Technical Report
1 June 1999

Chu–Vandermonde convolution identities and reduced Kergin interpolation

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

A binomial identity which counts the number of 'allowable' submultisets of any finite multiset of points in $\mathbb{R}^d$ is proved. For example, if the multiset consists of n points in general position each repeated m times, then the identity gives

$$\binom{mn-1+d}{d} = \sum_{j=1}^{d} \binom{n}{j} \sum_{k_1=1}^{m} \cdots \sum_{k_j=1}^{m} \binom{k_1+\cdots+k_j-1+d-j}{d-j}.$$

This reduces to the Chu–Vandermonde convolution identity when $m = 1$. The most general form of the result involves the dimension of certain spaces of simplex splines, and is closely related to finding the 'reduced' interpolation conditions for Kergin interpolation at a finite multiset of points in $\mathbb{R}^d$.

Key Words: binomial identity, Chu–Vandermonde identity, simplex spline, Kergin interpolation, Kergin–Fejèr interpolation, multivariate interpolation

AMS (MOS) Subject Classifications: primary 05A19, 41A05, 41A10, 65D07, secondary 05A10, 41A35, 41A63, 65D05,

1. Introduction – a curious conjecture

This paper concerns certain combinatorial–geometric properties of the submultisets of a finite multiset Θ of points in $\mathbb{R}^d$. The affine and convex hulls of a multiset Θ are those of the underlying set, and will be denoted by $\mathrm{aff}(\Theta)$ and $\mathrm{conv}(\Theta)$ respectively. The (affine) dimension of a submultiset $\tilde{\Theta} \subset \Theta$ is defined to be the dimension of the affine hull of underlying set: $\dim(\tilde{\Theta}) := \dim(\mathrm{aff}(\tilde{\Theta}))$. For example, if $\tilde{\Theta}$ 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 a multiset (or the 'number of times it occurs') is its repetition number. See the recent handbook [GO97] for general background.

Definition 1.1. *A submultiset $\tilde{\Theta} \subset \Theta$ is said to be* **allowable** *if*
(a) $\Theta \cap \mathrm{conv}(\tilde{\Theta}) \subset \tilde{\Theta}$, *i.e., all points of Θ in $\mathrm{conv}(\tilde{\Theta})$ are in $\tilde{\Theta}$.*
(b) *points of Θ in the relative interior of $\mathrm{conv}(\tilde{\Theta})$ are repeated in $\tilde{\Theta}$ with full multiplicity.*
For short we will write $\tilde{\Theta} \subset^* \Theta$.

Our aim is to understand why the following conjecture is true in so many cases, and in what way it should be modified in those cases where it is not. The size of a multiset Θ is its number of elements (counting repetitions), and is denoted by $\#\Theta$.

Conjecture 1.2. *For any multiset Θ of n points in $\mathbb{R}^d$*

$$\binom{n-1+d}{d} = \sum_{i=1}^{n} \sum_{\substack{\tilde{\Theta} \subset^* \Theta \\ \#\tilde{\Theta}=i}} \binom{i+d-\dim(\tilde{\Theta})-2}{d-\dim(\tilde{\Theta})-1}, \tag{1.3}$$

where the summation in (1.3) *is over all allowable submultisets (of size i).*

It turns out there is a related result (Theorem 2.6), valid for all Θ, which reduces to this conjecture in many cases (where it is true), and to modified versions of it in other cases. This result involves the dimension of certain spaces of simplex splines, and comes from finding a basis of 'reduced' interpolation conditions for Kergin interpolation at a finite multiset of points in $\mathbb{R}^d$. First we give some motivating examples of Conjecture 1.2.

Example 1 (All points coincident)

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

$$\binom{n-1+d}{d} = \sum_{i=1}^{n} \binom{i+d-2}{d-1}. \tag{1.4}$$

Example 2 (Points in general position)

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

$$\binom{n-1+d}{d} = \sum_{i=1}^{d} \binom{n}{i} \binom{i-1+(d-i+1)-1}{(d-i+1)-1} = \sum_{i=1}^{d} \binom{n}{i} \binom{d-1}{d-i}, \tag{1.5}$$

which is the *Chu–Vandermonde convolution identity.*

Example 3 (All points on a line in $\mathbb{R}^d$, with no repetitions)

If there are n points in $\mathbb{R}^d$ all distinct and lying on a line, then there are exactly $n+1-i$ allowable subsets of size i (consecutive points), which leads to

$$\binom{n-1+d}{d} = n + \sum_{i=2}^{n} (n+1-i)\binom{i+d-3}{d-2}, \tag{1.6}$$

and

$$\binom{n-1+d}{d} = \sum_{i=1}^{n} (n+1-i)\binom{i+d-3}{d-2}, \qquad d \geq 2. \tag{1.7}$$

These identities can be proved by induction.

Example 4 (All points on a line in $\mathbb{R}^d$, with multiplicities)

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 $m_1 + \cdots + m_r = n$. 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 (1.4) to count the submultisets of one point repeated, we obtain the binomial identity

$$\begin{aligned}\binom{n-1+d}{d} = &\sum_{i=1}^{r} \binom{m_i-1+d}{d} \\ &+ \sum_{1\leq i<j\leq r} \sum_{k_1=1}^{m_i} \sum_{k_2=1}^{m_j} \binom{k_1 + m_{i+1} + m_{i+2} + \cdots + m_{j-1} + k_2 - 1 + d - 2}{d-2}.\end{aligned} \tag{1.8}$$

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

$$\binom{n-1+d}{d} = n + \sum_{1\leq i<j\leq n} \binom{j-i+d-2}{d-2},$$

and

$$\binom{n-1+d}{d} = \sum_{1\leq i\leq j\leq n} \binom{j-i+d-2}{d-2}, \qquad d \geq 2.$$

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

$$\binom{n+1}{2} = \binom{m_1 + \cdots + m_r + 1}{2} = \sum_{i=1}^{r} \binom{m_i+1}{2} + \sum_{1\leq i<j\leq r} m_i m_j, \tag{1.9}$$

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

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

Example 5 (Points in general position in $\mathbb{R}^d$, each repeated m times)

Suppose that Θ consists of n 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 $\binom{n}{i_1}$ 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

$$\begin{aligned}&\binom{mn-1+d}{d}\\ &= \sum_{i_1=1}^{d}\binom{n}{i_1}\sum_{i_2=0}^{i_1}\binom{i_1}{i_2}\sum_{i_3=0}^{i_2}\binom{i_2}{i_3}\cdots\sum_{i_m=0}^{i_{m-1}}\binom{i_{m-1}}{i_m}\binom{d+i_2+i_3+\ldots+i_m-1}{d-i_1}.\end{aligned} \tag{1.10}$$

It can be verified that this identity does indeed hold.

