Acta Numerica (1993), *pp.* 65–109

Multivariate piecewise polynomials

C. de Boor*
Center for Mathematical Sciences
University of Wisconsin-Madison
Madison WI 53705 USA
E-mail: deboor@cs.wisc.edu

CONTENTS

1. Introduction

This article was supposed to be on 'multivariate splines'. An informal survey, taken recently by asking various people in Approximation Theory what they consider to be a 'multivariate spline', resulted in the answer that a multivariate spline is a possibly smooth, piecewise polynomial function of several arguments. In particular, the potentially very useful thin-plate spline was thought to belong more to the subject of radial basis functions than in the present article. This is all the more surprising to me since I am convinced that the variational approach to splines will play a much greater role in multivariate spline theory than it did or should have in the univariate theory. Still, as there is more than enough material for a survey of multivariate piecewise polynomials, this article is restricted to this topic, as is indicated by the (changed) title.

* Supported by the United States Army, the National Science Foundation, and the Alexander von Humboldt Stiftung.

The available material concerning the space

$$\Pi^\rho_{k,\Delta} = \Pi^\rho_{k,\Delta}(\mathbb{R}^d)$$

of all **pp** (:= piecewise polynomial) functions in $C^{(\rho)}(\mathbb{R}^d)$ of degree $\le k$ with some partition Δ is quite vast, as is evidenced by the bibliography Franke and Schumaker (1987) (which contains over 1100 items, yet, e.g., only skims the available engineering literature on finite elements) and the supplementary bibliographies in Schumaker (1988, 1991). This means that, in an article such as this, it is only possible to sketch some of the ideas underlying some of the recent developments in this area.

After a short section on notation, the major topics addressed here are:

(i) the BB-form;
(ii) the dimension of $\Pi^\rho_{k,\Delta}$;
(iii) polyhedral splines;
(iv) the Strang-Fix condition;
(v) upper bounds for the approximation power of $\Pi^\rho_{k,\Delta}$.

Of these, the **BB-** (:= Bernstein-Bézier-) form is perhaps the most immediately useful. Although approximation theorists became aware of it (through the work of Farin and others in CAGD) in the early 1980's, it should be much better known. For example, people in Finite Elements could benefit greatly from its use. For this reason, I am giving a rather leisurely introduction to it, in the generality of functions of several (rather than just one or two) variables.

The second topic, the dimension of $\Pi^\rho_{k,\Delta}$, has been a major topic since Strang published some conjectures concerning the bivariate case. It turned out to be a hard problem, perhaps solvable only for 'generic' partitions if at all. However, it gives me the opportunity to illustrate further the use of the BB-form in the process of indicating the difficulty of the problem.

Much effort has been expended in the last 15 years to understand and make use of polyhedral splines, especially simplex splines and box splines. These are multivariate generalizations of Schoenberg's highly successful univariate B-spline. Although some beautiful mathematics has been, and is still being, generated in pursuit of a better understanding, these multivariate B-splines have not yet become standard tools for approximation. However (or, perhaps, because of this), it is important to be aware of the basic idea underlying them, if only because it is the only general principle available at present for the construction of compactly supported pp functions of two or more arguments of degree $\le k$ and in $C^{(\rho)}$ for ρ 'near' k. Also, the recent introduction, by Dahmen, Micchelli and Seidel, of what looks in hindsight to be the 'right' construction principle for a basis of simplex splines suitable for a given triangulation, awakens new hope for the ultimate usefulness of polyhedral splines.

The Strang-Fix condition (as it is called in Approximation Theory) relates the approximation power of the space spanned by the integer translates of some compactly supported function φ to the behavior of its Fourier transform $\widehat{\varphi}$ 'at' the discrete set $2\pi\mathbb{Z}^d\backslash 0$. Since its formulation in the early 1970's as the result of a mathematical analysis of the Finite Element Method, it has been the main tool for the determination of approximation orders for shift-invariant pp spaces (such as those generated from box splines, or those on regular partitions). Recent understanding of the structure of shift-invariant spaces has led to a better understanding of what underlies the Strang-Fix condition.

The last section provides a simple discussion of the basic technique for determining upper bounds for the approximation power of a pp space.

The omission of any discussion of *parametric* pp functions, such as curves and surfaces, is likely to be remedied by an entire article on this topic, perhaps in the next volume of this journal. It is to be hoped that another major omission in the context of splines, the discussion of thin-plate splines and other radial functions, will be similarly remedied. Finally, the discussion of numerical methods for approximation by multivariate pp functions is better postponed to a time when these are better understood.

Incidentally, with the exception of numerical methods and, perhaps, the dimension question, none of the topics mentioned (as being discussed or omitted here) appears in the early survey Birkhoff and de Boor (1965) on piecewise polynomial interpolation and approximation.

Finally, a comment concerning the term 'multivariate'. To the annoyance and confusion of statisticians, the term 'multivariate' has become standard in Approximation Theory for what statisticians (and, perhaps, others) would call 'multivariable'. It is too late to change this.

2. Polynomials

The collection of all polynomials in d arguments is denoted here by

$$\Pi \;=\; \Pi(\mathbb{R}^d).$$

For multivariate polynomials, multi-index notation is standard. A **multi-index** is, by definition, any vector with nonnegative integer entries. The **length** of such a multi-index α is the sum of its entries,

$$|\alpha| := \sum_i \alpha(i).$$

Further, $\alpha \le \beta$ iff $\alpha(i) \le \beta(i)$ for all i, and $\alpha < \beta$ iff $\alpha \le \beta$ yet $\alpha \ne \beta$.

With $x(i)$ the ith component of $x \in \mathbb{R}^d$, one uses the abbreviation

$$x^\alpha := \prod_{i=1}^d x(i)^{\alpha(i)}, \quad x \in \mathbb{R}^d, \alpha \in \mathbb{Z}_+^d.$$

The notation

$$(\cdot)^\alpha : \mathbb{R}^d \to \mathbb{R} : x \mapsto x^\alpha$$

for the **monomial of degree** α is convenient (though nonstandard). With $\alpha \in \mathbb{Z}_+^d$,

$$\Pi_\alpha := \Pi_{\le\alpha} := \operatorname{span}\{(\cdot)^\beta : \beta \le \alpha\}$$

is the space of all polynomials of **degree** $\le \alpha$. For any integer k,

$$\Pi_k := \Pi_{\le k} := \operatorname{span}\{(\cdot)^\beta : |\beta| \le k\}$$

is the space of all polynomials of **total** degree $\le k$. The spaces $\Pi_{<\alpha}$ and $\Pi_{<k}$ are defined analogously.

Many expressions simplify if one uses the **normalized power function**

$$[\![\cdot]\!]^\alpha : x \mapsto x^\alpha/\alpha!,$$

with

$$\alpha! := \prod_i \alpha(i)!,$$

with the understanding that $[\![\cdot]\!]^\alpha = 0$ if $\alpha \in \mathbb{Z}^d \backslash \mathbb{Z}_+^d$. For example, with $\alpha, \xi, \upsilon, \zeta \in \mathbb{Z}_+^d$, the **Multinomial Theorem** takes the simple form

$$[\![x + y + \cdots + z]\!]^\alpha = \sum_{\xi+\upsilon+\cdots+\zeta=\alpha} [\![x]\!]^\xi [\![y]\!]^\upsilon \cdots [\![z]\!]^\zeta. \tag{2.1}$$

The multinomial theorem is immediate (by induction on the number of summands in the sum on the left-hand side) once one knows it for two summands. For two summands, though, it is just the special case $p = [\![\cdot]\!]^\alpha$ of the *Taylor expansion*

$$p(x+y) = \sum_\xi [\![x]\!]^\xi \, D^\xi p(y),$$

in which

$$D^\xi := D_1^{\xi(1)} \cdots D_d^{\xi(d)},$$

(with D_i differentiation with respect to the ith argument), hence

$$D^\xi [\![\cdot]\!]^\alpha(y) = [\![y]\!]^{\alpha-\xi}.$$

A more sophisticated example is provided by the **Leibniz-Hörmander**

formula

$$p(sD)(fg) = \sum_\beta \left(\left([\![D]\!]^\beta p\right)(sD)f\right)[\![sD]\!]^\beta g$$

concerning the differentiation of the product fg of two functions, in which s is an arbitrary scalar, and p an arbitrary polynomial, $p = \sum_\alpha [\![\cdot]\!]^\alpha c(\alpha)$ say, therefore

$$p(D) := \sum D^\alpha/\alpha!\, c(\alpha)$$

the corresponding constant-coefficient differential operator.

Since $D^\alpha [\![\cdot]\!]^\beta(0) = \delta_{\alpha\beta}$, Π_k has dimension

$$\dim \Pi_k = \#\{\alpha \in \mathbb{Z}_+^d : |\alpha| \le k\} = \binom{k+d}{d};$$

the last equality can be verified, e.g., with the aid of the invertible map

$$\{\alpha \in \mathbb{Z}_+^d : |\alpha| \le k\} \to \binom{\{1,\ldots,d+k\}}{d} : \alpha \mapsto \Big\{\sum_{i\le j}(\alpha(i)+1) : j = 1,\ldots,d\Big\},$$

with $\binom{X}{d}$ the collection of all d-sets, i.e., all subsets of cardinality d, in X.

While there are various *uni*variate polynomial forms available, there is, aside from the (possibly shifted and/or normalized) power form, only one *multi*variate polynomial form in general use, namely the BB-form, to be discussed next. In particular, the equivalent of a Chebyshev form (or similar form of good condition with respect to the max-norm on some domain) is, as yet, not readily available. The BB-form illustrates that it is often good to give up on the power form altogether in favor of forms which employ more general homogeneous polynomials of the form $x \mapsto \prod_{y\in Y} y^T x$, with

$$y^T x := \sum_i y(i)x(i)$$

the standard inner product.

3. BB-form

The BB-form is, at present, the most effective polynomial form for work with pp functions on a simplicial partition (or, more generally, a simploidal partition). For, the BB-form of a polynomial, with respect to a given simplex

$$\langle V\rangle := \operatorname{conv}(V)$$

spanned by some $(d+1)$-set $V \subset \mathbb{R}^d$, is symmetric with respect to the vertices of that simplex, and readily provides information about the behavior of the polynomial on all the faces $\langle W\rangle$, $W \subset V$, of that simplex. This facilitates the smooth matching of two polynomial pieces across the intersection of their respective simplicial cells. For more details than are (or can be) offered here, see Farin (1986) (which concentrates on the bivariate case) as well as de Boor (1987). The presentation here is based on the latter, albeit with certain changes in notation. For the use of the BB-form in the treatment of finite elements, see, e.g., Luscher (1987).

The BB-form can be viewed as a generalization of the standard representation

$$p = \sum_{v\in V} \xi_v p(v)$$

of the linear interpolant to data given at a $(d+1)$-point set $V \subset \mathbb{R}^d$ in general position, with $\xi_v = \xi_{v,V}$ the unique linear polynomial which takes the value 1 at v and vanishes on

$$V\backslash v \ := \ \{w \in V : w \neq v\}.$$

In this connection, 'general position' is tautological since it means nothing more than that such a representation exists for every $p \in \Pi_1(\mathbb{R}^d)$, hence is necessarily unique since $\dim \Pi_1(\mathbb{R}^d) = d+1 = \#V$.

The $(d+1)$-vector

$$\xi_V(x) := (\xi_v(x))_{v\in V}$$

provides the **barycentric coordinates** of x with respect to the point set V. Equivalently, $\xi_V(x)$ is the unique solution of the linear system

$$\sum_{v\in V} \xi_v(x)\,(v,1) \ = \ (x,1) \quad \in \mathbb{R}^{d+1}, \tag{3.1}$$

and this provides the opportunity to write out a formula for its components $\xi_v(x)$ as a ratio of determinants and so explains the alternative name **areal coordinates**.

The BB-form for $p \in \Pi_k$ employs all possible products of k of the linear polynomials ξ_v, $v \in V$, with repetitions permitted, i.e., all the functions

$$\xi_V^\alpha : x \mapsto \xi_V(x)^\alpha$$

with α any multi-index (indexed by V, i.e., in $\mathbb{Z}_+^V$) of length k. However, it turns out to be very convenient to use the particular normalization

$$B_\alpha := B_{\alpha,V} := \binom{|\alpha|}{\alpha} \xi_V^\alpha = |\alpha|!\,[\![\xi_V]\!]^\alpha,$$

which arises when we apply the multinomial theorem (2.1) to obtain

$$1 = k![\![\sum_{v\in V} \xi_v(x)]\!]^k = \sum_{|\alpha|=k} B_\alpha(x). \tag{3.3}$$

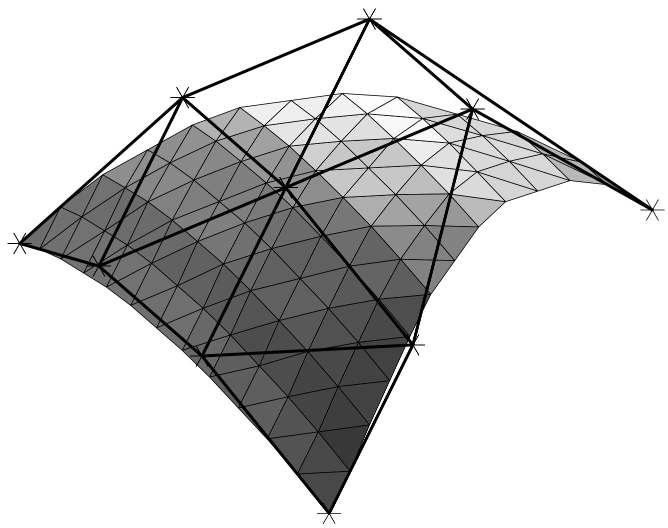

Fig. 3.2. A cubic patch and its control net.

The fact that

$$\#\{\alpha \in \mathbb{Z}_+^V : |\alpha| = k\} = \#\{\beta \in \mathbb{Z}^d : |\beta| \le k\} = \dim \Pi_k$$

implies that *the collection* $(B_\alpha)_{|\alpha|=k}$ *is a basis for* Π_k since (i) any $p \in \Pi_k$ can be written as a linear combination of products of k linear polynomials (e.g., the linear polynomials $x \mapsto x(i)$, $i = 1, \ldots, d$ and $x \mapsto 1$); and (ii) any linear polynomial can be written as a linear combination of the ξ_v, $v \in V$, hence $\Pi_k \subseteq \operatorname{span}\{B_\alpha : |\alpha| = k\}$. The resulting representation

$$p = \sum_{|\alpha|=k} B_\alpha \, b_{p,V}(\alpha)$$

for $p \in \Pi_k$ constitutes the BB-form (a form associated with the names of Bernstein (Lorentz (1953; p. 51)), de Casteljau (1963, 1985), Bézier (1970, 1977), Farin (1977, 1986, 1988), and perhaps others).

