On multivariate polynomial interpolation

Carl de Boor[1] & Amos Ron

ABSTRACT

We provide a map $\boldsymbol{\Theta} \mapsto \Pi_{\boldsymbol{\Theta}}$ which associates each finite set $\boldsymbol{\Theta}$ of points in $\mathbb{C}^s$ with a polynomial space $\Pi_{\boldsymbol{\Theta}}$ from which interpolation to arbitrary data given at the points in $\boldsymbol{\Theta}$ is possible and uniquely so. Among all polynomial spaces Q from which interpolation at $\boldsymbol{\Theta}$ is uniquely possible, our $\Pi_{\boldsymbol{\Theta}}$ is of smallest degree. It is also D- and scale-invariant. Our map is monotone, thus providing a Newton form for the resulting interpolant. Our map is also continuous within reason, allowing us to interpret certain cases of coalescence as Hermite interpolation. In fact, our map can be extended to the case where, with each $\theta \in \boldsymbol{\Theta}$, there is associated a polynomial space P_θ, and, for given smooth f, a polynomial $q \in Q$ is sought for which

$$p(D)(f-q)(\theta) = 0, \quad \text{all } p \in P_\theta,\ \theta \in \Theta.$$

We obtain $\Pi_{\boldsymbol{\Theta}}$ as the "scaled limit at the origin" $(\exp_{\boldsymbol{\Theta}})_\downarrow$ of the exponential space $\exp_{\boldsymbol{\Theta}}$ with frequencies $\boldsymbol{\Theta}$, and base our results on a study of the map $H \mapsto H_\downarrow$ defined on subspaces H of the space of functions analytic at the origin. This study also allows us to determine the local approximation order from such H and provides an algorithm for the construction of $H_\downarrow$ from any basis for H.

AMS (MOS) Subject Classifications: primary 41A05, 41A63, 41A10; secondary 65D05, 41A30

Key Words: exponentials, polynomials, multivariate, interpolation, Newton form, Birkhoff interpolation

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

[1] supported in part by the National Science Foundation under Grant No. DMS-8701275 and by the United States Army under Contract No. DAAL03-87-K-0030

On multivariate polynomial interpolation

Carl de Boor & Amos Ron

1. Introduction

The generalization of univariate polynomial interpolation to the multivariate context is made difficult by the fact that one has to decide just which of the many of its nice properties to preserve, as it is impossible to preserve them all. Particularly annoying is the fact that the dimensions of standard multivariate polynomial spaces, such as Π_k, make up only a small subset of $\mathbb{Z}$, hence we cannot hope to interpolate uniquely at an arbitrary pointset $\mathbf{\Theta} \subset \mathbb{C}^s$ from an appropriate Π_k. Further, even when we have $\dim \Pi_k$ points at hand, they may fail to be total for Π_k, hence interpolation at these points from Π_k may still not be possible.

For these reasons, generalizations have stressed some aspects of polynomial interpolation and ignored others. For example, there are various efforts (see, e.g., [CY], [GM]) to identify finite sets $\mathbf{\Theta}$ for which it is easy to construct polynomial Lagrange functions, i.e., polynomials p_θ with $p_\theta(\tau) = \delta_{\theta,\tau}$. Except for special circumstances, it is usually hard to ascertain the **degree** of the resulting interpolant or the maximal k for which Π_k lies in the range of this interpolant. A totally different effort, associated with the name Kergin (see [K], [M]), retains the fact that, with an arbitrary set $\mathbf{\Theta}$ of cardinality $k+1$, we interpolate from Π_k. The additional degrees of freedom available in a multivariate context Kergin disposes of in such a way that 'natural' meanvalue theorems continue to hold.

In this paper, we take a different tack. Given any finite set $\mathbf{\Theta}$, we determine a corresponding polynomial space $\Pi_{\mathbf{\Theta}}$ from which interpolation at $\mathbf{\Theta}$ is 'correct', i.e., is possible and uniquely so. We show that $\Pi_{\mathbf{\Theta}}$ is translation- and scale-invariant, and that it is a polynomial space of least degree from which interpolation at $\mathbf{\Theta}$ is correct. We also show that the resulting map $\mathbf{\Theta} \mapsto \Pi_{\mathbf{\Theta}}$ is monotone (as a map from sets to sets), making it natural to introduce a Newton form for the resulting interpolant. Further, we show that the map can be extended in a natural way to Hermite interpolation, where we allow some of the θ to coalesce.

In fact, given arbitrary finite-dimensional polynomial spaces P_θ, we provide such a polynomial space of least degree from which "generalized Birkhoff-Hermite" interpolation is correct, i.e., over which the linear space spanned by the linear functionals of the form $[\theta]p(D)$, with $p \in P_\theta$ and $\theta \in \mathbf{\Theta}$, is minimally total. Here, $[\theta]f := f(\theta)$, and, to recall, a space Λ of linear functionals is **total** for H if the only $h \in H$ for which $\lambda h = 0$ for all $\lambda \in \Lambda$ is $h = 0$.

The following notation and terminology will be used throughout. The collection of all polynomials on $\mathbb{C}^s$ (or whatever other space the context might indicate) is denoted by Π; Π_k denotes the collection of all those polynomials of (total) degree $\le k$, i.e., $\Pi_k := \operatorname{span}\big(()^\alpha\big)_{|\alpha|\le k}$, with $()^\alpha : x \mapsto x^\alpha$. We also find it convenient to use $\Pi_{<k}$ for the space of polynomials of degree $< k$. For any $p \in \Pi$, we denote by $p(D)$ the corresponding constant coefficient differential operator; in particular, $D^\alpha := \prod_j (D_j)^{\alpha(j)}$, with D_j differentiation with respect to the jth argument. We make good use of the representation of the linear functional $[\theta]p(D)$ on Π as $q \mapsto p^* E^\theta q = q^*(e_\theta p)$, with

$$q^* p := (q(D)p)(0) = \sum_\alpha D^\alpha q(0) D^\alpha p(0)/\alpha!,$$

with $e_\theta : x \mapsto \exp(\langle \theta, x\rangle)$, and with E the *shift*, i.e, $E^\theta f := f(\cdot + \theta)$. We also use the "least term" of a function f analytic at the origin, i.e., the homogeneous polynomial $f_\downarrow$ of largest degree j for which $f(x) = f_\downarrow(x) + o(\|x\|^j)$ as $x \to 0$. This notion makes (formal) sense on the larger space of formal power series, as does the notion of a D-invariant subspace, i.e., a subspace invariant under differentiation.

