Topics in Multivariate Approximation and Interpolation
K. Jetter et al., Editors

Polynomial Interpolation in Several Variables: Lattices, Differences, and Ideals

Tomas Sauer

Lehrstuhl für Numerische Mathematik, Justus–Liebig–Universität Gießen, Heinrich–Buff–Ring 44, D–35192 Gießen, Germany

Abstract

When passing from one to several variables, the nature and structure of polynomial interpolation changes completely: the solvability of the interpolation problem with respect to a given finite dimensional polynomial space, like all polynomials of at most a certain total degree, depends not only on the number, but significantly on the geometry of the nodes. Thus the construction of interpolation sites suitable for a given space of polynomials or of appropriate interpolation spaces for a given set of nodes become challenging and nontrivial problems. The paper will review some of the basic constructions of interpolation lattices which emerge from the geometric characterization due to Chung and Yao. Depending on the structure of the interpolation problem, there are different representations of the interpolation polynomial and several formulas for the error of interpolation, reflecting the underlying point geometry and employing different types of differences. In addition, we point out the close relationship with constructive ideal theory and degree reducing interpolation, whose most prominent representer is the least interpolant, introduced by de Boor et al.

Key words: Polynomial interpolation, geometric characterization, finite difference, ideal basis

2000 MSC: 41A05, 41A63, 13P10

Email address: tomas.sauer@math.uni-giessen.de (Tomas Sauer).

1. Introduction

One of the basic but fundamental tasks in the process of trying to understand the nature of multivariate polynomial interpolation is to provide extensions of the well-known and classical theory for univariate interpolation. As tautologic as this may sound, finding such extensions is often highly nontrivial, and in many cases it is not even agreed upon what the natural or even a reasonable counterpart of a (more or less) related univariate concept or result is. This makes polynomial interpolation a much wider field in several variables than in one variable with contributions ranging from analysis to algebraic geometry and commutative algebra and there are "generalizations" that do not even have any counterpart in one variable. In this paper, I will try to summarize some old and derive some new facts on three specific aspects of multivariate polynomial interpolation: the question of lattice generation, some forms of finite and divided differences that occur as coefficients in the interpolation polynomials and error formulas, and some algebraic aspects, especially the close connection between degree reducing interpolation and the "algorithmic" ideal bases like Gröbner and H–bases. Such bases are essential for computational methods in ideal theory based on computing a remainder of division by an ideal and thus are the basis on which algorithms are built that solve ideal theoretic problems like the question whether a polynomial is contained in a given ideal or not.

Due to its selective scope, this paper cannot be a survey and has no ambition to be one. Readers interested in surveys on multivariate polynomial interpolation can still find a lot of information in [33,31,44,45]. The selection of topics in this paper, however, is unbiased, incomplete and guided entirely by personal preferences. In particular, a lot of valuable and interesting work, like interpolation on spheres and other manifolds, applications in the spline or finite element context or the connection to polynomial system solving will be mentioned not at all or just briefly in passing.

On the other hand, this is also no research paper, though it contains some new results together with full proofs, as quite some part of it just lists (more or less widely) known facts and tries to put them into a context. Once in a while I will even mix in a short proof into these reviews – mainly when I feel that it could be illustrative or give an idea of the type of arguments used at that point. The chapter on algebraic concepts, on the other hand, will be equipped more substantially with proofs, some for the aforementioned reason, but mainly because it contains new material that simply needs to be proved. Also the list of references does not claim completeness. Though it is long and, to some extent, exotic, it is by no means exhaustive and only contains items which are referenced somewhere in the text. I apologize to everyone whose work has not been mentioned here, but *"... every story one chooses to tell is a kind of censorship, it prevents the telling of other tales ..."* [64].

According to [3], the name *interpolation* was introduced by Wallis in 1655. At this time, the goal of interpolation was to estimate the value of a function at a certain point based on the value of the function at other points; the typical application was to fill in values of functions available only in tabulated form like the logarithm. In fact, even the later papers by Aitken [1] and Neville [58] are still mainly driven by this application of interpolation. But also in several variables polynomials still play a fundamental role in the *local* approximation of smooth functions, and interpolation is one of the simplest projectors to polynomial spaces, though it is significantly more intricate than in the univariate case. The "modern" point of view of interpolation as means of reconstructing curves or surfaces from "measurements", on the other hand, has only been taken much more recently in the time of CAGD, and it is a fair question whether interpolation by polynomials can play a reasonable role at all in applications as the polynomials' oscillating nature and the sensitivity of polynomial interpolation to the node configuration really tend to accumulate in the multivariate case. In my opinion, it is more the mathematical beauty of polynomials and the interplay with other fields of mathematics, in particular with computer algebra via constructive ideal theory, that still makes interpolation by algebraic polynomials in several variables a worthwhile topic to study.

While polynomial interpolation in one variable is classical material in almost any textbook on Numerical Analysis, facts about interpolation in several variables are hard to find, and, if at all, it is more likely to find them in the "classics" like Steffensen's book [76] or in [36]. And though the oldest paper on polynomial interpolation in several variables, due to Kronecker [38], considers nodes that are *not* on a tensor product grid, most of the later textbook material only covers *tensor product* interpolation or the structurally similar interpolation on the triangular grid. In view of that, one might be tempted to conjecture that there is something peculiar about interpolation by polynomials in two and more variables, even if this was explicitly denied in [2], cf. [33]. As a general disclaimer that warns of some of the side effects of polynomial interpolation, we will begin by pointing out some of these peculiarities in Section 2. A more positive approach will be taken in Section 3 on lattice generation by means of intersection of hyperplanes; the classical paper by Chung and Yao reviewed there is still the main source for the explicit construction of point configurations that allow for unique interpolation by a total degree space. Section 4 deals with errors, not in the sense of mathematical flaws or roundoff errors in numerical computations, but with the global deviation of an interpolation polynomial from the function it interpolates. This concept is related to generalizations of the concept of divided differences and some of them are given and put into relation with each other. Section 5 simply makes use of the fact that polynomials can also be multiplied to provide an algebraic background in terms of graded rings and polynomial ideals for degree reducing interpolation spaces, pointing out their intimate relationship with normal forms modulo Γ–bases, a concept that simultaneously generalizes Gröbner bases and H–bases. The final Section 6 uses some more special though still sufficiently general types of gradings to arrive at the conclusion that tensor product data is best interpolated by tensor product spaces and to give a brief ideal theoretic view at error formulas. Interested? So let us start ...

2. Some Basics on Multivariate Polynomial Interpolation

The general form of the ***polynomial interpolation problem*** is as follows: given a finite linearly independent set Θ of functionals and an associated vector $Y = (y_\theta \,:\, \theta \in \Theta)$ of prescribed values, find a polynomial f such that

$$\Theta f = Y, \qquad \text{i.e.} \qquad \theta f = y_\theta, \quad \theta \in \Theta. \tag{1}$$

Here we consider polynomials in d variables, $d \geq 1$, involving, for some field $\mathbb{K}$, the *ring* $\Pi := \mathbb{K}[x] = \mathbb{K}[x_1, \dots, x_d]$ of polynomials in d variables, writing $x = (x_1, \dots, x_d)$. It is important to make two comments on the underlying algebra here: first, the field $\mathbb{K}$ should be infinite, with the three "role models" $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$; interpolation on finite fields is a very interesting topic with important applications in coding theory, but structurally it is significantly different from what we will consider here. Second, we will make use of the ***ring*** structure of polynomials when employing concepts from ideal theory. The polynomial interpolation problem by itself can easily be written as a linear algebra problem with respect to the vector space of polynomials, but the additional multiplicative structure of the ring will allow us to draw further conclusions on the behavior of interpolation.

Clearly, the interpolation problem as stated in (1) can be expected to have many solutions as without further restrictions we will be permitted to choose from the infinite dimensional vector space Π an interpolation polynomial f that must only match a finite number of conditions. In view of this it makes perfect sense to consider spaces that allow for *unique* interpolation, i.e., finite dimensional subspaces of Π whose dimension coincides with $\#\Theta$.

Definition 1. Let $\mathcal{P}$ be a subspace of Π. The polynomial interpolation problem with respect to Θ is said to be *poised for $\mathcal{P}$* or *correct for $\mathcal{P}$*, if for any $Y \in \mathbb{K}^\Theta$ there exists a *unique* $f \in \mathcal{P}$ such that $\Theta f = Y$.

There is a somewhat tautologic characterization of poisedness in terms of algebraic geometry and linear algebra. To that end, we will denote by $P \subset \mathcal{P}$ a basis for the finite dimensional space $\mathcal{P}$.

Theorem 2. *For $\mathcal{P} \subset \Pi$ and a finite set $\Theta \subset \Pi'$ the following statements are equivalent:*

(i) The polynomial interpolation problem with respect to Θ is poised for $\mathcal{P}$.

(ii) $\dim \mathcal{P} = \#\Theta$ *and the matrix*

$$\Theta P = [\theta p \,:\, \theta \in \Theta, p \in P] \in \mathbb{K}^{\Theta \times P} \tag{2}$$

satisfies $\det \Theta P \neq 0$ *for any basis P of $\mathcal{P}$.*

(iii) $\ker \Theta \cap \mathcal{P} = \{0\}$.

Usually, condition (iii) of the above theorem is phrased as *there is no algebraic hypersurface in $\mathcal{P}$ that contains Θ*, at least when Θ consists of point evaluation functionals or as points with multiplicities – we will get to this issue in the next chapter. Unfortunately, none of the two criteria above is easier to verify than the uniqueness of polynomial interpolation, at least not in general.

One remark on the notation in (2) which will be used throughout the paper: Θ and P are sets which can be used to index the elements of vectors or matrices like in (2). Moreover, each set gives rise to a natural vector, e.g., $P = (p : p \in P)$, which can be multiplied to such matrices. This notation allows, for example, to conveniently write

$$L = (\Theta P)^{-1} \, P = \sum_{p \in P} (\Theta P)_p^{-1} \, p \qquad \Rightarrow \qquad \Theta L = (\Theta P)^{-1} \, \Theta P = I,$$

for the dual basis L of Θ without having to state explicitly the cardinalities of Θ and P – except that, in order have an inverse, the matrix ΘP must be a square one, that is, $\#\Theta = \#P$.

2.1. Lagrange and the trouble with Hermite

It is common practice to classify interpolation problems according to the type of the functionals $\theta \in \Theta$ that have to be matched by the polynomial. If those functionals are point evaluations, i.e., $\theta = \delta_\xi$, $\xi \in \mathbb{K}^d$, the interpolation problem is called a *Lagrange interpolation problem* and the points are called *nodes* or *(interpolation) sites*. If functionals are consecutive derivatives at some points, one usually speaks of a *Hermite interpolation problem* while the remaining problems are usually named *Hermite–Birkhoff interpolation problems*. At least, this is the terminology for the univariate case. In several variables it is not even agreed upon what a Hermite interpolation problem should be. The most convincing generalization requires that Θ is an *ideal interpolation scheme* in the sense of Birkhoff [5] which means that the set $\ker \Theta = \{f \in \Pi \, : \, \Theta f = 0\}$ of all polynomials annihilated by all the functionals in Θ forms an ideal in Π. Substantial investigations of Hermite interpolation in connection with ideal schemes have been performed even earlier by Möller [52,53] who already proposed ideal schemes as the natural generalization of Hermite interpolation problem. And indeed the concept of ideal interpolation schemes provides precisely what one would expect, namely interpolation of derivatives at different nodes. More precisely, it was shown in [11,48] that a set Θ of functionals defines an ideal interpolation scheme if and only if there exists a finite subset $\Xi \subset \mathbb{K}^d$ and finite dimensional polynomial subspaces $\mathcal{Q}_\xi \subset \Pi$, $\xi \in \Xi$, each of them closed under differentiation, such that

$$\operatorname{span} \Theta = \operatorname{span} \{\delta_\xi \circ q(D) \, : \, q \in \mathcal{Q}_\xi, \xi \in \Xi\}, \qquad \bigcup_{j=1}^{d} \frac{\partial}{\partial x_j} \mathcal{Q}_\xi \subseteq \mathcal{Q}_\xi, \quad \xi \in \Xi,$$

where δ_ξ denotes the point evaluation functional at ξ. In fact, this is nothing else than interpolation *with multiplicities* at the finite set Ξ! To that end, one has to take into account that the multiplicity of a zero of a function or a common zero of finitely many functions is no more a number, i.e., a matter of counting, in two and more variables: it is a ***structural*** quantity, most conveniently expressed as a *D–invariant* space of polynomials, that is a finite dimensional subspace $\mathcal{Q} \subset \Pi$ closed under differentiation.

In the univariate case, Hermite interpolation problems are introduced rather conveniently as the limit of Lagrange interpolation problems with coalescing nodes; alternatively, this could also be seen as starting with the interpolation of (divided or at least reasonably normalized) differences which then converge to derivatives in the dual of the interpolation space. But the main point in the univariate case is that the Lagrange interpolation problem with respect to $n+1$ distinct sites is poised for Π_n, the polynomials of degree up to n, and this also remains a valid interpolation space upon coalescence. This property holds true no more in two or more variables as a very simple example shows: consider, for $h \in \mathbb{R} \setminus \{0\}$, the six point evaluation functionals that map f to

$$f(0,0),\ f(h,0),\ f(0,h),\ f(1,1),\ f(1+h,1),\ f(1,1+h)$$

which give rise to a poised Lagrange interpolation problem for

$$\Pi_2 = \operatorname{span}\left\{1, x, y, x^2, xy, y^2\right\}$$

as long as $h \neq 0$. More precisely, the Vandermonde determinant of this interpolation problem can be easily computed to be $-4h^5$ which, on the other hand, already indicates that there may be some trouble with the limit problem that has $\mathcal{Q}_{(0,0)} = \mathcal{Q}_{(1,1)} = \Pi_1$, interpolating point values and first derivatives at the two points. And indeed, this Hermite interpolation problem is not poised any more for Π_2 which can be verified by direct computations, but also follows from a slightly more general principle due to [73]: interpolating all derivatives of order k and k' at $\xi, \xi' \in \Xi$ by means of Π_n requires that $n > k + k'$. But it can even be seen directly from the discretization of the limit functionals in terms of the directional divided differences

$$\begin{aligned}
&f(0,0),\ \frac{f(h,0)-f(0,0)}{h},\ \frac{f(0,h)-f(0,0)}{h},\\
&f(1,1),\ \frac{f(1+h,1)-f(1,1)}{h},\ \frac{f(1,1+h)-f(1,1)}{h},
\end{aligned}$$

that trouble is to be expected: the Vandermonde determinant associated to the interpolation problem with respect to these functionals is $-4h$ then and this still converges to zero for $h \to 0$.

To make it clear: if the Hermite interpolation problem is poised for some subspace $\mathcal{P}$ of Π, then so are almost all Lagrange interpolation problems where partial derivatives are replaced by the associated divided differences; this is a simple consequence of the continuity of the determinant. However, there is no converse any more to the above statement as even the poisedness of ***all*** Lagrange interpolation problems does not guarantee the poisedness of the limit Hermite problem, and this is an exclusively multivariate phenomenon.

