Multivariate Chebyshev polynomials and FFT-like algorithms

Dissertation

zur Erlangung des naturwissenschaftlichen Doktorgrades

am Institut für Mathematik der

Julius-Maximilians-Universität Würzburg

vorgelegt von

Bastian Seifert

aus

Erlangen, Deutschland

Würzburg, 2020

Eingereicht: Juni 2019
Betreuer: Prof. Dr. Knut Hüper, Julius-Maximilians-Universität Würzburg

Gutachter:

Prof. Dr. Knut Hüper, Julius-Maximilians-Universität Würzburg, Deutschland

Prof. Dr. Hans Z. Munthe-Kaas, Universitetet i Bergen, Norwegen

Prof. Dr. Markus Püschel, Eidgenössische Technische Hochschule Zürich, Schweiz

Tag der mündlichen Prüfung: 18. Juni 2020

Acknowledgments

I would like to thank to all those who have enabled the completion of this thesis. First of all I am grateful to my advisor, Knut Hüper, for all the nice discussions, support in grant applications, improvement of my writing skills, and especially for teaching me how to write math papers, which engineers are able to understand.

Furthermore I would like to thank Christian Uhl for the opportunity to work with him at Ansbach, the financial support, especially the many conference visits, and academic freedom obtained during my time at the Ansbach University of Applied Sciences. This time led to many interesting ideas, of which not all made it into this thesis, and improved my skills in communicating with engineers. During the creation of this thesis I was supported by the European Regional Development Fund (ERDF) within the project "Technologie-Transferzentrum für den Mittelstand (TZM)".

I am grateful to Hans Munthe-Kaas for the nice hospitality and support as well as various fruitful discussions during a visit to Bergen. I like to thank the Deutscher Akademischer Austauschdienst (DAAD) for supporting me during the visit to Bergen under the Vortragsreisenprogramm.

I am grateful to Daan Huybrechs for fruitful discussions and comments on matrix-valued Chebyshev polynomials.

Last but not least I'd like to thank my family and friends for all their support in my life and during the writing of this thesis.

Ansbach, June 2019

Bastian Seifert

Contents

List of Illustrations

Figures

Tables

List of Symbols

$\mathscr{A}, \mathscr{B}$	algebras
M, N	modules
$\cdot \triangleright \cdot$	algebra action on a module
$\hookrightarrow$	inclusion
$\mathrm{diag}(a_i)$	diagonal matrix with entries a_i
$\mathcal{F}$	Fourier transform matrix
Π^d	polynomials in d variables
$\mathsf{V}(I)$	variety of the ideal I
$\mathsf{I}(V)$	ideal of the variety V
Q	root lattice
α	a simple root
$Q^\vee$	coroot lattice
$\alpha^\vee$	a simple coroot
P	weight lattice
λ	a weight
ω	a fundamental weight
$P^\vee$	coweight lattice
$\lambda^\vee$	a coweight
$\omega^\vee$	a fundamental coweight
W	Weyl group
σ_α	reflection through the hyperplane perpendicular to the root α
m_i	mark of a root system
$m_i^\vee$	comark of a root system
T_λ	multivariate Chebyshev polynomial of the first kind of weight λ
U_λ	multivariate Chebyshev polynomial of the second kind of weight λ
$\deg_{\mathrm{m}}$	m-degree

Chapter 1

Introduction

Abstract treatment of concepts in application areas has lead to various advantages in the application areas as well as in pure mathematics. In this dissertation we use an abstract treatment of concepts in signal processing using the language of algebra leading to algebraic signal processing theory. Indeed all the basic concepts of linear discrete-time signal processing, like signals, filters or the Fourier transform, have counterparts in algebra especially in the representation theory of algebras, i.e. modules, algebras and decomposition into irreducible submodules, respectively. Of course the connection between signal processing and algebra is now classical as it was already used for the derivation of fast Fourier transform algorithms by Nicholson [83] and Winograd [126] in the 1970s. But a conceptual algebraic structure capturing all the relevant aspects of linear signal processing, the algebraic signal model, has only been identified recently by Püschel and Moura [90].

This concept allows for the derivation of a large class of signal models and discrete signal transforms using polynomial algebras. In [109] we used the algebra-geometry dictionary, using Hilbert's Nullstellensatz [35] and the Serre-Swan theorem [112,115], to motivate a geometric interpretation of algebraic signal models. The geometric interpretation in turn implies that the underlying geometric objects, an algebraic variety together with a vector bundle, are intrinsically complicated objects. Since we are interested mainly in discrete signals we are interested in discrete varieties, as well. Another branch of mathematics interested in objects with connections to discrete varieties is orthogonal polynomials in several variables. Indeed we show that there is a connection between Gaußian cubature formulae based on orthogonal polynomials and the existence of a unitary Fourier transform for a corresponding signal model. This yields a multivariate generalization of the Gauß-Jacobi procedure [131] for the unitarization of signal transforms.

The content of Chapter 2 is as follows. First we recall the motivation and definition of algebraic signal models from [90] and state their geometric interpretation. Then Fourier transforms for such signal models, based on the Chinese remainder theorem, are recalled. The second part of that chapter contains the first new result - Theorem 2.14 - the multivariate Gauß-Jacobi-procedure based on a multivariate Christoffel-Darboux formula derived by Xu [129]. For this we first recall some facts about multivariate orthogonal polynomials. Then we prove the multivariate Gauß-Jacobi-procedure and show that the sufficient condition for its applicability is the same condition as for the existence of a Gaußian cubature formula.

If one has an algebraic model of discrete signals one likes to have a fast algorithm for the calculation of its Fourier transform. Even though the most popular fast algorithm for the calculation of the standard discrete Fourier transform, the fast Fourier transform (FFT), was basically already known by Gauß [24], it was only popularized after its independent rediscovery by Cooley and Tukey [13] one hundred years later, see [33] for a historic overview. Due to its numerous applications the FFT has been termed to be one of the most important algorithms of the twentieth century [101]. The usage of algebra in its derivation dates back at least until the works of Nicholson [83]. Algebraic approaches to FFT-like algorithms split into two main directions: group algebra and polynomial algebra approaches. The interpretation of the fast Fourier transform in terms of the cyclic group Z_n was introduced in [83]. This group-based approach allows for a generalization of FFT algorithms to non-abelian groups as derived by Diaconis and Rockmore [20] and has inspired the approach to FFTs on compact groups by Maslen [72]. The polynomial algebra approach relies on the insight that there exists an isomorphism of algebras $\mathbb{C}[Z_n] \cong \mathbb{C}[x]/\langle x^n - 1 \rangle$. This approach allows to study another large class of FFT algorithms [4, 34, 46] relying on ideas of Nussbaumer [84] and Winograd [127]. The full polynomial algebra approach was worked out in [89–93] leading to the advent of algebraic signal processing theory. See [90, Sect. I-B] for more historical remarks and references.

One main difference in algebraic signal processing when compared to other recent approaches, like the decomposition of semi-simple algebras using Bratelli-diagrams investigated by Maslen, Rockmore, and Wolff [71], is that in algebraic signal processing one decomposes *modules*. This is motivated by the fact that in algebraic signal processing theory the signals are modeled as a module over the algebra of filters. This approach then leads to explicit matrix factorizations.

In the algebraic signal processing theory one can identify up to now three approaches for the derivation of fast algorithms for Fourier transforms of algebraic signal models based on polynomials. Even though all three approaches are essentially based on the Chinese remainder theorem and a stepwise partial decomposition, the different details lead to algorithms of different structure and different computational cost [89, 92, 121].

The first one is based on a factorization of polynomials $f(x) = g(x) \cdot h(x)$. This approach requires no special conditions on the polynomial f but might lead to suboptimal $O(n \log^2 n)$ algorithms.

The second approach is based on the decomposition property $f(x) = p(q(x))$ of certain polynomials. This approach gives optimal $O(n \log n)$ algorithms. Unfortunately in one variable the only families of polynomials posessing this property are, up to affine-linear coordinate changes, the monomials x^n and the Chebyshev polynomials $T_n(x)$ [98, Ch. 4]. Hence the only $O(n \log n)$ algorithms for signal models based on polynomials in one variable derivable by this method are the Cooley-Tukey-type algorithms for the trigonometric, i.e. sine and cosine transforms associated to Chebyshev polynomials, and the discrete Fourier transform, associated to the monomials. In several variables it is not known [119] if there are, up to affine-linear coordinate changes, any examples of polynomials with this property except the monomials and multivariate Chebyshev polynomials. The second approach in combination with multivariate Chebyshev polynomials was used to derive fast algorithms for undirected hexagonal [96] and BCC lattices [111].

As there are other algorithms for the discrete Fourier transform, like the Britanak-Rao-FFT [9], one might wonder if these algorithms can be derived using algebraic signal processing theory. This question was solved using the third approach, which generalises the second approach [104]. Here one relies on induced modules. This approach raises the level of abstraction by not relying on properties of the polynomials but on properties of the signal modules. Module induction is based on an algebra $\mathscr{A}$ with subalgebra $\mathscr{B}$ and a finite set $T \subset \mathscr{A}$, the transversal, such that $\mathscr{A}$ is the direct sum of copies of $\mathscr{B}$ shifted, i.e. multiplied, by the elements of the transversal. The induced $\mathscr{A}$-module N of a $\mathscr{B}$-module M is the direct sum of shifted, i.e., acted on, copies of M by the elements of the transversal. In [104] this approach was worked out for polynomial algebras in one variable with regular modules, i.e. the algebra considered as a module over itself. As applications general-radix algorithms for the Britanak-Rao [9] and the Wang-FFT [122] were deduced.

In Chapter 3 we first show how to generalize the induction-based FFT theorem from one variable to several variable polynomial algebras, see Theorem 3.1. Then we investigate the decomposition property used in [96] in more detail and investigate if, like in the univariate case, the decomposition property always yields the existence of an induction. This is not the case as easy geometric considerations, i.e., counting the points of the underlying discrete varieties, show, cf. Proposition 3.12. These considerations lead to a more general FFT theorem, our Theorem 3.13, based on the decomposition property, than the one used in [96].

Now the question arises if the FFT theorem based on decomposition is significant to obtain new, fast algorithms. As aforementioned in the univariate case the only, up to affine-linear change of variables, families of polynomials which obey the decomposition property are the monomials x^n and the Chebyshev polynomials of the first kind $T_n(x)$ [98, Ch. 4]. Since the monomials are associated to standard discrete Fourier transform and the Chebyshev polynomials are associated to the discrete cosine and sine transforms this limits the applicability of the decomposition FFT theorem. This leads to the question whether in several variables the situation is different. The first task in this direction is to check whether there are other generalizations of the Chebyshev polynomials to several variables than the trivial product of two univariate Chebyshev polynomials in different variables and the one considered in [95, 96]. Luckily there is a rather mature theory of multivariate Chebyshev polynomials associated to root systems available [36]. This theory gives a nice geometric description of the decomposition property in terms of foldable figures and a stretching-and-folding mechanism, as well.

Chapter 4 starts by recalling the definition and some properties of the multivariate Chebyshev polynomials. The shifts of corresponding signal models lead to the recognition that these signal models are associated to the weight lattices of semisimple Lie groups. This is interesting because some of these weight lattices are optimal regular sampling lattices, i.e., lattices where one can get maximal information with a minimal number of sampling points, cf. [12, 85]. The chapter proceeds by deriving analogues of the fast cosine transform for some special cases: the A_2 transform, connected to the Lie group $\mathsf{SU}(3)$, yields the cosine transform on a directed hexagonal lattice, the C_2 transform, associated to $\mathsf{SO}(5)$, leads to a cosine transform on a directed lattice of triangles, and the A_3 transform, connected to $\mathsf{SU}(4)$, gives a cosine transform on the body-centered cubic lattice. The A_2 transform was introduced by Püschel and Rötteler [95, 96], while the C_2 and A_3 transform were introduced in [110, 111].

In the last section of Chapter 4 a method for the calculation of generating functions for the matrix-valued Chebyshev polynomials is deduced. This part is not thoroughly connected to the rest of the thesis. While studying multivariate Chebyshev polynomials we had some inspiring conversations with Hans Munthe-Kaas and Daan Huybrechs about these polynomials and they showed me their preprint [44] in which matrix-valued and multivariate Chebyshev polynomials associated to representations of Weyl groups were introduced. When we investigated these polynomials we stumbled upon a way to generalize the method for the calculation of generating functions of multivariate Chebyshev polynomials presented in Damskinsky, Kulish and Sokolov [18, 19, 114] to the matrix-valued setting. Even though we are still lacking an existence proof of the matrix-valued polynomials the examples provided through the generating functions show that they will be interesting subjects to study and apply in the future.

We also want to mention some historical facets of multivariate Chebyshev polynomials and their applications. To our knowledge the first person to study multivariate Chebyshev polynomials was Lidl [21, 64–66] motivated by questions about permutation polynomials. Koornwinder [52, 54] studied general orthogonal polynomials in two variables containing the bivariate Chebyshev polynomials as special cases. The decomposition property was first observed by Ricci [97]. Hoffman and Withers [36] finally gave the geometric approach using the stretching and folding property and the connection to root systems. Beerends [2] studied the connection to the Laplace-Beltrami operator of certain symmetric spaces. Lyakhovsky investigated the connection to singular elements of Lie groups [68, 69] much more recently. A very general approach to orthogonal polynomials in several variables is due to Heckman and Opdam [31, 32].

Applications of the multivariate Chebyshev polynomials are somewhat rare but exist. Borzov and Damaskinsky [5, 6] studied the quantum harmonic oscillator obtained from the bivariate Chebyshev polynomials of type A_2. These considerations led to the derivation and calculation of the generating functions of the multivariate Chebyshev polynomials by Damaskinsky, Kulish, and Sokolov [18, 19, 114]. Some generating functions were calculated by Czyzycki, Hrivnák and Patera [17], as well.

In the context of algebraic geometry, multivariate Chebyshev polynomials were used to derive surfaces with many singularities by Breske, Labs, and van Straten [8].

In dynamical systems theory multivariate Chebyshev polynomials were used to define chaotic mappings by Withers [128] and Uchimura [117] while Nakane [81] studied the external rays of these mappings.

Klimyk and Patera [48] considered Weyl group orbit functions, which are in fact just multivariate Chebyshev polynomials with a different normalization. This article lead to a further studies by the school of eastern european mathematicians around Hrivnák and Patera [29, 82]. As applications, cubature rules [28, 41, 76] and discretizations of tori of compact simple Lie groups [38, 40, 42] were studied by this school. Furthermore they studied discrete orbit transforms [10, 16, 39] but did not make the connection to the algebraic signal processing theory and hence could not derive fast algorithms for their calculation.

Another group studying discrete Fourier transforms connected to multivariate Chebyshev polynomials and cubature formulas is Li, Sun, and Xu [60–63].

In the context of spectral approximations on triangles and spectral methods applications were investigated by Munthe-Kaas in collaboration with Nome, Ryland, and Sorevik [77–79, 103].

The ideas towards the definition of matrix-valued multivariate Chebyshev polynomials by Huybrechs and Munthe-Kaas [44] were inspired by a preprint of Hoffman and Withers [37]. In general, matrix-valued and multivariate orthogonal polynomials have found attention only recently due to the work of Grünbaum, Pacharoni, and Tirao [27]. A connection between matrix-valued and multivariate orthogonal polynomials to Gelfand pairs was studied by Koelink, van Pruijssen and Román [49–51] inspired by ideas of Koornwinder on vector-valued polynomials [53].

We summarize the main contributions of this thesis. The first main result is the multivariate Gauß-Jacobi procedure in Theorem 2.14 and its connection to Gaußian cubature formulae. The second main contribution is the multivariate FFT theorem for signal modules based on multivariate polynomials, Theorem 3.1 and Theorem 3.13. The main contribution of Chapter 4 is that we show that the multivariate Chebyshev polynomials give rise to well-behaved examples of the general theory. Furthermore the geometric mechanism underlying these algorithms are identified. Finally Theorem 4.15 contains a procedure to derive generating functions of matrix-valued and multivariate Chebyshev polynomials.

Some of the results in this thesis have been published in [109–111]. The main results of this thesis are contained in the single author publication [109].

If we omit entries in a matrix the reader should read a zero instead. The last numbers of the entries in the bibliography refer to the pages where the reference was used.

Chapter 2

Algebraic signal processing

This chapter serves two purposes. First in Sect. 2.1 we discuss theoretical signal processing from an engineer's perspective to motivate the notions and definitions of algebraic signal processing introduced in [90]. Afterwards the basic definitions of algebraic signal processing are given together with some examples. This includes the notion of an algebraic signal model and corresponding generalized Fourier transforms. This part contains no new results, except the observation that there is a geometric counterpart to the signal models we consider.

The second part in Sect. 2.2 recalls some facts about multivariate orthogonal polynomials. Then we generalize the Gauß-Jacobi procedure for the derivation of orthogonal Fourier transforms for signal models relying on univariate polynomials to signal models relying on multivariate polynomials. Furthermore we show that the sufficient condition for the existence of an orthogonal Fourier transform is the same condition as for the existence of a Gaußian cubature formula. Hence we have a motivation to use a class of polynomials, which were used to derive multivariate Gaußian cubature formulas, as building blocks for examples of signal models with orthogonal Fourier transform.

We assume that the reader is familiar with basic representation theory of algebras and commutative algebra. References on these subjects are [15, 59, 102].

2.1 Algebraic signal processing theory

We motivate the definition of an algebraic signal model as explained in [90]. The main objects in linear signal processing are signals and filters. In the processing of time-dependent signals one possibility to filter signals is using special electrical circuits, termed filters. Now one can put these circuits in series, in parallel, and one can amplify them. These operations can be interpreted as a multiplication, an addition, and a scalar multiplication, respectively. Since the multiplication, i.e. putting filters in series, is bilinear with respect to the addition, i.e. putting in parallel, and the scalar multiplication, i.e. amplification, one obtains the structure of an algebra. On the other hand one can add two signals and one can apply a filter to a signal and obtains a new signal. Since one has the interpretation of the filters as an algebra this yields that the signals form a module over the algebra of filters. In this way, filters are linear mappings on signals.

The next thing to recall is that the electrical circuits realizing the filters can mathematically be described using ordinary differential equations. The standard method in engineering sciences for transforming ordinary differential equations to algebraic equations, from which one can obtain sometimes more information, is the Laplace transform

$$f \mapsto \mathcal{L}(f)(s) = \int_0^\infty f(t)\mathrm{e}^{-st}\,\mathrm{d}t, \tag{2.1}$$

for $s \in \mathbb{C}$. This indicates that the filters can be interpreted as elements of polynomial algebras.

Now we turn to the discrete world. One method to go from analog, i.e. continuous, signals to discrete signals is by sampling them, i.e. getting the values at fixed time-step values. That is one associates to a continuous signal $q(t)$ a sequence of numbers by the rule

$$q_k = q(k \cdot T_A), \tag{2.2}$$

where T_A is the sampling period. Now the Laplace transform of this discrete signal is

$$\mathcal{L}(q)(s) = \sum_{k=0}^{\infty} q(k \cdot T_A)\mathrm{e}^{-kT_A s}. \tag{2.3}$$

Substituting $z = \mathrm{e}^{-kT_A s}$ one obtains the Laurent series

$$\sum_{k=0}^{\infty} q(k \cdot T_A) z^{-k} = \sum_{k=0}^{\infty} q_k z^{-k}. \tag{2.4}$$

This transformation is called the z-transform and used in the analysis of discrete systems. Now in real-world applications one often is concerned with only finite signals, so one is also interested in a finite z-transform of a finite discrete signal q_k which is of the form

$$\sum_{k=0}^{n-1} q_k z^{-k}, \tag{2.5}$$

which results in a polynomial in z^{-k}. Hence the finite z-transform is a bijection between a set of numbers, the signal samples, and polynomials in z^{-1}, which will turn out to be elements of a signal module, when considered with the filter operations.

We have motivated the following definition of an algebraic signal model, which is the foundation of algebraic signal processing [90].

Definition 2.1 (Algebraic signal model) *An algebraic signal model is a triple $(\mathscr{A}, M, \Phi)$ consisting of an algebra $\mathscr{A}$, the algebra of filters, an $\mathscr{A}$-module M, the module of signals, and a bijective map $\Phi\colon \mathbb{K}^n \longrightarrow M$ for $n \in \mathbb{N} \cup \{\infty\}$ and some field $\mathbb{K}$.*

Remark 2.2 In this thesis only the fields $\mathbb{K} = \mathbb{R}, \mathbb{C}$ will be used. In principle one does not need to restrict to these ground fields. Indeed signal processing and Fourier transforms using finite fields might be of interest in some applications as these can be used for infinite precision calculations, see e.g. [67]. Nonetheless we will only consider algebras over $\mathbb{C}$ in this thesis since this simplifies some arguments and definitions.

For explicit computations we additionally require that we can choose a basis of the module as a $\mathbb{C}$-vector space. The elements of the basis play a role in interpreting concepts from signal processing, as well. Indeed the basis elements b_i are the impulses and the impulse response of a filter $h \in \mathscr{A}$ is $h \triangleright b_i$ for each basis elements.

The first example is the classical finite discrete time signal model [93].

Example 2.3 Consider the algebraic signal model with algebra $\mathscr{A} = \mathbb{C}[x]/\langle\, x^n - 1\,\rangle$, the regular module $M = \mathscr{A}$, and the z-transform $\Phi \colon \mathbb{C}^n \longrightarrow M$ given by

$$(s_0, \dots, s_{n-1}) \mapsto \sum_{k=0}^{n-1} s_k x^k. \tag{2.6}$$

If one replaces x by z^{-1} one obtains the finite discrete time signal model from theoretical electrical engineering, but from an algebra point of view it is more common to have polynomials in x than in z^{-1}.

The next example shows that one has to be more careful if one considers not necessarily finite signals [90].

Example 2.4 We consider infinite discrete signals. As algebra one chooses $\mathscr{A} = \ell^1(\mathbb{Z})$, which from an engineering perspective corresponds to filters being BIBO-stable, i.e. with bounded input and output. As signal module one chooses $M = \ell^2(\mathbb{Z})$, which can identified with Laurent series $M = \{s = \sum_{n\in\mathbb{Z}} s_n z^{-n} \mid s = (\dots, s_{-1}, s_0, s_1, \dots) \in \ell^2(\mathbb{Z})\}$. From a signal processing perspective these signals are the signals with finite energy. The z-transform corresponds to the choice of basis of M as $(\dots, z^1, z^0, z^{-1}, \dots)$ and is thus

$$\begin{aligned} \Phi \colon \ell^2(\mathbb{Z}) &\longrightarrow M, \\ (\dots, s_{-1}, s_0, s_1, \dots) &\mapsto \sum_{n\in\mathbb{Z}} s_n z^{-n}. \end{aligned} \tag{2.7}$$

That M is indeed an $\ell^1(\mathbb{Z})$-module follows from the general fact, that $\ell^p(\mathbb{Z})$ for any $p < \infty$ is an $\ell^1(\mathbb{Z})$-module with module action given by convolution. This can be proven as follows. Let $h \in \ell^1(\mathbb{Z})$, $s \in \ell^p(\mathbb{Z})$ and let $t = h * s$. One needs to show that $t \in \ell^p(\mathbb{Z})$. The nth part of t can be estimated using triangle and Hölder inequalities as

$$\begin{aligned} |t_n| &= |\sum_{k\in\mathbb{Z}} h_k s_{n-k}| \\ &\le \sum_{k\in\mathbb{Z}} |s_{n-k}||h_k|^{1/p}|h_k|^{1-1/p} \\ &\le \left(\sum_{k\in\mathbb{Z}} |s_{n-k}|^p |h_k|\right)^{1/p} \left(\sum_{k\in\mathbb{Z}} |h_k|\right)^{1-1/p}. \end{aligned} \tag{2.8}$$

The second factor is independent of n and h is absolute convergent hence

$$\begin{aligned}\sum_{n\in\mathbb{Z}} |t_n|^p &\leq \sum_{n\in\mathbb{Z}} \left(\left(\sum_{k\in\mathbb{Z}} |s_{n-k}|^p |h_k| \right) \left(\sum_{k\in\mathbb{Z}} |h_k| \right)^{p-1} \right) \\ &= \left(\sum_{k\in\mathbb{Z}} |h_k| \right)^{p-1} \left(\sum_{n\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} |s_{n-k}|^p |h_k| \right) \\ &= \left(\sum_{k\in\mathbb{Z}} |h_k| \right)^{p-1} \sum_{k\in\mathbb{Z}} |h_k| \sum_{n\in\mathbb{Z}} |s_{n-k}|^p \\ &= \left(\sum_{k\in\mathbb{Z}} |h_k| \right)^{p} \sum_{n\in\mathbb{Z}} |s_n|^p < \infty.\end{aligned} \tag{2.9}$$

So $t \in \ell^p(\mathbb{Z})$.

The next crucial notion to introduce is that of a shift and shift-invariance [90].

Definition 2.5 (Shift) *For an algebraic signal model $(\mathscr{A}, M, \Phi)$ a chosen set of generators of the algebra $x_1, \ldots, x_n$ is called the shifts of the signal model. A filter $h \in \mathscr{A}$ is called shift-invariant if $h \cdot x_i = x_i \cdot h$ for all x_i. A signal model is called shift-invariant if all the filters are shift-invariant.*

Since the shifts generate the algebra, a signal model is shift invariant if and only if the algebra is commutative. Since we want to investigate the algebraic essence of signal processing, we only consider algebras generated by the shifts algebraically, i.e. we will not consider algebras where one has to complete using some norm to generate the algebra.

Shift-invariant signal models are precisely those with a polynomial algebra as filter algebra [90]. This can be seen by observing that the shifts in a shift-invariant signal model commute and recalling that they are the generators of the algebra. This does not spoil the usage of group algebras of commutative groups in signal processing, since for any commutative group G one has $G \cong C_{n_1} \times \cdots \times C_{n_d}$, where C_{n_i} are cyclic groups of order n_i, and hence

$$\mathbb{C}[G] \cong \mathbb{C}[C_{n_1} \times \cdots \times C_{n_d}] \cong \mathbb{C}[x_1]/\langle\, x_1^{n_1} - 1 \,\rangle \otimes \cdots \otimes \mathbb{C}[x_d]/\langle\, x_d^{n_t} - 1 \,\rangle. \tag{2.10}$$

From a signal processing perspective this might explain why non-commutative groups have not found that many applications in signal processing, since shift-invariance is a desirable property. This is the case since in standard applications of time signal processing one desires that the filters behave the same no matter at which time one applies them.

The shifts can be used to visualize the signal model. The visualization is motivated by the following considerations. Assume there exists a basis $b_t \in M$ of M as a $\mathbb{C}$-vector space, the impulses of the signal model. Note that these basis elements are determined by the z-transform Φ, which maps the canonical basis of $\mathbb{C}^n$ to these basis elements. Now the shifts, as generators of the filter algebra, act on the basis elements resulting in a sum of basis elements. The non-zero coefficients tell which basis elements can be connected by the shifts.

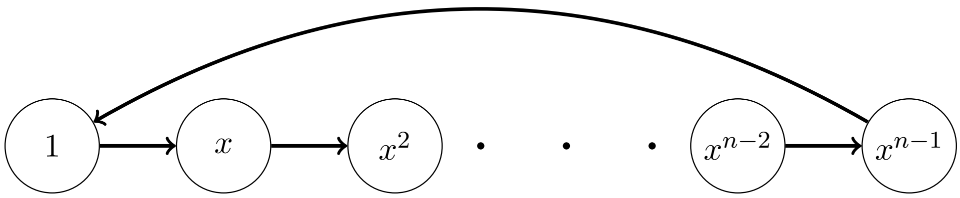

Figure 2.1: The visualization graph of the finite discrete time signal model.

Definition 2.6 (Visualization graph) *The visualization graph of an algebraic signal model is the graph obtained by adding a node for each basis element of the signal module and adding an edge from b_i to b_j if the coefficient of b_j in the sum $x_i \triangleright b_i = \sum s_k b_k$ is non-zero for at least one shift x_i of the filter algebra, weighted with the sum of these coefficients.*

Note that we will omit the weights in general if we draw the visualization graphs. The weights are important if one is interested in going from algebraic signal processing to graph signal processing, cf. [106, 107]. Furthermore we will omit the boundary connections when this would lead to visual confusion and not add much additional information. Especially in two and three dimensional space signal models we tend to omit the boundary conditions.

For example the visualization graph of the signal model from Example 2.3 is shown in Fig. 2.1. It clearly shows the time character of the signal model as it is a directed graph. Furthermore the periodic extension implicitly assumed by taking only finite signals appears in the visualization graph. Now there are signals which are not time-dependent but space-dependent, like images. The next example shows how one builds a space signal model [91]. The space-dependence will appear in the visualization graph, which will be an undirected one unlike the directed graph for time-dependent signals.

Example 2.7 Consider the Chebyshev polynomials of the first kind given by

$$T_n(\cos\theta) = \cos(n\theta), \tag{2.11}$$

which are polynomials in $x = \cos(\theta)$ for $\theta \in [0, \pi]$. They obey the recurrence relation

$$xT_n(x) = \tfrac{1}{2}T_{n-1}(x) + \tfrac{1}{2}T_{n+1}(x). \tag{2.12}$$

Consider the signal model with filter algebra $\mathscr{A} = \mathbb{C}[x]/\langle\, T_n(x)\,\rangle$, regular module $M = \mathscr{A}$, and z-transform $\Phi(s) = \sum_{i=0}^{n-1} s_i T_i(x)$. Then x is a generator of $\mathscr{A}$ and hence a shift. By (2.12) one has the visualization of the signal model shown in Fig. 2.2. Now in reality images are two-dimensional. So how does one obtain a signal model with a two-dimensional visualization graph? One simply takes the tensor product of the space signal model with itself, i.e. as algebra one takes

$$\mathscr{A}_2 = \mathbb{C}[x]/\langle\, T_n(x)\,\rangle \otimes \mathbb{C}[y]/\langle\, T_m(y)\,\rangle, \tag{2.13}$$

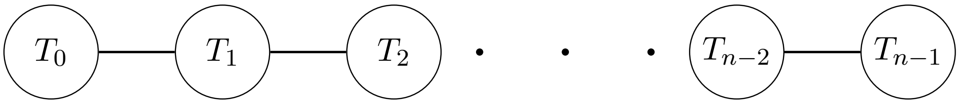

Figure 2.2: Visualization of the space signal model.

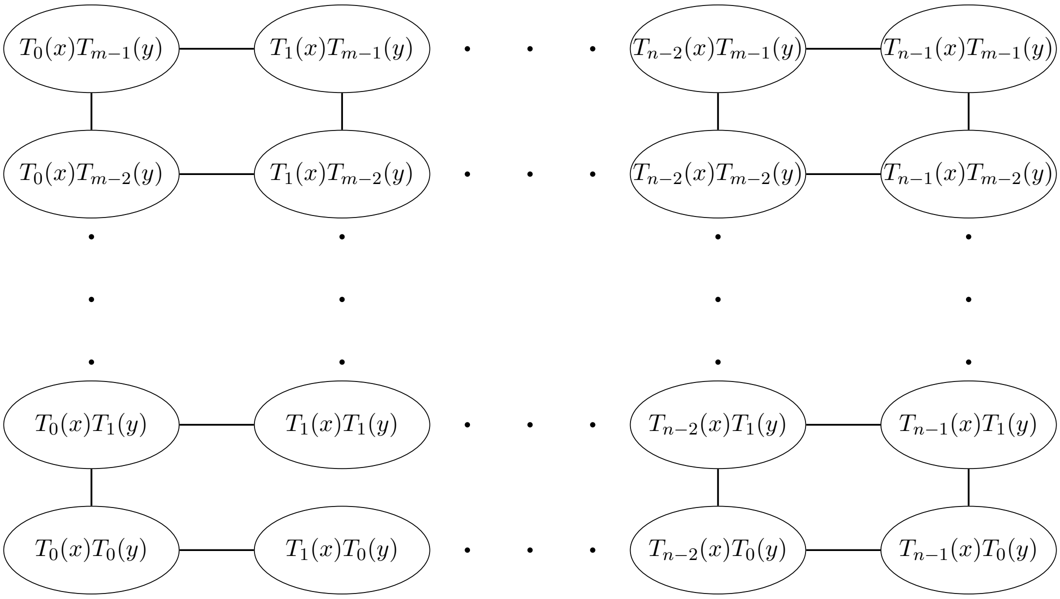

Figure 2.3: The visualization graph of the tensor product of two space signal models resembles a rectangular grid.

the module is again the regular one, and as z-transform one uses

$$\begin{aligned} \Phi_2 \colon \mathbb{C}^{n\times m} &\mapsto M \\ \Phi_2(s) &= \sum_{i=0}^{n}\sum_{j=0}^{m} s_{i,j}T_i(x)T_j(y). \end{aligned} \tag{2.14}$$

The corresponding visualization graph is shown in Fig. 2.3 and resembles the structure of an image sampled on a rectangular grid.

A crucial technique in signal processing is the decomposition of a signal into its frequency (or spectral) components using the Fourier transformation. In algebraic signal processing theory this can be captured using module theory. Since the signals are modeled using a module one likes to determine a module description of the spectral components. Simple modules, i.e., modules which have as only submodules the zero module and itself, are the notation which corresponds to the spectral components on the module level. A module is called semisimple if it can be decomposed into simple modules. This motivates the following definition [90].

