A Leibniz formula for multivariate divided differences

Carl de Boor

8oct02

Abstract The Leibniz formula, for the divided difference of a product, and Opitz's formula, for the divided difference table of a function as the result of evaluating that function at a certain matrix, are shown to be special cases of a formula available for the coefficients, with respect to any basis, of an 'ideal' or 'Hermite' polynomial interpolant, in any number of variables.

1. Introduction. The so-called Leibniz formula

$$\Delta(x_i,\ldots,x_j)(fg) = \sum_{k=i}^{j} \Delta(x_i,\ldots,x_k)f\ \Delta(x_k,\ldots,x_j)g, \tag{1.1}$$

for the divided difference of a product in terms of the divided differences of the factors, has played a major role in the development of spline theory; it was an essential tool in the derivation of the B-spline recurrence relations. My earliest reference for it now is [P33 : p. 12] who refers, for the case of uniform spacing, to [J20] where, on page 105, that formula is referred to as 'bekannt'. Nevertheless, the formula is generally credited (see, e.g., [O64]) to Steffensen, because of his paper [S39].

In this note, the algebraic background of the Leibniz formula is explored, showing the formula to be equivalent to Opitz's formula (from [O64]; see (2.1) below) that gives the divided difference table of any polynomial as the result of applying that polynomial to a certain matrix. This, in turn, is shown to be a particular consequence of the fact that, in G. Birkhoff's [B79] terminology, polynomial interpolation is an 'ideal' interpolation scheme. This insight is used to explore Leibniz (and Opitz) formulæ for certain *multivariate* polynomial interpolation schemes and their associated divided differences.

This note is laid out as follows. In Section 2, the connection between the Leibniz formula and the Opitz formula is recalled, along with Opitz's way of deriving them. The next section brings a brief discussion of the basic features of 'ideal' interpolation, i.e., linear projectors on the space of polynomials (in one or several variables, real or complex) whose kernel is a polynomial ideal. Section 4 provides the Opitz formula in the general setting of 'ideal' interpolation, and the truncated Taylor series serves as a trivial illustration. The nontrivial details for both the Opitz and the Leibniz formula are fully worked out for Chung-Yao interpolation, in Section 6. Such formulas for other divided differences are outlined in Section 7. The final section points out that this paper's restriction to interpolation to polynomials is easily removed.

For ready reference, here is the (mostly, but not entirely, standard) notation used in this note. $\alpha \in \mathbb{Z}_+^d$ denotes a **multiindex** or, more precisely, a d-index, i.e., a d-vector with nonnegative integer entries; $|\alpha| := \sum_j \alpha(j)$ is its **length** (or '**degree**'); also, $\alpha! := \prod_j \alpha(j)!$. There being no standard notation for it, I use

$$()^\alpha : \mathbb{F}^d \to \mathbb{F} : x \mapsto x^\alpha := \prod_j x(j)^{\alpha(j)}$$

for the monomial of **multidegree** α. Here, $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$, though usually it is $\mathbb{C}$. With this,

$$\Pi_{\mathbb{I}} := \operatorname{span}(()^\alpha : \alpha \in \mathbb{I}), \quad \mathbb{I} \subset \mathbb{Z}_+^d,$$

with the special cases

$$\Pi := \Pi(\mathbb{F}^d) := \Pi_{\mathbb{Z}_+^d}, \quad \Pi_k := \operatorname{span}(()^\alpha : |\alpha| \le k).$$

The *ad hoc* abbreviation

$$\widehat{p}(\alpha) := (D^\alpha p)(0)/\alpha!, \qquad p \in \Pi,\ \alpha \in \mathbb{Z}_+^d,$$

with

$$D^\alpha := \prod_j D_j^{\alpha(j)}$$

and D_j differentiation with respect to the jth argument, is convenient. Analogously,

$$()_j : x \mapsto x(j), \quad j = 1{:}d,$$

while

$$()_0 : x \mapsto 1.$$

In the dual, Π', of Π, *evaluation at some point* $v \in \mathbb{F}^d$ is singled out, i.e., the linear functional

$$\epsilon_v : \Pi \to \mathbb{F} : p \mapsto p(v),$$

and, more generally, $\epsilon_v q(D) : p \mapsto (q(D)p)(v)$ for $q \in \Pi$, with

$$q(D) := \sum_\alpha \widehat{q}(\alpha) D^\alpha.$$

Also,

$$Q(D) := \{q(D) : q \in Q\}, \qquad Q \subset \Pi,$$

and

$$\Lambda_\perp := \ker \Lambda := \cap_{\lambda\in\Lambda} \ker \lambda, \qquad \Lambda \subset \Pi'.$$

