LECTURE NOTES: SEMIDEFINITE PROGRAMS AND HARMONIC ANALYSIS

FRANK VALLENTIN

ABSTRACT. Lecture notes for the tutorial at the workshop *HPOPT 2008 — 10th International Workshop on High Performance Optimization Techniques (Algebraic Structure in Semidefinite Programming)*, June 11th to 13th, 2008, Tilburg University, The Netherlands.

CONTENTS

1. Introduction

Semidefinite programming is a vast extension of linear programming and has a wide range of applications: combinatorial optimization and control theory are the most famous ones. Although semidefinite programming has an enormous expressive power in formulating convex optimization problems it has a few practical drawbacks: Robust and efficient solvers, unlike their counterparts for solving linear programs, are currently not available. So it is crucial to exploit the problems' structure to be able to perform computations.

In the last years many results in the area of semidefinite programming were obtained for problems which have symmetry. The underlying principle which was used here is the following: One simplifies the original semidefinite program which is invariant under a group action by applying an algebra isomorphism that maps a "large" matrix algebra to a "small" matrix algebra. Then it is sufficient to solve the semidefinite program using the smaller matrices.

The aim of the tutorial is to give a general and explicit procedure to simplify semidefinite programs which are invariant under the action of a group.

A *(complex) semidefinite program* is an optimization problem of the form

$$(1) \qquad \max\{\langle C, K\rangle : \langle A_i, K\rangle = b_i,\ i = 1, \ldots, n,\ \text{and}\ K \succeq 0\},$$

where $A_i \in \mathbb{C}^{V \times V}$, and $C \in \mathbb{C}^{V \times V}$ are Hermitian matrices whose rows and columns are indexed by a finite set V, $(b_1, \ldots, b_n)^t \in \mathbb{R}^n$ is a given vector and $K \in \mathbb{C}^{V \times V}$ is a variable Hermitian matrix and where "$K \succeq 0$" means that K is positive semidefinite. Here $\langle C, K\rangle = \text{trace}(CK)$ denotes the trace inner product between matrices. In the following we denote the matrix entry (x, y) of K by $K(x, y)$ instead of the more familiar notion K_{xy}. This notation will make our treatment of infinite matrices more natural.

The punch line of the lecture is the following: Suppose that the semidefinite program (1) is invariant under a group Γ of permutations of V: If K is feasible for (1) then also K_u, defined by $K_u(x, y) = K(ux, uy)$, is feasible for all $u \in \Gamma$, and the objective values of K and K_u coincide. It is crucial to observe that to solve (1) it suffices to consider only those K which satisfy $K(x, y) = K(ux, uy)$ for all $u \in \Gamma$ because one can *symmetrize* every optimal solution K to obtain a Γ-invariant one:

$$\frac{1}{|\Gamma|} \sum_{u \in \Gamma} K(ux, uy).$$

Now one can apply Bochner's characterization of Γ-invariant positive semidefinite matrices. This is a classical result in harmonic analysis. It says that one can represent any Γ-invariant K by a block-diagonal matrix $\text{diag}(F_1, \ldots, F_l)$. One has the representation

$$K(x, y) = \sum_{k=1}^{l} \langle F_k, Z_k(x, y)\rangle,$$

where F_k are positive semidefinite matrices of size $m_k \times m_k$ and $Z_k(x, y) \in \mathbb{C}^{m_k \times m_k}$ are fixed basis matrices (At this point the notation $Z_k(x, y)$ might be confusing, it specifies a matrix and not a matrix entry, so it should be better something like $Z_k^{(x,y)}$ but this looks awkward.). Here the parameters l and $m_1, \ldots, m_l$ as well as

the basis matrices Z_k depend on the group Γ. The original semidefinite program (1) simplifies to

$$\max\left\{\left\langle C, \sum_{k=1}^{l}\langle F_k, Z_k\rangle\right\rangle \;:\; \left\langle A_i, \sum_{k=1}^{l}\langle F_k, Z_k\rangle\right\rangle = b_i, \; i = 1, \ldots, n,\right.$$
$$\left. F_1, \ldots, F_l \succeq 0\right\},$$

and in some cases this simplification results into a huge saving. The advantage is that instead of dealing with matrices of size $|V| \times |V|$ one has to deal with block diagonal matrices with l block matrices of size $m_1, \ldots, m_l$, respectively. In many applications the sum $m_1 + \cdots + m_l$ is much smaller than $|V|$ and in particular many practical solvers take advantage of the block structure to speed up the numerical calculations.