Definition 2.8 (Fourier transform and spectrum) *Assume the signal module M of an algebraic signal model is semisimple, i.e. there exist simple $\mathscr{A}$-modules*

M_w such that $M \cong \bigoplus_w M_w$. Then any isomorphism

$$\mathcal{F}\colon M \longrightarrow \bigoplus_w M_w \tag{2.15}$$

is called a Fourier transform of the signal model. The frequency spectrum of a signal $s \in M$ is the image of the signal under the Fourier transform $\mathcal{F}(s) = (s_w)_{w \in W}$.

For real-world applications one often likes to realize the Fourier transform as a matrix. For this one has to choose bases in M, which was already done via the z-transform Φ in the signal model, and in each spectral component M_w.

For the decomposition into simple submodules we will rely on the Chinese remainder theorem. This theorem states that for a ring R with an ideal $I = I_1 \times \cdots \times I_n$ being a product of coprime ideals one has

$$R/I \cong R/I_1 \times \cdots \times R/I_n. \tag{2.16}$$

Likewise for an R-module M one has

$$M/MI \cong M/MI_1 \times \cdots \times M/MI_n. \tag{2.17}$$

The reader not familiar with the Chinese remainder theorem, especially in its general version for rings and modules or the version for polynomial algebras, is directed to Appendix B where it is treated in some detail.

We denote by $\Pi^n(x) = \mathbb{C}[x_1, \dots, x_n]$ the space of all polynomials in n indeterminates. Let the filter algebra be of the form $\mathscr{A} = \Pi^n(x)/I$ for some radical, zero-dimensional ideal I (the rational behind this will be explained after we investigated some examples). Denote by $\mathsf{V}(I) = \{\alpha \in \mathbb{C}^n \mid p(\alpha) = 0 \text{ for all } p \in I\}$. Then by the Chinese remainder theorem we have

$$\Pi^n(x)/I \cong \bigoplus_{\alpha \in \mathsf{V}(I)} \Pi^n/\langle\, x_1 - \alpha_1, \dots, x_n - \alpha_n \,\rangle, \tag{2.18}$$

as all the $\langle\, x_1 - \alpha_1, \dots, x_n - \alpha_n \,\rangle$ are coprime as maximal ideals. The corresponding Fourier transform for the signal model with regular module is realized by the map

$$p(x_1, \dots, x_n) \mapsto \Big[p(\alpha)\Big]_{\alpha \in \mathsf{V}(I)}. \tag{2.19}$$

If we choose a basis B in the module $M = \Pi^n/I$ and the basis $\{1\}$ consisting of one only in each $\Pi^n/\langle\, x_1 - \alpha_1, \dots, x_n - \alpha_n \,\rangle$, the Fourier transform can be realized as multiplication by the matrix

$$P_{B,M} = [b(\alpha)]_{b \in B, \alpha \in \mathsf{V}(I)}. \tag{2.20}$$

If other bases than $\{1\}$ are used in each $\Pi^n(x)/\langle\, x_1 - \alpha_1, \dots, x_n - \alpha_n \,\rangle$, e.g. $\{a_i\}$ for $a_i \in \mathbb{C}$, the matrix changes to

$$P_{B,M} = \operatorname{diag}\left(\frac{1}{a_i}\right)[b(\alpha)]_{b \in B, \alpha \in \mathsf{V}(I)}, \tag{2.21}$$

where $\operatorname{diag}\left(\frac{1}{a_i}\right)$ is the diagonal matrix containing the reciprocals of the basis elements. We now give some examples of Fourier transforms in the case of signal models relying on univariate polynomials from [90, 121].

Example 2.9 *i.)* For the finite discrete time signal from Example 2.3 one has that $x^n - 1 = \prod(x - \mathrm{e}^{2\pi\mathrm{i}k/n})$. Thus for the ideal one has $\langle\, x^n - 1 \,\rangle = \bigcap \langle\, x - \mathrm{e}^{2\pi\mathrm{i}k/n} \,\rangle$. In this case the Chinese remainder theorem yields

$$\mathbb{C}[x]/\langle\, x^n - 1 \,\rangle \cong \bigoplus \mathbb{C}[x]/\langle\, x - \mathrm{e}^{2\pi\mathrm{i}k/n} \,\rangle. \tag{2.22}$$

If we choose $\{1\}$ as basis in each spectral component one obtains, since as basis for the module we had choosen $\{1, x, \dots, x^{n-1}\}$ via the z-transform, the matrix

$$\mathsf{DFT}_n = [\mathrm{e}^{2\pi\mathrm{i}k\ell/n}]_{k,\ell}. \tag{2.23}$$

This is precisely the well-known discrete Fourier transform. The choice of basis in the spectral components as $\{1\}$ is somewhat arbitrary. Indeed we can choose any other basis $a_i \in \mathbb{C}$ in the spectral component and obtain a scaled discrete Fourier transform

$$\mathcal{F} = D \cdot [\mathrm{e}^{2\pi\mathrm{i}k\ell/n}]_{k,\ell}, \tag{2.24}$$

where $D = \operatorname{diag}\left(\frac{1}{a_i}\right)$ is the diagonal matrix containing the reciprocals of the basis elements.

ii.) If one replaces the field $\mathbb{C}$ in the finite discrete time signal model by the field $\mathbb{R}$ one considers only real filters and signals. The main difference appears in the spectrum and the Fourier transform. This is due to $\mathbb{R}$ not being a splitting field of $x^n - 1$. In fact over $\mathbb{R}$ the polynomial $x^n - 1$ has irreducible factors of degree 1, which are $x-1$ and $x+1$ if n is even, and of degree 2, which are of the form $x^2 - 2\cos(2k\pi/n) + 1$. These irreducible factors correspond to one and two dimensional spectral components. In the two-dimensional spectral components one has more freedom to choose a basis. Since the degree two irreducible factors arise from the combination of two complex conjugated factors in the complex case one can obtain a real version of the discrete Fourier transform by combining these. If one orders the complex zeros in such a way that the zeros 1 and -1 appear at positions 1 and $\frac{n}{2}$ (if n is even) and the conjugate zeros at the positions k and $n-k$ one can multiply the DFT_n with an x-shaped matrix X to obtain a Fourier transform of the real discrete time signal model. The matrix X has the form

$$X = \begin{bmatrix} * & 0 & \dots & \dots & 0 \\ 0 & * & & & * \\ \vdots & & \ddots & \iddots & \\ \vdots & & \iddots & \ddots & \\ 0 & * & & & * \end{bmatrix}, \tag{2.25}$$

where at positions $(1,1)$ and $(n/2, n/2)$ one has entries corresponding to the choice of basis in the one dimensional spectral components. The positions $(k,k), (k, n-k), (n-k, k)$, and $(n-k, n-k)$ form a 2×2 block which maps a pair of complex conjugated zeros to the two dimensional real component. Since these real components are two dimensional one has a larger freedom of choice. For example the choice of each block as

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{bmatrix}, \tag{2.26}$$

mapping $(a+b\mathrm{i}, a-b\mathrm{i})$ to (a, b) results in the real discrete Fourier transform [22]

$$\begin{aligned} \mathsf{RDFT}_n &= [r_{k,\ell}]_{0\le k,\ell<n} \\ &\text{with} \\ r_{k,\ell} &= \begin{cases} \cos\frac{2\pi k\ell}{n} & 0 \le k \le n/2 \\ -\sin\frac{2\pi k\ell}{n} & n/2 < k < n. \end{cases} \end{aligned} \tag{2.27}$$

The choice of each block as

$$\frac{1}{2}\begin{bmatrix} 1+\mathrm{i} & 1-\mathrm{i} \\ 1-\mathrm{i} & 1+\mathrm{i} \end{bmatrix}, \tag{2.28}$$

which maps $(a+b\mathrm{i}, a-b\mathrm{i})$ to $(a-b, a+b)$ leads to the discrete Hartley transform [7]

$$\mathsf{DHT}_n = \left[\cos\frac{2\pi k\ell}{n} + \sin\frac{2\pi k\ell}{n}\right]_{0\le k,\ell<n}. \tag{2.29}$$

iii.) The Fourier transform of the space model of Example 2.7 is connected to the discrete cosine transforms. First observe that the n zeros of nth Chebyshev polynomial are given by $\{\cos\frac{(k+1/2)\pi}{n} \mid 0 \le k < n\}$. By the definition (2.11) of the first kind Chebyshev polynomials and by choosing the basis $\{1\}$ in each one-dimensional component one obtains

$$\mathsf{DCT}_n = \left[\cos\frac{\ell(k+1/2)\pi}{n}\right]_{k,\ell}. \tag{2.30}$$

This is the discrete cosine transform of type 3. The other 15 types of discrete cosine and sine transforms can be obtained via similar considerations with Chebyshev polynomials of the second, third and fourth kind. See [89, 90] for details on this.

The dictionary between signal processing concepts and algebraic concepts developed in algebraic signal processing is summarized in Table 2.1.

Since in the following we are especially interested in signal models based on polynomial algebras in several variables, we have to investigate the caveats of going from one to several variables. It turns out that an algebro-geometric point of view provides some of the insights needed. For example, we have chosen a zero-dimensional, radical ideal for concrete examples of the Fourier transform (2.18) and the rational behind this is obvious from the algebraic geometry point of view. Furthermore this interpretation gives a geometric point of view on signal processing as well. We recall just as much of algebraic geometry as we need in the sequel. The reader interested in diving deeper into the connection between algebra and geometry is kindly directed to textbooks on algebraic geometry like [113].

Let $\Pi^d = \Pi^d(x) = \mathbb{C}[x_1, \dots, x_d]$ denote the algebra of polynomials in d variables with complex coefficients. In general, a polynomial algebra is of the form Π^d/I, where I is an ideal of the algebra Π^d. Such an ideal I is generated by a set of polynomials $I = \langle\, p_1, \dots, p_n \,\rangle$. The area of algebraic geometry relates such polynomials algebras to geometric structures, which is the algebraic varieties. With an ideal $I \subseteq \Pi^d$ one

signal processing concept	algebraic concept
filter	$h \in \mathscr{A}$ (algebra)
signal	$s = \sum s_t b_t \in M$ (module)
z-transform	$\Phi\colon \mathbb{C}^N \to M$ (bijection)
filtering	$h \triangleright s$ (algebra action on module)
impulse	$b_t \in M$ basis element
impulse response of $h \in \mathscr{A}$	$h \triangleright b_t \in M$
Fourier transform	$\mathcal{F}\colon M \longrightarrow \bigoplus_w M_w$
spectrum of a signal	$\mathcal{F}(s) = (s_w)_{w \in W}$
shift	generator of $\mathscr{A}$
shift-invariance	$\mathscr{A}$ commutative

Table 2.1: The dictionary between signal processing and algebraic concepts [90].

can associate an algebraic variety, a set of points in $\mathbb{A}^d = \mathbb{C}^d$ (interpreted as only the points with no vector space structure), via the rule

$$\mathsf{V}(I) = \{\alpha \in \mathbb{A}^d \mid p(\alpha) = 0 \text{ for all } p \in I\}. \tag{2.31}$$

For the converse one can associate an algebraic variety to an ideal via

$$\mathsf{I}(V) = \{p \in \Pi^d \mid p(\alpha) = 0 \text{ for all } \alpha \in V\}. \tag{2.32}$$

Now one would like to have the identity $\mathsf{I}(\mathsf{V}(I)) = I$, since the association of ideals to varieties and vice versa respects the morphisms between these objects as well, so one would obtain an equivalence of categories. Unfortunately this is not the case as the simple example $\mathsf{I}(\mathsf{V}(\langle\, x^2 \,\rangle)) = \langle\, x \,\rangle$ shows. In one variable the difference between $\langle\, x^2 \,\rangle$ and $\langle\, x \,\rangle$ is that the polynomial x is separable while x^2 is not. In several variables the notion of separable yields that of a radical ideal. Given an ideal I, its radical is given by

$$\sqrt{I} = \{p \in \Pi^d \mid p^n \in I \text{ for some } n \in \mathbb{N}\}. \tag{2.33}$$

Since we are interested in discrete transforms, we would like to have a discrete spectrum, as well. To have a discrete spectrum one has to have $I = \bigcap_{\alpha \in \mathsf{V}(I)} \langle\, x_1 - \alpha_1, \ldots, x_d - \alpha_d \,\rangle$, i.e., the variety is a finite set. Such an ideal is called zero-dimensional and the quotient space Π^d / I is then finite-dimensional.

For the geometric interpretation of the signal module we have to investigate the notion of sections of a vector bundle. Let V be an algebraic variety. The trivial bundle is the product $E = V \times \mathbb{C}^n$ with the projection $\rho(x, v) \mapsto x$ such that $\rho^{-1}(x)$ has a vector space structure. A vector bundle is now a total space E together with a projection $\pi\colon E \to V$ such that $E_x = \pi^{-1}(x) \cong \mathbb{C}^n$ has a vector space structure for each $x \in V$ and is locally trivial. Here local triviality means that there is an open cover of $V = \bigcup U_\alpha$ such that the restriction of π to each U_α is isomorphic to the

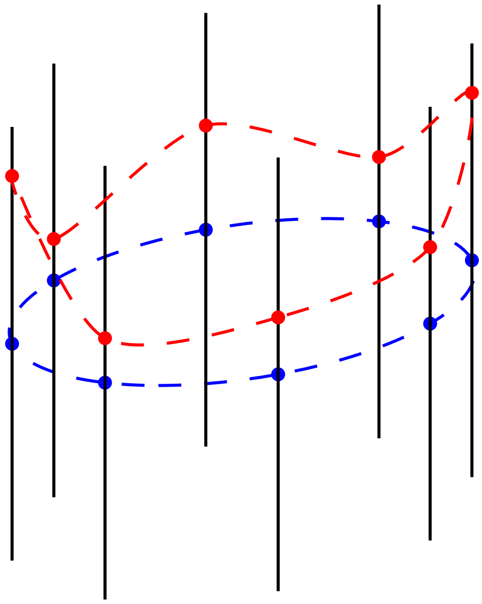

Figure 2.4: Sections of a vector bundle (red) over points on a circle (blue) form the signals of the finite, discrete time signal model.

trivial bundle, i.e., there exists a morphism ϕ_α such that the diagram

$$\begin{tikzcd} \pi^{-1}(U_\alpha) \arrow[rr, "\phi_\alpha"] \arrow[dr, "\pi"'] & & U_\alpha \times \mathbb{C}^n \arrow[dl, "\rho"] \\ & U_\alpha & \end{tikzcd} \tag{2.34}$$

commutes. A section of a vector bundle $\pi\colon E \to V$ is a morhpism $s\colon V \to E$ such that $\pi \circ s = \mathsf{id}$. The sections form a module over the coordinate algebra $\Pi^d/\mathsf{I}(V)$. In converse the famous Serre-Swan theorem [112, 115] states that to each module over the coordinate algebra one can associate a vector bundle. In this sense vector bundles and modules over the coordinate algebra are the same thing. Hence the geometric point of view on an algebraic signal model is that of an algebraic variety together with a vector bundle. As example the geometric picture corresponding to the finite discrete time signal model is shown in Fig. 2.4.

2.2 Orthogonal polynomials and orthogonal transforms

One property the monomials x^n and the Chebyshev polynomials T_n share is that both form families of univariate orthogonal polynomials. A family of polynomials $p_0, p_1, \ldots, p_n, \ldots$ is an orthogonal polynomial sequence if any two members are orthogonal with respect to some inner product and each p_n has degree n. The orthogonal polynomials are interesting for considerations in signal processing since there are exact n zeros of p_n, which are all simple. Hence the ideals they generate are radical. Furthermore one of their properties, the Christoffel-Darboux formula, can be used to derive unitary versions of the Fourier transforms of signal models defined by them [90, 105]. In the literature this is termed the Gauß-Jacobi procedure [131] and works for every family of orthogonal polynomials in one variable.

In several variables things get more complicated. The first thing one observes is that for $n > 0$ the set of polynomials $\mathcal{V}_n^d \subseteq \Pi_{\mathbb{R}}^d = \mathbb{R}[x_1, \dots, x_d]$ of total degree n which are orthogonal to all polynomials of lower degree form a vector space of dimension greater than one. Hence it is unlikely that there exists a distinguished basis of $\mathcal{V}_n^d$ which behaves better than the others. Indeed it turns out that it is advantageous to formulate the theory in terms of the spaces $\mathcal{V}_n^d$ and not for a special basis. The matrix-vector-notation needed for this formulation was introduced by Kowalski [56–58] and the formulation was worked out by Xu [129]. The resulting notion of multivariate orthogonal polynomials and the corresponding multivariate Christoffel-Darboux formula can be used to deduce a several variable analogon of the Gauß-Jacobi procedure. Unfortunately it turns out that its applicability is restricted to a small class of orthogonal polynomials, since the condition for its applicability is the same as for the existence of a Gaußian cubature formula, which rarely exist in higher dimensions. We start by recalling the basic definitions and properties of orthogonal polynomials in several variables.

Consider an inner product on $\Pi^d = \Pi_{\mathbb{R}}^d$ given by a non-negative weight function w, i.e.

$$\langle p, q \rangle = \int_{\mathbb{R}^d} p(x)q(x)w(x)\,\mathrm{d}x, \tag{2.35}$$

for $p, q \in \Pi^d$. Two polynomials p, q are called orthogonal if $\langle p, q \rangle = 0$. Let $\mathcal{V}_k^d \subseteq \Pi_k^d$ be the subset of the polynomials of total degree k that are orthogonal to all polynomials in Π_{k-1}^d together with zero. This yields the direct sum

$$\Pi^d = \bigoplus_{k=0}^{\infty} \mathcal{V}_k^d. \tag{2.36}$$

The dimension of $\mathcal{V}_k^d$ is $r_k = \binom{k+d-1}{k}$. To a basis $\{p_1^k, \dots, p_{r_k}^k\}$ of $\mathcal{V}_k^d$ associate the vector

$$\mathbb{P}_k(x) = [p_1^k(x), \dots, p_{r_k}^k(x)]^\top. \tag{2.37}$$

For any $x \in \mathbb{R}^d$ denote

$$x\mathbb{P}_k(x) = [xp_1^k(x), \dots, xp_{r_k}^k(x)]^\top. \tag{2.38}$$

The following theorem is a multivariate version of Favard's theorem [23] and shows that, like in the univariate case, orthogonality is equivalent to a three-term recurrence relation.

Theorem 2.10 *Let $\{\mathbb{P}_k\}_{k=0}^{\infty}$ be an arbitrary sequence in Π^d such that $\mathbb{P}_k$ spans Π_k^d for each k. Then one has equivalence of the following statements:*

i.) There exists a positive definite linear functional which makes $\{\mathbb{P}_k\}_{k=0}^{\infty}$ an orthogonal basis of Π^d.

ii.) For $n \geq 0$ and $1 \leq i \leq d$ there exist matrices $A_{k,i}, B_{k,i}$ and $C_{k,i}$ of size $r_k \times r_{k+1}, r_k \times r_k$, and $r_k \times r_{k-1}$, respectively, and $\operatorname{rank} A_{k,i} = r_{k+1}$ *and* $\operatorname{rank} C_{k,i} = r_{k-1}$ *such that one has the three-term recurrence relation*

$$x_i\mathbb{P}_k(x) = A_{k,i}\mathbb{P}_{k+1}(x) + B_{k,i}\mathbb{P}_k(x) + C_{k,i}\mathbb{P}_{k-1}(x). \tag{2.39}$$

PROOF: See [129, Thm. 2]. □

Note that the multivariate Favard theorem is not as strong as the univariate one. From the three-term recurrence relation one can deduce the multivariate Christoffel-Darboux formula.

Theorem 2.11 (Xu-Christoffel-Darboux formula) *Let* $\{P_k(x)\}_{k=0}^{\infty}$ *satisfy a three term recurrence relation of the form* (2.39) *with* $\mathbb{P}_{-1} = 0$ *and let there exist symmetric and invertible matrices* H_k *such that* $B_{k,i}H_k$ *is symmetric and*

$$A_{k,i}H_{k+1} = H_k C_{k+1,i}^{\top} \tag{2.40}$$

for any $1 \leq i \leq d$ *and* $k > 0$. *Then for any* n *and* $1 \leq i \leq d$ *one has*

$$\begin{aligned}&\sum_{k=0}^{n} \mathbb{P}_k^{\top}(x) H_k^{-1} \mathbb{P}_k(y) \\ &= \begin{cases} \frac{(A_{n,i}\mathbb{P}_{n+1}(x))^{\top} H_n^{-1} \mathbb{P}_n(y) - \mathbb{P}_n^{\top}(x) H_n^{-1} A_{n,i} \mathbb{P}_{n+1}(y)}{x_i - y_i} & \text{if } x \neq y, \\ \mathbb{P}_n^{\top}(x) H_n^{-1} A_{n,i} \frac{\partial}{\partial x_i} \mathbb{P}_{n+1}(x) - (A_{n,i}\mathbb{P}_{n+1}(x))^{\top} H_n^{-1} \frac{\partial}{\partial x_i} \mathbb{P}_n(x) & \text{if } x = y. \end{cases}\end{aligned} \tag{2.41}$$

PROOF: See [129, Thm. 3]. □

The matrices H_k can be obtained from the positive definite linear functional $\mathcal{L}$ of Theorem 2.10, *i.)*, e.g. the inner product $\mathcal{L}(\,\cdot\,,\,\cdot\,) = \langle\,\cdot\,,\,\cdot\,\rangle$, via $H_k = \mathcal{L}(\mathbb{P}_k\mathbb{P}_k^{\top})$. This shows that the value of the sum $\sum_{k=0}^{n} \mathbb{P}_k^{\top}(x) H_k^{-1} \mathbb{P}_k(y)$ is independent from the choice of the bases $\mathbb{P}_k$ in $\mathcal{V}_k^d$, as it follows from the identity $\mathbb{P}_k^{\top}\mathcal{L}(\mathbb{P}_k\mathbb{P}_k^{\top})\mathbb{P}_k = \mathbb{Q}_k^{\top}\mathcal{L}(\mathbb{Q}_k\mathbb{Q}_k^{\top})\mathbb{Q}_k$ for any choice of bases $\mathbb{P}$ and $\mathbb{Q}$ in $\mathcal{V}_k^d$, cf. [129, Thm. 4].

Note that even though the right-hand side of (2.41) suggests that it depends on i, the left-hand side shows that it is actually independent of i.

We now extend the point of view, that orthogonality does not hold in terms of particular bases of $\mathcal{V}_n^d$ but in terms of the subspaces $\mathcal{V}_n^d$, to algebraic signal models.

Definition 2.12 *Two signal models* $(\mathscr{A}, M, \Phi_1)$ *and* $(\mathscr{A}, M, \Phi_2)$, *with bases of the modules given by sets of orthogonal polynomials* $\mathbb{P}_1$ *and* $\mathbb{P}_2$, *are called* insignificantly different *if* $\mathbb{P}_1$ *and* $\mathbb{P}_2$ *are orthogonal with respect to the same positive definite linear functional.*

The next thing we need to discuss are the common zeros of multivariate orthogonal polynomials. Since we are concerned with multivariate polynomials the zero set, i.e., the variety, of one polynomial is of dimension $d - 1$. Since we are interested in discrete sets it is obvious that we should consider the intersection of as many varieties such that we obtain a discrete set. Bézout's theorem suggests that in general we should consider r_n polynomials. As we need the zeros to be simple as well, we are interested in very special situations. The common zeros of $\mathbb{P}_n$ turn out to be as well-behaved as possible, if they exist.

Theorem 2.13 *All common zeros of $\mathbb{P}_n$ are real, distinct and simple, i.e. at least one $\frac{\partial}{\partial x_i}\mathbb{P}_n(x)$ is not vanishing. There exists at most $\dim \Pi^d_{n-1}$ common zeros and this bound is reached if one has for orthonormal polynomials $\mathbb{P}_n$ the equality*

$$A_{n-1,i}A^\top_{n-1,j} = A_{n-1,j}A^\top_{n-1,i} \tag{2.42}$$

for all $1 \leq i, j \leq d$.

Proof: See [130, Thm. 2.13]. □

We are now ready to prove the multivariate Gauß-Jacobi procedure as our first new theorem.

Theorem 2.14 *Let $\{\mathbb{P}_k\}_{k=0}^\infty$ be a sequence of multivariate orthogonal polynomials. Consider a signal model with underlying variety $V = \langle\, \mathbb{P}_n \,\rangle$ such that $|V| = \dim \Pi^d_{n-1}$ and let the choice of basis via Φ be given by $\{\mathbb{P}_k\}_{k=0}^{n-1}$. Then there exists an insignificantly different signal model which has an orthogonal Fourier transform.*

Proof: We can assume that $\{\mathbb{P}_k\}$ are orthonormal since the choice of an orthonormal instead of an orthogonal basis only leads to an insignificantly different signal model.

The common zeros of $\mathbb{P}_n$ are real and simple and thus at least one partial derivative of $\mathbb{P}_n$ evaluated at one such zero is non-zero. In the one-dimensional irreducible component corresponding to $\alpha \in V$ choose as basis

$$\sqrt{\mathbb{P}^\top_{n-1}(\alpha)A_{n-1,i}\tfrac{\partial}{\partial x_i}\mathbb{P}_n(\alpha)}.$$

Note that this choice is independent of the index i. This choice of basis leads to an orthogonal Fourier transform $\mathcal{F}^{\mathsf{orth}}$ as can be seen as follows. First observe that $\mathcal{F}^{\mathsf{orth}} = \mathcal{F}\sqrt{D}$, where $\mathcal{F}$ is the Fourier matrix where in each spectral component the basis $\{1\}$ had been choosen, i.e. it is of the form

$$\mathcal{F} = (p(\alpha))_{p\in\{\mathbb{P}_k\}_k, \alpha\in V},$$

and $\sqrt{D}$ is the diagonal matrix

$$\sqrt{D} = \operatorname{diag}\Big(1\Big/\sqrt{\mathbb{P}^\top_{n-1}(\alpha)A_{n-1,i}\tfrac{\partial}{\partial x_i}\mathbb{P}_n(\alpha)}\ \Big|\ \alpha \in V\Big).$$

One then has

$$(\mathcal{F}^{\mathsf{orth}})^\top \cdot \mathcal{F}^{\mathsf{orth}} = \sqrt{D}\cdot\mathcal{F}^\top\cdot\mathcal{F}\cdot\sqrt{D}.$$

The product $\mathcal{F}^\top\cdot\mathcal{F}$ has entries of the form $\sum_{k=0}^{n-1}\mathbb{P}_k(\alpha)^\top\mathbb{P}_k(\beta)$ for $\alpha, \beta \in V$. Now by the Xu-Christoffel-Darboux formula (2.41) and the fact that V consists of common zeros of all the elements of $\mathbb{P}_n$, all non-diagonal elements vanish. The diagonal elements are of the form $\mathbb{P}^\top_{n-1}(\alpha)A_{n-1,i}\frac{\partial}{\partial x_i}\mathbb{P}_n(\alpha)$. These elements do not vanish since the common zeros of $\mathbb{P}_n$ are simple. Consequently, the matrix D is well-defined and it follows that $\mathcal{F}^{\mathsf{orth},\top}\cdot\mathcal{F}^{\mathsf{orth}} = \mathbb{1}$. □

Unfortunately the multivariate Gauß-Jacobi procedure is rarely applicable because of the condition on the number of common zeros in Theorem 2.14, which is the same as for the existence of a Gaußian cubature formula. A cubature formula of degree $2n-1$ is a finite sum that approximates an integral and is exact for all $p \in \Pi^d_{2n-1}$, i.e.,

$$\int_{\mathbb{R}^d} p(x)\,\mathrm{d}\mu = \sum_{k=1}^{N} w_k p(x_k), \tag{2.43}$$

and there is a $p^* \in \Pi^d_{2n}$ such that the equality does not hold. The $w_k \in \mathbb{R}$ are the weights of the cubature formula while the $x_k \in \mathbb{R}^d$ are the nodes. The number of nodes N fulfills the lower bound

$$N \geq \dim \Pi^d_{n-1}. \tag{2.44}$$

If the lower bound is actually reached the cubature formula is called Gaußian. The following theorem due to Mysovskikh [80] characterizes Gaußian cubature formulae in terms of common zeros of multivariate orthogonal polynomials.

Theorem 2.15 *A Gaußian cubature formula exists if and only if $\mathbb{P}_n$ has $\dim \Pi^d_{n-1}$ common zeros.*

PROOF: See [80]. □

While in the univariate case Gaußian quadrature formulae always exists, this is not the case in the multivariate setting. For example if the measure μ of the integral 2.43 is centrally symmetric then no Gaußian cubature formula exists, cf. [130, Cor. 2.16]. This rules out typical regions of multivariate orthogonality like cubes, balls or simplices. But since the approximation of integrals is a well-studied topic, there exist examples of Gaußian cubature formulae and we can use the corresponding orthogonal polynomials to construct examples of signal models with orthogonal transforms.

The first class of examples of the construction of Gaußian cubatures where presented by Schmid and Xu [108] in the bivariate case. Their construction relied on bivariate polynomials introduced by Koornwinder [52]. For these historical reasons we study the signal model of these polynomials in the following example. Indeed it will turn out that these Koornwinder polynomials in certain special cases coincide (up to normalization) with certain multivariate Chebyshev polynomials investigated in Chapter 4.

Example 2.16 Let $n \in \mathbb{N}_0$ and $u = x + y$ as well as $v = x \cdot y$. Let $\{p_n\}$ be orthogonal polynomials with respect to a weight function w on an interval I and denote by $x_{k,n}$ the zeros of p_n. Let the p_n satisfy the recurrence relation $xp_n = a_n p_{n+1} + b_n p_n + c_n p_{n-1}$. The bivariate orthogonal Koornwinder polynomials are then defined as

$$P_k^{n,-1/2}(u,v) = \begin{cases} p_n(x) \cdot p_k(y) + p_n(y) \cdot p_k(x) & \text{if } k < n, \\ \sqrt{2} p_n(x) \cdot p_n(y) & \text{if } k = n, \end{cases} \tag{2.45}$$

and

$$P_k^{n,1/2} = \frac{p_{n+1(x)} \cdot p_k(y) - p_{n+1}(y) \cdot p_k(y)}{x - y}. \tag{2.46}$$

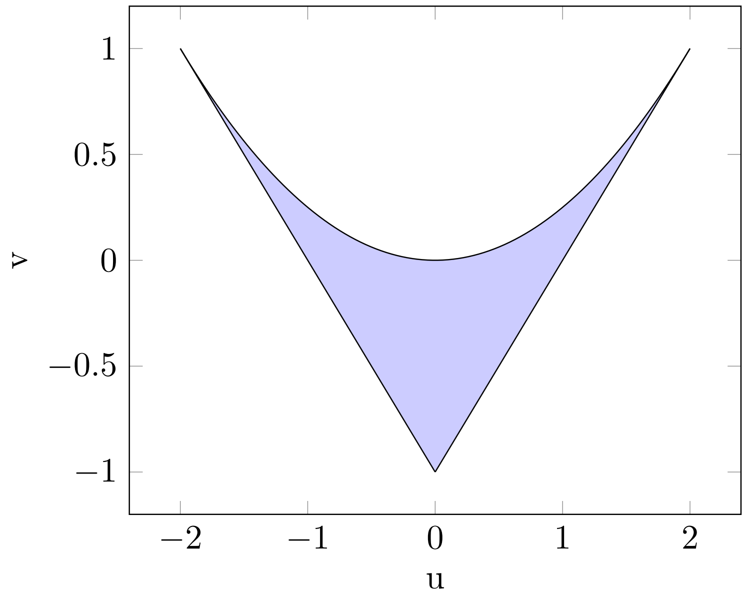

Figure 2.5: The region R of orthogonality of the Koornwinder polynomials for $I = [-1, 1]$.