2. The Opitz formula. In his short note [O64], describing a talk submitted but not given, G. Opitz introduces 'Steigungsmatrizen' (lit.: 'divided difference matrices') as matrices of the form

$$S[f; X] := f(A_X),$$

with f a (univariate) polynomial or rational function or, more generally, a suitable limit of such functions, and, correspondingly, $f(A_X)$ the 'value' of f at the matrix A_X, with

$$A_X := \begin{bmatrix} x_1 & 1 & & & \\ & x_2 & 1 & & \\ & & x_3 & \ddots & \\ & & & \ddots & 1 \\ & & & & x_n \end{bmatrix},$$

and with $X := (x_1, \ldots, x_n)$ a sequence of pairwise distinct complex numbers. The notation $S[f; X]$ for these 'Steigungsmatrizen' is his. Using the (obvious) eigenstructure of A_X, Opitz readily concludes that, for each i, j,

$$S[f; X](i, j) = \Delta(x_i, \ldots, x_j) f, \tag{2.1}$$

i.e., the divided difference of f at $(x_i, \ldots, x_j)$ (in W. Kahan's felicitous notation[1]), hence the name 'Steigungsmatrix'. Here, as is customary, $\Delta(x_i, \ldots, x_j) := 0$ for $i > j$.

In other words, $f(A_X)$ is (or, the upper triangular part of $f(A_X)$ provides) the divided difference table for f with respect to the sequence X, and, as Opitz points out, its calculation in this fashion from A_X is less affected by loss of significance than is the direct construction of the divided difference table by the repeated formation of divided differences. In fact, it can be used for the symbolic calculation of divided differences; see, e.g., [KF85], and, most recently, [RR01].

Further, Opitz observes that the map

$$f \mapsto S[f; X]$$

is linear as well as multiplicative, hence a ring homomorphism, from the ring of functions under pointwise addition and multiplication into the ring of matrices of order n. In particular,

$$(fg)(A_X) = f(A_X) g(A_X).$$

[1] I am using here Kahan's notation not only because it is quite literal, but because the standard notation, $[x_i, \ldots, x_j]$, has already other uses, e.g., the matrix with columns $x_i, \ldots, x_j$ or, in the case $j = i + 1$, the closed interval with endpoints x_i, x_j.

Because of (2.1), this is equivalent to the Leibniz formula, (1.1), i.e., to

$$\Delta(x_i,\ldots,x_j)(fg) = \sum_{k=i}^{j} \Delta(x_i,\ldots,x_k)f\ \Delta(x_k,\ldots,x_j)g.$$

Further, if we take (2.1) as the *definition* of $S[f;X]$, then the Leibniz formula implies that $f \mapsto S[f;X]$ is a ring homomorphism and so, in particular, $S[f;X] = f(A_X)$.

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

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

(with $P' : F' \to F' : \lambda \mapsto \lambda P$ the dual of P). In other words, given that $\ker P = \operatorname{ran}(\mathrm{id} - P)$, we have

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

In this way, $\operatorname{ran} P'$ provides the *interpolation conditions* matched by P. Not surprisingly, there are exactly as many independent conditions as there are degrees of freedom, i.e.,

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

Now we take

$$F = \Pi,$$

the ring of polynomials in d (complex) variables. In [B79], Garrett Birkhoff defined **ideal interpolation** as any linear projector P on Π whose nullspace or kernel is an ideal. In the interest of brevity, and without passing judgement, we will call such a projector **ideal**. However, Birkhoff seemed not to have been aware of the fact that ideal projectors had already been looked at carefully before that, by Möller in [Moe76], who called them 'Hermite interpolation', for the following reason.

As is well-known (and, in this formulation, probably due to Gröbner ; see [G70 : p. 176]), a nonempty subset I of Π is an ideal of finite codimension if and only if

$$I = \cap_{v\in\mathcal{V}} \ker(\epsilon_v Q_v(D))$$

for some finite subset $\mathcal{V}$ of $\mathbb{C}^d$ (necessarily the ideal's variety) and some nontrivial D-invariant finite-dimensional polynomial subspaces Q_v, necessarily given by

$$Q_v := \{q \in \Pi : ((D^\alpha q)(D)p)(v) = 0, \alpha \in \mathbb{Z}_+^d, p \in I\}.$$

In other words, as Möller rightly stresses, ideal interpolation is characterized by the fact that its interpolation conditions involve values and, possibly, also derivatives at certain sites, subject only to the condition that if the linear functional $\epsilon_v q(D)$ is matched, then so are all 'lower' derivatives, i.e., every $\epsilon_v(D^\alpha q)(D)$ for $\alpha \in \mathbb{Z}_+^d$.

