UNIVERSITY OF WISCONSIN-MADISON
CENTER FOR THE MATHEMATICAL SCIENCES

On the error in multivariate polynomial interpolation

C. de Boor [1]

Abstract. Simple proofs are provided for two properties of a new multivariate polynomial interpolation scheme, due to Amos Ron and the author, and a formula for the interpolation error is derived and discussed.

AMS (MOS) Subject Classifications: 41A63, 41A05; 41A10

Author's affiliation and address:
Center for the Mathematical Sciences
University of Wisconsin-Madison
610 Walnut St.
Madison WI 53705

[1] supported by the National Science Foundation under Grant No. DMS-9000053 and by the United States Army under Contract No. DAAL03-90-G-0090

On the error in multivariate polynomial interpolation

C. de Boor [1]

Dedicated to Garrett Birkhoff on the occasion of his 80th birthday

In interpolation, one hopes to determine, for g defined (at least) on a given pointset Θ, a function f from a given collection F which agrees with g on Θ. If, for arbitrary g, there is exactly one $f \in F$ with $f = g$ on Θ, then one calls the pair $\langle F, \Theta\rangle$ **correct**. (Birkhoff [Bi79] and others would say that, in this case, the problem of interpolating from F to data on Θ is **well set**.) Assuming that F is a finite-dimensional linear space, correctness of $\langle F, \Theta\rangle$ is equivalent to having

$$\dim F = \#\Theta = \dim F_{|\Theta} \tag{0.1}$$

(with $F_{|\Theta} := \{f_{|\Theta} : f \in F\}$ the set of restrictions of $f \in F$ to Θ).

Multivariate interpolation has to confront what one might call 'loss of Haar', i.e., the fact that, for every linear space F of continuous functions on $\mathbb{R}^d$ with $d > 1$ and $1 < \dim F < \infty$, there exist pointsets $\Theta \subset \mathbb{R}^d$ with $\dim F = \#\Theta > \dim F_{|\Theta}$. This observation rests on the following argument (see, e.g., the cover of [L66] or p.25 therein): For any basis $\Phi = (\phi_1, \ldots, \phi_n)$ for F, and any continuous curve $\gamma : [0\,\mathrm{d}1] \to (\mathbb{R}^d)^n : t \mapsto (\gamma_1(t), \ldots, \gamma_n(t))$, the function $g : t \mapsto \det(\phi_j(\gamma_i(t)))$ is continuous. Since $n > 1$ and $d > 1$, we can so choose the curve γ that, e.g., $\gamma(1) = (\gamma_2(0), \gamma_1(0), \gamma_3(0), \ldots, \gamma_n(0))$, while, for any t, the n entries of $\gamma(t)$ are pairwise distinct. Since then $g(1) = -g(0)$, we must have $g(t) = 0$ for some $t \in [0\,\mathrm{d}1]$, hence F is of dimension $< n$ when restricted to the corresponding pointset $\Theta := \{\gamma_1(t), \ldots, \gamma_n(t)\}$.

As a consequence, it is not possible for $n, d > 1$ (as it is for $n = 1$ or $d = 1$) to find an n-dimensional space of continuous functions which is correct for every n-point set $\Theta \in \mathbb{R}^d$. Rather, one has to choose such a correct interpolating space in dependence on the pointset.

A particular choice of such a *polynomial* space Π_Θ for given Θ has recently been proposed in [BR90], a list of its many properties has been offered and proved in [BR90-92], its computational aspects have been detailed in [BR91], and its generalization, from interpolation at a set of n points in $\mathbb{R}^d$ to interpolation at n arbitrary linearly independent linear functionals on the space

$$\Pi = \Pi(\mathbb{R}^d)$$

of all polynomials on $\mathbb{R}^d$, has been treated in much detail in [BR92].

The present short note offers some discussion concerning the error in this new polynomial interpolation scheme, and provides a short direct proof of two relevant properties of the interpolation scheme, whose proof was previously obtained, in [BR91-92], as part of more general results.

