THE ALGEBRAIC STRUCTURE IN SIGNAL PROCESSING: TIME AND SPACE

Markus Püschel and José M. F. Moura

Electrical and Computer Engineering Carnegie Mellon University Email: {pueschel,moura}@ece.cmu.edu

ABSTRACT

The assumptions underlying linear signal processing (SP) produce more structure than vector spaces. We capture this structure by describing the space of filters as an algebra and the space of signals as the associated module. We formulate an algebraic approach to SP that is axiomatically based on the concept of a *signal model*. Signal models for time are visualized as directed graphs. We construct corresponding models for undirected graphs, which we hence call space models, and show that, in particular, the 16 DCTs and DSTs are Fourier transforms for these finite space models. Finally, we discuss the extension of our theory to separable and nonseparable 2-D SP.

1. INTRODUCTION

Linear signal processing (henceforth simply called SP) is a welldeveloped theory and a comprehensive set of tools for processing 1-D time signals, continuous or discrete, of infinite or finite duration. Each of these four cases has specific notions of filtering, Fourier transform, and other key concepts needed in SP.

Many other transforms are used in SP, most prominently the class of trigonometric transforms that include the discrete cosine and sine transforms (DCTs and DSTs). These compute a spectrum of sorts, so conceptually they are Fourier transforms; but, questions remain: for which type of signals? what are their infinite counterparts? and, just as for the DFT, what are the associated notions of filtering, convolution, and spectrum? The current derivation of the DCTs based on random processes, [1], or their interpretation as Karhunen-Loève transforms of random fields, [2], do not answer these questions.

In 2-D, SP is usually developed in a separable way, which assumes the signal resides on a rectangular lattice. For many applications, this is appropriate. Mersereau et al. [3] provide some answers for more general lattices. We would like, for example, to explore generic questions including what are the DCT equivalents on nonseparable lattices.

In this paper, we give answers to these questions in an overview on recent work [4, 5]. Our approach is in two steps. First, we identify the algebraic structure inherent to SP that goes beyond vector spaces to algebras and modules. We argue that polynomial algebras are key concepts in SP. Second, we formulate a general framework for SP, based on the notion of a *signal model* that we define as a triple of an algebra of filters, a module of signals, and a mapping Φ that generalizes the z-transform. Once a model is given, other concepts such as the Fourier transform are automatically defined. We describe the signal models for infinite and finite time SP and visualize them as directed graphs. Then we identify the models underlying the DCTs and DSTs and show that their visualizations are undirected. We thus call them space models. Finally, we briefly discuss the extension of our approach to 2-D SP for separable and nonseparable time and space models.

Organization. We introduce the concept of algebra and module in Section 2. In Section 3, we motivate why linear SP is algebraic in nature and introduce the concept of a signal model that underlies our algebraic approach. In Sections 4 and 5, we show the signal models for infinite and finite SP in time and space. Section 6 briefly discusses the extension to separable and nonseparable 2-D SP. Finally, section 7 concludes the paper.

2. BACKGROUND: ALGEBRAS AND MODULES

In this section, we introduce two concepts from abstract algebra: Algebra and module. These are needed to identify the algebraic structure in linear signal processing. Both are vector spaces, but possess additional structure. In Section 3, we will explain that in linear signal processing filter spaces are algebras and signal spaces are modules. For every given pair of algebra and module, there is an associated notion of Fourier transform that we introduce. We focus on regular modules (the algebra and the module are the same set but with different properties) of polynomial algebras, which are particularly relevant in signal processing, as we will explain later.

Algebra (filter space). An *algebra* \mathcal{A} is a vector space that is also a ring, i.e., it permits multiplication and the distributivity law holds. Examples include the field of complex numbers \mathbb{C} , the space of polynomials in one variable $\mathbb{C}[x]$, and the space of Laurent series in x with absolute summable coefficient sequences. Another example, particularly important in SP, is introduced next. Let $p(x) \in \mathbb{C}[x]$ be a fixed polynomial of degree $\deg(p) = n$, then the set of polynomials

$$\mathbb{C}[x]/p(x) = \{q(x) \in \mathbb{C}[x] \mid \deg(q) < \deg(p)\}$$
(1)

of degree smaller than n with addition and multiplication modulo p is called a *polynomial algebra*. In $\mathbb{C}[x]/p(x)$, p(x) = 0 holds. As a vector space, $\mathbb{C}[x]/p(x)$ has dimension $\deg(p) = n$. We write elements of \mathcal{A} as h, the common letter used for filters in SP.