Since an ideal projector is, in a sense, aware of the multiplicative structure of Π, we would expect insights from considering its interaction with multiplication. The following lemma gives this interaction a handy formulation.

Lemma. *A linear projector P on Π is ideal if and only if*

$$P(pq) = P(pPq), \quad \forall p, q \in \Pi. \tag{3.1}$$

Proof. The condition (3.1) 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. □

It is standard in Algebraic Geometry (see, e.g., [CLO98 : p. 51ff]) to consider, on the quotient ring

$$\Pi/\mathcal{I} := \{f + \mathcal{I} : f \in \Pi\}$$

of the polynomials over the ideal $\mathcal{I}$ and for an arbitrary polynomial p, the map

$$\Pi/\mathcal{I} \to \Pi/\mathcal{I} : f + \mathcal{I} \mapsto pf + \mathcal{I}.$$

In our context, it is more convenient to consider, equivalently, the map

$$M_p : \operatorname{ran} P \to \operatorname{ran} P : f \mapsto P(pf). \tag{3.2}$$

Evidently,

$$M_p \in L(\operatorname{ran} P),$$

i.e., M_p is a linear map on $\operatorname{ran} P$. Further, (3.1) implies that, for arbitrary $p, q \in \Pi$ and $f \in \operatorname{ran} P$,

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

It follows that the map

$$m : \Pi \to L(\operatorname{ran} P) : p \mapsto M_p \tag{3.3}$$

is a ring homomorphism onto the commutative algebra generated by the specific linear maps

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

in terms of which

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

with

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

independent of the order in which this product is formed from its factors.

It follows, directly from (3.1), that

$$p(M)P()_0 = P(p\,P()_0) = Pp, \quad p \in \Pi. \tag{3.4}$$

Such a formula plays a major role in Mourrain's intriguing paper [Mou99], though it is proved there, consistent with that paper's setting, only for P whose range, $B := \operatorname{ran} P$, is **connected to 1**, meaning that each $b \in B$ can be written in the form $\sum_{j=0}^d ()_j b_j$ with each b_j in $B \cap \Pi_{<\deg b}$, hence, in particular, $()_0 \in B$, and (3.4) simplifies to $p(M)()_0 = Pp$.

(3.4) implies that $\ker m \subset \ker P$, while, if $p \in \ker P$, then $p(M)f = P(pf) = P(fPp) = P0 = 0$ for all $f \in \operatorname{dom} p(M) = \operatorname{ran} P$, i.e., $p(M) = 0$. Thus, altogether,

$$\ker m = \ker P. \tag{3.5}$$

4. A general Opitz formula. If now

$$V : \mathbb{F}^n \to \operatorname{ran} P : a \mapsto \sum_j v_j a(j) =: [v_1, \ldots, v_n]a$$

is any basis for $\operatorname{ran} P$, i.e., $V = [v_1, \ldots, v_n]$ is an invertible linear map, then the matrix representation for $M_p = p(M)$ with respect to this basis is

$$\widehat{M_p} = V^{-1} M_p V = p(\widehat{M}), \tag{4.1}$$

with

$$\widehat{M_j} = V^{-1} M_j V, \quad j = 1{:}d.$$

Consequently,

$$P(pv_j) = p(M)v_j = Vp(\widehat{M})(:, j), \quad p \in \Pi. \tag{4.2}$$

In particular,

$$Pp = p(M)P()_0 = Vp(\widehat{M})a_0, \quad p \in \Pi,$$

with $a_0 := V^{-1}P()_0$ the coordinates of $P()_0$ with respect to V.

(4.1), (4.2) is the promised generalization of Opitz's formula.

To make the connection with (2.1), take, in particular, $d = 1$, and let $P = P_n$ be the linear projector of interpolation from polynomials of degree $< n$ to data at the distinct sites $x_1, \ldots, x_n$. Choosing, specifically, for V the Newton basis

$$v_j := \prod_{j<k\le n} (\cdot - x_k), \quad j = 1{:}n,$$

we compute the jth column of $\widehat{M} := \widehat{M_1}$ as the coordinates, with respect to V, of

$$M_1 v_j = P_n(()_1 v_j) = P_n(x_j v_j + (\cdot - x_j)v_j) = x_j v_j + P_n v_{j-1} = x_j v_j + \begin{cases} v_{j-1}, & j > 1; \\ 0 & \text{otherwise}, \end{cases}$$

hence

$$\widehat{M} = \begin{bmatrix} x_1 & 1 & & & \\ & x_2 & 1 & & \\ & & x_3 & \ddots & \\ & & & \ddots & 1 \\ & & & & x_n \end{bmatrix} = A_X.$$

