HAL open science

Semidefinite programming, harmonic analysis and coding theory

Christine Bachoc

▶ To cite this version:

Christine Bachoc. Semidefinite programming, harmonic analysis and coding theory. 2009. hal-00419745v2

HAL Id: hal-00419745
https://hal.science/hal-00419745v2

Preprint submitted on 8 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SEMIDEFINITE PROGRAMMING, HARMONIC ANALYSIS AND CODING THEORY

CHRISTINE BACHOC

ABSTRACT. These lecture notes where presented as a course of the CIMPA summer school in Manila, July 20-30, 2009, *Semidefinite programming in algebraic combinatorics*. This version is an update of June 2010.

CONTENTS

1. Introduction 2
1.1. Notations: 3
2. Linear representations of finite groups 3
2.1. Definitions 3
2.2. Examples 4
2.3. Irreducibility 5
2.4. The algebra of G-endomorphisms 7
2.5. Orthogonality relations 7
2.6. Characters 8
2.7. Induced representation and Frobenius reciprocity 9
2.8. Examples from coding theory 11
3. Linear representations of compact groups 13
3.1. Finite dimensional representations 14
3.2. Peter Weyl theorem 14
3.3. Examples 16
4. Harmonic analysis of compact spaces 17
4.1. Commuting endomorphisms and zonal matrices. 17
4.2. Examples: G-symmetric spaces. 21
4.3. Positive definite functions and Bochner theorem 22
5. Explicit computations of the matrices $E_k(x, y)$ 26
5.1. 2-point homogeneous spaces. 26
5.2. $X = \{1, \ldots, q\}$ under the action of S_q 26
5.3. The q-Hamming space 26
5.4. Other symmetric spaces 29
5.5. Three cases with non trivial multiplicities 30
6. An SDP upper bound for codes from positive definite functions 38
6.1. The 2-point homogeneous spaces 39
6.2. Symmetric spaces 40
6.3. Other spaces with true SDP bounds 41

Date: today.

1991 *Mathematics Subject Classification.* 52C17, 90C22.

Key words and phrases. spherical codes, kissing number, semidefinite programming, orthogonal polynomials.

7. Lovász theta 41
7.1. Introduction to Lovász theta number 41
7.2. Symmetrization and the q-gones 42
7.3. Relation with Delsarte bound and with $m(X, \delta)$ 44
8. Strengthening the LP bound for binary codes 44
References 46

1. Introduction

In coding theory, the so-called linear programming method, introduced by Philippe Delsarte in the seventies [16] as proved to be a very powerful method to solve extremal problems. It was initially developed in the framework of association schemes and then extended to the family of 2-point homogeneous spaces, including the compact real manifolds having this property (see [18], [24], [13, Chapter 9]). Let us recall that a 2-point homogeneous space is a metric space on which a group G acts transitively, leaving the distance d invariant, and such that, for $(x, y) \in X^2$, there exists $g \in G$ such that $(gx, gy) = (x', y')$ if and only if $d(x, y) = d(x', y')$. The Hamming space H_n and the unit sphere of the Euclidean space S^{n-1} are core examples of such spaces which play a major role in coding theory. To such a space is associated a sequence of orthogonal polynomials $(P_k)_{k\geq 0}$ such that, for all $C \subset X$,

$$\sum_{(c,c')\in C^2} P_k(d(c, c')) \geq 0.$$

These inequalities can be understood as linear constraints on the distance distribution of a code and are at the heart of the LP method.

The applications of this method to the study of codes and designs are numerous: very good upper bounds for the number of elements of a code with given minimal distance can be obtained with this method, including a number of cases where this upper bound is tight and leads to a proof of optimality and uniqueness of certain codes, as well as to the best known asymptotic bounds (see [16], [30], [24], [13, Chapter 9], [28]).

In recent years, the development of the theory of error correcting codes has introduced many other spaces with interesting applications. To cite a few, codes over various alphabets associated to various weights, quantum codes, codes for the multi antenna systems of communications involving more complicated manifolds like the Grassmann spaces, have successively focused attention. For these spaces there was a need for a generalization of the classical framework of the linear programming method. This generalization was developed for some of these spaces, see [44], [45], [2], [37]. It turns out that in each of these cases, a certain sequence of orthogonal polynomials enters into play but unlike the classical cases, these polynomials are multivariate.

Another step was taken when A. Schrijver in [40] succeeded to improve the classical LP bounds for binary codes with the help of semidefinite programming. To that end he exploited *SDP constraints on triples of points* rather than on pairs, arising from the analysis of the Terwilliger algebra of the Hamming scheme. His method was then adapted to the unit sphere [4] in the framework of the representations of the orthogonal group. The heart of the method is to evidence matrices

$Z_k(x, y, z)$ such that for all $C \subset X$,

$$\sum_{(c,c',c'')\in C^3} Z_k(c, c', c'') \succeq 0.$$

Another motivation for the study of SDP constraints on k-tuples of points can be found in coding theory. It appears that not only functions on pairs of points such as a distance function $d(x, y)$ are of interest, but also functions on k-tuples have relevant meaning, e.g. in connection with the notion of list decoding.

In these lecture notes we want to develop a general framework based on harmonic analysis of compact groups for these methods. In view of the effective applications to coding theory, we give detailed computations in many cases. Special attention will be paid to the cases of the Hamming space and of the unit sphere.

Section 2 develops the basic tools needed in the theory of representations of finite groups, section 3 is concerned with the representations of compact groups and Peter Weyl theorem. Section 4 discusses the needed notions of harmonic analysis: the zonal matrices are introduced and the invariant positive definite functions are characterized with Bochner theorem. Section 5 is devoted to explicit computations of the zonal matrices. Section 6 shows how the determination of the invariant positive definite functions leads to an upper bound for codes with given minimal distance. Section 7 explains the connection with the so-called Lovász theta number. Section 8 shows how SDP bounds can be used to strengthen the classical LP bounds, with the example of the Hamming space.

1.1. **Notations:** for a matrix A with complex coefficients, A^* stands for the transposed conjugate matrix. A squared matrix is said to be hermitian if $A^* = A$ and positive semidefinite if it is hermitian and all its eigenvalues are non negative. This property is denoted $A \succeq 0$. We follow standard notations for sets of matrices: the set of $n \times m$ matrices with coefficients in a field K is denoted $K^{n\times m}$; the group of $n \times n$ invertible matrices by $\mathrm{Gl}(K^n)$; the group $U(\mathbb{C}^n)$ of unitary matrices, respectively $O(\mathbb{R}^n)$ of orthogonal matrices is the set of matrices $A \in \mathrm{Gl}(C^n)$, respectively $A \in \mathrm{Gl}(\mathbb{R}^n)$ such that $A^* = A^{-1}$. The space $\mathbb{C}^{n\times m}$ is endowed with the standard inner product $\langle A, B\rangle = \mathrm{Trace}(AB^*) = \sum_{i,j} A_{i,j}\overline{B_{i,j}}$. The number of elements of a finite set X is denoted $\mathrm{card}(X)$ of $|X|$.

2. Linear representations of finite groups

In this section we shortly review the basic notions of group representation theory that will be needed later. There are many good references for this theory e.g. [41], or [38] which is mainly devoted to the symmetric group.

2.1. **Definitions.** Let G be a finite group. A (complex linear) representation of G is a finite dimensional complex vector space V together with a homomorphism ρ:

$$\rho : G \to \mathrm{Gl}(V)$$

where $\mathrm{Gl}(V)$ is the general linear group of V, i.e. the set of linear invertible transformations of V. The degree of the representation (ρ, V) is by definition equal to the dimension of V.

Two representations of G say (ρ, V) and (ρ', V') are said to be equivalent or isomorphic if there exists and isomorphism $u : V \to V'$ such that, for all $g \in G$,

$$\rho'(g) = u\rho(g)u^{-1}.$$

For example, the choice of a basis of V leads to a representation equivalent to (ρ, V) given by $(\rho', \mathbb{C}^d)$ where $d = \dim(V)$ and $\rho'(g)$ is the matrix of $\rho(g)$ in the chosen basis. In general, a representation of G such that $V = \mathbb{C}^d$ is called a matrix representation.

The notion of a G-module is equivalent to the above notion of representation and turns out to be very convenient. A G-module, or a G-space, is a finite dimensional complex vector space V such that for all $g \in G$, $v \in V$, $gv \in V$ is well defined and satisfies the obvious properties: $1v = v$, $g(hv) = (gh)v$, $g(v+w) = gv+gw$, $g(\lambda v) = \lambda(gv)$ for $g, h \in G$, $v, w \in V$, $\lambda \in \mathbb{C}$. In other words, V is endowed with a structure of $\mathbb{C}[G]$-module. One goes from one notion to the other by the identification $gv = \rho(g)(v)$. The notion of equivalent representations corresponds to the notion of isomorphic G-modules, an isomorphism of G-modules being an isomorphism of vector spaces $u : V \to V'$ such that $u(gv) = gu(v)$. Note that here the operations of G on V and V' are denoted alike, which may cause some confusion.

2.2. **Examples.**

- The trivial representation $\mathbf{1}$: $V = \mathbb{C}$ and $gv = v$.
- Permutation representations: let X be a finite set on which G acts (on the left). Let $V_X := \oplus_{x \in X} \mathbb{C} e_x$. A natural action of G on V_X is given by $ge_x = e_{gx}$, and defines a representation of G, of degree $|X|$. The matrices of this representation (in the basis $\{e_x\}$) are permutation matrices.
  - The symmetric group S_n acts on $X = \{1, 2, \ldots, n\}$. This action defines a representation of degree n of S_n.
  - For all w, $1 \leq w \leq n$, S_n acts on the set X_w of subsets of $\{1, 2, \ldots, n\}$ of cardinal w. In coding theory an element of X_w is more likely viewed as a binary word of length n and Hamming weight w. The spaces X_w are called the Johnson spaces and denoted J_n^w.
- The regular representation is obtained with the special case $X = G$ with the action of G by left multiplication. In the case $G = S_n$ it has degree $n!$.. It turns out that the regular representation contains all building blocks of all representations of G.
- Permutation representations again: if G acts transitively on X, this action can be identified with the left action of G on the left cosets $G/H = \{gH : g \in G\}$ where $H = \mathrm{Stab}(x_0)$ is the stabilizer of a base point.
  - The symmetric group S_n acts transitively on $X = \{1, 2, \ldots, n\}$ and the stabilizer of one point (say n) can be identified with the symmetric group S_{n-1} acting on $\{1, \ldots, n-1\}$.
  - The action of S_n on J_n^w is also transitive and the stabilizer of one point (say $1^w 0^{n-w}$) is the subgroup $S_{\{1,\ldots,w\}} \times S_{\{w+1,\ldots,n\}}$ isomorphic to $S_w \times S_{n-w}$.
  - The Hamming space $H_n = \{0, 1\}^n = \mathbb{F}_2^n$ affords the transitive action of $G = T \rtimes S_n$ where T is the group of translations $T = \{t_u : u \in H_n\}$, $t_u(v) = u + v$ and S_n permutes the coordinates. The stabilizer of 0^n is the group of permutations S_n.
- Another way to see the permutation representations is the following: let

$$\mathcal{C}(X) := \{f : X \to \mathbb{C}\}$$

be the space of functions from X to $\mathbb{C}$. The action of G on X extends to a structure of G-module on $\mathcal{C}(X)$ given by:

$$gf(x) := f(g^{-1}x).$$

For the Dirac functions δ_y ($\delta_y(x) = 1$ if $x = y$, 0 otherwise), the action of G is given by $g\delta_y = \delta_{gy}$ thus this representation is isomorphic to the permutation representation defined by X. This apparently more complicated presentation of permutation representations has the advantage to allow generalization to infinite groups acting on infinite spaces as we shall encounter later.

2.3. **Irreducibility.** Let V be a G-module (respectively a representation (ρ, V) of G). A subspace $W \subset V$ is said to be G-invariant (or G-stable, or a G-submodule, or a subrepresentation of (ρ, V)), if $gw \in W$ (respectively $\rho(g)(w) \in W$) for all $g \in G$, $w \in W$.
Example: $V = V_G$ and $W = \mathbb{C}e_G$ with $e_G = \sum_{g\in G} e_g$. The restriction of the action of G to W is the trivial representation.

A G-module V is said to be irreducible if it does not contain any subspace W, $W \neq \{0\}, V$, invariant under G. Otherwise it is called reducible.
Example: The G-invariant subspaces of dimension 1 are necessarily irreducible. If G is abelian, a G-module of dimension greater than 1 cannot be irreducible, because endomorphisms that pairwise commute afford a common basis of eigenvectors.

The main result is then the decomposition of a G-module into the direct sum of irreducible submodules:

Theorem 2.1 (Maschke's theorem)**.** *Any G-module $V \neq \{0\}$ is the direct sum of irreducible G-submodules $W_1, \dots, W_k$:*

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k. \tag{1}$$

Proof. By induction, it is enough to prove that any G-submodule W of V affords a supplementary subspace which is also G-invariant. The main idea is to construct a G-invariant inner product and then prove that the orthogonal of W for this inner product makes the job.

We start with an inner product $\langle x, y\rangle$ defined on V. There are plenty of them since V is a finite dimensional complex vector space. For example we can choose an arbitrary basis of V and declare it to be orthonormal. Then we average this inner product on G, defining:

$$\langle x, y\rangle' := \sum_{g\in G} \langle gx, gy\rangle.$$

It is not difficult to check that we have defined a inner product which is G-invariant. It is also easy to see that

$$W^\perp := \{v \in V : \langle v, w\rangle' = 0 \text{ for all } w \in W\}$$

is G-invariant, thus we have the decomposition of G-modules:

$$V = W \oplus W^\perp$$

□

It is worth to notice that the above decomposition may not be unique. It is clear if one thinks of the extreme case $G = \{1\}$ for which the irreducible subspaces are simply the one dimensional subspaces of V. The decomposition of V into the direct sum of subspaces of dimension 1 is certainly not unique (if $\dim(V) > 1$ of course). But uniqueness is fully satisfied by the decomposition into isotypic subspaces. In order to define them we take the following notation: let $\mathcal{R}$ be a complete set of pairwise non isomorphic irreducible representations of G. We have seen that any G-module affords a G-invariant inner product so the action of G on R is expressed by unitary matrices in a given orthonormal matrix of R. According to the context we view R either as a G-module or as a homomorphism $g \mapsto R(g) \in U(\mathbb{C}^n)$. It will turn out that there is only a finite number of them but we have not proved it yet. The isotypic subspace $\mathcal{I}_R$ of V associated to $R \in \mathcal{R}$ is defined, with the notations of (1), by:

$$\mathcal{I}_R := \oplus_{W_i \simeq R} W_i. \tag{2}$$

Theorem 2.2. *Let $R \in \mathcal{R}$. The isotypic spaces $\mathcal{I}_R$ do not depend on the decomposition of V as the direct sum of G-irreducible subspaces. We have the canonical decomposition*

$$V = \oplus_{R \in \mathcal{R}} \mathcal{I}_R.$$

Any G-subspace $W \subset V$ such that $W \simeq R$ is contained in $\mathcal{I}_R$ and any G-irreducible subspace of $\mathcal{I}_R$ is isomorphic to R. A decomposition into irreducible subspaces of $\mathcal{I}_R$ has the form

$$\mathcal{I}_R = W_1 \oplus \cdots \oplus W_{m_R}$$

with $W_i \simeq R$. Such a decomposition is not unique in general but the number m_R does not depend on the decomposition and is called the multiplicity of R in V.

Moreover, if V is endowed with a G-invariant inner product, then the isotypic spaces are pairwise orthogonal.

Proof. We start with a lemma which points out a very important property of irreducible G-modules.

Lemma 2.3 (Schur Lemma)**.** *Let R_1 and R_2 two irreducible G-modules and let $\varphi : R_1 \to R_2$ be a G-homomorphism. Then either $\varphi = 0$ or φ is an isomorphism of G-modules.*

Proof. The subspaces $\ker \varphi$ and $\operatorname{im} \varphi$ are G-submodules of respectively R_1 and R_2 thus they are equal to either $\{0\}$ or R_i. □

We go back to the proof of the theorem. We start with the decomposition (1) of V and the definition (2) of $\mathcal{I}_R$, a priori depending on the decomposition. Let $W \subset V$, a G-submodule isomorphic to R. We apply Lemma 2.3 to the projections p_{W_i} and conclude that either $p_{W_i}(W) = \{0\}$ or $p_{W_i}(W) = W_i$ and this last case can only happen if $W \simeq W_i$. It proves that $W \subset \mathcal{I}_R$ and that a G-irreducible subspace of $\mathcal{I}_R$ can only be isomorphic to R. It also proves that

$$\mathcal{I}_R = \sum_{W \subset V, W \simeq R} W$$

hence giving a characterization of $\mathcal{I}_R$ independent of the initial decomposition. The number m_R must satisfy $\dim(\mathcal{I}_R) = m_R \dim(R)$ so it is independent of the decomposition of $\mathcal{I}_R$.

If V is equipped with a G-invariant inner product, we consider orthogonal projections. Schur Lemma shows that $P_W(W') = \{0\}$ or $= W$ if W and W' are irreducible. Thus if they are not G-isomorphic, W and W' must be orthogonal.

□

2.4. **The algebra of G-endomorphisms.** Let V be a G-module. The set of G-endomorphisms of V is an algebra (for the laws of addition and composition) denoted $\mathrm{End}_G(V)$. The next theorem describes the structure of this algebra.

Theorem 2.4. *If $V \simeq \oplus_{R\in\mathcal{R}} R^{m_R}$, then*

$$\mathrm{End}_G(V) \simeq \prod_{R\in\mathcal{R}} \mathbb{C}^{m_R\times m_R}.$$

Proof. The proof is in three steps: we shall assume first $V = R$ is irreducible, then $V \simeq R^m$, then the general case. Schur Lemma 2.3 is the main tool here.

If V is irreducible, let $\varphi \in \mathrm{End}_G(V)$. Since V is a complex vector space, φ has got an eigenvalue λ. Then $\varphi - \lambda\,\mathrm{Id}$ is a G-endomorphism with a non trivial kernel so from Schur Lemma $\varphi - \lambda\,\mathrm{Id} = 0$. We have proved that

$$\mathrm{End}_G(V) = \{\lambda\,\mathrm{Id}, \lambda \in \mathbb{C}\} \simeq \mathbb{C}.$$

We assume now that $V \simeq R^m$ and we fix a decomposition $V = W_1 \oplus \cdots \oplus W_m$. For all $1 \le i \le j \le m$, let $u_{j,i} : W_i \to W_j$ an isomorphism of G-modules such that the relations

$$u_{k,j} \circ u_{j,i} = u_{k,i} \text{ and } u_{i,i} = \mathrm{Id}$$

hold for all i, j, k. Let $\varphi \in \mathrm{End}_G(V)$; we associate to φ an element of $\mathbb{C}^{m\times m}$ in the following way. From previous discussion of the irreducible case it follows that for all i, j there exists $a_{i,j} \in \mathbb{C}$ such that, for all $v \in W_i$,

$$p_{W_j} \circ \varphi(v) = a_{j,i} u_{j,i}(v).$$

The matrix $A = (a_{i,j})$ is the matrix associated to φ. The proof that the mapping $\varphi \mapsto A$ is an isomorphism of algebras carries without difficulties and is left to the reader.

In the general case, $V = \oplus_{R\in\mathcal{R}} \mathcal{I}_R$. Let $\varphi \in \mathrm{End}_G(V)$. It is clear that $\varphi(\mathcal{I}_R) \subset \mathcal{I}_R$ thus

$$\mathrm{End}_G(V) = \oplus_{R\in\mathcal{R}} \mathrm{End}_G(\mathcal{I}_R)$$

and we are done. □

It is worth to notice that $\mathrm{End}_G(V)$ is a commutative algebra if and only if all the multiplicities m_R are equal to either 0 or 1. In this case we say that V is multiplicity free. It is also the unique case when the decomposition into irreducible subspaces (1) is unique.

