(\cdot - \mathbf{a})): l \in C(\mathbf{R}), \mathbf{a} \in \mathbf{R}^d\} = \overline{\operatorname{span}}\{\sigma(ch(\cdot - \mathbf{a}) + d): c, d \in \mathbf{R}, \mathbf{a} \in \mathbf{R}^d\}$$

in the space of continuous real-valued functions defined on $\mathbf{R}^d$, endowed with the topology of uniform convergence on compact sets.

Proof. Obviously

$$\overline{\operatorname{span}}\{\sigma(ch(\cdot - \mathbf{a}) + d): c, d \in \mathbf{R}, \mathbf{a} \in \mathbf{R}^d\} \subseteq \overline{\operatorname{span}}\{l(h(\cdot - \mathbf{a})): l \in C(\mathbf{R}), \mathbf{a} \in \mathbf{R}^d\}.$$

Choose $f \in \overline{\operatorname{span}}\{l(h(\cdot - \mathbf{a})): l \in C(\mathbf{R}), \mathbf{a} \in \mathbf{R}^d\}$. For any compact subset K of $\mathbf{R}^d$, and $\varepsilon > 0$ an $m \in \mathbf{Z}_+$, $l_i \in C(\mathbf{R})$, and $\mathbf{a}^i \in \mathbf{R}^d$, $i = 1, \ldots, m$, exist such that

$$\left| f(\mathbf{x}) - \sum_{i=1}^{m} l_i(h(\mathbf{x} - \mathbf{a}^i)) \right| < \varepsilon$$

for all $\mathbf{x} \in K$. Now

$$\{h(\mathbf{x} - \mathbf{a}^i): \mathbf{x} \in K\} \subseteq [\alpha_i, \beta_i]$$

for some $-\infty < \alpha_i \leq \beta_i < \infty$, $i = 1, \ldots, m$. Since σ is not a polynomial

$$\operatorname{span}\{\sigma(c \cdot + d): c, d \in \mathbf{R}\}$$

is dense in the space of continuous real-valued functions defined on $[\alpha_i, \beta_i]$, $i = 1, \ldots, m$. Thus, for each i,

$$\left| l_i(t) - \sum_{j=1}^{n_i} b_{ij}\sigma(c_j t + d_j) \right| < \frac{\varepsilon}{m}$$

for all $t \in [\alpha_i, \beta_i]$, and some choice of n_i, b_{ij}, c_j, and d_j. It now follows that

$$\left| f(\mathbf{x}) - \sum_{i=1}^{m} \sum_{j=1}^{n_i} b_{ij}\sigma(c_j h(\mathbf{x} - \mathbf{a}^i) + d_j) \right| < 2\varepsilon$$

for all $\mathbf{x} \in K$. ∎

Acknowledgments. The research of A. Pinkus was supported by the fund for the promotion of research at the Technion. This paper has gone through numerous changes and revisions. We discussed the contents of this and the previous versions with many people, and learnt from many of them. Professor V. P. Palamodov is responsible for the idea used in Lemma 7 and Proposition 8 which substantially simplified one of our previous proofs. We would also like to thank Professors C. de Boor, V. Ya. Lin, and the referee.

References

1. L. Leshno, V. Ya. Lin, A. Pinkus, S. Schocken (1993): *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function.* Neural Networks, **6**:861–867.
2. A. G. Vitushkin, G. M. Henkin (1967): *Linear superpositions of functions.* Uspekhi Mat. Nauk, **22**:77–124. (English translation (1967): Russian Math. Surveys, **22**:77–125.)

A. Pinkus
Department of Mathematics
Technion
32000 Haifa
Israel

B. Wajnryb
Department of Mathematics
Technion
32000 Haifa
Israel