Minimally supported error representations and approximation by the constants

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

Distribution theory is used to construct *minimally supported* Peano kernel type representations for linear functionals such as the error in multivariate Hermite interpolation. The simplest case is that of representing the error in approximation to f by the constant polynomial $f(a)$ in terms of integrals of the first order derivatives of f. This is discussed in detail. Here it is shown that suprisingly there exist many representations which are not minimally supported, and involve the integration of first order derivatives over multidimensional regions. The distance of smooth functions from the constants in the uniform norm is estimated using our representations for the errror.

1. Introduction

The motivation of this paper is the following general problem. Suppose that P is a linear projector onto $\Pi_k := \Pi_k(\mathbb{R}^d)$ (the polynomials of degree $\leq k$ in $\mathbb{R}^d$) such as Lagrange interpolation at some suitable set of points in $\mathbb{R}^d$. Let $D^{k+1}f$ denote the $(k+1)$th derivative of f, which can be represented by

$$(D^\alpha f)_{|\alpha|=k+1}$$

the sequence of $(k+1)$–order partial derivatives of f. Since the kernel of D^{k+1} is Π_k, it is possible to express the error in approximation by P at $x \in \mathbb{R}^d$ in the form

$$f(x) - Pf(x) = \mathcal{R}_x(D^{k+1}f), \tag{1.1}$$

for sufficiently smooth f, where $\mathcal{R}_x$ is some linear functional defined on the $(k+1)$th derivatives. This can be viewed as a nonconstructive multivariate version of Peano's kernel theorem. Recently there has been interest in obtaining explicit representations of $\mathcal{R}_x(D^{k+1}f)$ as a sum of integrals of $(k+1)$th derivatives of f over various regions in $\mathbb{R}^d$ from which estimates of the error $f - Pf$ in terms of some appropriate norm on $D^{k+1}f$ can be obtained, and even more general forms of this problem (see, e.g., de Boor [2], Waldron [12]).

This paper uses distribution theory to analyse representations like (1.1), with particular emphasis on the simplest case when $k = 0$

$$Pf := f(a), \quad a \in \mathbb{R}^d.$$

In Section 2, the 'support' of a representation of $\mathcal{R}_x(D^{k+1}f)$ is defined, and a family of representations which are 'minimally supported' is constructed. In Section 3, another family of minimally supported error representations for the case $Pf := f(a)$ is given, and these are used to show that there exist representations which involve integrals of first order derivatives over a wide variety of multidimensional regions. This striking example shows that even in the most simple cases when one extends the Peano kernel theory from the univariate to multivariate setting there is a large choice in the derivatives that appear and the regions over which they are integrated. In Section 4, these representations are used to estimate the L_∞–distance of smooth functions from the constants, and L_p–estimates are discussed.

2. Minimally supported error representations

Some basic results and notation from distribution theory will be used (see, e.g., Barros–Neto [3]). Suppose that the error in approximation to f by Pf can be expressed as

$$f(x) - Pf(x) = \sum_i \mathcal{R}_i(D^{k+1}f), \qquad \forall f \in C^{k+1}(\mathbb{R}^d), \tag{2.1}$$

1

where $\{\mathcal{R}_i\}$ is a finite collection of linear functionals for which $\mathcal{R}_i \circ D^{k+1}$ is continuous on $C^{k+1}(\mathbb{R}^d)$, i.e., in conventional distribution terminology $\mathcal{R}_i \circ D^{k+1}$ is a compactly supported distribution of order $k+1$, or, for short, $\mathcal{R}_i \circ D^{k+1} \in \mathcal{E}'_{k+1}(\mathbb{R}^d)$. The **support** of $\mathcal{R}_i$ is defined to be the support of the distribution $\mathcal{R}_i \circ D^{k+1}$, i.e, the compact set

$$\operatorname{supp} \mathcal{R}_i := \operatorname{supp}(\mathcal{R}_i \circ D^{k+1}),$$

and the **support of the representation** (2.1) is the compact set

$$K := \bigcup_i \operatorname{supp} \mathcal{R}_i := \bigcup_i \operatorname{supp}(\mathcal{R}_i \circ D^{k+1}). \tag{2.2}$$

A representation (2.1) with support K is said to be **minimally supported** if there is no other representation with support properly contained within K.

Consider first the case

$$f(x) - f(a) = \sum_i \mathcal{R}_i(Df), \qquad \forall f \in C^1(\mathbb{R}^d). \tag{2.3}$$

If a distribution is expressed as a finite sum

$$\lambda = \sum_i \lambda_i, \qquad \lambda_i \in \mathcal{D}'(\mathbb{R}^d),$$

then it can easily be shown that

$$\operatorname{supp} \lambda \subset \bigcup_i \operatorname{supp} \lambda_i. \tag{2.4}$$

In particular, the support of a representation of the form (2.3) must contain $\{a, x\}$. In addition, it can be shown that this support must be connected, and so the representation

$$f(x) - f(a) = \int_0^1 D_{x-a} f(a + t(x - a)) \, dt, \tag{2.5}$$

involving integration of $D_{x-a}f$ over the line segment joining a and x is minimally supported on this line segment. We now prove a general form of this result. The differential operator induced by $q \in \Pi_k$ is written $q(D)$, and the derivative of f in the direction $\xi \in \mathbb{R}^d$ as $D_\xi f$.