The paper is laid out as follows. After a short discussion of basic properties of the "leading term" $f_\uparrow$ and the "least term" $f_\downarrow$ of f, we show in Section 3 that, for any finite-dimensional space H of sufficiently smooth functions, $H_\downarrow := \mathrm{span}\{f_\downarrow : f \in H\}$ is a scale-invariant polynomial space of the same dimension as H which determines the "local approximation order" from H. We also establish other properties of $H_\downarrow$, such as the fact that $p_\uparrow(D)H_\downarrow \subset (p(D)H)_\downarrow$, of use later, and explore $H_\downarrow$ for the space $H = \exp_{\boldsymbol\Theta}$ spanned by the exponentials $e_\theta, \theta \in \boldsymbol\Theta$. Section 4 is devoted to the important observation that, among all polynomial spaces P for which the corresponding space $\overline{P}^* := \{\overline{p}^* : p \in P\}$ of linear functionals is minimally total for H, $H_\downarrow$ is of least degree in the sense that, for all j, $\dim(P \cap \Pi_j) \le \dim(H_\downarrow \cap \Pi_j)$. The resulting linear projector given by H and $\overline{H}_\downarrow^*$ is exploited in Section 5 in the derivation of an algorithm for the construction of $H_\downarrow$ from any basis for H. The dependence of $H_\downarrow$ on H is explored in Section 6; the main result is that the map $H \mapsto H_\downarrow$ is continuous at H if and only if, for some m, $\Pi_{<m} \subseteq H_\downarrow \subset \Pi_m$, a property of H which we term "regular", for want of a better word.

The remainder of the paper is devoted to the specific choice $H = \sum_{\theta\in\boldsymbol\Theta} e_\theta P_\theta$ of exponentials. The fact that its least part $H_\downarrow$ supplies correct conditions for interpolation from H is used in Section 7 to conclude by duality, as in [DR], that it is possible to interpolate, and uniquely so, from the polynomial space $H_\downarrow$ using the interpolation conditions $[\theta]p(D), p \in P_\theta, \theta \in \boldsymbol\Theta$. The special case $P_\theta = \Pi_0$, all θ, leads in Section 8 to Lagrange interpolation from $\Pi_{\boldsymbol\Theta} := (\exp_{\boldsymbol\Theta})_\downarrow$, with the algorithm from Section 5 providing information needed for the Newton form for the interpolant. The connection between coalescence of such interpolation points and osculatory interpolation is explored in the final section.

In a subsequent paper, we verify that various forms of multivariate Lagrange interpolation now in the literature are special cases of the scheme proposed here. In a different paper, we use $H_\downarrow$ to simplify and extend results from box spline and exponential box spline theory.

2. The least term of an analytic function

We denote by $p_\uparrow$ the **leading term** of the polynomial p. For $p \neq 0$, this is the (unique) *homogeneous* polynomial for which

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

For completeness, we take the zero polynomial to be its own leading term. We note that $(pq)_\uparrow = p_\uparrow q_\uparrow$.

We also single out the **least term** $f_\downarrow$ (read 'f least') of a polynomial or, more generally, a function f analytic at the origin, and mean by this

$$f_\downarrow := T_j f,$$

with j the smallest integer for which $T_j f \neq 0$, with $T_j f$ the Taylor polynomial of degree $< j$ for f at the origin, i.e.,

$$T_j f := \sum_{|\alpha|<j} [\![\,]\!]^\alpha D^\alpha f(0), \tag{2.1}$$

and with $[\![\,]\!]^\alpha : x \mapsto x^\alpha/\alpha!$ the normalized power function. For completeness, we take the zero function to be its own least term. We note that

$$(fg)_\downarrow = f_\downarrow g_\downarrow \tag{2.2}$$

and that, for any invertible matrix A,

$$(f \circ A)_\downarrow = f_\downarrow \circ A. \tag{2.3}$$

We are interested in $f_\downarrow$ because it describes the behavior of f near the origin. Precisely, $f_\downarrow$ is the homogeneous polynomial of largest degree j for which

$$f(x) = f_\downarrow(x) + o(\|x\|^j) \quad \text{as } x \to 0. \tag{2.4}$$

Consequently, with $j := \deg f_\downarrow$,

$$f_\downarrow = \lim_{t\to 0} f(t\cdot)/t^j, \tag{2.5}$$

in the pointwise sense, as follows readily from L'Hôpital's rule. In fact, (2.5) remains true if we take the limit in the sense of formal power series, i.e., in the sense that, for all α,

$$[0]D^\alpha f_\downarrow = \lim_{t\to 0} [0](D^\alpha f)(t\cdot)/t^j.$$

3. The limit at the origin of a space of functions analytic there

In this section, we consider subspaces of the space

$$A_0$$

of all functions analytic at the origin, with the topology of formal power series. For any subspace H of A_0, we consider its "limit at the origin", i.e. (with (2.5)), the polynomial space

$$H_\downarrow := \operatorname{span}\{f_\downarrow : f \in H\}. \tag{3.1}$$

We note that $H_\downarrow$ *is scale-invariant* since it is spanned by homogeneous polynomials.

We were led to $H_\downarrow$ in the analysis of the **local approximation order** from H. By definition, this is the largest integer d for which, for every $f \in C^\infty(\mathbb{R}^s)$, there exists $h \in H$ so that

$$(f - h)(x) = O(\|x\|^d) \quad \text{as } x \to 0.$$

The following lemma is of use in the discussion of approximation order.

(3.2) Lemma. *If $\Pi_{<k} \subset H_\downarrow$, then there exists a continuous linear projector $T_{k,H}$ on A_0 into H for which $T_k T_{k,H} = T_k$.*

Proof: If $\Pi_{<k} \subset H_\downarrow$, then, for each $|\alpha| < k$, there exists $f_\alpha \in H$ with $(f_\alpha)_\downarrow = [\![]\!]^\alpha$. It follows that the matrix $(D^\alpha f_\beta(0))_{|\alpha|,|\beta|<k}$ is unit triangular, hence invertible. This implies that we can find $(g_\alpha)_{|\alpha|<k}$ in H dual to $\big(f \mapsto D^\alpha f(0)\big)_{|\alpha|<k}$ (i.e., satisfying $D^\alpha g_\beta(0) = \delta_{\alpha\beta}$), and this implies that, for each $f \in A_0$, $h := T_{k,H} f := \sum_\alpha g_\alpha D^\alpha f(0)$ is in H and satisfies $T_k h = T_k f$. $\square$

(3.3) Corollary. *The local approximation order of a finite-dimensional subspace H of A_0 equals the largest integer d for which $\Pi_{<d} \subset H_\downarrow$.*

Proof: Let d be the local approximation order from H.

Having $(f-h)(x) = O(\|x\|^d)$ as $x \to 0$ is the same as having $\deg(f-h)_\downarrow \geq d$. If, in particular, $f \in \Pi_{<d}$, then this can only happen if $h_\downarrow = f_\downarrow$. Since $\Pi_{<d} = (\Pi_{<d})_\downarrow$, this shows that $\Pi_{<d} \subset H_\downarrow$.

Conversely, if $\Pi_{<k} \subset H_\downarrow$, then, by the lemma, there is a linear map $T_{k,H}$ into H with $T_k(1 - T_{k,H}) = 0$. This implies that, for any $f \in A_0$, $h := T_{k,H}f$ is in H and satisfies $(f-h)(x) = O(\|x\|^k)$, hence $d \geq k$. □

Further study of $H_\downarrow$ led us to the results on polynomial interpolation to be detailed in subsequent sections. In preparation, we now discuss various properties of $H_\downarrow$.