Example 6 (Points in general position in $\mathbb{R}^d$, with multiplicities)

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

$$\binom{n-1+d}{d} = \sum_{j=1}^{d}\sum_{1\leq i_1<i_2<\cdots<i_j\leq r}\sum_{k_1=1}^{m_{i_1}}\cdots\sum_{k_j=1}^{m_{i_j}}\binom{k_1+\cdots+k_j-1+d-j}{d-j}, \tag{1.11}$$

which can be verified. In the case of the previous example ($r = n$, $m_1, \ldots, m_r = m$), this gives

$$\binom{mn-1+d}{d} = \sum_{j=1}^{d}\binom{n}{j}\sum_{k_1=1}^{m}\cdots\sum_{k_j=1}^{m}\binom{k_1+\cdots+k_j-1+d-j}{d-j}.$$

Compare this with (1.10), which was obtained by counting the allowable submultisets in a different way.

Example 7 (Points in $\mathbb{R}$ and $\mathbb{R}^2$)

The conjecture holds when $d = 1$ and $d = 2$. For $d = 1$, the submultisets which contribute to the summation contain exactly one point repeated a number of times. There

are n such multisets, and each contributes 1 to the sum. For $d = 2$, two types of allowable submultisets contribute to the summation. Suppose that the distinct points in $\mathbb{R}^2$ are $x_1, \ldots, x_r$, and they occur with multiplicities $m_1, \ldots, m_r$, where $m_1 + \cdots + m_r = n$. Then the allowable submultisets of dimension 0 consist of x_i repeated $j = 1, \ldots, m_i$ times, and contribute $\binom{m_i+1}{2}$ to the summation. The allowable subsets of dimension 1 are obtained by taking two distinct points $x_i \neq x_j$, the first with multiplicity $1, \ldots, m_i$ and the second with multiplicity $1, \ldots, m_j$, together with all the points that lie on the line segment joining them (up to full multiplicity). Each such submultiset contributes 1 to the sum, and so the conjecture becomes

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

which is (1.9). This can be proved by a simple inductive argument.

Example 8 (An example where the conjecture fails in $\mathbb{R}^3$)

For $d \geq 3$, there are examples where the conjecture fails (as stated). The first of these is for 4 distinct points in $\mathbb{R}^3$ which lie in a plane, of which no three are collinear. In this case $\binom{n-1+d}{d} = 20$. All submultisets of sizes 1, 2 are all allowable, contributing $4 \times 1 + 6 \times 2 = 16$ to the summation. 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 summation of $16 + 4 \times 1 + 1 \times 1 \neq 20$. If Definition 1.1 is modified so that the convex hull of an allowable submultiset must be a simplex (of some dimension), then the previous examples continue to hold, as does the conjecture for any multiset of ≤ 4 points in $\mathbb{R}^3$. Though, as the example of $n \geq 5$ points which lie on a circle in $\mathbb{R}^3$ shows, this is through chance, rather than finding the "correct version" of the conjecture. This particular example (n points on a circle in $\mathbb{R}^d$, $d \geq 2$) will be analysed further in the next section.

2. Why the conjecture fails – a closely related result

We now prove a related result (Theorem 2.6), which allows us to understand why the conjecture holds in the first seven examples, and fails in Example 8.

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

$$f \mapsto \int_\Theta f := \int_0^1 \int_0^{s_1} \ldots \int_0^{s_{k-1}} f(\theta_0 + s_1(\theta_1 - \theta_0) + \cdots + s_k(\theta_k - \theta_{k-1}))\, ds_k \cdots ds_2\, ds_1. \quad (2.1)$$

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

$$\int_{\mathrm{aff}(\Theta)} M(\cdot|\Theta) f = \int_{\mathrm{conv}(\Theta)} M(\cdot|\Theta) f = k! \int_\Theta f, \quad (2.2)$$

which ensures it integrates to 1. Properties of these splines, such as their smoothness, approximation properties, computational algorithms, etc, have been studied in great detail (see, e.g., [BHS93]). We will only be interested in the dimension of a particular spline space $\mathcal{S}_i(\Theta)$, which occurs in Theorem 2.6. This space of functions on $\mathrm{aff}(\Theta)$ is defined as

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

where

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

When $\mathcal{A}_i(\Theta) = \{\tilde{\Theta} \subset \Theta : \#\tilde{\Theta} = i\}$, as is the case for Θ in general position in $\mathrm{aff}(\Theta)$, Dahmen and Micchelli [DM83:Th.1] have shown that