2.5. **Orthogonality relations.** Another important result which is a consequence of Schur lemma is the orthogonality relations between the matrix coefficients of the elements of $\mathcal{R}$:

Theorem 2.5. *For $R \in \mathcal{R}$, let $d_R := \dim(R)$. For all $R, S \in \mathcal{R}$, i, j, k, l,*

$$\langle R_{i,j}, S_{k,l}\rangle = \frac{1}{d_R}\delta_{R,S}\delta_{i,k}\delta_{j,l}.$$

Proof. For $A \in \mathbb{C}^{d_R \times d_S}$, let

$$A' = \frac{1}{|G|} \sum_{g \in G} R(g) A S(g)^{-1}.$$

This matrix satisfies $R(g)A' = A'S(g)$ for all $g \in G$. In other words it defines an homomorphism of G-modules from $(\mathbb{C}^{d_S}, S)$ to $(\mathbb{C}^{d_R}, R)$. Schur lemma shows that if $S \neq R$, $A' = 0$ and if $S = R$, $A' = \lambda \mathrm{Id}$. Computing the trace of A' shows that $\lambda = \mathrm{Trace}(A)/d_R$. Taking $A = E_{i,j}$ the elementary matrices, with the property that $S(g)^{-1} = S(g)^*$, leads to the announced formula. □

2.6. **Characters.** The character of a representation (ρ, V) of G is the function $\chi_\rho : G \to \mathbb{C}$ defined by

$$\chi_\rho(g) = \mathrm{Trace}(\rho(g)).$$

As a consequence of the standard property of traces of matrices $\mathrm{Trace}(AB) = \mathrm{Trace}(BA)$, the character of a representation only depends on its equivalence class, and it is a complex valued function on G which is constant on the conjugacy classes of G (such a function is called a class function). The inner product of any two χ, $\psi \in \mathcal{C}(G)$ is defined by

$$\langle \chi, \psi \rangle := \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\psi(g)}.$$

We have the very important orthogonality relations between characters:

Theorem 2.6 (Orthogonality relations of the first kind)**.** *Let χ and χ' be respectively the characters of two irreducible representations (ρ, V) and (ρ', V') of G. Then*

$$\langle \chi, \chi' \rangle = \begin{cases} 1 & \text{if } \rho \simeq \rho' \\ 0 & \text{otherwise.} \end{cases}$$

Proof. It is a straight forward consequence of Theorem 2.5, since the trace of a representation is the sum of the diagonal elements of any equivalent matrix representation. □

A straightforward consequence of the above theorem is that $\langle \chi_\rho, \chi_R \rangle = m_R$ for all $R \in \mathcal{R}$. This property is a very convenient tool to study the irreducible decomposition of a given representation (ρ, V) of G; in particular it shows that a representation is irreducible if and only if its character χ satisfies $\langle \chi, \chi \rangle = 1$. In the case of the regular representation it leads to the following very important result:

Theorem 2.7. *[Decomposition of the regular representation]*

$$\mathcal{C}(G) \simeq \oplus_{R \in \mathcal{R}} R^{\dim(R)}$$

Proof. Compute the character of the regular representation. □

A consequence of the above theorem is the finiteness of the number of irreducible representations of a given finite group, together with the formula

$$|G| = \sum_{R \in \mathcal{R}} (\dim(R))^2$$

which shows e.g. completeness of a given set of irreducible G-modules.

A second consequence of the orthogonality relations is that a representation of G is uniquely characterized up to isomorphism by its character.

Theorem 2.8.

$$(\rho, V) \simeq (\rho', V') \Longleftrightarrow \chi_\rho = \chi_{\rho'}.$$

Proof. If $\chi_\rho = \chi_{\rho'}$, the multiplicities of an irreducible representation of G are the same in V and V', hence $V \simeq_G V'$. □

Let us denote by $R(G)$ the subspace of elements of $\mathcal{C}(G)$ which are constant on the conjugacy classes $C_1, \ldots, C_s$ of G. The dimension of $R(G)$ is obviously the number s of conjugacy classes of G. We have seen that the characters χ_R of the irreducible representations of G belong to $R(G)$ and form an orthonormal family. It turns out that they in fact form a basis of $R(G)$, which in other words means that the number of irreducible representations of G is exactly equal to its number of conjugacy classes.

Theorem 2.9.

$$R(G) = \oplus_{R \in \mathcal{R}} \mathbb{C}\chi_R.$$

Proof. It is clear that $\mathcal{C}(G) = \mathbb{C}[G]\delta_e$. Thus $\mathrm{End}_G(\mathcal{C}(G)) \simeq \mathbb{C}[G]$. In particular, the center of $\mathrm{End}_G(\mathcal{C}(G))$ is isomorphic to the center $Z(\mathbb{C}[G])$ of $\mathbb{C}[G]$. It is easy to verify that the center of $\mathbb{C}[G]$ is the vector space spanned by the elements $\lambda_i := \sum_{g \in C_i} g$ associated to each conjugacy class C_i of G, thus $Z(\mathbb{C}[G])$ is of dimension s the number of conjugacy classes of G. On the other hand, as a consequence of Theorem 2.7 and Theorem 2.4, we have $\mathrm{End}_G(\mathcal{C}(G)) \simeq \prod_{R \in \mathcal{R}} \mathbb{C}^{d_R \times d_R}$ where $d_R = \dim(R)$. Thus the center of $\mathrm{End}_G(\mathcal{C}(G))$ is isomorphic to $\mathbb{C}^{|\mathcal{R}|}$ and we have proved that the number of G-irreducible modules is equal to the number of conjugacy classes of G. □

Remark 2.10. *There is not in general a natural bijection between the set of conjugacy classes of G and the set of its irreducible representations. However, in the special case of the symmetric group S_n, such a correspondance exists. The conjugacy classes are naturally indexed by the partitions λ of n and to every partition λ of n is associated an irreducible module S^λ also called a Specht module (see* [38]*).*

2.7. **Induced representation and Frobenius reciprocity.** Induction is a way to construct representations of a group G from representations of its subgroups. Looking at the irreducible subspaces of representations that are induced from subgroups is a very convenient way to find new irreducible representations of a group G. Induction is an operation on representations which is dual to the easier to understand restriction. If V is a G-module and H is a subgroup of G, the restriction $\mathrm{Res}_H^G(V)$ is simply the space V considered as a $\mathbb{C}[H]$-module. If V is an H-module, we define $\mathrm{Ind}_H^G(V)$ to be the $\mathbb{C}[G]$-module

$$\mathrm{Ind}_H^G(V) := \mathbb{C}[G] \otimes_{\mathbb{C}[H]} V.$$

Here we exploit the bi-module structure of $\mathbb{C}[G]$ (the tensor product over $\mathbb{C}[H]$ means that $\lambda\mu \otimes v = \lambda \otimes \mu v$ when $\mu \in \mathbb{C}[H]$). A more explicit (but less intrisic) formulation for this construction is the following: let $\{x_1, \ldots, x_t\}$ be a complete system of representatives of G/H, so that $G = x_1 H \cup \cdots \cup x_t H$. Then

$$\mathrm{Ind}_H^G(V) = \oplus_{i=1}^t x_i V$$

where the left action of G is as follows: for all i, there is j and $h \in H$ both depending on g such that $gx_i = x_j h$. Then $gx_i v := x_j(hv)$ where $hv \in V$. A

third construction of $\mathrm{Ind}_H^G(V)$ is the following:

$$\mathrm{Ind}_H^G(V) = \{f : G \to V \text{ such that } f(gh) = h^{-1}f(g)\}.$$

The equivalence of these three formulations is a recommended exercise !

Example: The permutation representation of G on $X = G/H$ is nothing else than the induction of the trivial representation of H. In short, $\mathcal{C}(X) = \mathrm{Ind}_H^G \mathbf{1}$.

Since the induction of two isomorphic H-modules leads to isomorphic G-modules and similarly for the restriction, these operations act on the characters thus we denote similarly $\mathrm{Res}_H^G \chi$, $\mathrm{Ind}_H^G \chi$ the characters of the corresponding modules.

Lemma 2.11. *Let χ be a character of H. The induced character $\mathrm{Ind}_H^G \chi$ is given by the formula:*

$$\mathrm{Ind}_H^G \chi(g) = \frac{1}{|H|} \sum_{\substack{x \in G \\ x^{-1}gx \in H}} \chi(x^{-1}gx).$$

Proof. We take a decomposition $\mathrm{Ind}_H^G(V) = x_1V \oplus \cdots \oplus x_tV$ where $\{x_1, \ldots, x_t\}$ are representatives of G/H. Since $gx_iv = x_jhv$ with the notations above, $g(x_iV) \subset x_jV$ and the block x_iV will contribute to the trace of $x \mapsto gx$ only when $j = i$, which corresponds to the case when $x_i^{-1}gx_i \in H$. Then, the multiplication by g on x_iV acts like the multiplication by $h = x_i^{-1}gx_i$ on V. Thus we have

$$\begin{aligned}\mathrm{Ind}_H^G \chi(g) &= \sum_{\substack{1 \le i \le t \\ x_i^{-1}gx_i \in H}} \chi(x_i^{-1}gx_i) \\ &= \frac{1}{|H|} \sum_{\substack{x \in G \\ x^{-1}gx \in H}} \chi(x^{-1}gx).\end{aligned}$$

□

The duality between the operations of restriction and induction is expressed in the following important theorem:

Theorem 2.12 (Frobenius reciprocity)**.** *Let H be a subgroup of G and let χ and ψ be respectively a character of H and a character of G. Then*

$$\langle \mathrm{Ind}_H^G \chi, \psi \rangle = \langle \chi, \mathrm{Res}_H^G \psi \rangle.$$

Proof. Let $\tilde{\chi} : G \to \mathbb{C}$ be defined by: $\tilde{\chi}(g) = \chi(g)$ if $g \in H$ and $\tilde{\chi}(g) = 0$ if $g \notin H$ (of course $\tilde{\chi}$ is not a character of G). We compute $\langle \mathrm{Ind}_H^G \chi, \psi \rangle$:

$$\begin{aligned}
\langle \mathrm{Ind}_H^G \chi, \psi \rangle &= \frac{1}{|G|} \sum_{x \in G} \mathrm{Ind}_H^G \chi(g) \overline{\psi(g)} \\
&= \frac{1}{|G||H|} \sum_{g \in G} \big(\sum_{x \in G} \tilde{\chi}(x^{-1} g x) \big) \overline{\psi(g)} \\
&= \frac{1}{|G||H|} \sum_{x \in G} \big(\sum_{g \in G} \tilde{\chi}(x^{-1} g x) \overline{\psi(g)} \big) \\
&= \frac{1}{|G||H|} \sum_{x \in G} \big(\sum_{g' \in G} \tilde{\chi}(g') \overline{\psi(x g' x^{-1})} \big) \\
&= \frac{1}{|G||H|} \sum_{x \in G} \big(\sum_{g' \in G} \tilde{\chi}(g') \overline{\psi(g')} \big) \\
&= \frac{1}{|H|} \sum_{h \in H} \chi(h) \overline{\psi(h)} = \langle \chi, \mathrm{Res}_H^G \psi \rangle.
\end{aligned}$$

□

2.8. **Examples from coding theory.** In coding theory we are mostly interested in the decomposition of $\mathcal{C}(X)$ under the action of $G = \mathrm{Aut}(X)$ for various spaces X. We recall that the action of G on $f \in \mathcal{C}(X)$ is given by $(gf)(x) = f(g^{-1}x)$. The space $\mathcal{C}(X)$ is endowed with the inner product

$$\langle f, f' \rangle = \frac{1}{|X|} \sum_{x \in X} f(x) \overline{f'(x)}.$$

which is G-invariant.

2.8.1. **The binary Hamming space H_n:** recall that $G = T \rtimes S_n$. Let, for $y \in H_n$, $\chi_y \in \mathcal{C}(H_n)$ be defined by $\chi_y(x) = (-1)^{x \cdot y}$. The set $\{\chi_y, y \in H_n\}$ is exactly the set of irreducible characters of the additive group $\mathbb{F}_2^n$, and form an orthonormal basis of $\mathcal{C}(H_n)$. The computation of the action of G on χ_y shows that for $\sigma \in S_n$, $\sigma \chi_y = \chi_{\sigma(y)}$ and for $t_u \in T$, $t_u \chi_y = (-1)^{u \cdot y} \chi_y$. Let, for $0 \leq k \leq n$,

$$P_k := \perp_{y, wt(y)=k} \mathbb{C} \chi_y$$

Thus P_k is a G-invariant subspace of $\mathcal{C}(H_n)$ of dimension $\binom{n}{k}$ and we have the decomposition

$$\mathcal{C}(H_n) = P_0 \perp P_1 \perp \cdots \perp P_n. \tag{3}$$

The computation $\langle \chi_{P_k}, \chi_{P_k} \rangle = 1$ where χ_{P_k} is the character of the G-module P_k shows that these modules are G-irreducible.

Now we introduce the Krawtchouk polynomials. The element $Z_k := \sum_{wt(y)=k} \chi_y$ of $\mathcal{C}(H_n)$ is S_n-invariant. In other words, $Z_k(x)$ only depends on $wt(x)$. We define

the Krawtchouk polynomial K_k for $0 \leq k \leq n$ by

$$K_k(w) := Z_k(x) = \sum_{wt(y)=k} (-1)^{x \cdot y} \text{ where } wt(x) = w \tag{4}$$

$$= \sum_{i=0}^{w} (-1)^i \binom{w}{i} \binom{n-w}{k-i}. \tag{5}$$

We review some properties of these polynomials:

(1) $\deg(K_k) = k$
(2) $K_k(0) = \binom{n}{k}$
(3) Orthogonality relations: for all $0 \leq k \leq l \leq n$

$$\frac{1}{2^n} \sum_{w=0}^{n} \binom{n}{w} K_k(w) K_l(w) = \delta_{k,l} \binom{n}{k}$$

The last property is just a reformulation of the orthogonality of the $Z_k \in P_k$, since, if $f, f' \in \mathcal{C}(H_n)$ are S_n-invariant, and $\tilde{f}(w) := f(x)$, $wt(x) = w$,

$$\langle f, f' \rangle = \frac{1}{2^n} \sum_{x \in H_n} f(x) f'(x)$$
$$= \frac{1}{2^n} \sum_{w=0}^{n} \binom{n}{w} \tilde{f}(w) \tilde{f}'(w).$$

The above three properties characterize uniquely the Krawtchouk polynomials.

Let $C \subset H_n$ be a binary code. Let $\mathbf{1}_C$ be the characteristic function of C. The obvious inequalities hold:

$$0 \leq k \leq n, \quad \sum_{wt(y)=k} \langle \mathbf{1}_C, \chi_y \rangle^2 \geq 0. \tag{6}$$

Since the decomposition of $\mathbf{1}_C$ over the basis χ_y reads

$$\mathbf{1}_C = \sum_{y \in H_n} \langle \mathbf{1}_C, \chi_y \rangle \chi_y.$$

the above inequalities are indeed reformulations of the non negativity of the squared norm of the projections $p_{P_k}(\mathbf{1}_C)$. They express in terms of the Krawtchouk polynomials:

$$0 \leq k \leq n, \quad \frac{1}{2^{2n}} \sum_{(x,x') \in C^2} K_k(d_H(x, x')) \geq 0 \tag{7}$$

or equivalently in terms of the distance distribution of the code C: if

$$A_w(C) := \frac{1}{|C|} |\{(x, x') \in C^2 : d_H(x, x') = w\}|$$

then

$$0 \leq k \leq n, \quad \frac{|C|}{2^{2n}} \sum_{w=0}^{n} A_w(C) K_k(w) \geq 0.$$

These inequalities are the basic inequalities involved in Delsarte linear programming method. We shall encounter similar inequalities in a very general setting.

In the special case when C is linear, we have

$$\langle \mathbf{1}_C, \chi_y \rangle = \frac{|C|}{2^n} \mathbf{1}_{C^\perp}(y)$$

so that we recognise the identity

$$\sum_{wt(y)=k} \langle \mathbf{1}_C, \chi_y \rangle^2 = \frac{|C|}{2^{2n}} \sum_{w=0}^{n} A_w(C) K_k(w)$$

to be the Mac Williams identity

$$A_k(C^\perp) = \frac{1}{|C|} \sum_{w=0}^{n} A_w(C) K_k(w).$$

The more general q-ary Hamming space affords similar results; it is treated in 5.3.

2.8.2. **The Johnson spaces J_n^w:** the group is $G = S_n$. Here, we shall see at work a standard way to evidence G-submodules as kernels of G-endomorphisms. For details we refer to [17] where the q-Johnson spaces are given a uniform treatment. We introduce the applications

$$\begin{aligned} \delta : \mathcal{C}(J_n^w) &\to \mathcal{C}(J_n^{w-1}) \\ f \mapsto \delta(f) : \delta(f)(x) &:= \sum_{y \in J_n^w,\ x \subset y} f(y) \end{aligned}$$

and

$$\begin{aligned} \psi : \mathcal{C}(J_n^{w-1}) &\to \mathcal{C}(J_n^w) \\ f \mapsto \psi(f) : \psi(f)(x) &:= \sum_{y \in J_n^{w-1},\ y \subset x} f(y) \end{aligned}$$

Both of these applications commute with the action of G. They satisfy the following properties: $\langle f, \psi(f') \rangle = \langle \delta(f), f' \rangle$, ψ is injective and δ is surjective. Therefore the subspace of $\mathcal{C}(J_n^w)$:

$$H_w := \ker \delta$$

is a G-submodule of dimension $\binom{n}{w} - \binom{n}{w-1}$ and we have the orthogonal decomposition

$$\mathcal{C}(J_n^w) = H_w \perp \psi(\mathcal{C}(J_n^{w-1})) \simeq H_w \perp \mathcal{C}(J_n^{w-1}).$$

By induction we obtain a decomposition

$$\mathcal{C}(J_n^w) \simeq H_w \perp H_{w-1} \perp \cdots \perp H_0$$

which can be proved to be the irreducible decomposition of $\mathcal{C}(J_n^w)$ (see 5.3.1).

3. Linear representations of compact groups

In this section we enlarge the discussion to the representation theory of compact groups. For this section we refer to [12].

3.1. **Finite dimensional representations.** The theory of finite dimensional representations of finite groups extends nicely and straightforwardly to compact groups. A finite dimensional representation of a compact group G is a continuous homomorphism $\rho : G \to \mathrm{Gl}(V)$ where V is a complex vector space of finite dimension.

A compact group G affords a Haar measure, which is a regular left and right invariant measure. We assume this measure to be normalized, i.e. the group G has measure 1. With this measure the finite sums over elements of a finite group can be replaced with integrals; so the crucial construction of a G-invariant inner product in the proof of Maschke theorem extends to compact groups with the formula

$$\langle x, y\rangle' := \int_G \langle gx, gy\rangle dg.$$

Hence Maschke theorem remains valid for finite dimensional representations. We keep the notation $\mathcal{R}$ for a set of representatives of the finite dimensional irreducible representations of G, chosen to be representations with unitary matrices. A main difference with the finite case is that $\mathcal{R}$ is not finite anymore.

3.2. **Peter Weyl theorem.** Infinite dimensional representations immediately occur with the generalization of permutation representations. Indeed, if G acts continuously on a space X, it is natural to consider the action of G on the space $\mathcal{C}(X)$ of complex valued continuous functions on X given by $(gf)(x) = f(g^{-1}x)$ to be a natural generalization of permutation representations. A typical example of great interest in coding theory is the action of $G = O(\mathbb{R}^n)$ on the unit sphere of the Euclidean space:

$$S^{n-1} := \{x \in \mathbb{R}^n \ : \ x \cdot x = 1\}.$$

The regular representation, which is the special case $\mathcal{C}(G)$, with the left action of G on itself, can be expected to play an important role similar to the finite case. It is endowed with the inner product

$$\langle f, f'\rangle := \int_G f(g)\overline{f'(g)}dg.$$

For $R \in \mathcal{R}$, the matrix coefficients $g \to R_{i,j}(g)$ belong to unitary matrices. Theorem 2.5 establishing the orthogonality relations between the matrix coefficients of the elements of $\mathcal{R}$ remains valid; thus they form an orthogonal system in $\mathcal{C}(G)$. The celebrated Peter Weyl theorem asserts that these elements span a vector space which is dense in $\mathcal{C}(G)$ for the topology of uniform convergence.