Each $P_k^{n,\pm 1/2}$ is a polynomial of total degree n. They are orthogonal polynomials on $R = \{(u, v) \mid (x, y) \in I \times I \text{ such that } x < y\}$, cf. Fig.2.5, with respect to the weight function $W(u, v) = (u^2 - 4v)^{\pm\frac{1}{2}} w(x)w(y)$. For more pleasant formulas we restrict to ortho*normal* polynomials, but the reasoning is the same for orthogonal polynomials. Note that for orthonormal polynomials one has $c_n = a_{n-1}$. The polynomials (2.45) and (2.46) obey the three-term recurrence formula (2.39) with matrices

$$A_{n,1}^{\gamma} = a_{n,\gamma} \begin{bmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ & & f_\gamma & 0 \end{bmatrix}, \tag{2.47}$$

and

$$A_{n,2}^{\gamma} = a_{n,\gamma} \begin{bmatrix} b_0 & a_0 & & & & 0 \\ a_0 & b_1 & a_1 & & & \vdots \\ & \ddots & \ddots & \ddots & & \vdots \\ & & a_{n-2} & b_{n-1} & a_{n-1} & 0 \\ & & & f_\gamma a_{n-1} & f_\gamma b_n & a_n \end{bmatrix}, \tag{2.48}$$

with $f_{-\frac{1}{2}} = \sqrt{2}, a_{n,-\frac{1}{2}} = a_n, f_{\frac{1}{2}} = 1$, and $a_{n,\frac{1}{2}} = a_{n+1}$, as well as the matrices

$$B_{n,1}^{\gamma} = \begin{bmatrix} b_0 & a_0 & & & 0 \\ a_0 & b_1 & a_1 & & \vdots \\ & \ddots & \ddots & \ddots & \vdots \\ & & a_{n-2} & b_{n-1} & f_\gamma a_{n-1} \\ & & & f_\gamma a_{n-1} & b_n \end{bmatrix} + b_{n,\gamma} \mathbb{1}_{n+1}, \tag{2.49}$$

$$B_{n,2}^{-1/2} = b_{n,\gamma} \begin{bmatrix} b_0 & a_0 & & & \\ a_0 & b_1 & a_0 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-2} & b_{n-1} & f_\gamma a_{n-1} \\ & & & f_\gamma a_{n-1} & b_n \end{bmatrix} + a_{n-1}^2 \begin{bmatrix} 0 & \dots & & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & 0 & & \vdots \\ \vdots & & & 1 & 0 \\ 0 & \dots & & 0 & 0 \end{bmatrix}, \tag{2.50}$$

and

$$B_{n,2}^{1/2} = b_{n,\gamma} \begin{bmatrix} b_0 & a_0 & & & \\ a_0 & b_1 & a_0 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-2} & b_{n-1} & f_\gamma a_{n-1} \\ & & & f_\gamma a_{n-1} & b_n \end{bmatrix} - a_n^2 \begin{bmatrix} 0 & & \dots & & 0 \\ \vdots & \ddots & & & \vdots \\ & & & & \\ & & & 0 & 0 \\ 0 & \dots & & & 1 \end{bmatrix}, \tag{2.51}$$

with $b_{n,-1/2} = b_n$ and $b_{n,1/2} = b_{n+1}$. The $\dim \Pi_{n-1}^2$ common zeros of $\mathbb{P}_n^{-1/2}$ are given by $(x_{k,n} + x_{j,n}, x_{k,n} \cdot x_{j,n})$ for $j \leq k$ and the common zeros of $\mathbb{P}_n^{1/2}$ are given by $(x_{k,n+1} + x_{j,n+1}, x_{k,n+1} \cdot x_{j,n+1})$ for $j < k$.

Now consider the signal model with algebra $\mathscr{A} = \Pi^2 / \langle\, \mathbb{P}_n \,\rangle$, regular module $M = \mathscr{A}$, and z-transform

$$\begin{aligned} \Phi \colon \mathbb{C}^{\dim \Pi_{n-1}^2} &\longrightarrow M, \\ (s_{k,\ell}) &\mapsto \sum_{k<\ell<n} s_{k,\ell} P_k^{\ell}. \end{aligned} \tag{2.52}$$

By the matrices for the three-term recurrence relations (2.47), (2.48), (2.49), (2.50), and (2.51) the visualization in Fig. 2.6 is obtained. Since the number of common zeros of $\mathbb{P}_n$, i.e. $|\mathsf{V}(\langle\, \mathbb{P}_n \,\rangle)|$, is $\dim \Pi_{n-1}^2$ and we have orthonormal polynomials, the Fourier transform of these signal models are orthogonal by Theorem 2.13. Note that one can not state the Fourier transformation explicitly, since no closed form for the zeros of the univariate orthogonal polynomials is available.

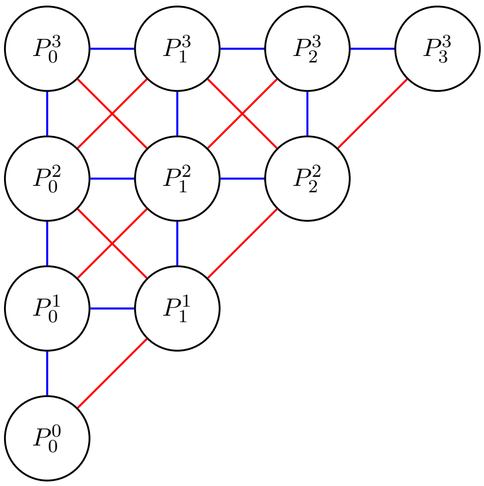

Figure 2.6: Visualization of the signal models relying on the Koornwinder polynomials for $n = 4$. Boundary conditions and weights are omitted. Red corresponds to the u-shift, while blue corresponds to the v-shift.

Chapter 3

FFT-like algorithms

The discrete Fourier transform would probably not have found not as many applications as it currently has, if there would not exists a fast algorithm for its computation. The fast Fourier transform (FFT) has thus been termed to be one of the most important algorithms of the twentieth century, cf. [101].

This chapter is concerned with two approaches to fast Fourier transform algorithms for algebraic signal models. The first is relying on induced modules and generalizes the method developed in [104] for signal models based on univariate polynomials to the multivariate case. The second approach relies on a decomposition property of special polynomials. In the univariate case the second approach is a special case of the first one [104].

The general idea of FFT algorithms is a divide-and-conquer approach. That is one breaks the problem of calculating the Fourier transform into smaller pieces, i.e. Fourier transforms of smaller sizes. By a recursive application of this method one obtains a substantially faster method for computation of the Fourier transform than the naive approach of evaluating the matrix-vector products directly.

In a mathematically more rigorous setting this divide-and-conquer approach can be implemented as factorization of the Fourier transform matrix into sparse matrices, cf. [92]. Since in algebraic signal processing theory one likes to obtain everything on the level of algebras and modules, the question arises how the factorization into sparse matrices translates into the structure of the modules. Indeed in this chapter it is shown that the sparse factorizations stem from partial decompositions of the signal module, cf. [92, 96, 104].

3.1 Fast Fourier transform via induction

We need to recall some tools from the representation theory of algebras. Let $\mathscr{A}$ be an algebra and $\mathscr{B}$ be a subalgebra of $\mathscr{A}$. A finite set $T \subseteq \mathscr{A}$ is called a transversal of $\mathscr{B}$ in $\mathscr{A}$ if

$$\mathscr{A} = \bigoplus_{t \in T} t\mathscr{B} \tag{3.1}$$

as vector spaces. In the sequel $\cdot \triangleright \cdot$ denotes the action of an algebra on a module. If M is a $\mathscr{B}$-module the module $\bigoplus_{t\in T} t \triangleright M$, with vector space direct sum, is called the T-induced $\mathscr{A}$-module of M.

The main idea behind fast algorithms using induction is stated in the following commutative diagram. For ease of notation the modules in the diagram are assumed to be regular ones, even though the concept works for any module, cf. the Chinese remainder theorem for modules B.1.

$$\begin{tikzcd}
N \arrow[rr, "\text{change of basis}"] \arrow[d, "\text{Fourier}"] & & \bigoplus_t t \triangleright M_t \arrow[d, "\text{Fourier}"] \\
\bigoplus_i \Pi^n/\langle\, x - \alpha_i \,\rangle & & \bigoplus_t \bigoplus_j \Pi^n(y)/\langle\, y - \beta_{t,j} \,\rangle \arrow[ll, "\text{project and scale}"']
\end{tikzcd} \tag{3.2}$$

The diagram (3.2) shows that, possibly using a change of basis, one can break the problem of calculating the Fourier transform of an induced module into the less computational costly problem of calculating the Fourier transform of lower-dimensional modules. Repeated application of this idea will, under certain mild assumptions, lead to a fast recursive algorithm. This is the content of the following theorem, which generalizes [104, Thm. 5.1].

Theorem 3.1 (FFT algorithms via induced modules) *Let $\mathscr{A}$ be an algebra and let $\mathscr{B}$ be a subalgebra of $\mathscr{A}$. Let $M_t = \Pi^d(y)/J_t$ be a set of $\mathscr{B}$-modules such that $N = \bigoplus_{t\in T} t \triangleright M_t = \Pi^n/I$, with $T = \{t_1, \dots, t_w\} \subset \mathscr{A}$ a finite set, is an $\mathscr{A}$-module. Assume the action of $t \in T$ on M_t is by multiplication with a polynomial t. Let $r_t\colon V(I) \longrightarrow V(J_t)$ be a surjective map between the corresponding varieties. The Fourier transform of N with respect to a basis b_N can be decomposed as*

$$P_{b_N,N} = [D_1R_1|\dots|D_wR_w]\left[\bigoplus_{t\in T} P_{b_t,t\triangleright M}\right] B^{b_N}_{\oplus b_t}, \tag{3.3}$$

where $B^{b_N}_{\bigoplus b_t}$ is the change of basis from the basis b_N to the concatenation of bases of the $t \triangleright M_t$, the matrices $P_{b_t,t\triangleright M_t}$ are the Fourier transforms of the $t \triangleright M_t$, the matrices R_t are matrices with entries $[R_t]_{\alpha\in \mathsf{V}(I),\beta\in\mathsf{V}(J)}$ being 1 if $r_t(\alpha) = \beta$ and 0 otherwise, and the $D_t = \operatorname{diag}(t(\alpha)|\alpha \in \mathsf{V}(I))$.

Proof: First note that since t acts as multiplication by a polynomial any element of $t \triangleright M_t$ can be written as $t \cdot b$ (by identifying t with the polynomial as which t acts by multiplication) and denote the chosen basis without the t as b_t. Then the claim follows from the following unwinding of definitions

$$\begin{aligned}
P_{b_N,N} &= [b(\alpha)]_{b\in b_N,\alpha\in\mathsf{V}(I)} \\
&= [t(\alpha)b(r(\alpha))_{b\in\oplus b_t,\alpha\in\mathsf{V}(I)}|t\in T]B^{b_N}_{\oplus b_t} \\
&= [\operatorname{diag}(t(\alpha)|\alpha\in\mathsf{V}(I))R_t(b(\beta))_{b\in b_t,\beta\in\mathsf{V}(J_t)}]B^{b_N}_{\oplus b_t} \\
&= [D_1R_1|\dots|D_wR_w]\left[\bigoplus_{t\in T} P_{b_t,t\triangleright M_t}\right] B^{b_N}_{\oplus b_t}.
\end{aligned}$$

Here $(b(r(\alpha))_{b\in b_t,\alpha\in\mathsf{V}(I)} = R_t(b(\beta))_{b\in b_t,\beta\in\mathsf{V}(J_t)}$ follows since $r\colon \mathsf{V}(I) \longrightarrow \mathsf{V}(J_t)$ is onto and R_t keeps track of this map. The result follows. $\square$

Since the theorem captures very general situations, it is not possible to derive a better general statement than $O(n^2)$ for the computational cost of the obtained algorithms. This is since in general we do have only the trivial $O(n^2)$ estimate for the computational cost of the matrices B and R_i. But if we assume them to be of linear cost and if we can find a suitable descending chain of submodules these algorithms are of cost $O(k\log(k))$, where $k = |\mathsf{V}(I)|$. Then the following proposition is a simple consequence of the Akra-Bazzi-Theorem [1], a refined version of the Master Theorem for divide and conquer recurrences [3], cf. App. C for some more information on this theorem.

Proposition 3.2 *Consider the decomposition of the Fourier transform from Theorem 3.1 and assume one has a descending chain of submodules, where in each step we have a split in at least two submodules. If the basis change matrices B and the M_i in each step are $O(k)$ then the decomposition is $O(k \cdot \log(k))$ with $k = |\mathsf{V}(I)|$.*

Finding a descending chain of submodules is no problem as one can collect random points of the variety but this typically leads to neither sparse B nor sparse R_i. Hence the main difficulty for an effective application of the theorem is finding good examples.

We want to investigate, how we can ensure the existence of a transversal. We start by characterizing subalgebras generated by exactly the number of variables many generators. This is done in terms of the image of the variety under the image of the generators of the subalgebra.

Proposition 3.3 *Let $\mathscr{B} \subseteq \mathscr{A} = \Pi^n/I$ be a finitely generated subalgebra, s.t. $\mathscr{B} = \langle\, r_1, \ldots, r_n \,\rangle$ for $r_i \in \mathscr{A}$. Then as algebras*

$$\mathscr{B} \cong \Pi^n(y)/J, \tag{3.4}$$

where $J = \mathsf{I}((r_1, \ldots, r_n)(\mathsf{V}(I)))$ is the ideal of the image of $\mathsf{V}(I)$ under the generators of $\mathscr{B}$ in $\mathbb{C}[y_1, \ldots, y_n]$.

Proof: To proof (3.4), we show that both sides have the same dimension and the kernel of an algebra homomorphism between them is trivial.

Denote the finite variety by $\{\alpha_1, \ldots, \alpha_k\} = \mathsf{V}(I)$. Let $\{\beta_1, \ldots, \beta_\ell\}$ be the image of these points under $(r_1, \ldots, r_n)$. Then $\ell \leq k$.

Claim 3.4 $\dim \mathscr{B} = \ell$.

Proof: We prove Claim 3.4. We can write

$$I = \prod_i \langle\, x_1 - \alpha_{i,1}, \ldots, x_n - \alpha_{i,n} \,\rangle.$$

Each of the $\langle\, x_1 - \alpha_{i,1}, \ldots, x_n - \alpha_{i,n} \,\rangle$ is maximal, hence they are all coprime and we can use the Chinese remainder theorem B.2 to decompose $\mathscr{A}$. Denote by

$$\mathcal{F}\colon \mathscr{A} \longrightarrow \bigoplus_i \Pi^n/\langle\, x_1 - \alpha_{i,1}, \ldots, x_n - \alpha_{i,n} \,\rangle$$

the isomorphism from equation (2.21). The diagram

$$\begin{tikzcd}
\ker(\mathrm{pr}_{\mathscr{B}}) \arrow[d] & & \ker(\mathrm{pr}) \arrow[d] \\
\mathscr{A} \arrow[rr, "\mathcal{F}"] \arrow[d, "\mathrm{pr}_{\mathscr{B}}"'] & & \bigoplus_i \Pi^n/\langle\, x_1-\alpha_{i,1},\dots,x_n-\alpha_{i,n}\,\rangle \arrow[d, "\mathrm{pr}"] \\
\mathscr{B} \arrow[rr, "\mathcal{F}"] \arrow[d] & & \mathcal{F}(\mathscr{B}) \arrow[d] \\
0 & & 0
\end{tikzcd}$$

commutes. Hence it suffices to determine the dimension of $\ker(\mathrm{pr})$, to determine the dimension of $\mathscr{B}$. But the dimension of $\ker(\mathrm{pr})$ is given by the number of α_i, which get mapped to the same β_j under the r_i, so $\dim\ker(\mathrm{pr}) = k-\ell$. Henceforth $\dim\mathscr{B} = \dim\mathcal{F}(\mathscr{B}) = k - \dim\ker(\mathrm{pr}) = \ell$. Hence $\dim\mathscr{B} = \dim\Pi^n(y)/J$, as J is radical and $|\mathsf{V}(J)| = \ell$. This proves claim 3.4. $\triangledown$

Consider the algebra homomorphism

$$\kappa\colon \mathscr{B} \longrightarrow \Pi^n(y)/J,$$
$$r_i \mapsto y_i,$$

which maps generators to generators. We have the short exact sequence

$$0 \longrightarrow \kappa^{-1}(J) \longrightarrow \mathscr{B} \xrightarrow{\ \kappa\ } \Pi^n(y)/J \longrightarrow 0$$

and hence $\Pi^n(y)/J \cong \mathscr{B}/\kappa^{-1}(J)$. So we still need to show:

Claim 3.5 $\ker(\kappa) = \{0\}$.

PROOF: We prove Claim 3.5. It suffices to show, that the r_i vanish on the ideal $\ker(\kappa) = \kappa^{-1}(J)$. As J is the ideal of the points β_i, it can be written as

$$J = \prod_j \langle\, y_1-\beta_{j,1},\dots,y_n-\beta_{j,n}\,\rangle.$$

So $\kappa^{-1}(J) = \prod_i\langle\, r_1-\beta_{i,1},\dots,r_n-\beta_{i,n}\,\rangle$. Now the isomorphism $\mathcal{F}$ maps the r_i to the β_i, as the β_i are the image of the α_i under r_i and $\mathcal{F}$ is, by (2.21), just inserting α_i into the polynomials. So $\mathcal{F}(\kappa^{-1}(J)) = \{0\}$, hence $\kappa^{-1}(J) = \{0\}$ in $\mathscr{A}$ and evidently in $\mathscr{B}$ as well, as $\mathscr{B}$ is a subalgebra of $\mathscr{A}$. Hence the Claim 3.5 is proved. $\triangledown$

By Claim 3.4 and Claim 3.5 we have proved the proposition. $\square$

Hence in this case there always exists a transversal of $\mathscr{B}$ in $\mathscr{A}$, as one can choose each $t \in T$ such that $t(a_i) = 0$ and $t(a_\ell) \neq 0$ for one $a_\ell \in \mathsf{V}(I)$. Then each $t\mathscr{B}$ has dimension 1, and hence $\dim\bigoplus_{t\in T} t\mathscr{B} = \dim\bigoplus_{a\in\mathsf{V}(I)} \Pi^n(x)/\langle\, x_1-a_1,\dots,x_n-a_n\,\rangle$.

Thus they are isomorphic as vector spaces and by (2.18) to $\mathscr{A}$ as well. Note that this choice is a useless one for the development of fast algorithms, as we have no intermediate steps and hence one does not obtain a recursive structure which can be exploited for speeding up calculations. Nonetheless this is a necessary remark, as now we can always assume a transversal existent.

Choose a transversal T of the subalgebra $\mathscr{B}$ in $\mathscr{A}$. The next step is to show that the structure of the $\mathscr{B}$-modules $t \triangleright M$ for $\mathscr{B}$-modules of the form $M = \Pi^n(y)/J$ with zero-dimensional, radical ideal J and $t \in T$ is again a polynomial module. Hence one gets a descending chain of submodules where one can easily the describe the corresponding Fourier transforms.

Proposition 3.6 *Let $\mathscr{A}$ be an algebra with subalgebra $\mathscr{B}$ and let T be a finite transversal of $\mathscr{B}$ in $\mathscr{A}$. Let $M = \Pi^n(y)/J$ be a $\mathscr{B}$-module and $\bigoplus t \triangleright M = \Pi^n/I$ the induced $\mathscr{A}$-module. There exists a map $r\colon \mathsf{V}(I) \longrightarrow \mathsf{V}(J)$. The action of the transversal elements leads to $\mathscr{B}$-modules of the form*

$$t \triangleright M \cong \mathbb{C}[y_1, \dots, y_n]/J_t, \tag{3.5}$$

where $J_t = \mathsf{I}(\{r(\alpha) \mid \alpha \in \mathsf{V}(I) \text{ and } t_p(\alpha) \neq 0\}$.

PROOF: The existence of the map r is clear, since $\mathscr{B}$ is a subalgebra of $\mathscr{A}$. Hence T must contain 1 and thus M is a submodule of $\bigoplus t \triangleright M$. Therefore r can be choosen as a projection of $\mathsf{V}(I)$ onto its subset $\mathsf{V}(J)$.

It suffices to show that the $\mathscr{B}$-modules on both sides of 3.5 are of equal dimension. The isomorphism from the Chinese remainder theorem for $\bigoplus t \triangleright M$ leads for the subset $t \triangleright M$ to

$$t_p p \mapsto (t_p(\alpha) p(r(\alpha)))_{\alpha \in \mathsf{V}(I)},$$

for any $p \in M$. Denote by $[\alpha]$ the equivalence class of $\alpha \in \mathsf{V}(I)$ which map to the same $\beta \in \mathsf{V}(J)$. The dimension of $t \triangleright M$ is $|\mathsf{V}(J)|$ minus one for each $[\alpha]$ where $t_p(\alpha) = 0$. Restricting to J_t hence does not change the dimension. The proposition is proven. □

Remark 3.7 Note that the map $r\colon \mathsf{V}(I) \longrightarrow \mathsf{V}(J)$ from Prop. 3.6 can be explicitly determined if N and M are regular and M is a *subalgebra* of N. Then the map is just the set of generators $r = (r_1, \dots, r_n)$ from Prop. 3.3.

The first example is the classical FFT algorithm for the Fourier transform of the discrete finite time signal model from Example 2.3 and is from [104].

Example 3.8 The module $\mathbb{C}[x]/\langle\, x^4 - 1 \,\rangle$ can be represented as $1 \triangleright \mathbb{C}[y]/\langle\, y^2 - 1 \,\rangle \oplus x \triangleright \mathbb{C}[y]/\langle\, y^2 - 1 \,\rangle$ with transversal $T = \{1, x\}$. The algebra action is multiplication modulo $\langle\, x^4 - 1 \,\rangle$ and $y = x^2$. The change of basis B is from $\{1, x, x^2, x^3\}$ to $\{1, x^2, x, x^3\} = \{1 \triangleright 1, 1 \triangleright y, x \triangleright 1, x \triangleright y\}$ and thus

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The direct sum is $\mathbb{C}[y]/\langle\, y^2-1 \,\rangle \oplus x\mathbb{C}[y]/\langle\, y^2-1 \,\rangle$ as modules leading to the matrix

$$\mathsf{DFT}_2 \oplus \mathsf{DFT}_2 = \begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix}.$$

The matrices R_1 and R_2 keeping track of the map between the varieties are given by

$$R_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

since $r = x^2$ maps $\mathsf{V}(\langle\, x^4-1 \,\rangle) = \{1, \mathrm{i}, -1, -\mathrm{i}\}$ onto $\{1, -1, 1, -1\}$. The diagonal matrix D_1 is the identity since the polynomial 1 evaluates always to 1, while

$$D_2 = \begin{bmatrix} 1 & & & \\ & \mathrm{i} & & \\ & & -1 & \\ & & & -\mathrm{i} \end{bmatrix},$$

since the polynomial x corresponds to the identity map. Hence we obtain

$$[D_1M_1 \mid D_2M_2] = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \mathrm{i} \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\mathrm{i} \end{array}\right].$$

We obtained the sparse matrix factorization

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\mathrm{i} \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & \mathrm{i} \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.6}$$

and thus a fast algorithm for the computation of the Fourier transform of the discrete, finite time signal model. Indeed for the calculation of a matrix-vector product using the left-hand side of (3.6) one needs 12 complex additions and 8 complex multiplications, while using the right-hand side the same result is obtained using only 8 complex additions and 5 complex multiplications.

We now give an example, i.e. the FFT on a directed hexagonal lattice, which shows how one can derive FFTs on various lattices from the literature. The derivation of FFTs on regular directed lattices was first obtained in [75]. See [132] for more concrete examples using the classical derivation. The example illustrates a reverse engineering approach to obtain these algorithms by algebraic signal processing theory, as well.

Example 3.9 We reverse engineer the FFT of a directed hexagonal lattice from Mersereau [74] by algebraic signal processing theory. Assume $N = 2^k$ for some $k > 1$. Recall from [74] that the discrete Fourier transform for a signal s_{n_1,n_2} sampled on an hexagonal lattice is given as

$$\mathcal{F}(s)_{k_1,k_2} = \sum_{n_1=0}^{3N-1} \sum_{n_2=0}^{N-1} s_{n_1,n_2} \exp(-\tfrac{-\pi \mathrm{i}}{3N}((2n_1 - n_2)(2k_1 - k_2) + 6n_2k_2)). \tag{3.7}$$

From this formula and the definition of Fourier transforms corresponding to zero-dimensional varieties (2.19) it is evident that the variety is given by the points

$$\big\{\big(\exp\big(\tfrac{-\pi\mathrm{i}(2k_1-k_2)}{3N}\big), \exp\big(\tfrac{2\pi\mathrm{i}k_2}{N}\big)\big) \mid k_1 = 0, \dots, 3N-1; \quad k_2 = 0, \dots, N-1\big\}. \tag{3.8}$$

The basis is determined by (3.7), as well, and consists of elements $x^{2n_1-n_2}y^{n_2}$ for $n_1 = 0, \dots, 3N-1$ and $n_2 = 0, \dots, N-1$.

The vector space underlying the module is hence given by

$$M = \mathbb{C}[x^2, xy]/\langle\, y^N - 1, x^{3N} - y^{N/2} \,\rangle. \tag{3.9}$$

Now we have to expose for which algebra we can find a module structure, such that we get a hexagonal model and an FFT-like algorithm. Unlike one might speculate at first, one realizes the module structure of M not as a module over a polynomial algebra in two variables but in three indeed. For this consider the algebra $\mathscr{A} = \mathbb{C}[X_1, X_2, X_3]/\langle\, X_1^{3N} - 1, X_2^{N/2} - 1, X_3^{N/2} - 1 \,\rangle$, with actions on M given by

$$\begin{aligned} X_1 \triangleright p(x,y) &= x^2 \cdot p(x,y), \\ X_2 \triangleright p(x,y) &= xy \cdot p(x,y), \\ X_3 \triangleright p(x,y) &= x^{-1}y \cdot p(x,y). \end{aligned} \tag{3.10}$$

The resulting visualization graph of the signal model is shown in Fig. 3.1.

The signal module can be decomposed into submodules. The choice of lattice cosets in [74] corresponds to the choice of the submodule $S = \mathbb{C}[r^2, rs]/\langle\, s^{N/2} - 1, r^{3N/2} - s^{N/4} \,\rangle$ with $r = x^2$ and $s = y^2$. We need to find a subalgebra and transversal of the underlying algebra, which results in the induced module of S being M. Consider the subalgebra $\mathscr{B} = \mathbb{C}[Y_1, Y_2, Y_3]/\langle\, Y_1^{3N/2} - 1, Y_2^{N/4} - 1, Y_3^{N/4} - 1 \,\rangle$. A transversal of $\mathscr{B}$ in $\mathscr{A}$ is $\{1, X_1, X_2, X_3\}$. The action of the transversal elements on S is realized by multiplication with the polynomials $\{1, x^2, xy, x^{-1}y\}$. Then one obtains

$$M = S + x^2S + xyS + x^{-1}yS. \tag{3.11}$$

The sublattice corresponding to the transversal element 1 is depicted in Fig. 3.2. From the structure of the submodule and the transversal it is obvious that the change of basis to the induced module is a permutation matrix, hence is sparse.

None of the elements of $\mathsf{V}(\langle\, y^N - 1, x^{3N} - y^{N/2} \,\rangle)$ gets mapped to zero by an element of the transversal. The preimage of each point of $V(\langle\, y^{N/2}-1, x^{3N/2}-y^{N/4} \,\rangle)$ consists at most of four points of $V(\langle\, y^N - 1, x^{3N} - y^{N/2} \,\rangle)$. Hence each row of M has at most four non-zero entries, thus M has $O(n)$ entries and is sparse. Thus by Prop. 3.2 we have indeed a fast algorithm.

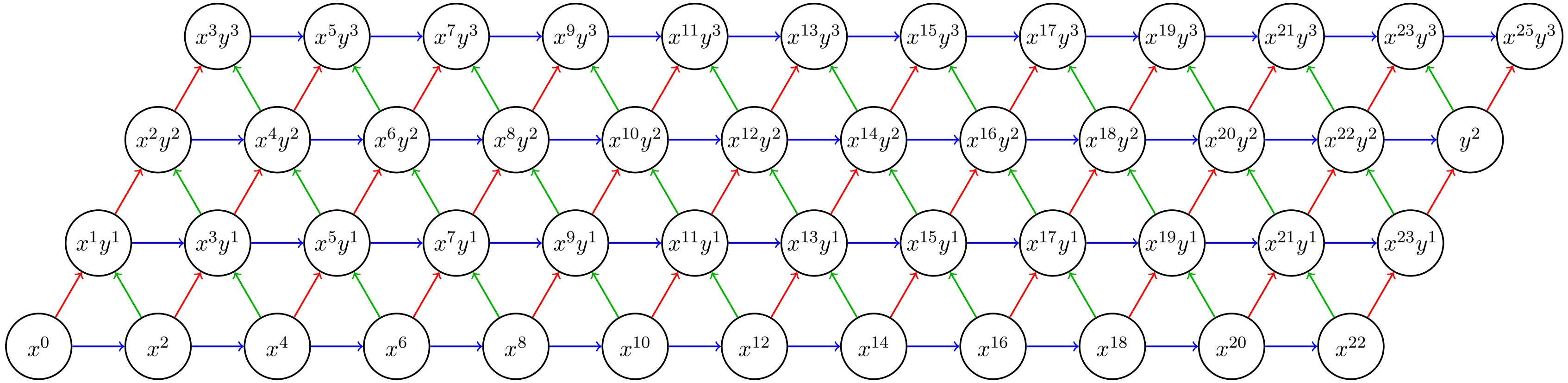

Figure 3.1: Signal model of a directed hexagonal lattice for $N = 4$. The boundary conditions are omitted. Shifts of X are blue, shifts of Y are red, and shifts of Z are green colored.

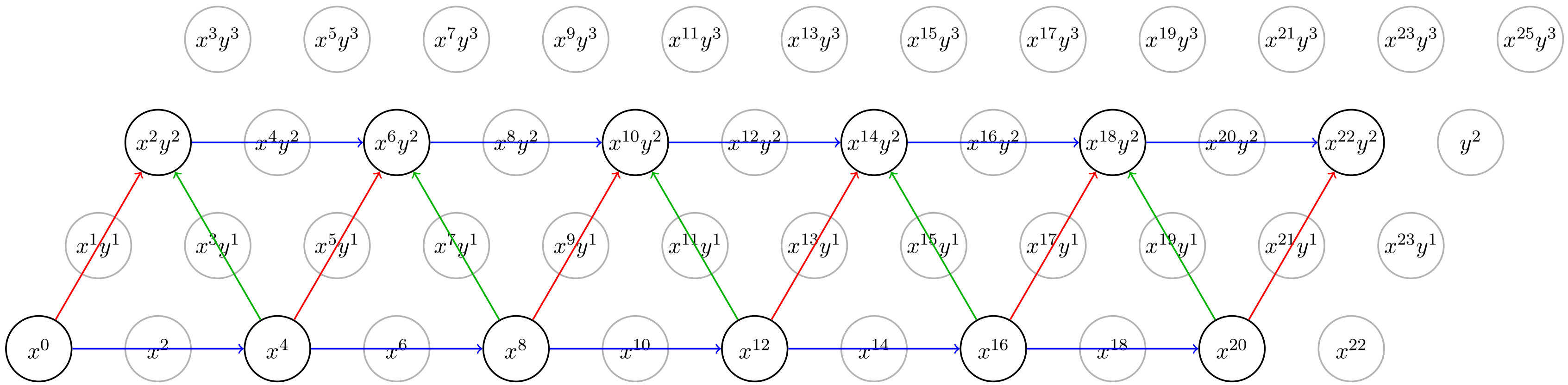

Figure 3.2: Sublattice of the hexagonal lattice corresponding to the transversal element 1.

Remark 3.10 In [94] a signal model for the directed quincunx lattice was introduced. This signal model used a basis similar to the one we used in Example 3.9. These examples show that the algebra action on the module is indeed crucial for the signal model.

For a fast recursive algorithm one needs a chain of descending submodules. The decomposition property $p(x) = q(r(x))$ is very useful for the development of fast algorithms in one variable as from the following propositions one obtains a nice chain of subalgebras, cf. [89, 92]. The several variables analog of the decomposition property reads

$$(p_1, \ldots, p_n) = (q_1(r_1, \ldots, r_n), \ldots, q_n(r_1, \ldots, r_n)). \tag{3.12}$$

Since this notation is rather opulent, we abbreviate $\langle\ p\ \rangle = \langle\ p_1, \ldots, p_n\ \rangle$ and $\langle\ q(r)\ \rangle = \langle\ q_1(r_1, \ldots, r_n), \ldots, q_n(r_1, \ldots, r_n)\ \rangle$ if confusion with the one-variable case can be avoided by context. The decomposition property yields the existence of sufficiently well-behaved submodules.

Proposition 3.11 *Assume the zero-dimensional radical ideal $I = \langle\ p_1, \ldots, p_n\ \rangle$ satisfies*

$$\langle\ p\ \rangle = \langle\ q(r)\ \rangle.$$

Then $\langle\ r\ \rangle \cong \Pi^n(y)/\langle\ q\ \rangle$.

Proof: The mapping $(r_1, \ldots, r_n)$ maps $\mathsf{V}(I)$ to the variety of the $q_1, \ldots, q_n$, i.e. $(r_1, \ldots, r_n)(\mathsf{V}(I)) = \mathsf{V}(\langle\ q\ \rangle)$, as $\langle\ p\ \rangle = \langle\ q(r)\ \rangle$. By Proposition 3.3 one has $\langle\ r\ \rangle \cong \Pi^n(y)/\langle\ q\ \rangle$. Thus the proposition is proven. □

In the univariate case one can always obtain a transversal of the algebra $\langle\ r(x)\ \rangle \cong \mathbb{C}[y]/\langle\ q(y)\ \rangle$ from a basis of $\mathbb{C}[x]/\langle\ r(x)\ \rangle$ since the number of common zeros of univariate, separable $r - \alpha$ is independent of α. In the multivariate case this is not always the case. The next proposition formalizes this in terms of the appearing varieties.