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

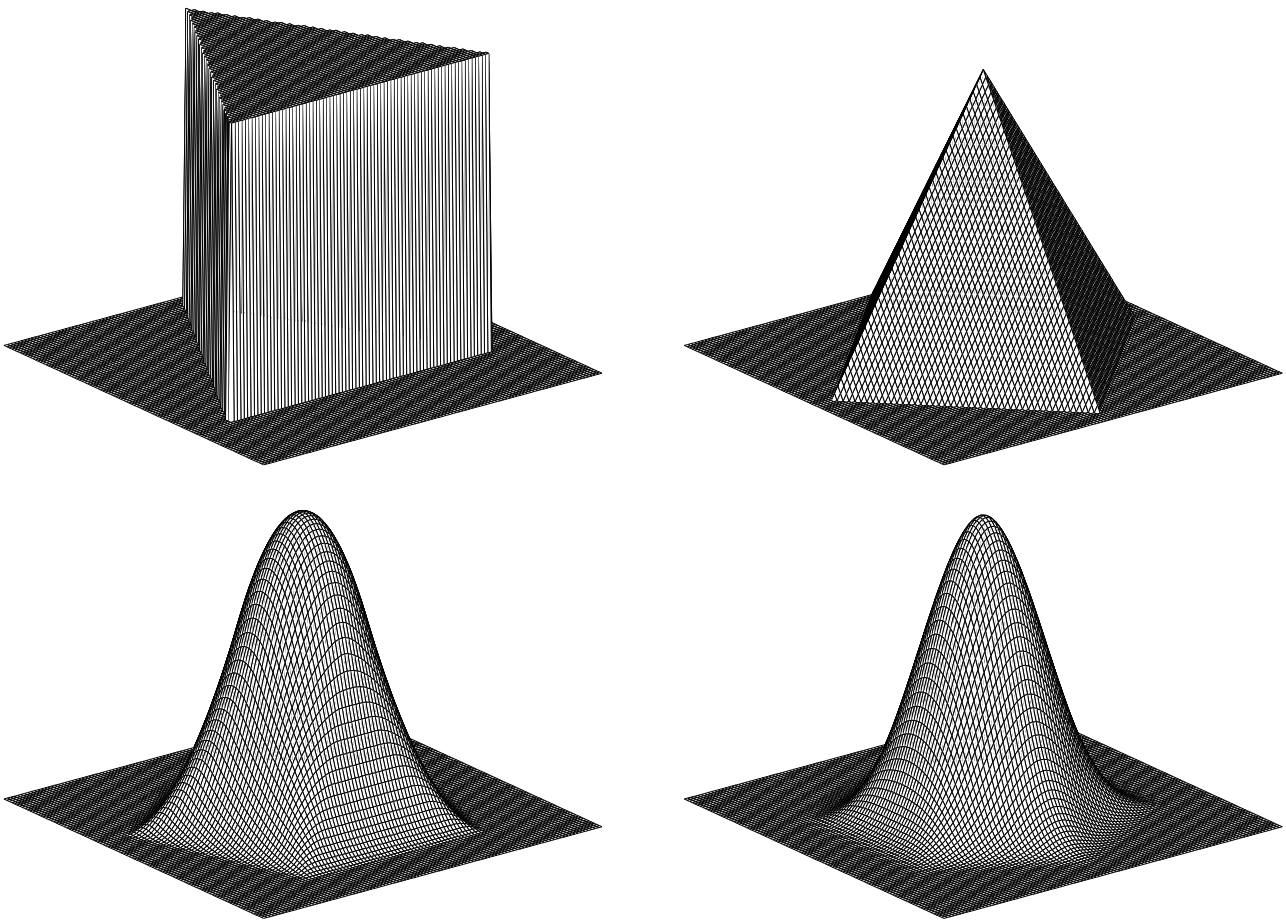

Fig. 2.1. Simplex splines for points equally spaced on a circle in $\mathbb{R}^2$: 3 knots (discontinous piecewise constant function), 4 knots (continuous piecewise linear function), 5 knots (C^1 piecewise quadratic function), and 6 knots (C^2 piecewise cubic function).

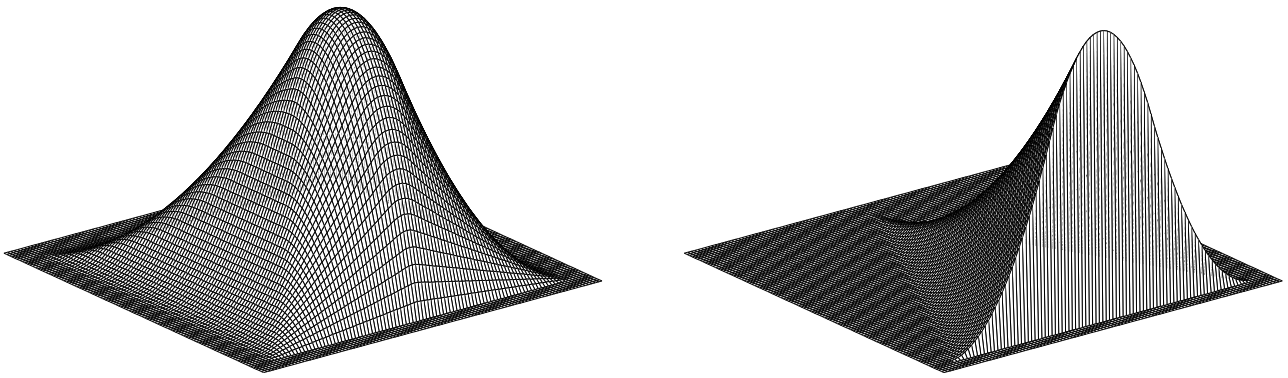

Fig. 2.2. Simplex splines for 5 points in $\mathbb{R}^2$ which are not in general position (piecewise quadratics). In the first 3 knots are collinear, and the spline is only C^0 across this line. In the second 4 knots are collinear, and the spline is discontinuous across this line.

Lemma 2.3. *Let Ψ consist of j affinely independent points with multiplicities $m_1, \ldots, m_j$. Then for $i = j, \ldots, \#\Psi$*

$$\dim(\mathcal{S}_i(\Psi)) = \#\mathcal{A}_i(\Psi) = c_i(m_1, \ldots, m_j), \tag{2.4}$$

where

$$c_i(m_1, \ldots, m_j) := \textit{coefficient of } t^i \textit{ in } \prod_{k=1}^{j} (t + t^2 + \cdots + t^{m_k}). \tag{2.5}$$

Proof: The submultisets in $\mathcal{A}_i(\Psi)$ consist of those $\tilde{\Psi}$ 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 (2.5), or by using the principle of inclusion and exclusion.

Let $\xi_1, \ldots, \xi_j$ be the barycentric coordinates of a point in $\mathrm{aff}(\Psi)$ relative to the j points, and $k_1, \ldots, k_j$ be the multiplicities of these points as they occur in $\tilde{\Psi} \in \mathcal{A}_i(\Psi)$. Then the restriction of $M(\cdot|\tilde{\Psi})$ to the simplex $T := \mathrm{conv}(\Psi)$ 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 $M(\cdot|\tilde{\Psi})$ are linearly independent, and so $\dim(\mathcal{S}_i(\Psi)) = \#\mathcal{A}_i(\Psi)$. □

Here is a combinatorial version of the related result. 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.

Theorem 2.6 (Combinatorial version). *For any multiset Θ of n points in $\mathbb{R}^d$*

$$\binom{n-1+d}{d} = \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^* \Theta = \Psi}} \sum_{i=j}^{\#\Psi} \binom{i-1+d-j}{d-j} \dim(\mathcal{S}_i(\Psi)), \tag{2.7}$$

where

$$1 \le \dim(\mathcal{S}_i(\Psi)) \le \binom{\#\Psi - i + j - 1}{j-1}. \tag{2.8}$$

The upper bound in (2.8) is attained when $\mathcal{A}_i(\Psi) = \{\tilde{\Psi} \subset \Psi : \#\tilde{\Psi} = i\}$. 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 $\le i - j$.

Before proving this, we give some examples where it reduces to Conjecture 1.2.