Since ξ_v vanishes at all the points in $V \backslash v$ and is linear, it vanishes on the simplex $\langle V \backslash v \rangle$ spanned by these points. It follows that, for any subset U of V, the restriction of B_α to $\langle U \rangle$ is not the zero function if and only if $\operatorname{supp} \alpha \subseteq U$. In particular, the only B_α not zero on $\{v\} = \langle \{v\} \rangle$ is the one with $\alpha = k\mathbf{i}_v$, where

$$\mathbf{i}_v(u) := \delta_{vu}, \quad v \in V.$$

With (3.3), this implies that $B_{k\mathbf{i}_v}(v) = 1$, hence further that

$$p(v) = b_{p,V}(k\mathbf{i}_v), \quad v \in V.$$

This fact and others have made it customary to associate, more generally, the coefficient $b_{p,V}(\alpha)$ with the corresponding **domain point**

$$V_\alpha := \sum_{v\in V} v\,\alpha(v)/|\alpha|,$$

thereby obtaining the **(Bézier) control net**

$$C_p \ := \ C_{p,V,k} := (V_\alpha, b_{p,V}(\alpha))_{|\alpha|=k}$$

for p. Note that $\operatorname{supp}\alpha \subseteq U$ for some $U \subseteq V$ if and only if $V_\alpha \in \langle U\rangle$. Hence, on $\langle U\rangle$, p is entirely determined by $b_{p,V}(\alpha)$ with $V_\alpha \in \langle U\rangle$. To put it differently, the restriction of p to $\langle U\rangle$ has the control net

$$C_{p,U,k} = (V_\alpha, b_{p,V}(\alpha))_{|\alpha|=k, V_\alpha\in\langle U\rangle}.$$

In particular, if

$$f = \begin{cases} p & \text{on } \langle V\rangle, \\ q & \text{on } \langle W\rangle, \end{cases} \tag{3.4}$$

for some $p, q \in \Pi_k$, then f is continuous on $\langle V\rangle \cup \langle W\rangle$ if and only if

$$\forall\{\alpha \in \mathbb{Z}_+^{V\cup W} : |\alpha| = k, \operatorname{supp}\alpha \subset V\cap W\} \quad b_{p,V}(\alpha_{|V}) = b_{q,W}(\alpha_{|W}).$$

Thus, if f is a *continuous* pp function of degree $\le k$ on some **complex** (:= partition of some domain $G \subset \mathbb{R}^d$ into simplices) Δ, in formulæ:

$$f \in \Pi^0_{k,\Delta},$$

then it is uniquely describable in terms of its **BB-net**, b_f. This is, by definition, the mesh-function, defined on the union of all the domain points V_α, $|\alpha| = k$, $\langle V\rangle \in \Delta$, which, for each $\langle V\rangle \in \Delta$, agrees with $b_{p,V}$ on the points in $\langle V\rangle$.

It is well worth stressing that, as d increases, the ratio of domain points in the boundary of a $\langle V\rangle$ over the total number of domain points in $\langle V\rangle$ increases for fixed k, reaching the limiting value 1 as soon as $d > k$. In effect, with increasing d, the polynomial pieces in a pp function of fixed degree $\le k$ become increasingly 'superficial', with more and more of their degrees of freedom needed just to maintain continuity.

3.1. The BB-form as a k-fold difference

For a discussion of a smoother join as well as for its own sake, we need to know how to differentiate the BB-form. For this, and for various other properties, we observe the following striking

Fact 3.5 *For $\omega \in \mathbb{R}^V$, let ωE denote the 'difference operator' which acts on the mesh-function $c : \mathbb{Z}^V \to \mathbb{R}$ by the rule*

$$(\omega E)c := \sum_{v\in V} \omega(v) c(\cdot + \mathbf{i}_v).$$

Then

$$\sum_{|\alpha|=k} B_\alpha(x)c(\alpha) = (\xi_V(x)E)^k c(0).$$

Indeed,

$$(\xi_V(x)E)^k c(0) = \sum_{u\in V}\sum_{v\in V}\cdots\sum_{w\in V} \xi_u(x)\xi_v(x)\cdots\xi_w(x)\;\; c(\mathbf{i}_u+\mathbf{i}_v+\cdots+\mathbf{i}_w)$$

with exactly k summations, hence all summands are of the form $\xi_V(x)^\alpha c(\alpha)$ for some $\alpha\in\mathbb{Z}_+^V$ with $|\alpha|=k$, and this particular summand occurs exactly $\binom{k}{\alpha}$ times. See Figure 3.13 for an illustration.

With this,

$$\sum_{|\alpha|=k} B_\alpha\, b_{p,V}(\alpha) \;=p\; = (\xi_V E)^k b_{p,V}(0), \qquad p\in\Pi_k.$$

With this formula in hand, differentiation of the BB-form requires nothing more than the chain rule, as follows. If $y\in\mathbb{R}^d\backslash 0$, then

$$D_y p = D_y(\xi_V E)^k c(0) = k(\xi_V E)^{k-1}(D_y\xi_V E)c(0).$$

We obtain the vector $D_y\xi_V$ by the observation that, by (3.1), $\xi_V(x+ty)-\xi_V(x)=t\eta_V(y)$, with $\eta_V(y)\in\mathbb{R}^V$ the unique solution of

$$\sum_{v\in V}\eta_v(y)\,(v,1)=(y,0).$$

Hence, altogether,

$$\begin{aligned} D_y p \;&=\; k\sum_{|\alpha|=k-1} B_\alpha(\eta_V(y)E)b_{p,V}(\alpha) \\ &\text{for } p\in\Pi_k \text{ and } \sum_{v\in V}\eta_v(y)\,(v,1) := (y,0)\in\mathbb{R}^{d+1}\backslash 0. \end{aligned} \tag{3.6}$$

For example, for two distinct points $v,u\in V$,

$$\eta_V(v-u)=\mathbf{i}_v-\mathbf{i}_u,$$

hence

$$b_{D_{v-u}p,V}(\alpha)=\frac{b_{p,V}(\alpha+\mathbf{i}_v)-b_{p,V}(\alpha+\mathbf{i}_u)}{1/k}.$$

Repeated application of (3.6) provides the BB-form for any derivative of p of the form $D_Y p$ with Y any finite subset of $\mathbb{R}^d\backslash 0$ and

$$D_Y := \prod_{y\in Y} D_y.$$

3.2. Smooth matching of polynomial pieces

Since we now know how to obtain the BB-form of any derivative of a polynomial p from the BB-form of p, we can describe the matching of derivatives across the common interface $\langle V\cap W\rangle$ of two simplicial cells $\langle V\rangle$ and $\langle W\rangle$. Simply put, the derivative in question of the polynomial p on $\langle V\rangle$ and the polynomial q on $\langle W\rangle$ must agree on $\langle V\cap W\rangle$, i.e., their corresponding control points with domain point in $\langle V\cap W\rangle$ must agree.

It is not hard to write specific smoothness conditions in the form of an equality between the expressions, obtained by application of (3.6), for the relevant control points (see, e.g., Chui and Lai (1987) and Chui (1988, Theorem 5.1) or Farin (1986)) of the relevant derivatives. However, if the goal is a $C^{(\rho)}$-match, i.e., a matching of all derivatives of order $\le\rho$, then the uniformity of the BB-form permits a more unexpected formulation of the corresponding smoothness conditions, as follows.

For $p\in\Pi_k$ and $\beta\in\mathbb{Z}_+^V$ with $|\beta|\le k$, let

$$p_\beta := \sum_{|\gamma|=k-|\beta|} b_{p,V}(\beta+\gamma)B_\gamma.$$

These are the **subpolynomials** introduced in de Boor (1987); see also Farin (1986; (2.5)). For example, if $|\beta|=k$, then p_β is the constant polynomial with value $b_{p,V}(\beta)$. Consequently, (3.4) is continuous if and only if $p_\beta=q_\beta$ for all β with $\operatorname{supp}\beta\subset V\cap W$ and $|\beta|=k$. As another example, if $|\beta|=k-1$, then p_β is the linear polynomial whose value at $v\in V$ is $b_{p,V}(\beta+\mathbf{i}_v)$, and, for any y, its derivative D_yp_β is the constant $b_{D_yp,V}(\beta)$. Consequently, (3.4) is in $C^{(1)}$ if and only if $p_\beta=q_\beta$ for all β with $\operatorname{supp}\beta\subset V\cap W$ and $|\beta|=k-1$.

Here is the general theorem.

Theorem 3.7 *The pp function f, defined in (3.4), is in $C^{(r)}$ for some $r\le k$ if and only if*

$$\forall\{\beta\in\mathbb{Z}_+^V:\operatorname{supp}\beta\subset V\cap W, |\beta|=k-r\}\quad p_\beta=q_\beta. \tag{3.8}$$

In particular, since $q_\beta(w)=b_{q,W}(\beta+r\mathbf{i}_w)$ for each such β and each $w\in W$, $C^{(r)}$-continuity requires that

$$\forall\{\beta\in\mathbb{Z}_+^V:\operatorname{supp}\beta\subset V\cap W, |\beta|=k-r, w\in W\backslash V\}\quad b_{q,W}(\beta+r\mathbf{i}_w)=p_\beta(w). \tag{3.9}$$

Conversely, if our f is already in $C^{(r-1)}$, hence $p_\beta=q_\beta$ for all β with $\operatorname{supp}\beta\subset V\cap W$ and $|\beta|=k-r+1$, then the conditions (3.8) are equivalent to the conditions (3.9). In particular, (3.9) supplies a complete and independent set of conditions for $C^{(r)}$-continuity across $\langle V\cap W\rangle$ in the presence of $C^{(r-1)}$-continuity. Consequently, the union over $r=0,\ldots,\rho$ of these conditions constitutes a complete and independent set of conditions for $C^{(\rho)}$-continuity across $\langle V\cap W\rangle$.

Note the remarkable *uniformity* of the conditions (3.9): The weights in the right-hand side $p_\beta(w)$, considered as a linear combination of the BBcoefficients $b_{p,V}(\alpha)$ for p, depend only on w and r (and V) and not on β or k.

Note also that the smoothness conditions of order r, i.e., the conditions (3.9), involve only control points of f in the first r 'layers' along $\langle V \cap W \rangle$. Note finally, that we might have, equally well, used the complementary conditions

$$\forall \{\beta \in \mathbb{Z}_+^V : \operatorname{supp} \beta \subset V \cap W, |\beta| = k - r, v \in V \backslash W\} \quad b_{p,V}(\beta + r\mathbf{i}_v) = q_\beta(v). \tag{3.10}$$

In effect, the subpolynomials $p_\beta = q_\beta$ with $\operatorname{supp} \beta \subset V \cap W$ and $|\beta| = k - r$ give a complete description of the behavior of all derivatives of f of order $\le r$ on $\langle V \cap W \rangle$, and enforcement of (3.9) and (3.10) makes certain that the corresponding derivatives of p and q agree with these of f on $\langle V \cap W \rangle$.

It is this remarkably explicit geometric connection between the control points and the behavior 'near' any particular face of $\langle V \rangle$ that makes the BB-form so attractive for work with pp functions.

The simplest nontrivial case, $r = 1$, is of particular practical interest. It requires that, for each $\beta \in \mathbb{Z}_+^V$ with $\operatorname{supp} \beta \subset V \cap W$ and $|\beta| = k - 1$, $q_\beta(w) = p_\beta(w)$, i.e., that the control point $(W_{\beta + \mathbf{i}_w}, b_{q,W}(\beta + \mathbf{i}_w))$ lie on the (hyper)plane spanned by the control points $(V_{\beta + \mathbf{i}_v}, b_{p,V}(\beta + \mathbf{i}_v))$, $v \in V$, a particularly nice geometric interpretation rightfully stressed in the CAGD literature (see, e.g., Boehm, Farin and Kahmann (1984), Farin (1988) and Hoschek and Lasser (1989)).

3.3. Simple examples

As an illustration of the strength and efficiency of the BB-form, here is a discussion of three standard topics concerning bivariate pp functions.

Quintic Hermite interpolant In bivariate quintic Hermite interpolation, one matches value and first and second derivatives at three points, thus using up eighteen of the available $21 = \binom{7}{2} = \binom{k+d}{d}$ degrees of freedom, and then uses the remaining three degrees of freedom for a possible $C^{(1)}$-join with neighboring quintic patches. Here are the details, well known, but particularly evident when discussed in terms of the BB-form.

Let $d = 2$, $V = \{u, v, w\}$, $k = 5$, and $p \in \Pi_k$.

Then $p(u) = b_p(5\mathbf{i}_u)$.

Further, for $\nu \in V \backslash u$,

$$D_{\nu - u} p(u) = b_{D_{\nu - u} p}(4\mathbf{i}_u) = 5(b_p(4\mathbf{i}_u + \mathbf{i}_\nu) - b_p(4\mathbf{i}_u + \mathbf{i}_u)),$$

showing that the coefficients $b_p(4\mathbf{i}_u + \mathbf{i}_\nu)$, $\nu \in V \backslash u$, are determined by $D_{\nu - u} p(u)$, $\nu \in V \backslash u$ and *vice versa* (once $p(u) = b_p(5\mathbf{i}_u)$ is known). More than that, it shows that *the tangent plane to p at u is the plane spanned*

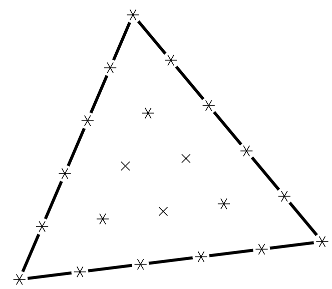

Fig. 3.11. The quintic Hermite interpolant.

by the control points at and next to u. (This discussion actually applies for arbitrary d and k.)

Finally, with $\mu, \nu \in V\backslash u$, all second derivatives are linear combinations of the second derivatives of the form $D_{\mu,\nu}$, of which there are exactly as many as there are distinct points of the form $3\mathbf{i}_u + \mathbf{i}_\mu + \mathbf{i}_\nu$, i.e., control points in the second layer of control points near u, and, correspondingly, with the tangent plane at u already determined, the specification of all second derivatives of p at u is equivalent to the specification of all the control points in that second layer. (Again, this discussion applies for arbitrary d and arbitrary k.)

In other words, the behavior of all derivatives of p at u of order ≤ 2 is determined by the subpolynomial

$$p_{3\mathbf{i}_u} = \sum_{|\gamma|=2} b_p(3\mathbf{i}_u + \gamma) B_\gamma,$$

and it involves the control points in the zeroth, first and second layer for u. Since $d = 2$ and $k = 5$, this 'triangle' of control points associated with u has no intersection with the corresponding coefficient 'triangles' associated with the other vertices. This implies that one can freely specify value, first and second derivatives of $p \in \Pi_5$ at each of these three vertices, and this specifies the 18 control points in those 'triangles', and leaves free exactly one control point per edge. This control point is in the first layer for that edge, hence determines the middle control point for that edge for any particular first derivative of p. Equivalently, for the control point associated in this way with the edge $\langle u, v\rangle$, it is the only piece of information for the (linear) subpolynomial $p_{2\mathbf{i}_u+2\mathbf{i}_v}$ not yet specified (and this is the only linear subpolynomial p_β with $\operatorname{supp}\beta \subset \langle\{u, v\}\rangle$ not yet completely specified). Consequently, if the control point is determined in such a way that it equals the corresponding control point of the same derivative of a quintic Hermite interpolant (to the same vertex data) in the triangle sharing this edge, then the two quintic polynomials form a $C^{(1)}$ pp function. This can be achieved, e.g., by specifying the normal derivative at the midpoint of that edge (or any other particular, transversal, derivative).

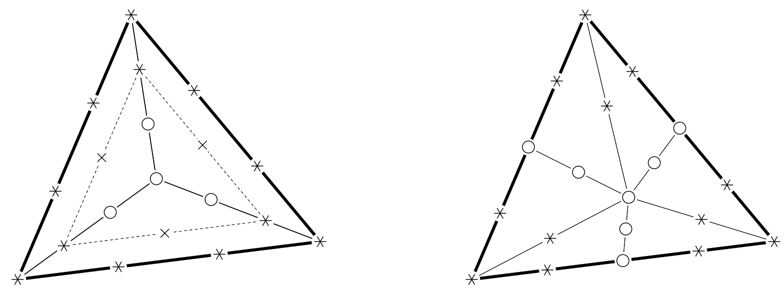

Fig. 3.12. The Clough-Tocher split and the Powell-Sabin split.

To re-iterate, the point of this example (and the two to follow) is not to derive a new result, but to show how easily these known results are derivable in the language of the BB-form.

Clough-Tocher Here, one subdivides a given triangle arbitrarily into three, by connecting its vertices to an arbitrarily chosen point in the interior. Prescribing the tangent plane at each vertex determines the vertex control points and the next-to-vertex control points (see the points marked $*$ in Figure 3.12).

