Multivariate refinable functions, differences and ideals —

a simple tutorial

Tomas Sauer

Lehrstuhl für Numerische Mathematik, Justus–Liebig–Universität Gießen, Heinrich–Buff–Ring 44, D–35392 Gießen, Germany

Received 10 October 2006; received in revised form 22 May 2007

Abstract

This paper summarizes the algebraic quotient ideal approach to polynomial generation by refinable functions and connects it to Strang–Fix conditions and factorization with respect to difference operators. Motivated by the latter one, we also consider vector subdivision schemes with matrix valued coefficients and review some of their properties.

© 2008 Published by Elsevier B.V.

MSC: 65T60; 68W30

Keywords: Subdivision; Quotient ideal; Strang–Fix conditions

1. Introduction

A compactly supported function \( f : \mathbb{R}^s \to \mathbb{R} \) is called \textit{refinable} if there exist coefficients \( a(\alpha), \alpha \in \mathbb{Z}^s \), such that \( f \) can be represented as a linear combination of shifted and dilated copies of itself:

\[
f = \sum_{\alpha \in \mathbb{Z}^s} f(2 \cdot -\alpha) a(\alpha). \tag{1}\]

Being the building block of a multiresolution analysis (MRA), those functions have been the subject of considerable research and it is practically impossible to list all the literature that deals with refinable functions, so that I only list the books [8,12,26] here and recommend them for further references.

In most cases, in fact, practically whenever \( f \) is not a cardinal B-spline, the refinable function \( f \) is not given explicitly, but only in terms of the mask coefficients \( a = (a(\alpha) : \alpha \in \mathbb{Z}^s) \). Therefore, it is important to have criteria available that describe properties of the function \( f \) in terms of the mask \( a \). For compactly supported \( f \), a property of particular importance is the containment of polynomials of a certain degree in the shift invariant space generated by \( f \), in symbols

\[
\Pi_n \subset \mathbb{S}(f) = \left\{ \sum_{\alpha \in \mathbb{Z}^s} f(\cdot - \alpha) c(\alpha) : c(\alpha) \in \mathbb{R}, \alpha \in \mathbb{Z}^s \right\}, \tag{2}\]
where

$$\Pi_n = \text{span} \{ c^\alpha = z_1^{\alpha_1} \cdots z_s^{\alpha_s} : n \geq |\alpha| = \alpha_1 + \cdots + \alpha_s \},$$

is the vector space of all complex polynomials in $\Pi = \mathbb{C}[z_1, \ldots, z_s]$ of total degree at most $n$, since this allows for good approximation of smooth functions from $\mathcal{S}(f)$, more precisely, for $g \in C^{(n+1)}(\mathbb{R}^s)$ one has that

$$\inf_{c_\alpha} \left\| g - \sum_{\alpha \in \mathbb{Z}^s} f(2^n \cdot -\alpha) c(\alpha) \right\| \leq C 2^{-rn}.$$  

Recall that the idea behind the proof of such approximation results is to first approximate $g$ at a fixed point by a Taylor polynomial of order $n$ which is reproduced by $\mathcal{S}(f)$ and then to use the locality of $f$ to approximate the small error term of the Taylor expansion. We will not go into further details here, but only use this argument as a motivation or justification why we are interested in describing the polynomial (re)production property (2).

The intention of this paper is to describe and explain the algebraic property of the mask that ensures polynomial reproduction in a simple way and to point out that these properties are natural though nontrivial generalizations of their univariate counterpart. It will also lead us to the concept of vector subdivision which will be described as well.

The intention of this paper is not to be a survey on conditions that describe polynomial reproduction by refinable functions (for example, the sum rules are missing completely), so that the list of references is definitely incomplete as I will describe one possible approach only.

2. Symbols, transforms, filters

Let us begin with a little bit of notation. By $\ell(\mathbb{Z}^d)$ we denote the totality of all multiindexed sequences $c = (c(\alpha) : \alpha \in \mathbb{Z}^d)$, use $\ell_p(\mathbb{Z}^d)$ for those with finite $p$-norm, $1 \leq p \leq \infty$, and write $\ell_0(\mathbb{Z}^d)$ for the compactly supported sequences. Moreover, we will use $C_0(\mathbb{R}^d)$ for the compactly supported continuous functions defined on $\mathbb{R}^d$; while much of the theory can be extended to the $L_p$-case as well, we will focus on continuous functions here for the sake of simplicity.

For $f \in C_0(\mathbb{R}^d)$ and $c \in \ell_p(\mathbb{Z}^d)$ we define the (sometimes called “semidiscrete”) convolution

$$f * c = \sum_{\alpha \in \mathbb{Z}^d} f(\cdot - \alpha) c(\alpha);$$

(3)

since $f$ is compactly supported and continuous, the above sum yields a well-defined continuous function that also belongs to $L_p(\mathbb{R}^d)$ provided that $c \in \ell_p(\mathbb{Z}^d)$, $1 \leq p < \infty$. If $c \in \ell_{\infty}(\mathbb{Z}^d)$ then $f * c$ is a uniformly continuous and uniformly bounded function. The algebraic span of $f$’s integer translates is $\mathcal{S}(f) = f * \ell(\mathbb{Z}^d)$.

To a sequence $c \in \ell(\mathbb{Z}^d)$ we associate its symbol which is the formal Laurent series

$$c^s(z) = \sum_{\alpha \in \mathbb{Z}^d} c(\alpha) z^\alpha, \quad z \in \mathbb{C}_x, \quad \mathbb{C}_x = \mathbb{C} \setminus \{0\}.$$  

If $c \in \ell_0(\mathbb{Z}^d)$ has finite support, then $c^s$ is a Laurent polynomial, hence a finite sum

$$c^s(z) = \sum_{\alpha \in \mathbb{Z}^d} c(\alpha) z^\alpha = \sum_{\beta \in \mathbb{N}^d} c(\alpha + \beta) z^\alpha = z^\beta p(z), \quad p \in \Pi,$$

for an appropriate $\beta \in \mathbb{Z}^d$. In other words:

Any Laurent polynomial can be written as the product of a monomial and a polynomial.