Example 9 (Points in general position in an affine subspace of $\mathbb{R}^d$)

Suppose that Θ consists of $n > s$ (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

$$\binom{n-1+d}{d} = \sum_{j=1}^{s} \binom{n}{j}\binom{d-1}{d-j} + \sum_{i=s+1}^{n} \binom{i+d-s-2}{d-s-1}\binom{n-i+s}{s}, \quad n > s,\ 1 \le s \le d.$$

In the special cases of Examples 2 and 3 ($s = d$ and $s = 1$), this reduces to (1.5) and (1.6). Another special case is the following.

Example 10 (Points on a circle in $\mathbb{R}^d$)

This a more general form of Example 8, for which Conjecture 1.2 failed (4 points on a circle in $\mathbb{R}^3$). Let Θ be $n > 2$ points in general position in some plane in $\mathbb{R}^d$, $d \ge 2$, e.g., points on some circle. Then the Ψ contributing to the sum are those of size 1, 2 and n, which gives the identity

$$\binom{n-1+d}{d} = \binom{n}{1} + \binom{n}{2}(d-1) + \sum_{i=3}^{n} \binom{i+d-4}{d-3}\binom{n-i+2}{2}.$$

For this example Conjecture 1.2 is

$$\binom{n-1+d}{d} = \binom{n}{1} + \binom{n}{2}(d-1) + \sum_{i=3}^{n} \binom{n}{3}\binom{i+d-4}{d-3},$$

which *over counts* the submultisets of size $i = 3, \ldots, n$ for large n, and so is false (e.g., for 4 points on a circle in $\mathbb{R}^3$).

Example 11 (Points in general position in an affine subspace of $\mathbb{R}^d$, with multiplicities)

Suppose that Θ consists of $r > s$ points in general position in an s–dimensional affine subspace of $\mathbb{R}^d$, which are repeated with multiplicities $m_1, \ldots, m_r$, $m_1 + \cdots + m_r = n$. Then the submultisets Ψ which contribute to the summation are those consisting of $j = 1, \ldots, s$ of the points with full multiplicity, and $\Psi = \Theta$. This leads to

$$\begin{aligned}\binom{n-1+d}{d} = \sum_{j=1}^{s} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le r} \sum_{i=j}^{m_{i_1}+\cdots+m_{i_j}} \binom{i-1+d-j}{d-j} c_i(m_{i_1}, \ldots, m_{i_j}) \\ + \sum_{i=s+1}^{n} \binom{i+d-s-2}{d-s-1} \dim(\mathcal{S}_i(\Theta)),\end{aligned} \tag{2.9}$$

where $c_i(m_{i_1}, \ldots, m_{i_j})$ is given by (2.5). For $s = 1$ (points on a line) this reduces to

$$\binom{n-1+d}{d} = \sum_{k=1}^{r} \sum_{i=1}^{m_k} \binom{i+d-2}{d-1} + \sum_{i=2}^{n} \binom{i+d-3}{d-2} \dim(\mathcal{S}_i(\Theta)), \tag{2.10}$$

which is different from the formula (1.8) obtained by counting in a different way. Both these formulæ are true, and in Corollary 2.12 we show why they are equivalent. The numbers $\dim(\mathcal{S}_i(\Theta))$ occuring in (2.10) depend on the multiplicities $m_1, \ldots, m_r$, and are easily computed. If the points in Θ are ordered $\theta_1 \le \theta_2 \le \cdots \le \theta_n$ (as they appear on the line), then $\mathcal{S}_i(\Theta)$ has a B–spline basis (see, e.g., [BHS93])

$$\{M(\cdot|\theta_j, \theta_{j+1}, \ldots, \theta_{j+i-1}) : \theta_j < \theta_{j+i-1},\ j = 1, 2, \ldots, n-i+1\},$$

and so

$$\begin{aligned}\dim(\mathcal{S}_i(\Theta)) &= \#\{j : \theta_j < \theta_{j+i-1},\ j = 1, 2, \ldots, n-i+1\} \\ &= n + 1 - i - \sum_{k=1}^{r} \max\{m_k - i, 0\}.\end{aligned} \tag{2.11}$$

When $r = n$ ($m_i = 1$, $\forall i$), this gives $\dim(\mathcal{S}_i(\Theta)) = n + 1 - i$, and (2.10) reduces to (1.6). For the case $s = d$ (Example 6) the formula (2.9) gives

$$\binom{n-1+d}{d} = \sum_{j=1}^{d} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le r} \sum_{i=j}^{m_{i_1}+\cdots+m_{i_j}} \binom{i-1+d-j}{d-j} c_i(m_{i_1}, \ldots, m_{i_j}).$$

Compare this with (1.11).

We now give the link between Conjecture 1.2 and Theorem 2.6.

Corollary 2.12 (Relationship with the conjecture). *Suppose for each* $\Psi \subset \Theta$, *with* $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$, *and* $\dim(\Psi) = j - 1 < d$, *that*

$$\#\{\tilde{\Psi} \in \mathcal{A}_i(\Psi) : \tilde{\Psi} \textit{ is allowable}\} = \dim(\mathcal{S}_i(\Psi)), \qquad i = j, \ldots, \#\Psi. \tag{2.13}$$

Then equation (1.3) *of* Conjecture 1.2 *holds. In particular, this happens when*

(a) Θ *consists of points in general position in an* s*–dimensional affine subspace of* $\mathbb{R}^d$, *with repetitions allowed, where* $s = 0$, $s = 1$ *or* $s = d$.

(b) Θ *is any configuration of points in* $\mathbb{R}$ *or* $\mathbb{R}^2$.

Proof: Every allowable $\tilde{\Theta} \subset \Theta$ which contributes to the summation (1.3) belongs to a unique Ψ with $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$ and $\dim(\Psi) = j - 1 < d$, namely $\Psi := \mathrm{aff}(\tilde{\Theta}) \cap^* \Theta$. Hence

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

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

If (2.13) is satisfied, then the RHS (right hand side) of the above is term by term equal to the RHS of (2.7), and so the conjecture holds.

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 $\tilde{\Psi} \in \mathcal{A}_i(\Psi)$ is allowable, and so Lemma 2.3 gives

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

which is (2.13). If $\dim(\Psi) = s$, then there is exactly one Ψ with $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$, namely $\Psi = \Theta$, and (2.13) is only required to hold when $s < d$. By (2.11), it holds for $s = 1$ because

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

For $s = 0$, Θ is a single point repeated, and (2.11) 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 (2.13) holds for $j - 1 = 0, 1$, which again follows from (a). □

It is possible to modify the definition of 'allowable' submultisets, and the way they are counted, so that Conjecture 1.2 is true in appropriate situations. The basic result in this direction is the following.