2.2. The loss of Haar

Depending on the point of view, the most scary or the most challenging aspect of multivariate polynomial interpolation is the fact that geometry really begins to matter now. In one variable, the Lagrange interpolation problem with respect to $n+1$ nodes is poised for Π_n, so all that has to be done was to match the number of interpolation sites and the dimension of the space. Univariate polynomials up to a certain degree form a *Haar space*: they allow for unique interpolation at an appropriate *number* of distinct nodes, regardless of their *position*. This is no more true in two and more variables, cf. [40], and generally Mairhuber's theorem [47] makes it clear that Haar spaces exist essentially for two different topologies: the interval and the unit circle which can be seen as the periodic relative of the interval. Consequently, the loss of Haar (and I do not claim to have invented this terminology) is unavoidable in $\mathbb{K}^d$, $d > 1$.

In particular, this means that for no polynomial subspace $\mathcal{P}$ of Π the interpolation problem can be poised with respect to *all* interpolation conditions Θ, not even if Θ consists only of point evaluation functionals. On the other hand, if $\Theta = \{\delta_\xi \ : \ \xi \in \Xi\}$ defines a Lagrange interpolation problem and P is a basis of an interpolation space $\mathcal{P}$, then the matrix

$$P(\Xi) := \Theta P = [p(\xi) = \delta_\xi p \ : \ \xi \in \Xi, p \in P]$$

is a nonzero polynomial in the $d \times \#\Xi$ variables $(\xi \ : \ \xi \in \Xi)$ and therefore vanishes at most on a set of measure zero. This fact we record in the following statement.

Proposition 3. *If for a subspace $\mathcal{P}$ of Π there exists a set Ξ of* $\dim \mathcal{P}$ *interpolation sites in $\mathbb{K}^d$ such that the interpolation problem with respect to Ξ is poised for $\mathcal{P}$ then* $\det P(\Xi) \neq 0$ *for any basis P of $\mathcal{P}$ and almost any $\Xi \subset \mathbb{K}^d$, $\#\Xi = \dim \mathcal{P}$.*

This property of Lagrange interpolation problems is called *almost poisedness* of the interpolation problem for $\mathcal{P} \subset \Pi$ and is the "proper" counterpart for the Haar space property. In particular, almost poisedness again simply requires the dimension of the interpolation space to match the cardinality of the set of interpolation nodes.

Proposition 4. *For every finite dimensional subspace $\mathcal{P}$ of Π the* Lagrange interpolation problem *is poised with respect to almost any $\Xi \subset \mathbb{K}^d$ as long as $\#\Xi =$* $\dim \mathcal{P}$.

Proof. We first show by induction on $N := \#P = \dim \mathcal{P}$ that there always exists $\Xi \subset \mathbb{K}^d$ such that $\det P(\Xi) \neq 0$ from which the claim follows immediately by applying Proposition 3. If $N = 1$ then P consists of a single nonzero polynomial which clearly does not vanish at some $\xi \in \mathbb{K}^d$. To advance the induction hypothesis from N to $N+1$, we choose $p \in P$, set

$$P' := P \setminus \{p\}$$

and note that by the induction hypothesis there exists $\Xi' \subset \mathbb{K}^d$, $\#\Xi' = \dim \mathcal{P} - 1$, such that $\det P'(\Xi') \neq 0$. Therefore, the polynomial

$$q := p - p(\Xi')^T \, P'(\Xi')^{-1} \, P'$$

vanishes at Ξ', but is nonzero by linear independence, thus $q(\xi) \neq 0$ for some point $\xi \in \mathbb{K}^d \setminus \Xi'$. With $\Xi = \Xi' \cup \{\xi\}$ we then get that

$$\det P(\Xi) = \det \{q, P'\} (\{\xi, \Xi'\}) = q(\xi) \, \det P'(\Xi') \neq 0,$$

which completes the proof. □

So it seems as if almost poisedness is the property to go for and therefore all the difference between the univariate and the multivariate case is something that happens just on a set of measure zero? Unfortunately, this is not the case, as there is also the issue of *never poisedness* when derivatives enter the scene: there are *Hermite* interpolation problems that will never be poised for a given interpolation space of fitting dimension, regardless of how we choose the nodes. And we already know an example: the innocent interpolation of first derivatives at two points in $\mathbb{R}^2$ from Π_2. This issue, that definitely adds a different flavor to Birkhoff interpolation problems in several variables, has been addressed to some extent in [73] for Hermite problems, but the much more systematic and substantial treatment has been done by Lorentz and Lorentz in [41–43] and is mostly summarized in [44] to which the reader is referred for additional information.

2.3. Let us keep it simple

We have seen that there is a difference between Lagrange and Hermite interpolation problems in two and more variables and that the simple and intuitive way to relate them by means of coalescence cannot be followed so easily – and we haven't even mentioned the question of points coalescing along lines or curves and similar issues yet. In fact, though many of the facts and concepts presented in what follows also continue to persist for Hermite interpolation (in Möller's, Birkhoff's and de Boor's sense of ideal interpolation; and I consider it justified to vote with this majority myself), the main focus will be on Lagrange interpolation and potential or possible extensions to Hermite problems will only be mentioned in passing. Nevertheless, I can assure the ambitious reader that Lagrange interpolation in several variables is by itself a sufficiently intricate issue.

3. Lattice Generation – The Legacy of Chung and Yao

One of the consequences of the "loss of Haar" is that there is no a priori certainty that a given node configuration (or a given set of *data sites*) really admits a poised interpolation problem for a fixed polynomial subspace. This problem is frequently "resolved" for "applications" by making use of the handwaving argument that essentially all this fuss is just about a set of measure zero only, but unfortunately any interpolation problem "close" to such configurations becomes terribly ill–conditioned and thus useless. Consequently, it makes sense to ask for explicit constructions of data sites with respect to which the interpolation problem is poised for a given subspace $\mathcal{P}$. The most common choice for $\mathcal{P}$ is Π_n, the space of all polynomials of total degree at most n. Recall that a polynomial f is a *finite* sum of the form

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^d} f_\alpha x^\alpha, \qquad f_\alpha \in \mathbb{K}, \qquad \# \{\alpha \ : \ f_\alpha \neq 0\} < \infty,$$

for which the *total degree* $\deg f$ is defined as

$$\deg f := \max \{|\alpha| \ : \ f_\alpha \neq 0\}, \qquad |\alpha| := \sum_{j=1}^d \alpha_j.$$

Then $\Pi_n := \{f \in \Pi \ : \ \deg f \leq n\}$ is a finite dimensional $\mathbb{K}$–vector space of dimension $\binom{n+d}{n}$. At present, I just want to mention that the total degree is a very natural, but by far not the only way to extend the notion of degree to polynomials in several variables; we will encounter the more general concept of *graded rings* later in Section 5. For the time being, however, we will focus on the problem to construct sets $\Xi \subset \mathbb{K}^d$ of $\binom{n+d}{n}$ distinct nodes such that the interpolation problem with respect to Ξ is poised for Π_n.

3.1. The geometric characterization

The geometric characterization due to Chung and Yao [22] is based on the intersection of hyperplanes. Recall that a hyperplane $H \subset \mathbb{K}^d$ is the zero set of a nondegenerate affine function $h \in \Pi_1$, i.e.,

$$H = \left\{x \in \mathbb{K}^d \ : \ 0 = h(x) = v^T x + c\right\}, \qquad v \in \mathbb{K}^d \setminus \{0\}, \, c \in \mathbb{K}.$$

A set $\Xi \subset \mathbb{K}^d$, $\#\Xi = \binom{n+d}{d}$, satisfies the *geometric characterization* of degree n if for any $\xi \in \Xi$ there exist hyperplanes $H_{\xi,1}, \dots, H_{\xi,n}$ such that

$$\xi \notin \bigcup_{j=1}^n H_{\xi,j}, \qquad \Xi \setminus \{\xi\} \subset \bigcup_{j=1}^n H_{\xi,j}. \tag{3}$$

The geometric characterization is clearly a sufficient condition for the poisedness of the interpolation problem with respect to Ξ for Π_n, as it implies that the *Lagrange fundamental polynomials*

$$\ell_\xi := \prod_{j=1}^n \frac{h_{\xi,j}}{h_{\xi,j}(\xi)}, \qquad \xi \in \Xi, \tag{4}$$

are well defined and satisfy $\ell_\xi(\xi') = \delta_{\xi,\xi'}$ for $\xi, \xi' \in \Xi$. However, the fundamental polynomials have an additional property: they are *factorizable* into affine factors, a property quite "normal" for univariate polynomials but very rare for multivariate ones. On the other hand, if the fundamental polynomial ℓ_ξ which must be of total degree precisely n, can be factored into n affine factors then the zero set of this polynomial is a union of n hyperplanes not containing ξ while the remaining sites in $\Xi \setminus \{\xi\}$ must be distributed over these hyperplanes. This, however, is nothing else than (3) and so we have proved the following result.

Theorem 5. *Let $\Xi \subset \mathbb{K}^d$ be a set of $\binom{n+d}{d}$ distinct points. The Lagrange fundamental polynomials $\ell_\xi \in \Pi_n$, $\xi \in \Xi$, defined by $\ell_\xi(\xi') = \delta_{\xi,\xi'}$ can be factorized into affine factors if and only if Ξ satisfies the geometric characterization.*

Though condition (3) even characterizes a certain type of point configurations for interpolation, it mainly serves the purpose of being a ***sufficient*** condition for poisedness or as device for constructing node configurations which lead to poised interpolation problems. However, this appears paradoxical in some sense: on the one hand, the Lagrange interpolation problem with respect to Π_n is poised for almost any choice Ξ of cardinality $\binom{n+d}{d}$, on the other hand, the most popular constructions of such interpolation sets place the sites appropriately on (intersections of) finitely many hyperplanes, i.e., on sets of measure zero. Nevertheless, these point configurations offer a lot of appealing properties that have been discovered in the course of a substantial amount of research, most of which is summarized and explained in [26].

3.2. Natural lattices

The first specific method in [22] to construct point configurations that satisfy the geometric characterization is the approach of ***natural lattices***. Here, one chooses, depending of the order $n \in \mathbb{N}_0$ of the lattice, a set $\mathcal{H}$ of $n+d$ hyperplanes ***in general position***. The latter means that any d of the hyperplanes intersect in a point (which can be assured by requiring that all the normal vectors to the hyperplanes are in general position, i.e., that any d of them are linearly independent) and that no $d+1$ hyperplanes have a point in common. That is, any set $\mathcal{K} \subset \mathcal{H}$ of cardinality d, or, in the convenient shorthand notation from [8], any $\mathcal{K} \in \binom{\mathcal{H}}{d}$, defines a unique point $\xi_{\mathcal{K}}$ and for $\mathcal{K}, \mathcal{K}' \in \binom{\mathcal{H}}{d}$, $\mathcal{K} \neq \mathcal{K}'$, we have that $\xi_{\mathcal{K}} \neq \xi_{\mathcal{K}'}$. Since there are $\binom{n+d}{d}$ such subsets $\mathcal{K} \in \binom{\mathcal{H}}{d}$, this process defines a set Ξ of $\binom{n+d}{d}$ distinct nodes ξ_K which satisfy the geometric characterization.

There has been a lot of work on natural lattices, in particular by Mariano Gasca and his collaborators; for a fairly complete list of references and a more detailed description see [26,31].

3.3. Principal lattices

Although the very first explicit work on multivariate polynomial interpolation [38], cf. [33], that I could track down so far is concerned with what is called the *complete intersection case* nowadays, the "classical", "early" point configurations for bivariate and multivariate interpolation have been the tensor product and the structurally similar triangular grid. These grids can be seen as the intersection of a set of lines, for example those parallel to the y–axis, with "transversal" lines, for example those parallel to the x–axis. This is the geometric idea behind the concept of ***principal lattices*** that start with d families of $n+1$ distinct hyperplanes, $H_{j,k}$, $j = 1, \dots, d$, $k = 0, \dots, n$. For $|\alpha| \le n$, the interpolation sites are now chosen as the intersections

$$\xi_\alpha := \bigcap_{j=1}^{d} H_{j,\alpha_j},$$

with the ***assumption*** that these points are distinct. Since the polynomials

$$p_\alpha := \prod_{j=1}^{d} \prod_{k=0}^{\alpha_j - 1} \frac{h_{j,k}}{h_{j,k}(\xi_\alpha)}, \qquad |\alpha| \le n, \tag{5}$$

satisfy $p_\alpha(\xi_\beta) = \delta_{\alpha,\beta}$, $|\beta| \le |\alpha| \le n$, if follows that for $P = \{p_\alpha : |\alpha| \le n\}$ and $\Xi = \{\xi_\alpha : |\alpha| \le n\}$ the matrix $[p_\beta(\xi_\alpha) : |\alpha|, |\beta| \le n]$, being a rearrangement of $P(\Xi)$, is block upper triangular with identities on the diagonal and thus invertible.

Geometrically, principal lattices define the nodes by intersection of their duals, namely the hyperplanes, which are the same as the interpolation sites in one variable. In this respect principal lattices form a very natural generalization of univariate polynomial interpolation, and indeed some univariate features can be recovered in principal lattices. For example, as shown in [32] for the bivariate case, coalescence of the hyperplanes always leads to a poised Hermite interpolation problem, but then several sites become "Hermite sites" simultaneously. It also turns out that a geometric variant of the Aitken–Neville method (by means of barycentric coordinates and affine combinations) works only if the interpolation nodes form a principal lattice, see [75].

There is a multitude of constructions for principal lattices, many using ***pencils of hyperplanes*** as introduced by Lee and Phillips [39] and elaborated in a series of papers by Gasca and Mühlbach [28–30]. This approach nicely connects to projective geometry and again much of this work is summarized and referenced in [31]. However, it is worthwhile to also look at the recent [20] which employs the algebraic concept of ***addition of points on a cubic***.

3.4. There are still open questions

Let us consider natural and principal lattices in $d = 2$ for the moment, that is, the hyperplanes simply become lines in the plane. In the first case of natural lattices, any line intersects with any of the $n+1$ other lines and so even ***any*** line contains $n+1$ interpolation sites. In the case of principal lattices, on the other hand, the lines $H_{1,0}$ and $H_{2,0}$ also have $n+1$ interpolation sites lying on them, namely $(0,k)$ and $(k,0)$, $k = 0, \dots, n$. This evidence is a first suggestion of the fascinating observation is that ***all*** lattices satisfying the geometric characterization in $d = 2$ seem to have this particularly nice structure which has been conjectured in [27].

Conjecture 6 (Gasca). *If $\Xi \subset \mathbb{R}^2$ satisfies the geometric characterization of degree n in $d = 2$, then there always exists a line H that contains $n+1$ sites.*

This conjecture, which still remains unproven at the time of writing, would allow for a recursive "Newton" approach to interpolation on natural lattices: if H is any line containing $n+1$ nodes then the removal of H and the $n+1$ sites on it leaves a set Ξ' that satisfies the geometric characterization of degree $n-1$ and can be further decomposed by the same argument. The conjecture is trivial for $n = 1$, easily proved by means of a counting argument for $n = 2$ and there also exists a quite complicated proof for $n = 3$, see [17]. The most recent work on Conjecture 6, to be found in [18,19], does not yet prove the conjecture but gives a classification of node sets that satisfy the geometric condition and actually indicates that probably an even stronger version of Conjecture 6 may hold.