Theorem 2.6 (minimally supported representations). *Let $\lambda \in \mathcal{D}'(\mathbb{R}^d)$ be a distribution of the form*

$$\lambda(f) := \sum_{j=1}^m q_j(D) f(a_j), \qquad \forall f \in C^{k+1}(\mathbb{R}^d), \tag{2.7}$$

where $0 \neq q_j \in \Pi_k$, and $a_j \in \mathbb{R}^d$ are distinct points. If λ vanishes on Π_k, then there exist a finite number of linear maps $\mathcal{R}_i$ with $\mathcal{R}_i \circ D^{k+1} \in \mathcal{E}'_{k+1}(\mathbb{R}^d)$, such that

$$\lambda(f) = \sum_i \mathcal{R}_i(D^{k+1}f), \qquad \forall f \in C^{k+1}(\mathbb{R}^d). \tag{2.8}$$

2

The support of any representation of the form (2.8), i.e., the compact set

$$K := \bigcup_i \operatorname{supp}(\mathcal{R}_i \circ D^{k+1}), \qquad (2.9)$$

contains the points $\{a_j\}$. Suppose in addition that no (nontrivial) proper subsum of (2.7) vanishes for all $f \in \Pi_k$. Then the support K is connected, and for any $c \in \mathbb{R}^d$, it is possible to choose $\mathcal{R}_i$ so that K is the union of the line segments from c to the a_j, thereby obtaining a minimally supported representation (2.8).

Proof: *Properties.* If a representation (2.8) exists, then it follows from (2.4) that

$$\operatorname{supp} \lambda = \{a_j\} \subset K.$$

Suppose that K is not connected, and let K_1 be the component of K containing a_1. Then the condition that no subsum of (2.7) vanishes for all $f \in \Pi_k$ implies that there exists a polynomial $p \in \Pi_k$ with

$$\sum_{a_j \in K_1} q_j(D)p(a_j) \neq 0. \qquad (2.10)$$

Let $f \in \mathcal{D}(\mathbb{R}^d)$ be a function which is equal to p on a small neighbourhood of K_1 and is zero on a small neighbourhood of the other components of K. Then $D^{k+1}f = 0$ on a neighbourhood of K, which implies

$$\sum_i \mathcal{R}_i(D^{k+1}f) = 0,$$

whilst

$$\lambda(f) = \sum_{a_j \in K_1} q_j(D)f(a_j) + \sum_{a_j \in K \setminus K_1} q_j(D)f(a_j) = \sum_{a_j \in K_1} q_j(D)p(a_j) \neq 0,$$

which contradicts (2.8). Hence K must be connected.

Existence. We now construct minimally supported representations (2.8). This is done by using a formula of the form (2.8) for the error at $x \in \mathbb{R}^d$ in approximation by $\mathcal{T}_{k,c}$ (the operator of Taylor interpolation from Π_k at $c \in \mathbb{R}^d$) to convert (2.7) into such a representation. For $q \in \Pi_\ell^0(\mathbb{R}^d)$, $0 \leq \ell \leq k$ (the homogeneous polynomials of degree ℓ)

$$q(D)(\mathcal{T}_{k,c}f) = \mathcal{T}_{k-l,c}(q(D)f),$$

and so by the univariate integral remainder theorem for Taylor interpolation

$$q(D)(f - \mathcal{T}_{k,c}f)(x) = \frac{1}{k!}\int_0^1 (1-t)^k (D_{x-c}^{k+1-\ell}q(D)f)(c+t(x-c))\,dt =: \mathcal{R}(D^{k+1}f). \qquad (2.11)$$

It is easily verified that $\mathcal{R} \circ D^{k+1} \in \mathcal{E}'_{k+1}(\mathbb{R}^d)$, and for $q \neq 0$ the support of $\mathcal{R}$ is the line segment from x to c. Express each q_j as a sum of its homogeneous terms, i.e.,

$$q_j = \sum_{\ell=0}^{k} q_j^\ell, \qquad q_j^\ell \in \Pi_\ell^0(\mathbb{R}^d). \tag{2.12}$$

Since $\lambda \in \mathcal{E}'_{k+1}(\mathbb{R}^d)$ and vanishes on Π_k, (2.7) can be expanded using (2.11) and (2.12) as follows

$$\lambda(f) = \lambda(f - \mathcal{T}_{k,c}f) = \sum_{j=1}^{m} \sum_{\ell=0}^{k} q_j^\ell(D)(f - \mathcal{T}_{k,c}f)(a_j)$$

$$= \frac{1}{k!} \sum_{j=1}^{m} \sum_{\ell=0}^{k} \int_0^1 (1-t)^k (D_{a_j-c}^{k+1-\ell} q_j^\ell(D) f)(c + t(a_j - c)) \, dt, \tag{2.13}$$

which is a representation of the form (2.8). The support of the representation (2.13) is the union of the line segments from c to the a_j. No proper subset of this support which contains $\{a_j\}$ is connected, and so (2.13) is a minimally supported representation. $\qquad\square$

Theorem 2.6 is a multivariate extension of the Peano kernel theorem for functionals of the form (2.7). Similar extensions for more general linear functionals can be obtained along the same lines. This result applies to a number of linear functionals of practical interest, such as the pointwise error in multivariate Hermite interpolation from $\Pi_k(\mathbb{R}^d)$. Here is a typical example.

