Polynomial interpolation in several variables

Mariano Gasca, [a,*] Thomas Sauer [b,**]

[a] *Department of Applied Mathematics, University of Zaragoza, 50009 Zaragoza, Spain,*
E-mail: gasca@posta.unizar.es
[b] *Mathematisches Institut, Universität Erlangen–Nürnberg, Bismarckstr* $1\frac{1}{2}$*, D–91054 Erlangen, Germany*
E-mail: sauer@mi.uni-erlangen.de

This is a survey of the main results on multivariate polynomial interpolation in the last twenty five years, a period of time when the subject experienced its most rapid development. The problem is considered from two different points of view: the construction of data points which allow unique interpolation for given interpolation spaces as well as the converse. In addition, one section is devoted to error formulas and another one to connections with Computer Algebra. An extensive list of references is also included.
Keywords: Interpolation, multivariate polynomials, Newton approach, divided differences, Gröbner bases, H–bases

1. Introduction

Interpolation is the problem of constructing a function p belonging to a (simple) finite dimensional linear space from a given set of data. Usually, the interpolation data are obtained by sampling another (more difficult) function and, in that case, it is said that p *interpolates* f, in the sense that both functions coincide on that data set. The simplest context to study here is interpolation by univariate polynomials. Therefore it is no surprise that interpolation by univariate polynomials is a very classical topic. However, interpolation by polynomials of several variables is much more intricate and is a subject which is currently an active area of research. In this paper we want to describe some recent developments in polynomial interpolation, especially those which lead to the construction

* Partially supported by DGES Spain, PB 96-0730
** Partially supported by DGES Spain, PB 96-0730 and Programa Europa CAI-DGA, Zaragoza, Spain

of the interpolating polynomial, rather than verification of its mere existence.

Let us denote by $x = (\xi_1, \ldots, \xi_d)$ any point of $\mathbb{R}^d$ and by Π^d the space of all d–variate polynomials with real coefficients. The subspace of polynomials of total degree at most n, denoted by Π_n^d, is formed by polynomials

$$p(x) = \sum_{\alpha \in \mathbb{N}_0^{d,n}} a_\alpha x^\alpha, \tag{1.1}$$

where $\mathbb{N}_0^{d,n}$ is the set of all (integer lattice) points $\alpha = (\alpha_1, \ldots, \alpha_d)$, $\alpha_i \geq 0$, $i = 1, \ldots, d$, with $|\alpha| = \alpha_1 + \cdots + \alpha_d \leq n$. In addition, the coefficients a_α, $\alpha \in \mathbb{N}_0^{d,n}$, are real constants and $x^\alpha = \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d}$. We also use the notation $\mathbb{H}^{d,n} := \mathbb{N}_0^{d,n} \setminus \mathbb{N}_0^{d,n-1}$.

This survey will be mainly concerned with the problem of finding a polynomial $p \in \Pi^d$ such that the values of p and/or some of its derivatives are prescribed real numbers at points $x_1, \ldots, x_N$ of $\mathbb{R}^d$. When derivatives are not interpolated, the problem is referred to as the *Lagrange interpolation problem* and can be stated in the following form:

Given a finite number of points $x_1, \ldots, x_N$, some real constants $y_1, \ldots, y_N$ and a subspace V of Π^d, find a polynomial $p \in V$, such that

$$p(x_j) = y_j, \qquad j = 1, \ldots, N. \tag{1.2}$$

The interpolation points x_i are also called *nodes* and V is the *interpolation space*.

1.1. The univariate case

There is a well–developed and extensive classical theory of univariate Lagrange polynomial interpolation. In this context the *Hermite interpolation problem* arises as a limiting case when some of the interpolation points coalesce, giving rise to derivatives of consecutive orders. The Lagrange problem with N different nodes $x_i \in \mathbb{R}$ or the Hermite problem, with m_i derivatives of consecutive orders $0, 1, \ldots, m_i - 1$ at each node x_i and $N = \sum_i m_i$, have always a unique solution in the space Π_{N-1}^1 of univariate polynomials of degree not greater than $N - 1$.

If there are some "gaps" in the order of the derivatives at some interpolation point, the problem is called a *Birkhoff interpolation problem*. In the univariate case, this problem has a well developed theory, see [77] for conditions ensuring its solvability. Multivariate Birkhoff interpolation will enter our discussion only

in a peripheral manner. The book [78] is the best source of information about this subject.

Returning to the univariate Lagrange interpolation problem, we recall that the *Lagrange formula*

$$p(x) = \sum_{i=1}^{N} y_i \ell_i(x), \tag{1.3}$$

where

$$\ell_i(x) = \prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{x - x_j}{x_i - x_j}, \qquad i = 1, \ldots, N, \tag{1.4}$$

explicitly provides the solution of the problem. Alternatively, a recursive form is obtained from the *Newton formula*, which makes use of *divided differences*. Specifically, we have that

$$p(x) = \sum_{i=1}^{N} f\,[x_1, \ldots, x_i] \prod_{j=1}^{i-1} (x - x_j), \tag{1.5}$$

where the divided difference $f[x_j, \ldots, x_k]$, $j \leq k$, is recursively defined by the equations

$$\begin{aligned} f\,[x_j] &= f\,(x_j) \\ f\,[x_j, \ldots, x_k] &= \frac{f\,[x_j, \ldots, x_{k-1}] - f\,[x_{j+1}, \ldots, x_k]}{x_j - x_k}. \end{aligned} \tag{1.6}$$

An important advantage of the Newton formula is that it can be easily extended to include the Hermite case.

Let us also mention the *Neville–Aitken formula*

$$p(x) = \frac{(x - x_1)p_1(x) - (x - x_N)p_2(x)}{x_N - x_1}, \tag{1.7}$$

where p_1, p_2 solve the interpolation problems at the nodes $x_2, \ldots, x_N$ and $x_1, \ldots, x_{N-1}$, respectively.

These formulas suggest a strategy of constructing the interpolating polynomial at N nodes from the solutions of some interpolation problems depending on less data. The Lagrange formula uses the solutions of N interpolation problems,

each of them with only one interpolation point. The Newton formula, written in the form

$$p(x) = \sum_{i=1}^{N-1} f[x_1, \ldots, x_i] \prod_{j=1}^{i-1} (x - x_j) + f[x_1, \ldots, x_N] \prod_{j=1}^{N-1} (x - x_j), \qquad (1.8)$$

tells us what has to be added to the solution of the problem with the $N-1$ points $x_1, \ldots, x_{N-1}$ to get the solution of the problem with the N points $x_1, \ldots, x_N$. Finally, the Neville–Aitken formula tells us how this solution can be obtained by combining the solutions of the two problems corresponding to the data points $x_2, \ldots, x_N$ and $x_1, \ldots, x_{N-1}$. A general interpolation formula including all these cases can be found in [46], where an application to bivariate problems is also given.

1.2. The multivariate case is a more difficult problem

Definition 1.1. Let V be an N–dimensional linear space of continuous functions. The Lagrange interpolation problem (1.2), for the points $x_1, \ldots, x_N \in \mathbb{R}^d$, is called *poised* in V if, for any given data $y_1, \ldots, y_N \in \mathbb{R}$, there exists a function $f \in V$ such that $f(x_j) = y_j$, $j = 1, \ldots, N$. When the Lagrange interpolation problem for any N distinct points in $\mathbb{R}^d$ is poised in V, then V is called a *Haar space* of order N.

Haar spaces exist in abundance for $d = 1$. The situation for $d > 1$ is dramatically different. In fact in this case there are no Haar spaces of dimension greater than one. For refinements of this important result see [41,76,81], and for useful concepts related to the poisedness of interpolation problems, for example almost regularity and singularity, see [78].

In particular, whenever $\mathcal{P}$ is a finite dimensional space of polynomials in d variables of dimension $N > 1$, there always exist nodes $x_1, \ldots, x_N$ and a nontrivial polynomial $p \in \mathcal{P}$ such that $p(x_j) = 0$, $j = 1, \ldots, N$. Consequently, deciding if the interpolation problem (1.2) is poised for $\mathcal{P}$ is difficult.

On the other hand, if we allow in the definition of a Haar space of order N, where N is the number of interpolation points, that V may have dimension greater than N, then their existence is ensured for any d. For example, in this sense, the space of polynomials of total degree at most $N-1$ on $\mathbb{R}^d$, which we denote by Π_{N-1}^d, is a Haar space of order N. To see this, let $x_1, \ldots, x_N$ be any N distinct points in $\mathbb{R}^d$. For each point x_i we choose a hyperplane H_i (identified

with its defining affine function) containing x_i, but not x_j, $j \neq i$. Then the polynomial

$$p = \sum_{i=1}^{N} y_i \prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{H_j}{H_j(x_i)} \tag{1.9}$$

belongs to Π_{N-1}^d and satisfies (1.2). However, for $d > 1$, the Haar space of order N of least dimension is yet to be determined and only known for a few special cases.

So far we have only considered Lagrange interpolation. The meaning of Hermite interpolation in the multivariate case is richer in variety and depth. Before addressing this problem, it is convenient to establish some notation for the differential operators which will appear in the paper. For $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ we denote by D^α the differential operator

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f(x)}{\partial \xi_1^{\alpha_1} \cdots \partial \xi_d^{\alpha_d}}, \quad x = (\xi_1, \ldots, \xi_d),$$

and, for a polynomial p given as in (1.1), we write

$$p\,(D) = \sum_{\alpha \in \mathbb{N}_0^{d,N}} a_\alpha D^\alpha \tag{1.10}$$

for the associated differential operator. If v is a point in $\mathbb{R}^d$, we denote by D_v the directional derivative operator, which corresponds to the linear polynomial $p(x) = v \cdot x$, $x \in \mathbb{R}^d$, where $\cdot$ denotes the euclidian product in $\mathbb{R}^d$. Likewise, repeated directional derivative with respect to v, denoted by D_v^n, corresponds to the polynomial $p(x) = (v \cdot x)^n$, $x \in \mathbb{R}^d$.

The idea of Hermite interpolation is clear in the univariate case, namely, when some of the interpolation points coalesce, the interpolating polynomials converge to the Hermite interpolation polynomial which interpolates function values and derivatives. In general this does not hold true in two and more variables. Specifically, we consider the problem of interpolating f and its gradient at two distinct points $x_1, x_2 \in \mathbb{R}^2$. This problem is the limit of the Lagrange interpolation problem at the six points x_j, $x_j + he_1$, $x_j + he_2$, $j = 1, 2$, which is poised with respect to Π_2^2 for all $h \neq 0$ and almost all choices of x_1, x_2. However, the original Hermite interpolation problem is *never* poised in Π_2^2, for any choice of x_1, x_2.

An interpolation problem is called *singular* for a given space if the problem is not poised for any set of nodes (note that Lagrange interpolation problems are never singular). The Hermite problem above is singular in Π_2^2. Many results related to regularity, almost regularity and singularity of interpolation problems can be found in [6,78,108] and [73] in this volume.

There is no general agreement in the multivariate case on the definition of "Hermite interpolation", see also Theorem 6.1 and the discussion there. However it is very common to associate this name with problems whose data are function evaluations and derivatives at the same points, especially those which are of the type

$$D^{\alpha} f\left(x_i\right), \qquad \alpha \in \mathbb{N}_0^{d,q_i}, \quad i = 1, \ldots, N. \tag{1.11}$$

That is, there are $\binom{q_i+d}{d}$ interpolation data associated to the point x_i. When the number q_i is the same for each x_i we call this problem a *uniform* Hermite interpolation problem.

Very few books on Numerical Analysis include a section on multivariate interpolation. Curiously, three texts of the 1960s [69,3,64] do treat the subject, always inspired in the classical book by Steffensen [111], but more recent books ignore it, with the exception of [67]. See also [35] in the context of the finite element approach. Multivariate interpolation is only briefly mentioned in classical texts on Approximation Theory, for example [41] and [76]. However, there are at least two monographs [78,6] and several surveys, [44,50,72,7,9,110] among others, devoted totally or partially to the subject.

2. Construction of sets of interpolation points

Since the poisedness of multivariate polynomial interpolation depends on the geometric structure of the points at which one interpolates, there has been interest in identifying points and polynomial subspaces, for example Π_n^d, for which interpolation is poised. This problem has important applications in the context of finite element analysis where the construction of polynomial interpolants with good approximation order is crucial. We review in this section the best–known techniques for choosing interpolation points.