Consider now $p(M)v_j = P_n(pv_j)$. Certainly, $(P_j p)v_j$ is in $\operatorname{ran} P_n$ and matches pv_j at all the x_i, hence must equal $P_n(pv_j)$. Therefore,

$$p(M)v_j = \left(\sum_{k=1}^{j} \prod_{k<h\le j} (\cdot - x_h) \ \Delta(x_k, \ldots, x_j)p \right) v_j = \sum_{k=1}^{j} v_k \Delta(x_k, \ldots, x_j)p.$$

Consequently,

$$p(\widehat{M})(k, j) = \Delta(x_k, \ldots, x_j)p, \qquad k, j = 1{:}n.$$

Since Opitz [O64] bases his derivations on the eigenstructure of the matrix A_X, it seems appropriate to point out that it is standard in Algebraic Geometry (see, e.g., [CLO98 : p. 54ff]) to consider the eigenstructure of the linear maps M_p (defined in (3.2)). To be sure, it is their dual, more precisely the matrix $\mathcal{M}_p$, called a **multiplication table** and defined implicitly by

$$\langle ()^\alpha p \rangle =: \sum_{\beta \in \mathbb{I}} \mathcal{M}_p(\alpha, \beta) \langle ()^\beta \rangle, \qquad \alpha \in \mathbb{I}$$

(with $\langle f \rangle := f + \mathcal{I}$ and $\mathbb{I}$ the set of multidegrees that don't occur among the multidegrees of elements of the ideal) whose eigenstructure is given, by H. Stetter and his collaborators, a major role in the solving of polynomial systems; see, e.g., [AS88], [MoeS95]. But I find it more convenient to deal with the linear maps M_p.

The bare facts are these: For each v in the variety $\mathcal{V} := \mathcal{V}(\ker P)$ of the ideal $\ker P$, $\epsilon_v \in \operatorname{ran} P'$, hence, for every $f \in B := \operatorname{ran} P$,

$$\epsilon_v M_p f = \epsilon_v P(pf) = \epsilon_v(pf) = p(v)\epsilon_v f,$$

and this shows ϵ_v (or, more precisely, $\epsilon_v|_B$) to be a left eigenvector of M_p, with corresponding eigenvalue $p(v)$. Hence, if we are dealing with Lagrange interpolation (as is the case in [O64] at the outset), i.e., if $(\epsilon_v : v \in \mathcal{V})$ spans $\operatorname{ran} P'$, then M_p is diagonalizable, and $\{p(v) : v \in \mathcal{V}\}$ is its spectrum. In that case, a right eigenbasis for M_p is the basis $(\ell_v : v \in \mathcal{V})$ of $\operatorname{ran} P$ dual to $(\epsilon_v : v \in \mathcal{V})$, i.e., $\ell_v(w) = \delta_{vw}$, the Lagrange basis. Further, $\{p(v) : v \in \mathcal{V}\}$ is also the spectrum of M_p in the general case, with each $q \in Q_v$ that is not in $\sum_{j=1}^d D_j Q_v$ giving rise to a (right) eigenvector of M_p for the eigenvalue $p(v)$.

5. An Example: the truncated Taylor series. As a first (and trivial) d-variate example with $d > 1$, consider $P = T_k$, the linear map on Π that associates with $p \in \Pi$ its Taylor expansion

$$T_k p := \sum_{|\alpha|<k} ()^\alpha D^\alpha p(0)/\alpha!$$

of order k. Evidently,

$$\operatorname{ran} T_k' = \epsilon_0 \Pi_{<k}(D),$$

thus

$$\ker T_k = \operatorname{ideal}(()^\alpha : |\alpha| = k).$$

In particular, with

$$V_{<k} := [()^\alpha : |\alpha| < k]$$

the power basis for $\Pi_{<k} = \operatorname{ran} T_k$, we find $()_j()^\alpha \in \operatorname{ran} T_k$ iff $|\alpha| < k-1$, while, for $|\alpha| = k-1$, $P(()_j()^\alpha) = 0$. Hence, with $\iota_j := (\delta_{ij} : i = 1{:}d)$,

$$\widehat{M}_j(\alpha, \beta) = \delta_{\beta+\iota_j-\alpha}, \qquad |\alpha|, |\beta| < k,$$

a strictly lower triangular matrix in any total ordering of $\mathbb{Z}_+^d$ that respects 'degree', i.e., for which $|\alpha| < |\beta| \implies \alpha < \beta$. It reflects the evident fact that the action of $\widehat{M}_j$ is to shift the coefficient function

$$\widehat{p} : \alpha \to \widehat{p}(\alpha) = D^\alpha p(0)/\alpha!$$

by ι_j, i.e.,

$$\widehat{M_j p} = \widehat{p}(\cdot - \iota_j),$$

dropping off those terms that are, thereby, pushed outside the relevant index set, $\{\alpha : |\alpha| < k\}$.