Proposition 3.12 *Consider $\langle\ p_1, \ldots, p_n\ \rangle = \langle\ q_1(r_1, \ldots, r_n), \ldots, q_n(r_1, \ldots, r_n)\ \rangle$ with zero-dimensional variety. If $|\mathsf{V}(\langle\ p\ \rangle)| \neq |\mathsf{V}(\langle\ r\ \rangle)| \cdot |\mathsf{V}(\langle\ q\ \rangle)|$ then no basis of $\Pi^n/\langle\ r\ \rangle$ is a transversal of $\Pi^n(y)/\langle\ q\ \rangle$ in $\Pi^n/\langle\ p\ \rangle$. If $|\mathsf{V}(\langle\ p\ \rangle)| = |\mathsf{V}(\langle\ r\ \rangle)| \cdot |\mathsf{V}(\langle\ q\ \rangle)|$ then any basis of $\Pi^n/\langle\ r\ \rangle$ is a transversal of $\Pi^n(y)/\langle\ q\ \rangle$ in $\Pi^n/\langle\ p\ \rangle$.*

Proof: If $|\mathsf{V}(\langle\ p\ \rangle)| \neq |\mathsf{V}(\langle\ r\ \rangle)| \cdot |\mathsf{V}(\langle\ q\ \rangle)|$ the dimensions of $\Pi^n(y)/\langle\ q\ \rangle$ and $\Pi^n/\langle\ r\ \rangle$ do not multiply to the dimension of $\Pi^n/\langle\ p\ \rangle$ so a basis of $\Pi^n/\langle\ r\ \rangle$ can not be a transversal of $\Pi^n(y)/\langle\ q\ \rangle$.

For the second part observe that if $\{Q_1, \ldots, Q_{q_d}\}$ is a basis of $\Pi^n(y)/\langle\ q\ \rangle$ and $\{R_1, \ldots, R_{r_d}\}$ is a basis of $\Pi^n/\langle\ r\ \rangle$ then

$$\begin{bmatrix} R_1 Q_1(r_1, \ldots, r_n) & \ldots & R_1 Q_{q_d}(r_1, \ldots, r_n) \\ \vdots & & \vdots \\ R_{r_d} Q_1(r_1, \ldots, r_n) & \ldots & R_{r_d} Q_{q_d}(r_1, \ldots, r_n) \end{bmatrix}$$

is a basis of $\Pi^n/\langle\ p\ \rangle$ if $|\mathsf{V}(\langle\ p\ \rangle)| = |\mathsf{V}(\langle\ r\ \rangle)| \cdot |\mathsf{V}(\langle\ q\ \rangle)|$. Hence $\{R_1, \ldots, R_{r_d}\}$ is a transversal of $\Pi^n(y)/\langle\ q\ \rangle$ in $\Pi^n/\langle\ p\ \rangle$. □

This renders some of the ideals obeying the decomposition property (3.12) useless for their application with the decomposition Theorem 3.1 for Fourier transforms. The next section shows that this property is useful nonetheless as there is another decomposition theorem for the Fourier transform.

3.2 Fast Fourier transform via decomposition

As we have seen in the previous section in the multivariate case the decomposition property (3.12) does, unlike in the univariate case, not always yield the existence of an induction. But we still can decompose a module whose quotient ideal obeys the decomposition property step-wise. Instead of the diagram (3.2) the following commutative diagram explains the rationale behind the FFT algorithms based solely on the decomposition property in case of regular modules

$$\begin{tikzcd}
\Pi^n/\langle\, p \,\rangle \arrow[rr, "\text{change of basis}"] \arrow[dd, "\text{Fourier}"] & & \Pi^n/\langle\, q(r) \,\rangle \arrow[d, "\text{Fourier}"] \\
 & & \bigoplus_{\alpha\in\mathsf{V}(\langle\, q \,\rangle)} \Pi^n/\langle\, r-\alpha \,\rangle \arrow[d, "\text{Fourier}"] \\
\bigoplus_{\gamma\in\mathsf{V}(\langle\, p \,\rangle)} \Pi^n/\langle\, x-\gamma \,\rangle & & \bigoplus_{\alpha\in\mathsf{V}(\langle\, q \,\rangle)} \bigoplus_{\beta\in\mathsf{V}(\langle\, r-\alpha \,\rangle)} \Pi^n/\langle\, x-\beta \,\rangle \arrow[ll, "\text{permute}"']
\end{tikzcd} \tag{3.13}$$

The following theorem is a generalized version of [96, Thm. 3] for the situation where one does not assume that the sizes of the varieties of the decomposed ideals multiply to the size of the original variety. The proof is essentially diagram (3.13), but the formulation is a bit technical. Consider $\mathscr{A} = \Pi^n/\langle\, p \,\rangle$ such that the ideal obeys the decomposition property, i.e. $\langle\, p \,\rangle = \langle\, q(r) \,\rangle$, and consider the signal model with regular module $N = \mathscr{A}$. Let $k = |\mathsf{V}(\langle\, q \,\rangle)|$. Denote by $M_\alpha = \Pi^n/\langle\, r-\alpha \,\rangle$ for $\alpha \in \mathsf{V}(\langle\, q \,\rangle)$. Denote by $d_i = \dim M_{\alpha_i}$ the dimension of the submodule to α_i ordered with respect to size, i.e. $d_1 > \cdots > d_k$. In N one can choose, using the decomposition property, a basis of the form

$$\left\{\begin{matrix}
t_1u_1(r(x)) & \dots & t_{d_k}u_1(r(x)) & t_{d_k+1}u_1(r(x)) & \dots & t_{d_1}u_1(r(x)), \\
t_1u_2(r(x)) & \dots & t_{d_k}u_2(r(x)) & \dots & t_{d_2}u_2(r(x)), & \\
\vdots & & \vdots & & & \\
t_1u_k(r(x)) & \dots & t_{d_k}u_k(r(x)) & & &
\end{matrix}\right\}, \tag{3.14}$$

where the $u_1, \dots, u_k$ form a basis of $\Pi^n/\langle\, q \,\rangle$. The elements $t_iu_\ell(r(x))$ with $i > d_j$ are then called the *excess elements* of M_{α_j}. Now we can formulate the theorem, which generalizes [96, Thm. 3] if the subalgebras are not all of the same dimension.

Theorem 3.13 (FFT algorithms via decomposition) *Let $\mathscr{A} = \Pi^n/\langle\, p \,\rangle$ such that $\langle\, p \,\rangle = \langle\, q(r) \,\rangle$ and consider the signal model with regular module $N = \mathscr{A}$. Let $k = |\mathsf{V}(\langle\, q \,\rangle)|$. Denote by $M_\alpha = \Pi^n/\langle\, r-\alpha \,\rangle$ for $\alpha \in \mathsf{V}(\langle\, q \,\rangle)$. Denote for*

$i = 1, \ldots, k$ by $d_i = \dim M_{\alpha_i}$, ordered with respect to size. The Fourier transform of N with respect to a basis b can then be decomposed as

$$P_{b,N} = P \cdot \left(\bigoplus_i P_{M_{\alpha_i}} \right) \cdot T \cdot B, \tag{3.15}$$

where P is permutation matrix, $P_{M_{\alpha_i}}$ are the Fourier transforms of each M_{α_i}, B is a change of basis between bases of N. Denote by $(c_{i,j})$ the entries of the Fourier transform matrix of $\Pi^n/\langle\, q \,\rangle$ with respect to the basis $\{u_1, \ldots, u_k\}$ of (3.14). The matrix T is a block matrix of the form

$$\begin{bmatrix} c_{i,j}\mathbb{1}_{\min(d_i,d_j)} & D_{d_i,d_j} \\ 0^\top_{d_j,d_i} & 0 \end{bmatrix}_{i,j=1,\ldots,k}, \tag{3.16}$$

where $0_{d_i,d_j}$ is the (possibly empty) $d_i \times \max(0, d_j - d_i)$ zero matrix and D_{d_i,d_j} is a (possibly empty) $d_i \times \max(0, d_j - d_i)$ matrix containing the decomposition of the excess basis elements.

Proof: Consider the diagram (3.13). The change of basis is from b to the basis 3.14. The isomorphism $\Pi^n/\langle\, q(r) \,\rangle \longrightarrow \bigoplus_i M_{\alpha_i}$ is, using that basis, realized by T. By the decomposition property the zeros of the M_{α_i} are equal to the zeros of N, except possibly in a different ordering. The theorem follows. □

If all M_α are of equal dimension the matrix T is just equal to the tensor product of the Fourier transform of $\Pi^n/\langle\, q \,\rangle$ with $\mathbb{1}_k$. Reasoning analogously as for Prop. 3.2 one gets again a fast algorithm if the change of basis is sparse and one has a descending chains of submodules with the decomposition property.

The first example shows that one obtains indeed different algorithms from Theorem 3.1 and from Theorem 3.13, cf. [89, Example 7.4].

Example 3.14 The monomials obey the decomposition property since $x^n = (x^m)^r$ if $n = m \cdot r$. For the module $M = \mathbb{C}[x]/\langle\, x^4 - 1 \,\rangle$ with basis $\{1, x, x^2, x^3\}$ to change to the representation $\mathbb{C}[x]/\langle\, (x^2)^2 - 1 \,\rangle$ one needs no change of basis, i.e. $B = \mathbb{1}$, since the basis is already of the desired form. Since in the one variable case by the fundamental theorem of algebra every $x^2 - \alpha$ for any $\alpha \in \mathbb{C}$ has the same number of zeros the matrix T from (3.15) is given as the tensor product of the Fourier transform of $\mathbb{C}[x]/\langle\, x^2 - 1 \,\rangle$ with the 2×2 identity matrix, i.e.

$$T = \mathsf{DFT}_2 \otimes \mathbb{1}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}. \tag{3.17}$$

This decomposes $\mathbb{C}[x]/\langle\, (x^2)^2 - 1 \,\rangle$ into $\mathbb{C}[x]/\langle\, x^2 - 1 \,\rangle \oplus \mathbb{C}[x]/\langle\, x^2 + 1 \,\rangle$. The first submodule can be decomposed with the standard discrete Fourier transform of size 2 while the second submodule has as Fourier transform $\begin{bmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{bmatrix}$, since $x^2+1 = (x-\mathrm{i})(x+\mathrm{i})$. Hence one obtains as decomposition for the direct sum

$$\begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & \mathrm{i} \\ & & 1 & -\mathrm{i} \end{bmatrix}. \tag{3.18}$$

For the permutation observe that for the discrete Fourier transform of size 4 the zeros are ordered as $\{1, -\mathrm{i}, -1, \mathrm{i}\}$ while until now we obtained the zeros ordered as $\{1, -1, -\mathrm{i}, \mathrm{i}\}$. Hence we finally need to multiply with the permutation matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.19}$$

Thus the complete decomposition (3.15) reads in this case

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & \mathrm{i} \\ & & 1 & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}. \tag{3.20}$$

Comparing (3.20) with (3.6) one observes that we indeed get two different sparse factorizations leading to two different fast algorithms.

Unfortunately the decomposition property is a rare one. Indeed in one variable the only families which have this property are, up to affine-linear coordinate changes, the monomials x^k and the Chebyshev polynomials of the first kind T_k, cf. [98, Ch. 4]. In Chapter 4 we will investigate a larger class of families with the decomposition property generalizing the univariate Chebyshev polynomials.

As a first multivariate example we propose a fast algorithm for a special case of the Koornwinder polynomial signal models from Example 2.16.

Example 3.15 Consider a family of orthogonal polynomials $\{p_n\}$ which obey the decomposition property, i.e. $p_n(x) = p_m(p_r(x))$ if $n = m \cdot r$. By [98, Ch. 4] we can assume that these orthogonal polynomials are the Chebyshev polynomials of the first kind. The non-normalized Koornwinder polynomials to the value $-\frac{1}{2}$, i.e. $P_k^n = P_k^{n,-1/2}$ if $k < n$ and $P_n^n(u,v) = p_n(x) \cdot p_n(y)$, with $u = x + y$ and $v = x \cdot y$, obey the decomposition property since one has

$$P_0^r(u,v) = p_r(x) + p_r(y)$$

and

$$P_r^r(u,v) = p_r(x) \cdot p_r(y)$$

and thus it follows that for $n = m \cdot r$ one has

$$\begin{aligned} P_k^m(P_0^r(u,v), P_r^r(u,v)) &= p_m(p_r(x))p_k(p_r(x)) + p_m(p_r(y))p_k(p_r(x)) \\ &= p_n(x)p_{k\cdot r}(x) + p_n(y)p_{k\cdot r}(x) \\ &= P_{k\cdot r}^n(u,v) \end{aligned}$$

if $k < n$ and

$$\begin{aligned} P_m^m(P_0^r(u,v), P_r^r(u,v)) &= p_m(p_r(x)) \cdot p_m(p_r(y)) \\ &= p_n(x) \cdot p_n(y) \\ &= P_n^n(u,v), \end{aligned}$$

if $k = n$. We want to give an example as explicit as possible, hence we investigate the signal model with $\mathscr{A} = \mathbb{R}[u,v]/\langle\, \mathbb{P}_4 \,\rangle$, $M = \mathscr{A}$, and

$$\Phi(s) = \sum_{k \leq n < 4} s_{k,n} P_k^n. \tag{3.21}$$

The polynomials of the basis of the module are explicitly given through the following list

$$\begin{aligned} P_0^0 &= 1, \\ P_0^1 &= u, \\ P_1^1 &= v, \\ P_0^2 &= 2u^2 - 4v - 2, \\ P_1^2 &= 2uv - u, \\ P_2^2 &= 4v^2 - 2u^2 + 4v + 1, \\ P_0^3 &= 4u^3 - 12uv + 3u, \\ P_1^3 &= 4u^2v - 8v^2 - 6v, \\ P_2^3 &= 8uv^2 - 4u^3 + 6uv + 3u, \\ P_3^3 &= 16v^3 - 12vu^2 + 24v^2 + 9v. \end{aligned} \tag{3.22}$$

For using the Chinese remainder theorem B.3 with the decomposition property, and thus for application of Theorem 3.13 for the decomposition of the corresponding Fourier transform, we do use a basis change to the basis

$$[P_0^0, P_0^1, P_1^1, P_1^2, P_0^2, P_0^1P_0^2, P_1^1P_0^2, P_2^2, P_0^1P_2^2, P_1^1P_2^2]. \tag{3.23}$$

Observing that one has

$$\begin{aligned} P_0^3 &= 2P_0^1P_0^2 - 2P_1^2 + 5P_0^1, \\ P_1^3 &= 2P_1^1P_0^2 - 2P_1^1, \\ P_2^3 &= 2P_0^1P_2^2 - P_1^2, \\ P_3^3 &= 4P_1^1P_2^2 - 2P_1^1P_0^2 + P_1^1, \end{aligned} \tag{3.24}$$

and the other elements being in the old as well as in the new basis, one obtains the sparse change of basis

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}. \tag{3.25}$$

One observes that $\mathsf{V}(\langle\ \mathbb{P}_4\ \rangle) = \mathsf{V}(\langle\ P_0^4, P_4^4\ \rangle)$ decomposes into three subvarieties using the decomposition

$$\begin{aligned} P_0^4 &= P_0^2(P_0^2, P_2^2), \\ P_4^4 &= P_2^2(P_0^2, P_2^2), \end{aligned}$$

one consisting of four points and two consisting of three points. Indeed one has

$$\mathsf{V}(\langle\ \mathbb{P}_n\ \rangle) = \big\{\big(\cos\big(\tfrac{k+\frac{1}{2}}{2}\pi\big) + \cos\big(\tfrac{j+\frac{1}{2}}{2}\pi\big),\ \cos\big(\tfrac{k+\frac{1}{2}}{2}\pi\big)\cdot\cos\big(\tfrac{j+\frac{1}{2}}{2}\pi\big) \mid 0 \le j \le k < n\big\}, \tag{3.26}$$

cf. Example 2.16, since the zeros of the Chebyshev polynomials of the first kind are given by $\cos(\frac{k+1/2}{2}\pi)$. Hence the Fourier transform of this signal model is given as

$$\mathcal{F}_4 = (P_i^k(\alpha))_{0\le i\le k\le 3, \alpha\in\mathsf{V}(\langle\ \mathbb{P}_4\ \rangle)}. \tag{3.27}$$

We are interested in the varieties corresponding to the partial decomposition

$$\langle\ P_0^4, P_4^4\ \rangle = \bigcap_{0\le j\le k<2} \langle\ P_0^2 - \cos(\tfrac{k+\frac{1}{2}}{2}\pi) - \cos(\tfrac{j+\frac{1}{2}}{2}\pi), P_2^2 - \cos(\tfrac{k+\frac{1}{2}}{2}\pi)\cdot\cos(\tfrac{j+\frac{1}{2}}{2}\pi)\ \rangle \tag{3.28}$$

One obtains in (x, y)-coordinates

$$\begin{aligned} &\mathsf{V}(\langle\ P_0^2 - \cos(\tfrac{\pi}{4}) - \cos(\tfrac{\pi}{4}), P_2^2 - \cos(\tfrac{\pi}{4})\cdot\cos(\tfrac{\pi}{4}\pi)\ \rangle) \\ &\quad = \{(\cos\tfrac{\pi}{8}, \cos\tfrac{\pi}{8}), (\cos\tfrac{7\pi}{8}, \cos\tfrac{\pi}{8}), (\cos\tfrac{7\pi}{8}, \cos\tfrac{7\pi}{8})\}, \\ \\ &\mathsf{V}(\langle\ P_0^2 - \cos(\tfrac{3\pi}{4}) - \cos(\tfrac{\pi}{4}), P_2^2 - \cos(\tfrac{3\pi}{4})\cdot\cos(\tfrac{\pi}{4}\pi)\ \rangle) \\ &\quad = \{(\cos\tfrac{3\pi}{8}, \cos\tfrac{\pi}{8}), (\cos\tfrac{5\pi}{8}, \cos\tfrac{\pi}{8}), (\cos\tfrac{7\pi}{8}, \cos\tfrac{3\pi}{8}), (\cos\tfrac{7\pi}{8}, \cos\tfrac{5\pi}{8})\}, \\ \\ &\mathsf{V}(\langle\ P_0^2 - \cos(\tfrac{3\pi}{4}) - \cos(\tfrac{3\pi}{4}), P_2^2 - \cos(\tfrac{3\pi}{4})\cdot\cos(\tfrac{3\pi}{4}\pi)\ \rangle) \\ &\quad = \{(\cos\tfrac{3\pi}{8}, \cos\tfrac{3\pi}{8}), (\cos\tfrac{5\pi}{8}, \cos\tfrac{3\pi}{8}), (\cos\tfrac{5\pi}{8}, \cos\tfrac{5\pi}{8})\}. \end{aligned} \tag{3.29}$$

The Fourier transform of $M = \mathbb{R}[u,v]/\langle\, \mathbb{P}_2 \,\rangle$ with basis $[P_0^0, P_0^1, P_1^1]$ is

$$
\begin{aligned}
\mathcal{F}_2 &= \begin{bmatrix} 1 & \cos(\frac{\pi}{4}) + \cos(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \cdot \cos(\frac{\pi}{4}) \\ 1 & \cos(\frac{3\pi}{4}) + \cos(\frac{\pi}{4}) & \cos(\frac{3\pi}{4}) \cdot \cos(\frac{\pi}{4}) \\ 1 & \cos(\frac{3\pi}{4}) + \cos(\frac{3\pi}{4}) & \cos(\frac{3\pi}{4}) \cdot \cos(\frac{3\pi}{4}) \end{bmatrix} \\
&= \begin{bmatrix} 1 & \sqrt{2} & \frac{1}{2} \\ 1 & 0 & -\frac{1}{2} \\ 1 & -\sqrt{2} & \frac{1}{2} \end{bmatrix}.
\end{aligned}
\tag{3.30}
$$

For the partial decomposition observe that all elements of the new basis are of either of the forms $P_i^j P_0^0, P_i^j P_0^2, P_i^j P_2^2$, but the element P_1^2 is an excess element for the varieties consisting of three points. As matrix T from (3.15) one obtains in this case

$$
\left[\begin{array}{cccc|ccc|ccc}
1 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 1 & 0 & \frac{1}{\sqrt{2}} & 0 & \sqrt{2} & 0 & 0 & \frac{1}{2} & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & \frac{1}{2} \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
1 & 0 & 0 & 0 & -\sqrt{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 1 & 0 & \frac{-1}{\sqrt{2}} & 0 & -\sqrt{2} & 0 & 0 & \frac{1}{2} & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -\sqrt{2} & 0 & 0 & \frac{1}{2}
\end{array}\right]
\tag{3.31}
$$

Note that the 4×4 blocks appear in the second block row instead of the first, as one would expect from Theorem 3.13, since from (3.30) one obtains the sizes of the subvarieties (3.29) in order $(3, 4, 3)$. As basis in each of the three dimensional spaces one obtains from the partial decomposition (3.31) the elements $[P_0^0, P_0^1, P_1^1]$, while for the four dimensional space one obtains $[P_0^0, P_0^1, P_1^1, P_1^2]$. This results in the direct sum

$$\begin{bmatrix} 1 & 2\cos\frac{\pi}{8} & \cos^2\frac{\pi}{8} \\ 1 & 0 & -\cos^2\frac{\pi}{8} \\ 1 & -2\cos\frac{\pi}{8} & \cos^2\frac{\pi}{8} \end{bmatrix}$$
$$\bigoplus$$
$$\begin{bmatrix} 1 & \cos\frac{\pi}{8}+\sin\frac{\pi}{8} & \cos\frac{\pi}{8}\sin\frac{\pi}{8} & -\cos\frac{\pi}{8}-\sin\frac{\pi}{8}+2\cos\frac{\pi}{8}\sin\frac{\pi}{8}\left(\cos\frac{\pi}{8}+\sin\frac{\pi}{8}\right) \\ 1 & \cos\frac{\pi}{8}-\sin\frac{\pi}{8} & -\cos\frac{\pi}{8}\sin\frac{\pi}{8} & -\cos\frac{\pi}{8}+\sin\frac{\pi}{8}-2\cos\frac{\pi}{8}\sin\frac{\pi}{8}\left(\cos\frac{\pi}{8}-\sin\frac{\pi}{8}\right) \\ 1 & -\cos\frac{\pi}{8}-\sin\frac{\pi}{8} & -\cos\frac{\pi}{8}\sin\frac{\pi}{8} & \cos\frac{\pi}{8}-\sin\frac{\pi}{8}-2\cos\frac{\pi}{8}\sin\frac{\pi}{8}\left(-\cos\frac{\pi}{8}+\sin\frac{\pi}{8}\right) \\ 1 & -\cos\frac{\pi}{8}-\sin\frac{\pi}{8} & \cos\frac{\pi}{8}\sin\frac{\pi}{8} & \cos\frac{\pi}{8}+\sin\frac{\pi}{8}+2\cos\frac{\pi}{8}\sin\frac{\pi}{8}\left(-\cos\frac{\pi}{8}-\sin\frac{\pi}{8}\right) \end{bmatrix}$$
$$\bigoplus$$
$$\begin{bmatrix} 1 & 2\sin\frac{\pi}{8} & \sin^2\frac{\pi}{8} \\ 1 & 0 & -\sin^2\frac{\pi}{8} \\ 1 & -2\sin\frac{\pi}{8} & \sin^2\frac{\pi}{8} \end{bmatrix}. \tag{3.32}$$

The last thing we need for the decomposition 3.15 is a permutation matrix. Going from (3.29) to (3.26), with $n = 4$, is done using the permutation

$$P = \begin{bmatrix} 1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0 \end{bmatrix}. \tag{3.33}$$

Hence the overall computational cost of this decomposition is as follows. From the change of basis (3.25) one has 6 (real) additions and 9 (real) multiplications, from the partial decomposition (3.31) one gets 17 additions and 17 multiplications, and for the direct sum of the smaller transforms (3.32) one obtains 22 additions and 22 multiplications. This results in an overall of 45 additions and 48 multiplications, while the naive matrix-vector product approach requires 78 additions and 88 multiplications. This results in saving of approximately 44 % of overall operations.

Chapter 4

Multivariate Chebyshev polynomials and generalized cosine transforms

The previous chapters suggest that nice examples of multivariate polynomials as basis for algebraic signal models should possess two properties. The first one is the decomposition property (3.12), so one can either use Theorem 3.1 or Theorem 3.13 to obtain a fast algorithm for the computation of their Fourier transforms. The second one is that the number of common zeros of $\mathbb{P}_n$ should be equal to $\dim \Pi^d_{n-1}$ so that one can use Theorem 2.14 to obtain an orthogonal Fourier transform. Unfortunately both properties occur rarely and the appearance of both properties at once is even more rare.

In this chapter we study generalizations of the Chebyshev polynomials, used for the derivation of the discrete space signal model in Example 2.7, to multivariate Chebyshev polynomials. The construction of these polynomials is based on a geometric stretching and folding property which was introduced by Hoffman and Withers [36]. The generalizations of the first kind Chebyshev polynomials turn out to be subject to the decomposition property while the multivariate analogues of the second kind Chebyshev polynomials have, by adjusting the grading, enough common zeros, cf. [76], so one can apply the multivariate Gauß-Jacobi-procedure of Theorem 2.14. After the definition and study of some of the properties of the multivariate Chebyshev polynomials in Sect. 4.1, some of the corresponding fast transforms are studied in detail in Sect. 4.2. The focus on these special cases is motivated as follows. The first example is that corresponding to the root system of type A_2. There one obtains a hexagonal space signal model and was first motivated in [95] as cosine transforms on hexagonally sampled images. The fast algorithm for its computation was first obtained in [96] using the decomposition property and a variant of Theorem 3.13. The hexagonal lattice is the optimal regular sampling lattice in two dimensions requiring 13.4 % less sampling points to obtain the same information as a regular rectangular sampling lattice. Similar we considered the signal model corresponding to the root system of A_3 in [111] as the signal model corresponds to a body-centered cubic (BCC) lattice (though we wrongly identified the lattice of the signal model with the reciprocal lattice in that paper). The BCC lattice is the optimal regular lattice for

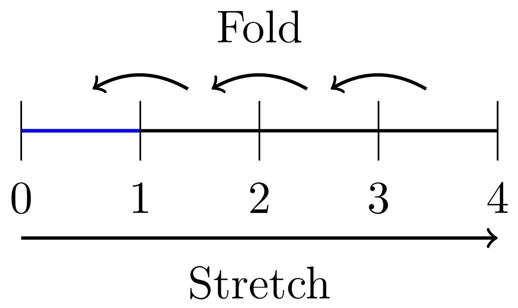

Figure 4.1: The stretching and folding property of the univariate Chebyshev polynommials of the first kind.

sampling band-limited signals above the Nyquist limit, while below the Nyquist limit it is the face-centered cubic (FCC) lattice, the reciprocal lattice of the BCC lattice, cf. [73, 85, 118]. The BCC lattice requires 29.3 % less sampling points to obtain the same information as a regular rectangular samplig lattice. Furthermore alias errors are reduced best [118]. See also the book of Conway and Sloane [12] for more details on lattices and their properties.

The last example, the signal model associated to Chebyshev polynomials to the root system C_2, was motivated by us in [110] as it gives a cosine transform for signals sampled on a lattice of triangles. Furthermore in this special case we have enough common zeros to apply the Gauß-Jacobi procedure to obtain an orthogonal transform.

As one can observe that the construction of the multivariate Chebyshev polynomials relies on one-dimensional representations of Weyl groups, it is natural to ask what happens if one uses representations that are of higher dimension. Basically the same construction, with some technical adjustments, lead to matrix-valued Chebyshev polynomials associated to these representations as was observed by Huybrechs and Munthe-Kaas [44]. In Sect. 4.3 we generalize a method for the derivation of generating functions of multivariate Chebyshev polynomials introduced by Damaskinsky, Kulish and Sokolov [18] to the matrix-valued case. Furthermore using extensive computer algebra some examples are calculated. Since in general the existence of the matrix-valued polynomials is not proven, this is a strong indicator towards their existence.

4.1 Multivariate Chebyshev polynomials associated to root systems

The general construction of multivariate Chebyshev polynomials of the first kind by Hoffman and Withers [36] relies on the following geometric interpretation of the decomposition property. The map

$$\cos^{-1} \circ T_n \circ \cos \tag{4.1}$$

stretches the interval $[0, 1]$ and folds it back at the integers, cf. 4.1. This is called a stretching and folding operation. Hence a natural generalization of Chebyshev polynomials should be associated to some region which can be stretched and then folded back to itself. In [36] it was shown that the foldable figures in higher dimensions are in one-to-one correspondence to the Weyl groups of root systems.

For this generalization we thus need to recall the notions of Weyl groups and root systems, both stemming from Lie theory. Indeed the root systems were introduced by Killing [47] for the classification of the complex, semi-simple Lie algebras. See [43] for a thorough investigation of root systems and associated groups as well as classification proofs and references to the literature.

Definition 4.1 *A crystallographic root system in a finite-dimensional euclidean space* $(\mathbb{R}^d, \langle \cdot\, , \cdot \rangle)$ *is a finite set* R *of non-zero vectors, the so-called roots, which span* $\mathbb{R}^d$ *subject to the conditions*

i.) $r \cdot \alpha \in R$ *then* $r = \pm 1$ *for all* $\alpha \in R$*,*

ii.) closedness under reflections through the hyperplanes perpendicular to the roots, i.e.,

$$\sigma_\alpha(\beta) = \beta - 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha \in R \tag{4.2}$$

for all $\alpha, \beta \in R$*,*

iii.) for any $\alpha, \beta \in R$ *we have* $2\frac{\langle \alpha,\beta \rangle}{\langle \alpha,\alpha \rangle} \in \mathbb{Z}$*.*

The set of integer linear combinations of the roots is termed the root lattice

$$Q = \operatorname{span}_{\mathbb{Z}} R \subseteq \mathbb{R}^d \tag{4.3}$$

of the root system. The coroot of a root $\alpha \in R$ *is*

$$\alpha^\vee = \frac{2}{\langle \alpha, \alpha \rangle}\alpha. \tag{4.4}$$

The coroots form a root system which is denoted by $R^\vee$*. The coroot lattice* $Q^\vee$ *is the* $\mathbb{Z}$*-span of the coroots.*

There are at most two different root lengths (root lengths are defined as lengths as vectors) for an irreducible root system, i.e., one which is not a combination of root systems with mutually orthogonal spaces. The irreducible root systems can be classified using Coxeter-Dynkin diagrams. There are four infinite series A_n, B_n, C_n, D_n, cf. Fig. 4.2, and five exceptional root systems E_6, E_7, E_8, F_4, G_2, cf. Fig. 4.3. The four infinite series correspond to the special unitary, special orthogonal and symplectic Lie algebras as summarized in Table 4.1. One can choose a basis $\Delta = \{\alpha_1, \dots, \alpha_d\} \subseteq R$ of the root system such that one has $\alpha = \sum_{i=1}^d c_j \alpha_j$ with all $c_j \in \mathbb{Z}$ of the same sign. The α_i are called simple roots. The simple roots divide the root system into positive roots R^+ and negative roots R^-. The simple roots introduce a partial order on the roots, as well. The partial order is defined by $\lambda \succeq \mu$ if the expansion of $\lambda - \mu$ in simple roots has non-negative coefficients only. Then λ is called higher than μ. The highest root has the form

$$\alpha_0 = m_1\alpha_1 + \dots + m_d\alpha_d, \tag{4.5}$$

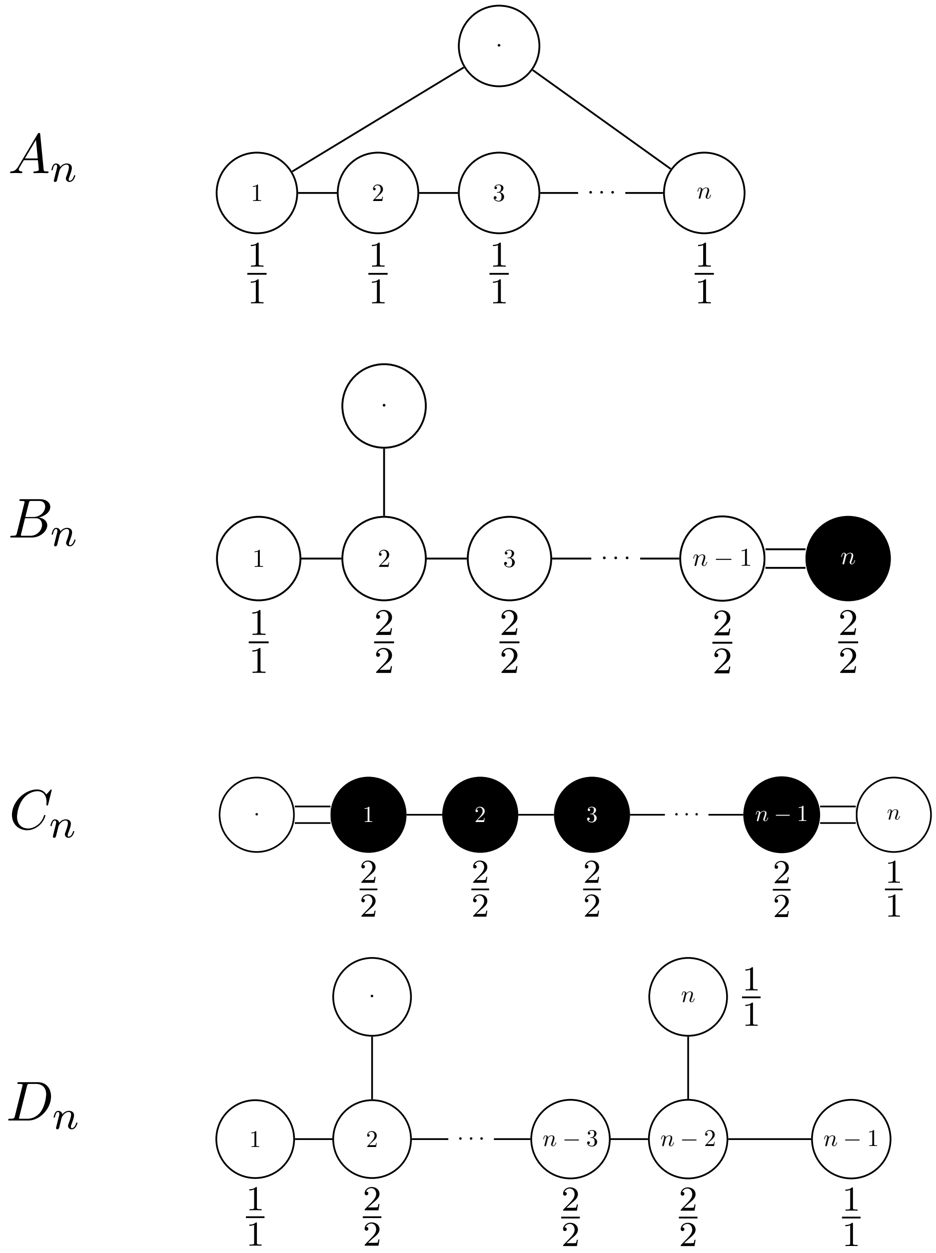

Figure 4.2: Affine Coxeter-Dynkin diagrams for the reduced, crystallographic root systems. The dotted node corresponds to the lowest root $-\alpha_0$, the numbered nodes to the simple roots α_i. Open circles are long roots, while filled nodes indicate short roots. The marks and comarks are shown below the nodes as $\frac{m_i}{m_i^\vee}$. The angle between two roots depends on the multiplicity k of the edge between them and is given as $4\cos^2\theta = k$ and $\cos\theta \leq 0$, i.e., $\frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}$ with length ratio being arbitrary, $1, \sqrt{2}, \sqrt{3}$ for $k = 0, 1, 2, 3$, respectively. This figure shows the four infinite series, where n starts at 1 for A_n, at 2 for B_n, at 3 for C_n, and at 4 for D_n.

root system	compact Lie algebra
A_n	$\mathfrak{su}_{n+1}$
B_n	$\mathfrak{so}_{2n+1}$
C_n	$\mathfrak{sp}_n$
D_n	$\mathfrak{so}_{2n}$

Table 4.1: Root systems and corresponding compact Lie algebras.

