On interpolation by radial polynomials

C. de Boor

Happy 60th and beyond, Charlie!

Abstract A lemma of Micchelli's, concerning radial polynomials and weighted sums of point evaluations, is shown to hold for arbitrary linear functionals, as is Schaback's more recent extension of this lemma and Schaback's result concerning interpolation by radial polynomials. Schaback's interpolant is explored.

In his most-cited paper, [M], Micchelli supplies the following interesting auxiliary lemma (his Lemma 3.1).

(1) Lemma. *If* $\sum_{i=1}^n c_i p(x_i) = 0$ *for all* $p \in \Pi_{<k}(\mathbb{R}^d)$, *then*

$$(-1)^k \sum_{i=1}^n \sum_{j=1}^n c_i c_j \|x_i - x_j\|^{2k} \ge 0,$$

where equality holds if and only if

$$\sum_{i=1}^n c_i p(x_i) = 0, \quad p \in \Pi_{\le k}(\mathbb{R}^d).$$

Here, $(x_1, \ldots, x_n)$ is a sequence in $\mathbb{R}^d$, and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$. Further, with

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

a convenient if nonstandard notation for the power function,

$$\Pi_{\le k} := \Pi_{\le k}(\mathbb{R}^d) := \text{span}(()^\alpha : \alpha \in \mathbb{Z}_+^d, |\alpha| \le k)$$

is the collection of all polynomials in d real variables of (total) degree $\le k$, and

$$\Pi_{<k} := \Pi_{<k}(\mathbb{R}^d)$$

those of degree $< k$. The collection of all polynomials in d real variables will be denoted here, correspondingly, by $\Pi = \Pi(\mathbb{R}^d)$.

Micchelli follows the lemma by the following

(2) Remark. *Applying the Lemma inductively shows that the conditions*

$$\sum_{i=1}^n c_i p(x_i) = 0, \quad p \in \Pi_{\le k}(\mathbb{R}^d),$$

and

$$(-1)^k \sum_{i=1}^n \sum_{j=1}^n c_i c_j q(\|x_i - x_j\|^{2k}) = 0, \quad q \in \Pi_{\le k}(\mathbb{R}),$$

are equivalent.

The essence of the proof is, perhaps, the observation (implicit in Micchelli's proof) that

$$\begin{aligned} \|x-y\|^{2k} &= (\|x\|^2 - 2\sum_i x(i)y(i) + \|y\|^2)^k \\ &= \sum_{a+|\beta|+c=k} \frac{k!}{a!\beta!c!} \|x\|^{2a}(-2)^{|\beta|} x^\beta y^\beta \|y\|^{2c} \\ &= \sum_{a+b+c=k} (-2)^b \sum_{|\beta|=b} \frac{k!}{a!\beta!c!} p_{a,\beta}(x) p_{c,\beta}(y), \end{aligned} \tag{3}$$

with

$$p_{a,\beta}(x) := \|x\|^{2a} x^\beta$$

and with a, b, c nonnegative integers, $\beta \in \mathbb{Z}_+^d$, and

$$|\beta| := \sum_j \beta(j), \quad \beta! := \beta(1)! \cdots \beta(d)!.$$

Each summand in the final sum of (3) is thus the product of a (homogeneous) polynomial in x of degree $2a+b$ and a (homogeneous) polynomial in y of degree $2c+b$. Hence, if

$$\lambda \perp \Pi_{<k},$$

i.e., λ is a linear functional on Π that vanishes on $\Pi_{<k}$, then the tensor product of λ with itself, i.e., the linear map

$$\lambda \otimes \lambda : \Pi \otimes \Pi \to \mathbb{R} : ()^\alpha \otimes ()^\beta \mapsto \lambda()^\alpha \, \lambda()^\beta,$$

annihilates all the summands with $2a+b<k$ *or* $2c+b<k$. As to any other summands, they must have $2a+b \ge k$ *and* $2c+b \ge k$, hence