Correspondingly (or directly by (4.2)), the αth column of $p(\widehat{M})$ is obtained from $\widehat{p}$ by a shift of $\widehat{p}$ by α, again dropping off those terms that are, thereby, pushed outside $\{\alpha : |\alpha| < k\}$, i.e.,

$$p(\widehat{M})(:, \alpha) = \widehat{p}(\cdot - \alpha).$$

In particular, for any $p, q \in \Pi$,

$$\widehat{(pq)}(\alpha) = (pq)(\widehat{M})(\alpha, 0) = p(\widehat{M})q(\widehat{M})(\alpha, 0) = \sum_{\beta \le \alpha} \widehat{p}(\alpha - \beta)\widehat{q}(\beta),$$

the familiar Leibniz formula for the derivative of a product.

6. An Example: Chung-Yao interpolation. In [CY77], Chung and Yao introduced the eponymous multivariate polynomial interpolation scheme. This scheme provides interpolation from Π_k to data at the sites

$$\Theta_{\mathbb{H}} := \{\theta_H : H \in \binom{\mathbb{H}}{d}\},$$

with $\mathbb{H}$ a set of $d+k$ hyperplanes in $\mathbb{R}^d$ in general position and θ_H the unique point common to the d hyperplanes in such an $H \in \binom{\mathbb{H}}{d}$. Chung and Yao [CY77] show that such interpolation is possible and uniquely so, by exhibiting the interpolant $P_{\mathbb{H}}g$ to g in Lagrange form.

[dB95] (see [dB97] for details) provides the following Newton form for $P_{\mathbb{H}}g$:

$$P_{\mathbb{H}}g = \sum_{j=0}^{k} \sum_{K\in\binom{\mathbb{H}_{j-1}}{d-1}} p_{j-1,K}\,[\Theta_{\mathbb{H}_j,K} \mid n_K,\ldots,n_K]g, \tag{6.1}$$

with the various terms occurring here defined as follows.

$$\mathbb{H}_{-1} \subset \cdots \subset \mathbb{H}_k := \mathbb{H}$$

is any increasing sequence of subsets of $\mathbb{H}$ with $\#\mathbb{H}_j = d+j$, all j. Further,

$$p_{j,K} := \prod_{h\in\mathbb{H}_j\backslash K} \frac{h}{h_\uparrow(n_K)},$$

with h denoting a hyperplane as well as a particular linear polynomial whose zero set coincides with that hyperplane, and $h_\uparrow$ its *leading term*, i.e., its linear homogeneous part. Also,

$$\Theta_{\mathbb{K},K} := \Theta_{\mathbb{K}} \cap l_K,$$

with

$$l_K := \cap_{h\in K} h$$

the straight line common to the $d-1$ hyperplanes in K, while

$$n_K$$

is an arbitrary nontrivial vector parallel to that line. Last, but certainly not least,

$$[X \mid \Xi]g := \int_{[X]} D_\Xi g$$

is the multivariate divided difference (notation) introduced in [dB95]. Here, $X = (x_0,\ldots,x_n)$ and $\Xi = (\xi_1,\ldots,\xi_n)$ are arbitrary sequences in $\mathbb{R}^d$, the first one having one more entry than the second, $D_\Xi := D_{\xi_1}\cdots D_{\xi_n}$ is the composition of directional derivatives $D_\xi := \sum_j \xi(j)D_j$, and

$$f \mapsto \int_{[x_0,\ldots,x_n]} f := \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} f(x_0 + s_1\nabla x_1 + \cdots + s_n\nabla x_n)\,\mathrm{d}s_n \cdots \mathrm{d}s_1 \tag{6.2}$$

(with $\nabla x_j := x_j - x_{j-1}$) is termed, by Micchelli in [Mi79], the *divided difference functional on* $\mathbb{R}^d$ and is familiar from the Genocchi-Hermite formula for the univariate divided difference. $[X \mid \Xi]$ is symmetric in the 'sites' $x \in X$, and is linear and symmetric in the 'directions' $\xi \in \Xi$, and satisfies the recurrence

$$[X \mid \Xi][X', \cdot \mid \Xi'] = [X, X' \mid \Xi, \Xi'].$$

Let now

$$V := [p_{j,K} : (j,K) \in \mathbb{I}], \qquad \text{with } \mathbb{I} := \{(j,K) : K \in \binom{\mathbb{H}_j}{d-1}, j = -1{:}(k-1)\},$$

be the corresponding 'Newton' basis for $\operatorname{ran} P = \Pi_k$. For $j = 0{:}k$, let h_j be the sole element of $\mathbb{H}_j\backslash\mathbb{H}_{j-1}$, pick $K \in \binom{\mathbb{H}_{j-1}}{d-1}$, and let $H := K \cup h_j$. Then

$$(x - \theta_H) = \sum_{h\in H} n_{H\backslash h} \frac{h(x)}{h_\uparrow(n_{H\backslash h})}.$$