Module (signal space). Assume that an algebra \mathcal{A} has been chosen. An \mathcal{A} -module, \mathcal{M} is a vector space that permits an operation of the \mathcal{A} on \mathcal{M} , written as "." (multiplication), such that the distributivity law holds (as well as some other properties that we omit). Formally, for $h \in \mathcal{A}$, $s, s_1, s_2 \in \mathcal{M}$,

$$h \cdot s \in \mathcal{M}, \quad h \cdot (s_1 + s_2) = h \cdot s_1 + h \cdot s_2.$$

For a given algebra, one choice of module is the *regular module* $\mathcal{M} = \mathcal{A}$ with the operation of \mathcal{A} on \mathcal{M} being the multiplication in \mathcal{A} . Almost all modules considered in this paper are regular.

We will focus on algebras and modules that are \mathbb{C} -vector spaces in this paper, but the theory extends to other basefields. We denote signals in \mathcal{M} as *s*.

This work was supported by NSF through award 0310941.

Fourier transform. Assume that an algebra \mathcal{A} and an \mathcal{A} -module \mathcal{M} are given. Then the associated Fourier transform (if it exists) is a linear mapping¹ that decomposes \mathcal{M} into a direct sum of smallest submodules invariant under \mathcal{A} . We omit a formal definition, as we are only interested in one special case discussed next.

Let $\mathcal{A} = \mathbb{C}[x]/p(x)$ and let $\mathcal{M} = \mathcal{A}$ be the regular module. Further, assume p(x) is of degree n and separable, i.e., it has pairwise distinct zeros $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$. Then, the associated Fourier transform is given by the Chinese remainder theorem as

$$\Delta : \mathbb{C}[x]/p(x) \to \mathbb{C}[x]/(x-\alpha_0) \oplus \ldots \oplus \mathbb{C}[x]/(x-\alpha_{n-1}),$$

$$s(x) \mapsto (s(\alpha_0), \ldots, s(\alpha_{n-1})).$$

Since Δ is a linear mapping, we can express it as a matrix w.r.t. chosen bases. Let $b = (p_0, \ldots, p_{n-1})$ be a basis of \mathcal{M} , and choose (x^0) as basis in each $\mathbb{C}[x]/(x - \alpha_k)$. Then Δ corresponds to

$$\mathcal{P}_{b,\alpha} = [p_{\ell}(\alpha_k)]_{0 \le k, \ell < n},\tag{2}$$

which we also call a Fourier transform for the \mathcal{A} -module \mathcal{M} with basis b. Choosing arbitrary bases $(a_k x^0)$, $a_k \in \mathbb{C}$, in $\mathbb{C}[x]/(x - \alpha_k)$ yields the most general form of Fourier transform for \mathcal{M} (with basis b):

$$\mathcal{F} = \operatorname{diag}_{0 \le k \le n} (1/a_k) \mathcal{P}_{b,\alpha}.$$
(3)

3. THE ALGEBRAIC STRUCTURE IN SIGNAL PROCESSING

This section introduces the *signal model* that formalizes the connection between algebras and modules and SP.

SP is algebraic. We identify the basic assumptions imposed on the filter and signal space in SP. Clearly, the set of filters is a vector space (scalar multiplication = amplification by $\alpha \in \mathbb{C}$; addition = parallel connection) but is even an algebra \mathcal{A} , since sequential connection of filters, i.e., multiplication of filters, is permissible. The set of signals is also a vector space in linear SP. Further, filters operate on this space, i.e., filtering a signal produces a signal, making the signal space an \mathcal{A} -module \mathcal{M} , where \mathcal{A} is the chosen set of filters.

