ELSEVIER

Journal of Computational and Applied Mathematics 122 (2000) 23–35

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

On the history of multivariate polynomial interpolation

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

[a] *Department of Applied Mathematics, University of Zaragoza, 50009 Zaragoza, Spain*
[b] *Institute of Mathematics, University Erlangen-Nürnberg, Bismarckstr. $1\frac{1}{2}$, D-91054 Erlangen, Germany*

Received 7 June 1999; received in revised form 8 October 1999

Abstract

Multivariate polynomial interpolation is a basic and fundamental subject in Approximation Theory and Numerical Analysis, which has received and continues receiving not deep but constant attention. In this short survey, we review its development in the first 75 years of this century, including a pioneering paper by Kronecker in the 19th century.

1. Introduction

Interpolation, by polynomials or other functions, is a rather old method in applied mathematics. This is already indicated by the fact that, apparently, the word "interpolation" itself has been introduced by J. Wallis as early as 1655 as it is claimed in [13]. Compared to this, polynomial interpolation in *several variables* is a relatively new topic and probably only started in the second-half of the last century with work in [6,22]. If one considers, for example, the *Encyklopädie der Mathematischen Wissenschaften* [13] (Encyclopedia of Math. Sciences), originated by the *Preußische Akademie der Wissenschaften* (Prussian Academy of Sciences) to sum up the "state of art" of mathematics at its time, then the part on interpolation, written by J. Bauschinger (Bd. I, Teil 2), mentions only one type of multivariate interpolation, namely (tensor) products of sine and cosine functions in two variables, however, without being very specific. The French counterpart, the *Encyclopédie de Sciences Mathematiques* [14], also contains a section on interpolation (Tome I, vol. 4), where Andoyer translated and extended Bauschinger's exposition. Andoyer is even more

* Corressponding author.
E-mail addresses: gasca@posta.unizar.es (M. Gasca), sauer@mi.uni-erlangen.de (T. Sauer).

0377-0427/00/$ - see front matter
PII: S0377-0427(00)00353-8

explicit with his opinion on multivariate polynomial interpolation, by making the following statement which we think that time has contradicted:

> Il est manifeste que l'interpolation des fonctions de plusiers variables ne demande aucun principe nouveau, car dans tout ce qui précède le fait que la variable indépendante était unique n'a souvent joué aucun rôle. [1]

Nevertheless, despite of Andoyer's negative assessment, multivariate polynomial interpolation has received not deep but constant attention from one part of the mathematical community and is today a basic subject in Approximation Theory and Numerical Analysis with applications to many mathematical problems. Of course, this field has definitely been influenced by the availability of computational facilities, and this is one of the reasons that more papers have been published about this subject in the last 25 years than in the preceding 75 ones.

To our knowledge, there is not any paper before the present one surveying the early papers and books on multivariate polynomial interpolation. Our aim is a first, modest attempt to cover this gap. We do not claim to be exhaustive and, in particular, recognize our limitations with respect to the Russian literature. Moreover, it has to be mentioned that the early results on multivariate interpolation usually appear in the context of many different subjects. For example, papers on cubature formulas frequently have some part devoted to it. Another connection is Algebraic Geometry, since the solvability of a multivariate interpolation problem relies on the fact that the interpolation points do not lie on an algebraic surface of a certain type. So it is difficult to verify precisely if and when a result appeared somewhere for the first time or if it had already appeared, probably even in an implicit way, in a different context. We remark that another paper in this volume [25] deals, complementarily, with recent results in the subject, see also [16].

Along the present paper we denote by Π_k^d the space of d-variate polynomials of total degree not greater than k.

2. Kronecker, Jacobi and multivariate interpolation

Bivariate interpolation by the tensor product of univariate interpolation functions, that is when the variables are treated separately, is the classical approach to multivariate interpolation. However, when the set of interpolation points is not a Cartesian product grid, it is impossible to use that idea. Today, given any set of interpolation points, there exist many methods [2] to construct an adequate polynomial space which guarantees unisolvence of the interpolation problem. Surprisingly, this idea of constructing an appropriate interpolation space was already pursued by Kronecker [22] in a widely unknown paper from 1865, which seems to be the first treatment of multivariate polynomial interpolation with respect to fairly arbitrary point configurations. Besides the mathematical elegance of this approach, we think it is worthwhile to devote some detailed attention to this paper and to resolve its main ideas in today's terminology, in particular, as it uses the "modern" approach of connecting polynomial interpolation to the theory of polynomial ideals.

---

[1] It is clear that the interpolation of functions of several variables does not demand any new principles because in the above exposition the fact that the variable was unique has not played frequently any role.

[2] See [16,25] for exposition and references.