$$2k = 2(a+b+c) = 2a+b+2c+b \ge 2k,$$

therefore

$$2a+b = k = 2c+b.$$

These are the summands in which the polynomial in x equals the polynomial in y and, moreover, $k-b$ is even. Thus, altogether (and in faulty but understandable notation),

$$(\lambda \otimes \lambda)\|x-y\|^{2k} = (-1)^k \sum_{k-b \text{ even}} 2^b \frac{k!}{(((k-b)/2)!)^2} \sum_{|\beta|=b} (\lambda p_{(k-b)/2,\beta})^2/\beta!.$$

In particular, $(\lambda\otimes\lambda)((-1)^k\|x-y\|^{2k})\ge 0$ with equality if and only if λ vanishes on the sequence

$$(4)\qquad (p_{(k-b)/2,\beta} : |\beta| = b; b\in \mathbb{Z}_+,\ k-b\ \text{ even}).$$

But since each $p_{(k-b)/2,\beta}$ here is homogeneous of degree k while, in particular, $k-b$ is even when $b=k$ hence each $()^\beta$ with $|\beta|=k$ appears in (4), this last condition is equivalent to having λ vanish on $\mathrm{span}(()^\beta : |\beta| = k, \beta\in\mathbb{Z}_+^d)$, hence on $\Pi_{\le k}$.

Altogether, this proves the following generalization of Micchelli's Lemma and Remark.

(5) Proposition. *If $\lambda\perp\Pi_{<k}$, then*

$$(-1)^k(\lambda\otimes\lambda)\|x-y\|^{2k}\ge 0$$

with equality iff $\lambda\perp\Pi_{\le k}$.

(6) Corollary. $\lambda\perp\Pi_{\le k}$ *if and only if*

$$(\lambda\otimes\lambda)\|x-y\|^{2r}=0,\quad 0\le r\le k.$$

Note the following useful

(7) Corollary. *The bilinear form*

$$\langle,\rangle_k : \Pi'\otimes\Pi' : (\lambda,\mu)\mapsto (-1)^k(\lambda\otimes\mu)\|x-y\|^{2k}$$

on the algebraic dual Π' of Π is an inner product on any algebraic complement of

$$\perp\Pi_{\le k} := \{\mu\in\Pi' : \mu\perp\Pi_{\le k}\}$$

in $\perp\Pi_{<k}$.

Schaback [S] reiterates Micchelli's results and extends them as follows (Lemmata 8 and 9 of [S], though only for linear functionals λ that are linear combinations of point evaluations).

(8) Lemma. *If $\lambda\perp\Pi_{\le k}$, then, for all $k\le 2\ell$,*

$$x\mapsto\lambda\|x-\cdot\|^{2\ell}$$

has degree $<2\ell-k$.

Conversely, if, for some $k\le 2\ell$,

$$x\mapsto\lambda\|x-\cdot\|^{2\ell}$$

has degree $<2\ell-k$, then $\lambda\perp\Pi_{\le k}$.

To be sure, the first assertion follows directly from the basic identity (3) since, by that identity, application of such λ to $\|x-\cdot\|^{2\ell}$ kills all summands with $2c+|\beta|\le k$, leaving only those with $2\ell-2a-|\beta| = 2c+|\beta|>k$, i.e., with $2\ell-k>2a+|\beta|$.

For the second assertion, simply "apply the same idea as in the proof of Micchelli's lemma", to quote [S]. Arguing perhaps differently, rewrite (3) in terms of polynomial degrees in x to get

$$\|x-y\|^{2\ell} = \sum_{j=0}^{2\ell} \sum_{2a+b=j} (-2)^b \sum_{|\beta|=b} \frac{\ell!}{a!\beta!c!} p_{a,\beta}(x) p_{c,\beta}(y).$$

