Technical Report
17 August 2001

Mean value interpolation for points in general position

Shayne Waldron

Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, New Zealand
e–mail: waldron@math.auckland.ac.nz (http:www.math.auckland.ac.nz/~waldron)

ABSTRACT

The reduced interpolation conditions for mean value interpolation at a multiset of points in $\mathbb{R}^d$ are given explicitly, i.e., the interpolation conditions are expressed using (tangential) derivatives of the lowest possible orders averaged over the smallest possible (polytope) regions. For points in general position the result takes a particularly simple form. Here we give the 'Lagrange' form of the extension of the mean value interpolant to less smooth functions, and give an application of it to the finite element method. For points which are not in general position the statement and proof of the result are more technical. These are closely related to an interesting binomial identity which counts the number of 'allowable' submultisets of a finite multiset of points in $\mathbb{R}^d$. For example, if the multiset consists of r points in general position each repeated m times, then the identity gives

$$\binom{mr-1+d}{d} = \sum_{j=1}^{d} \binom{r}{j} \sum_{k_1=1}^{m} \cdots \sum_{k_j=1}^{m} \binom{k_1+\cdots+k_j-1+d-j}{d-j},$$

which reduces to the Chu–Vandermonde convolution identity when $m = 1$. From our construction we are able to define multivariate analogues of Hermite–Fejér interpolation. We also outline the numerical calculation of the reduced functionals and the corresponding mean value interpolants.

Key Words: mean value interpolation, Kergin interpolation, Hakopian interpolation, multivariate interpolation, finite element method, binomial identity, Chu–Vandermonde identity, simplex spline, Kergin–Fejér interpolation

AMS (MOS) Subject Classifications: primary 05A19, 41A05, 65D07, secondary 41A10, 41A63,

1. Introduction

Kergin interpolation $\mathcal{K}_\Theta$ is a natural multivariate extension of Hermite interpolation, much studied since its introduction in [K80]. It is the first of the **mean value interpolation** operators

$$\mathcal{H}_\Theta^{(m)} : C^{n-m}(\mathbb{R}^d) \to \Pi_{n-m}(\mathbb{R}^d), \qquad 0 \le m \le d-1$$

which depend on the $n+1 > m$ (possibly coincident) 'interpolation' points $\Theta \subset \mathbb{R}^d$. The last of this family $\mathcal{H}_\Theta^{(d-1)}$ is known as the **Hakopian map** (see [H81]). They are most easily defined using the linear functional

$$f \mapsto \int_{[\theta_0,\ldots,\theta_k]} f := \frac{1}{k!\,\mathrm{vol}_k(S)} \int_S f \circ A, \tag{1.1}$$

where S is any k–simplex in $\mathbb{R}^s$ with (k–dimensional) volume $\mathrm{vol}_k(S)$, and $A : \mathbb{R}^s \to \mathbb{R}^d$ is any affine map taking the $k+1$ vertices of S onto the points $\theta_0, \ldots, \theta_k$. The change of variables formula shows that (1.1) does not depend on the particular choice of S and A, e.g., one can take

$$\int_{[\theta_0,\ldots,\theta_k]} f = \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{k-1}} f(\theta_0 + s_1(\theta_1 - \theta_0) + \cdots + s_k(\theta_k - \theta_{k-1}))\, ds_k \cdots ds_2\, ds_1.$$

Then the linear projector $\mathcal{H}_\Theta^{(m)}$ can be defined via its Newton form (see Goodman [G83])

$$\mathcal{H}_\Theta^{(m)} f(x) := m! \sum_{j=m+1}^{n+1} \sum_{\substack{\Psi \subset \Theta_{j-1} \\ \#\Psi = m}} \int_{\Theta_j} D_{x-\Theta_{j-1}\setminus\Psi} f. \tag{1.2}$$

Here Ψ is an m–subsequence of Θ_{j-1}, which is the first $j-1$ points of $\Theta = [\theta_0, \ldots, \theta_n]$, and $D_{x-\Theta_{j-1}\setminus\Psi} f$ is the $(j-1-m)$–th order derivative

$$D_{x-\Theta_{j-1}\setminus\Psi} f := \Big(\prod_{\substack{\theta \in \Theta_{j-1} \\ \theta \notin \Psi}} D_{x-\theta} \Big) f, \qquad D_y f := y_1 D_1 f + \cdots + y_d D_d f.$$

Desirable properties of mean value interpolation (cf [BHS93:Ch.12,13]) include:

Interpolation: The following mean values are matched

$$\int_\Psi \mathcal{H}_\Theta^{(m)} f = \int_\Psi f, \qquad \Psi \subset \Theta,\ \#\Psi = m+1.$$

These are just the function values $f(\theta_i)$ in the case of Kergin interpolation.

Symmetry in Θ: $\mathcal{H}_\Theta^{(m)}$ does not depend on the ordering of the points in Θ.

PDE Faithfulness: For *any* homogeneous constant coefficient partial differential operators $p(D)$ and $q(D)$ of degree $k \le n-m$,

$$p(D)f = q(D)g \quad \Longrightarrow \quad p(D)(\mathcal{H}_\Theta^{(m)} f) = q(D)(\mathcal{H}_\Theta^{(m)} g).$$

Continuity and Coalescence: For each $f \in C^{n-m}(\mathbb{R}^d)$, the map

$$\mathbb{R}^{d\times(n+1)} \to \Pi_{n-m}(\mathbb{R}^d) : \Theta \mapsto \mathcal{H}_\Theta^{(m)} f$$

is continuous. Moreover, if $\Theta = [a,\ldots,a]$, then substitution into (1.2) shows that $\mathcal{H}_\Theta^{(m)} f$ is the degree $n-m$ Taylor polynomial of f centred at a.

Because of the coalescence property, the mean value interpolant is defined in general only for C^{n-m} functions. However, for points in general position, we will show that $\mathcal{H}_\Theta^{(m)}$ can be extended to C^{d-1-m} functions. In Theorem 3.1 we give an explicit description of the interpolation conditions of this continuous extension. These 'reduced' functionals are decomposed into their dependence on derivatives of lowest possible orders (averaged over simplices with vertices from Θ which have the lowest possible dimension). Since these 'nodal variables' are of low smoothness, they invite use in finite elements. This application is considered in Section 4. In Section 5, we describe the extension of $\mathcal{H}_\Theta^{(m)}$ to functions with the least smoothness possible for an arbitrary multiset Θ. This more technical result is closely related to an interesting binomial identity which counts the number of 'allowable' submultisets of Θ. For example, if the multiset consists of r points in general position each repeated m times, then the identity gives

$$\binom{mr-1+d}{d} = \sum_{j=1}^{d} \binom{r}{j} \sum_{k_1=1}^{m} \cdots \sum_{k_j=1}^{m} \binom{k_1+\cdots+k_j-1+d-j}{d-j}, \tag{1.3}$$

which reduces to the Chu–Vandermonde convolution identity when $m=1$. From this construction we are able to define multivariate analogues of Hermite–Fejér interpolation. In Section 7, we outline the numerical calculation of the reduced functionals and the corresponding mean value interpolants.

The emphasis of this paper is on the explicit construction of the so called reduced functionals, and applications. Already, in [K80:Remark 4.8] it was known that Kergin interpolation $\mathcal{H}_\Theta^{(0)}$ had a continuous extension to C^{d-1} when the points Θ were in general position in $\mathbb{R}^d$. Not until [BC97] were the reduced functionals described in the bivariate case ($d=2$) and a formula for the continuous extension given. A complete discussion of the relationship of the results presented here with the existing literature is given at the end of Section 3.

2. Preliminary results

Here we give a version of the extension [M80:Th.6], which is in the spirit of the following manifestation of the fundamental theorem of calculus

$$\int_{[\Theta,v,w]} D_{v-w} f = \int_{[\Theta,v]} f - \int_{[\Theta,w]} f. \tag{2.1}$$

This is used to 'reduce' the interpolation conditions of the mean value interpolant.

As usual, let $\Pi_k(V)$ and $\Pi_k^0 := \Pi_k^0(V)$ denote the polynomials defined on a subspace $V \subset \mathbb{R}^d$ that are of degree $\le k$ and homogeneous of degree k, respectively. Let $\mathrm{aff}(\Psi)$ denote the affine hull of $\Psi \subset \mathbb{R}^d$, and $\mathrm{aff}(\Psi)^\perp$ the linear space of vectors in $\mathbb{R}^d$ orthogonal to it. Then the derivatives of order k which are (in directions) orthogonal to $\mathrm{aff}(\Psi)$ can be described by

$$\{q(D) : q \in \Pi_k^0(\mathrm{aff}(\Psi)^\perp)\} = \mathrm{span}\{D_{\mathrm{n}_1} D_{\mathrm{n}_2} \cdots D_{\mathrm{n}_k} : \mathrm{n}_1, \ldots, \mathrm{n}_k \perp \mathrm{aff}\,\Psi\},$$

where $q(D)$ is the differential operator induced by q. If Ψ has affine dimension s, then this space has dimension

$$\dim(\Pi_k^0(\mathrm{aff}(\Psi)^\perp)) = \dim(\Pi_k^0(\mathbb{R}^{d-s})) = \binom{k+d-s-1}{d-s-1}. \tag{2.2}$$

Lemma 2.3 (cf [M80:Th.6]). *Suppose Θ is a sequence of two or more points in $\mathbb{R}^d$. Then the linear functional*

$$\lambda : f \mapsto \int_\Theta h(D) f, \qquad h \in \Pi_k^0(\mathbb{R}^d), \quad k \ge 1$$

belongs to $A + B$ where

$$A := \mathrm{span}\{f \mapsto \int_\Psi p(D) f : \Psi \subset \Theta,\ \#\Psi = \#\Theta - 1,\ p \in \Pi_{k-1}^0(\mathbb{R}^d)\},$$
$$B := \mathrm{span}\{f \mapsto \int_\Theta q(D) f : q \in \Pi_k^0(\mathrm{aff}(\Theta)^\perp)\}.$$

Proof: It is sufficient to assume that $h(D) = D_{\xi_1} D_{\xi_2} \cdots D_{\xi_k}$ (a spanning set). Express ξ_1 as

$$\xi_1 = y + \mathrm{n}_1, \qquad y \parallel \mathrm{aff}(\Theta), \quad \mathrm{n}_1 \in \mathrm{aff}(\Theta)^\perp.$$

Since y is parallel to the affine hull of $\Theta := [\theta_1, \ldots, \theta_r]$, it can be expressed in the form

$$y = \textstyle\sum_{i=1}^r \mu_i \theta_i, \qquad \mu_i \in \mathbb{R}, \quad \sum_{i=1}^r \mu_i = 0.$$

Thus, by [M80:Th.6], we obtain that

$$\int_\Theta h(D) f = -\sum_{i=1}^r \mu_i \int_{\Theta \backslash \theta_i} D_{\xi_2} \cdots D_{\xi_k} f + \int_\Theta D_{\mathrm{n}_1} D_{\xi_2} \cdots D_{\xi_k} f. \tag{2.4}$$

Hence λ can be expressed as the sum of an element of A and the linear functional

$$f \mapsto \int_\Theta D_{\mathrm{n}_1} D_{\xi_2} \cdots D_{\xi_k} f.$$

Repeating this process (on the above functional) an additional $k-1$ times (for $\xi_2, \ldots, \xi_k$) shows that λ is the sum of an element of A and a linear functional of the form

$$f \mapsto \int_\Theta D_{\mathrm{n}_1} D_{\mathrm{n}_2} \cdots D_{\mathrm{n}_k} f, \qquad \mathrm{n}_1, \ldots, \mathrm{n}_k \perp \mathrm{aff}\,\Theta,$$

which is an element of B. $\square$

We will also need the following combinatorial result.