Theorem 3.1. *[Peter Weyl theorem] The finite linear combinations of the functions $R_{i,j}$ are dense in $\mathcal{C}(G)$ for the topology of uniform convergence.*

Proof. We give a sketch of the proof:

(1) If V is a finite dimensional subspace of $\mathcal{C}(V)$ which is stable by right translation (i.e. by $gf(x) = f(xg)$) and $f \in V$, then f is a linear combination of a finite number of the $R_{i,j}$: according to previous discussion, there is a decomposition $V = W_1 \oplus \cdots \oplus W_n$ such that W_k is irreducible. If $W_k \simeq R$, there exists a basis $e_1, \ldots, e_{d_R}$ of W_k in which the action of G has matrices R. Explicitly,

$$e_j(hg) = \sum_{i=1}^{d_R} R_{i,j}(g)e_i(h).$$

Taking $h = 1$, we obtain $e_j = \sum_{i=1}^{d_R} e_i(1) R_{i,j}$.

(2) The idea is to approximate $f \in \mathcal{C}(G)$ by elements of such subspaces, constructed from the eigenspaces of a compact selfadjoint operator. We introduce the convolution operators: let $\phi \in \mathcal{C}(G)$,

$$T_\phi(f)(g) = (\phi * f)(g) = \int_G \phi(gh^{-1}) f(h) dh.$$

(3) Since G is compact, f is uniformly continuous; this property allows to choose ϕ such that $\|f - T_\phi(f)\|_\infty$ is arbitrary small.

(4) The operator T_ϕ is compact and can be assumed to be selfadjoint. The spectral theorem for such operators on Hilbert spaces (here $L^2(G)$) asserts that the eigenspaces $V_\lambda := \{f : T_\phi f = \lambda f\}$ for $\lambda \neq 0$ are finite dimensional and that the space is the direct Hilbert sum $\oplus_\lambda V_\lambda$. For $t > 0$, the subspaces $V_t := \oplus V_{\lambda,\, |\lambda| > t}$ have finite dimension (i.e. there is only a finite number of eigenvalues λ with $|\lambda| > t > 0$).

(5) The operator T_ϕ commutes with the action of G by right translation thus the subspaces V_λ are stable under this action.

(6) Let f_λ be the projection of f on V_λ. The finite sums $f_t := \sum_{|\lambda|>t} f_\lambda$ converge to $f - f_0$ for the L^2-norm when $t \to 0$.

(7) Moreover, for all $f \in \mathcal{C}(V)$, $\|T_\phi(f)\|_\infty \leq \|\phi\|_\infty \|f\|_2$. Thus, $T_\phi(f_t)$ converges *uniformly* to $T_\phi(f - f_0) = T_\phi(f)$. Finally, $T_\phi(f_t) \in V_t$ and V_t is finite dimensional and invariant under the action of G by right translations, thus by (1) $T_\phi(f_t)$ is a linear combinations of the $R_{i,j}$.

□

If $d_R = \dim(R)$, the vector space spanned by $\{\overline{R_{i,j}}, i = 1, \ldots, d_R\}$ is G-invariant and isomorphic to R. So Peter-Weyl theorem means that the decomposition of the regular decomposition is

$$\mathcal{C}(G) = \perp_{R \in \mathcal{R}} \mathcal{I}_R$$

where $\mathcal{I}_R \simeq R^{d_R}$, generalizing Theorem 2.7 (one has a better understanding of this decomposition with the action of $G \times G$ on G given by $(g, g')h = ghg'^{-1}$. For this action $\mathcal{C}(G) = \oplus_{R \in \mathcal{R}} R \otimes R^*$ where R^* is the contragredient representation, and $R \otimes R^*$ is $G \times G$-irreducible).

Since uniform convergence is stronger than L^2 convergence, we also have as a consequence of Peter Weyl theorem that the matrix coefficients $R_{i,j}$ (suitably rescaled) form an orthonormal basis of $L^2(G)$ in the sense of Hilbert spaces.

A slightly more general version of Peter Weyl theorem deals with the decomposition of $\mathcal{C}(X)$ where X is a compact space on which G acts homogeneously. If G_{x_0} is the stabilizer of a base point $x_0 \in X$, then X can be identified with the quotient space G/G_{x_0}. The Haar measure on G gives rise to a G-invariant regular measure μ on X and $\mathcal{C}(X)$ is endowed with the inner product

$$\langle f, f' \rangle := \frac{1}{\mu(X)} \int_X f(x) \overline{f'(x)} d\mu(x).$$

The space $\mathcal{C}(X)$ can be identified with the space $\mathcal{C}(G)^{G_{x_0}}$ of G_{x_0}-invariant (for the right translation) functions thus $\mathcal{C}(X)$ affords a decomposition of the form

$$\mathcal{C}(X) \simeq \perp_{R \in \mathcal{R}} R^{m_R}$$

for some integers m_R, $0 \leq m_R \leq d_R$, in the sense of uniform as well as L^2 convergence.

A more serious generalization of the above theorem deals with the unitary representations of G. These are the continuous homomorphisms from G to the unitary group of a Hilbert space.

Theorem 3.2. *Let $\pi : G \to U(H)$ be a continuous homomorphism from G to the unitary group of a Hilbert space H. Then H is a direct Hilbert sum of finite dimensional irreducible G-modules.*

Proof. The idea is to construct in H a G-subspace of finite dimension and then to iterate with the orthogonal complement of this subspace. To that end, for a fixed $v \in H$, one chooses $f \in \mathcal{C}(G)$ such that $\int_G f(g)(\pi(g)v)dg \neq 0$. From Peter Weyl theorem, f can be assumed to be a finite linear combination of the $R_{i,j}$. In other words, there exists a finite dimensional unitary representation (ρ, V) and $e_1, e_2 \in V$ such that $f(g) = \langle \rho(g^{-1})e_1, e_2\rangle_V$. The operator $T : V \to H$ defined by

$$T(x) = \int_G \langle \rho(g^{-1})x, e_2\rangle_V (\pi(g)v)dg$$

commutes with the actions of G and is non zero. Thus its image is a non zero G-subspace of finite dimension of H.

□

3.3. **Examples.**

3.3.1. **The unit sphere S^{n-1}:** it is the basic example. The orthogonal group $G = O(\mathbb{R}^n)$ acts homogeneously on S^{n-1}. The stabilizer G_{x_0} of x_0 can be identified with $O(x_0^\perp) \simeq O(\mathbb{R}^{n-1})$. Here $\mu = \omega$ is the Lebesgue measure on S^{n-1}. We set $\omega_n := \omega(S^{n-1})$. The irreducible decomposition of $\mathcal{C}(S^{n-1})$ is as follows:

$$\mathcal{C}(S^{n-1}) = H_0^n \perp H_1^n \perp \dots H_k^n \perp \dots$$

where H_k^n is isomorphic to the space Harm_k^n of harmonic polynomials:

$$\mathrm{Harm}_k^n := \{P \in \mathbb{C}[X_1, \dots, X_n]_k : \Delta P = 0, \Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}\}$$

The space Harm_k^n is a $O(\mathbb{R}^n)$-module because the Laplace operator Δ commutes with the action of the orthogonal group and it is moreover irreducible. Its dimension equals $h_k^n := \binom{n+k-1}{k} - \binom{n+k-3}{k-2}$. The embedding of Harm_k^n into $\mathcal{C}(S^{n-1})$ is the obvious one, to the corresponding polynomial function in the n coordinates.

3.3.2. **The action of stabilizers of many points:** for our purposes we are interested in the decomposition of some spaces $\mathcal{C}(X)$, X homogeneous for G, for the action of a subgroup H of G, typically $H = G_{x_1,\dots,x_s}$ the stabilizer of s points. In order to describe it, it is enough to study the decomposition of the G-irreducible submodules of $\mathcal{C}(X)$ under the action of H; thus we have to decompose only finite dimensional spaces. However, because the same irreducible representation of H may occur in infinitely many of the G-isotypic subspaces, it happens that the H-isotypic subspaces are not of finite dimension. A typical example is given by $X = S^{n-1}$, $G = O(\mathbb{R}^n)$ and $H = G_e \simeq O(\mathbb{R}^{n-1})$. It is a classical result

that for the restricted action to H the decomposition of Harm_k^n into H-irreducible subspaces is given by:

$$\mathrm{Harm}_k^n \simeq \bigoplus_{i=0}^{k} \mathrm{Harm}_i^{n-1}\,. \tag{8}$$

Hence, each of the H_k^n in (3.3.1) decomposes likewise:

$$H_k^n = H_{0,k}^n \perp H_{1,k}^n \perp \ldots \perp H_{k,k}^n$$

where $H_{i,k}^n \simeq \mathrm{Harm}_i^{n-1}$. We have the following picture, where the H-isotypic components appear to be the rows of the second decomposition.

$$\begin{array}{lllllllll}
\mathcal{C}(S^{n-1}) & =_G & H_0^n & \perp & H_1^n & \perp & \ldots & \perp & H_k^n & \perp \ldots \\
 & =_H & H_{0,0}^n & \perp & H_{0,1}^n & \perp & \ldots & \perp & H_{0,k}^n & \perp \ldots \\
 & & & \perp & H_{1,1}^n & \perp & \ldots & \perp & H_{1,k}^n & \perp \ldots \\
 & & & & & \multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\
 & & & & & & & \perp & H_{k,k}^n & \perp \ldots
\end{array}$$

4. Harmonic analysis of compact spaces

We take notations for the rest of the lecture notes. X is a compact space (possibly finite) on which a compact group (possibly finite) G acts continuously. Moreover, X is endowed with a G-invariant Borel regular measure μ for which $\mu(X)$ is finite. If X itself is finite, the topology is the discrete topology and the measure is the counting measure. In the previous sections we have discussed the decomposition of the permutation representation $\mathcal{C}(X)$. In order to lighten the notations, we assume that G has a countable number of finite dimensional irreducible representations (it is the case if G is a group of matrices over the reals since then $L^2(G)$ is a separable Hilbert space), and we let $\mathcal{R} = \{R_k, k \geq 0\}$, where R_0 is the trivial representation. We let $d_k := \dim(R_k)$. From Theorem 3.2, we have a decomposition

$$\mathcal{C}(X) \subset L^2(X) = \oplus_{k\geq 0, 1\leq i\leq m_k} H_{k,i} \tag{9}$$

where $H_{k,i} \subset \mathcal{C}(X)$, $H_{k,i} \simeq R_k$, $0 \leq m_k \leq +\infty$ (the case $m_k = 0$ means that R_k does not occur, the case $m_k = +\infty$ may occur if G is not transitive on X). The isotypic subspaces are pairwise orthogonal and denoted $\mathcal{I}_k$:

$$\mathcal{I}_k = \oplus_{i=1}^{m_k} H_{k,i}$$

We take the subspaces $H_{k,i}$ to be also pairwise orthogonal. For all k, i, we choose an orthonormal basis $e_{k,i,1}, \ldots, e_{k,i,d_k}$ of $H_{k,i}$ such that in this basis the action of $g \in G$ is expressed by the unitary matrix $R_k(g)$. The set $\{e_{k,i,s}\}$ is an orthonormal basis in the Hilbert sense.

4.1. **Commuting endomorphisms and zonal matrices.** In this subsection we want to give more information on the algebra $\mathrm{End}_G(\mathcal{C}(X))$ of commuting continuous endomorphisms of $\mathcal{C}(X)$. We introduce, for $K \in \mathcal{C}(X^2)$, the operators T_K, called Hilbert-Schmidt operators:

$$T_K(f)(x) = \frac{1}{\mu(X)} \int_X K(x,y) f(y) d\mu(y).$$

It is easy to verify that $T_K \in \mathrm{End}_G(\mathcal{C}(X))$ if K is G-invariant, i.e. if $K(gx, gy) = K(x,y)$ for all $g \in G$, $(x,y) \in X^2$. A continuous function $K(x,y)$ with this

property is also called a zonal function. It is also easy, but worth to notice that $T_K \circ T_{K'} = T_{K*K'}$ where $K * K'$ is the convolution of K and K':

$$(K * K')(x, y) := \int_X K(x, z)K'(z, y)d\mu(z).$$

Let

$$\mathcal{K} := \{K \in \mathcal{C}(X^2) : K(gx, gy) = K(x, y) \text{ for all } g \in G, (x, y) \in X^2\}.$$

The triple $(\mathcal{K}, +, *)$ is a $\mathbb{C}$-algebra (indeed a $\mathbb{C}^*$-algebra, with $K^*(x, y) := \overline{K(y, x)}$). Thus we have an embedding $\mathcal{K} \to \mathrm{End}_G(\mathcal{C}(X))$.

Assume $V \subset \mathcal{C}(X)$ is a finite dimensional G-subspace such that $V = W_1 \perp \cdots \perp W_m$ with $W_i \simeq R \in \mathcal{R}$. By the same proof as the one of Theorem 2.4, $\mathrm{End}_G(V) \simeq \mathbb{C}^{m\times m}$. More precisely, we have seen that, if $u_{j,i} : W_i \to W_j$ are G-isomorphisms, such that $u_{k,j} \circ u_{j,i} = u_{k,i}$, then an element $\phi \in \mathrm{End}_G(V)$ is associated to a matrix $A = (a_{i,j}) \in \mathbb{C}^{m\times m}$ such that, for all $f \in V$, with $p_{W_i}(f) = f_i$,

$$\phi(f) = \sum_{i,j=1}^{m} a_{j,i}u_{j,i}(f_i).$$

For all $1 \le i \le m$, let $(e_{i,1}, \dots, e_{i,d})$, $d = \dim(R)$, be an orthonormal basis of W_i such that in this basis the action of $g \in G$ is expressed by the unitary matrix $R(g)$. We define

$$E_{i,j}(x, y) := \sum_{s=1}^{d} e_{i,s}(x)\overline{e_{j,s}(y)}.$$

Then we have:

Lemma 4.1. *The above defined functions $E_{i,j}$ satisfy:*

(1) *$E_{i,j}$ is zonal: $E_{i,j}(gx, gy) = E_{i,j}(x, y)$.*
(2) *Let $T_{i,j} := T_{E_{i,j}}$. Then $T_{j,i}(W_i) = W_j$ and $T_{j,i}(W_k) = 0$ for $k \neq i$.*
(3) *$T_{i,j} \circ T_{j,k} = T_{i,k}$.*

Proof. (1) From the construction, we have

$$e_{i,s}(gx) = \sum_{t=1}^{d} \overline{R_{s,t}(g)}e_{i,t}(x)$$

thus

$$\begin{aligned}
E_{i,j}(gx, gy) &= \sum_{s=1}^{d} e_{i,s}(gx)\overline{e_{j,s}(gy)} \\
&= \sum_{s=1}^{d}\sum_{k,l=1}^{d} \overline{R_{s,k}(g)}R_{s,l}(g)e_{i,k}(x)\overline{e_{j,l}(y)} \\
&= \sum_{k,l=1}^{d} \Big(\sum_{s=1}^{d} \overline{R_{s,k}(g)}R_{s,l}(g)\Big)e_{i,k}(x)\overline{e_{j,l}(y)} \\
&= \sum_{k}^{d} e_{i,k}(x)\overline{e_{j,k}(y)} = E_{i,j}(x, y)
\end{aligned}$$

where the second last equality holds because $R(g)$ is a unitary matrix.

(2) We compute $T_{j,i}(e_{k,t})$:

$$\begin{aligned} T_{j,i}(e_{k,t})(x) &= \frac{1}{\mu(X)} \int_X \big(\sum_{s=1}^{d} e_{j,s}(x)\overline{e_{i,s}(y)}\big) e_{k,t}(y) d\mu(y) \\ &= \frac{1}{\mu(X)} \sum_{s=1}^{d} e_{j,s}(x) \int_X \overline{e_{i,s}(y)} e_{k,t}(y) d\mu(y) \\ &= \sum_{s=1}^{d} e_{j,s}(x) \langle e_{k,t}, e_{i,s} \rangle \\ &= \sum_{s=1}^{d} e_{j,s}(x) \delta_{k,i} \delta_{t,s} = \delta_{k,i} e_{j,t}(x). \end{aligned}$$

(3) Similarly one computes that

$$E_{i,j} * E_{l,k} = \delta_{j,l} E_{i,k}.$$

□

The $E_{i,j}(x,y)$ put together form a matrix $E = E(x,y)$, that we call the zonal matrix associated to the G-subspace V:

$$E(x,y) := \big(E_{i,j}(x,y)\big)_{1 \le i,j \le m}. \tag{10}$$

At this stage is is natural to discuss the dependence of this matrix on the various ingredients needed for its definition.

Lemma 4.2. *We have*

(1) *$E(x,y)$ is unchanged if another orthonormal basis of W_i is chosen (i.e. if another unitary representative of the irreducible representation R is chosen).*

(2) *$E(x,y)$ is changed to $AE(x,y)A^*$ for some matrix $A \in \mathrm{Gl}(\mathbb{C}^m)$ if another decomposition (not necessarily with orthogonal spaces) $V = W_1' \oplus \cdots \oplus W_m'$ is chosen.*

Proof. (1) Let $(e_{i,1}', \ldots, e_{i,d}')$ be another orthonormal basis of W_i and let U_i be unitary $d \times d$ matrices such that

$$(e_{i,1}', \ldots, e_{i,d}') = (e_{i,1}, \ldots, e_{i,d}) U_i.$$

Since we want the representation R to be realized by the same matrices in the basis $(e_{i,1}', \ldots, e_{i,d}')$ when i varies, we have $U_i = U_j = U$. Then, with obvious notations,

$$\begin{aligned} E_{i,j}'(x,y) =& (e_{i,1}'(x), \ldots, e_{i,d}'(x))(e_{i,1}'(y), \ldots, e_{i,d}'(y))^* \\ =& (e_{i,1}(x), \ldots, e_{i,d}(x)) UU^* (e_{i,1}(y), \ldots, e_{i,d}(y))^* \\ =& (e_{i,1}(x), \ldots, e_{i,d}(x))(e_{i,1}(y), \ldots, e_{i,d}(y))^* \\ =& E_{i,j}(x,y). \end{aligned}$$

(2) If $V = W_1 \perp \cdots \perp W_m = W_1' \perp \cdots \perp W_m'$ with basis $(e_{i,1}, \ldots, e_{i,d})$ of W_i and $(e_{i,1}', \ldots, e_{i,d}')$ of W_i' in which the action of G is by the same matrices $R(g)$, let $\phi \in \mathrm{End}(V)$ be defined by $\phi(e_{i,s}) = e_{i,s}'$. Clearly

ϕ commutes with the action of G; if $u_{j,i}$ is defined by $u_{j,i}(e_{i,s}) = e_{j,s}$ then we have seen that, for some matrix $A = (a_{i,j})$, $e'_{i,s} = \phi(e_{i,s}) = \sum_{j=1}^m a_{j,i} e_{j,s}$. Moreover A is invertible. It is unitary if the spaces W'_i are pairwise orthogonal. With the notations $E(x) := (e_{i,s}(x))$, we have

$$E(x,y) = E(x)E(y)^* \text{ and } E'(x) = A^t E(x)$$

thus

$$E'(x,y) = A^t E(x,y)\overline{A}.$$

□

Going back to $\phi \in \mathrm{End}_G(V)$, from Lemma 4.1 we can take $u_{j,i} = T_{j,i}$ and we have the expression

$$\phi = \sum_{i,j=1}^m a_{j,i} T_{j,i} = T_{\langle A, \overline{E}\rangle}.$$

We take the following notation: the space of linear combinations of elements of the form $f(x)\overline{g(y)}$ for $(f,g) \in V^2$ is denoted $V^{(2)}$. We have proved the following:

Proposition 4.3. *Let*

$$\mathcal{K}_V := \{K \in V^{(2)} : K(gx, gy) = K(x,y) \text{ for all } g \in G, (x,y) \in X^2\}.$$

The following are isomorphisms of $\mathbb{C}$-algebras:

$$\begin{array}{rl} \mathcal{K}_V & \to \mathrm{End}_G(V) \\ K & \mapsto T_K \end{array} \qquad \begin{array}{rl} \mathbb{C}^{m\times m} & \to \mathrm{End}_G(V) \\ A & \mapsto T_{\langle A, \overline{E}\rangle}. \end{array}$$

Moreover, $\mathrm{End}_G(\mathcal{C}(X))$ is commutative iff $\mathcal{K}$ is commutative iff $m_k = 0, 1$ for all $k \geq 0$.