As simple, obvious and innocent as this observation may appear, it will turn out to be very useful as it permits us to use polynomial ideal theory later, which keeps things very simple. Despite of their similar nature, polynomials and Laurent polynomials are surprisingly different algebraic objects, and it is not difficult to see why: just recall that a unit in a ring is an element that has a multiplicative inverse in the ring, for example the numbers $\pm 1$ in $\mathbb{Z}$, the general linear group in the ring of square matrices or everything except zero in a field like $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. So the units can be expected to play an important role for the ring and to some extent describe it, but while the units in $\Pi = \mathbb{C}[z_1, \ldots, z_s]$, the ring

...
of polynomials in $s$ variables, are precisely the nonzero constant polynomials, any nonzero multiple of a monomial $z^\alpha, \alpha \in \mathbb{Z}^s$, is a unit in $A = \mathbb{C}[z_1^{\pm 1}, \ldots, z_s^{\pm 1}]$, the ring of Laurent polynomials, so that the latter one is spanned (as a vector space) by its units! This has some quite unpleasant consequences.

**Theorem 1.** The only notion of degree that Laurent polynomials can be endowed with is the trivial one: All Laurent polynomials have degree zero.

But let us turn back to the analysis of refinable functions, i.e., to functions which satisfy the refinement equation (1), written in terms of the convolution as $f = f \ast a (2 \cdot)$. Again, we stress the fact that both $f$ and $a$ are supposed to be of compact or finite support, respectively. Taking the Fourier transform of the refinement equation leads to the equivalent description

$$\hat{f}(\xi) = \frac{1}{2^s} \hat{f} \left( \frac{\xi}{2} \right) \hat{a} \left( \frac{\xi}{2} \right), \quad \hat{a}(\xi) := a^* \left( e^{i\xi} \right), \quad \xi \in \mathbb{R}^s. \quad (4)$$

We will return to this identity later, but for the moment we look at another simple consequence of the refinement equation, namely

$$f \ast c = \sum_{\alpha \in \mathbb{Z}^s} f \cdot (-\alpha) c(\alpha) = \sum_{\alpha \in \mathbb{Z}^s} \sum_{\beta \in \mathbb{Z}^s} f \left( 2(-\alpha) - \beta \right) a(\beta)c(\alpha)$$

$$= \sum_{\beta \in \mathbb{Z}^s} f \left( 2 \cdot -\beta \right) \left( \sum_{\alpha \in \mathbb{Z}^s} a (\beta - 2\alpha) c(\alpha) \right) =: f \ast S_a c(2 \cdot),$$

which defines the subdivision operator $S_a$. A very straightforward computation also shows us that

$$(S_a c)^* = \sum_{\alpha \in \mathbb{Z}^s} \sum_{\beta \in \mathbb{Z}^s} a (\alpha - 2\beta) c(\beta) z^\alpha$$

$$= \sum_{\beta \in \mathbb{Z}^s} \left( \sum_{\alpha \in \mathbb{Z}^s} a (\alpha - 2\beta) z^{\alpha - 2\beta} \right) c(\beta) z^{2\beta}$$

$$= \sum_{\alpha \in \mathbb{Z}^s} a (\alpha) z^{\alpha} \sum_{\beta \in \mathbb{Z}^s} c(\beta) z^{2\beta} = a^* (z) c^* \left( z^2 \right),$$

a formula that can be interpreted in terms of signal processing as an upsampling of $c$ followed by a convolution with the filter $a$, see [25,26]. The subdivision operator is intimately connected with another operator, namely the transition operator $T_a$, defined as

$$T_a f = \sum_{\alpha \in \mathbb{Z}^s} f \left( 2 \cdot -\alpha \right) a (\alpha),$$

which has a refinable function $f$ as a fixpoint: $T_a f = f$ iff $f$ satisfies (1). In fact, the iteration of both operators can be used to compute refinable functions, either in a discrete way by interpreting the entries of $S_a^\ell b$ as the approximate values of $f$ at the grid $2^{-\ell} \mathbb{Z}^s$, or by considering the sequence of functions $T_a^\ell g$ for an appropriate initial function $g$, namely a test function in the sense of [7]. Without becoming too specific here, one can say that the subdivision iteration $S_a^\ell c$ converges to $f \ast c$ for any $c \in \ell_\infty (\mathbb{Z}^s)$ if and only if the iteration $T_a^\ell g$ converges to $f$ for one (or for any) test function $g$ (if it works for one, it works for each of them), and that the reason for this is comparatively simple: the operators $S_a$ and $T_a$ are mutual adjoints with respect to the bilinear form

$$(\cdot, \cdot) : C_00 (\mathbb{R}^s) \times \ell_\infty (\mathbb{Z}^s) \to C (\mathbb{R}^s), \quad (f, c) := f \ast c.$$

Let me close this section by mentioning that it contains none of the “scary” truly multivariate concepts. All we did and used so far were merely formal extensions of the univariate formulas, essentially a cut-and-paste transition from integer indices to multiindices. Even the structural difference between polynomials and Laurent polynomials is present in the univariate case as well.
3. The Strang–Fix-conditions and the mask

The Strang–Fix conditions describe in terms of the Fourier transform when the space $S(f)$ contains polynomials of degree up to $n$. Though we only state it only as a sufficient condition here, it is almost a characterization, or, to be precise, it is one if $f$ satisfies the additional condition

$$0 \neq (f \ast 1)(0) = \sum_{\alpha \in \mathbb{Z}^s} f(\alpha).$$

The result is as follows.

**Theorem 2 (Strang–Fix Conditions).** If the function $f$ satisfies the Strang–Fix conditions of order $n$:

1. $\hat{f}(0) \neq 0$,
2. for $|\beta| \leq n$ and $\alpha \in \mathbb{Z}^s \setminus \{0\}$ one has $D^\beta \hat{f}(2\pi \alpha) = 0$,

then $\Pi_\alpha \subseteq S(f)$.

The proof of this result is based on the very nice idea to apply the Poisson summation formula on the convolution $f \ast ()^\beta$ between $f$ and a monomial sequence $(\alpha^\beta : \alpha \in \mathbb{Z}^s)$ to obtain

$$p(x) := \sum_{\alpha \in \mathbb{Z}^s} f(x - \alpha) \alpha^\beta = \sum_{\alpha \in \mathbb{Z}^s} |\alpha| \sum_{\gamma \leq \beta} \left( \begin{array}{c} \beta \\ \gamma \end{array} \right) D^\gamma \hat{f}(2\pi \alpha) (ix)^{\beta - \gamma} e^{2\pi i \alpha^T x}.$$ 

Now the function on the right-hand side has the periodic part $e^{2\pi i \alpha^T x}$ and can be a polynomial (if and only if all terms with $\alpha \neq 0$ vanish)—by an inductive argument, this leads to the conditions in **Theorem 2**.