Suppose that

$$\lambda(f) := f(x) - \mathcal{L}_\Theta f(x) = f(x) - \sum_{v \in \Theta} \ell_v(x) f(v), \qquad x \notin \Theta \tag{2.14}$$

is the error in linear interpolation (interpolation by linear polynomials) at the points $\Theta \subset \mathbb{R}^d$, where ℓ_v are the Lagrange polynomials (which are the barycentric coordinates). This satisfies the conditions of Theorem 2.6 with $\Pi_k = \Pi_1$. Taking $c = x$ in the expansion (2.13) gives the *multipoint Taylor formula* of Ciarlet and Wagschal [4:Th.2], while the choice $c = w \in \Theta$ gives (in the notation used in [12])

$$f(x) - \mathcal{L}_\Theta f(x) = \int_{[w,w,x]} D_{x-w}^2 f - \sum_{\substack{v \in \Theta \\ v \neq w}} \ell_v(x) \int_{[w,w,v]} D_{v-w}^2 f, \tag{2.15}$$

which is a new formula. There also exist representations of (2.14) which are not minimally supported, see, e.g., Gregory [9] or Waldron [13]. Simpler examples of representations which are not minimally supported are examined in the next section.
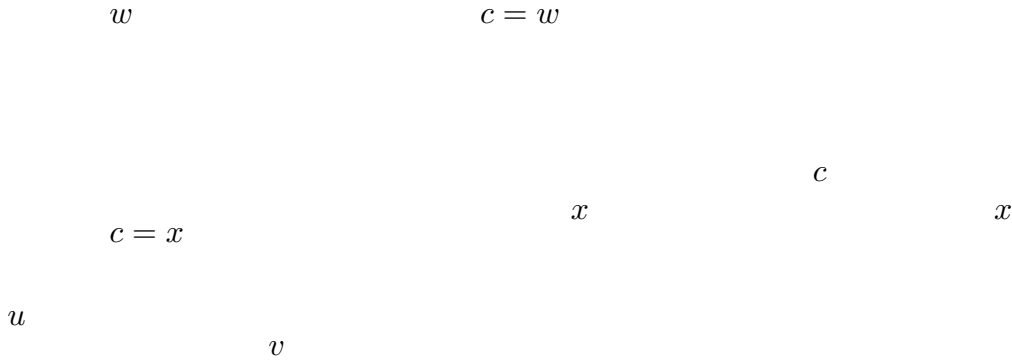
4

$$w \qquad\qquad c = w$$

$$c$$

$$x \qquad\qquad\qquad x$$

$$c = x$$

$$u$$

$$v$$

Fig. 2.1. The support K of minimally supported error formulæ for linear interpolation at the vertices of a triangle (shaded) given by (2.13). First the multipoint Taylor formula, followed by formula (2.15), and an example where $c \notin \{\Theta, x\}$.

3. Error representations for approximation by the constants

In this section we consider error representations of the form (2.3) for approximation by the constant $Pf := f(a)$, i.e., the linear functional

$$\lambda(f) := f(x) - f(a), \qquad x \neq a.$$

First we give a family of minimally supported representations for λ which contains those given by (the proof of) Theorem 2.6 as particular cases. These representations are then "averaged" to obtain ones which are supported on a variety of multidimensional regions. This example illustrates a basic difficulty in developing a constructive and useful multivariate Peano kernel theory, namely deciding

- *Which $(k + 1)$–order derivatives should appear in (2.1)?*
- *What regions should they be integrated over?*

when it is known that many such formulæ exist.

A *path* in $\mathbb{R}^d$ is a continuous, piecewise differentiable map $\gamma : [0, 1] \to \mathbb{R}^d$. Let $\mathbf{T}$ denote the unit tangent vector along the curve γ, i.e.,

$$\mathbf{T} := \frac{D\gamma}{\|D\gamma\|},$$

length(γ) the length of γ, and ds the arc length element. A restatement of the fundamental theorem of calculus shows that there is a representation (2.3) which is supported on

$$K = \gamma^* := \gamma([0, 1]),$$

where γ is any path from a to x, i.e., $\gamma(0) = a$, $\gamma(1) = x$. Here ∇f is the gradient of f.

Lemma 3.1. *Let $\gamma : [0,1] \to \mathbb{R}^d$ be a path from a to x, then*

$$\lambda(f) := f(x) - f(a) = \int_\gamma D_{\mathbf{T}} f \, ds, \qquad \forall f \in C^1(\gamma^*), \tag{3.2}$$

and in particular,

$$|f(x) - f(a)| \leq \left(\max_{\gamma^*} \|\nabla f\| \right) \text{length}(\gamma). \tag{3.3}$$

Proof:　Let $g := f \circ \gamma$, then by the fundamental theorem of calculus

$$f(x) - f(a) = g(1) - g(0) = \int_0^1 Dg = \int_0^1 (D_{\mathbf{T}} f \circ \gamma) \|D\gamma\| = \int_\gamma D_{\mathbf{T}} f \, ds,$$

and hence

$$|f(x) - f(a)| \leq \int_\gamma |D_{\mathbf{T}} f| \, ds \leq \int_\gamma \|\nabla f\| \, ds \leq \left(\max_{\gamma^*} \|\nabla f\| \right) \text{length}(\gamma).$$