2.1. Interpolation by tensor product

Tensor product interpolation is the oldest method of extending the univariate theory. The interpolation points and space are obtained by tensor product of the univariate ones. The Lagrange formula and the Newton formula, with divided differences are easily extended to this problem, as it can be found in [69,3,64].

2.2. Interpolation space Π_n^d: from regular grids to natural lattices and pencils of hyperplanes

In this subsection we focus on various methods to choose points $x_1, \ldots, x_N$ in $\mathbb{R}^d$ such that the interpolation problem with respect to these points is poised in Π_n^d and moreover the Lagrange formula can be easily constructed. Clearly, this requires that $N = \dim \Pi_n^d = \binom{n+d}{d}$.

The first and most natural approach to choose such interpolation nodes is the triangular grid of the unit simplex formed by the points in $\frac{1}{n}\mathbb{N}_0^{d,n}$. In the bivariate case, this configuration has already been discussed in [4], [111] and in classical textbooks on numerical analysis, for example in [69,64]. These books also deal with the more general case of arrays formed by points (x_i, y_j), $0 \le i+j \le n$, where $\{x_i\}, \{y_j\}$, $i, j = 0, \ldots, n$, are two sets of $n+1$ distinct points. A Newton formula with bivariate (tensor product) divided differences is provided for this case. The bivariate array is triangular when x_i and y_j are ordered and uniformly spaced, and in this case a Lagrange formula for the interpolating polynomial is given in [69]. Note that this situation considers the "lower left" triangle in a rectangular grid and therefore still allows for the application of tensor product methods.

The triangular case appeared later in [98], where the notion of *principal lattices*, affinely equivalent to the triangular sets was introduced. It was this paper which apparently motivated the construction in the paper [33], written by Chung and Yao. Their approach is based on the idea of taking the intersections of hyperplanes as interpolation nodes, so that products of affine functions can be used to find the interpolation polynomial *explicitly*, which guarantees poisedness. According to [33], a set of $N = \binom{n+d}{d}$ points $\mathcal{X} = \{x_1, \ldots, x_N\}$ in $\mathbb{R}^d$ satisfies the GC condition (Geometric Characterization) if for each point x_i there exist hyperplanes G_{il}, $l = 1, 2, \ldots, n$, such that x_i is not on any of these hyperplanes,

and all points of $\mathcal{X}$ lies on at least one of them. Equivalently, we have that

$$x_j \in \bigcup_{l=1}^{n} G_{il} \Longleftrightarrow j \neq i \quad , \quad i,j = 1,2,\ldots,N. \tag{2.1}$$

In this case, the solution of the Lagrange interpolation problem is given explicitly by

$$p = \sum_{i=1}^{N} f(x_i) \prod_{j=1}^{n} \frac{G_{ij}}{G_{ij}(x_i)}, \tag{2.2}$$

where $G_{ij}(\cdot) = 0$ is the defining equation of G_{ij}.

In general, it is difficult to decide if a given set of points satisfies the GC condition, even for $\mathbb{R}^2$. However, there are several well-known interesting special cases. For example, let $r_0, r_1, \ldots, r_{n+1}$ be $n+2$ straight lines in $\mathbb{R}^2$ such that any two of them r_i, r_j intersect at exactly one point x_{ij} and these points have the property that $x_{ij} \neq x_{kl} \Longleftrightarrow \{i,j\} \neq \{k,l\}$. Then the set $\mathcal{X} = \{x_{ij} : 0 \leq i < j \leq n+1\}$ satisfies the GC condition and formula (2.2) reads as

$$p = \sum_{i=0}^{n} \sum_{j=i+1}^{n+1} f(x_{ij}) \prod_{\substack{k=0 \\ k\neq i,j}}^{n+1} \frac{r_k}{r_k(x_{ij})}. \tag{2.3}$$

The set $\mathcal{X}$ is called a *natural lattice* of order n in [33].

Other examples of sets with the GC condition, are provided in [71]. To this end, we recall that a *pencil of order* n in $\mathbb{R}^d$ is a family of $n+1$ hyperplanes which either all intersect in an affine subspace of codimension 2 or are all parallel. The intersection (in the projective sense) of the hyperplanes of a pencil is called its *center*. We consider $d+1$ pencils of order n in $\mathbb{R}^d$ with centers $C_1, \ldots, C_{d+1}$ not contained in a hyperplane of the d-dimensional projective space $\mathbb{P}^d$, with the additional condition that there exist $\binom{n+d}{d}$ points, each of them lying precisely on $d+1$ hyperplanes, one from each pencil. In [71], the set of these points is called a $(d+1)$*–pencils–lattice* of order n. Some examples of these lattices and a method of constructing them can be found in [71] and [96], see also Figure 1.

When d of the $d+1$ pencils in $\mathbb{R}^d$ have pairwise orthogonal hyperplanes and the remaining one has a finite center, Lee and Phillips [70] called this lattice a *geometric mesh*. For $x = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$ and q a real number different from 0, 1 and -1, the *standard geometric mesh generated by* x *and* q is the set

$$\left\{ x_\alpha : x_\alpha \in \mathbb{R}^d, x_\alpha = (q^{\alpha_1}\xi_1, \ldots, q^{\alpha_d}\xi_d) \right\}, \tag{2.4}$$

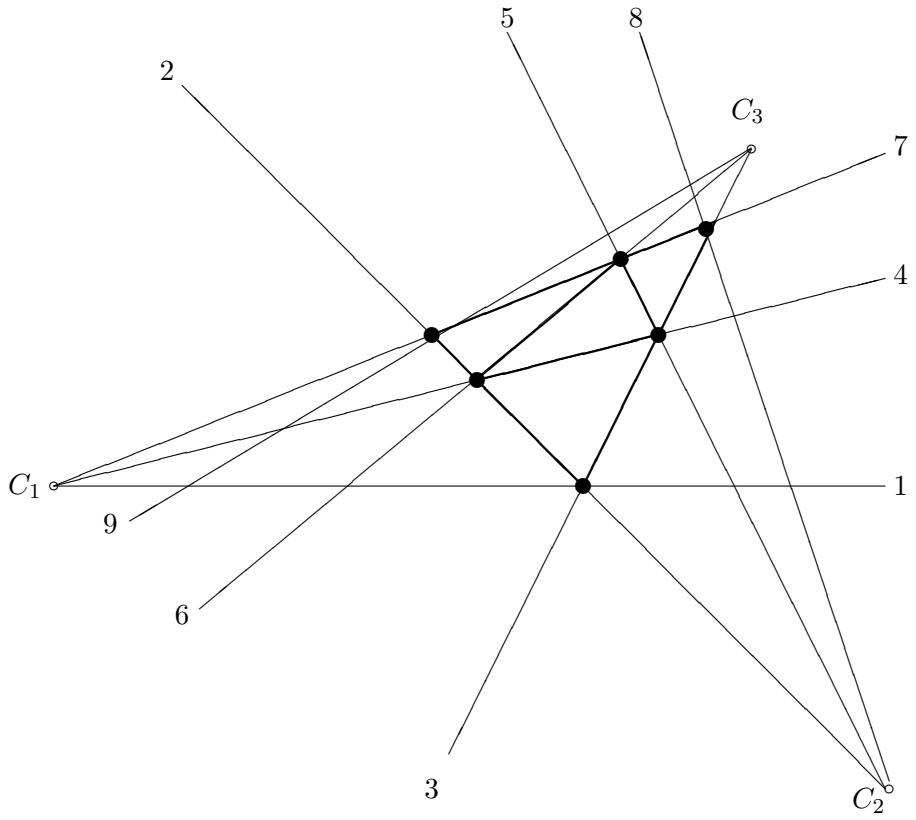

Fig.1: a 3–pencils–lattice of order 2 (6 interpolation points •) in $\mathbb{R}^2$ with 3 finite centers C_1, C_2, C_3

where

$$\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d, \qquad 0 \le \alpha_1 \le \ldots \le \alpha_d \le n. \tag{2.5}$$

The $(d+1)$-pencils-lattices of order n in $\mathbb{R}^d$ with $d+1$ infinite centers are called in [70] *regular meshes*, and are, obviously, the principal lattices of [98]. They appear as the limits of $(d+1)$-pencils-lattices when all the centers tend to infinity in such a way that they determine uniquely the infinite plane in $\mathbb{P}^d$.

As suggested by Figure 1, one can show in general that every $(d+1)$–pencils–lattice satisfies the GC condition. Hence the Lagrange interpolation problem defined by such a lattice is poised in Π_n^d and a Lagrange formula is available.

J. R. Busch [23] considered an extended version of the GC condition to deal with Hermite interpolation problems in $\mathbb{R}^d$ which he refers to as the HGC condition. Let $S = \{x_1, \ldots, x_M\}$ be a set of M distinct points of $\mathbb{R}^d$. We associate to each x_i a positive integer m_i and assume there exists n satisfying

$$\binom{n+d}{d} = \sum_{i=1}^{M} \binom{m_i - 1 + d}{d}.$$

We say that the HGC condition is satisfied if for each i $(1 \leq i \leq M)$ there exist $n_i = n - m_i + 1$ hyperplanes $G_{i1}, \ldots, G_{in_i}$ such that:

$$x_i \in \bigcup_{k=1}^{n_j} G_{jk} \Longleftrightarrow j \neq i \quad , \quad i, j = 1, \ldots, M. \tag{2.6}$$

The associated Hermite problem has all derivatives of orders up to $m_i - 1$ (that is $\binom{m_i - 1 + d}{d}$ derivatives) as interpolation data at x_i.

The most interesting example of set satisfying this condition arises from the extension of the concept of natural lattice of order n in $\mathbb{R}^d$. Specifically, in the bivariate case, the set consists of the intersection points of $n + 2$ lines such that the intersection of any two of them reduces to one point (we do not assume that different choices give different points). When three or more of the lines intersect at the same point, the associated interpolation problem becomes a Hermite problem.

It should be noted here that the explicit expression of the Lagrange formula for HGC lattices is quite complicated, see [73]. In fact, the examples given by Busch can be obtained alternatively from the Newton approach which we shall describe below.

We should also mention that the Lagrange formula of the interpolating polynomial is used in the finite element method. In this context, any interested reader should be familiar with [34] and [35]. Moreover, the papers [72], [73] contain additional important references for this subject.

2.3. Choosing nodes and space

A special case of Bezout's theorem states that two planar curves of degree m and n, with no common component, intersect each other at exactly mn real or complex points, counting multiplicities, [114]. This classical result is the basis of a decomposition method for the multivariate interpolation problem, which has been used by many authors, for example Guenter and Roetman in [59](see also

[73]). The idea is to decompose an interpolation problem in Π_n^d into two simpler problems, one in Π_n^{d-1} and the other one in Π_{n-1}^d. If the two corresponding smaller problems are poised, then the initial problem is poised. To implement this procedure requires some conditions on the geometry of the points.

This idea was one of the motivations of Gasca and Maeztu in [47] for the extension of the univariate Newton interpolation formula to bivariate problems. One of the original features of that paper was to determine each interpolation point as the intersection of two straight lines and use the equations of the lines as factors in a basis of Π_n^2, similarly to the univariate basis

$$1, x - x_0, (x - x_0)(x - x_1), \ldots, (x - x_0)(x - x_1) \cdots (x - x_{n-1}). \tag{2.7}$$

An important property of this univariate basis is that the interpolation conditions give rise to a *triangular* system, and this happens also in the bivariate case in [47]. Another feature of [47] is that Hermite problems are solved as easily as Lagrange problems, as it happens in the univariate Newton formula. We recall some of these observations here.