1. The interpolation scheme

To recall from [BR90-91], the interpolation scheme is stated in terms of a pairing, between Π and the space

$$A_0$$

of all functions on $\mathbb{R}^d$ with a convergent power series (in fact, the larger space Π' of all formal power series would work as well). Here is the pairing:

$$\langle p, g\rangle := p(D)g(0) = \sum_{\alpha\in\mathbb{Z}_+^d} D^\alpha p(0)\ D^\alpha g(0)\ /\alpha!. \tag{1.1}$$

The sum is over all **multi-indices** α, i.e., over all d-vectors with nonnegative integer entries, D^α denotes the partial derivative $D_1^{\alpha(1)}\cdots D_d^{\alpha(d)}$, and $\alpha! := \alpha(1)!\cdots\alpha(d)!$. Further, we will use the standard abbreviation

$$|\alpha| := \alpha(1)+\cdots+\alpha(d), \quad \alpha\in\mathbb{Z}_+^d,$$

and the nonstandard, but convenient, notation

$$()^\alpha : \mathbb{R}^d \to \mathbb{R} : x \mapsto x^\alpha := x(1)^{\alpha(1)}\cdots x(d)^{\alpha(d)}$$

for the α-**power function** or α-**monomial**. E.g., if p is the polynomial $\sum_\alpha c(\alpha)()^\alpha$, then $p(D)$ is the constant coefficient differential operator $\sum_\alpha c(\alpha)D^\alpha$.

The pairing is set up so that *the linear functional*

$$\delta(\theta) : \Pi \to \mathbb{R} : p \mapsto p(\theta)$$

of evaluation at θ is represented with respect to this pairing by the **exponential with frequency** $\theta\in\mathbb{R}^d$, i.e., by the function

$$e_\theta : \mathbb{R}^d \to \mathbb{R} : x \mapsto e^{\theta x}$$

(with $\theta x := \sum_j \theta(j)x(j)$ the usual scalar product). Indeed, since $D^\alpha e_\theta(0) = \theta^\alpha$, one computes

$$\langle p, e_\theta\rangle = \sum_\alpha D^\alpha p(0)\ \theta^\alpha/\alpha! = p(\theta). \tag{1.2}$$

Further, the pairing is graded, in the following sense. For $g\in A_0 \supset \Pi$ and $k\in\mathbb{Z}_+$, denote by

$$g^{[k]} := \sum_{|\alpha|=k} D^\alpha g(0)()^\alpha/\alpha!$$

its kth order term, i.e., the sum of all terms of exact order k in its power series expansion. In these terms, $p^{[k]}$ interacts in $\langle p, g\rangle$ only with the corresponding $g^{[k]}$. In particular, if we denote by $p_\uparrow$ the **leading** term of $p\in\Pi$, i.e., the nonzero term $p^{[k]}$ with maximal k, then

$$\langle p_\uparrow, p\rangle = \langle p_\uparrow, p_\uparrow\rangle > 0,$$

except when $p = 0$, in which case, by convention, $p_\uparrow := 0$. Correspondingly, we denote by $g_\downarrow$ the **least** or **initial** term of $g \in A_0$, i.e., the nonzero term $g^{[k]}$ with minimal k, and conclude correspondingly that

$$\langle g_\downarrow, g\rangle = \langle g_\downarrow, g_\downarrow\rangle > 0,$$

except when $g = 0$, in which case, by convention, $g_\downarrow := 0$.