Let H be a finite-dimensional subspace of A_0. We observe that, for $f \in H$, $\deg f_\downarrow = j$ if and only if $f \in (\ker_H T_j)\backslash(\ker_H T_{j+1})$, i.e., if and only if $f_\downarrow \in T_{j+1}(\ker_H T_j)\backslash 0$, with

$$\ker_H T_j := \ker(T_{j|H}). \tag{3.4}$$

Since also

$$H = \ker_H T_0 \supseteq \ker_H T_1 \supseteq \ker_H T_2 \supseteq \cdots,$$

hence

$$\dim T_{j+1}(\ker_H T_j) = \dim \ker_H T_j - \dim \ker_H T_{j+1},$$

we conclude that

$$H_\downarrow = \sum_{j=0}^{\infty} T_{j+1}(\ker_H T_j) = \bigoplus_{j=0}^{\infty} T_{j+1}(\ker_H T_j)$$

and

$$\dim(H_\downarrow \cap \Pi_k) = \sum_{j=0}^{k} \dim \ker_H T_j - \dim \ker_H T_{j+1} = \dim \ker_H T_0 - \dim \ker_H T_{k+1} = \dim T_{k+1}(H).$$

We have proved:

(3.5) Proposition. *$H_\downarrow$ is a scale-invariant space of polynomials of the same dimension as H. In fact, for every j,*

$$\dim(H_\downarrow \cap \Pi_{<j}) = \dim T_j(H). \tag{3.6}$$

Also, $(H_\downarrow)_\downarrow = H_\downarrow$, and $(T_jH)_\downarrow = H_\downarrow$ for all sufficiently large j, and $H = H_\downarrow$ in case H is a scale-invariant polynomial space.

For the particular space $H := \operatorname{span}\big(1 + ()^{1,0}, ()^{0,1}\big)$, one computes that $H_\downarrow = \operatorname{span}\big(1, ()^{0,1}\big) \neq H = T_2(H)$, thus illustrating that $T_j(H)$ and $H_\downarrow \cap \Pi_{<j}$ need not be equal in general (even though they are always of the same dimension).

Next, consider the effect of multiplying all the elements of H by some $f \in A_0$, i.e., the relationship between H and

$$fH := \{fg \mid g \in H\}.$$

We deduce from (2.2) the following observation.

(3.7) Proposition. *For any $f \in A_0$ satisfying $f(0) \neq 0$,*

$$(fH)_\downarrow = H_\downarrow.$$

The interaction of differentiation with the map $H \mapsto H_\downarrow$ is determined by the fact that, for any $p \in \Pi$ and any $f \in A_0$,

$$p(D)f = p_\uparrow(D)f_\downarrow + \text{ terms of higher degree.} \tag{3.8}$$

This implies that $(p(D)f)_\downarrow = p_\uparrow(D)f_\downarrow$ in case $p_\uparrow(D)f_\downarrow \neq 0$ and so proves the following.

(3.9) Proposition. *For every $p \in \Pi$,*

$$p_\uparrow(D)H_\downarrow \subset (p(D)H)_\downarrow. \tag{3.10}$$

(3.11) Corollary. *If $p(D)$ annihilates H, then $p_\uparrow(D)$ annihilates $H_\downarrow$.*

(3.12) Corollary. *If H is D-invariant, then so is $H_\downarrow$.*

Proof: For every $y \in \mathbb{R}^s$, we have $D_y(H_\downarrow) \subset (D_yH)_\downarrow$ by (3.9)Proposition, while $(D_yH)_\downarrow \subset H_\downarrow$ since $D_yH \subset H$ by assumption. □

If H consists of functions analytic on some domain G, then it makes sense to consider the "limit of H at z" for any $z \in G$. If H is D-invariant, hence translation-invariant, we expect all these limits to coincide. The Corollary confirms this expectation.

As an **example**, consider the space

$$\exp_{\boldsymbol{\Theta}} := \operatorname{span}\{e_\theta : \theta \in \boldsymbol{\Theta}\} \tag{3.13}$$

of simple exponentials with **frequencies $\boldsymbol{\Theta}$**. Here and below,

$$e_\theta : x \mapsto \exp(\langle \theta, x \rangle).$$

Since $\exp_{\boldsymbol{\Theta}}$ is D-invariant, its limit at any point is just $(\exp_{\boldsymbol{\Theta}})_\downarrow$. For its construction, we can proceed as follows. Define

$$p_{a,j} := \sum_{\theta \in \boldsymbol{\Theta}} (\theta\cdot)^j a(\theta),$$

with

$$(\theta\cdot)^j : x \mapsto \langle \theta, x \rangle^j / j!.$$

Then

$$\Pi_{\boldsymbol{\Theta}} := (\exp_{\boldsymbol{\Theta}})_\downarrow = \bigoplus_j \{p_{a,j} : p_{a,i} = 0 \text{ for } i < j\}. \tag{3.14}$$

The D-invariance of $\Pi_{\boldsymbol{\Theta}}$ can also be seen directly, as follows: For $y \in \mathbb{R}^s$, $D_y p_{a,j} = p_{ay,j-1}$, with $ay := \big(a(\theta)\langle \theta, y \rangle\big)_{\theta \in \boldsymbol{\Theta}}$. Hence if $p_{a,i} = 0$ for all $i < j$, then $p_{ay,i-1} = D_y p_{a,i} = 0$ for all $i - 1 < j - 1$, therefore $D_y p_{a,j} \in \Pi_{\boldsymbol{\Theta}}$ if $p_{a,j} \in \Pi_{\boldsymbol{\Theta}}$.

4. The dual of H

As it turns out, the construction of $H_\downarrow$ can be carried out by a bootstrap procedure which uses interpolation from H. For this reason (and others), we now show that the dual of H can be represented by $H_\downarrow$.

Abstractly, interpolation from H can be described as the task of finding, for given $f \in A_0$, an $h \in H$ for which $\lambda h = \lambda f$ for all λ in some linear space Λ of linear functionals on A_0. We call Λ the (space of) interpolation conditions for this particular interpolation problem. We call the problem **correct** if there is, for each f, exactly one solution h.

For completeness, we recall (without proof) the following well known characterizations of correctness.

(4.1) Lemma. *Let H and Λ be finite-dimensional linear subspaces of a linear space X (over $\mathbb{C}$) and its dual, respectively. Then the following are equivalent.*

(i) The interpolation problem given by H and Λ is correct.
(ii) With $(\lambda_j)_1^n$ any basis for Λ, the linear map $H \to \mathbb{C}^n : h \mapsto \big(\lambda_j h\big)_1^n$ is one-one and onto.
(iii) Λ is minimally total for H.
(iv) Λ can be used to represent the dual H' of H in the sense that the map $\Lambda \to H' : \lambda \mapsto \lambda_{|H}$ is one-one and onto.

If the interpolation problem given by H and Λ is correct, then it defines a linear projector, $P := P_{H,\Lambda}$ say, by the rule that, for any f, $Pf \in H$ and $\lambda Pf = \lambda f$ for all $\lambda \in \Lambda$. We will be interested later in the dependence of $P_{H,\Lambda}$ on H and Λ. For this, we remark that P can be written in the form

$$P = V(MV)^{-1}M,$$

with V any coordinate map for H (i.e., $V : \mathbb{C}^n \to H : a \mapsto \sum_j v_j a(j)$ for some basis $\big(v_j\big)$ for H), and M the dual of any coordinate map for Λ (i.e., $M : X \to \mathbb{C}^n : f \mapsto \big(\lambda_j f\big)$ for some basis $\big(\lambda_j\big)$ for Λ). This implies that $P_{H,\Lambda}$ is close to $P_{H',\Lambda'}$ in case H' and Λ' have bases close to bases of H and Λ respectively.