$\square$

Thus, if γ has no self intersections, then Theorem 2.6 implies that (3.2) provides a minimally supported error representation. A number of these representations are well–known, and have been used in numerical analysis. For example, if γ is the line segment from a to x, then (3.2) is the univariate Taylor error formula (2.5) for approximation by a constant, while if γ is the path in $\mathbb{R}^2$ consisting of the line segment from (a,b) to (x,b) followed the line segment from (x,b) to (x,y), then (3.2) becomes

$$f(x,y) - f(a,b) = \int_a^x D_1 f(s,b) \, ds + \int_b^y D_2 f(x,t) \, dt. \tag{3.4}$$

The formula (3.4) is an example of the Taylor formula given by Sard [11:p163] for certain spaces, denoted by $B_{p,q}$, of functions defined on some rectangle I in $\mathbb{R}^2$. This example is for the space $B_{0,1} = B_{0,1}(I; a, b)$, where $(a,b) \in I$. If γ^* is the union of the line segments from c to a and x, then (3.2) is the representation (2.13).

<div style="display: flex; justify-content: space-around;">
$x$$x$$x$$x$
</div>

<div style="display: flex; justify-content: space-around;">
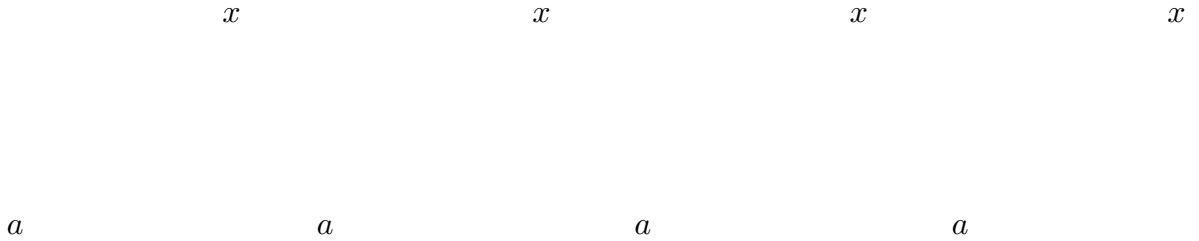$a$$a$$a$$a$
</div>

Fig. 3.1. The support $K = \gamma^*$ of some error representations given by (3.2). The first is (2.5) the 'univariate Taylor formula for approximation by a constant', and

the second is (3.4), which was used by Sard. All but the last (which has a self intersection) are minimal supports.

Given the error formulæ (3.2) of Lemma 3.1, it is not difficult to imagine taking an average (or even weighted average) of some family of them to obtain an error representation (2.3) which is supported on some multidimensional region. We illustrate this idea in the bivariate case with the construction of a representation for the error $f(x) - f(a)$ in terms of Df which is supported on the rectangle aligned in the coordinate directions with opposite vertices given by a and x. To the author at least, the existence of such a formula is suprising. In principle, an error representation supported on any compact path–connected set containing a and x could be obtained in this way.

For simplicity, suppose that $a = (0,0)$ and (x, y) is in the first quadrant. Let γ_λ be the path given by the line segments from $(0,0)$ to $\xi_\lambda := ((1-\lambda)x, \lambda y)$ and ξ_λ to (x, y). Taking the average over $\lambda_1 \le \lambda \le \lambda_2$ of the corresponding representation (3.2) gives

$$f(x, y) - f(0,0) = \frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} \int_{\gamma_\lambda} D_{\mathbf{T}} f \, ds \, d\lambda, \tag{3.5}$$

which is supported on Q the quadrilateral with vertices $(0,0)$, ξ_{λ_1}, (x, y), ξ_{λ_2} (ordered in the positive angular direction).
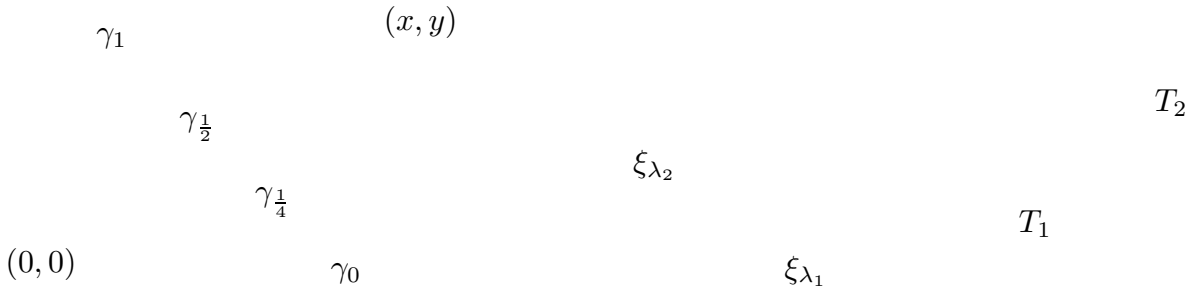

Fig. 4.1. The family of curves γ_λ ($\lambda = 0, 1/4, 1/2, 1$), followed by the support of the formula (3.5) for $0 < \lambda_1 < \lambda_2 < 1/2$, and the support $Q = T_1 \bigcup T_2$ of the formula obtained by taking $\lambda_1 = 0$, $\lambda_2 = 1$.