Kronecker's method to construct an interpolating polynomial assumes that the disjoint nodes $z_1,\dots,z_N \in \mathbb{C}^d$ are given in *implicit* form, i.e., they are (all) the common *simple* zeros of d polynomials $f_1,\dots,f_d \in \mathbb{C}[z] = \mathbb{C}[\zeta_1,\dots,\zeta_d]$. Note that the nonlinear system of equations

$$f_j(\zeta_1,\dots,\zeta_d) = 0, \quad j = 1,\dots,d, \tag{1}$$

is a *square* one, that is, the number of equations and the number of variables coincide. We are interested in the *finite* variety V of solutions of (1) which is given as

$$V := \{z_1,\dots,z_N\} = \{z \in \mathbb{C}^d : f_1(z) = \cdots = f_d(z) = 0\}. \tag{2}$$

The *primary decomposition* according to the variety V allows us to write the ideal $\mathscr{I}(V) = \{p : p(z) = 0,\ z \in V\}$ as

$$\mathscr{I}(V) = \bigcap_{k=1}^{N} \langle \zeta_1 - \zeta_{k,1},\dots,\zeta_d - \zeta_{k,d}\rangle,$$

where $z_k = (\zeta_{k,1},\dots,\zeta_{k,d})$. In other words, since $f_k \in \mathscr{I}(V)$, $k = 1,\dots,d$, any of the polynomials $f_1,\dots,f_d$ can be written, for $k = 1,\dots,N$, as

$$f_j = \sum_{i=1}^{d} g_{i,j}^k(\cdot)(\zeta_i - \zeta_{k,i}), \tag{3}$$

where $g_{i,j}^k$ are appropriate polynomials. Now consider the $d \times d$ square matrices of polynomials

$$G_k = [g_{i,j}^k : i,j = 1,\dots,d], \quad k = 1,\dots,N$$

and note that, due to (3), and the assumption that $f_j(z_k) = 0$, $j = 1,\dots,d$, $k = 1,\dots,N$, we have

$$0 = \begin{bmatrix} f_1(Z_j) \\ \vdots \\ f_d(z_j) \end{bmatrix} = G_k(z_j) \begin{bmatrix} (\zeta_{j,1} - \zeta_{k,1}) \\ \vdots \\ (\zeta_{j,d} - \zeta_{k,d}) \end{bmatrix}, \quad k = 1,\dots,N. \tag{4}$$

Since the interpolation nodes are assumed to be *disjoint*, this means that for all $j \neq k$ the matrix $G_k(z_j)$ is *singular*, hence the determinant of $G_k(z_j)$ has to be zero. Moreover, the assumption that $z_1,\dots,z_N$ are *simple* zeros guarantees that $\det G_k(z_k) \neq 0$. Then, Kronecker's interpolant takes, for any $f : \mathbb{C}^d \to \mathbb{C}$, the form

$$Kf = \sum_{j=1}^{N} f(z_j) \frac{\det G_k(\cdot)}{\det G_k(z_k)}. \tag{5}$$

Hence,

$$\mathscr{P} = \operatorname{span}\left\{ \frac{\det G_k(\cdot)}{\det G_k(z_k)} : k = 1,\dots,N \right\}$$

is an interpolation space for the interpolation nodes $z_1,\dots,z_N$. Note that this method does not give only *one* interpolation polynomial but in general *several different* interpolation spaces, depending on how the representation in (3) is chosen. In any way, note that for each polynomial $f \in \mathbb{C}[z]$ the difference

$$f - \sum_{j=1}^{N} f(z_j) \frac{\det G_k(z)}{\det G_k(z_k)}$$

belongs to the ideal $\langle f_1,\dots,f_d\rangle$, hence there exist polynomials $q_1,\dots,q_d$ such that

$$f-\sum_{j=1}^{N} f(z_j)\frac{\det G_k(z)}{\det G_k(z_k)}=\sum_{j=1}^{d} q_j f_j. \tag{6}$$

Moreover, as Kronecker points out, the "magic" polynomials $g_{i,j}^k$ can be chosen such that their leading homogeneous terms, say $G_{i,j}^k$, coincide with the leading homogeneous terms of $(1/\deg f_j)\partial f_j/\partial\zeta_i$. If we denote by F_j the leading homogeneous term of f_j, $j=1,\dots,d$, then this means that

$$G_{i,j}^k=\frac{1}{\deg F_j}\frac{\partial F_j}{\partial\zeta_i},\quad i,j=1,\dots,d,\quad k=1,\dots,N. \tag{7}$$

But this implies that the homogeneous leading term of the "fundamental" polynomials $\det G_k$ coincides, after this particular choice of $g_{i,j}^k$, with

$$g=\frac{1}{\deg f_1\cdots\deg f_d}\det\left[\frac{\partial F_j}{\partial\zeta_i}: i,j=1,\dots,d\right],$$

which is *independent of* k now; in other words, there exist polynomials $\hat g_k$, $k=1,\dots,N$, such that $\deg\hat g_k<\deg g$ and $\det G_k=g+\hat g_k$. Moreover, g is a homogeneous polynomial of degree at most $\deg f_1+\cdots+\deg f_d-d$. Now, let p be any polynomial, then

$$Kp=\sum_{j=1}^{N} p(z_j)\frac{\det G_j(\cdot)}{\det G_j(z_j)}=g\sum_{j=1}^{N}\frac{p(z_j)}{\det G_j(z_j)}+\sum_{j=1}^{N}\frac{p(z_j)}{\det G_j(z_j)}\hat g_j. \tag{8}$$

