# Ideal Interpolation

Carl de Boor

**Abstract.** A linear interpolation scheme is termed 'ideal' when its errors form a polynomial ideal. The paper surveys basic facts about ideal interpolation and raises some questions.

Ideal interpolation is, by definition, given by a linear projector on the space $\Pi$ of polynomials whose kernel is a polynomial ideal. It is therefore also any linear map, as used in algebra, that associates a polynomial with its normal form with respect to a polynomial ideal. This article lists (and mostly proves) basic facts about ideal interpolation and raises some questions.

## §1. Definition and Basic Algebraic Facts

If $P$ is a linear projector of finite rank on the linear space $X$ over the commutative field $\mathbb{F}$ with algebraic dual $X'$, then we can think of it as providing a linear interpolation scheme on $X$: For each $g \in X$, $f = Pg$ is the unique element of $\operatorname{ran} P := P(X)$ for which

$$\lambda f = \lambda g, \quad \forall \lambda \in \operatorname{ran} P' = \{\lambda \in X' : \lambda P = \lambda\},$$

with $P'$ the dual of $P$, i.e., the linear map $X' \to X' : \lambda \mapsto \lambda P$. In other words, given that $\ker P := \{g \in X : Pg = 0\} = \operatorname{ran}(\mathrm{id} - P)$,

$$\operatorname{ran} P' = (\ker P)^{\perp} := \{\lambda \in X' : \ker P \subset \ker \lambda\},$$

the set of interpolation conditions matched by $P$. Not surprisingly, there are exactly as many independent conditions as there are degrees of freedom, i.e.,

$$\dim \operatorname{ran} P = \dim \operatorname{ran} P'.$$



Proceedings Title
XXX (eds.), pp. 59–91.

ISBN 0-9728482-x-x

Put into more practical terms, if the column maps

$$V : \mathbb{F}^n \to X : a \mapsto \sum_{j=1}^{n} v_j a(j) =: [v_1, \ldots, v_n]a$$

and

$$\Lambda : \mathbb{F}^n \to X' : a \mapsto \sum_{j=1}^{n} \lambda_j a(j) =: [\lambda_1, \ldots, \lambda_n]a,$$

into $X$ and $X'$ respectively, are such that their Gram matrix

$$\Lambda^{\mathrm{t}} V := (\lambda_i v_j : i, j = 1{:}n)$$

is invertible, then, in particular, both $V$ and $\Lambda$ are 1-1, hence bases for their respective ranges and there is, for given $b \in \mathbb{F}^n$, exactly one element, call it $Va$, of $\operatorname{ran} V$ that satisfies the equation

$$\Lambda^{\mathrm{t}}(Va) = b,$$

thus giving rise to the map

$$P = V(\Lambda^{\mathrm{t}} V)^{-1} \Lambda^{\mathrm{t}}$$

on $X$, evidently a linear projector, that associates $g \in X$ with the unique element $f = Pg$ in $\operatorname{ran} V = \operatorname{ran} P$ for which

$$\Lambda^{\mathrm{t}} f := (\lambda_i f : i = 1{:}n)$$

agrees with $\Lambda^{\mathrm{t}} g$, hence $\lambda f = \lambda g$ for all $\lambda \in \operatorname{ran} \Lambda = \operatorname{ran} P'$.

Consider now, in particular, the linear space

$$\Pi \subset (\mathbb{F}^d \to \mathbb{F})$$

of all $\mathbb{F}$-valued polynomials in $d$ real ($\mathbb{F} = \mathbb{R}$) or complex ($\mathbb{F} = \mathbb{C}$) variables. It will be important that $\Pi$ is also a ring under pointwise multiplication,

$$(pq)(x) := p(x)q(x), \quad p, q \in \Pi, \ x \in \mathbb{C}^d.$$

In [Bi], Garrett Birkhoff defined ideal interpolation as a linear projector $P$ on $\Pi$ whose nullspace or kernel is an ideal, i.e., not only closed under addition and multiplication by scalars but also under (pointwise) multiplication by arbitrary polynomials. *Lagrange interpolation* is mentioned by Birkhoff as a particular example. However, ideal projectors are already looked at carefully in [M76], where they are called 'Hermite interpolation'.

Ideal projectors are, in a sense, aware of the multiplicative structure of $\Pi$, hence we would expect insights from considering their interaction with multiplication, as exhibited by the following very handy fact.

**Lemma 1.1 ([B03]).** *A linear projector* $P$ on $\Pi$ is ideal if and only if*

$$P(pq) = P(pPq), \quad \forall p, q \in \Pi. \tag{1.2}$$

**Proof:** The condition (1.2) is equivalent to having

$$P(\Pi(\mathrm{id} - P)(\Pi)) = \{0\},$$

and, since $P$ is a linear projector hence $(\mathrm{id} - P)(\Pi) = \ker P$, this is equivalent to

$$\Pi \ker P \subset \ker P,$$

hence, given that $\ker P$ is a linear subspace, to $\ker P$ being an ideal. □

An ideal projector is completely determined by its action on a subspace only slightly larger than its range. This is readily seen by the following considerations.

Each ideal projector $P$ induces a map,

$$M : \Pi \to L(\mathrm{ran}\, P) : p \mapsto M_p, \tag{1.3}$$

on $\Pi$ into the space $L(\mathrm{ran}\, P)$ of linear maps on $\mathrm{ran}\, P$, by the prescription

$$M_p : \mathrm{ran}\, P \to \mathrm{ran}\, P : f \mapsto P(pf), \quad p \in \Pi. \tag{1.4}$$

Indeed, $M_p$ so defined is a linear map on $\mathrm{ran}\, P$, and depends linearly on $p$, hence the map $M$ is well-defined and is linear. More than that, for arbitrary $p, q \in \Pi$ and $f \in \mathrm{ran}\, P$,

$$M_q M_p f - M_{qp} f = P(qP(pf)) - P(qpf) = 0,$$

the last equality by (1.2), hence $M$ is also a homomorphism, on the ring $\Pi$ into $L(\mathrm{ran}\, P)$ considered as a ring with respect to map composition as multiplication. Also, since $\Pi$ is a commutative ring, so is $\mathrm{ran}\, M$, even though it is a subring of the *non*commutative ring $L(\mathrm{ran}\, P)$.

The ring $\Pi$ is generated by the specific polynomials

$$()_j := ()^{\varepsilon_j}, \quad \varepsilon_j := (\delta_{jk} : k = 1{:}d), \quad j = 0{:}d,$$

with

$$()^\alpha : \mathbb{F}^d \to \mathbb{F} : x \mapsto x^\alpha := \prod_j x(j)^{\alpha(j)}, \quad \alpha \in \mathbb{Z}_+^d,$$

* dec05: sufficient to assume that $P$ is a linear map

a convenient if nonstandard notation for the monomials. Consequently, $\operatorname{ran} M$ is generated by the specific linear maps

$$M_j : \operatorname{ran} P \to \operatorname{ran} P : f \mapsto P(()_j f), \quad j = 0{:}d, \tag{1.5}$$

in terms of which

$$M_p = p(M) := \sum_\alpha \widehat{p}(\alpha)\, M^\alpha, \quad p \in \Pi,$$

with

$$p =: \sum_\alpha \widehat{p}(\alpha)()^\alpha,$$

and with

$$M^\alpha := \prod_j (M_j)^{\alpha(j)} = M_{()^\alpha}$$

independent of the order in which this product is formed from its factors. (Since the map $M$ cannot be composed with itself, hence a polynomial in $M$ makes no sense, it may be excusable to use, as I have done here, the notations $M^\alpha$ and $p(M)$ for a related but different purpose.)

By way of background, the transpose of the matrix representations of the $M_j$ with respect to a monomial basis for $\operatorname{ran} P$ (if any) are known as 'multiplication tables' and the maps $M_f$ as 'multiplication maps'; see [CLO98: p.51ff]. The latter term derives from the fact that it is customary (see, e.g., [CLO98] and [AS]) to think of $M$ as mapping into $L(\Pi/\ker P)$ (rather than into $L(\operatorname{ran} P)$) and, in that setting, $M_p$ models multiplication by $p + \ker P$ in the algebra $\Pi/\ker P$, i.e., carries the coset $q + \ker P$ to the coset $pq + \ker P = (p + \ker P)(q + \ker P)$.

It follows, directly from (1.2), that

$$p(M)P()_0 = P(p\,P()_0) = P(p\,()_0) = Pp, \quad p \in \Pi. \tag{1.6}$$

This representation of $P$ has been used in [B03] to uncover the close connection between the Opitz formula and the Leibniz formula for univariate divided differences and to prove such formulæ for certain multivariate divided differences.

**Proposition 1.7.** *If we know the ideal projector $P$ on $()_0$ and on*

$$\Pi_1(\operatorname{ran} P) := \sum_{j=0}^d ()_j \operatorname{ran} P,$$

*then we know $P$ everywhere.*

**Proof:** As soon as we know $P$ on $\Pi_1(\operatorname{ran} P)$, we can compute the linear maps $M_j$, hence can compute $p(M)$ for any $p \in \Pi$ and, with that, can determine $Pp$ from (1.6) provided we also know $P()_0$. □

**Example** As an example, consider the following situation, discussed in [Sh] in the bivariate case: $P$ is an ideal projector with range

$$F := \operatorname{ran}[()_1^j : j = 0{:}n{-}1],$$

and $\mathbb{F} = \mathbb{C}$ hence $()_1^n - P()_1^n$, considered as a univariate polynomial, has $n$ zeros counting multiplicities. Assume, finally, that these zeros are all simple, hence

$$(()_1^n - P()_1^n)(x) =: \prod_{j=1}^{n}(x(1) - \tau(j))$$

defines the sequence $\tau$ with pairwise distinct entries. Set

$$z_j := (\tau_j, (P()_2)(\tau_j), \ldots, (P()_d)(\tau_j)), \quad j = 1{:}n.$$

Then any $p \in F$ vanishing on $z$ is necessarily zero, hence since $z$ has $n$ entries and $\dim F = n$, there is, for each $p \in \Pi$, exactly one element of $F$, call it $Rp$, that agrees with $p$ on $z$. I claim that $R = P$ and, by Proposition 1.7, need to check this only for $()^\alpha$ with $\alpha(1) < n$, $\alpha(2{:}d) = (\delta_{ij} : j = 2{:}d)$, $i = 2{:}d$, since it is already evident for $\alpha = (n, 0, \ldots, 0)$, hence for $\alpha = (m, 0, \ldots, 0)$ for all $m \in \mathbb{N}$, by Proposition 1.7 (since $()_1^n$ spans an algebraic complement of $F$ in $\Pi_1(F)$ when considering only the ring of univariate polynomials). For the check, notice that

$$(R()_i)(z_j) = ()_i(z_j) = z_j(i) = (P()_i)(\tau_j) = (P()_i)(z_j),$$

hence $R = P$ on $()_i$ for $i = 2{:}d$. With that, for any $j$,

$$P(()_1^j()_i) = P(()_1^j P()_i) = R(()_1^j R()_i) = R(()_1^j()_i),$$

the middle equality since $P()_i = R()_i \in F$, while the other two equalities follow from $P$ and $R$ being ideal. □