Now, $(-2)^b = (-1)^j 2^b$ since $j-b$ here is always even. Also (using $a+b+c=\ell$), $p_{a,\beta} = \|\cdot\|^{2(j-\ell)} p_{c,\beta}$. Therefore,

$$\|x-y\|^{2\ell} = \sum_{j=0}^{2\ell} (-1)^j \|x\|^{2(j-\ell)} \sum_{2a+b=j} 2^b \sum_{|\beta|=b} \frac{\ell!}{a!\beta!c!} p_{c,\beta}(x) p_{c,\beta}(y).$$

Hence, if now $x \mapsto \lambda\|x-\cdot\|^{2\ell}$ is of degree $< 2\ell-k$, then each of the sums

$$\sum_{2a+b=j} 2^b \sum_{|\beta|=b} \frac{\ell!}{a!\beta!c!} p_{c,\beta} \lambda p_{c,\beta}, \quad 2\ell-k \le j,$$

must be zero, hence so must be the value of λ on each such sum, i.e.,

$$0 \;=\; \sum_{2a+b=j} 2^b \sum_{|\beta|=b} \frac{\ell!}{a!\beta!c!} (\lambda p_{c,\beta})^2, \quad 2\ell-k \le j.$$

This implies that

$$\lambda p_{c,\beta} = 0, \quad |\beta| = b,\ 2\ell-k \le 2a+b,\ \ell = a+b+c,\ 0 \le a,b,c,$$

hence, for the particular choice $\ell = a+b$, therefore $c=0$ and $2\ell-2a-b \le k$, i.e., $b \le k$,

$$\lambda p_{0,\beta} = 0, \quad |\beta| \le k.$$

But that says that $\lambda \perp \Pi_{\le k}$.

In what follows, it is convenient to consider, for any sequence or indexed 'set' $(v_j : j \in J)$ of vectors in some linear space Y over the scalar field $\mathbb{F}$, the corresponding map

$$[v_j : j \in J] : \mathbb{F}_0^J \to Y : c \mapsto \sum_{j \in J} c(j) v_j,$$

with $\mathbb{F}_0^J$ denoting all scalar-valued functions on J with finite support. Note that the v_j enter the description of this linear map in exactly the manner in which the columns of an (m,n)-matrix A enter the description of the corresponding linear map $\mathbb{F}^n \to \mathbb{F}^m : c \mapsto Ac$, hence it seems reasonable to call $[v_j : j \in J]$ the **column map with columns** v_j. We can think of $[v_j : j \in J]$ as a row, much as we can think of a matrix as the row of its columns.

The sequence $(v_j : j \in J)$ is a basis for Y exactly when $[v_j : j \in J]$ is invertible, in which case one might just as well refer to the latter as a basis for Y. Further, if $(\lambda_i : i \in I)$ is an indexed 'set' in the dual, Y', of Y, then it is convenient to consider the corresponding map

$$[\lambda_i : i \in I]^{\mathrm{t}} : Y \to \mathbb{F}^I : y \mapsto (\lambda_i y : i \in I),$$

calling it the **row map with rows** λ_i (hence the use of the transpose sign) since the λ_i enter the description of this linear map in exactly the same manner in which the rows of the transpose A^{t} (i.e., the columns of A) of an (m, n)-matrix A enter the description of the corresponding linear map $\mathbb{F}^m \to \mathbb{F}^n : c \mapsto A^{\mathrm{t}}c$. We can think of $[\lambda_i : i \in I]^{\mathrm{t}}$ as a column, much as we can think of a matrix as the column of its rows. With this, we are ready to think, as we may, of the composition of such a row map with such a column map as a matrix, i.e., the **Gram matrix** or **Gramian**

$$[\lambda_i : i \in I]^{\mathrm{t}}[v_j : j \in J] = (\lambda_i v_j : i \in I, j \in J).$$

This use of the superscript $^{\mathrm{t}}$ also seems consistent with the notation $x^{\mathrm{t}}y := \sum_{j=1}^{d} x(j)y(j)$ for the scalar product of $x, y \in \mathbb{R}^d$, given that it is standard to think of the elements of $\mathbb{R}^d$ as columns. For completeness, we note that the composition $[v_j : j \in J][\lambda_i : i \in I]^{\mathrm{t}}$ makes sense only when $I = J$, in which case it is a linear map from Y to Y and the most general such in case Y is finite-dimensional.