Corollary 2.14 (Different allowable submultisets). *For each $\Psi \subset \Theta$ which satisfies $\mathrm{aff}(\Psi) \cap^* \Theta = \Psi$, $\dim(\Psi) = j-1 < d$, and each $i = j, \ldots, \#\Psi$, choose a collection of submultisets*

$$\mathcal{B}_i(\Psi) \subset \{\tilde{\Theta} \subset \Psi : \#\tilde{\Theta} = i\}, \tag{2.15}$$

for which $\{M(\cdot|\tilde{\Theta}) : \tilde{\Theta} \in \mathcal{B}_i(\Psi)\}$ is a basis for $\mathcal{S}_i(\Psi)$. If

$$\mathcal{A} := \bigcup_{j=1}^{d} \bigcup_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \bigcup_{i=j}^{\#\Psi} \mathcal{B}_i(\Psi) \tag{2.16}$$

is deemed to be the set of allowable submultisets, then Conjecture 1.2 holds.

Proof: Since $\{M(\cdot|\tilde{\Theta}) : \tilde{\Theta} \in \mathcal{B}_i(\Psi)\}$ is a basis for $\mathcal{S}_i(\Psi)$, each $\tilde{\Theta} \in \mathcal{B}_i(\Psi)$ has $\dim(\tilde{\Theta}) = j-1$, and $\#\mathcal{B}_i(\Psi) = \dim(\mathcal{S}_i(\Psi))$, which gives

$$\begin{aligned}\sum_{i=1}^{n} \sum_{\substack{\tilde{\Theta} \subset^* \Theta \\ \#\tilde{\Theta}=i}} \binom{i+d-\dim(\tilde{\Theta})-2}{d-\dim(\tilde{\Theta})-1} &= \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \sum_{\tilde{\Theta} \in \mathcal{B}_i(\Psi)} \binom{i-1+d-j}{d-j} \\ &= \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \binom{i-1+d-j}{d-j} \dim(\mathcal{S}_i(\Psi)).\end{aligned}$$

Hence, by (2.7) the conjecture holds. □

We now give an alternative statement of Theorem 2.6 (involving Kergin interpolation), and its proof. Let $\Pi_k(S)$ and $\Pi_k^0(S)$ denote the polynomials defined on S an s–dimensional affine subspace of $\mathbb{R}^d$ that are of degree $\le k$ and homogeneous of degree k, respectively. The dimensions of these spaces are

$$\dim(\Pi_k(S)) = \binom{k+s}{s}, \qquad \dim(\Pi_k^0(S)) = \binom{k+s-1}{s-1}. \tag{2.17}$$

The operator of **Kergin interpolation** at a multiset Θ of n points in $\mathbb{R}^d$ is the linear projector $\mathcal{K}_\Theta$ of $C^{n-1}(\mathbb{R}^d)$ onto $\Pi_{n-1}(\mathbb{R}^d)$ defined by the conditions

$$\int_{\tilde{\Theta}} q(D)\mathcal{K}_\Theta f = \int_{\tilde{\Theta}} q(D) f, \qquad \forall f \in C^{n-1}(\mathbb{R}^d), \tag{2.18}$$

where $\tilde{\Theta} \subset \Theta$, and $q(D)$ is any constant coefficient homogeneous differential operator of order $\#\tilde{\Theta} - 1$ in d variables, i.e., $q \in \Pi^0_{\#\tilde{\Theta}-1}(\mathbb{R}^d)$. It is convenient to write (2.18) as

$$\lambda(\mathcal{K}_\Theta f) = \lambda(f), \qquad \forall \lambda \in \Lambda_\Theta, \quad \forall f \in C^{n-1}(\mathbb{R}^d), \tag{2.19}$$

where Λ_Θ is the space of **interpolation conditions**

$$\Lambda_\Theta := \mathrm{span}\{f \mapsto \int_{\tilde\Theta} q(D)f : \tilde\Theta \subset \Theta,\ q \in \Pi^0_{\#\tilde\Theta-1}(\mathbb{R}^d)\} \tag{2.20}$$

(a space of distributions on $\mathbb{R}^d$). Since $\mathcal{K}_\Theta$ reproduces $\Pi_{n-1}(\mathbb{R}^d)$,

$$\dim(\Lambda_\Theta) = \dim(\Lambda_\Theta|_{\Pi_{n-1}(\mathbb{R}^d)}) = \dim(\Pi_{n-1}(\mathbb{R}^d)) = \binom{n-1+d}{d}. \tag{2.21}$$

For $\Theta = [\theta_1, \theta_2, \ldots, \theta_n]$ an explicit formula for $\mathcal{K}_\Theta$ is given by the *Newton form*

$$\begin{aligned}\mathcal{K}_\Theta f(x) = f(\theta_1) + \int_{[\theta_1,\theta_2]} D_{x-\theta_1} f + \int_{[\theta_1,\theta_2,\theta_3]} D_{x-\theta_1} D_{x-\theta_2} f + \cdots \\ \cdots + \int_{[\theta_1,\theta_2,\ldots,\theta_n]} D_{x-\theta_1} D_{x-\theta_2} \cdots D_{x-\theta_n} f,\end{aligned} \qquad \forall f \in C^{n-1}(\mathbb{R}^d), \tag{2.22}$$

where $D_y f$ denotes the derivative of f in the direction $y \in \mathbb{R}^d$. The Kergin operator has the property (which uniquely defines it) that

$$\mathcal{K}_\Theta(g \circ u^*) = (H_{u^*\Theta} g) \circ u^*, \qquad \forall u \in \mathbb{R}^d, \quad \forall g \in C^{n-1}(\mathbb{R}),$$

where $u^* : \mathbb{R}^d \to \mathbb{R} : x \mapsto \langle u, x\rangle := \sum_i u_i x_i$, and $H_{u^*\Theta}$ is the operator of *Hermite interpolation* at the multiset $u^*\Theta$ of points in $\mathbb{R}$ (see Micchelli [M80] for details).