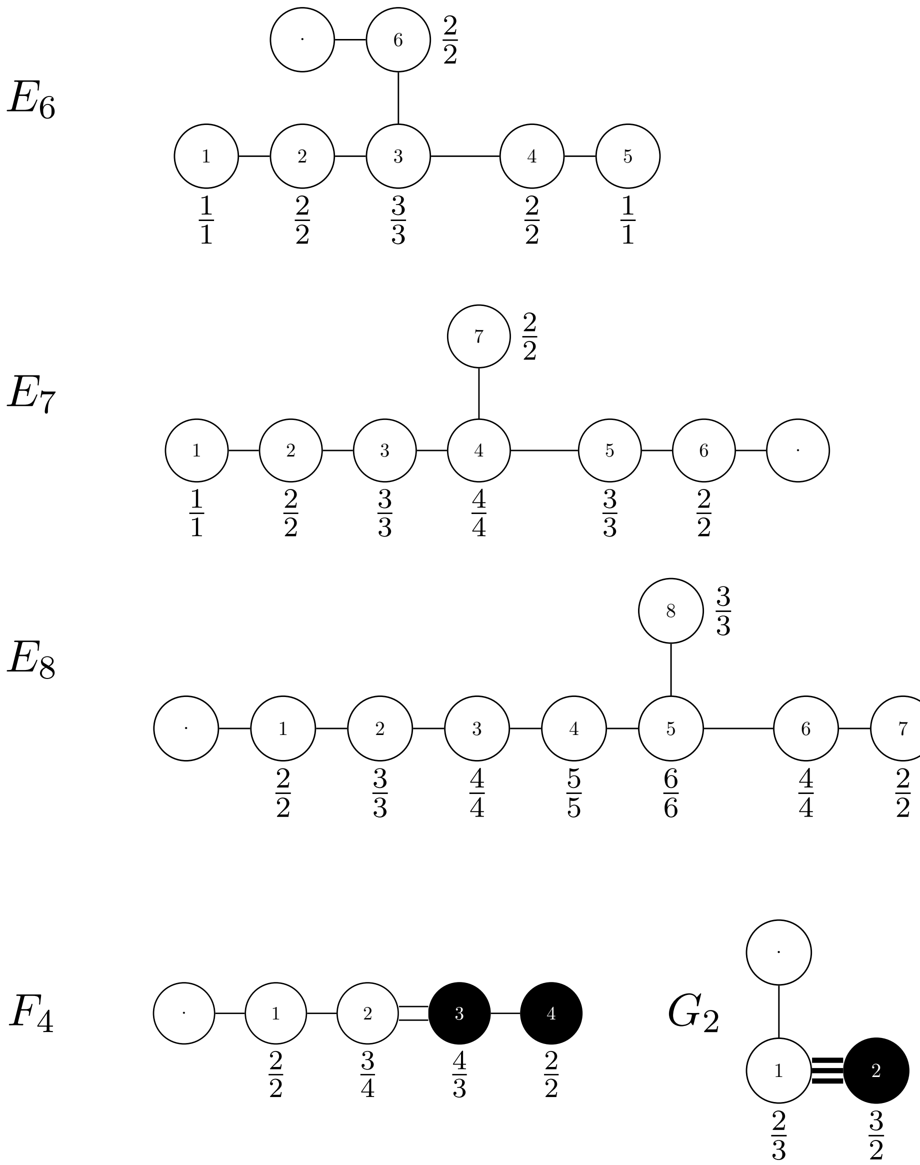

Figure 4.3: The five exceptional affine Coxeter-Dynkin diagrams. Notation as in Fig. 4.2.

with positive integers m_i. The m_i are called the marks of the root system. The marks of the coroot system are called the comarks of the initial root system and denoted by $m_1^\vee, \dots, m_d^\vee$.

The Weyl group of a root system R is the group generated by the reflections

$$W = \langle\, \sigma_\alpha \mid \alpha \in R \,\rangle. \tag{4.6}$$

The $\mathbb{Z}$-dual of Q is the coweight lattice $P^\vee$, while the $\mathbb{Z}$-dual of $Q^\vee$ is the weight lattice P. The generators of P are the fundamental weights ω_j and the generators of $P^\vee$ are the fundamental coweights $\omega_j^\vee$. Between these lattices the following relations hold

$$\begin{tikzcd}
\{\alpha_1,\dots,\alpha_d\} & \subseteq & Q \arrow[d, hook] \arrow[drr, leftrightarrow] & & Q^\vee \arrow[d, hook] & \supseteq & \{\alpha_1^\vee,\dots,\alpha_d^\vee\} \\
\{\omega_1,\dots,\omega_d\} & \subseteq & P \arrow[urr, leftrightarrow] & & P^\vee & \supseteq & \{\omega_1^\vee,\dots,\omega_d^\vee\}
\end{tikzcd} . \tag{4.7}$$

The coroot lattice acts on $\mathbb{R}^d$ by translation and the affine Weyl group is the semi-direct product

$$W_{\text{aff}} = W \ltimes Q^\vee. \tag{4.8}$$

The simplex $F = \mathbb{R}^d/Q^\vee$ tiles $\mathbb{R}^d$ under the action of the affine Weyl group and is called the fundamental Weyl chamber. One can describe the fundamental Weyl chamber as the convex hull

$$F = \mathsf{conv}\left\{0, \frac{\omega_1^\vee}{m_1}, \dots, \frac{\omega_d^\vee}{m_d}\right\} \tag{4.9}$$

of the fundamental coweights scaled by the marks. This fundamental region replaces the interval $[0,1]$ as stretching and folding region for multivariate Chebyshev polynomials. In Fig. 4.4 the root systems of type A_2 and C_2 are shown together with the simple scaled coweights and the fundamental domains. The dual pairing $(\,\cdot\,,\,\cdot\,)\colon P \times \mathbb{R}^d/Q^\vee \longrightarrow \mathbb{C}$ is given by

$$(\lambda, \theta) = \exp(2\pi \mathrm{i} \langle \lambda, \theta \rangle). \tag{4.10}$$

The Weyl group, which is isomorphic to a group of integer matrices, acts on P and $\mathbb{R}^d/Q^\vee$. Symmetrization of the dual pairing with respect to the corresponding Weyl group now leads to the definition of multivariate Chebyshev polynomials of the first kind [36].

Definition 4.2 *Let W be a Weyl group of a root system R with weight lattice P and coroot lattice $Q^\vee$. The multivariate Chebyshev polynomial of the first kind of weight $\lambda \in P$ is*

$$T_\lambda(x_1, \dots, x_d) = T_\lambda(\theta) = \frac{1}{|W|} \sum_{w \in W} (\lambda, w\theta) \tag{4.11}$$

for $\theta \in F$. The multivariate Chebyshev polynomials are polynomials in the variables

$$x_k = T_{\omega_k}(\theta), \tag{4.12}$$

with $\theta \in F$.

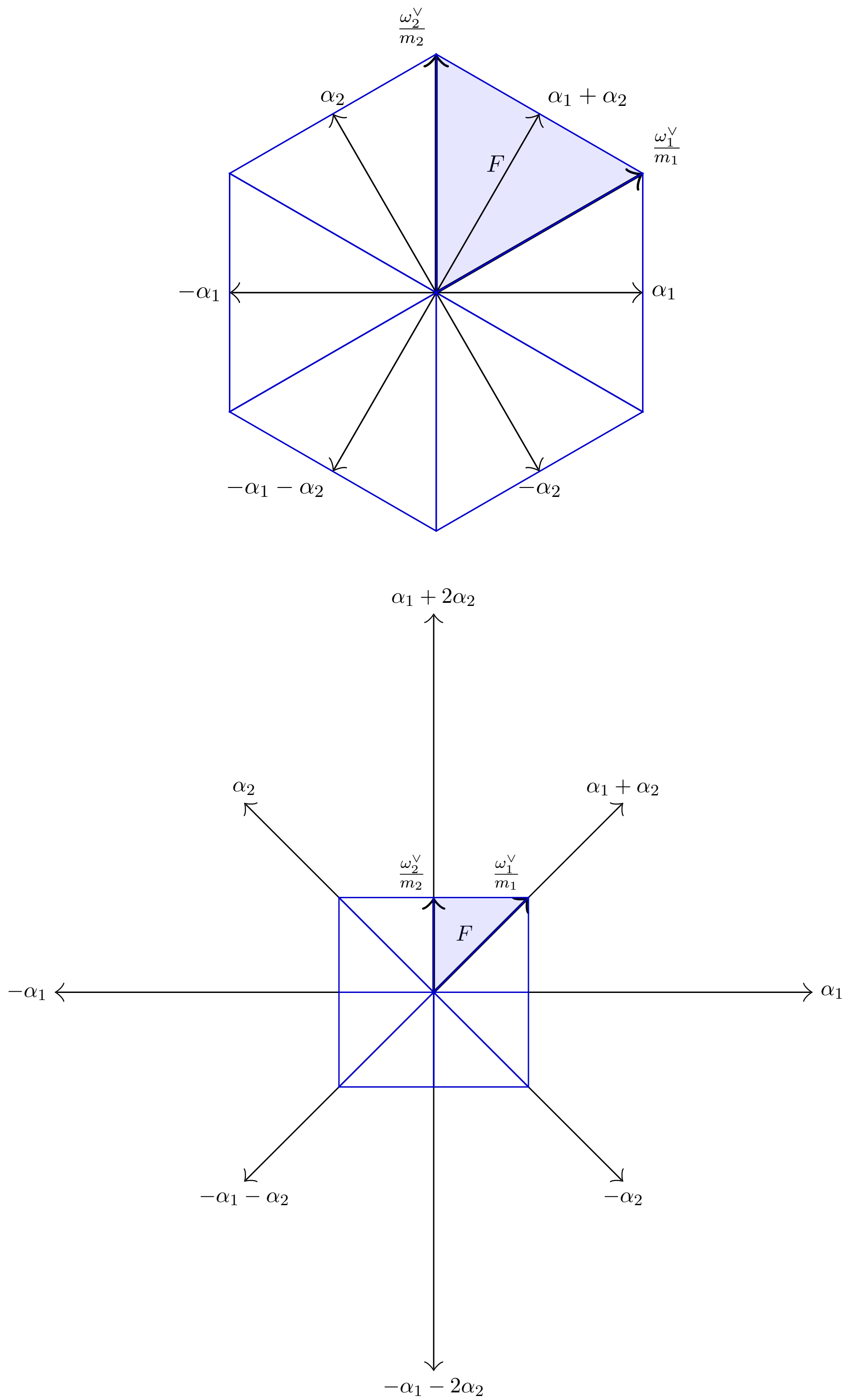

Figure 4.4: The root systems of type A_2 (upper) and C_2 (lower) together with the fundamental region F (shaded region) and the image of F under the action of the Weyl group.

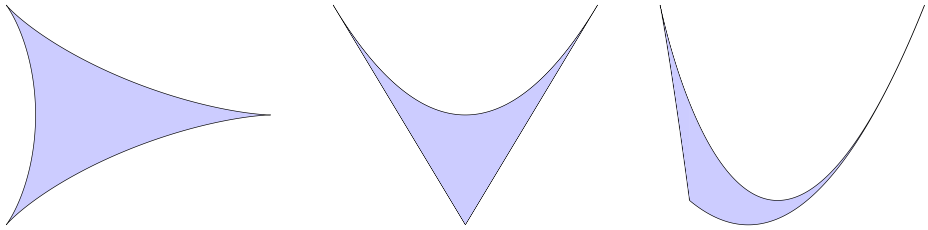

Figure 4.5: The image of the fundamental region F under the variable change in case of A_2, C_2, and G_2, respectively.

That the multivariate Chebyshev polynomials are indeed polynomials follows from a theorem of Chevalley [11], which considers invariants of finite groups generated by reflections.

For the Weyl group $W(A_1)$ of type A_1 one gets back the original definition of Chebyshev polynomials, as the root system is then $R_{A_1} = \{1, -1\}$ and is equal to the coroot lattice, the only simple root is $\{1\}$, the Weyl group is $W(A_1) = \{1, -1\}$, the root and coroot lattice is $\mathbb{Z}$, the weight lattice is $P = \mathbb{Z}$. This leads to $F = [0, 1]$ and $T_n(x) = \frac{1}{2}(\exp(2\pi \mathrm{i} n\theta) + \exp(-2\pi \mathrm{i} n\theta) = \cos(n\theta)$.

From the definition (4.11) one can deduce that the obtained variables are real-valued or consist of pairs of complex-conjugates. That is if we complex-conjugate the x_j in the θ-domain we obtain either the identity or a permutation between the x_j, cf. [76, Sect. 6]. The only cases where one does not obtain real variables are A_n, D_{2n+1}, and E_6. In these cases one obtains the following pairs

$$
\begin{array}{rl}
A_n : & \begin{array}{|c||c|c|c|c|c|} \hline x & x_1 & x_2 & \cdots & x_{n-1} & x_n \\ \hline \overline{x} & x_n & x_{n-1} & \cdots & x_2 & x_1 \\ \hline \end{array} \\
D_{2n+1} : & \begin{array}{|c||c|c|c|c|c|c|} \hline x & x_1 & x_2 & \cdots & x_{2n-1} & x_{2n} & x_{2n+1} \\ \hline \overline{x} & x_1 & x_2 & \cdots & x_{2n-1} & x_{2n+1} & x_{2n} \\ \hline \end{array} \\
E_6 : & \begin{array}{|c||c|c|c|c|c|c|} \hline x & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \hline \overline{x} & x_5 & x_4 & x_3 & x_2 & x_1 & x_6 \\ \hline \end{array} .
\end{array}
\tag{4.13}
$$

The simplex F gets transformed under the variable change $x_k = T_{\omega_k}(\theta)$ to a cusped region. For example in case of the root system A_2 the fundamental region F is an equilateral triangle which gets transformed to a deltoid under (x_1, x_2) (and, to be rigorous, shown in $(\frac{1}{2}(x_1 + x_2), \frac{1}{2\mathrm{i}}(x_1 - x_2))$ coordinates, cf. (4.13)). In Fig. 4.5 the cusped regions for the irreducible two-dimensional root systems are shown. The following proposition contains the properties of the first kind multivariate Chebyshev we need in the sequel, cf. [77, 97].

Proposition 4.3 *The multivariate Chebyshev polynomials of the first kind associated to a Weyl group W are subject to*

i.) invariance with respect to the action of the Weyl group on the weight indices

$$T_{w^\top \lambda} = T_\lambda \tag{4.14}$$

and invariance with respect to the affine Weyl group on the argument in the fundamental domain

$$T_\lambda(T_{\omega_1}(w\theta),\ldots,T_{\omega_d}(w\theta)) = T_\lambda(T_{\omega_1}(\theta),\ldots,T_{\omega_d}(\theta)), \tag{4.15}$$

ii.) the shift property

$$T_{\lambda_1}T_{\lambda_2} = \frac{1}{|W|}\sum_{w\in W} T_{\lambda_1+w^\top\lambda_2}, \tag{4.16}$$

iii.) the decomposition property

$$T_{k\lambda} = T_\lambda(T_{k\omega_1},\ldots,T_{k\omega_d}) \tag{4.17}$$

for $\lambda\in P$ and $k\in\mathbb{Z}$.

Proof: For the first part of *i.)* observe that

$$\begin{aligned}
T_{v^\top\lambda} &= \frac{1}{|W|}\sum_{w\in W}(v^\top\lambda, w\theta)\\
&= \frac{1}{|W|}\sum_{w\in W}(\lambda, vw\theta)\\
&= \frac{1}{|W|}\sum_{u\in W}(\lambda, u\theta)\\
&= T_\lambda,
\end{aligned}$$

with $u=vw$. For the second part one has

$$\begin{aligned}
T_\lambda((\omega_1,v\theta),\ldots,(\omega_d,v\theta)) &= \frac{1}{|W|}\sum_{w\in W}(\lambda, wv\theta)\\
&= \frac{1}{|W|}\sum_{u\in W}(\lambda, u\theta)\\
&= T_\lambda((\omega_1,\theta),\ldots,(\omega_d,\theta)),
\end{aligned}$$

with $u=wv$.

For the part *ii.)* one has

$$\begin{aligned}
T_{\lambda_1}T_{\lambda_2} &= \left(\frac{1}{|W|}\sum_{u\in W}(\lambda_1,u\theta)\right)\left(\frac{1}{|W|}\sum_{v\in W}(\lambda_2,v\theta)\right)\\
&= \left(\frac{1}{|W|}\sum_{u\in W}(\lambda_1,u\theta)\right)\left(\frac{1}{|W|}\sum_{v\in W}(v^\top\lambda_2,\theta)\right)\\
&= \frac{1}{|W|^2}\sum_{u,v\in W}(\lambda_1,u\theta)(v^\top\lambda_2,\theta)\\
&= \frac{1}{|W|^2}\sum_{u,v\in W}(\lambda_1+(vu^{-1})^\top\lambda_2,u\theta)\\
&= \frac{1}{|W|}\sum_{w\in W}\frac{1}{|W|}\sum_{u\in W}(\lambda_1+w^\top\lambda_2,u\theta)\\
&= \frac{1}{|W|}\sum_{w\in W}T_{\lambda_1+w^\top\lambda_2},
\end{aligned}$$

where we can write w instead of vu^{-1} since we are summing over the whole group.

The part *iii.)* follows since one has formally

$$\begin{aligned}T_{k\omega_i}(\theta) &= \tfrac{1}{|W|}\sum_{w\in W}(k\omega_i, w\theta)\\ &= \tfrac{1}{|W|}\sum_{w\in W}(\omega_i, wk\theta)\\ &= T_{\omega_i}(k\theta),\end{aligned}$$

and thus

$$\begin{aligned}T_\lambda(T_{k\omega_1}(\theta),\ldots,T_{k\omega_d}(\theta)) &= T_\lambda(T_{\omega_1}(k\theta),\ldots,T_{\omega_d}(k\theta))\\ &= \tfrac{1}{|W|}\sum_{w\in W}(\lambda, k\theta)\\ &= \tfrac{1}{|W|}\sum_{w\in W}(k\lambda, \theta)\\ &= T_{k\lambda}(\theta).\end{aligned}$$

We have proven the proposition. □

The properties i.) and ii.) of Prop. 4.3 yield a recursion relation if one uses the shift relation with the $x_k = T_{\omega_k}$.

Since the multivariate Chebyshev polynomials are subject to the decomposition property they give rise to examples of signal models with a fast Fourier transform algorithm via Theorem 3.13. In certain cases one can even get an induction of modules and thus can apply Theorem 3.1 to get a fast Fourier transform algorithm. We will investigate some examples in Sect. 4.2. Unfortunately one does in general not get orthogonal Fourier transforms by the multivariate Gauß-Jacobi procedure 2.14 since one has not enough common zeros for all multivariate Chebyshev polynomials of the same degree. Indeed the following proposition due to Li, Sun, and Xu shows that there is in general no hope to obtain enough common zeros.

Proposition 4.4 *The multivariate Chebyshev polynomials of the first kind associated to the root system A_2 of degree n have no common zero.*

PROOF: See [60]. □

But already in one variable there is a second kind of Chebyshev polynomials. Their multivariate counterparts have been studied in the context of cubature formulas in [76]. The A_2 second kind Chebyshev polynomials were used in [95] signal processing for the spatial hexagonal lattice. Recall that the second kind Chebyshev polynomials in one variable are defined as

$$U_k(x) = U_k(\cos\theta) = \frac{\sin(k+1)\theta}{\sin\theta} = \frac{\mathrm{e}^{2\pi\mathrm{i}(k+1)\theta} - \mathrm{e}^{2\pi\mathrm{i}(k+1)\theta}}{\mathrm{e}^{2\pi\mathrm{i}\theta} - \mathrm{e}^{-2\pi\mathrm{i}\theta}}. \tag{4.18}$$

Hence the second kind Chebyshev polynomials are anti-symmetrized exponentials with an index shift, divided by the minimal anti-symmetric sum. The $+1$ index shift has an interpretation as the sum of fundamental weights. The definition of the multivariate Chebyshev polynomials is as follows.

Definition 4.5 *Let W be a Weyl group of a root system R with weight lattice P and coroot lattice $Q^\vee$. Let $\delta = \sum_\omega \omega$ be the sum of fundamental weights. Then the multivariate Chebyshev polynomials of the second kind of weight $\lambda \in P$ are defined as*

$$U_\lambda(x_1, \dots, x_k) = \frac{\sum_{w \in W} \det(w)(\lambda + \delta, w\theta)}{\sum_{w \in W} \det(w)(\delta, w\theta)}. \tag{4.19}$$

They are polynomials in the variables x_i of (4.12).

Actually these are just the characters of semi-simple Lie algebras, as follows from the Weyl character formula in combination with the Weyl denominator formula [123–125].

That the second kind Chebyshev polynomials give rise to Gaußian cubature formulas has been shown for root systems of type A_n in [63]. For the other root systems one needs to adjust the grading on the polynomials as shown in [76]. Then one can even associate the common zeros with elements of finite order of the corresponding Lie group. This adjusted grading is useful for the first kind Chebyshev polynomial, as well, since only with this grading the $T_{k\omega_i}$ form a Gröbner basis for the ideal they generate.

Definition 4.6 *Let $\lambda \in P$ then its m-degree is given by*

$$\deg_{\mathrm{m}}(\lambda) = \langle \lambda, \alpha_0^\vee \rangle. \tag{4.20}$$

A monomial $x_1^{\lambda_1} \dots x_d^{\lambda_d}$ has m-degree $\deg_{\mathrm{m}}(\lambda) = \deg_{\mathrm{m}}((\lambda_1, \dots, \lambda_d))$.

The m-graded lexicographical ordering *on the monomials is defined by ordering the monomials with respect to* $\deg_{\mathrm{m}}$ *and then breaking ties by the lexicographical order on the variables.*

In case of type A_n all marks are equal to 1, so in these cases the m-degree coincides with the standard degree. The leading monomial of the Chebyshev polynomial T_λ with respect to the m-graded lexicographical ordering is $x_1^{\lambda_1} \dots x_d^{\lambda_d}$. By the recursion relations obtained from the shift relation the leading monomials with respect to m-graded lexicographical ordering of the polynomials $T_{n\omega_1}, \dots, T_{n\omega_d}$ are disjoint. Hence they form a Gröbner basis for the ideal they generate with respect to the m-graded lexicographical ordering.

4.2 Examples of fast cosine transforms on weight lattices

This section is concerned with the derivation of fast transform algorithms for analogues of the discrete cosine transform, as derived in Examples 2.7 and 2.9, *iii.)*, on other undirected lattices in higher dimensions. The general approach is to take some - depending on the number of common zeros they have - of the multivariate Chebyshev polynomials of specified degree to define a zero-dimensional variety. As signal module we will in these examples always choose the regular module. Then the multivariate Chebyshev polynomials of lower degree will form the basis for our signal model. By the shift relation of Prop. 4.3, *ii.)*, one obtains that the lattices in the visualization graph correspond to the associated weight lattices of the root system. Using the decomposition property (4.17) the fast algorithms are derived using Theorem 3.13. In the cases were we can obtain a transversal we derive fast algorithms using Theorem 3.1 instead. Furthermore we picture the underlying geometric principle, the stretching and folding, of these algorithms. The actual details of course differ between the different root systems. But the given examples show that it is in any case possible to derive fast algorithms for suitably defined cosine transforms on weight lattices.

The change of basis matrices can be described in general for each root system but require lengthy calculations using the recurrence relations and the invariance under the action of the Weyl group, i.e. Prop. 4.3, *4.14.)* and *ii.)*. Hence a Mathematica® [45] module was developed which is able to perform these algebraic calculations. It is available in a Github repository under the URL-address https://github.com/bseifert-HSA/basis-change-Chebyshev-transforms.

Consider the situation that the size of the subvarieties multiply and one wants to derive a fast algorithm using induced modules via Theorem 3.1. Then one has to ensure that the matrices R_i are sparse, as well. In the case of the Chebyshev polynomials this can be shown as follows. Denote the chosen subset of Chebyshev polynomials of degree n by $\mathbb{T}_n$. If one has $n = r \cdot m$ then by the stretching and folding property the fundamental domain gets stretched by the $T_{r\omega_i}$ to r^d copies of it and then gets folded back to the fundamental domain. Thus each common zero of $\mathbb{T}_m$ has at most r^d preimages under the $T_{r\omega_i}$ in the common zeros of $\mathbb{T}_n$. Thus the R_i have $O(|\mathsf{V}(\langle\, \mathbb{T}_n \,\rangle)|)$ entries, i.e. are sparse.

The first example is that of A_2, where the signal model implements a signal model of space signals sampled on a hexagonal lattice. This connection was first investigated in [95] and a fast algorithm derived in [96] using a variant of Theorem 3.13. Since in this case we will see that one has n^2 common zeros for $n = m \cdot r$ one obtains $n^2 = m^2 r^2$, so Theorem 3.1 in combination with Proposition 3.12 is applicable as well and yields another fast algorithm for computing the Fourier transform of the signal model. We present the alternative algorithm based on induced modules in some detail.

Example 4.7 We consider the signal model associated to Chebyshev polynomials of type A_2. The root system and fundamental region are shown in Fig. 4.4. In this case the shift relation reads

$$\begin{aligned} x \cdot T_{\lambda_1,\lambda_2} &= \tfrac{1}{3}(T_{\lambda_1+1,\lambda_2} + T_{\lambda_1,\lambda_2-1} + T_{\lambda_1-1,\lambda_2+1}), \\ y \cdot T_{\lambda_1,\lambda_2} &= \tfrac{1}{3}(T_{\lambda_1,\lambda_2+1} + T_{\lambda_1-1,\lambda_2} + T_{\lambda_1+1,\lambda_2-1}). \end{aligned} \tag{4.21}$$

A set of sufficient starting conditions for running the recursion is

$$\begin{aligned} T_{0,0} &= 1, \\ T_{1,0} &= x, \\ T_{0,1} &= y, \\ T_{1,1} &= \tfrac{3}{2}xy - \tfrac{1}{2}. \end{aligned} \tag{4.22}$$

We consider the signal model consisting of

$$\begin{aligned} \mathscr{A} &= \mathbb{C}[x,y]/\langle\, T_{0,n}, T_{n,0} \,\rangle, \\ M &= \mathscr{A} \text{ (as regular module)}, \end{aligned} \tag{4.23}$$

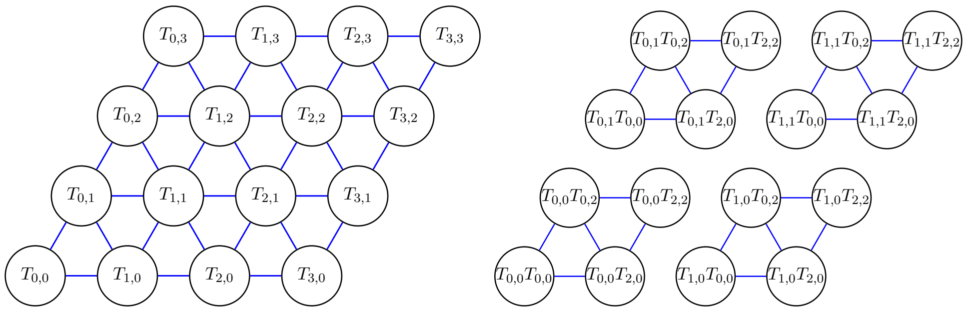

Figure 4.6: Visualization of the signal model for Chebyshev polynomials of type A_2, an undirected hexagonal lattice (left), and after representing the module as induction (right).

and

$$\begin{aligned} \Phi\colon \mathbb{C}^{n^2} &\longrightarrow M, \\ \Phi(s) &= \sum_{k,\ell=0}^{n} s_{k,\ell} T_{k,\ell}. \end{aligned} \tag{4.24}$$

The visualization unveils the hexagonal lattice underlying this signal model, cf. Fig. 4.6, obtained from the shift relations (4.21). Indeed one obtains an undirected hexagonal lattice, i.e. a space signal model. As already mentioned in the introductory part of this chapter, sampling using a hexagonal lattice as the advantage of obtaining the same information with 13.4 % less sampling points than one would need if one would use a regular rectangular lattice. In [96] the n^2 common zeros of $T_{n,0}$ and $T_{0,n}$ were described elementary. We proposed a geometric description of the common zeros in [109], as this shows the geometric mechanisms underlying the decomposition more clearly. The preimage of 0 in the θ-domain is $\frac{1}{3}\omega_1^\vee + \frac{1}{3}\omega_2^\vee$. Through the stretching-and-folding property and the condition that the common zeros be in the fundamental domain F one obtains

$$\begin{aligned} \mathsf{V}(\langle\, T_{n,0}, T_{0,n} \,\rangle) = \{ &\frac{1+3j}{3n}\omega_1^\vee + \frac{1+3k}{3n}\omega_2^\vee \mid 2 + 3(j+k) < 3n \} \\ \cup \{ &\frac{2+3j}{3n}\omega_1^\vee + \frac{2+3k}{3n}\omega_2^\vee \mid 4 + 3(j+k) < 3n \}, \end{aligned} \tag{4.25}$$

with $j,k = 0,\dots,n-1$. For each $\alpha \in \mathsf{V}(\langle\, T_{r,0}, T_{0,r} \,\rangle)$ one has $|\mathsf{V}(\langle\, T_{m,0} - \alpha_1, T_{0,m} - \alpha_2 \,\rangle)| = m^2$. The geometric mechanism of the distribution of the common zeros is illustrated in Fig. 4.7. Since for $n = r \cdot m$ one thus has $|\mathsf{V}(\langle\, T_{n,0}, T_{0,n} \,\rangle)| = |\mathsf{V}(\langle\, T_{r,0}, T_{0,r} \,\rangle)| \cdot |\mathsf{V}(\langle\, T_{m,0}, T_{0,m} \,\rangle)|$ and all the subalgebras $\mathbb{C}[x,y]/\langle\, T_{m,0} - \alpha_1, T_{0,m} - \alpha_2 \,\rangle$ are of equal dimension, by Prop. 3.12 any basis of $\mathbb{C}[x,y]/\langle\, T_{r,0}, T_{0,r} \,\rangle$ is a transversal of $\mathbb{C}[x,y]/\langle\, T_{m,0}, T_{0,m} \,\rangle$. We pick as basis of $\mathbb{C}[x,y]/\langle\, T_{r,0}, T_{0,r} \,\rangle$ the Chebyshev polynomials up to degree $r-1$. We consider only the case $r = 2$ as otherwise the following case analysis would become even longer. But using our Mathematica® module it is possible to obtain other variants, as well. Thus the change of basis is

$$(T_{0,0}, \dots, T_{n,n}) \to (T_{0,0}T_{0,0}(T_{m,0}, T_{0,m}), \dots, T_{m-1,m-1}T_{1,1}(T_{m,0}, T_{0,m}).$$

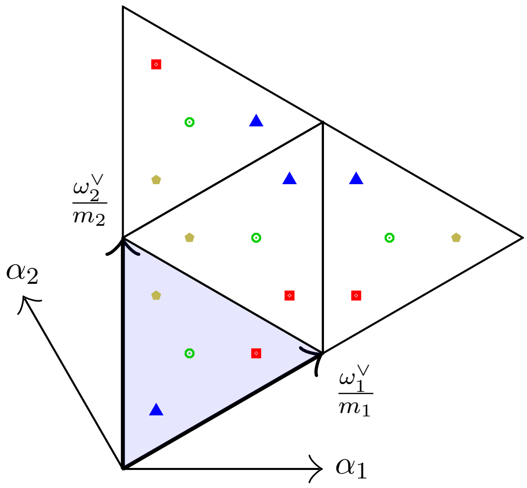

Figure 4.7: The common zeros of $\langle\ T_{4,0}, T_{0,4}\ \rangle$ are shown after being stretched by a factor of 2. The action of the affine Weyl group, folding the stretched triangle back to the fundamental domain, is indicated. The different colors indicate which common zeros of $\langle\ T_{4,0}, T_{0,4}\ \rangle$ are common zeros of which $\langle\ T_{2,0} - \alpha_1, T_{0,2} - \alpha_2\ \rangle$.