## §2. A Basis for the Ideal $\ker P$

By (1.6), $\ker M \subset \ker P$, while, if $p \in \ker P$, then $p(M)f = P(pf) = P(fPp) = P0 = 0$ for all $f$ in $\operatorname{ran} P$ which is the domain of $p(M)$, hence then $p(M) = 0$. Thus, altogether,

$$\ker M = \ker P. \tag{2.1}$$

Hence, by Proposition 1.7, we should be able to derive $\ker P$ from $()_0 - P()_0$ and the action of the restriction

$$N := P|_{\Pi_1(F)}$$

of $P$ to $\Pi_1(F)$, with

$$F := \operatorname{ran} P.$$

**Proposition 2.2.** *If* $()_0 \in \operatorname{ran} P$, *then*

$$\ker P = \operatorname{ideal}(\ker N) =: \mathcal{I}. \tag{2.3}$$

**Proof:** Since $\ker N = \ker P \cap \Pi_1(F)$ and $\ker P$ is an ideal, we immediately have

$$\ker P \supseteq \mathcal{I}.$$

For the converse containment, let

$$\Pi_k(S) := \sum_{|\alpha|\le k} ()^\alpha S, \qquad \emptyset \ne S \subset \Pi.$$

Then, for any additive subset $S$ of $\Pi$, we have

$$\Pi_{r+s}(S) = \Pi_r(\Pi_s(S)).$$

In particular,

$$\Pi_k := \Pi_k(\mathbb{F}) = \Pi_1(\Pi_{<k}), \quad \text{with} \;\; \Pi_{<k} := \Pi_{k-1}.$$

Specifically, $\cup_k \Pi_k(F) = \Pi$ since we assumed $F = \operatorname{ran} P$ to contain $()_0$. Therefore, we know that $\ker P \subseteq \mathcal{I}$ once we show, by induction on $k$, that

$$p \in \ker P \cap \Pi_k(F) \quad \Longrightarrow \quad p \in \mathcal{I}.$$

For $k = 1$, this is so by definition of $\mathcal{I}$. So assuming it to hold for all $k < h$, let $p \in \ker P \cap \Pi_h(F)$. Then

$$p = \sum_{j=0:d} ()_j p_j$$

with $p_j \in \Pi_{<h}(F)$, hence $(\mathrm{id} - P)p_j$ is in $\Pi_{<h}(F) + F = \Pi_{<h}(F)$ as well as in $\ker P$, hence in $\mathcal{I}$ by induction hypothesis. Thus,

$$p \in \sum_j ()_j (Pp_j + \mathcal{I}) = \sum_j ()_j Pp_j + \mathcal{I},$$

while, by (1.2), $P\sum_j ()_j Pp_j = P\sum_j ()_j p_j = Pp = 0$, hence $\sum_j ()_j Pp_j \in \ker P \cap \Pi_1(F)$, therefore in $\mathcal{I}$. □

It follows that $\ker P$ is generated, as an ideal, by any (vector-space) basis for $\ker P \cap \Pi_1(F)$. Further, such a basis is readily obtained in the form

$$(b - Nb : b \in B),$$

with $B$ any basis for an algebraic complement of $F$ in $\Pi_1(F)$. As the example of bivariate tensor-product interpolation to gridded data shows, the resulting (ideal) basis may be far from minimal.

## §3. Mourrain's Condition

Proposition 2.2 (though not the proof given here) is essentially due to Mourrain [Mo] who proved it under the additional assumption that $F$ satisfy what I will call here

**(3.1) Mourrain's condition.** *For $f \in F$, $f \in \Pi_1(F \cap \Pi_{<\deg f})$; i.e., in Mourrain's words, $F$ is connected to 1.*

Here, as usual, for $p \in \Pi\backslash 0$,

$$\deg p := \min\{k : p \in \Pi_k\} = \max\{|\alpha| : \widehat{p}(\alpha) \neq 0\}, \quad \text{with } |\alpha| := \sum_j \alpha(j).$$

Mourrain's condition implies that $()_0 \in F$ but is, offhand, much stronger. For example, in the univariate case, (3.1) implies that $F = \Pi_k$ for some $k$, hence also that $F$ is $D$-invariant, i.e., closed under differentiation. See [B05b] for the fact that, in the multivariate case, (3.1) and $D$-invariance are not related.

Mourrain [Mo] investigates the following problem: *Given a finite-dimensional linear subspace $F$ of $\Pi$ and a linear projector $N$ on $\Pi_1(F)$ with range $F$, provide necessary and sufficient conditions on $N$ to be the restriction to $\Pi_1(F)$ of an ideal projector $P$ with range $F$.*

There is at most one such ideal projector since, by Proposition 2.2, its kernel is necessarily the ideal generated by $\ker N$. Mourrain shows the *existence* of such an ideal projector under the (obviously necessary) assumption that the linear maps

$$M_j : F \to F : f \mapsto N(()_j f), \quad j = 1{:}d,$$

commute, but only for an $F$ that satisfies (3.1).

**Theorem 3.2 ([Mo]).** *Let $F$ be a finite-dimensional linear subspace of $\Pi$ satisfying Mourrain's condition, (3.1). Let $N$ be a linear projector on $\Pi_1(F)$ with range $F$. Then, the following are equivalent:*

(a) *$N$ is the restriction to $\Pi_1(F)$ of an ideal projector $P$ with range $F$.*

(b) *The linear maps $M_j : F \to F : f \mapsto N(()_j f)$, $j = 1{:}d$, commute.*

*Further, if either holds, hence both hold, then* $\ker P = \text{ideal}(\ker N)$.

**Proof:** It only remains to prove that (b) implies (a). With the $M_j$ commuting, we can define

$$R : \Pi \to \Pi : p \mapsto p(M)()_0$$

and find it to be a linear map into $F$, but it is, offhand, not clear that it coincides with $N$ on $F$, nor that it is a projector.

To begin with, we know for sure that $R$ and $N$ agree on $\Pi_0 \subseteq F$. If $C$ is a linear subspace of $F$ for which we already know that $R = N$ on it, then, for any $f =: \sum_j ()_j c_j \in \Pi_1(C)$,

$$Nf = \sum_j N(()_j c_j) = \sum_j M_j c_j = \sum_j M_j c_j(M)()_0 = f(M)()_0 = Rf,$$

hence we also know it for $\Pi_1(C)$. So, starting with $C = \Pi_0$, we can iterate $C \leftarrow \Pi_1(C) \cap F$, and in this way generate an increasing sequence of subspaces. Since $F$ is finite-dimensional, this leads to the linear subspace $C_*$ of $F$ containing $()_0$ and satisfying $C_* = \Pi_1(C_*) \cap F$, and, on it, $R = N$, but it is not clear that $C_* = F$.

It is exactly this difficulty that Mourrain's condition, (3.1), is designed to deal with. For, Mourrain's condition certainly ensures that $C_* = F$, hence that $R$ extends $N$, i.e., $R = N$ on $\Pi_1(F)$. Since $\operatorname{ran} R \subset F \subset \Pi_1(F)$, this also implies that $R$ is a linear projector, with range $F$. $\square$

For a simple univariate example, consider $F = \operatorname{ran}[()^0, ()^2] \subset \Pi \subset (\mathbb{F} \to \mathbb{F})$, for which $\Pi_1(\operatorname{ran}[()^0]) \cap F = \operatorname{ran}[()^0]$, hence Mourrain's condition fails spectacularly. At the same time, let $N$ be the linear projector on $\Pi_1(F) = \Pi_3$ specified by

$$N((()^0, ()^1, ()^2, ()^3) = (()^0, ()^0, ()^2, 0).$$

$N$ is indeed a linear projector, with range equal to $F$, but $\ker N$ contains both $()^1 - ()^0$ and $()^3$ and, as these are relatively prime, $\operatorname{ideal}(\ker N) = \Pi$. Hence, while the $M_j$ trivially commute (there being only one), no extension of $N$ to an ideal projector exists.

To be sure, since the question of whether a projector is ideal only depends on its nullspace, it is easy to construct an ideal projector having this particular $F$ as its range. Simply take $\operatorname{ran} P' = \operatorname{ran}[\delta_0, \delta_1]$ (with $\delta_v : f \mapsto f(v)$). Then $N := P|_{\Pi_1(F)}$ is given by the recipe

$$N((()^0, ()^1, ()^2, ()^3) = (()^0, ()^2, ()^2, ()^2).$$

Now $\ker N = \operatorname{ran}[()^2 - ()^1, ()^3 - ()^2 = ()^1(()^2 - ()^1)]$, hence $\operatorname{ideal}(\ker N) = \operatorname{ideal}(()^2 - ()^1) = \ker P$. This confirms Proposition 2.2. In effect, $N$ has an extension to an ideal projector with the same range if and only if

$$F \cap \operatorname{ideal}(\ker N) = \{0\}.$$

See [B05b] for an example showing that, in Theorem 3.2, Mourrain's condition cannot be replaced by $D$-invariance.

As a historical aside, Hakopian and Tonoyan announced in 1998 (see, e.g., [HT98]) the following closely related result which is fully detailed and further extended in [HT02].

**Proposition.** *Let*

$$f_\alpha =: ()^\alpha - \sum_{\beta\in J} ()^\beta a_{\alpha,\beta}, \quad \alpha \in I,$$

*with* $I := \{\alpha : |\alpha| = k+1\}$, $J := \{\beta : |\beta| \le k\}$, *and let* $\mathcal{A}_j$ *be the matrices defined by the quadratic system equivalent to* $f_\alpha(x) = 0$, $\alpha \in I$, *namely*

$$\mathcal{A}_j \mathbf{x} = x(j)\mathbf{x}, \quad j = 1{:}d,$$

*with* $\mathbf{x} := (()^\beta : \beta \in J)$, *hence*

$$\mathcal{A}_j(\alpha,\beta) = \begin{cases} a_{\alpha,\beta}, & \alpha+\varepsilon_j \in J; \\ \delta_{\alpha+\varepsilon_j,\beta}, & \text{otherwise.} \end{cases}$$

*Then, the polynomial system* $f_\alpha(x) = 0$, $\alpha \in I$, *has at most* $\#J$ *solutions, with equality if and only if the matrices* $\mathcal{A}_j$ *commute and are diagonalizable.*

Hakopian and Tonoyan came to this result as part of their effort to derive, for a given system of partial differential equations, an equivalent first-order system; see, e.g., [HT04]. In that context, they trace the commuting condition back to Frobenius, [F]. They also prove this result in the more general context when $J$ is a 'lower' set (as defined in the next section).

To be sure, $\mathcal{A}_j$ is the transpose of the matrix representation of $M_j$ with respect to the monomial basis of $F = \Pi_J$; the eigenstructure of the $M_j$ is discussed in section 6.

## §4. Normal Forms

Mourrain's intent in [Mo] is to construct a convenient "normal form" for the ideal

$$\mathcal{I} := \text{ideal}(G)$$

generated by a given finite set $G$ of polynomials. This is a basic task in computational algebraic geometry (see, e.g., [CLO92] where the material discussed in this section can be found) and is traditionally performed with the aid of a Gröbner basis for the ideal. This, in turn, involves a so-called monomial order, i.e., an ordering $<$ on the set $\mathbb{Z}_+^d$ of multi-indices that respects addition, i.e.,

$$\forall \alpha,\beta,\gamma \in \mathbb{Z}_+^d \quad \alpha < \beta \quad \Longrightarrow \quad \alpha+\gamma < \beta+\gamma, \tag{4.1}$$

and is a well-ordering, meaning that every subset of $\mathbb{Z}_+^d$ has a smallest element. Standard examples are the Lexicographic Order (`lex`) in which

$\alpha < \beta$ means that the *first* nonzero entry in $\beta - \alpha$ is positive, and the Graded Reverse Lexicographic Order (grevlex) in which $\alpha < \beta$ if, either $|\alpha| < |\beta|$, or else $|\alpha| = |\beta|$ and the *last* nonzero entry in $\beta - \alpha$ is positive.

