TECHNION -- ISRAEL INSTITUTE OF TECHNOLOGY
DEPARTMENT OF MATHEMATICS

# Multipoint Taylor formulæ

Shayne Waldron* (waldron@math.auckland.ac.nz)

Technical Summary Report
March 1996

## ABSTRACT

The main result of this paper is an abstract version of the *Kowalewski – Ciarlet – Wagschal multipoint Taylor formula* for representing the pointwise error in multivariate Lagrange interpolation. Several applications of this result are given in the paper. The most important of these is the construction of a multipoint Taylor error formula for a *general finite element*, together with the corresponding $L_p$–error bounds. Another application is the construction of a family of error formulæ for linear interpolation (indexed by real measures of unit mass) which includes some recently obtained formulæ. It is also shown how the problem of constructing an error formula for Lagrange interpolation from a $D$–invariant space of polynomials with the property that it involves only derivatives which annihilate the interpolating space can be reduced to the problem of finding such a formula for a 'simpler' one–point interpolation map.

**AMS (MOS) Subject Classifications:** primary 41A65, 41A80, 65D05, secondary 41A05, 41A10, 41A44, 41A55

**Key Words:** multipoint Taylor formula, Kowalewski's remainder, Lagrange interpolation, linear interpolation on a triangle, finite elements, multivariate Hardy's inequality, Taylor interpolation, one–point interpolation, least solution

---

* Partially supported by the Steenbock Professorship of Carl de Boor and the Israel Council for Higher Education.

# 1. Introduction

### Overview

This paper is concerned with *explicit* representations of the linear map $\mathcal{Q}$ which is defined *implicitly* as the unique map which makes the following diagram commute

$$\begin{tikzcd} X \arrow[rr, "\mathcal{U}"] \arrow[dr, "\mathcal{R}"'] & & \operatorname{ran}\mathcal{U} \arrow[dl, dotted, "\mathcal{Q}"] & \subset Y \\ & Z & & \end{tikzcd}$$

where $\mathcal{U} : X \to Y$, $\mathcal{R} : X \to Z$ are linear maps with

$$\ker \mathcal{U} \subset \ker \mathcal{R}.$$

Maps $\mathcal{Q}$ of this type occur frequently in approximation theory and numerical analysis. For example, if $\mathcal{U} = D^{k+1}$ and $\mathcal{R}$ is a linear functional whose kernel ($\ker \mathcal{R}$) contains $\Pi_k$ (the polynomials of degree $\leq k$), then $\mathcal{Q}$ gives the *Peano kernel representation*

$$\mathcal{R}f = \mathcal{Q}(D^{k+1}f).$$

There is particular interest in certain such maps $\mathcal{Q}$ whose norm provides the best constant in error bounds for various numerical schemes. The most notable of these are the *finite element method* and *numerical differentiation and integration rules* (see, e.g., Meinguet [M84] and the references therein).

The main result of this paper is that for certain maps of the form

$$\mathcal{R} = \mathcal{A}(1 - \mathcal{L})$$

part of the flexibility in Meinguet's abstract method for representing $\mathcal{Q}$ (detailed below) can be used to satisfy conditions that ensure the representation for $\mathcal{Q}$ take a simple form. A precise statement of this result is given in Theorem 2.2.

The prototype of such representations is the *multipoint Taylor formula* for Lagrange interpolation (see Ciarlet and Wagschal [CW71]). In Section 3, this is discussed and then extended (in its unexpanded form) to *all* linear operators that reproduce $\Pi_k$.

In Section 4, this extension is discussed within the setting of *finite elements.* It is expanded in terms of the *nodal variables* to obtain the analogue of the original presentation of the multipoint Taylor formula.

In Section 5, $L_p$–error bounds for finite elements are obtained from these 'expanded' multipoint Taylor formulæ by using the *multivariate Hardy's inequality* recently given by the author (and intended for precisely this type of situation).

In Section 6, to emphasise the wide applicability of the main 'abstract result' and to illustrate some of its more subtle points, it is used to obtain a family of formulæ for the error in linear interpolation which is indexed by real measures of unit mass. These include many of the known formulæ (some of which have only recently been obtained).

Finally, in Section 7, a general form of Taylor interpolation is introduced which allows the results of the previous sections to be extended. This 'one–point interpolation from an arbitrary polynomial space' has close connections with the *least solution* of de Boor and Ron [BR92] to the 'polynomial interpolation problem'. It is shown how error formulæ for the least solution can be obtained from those for a simpler one–point interpolation problem. Representing the error in both of these maps involves some deep questions in the theory of multivariate Peano kernels.

### Meinguet's abstract method for representing $\mathcal{Q}$

Given the situation above, i.e.,

$$\mathcal{Q}\mathcal{U} = \mathcal{R} \tag{1.1}$$

with $\mathcal{R}, \mathcal{U}$ known linear maps, how can an *explicit* representation of $\mathcal{Q}$ be obtained?

The approach most often used is to take

$$\mathcal{V} : \operatorname{ran} \mathcal{U} \to X$$

a *right inverse* for $\mathcal{U}$, i.e., a map with

$$\mathcal{U}\mathcal{V} = 1,$$

and right multiply (1.1) by $\mathcal{V}$ to obtain

$$\mathcal{Q} = \mathcal{R}\mathcal{V}. \tag{1.2}$$

This is *stage* (i) of what is referred to in Meinguet [M84] as the *abstract method* for representing $\mathcal{Q}$. The prototypical example of (1.2) is the Peano kernel representation.

*Stage* (ii) of the abstract method is the observation that if the restriction of $\mathcal{R}$ to the range of $\mathcal{V}$ can be written as a finite sum

$$\mathcal{R}|_{\operatorname{ran} \mathcal{V}} = \sum_j \mathcal{S}_j,$$

of what are referred to as *standard* linear mappings $\mathcal{S}_j$, then (1.2) can be expanded as

$$\mathcal{Q} = \sum_j \mathcal{S}_j \mathcal{V}. \tag{1.3}$$

Now assume that $\mathcal{V}$ is linear and $Y$, $Z$ are normed linear spaces. In the final stage of the abstract method, *stage* (iii), the triangle inequality is applied to (1.3) to obtain

$$\|\mathcal{Q}\| \leq \sum_j \|\mathcal{S}_j \mathcal{V}\|. \tag{1.4}$$

If $\mathcal{R}$,$\mathcal{U}$ are continuous and $\mathcal{U}$ is open, then $\mathcal{Q}$ is a continuous linear map. This was first observed by Sard [S48:Th.1] when $X, Y, Z$ are Banach spaces, and is referred to as *Sard's factorisation theorem* by Atteia [At92:Th.2.1,p98] who states it for topological vector spaces.

Meinguet does not suggest how best to choose $\mathcal{V}$ and $\{\mathcal{S}_j\}$, other than that it should be done in such a way that bounding each $\|\mathcal{S}_j\mathcal{V}\|$ is simpler than bounding $\|\mathcal{R}\mathcal{V}\|$ directly. He gives many examples in [M75], ..., [M84]. Earlier examples of the abstract method in the multivariate setting include Èzrohi [E57] and Sard [S63].

In each case known to the author the right inverse $\mathcal{V}$ is *linear* and comes via an associated linear projector $\mathcal{P}$ onto $\ker\mathcal{U}$ as described below.

**Proposition 1.5 (cf [M84]).** *Let $\mathcal{U} : X \to Y$ be a linear map. There is a 1–1 (linear) correspondence between linear right inverses*

$$\mathcal{V} : \operatorname{ran}\mathcal{U} \to X$$

*for $\mathcal{U}$ and linear projectors*

$$\mathcal{P} : X \to \ker\mathcal{U}$$

*onto* $\ker\mathcal{U}$*, which is given by*

$$\mathcal{P} = 1 - \mathcal{V}\mathcal{U}.$$

## 2. The main result

Next is the main result, that if

$$\mathcal{R} = \mathcal{A}(1 - \mathcal{L}), \tag{2.1}$$

then it may be possible to choose $\mathcal{V}$ so that the representation of $\mathcal{Q}$ given by Meinguet's *abstract method* takes a simple form. Later it will be shown how many practical estimates follow from this 'abstract result' for operators of the form (2.1), such as the *error* in the *finite element method* (see Examples 1 and 2 of Section 4). The reader might like to keep in mind the simple example

$$\mathcal{R}f = \delta_x(1 - L)f = f(x) - Lf(x)$$

where $L$ is the operator of Lagrange interpolation from $\Pi_k$ and $\mathcal{U} = D^{k+1}$, $\ker\mathcal{U} = \Pi_k$.

**Theorem 2.2.** *Let $\mathcal{Q}$ be the linear map which is defined implicitly as the unique map that makes the following diagram commute*

$$\begin{tikzcd} X \arrow[r, "\mathcal{U}"] \arrow[dr, "\mathcal{R}:=\mathcal{A}(1-\mathcal{L})"'] & \operatorname{ran}\mathcal{U} \arrow[d, dotted, "\mathcal{Q}"] & \subset Y \\ & Z & \end{tikzcd}$$

*where* $\mathcal{U} : X \to Y$, $\mathcal{A} : X \to Z$, $\mathcal{L} : X \to X$ *are linear maps with*

$$\ker \mathcal{U} \subset \ker \mathcal{R}.$$

*If* $\mathcal{V} : \operatorname{ran} \mathcal{U} \to X$ *is a right inverse of* $\mathcal{U}$*, chosen so that*

$$\mathcal{AVU} = 0, \tag{2.3}$$

*then*

$$\mathcal{Q} = -\mathcal{ALV}. \tag{2.4}$$

**Proof:** Expanding (1.2) using $\mathcal{R} = \mathcal{A}(1 - \mathcal{L})$ gives

$$\mathcal{Q} = \mathcal{AV} - \mathcal{ALV}. \tag{2.5}$$

If $\mathcal{AVU} = 0$ (which is equivalent to $\mathcal{AV} = 0$), then (2.5) reduces to (2.4). □

Above, the condition $\mathcal{AVU} = 0$ was favoured over the equivalent one that $\mathcal{AV} = 0$. This is because usually the right inverse $\mathcal{V}$ is a *linear* map which comes via an associated linear projector $\mathcal{P} := 1 - \mathcal{VU}$ (see Proposition 1.5), and this choice makes it clear that (2.3) can be rewritten in the (often more convenient) form

$$\mathcal{A} = \mathcal{AP}. \tag{2.6}$$

At times, such as when $\mathcal{U}$ is the *multivariate* differentiation operator $D^{k+1}$, it is convenient to use Theorem 2.2 without any direct reference to $\mathcal{U}$ or $\mathcal{V}$. The following corollary allows this.