But what is the use of the Strang–Fix conditions for our purposes, taking into account that in the preceding chapter we stressed the fact that the function $f$ is usually not known explicitly, that all available information is really the mask $a$? Indeed, the fact that the paper [24] by Strang and Fix appeared long before refinable functions were studied seriously and a brief look at its title should convince us that the Strang–Fix conditions were not designed for refinable functions. However, it is not difficult to “transfer” them to the mask, more precisely, to its Fourier transform by making use of refinability. To that end, just look at the Strang–Fix conditions of order zero. $\hat{f}(2\pi \alpha) = 0$, $\alpha \in \mathbb{Z}^s \setminus \{0\}$, and substitute the refinement Eq. (4) to obtain

$$0 = \hat{f}(2\pi \alpha) = \frac{1}{2^s} \hat{f}(\pi \alpha) \hat{a}(\pi \alpha).$$

We now write $\alpha = \epsilon + 2\pi \beta$, $\beta \in \mathbb{Z}^s$, where $\epsilon \in \{0, 1\}^s$ indicates the parity of the components of $\alpha$, and use the $2\pi$-periodicity of $\hat{a}$ to get

$$0 = \hat{f}(2\pi \alpha) = \frac{1}{2^s} \hat{f}(\pi \epsilon + 2\pi \beta) \hat{a}(\pi \epsilon + 2\pi \beta) = \frac{1}{2^s} \hat{f}(\pi \epsilon + 2\pi \beta) \hat{a}(\pi \epsilon).$$

Hence, if we suppose that

$$\left( \hat{f}(\pi \epsilon + 2\pi \alpha) : \alpha \in \mathbb{Z}^s \right) \neq 0, \quad \epsilon \in E := \{0, 1\}^s \setminus \{0\},$$

the above observation immediately yields the fundamental condition

$$0 = \hat{a}(\epsilon \pi) = a^\epsilon \left( e^{2\pi \epsilon} \right), \quad \epsilon \in E,$$

on the mask $a$. By an inductive application of this idea and a little bit of Leibniz rule it is not difficult any more to prove the following result.

**Theorem 3.** For an $a$-refinable function $f \in C_{00}(\mathbb{R}^s)$ which satisfies (5) the following statements are equivalent:

1. $\Pi_\alpha \subset S(f)$,
2. $f$ satisfies the Strang–Fix conditions of order $n$,
3. $D^\beta \hat{a}(\pi \epsilon) = 0$, $|\beta| \leq 0$, $\epsilon \in E$,
4. $D^\beta a^\epsilon(z) = 0$, $|\beta| \leq 0$, $z \in e^{i\pi E} := \{-1, 1\}^s \setminus \{(1, \ldots, 1)\}$. 


So finally we are in business. The last criterion in Theorem 3 gives an algebraic condition on $a$, more precisely on $a^*$ that has to be satisfied in order to achieve polynomial (re)production and this condition will eventually become the “truly multivariate” one. But before we focus on this in the paragraphs to follow, we close with a comment on (5): indeed, these conditions are a subset of the conditions

$$(\hat{f}(\xi + 2\pi \alpha) : \alpha \in \mathbb{Z}^3) \neq 0, \quad \xi \in \mathbb{R}^3,$$

for the stability of $f$ or the linear independence of its integer translates, see [10], and without this condition it is generally impossible to describe polynomial generation in terms of the mask.

**Example 4.** The standard counterexample is the even univariate function $f = \chi_{[0,2]}$; it is refinable with $a^*(z) = 1 + z^2$ and has the Fourier transform $\hat{f}(\xi) = 2e^{-i\xi} \sin \frac{\xi}{2}$ so that

$$(\hat{f}(\pi + 2k\pi) : k \in \mathbb{Z}) = (\hat{f}((2k + 1)\pi) : k \in \mathbb{Z}) = 0.$$

This function generates constants as $f * \frac{1}{2} = 1$ while $a^*(-1) = 2 \neq 0$, thus showing that the condition (5) is really needed in Theorem 3.

4. Ideals and smallest masks

Finally, Theorem 3 brings us to a situation where there is a difference between the univariate and the multivariate case: in one variable the condition is simply the presence of an $n$-fold zero of $\hat{a}$ at $\pi$ and of $a^*$ at $-1$, respectively, which corresponds to the fact that $a^*(z) = (z + 1)^n b^*(z)$ for some other finitely supported mask $b \in \ell_0^0(\mathbb{Z}^3)$.

The multivariate case, however, is different as being factorizable is a rare property among polynomials in several variables. But there is still structure, of course. To explore this structure, let us consider $a^*$ and make the simplifying assumption that $a^*$ is a polynomial. In fact, this is not a serious restriction as the multiplication of $a^*$ with any monomial corresponds to a shift of the mask and then a shifted copy of $f$ would be refinable with respect to the shifted mask. According to Theorem 3, the polynomial $a^*$ must now belong to the set $\mathcal{I}$ of all polynomials that vanish of order $n$ at the points from $e^{i\pi E}$. Clearly, if two polynomials $p, p'$ belong to $\mathcal{I}$ then so do their sum as derivatives and point evaluations are linear functionals. Moreover, if $p \in \mathcal{I}$ belongs to this set and $q \in \mathcal{II}$ is an arbitrary polynomial, then the Leibniz rule yields for any $z \in e^{i\pi E}$ and $|\beta| \leq n$ that

$$D^\beta (pq)(z) = \sum_{\gamma \leq \beta} \binom{\beta}{\gamma} D^\gamma p(z) D^{\beta - \gamma} q(z) = 0,$$

hence $pq \in \mathcal{I}$. In other words, the set from which we have to choose our symbol $a^*$ is an ideal in $\mathcal{II}$ and this is good since Hilbert’s famous Basisatz tells us that such an ideal has to have a finite basis, thus can be described by finite information.

**Definition 5.** A set $P \subset \mathcal{II}$ is called a basis for the ideal $\mathcal{I}$ if

$$\mathcal{I} = \langle P \rangle = \left\{ \sum_{p \in P} q_p p : q_p \in \mathcal{II} \right\},$$

where $\langle P \rangle$ is the ideal generated by $P$.