To deduce the basis change one has to investigate the cases of indices ranging from $(0,0)$ to (m,m), from $(m+1,0)$ to $(2m,m)$, from $(0,m+1)$ to $(m,2m)$, and from $(m+1,m+1)$ to $(2m,2m)$. Let $k,\ell < m$. The orbit of (k,ℓ) under the Weyl group $W(A_2)$ is

$$\{(-k-\ell,,k),(-k,k+\ell),(-\ell,-k),(\ell,-k-\ell),(k,\ell),(k+\ell,-\ell)\}.$$

Hence one has additionally to distinguish between the cases were $k=0$, $\ell=0$, $k+\ell<m$, $k+\ell=m$, and $k+\ell>m$.

In region I, i.e. indices ranging from $(0,0)$ to (m,m), nothing has to be done

$$T_{k,\ell} = T_{k,\ell}$$

For region II, i.e. indices ranging from $(m+1,0)$ to $(2m,m)$, one obtains

$$T_{m+k,\ell} = \begin{cases} T_{m,0} & k,\ell = 0, \\ -2T_{m-k,k} + 3T_{k,0}T_{m,0} & \ell = 0, \\ -\frac{1}{2}T_{m-\ell,0} + \frac{3}{2}T_{0,\ell}T_{m,0} & k = 0, \\ -T_{m-k,k+\ell} - T_{m-k-\ell,k} + 3T_{k,\ell}T_{m,0} & k+\ell < m, \\ -\frac{1}{2}T_{0,k} - \frac{3}{2}T_{\ell,0}T_{0,m} + 3T_{k,\ell}T_{m,0} & k+\ell = m, \\ T_{\ell,2m-k-\ell} - 3T_{m-k,k+\ell-m}T_{0,m} + 3T_{k,\ell}T_{m,0} & k+\ell > m. \end{cases}$$

Region III, i.e. indices ranging from $(0,m+1)$ to $(m,2m)$, can be calculated as follows

$$T_{k,m+\ell} = \begin{cases} T_{0,m} & k,\ell = 0, \\ -\frac{1}{2}T_{0,m-k} + \frac{3}{2}T_{k,0}T_{0,m} & \ell = 0, \\ -2T_{\ell,m-\ell} + 3T_{0,\ell}T_{0,m} & k = 0, \\ -T_{\ell,m-k-\ell} - T_{k+\ell,m-\ell} + 3T_{k,\ell}T_{0,m} & k + l < m, \\ -\frac{1}{2}T_{\ell,0} - \frac{3}{2}T_{0,-m-\ell}T_{m,0} + 3T_{k,\ell}T_{0,m} & k + \ell = m, \\ T_{2m-k-\ell,k} - 3T_{k+\ell-m,m-\ell}T_{m,0} + 3T_{k,\ell}T_{0,m} & k + \ell > m. \end{cases}$$

For region IV, i.e. indices ranging from $(m+1, m+1)$ to $(2m, 2m)$, one obtains

$$T_{m+k,m+\ell} = \begin{cases} T_{m,m} & k,\ell = 0, \\ T_{k,0} - 3T_{m-k,k}T_{0,m} + 3T_{k,0}T_{m,m} & \ell = 0, \\ T_{0,\ell} - 3T_{\ell,m-\ell}T_{m,0} + 3T_{0,\ell}T_{m,m} & k = 0, \\ \begin{aligned} &2T_{k,\ell} - T_{m-\ell,m-k} - 3T_{m-k,k+\ell}T_{0,m} \cdots \\ &- 3T_{k+\ell,m-\ell}T_{m,0} + 6T_{k,\ell}T_{m,m} \end{aligned} & k + \ell < m, \\ T_{k,\ell} - \frac{3}{2}T_{0,k}T_{0,m} - \frac{3}{2}T_{\ell,0}T_{m,0} + 6T_{k,\ell}T_{m,m} & k + \ell = m, \\ \begin{aligned} &- T_{m-\ell,m-k} + 2T_{k,\ell} - 3T_{m-k,k+\ell-m}T_{m,0} \cdots \\ &+ 3T_{\ell,2m-k-\ell}T_{0,m} + 3T_{2m-k-\ell,k}T_{m,0} \cdots \\ &- 3T_{k+\ell-m,m-\ell}T_{0,m} + 6T_{k,\ell}T_{m,m} \end{aligned} & k + \ell > m \end{cases}$$

by keeping in mind that $T_{2m,0} = T_{0,2m} = 0$ in the signal module.

Since the basis change and the R_i are sparse one obtains by Prop. 3.2 and Theorem 3.1 a $O(n^2 \log(n))$ algorithm for the computation of the Fourier transform of the A_2 signal model. This is substantially faster than the naive $O(n^4)$-approach.

The next example considers a signal model associated to the root system A_3. In [111] a fast algorithm for the Fourier transform of the signal model relying on Theorem 3.13 was derived. Note that we have in case of A_3 that the varieties consist of n^3 points, so we can apply Theorem 3.1 in combination with Proposition 3.12 as well to obtain another fast algorithm.

Example 4.8 The next example is considered with the signal model associated to Chebyshev polynomials of type A_3. The root system and the fundamental region are shown in the upper part of Fig 4.8. The recursion relations reads in this case

$$\begin{aligned} x \cdot T_{\lambda_1,\lambda_2,\lambda_3} &= \tfrac{1}{4}(T_{\lambda_1+1,\lambda_2,\lambda_3} + T_{\lambda_1,\lambda_2,\lambda_3-1} + T_{\lambda_1,\lambda_2-1,\lambda_3-1} + T_{\lambda_2-1,\lambda_2+1,\lambda_3}), \\ y \cdot T_{\lambda_1,\lambda_2,\lambda_3} &= \tfrac{1}{6}(T_{\lambda_1,\lambda_2+1,\lambda_3} + T_{\lambda_1+1,\lambda_2,\lambda_3-1} + T_{\lambda_1+1,\lambda_2-1,\lambda_3+1} \\ &\qquad + T_{\lambda_1-1,\lambda_2+1,\lambda_3-1} + T_{\lambda_1-1,\lambda_2,\lambda_3+1} + T_{\lambda_1,\lambda_2-1,\lambda_3}), \\ z \cdot T_{\lambda_1,\lambda_2,\lambda_3} &= \tfrac{1}{4}(T_{\lambda_1,\lambda_2,\lambda_3+1} + T_{\lambda_1,\lambda_2+1,\lambda_3-1} + T_{\lambda_1+1,\lambda_2-1,\lambda_3} + T_{\lambda_1-1,\lambda_2\lambda_3}), \end{aligned} \tag{4.26}$$

with starting conditions

$$\begin{aligned}
T_{0,0,0} &= 1,\\
T_{1,0,0} &= x,\\
T_{0,1,0} &= y,\\
T_{0,0,1} &= z,\\
T_{1,1,0} &= 2xy - z,\\
T_{1,0,1} &= \tfrac{4}{3}xz - \tfrac{1}{3},\\
T_{0,1,1} &= 2yz - x,\\
T_{1,1,1} &= 4xyz - 2x^2 - 2z^2 + y.
\end{aligned} \tag{4.27}$$

Thus the neighbourhood of each node in the signal model looks as depicted in the lower part of Fig. 4.8. This structure resembles a body centered cubic lattice. The Voronoi cell of the lattice points is a truncated octahedron and shown for the central point. We consider the signal model consisting of

$$\begin{aligned}
\mathscr{A} &= \mathbb{C}[x,y,z]/\langle\, T_{0,0,n}, T_{0,n,0}, T_{n,0,0}\,\rangle,\\
M &= \mathscr{A} \text{ (as regular module)},
\end{aligned} \tag{4.28}$$

and

$$\begin{aligned}
\Phi\colon \mathbb{C}^{n^3} &\longrightarrow M,\\
\Phi(s) &= \sum_{k,\ell,p=0}^{n} s_{k,\ell,p} T_{k,\ell,p}.
\end{aligned} \tag{4.29}$$

Denote $\mathbb{T}_n = \langle\, T_{0,0,n}, T_{0,n,0}, T_{n,0,0}\,\rangle$. In [111] we used an elementary ad-hoc description of the variety $\mathsf{V}(\mathbb{T}_n)$. Here we will give a more geometric description as for the A_2 case. In terms of the coweights we have the n^3 common zeros of the $T_{n\omega_i}$ be given as

$$\begin{aligned}
\mathsf{V}(\mathbb{T}_n) = &\left\{\frac{1+4j}{4n}\omega_1^\vee + \frac{1+4k}{4n}\omega_2^\vee + \frac{1+4\ell}{4n}\omega_3^\vee \mid 3 + 4(k+j+\ell) < 4n\right\}\\
\cup &\left\{\frac{2+4j}{4n}\omega_1^\vee + \frac{1+4k}{4n}\omega_2^\vee + \frac{2+4\ell}{4n}\omega_3^\vee \mid 5 + 4(k+j+\ell) < 4n\right\}\\
\cup &\left\{\frac{3+4j}{4n}\omega_1^\vee + \frac{2+4k}{4n}\omega_2^\vee + \frac{1+4\ell}{4n}\omega_3^\vee \mid 6 + 4(k+j+\ell) < 4n\right\}\\
\cup &\{\frac{1+4j}{4n}\omega_1^\vee + \frac{2+4k}{4n}\omega_2^\vee + \frac{3+4\ell}{4n}\omega_3^\vee \mid 6 + 4(k+j+\ell) < 4n\}\\
\cup &\left\{\frac{2+4j}{4n}\omega_1^\vee + \frac{3+4k}{4n}\omega_2^\vee + \frac{2+4\ell}{4n}\omega_3^\vee \mid 7 + 4(k+j+\ell) < 4n\right\}\\
\cup &\left\{\frac{3+4j}{4n}\omega_1^\vee + \frac{3+4k}{4n}\omega_2^\vee + \frac{3+4\ell}{4n}\omega_3^\vee \mid 9 + 4(k+j+\ell) < 4n\right\}
\end{aligned} \tag{4.30}$$

Since for $n = r \cdot m$ one thus has $|\mathsf{V}(\mathbb{T}_n)| = |\mathsf{V}(\mathbb{T}_r)| \cdot |\mathsf{V}(\mathbb{T}_m)|$ and all the subalgebras $\mathbb{C}[x,y,z]/\langle\, T_{m,0,0} - \alpha_1, T_{0,m,0} - \alpha_2, T_{0,0,m} - \alpha_3\,\rangle$ are hence of equal dimension.

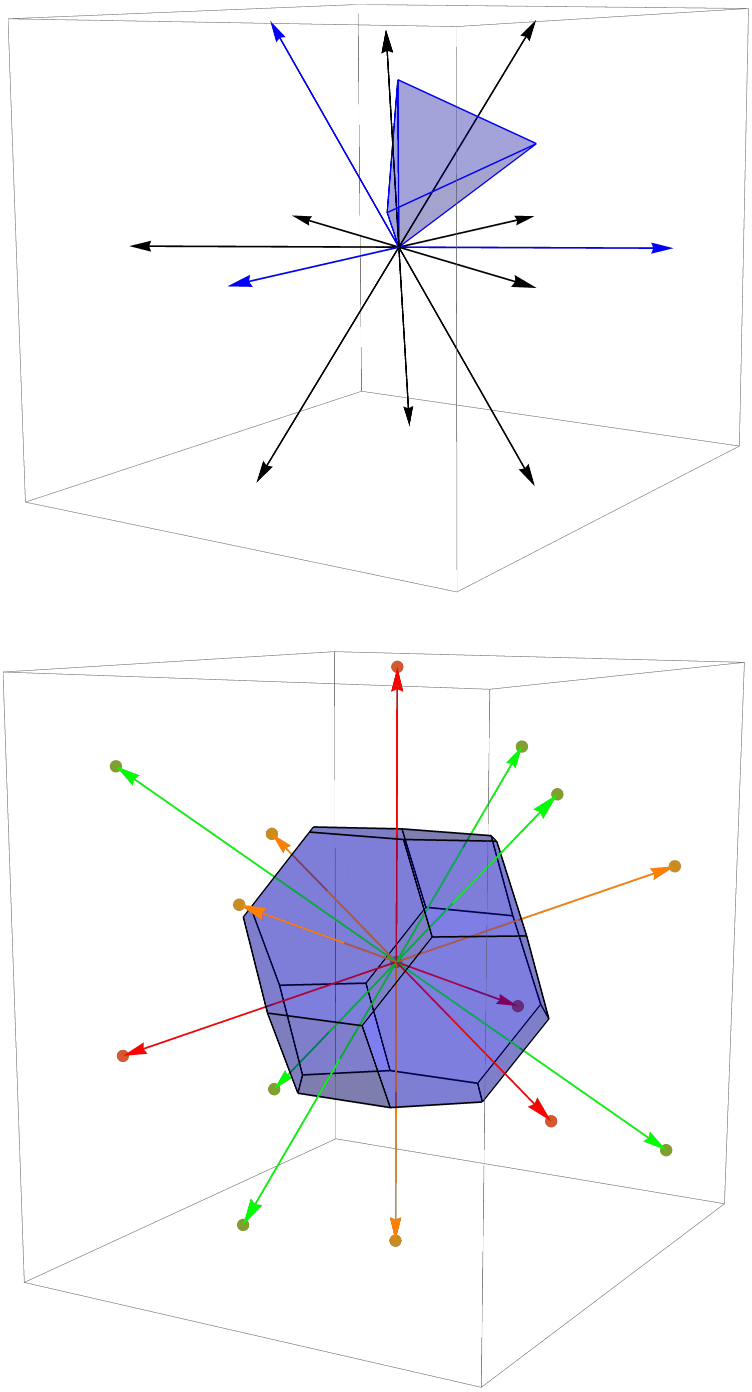

Figure 4.8: The upper part shows root system A_3 with the simple roots in blue and the fundamental region. In the lower part the neighbourhood of each node in the visualization graph of the A_3 Chebyshev signal model resembles the neighbourhood in an BCC lattice. The shifts of x, y, and z are shown in orange, green, and red, respectively. The Voronoi cell of the lattice - a truncated octahedron - is shown for the center point.

Moreover, by Prop. 3.12 any basis of $\mathbb{C}[x,y,z]/\mathbb{T}_r$ is a transversal of $\mathbb{C}[x,y,z]/\mathbb{T}_m$. We pick as basis of $\mathbb{C}[x,y,z]/\mathbb{T}_r$ the Chebyshev polynomials up to degree $r-1$.

That the basis change in this example is sparse follows from explicit calculations, which can be found online under https://github.com/bseifert-HSA/basis-change-Chebyshev-transforms/tree/master/A3. Since the results are very lengthy, we omit them here. We just mention that one has to subdivide the basis into 8 regions and the case analysis consists of 45 cases in each region.

Since one has a sparse basis change and sparse matrices R_i one obtains, similar to the case A_2, a $O(n^3\log(n))$ algorithm, which is substantially faster than the naive $O(n^6)$ approach.

The next example, based on Chebyshev polynomials of type C_2, which are a special case of the Koornwinder polynomials in Example 2.16 with different normalization, was presented by us in [109,110]. This example was the starting point to investigate the multivariate Gauß-Jacobi procedure from Theorem 2.14 since it was the first example we became aware of where this method did work in [110], where we used an elementary description of the common zeros. A description of the fast algorithm for this signal model was first presented in [109] with less details.

Example 4.9 In case C_2 the shift relations are

$$\begin{aligned} x_1\cdot T_{k,\ell} &= \tfrac{1}{4}(T_{k+1,\ell}+T_{k-1,\ell}+T_{k-1,\ell+2}+T_{k+1,\ell-2}),\\ x_2\cdot T_{k,\ell} &= \tfrac{1}{4}(T_{k,\ell+1}+T_{k,\ell-1}+T_{k-1,\ell+1}+T_{k+1,\ell-1}). \end{aligned} \tag{4.31}$$

A set of sufficient starting conditions for running the recurrence relation is

$$\begin{aligned} T_{0,0} &= 1,\\ T_{1,0} &= x_1,\\ T_{0,1} &= x_2,\\ T_{1,1} &= 2x_1x_2-x_1. \end{aligned} \tag{4.32}$$

The weight vector for the total m-degree lexicographical ordering of the monomials is $(1,2)$. That is $\deg_{\mathrm{m}}(x_1)=1$ and $\deg_{\mathrm{m}}(x_2)=2$.

Denote by $\mathbb{T}_n=\{T_{k,\ell}\mid k+\ell=n\}$. Then one has a three-term recurrence of the form

$$x_i\mathbb{T}_k = A_{k,i}\mathbb{T}_{k+1}+B_{k,i}\mathbb{T}_k+C_{k,i}\mathbb{T}_{k-1}, \tag{4.33}$$

were the matrices $A_{k,i}, B_{k,i}$, and $C_{k,i}$ can be deduced from the shift relations (4.31). For example for the x_1 shift one obtains

$$A_{k,1} = \begin{bmatrix} 0 & 1/2 & 0 & \dots & & 0\\ 1/4 & 0 & 1/4 & 0 & \dots & 0\\ 0 & \ddots & \ddots & \ddots & & \vdots\\ \dots & 0 & 1/4 & 0 & 1/4 & 0\\ 0 & \dots & 0 & 1/2 & 0 & 1/4 \end{bmatrix}, \tag{4.34}$$

$$B_{k,1} = \begin{bmatrix} 0 & \dots & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 1/4 & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix}, \tag{4.35}$$

and

$$C_{k,1} = \begin{bmatrix} 0 & 1/2 & 0 & \dots & 0 \\ 1/4 & 0 & 1/4 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ & & 1/4 & 0 & 1/4 \\ \vdots & & 0 & 1/4 & 0 \\ 0 & \dots & & 0 & 1/4 \end{bmatrix}, \tag{4.36}$$

with special case $B_{1,1} = \left[\begin{smallmatrix} 1/2 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. The multivariate Christoffel-Darboux formula (2.41) can be realized using the matrices $H_0 = \frac{1}{2}$ and $H_k = \mathsf{diag}(\frac{1}{8}, \frac{1}{16}, \dots, \frac{1}{16}, \frac{1}{8})$.

We consider the signal model consisting of

$$\begin{aligned} \mathscr{A} &= \mathbb{R}[x_1, x_2]/\langle\ \mathbb{T}_n\ \rangle, \\ M &= \mathscr{A} \text{ (as regular module)}, \end{aligned} \tag{4.37}$$

and

$$\begin{aligned} \Phi \colon \mathbb{R}^{\frac{n(n+1)}{2}} &\longrightarrow M, \\ \Phi(s) &= \sum_{k+\ell<\frac{n(n+1)}{2}} s_{k,\ell} T_{k,\ell}. \end{aligned} \tag{4.38}$$

The signal model has a visualization, which resembles a triangle, cf. Fig. 4.9. We present a geometric description of the common zeros using the coweights. That is the common zeros are given as

$$\mathsf{V}(\langle\ \mathbb{T}_n\ \rangle) = \left\{ \tfrac{2j+1}{2n}\omega_1^\vee + \tfrac{k}{2n}\omega_2^\vee \;\middle|\; j,k = 0,\dots,n-1,\ j+k<n \right\}. \tag{4.39}$$

This results in $\frac{n(n+1)}{2}$ common zeros.

We derive a fast algorithm in case $n = 2\cdot m$. Since for $n = 2\cdot m$ it is $\frac{n(n+1)}{2} \neq \frac{2(2+1)}{2}\cdot\frac{m(m+1)}{2}$ one does not get an induction via the decomposition.

Due to the decomposition property 4.3, *iii.)*, the map

$$(x_1, x_2) \mapsto (T_{2,0}(x_1,x_2), T_{0,2}(x_1,x_2)) \tag{4.40}$$

maps the variety $\mathsf{V}(\mathbb{T}_n)$ to the variety $\mathsf{V}(\mathbb{T}_m)$. The map

$$(x_1, x_2) \mapsto (T_{m,0}(x_1,x_2), T_{0,m}(x_1,x_2) \tag{4.41}$$

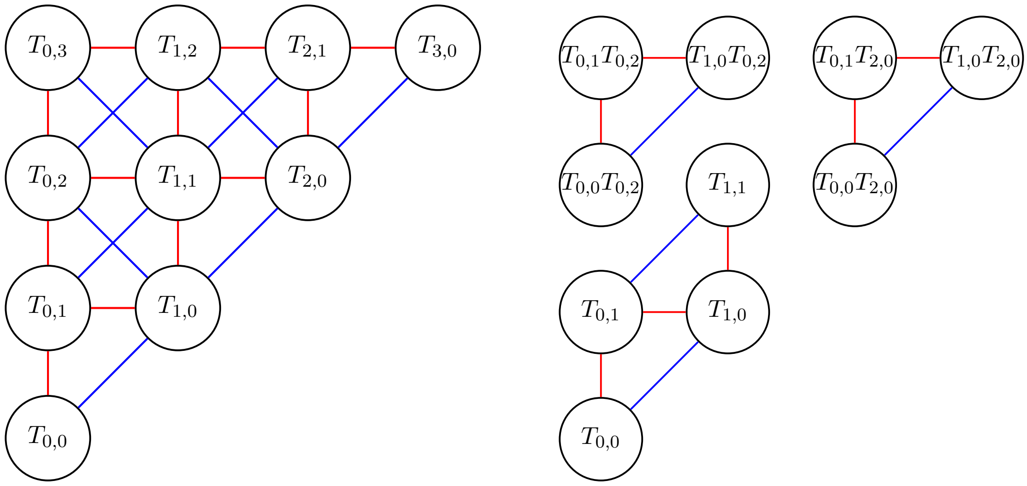

Figure 4.9: Visualization of the signal model for Chebyshev polynomials of type C_2 on the left. The shifts of x_1 are blue and the shifts of x_2 are red colored. On the right the decomposed lattice after the basis change is shown.

stretches the fundamental region F by a factor of m and folds it back under the affine Weyl group. After the stretching operation one obtains m^2 copies of the fundamental region. Thus if the stretched zero is in the interior of F after folding back one obtains m^2 common zeros. If the stretched zeros is on the boundary of F after folding back always two of the copies of F in the interior of the stretched region share these common zeros. Hence in this case one only obtains $\frac{m(m+1)}{2}$ common zeros. Of the common zeros of $\mathbb{T}_2$ there is one in the interior and two on the boundary. Since these are the images under the stretching and folding operation one obtains two subalgebras with $\frac{m(m+1)}{2}$ common zeros and one subalgebra with m^2 common zeros. Using Mathematica® the following general basis change is obtained. The basis change is from $(T_{k,\ell} \mid k+\ell < n)$ to $(T_{k,\ell}T_{t,p}(T_{m,0}, T_{0,m}) \mid k+\ell < m, t+p < 2)$. For this one as to distinguish between three regions of the indices ranging from $(0,0)$ to (m,m), from $(m+1,0)$ to $(2m,m)$, and from $(0,m+1)$ to $(m,2m)$. Let $k,\ell < m$.

The orbit of (k,ℓ) under the Weyl group $W(C_2)$ is

$$\{(-k-2\ell,\ell),(-k-2\ell,k+\ell),(-k,-\ell),(-k,k+\ell),\\ (k,-k-\ell),k,\ell,(k+2\ell,-k-\ell),(k+2\ell,-\ell)\}.$$

Hence one has to distinguish the eight cases

$$\begin{aligned}
&\ell = 0,\\
&k = 0 \text{ and } 2\cdot\ell < m,\\
&k = 0 \text{ and } 2\cdot\ell = m,\\
&k = 0 \text{ and } 2\cdot\ell > m,\\
&k + 2\cdot\ell < m,\\
&k + 2\cdot\ell = m,\\
&k + 2\cdot\ell > m,
\end{aligned}$$

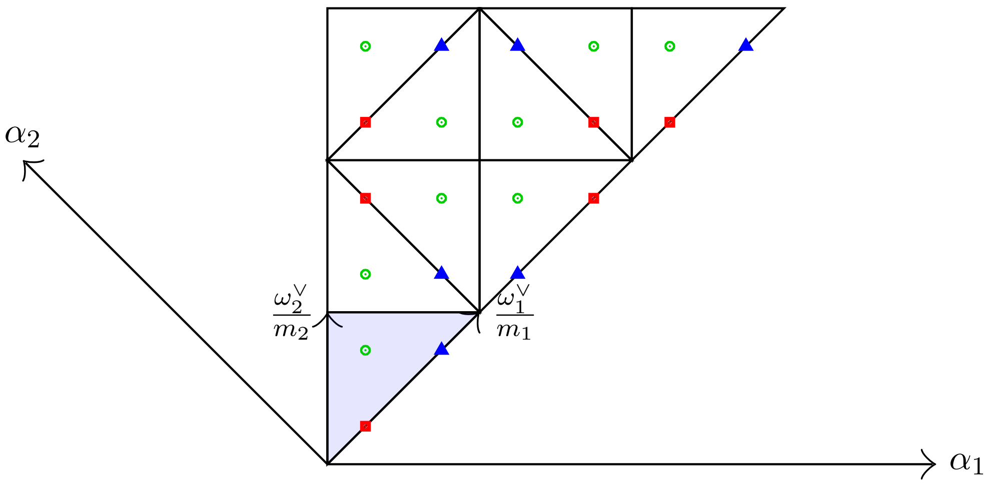

Figure 4.10: The three classes of common zeros of for the skew transforms in case $n = 2 \cdot 3$. The common zeros of $\mathbb{T}_6$ are shown after being stretched by a factor of 3.The action of the affine Weyl group, folding the stretched triangle back to the fundamental domain, is indicated. The different colors indicate which common zeros of $\langle\, T_{6,0}, T_{0,6} \,\rangle$ are common zeros of which $\langle\, T_{3,0} - \alpha_1, T_{0,3} - \alpha_2 \,\rangle$.

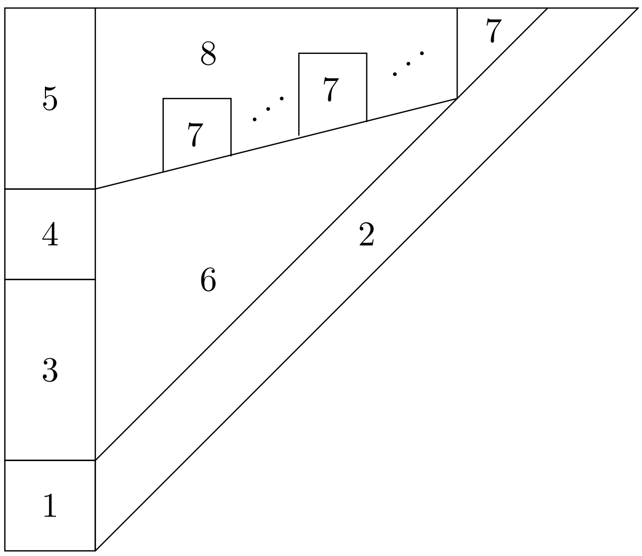

Figure 4.11: The different positions of the basis elements correspond to one of the eight cases.

as visualized in Fig. 4.11. Note that since for any basis elements index (t, p) one has $t + p < n = 2m$ one always has $k + \ell < m$ in the sequel.

For region I, i.e. indices ranging from $(0, 0)$ to (m, m), one has

$$T_{k,\ell} = T_{k,\ell}$$

since every element is in the new basis as well as in the old. For region II, i.e. indices ranging from $(m + 1, 0)$ to $(2m, m)$, one obtains

$$T_{m+k,\ell} = \begin{cases} T_{m,0} & k,\ell = 0, \\ -2T_{-k+m,k} - T_{-k+m,0} + 4T_{k,0} \cdot T_{m,0} & \ell = 0, \\ -T_{-2\cdot\ell+m,\ell} + 2T_{0,\ell} \cdot T_{m,0} & k = 0, 2\ell < m, \\ -T_{0,\ell} + 2T_{0,\ell} \cdot T_{m,0} & k = 0, 2\ell = m, \\ -T_{2\cdot\ell-m,-\ell+m} + 2T_{0,\ell} \cdot T_{m,0} & k = 0, 2\ell > m, \\ \begin{array}{l} - T_{-k+m,k+\ell} - T_{-k-2\cdot\ell+m,\ell} \cdots \\ - T_{-k-2\cdot\ell+m,k+\ell} + 4T_{k,\ell} \cdot T_{m,0} \end{array} & k + 2\ell < m, \\ -T_{0,\ell} - T_{0,k+\ell} - T_{2\cdot\ell,k+\ell} + 4T_{k,\ell} \cdot T_{m,0} & k + 2\ell = m, \\ \begin{array}{l} - T_{k+2\cdot\ell-m,-\ell+m} - T_{k+2\cdot\ell-m,-k-\ell+m} \cdots \\ - T_{-k+m,k+\ell} + 4T_{k,\ell} \cdot T_{m,0} \end{array} & k + 2\ell > m. \end{cases}$$

And finally for region III, i.e. indices ranging from $(0, m + 1)$ to $(m, 2m)$, one gets

$$T_{k,m+\ell} = \begin{cases} T_{0,m} & k,\ell = 0, \\ -T_{k,-k+m} + 2T_{k,0} \cdot T_{0,m} & \ell = 0, \\ -T_{0,-\ell+m} - 2T_{2\cdot\ell,-\ell+m} + 4T_{0,\ell} \cdot T_{0,m} & k = 0, 2\ell < m, \\ T_{0,\ell} - 4T_{0,\ell}T_{m,0} + 4T_{0,\ell} \cdot T_{0,m} & k = 0, 2\ell = m, \\ \begin{array}{l} 2T_{0,\ell} + T_{0,m-\ell} + 2T_{2m-2\ell,\ell} \\ - 8T_{2\ell-m,m-\ell} + 4T_{0,\ell} \cdot T_{0,m} \end{array} & k = 0, 2\ell > m, \\ \begin{array}{l} - T_{k,-k-\ell+m} - T_{k+2\cdot\ell,-\ell+m} \\ - T_{k+2\cdot\ell,-k-\ell+m} + 4T_{k,\ell} \cdot T_{0,m} \end{array} & k + 2\ell < m, \\ T_{k,\ell} - 2T_{0,\ell}T_{m,0} - 2T_{0,m-\ell}T_{m,0} + 4T_{k,\ell}T_{0,m} & k + 2\ell = m, \\ \begin{array}{l} T_{k,m-k-\ell} + 2T_{k,\ell} + T_{2\cdot m-k-2\cdot\ell,k+\ell} \cdots \\ - 4T_{k+2\cdot\ell-m,-\ell+m} \cdot T_{m,0} + T_{2\cdot m-k-2\cdot\ell,\ell} \cdots \\ - 4T_{k+2\cdot\ell-m,-k-\ell+m} \cdot T_{m,0} + 4T_{k,\ell} \cdot T_{0,m} \end{array} & k + 2\ell > m. \end{cases}$$

Since the obtained change of basis is sparse, we get a fast $O(n^2 \log n)$ algorithm by Theorem 3.13.

Since the number of common zeros of $\mathbb{T}_n$ equals $\dim \Pi_2^{n-1}$ Theorem 2.14 implies the existence of an orthogonal transform. Denote the diagonal matrix with inverted entries by

$$D_n = \mathsf{diag}\Big(1 \Big/ \Big(\mathbb{T}_{n-1}^\top(x) H_{n-1}^{-1} A_{n-1,1} \tfrac{\partial}{\partial x_1} \mathbb{T}_n(x)\Big)\Big), \tag{4.42}$$

and by

$$H_n^{\oplus} = \bigoplus_{k=0}^{n-1} H_k^{-1} \tag{4.43}$$

the direct sum of the H_k^{-1} matrices. Reasoning analogously to the proof of Theorem 2.14 an orthogonal version of the transform is given by

$$\mathcal{F}^{\mathsf{orth}} = \sqrt{H_n^{\oplus}} \cdot \mathcal{F} \cdot \sqrt{D_n}. \tag{4.44}$$

The matrix $H_n^{\oplus}$ is needed here since we do not have orthonormal but only orthogonal polynomials.