**Corollary 2.7.** *Let*

$$\mathcal{R} := \mathcal{A}(1 - \mathcal{L}) : X \to Z$$

*where* $\mathcal{A} : X \to Z$, $\mathcal{L} : X \to X$ *are linear maps. If* $\mathcal{P} : X \to X$ *is a linear projector with*

$$\operatorname{ran} \mathcal{P} \subset \ker \mathcal{R}, \tag{2.8}$$

*and*

$$\mathcal{A} = \mathcal{AP}, \tag{2.9}$$

*then*

$$\mathcal{R} = -\mathcal{AL}(1 - \mathcal{P}). \tag{2.10}$$

**Proof:** This corollary follows by taking $\mathcal{U} = 1 - \mathcal{P}$ and choosing $\mathcal{V}$ to be the identity. However, it may be more instructive to give the direct proof that:

$$\begin{aligned}
\mathcal{R} &= \mathcal{R}(1 - \mathcal{P}) \qquad (\text{since } \operatorname{ran} \mathcal{P} \subset \ker \mathcal{R}) \\
&= \mathcal{A}(1 - \mathcal{L})(1 - \mathcal{P}) = \mathcal{A}(1 - \mathcal{P}) - \mathcal{AL}(1 - \mathcal{P}) \\
&= -\mathcal{AL}(1 - \mathcal{P}) \qquad (\text{since } \mathcal{A} = \mathcal{AP}).
\end{aligned}$$

□

## Examples

As far as the author is aware, there are two examples of (what are effectively) representations of $\mathcal{Q}$ of the form (2.4). One is the multipoint Taylor formula for the error in Lagrange interpolation, and certain kinds of Hermite interpolation, from a space containing $\Pi_k$. This is discussed and then extended in the next section. The other is Gregory's error formula for linear interpolation on a (standard) triangle (see next section).

# 3. Multipoint Taylor formulæ for linear operators that reproduce $\Pi_k$

In this section it is shown that the *multipoint Taylor formula* for the error in a Lagrange map (see below) is a natural example of equation (2.4) of Theorem 2.2, and it is extended (in its unexpanded form) to all linear operators that reproduce $\Pi_k$.

## Lagrange maps

We say that *Lagrange interpolation* from a space of functions $P$ to data at a finite set of points $\Theta \subset \mathbb{R}^n$ is *correct* if for each function $f$ (defined at least on $\Theta$) there is a unique $p \in P$ with

$$p(v) = f(v), \quad \forall v \in \Theta.$$

The associated linear projector $f \mapsto p$ is called the **Lagrange map** (given by the space of interpolants $P$ and points of interpolation $\Theta$), and it is denoted by

$$\mathcal{L}_{P,\Theta} : f \mapsto p.$$

The *Lagrange form* of a Lagrange map $\mathcal{L}_{P,\Theta}$ is

$$\mathcal{L}_{P,\Theta} f = \sum_{v \in \Theta} f(v)\, \ell_v, \tag{3.1}$$

where (3.1) uniquely defines

$$\ell_v := \ell_{v,P,\Theta} \in P,$$

the *Lagrange function* for $v \in \Theta$.

Lagrange maps with a space of interpolants that contains $\Pi_k$ give rise to many commonly used finite elements, and are also frequently used to interpolate to scattered data. Particular examples of the latter, which have received much attention lately, are maps where the interpolants include *radial basis functions*, and the *least solution* to the polynomial interpolation problem (see Section 7).

## The multipoint Taylor formula

The multipoint Taylor formula is a representation of the error

$$\mathcal{R} := \mathcal{A}(1 - \mathcal{L})$$

in a Lagrange map

$$\mathcal{L} := \mathcal{L}_{P,\Theta}$$

for which $\Pi_k \subset P$, and where $\mathcal{A}$ is any of the derivatives

$$\mathcal{A} := \mathcal{A}_{g,x} := \delta_x g(D) : f \mapsto g(D)f(x), \quad g \in \Pi_k.$$

Here, and whenever *point evaluation* at $x$ is being thought of as a linear map on functions (and so should be applied from the left) it will be denoted by

$$\delta_x : f \mapsto f(x).$$

The differential operator induced by $g \in \Pi_k$ is written $g(D)$.

If $\mathcal{P}$ is a linear projector onto $\Pi_k$ which interpolates $\mathcal{A}$, i.e., for which

$$g(D)f(x) = g(D)\mathcal{P}f(x), \quad \forall f, \tag{3.2}$$

then by Corollary 2.7

$$\mathcal{R} = -\mathcal{A}\mathcal{L}(1 - \mathcal{P}). \tag{3.3}$$

Here, the intent is to apply Theorem 2.2 with $\mathcal{U} = D^{k+1}$ (the $k+1$ order multivariate differentiation operator) to obtain expansions of $\mathcal{R}(f)$ in terms of $D^{k+1}f$.

The linear functionals

$$\{\mathcal{A}_{g,x} : g \in \Pi_k\}$$

are the interpolation conditions for the operator of *Taylor interpolation* from $\Pi_k$ at $x$, which we denote $\mathcal{T}_{k,x}$. This operator is a natural choice for $\mathcal{P}$ since it is the *only* linear projector (onto $\Pi_k$) that interpolates $\mathcal{A} = \mathcal{A}_{g,x}$ for all $g \in \Pi_k$. For this choice the equation (3.3) becomes

$$g(D)(f - \mathcal{L}_{P,\Theta}f)(x) = -g(D)(\mathcal{L}_{P,\Theta}\mathcal{R}_{k,x}f)(x), \quad \forall f, \tag{3.4}$$

where

$$\mathcal{R}_{k,x} := 1 - \mathcal{T}_{k,x}$$

is the *error* (remainder) in Taylor interpolation from $\Pi_k$ at $x$, which can be expressed as

$$\mathcal{R}_{k,x}f(v) = \frac{1}{k!}\int_0^1 (1-t)^k D_{v-x}^{k+1} f(x + t(v-x))\, dt. \tag{3.5}$$

Here $D_y f$ denotes the derivative of $f$ in the direction $y$. For a given linear functional $\mathcal{A}$, such as point evaluation at $x$, there are (for $k > 0$) many linear projectors satisfying (3.2). This is illustrated in Section 6 for linear interpolation, i.e., $P = \Pi_1$.

The careful reader will notice that the only property of $\mathcal{L}_{P,\Theta}$ used so far is that it is a linear operator which reproduces $\Pi_k$, and so, in particular, (3.4) holds with $\mathcal{L}_{P,\Theta}$ replaced by any such operator. This is a crucial observation, which leads to Theorem 3.15 and is discussed further in this section.

By expressing the Lagrange interpolant $\mathcal{L}_{P,\Theta}(\mathcal{R}_{k,x}f)$ in Lagrange form (3.1) the equation (3.4) can be expanded as

$$g(D)(f - \mathcal{L}_{P,\Theta}f)(x) = -\sum_{v\in\Theta} \mathcal{R}_{k,x}f(v)\,(g(D)\ell_v)(x). \tag{3.6}$$

This formula is referred to by Meinguet in [M84:p98] as the *Kowalewski – Ciarlet – Wagschal multipoint Taylor formula* out of deference to its origins which are as follows.

### History of the (Kowalewski – Ciarlet – Wagschal) multipoint Taylor formula

The earliest occurrence of (3.6) is **Kowalewski's remainder** (see [K32:p21-24], or Davis's book [D75:p71]), which is the *univariate* case when the linear functional $\mathcal{A}$ is point evaluation at $x$.

In Ciarlet and Wagschal [CW71] the *multivariate* form of (3.6) was given for the special cases of interpolation by linear polynomials (Theorem 2) and by quadratic polynomials for derivatives $\mathcal{A} = \delta_x g(D)$ up to order 1 (Theorem 4). The term *multipoint Taylor formula* was coined in that paper. The subsequent paper of Ciarlet and Raviart [CR72] proved the general formula (3.6). A discussion of the relationship of that proof with Theorem 2.2 and of the choice of the term 'multipoint Taylor formula' is given below.

The special case of formula (3.6) where $\mathcal{A}$ is point evaluation at $x$ and $\Theta$ consists of the *simplex points* ($\dim \Pi_k$ equally spaced points on a simplex) was proved independently by Nicolaides in [N72:Th.5.1] and [N73:Th.2.1].

### The Ciarlet–Raviart proof of the multipoint Taylor formula

To understand the proof of (3.6) given in [CR72] it is instructive to consider the following alternative simple proof found by the author (in trying to do so), which invited the generalisation to Theorem 2.2.

Since

$$f = \mathcal{T}_{k,x}f + \mathcal{R}_{k,x}f,$$

and $\mathcal{L}_{P,\Theta}$ *reproduces* $\Pi_k$, i.e.,

$$\mathcal{L}_{P,\Theta}g = g, \qquad \forall g \in \Pi_k,$$

it follows that

$$\mathcal{L}_{P,\Theta}f = \mathcal{L}_{P,\Theta}\mathcal{T}_{k,x}f + \mathcal{L}_{P,\Theta}\mathcal{R}_{k,x}f = \mathcal{T}_{k,x}f + \mathcal{L}_{P,\Theta}\mathcal{R}_{k,x}f. \tag{3.7}$$

In particular, since $\mathcal{T}_{k,x}f$ is the Taylor interpolant from $\Pi_k$ to $f$ at $x$, applying $\delta_x g(D)$ to each side of (3.7) gives

$$g(D)\mathcal{L}_{P,\Theta}f(x) = g(D)f(x) + (g(D)\mathcal{L}_{P,\Theta}\mathcal{R}_{k,x}f)(x), \tag{3.8}$$

which (up to a rearrangement) is (3.4) and can be expanded to (3.6). Again, the careful reader will observe that (3.7) holds with $\mathcal{L}_{P,\Theta}$ replaced by any linear operator that reproduces $\Pi_k$.

The proof of (3.6) given in [CR72] uses the $m$–th order differentiation operator, which (for simplicity) can be thought of as the map

$$D^m : f \mapsto (D^\alpha f : |\alpha| = m)$$

to the sequence of partial derivatives of order $m$. There the Taylor formula

$$\begin{aligned} f(v) &= \mathcal{T}_{k,x}f(v) + \mathcal{R}_{k,x}f(v) \\ &= f(x) + D_{v-x}f(x) + \cdots + \frac{1}{k!}D^k_{v-x}f(x) + \mathcal{R}_{k,x}f(v) \end{aligned} \tag{3.9}$$

is used to expand the Lagrange form of $\mathcal{L}_{P,\Theta}$, and this expansion is differentiated to obtain

$$D^m\mathcal{L}_{P,\Theta}f(x) = \sum_{l=0}^{k}\frac{1}{l!}\sum_{v\in\Theta}D^l_{v-x}f(x)\,(D^m\ell_v)(x) + \sum_{v\in\Theta}\mathcal{R}_{k,x}f(v)\,(D^m\ell_v)(x). \tag{3.10}$$

It is then proved that the double sum in (3.10) reduces to

$$\sum_{l=0}^{k}\frac{1}{l!}\sum_{v\in\Theta}D^l_{v-x}f(x)\,(D^m\ell_v)(x) = D^m f(x), \quad 0 \le m \le k, \tag{3.11}$$

thereby proving (3.6), where $D^m$, $0 \le m \le k$ plays the role of $g(D)$, $g \in \Pi_k$.

With hindsight, since

$$\sum_{l=0}^{k}\frac{1}{l!}D^l_{v-x}f(x) = \mathcal{T}_{k,x}f(v),$$

it is seen that (3.11) is an expansion into Lagrange form of

$$D^m(\mathcal{L}_{P,\Theta}\mathcal{T}_{k,x}f)(x) = D^m f(x), \qquad 0 \le m \le k,$$

an identity which holds because

$$\mathcal{L}_{P,\Theta}(\mathcal{T}_{k,x}f) = \mathcal{T}_{k,x}f,$$

and not because of the specific nature of the linear operator $\mathcal{L}_{P,\Theta}$. Instead, the proof of (3.11) given in [CR72] involves showing (by induction) what is described as 'the (somehow unexpected) result' that:

$$\frac{1}{l!}\sum_{v\in\Theta}D^l_{v-x}f(x)\,(D^m\ell_v)(x) = \begin{cases} 0, & 0 \le l < m; \\ D^m f(x), & l = m; \\ 0, & m < l \le k. \end{cases} \tag{3.12}$$

**Gregory's error formula for linear interpolation on a triangle**

For the multivariate case, in addition to (3.5), there are many expansions for the error in Taylor interpolation $\mathcal{R}_{k,x}f(v)$ in terms of $D^{k+1}f$ that could be substituted into (3.6), each giving different sums of *standard* linear mappings (1.3) (and thus leading to possibly better norm bounds).