Combining (8) with (6) then yields the existence of polynomials $q_1,\dots,q_d$ such that

$$p=g\sum_{j=1}^{N}\frac{p(z_j)}{\det G_j(z_j)}+\sum_{j=1}^{N}\frac{p(z_j)}{\det G_j(z_j)}\hat g_j+\sum_{j=1}^{d} q_j f_j$$

and comparing homogeneous terms of degree $\deg g$ Kronecker realized that either, for any p such that $\deg p<\deg g$,

$$\sum_{j=1}^{N}\frac{p(z_j)}{\det G_j(z_j)}=0 \tag{9}$$

or there exist homogeneous polynomials $h_1,\dots,h_d$ such that

$$g=\sum_{j=1}^{d} h_j\det F_j. \tag{10}$$

The latter case, Eq. (10), says (in algebraic terminology) that there is a *syzygy* among the leading terms of the polynomials F_j, $j=1,\dots,d$, and is equivalent to the fact that $N<\deg f_1\cdots\deg f_d$, while (9) describes and even characterizes the *complete intersection case* that $N=\deg f_1\cdots\deg f_d$. In his paper, Kronecker also mentions that the condition (10) has been overlooked in [21]. Jacobi dealt there with the common zeros of two bivariate polynomials and derived *explicit* representations for the functional

$$[z_1,\dots,z_N]f:=\sum_{j=1}^{N}\frac{f(z_j)}{\det G_j(z_j)}, \tag{11}$$

which behaves very much like a divided difference, since it is a combination of point evaluations which, provided that (9) hold true, annihilates $\Pi^d_{\deg g-1}$.

In addition, Kronecker refers to a paper [6] which he says treats the case of symmetric functions, probably elementary symmetric polynomials. Unfortunately, this paper is unavailable to us so far.

3. Bivariate tables, the natural approach

Only very few research papers on multivariate polynomial interpolation were published during the first part of this century. In the classical book *Interpolation* [45], where one section (Section 19) is devoted to this topic, the author only refers to two related papers, recent at that time (1927), namely [27,28]. The latter one [28], turned out to be inaccessible to us, unfortunately, but it is not difficult to guess that it might have pursued a tensor product approach, because this is the unique point of view of [45] (see also [31]).

The formulas given in [27] are Newton formulas for tensor product interpolation in two variables, and the author, Narumi, claims (correctly) that they can be extended to "many variables". Since it is a tensor product approach, the interpolation points are of the form (x_i, y_j), $0 \leqslant i \leqslant m$, $0 \leqslant j \leqslant n$, with x_i, y_j arbitrarily distributed on the axes OX and OY, respectively. Bivariate divided differences for these sets of points are obtained in [27], by recurrence, separately for each variable. With the usual notations, the interpolation formula from [27] reads as

$$p(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{n} f[x_0,\dots,x_i; y_0,\dots,y_j] \prod_{h=0}^{i-1}(x-x_h) \prod_{k=0}^{j-1}(y-x_k), \tag{12}$$

where empty products have the value 1. Remainder formulas based on the mean value theorem are also derived recursively from the corresponding univariate error formulas in [27]. For f sufficiently smooth there exist values ξ, ξ', η, η' such that

$$R(x,y) = \frac{\partial^{m+1} f(\xi,y)}{\partial x^{m+1}} \frac{\prod_{h=0}^{m}(x-x_h)}{(m+1)!} + \frac{\partial^{n+1} f(x,\eta)}{\partial y^{n+1}} \frac{\prod_{k=0}^{n}(y-y_k)}{(n+1)!}$$

$$- \frac{\partial^{m+n+2} f(\xi',\eta')}{\partial x^{m+1} \partial y^{n+1}} \frac{\prod_{h=0}^{m}(x-x_h)}{(m+1)!} \frac{\prod_{k=0}^{n}(y-y_k)}{(n+1)!}. \tag{13}$$

The special case of equidistant points on both axes is particularly considered in [27], and since the most popular formulas at that time were based on finite differences with equally spaced arguments, Narumi shows how to extend Gauss, Bessel and Stirling univariate interpolation formulas for equidistant points to the bivariate case by tensor product. He also applies the formulas he obtained to approximate the values of bivariate functions, but he also mentions that some of his formulas had been already used in [49].

In [45], the Newton formula (12) is obtained in the same way, with the corresponding remainder formula (13). Moreover, Steffensen considers a more general case, namely when for each i, $0 \leqslant i \leqslant m$, the interpolation points are of the form $y_0,\dots,y_{n_i}$, with $0 \leqslant n_i \leqslant n$. Now with a similar argument the interpolating polynomial becomes

$$p(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{n_i} f[x_0,\dots,x_i; y_0,\dots,y_j] \prod_{h=0}^{i-1}(x-x_h) \prod_{k=0}^{j-1}(y-x_k) \tag{14}$$

with a slightly more complicated remainder formula. The most interesting particular cases occur when $n_i = n$, which is the Cartesian product considered above, and when $n_i = m - i$. This *triangular* case (triangular not because of the geometrical distribution of the interpolation points, but of the indices (i,j)), gives rise to the interpolating polynomial

$$p(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{m-i} f[x_0,\dots,x_i; y_0,\dots,y_j] \prod_{h=0}^{i-1}(x - x_h) \prod_{k=0}^{j-1}(y - x_k), \tag{15}$$

that is

$$p(x,y) = \sum_{0 \leqslant i+j \leqslant m} f[x_0,\dots,x_i; y_0,\dots,y_j] \prod_{h=0}^{i-1}(x - x_h) \prod_{k=0}^{j-1}(y - x_k). \tag{16}$$

Steffensen refers for this formula to Biermann's lecture notes [4] from 1905, and actually it seems that Biermann has been the first who considered polynomial interpolation on the triangular grid in a paper [3] from 1903 (cf. [44]) in the context of cubature.