Set now (as in [BR90])

$$\Pi_\Theta := \mathrm{span}\{g_\downarrow : g \in \mathrm{Exp}_\Theta\}, \quad \text{with } \mathrm{Exp}_\Theta := \mathrm{span}\{e_\theta : \theta \in \Theta\}.$$

Then $\Pi_\Theta \to \mathbb{R}^\Theta : p \mapsto p_{|\Theta}$ is 1-1: For, if $p_{|\Theta} = 0$, then, by (1.2), $\langle p, g\rangle = 0$ for all $g \in \mathrm{Exp}_\Theta$. If now $p \neq 0$, then necessarily $p_\uparrow = g_\downarrow$ for some $g \in \mathrm{Exp}_\Theta \backslash 0$, and then $0 = \langle p, g\rangle = \langle p_\uparrow, g_\downarrow\rangle = \langle p_\uparrow, p_\uparrow\rangle > 0$, a contradiction. It follows that $\dim \Pi_\Theta = \dim(\Pi_\Theta)_{|\Theta} \leq \#\Theta$.

On the other hand, it is possible to show, by a variant of the Gram-Schmidt orthogonalization process started from the basis $(e_\theta)_{\theta\in\Theta}$ for Exp_Θ, the existence of a sequence $(g_1, \dots, g_n)$ in Exp_Θ (with $n := \#\Theta$), for which

$$\langle g_{i\downarrow}, g_j\rangle = 0 \iff i \neq j. \tag{1.3}$$

This shows, in particular, that $(g_{1\downarrow}, \dots, g_{n\downarrow})$ is independent (and in Π_Θ), hence that $\dim \Pi_\Theta \geq n = \#\Theta$.

Consequently, $\langle \Pi_\Theta, \Theta\rangle$ is correct. More than that, *for arbitrary* $f \in \Pi$,

$$I_\Theta f := \sum_j g_{j\downarrow} \frac{\langle f, g_j\rangle}{\langle g_{j\downarrow}, g_j\rangle} \tag{1.4}$$

is the unique element in Π_Θ *which agrees with* f *at* Θ. Indeed, it follows from (1.3) that, for $i = 1, \dots, n$, $\langle I_\Theta f, g_i\rangle = \langle f, g_i\rangle$. Since $(g_1, g_2, \dots, g_n)$ is independent (by (1.3)), hence a basis for Exp_Θ (as Exp_Θ is spanned by n elements), we conclude with (1.2) that, for all $\theta \in \Theta$, $I_\Theta f(\theta) = \langle I_\Theta f, e_\theta\rangle = \langle f, e_\theta\rangle = f(\theta)$.

Remark. Since Exp_Θ, as the span of $n = \#\Theta$ functions, is trivially of dimension $\leq \#\Theta$, we seem to have just proved that

$$\dim \mathrm{Exp}_\Theta \;=\; \#\Theta. \tag{1.5}$$

This is misleading, though, since the proof of the existence of that sequence $(g_1, g_2, \dots, g_n)$ in Exp_Θ satisfying (1.3) uses (1.5).

Note that, with $g_j =: \sum_{\theta\in\Theta} B(j,\theta) e_\theta$,

$$\langle f, g_j\rangle = \sum_{\theta\in\Theta} B(j,\theta) f(\theta). \tag{1.6}$$

Thus, with (1.6) as a definition for $\langle f, g_j\rangle$ in case $f \notin \Pi$, (1.4) provides the polynomial interpolant from Π_Θ at Θ to arbitrary f defined (at least) on Θ.

2. Simple proofs of some properties of I_Θ

As shown in [BR90-92], the interpolation scheme I_Θ has many desirable properties. Some of these follow directly from the definition of Π_Θ: For example, $\Pi_\Theta \subset \Pi_{\Theta'}$ in case $\Theta \subset \Theta'$ (leading to a Newton form for the interpolant). Also, for any $r > 0$ and any $c \in \mathbb{R}^d$, $\Pi_{r\Theta+c} = \Pi_\Theta$. The translation-invariance, $\Pi_{\Theta+c} = \Pi_\Theta$, implies that Π_Θ is D-invariant, i.e.,

$$\forall\{p \in \Pi_\Theta, \alpha \in \mathbb{Z}_+^d\} \;\; D^\alpha p \in \Pi_\Theta. \tag{2.1}$$

Further, for any invertible matrix C, $\Pi_{C\Theta} = \Pi_\Theta \circ C^t$ (with C^t the transposed of C). Also, Π_Θ depends continuously on Θ (to the extent possible, limits on this being imposed by 'loss of Haar'), and I_Θ converges to appropriate Hermite interpolation if elements of Θ are allowed to coalesce in a sufficiently nice manner.