Gregory's well known formulæ for the error in linear interpolation at

$$\Theta := \{(1,0),\ (0,1),\ (0,0)\}$$

is an example of this. In Gregory [G75] the formulæ (3.6) for remainders

$$f(x) - \mathcal{L}_{\Pi_1,\Theta}f(x), \qquad D^{(1,0)}(f - \mathcal{L}_{\Pi_1,\Theta}f)(x), \qquad D^{(0,1)}(f - \mathcal{L}_{\Pi_1,\Theta}f)(x) \tag{3.13}$$

are expanded using a formula for $\mathcal{R}_{1,x}f(v)$ different from (3.5). To get an appreciation for the detail involved, the formula used, which is in the spirit of Sard [S63:Th.8,p163], is that for $x, v \in T$ (the triangle with vertices $\Theta$)

$$\begin{aligned}
\mathcal{R}_{1,x}f(v) = \chi_{A_1}(v)\Big[&\int_{x_1}^{v_1}(v_1-s)D^{(2,0)}f(s,x_2)\,ds + \int_{x_2}^{v_2}\int_{x_1}^{v_1} D^{(1,1)}f(s,t)\,ds\,dt \\
&+ \int_{x_2}^{v_2}(v_2-t)D^{(0,2)}f(x_1,t)\,dt\Big] \\
+ \chi_{A_2}(v)\Big[&\int_{x_1}^{v_1}\int_{x_1}^{t} D^{(2,0)}f(s,1-t)\,ds\,dt + \int_{x_1}^{v_1}\int_{x_2}^{1-s} D^{(1,1)}f(x_1,t)\,dt\,ds \\
&+ \int_{x_1}^{v_1}\int_{1-s}^{v_2} D^{(1,1)}f(s,t)\,dt\,ds + \int_{x_2}^{v_2}(v_2-t)D^{(0,2)}f(x_1,t)\,dt\Big] \\
+ \chi_{A_3}(v)\Big[&\int_{x_1}^{v_1}(v_1-s)D^{(2,0)}f(s,x_2)\,ds + \int_{x_2}^{v_2}\int_{1-t}^{v_1} D^{(1,1)}f(s,t)\,ds\,dt \\
&+ \int_{x_2}^{v_2}\int_{x_1}^{1-t} D^{(1,1)}f(s,x_2)\,ds\,dt + \int_{x_2}^{v_2}\int_{x_2}^{s} D^{(0,2)}f(1-s,t)\,dt\,ds\Big],
\end{aligned} \tag{3.14}$$

where $\chi_{A_i}$ is the *characteristic function* of the set $A_i$ which is defined by

$$A_i := \begin{cases} T \setminus (A_2 \cup A_3), & i = 1; \\ \{v \in\ T : v_1 > 1 - x_2\}, & i = 2; \\ \{v \in\ T : v_2 > 1 - x_1\}, & i = 3. \end{cases}$$

Naturally, the formulæ for the remainders (3.13) so obtained are of a similar complexity to (3.14). Gregory was successful in obtaining useful $L_p$–error bounds from these formulæ.

**The (unexpanded) multipoint Taylor formula**
**for linear operators that reproduce $\Pi_k$**

We end this section with the general result referred to previously.

**Theorem 3.15.** *Suppose that $\mathcal{L}$ is a linear operator that reproduces $\Pi_k$. Then, for sufficiently smooth $f$*

$$\mathcal{L}f = \mathcal{T}_{k,x}f + \mathcal{L}\mathcal{R}_{k,x}f. \tag{3.16}$$

*In particular, for $g \in \Pi_k$*

$$g(D)(f - \mathcal{L}f)(x) = \mathcal{Q}(D^{k+1}f), \tag{3.17}$$

*where*

$$\mathcal{Q}(D^{k+1}f) := -g(D)(\mathcal{L}\mathcal{R}_{k,x}f)(x). \tag{3.18}$$

**Proof:** The proof is exactly as for (3.7) and (3.8). Since $\mathcal{T}_{k,x}f \in \Pi_k$ and $\mathcal{L}$ reproduces $\Pi_k$,

$$\mathcal{L}f = \mathcal{L}\mathcal{T}_{k,x}f + \mathcal{L}\mathcal{R}_{k,x}f = \mathcal{T}_{k,x}f + \mathcal{L}\mathcal{R}_{k,x}f.$$

Apply $\delta_x g(D)$ to each side of this, to obtain (after rearrangement) that

$$g(D)(f - \mathcal{L}f)(x) = -g(D)(\mathcal{L}\mathcal{R}_{k,x}f)(x).$$

From (3.5) it follows that $-g(D)(\mathcal{L}\mathcal{R}_{k,x}f)(x)$ depends only on $D^{k+1}f$, as is claimed by the definition of (3.18). □

### Terminology

The term *multipoint Taylor formula* for (3.6) was adopted in [CW71], not because it involves the error $\mathcal{R}_{k,x}f(v)$ for *many* points $v \in \Theta$, but rather because (3.6) can be rewritten as

$$g(D)f(x) = \sum_{v\in\Theta} f(v)\,(g(D)\ell_v)(x) + \mathrm{error}(D^{k+1}f),$$

which expresses the derivative $g(D)f(x)$ in terms of the value of $f$ at the *multi*ple *point*s $v \in \Theta$ plus an error term. Since a similar phenomenon will be observed for the linear projectors $\mathcal{L}$ considered in the next section, equation (3.17) will be referred to as the **unexpanded multipoint Taylor formula** for $\mathcal{L}$. The adjective *unexpanded* is used since, as we have just seen, the error $\mathcal{R}_{k,x}f$ occurring in (3.18) can be expanded in many ways (as can $\mathcal{L}$ itself). These expansions are explored within the setting of finite elements in the next section.

## 4. Error formulæ for finite elements

In this section the *expanded* form of the multipoint Taylor formula (3.6) is extended to *all* finite elements, with the examples of the *cubic Hermite $n$-simplex* and *Wilson's brick* being treated in detail.

### Finite elements

The following (standard) definitions can be found in Brenner and Scott [BS94:Ch.3].

**Definition 4.1.** *Let*
(i) $K \subset \mathbb{R}^n$ *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 (e.g., those from some Sobolev space). 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

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

A typical example of a finite element is that (with its local interpolant) given by a Lagrange map $\mathcal{L}_{P,\Theta}$, i.e., a **Lagrange finite element** (see, e.g., Ciarlet and Lions [CL91:p95]). Here, the *element domain* is usually the convex hull of $\Theta$, the *shape functions* are $P$, and the *nodal variables* are the point evaluations $\{\delta_v\}_{v \in \Theta}$. In this case the *nodal basis* consists of the Lagrange functions and (4.3) is exactly the Lagrange form (3.1) of $\mathcal{L}_{P,\Theta}$.

## Expanding the (unexpanded) multipoint Taylor formula for a finite element in terms of the nodal basis

In Section 3, when $\mathcal{L}$ was a Lagrange map $\mathcal{L}_{P,\Theta}$, equation (3.17) was expanded in terms of the Lagrange functions (the nodal basis) to obtain the Kowalewski – Ciarlet – Wagschal multipoint Taylor formula (3.6). In exactly the same way, when $\mathcal{L} = \mathcal{I}_K$ is any finite element it is possible to expand (3.17) in terms of the nodal basis $\{\phi_i\}$ as follows.

**Theorem 4.4.** *Suppose that* $(K, P, N)$ *is a finite element with* $\Pi_k \subset P \subset C^k$*. Then, for sufficiently smooth* $f$ *and* $g \in \Pi_k$*,*

$$g(D)(f - \mathcal{I}_K f)(x) = -\sum_i N_i(\mathcal{R}_{k,x} f)\,(g(D)\phi_i)(x), \tag{4.5}$$

*where* $N_i(\mathcal{R}_{k,x} f)$ *can be expressed in terms of* $D^{k+1} f$ *using Lemma 4.11 (see below).*

**Proof:** Since $\mathcal{L} := \mathcal{I}_K$ is a linear projector that reproduces $\Pi_k$, Theorem 3.15 gives (3.17). Expanding $\mathcal{I}_K(\mathcal{R}_{k,x} f)$ as it occurs in (3.17) using (4.3) gives (4.5). The expansion of $N_i(\mathcal{R}_{k,x} f)$ in terms of $D^{k+1} f$ is discussed below. □

For the reasons just mentioned, equation (4.5) (and expansions of it by Lemma 4.11) will be referred to as the **expanded multipoint Taylor formula** for the finite element $(K, P, N)$ (or, equivalently, for its local interpolant $\mathcal{I}_K$).

Using variations of the Ciarlet–Raviart proof of the multipoint Taylor formula (discussed earlier) the equivalent of (4.5) has been given for certain finite elements with nodal variables of the form

$$f \mapsto D_{\xi_1} \cdots D_{\xi_d} f(v), \qquad 0 \le d \le k, \tag{4.6}$$

which are often called *Hermite finite elements.* In Ciarlet and Raviart [CR72:Th.3] it is proved for $P = \Pi_k$ and nodal variables with $0 \le d \le 1$, with the result involving nodal variables with $2 \le d \le k$ being described as 'long and cumbersome' and hence omitted. In Gout [Go77:p414] the result is shown for $\Pi_k \subset P$ (without restriction on $0 \le d \le k$).

### Expressing $N_i(\mathcal{R}_{k,x} f)$ in terms of $D^{k+1} f$

The majority of finite elements used in numerical modelling have nodal variables of the form

$$f \mapsto q(D) f(v), \tag{4.7}$$

where $q \in \Pi_d^0$ (the space of *homogeneous* polynomials of degree $d$), $0 \le d \le k$. This is a (slightly) more general form than (4.6). Indeed, the very term *nodal variable* (*degree of freedom* is also used) comes from the fact that (4.6) and (4.7) involve the *node* $v$. A nodal variable of the form (4.7) will be said to have the *usual form.*

The exceptions given in Ciarlet and Lions monograph [CL91] are *Wilson's brick* (page 214), and the *Fraeijs de Veubeke triangle* (page 309). In addition to nodal variables of the usual form (4.7), these involve some of the form

$$f \mapsto \int_T q(D) f(v)\, dv, \tag{4.8}$$

where $T$ is the element domain $K$, or one of its faces, and $dv$ is Lebesgue measure.