Since the triangular case corresponds to looking at the "lower triangle" of the tensor product situation only, this case can be resolved by tensor product methods. In particular, the respective error formula can be written as

$$R(x,y) = \sum_{i=0}^{m+1} \frac{\partial^{m+1} f(\xi_i,\eta_i)}{\partial x^i \partial y^{m+1-i}} \frac{\prod_{h=0}^{i-1}(x - x_h)}{i!} \frac{\prod_{k=0}^{m-i}(y - y_k)}{(m-i+1)!}. \tag{17}$$

In the case of Cartesian product Steffensen also provides the Lagrange formula for (12), which can be obviously obtained by tensor product of univariate formulas.

Remainder formulas based on intermediate points (ξ_i,η_i) can be written in many different forms. For them we refer to Stancu's paper [44] which also contains a brief historical introduction where the author refers, among others, to [3,15,27,40,41]. Multivariate remainder formulas with Peano (spline) kernel representation, however, have not been derived until very recently in [42] and, in particular, in [43] which treats the triangular situation.

4. Salzer's papers: from bivariate tables to general sets

In 1944, Salzer [33] considered the interpolation problem at points of the form $(x_1 + s_1h_1,\dots,x_n + s_nh_n)$ where

(i) $(x_1,\dots,x_n)$ is a given point in $\mathbb{R}^n$,
(ii) $h_1,\dots,h_n$ are given real numbers,
(iii) $s_1,\dots,s_n$ are nonnegative integers summing up to m.

This is the multivariate extension of the triangular case (16) for equally spaced arguments, where finite differences can be used. Often, different names are used for the classical Newton interpolation formula in the case of equally spaced arguments using forward differences: Newton–Gregory, Harriot–Briggs, also known by Mercator and Leibnitz, etc. See [18] for a nice discussion of this issue. In [33], Salzer takes the natural multivariate extension of this formula considering the polynomial $q(t_1,\dots,t_n) := p(x_1 + t_1h_1,\dots,x_n + t_nh_n)$ of total degree not greater than m in the variables $t_1,\dots,t_n$, which interpolates a function $f(x_1+t_1h_1,\dots,x_n+t_nh_n)$ at the points corresponding to $t_i=s_i$, $i=1,\dots,n$,

where the s_i are all nonnegative integers such that $0 \leqslant s_1 + \cdots + s_n \leqslant m$. The formula, which is called in [33] a *multiple Gregory–Newton formula*, is rewritten there in terms of the values of the function f at the interpolation points, i.e., in the form

$$q(t_1,\ldots,t_n) = \sum_{s_1+\cdots+s_n\leqslant m} \binom{t_1}{s_1}\cdots\binom{t_n}{s_n}\binom{m-t_1-\cdots-t_n}{m-s_1-\cdots-s_n} f(x_1+s_1h_1,\ldots,x_n+s_nh_n). \quad (18)$$

Note that (18) is the Lagrange formula for this interpolation problem. Indeed, each function

$$\binom{t_1}{s_1}\cdots\binom{t_n}{s_n}\binom{m-t_1-\cdots-t_n}{m-s_1-\cdots-s_n} \quad (19)$$

is a polynomial in $t_1,\ldots,t_n$ of total degree m which vanishes at all points $(t_1,\ldots,t_n)$ with t_i nonnegative integers $0 \leqslant t_1 + \cdots + t_n \leqslant m$, except at the point $(s_1,\ldots,s_n)$, where it takes the value 1. In particular, for $n = 1$ we get the well-known univariate Lagrange polynomials

$$\ell_s(t) = \binom{t}{s}\binom{m-t}{m-s} = \prod_{\substack{0\leqslant i\leqslant m,\\ i\neq s}} \frac{t-i}{s-i}$$

for $s = 0,\ldots,m$.

Salzer used these results in [34] to compute tables for the polynomials (18) and, some years later in [35], he studied in a similar form how to get the Lagrange formula for the more general case of formula (16), even starting with this formula. He obtained the multivariate Lagrange polynomials by a rather complicated expression involving the univariate ones.

It should be noted that several books related to computations and numerical methods published around this time include parts on multivariate interpolation to some extent, surprisingly, more than most of the recent textbooks in Numerical Analysis. We have already mentioned Steffensen's book [45], but we should also mention Whittaker and Robinson [51, pp. 371–374], Mikeladze [26, Chapter XVII] and especially Kunz [23, pp. 248–274], but also Isaacson and Keller [20, pp. 294–299] and Berezin and Zhidkov [2, pp. 156–194], although in any of them not really much more than in [45] is told.