In this lecture we develop the theory of "block-diagonalization" of semidefinite programs starting from basic principles. We illustrate this symmetry reduction for one specific semidefinite program: The theta function for distance graphs on compact metric spaces. This is a generalization of the Lovász theta function for finite graphs. To focus on this example has the following benefits: It shows the geometric core of the process of symmetry reduction. In this context it is natural to connect semidefinite programs with harmonic analysis and the theory of group representation. In particular this provides the possibility to consider infinite-dimensional semidefinite programs which in many cases is a very convenient framework having potential for future research.

Then, this lecture gives the background for the recent development of semidefinite programming bounds for combinatorial and geometric packing problems initiated by Schrijver [82], and further developed by Laurent [60], Gijswijt, Schrijver, Tanaka [39], Gijswijt [38], Bachoc, Vallentin [5], [6], [7], [8], Musin [71]. It also gives the background the developments dealing with the more complicated case of noncompact metric spaces, see Cohn, Elkies [25], de Oliveira Filho, Vallentin [78]

We want to stress (once more):

> The techniques we present here apply to general semidefinite programs which are invariant under a symmetry group.

Then, one important remark: *For comments and suggestions concerning these notes the author would be very grateful.*

2. The theta function for distance graphs

In this section we study the theta function of distance graphs in compact metric spaces. These graphs can have infinitely many vertices. The theta function is a generalization of the Lovász theta function for finite graphs, which originally was introduced by Lovász in the celebrated paper [64]. The Lovász theta function gives an upper bound for the stability number which one can efficiently compute using semidefinite programming. The generalization was studied by Bachoc, Nebe, de Oliveira Filho, Vallentin [9]. The generalization also gives an upper bound for the stability number which one can compute using semidefinite programming. The main difference to the Lovász theta function is that one has to solve an infinite-dimensional

semidefinite program if the graph has infinitely many vertices. However, we show that if the distance graph is symmetric, solving this infinite-dimensional semidefinite program is feasible.

Outline of this section:

In Section 2.1 we provide the necessary definitions from graph theory (stability number and distance graphs). We show that finite graphs are distance graphs and we present the two main examples: distance graphs on the Hamming cube (related to error-correcting codes) and distance graphs on the unit sphere (related to discrete geometry).

In Section 2.2 we discuss two possible formulations of the Lovász theta function.

Before we can formulate the generalization of the theta function in Section 2.4 we review an infinite-dimensional generalization of positive semidefinite matrices, so-called positive Hilbert-Schmidt kernels, in Section 2.3. There we in particular focus on the spectral decomposition theorem which turns out to be central in the following.

In Section 2.5 we show how to exploit the symmetry of distance graphs in order to simplify the computation of the theta function. For this we introduce the automorphism group of a distance graph and explain the essential tool of group invariant integration.

2.1. **Distance graphs.** We start with some basic definitions from graph theory. Let $G = (V, E)$ be an *undirected graph* given by a finite set V of *vertices* and a subset $E \subseteq \binom{V}{2}$ of two-element subsets of V called *edges*. Two vertices x, y with $\{x, y\} \in E$ are called *adjacent* and a family of vertices $(x_1, \ldots, x_n)$ in which every two consequent elements are connected is called a *path of length $n - 1$*.

A *stable set* (sometimes also called *independent set*) of a graph G is a finite subset of the vertex-set in which no two vertices are adjacent. The *stability number* (or *independence number*) of a graph is the maximum cardinality of a stable set of G:

$$\alpha(G) = \max\{|C| : C \subseteq V, \ \{x, y\} \notin E \text{ for all } x, y \in C\}.$$

If the graph has infinitely many vertices it may happen that there is no maximum: The graph with vertex-set $V = S^{n-1} = \{x \in \mathbb{R}^n : x \cdot x = 1\}$, in which two points are adjacent whenever they are orthogonal, has stable sets of arbitrary cardinality. In this case, it makes sense to replace the maximal cardinality of a finite stable set by the maximal measure of a measurable stable set. Then in case of the circle S^1 the stability number equals π. This approach has been worked out by Bachoc, Nebe, de Oliveira Filho, Vallentin [9], but here we will stick to the case when $\alpha(G)$ is finite.

For the definition of a distance graph we use the triple (V, d, μ). Here V is a metric space with the distance function $d : V \times V \to \mathbb{R}$ which is equipped with the finite, regular Borel measure μ. We assume that V is separable and compact. Usually it does not cause confusion if one uses only V to specify the triple (V, d, μ) and we refer to V simply as a *compact metric space*. The adjacency relation of a distance graph only depends on the distance map d:

Definition 2.1. *Let V be a compact metric space and let I be a subset of the distances which may occur among distinct points in V. The* distance graph $G(V, I)$ *is the graph with vertex-set V and in which two vertices are adjacent if their distance lies in I. So the edge-set is $E = \big\{\{x, y\} : d(x, y) \in I\big\}$.*

The first example shows that finite graphs are distance graphs.