By considering the error in Taylor interpolation we can now give expressions for $N_i(\mathcal{R}_{k,x} f)$ in terms of $D^{k+1} f$ for nodal variables $N_i$ of either form. These will be presented using the following instance of the *divided difference functional on* $\mathbb{R}^n$ (see (5.6) for the general definition). For $0 \le d \le k$, let

$$\int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} f := \frac{1}{(k-d)!} \int_0^1 (1-t)^{k-d} f(x + t(v-x))\, dt, \tag{4.9}$$

which allows (3.5) to be rewritten as

$$\mathcal{R}_{k-d,x} f(v) = \int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} D_{v-x}^{k+1-d} f. \tag{4.10}$$

**Lemma 4.11.** *Suppose that $\lambda = N_i$ is a nodal variable of the usual form, i.e.,*

$$\lambda(f) := q(D)f(v),$$

*where $q \in \Pi_d^0$, $0 \le d \le k$. Then, for sufficiently smooth $f$,*

$$\lambda(\mathcal{R}_{k,x} f) = \int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} D_{v-x}^{k+1-d} q(D) f. \tag{4.12}$$

*In particular, if $f$ is $C^{k+1}$ in a neighbourhood of the line segment from $x$ to $v$, then there exists $\xi$ belonging to the line segment from $x$ to $v$ for which*

$$\lambda(\mathcal{R}_{k,x} f) = \frac{1}{(k+1-d)!} (D_{v-x}^{k+1-d} q(D)\, f)(\xi). \tag{4.13}$$

*If $\lambda = N_i$ is a nodal variable which is of the form (4.8), i.e.,*

$$\lambda(f) := \int_T q(D) f\,(v)\,dv,$$

*then, for sufficiently smooth $f$,*

$$\lambda(\mathcal{R}_{k,x} f) = \int_T \left( \int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} D_{v-x}^{k+1-d} q(D) f \right) dv. \tag{4.14}$$

**Proof:** The proof relies upon the fact that

$$q(D)(\mathcal{T}_{k,x} f) = \mathcal{T}_{k-d,x}(q(D)f), \tag{4.15}$$

which is a special case of a more general result for the scale of *mean value interpolations* (see Waldron [W96$_1$:Prop.5.11]). Here is a direct proof. If $f \in \Pi_k$, then $q(D)f \in \Pi_{k-d}$ and it is obvious that (4.15) holds. But both sides of (4.15) depend only on the derivatives up to order $k$ of $f$ at $x$, and so it holds for all (smooth) $f$.

From (4.15) it follows that

$$q(D)(\mathcal{R}_{k,x} f) = \mathcal{R}_{k-d,x}(q(D)f),$$

and hence, by (4.10),

$$\lambda(\mathcal{R}_{k,x} f) = q(D)(\mathcal{R}_{k,x} f)(v) = \mathcal{R}_{k-d,x}(q(D)f) = \int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} D_{v-x}^{k+1-d} q(D) f,$$

giving (4.12). Since

$$\int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} 1 = \frac{1}{(k+1-d)!},$$

the mean value theorem (for integrals) can be applied to (4.12) to obtain (4.13). Lastly, integrating (4.12) over $v \in T$ gives (4.14). □

Now we illustrate the use of Theorem 4.4 with three examples.

**Example 1: The cubic Hermite $n$-simplex**

Here we give the expanded multipoint Taylor error formula for the *cubic Hermite $n$-simplex.* This finite element was introduced (under the name *approximation of type* 3) in Ciarlet and Wagschal's paper [CW71]. It is defined as follows.

Let $K$ be an $n$-simplex in $\mathbb{R}^n$ with vertices $\{a_i\}_{i=1}^{n+1}$, and let

$$a_{ijk} := (a_i + a_j + a_k)/3, \qquad 1 \le i < j < k \le n+1.$$

Let $N$ be the set of linear functionals consisting of point evaluation at $a_i$, $1 \le i \le n+1$, and $a_{ijk}$, $1 \le i < j < k \le n+1$, together with

$$N_{ij} : f \mapsto D_{a_j - a_i} f(a_i), \qquad 1 \le i, j \le n+1, \quad i \ne j.$$

Then $(K, \Pi_3, N)$ is a finite element called the **cubic Hermite $n$-simplex**.

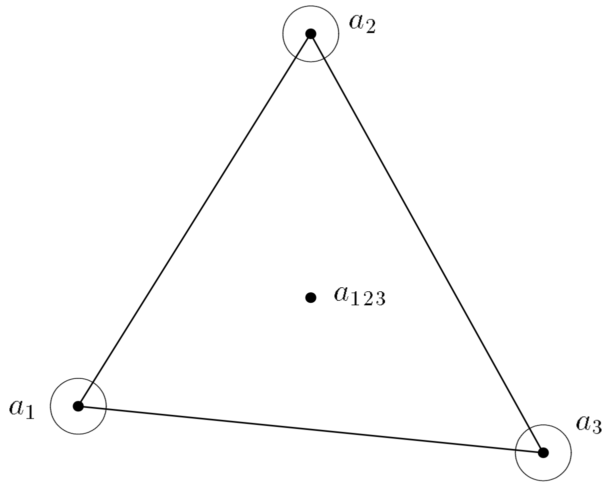


**Fig. 4.1.** The cubic Hermite triangle ($\dim \Pi_3(\mathbb{R}^2) = 10$).

Its local interpolant is given by

$$\mathcal{H}f := \mathcal{I}_K f = \sum_i f(a_i)\,\phi_i + \sum_{i<j<k} f(a_{ijk})\,\phi_{ijk} + \sum_{i \ne j} D_{a_j - a_i} f(a_i)\,\phi_{ij},$$

where the nodal basis has the explicit representation

$$\begin{aligned}
\phi_i &:= -2\lambda_i^3 + 3\lambda_i^2 - 7\lambda_i \sum_{\substack{j<k \\ j,k \ne i}} \lambda_j \lambda_k, \\
\phi_{ijk} &:= 27\lambda_i \lambda_j \lambda_k, \\
\phi_{ij} &:= \lambda_i \lambda_j (2\lambda_i + \lambda_j - 1),
\end{aligned}$$

with $\{\lambda_i\}$ the barycentric coordinates corresponding to the points $\{a_i\}$.

Taking $q(D) = D_{a_j - a_i}$ and $v = a_i$ in Lemma 4.11 gives the expansion

$$N_{ij}(\mathcal{R}_{3,x} f) = \int_{[x,x,x,a_i]} D^3_{a_i - x} D_{a_j - a_i} f.$$

Thus, applying Theorem 4.4 to $\mathcal{I}_K = \mathcal{H}$ (with $k = 3$) gives the error formula: for $g \in \Pi_3$

$$\begin{aligned} g(D)(f - \mathcal{H}f)(x) = &- \sum_i \Big( \int_{[x,x,x,x,a_i]} D^4_{a_i - x} f \Big) (g(D)\phi_i)(x) \\ &- \sum_{i<j<k} \Big( \int_{[x,x,x,x,a_{ijk}]} D^4_{a_{ijk} - x} f \Big) (g(D)\phi_{ijk})(x) \\ &- \sum_{i \neq j} \Big( \int_{[x,x,x,a_i]} D^3_{a_i - x} D_{a_j - a_i} f \Big) (g(D)\phi_{ij})(x), \end{aligned} \tag{4.16}$$

which holds for all sufficiently smooth $f$.

In the next section it is shown how to obtain $L_p$-error bounds from this and other such formulæ.

### Example 2: Wilson's brick

The second example involves nodal variables which are not of the usual form. Let $K$ be a rectangle in $\mathbb{R}^3$ aligned in the coordinate directions $e_1$, $e_2$, $e_3$, with its length in these directions being $2h_1$, $2h_2$, $2h_3$ respectively, and its (eight) vertices being $v \in \Theta$. Let

$$P := \Pi_2 \oplus \operatorname{span}\{(\,)^{1,1,1}\},$$

where $(\,)^{1,1,1} : (x_1, x_2, x_3) \mapsto x_1 x_2 x_3$. Then, **Wilson's brick** (see [WTDG73]) is the finite element $(K, P, N)$, where the nodal variables $N$ consist of the eight point evaluations $\{\delta_v : v \in \Theta\}$ together with

$$N_j : f \mapsto \frac{h_j^2}{h_1 h_2 h_3} \int_K D_j^2 f, \qquad 1 \le j \le 3.$$

Here $D_j^2 f$ denotes $D^2_{e_j} f$, the second partial derivative of $f$ in the coordinate direction $e_j$.

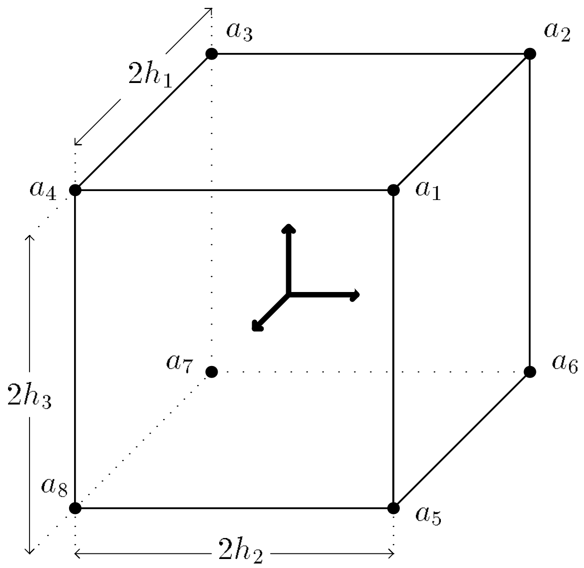


**Fig. 4.2.** Wilson's brick, $n = 3$ ($\dim P = 11$).

The local interpolant for Wilson's brick is

$$\mathcal{W}f := \mathcal{I}_K f := \sum_{v\in\Theta} f(v)\,\phi_v + \sum_j \left(\frac{h_j^2}{h_1h_2h_3}\int_K D_j^2 f\right)\phi_j.$$

Formulæ for the nodal basis $\{\phi_v : v \in \Theta\} \cup \{\phi_j : 1 \le j \le 3\}$ can be found in [CL91:p214].

Applying Theorem 4.4 to $\mathcal{I}_K = \mathcal{W}$ (using (4.14) to express $N_j(\mathcal{R}_{2,x}f)$ in terms of $D^3 f$) gives the following error formula: for $g \in \Pi_2$

$$\begin{aligned} g(D)(f - \mathcal{W}f)(x) = &- \sum_{v\in\Theta}\left(\int_{[x,x,x,v]} D^3_{v-x} f\right)(g(D)\phi_v)(x) \\ &- \sum_j \frac{h_j^2}{h_1h_2h_3}\int_K \left(\int_{[x,v]} D_{v-x} D_j^2 f\right) dv\ (g(D)\phi_j)(x), \end{aligned} \tag{4.17}$$

for all sufficiently smooth $f$.

### Example 3: Numerical differentiation and integration rules

Interpolatory (numerical) differentiation and integration rules based on point evaluations at $v \in \Theta$ and exact for $\Pi_k$ are of the form

$$\sum_{v\in\Theta} w(v)\,f(v) = \mathcal{A}\mathcal{L}_{P,\Theta} f,$$

where $\mathcal{A}$ is the derivative or integral to be approximated, $\mathcal{L}_{P,\Theta}$ is a Lagrange map, and the *weights* $w(v)$ are given by

$$w(v) = \mathcal{A}(\ell_v), \quad v \in \Theta.$$

The error in approximation by such a rule is

$$\mathcal{A}(1 - \mathcal{L}_{P,\Theta})f, \tag{4.18}$$

which is of the form (2.1). For a *differentiation* rule, i.e.,

$$\mathcal{A}f := g(D)f(x), \qquad g \in \Pi_k,$$

this is the pointwise error in a Lagrange map, for which there is the Kowalewski – Ciarlet – Wagschal multipoint Taylor formula (3.6). For differentiation rules based on data that includes linear functionals other than point evaluations the expanded multipoint Taylor formula for a finite element (4.5) applies.