Signal model. Signals do not arise as elements of modules, but as sequences **s** of numbers. We treat the discrete case, where signals arise as infinite or finite sequences over an index domain *I*. Examples include infinite signals $\mathbf{s} \in \ell^2(\mathbb{Z}) \leq \mathbb{C}^{\mathbb{Z}}$, or finite signals in $\mathbf{s} \in \mathbb{C}^n$. In a natural way, one can impose a vector space structure on these sequences; in contrast, there may be many possible choices of filter space \mathcal{A} and filtering operation. To assign such a choice, we introduce the concept of signal model.²

Definition 1 (Signal model) Let $V \leq \mathbb{C}^I$ be a vector space. We call the triple $(\mathcal{A}, \mathcal{M}, \Phi)$ a signal model for V, if \mathcal{A} is an algebra, \mathcal{M} an \mathcal{A} -module of the same dimension as V and with basis $b = (b_\ell)_{\ell \in I}$, and Φ is the bijective (one-to-one) mapping

$$\begin{array}{rcl} \Phi : & V & \to & \mathcal{M}, \\ \mathbf{s} = (s_{\ell})_{\ell \in I} & \mapsto & s = \sum_{\ell \in I} s_{\ell} b_{\ell} \in \mathcal{M} \end{array}$$

This definition shows that, with the chosen signal model, **s** becomes the coordinate vector of $s \in \mathcal{M}$ w.r.t. the chosen basis *b* of \mathcal{M} . The bijective mapping Φ generalizes the idea of a *z*-transform



Fig. 1. Visualization of the infinite discrete time model (4).

as explained below. Once a signal model is chosen, filtering is well-defined (the operation of \mathcal{A} on \mathcal{M}) and the main ingredients for SP follow: spectrum, Fourier transform (see Section 2), frequency response, as well as others [4].

Example. Consider the space of finite energy sequences $\mathbf{s} \in V = \ell^2(\mathbb{Z})$. A signal model for V is given by

$$\mathcal{A} = \{ \sum_{\ell \in \mathbb{Z}} h_{\ell} x^{\ell} \mid (h_{\ell}) \in \ell^{1}(\mathbb{Z}) \}, \\ \mathcal{M} = \{ \sum_{\ell \in \mathbb{Z}} h_{\ell} x^{\ell} \mid (h_{\ell}) \in \ell^{2}(\mathbb{Z}) \}, \\ \Phi = V \to \mathcal{M}, \ \mathbf{s} \mapsto s = s(x) = \sum_{\ell \in \mathbb{Z}} s_{\ell} x^{\ell}.$$

$$(4)$$

The mapping Φ is the ordinary *z*-transform³. For this reason we call this signal model the infinite discrete time model. Note that the module is not regular, since $\mathcal{A} \neq \mathcal{M}$.⁴

Shift. An important concept in our general framework is the shift operator. We formally define the shift as the chosen generator of the algebra \mathcal{A} . For example, the 1-D model (4) has only one shift x. With this definition, we can classify all models that lead to shift-invariant SP as those with commutative algebras. If the algebra has only one shift, it is necessarily commutative as in the case of (4).

In the case of finite sequences, e.g., $\mathbf{s} \in \mathbb{C}^n$, \mathcal{M} and \mathcal{A} are finite-dimensional. Finite-dimensional commutative algebras are precisely the polynomial algebras; in the case of one shift x they have the form in (1).

In the remainder of the paper, we identify several signal models that occur in SP. Lacking space, we sketch only their derivation and focus on providing intuition about them. A detailed exposition is found in [4, 5].

4. 1-D TIME MODELS

Infinite time. In (4), we presented the signal model adopted for infinite discrete time. The property of modeling *time* is best explained by the operation of the shift x on the basis $(x^{\ell})_{\ell \in \mathbb{Z}}$ of the signal module \mathcal{M} : $x \cdot x^{\ell} = x^{\ell+1}$. The operation is graphically displayed in Fig. 1. We call the graph a *visualization* of the signal model (4). The graph is directed, an inherent property of time.

Finite time. For finite time, the model usually adopted is

$$\mathcal{A} = \mathbb{C}[x]/(x^n - 1),$$

$$\mathcal{M} = \mathbb{C}[x]/(x^n - 1),$$

$$\Phi = \mathbb{C}^n \to \mathcal{M}, \ \mathbf{s} \mapsto s = s(x) = \sum_{0 \le \ell < n} s_\ell x^\ell.$$
(5)

We call Φ the finite z-transform. Filtering in this model is multiplication of polynomials modulo $x^n - 1$, which is equivalent to circular convolution [6]. The Fourier transform is obtained from (2) and is the DFT_n. The periodic boundary condition, and the periodic signal extension, known to be associated with the DFT, are encoded in the polynomial $x^n - 1$. Namely, $x^n - 1 = 0$ implies $x^n = 1$. The visualization of the model is the circle in Fig. 2. It is a time model since the graph is directed.