Perhaps the two most striking properties are that (i) I_Θ is degree-reducing, and (ii) $\Pi_\Theta = \cap_{p_{|\Theta}=0} \ker p_\uparrow(D)$. These properties are proved in [BR90-92] as part of more general results. Because of the evident and expected importance of these results, it seems useful to provide direct proofs, which I now do.

The **minimum-degree property**:

$$\forall\{p \in \Pi\} \;\; \deg I_\Theta p \le \deg p, \tag{2.2}$$

follows immediately from (1.4) since $\langle p, g_j\rangle = 0$ whenever $\deg p < \deg g_{j\downarrow}$. (It is stressed in [Bi79] that univariate Lagrange interpolation has this property.)

In fact, *the inequality in (2.2) is strict if and only if* $p_\uparrow \perp \Pi_\Theta$, as will be established during the proof of the second property. Here and below, I find it convenient to write $p \perp G$ (and say that 'p is perpendicular to G') in case $\langle p, g\rangle = 0$ for all $g \in G$, with $p \in \Pi$ and $G \subset A_0$.

(2.3)Proposition ([BR91-92]). $\Pi_\Theta = \cap_{p_{|\Theta}=0} \ker p_\uparrow(D)$.

Proof: I begin with a proof of the following string of equivalences and implications:

$$\begin{aligned} p_{|\Theta} = 0 \quad &\Longleftrightarrow \quad p \perp \mathrm{Exp}_\Theta \\ &\Longrightarrow \quad p_\uparrow \perp \Pi_\Theta \\ &\Longleftrightarrow \quad \forall\{q \in \Pi_\Theta\}\; p_\uparrow(D)q(0) = 0 \\ &\Longleftarrow \quad \forall\{q \in \Pi_\Theta\}\forall\{\alpha\}\; p_\uparrow(D)D^\alpha q(0) = 0 \\ &\Longleftrightarrow \quad \forall\{q \in \Pi_\Theta\}\; p_\uparrow(D)q = 0. \end{aligned} \tag{2.4}$$

The first equivalence follows from (1.2), the second relies on the definition of orthogonality, and the third uses the facts that $p(D)D^\alpha = D^\alpha p(D)$ (for any $p \in \Pi$), and that the polynomial $p_\uparrow(D)q$ is the zero polynomial iff all its Taylor coefficients are zero.

The '$\Longrightarrow$' follows from the observation that if $\langle p, g\rangle = 0$, then $\langle p_\uparrow, g_\downarrow\rangle = 0$, either because $\deg p_\uparrow \ne \deg g_\downarrow$, or else because, in the contrary case, $\langle p, g\rangle = \langle p_\uparrow, g_\downarrow\rangle$. Finally, the '$\Longleftarrow$' is trivial.

The '$\Longleftarrow$' can actually be replaced by '$\Longleftrightarrow$' since Π_Θ is D-invariant, by (2.1). Also, the '$\Longrightarrow$' can be reversed in the following way:

$$\forall\{\Pi \ni f \perp \Pi_\Theta\}\ \exists\{p \perp \mathrm{Exp}_\Theta\}\ \ p_\uparrow = f_\uparrow. \tag{2.5}$$

For, if f is a polynomial perpendicular to Π_Θ, of degree k say, then $I_\Theta f$ is necessarily of degree $< k$, since, in the formula (1.4), the terms $\langle f, g_j\rangle$ for $\deg g_{j\downarrow} > k$ are trivially zero while, for $\deg g_{j\downarrow} = k$, we have $\langle f, g_j\rangle = \langle f, g_{j\downarrow}\rangle$ and this vanishes since $f \perp \Pi_\Theta$. Consequently $p := f - I_\Theta f$ is a polynomial with the same leading term as f and perpendicular to Exp_Θ.