4. Error, Formulas, and Differences

Clearly, interpolation polynomials coincide with the function f to be interpolated at the nodes. At least they should, even if they usually refuse to do so in numerical computations once the degree exceeds certain bounds, cf. [65], but also [59,60] for some further aspects of roundoff error analysis for multivariate polynomials. In addition to the unavoidable numerical difficulties, it is important to understand the approximation behavior of the interpolation polynomials, i.e., the behavior of the error of interpolation $f - L_n f$. Here, for simplicity, f is a sufficiently smooth function defined on some compact set $\Omega \subset \mathbb{R}^d$ (we neglect the more general fields for the sake of analysis), and L_n is the interpolation operator with respect to some set $\Xi \subset \Omega$ such that the interpolation problem with respect to Ξ is poised for Π_n.

4.1. The Newton approach and the finite difference

The first step towards the error formula was originally motivated by a very simple observation: the Vandermonde matrix

$$[\xi^\alpha \ : \ \xi \in \Xi, \ |\alpha| \leq n]$$

with respect to the canonical *monomial basis* of Π_n indexes the nodes and the basis elements in two different ways. So why not re-index the nodes as $\Xi = \{\xi_\alpha \ : \ |\alpha| \leq n\}$? Clearly, this is always possible in many ways, so that we make the additional requirement that the interpolation problem with respect to the node subsets $\Xi_k := \{\xi_\alpha \ : \ |\alpha| \leq k\} \subseteq \Xi$, $k = 0, \ldots, n$, is poised for Π_k. The "dual" polynomials for this arrangement of nodes are the *Newton fundamental polynomials* p_α, $|\alpha| \leq n$, defined by the requirement that

$$p_\alpha \in \Pi_{|\alpha|}, \qquad p_\alpha(\xi_\beta) = \delta_{\alpha,\beta}, \quad |\beta| \leq |\alpha| \leq n. \tag{6}$$

The *Newton basis* $P = \{p_\alpha : |\alpha| \le n\}$ of Π_n has the property that the matrix $[p_\beta(\xi_\alpha) : |\alpha|, |\beta| \le n]$ is block upper triangular with identity matrices on the diagonal. Moreover, the polynomials in (6), which generalize the ones from (5), can be computed as a by–product of the indexing process of Ξ, cf. [6,72], either by means of Gauss elimination or a Gram–Schmidt orthogonalization process; mathematically, both approaches are equivalent, but they differ in implementation details. Even the polynomials p_α do not depend uniquely on Ξ: in the generic case, even *any* $\xi \in \Xi$ can be chosen as ξ_α and the polynomials have to adapt accordingly. This is the point where some subtle problems arise with Hermite interpolation: for a reasonable generalization of the Newton basis it would be necessary to index Θ as $\{\theta_\alpha : |\alpha| \le n\}$ in such a way that, for $k = 0, \dots, n$,

(i) the interpolation problem with respect to Θ_k is poised for Π_k *and*

(ii) $\ker \Theta_k$ is also an ideal.

It is not difficult to ensure each of these properties *separately*, which corresponds to row or column permutations in Gaussian elimination, but how to satisfy them simultaneously is a tricky issue. Recall that the above conditions permit "putting the interpolation problem into block" [73], hold trivially for Lagrange interpolation problems (since condition ii is always satisfied then) and were essential for extending the Newton approach from Lagrange to Hermite interpolation problems in [73]. At the present it is not clear whether or not this can be done for a general set of functionals, but evidence fortunately is more on the optimist's than on the pessimist's side.

Conjecture 7. *Any finite set $\Theta \subset \Pi'$ of linear functionals which admits a poised ideal interpolation problem for Π_n can be graded as $\Theta_0 \subset \Theta_1 \subset \cdots \subset \Theta_n = \Theta$ such that for $k = 0, \dots, n$ the interpolation with respect to Θ_k is poised for Π_k and* $\ker \Theta_k$ *is an ideal in* Π.

As shown in [72], the interpolant of degree n to f can now be written in terms of the Newton basis as

$$L_n f = \sum_{|\alpha| \le n} \lambda\left[\Xi_{|\alpha|-1}, \xi_\alpha\right] f \cdot p_\alpha, \tag{7}$$

where the *finite differences* $\lambda\left[\Xi_k, x\right] f$ satisfy the recurrence relation

$$\lambda[x] f = f(x), \tag{8}$$

$$\lambda\left[\Xi_k, x\right] f = \lambda\left[\Xi_{k-1}, x\right] - \sum_{|\alpha|=k} \lambda\left[\Xi_{k-1}, \xi_\alpha\right] f \cdot p_\alpha(x), \qquad k = 0, \dots, n, \tag{9}$$

and

$$\lambda\left[\Xi_k, x\right] = \left(f - L_k f\right)(x), \qquad x \in \mathbb{R}^d. \tag{10}$$

Substituting (10) into (7) shows that the Newton approach to polynomial interpolation is repeated interpolation of errors:

$$L_n f = L_{n-1} + \sum_{|\alpha|=n} (f - L_{n-1})(\xi_\alpha) \cdot p_\alpha.$$

The difference $\lambda[\Xi_k, x] f$ is called *finite* difference for the simple reason that it is **not** a *divided* difference. This is most easily confirmed by looking at (10) from which it follows that

$$\lim_{x \to \Xi_k} \lambda[\Xi_k, x] = 0,$$

where $x \to \Xi_k$ means that $x \to \xi_\alpha$ for some $|\alpha| \leq k$. Another, even more instructive, example is to consider $\xi_\alpha = h\alpha$, $h > 0$, as then

$$\lambda[\Xi_{k-1}, \xi_\alpha] f = \Delta_h^\alpha f(0), \qquad |\alpha| \leq n,$$

with the standard forward difference Δ_h^α with step width h. The associated *divided* difference, on the other hand, would be $h^{-|\alpha|}\Delta_h^\alpha$ in this case, hence, it is precisely the division that is lacking in $\lambda\left[\Xi_{|\alpha|-1}, \xi_\alpha\right] f$. Also in the univariate case there is a big difference between the differences: denoting by $[x_0, \dots, x_n] f$ the divided difference, we get that

$$\lambda[x_0, \dots, x_n] f = (x_n - x_0) \cdots (x_n - x_{n-1}) \, [x_0, \dots, x_n] f.$$

4.2. Error formulas

The error formulas we consider here are describing the error of interpolation $f - L_n f$ in terms of integrating $(n+1)$–st derivatives of the function f against linear combinations of simplex splines as introduced by Micchelli in [50,51]. Recall that the *simplex spline* for the finite knot set $T \subset \mathbb{R}^d$, most conveniently written as a matrix $\mathrm{T} = [\tau : \tau \in \mathrm{T}] \subset \mathbb{R}^{d \times \#\mathrm{T}}$, is the density of the functional

$$f \mapsto \int_{[T]} f = \int_{\mathbb{S}_{\#T}} f(\mathrm{T}\, u)\, du, \tag{11}$$

where

$$\mathbb{S}_m := \left\{ u = (u_0, \dots, u_m) \in \mathbb{R}^{m+1} \; : \; u_j \geq 0, \sum_{j=0}^m u_j = 1 \right\}$$

denotes the standard m–simplex. Based on simplex splines, de Boor introduced in [7] a concept of a *multivariate divided difference* for $\mathrm{T} = [\tau_j \; : \; j = 0, \dots, n] \in \mathbb{R}^{d \times n+1}$ and $\mathrm{H} = [\eta_j \; : \; j = 1, \dots, n] \in \mathbb{R}^{d \times n}$ whose columns are interpreted as nodes and associated directions, respectively; the difference is then defined as

$$[\mathrm{T}; \mathrm{H}] f = \int_{[\mathrm{T}]} D_{\eta_1} \cdots D_{\eta_n} f.$$

Recalling the classical Hermite–Genocchi formula, which expresses the univariate divided difference in terms of a B–spline integral and is the origin from which many property of simplex splines were derived, cf. [51], one uses as "directions" differences of successive nodes, and so we specialize de Boor's divided difference to

$$[\mathrm{T}]\, f := [\mathrm{T}; \mathrm{T}D]\, f, \qquad D = \begin{bmatrix} -1 & & \\ 1 & \ddots & \\ & \ddots & -1 \\ & & 1 \end{bmatrix} \in \mathbb{R}^{n+1 \times n},$$

which will become one essential building block for the remainder formula. The other is the concept of a *path* of length n which is a vector $\mu = \left(\mu^0, \ldots, \mu^n\right)$ of multiindices such that $\left|\mu^k\right| = k$, $k = 0, \ldots, n$. The set of all paths of length n will be denoted by M_n. Associated to a path $\mu \in \mathrm{M}_n$ and a properly indexed set $\Xi = \{\xi_\alpha \ : \ |\alpha| \le n\}$ we obtain a matrix $\Xi_\mu := \left[\xi_{\mu^j} \ : \ j = 0, \ldots, n\right]$ of sites visited along the path as well as the number

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

With this notation at hand, the error formula from [72] takes the convenient form

$$\left(f - L_n f\right)(x) = \sum_{\mu \in \mathrm{M}_n} p_{\mu^n}(x)\, \pi_\mu\, \left[\Xi_\mu, x\right] f. \tag{12}$$

This formula illustrates how significantly things change when passing from univariate to multivariate polynomial interpolation. In one variable, there is precisely one path and (12) is the well–known error formula with B–splines as Peano kernels that can be encountered in many textbooks as for example in [25,36]. The multivariate version of this formula, however, contains

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

terms in the summation which already becomes $(n+1)!$ for $d = 2$ and grows much faster for higher values of d. In particular, these numbers by far exceed even the "dimension curse" factor d^n which is widely accepted as an unavoidable growth rate in multivariate problems. But the number of terms in (12) appears even more peculiar when comparing it to another formula for the error of multivariate polynomial interpolation which is due to Ciarlet and Raviart [23] and uses the Lagrange fundamental polynomials ℓ_α, $|\alpha| \le n$, to express the error as

$$\left(f - L_n f\right)(x) = \sum_{|\alpha| \le n} \ell_\alpha(x)\, \left[\xi_\alpha^n, x; (x - \xi_\alpha)^n\right] f, \tag{13}$$

where exponentiation indicates n–fold repetition of a column in the respective matrix. This formula, based on a multipoint Taylor expansion, has become an important tool in the analysis of finite elements and it has the apparent advantage that the number of terms in the sum equals the number of interpolation sites, a number much smaller than the number of paths – except, of course, in the univariate case where it becomes a formula due to Kowalewski, [37, eq. (24), p. 24], but not the "standard" error formula. In this respect, (13) is *no* (direct) generalization of the univariate formula, while (12) definitely is. Another fundamental difference is that the simplex spline integrals in (12) normally run over nondegenerate convex sets in $\mathbb{R}^d$, while the one in (13) only integrates along lines. The surprising fact that these two expressions nevertheless describe precisely the same quantity, namely the error of interpolation, shows that there is still quite a bit of magic going on among simplex splines.

4.3. The error on lattices

For interpolation sites located on natural or principal lattices, one can also hope for the associated error formulas to take a very special form. We begin with the error by interpolation on natural lattices which is expressed in a beautiful and very compact way in the formula

$$(f - L_n f)(x) = \sum_{\mathcal{K} \in \binom{\mathcal{H}}{d}} p_{\mathcal{K}}(x) \left[\Xi_{\mathcal{K}}, x; D_{\eta_{\mathcal{K}}}^{n+1}\right] f \tag{14}$$

from [8]. Apparently, some of the ingredients in this formula we haven't encountered yet, so let us figure out what they are. In fact, the notation $\mathcal{K} \in \binom{\mathcal{H}}{d}$ has appeared in Section 3.2 and it stands for choosing a subset $\mathcal{K}$ of $\mathcal{H}$ consisting of d hyperplanes; since they are in general position, they intersect in a line with normalized directional vector η_K. On the other hand, the complement $\mathcal{H} \setminus \mathcal{K}$ consists of $n + d - d = n$ hyperplanes whose associated normalized affine polynomials multiply to $p_{\mathcal{K}}$, and each of these hyperplanes intersects with the line defined by $\mathcal{K}$ in a different point – these n nodes are collected in $\Xi_{\mathcal{K}}$. Note that this is quite much in the spirit of the formula in (12): since they are all lining up along the line in direction $\eta_{\mathcal{K}}$, the difference of any two nodes in $\Xi_{\mathcal{K}}$ is clearly a multiple of η_k, only the directional derivative in the direction $x - \xi$ for some $\xi \in \Xi_{\mathcal{K}}$ is missing relative to (12); apparently it is compensated in the $p_{\mathcal{K}}$.

Also for principal lattices there exists a special form for the error formula which has been developed for the bivariate case in [32], the general situation will be considered in the forthcoming [21]. The expression takes the form

$$(f - L_n f)(x) = \sum_{\mu \in \mathrm{M}_n^*} \frac{p_{\mu^n}(x)}{\theta_\mu} \left[\Xi_\mu, x; \mathrm{H}_\mu, x - \xi_{\mu^n}\right] f, \tag{15}$$

where M_n^* denotes the (much smaller) set of all paths μ such that $\mu^{k+1} - \mu^k \in \mathbb{N}_0^d$, H_μ is the set of *normalized* directions that are followed along this path while θ_μ – the "new" quantity – is the product of the "angles" encountered along the path. Without going into the intricate details, cf. [32], it is just worthwhile to mention here that the identity (15) highlights the crucial geometric parameters in this type of interpolation, namely the angle of intersection between the hyperplanes that define the lattice.

Another common property of the formulas (14) and (15) is that, in accordance with line geometry, the formulas are invariant under scaling – the distance between the lines is irrelevant in the error of interpolation. This, in turn, is the major tool that allows us for such special configurations to pass from Lagrange to Hermite interpolation problems by means of coalescing lines (not points!), see [32] for the hairy details.