Lemma 2.5 (Chu–Vandermonde convolution). *For $0 \le m \le d-1$, $n \ge m$,*

$$\dim(\Pi_{n-m}(\mathbb{R}^d)) = \binom{n-m+d}{d} = \sum_{j=m+1}^{d} \binom{d-m-1}{d-j}\binom{n+1}{j}. \tag{2.6}$$

Proof: The right hand side counts the number of ways of choosing d objects from $n-m+d$ after having split them into two groups of sizes $d-m-1$ and $n+1$, and so equals the left hand side. □

3. The extension of $\mathcal{H}_\Theta^{(m)}$ to C^{d-1-m} for points in general position

A sequence of points is said to be in general position in $\mathbb{R}^d$ if every subsequence has the maximum affine dimension that its cardinality allows, e.g., points are in general position in $\mathbb{R}^2$ if no three of them are collinear.

We now show that if Θ is in general position in $\mathbb{R}^d$, then the mean value interpolation operators can be extended from C^{n-m} to C^{d-1-m} functions. The argument given below can be viewed as the multivariate analogue of taking the Newton form for Hermite interpolation at points in general position in $\mathbb{R}$ (no points repeated) and using the divided difference identity (cf (2.1) and Lemma 2.3) to obtain the Lagrange form.

A finite rank linear projector $P : X \to X$ is uniquely determined by its range V, and the range of its dual Λ, called its **interpolation conditions**. That is Pf is the unique element of V matching Λ, i.e.,

$$\lambda(Pf) = \lambda(f), \qquad \forall \lambda \in \Lambda,$$

where $\dim(V) = \dim(\Lambda)$. Denote by $\operatorname{conv}(\Theta)$ the convex hull of Θ.

Theorem 3.1 (Reduced functionals). *Suppose the $n+1$ points of Θ are in general position in $\mathbb{R}^d$. Then the mean value interpolation map has a continuous extension*

$$\mathcal{H}_\Theta^{(m)} : C^{d-1-m}(\operatorname{conv}(\Theta)) \to \Pi_{n-m}(\mathbb{R}^d),$$

which is uniquely determined by the interpolation conditions

$$\Lambda := \bigoplus_{j=m+1}^{d} \bigoplus_{\substack{\Psi\subset\Theta \\ \#\Psi=j}} \Lambda_\Psi^{(m)}, \tag{3.2}$$

where

$$\Lambda_\Psi^{(m)} := \operatorname{span}\{f \mapsto \int_\Psi q(D)f : q \in \Pi^0_{j-m-1}(\operatorname{aff}(\Psi)^\perp)\}. \tag{3.3}$$

Proof: First we show the interpolation conditions of $\mathcal{H}_\Theta^{(m)}$ are contained in Λ. By using (for simplicity) the basis $\{D^\alpha : \alpha \in \mathbb{Z}_+^d, |\alpha| \le n-m\}$ for the space of constant coefficient differential operators of order $\le n-m$, the Newton form (1.2) can be expanded

$$\mathcal{H}_\Theta^{(m)} f(x) = \sum_{j=m+1}^{n+1} \sum_{|\alpha|=j-1-m} p_\alpha(x)\lambda_\alpha(f), \qquad \lambda_\alpha : f \mapsto \int_{\Theta_j} D^\alpha f, \tag{3.4}$$

where $p_\alpha \in \Pi_{n-m}(\mathbb{R}^d)$. Since $\mathcal{H}_\Theta^{(m)}$ is a linear projector onto $\Pi_{n-m}(\mathbb{R}^d)$, and the number of terms in (3.4) is

$$\sum_{k=0}^{n-m} \#\{D^\alpha : |\alpha| = k\} = \dim(\Pi_{n-m}(\mathbb{R}^d)),$$

the polynomials (p_α) are a basis for $\Pi_{n-m}(\mathbb{R}^d)$ with dual basis (λ_α), and (λ_α) is a basis for the interpolation conditions. Since the j points of Θ_j are in general position,

$$\dim(\Pi_k^0(\operatorname{aff}(\Psi)^\perp)) = 0, \qquad \Psi \subset \Theta_j,\ \#\Psi \ge d+1,$$

and so $j-1-m$ applications of Lemma 2.3 to $\lambda_\alpha : f \mapsto \int_{\Theta_j} D^\alpha f$ gives

$$\lambda_\alpha \in \operatorname{span}\{f \mapsto \int_\Psi q(D)f : \Psi \subset \Theta_j,\ m+1 \le \#\Psi \le d,\ q \in \Pi^0_{\#\Psi-m-1}(\operatorname{aff}(\Psi)^\perp)\} \subset \Lambda.$$

Since $\lambda_\alpha \in \Lambda$ is continuous on $C^{d-1-m}(\operatorname{conv}(\Theta))$, the desired extension of $\mathcal{H}_\Theta^{(m)}$ is given by

$$\mathcal{H}_\Theta^{(m)} f := \sum_{j=m+1}^{n+1} \sum_{|\alpha|=j-1-m} p_\alpha \lambda_\alpha(f), \qquad \forall f \in C^{d-1-m}(\operatorname{conv}(\Theta)),$$

and it is unique, since (the restriction of) $C^{n-m}(\mathbb{R}^d)$ is dense in $C^{d-1-m}(\operatorname{conv}(\Theta))$.

It remains only to show that the sum in (3.2) is direct. By (2.2) and Lemma 2.5,

$$\dim(\Lambda) \le \sum_{j=m+1}^{d} \sum_{\substack{\Psi\subset\Theta \\ \#\Psi=j}} \dim(\Lambda_\Psi^{(m)}) \le \sum_{j=m+1}^{d} \binom{d-m-1}{d-j}\binom{n+1}{j} = \dim(\Pi_{n-m}(\mathbb{R}^d)).$$

Since $\dim(\Lambda) = \dim(\Pi_{n-m}(\mathbb{R}^d))$, there must be equality throughout above, and so by a dimension count the sum is direct. $\square$

The Lagrange form

For each Ψ, let $B_\Psi = B_\Psi^{(m)}$ be some basis for $\Pi^0_{\#\Psi-m-1}(\mathrm{aff}(\Psi)^\perp)$. Then the extension can be represented in the **Lagrange form**

$$\mathcal{H}_\Theta^{(m)} f(x) = \sum_{j=m+1}^{d} \sum_{\substack{\Psi\subset\Theta \\ \#\Psi=j}} \sum_{q\in B_\Psi} p_{\Psi,q}(x) \int_\Psi q(D) f, \qquad \forall f \in C^{d-1-m}(\mathrm{conv}(\Theta)), \tag{3.5}$$

where $(p_{\Psi,q})$ is the basis of $\Pi_{n-m}(\mathrm{I\!R}^d)$ which is dual to the basis of interpolation conditions

$$\{f \mapsto \int_\Psi q(D) f : \Psi \subset \Theta,\ m+1 \le \#\Psi \le d,\ q \in B_\Psi\}.$$

We refer to (3.5) as the Lagrange form, as opposed to the Newton form (1.2), since it involves derivatives of f of the lowest possible orders. In [BW01] it is shown that the 'Lagrange polynomials' $p_{\Psi,q}$ are multi–ridge functions of $d+m+1-\#\Psi$ variables.

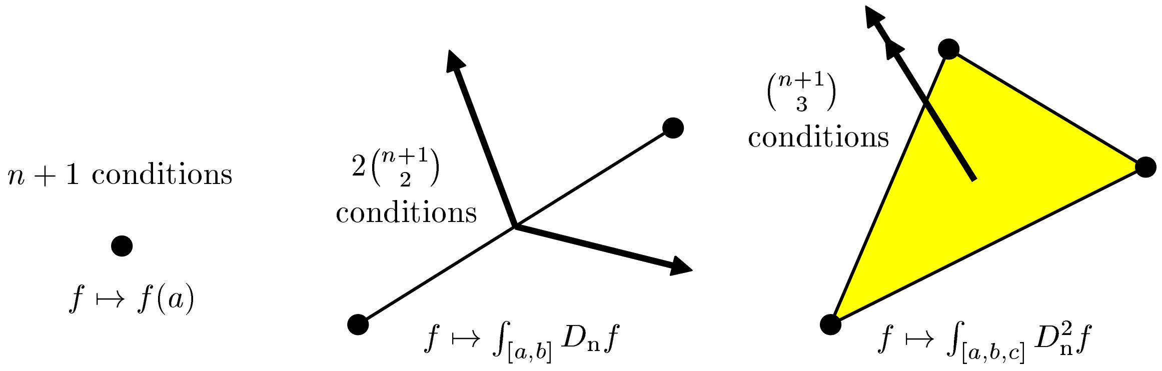

Fig. 3.1. The $\binom{n+3}{3}$ reduced interpolation conditions for the Lagrange form of Kergin interpolation at $n+1$ points in general position in $\mathrm{I\!R}^3$.

Examples

In $\mathrm{I\!R}$, points in general position are simply distinct points, and so $\mathcal{H}_\Theta^{(0)}$ is Lagrange interpolation at Θ, with (1.2) giving the classical Newton form, via the Hermite–Genocchi formula for divided differences

$$[\theta_0, \ldots, \theta_k] f = \int_{[\theta_0,\ldots,\theta_k]} D^k f. \tag{3.6}$$

In (5.6) all the Ψ consist of a single point, with $\Pi^0_{\#\Psi-0-1}(\mathrm{aff}(\Psi)^\perp) = \Pi_0(\mathrm{I\!R}^d)$, for which it is natural to choose the basis $B_\Psi = \{1\}$. Then (3.5) is the classical Lagrange form.

In $\mathrm{I\!R}^2$, the extension of Kergin interpolation $\mathcal{K}_\Theta := \mathcal{H}_\Theta^{(0)}$ is defined on C^1 functions, and the subsets Ψ in (3.2) consist of either one or two points. For one point subsets we have function evaluation, while for subsets $\Psi = \{\theta_s, \theta_t\}$, the average over the line segment from θ_s to θ_t of the derivative in the direction $(\theta_t-\theta_s)^\perp$ orthogonal to it is matched. By choosing the bases $B_\Psi=\{1\}$ and $B_\Psi(D)=\{D_{(\theta_t-\theta_s)^\perp}\}$ in each case respectively, the Lagrange form (3.5) for $\Theta=(\theta_0,\ldots,\theta_n)$ becomes

$$\mathcal{K}_\Theta f=\sum_{r=0}^n p_r f(\theta_r)+\sum_{0\le s\le t\le n} p_{st}\int_{[\theta_s,\theta_t]} D_{(\theta_t-\theta_s)^\perp} f, \qquad \forall f\in C^1(\operatorname{conv}\Theta). \tag{3.7}$$

This extension, together with some beautiful formulæ (involving the identification of $\mathbb{R}^2$ with $\mathbb{C}$) for the Lagrange polynomials $p_r, p_{st}\in\Pi_n(\mathbb{R}^2)$, was recently obtained by Bos and Calvi [BC97:Th.2.3].

In $\mathbb{R}^3$, the extension $\mathcal{K}_\Theta$ is defined for C^2 functions. In addition to reduced interpolation conditions like those in (3.7), the data $\int_{[a,b,c]} D^2_{\mathrm{n}} f$ is matched, where n is a normal to the plane through $a,b,c\in\Theta$ (see Fig. 3.1). For four points these 20 reduced functionals were derived in Lorentz [L92:Th.13.2.4] (also see the applications below).