Schaback [S] considers, for the n-dimensional space M spanned by evaluation at the elements of the given n-set X in $\mathbb{R}^d$, a basis $\Lambda = [\lambda_1, \ldots, \lambda_n]$ **graded (by degree)** (he calls any such a 'discrete moment basis') in the sense that the sequence $(\kappa_i : i = 1, \ldots, n)$, with

$$\kappa_i := \max\{k : \lambda_i \perp \Pi_{<k}\}, \quad \text{all } i,$$

is nondecreasing, and, for each k, $[\lambda_i : \kappa_i \geq k]$ is a basis for $\mathrm{M} \cap \perp\Pi_{<k}$.

One readily obtains such a basis from any particular basis $[\mu_1, \ldots, \mu_n]$ for M, by applying Gauss elimination with row interchanges to the Gram matrix

$$(\mu_i()^\alpha : i = 1, \ldots, n, \alpha \in \mathbb{Z}_+^d) = [\mu_1, \ldots, \mu_n]^{\mathrm{t}}V, \tag{9}$$

with the columns of

$$V := [()^\alpha : \alpha \in \mathbb{Z}_+^d] \;=:\; [()^{\alpha_j} : j = 1, 2, \ldots]$$

so ordered that $j \mapsto |\alpha_j|$ is nondecreasing.

Indeed, Gauss elimination applied to an onto Gram matrix such as (9) can be interpreted as providing an invertible matrix L (the product of a permutation matrix with a lower triangular matrix) and thereby the basis

$$[\lambda_1, \ldots, \lambda_n] := [\mu_1, \ldots, \mu_n](L^{-1})^{\mathrm{t}}$$

for M, and a subsequence $(\beta_1, \ldots, \beta_n)$ of $(\alpha_j : j = 1, 2, \ldots)$ so that, for each i, the first nonzero entry in row i of $[\lambda_1, \ldots, \lambda_n]^{\mathrm{t}} V$, i.e., in $\lambda_i V$, is the β_ith. Since, by assumption, the map $j \mapsto |\alpha_j|$ is nondecreasing, this implies that

$$\kappa_i = \max\{k : \lambda_i \perp \Pi_{<k}\} = |\beta_i|, \quad i = 1, \ldots, n,$$

and, in particular, $i \mapsto \kappa_i$ is nondecreasing. Further, if $\sum_i c(i)\lambda_i \perp \Pi_{<k}$, then, since $\kappa_i = |\beta_i|$, also $\sum_{|\beta_i|<k} c(i)\lambda_i \perp \Pi_{<k}$. But this implies that

$$[c(i) : |\beta_i| < k] B = 0,$$

with the matrix

$$B := (\lambda_i ()^{\beta_j} : |\beta_i|, |\beta_j| < k)$$

square upper triangular with nonzero diagonal entries, hence invertible, and therefore $c(i) = 0$ for $|\beta_i| < k$. This proves that, for each k, $[\lambda_i : |\beta_i| \geq k]$ is a basis for $\mathrm{M} \cap \perp\Pi_{<k}$.