Example 2.2. *Let $G = (V, E)$ be a finite graph. By $d : V \times V \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ we denote the length of a shortest path connecting two vertices x and y in G where we set $d(x, y) = \infty$ whenever there is no connection at all. If the graph is connected, that is if d does not attain the value ∞, then d defines a metric on V. The set V comes with the uniform measure μ; we have $\mu(A) = |A|$ for any $A \subseteq V$. The graph G is a distance graph. Two vertices are adjacent whenever their distance is equal to one. Hence, the adjacency relation only depends on the distance map.*

The second example comes from engineering, error correcting codes. In fact, it can be viewed as a special case of the first example. Error correcting codes have great practical importance for communication across noisy channels and for storing and retrieving information on media. The idea is that the sender adds redundant data to its messages which allows the receiver to detect and correct errors under the assumption that there is not unreasonable amount of noise. In particular there is no need to ask the sender for additional data. The book MacWilliams, Sloane [66] is the definitive reference in algebraic coding theory.

Example 2.3. *Let $\{0, 1\}^n$ be the set of binary strings of length n. It is also called the n-dimensional* Hamming cube*. The Hamming cube is a metric space. The distance between two binary strings is measured by the* Hamming distance δ*, which is the number of entries in which they differ:*

$$\delta((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = |\{i \in \{1, \ldots, n\} : x_i \neq y_i\}|.$$

The Hamming norm *(or* Hamming weight*) of a vector $x \in \{0, 1\}^n$ is $\|x\| = \delta(x, 0)$.*

A central parameter is $A(n, d)$, the maximal cardinality of a subset $C \subseteq \{0, 1\}^n$ such that every two points in it have Hamming distance at least d. In other words, $A(n, d)$ is the stability number of the distance graph $G(\{0, 1\}^n, \{1, \ldots, d - 1\})$. A stable set in this graph can correct $\lfloor \frac{d-1}{2} \rfloor$ errors. This value has been determined for various parameters of n and d, but in general it is unknown to large extend. Brouwer [21] maintains a list of lower and upper bounds for $A(n, d)$ for $n \leq 28$.

The third example comes from discrete geometry.

Example 2.4. *We consider the infinite graph $G(S^{n-1}, (0, \theta))$ whose vertex-set consists of all the points on the $(n - 1)$-dimensional unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : x \cdot x = 1\}$. The distance between two points on the unit sphere can be measured by the spherical distance d. For $x, y \in S^{n-1}$ we have $d(x, y) = \arccos(x \cdot y)$, where $x \cdot y$ denotes the Euclidean inner product. E.g. antipodal points have spherical distance $d(x, -x) = \pi$. The unit sphere comes with a measure, the surface area ω which is induced by the Lesbegue measure on $\mathbb{R}^n$. We have $\omega(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$.*

In the distance graph $G(S^{n-1}, (0, \theta))$ each vertex x is adjacent to a spherical cap centered at x. Stable sets of $G(S^{n-1}, (0, \theta))$ are called *spherical codes with minimal angular distance θ*. They are of special interest in information theory. The stability number of this graph is also denoted by $A(n, \theta)$.

The kissing number problem is equivalent to the problem of finding $A(n, \pi/3)$. In geometry, the *kissing number problem* asks for the maximum number τ_n of unit spheres that can simultaneously touch the unit sphere in n-dimensional Euclidean space without pairwise overlapping. The touching points form a spherical code with

minimal angular distance $\pi/3$. The value of τ_n is only known for $n = 1, 2, 3, 4, 8, 24$. While its determination for $n = 1, 2$ is trivial, it is not the case for other values of n. The case $n = 3$ was the object of a famous discussion between Isaac Newton and David Gregory in 1694. For a historical perspective of this discussion we refer to Casselman [24]. The first valid proof of the fact "$\tau_3 = 12$", like in the icosahedron configuration, was only given in 1953 by Schütte and van der Waerden [83]. Odlyzko and Sloane [77], and independently Levenshtein [63], proved $\tau_8 = 240$ and $\tau_{24} = 196560$ which are respectively the number of shortest vectors in the root lattice E_8 and in the Leech lattice. In 2003, Musin [70] succeeded to prove the conjectured value $\tau_4 = 24$, which is the number of shortest vectors in the root lattice D_4. See also the survey Pfender, Ziegler [73]. The known lower and upper bounds for τ_n, with $n \leq 10$, are given in Bachoc, Vallentin [5].