In any case, the argument given so far shows that $\Pi_\Theta \subset \cap_{p_{|\Theta}=0} \ker p_\uparrow(D)$. To show equality, note that $\dim \Pi_\Theta = \#\Theta < \infty$, hence $\Pi_\Theta \subset \Pi_k$ for some k. Thus, for any $|\alpha| = k+1$, $\deg I_\Theta()^\alpha < \deg()^\alpha = k+1$, hence

$$\big(()^\alpha - I_\Theta()^\alpha\big)_\uparrow = ()^\alpha,$$

therefore $\bigcap_{p_{|\Theta}=0} \ker p_\uparrow(D) \subset \bigcap_{|\alpha|=k+1} \ker D^\alpha = \Pi_k \subset \Pi$. Further, if $q \in \Pi$, then $p := q - I_\Theta q$ is a polynomial of degree $\le \deg q$ (by (2.2)) and $p_\uparrow(D)(\Pi_\Theta) = 0$ (since $p_{|\Theta} = 0$), therefore $p_\uparrow(D)p = p_\uparrow(D)q$. Hence, if $q \in \cap_{p_{|\Theta}=0} \ker p_\uparrow(D)$, then $p_\uparrow(D)p = 0$, hence $p = 0$, i.e., $q \in \Pi_\Theta$. □

3. Error

The standard error formula for univariate polynomial interpolation is based on the Newton form, i.e., on the 'correction' term $[\theta_1, \theta_2, \ldots, \theta_k, x]f\ \prod_{j=1}^k (\cdot - x_j)$ which is added to the polynomial interpolating to f at $\theta_1, \theta_2, \ldots, \theta_k$ in order to obtain the polynomial interpolating to f at $\theta_1, \theta_2, \ldots, \theta_k, x$. An analogous formula is available for the error $f - I_\Theta f$ in our multivariate polynomial interpolant. For its description, it is convenient to use the **dual of I_Θ with respect to the pairing (1.1)**, i.e., the map

$$I_\Theta^* : A_0 \to A_0 : g \mapsto \sum_j g_j \frac{\langle g_{j\downarrow}, g\rangle}{\langle g_{j\downarrow}, g_j\rangle}.$$

(3.1)Proposition. *For any $x \in \mathbb{R}^d$ and any $f \in \Pi$,*

$$f(x) - (I_\Theta f)(x) \ = \ \langle f, \varepsilon_{\Theta,x}\rangle \tag{3.2}$$

with

$$\varepsilon_{\Theta,x} := e_x - I_\Theta^* e_x = e_x - \sum_j g_j \frac{\langle g_{j\downarrow}, e_x\rangle}{\langle g_{j\downarrow}, g_j\rangle}. \tag{3.3}$$

Proof: Since e_x represents the linear functional $\delta(x)$ of evaluation at x with respect to (1.1), $I_\Theta^* e_x$ is the exponential which represents the linear functional $\delta(x) I_\Theta$ with respect (1.1). □

(3.4)Corollary. *The exponential $\varepsilon_{\Theta,x}$ represents $\delta(x)$ on the ideal*

$$\mathrm{ideal}(\Theta) := \ker I_\Theta = \{f \in \Pi : f_{|\Theta} = 0\},$$

and is orthogonal to Π_Θ, hence so are all its homogeneous components $\varepsilon^{[k]}_{\Theta,x}$.

Proof: If $f_{|\Theta} = 0$, then $f(x) = f(x) - I_\Theta f(x) = \langle f, \varepsilon_{\Theta,x}\rangle$.