Describing an ideal usually means giving a basis for it and such bases are what your favorite Computer Algebra system like Maple, Mathematica, MuPAD, Singular or CoCoA can store and manipulate. Hence, the primary goal of this section will be to find bases for the family of ideals

$$\mathcal{I}_n := \left\{ p \in \mathcal{II} : D^\beta p(z) = 0, |\beta| \leq n, z \in e^{i\pi E} \right\}, \quad n \in \mathbb{N},$$

which clearly satisfy $\mathcal{I}_0 \supset \mathcal{I}_1 \supset \mathcal{I}_2 \supset \cdots$. In doing so for the case $n = 0$, we begin with a slightly different ideal, namely

$$\mathcal{J} := \left\{ p \in \mathcal{II} : p(z) = 0, z \in \{-1, 1\}^3 \right\}.$$
and show, like it had already been done in [2], that it is easy to find a basis for this ideal, namely the functions 
\((z_j^2 - 1)\), \(j = 1, \ldots, s\). Obviously, they all vanish on \(\{\pm 1\}^2\), while any polynomial \(p \in \Pi\) can be written as the sum of an element from \(\mathcal{J}\) and a polynomial that interpolates on \(\{\pm 1\}^2\), and this interpolation polynomial can be taken from the space spanned by \(z^\alpha, \alpha \in \{0, 1\}^s\). In fact, the interpolation polynomial takes the explicit form

\[ Lp = \sum_{\alpha \in \{0, 1\}^s} p((-1)^{\alpha_1}, \ldots, (-1)^{\alpha_s}) \prod_{j=1}^{s} \left(\frac{(1 + z_j)^{1-\alpha_j} - (1 - z_j)^{\alpha_j}}{2}\right), \quad p \in \Pi, \]

which is easily verified by substitution. With this observation at hand, it is not difficult to show that

\[ \mathcal{J} = \langle z^2 - 1 \rangle := \left\langle \left(z_j^2 - 1 \right) : j = 1, \ldots, s \right\rangle, \] (8)

hence, any polynomial \(p\) that vanishes on \(\{\pm 1\}^2\) has the form

\[ p(z) = \sum_{j=1}^{s} (z_j^2 - 1) q_j(z) \] (9)

and while in the univariate case this simply means \(p(z) = (z^2 - 1) q(z)\), the multivariate situation is obviously more intricate.

Before we return to our original problem of determining \(\mathcal{I}_n\), we take a little excursion to computational ideal theory. Having a basis for an ideal means that we can parameterize the ideal by means of the basis elements and polynomial “coefficients” which gives us a way to systematically generate all polynomials that fulfill our side conditions. But for this process it is desirable to obtain the ideal masks in increasing complexity, i.e., in increasing support size as this affects the filter length as well as the support of the refinable function. In doing so, it would be preferable if we could control the degree of the result completely in terms of the parameters, i.e., if

\[ \mathcal{J} \cap \Pi_m = \sum_{j=1}^{s} (z_j^2 - 1) \Pi_{m-2}. \]

so that we get all masks of total degree at most \(m\) by using as parameters polynomials of degree at most \(m - 2\). In general, it seems to be very reasonable to look for ideal bases \(P\) which have the crucial property that

\[ \langle P \rangle \cap \Pi_m = \sum_{p \in P} p \Pi_{m-\deg p}; \]

such bases have been introduced by Macaulay as early as 1916 and they are called \(H\)-bases. For our purposes here, it suffices to know the following:

1. Not any basis of an ideal is automatically an \(H\)-basis, but for any given ideal there usually exists a multitude of \(H\)-bases, among them always some Gröbner bases, for example the graded lexicographic “\texttt{gradlex}” one, cf. [6].
2. Such bases can be efficiently computed from an initial basis of the ideal and at least the computation of the \texttt{gradlex} Gröbner basis is included in any Computer Algebra system.
3. There also exist \(H\)-bases that are no Gröbner bases, see [19], but which can nevertheless be determined computationally and often offer more symmetry and numerical stability, see [16,17].

So we could now feed the basis vector \([z^2 - 1]\) into the computer which would, to our great surprise, return this very basis as a Gröbner basis, in fact even independently of the term orders we could choose. Actually, there is nothing wrong with the Computer Algebra systems, as one can prove that \([z^2 - 1]\) is a universal basis for the ideal which even holds true for all bases dual to tensor-product-like interpolation schemes, cf. [22].

Though we have learned a bit about ideals now, we are still not yet at the point of describing the ideal \(\mathcal{I}_n\) which we are interested in, we can not even describe \(\mathcal{I}_0\) so far. But we only need one more concept to arrive there, namely the notion of a quotient ideal. Given two ideals \(A, B \subset \Pi\) the quotient ideal \(A : B\) is defined as

\[ A : B := \{ p \in \Pi : pB \subseteq A \}, \] (10)
i.e., the set of all polynomials such that their product with any element of \(B\) is contained in \(A\). Since \(A\) is an ideal, we immediately observe that \(A \subset A : B\), and it is also straightforward to see that \(A : B\) is an ideal again. So what is the use of quotient ideals? Very simple, if we use the prime ideal, or maximal ideal as it is usually called in the context of Banach algebras [27],

\[
\langle z - 1 \rangle = \langle z_j - 1 : j = 1, \ldots, s \rangle
\]
of all polynomials vanishing at the “extra point” \((1, \ldots, 1)\) then it immediately follows that

\[
I_0 = \langle z^2 - 1 \rangle : \langle z - 1 \rangle.
\]  
(11)

Why? Just consider the ideal on the right-hand side: it consists of all polynomials which vanish on \(\pm 1\) after being multiplied with an arbitrary polynomial that vanishes at \((1, \ldots, 1)\), and these are precisely the polynomials which vanish on \(e^{\pi i E} = \{ \pm 1 \}^s \setminus \{(1, \ldots, 1)\}\), hence precisely what we want in \(I_0\), so the quotient ideal is the object to go for. This is not surprising as in general the variety associated with a quotient ideal is essentially (it really is the so-called Zariski closure, but for zero-dimensional ideals, i.e., finite varieties, this is the same) the difference of the associated varieties, cf. [6].

With some more effort it is then possible to describe the ideal \(I_n\) of all symbols that allow for polynomial reproduction as follows.