For the Hakopian map $\mathcal{H}_\Theta^{(d-1)}$, only subsets Ψ with d elements are taken in (3.2), and so the original presentation (which required the points be in general position) is obtained.

In Cavaretta, Goodman, Micchelli and Sharma [CGMS83] the extension of Theorem 3.1 is given in the cases $m=d-2$ (Theorem 2) and $m=d-3$ (Theorem 3), together with some explicit formulæ for the corresponding Lagrange polynomials (Theorems 4 and 5).

4. An application to finite elements

The extended mean value interpolation maps of Theorem 3.1 give rise to a family of finite elements. First we recall some definitions (see, e.g., Brenner and Scott [BS94]).

Definition 4.1. *Let*
(i) $K\subset\mathbb{R}^d$ *be a domain with piecewise smooth boundary (the* **element domain***),*
(ii) P *be a finite-dimensional space of functions on* K *(the* **shape functions***), and*
(iii) $N=\{N_i\}$ *be a basis for* P' *(the* **nodal variables***).*
Then (K,P,N) *is called a* **finite element**.

It is assumed that the nodal variables $\{N_i\}$ are defined on sufficiently smooth functions. For a finite element (K,P,N), the basis for P dual to $\{N_i\}$ is called the **nodal basis** for P. It will be denoted by $\{\phi_i\}$.

Definition 4.2. *Given a finite element* (K,P,N)*, its* **local interpolant** *(to f) is*

$$\mathcal{I}_K f:=\mathcal{I}_{(K,P,N)}f:=\sum_i N_i(f)\,\phi_i. \tag{4.3}$$

Observe that the map $\mathcal{I}_K$ defined by (4.3) is the linear projector onto P with interpolation conditions N, i.e., for which

$$N_i(\mathcal{I}_K f)=N_i(f), \qquad \forall i,\ \forall f.$$

If $\Theta\subset\mathbb{R}^d$ is a sequence of $n+1$ points in general position, $n\ge d$, then $K:=\operatorname{conv}(\Theta)$ is an element domain, and (3.5) is the local interpolant of a finite element with shape functions $P=\Pi_{n-m}(\mathbb{R}^d)$, and nodal variables given by the interpolation conditions (3.2). We now state this result and give some examples.

Theorem 4.4 (Finite elements). *Suppose $\Theta \subset \mathbb{R}^d$ is a sequence of $n+1$ points in general position, $n \geq d > m \geq 0$. Let $K := \mathrm{conv}(\Theta)$, $P := \Pi_{n-m}(\mathbb{R}^d)$, and N be the nodal variables given by*

$$f \mapsto \int_\Psi q(D)f, \qquad \Psi \subset \Theta, \quad m+1 \leq \#\Psi \leq d, \quad q \in B_\Psi, \tag{4.5}$$

where B_Ψ is some basis for $\Pi^0_{\#\Psi-m-1}(\mathrm{aff}(\Psi)^\perp)$. Then (K,P,N) is a finite element with its local interpolant given by (3.5).

For the general theory of the error in such a local interpolant see [C78:Ch.3]. This effectively says that the reproduction of the polynomials $\Pi_{n-m}(\mathbb{R}^d)$ gives approximation order h^{m-n+1} where h is the diameter of the element domain K. Estimates of the constants involved here can be found using an error formula for the local interpolant, involving only the $(n-m+1)$–th derivative of f, which was given in [W97$_1$:Th.5.12], and applying to it a multivariate version of Hardy's inequality to get L_p–estimates for $p > d$ (see [W97$_2$:Th.4.3.1]).

Examples

The simplest bivariate example is $\Theta = (\theta_0, \theta_1, \theta_2)$ the vertices of a triangle K, and $m = 0$ (Kergin). Here $P = \Pi_2(\mathbb{R}^2)$, and the data matched by the nodal variables (4.5) is

$$f(\theta_0),\ f(\theta_1),\ f(\theta_2),\ \int_{\theta_0}^{\theta_1} D_{(\theta_1-\theta_0)^\perp} f,\ \int_{\theta_1}^{\theta_2} D_{(\theta_2-\theta_1)^\perp} f,\ \int_{\theta_0}^{\theta_2} D_{(\theta_2-\theta_0)^\perp} f, \tag{4.6}$$

with $\int_a^b$ denoting the line integral over the segment from a to b. This finite element is given in Lorentz [L92:Th.13.2.2,Cor.13.2.3]. For $\Theta = (\theta_0, \theta_1, \theta_2, \theta_3)$ the vertices of a convex quadrilateral K, $P = \Pi_3(\mathbb{R}^2)$, and the 10 nodal variables match the values

$$f(\theta_r), \quad 0 \leq r \leq 3, \qquad \int_{\theta_s}^{\theta_t} D_{(\theta_t-\theta_s)^\perp} f, \quad 0 \leq s < t \leq 3,$$

and so forth. See Bos and Calvi [BC97:Th.2.3] for a description of the local interpolant.

In $\mathbb{R}^3$, the first example is $\Theta = (\theta_0, \theta_1, \theta_2, \theta_3)$ the vertices of a tetrahedron K. Here $P = \Pi_3(\mathbb{R}^3)$, and the data matched by the nodal variables is the 4 function values $f(\theta_r)$, $0 \leq r \leq 3$, the 12 values of the form

$$\int_{\theta_s}^{\theta_t} D_{(\theta_t-\theta_s)^\perp} f, \quad 0 \leq s < t \leq 3,$$

where $D_{(\theta_t-\theta_s)^\perp} f$ now runs over some basis for the 2–dimensional space of derivatives in directions orthogonal to the line through θ_t and θ_s, and the 4 values

$$\int_{\mathrm{conv}(\theta_r,\theta_s,\theta_t)} D_{\mathrm{n}}^2 f, \quad 0 \leq r < s < t \leq 3,$$

where the integration above is over the triangle with vertices $\{\theta_r, \theta_s, \theta_t\}$, and n is some vector orthogonal to this triangle. This finite element is given in [L92:Th.13.2.4,Cor.13.2.5].

It might be hoped to obtain more practicable finite elements by replacing the averages occuring in the nodal variables by averages over sets of smaller affine dimension (ideally point evaluations). For example, **Morley's** finite element (see Ciarlet [C78:p.374]) can be obtained by replacing (4.6) by

$$f(\theta_0), f(\theta_1), f(\theta_2),\ D_{(\theta_1-\theta_0)^\perp} f\Big(\frac{\theta_0+\theta_1}{2}\Big),\ D_{(\theta_2-\theta_1)^\perp} f\Big(\frac{\theta_1+\theta_2}{2}\Big),\ D_{(\theta_2-\theta_0)^\perp} f\Big(\frac{\theta_0+\theta_2}{2}\Big).$$

The author is currently unaware of any general results in this direction.

5. The reduced functionals for $\mathcal{H}_\Theta^{(m)}$ where Θ is any multiset

Here we find the reduced functionals for $\mathcal{H}_\Theta^{(m)}$ where Θ is any sequence (multiset) of $n+1$ points in $\mathbb{R}^d$. By considering the basis (λ_α) of (3.4), it follows the interpolation conditions of $\mathcal{H}_\Theta^{(m)}$ are contained in

$$\sum_{\substack{\Phi\subset\Theta \\ \#\Phi\ge m+1}} L^{(m)}(\Phi), \qquad L^{(m)}(\Phi) := \operatorname{span}\{f \mapsto \int_\Phi q(D)f : q \in \Pi^0_{i-1-m}(\mathbb{R}^d), i = \#\Phi\}.$$

As before, our plan is to 'reduce' the spaces $L_\Theta^{(m)}(\Phi)$ as much as possible, so that we obtain a direct sum decomposition of reduced functionals. Since now the points need not be in general position this is more delicate, requiring the introduction of certain spaces of simplex splines to describe the averaging of the tangential derivatives. We are also unable to appeal to the Chu-Vandermonde convolution to show the sum we obtain is direct, though once proved we show that it implies a most interesting generalisation of the Chu-Vandermonde convolution identity. The result for Hakopian interpolation ($m = d-1$) has already been obtained in [H84], although there the directness of the sum is not proved.

The (affine) dimension of a submultiset $\Psi \subset \Theta$ is defined to be the dimension of the affine hull of the underlying set, i.e., $\dim(\Psi) := \dim(\operatorname{aff}(\Psi))$. For example, if Ψ consists of two distinct points (giving a line), each repeated a number of times, then its dimension is 1. The *multiplicity* of a point in multiset is its repetition number. Let $S \cap^* \Theta$ denote the intersection of an affine space $S \subset \mathbb{R}^d$ with a multiset $\Theta \subset \mathbb{R}^d$ *counting multiplicities.* For example, $\Theta \cap^* \mathbb{R}^d$ is the multiset Θ, whilst $\Theta \cap \mathbb{R}^d$ is the underlying set.

The **simplex spline** corresponding to a finite multiset (or sequence) $\Theta := [\theta_0, \ldots, \theta_k]$ of points in $\mathbb{R}^d$ (sometimes called *knots*) is defined to be the representer of the distribution

$$f \mapsto \int_\Theta f, \qquad \int_{[]} f := 0,$$

defined by (1.1). It is a nonnegative piecewise polynomial function $M(\cdot|\Theta)$ defined on $\mathrm{aff}(\Theta)$, supported on $\mathrm{conv}(\Theta)$, with the normalisation

$$\int_{\mathrm{aff}(\Theta)} M(\cdot|\Theta)f = \int_{\mathrm{conv}(\Theta)} M(\cdot|\Theta)f = k!\int_{\Theta} f, \tag{5.1}$$

which ensures it integrates to 1. Properties of these splines have been studied in great detail (see, e.g., [BHS93]). We are interested only in a particular spline space $\mathcal{S}_i(\Psi)$, which occurs in our construction of the reduced functionals. This space of functions on $\mathrm{aff}(\Psi)$ is defined as

$$\mathcal{S}_i(\Psi) := \mathrm{span}\{M(\cdot|\Phi) : \Phi \in \mathcal{A}_i(\Psi)\}, \qquad i = 1, \ldots, \#\Psi,$$

where

$$\mathcal{A}_i(\Psi) := \{\Phi \subset \Psi : \#\Phi = i,\ \dim(\Phi) = \dim(\Psi)\}.$$