That leaves the points marked x still undetermined, hence allows matching of some transversal derivative at some point. Traditionally, this has been the normal derivative at the midpoint, with the value either given, or else estimated from the vertex information. In this way, value and first derivatives along an edge are entirely determined by information specified on that edge. Hence $C^{(1)}$ matching across that edge is ensured provided the abutting triangle is handled in the same way.

That leaves the control points marked o . These must be determined so that the $C^{(1)}$ conditions hold across the interior edges. At this point, the uniformity of the BB-form comes into play, as follows. One determines the unknown control points to be the control points (with respect to the triangle(s) to which they are assigned) of the unique quadratic polynomial for which the six points on the dot-dashed triangle are the control points (with respect to the triangle to which they are assigned, i.e., the dot-dashed triangle). This can be done by one application of the de Casteljau algorithm to evaluate the given BB-form of this quadratic polynomial at the 'dividing' point chosen in the interior; see the next subsection for details. The resulting control points will satisfy the $C^{(1)}$-conditions since they represent a piecewise quadratic which is even in $C^{(2)}$. In particular, the resulting piecewise cubic is $C^{(2)}$ at the interior vertex (in addition to being $C^{(1)}$ everywhere).

Powell–Sabin There is a corresponding construction of a piecewise quadratic $C^{(1)}$ element, the Powell-Sabin macro-element. Here, one subdivides the triangle into six pieces, starting with some interior point as an additional vertex, but connecting it not only to the vertices, but also to a point on each edge. But, as we shall see, this has to be done just right, to ensure a $C^{(1)}$ match between such macro-elements.

As before, prescription of the tangent plane at each (exterior) vertex pins down vertex and next-to-vertex control points (marked $*$ in Figure 3.12), leaving a 'Y' of control points (marked o). The $C^{(1)}$-conditions across the interior edges determine all but the interior vertex one, and that will necessarily have to lie in the plane spanned by the three control points next to it.

With this, the element is $C^{(1)}$, and any first derivative is piecewise linear along an (exterior) edge, with its extreme values determined explicitly by the given tangent planes at the two vertices of interest. The middle corner of this piecewise linear function is also determined by this information, but in ways that depend strongly on the choice of that the interior vertex and the additional vertex on the edge, as well as on the particular derivative direction. Since only one particular transversal derivative needs to be matched in order to achieve $C^{(1)}$ across the edge, choose a particular direction and then make certain that the interior and the additional edge vertices are so chosen that this particular transversal derivative is actually linear (i.e., has no active interior vertex). Powell and Sabin do this by choosing the midpoint of the edge as the edge vertex and, correspondingly, the interior vertex as the intersection of midpoint normals, i.e., as the center of the circumscribed circle. This makes the derivative in the direction normal to the edge linear.

More generally, pick, in each macro-triangle to be, the interior vertex in such a way (e.g., as the center of the inscribed circle) that the line from it to the corresponding point in any neighboring triangle cuts the common edge at some point strictly between the two common vertices, and use this intersection point as the additional vertex on that edge. Then the three new control points along that midline, as the average of two triples of points with each triple on a straight line, lie themselves on a straight line, thus ensuring $C^{(1)}$.

3.4. Evaluation of the BB-form

As a final advertisement for the BB-form, I discuss the de Casteljau algorithm (de Casteljau (1963)) for its evaluation. This algorithm obtains the value $p(x)$ by carrying out the k-fold application of the difference operator $\xi_V(x)E$ to the mesh-function b_p, as described in Fact 3.5. Since only the value of $(\xi_V(x)E)^k b_p$ at 0 is wanted, we only require $(\xi_V(x)E)^{k-1} b_p$ at α with $|\alpha| = 1$, $(\xi_V(x)E)^{k-2} b_p$ at α with $|\alpha| = 2$, ..., b_p at α with $|\alpha| = k$. It is instructive to visualize the entire discrete $(d+1)$-simplex of mesh points α involved here, as is done in Figure 3.13. For $j = k-1, k-2, \ldots, 0$, the

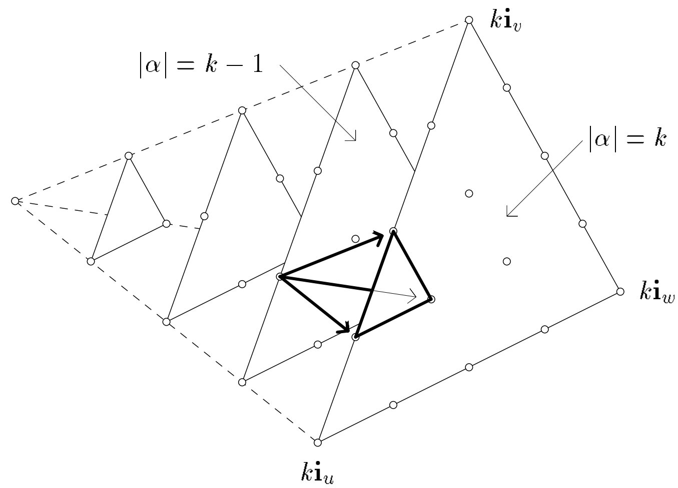

Fig. 3.13. The mesh-point simplex for evaluation.

algorithm derives the 'layer' of values associated with $|\alpha| = j$ from the layer associated with $|\alpha| = j+1$, with each value computed as exactly the same averave of the corresponding d-simplex of values in the next layer.

As a remarkable bonus, the calculations provide (Goldman (1983)), simultaneously, the BB-form for p with respect to $W := (V\backslash v) \cup x$ for any particular $v \in V$: If we denote by c the mesh-function whose values at α, $|\alpha| \le k$, are being generated during the algorithm from the numbers $c(\alpha) := b_{p,V}(\alpha)$, $|\alpha| = k$, then

$$b_{p,W}(\alpha + (k - |\alpha|)\mathbf{i}_v) = c(\alpha), \quad \alpha(v) = 0.$$

This is another effect of the uniformity of the BB-form. As we evaluate the BB-form of some polynomial at some point, we are simultaneously evaluating all associated subpolynomials at the same point. On the other hand, the coefficient $b_{p,V}(\alpha)$ is the value at v of the subpolynomial $p_{\alpha-\alpha(v)\mathbf{i}_v}$. See the discussion of the Clough-Tocher element in the preceding section for a ready application of this.

The evaluation at x of a particular derivative, of the form D_Y with the entries of the sequence Y taken from V, proceeds similarly, except that, during the first $\#Y$ steps, one applies the difference operators $\eta_V(y)E$ corresponding to the entries y of Y, and uses the 'evaluation' difference operator $\xi_V(x)E$ only for the remaining $k - \#Y$ steps. Of course, since any two such difference operators commute, one is entitled to apply the relevant difference operators in any order. In particular, it might be most efficient and stable to apply the $k - \#Y$ 'evaluation' operators first, leaving the application of the 'differentiation' operators for the remaining $\#Y$ layers, which are smaller.

Finally, the de Casteljau algorithm in no way relies on the fact (except, perhaps in the argument for its stability) that the weights ω in the difference operator ωE sum to one. If we employ it with some arbitrary weight vector ω instead of with $\xi_V(x)$, we obtain the number

$$H_p(\omega) := \sum_{|\alpha|=k} b_p(\alpha) k! [\![\omega]\!]^\alpha,$$

i.e., the value at ω of the unique *homogeneous* polynomial H_p on $\mathbb{R}^{d+1}$ for which $H_p(\xi_V(x)) = p(x)$ for all $x \in \mathbb{R}^d$. In conjunction with (3.6), this leads to the formulæ

$$\frac{k!}{(k-\rho)!} \sum_{|\alpha|=k-\rho} H_{p_\alpha}(\eta_V(y)) B_\alpha = D_y^\rho p = \frac{k!}{(k-\rho)!} \sum_{|\alpha|=k-\rho} p_\alpha \, H_{B_\alpha}(\eta_V(y))$$

of Farin (1986; Thm. 2.4, Cor. 2.5), sometimes stated with somewhat less care.

4. The space $\Pi_{k,\Delta}^\rho$

While automotive and aerospace engineers have been working with tensor product spline functions since the early 1960's and structural engineers have been working with pp finite elements just as long, mathematicians in Approximation Theory began to study spaces of multivariate pp functions of non-tensor product type seriously only in the 1970's.

The initial focus was the 'spline' space

$$\Pi_{k,\Delta}^\rho$$

(also denoted by $\mathcal{S}_k^\rho(\Delta)$) of all pp functions of degree $\le k$ in $C^{(\rho)}$ with partition Δ. Here, in full generality, Δ is a collection of 'cells', i.e., closed convex sets δ, with pairwise disjoint, nonempty interiors, whose union is some domain $G \subset \mathbb{R}^d$ of interest, and $\Pi_{k,\Delta}^\rho$ consists of exactly all those $f \in C^{(\rho)}(G)$ for which $f_{|\delta} \in \Pi_k(\delta)$ for all $\delta \in \Delta$. Any such space is contained in the space

$$\Pi_{k,\Delta} =: \Pi_{k,\Delta}^{-1}$$

of all pp functions of degree $\le k$ with partition Δ. However, as soon as we impose some smoothness condition, i.e., as soon as $\rho \ge 0$, the 'cells' of Δ are chosen to be **polytopes**, i.e., the convex hull of a finite set (the **vertex set** for the cell), since the task of matching polynomial pieces across the common boundary of two such cells becomes too difficult otherwise. Further, the partition Δ is taken to be **regular** in the sense that the intersection of two cells is the convex hull of the intersection of their vertex sets. In the simplest case, Δ is a **complex**, i.e., a regular partition consisting of simplices. Such a partition is often called a **triangulation** even when $d > 2$.

Initially, there were high hopes that it would be possible to generate a theory of these spaces to parallel the theory of univariate splines (as recorded, e.g., in Schoenberg (1969), de Boor (1976, 1978), Schumaker (1981) and Powell (1981)). For example, here is a list of desirable goals, from Schumaker (1988, 1991):

1. Explicit formulæ for the dimension of spline spaces;
2. Explicit bases consisting of locally-supported elements;
3. Convenient algorithms for storing and evaluating the splines, their derivatives, and integrals;
4. Estimates of the approximation power of spline spaces;
5. Conditions under which interpolation is well-defined;
6. Algorithms for interpolation and approximation.

However, the experience gained so far has led to some doubt as to whether these goals are likely to be achieved fully even in the bivariate case.

It is also not clear whether the restriction to polynomials of total degree $\le k$ is reasonable *a priori*. On a cell which is the cartesian product $\delta_1 \times \delta_2$ of lower-dimensional cells δ_1 and δ_2, it seems, offhand, more reasonable to use elements from the tensor product $\Pi_k(\delta_1) \otimes \Pi_k(\delta_2)$ of polynomials of total degree $\le k$ on those lower-dimensional sets. For example, in a bivariate context, a typical practical partition involves triangles and quadrilaterals, and, in such a setting, the restriction to polynomials of total degree $\le k$ seems reasonable only if one first refines the partition, by subdividing each quadrilateral into triangles. This does have the advantage of uniformity and, if properly done, may produce partitions which support locally supported smooth pp functions of smaller degree than did the original partition. In fact, for a general partition, this is certain to be so if even the triangles are subdivided appropriately. On the other hand, as of this writing and as a consequence of the early dominance of tensor product methods, most commercially used software packages for surface design and manufacturing can only handle partitions with quadrilateral cells and, correspondingly, bicubic, or biquintic, polynomial pieces.

4.1. The dimension of $\Pi^\rho_{k,\Delta}$

When $\rho = -1$, then $\dim \Pi^\rho_{k,\Delta} = \dim \Pi_k(\mathbb{R}^d) \cdot \#\Delta$. However, already for $\rho = 0$, there is no hope for a formula for $\dim \Pi^\rho_{k,\Delta}$, except in the simplest case, when Δ is a triangulation. In this case, the BB-nets for the polynomial pieces of $f \in \Pi^0_{k,\Delta}$ associated with two neighboring cells, $\langle V\rangle$ and $\langle W\rangle$, necessarily agree at all domain points in the intersection $\langle V\rangle \cap \langle W\rangle = \langle V \cap$

$W\rangle$. Consequently, the map

$$f \mapsto b_f$$

from f to its BB-net sets up a 1-1 correspondence between $\Pi^0_{k,\Delta}$ and all scalar-valued functions on the mesh

$$A_{k,\Delta} := \{V_\alpha : |\alpha| = k, \langle V\rangle \in \Delta\}.$$

In particular,

$$\dim \Pi^0_{k,\Delta} \;=\; \#A_{k,\Delta}.$$

For $\rho > 1$, one would think of $\Pi^\rho_{k,\Delta}$ as the linear subspace of $\Pi^0_{k,\Delta}$ singled out by the $C^{(\rho)}$-conditions across facets, hence could, in principle, determine its dimension as the difference between $\dim \Pi^0_{k,\Delta}$ and the *rank* of the collection of $C^{(\rho)}$-conditions. While, as we have seen, it is easy to specify this rank for the collection of all $C^{(\rho)}$-conditions across *one* facet, it is, in general, very difficult to determine the rank of all conditions, as a simple example below will illustrate. Already for $\rho = 1$, there are real difficulties in ascertaining $\dim \Pi^\rho_{k,\Delta}$. Strang's articles (1973, 1974) called attention to this by providing a conjecture concerning $\dim \Pi^\rho_{k,\Delta}$ in the *bi*variate case, namely that the lower bound in the following theorem, due to Schumaker, is the exact dimension for 'generic' triangulations.

Theorem 4.1 *Let Δ be a finite triangulation in $\mathbb{R}^2$, let V_I, E_I denote the collection of its* interior *vertices and edges, respectively. Further, for each $v \in V_I$, let E_v denote the collection of all edges having v as an endpoint, and denote by $\tilde{E}_v \subset E_v$ those with different slopes.*

Then

$$\dim \Pi^\rho_{k,\Delta} - (\dim \Pi_k + \dim \Pi_{k-\rho-1} \cdot \#E_I - (k^2+3k-\rho^2-3\rho)/2 \cdot \#V_I) \in [\sigma \,.\,.\, \tilde{\sigma}],$$

with

$$\sigma \;:=\; \sum_{v\in V_I} \sum_{j=1}^{k-\rho} (\rho + j + 1 - j \cdot \#E_v)_+$$

and $\tilde{\sigma}$ defined in the same way, but with E_v replaced by $\tilde{E}_v$.

(Here and elsewhere, $[a \,.\,.\, b]$ specifies the (closed) interval with endpoints a and b, since the more customary notation $[a, b]$ is also used for the divided difference at two points as well as for the matrix with columns a and b.) See Schumaker (1979 (1984)) for a proof of the lower (upper) bound.