As in [47], [28] or [44], an *interpolation system* in $\mathbb{R}^2$ is a set

$$\{(r_i, r_{ij}, x_{ij}) : (i, j) \in I\}, \tag{2.8}$$

where I is a lexicographically ordered index set

$$I = \{(i, j) \ : \ j = 0, \ldots, m_i, \, i = 0, \ldots, n\}, \tag{2.9}$$

and r_i, r_{ij} are straight lines with exactly one intersection point x_{ij}. Again, we use the notation r_i, r_{ij} to represent either the straight lines or the affine functions that (up to a constant factor) give rise to the equations of the lines. It should be noted that repetitions of lines (and also of points) are accepted with the only condition (mentioned above) that the line r_i, for any given i, intersects the lines r_{ij}. The interpolation system (2.8) is associated to an interpolation problem, whose interpolation data are defined by the linear functionals

$$L_{ij} f = D_{\rho_i}^{t_{ij}} D_{\rho_i}^{s_{ij}} f\,(x_{ij}), \tag{2.10}$$

where t_{ij} (respectively s_{ij}) is the number of lines in the list

$$r_0, r_1, \ldots, r_{i-1}, r_{i0}, r_{i1}, \ldots, r_{i,j-1} \tag{2.11}$$

which contain x_{ij} and coincide (resp. do not coincide) with r_i, and where ρ_i, ρ_{ij} are vectors in the directions of the lines r_i, r_{ij}. The directional derivatives in

(2.10) are evaluated at the intersection point x_{ij}. In the simplest case, $t_{ij} = s_{ij} = 0$ for all (i,j), only evaluations of f appear as data and the problem is a Lagrange interpolation problem, but in special cases it is an Hermite or even Hermite-Birkhoff (in the sense of gaps in the order of the derivatives) interpolation problem.

In this method the interpolation space is spanned by the polynomials

$$\phi_{ij} = r_0 r_1 \cdots r_{i-1} r_{i0} r_{i1} \cdots r_{i,j-1}, \quad (i,j) \in I, \tag{2.12}$$

(as usual, empty products due to negative subindices are taken as 1) and with this basis, a natural extension of the univariate one in (2.7), the interpolation conditions give rise to a triangular linear system which can be easily and recursively solved. In fact, as proved in [47], for $(k,l),(i,j) \in I$ one has

$$L_{ij}\phi_{kl} = 0, \quad (k,l) > (i,j), \tag{2.13}$$

$$L_{ij}\phi_{ij} \neq 0, \tag{2.14}$$

with the lexicographical order.

Therefore, with this method the interpolation space is associated to the geometric distribution of the interpolation points along straight lines. A simple argument shows that the interpolation space, spanned by the polynomials (2.12), is Π_n^2 if and only if there are $n+1$ data on the line r_0, n on r_1 and so on. Observe that this distribution is the one suggested above as a consequence of Bezout theorem and that the Gasca-Maeztu method provides a Newton formula for it. This configuration of points was called *DH-sets* by Chui and Lai in [31]. In the same paper DH-sets in $\mathbb{R}^k$ were defined similarly using hyperplanes.

Coming back to the GC condition mentioned in the preceding subsection, one observes that *all* the known examples of sets satisfying this condition (see [33]) are DH-sets. In fact it was conjectured in [47] that the GC condition with $\binom{n+2}{2}$ points in $\mathbb{R}^2$ implies that $n+1$ points are collinear. The conjecture has been proved affirmatively until $n = 4$ by J. R. Busch in [24] but it remains unproved (although highly probable) in general. On the contrary, it is very easy to find (see [47]) DH-sets which do not satisfy the GC condition. In summary, accepting the conjecture as true, as it is at least for natural lattices, principal lattices, 3-pencils lattices, etc. in $\mathbb{R}^2$, sets satisfying the GC condition would be a subclass of DH-sets. This subclass has very simple Lagrange formulas but the solution can also be easily found with the Newton approach of [47]. For the rest of DH-sets which

do not satisfy the GC condition the problem is solved by this Newton formula because they do not have a simple Lagrange formula. We remark in passing that there is a certain confusion in the literature between the GC condition and the notion of natural lattices. As we have mentioned, there are sets with the GC condition which are not natural lattices, but which are DH-sets.

The triangularity of the matrix of the linear system of the interpolation problem obtained with the Gasca-Maeztu method in [47] was used in [48] to compute bivariate Vandermonde and confluent Vandermonde determinants. These determinants were also computed in a different form in [31].

Except for DH-sets, for which the interpolation space is Π_n^2, the space spanned by (2.12) does not have an alternate characterization. One of the cases when that happens is when the lines r_i are all parallel to one of the coordinate axes, the lines r_{ij} parallel to the other axis, and the index set (2.9) satisfies

$$m_0 \geq \cdots \geq m_n. \tag{2.15}$$

In that case the polynomial space is spanned by the monomials

$$x^i y^j, \qquad j = 0, \ldots, m_i, \quad i = 0, \ldots, n, \tag{2.16}$$

whose set of exponents (i, j) forms what is called a *lower set*. That is, if (i, j) belongs to the set, then so too does (k, l) with $k \leq i$ and $l \leq j$. H. Werner [115] worked with these sets and they are frequently used in [78] and [94]. Lower sets are also important in the problem of minimal degree interpolation, see the next section and also [104].

An interesting interpolation problem which generalizes the above setup was given in [80] and has been analyzed in more detail in [53]. In these papers a *reversible system* is defined as an interpolation system (2.8), (2.9) which satisfies (2.15) and the following conditions:

1. For each j, r_{ij} does not depend on i, that is $r_{ij} = r'_j$ for any i.

2. If $x_{lj} = x_{ij}$ for $(l, j), (i, j)$ both in I, then the lines r_h, r_i are coincident.

3. If $x_{ik} = x_{il}$ for $(i, k), (i, l)$ both in I, then the lines r'_k, r'_l are coincident.

The name "reversible system" stems from the fact that the interpolation problem is symmetric in r_j and r'_j, that is, the interpolation system remains the same if their roles are interchanged.

In this case the interpolation data (2.10) associated to the interpolation system become

$$L_{ij}f = D^{s_i}_{\rho'_j} D^{t_j}_{\rho_i} f\,(u_{ij})\,, \tag{2.17}$$

and the interpolation space is spanned by the polynomials (see (2.12):

$$\phi_{ij} = \prod_{l=0}^{i-1} r_l \prod_{k=0}^{j-1} r'_k, \qquad j = 0,\dots,m_i, \quad i = 0,\dots,n, \tag{2.18}$$

as a generalization of (2.16). In [53] the poisedness of the problem is proved and moreover the minimality of the total degree of the interpolation space as well as formulas for the interpolation error (see sections 3 and 4 below) are studied there.

Recall that the idea of the Aitken–Neville scheme is to combine the solutions of two simpler but *similar* problems to obtain the solution of the initial problem: it was developed independently by A. C. Aitken [1] and E. H. Neville [97] to avoid the explicit use of divided differences for Lagrange interpolation. Here the emphasis is put on the similarity of the two smaller problems, in contrast to the preceeding decomposition. An extension of the Aitken–Neville formula to bivariate interpolation was obtained by Thacher and Milne in [112]. For example, consider an interpolation problem with ten interpolation points, namely those of the principal lattice $S = \mathbb{N}_0^{2,3}$, and the interpolation space Π_3^2. The solution p_S of this problem is obtained in [112] from the solutions $p_{S_i}, i = 1,2,3$, in Π_2^2 of 3 interpolation problems on principal lattices $S_i \subset S$ with 6 points each: $S_1 = \mathbb{N}_0^{2,2}$, $S_2 = \{(i,j) : (i,j) \in S,\, i > 0\}$, $S_3 = \{(i,j) : (i,j) \in S,\, j > 0\}$. Therefore, one has

$$p_S = l_1 p_{S_1} + l_2 p_{S_2} + l_3 p_{S_3}, \tag{2.19}$$

where l_i, $i = 1,2,3$ are appropriately chosen affine polynomials. In fact in this case these polynomials are the barycentric coordinates relative to the simplex $(0,0)$, $(3,0)$, $(0,3)$.

In [45] Gasca and Lebrón gave a general framework for this type of decomposition that in practice can only be done for sets with special structures, in particular, principal lattices and rectangular grids, see also [51,60]. For the application of some of these ideas in building interpolating wavelets see [90].

Let us also mention that computational issues for these interpolation schemes have been studied in [29,30,74,100].

3. Construction of interpolation spaces

In this section, we start with interpolation points and identify linear spaces of polynomials with which we can interpolate at these points. This idea was pursued by Kronecker [68] as early as 1865. Here we will focus on recent activities concerning this problem.

3.1. Least Interpolation

Except for the work on *Gröbner bases*, which will be presented in section 6, there has been little effort put into the construction of interpolation spaces for arbitrary interpolation nodes until recently.

In 1990, de Boor and Ron [14], constructed, for a given set of nodes, an interesting space of polynomials with which one can interpolate. We recall their construction now. To this end, we let $\mathcal{R}$ denote the algebra of formal power series with real coefficients. We introduce the map $\lambda : \mathcal{R} \to \Pi^d$ which associates to each power series its nonzero homogeneous term of *minimal* degree. For $f = \sum_{\alpha\in\mathbb{N}_0^d} f_\alpha x^\alpha$, this map is defined by setting

$$\lambda(f) = \min_{n\in\mathbb{N}_0} \left\{ p_n : p_n = \sum_{|\alpha|=n} f_\alpha x^\alpha,\ p_n \neq 0 \right\}. \tag{3.1}$$

Note that for any functional $\vartheta \in \Pi'$ there exists a formal power series $f_\vartheta \in \mathcal{R}$ such that, for any $p \in \Pi$, one has

$$\theta(p) = (p(D) f_\theta)(0).$$

For example, for $x \in \mathbb{R}^d$, the point evaluation δ_x is represented by $f_x(y) = e^{x\cdot y}$.

Definition 3.1. Let $\mathcal{X} = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be a finite set of distinct nodes. The *least interpolation space* $\mathcal{P}_l(\mathcal{X})$ is defined as

$$\mathcal{P}_l(\mathcal{X}) = \left\{ \lambda(f) : f = \sum_{j=1}^{N} a_j f_{x_j},\ a_j \in \mathbb{R} \right\}. \tag{3.2}$$

The space $\mathcal{P}_l(\mathcal{X})$ has the following properties.

Theorem 3.2 [14,16]. For any finite set of distinct points $\mathcal{X} = \{x_1, \ldots, x_N\}$, $\mathcal{P}_l(\mathcal{X})$ is an interpolation space which is degree reducing. That is, for any $q \in \Pi^d$ the interpolant $p \in \mathcal{P}_l(\mathcal{X})$, defined by $p(x_j) = q(x_j)$, satisfies $\deg p \leq \deg q$.

The issue of an algorithmic construction of $\mathcal{P}_l$ has been considered in [15,8] introducing and using the technique of *Gauss elimination by segments*. We will compare this technique to work in Computational Algebra in section 6. We also remark that the least interpolation approach can be extended to "ideal interpolation schemes" [5], which might also be understood as a very reasonable multivariate notion of Hermite interpolation. These perspectives again are strongly connected to the algebra of polynomial ideals.

3.2. Minimal degree interpolation

Motivated by the properties of the least interpolant stated in Theorem 3.2, we consider the following definition.

Definition 3.3. Let $\mathcal{X} = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be a finite set of distinct nodes. A linear space $\mathcal{P}(\mathcal{X}) \subset \Pi^d$ is called a *minimal degree interpolation space* with respect to $\mathcal{X}$ if it is a degree reducing interpolation space.

The name "minimal degree interpolation space" stems from the fact that, setting $n = \min\left\{k : k \in \mathbb{N}_0,\, \mathcal{P} \subset \Pi_k^d\right\}$, there exists no subspace of Π_{n-1}^d which allows unique interpolation. To see this, choose the *Lagrange basis* $\ell_1, \ldots, \ell_N$ for $\mathcal{P}$, defined by $\ell_j(x_k) = \delta_{jk}$, $j, k = 1, \ldots, N$. Then one of these basis polynomials, say ℓ_1, must be of degree n. However, the assumption that there exists some subspace of Π_{n-1}^d which allows unique interpolation implies the existence of a polynomial $q \in \Pi_{n-1}^d$ which satisfies $q(x_j) = \delta_{1j}$ and therefore the interpolant with respect to q in $\mathcal{P}$ is ℓ_1. But then $\deg \ell_1 > \deg q$ which contradicts degree reduction property of $\mathcal{P}$.

For a given set $\mathcal{X}$ of nodes there usually exists a multitude of minimal degree interpolation spaces with one (important) exception: For a given set $\mathcal{X}$ of nodes the minimal degree interpolation space $\mathcal{P}(\mathcal{X})$ is unique if and only if $\mathcal{P}(\mathcal{X}) = \Pi_n^d$ for some $n \in \mathbb{N}_0$. We now introduce the notion of a Newton basis for an interpolation space.