The space of differential operators of order k which involve only derivatives in directions orthogonal to $\mathrm{aff}(\Theta)$ are defined as follows

$$D^k_{\mathrm{aff}(\Theta)^\perp} := \mathrm{span}\{D_{y_1} D_{y_2} \cdots D_{y_k} : y_1, \ldots, y_k \perp \mathrm{aff}(\Theta)\}.$$

The dimension of this space is

$$\dim(D^k_{\mathrm{aff}(\Theta)^\perp}) = \dim(\Pi^0_k(\mathrm{aff}(\Theta)^\perp) = \binom{k+d-\dim(\Theta)-1}{d-\dim(\Theta)-1}. \tag{2.23}$$

Let $L(\tilde\Theta)$ be the space of linear functionals

$$L(\tilde\Theta) := \mathrm{span}\{f \mapsto \int_{\tilde\Theta} q(D)f : q \in \Pi^0_{\#\tilde\Theta-1}(\mathbb{R}^d)\}, \qquad \tilde\Theta \subset \Theta,$$

where $L(\tilde\Theta) := \{0\}$ when $\tilde\Theta$ is empty. Note that $\Lambda_\Theta = \cup_{\tilde\Theta\subset\Theta} L(\tilde\Theta)$.

Lemma 2.24 (Micchelli [M80:Th.6]). *Suppose that $\lambda \in \Lambda_\Theta$ belongs to $L(\tilde\Theta)$, where $i = \#\tilde\Theta$. Then*

$$\lambda \in \Lambda_{\Psi,i} + \sum_{\substack{\Theta^* \subset \tilde\Theta \\ \#\Theta^* = i-1}} L(\Theta^*), \qquad \Psi := \mathrm{aff}(\tilde\Theta) \cap^* \Theta. \tag{2.25}$$

Theorem 2.6 (Interpolation version). *Let Θ be any multiset of n points in $\mathbb{R}^d$. The space of interpolation conditions for Kergin interpolation at Θ can be written as a direct sum*

$$\Lambda_\Theta = \bigoplus_{j=1}^{d} \bigoplus_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \bigoplus_{i=j}^{\#\Psi} \Lambda_{\Psi,i}, \tag{2.26}$$

where

$$\Lambda_{\Psi,i} := \mathrm{span}\{f \mapsto \int_{\tilde{\Theta}} q(D)f : q(D) \in D^{i-1}_{\mathrm{aff}(\Psi)^\perp},\ \tilde{\Theta} \in \mathcal{A}_i(\Psi)\}, \tag{2.27}$$

and

$$\dim(\Lambda_{\Psi,i}) = \binom{i-1+d-j}{d-j} \dim(\mathcal{S}_i(\Psi)), \qquad j-1 := \dim(\Psi). \tag{2.28}$$

Proof: It is clear that Λ_Θ, which is defined by (2.20), satisfies the inclusion

$$\sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \Lambda_{\Psi,i} \subset \Lambda_\Theta.$$

Hence it is sufficient to show that

$$\Lambda_\Theta \subset \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \Lambda_{\Psi,i}, \tag{2.29}$$

and the sum in (2.29) is direct.

Repeatedly applying (2.25) to $L(\tilde{\Theta})$ gives

$$L(\tilde{\Theta}) \subset \sum_{\Theta^* \subset \tilde{\Theta}} \Lambda_{\mathrm{aff}(\Theta^*)\cap^*\Theta, \#\Theta^*},$$

and so

$$\Lambda_\Theta = \sum_{\tilde{\Theta}\subset\Theta} L(\tilde{\Theta}) \subset \sum_{\tilde{\Theta}\subset\Theta} \sum_{\Theta^*\subset\tilde{\Theta}} \Lambda_{\mathrm{aff}(\Theta^*)\cap^*\Theta,\#\Theta^*} \subset \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \sum_{i=j}^{\#\Psi} \Lambda_{\Psi,i},$$

which is (2.29).

Now we show that the sum in (2.29) is direct, i.e.,

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

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 $\Lambda_{\Psi,i}$ with a space of functions $F_{\Psi,i}$. Let $\mathrm{aff}(\Psi)^\perp$ denote the orthogonal complement of $\mathrm{aff}(\Psi)$ in $\mathbb{R}^d$. Since $\mathbb{R}^d = \mathrm{aff}(\Psi)^\perp \oplus \mathrm{aff}(\Psi)$, the tensor product

$$F_{\Psi,i} := \Pi^0_{i-1}(\mathrm{aff}(\Psi)^\perp) \otimes \mathcal{S}_i(\Psi) := \mathrm{span}\{f = p \otimes s : p \in \Pi^0_{i-1}(\mathrm{aff}(\Psi)^\perp),\ s \in \mathcal{S}_i(\Psi)\}, \quad (2.30)$$

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 (2.2), that

$$\Lambda_{\Psi,i} = \mathrm{span}\{f \mapsto \int_{\mathrm{aff}(\Psi)} (p(D)f)s : p(D) \in D^{i-1}_{\mathrm{aff}(\Psi)^\perp},\ s \in \mathcal{S}_i(\Psi)\}. \quad (2.31)$$

Let $P_{\Psi,i} : \Lambda_{\Psi,i} \to F_{\Psi,i}$ 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 $\Lambda_{\Psi,i}$ and $F_{\Psi,i}$ are isomorphic, and in particular, by (2.23),

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

which is (2.28). In view these isomorphisms, it is sufficient to show that the sum

$$F_{\Psi_1,i_1} + F_{\Psi_2,i_2}, \qquad (\Psi_1,i_1) \neq (\Psi_2,i_2), \quad \mathrm{aff}(\Psi_1) \cap \mathrm{aff}(\Psi_2) \neq \phi, \quad (2.32)$$

is direct, i.e., that

$$f \in F_{\Psi_1,i_1} \cap F_{\Psi_2,i_2}, \quad \Longrightarrow \quad f = 0.$$

This we do in two cases. Note, by (2.30), that

$$\mathrm{supp}(f) = \mathrm{supp}(f|_{\mathrm{conv}(\Psi)}) + \mathrm{aff}(\Psi)^\perp, \qquad \forall f \in F_{\Psi,i}, \quad (2.33)$$

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

$$L \text{ is orthogonal to } \mathrm{aff}(\Psi) \quad \Longrightarrow \quad f|_L \in \Pi^0_{i-1}(L). \quad (2.34)$$

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 (2.34), the restriction of f to each line orthogonal to $\mathrm{aff}(\Psi_1)$ is a homogeneous polynomial of degree $i_1 - 1$ and $i_2 - 1$, 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 (2.33), 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

$$K + \mathrm{aff}(\Psi_1)^\perp = K + \mathrm{aff}(\Psi_2)^\perp. \tag{2.35}$$

If $f \neq 0$, then $K \neq \phi$ is compact, and (2.35) implies

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

which contradicts our assumption.