Perhaps the simplest example indicating that it is not possible to be more precise than this is provided by consideration of $\dim \Pi^1_{2,\Delta}$, with the partition Δ obtained by connecting the four points of a (convex) quadrilateral with some point in its interior. Assume first that the interior point was chosen 'generically', in which case the four interior edges for Δ have four distinct

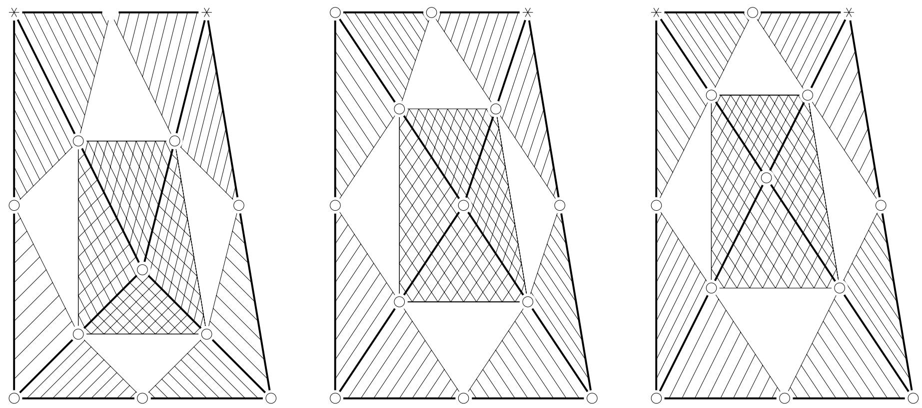

Fig. 4.2. Generic and related nongeneric partitions.

slopes, as in the left half of Figure 4.2. In search for some $f \in \Pi^1_{2,\Delta} \backslash \Pi_2$, we consider the BB-net for f. We assume without loss that f vanishes on the bottom triangle, and have indicated this in Figure 4.2 by drawing a '$\circ$' at the six domain points in that triangle for the BB-net for f. Now, as discussed in the last paragraph of subsection 3.2 above, $C^{(1)}$-continuity requires the coplanarity of the four control points associated with each of the shaded quadrilaterals. In particular, this forces all the control points in the first layer outside the edges of the bottom triangle to be zero, and this is also indicated in the figure. Offhand, the control points associated with the two top corners are freely choosable *except* that the control point associated with the midpoint of the top edge (the one left blank) must lie on the plane spanned by the three control points to the left as well as on the plane spanned by the three control points to the right. In the generic case, this imposes one constraint on the two vertex control points, and we conclude that $\dim \Pi^1_{2,\Delta} = 7$ in this case.

The same conclusion can be reached when the interior vertex lies on one but not the other of the two diagonals of the quadrilateral, as shown in the middle of Figure 4.2. In terms of that figure, the domain point in the middle of the upper edge lies on the straight line through the domain points of the two zero control points to the right of it, hence the corresponding control point must be zero. Since its domain point does *not* lie on the straight line through the domain points of the two zero control points to the left of it, this implies that also the remaining control point associated with the upper left shaded quadrilateral, the vertex control point, must be zero. The other upper vertex control point, however, is freely choosable.

Finally, if that interior vertex happens to be the intersection of the two diagonals of the quadrilateral (as shown in the right of Figure 4.2), then the argument just given shows that the control point associated with the middle of the upper edge must be zero, and both upper vertex control points are freely choosable. Hence, $\dim \Pi^1_{2,\Delta} = 8$ in this case.

For comparison, for this particular example, we have just one interior vertex, v, and $\#E_v = 4$, while, in the three distinct cases, $\#\tilde{E} = 4, 3, 2$. Correspondingly, $\sigma = (1 + 1 + 1 - 4)_+ = 0$, while $\tilde{\sigma} = 0, 0, 1$ in the three cases. Thus, for this example and in these three cases, the theorem is sharp in the sense that it amounts to the assertion that

$$(7,7,8) - 7 \in [0\,..\,(0,0,1)].$$

The arguments used in this example illustrate how, in general, one might go about to determine $\dim \Pi^\rho_{k,\Delta}$. As already stressed, one rightly thinks of $\Pi^\rho_{k,\Delta}$ as the subspace of $\Pi^0_{k,\Delta}$ characterized by the $C^{(\rho)}$-conditions. A pp function on the triangulation Δ is in $C^{(\rho)}$ precisely when it is in $C^{(\rho)}$ on any two simplices of Δ which share a whole facet, i.e., whose vertex sets differ only by one point. For this reason, $\Pi^\rho_{k,\Delta}$ is linearly isomorphic to all the mesh-functions b_f on $A_{k,\Delta}$ which, for each such simplex pair, satisfy the corresponding conditions (3.9) across their common facet for $r = 1, \ldots, \rho$. Moreover, for each such facet, this provides a maximally linearly independent set of $C^{(\rho)}$-conditions imposed across *one* such facet. However, conditions across different (but neighboring) facets may well be linearly dependent. For example, Figure 4.2 shows four $C^{(1)}$-conditions involving the control point at the interior vertex. Yet, since they all require that their respective control points lie on a certain plane, it takes just two such conditions to ensure that all five control points involved lie on the same plane, hence the other two conditions must be dependent on them. Unfortunately, it is in general impossible to provide a basis for the collection of *all* smoothness conditions imposed. This has made it a challenge (unsolved so far and not likely to be solved in any generality) to determine the dimension of $\Pi^\rho_{k,\Delta}$ when $\rho > 0$.

As the example shows, there is no hope to express $\dim \Pi^\rho_{k,\Delta}$ entirely in such combinatorial terms as the number of (interior or boundary) vertices, edges, triangles. However, even the hope that, as in this case, the counting of such things as nonparallel edges incident to a vertex might suffice is dashed by a more subtle example due to Morgan and Scott in 1977 (Morgan and Scott (1990)), which uses the partition Δ obtained by placing a scaled and reflected copy of an equilateral triangle concentrically inside that triangle and connecting each vertex of the inner triangle to the two closer vertices of the outer triangle. As Morgan and Scott show (and use of the BB-net would show more readily), for this Δ, $\dim \Pi^1_{2,\Delta} = 7$ while, for any generic perturbation Δ' of Δ, $\dim \Pi^1_{2,\Delta'} = \dim \Pi_2 = 6$.

Since the arguments for Theorem 4.1 make essential use of the fact that one knows how to construct bases for arbitrary univariate spline spaces, while we do not know how to do this in general for bivariate spline spaces, it is unlikely that one can obtain even the trivariate analogon of Theorem 4.1. An observation of Alfeld (in Alfeld, Schumaker and Sirvent (1992), see Schumaker (1991)) makes this precise. The latter reference gives a very good summary of what is presently known about $\dim \Pi^\rho_{k,\Delta}$. In particular, the recent paper Alfeld, Whiteley and Schumaker (199x) gives first specific results concerning the dimension of *tri*variate spline spaces. In addition, Billera and his colleagues initiated and pursued an investigation of $\dim \Pi^\rho_{k,\Delta}$ for arbitrary d with tools from Homological Algebra, which, however, forces them to consider only the case of a 'generic' Δ (which is difficult enough); see Billera (1988, 1989), Billera and Haas (1987) and Billera and Rose (1989, 1991). For example, Billera (1988) shows Strang's conjecture for $\rho = 1$ to be correct 'generically', using a specific construction of Whiteley (1991) to make certain that a certain determinant is not identically zero, hence must be generically nonzero.

Those with an urge to get a feeling for the difficulties one might encounter in considering arbitrary partitions should try the still unsolved problem of providing a formula for $\dim \Pi^1_{3,\Delta}(\mathbb{R}^2)$ for arbitrary Δ.

4.2. Subspaces of $\Pi^\rho_{k,\Delta}$

It is not only the difficulty of determining $\dim \Pi^\rho_{k,\Delta}$, hence of constructing bases for $\Pi^\rho_{k,\Delta}$, that makes the full space more of a challenge than of real interest. For certain partitions, $\Pi^\rho_{k,\Delta}$ contains elements of no use for approximation (such as the **half-space** spline $\mathbb{R}^d \to \mathbb{R} : x \mapsto (\langle y, x\rangle - c)^k_+$, with y a certain element of $\mathbb{R}^d$ and c some constant). Also, if k is large enough compared with ρ, then there are often subspaces of $\Pi^\rho_{k,\Delta}$ with the same 'approximation power' as $\Pi^\rho_{k,\Delta}$ itself.

For example, in the Finite Element method, bivariate pp spaces studied by Ženišek (1970, 1973, 1974) and recently termed super-spline spaces in Chui and Lai (1987) consist of all elements of $\Pi^\rho_{k,\Delta}$ which, at each vertex, are in $C^{(2\rho)}$. In terms of the BB-net, the motivation (as explained, e.g., in Farin (1986)) for consideration of such subspaces is simple: if, for some $\delta \in \Delta$, we want to determine the polynomial piece $p = f_{|\delta}$ on δ so as to have a $C^{(\rho)}$-join with its neighboring pieces, then its first ρ layers of control points along each edge of δ are determined by the polynomial piece adjoining that edge. However, certain of these control points are in the first ρ layers of two edges, hence in danger of being overdetermined. For any two edges, these endangered control points are contained in the first 2ρ layers for the vertex common to those two edges (and in no smaller set of layers). Hence, the enforcement of $C^{(2\rho)}$-continuity at the vertices ensures consistency for the competing smoothness conditions.

There are certain questions to be raised here. First, it has become popular, because of the success of the multigrid method, to work with a sequence of spaces, each obtained from the previous one by *refinement*, typically looking at the space of the same type on a refinement of the triangulation of the preceding one. If the spaces involved are super-spline spaces, then, because of the higher smoothness requirement at the vertices, the finer space will fail to contain the rougher space. Also, the degree k must be large enough so that the only questions of consistency of the smoothness conditions are of the kind described. For $d = 2$, this means that $k \geq 4\rho + 1$. Analogous considerations for arbitrary d (though not using BB-nets) led Le Méhauté (1990) to the conclusion that $k \geq 2^d\rho + 1$ was necessary (and sufficient) to provide such a super-spline space, in which an approximation can be constructed in a totally local way, with the approximant f on the simplex δ depending only on data on δ.

Such degrees are daunting. One response is to give up on using arbitrary triangulations, but use instead triangulations Δ obtained, e.g., by proper refinement of a given triangulation. The standard example is the Clough-Tocher element (although, because of its greater smoothness at its interior vertex, the space spanned by it does not properly refine, either). The extreme case of partitions (in general, they are not even triangulations) which will support compactly supported pp functions of low degree compared with the required smoothness are those provided by the multivariate B-spline construct to be discussed next.

5. Multivariate B-splines

The central role ultimately played by the univariate B-splines of (Curry and) Schoenberg (1946, 1966) in univariate spline theory (as illustrated, e.g., in Schoenberg (1969), de Boor (1976), or Schumaker (1981)) provided the impetus for the study of a certain multivariate generalization. Offhand, this generalization is based on preserving the somewhat obscure property of the univariate B-spline illustrated in Figure 5.1 and originally proved, in Curry and Schoenberg (1966), for the purpose of showing that the univariate B-spline is log-concave. Here are the details.

The **univariate B-spline $M(\cdot|\Theta)$ with knot sequence** $\Theta = (\theta_0, \ldots, \theta_s)$ is, by one of its definitions, the Peano kernel for the divided difference (functional) $[\theta_0, \ldots, \theta_s]$, i.e., it is the unique function for which

$$[\theta_0, \ldots, \theta_s]f \;=\; \int_{\mathbb{R}} M(t|\Theta)\, D^s f(t)\, dt/s!$$

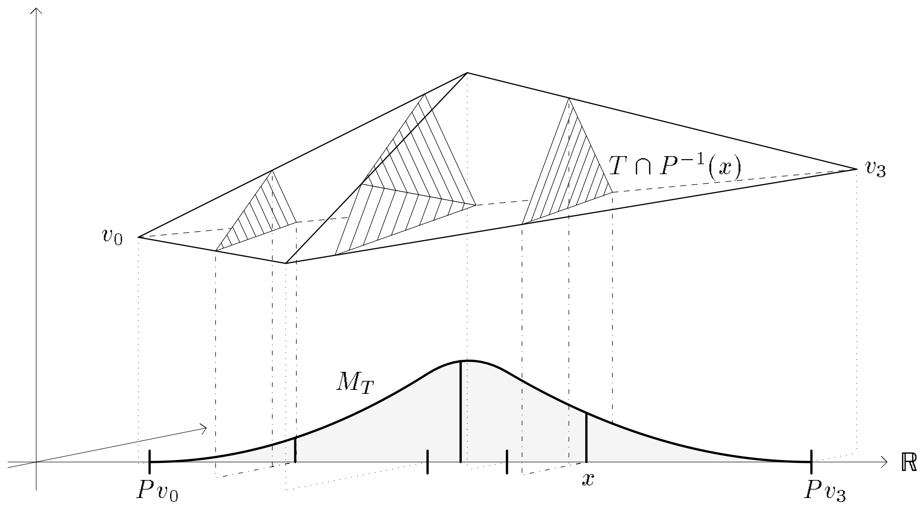

Fig. 5.1. A quadratic B-spline as the shadow of a tetrahedron.

for all sufficiently smooth functions f. On combining this with the Hermite-Genocchi formula (Nörlund (1924)) for the divided difference, Schoenberg obtains the equation

$$\int_{\mathbb{R}} M(t|\Theta) D^s f(t) dt/s! =$$

$$\int_{\Theta} D^s f := \int_0^1 \int_0^{\tau_1} \cdots \int_0^{\tau_{s-1}} D^s f(\theta_0 + \tau_1 \nabla\theta_1 + \cdots + \tau_s \nabla\theta_s) d\tau_s \cdots d\tau_2 \, d\tau_1$$

(with $\nabla\theta_j := \theta_j - \theta_{j-1}$, as usual). This equation implies that $M(t|\Theta)$ is the $(s-1)$-dimensional volume of the set

$$\{\tau \in T_s : \theta_0 + \tau_1\nabla\theta_1 + \cdots + \tau_s\nabla\theta_s = t\},$$

with T_s the **standard s-simplex**

$$T_s := \{\tau \in \mathbb{R}^s : 1 \geq \tau_1 \geq \tau_2 \geq \cdots \geq \tau_s \geq 0\}.$$

This simplex has vertices $v_j := \sum_{i=1}^{j} \mathbf{i}_i$, $j = 0, \ldots, s$. Hence,

$$P : \mathbb{R}^s \to \mathbb{R} : \tau \mapsto \theta_0 + \tau_1\nabla\theta_1 + \cdots + \tau_s\nabla\theta_s$$

is the affine map which carries v_j to θ_j, all j. Consequently, $M(\cdot|\Theta)$ represents the distribution (aka continuous linear functional on $C(\mathbb{R})$)

$$f \mapsto \int_{T_s} f \circ P$$

which carries f to the sum over T_s of its extension $f \circ P$ to a function on $\mathbb{R}^s$. This is illustrated in Figure 5.1 for $s = 3$.

Once this is recognized, there is much scope for generalization (initiated in Schoenberg (1965) and followed up in de Boor (1976), Micchelli (1980), de Boor and DeVore (1983) and de Boor and Höllig (1982)), as follows. For a given (convex) body B in $\mathbb{R}^s$ and a given affine map $P : \mathbb{R}^s \to \mathbb{R}^d$, one defines the corresponding B-spline M_B as the distribution $f \mapsto \int_B f \circ P$. M_B is nonnegative and has support $P(B)$. M_B is a function exactly when $P(B) \subset \mathbb{R}^d$ has interior, but is always a function on affine($P(B)$). When B is a polytope (i.e., the convex hull of some *finite* set), then M_B is called a **polyhedral** spline. A polyhedral spline is pp, with the junction places the images under P of the $(d-1)$-dimensional faces of B. This is most readily seen by using Stokes' theorem, as follows.

After a shift, if need be, we can assume that P is a linear map. Then

$$D_z(f \circ P) = (D_{Pz} f) \circ P.$$

Further, with M_B merely a distribution, $D_y M_B$ is defined by integration by parts,

$$D_y M_B f = -M_B(D_y f).$$

Therefore, for arbitrary $y \in \mathbb{R}^d$ and for any $z \in P^{-1}\{y\}$,

$$\begin{aligned}(D_{Pz} M_B) f &= -\int_B (D_{Pz} f) \circ P = -\int_B D_z(f \circ P) \\ &= -\int_{\partial B} z^T n \,(f \circ P) = -\sum_{F \in B^{(s-1)}} z^T n_F \; M_F f.\end{aligned} \tag{5.2}$$

Here, ∂B is the (oriented) boundary of B. Since B is a polytope, ∂B is the essentially disjoint union of the collection $B^{(s-1)}$ of **facets** (i.e., $(s-1)$-dimensional faces) of B. Further, n is the outward unit normal, and n_F is its constant value on the facet F.