This implies that

$$\begin{aligned} x p_{j-1,K}(x) &= (\theta_H \ + \ (x-\theta_H))p_{j-1,K}(x) \\ &= \theta_H p_{j-1,K}(x) + \sum_{h\in H} n_{H\backslash h}\left(\prod_{h'\in \mathbb{H}_j\backslash H} \frac{h'_\uparrow(n_{H\backslash h})}{h'_\uparrow(n_K)}\right) p_{j,H\backslash h}(x). \end{aligned}$$

Notice that each of the $p_{j,H\backslash h}$ in the sum over H vanishes on $\Theta_{\mathbb{H}_j}$. In particular, for $j=k$, the sum over H vanishes for every $x\in\Theta_{\mathbb{H}}$. It follows that, for $i=1{:}d$, the matrix representation $\widehat{M_i}$ for $M_i : f\mapsto P(()_i f)$ with respect to the 'Newton' basis V is 'lower triangular' and quite sparse, with the column corresponding to $p_{j-1,K}$ having nonzero entries only on the diagonal, where it has the value $\theta_{h_j\cup K}(i)$, and at the entries, if any, corresponding to $p_{j,h_j\cup K\backslash h}$ for $h\in h_j\cup K$.

Now, what about $f(\widehat{M})$ for arbitrary $f\in\Pi$? The polynomial $fp_{j-1,K}$ vanishes on $\Theta_{\mathbb{H}_{j-1}}$, hence depends only on f restricted to $\Theta_{\mathbb{H}}\backslash\Theta_{\mathbb{H}_{j-1}}$. However, this dependence is hardly simple. Formally, we have

$$f(\widehat{M})((j,K),(j',K')) = [\Theta_{\mathbb{H}_{j+1},K} \mid n_K,\ldots,n_K](fp_{j',K'}), \quad (j,K),(j',K')\in\mathbb{I}.$$

The fact that $f(\widehat{M})$ is lower-triangular, in any ordering of the index set $\mathbb{I}$ that refines the natural partial ordering provided by the first components, is evident.

With this, from the fact that $(fg)(\widehat{M}) = f(\widehat{M})g(\widehat{M})$, we get the following 'Leibniz formula':

$$(6.3)\quad [\Theta_{\mathbb{H}_j,K}\mid n_K,\ldots,n_K](fg) = \sum_{(j',K')\in\mathbb{I};j'<j} [\Theta_{\mathbb{H}_j,K}\mid n_K,\ldots,n_K](fp_{j',K'})\,[\Theta_{\mathbb{H}_{j'+1},K'}\mid n_{K'},\ldots,n_{K'}]g.$$

Note that the second factor depends only on g on the sites $\Theta_{\mathbb{H}_{j'+1}}$, while the first factor depends only on f on the sites in $\Theta_{\mathbb{H}_j}\backslash\Theta_{\mathbb{H}_{j'}}$. In particular, the first factor is trivially zero when $j'\ge j$, hence the sum's restriction to $j'<j$.

Note also, by way of a check, that, for $d=1$, $\mathbb{H}$ consists of pairwise distinct points, with $\mathbb{H}_j$ containing $j+1$ points, $h_0,\ldots,h_j$, say. Further, $K=\emptyset$ is the sole element of $\binom{\mathbb{H}_j}{d-1}$, and $l_\emptyset=\mathbb{R}$, hence we may choose ι_1 for $n_\emptyset$ and, with that,

$$[\Theta_{\mathbb{H}_j,K}\mid n_K,\ldots,n_K] = \Delta(h_0,\ldots,h_j),$$

by the Genocchi-Hermite formula, while, as observed earlier,

$$\Delta(h_0,\ldots,h_j)(\prod_{i<j'}(\cdot-h_i)f) = \Delta(h_{j'},\ldots,h_j)f.$$

This verifies that, indeed, (6.3) reduces to (1.1) when $d=1$.