Note $\mathcal{S}_i(\Psi) = 0$ if $i + 1 < \dim(\Psi)$. When $\mathcal{A}_i(\Psi) = \{\Phi \subset \Psi : \#\Phi = i\}$, as is the case for Ψ in general position in $\mathrm{aff}(\Psi)$, Dahmen and Micchelli [DM83:Th.1] have shown that

$$\dim(\mathcal{S}_i(\Psi)) = \binom{\#\Psi - i + \dim(\Psi)}{\dim(\Psi)}.$$

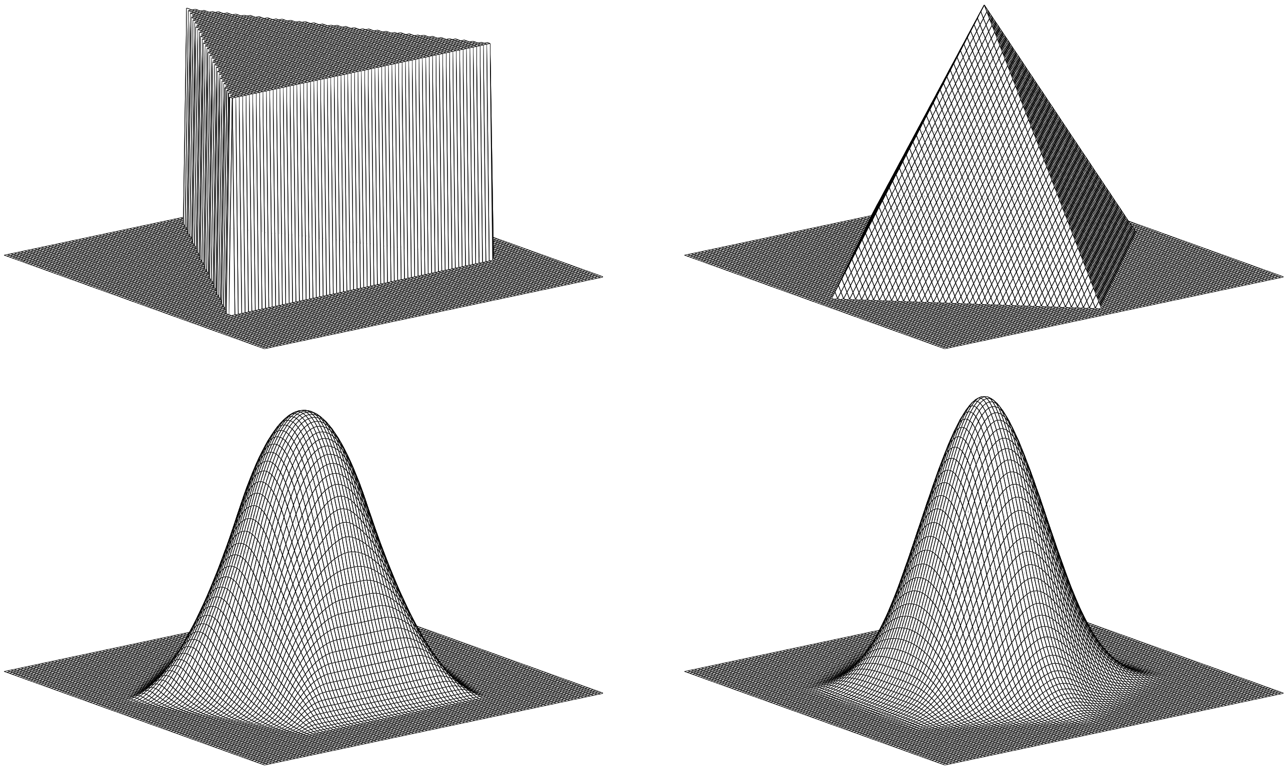

Fig. 5.1. Simplex splines for points equally spaced on a circle in $\mathbb{R}^2$: 3 knots (discontinous piecewise constant function), 4 knots (continuous piecewise linear function), 5 knots (C^1 piecewise quadratic function), and 6 knots (C^2 piecewise cubic function).

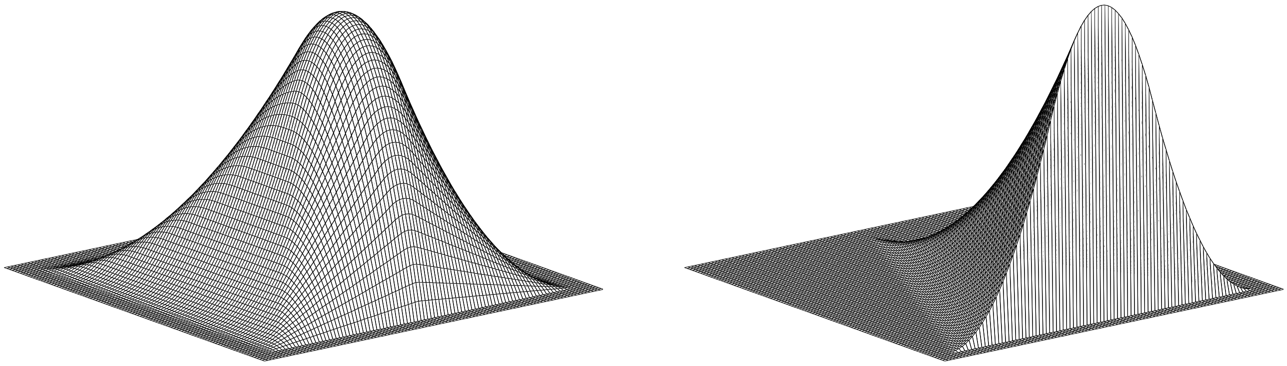

Fig. 5.2. Simplex splines for 5 points in $\mathbb{R}^2$ which are not in general position (piecewise quadratics). In the first 3 knots are collinear, and the spline is only C^0 across this line. In the second 4 knots are collinear, and the spline is discontinuous across this line.

Lemma 5.2. *Let Ψ consist of j affinely independent points with multiplicities $m_1,\ldots,m_j$. Then for $i=j,\ldots,\#\Psi$*

$$\dim(\mathcal{S}_i(\Psi)) = \#\mathcal{A}_i(\Psi) = c_i(m_1,\ldots,m_j), \tag{5.3}$$

where

$$c_i(m_1,\ldots,m_j) := \textit{coefficient of } t^i \textit{ in } \prod_{k=1}^{j}(t+t^2+\cdots+t^{m_k}). \tag{5.4}$$

Proof: The submultisets in $\mathcal{A}_i(\Psi)$ consist of those Φ of size i which contain each of the j points at least once. There is not a general formula for $c_i(m_1,\ldots,m_j)$, the number of these, but it can be calculated from the generating function (5.4), or by using the principle of inclusion and exclusion.

Let $\xi_1,\ldots,\xi_j$ be the barycentric coordinates of a point in $\operatorname{aff}(\Psi)$ relative to the j points, and $k_1,\ldots,k_j$ be the multiplicities of these points as they occur in $\Phi\in\mathcal{A}_i(\Psi)$. Then the restriction of $M(\cdot|\Phi)$ to the simplex $T:=\operatorname{conv}(\Psi)$ is a scalar multiple of

$$\xi_1^{k_1}\xi_2^{k_2}\cdots\xi_j^{k_j} \quad (k_1+\cdots+k_j=i).$$

It is well known that these polynomials of degree i which appear in the definition of the Bernstein operator (of degree i for the simplex T) are linearly independent. Hence the $M(\cdot|\Phi)$ are linearly independent, and so $\dim(\mathcal{S}_i(\Psi)) = \#\mathcal{A}_i(\Psi)$. □

Theorem 5.5 (Reduced functionals). *Let Θ be any sequence (multiset) of $n+1$ points in $\mathbb{R}^d$. Then $\Lambda_\Theta^{(m)}$ the interpolation conditions for $\mathcal{H}_\Theta^{(m)}$ can be written as the direct sum*

$$\Lambda_\Theta^{(m)} = \bigoplus_{j=1}^{d} \bigoplus_{\substack{\Psi\subset\Theta,\#\Psi\ge m+1\\ \dim(\Psi)=j-1\\ \operatorname{aff}(\Psi)\cap^*\Theta=\Psi}} \bigoplus_{i=m+1}^{\#\Psi} \Lambda_{\Psi,i}^{(m)}, \tag{5.6}$$

where

$$\Lambda^{(m)}_{\Psi,i} := \mathrm{span}\{f \mapsto \int_\Phi q(D)f : q \in \Pi^0_{i-1-m}(\mathrm{aff}(\Psi)^\perp),\ \Phi \in \mathcal{A}_i(\Psi)\}, \tag{5.7}$$

and

$$\dim(\Lambda^{(m)}_{\Psi,i}) = \binom{i-1-m+d-j}{d-j} \dim(\mathcal{S}_i(\Psi)), \qquad j-1 := \dim(\Psi). \tag{5.8}$$

Proof: Suppose $\lambda \in L^{(m)}(\Phi)$, $\Phi \subset \Theta$, $\#\Phi = i \geq m+1$, then, by Lemma 2.3,

$$\lambda \in \Lambda^{(m)}_{\Psi,i} + \sum_{\substack{\Xi\subset\Phi \\ \#\Xi=i-1}} L^{(m)}(\Xi), \qquad \Psi := \mathrm{aff}(\Phi) \cap^* \Theta. \tag{5.9}$$

Note $\#\Psi \geq i$ and $\Lambda^{(m)}_{\Psi,i} = 0$ if $\dim(\Psi) = d$. Since $\Lambda^{(m)}_\Theta \subset \sum_{\substack{\Phi\subset\Theta \\ \#\Phi\geq m+1}} L^{(m)}(\Phi)$, repeated application of (5.9) gives

$$\Lambda^{(m)}_\Theta \subset \sum_{j=1}^{d} \sum_{\substack{\Psi\subset\Theta,\#\Psi\geq m+1 \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=m+1}^{\#\Psi} \Lambda^{(m)}_{\Psi,i}. \tag{5.10}$$

The reverse inclusion follows from (5.7), and so it suffices to show the sum in (5.10) is direct, i.e.,

$$\lambda \in \Lambda^{(m)}_{\Psi_1,i_1} \cap \Lambda^{(m)}_{\Psi_2,i_2}, \quad (\Psi_1,i_1) \neq (\Psi_2,i_2) \quad \Longrightarrow \quad \lambda = 0.$$

If $\mathrm{aff}(\Psi_1) \cap \mathrm{aff}(\Psi_2) = \phi$, then by considering the support of the distribution λ, we have

$$\mathrm{supp}(\lambda) \subset \mathrm{aff}(\Psi_1) \cap \mathrm{aff}(\Psi_2) = \phi \quad \Longrightarrow \quad \lambda = 0.$$

Hence, it remains to consider the case $\mathrm{aff}(\Psi_1) \cap \mathrm{aff}(\Psi_2) \neq \phi$. For this, our proof relies on the identification of $\Lambda^{(m)}_{\Psi,i}$ with a space of functions $F^{(m)}_{\Psi,i}$. Since $\mathbb{R}^d = \mathrm{aff}(\Psi)^\perp \oplus \mathrm{aff}(\Psi)$, the tensor product

$$\begin{aligned} F^{(m)}_{\Psi,i} &:= \Pi^0_{i-1-m}(\mathrm{aff}(\Psi)^\perp) \otimes \mathcal{S}_i(\Psi) \\ &:= \mathrm{span}\{f = p \otimes s : p \in \Pi^0_{i-1-m}(\mathrm{aff}(\Psi)^\perp),\ s \in \mathcal{S}_i(\Psi)\}, \end{aligned} \tag{5.11}$$

with

$$f(x+y) := (p \otimes s)(x,y) = p(x)s(y), \qquad \forall x \in \mathrm{aff}(\Psi)^\perp,\ \forall y \in \mathrm{aff}(\Psi),$$

is a space of functions on $\mathbb{R}^d$. Note, by (5.1), that

$$\Lambda^{(m)}_{\Psi,i} = \mathrm{span}\{f \mapsto \int_{\mathrm{aff}(\Psi)} (p(D)f)s : p \in \Pi^0_{i-1-m}(\mathrm{aff}(\Psi)^\perp),\ s \in \mathcal{S}_i(\Psi)\}. \tag{5.12}$$