4.4. Different differences

We have made the point before that the finite difference from (8) is *not* a divided difference, simply because it lacks division. So the question remains what a *divided difference* is in the context of multivariate polynomial interpolation. One way to generalize the notion of divided difference is to retreat to saying that it is the leading coefficient of the interpolation polynomial, i.e., $L_n f(x) = [x_0, \ldots, x_n] \, f \, x^n + q(x)$, $q \in \Pi_{n-1}$. This notion of a divided difference has been investigated in [62] and also, for special configurations, in [75]. On the other hand, until we will know "better" in the next section, the leading coefficient of a polynomial is now a *vector* of coefficients. More precisely, setting $X^k = (X_\alpha = (\cdot)^\alpha \; : \; |\alpha| = k)$, we can write a polynomial of degree n as

$$f = \sum_{k=0}^{n} f_k^T \, X^k, \qquad f_k = (f_\alpha \; : \; |\alpha| = k) \, .$$

With $P^k = (p_\alpha \; : \; |\alpha| = k)$ and $\lambda \, [\Xi_k] \, f = (\lambda \, [\Xi_{k-1}, \xi_\alpha] \, f \; : \; |\alpha| = k)$ we thus get from the Newton formula (7) that

$$L_n f = \sum_{k=0} \lambda \, [\Xi_k] \, f^T \, P^k = \sum_{k=0}^{n} c_k^T \, X^k,$$

where c_n would be the desired divided difference. If M_α denotes the *monomial Newton basis polynomial,* defined by the requirements

$$M_\alpha - (\cdot)^\alpha \in \Pi_{|\alpha|-1}, \qquad M_\alpha \, (\xi_\beta) = 0, \quad |\beta| < |\alpha|,$$

then there exist invertible matrices A_k such that $P^k = A_k M^k$, $k = 0, \ldots, n$, and therefore

$$L_n f = \sum_{k=0} \lambda \, [\Xi_k] \, f^T \, P^k = \sum_{k=0} \lambda \, [\Xi_k] \, f^T \, A_k M^k = \sum_{k=0} \left(A_k^T \, \lambda \, [\Xi_k] \, f \right)^T \, M^k,$$

so that $\Delta \, [\Xi_k] \, f := A_k^T \, \lambda \, [\Xi_k] \, f$ gives the divided difference of order k:

$$L_n f = \Delta \, [\Xi_k] \, f^T \, X^k + q, \qquad q \in \Pi_{n-1}. \tag{16}$$

A simple recursive application of this argument immediately gives another Newton formula, namely

$$L_n f = \sum_{k=0}^{n} \Delta \, [\Xi_k] \, f^T \, M^k. \tag{17}$$

Taking into account (12) and the fact that $\lambda [\Xi_n] f = ((f - L_{n-1} f)(\xi_\alpha) \; : \; |\alpha| = n)$ we thus obtain for $|\alpha| = n$ the spline integral representation

$$\begin{aligned}\Delta [\Xi_n]_\alpha f &= \sum_{|\beta|=n} (A_k)_{\beta,\alpha} \sum_{\mu \in M_{n-1}} p_{\mu^{n-1}} (\xi_\beta) \pi_\mu [\Xi_\mu, \xi_\beta] f \\ &= \sum_{\mu \in M_n} (A_k)_{\mu^n,\alpha} \pi_\mu [\Xi_\mu] f = \sum_{\mu \in M_n} \frac{1}{\alpha!} \frac{\partial^n p_{\mu^n}}{\partial x^\alpha} \pi_\mu [\Xi_\mu] f,\end{aligned}$$

that is,

$$\Delta [\Xi_n] f = D^n \sum_{\mu \in M_n} p_{\mu^n} \pi_\mu [\Xi_\mu] f, \qquad D^n = \left[\frac{1}{\alpha!} \frac{\partial^n}{\partial x^\alpha} \; : \; |\alpha| = n \right]. \tag{18}$$

This, of course, is mainly a starting point for a study of the divided difference, understood as leading homogeneous coefficients of interpolation polynomials; at this point, however, we will not pursue this issue any further and just refer to [62], where the interested reader can find recurrences, for example.

5. The Algebraic Approach

This will become the longest part of this article and the one that contains the largest amount of new material, thus also the largest amount of proofs. While so far all we considered dealt exclusively with the vector space structure of polynomials by regarding interpolation as a linear algebra problem, we will now make more and more use of the ring structure of polynomials, taking into account multiplication and ideals.

The starting point here is the question of finding appropriate interpolation spaces for a given set Θ of functionals; though I will try to keep the presentation as general as possible here, it is not wrong to think of Lagrange interpolation problems, but definitely Θ is now supposed to be an ideal interpolation problem, i.e., $\ker \Theta$ is an ideal in $\Pi = \mathbb{K}[x]$. Of course, as mentioned before, there are still many subspaces $\mathcal{P}$ of Π such that the interpolation problem with respect to Θ is poised for $\mathcal{P}$. In order to find "good" interpolation spaces among this manifold, it seems to make very good sense to keep the degree low ...

5.1. Degree reducing interpolation

First, recall that polynomials of high degree are not very desirable, especially not from a numerical point of view: the memory consumption for storing them is higher, they tend to be more critical and unstable in all computations and the higher the degree, the more the polynomials feel free to follow their desire to oscillate. The first requirement to be made on a reasonable interpolation $\mathcal{P}$ space is that it should be of *minimal degree*, i.e., that any other interpolation space $\mathcal{Q}$ should satisfy $\deg \mathcal{Q} \geq \deg \mathcal{P}$, where $\deg \mathcal{P} := \max \{\deg p \; : \; p \in \mathcal{P}\}$. Spaces of minimal degree, however, can still provide unwanted effects as the following simple example shows: the Lagrange interpolation problem with respect to $\Xi = \{(1,0),(0,1)\}$ is poised for $\mathcal{P} = \operatorname{span}\{x,y\}$, and since $\deg \mathcal{P} = 1$, this is an interpolation space of minimal degree. On the other hand, when interpolating the constant polynomial that has degree 0, one obtains the polynomial $x+y$ which is obviously of degree 1. In this sense, the interpolation operator $L_{\mathcal{P}} = L_{\mathcal{P},\Theta}$ with respect to Θ and $\mathcal{P}$, defined as

$$\Theta L_{\mathcal{P}} f = \Theta f, \qquad f \in \Pi,$$

does not behave "well" for polynomials. In response to this undesirable phenomenon, we will restrict ourselves to interpolation spaces that treat the degree of the polynomial to be interpolated in a more respectful way.

Definition 8. A subspace $\mathcal{P}$ of Π is called a *degree reducing interpolation space* with respect to $\Theta \subset \Pi'$ if the interpolation problem with respect to Θ is poised for $\mathcal{P}$ and if $\deg L_{\mathcal{P}} f \leq \deg f$ for any $f \in \Pi$.

Degree reduction is a key property of the *least interpolant* introduced by de Boor and Ron in [10] and further elaborated in [12,13]. Motivated by their investigations which also included more general (total) degree reducing interpolation spaces sharing some (but not all) of the properties of the least interpolant as well as a comparison between minimal degree and degree reducing interpolation, a "direct" approach to degree reducing interpolation was pursued later in [66], deriving some properties directly from the requirement of degree reduction.

Degree reduction is really a stronger property than being of minimal degree. Let us formally state this and recall one simple proof for the reader's convenience, and keep in mind that the above example shows that the inclusion is indeed a strict one.

Proposition 9. *Any degree reducing interpolation space $\mathcal{P}$ is also of minimal degree.*

Proof. Suppose that $\mathcal{Q} \subset \mathcal{P}$ were an interpolation space with $\deg \mathcal{Q} < \deg \mathcal{P}$ and let $\ell_\theta \in \mathcal{Q}$, $\theta \in \Theta$, denote the fundamental polynomials: $\theta \ell_{\theta'} = \delta_{\theta,\theta'}$, $\theta, \theta' \in \Theta$. Then $\{L_{\mathcal{P}} \ell_\theta \,:\, \theta \in \Theta\}$ is a basis of $\mathcal{P}$ and by degree reduction it follows that $\deg L_{\mathcal{P}} \ell_\theta \leq \deg \ell_\theta \leq \deg \mathcal{Q} < \deg \mathcal{P}$, which is a contradiction since $\deg \mathcal{P} \geq \max_\theta L_{\mathcal{P}} \ell_\theta$. $\square$

So far we have always made the implicit assumption that **the** degree of a multivariate polynomial is its total degree; this is natural from the analytic and even the numerical point of view (as [59] shows, for example, that the roundoff error in evaluating a polynomial depends linearly on the total degree), but in no way the only possible choice. Quite the contrary, the notion of *multidegree* based on term orders, cf. [24], plays an important role in Computer Algebra ever since the introduction [15] of Gröbner bases. Moreover, it will turn out that there is a very close relationship between degree reducing interpolation and good ideal bases which persists for a much more general concept of degree and therefore it will be worthwhile to assume a more generous point of view of to what degree we consider something the degree of a polynomial.

5.2. Grading rings

The most natural generalization of the concept of degree, not only to multivariate polynomials, but to arbitrary rings, is that of a *graded ring*. To comprehend the appealing and intuitive idea behind this concept, it is instructive to begin with a look at the total degree. First, note that the degree of a polynomial essentially depends on a *homogeneous* polynomial, namely its leading part. Indeed, if $f \in \Pi_n$ is written as

$$f(x) = \sum_{|\alpha|=n} f_\alpha \, x^\alpha + \sum_{|\alpha|<n} f_\alpha \, x^\alpha =: f_k(x) + f_{<n}(x), \qquad x \in \mathbb{K}^d,$$

then the polynomial $f_{<k}$ does not influence the degree of f which depends entirely of the homogeneous part f_k. Moreover, adding two homogeneous polynomials of the same degree either gives the zero polynomial or another homogeneous polynomial of *precisely* the same (total) degree, and, finally, the multiplication of two homogeneous polynomials (and thus of two arbitrary polynomials) corresponds to adding the degrees. In other words: the degree should be an additive structure that defines "homogeneous" polynomials and the multiplication of such objects should correspond to the addition of degrees.

And this is precisely the concept of a graded ring. The additive structure is a semigroup Γ with neutral element 0, that is, a *monoid*, and the structure of Γ is connected to our ring $\Pi = \mathbb{K}[x]$ via the *direct sum decomposition*

$$\Pi = \bigoplus_{\gamma \in \Gamma} \Pi_\gamma^0, \qquad \text{i.e.,} \qquad f = \sum_{\gamma \in \Gamma} f_\gamma, \quad f \in \Pi, \, f_\gamma \in \Pi_\gamma^0, \tag{19}$$

where each Π_γ^0 is an abelian group under addition (one can add and subtract homogeneous polynomials) and multiplication of homogeneous polynomials is reflected by addition in Γ:

$$\Pi_\gamma^0 \cdot \Pi_{\gamma'}^0 \subseteq \Pi_{\gamma+\gamma'}^0, \qquad \gamma, \gamma' \in \Gamma. \tag{20}$$

In fact, we can even retreat to more familiar concepts in the case of Π: the homogenous spaces Π_γ^0 must be $\mathbb{K}$ vector spaces. This follows by first noticing that (20) yields that $1 \in \Pi_0^0$ and then realizing that any *unit* in $f \in \Pi^*$, i.e., any element such that $f^{-1} \in \Pi$, must have degree 0 again, as follows from considering

$$\Pi_\gamma^0 = 1 \cdot \Pi_\gamma^0 = \left(f \, f^{-1}\right) \Pi_\gamma^0 = \sum_{\eta, \eta' \in \Gamma} f_\eta \, f_{\eta'}^{-1} \Pi_{\gamma+\eta+\eta'}^0.$$

It is worthwhile to mention that this is a general property of graded rings: units have degree zero. As innocent and natural as this sounds, it has the somewhat surprising consequence that Laurent polynomials (which are analytically very similar to polynomials, but are spanned by units – all the monomials are invertible) only admit the *trivial grading* $\Gamma = \{0\}$; this makes the construction of Gröbner- and H–bases more intricate, cf. [57].

We also denote by

$$\Pi_\gamma := \bigoplus_{\eta \le \gamma} \Pi^0_\gamma, \qquad \gamma \in \Gamma,$$

the vector space of all polynomials of (Γ–)degree at most γ. Note that, depending on the underlying grading, these spaces can be finite dimensional (like with the grading by total degree) or infinite dimensional (the grading by a purely lexicographic term order serves as the standard example here).

To step forward from grading to degree, we need some additional structure of Γ, namely a *well–ordering* "$<$" that is compatible with the semigroup structure. Recall that "well–ordering" means that any strictly descending chain $\gamma > \gamma' > \cdots$ must be finite, i.e., any subset of Γ has a smallest element and that "compatibility" is equivalent to $\gamma < \gamma'$ implying $\gamma + \gamma'' < \gamma' + \gamma''$ for all $\gamma, \gamma', \gamma'' \in \Gamma$. It is not hard to guess, and not much harder to prove, that this property implies $0 \le \Gamma$. To any polynomial $f \in \Pi$ which can be written as a *finite* sum (19) of its homogeneous components, we can now associate the *degree*

$$\delta(f) := \max \{\gamma \in \Gamma \ : \ f_\gamma \neq 0\}, \qquad \delta : \Pi \to \Gamma,$$

and the *leading term*

$$\Lambda(f) := f_{\delta(f)} \neq 0, \qquad \Lambda : \Pi \to \Pi^0 := \bigcup_{\gamma \in \Gamma} \Pi^0_\gamma,$$

which leaves, as usually $\delta(0)$ and $\Lambda(0)$ undefined.

So the degree of a polynomial now depends on two ingredients: the monoid Γ and the well ordering "$<$". In the case of total degree, the monoid is $\mathbb{N}_0$, the well–ordering is the canonical one and the homogenous spaces are

$$\Pi^0_n = \operatorname{span}_{\mathbb{K}} \{x^\alpha \ : \ |\alpha| = n\}, \qquad k \in \mathbb{N}_0.$$

A well–ordering on $\mathbb{N}_0^d$ is called a *term order* where classical examples are the *lexicographic* term order

$$\alpha \prec_\ell \beta \qquad \Leftrightarrow \qquad \alpha_j = \beta_j,\ j = 1, \dots, k-1, \quad \alpha_k < \beta_k,$$

and the *graded lexicographical* term order

$$\alpha \prec_g \beta \qquad \Leftrightarrow \qquad |\alpha| < |\beta| \text{ or } |\alpha| = |\beta|,\ \alpha \prec_\ell \beta.$$

In fact, the graded lexicographical term order induces a grading that can be seen as a refinement of the grading by total degree. In the case of term orders, we have "homogeneous" spaces

$$\Pi^0_\alpha = \operatorname{span}_{\mathbb{K}} \{x^\alpha\}, \qquad \alpha \in \mathbb{N}_0^d,$$

which are of dimension 1. Some more esoteric gradings can be defined by $\Gamma = \mathbb{N}_0$ and

$$\Pi^0_k := \operatorname{span}_{\mathbb{K}} \left\{x^\alpha \ : \ \omega^T \alpha = k\right\}, \qquad k \in \mathbb{N}_0,\ 0 \neq \omega \in \mathbb{N}_0^d, \tag{21}$$

or, driving this idea even further, $\Gamma = \mathbb{N}_0^m$, equipped with some term order, and

$$\Pi_\beta := \operatorname{span}_{\mathbb{K}} \{x^\alpha \ : \ M\alpha = \beta\}, \qquad \beta \in \mathbb{N}_0^m, 0 \neq M \in \mathbb{N}_0^{m\times d}.$$

The latter, however, is much more than academic eccentricity and leads to the concepts like the "Gröbner fan" and the "Gröbner walk", see [63] as a starting point. Finally, we single out two classes of gradings which will be useful in the sequel.

Definition 10. A grading is called *strict* if $\Pi_0^0 = \mathbb{K}$ and *monomial* if all homogeneous spaces Π_γ^0 are spanned by monomials.