For an *integration* rule, i.e.,

$$\mathcal{A}f := \int_\Omega f,$$

where $\Omega$ is some region of volume $\mathrm{vol}(\Omega)$, it is also possible to apply Theorem 2.2. Let $\mathcal{P}$ be some linear projector onto $\Pi_k$ satisfying

$$\int_\Omega f = \int_\Omega \mathcal{P}f, \tag{4.19}$$

which is condition (2.9), and denote its error by $\mathcal{E} := 1 - \mathcal{P}$. Then, Corollary 2.7 gives the error formula

$$\begin{aligned} \int_\Omega f - \sum_{v\in\Theta} w(v)\, f(v) &= -\mathcal{A}\mathcal{L}_{P,\Theta}\mathcal{E}f \\ &= -\mathcal{A}\big(\sum_{v\in\Theta} \mathcal{E}(v)\ell_v\big) \\ &= -\sum_{v\in\Theta} w(v)(\mathcal{E}f)(v). \end{aligned} \tag{4.20}$$

The hope is then to express the quantities $(\mathcal{E}f)(v)$, $v \in \Theta$ occuring above in terms of $D^{k+1}f$. The practical difficulty with this scheme is finding (appropriate) projectors $\mathcal{P}$ which satisfy (4.19). At first glance it might seem the averaged Taylor interpolant

$$\mathcal{P}f := \frac{1}{\mathrm{vol}(\Omega)} \int_\Omega \mathcal{T}_{k,x} f\, dx$$

would be a good choice, however this does not satisfy (4.19). The author has not pursued these questions further. There are a number of linear projectors $\mathcal{P}$ matching certain integrals, together with error representations $\mathcal{E}f = \mathcal{Q}(D^{k+1}f)$, the best known being the scale of *mean value interpolations* (see, e.g., [W96$_1$]), and so it would seem that at least in some situations such a theory might be feasible.

## 5. $L_p$–error bounds for finite elements

In this section, we use the *multivariate Hardy's inequality* (recently introduced by the author) to obtain $L_p$–error bounds from the expanded multipoint Taylor formula for a finite element. To do this, the following facts about Sobolev spaces will be required.

## Sobolev spaces

Let $W_p^k(\Omega)$ be the *Sobolev space* of functions defined on $\Omega$ (a bounded open set in $\mathbb{R}^n$ with a *Lipschitz* boundary) with derivatives up to order $k$ in $L_p(\Omega)$ equipped with the usual topology (see, e.g., Adams [Ad75]). It is convenient to include the condition that $\Omega$ have a Lipschitz boundary in the definition, so that Sobolev's embedding theorem(s) can be applied. We will only need the following consequence of Sobolev's embedding theorem(s) (which can be found in any text on Sobolev spaces). If $k+1-d-n/p>0$, then

$$W_p^{k+1}(\Omega) \subset C^d(\bar{\Omega}), \tag{5.1}$$

where $\bar{\Omega}$ is the closure of $\Omega$.

To measure the size of its $k$–th order derivative, it is convenient to associate with each $f \in W_p^k(\Omega)$ the function $|D^k f| \in L_p(\Omega)$, given by the rule

$$|D^k f|(x) := \sup_{\substack{\xi_1,\ldots,\xi_k \in \mathbb{R}^n \\ \|\xi_i\| \le 1}} |D_{\xi_1} \cdots D_{\xi_k} f(x)| = \sup_{\substack{u \in \mathbb{R}^n \\ \|u\|=1}} |D_u^k f(x)|, \tag{5.2}$$

where the derivatives $D_{\xi_1} \cdots D_{\xi_k} f$ are computed from any (fixed) choice of representatives for the $k$–th order partial derivatives of $f$. The equality of the two suprema above follows from a classical result of Banach on the norm of a symmetric multilinear map. This definition of $|D^k f|$ is consistent with its standard univariate interpretation. From (5.2), it is easy to see that $|D^k f|$ is well–defined and satisfies

$$|D_{\xi_1} \cdots D_{\xi_k} f| \le |D^k f| \, \|\xi_1\| \cdots \|\xi_k\|, \qquad \text{a.e.}, \tag{5.3}$$

for $\xi_1, \ldots, \xi_k \in \mathbb{R}^n$. The $L_p(\Omega)$–norm of $|D^k f|$ gives a seminorm on $W_p^k(\Omega)$

$$f \mapsto |f|_{k,p,\Omega} := \| \, |D^k f| \, \|_{L_p(\Omega)}. \tag{5.4}$$

This coordinate–independent seminorm (5.4) is ideal for the analysis that follows because of (5.3). It is convenient to generalise (5.3) slightly. For $q \in \Pi_d^0$, let

$$\|q(D)\| := \max_{f \in \Pi_d^0 \setminus 0} \frac{|q(D)f(0)|}{|D^d f|(0)},$$

which defines norm on the differential operators $\{q(D) : q \in \Pi_d^0\}$. It can easily be shown that: for $f \in W_p^k(\Omega)$

$$|D_{\xi_1} \cdots D_{\xi_{k-d}} q(D) f| \le \|\xi_1\| \cdots \|\xi_{k-d}\| \, \|q(D)\| \, |D^k f|. \tag{5.5}$$

## The multivariate Hardy's inequality

The multivariate Hardy's inequality involves the following linear functional called the *divided difference functional on* $\mathbb{R}^n$ by Micchelli in [Mi80].

**Definition.** *For $\Theta := [\theta_0, \ldots, \theta_k]$ any sequence of $k+1$ points in $\mathbb{R}^n$, let*

$$f \mapsto \int_\Theta f := \int_0^1 \int_0^{s_1} \ldots \int_0^{s_{k-1}} f(\theta_0 + s_1(\theta_1 - \theta_0) + \cdots + s_k(\theta_k - \theta_{k-1}))\, ds_k \cdots ds_2\, ds_1, \quad (5.6)$$

*with the convention that $\int_{[]} f := 0$.*

This linear functional was introduced (with a change of variables) in (4.9), for the special case $\Theta = [x, \ldots, x, v]$. Let

$$(a)_n := (a)(a+1)(a+2)\cdots(a+n-1),$$

be the shifted factorial function, and $\#\Theta$ denote the cardinality of $\Theta$.

**Multivariate Hardy's inequality 5.7 ([W96$_2$:Th.3.2.1]).** *Let $\Theta$ be a finite sequence in $\mathbb{R}^n$, and let $\Omega$ be an open set in $\mathbb{R}^n$ for which $\bar{\Omega}$ is starshaped with respect to $\Theta$. If $m - n/p > 0$, then the rule*

$$H_{m,\Theta} f(x) := \int_{[\underbrace{x,\ldots,x}_{m},\Theta]} f \quad (5.8)$$

*induces a positive bounded linear map $H_{m,\Theta} : L_p(\Omega) \to L_p(\Omega)$ with norm*

$$\|H_{m,\Theta}\|_{L_p(\Omega)} \le \frac{1}{(m-1)!(m-n/p)_{\#\Theta}} \to \infty \quad \text{as} \quad m - n/p \to 0^+. \quad (5.9)$$

*This upper bound for $\|H_{m,\Theta}\|_{L_p(\Omega)}$ is sharp when $\Theta$ involves only one point, i.e., when*

$$\Theta = [v, \ldots, v],$$

*and when $p = \infty$ (with the norm taken on only for the constant functions).*

When $n = 1$, $m = 1$, $\Theta = [0]$ and $\Omega = (0, \infty)$ the inequality (5.9) is the well–known Hardy's inequality: for $p > 1$

$$\| \, x \mapsto \frac{1}{x} \int_0^x f \, \|_{L_p(0,\infty)} \le \frac{p}{p-1} \|f\|_{L_p(0,\infty)}, \quad \forall f \in L_p(0,\infty),$$

and so (5.9) is referred to as the *multivariate (form of) Hardy's inequality.*

### $L_p$–error bounds from the multivariate Hardy's inequality

Next, $L_p$–error bounds are obtained from the expanded multipoint Taylor formula for a finite element using the multivariate Hardy's inequality. Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ (with Lipschitz boundary), and

$$h_{v,\Omega} := \sup_{x \in \Omega} \|v - x\| \le h := \operatorname{diam} \Omega \qquad \text{(the diameter of } \Omega).$$

**Lemma 5.10.** *Suppose that $\bar\Omega$ is starshaped with respect to $v$, and $q \in \Pi_d^0$, $0 \le d \le k$. If $k+1-d-n/p > 0$ and $f \in W_p^{k+1}(\Omega)$, then*

$$\| \, x \mapsto (q(D)\mathcal{R}_{k,x} f)(v) \, \|_{L_p(\Omega)} \le \frac{(h_{v,\Omega})^{k+1-d} \|q(D)\|}{(k-d)!(k+1-d-n/p)} |f|_{k+1,p,\Omega}. \tag{5.11}$$

**Proof:** Since $k+1-d-n/p>0$, it follows from Lemma 4.11 and the multivariate Hardy's inequality 5.7 that

$$Sf : x \mapsto (q(D)\mathcal{R}_{k,x} f)(v) \tag{5.12}$$

defines a function in $L_p(\Omega)$. From (5.5) we obtain that

$$\begin{aligned} |D_{v-x}^{k+1-d} q(D) f| &\le \|x-v\|^{k+1-d} \|q(D)\| \, |D^{k+1} f| \\ &\le (h_{v,\Omega})^{k+1-d} \|q(D)\| \, |D^{k+1} f|, \end{aligned}$$

in $L_p(\Omega)$. Thus, by Lemma 4.11 and the multivariate Hardy's inequality 5.7, we obtain

$$\begin{aligned} \|Sf\|_{L_p(\Omega)} &\le (h_{v,\Omega})^{k+1-d} \|q(D)\| \left\| x \mapsto \int_{[\underbrace{x,\ldots,x}_{k+1-d},v]} |D^{k+1} f| \right\|_{L_p(\Omega)} \\ &\le (h_{v,\Omega})^{k+1-d} \|q(D)\| \frac{1}{(k-d)!(k+1-d-n/p)} \| \, |D^{k+1} f| \, \|_{L_p(\Omega)} \\ &= \frac{(h_{v,\Omega})^{k+1-d} \|q(D)\|}{(k-d)!(k+1-d-n/p)} |f|_{k+1,p,\Omega}, \end{aligned}$$

which is (5.11). $\square$

We now extend the $L_p$–error bounds of Arcangeli and Gout [AG76] and Gout [Go77] to a general finite element and for a much wider class of seminorms, defined as follows. A family of seminorms $|\cdot|_{p,\Omega}$ on $W_p^k(\Omega)$, $1 \le p \le \infty$ is said to be of the *usual form* if either

$$|f|_{p,\Omega} := \| \left( \|g_j(D) f\|_{L_p(\Omega)} : j = 1, \ldots m \right) \|_{\mathbb{R}^m}, \tag{5.13}$$

where the $g_j \in \Pi_k(\mathbb{R}^n)$ are fixed and $\|\cdot\|_{\mathbb{R}^m}$ is any norm on $\mathbb{R}^m$, or

$$|\cdot|_{p,\Omega} := |\cdot|_{j,p,\Omega}, \qquad 0 \le j \le k. \tag{5.14}$$

**Theorem 5.15.** *Suppose that $(K,P,N)$ is a finite element with $\Pi_k \subset P \subset C^k(\bar\Omega)$, and nodal variables $\{N_i\}$ of the usual form*

$$N_i(f) := q_i(D) f(v_i),$$

*where $q_i \in \Pi_{d_i}^0$, $0 \le d_i \le k$, and $v_i \in \mathbb{R}^n$. Let $d := \max_i d_i$. If $k+1-d-n/p > 0$, then*

$$|f - \mathcal{I}_K f|_{p,\Omega} \le \sum_i \frac{(h_{v_i,\Omega})^{k+1-d_i} \|q_i(D)\| \, |\phi_i|_{\infty,\Omega}}{(k-d_i)!(k+1-d_i-n/p)} |f|_{k+1,p,\Omega}, \qquad \forall f \in W_p^{k+1}(\Omega), \tag{5.16}$$