Any such ordering admits the definition of the corresponding polynomial degree:

$$\mathrm{Deg} : \Pi\backslash 0 \to \mathbb{Z}_+^d : p \mapsto \max \operatorname{supp} \widehat{p},$$

with (4.1) ensuring that

$$\mathrm{Deg}(pq) = \mathrm{Deg}(p) + \mathrm{Deg}(q). \tag{4.2}$$

Note that, in this, the degree of the zero polynomial is undefined. Perhaps a mathematically cleaner definition of $\mathrm{Deg}(p)$ would be the *set* $\{\alpha \in \mathbb{Z}_+^d : \alpha \le \max \operatorname{supp} \widehat{p}\}$ which now has the empty set as the natural definition of $\mathrm{Deg}(0)$ yet still satisfies (4.2) (since $A + \emptyset = \emptyset$).

With respect to such an ordering, one then constructs a Gröbner basis $G$ for $\mathcal{I}$, meaning that $G$ is a finite subset of $\mathcal{I}$ with the property that

$$\forall p \in \mathcal{I}, \quad p \in \sum_{g \in G} g\, \Pi_{\le \mathrm{Deg}(p) - \mathrm{Deg}(g)}.$$

Here and below, for any subset $\Gamma$ of $\mathbb{Z}_+^d$ (including subsets merely specified by the condition its elements are to satisfy),

$$\Pi_\Gamma := \operatorname{ran}[()^\gamma : \gamma \in \Gamma].$$

Actually, a simpler definition in use identifies a Gröbner basis for $\mathcal{I}$ as a finite subset $G$ of $\mathcal{I}$ with

$$\bigcup_{g \in G} (\mathrm{Deg}(g) + \mathbb{Z}_+^d) \supset \{\mathrm{Deg}(f) : f \in \mathcal{I}\} =: \mathrm{Deg}(\mathcal{I}).$$

Note that, directly from (4.2),

$$\mathrm{Deg}(\mathcal{I}) = \mathrm{Deg}(\mathcal{I}) + \mathbb{Z}_+^d,$$

showing $\mathrm{Deg}(\mathcal{I})$ to be an upper set. But (by Dickson's Lemma), any upper set $U$ in $\mathbb{Z}_+^d$ is necessarily of the form

$$U = \operatorname{extr}(U) + \mathbb{Z}_+^d,$$

with

$$\operatorname{extr}(U) := \{\alpha \in U : U\backslash\alpha \text{ is upper}\}$$

its necessarily finite set of extreme points. This proves the existence of Gröbner bases. A naive definition of the *normal form mod* $\mathcal{I}$ for $p \in \Pi$ is the element $r$ of $p+\mathcal{I}$ of minimal Deg. However, there is, offhand, nothing to prevent $\mathcal{I}$ from containing $f \neq 0$ with $\text{Deg}(f) < \text{Deg}(r)$, and then also $(r+f)/2$ is a different element of $p+\mathcal{I}$ of minimal degree.

So, a better definition is the following. The normal form mod $\mathcal{I}$ for $p \in \Pi$ is the unique element in

$$(p+\mathcal{I}) \cap \Pi_{\setminus \text{Deg}(\mathcal{I})}.$$

Indeed, if both $r$ and $s$ are in this intersection, then their difference is in $\mathcal{I}$, yet, if $r-s$ were nonzero, then $\text{Deg}(r-s) \notin \text{Deg}(\mathcal{I})$. This shows uniqueness.

As to existence, let

$$F := \Pi_{\setminus \text{Deg}(\mathcal{I})} \;=\; \text{ran}[()^\alpha : \alpha \notin \text{Deg}(\mathcal{I})].$$

Then, as we just pointed out, $F$ and $\mathcal{I}$ are linear subspaces of $\Pi$ with trivial intersection,

$$F \cap \mathcal{I} = \{0\}.$$

Further if, in the monomial order, the left shadow

$$\mathbb{Z}_{\leq\alpha} := \{\beta \in \mathbb{Z}_+^d : \beta \leq \alpha\}$$

of every $\alpha$ is finite (as is the case, e.g., in `grevlex`), then, for arbitrary $p \in \Pi$, the following elimination algorithm produces an $r \in F$ with $p - r \in \mathcal{I}$.

**Division by $G$.**
*Input:* $p \in \Pi$, $G$.
$r \leftarrow p$.
*for* $\alpha = \text{argmax}(\text{Deg}(G) \cap \text{supp}\,\widehat{r})$, *and* $g \in G$ *so that* $\alpha = \text{Deg}(g)$, $r \leftarrow r - (\widehat{r}(\alpha)/\widehat{g}(\alpha))g$.
*Output: The resulting $r$ is "the remainder of the division of $p$ by $G$".*

Indeed, for a monomial ordering such as `grevlex`, the entire calculation takes place on the *finite* index set $\mathbb{Z}_{\leq \text{Deg}(p)}$, hence necessarily stops after finitely many steps, at which point, assuming we chose $G$ to be $\mathcal{I}$, $r \in F$ while, at every step, $p - r \in \mathcal{I}$.

For a monomial ordering, such as `lex`, in which left shadows can be infinite, a more subtle argument is required to prove that, nevertheless, the elimination algorithm terminates in finitely many steps. This more subtle argument leads naturally to the creation of a Gröbner basis $G$ for $\mathcal{I}$ and its use in more refined versions of the elimination algorithm; see, e.g., [CLO92].

In any case, taking this for granted, we conclude that

$$\Pi = F \oplus \mathcal{I},$$

with the normal form for $p$ mod $\mathcal{I}$ nothing but the projection of $p$ to $F$ along $\mathcal{I}$, i.e., the image of $p$ under the ideal projector with range $F$ and kernel $\mathcal{I}$.

**Proposition 4.3.** *Each ideal $\mathcal{I}$ of finite codimension has a $D$-invariant algebraic complement spanned by monomials, hence also satisfying Mourrain's condition.*

**Proof:** Take for $F$ the space $F = \Pi_{\backslash \operatorname{Deg}(\mathcal{I})}$ just constructed, already shown to be an algebraic complement for $\mathcal{I}$. It is monomial, in the sense that it is spanned by monomials, but, with that, $F$ is also $D$-invariant, since $\operatorname{Deg}(\mathcal{I}) = \operatorname{Deg}(\mathcal{I}) + \mathbb{Z}_+^d$, hence

$$\alpha \notin \operatorname{Deg}(\mathcal{I}) \quad \Longrightarrow \quad (\alpha - \mathbb{Z}_+^d) \cap \operatorname{Deg}(\mathcal{I}) = \emptyset.$$

In other words, $\mathbb{Z}_+^d \backslash \operatorname{Deg}(\mathcal{I})$ (like the complement of any upper set) is a lower set. This also implies that $F$ satisfies Mourrain's condition (3.1). □

Now, Mourrain's point is that the construction of a Gröbner basis is, in general, time-consuming, as is working term by term. Can we, he asks (as have others before him), construct the normal form by some other, perhaps more efficient, way? If $G$ spans an algebraic complement of some polynomial space $F$ within $\Pi_1(F)$, and if this $F$ satisfies his condition (3.1) and is complementary to $\mathcal{I} = \operatorname{ideal}(G)$, then, as we saw, for any $p \in \Pi$, its normal form mod $\mathcal{I}$ is the polynomial $p(M)()_0$, with the $M_j$ determined as above from the linear projector $N$ on $\Pi_1(F)$ with range $F$ whose kernel is $\operatorname{span}(G)$.

Mourrain also investigates the question of just what to do if we have to start with some arbitrary finite $G$, and develops an algorithm for constructing an H-basis for $\mathcal{I} = \operatorname{ideal}(G)$, i.e., a finite subset $H$ of $\mathcal{I}$ for which $\{h_\uparrow : h \in H\}$ is a basis for the homogeneous ideal

$$\mathcal{I}_\uparrow := \operatorname{ideal}(p_\uparrow : p \in \mathcal{I}),$$

with $p_\uparrow$ (also denoted $L(p)$ or, in conflict with other notation used here, $\Lambda(p)$, and called the leading term of $p$) uniquely determined (for $p \neq 0$) by the requirements that it be homogeneous and satisfy

$$\deg(p - p_\uparrow) < \deg p.$$

Lack of time and space prevents me from pursuing this further here. For H-bases in connection with multivariate polynomial interpolation, see [B94], [MSa], [MSb], [MSc], [S98], [S01], [S02], [S05].

## §5. The Nature of $\operatorname{ran} P'$

We now take a look at the interpolation conditions for the ideal projector $P$, under the assumptions that $P$ is of finite rank and that $\mathbb{F} = \mathbb{C}$.

Polynomial ideals arise naturally in the study of the common zeros of a collection $G$ of polynomials, i.e., the set

$$\mathcal{V}(G) := \{v \in \mathbb{C}^d : g(v) = 0, g \in G\}.$$

Any finite weighted sum

$$\sum_{g \in G} a_g g$$

of elements $g$ of $G$ will have these same zeros, even if we use for the weights $a_g$ not just scalars but polynomials. In other words,

$$\mathcal{V}(G) = \mathcal{V}(\text{ideal}(G)).$$

To what an extent is an ideal $\mathcal{I}$ characterized by its variety, $\mathcal{V}(\mathcal{I})$? A partial answer is provided by

**Hilbert's Nullstellensatz.** *If $p \in \Pi$ vanishes on $\mathcal{V}(\mathcal{I})$, then some power of $p$ lies in $\mathcal{I}$.*

So, while there is no 1-1 correspondence between varieties and ideals, the connection is, nevertheless, quite close.

In particular, Hilbert's Nullstellensatz is a kind of multivariate fundamental theorem of algebra: for, if $\mathcal{V}(\mathcal{I})$ is empty, then, e.g., the polynomial $()_0$ vanishes on that variety, hence must be in $\mathcal{I}$, therefore so must be $()_0 \cdot \Pi = \Pi$. In other words, any proper ideal has zeros.

In particular, assuming our ideal projector, $P$, not to be trivial, its kernel

$$\mathcal{I} := \ker P$$

is a proper ideal, hence has zeros. Let

$$v \in \mathcal{V} := \mathcal{V}(\mathcal{I}).$$

This says that the linear functional

$$\delta_v : p \mapsto p(v)$$

vanishes on $\mathcal{I} = \ker P$, hence is in $\operatorname{ran} P'$, i.e., provides an interpolation condition for $P$. More than that,

$$[\delta_v : v \in \mathcal{V}]$$

is 1-1, hence,

$$\#\mathcal{V} \le \dim \Pi/\mathcal{I}. \tag{5.1}$$

But, and this is a subtlety, there need not be equality here. This is already hinted at by Hilbert's Nullstellensatz which only requires a sufficiently high power of $p$ to lie in $\mathcal{I}$. Now, if $p(v) = 0$, then also $p^{*k}(v) := (p(v))^k = 0$, but (for $k > 1$) $v$ is more of a zero of $p^{*k}$ in the sense that $p^{*k}(z)$ goes to zero faster than $p(z)$ as $z \to v$. Various derivatives of $p^{*k}$ are zero at $v$ as well. So, as the Nullstellensatz hints at, in order for $p$ to belong to $\mathcal{I}$, it must vanish at each $v \in \mathcal{V}(\mathcal{I})$ to the right 'order' or multiplicity.