Recall that we always have $\mathbb{K} \subseteq \Pi_0^0$ and so strict gradings are those where *exactly* the constants have degree zero; an example for a non–strict grading is obtained by setting

$$\Gamma = \mathbb{N}_0, \qquad \Pi_k^0 = \operatorname{span}_{\mathbb{K}} \{x^\alpha \ : \ \alpha_1 = k\}, \quad k \in \mathbb{N}_0, \tag{22}$$

choosing $\omega = (1, 0, \ldots, 0)$ in (21). This yields a grading that only cares for the first variable x_1. Non–monomial gradings are a little bit more obscure and not really to be found in the literature, but nevertheless do exist. It has to be noted, however, that all the "standard" gradings are strict monomial ones.

With this terminology at hand, we can literally repeat the definition of degree reducing interpolation. To that end, we introduce the shorthand notation

$$f \preceq g \qquad \Leftrightarrow \qquad \delta(f) \leq \delta(g), \tag{23}$$

which determines a well–ordering on Π.

Definition 11. A subspace $\mathcal{P} \subset \Pi$ is called a *degree reducing interpolation space* for $\Theta \subset \Pi'$ if the interpolation problem with respect to Θ is poised for $\mathcal{P}$ and if the interpolation operator satisfies $L_{\mathcal{P}} f \preceq f$, $f \in \Pi$.

Since the proof of Proposition 9 works for any arbitrary grading, it follows immediately that degree reducing interpolation spaces are always of minimal degree, i.e., $\mathcal{P} \preceq \mathcal{Q}$ for any interpolation space $\mathcal{Q} \subset \Pi$. total

5.3. A good basis for idealism

Finally, now arises the time when ideals really matter. An *ideal* $\mathcal{I} \subseteq \Pi$ is a subset, in fact a vector subspace, of Π that is closed under addition, $\mathcal{I} + \mathcal{I} = \mathcal{I}$, and closed under multiplication by arbitrary polynomials, $\mathcal{I} \cdot \Pi = \mathcal{I}$. Consequently, if $F \subseteq \mathcal{I}$ then also $\langle F \rangle \subseteq \mathcal{I}$, where

$$\langle F \rangle := \left\{ \sum_{f\in F} g_f \, f \ : \ g_f \in \Pi \right\},$$

is the *ideal generated by* F. A set $F \subset \Pi$ is called a *basis* of $\mathcal{I}$ if $\mathcal{I} = \langle F \rangle$. Recall that Hilbert's famous *Basissatz* [35] states that every polynomial ideal, i.e., any ideal in Π, has a *finite* basis. This is the background and the theoretical justification for the way in which polynomial ideals are represented in computer algebra, namely as finite (unordered) sets of polynomials that generate the ideal. Note, however, that though the word "basis" is used, there is in general nothing like independence or uniqueness of representation as one would expect from linear algebra – in this respect an ideal basis is only a generating system for the ideal.

A finite set G is called a Γ–basis for the ideal $\langle G \rangle$ if

$$f \in \langle G \rangle \qquad \Leftrightarrow \qquad f = \sum_{g \in G} f_g \, g, \quad f_g \, g \preceq f, \, g \in G. \tag{24}$$

The representation of f on the right hand side of (24) – relevant here is the degree restriction on the terms in the sum – is called a *Γ–representation* and can be understood as a *non–redundant* representation: if there were some g such that $f_g \, g \succ f$, then all terms in $f_g \, g$ whose degree exceeds that of f must be canceled by the rest of the sum, thus is redundant and therefore should better not appear from the beginning. As pointed out in [70], Γ–bases exist for any grading Γ and can be constructed (more or less) efficiently. At least two instances of Γ–bases are classical: when the underlying grading is the one by total degree, they become *H–bases*, due to Macaulay [46], in the case of term orders, on the other hand, they are *Gröbner bases.* These two notions are not mutually exclusive as there exist Gröbner bases that are H–bases as well, but there are also H–bases which are no Gröbner bases and vice versa.

Next, there is another characterization of Γ–bases that will become useful in the sequel. To that end, we need the notion of the *homogeneous ideal*

$$\langle F \rangle^0 = \bigcup_{\gamma \in \Gamma} \left\{ \sum_{f \in F} p_f \, f \ : \ p_f \in \Pi_{\gamma - \delta(f)} \right\} \subset \Pi^0$$

generated by a finite set $F \subset \Pi^0$ of *homogeneous* polynomials. In the case of term orders, the homogeneous ideals are *monomial ideals* which are the essence of Dickson's Lemma, cf. [24]. Homogeneous ideals in connection with dehomogeneization, on the other hand, play a fundamental role in the original definition of H–bases, cf. [34]. The characterization of Γ–bases in terms of homogeneous ideals is now as follows.

Proposition 12. *A set $G \subset \Pi$ is a Γ–basis of an ideal $\mathcal{I} \subsetneq \Pi$ if and only if $\Lambda(\mathcal{I}) = \langle \Lambda(G) \rangle^0$.*

Proof. If G is a Γ–basis then define, for a Γ–representation (24), $G^* \subset G$ as $G^* = \{ g \in G \ : \ \delta(f_g \, g) = \delta(f) \}$ and obtain

$$\Lambda(f) = \sum_{g \in G^*} \Lambda(f_g \, g) = \sum_{g \in G^*} \Lambda(f_g) \, \Lambda(g) \in \langle \Lambda(G^*) \rangle^0 \subseteq \langle \Lambda(G) \rangle^0 .$$

Conversely, if $\Lambda(\mathcal{I}) = \langle \Lambda(G) \rangle^0$ then we represent, for $f \in \mathcal{I}$,

$$\Lambda(f) = \sum_{g \in G} f_g \, \Lambda(g), \qquad f_g \in \Pi^0_{\delta(f)-\delta(g)}, \, g \in G,$$

and replace f by

$$f' = f - \sum_{g \in G} f_g \, g, \qquad \text{i.e.} \qquad f' \prec f.$$

Since "<" is a well–ordering on Γ we can rest assured that finitely many applications of this process lead to a Γ–representation of f with respect to G. □

The main computational advantage of Γ–bases, however, and as stated in Buchberger's thesis [14], the main intention of Gröbner when addressing the research topic for Buchberger's thesis, rests in the fact that they enable us to efficiently *compute* remainders modulo an ideal. As a first step towards the notion of a *reduced* polynomial, we define for $F \subset \Pi$ and $\gamma \in \Gamma$ the vector space

$$V_\gamma(F) := \langle \Lambda(F) \rangle^0 \cap \Pi^0_\gamma.$$

Moreover, let $(\cdot,\cdot) : \Pi \to \mathbb{K}$ denote an inner product on Π such that $\left(\Pi^0_\gamma, \Pi^0_{\gamma'}\right) = 0$ for $\gamma \neq \gamma' \in \Gamma$ and write $f \perp g$ if $(f,g) = 0$. If Γ is a monomial grading, then canonical choices for

$$f(x) = \sum_{\alpha \in \mathbb{N}^d_0} f_\alpha \, x^\alpha, \qquad g(x) = \sum_{\alpha \in \mathbb{N}^d_0} g_\alpha \, x^\alpha,$$

are

$$(f,g) = \sum_{\alpha \in \mathbb{N}^d_0} f_\alpha \, g_\alpha \qquad \text{or} \qquad (f,g) = (f(D)g)(0) = \sum_{\alpha \in \mathbb{N}^d_0} \alpha! \, f_\alpha \, g_\alpha.$$

The latter inner product can be found in the literature under various names, but in the context of polynomial interpolation it has been established by de Boor and Ron [10–13] as a useful tool to characterize the *least interpolant.* Also observe that in the case of a term order the inner product is irrelevant as the orthogonality between monomial spaces of different degree already defines it up to the positive constants (x^α, x^α), $\alpha \in \mathbb{N}^d_0$.

A polynomial $g \in \Pi$ is called *reduced* with respect to $F \subset \Pi$ if any of its homogeneous components g_γ is in the orthogonal complement of the respective $V_\gamma(F)$, that is

$$g \in W(F) := \bigoplus_{\gamma \in \Gamma} W_\gamma(F), \qquad \Pi^0_\gamma = V_\gamma(F) \oplus W_\gamma(F), \quad V_\gamma(F) \perp W_\gamma(F). \tag{25}$$

To illustrate what this property means, we have a brief look at the univariate case. Since univariate polynomials form a principal ideal ring, we can assume that $F = \{f\}$ and get that $V_k(F)$ is either $\mathbb{K} \cdot (\cdot)^k$ if $k \geq \deg f$ or $\{0\}$ otherwise. Consequently, a *univariate* polynomial is reduced with respect to f, $\deg f = k$, if and only if it belongs to $\Pi_{k-1} \simeq \Pi / \langle F \rangle$.

A polynomial $r \in \Pi$ is called a ***remainder of division*** of f with respect to $G \subset \Pi$ if it is a reduced remainder of a Γ–representation of f, that is,

$$f = \sum_{g \in G} f_g \, g + r, \qquad f_g \, g \preceq f, \quad r \in W(G). \tag{26}$$

The representation in (26) can be computed more or less efficiently by means of *reduction*: suppose that $\delta(f) = \gamma$, compute the orthogonal projection of $\Lambda(f)$ to $V_\gamma(G)$, yielding

$$\Lambda(f) = \sum_{g \in G} f_{g,\gamma} \, \Lambda(g) + r_\gamma, \qquad r_\gamma \in W_\gamma(G), \tag{27}$$

and continue this process with

$$f' := f - \sum_{g \in G} f_{g,\gamma} \, g - r_\gamma, \qquad \text{hence,} \qquad f' \prec f. \tag{28}$$

This process terminates after finitely many steps and yields the terms in (26) as

$$f_g = \sum_{\gamma \le \delta(f)} f_{g,\gamma}, \quad g \in G, \qquad r = \sum_{\gamma \le \delta(f)} r_\gamma.$$

The use of orthogonal projections, introduced in [68] for the grading by total degree, provides the "gradual" notion of divisibility needed for non–term–order gradings: while a monomial x^α either does divide another monomial x^β or is does not, homogeneous terms are much more indifferent towards each other in terms of divisibility. For example, the bivariate polynomial $x^2 + y^2$ does not divide $x^3 + y^3$, but nevertheless they have something "in common" as well. From a computational point of view, it should be clear that division of monomials is definitely much simpler than to determine orthogonal projection though the latter is "only" a problem from linear algebra; an efficient method to compute this projection for the inner product $f(D)g(0)$ has been given recently in [61]. On the other hand, H–bases can be very helpful in avoiding ***representation singularities*** in polynomial system solving, cf. [56].

To finish the "short course" on Γ–bases, note that the reduction algorithm looks very similar to what was done in the proof of Proposition 12, so Γ–bases can be expected to play a fundamental role in reduction. In fact, they turn out to be absolutely necessary for reduction to work at all. This is due to the subtle fact that the representation of $\Lambda(f)$ in (27) need not be unique and, in general, will not be because of the ubiquitous syzygies. And while this does not affect the leading term $r_{\delta(f)}$ of the remainder, different choices of this representation can lead to different polynomials f' in (28) which may very well affect lower order terms of r. Consequently, the remainder of division can depend on how the orthogonal projections are expanded – except in the case when we divide by a Γ–basis. Specifically, we have the following result, cf. [70].

Theorem 13. *A finite set $G \subset \Pi$ is a Γ–basis if and only if for any $f \in \Pi$ and any pair of Γ–representations*

$$f = \sum_{g \in G} f_g \, g + r = \sum_{g \in G} f_g' \, g + r', \qquad r, r' \in W(G),$$

one has $r = r'$.

In other words, if G is a Γ–basis, we can speak of **the** remainder of the division of f by G, or of the *normal form* $\nu_G(f) := r$ of f modulo G. This immediately implies that $\nu_G(f) = 0$ if and only $f \in \langle G \rangle$ giving an algorithmic criterion to decide the *ideal membership problem.* But Γ–bases offer even more: the normal form depends entirely on the ideal (and of course on the grading and the inner product), but not on the actual Γ–basis!

Proposition 14. *If $G, G' \subset \Pi$ are two Γ–bases for the same ideal $\langle G \rangle = \langle G' \rangle$ then $\nu_G = \nu_{G'}$.*

Proof. We expand $G' \ni g' = \sum_g p_{g',g} \, g$ in terms of G and substitute this into a Γ–representation of $f \in \Pi$ with respect to G',

$$f = \sum_{g' \in G'} f_{g'} \, g' + r' = \sum_{g \in G} \left(\sum_{g' \in G'} f_{g'} p_{g',g} \right) g + r',$$

and a comparison of degrees yields that the term on the right hand side is a Γ–representation of f with respect to G. Consequently, Theorem 13 implies that $r' = r$. $\square$

5.4. What remains is interpolation

Proposition 14 states that for a given grading and a given inner product (which can be seen as a part of the grading), there is a *unique* normal form $\nu_{\mathcal{I}}(f)$ for a polynomial f modulo an ideal $\mathcal{I}$, and the polynomial $\nu_{\mathcal{I}}(f)$ is the canonical representer of the equivalence class $[f] = f + \mathcal{I}$ in $\Pi/\mathcal{I}$. This finally returns us to interpolation as it says that normal forms are natural interpolants for ideal interpolation schemes.

Theorem 15. *If $\Theta \subset \Pi'$ has the property that $\mathcal{I}(\Theta) := \ker \Theta$ is an ideal in Π, then $\mathcal{P}^* := \nu_{\mathcal{I}(\Theta)}(\Pi)$ is a degree reducing interpolation space for Θ and the interpolation operator takes the form $L_{\mathcal{P}^*} = \nu_{\mathcal{I}(\Theta)}$.*

Proof. Since $f - \nu_{\mathcal{I}(\Theta)}(f) \in \mathcal{I}(\Theta)$, the mapping $\nu_{\mathcal{I}(\Theta)}$ is an interpolation operator whose linearity follows from the uniqueness of remainders and the fact that $f + f'$ has the Γ–representation

$$f+f'=\sum_{g\in G} h_g\, g+r+r', \qquad r+r'\in W(G), \quad h_g\in\Pi.$$

Degree reduction of the normal form operator, on the other hand, is seen directly from the reduction algorithm. $\square$

Hence, normal forms are canonical degree reducing interpolants and $\mathcal{P}^*=\nu_{\mathcal{I}(\Theta)}(\Pi)$ is the canonical degree reducing interpolation space. Moreover, reduction even allows us to compute an interpolant to a polynomial f if the interpolation problem is only known in a *implicit* way, namely in terms of a basis of $\mathcal{I}(\Theta)$. On the other hand, in the case of a Lagrange interpolation problem, that is, if a finite node set $\Xi\subset\mathbb{K}^d$ is given, we can easily give the Lagrange fundamental polynomials *explicitly*. To that end, let $v\in\mathbb{K}^d$ be any vector such that $v^T(\xi-\xi')\neq 0$, $\xi,\xi'\in\Xi$ – such vectors exist in abundance, in fact there are only finitely many (normalized) vectors that do not have this property. Now simply define for $\xi\in\Xi$ the polynomials

$$\widehat{\ell}_\xi := \prod_{\xi'\in\Xi\setminus\{\xi\}} \frac{v^T(\cdot-\xi')}{v^T(\xi-\xi')}, \qquad \ell_\xi:=\nu_{\mathcal{I}(\Xi)}(\widehat{\ell}_\xi), \tag{29}$$

and ℓ_ξ, $\xi\in\Xi$, is the Lagrange basis of $\mathcal{P}^*=\nu_{\mathcal{I}(\Xi)}(\Pi)$, so that the interpolant takes the form

$$L_{\mathcal{P}^*}f=\sum_{\xi\in\Xi} f(\xi)\,\nu_{\mathcal{I}(\Xi)}\left(\prod_{\xi'\in\Xi\setminus\{\xi\}} \frac{v^T(\cdot-\xi')}{v^T(\xi-\xi')}\right).$$

But this tells us even more: the basis is obtained by taking the normal forms of polynomials of total degree $\#\Xi-1$, leading to the following observation.