Schaback then considers the polynomials

$$w_j : x \mapsto \lambda_j \|x - \cdot\|^{2\kappa_j}, \quad j = 1, \ldots, n.$$

Note that, by (8)Lemma,

$$\deg w_j = 2\kappa_j - \kappa_j = \kappa_j.$$

This implies that the Gram matrix

$$\Lambda^{\mathrm{t}} W = (\lambda_i w_j : i, j = 1, \ldots, n)$$

is block upper triangular since

$$(\Lambda^{\mathrm{t}} W)(i,j) = \lambda_i w_j = (\lambda_i \otimes \lambda_j)\|x - y\|^{2\kappa_j} = (-1)^{\kappa_j} \langle \lambda_i, \lambda_j \rangle_{\kappa_j}$$

is zero as soon as $\kappa_i > \kappa_j$. Further, for each k and with

$$I_k := \{i : \kappa_i = k\},$$

the diagonal block

$$\Lambda^{\mathrm{t}} W(I_k, I_k) = (-1)^k (\langle \lambda_i, \lambda_j \rangle_k : i, j \in I_k)$$

is invertible, by (7)Corollary and the fact that $[\lambda_i : i \in I_k]$ is a basis for an algebraic complement of $\perp\Pi_{\leq k}$ in $\perp\Pi_{<k}$. We will use later that this implies that W is a **graded** basis for

$$F := \operatorname{ran} W := \{\sum_j a(j) w_j : a \in \mathbb{R}^n\}$$

in the sense that $j \mapsto \deg w_j$ is nondecreasing and, for each k, $[w_j : \deg w_j < k]$ is a basis for $F \cap \Pi_{<k}$.

For the moment, we only use the conclusion (which Schaback draws for the case that M is spanned by point evaluations) that $\Lambda^{\mathrm{t}}W$ is invertible, hence

$$P = P_S := W(\Lambda^{\mathrm{t}}W)^{-1}\Lambda^{\mathrm{t}}$$

is the linear projector that associates with each $p \in \Pi$ the unique element $f \in F$ that **matches** p **at** M in the sense that

$$\mu f = \mu p, \quad \mu \in \mathrm{M}.$$

Schaback also observes that P_S is of minimal degree in the sense that F minimizes

$$\deg G := \max\{\deg g : g \in G\}$$

among all polynomial subspaces G that are **correct for** M in the sense that, for every $p \in \Pi$, they contain a unique match at M.

However, much more is true. P_S is of minimal degree in the strong sense (of, e.g., [BR2]) that it is **degree-reducing**, meaning that

$$\deg P_S p \le \deg p, \quad p \in \Pi.$$

This condition is shown, in [BR2], to be equivalent to the following property more explicitly associated with the words 'minimal degree':

Definition. *The finite-rank linear projector P on Π is* **of minimal degree** :=

$$\dim(G \cap \Pi_{<k}) \le \dim(\operatorname{ran} P \cap \Pi_{<k}), \quad k \in \mathbb{N},$$

for all linear subspaces G of Π that are correct for $\perp \ker P$.

Proposition. *P_S is of minimal degree.*

Proof: Let G be a linear subspace of Π correct for M. Then, G is n-dimensional and, for any bases Λ of M and W of G, respectively, the Gramian $\Lambda^{\mathrm{t}}W$ is invertible.

Choose, in particular, Λ to be a graded basis for M and W to be a graded basis for G. Then, for any $k \in \mathbb{N}$, the first $\dim(G \cap \Pi_{<k}) = \#\{j : \deg w_j < k\}$ columns of the Gramian $\Lambda^{\mathrm{t}}W = (\lambda_i w_j)$ have nonzero entries only in the first $\#\{i : \kappa_i < k\} = n - \#\{i : \kappa_i \ge k\} = n - \dim(\mathrm{M} \cap \perp\Pi_{<k})$ rows. The invertibility of the Gramian therefore implies that

$$\dim(G \cap \Pi_{<k}) \le n - \dim(\mathrm{M} \cap \perp\Pi_{<k}).$$

On the other hand, there is equality here when $G = \operatorname{ran} P_S$ since, as observed earlier, Schaback's w_j form a graded basis while $\kappa_i = \deg w_i$, all i, hence $n - \dim(\mathrm{M} \cap \perp\Pi_{<k}) = \#\{i : \kappa_i < k\} = \#\{j : \deg w_j < k\}$. $\square$

The proof shows that a linear projector P on Π is of minimal degree if and only if

$$\dim(\operatorname{ran} P \cap \Pi_{<k}) \;+\; \dim(\perp \ker P \cap \perp\Pi_{<k}) = \dim \operatorname{ran} P, \quad \text{all } k.$$

The polynomial interpolant $P_S p$ to p at M is, in general (see below), not the least interpolant $P_{BR} p$ of [BR2] to p at M, hence we are free to give it a name, and **Schaback interpolant** seems entirely appropriate (hence the suffix S).