2.2. **Original formulations for theta.** In the original paper [64], the Lovász theta function was given as the solution of a semidefinite program which involves a positive semidefinite matrix $K \in \mathbb{R}^{V \times V}$ whose rows and columns are indexed by the finite vertex-set of a graph $G = (V, E)$. In [64, Theorem 3, Theorem 4] Lovász gave the following two formulations of the theta function, which are equivalent by semidefinite programming duality:

$$\vartheta(G) = \max \Big\{ \sum_{x \in V} \sum_{y \in V} K(x, y) \ : \ K \in \mathbb{R}^{V \times V} \text{ is positive semidefinite,}$$

(2)
$$\sum_{x \in V} K(x, x) = 1,$$

$$K(x, y) = 0 \text{ if } \{x, y\} \in E \Big\},$$

and

$$\vartheta(G) = \min \Big\{ \lambda \ : \ K \in \mathbb{R}^{V \times V} \text{ is positive semidefinite,}$$

(3)
$$K(x, x) = \lambda - 1 \text{ for all } x \in V,$$

$$K(x, y) = -1 \text{ if } \{x, y\} \notin E \Big\}.$$

There are many equivalent definitions of the theta function. Possible alternatives are reviewed by Knuth [58]. One can rephrase (3) by saying that λ is the minimum of the largest eigenvalue of any symmetric matrix $K \in \mathbb{R}^{V \times V}$ such that $K(x, y) = 1$ whenever $x = y$ or if x and y are not adjacent.

Exercise 2.5. *Prove this statement.*

It is easy to see that the theta function gives an upper bound for the stability number of a finite graph: Let $C \subseteq V$ be a stable set of maximal cardinality. Consider the column vector $1_C \in \mathbb{R}^V$ which is the characteristic vector of C defined by $1_C(x) = 1$ if $x \in C$ and $1_C(x) = 0$ otherwise. Then, the rank-1 matrix $K = 1/|C|1_C(1_C)^t$, which componentwise is $K(x, y) = 1/|C|1_C(x)1_C(y)$, is feasible for (2). Thus, $\vartheta(G) \geq \alpha(G)$.

One can strengthen the equalities in (2) and (3) by the inequalities $K(x, y) \geq 0$ and $K(x, y) \leq -1$ respectively. This strengthening was introduced by Schrijver [81]. Sometimes it is denoted by $\vartheta'(G)$ and we have $\vartheta(G) \geq \vartheta'(G) \geq \alpha(G)$.

Using semidefinite programming one can compute the theta function in polynomial time, in the sense that one can approximate it with any given precision. In general the upper bound given by the theta function is weak: Feige [36] proved that there exists a constant c and an infinite family of graphs on n vertices for

which $\vartheta(G)/\alpha(G) > n/2^{c\sqrt{\log n}}$ holds. So in particular it does not give a $n^{1-\varepsilon}$-approximation for any fixed $\varepsilon > 0$. It is weak for a good reason: Håstad [46] showed that for any fixed $\varepsilon > 0$ one cannot approximate the stability number of a general graph with n vertices within a factor of $n^{1-\varepsilon}$ in polynomial time unless any problem in NP can be solved in expected, probabilistic polynomial time. However, as we will see later, it is sometimes surprisingly good, especially for symmetric graphs.

2.3. **Positive Hilbert-Schmidt kernels.** To be able to define the theta function for infinite graphs we need a concept of positive semidefinite matrices with infintely many rows and columns. We use positive Hilbert-Schmidt kernels for this, which are well-studied objects in functional analysis. Many familiar facts about positive semidefinite matrices can be generalized to positive Hilbert-Schmidt kernels. Historically, the central results of the theory of Hilbert-Schmidt kernels have been developed at the beginning of the 20th century mainly by Fredholm, Hilbert, Mercer, Schmidt. The classical text books Courant, Hilbert [30] and Riesz, Sz.-Nagy [75] are beautiful expositions. A modern treatment is contained for example in the comprehensive books by Reed, Simon [74].

Let us list some basic properties of positive Hilbert-Schmidt kernels. Here it is sometimes enlightening to compare positive Hilbert-Schmidt kernels with the more familiar positive semidefinite matrices.