Corollary 16. *For any finite node set $\Xi\subset\mathbb{K}^d$ the normal form interpolation space $\mathcal{P}^*$ satisfies $\mathcal{P}^*=\nu_{\mathcal{I}(\Xi)}(\Pi_{\#\Xi-1})$.*

In other words, the minimal degree interpolation space is spanned by the polynomials $p_\alpha=\nu_{\mathcal{I}(\Xi)}((\cdot)^\alpha)$ and the finite matrix

$$P(\Xi)=[p_\alpha(\xi)\ :\ \xi\in\Xi,\ |\alpha|\le \#\Xi-1]$$

has the maximal rank $\#\Xi$. In contrast to the approach in [9], where infinite matrices are considered, the linear system associated to a Lagrange interpolation problem can therefore always be assumed to be finite, even if the number of polynomials exceeds the number of points significantly.

Normal forms are closely tied to the *Hilbert function* of the ideal $\mathcal{I}(\Xi)$. Though the Hilbert function of an ideal is usually defined and considered for the homogeneous grading, cf. [34,54,55], it can be extended in quite a straightforward way to arbitrary gradings, which we will describe next. For an ideal $\mathcal{I}\subset\Pi$ we set $W_\gamma(\mathcal{I})=W_\gamma(G)$, G any Γ–basis of $\mathcal{I}$, and define the *homogeneous* Hilbert function of $\mathcal{I}$ as

$$H^0_{\mathcal{I}}(\gamma)=\dim W_\gamma(\mathcal{I})=\dim\Pi^0_\gamma-\dim\left(\Lambda(\mathcal{I})\cap\Pi^0_\gamma\right), \qquad \gamma\in\Gamma, \tag{30}$$

Moreover, the "summarized" function

$$H_{\mathcal{I}}(\gamma) = \sum_{\gamma' \leq \gamma} H_{\mathcal{I}}^0(\gamma'), \qquad \gamma \in \Gamma,$$

is called the *affine* Hilbert function of $\mathcal{I}$. The latter one is monotonically increasing, i.e., $H_{\mathcal{I}}(\gamma) \leq H_{\mathcal{I}}(\gamma')$ if $\gamma \leq \gamma'$. Also note that, strictly speaking the second identity in (30) is only valid if Π_γ^0 is of finite dimension which is the case, for example, for the grading by total degree, but not for the one from (22). The affine Hilbert function describes the dimensions of the nested interpolation spaces $\mathcal{P}_\gamma^* = \mathcal{P}^* \cap \Pi_\gamma$, $\gamma \in \Gamma$.

Proposition 17. *For $\gamma \in \Gamma$ we have that*

$$H_{\mathcal{I}(\Xi)}^0(\gamma) = \dim\left(\mathcal{P}^* \cap \Pi_\gamma^0\right), \qquad H_{\mathcal{I}(\Xi)}(\gamma) = \dim \mathcal{P}_\gamma^*. \tag{31}$$

Proof. We first note that $\mathcal{P}^* = \nu_{\mathcal{I}(\Xi)}(\Pi)$ is spanned by homogeneous polynomials: any polynomial $f \in W_\gamma(\mathcal{I}) \subset \Pi_\gamma^0$ has the property that $f = \nu_{\mathcal{I}(\Xi)}(f)$. Hence $H_{\mathcal{I}(\Xi)}^0(\gamma) \leq \dim\left(\mathcal{P} \cap \Pi_\gamma^0\right)$. On the other hand, any homogeneous $p \in \mathcal{P}^* \cap \Pi_\gamma^0$ also satisfies $p = \nu_{\mathcal{I}(\Xi)}(p) \in W_\gamma(\mathcal{I})$ and therefore $H_{\mathcal{I}(\Xi)}^0(\gamma) \geq \dim\left(\mathcal{P} \cap \Pi_\gamma^0\right)$ as well. $\square$

The understanding obtained so far of degree reducing interpolation for graded rings and the prominent role that normal forms play, finally allows us to define the notion of a *Newton basis* for arbitrary gradings, and even to construct such bases. To that end, we start off with the *homogeneous bases* $P_\gamma^0 \subset \mathcal{P}^*$, $\gamma \leq \delta(\mathcal{P}^*)$, which are defined by the requirement that

$$\operatorname{span} P_\gamma^0 = \mathcal{P}_\gamma^* \cap \Pi_\gamma^0 = W_\gamma(\mathcal{I}(\Xi)). \tag{32}$$

Since $\dim \mathcal{P}^* = \#\Xi < \infty$ there is a finite set $\Gamma^* \subset \Gamma$ such that P_γ^0, $\gamma \in \Gamma^*$, still form a homogeneous basis for $\mathcal{P}^*$. Moreover, we can assume that $0 \in \Gamma^*$ as $W_0(\mathcal{I}(\Xi)) = \{0\}$ implies that $V_0(\mathcal{I}(\Xi)) = \Pi_0^0$ and thus $\mathcal{I}(\Xi) = \Pi$, i.e., $\Xi = \emptyset$.

The construction of the Newton basis proceeds as follows: beginning with $0 = \min \Gamma^*$ and taking into account that the $P_0^0(\Xi)$ has rank $H_{\mathcal{I}(\Xi)}(0)$, we can find $\Xi_0 \subset \Xi$ such that $\det P_0^0(\Xi_0) \neq 0$, and so $N_0 := P_0^0(\Xi_0)^{-1} P_0^0$ satisfies $N_0(\Xi_0) = I$.

Suppose now that for the k smallest elements $0 = \gamma^0 < \cdots < \gamma^{k-1}$ of Γ^* we have constructed $\Xi_0, \ldots, \Xi_{\gamma^{k-1}}$ such that

$$\det P_{k-1}(\Xi_{k-1}) \neq 0, \qquad P_{k-1} := \bigcup_{j=0}^{k-1} N_{\gamma^j}, \qquad \Xi_{k-1} := \bigcup_{j=0}^{k-1} \Xi_{\gamma^j},$$

and set $\gamma = \gamma^k = \min\left\{\gamma' \in \Gamma \,:\, \gamma' > \gamma^{k-1}\right\}$. Then the polynomials in the vector

$$P' := P_\gamma^0 - P_\gamma^0(\Xi_{k-1})^T P_{k-1}(\Xi_{k-1})^{-T} P_{k-1}$$

are linearly independent, vanish on Ξ_{k-1}, and their leading terms $\Lambda(P')$ span $W_\gamma(\mathcal{I}(\Xi))$. In addition, the rank of the matrix $P'(\Xi \setminus \Xi_{k-1})$ coincides with the cardinality of P'. Consequently, there is a subset Ξ_γ of $\Xi \setminus \Xi_{k-1}$ such that $\det P'(\Xi_\gamma) \neq$

0, and once we define the polynomial vector $N_\gamma = P'(\Xi_\gamma)^{-1} P'$ it follows that $N_\gamma(\Xi_\gamma) = I$. This process, first used in [16] in the case of monomial gradings based on term orders, is nothing but the Γ–equivalent of *Gauß elimination by segments*, see [6], or the Gram–Schmidt process in [66], and creates a decomposition of Ξ into Ξ_γ and a Newton basis N_γ, $\gamma \in \Gamma^*$, such that

$$N_\gamma(\Xi_{\gamma'}) = \delta_{\gamma',\gamma} I, \quad \gamma' \leq \gamma, \qquad \operatorname{span} \Lambda(N_\gamma) = \operatorname{span} \Lambda\left(P_\gamma^0\right) = W_\gamma(\mathcal{I}(\Xi)).$$

Once more the normal form space $\mathcal{P}^*$ serves as a prototype for degree reducing interpolation spaces in the sense that we can even *characterize* degree reducing interpolation in terms of the existence of a Newton basis. For the grading by total degree, this result has been given in [66], the presentation here, however, is strongly influenced by discussions with Carl de Boor on [67] and his comments made in [9].

Theorem 18. *A polynomial subspace $\mathcal{P} \subset \Pi$ is a degree reducing interpolation space for a finite node set Ξ if and only if there exists a* Newton basis $N_\gamma \subset \mathcal{P}$, *$\gamma \in \Gamma^* \subset \Gamma$, and a decomposition of Ξ into Ξ_γ, $\gamma \in \Gamma^*$, such that*

$$N_\gamma(\Xi_{\gamma'}) = \delta_{\gamma,\gamma'}, \quad \gamma' \leq \gamma, \qquad \Pi_\gamma^0 = \operatorname{span} \Lambda(N_\gamma) + V_\gamma(\mathcal{I}(\Xi)), \quad \gamma \in \Gamma. \tag{33}$$

Proof. Suppose that $\mathcal{P}$ is a degree reducing interpolation space. Together with $L_\mathcal{P}\mathcal{I}(\Xi) = \nu_{\mathcal{I}(\Xi)}(\mathcal{I}(\Xi)) = 0$ and Theorem 15 this implies for any $f \in \Pi$ that

$$\delta(L_\mathcal{P} f) = \delta\left(L_\mathcal{P}\nu_{\mathcal{I}(\Xi)}(f)\right) \leq \delta\left(\nu_{\mathcal{I}(\Xi)}(f)\right) = \delta\left(\nu_{\mathcal{I}(\Xi)}(L_\mathcal{P} f)\right) \leq \delta(L_\mathcal{P} f),$$

i.e.,

$$\delta(L_\mathcal{P} f) = \delta\left(\nu_{\mathcal{I}(\Xi)}(f)\right) = \delta(L_{\mathcal{P}^*} f). \tag{34}$$

Let N_γ^*, $\gamma \in \Gamma^*$, denote a Newton basis of $\mathcal{P}^*$, constructed as above, and set $N_\gamma = L_\mathcal{P} N_\gamma^*$, $\gamma \in \Gamma^*$. The interpolation property of $L_\mathcal{P}$ then yields that

$$N_\gamma(\Xi_{\gamma'}) = N_\gamma^*(\Xi_{\gamma'}) = \delta_{\gamma,\gamma'} I, \qquad \gamma' \leq \gamma.$$

Hence, for any $\xi \in \Xi_\gamma$ there exist $N_\xi \in N_\gamma$ and $N_\xi^* \in N_\gamma^*$ such that $N_\xi - N_\xi^* \in \mathcal{I}(\Xi)$ which is improved by (34) into

$$N_\xi - N_\xi^* \in \mathcal{I}(\Xi) \cap \Pi_\gamma \qquad \text{i.e.,} \qquad \Lambda(N_\xi) - \Lambda\left(N_\xi^*\right) \in V_\gamma(\mathcal{I}(\Xi)).$$

Since $\Pi_\gamma^0 = \operatorname{span} \Lambda\left(N_\gamma^*\right) + V_\gamma(\mathcal{I}(\Xi))$ for all $\gamma \in \Gamma$, this implies (33).

Conversely, suppose that $\mathcal{P}$ has a basis that satisfies (33), hence, $\mathcal{P}$ is an interpolation space. Moreover, the sets $N_\gamma^* \subset \mathcal{P}_\gamma^*$, $\gamma \in \Gamma^*$, defined by $N_\gamma^* = \nu_{I(\Xi)}(N_\gamma)$ also satisfy $N_\gamma^*(\Xi_{\gamma'}) = \delta_{\gamma,\gamma'} I$, $\gamma, \gamma' \in \Gamma^*$, and thus form a Newton basis for the normal form interpolation space $\mathcal{P}^*$. Now, for $f \in \Pi$,

$$L_{\mathcal{P}^*} f = \sum_{\gamma \in \Gamma^*} f_\gamma^T N_\gamma^* \qquad \Leftrightarrow \qquad L_\mathcal{P} f = \sum_{\gamma \in \Gamma^*} f_\gamma^T N_\gamma$$

and since the leading terms of both sets, N_γ and N_γ^*, each span Π_γ^0 jointly with $V_\gamma(\mathcal{I}(\Xi))$, they consist of $H_{\mathcal{I}(\Xi)}^0(\gamma)$ linearly independent elements, which leads to the conclusion that $\delta(L_\mathcal{P}(f)) = \delta(L_{\mathcal{P}^*}(f)) \leq \delta(f)$; in other words, $\mathcal{P}$ is a degree reducing interpolation space. $\square$

From the proof of Theorem 18 we can draw another interesting conclusion which further justifies the notation "Γ^*".

Corollary 19. *Let $\Xi \subset \mathbb{K}^d$. Then the index set γ^* for the Newton basis is the same for all degree reducing interpolation spaces and takes the form*

$$\Gamma^* = \left\{ \gamma \in \Gamma \ : \ H^0_{\mathcal{I}(\Xi)}(\gamma) \neq 0 \right\}.$$

Allright, the existence of a Newton basis characterizes degree reducing interpolation spaces and of course those will depend on the parameters involved with the grading, that is, the monoid with its total order and the inner product that is used for the orthogonal projection. We should first note that the normal form interpolant is singled out by requiring that its Newton basis makes the decomposition in (33) an *orthogonal* one. But even then there is in general a freedom of choice as different decompositions of Ξ into Ξ_γ, $\gamma \in \Gamma^*$, lead to different bases. This is, however, not a multivariate phenomenon as even the univariate Newton polynomials

$$\frac{(x-\xi_0)\cdots(x-\xi_{k-1})}{(\xi_k-\xi_0)\cdots(\xi_k-\xi_{k-1})}, \qquad k = 0,\dots,n, \qquad \Xi = \{\xi_0,\dots,\xi_n\},$$

depend on the ordering of the points in Ξ which need not be in any "natural", for example ascending, order.

It is clear that if N_γ is a Newton basis for an interpolation space $\mathcal{P}$, then so is

$$N'_\gamma = N_\gamma + \sum_{\gamma' \le \gamma} V_\gamma\left(\mathcal{I}(\Xi)\right), \qquad \gamma \in \Gamma^*,$$

that is, even after fixing the decomposition Ξ_γ, the elements of a Newton basis are only defined up to addition of an ideal element of at most the same degree – and this *is* a truly multivariate feature. On the other hand, the above modification is also the only way to pass from one Newton basis to another so that they are normal forms up to addition of ideal elements of restricted degree.