In [36,37], Salzer introduced a concept of bivariate divided differences abandoning the idea of iteration for each variable x and y taken separately. Apparently, this was the first time (in spite of the similarity with (11)), that bivariate divided differences were explicitly defined for irregularly distributed sets of points. Divided differences with repeated arguments are also considered in [37] by coalescence of the ones with different arguments. Since [36] was just a first attempt of [37], we only explain the latter one. Salzer considers the set of monomials $\{x^i y^j\}$, with i,j nonnegative integers, ordered in a graded lexical term order, that is,

$$(i,j) < (h,k) \Leftrightarrow i + j < h + k \quad \text{or} \quad i + j = h + k,\ i > h. \quad (20)$$

Hence, the monomials are listed as

$$\{1, x, y, x^2, xy, y^2, x^3, \ldots\}. \quad (21)$$

For any set of $n+1$ points (x_i, y_i), Salzer defines the associated divided difference

$$[01\ldots n]f := \sum_{k=0}^{n} A_k f(x_k, y_k), \quad (22)$$

choosing the coefficients A_k in such a form that (22) vanishes when f is any of the first n monomials of list (21) and takes the value 1 when f is the $(n+1)$st monomial of that list. In other words, the coefficients A_k are the solution of the linear system

$$\sum_{k=0}^{n} A_k x_k^i y_k^j = 0, \quad x^i y^j \text{ any of the first } n \text{ monomials of (21)},$$
$$\sum_{k=0}^{n} A_k x_k^i y_k^j = 1, \quad x^i y^j \text{ the } (n+1)\text{th monomial of (21)}. \tag{23}$$

These generalized divided differences share some of the properties of the univariate ones but not all. Moreover, they have some limitations, for example, they exist only if the determinant of the coefficients in (23) is different from zero, and one has no control of that property in advance. On the other hand, observe that for example the simple divided difference with two arguments (x_0, y_0) and (x, y), which is

$$\frac{f(x,y) - f(x_0,y_0)}{x - x_0},$$

gives, when applied to $f(x,y) = xy$, the rational function

$$\frac{xy - x_0 y_0}{x - x_0}$$

and not a polynomial of lower degree. In fact, Salzer's divided differences did not have great success. Several other definitions of multivariate divided differences had appeared since then, trying to keep as many as possible of the good properties of univariate divided differences, cf. [16].

5. Reduction of a problem to other simpler ones

Around the 1950s an important change of paradigm happened in multivariate polynomial interpolation, as several people began to investigate more general distributions of points, and not only (special) subsets of Cartesian products. So, when studying cubature formulae [32], Radon observed the following in 1948: if a bivariate interpolation problem with respect to a set $T \subset \mathbb{R}^2$ of $\binom{k+2}{2}$ interpolation points is unisolvent in Π_k^2, and U is a set of $k+2$ points on an arbitrary straight line $\ell \subset \mathbb{R}^2$ such that $\ell \cap T = \emptyset$, then the interpolation problem with respect to $T \cup U$ is unisolvent in Π_{k+1}^2. Radon made use of this observation to build up point sets which give rise to unisolvent interpolation problems for Π_m recursively by degree. Clearly, these interpolation points immediately yield interpolatory cubature formulae.

The well-known Bézout theorem, cf. [50], states that two planar algebraic curves of degree m and n, with no common component, intersect each other at exactly mn points in an algebraic closure of the underlying field, counting multiplicities. This theorem has many interesting consequences for bivariate interpolation problems, extensible to higher dimensions. For example, no unisolvent interpolation problem in Π_n^2 can have more than $n+1$ collinear points. Radon's method in [32] is a consequence of this type of observations, and some other more recent results of different authors can also be deduced in a similar form, as we shall see later.

Another example of a result which shows the more general point of view taken in multivariate interpolation at that time is due to Thacher Jr. and Milne [47] (see also [48]). Consider two univariate interpolation problems in Π^1_{n-1}, with $T_1,\ T_2$ as respective sets of interpolation points, both of cardinality n. Assume that $T_1 \cap T_2$ has cardinality $n-1$, hence $T = T_1 \cup T_2$ has cardinality $n+1$. The univariate Aitken–Neville interpolation formula combines the solutions of the two smaller problems based on T_1 and T_2 to obtain the solution in Π^1_n of the interpolation problem with T as the underlying set of interpolation points. The main idea is to find a *partition of unity*, in this case affine polynomials ℓ_1, ℓ_2, i.e., $\ell_1 + \ell_2 = 1$, such that

$$\ell_1(T_2 \setminus T_1) = \ell_2(T_1 \setminus T_2) = 0$$

and then combine the solutions p_1, p_2 with respect to $T_1,\ T_2$, into the solution $\ell_1\, p_1 + \ell_2 p_2$ with respect to T. This method was developed in the 1930s independently by Aitken and Neville with the goal to avoid the explicit use of divided differences in the computation of univariate Lagrange polynomial interpolants.