Since I^*_Θ is the dual to the linear projector of interpolation from Π_Θ, its interpolation conditions are of the form $\langle p, \cdot\rangle$ with $p \in \Pi_\Theta$. Hence $\varepsilon_{\Theta,x}$, as the error $e_x - I^*_\Theta e_x$, must be perpendicular to Π_Θ, and this, incidentally, can also be written as

$$p(D)\varepsilon_{\Theta,x}(0) = 0, \quad \forall\, p \in \Pi_\Theta.$$

Finally, since Π_Θ is spanned by homogeneous polynomials, $f \perp \Pi_\Theta$ implies that $f^{[k]} \perp \Pi_\Theta$ for all $k \in \mathbb{Z}_+$. □

(3.5)Lemma ([BR90]). *For any $x \notin \Theta$, $\Pi_{\Theta\cup x} = \Pi_\Theta + \mathrm{span}\{p_{\Theta,x}\}$, with*

$$p_{\Theta,x} := (e_x - I^*_\Theta e_x)_\downarrow \tag{3.6}$$

the initial term of $\varepsilon_{\Theta,x}$.

Proof: First, $p_{\Theta,x} \in \Pi_{\Theta\cup x}$ since it is the initial term of some element of $\mathrm{Exp}_{\Theta\cup x}$. Further, $p_{\Theta,x} \neq 0$ since $p_{\Theta,x} = 0$ would imply that $e_x \in \mathrm{Exp}_\Theta$, hence $x \in \Theta$ by (1.5). Therefore

$$0 < \langle p_{\Theta,x}, p_{\Theta,x}\rangle = \langle p_{\Theta,x}, \varepsilon_{\Theta,x}\rangle = \langle p_{\Theta,x} - I_\Theta p_{\Theta,x}, e_x\rangle = (p_{\Theta,x} - I_\Theta p_{\Theta,x})(x), \tag{3.7}$$

showing that $p_{\Theta,x} - I_\Theta p_{\Theta,x} \neq 0$, hence $p_{\Theta,x} \notin \Pi_\Theta$. □

(3.8)Corollary. *For any ordering $\Theta = (\theta_1, \theta_2, \ldots, \theta_n)$ and with $\Theta_j := (\theta_1, \theta_2, \ldots, \theta_j)$,*

$$I_\Theta f = \sum_{j=1}^{n} \bigl(p_{\Theta_j} - I_{\Theta_{j-1}} p_{\Theta_j}\bigr) \frac{\langle f, \varepsilon_{\Theta_j}\rangle}{\langle p_{\Theta_j}, p_{\Theta_j}\rangle}. \tag{3.9}$$

Proof: The proof is by induction on $\#\Theta$, starting with the case $n = 0$, i.e., $\Theta = \{\}$, for which the definition $I_{\{\}} := 0$ is suitable. For any finite Θ and $x \notin \Theta$ and any f, we know from the lemma that

$$p \;:=\; I_\Theta f + \bigl(p_{\Theta,x} - I_\Theta p_{\Theta,x}\bigr) \frac{\langle f, \varepsilon_{\Theta,x}\rangle}{\langle p_{\Theta,x}, p_{\Theta,x}\rangle} \tag{3.10}$$

is in $\Pi_{\Theta\cup x}$, and from (3.7) and Proposition 3.1, that $p(x) = f(x)$, while evidently $p = f$ on Θ, hence p must be the polynomial $I_{\Theta\cup x} f$. Thus if (3.9) holds for Θ, it also holds for $\Theta \cup x$. □

Such a **Newton form** for $I_\Theta f$ was derived in a somewhat different manner in [BR90]. Note that

$$q_{\Theta,x} := (p_{\Theta,x} - I_\Theta p_{\Theta,x})/\langle p_{\Theta,x}, p_{\Theta,x}\rangle$$

is the unique element of $\Pi_{\Theta\cup x}$ which vanishes at Θ and takes the value 1 at x. But there does not appear to be in general (as there is in the univariate case) a scaling $sq_{\Theta,x}$ which makes its coefficient $\langle f, \varepsilon_{\Theta,x}\rangle/s$ in (3.10) independent of the way $\Theta \cup x$ has been split into Θ and x. The only obvious exception to this is the case when $\Pi_\Theta = \Pi_k :=$ the collection of polynomials of total degree $\le k$. Thus, only for this case does one obtain from I_Θ a ready multivariate divided difference.