¹More specifically an \mathcal{A} -module homomorphism.

²The definition can be stated without explicitly giving the basis, but it is more intuitive for the purpose of this paper in this form.

 $^{{}^{3}}$ We set $x = z^{-1}$ for notational convenience later when considering more general models.

⁴Convolving two ℓ^2 sequences does in general not yield an ℓ^2 sequence.



Fig. 2. Visualization of the finite discrete time model (5).

Table 1. Four types of Chebyshev polynomials. The range for the zeros is $0 \le k < n$. In the trigonometric closed form $\cos \theta = x$.

	n=0,1	closed form	symmetry	zeros
T_n	1, x	$\cos(n\theta)$	$T_{-n} = T_n$	$\cos \frac{(k+\frac{1}{2})\pi}{n}$
U_n	1, 2x	$\frac{\sin(n+1)\theta}{\sin\theta}$	$U_{-n} = -U_{n-2}$	$\cos \frac{(k+1)\pi}{n+1}$
V_n	1, 2x - 1	$\frac{\cos(n+\frac{1}{2})\theta}{\cos\frac{1}{2}\theta}$	$V_{-n} = V_{n-1}$	$\cos\frac{(k+\frac{1}{2})\pi}{n+\frac{1}{2}}$
W_n	1, 2x + 1	$\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}$	$W_{-n} = -W_{n-1}$	$\cos\frac{(k+1)\pi}{n+\frac{1}{2}}$

5. 1-D SPACE MODELS

The models in the previous section were termed time models; their graph visualizations are directed. In this section, we consider models with no inherent direction, i.e., visualized by undirected graphs. We call such models space models.⁵ It turns out that the 16 DCTs and DSTs are Fourier transforms for space models.

$$\begin{array}{c} \bullet \stackrel{\frac{1}{2}}{\longleftarrow} \bullet \stackrel{\frac{1}{2}}{\longrightarrow} \bullet \\ p_{\ell-1} \quad p_{\ell} \quad p_{\ell+1} \end{array}$$

Fig. 3. The space shift $x \cdot p_{\ell}$.

Infinite case. To construct a space model, we choose the *symmetric* shift operator x shown in Fig. 3. We now consider the polynomials p_{ℓ} . Fig. 3 implies $xp_{\ell} = (p_{\ell-1} + p_{\ell+1})/2$ or

$$p_{\ell+1} = 2xp_{\ell} - p_{\ell-1},\tag{6}$$

which is exactly the recurrence for the Chebyshev polynomials [7]. The exact form is determined by the choice of p_0 and p_1 as initial polynomials of degree 0 and 1. Then p_{ℓ} is a polynomial of degree ℓ . Running (6) in the other direction yields also polynomials for the negative indices. Thus, only the polynomials with positive subscripts are linearly independent: The model is for rightsided sequences and has a left boundary. The left side (negative subscript) is the signal extension. We normalize $p_0 = 1$ and choose p_1 such that the resulting signal extension is simple, i.e., has a symmetry. This happens exactly in the four cases shown in Table 1 (see the symmetry column).⁶

In summary, we get four infinite space models for rightsided sequences, which we collectively state as follows. To ensure convergence in all cases, we require $V = \ell^1(\mathbb{N})$.

$$\mathcal{A} = \{ \sum_{\ell \in \mathbb{N}} h_{\ell} T_{\ell} \mid (h_{\ell}) \in \ell^{1}(\mathbb{N}) \},$$

$$\mathcal{M} = \{ \sum_{\ell \in \mathbb{N}} h_{\ell} C_{\ell} \mid (h_{\ell}) \in \ell^{1}(\mathbb{N}) \},$$

$$\Phi = V \to \mathcal{M}, \mathbf{s} \mapsto s = s(x) = \sum_{\ell \in \mathbb{N}} s_{\ell} C_{\ell}.$$
(7)

We call Φ the C-transform for $C \in \{T, U, V, W\}$, i.e., C is one of the four Chebyshev polynomials in Table 1. There is a degree of



Fig. 4. Visualization of the four infinite space models (7).