It was exactly this idea which Thatcher and Milne extended to the multivariate case in [47]. Let us sketch their approach in the bivariate case. For example, consider an interpolation problem with 10 interpolation points, namely, the set $T = \{(i,j) : 0 \leqslant i + j \leqslant 3\}$, where i, j are nonnegative integers, and the interpolation space Π^2_3. The solution p_T of this problem is obtained in [47] from the solutions $p_{T_k} \in \Pi^2_2,\ k = 1,2,3$, of the 3 interpolation problems with respect to the six-point sets $T_k \subset T,\ k = 1,2,3,$ where

$$T_1 = \{(i,j) : 0 \leqslant i + j \leqslant 2\},$$

$$T_2 = \{(i,j) : (i,j) \in T,\ i > 0\},$$

$$T_3 = \{(i,j) : (i,j) \in T,\ j > 0\}.$$

Then,

$$p_T = \ell_1\, p_{T_1} + \ell_2\, p_{T_2} + \ell_3\, p_{T_3},$$

where $\ell_k,\ k = 1,2,3$ are appropriate polynomials of degree 1. In fact, in this case these polynomials are the barycentric coordinates relative to the simplex (0, 0), (3, 0), (0, 3) and thus a partition of unity. In [47] the problem is studied in d variables and in that case $d+1$ "small" problems, with respective interpolation sets $T_k,\ k = 1,\ldots,d$, with a simplicial structure (the analogue of the triangular grid), are used to obtain the solution of the full problem with $T = T_1 \cap \cdots \cap T_{d+1}$ as interpolation points.

In 1970, Guenter and Roetman [19], among other observations, made a very interesting remark, which connects to the Radon/Bézout context and deserves to be explained here. Let us consider a set T of $\binom{m+d}{d}$ points in $\mathbb{R}^d$, where exactly $\binom{m+d-1}{d-1}$ of these points lie on a hyperplane H. Then $T \setminus H$ consists of $\binom{m-1+d}{d}$ points. Let us denote by $\Pi^m_{d,H}$ the space of polynomials of Π^m_d with the variables restricted to H, which is isomorphic to Π^m_{d-1}. If the interpolation problems defined by the sets $T \setminus H$ and $T \cap H$ are unisolvent in the spaces Π^{m-1}_d and $\Pi^m_{d,H}$, respectively, then the interpolation problem defined by T is unisolvent in Π^m_d. In other words, the idea is to decompose, whenever possible, a problem of degree m and d variables into two simpler problems, one of degree m and $d-1$ variables and the other one with degree $m-1$ and d variables.

6. The finite element approach

In 1943, Courant [11] suggested a finite difference method applicable to boundary value problems arising from variational problems. It is considered one of the motivations of the finite element method, which emerged from the engineering literature along the 1950s. It is a variational method of approximation which makes use of the Rayleigh–Ritz–Galerkin technique. The method became very successful, with hundreds of technical papers published (see, e.g., the monograph [52]), even before its mathematical basis was completely understood at the end of the 1960s.

Involved in the process of the finite element method there are local polynomial interpolation problems, generally for polynomials of low degree, thus, with only few interpolation data. The global solution obtained by solving all the local interpolation problems is a piecewise polynomial of a certain regularity, depending on the amount and type of interpolation data in the common boundary between pieces. Some of the interest in multivariate polynomial interpolation along the 1960/1970s was due to this method. Among the most interesting mathematical papers of that time in Finite Elements, we can mention [53,5], see also the book [46] by Strang and Fix, but, in our opinion, the most relevant papers and book from the point of view of multivariate polynomial interpolation are due to Ciarlet et al., for example [7–9].

In 1972, Nicolaides [29,30] put the classical problem of interpolation on a simplicial grid of $\binom{m+d}{d}$ points of $\mathbb{R}^d$, regularly distributed, forming what he called a *principal lattice*, into the finite element context. He actually used barycentric coordinates for the Lagrange formula, and moreover gave the corresponding error representations, see also [7]. However, much of this material can already be found in [3]. In general, taking into account that these results appeared under different titles, in a different context and in journals not accessible everywhere, it is not so surprising any more, how often the basic facts on the interpolation problem with respect to the simplicial grid had been rediscovered.

7. Hermite problems

The use of partial or directional derivatives as interpolation data in the multivariate case had not received much attention prior to the finite element method, where they were frequently used. It seems natural to approach partial derivatives by coalescence, as in univariate Hermite interpolation problems. However, things are unfortunately much more complicated in several variables. As it was already pointed out by Salzer and Kimbro [39] in 1958, the Hermite interpolation problem based on the values of a bivariate function $f(x, y)$ at two distinct points $(x_1, y_1), (x_2, y_2)$ and on the values of the partial derivatives $\partial f/\partial x$, $\partial f/\partial x$ at each of these two points is not solvable in the space Π_2^2 for any choice of points, although the number of interpolation conditions coincides with the dimension of the desired interpolation space. Some years later, Ahlin [1] circumvented some of these problems by using a tensor product approach: k^2 derivatives $\partial^{p+q} f/\partial x^p \partial y^q$ with $0 \leqslant p, q \leqslant k-1$ are prescribed at the n^2 points of a Cartesian product. The interpolation space is the one spanned by $x^\alpha y^\beta$ with $0 \leqslant \alpha, \beta \leqslant nk-1$ and a formula for the solution is easily obtained.