If $\lambda_1 = 0$, $\lambda_2 = 1$, then Q is the rectangle with vertices $(0,0)$, $(x, 0)$, (x, y), $(0, y)$. For this case we now compute the weight functions w_1, w_2 that occur when (3.5) is rewritten in the form

$$f(x, y) - f(0,0) = \int_Q w_1 \, D_1 f + \int_Q w_2 \, D_2 f. \tag{3.6}$$

Since

$$f(x, y) - f(0,0) = \{f(x, y) - f(\xi_\lambda)\} + \{f(\xi_\lambda) - f(0,0)\},$$

7

by using (2.5), (3.5) can be rewritten as

$$f(x,y) - f(0,0) = \int_0^1 \int_0^1 \Big((1-\lambda)x D_1 f(t(1-\lambda)x, t\lambda y) + \lambda y D_2 f(t(1-\lambda)x, t\lambda y) \Big) \, dt \, d\lambda$$

$$+ \int_0^1 \int_0^1 \Big(\lambda x D_1 f((1-\lambda)x + t\lambda x, \lambda y + t(1-\lambda)y)$$

$$+ (1-\lambda)y D_2 f((1-\lambda)x + t\lambda x, \lambda y + t(1-\lambda)y) \Big) \, dt \, d\lambda. \tag{3.7}$$

The first of these iterated integrals is of Df over the triangle T_1 with vertices $(0,0)$, $(x,0)$, $(0,y)$. In this integral make the change of variables:

$$u = t(1-\lambda)x, \qquad v = t\lambda y,$$

to obtain

$$\int_{T_1} \frac{xyu}{(yu+xv)^2} D_1 f(u,v) + \frac{xyv}{(yu+xv)^2} D_2 f(u,v) \, du \, dv.$$

Using a symmetry argument, this can be used to express the second iterated integral in (3.7) as an integral over T_2 the triangle with vertices $(x,0)$, (x,y), $(0,y)$. In this way, we obtain that the weight functions w_1, w_2 occuring in (3.6) are given by

$$w_1(u,v) = xy \begin{cases} \frac{u}{(yu+xv)^2}, & (u,v) \in T_1 \\ \frac{x-u}{(y(x-u)+x(v-y))^2}, & (u,v) \in T_2 \end{cases}$$

and

$$w_2(u,v) = xy \begin{cases} \frac{v}{(yu+xv)^2}, & (u,v) \in T_1 \\ \frac{y-v}{(y(x-u)+x(v-y))^2}, & (u,v) \in T_2. \end{cases}$$

These integrable functions $w_i : Q \to \mathbb{R}$ are supported on Q and continuous at every point except $(0,0)$ and (x,y), near which they are unbounded.

4. The distance of smooth functions from the constants

In this section we use the (minimally supported) error formulæ of Section 3 to obtain estimates of the $L_\infty(\Omega)$–distance of smooth functions from Π_0 (the constants). In contrast to previous results, these estimates do not require special conditions on the domain Ω, such as being convex or starshaped.

Specifically, we are interested in the best constant C_Ω in the estimate

$$\text{dist}_{L_\infty(\Omega)}(f, \Pi_0) \leq C_\Omega \|Df\|_{L_\infty(\Omega)}, \tag{4.1}$$

for all sufficiently smooth f, such as f from $W_1^\infty(\Omega)$, $C^1(\Omega)$, or $C^1(\bar{\Omega})$. This turns out to be an involved question, with C_Ω depending not only on the geometry of the domain

Ω, but also, in some cases, on the definition of sufficiently smooth used. We will deal primarily with the most technical and interesting case when $f \in W_1^\infty(\Omega)$.

Suppose that $\Omega \subset \mathbb{R}^d$ is a *domain*, i.e., is open and path–connected. Our estimates of C_Ω in (4.1) will be described in terms of d_Ω the *intrisic* (or *geodesic*) metric on Ω,

$$d_\Omega(a, x) := \text{the infimum of the lengths of the paths from } a \text{ to } x \text{ that lie in } \Omega.$$

If the line segment from a to x lies in $\bar{\Omega}$, then $d_\Omega(a, x) = \|x - a\|$, otherwise $d_\Omega(a, x) > \|x - a\|$. We define an "intrinsic" radius of Ω about a, which measures the length of the longest path needed to connect a with a point in Ω, as follows

$$r(a, \Omega) := \sup_{x \in \Omega} d_\Omega(a, x) \tag{4.2}$$

(which may be infinite, even when Ω is bounded), and let

$$r(\Omega) := \inf_{a \in \Omega} r(a, \Omega) = \inf_{x \in \Omega} \sup_{y \in \Omega} d_\Omega(x, y). \tag{4.3}$$

It follows from the triangle inequality that

$$r(\Omega) \leq \sup_{x,y \in \Omega} d_\Omega(x, y) \leq 2r(\Omega). \tag{4.4}$$