Let $P^{(m)}_{\Psi,i} : \Lambda^{(m)}_{\Psi,i} \to F^{(m)}_{\Psi,i}$ be the linear map $\lambda \mapsto p \otimes s$, where $\lambda(f) = \int_{\mathrm{aff}(\Psi)}(p(D)f)s$. This map is clearly onto, and it is 1–1 since

$$p \otimes s = 0,\ s \neq 0 \quad \Longrightarrow \quad p(x)s(y) = 0 \quad \Longrightarrow \quad p(x) = 0,\ \forall x \in \mathrm{aff}(\Psi)^\perp \quad \Longrightarrow \quad p = 0.$$

Hence the linear spaces $\Lambda^{(m)}_{\Psi,i}$ and $F^{(m)}_{\Psi,i}$ are isomorphic, and in particular,

$$\begin{aligned}\dim(\Lambda^{(m)}_{\Psi,i}) &= \dim(F^{(m)}_{\Psi,i}) = \dim(\Pi^0_{i-1}(\mathrm{aff}(\Psi)^\perp))\dim(\mathcal{S}_i(\Psi)) \\ &= \binom{i-1-m+d-j}{d-j}\dim(\mathcal{S}_i(\Psi)),\end{aligned}$$

which is (5.8). In view these isomorphisms, it is sufficient to show that the sum

$$F^{(m)}_{\Psi_1,i_1} + F^{(m)}_{\Psi_2,i_2}, \qquad (\Psi_1,i_1)\neq(\Psi_2,i_2), \quad \mathrm{aff}(\Psi_1)\cap\mathrm{aff}(\Psi_2)\neq\phi, \tag{5.13}$$

is direct, i.e., that

$$f\in F^{(m)}_{\Psi_1,i_1}\cap F^{(m)}_{\Psi_2,i_2}, \quad \Longrightarrow \quad f=0.$$

This we do in two cases. Note, by (5.11), that

$$\mathrm{supp}(f) = \mathrm{supp}(f|_{\mathrm{conv}(\Psi)}) + \mathrm{aff}(\Psi)^\perp, \qquad \forall f\in F^{(m)}_{\Psi,i}, \tag{5.14}$$

and if L is a line in $\mathbb{R}^d$

$$L \text{ is orthogonal to } \mathrm{aff}(\Psi) \quad \Longrightarrow \quad f|_L\in\Pi^0_{i-1-m}(L). \tag{5.15}$$

Case 1: $\mathrm{aff}(\Psi_1)^\perp = \mathrm{aff}(\Psi_2)^\perp$, i.e., $\mathrm{aff}(\Psi_1)$ and $\mathrm{aff}(\Psi_2)$ are parallel. The condition $\mathrm{aff}(\Psi_1)\cap\mathrm{aff}(\Psi_2)\neq\phi$ implies $\mathrm{aff}(\Psi_1)=\mathrm{aff}(\Psi_2)$, $i_1\neq i_2$. By (5.15), the restriction of f to each line orthogonal to $\mathrm{aff}(\Psi_1)$ is a homogeneous polynomial of degree i_1-1-m and i_2-1-m, i.e., the zero polynomial ($i_1\neq i_2$). Since $\mathrm{aff}(\Psi_1)\neq\mathbb{R}^d$ ($\dim(\Psi_1)=j-1<d$), this implies $f=0$.

Case 2: $\mathrm{aff}(\Psi_1)^\perp\neq\mathrm{aff}(\Psi_2)^\perp$. If B is a ball in $\mathbb{R}^d$ containing Ψ_1 and Ψ_2, then

$$\mathrm{supp}(f|_{\mathrm{conv}(\Psi_i)})\subset\mathrm{supp}(f)\cap B =: K, \qquad i=1,2.$$

Since $\mathrm{aff}(\Psi_i)^\perp$ is a linear subspace, using (5.14), this implies

$$\mathrm{supp}(f)=\mathrm{supp}(f|_{\mathrm{conv}(\Psi_i)})+\mathrm{aff}(\Psi_i)^\perp\subset K+\mathrm{aff}(\Psi_i)^\perp\subset\mathrm{supp}(f)+\mathrm{aff}(\Psi_i)^\perp=\mathrm{supp}(f),$$

giving

$$K+\mathrm{aff}(\Psi_1)^\perp = K+\mathrm{aff}(\Psi_2)^\perp. \tag{5.16}$$

If $f\neq0$, then $K\neq\phi$ is compact, and (5.16) implies

$$\mathrm{aff}(\Psi_1)^\perp=\mathrm{aff}(\Psi_2)^\perp,$$

which contradicts our assumption.

This completes the proof. □

Example 1 (Points in general position). For Θ in general position, $\mathrm{aff}(\Psi)\cap^*\Theta=\Psi$, $\dim(\Psi)=\#\Psi+1$, $\mathcal{A}_i(\Psi)=0$ if $i<\#\Psi$ and $\mathcal{A}_i(\Psi)=\{\Psi\}$ if $i=\#\Psi$, so (5.6) becomes

$$\Lambda_\Theta^{(m)}=\bigoplus_{j=m+1}^{d}\bigoplus_{\substack{\Psi\subset\Theta\\ \#\Psi=j}}\Lambda_{\Psi,j}^{(m)}, \tag{5.17}$$

where $\Lambda_{\Psi,j}^{(m)}=\Lambda_\Psi^{(m)}$ of (3.3), and we recover Theorem 3.1.

Example 2 (All points coincident). If all points are the same, i.e., $\Theta=[a,\ldots,a]$, only $\Psi=\Theta$ satisfies $\mathrm{aff}(\Psi)\cap^*\Theta=\Psi$. Since $\mathrm{aff}(\Psi)^\perp=\mathrm{IR}^d$, and $\int_\Phi f$, $\Phi\subset\Theta$, $\#\Phi\ge 1$ is a multiple of $f(a)$, the direct sum (5.6) becomes

$$\bigoplus_{i=m+1}^{n+1}\Lambda_{\Theta,i}^{(m)},\qquad \Lambda_{\Theta,i}^{(m)}=\mathrm{span}\{f\mapsto (q(D)f)(a):q\in\Pi_{i-1-m}^0(\mathrm{IR}^d)\},$$

i.e., the interpolation conditions for Taylor interpolation of degree $n-m$ at a.

Example 3 (Hermite interpolation). Let Θ consist of the points $x_1,\ldots,x_r$ in IR repeated with multiplicities $m_1,\ldots,m_r$. Then the Ψ in (5.6) consist of a point x_k repeated m_k times, giving

$$\Lambda_\Theta^{(0)}=\bigoplus_{k=1}^{r}\bigoplus_{i=1}^{m_k}\Lambda_{\underbrace{[x_k,\ldots,x_k]}_{m_k},i}=\bigoplus_{k=1}^{r}\bigoplus_{i=1}^{m_k}\mathrm{span}\{f\mapsto D^{i-1}f(x_k)\}, \tag{5.18}$$

which is the usual presentation of the interpolation conditions for Hermite interpolation. By Hermite–Genocchi (3.6), the interpolation conditions (λ_α) in (3.4) are equivalent to

$$\Lambda_\Theta^{(0)}=\mathrm{span}\{f\mapsto[\theta_1,\ldots,\theta_r]f:[\theta_1,\ldots,\theta_r]\subset\Theta,\ r=1,\ldots,n\},$$

where $[\theta_1,\ldots,\theta_r]f$ is the divided difference of f at the sequence of points $\theta_1,\ldots,\theta_r$.

Example 4 (Points on a circle in IR^d). Let Θ consist of $n+1$ points in general position in some plane in IR^d, $d\ge 2$, e.g., points on some circle. Then the Ψ in (5.6) are those of size 1, 2 and $n+1$. For the Ψ of sizes $j=1,2$, the corresponding (reduced) interpolation conditions are

$$\left.f(\theta_r),\qquad \binom{n+1}{1}\text{ conditions,}\right\}\ j=1$$

$$\left.\int_{[\theta_r,\theta_t]}D_{(\theta_t-\theta_r)^\perp}f,\qquad \binom{n+1}{2}(d-1)\text{ conditions,}\right\}\ j=2$$

where $\theta_r,\theta_t\in\Theta$, $\theta_r\ne\theta_t$. The only other Ψ occuring is $\Psi=\Theta$, for $j=3$. When Θ consists of $n+1$ points in general position in an s–dimensional subspace of IR^d, it is possible to give an explicit subset of $\mathcal{A}_i(\Theta)$ whose corresponding splines are a basis of $\mathcal{S}_i(\Theta)$. This *s–generating sequence* of [DM83:Th.4.2] is defined as follows

$$\mathcal{B}_i(\Theta) := \{[\theta_1, \ldots, \theta_{i-s}] \cup \Phi : \Phi \subset [\theta_{i-s+1}, \ldots, \theta_{n+1}],\ \#\Phi = s\},$$

where $\Theta = [\theta_1, \ldots, \theta_{i-s}, \theta_{i-s+1}, \ldots, \theta_{n+1}]$. In particular, for our points on a circle ($s = 2$) the remaining interpolation conditions (for $\Psi = \Theta$) can be expressed as: for $i = 3, \ldots, n+1$

$$\left.\int_{[\theta_1,\ldots,\theta_{i-2},\theta_r,\theta_t]} D_\perp^{i-1} f, \qquad \binom{i+d-4}{d-3}\binom{n+3-i}{2} \text{ conditions}\right\} j = 3,$$

where $i - 2 < r < t \leq n + 1$.

Example 5 (Hakopian interpolation). For Hakopian interpolation ($m = d - 1$) with $\mathrm{aff}(\Theta) = d$ the result was obtained in [H84]. In the notation used there our sum $\oplus_{j=1}^d$ is described in two parts. For $\oplus_{j=2}^d$ the set of Ψ ($\dim \Psi = j-1$) in the second sum corresponds to Hakopian's set $\mathcal{L}_p$, $1 \leq p = j - 1 \leq d - 1$ of all p–dimensional hyperplanes spanned by points from Θ, and for a fixed $L_p^q \in \mathcal{L}_p$ the $B^p_{\alpha, L_p^q \cap \{x^0,\ldots,x^r\}}$ are bases of simplex splines for our $S_i(\Psi)$, $i = d, \ldots, \#\psi$. For $\oplus_{j=1}$ only the points a with multiplicity $r := \#\Psi \geq d$, $\Psi := \{a\} \cap^* \Theta$ contribute interpolation conditions, namely

$$\Lambda_{\Psi,i}^{(d-1)} = \mathrm{span}\{f \mapsto (q(D)f)(a) : q \in \Pi_{i-d}(\mathbb{R}^d)\}, \qquad i = d, \ldots, r$$

which are clearly the same as those given by Hakopian. Hakopian states than an inductive argument shows the number of interpolation conditions equals $\dim(\Pi_{n-d+1}(\mathbb{R}^d)) = \binom{n+1}{d}$, and so the sum is direct, but it is not included.

Example 6 (Extensions of $\mathcal{H}_\Theta^{(m)}$). The linear functionals in (5.6) are called *reduced interpolation conditions* because the derivatives in them are tangential to the polytopes they are averaged over, and so the order of these derivatives cannot be decreased any further by application of Lemma 2.3. As in Theorem 3.1 the mean value interpolation operator $\mathcal{H}_\Theta^{(m)}$ can be extended continuously to any functions these reduced functionals are defined on. In particular, there is an extension to $C^r(\mathbb{R}^d) \to \Pi_{n-m}(\mathbb{R}^d)$, where

$$r := \max_{\substack{\Psi \subset \Theta, \#\Psi \geq m+1 \\ \dim(\Psi) < d \\ \mathrm{aff}(\Psi) \cap^* \Theta = \Psi}} (\#\Psi - 1 - m).$$

6. Chu–Vandermonde convolution identities

Since the sum (5.6) is direct we obtain the following dimension count.