**Theorem 6 ([20]).** For \(n \in \mathbb{N}_0\) we have that

\[
I_n = \left( \langle z^2 - 1 \rangle : \langle z - 1 \rangle \right)^{n+1} = \langle z^2 - 1 \rangle^{n+1} : \langle z - 1 \rangle^{n+1}.
\]  
(12)

Of course, the question is what we can make out of the representation (12) that we cannot get from (7), or, to phrase it differently, where do the ideal bases become useful? The first immediate observation is that in the univariate case the (principal) ideals can simply be divided giving

\[
\left( \langle z^2 - 1 \rangle : \langle z - 1 \rangle \right)^{n+1} = \left( \langle z^2 - 1 \rangle : \langle z - 1 \rangle \right)^{n+1} = \left( \langle z - 1 \rangle \right)^{n+1}
\]

and thus we rediscover the good old factor \(z + 1\). In several variables, however, this is no more a valid manipulation.

It is indeed possible to compute \(H\)-bases for the quotient ideals and thus to identify minimally supported masks. Indeed, bases for \(I_0\) and \(I_1\) can be found for arbitrary \(s\) in [19] while for \(s = 2\) the unique minimal degree elements of \(I_n\) have been identified in [18] as autoconvolutions of three-directions box splines, cf. [1]. More precisely, the minimally supported bivariate masks are

\[
\begin{align*}
a_0^s(z) &= \frac{1}{2} (z_1 + 1) (z_2 + 1) \\
a_1^s(z) &= \frac{1}{4} (z_1 + 1) (z_2 + 1) (z_1 + z_2) \\
a_2^s(z) &= \frac{1}{16} (z_1 + 1)^2 (z_2 + 1)^2 (z_1 + z_2) = \frac{1}{2} a_0^s(z) a_1^s(z) \\
a_3^s(z) &= \frac{1}{32} (z_1 + 1)^2 (z_2 + 1)^2 (z_1 + z_2)^2 = \frac{1}{2} (a_1^s(z))^2 \\
&\vdots \\
a_{2m}^s(z) &= \frac{1}{2} a_0^s(z) (a_1^s(z))^m \\
a_{2m+1}^s(z) &= \frac{1}{2} (a_1^s(z))^{m+1},
\end{align*}
\]

so that the total degree of the “smallest” mask polynomials relative to the order \(n\) of polynomial reproduction is \(\frac{3}{2} n + 3\) if \(m\) is odd and \(\frac{3}{2} n + 2\) if \(n\) is even. The associated refinable functions are either autoconvolutions of the three-directions box spline or autoconvolutions convolved with the characteristic function of the unit square.
Actually, the appearance of convolutions is no surprise: if \( f \) and \( g \) are refinable relative to \( a \) and \( b \), respectively, then
\[
(f \ast g) \hat{\cdot} (\xi) = \hat{f}(\xi)\hat{g}(\xi) = \left[ \hat{a}\left(\frac{\xi}{2}\right)\hat{b}\left(\frac{\xi}{2}\right) \right] \left[ f\left(\frac{\xi}{2}\right)\hat{g}\left(\frac{\xi}{2}\right) \right] \\
= (a \ast b) \hat{\cdot} \left(\frac{\xi}{2}\right) (f \ast g) \hat{\cdot} \left(\frac{\xi}{2}\right),
\]
hence, the convolution of the functions is refinable with respect to the convolution of the masks which also means that the resulting symbol is the product of the symbols. But if \( a \in \mathcal{I}_m \) and \( b \in \mathcal{I}_n \), then it immediately follows (once more from the Leibniz rule) that \( ab \in \mathcal{I}_{m+n} \), so that the degrees of polynomial reproduction add up when taking a convolution of the masks. But note that the univariate strategy of convolving with the characteristic function of the interval to increase the polynomial reproduction by one is far from optimal in the bivariate case as in order to increase the degree of reproduction by one, the degree of the symbol is increased by 2 while the above optimal process only increases it by \( \frac{3}{2} \). This gap is to be expected to increase when the number \( s \) of variables grows as it was shown in [19] that the degree of \( a^*_1 \) is always \( s + 1 \) while that of \( a^*_0 \), representing the characteristic function again, is already \( s \), so that the symbol representing an \( n \)-fold autocorrelation of the characteristic function of the unit cube, an \( s \)-linear tensor product \( B \)-spline, already has degree \( ns \).

5. Differences and a first encounter with the vector case

The quotient ideal formula for \( \mathcal{I}_n \) has another interesting consequence which we point out here only in the case of \( \mathcal{I}_0 \). If \( a^* \in \mathcal{I}_0 = (z^2 - 1) : (z^s - 1) \), it follows that \( pa^* \in (z^s - 1) \) for any \( p \in (z - 1) \), in particular for \( p(z) = z_j - 1 \), \( j = 1, \ldots, s \). Since \( \{z^2 - 1\} \) is an \( H \)-basis, moreover we conclude that there exist sequences \( b_{jk} \in \ell_{00}(\mathbb{Z}^s) \) with symbols \( b^*_jk \in \Pi_{\deg a^*-1} \), therefore of \textit{smaller support}, such that
\[
(z_j - 1) a^*(z) = \sum_{k=1}^{s} b^*_jk(z) (z_k^2 - 1), \quad j = 1, \ldots, s.
\]
Using
\[
[z - 1] = \begin{bmatrix} z_1 - 1 \\ \vdots \\ z_s - 1 \end{bmatrix}, \quad [z^2 - 1] = \begin{bmatrix} z_1^2 - 1 \\ \vdots \\ z_s^2 - 1 \end{bmatrix}
\]
it is only convenient to rewrite (13) in vector form as
\[
[z - 1]a^*(z) = B^*(z) [z^2 - 1],
\]
with the finitely supported matrix sequence \( B \in \ell_{00}^{s \times s}(\mathbb{Z}^s) \) and its symbol \( B^*(z) \in \Pi_{\deg a^*-1}^{s \times s} \). To get another interpretation of this identity, we introduce the \textit{backwards difference operator} \( \nabla : \ell(\mathbb{Z}^s) \rightarrow \ell(\mathbb{Z}^s) \) as
\[
\nabla c = \begin{bmatrix} \nabla_1 c \\ \vdots \\ \nabla_s c \end{bmatrix} = \begin{bmatrix} c(\cdot - \epsilon_1) - c(\cdot) \\ \vdots \\ c(\cdot - \epsilon_s) - c(\cdot) \end{bmatrix}
\]
whose symbol is easily seen to be \( \nabla^*(z) = [z - 1] \). Hence, combining (14) with the fact that \( (S_n c)^* (z) = a^*(z)c^*(z^2) \), we can describe \( a^* \in \mathcal{I}_0 \) also as
\[
\nabla S_n = S_B \nabla,
\]
that is, as a commuting property between the subdivision operators and the difference operator. But even more is true: (15) is equivalent to the fact that the subdivision operator maps constant sequences to constant sequences, hence to \( a^* \in \mathcal{I}_0 \). However, there is a fundamental difference between the univariate case and \( s > 1 \) once more: the univariate
difference operator is surjective, i.e., for any \( \ell \left( \mathbb{Z}^s \right) = \nabla \ell \left( \mathbb{Z}^s \right) \) while for \( s > 1 \) there are conditions that relate the components of \( \nabla c \) and so make \( \nabla \ell \left( \mathbb{Z}^s \right) \) a proper subspace of \( \ell^s \left( \mathbb{Z}^s \right) \): since