Unless the ordering $(\theta_1, \theta_2, \ldots, \theta_n)$ is carefully chosen (e.g., as in the algorithm in [BR91]), there is no reason for the corresponding sequence $(\deg p_{\Theta_1}, \ldots, \deg p_{\Theta_n})$ to be nondecreasing. In particular, $\deg p_{\Theta,x}$ may well be smaller than $\deg I_\Theta f$. For example, if x is not in the affine hull of Θ, then $\deg p_{\Theta,x} = 1$. This means that the **order of the interpolation error**, i.e., the largest integer k for which $f(x) - I_\Theta f(x) = 0$ for all $f \in \Pi_{<k}$, may well change with x, since it necessarily equals $\deg p_{\Theta,x}$. The only exception to this occurs when $\Pi_\Theta = \Pi_k$ for some k. More generally, $\deg p_{\Theta,x}$ is a continuous function of x, hence constant, in some neighborhood of the point ξ if the pointset $\Theta \cup \xi$ is **regular** in the sense of [BR90], i.e., if

$$\Pi_{<k} \subseteq \Pi_{\Theta\cup\xi} \subset \Pi_k$$

for some k. (To be precise, [BR90] calls $\mathrm{Exp}_{\Theta\cup\xi}$ rather than $\Theta \cup \xi$ regular in this case.)

(3.11)Proposition. $k := \deg p_{\Theta,x} = \min\{\deg p : p(x) \ne 0, p \in \mathrm{ideal}(\Theta)\}$.

Proof: Let $k' := \min\{\deg p : p(x) \ne 0, p \in \mathrm{ideal}(\Theta)\}$. If $p \in \mathrm{ideal}(\Theta)$, then $p(x) = \langle p, \varepsilon_{\Theta,x}\rangle$ by Corollary 3.4, therefore $p(x) \ne 0$ implies $k \le \deg p$. Thus $k \le k'$. On the other hand, $q := p_{\Theta,x} - I_\Theta p_{\Theta,x}$ has degree $\le \deg p_{\Theta,x}$ (by (2.2); in fact, we already know from the proof of Proposition 2.3 that $\deg q = \deg p_{\Theta,x}$, but we don't need that here) and does not vanish at x, by (3.7), hence also $k' \le \deg q \le k$. □

The derivation from (3.2) of useful error *bounds* requires suitable bounds for expressions like

$$\sum_{|\alpha|\ge k} |D^\alpha f(0)|^2/\alpha!$$

in terms of norms like $\sum_{|\alpha|=k} \|D^\alpha f\|(L_p(B))$, with $k = \deg p_{\Theta,x}$ and B containing $\Theta \cup x$. Presumably, one would first shift the origin to lie in B, in order to keep the constants small, and so as to benefit from the fact that $\varepsilon_{\Theta,x}$ vanishes to order k at 0.

In view of Proposition 2.3, *integral representations* for the interpolation error $f - I_\Theta f$ should be obtainable from the results of K. Smith, [K70], using as differential operators the collection $p_\uparrow(D)$, $p \in P$, with P a minimal generating set for ideal(Θ).