We now compare P_S and P_{BR} in the specific context of [S], i.e., when M is spanned by evaluation on some n-set X in $\mathbb{R}^d$. In addition to being of minimal degree, each interpolant is constant in any direction perpendicular to the affine hull

$$\flat X$$

of X (or, **flat** spanned by X), i.e., both satisfy

$$Pf(x) = (Pf)(P_X x),$$

with P_X the orthoprojector of $\mathbb{R}^d$ onto $\flat X$.

For the Schaback interpolant, this follows from the fact that, for any x, and any $y \in \flat X$,

$$\|x - y\|^2 = \|P_X x - y\|^2 + \|x - P_X x\|^2,$$

hence, since $\lambda_j \perp \Pi_{<\kappa_j}$,

$$w_j(x) = \lambda_j(\|P_X x - \cdot\|^2 + \|x - P_X x\|^2)^{\kappa_j} = \lambda_j \|P_X x - \cdot\|^{2\kappa_j},$$

by (8)Lemma.

This readily implies that *both interpolants coincide in case X is contained in a 1-dimensional flat.*

Also, *both projectors commute with translation*, i.e.,

$$Pp(\cdot + y) = (Pp)(\cdot + y),$$

and interact with any unitary change of variables as follows:

$$Pp(A\cdot) = (Pp)(A^{\mathrm{t}}\cdot) \tag{10}$$

for all real unitary matrices A, as follows for Schaback's projector directly from the observation that, for all such A and any x and y,

$$\|x - Ay\|^2 = \|A(A^{\mathrm{t}}x - y)\|^2 = \|A^{\mathrm{t}}x - y\|^2.$$

However, while (10) holds for $P = P_{BR}$ and arbitrary invertible A (see, e.g., [BR1]), this is not in general so for $P = P_S$, due to the fact that it is the 'kernel' $(x, y) \mapsto \exp(x^{\mathrm{t}}y)$ on which P_{BR} is based rather than Schaback's $(x, y) \mapsto \|x - y\|^{2\ell}$. We would therefore expect the two interpolants in general to differ when $\dim \flat X > 1$ and $\operatorname{ran} P$ isn't just some $\Pi_{<k}$.

The simplest such example occurs when $X = \{x_1, \ldots, x_4\}$ is a 4-set in $\mathbb{R}^2$ that spans $\mathbb{R}^2$. In that case, the range of each projector is of the form

$$\Pi_{<2} + w\mathbb{R}$$

for some homogeneous quadratic polynomial w. For the least interpolant, [BR1] gives w as

$$x \mapsto \lambda(x^{\mathrm{t}} \cdot)^2,$$

with

$$\lambda : f \mapsto \sum_{j=1}^{4} a(j) f(x_j)$$

such that

$$\sum a(j) x_j = 0, \quad \sum_j a(j) = 0. \tag{11}$$

Since X spans $\mathbb{R}^2$, this says that $\lambda \perp \Pi_{<2}$ and, since $\#X = 4$, λ is, up to scalar multiples, the unique such element of M. That means that Schaback's w is the leading term of

$$x \mapsto \lambda \|x - \cdot\|^4.$$

Since $\lambda \perp \Pi_{<2}$, and $\|x - x_j\|^4 = (\|x\|^2 - 2x^{\mathrm{t}} x_j + \|x_j\|^2)^2$, this leading term is

$$x \mapsto \lambda(2\|x\|^2 \| \cdot \|^2 + 4(x^{\mathrm{t}} \cdot)^2),$$

and this is, offhand, not just a scalar multiple of the least's w. For, there is no reason to believe that $\lambda \| \cdot \|^2 = 0$.