Proof. The isomorphisms are clear from previous discussion. For the last assertion, it is enough to point out that

$$\mathrm{End}_G(\mathcal{C}(X)) = \prod_{k\geq 0} \mathrm{End}_G(\mathcal{I}_k).$$

□

Remark 4.4. *Proposition 4.3 shows in particular that $\mathcal{K}_V$ and $\mathrm{End}_G(V)$ have the same dimension. It is sometimes easy to calculate the dimension of $\mathcal{K}_V$; for example if X is a finite set and $V = \mathcal{C}(X)$, then $\dim(\mathcal{K}_V)$ is exactly equal to the number of orbits of G acting on X^2. On the other hand, in this case, the dimension of $\mathrm{End}_G(V)$ is the sum of the squares of the multiplicities in $\mathcal{C}(X)$ of the irreducible representations of G. For the binary Hamming space treated in 2.8.1, the orbits of G acting on X^2 are in one to one correspondance with the values taken by the Hamming distance, i.e. there are $(n+1)$ such orbits. Thus, once we have obtained the decomposition $\mathcal{C}(H_n) = P_0 \perp \cdots \perp P_n$, because this decomposition involves allready $(n+1)$ subspaces, we can conclude readily that these subspaces are irreducible. This reasoning applies also to the Johnson space 2.8.2 and to the more general q-Hamming space 5.3. A variant of this method is as follows: if we suspect $V \subset \mathcal{C}(X)$ to be irreducible, then it is enough to prove that $\mathcal{K}_V$ has dimension 1. See in 5.3.1 for an illustration.*

4.2. **Examples: G-symmetric spaces.**

Definition 4.5. *We say that X is G-symmetric if for all $(x, y) \in X^2$, there exists $g \in G$ such that $gx = y$ and $gy = x$. In other words, (x, y) and (y, x) belong to the same orbit of G acting on X^2.*

A first consequence of Proposition 4.3 is that G-symmetric spaces have multiplicity free decompositions.

Proposition 4.6. *If X is G-symmetric then $m_k = 0, 1$ for all $k \geq 0$ and $E_k(x, y)$ is real symmetric.*

Proof. For all $K \in \mathcal{K}$, $K(x, y) = K(y, x)$. Thus $\mathcal{K}$ is commutative: indeed,

$$\begin{aligned}(K' * K)(x, y) &= \frac{1}{\mu(X)} \int_X K'(x, z)K(z, y)d\mu(z) \\ &= \frac{1}{\mu(X)} \int_X K'(z, x)K(y, z)d\mu(z) \\ &= (K * K')(y, x) = (K * K')(x, y).\end{aligned}$$

Moreover $\overline{E_k(x, y)} = E_k(x, y) = E_k(y, x)$. □

4.2.1. 2-**point homogeneous spaces:** these spaces are prominent examples of G-symmetric spaces.

Definition 4.7. *A metric spaces (X, d) is said to be 2-point homogeneous for the action of G if G is transitive on X, leaves the distance d invariant, and if, for $(x, y) \in X^2$,*

$$\textit{there exists } g \in G \textit{ such that } (gx, gy) = (x', y') \iff d(x, y) = d(x', y').$$

Examples of such spaces of interest in coding theory are numerous: the Hamming and Johnson spaces, endowed with the Hamming distance, for the action of respectively $T \rtimes S_n$ and S_n; the unit sphere S^{n-1} for the angular distance $\theta(x, y)$ and the action of the orthogonal group. It is a classical result that, apart from S^{n-1}, the projective spaces $\mathbb{P}^n(K)$ for $K = \mathbb{R}, \mathbb{C}, \mathbb{H}$, and $\mathbb{P}^2(\mathbb{O})$, are the only real compact 2-point homogeneous spaces.

There are more examples of finite 2-point homogeneous spaces, we can mention among them the q-Johnson spaces. The q-Johnson space $J_n^w(q)$ is the set of linear subspaces of $\mathbb{F}_q^n$ of fixed dimension w, with the action of the group $\mathrm{Gl}(\mathbb{F}_q^n)$ and the distance $d(x, y) = \dim(x + y) - \dim(x \cap y)$. We come back to this space in the next section.

There are other symmetric spaces occurring in coding theory:

4.2.2. **The Grassmann spaces:** $X = \mathcal{G}_{m,n}(K)$, $K = \mathbb{R}, \mathbb{C}$, i.e. the set of m-dimensional linear subspaces of K^n, with the homogeneous action of $G = O(\mathbb{R}^n)$ (respectively $U(\mathbb{C}^n)$). This space is G-symmetric but not 2-point homogeneous (if $m \geq 2$). The orbits of G acting on pairs $(p, q) \in X^2$ are characterized by their principal angles [21]. The principal angles of (p, q) are m angles $(\theta_1, \ldots, \theta_m) \in [0, \pi/2]^m$ constructed as follows: one iteratively constructs an orthonormal basis $(e_1, \ldots, e_m)$ of p and an orthonormal basis $(f_1, \ldots, f_m)$ of q such that, for $1 \leq$

$i \leq m$,

$$\begin{aligned}\cos\theta_i &= \max\{|(e,f)| \ : \ & e \in p,\ f \in q,\\ & & (e,e) = (f,f) = 1,\\ & & (e,e_j) = (f,f_j) = 0 \text{ for } 1 \leq j \leq i-1\}\\ &= |(e_i,f_i)|\end{aligned}$$

The we have (see [21]):

$$\begin{gathered}\text{there exists } g \in G \text{ such that } (gp,gq) = (p',q')\\ \Longleftrightarrow\\ (\theta_1(p,q),\dots,\theta_m(p,q)) = (\theta_1(p',q'),\dots,\theta_m(p',q')).\end{gathered}$$

4.2.3. **The ordered Hamming space:** $X = (\mathbb{F}_2^r)^n$ (for the sake of simplicity we restrict here to the binary case). Let $x = (x_1,\dots,x_n) \in X$ with $x_i \in \mathbb{F}_2^r$. For $y \in \mathbb{F}_2^r$, the ordered weight of y, denoted $w_r(y)$, is the right most non zero coordinate of y. The ordered weight of $x \in X$ is $w_r(x) := \sum_{i=1}^n w_r(x_i)$ and the ordered distance of two elements $(x,y) \in X^2$ is $d_r(x,y) = w_r(x-y)$. Moreover we define the shape of (x,y):

$$\mathrm{shape}(x,y) := (e_0,e_1,\dots,e_r) \text{ where } \begin{cases} 1 \leq i \leq r, e_i := \mathrm{card}\{j \ : \ w_r(x_j) = i\}\\ e_0 := n - (e_1 + \cdots + e_r).\end{cases}$$

Another expression of $w_r(x)$ is $w_r(x) = \sum_i ie_i$.

If B is the group of upper triangular matrices in $\mathrm{Gl}(\mathbb{F}_2^r)$, and B_{aff} the group of affine transformations of $\mathbb{F}_2^r$ combining the translations by elements of $\mathbb{F}_2^r$ with B, the group $G := B_{\text{aff}}^n \rtimes S_n$ acts transitively on X. Since B acting on $\mathbb{F}_2^r$ leaves w_r invariant, it is clear that the action of G on X leaves the shape $\mathrm{shape}(x,y)$ invariant. More precisely, the orbits of B on $\mathbb{F}_2^r$ are the sets $\{y \in \mathbb{F}_2^r : w_r(x) = i\}$ and, consequently, the orbits of G acting on X^2 are characterized by the so-called shape of (x,y). Since obviously $\mathrm{shape}(x,y) = \mathrm{shape}(y,x)$ it is a symmetric space. This space shares many common features with the Grassmann spaces, especially from the point of view of the linear programming method (see [2], [9], [31]).

4.2.4. **The space $X = \Gamma$ under the action of $G = \Gamma \times \Gamma$:** the action of G is by $(\gamma,\gamma')x = \gamma x \gamma'^{-1}$. Then two pairs (x,y) and (x',y') are in the same orbit under the action of G iff xy^{-1} and $x'y'^{-1}$ are in the same conjugacy class of Γ. Obviously (x,y) and (y^{-1},x^{-1}) are in the same G-orbit. We are not quite in the case of a G-symmetric space however the proof of the commutativity of $\mathcal{K}$ of Proposition 4.6 remains valid because the variable change $x \to x^{-1}$ leaves the Haar measure invariant.

4.3. **Positive definite functions and Bochner theorem.**

Definition 4.8. *A positive definite continuous function on X is a function $F \in \mathcal{C}(X^2)$ such that $F(x,y) = \overline{F(y,x)}$ and one of the following equivalent properties hold:*

(1) *For all n, for all $(x_1,\dots,x_n) \in X^n$, for all $(\alpha_1,\dots,\alpha_n) \in \mathbb{C}^n$,*

$$\sum_{i,j=1}^n \alpha_i F(x_i,x_j)\overline{\alpha_j} \geq 0.$$

(2) *For all* $\alpha \in \mathcal{C}(X)$,

$$\int_{X^2} \alpha(x)F(x,y)\overline{\alpha(y)}d\mu(x,y) \geq 0.$$

This property will be denoted $F \succeq 0$.

The first property means in other words that, for all choice of a finite set of points $(x_1, \ldots, x_n) \in X^n$, the matrix $(F(x_i, x_j))_{1\leq i,j\leq n}$ is hermitian positive semidefinite. The equivalence of the two properties results from compactness of X. Note that, if X is finite, F is positive definite iff the matrix indexed by X, with coefficients $F(x, y)$, is positive semidefinite.

We want to characterize those functions which are G-invariant. This characterization is provided by Bochner in [11] in the case when the space X is G-homogeneous. It is clear that the construction of previous subsection provides positive definite functions. Indeed,

Lemma 4.9. *if* $A \succeq 0$, *then* $\langle A, \overline{E}\rangle$ *is a* G*-invariant positive definite function.*

Proof. Let $\alpha(x) \in \mathcal{C}(X)$. We compute

$$\begin{aligned}
\int_{X^2} \alpha(x)\langle A, \overline{E}\rangle\overline{\alpha(y)}d\mu(x,y) &= \int_{X^2} \sum_{i,j=1}^{m} A_{i,j}\alpha(x)E_{i,j}(x,y)\overline{\alpha(y)}d\mu(x,y) \\
&= \sum_{i,j=1}^{m} A_{i,j} \int_{X^2} \alpha(x)E_{i,j}(x,y)\overline{\alpha(y)}d\mu(x,y) \\
&= \sum_{i,j=1}^{m}\sum_{s=1}^{d} A_{i,j} \int_{X^2} \alpha(x)e_{i,s}(x)\overline{e_{j,s}(y)\alpha(y)}d\mu(x,y) \\
&= \sum_{i,j=1}^{m}\sum_{s=1}^{d} A_{i,j}\alpha_{i,s}\overline{\alpha_{j,s}} \\
&= \sum_{s=1}^{d}\sum_{i,j=1}^{m} \alpha_{i,s}A_{i,j}\overline{\alpha_{j,s}} \geq 0
\end{aligned}$$

where $\alpha_{i,s} := \int_X \alpha(x)e_{i,s}(x)d\mu(x)$. □

Remark 4.10. *The following properties are equivalent, for a* $m \times m$ *matrix function* $E(x, y)$*:*

(1) *For all* $A \succeq 0$, $\langle A, \overline{E(x,y)}\rangle \succeq 0$
(2) *For all* $(x_1, \ldots, x_n) \in X^n$, $(\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$, $\sum_{i,j} \alpha_i E(x_i, x_j)\overline{\alpha_j} \succeq 0$.

The proof is left to the reader as an exercise (hint: use the fact that the cone of positive semidefinite matrices is self dual).

To start with, we extend the notations of the previous subsection. We define matrices $E_k = E_k(x, y)$ associated to each isotypic component $\mathcal{I}_k$, of size $m_k \times m_k$ (thus possibly of infinite size) with coefficients $E_{k,i,j}(x, y)$ defined by:

$$E_{k,i,j}(x,y) := \sum_{s=1}^{d_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}.$$

If $F_k = (f_{k,i,j})_{1\leq i,j\leq m_k}$ is hermitian, and if $\sum_{i,j} |f_{k,i,j}|^2 < +\infty$, the sum

$$\langle F_k, \overline{E_k}\rangle := \sum_{i,j} f_{k,i,j} E_{k,i,j}$$

is L^2-convergent since the elements $e_{k,i,s}(x)\overline{e_{l,j,t}(y)}$ form a complete system of orthonormal elements of $\mathcal{C}(X^2)$. We say F_k is positive semidefinite ($F_k \succeq 0$) if $\sum_{i,j} \overline{\lambda_i} f_{k,i,j}\lambda_j \geq 0$ for all $(\lambda_i)_{1\leq i\leq m_k}$ such that $\sum |\lambda_i|^2 < +\infty$. Then, with the same proof as the one of Lemma 4.9, the function $\langle F_k, \overline{E_k}\rangle$ is positive definite if $F_k \succeq 0$. The following theorem provides a converse statement (see [11]).

Theorem 4.11. *$F \in \mathcal{C}(X^2)$ is a G-invariant positive definite function if and only if*

$$F(x,y) = \sum_{k\geq 0} \langle F_k, \overline{E_k(x,y)}\rangle \tag{11}$$

where, for all $k \geq 0$,

$$F_k = \frac{1}{d_k \mu(X^2)} \int_{X^2} F(x,y)\overline{E_k(x,y)} d\mu(x,y) \succeq 0,$$

and the sum converges to F for the L^2 topology. If moreover G acts homogeneously on X, the sum (11) *itself converges uniformly.*

Proof. The elements $e_{k,i,s}(x)\overline{e_{l,j,t}(y)}$ form a complete system of orthonormal elements of $\mathcal{C}(X^2)$. Hence F has a decomposition

$$F(x,y) = \sum_{k,i,s,l,j,t} f_{k,i,s,l,j,t} e_{k,i,s}(x)\overline{e_{l,j,t}(y)}$$

where the convergence of the sum is L^2. The condition $F(gx, gy) = F(x,y)$ translates to:

$$f_{k,i,u,l,j,v} = \sum_{s,t} f_{k,i,s,l,j,t} R_{k,u,s}(g)\overline{R_{l,v,t}(g)}.$$

Integrating on $g \in G$ and applying the orthogonality relations of Theorem 2.5 shows that $f_{k,i,u,l,j,v} = 0$ if $k \neq l$ or $u \neq v$. Moreover it shows that $f_{k,i,u,k,j,u}$ does not depend on u. The resulting expression of F reads:

$$F(x,y) = \sum_{k\geq 0} \Big(\sum_{i,j} f_{k,i,j} E_{k,i,j}(x,y) \Big)$$

and

$$d_k f_{k,i,j} = \frac{1}{\mu(X^2)} \int_{X^2} F(x,y)\overline{E_{k,i,j}(x,y)} d\mu(x,y),$$

which is the wanted expression, with $F_k := (f_{k,i,j})_{1\leq i,j\leq m_k}$.

Now we show that $F_k \succeq 0$. Let, for k, s fixed, $\alpha(x) = \sum_i \alpha_i \overline{e_{k,i,s}(x)}$, with $\sum_i |\alpha_i|^2 < +\infty$. By density, property (2) of Definition 4.8 holds for $\alpha \in L^2(X)$. We compute like in the proof of Lemma 4.9

$$\int_{X^2} \alpha(x) F(x,y)\overline{\alpha(y)} d\mu(x,y) = \sum_{i,j=1}^{m_k} \alpha_i f_{k,i,j} \overline{\alpha_j}$$

thus $F_k \succeq 0$.

In the case of X being G-homogeneous, the uniform convergence of the sum in (11) is proved in [11].

□

In order to reduce linear programs involving G-invariant positive definite functions to finite dimensional semidefinite programs, we need to be able to approximate such functions uniformly with finite sums of the type (11), in other words by functions built form finite dimensional subspaces of $\mathcal{C}(X)$. A necessary condition is thus that all continuous functions on X are uniformly approximated by elements of some sequence of finite dimensional subspaces of $\mathcal{C}(X)$. Such subspaces are usually provided by the polynomial functions of bounded degree, when it makes sense. More generally, let us assume that there exists a sequence $(V_d)_{d\geq 0}$ of finite dimensional G-subspaces of $\mathcal{C}(X)$ such that $V_d \subset V_{d+1}$, and $\cup_{d\geq 0} V_d$ is dense in $\mathcal{C}(X)$ for the topology of uniform convergence. For example, Peter-Weyl theorem provides such subspaces when X is Γ-homogeneous, for a compact group Γ containing G. Then we have the following result:

Theorem 4.12. *Under the above assumptions, if moreover X is homogeneous under a larger compact group Γ, and if the irreducible subspaces $H_{k,i}$ are chosen so that $H_{k,i} \subset V_d$ for all $1 \leq i \leq m_{d,k}$ where $m_{d,k}$ is the multiplicity of R_k in V_d, then a G-invariant positive definite function $F \in \mathcal{C}(X^2)$ is the uniform limit of a sequence of positive definite functions $F_d \in V_d \otimes V_d$ thus of the form*

$$F_d(x,y) = \sum_{k\geq 0} \langle F_{d,k}, \overline{E_k(x,y)} \rangle \tag{12}$$

where $F_{d,k}$ is a matrix of size $m_{d,k}$ (and thus the sum has a finite number of non zero terms).

Proof. We proceed like in the proof of Peter Weyl theorem. Compact self-adjoint Hilbert-Schmidt operators on $\mathcal{C}(X^2)$ are of the form

$$T_K(F)(x,y) = \int_{X^2} K((x,y),(z,t))F(z,t)d\mu(z,t).$$

We start to construct K such $T_K(F) \succeq 0$ and $\|T_K(F) - F\|_\infty$ is arbitrary small. The first condition is fulfilled if K can be expressed in the form $K((x,y),(z,t)) = K_0(x,z)\overline{K_0(y,t)}$ where $K_0(x,z) = \overline{K_0(z,x)}$. We take ϕ_0 a continuous function on Γ; if ϕ_0' denotes the left and right average of ϕ_0 over Γ_0 (where $X = \Gamma/\Gamma_0$), we take $K_0(x,y) = \phi_0'(\gamma^{-1}\delta)$ for any $\gamma \in x$, $\delta \in y$). Then with a suitable choice of ϕ_0, $\|T_K(F) - F\|_\infty \leq \epsilon$ (thanks to uniform continuity of F, it is enough that ϕ_0 has support contained in some prescribed open neighborhood of 1, takes values between 0 and 1, satisfies $\phi_0(\gamma) = \phi_0(\gamma^{-1})$ and $\int_\Gamma \phi_0 = |\Gamma_0|$). Moreover, K_0 is Γ-invariant.

We can find $d \geq 0$ and $L_0(x,y) \in V_d \otimes V_d$ such that $L_0(x,z) = \overline{L_0(z,x)}$ and $\|L_0 - K_0\|_\infty$ is arbitrary small. Replacing L_0 by its average on G will not change these three properties of L_0. Then, if $L((x,y),(z,t)) := L_0(x,z)\overline{L_0(y,t)}$, $T_L(F)$ comes arbitrary close to $T_K(F)$ for $\| \|_\infty$ and $T_L(F) \in V_d \otimes V_d$. Now, $T_L(F) \succeq 0$, is invariant under G and belongs to the finite dimensional space $V_d \otimes V_d$ thus it has the announced form from Theorem 4.11.