4.3 Generating functions of matrix-valued and multivariate Chebyshev polynomials

When we considered the multivariate Chebyshev polynomials of the second kind in Def. 4.5, we gave an ad-hoc explanation on the appearance of the sum of fundamental weights and the division by the minimal anti-symmetric sum. In this section we will consider these things from a more conceptual point of view, even valid for all representations and not just one-dimensional ones. This construction is due to [44], which was inspired by [37]. For this one observes that the same construction as in (4.11) enriched by a representation $\rho\colon W \longrightarrow \mathsf{GL}(V_\rho)$ of the Weyl group

$$(\lambda, \theta)_\rho = \sum_{w \in W} \rho(w) \exp(2\pi \mathrm{i} \langle \lambda, w^\top \theta \rangle) \tag{4.45}$$

leads to equivariant sums

$$(\lambda, w^\top \theta)_\rho = \rho(w)^\top (\lambda, \theta)_\rho. \tag{4.46}$$

One is interested in finding a projection from these equivariant sums to invariant sums. The correct definition needs some technicalities first derived in [44].

One starts by investigating under which conditions the matrix-valued sums (4.45) possess rank deficiency. This can occur either if $\theta = w^\top \theta$ for some $\theta \in F$, which can be avoided by choosing a subset of F as domain of definition, or if $(\lambda, \theta)_\rho = (\lambda, w^\top \theta)_\rho$ for all $\theta \in F$. Both cases imply, using equivariance,

$$(\lambda, \theta)_\rho = (\lambda, w^\top \theta)_\rho = \rho(w)^\top (\lambda, \theta)_\rho. \tag{4.47}$$

That is, the rows of $(\lambda, \theta)_\rho$ lie in the eigenspace of $\rho(w)$ corresponding to the eigenvalue 1. This motivates to consider only special weights.

A weight $\lambda = \sum \lambda_i \omega_i$ has signature $I \subseteq \{1, 2, \ldots, d\}$ if $\lambda_i \neq 0$ if and only if $i \in I$. Conversely to every $I \subseteq \{1, \ldots, d\}$ the associated primal weight is $\lambda_I = \sum_{i \in I} \omega_i$. The intersection of the 1-eigenspaces of a signature I are defined as

$$V_I = \bigcap_{i \in I} V_{\{i\}} \tag{4.48}$$

with

$$V_{\{i\}} = \{x \in V_\rho \mid \rho(\sigma_{\alpha_i})x = x\}. \tag{4.49}$$

Denote by $I^c = \{1, \ldots, d\} \setminus I$. Then for any $j \in I^c$ one has

$$(\lambda, \theta)_\rho = (\lambda, \sigma_{\alpha_j}\theta). \tag{4.50}$$

Thus

$$\operatorname{rank}(\lambda, \theta)_\rho \leq V_{I^c} \tag{4.51}$$

and in turn

$$(\lambda, \theta)_\rho = 0 \tag{4.52}$$

if $V_{I^c} = \{0\}$. This justifies the following definition.

Definition 4.10 *A weight $\lambda \in P$ and its signature I are called admissible if $V_{I^c} \neq \{0\}$. The set of minimal admissible signatures is denoted by $\mathcal{A}$.*

For example for the trivial representation all weights are admissible, while for the determinant representation the only admissible signature is $\{1, \ldots, d\}$. Thus for the determinant representation the admissible weights in the positive cone are of the form $\lambda + \delta$, with $\delta = \sum_i \omega_i = \frac{1}{2} \sum \alpha_i$ the half-sum of positive roots, well-known from Weyls character formula [123].

For the second kind Chebyshev polynomials the projection is hence given as $(\delta, \theta)_{\text{det}}$. But δ is the primal weight of the minimal admissible set. The projection from the equivariant to the invariant matrix-sums is hence defined as the minimal function

$$S^\rho(x_1(\theta), \ldots, x_d(\theta)) = \sum_{I \in \mathcal{A}} (\lambda_I, \theta)_\rho. \tag{4.53}$$

Now the definition of the multivariate and matrix-valued Chebyshev polynomials is as follows.

Definition 4.11 *Let ρ be a representation of a Weyl group and $\lambda \in P$ an admissible weight. The λth Chebyshev polynomial A^ρ_λ is the solution to*

$$S^\rho(x_1(\theta), \ldots, x_d(\theta)) \cdot A^\rho_\lambda(x_1(\theta), \ldots, x_d(\theta)) = (\lambda, \theta)_\rho. \tag{4.54}$$

Remark 4.12 Note that it is up to now not known if S^ρ is invertible for all representations. The examples given below strongly indicate that the minimal functions should be invertible in general. The investigation of a proof of the invertibility of S^ρ is work in progress.

The multivariate and matrix-valued Chebyshev polynomials are polynomials in the variables (4.12) as follows from a theorem by Kostant [55], which is the matrix version of the theorem by Chevalley [11] which can be used in the scalar-valued case to show that one has indeed polynomials . They are subject to the shift relation

$$T_{\lambda_1} \cdot A^\rho_{\lambda_2} = \frac{1}{|W|} \sum_{w \in W} A^\rho_{\lambda_1 + w\lambda_2} \tag{4.55}$$

and are invariant in θ since they are polynomials in (4.12) but equivariant in the index

$$A^{\rho}_{w\lambda} = A^{\rho}_{\lambda} \cdot \rho(w)^{\top}. \tag{4.56}$$

Since one can choose different representations with the same properties, there is a notion of equivalence for representations. Two representations ρ_1 and ρ_2 are called equivalent if there exists a matrix E such that

$$E \cdot \rho_1 \cdot E^{-1} = \rho_2. \tag{4.57}$$

This motivates the introduction of the following definition of equivalent Chebyshev polynomials.

Definition 4.13 *Two Chebyshev polynomials to representations ρ_1 and ρ_2 are called equivalent if there exists an invertible matrix E such that*

$$E \cdot A^{\rho_1}_{\lambda} \cdot E^{-1} = A^{\rho_2}_{\lambda}. \tag{4.58}$$

Equivalent representations lead to equivalent Chebyshev polynomials.

Lemma 4.14 *Let ρ_1 and ρ_2 be two equivalent representations. Then the Chebyshev polynomials $A^{\rho_1}_{\lambda}$ and $A^{\rho_2}_{\lambda}$ are equivalent for any λ.*

Proof: Let E be an invertible matrix such that $E \cdot \rho_1 \cdot E^{-1} = \rho_2$. Then

$$\begin{aligned}
E(\lambda, \theta)_{\rho_1} E^{-1} &= \sum_{w \in W} E \rho_1(w) E^{-1} \exp(2\pi \mathrm{i} \langle \lambda, w^{\top} \theta \rangle) \\
&= \sum_{w \in W} \rho_2(w) \exp(2\pi \mathrm{i} \langle \lambda, w^{\top} \theta \rangle) \\
&= (\lambda, \theta)_{\rho_2},
\end{aligned}$$

and hence $E \cdot S^{\rho_1} E^{-1} = S^{\rho_2}$, as well. The Chebyshev polynomial $A^{\rho_2}_{\lambda}$ is the solution to

$$S^{\rho_2} A^{\rho_2}_{\lambda} = (\lambda, \theta)_{\rho_2}.$$

Conjugation of the definining equation

$$S^{\rho_1} A^{\rho_1}_{\lambda} = (\lambda, \theta)_{\rho_1}$$

for $A^{\rho_1}_{\lambda}$ with E leads to

$$E\big(S^{\rho_1} A^{\rho_1}_{\lambda}\big) E^{-1} = E(\lambda, \theta)_{\rho_1} E^{-1}$$

inserting $\mathbb{1} = E E^{-1}$ one obtains

$$\big(E S^{\rho_1} E^{-1}\big)\big(E A^{\rho_1}_{\lambda} E^{-1}\big) = E(\lambda, \theta)_{\rho_1} E^{-1}$$

and thus one gets

$$S^{\rho_2}\big(E A^{\rho_1}_{\lambda} E^{-1}\big) = (\lambda, \theta)_{\rho_2}.$$

Thus $E A^{\rho_1}_{\lambda} E^{-1}$ and $A^{\rho_2}_{\lambda}$ are solutions to the same equation and hence are equal. $\square$

We are now ready to deduce the generating functions. The formula for the generating functions and the proof are generalizations of ideas given in [18] for the scalar-valued multivariate Chebyshev polynomials to the matrix-valued case.

Theorem 4.15 *Let $\rho\colon W \longrightarrow \mathsf{GL}(V)$ denote the representation of a Weyl group. Denote by $W_r = \{w \in W \mid \rho(w) = r\}$ and by*

$$M_{r,k} = \operatorname{diag}(\exp(2\pi \mathrm{i} f_{1,k}(\theta)), \ldots, \exp(2\pi \mathrm{i} f_{|W_r|,k}(\theta))) \tag{4.59}$$

the diagonal matrix of action entries, i.e. $f_{i,k}(\theta)$ is the kth component of $w_i \in W_r$ acting on θ. Assume that S^ρ is invertible. The generating function for the Chebyshev polynomials associated to ρ is then given by

$$S_\rho(\theta)^{-1} F_\rho(p_1, \ldots, p_d) \tag{4.60}$$

where the nominator is

$$F_\rho(p_1, \ldots, p_d) = \sum_{r \in \rho(W)} r \cdot \mathsf{tr}(R_{r,p_1} \cdots R_{r,p_d}) \tag{4.61}$$

with

$$R_{r,p_i} = (\mathbb{1}_{|W_r|} - p_k M_{r,k})^{-1}. \tag{4.62}$$

Proof: The symmetrized pairing obeys

$$\begin{aligned}
(\lambda, \theta)_\rho &= \sum_{w \in W} \rho(w) \exp(2\pi \mathrm{i} \langle \lambda, w^\top \theta \rangle) \\
&= \sum_{r \in \rho(W)} r \cdot \sum_{w \in W_r} \exp(2\pi \mathrm{i} \langle \lambda, w^\top \theta \rangle \\
&= \sum_{r \in \rho(W)} r \sum_{k=1}^{|W_r|} \exp\left(2\pi \mathrm{i} \sum_{i=1}^{d} f_{k,i}(\theta) \lambda \right) \\
&= \sum_{r \in \rho(W)} r \, \mathsf{tr}(M_r)
\end{aligned}$$

with

$$M_r = \operatorname{diag}\left(\exp\left(2\pi \mathrm{i} \sum_{i=1}^{d} f_{1,i}(\theta) \lambda \right), \ldots, \left(2\pi \mathrm{i} \sum_{i=1}^{d} f_{|W_r|,i}(\theta) \lambda \right) \right).$$

Now the matrix M_r can be written as a product

$$M_r = \prod_{k=1}^{d} M_{r,k}^{\lambda_k}$$

with

$$M_{r,k} = \operatorname{diag}(\exp(2\pi \mathrm{i} f_{1,k}(\theta)), \ldots, \exp(2\pi \mathrm{i} f_{|W_r|,k}(\theta))).$$

Observe that

$$M_{r,k}^{\lambda_k} = \frac{1}{\lambda_k!} \frac{\mathrm{d}^{\lambda_k}}{\mathrm{d}p_k^{\lambda_k}} R_{r,p_k} \big|_{p_k = 0}.$$

Thus one obtains

$$(\lambda, \theta)_\rho = \frac{1}{\lambda!} \frac{\partial^\lambda}{\partial p_1^{\lambda_1} \cdots \partial p_d^{\lambda_d}} \sum_{r \in \rho(W)} r \cdot \mathsf{tr}(R_{r,p_1} \cdots R_{r,p_d})\Big|_{p_1 = \cdots = p_d = 0}.$$

Since $A_\lambda^\rho(\theta) = S^\rho(\theta)^{-1} \cdot (\lambda, \theta)_\rho$ the proof is finished. □

Remark 4.16 Observe that the generating functions derived using Theorem 4.15 are still in θ-coordinates. But it is, using computer algebra, much easier to convert the occurring generating functions to the x-coordinates than to find a closed-form of the Chebyshev polynomials in x-coordinates.

We give now two examples of the generating functions of matrix-valued Chebyshev polynomials associated to two-dimensional representations of Weyl groups. The calculations were performed in the computer algebra system Mathematica® and the representations where chosen so that Mathematica® could handle the occurring variable changes more easily. The Mathematica® notebooks used for the calculation are available online under the url https://github.com/bseifert-HSA/generating-functions-matrix-chebyshevs.

Example 4.17 A faithful two-dimensional representation of the Weyl group of the root system A_2 is given by the following six matrices

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \right\}. \tag{4.63}$$

There are two minimal admissible sets $\mathcal{A} = \{\{1\}, \{2\}\}$. Denote by

$$d_{A_2} = (-1 + p^3 + 3px - 3p^2 y) \cdot (-1 + q^3 - 3q^2 x + 3qy). \tag{4.64}$$

Using Theorem 4.15 we obtain the following entries of the matrix-valued generating function

$$\begin{aligned} F_{1,1}(p,q) &= \tfrac{2}{d_{A_2}} \cdot \big(q + 2q^2 + p(-4 + q(q - 3x - 6qx + 6y) \\ &\qquad + p^2(-2 + q(2 - 3qx + 6y)))\big), \\ F_{1,2}(p,q) &= \tfrac{2}{d_{A_2}} \cdot \big((-2q(1 + 2q) + p(2 + q(q + 3x + 12qx - 6y)) \\ &\qquad - p^2(1 + 2q)(-1 + 3qy))\big), \\ F_{2,1}(p,q) &= \tfrac{2}{d_{A_2}} \big(q(2 + q) + p^2(-4 + q - 6q^2 x + 12qy) \\ &\qquad + p(-2 + q(q(2 - 3x) - 6x + 3y))\big), \\ F_{2,2}(p,q) &= \tfrac{2}{d_{A_2}} \big(2q(2 + q) - p(1 + q(6x + q(2 + 6x) - 3y)) \\ &\qquad + p^2(2 + q)(-1 + 3qy)\big). \end{aligned} \tag{4.65}$$

Example 4.18 As before we calculate the generating function for a representation of the Chebyshev polynomials of type B_2. A faithful two-dimensional representation of the Weyl group of B_2 is given by the matrices

$$\begin{aligned}\Big\{&\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}-1&0\\0&-1\end{bmatrix},\begin{bmatrix}-1&-1\\0&1\end{bmatrix},\begin{bmatrix}-1&-1\\2&1\end{bmatrix},\\ &\begin{bmatrix}-1&0\\2&1\end{bmatrix},\begin{bmatrix}1&0\\-2&-1\end{bmatrix},\begin{bmatrix}1&1\\-2&-1\end{bmatrix},\begin{bmatrix}1&1\\0&-1\end{bmatrix}\Big\}.\end{aligned} \tag{4.66}$$

Denote by

$$d_{B_2}=((-1+p^2)^2+4p(1+p)^2x-16p^2y^2)\cdot\cdot(1+q(q(2+q^2+4x)-4(1+q^2)y)) \tag{4.67}$$

The generating function is given by

$$\begin{aligned}F_{1,1}(p,q)=\tfrac{8}{d_{B_2}}\big(&-q(1+q)^2+p(1+q)(2+q(-1+q+4x+4qx-4y))\\&+p^2\big(4+8y+q(1+4x+q(6+q+4(4+q)x)-12y\\&\quad-4q(-2+q-4x)y-16(2+q+q^2)y^2)\big)\\&-p^3(-2+q(1+8y+q(q+4qy-4(1+2x+y))))\big),\\F_{1,2}(p,q)=\tfrac{8}{d_{B_2}}\big(&p-q(1+q)^2+p^3(1+q(q+2q(2+q)x-4y))\\&+pq(q+4q(2+q)x+2(x-2y))\\&+p^2\big(2+4y+q(1+2x+q(4+q+6(2+q)x)\\&\quad-4y-4q(-1+q-2x)y-16(1+q+q^2)y^2)\big)\big),\\F_{2,1}(p,q)=\tfrac{16}{d_{B_2}}\big(&q(1+q+q^2)-p(1+q(1+4x+q(q+4x+4qx-4y)-2y))\\&+p^2\big(-2-8y+q(-1-q(3+q)-4(1+q)^2x\\&\quad+2(3+q(-4+q-8x))y+8(4+q+2q^2)y^2)\big)\\&+p^3(-1+q(1+4y+q(-2+q-4x+2(-2+q)y)))\big),\\F_{2,2}(p,q)=\tfrac{8}{d_{B_2}}\big(&-p+2q(1+q+q^2)-pq(q+8q(1+q)x+4(x-y))\\&-2p^2\big(1+4y+q(1+2x+q(2+q+6(1+q)x)\\&\quad-2y-2q(-2+q-4x)y-8(2+q+2q^2)y^2)\big)\\&-p^3(1+q(q+4q(1+q)x-4y))\big).\end{aligned} \tag{4.68}$$

Chapter 5

Conclusion and Future work

This thesis has made several contributions to the development of the algebraic signal processing theory in the case of polynomial algebras in several variables. First we connected the algebraic signal processing theory to algebraic geometry by interpreting the filter algebra as a space and the signal module as a vector bundle over this space. This enabled us to give sufficient conditions on ideals to induce well-defined Fourier transforms by the Chinese remainder theorem.

The second contribution was the connection between algebraic signal processing and the theory of multivariate orthogonal polynomials expanding prior work. This gave a generalization of the Gauß-Jacobi procedure for the derivation of orthogonal polynomials in the multivariate case. Unfortunately the multivariate Gauß-Jacobi procedure is tightly connected to the existence of Gaußian cubature formulae, which rarely exist.

We proceeded to generalize prior results on induction- and decomposition-based approaches to FFT-like algorithms in the multivariate case. The generalization of the induction-based approach allowed to include the fast Fourier transform algorithms on various directed lattices by Merseareau and Speake [75] into the algebraic theory. The generalization of the decomposition-based approach allowed for the inclusion of polynomial algebras with different-sized subalgebras.

We then investigated the connection to Lie theory. The generalizations of Chebyshev polynomials based on the reflection groups associated to the root systems were investigated. Since the multivariate Chebyshev polynomials are indexed by the elements of the weight lattices of semi-simple Lie algebras, they gave rise to signal models on these lattices. Furthermore they obey the decomposition property and thus their exist fast algorithms, for which we gave a geometric interpretation.

A further generalization of the multivariate Chebyshev polynomials leads to matrix-valued Chebyshev polynomials associated to representations of the reflections groups of the root lattices. We deduced a general scheme for the calculation of the generating functions of matrix-valued and multivariate Chebyshev polynomials and gave two examples.

Of course there are things open to investigate in the future. First and foremost there is still missing an existence proof of the matrix-valued Chebyshev polynomials 4.54. For this one has to show that the minimal function S_ρ does not vanish in the interior of the fundamental region. Daan Huybrechs suggested the following

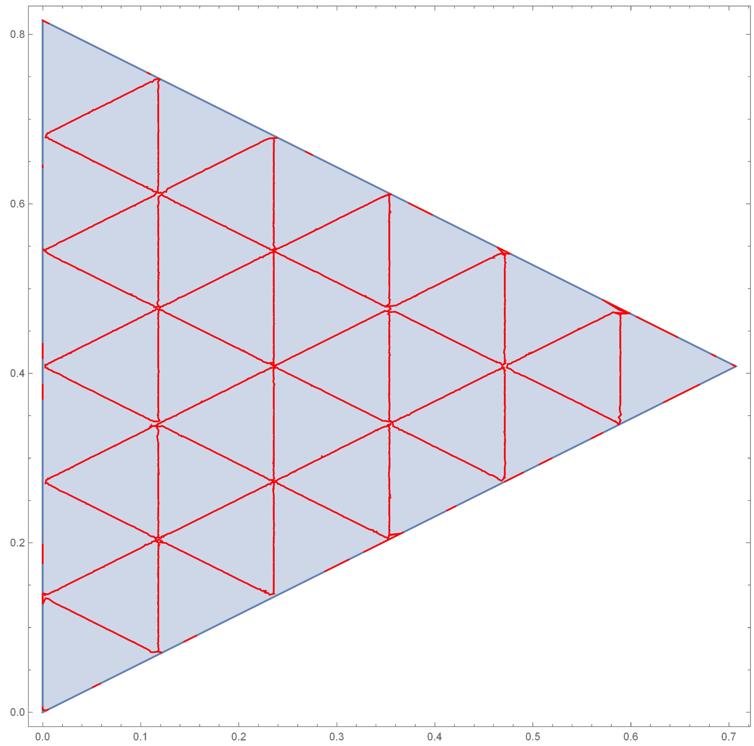

Figure 5.1: Nodal lines, i.e. set of points where the determinant vanishes, of the 36th matrix-valued eigenfunction with A_2-symmetries on the equilateral triangle.

approach in private communication. First one observes that S_ρ is element-wise an eigenfunction of the Laplacian on F, with certain jumping conditions on the boundary. In the case of $\rho = \det$ the boundary conditions are the homogeneous Dirichlet boundary conditions and in case of $\rho = \mathsf{id}$ one has Neumann boundary conditions. So in the one-dimensional case that S_ρ is invertible follows from the Courant nodal theorem [14]. Hence one possible approach to prove that S_ρ is invertible is to derive a theory of matrix element-wise eigenfunctions of the Laplacian with jumping boundary conditions and proof an analogue of the Courant nodal theorem in this setting. One might wonder why the theory of matrix-valued eigenfunctions of Laplacians has not been considered, yet. One possible explanation to this is that the proof of the Courant nodal line relies on a patching lemma, which states that the eigenfunctions of subdomains, bounded by the nodal domains of the nth eigenfunction, corresponding to the smallest eigenvalue is the nth eigenfunction of the original domain. Now in the matrix-valued setting this would imply that the nodal domains have the same symmetries as the original domain. Hence this restricts the possible domains for matrix-valued eigenfunctions with jumping boundary conditions to the foldable figures. First numerical experiments support this conjecture, cf. Fig. 5.1.

The second thing missing is a general method to deduce fast algorithms for the inverse Fourier transforms. In the one-variable case one can rely on the Gauß-Jacobi procedure if the signal model is based on orthogonal polynomials. Then one can define an orthogonal transform whose inverse can be calculated with the same algorithm as the original transform. In several variables this is only possible in certain special cases, as follows from Theorem 2.15. Hence a general method for the calculation of the inverse Fourier transforms is desirable.

Moreover the connection to algebraic geometry is still somewhat superficial and needs to be tightened. We conjecture that the usage of more advanced tools, e.g. cohomology, could be beneficial. Indeed in topological signal processing sheaf cohomology has found applications in sensor fusion [100] and general sampling theorems [99]. General sheaf theory has found applications in network analysis [30], as well.

Another idea from a current trend in algebraic geometry is tropicalization. In tropical geometry [70] one studies piecewise-linear shadows of algebraic curves by replacing the field $\mathbb{R}$ with the semi-field $(\mathbb{R} \cup \{-\infty\}, \max, +)$. This theory has applications in economics and dynamic programming and has recently found some applications in signal processing [26, 116]. Hence establishing a connection to algebraic signal processing theory seems to be reasonable goal.

Furthermore there exist non-equispaced, i.e., on arbitrary grids, Fourier transforms on the circle, the sphere and the rotation group [86,120]. These non-equispaced Fourier transforms can be used to deduce fast global optimization methods on these manifolds [25]. Since the multivariate Chebyshev polynomials have strong connections to the representation theory of Lie groups it is natural to wonder if one can generalize these methods to compact Lie groups. One step into this direction would be to translate the interpolation used to calculate the non-equispaced Fourier transform into the algebraic setting. A possible tool to implement this could be deformation theory.

Finally there are of course other algebras than polynomial algebras, which could be of interest. The first class to consider are algebras associated to graphs, i.e. the ones generated by the adjacency matrix of a graph. This is one of the approaches to graph signal processing [106, 107], which has recently found many applications in the applied sciences. Another class of algebras to be considered are algebras associated to lattices of partially ordered sets or hypergraphs. First steps in this directions can be found in [87, 88].

Appendix A

Gröbner bases

This appendix summarizes some results on Gröbner bases. See [15] for more details and proofs of the results, as well as references to the literature.

Let $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ be a multi-index. Any polynomial $p \in \mathbb{C}[x_1, \ldots, x_d]$ can be written as

$$p = \sum_\alpha c_\alpha x^\alpha = \sum_\alpha x_1^{\alpha_1} \ldots x_d^{\alpha_d}, \tag{A.1}$$

with $c_\alpha \in \mathbb{C}$ and α ranging over a finite subset of $\mathbb{N}_0^d$. The c_α are called the coefficients of p and the $c_\alpha x^\alpha$ are called the terms of p.

The ideal generated by $p_1, \ldots, p_s \in \mathbb{C}[x_1, \ldots, x_d]$ is

$$\langle\, p_1, \ldots, p_s \,\rangle = \left\{ \sum_i h_i p_i \mid h_i \in \mathbb{C}[x_1, \ldots, x_d] \right\}. \tag{A.2}$$

The polynomials $p_1, \ldots, p_s$ are called a basis of the ideal $\langle\, p_1, \ldots, p_s \,\rangle$. The Hilbert's basis theorem, cf. [15, Ch. 2, §5, Thm. 4], states that every ideal of $\mathbb{C}[x_1, \ldots, x_d]$ is finitely generated, i.e. of the form A.2. Note that the basis of an ideal is not uniquely determined.

One is interested in special bases of an ideal, which possess good computational properties. These special bases are the Gröbner bases. For their definition, we need to recall the notion of a monomial ordering. First one observes that there is a one-to-one correspondence between $\mathbb{N}_0^d$ and the monomials of $\mathbb{C}[x_1, \ldots, x_d]$ via

$$(\alpha_1, \ldots, \alpha_d) \leftrightarrow x_1^{\alpha_1} \cdots x_d^{\alpha_d}. \tag{A.3}$$

Hence one relies on the definition of an ordering on $\mathbb{N}_0^d$ to define orderings on the monomials, by setting $x^\alpha > x^\beta$ if $\alpha > \beta$.

Definition A.1 (Monomial ordering) *A monomial ordering on $\mathbb{C}[x_1, \ldots, x_d]$ is a relation $>$ on $\mathbb{N}_0^d$ which satisfies*

i.) the relation $>$ is a total order on $\mathbb{N}_0^d$, i.e. exactly one of the statements

$$\alpha > \beta, \quad \alpha = \beta, \quad \beta > \alpha \tag{A.4}$$

is true for any $\alpha, \beta \in \mathbb{N}_0^d$, *ii.)* *if* $\alpha > \beta$ *and* $\gamma \in \mathbb{N}_0^d$ *then* $\alpha + \gamma > \beta + \gamma$,

iii.) *the relation* $>$ *is a well-ordering of* $\mathbb{N}_0^d$, *i.e. every non-empty subset of* $\mathbb{N}_0^d$ *contains a smallest element.*

Definition A.2 *Let* $p = \sum_\alpha c_\alpha x^\alpha \neq 0 \in \mathbb{C}[x_1, \ldots, x_d]$ *and* $>$ *be a monomial order, then*

i.) *the multi-degree of the polynomial* p *is*

$$\text{multideg}(p) = \max\{\alpha \in \mathbb{N}_0^d \mid c_\alpha \neq 0\}, \tag{A.5}$$

ii.) *the leading coefficient of* p *is*

$$\text{LC}(p) = c_{\text{multideg}(p)}, \tag{A.6}$$

iii.) *the leading monomial of* p *is*

$$\text{LM}(p) = x^{\text{multideg}(p)}, \tag{A.7}$$

iv.) *the leading term of* p *is*

$$\text{LT}(p) = \text{LC}(p) \cdot \text{LM}(p). \tag{A.8}$$

Definition A.3 *Let* $I \subseteq \mathbb{C}[x_1, \ldots, x_d]$ *be a non-zero ideal and let* $>$ *be a monomial oder, then*

i.) *the leading terms of the ideal* I *are*

$$\text{LT}(I) = \{cx^\alpha \mid \text{ there exists } p \in I \setminus \{0\} \text{ such that } \text{LT}\, p = cx^\alpha\}, \tag{A.9}$$

ii.) *the ideal of leadings terms is* $\langle\, \text{LT}\, I \,\rangle$.

Note that in general $\langle\, \text{LT}\, p_1, \ldots, \text{LT}\, p_s \,\rangle \neq \langle\, \text{LT}\, \langle\, p_1, \ldots, p_s \,\rangle\, \rangle$. The bases for an ideal for which equality holds are the Gröbner bases.

Definition A.4 *A subset* $G = \{g_1, \ldots, g_s\} \subseteq I$ *such that*

$$\langle\, \text{LT}\, g_1, \ldots, \text{LT}\, g_s \,\rangle = \langle\, \text{LT}\, I \,\rangle \tag{A.10}$$

is called a Gröbner basis.

Proposition A.5 *Every ideal has a Gröbner basis. Any Gröbner basis of an ideal is a basis for the ideal.*

What is so remarkable about Gröbner bases is that the remainder of division by them is unique. For the division algorithm in several variables see Algorithm 1.

Proposition A.6 *Let* $I \subseteq \mathbb{C}[x_1, \ldots, x_d]$ *be an ideal and* $G = \{g_1, \ldots, g_s\}$ *be a Gröbner basis for* I. *If* $p \in \mathbb{C}[x_1, \ldots, x_d]$ *then there is a unique* $r \in \mathbb{C}[x_1, \ldots, x_d]$ *such that*

i.) *No term of* r *is divided by any of the* $\text{LT}\, g_1, \ldots, \text{LT}\, g_s$,

ii.) *there is a* $g \in I$ *such that* $p = g + r$.

Algorithm 1 The division algorithm for polynomials in several variables.

function DIVISION($G = \{g_1, \dots, g_s\}, p$)
    $q_i = 0, r = 0, tmp = p$
    while $tmp \neq 0$ **do**
        $i = 1$
        division_occured = false
        while $i < s$ and division_occured = false **do**
            if $\mathrm{LT}(g_i)$ divides $\mathrm{LT}(tmp)$ **then**
                $q_i = q_i + \frac{\mathrm{LT}(tmp)}{\mathrm{LT}(g_i)}$
                $tmp = tmp - \frac{\mathrm{LT}(tmp)}{\mathrm{LT}(g_i)} \cdot g_i$
                division_occured = true
            else
                $i = i + 1$
            end if
        end while
        if division_occured = false **then**
            $r = r + \mathrm{LT}(tmp)$
            $tmp = tmp - \mathrm{LT}(tmp)$
        end if
    end while
    return $q_1, \dots, q_s, r$ such that $p = q_1 g_1 + \dots q_s g_s + r$
end function

Note that the quotients q_i from Algorithm 1 change if one changes the implicit chosen ordering of the g_i.

An easy criterion for deciding, if a basis of an ideal is a Gröbner basis, is the Buchberger criterion, which relies on the notion of S-polynomials.

Definition A.7 *Let* $f, g \in \mathbb{C}[x_1, \dots, x_d] \setminus \{0\}$. *Let* $\alpha = \mathrm{multideg}(f)$ *and* $\beta = \mathrm{multideg}(g)$. *The least common multiple of* $\mathrm{LM}(f)$ *and* $\mathrm{LM}(g)$ *is* x^γ *with* $\gamma_i = \max(\alpha_i, \beta_i)$. *The* S*-polynomial of* f *and* g *is*

$$S(f,g) = \frac{x^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{x^\gamma}{\mathrm{LT}(g)} \cdot g. \tag{A.11}$$

Theorem A.8 (Buchberger criterion) *Let* I *be an ideal. A basis* $G = \{g_1, \dots, g_s$ *of* I *is a Gröbner basis of* I *if and only if for all* $i \neq j$ *the remainder of division of* $S(g_i, g_j)$ *by* G *is zero.*

The Buchberger criterion is for example fulfilled if the leading monomials of the g_i are disjoint, i.e. if $\mathrm{LM}(g_i) = x^{\alpha_i}$ then $a_{i,k} \neq 0$ if and only if $a_{j,k} = 0$ for all $j \neq i$.

In this thesis we use the fact that Gröbner bases solve the ideal membership problem. That is $f \in I = \langle g_1, \dots, g_s \rangle$, for $G = \{g_1, \dots, g_s\}$ a Gröbner basis, if and only if the remainder of division of f by G is zero.

Appendix B

Chinese remainder theorem

This appendix recalls the Chinese remainder theorem for rings and modules with a special focus on polynomial algebras. We will especially focus on the explicit form for polynomials in several variables.

Theorem B.1 (Chinese remainder theorem) *Let R be a commutative ring and $I = I_1 \cap \cdots \cap I_n$ be the intersection of pairwise comaximal ideals. Then*

$$R/I \cong R/I_1 \times \cdots \times R/I_n. \tag{B.1}$$

If M is an R-module then

$$M/MI \cong M/MI_1 \times \cdots \times M/MI_n. \tag{B.2}$$

PROOF: We consider $n = 2$, the general fact follows by induction. We start by showing that for comaximal ideals I_1, I_2 one has $I_1 I_2 = I_1 \cap I_2$. For this observe that any $x \in I_1 I_2$ is of the form $\sum_{i_1 \in I_1, i_2 \in I_2} i_1 i_2$. But then $x \in I_1$ since the i_1 are and $x \in I_2$ since the i_2 are and the R is commutative. Hence $x \in I \cap J$ so $IJ \subseteq I \cap J$. Since I_1, I_2 are comaximal one has $R = I_1 + I_2$. Now observe that

$$I_1 \cap I_2 = R(I_1 \cap I_2) = (I_1 + I_2)(I_1 \cap I_2) = I_1(I_1 \cap I_2) + I_2(I_1 \cap I_2) \subseteq I_1 I_2 + I_2 I_1 = I_1 I_2$$

by commutativity. Hence $I_1 I_2 = I_1 \cap I_2$ follows.