Definition 3.4 [104]. A set of polynomials $\left\{p_\alpha : \alpha \in I \subset \mathbb{N}_0^d\right\}$ is called a *Newton basis* with respect to the set $\mathcal{X}$ of distinct interpolation nodes if $\mathcal{X}$ can be indexed as $\mathcal{X} = \{x_\alpha : \alpha \in I\}$ such that

1. for any $\alpha, \beta \in I$ with $|\beta| \le |\alpha|$, one has

$$p_\alpha(x_\beta) = \delta_{\alpha,\beta}; \tag{3.3}$$

2. for any $n \in \mathbb{N}_0$ there is a decomposition

$$\Pi_n^d = \text{span } \{p_\alpha : |\alpha| \le n\} \oplus \left\{q \in \Pi_n^d : q(\mathcal{X}) = 0\right\}. \tag{3.4}$$

These *graded* Newton bases allow for an algorithmic approach to multivariate polynomial interpolation and a recursive construction of the solution of the problem. Moreover, the concept of a Newton basis is actually equivalent to minimal degree interpolation.

Theorem 3.5 [104]. A subspace $\mathcal{P} \subset \Pi^d$ has a Newton basis with respect to $\mathcal{X}$ if and only if it is a minimal degree interpolation space with respect to $\mathcal{X}$.

This enables us to resolve the question of uniqueness of minimal degree interpolation spaces: let $\mathcal{P} \subset \Pi^d$ be a minimal degree interpolation space with respect to the node set $\mathcal{X}$ and let p_α, $\alpha \in I$, be a Newton basis for $\mathcal{P}$. Then the set of polynomials

$$\left\{p_\alpha + q_\alpha : \alpha \in I,\, q_\alpha \in \Pi_{|\alpha|}^d,\, q_\alpha(\mathcal{X}) = 0\right\}$$

is another Newton basis with respect to $\mathcal{X}$ and any Newton basis can be obtained in this way. Hence, the Newton basis and the minimal degree interpolation space $\mathcal{P}$ are unique if and only if $\Pi_n^d \cap \{q : q(\mathcal{X}) = 0\} = \{0\}$. Observe that the concept of Newton basis given in this section is different from that of subsection 2.3, where, instead of (3.3) and (3.4), only (2.13) and (2.14) were assumed.

4. Remainder formulas

Remainder formulas for polynomial interpolation give explicit representations for the interpolation error. Let L_n denote the interpolation operator

$$L_n : C(\mathbb{R}^d) \to \Pi_n^d, \tag{4.1}$$

defined by

$$L_n f(x_j) = f(x_j), \qquad j = 1, \ldots, \binom{n+d}{d},$$

where we assume that the nodes $x_j \in \mathbb{R}^d$ make the interpolation problem poised. In this section we study multivariate generalizations of the formula

$$f(x) - L_n f(x) = f[x, x_0, \ldots, x_n] \prod_{j=0}^{n} (x - x_j)$$
$$= \left(\prod_{j=0}^{n} (x - x_j) \right) \int_{\mathbb{R}} f^{(n+1)}(t) M(t|x_0, \ldots, x_n)\ dt, \quad (4.2)$$

where $f \in C^{n+1}(\mathbb{R})$ and $M(\cdot|x_0, \ldots, x_n)$ denotes the B–spline with knots $x_0, \ldots, x_n$, normalized such that

$$\int_{\mathbb{R}} M(t|x_0, \ldots, x_n)\ dt = \frac{1}{(n+1)!}.$$

Formula (4.2) also holds for repeated nodes. In particular, in the case that $x_0 = \cdots = x_n$, (4.2) yields an error formula for the *Taylor polynomial* $T_n f$,

$$f(x) - T_n f(x) = (x - x_0)^{n+1} \int_{\mathbb{R}} f^{(n+1)}(t) M(t|x_0, \ldots, x_0)\ dt. \quad (4.3)$$

For the multivariate case $d > 1$, we use the *simplex spline*, introduced by Micchelli, cf. [85,86,89]. Recall that the simplex spline $M(\cdot|x_0, \ldots, x_n)$ with knots $x_0, \ldots, x_n \in \mathbb{R}^d$ is defined as the distribution which satisfies

$$\int_{\mathbb{R}^d} f(t) M(t|x_0, \ldots, x_n)\ dt = \int_{\Delta_n} f(u_0 x_0 + \cdots + u_n x_n)\ du, \qquad f \in C\left(\mathbb{R}^d\right), \quad (4.4)$$

where

$$\Delta_n = \left\{ u : u = (u_0, \ldots, u_n),\ u_j \geq 0,\ \sum_{j=0}^{n} u_j = 1 \right\}$$

is the standard n–simplex. Following the notation introduced in [86], we are pleased to write

$$\int_{[x_0, \ldots, x_n]} f := \int_{\mathbb{R}^d} f(t) M(t|x_0, \ldots, x_n)\ dt. \quad (4.5)$$

We remark that distributions of the form

$$[x_0, \ldots, x_n; y_1, \ldots, y_n] f := \int_{[x_0, \ldots, x_n]} D_{y_1} \cdots D_{y_n} f, \qquad x_0, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}^d,$$

are called a *multivariate divided difference* in [10]; see also [85,86] where it has been pointed out how these distributions provide generalizations of univariate

divided differences (see also [58, p. 671] for another possibility based on an observation by Jacobi going back to 1835).

With this notation at hand and the distributional properties of the simplex spline, the multivariate remainder formula for the Taylor polynomial becomes

$$f(x) - T_n f(x) = \int\limits_{[x,x_0,\ldots,x_0]} D^{n+1}_{x-x_0} f, \qquad x_0, x \in \mathbb{R}^d. \tag{4.6}$$

Note that this representation is in essence a *univariate* formula since it is combined directly from univariate Taylor polynomials along the line segment $[x_0, x]$, cf. [107].

Let us now consider different remainder formulas for Lagrange interpolation.

4.1. Multipoint Taylor expansions and Ciarlet's formula

We begin with an observation by Ciarlet which has first been stated in [36] and has later been extended in [34]. This remainder formula is based on using the error representation (4.6) for the Taylor polynomial $T_n f$. Indeed, let $x_1, \ldots, x_N \in \mathbb{R}^d$, $N = \binom{n+d}{d}$, be nodes which allow unique polynomial interpolation of degree n. Then, there exist unique *Lagrange fundamental polynomials* p_j, $j = 1, \ldots, N$, of degree n, such that $p_j(x_k) = \delta_{jk}$, $j, k = 1, \ldots, N$. Moreover, fix $x \in \mathbb{R}^d$ and let $T_n f$, $j = 1, \ldots, N$, denote the Taylor polynomial of order n at x. Choosing $x = x_j$ in (4.6), we have

$$f(x_j) = T_n f(x_j) + \int\limits_{[x_j,x,\ldots,x]} D^{n+1}_{x_j-x} f$$

and since T_n and L_n are projections on Π^d_n we conclude that

$$\begin{aligned} L_n f &= \sum_{j=1}^{N} f(x_j)\ p_j = \sum_{j=1}^{N} T_n f(x_j)\ p_j + \sum_{j=1}^{N} p_j \int\limits_{[x_j,x,\ldots,x]} D^{n+1}_{x_j-x} f \\ &= T_n f + \sum_{j=1}^{N} p_j \int\limits_{[x_j,x,\ldots,x]} D^{n+1}_{x_j-x} f. \end{aligned}$$

Hence, evaluating at x, and noting that $T_n f(x) = f(x)$, we obtain that

$$f(x) - L_n f(x) = \sum_{j=1}^{N} p_j(x) \int\limits_{[x_j,x,\ldots,x]} D^{n+1}_{x-x_j} f, \tag{4.7}$$

which is Ciarlet's error formula.

For later comparison, let us make some remarks on this formula. First, the number of terms in this sum always equals the number of interpolation points, independently of the number of variables. This shows that Ciarlet's formula is *not* a straightforward generalization of (4.2) since this formula consists of only one term. Moreover, the splines appearing in Ciarlet's formula are essentially *univariate* B–splines and the integration always is over line segments only. Hence, the function f under consideration has to be smooth on a domain which is star shaped with respect to the interpolation points.

4.2. The Newton approach and its remainder formula

More than twenty years after the appearance of Ciarlet's formula, which became an essential tool in the analysis of finite elements [35], a different approach to remainder formulas was developed in [107]. This method is based on a multivariate Newton approach. Let us describe the setup. Suppose that distinct points $x_1, \ldots, x_N \in \mathbb{R}^d$, $N = \binom{n+d}{d}$, are given which admit unique polynomial interpolation of total degree at most n. The first step is the construction of the *Newton fundamental polynomials* p_α, $\alpha \in \mathbb{N}_0^{d,n}$, (see Definition 3.4),

$$p_\alpha(x_\beta) = \delta_{\alpha,\beta}, \qquad |\beta| \le |\alpha|,$$

where $\left\{x_\alpha : \alpha \in \mathbb{N}_0^{d,n}\right\}$, is a re–ordering of the points $x_1, \ldots, x_N$. The construction of these polynomials and the re–ordering of the points can be effectively done by a Gram–Schmidt procedure, see [103]. The dual functionals with respect to this basis are the *finite difference* functionals which we define recursively, for $x \in \mathbb{R}^d$ and $f \in C\left(\mathbb{R}^d\right)$, as

$$\begin{aligned} &\lambda_0[x]f = f(x), \\ &\lambda_k[x]f = \lambda_{k-1}f[x] - \sum_{\alpha\in\mathbb{H}^{d,k}} \lambda_{k-1}[x_\alpha]\, f\, p_\alpha(x), \qquad k \ge 1. \end{aligned}$$

Theorem 4.1 [107]. For any poised interpolation problem in Π_n^d we have that

$$L_n f = \sum_{\alpha\in\mathbb{N}_0^{d,n}} \lambda_{|\alpha|}[x_\alpha]\, f\, p_\alpha \tag{4.8}$$

and

$$f - L_n f = \lambda_{n+1}[\cdot]f. \tag{4.9}$$

Because of equation (4.9), a remainder formula is obtained by finding a representation for the finite difference $\lambda_{n+1}[\cdot]f$ in terms of derivatives of f. This has been done in [107] and to state the result we introduce a new concept. By a *path* μ of length n we mean a vector

$$\mu = (\mu_0, \ldots, \mu_n), \qquad \mu_j \in \mathbb{N}_0^d, \quad |\mu_j| = j, \quad j = 0, \ldots, n,$$

of integer points (of increasing length). We denote the set of all paths of length n by Λ_n. Associated to any such path $\mu \in \Lambda_n$ is a set of interpolation points

$$X_\mu = \left\{ x_{\mu_j} \in \mathbb{R}^d : j = 0, \ldots, n \right\},$$

an nth order homogeneous differential operator

$$D_\mu^n = D_{x_{\mu_n} - x_{\mu_{n-1}}} \cdots D_{x_{\mu_1} - x_{\mu_0}}$$

and a number

$$\pi_\mu = \prod_{j=0}^{n-1} p_{\mu_j}\left(x_{\mu_{j+1}}\right).$$

With this notation at hand, we can state, for $x \in \mathbb{R}$ and $f \in C^{n+1}\left(\mathbb{R}^d\right)$, the following result from [107], see also [11],

$$f(x) - L_n f(x) = \sum_{\mu \in \Lambda_n} p_{\mu_n}(x) \pi_\mu \int\limits_{[X_\mu, x]} D_{x - x_{\mu_n}} D_\mu^n f. \tag{4.10}$$

Let us compare this formula (4.10) with formula (4.7). First, we note that for $d = 1$ the above (4.10) becomes (4.2) and is different from (4.7). For $d \geq 2$, however, the situation changes. The number of terms in the sum in (4.10) is

$$\prod_{j=0}^{n} \binom{d-1+j}{d}$$

which exceeds N, the number of terms in (4.7). On the other hand, (4.10) contains $\binom{n+d-1}{d-1}$ terms which depend on the point x and this number of terms is certainly less than N. Another difference between these two formulas is that (4.10) contains integration over "truly" d–dimensional domains provided that the convex hull of X_μ has nonzero d–dimensional volume for some path μ, a property satisfied by any generic selection of points.