Fig. 5. Visualization of the finite discrete space model (8) associated with the $DCT-2_n$.

freedom regarding the basis to choose in \mathcal{A} . Because of the known property $T_k C_{\ell} = (C_{\ell-k} + C_{\ell+k})/2$, we choose the *T* polynomials for simpler computation. Filtering, or convolution, in this model is the multiplication

$$h \cdot s = \left(\sum_{\ell \in \mathbb{N}} h_{\ell} T_{\ell}\right) \left(\sum_{\ell \in \mathbb{N}} s_{\ell} C_{\ell}\right),$$

which yields again a series in C_{ℓ} , or a signal.

The four models are visualized in Fig. 4. The graphs are undirected (since these are space models) and the left boundary condition (third column in Table 1) is expressed through redirecting the arrow that would go to C_{-1} , based on the boundary condition (second column in Table 1).

Finite space. We introduce the finite space models using an example. We start with the infinite V-transform (7), which implies a left boundary condition of $V_{-1} = V_0$. We cut the basis to length $n: b = (V_0, \ldots, V_{n-1})$ and introduce (choose) the mirrored boundary condition on the right side: $V_n = V_{n-1}$ or $V_n - V_{n-1} = 2(x - 1)U_{n-1} = 0$ (the 2 can be omitted). This yields the signal model (for $V = \mathbb{C}^n$)

$$\mathcal{A} = \mathbb{C}[x]/(x-1)U_{n-1},$$

$$\mathcal{M} = \mathbb{C}[x]/(x-1)U_{n-1},$$

$$\Phi = \mathbb{C}^n \to \mathcal{M}, \mathbf{s} \mapsto s = s(x) = \sum_{\ell \in \mathbb{N}} s_{\ell} V_{\ell}.$$
(8)

The model is visualized in Fig. 5.

The Fourier transform (2) for this model is given by

$$\mathcal{P}_{b,\alpha} = \left[\frac{\cos k(\ell+1/2)\pi/n}{\cos k\pi/(2n)}\right]_{0 \le k, \ell < n},$$

using Table 1 to find the zeros $\alpha_k = \cos(k + 1/2)\pi/n$. Proper scaling shows that the DCT, type 2, is also a Fourier transform for this model of the form (3):

DCT-2_n = diag_{0 < k < n} (cos
$$k\pi/(2n)$$
) $\mathcal{P}_{b,\alpha}$.

All possible choices of the left boundary conditions in Fig. 7 with their mirrored versions for the right side yield a total of 16 finite space models, corresponding to the 16 types of DCTs and

⁵This differs from current practice in which the distinction between time and space is one of 1-D versus 2-D.

⁶See [5] for a more rigorous explanation of "simple" signal extension.

C $C_n - C_{n-2}$ C_n $C_n - C_{n-1}$ $C_n + C_{n-1}$ DCT-1 DCT-3 DCT-7 TDCT-5 $(x^2 - 1)U_{n-2}$ T_n $(x-1)W_{n-1}$ $(x+1)V_{n-1}$ U DST-3 DST-1 DST-7 DST-5 T_n U_n V_n W_n DCT-6 DCT-8 DCT-2 DCT-4 V $(x-1)W_{n-1}$ V_n $(x-1)U_{n-1}$ $2T_n$ DST-6 DST-2 DST-8 DST-4 W $2T_n$ $(x+1)V_{n-1}$ W_n $(x+1)U_{n-1}$

Table 2. Signal models associated with the 16 DCTs and DSTs.

DSTs. An overview is provided in Table 2. For a given DCT or DST, let p(x) be the polynomial below it, and let $C \in \{T, U, V, W\}$ be the type of Chebyshev polynomial given in the first column. Then the associated model is given by $\mathcal{A} = \mathcal{M} = \mathbb{C}[x]/p(x)$, and $\Phi: \mathbf{s} \mapsto \sum_{0 \le \ell < n} s_{\ell} C_{\ell}$, which we call a finite *C*-transform. The choice of *C* encodes the left b.c., and the choice of a column fixes the right b.c. In some cases a scaling diagonal is necessary, i.e., the DCT/DST is of the form (3). More details are in [5, 8].