Iteration of this recurrence relation shows that any derivative of M_B of order $> s-d$ is a linear combination of distributions of the form M_F with F itself less than d-dimensional. Hence, on any connected component of the complement of the set

$$\bigcup_{F \in B^{(d-1)}} P(F),$$

(with $B^{(d-1)}$ the collection of all $(d-1)$-dimensional faces of B), M_B is a polynomial of degree $\le k := s-d$. Further, if the polytope B is in general position and P is onto $\mathbb{R}^d$, then any d-face of B is mapped by P to a set with interior, hence all derivatives of M_B of order $\le s-d$ are L_∞ functions. This means that, generically, M_B is pp of degree $\le s-d$ and in $C^{(s-d-1)}$. However, in the interest of obtaining a relatively simple partition (or a partition which is not too different from a given one), one may have to choose B in a special way, and then M_B may not be maximally smooth.

For, as the argument shows, M_B is in $C^{(s-m-1)}$, with m the smallest integer for which P maps every $F \in B^{(m)}$ to a set with interior.

For example, if $B = [0\,..\,1]^s$ is the s-dimensional unit cube, and $\theta_j := P\mathbf{i}_j$, $j = 1,\ldots,s$, and $\theta_0 := P0 = 0$, then the *bivariate* B-spline M_B may have discontinuities in some derivative across any image under P of an edge of B, i.e., across any segment of the form $[\sum_{\theta\in U}\theta\,..\,\sum_{\theta\in W}\theta]$, with U, W arbitrary subsequences of the sequence $(\theta_0,\ldots,\theta_s)$. If each of these segments is also required to be part of the so-called square mesh (or, **two-direction** mesh) (formed by all the lines of the form $\{x \in \mathbb{R}^2 : x(j) = h\}$ with $j \in \{1,2\}$ and $h \in \mathbb{Z}$), then, up to scaling and certain translations, each θ_j is necessarily one of the two unit vectors $\mathbf{i}_1$, $\mathbf{i}_2$. This implies that some face of B of dimension $\lceil s/2 \rceil$ is mapped by P to a set without (2-dimensional) interior, hence M_B is, at best, in $C^{(s/2-2)}$ if s is even. The situation is slightly better for the **three-direction** mesh (formed by all lines of the form $\{x \in \mathbb{R}^2 : x(j) = h\}$ with $j \in \{1,2,3\}$ and $h \in \mathbb{Z}$, and $x(3) := x(1) - x(2)$). Now, θ_j may, in addition to $\mathbf{i}_1$ and $\mathbf{i}_2$, also take on the value $\mathbf{i}_3 := \mathbf{i}_1 + \mathbf{i}_2$. In fact, if $s = 3$ and $\theta_j = \mathbf{i}_j$, $j = 1,2,3$, then the resulting M_B is the hat function, the standard linear finite element at times associated with Courant because of Courant (1943).

Of course, one uses not just one polyhedral spline but linear combinations of sufficiently many to effect good approximation. At a minimum, this means that, after normalization if need be, such a collection $(M_B)_{B\in\mathcal{B}}$ of polyhedral splines should form a **partition of unity**, i.e., satisfy

$$\sum_{B\in\mathcal{B}} M_B = 1.$$

This is quite easy to achieve, as follows. Simply choose the collection $\mathcal{B}$ so that its elements are pairwise essentially disjoint, and their union is a set of the form $\mathbb{R}^d \times C$ for some suitable (convex) $(s-d)$-dimensional set C. For, in that case, $\sum_{B\in\mathcal{B}} M_B(x) = \mathrm{vol}_{s-d}(C)$, while $M_B \geq 0$ in any case. If $B = [0\,..\,1]^s$ (hence M_B is a 'box spline') and P is given by an integer matrix, then the collection $M_B(\cdot - j)$, $j \in \mathbb{Z}^d$, of all integer shifts can be shown to be a partition of unity. Standard arguments concerning approximation order (see the next section) require, more generally, that it be possible to write every $p \in \Pi_{<r}$ as a linear combination of the M_B, $B \in \mathcal{B}$, and this is clearly satisfied for $r = 1$ in case $(M_B)_{B\in\mathcal{B}}$ forms a partition of unity. Much work has gone into constructing $\mathcal{B}$ for which r is large, preferably as large as $s - d + 1$ (it could be no larger), or, alternatively, into determining the largest possible such r for a given $\mathcal{B}$.

It is also important to have the means for reliable evaluation of such a polyhedral spline. It was only after the discovery of stable recurrence relations that univariate B-splines became an effective computational tool. In the same way, work on polyhedral splines only flourished after Micchelli

(1980) established stable recurrence relations for simplex splines. The following generalization, to arbitrary polyhedral splines, was given in de Boor and Höllig (1982); it connects M_B to the M_F with F a facet of B:

$$(s-d)M_B(Pz) \ = \sum_{F\in B^{(s-1)}} (z-a_F)^T n_F \ M_F(Pz), \tag{5.3}$$

with a_F an arbitrary point in affine(F). But there are only very few bodies B for which such a facet F is again a body of the same kind: the simplex, the cube or 'box', and the (polyhedral) cone. The corresponding B-splines are called, correspondingly, **simplex** spline, **box** spline, and **cone** spline (the last introduced in Dahmen (1979)). Each of these can be described entirely in terms of $P(B)$. In other words, any such B-spline is (a shift of) M_B with B a standard simplex, e.g., $\langle 0, \mathbf{i}_1, \ldots, \mathbf{i}_s \rangle$, a standard box $\square := [0\,..\,1]^s$, or a standard cone $\mathbb{R}^s_+$, and P a suitable linear map (which is specified as soon as we know $P\mathbf{i}_j$ for all j).

A first survey of multivariate B-splines is given in Dahmen and Micchelli (1983), an introduction to both simplex splines and box splines is given in Höllig (1986). The only book so far devoted entirely to multivariate B-splines is de Boor, Höllig and Riemenschneider (1992), a book on box splines. Box splines also figure prominently in the survey Chui (1988).

The first multivariate B-spline (and for some still the only one worthy of this appellation) was the simplex spline. If $v_0, \ldots, v_s$ is the sequence of vertices of the underlying simplex, then $M_{\langle v_0,\ldots,v_s\rangle}$ is, up to a scale factor, uniquely determined by the sequence $\Theta := (Pv_j)_j$. For this reason, it has become standard to denote the typical simplex spline by

$$M(\cdot|\Theta),$$

with Θ some finite sequence in $\mathbb{R}^d$ (the images under P of the vertices of the underlying simplex) and to choose the underlying simplex to have unit volume, whence $\int_{\mathbb{R}^d} M(\cdot|\Theta) = 1$. This is entirely consistent with the notation $M(\cdot|\Theta)$ used earlier for the univariate B-spline.

The relative neglect simplex splines have experienced in spite of the fact that they were the first multivariate B-splines to be considered may have several reasons.

Box splines, like their univariate antecedents, the cardinal B-splines (see Schoenberg's monograph (1969)), lead very quickly to a rich mathematical theory, as exemplified by the beautiful results of Dahmen and Micchelli (announced in Dahmen and Micchelli (1984)). This theory concerns mainly the shift-invariant space spanned by the integer translates of one box spline, and these are pp spaces with a regular partition Δ, and this regularity makes them amenable to Fourier transform techniques.

In contrast, the simplex splines were expected to be the multivariate equivalent of the general univariate B-spline, of use in the understanding and handling of *arbitrary* multivariate spline spaces. Since any polytope is the essentially disjoint union of simplices, any multivariate B-spline is a linear combination of simplex splines. However, use of the recurrence relations for the evaluation of simplex splines turned out to be much more expensive than had been hoped, for the simple reason (Grandine (1986)) that the recurrrence relation connects a d-variate simplex spline to at least $d+1$ simplex splines of one order less, while it connects it to at most two simplex splines of one order higher. Further, as already pointed out, for an arbitrary partition Δ and positive ρ, $\Pi^{\rho}_{k,\Delta}$ may not contain any compactly supported element unless k is very much larger than ρ. This means that, for k 'close' to ρ, only some suitably chosen refinement Δ' of Δ may support enough simplex splines so that their span has some approximation power. Unfortunately, the first scheme proposed for this (in Goodman and Lee (1981), Dahmen and Micchelli (1982) and Höllig (1982)) did not lead to a spline space with easily constructed quasi-interpolant schemes. However, very recently, a scheme has become available, in Dahmen, Micchelli and Seidel (1992), that, in hindsight, appears to be the 'right' one. It is based on the multivariate 'B-patch' of Seidel (1991). Given a triangulation Δ, it provides a suitable basis of simplex splines for the space $\Pi^{k-1}_{k,\Delta'}$, with Δ' obtained, in effect, as the roughest partition that contains all the cells for the simplex splines employed, thus known, at least in principle, once these simplex splines are in hand. These simplex splines are all possible ones of the form

$$M(\cdot|V^{\beta}),$$

where

(i) V is a $(d+1)$-set with $\langle V\rangle \in \Delta$;
(ii) $\beta \in \mathbb{Z}^V_+$ with $|\beta| = k$;
(iii) $V^{\beta} := \{v_j : 0 \le j \le \beta(v); v \in V\}$;
(iv) the points v_j are obtained, by choosing, for each v in the vertex set $V(\Delta) := \cup_{\langle V\rangle\in\Delta} V$ for Δ, k additional points $v_1, \ldots, v_k \in \mathbb{R}^d$, and setting $v_0 := v$.

The only condition imposed upon the choice of these additional points v_j, $j = 1, \ldots, k$, $v \in V(\Delta)$, is the following. For any $(d+1)$-set V with $\langle V\rangle \in \Delta$,

$$\Omega_{V,k} := \bigcap\{\langle (v_{\beta(v)})_{v\in V}\rangle : \beta \in \mathbb{Z}^V_+; |\beta| \le k\} \neq \emptyset.$$

Under these assumptions, Seidel (1992) proves that, for any $f \in \Pi^{k-1}_{k,\Delta}$,

$$f = \sum_{V,\beta} M(\cdot|V^{\beta}) w(V,\beta) F_V(V^{\beta-\mathbf{i}}), \qquad (5.5)$$

with $w(V,\beta)$ certain explicitly known normalizing factors, with

$$\beta - \mathbf{i} : v \mapsto \beta(v) - 1,$$

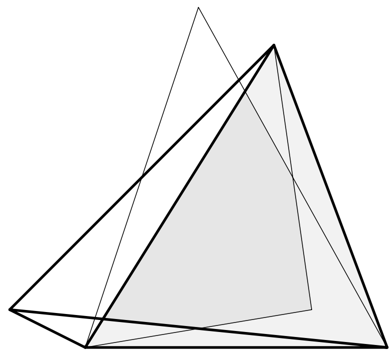

Fig. 5.4. With $k = 1$, the triangle $\langle V\rangle$ (lightly shaded and partially covered by) the set $\Omega_{V,k}$ (strongly shaded), and the meshlines (heavy) for one of the three related simplex splines.

hence $\#V^{\beta-\mathbf{i}} = |\beta| = k$, and with F_V the **blossom** of the polynomial which agrees with f on the cell $\langle V\rangle \in \Delta$. This means that F_V is the unique symmetric multi-linear form with k arguments for which

$$f(x) = F_V(x, x, \ldots, x), \quad \forall x \in \langle V\rangle.$$

The proof uses the validity of this result for any $f \in \Pi_k$, as established in Dahmen *et al.* (1992).

This is a most surprising and unexpected result. It captures completely the now standard formula for the coefficients in the B-spline expansion of an arbitrary *uni*variate spline as stated in de Casteljau (1963) and beautifully explained in Ramshaw (1987, 1989). It is to be hoped that the computational aspects of this formulation are equally favorable.

6. Approximation order

The treatment of approximation order given here follows in part the survey article de Boor (1992). The approximation power of a subspace S of $\Pi_{k,\Delta}$ is, typically, measured in terms of the **mesh(size)**

$$|\Delta| := \sup_{\delta\in\Delta} \operatorname{diam} \delta$$

of the partition Δ and the smoothness of the function f being approximated. The typical result is a statement of the following sort:

$$\operatorname{dist}(f, S) \le \operatorname{const}|\Delta|^r \|D^r f\|,$$

in which $\|D^r f\|$ is some appropriate measure of the derivatives of order r of f, and const is independent of f and Δ, provided Δ is chosen from some appropriate class of partitions. For example, the constant may, offhand, depend on the **uniformity measure**

$$R_\Delta := \sup_{\delta\in\Delta} \inf\{M/m : B_m(x) \subset \delta \subset B_M(y)\}$$

(with $B_m(x)$ the open ball with center x and radius m), hence be independent of Δ only if Δ is restricted to have $R_\Delta \le R$ for some finite R.

A particularly simple version of the approximation order of S is the following. One considers not just S, but the entire **scale** $(\sigma_h S)_h$ with

$$\sigma_h S \ := \ \{f(\cdot/h) : f \in S\},$$

and says that S **has (exact) approximation order** r and writes

$$\mathbf{ao}(S) \ = \ r,$$

provided

(i) for all 'smooth' f, $\operatorname{dist}(f, \sigma_h S) = O(h^r)$;
(ii) for some 'smooth' f, $\operatorname{dist}(f, \sigma_h S) \ne o(h^r)$.

By itself, (i) provides a *lower* bound for $\mathbf{ao}(S)$, and such lower bounds are usually established by exhibiting a particular approximation scheme, Q_h say, for which $\operatorname{ran} Q_h$ (= the range of Q_h) lies in $\subset \sigma_h S$, and $\|f - Q_h f\| \le \operatorname{const} h^r \|D^r f\|$. So-called quasi-interpolants are a favorite choice for the Q_h, of which more below.

By itself, (ii) provides an *upper* bound on $\mathbf{ao}(S)$, and there seems to be only duality (as made clear below) to establish such upper bounds.

Of course, for completeness, this definition requires specification of the norm in which the distance is to be measured, i.e., the normed linear space X in which the approximation is to take place. Typically, it is $L_p(G)$, with G some suitable subset of $\mathbb{R}^d$, and $p = 1, 2$ or ∞. It also requires a definition of 'smooth'. Often, it is sufficient to mean 'polynomial' or 'complex exponential'. However, it usually means that some norm involving certain derivatives is finite.

Somewhat more generally, one considers an indexed family $(S_h)_h$ of spaces, and denotes its approximation order, correspondingly, by $\mathbf{ao}((S_h)_h)$ to stress the fact that it is not (necessarily) obtained by scaling. In the latter situation, it turns out to be helpful to consider S_h to be of the form

$$S_h =: \sigma_h S^h.$$

If S^h is independent of h, we are back to the scaling case which, therefore, is also referred to as the **stationary** case, to distinguish it from the more general **nonstationary** case.

Questions of approximation order, particularly from (multivariate) pp spaces, have been dominated by what in Approximation Theory is called the

Strang-Fix theory, which, on careless reading, seems to imply that $\mathbf{ao}((S_h)_h)$ cannot be $\geq r$ unless $\Pi_{<r} \subset \cap_h S^h$. In fact, such a conclusion can only be reached in the stationary case, and even there only for very special situations. See Example 6.4 below for a simple counterexample; Ron (1991, 1992) and Beatson and Light (1992) treat approximation order specifically in the absence of polynomial reproduction. A similarly careless reading has also led to the *wrong* conclusion that, if all of $\Pi_{<r}$ is contained in each S^h locally, uniformly in h, then $\mathbf{ao}((S_h)_h) \geq r$. Even in the stationary case, the situation is more subtle, as is indicated in the subsections to follow. A first counter-example to that careless reading was given in de Boor and Höllig (1983).

In any event, the Strang-Fix theory applies only to the stationary case $S_h = \sigma_h S$, with S a *shift-invariant* space.