Concretely, we are interested in using linear functionals of the form

$$\overline{p}^* : f \mapsto \big(\overline{p}(D)f\big)(0) = \sum_\alpha \overline{D^\alpha p(0)} D^\alpha f(0)/\alpha! \tag{4.2}$$

with $p \in \Pi$. These are continuous linear functionals on A_0 and even on $C^k(0)$ (over the complex scalars). The map $p \mapsto \overline{p}^*$ is skew-linear and one-one, hence provides a skew-linear embedding of Π in the dual of A_0.

(4.3) Proposition. *For any finite-dimensional linear subspace H of A_0, the linear space $\overline{H}_\downarrow^*$ is minimally total for H.*

Proof: For any $f \in H\backslash 0$, $p := f_\downarrow \in H_\downarrow$ and $\overline{p}^* f = \overline{p}^* p > 0$. This implies that the only $f \in H$ with $p^* f = 0$ for all $p \in \overline{H}_\downarrow$ is $f = 0$, i.e., $\overline{H}_\downarrow^*$ is total for H. On the other hand, since $\dim \overline{H}_\downarrow^* = \dim H_\downarrow = \dim H$, no proper subspace of $\overline{H}_\downarrow^*$ could be total for H. □

We conclude that $\overline{H}_\downarrow^*$ can be used to represent the dual of H.

Of course, there exist polynomial spaces other than $H_\downarrow$ that provide correct interpolation conditions for H. The next result shows that, compared to all these spaces P, $H_\downarrow$ is **of least degree** in the sense that

$$\dim(H_\downarrow \cap \Pi_j) \geq \dim(P \cap \Pi_j) \qquad \text{all } j. \tag{4.4}$$

(4.5) Theorem. *Among all polynomial spaces P for which $\overline{P}^*$ is minimally total for H, $H_\downarrow$ is of least degree.*

Proof: Let B be a basis for $P \cap \Pi_j$. Since $\overline{P}^*$ is minimally total for H, the sequence $\overline{B}^*$ must be linearly independent over H. On the other hand, for any $p \in \Pi_j$, $\overline{p}^* = \overline{p}^* T_{j+1}$. Hence, $\overline{B}^*$ must already be linearly independent over $T_{j+1}(H)$. Therefore, with (3.6),

$$\dim(H_\downarrow \cap \Pi_j) = \dim T_{j+1}(H) \geq \#B = \dim(P \cap \Pi_j).$$

□

5. The construction of $H_\downarrow$

Denote by T_H the linear projector given by H and $\overline{H}_\downarrow^*$, i.e., the linear map characterized by the fact that its range is H and that $\overline{p}^* T_H f = \overline{p}^* f$ for all $p \in H_\downarrow$. The following strong monotonicity property will be of use in the construction of $H_\downarrow$ for specific choices of H.

(5.1) Proposition. *If $K = H + \operatorname{span}\{h\}$, then $K_\downarrow = H_\downarrow + \operatorname{span}\{(h - T_H h)_\downarrow\}$.*

Proof: There is nothing to prove in case $h \in H$, so assume that $h \notin H$, hence $k := h - T_H h \neq 0$, therefore also $k_\downarrow \neq 0$, while $k_\downarrow \in K_\downarrow$. Since $H_\downarrow \subset K_\downarrow$ and $\dim K_\downarrow = \dim K = \dim H + 1 = \dim H_\downarrow + 1$, we therefore need only show that $k_\downarrow \notin H_\downarrow$.

For this, observe that, by construction, $\overline{p}^* k = 0$ for all $p \in H_\downarrow$. Hence, from (3.8), $\overline{p}^* k_\downarrow = 0$ for all *homogeneous* $p \in H_\downarrow$. If now $k_\downarrow \in H_\downarrow$, then, in particular, $\overline{k}_\downarrow^* k_\downarrow = 0$, i.e., $k_\downarrow = 0$, a contradiction. □

This proposition suggests a simple Gram-Schmidt-like algorithm for the conversion of a basis $\big(p_j\big)$ for H into a basis $\big(q_j\big)$ for H to which $\big(r_j\big) := \big(q_{j\downarrow}\big)$ is bi-orthogonal with respect to the (complex) pairing

$$\langle , \rangle : \Pi \times A_0 \to \mathbb{C} : (p, q) \mapsto \overline{p}^* q,$$

hence provides the homogeneous orthogonal basis $\big(r_j\big)$ for $H_\downarrow$. The idea is simple. Assume that we have already determined such a basis $(q_j)_{j<k}$ for $H_k := \operatorname{span}\big(p_j\big)_{j<k}$. Then we compute

$$q_k := (1 - T_{H_k}) p_k = p_k - \sum_{j<k} q_j \frac{\langle r_j, p_k \rangle}{\langle r_j, q_j \rangle}$$

and note that $q_k \neq 0$ since $p_k \notin H_k = \operatorname{ran} T_{H_k}$. Consequently,

$$r_k := q_{k\downarrow}$$

is not zero, and $\langle r_k, q_k\rangle \neq 0$. By (5.1)Proposition, we know that $(H_{k+1})_\downarrow = \mathrm{span}\,\bigl(r_j\bigr)_{j\le k}$. Also, by construction,

$$\langle r_j, q_k\rangle = 0 \qquad j < k. \tag{5.2}$$

But there is, off-hand, no reason to expect that $\langle r_k, q_j\rangle = 0$ for $j < k$. Of course, if $\deg r_k < \deg r_j$, then we have $\langle r_k, q_j\rangle = 0$ trivially. If $\deg r_k = \deg r_j$, then $\langle r_k, q_j\rangle = \langle r_j, q_k\rangle = 0$. But for $\deg r_k > \deg r_j$, we may well have $\langle r_k, q_j\rangle \neq 0$. In that case, we simply modify q_j appropriately, setting

$$q_j := q_j - q_k \frac{\langle r_k, q_j\rangle}{\langle r_k, q_k\rangle} \qquad \text{if } \deg r_k > \deg r_j. \tag{5.3}$$

By (5.2), this does not change the bi-orthogonality of $(r_j)_{j<k}$ and $(q_j)_{j<k}$. Also, since it modifies q_j at terms of order $\deg r_k$ and higher, it does not change $q_{j\downarrow}$, i.e., *it does not change the fact that* $r_j = q_{j\downarrow}$ *for* $j < k$. In this way, we have now at hand a basis of the promised sort for H_{k+1}.

For easy reference, we collect the result of the last paragraph in the following.

(5.4) Algorithm. *Given the basis* $\bigl(p_j\bigr)$ *of the finite-dimensional subspace* H *of* A_0.

For $k = 1, 2, \ldots$, *carry out the following three steps:*

Step 1. $$q_k \leftarrow p_k - \sum_{j<k} q_j \frac{\langle r_j, p_k\rangle}{\langle r_j, q_j\rangle}$$

Step 2. $$r_k \leftarrow q_{k\downarrow}$$

Step 3. $$q_j \leftarrow q_j - q_k \frac{\langle r_k, q_j\rangle}{\langle r_k, q_k\rangle} \qquad \text{if } \deg r_k > \deg r_j.$$

Then $\bigl(r_j\bigr)$ *is bi-orthogonal to* $\bigl(q_j\bigr)$ *and provides a homogeneous orthogonal basis for* $H_\downarrow$.