By $\mathcal{C}(V)$ we denote the set of complex-valued continuous functions $f : V \to \mathbb{C}$ which is an inner product space by

$$(4) \qquad (f, g) = \int_V f(x)\overline{g(x)}d\mu(x),$$

and a normed space by $\|f\| = \sqrt{(f,f)}$. We say that a family of continuous functions $e_1, e_2, \ldots \in \mathcal{C}(V)$ is an *orthonormal system* if

$$(e_k, e_k) = 1 \text{ and } (e_k, e_l) = 0, \text{ whenever } k \neq l.$$

We say that it is *complete* if every continuous function can be approximated arbitrarily well by finite linear combinations in terms of convergence in the mean, i.e. convergence with respect to the norm $\|\cdot\|$ introduced above. Let $e_1, e_2, \ldots, e_d$ be an orthonormal system, then we have the fundamental *Bessel's inequality*

$$0 \leq \|f - \sum_{k=1}^{d}(f, e_k)e_k\|^2 = \|f\|^2 - \sum_{k=1}^{d}|(f, e_k)|^2.$$

By $\mathcal{C}(V \times V)$ we denote the set of continuous functions $K : V \times V \to \mathbb{C}$. The elements of $\mathcal{C}(V \times V)$ are traditionally called *kernels* because they appear as integral kernels in the theory of integral equations: For instance the *homogeneous Fredholm integral equation of the second type* is an integral equation of the form

$$(5) \qquad \lambda f(x) - \int_V K(x, y)f(y)d\mu(y) = 0,$$

where $K \in \mathcal{C}(V \times V)$ is a given integral kernel, and $\lambda \in \mathbb{C}$ and $f \in \mathcal{C}(V)$ are to be determined. In fact, solutions to (5) are given by *eigenvalues* and *eigenfunctions* of the linear map

$$T_K : \mathcal{C}(V) \to \mathcal{C}(V), \quad T_K(f)(x) = \int_V K(x, y)f(y)d\mu(y),$$

i.e. nontrivial $f \in \mathcal{C}(V)$ and $\lambda \in \mathbb{C}$ with $T_K(f) = \lambda f$.

In the following we only consider *symmetric* or *Hermitian* kernels. They satisfy $K(x, y) = \overline{K(y, x)}$ for all $x, y \in V$. A kernel $K \in \mathcal{C}(V \times V)$ is called *positive* if for any nonnegative integer m, any points $x_1, \ldots, x_m \in V$, and any complex numbers $u_1, \ldots, u_m$, we have

$$\sum_{i=1}^{m} \sum_{j=1}^{m} K(x_i, x_j) u_i \overline{u_j} \geq 0.$$

So any finite matrix $(K(x_i, x_j))_{i,j}$ which one extracts from K has to be positive semidefinite. We refer to positive, symmetric, continuous kernels as *positive Hilbert-Schmidt kernels* or simply as *positive kernels*.

Let us turn to the spectral decomposition of positive Hilbert-Schmidt kernels $K \in \mathcal{C}(V \times V)$ where we first recall what happens for positive semidefinite matrices $A \in \mathbb{C}^{n \times n}$ (or in a more snobbish notation $A \in \mathcal{C}(\{1, \ldots, n\} \times \{1, \ldots, n\})$).

Positive semidefinite matrices $A \in \mathbb{C}^{n \times n}$ can be diagonalized: There is an orthonormal basis $e_1, \ldots, e_n$ of $\mathbb{C}^n$ consisting of eigenvectors of A, all eigenvalues $\lambda_1, \ldots, \lambda_n$ are nonnegative real numbers, and the (x, y) entry of A can be written as

$$A(x, y) = \sum_{k=1}^{n} \lambda_k e_k(x) \overline{e_k(y)}.$$

The *rank* of A is the number of nonzero eigenvalues, counted with multiplicity. Suppose that the eigenvalues are ordered in descending order

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0,$$

then they can be determined from the following max-min characterization:

$$\lambda_k = \max_{S_k} \min_{f \in S_k, (f, f)=1} (Af, f),$$

where S_k runs through all k-dimensional subspaces of $\mathbb{C}^n$. Geometrically, the eigenvectors form the principal axes of the ellipsoid $\{f \in \mathbb{C}^n : (Af, f) = 1\}$ defined by A. The square roots of the reciprocals of the corresponding eigenvalues give the length of these axes.

Essentially the same statements hold true for positive Hilbert-Schmidt kernels. The main difference between positive semidefinite matrices and positive kernels is that the latter have infinitely many, different eigenvalues.