Therefore, we see that these two remainder formulas are structurally different and provide us with alternate ways to describe the error in polynomial

interpolation. We remark that formula (4.10) was used in [38] to prove the convergence of trust region algorithms for unconstraint derivative–free optimization.

It is possible to extend these results to minimal degree interpolation, but the resulting formulas are intricate. We refer the interested reader to [104].

4.3. An error formula for interpolation on natural lattices

So far, we have considered remainder formulas which apply to any distinct points which admit polynomial interpolation. In this and the next subsection we present two results which pertain to restricted sets of interpolation points.

The first one is an elegant formula for the error in natural lattices interpolation which has been derived by de Boor [10,12]. The setup here is a set $\mathcal{H}$ of $n+d$ hyperplanes in $\mathbb{R}^d$ which are in *general position*, that is, any d hyperplanes $H_1, \ldots, H_d \in \mathcal{H}$ intersect in exactly one point and any two different selections of d hyperplanes intersect in different points.

We define

$$\mathcal{H}_k = \{\mathcal{K} : \mathcal{K} \subset \mathcal{H},\, \#K = k\}, \qquad k = 1, \ldots, d. \tag{4.11}$$

Since the hyperplanes are in general position we conclude that, for every $\mathcal{K} \in \mathcal{H}_d$ there exists a point $x_{\mathcal{K}}$ defined by

$$x_{\mathcal{K}} = \bigcap_{H \in \mathcal{K}} H.$$

Note that $x_{\mathcal{K}} \notin \mathcal{H} \setminus \mathcal{K}$, that is

$$H(x_{\mathcal{K}}) \neq 0, \qquad H \in \mathcal{H} \setminus \mathcal{K}.$$

Here we again identify the hyperplane H with the affine function $H \in \Pi_1^d$ such that $x \in H \Leftrightarrow H(x) = 0$. It has been proved in [33] that the points $x_{\mathcal{K}}$, $\mathcal{K} \in \mathcal{H}_d$, form a natural lattice and hence allow unique interpolation from Π_n^d. The Lagrange representation of the interpolant takes the simple form

$$L_n f = \sum_{\mathcal{K} \in \mathcal{H}_d} f(x_{\mathcal{K}}) \prod_{H \in \mathcal{H} \setminus K} \frac{H}{H(x_{\mathcal{K}})},$$

which becomes (2.3) in the bivariate case.

Since the hyperplanes are in general position, for any $\mathcal{K} \in \mathcal{H}_{d-1}$ there exists a line $L_{\mathcal{K}}$ defined as

$$L_{\mathcal{K}} = \bigcap_{H \in \mathcal{K}} H.$$

Let $v_{\mathcal{K}}$ be a vector of euclidian length 1 in the direction of the line $L_{\mathcal{K}}$. Moreover, note that for any $H \in \mathcal{H} \setminus \mathcal{K}$ the line $L_{\mathcal{K}}$ intersects H in the point $x_{\mathcal{K}\cup\{H\}}$. This leads us to introduce, for $\mathcal{K} \in \mathcal{H}_{d-1}$, the set of $n+1$ points

$$X_{\mathcal{K}} := \{L_{\mathcal{K}} \cap H : H \in \mathcal{H} \setminus \mathcal{K}\}.$$

Likewise, we define the polynomial

$$p_{\mathcal{K}} := \prod_{H\in\mathcal{H}\setminus K} \frac{H}{H(v_{\mathcal{K}}) - H(0)}, \qquad \mathcal{K} \in \mathcal{H}_{d-1}. \tag{4.12}$$

Note that this polynomial is independent of how we normalize the affine polynomials H associated with its hyperplane. Moreover, the denominator appearing in (4.12) is nonzero due to the assumption that the hyperplanes in $\mathcal{H}$ are in general position. We also observe that $p_{\mathcal{K}}(x_{\mathcal{J}}) = 0$ for all $\mathcal{K} \in \mathcal{H}_{d-1}$ and $\mathcal{J} \in \mathcal{H}_d$.

With this notation at hand, de Boor's error representation for the Chung–Yao interpolant takes the form

$$f - L_n f = \sum_{\mathcal{K}\in\mathcal{H}_{d-1}} p_{\mathcal{K}} \int_{[\cdot,X_{\mathcal{K}}]} D^{n+1}_{v_{\mathcal{K}}} f. \tag{4.13}$$

Note that formula (4.13) uses only directional derivatives along the lines $L_{\mathcal{K}}$ formed by the intersection of $d-1$ hyperplanes from $\mathcal{H}$. Moreover, if $x \in L_{\mathcal{K}}$ for some $\mathcal{K} \in \mathcal{H}_{d-1}$ then we obtain the *univariate* remainder formula. For general points $x \in \mathbb{R}^d$, however, the domains of all the integrals in (4.13) are non–degenerate triangles. The number of terms in the sum on the right hand side of (4.13) is now $\binom{n+d}{d-1}$ which equals $\dim \Pi^{d-1}_{n+1}$, and which is less than in (4.10). We also remark that the polynomials p_K, $K \in \mathcal{H}_{d-1}$, form an H–basis (cf. [93]) for the ideal

$$\left\{p \in \Pi^d : p(x_{\mathcal{K}}) = 0,\, \mathcal{K} \in \mathcal{H}_d\right\}.$$

More on these algebraic aspects later. For an application of (4.13) in a special case see [109].

4.4. The error in reversible systems interpolation

The next error representation formula to be considered is for interpolation by *reversible systems* as described in subsection 2.3. To this end, we analyze the underlying geometry in more detail. Let ρ_i, ρ'_j and η_i, η'_j denote vectors of euclidian norm 1 which are parallel and perpendicular to the lines r_i, r'_j, respectively. We

define $M = \max\{i+j : (i,j) \in I\}$ and, by choosing arbitrary additional lines, we extend our system of straight lines to $r_0, \ldots, r_{M+1}$ and $r'_0, \ldots, r'_{M+1}$. This then provides polynomials ϕ_{ij} for *any* $0 \le i, j \le M+1$. Next, we consider paths $\mu = (\mu_0, \ldots, \mu_{k+1})$ such that $\mu_0, \ldots, \mu_k \in I$, $\mu_{k+1} \notin I$ and, for all j, $\mu_{j+1} = \mu_j + e_1$ or $\mu_{j+1} = \mu_j + e_2$. We denote the totality of all such paths (of length at most $M+1$) by Λ^* and note that $\#\Lambda^* \le 2^{M+1}$. To each such path $\mu \in \Lambda^*$ we define directions

$$\eta_j(\mu) := \begin{cases} \eta_l & \text{if} \quad \mu_{j+1} = (l+1, l')\,,\ \mu_j = (l, l')\,, \\ \eta'_{l'} & \text{if} \quad \mu_{j+1} = (l, l'+1)\,,\ \mu_j = (l, l')\,, \end{cases} \qquad j = 0, \ldots, k,$$

and

$$\rho_j(\mu) := \begin{cases} \rho_l & \text{if} \quad \mu_{j+1} = (l+1, l')\,,\ \mu_j = (l, l')\,, \\ \rho'_{l'} & \text{if} \quad \mu_{j+1} = (l, l'+1)\,,\ \mu_j = (l, l')\,, \end{cases} \qquad j = 0, \ldots, k.$$

We write θ_μ for the accumulated angles between these two systems of lines,

$$\theta_\mu = \prod_{j=0}^{k} \eta_j(\mu) \cdot \rho_j(\mu),$$

and obtain the formula

$$f - Lf = \sum_{\mu \in \Lambda^*} \frac{\phi_{\mu_{k+1}}}{\theta_\mu} \int_{[X_\mu, \cdot]} D_{\rho_k(\mu)} \cdots D_{\rho_0(\mu)} f. \tag{4.14}$$

This formula is another instance of a phenomenon known from classical geometry which states: "there is no distance between lines". That is, the only geometrical quantity describing the relation between the lines is their angle of intersection. Because of this property, (4.14) also holds true for the corresponding Hermite interpolation problem.

We also remark that the two formulas (4.13) and (4.14) have a rather simple form and give rise to remainder estimates which strongly mimic the univariate case. However, one should have in mind that these two situations use the strong assumption that the Lagrange fundamental polynomials can be factored into a product of affine polynomials.

5. Kergin Interpolation

The interpolation schemes considered so far have one important property in common: the dimension of the interpolation spaces coincides with the number of interpolation conditions. A different approach has been taken by Kergin [66]

who constructed, for given nodes $x_0, \dots, x_n \in \mathbb{R}^d$, an interpolant of degree n. This interpolant becomes unique because of additional interpolation constraints. Precisely, we have the following result.

Theorem 5.1 (Kergin). For any points $x_0, \dots, x_n \in \mathbb{R}^d$ there is a unique mapping $P : C^n\left(\mathbb{R}^d\right) \to \Pi_n^d$ with the property that for any $f \in C^n\left(\mathbb{R}^d\right)$, any constant coefficient homogeneous differential operator $q(D)$, $q \in \Pi_n^d$, and any subset $J \subseteq \{0, \dots, n\}$, $\#J = \deg q + 1$, there exists a point x in the convex hull $[x_j : j \in J]$ such that

$$(q(D)Pf)(x) = (q(D)f)(x). \tag{5.1}$$

Due to the complexity of this condition it is not surprising that the main part of [66] consists of showing the *existence* of the above interpolation operator which was done in a non–constructive way. This issue was resolved constructively by Micchelli [86], see also [87]. It turns out that Kergin interpolation can actually be considered as an extension of the approach (1.9). Indeed, substitution of the univariate *Hermite–Genocchi formula*

$$f[x_0, \dots, x_n] = \int_{[x_0,\dots,x_n]} f^{(n)}, \qquad f \in C^n(\mathbb{R}),$$

which expresses the divided difference as a B–Spline integral, into the Newton formula (1.5) allows us to represent the interpolation polynomial $L_n f$ as

$$L_n f(x) = \sum_{j=0}^{n} \int_{[x_0,\dots,x_j]} f^{(j)} \prod_{k=0}^{j-1} (x - x_k), \qquad x \in \mathbb{R},$$

which can be rewritten in a "fancy" way as

$$L_n f(x) = \sum_{j=0}^{n} \int_{[x_0,\dots,x_j]} D_{x-x_0} \cdots D_{x-x_{j-1}} f, \qquad x \in \mathbb{R}. \tag{5.2}$$

Now, formula (5.2) already gives the Kergin interpolant P with the above properties by simply replacing the qualifier "$x \in \mathbb{R}$" by "$x \in \mathbb{R}^d$".

Theorem 5.2 (Micchelli [86]).The Kergin interpolant P is given as

$$Pf(x) = \sum_{j=0}^{n} \int_{[x_0,\dots,x_j]} D_{x-x_0} \cdots D_{x-x_{j-1}} f, \qquad x \in \mathbb{R}^d, \tag{5.3}$$

and the error of interpolation takes the form

$$(f - Pf)(x) = \int_{[x,x_0,\ldots,x_n]} D_{x-x_0} \cdots D_{x-x_n} f, \qquad x \in \mathbb{R}^d. \tag{5.4}$$

For more information on Kergin interpolation we particularly recommend the monograph [6] and the references therein.

6. Algebraic aspects

There is a strong connection between polynomial interpolation in several variables and the theory of polynomial ideals which has actually led to parallelism in the development of results. In this section we want to describe the main ideas and relate them.