□

Now the main deal is to compute explicitly the matrices $E_k(x,y)$ for a given space X. The next section gives explicit examples of such computation.

5. Explicit computations of the matrices $E_k(x,y)$

We keep the same notations as in previous section. Since the matrices $E_k(x,y)$ are G-invariant, their coefficients are functions of the orbits of G acting on X^2. So the first task is to describe these orbits. Let us assume that these orbits are parametrized by some variables $u=(u_i)$. Then we seek for explicit expressions of the form

$$E_k(x,y)=Y_k(u(x,y)).$$

The measure μ induces a measure on the variables that describe these orbits, for which the coefficients of E_k are pairwise orthogonal. This property of orthogonality turns to be very useful, if not enough, to calculate the matrices E_k.

The easiest case is when the space X is 2-point homogeneous for the action of G, because in this case the orbits of pairs are parametrized by a single variable $t:=d(x,y)$. Moreover we have already seen that in this case, the decomposition of $\mathcal{C}(X)$ is multiplicity free so the matrices $E_k(x,y)$ have a single coefficient.

5.1. 2-**point homogeneous spaces.** We summarize the results we have obtained so far:

$$\mathcal{C}(X)=\oplus_{k\geq 0}H_k$$

where H_k are pairwise orthogonal G-irreducible subspaces; to each H_k is associated a continuous function $P_k(t)$ such that $E_k(x,y)=P_k(d(x,y))$ and

$$F\succeq 0\Longleftrightarrow F=\sum_{k\geq 0}f_kP_k(d(x,y)) \text{ with } f_k\geq 0.$$

$P_k(t)$ is called the zonal function associated to H_k. Since the subspaces H_k are pairwise orthogonal, the functions $P_k(t)$ are pairwise orthogonal for the induced measure. This property of orthogonality is in general enough to determine them in a unique way. We can also notice here that $P_k(0)=d_k$. This value is obtained with the integration on X of the formula $P_k(0)=\sum_{s=1}^{d_k}e_{k,1,s}(x)\overline{e_{k,1,s}(x)}$.

5.2. $X=\{1,\dots,q\}$ **under the action of** S_q**.** This is a very easy case, which will play a role in the study of the q-Hamming space. Since the constant function $\mathbf{1}$ is S_q-invariant, we have the S_q decomposition $\mathcal{C}(X)=\mathbb{C}\,\mathbf{1}\perp L$. Obviously, the action of S_q on X^2 has two orbits: the set of pairs (i,i), and the set of pairs (i,j) for $i\neq j$. Thus, from Proposition 4.3 and Remark 4.4, L is irreducible. We let $z_0:=\mathbf{1}$ and choose an orthonormal basis $(z_1,\dots,z_{q-1})$ of L. We want to compute the zonal function E_L associated to L. We have by definition $E_L(x,y)=\sum_{i=1}^{q-1}z_i(x)\overline{z_i(y)}$ and E_L takes only two different values: one for $x=y$ and one for $x\neq y$. We have $E_L(0,0)=\dim(L)=q-1$ and we can compute $E_L(0,1)$ easily using the fact that $E_L(0,y)$ is orthogonal to z_0 thus $\sum_{y=1}^{q}E_L(0,y)=0=E_L(0,0)+(q-1)E_L(0,1)$. Thus $E_L(0,1)=-1$.

5.3. **The** q**-Hamming space.** In the binary case we have already calculated the functions $P_k(t)$ in 2.8.1. Indeed, the irreducible subspaces P_k afford the orthonormal basis $\{\chi_z, wt(z)=k\}$. So,

$$E_k(x,y)=\sum_{wt(z)=k}\chi_z(x)\chi_z(y)=\sum_{wt(z)=k}(-1)^{z\cdot(x+y)}=K_k(d_H(x,y))$$

from (4). Now we treat the more general q-Hamming space. This is the space $H_{n,q}=F^n$ where F is a finite set with q elements denoted $F=\{a_0,a_1,\dots,a_{q-1}\}$.

The semidirect product $G = S_q^n \rtimes S_n$ acts on $H_{n,q}$ and leaves the Hamming distance invariant. Here the permutation group S_q acts on F by $\tau a_i = a_{\tau(i)}$ while the permutation group S_n acts on $H_{n,q}$ by $\sigma(x_1, \dots, x_n) = (x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$. Moreover G acts on $H_{n,q}$ 2-point homogeneously. The action of S_q on $\mathcal{C}(F)$ is studied in 5.2 and we take the same notations. We define $\phi = (\phi_1, \dots, \phi_n) \in \mathcal{C}(H_{n,q})$ where $\phi_i \in \{z_0, z_1, \dots, z_{q-1}\}$ by: $\phi(x) = \prod_{i=1}^n \phi_i(x_i)$. These elements ϕ form an orthonormal system: it is easy to see that

$$\langle \phi, \psi \rangle = \prod_{i=1}^n \langle \phi_i, \psi_i \rangle.$$

We define the weight of ϕ by: $wt(\phi) := |\{1 \le i \le n : \phi_i \neq z_0\}|$. For $0 \le k \le n$, let P_k be the subspace generated by the set of ϕ with $wt(\phi) = k$. The dimension of P_k is the number of such ϕ, which is equal to $(q-1)^k \binom{n}{k}$ and we have the decomposition

$$\mathcal{C}(H_n) = P_0 \perp P_1 \perp \cdots \perp P_n. \tag{13}$$

An element $\tau \in S_q$ act trivially on z_0 and sends z_i for $i \neq 0$ to a linear combination of $z_1, \dots, z_{q-1}$. Thus for all $g \in G$, $g\phi$ is a linear combination of ψ's with the same weight as ϕ and G stabilizes P_k. The action of G on pairs of elements of $H_{n,q}$ has exactly $(n+1)$ orbits corresponding to the $(n+1)$ values $0, 1, \dots, n$ that the Hamming distance takes thus we can conclude that P_k is irreducible from Proposition 4.3. Now we compute the zonal function $E_k(x, y)$ attached to P_k. By definition we have

$$E_k(x, y) = \sum_{\phi, wt(\phi)=k} \phi(x)\overline{\phi(y)}$$

and we want to calculate P_k such that $P_k(t) = E_k(x, y)$ for any (x, y) with $d(x, y) = t$. We set $x = (a_1, \dots, a_1, a_0, \dots, a_0)$ where t coordinates of x are equal to a_1 and $y = (a_0, \dots, a_0)$. For all ϕ, we let $i := |\{j : 1 \le j \le n : x_j = a_1 \text{ and } \phi_j \neq z_0\}$ and reorder the set of $\phi \in P_k$ according to i.

$$\begin{aligned}
P_k(t) &= \sum_{i=0}^k \binom{t}{i}\binom{n-t}{k-i} \sum_{j_1, \dots, j_k \neq 0} \prod_{u=1}^i z_{j_u}(a_1)\overline{z_{j_u}(a_0)} \prod_{u=i+1}^k z_{j_u}(a_0)\overline{z_{j_u}(a_0)} \\
&= \sum_{i=0}^k \binom{t}{i}\binom{n-t}{k-i} \Big(\sum_{s=1}^{q-1} z_s(a_1)\overline{z_s(a_0)}\Big)^i \Big(\sum_{s=1}^{q-1} z_s(a_0)\overline{z_s(a_0)}\Big)^{k-i} \\
&= \sum_{i=0}^k \binom{t}{i}\binom{n-t}{k-i} E_L(a_1, a_0)^i E_L(a_0, a_0)^{k-i} \\
&= \sum_{i=0}^k \binom{t}{i}\binom{n-t}{k-i} (-1)^i (q-1)^{k-i}
\end{aligned}$$

with the notations and results of 5.2. $P_k(t)$ is equal to the Krawtchouck polynomial $K_k^{n,q}(t)$ of parameters q and n which satisfies the following characteristic properties:

(1) $\deg(K_k^{n,q}) = k$
(2) $K_k^{n,q}(0) = (q-1)^k \binom{n}{k}$

(3) Orthogonality relations: for all $0 \leq k \leq l \leq n$

$$\frac{1}{q^n}\sum_{w=0}^{n}\binom{n}{w}K_k^{n,q}(w)K_l^{n,q}(w) = \delta_{k,l}\binom{n}{k}(q-1)^k.$$

The orthogonality relations are direct consequences of the orthogonality of the subspaces P_k.

5.3.1. **The Johnson space J_n^w:** with the notations of subsection 2.8.2, we have shown the decomposition

$$\mathcal{C}(J_n^w) \simeq H_w \perp H_{w-1} \perp \cdots \perp H_0$$

but not yet the irreducibility of H_i. So far their might by several $P_{i,j}$, $j = 1, \dots$ associated to H_i. The zonal functions express as functions of $t := |x \cap y|$ the number of common ones in x and y. The orthogonality relation is easy to compute:

$$\begin{aligned}\sum_{x\in X} f(|x\cap y|)\overline{f'(|x\cap y|)} &= \sum_{i=0}^{n}\mathrm{card}\{x \ : \ |x\cap y| = i\}f(i)\overline{f'(i)}\\ &= \sum_{i=0}^{w}\binom{w}{i}\binom{n-w}{w-i}f(i)\overline{f'(i)}\\ &= \sum_{i=0}^{w}\binom{w}{i}\binom{n-w}{i}f(w-i)\overline{f'(w-i)}.\end{aligned}$$

By induction on k one proves that $P_{k,j}$ has degree at most k in t. The conditions:

(1) $\deg(Q_k) = k$
(2) $Q_k(0) = 1$
(3) for all $0 \leq k < l \leq n$

$$\sum_{i=0}^{w}\binom{w}{i}\binom{n-w}{i}Q_k(i)Q_l(i) = 0$$

determine a unique sequence $(Q_0, Q_1, \dots, Q_w)$. Thus there is only one $P_{k,j}$ for each k and it is equal to $h_kQ_k(w-t)$. The polynomials Q_k defined above belong to the family of Hahn polynomials.

5.3.2. **The sphere S^{n-1}:** the distance on the sphere is the angular distance $\theta(x,y)$. It appears more convenient to express the functions in the variable $t = x\cdot y = \cos\theta(x,y)$. A standard calculation shows that

$$\int_{S^{n-1}} f(x\cdot y)d\mu(y) = c_n\int_{-1}^{1} f(t)(1-t^2)^{\frac{n-3}{2}}dt$$

for some irrelevant constant c_n. The conditions:

- $\deg(P_k^n) = k$
- $P_k^n(1) = 1$
- For all $k \neq l$, $\int_{-1}^{1} P_k^n(t)P_l^n(t)(1-t^2)^{\frac{n-3}{2}}dt = 0$

define a unique sequence of polynomials by standard arguments (i.e. obtained by Gram Schmidt orthogonalization of the basis $(1, t, \dots, t^k, \dots)$), it is the sequence of so-called Gegenbauer polynomials with parameter $n/2-1$ [43]. The decomposition 3.3.1 of $\mathcal{C}(S^{n-1})$ shows that, to each $k \geq 0$ the function $P_k(x\cdot y)$ associated

to $H_k^n \simeq \mathrm{Harm}_k^n$ is polynomial in $x \cdot y$ and satisfies the above conditions except the normalization of $P_k(1)$ thus we have $P_k(t) = h_k^n P_k^n(t)$.

5.3.3. **Other 2-point homogeneous spaces:** as it is shown in the above examples, a sequence of orthogonal polynomials in one variable is associated to each such space. In the case of the projective spaces, it is a sequence of Jacobi polynomials. We refer to [24], [28], [48] for their determination in many cases and for the applications to coding theory.

5.4. **Other symmetric spaces.** Now we turn to other cases of interest in coding theory, where the space X is symmetric but not necessarily 2-point homogeneous. Since the decomposition of $\mathcal{C}(X)$ is multiplicity free, the matrices $E_k(x, y)$ still have a single coefficient which is a member of a sequence of orthogonal polynomials, but this time multivariate. The first case ever studied (at least to my knowledge) is the case of the non binary Johnson spaces [44], its associated functions are two variables polynomials, a mixture of Hahn and Eberlein polynomials. We briefly discuss a few of these cases.

5.4.1. **The Grassmann spaces:** [2] the orbits of X^2 are parametrized by the principal angles $(\theta_1, \ldots, \theta_m)$ (4.2.2). The appropriate variables are the $y_i := \cos^2 \theta_i$. The decomposition of $\mathcal{C}(\mathcal{G}_{m,n})$ under $O(\mathbb{R}^n)$ (respectively $U(\mathbb{C}^n)$) together with the computation of the corresponding sequence of orthogonal polynomials was performed in [23]. We focus here on the real case. We recall that the irreducible representations of $\mathrm{O}(\mathbb{R}^n)$ are (up to a power of the determinant) naturally indexed by partitions $\kappa = (\kappa_1, \ldots, \kappa_n)$, where $\kappa_1 \geq \cdots \geq \kappa_n \geq 0$ (we may omit the last parts if they are equal to 0). Following [22], let them be denoted by V_n^κ. For example, $V_n^{()} = \mathbb{C}\,\mathbf{1}$, and $V_n^{(k)} = \mathrm{Harm}_k$.

The length $\ell(\kappa)$ of a partition κ is the number of its non zero parts, and its degree $\deg(\kappa)$ also denoted by $|\kappa|$ equals $\sum_{i=1}^n \kappa_i$.

Then, the decomposition of $\mathcal{C}(\mathcal{G}_{m,n})$ is as follows:

$$\mathcal{C}(\mathcal{G}_{m,n}) \simeq \oplus V_n^{2\kappa}$$

where κ runs over the partitions of length at most m and 2κ stands for partitions with even parts. We denote by $P_\kappa(y_1, \ldots, y_m)$ the zonal function associated to $V_n^{2\kappa}$. It turns out that the P_κ are symmetric polynomials in the m variables $y_1, \ldots, y_m$, of degree $|\kappa|$, with rational coefficients once they are normalized by the condition $P_\kappa(1, \ldots, 1) = 1$. Moreover, the set $(P_\kappa)_{|\kappa| \leq k}$ is a basis of the space of symmetric polynomials in the variables $y_1, \ldots, y_m$ of degree at most equal to k, which is orthogonal for the induced inner product calculated in [23],

$$d\mu = \lambda \prod_{\substack{i,j=1 \\ i<j}}^{m} |y_i - y_j| \prod_{i=1}^{m} y_i^{-1/2} (1 - y_i)^{n/2 - m - 1/2} dy_i$$

(One recognizes a special case of the orthogonal measure associated to *generalized Jacobi polynomials* ([25]).

5.4.2. **The ordered Hamming space:** it follows from the discussion in 4.2.3 that the variables of the zonal functions are the $(e_0, e_1, \dots, e_r)$. Elaborating on the computation explained above for the Johnson space, one can see that in the case of finite spaces, the weights of the induced measure are given by the number of elements of the orbits of X under the action of $\mathrm{Stab}(e)$ for any $e \in X$. Taking $e = 0^{rn}$, thus $\mathrm{Stab}(e) = B^n \rtimes S_n$, and the orbit of x is the set of elements with the same shape $(f_0, \dots, f_r)$ as x. The number of such elements is $\binom{n}{f_0 \dots f_r} 2^{\sum_i (i-1)e_i}$. These are the weights associated to the multivariate Krawtchouk polynomials.

5.4.3. **The space $X = \Gamma$ under the action of $G = \Gamma \times \Gamma$:** we need an explicit parametrization of the conjugacy classes of Γ, which is afforded by very few groups. Famous examples (if not the only ones) are provided by the permutation groups and the unitary groups. In the first case the parametrization is by the decomposition in disjoint cycles and in the second case it is by the eigenvalues. The decomposition of $\mathcal{C}(X)$ is given by Peter Weyl theorem

$$\mathcal{C}(\Gamma) = \sum_{R \in \mathcal{R}} R \otimes R^*$$

and the associated functions $P_R(x, y)$ are the characters:

$$P_R(x, y) = \chi_R(xy^{-1}).$$

In both cases (S_n and $U(\mathbb{C}^n)$) the irreducible representations are indexed by partitions λ and there are explicit expressions for P_λ. In the case of the unitary group $P_\lambda(xy^{-1})$ are the so-called Schur polynomials evaluated at the eigenvalues of xy^{-1}.

5.5. **Three cases with non trivial multiplicities.** So far the computation of the matrices $E_k(x, y)$ in cases of non trivial multiplicities has been worked out in very few cases. We shall discuss three very similar cases, namely the unit sphere of the Euclidean sphere ([4]), the Hamming space ([46]), and the projective geometry over $\mathbb{F}_q$ ([7]), where the group considered is the stabilizer of one point. In the case of the Hamming space, this computation amounts to the computation of the Terwilliger algebra of the association scheme and was performed initially by A. Schrijver in [40], who treated also the non binary Hamming space [20]. The framework of group representations was used in [46] to obtain the semidefinite matrices of [40] in terms of orthogonal polynomials. We present here the uniform treatment of the Hamming space and of the projective geometry in the spirit of [17] adopted in [7]. We also generalize to the case of the stabilizer of many points in the spherical case and enlighten the connection with the positive definite functions calculated in [34].

5.5.1. **The unit sphere S^{n-1}, with $G := \mathrm{Stab}(e, O(\mathbb{R}^n))$.** We continue the discussion initiated in 3.3.2 and we follow [4]. Let $E_k^n(x, y)$ be the zonal matrix associated to the isotypic subspace $\mathcal{I}_k$ related to Harm_k^{n-1} and to its decomposition described in 3.3.2:

$$\mathcal{I}_k = H_{k,k}^n \perp H_{k,k+1}^n \perp \dots$$

We index E_k^n with $i, j \geq 0$ so that $E_{k,i,j}^n(x, y)$ is related to the spaces $H_{k,k+i}^n$, $H_{k,k+j}^n$. The orbits of G on pairs of points $(x, y) \in X^2$ are characterized by the

values of the three inner products $u := e \cdot x$, $v := e \cdot y$ and $t := x \cdot y$. Thus (u, v, t) are the variables of the zonal matrices and we let:

$$E_k^n(x, y) = Y_k^n(u, v, t).$$

Theorem 5.1. *[*[4]*]*

$$Y_{k,i,j}^n(u, v, t) = \lambda_{k,i}\lambda_{k,j} P_i^{n+2k}(u) P_j^{n+2k}(v) Q_k^{n-1}(u, v, t), \tag{14}$$

where

$$Q_k^{n-1}(u, v, t) := \big((1-u^2)(1-v^2)\big)^{k/2} P_k^{n-1}\Big(\frac{t-uv}{\sqrt{(1-u^2)(1-v^2)}}\Big),$$

and $\lambda_{k,i}$ are some real constants.