Let $K \in \mathcal{C}(V \times V)$ be a positive Hilbert-Schmidt kernel. To get the flavour of the reasoning one needs when working with Hilbert-Schmidt kernels we show using Bessel's inequality that the multiplicity of every nonzero eigenvalue is finite: Let $e_1, \ldots, e_d$ be an orthonormal system of eigenfunctions of a nonzero eigenvalue λ. For fixed $x \in V$ we apply Bessel's inequality to the function $y \mapsto K(x, y)$ and get

$$\int_V |K(x, y)|^2 d\mu(y) \geq \sum_{k=1}^{d} \left| \int_V K(x, y) e_k(y) d\mu(y) \right|^2 = \sum_{k=1}^{d} |\lambda e_k(x)|^2.$$

Integrating both sides again yields an upper bound for d:

$$\int_V \int_V |K(x, y)|^2 d\mu(y) d\mu(x) \geq |\lambda|^2 d.$$

Furthermore, one can show that the series of squared eigenvalues converges to 0. So there are two possibilities: K has finitely many or infinitely many nonzero eigenvalues. In the latter case, the only accumulation point of the eigenvalues is 0. The *rank* of a kernel is the number of nonzero eigenvalues counted with multiplicity.

We have the *spectral decomposition* of a positive Hilbert-Schmidt kernel:

$$K(x,y) = \sum_k \lambda_k e_k(x) \overline{e_k(y)}.$$

Here λ_k are the eigenvalues of K, which are all nonnegative, and the e_k form a complete orthonormal system. In the case of infinitely many different eigenvalues the right hand side converges absolutely and uniformly to the left hand side. In the case of finitely many different eigenvalues, the finite rank case, the series degenerates to a finite sum. In the literature this statement about the spectral decomposition of positive Hilbert-Schmidt kernels is often called *Mercer's theorem.*

The max-min characterization of the eigenvalues of K reads as follows: Suppose that the eigenvalues of K are ordered in descending order

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq 0,$$

then they can be determined from the following max-min characterization:

$$\lambda_k = \max_{S_k} \min_{f \in S_k, (f,f)=1} (T_K(f), f),$$

where S_k runs through all k-dimensional subspace of $\mathcal{C}(V)$. In fact the max-min characterization of the eigenvalues is used in Courant, Hilbert [30] and Riesz, Sz.-Nagy [75] to prove the spectral decomposition theorem.

Exercise 2.6. *Determine the eigenvalues of the inner product kernel $K : S^{n-1} \times S^{n-1} \to \mathbb{C}$ defined by $K(x,y) = x \cdot y$. Give the spectral decomposition of it in the case $n = 2$. (Maybe you want to come back to this exercise after Section 5.)*

2.4. **The theta function for infinite graphs.** Now we are ready to generalize the theta function to possibly infinite distance graphs $G(V, I)$. It will turn out that for the combinatorial and geometric packing problems we consider it is convenient to work with the second formulation (3); possibilities to work with the first formulation are considered in Bachoc, Nebe, de Oliveira Filho, Vallentin [9]. In the second formulation we simply replace the condition that K is a positive semidefinite matrix by the condition that K is a positive kernel.

$$
\begin{aligned}
\vartheta(G(V,I)) = \min \big\{ \lambda \ \ : \ \ & K \in \mathcal{C}(V \times V) \text{ is positive,} \\
& K(x,x) = \lambda - 1 \text{ for all } x \in V, \\
& K(x,y) = -1 \text{ if } \{x,y\} \notin E \big\}.
\end{aligned}
$$

(6)

When we are working with an infinite-dimensional semidefinite program we can no longer use the simple duality argument given in Section 2.2 to show that the theta function gives an upper bound for the stability number. So we have to provide a new (simple) one: Instead of constructing a feasible kernel from a stable set and going to the dual semidefinite program, we show that any feasible kernel can be used as a test function on any stable set C to bound the cardinality of C.

Theorem 2.7. *Let $G(V, I)$ be a distance graph. Then, the theta function is an upper bound for the stability number: $\vartheta(G(V, I)) \geq \alpha(G(V, I))$.*

Proof. Let C be a stable set of $G(V, I)$ and let K be a kernel which satisfies the conditions in (6). We consider the $|C| \times |C|$ matrix $(K(c, c'))_{(c,c')\in C^2}$ which we extract from K. Because this matrix is positive semidefinite we have

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

On the other hand,

$$\sum_{(c,c')\in C^2} K(c, c') = \sum_c K(c, c) + \sum_{c \neq c'} K(c, c') \leq |C|K(c, c) - |C|(|C| - 1),$$

so that $|C| - 1 \leq K(c, c)$ and the statement follows. $\square$

In the same way one can define $\vartheta'(G(V, I))$ by strengthening the condition $K(x, y) = 1$ to $K(x, y) \leq -1$. Then, one has

$$\vartheta(G(V, I)) \geq \vartheta'(G(V, I)) \geq \alpha(G(V, I)).$$

2.5. **Exploiting symmetry.** *We should start to speak about symmetry.* In this section we show the basic step to simplify the computation of the theta function for symmetric distance graphs. Note that this basic step can be applied to other semidefinite programs which are invariant under the action of a group. Once one has identified this symmetry, this basic step, as well as the following steps, can be applied to some extend mechanically.