\[
(\nabla c)^v_j (z) = (z_j - 1) e^v(z), \quad j = 1, \ldots, s,
\]

it immediately follows that

\[
(z_k - 1) (\nabla c)^v_j (z) - (z_j - 1) (\nabla c)^v_k (z) = 0, \quad j, k = 1, \ldots, s.
\]  

(16)

In algebraic terms, these dependency relations are known as syzygies, but here they just tell us that most vector sequences are not in the image of the difference operator. This has its consequences when the convergence of subdivision schemes is considered since it requires the determination of a restricted spectral radius in order to characterize convergence, see [3,11,23].

This also explains the seemingly contradictitory fact that usually there are plenty of choices for \( B^s \in \Pi_{\deg a^s - 1} \) such that (14) is satisfied as these different choices only are considered on the image of \( \nabla \) or, in other words, they are only seen through the differences.

For arbitrary \( n \in \mathbb{N} \) the decomposition with respect to the difference operators works practically the same way! The ideal \( \left( z^2 - 1 \right)^n \) is generated by the basis elements

\[
\left( z^2 - 1 \right)^{\alpha} = \left( z_1^2 - 1 \right)^{\alpha_1} \cdots \left( z_s^2 - 1 \right)^{\alpha_s}, \quad |\alpha| \leq n,
\]

hence,

\[
\left( z^2 - 1 \right)^n = \left( \left( z^2 - 1 \right)^{\alpha} : |\alpha| \leq n \right)
\]

\[
\langle z - 1 \rangle^n = \left( \langle z - 1 \rangle^{\alpha} : |\alpha| \leq n \right),
\]

and like in (13) we have that

\[
\langle z - 1 \rangle^{\alpha} a^v(z) = \sum_{|\beta| = n} b^{v}_{\alpha \beta}(z) \left( z^2 - 1 \right)^{\beta}, \quad |\alpha| = n,
\]

or, in other words,

\[
\nabla^n S = S_B \nabla^n, \quad \nabla^n = \left[ \nabla^\alpha = \nabla^\alpha_1 \cdots \nabla^\alpha_s : |\alpha| = n \right].
\]

In summary, the notion of ideals and the use of multiindices allow us to write polynomial generation properties of \( f \) with respect to the mask in a compact and relatively simple way, although conceptually the situation is more intricate. The results themselves look very much like what they are supposed to be.

Nevertheless, there is also a connection to analysis. To explore it, we return to the \( a \)-refinable function \( f \), suppose that it is differentiable and consider the gradient

\[
\nabla f = \left[ \frac{\partial}{\partial x_j} f : j = 1, \ldots, s \right],
\]

more precisely, its Fourier transform

\[
(\nabla f)^\wedge (\xi) = \left[ i \xi_j \widehat{f}(\xi) : j = 1, \ldots, s \right] = [i \xi] \hat{f}(\xi).
\]

The gradient defines a diagonal matrix \( \hat{f}(\xi) D(i \xi) \), where we will use \( D(x) \) for the diagonal matrix defined by the vector \( x \in \mathbb{R}^s \) and it satisfies

\[
D(i \xi) \hat{f}(\xi) = \frac{1}{2^s} D(i \xi) \hat{\alpha} \left( \frac{\xi}{2} \right)
\]

\[
= \frac{1}{2^s} D \left( \frac{i \xi}{e^{i \xi/2} - 1} \right) \hat{f} \left( \frac{\xi}{2} \right) D \left( e^{i \xi/2} - 1 \right) \hat{\alpha} \left( \frac{\xi}{2} \right)
\]

\[
= \frac{1}{2^s} D \left( \frac{i \xi}{e^{i \xi/2} - 1} \right) \hat{f} \left( \frac{\xi}{2} \right) D \left( \frac{\xi}{2} \right) \left[ e^{i \xi} - 1 \right],
\]

\[
D \left( \frac{i \xi}{e^{i \xi/2} - 1} \right) \hat{f} \left( \frac{\xi}{2} \right) D \left( \frac{\xi}{2} \right) \left[ e^{i \xi} - 1 \right],
\]
hence, after right multiplication with \( D \left( e^{i\xi} - 1 \right)^{-1} \)
\[
D \left( \frac{i\xi}{e^{i\xi} - 1} \right) \hat{f}(\xi) = \frac{1}{2^s} D \left( \frac{i\xi}{e^{i\xi/2} - 1} \right) \hat{f} \left( \frac{\xi}{2} \right) D \left( \frac{\xi}{2} \right) \left[ e^{i\xi} - 1 \right] D \left( e^{i\xi} - 1 \right)^{-1},
\]
so that transposition and multiplication of the above identity with the column vector \( \mathbf{1} = [1, \ldots, 1] \) yields that the modified gradient \( \hat{g}(\xi) = \left[ \frac{i\xi}{e^{i\xi} - 1} \right] \hat{f}(\xi) \) satisfies the refinement equation
\[
\hat{g}(\xi) = \frac{1}{2^s} D \left( e^{i\xi} - 1 \right)^{-1} D \left( \left[ e^{i\xi} - 1 \right]^T \hat{B} \left( \frac{\xi}{2} \right) \right) \hat{g} \left( \frac{\xi}{2} \right). \tag{18}
\]
Since the backward differences satisfy \( \nabla_j^a (z) = z_j - 1 \), the function \( g \) above has a simple explanation, namely as \( g = \left[ \nabla_j^{-1} \frac{\partial}{\partial x_j} f : j = 1, \ldots, s \right] \). Also recall that the inverse of a difference operator is a summation operator.