6.1. Shift-invariance

A collection S of functions on $\mathbb{R}^d$ is called **shift-invariant** if it is invariant under any translation by an integer, i.e., if

$$g \in S \quad \Longrightarrow \quad g(\cdot + \alpha) \in S \text{ for all } \alpha \in \mathbb{Z}^d.$$

For example, the space $\Pi^{\rho}_{k,\Delta}$ is shift-invariant in case Δ is shift-invariant in the sense that

$$\Delta + \alpha = \Delta \text{ for all } \alpha \in \mathbb{Z}^d.$$

Examples of interest include the three- and four-direction mesh popular in the bivariate box spline literature.

With $\ell_0(\mathbb{Z}^d)$ the collection of all *finitely* supported sequences $c : \mathbb{Z}^d \mapsto \mathbb{R}$, the simplest (nontrivial) example of a shift-invariant space is the space

$$\mathcal{S}_0(\varphi) := \Big\{ \sum_{\alpha \in \mathbb{Z}^d} \varphi(\cdot - \alpha)\, c(\alpha) : c \in \ell_0(\mathbb{Z}^d) \Big\}$$

of all finite linear combinations of the shifts of one (nontrivial) function, φ. This is the **shift-invariant space generated by** φ since it is the smallest shift-invariant space containing φ. Following de Boor, DeVore and Ron (1991), its closure, in whatever norm the context suggests, is denoted by

$$\mathcal{S}(\varphi) := \overline{\mathcal{S}_0(\varphi)}$$

and called the **principal shift-invariant**, or **PSI**, space generated by φ. For example, approximation by box splines has been discussed almost entirely in terms of the scale $(\sigma_h \mathcal{S}(\varphi))_h$ with φ a box spline.

More generally, if Φ is a finite collection of functions on $\mathbb{R}^d$, then one defines

$$\mathcal{S}_0(\Phi) := \sum_{\varphi\in\Phi} \mathcal{S}_0(\varphi)$$

and calls

$$\mathcal{S}(\Phi) := \overline{\mathcal{S}_0(\Phi)}$$

the **finitely generated shift-invariant**, or **FSI**, space, and calls Φ its set of **generators**. The structure of PSI and FSI spaces in $L_2(\mathbb{R}^d)$ is detailed in de Boor *et al.* (1991, 1992a), with particular emphasis on the construction of generating sets for a given FSI space having good properties (such as 'stability' or 'linear independence').

It is natural to consider approximations from $\mathcal{S}(\varphi)$ in the form

$$\varphi{*}c := \sum_{\alpha\in\mathbb{Z}^d} \varphi(\cdot - \alpha)\, c(\alpha) \tag{6.1}$$

for a suitable coefficient sequence c. However, offhand, such a sum makes sense only for finitely supported c, and one of the technical difficulties in ascertaining the approximation order of $\mathcal{S}(\varphi)$ derives from the fact that, in general, $\mathcal{S}(\varphi)$ may contain elements which cannot be represented in the form $\varphi{*}c$ for some sequence c, with the series $\varphi{*}c$ converging in norm. This is a problem even in the present context, where φ is, typically, some pp finite element and, in particular, compactly supported, hence the sum (6.1) converges pointwise (and even uniformly on compact sets) for arbitrary c. To give a simple example, from de Boor DeVore Ron (1992a), take for φ the Haar function, specifically $\varphi := \chi_{[-1..0)} - \chi_{[0..1)}$, with χ_I the characteristic function of the set I. Then $\mathcal{S}(\varphi) = \Pi_{0,\mathbb{Z}} \cap L_2(\mathbb{R})$ and, in particular, $\chi_{[0..1)} \in \mathcal{S}(\varphi)$. However, if the equation $\chi_{[0..1)} = \varphi * c$ is to hold even only in some weak sense, e.g., in the sense of pointwise convergence, then necessarily $c(\alpha) = c(0) + (\alpha - .5)^0_+$, all $\alpha \in \mathbb{Z}$, and $\varphi * c$ fails to converge in norm.

6.2. Quasi-interpolants

In the spline and finite-element literature, lower bounds for $\mathbf{ao}((S_h)_h)$ are usually obtained with the aid of a corresponding sequence $(Q_h)_h$ of linear maps, with $\operatorname{ran} Q_h \subseteq S_h$, which is a 'good quasi-interpolant sequence of order r' in the sense of the following definition.

Definition 6.2 *$(Q_h)_h$ is a* **good quasi-interpolant sequence of order** *r if it satisfies the following two conditions:*

(i) **uniformly local**: *For some h-independent finite ball B and all $x \in G$, $|(Q_h f)(x)| \le \text{const} \|f_{|x+hB}\|$;*
(ii) **polynomial reproduction**: *$Q_h f = f$ for all $f \in \Pi_{<r}$.*

For example, if $(\varphi)_{\varphi\in\Phi}$ is a **stable and local partition of unity**, i.e.,

$$\|\sum_{\varphi\in\Phi}|\varphi|\|_\infty < \infty, \quad \sup_{\varphi\in\Phi} \operatorname{diam}\operatorname{supp}\varphi < \infty, \quad \sum_{\varphi\in\Phi}\varphi = 1,$$

then $(\sigma_h Q \sigma_{1/h})_h$ with

$$Q : f \mapsto \sum_{\varphi\in\Phi} \varphi\, f(\tau_\varphi)$$

and $\tau_\varphi \in \operatorname{supp}\varphi$, all $\varphi\in\Phi$, is a good quasi-interpolant sequence of order 1.

As a more substantial example, it is part of the attraction of (5.5) that it provides an expansion of any $f\in\Pi_k$ in the form

$$f \;=\; \sum_{V,\beta} M(\cdot|V^\beta)w(V,\beta)\lambda_{V,\beta}(f), \tag{6.3}$$

with each $\lambda_{V,\beta}$ an explicitly known linear functional on Π_k. In particular (see Dahmen *et al.* (1992)) it is possible, as in the univariate case, to specify points $\tau_{V,\beta}$ so that the **Schoenberg operator**

$$Qf := \sum_{V,\beta} M(\cdot|V^\beta)w(V,\beta)f(\tau_{V,\beta})$$

reproduces every $f\in\Pi_1$. Since $\tau_{V,\beta}$ necessarily lies in the support of $M(\cdot|V^\beta)$ and this support is compact (and of the size of $\langle V\rangle$), it follows that $(\sigma_h Q\sigma_{1/h})_h$ is a good quasi-interpolant sequence of order 2. In fact, Dahmen *et al.* (1992) are able to lift the entire univariate quasi-interpolation argument (see, e.g., de Boor (1976)) to their multivariate setting, by showing the uniform linear independence of the functions $M(\cdot|V^\beta)w(V,\beta)$ which, in conjunction with (5.5), implies that any norm-preserving extension of $\lambda_{V,\beta}$ from $\Pi_k(\langle V\rangle)$ to some linear functional $\mu_{V,\beta}$, all V and β, provides a bounded linear projector

$$P : f\mapsto \sum_{V,\beta} M(\cdot|V^\beta)w(V,\beta)\mu_{V,\beta}(f)$$

onto the span of the simplex splines involved, and now, $(\sigma_h P\sigma_{1/h})_h$ is a good quasi-interpolant sequence of order $k+1$.

The term 'quasi-interpolant' is used in the finite element literature (see, e.g., Strang and Fix (1973)) to stress the fact that $Q_h f$ does not necessarily match function values at all the nodes of the finite elements used, but 'merely' reproduces certain polynomials. For a recent survey of the use of quasi-interpolants in spline theory, see de Boor (1990).

To recall, the standard use made of such a good quasi-interpolant sequence is to observe that, for arbitrary f and arbitrary $g\in\Pi_{<r}$,

$$|f(x) - Q_h f(x)| = |(1-Q_h)(f-g)(x)| \le \operatorname{const}\|(f-g)_{|x+hB}\|,$$

which provides a bound on $\|f-Q_h f\|$ in terms of how well f can be approximated from $\Pi_{<r}$ on a set of the form $x + hB$, giving the error bound $\text{const}_B h^r \|D^r f\|$ in which $\|D^r f\|$ measures the 'size' of the rth derivatives of f and which provides the desired $O(h^r)$. If our space X is L_p for some $p < \infty$, then this argument has to be fleshed out a bit (see, e.g., Jia and Lei (1991)).

There are certain costs associated with the quasi-interpolant approach, even when one only considers shift-invariant spaces with compactly supported generators. For example, it works, offhand, only with integer values of r. Also, offhand, it requires that $\cap_h S_h$ contain some non-trivial polynomial space. The artificiality of this last restriction is nicely illustrated by the following simple example, from Dyn and Ron (1990):

Example 6.4. Let $d = 1$, $p = \infty$, and let S_h be the span of the $h\mathbb{Z}$-translates of the piecewise linear function

$$\varphi_h : x \mapsto \begin{cases} x+1\,, & 0 \le x < h\,; \\ 0\,, & \text{otherwise.} \end{cases}$$

Thus S_h consists of certain piecewise linear functions, with breakpoint sequence $h\mathbb{Z}$, but the only polynomial (hence the only analytic function) it contains is the zero polynomial. In particular, it is not possible to construct a quasi-interpolant of positive order for it. Nevertheless, the approximation

$$Q_h f := \sum_{j \in h\mathbb{Z}} \varphi_h(\cdot - j) f(j)$$

has the error

$$f - Q_h f = f - \sum_{j \in h\mathbb{Z}} \chi_h(\cdot - j) f(j) \quad + \quad \sum_{j \in h\mathbb{Z}} (\chi_h - \varphi_h)(\cdot - j) f(j),$$

with χ_h the characteristic function of the interval $[0\,..\,h)$. Since $\|\chi_h - \varphi_h\|_\infty = h$,

$$\|f - Q_h f\|_\infty \le \omega_f(h) + \|f\|_\infty h,$$

where ω_f is the modulus of continuity of f. It follows that $Q_h f$ converges to f uniformly in case f is uniformly continuous and bounded. More than that, if f has a bounded first derivative, then $\|f - Q_h\|_\infty \le (\|Df\|_\infty + \|f\|_\infty)h$, giving approximation order 1 in the uniform norm.

This example could still be treated by an appropriate generalization of the notion of quasi-interpolant. Specifically, one could consider a good quasi-interpolant sequence (Q_h) of positive **local** order r, meaning that (Q_h) is uniformly local and that

$$Q_h f = f + O(\|f_{|B}\|\,|h|^r)$$

on hB for any $f \in \Pi_{<r}$. However, the point is made that a sequence $(S_h)_h$

of spaces does not need to contain a nontrivial polynomial space in order to have positive approximation order.

Finally, the quasi-interpolant approach is of no help with upper bounds.

6.3. The Strang-Fix condition

The literature on $\mathbf{ao}(\mathcal{S}(\varphi))$ for a compactly supported φ has been dominated by the Strang-Fix condition. It concerns the behavior of the Fourier transform

$$\widehat{\varphi} : \xi \mapsto \int_{\mathbb{R}^d} \varphi \, e_{-\xi}$$

of φ at the points of $2\pi\mathbb{Z}^d$. Here and below,

$$e_\theta : \mathbb{R}^d \to \mathbb{C} : x \mapsto \exp(i\theta^T x)$$

denotes the exponential function (with purely imaginary frequency $i\theta$). In one of its many versions, the Strang-Fix condition reads as follows.

Definition 6.5 *We say that φ satisfies SF_r in case*

(i) $\widehat{\varphi}(0) = 1$;
(ii) For all multi-indices α satisfying $|\alpha| < r$ we have $D^\alpha \widehat{\varphi} = 0$ on $2\pi\mathbb{Z}^d \backslash 0$.

Its importance derives from the following theorem (see Schoenberg (1946) for $d = 1$ and Strang and Fix (1973) for the general case), in which we use the convenient notation

$$\varphi *' f := \sum_{j \in \mathbb{Z}^d} \varphi(\cdot - j) f(j)$$

for the **semidiscrete convolution** of the two functions φ and f even if it requires further discussion of just what exactly is meant by it when the sum is not (locally) finite. Also, for any set X of functions on $\mathbb{R}^d$, we denote by

$$X_{\mathbf{c}}$$

the compactly supported functions in X.

Theorem 6.6 *For $\varphi \in L_1(\mathbb{R}^d)_{\mathbf{c}}$, the following are equivalent:*

(a) $\varphi *'$ is **degree-preserving on** $\Pi_{<r}$, *i.e., $\varphi *' p \in p + \Pi_{<\deg p}$, for all p in $\Pi_{<r}$;*
(b) φ satisfies SF_r.

The proof is via the Poisson summation formula (for which see, e.g., Stein and Weiss (1971; p. 252)). Starting with Strang and Fix (1973), the theorem is used to construct a good quasi-interpolant sequence (Q_h) of order r with $\operatorname{ran} Q_h \subseteq \sigma_h \mathcal{S}(\varphi)$. More than that, it forms part of an argument that seems to show that $\mathbf{ao}(\mathcal{S}(\varphi)) \geq r$ if and only if $\varphi/\widehat{\varphi}(0)$ satisfies SF_r. The precise statement of this equivalence for $X = L_2(\mathbb{R}^d)$ (see Strang and Fix (1973)) involves, unfortunately, a restricted notion of approximation order called 'controlled' approximation.

For $X = L_2(\mathbb{R}^d)$, the recent paper de Boor *et al.* (1991) contains a complete characterization of the approximation order of a not necessarily stationary scale of closed shift-invariant spaces. A crucial ingredient is the following theorem from the same reference, in which $P_S f$ denotes the orthogonal projection of f onto S, hence $\operatorname{dist}(f, S) = \|f - P_S f\|$.

Theorem 6.7 *Let S be a closed shift-invariant subspace of $L_2(\mathbb{R}^d)$, and let $f, g \in L_2(\mathbb{R}^d)$. Then*

$$\operatorname{dist}(f, S) \;\le\; \operatorname{dist}(f, \mathcal{S}(P_S g)) \;\le\; \operatorname{dist}(f, S) + 2\operatorname{dist}(f, \mathcal{S}(g)).$$

This theorem shows that the approximation power of a general shift-invariant subspace of L_2 is already attained by one of its PSI subspaces, provided one can, for given r, supply an element $g \in L_2(\mathbb{R}^d)$ for which $\mathbf{ao}(\mathcal{S}(g)) > r$. But that is easy to do, as follows.

Lemma 6.8 *There are simple functions g (e.g., the inverse Fourier transform of the characteristic function of some small neighborhood of the origin) for which, for any r,*

$$\operatorname{dist}(f, \sigma_h \mathcal{S}(g)) \;=\; o(h^r \|f\|_{W_2^r(\mathbb{R}^d)}).$$

Here,

$$\|f\|_{W_2^r(\mathbb{R}^d)} := \|(1 + |\cdot|)^r \widehat{f}\|_2.$$

For a directed family $(\sigma_h S^h)_h$ with each S^h a PSI space, de Boor *et al.* (1991) provides the following characterization of the approximation order, in which

$$\Lambda_\varphi := 1 - \frac{|\widehat{\varphi}|^2}{[\widehat{\varphi}, \widehat{\varphi}]} = \frac{\sum_{\alpha \in \mathbb{Z}^d \backslash 0} |\widehat{\varphi}(\cdot + 2\pi\alpha)|^2}{\sum_{\alpha \in \mathbb{Z}^d} |\widehat{\varphi}(\cdot + 2\pi\alpha)|^2},$$

$\mathbb{T}^d$ is the d-dimensional torus, i.e.,

$$\mathbb{T}^d \;:=\; [-\pi \mathinner{..} \pi]^d$$

with the appropriate identification of boundary points, and

$$[f, g] : \mathbb{T}^d \to \mathbb{C} : x \mapsto \sum_{\alpha \in \mathbb{Z}^d} f(x + 2\pi\alpha)\overline{g}(x + 2\pi\alpha)$$

is the very convenient **bracket product** of $f, g \in L_2(\mathbb{R}^d)$.