7. Other divided differences. Let T be an arbitrary finite subset of $\mathbb{C}^d$ and assume that the polynomial subspace B is correct for it in the sense that

$$\Lambda^{\mathrm{t}}_{\mathrm{T}} : B\to\mathbb{C}^{\mathrm{T}} : b\mapsto b|_{\mathrm{T}}$$

is 1-1 and onto. Then, with

$$W:\mathbb{C}^W\to B : a\mapsto\sum_{w\in W} a(w)w$$

an arbitrary basis for B (using W to denote both the basis and the associated basis map), the Gram matrix

$$\Lambda^{\mathrm{t}}_{\mathrm{T}}W = (w(\tau) : \tau\in\mathrm{T}, w\in W)$$

is invertible, hence, for any particular ordering of the basis W, there is some ordering of T so that

$$\Lambda^{\mathrm{t}}_{\mathrm{T}}W = LU,$$

with L lower triangular and U unit upper triangular (in the chosen orderings of T and W). Then one is free to call

$$\lambda(\tau_1, \ldots, \tau_i) := \sum_k L^{-1}(i,k)\epsilon_{\tau_k} = \sum_{k\le j} L^{-1}(i,k)\epsilon_{\tau_k}$$

the 'divided difference' at the sequence $(\tau_1, \ldots, \tau_i)$, and to call, correspondingly, the polynomials

$$v_j := \sum_k w_k U^{-1}(k,j) = \sum_{k\le j} w_k U^{-1}(k,j)$$

'Newton polynomials', and to call

$$\sum_j v_j\ \lambda(\tau_1, \ldots, \tau_j) f$$

the 'Newton form' of the interpolant from B to f at T. Assuming that B contains the constant function and that, in fact, $v_1 = ()_0$, it then follows that

$$\lambda(\tau_1, \ldots, \tau_j)(fg) = \sum_{k=1}^{j} \lambda(\tau_1, \ldots, \tau_j)(f v_k)\, \lambda(\tau_1, \ldots, \tau_k) g,$$

with $\lambda(\tau_1, \ldots, \tau_j)(f v_k)$ only depending on f at $\tau_k, \ldots, \tau_j$. The role reversal of f and g here as compared to (1.1) is due to the fact that the 'Newton' basis here is ordered differently than there.

It is in this manner, or, perhaps, in a more relaxed block-triangular way, that one could provide some kind of Leibniz formula and even an Opitz formula in the context of more general schemes of multivariate polynomial interpolation, e.g., the least interpolant of [dBR90], or the Sauer-Xu formulation [SX95].

The divided difference introduced by Rabut in [R01] does not quite fit this pattern. For, while Rabut does define divided differences as the coefficients of the interpolating polynomial, he sticks to the power basis

$$V_k := [()^\alpha : |\alpha| \le k]$$

rather than some kind of multivariate Newton basis. Precisely, with T some pointset in $\mathbb{R}^d$ correct for interpolation from Π_k, hence

$$P := V_k (\Lambda_\mathrm{T}^\mathrm{t} V_k)^{-1} \Lambda_\mathrm{T}^\mathrm{t}$$

well-defined, he denotes the (T, α)**-divided difference of** f by

$$f[\mathrm{T}]^\alpha$$

and defines it implicitly by

$$Pf =: \sum_\alpha ()^\alpha f[\mathrm{T}]^\alpha.$$

With this definition, it follows from (4.2) that

$$(p[\mathrm{T}]^\alpha : |\alpha| \le k) = p(\widehat{M})(:,0), \quad p \in \Pi,$$

hence that

$$(pq)[\mathrm{T}]^\alpha = \sum_\beta (p()^\beta)[\mathrm{T}]^\alpha\, q[\mathrm{T}]^\beta = \sum_{\beta\le\alpha} (p()^\beta)[\mathrm{T}]^\alpha\, q[\mathrm{T}]^\beta.$$

However, since $f[\mathrm{T}]^\alpha$ depends on f on all of T, the first factor in each summand still depends, offhand, on p on all of T.

In Rabut's setting, the matrix representation $\widehat{M}_j$ of

$$M_j : \Pi_k \to \Pi_k : p \mapsto P(()_j p)$$

is, in principle, not that hard to work out. For $|\alpha| \le k$, we have $()_j()^\alpha \in \Pi_k$ if and only if $|\alpha| < k$. Therefore

$$\widehat{M}_j(\alpha,\beta) = \begin{cases} \delta_{\beta+\iota_j-\alpha}, & |\beta| < k; \\ ()^{\beta+\iota_j}[\mathrm{T}]^\alpha, & |\beta| = k. \end{cases}$$

However, this still leaves the particular details of the specific divided differences $()^{\beta+\iota_j}[\mathrm{T}]^\alpha$ for $|\beta| = k$ to be supplied. At this point, I do not know whether it would be worthwhile to make that effort.

8. Extensions. In contrast to the standard literature on polynomial interpolation and divided differences, I have restricted here attention to interpolation to polynomials. However, since a polynomial interpolant only depends on the values at the interpolation sites of the function being interpolated, interpolation extends immediately to any function having values at least at the interpolation sites, and this leads to a natural extension, to such functions, of whatever divided difference notion or polynomial interpolation scheme is used.