Let $\text{diam}\,\Omega$ be the (Euclidean) diameter of Ω. It can easily be shown that

$$\frac{1}{2}\,\text{diam}\,\Omega \leq r(\Omega) \leq r(a, \Omega), \tag{4.5}$$

and that if Ω is starshaped with respect to a point a, then

$$r(\Omega) \leq r(a, \Omega) \leq \text{diam}\,\Omega. \tag{4.6}$$

x

$d_\Omega(a, x)$

$2r(\Omega)$

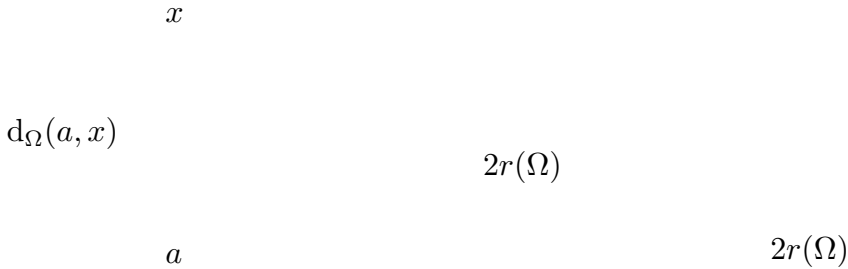a $2r(\Omega)$

Fig 4.2 The distance $d_\Omega(a, x)$ for a domain Ω (shaded), followed by the distance $2r(\Omega)$ for a convex and nonconvex bounded domain Ω.

For any domain Ω, the Sobolev space $W^1_\infty(\Omega)$ consists of the bounded *locally Lipschitz* functions on Ω, i.e., those bounded functions f satisfying an inequality

$$|f(x) - f(y)| \leq M \|x - y\|, \qquad M < \infty, \tag{4.7}$$

for all points x, y for which $[x, y]$ the line segment from x to y is contained in Ω. It is easily shown that the seminorm on $W^1_\infty(\Omega)$ defined by

$$\|Df\|_{L_\infty(\Omega)} := \text{the } L_\infty(\Omega)\text{–norm of } \|\nabla f\| \tag{4.8}$$

coincides with the smallest constant M in the inequality (4.7), i.e.,

$$M := \sup_{\substack{[x,y] \subset \Omega \\ x \neq y}} \frac{|f(x) - f(y)|}{\|x - y\|}. \tag{4.9}$$

We are now able to give the main result of this section.

Theorem 4.10 (approximation by constants). *Suppose that $\Omega \subset \mathbb{R}^d$ is open and path–connected. Then there exists a finite constant C_Ω for which*

$$\text{dist}_{L_\infty(\Omega)}(f, \Pi_0) := \inf_{c \in \Pi_0} \|f - c\|_{L_\infty(\Omega)} \leq C_\Omega \|Df\|_{L_\infty(\Omega)}, \qquad \forall f \in W^1_\infty(\Omega), \tag{4.11}$$

if and only if

$$\sup_{x,y \in \Omega} d_\Omega(x, y) < \infty. \tag{4.12}$$

The best such constant satisfies

$$\frac{1}{4} \text{diam}\, \Omega \leq \frac{1}{2} r(\Omega) \leq \frac{1}{2} r(a, \Omega) \leq C_\Omega \leq r(\Omega), \qquad \forall a \in \Omega. \tag{4.13}$$

For $d > 1$, the quotient $C_\Omega / \text{diam}\, \Omega$ can be arbitrarily large for a domain satisfying (4.12). However, for Ω convex

$$C_\Omega = \frac{1}{2} \text{diam}\, \Omega, \tag{4.14}$$

and for Ω starshaped

$$\frac{1}{4} \text{diam}\, \Omega \leq C_\Omega \leq \text{diam}\, \Omega. \tag{4.15}$$

Proof: First, we prove the lower bound on C_Ω, and the necessity of the condition (4.12). Fix $a \in \Omega$. Let

$$f := d_\Omega(a, \cdot), \tag{4.16}$$

which is bounded if and only if $r(a, \Omega) < \infty$, and (by the triangle inequality) satisfies

$$M := \sup_{\substack{[x,y] \subset \Omega \\ x \neq y}} \frac{|f(x) - f(y)|}{\|x - y\|} = 1. \tag{4.17}$$

10

Thus, if (4.12) holds, then $f \in W_\infty^1(\Omega)$ and satisfies

$$\inf_\Omega f = f(a) = 0, \qquad \sup_\Omega f = r(a,\Omega) < \infty, \qquad \|Df\|_{L_\infty(\Omega)} = 1,$$

so that

$$\frac{\text{dist}_{L_\infty(\Omega)}(f, \Pi_0)}{\|Df\|_{L_\infty(\Omega)}} = \frac{1}{2} r(a,\Omega) \leq C_\Omega,$$

which together with (4.5) gives the lower bounds of (4.13). Suppose that (4.12) does not hold, i.e.,

$$\sup_{x,y \in \Omega} \mathrm{d}_\Omega(x,y) = \infty, \tag{4.18}$$

then f may be unbounded, and so not belong to $W_\infty^1(\Omega)$. In this case, we slightly modify f so that it belongs to $W_\infty^1(\Omega)$, and thereby show that no finite constant can exist in (4.11). By (4.18), we can choose $a, x \in \Omega$ for which

$$\mathrm{d}_\Omega(a,x) \geq N,$$

for any $N > 0$. Let f be defined by

$$f(y) := \min\{\mathrm{d}_\Omega(a,y), N\}.$$

This f satisfies (4.17), and

$$\inf_\Omega f = f(a) = 0, \qquad \sup_\Omega f = N, \qquad \|Df\|_{L_\infty(\Omega)} = 1,$$

so it belongs to $W_\infty^1(\Omega)$, and we obtain

$$\frac{\text{dist}_{L_\infty(\Omega)}(f, \Pi_0)}{\|Df\|_{L_\infty(\Omega)}} = \frac{1}{2} N \leq C_\Omega.$$

Hence, when (4.12) fails there exists no finite constant in (4.11).