Even this notion of 'order' or multiplicity is subtle. It isn't just that

$$p(z) = O(|z - v|^k)$$

for some $k$. The full story is the following.

**"Lefranc's Nullstellensatz" [Le].** *For an arbitrary polynomial ideal $\mathcal{I}$ in $\Pi = \Pi(\mathbb{C}^d)$,*

$$\mathcal{I} = \bigcap_v (\mathcal{I}\perp^v)\perp_v, \tag{5.2}$$

*where, for any $S \subset \Pi$,*

$$S\perp^v := \{q \in \Pi : q(D)s(v) = 0, s \in S\}$$

*and*

$$S\perp_v := \{p \in \Pi : s(D)p(v) = 0, s \in S\}.$$

**Corollary.** *For an ideal projector $P$ with $\mathcal{I} = \ker P$ of finite codimension,*

$$\operatorname{ran} P' = \sum_v \delta_v Q_v(D),$$

*with*

$$Q_v := \mathcal{I}\perp^v = \{q \in \Pi : q(D)f(v) = 0, f \in \mathcal{I}\}.$$

Actually, the corollary can already be found in basic algebra books, e.g., [G70: p.168ff], but see already [G49] and the very nice overview article [G50]. Gröbner attributes the idea to Macaulay, e.g., [Ma: p.64ff], though it is described there in a different language (i.e., in terms of *inverse systems*) and there credit for first defining multiplicity correctly is given to Lasker [La] (who, however, defines it only as a number, namely the length (i.e., the codimension) of the associated primary ideal).

The space $Q_v = \mathcal{I}\perp^v$ is called the multiplicity space of $\mathcal{I}$ at $v$ (or, less descriptively, the Max Noether space of $\mathcal{I}$ at $v$; see [MT]). $Q_v$ is a linear subspace of $\Pi$, of the same dimension as the linear subspace

$$\delta_v Q_v(D) := \{f \mapsto q(D)f(v) : q \in Q_v\}$$

of $\Pi'$ that it supplies, and, obviously,

$$\delta_v Q_v(D) \subset \mathcal{I}^\perp = (\ker P)^\perp = \operatorname{ran} P'.$$

In other words, any ideal interpolant has interpolation conditions of the form

$$\delta_v q(D)$$

for certain sites $v$ and certain polynomials $q$. But much more is true. Since each of the spaces $\delta_v Q_v(D)$ lies in $\operatorname{ran} P'$, each must, in particular, be finite-dimensional. Also, since any finite sum of the form

$$\sum_v \delta_v Q_v(D)$$

is necessarily direct, there can be only finitely many nontrivial $Q_v$ here. But the most important fact is that each $Q_v$ is necessarily $D$-invariant. Is that obvious?

It can be verified in many ways. Perhaps the simplest is the following which uses the intriguing formula

$$q(D)f(0) = \sum_\alpha D^\alpha q(0) D^\alpha f(0)/\alpha! =: q * f, \tag{5.3}$$

which, quite rightly, has made its appearance in various papers concerning multivariate polynomials but under various names (see, e.g., [S05: above Theorem 6.1]). It is the unique bilinear form on $\Pi \times \Pi$ for which

$$(rq) * f = q * (r(D)f), \quad r, q, f \in \Pi. \tag{5.4}$$

(5.4) follows directly from (5.3) while, for the verification of (5.3), note that it is linear in $q$ and $f$, hence can be verified by checking it for

$$q = [\![]\!]^\beta : x \mapsto x^\beta/\beta!,$$

the conveniently normalized power function, and $f = [\![]\!]^\gamma$. For these, $D^\alpha q(0) = [\![0]\!]^{\beta-\alpha} = \delta_{\beta,\alpha}$, hence

$$\sum_\alpha D^\alpha q(0) D^\alpha f(0)/\alpha! = \delta_{\alpha,\beta}\delta_{\alpha,\gamma}/\alpha! = \delta_{\beta,\gamma}/\beta!,$$

while

$$\delta_0([\![D]\!]^\beta [\![]\!]^\gamma) = \delta_0 [\![]\!]^{\gamma-\beta}/\beta! = \delta_{\gamma,\beta}/\beta!\ .$$

Note the symmetry, i.e.,

$$q * f = f * q,$$

hence, by symmetry, also

$$(r(D)q) * f = q * (rf).$$

Therefore, with

$$E^v : f \mapsto f(\cdot + v)$$

the translation by $v$, we have, for $q \in Q_v$, $f \in \mathcal{I}$, and $r \in \Pi$,

$$\begin{aligned}(r(D)q)(D)f(v) &= r(D)q * E^v f \\ &= q * (rE^v f) \\ &= q * E^v((E^{-v}r)f) \\ &= q(D)((E^{-v}r)f)(v) = 0,\end{aligned}$$

since $E^{-v}r \in \Pi$ and therefore $(E^{-v}r)f \in \mathcal{I}$.

With each $Q_v$ now known to be $D$-invariant, we know that it contains all constant polynomials if it is nontrivial. Hence, each nontrivial $Q_v$ supplies, in particular, the interpolation condition $\delta_v$. Correspondingly,

$$\mathcal{V}(\mathcal{I}) = \{v : Q_v \neq \{0\}\}$$

is the variety of the ideal $\mathcal{I}$, i.e., the set of zeros common to all polynomials in $\mathcal{I}$. But, in general, we have not just the matching of function values, but also the matching of some derivative information, with the important restriction that, if $\delta_v q(D)$ is being matched, then so is $\delta_v(D^\alpha q)(D)$ for all $\alpha$.

In the univariate case, there is only one $D$-invariant polynomial subspace of dimension $k$, namely $\Pi_{<k}$, the polynomials of order $k$. But this says that, in the univariate case, ideal interpolation is Hermite interpolation. For that reason*, we also use the term Hermite interpolation for the projector $P$ in the multivariate case when the interpolation conditions are of the form

$$\operatorname{ran} P' = \sum_{z \in Z} \delta_z Q_z(D) \tag{5.5}$$

for some finite set $Z$, with each $Q_z$ a $D$-invariant finite-dimensional polynomial space.

Is any such Hermite interpolation ideal?

*If $Q$ is any $D$-invariant linear subspace of $\Pi$, then, for arbitrary $z$, $Q\perp_z$ is an ideal*: For, if $q \in Q$ and $f \in Q\perp_z$, then, for arbitrary $r \in \Pi$,

$$(rf) * E^z q = f * r(D)(E^z q) = f * E^z(r(D)q) = 0,$$

* dec05: not good enough any more

since then $r(D)q \in Q$, hence also $rf \in Q\perp_z$. But this says that

$$(\sum_z \delta_z Q_z(D))_\perp = \bigcap_z (\delta_z Q_z(D))_\perp = \cap_z Q_z \perp_z$$

is the intersection of ideals, hence an ideal. In other words, *Hermite interpolation is characterized by the fact that it is ideal.*

Apparently, the first to use 'Hermite interpolation' in this sense in the multivariate context is H. M. Möller; see [M76], [M77] which predate [Bi] and, in contrast to [Bi], describe $\operatorname{ran} P'$.

In [BR90] and, regrettably, not yet aware of Möller's work, we defined 'Birkhoff-Hermite interpolation' to mean a linear projector $P$ on $\Pi$ satisfying (5.5) with each $Q_z$ dilation-invariant (i.e., $q \in Q_z$ and $h > 0$ implies $q(\cdot h) \in Q_z$ or, what is the same, $Q_z$ is spanned by homogeneous polynomials), and restricted the term 'Hermite interpolation' to such $P$ for which each $Q_z$ is also $D$-invariant. Note that Hakopian and his colleagues reserve the term 'Hermite interpolation' for $P$ for which $\operatorname{ran} P = \Pi_k$ for some $k$ while $\operatorname{ran} P'$ is given by (5.5), with $Q_z = \Pi_{k_z}$, all $z$; see, e.g., [H], [BHS]. Earlier, [Lo92] called such interpolation 'Hermite interpolation of type total degree' but also considered 'Hermite interpolation of type tensor product', in which each $Q_z$ is of the form $\Pi_{\le\alpha}$ for some $z$-dependent $\alpha$; see [LL] for an early paper and [Lo00] for a recent survey. Further, [SX95b] use 'Hermite interpolation' to mean $P$ with $\operatorname{ran} P'$ of the form (5.5) with each $Q_z$ spanned by polynomials of the form

$$\langle \cdot, Y \rangle := \prod_{y \in Y} \langle \cdot, y \rangle,$$

and containing, with each such $\langle \cdot, (y_1, \ldots, y_r) \rangle$, also $\langle \cdot, (y_1, \ldots, y_{r-1}) \rangle$. Here,

$$\langle x, y \rangle := \sum_i x(i) y(i).$$

Such a $Q_z$ may fail to be $D$-invariant unless it contains, with each $\langle \cdot, Y \rangle$, also $\langle \cdot, Y \backslash y \rangle$ for every $y \in Y$. [SX95b] call their 'Hermite interpolation' regular in case all the $Q_z$ are $D$-invariant (hence the interpolation is ideal). This raises the question whether any $D$-invariant space has such a spanning set, for only then would such 'regular Hermite interpolation' be exactly the same as what we have called here 'Hermite interpolation'.

The above characterization of ideal interpolation implies that Kergin interpolation (see, e.g., [K] and [Mi]) is ideal only when it is a Taylor projector, i.e., when it involves only one site. In the same vein, the various mean-value interpolation schemes developed by Hakopian (see, e.g., [BHS]) fail to be ideal except when the underlying simplex degenerates to a point.

## §6. When is Hermite Interpolation Lagrange Interpolation?

It is evident that Hermite interpolation is Lagrange interpolation exactly when there is equality in (5.1), i.e., when

$$\#\mathcal{V}(\ker P) = \dim \operatorname{ran} P,$$

or, equivalently, when the ideal $\ker P$ is radical. There is a pretty characterization of this in terms of the linear maps $M_j$, $j = 1{:}d$, introduced in (1.5). This characterization is in terms of the eigenstructure of the $M_j$. Since the $M_j$ commute, they have a joint set of eigenvectors. The following lemma is standard (see, e.g., [CLO98: p.54]) but is proved here for the reader's convenience.