Proof. We need an explicit construction of the spaces $H_{k,k+i}^{n-1}$. We refer to [1, Ch. 9.8]. For $x \in S^{n-1}$, let

$$x = ue + \sqrt{1-u^2}\zeta,$$

where $u = x \cdot e$ and ζ belongs to the unit sphere S^{n-2} of $(\mathbb{R}e)^\perp$. With $f \in H_k^{n-1} \subset \mathcal{C}(S^{n-2})$ we associate $\varphi(f) \in \mathcal{C}(S^{n-1})$ defined by:

$$\varphi(f)(x) = (1-u^2)^{k/2} f(\zeta).$$

Moreover, we recall that H_k^n is a subspace of the space $\mathrm{Pol}_{\leq k}(S^{n-1})$ of polynomial functions in the coordinates of degree at most k. Note that the multiplication by $(1-u^2)^{k/2}$ forces $\varphi(f)$ to be a polynomial function in the coordinates of x. Clearly φ commutes with the action of G. Hence $\varphi(H_k^{n-1})$ is a subspace of $\mathrm{Pol}_{\leq k}(S^{n-1})$ which is isomorphic to Harm_k^{n-1}. It is clear that these spaces are pairwise orthogonal. More generally, the set $\{\varphi(f)P(u) : f \in \mathrm{Harm}_k^{n-1}, \deg P \leq i\}$ is a subspace of $\mathrm{Pol}_{\leq k+i}(S^{n-1})$ which is isomorphic to $i+1$ copies of Harm_k^{n-1}. By induction on k and i there exist polynomials $P_i(u)$ of degree i such that $H_{k,k+i}^{n-1} := \varphi(H_k^{n-1})P_i(u)$ is a subspace of H_{k+i}^n. This construction proves the decomposition (8). Moreover, we can exploit the fact that the subspaces $H_{k,l}^{n-1}$ are pairwise orthogonal to prove an orthogonality relation between the polynomials P_i. Then this orthogonality relation will enable us to identify the polynomials P_i with Gegenbauer polynomials, up to the multiplication by a constant factor. Let us recall that the measures on S^{n-1} and on S^{n-2} are related by:

$$d\omega_n(x) = (1-u^2)^{(n-3)/2} du d\omega_{n-1}(\zeta).$$

Whenever $i \neq j$ we have for all $f \in H_k^{n-1}$

$$\begin{aligned} 0 &= \frac{1}{\omega_n}\int_{S^{n-1}} \varphi(f)P_i(u)\overline{\varphi(f)P_j(u)}d\omega_n(x) \\ &= \frac{1}{\omega_n}\int_{S^{n-1}} |f(\zeta)|^2(1-u^2)^k P_i(u)\overline{P_j(u)}d\omega_n(x) \\ &= \frac{1}{\omega_n}\int_{S^{n-2}} |f(\zeta)|^2 d\omega_{n-1}(\zeta)\int_{-1}^1 (1-u^2)^{k+(n-3)/2} P_i(u)\overline{P_j(u)}du, \end{aligned}$$

from which we derive that

$$\int_{-1}^1 (1-u^2)^{k+(n-3)/2} P_i(u)\overline{P_j(u)}du = 0;$$

hence the polynomials $P_i(u)$ are proportional to $P_i^{n+2k}(u)$ (thus with real coefficients..). We obtain an orthonormal basis of $H_{k,k+i}^{n-1}$ from an orthonormal basis $(f_1,\dots,f_h)$ of H_k^{n-1} by taking $e_{k,i,s}=\lambda_{k,i}\varphi(f_s)P_i^{n+2k}(u)$ for a suitable normalizing factor $\lambda_{k,i}>0$. With these basis we can compute $E_{k,i,j}^n$:

$$\begin{aligned}
E_{k,i,j}^n(x,y) &= \sum_{s=1}^{h_k^{n-1}} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}\\
&= \sum_{s=1}^{h_k^{n-1}} \lambda_{k,i}(1-u^2)^{k/2}f_s(\zeta)P_i^{n+2k}(u)\lambda_{k,j}(1-v^2)^{k/2}\overline{f_s(\xi)}P_j^{n+2k}(v)\\
&= \lambda_{k,i}\lambda_{k,j}P_i^{n+2k}(u)P_j^{n+2k}(v)\big((1-u^2)(1-v^2)\big)^{k/2}\sum_{s=1}^{h_k^{n-1}} f_s(\zeta)\overline{f_s(\xi)}\\
&= \lambda_{k,i}\lambda_{k,j}P_i^{n+2k}(u)P_j^{n+2k}(v)\big((1-u^2)(1-v^2)\big)^{k/2}h_k^{n-1}P_k^{n-1}(\zeta\cdot\xi),
\end{aligned}$$

where we have written $y=ve+\sqrt{1-v^2}\xi$ and where the last equality results from the analysis of zonal functions of S^{n-1}. Since

$$\zeta\cdot\xi=(t-uv)/\sqrt{(1-u^2)(1-v^2)},$$

we have completed the proof. □

5.5.2. **The unit sphere S^{n-1} with the action of $G:=\mathrm{Stab}(e_1,\dots,e_s,O(\mathbb{R}^n))$.** We assume that $(e_1,\dots,e_s)$ is a set of orthonormal vectors. The group $G:=\mathrm{Stab}(e_1,\dots,e_s,O(\mathbb{R}^n))$ is isomorphic to $O(\mathbb{R}^{n-s})$. The orbit of a pair $(x,y)\in X^2$ under G is characterized by the data: $t:=x\cdot y$, $u:=(x\cdot e_1,\dots,x\cdot e_s)$, $v:=(y\cdot e_1,\dots,y\cdot e_s)$. The decomposition (8) applied recursively shows that $\mathcal{C}(S^{n-1})$ decomposes as the sum of G-irreducible subspaces $H_{\underline{k}}$ where $\underline{k}=(k_0,\dots,k_s)$, $k_0\le k_1\le\dots\le k_s$, with the properties:

$$H_{\underline{k}}\subset H_{\underline{k}^{(r)}}\subset \mathrm{Pol}_{k_s},\quad H_{\underline{k}}\simeq \mathrm{Harm}_{k_0}^{n-s}$$

where $\underline{k}^{(r)}=(k_{s-r+1},\dots,k_s)$. Thus, for a given k_0, the multiplicity of the isotypic component $\mathcal{I}_{k_0}^d$ associated to $\mathrm{Harm}_{k_0}^{n-s}$ in $\mathrm{Pol}_{\le d}$ is the number of elements of

$$K_d:=\{(k_1,\dots,k_s)\ :\ k_0\le k_1\le\dots\le k_s\le d\}.$$

We construct the spaces $H_{\underline{k}}$ like in the proof of Theorem 5.1: for $x\in S^{n-1}$, let

$$x=u_1e_1+\dots+u_se_s+\sqrt{1-|u|^2}\zeta$$

where $u=(u_1,\dots,u_s)$ and $|u|^2=\sum_{i=1}^s u_i^2$. Let $\varphi: H_{k_0}^{n-s}\to\mathcal{C}(S^{n-1})$ be defined by $\varphi(f)(x)=(1-|u|^2)^{k_0/2}f(\zeta)$. Then $\varphi(H_{k_0}^{n-s})=H_{k_0^{s+1}}$ where $k_0^{s+1}=(k_0,k_0,\dots,k_0)$ and we set, for $\underline{l}=(l_1,\dots,l_s)$, $H_{k_0,\underline{l}}:=u_1^{l_1}\dots u_s^{l_s}H_{k_0^{s+1}}$. It is clear that $H_{k_0,\underline{l}}\simeq_G \mathrm{Harm}_{k_0}^{n-s}$ and that $H_{k_0,\underline{l}}\subset\mathrm{Pol}_d$ if $l_1+\dots+l_s\le d-k_0$ thus, since

$$K'_d:=\{l=(l_1,\dots,l_s)\ :\ l_i\ge 0,\ l_1+\dots+l_s\le d-k_0\}$$

has the same number of elements as K_d,

$$\mathcal{I}_{k_0}^d=\oplus_{\underline{l}\in K'_d}H_{k_0,l}.$$

This sum is not orthogonal but we can still use it to calculate E_{k_0}, the change will be to $AE_k(x,y)A^*$ for some invertible matrix A. The same calculation as in Theorem 5.1 shows that, (up to a change to some AY_kA^*):

$$Y_{k,\underline{i},\underline{j}}(u,v,t) = u^{\underline{i}-k}v^{\underline{j}-k}Q_k^{n-s}(u,v,t)$$

with the notations: $u^{\underline{i}-k} := u_1^{i_1-k}u_2^{i_2-k}\dots u_s^{i_s-k}$ and

$$Q_k^{n-s}(u,v,t) = \big((1-|u|^2)(1-|v|^2)\big)^{k/2}P_k^{n-s}\Big(\frac{t-(u\cdot v)}{\sqrt{(1-|u|^2)(1-|v|^2)}}\Big).$$

With Bochner Theorem 4.11 we recover the description of the multivariate positive definite functions on the sphere given in [34].

5.5.3. **The Hamming space and the projective geometry.** The set of all $\mathbb{F}_q$-linear subspaces of $\mathbb{F}_q^n$, also called the projective geometry, is denoted by $\mathcal{P}(n,q)$. The linear group $\mathrm{Gl}(n,\mathbb{F}_q)$ acts on $\mathcal{P}(n,q)$. The orbits of this action are the subsets of subspaces of fixed dimension, i.e. the q-Johnson spaces. If the Hamming space $\mathbb{F}_2^n$ is considered together with the action of the symmetric group S_n, the orbits of this action are the Johnson spaces. In [17] the Johnson space and the q-Johnson spaces are treated in a uniform way from the point of view of the linear programming method, the latter being viewed as q-analogs of the former. Thus the Johnson space corresponds to the value $q=1$. In particular the zonal polynomials are computed and they turn to be q-Hahn polynomials. Here we want to follow the same line for the determination of the zonal matrices $E(x,y)$ in both cases.

We take the following notations: if q is a power of a prime number, we let $X=\mathcal{P}(n,q)$ and $G=\mathrm{Gl}(n,\mathbb{F}_q)$, and, if $q=1$, we let X be the Hamming space, identified with the set of subsets of $\{1,\dots,n\}$, and $G=S_n$ the symmetric group with its standard action on X. Let

$$|x| := \begin{cases} wt(x) & \text{if } q=1 \\ \dim(x) & \text{if } q>1 \end{cases}$$

For all $w=0,\dots,n$, the space X_w is defined by

$$X_w = \{x\in X : |x| = w\}.$$

These subsets of X are exactly the orbits of G. The distance on X is given in every case by the formula

$$d(x,y) = |x|+|y|-2|x\cap y|. \tag{15}$$

The restriction of the distance d to X_w equals $d(x,y)=2(w-|x\cap y|)$ and it is a well known fact that G acts 2-points homogeneously on X_w. It is not difficult to see that the orbit of a pair (x,y) under the action of G is characterized by the triple $(|x|,|y|,|x\cap y|)$.

Following the notations of [17], the q-binomial coefficient $\left[{n \atop w}\right]$ expresses the cardinality of X_w. We have

$$\begin{bmatrix} n \\ w \end{bmatrix} = \begin{cases} \displaystyle\prod_{i=0}^{n-1}\frac{n-i}{w-i} = \binom{n}{w} & \text{if } q=1 \\ \displaystyle\prod_{i=0}^{n-1}\frac{q^{n-i}-1}{q^{w-i}-1} & \text{if } q>1 \end{cases}$$

In terms of the variable

$$[x] = q^{1-x}\begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{cases} x & \text{if } q = 1 \\ \dfrac{q^{-x}-1}{q^{-1}-1} & \text{if } q > 1 \end{cases},$$

we have

$$\begin{bmatrix} n \\ w \end{bmatrix} = q^{w(n-w)} \prod_{i=0}^{w-1} \frac{[n-i]}{[w-i]} = q^{w(n-w)} \frac{[n]!}{[w]![n-w]!}.$$

We have the obvious decomposition into pairwise orthogonal G-invariant subspaces:

$$\mathcal{C}(X) = \mathcal{C}(X_0) \perp \mathcal{C}(X_1) \perp \cdots \perp \mathcal{C}(X_n).$$

The decomposition of $\mathcal{C}(X_w)$ into G-irreducible subspaces is described in [17]. We have

$$\mathcal{C}(X_w) = H_{0,w} \perp H_{1,w} \perp \cdots \perp H_{\min(w,n-w),w}$$

where the $H_{k,w}$ are pairwise isomorphic for equal k and different w. and pairwise non isomorphic for different k. The picture looks like:

$$\begin{array}{lllllllll}
\mathcal{C}(X) = & \mathcal{C}(X_0) \perp & \mathcal{C}(X_1) \perp & \dots & \perp \mathcal{C}(X_{\lfloor \frac{n}{2} \rfloor}) \perp & \dots & \perp \mathcal{C}(X_{n-1}) & \perp \mathcal{C}(X_n) \\
 & H_{0,0} \perp & H_{0,1} \perp & \dots & \perp H_{0,\lfloor \frac{n}{2} \rfloor} \perp & \dots & \perp H_{0,n-1} & \perp H_{0,n} \\
 & & H_{1,1} \perp & \dots & & & \perp H_{1,n-1} & \\
 & & & \ddots & \vdots & & & \\
 & & & & H_{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor} & & &
\end{array}$$

where the columns represent the decomposition of $\mathcal{C}(X_w)$ and the rows the isotypic components of $\mathcal{C}(X)$, i.e. the subspaces $\mathcal{I}_k := H_{k,k} \perp H_{k,k+1} \perp \cdots \perp H_{k,n-k}$, $0 \le k \le \lfloor \frac{n}{2} \rfloor$, with multiplicity $m_k = (n-2k+1)$.

Let, for all (k,i) with $0 \le k \le i \le n-k$,

$$\begin{array}{llll}
\psi_{k,i}: & \mathcal{C}(X_k) & \to & \mathcal{C}(X_i) \\
 & f & \mapsto & \psi_{k,i}(f) : \psi_{k,i}(f)(y) = \sum_{\substack{|x|=k \\ x \subset y}} f(x)
\end{array}$$

and

$$\begin{array}{llll}
\delta_k: & \mathcal{C}(X_k) & \to & \mathcal{C}(X_{k-1}) \\
 & f & \mapsto & \delta_k(f) : \delta_k(f)(z) = \sum_{\substack{|x|=k \\ z \subset x}} f(x)
\end{array}$$

Obviously, these transformations commute with the action of G. The spaces $H_{k,i}$ are defined by: $H_{k,k} = \ker \delta_k$ and $H_{k,i} = \psi_{k,i}(H_{k,k})$. Moreover,

$$h_k := \dim(H_{k,k}) = \begin{bmatrix} n \\ k \end{bmatrix} - \begin{bmatrix} n \\ k-1 \end{bmatrix}.$$

We need later the following properties of $\psi_{k,i}$:

Lemma 5.2. *If $f, g \in H_{k,k}$,*

$$\langle \psi_{k,i}(f), \psi_{k,i}(g) \rangle = \begin{bmatrix} n-2k \\ i-k \end{bmatrix} q^{k(i-k)} \langle f, g \rangle. \tag{16}$$

Moreover,

$$\psi_{i,j} \circ \psi_{k,i} = \begin{bmatrix} j-k \\ i-k \end{bmatrix} \psi_{k,j} \tag{17}$$

Proof. [17, Theorem 3] proves (16). The relation (17) is straightforward: if $|z| = j$,

$$\begin{aligned}\psi_{i,j}(\psi_{k,i}(f))(z) &= \sum_{\substack{|y|=i\\ y\subset z}} \psi_{k,i}(f)(y) = \sum_{\substack{|y|=i\\ y\subset z}} \Big(\sum_{\substack{|x|=k\\ x\subset y}} f(x)\Big)\\ &= \sum_{\substack{|x|=k\\ x\subset z}} \Big(\sum_{\substack{|y|=i\\ x\subset y\subset z}} 1\Big) f(x) = \sum_{\substack{|x|=k\\ x\subset z}} \begin{bmatrix} j-k\\ i-k\end{bmatrix} f(x)\\ &= \begin{bmatrix} j-k\\ i-k\end{bmatrix} \psi_{k,j}(f)(z).\end{aligned}$$

□

Now we want to calculate the matrices E_k of size $m_k = (n-2k+1)$ associated to each isotypic space $\mathcal{I}_k$. We fix an orthonormal basis $(e_{k,k,1}, \dots, e_{k,k,h_k})$ of $H_{k,k}$ and we define $e_{k,i,s} := \psi_{k,i}(e_{k,k,s})$. It is clear from the definitions above that $e_{k,i,s}$ can be assumed to take real values. From (16), for fixed k and i, they form an orthogonal basis of $H_{k,i}$ with square norm equal to $\begin{bmatrix} n-2k\\ i-k\end{bmatrix} q^{k(i-k)}$. Normalizing them would conjugate E_k by a diagonal matrix, so we can omit to do it. The matrix E_k is indexed with i, j subject to $k \le i, j \le n-k$. From the construction, we have $E_{k,i,j}(x, y) = 0$ if $|x| \ne i$ or $|y| \ne j$; since the matrix E_k is zonal, we can define $P_{k,i,j}$ by

$$E_{k,i,j}(x, y) = P_{k,i,j}(i - |x \cap y|)$$

and our goal is to calculate the $P_{k,i,j}$. It turns out that these functions express in terms of the so-called q-Hahn polynomials.

We define the q-Hahn polynomials associated to the parameters n, i, j with $0 \le i \le j \le n$ to be the polynomials $Q_k(n, i, j; x)$ with $0 \le k \le \min(i, n-j)$ uniquely determined by the properties:

- Q_k has degree k in the variable $[x]$.
- $(Q_k)_k$ is a sequence of polynomials orthogonal for the weights
$$0 \le u \le i \quad w(n, i, j; u) = \begin{bmatrix} i\\ u\end{bmatrix}\begin{bmatrix} n-i\\ j-i+u\end{bmatrix} q^{u(j-i+u)}$$
- $Q_k(0) = 1$

The polynomials Q_k defined in [17] and 5.3.1 correspond up to multiplication by h_k to the parameters (n, w, w) and, with the notations of [19], according to Theorem 2.5, again up to a multiplicative factor, $Q_k(n, i, j; x) = E_m(i, n-i, j, i-x; q^{-1})$. The combinatorial meaning of the above weights is the following:

Lemma 5.3. [19, Proposition 3.1] *Given $x \in X_i$, the number of elements $y \in X_j$ such that $|x \cap y| = i - u$ is equal to $w(n, i, j; u)$.*

Theorem 5.4. *If $k \le i \le j \le n-k$, $|x| = i$, $|y| = j$,*

$$E_{k,i,j}(x, y) = |X| h_k \frac{\begin{bmatrix} j-k\\ i-k\end{bmatrix}\begin{bmatrix} n-2k\\ j-k\end{bmatrix}}{\begin{bmatrix} n\\ j\end{bmatrix}\begin{bmatrix} j\\ i\end{bmatrix}} q^{k(j-k)} Q_k(n, i, j; i - |x \cap y|)$$

If $|x| \ne i$ or $|y| \ne j$, $E_{k,i,j}(x, y) = 0$.

Proof. We proceed in two steps: the first step (18) calculates $P_{k,i,j}(0)$ and the second step (19) obtains the orthogonality relations.