We must mention that Salzer came back to bivariate interpolation problems with derivatives in [38] studying *hyperosculatory interpolation* over Cartesian grids, that is, interpolation problems where all partial derivatives of first and second order and the value of the function are known at the

interpolation points. Salzer gave some special configurations of points which yield solvability of this type of interpolation problem in an appropriate polynomial space and also provided the corresponding remainder formulae.

Nowadays, Hermite and Hermite–Birkhoff interpolation problems have been studied much more systematically, see [16,25] for references.

8. Other approaches

In 1966, Coatmelec [10] studied the approximation of functions of several variables by linear operators, including interpolation operators. At the beginning of the paper, he only considered interpolation operators based on values of point evaluations of the function, but later he also used values of derivatives. In this framework he obtained some qualitative and quantitative results on the approximation order of polynomial interpolation. At the end of [10], Coatmelec also includes some examples in $\mathbb{R}^2$ of points which are distributed irregularly along lines: $n+1$ of the points on a line r_0, n of them on another line r_1, but not on r_0, and so on until 1 point is chosen on a line r_n but not on $r_0 \cup \cdots \cup r_{n-1}$. He then points out the unisolvence of the corresponding interpolation problem in Π_n^2 which is, in fact, again a consequence of Bézout's theorem as in [32].

In 1971, Glaeser [17] considers Lagrange interpolation in several variables from an abstract algebraic/analytic point of view and acknowledges the inconvenience of working with particular systems of interpolation points due to the possibility of the nonexistence of a solution, in contrast to the univariate case. This is due to the nonexistence of polynomial spaces of dimension $k > 1$ in more than one variable such that the Lagrange interpolation problem has a unique solution for any system of k interpolation points. In other words, there are no nontrivial Haar (or Chebychev) spaces any more for two and more variables, cf. [12] or [24]. In [17], polynomial spaces with dimension greater than the number of interpolation conditions are considered in order to overcome this problem. Glaeser investigated these *underdetermined* systems which he introduced as *interpolation schemes* in [17] and also studied the problem of how to particularize the affine space of all solutions of a given interpolation problem in order to obtain a unique solution. This selection process is done in such a way that it controls the variation of the solution when two systems of interpolation points are very "close" to each other, with the goal to obtain a continuous selection process.

Acknowledgements

1. We thank Carl de Boor for several references, in particular, for pointing out to us the paper [32] with the result mentioned at the beginning of Section 4, related to Bezout's theorem. We are also grateful the help of Elena Ausejo, from the group of History of Sciences of the University of Zaragoza, in the search of references.
2. M. Gasca has been partially supported by the Spanish Research Grant PB96 0730.
3. T. Sauer was supported by a DFG Heisenberg fellowship, Grant SA 627/6.

References

[1] A.C. Ahlin, A bivariate generalization of Hermite's interpolation formula, Math. Comput. 18 (1964) 264–273.
[2] I.S. Berezin, N.P. Zhidkov, Computing Methods, Addison-Wesley, Reading, MA, 1965 (Russian version in 1959).
[3] O. Biermann, Über näherungsweise Kubaturen, Monatshefte Math. Phys. 14 (1903) 211–225.
[4] O. Biermann, Vorlesungen über Mathematische Näherungsmethoden, Vieweg, Braunschweig, 1905.
[5] G. Birkhoff, M.H. Schultz, R.S. Varga, Piecewise Hermite interpolation in one and two variables with applications to partial differential equations, Numer. Math. 11 (1968) 232–256.
[6] W. Borchardt, Über eine Interpolationsformel für eine Art symmetrischer Funktionen und deren Anwendung, Abh. d. Preuß. Akad. d. Wiss. (1860) 1–20.
[7] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[8] P.G. Ciarlet, P.A. Raviart, General Lagrange and Hermite interpolation in R^n with applications to finite element methods, Arch. Rational Mech. Anal. 46 (1972) 178–199.
[9] P.G. Ciarlet, C. Wagschal, Multipoint Taylor formulas and applications to the finite element method, Numer. Math. 17 (1971) 84–100.
[10] C. Coatmelec, Approximation et interpolation des fonctions differentiables de plusieurs variables, Ann. Sci. Ecole Norm. Sup. 83 (1966) 271–341.
[11] R. Courant, Variational methods for the solution of problems of equilibrium and vibrations, Bull. Amer. Math. Soc. 49 (1943) 1–23.
[12] P.J. Davis, Interpolation and Approximation, Blaisdell, Walthan, MA, 1963 (2nd Edition, Dover, New York, 1975).
[13] Encyklopädie der mathematischen Wissenschaften, Teubner, Leipzig, pp. 1900–1904.
[14] Encyclopédie des Sciences Mathematiques, Gauthier-Villars, Paris, 1906.
[15] I.A. Ezrohi, General forms of the remainder terms of linear formulas in multidimensional approximate analysis I, II Mat. Sb. 38 (1956) 389–416 and 43 (1957) 9–28 (in Russian).
[16] M. Gasca, T. Sauer, Multivariate polynomial interpolation, Adv. Comput. Math., 12 (2000) 377–410.
[17] G. Glaeser, L'interpolation des fonctions differentiables de plusieurs variables, in: C.T.C. Wall (Ed.), Proceedings of Liverpool Singularities Symposium II, Lectures Notes in Mathematics, Vol. 209, Springer, Berlin, 1971, pp. 1–29.
[18] H.H. Goldstine, A History of Numerical Analysis from the 16th Through the 19th Century, Springer, Berlin, 1977.
[19] R.B. Guenter, E.L. Roetman, Some observations on interpolation in higher dimensions, Math. Comput. 24 (1970) 517–521.
[20] E. Isaacson, H.B. Keller, Analysis of Numerical Methods, Wiley, New York, 1966.
[21] C.G.J. Jacobi, Theoremata nova algebraica circa systema duarum aequationum inter duas variabiles propositarum, Crelle J. Reine Angew. Math. 14 (1835) 281–288.
[22] L. Kronecker, Über einige Interpolationsformeln für ganze Funktionen mehrerer Variabeln. Lecture at the academy of sciences, December 21, 1865, in: H. Hensel (Ed.), L. Kroneckers Werke, Vol. I, Teubner, Stuttgart, 1895, pp. 133–141. (reprinted by Chelsea, New York, 1968).
[23] K.S. Kunz, Numerical Analysis, McGraw-Hill, New York, 1957.
[24] G.G. Lorentz, Approximation of Funtions, Chelsea, New York, 1966.
[25] R. Lorentz, Multivariate Hermite interpolation by algebaic polynomials: a survey, J. Comput. Appl. Math. (2000) this volume.
[26] S.E. Mikeladze, Numerical Methods of Mathematical Analysis, Translated from Russian, Office of Tech. Services, Department of Commerce, Washington DC, pp. 521–531.
[27] S. Narumi, Some formulas in the theory of interpolation of many independent variables, Tohoku Math. J. 18 (1920) 309–321.
[28] L. Neder, Interpolationsformeln für Funktionene mehrerer Argumente, Skandinavisk Aktuarietidskrift (1926) 59.
[29] R.A. Nicolaides, On a class of finite elements generated by Lagrange interpolation, SIAM J. Numer. Anal. 9 (1972) 435–445.
[30] R.A. Nicolaides, On a class of finite elements generated by Lagrange interpolation II, SIAM J. Numer. Anal. 10 (1973) 182–189.
[31] K. Pearson, On the construction of tables and on interpolation, Vol. 2, Cambridge University Press, Cambridge, 1920.
[32] J. Radon, Zur mechanischen Kubatur, Monatshefte Math. Physik 52 (1948) 286–300.