**Lemma 6.1.** *For any $p \in \Pi$, the spectrum of $p(M)$ is*

$$\operatorname{spect}(p(M)) = p(\mathcal{V}).$$

**Proof:** We continue to take for granted that $[\delta_v : v \in \mathcal{V}]$ is 1-1, i.e., that

$$\Pi \to \mathbb{C}^{\mathcal{V}} : p \mapsto p|_{\mathcal{V}} \text{ is onto.} \tag{6.2}$$

Take $p \in \Pi$, $\mu \in \mathbb{C}$, and consider

$$p(M) - \mu \mathrm{id} := q(M).$$

If $\mu \notin p(\mathcal{V})$, then $q$ does not vanish on $\mathcal{V}$, therefore, by (6.2), for some polynomial $r$, $()_0 - qr$ vanishes on $\mathcal{V}$, hence, by Hilbert's Nullstellensatz, some power of it, say the $k$th, lies in $\ker P = \ker M$. This says that

$$0 = (()_0 - qr)^k(M) = (M^0 - q(M)r(M))^k = \mathrm{id} - q(M)A$$

for some $A \in L(\operatorname{ran} P)$, showing $q(M) = p(M) - \mu \mathrm{id}$ to be invertible (since $\operatorname{ran} P$ is finite-dimensional).

If, on the other hand, $\mu = p(v)$ for some $v \in \mathcal{V}$, then, for all $q \in \operatorname{ran} P$,

$$\delta_v M_p q = \delta_v P(pq) = \delta_v(pq) = \mu \delta_v q,$$

showing $\delta_v$ to be a left eigenvector for $M_p$ for the eigenvalue $\mu = p(v)$ (this is Stetter's insight; see [AS]). □

**Proposition 6.3 ([MSt]).** *The ideal projector $P$ with $F := \operatorname{ran} P$ is Lagrange interpolation (i.e., $\#\mathcal{V} = \dim F$) if and only if the $M_j$ are diagonalizable.*

**Proof:** If $\#\mathcal{V} = \dim F$, then, since $\dim \operatorname{ran} P' = \dim F$, $[\delta_v : v \in \mathcal{V}]$ is an eigenbasis for $M_p'$ (for any $p$). Correspondingly, its dual basis in $F$, i.e., the basis $[\ell_v : v \in \mathcal{V}]$ with

$$\ell_v(w) = \delta_{vw}, \quad v, w \in \mathcal{V},$$

is an eigenbasis for $M_p$ (again for any $p$); it is evidently the Lagrange basis for interpolation from $F$ at $\mathcal{V}$.

Conversely, let $V : \mathbb{C}^n \to \operatorname{ran} P$ be an eigenbasis for the $M_j$. Then, the map

$$\Pi \to \mathbb{C}^{n\times n} : p \mapsto V^{-1}p(M)V$$

is linear and, by (2.1), has $\ker P$ as its kernel. In other words, with $\lambda_{ij}$ the map that carries $p \in \Pi$ to the $(i,j)$-entry of the matrix $V^{-1}p(M)V$, we have

$$\ker P = \cap_{i,j} \ker \lambda_{ij},$$

hence $(\lambda_{ij} : i,j = 1{:}n)$ spans $\operatorname{ran} P'$. But, since $V$ is an eigenbasis for the $M_j$, all the matrices $V^{-1}p(M)V$ are diagonal, hence only the $\lambda_{ii}$ are nontrivial and, since there are only $n := \dim \operatorname{ran} P'$ of them, they must form a basis for $\operatorname{ran} P'$. In particular, there must exist $p \in \Pi$ for which $\#\{\lambda_{ii}p : i = 1{:}n\} = n$. Since $\{\lambda_{ii}p : i = 1{:}n\} = \operatorname{spect}(p(M)) = \{p(v) : v \in \mathcal{V}\}$, this implies that $\#\mathcal{V} = n$. □

As the simplest example, consider $P : p \mapsto p(0)()^0 + Dp(0)()^1$. We compute the matrix representation for $M_1$ with respect to the standard basis, $[()^0, ()^1]$, for $\operatorname{ran} P = \Pi_1 \subset (\mathbb{F} \to \mathbb{F})$:

$$M_1()^0 = P()^1 = ()^1; \quad M_1()^1 = P()^2 = 0,$$

hence

$$\widehat{M_1} = [\varepsilon_2, 0] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

the simplest example of a defective matrix.

It seems that Auzinger and Stetter [AS] were the first to propose to use the eigenstructure of the $M_j$ for the calculation of $\mathcal{V}$. This requires, in principle, nothing more than the calculation of a matrix $\widehat{M_j}$ similar to $M_j$, and this can be obtained in many ways, e.g., by computing the representation of $M_j$ with respect to some basis $W$ of $\operatorname{ran} P$. From this, one can, in principle, compute a basis $U$ consisting of (generalized) eigenvectors for any particular $M_j$, and, with that in hand, can now compute $\widehat{M_j} := U^{-1}M_jU$ for every $j$, hence know, in particular, not only $v(j)$ for all $j$, but even the points $v$ themselves, since one then knows the $\lambda_{ii}$ at least on $\Pi_1$.

However, Auzinger and Stetter go for the eigenvectors of the *transpose* of $\widehat{M_j}$, as these are necessarily of the form $\delta_v U = (u(v) : u \in U)$. Actually, [AS] focus on the *left* eigenvector $a_v$ of the matrix $\widehat{M_p}$ belonging to the eigenvalue $p(v)$ since it is necessarily (a scalar multiple of) $\delta_v W$, hence has $w(v)$, $w \in W$, as its entries. If now $W$ can be chosen to contain $()_j$, $j = 1{:}d$, then $a_v$ contains the very coordinates of $v$. If $W$ cannot be so chosen, still there are then techniques for teasing out $v$ from the vector $a_v$; see [St], [MSt].

## §7. Is Hermite Interpolation the Limit of Lagrange Interpolation?

While one is, of course, free to give names to hitherto unnamed concepts and constructs, use of an established name in a new or more general context needs justification. Since it is an integral and often used aspect of univariate Hermite interpolation that it is the (pointwise) limit of Lagrange interpolation, it is fair to ask whether multivariate ideal interpolation is also the limit of Lagrange interpolation. This question was already raised in [BR90], within the restricted meaning of 'Hermite interpolation' used there, but has yet to be answered even in that restricted context.

To be sure, pointwise convergence of maps on a linear space depends on the notion of limit in that space to be employed. On $\Pi$, we use uniform convergence on compact sets or, what is the same, coefficient-wise convergence, i.e.,

$$\lim_{n\to\infty} p_n = p \quad \iff \quad \forall \alpha \in \mathbb{Z}_+^d \quad \lim_{n\to\infty} \widehat{p}_n(\alpha) = \widehat{p}(\alpha).$$

**Proposition 7.1.** *The pointwise limit of ideal projectors is ideal.*

**Proof:** Since the property of being ideal can be characterized pointwise (see Lemma 1.1), it is preserved under pointwise convergence. □

Since a linear projector is determined by its range and the range of its dual, the pointwise convergence of a sequence $(P_n : n \in \mathbb{N})$ of (finite-rank) linear projectors is equivalent to the convergence of their ranges and the ranges of their duals. Thus, we are interested in what limits, if any, can linear spaces, spanned by finitely many point evaluations, have as the evaluation sites all coalesce at one site, $v$. The above proposition implies that, if there is a limiting space, it is necessarily of the form $\delta_v Q_v(D)$ for some $D$-invariant space $Q_v$. But the space $Q_v$ will crucially depend on just how the evaluation sites coalesce. Here is an example, from [BR90].

**Proposition 7.2.** *Let $v$ and $\mathrm{T}$ be a point, respectively a finite subset, in $\mathbb{Z}^d$. Then*

$$\lim_{h\to 0} \operatorname{ran}[\delta_{v+h\tau} : \tau \in \mathrm{T}] \;=\; \delta_v \Pi_{\mathrm{T}}(D),$$

*with*

$$\Pi_{\mathrm{T}} \;:=\; \bigcap_{p|_{\mathrm{T}}=0} \ker p_{\uparrow}(D).$$

**Proof:** Assume without loss that $v = 0$. Then the general element of $\operatorname{ran}[\delta_{v+h\tau} : \tau \in \mathrm{T}]$ is of the form

$$\lambda_h : p \mapsto \lambda p(h\cdot), \quad \text{with} \;\; \lambda := \sum_{\tau\in\mathrm{T}} c(\tau)\delta_\tau.$$

We compute

$$\begin{aligned}\lambda_h p = \lambda p(h\cdot) &= \sum_{\tau\in\mathrm{T}} c(\tau) \sum_\alpha (h\tau)^\alpha \widehat{p}(\alpha)\\ &= \sum_j h^j \sum_{|\alpha|=j} \underbrace{\sum_{\tau\in\mathrm{T}} c(\tau)\tau^\alpha}_{\lambda()^\alpha}\ \widehat{p}(\alpha)\\ &= \sum_{j\ge\operatorname{ord}\lambda} h^j \sum_{|\alpha|=j} \lambda()^\alpha \widehat{p}(\alpha)\end{aligned}$$

with

$$\operatorname{ord}\lambda := \min\{|\alpha| : \lambda()^\alpha \ne 0\}.$$

Therefore

$$\begin{aligned}\lim_{h\to 0} \lambda_h p/h^{\operatorname{ord}\lambda} = \sum_{|\alpha|=\operatorname{ord}\lambda} \lambda()^\alpha \widehat{p}(\alpha) &= \sum_{|\alpha|=\operatorname{ord}\lambda} \lambda()^\alpha \frac{1}{\alpha!} D^\alpha p(0)\\ &= q(D)p(0),\end{aligned}$$

with

$$q := \sum_{|\alpha|=\operatorname{ord}\lambda} \sum_{\tau\in\mathrm{T}} c(\tau)\frac{\tau^\alpha}{\alpha!}\ ()^\alpha \ = ???$$

a certain polynomial. Note that, in the univariate case, this sum would only have one term in it and, correspondingly, the limit is just a scalar multiple of the $(\operatorname{ord}\lambda)$-th derivative at the origin, just as expected. In the multivariate case, things are more complicated. Yet, as we look further into this polynomial $q$, we'll also discover real beauty.

What does the term $\tau^\alpha/\alpha!$ remind you of? The *exponential function*! In fact, you recall

$$\mathrm{e}_\tau : x \mapsto \mathrm{e}^{\langle\tau,x\rangle} = \sum_j \langle\tau,x\rangle^j/j! = \sum_\alpha \frac{\tau^\alpha}{\alpha!}\ x^\alpha,$$

the exponential with frequency $\tau$. So, with the definitions

$$f := \sum_{\tau\in\mathrm{T}} c(\tau)\mathrm{e}_\tau = \sum_j \underbrace{\sum_{|\alpha|=j} \sum_{\tau\in\mathrm{T}} c(\tau)\frac{\tau^\alpha}{\alpha!}\ ()^\alpha}_{=:\ f^{[j]}},$$

we see again $q$:

$$q = f^{[\operatorname{ord}\lambda]}.$$

In other words: if we organize $f = \sum_\tau c(\tau)\mathrm{e}_\tau$ into its *homogeneous terms*,

$$f = f^{[0]} + f^{[1]} + \cdots,$$

then we find that $f^{[\operatorname{ord}\lambda]}$ is the first such term that is non-zero. For that reason, we call it the least or initial term of $f$, and denote it by

$$f_\downarrow.$$

$f_\downarrow$ is the unique homogeneous polynomial for which

$$\operatorname{ord}(f - f_\downarrow) > \operatorname{ord} f.$$

It follows that $\lim_{h\to 0} \operatorname{ran}[\delta_{v+h\tau} : \tau \in \mathrm{T}]$ contains $\delta_0(\mathrm{Exp}_\mathrm{T})_\downarrow(D)$, with

$$\mathrm{Exp}_\mathrm{T} := \operatorname{ran}[\mathrm{e}_\tau : \tau \in \mathrm{T}]$$

and

$$F_\downarrow := \operatorname{span}(f_\downarrow : f \in F)$$

for any linear subspace $F$ of

$$\Pi' \;\sim\; \mathcal{P} \;:=\; \mathbb{F}[[x]], \tag{7.3}$$

the space of formal power series in $d$ variables $x(1), \ldots, x(d)$ with coefficients in $\mathbb{F}$.