Lemma 5.5. *With the above notations,*

$$P_{k,i,j}(0) = |X| h_k \frac{\begin{bmatrix} j-k \\ i-k \end{bmatrix} \begin{bmatrix} n-2k \\ j-k \end{bmatrix}}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} q^{k(j-k)}. \tag{18}$$

Proof. We have $P_{k,i,j}(0) = E_{k,i,j}(x,y)$ for all x, y with $|x| = i$, $|y| = j$, $x \subset y$. Hence

$$\begin{aligned} P_{k,i,j}(0) &= \frac{1}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{\substack{|x|=i, |y|=j \\ x \subset y}} E_{k,i,j}(x,y) \\ &= \frac{1}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{\substack{|x|=i, |y|=j \\ x \subset y}} \sum_{s=1}^{h_k} e_{k,i,s}(x) e_{k,j,s}(y) \\ &= \frac{1}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{s=1}^{h_k} \sum_{|y|=j} \Big(\sum_{\substack{|x|=i \\ x \subset y}} e_{k,i,s}(x) \Big) e_{k,j,s}(y) \\ &= \frac{1}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{s=1}^{h_k} \sum_{|y|=j} \psi_{i,j}(e_{k,i,s})(y) e_{k,j,s}(y) \end{aligned}$$

Since, from (17)

$$\psi_{i,j}(e_{k,i,s}) = \psi_{i,j} \circ \psi_{k,i}(e_{k,k,s}) = \begin{bmatrix} j-k \\ i-k \end{bmatrix} \psi_{k,j}(e_{k,k,s}) = \begin{bmatrix} j-k \\ i-k \end{bmatrix} e_{k,j,s},$$

we obtain

$$\begin{aligned} P_{k,i,j}(0) &= \frac{1}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{s=1}^{h_k} \sum_{|y|=j} \begin{bmatrix} j-k \\ i-k \end{bmatrix} e_{k,j,s}(y) e_{k,j,s}(y) \\ &= \frac{\begin{bmatrix} j-k \\ i-k \end{bmatrix}}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} \sum_{s=1}^{h_k} |X| \langle e_{k,j,s}, e_{k,j,s} \rangle = |X| h_k \frac{\begin{bmatrix} j-k \\ i-k \end{bmatrix} \begin{bmatrix} n-2k \\ j-k \end{bmatrix}}{\begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} j \\ i \end{bmatrix}} q^{k(j-k)} \end{aligned}$$

from (16). □

Lemma 5.6. *With the above notations,*

$$\sum_{u=0}^{i} w(n,i,j;u) P_{k,i,j}(u) P_{l,i,j}(u) = \delta_{k,l} |X|^2 h_k \frac{\begin{bmatrix} n-2k \\ i-k \end{bmatrix} \begin{bmatrix} n-2k \\ j-k \end{bmatrix} q^{k(i+j-2k)}}{\begin{bmatrix} n \\ i \end{bmatrix}}. \tag{19}$$

Proof. We compute $\Sigma := \sum_{y \in X} E_{k,i,j}(x,y) E_{l,i',j'}(y,z)$.

$$\begin{aligned}
\Sigma &= \sum_{y\in X}\sum_{s=1}^{h_k}\sum_{t=1}^{h_l} e_{k,i,s}(x)e_{k,j,s}(y)e_{l,i',t}(y)e_{l,j',t}(z)\\
&= \sum_{s=1}^{h_k}\sum_{t=1}^{h_l} e_{k,i,s}(x)e_{l,j',t}(z)\Big(\sum_{y\in X} e_{k,j,s}(y)e_{l,i',t}(y)\Big)\\
&= \sum_{s=1}^{h_k}\sum_{t=1}^{h_l} e_{k,i,s}(x)e_{l,j',t}(z)|X|\langle e_{k,j,s}, e_{l,i',t}\rangle\\
&= \sum_{s=1}^{h_k}\sum_{t=1}^{h_l} e_{k,i,s}(x)e_{l,j',t}(z)|X|\begin{bmatrix} n-2k\\ j-k\end{bmatrix} q^{k(j-k)}\delta_{k,l}\delta_{j,i'}\delta_{s,t}\\
&= \delta_{k,l}\delta_{j,i'}|X|\begin{bmatrix} n-2k\\ j-k\end{bmatrix} q^{k(j-k)}\sum_{s=1}^{h_k} e_{k,i,s}(x)e_{l,j',s}(z)\\
&= \delta_{k,l}\delta_{j,i'}|X|\begin{bmatrix} n-2k\\ j-k\end{bmatrix} q^{k(j-k)}E_{k,i,j'}(x,z).
\end{aligned}$$

We obtain, with $j = i'$, $j' = i$, $x = z \in X_i$, taking account of $E_{l,j,i}(y,x) = E_{l,i,j}(x,y)$,

$$\sum_{y\in X_j} E_{k,i,j}(x,y)E_{l,i,j}(x,y) = \delta_{k,l}|X|\begin{bmatrix} n-2k\\ j-k\end{bmatrix} q^{k(j-k)}E_{k,i,i}(x,x).$$

The above identity becomes in terms of $P_{k,i,j}$

$$\sum_{y\in X_j} P_{k,i,j}(i-|x\cap y|)P_{l,i,j}(i-|x\cap y|) = \delta_{k,l}|X|\begin{bmatrix} n-2k\\ j-k\end{bmatrix} q^{k(j-k)}P_{k,i,i}(0).$$

Taking account of (18) and Lemma 5.3, we obtain (19). □

To finish the proof of Proposition 5.4, it remains to prove that $P_{k,i,j}$ is a polynomial of degree at most k in the variable $[u] = [|x\cap y|]$. It follows from the reasons invoked in [17] in the case $i = j$ (see the proof of Theorem 5). □

Remark 5.7. *In the case $q = 1$, i.e. the Hamming space, we could have followed the same line as for the sphere in order to decompose $\mathcal{C}(H_n)$ under the action of G. We could have started from the decomposition of $\mathcal{C}(H_n)$* (3) *under the action of $\Gamma := T \rtimes S_n = \mathrm{Aut}(H_n)$ and then we could have decomposed each space P_k under the action of $G = \mathrm{Stab}(0^n, \Gamma)$. But we have a G-isomorphism from $\mathcal{C}(X_w) = \mathcal{C}(J_n^w)$ to P_w given by:*

$$\begin{aligned}
\mathcal{C}(J_n^w) &\to P_w\\
f &\mapsto \sum_{wt(y)=w} f(y)\chi_y
\end{aligned}$$

Note that the inverse isomorphism is the Fourier transform on $(\mathbb{Z}/2\mathbb{Z})^n$. So we pass from one to the other decomposition of $\mathcal{C}(H_n)$ through Fourier transform.

6. An SDP upper bound for codes from positive definite functions

In this section we want to explain how the computation of the continuous G-invariant positive definite functions on X can be used for applications to coding theory. In coding theory, it is of great importance to estimate the maximal number of elements of a finite subset C of a space X, where C is submitted to some constraints. Typically X is a metric space with G-invariant distance $d(x,y)$ and the constraints are related to the values taken by the distance on pairs of elements of C. In the following we concentrate on the basic case where the requirement is that the distance takes non zero values at least equal to some minimum δ. We denote by D the set of all values taken by $d(x,y)$ and we define $D_{\geq\delta} = D \cap [\delta, +\infty[$ and

$$A(X,\delta) := \max\{\mathrm{card}(C) \ : \ d(c,c') \geq \delta \text{ for all } c \neq c', (c,c') \in C^2\}.$$

We first focus on an upper bound for $A(X,\delta)$, which is obtained very obviously from the optimal value of the following program:

Definition 6.1.

$$\begin{aligned} m(X,\delta) \quad = \inf\Big\{ \quad t: \quad & F \in \mathcal{C}(X^2),\ \overline{F} = F,\ F \succeq 0 \\ & F(x,x) \leq t-1, \\ & F(x,y) \leq -1 \quad d(x,y) \geq \delta \Big\} \end{aligned} \tag{20}$$

Then we obtain an upper bound for $A(X,\delta)$:

Theorem 6.2.

$$A(X,\delta) \leq m(X,\delta).$$

Proof. For a feasible solution F, and for $C \subset X$ with $d(C) \geq \delta$ we have

$$0 \leq \sum_{(c,c')\in C^2} F(c,c') \leq (t-1)|C| - |C|(|C|-1)$$

thus $|C| \leq t$. □

Now the group G comes into play. From a feasible solution F one can construct a G-invariant feasible solution F' with the same objective value:

$$F'(x,y) = \int_G F(gx,gy)dg$$

thus we can add to the conditions defining the feasible solutions of $m(X,\delta)$ that F is G-invariant. Then we can apply Bochner characterization of the G-invariant positive definite functions (Theorem 4.11). Moreover we have also seen in Theorem 4.12 that if X is a homogeneous space, the finite sums of type (12) are arbitrary close for $\|\|_\infty$ to the G-invariant positive definite functions on X, so we can replace F by an expression of the form (12) in the SDP $m(X,\delta)$. Moreover, we replace $E_k(x,y)$ with its expression $Y_k(u(x,y))$ in terms of the orbits of pairs and we take account of the fact that $\overline{F} = F$. All together, with the notations of subsection 4.3 we obtain the (finite) semidefinite programs:

$$\begin{aligned} m^{(d)}(X,\delta) \quad = \inf\Big\{ \quad t: \quad & F_0 \succeq 0, \ldots, F_k \succeq 0, \ldots \\ & \textstyle\sum_{k\geq 0} \langle F_k, \tilde{Y}_k(u(x,x)) \leq t-1, \\ & \textstyle\sum_{k\geq 0} \langle F_k, \tilde{Y}_k(u(x,y)) \leq -1 \quad d(x,y) \geq \delta \Big\} \end{aligned} \tag{21}$$

where the matrices F_k are real symmetric, with size $m_{d,k}$, and $\tilde{Y}_k(u(x,y)) = Y_k(u(x,y)) + \overline{Y_k(u(x,y))}$. We insist that in the above program only a finite number of integers k are to be taken account of because $m_{d,k} \neq 0$ for a finite number of integers k. Thus we have $m(X,\delta) \leq m^{(d)}(X,\delta)$ and

$$\lim_{d \to +\infty} m^{(d)}(X,\delta) = m(X,\delta).$$

6.1. **The 2-point homogeneous spaces.** We recall that a sequence of orthogonal functions $(P_k)_{k\geq 0}$ is associated to X such that the G-invariant positive definite functions have the expressions

$$F(x,y) = \sum_{k\geq 0} f_k P_k(d(x,y)) \text{ with } f_k \geq 0.$$

Then

$$m(X,\delta) = \inf \{ \, 1 + \textstyle\sum_{k\geq 1} f_k \; : \quad f_k \geq 0, \\ 1 + \textstyle\sum_{k\geq 1} f_k P_k(i) \leq 0 \text{ for all } i \in D_{\geq \delta} \, \}$$

We restate Theorem 6.2 in the classical form of Delsarte linear programming bound:

Theorem 6.3. *Let $F(t) = f_0 + f_1 P_1(t) + \cdots + f_d P_d(t)$. If $f_k \geq 0$ for all $0 \leq k \leq d$ and $f_0 > 0$, and if $F(t) \leq 0$ for all $t \in D_{\geq \delta}$, then*

$$A(X,\delta) \leq \frac{f_0 + f_1 + \cdots + f_d}{f_0}.$$

Example: $X = S^7$, $d(x,y) = \theta(x,y)$, $d(C) = \pi/3$. This value of the minimal angle corresponds to the kissing number problem. A very good kissing configuration is well known: it is the root system E_8, also equal to the set of minimal vectors of the E_8 lattice. It has 240 elements and the inner products take the values ± 1, 0, $\pm 1/2$. We recall that the zonal polynomials associated to the unit sphere are proportional to the Gegenbauer polynomials P_k^n in the variable $x \cdot y$. If $P(t)$ obtains the tight bound 240 in Theorem 6.3, then we must have $P(t) \leq 0$ for $t \in [-1, 1/2]$ and $P(-1) = P(\pm 1/2) = P(0) = 0$ (as part of the *complementary slackness conditions*). The simplest possibility is $P = (t - 1/2)t^2(t + 1/2)^2(t + 1)$. One can check that

$$\frac{320}{3} P = P_0^8 + \frac{16}{7} P_1^8 + \frac{200}{63} P_2^8 + \frac{832}{231} P_3^8 + \frac{1216}{429} P_4^8 + \frac{5120}{3003} P_5^8 + \frac{2560}{4641} P_6^8$$

and that

$$\frac{P(1)}{f_0} = 240.$$

Thus the kissing number in dimension 8 is equal to 240. This famous proof is due independently to Levenshtein [27] and Odlysko and Sloane [35]. A proof of uniqueness derives from the analysis of this bound ([10]). For the kissing number problem, this miracle reproduces only for dimension 24 with the set of shortest vectors of the Leech lattice. For the other similar cases in 2-point homogeneous spaces we refer to [28].

It is not always possible to apply the above "guess of a good polynomial" method. In order to obtain a more systematic way to apply Theorem 6.3, one can of course restrict the degrees of the polynomials to some reasonable value, but needs also to overcome the problem that the conditions $F(t) \leq 0$ for $t \in [-1, 1/2]$

represent infinitely many linear inequalities. One possibility is to sample the interval and then a posteriori study the extrema of the approximated optimal solution found by an algorithm that solves the linear program with finitely many unknowns and inequalities. It is the method adopted in [35], where upper bounds for the kissing number in dimension $n \leq 30$ have been computed. We want to point out that polynomial optimization methods using SDP give another way to handle this problem. A polynomial $Q(t) \in \mathbb{R}[t]$ is said to be a sum of squares if $Q = \sum_{i=1}^{r} Q_i^2$ for some $Q_i \in \mathbb{R}[t]$. Being a sum of squares is a SDP condition since it amounts to ask that

$$Q = (1, t, \ldots, t^k) F (1, t, \ldots, t^k)^* \text{ with } F \succeq 0.$$

Here k is an upper bound for the degrees of the polynomials Q_i. Now we can relax the condition that $F(t) \leq 0$ for $t \in [-1, 1/2]$ to $F(t) = -Q(t) - Q'(t)(t+1)(t-1/2)$ with Q and Q' being sums of squares. A theorem of Putinar claims that in fact the two conditions are equivalent (but the degree of the polynomials under the squares are unknown).

A very nice achievement of the linear programming method in 2-point homogeneous spaces is the derivation of an asymptotic upper bound for the rate of codes (i.e. for the quotient $\log \operatorname{card}(C)/\dim(X)$) obtained from the so-called Christoffel-Darboux kernels. This method was first discovered for the Hamming and Johnson spaces [30] and then generalized to the unit sphere [24] and to all other 2-point homogeneous spaces [28]. It happens to be the best known upper bound for the asymptotic range. In [24] an asymptotic bound is derived for the density of sphere packings in Euclidean space which is also the best known.

6.2. **Symmetric spaces.** For these spaces, which are not 2-point homogeneous, there may be several distance functions of interest which are G-invariant. For example, the analysis of performance of codes in the Grassmann spaces for the MIMO channel [14] involves both the chordal distance:

$$d_c(p, q) := \sqrt{\sum_{i=1}^{m} \sin^2 \theta_i(p, q)}$$

and the product pseudo distance (it is not a distance in the metric sense):

$$d_p(p, q) := \prod_{i=1}^{m} \sin \theta_i(p, q).$$

The reformulation of Theorem 6.2 leads to a theorem of the type 6.3 for any symmetric function of the $y_i := \cos \theta_i$ with the Jacobi polynomials $P_\mu(y_1, \ldots, y_m)$ instead of the P_k. For a general symmetric space, a theorem of the type 6.3 is obtained, where the sequence of polynomials $P_k(t)$ is replaced by a sequence of multivariate polynomials, and the set D_δ is replaced by some compact subspace of the domain of the variables of the zonal functions, i.e. of the orbits of G acting on pairs. Then one can derive explicit upper bounds, see [45] for the permutation codes, [2] for the real Grassmann codes, [37] and [14] for the complex Grassmann codes, [15] for the unitary codes, [9] and [31] for the ordered codes. Moreover an asymptotic bound is derived in [2] and [9].

6.3. **Other spaces with true SDP bounds.** An example where the bound (20) does not boil down to an LP is provided by the spaces $\mathcal{P}(n,q)$ endowed with the distance (15) for which the matrices E_k are computed in section 5.5.3 (see [7]). In this case the group G is the largest group that acts on the SDP.

Indeed, it is useless to restrict the symmetrization of the program (20) to some subgroup of the largest group G that preserves (X,d). However, another interesting possibility is to change the restricted condition $d(x,y) \geq \delta$ in $A(X,\delta)$ for the conditions:

$$d(x,y) \geq \delta,\ d(x,e) \leq r,\ d(y,e) \leq r \tag{22}$$

where $e \in X$ is a fixed point. Then the new $A(X,e,r,\delta)$ is the maximal number of elements of a code with minimal distance δ in the ball $B(e,r) \subset X$. Here the group that leaves the program invariant is $\mathrm{Stab}(e,G)$. The corresponding bounds for codes in spherical caps where computed in [6] using the expressions of the zonal matrices of 5.5.1.

We end this section with some comments on these SDP bounds. We have indeed generalized the framework of the classical LP bounds but the degree of understanding of the newly defined bounds is far from the one of the classical LP bounds after the work done since [17], see e.g. [28]. It would be very interesting to have a better understanding of the best functions F that give the best bounds, to analyse explicit bounds and to analyse the asymptotic range, although partial results in these directions have already been obtained. The fact that one has to deal with multivariate polynomials introduces great difficulties when one tries to follow the same lines as for the classical one variable cases. A typical example is provided by the configuration of 183 points on the half sphere that seems numerically to be an optimal configuration for the one sided kissing number, and for which we failed to find the proper function F leading to a tight bound (see [7]).

7. Lovász theta

In this section we want to establish a link between the program (20) and the so-called Lovász theta number. This number was introduced by Lovász in the seminal paper [29] in order to compute the capacity of the pentagon. This remarkable result is the first of a long list of applications. This number is the optimal solution of a semidefinite program, thus is "easy to calculate", and offers an approximation of invariants of graphs that are "hard to calculate". Since then many other SDP relaxations of hard problems have been proposed in graph theory and in other domains.

7.1. **Introduction to Lovász theta number.** A graph $\Gamma = (V,E)$ is a finite set V of vertices together with a finite set E of edges, i.e. $E \subset V^2$. An independence set S is a subset of V such that $S^2 \cap E = \emptyset$. The independence number $\alpha(\Gamma)$ is the maximum of the number of elements of an independence set. It is a hard problem to determine the independence number of a graph. The connection with coding theory is as follows: a code C of a finite space X with minimal distance $d(C) \geq \delta$ is an independence set of the graph $\Gamma(X,\delta)$ which vertex set is equal to X and which edge set is equal to $E_\delta := \{(x,y) \in X^2 \ :\ d(x,y) \in]0,\delta[\}$. Thus the determination of $A(X,\delta)$ is the same as the determination of the independence number of this graph.

Among the many definitions of Lovász theta, we choose one which generalizes nicely to infinite graphs. For $S \subset V$, let $\mathbf{1}_S$ be the characteristic function of S. Let

$$M(x,y) := \frac{1}{|S|}\,\mathbf{1}_S(x)\,\mathbf{1}_S(y).$$

The following properties hold for M:

(1) $M \in \mathbb{R}^{n\times n}$, where $|V| = n$, and M is symmetric
(2) $M \succeq 0$
(3) $\sum_{x\in V} M(x,x) = 1$
(4) $M(x,y) = 0$ if $(x,y) \in E$
(5) $\sum_{(x,y)\in V^2} M(x,y) = |S|$.