This completes the proof. □

Notice that $\Lambda_{\Psi,i}$ depends only on Ψ and i, and not on the other points in Θ. The combinatorial version is then obtained as follows.

Proof (Combinatorial version): Since the sum (2.26) is direct,

$$\dim(\Lambda_\Theta) = \sum_{j=1}^{d} \sum_{\substack{\Psi \subset \Theta \\ \dim(\Psi)=j-1 \\ \mathrm{aff}(\Psi) \cap {}^*\Theta = \Psi}} \sum_{i=j}^{\#\Psi} \dim(\Lambda_{\Psi,i}),$$

which by (2.21) and (2.28) gives (2.7).

The estimate (2.8) follows from the proof of Lemma 4.1 in Dahmen and Micchelli [DM83], where it is shown

$$\begin{aligned} \dim(\mathcal{S}_i(\Theta)) &= \dim \mathrm{span}\{x \mapsto \prod_{\theta \in \tilde{\Theta}} (1 + \theta^* x) : \tilde{\Theta} \subset \Theta,\ \#\tilde{\Theta} = n - i\} \\ &\leq \dim(\Pi_{n-i}(\mathbb{R}^s)) = \binom{n-i+s}{s}, \end{aligned}$$

together with the asserted equality.

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

The linear functionals in the direct sum decomposition (2.26) might well be termed the **reduced interpolation conditions** for Kergin interpolation, as the degrees of the derivatives occuring in them can not be decreased any further by the application of Lemma 2.24. We conclude this section with some examples of (2.26) which illustrate this.

Example 12 (Hermite interpolation)

Let Θ consist of the points $x_1, \ldots, x_r$ in $\mathbb{R}$ repeated with multiplicities $m_1, \ldots, m_r$. Then the Ψ in (2.26) consist of a point x_k repeated m_k times, giving

$$\Lambda_\Theta = \bigoplus_{k=1}^{r} \bigoplus_{i=1}^{m_k} \Lambda_{\underbrace{[x_k,\ldots,x_k]}_{m_k},i} = \bigoplus_{k=1}^{r} \bigoplus_{i=1}^{m_k} \operatorname{span}\{f \mapsto D^{i-1}f(x_k)\}, \tag{2.36}$$

which is the usual presentation of the interpolation conditions for Hermite interpolation. Using the Hermite–Genocchi formula, the presentation (2.20) is equivalent to

$$\Lambda_\Theta = \operatorname{span}\{f \mapsto [\theta_1, \ldots, \theta_r]f : [\theta_1, \ldots, \theta_r] \subset \Theta,\ r = 1, \ldots, n\},$$

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

Example 13 (Kergin interpolation at points in general position in $\mathbb{R}^2$)

Let Θ consist of points $x_1, x_2, \ldots, x_n$ which are in general position in $\mathbb{R}^2$. Then the Ψ in (2.26) consist of one or two points, giving the direct sum of 1–dimensional subspaces

$$\Lambda_\Theta = \left(\bigoplus_{k=1}^{n} \operatorname{span}\{f \mapsto f(x_k)\} \right) \oplus \left(\bigoplus_{1 \le s < t \le n} \operatorname{span}\{f \mapsto \int_{[x_s,x_t]} D_{(x_t - x_s)^\perp} f\} \right),$$

where $(x_t - x_s)^\perp$ is any vector orthogonal to the line through x_s and x_t. See Bos and Calvi [BC97] for an explicit representation of $\mathcal{K}_\Theta f$ in terms of these reduced interpolation conditions.

Example 14 (Kergin–Fejèr interpolation in $\mathbb{R}^2$)

The *Hermite–Fejèr interpolation* operator at distinct points $x_1, x_2, \ldots, x_n$ in $\mathbb{R}$ is a modified form of Hermite interpolation at these points repeated with multiplicity 2, where the function values at $x_1, \ldots, x_n$ are matched, but the remaining reduced interpolation conditions from (2.36) are taken to be zero (i.e., the derivative of the interpolant is zero at the points $x_1, \ldots, x_n$).

For any Θ, the linear projector $\mathcal{K}_\Theta : C^{n-1}(\mathbb{R}^d) \to \Pi_{n-1}(\mathbb{R}^d)$ can be written as

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

where $\mathcal{K}_{\Theta,\Psi,i}$ is the linear projector defined by the space of interpolants

$$\Pi_{\Theta,\Psi,i} := \{p \in \Pi_{n-1}(\mathbb{R}^d) : p \in \ker(\Lambda_{\Psi^*,i^*}),\ (\Psi^*,i^*) \neq (\Psi,i)\},$$

and the space of interpolation conditions $\Lambda_{\Psi,i}$. This 'direct sum' satisfies

$$\begin{aligned}
&\mathcal{K}_{\Theta,\Psi_1,i_1}\mathcal{K}_{\Theta,\Psi_2,i_2} = 0, \quad (\Psi_1,i_1) \neq (\Psi_2,i_2),\\
&(\mathcal{K}_{\Theta,\Psi,i})^2 = \mathcal{K}_{\Theta,\Psi,i},\\
&\mathcal{K}_{\Theta,\Psi,i}(p) = p, \quad \forall p \in \Pi_{\Theta,\Psi,i}.
\end{aligned}$$

The polynomial spaces $\Pi_{\Theta,\Psi,i}$ are dual to the $\Lambda_{\Psi,i}$, and the dual of (2.26) is

$$\Pi_{n-1}(\mathbb{R}^d) = \bigoplus_{j=1}^{d} \bigoplus_{\substack{\Psi\subset\Theta\\ \dim(\Psi)=j-1\\ \mathrm{aff}(\Psi)\cap^*\Theta=\Psi}} \bigoplus_{i=j}^{\#\Psi} \Pi_{\Theta,\Psi,i}. \tag{2.37}$$

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

$$\mathcal{F}_\Theta := \bigoplus_{k=1}^{n} \mathcal{K}_{\Theta,[x_k,x_k],1}, \qquad \Theta := \{x_1,x_1,x_2,x_2,\ldots,x_n,x_n\}. \tag{2.38}$$

By analogy, we define **Kergin–Fejèr interpolation** in $\mathbb{R}^n$ by (2.38), when $x_1,\ldots,x_n$ are distinct points in $\mathbb{R}^n$. This operator has a continuous extension to functions defined at the points $x_1,\ldots,x_n$, and maps into an n–dimensional subspace of $\Pi_{2n-1}(\mathbb{R}^d)$. In order to compute this operator it is necessary to find the reduced interpolation conditions for Kergin interpolation at Θ. We now use (2.26) to do this for the case of points in general position in $\mathbb{R}^2$.