[33] H.E. Salzer, Note on interpolation for a function of several variables, Bull. AMS 51 (1945) 279–280.
[34] H.E. Salzer, Table of coefficients for interpolting in functions of two variables, J. Math. Phys. 26 (1948) 294–305.
[35] H.E. Salzer, Note on multivariate interpolation for unequally spaced arguments with an application to double summation, J. SIAM 5 (1957) 254–262.
[36] H.E. Salzer, Some new divided difference algorithms for two variables, in: R.E. Langer (Ed.), On Numerical Approximation, University of Wisconsin Press, Madison, 1959, pp. 61–98.
[37] H.E. Salzer, Divided differences for functions of two variables for irregularly spaced arguments, Numer. Math. 6 (1964) 68–77.
[38] H.E. Salzer, Formulas for bivariate hyperosculatory interpolation, Math. Comput. 25 (1971) 119–133.
[39] H.E. Salzer, G.M. Kimbro, Tables for Bivariate Osculatory Interpolation over a Cartesian Grid, Convair Astronautics, 1958.
[40] A. Sard, Remainders: functions of several variables, Acta Math. 84 (1951) 319–346.
[41] A. Sard, Remainders as integrals of partial derivatives, Proc. Amer. Math. Soc. 3 (1952) 732–741.
[42] T. Sauer, Yuan Xu, On multivariate Lagrange interpolation, Math. Comput. 64 (1995) 1147–1170.
[43] T. Sauer, Yuan Xu, A case study in multivariate Lagrange interpolation, in: S.P. Singh (Ed.), Approximation Theory, Wavelets and Applications, Kluwer Academic Publishers, Dordrecht, 1995, pp. 443–452.
[44] D.D. Stancu, The remainder of certain linear approximation formulas in two variables, J. SIAM Numer. Anal. 1 (1964) 137–163.
[45] I.F. Steffensen, Interpolation, Chelsea, New York, 1927 (2nd Edition, 1950).
[46] G. Strang, G.J. Fix, An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, NJ, 1973.
[47] H.C. Thacher Jr., W.E. Milne, Interpolation in several variables, J. SIAM 8 (1960) 33–42.
[48] H.C. Thacher Jr., Derivation of interpolation formulas in several independent variables, Ann. N.Y. Acad. Sci. 86 (1960) 758–775.
[49] T.N. Thiele, Interpolationsrechnung, Teubner, Leipzig, 1909.
[50] R.S. Walker, Algebraic Curves, Springer, Berlin, 1978.
[51] E.T. Whittaker, G. Robinson, Calculus of Observations, 4th Edition, Blackie and Sons, London, 1944.
[52] O.C. Zienkiewicz, The Finite Element Method in Structural and Continuum Mechanics, McGraw-Hill, London, 1967.
[53] M. Zlamal, On the finite element method, Numer. Math. 12 (1968) 394–409.