In the case \( s = 1 \) (18) immediately becomes \( \hat{g}(\xi) = \frac{1}{2} \hat{B} \left( \xi/2 \right) \hat{g}(\xi/2) \), i.e., the function \( g \) is \( b \)-refinable. In the multivariate case, however, the extra terms in (18) indicate that \( g \) is not “fully” refinable relative to \( B \). Indeed, this can be understood as restricted refinability, that is, as a form of refinability that only works on the image of the difference operator \( \nabla \). For details, see [23].

In conclusion, the algebra is an extra effort to be incorporated into this theory, but it is natural and it allows us to understand the truly multivariate features that we encounter with refinable functions in several variables. But this is enough on algebra for the time being, we now turn our interest back to more elementary properties of the mask.

6. Vector subdivision and the rank

The preceding chapter, in particular the property \( \nabla S_0 = S_B \nabla \) created a direct connection between scalar subdivision schemes and vector subdivision schemes based on matrix masks. These schemes are not only of interest as “derivative schemes” or better difference schemes in view of \( \nabla S_0 = S_B \nabla \) and the above refinable gradient function, they are also interesting subdivision schemes in their own way.

So now the mask is a sequence \( A \in \ell^N_{00} (\mathbb{Z}^s) \) which acts either on \( N \)-vectors \( c \in \ell^N (\mathbb{Z}^s) \) or even on matrix valued sequences \( C \in \ell^N \times N (\mathbb{Z}^s) \), as
\[
S_A c = \sum_{\alpha \in \mathbb{Z}^s} A (\cdot - 2\alpha) c (\alpha) \quad \text{or} \quad S_A C = \sum_{\alpha \in \mathbb{Z}^s} A (\cdot - 2\alpha) C (\alpha)
\]
respectively, where, by acting on the columns of \( C \) separately, we can always restrict ourselves to the vector case. If the subdivision scheme converges, it has limits of the form
\[
f_c = F \ast c = \sum_{\alpha \in \mathbb{Z}^s} F (\cdot - \alpha) c (\alpha),
\]
where \( F \) is an \( N \times N \)-matrix valued function and, in addition, refinable with respect to \( A \), which can be written as
\[
F = F \ast A (2\cdot) = \sum_{\alpha \in \mathbb{Z}^s} F (2\cdot - \alpha) A (\alpha).
\]
At this point, a short word of warning seems appropriate: all the extensions so far are and appear quite straightforward and formal, and it seems that the passing from scalar univariate subdivision to multivariate vector subdivision is entirely a matter of replacing indices by Greek letters and using uppercase boldface ones for the symbol and its coefficients. However, there is a major difference that comes from the fact that in general matrices do not commute, in fact commuting families of matrices are a very rare thing, see [13]. Nevertheless, many of the identities in the preceding chapters still hold for the vector case as well, even without commuting, which was actually the reason why they were written in that slightly nonstandard way.

There is a simple but important equivalence between matrix subdivision schemes and matrix refinement equations, namely similarity. For any nonsingular matrix \( T \in \mathbb{R}^{N \times N} \), the similarity transform maps \( A \in \ell^N_{00} (\mathbb{Z}^s) \) to \( A' \) by taking
\[
A' = T^{-1} AT = \left( T^{-1} A (\alpha) T : \alpha \in \mathbb{Z}^s \right).
\]
Since \( S_A' = S_{T^{-1}AT} = T^{-1}S_A'T \), the scheme \( S_A' \) converges if and only if the original one converges and the limit function is \( F' \ast c \), where \( F' = T^{-1}FT \) is easily seen to be refinable with respect to \( A' \). Among the equivalence class of similar schemes we can always choose the reprenter that is most appropriate for our purposes.

Much of what we considered before could also be seen as the question when the subdivision operator is able to (re)produce polynomial sequences of a certain degree. The simplest form of this question is when the subdivision operator preserves constant sequences, i.e., when \( S_Ac = c \) for a constant sequence \( c \). By straightforward computations we thus get for \( \alpha \in \mathbb{Z}^s \) and the unique \( \epsilon \in \{0, 1\}^s \cap (\alpha + 2\mathbb{Z}^s) \) that

\[
c = c(\alpha) = S_Ac(\alpha) = \sum_{\beta \in \mathbb{Z}^s} A(\alpha - 2\beta) c = \left( \sum_{\beta \in \mathbb{Z}^s} A (\epsilon + 2\beta) \right) c.
\]

Consequently, a constant sequence \( c \) is reproduced if and only if it is a joint eigenvector of the \( 2^s \) matrices

\[
A_\epsilon = \sum_{\beta \in \mathbb{Z}^s} A (\epsilon + 2\beta), \quad \epsilon \in \{0, 1\}^s,
\]

with respect to the eigenvalue 1. Thus, it makes sense to consider the joint one eigenspaces

\[
\mathcal{E}(A) = \left\{ y \in \mathbb{R}^N : A_\epsilon y = y, \epsilon \in \{0, 1\}^s \right\}
\]

and call the dimension of \( \mathcal{E}(A) \) the rank of \( A \), a quantity introduced (for the univariate case) in [14,15]. This number can vary between 0 and \( N \), but the first case has to be excluded as it corresponds to no convergent subdivision scheme—any convergent subdivision scheme has its rank between 1 and \( N \).