Next, we prove the upper bound on C_Ω by using (3.3) for $f \in W_\infty^1(\Omega)$. Suppose that (4.12) holds, and so, by (4.4), $r(\Omega) < \infty$. For $a, x \in \Omega$, let γ be a path in Ω from a to x. Since $f \circ \gamma$ is a Lipschitz function, (3.2) holds, and from it we obtain that

$$|f(x) - f(a)| \leq \int_\gamma |D_{\mathbf{T}} f| \, ds \leq \int_\gamma \|Df\|_{L_\infty(\gamma^*)} \, ds \leq \|Df\|_{L_\infty(\Omega)} \, \text{length}(\gamma). \tag{4.19}$$

Taking the infimum of the righthand side of (4.19) over all possible γ then gives

$$|f(x) - f(a)| \leq \|Df\|_{L_\infty(\Omega)} \, \mathrm{d}_\Omega(a,x),$$

so that

$$\|f - f(a)\|_{L_\infty(\Omega)} \leq \|Df\|_{L_\infty(\Omega)} \, r(a,\Omega). \tag{4.20}$$

Hence by (4.20)

$$\operatorname{dist}_{L_\infty(\Omega)}(f, \Pi_0) \leq \inf_{a \in \Omega} \|f - f(a)\|_{L_\infty(\Omega)} \leq \|Df\|_{L_\infty(\Omega)} \, r(\Omega),$$

giving the upper bound $C_\Omega \leq r(\Omega)$.

The remaining estimates of C_Ω follow from (4.5) and (4.6). For example, if Ω is convex, then

$$C_\Omega \leq r(\Omega) = \frac{1}{2} \operatorname{diam} \Omega,$$

while choosing $a \in \Omega$ with $\sup_{x \in \Omega} \|x - a\| = \operatorname{diam} \Omega - \varepsilon$, $\varepsilon > 0$, gives

$$\frac{1}{2}(\operatorname{diam} \Omega - \varepsilon) = \frac{1}{2} r(a, \Omega) \leq C_\Omega.$$

Finally, we show for $d > 1$ that $C_\Omega / \operatorname{diam} \Omega$ can be arbitrarily large. Let γ be some path of length L (which can be as large as desired) contained in some ball of diameter 1, e.g., a suitably tight spiral. For $\varepsilon > 0$, let Ω be a domain obtained from γ by taking the union over $a \in \gamma^*$ of open balls of centre a and radius $< \varepsilon$ chosen so that

$$L - \varepsilon < \sup_{x,y \in \Omega} \, \mathrm{d}_\Omega(x, y) < L + \varepsilon.$$

For this domain, it follows from (4.13) and (4.4) that

$$\frac{C_\Omega}{\operatorname{diam} \Omega} \geq \frac{r(\Omega)/2}{\operatorname{diam} \Omega} \geq \frac{\sup_{x,y \in \Omega} \mathrm{d}_\Omega(x, y)/4}{\operatorname{diam} \Omega} \geq \frac{(L - \varepsilon)/4}{1 + 2\varepsilon},$$

which for suitably large L and small ε can be as large as desired. $\qquad\square$

The best constant C_Ω in (4.1) also depends on the definition of sufficiently smooth used, as we now show by example.

Let $\Omega := (0, 1) \times (-\delta, \delta) \subset \mathbb{R}^2$, for small $\delta > 0$. For f from $W_\infty^1(\Omega)$ and $C^1(\bar{\Omega})$ the best constant C_Ω is the same, and close to $1/2$ (approximate f by $f(1/2, 0)$ to estimate C_Ω from above, and use $f(x) := x$ to get the lower bound). Now make a slit along $(\delta, 1) \times \{0\}$ to obtain $\Omega^* := \Omega \setminus (\delta, 1) \times \{0\}$. This effectively gives a strip of length 2 instead of 1 as previously. For f from $W_\infty^1(\Omega)$ the approximation $f(\delta/2, 0)$ is nearly optimal, and multiplies the previous constant by almost 2, i.e., $C_{\Omega^*} \approx 2C_\Omega$. But for f from $C^1(\bar{\Omega}^*) = C^1(\bar{\Omega})$ the best constant remains the same, i.e., $C_{\Omega^*} = C_\Omega$.