4. A generalization and Birkhoff's ideal interpolation schemes

In [Bi79], Birkhoff gives the following abstract description of interpolation schemes. With X some space of function on some domain T into some field F and closed under pointwise multiplication, and Φ a collection of functionals (i.e., F-valued functions) on X, there is associated the **data map**

$$\delta(\Phi) : X \to F^{\Phi} : g \mapsto (\phi g)_{\phi \in \Phi}$$

(for which Birkhoff uses the letter α). Birkhoff calls any right inverse I of $\delta(\Phi)$ an **interpolation scheme on** Φ. (To be precise, Birkhoff talks about maps $I : \Phi \to X$ which are to be right inverses for $\delta(\Phi)$, and uses F^Y with $Y \subset T$ as an example for Φ, but the intent is clear.) He observes that $P := I\delta(\Phi)$ is necessarily a projector, i.e., idempotent.

He calls the pair $(\delta(\Phi), I)$, or, better, the resulting projector $P := I\delta(\Phi)$, an **ideal** interpolation scheme in case

(i) $\delta(\Phi)I = \mathrm{id}$;
(ii) both $\delta(\Phi)$ and I are linear (hence P is linear);
(iii) $\ker P$ is an **ideal**, i.e., closed under pointwise multiplication by any element from X.

For linear $\delta(\Phi)$ and I, $(\delta(\Phi), I)$ is ideal if and only if $\ker \delta(\Phi)$ is an ideal (since $\ker P = \ker \delta(\Phi)$ regardless of I). Thus any linear scheme for which the data map is a restriction map $f \mapsto f_{|\Theta}$ (such as the map I_Θ discussed in the preceding sections) is trivially ideal.

In these terms, the generalization of I_Θ treated in [BR92] deals with the situation when $T = \mathbb{R}^d$ and $X = \Pi = \Pi(\mathbb{R}^d)$, and $\Phi : f \mapsto (\phi f)_{\phi \in \Phi}$ for some finite, linearly independent, collection of linear functionals on Π (with a further extension, to infinite Φ, also analysed). The algebraic dual Π' can be represented by the space of formal power series (in d indeterminates), and the pairing (1.1) has a natural extension to $\Pi \times \Pi'$.

In this setting,

$$\ker \delta(\Phi) = \Lambda_\perp := \{p \in \Pi : p \perp \Lambda\},$$

with

$$\Lambda := \mathrm{span}\, \Phi.$$

(4.1)Proposition ([BR92]). $\ker \delta(\Phi)$ *is an ideal if and only if* Λ *is* D*-invariant.*

The proof uses nothing more than the observation that

$$\langle ()^\alpha p, \phi \rangle = \langle p, D^\alpha \phi \rangle.$$

As an example, if Φ is a linearly independent subset of $\cup_{\theta \in \Theta} e_\theta \Pi$, then $\phi \in \Phi$ is of the form

$$f \mapsto p(D) f(\theta)$$

for some $\theta \in \Theta$ and $p \in \Pi$. Correspondingly, $\Lambda = \mathrm{span}\, \Phi = \sum_{\theta \in \Theta} e_\theta P_\theta$ for certain polynomial spaces P_θ. Hence, $\ker \delta(\Phi)$ is an ideal iff each P_θ is D-invariant. In particular, Hermite interpolation at finitely many points is ideal, while G.D. Birkhoff interpolation is, in general, not.

References

[B] G. Birkhoff (1979), "The algebra of multivariate interpolation", in *Constructive approaches to mathematical models* (C. V. Coffman and G. J. Fix, eds), Academic Press (New York), 345–363.

[B] C. de Boor and A. Ron (1990), "On multivariate polynomial interpolation", *Constr. Approx.* **6**, 287–302.

[BR91] C. de Boor and Amos Ron (199x), "Computational aspects of polynomial interpolation in several variables", *Math. Comp.* **xx**, xxx–xxx.

[BR92] C. de Boor and A. Ron (1992), "The least solution for the polynomial interpolation problem", *Math. Z.* **xx**, xxx–xxx.

[L66] G. G. Lorentz (1966), *Approximation of Functions*, Holt, Rinehart and Winston (New York).

[K70] K. T. Smith (1970), "Formulas to represent functions by their derivatives", *Math. Ann.* **188**, 53–77.