Now define the homomorphism

$$\phi\colon R \longrightarrow R/I_1 \times R/I_2$$

by $\phi(r) = (r + I_1, r + I_2)$. For (B.1) to hold one has to show that the sequence

$$0 \longrightarrow I_1 \cap I_2 \longrightarrow R \xrightarrow{\phi} R/I_1 \times R/I_2 \longrightarrow 0$$

is exact. That is, one has to show that ϕ is surjective. We have to find to each $(r_1 + I_1, r_2 + I_2)$ an $x \in R$ such that $\phi(x) = (r_1 + I_1, r_2 + I_2)$. Since $R = I_1 + I_2$ one can write $r_1 = i_{1,1} + i_{2,1}$ and $r_2 = i_{1,2} + i_{2,2}$ with $i_{1,1}, i_{1,2} \in I_1$ and $i_{2,1}, i_{2,2} \in I_2$. Now let $x = i_{1,2} + i_{2,1}$. Then $x - r_1 = i_{1,2} - i_{1,1} \in I_1$ so $x + I_1 = r_1 + I_1$ and $x - r_2 = i_{2,1} - i_{2,2} \in I_2$ so $x + I_2 = r_2 + I_2$. Hence $\phi(x) = (r_1 + I_1, r_2 + I_2)$ and surjectivity of ϕ is proven. Thus the Chinese remainder theorem for commutative rings holds.

The Chinese remainder theorem for modules (B.2) is now just a simple consequence of its version for rings since one has the canonical isomorphism $R/I \otimes A \cong A/AI$ and tensor products commute with finite direct products. $\square$

This general form of the Chinese remainder theorem does not tell how to explicitly calculate the isomorphism. As we are mainly interested in polynomial algebras, we specialize the isomorphism from the theorem to the case of polynomial algebras. We investigate first the complete decomposition and then what happens if one has a decomposition.

Proposition B.2 *Consider Π^n/I for a zero-dimensional radical ideal I. Then $I = \prod_{\alpha \in \mathsf{V}(I)} \langle\, x_1 - \alpha_1, \ldots, x_n - \alpha_n \,\rangle$ with all $\langle\, x_1 - \alpha_1, \ldots, x_n - \alpha_n \,\rangle$ comaximal. Then the isomorphism from the Chinese remainder theorem is realized by the evaluation homomorphism $p \mapsto (p(\alpha))_{\alpha \in \mathsf{V}(I)}$.*

Proof: We have to show that $p - p(\alpha) \in \langle\, x_1 - \alpha_1, \ldots, x_n - \alpha_n \,\rangle$. Denote by $g(x_1, \ldots, x_n) = p(x_1, \ldots, x_n) - p(\alpha)$. Then $g(\alpha) = 0$. A Gröbner basis for $\langle\, x_1 - \alpha_1, \ldots, x_n - \alpha_n \,\rangle$ is $\{x_1 - \alpha_1, \ldots, x_n - \alpha_n\}$. Now g can be written as $g = p_1(x_1 - \alpha_1) + \cdots + p_n(x_n - \alpha_n) + r$ for polynomials p_i, where none of the $x_i - \alpha_i$ divides r, so r is constant. Since $g(\alpha) = p_1 \cdot 0 + \cdots + p_n \cdot 0 + r = 0$ it follows that $r = 0$ and thus $p - p(\alpha) \in \langle\, x_1 - \alpha_1, \ldots, x_n - \alpha_n \,\rangle$. $\square$

Next we show how the Chinese remainder theorem is realized if one uses only a partial composition. This version is used in Sect. 3.2.

Proposition B.3 *Consider the algebra Π^n/I for a zero-dimensional radical ideal $I = \langle\, p_1, \ldots, p_n \,\rangle$ which obeys the decomposition property*

$$\langle\, p_1, \ldots, p_n \,\rangle = \langle\, q_1(r_1, \ldots, r_n), \ldots, q_n(r_1, \ldots, r_n) \,\rangle. \tag{B.3}$$

Assume the $r_i - \alpha_i$ form Gröbner bases for the ideals they generate. Then the ideals $\langle\, r_1 - \alpha_1, \ldots, r_n - \alpha_n \,\rangle$ are comaximal. The isomorphism from the Chinese remainder theorem is realized by replacing each occurrence of r_i by α_i.

Proof: The comaximality of the ideals $\langle\, r_1 - \alpha_1, \ldots, r_n - \alpha_n \,\rangle$ follows from the radicality of I. Denote by $p(r_i :> \alpha_i)$ the polynomial were every occurrence of r_i has been replaced by α_i. We show that $g = p - p(r_i :> \alpha_i) \in \langle\, r_1 - \alpha_1, \ldots, r_n - \alpha_n \,\rangle$. Observe that due to the Gröbner basis property we can write $g = p_1(r_1 - \alpha_1) + \cdots + p_n(r_n - \alpha_n) + r$, where r is polynomial not divided by any of the $r_i - \alpha_i$. Now observe again that $g(r_i :> \alpha_i) = 0$ and thus $r = 0$. Hence $p - p(r_i :> \alpha_i) \in \langle\, r_1 - \alpha_1, \ldots, r_n - \alpha_n \,\rangle$. $\square$

Appendix C

The Akra-Bazzi theorem and computational costs of matrices

The Akra-Bazzi theorem [1] is a method to determine the asymptotic behaviour of algorithms, whose asymptotic complexity can be described using recurrence equations. In this appendix we investigate the application of the Akra-Bazzi theorem on matrix factorizations, i.e. we state what we mean by the computational cost of a matrix factorization and how to determine the asymptotic computationally cost if one has a recurrence formula for the factorization.

We use the following two Landau symbols. The first one is the asymptotic upper bound which tells that a function f grows at most as fast as another function g, it is denoted by $f \in O(g)$ and defined as $\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty$. The second one is the asymptotic sharp bound which tells that a function f is bounded from above and from below by a function g, it is denoted by $f \in \Theta(g)$ and defined as $0 < \liminf_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| \leq \limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty$. The asymptotic sharp bound implies $f \in O(g)$ as well as $g \in O(f)$.

The Akra-Bazzi theorem is now as follows.

Theorem C.1 (Akra-Bazzi) *Let $T\colon \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ be a function which obeys the recurrence relation*

$$T(x) = g(x) + \sum_{i=1}^{k} a_i T(b_i x + h_i(x)) \tag{C.1}$$

for all $x \geq x_0$, where x_0 is some constant, such that

- *the recurrence* (C.1) *can be resolved uniquely,*
- *$a_i > 0$ and $b_i \in [0,1]$ are constants,*
- *$|g'(x)| \in O(x^c)$ for some constant c,*
- *$|h_i(x)| \in O\Big(\frac{x}{(\log x)^2}\Big)$.*

Then one has

$$T(x) \in \Theta\bigg(x^p\bigg(1 + \int_1^x \tfrac{g(x)}{u^{p+1}}\,\mathrm{d}u\bigg)\bigg), \tag{C.2}$$

with $p \in \mathbb{R}^+$ *such that* $\sum_{i=1}^{k} a_i b_i^p = 1$.

Now for its applications to matrix factorizations consider the following. The *computational cost* of a matrix is defined to be the number of complex additions and complex multiplications to form the matrix-vector product of the matrix with an arbitrary vector. That is, for an $n \times n$ matrix one has for each row $n-k$ complex multiplications, where k is the number of entries being 0 or 1, and $n-1-j$ additions, where j is the number of entries being 0. The overall computational cost of a matrix are considered as the sum of the number of complex additions and multiplications. So an $n \times n$ matrix where all entries are non-zero is of computational cost $O(n^2)$.

As an example consider a recursive matrix factorization, where in each step the $n \times n$ matrix factors into a product of matrices of computational cost $O(n)$ and a block diagonal matrix with blocks of size $n/2 \times n/2$. E.g. consider $P_1, P_2, B, M \in \mathbb{C}^{n \times n}$ such that

$$M = P_1 \cdot B \cdot P_2 \tag{C.3}$$

with $P_1, P_2 \in O(n)$ and B a block diagonal matrix with blocks of size $n/2 \times n/2$. Then one obtains a recurrence equation for the computational cost of the form

$$T(n) = T(n/2) + T(n/2) + c \cdot n, \tag{C.4}$$

where c is the number of matrices of cost $O(n)$. Since $1/2 + 1/2 = 1$ we obtain from the Akra-Brazzi Theorem C.1 that

$$T(n) \in \Theta(n \cdot \log n), \tag{C.5}$$

since one has $\int_1^n \frac{u}{u^2} \,\mathrm{d}u = \log n$.

Bibliography

[1] Akra, M., Bazzi, L.: *On the solution of linear recurrence equations.* Computational Optimization and Applications **10**.2 (1998), 195–210. 37, 89

[2] Beerends, R.: *Chebyshev polynomials in several variables and the radial part of the Laplace-Beltrami operator.* Trans. Amer. Math. Soc. **328** (1991), 779–814. 14

[3] Bentley, J. L., Haken, D., Saxe, J. B.: *A general method for solving divide-and-conquer recurrences.* ACM SIGACT News **12**.3 (1980), 36–44. 37

[4] Beth, T.: *Verfahren der schnellen Fouriertransformationen.* Teubner Verlag, 1984. 12

[5] Borzov, V. V., Damaskinsky, E. V.: *Chebyshev-Koornwinder oscillator.* Theor. Math. Phys. **175**.3 (2013), 763–770. 14

[6] Borzov, V. V., Damaskinsky, E. V.: *The algebra of two dimensional generalized Chebyshev-Koornwinder oscillator.* J. Math. Phys. **55**.10 (2014), 103505. 14

[7] Bracewell, R. N.: *Discrete Hartley transform.* J. Opt. Soc. Am. **73**.12 (1983), 1832–1835. 25

[8] Breske, S., Labs, O., van Straten, D.: *Real Line Arrangements and Surfaces with Many Real Nodes.* In: Jüttler, B., Piene, R. (eds.): *Geometric Modeling and Algebraic Geometry*, 47–54. Springer, 2008. 14

[9] Britanak, V., Rao, K. R.: *The fast generalized discrete Fourier transforms: A unified approach to the discrete sinusoidal transforms computation.* Signal Processing **79** (1999), 135–150. 13

[10] Chadzitaskos, G., Hakova, L., Kajinek, O.: *Weyl Group Orbit Functions in Image Processing.* Applied Mathematics **5**.3 (2014), 501–511. 15

[11] Chevalley, C.: *Invariants of Finite Groups Generated by Reflections.* Amer. J. Math **77**.4 (1955), 778–782. 58, 74

[12] Conway, J., Sloane, N.: *Sphere Packings, Lattices and Groups.* Springer, third. edition, 1999. 13, 52

[13] Cooley, J. W., Tukey, J. W.: *An algorithm for the machine calculation of complex Fourier series.* Math. Comput. **19** (1965), 297–301. 12

[14] Courant, R.: *Ein allgemeiner Satz zur Theorie der Eigenfunktionen selbstadjungierter Differentialausdrücke.* Nachr. Ges. Wiss. Göttingen Math. Phys. Kl. (1923), 81–84. 80

[15] Cox, D. A., Little, J., O'Shea, D.: *Ideals, Varieties, and Algorithms.* Springer, fourth. edition, 2015. 17, 83

[16] Czyzycki, T., Hrivnák, J.: *Generalized discrete orbit function transforms of affine Weyl groups.* J. Math. Phys. **55** (2014), 113508. 15

[17] Czyzycki, T., Hrivnák, J., Patera, J.: *Generating Functions for Orthogonal Polynomials of A_2, C_2 and G_2.* Symmetry **10** (2018), 354. 14

[18] Damaskinsky, E. V., Kulish, P. P., Sokolov, M. A.: *On calculation of generating functions of Chebyshev polynomials in several variables.* J. Math. Phys. **56** (2015), 063507. 14, 52, 76

[19] Damaskinsky, E. V., Sokolov, M. A.: *The generating function of bivariate Chebyshev polynomials associated with the Lie algebra G_2.* Theor. Math. Phys. **192**.2 (2017), 1115–1128. 14

[20] Diaconis, P., Rockmore, D.: *Efficient computation of the Fourier transfrom on finite groups.* Amer. Math. Soc. **3**.2 (1990), 297–332. 12

[21] Eier, R., Lidl, R.: *A Class of Orthogonal Polynomials in k Variables.* Math. Ann. **260** (1982), 93–99. 14

[22] Ersoy, O.: *Real Discrete Fourier Transform.* IEEE Trans. Acoust., Speech, Signal Process. **33**.4 (1985), 880–882. 25

[23] Favard, J.: *Sur les polynomes de Tchebicheff.* C. R. Acad. Sci. Paris **200** (1935), 2052–2053. 28

[24] Gauss, C. F.: *Theoria interpolationis methodo nova tractata.* In: *Werke, Band 3.* Königliche Gesellschaft der Wissenschaften, 1866. posthum. 12

[25] Gräf, M., Hielscher, R.: *Fast Global Optimization on the Torus, the Sphere, and the Rotation Group.* SIAM J. Optim. **25**.1 (2015), 540–563. 81

[26] Gripon, V.: *Tropical graph signal processing.* In: *51st Asilomar Conference on Signals, Systems, and Computers*, 50–54, Oct 2017. 81

[27] Grünbaum, F. A., Pacharoni, I., Tirao, J.: *Matrix-valued spherical functions associated to the complex projective plane.* J. Funct. Anal. **188** (2002), 350–441. 15

[28] Hakova, L., Hrivnák, J., Motlochová, L.: *On cubature rules associated to Weyl group orbit functions.* Acta Polytechnica **56**.3 (2016), 202–213. 14

[29] HAKOVA, L., HRIVNÁK, J., PATERA, J.: *Four families of Weyl group orbit functions of B_3 and C_3.* J. Math. Phys. **54**.8 (2013), 083501. 14

[30] HANSEN, J., GHRIST, R.: *Learning Sheaf Laplacians from Smooth Signals.* In: *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5446–5450, May 2019. 81

[31] HECKMAN, G. J.: *Root systems and hypergeometric functions II.* Comp. Math. **64** (1987), 353–373. 14

[32] HECKMAN, G. J., OPDAM, E. M.: *Root systems and hypergeometric functions I.* Comp. Math. **64** (1987), 329–352. 14

[33] HEIDEMAN, M., JOHNSON, D., BURRUS, C.: *Gauss and the history of the fast fourier transform.* IEEE ASSP Magazine **1**.4 (1984), 14–21. 12

[34] HEIDEMAN, M. T., BURRUS, C. S.: *On the number of multiplications necessary to compute a length-2^n DFT.* IEEE Trans. Acoust., Speech, Signal Proc. **34**.1 (1986), 91–95. 12

[35] HILBERT, D.: *Über die vollen Invariantensysteme.* Math. Ann. **42** (1893), 313–337. 11

[36] HOFFMAN, M. E., WITHERS, W. D.: *Generalized Chebyshev Polynomials associated with affine Weyl groups.* Trans. Amer. Math. Soc. **308** (1988), 91–104. 13, 14, 51, 52, 53, 56

[37] HOFFMAN, M. E., WITHERS, W. D.: *Linear Characters of Weyl Groups and Associated Multivariable Orthogonal Polynomials.* unpublished, thankfully provided by Hans Munthe-Kaas, 1988. 15, 73

[38] HRIVNÁK, J., JURÁNEK, M.: *On E-discretization of tori of compact simple Lie groups. II.* J. Math. Phys. **58**.10 (2017), 103504. 14

[39] HRIVNÁK, J., MOTLOCHOVÁ, L.: *Discrete cosine and sine transforms generalized to honeycomb lattice.* J. Math. Phy. **59**.6 (2018), 063503. 15

[40] HRIVNÁK, J., MOTLOCHOVÁ, L., PATERA, J.: *On Discretization of Tori of Compact Simple Lie Groups II.* J. Phys. A **45** (2012), 255201. 14

[41] HRIVNÁK, J., MOTLOCHOVÁ, L., PATERA, J.: *Cubature Formulas of Multivariate Polynomials Arising from Symmetric Orbit Functions.* Symmetry **8**.7 (2016), 63. 14

[42] HRIVNÁK, J., PATERA, J.: *On discretization of tori of compact simple Lie groups.* J. Phy. A.: Math. Theor. **42**.38 (2009), 385208. 14

[43] HUMPHREYS, J. E.: *Reflection Groups and Coxeter Groups.* Cambridge University Press, 1990. 53

[44] HUYBRECHS, D., MUNTHE-KAAS, H.: *Matrix-valued and multivariate orthogonal polynomials of Chebyshev type.* unpublished, thankfully provided by Hans Munthe-Kaas, 2013. 14, 15, 52, 73

[45] Inc., W. R.: *Mathematica, Version 12.0.* Champaign, IL, 2019. 62

[46] Johnson, H. W., Burrus, C. S.: *On the structure of efficient DFT algorithms.* IEEE Trans. Acoust., Speech, Signal Proc. **34**.1 (1985), 248–254. 12

[47] Killing, W.: *Die Zusammensetzung der stetigen endlichen Transformationsgruppen, I.* Math. Ann. **31** (1888), 252–290. 53

[48] Klimyk, A., Patera, J.: *Orbit Functions.* SIGMA **2** (2006). 14

[49] Koelink, E., van Pruijssen, M., Román, P.: *Matrix-valued orthogonal polynomials related to* $(SU(2)\times SU(2), diag)$. Int. Math. Res. Not. **2012** (2012), 5673–5730. 15

[50] Koelink, E., van Pruijssen, M., Román, P.: *Matrix-valued orthogonal polynomials related to* $(SU(2)\times SU(2), diag)$ *II.* Publ. RIMS Kyoto **49** (2013), 271–312. 15

[51] Koelink, E., van Pruijssen, M., Román, P.: *Matrix elements of irreducible representations of* $SU(n+1)\times SU(n+1)$ *and multivariable matrix valued orthogonal polynomials.* arXiv:1706.01927, 2017. 15

[52] Koornwinder, T.: *Two-variable analogues of the classical orthogonal polynomials.* In: Askey, R. A. (eds.): *Theory and Applications of Special Functions.* Academic Press, 1975. 14, 31

[53] Koornwinder, T.: *Matrix elements of irreducible representations of* $SU(2)\times SU(2)$ *and vector-valued orthogonal polynomials.* SIAM J. Math. Anal. **16** (1985), 602–613. 15

[54] Koornwinder, T. H.: *Orthogonal polynomials in two variables which are eigenfunctions of two algebraically independent partial differential operators. I - IV.* Indagationes Mathematicae (Proceedings) **77**.1-4 (1974), 48–58, 59–66, 357–369, 370–381. 14

[55] Kostant, B.: *Lie group representations on polynomial rings.* Amer. J. Math. **85** (1963), 327–404. 74

[56] Kowalski, M. A.: *Orthogonality and recursion formulas for polynomials in n variables.* SIAM J. Math. Anal. **13** (1982), 316–323. 28

[57] Kowalski, M. A.: *The recursion formulas for orthogonal polynomials in n variables.* SIAM J. Math. Anal. **13** (1982), 309–315. 28

[58] Kowalski, M. A.: *Algebraic characterization of orthogonality in the space of polynomials.* Lecture Notes in Math. **1171** (1985), 101–110. 28

[59] Lang, S.: *Algebra.* Springer, 2002. 17

[60] Li, H., Sun, J., Xu, Y.: *Discrete Fourier Analysis, Cubature, and Interpolation on a Hexagon and a Triangle.* SIAM J. Numer. Anal. **46**.4 (2008), 1653–1681. 15, 60

[61] Li, H., Sun, J., Xu, Y.: *Discrete Fourier Analysis and Chebyshev Polynomials with G_2 Group.* SIGMA **8** (2012), 29 pp. 15

[62] Li, H., Xu, Y.: *Discrete Fourier analysis on a dodecahedron and a tetrahedron.* Math. Comp. **78** (2009), 999–1029. 15

[63] Li, H., Xu, Y.: *Discrete Fourier Analysis on Fundamental Domain and Simplex of A_d Lattice in d-Variables.* J. Fourier Anal. Appl. **16**.3 (2010), 383–433. 15, 61

[64] Lidl, R.: *Tschebyscheffpolynome und die dadurch dargestellten Gruppen.* Monatsh. Math. **77** (1973), 132–147. 14

[65] Lidl, R.: *Tschebyscheffpolynome in mehreren Variablen.* J. reine angew. Math. **273** (1975), 178–198. 14

[66] Lidl, R., Wells, C.: *Chebyshev polynomials in several variables.* J. reine angew. Math. **255** (1972), 104–11. 14

[67] Lima, J. B., Campello de Souza, R. M.: *Finite filed trigonometric transforms.* Applicable Algebra in Engineering, Communication and Computing **23** (2011), 393–411. 18

[68] Lyakhovsky, V. D.: *Multivariate Chebyshev polynomials in terms of singular elements.* Theor. Math. Phys. **175**.3 (2013), 797–805. 14

[69] Lyakhovsky, V. D., Uvarov, P. V.: *Multivariate Chebyshev polynomials.* J. Phy. A.: Math. Theor. **46** (2013), 125201. 14

[70] Maclagan, D., Sturmfels, B.: *Introduction to Tropical Geometry.* American Mathematical Society, 2015. 81

[71] Maslen, D., Rockmore, D. N., Wolff, S.: *The efficient computation of Fourier transforms of semisimple algebras.* J. Fourier Anal. Appl. (2017). 12

[72] Maslen, D. K.: *Efficient Computation of Fourier Transforms on Compact Groups.* J. Fourier Anal. Appl. **4**.1 (1998), 19–52. 12

[73] Meng, T., Entezari, A., Smith, B., Möller, T., Weiskopf, D., Kirkpatrick, A.: *Visual comparability of 3D regular sampling and reconstruction.* IEEE Trans. Vis. Comput. Graphics **17**.10 (2011), 1420–1432. 52

[74] Mersereau, R. M.: *The processing of hexagonally sampled two-dimensional signals.* Proc. IEEE **67** (1979), 930–949. 41

[75] Mersereau, R. M., Speake, T. C.: *A Unified Treatment of Cooley-Tukey Algorithms for the Evaluation of the Multidimensional DFT.* IEEE Trans. Acoustics, Speech, and Signal Processing **29**.5 (1981), 1011–1018. 40, 79

[76] Moody, R. V., Patera, J.: *Cubature formulae for orthogonal polynomials in terms of elements of finite order of compact simple Lie groups.* Adv. Appl. Math. **47**.3 (2011), 509–535. 14, 51, 58, 60, 61

[77] MUNTHE-KAAS, H.: *On group Fourier analysis and symmetry preserving discretizations of PDEs.* J. Phys. A: Math. Gen. **39** (2006), 5563–5584. 15, 58

[78] MUNTHE-KAAS, H., NOME, M., RYLAND, B. N.: *Through the Kaleidoscope: Symmetries, Groups and Chebyshev-Approximations from a Computational Point of View.* In: CUCKER, F., KRICK, T., PINKUS, A., SZANTO, A. (EDS.): *Foundations of Computational Mathematics, Budapest 2011*, 188–229. Cambridge Universtiy Press, 2012. 15

[79] MUNTHE-KAAS, H., SOREVIK, T.: *Multidimensional pseudo-spectral methods on lattice grids.* Applied Numerical Mathematics **62** (2012), 155–165. 15

[80] MYSOVSKIKH, I. P.: *Numerical characteristics of orthogonal polynomials in two variables.* Vestnik Leningrad Univ. Math. **3** (1976), 323–332. 31

[81] NAKANE, S.: *External rays for polynomial maps of two variables associated with Chebyshev maps.* J. Math. Anal. Appl. **338** (2008), 552–562. 14

[82] NESTERENKO, M., PATERA, J., TERESZKIEWICZ, A.: *Orbit functions of $SU(n)$ and Chebyshev polynomials.* In: *Proceedings of the 5th Workshop "Group Analysis of Differential Equations & Integrable Systems"*, 133–151, 2010. 14

[83] NICHOLSON, P. J.: *Algebraic theory of finite Fourier transforms.* Journal of Computer and System Sciences **5** (1971), 524–547. 11, 12

[84] NUSSBAUMER, H. J.: *Fast Fourier Transformation and Convolution Algorithms.* Springer, 1982. 12

[85] PETERSON, D. P., MIDDLETON, D.: *Sampling and reconstruction of waenumber-limited functions in N-dimensional Euclidean spaces.* Information and Control **5**.4 (1962), 279–323. 13, 52

[86] POTTS, D., PRESTIN, J., VOLLRATH, A.: *A fast algorithm for nonequispaced Fourier transforms on the rotation group.* A. Numer Algor **52** (2009), 355–384. 81

[87] PÜSCHEL, M.: *A discrete signal processing framework for set functions.* In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 4359–4363, April 2018. 81

[88] PÜSCHEL, M.: *A Discrete Signal Processing Framework for Meet/join Lattices with Applications to Hypergraphs and Trees.* In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 5371–5375, May 2019. 81

[89] PÜSCHEL, M., MOURA, J.: *The algebraic approach to the discrete cosine and sine transforms and their fast algorithms.* SIAM J. Comput. **32**.5 (2003), 1280–1316. 12, 25, 43, 45

[90] PÜSCHEL, M., MOURA, J.: *Algebraic Signal Processing Theory.* arXiv:cs/0612077v1, 2006. 8, 11, 12, 17, 18, 19, 20, 22, 23, 25, 26, 28

[91] Püschel, M., Moura, J.: *Algebraic Signal Processing Theory: 1-D Space.* Signal Processing, IEEE Trans. **56**.8 (2008), 3586–3599. 12, 21

[92] Püschel, M., Moura, J.: *Algebraic Signal Processing Theory: Cooley-Tukey Type algorithms for DCTs and DSTs.* Signal Processing, IEEE Trans. **56**.4 (2008), 1502–1521. 12, 35, 43

[93] Püschel, M., Moura, J.: *Algebraic Signal Processing Theory: Foundation and 1-D Time.* Signal Processing, IEEE Trans. **56**.8 (2008), 3572–3585. 12, 19

[94] Püschel, M., Rötteler, M.: *Fourier transform for the directed quincunx lattice.* In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, 401–404, March 2005. 43

[95] Püschel, M., Rötteler, M.: *Algebraic signal processing theory: 2-D spatial hexagonal lattice.* IEEE Trans. Image Process. **16**.6 (2007), 1506–1521. 13, 14, 51, 60, 62

[96] Püschel, M., Rötteler, M.: *Algebraic signal processing theory: Cooley-Tukey type algorithms on the 2-D hexagonal spatial lattice.* Applicable Algebra in Engineering, Communication and Computing **19**.3 (2008), 259–292. 13, 14, 35, 44, 51, 62, 63

[97] Ricci, P. E.: *Una proprietà iterativa dei polinomi di Chebyshev di prima specie in più variabili.* Rendiconti di matematica e delle sue applicazioni **6** (1986), 555–563. 14, 58

[98] Rivlin, T. J.: *The Chebyshev Polynomials.* Wiley Interscience, 1974. 12, 13, 46

[99] Robinson, M.: *Topological Signal Processing.* Springer, 2014. 81

[100] Robinson, M.: *Sheaves are the canonical data structure for sensor integration.* Information Fusion **36** (2017), 208–224. 81

[101] Rockmore, D. N.: *The FFT: an algorithm the whole family can use.* Computing in Science Engineering **2**.1 (2000), 60–64. 12, 35

[102] Roman, S.: *Advanced Linear Algebra.* Springer, 2008. 17

[103] Ryland, B. N., Munthe-Kaas, H.: *On Multivariate Chebyshev Polynomials and Spectral Approximations on Triangles.* In: Hesthaven, J. S., Ronquist, E. M. (eds.): *Spectral and High Order Methods for Partial Differential Equations*, 19–41. Springer, 2011. 15

[104] Sandryhaila, A., Kovacevic, J., Püschel, M.: *Algebraic Signal Processing Theory: Cooley-Tukey-Type Algorithms for Polynomial Transforms Based on Induction.* SIAM. J. Matrix Anal. & Appl. **32**.2 (2011), 364–384. 13, 35, 36, 39

[105] Sandryhaila, A., Kovacevic, J., Püschel, M.: *Algebraic Signal Processing Theory: 1-D Nearest Neighbor Models.* Signal Processing, IEEE Trans. **60**.5 (2012), 2247–2259. 28

[106] Sandryhaila, A., Moura, J.: *Big Data Analysis with Signal Processing on Graphs: Representation and processing of massive data sets with irregular structure.* Signal Processing Magazine, IEEE **31**.5 (2014), 80–90. 21, 81

[107] Sandryhaila, A., Moura, J.: *Discrete Signal Processing on Graphs: Frequency Analysis.* IEEE Transactions on Signal Processing **62**.12 (2014), 3042 – 3054. 21, 81

[108] Schmid, H., Xu, Y.: *On bivariate Gaussian cubature formula.* Proc. Amer. Math. Soc. **122**.3 (1994), 833–841. 31

[109] Seifert, B.: *FFT and orthogonal discrete transform on weight lattices of semi-simple Lie groups.* arXiv:1901.06254, 2019. 11, 15, 63, 68

[110] Seifert, B., Hüper, K.: *The discrete cosine transform on triangles.* In: *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5023–5026, May 2019. 14, 15, 52, 68

[111] Seifert, B., Hüper, K., Uhl, C.: *Fast cosine transform for FCC lattices.* In: *13th APCA International Conference on Control and Soft Computing (CONTROLO)*, 207–212, June 2018. 13, 14, 15, 51, 65, 66

[112] Serre, J.-P.: *Faisceaux algébriques cohérents.* Ann. Math **61**.2 (1955), 197–278. 11, 27

[113] Shafarevich, I.: *Basic Algebraic Geometry 1.* Springer, 1994. 25

[114] Sokolov, M.: *Generating Functions of Chebyshev Polynomials in Three Variables.* J. Math. Sci. **213**.5 (2016), 786–794. 14

[115] Swan, R. G.: *Vector Bundles and Projective Modules.* Trans. Amer. Math. Soc. **105**.2 (1962), 264–277. 11, 27

[116] Theodosis, E., Maragos, P.: *Tropical Modeling of Weighted Transducer Algorithms on Graphs.* In: *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8653–8657, May 2019. 81

[117] Uchimura, K.: *The dynamical systems associated with Chebyshev polynomials in two variables.* Int. J. Bifur. Chaos **6** (1996), 2611–2618. 14

[118] Vad, V., Csébfalvi, B., Rautek, P., Gröller, E.: *Towards an Unbiased Comparison of CC, BCC, and FCC Lattices in Terms of Prealiasing.* Computer Graphics Forum **33**.3 (2014), 81–90. 52

[119] Vesolov, A. P.: *What Is an Integrable Mapping?* In: Zakharov, V. E. (eds.): *What Is Integrability?*, 251–272. Springer, 1991. 12

[120] Vollrath, A.: *The Nonequispaced Fast SO(3) Fourier Transform, Generalisations and Applications.* PhD thesis, Universität zu Lübeck, 2010. 81

[121] Voronenko, Y., Püschel, M.: *Algebraic Signal Processing Theory: Cooley-Tukey Type Algorithms for Real DFTs.* IEEE Transactions on Signal Processing **57**.1 (2009), 205–222. 12, 23

[122] Wang, Z.: *Fast algorithms for the discrete W transform and for the discrete Fourier transform.* IEEE Trans. Acoust., Speech, Signal Proc. **32**.4 (1984), 803–816. 13

[123] Weyl, H.: *Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. I.* Mathematische Zeitschrift **23** (1925), 271–309. 61, 74

[124] Weyl, H.: *Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. II.* Mathematische Zeitschrift **24** (1926), 328–376. 61

[125] Weyl, H.: *Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. III.* Mathematische Zeitschrift **24** (1926), 377–395. 61

[126] Winograd, S.: *On computing the discrete Fourier transform.* Math. Comp. **32** (1978), 175–199. 11

[127] Winograd, S.: *On the multiplicative complexity of the discrete Fourier transform.* Advances in Mathematics **32** (1979), 83–117. 12

[128] Withers, W. D.: *Folding Polynomials and Their Dynamics.* American Mathematical Monthly **95**.5 (1988), 399–413. 14

[129] Xu, Y.: *On multivariate orthogonal polynomials.* SIAM J. Math. Anal. **24**.3 (1993), 783–794. 11, 28, 29

[130] Xu, Y.: *Orthogonal polynomials of several variables.* arXiv:1701.02709, 2017. 30, 31

[131] Yemini, Y., Pearl, J.: *Asymptotic properties of discrete unitary transforms.* IEEE Trans. on Pattern Analysis and Machine Intelligence **1**.4 (1979), 366–371. 11, 28

[132] Zheng, X., Gu, F.: *Fast Fourier Transform on FCC and BCC Lattices with Outputs on FCC and BCC Lattices Respectively.* J. Math. Imaging Vis. **49**.3 (2014), 530–550. 40