Corollary 6.1 (Chu–Vandermonde). *For any multiset* Θ *of* $n+1$ *points in* $\mathbb{R}^d$ *and* $0 \le m \le d-1$,

$$\binom{n-m+d}{d} = \sum_{j=1}^{d} \sum_{\substack{\Psi\subset\Theta,\#\Psi\ge m+1\\ \dim(\Psi)=j-1\\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=m+1}^{\#\Psi} \binom{i-1-m+d-j}{d-j} \dim(\mathcal{S}_i(\Psi)), \tag{6.2}$$

where

$$1 \le \dim(\mathcal{S}_i(\Psi)) \le \binom{\#\Psi - i + j - 1}{j-1}. \tag{6.3}$$

The upper bound in (6.3) *is attained when* $\mathcal{A}_i(\Psi) = \{\Phi \subset \Psi : \#\Phi = i\}$*. This happens if the underlying point set of* Ψ *is in general position in the* $(j-1)$*–dimensional affine space that it spans, and the number of repeated points* $\Psi \setminus (\mathrm{aff}(\Psi) \cap \Psi)$ *is* $\le i-j$*.*

Proof: Since (3.2) is direct, $\dim(\Lambda_\Theta^{(m)}) = \dim(\Pi_{n-m}(\mathbb{R}^d)) = \binom{n-m+d}{d}$ equals the sum of the dimensions of the $\Lambda_{\Psi,i}^{(m)}$ which are given by (5.8). The estimate (6.3) follows from the proof of Lemma 4.1 in Dahmen and Micchelli [DM83], where it is shown

$$\dim(\mathcal{S}_i(\Theta)) = \dim\operatorname{span}\{x \mapsto \prod_{\theta\in\Phi}(1+\theta^* x) : \Phi \subset \Theta,\ \#\Phi = n-i\}$$
$$\le \dim(\Pi_{n-i}(\mathbb{R}^s)) = \binom{n-i+s}{s},$$

together with the asserted equality. Furthermore, if the stated condition holds, then every $\Phi \subset \Psi$ of size i contains j affinely independent points, and so belongs to $\mathcal{A}_i(\Psi)$, i.e., $\mathcal{A}_i(\Psi) = \{\Phi \subset \Psi : \#\Phi = i\}$. □

When the points Θ are in general position we obtain from (5.17) the Chu-Vandermonde convolution (2.6). It is possible to interpret (6.2) as a way of counting the 'allowable' submultisets of a finite multset of points Θ in $\mathbb{R}^d$. We will illustrate this for $m = 0$.

Definition 6.4. *A submultiset* $\Phi \subset \Theta$ *is said to be* **allowable** *if*
(a) $\Theta \cap \mathrm{conv}(\Phi) \subset \Phi$*, i.e., all points of* Θ *in* $\mathrm{conv}(\Phi)$ *are in* Φ*.*
(b) *points of* Θ *in the relative interior of* $\mathrm{conv}(\Phi)$ *are repeated in* Φ *with full multiplicity.*
For short we will write $\Phi \subset^* \Theta$*.*

Theorem 6.5 (Allowable submultisets). *Suppose* Θ *is a multiset of* $n+1$ *points in* $\mathbb{R}^d$*, and either*
(a) *The points* $\Theta \cap \mathbb{R}^d$ *are in general position in an* s*–dimensional affine subspace of* $\mathbb{R}^d$*, where* $s = 0$*,* $s = 1$ *or* $s = d$*.*
(b) $\mathbb{R}^d = \mathbb{R}$ *or* $\mathbb{R}^d = \mathbb{R}^2$*.*
then

$$\binom{n+d}{d} = \sum_{i=1}^{n+1} \sum_{\substack{\Phi\subset^*\Theta\\ \#\Phi=i}} \binom{i+d-\dim(\Phi)-2}{d-\dim(\Phi)-1}, \tag{6.6}$$

where the sum in (6.6) *is over all allowable submultisets (of size i).*

Proof: Every allowable $\Phi \subset \Theta$ which contributes to the sum (6.6) belongs to a unique Ψ with $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$ and $\dim(\Psi) = j-1 < d$, namely $\Psi := \mathrm{aff}(\Phi) \cap^* \Theta$. Hence

$$\sum_{i=1}^{n+1} \sum_{\substack{\Phi \subset^* \Theta \\ \#\Phi = i}} \binom{i+d-\dim(\Phi)-2}{d-\dim(\Phi)-1}$$
$$= \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \binom{i-1+d-j}{d-j} \#\{\Phi \in \mathcal{A}_i(\Psi) : \Phi \text{ is allowable}\}.$$

Thus, (6.2) with $m = 0$ implies (6.6), provided

$$\#\{\Phi \in \mathcal{A}_i(\Psi) : \Phi \text{ is allowable}\} = \dim(\mathcal{S}_i(\Psi)), \qquad i = j, \ldots, \#\Psi. \tag{6.7}$$

We now show (6.7) holds in the cases (a) and (b).

Case (a): Suppose (a) holds. Then the possible values of $\dim(\Psi)$ are $0, 1, \ldots, s$. Those Ψ with $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$, and $\dim(\Psi)+1 = j = 1, 2, \ldots, s$, consist of j affinely independent points repeated with full multiplicity. Fix $i = j, \ldots, \#\Psi$. Then each $\Phi \in \mathcal{A}_i(\Psi)$ is allowable, and so Lemma 5.2 gives

$$\{\Phi \in \mathcal{A}_i(\Psi) : \Phi \text{ is allowable}\} = \#\mathcal{A}_i(\Psi) = \dim(\mathcal{S}_i(\Psi)),$$

which is (6.7). If $\dim(\Psi) = s$, then there is exactly one Ψ with $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$, namely $\Psi = \Theta$, and so (6.7) is only required to hold when $s < d$. For $s = 1$, by taking a B–spline basis (see, e.g., [BHS93]) for Θ ordered $\theta_1 \le \theta_2 \le \cdots \le \theta_{n+1}$ (as they appear on the line) it follows

$$\begin{aligned}\#\{\Phi \in \mathcal{A}_i(\Theta) : \Phi \text{ is allowable}\} &= \#\{(\theta_j, \ldots, \theta_{j+i-1}) : \theta_j < \theta_{j+i-1},\ j = 1, \ldots, n+2-i\} \\ &= \dim(\mathcal{S}_i(\Theta)), \qquad i = 2, \ldots, n+1,\end{aligned}$$

as required. For $s = 0$, Θ is a single point repeated, and (6.7) holds because the one submultiset of each size i is allowable.

Case (b): The case of $\mathbb{R}$ follows from (a), take $d = 1$, and $s = 0, 1$. For $\mathbb{R}^2$, it is necessary to show that (6.7) holds for $j - 1 = 0, 1$, which again follows from (a). □

Here are some examples of the binomial identities given by Theorem 6.5.

Example 1 (Points in general position). If the points are in general position, then every submultiset of size i up to d is allowable, with $\dim(\Phi) = i - 1$. Those of size $d+1$ or larger have $\dim(\Phi) = d$, and so contribute zero to the sum. Hence

$$\binom{n+d}{d} = \sum_{i=1}^{d} \binom{n+1}{i} \binom{i-1+(d-i+1)-1}{(d-i+1)-1} = \sum_{i=1}^{d} \binom{n+1}{i} \binom{d-1}{d-i}, \tag{6.8}$$

which is the Chu–Vandermonde identity.

Example 2 (All points coincident). If all the points coincide, then there is exactly one allowable submultiset of size i, giving the binomial identity

$$\binom{n+d}{d} = \sum_{i=1}^{n+1} \binom{i+d-2}{d-1}. \tag{6.9}$$

Example 3 (All points on a line in $\mathbb{R}^d$, with no repetitions). If there are $n+1$ points in $\mathbb{R}^d$ all distinct and lying on a line, then there are exactly $n+2-i$ allowable subsets of size i (consecutive points), which leads to

$$\binom{n+d}{d} = (n+1) + \sum_{i=2}^{n+1} (n+2-i) \binom{i+d-3}{d-2}, \tag{6.10}$$

and

$$\binom{n+d}{d} = \sum_{i=1}^{n+1} (n+2-i) \binom{i+d-3}{d-2}, \qquad d \geq 2. \tag{6.11}$$

Example 4 (All points on a line in $\mathbb{R}^d$, with repetitions). Suppose that the points in Θ lie on a line in $\mathbb{R}^d$, say in the order $x_1, \ldots, x_r$, occurring with multiplicities $m_1, \ldots, m_r$, where $m_1 + \cdots + m_r = n+1$. Then the allowable submultisets are either one point repeated a number of times, or start with a point x_i (repeated $1, \ldots, m_i$ times) contain all points $x_{i+1}, x_{i+2}, \ldots, x_{j-1}$ with full multiplicity and the point x_j (repeated $1, \ldots, m_j$ times). Using (6.9) to count the submultisets of one point repeated, we obtain the identity

$$\begin{aligned} \binom{n+d}{d} = & \sum_{i=1}^{r} \binom{m_i - 1 + d}{d} \\ & + \sum_{1 \leq i < j \leq r} \sum_{k_1=1}^{m_i} \sum_{k_2=1}^{m_j} \binom{k_1 + m_{i+1} + m_{i+2} + \cdots + m_{j-1} + k_2 - 1 + d - 2}{d-2}. \end{aligned} \tag{6.12}$$

When all the points occur with multiplicity 1 (Example 3), this reduces to

$$\binom{n+d}{d} = (n+1) + \sum_{1 \leq i < j \leq n+1} \binom{j-i+d-2}{d-2}. \tag{6.13}$$

For points in $\mathbb{R}^2$, we get

$$\binom{n+2}{2} = \binom{m_1 + \cdots + m_r + 1}{2} = \sum_{i=1}^{r} \binom{m_i+1}{2} + \sum_{1 \leq i < j \leq r} m_i m_j, \tag{6.14}$$

and for two points occuring with multiplicities m and $n+1-m$ respectively, we get

$$\binom{n+d}{d} = \binom{m-1+d}{d} + \binom{n-m+d}{d} + \sum_{i=1}^{m} \sum_{j=1}^{n+1-m} \binom{i+j-1+d-2}{d-2}. \tag{6.15}$$

Example 5 (Points in general position in $\mathbb{R}^d$, each repeated m times). Suppose that Θ consists of r points in general position in $\mathbb{R}^d$, each repeated m times. Then the allowable submultisets which contribute to the sum consist of up to d points, each repeated with some multiplicity ($\leq m$). We count these as follows. The $i_1 = 1, \ldots, d$ distinct points can be chosen in $\binom{r}{i_1}$ ways. Of these, i_2 of them occur with multiplicity at least two, this selection can be done in $\binom{i_1}{i_2}$ ways, and i_3 of these i_2 points occur with multiplicity three or more, and so forth, giving

$$\begin{aligned}&\binom{mr-1+d}{d} \\ &= \sum_{i_1=1}^{d} \binom{r}{i_1} \sum_{i_2=0}^{i_1} \binom{i_1}{i_2} \sum_{i_3=0}^{i_2} \binom{i_2}{i_3} \cdots \sum_{i_m=0}^{i_{m-1}} \binom{i_{m-1}}{i_m} \binom{d+i_2+i_3+\ldots+i_m-1}{d-i_1}.\end{aligned} \tag{6.16}$$

Example 6 (Points in general position in $\mathbb{R}^d$, with repetitions). Suppose that Θ consists of r points in general position in $\mathbb{R}^d$, repeated with multiplicities $m_1, \ldots, m_r$, where $m_1 + \cdots + m_r = n+1$. Then the allowable submultisets which contribute to the sum consist of up to d points, each repeated with some multiplicity. This gives the identity

$$\binom{n+d}{d} = \sum_{j=1}^{d} \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq r} \sum_{k_1=1}^{m_{i_1}} \cdots \sum_{k_j=1}^{m_{i_j}} \binom{k_1+\cdots+k_j-1+d-j}{d-j}. \tag{6.17}$$

In the case of the previous example ($m_1, \ldots, m_r = m$), this gives

$$\binom{mr-1+d}{d} = \sum_{j=1}^{d} \binom{r}{j} \sum_{k_1=1}^{m} \cdots \sum_{k_j=1}^{m} \binom{k_1+\cdots+k_j-1+d-j}{d-j}, \tag{6.18}$$

which is (1.3). Compare this with (6.16), which was obtained by counting the allowable submultisets in a different way.