On the other hand, each $\operatorname{ran}[\delta_{v+h\tau} : \tau \in \mathrm{T}]$ has dimension equal to $\#\mathrm{T}$, hence its limit as $h \to 0$ can have dimension at most $\#\mathrm{T}$, while (see [BR90]) $\dim F_\downarrow = \dim F$ and $\dim \mathrm{Exp}_\mathrm{T} = \#\mathrm{T}$. Therefore

$$\lim_{h\to 0} \operatorname{ran}[\delta_{v+h\tau} : \tau \in \mathrm{T}] = \delta_0(\mathrm{Exp}_\mathrm{T})_\downarrow(D).$$

Finally (see [BR92a] and [BR92b]; for a direct proof, see [B92]),

$$(\mathrm{Exp}_\mathrm{T})_\downarrow = \bigcap_{p|_\mathrm{T}=0} \ker p_\uparrow(D). \quad \square$$

The equivalence of $\Pi'$ with $\mathcal{P}$ claimed in (7.3) can be established in several ways. For our purposes, it is convenient to do it via the natural extension of the bilinear form (5.3) to

$$\mathcal{P} \times \Pi \to \mathbb{F} : (f, p) \mapsto f * p = \sum_\alpha \widehat{f}(\alpha)\alpha!\widehat{p}(\alpha).$$

Note that, for any $v \in \mathbb{F}^d$ and any $p \in \Pi$,

$$e_v * p = \sum_\alpha v^\alpha \widehat{p}(\alpha) = p(v).$$

In other words, the exponential function with frequency $v$ represents evaluation at $v$ with respect to this pairing. In particular, given that we were interested in finding $\lim_{h\to 0} \sum_\tau c(\tau)\delta_{h\tau}$, the appearance of the exponential function in the above proof is not accidental.

Note further that $\Pi_T$ is not only $D$-invariant (as the intersection of kernels of constant-coefficient differential operators) but also dilation-invariant (as the span of homogeneous polynomials). In contrast, in general, the multiplicity spaces $Q_v$ for an ideal projector need only be $D$-invariant. Here is a further example, from [BR90], to show how such a $\delta_v Q_v(D)$ may, nevertheless, be the limit of spaces spanned by point evaluations.

Let $T_h := \{\xi_- := (-h, h^2),\ 0,\ \xi_+ := (h, h^2)\} \subset \mathbb{F}^2$ and set $M_h := \mathrm{ran}[\delta_\tau : \tau \in T_h]$. Then, with $\xi_0 := (0, h^2)$, $M_h$ contains

$$(\delta_{\xi_+} + \delta\xi_- - 2\delta_0)/h^2 = (\delta_{\xi_+} - 2\delta_{\xi_0} + \delta_{\xi_-})/h^2 \ +\ 2(\delta_{\xi_0} - \delta_0)/h^2,$$

and this evidently converges to $\delta_0(D_1^2 + 2D_2)$ as $h \to 0$, while certainly $(\delta_{\xi_+} - \delta_{\xi_-})/h$ is in $M_h$ and converges to $\delta_0 D_1$, and $\delta_0$ is in $M_h$ for all $h$. This shows that the 3-dimensional space $\delta_0 Q_0(D)$ with $Q_0 := \mathrm{ran}[()^0, ()^{1,0}, ()^{2,0} + 2()^{0,1}]$ is in $\lim_{h\to 0} M_h$, hence must coincide with it since each $M_h$ is only 3-dimensional. Note that $Q_0$ is $D$-invariant but not dilation-invariant.

**Conjecture.** *A linear projector on $\Pi \subset (\mathbb{C}^d \to \mathbb{C})$ is ideal if and only if it is the (pointwise) limit of Lagrange interpolation.*

Some people have told me that this conjecture is obviously true, because of known results concerning the resolution of singularities. On the other hand, Geir Ellingsrud has pointed out to me that this conjecture must fail for $d > 2$, because of results by Iarrobino (see [I]) concerning the dimension of the manifold of ideals of codimension $k$ with $k$ points in their variety as compared with the dimension of the manifold of ideals of codimension $k$ with variety $\{0\}$. But, lacking as yet a sufficiently good background in Algebraic Geometry, I have not yet understood his reasoning. In any case, Ellingsrud's remark does not contradict the following, very recent, response, by Boris Shekhtman, to the above conjecture.

**Proposition 7.4 ([Sh]).** *Any ideal projector on $\Pi \subset (\mathbb{C}^2 \to \mathbb{C})$ with range the polynomials of degree $\le k$ (for some $k$) is the pointwise limit of Lagrange interpolation projectors.*

**Proof outline:** Let $F = \Pi_k$ be the range of the ideal projector $P$, and recall from Proposition 6.3 that $P$ is Lagrange interpolation iff the linear maps $M_j : F \to F : f \mapsto P(()_j f)$ are diagonalizable. Since $F$ is finite-dimensional, the diagonalizable linear maps on $F$ are dense in $L(F)$. Hence we are looking for an indication that the set of all ideal projectors with range $F$ is open in some sense.

From Proposition 1.7, we know that $P$ is characterized by its action on $\Pi_1(F) = \Pi_{k+1}$, hence by the polynomials

$$h_\alpha := P()^\alpha \subset \operatorname{ran} P = \Pi_k, \quad |\alpha| = k+1,$$

since $P()^\alpha = ()^\alpha$ for $|\alpha| \le k$. On the other hand, while any choice of the $h_\alpha$ gives rise to a linear projector $N$ on $\Pi_1(F)$ with range $F = \Pi_k$, not all of them are the restriction to $\Pi_1(F)$ of an ideal projector with range $F$. Since $F$ evidently satisfies Mourrain's condition (3.1), we know from Theorem 3.2 that $N$ is the restriction of an ideal projector with range $F$ if and only if $M_i M_j = M_j M_i$ on all relevant $()^\alpha$, i.e.,

$$N(()_i N(()_j ()^\alpha)) = N(()_j N(()_i ()^\alpha)), \quad |\alpha| \le k, \ 1 \le i < j \le d.$$

This equality holds trivially for $|\alpha| < k$ since then $N(()_i ()^\alpha) = ()_i ()^\alpha$. Further, for $|\alpha| = k$, $N(()_i ()^\alpha) = h_{\varepsilon_i + \alpha}$, hence, altogether, the condition is that

$$()_i h_{\varepsilon_j+\alpha} - ()_j h_{\varepsilon_i+\alpha} \in \ker N, \quad |\alpha| = k, \ i < j.$$

But $(()^\beta - h_\beta : |\beta| = k+1)$ is evidently linearly independent (since $h_\beta \in \Pi_k$) and has $\dim \ker N$ terms and is in $\ker N$, hence is a basis for $\ker N$. Therefore, the choice $(h_\beta : |\beta| = k+1)$ specifies an ideal projector with range $\Pi_k$ if and only if there are matrices $C_{ij}$ (necessarily unique) so that

$$()_i h_{\varepsilon_j+\alpha} - ()_j h_{\varepsilon_i+\alpha} = \sum_{|\beta|=k+1} C_{ij}(\alpha,\beta)(()^\beta - h_\beta), \quad |\alpha| = k, \ i < j. \tag{7.5}$$

Now, in the bivariate case actually under discussion, there is just one choice for $(i, j)$, namely $(1, 2)$, hence $(h_\beta : |\beta| = k+1)$ in $\Pi_k$ gives rise to an ideal projector with range $\Pi_k$ if and only if there is some matrix $C$ so that

$$()_1 h_{\varepsilon_2+\alpha} - ()_2 h_{\varepsilon_1+\alpha} = \sum_{|\beta|=k+1} C(\alpha,\beta)(()^\beta - h_\beta), \quad |\alpha| = k. \tag{7.6}$$

It is this equation, Shekhtman derives and looks at. He treats it as an equation for the vector

$$h := (h_\beta : |\beta| = k+1),$$

hence writes it in the form

$$Ah - C(b - h) = 0, \tag{7.7}$$

with

$$b := (()^\beta : |\beta| = k+1)$$

and

$$Ah := (()_1 h_{\varepsilon_2+\alpha} - ()_2 h_{\varepsilon_1+\alpha} : |\alpha| = k),$$

hence

$$Ab = 0,$$

therefore (7.6) is equivalent to

$$(A+C)(h-b) = 0. \tag{7.8}$$

Now, given that $A + C$ has one more column than it has rows, it follows, by a standard formula, that

$$h := (()^\beta - (-1)^\beta \det(A+C)(:, \backslash\beta) : |\beta| = k+1) \tag{7.9}$$

solves (7.8), hence (7.6), with $(-1)^\beta$ equal to 1 or $-1$ depending on the parity of column $\beta$ in the columns of $A$ or $C$. Assume, in particular, the columns so ordered that, for $\beta =: (j, k+1-j)$, $j$ increases as we traverse the columns from left to right. Then it is immediate that $\det A(:, \backslash\beta) = (-1)^\beta ()^\beta$, hence this $h$ is in $\Pi_k$, as required. This shows that each choice of the scalar matrix $C$ gives rise to an ideal projector. It also shows that each $\det A(:, \backslash\beta)$ is nonzero almost everywhere, hence $A+C$ is onto almost everywhere and, therefore, $\ker(A+C)$ is 1-dimensional almost everywhere. In other words, for given $C$, $h$ uniquely solves (7.6). In particular, (7.9) provides a parametrization of the set of all ideal projectors with range $\Pi_k$. (Appreciate the fact that the number of entries of $C$ equals $2 \dim \Pi_k$, i.e., the degrees of freedom that uniquely pin down any particular bivariate Lagrange projector with range $\dim \Pi_k$.)

Now notice that (7.9) describes the solution $h$ as a polynomial function in the entries of the (scalar) matrix $C$. Hence, with $\Lambda$ a basis for $\operatorname{ran} P'$ and $n := \dim F = \dim \operatorname{ran} \Lambda$, the determinant of the Gram matrix

$$\Lambda^{\mathrm{t}}[()_1^j : j < n]$$

is also a polynomial in the entries of $C$, and is nonzero for some choice of $C$. Hence, every neighborhood of our ideal projector $P$ contains an ideal projector $R$ with range $\Pi_k$ and such that, for any basis M for $\operatorname{ran} R'$, the Gram matrix $\mathrm{M}^{\mathrm{t}}[()_1^j : j < n]$ is invertible, hence there is a linear projector $S$ with $\operatorname{ran} S = \operatorname{ran}[()_1^j : j < n]$ and $\operatorname{ran} S' = \operatorname{ran} R'$, hence an ideal projector. By perturbing, if necessary, the zeros of the polynomial $()_1^n - S()_1^n$ (considered as a univariate polynomial), we obtain (see the example following Proposition 1.7) an interpolating ideal projector $T$ as close to $S$ as we would like, and, with that, the linear projector $U$ with range $\Pi_k$ and $\operatorname{ran} U' = \operatorname{ran} S'$ is well-defined and an interpolating projector as close to $P$ as we would like. □

Actually, Shekhtman's argument proves the conjecture for an arbitrary bivariate ideal projector $P$, i.e., we have the following.