Definition 7.1. *The theta number of the graph* $\Gamma = (V,E)$ *with* $V = \{1,2,\dots,n\}$ *is*

$$\begin{aligned} \vartheta(\Gamma) \quad = \max\Big\{ \textstyle\sum_{i,j} B_{i,j} : \quad & B \in \mathbb{R}^{n\times n},\ B \succeq 0 \\ & \textstyle\sum_i B_{i,i} = 1, \\ & B_{i,j} = 0 \quad (i,j) \in E \Big\} \end{aligned} \tag{23}$$

The dual program for ϑ has the same optimal value and is equal to:

$$\begin{aligned} \vartheta(\Gamma) \quad = \min\Big\{ \quad t : \quad & B \succeq 0 \\ & B_{i,i} = t-1, \\ & B_{i,j} = -1 \quad (i,j) \notin E \Big\} \end{aligned} \tag{24}$$

The complementary graph of Γ is denoted $\overline{\Gamma}$. The chromatic number $\chi(\Gamma)$ is the minimum number of colors needed to color the vertices so that no two connected vertices receive the same color. In other words it is a minimal partition of the vertex set with independence sets. Then the so-called Sandwich theorem holds:

Theorem 7.2.

$$\alpha(\Gamma) \le \vartheta(\Gamma) \le \chi(\overline{\Gamma})$$

Proof. The discussion prior to the theorem proves the first inequality. For the second inequality, let $c : V \to \{1,\dots,k\}$ be a coloring of $\overline{\Gamma}$. Then the matrix C with $C_{i,j} = -1$ if $c(i) \neq c(j)$, $C_{i,i} = k-1$ and $C_{i,j} = 0$ otherwise provides a feasible solution of (24). □

7.2. **Symmetrization and the q-gones.** Now we assume that G is (a subgroup of) the automorphism group $\mathrm{Aut}(\Gamma)$ of the graph. Then, G acts also on the above defined semidefinite programs. Averaging on G allows to construct a G-invariant optimal feasible solution B' from any optimal feasible solution B with the same objective value:

$$B'_{i,j} := \frac{1}{|G|}\sum_{g\in G} B_{g(i),g(j)}.$$

Thus one can restrict in the above programs to the G-invariant matrices. Then one can exploit the method developed in previous sections, in order to obtain a description of the G-invariant $B \succeq 0$ form the decomposition of the space $\mathcal{C}(V)$ under the action of G. We illustrate the method in the case of the q-gone C_q. There we have $V = G = \mathbb{Z}_q$ the group of integers modulo q. Let ζ_q be a fixed primitive root of 1 in $\mathbb{C}$. Let $\chi_k : \mathbb{Z}_q \to \mathbb{C}^*$ be defined by $\chi_k(x) = \zeta_q^{kx}$. The characters of $\mathbb{Z}_q$ are the χ_k for $0 \le k \le q-1$ and we have the decomposition

$$\mathcal{C}(\mathbb{Z}_q) = \oplus_{k=0}^{q-1}\mathbb{C}\chi_k.$$

According to Theorem 4.11, the G-invariant positive definite functions on V are exactly the functions $F(x,y)$ of the form:

$$F(x,y) = \sum_{k=0}^{q-1} f_k \chi_k(x)\overline{\chi_k(y)} = \sum_{k=0}^{q-1} f_k \zeta_q^{k(x-y)}$$

with $f_k \geq 0$. The ones taking real values have the form

$$F(x,y) = \sum_{k=0}^{\lfloor q/2 \rfloor} f_k \cos((x-y)2k\pi/q), \quad f_k \geq 0.$$

When one replaces in ϑ the expression $B_{i,j} = F(i,j)$, the SDP transforms into a LP on the variables f_k. More precisely, we compute $\sum_{(x,y)\in V^2} F(x,y) = q^2 f_0$ and $\sum_{x\in V} F(x,x) = q\sum_k f_k$. Thus we obtain (after a change of qf_k to f_k):

$$\begin{aligned} \vartheta(C_q) \quad = \max\big\{ qf_0 : \quad & f_k \geq 0,\ 0 \leq k \leq \lfloor q/2 \rfloor, \\ & \sum_{k=0}^{\lfloor q/2 \rfloor} f_k = 1, \\ & \sum_{k=0}^{\lfloor q/2 \rfloor} f_k \cos(2k\pi/q) = 0 \end{aligned}$$

The optimal value of this very simple linear program, is obtained for $f_1 = f_2 = \cdots = f_{\lfloor q/2 \rfloor - 1} = 0$, and equals

$$\vartheta(\mathbb{C}_q) = \begin{cases} \frac{q}{2} & \text{if } q \text{ is even} \\ \frac{q\cos(\pi/q)}{1+\cos(\pi/q)} & \text{if } q \text{ is odd .} \end{cases}$$

Note that when q is even, the independence number of the q-gone is exactly $q/2$. If the independence number of a graph as simple as the q-gone is not a great deal (it is of course equal to $\lfloor q/2 \rfloor$), a more challenging issue is to determine its capacity. In general, the capacity $C(\Gamma)$ of a graph Γ is defined to be

$$C(\Gamma) = \lim_{n\to+\infty} \alpha(\Gamma^n)^{1/n}.$$

Here the graph Γ^n is defined as follows: its vertex set is equal to V^n and an edge connects $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$ iff for all $1 \leq i \leq n$ either $x_i = y_i$ or $(x_i,y_i) \in E$. Introduced by Shannon in 1956, this number represents the effective size of an alphabet used to transmit information through the channel associated to the graph Γ (where two symbols are undistinguable if they are connected by an edge). If the capacity of a graph is in general very difficult to calculate, the theta number of a graph provides an upper bound for it because $\vartheta(\Gamma^n) = \vartheta(\Gamma)^n$ (see [29]). This upper bound is an equality for the pentagon since on one hand $\vartheta(C_5) = \sqrt{5}$ from our previous computation, and on the other hand it is easy to see that $\alpha((C_5)^2) = 5$ (while $\alpha(C_5) = 2$); this is the way taken by Lovász in [29] to prove that $C(C_5) = \sqrt{5}$. The determination of the capacity of the q-gone for q odd and greater than 5 is still opened.

7.3. **Relation with Delsarte bound and with $m(X,\delta)$.** We introduce a slightly stronger bound for $\alpha(\Gamma)$ with ϑ' and its dual form:

$$\begin{aligned} \vartheta'(\Gamma) \quad = \max\big\{ \textstyle\sum_{i,j} B_{i,j} : \quad & B \succeq 0,\ B \geq 0 \\ & \textstyle\sum_i B_{i,i} = 1, \\ & B_{i,j} = 0 \quad (i,j) \in E \big\} \end{aligned} \tag{25}$$

$$\begin{aligned} \vartheta'(\Gamma) \quad = \min\big\{ \quad t : \quad & B \succeq 0 \\ & B_{i,i} \ \leq \ t-1, \\ & B_{i,j} \ \leq \ -1 \quad (i,j) \notin E \big\} \end{aligned} \tag{26}$$

Since $M(x,y) \geq 0$, we still have that $\alpha(\Gamma) \leq \vartheta'(\Gamma)$. Again one can restrict in the above programs to the G-invariant matrices. It was recognized independently by McEliece, Rodemich, Rumsey, and Schrijver [39] that Delsarte bound of Theorem 6.3 for $A(H_n,\delta)$ is equal to ϑ' for the graph $\Gamma(X,\delta)$, once the feasible set is restricted to the $\mathrm{Aut}(H_n)$-invariant matrices, and similarly for the other finite 2-point homogeneous spaces. Indeed, by virtue of Theorem 4.11, the matrices B turn to be of the form $B(x,y) = \sum_{k\geq 0} f_k P_k(d(x,y))$. This symmetrization process is of great importance, not only because it has the great advantage to change an SDP to an LP, but also because it does change the complexity of the problem. Indeed, there are algorithms with polynomial complexity that do compute approximations of the optimal value of SDP's, thus algorithms with polynomial complexity *in the number of vertices* of Γ for ϑ. But the graphs arising from coding theory have in general an exponential number of vertices, e.g. 2^n for the Hamming graph. It is important to insist that the symmetrized theta has polynomial complexity in n.

Now we can see that the program $m(X,\delta)$ (20) is a natural generalization of ϑ' for metric spaces under the assumptions of Section 4. We refer to [8] for a more general discussion about generalized theta where also chromatic numbers are involved.

8. Strengthening the LP bound for binary codes

In this section we explain how the zonal matrices $E_k(x,y)$ related to the binary Hamming space computed in 5.5.3 are exploited in [40] in order to strengthen the LP bound. We shall work with the primal programs so we start to recall the primal version of (20) in the case of the Hamming space.

We recall that the sequence of orthogonal functions $(P_k)_{0\leq k\leq n}$ with $P_k = K_k$ the Krawtchouk polynomials is associated to H_n such that $P_k(d(x,y)) \succeq 0$. As a consequence, we have for all $k \geq 0$

$$\sum_{(c,c')\in C^2} P_k(d(c,c')) \geq 0.$$

We introduce the variables x_i, for $i \in [0\ldots n]$

$$x_i := \frac{1}{\mathrm{card}(C)} \mathrm{card}\{(c,c') \in C^2 \ : \ d(c,c') = i\}. \tag{27}$$

They satisfy the properties:

(1) $x_0 = 1$
(2) $x_i \geq 0$
(3) $\sum_i x_i P_k(i) \geq 0$ for all $k \geq 0$
(4) $x_i = 0$ if $i \in [1\ldots\delta-1]$

(5) $\mathrm{card}(C) = \sum_i x_i$.

With these properties which are linear inequalities, we obtain the following linear program which is indeed the dual of (20):

$$\sup \left\{ 1 + \sum_{i=\delta}^{n} x_i \ : \quad \begin{array}{l} x_i \geq 0, \\ 1 + \sum_{i=\delta}^{n} x_i P_k(i) \geq 0 \text{ for all } 1 \leq k \leq n \end{array} \right\}$$

where we have taken into account $P_0 = 1$.

We recall that to every $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, we have associated a matrix $E_k(x,y) \succeq 0$ of size $n - 2k + 1$. In particular, for all $C \subset H_n$ (see the remark 4.10),

$$\sum_{(c,c') \in C^2} E_k(c,c') \succeq 0.$$

These constraints are not interesting for pairs because they are not stronger than the linear inequalities coming from the Krawtchouk polynomials. They are only interesting if triples of points are involved: namely we associate to $(x,y,z) \in H_n^3$ the matrices

$$F_k(x,y,z) := E_k(x-z, y-z).$$

We have for all $C \subset H_n$, and for all $z \in H_n$,

$$\sum_{(c,c') \in C^2} F_k(c,c',z) \succeq 0$$

which leads to the two positive semidefinite conditions:

$$\left\{ \begin{array}{l} \sum_{(c,c',c'') \in C^3} F_k(c,c',c'') \succeq 0 \\ \sum_{(c,c') \in C^2,\ c'' \notin C} F_k(c,c',c'') \succeq 0 \end{array} \right. \tag{28}$$

Theorem 5.4, expresses the coefficients of $E_k(x-z, y-z)$ in terms of $wt(x-z)$, $wt(y-z)$, $wt(x-y)$; so with $a := d(y,z)$, $b := d(x,z)$, $c := d(x,y)$, we have for some matrices $T_k(a,b,c)$,

$$F_k(x,y,z) = T_k(a,b,c).$$

We introduce the unknowns $x_{a,b,c}$ of the SDP. Let

$$\Omega := \left\{ (a,b,c) \in [0 \ldots n]^3 : \quad \begin{array}{l} a+b+c \equiv 0 \mod 2 \\ a+b+c \leq 2n \\ c \leq a+b \\ b \leq a+c \\ a \leq b+c \end{array} \right\}$$

It is easy to check that $\Omega = \{(d(y,z), d(x,z), d(x,y)) \ : \ (x,y,z) \in H_n^3\}$. Let, for $(a,b,c) \in \Omega$,

$$x_{a,b,c} := \frac{1}{\mathrm{card}(C)} \mathrm{card}\{(x,y,z) \in C^3 : d(y,z) = a, d(x,z) = b, d(x,y) = c\}.$$

Note that

$$x_{0,c,c} = \frac{1}{\mathrm{card}(C)} \mathrm{card}\{(x,y) \in C^2 : d(x,y) = c\}$$

thus the old variables x_i (27) of the linear program are part of these new variables. We need a last notation: let

$$\begin{aligned} t(a,b,c) \ &:= \mathrm{card}\{z \in H_n : d(x,z) = b \text{ and } d(y,z) = a\} \text{ for } d(x,y) = c \\ &= \tbinom{c}{i}\tbinom{n-c}{a-i} \text{ where } a - b + c = 2i \end{aligned}$$

Then, if C is a binary code with minimal distance at least equal to δ, the following inequalities hold for $x_{a,b,c}$:

(1) $x_{0,0,0} = 1$
(2) $x_{a,b,c} \geq 0$
(3) $x_{a,b,c} = x_{\tau(a),\tau(b),\tau(c)}$ for all permutation τ of $\{a, b, c\}$
(4) $x_{a,b,c} \leq t(a, b, c)x_{0,c,c}$, $x_{a,b,c} \leq t(b, c, a)x_{0,a,a}$, $x_{a,b,c} \leq t(c, a, b)x_{0,b,b}$.
(5) $\sum_{a,b,c} T_k(a, b, c)x_{a,b,c} \succeq 0$ for all $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$
(6) $\sum_{a,b,c} T_k(a, b, c)(t(a, b, c)x_{0,c,c} - x_{a,b,c}) \succeq 0$ for all $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$
(7) $x_{a,b,c} = 0$ if a, b or $c \in]0, \delta[$.
(8) $\mathrm{card}(C) = \sum_c x_{0,c,c}$.

Conditions (5) and (6) are equivalent to (28). Condition (7) translates the assumption that $d(C) \geq \delta$. Thus an upper bound on $\mathrm{card}(C)$ is obtained with the optimal value of the program that maximizes $\sum_c x_{0,c,c}$ under the constraints (1) to (7). This upper bound is at least as good as the LP bound because the SDP program does contain the LP program of 6.1. Indeed, the sum of the two SDP conditions (28) is equivalent to

$$\sum_{z \in H_n} E_k(x - z, y - z) \succeq 0.$$

We claim that this set of conditions when $k = 0, 1, \ldots, \lfloor \frac{n}{2} \rfloor$ is equivalent to the set of conditions $P_k(d(x, y)) \succeq 0$ for $k = 0, \ldots, n$. Indeed let $B_k(x, y) := \sum_{z \in H_n} E_k(x - z, y - z)$. Up to a change of $B_k(x, y)$ to $AB_k(x, y)A^*$, we assume that E_k was constructed using the decomposition of $\mathcal{C}(H_n)$ first under $\Gamma := T \rtimes S_n = \mathrm{Aut}(H_n)$ then under G (see Remark 5.7). Clearly B_k is Γ-invariant. Since $x \to E_{k,i,j}(x, y) \in P_i$ and P_i is a Γ-module, also $x \to B_{k,i,j}(x, y) \in P_i$ and similarly $y \to \overline{B_{k,i,j}(x, y)} \in P_j$. But P_i and P_j are non isomorphic Γ-modules for $i \neq j$ thus $B_{k,i,j}(x, y) = 0$ for $i \neq j$. Since P_i is Γ-irreducible, $B_{k,i,i}(x, y) = \lambda_i P_i(d(x, y))$ for some $\lambda_i > 0$ that can be computed with $B_k(x, x)$. So we have proved that the linear program associated to H_n like in 6.1 is contained in the SDP program obtained from the above conditions (1) to (7). Moreover it turns out that in some explicit cases of small dimension the SDP bound is strictly better than the LP bound (see [40]).

A similar strengthening of the LP bound for the Johnson space and for the spaces of non binary codes where obtained in [40] and [20]. In the case of the spherical codes, for the same reasons as for the LP bound, one has to deal with the dual program, see [4].

REFERENCES

[1] G.E. Andrews, R. Askey, R. Roy, *Special functions*, Cambridge University Press, 1999.
[2] C. Bachoc, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52-5** (2006), 2111-2125.
[3] C. Bachoc, Y. Ben-Haim, S. Litsyn, *Bounds for codes in products of spaces, Grassmann and Stiefel manifolds*, IEEE Trans. Inf. Th., 54-3 (2008), 1024-1035.
[4] C. Bachoc, F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. 21 (2008), 909-924.
[5] C. Bachoc, F. Vallentin, *Optimality and uniqueness of the (4,10,1/6) spherical code*, Journal of Combinatorial Theory, Series A 116 (2009), 195-204.
[6] C. Bachoc, F. Vallentin, *Semidefinite programming, multivariate orthogonal polynomials, and codes in spherical caps*, special issue in the honor of Eichii Bannai, European Journal of Combinatorics **30** (2009), 625-637.

[7] C. Bachoc, F. Vallentin, *More semidefinite programming bounds*, Proceeding of DMHF 2007 (Fukuoka, 2007).
[8] C. Bachoc, G. Nebe, F.M. De Oliveira Filho, F. Vallentin, *Lower bounds for measurable chromatic numbers*, arXiv:math.MG/0801.1059, to appear in Geometric and Functional Analysis.
[9] A. Barg, P. Purkayastha, *Bounds on ordered codes and orthogonal arrays*, Moscow Math. Journal, no. 2 (2009)
[10] E. Bannai, N.J.A. Sloane, *Uniqueness of certain spherical codes*, Canad J. Math. **33** (1981), 437–449.
[11] S. Bochner, *Hilbert distances and positive definite functions*, Annals of Mathematics, **42**-3 (1941), 647-656.
[12] D. Bump, *Lie Groups*, Springer, GTM 225 (2004).
[13] J.H. Conway, N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 1988.
[14] J. Creignou *Mathématiques pour les télécommunications multi-antennes*, Thèse, Université Bordeaux 1, 2008.
[15] J. Creignou, H. Diet, *Linear programming bounds for unitary space time codes*, ISIT 2009, math.arXiv:0803.1227
[16] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.
[17] P. Delsarte, *Hahn polynomials, discrete harmonics and t-designs*, SIAM J. Appl. Math. **34**-1 (1978)
[18] P. Delsarte, J.M. Goethals, J.J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.
[19] R.J. Duffin, *Infinite Programs*, in: Linear inequalities and related systems, (H.W. Kuhn, A.W. Tucker eds.), Princeton Univ. Press, 1956, 157–170.
[20] D.C. Gijswijt, A. Schrijver, H. Tanaka, *New upper bounds for nonbinary codes*, J. Combin. Theory Ser. A **13** (2006), 1717–1731.
[21] G. H. Golub and C. F. van Loan, *Matrix computations*, 2nd edition., 1989, John Hopkins university press.
[22] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*, Encyclopedia of Mathematics and its Applications 68, 1998, Cambridge University Press.
[23] A. T. James and A. G. Constantine, "Generalized Jacobi polynomials as spherical functions of the Grassmann manifold", *Proc. London Math. Soc.* vol. 29 no. 3, 1974, 174-192.
[24] G.A. Kabatiansky, V.I. Levenshtein, *Bounds for packings on a sphere and in space*, Problems of Information Transmission **14** (1978), 1–17.
[25] M. Lassalle, "Polynômes de Jacobi généralisés", *C. R. Acad. Sci. Paris Sr. I Math.* vol. 312 no. 6, 1991, 425-428.
[26] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2001), 796–817.
[27] V.I. Levenshtein, *On bounds for packing in n-dimensional Euclidean space*, Soviet Math. Dokl. **20** (1979), 417–421.
[28] V. I. Levenshtein, "Universal bounds for codes and designs", in *Handbook of Coding Theory*, eds V. Pless and W. C. Huffmann, Amsterdam: Elsevier, 1998, 499-648.
[29] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **IT-25** (1979), 1-5
[30] R. J. McEliece, E. R. Rodemich, H. Rumsey, L. Welch, *New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities*, IEEE Trans. Inform. Theory **IT-23** (1977), 157-166.
[31] W.J. Martin, D.R. Stinson, *Association schemes for ordered orthogonal arrays and (T, M, S)-nets*, Canad. J. Math. **51**-2 (1999), 326–346.
[32] H. D. Mittelmann, F. Vallentin, *High accuracy semidefinite programming bounds for kissing numbers*, arXiv.math:0902.1105
[33] O.R. Musin, *The kissing number in four dimensions*, to appear in Annals of Mathematics.
[34] O.R. Musin, *Multivariate positive definite functions on spheres*, arxiv.math:0701083
[35] A.M. Odlyzko, N.J.A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26** (1979), 210–214.
[36] A. Roy, A. J. Scott *Unitary designs and codes*, preprint, arXiv:0809.3813
[37] A. Roy *Bounds for codes and designs in complex subspaces*, preprint, arXiv:0806.2317
[38] B. E. Sagan, *The symmetric group. representations, combinatorial algorithms and symmetric functions*, Springer, GTM 203, 2001.

[39] A. Schrijver, *A comparaison of the Delsarte and Lovász bound*, IEEE Trans. Inform. Theory **IT-25** (1979), 425-429

[40] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), 2859–2866.

[41] J.-P. Serre, *Représentations linéaires des groupes finis*

[42] K. Schütte, B.L. van der Waerden, *Das Problem der dreizehn Kugeln*, Math. Ann. **125** (1953) 325–334.

[43] G. Szegö, *Orthogonal polynomials*, American Mathematical Society, 1939.

[44] H. Tarnanen, M. Aaltonen, J.-M. Goethals, *On the nonbinary Johnson scheme*. European J. Combin. **6**-3 (1985),279–285.

[45] H. Tarnanen. *Upper bounds on permutation codes via linear programming*, Europ J. Combinatorics **20** (1999) 101–114.

[46] F. Vallentin, *Lecture notes: Semidefinite programs and harmonic analysis*, arXiv.math:0809.2017

[47] F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and Appl. 430 (2009), 360-369.

[48] N.Ja. Vilenkin, A.U. Klimyk, *Representation of Lie Groups and Special Functions, Volume 2*, Kluwer Academic Publishers, 1993.

C. BACHOC, INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UNIVERSITÉ BORDEAUX I, 351, COURS DE LA LIBÉRATION, 33405 TALENCE FRANCE

E-mail address: bachoc@math.u-bordeaux1.fr