*where* $|\cdot|_{p,\Omega}$ *is any family of seminorms on* $W_p^k(\Omega)$ *of the usual form.*

**Proof:** Let $g_j \in \Pi_k$. Since $k+1-d-n/p>0$, Sobolev's embedding theorem (5.1) implies that $\mathcal{I}_K$ is well defined on $W_p^{k+1}(\Omega)$. It follows from the multivariate Hardy's inequality 5.7 and Theorem 4.4 that, for $f \in W_p^{k+1}(\Omega)$, the error satisfies

$$g_j(D)(f-\mathcal{I}_K f)(x) = -\sum_i (q_i(D)\mathcal{R}_{k,x}f)(v_i)\ (g_j(D)\phi_i)(x)$$

in $L_p(\Omega)$, giving

$$|g_j(D)(f-\mathcal{I}_K f)(x)| \le \sum_i |(q_i(D)\mathcal{R}_{k,x}f)(v_i)|\, \|g_j(D)\phi_i\|_{L_\infty(\Omega)}. \tag{5.17}$$

Since $\|\cdot\|_{L_p(\Omega)}$ is a monotone norm, applying Lemma 5.10 to this (as a function of $x$) gives

$$\|g_j(D)(f-\mathcal{I}_K f)\|_{L_p(\Omega)} \le \sum_i \frac{(h_{v_i,\Omega})^{k+1-d_i}\|q_i(D)\|}{(k-d_i)!(k+1-d_i-n/p)} \|g_j(D)\phi_i\|_{L_\infty(\Omega)}\ |f|_{k+1,p,\Omega}. \tag{5.18}$$

Taking the $\|\cdot\|_{\mathbb{R}^m}$ norm of the inequality for $m$–vectors (indexed by $j$) which is given coordinatewise by (5.18) gives (5.16) for a family of seminorms of the form (5.13). Take the supremum of each side of (5.17) over the set of differential operators $\{D_u^j : \|u\|=1\}$ to obtain

$$|D^j(f-\mathcal{I}_K f)|(x) \le \sum_i |(q_i(D)\mathcal{R}_{k,x}f)(v_i)|\ |\phi_i|_{j,\infty,\Omega}, \qquad \text{a.e. } x.$$

Next, take the $L_p(\Omega)$–norm of each side of this and apply Lemma 5.10 to get

$$|f-\mathcal{I}_K f|_{j,p,\Omega} \le \sum_i \frac{(h_{v_i,\Omega})^{k+1-d_i}\|q_i(D)\|\ |\phi_i|_{j,\infty,\Omega}}{(k-d_i)!(k+1-d_i-n/p)}\ |f|_{k+1,p,\Omega},$$

which is (5.16) for seminorms of the form (5.14). □

*Remark.* If $(K,P,N)$ involves nodal variables of the form (4.8), i.e.,

$$\lambda : f \mapsto \int_T q(D)f\,dv,$$

then the argument of Theorem 5.15 can be modified by using the integral form of Minkowski's inequality to bring $\|\cdot\|_{L_p(\Omega)}$ inside the integral $\int_T(\cdot)\,dv$. This leads to estimates of the form

$$|\ x \mapsto \lambda(\mathcal{R}_{k,x}f)\phi(x)\ |_{p,\Omega} \le \sum_i \frac{h^{k+1-d}\|q(D)\|\,\mathrm{vol}(T)|\phi|_{\infty,\Omega}}{(k-d)!(k+1-d-n/p)}\ |f|_{k+1,p,\Omega} \tag{5.19}$$

in (5.16).

Now we illustrate Theorem 5.15 with some examples where for simplicity $h_{v_i,\Omega}$ is replaced by the (possibly larger) constant $h$.

If $\mathcal{I}_K$ is a Lagrange map $\mathcal{L}_{P,\Theta}$, then (5.16) reduces to

$$|f - \mathcal{L}_{P,\Theta} f|_{p,\Omega} \le \frac{1}{k!(k+1-n/p)} \left( \sum_{v\in\Theta} |\ell_v|_{\infty,\Omega} \right) h^{k+1} |f|_{k+1,p,\Omega}, \tag{5.20}$$

which was proved by Arcangeli and Gout in [AG76:Th.1–1] when $|\cdot|_{p,\Omega}$ is a seminorm of the form (5.14).

Following the work of [AG76], Gout [Go77] obtained Theorem 5.15 for finite elements with nodal variables of the form (4.6) and seminorms of the form (5.14). Both of these works are based on (the equivalent of) the multivariate Hardy's inequality as used in the proof of Theorem 5.15.

For the cubic Hermite $n$–simplex (5.16) reduces to: for $p > n/3$ and $f \in W_p^4(\Omega)$

$$|f - \mathcal{H} f|_{p,\Omega} \le \left( \sum_i \frac{|\phi_i|_{\infty,\Omega}}{3!(4-n/p)} + \sum_{i<j<k} \frac{|\phi_{ijk}|_{\infty,\Omega}}{3!(4-n/p)} + \sum_{i\ne j} \frac{|\phi_{ij}|_{\infty,\Omega}}{2!(3-n/p)} \right) h^4 \; |f|_{4,p,\Omega},$$

where $|\cdot|_{p,\Omega}$ is any of the usual seminorms on $W_p^3(\Omega)$. While for Wilsons's brick, using (5.19), we obtain: for $p > n$ and $f \in W_p^3(\Omega)$

$$|f - \mathcal{W} f|_{p,\Omega} \le \left( \sum_{v\in\Theta} \frac{|\phi_v|_{\infty,\Omega}}{2!(3-n/p)} + \sum_j \frac{|\phi_j|_{\infty,\Omega}}{0!(1-n/p)} \right) h^3 \; |f|_{3,p,\Omega},$$

where $|\cdot|_{p,\Omega}$ is any of the usual seminorms on $W_p^2(\Omega)$.

## 6. Error formulæ for linear interpolation

In this section, Theorem 2.2 is used to construct a family of formulæ for pointwise error in linear interpolation. By *linear interpolation* we mean interpolation from the space of linear polynomials $\Pi_1$ at $\Theta$ a set of $n+1$ (affinely independent) points in $\mathbb{R}^n$, i.e., a Lagrange map

$$\mathcal{L}_\Theta := \mathcal{L}_{\Pi_1,\Theta}.$$

If $\mathcal{V}$ is a linear right inverse for

$$\mathcal{U} := D^2,$$

and the corresponding linear projector $\mathcal{P} := 1 - \mathcal{V}\mathcal{U}$ onto $\Pi_1 (= \ker D^2)$ satisfies

$$\mathcal{P} f(x) = f(x), \tag{6.1}$$

then by Theorem 2.2

$$f(x) - \mathcal{L}_\Theta f(x) = -(\mathcal{L}_\Theta \mathcal{V}(D^2 f))(x) = -\sum_{v\in\Theta} \mathcal{V}(D^2 f)(v)\,\ell_v(x). \tag{6.2}$$

Earlier, the choice $\mathcal{P} = \mathcal{T}_{1,x}$ was made to obtain from (6.2) that

$$f(x) - \mathcal{L}_\Theta f(x) = -\sum_{v\in\Theta} \left( \int_{[x,x,v]} D^2_{v-x} f \right) \ell_v(x), \tag{6.3}$$

which is the expanded multipoint Taylor formula (4.5) for $g = 1$ (originally given in [CW71]).

In addition to $\mathcal{P} = \mathcal{T}_{1,x}$, there are many other linear projectors onto $\Pi_1$ that satisfy (6.1). When a representation for $\mathcal{V}$ the corresponding right inverse of $D^2$ is available, each of these leads to a different error formula (6.2). For example, one might take $\mathcal{P}$ to be *Kergin interpolation* at $\{x, \xi\}$ with $\xi \in \mathbb{R}^n$ (see, e.g., Waldron [W96$_1$]), for which

$$\mathcal{V}(D^2 f) = \int_{[\cdot,x,\xi]} D_{\cdot-x} D_{\cdot-\xi} f. \tag{6.4}$$

When $\xi = x$ this is Taylor interpolation at $x$ (compare (6.4) with (4.10) for $k - d = 1$). More generally, one can take a weighted average over $\xi$ of such approximations to obtain the following family of formulæ of the type (6.2).

**Theorem 6.5.** *Let $\mu$ be a real measure on $\mathbb{R}^n$ of mass 1 (and finite variation). Then, the error in linear interpolation to $\Theta$ satisfies*

$$f(x) - \mathcal{L}_\Theta f(x) = -\sum_{v\in\Theta} \int_{\mathbb{R}^n} \left( \int_{[v,x,\xi]} D_{v-x} D_{v-\xi} f \right) d\mu(\xi)\,\ell_v(x), \tag{6.6}$$

*for sufficiently smooth $f$.*

**Proof:** The convergence of the integrals of the operators used below (which can be interpreted pointwise) follows from the standard theory of integration.

Let $\mathcal{K}_{\{x,\xi\}}$ denote the operator of Kergin interpolation to $\{x, \xi\}$ which has the representation

$$\mathcal{K}_{\{x,\xi\}} f = f(x) + \int_{[x,\xi]} D_{\cdot-x} f. \tag{6.7}$$

Integrating the right hand side of (6.7) against the measure $\mu$ shows that

$$\mathcal{P} f := \int \mathcal{K}_{\{x,\xi\}} f \, d\mu(\xi)$$

defines a linear projector onto $\Pi_1$ which satisfies (6.1). Similarly, the right inverses corresponding to $\mathcal{K}_{\{x,\xi\}}$ given by (6.4) can be integrated against $\mu$ to get

$$\mathcal{V}(D^2 f) = \int_{\mathbb{R}^n} \int_{[\cdot,x,\xi]} D_{\cdot-x} D_{\cdot-\xi} f \, d\mu(\xi), \tag{6.8}$$

the right inverse corresponding to $\mathcal{P}$. Substitute (6.8) in (6.2) to obtain (6.6). $\square$

Certain choices of $\mu$ in Theorem 6.5 as a measure supported on a finite number of points give error formulæ that are known. Here are these examples which were obtained (some very recently) by a variety of methods.

**Example 1:** Suppose that

$$\mu = \delta_x,$$

i.e., point evaluation at $x$, then (6.6) gives the 'classical' multipoint Taylor formula (6.3).

**Example 2:** Suppose that

$$\mu = \delta_w,$$

i.e., point evaluation at $w \in \Theta$, then (6.6) gives

$$f(x) - \mathcal{L}_\Theta f(x) = \sum_{\substack{v \in \Theta \\ v \neq w}} \left( \int_{[x,w,v]} D_{x-v} D_{v-w} f \right) \ell_v(x), \tag{6.9}$$

which is a special case of the *Sauer–Xu formula* [SX95:Cor.3.11] (also see de Boor [B96]).