**Corollary.** *Any ideal projector on* $\Pi \subset (\mathbb{C}^2 \to \mathbb{C})$ *is the pointwise limit of Lagrange interpolation projectors.*

**Proof outline*:** We know from Proposition 4.3 that the ideal $\ker P$ is complemented by a polynomial space $\Pi_\Gamma$ with $\Gamma$ a lower set (i.e., the complement of an upper set), hence there is a linear projector $R$ with $\operatorname{ran} R = \Pi_\Gamma$ and $\ker R = \ker P$, hence $R$ is ideal. Further, for any projector $S$ close enough to $R$, $\operatorname{ran} S'$ must be close enough to $\operatorname{ran} R' = \operatorname{ran} P'$ so that there is a projector $T$ with $\operatorname{ran} T = \operatorname{ran} P$ and $\operatorname{ran} T' = \operatorname{ran} S'$ and $T$ is close to $P$. It is therefore sufficient to consider $P$ with $F := \operatorname{ran} P = \Pi_\Gamma$ and $\Gamma$ a lower set. For such $P$, $\operatorname{ran} P$ satisfies Mourrain's condition (3.1), hence we may proceed as in Shekhtman's proof, except for the following wrinkle. It now may happen for some

$$\alpha \in \partial\Gamma := \{\alpha \in \Gamma : \exists j \;\; \varepsilon_j + \alpha \notin \Gamma\}$$

that, e.g., $\varepsilon_1 + \alpha \notin \Gamma$ while $\varepsilon_2 + \alpha \in \Gamma$ (something not possible when $F = \Pi_k = \Pi_{\{\alpha:|\alpha|\le k\}}$). In this case, $()_1 M_2 ()^\alpha = ()^{\varepsilon_1+\varepsilon_2+\alpha}$, hence the condition $M_1 M_2 = M_2 M_1$ on $()^\alpha$ now reads that $()^{\varepsilon_1+\varepsilon_2+\alpha} - ()_2 h_{\varepsilon_1+\alpha} \in \ker N$. But this is equivalent to the condition that

$$h_{\varepsilon_1+\varepsilon_2+\alpha} - ()_2 h_{\varepsilon_1+\alpha} \in \ker N,$$

since $()^{\varepsilon_1+\varepsilon_2+\alpha} - h_{\varepsilon_1+\varepsilon_2+\alpha} \in \ker N$. This means that the condition for the projector $N$ on $\Pi_1(F)$ with range $F$, specified by the sequence

$$h := (h_\beta =: N()^\beta : \beta \in \partial(\backslash\Gamma))$$

in $F$, with

$$\partial(\backslash\Gamma) := (\{\varepsilon_1, \varepsilon_2\} + \partial\Gamma)\backslash\Gamma,$$

to be the restriction to $\Pi_1(F)$ of some ideal projector with range $F = \Pi_\Gamma$ is still (7.7), except that now

$$A(\alpha,\beta) = \left\{ \begin{array}{ll} ()_1, & \varepsilon_2+\alpha=\beta; \\ 1, & \varepsilon_2+\alpha\in\Gamma \text{ and } \varepsilon_1+\varepsilon_2+\alpha=\beta; \\ -()_2, & \varepsilon_1+\alpha=\beta; \\ -1, & \varepsilon_1+\alpha\in\Gamma \text{ and } \varepsilon_2+\varepsilon_1+\alpha=\beta; \\ 0, & \text{otherwise} \end{array} \right\}, \quad \left\{ \begin{array}{l} \alpha\in\partial\Gamma, \\ \beta\in\partial(\backslash\Gamma) \end{array} \right\},$$

with the fact that $\Gamma$ is a lower set guaranteeing that, for each such $(\alpha, \beta)$, exactly one of these cases obtains. In particular, $A$ still has one more column than it has rows, and all other assertions made in the proof about

* dec05: this outline seems not realizable; rather, Shekhtman had to dig into Algebraic Geometry to prove the corollary.

$A$ remain valid, including that $Ab = 0$ and $\det A(:, \backslash\beta) = (-1)^{\beta}()^{\beta}$ for all $\beta \in \partial(\backslash\Gamma)$, hence the rest of Shekhtman's proof goes through without change, – except that, as Shekhtman has pointed out to me, in this more general situation, his formula (7.9) will not give rise to a sequence $h$ in $F$ for every choice of the matrix $C$. Rather, $h$ will be in $F$ if and only if certain submatrices of $C$ are zero. □

Returning to our 0-dimensional polynomial ideal $\mathcal{I}$, it is customary to refer to the dimension of $Q_v = \mathcal{I}\perp^v$ as the multiplicity of $v$ as a point in $\mathcal{V}(\mathcal{I})$. But it is clear that, in the multivariate context, this provides too little information. It is the space $Q_v$ itself that carries the detailed information.

[G50] contains a whole section devoted to the pitfalls to be avoided by anyone wishing to explain the multiplicity of a zero of an ideal in terms of coalescing point evaluations. Specifically, it is pointed out there that it is not possible to *define* multiplicity by the number of point evaluations that might be coalescing there since that number will surely depend on the particular sequence chosen. In particular, there are cases of higher-dimensional ideals (hence their variety is not finite) that can be approximated in some nice geometric sense by 0-dimensional ideals, perhaps even with a bound on the cardinality of their varieties. A footnote refers to a private communication from Burau who states that, nevertheless, he had been able to arrive in this way at a satisfactory definition of multiplicity that, not surprisingly, was equivalent to the present, ideal-theoretic one.

## §8. The Choice of $\operatorname{ran} P$

A projector's property of being ideal is entirely determined by its kernel, the ideal $\mathcal{I}$. For a given nontrivial ideal $\mathcal{I}$ or, equivalently, a given 'ideal' space $\mathcal{I}^{\perp}$ of interpolation conditions, there are infinitely many ideal projectors, one for each choice of an algebraic complement of $\mathcal{I}$ as $\operatorname{ran} P$.

One popular choice for $\operatorname{ran} P$ is to ensure that $P$ be degree-reducing, meaning that

$$\deg Pp \leq \deg p, \quad p \in \Pi.$$

This is called of least degree in [BR90], and of minimal degree in [BR92a] and [BR92b], and [S97] is entirely devoted to this notion, with a highlight the proof that *every 0-dimensional ideal has an algebraic complement that is spanned by monomials and is D-invariant and whose corresponding projector is degree-reducing.*

As is pointed out in [B05a] (thus providing another proof for Proposition 4.3), such an algebraic complement can be obtained by Gauss elimination with partial pivoting, applied to the Gram matrix

$$\Lambda^{\mathrm{t}} V,$$

with $\Lambda$ a column map into $\Pi'$ for which $\ker \Lambda^t$ is the ideal and

$$V := [()^\alpha : \alpha \in \mathbb{Z}_+^d]$$

such that the order $<$ on $\mathbb{Z}_+^d$ corresponding to the order of the columns of $V$ respects addition, i.e., satisfies (4.1), and respects degree, i.e., $|\alpha| < |\beta| \Longrightarrow \alpha < \beta$. If $\beta_1 < \cdots < \beta_n$ is the sequence of indices of the bound columns of $\Lambda^t V$ as determined by Gauss elimination, then $\operatorname{ran}[()^{\beta_i} : i = 1{:}n]$ is that desired algebraic complement.

A quite different choice for $\operatorname{ran} P$ may result from the wish for a particularly 'nice' error formula. One reason for choosing ideal interpolation in the first place is the resulting possibility of writing the error in the form

$$f - Pf \ = \ \sum_{b \in B} b \, q_{b,f}$$

with $B$ a minimal basis for the ideal $\ker P$, and $q_{b,f}$ suitable polynomials depending on $f \in \Pi$. (This nice feature of ideal interpolation is also recognized implicitly in [SX95b] where it motivates the restriction to 'regular' Hermite interpolation.)

In the univariate case, the standard error formula takes the form

$$f - Pf \ = \ b \Delta(\tau_1, \ldots, \tau_n, \cdot) f,$$

with

$$b := (\cdot - \tau_1) \cdots (\cdot - \tau_n)$$

the monic polynomial that vanishes at the interpolation sites to the appropriate multiplicity, i.e., the monic polynomial that generates the ideal $\ker P$, and $\Delta(\tau_1, \ldots, \tau_n, x) f$ the divided difference of $f$ at the sites $\tau_1, \ldots, \tau_n, x$, hence a polynomial in $x$ that depends linearly on $D^n f$. More precisely,

$$\Delta(\tau_1, \ldots, \tau_n, x) f = \int K(\cdot | \tau_1, \ldots, \tau_n, x) D^n f$$

for a certain function $K$, namely a B-spline with knots $\tau_1, \ldots, \tau_n, x$. Since $D^n = b_\uparrow(D)$, one may therefore hope, in the multivariate case, for an error formula of the form

$$f(x) - Pf(x) \ = \ \sum_{b \in B} b(x) I_{x,b}(b_\uparrow(D) f) \tag{8.1}$$

with $B$ a minimal generating set for $\mathcal{I}$ and with each $I_{x,b}$ some linear integral operator. Since $\operatorname{ran} P$ comprises exactly those polynomials for which $f - Pf = 0$, this would imply

$$\bigcap_{p \in \ker P} \ker p_\uparrow(D) = \bigcap_{b \in B} \ker b_\uparrow(D) \ \subseteq \operatorname{ran} P,$$

the equality holding because $B$ is a basis for the ideal $\ker P$. But since $\operatorname{ran} P$ is complementary to the ideal $\ker P$, this would imply

$$\bigcap_{p \in \ker P} \ker p_{\uparrow}(D) = \operatorname{ran} P.$$

But this implies (see [BR92a]) that $P$ is necessarily the least projector for the given interpolation conditions $(\ker P)^{\perp}$, as introduced in [BR92a] for arbitrary (finite-dimensional) spaces of interpolation conditions. I resist the urge to call the linear projector with

$$\operatorname{ran} P_{\mathcal{I}} = \bigcap_{p \in \mathcal{I}} \ker p_{\uparrow}(D) \quad \text{and} \quad \ker P_{\mathcal{I}} = \mathcal{I}$$

a 'least ideal projector', and call it least Hermite interpolation instead.

As a simple example, consider interpolation at $\Sigma \times \mathrm{T}$, with $\Sigma$ and $\mathrm{T}$ finite subsets of $\mathbb{F}$. The ideal $\mathcal{I}$ of all bivariate polynomials vanishing on $\Sigma \times \mathrm{T}$ is generated by the two polynomials

$$b_{\sigma} : x \mapsto \prod_{\sigma \in \Sigma} (x(1) - \sigma), \quad b_{\tau} : x \mapsto \prod_{\tau \in \mathrm{T}} (x(2) - \tau).$$

Correspondingly, with

$$m := \deg b_{\sigma}, \quad n := \deg b_{\tau},$$

the least choice for the space from which to interpolate in this case is the standard one, i.e.,

$$\begin{aligned} \operatorname{ran} P_{\mathcal{I}} &= \ker (b_{\sigma})_{\uparrow}(D) \cap \ker (b_{\tau})_{\uparrow}(D) \\ &= \ker D_1^m \cap \ker D_2^n = \operatorname{ran}[()^{\alpha} : \alpha(1) < \deg b_{\sigma}, \alpha(2) < \deg b_{\tau}]. \end{aligned}$$

However, the standard formula for the error in such tensor-product interpolation to $f$ involves not only $D_1^m f$ and $D_2^n f$ but also the higher mixed derivative $D^{m,n} f$. Nevertheless, it is possible (see [B97]) to derive an error formula for this particular, and even for general multivariate, tensor product interpolation, of the form (8.1), with $B$ the 'natural' basis for $\mathcal{I}$.