Example 7 (Points in general position in an affine subspace of $\mathbb{R}^d$). Suppose that Θ consists of $n+1 > s$ (distinct) points which are in general position in an s–dimensional affine subspace of $\mathbb{R}^d$. Then the submultisets Ψ which contribute to the summation are all those of size j up to s, and $\Psi = \Theta$. This gives

$$\begin{aligned}\binom{n+d}{d} = \sum_{j=1}^{s} \binom{n+1}{j} \binom{d-1}{d-j} + \sum_{i=s+1}^{n+1} \binom{i+d-s-2}{d-s-1} \binom{n+1-i+s}{s}, \\ n+1 > s,\ 1 \leq s \leq d.\end{aligned} \tag{6.19}$$

In the special cases Example 1 and 3 ($s = d$ and $s = 1$), this reduces to (6.8) and (6.10).

Example 8 (Different allowable submultisets). For $d \geq 3$, the first example where (6.6) does not hold is 4 distinct points in $\mathbb{R}^3$ which lie in a plane, of which no three are collinear. In this case $\binom{n+d}{d} = 20$. All submultisets of sizes 1, 2 are allowable, contributing $4 \times 1 + 6 \times 2 = 16$ to the sum. If none of the points is in the convex hull of the other three (take the vertices of a square for example), then there are 4 allowable submultisets of size 3 and 1 allowable submultiset of size 4, giving total sum of $16 + 4 \times 1 + 1 \times 1 \neq 20$. It is possible to modify the definition of 'allowable' to include such configurations. For example, from the proof of Theorem 6.5, it follows one can take any definition of 'allowable' for which (6.7) holds.

7. Kergin–Fejèr interpolation in $\mathbb{R}^d$

The **Hermite–Fejèr interpolation** operator at distinct points $x_1, x_2, \ldots, x_r$ in $\mathbb{R}$ is a modified form of Hermite interpolation at these points repeated with multiplicity 2, where the function values at $x_1, \ldots, x_r$ are matched, but the remaining reduced interpolation conditions from (5.18) are taken to be zero (i.e., the derivative of the interpolant is zero at the points $x_1, \ldots, x_n$). Let $\mathcal{K}_\Theta := \mathcal{H}_\Theta^{(0)}$ be the Kergin interpolation operator.

For any Θ, the linear projector $\mathcal{K}_\Theta : C^n(\mathbb{R}^d) \to \Pi_n(\mathbb{R}^d)$ can be written as

$$\mathcal{K}_\Theta = \bigoplus_{j=1}^{d} \bigoplus_{\substack{\Psi \subset \Theta \\ \dim(\Psi) = j-1 \\ \operatorname{aff}(\Psi) \cap^* \Theta = \Psi}} \bigoplus_{i=1}^{\#\Psi} \mathcal{K}_{\Theta,\Psi,i}, \qquad \mathcal{K}_{\Theta,\Psi,i} : C^n(\mathbb{R}^d) \to \Pi_{\Theta,\Psi,i}, \tag{7.1}$$

where $\mathcal{K}_{\Theta,\Psi,i}$ is the linear projector defined by the interpolation conditions $\Lambda_{\Psi,i} := \Lambda_{\Psi,i}^{(0)}$ of (5.7), and the space of interpolants

$$\Pi_{\Theta,\Psi,i} := \{p \in \Pi_n(\mathbb{R}^d) : p \in \ker(\Lambda_{\Phi,k}),\ \forall (\Phi, k) \neq (\Psi, i)\},$$

This 'direct sum' satisfies

$$\mathcal{K}_{\Theta,\Psi_1,i_1} \mathcal{K}_{\Theta,\Psi_2,i_2} = 0, \quad (\Psi_1, i_1) \neq (\Psi_2, i_2).$$

The polynomial spaces $\Pi_{\Theta,\Psi,i}$ are dual to the $\Lambda_{\Psi,i}$, and the dual of (5.6) is

$$\Pi_n(\mathbb{R}^d) = \bigoplus_{j=1}^{d} \bigoplus_{\substack{\Psi \subset \Theta \\ \dim(\Psi) = j-1 \\ \operatorname{aff}(\Psi) \cap^* \Theta = \Psi}} \bigoplus_{i=1}^{\#\Psi} \Pi_{\Theta,\Psi,i}. \tag{7.2}$$

In the above terms, the Hermite–Fejèr interpolation operator is

$$\mathcal{F}_\Theta := \bigoplus_{k=1}^{r} \mathcal{K}_{\Theta,[x_k,x_k],1}, \qquad \Theta := \{x_1, x_1, x_2, x_2, \ldots, x_r, x_r\}. \tag{7.3}$$

By analogy, we define **Kergin–Fejèr interpolation** in $\mathbb{R}^d$ by (7.3), when $x_1, \ldots, x_r$ are distinct points in $\mathbb{R}^d$. This operator has a continuous extension to functions defined at the points $x_1, \ldots, x_r$, and maps into an r–dimensional subspace of $\Pi_{2r-1}(\mathbb{R}^d)$. In order to compute this operator it is necessary to find the reduced interpolation conditions for Kergin interpolation at Θ, which we now do for points in general position in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Example 1 (Kergin–Fejèr interpolation in $\mathbb{R}^2$). The Ψ in the direct sum (7.1) have affine dimension 0 or 1. Those of dimension 0 are of the form $\Psi = [x_k, x_k]$, and give the matching of function values and first order derivatives at x_k, i.e., $f(x_k)$ and $\nabla f(x_k)$. Those of dimension 1 ($j = 2$) are of the form $\Psi = [x_s, x_s, x_t, x_t]$, with

$$\mathcal{A}_2(\Psi) = \{[x_s, x_t]\}, \quad \mathcal{A}_3(\Psi) = \{[x_s, x_s, x_t], [x_s, x_t, x_t]\}, \quad \mathcal{A}_4(\Psi) = \{[x_s, x_s, x_t, x_t]\}.$$

With $(x_t - x_s)^\perp$ retaining its previous definition, we can now write down the reduced interpolation conditions (as applied to f)

$$\begin{aligned}
&f(x_k), \quad \nabla f(x_k), && 3n \text{ conditions}, \\
&\int_{[x_s,x_t]} D_{(x_t-x_s)^\perp} f, \quad s \neq t, && \tfrac{1}{2}n(n-1) \text{ conditions}, \\
&\int_{[x_s,x_s,x_t]} D^2_{(x_t-x_s)^\perp} f, \quad s \neq t, && n(n-1) \text{ conditions}, \\
&\int_{[x_s,x_s,x_t,x_t]} D^3_{(x_t-x_s)^\perp} f, \quad s \neq t, && \tfrac{1}{2}n(n-1) \text{ conditions},
\end{aligned}$$

giving a total of $\dim(\Pi_{2r-1}(\mathbb{R}^2)) = \binom{2r+1}{2}$ linearly independent conditions.

Example 2 (Kergin–Fejèr interpolation in $\mathbb{R}^3$). The Ψ in (7.1) are of dimensions $0, 1, 2$, and are of the forms

$$\Psi = [x_s, x_s], \quad \Psi = [x_s, x_s, x_t, x_t], \quad \Psi = [x_s, x_s, x_t, x_t, x_u, x_u],$$

respectively (where s, t, u are distinct). Using $D_\perp$ to indicate a directional derivative in a direction orthogonal to the affine hull of the points that occur in the integral, we can express the reduced interpolation conditions (applied to f) as

$$\left.\begin{aligned}
&f(x_s), \quad \nabla f(x_s), && 4r \text{ conditions},
\end{aligned}\right\} j = 1$$

$$\left.\begin{aligned}
&\int_{[x_s,x_t]} D_\perp f, && 2\binom{r}{2} \text{ conditions}, \\
&\int_{[x_s,x_s,x_t]} D^2_\perp f, && 6\binom{r}{2} \text{ conditions}, \\
&\int_{[x_s,x_s,x_t,x_t]} D^3_\perp f, && 4\binom{r}{2} \text{ conditions},
\end{aligned}\right\} j = 2$$

$$\left.\begin{array}{lll}\displaystyle\int_{[x_s,x_t,x_u]} D_\perp^2 f, & \binom{r}{3} & \text{conditions,} \\ \displaystyle\int_{[x_s,x_s,x_t,x_u]} D_\perp^3 f, & 3\binom{r}{3} & \text{conditions,} \\ \displaystyle\int_{[x_s,x_s,x_t,x_t,x_u]} D_\perp^4 f, & 3\binom{r}{3} & \text{conditions,} \\ \displaystyle\int_{[x_s,x_s,x_t,x_t,x_u,x_u]} D_\perp^5 f, & \binom{r}{3} & \text{conditions,}\end{array}\right\} j=3$$

The author has not investigated to what extent the convergence properties (for large numbers of nodes) of Kergin–Fejèr interpolation in $\mathbb{R}^2$ and $\mathbb{R}^3$ are similar those of the well understood case of Hermite–Fejèr interpolation in $\mathbb{R}$. Any results in this direction would be of interest.

8. Computational considerations

In this final section, we outline how the subspaces $\Lambda_{\Psi,i}$, their duals $\Pi_{\Theta,\Psi,i}$, and the related polynomial interpolants $\mathcal{K}_{\Theta,\Psi,i}f$ of (7.1) can be computed numerically. Similar considerations apply to the other mean value interpolation operators.

Identification of the Ψ

The first step is to identify the submultisets Ψ in the direct sum decomposition (3.2), if they have not already been given.

In exact arithmetic this can be done simply by, at the j–th stage, taking the Ψ of dimension $j-1$ and adding new points to them (as many as possible) to get the Ψ of dimensional j. These calculations should be arranged efficiently. First determine the Ψ of dimension 0, i.e., find the multiplicities of each of the distinct points. At the j–th stage, take each Ψ of dimension $j-1$ and add to it a new point θ (up to full multiplicity) which gives a j–dimensional affine subspace different from one already found at this stage (i.e., the j–dimensional multiset $\Psi\cup\{\theta,\ldots,\theta\}$ is not contained some $\tilde{\Psi}$ of dimension j already obtained), then add to $\Psi\cup\{\theta,\ldots,\theta\}$ as many of the remaining points as is possible without changing its dimension, to obtain a new $\tilde{\Psi}$ of dimension j. Stop as soon as all the Ψ have been found, i.e., when $j=d$ or $\Psi=\Theta$.