**Example 3:** Suppose that

$$\mu = \sum_{w \in \Theta} \ell_w(x) \delta_w$$

a weighted average of the point evaluations at $w \in \Theta$, then (6.6) gives

$$f(x) - \mathcal{L}_\Theta f(x) = \sum_{v \in \Theta} \sum_{w \in \Theta} \left( \int_{[x,w,v]} D_{x-v} D_{v-w} f \right) \ell_v(x) \ell_w(x). \tag{6.10}$$

The 'symmetrised' form of (6.10) is

$$f(x) - \mathcal{L}_\Theta f(x) = \sum_{\substack{\{v,w\} \subset \Theta \\ v \neq w}} \left( \int_{[x,w,v]} D_{w-v} D_{v-w} f \right) \ell_v(x) \ell_w(x), \tag{6.11}$$

which was given by Waldron in [W96$_3$].

**Example 4:** Suppose that

$$\mu = \frac{1}{n+1} \sum_{w \in \Theta} \delta_w$$

the average of the point evaluations at $w \in \Theta$, then (6.6) gives

$$f(x) - \mathcal{L}_\Theta f(x) = \frac{1}{n+1} \sum_{v \in \Theta} \sum_{w \in \Theta} \left( \int_{[x,v,w]} D_{v-x} D_{w-v} f \right) \ell_v(x), \tag{6.12}$$

which has the 'symmetrised' form

$$f(x) - \mathcal{L}_\Theta f(x) = \frac{1}{n+1} \sum_{\substack{\{v,w\}\subset\Theta \\ v\neq w}} \int_{[x,w,v]} \ell_v(x) D_{v-x} D_{w-v} f + \ell_w(x) D_{w-x} D_{v-w} f. \quad (6.13)$$

Both of these formulæ are new.

It is easy to see from these examples how many new formulæ for the error in linear interpolation can be obtained by choosing different $\mu$ in Theorem 6.5. This section was only intended to give an indication of how effectively Theorem 2.2 can be applied to obtain error formulæ. The generalisation of Theorem 6.5 to the error in interpolation from $\Pi_k$, where $\mu$ becomes a real measure on $(\mathbb{R}^n)^k$ of mass 1, is left to the reader. In the last section another promising application of Theorem 2.2 is outlined.

## 7. Error formulæ for Lagrange maps from those for one–point interpolation

In this section, we give a generalisation of Theorem 3.15 where $\Pi_k$ is replaced by a polynomial space $Q$. It is shown that the problem of constructing an error formula for Lagrange interpolation from a $D$–invariant polynomial space $P$ (such as a tensor product space, or more generally the least solution) that involves only derivatives which annihilate $P$ can be reduced to the simpler problem of obtaining such a formula for 'one–point' (Taylor) interpolation from $P$ at the origin (which has been solved in a number of cases). These results require the following facts about one–point interpolation.

### One–point interpolation

We say that *one–point interpolation* from a (finite dimensional) space of polynomials $Q$ to derivatives $Q(D)$ at a point $x \in \mathbb{R}^n$ is *correct* if for each function $f$ (smooth enough at $x$) there is a unique $q \in Q$ with

$$p(D)q(x) = p(D)f(x), \qquad \forall p \in Q. \quad (7.1)$$

The associated linear projector $f \mapsto q$ is called the **one–point interpolation map** (from $Q$ at $x$). It is denoted by

$$\mathcal{T}_{Q,x} : f \mapsto q, \quad (7.2)$$

and its error (remainder) by

$$\mathcal{R}_{Q,x} : f \mapsto f - q. \quad (7.3)$$

For $Q = \Pi_k$ this one–point interpolation map is simply the Taylor map from $\Pi_k$ at $x$, which was discussed in Section 3. There the (shorter) notations

$$\mathcal{T}_{k,x} = \mathcal{T}_{\Pi_k,x}, \qquad \mathcal{R}_{k,x} = \mathcal{R}_{\Pi_k,x}$$

were used.

One–point interpolation at the origin, which will be denoted by

$$\mathcal{T}_Q := \mathcal{T}_{Q,0} \qquad (\text{with } \mathcal{R}_Q := \mathcal{R}_{Q,0}),$$

is always correct (even if $Q$ is not a homogeneous space of polynomials). To see this, and other properties of the set

$$\mathrm{OPI}(Q) := \{x \in \mathbb{R}^n : \text{One–point interpolation from } Q \text{ at } x \text{ is correct}\},$$

we equip $\Pi$ (the space of polynomials) with the inner product

$$\langle f, p\rangle := p(D)f(0), \tag{7.4}$$

which was used so effectively in de Boor and Ron [BR92] to study multivariate polynomial interpolation. With the polynomials thought of as an incomplete inner product space in this way we have the following.

**Proposition 7.5.** *One–point interpolation from $Q$ at $x \in \mathbb{R}^n$ is correct for* a.e. $x$. *If $Q$ is $D$–invariant (invariant under differentiation), then one–point interpolation is correct for all $x \in \mathbb{R}^n$, i.e.,*

$$\mathrm{OPI}(Q) = \mathbb{R}^n,$$

*and*

$$\mathcal{T}_{Q,x} = \tau_{-x}\mathcal{T}_Q\tau_x, \tag{7.6}$$

*where $\tau_x$ is the translation operator*

$$\tau_x : f \mapsto f(\cdot + x).$$

*Furthermore, one–point interpolation at the origin is always correct, and its restriction*

$$\mathcal{T}_Q|_\Pi : \Pi \to Q$$

*to the polynomials is the orthogonal projection onto $Q$.*

**Proof:** The orthogonal projection of $f \in \Pi$ onto $Q$ is, by definition, the unique $q \in Q$ for which

$$\langle f - q, Q\rangle = 0. \tag{7.7}$$

By (7.4) it is seen that (7.7) is precisely (7.1), i.e., the interpolation conditions for one–point interpolation from $Q$ at the origin, which gives the last statement.

Let $\{q_i\}$ be a basis for $Q$. One–point interpolation from $Q$ at $x$ is correct if and only if the Gramian matrix

$$G(x) := [q_i(D)q_j(x)]_{i,j} \tag{7.8}$$

is invertible. Thus, one–point interpolation from $Q$ at $x$ is not possible *only* when $x$ is a zero of the polynomial

$$p : x \mapsto \det(G(x)).$$

But, we have shown that one–point interpolation at the origin is correct, so $p(0) \neq 0$, and $p$ is a nonzero polynomial, which therefore has zero set

$$Z(p) = \mathbb{R}^n \setminus \mathrm{OPI}(Q)$$

of measure zero. In other words, one–point interpolation from $Q$ is correct at a.e. $x$.

Suppose that $Q$ is $D$–invariant. Then, $Q$ is translation–invariant (see de Boor [B87]), and so

$$\tau_{-x} \mathcal{T}_Q \tau_x \tag{7.9}$$

is a linear projector onto $Q$. Next, we show that (7.9) interpolates derivatives $Q(D)$ at $x$, so that one–point interpolation from $Q$ at $x$ is correct and given by (7.6). This argument uses the facts that differentiations and translations commute, and

$$\delta_x \tau_y = \delta_{x+y}.$$

Let $p \in Q$, then

$$\begin{aligned}
\delta_x p(D)(\tau_{-x} \mathcal{T}_Q \tau_x f) &= \delta_x \tau_{-x} p(D) \mathcal{T}_Q \tau_x f \\
&= \delta_0 p(D) \mathcal{T}_Q \tau_x f \\
&= \delta_0 p(D) \tau_x f \qquad (\mathcal{T}_Q \text{ interpolates } \delta_0 p(D)) \\
&= \delta_0 \tau_x p(D) f \\
&= \delta_x p(D) f,
\end{aligned}$$

so that $\tau_{-x} \mathcal{T}_Q \tau_x$ matches $p(D)f(x)$. □

The 'almost everywhere' in Proposition 7.5 is needed. Suppose that $Q$ is the space spanned by the (univariate) polynomial

$$q_1 : x \mapsto 1 + ax, \qquad 0 \neq a \in \mathbb{R}.$$

Then, the Gramian matrix (7.8) is

$$G(x) = [1 + a^2 + ax],$$

and so one–point interpolation from $Q$ to $x \in \mathbb{R}$ is *not* correct when

$$x = -\frac{1 + a^2}{a},$$

and is correct for all other values of $x$.

### More general multipoint Taylor formulæ for Lagrange maps and other linear projectors

We are now able to give a generalisation of Theorem 3.15 where $\Pi_k$ is replaced by a (finite–dimensional) polynomial space $Q$.

**Theorem 7.10.** *Suppose that $\mathcal{L}$ is a linear operator that reproduces a polynomial space $Q$. Then, for $x \in \mathrm{OPI}$ and sufficiently smooth $f$*

$$\mathcal{L}f = \mathcal{T}_{Q,x}f + \mathcal{L}\mathcal{R}_{Q,x}f. \tag{7.11}$$

*In particular, for $g \in Q$ and $x \in \mathrm{OPI}(Q)$*

$$g(D)(f - \mathcal{L}f)(x) = -g(G)(\mathcal{L}\mathcal{R}_{Q,x}f)(x). \tag{7.12}$$

*If $\mathcal{L}$ is a Lagrange map $\mathcal{L}_{P,\Theta}$, then (7.12) can be expanded as*

$$g(D)(f - \mathcal{L}_{P,\Theta}f)(x) = -\sum_{v\in\Theta} \mathcal{R}_{Q,x}f(v)\,(g(D)\ell_v)(x). \tag{7.13}$$

**Proof:** The proof is exactly as for Theorem 3.15 with $\Pi_k$ replaced by $Q$ and using the fact that

$$g(D)(\mathcal{T}_{Q,x}f) = g(D)f(x), \qquad \forall x \in \mathrm{OPI}(Q).$$

□

Proposition 7.5 gives the following useful corollary for Lagrange interpolation from a $D$–invariant space.

**Corollary 7.14.** *Suppose that $P$ is a (finite–dimensional) $D$–invariant space (necessarily a polynomial space) for which Lagrange interpolation at $\Theta \subset \mathbb{R}^n$ is correct, and $\mathcal{L}_{P,\Theta}$ is the corresponding Lagrange map. Then, for $g \in P$ and $x \in \mathbb{R}^n$*

$$g(D)(f - \mathcal{L}_{P,\Theta}f)(x) = -\sum_{v\in\Theta} (\tau_{-x}\mathcal{R}_P\tau_x f)(v)\,(g(D)\ell_v)(x), \tag{7.15}$$

*for sufficiently smooth $f$.*

**Proof:** It is shown in de Boor [B87] that a finite–dimensional, $D$–invariant space is necessarily a polynomial space. By (7.6),

$$\mathcal{R}_{P,x} = 1 - \mathcal{T}_{P,x} = \tau_{-x}(1 - \mathcal{T}_P)\tau_x = \tau_{-x}\mathcal{R}_P\tau_x,$$

which is substituted into (7.13) to get (7.15), which holds for all $x \in \mathbb{R}^n = \mathrm{OPI}(P)$. □

Thus, when $P$ is $D$–invariant, to obtain an error formula for $\mathcal{L}_{P,\Theta}$ that involves only derivatives that kill $P$, it is sufficient to do so for the 'simpler' problem of one–point interpolation at the origin. A particular $D$–invariant choice of great interest is the *least solution* of de Boor and Ron [BR92], which associates with any (finite) set of points $\Theta \subset \mathbb{R}^n$ a polynomial space, denoted by

$$P = \Pi_\Theta,$$

for which Lagrange interpolation at $\Theta$ is correct. In addition to being $D$–invariant the *least space* $\Pi_\Theta$ has many other desirable properties including *constructability* and being a *minimal–degree solution*, i.e.,

$$\dim(\Pi_\Theta \cap \Pi_k) \geq \dim(P \cap \Pi_k), \qquad \forall k,$$

for all polynomial spaces $P$ for which Lagrange interpolation at $\Theta$ is correct.