But (8.1) fails the next test, Chung-Yao interpolation, for which the error formula, derived in [B97], is of the slightly more complicated form

$$f(x) - Pf(x) = \sum_{b \in B} b(x) I_{b,x}(\tilde{b}_{\uparrow}(D) f), \tag{8.2}$$

with $(\tilde{b} : b \in B)$ also a (minimal) basis for $\mathcal{I}$ and such that $\tilde{b}_{\uparrow}(D) c = \delta_{b,c}$ for $b, c \in B$.

One may therefore hope for an error formula of the form (8.2) for arbitrary least Hermite interpolation (a hope first expressed in [B97]). But, already for general Lagrange interpolation from $\Pi_k$, this is still only a hope, as the Sauer-Xu error formula for that case (see [SX95a]) does not readily convert into the form (8.2).

To be sure, while I have restricted attention to interpolation on $\Pi$, it is easy to extend ideal interpolation to more general functions, namely to all functions $f$ smooth enough 'at' each interpolation site $v$ so that $q(D)f$ is defined there for all $q \in Q_v$, $v \in \mathcal{V}$. Given the density of polynomials in various function spaces, also the error formulas (known or yet to be derived) extend similarly to (smooth enough) functions other than polynomials. On the other hand, the restriction here to interpolation on $\Pi$ makes possible a simple, purely algebraic, treatment of the essential aspects of the polynomial interpolation discussed.

**Acknowledgments.** This article has benefited from a critical reading by Tomas Sauer. Special thanks are due to Boris Shekhtman for making available his results as soon as he got them and helping me understand his arguments.

## References

[AS] Auzinger, W. and H. Stetter, An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations, in *Proc. Internat. Conf. Numer. Math., ISNM vol. 86,* xxx (ed), Birkhäuser, Basel, 1988, 12–30.

[Bi] Birkhoff, G., The algebra of multivariate interpolation, in *Constructive Approaches to Mathematical Models,* C. V. Coffman and G. J. Fix (eds), Academic Press, New York, 1979, 345–363.

[BHS] Bojanov, B. D., H. A. Hakopian, and A. A. Sahakian, *Spline Functions and Multivariate Interpolations,* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[B92] Boor, C. de, On the error in multivariate polynomial interpolation, Applied Numerical Mathematics **10** (1992), 297–305.

[B94] Boor, C. de, Gauss elimination by segments and multivariate polynomial interpolation, in *Approximation and Computation: A Festschrift in Honor of Walter Gautschi,* R. V. M. Zahar (ed), ISNM 119, Birkhäuser Verlag, Basel-Boston-Berlin, 1994, 1–22.

[B97] Boor, C. de, The error in polynomial tensor-product, and in Chung-Yao, interpolation, in *Surface Fitting and Multiresolution Methods,* A. LeMéhauté, C. Rabut, and L. L. Schumaker (eds), Vanderbilt University Press, Nashville TN, 1997, 35–50.

[B03] Boor, C. de, A Leibniz formula for multivariate divided differences, SIAM J. Numer. Anal. **41(3)** (2003), 856–868.

[B05a] Boor, C. de, Interpolation from spaces spanned by monomials, Advances in Comp. Math., to appear.

[B05b] Boor, C. de, Ideal interpolation: Mourrain's condition *vs* $D$-invariance, in *Proceedings of conference in honor of Z. Ciesielski*, A. Kamont (ed), Banach Center Publications, Poland, 2005, to appear.

[BR90] Boor, C. de and A. Ron, On multivariate polynomial interpolation, Constr. Approx. **6** (1990), 287–302.

[BR91] Boor, C. de and A. Ron, On polynomial ideals of finite codimension with applications to box spline theory, J. Math. Anal. Appl. **158** (1991), 168–193.

[BR92a] Boor, C. de and A. Ron, The least solution for the polynomial interpolation problem, Math. Z. **210** (1992), 347–378.

[BR92b] Boor, Carl de and Amos Ron, Computational aspects of polynomial interpolation in several variables, Math. Comp. **58(198)** (1992), 705–727.

[CLO92] Cox, David, John Little, and Donal O'Shea, *Ideals, Varieties, and Algorithms*, Undergraduate Texts in Math., Springer-Verlag, New York, 1992.

[CLO98] Cox, David, John Little, and Donal O'Shea, *Using Algebraic Geometry*, Grad. Texts in Math., Springer-Verlag, New York, 1998.

[F] Frobenius, G., Über das Pfaffsche Problem, J. Reine Angew. Math. **82** (1877), 230–315.

[G49] Gröbner, W., *Moderne Algebraische Geometrie*, Wien, Austria, 1949.

[G50] Gröbner, W., Über den Multiplizitätsbegriff in der algebraischen Geometrie, Math. Notes **4(1-6)** (1950/51), 194–201.

[G70] Gröbner, W., *Algebraische Geometrie II*, B.I.-Hochschultaschenbuch, Bibliographisches Institut Mannheim, Germany, 1970.

[H] Hakopian, Hakop A., On the regularity of multivariate Hermite interpolation, J. Approx. Theory **105(1)** (2000), 1–18.

[HT98] Hakopian, H. and M. Tonoyan, On two partial differential and one algebraic systems, Internat. Congr. Mathem., Short Communications, p.205,1998.

[HT02] Hakopian, Hakop and Mariam Tonoyan, On a multivariate theory, in *Approximation Theory*, B. D. Bojanov (ed), DARBA, Sofia, 2002, 212–230.

[HT04] Hakopian, Hakop A. and Mariam G. Tonoyan, Partial differential analogs of ordinary differential equations and systems, NY J. Math. **10** (2004), 89–116.

[I] Iarrobino, A., Reducibility of the families of 0-dimensional schemes on a variety, Inventiones math. **15** (1972), 72–77.

[K] Kergin, P., A natural interpolation of $C^k$ functions, J. Approx. Theory **29** (1980), 278–293.

[La] Lasker, E., Zur Theorie der Moduln und Ideale, Math. Ann. **60** (1905), 20–111?.

[Le] Lefranc, M., Analyse Spectrale sur $\mathrm{Z_n}$, C. R. Acad. Sci. Paris **246** (1958), 1951–1953.

[LL] Lorentz, G. G. and R. A. Lorentz, Multivariate interpolation, in *Rational Approximation and Interpolation,* P. R. Graves-Morris *et al* (eds), Lecture Notes in Math. 1105, Springer-Verlag, Berlin, 1984, 136–144.

[Lo92] Lorentz, R. A., *Multivariate Birkhoff Interpolation,* Lecture Notes in Mathematics, No. 1516, Springer Verlag, Heidelberg, 1992.

[Lo00] Lorentz, R. A., Multivariate Hermite interpolation by algebraic polynomials: a survey, J. Assoc. Comput. Mach. **122(1-2)** (2000), 167–201.

[Ma] Macaulay, F. S., *The Algebraic Theory of Modular Systems,* Cambridge Tracts in Math. and Math. Physics, no. 19, Cambridge University Press, Cambridge, England, 1916.

[Mi] Micchelli, C. A., A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B-splines and Lagrange interpolation, Rocky Mountain J. Math. **10** (1980), 485–497.

[M76] Möller, H. M., Mehrdimensionale Hermite-Interpolation und numerische Integration, Math. Z. **148** (1976), 107–118.

[M77] Möller, H. M., Hermite interpolation in several variables using ideal-theoretic methods, in *Constructive Theory of Functions of Several Variables, Oberwolfach 1976*, W. Schempp and K. Zeller (eds.), Springer Lecture Notes in Math. 571, Springer-Verlag, Berlin, 1977, 155–163.

[MSa] Möller, H. Michael and Thomas Sauer, H–bases i: the foundation, in *Curve and Surface Fitting: Saint-Malo 99,* A. Cohen, C. Rabut, and L. L. Schumaker (eds), Vanderbilt University Press, Nashville TN, 2000, 325–332.

[MSb] Möller, H. Michael and Thomas Sauer, H–bases ii: applications to numerical problems, in *Curve and Surface Fitting: Saint-Malo 99,* A. Cohen, C. Rabut, and L. L. Schumaker (eds), Vanderbilt University Press, Nashville TN, 2000, 333–342.

[MSc] Möller, H. Michael and Thomas Sauer, H–bases for polynomial interpolation and system solving, Advances in Comp. Math. **12** (2000), 335–362.

[MSt] Möller, H. Michael and Hans J. Stetter, Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems, Numer. Math. **70** (1995), 311–329.

[MT] Möller, H. Michael and Ralf Tenberg, Multivariate polynomial system solving using intersections of eigenspaces, J. Symbol. Comput. **30** (2001), 1–19.

[Mo] Mourrain, B., A new criterion for normal form algorithms, in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 13th Intern. Symp., AAECC-13, Honolulu, Hawaii USA, Nov.'99, Proc.*, Mark Fossorier, Hideki Imai, Shu Lin, Alan Pol (eds.), Springer Lecture Notes in Computer Science, 1719, Springer-Verlag, Heidelberg, 1999, 430–443.

[S97] Sauer, T., Polynomial interpolation of minimal degree, Numer. Math. **78(1)** (1997), 59–85.

[S98] Sauer, T., Polynomial interpolation of minimal degree and Gröbner bases, in *Groebner Bases and Applications (Proc. of the Conf. 33 Years of Groebner Bases), volume 251 of London Math. Soc. Lecture Notes,* B. Buchberger, F. Winkler (eds), Cambridge University Press, Cambridge, 1998, 483–494.

[S01] Sauer, T., Gröbner bases, H–bases and interpolation, Trans. Amer. Math. Soc. **353** (2001), 2293–2308.

[S02] Sauer, T., Ideal bases for graded polynomial rings and applications to interpolation, in *Multivariate Approximation and Interpolation with Applications (Monograph. Acad. Cienc. Zaragoza vol. 20),* M. Gasca (ed), Academia de Ciencias, Zaragoza, 2002, 97–110.

[S05] Sauer, T., Polynomial interpolation in several variables: lattices, differences, and ideals, in *Topics in Multivariate Approximation and Interpolation*, K. Jetter et al. (eds.), Elsevier, to appear.

[SX95a] Sauer, T. and Yuan Xu, On multivariate Lagrange interpolation, Math. Comp. **64** (1995), 1147–1170.

[SX95b] Sauer, T. and Yuan Xu, Multivariate Hermite interpolation, Advances in Comp. Math. **4** (1995), 207–259.

[Sh] Shekhtman, Boris, On ideal projectors on bivariate polynomials, ms, 1dec,2004.

[St] Stetter, Hans J., Matrix eigenproblems are at the heart of polynomial system solving, SIGSAM Bull. **30(4)** (1996), 22-25.

Carl de Boor
POB 1076, 81 Caprice Lane
Eastsound WA 98245 USA
deboor@cs.wisc.edu