If computations are only done to machine precision, then some tolerance, say `tol`, should be used when deciding whether or not a point lies in the affine hull of some set of points. The issues involved here are perhaps best illustrated by an example. Consider two points $\Theta := \{\theta_1,\theta_2\}$ in the plane which are close, but not equal. Though the map $(\theta_1,\theta_2)\mapsto\mathcal{K}_{[\theta_1,\theta_2]}f$ is continuous for $f\in C^1(\mathbb{R}^2)$, the computation of $\mathcal{K}_{[\theta_1,\theta_2]}f$ from the reduced interpolation conditions suffers a loss of significance as $\theta_1,\theta_2\to\theta$. Here the $\Lambda_{\Psi,i}$ are 1-dimensional

$$\Lambda_{[\theta_i],1}=\operatorname{span}\{f\mapsto f(\theta_i)\},\quad i=1,2,\qquad \Lambda_{[\theta_1,\theta_2],2}=\operatorname{span}\{f\mapsto\int_{[\theta_1,\theta_2]} D_{(\theta_2-\theta_1)^\perp}f\},$$

with corresponding dual polynomial spaces

$$\Pi_{\Theta,[\theta_1],1} = \mathrm{span}\{(\theta_2-\theta_1)^T(\cdot-\theta_2)\}, \quad \Pi_{\Theta,[\theta_2],1} = \mathrm{span}\{(\theta_2-\theta_1)^T(\cdot-\theta_1)\},$$
$$\Pi_{\Theta,[\theta_1,\theta_2],2} = \mathrm{span}\{(\theta_2-\theta_1)^{\perp T}(\cdot-\frac{\theta_2+\theta_1}{2})\}.$$

This gives the Kergin interpolant

$$\begin{aligned}\mathcal{K}_{[\theta_1,\theta_2]}f(x) &= f(\theta_1)\frac{(\theta_1-\theta_2)^T(x-\theta_2)}{\|\theta_2-\theta_1\|^2} + f(\theta_2)\frac{(\theta_2-\theta_1)^T(x-\theta_1)}{\|\theta_2-\theta_1\|^2} \\ &\quad + \int_{[\theta_1,\theta_2]} D_{(\theta_2-\theta_1)^\perp} f\, \frac{(\theta_2-\theta_1)^{\perp T}\,(2x-(\theta_1+\theta_2))}{2\|(\theta_2-\theta_1)^\perp\|^2}.\end{aligned} \tag{8.1}$$

There is a loss of significance in the computed values of each of the terms in (8.1) when θ_1, θ_2 are close to θ. Here the calculated value of $\mathcal{K}_{[\theta_1,\theta_2]}f$ should be close to $\mathcal{K}_{[\theta,\theta]}f$ (the tangent plane to f at θ), which can be computed from $f(\theta)$, $\nabla f(\theta)$ without loss of significance. This behaviour accurately reflects the general situation.

In general, any points $\theta \in \Theta$ at a distance less than `tol` from the affine hull of a Ψ being computed should be included in Ψ (without increasing its affine dimension), and the Kergin interpolant calculated from the corresponding reduced interpolation conditions. To do this, translate the points in Ψ (and θ) so that the affine subspace they span becomes a linear subspace (subtract one of the points, or better their average, from each of them). Form the matrix $A = [\Psi, \theta]$ with columns given by the points in Ψ and θ, and find a QR–factorisation $A = QR$. The absolute value of the last diagonal entry of R is the distance of θ from the affine subspace $\mathrm{aff}(\Psi)$, and if this is $<$ `tol`, then θ should be included in Ψ. If the translation of subtracting to first point in Ψ is adopted, then the Q and R can simply be updated one column at a time as each point is added to Ψ. Alternatively, one could use a standard package such as MATLAB's `rank(A,tol)` to compare the numerical rank of $A = [\Psi]$ and $A = [\Psi, \theta]$.

Finding the interpolant $\mathcal{K}_\Theta f$, $\Lambda_{\Psi,i}$ and $\Pi_{\Theta,\Psi,i}$

Now suppose the Ψ in the direct sum have been obtained, either exactly or numerically (as just discussed). To find the Kergin interpolant $\mathcal{K}_\Theta f$, or the interpolants $\mathcal{K}_{\Theta,\Psi,i}f$ (where some of the interpolation conditions are zero) it is necessary to solve a linear system for the coefficients $\{a_k\}$ of the interpolant $p = \sum_k a_k p_k$ relative to some choice of basis $\{p_k\}$ for $\Pi_{n-1}(\mathbb{R}^d)$, i.e.,

$$\sum_k a_k \lambda(p_k) = \lambda(p) = \lambda(f), \tag{8.2}$$

for λ from some spanning set for the reduced interpolation conditions. One could take $\{p_k\}$ to be the basis of shifted monomials with centre the average of the points in Θ. Choosing a spanning set of $\lambda \in \Lambda_{\Psi,i}$, as defined by (5.7), can be done as follows. Let $N := \{u_1, u_2, \ldots, u_{d-j+1}\}$ be an orthonormal basis for $\mathrm{aff}(\Psi)^\perp$, and choose the $p(D)$ in (5.12) to be of the form $p(D) = D_{y_1}D_{y_2}\cdots D_{y_{i-1}}$, where $y_1, \ldots, y_{i-1} \in N$. For each such choice, taking s from a basis for $\mathcal{S}_i(\Psi)$ leads to a basis of λ for $\Lambda_{\Psi,i}$. In the absence of a known basis for $\mathcal{S}_i(\Psi)$, it is natural to take s from the spanning set $\{M(\cdot|\Phi) : \Phi \in \mathcal{A}_i(\Theta)\}$, which in general leads to a system (8.2) of more equations than unknowns. If solved numerically, this system should be solved using the least squares method. Even if a basis for $\mathcal{S}_i(\Psi)$ is available, as is outlined in some special cases in Example 16, the author suggests finding the least squares solution of the (possibly) larger system of equations (which is symmetric in the points Ψ, and likely to give a more accurate answer).

The dual polynomial spaces $\Pi_{\Theta,\Psi,i}$ can be found as follows. Order the $\Lambda_{\Psi,i}$ as $\Lambda_1, \Lambda_2, \ldots$, and suppose that a spanning set $\lambda_1^{(k)}, \ldots \lambda_{m_k}^{(k)}$ is given for each Λ_k, as above. Let $\Lambda : \Pi_{n-1}(\mathbb{R}^d) \to \mathbb{R}^m$, $m := \sum_k m_k$, be the data map given by

$$\Lambda(f) := \begin{bmatrix} \Lambda_1(f) \\ \Lambda_2(f) \\ \vdots \end{bmatrix}, \qquad \Lambda_k(f) := \begin{bmatrix} \lambda_1^{(k)}(f) \\ \vdots \\ \lambda_{m_k}^{(k)}(f) \end{bmatrix},$$

and $P : \mathbb{R}^\ell \to \Pi_{n-1}(\mathbb{R}^d)$, $\ell := \binom{n-1+d}{d}$, be the basis map $(a_k) \mapsto \sum_k a_k p_k$. Apply elementary column operations to the $m \times \ell$ matrix $A := \Lambda P$ to obtain a *block diagonal* form

$$AM = (\Lambda P)M = \begin{bmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & C_n \end{bmatrix}, \qquad M \in \mathbb{R}^{\ell\times\ell} \text{ invertible},$$

where the blocks C_k are of size $m_k \times \ell_k$, $\ell_k := \dim(\Lambda_k)$, and have linearly independent columns. This can be done by using Gauss elimination (with pivoting) to find a column echelon form for A, or using *Gauss elimination by segments* applied to the 'segmented matrix'

$$A^T = [(\Lambda_1 P)^T, (\Lambda_2 P)^T, \ldots] \in \mathbb{R}^{\ell\times m}$$

(as detailed in de Boor [B94]). The sequence of ℓ_k polynomials

$$P(Me_i), \qquad k = \ell_1 + \cdots + \ell_{k-1} + 1, \ldots, \ell_1 + \cdots + \ell_k$$

is a basis for Λ_k (above Me_i is the i–th column of the matrix M).

A typical use of the dual polynomial spaces is in the computation of the Kergin–Fejèr interpolant $\mathcal{F}_\Theta(f)$ from the point evaluations of f at $\{x_1, \ldots, x_r\}$, which requires a basis for the r–dimensional space

$$\bigoplus_{k=1}^{r} \Pi_{\Theta,[x_k,x_k],1} \subset \Pi_{2r-1}(\mathbb{R}^d), \qquad \Theta := \{x_1, x_1, \ldots, x_r, x_r\}.$$

Code for the computations outlined in this section can be obtained from the author.

Acknowledgement

Thanks to Thomas Grandine (Seattle) for providing the data for the plots of simplex splines in Figures 5.1 and 5.2 from his implementation of [Gr86].

References

[BHS93] B. D. Bojanov, H. A. Hakopian, and A. A. Sahakian, "Spline functions and multivariate interpolations", Kluwer Academic Publishers, , 1993.

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

[BC97] L. Bos and J.-P. Calvi, Kergin interpolants at the roots of unity approximate C^2 functions, *J. Anal. Math.* **72** (1997), 203–221.

[BW01] L. Bos and S. Waldron, On the structure of Kergin interpolation for points in general position, To appear, 2001.

[BS94] S. C. Brenner and L. R. Scott, "The mathematical theory of finite element methods", Springer-Verlag, New York, 1994.

[CGMS83] A. S. Cavaretta, T. N. T. Goodman, C. A. Micchelli, and A. Sharma, Multivariate interpolation and the Radon transform, part III: Lagrange representation, *Canad. Math. Soc. Conf. Proceedings* **3** (1983), 37–50.

[C78] P. G. Ciarlet, "The finite element method for elliptic problems", North Holland, , 1978.

[DM83] W. A. Dahmen and C. A. Micchelli, On the linear independence of multivariate B-splines II: complete configurations, *Math. Comp.* **41(163)** (1983), 143–163.

[G83] T. N. T. Goodman, Interpolation in minimum semi-norm, and multivariate B-splines, *J. Approx. Theory* **37** (1983), 212–223.

[Gr86] T. A. Grandine, The stable evaluation of multivariate simplex splines, *Math. Comp.* **50** (1988), 197–205.

[H81] H. Hakopian, Les differences divisées de plusieurs variables et les interpolations multidimensionnelles de types Lagrangien et Hermitien, *C. R. Acad. Sci. Paris Ser. I* **292** (1981), 453-456.

[H84] Hakop Hakopian, On multivariate spline functions, B-spline bases and polynomial interpolation II, *Studia Math.* **79** (1984), 91–102.

[K80] P. Kergin, A natural interpolation of C^k functions, *J. Approx. Theory* **29** (1980), 278–293.

[L92] R. A. Lorentz, "Multivariate Birkhoff Interpolation", Lecture Notes in Mathematics, No. 1516, Springer Verlag, Heidelberg, 1992.

[M80] C. A. Micchelli, A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B-splines and Lagrange interpolation, *Rocky Mountain J. Math.* **10** (1980), 485–497.

[W97$_1$] S. Waldron, Integral error formulæ for the scale of mean value interpolations which includes Kergin and Hakopian interpolation, *Numer. Math.* **77(1)** (1997), 105–122.

[W97$_2$] S. Waldron, A multivariate form of Hardy's inequality and L_p-error bounds for multivariate Lagrange interpolation schemes, *SIAM J. Math. Anal.* **28(1)** (1997), 233–258.