For the calculations, it is useful to observe that only inner products with the homogeneous polynomials r_j are required. This means, in particular, that the calculation could be carried out with $T_m p_k$ rather than p_k, for some m which is determinable *a priori* in case H is D-invariant.

Numerically, the calculation is challenging only because it requires the determination of the least part r_k of q_k. When using finite-precision arithmetic, it may be necessary to replace 'least part' by 'significant least part' in order to avoid use of a least part that turned out not to be zero only because of the noise in the calculation. Concretely, this means that one takes r_k to be the homogeneous part of q_k of lowest degree which is not significantly smaller than the corresponding part of p_k. Considerations of this kind could be used to establish under what circumstances $H_\downarrow$ depends continuously on H. In the next section, we choose to settle this question by a more direct route.

6. Continuity of the map $H \mapsto H_\downarrow$

In the discussion later of Hermite interpolation as the limit of Lagrange interpolation, it will be important to understand to what an extent $H_\downarrow$ depends continuously on H. Precisely, we wish to know under what circumstances $\lim_{t\to 0}\bigl(H(t)_\downarrow\bigr) = \bigl(\lim_{t\to 0} H(t)\bigr)_\downarrow$. Since A_0 is a metric linear space, we use the gap between subspaces as a means of defining the statement $H = \lim_{t\to 0} H(t)$. For this, recall the standard definition of the **gap**

$$\mathrm{gap}(H, K) := \max\{\mathrm{dist}\,(H \cap B, K), \mathrm{dist}\,(K \cap B, H)\}$$

between subspaces H, K of the metric linear space X. Here, $\operatorname{dist}(Y,Z) := \sup_{y\in Y} \inf_{z\in Z} \operatorname{dist}(y,z)$, and $B := B_1(0) := \{x \in X : \operatorname{dist}(x,0) < 1\}$.

If H is, in particular, a finite-dimensional subspace, then $H = \lim H(t)$ iff, for some (every) basis (h_j) of H and all small enough t, there is a corresponding basis $(h_j(t))$ for $H(t)$ for which $h_j = \lim_{t\to 0} h_j(t)$ for all j. For example, by (2.5),

$$\lim_{t\to 0}\{h(t\cdot) : h \in H\} \;=\; H_\downarrow. \tag{6.1}$$

As an illustration of the possible lack of continuity in the map $H \mapsto H_\downarrow$, consider $H(t) = \exp_{\boldsymbol{\Theta}(t)}$ with $s = 2$ and $\boldsymbol{\Theta}(t) := \{(-1,0),(0,t),(1,0)\}$. Then $H(0) = \lim_{t\to 0} H(t)$ and, for $t \neq 0$, $H(t)_\downarrow = \Pi_1$, while $H(0)_\downarrow = \Pi_2(\mathrm{I\!R}) \circ ()^{1,0}$. We will show that this example of a discontinuity is prototypical.

For want of a better word, we call the linear subspace H of A_0 **regular** in case $H_\downarrow$ is of least degree, i.e., $\dim H_\downarrow \cap \Pi_m = \max\{\dim H_\downarrow, \dim \Pi_m\}$ for all m. Equivalently, H is regular iff $H_\downarrow \cap \Pi_m = H_\downarrow$ or Π_m, for all m. Thus, H is regular iff, for some m, $\Pi_{<m} \subseteq H_\downarrow \subset \Pi_m$. We note in passing that this m provides the local approximation order of H.

The importance of this notion of regularity for the continuity of the map $H \mapsto H_\downarrow$ is illustrated by the following.

(6.2) Lemma. *If each $H(t)$ is regular and $P = \lim H(t)_\downarrow$ exists, then P is regular.*

Proof: Since $H(t)_\downarrow$ converges, its dimension is eventually constant, and, since each is regular, this implies the existence of some m so that, for all small t, each $H(t)_\downarrow$ is scale-invariant and satisfies $\Pi_{<m} \subseteq H(t)_\downarrow \subset \Pi_m$, hence their limit P satisfies the same conditions. □

(6.3) Lemma. *The finite-dimensional subspace H is regular if and only if, for some regular, scale-invariant polynomial space P, $\overline{P}^*$ is minimally total for H.*

Proof: Since $(H_\downarrow)_\downarrow = H_\downarrow$, H is regular iff $H_\downarrow$ is regular, hence (4.3)Theorem proves the "only if".

For the converse, let P be a regular scale-invariant polynomial space. Then, with m such that $\dim \Pi_{<m} \le \dim P < \dim \Pi_m$, we must have $\Pi_{<m} \subseteq P_\downarrow = P \subset \Pi_m$. Assume that $\overline{P}^*$ is minimally total for H. Then H contains a dual set $(h_\alpha)_{|\alpha|<m}$ for $(\overline{p}^*_\alpha)_{|\alpha|<m}$, with $p_\alpha := ()^\alpha$. Consequently, for all $|\alpha|, |\beta| < m$, $D^\beta h_\alpha(0) = c_\alpha \delta_{\alpha\beta}$ for some $c_\alpha \neq 0$, hence $h_{\alpha\downarrow} = c_\alpha ()^\alpha$. This implies that $\Pi_{<m} \subset H_\downarrow$. To see that $H_\downarrow \subset \Pi_m$, observe that, since $\overline{P}^*$ is total for H, we can find, for any $h \in H\backslash 0$, some $p \in P$ so that $\overline{p}^* h \neq 0$, and, since $P \subset \Pi_m$, this implies that $h_\downarrow \in \Pi_m$. □

(6.4) Theorem. *The set of regular (finite-dimensional) subspaces of A_0 is open and dense (in the space of all finite-dimensional subspaces, and in the "gap topology").*

Proof: Assume that H is regular. We have to prove that all K in some neighborhood of H are also regular. But this follows from (6.3)Lemma and from the fact that if Λ is minimally total for H, then it is minimally total for all nearby K. This proves that the set of regular H is open.

To prove that the set of regular subspaces is dense, let H be an arbitrary finite-dimensional subspace and let P be an arbitrary regular scale-invariant polynomial space of the same dimension as H. Consider the map $R : P \to H' : p \mapsto \overline{p}^*{}_{|H}$. If R is 1-1, H itself is regular, by (6.3)Lemma. Otherwise, choose an orthonormal basis $(r_j)_1^n$ for P whose first m terms span $\ker R$. Then H contains a dual set $(h_j)_{m+1}^n$ for $(r_j)_{m+1}^n$, and this can be extended to a basis $(h_j)_1^n$ for H. Define $k_j := h_j + \varepsilon r_j$, all j. Then

$$\langle r_i, k_j\rangle = \begin{cases} \varepsilon\delta_{ij}, & i,j \le m;\\ 0, & i \le m < j;\\ (1+\varepsilon)\delta_{ij}, & i,j > m,\end{cases}$$

i.e., the matrix $\left(\langle r_i, k_j\rangle\right)_{i,j=1}^n$ is lower block triangular with diagonal blocks εI_m and $(1+\varepsilon)I_{m-n}$, hence invertible for all positive ε. Consequently, $\overline{P}^*$ is minimally total for $K := \operatorname{span}\left(h_j + \varepsilon r_j\right)_1^n$, for all positive ε. This shows, with (6.3)Lemma, that every neighborhood of H contains regular subspaces. □

Remark. The proof actually shows that, in the set of all n-dimensional subspaces of A_0, those for which a fixed regular polynomial space is minimally total form an open and dense set.