We first recall some terminology from algebraic geometry. Instead of finite sets of points we now speak of *zero dimensional varieties*. To explain this notion, let $\mathbb{K}$ be an infinite field. For any finite set of polynomials, $\mathcal{F} \subset \Pi$, we define an *algebraic variety* $V(\mathcal{F})$ by setting

$$V(\mathcal{F}) = \{x \in \mathbb{K} : \mathcal{F}(x) = (f(x) : f \in \mathcal{F}) = 0\}.$$

Conversely, we can associate to any set $V \subset \mathbb{K}^d$ a set $\mathcal{I}(V)$ of polynomials, defined by

$$\mathcal{I}(V) = \{p \in \Pi : p(x) = 0,\ x \in V\}. \tag{6.1}$$

Note that $\mathcal{I}(V)$ is an ideal, i.e., it is closed under addition and multiplication by Π. Moreover, if $\#V < \infty$, then we already know what to do in order to find $\mathcal{I}(V)$: we find a minimal degree interpolation space with respect to V and, writing p_f for the interpolation polynomial with respect to $f \in \Pi$, we have that

$$\mathcal{I}(V) = \{f - p_f : f \in \Pi\}. \tag{6.2}$$

Clearly, any polynomial on the right hand side of (6.2) vanishes at V and any polynomial in $f \in \mathcal{I}(V)$ can be written that way – then even $p_f = 0$. Note, however, that it is crucial here that V is a *minimal degree* interpolation space. But we can also use this procedure to construct a *basis* for the ideal. For this

purpose we recall that a finite set $\mathcal{F} \subset \Pi$ of polynomials is called a *basis* for an ideal $\mathcal{I} \subset \Pi$ if

$$\mathcal{I} = \langle \mathcal{F} \rangle = \langle f : f \in \mathcal{F} \rangle = \left\{ \sum_{f \in \mathcal{F}} \phi_f f : \phi_f \in \Pi,\ f \in \mathcal{F} \right\}. \tag{6.3}$$

Recall that finiteness does not imply any restriction here since Hilbert's *Basissatz* tells us that *any* polynomial ideal has a finite basis. Now, the construction for an ideal basis by interpolation works as follows: we start with any basis of Π, i.e., a countable set $\{f_j : j \in \mathbb{N}_0\}$ of polynomials and an initial basis $\mathcal{F}_{-1} = \emptyset$ for the ideal. For any $j \in \mathbb{N}_0$, we check if $q_j := f_j - p_{f_j} = 0$ or if $q_j \in \langle \mathcal{F}_{j-1} \rangle$. If the answer is "no" in both cases, we set $\mathcal{F}_j = \mathcal{F}_{j-1} \cup \{q_j\}$, otherwise we continue with $\mathcal{F}_j = \mathcal{F}_{j-1}$. Since the sequence $\langle \mathcal{F}_j \rangle$, $j \in \mathbb{N}_0$, is an increasing sequence of ideals and since polynomial rings are Noetherian, this process has to terminate after a finite number of steps yielding even a *minimal basis* for the ideal $\mathcal{I}(V)$.

This simple construction immediately shows that there must be an intimate relation between polynomial interpolation and the construction of ideal bases. Indeed, there has been a parallel development of ideas in Computer Algebra on the one hand and in Numerical Analysis (in particular, the theory of multivariate polynomial interpolation) where both sides have been unaware of each other. In the remainder of this section we want to show and connect some of these approaches. For a more detailed exposition see also [93].

6.1. The setup – ideal interpolation schemes

A general *interpolation scheme* is given by a finite set $\Theta \subset \Pi'$ of continuous linear functionals, defined on the algebra of polynomials. To exclude trivial degenerate cases, we demand Θ to be *linearly independent*, i.e.,

$$\sum_{\vartheta \in \Theta} c_\vartheta \vartheta(p) = 0 \quad \forall p \in \Pi \qquad \Leftrightarrow \qquad c_\vartheta = 0, \quad \vartheta \in \Theta.$$

Following Birkhoff [5], we call Θ an *ideal interpolation scheme* if the set of all homogeneous solutions of the interpolation problem,

$$\ker \Theta = \{f \in \Pi : \Theta(f) = (\vartheta(f) : \vartheta \in \Theta) = 0\} \subset \Pi$$

is an ideal in Π. These interpolation schemes can be characterized completely in the following way, cf.[82,16]

Theorem 6.1. A finite set $\Theta \subset \Pi'$ is an ideal interpolation scheme if and only if Θ can be represented as

$$\Theta(f) = \{(q_{j,k}(D)f)(x_j) : k = 0, \dots, m_j - 1,\ j = 1, \dots, n\}, \qquad f \in \Pi,$$

where the polynomial spaces

$$\mathcal{Q}_j = \text{span } \{q_{j,k} : k = 0, \dots, m_j - 1\} \subset \Pi, \qquad j = 1, \dots, n,$$

are closed under differentiation, i.e., $q(D)\mathcal{Q}_j \subseteq \mathcal{Q}_j$, $q \in \Pi$.

In view of this result, the notion of "ideal interpolation scheme" is a very reasonable generalization of the univariate Hermite interpolation scheme, in particular, since, for $d > 1$, m–fold (common) zeros of (finite sets of) polynomials correspond to m–dimensional spaces of polynomials which are closed under differentiation, cf. [57]. For example, a triple zero of $f \in \Pi$ at a point $x_0 \in \mathbb{R}^2$ could either mean that

$$f(x_0) = \frac{\partial f}{\partial \xi_1}(x_0) = \frac{\partial f}{\partial \xi_2}(x_0) = 0$$

or that there exists a nonzero $y \in \mathbb{R}^2$ such that

$$f(x_0) = D_y f(x_0) = D_y^2 f(x_0) = 0.$$

Conversely, if $\mathcal{I}$ is a zero dimensional ideal, i.e., if the associated variety is finite, then there always exist a set Θ of *dual functionals* which may, however, be defined in a field extension of $\mathbb{K}$ – this corresponds to the fact that polynomials with real coefficients (like $p(x) = x^2 + 1$) might have complex zeros. And again, the spaces of "local" differential conditions from Theorem 6.1 are the natural generalization of the *multiplicity* of such a zero. We want to emphasize once more that in two and more variables multiplicity of a polynomial zero is a *structured* quantity and not just a matter of counting, cf. [57,84].

Now, suppose that we are given a zero dimensional ideal $\mathcal{I}$, then there exists a finite set Θ of linearly independent dual functionals such that

$$\mathcal{I} = \{p \in \Pi : \Theta(p) = 0\}.$$

These dual functionals define an ideal interpolation scheme and any two interpolation spaces $\mathcal{P}_1, \mathcal{P}_2 \subset \Pi$ (clearly, $\dim \mathcal{P}_1 = \dim \mathcal{P}_2 = \#\Theta$) are equivalent modulo $\mathcal{I}$, in the sense that for each polynomial $p \in \mathcal{P}_1$ there exists $q \in \mathcal{P}_2$ such that $p - q \in \mathcal{I}$ and reciprocally. More precisely, given any data $y \in \mathbb{K}^\Theta$, the affine

space of all solutions of the interpolation problem $\Theta(p) = y$ is given by $p^* + \mathcal{I}$, where p^* is any polynomial satisfying $\Theta(p^*) = y$.

6.2. Gauß elimination on the Vandermonde matrix

The straightforward approach to finding a solution of the interpolation problem

$$\Theta(p) = y, \qquad y \in \mathbb{K}^{\Theta}, \tag{6.4}$$

is to view it as a linear system and to determine (formally) bi–infinite coefficient vectors $c = \left(c_\alpha \in \mathbb{K} : \alpha \in \mathbb{N}_0^d\right)$ such that

$$\sum_{\alpha \in \mathbb{N}_0^d} c_\alpha \Theta\left(x^\alpha\right) = \left[\Theta\left(x^\alpha\right) : \alpha \in \mathbb{N}_0^d\right] \; c = y. \tag{6.5}$$

The matrix $\mathcal{V}_\Theta = \left[\Theta\left(x^\alpha\right) : \alpha \in \mathbb{N}_0^d\right]$ in equation (6.5) is called the *Vandermonde matrix* with respect to Θ. Since the functionals in Θ were assumed to be linearly independent, we know that $\operatorname{rank}\mathcal{V}_\Theta = \#\Theta$, hence there exist $\#\Theta$ linearly independent column vectors from $\mathcal{V}_\Theta$. In other words, there exists a set $A \subset \mathbb{N}_0^d$, $\#A = \#\Theta$, such that

$$\operatorname{rank}\ \left[\Theta\left(x^\alpha\right) : \alpha \in A\right] = \#\Theta.$$

Consequently, any such set A yields an interpolation space, namely the one spanned by $\{x^\alpha : \alpha \in A\}$. Usually, however, this set A is far from being unique, which raises the question about a good choice of A. This selection can be "automated" by using an appropriate version of Gaussian elimination as proposed in [22,82], see also [92]. The idea is to eliminate column–by–column, processing the multiindices with respect to some term order $\prec$. Recall that a term order is a total ordering on $\mathbb{N}_0^d$ which has 0 as minimal element and is compatible with addition in $\mathbb{N}_0^d$. In the process of elimination it can happen that a zero column, say with index α, is produced, that is, there exist coefficients c_β, $\beta \preceq \alpha$, $c_\alpha = 1$, such that

$$0 = \sum_{\beta \preceq \alpha} c_\beta \Theta\left(x^\alpha\right) = \Theta\left(\sum_{\beta \preceq \alpha} c_\beta x^\alpha\right).$$

This means that the polynomial

$$p(x) = \sum_{\beta \preceq \alpha} c_\beta x^\alpha$$

belongs to $\mathcal{I}$. But even more is true: recording these "unwanted" (in terms of polynomial interpolation) polynomials in a proper way yields a *Gröbner basis* (see [93] for details) for the ideal $\mathcal{I}$. And in fact, the goal in [22,82] was not to compute the interpolation polynomials but the Gröbner basis for ideal $\mathcal{I} = \ker \Theta$ which is implicitly defined by Θ. This approach and the connection to Gröbner bases has also been mentioned in [104]. Moreover, it is worthwhile to note that, though this approach formally performs elimination on an infinite matrix, it always works on a finite matrix only: as shown in [105] the set A of multiindices constructed above always satisfies

$$A \subset \left\{\alpha \in \mathbb{N}_0^d : |\alpha| \le \#\Theta - 1\right\}.$$

A different approach towards elimination "by segments" in the Vandermonde matrix has been taken in [15], collecting all monomials of the same total degree into one object and mutually orthogonalizing them to obtain uniqueness. The result of this elimination process is then the least interpolation space described above. We will, however, omit the technical details of this intricate method which is also described in [8] from the linear algebra point of view.

6.3. Choosing the right representer by reduction

We now turn back to the problem of choosing a "good" solution for the interpolation problem $\Theta(p) = y$, where $y \in \mathbb{K}^\Theta$ is some given data vector. As mentioned before, any two solutions p, p' of this problem differ exactly by an element from $\mathcal{I} = \ker \Theta$ and thus belong to the same equivalence class modulo $\mathcal{I}$. Hence, the problem of finding a "good" interpolation polynomial is equivalent to finding a "good" representer for this equivalence class.

One way of obtaining this representer is in fact motivated by the univariate case: let $x_0, \dots, x_n$ be (for the sake of simplicity) distinct points in $\mathbb{R}$ and let $p \in \Pi$ be *any* polynomial which satisfies

$$p(x_j) = y_j, \qquad j = 0, \dots, n,$$

for some given data y_j, $j = 0, \dots, n$. Of course, the "standard" interpolation polynomial would be the unique one of degree n which can be obtained from p "algebraically" by division with remainder. Indeed, if we write p in the form $p = q\omega + r$, $\omega = (x - x_0) \cdots (x - x_n)$, $\deg r < \deg \omega = n + 1$, then r is the desired polynomial and can be obtained by performing the Euclidean algorithm. But we can also interpret the above division in a slightly different way which allows for

generalization to the multivariate case: the fact that the remainder polynomial r has degree n is equivalent to the property that no multiple of the leading term x^{n+1} of ω divides *any* homogeneous term of r any more. Such a polynomial is called *reduced*.

Reduction is now the key to the multivariate case, where a division algorithm, straightforwardly extending the univariate Euclidean algorithm, can be defined to divide by a *finite family* of polynomials, cf. [39, p. 63 ff.] for an excellent introduction. The idea is to use some term order and try to cancel leading terms of a given polynomial by a monomial multiple of the leading term of some polynomial from the finite set of divisors to end up with a reduced remainder polynomial. Since in some steps of the algorithm there can be more than one divisor whose leading term divides the homogeneous target term, there can (and often will) be ambiguities due to which even the remainder may not be unique any more. There is, however, an important case where the reduced remainder is unique, namely if the divisors form a *Gröbner basis* of the ideal they generate. Recall that a finite set $\mathcal{P}$ of polynomials is a Gröbner basis for the ideal $\langle \mathcal{P} \rangle$ if any polynomial $f \in \langle \mathcal{P} \rangle$ can be written as

$$f = \sum_{p \in \mathcal{P}} q_p p,$$

where the degree (i.e., the $\prec$–maximal power with respect to a given term order $\prec$) of any term in the sum on the right hand side does not exceed the degree (in the same sense) of p. If the term order based notion of degree is replaced by the total degree, then the basis with the respective property is called an *H–basis*. H–bases have been introduced by Macaulay as early as 1916 [79], while Gröbner bases have been brought up by Buchberger in his doctoral thesis in 1965, cf. [19,21]. Actually, the first major application of Gröbner bases was to compute a unique representer for each equivalence class modulo the ideal $\mathcal{I}$, in order to generate a "multiplication table" modulo $\mathcal{I}$, that is, to describe the action of multiplication by a given polynomial $f \in \Pi$ as an automorphism on the equivalence class. As shown in [106] the idea of reduction can also be carried over to the grading by total degree, yielding a constructive approach to H–bases. For the details and more facts about H–bases, Gröbner bases and their applications, the reader is once more referred to [93].