Theorem 6.9 *For any $(\varphi_h)_h$ in $X := L_2(\mathbb{R}^d)$,*

$$\mathbf{ao}((\sigma_h \mathcal{S}(\varphi_h))_h) \ge r \quad\Longleftrightarrow\quad \sup{}_h \Big\| \frac{\Lambda_{\varphi_h}}{(h + |\cdot|)^{2r}} \Big\|_{L_\infty(\mathbb{T}^d)} < \infty.$$

This result focuses attention on the behavior of Λ_{φ_h} near 0, hence, if $\widehat{\varphi}_h$ is bounded away from zero near 0 (uniformly in h), it focuses attention on the ratios

$$\widehat{\varphi}_h(\cdot + 2\pi\alpha)/\widehat{\varphi}_h \ , \qquad \alpha \in \mathbb{Z}^d\backslash 0. \tag{6.10}$$

Here is a typical corollary (from the same reference) which shows the relationship of this characterization to the Strang-Fix condition.

Corollary 6.11 *If $\varphi \in L_2(\mathbb{R}^d)$, and $1/\widehat{\varphi}$ is essentially bounded near 0, and $\widehat{\varphi} \in W_2^\rho(U)$ for some $\rho > r + d/2$ and some neighborhood U of $2\pi\mathbb{Z}^d\backslash 0$, and if φ satisfies SF_r, then $\mathbf{ao}(\mathcal{S}(\varphi)) \ge r$.*

Finally, as a consequence of Theorem 6.7 (and a good understanding of the structure of FSI spaces), de Boor *et al.* (1992a) obtains the following result which finishes a job left undone in Strang and Fix (1973) (see de Boor *et al.* (1992a) for historical commentary).

Theorem 6.12 *The approximation order in $L_2(\mathbb{R}^d)$ of the FSI space $\mathcal{S}(\Phi)$ with $\Phi \subset L_2(\mathbb{R}^d)$ is already attained by some PSI space $\mathcal{S}(\varphi)$ with $\varphi \in \mathcal{S}_0(\Phi)$.*

In particular, if Φ consists of compactly supported functions, then the 'super element' φ of the theorem is also compactly supported. This follows, more explicitly, from a representation of the Fourier transform of $P_S g$ as a sum of the form $\sum_{\varphi\in\Phi} \tau_\varphi \widehat{\varphi}$, in which the τ_φ are ratios of 2π-periodic functions, each a linear combination of products of functions of the form $[\widehat{\varphi}, \widehat{\psi}]$ with $\phi, \psi \in \Phi \cup \{g\}$. Now, for any particular r, it is possible to choose g compactly supported and such that $\mathbf{ao}(()\mathcal{S}(g)) \ge r$, while all the elements of Φ are compactly supported by assumption. This means that, with such a choice for g, each τ_φ is the ratio of two trigonometric polynomials, hence, there are trigonometric polynomials T_g, T_φ, $\varphi \in \Phi$, so that $T_g \widehat{P_S g} = \sum_{\varphi\in\Phi} T_\varphi \widehat{\varphi}$. This implies that the inverse Fourier transform of $T_g\widehat{P_S g}$ is in $\mathcal{S}_0(\Phi)$ and generates the same shift-invariant space as does $P_S g$, hence may be taken as the desired 'super-element'.

The paper de Boor and Ron (1991) deals with approximation from PSI spaces in $L_\infty(\mathbb{R}^d)$. The results are surprisingly similar in form, even if, due to the greater difficulties expected in this norm, there is a gap between lower and upper bounds for the approximation order obtained.

The main tool is Ron's (1991) surprisingly simple observation that, since

$$\varphi *' f = f *' \varphi \qquad \forall\, f \in \mathcal{S}(\varphi) \tag{6.13}$$

(as hinted at in Chui, Jetter and Ward (1987)), therefore

$$\varphi *' e_\theta - e_\theta *' \varphi = \varphi *' (e_\theta - f) - (e_\theta - f) *' \varphi, \qquad \forall\, f \in \mathcal{S}(\varphi)$$

(recall that $e_\theta : x \mapsto \exp(i\theta^T x)$), and this leads to the conclusion that

$$\|\varphi *' e_\theta - e_\theta *' \varphi\|_\infty \le 2\|\varphi *'\|_\infty \operatorname{dist}_\infty(e_\theta, \mathcal{S}(\varphi)), \tag{6.14}$$

with

$$\|\varphi *'\|_\infty := \| \sum_{\alpha\in\mathbb{Z}^d} |\varphi(\cdot - \alpha)| \, \|_\infty.$$

Since (as pointed out by A. Ron)

$$\frac{\varphi *' e_\theta - e_\theta *' \varphi}{e_\theta} \quad \sim \quad c + \sum_{\alpha\in\mathbb{Z}^d\backslash 0} \widehat{\varphi}(\theta + 2\pi\alpha)\, e_\alpha,$$

and the left-hand side has the same norm as $\|\varphi *' e_\theta - e_\theta *' \varphi\|_\infty$, this throws new light on the connection between $\mathbf{ao}(\mathcal{S}(\varphi))$ in L_∞ and the behavior of $\widehat{\varphi}$ 'at' $2\pi\mathbb{Z}^d\backslash 0$, and provides both upper and lower bounds for $\mathbf{ao}((\mathcal{S}(\varphi_h))_h)$.

As to lower bounds, these are obtained (in de Boor and Ron (1991)) by the approximation

$$f(\theta) = \int_{\mathbb{R}^d} e_\theta \widehat{f}/(2\pi)^d \sim \int_{\mathbb{R}^d} \varepsilon_\theta \widehat{f}/(2\pi)^d$$

(and a related one), with

$$\varepsilon_\theta := \varphi *' e_\theta / \sum_{\alpha\in\mathbb{Z}^d} \varphi(\alpha) e_{-\alpha}$$

an approximation from $\mathcal{S}(\varphi)$ to e_θ suggested by (6.14). In particular, the following theorem is proved there, in which $\mathcal{S}(\varphi)$ is not the norm-closure of $\mathcal{S}_0(\varphi)$ in $L_\infty(\mathbb{R}^d)$ but, in effect, the largest shift-invariant space containing $\mathcal{S}_0(\varphi)$ and satisfying (6.13). Also, the 'size' of the rth derivatives of f is measured in terms of its Fourier transform, as follows. It is assumed that f is 'smooth' in the sense that its Fourier transform is a Radon measure for which

$$\|f\|_{(r)} := \|(1 + |\cdot|^r)\widehat{f}\|_1 < \infty,$$

with the suffix '1' intended to indicate that the total variation of the measure in question is meant.

Theorem 6.15 *Assume that* $\|\varphi_h *'\| < \infty$ *for every* h. *Then, for any positive* η,

$$\operatorname{dist}(f, \sigma_h \mathcal{S}(\varphi_h)) \;\le\; h^r\,(2\pi)^{-d}\,\|f\|_{(r)}\,A \;+\; o(h^r)$$

with

$$A := \sup_h \sum_{\alpha\in\mathbb{Z}^d\backslash 0} \Big\| \frac{1}{(h^r + |\cdot|^r)} \frac{\widehat{\varphi}_h(\cdot + 2\pi\alpha)}{\widehat{\varphi}_h} \Big\|_{L_\infty(B_\eta)}.$$

Since this theorem gives $\mathbf{ao}((\sigma_h \mathcal{S}(\varphi_h))_h) \ge r$ only if $A < \infty$, this focuses, once again, attention on the behavior near zero of each of the ratios

$$\widehat{\varphi}_h(\cdot + 2\pi\alpha)/\widehat{\varphi}_h\,, \qquad \alpha \in \mathbb{Z}^d\backslash 0$$

mentioned already in (6.10). Specifically, in the stationary case, if this ratio is a smooth function in a neighborhood of 0, then the finiteness of A would require the ratio to have a zero of order r at 0, and conversely, provided $\widehat{\varphi}$ has some decay. From this vantage point, the Strang-Fix condition SF_r is seen to be neither necessary nor sufficient for $\mathbf{ao}(\mathcal{S}(\varphi)) \ge r$, but to come close to being necessary and sufficient for appropriately restricted φ.

The striking observation (6.14) actually provides more immediately an *upper* bound on the approximation order (see Ron (1991)). The main result of de Boor and Ron (1991) concerning this is the following.

Theorem 6.16 *Let (φ_h) be an indexed collection of elements of $X := L_\infty(\mathbb{R}^d)$. Assume that $\sup_h \|\varphi_h *'\| < \infty$, and that $\theta \in \mathbb{R}^d$.*
If $\mathrm{dist}(e_\theta, \sigma_h \mathcal{S}(\varphi_h)) = O(h^r)$, then

$$\sum_{\alpha \in \mathbb{Z}^d \backslash 0} |\widehat{\varphi}_h(h\theta + 2\pi\alpha)|^2 \le \mathrm{const}_\theta \, h^{2r}.$$

In particular, then

$$|\widehat{\varphi}_h(h\theta + 2\pi\alpha)| \le \mathrm{const}_\theta h^r \text{ for all nonzero } \alpha \text{ in } \mathbb{Z}^d.$$

Note that nothing is said here about $\widehat{\varphi}_h(0)$ (which is particularly important if $\widehat{\varphi}_h(0)$ is zero). On the other hand, it is easy to recover from this the rest of SF_r in the stationary case, i.e., in case $\varphi_h = \varphi$, for all h.

6.4. Upper bounds

Upper bounds for $\mathbf{ao}((S_h)_h)$ have to be fashioned separately for each case. available. However, one always employs duality, which provides the following well-known observation.

If Y is a linear subspace of the normed linear space X, and $\lambda \in X^*$ with $\lambda \perp Y$ (i.e., λ is a continuous linear functional on X which vanishes on all of Y), then, for any $x \in X$ and any $y \in Y$, $\lambda x = \lambda(x-y) \le \|\lambda\| \|x-y\|$, hence $|\lambda x| \le \|\lambda\| \,\mathrm{dist}(x, Y)$. In other words,

$$\lambda \perp Y \quad \Longrightarrow \quad \mathrm{dist}(x, Y) \ge \frac{|\lambda x|}{\|\lambda\|}.$$

For example, Ron's upper-bound argument mentioned in the preceding subsection is based on the linear map $f \mapsto \varphi *' f - f *' \varphi$ which vanishes on all of $\mathcal{S}(\varphi)$.

As a more direct example, consider $\mathbf{ao}(S)$ for

$$X = L_\infty(G), \quad S = \Pi^\rho_{k,\Delta}.$$

Assume without loss of generality that G is the d-dimensional cube,

$$G = C := [-1 \,.\,.\, 1]^d,$$

let δ be any cell in the partition Δ, and let g be any nontrivial homogeneous polynomial of degree $k+1$. If γ is the error in the best $L_2(\delta)$-approximation to g from Π_k, then the mapping

$$\lambda : L_\infty \to \mathbb{R} : f \mapsto \int_\delta \gamma f$$

(i) is a bounded linear functional;
(ii) is orthogonal to S, since all λ sees of $f \in S$ is its restriction to δ, and on δ each $f \in S$ is just a polynomial of degree $\le k$;
(iii) satisfies $\lambda g = \int_\delta \gamma\gamma > 0$.

Now consider $\lambda_h f := \int_\delta \gamma f(h\cdot)$. Then

(i) λ_h is a bounded linear functional, with h-independent norm

$$\|\lambda_h\| = \int_\delta |\gamma| \;=\; \lambda\,\text{signum}(\gamma),$$

where $\text{signum}(\gamma) : x \mapsto \text{signum}(\gamma(x))$.
(ii) $\lambda_h \perp S_h := \sigma_h S$, since $g \in S_h$ is of the form $f(\cdot/h)$ for some $f \in S$.
(iii) Using the homogeneity of g, one computes that

$$\lambda_h g = \int_\delta \gamma g(h\cdot) = h^{k+1}\int_\delta \gamma g = h^{k+1}\lambda g$$

with $\lambda g > 0$.

So, altogether,

$$\text{dist}(g, S_h) \ge h^{k+1}(\lambda g/\lambda\,\text{signum}(\gamma)),$$

showing that $\mathbf{ao}(\Pi^\rho_{k,\Delta}) \le k+1$.

If we try the same argument for $p < \infty$, we hit a little snag. Take, in fact, p at the other extreme, $p = 1$. There is no difficulty with (ii) or (iii), but the conclusion is weakened because (i) now reads

(i)′ $\|\lambda_h\| = \sup_{f\in L_1} |\int_\delta \gamma f(h\cdot)|/\|f\|_1 \le \|\gamma_{|\delta}\|_\infty \sup_{f\in L_1(\delta)} \int_\delta |f(h\cdot)|/\|f\|_1$,

and the best we can say about that last supremum is that it is at most h^{-d} since $\int_\delta f(h\cdot) = \int_{h\delta} f/h^d$. Hence, altogether, $\|\lambda_h\| \le \text{const}/h^d$. Thus, now our bound reads

$$\text{dist}_1(g, S_h) \ge h^{k+1}\text{const}/(\text{const}/h^d) \ne o(h^{k+1+d})$$

which is surely correct, but not very helpful.

What we are witnessing here is the fact that the error in a max-norm approximation is indeed localized, i.e., it occurs at a point, while, for $p < \infty$, the error 'at a point' is less relevant; the error is more global; one needs to consider the error over a good part of G. Further, in the argument below, I need some kind of uniformity of the partition Δ, of the following (very weak) sort (in which $|A|$ denotes the d-dimensional volume of the set A, and C continues to denote the cube $[-1\,..\,1]^d$):

Assumption 6.17 *There exists an open set b and a locally finite set $I \subset \mathbb{R}^d$ (meaning that I meets any bounded set only in finitely many points) so that*

(α) $b + I$ *is the* disjoint *union of* $b + i$, $i \in I$, *with each* $b + i$ *lying in some* $\delta \in \Delta$ *(the possibility of several lying in the same* δ *is not excluded);*
(β) *for some* const > 0 *and all* n, $|(b + I) \cap nC| \geq \text{const}|nC|$.

For example, any uniform partition of $\mathbb{R}$ satisfies this condition. As another example, if $d = 2$ and Δ is the three-direction mesh, then Δ consists of triangles of two kinds, and taking b to be the interior of one of these triangles and $I = \mathbb{Z}^2$ guarantees (α), while (β) holds with const $= 1/2$. On the other hand, Shayne Waldron (a student at Madison) has constructed a neat example to show that the Assumption 6.17 is, in general, necessary for the conclusion that $\mathbf{ao}(\Pi^{\rho}_{k,\Delta}) \leq k + 1$. The example uses $\rho = -1$ and arbitrary k, $d = 1$, $G = [-1\,..\,1]$, $p = 1$, and Δ obtained from $\mathbb{Z}$ by subdividing $[j\,..\,j + 1]$ into $2^{|j|}$ equal pieces, $j \in \mathbb{Z}$.

With Assumption 6.17 holding, define λ as before, but with b replacing the element δ of Δ. Further, assume without loss that $C \subseteq G$, and define

$$\lambda_h f := \int_b \gamma \sum_{i \in I_h} f(h \cdot + i),$$

where

$$I_h := \{i \in I : b + i \subseteq C/h\}.$$

This gives

(i)$_1$

$$\|\lambda_h\| \leq \sup_{f \in L_1} \frac{\sum_{i \in I_h} \int_{b+i} |\gamma||f(h\cdot)|}{\sum_{i \in I_h} \int_{h(b+i)} |f|} = \|\gamma_{|b}\|_\infty / h^d,$$

using the fact that the union $b + I_h$ is disjoint.