(6.5) Corollary. *The map $H \mapsto H_\downarrow$ is continuous at H if and only if H is regular.*

Proof: Assume that $H \mapsto H_\downarrow$ is continuous at H. By (6.4)Theorem, we can find regular $H(t)$ with $\lim_{t\to 0} H(t) = H$, hence, by the assumed continuity, $H_\downarrow = \lim_{t\to 0} H(t)_\downarrow$, thus H is regular by (6.2)Lemma.

For the proof of the converse, assume, more precisely, that $\Pi_{<k} \subset H_\downarrow \subset \Pi_k$. By (3.2)Lemma, this implies the existence of a continuous linear projector $T_{k,H}$ on A_0 onto H which preserves Taylor polynomials of degree $< k$, i.e., which satisfies $T_k T_{k,H} = T_k$. Since $T_{k,H}$ is a continuous linear projector, it carries any subspace K sufficiently close to H 1-1 onto H. Consequently, for every $f \in H$, we can find in every such K an element g close to f which satisfies $T_k g = T_k f$. In particular, $g_\downarrow = f_\downarrow$ in case $\deg f_\downarrow < k$, hence $\Pi_{<k} \subset K_\downarrow$ for any such K. Further, if $\deg f_\downarrow = k$, then $T_k g = 0$, i.e., $\deg g_\downarrow \ge k$, hence, since g is close to f, $\deg g_\downarrow = k$ and $g_\downarrow$ is close to $f_\downarrow$. This shows that, for such K and every $p \in H_\downarrow$, there is some $q \in K_\downarrow$ close to it, and, since $\dim K_\downarrow = \dim H_\downarrow$ for such K, it follows that $K_\downarrow$ and $H_\downarrow$ are also close. □

7. Hermite-Birkhoff interpolation

In this section, we consider interpolation by polynomials using interpolation conditions of the form $[\theta]p(D)$, with $[\theta]$ the linear functional of point evaluation at θ and p a polynomial. More precisely, we want to interpolate from some polynomial space Q, using the interpolation conditions

$$\Lambda := \Lambda(\boldsymbol{\Theta}; (P_\theta)) := \sum_\theta [\theta] P_\theta(D) := \sum_\theta \{[\theta]p(D) : p \in P_\theta\}. \tag{7.1}$$

For the analysis of this problem, observe that, in terms of (4.2),

$$[\theta]p(D)q = p^* E^\theta q = q^*(e_\theta p), \qquad \forall p, q \in \Pi. \tag{7.2}$$

This implies that our interpolation problem, as specified by Q and $\Lambda = \Lambda(\boldsymbol{\Theta}; (P_\theta))$, is correct if and only if the dual problem of interpolation from $H := \sum_{\theta\in\boldsymbol{\Theta}} e_\theta P_\theta$ with interpolation conditions Q^* is correct. Therefore, (4.5)Theorem provides the following conclusions.

(7.3) Theorem. *Given any finite set $\boldsymbol{\Theta} \subset \mathbb{C}^s$ and corresponding finite-dimensional polynomial spaces P_θ for $\theta \in \boldsymbol{\Theta}$, let $H := \sum_\theta e_\theta P_\theta$. Then $\overline{H_\downarrow}$ is a polynomial space of least degree among all polynomial spaces from which interpolation with interpolation conditions $\Lambda(\boldsymbol{\Theta}; (P_\theta)) := \operatorname{span}\{[\theta]p(D) : p \in P_\theta; \theta \in \boldsymbol{\Theta}\}$ is correct.*

In univariate Hermite-Birkhoff interpolation [LJR], one matches certain derivatives rather than linear combinations of derivatives. Correspondingly, we will use the term **Hermite-Birkhoff interpolation** in the multivariate context of (7.3)Theorem in case all the spaces P_θ have a homogeneous basis, i.e., are scale-invariant. If all the spaces P_θ are, in addition, D-invariant, then we speak of **Hermite interpolation**.

If each P_θ is D-invariant, then so is $H := \sum_\theta e_\theta P_\theta$, therefore, by (3.12), so is $H_\downarrow$. In the univariate case, it follows that $H_\downarrow = \Pi_k$ for some k, since Π_k is the only D-invariant polynomial space of dimension $k+1$. Thus we obtain the wellknown fact that univariate Hermite interpolation from Π_k with $k+1$ conditions is always correct.

8. Lagrange interpolation

The special case

$$\mathbf{\Theta} \subset \mathbb{R}^s, \qquad P_\theta = \Pi_0, \text{all } \theta$$

in (7.3)Theorem is particularly striking. The claim here is that, with

$$\exp_{\mathbf{\Theta}} := \operatorname{span}(e_\theta)_{\theta \in \mathbf{\Theta}},$$

the polynomial space

$$\Pi_{\mathbf{\Theta}} := (\exp_{\mathbf{\Theta}})_\downarrow$$

is of least degree among all those from which interpolation at $\mathbf{\Theta}$ is uniquely possible. (We are using the fact that, for this case, $H = \overline{H}$.) See (3.14) for a recipe for generating $\Pi_{\mathbf{\Theta}}$.

(8.1) Example. As a simple **illustration**, consider $s = 2$.

For $\#\mathbf{\Theta} = 1$, $\Pi_{\mathbf{\Theta}} = \Pi_0$.

For $\#\mathbf{\Theta} = 2$, $\Pi_{\mathbf{\Theta}} = \Pi_1(\mathbb{R}) \circ (\lambda\cdot)$, with λ any nonzero vector parallel to the affine hull of $\mathbf{\Theta}$.

For $\#\mathbf{\Theta} = 3$, $\Pi_{\mathbf{\Theta}} = \Pi_2(\mathbb{R}) \circ (\lambda\cdot)$, with λ any nonzero vector parallel to the affine hull of $\mathbf{\Theta}$, in case that hull is a line. Otherwise, $\Pi_{\mathbf{\Theta}} = \Pi_1$.

For $\#\mathbf{\Theta} = 4$, $\Pi_{\mathbf{\Theta}} = \Pi_3(\mathbb{R}) \circ (\lambda\cdot)$, with λ any nonzero vector parallel to the affine hull of $\mathbf{\Theta}$, in case that hull is a line. Otherwise, $\Pi_1 \subset \Pi_{\mathbf{\Theta}} \subset \Pi_2$. In that case, we can compute the barycentric coordinates of one point with respect to the other three, say $\theta_4 = \sum_1^3 a(j)\theta_j$ with $\sum_1^3 a(j) = 1$. On setting $a(4) = -1$, we thereby obtain the (essentially unique) quadratic homogeneous element $\sum_1^4 a(j)(\theta_j\cdot)^2$ of $\Pi_{\mathbf{\Theta}}$. Note that its span is a continuous function of $\mathbf{\Theta}$ except when the four points become collinear, since $\exp_{\mathbf{\Theta}}$ fails to be regular only in that case.