We note that, once the signal model is explicitly known, many known, and also new, algorithms can be derived easily by manipulating the polynomial algebra rather than the transform matrix [8].

6. 2-D TIME AND SPACE MODELS

The introduced algebraic approach to 1-D SP is readily extended to higher-dimensional SP, and can be used to derive new SP schemes for nonseparable SP. We briefly discuss the 2-D case.

Separable 2-D signal models. Usually, 2-D SP is done in a separable way, which assumes the signal resides on a rectangular lattice with two shift operators operating orthogonally. In the time case (directed, Fig. 6(a)), the signal model consists of Laurent series in two variables (infinite case) or of the polynomial algebra $\mathbb{C}[x, y]/\langle x^n - 1, y^n - 1 \rangle$ (finite case) with basis $(x^k y^\ell)_{k,\ell}$. In the space case (undirected, Fig. 6(b)), the signal model consists analogously of series (infinite case) or polynomials (finite case) in $C_k(x)C_\ell(y)$. In the finite case, the polynomials are elements of polynomial algebras in two variables. For example, applying a 2-D DCT-2 to an $n \times n$ image (as done in JPEG image compression), assumes the space model $\mathcal{A} = \mathcal{M} = \mathbb{C}[x, y]/\langle (x - 1)U_{n-1}(x), (y - 1)U_{n-1}(y) \rangle, \Phi : \mathbf{s} \mapsto \sum_{0 < k, \ell < n} s_{k, \ell} V_k V_{\ell}$.

Nonseparable 2-D signal models. In 2-D, there are lattices that do not have 1-D counterparts. Examples include the hexagonal lattice or quincunx lattice. Time (directed) SP in these cases was developed by Mersereau et al. [3]. The algebraic approach enables the derivation of infinite and finite space signal models for these lattices. This includes the proper choice of boundary conditions, the exact form of the associated Fourier transform, and the derivation of its fast algorithm. See [9, 10, 11] for more details.

7. CONCLUSION

The paper presents an algebraic theory of signal processing in which the basic building block is the signal model: a triple of a space of filters (an algebra), a space of signals (a module), and a



Fig. 6. Shifts for the separable 2-D time (directed) and 2-D space (undirected) models. Both yield rectangular lattices.



Fig. 7. Shifts for non-separable 2-D space (undirected) models. They yield a hexagonal and a quincunx lattice, respectively.

bijective mapping (a generalization of the *z*-transform). Instantiations of the signal model give rise to infinite and finite discrete time and discrete space linear signal processing models with the corresponding linear transforms and other basic concepts.

8. REFERENCES

- [1] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press, 1990.*
- [2] J. M. F. Moura and M. G. S. Bruno, "DCT/DST and Gauss-Markov fields: Conditions for equivalence," *IEEE Trans. on Signal Processing*, vol. 46, no. 9, pp. 2571–2574, 1998.
- [3] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall, 1984.
- [4] M. Püschel and J. M. F. Moura, "Algebraic theory of signal processing: Foundation and 1-D time," submitted for publication.
- [5] M. Püschel and J. M. F. Moura, "Algebraic theory of signal processing: 1-D space," submitted for publication.
- [6] H. J. Nussbaumer, *Fast Fourier Transformation and Convolution Algorithms*, Springer, 2nd edition, 1982.
- [7] T. J. Rivlin, *The Chebyshev Polynomials*, Wiley Interscience, 1974.
- [8] M. Püschel and J. M. F. Moura, "The algebraic approach to the discrete cosine and sine transforms and their fast algorithms," *SIAM Journal of Computing*, vol. 32, no. 5, pp. 1280–1316, 2003.
- [9] M. Püschel and M. Rötteler, "The Discrete Triangle Transform," in *Proc. ICASSP*, 2004.
- [10] M. Püschel and M. Rötteler, "Cooley-Tukey FFT like fast algorithms for the discrete triangle transform," in *Proc. 11th IEEE DSP Workshop*, 2004.
- [11] M. Püschel and M. Rötteler, "Fourier transform for the spatial quincunx lattice," in *Proc. ICIP*, 2005.