6.4. Normal form interpolation

One can use the algebraic technique of reduction to generate interpolation spaces: associated to each equivalence class modulo ker Θ, hence associated to any family of solutions of an ideal interpolation scheme Θ with respect to a given right hand side y, there is a "standard" element of this equivalence class, i.e., a "standard" interpolant, which can be obtained by the reduction process. This representer is often called the *normal form* modulo $\mathcal{I} = \ker \Theta$. It can be shown that the set of all normal forms is a $\#\Theta$–dimensional linear subspace of Π which admits unique interpolation. Moreover, acting on Π, reduction and interpolation are the same operation. The "free parameter" in this process is the notion of degree; if the degree is chosen to be based on a term order, then one enters the Gröbner basis environment and generates an interpolation space which is generated by $\#\Theta$ monomials which are $\prec$–minimal, where $\prec$ is again the underlying term order. Moreover, as shown in [105], this approach can actually be interpreted as term order least interpolation using single monomials instead of homogeneous terms. In the total degree setting, on the other hand, the least interpolation space can be obtained by a properly chosen reduction process where one deals with H–Bases then. Conversely, if the set Θ of dual functionals is given, then the Gauß elimination techniques on the Vandermonde matrix yield, as a by–product, a Gröbner basis or an H–basis, respectively, depending on the elimination technique to be used.

Summarizing we can say that the strong duality between interpolation and the generation of "good" bases for zero dimensional ideals nicely connects ideas from both Numerical Analysis and Computer Algebra and allows for the application of methods and techniques from one field in the other.

7. Final remarks

7.1. Software for polynomial interpolation

Besides the theoretical study of multivariate polynomial interpolation there is still the problem of *how* to compute interpolation polynomials in a numerically efficient and stable way. Of course, the first way to do so is the "naive" one which first generates the Vandermonde matrix

$$\left[x_j^\alpha : j = 0, \ldots, N,\, \alpha \in A\right], \qquad \#A = N + 1,$$

and then applies a "standard solver" like the famous `LAPACK` [2] to it. Unfortunately, Vandermonde matrices are known to be rather ill-conditioned and a "blind" application of linear algebra cannot be expected to produce good results. This phenomenon, by the way, becomes more and more apparent as the number of variables increases, which is, for example the case in optimization problems, cf. [38]. From this perspective, it might be useful to have software available which makes use of the structural properties of polynomial interpolation to obtain algorithms with better numerical properties.

Right now there are, to our knowledge, two different packages which are freely available on the Internet and which correspond to two "competing" approaches to multivariate interpolation. The first one is a set of *m–files* (MATLAB command files) developed by C. de Boor on the basis of the computational treatment of least interpolation in [15], which can be downloaded via anonymous `ftp` from

`ftp://ftp.cs.wisc.edu/Approx`

The interested reader is also referred to [13] where also some important details of the implementation are provided.

The other software is **MPI**, a `C++` class library by T. Sauer based on the Newton approach from [107] which can be obtained from

`http://www.mi.uni-erlangen.de/~sauer/interpol`

The class library includes a templated class `polynomial` (i.e., it is possible to use coefficients of arbitrary type which supports the basic field operations) with the basic linear space operations and point evaluation. Interpolation polynomials are easily generated by submitting two arrays containing the nodes and the values to be interpolated there, and can be manipulated in various ways. For example, it is possible to add further interpolation points without having to compute the interpolation polynomial from scratch every time, an ability which is one of the very handy properties of any Newton–like scheme. Some numerical experiences with **MPI** are described in [103], details on the Horner scheme used in this software and its surprisingly good numerical properties are given in [100].

We will not try to make any comparison between these two packages since they are designed for completely different purposes and work in very different environments. An obvious fact, however, is that the MATLAB–routines are easier to handle and therefore are much more suitable for the casual user who is mainly

interested in interactively experimenting with polynomial interpolation without having to write and compile a program. On the other hand, the C++–routines in **MPI** are much more trimmed for efficiency, but in order to apply them clearly a certain background in programming is needed to use classes in an appropriate way.

The final question, however, is if polynomial interpolation is a reasonable method in practice at all. It is well–known from the univariate case that interpolation polynomials often oscillate up to a point which renders them useless – this is one of the reasons why spline functions are much more useful for the interpolation of large data sets. On the other hand, it is the *degree* of the polynomial which causes the oscillations (and also the difficulties emerging from roundoff errors, cf. [100]), so that in higher dimensions polynomial interpolation may be a reasonable tool for a "moderate" number of nodes. Up to now very little can be said on the usefulness of polynomial interpolation in serious applications so that we can only invite people to try the available routines and to report their experiences.

7.2. The bibliography

Here below we present a selection of over one hundred references which, in our opinion, is very representative of the development in the last 25 years of the subject, specially in the constructive approaches which have been mentioned in the preceding sections.

The list includes in some cases old papers which are related to the recent ones, and can be used as a source of further references of the different authors. For readers interested in multivariate polynomial interpolation we recommend, for example, to enlarge the bibliography cited here of de Boor, Ciarlet, Gasca, Goldman, Hakopian, Le Mehauté, R. Lorentz, Phillips, Sauer, and respective coauthors, who summarize the different approaches in different countries.

References

[1] A. G. Aitken. On interpolation by iteration of proportional parts without the use of differences. *Proc. Edinburgh Math. Soc.* **3** (1932), 56–76.

[2] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov and D. Sorensen. *LAPACK Users' Guide, Second Edition.* SIAM, 1995.

[3] I. S. Berezin and N. P. Zhidkov. *Computing Methods*. Addison–Wesley, 1965. Russian version in 1959.

[4] O. Biermann. Über näherungsweise Kubaturen. *Monatsh. Math. Phys.*, **14** (1903), 211–225.

[5] G. Birkhoff. The algebra of multivariate interpolation. In C.V. Coffman and G.J. Fix, editors, *Constructive Approaches to Mathematical Models*, pages 345–363. Academic Press Inc., 1979.

[6] B. Bojanov, H. A. Hakopian and A. Sahakian. *Spline functions and multivariate interpolation*, Kluwer Academic Publishers, 1993.

[7] C. de Boor. Topics in multivariate approximation theory. In *Topics in Numerical Analysis*, Lect. Not. Math. **965**, pages 39–78. Birkhäuser Verlag, 1982.

[8] C. de Boor. Gauss elimination by segments and multivariate polynomial interpolation. In R. V. M. Zahar, editor, *Approximation and Computation: A Festschrift in Honor of Walter Gautschi*, pages 87–96. Birkhäuser Verlag, 1994.

[9] C. de Boor. Polynomial interpolation in several variables. In R. de Millo and J. R. Rice editors, *Studies in Computer Science* , pages 87–119. Plenum Press, 1994.

[10] C. de Boor. A multivariate divided difference. In C. K. Chui and L. L. Schumaker, editors, *Approximation Theory VIII, Vol. 1: Approximation and Interpolation*, pages 87–96. World Scientific Publishing Co., 1995.

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

[12] C. de Boor. The error in polynomial tensor-product, and in Chung-Yao, interpolation. In A. Le Méhauté, C. Rabut, and L. L. Schumaker, editors, *Surface Fitting and Multiresolution Methods*, pages 35–50. Vanderbilt University Press, 1997.

[13] C. de Boor. Computational aspects of multivariate polynomial interpolation: Indexing the coefficients. This volume.

[14] C. de Boor and A. Ron. On multivariate polynomial interpolation. *Constr. Approx.*, **6** (1990), 287–302.

[15] C. de Boor and A. Ron. Computational aspects of polynomial interpolation in several variables. *Math. Comp.*, **58** (1992), 705–727.

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

[17] L. Bos. On certain configurations of points in $\mathbb{R}^n$ which are unisolvent for polynomial interpolation. *SIAM J. Approx. Theo.*, **64** (1991), 271–280.

[18] G. Brezinski. The Muhlbach-Neville Aitken algorithm and some extensions. *BIT*, **80** (1980), 444–451.

[19] B. Buchberger. Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aeq. Math.*, **4** (1970), 374–383.

[20] B. Buchberger. Gröbner bases: an algorithmic method in polynomial ideal theory. In N. K. Bose, editor, *Multidimensional System Theory*, pages 184–232, Reidel, 1985.

[21] B. Buchberger. An algorithmic criterion for the solvability of a system of algebraic equations. Translation of [19] by M. Abramson and R. Lumbert. In B. Buchberger and

F. Winkler, editors, *Groebner Bases and Applications (Proc. of the Conf. 33 Years of Groebner Bases)*, volume 251 of *London Math. Soc. Lecture Notes*, pages 535–545. Cambridge University Press, 1998.

[22] B. Buchberger and H. M. Möller. The construction of multivariate polynomials with preassigned zeros. In G. Goos and J. Hartmanis, editors, *Computer Algebra, EUROCAM '82, European Computer Algebra Conference*, pages 24–31, Springer Lecture Notes in Computer Science **144**, 1982.

[23] J. R. Busch. Osculatory interpolation in $\mathbb{R}^n$. *SIAM J. Num. Anal.*, **22** (1985), 107–113.

[24] J. R. Busch. A note on Lagrange interpolation in $\mathbb{R}^2$.. *Rev. Union Matem. Argent.*, **36** (1990), 33–38.

[25] A. L. Cavaretta, C. A. Micchelli and A. Sharma. Multivariate interpolation and the Radon transform. *Math. Z.*, **174** (1980), 263–279.

[26] A. L. Cavaretta, C. A. Micchelli and A. Sharma. Multivariate interpolation and the Radon transform, part II: Some further examples. In R. A. DeVore and K. Scherer, editors, *Quantitative Approximation*, pages 49–62, Academic Press, 1980.

[27] A. L. Cavaretta, T. Goodman, C. A. Micchelli and A. Sharma. Multivariate interpolation and the Radon transform, part III: Lagrange representation. In Z. Ditzian, A. Meir, S. D. Riemenschneider and A. Sharma, editors, *CMS Conference Proceedings Vol. 3*, pages 37–50, Canadian Mathematical Society 1983.

[28] J. M. Carnicer and M. Gasca. On finite element interpolation problems. In T. Lyche and L. L. Schumaker, editors, *Mathematical Methods in Computer Aided Geometric Design*, pages 105–113. Academic Press, 1989.

[29] J. M. Carnicer and M. Gasca. On the evaluation of multivariate Lagrange formulae. In W. Schempp edit.,*Multivariate Approximation Theory IV*, pages 65–72. Birkhauser V., 1989.

[30] J. M. Carnicer and M. Gasca. Evaluation of multivariate polynomials and their derivatives. *Math. Comp.*, **54** (1990), 231–244.

[31] C. K. Chui and M. J. Lai. Vandermonde determinants and Lagrange interpolation in $\mathbb{R}^s$. In B. L Lin, editor, *Nonlinear and Convex Analysis*. Marcel Dekker, 1988.

[32] C. K. Chui. Multivariate Splines. CBMS–NSF Regional Conference Series in Applied Mathematics **54**, SIAM, 1988.

[33] K. C. Chung and T. H. Yao. On lattices admitting unique Lagrange interpolation. *SIAM J. Num. Anal.*, **14** (1977), 735–743.

[34] P. G. Ciarlet and P. A. Raviart. General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods. *Arch. Rational Mech. Anal.*, **46** (1972), 178–199.

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

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

[37] C. Coatmelec. Approximation et interpolation des fonctions differentiables de plusieurs variables. *Ann. Sci. Ecole Norm. Sup.*, **83** (1966), 271–341.