Usually many symmetries of a distance graph come from the *automorphism group* of the underlying metric space V. It is the group of all permutations $u : V \to V$ which leave the distance map d invariant:

$$\mathrm{Aut}(V) = \{u : V \to V : d(x, y) = d(ux, uy) \text{ for all } x, y \in V\}.$$

Let us take a look at the automorphism group of the three examples 2.2–2.4:

Example 2.8. *The automorphism group of a metric space defined by the shortest paths in a connected graph $G = (V, E)$ equals the automorphism group of the underlying graph:*

$$\mathrm{Aut}(G) = \{u : V \to V : \{x, y\} \in E \text{ if and only if } \{ux, uy\} \in E\}.$$

Computing the automorphism group is not easy. Deciding whether the automorphism group is trivial is as difficult as the graph isomorphism problem. For this no polynomial time algorithm is known. The graph isomorphism problem is generally believed to lie in NP $\cap$ co-NP. So it is unlikely that it is NP-hard. For more information on the computational complexity of this problem we refer to the book Köbler, Schöning, Tóran [59]. On the practical side the program `nauty` of McKay [67] is a very useful tool for computing the automorphism group of a graph.

Example 2.9. *The automorphism group of the Hamming cube $\{0, 1\}^n$ has order $2^n n!$. It is generated by all $n!$ permutations of the n coordinates and all 2^n switches $0 \leftrightarrow 1$ which one can identify with the Hamming cube $\{0, 1\}^n$ itself where one considers addition modulo 2.*

Example 2.10. *The automorphism group of the unit sphere is the* orthogonal group. *It is the group of orthogonal matrices*

$$\mathrm{Aut}(S^{n-1}) = \mathrm{O}(\mathbb{R}^n) = \{u \in \mathbb{R}^{n \times n} : u^t u = I_n\},$$

where I_n is the $(n \times n)$-identity matrix.

The orthogonal group is generated by a reflection over some hyperplane and all rotations. It preserves the inner product (and thus the spherical distance) as well as the Euclidean distance. Another way to view it is to observe that the elements of $O(\mathbb{R}^n)$ map orthonormal systems of $\mathbb{R}^n$ to orthonormal systems.

$$\boxed{\text{Crucial observation}}$$

Now we come to a crucial observation: In the computation of the theta function (6) of a distance graph $G(V, I)$ one can restrict the semidefinite program to positive Hilbert-Schmidt kernels which are invariant under the automorphism group of V. The optimal objective value of this restricted semidefinite program remains unchanged, and the matrix sizes which are needed in the practical computation can be drastically smaller.

We say that a positive kernel $K \in \mathcal{C}(V \times V)$ is $\text{Aut}(V)$-*invariant* if $K(ux, uy) = K(x, y)$ holds for all $u \in \text{Aut}(V)$ and all $x, y \in V$. For instance, the positive kernel $K \in \mathcal{C}(S^{n-1} \times S^{n-1})$ which is defined by the inner product $K(x, y) = x \cdot y$ is $O(\mathbb{R}^n)$-invariant.

If the graph is finite, this crucial observation is easy. If K is a feasible solution of (6) so is its $\text{Aut}(V)$-invariant *group average*

$$(7) \qquad \tilde{K}(x, y) = \frac{1}{|\text{Aut}(V)|} \sum_{u \in \text{Aut}(V)} K(ux, uy).$$

This is also easy to verify, e.g. we have for $x, y \in V$ with $\{x, y\} \notin E$

$$\tilde{K}(x, y) = \frac{1}{|\text{Aut}(V)|} \sum_{u \in \text{Aut}(V)} K(ux, uy) = \frac{|\text{Aut}(V)|}{|\text{Aut}(V)|}(-1) = -1.$$

The matrix $\tilde{K}$ is positive semidefinite because it is a nonnegative sum of $|\text{Aut}(V)|$ positive semidefinite matrices $K_u(x, y) = K(ux, uy)$. Also the objective values of K and $\tilde{K}$ coincide.

If the distance graph is infinite we need a replacement for the finite sum (7). For this we use the invariant integral of the group $\text{Aut}(V)$. If $\text{Aut}(V)$ is finite, then the invariant integral is given by the sum (7). Generally, this invariant integral is defined for compact topological groups and it can be constructed from the Haar measure (see e.g. the original Haar [45], or the standard text on measure theory Taylor [88]). Lovász [65] gives an elementary, combinatorial construction of the invariant integral based on the marriage theorem of matching theory.