The rank of a subdivision scheme leads to a normal form of the mask. To that end, suppose that the subdivision scheme has rank \( r \leq N \) and let \( \mathcal{Y} = \{y_1, \ldots, y_r\} \) be an orthonormal basis of \( \mathcal{Y} \) which can be completed to an orthonormal basis \( \mathcal{V} = \{v_1, \ldots, v_N\} \) of \( \mathbb{R}^N \), i.e., \( v_j = y_j \), \( j = 1, \ldots, r \). The matrix \( V \) with columns \( v_1, \ldots, v_N \) is an orthogonal matrix and it follows that

\[
V^T A_\epsilon V = \begin{bmatrix} I & 0 \\ 0 & * \end{bmatrix}, \quad \epsilon \in \{0, 1\}^s,
\]

(19)

where the lower right part is relatively unimportant, cf. [21]. This property can again be expressed in terms of quotient ideals (no surprise), but things are slightly different. To that end, we assume that \( A \) already is in the “normal form” (19), so that we can choose \( V = I \), and again look at the symbol on \( \{-1, 1\}^s \). Making use of the identity

\[
A^s(\zeta) = \sum_{\epsilon \in \{0,1\}^s} \zeta^\epsilon A_\epsilon^s \left( \zeta^2 \right), \quad A_\epsilon^s(\zeta) = \sum_{\alpha \in \mathbb{Z}^s} A (\epsilon + 2\alpha) \zeta^\alpha,
\]

we find for \( \zeta \in \{-1, 1\}^s \) that

\[
A^s(\xi) = \sum_{\epsilon \in \{0,1\}^s} \xi^\epsilon A_\epsilon^s \left( \xi^2 \right) = \sum_{\epsilon \in \{0,1\}^s} \xi^\epsilon A_\epsilon = \begin{cases} \begin{bmatrix} 2I & 0 \\ 0 & * \end{bmatrix}, & \xi = (1, \ldots, 1), \\ \begin{bmatrix} 0 & 0 \\ 0 & * \end{bmatrix}, & \text{otherwise}, \end{cases}
\]

which meanwhile we know how to express in terms of ideals: the diagonal elements in the upper left part vanish on \( e^{i\pi E} \) and thus belong to \( \{z^2 - 1\} : \{z - 1\} \) while the off-diagonal there and all elements in the upper right and lower left quadrant vanish on all of \( \{\pm 1\}^s \) and thus belong to \( \{z^2 - 1\} \). This can be extended to higher-order polynomial (re)production, but the resulting matrices become quite complicated and, surprisingly, the upper right and the lower left block develop differently. Readers of a sufficiently masochistic nature can find the formula in [21, Proposition 6].

There is one more subtle difference between the scalar and the vector case and it concerns a necessary condition for convergence due to [7], requiring that

\[
(A_\epsilon - I) F = 0.
\]

(20)
In the scalar case this just means, as soon as \( f \neq 0 \), that
\[
\sum_{a \in \mathbb{Z}^s} a (\epsilon + 2a) = 1, \quad \epsilon \in \{0, 1\}^s,
\]
as it can be found in [2], while in the vector case it says that
\[
\mathcal{R}(F) \subset \mathcal{E}(A), \quad \mathcal{R}(F) = \bigcup_{x \in \mathbb{R}^s} F(x)\mathbb{R}^N,
\]
i.e., the range of \( F \) as a finite-dimensional subspace of \( \mathbb{R}^N \) must be contained in \( \mathcal{E}(A) \). Consequently, the connection between the function and the mask is a much stronger one, already in this context.

7. Special ranks

Like there are essentially three values of \( p \) in \( L_p \)-theory, namely \( p = 1, 2, \infty \) (one seldom finds papers dealing with specific results in \( L_{3\frac{1}{3}, 1592} \), for instance) there are two major special cases in vector subdivision: rank-1 schemes and full rank schemes.

Rank-1 schemes have been considered extensively in the context of multiwavelets, i.e., for constructing multiresolution analyses that are generated by a finite number of functions which are jointly refinable, \( B \)-splines with multiple knots being the most prominent univariate example in this respect. In fact, it was proved in [7] that the existence of a stable vector solution of the refinement equation, that is, the existence of a vector function \( f \) such that
\[
f^T = f^T \ast A (2 \cdot) \quad \text{and} \quad \| c \|_{\ell_\infty} \simeq \| f^T \ast c \|_{L_\infty},
\]
already implies that \( A \) must be of rank 1. The matrix function associated with this \( f \) is \( F = vf^T, v \in \mathbb{R}^N \setminus \{0\} \), a matrix function with all rows being linearly dependent. Note that \( F \) is also stable in the matrix sense, \( \| c \|_{\ell_\infty} \simeq \| F \ast c \|_{L_\infty} \), since even any nonzero row is stable as a vector function. The converse, however, is not true: there exist stable matrix functions whose individual rows are not stable! A discussion of Strang–Fix conditions for the multivariate rank-1 situation with a lot of further information can be found in the survey [9].

The other extreme case is full rank schemes where \( A_\epsilon = I, \epsilon \in \{0, 1\}^s \). Such schemes are surprisingly scalar in appearance, though in many cases the proofs are not. The most prominent instance of full rank schemes are interpolatory schemes which are characterized by the fact that \( S_\delta (2 \cdot) = c \), or, equivalently, \( F (2 \cdot) = \delta I \). Since the associated limit function is cardinal, \( F (\alpha) = \delta (\alpha) I, \alpha \in \mathbb{Z}^s \), the necessary condition (20) leaves only the full rank choice for \( A \). For an investigation of full rank interpolatory schemes see [4].

The cardinal function above is the special case of a full rank function which is an \( N \times N \) matrix function \( F \) with the property that \( \mathcal{R}(F) = \mathbb{R}^N \). These functions admit the characterizations of polynomial generation that very well match the scalar case. As an example, I just want to mention the following result from [5] which provides Strang–Fix conditions for full rank functions. The only side condition, replacing the \( (f \ast 1) (0) \neq 0 \) from above is that
\[
F_0 = (F \ast I) (0) = \sum_{a \in \mathbb{Z}^s} F(\alpha)
\]
is a nonsingular (and not nonzero!) matrix—which immediately implies full rank! If \( F_0 \) is nonsingular, then the normalized matrix valued function \( F_\ast = FF_0^{-1} \) is well-defined and takes the leading role in the following version of the Strang–Fix conditions.

**Theorem 7.** If \( F_0 \) is nonsingular then \( \mathcal{S}(F) \) contains all \( N \)-vector valued polynomials of degree \( n \) if and only if

1. \( \hat{F}_\ast (0) = (2\pi)^2 I, \)
2. \( D^\beta \hat{F}_\ast (2\pi \alpha) = 0, |\beta| \leq n, \alpha \in \mathbb{Z}^s \setminus \{0\} \).

This is, in my opinion at least, quite a similar and elegant formula compared to the Strang–Fix conditions from [9], but the simplicity simply comes from the situation: full rank is a good deal closer to the scalar case than rank 1. And everything in between is just a mess anyway . . .
Acknowledgements

I want to thank Costanza Conti for a careful reading of the first version of the manuscript and her comments which were – as usual – helpful and stimulating. I am also indebted to the people from Rome and Reggio who patiently endured my lecture based on this paper.

References