Corollary 20. *If $\delta(\mathcal{P}) < \min\{\delta(g) \ : \ g \in G\}$ for some Γ–basis G of $\mathcal{I}(\Xi)$ then the degree reducing interpolation space is unique and so is its Newton basis up to indexing of nodes.*

5.5. When Newton goes Γ ...

We close this section with a brief look at an appealing algebraic property of the Newton basis. To that end we fix $\gamma \in \Gamma^*$ and consider the node set

$$\Xi \supseteq \Xi' := \Xi'_\gamma = \bigcup_{\gamma' \le \gamma} \Xi_\gamma$$

and the associated ideal $\mathcal{I}(\Xi') \supseteq \mathcal{I}(\Xi)$. How can we obtain a Γ–basis G' for $\mathcal{I}(\Xi')$? By the inclusion relation between the ideals we can consider this as a basis extension problem, i.e., $G' = G + H$, $H \subset \Pi$, but what is H? Let $\mathcal{P}$ be a degree reducing interpolation space and consider $f \in \mathcal{I}(\Xi')$: since $f(\Xi') = 0$ the interpolant $L_{\mathcal{P}} f$ lies in the space spanned by $N_{\gamma'}$, $\gamma' > \gamma$, and so it follows that

$$f = L_{\mathcal{P}} f + q = \sum_{\gamma' > \gamma} f_\gamma^T N_\gamma + \sum_{g \in G} q_g \, g, \qquad q = \sum_{g \in G} q_g \, g \in \mathcal{I}(\Xi).$$

Since $L_{\mathcal{P}} f \preceq f$, all degrees appearing in the sums on the right hand side of this equation are $\leq \delta(f)$, hence this equation is in fact a Γ–representation. In other words: the set

$$G \cup \bigcup_{\gamma' > \gamma} N_\gamma$$

is a Γ–basis for $\mathcal{I}(\Xi')$. This observation is the motivation behind Conjecture 7. In order to make a Newton basis work as a supplement to a Γ–basis, it is necessary that one is dealing with an ideal structure on all levels: the decomposition of $\Theta \subset \Pi'$ into Θ_γ, $\gamma \in \Gamma^*$, must have the property that $\ker\left(\Theta'_\gamma\right)$ is an ideal for any $\gamma \in \Gamma^*$, where $\Theta'_\gamma = \bigcup_{\gamma' \leq \gamma} \Theta_{\gamma'}$. Whether or not such a decomposition exists is not known at the present moment, but it would definitely be useful in deriving remainder formulas, see [66,73,72]. Anyway, the generalization of Conjecture 7 is as follows.

Conjecture 21. *Any finite set $\Theta \subset \Pi'$ of linear functionals which admits an ideal interpolation scheme can be graded as*

$$\Theta = \bigcup_{\gamma \in \Gamma^*} \Theta_\gamma, \qquad \Theta'_\gamma := \bigcup_{\gamma' \leq \gamma} \Theta_{\gamma'}, \quad \gamma \in \Gamma^*$$

such that for $\gamma \in \Gamma^*$

(i) the interpolation problem with respect to Θ'_γ is poised for $\mathcal{P}^*_\gamma = \mathcal{P}^* \cap \Pi_\gamma$,

(ii) $\ker \Theta'_\gamma$ *is an ideal.*

6. More Special General Gradings

The approach by means of arbitrary gradings that we pursued in the preceding section provides the advantage of great generality and allows us to comprehend the general principle behind degree reducing interpolation. On the other hand, it is a very simple piece of wisdom that imposing further structure allows us to obtain more detailed results. To that end, we will now say goodbye to our general gradings and assume for this section that our grading is a *strict monomial* one, the concept that includes all the "popular" notions of degree and thus still offer sufficient generality. In particular, it covers the grading by total degree, by term orders and the gradings supported by the Gröbner fan.

6.1. Kernels from the right product line

So far we have not really put emphasis on the inner product used in the reduction process which nevertheless affects the behavior of the normal forms and thus the interpolation space and process quite significantly. It seems as if "the" inner product is the one first used for interpolation by de Boor and Ron, which is defined as

$$(f,g) = (f(D)g)(0) = \sum_{\alpha \in \mathbb{N}_0^d} f_\alpha \, g_\alpha \, \alpha!, \qquad f,g \in \Pi. \tag{35}$$

To be precise: this is only the *real* version of the inner product, in the case $\mathbb{K} = \mathbb{C}$, complex conjugation has to be added, but for simplicity I prefer being a realist here. The inner product in (35) seems to carry many names: Bombieri inner product or Fisher inner product, I have even been told by Ch. Dunkl that it has been used by Calderon in the context of harmonic polynomials. But whatever name it bears, this inner product turns out to be quite useful, not only in the practical problem of computing reductions [61], but also for an explicit description of the normal form interpolation space that has been given in [70], extending results e.g. from [13].

Theorem 22. *Suppose that $\Theta \subset \Pi'$ defines an ideal interpolation scheme and let G be a Γ–basis of $\mathcal{I}(\Theta)$, where the grading Γ is based on the inner product from (35). Then the canonical normal form interpolation space*

$$\mathcal{P}^* = \nu_{I(\Theta)}(\Pi) = \bigcap_{g \in G} \ker \Lambda(g)(D) \tag{36}$$

is closed under differentiation.

The "explicit" form (36) of the interpolation space $\mathcal{P}^*$ is well–known for the grading by total degree: it is precisely the *least interpolation* scheme from [10–13] which is thus the equivalence of normal form interpolation for any strict monomial grading as long as the associated inner product is chosen as in (35). Actually, there is even more behind this relationship. As has been shown in [70] for arbitrary gradings and in [67] in the Gröbner context, one can replay the whole formalism of leading terms for polynomials and least terms in power series for general gradings to obtain (36). In this respect, the argument is a purely algebraic one to identify normal forms as kernel elements of the differential operators $\Lambda(g)(D)$ and this, of course, has to take into account the nature of the inner product. On the other hand, the kernels of the differential operators can even be used to characterize Γ–bases, a fact proven in [69] as a tool to study quotient ideals that arise in connection with refinable functions.

Theorem 23. *A finite set $G \subset \Pi$ of polynomials is a Γ–basis for an ideal $\mathcal{I}$ if and only if*

$$\dim \Pi/\mathcal{I} = \dim \bigcap_{g \in G} \ker \Lambda(g)(D).$$

6.2. Lower sets and tensor product data

The fact that the total ordering "<" on Γ is compatible with addition makes it immediate that $\gamma \geq \gamma - \Gamma$ where, of course, $\gamma - \Gamma := \{\gamma' \in \Gamma \ : \ \gamma \in \gamma' + \Gamma\}$. Therefore, the span or direct sum of $\Pi_{\gamma'}$, $\gamma' \in \gamma - \Gamma$, is a natural subset of polynomials whose degree is bounded by γ. We now call a set $L \subset \Gamma$ a *lower set* if $\gamma \in L$ implies that $\gamma - \Gamma \subseteq L$. If, specifically, $\Gamma = \mathbb{N}_0^d$, then the lower sets take the form

$$\alpha - \Gamma = \alpha - \mathbb{N}_0^d = \left\{ \beta \in \mathbb{N}_0^d \;:\; \beta_j \le \alpha_j,\, j = 1, \ldots, d \right\}, \qquad \alpha \in \mathbb{N}_0^d.$$

Note that being a lower set is independent of the ordering, it is exclusively a property of the additive structure of the monoid Γ, the fact that lower sets interact properly with the ordering was an extra assumption that we had to make.

The "classical" gradings, i.e., the ones by total degree and by term orders, have an appealing property: the set Γ^* of degrees appearing in $\mathcal{P}^*$ (or, for the readers who prefer it in "algebraic": the degrees of the homogeneous components whose direct sum is $\mathcal{P}^*$) form a lower set, cf. [66,67]. For arbitrary strict monomial gradings the problem is more intricate: the simple example of the bivariate gradings based on

$$\Gamma = \mathbb{N}_0, \qquad \delta\left(x^\alpha\right) = \alpha_1 + n\,\alpha_2, \qquad n \in \mathbb{N},$$

have, for $\Xi = \{(0,0), (1,0), (0,1)\}$ their normal form spaces spanned by $\{1, x, y\}$ and these are homogeneous contributions are of degree $0, 1, n$ which is no lower set as soon as $n \ge 2$. However, this example clearly shows what really goes wrong here: there is a coordinate polynomial, namely y, that belongs to Π_n and this $n > 1$ can be decomposed into summands in $\mathbb{N}$; in other words,

$$n - \mathbb{N}_0 \ne \{n\} \qquad \text{or} \qquad n - \mathbb{N} = n - (\mathbb{N}_0 \setminus \{0\}) \ne \emptyset.$$

This can be overcome by requiring that for $j = 1, \ldots, d$

$$x_j \in \Pi_\gamma \qquad \Rightarrow \qquad \gamma - \Gamma = \{\gamma\}, \tag{37}$$

and indeed it can be shown that (37) even characterizes the situation where Γ^* is a lower set – provided that Γ is a strict monomial grading and that the monomials are all orthogonal with respect to the inner product, which is obviously the case for the one from (35). In a way, this observation and the above example can be understood in the sense that the grading should be "democratic" with respect to coordinate polynomials. Just note that in the example the coordinate polynomial y had the same degree as an nth power of the other coordinate polynomial x.

But let us continue with *tensor product sites* now, that is, with a very special, very classical, but also very organized and structural way of choosing the nodes. To that end, let $L \subset \mathbb{N}_0^d$ a finite lower set of multiindices and suppose that abscissae

$$\xi_{j,k} \in \mathbb{K}, \qquad j = 1, \ldots, d,\; k \in \mathbb{N}_0,$$

are given; we assume no ordering among them, and that k ranges over $\mathbb{N}_0$ is for convenience only. The tensor product sites are then given as

$$\Xi = \left\{ \xi_\alpha = \left(\xi_{1,\alpha_1}, \ldots, \xi_{d,\alpha_d}\right) \;:\; \alpha \in L \right\}. \tag{38}$$

This is slightly more general than the usual notion of tensor product sites that corresponds to choosing $L = \{\alpha \;:\; \alpha_j \le n_j\}$, $n_j \in \mathbb{N}$, $j = 1, \ldots, d$. On the other hand, this concept also includes "triangular" data by means of $L = \{\alpha \;:\; |\alpha| \le n\}$, $n \in \mathbb{N}_0$. This latter configuration, however, could also be interpreted as the principal lattice with the hyperplanes $H_{j,k} = x_j - \xi_{j,k}$.

There is a canonical interpolation space for node configurations as in (38) which is the associated space of "tensor product polynomials"

$$\mathcal{P} = \operatorname{span}_{\mathbb{K}} \{x^\alpha \ : \ \alpha \in L\}, \tag{39}$$

and this connection has made is possible to investigate the "triangular grid" mainly by means of univariate methods more than 100 years ago in [4], also leading remainder formulas by means of mean values, cf. [31,36,76]. The extension to lower sets was also studied by Werner [77] in the context of arbitrary Chebyshev systems. Deriving a remainder formula of the type (12) in terms of integrals against splines, on the other hand, was left to [74]. The splines appearing there were actually box splines and on the way of obtaining them the simplex spline decomposition of cube splines, cf. [49, p. 192ff], appeared in a very natural way.

An interesting property of the specific minimal degree interpolation procedures, in particular the *least interpolant* for which this fact is pointed out in [12], is that when applied to tensor product sites they are what one expects them to be, namely the tensor product space from (39). And so it is fair to ask whether this natural behavior is a specific aspect of particularly well-chosen interpolation spaces or whether this is a more general fact that holds for "any" choice of interpolation space. It turns out that the latter is true; more precisely, the following result holds true [71].

Theorem 24. *If Γ is a strict monomial grading and Ξ is of the form (38) then*

$$\nu_{\mathcal{I}(\Xi)}(\Pi) = \operatorname{span}\{x^\alpha \ : \ \alpha \in L\}.$$

In other words, any normal form interpolant with respect to a strict monomial grading – and this includes the least interpolant – *must* be a tensor product space. Of course, there always exist non–tensor interpolants as well, but in view of the preceding section those are more unnatural and the examples of such spaces always are very artificial in appearance. The proof of Theorem 24 is based on the fact that for tensor product sites Ξ from (38) the ideal $\mathcal{I}(\Xi)$ has a *universal* Γ*–basis*, that is, a set G which is a Γ–basis for all strict monomial gradings (and any inner product) simultaneously.

6.3. H–bases and the error formula revisited

Getting close to the end of our travel through some parts of the world of multivariate polynomial interpolation, we will enjoy a brief look on the ideal theoretic interpretation of error formulas. The starting point is the simple observation that for $f \in \Pi$ and an interpolation space $\mathcal{P}$ we clearly have that $\Theta\,(f - L_{\mathcal{P}} f) = 0$, i.e., $f - L_{\mathcal{P}} f \in \mathcal{I}(\Theta)$, regardless of whether we deal with Lagrange or Hermite interpolation schemes. Hence, as long as f is a polynomial and G is a Γ–basis of $\mathcal{I}(\Theta)$, there is a representation

$$f - L_{\mathcal{P}} f = \sum_{g \in G} P_g(f)\, g,$$

where any P_g is an operator that maps a polynomial f to a polynomial of degree at most $\delta(f) - \delta(g)$. Note that it is not all clear how to find such coefficient operators, and in general, they are in no way unique. Once an explicit formula for such operators P_g has been given – usually in terms of a simplex spline integral over certain derivatives of f or linear combinations thereof – we can then hope to extend this operator from polynomials to sufficiently smooth functions. In fact, smoothness would not even be needed for the error formulas from Section 4 as the formulas always hold in the *distributional* sense. On the other hand, note that unfortunately the temptation to assume duality, $P_g\,(g') = 0$, $g, g' \in G$, has to be resisted or at least been considered carefully: the representation with respect to a Γ–basis, not even with respect to a minimal Gröbner or H–basis, is usually *not* unique due to the unavoidable appearance of syzygies and so there is no a priori justification that the "good" operators have to be of this type. On the other hand, lack of uniqueness is not by itself a disadvantage as it allows for certain degrees of flexibility. Be it good or bad, the task to determine P_g a really hard one, and in full generality this is still an unsolved problem, though it is conjectured in [8] what a decent remainder formula should look like, albeit supported only by evidence for configurations that satisfy the very restrictive geometric characterization.

To interpret the error formulas from Section 4 in algebraic terms, we first note that the notion of degree employed there was that of *total degree* and so the Γ–bases will become H–bases. An almost perfect formula from the algebraic point of view is de Boor's formula (14). The $\binom{n+d}{d}$ polynomials $p_{\mathcal{K}}$ there even form a *minimal reduced H–basis* for $\mathcal{I}(\Xi)$ and the representation operators $P_{\mathcal{K}}$ are given explicitly as simplex spline integrals over a $n+1$–fold directional derivative which takes care of reducing the degree of polynomials by a proper amount.