In the univariate setting, if the interpolation involves 'repeated' sites, i.e., matching of certain 'consecutive' derivatives, then, correspondingly, the interpolation scheme and the divided differences extend to functions suitably differentiable at the interpolation sites. The same holds for multivariate ideal interpolation, except that, at present, it is not known whether every such Hermite interpolation scheme can be viewed as the limit of suitable Lagrange interpolation schemes, i.e., whether in this sense multivariate Hermite interpolation can be viewed as interpolation involving 'repeated' sites.

Acknowledgements I am grateful to H. Michael Möller for various discussions concerning the Algebraic Geometry literature and, in particular, for pointing me toward Hans J. Stetter, and to the latter for pointing out to me B. Mourrain's paper. I also thank the two referees for a most careful reading of the ms which further strengthened my belief that it is impossible to get anything right.

References

[AS88] W. Auzinger and H. Stetter, *An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations*, Numerical Mathematics, Singapore 1988 (Proc. Internat. Conf. Numer. Math., ISNM vol. 86) (Ravi P. Agarwal, Y. M. Chow, and S. J. Wilson, eds.), Birkhäuser, Basel, 1988, pp. 11–30.

[B79] G. Birkhoff, *The algebra of multivariate interpolation*, Constructive approaches to mathematical models (C. V. Coffman and G. J. Fix, eds.), Academic Press, New York, 1979, pp. 345–363.

[dB95] C. de Boor, *A multivariate divided difference*, Approximation Theory VIII, Vol. 1: Approximation and Interpolation (Charles K. Chui and Larry L. Schumaker, eds.), World Scientific Publishing Co., Inc., Singapore, 1995, pp. 87–96.

[dB97] C. de Boor, *The error in polynomial tensor-product, and in Chung-Yao, interpolation*, Surface Fitting and Multiresolution Methods (A. LeMéhauté, C. Rabut, and L. L. Schumaker, eds.), Vanderbilt University Press, Nashville TN, 1997, pp. 35–50.

[dBR90] C. de Boor and A. Ron, *On multivariate polynomial interpolation*, Constr. Approx., 6 (1990), pp. 287–302.

[CY77] K. C. Chung and T. H. Yao, *On lattices admitting unique Lagrange interpolations*, SIAM J. Numer. Anal., 14 (1977), pp. 735–743.

[CLO98] David Cox, John Little, and Donal O'Shea, *Using Algebraic Geometry*, Graduate Texts in Mathematics 185, Springer, New York, 1998.

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

[J20] M. E. Jacobsthal, *Mittelwertbildung und Reihentransformation*, Math. Z., 6 (1920), pp. 100–117.

[KF85] W. Kahan and R. Fateman, *Symbolic Computation of Divided Differences*, unpublished report, available at `http://www.cs.berkeley.edu/~fateman/papers/divdiff.pdf`, 1985.

[Mi79] C. A. Micchelli, *On a numerically efficient method for computing multivariate B-splines*, Multivariate Approximation Theory (W. Schempp and K. Zeller, eds.), Birkhäuser, Basel, 1979, pp. 211–248.

[Moe76] H. M. Möller, *Mehrdimensionale Hermite-Interpolation und numerische Integration*, Math. Z., 148 (1976), pp. 107–118.

[MoeS95] H. Michael Möller and Hans J. Stetter, *Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems*, Numer. Math., 70 (1995), pp. 311–329.

[Mou99] B. Mourrain, *A new criterion for normal form algorithms*, 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, pp. 430–443.

[O64] G. Opitz, *Steigungsmatrizen*, Z. Angew. Math. Mech., 44 (1964), pp. T52–T54.

[P33] Tibère Popoviciu, "Sur quelques proprietés des fonctions d'une ou de deux variables réelles", dissertation, presented to the Faculté des Sciences de Paris, published by Institutul de Arte Grafice "Ardealul" (Cluj, Romania), 1933.

[R01] C. Rabut, *Multivariate divided differences with simple knots*, SIAM J. Numer. Anal., 38(4) (2001), pp. 1294–1311.

[RR01] Thomas W. Reps and Louis B. Rall, *Computational divided differencing and divided-difference arithmetics*, Higher-Order and Symbolic Computations, xx (200x), pp. xxx–xxx.

[SX95] T. Sauer and Yuan Xu, *On multivariate Lagrange interpolation*, Math. Comp., 64 (1995), pp. 1147–1170.

[S39] J. F. Steffensen, *Note on divided differences*, Danske Vid. Selsk. Math.-Fys. Medd., 17(3) (1939), pp. 1–12.