[38] A. R. Conn, K. Scheinberg and Ph. L. Toint. On the convergence of derivative–free methods for unconstrained optimization. In M. D. Buhmann and A. Iserles, editors, *Approximation and Optimization. Tributes to M. J. D. Powell*, pages 83–109. Cambridge

University Press, 1997.

[39] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Undergraduate Texts in Mathematics. Springer–Verlag, 1992.

[40] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer Verlag, 1998.

[41] P. J. Davis. *Interpolation and approximation*. Dover Pub., 1975.

[42] N. Dyn and A. Ron. On multivariate polynomial interpolation. In J. C. Mason and M. E. Cox editors, *Algorithms for Approximation II*, pages 177–184. Chapman and Hall, 1990.

[43] D. Eisenbud. Commutative Algebra with a View Toward Algebraic Geometry. Graduate Texts in Mathematics, **150**, Springer 1994.

[44] M. Gasca. Multivariate polynomial interpolation. In W. Dahmen, M. Gasca, and C. A. Micchelli, editors, *Computation of Curves and Surfaces*, pages 215–236. Kluver Academic Publishers, 1990.

[45] M. Gasca and E. Lebrón. On Aitken–Neville formulae for multivariate interpolation. In E. L. Ortiz editor, *Numerical approximation of partial differential equations* , pages 133–140. North Holland, 1987.

[46] M. Gasca and A. López-Carmona. A general recurrence interpolation formula and its applications to multivariate interpolation. *J. Approx. Theory*, **34** (1982), 361–374.

[47] M. Gasca and J. I. Maeztu. On Lagrange and Hermite interpolation in $\mathbb{R}^k$. *Numer. Math.*, **39** (1982), 1–14.

[48] M. Gasca and J. J. Martínez. On the computation of multivariate confluent Vandermonde determinants and its applications . In R. Martin, editor, *The Mathematics of surfaces II*, pages 101–114, Oxford Univ. Press, 1987.

[49] M. Gasca and J. J. Martínez. On the solvability of bivariate Hermite–Birkhoff interpolation problems. *J, Comput. and Appl. Math.*, **32** (1990), 77–82.

[50] M. Gasca and G. Mühlbach. Multivariate interpolation: a survey with regard to extrapolation. In *Proceedings IMACS Transactions on scientific computing*, **vol. 1**, pages 255–278,Paris, (1988).

[51] M. Gasca and G. Mühlbach. Multivariate polynomial interpolation under projectivities II: Aitken-Neville formulas. *Numer. Algor.*, **2** (1992) 255–278.

[52] M. Gasca and V. Ramirez. Interpolation systems in $\mathbb{R}^k$. *J. Approx. Theory*, **42** (1984) 36–51.

[53] M. Gasca and T. Sauer. On bivariate Hermite interpolation with minimal degree polynomials. (1998). *SIAM J. Numer. Anal.*, to appear.

[54] H. V. Gevorgian, H. A. Hakopian and A. A. Sahakian. On the bivariate Hermite interpolation problem. *Constr. Approx.*, **11** (1995) 23–35.

[55] G. Glaeser. L'interpolation des fonctions differentiables de plusieurs variables. In *Proc. Liverpool Singularities Symposium II*, pages 1–29, Lect. Not. Math. **209**. Springer V. 1971.

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

[57] W. Gröbner. Algebraische Geometrie I, II. Bibliographisches Institut Mannheim, B.I–Hochschultaschenbücher, 1968 and 1970.

[58] P. Griffiths and J. Harris. Principles of Algebraic Geometry. John Wiley & Sons, New York, 1978.

[59] R. B. Guenter and E. L. Roetman. Some observations on interpolation in higher dimensions. *Math.Comp.*, **24** (1970) 517–521.

[60] A. W. Habib, R. N. Goldman and T. Lyche. A recursive algorithm for Hermite interpolation over a triangular grid. *J. Comp. Appl. Math.*, **73** (1996) 95–118.

[61] F. Hack. On bivariate Birkhoff interpolation. *J. Approx. Th.*, **49** (1987) 18–30.

[62] H. A. Hakopian. Multivariate divided differences and multivariate interpolation of Lagrange and Hermite type. *J. Approx. Th.*, **34** (1982) 286–305.

[63] K. Höllig and C. A. Micchelli. Divided differences, hyperbolic equations and lifting distributions. *Constr. Approx.*, **3** (1987) 143–156.

[64] E. Isaacson and H. B. Keller. *Analysis of Numerical Methods*. John Wiley & Sons, 1966.

[65] K. Jetter. Some contributions to bivariate interpolation and cubature. In C. K. Chui, L. L. Schumaker and J. D. Ward, editors, *Approximation Theory IV*, pages 533–538, Academic Press, 1983.

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

[67] D. Kincaid and W. Cheney. *Numerical Analysis*. Brooks/Cole Pub., 1991.

[68] L. Kronecker. Über einige Interpolationsformeln für ganze Funktionen mehrerer Variabeln. In H. Hensel, editor, *L. Kroneckers Werke*, volume I, pages 133–141. Chelsea Publishing Company, 1865. Lecture at the academy of sciences, December 21, 1865.

[69] K.S. Kunz. *Numerical Analysis*. McGraw-Hill Book Company, 1957.

[70] S. L. Lee and G. M. Phillips. Polynomial interpolation at points of a geometric mesh on a triangle. *Proc. Roy. Soc. Edinburgh*, **108A** (1988), 75–87.

[71] S. L. Lee and G. M. Phillips. Construction of lattices for Lagrange interpolation in Projective space. *Constructive Approximation*, **7** (1991), 283–297.

[72] A. Le Méhauté. A finite element approach to surface reconstruction. In W. Dahmen, M. Gasca, and C. A. Micchelli, editors, *Computation of Curves and Surfaces*, pages 237–274. Kluver Academic Publishers, 1990.

[73] A. Le Méhauté. On some aspects of multivariate polynomial interpolation. This volume.

[74] S. K. Lodha and R. N. Goldman. A unified approach to evaluation algorithms for multivariate polynomials. *Math. Comp.* **66** (1997) 1521–1553.

[75] J. Lorente and V. Ramírez. On interpolation systems and H-reducible interpolation problems. In C. K. Chui, L. L. Schumaker, and F. Utreras, editors, *Topics in multivariate approximation*, pages 153–163. Academic Press, 1987.

[76] G. G. Lorentz. *Approximation of functions*. Chelsea Publishing Company, 1966.

[77] G. G. Lorentz, K. Jetter and S. D. Riemenschneider. *Birkhoff interpolation*. Addison-Wesley,1983.

[78] R. A. Lorentz. *Multivariate Birkhoff Interpolation*. Number 1516 in Lecture Notes in Mathematics. Springer Verlag, 1992.

[79] F. S. Macaulay. *The Algebraic Theory of Modular Systems*. Number 19 in Cambridge Tracts in Math. and Math. Physics. Cambridge Univ. Press, 1916.

[80] J. I. Maeztu. Divided differences associated with reversible systems in $\mathbb{R}^2$. *SIAM J. Num. Anal.*, **19** (1982) 1032–1040.

[81] J. C. Mairhuber. On Haar's Theorem concerning Chebychev approximation problems having unique solutions. *Proceedings of the A.M.S.*, **7** (1956) 609–615.

[82] M. G. Marinari, H. M. Möller and T. Mora. Gröbner bases of ideals given by dual bases. In: Proceedings of ISAAC 1991.

[83] M. G. Marinari, H. M. Möller and T. Mora. Gröbner bases of ideals defined by functionals with an application to ideals of projective points. *J. AAECC*, **4** (1993) 103–145.

[84] M. G. Marinari, H. M. Möller and T. Mora. On multiplicities in polynomial system solving. *Trans. Amer. Math. Soc.*, **348** (1996) 3283–3321.

[85] C. A. Micchelli. On a numerically efficient method of computing multivariate B–splines. In W. Schempp and K. Zeller, editors, *Multivariate Approximation Theory*, pages 211–248. Birkhäuser, Basel, 1979.

[86] C. A. Micchelli. A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B–splines and Lagrange interpolation. *Rocky Mountain J. Math.*, **10** (1979), 485–497.

[87] C. A. Micchelli and P. Milman. A formula for Kergin interpolation in $\mathbb{R}^k$. *J. Approx. Th.*, **29** (1980) 294–296.

[88] C. A. Micchelli. Algebraic aspects of interpolation. In C. de Boor, editor,*Approximation Theory*, Proc. Symp. Appl. Math, **36**, pages 81–102, Amer. Math. Soc., 1986.

[89] C.A. Micchelli. *Mathematical Aspects of Geometric Modeling*, volume 65 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1995.

[90] C. A. Micchelli, T. Sauer, and Yuesheng Xu. A construction of refinable sets for interpolating wavelets. *Results in Mathematics*, **34** (1998), 359–372.

[91] H. M. Möller. Hermite interpolation in several variables using ideal-theoretic methods. In W. Schempp and K. Zeller editors, *Constructive Theory of Functions of Several Variables*, Lect. Not. Math.**571**, pages 155–163, 1977.

[92] H. M. Möller. Gröbner bases and Numerical Analysis. In B. Buchberger and F. Winkler, editors, *Groebner Bases and Applications (Proc. of the Conf. 33 Years of Groebner Bases)*, volume 251 of *London Math. Soc. Lecture Notes*, pages 159–179. Cambridge University Press, 1998.

[93] H. M. Möller and T. Sauer. H–bases for polynomial interpolation and system solving. This volume.

[94] G. Mühlbach. On multivariate interpolation by generalized polynomials on subsets of grids. *Computing*, **40** (1988).

[95] G. Mühlbach. The general Neville-Aitken algorithm and some applications. *Numer. Math.* **31** (1978), 97-110.

[96] G. Mühlbach and M. Gasca. Multivariate interpolation under Projectivities I. Lagrange and Newton interpolation formulas. *Numer. Algor.*, **1** (1991), 375–400.

[97] E. H. Neville. Iterative interpolation. *J. Indian Math. Soc.* **20** (1934), 87–120.

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

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

[100] J. M. Peña and T. Sauer. On the multivariate Horner scheme. *SIAM J. Numer. Anal.*, (1999), to appear.

[101] L. Robbiano. On the theory of graded structures. *J. Symbolic Computation*, **2** (1986), 139–170.

[102] H. E. Salzer. Divided differences for functions of two variables for irregularly spaced arguments. *Numer. Math.*, **6** (1964), 68–77.

[103] T. Sauer. Computational aspects of multivariate polynomial interpolation. *Advances Comput. Math.*, **3** (1995), 219–238.

[104] T. Sauer. Polynomial interpolation of minimal degree. *Numer. Math.*, **78** (1997), 59–85.

[105] T. Sauer. Polynomial interpolation of minimal degree and Gröbner bases. In B. Buchberger and F. Winkler, editors, *Groebner Bases and Applications (Proc. of the Conf. 33 Years of Groebner Bases)*, volume 251 of *London Math. Soc. Lecture Notes*, pages 483–494. Cambridge University Press, 1998.

[106] T. Sauer. Gröbner bases, H–bases and interpolation. *Trans. Amer. Math. Soc.*, to appear.

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

[108] T. Sauer and Yuan Xu. On multivariate Hermite interpolation. *Advances Comput. Math.*, **4** (1995), 207–259.

[109] T. Sauer and Yuan Xu. Regular points for Lagrange interpolation on the unit disk. *Numer. Algo.*, **12** (1996), 287-296.

[110] L. L. Schumaker. Fitting surfaces to scattered data. In G. G. Lorentz. C. K Chui and L. L. Schumaker, editors, *Approximation Theory 2.* Academic Press, 1976.

[111] I. F. Steffensen. *Interpolation.* Chelsea, New York, 1927.

[112] H. C. Thacher Jr. and W. E. Milne. Interpolation in several variables. *J. SIAM*, **8** (1960), 33–42.

[113] H. C. Thacher Jr. Derivation of interpolation formulas in several independent variables. *Annals New York Acad. Sc.*, **86** (1960), 758–775.

[114] R. S. Walker. *Algebraic curves.* Springer Verlag 1978.

[115] H. Werner. Remarks on Newton type multivariate interpolation for subsets of grids. *Computing*, **25** (1980), 181–191.