Remarks on L_p–estimates

From the beginnings of the theory of Sobolev spaces it has been known that for a suitably regular domain Ω there exists a finite constant $C_{p,\Omega}$ for which

$$\operatorname{dist}_{L_p(\Omega)}(f, \Pi_0) := \inf_{c \in \Pi_0} \|f - c\|_{L_p(\Omega)} \leq C_{p,\Omega} \, \|Df\|_{L_p(\Omega)}, \qquad \forall f \in W_p^1(\Omega), \qquad (4.21)$$

where $\|Df\|_{L_p(\Omega)}$ is some seminorm measuring the $L_p(\Omega)$–size of Df. The argument for this, which is by contradiction (see Morrey [10:Th.3.6.11]) , does not give any information about how the best such constant $C_{p,\Omega}$ depends on the geometry of the domain Ω (or on p). From now on, let

$$\|Df\|_{L_p(\Omega)} := \text{the } L_p(\Omega)\text{–norm of } \|\nabla f\|.$$

There have been many constructive estimates of $C_{p,\Omega}$ (and similar constants), see, e.g., Dupont and Scott [7], or Dahmen, DeVore and Scherer [5]. For the specific case (4.21), the best of these is the following. Suppose that $p > d$ and Ω is starshaped with respect to a. By a slight refinement of the estimate of Arcangeli and Gout [1:Th.1-1] one obtains

$$\text{dist}_{L_p(\Omega)}(f,\Pi_0) \leq \|f - f(a)\|_{L_p(\Omega)} \leq \frac{r(a,\Omega)}{1 - d/p}\, \|Df\|_{L_p(\Omega)}, \qquad \forall f \in W_p^1(\Omega), \qquad (4.22)$$

and Dechevski and Quak [6:p491] give the bound

$$\text{dist}_{L_p(\Omega)}(f,\Pi_0) \leq \|f - f(a)\|_{L_p(\Omega)} \leq \frac{\text{diam}\,\Omega}{1 - d/p} \sum_{i=1}^{d} \|D_i f\|_{L_p(\Omega)}, \qquad \forall f \in W_p^1(\Omega). \quad (4.23)$$

From (4.13) it is seen when $p = \infty$ the dependence of the best constant $C_{p,\Omega}$ on Ω is in terms of the equivalent quantities

$$r(\Omega) := \inf_{a \in \Omega} \sup_{x \in \Omega} \mathrm{d}_\Omega(a,x), \qquad d(\Omega) := \sup_{x,y \in \Omega} \mathrm{d}_\Omega(x,y),$$

rather than on $\text{diam}\,\Omega$. Presumably this behaviour also holds to some extent for other values of p (particularly large p). One might conjecture that for $1 \leq p \leq \infty$, there exist constants $A_{n,p}$ and $B_{n,p}$ (depending only on n and p), for which

$$A_{n,p}\, r(\Omega) \leq C_\Omega \leq B_{n,p}\, r(\Omega),$$

for all sufficiently regular domains Ω. It is not at all obvious how the proof of Theorem 4.10 could be modified to obtain such a result. However, for $f := \mathrm{d}_\Omega(a,\cdot)$ one does have

$$\|Df\|_{L_p(\Omega)} = (\text{measure of } \Omega)^{1/p},$$

and so, for the lower bound at least, there is some hope of obtaining at least a partial result.

Inequalities of the form (4.21) are closely related to several classical *Sobolev inequalities* from the theory of partial differential equations. These include *Friedrichs' inequality*

$$\|f - f_\Omega\|_{L_p(\Omega)} \leq C\, \|Df\|_{L_p(\Omega)}, \qquad \forall f \in W_p^1(\Omega),$$

where f_Ω is the constant function given by

$$f_\Omega := \frac{1}{\text{meas}\,\Omega} \int_\Omega f,$$

(it can easily be shown that $\|f - f_\Omega\|_{L_p(\Omega)} \leq 2\, \text{dist}_{L_p(\Omega)}(f,\Pi_0)$), and the *Poincaré inequality*

$$\|f\|_{L_p(\Omega)} \leq C\, \|Df\|_{L_p(\Omega)}, \qquad \forall f \in C_0^\infty(\Omega).$$

Most of the work on these involves determining conditions on the domain Ω that are necessary and sufficient for these inequalities to exist for an appropriately defined Sobolev space, see, e.g., Edmonds and Opic [8] where weighted Sobolev spaces are considered.

Acknowledgement

I would like to thank Pavel Shvartsman and Jean–Pierre Rosay for some helpful and interesting discussions related to Section 4.

References

[1] 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

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

[3] Barros–Neto, J.(1973): An introduction to the theory of distributions. Marcel Dekker, New York

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

[5] Dahmen, W., R. DeVore, and K. Scherer(1980): Multidimensional spline approximations. SIAM J. Numer. Anal. **17**, 380–402

[6] Dechevski, L. T. and E. Quak(1990): On the Bramble-Hilbert lemma. Numer. Func. Anal. Optim. **11 (5&6)**, 485–495

[7] Dupont, T. and R. Scott(1980): Polynomial approximation of functions in Sobolev spaces. Math. Comp. **34(150)**, 441–463

[8] Edmonds, D. E. and B. Opic(1993): Weighted Poincaré and Friedrichs inequalities. J. London Math. Soc. **47**, 79–96

[9] 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

[10] Morrey, C. B.(1966): Multiple integrals in the calculus of variations. Springer-verlag, Berlin

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

[12] Waldron, S.(1998): Multipoint Taylor formulæ. Numer. Math. **80(3)**, 461–494

[13] Waldron, S.(1998): The error in linear interpolation at vertices of a simplex. SIAM J. Numer. Anal. **35(3)**, 1191–1200