But also (12) is quite easily converted into an expression with H–bases in it by observing that

$$[\Xi_\mu, x]\, f = \sum_{j=1}^{d} (x_j - \xi_{\mu^n, j}) \left[\Xi_\mu, x; \Xi_\mu D, \frac{\partial}{\partial x_j} \right] f$$

and so, with $\mathrm{M}_n(\alpha) = \{\mu \in \mathrm{M}_n \ : \ \mu^n = \alpha\}$,

$$(f - L_n f)(x) = \sum_{|\alpha|=n} \sum_{j=1}^{d} (x_j - \xi_{\alpha,j})\, p_\alpha(x) \sum_{\mu \in \mathrm{M}_n(\alpha)} \pi_\mu \left[\Xi_\mu, x; \Xi_\mu D, \frac{\partial}{\partial x_j} \right] f.$$

The polynomials $(x_j - \xi_{\alpha,j})\, p_\alpha(x)$, $|\alpha| = n$, $j = 1, \ldots, d$, form an H–basis for $\mathcal{I}(\Theta)$, but while the above representation has the "proper" algebraic form, it definitely lacks the simplicity of (14): the H–basis is by no means minimal and there is a loss of symmetry due to different ways of indexing. But, on the other hand, the formula holds true for *any* node set Ξ with respect to which the interpolation problem is poised for Π_n and thus is valid in a sufficiently general context. If there are simpler, more symmetric formulas for arbitrary point configurations, or even something of the form conjectured in equation (5.1) of [8], remains an open question at the moment.

Acknowledgements

A lot of people have contributed to this paper, consciously or not, willingly or not, by discussions, encouragement, collaboration or simply by listening patiently. There are, however, four persons to whom I am particularly indebted: Yuan Xu, for starting the work on polynomial interpolation with me, Michael Möller for guiding me into the wonderful world of algebra, Mariano Gasca for all his help and encouragement and Carl de Boor who spent an enormous effort on controversial discussion from which I benefitted very much. Thanks to all of them!

In addition, Michael Möller and Carl de Boor provided me with a lot of helpful comments and suggestions for this present paper which helped a lot to improve content and presentation and which I am very grateful for.

References

1. Aitken, A. G., On interpolation by iteration of proportional parts, without the use of differences, *Proc. Edinburgh Math. Soc.* **3** (1932), 56–76.
2. Andoyer, H., Interpolation, in: *Encyclopédie de Sciences Mathématiques, Tôme I, Vol. 4*, Gauthier–Villars, Paris, 1906.
3. Bauschinger, J., Interpolation, in: *Encyklopädie der Mathematischen Wissenschaften, Bd. I, Teil 2*, Teubner, Leipzig, 1900, pp. 800–821.
4. Biermann, O., Über näherungsweise Kubaturen, *Monatsh. Math. Phys.* **14** (1903), 211–225.
5. Birkhoff, G., The algebra of multivariate interpolation, in: *Constructive Approaches to Mathematical Models* (C. Coffman and G. Fix, Eds.), Academic Press, Boston, 1979, pp. 345–363.
6. de Boor, C., Gauss elimination by segments and multivariate polynomial interpolation, in: *Approximation and Computation: A Festschrift in Honor of Walter Gautschi* (R. V. M. Zahar, Ed.), Birkhäuser, Basel, 1994, pp. 87–96.
7. de Boor, C., A multivariate divided difference, in: *Approximation Theory VIII, Vol. 1: Approximation and Interpolation* (C. K. Chui and L. L. Schumaker, Eds.), World Scientific, Singapore, 1995, pp. 87–96.
8. de Boor, C., The error in polynomial tensor-product, and in Chung-Yao, interpolation, in: *Surface Fitting and Multiresolution Methods* (A. Le Méhauté, C. Rabut and L. L. Schumaker, Eds.), Vanderbilt University Press, Nashville, 1997, pp. 35–50.
9. de Boor, C., Interpolation from spaces spanned by monomials, *Adv. Comput. Math.*, to appear.
10. de Boor, C. and Ron, A., On multivariate polynomial interpolation, *Constr. Approximation* **6** (1990), 287–302.
11. de Boor, C. and Ron, A., On polynomial ideals of finite codimension with applications to box spline theory, *J. Math. Anal. Appl.* **158** (1991), 168–193.
12. de Boor, C. and Ron, A., Computational aspects of polynomial interpolation in several variables, *Math. Comp.* **58** (1992), 705–727.

13. de Boor, C. and Ron, A., The least solution for the polynomial interpolation problem, *Math. Z.* **210** (1992), 347–378.
14. Buchberger, B., *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal*, PhD thesis, Innsbruck, 1965.
15. Buchberger, B., Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems, *Aequationes Math.* **4** (3) (1970), 374–383.
16. Buchberger, B. and Möller, H. M., The construction of multivariate polynomials with preassigned zeros, in: *Computer Algebra*, Proc. EUROCAM 1982 (G. Goos and J. Hartmanis, Eds.), Lecture Notes in Computer Science, Vol. 144, Springer, Berlin, 1982, pp. 24–31.
17. Busch, J. R., A note on Lagrange interpolation in $\mathbb{R}^2$, *Rev. Union Matem. Argent.* **36** (1990), 33–38.
18. Carnicer, J. M. and Gasca, M., A conjecture on multivariate polynomial interpolation, Rev. R. Acad. Cien., Serie A. Mat. **95** (2001), 145–153.
19. Carnicer, J. M. and Gasca, M., Classification of bivariate GC configurations for interpolation, *Adv. Comput. Math.* **20** (2004), 5–16.
20. Carnicer, J. M. and Gasca, M., Interpolation on lattices generated by cubic pencils, *Adv. Comput. Math.*, to appear.
21. Carnicer, J., Gasca, M. and Sauer, T., Interpolation lattices in several variables, preprint (2005).
22. Chung, K. C. and Yao, T. H., On lattices admitting unique Lagrange interpolation, *SIAM J. Numer. Anal.* **14** (1977), 735–743.
23. Ciarlet, P. G. and Raviart, P. A., General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods, *Arch. Rational Mech. Anal.* **46** (1972), 178–199.
24. Cox, D., Little, J. and O'Shea, D., *Ideals, Varieties and Algorithms*, 2nd Edition, Springer, Berlin. 1996.
25. Davis, P. J., *Interpolation and Approximation*, Dover, New York, 1975.
26. Gasca, M., Multivariate polynomial interpolation, in: *Computation of Curves and Surfaces* (W. Dahmen, M. Gasca and C. A. Micchelli, Eds.), Kluver Academic Publ., Dordrecht, 1990, pp. 215–236.
27. Gasca, M. and Maeztu, J. I., On Lagrange and Hermite interpolation in $\mathbb{R}^k$, *Numer. Math.* **39** (1982), 1–14.
28. Gasca, M. and Mühlbach, G., Multivariate polynomial interpolation under projectivities. Part I: Lagrange and Newton interpolation formulas, *Numer. Algorithms* **1** (1991), 375–400.
29. Gasca, M. and Mühlbach, G., Multivariate polynomial interpolation under projectivities II: Neville–Aitken formulas, *Numer. Algorithms* **2** (1992), 255–278.
30. Gasca, M. and Mühlbach, G., Multivariate polynomial interpolation under projectivities III: Remainder formulas, *Numer. Algorithms* **8** (1994), 103–109.
31. Gasca, M. and Sauer, T., Polynomial interpolation in several variables, *Adv. Comput. Math.* **12** (2000), 377–410.

32. Gasca, M. and Sauer, T., On bivariate Hermite interpolation with minimal degree polynomials, *SIAM J. Numer. Anal.* **37** (2000), 772–798.
33. Gasca, M. and Sauer, T., On the history of multivariate polynomial interpolation, *J. Comput. Appl. Math.* **122** (2000), 23–35.
34. Gröbner, W., *Algebraische Geometrie II*, B.I–Hochschultaschenbücher, Vol. 737, Bibliographisches Institut Mannheim, 1970.
35. Hilbert, D., Über die Theorie von algebraischen Formen, *Math. Ann.* **36** (1890), 473–534.
36. Isaacson, E. and Keller, H. B., *Analysis of Numerical Methods*, John Wiley & Sons, New York, 1966.
37. Kowalewski, G., *Interpolation und genäherte Quadratur*, Teubner, Berlin, 1932.
38. Kronecker, L., Über einige Interpolationsformeln für ganze Funktionen mehrerer Variabeln, in: *L. Kroneckers Werke* (H. Hensel, Ed.), Chelsea Publ. Company, New York 1966, pp. 133–141.
39. Lee, S. L. and Phillips, G. M., Polynomial interpolation at points of a geometric mesh on a triangle, *Proc. Roy. Soc. Edinburgh* **108A** (1988), 75–87.
40. Lorentz, G. G., *Approximation of Functions*, 2nd ed., Chelsea Publ. Company, New York, 1986.
41. Lorentz, G. and Lorentz, R., Solvability problems of bivariate interpolation. I, *Constr. Approximation* **2** (1986), 153–170.
42. Lorentz, G. and Lorentz, R., Solvability problems of bivariate interpolation. II. Applications, *Approx. Theory Appl.* **3** (1987), 79–97.
43. Lorentz, G. and Lorentz, R., Bivariate Hermite interpolation and applications to algebraic geometry, *Numer. Math.* **57** (1990), 669–680.
44. Lorentz, R. A., *Multivariate Birkhoff Interpolation*, Lecture Notes in Mathematics, Vol. 1516, Springer, Berlin, 1992.
45. Lorentz, R. A., Multivariate Hermite interpolation by algebraic polynomials: a survey, *J. Comput. Appl. Math.* **122** (2000), 167–201.
46. Macaulay, F. S., *The Algebraic Theory of Modular Systems*, Cambridge Tracts in Math. and Math. Physics, Vol. 19, Cambridge Univ. Press, Cambridge, 1916.
47. Mairhuber, J. C., On Haar's theorem concerning Chebychev approximation problems having unique solutions, *Proc. Amer. Math. Soc.* **7** (1956), 609–615.
48. Marinari, M. G., Möller, H. M. and Mora, T., Gröbner bases of ideals given by dual bases, in: *Proc. ISAAC*, 1991.
49. Micchelli, C. A., *Mathematical Aspects of Geometric Modeling*, CBMS-NSF Reg. Conf. Series in Applied Mathematics, Vol. 65, SIAM, Philadelphia, 1995.
50. Micchelli, C. A., On a numerically efficient method of computing multivariate B–splines, in: *Multivariate Approximation Theory* (W. Schempp and K. Zeller, Eds.), Birkhäuser, Basel, 1979, pp. 211–248.
51. Micchelli, C. A., A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B–splines and Lagrange interpolation, *Rocky Mountain J. Math.* **10** (1979), 485–497.
52. Möller, H. M., Mehrdimensionale Hermite–Interpolation und numerische Integration, *Math. Z.* **148** (1976), 107–118.

53. Möller, H. M., Hermite interpolation in several variables using ideal–theoretic methods, in: *Constructive Theory of Functions of Several Variables* (W. Schempp and K. Zeller, Eds.), Lecture Notes in Mathematics, Vol. 571, Springer, Berlin, 1977, pp. 155–163.
54. Möller, H. M. and Sauer, T., H–bases for polynomial interpolation and system solving, *Adv. Comput. Math.* **12** (2000), 335–362.
55. Möller, H. M. and Sauer, T., H–bases I: The foundation, in: *Curve and Surface Fitting*, Proc. Saint–Malo 1999 (A. Cohen, C. Rabut and L. L. Schumaker, Eds.), Vanderbilt University Press, Nashville, 2000, pp. 325–332.
56. Möller, H. M. and Sauer, T., H–bases II: Applications to numerical problems, in: *Curve and Surface Fitting*, Proc. Saint–Malo 1999 (A. Cohen, C. Rabut and L. L. Schumaker, Eds.), Vanderbilt University Press, Nashville, 2000, pp. 333–342.
57. Möller, H. M. and Sauer, T., Multivariate refinable functions of high approximation order via quotient ideals of Laurent polynomials, *Adv. Comput. Math.* **20** (2004), 205–228.
58. Neville, E. H., Iterative interpolation, *J. Indian Math. Soc.* **20** (1934), 87–120.
59. Peña, J. M. and Sauer, T., On the multivariate Horner scheme, *SIAM J. Numer. Anal.* **37** (2000), 1186–1197.
60. Peña, J. M. and Sauer, T., On the multivariate Horner scheme II: Running error analysis, *Computing* **65** (2000), 313–322.
61. Peña, J. M. and Sauer, T., Efficient polynomial reduction, *Adv. Comput. Math.*, to appear.
62. Rabut, C., Generalized divided difference with simple knots, *SIAM J. Numer. Anal.* **38** (2000), 1294–1311.
63. Robbiano, L., On the theory of graded structures, *J. Symbolic Computation* **2** (1986), 139–170.
64. Rushdie, S., *Shame*, Vintage Books, London, 1983.
65. Sauer, T., Computational aspects of multivariate polynomial interpolation, *Adv. Comput. Math.* **3** (3) (1995), 219–238.
66. Sauer, T., Polynomial interpolation of minimal degree, *Numer. Math.* **78** (1) (1997), 59–85.
67. Sauer, T., Polynomial interpolation of minimal degree and Gröbner bases, in: *Groebner Bases and Applications*, Proc. Conf. 33 Years of Groebner Bases (B. Buchberger and F. Winkler, Eds.), London Math. Soc. Lecture Notes, Vol. 251, Cambridge University Press, Cambridge, 1998, pp. 483–494.
68. Sauer, T., Gröbner bases, H–bases and interpolation, *Trans. Amer. Math. Soc.* **353** (2001), 2293–2308.
69. Sauer, T., Polynomial interpolation, ideals and approximation order of refinable functions, *Proc. Amer. Math. Soc.* **130** (2002), 3335–3347.
70. Sauer, T., Ideal bases for graded polynomial rings and applications to interpolation, in: *Multivariate Approximation and Interpolation with Applications* (M. Gasca, Ed.), Monograph. Academia de Ciencias de Zaragoza, Vol. 20, Academia de Ciencias Zaragoza, Zaragoza, 2002, pp. 97–110.

71. Sauer, T., Lagrange interpolation on subgrids of tensor product grids, *Math. Comp.* **73** (2004), 181–190.
72. Sauer, T. and Xu, Y., On multivariate Lagrange interpolation, *Math. Comp.* **64** (1995), 1147–1170.
73. Sauer, T. and Xu, Y., On multivariate Hermite interpolation, *Adv. Comput. Math.* **4** (4) (1995), 207–259.
74. Sauer, T. and Xu, Y., A case study in multivariate Lagrange interpolation, in: *Approximation Theory, Wavelets and Applications*, NATO–ASI Proceedings (S. Singh, Ed.), Kluwer Academic Publ., Dordrecht, 1995, pp. 443–452.
75. Sauer, T. and Xu, Y., The Aitken–Neville scheme in several variables, in: *Approximation Theory X* (C. K. Chui, L. L. Schumaker and J. Stöckler, Eds.), Vanderbilt University Press, Nashville, 2002, pp. 353–366.
76. Steffensen, I. F., *Interpolation*, Chelsea, New York, 1927.
77. Werner, H., Remarks on Newton type multivariate interpolation for subsets of grids, *Computing* **25** (1980), 181–191.