For the particular choice $\mathbf{\Theta} = \{0, \theta, \tau, \theta+\tau\}$ with $\theta = e_1, \tau = \alpha e_2$, the quadratic term becomes

$$((e_1\cdot)^2 + (\alpha e_2\cdot)^2 - ((e_1 + \alpha e_2)\cdot)^2) = -\alpha()^{1,1},$$

hence $\Pi_{\mathbf{\Theta}} = \Pi_{1,1}$, the space of bilinear polynomials. Correspondingly, the interpolation at $\mathbf{\Theta}$ from $\Pi_{\mathbf{\Theta}}$ is, in this case, the tensor product of linear interpolation, as one would hope. □

We now consider how $\Pi_{\mathbf{\Theta}}$ changes with $\mathbf{\Theta}$. Since $e_{\theta+a} = e_\theta e_a$, we conclude from (3.7)Proposition that

(8.2) $$\Pi_{\mathbf{\Theta}+a} = \Pi_{\mathbf{\Theta}}, \qquad \text{all } a \in \mathbb{R}.$$

Further, since the limit at 0 of any $H \subset A_0$ does not change under scaling,

(8.3) $$\Pi_{a\mathbf{\Theta}} = \Pi_{\mathbf{\Theta}}, \qquad \text{all } a \in \mathbb{R}\backslash 0.$$

Both of these facts could also be deduced directly from (3.14). More generally, if A is any matrix, then $e_{A\theta} = e_\theta \circ A^T$, hence, by (2.3),

$$\Pi_{A\mathbf{\Theta}} = \Pi_{\mathbf{\Theta}} \circ A^T$$

for any invertible matrix A.

More substantial changes in $\mathbf{\Theta}$ may change $\Pi_{\mathbf{\Theta}}$ substantially. In fact, the map $\mathbf{\Theta} \mapsto \Pi_{\mathbf{\Theta}}$ has jumps, as can be expected from (6.5)Corollary. The simplest possible example occurs with $s = 2$ and $\#\mathbf{\Theta} = 3$ (cf. (8.1)). Here $\Pi_{\mathbf{\Theta}} = \Pi_1$ except when $\mathbf{\Theta}$ is collinear, in which case $\Pi_{\mathbf{\Theta}} = \Pi_2(\mathbb{R}) \circ (\lambda \cdot)$ for any λ not zero and parallel to the affine hull of $\mathbf{\Theta}$. In terms of the Gram-Schmidt algorithm (5.4) for the construction of $\Pi_{\mathbf{\Theta}} = (\exp_{\mathbf{\Theta}})_{\downarrow}$, these jumps are due to the fact that, as $\mathbf{\Theta}$ changes, the degree of some $q_{k\downarrow}$ may jump even if we arrange the starting basis (p_j) for $\Pi_{\mathbf{\Theta}}$ in such a fashion that the degree of each $q_{k\downarrow}$ is as small as possible.

Connected with this is the fact that, near a $\mathbf{\Theta}$ at which the map jumps, our interpolation scheme is badly behaved. Put positively, in that case, it is a much stabler thing to use $\Pi_{\mathbf{\Theta}}$ when matching data at some $\mathbf{\Theta}'$ near such $\mathbf{\Theta}$, rather than using $\Pi_{\mathbf{\Theta}'}$. Thus in our simple example, it would be better to interpolate at $\mathbf{\Theta}$ from $\Pi_2(\mathbb{R}) \circ (\lambda \cdot)$ in case the points in $\mathbf{\Theta}$ are 'nearly collinear', rather than from $\Pi_{\mathbf{\Theta}}$ itself. In terms of the Gram-Schmidt algorithm (5.4), one would reject a 'near-zero' least term in favor of the next non-zero homogeneous term.

It seems natural to use $(e_\theta)_{\theta \in \mathbf{\Theta}}$ in the role of the initial basis $\big(p_j\big)$ in the Gram-Schmidt algorithm (5.4), making use of the fact that only the first few terms in the Taylor expansion of p_k are needed.

We can also use induction to construct the unique $I_{\mathbf{\Theta}} f \in \Pi_{\mathbf{\Theta}}$ which agrees with f at $\mathbf{\Theta}$, thereby obtaining its **Newton form**. For $\tau \notin \mathbf{\Theta}$, this gives

$$I_{\mathbf{\Theta} \cup \tau} f = I_{\mathbf{\Theta}} f + \frac{(1 - I_{\mathbf{\Theta}}) f(\tau)}{(1 - I_{\mathbf{\Theta}}) r_\tau(\tau)} (1 - I_{\mathbf{\Theta}}) r_\tau,$$

with $r_\tau := \big(e_\tau - T_{\mathbf{\Theta}} e_\tau\big)_{\downarrow}$ the least term of the error when interpolating e_τ from $\exp_{\mathbf{\Theta}}$ with respect to $(\Pi_{\mathbf{\Theta}})^*$. For, $r_\tau \in \Pi_{\mathbf{\Theta} \cup \tau} \backslash \Pi_{\mathbf{\Theta}}$ by (5.1)Proposition.

9. Osculation and coalescence

Our interpolation scheme $I_{\mathbf{\Theta}}$ depends on $\mathbf{\Theta}$ nicely enough to allow for the existence of a limit when some or all of the points in $\mathbf{\Theta}$ coalesce in a nice enough manner. In that case, the limiting situation often is Hermite interpolation, in the sense defined in Section 7.

Given that $I_{\mathbf{\Theta}}$ is characterized by $\Pi_{\mathbf{\Theta}}$ and $\exp_{\mathbf{\Theta}}$, it is natural, in light of the remarks following (4.1)Lemma, to study the limiting situation by considering the limits (if any) of these two spaces, as $\mathbf{\Theta}$ approaches some limiting set T.

In the univariate case, $\Pi_{\mathbf{\Theta}} = \Pi_{<\#\mathbf{\Theta}}$, hence it does not change (assuming that $\mathbf{\Theta}$ converges). Further, $\exp_{\mathbf{\Theta}}$ converges to

$$\exp_{\mathrm{T},\#\mathrm{T}} := \sum_{\tau \in \mathrm{T}} e_\tau \Pi_{<\#\tau}$$

with $\#\tau$ the **multiplicity** of τ, i.e., the number of points from $\mathbf{\Theta}$ which coalesce at τ. In particular, the limit always exists and does not at all depend on just how $\mathbf{\Theta}$ approaches T.

The multivariate situation is much more complicated. Neither $\Pi_{\mathbf{\Theta}}$ nor $\exp_{\mathbf{\Theta}}$ need to converge. If, for example, $s = 2$ and $\mathbf{\Theta} = \{0, \theta\}$ and θ alternates between the two axes as it approaches 0, then $\Pi_{\mathbf{\Theta}}$ alternates between the span of $()^0, ()^{1,0}$ and the span of $()^0, ()^{0,1}$, hence does not converge. Even if $\Pi_{\mathbf{\Theta}}$ and $\exp_{\mathbf{\Theta}}$ converge, their limits strongly depend on the manner in which $\mathbf{\Theta}$ approaches T. If, for example, $\theta = \theta' t + o(t)$ in the earlier example, then $\lim_{t \to 0} \Pi_{\mathbf{\Theta}} = \Pi_1(\mathbb{R}) \circ (\theta' \cdot) = \lim_{t \to 0} \exp_{\mathbf{\Theta}}$.