Hence, we have not worsened our situation here. Neither have we sacrificed (ii) because, by assumption, each $b+i$ lies in the interior of some $\delta \in \Delta$, and therefore $\int_b \gamma f(h \cdot + i) = 0$ for every $f \in S_h$. But we have materially improved the situation as regards (iii), for we now obtain

(iii)$_1$

$$\lambda_h g = \int_b \gamma \sum_{i \in I_h} g(h \cdot + i) = h^{k+1} \int_b \gamma \sum_{i \in I_h} g = h^{k+1} \text{const}\, \# I_h$$

with

$$\# I_h = |b + I_h|/|b| \geq \text{const}|C/h|/|b| = \text{const}/h^d.$$

With this, our conclusion is back to what we want:

$$\text{dist}\,_1(g, S_h) \geq (h^{k+1}\text{const}/h^d)/(\text{const}/h^d) \neq o(h^{k+1}).$$

Note that this lower bound on the distance only sees S as a space of pp's of degree $\le k$, hence is valid even when we take the biggest such space, i.e., the space $\Pi_{k,\Delta}$ of *all* pp functions of degree $\le k$ on the partition Δ. For this space, it is not hard to show that the approximation order is at least $k+1$, since approximations can be constructed entirely locally. Thus,

$$\mathbf{ao}(\Pi_{k,\Delta}) = k+1.$$

For this reason, this is called the **optimal** approximation order for a pp space of degree $\le k$.

Such a local construction of approximations is still possible for $\Pi^0_{k,\Delta}$, hence,

$$\mathbf{ao}(\Pi^\rho_{k,\Delta}) = k+1 \qquad \text{for} \quad \rho \le 0.$$

However, for $\rho > 0$, the story is largely unknown. Here are some working conjectures.

Conjecture (Ming-Yun Lai) *If* $\mathbf{ao}(\Pi^\rho_{k,\Delta}) = k+1$*, then* $\mathbf{ao}(\Pi^\rho_{k',\Delta}) = k'+1$ *for all* $k' \ge k$.

Conjecture $\mathbf{ao}(\Pi^\rho_{k,\Delta}) > 0 \implies \Pi^\rho_{k,\Delta}$ *contains elements with compact support.*

Conjecture $\mathbf{ao}(\Pi^\rho_{k,\Delta}) > 0 \implies \Pi^\rho_{k,\Delta}$ *contains a local partition of unity.*

First results (and more conjectures) can be found in de Boor and DeVore (1985) and Jia (1989).

Further illustrations of the use of duality in the derivation of upper bounds on $\mathbf{ao}(S)$ (albeit only for bivariate pp S) can be found in de Boor and Jia (199x) and its references. In particular, in conjunction with de Boor and Höllig (1988), it is proved there that, with Δ the three-direction mesh, the approximation order of $\Pi^\rho_{k,\Delta}$ (in the uniform norm) is $k+1$ (i.e., optimal) if and only if $k > 3\rho + 1$.

ACKNOWLEDGEMENT

It is a pleasure to thank Martin Buhmann for his many constructive comments on a draft of this paper.

REFERENCES

P. Alfeld, L.L. Schumaker and Maritza Sirvent (1992), 'On dimension and existence of local bases for multivariate spline spaces', *J. Approx. Theory* **70**, 243–264.

P. Alfeld, W. Whiteley and L.L. Schumaker (199x), 'The generic dimension of the space of C^1 splines of degree $d \ge 8$ on tetrahedral decompositions', *SIAM J. Numer. Anal.* **xx**, xxx–xxx.

R. K. Beatson and W. A. Light (1992), 'Quasi-interpolation in the absence of polynomial reproduction', Univ. of Leicester Mathematics Techn. Report 1992/15, May.

P.E. Bézier (1970), *Emploi des Machines à Commande Numérique*, Masson et Cie (Paris).

P.E. Bézier (1977), 'Essai de définition numérique des courbes et des surfaces expérimentale', dissertation, Univ. Paris.

L.J. Billera (1988), 'Homology of smooth splines: generic triangulations and a conjecture of Strang', *Trans. Amer. Math. Soc.* **310**, 325–340.

L.J. Billera (1989), 'The algebra of continuous piecewise polynomials', *Advances of Math.* **76(2)**, 170–183.

L.J. Billera and R. Haas (1987), 'Homology of divergence-free splines', Cornell.

L.J. Billera and L.L. Rose (1989), 'Gröbner basis methods for multivariate splines', in *Mathematical Methods in Computer Aided Geometric Design* (T. Lyche and L. Schumaker, eds), Academic Press (New York), 93–104.

Louis J. Billera and Lauren L. Rose (1991), 'A dimension series for multivariate splines', *Discrete & Computational Geometry* **6**, 107–128.

G. Birkhoff and C.R. de Boor (1965), 'Piecewise polynomial interpolation and approximation', in *Approximation of Functions* (H.L. Garabedian, ed), Elsevier (Amsterdam), 164–190.

W. Boehm, G. Farin and J. Kahmann (1984), 'A survey of curve and surface methods in CAGD', *Comput. Aided Geom. Design* **1**, 1–60.

C. de Boor (1976), 'Splines as linear combinations of B–splines, a survey', in *Approximation Theory II* (G. G. Lorentz, C. K. Chui and L. L. Schumaker, eds), Academic Press (New York), 1–47.

C. de Boor (1978), *A Practical Guide to Splines*, Springer Verlag (New York).

C. de Boor (1987), 'B–form basics', in *Geometric Modeling: Algorithms and New Trends* (G. E. Farin, ed), SIAM Publications (Philadelphia), 131–148.

C. de Boor (1990), 'Quasiinterpolants and approximation power of multivariate splines', in *Computation of Curves and Surfaces* (W. Dahmen, M. Gasca and C. Micchelli, eds), Kluwer, 313–345.

C. de Boor (1992), 'Approximation order without quasi-interpolants', in *Approximation Theory VII* (E. W. Cheney, C. Chui and L. Schumaker, eds), Academic Press (New York), xxx–xxx.

C. de Boor and R. DeVore (1983), 'Approximation by smooth multivariate splines', *Trans. Amer. Math. Soc.* **276**, 775–788.

C. de Boor and R. DeVore (1985), 'Partitions of unity and approximation', *Proc. Amer. Math. Soc.* **93**, 705–709.

C. de Boor, R. DeVore and A. Ron (1991), 'Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$', CMS-TSR **92-2**, University of Wisconsin (Madison).

C. de Boor, R. DeVore and A. Ron (1992a), 'The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$', CMS-TSR **92-8**, University of Wisconsin (Madison).

C. de Boor and K. Höllig (1982), 'Recurrence relations for multivariate B-splines', *Proc. Amer. Math. Soc.* **85**, 397–400.

C. de Boor and K. Höllig (1983), 'Approximation order from bivariate C^1-cubics: a counterexample', *Proc. Amer. Math. Soc.* **87**, 649–655.

C. de Boor and K. Höllig (1988), 'Approximation power of smooth bivariate pp functions', *Math. Z.* **197**, 343–363.

C. de Boor, K. Höllig and S.D. Riemenschneider (1992), *Box Splines*, Springer-Verlag (Berlin).

C. de Boor and Rong-Qing Jia (199x), 'A sharp upper bound on the approximation order of smooth bivariate pp functions', *J. Approx. Theory* **xx**, xxx–xxx.

C. de Boor and A. Ron (1991), 'Fourier analysis of approximation orders from principal shift-invariant spaces', *Constr. Approx.* **x**, xxx–xxx.

P. de Casteljau (1963), 'Courbes et Surfaces à Pôles', André Citroën Automobiles SA (Paris).

P. de Casteljau (1985), *Formes à Pôles*, Hermes (Paris).

C.K. Chui (1988), *Multivariate Splines*, CBMS-NSF Reg. Conf. Series in Appl. Math., vol. 54, SIAM (Philadelphia).

C.K. Chui, K. Jetter and J.D. Ward (1987), 'Cardinal interpolation by multivariate splines', *Math. Comput.* **48**, 711–724.

C.K. Chui and M.J. Lai (1987), 'On multivariate vertex splines and applications', in *Topics in Multivariate Approximation* (C. K. Chui, L. L. Schumaker and F. Utreras, eds), Academic Press (New York), 19–36.

R. Courant (1943), 'Variational methods for the solution of problems in equilibrium and vibrations', *Bull. Amer. Math. Soc.* **49**, 1–23.

H.B. Curry and I.J. Schoenberg (1966), 'On the Pólya frequency functions IV: the fundamental spline functions and their limits', *J. Analyse Math.* **17**, 71–107.

W. Dahmen (1979), 'Multivariate B-splines—Recurrence relations and linear combinations of truncated powers', in *Multivariate Approximation Theory* (W. Schempp and K. Zeller, eds), Birkhäuser (Basel), 64–82.

W. Dahmen and C.A. Micchelli (1982), 'On the linear independence of multivariate B-splines I. Triangulations of simploids', *SIAM J. Numer. Anal.* **19**, 992–1012.

W. Dahmen and C.A. Micchelli (1983), 'Recent progress in multivariate splines', in *Approximation Theory IV* (C. Chui, L. Schumaker and J. Ward, eds), Academic Press (New York), 27–121.

W. Dahmen and C.A. Micchelli (1984), 'Some results on box splines', *Bull. Amer. Math. Soc.* **11**, 147–150.

W. Dahmen, C.A. Micchelli and H.-P. Seidel (1992), 'Blossoming begets B-spline bases built better by B-patches', *Math. Comput.* **59**, 97–115.

N. Dyn and A. Ron (1990), 'Local approximation by certain spaces of multivariate exponential-polynomials, approximation order of exponential box splines and related interpolation problems', *Trans. Amer. Math. Soc.* **319**, 381–404.

G. Farin (1977), 'Konstruktion und Eigenschaften von Bézier-Kurven und Bézier Flächen', Diplom-Arbeit, Univ. Braunschweig.

G. Farin (1986), 'Triangular Bernstein-Bézier patches', *Comput. Aided Geom. Design* **3**, 83–127.

G. Farin (1988), *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press (New York).

Ivor D. Faux and Michael J. Pratt (1979), *Computational Geometry for Design and Manufacture*, Ellis Horwood (Chichester).

R. Franke and L.L. Schumaker (1987), 'A bibliography of multivariate approximation', in *Topics in Multivariate Approximation* (C. K. Chui, L. L. Schumaker and F. Utreras, eds), Academic Press (New York), 275–335.

R.N. Goldman (1983), 'Subdivision algorithms for Bézier triangles', *Computer-Aided Design* **15**, 159–166.

T.N. Goodman and S.L. Lee (1981), 'Spline approximation operators of Bernstein-Schoenberg type in one and two variables', *J. Approx. Theory* **33**, 248–263.

T.A. Grandine (1986), 'The computational cost of simplex spline functions', MRC Rpt. 2926, Univ. Wis. (Madison WI).

K. Höllig (1982), 'Multivariate splines', *SIAM J. Numer. Anal.* **19**, 1013–1031.

K. Höllig (1986), 'Multivariate splines', in *Approximation Theory, Proc. Symp. Appl. Math.* **36** (C. de Boor, ed), Amer. Math. Soc. (Providence, RI), 103–127.

J. Hoschek and D. Lasser (1989), *Grundlagen der geometrischen Datenverarbeitung*, B.G. Teubner (Stuttgart).

Rong-Qing Jia (1989), 'Approximation order of translation invariant subspaces of functions', in *Approximation Theory VI* (C. Chui, L. Schumaker and J. Ward, eds), Academic Press (New York), 349–352.

Rong-Qing Jia (1992), 'Approximation by multivariate splines: an application of Boolean methods', ms.

Rong-Qing Jia and Junjiang Lei (1991), 'Approximation by multiinteger translates of functions having global support', *J. Approx. Theory* **x**, xxx–xxx.

A. Le Méhauté (1990), 'A finite element approach to surface reconstruction', in *Computation of Curves and Surfaces* (W. Dahmen, M. Gasca and C. Micchelli, eds), Kluwer, 237–274.

G.G. Lorentz (1953), *Bernstein Polynomials*, Toronto Press (Toronto).

N. Luscher (1987), 'Die Bernstein-Bézier Technik in der Methode der finiten Elemente', dissertation, Univ. Braunschweig.

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

J. Morgan and R. Scott (1990), 'The dimension of the space of C^1 piecewise polynomials', Research Report UH/MD-78, University of Houston.

N.E. Nörlund (1924), *Vorlesungen über Differenzenrechnung*, Grundlehren Vol. XIII, Springer (Berlin).

M.J.D. Powell (1981), *Approximation Theory and Methods*, Cambridge University Press (Cambridge).

L. Ramshaw (1987), 'Blossoming: a connect-the-dots approach to splines', Techn. Rep., Digital Systems Research Center (Palo Alto).

L. Ramshaw (1989), 'Blossoms are polar forms', *Comput. Aided Geom. Design* **6**, 23–32.

A. Ron (1991), 'A characterization of the approximation order of multivariate spline spaces', *Studia Math.* **98**, 73–90.

A. Ron (1992), 'The L_2-approximation orders of principal shift-invariant spaces generated by a radial basis function', in *Numerical Methods in Approximation Theory, ISNM 105* (D. Braess, L.L. Schumaker, eds), Birkhäuser (Basel), 245–268.

I. J. Schoenberg (1946), 'Contributions to the problem of approximation of equidistant data by analytic functions, Part A: On the problem of smoothing or graduation, a first class of analytic approximation formulas', *Quart. Appl. Math.* **4**, 45–99.

I.J. Schoenberg (1965), 'Letter to Philip J. Davis', dated May 31.

I.J. Schoenberg (1969), 'Cardinal interpolation and spline functions', *J. Approx. Theory* **2**, 167–206.

L.L. Schumaker (1979), 'On the dimension of spaces of piecewise polynomials in two variables', in *Multivariate Approximation Theory* (W. Schempp and K. Zeller, eds), Birkhäuser (Basel), 396–412.

L.L. Schumaker (1981), *Spline Functions: Basic Theory*, Wiley (New York).

L.L. Schumaker (1984), 'Bounds on the dimension of spaces of multivariate piecewise polynomials', *Rocky Mountain J. Math.* **14**, 251–264.

L.L. Schumaker (1988), 'Constructive aspects of bivariate piecewise polynomials', in *The Mathematics of Finite Elements and Applications VI* (J.R. Whiteman, ed), Academic Press (London), 513–520.

L.L. Schumaker (1991), 'Recent progress on multivariate splines', in *Mathematics of Finite Elements and Applications VII* (J. Whiteman, ed), Academic Press (London), 535–562.

H.-P. Seidel (1991), 'Symmetric recursive algorithms for surfaces: B-patches and the de Boor algorithm for polynomials over triangles', *Constr. Approx.* **7**, 257–279.

H.-P. Seidel (1991), 'Representing piecewise polynomials as linear combinations of multivariate B-splines', ms.

E. Stein and G. Weiss (1971), *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press (Princeton).

G. Strang (1973), 'Piecewise polynomials and the finite element method', *Bull. Amer. Math. Soc.* **79**, 1128–1137.

G. Strang (1974), 'The dimension of piecewise polynomials, and one-sided approximation', in *Numerical Solution of Differential Equations* (G. A. Watson,, ed), Springer (Berlin), 144–152.

G. Strang and G. Fix (1973), 'A Fourier analysis of the finite element variational method', in *Constructive Aspects of Functional Analysis* (G. Geymonat, ed), C.I.M.E. II Ciclo 1971, 793–840.

W. Whiteley (1991), 'A matrix for splines', in *Progress in Approximation Theory* (P. Nevai and A. Pinkus, eds), Academic Press (Boston), 821–828.

A. Ženišek (1970), 'Interpolation polynomials on the triangle', *Numer. Math.* **15**, 283–296.

A. Ženišek (1973), 'Polynomial approximation on tetrahedrons in the finite element method', *J. Approx. Theory* **7**, 334–351.

A. Ženišek (1974), 'A general theorem on triangular C^m elements', *Rev. Française Automat. Informat. Rech. Opér., Anal. Numer.* **22**, 119–127.