E.g., with the specific choice

$$X = (0, \mathbf{i}_1, \mathbf{i}_2, z)$$

involving $\mathbf{i}_1 := (1, 0)$ and $\mathbf{i}_2 := (0, 1)$, we have

$$\lambda : f \mapsto f(z) - z(1) f(\mathbf{i}_1) - z(2) f(\mathbf{i}_2),$$

hence

$$\lambda \| \cdot \|^2 = \|z\|^2 - z(1) - z(2) = z(1)(z(1) - 1) + z(2)(z(2) - 1),$$

and this is zero only for special choices of z. To be sure, it is zero when $z = \mathbf{i}_1 + \mathbf{i}_2$, i.e., in case of gridded data, giving us then bilinear interpolation.

To be sure, graded bases of the space M, of linear functionals at which to match given values by polynomials, have been used before in multivariate polynomial interpolation. For example, the multivariate 'finite differences' introduced and used in [SX] are easily seen to form such a graded basis.

In particular, the construction of the least interpolant makes use of a graded basis Λ for M (constructed by a more stable variant of Gauss elimination, namely Gauss elimination 'by segments') but obtains the polynomial space G of interpolants as the span of the 'leasts' of the λ_i. To recall (e.g., from [BR2]), any $\lambda \in \Pi'$ is uniquely representable, with respect to the bilinear form

$$\mathbb{R}[\![x]\!] \otimes \Pi \to \mathbb{R} : (f,p) \mapsto \sum_{\alpha \in \mathbb{Z}_+^d} D^\alpha(f)(0) D^\alpha p(0)/\alpha!,$$

by the formal power series

$$\hat{\lambda} := \sum_{k=0}^{\infty} \hat{\lambda}^{[k]},$$

with

$$\hat{\lambda}^{[k]} := \sum_{|\alpha|=k} (\lambda()^\alpha)()^\alpha/\alpha!, \quad \text{all } k.$$

Then

$$\kappa := \max\{k : \lambda \perp \Pi_{<k}\} = \min\{k : \hat{\lambda}^{[k]} \neq 0\}$$

is known as the **order** of λ, and $\hat{\lambda}^{[\kappa]}$ is, by definition, the **least** of λ (with κ taken to be -1 when $\lambda = 0$ and, correspondingly, $\hat{\lambda}^{[-1]} = 0$). It is easy to see that $G = \operatorname{span}\{\hat{\lambda}_i^{[\kappa_i]} : i = 1, \dots, n\}$ depends only on M and not on the particular graded basis Λ for M used, and is spanned by homogeneous polynomials, hence, equivalently, is dilation-invariant. For the most striking properties of G (such as a finite list of constant coefficient differential operators whose joint kernel is G), see [BR1] or [BR2].

Acknowledgements Thanks are due to Tomas Sauer for a constructive reading of what I thought was the final draft. Further, on receiving a preprint of the present note, Robert Schaback informed me that, in the meantime, he, too, had extended (5)–(8) to arbitrary linear functionals, albeit with different proofs.

References

[BR1] C. de Boor and A. Ron (1990), "On multivariate polynomial interpolation", *Constr. Approx.* **6**, 287–302.

[BR2] C. de Boor and A. Ron (1992), "The least solution for the polynomial interpolation problem", *Math. Z.* **210**, 347–378.

[M] C. A. Micchelli (1986), "Interpolation of scattered data: distance matrices and conditionally positive definite functions", *Constr. Approx.* **2**, 11–22.

[SX] T. Sauer and Yuan Xu (1995), "On multivariate Lagrange interpolation", *Math. Comp.* **64**, 1147–1170.

[S] Robert Schaback (2002), "Multivariate Interpolation by Polynomials and Radial Basis Functions", ms, April. (available from Schaback's homepage)