The Ψ in the direct sum 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 $\dim(\Pi_{2n-1}(\mathbb{R}^2)) = \binom{2n+1}{2}$ linearly independent conditions.

Example 15 (Kergin–Fejèr interpolation in $\mathbb{R}^3$)

Let $\{x_1, x_2, \ldots, x_n\}$ be n points in general position in $\mathbb{R}^3$, and take

$$\Theta := \{x_1, x_1, x_2, x_2, \ldots, x_n, x_n\}.$$

As in the previous example, we will determine the reduced interpolation conditions for Kergin interpolation at Θ, which are required to for finding the Kergin–Fejèr interpolation operator $\mathcal{F}_\Theta$.

The Ψ in (2.26) are of dimensions $0, 1, 2$, and are of the forms

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

respectively (where r, s, t 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

$$\left.\begin{array}{ll} f(x_r), \quad \nabla f(x_r), & 4n \text{ conditions,} \end{array}\right\} \; j = 1$$

$$\left.\begin{array}{ll} \displaystyle\int_{[x_r, x_s]} D_\perp f, & 2\binom{n}{2} \text{ conditions,} \\ \displaystyle\int_{[x_r, x_r, x_s]} D_\perp^2 f, & 6\binom{n}{2} \text{ conditions,} \\ \displaystyle\int_{[x_r, x_r, x_s, x_s]} D_\perp^3 f, & 4\binom{n}{2} \text{ conditions,} \end{array}\right\} \; j = 2$$

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

for a total of $\dim(\Pi_{2n-1}(\mathbb{R}^3)) = \binom{2n+2}{3}$ linearly independent conditions.

The author has not yet investigated to what extent the convergence properties (for large numbers of nodes) of Kergin–Fejèr interpolation and Kergin interpolation in $\mathbb{R}^2$, $\mathbb{R}^3$ are similar those of the well understood case of interpolation in $\mathbb{R}$ (Hermite–Fejèr and Lagrange interpolation, respectively).

Example 16 (Interpolation conditions for points on a circle in $\mathbb{R}^d$)

Consider Example 10 once more. For the Ψ of sizes $j = 1$ and $j = 2$, the corresponding interpolation conditions are

$$\left.\begin{array}{ll} f(\theta_r), & \binom{n}{1} \text{ conditions,} \end{array}\right\} \; j = 1$$

$$\left.\begin{array}{ll} \displaystyle\int_{[\theta_r,\theta_t]} D_{(\theta_t-\theta_r)^\perp} f, & \binom{n}{2}(d-1) \text{ conditions,} \end{array}\right\} \; j = 2$$

where $\theta_r, \theta_t \in \Theta$, $\theta_r \neq \theta_t$. The only other Ψ occuring is $\Psi = \Theta$, for $j = 3$. When Ψ consists of n points in general position in an s–dimensional subspace of $\mathbb{R}^d$, it is possible to give an explicit choice of the 'different allowable submultisets' $\mathcal{B}_i(\Psi)$ of (2.15). This *s–generating sequence* of [DM83:Th.4.2] is defined as follows:

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

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

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

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

Example 17 (Extending Kergin interpolation to less smooth functions)

Using (2.19) the decompositon of Λ_Θ can be used to extend $\mathcal{K}_\Theta$ to a continuous linear map $C^{m-1}(\mathbb{R}^d) \to \Pi_{n-1}(\mathbb{R}^d)$, where

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

For example, if the points in Θ are in general position, then $m = d$ when $n = \#\Theta > d$.

3. Computational considerations

In this final section, we briefly outline the numerical implementation of finding the subspaces $\Psi_{\Psi,i}$, their duals $\Pi_{\Theta,\Psi,i}$, and the related polynomial interpolants $\mathcal{K}_{\Theta,\Psi,i} f$.

Identification of the Ψ

The first step is to identify the submultisets Ψ in the direct sum decomposition (2.26), 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

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

Solving any 3×3 linear system obtained from (2.19) by the choice of a basis for linear polynomials $\Pi_1(\mathbb{R}^2)$, gives the Kergin interpolant

$$\begin{aligned}\mathcal{K}_{[\theta_1,\theta_2]} f(x) = f(\theta_1)\frac{(\theta_1-\theta_2)^T(x-\theta_2)}{\|\theta_2-\theta_1\|^2} + f(\theta_2)\frac{(\theta_2-\theta_1)^T(x-\theta_1)}{\|\theta_2-\theta_1\|^2} \\ + \int_{[\theta_1,\theta_2]} D_{(\theta_2-\theta_1)^\perp} f \, \frac{(\theta_2-\theta_1)^{\perp T}\,(2x-(\theta_1+\theta_2))}{2\|(\theta_2-\theta_1)^\perp\|^2}.\end{aligned} \tag{3.1}$$

There is a loss of significance in the computed values of each of the terms in (3.1) 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 $< \texttt{tol}$, 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.,

$$\sum_k a_k \lambda(p_k) = \lambda(p) = \lambda(f), \tag{3.2}$$

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 (2.27), 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 (2.31) 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 $\{M(\cdot|\tilde{\Theta}) : \tilde{\Theta} \in \mathcal{A}_i(\Theta)\}$, which in general leads to a system (3.2) 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 $\{x_1, \ldots, x_n\}$, which requires a basis for the n–dimensional space

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

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

Acknowledgements

Thanks to Thomas Grandine (Seattle) for providing the data for the plots of simplex splines in Figures 2.1 and 2.2 from his implementation of [G86], and to Sangnam Nam (Madision) for a number of very helpful comments.

References

[BHS93] B. D. Bojanov, H. A. Hakopian, and A. A. Sahakian, "Spline Functions and Multivariate Interpolations", Kluwer Academic Publishers, Dordrecht, The Netherlands, 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.

[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.

[GO97] J. E. Goodman and J. O'Rourke, "Handbook of Discrete and Computational Geometry", CRC Press, Boca Raton, 1997.

[G86] T. A. Grandine, The stable evaluation of multivariate simplex splines, *Math. Comp.* **50** (1988), 197–205.

[H86] K. Höllig, Multivariate splines, *in* "Approximation Theory, Proc. Symp. Appl. Math. **36**" (C. de Boor Ed.), pp. 103–127, Amer.Math.Soc., Providence, 1986.

[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.