## References


[Ad75] Adams, R. A.(1975): Sobolev spaces. Academic Press, New York

[At92] Attéia, M.(1992): Hilbertian Kernels and Spline Functions. Elsevier Science Publishers, Amsterdam

[AG76] Arcangeli, R. and J. L. Gout(1976): Sur l'evaluation de l'erreur d'interpolation de Lagrange dans un ouvert de $\mathbb{R}^n$. Rev. Française Automat. Informat. Rech. Opér., Anal. Numer. **10(3)**, 5–27

[B87] Boor, C. de(1987): The polynomials in the linear span of integer translates of a compactly supported function. Constr. Approx. **3**, 199-208

[B96] Boor, C. de(1996): On the Sauer-Xu formula for the error in multivariate polynomial interpolation. Math. Comp. **65**, 1231–1234

[BR92] Boor, C. de and A. Ron(1992): The least solution for the polynomial interpolation problem. Math. Z. **210**, 347–378

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

[CL91] Ciarlet, P.G. and J. L. Lions(1991): Handbook of Numerical Analysis, v. II, Finite Element Methods (Part 1). North-Holland, Amsterdam

[CW71] Ciarlet, P. G. and C. Wagschal(1971): Multipoint Taylor formulas and applications to the finite element method. Numer. Math. **17**, 84–100

[CR72] Ciarlet, P. G. and P. A. Raviart(1972): General Lagrange and Hermite interpolation in $\mathbb{R}^N$ with applications to finite element methods. Arch. Rational Mech. Anal. **46**, 177–199

[D75] Davis, P. J.(1975): Interpolation and approximation. Dover, New York

[E57] Èzrohi, I. A.(1957): General forms of the remainder terms of linear formulæ in multidimensional approximate analysis II. Mat. Sb. (in Russian) **43(85)**, 9–28

[F74] Fraeijs de Veubeke, B.(1974): Variational principles and the patch test. Intern. J. Numer. Methods Eng. **8**, 783–801

[Go77] Gout, J. L.(1977): Estimation de l'erreur d'interpolation d'Hermite dans $\mathbb{R}^n$. Numer. Math. **28**, 407–429

[G75] Gregory, J. A.(1975): Error bounds for linear interpolation on triangles. In: J. Whiteman, ed, Mathematics of Finite Elements and Applications, 163–170. Academic Press, London

[K32] Kowalewski, G.(1932): Interpolation und genäherte Quadratur. B. G. Teubner, Berlin

[M75] Meinguet, J.(1975): Realistic estimates for generic constants in multivariate pointwise approximation. In: J. J. H. Miller, ed, Topics in Numerical Analysis II, 89–107. Academic Press, London

[M77] Meinguet, J.(1977): Structure et estimations de coefficients d'erreurs. Rev. Française Automat. Informat. Rech. Opér., Anal. Numer. **11**, 355–368

[M78] Meinguet, J.(1978): A practical method for estimating approximation errors in Sobolev spaces. In: D. C. Handscomb, ed, Multivariate Approximation, 169–187. Academic Press, London

[M79] Meinguet, J.(1979): A convolution approach to multivariate representation formulas. In: W. Schempp and K. Zeller, eds, Multivariate Approximation Theory, 198–210. Birkhäuser, Basel

[M81] Meinguet, J.(1981): From Dirac distributions to multivariate representation formulas. In: Z. Ziegler, ed, Approximation Theory and Applications, 225–248. Academic Press, New York

[M82] Meinguet, J.(1982): Sharp "a priori" error bounds for polynomial approximation in Sobolev spaces. In: W. Schempp and K. Zeller, eds, Multivariate Approximation Theory II, 255–274. Birkhäuser, Basel

[M84] Meinguet, J.(1984): A practical method for obtaining a priori error bounds in pointwise and mean-square approximation problems. In: S. P. Singh, J. H. W. Burry, and B. Watson, eds, Approximation Theory and Spline Functions, 97–125. Reidel, Dordrecht

[Mi80] Micchelli, C. A.(1980): A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B-splines and Lagrange interpolation. Rocky Mountain J. Math. **10**, 485–497

[N72] Nicolaides, R. A.(1972): On a class of finite elements generated by Lagrange interpolation. SIAM J. Numer. Anal. **9(3)**, 435–445

[N73] Nicolaides, R. A.(1973): On a class of finite elements generated by Lagrange interpolation II. SIAM J. Numer. Anal. **10(1)**, 182–189

[S48] Sard, A.(1948): Integral representations of remainders. Duke Math. J. **15**, 333–345

[S63] Sard, A.(1963): Linear approximation. Math. Survey 9, AMS,

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

[$W96_1$] Waldron, S.(199x): Integral error formulæ for the scale of mean value interpolations which includes Kergin and Hakopian interpolation. Numer. Math. **xx**, xxx–xx

[$W96_2$] Waldron, S.(1997): A multivariate form of Hardy's inequality and $L_p$-error bounds for multivariate Lagrange interpolation schemes. SIAM J. Math. Anal. **xx**, xxx–xxx

[$W96_3$] Waldron, S.(1996): The error in linear interpolation at vertices of a simplex. submitted

[WTDG73] Wilson, E. L., R. L. Taylor, W. P. Doherty, and J. Ghaboussi(1973): Incompatible displacement models. In: S. J. Fenves, N. Perrone, A. R. Robinson, W. C. Schnobrich, eds, Numerical and Computer Methods in Structural Mechanics, 43–57. Academic Press, New York

## 8. Some details not included in the paper

**1.** Direct proof of (7.5). First, suppose that $f \in \Pi_k$, then $q(D)f \in \Pi_{k-d}$, and hence it is obvious that

$$q(D)(\mathcal{T}_{k,x} f) = \mathcal{T}_{k-d}(q(D)f), \tag{8.1}$$

Since both sides of (8.1) depend only on the derivatives up to order $k$ of $f$ at $x$, it holds for all (smooth) $f$ (giving (7.5)).

**2.** Details about the convergence of the integrals of operators used in the proof of Theorem 6.5.

**3.** Proof of Proposition 1.5. Let $\mathcal{V} : \operatorname{ran}\mathcal{U} \to X$ be a linear right inverse, then

$$\mathcal{P}^2 = (1 - \mathcal{V}\mathcal{U})^2 = 1 - \mathcal{V}\mathcal{U} - \mathcal{V}\mathcal{U} + \mathcal{V}(\mathcal{U}\mathcal{V})\mathcal{U} = 1 - \mathcal{V}\mathcal{U} = \mathcal{P},$$

so that $\mathcal{P}$ is a linear projector onto $\operatorname{ran}\mathcal{P}$. If $x \in \ker\mathcal{U}$, then $\mathcal{P}x = x - \mathcal{V}\mathcal{U}x = x$, so $\ker\mathcal{U} \subset \operatorname{ran}\mathcal{P}$, and if $\mathcal{P}x = x - \mathcal{V}\mathcal{U}x \in \operatorname{ran}\mathcal{P}$, then $\mathcal{U}(x - \mathcal{V}\mathcal{U}x) = 0$, so $\operatorname{ran}\mathcal{P} \subset \ker\mathcal{U}$. Thus, $\operatorname{ran}\mathcal{P} = \ker\mathcal{U}$.

Conversely, suppose that $\mathcal{P} : X \to \ker\mathcal{U}$ is a linear projector onto $\ker\mathcal{U}$. Since $\ker(1-\mathcal{P}) = \ker\mathcal{U}$, the rule

$$\mathcal{V}(\mathcal{U}x) := x - \mathcal{P}x$$

defines a linear map $\mathcal{V} : \operatorname{ran}\mathcal{U} \to X$, which satisfies

$$\mathcal{U}\mathcal{V}(\mathcal{U}x) = \mathcal{U}(x - \mathcal{P}x) = \mathcal{U}x,$$

i.e., $\mathcal{U}\mathcal{V} = 1$ ($\mathcal{V}$ is a right inverse for $\mathcal{U}$).

**4.** Comment removed from the end of Section 7:

These connections are investigated in more detail in a future paper [W96$_4$]. There, error formulæ for one–point interpolation from a number of $D$–invariant spaces, including those spanned by monomials, (and the corresponding multipoint one–point formulæ) are given. There is a discussion about what derivatives should occur in these formulæ, which for interpolation to $\Theta$ is related to the construction of an $H$–basis for the ideal of polynomials vanishing at $\Theta$.

## Extra references not included in article


[At77] Atteia, M.(1977): Evaluation de l'error dans la méthode des élements finis. Numer. Math. **28**, 295–306

[BG76$_2$] Barnhill, R. E. and J. A. Gregory(401–408): Interpolation remainder theory from Taylor expansions on triangles. Numer. Math. **25**, 1976

[Bi03] Biermann, O.(1903): Über näherungsweise Cubaturen. Monatsh. für Math. Phys. **XIV**, 211–225

[CD90] Chen, W. and Z. Ditzian(1990): Mixed and directional derivatives. Proc. Amer. Math. Soc. **108**, 177–185[Norms of symmetric multilinear mappings]

[C78] Ciarlet, P. G.(1978): The finite element method for elliptic problems. North-Holland, Amsterdam

[Co66] Coatmélec, M. C.(1966): Approximation et interpolation des fonctions différentiables de plusiers variables. Ann. Sci. École Norm. Sup. **83**, 271–341

[D83] Durán, R. G.(1983): On polynomial approximation in Sobolev spaces. SIAM J. Numer. Anal. **29(5)**, 985–988

[FN74] Feinermann, R. P. and D. J. Newman(1974): Polynomial Approximation. Williams and Wilkins Co., Baltimore

[Ha28] Hardy, G. H.(1928): Notes on some points in the integral calculus LXIV. Messenger of Math. **57**, 12–16[Hardy's inequality]

[LK74] Leaf, G. K. and H. G. Kaper(1974): $L^\infty$-error bounds for multivariate Lagrange interpolation. SIAM J. Numer. Anal. **11(2)**, 363–381

[M67] Meinguet, J.(1967): Optimal approximation and error bounds in seminormed spaces. Numer. Math. **10**, 370–388

[MD77] Meinguet, J. and J. Descloux(1977): An operator-theoretical approach to error estimation. Numer. Math. **27**, 307–326

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

[N20] Narumi, S.(1920): Some formulas in the theory of interpolation of many independent variables. T̂hoku Math. J. **18**, 309–xxx

[St64] Stancu, D. D.(1964): The remainder of certain linear approximation formulas in two variables. SIAM J. Numer. Anal. Ser. B **1**, 137–163

[Ste27] Steffensen, J. F.(1927): Interpolation. Williams and Wilkins Co., Baltimore

[W96$_4$] Waldron, S.(199x): Error formulæ for generalised Taylor interpolation. preprint

[Wat79] Watkins, D. S.(1979): A generalisation of the Bramble–Hilbert lemma and applications to multivariate interpolation. J. Approx. Theory **26**, 219–231