Here is a more striking example. We take again $s = 2$, but take $\mathbf{\Theta} = \mathbf{\Theta}(t) = \{0, (t, 0), (t^2, t^3)\}$. Then $\mathbf{\Theta}$ is in general position, hence $\Pi_{\mathbf{\Theta}} = \Pi_1$, therefore also $\lim_{t \to 0} \Pi_{\mathbf{\Theta}(t)} = \Pi_1$. Further, each $\exp_{\mathbf{\Theta}}$ contains $()^0$. In addition, $\exp_{\mathbf{\Theta}}$ contains $t^{-1}(e_{(t,0)} - 1) \xrightarrow[t \to 0]{} ()^{1,0}$. Finally, $\exp_{\mathbf{\Theta}}$ contains

$$t^{-3}\big(e_{(t^2,t^3)} - t e_{(t,0)} - (1 - t)\big) \xrightarrow[t \to 0]{} ()^{0,1} - ()^{2,0}/2.$$

Since $\exp_{\boldsymbol{\Theta}}$ is three-dimensional, this shows that

$$\lim_{t\to 0} \exp_{\boldsymbol{\Theta}} = \operatorname{span}\big(()^0, ()^{1,0}, ()^{0,1} - ()^{2,0}/2\big) \;\ne\; \Pi_1 = \lim \Pi_{\boldsymbol{\Theta}}.$$

In particular, $\lim \exp_{\boldsymbol{\Theta}}$ is not even scale-invariant, hence the resulting limiting interpolation scheme is not Hermite or Hermite-Birkhoff interpolation by our earlier definition. But $\lim \exp_{\boldsymbol{\Theta}}$ is D-invariant (as it has to be since each $\exp_{\boldsymbol{\Theta}}$ is D-invariant). After changing the point (t^2, t^3) in $\boldsymbol{\Theta}$ to (t^2, t^4), we still have $\Pi_{\boldsymbol{\Theta}} = \Pi_1$, but now

$$\lim_{t\to 0} \exp_{\boldsymbol{\Theta}} = \operatorname{span}\big(()^0, ()^{1,0}, ()^{2,0}/2\big) = (\lim \exp_{\boldsymbol{\Theta}})_\downarrow \;\ne\; \Pi_1 = \lim(\exp_{\boldsymbol{\Theta}})_\downarrow.$$

This shows that formation of the 'least' does not, in general, commute with limit formation, even if both limits exist and are scale-invariant.

We postpone a full discussion of the general situation to a future paper and content ourselves here with the following very simple case.

We assume that, more precisely, $\boldsymbol{\Theta} = \boldsymbol{\Theta}(t)$ consists of the points $\theta(t)$, with $\theta(t) = \theta(0) + \theta'(0)t$. Consider first the special situation that $\boldsymbol{\Theta}(0)$ consists of one point only. Then, in considering $\lim_{t\to 0} \exp_{\boldsymbol{\Theta}(t)}$, we may as well assume that $\boldsymbol{\Theta}(0) = \{0\}$ (see (8.2)), hence $\boldsymbol{\Theta}(t) = \Xi t$, with $\Xi := \{\theta'(0) : \theta \in \boldsymbol{\Theta}\}$. Consequently, from (6.1), $\Pi_\Xi = \big(\exp_\Xi\big)_\downarrow = \lim_{t\to 0} \exp_{\Xi t}$, while $\Pi_{\Xi t} = \Pi_\Xi$. This implies

$$\lim \Pi_{\boldsymbol{\Theta}(t)} = \Pi_\Xi = \lim \exp_{\boldsymbol{\Theta}(t)} \tag{9.1}$$

in this simple situation.

The same argument handles the slightly more general situation described in the following proposition.

(9.2) Proposition. *Assume that each $\theta \in \boldsymbol{\Theta} = \boldsymbol{\Theta}(t)$ is of the form $\theta_0 + \theta_1 t$. Then, for each $\tau \in \mathrm{T} := \{\theta_0 : \theta \in \boldsymbol{\Theta}\}$, the set $\Xi_\tau := \{\theta_1 : \theta_0 = \tau\}$ has as many elements as there are $\theta \in \boldsymbol{\Theta}$ with $\theta_0 = \tau$, and $\lim \exp_{\boldsymbol{\Theta}} = \sum_{\tau\in\mathrm{T}} e_\tau \Pi_{\Xi_\tau}$.*

Proof: The exponential space $\exp_{\boldsymbol{\Theta}}$ is the direct sum of the spaces $e_\tau \exp_{\Xi_\tau t}$ and these converge to $e_\tau \Pi_{\Xi_\tau}$, by the earlier argument. □

(9.3) Corollary. *If, in addition, $\lim \exp_{\boldsymbol{\Theta}}$ is regular, then $I_{\boldsymbol{\Theta}}$ converges (e.g., pointwise on Π) to Hermite interpolation from $P := (\lim \exp_{\boldsymbol{\Theta}})_\downarrow$ with interpolation conditions $\sum_{\tau\in\mathrm{T}} [\tau]\Pi_{\Xi_\tau}(D)$.*

Proof: From (6.5)Corollary, we conclude that the assumed regularity of $\lim \exp_{\boldsymbol{\Theta}}$ ensures the convergence of $\Pi_{\boldsymbol{\Theta}}$ to P, and the rest follows from the remarks following (4.1)Lemma. □

We have concluded from (6.1) that the exponential space $\exp_{t\boldsymbol{\Theta}}$ approaches the polynomial space $\Pi_{\boldsymbol{\Theta}}$ as $t \to 0$. This polynomial space is finite-dimensional, scale-invariant, and D-invariant. It would be very nice to know whether every finite-dimensional scale- and D-invariant polynomial space arises in this way. For it would then be possible to view all (regular) osculation as coalescence.

References

[CY] K. C. Chung and T. H. Yao, On Lattices Admitting Unique Lagrange Interpolations, *SIAM J. Numer. Anal.* **14** (1977), 735-741.

[DR] N. Dyn and A. Ron, Local approximation by certain spaces of exponential polynomials, approximation order of exponential box splines, and related interpolation problems, *Trans. Amer. Math. Soc.* **xx** (198x), xxx–xxx.

[GM] M. Gasca & J.I. Maeztu, On Lagrange and Hermite interpolation in $\mathbf{R}^k$, *Numerische Mathematik* **39** (1982), 361–374.

[K] P. Kergin, *Interpolation of C^k Functions*, Ph.D. Thesis, University of Toronto, Canada, (1978); published as 'A Natural Interpolation of C^k Functions', *J. Approximation Theory* **29** (1980), 278–293.

[LJR] G.G. Lorentz, K. Jetter and S.D. Riemenschneider, *Birkhoff Interpolation*, Encyclopedia of Mathematics and Its Applications vol.**19**, Addison Wesley, Reading, (1983).

[M] C. A. Micchelli, A Constructive Approach to Kergin Interpolation in $\mathbf{R}^k$: Multivariate B-Splines and Lagrange Interpolation, *Rocky Mountain J. Math.* **10** (1980), 485–497.