Here we only give the defining properties of the invariant integral. For the explicit construction we refer the interested reader to the above mentioned literature. Actually one could establish the invariant integral in our three examples directly.

Let Γ be a compact topological group. A *topological group* is a group which is equipped with a topology so that multiplication and inversion are continuous functions. Examples of compact topological groups are all finite groups and the orthogonal group $O(\mathbb{R}^n)$. The topology here comes from the topology defined on $(n \times n)$-matrices. A topological group which is not compact is the group of translations $(\mathbb{R}^n, +)$.

We consider the space of complex-valued continuous functions on the group $\mathcal{C}(\Gamma)$. Then there is a unique map $\int_\Gamma : \mathcal{C}(\Gamma) \to \mathbb{C}$, the *invariant integral*, with the following properties:

(a) Linearity

$$\int_\Gamma \alpha f + \beta g = \alpha \int_\Gamma f + \beta \int_\Gamma g \quad \text{for all } f, g \in \mathcal{C}(\Gamma),\ \alpha, \beta \in \mathbb{C}.$$

(b) Monotonicity

$$\int_\Gamma |f| \geq 0 \quad \text{for all } f \in \mathcal{C}(\Gamma).$$

(c) Normalization

$$\int_\Gamma 1 = 1.$$

(d) Invariance

$$\int_\Gamma f(vu) = \int_\Gamma f(u) \quad \text{for all } v \in \Gamma.$$

Two quick examples: If Γ is a finite group, the invariant integral simply is

$$\int_\Gamma f = \frac{1}{|\Gamma|} \sum_{u \in \Gamma} f(u).$$

For the two-dimensional rotation group $\mathrm{SO}(\mathbb{R}^2) = \{u \in \mathrm{O}(\mathbb{R}^2) : \det u = 1\}$ the invariant integral is

$$\int_{\mathrm{SO}(\mathbb{R}^2)} f = \frac{1}{2\pi} \int_0^{2\pi} f\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right) d\theta.$$

Now the proper replacement of the group average given in (7) for general distance graphs $G(V, I)$ is

$$(8) \qquad\qquad \tilde{K}(x, y) = \int_{\mathrm{Aut}(V)} K(ux, uy),$$

where we take the invariant integral of the function $u \mapsto K(ux, uy) \in \mathcal{C}(\Gamma)$.

3. Harmonic analysis

In the last section we showed that in the computation of the theta function it suffices to consider invariant kernels only. If one chooses an appropriate basis of the invariant kernels, then it is possible to express positivity in an efficient way, namely as the positive semidefiniteness of a couple of smaller block matrices. To make this statement precise we have to use some harmonic analysis which we develop in this section starting from basic principles.

To find an appropriate basis we use basic tools from harmonic analysis and representation theory of compact groups which are usually associated with the Peter-Weyl theorem. In Section 3.1 we present the basic notions and in Section 3.2 we state and prove the Peter-Weyl theorem. In Section 3.3 we apply it to find Bochner's characterization of positive, invariant kernels. Throughout this section we develop the theory with the help of the spectral theory of positive kernels. Then, in the following sections we show how the Peter-Weyl theorem and Bochner's

characterization specializes for the computation of the theta functions of distance graphs on the Hamming cube and on the unit sphere.

Literature: The literature on harmonic analysis and group representations is huge. Probably the book by Serre [84] on the linear representation theory of finite groups is the most prominent. Other books which emphasize the connection of group representations of finite groups to probability is Diaconis [34], one emphasizing the connection to algebraic combinatorics is Sagan [79], and one connecting to chemistry, error-correcting codes, data analysis, graph theory, and probability is Terras [89]. Further good sources are: Bröcker and tom Dieck [20], Vinberg [93], Carter, Segal, Macdonald [23], Goodman, Wallach [40]. The articles Borel [16], Gross [41] and Slodowy [85] illuminate the historical background.

3.1. **Basic tools.** Let us recall the set-up: V is a compact metric space with measure μ, and $\Gamma = \mathrm{Aut}(V)$ is the automorphism group of V. We define the action of Γ on the linear space of complex-valued continuous functions $\mathcal{C}(V)$ by

$$uf(x) = f(u^{-1}x).$$

We assume that the measure μ is Γ-invariant ($\mu(uA) = \mu(A)$), so that we have an *invariant inner product* (4) on the space of continuous functions:

$$(f, g) = (uf, ug) \text{ for all } f, g \in \mathcal{C}(V) \text{ and } u \in \Gamma.$$

The action of Γ on the vector space $\mathcal{C}(V)$ is linear, that is we have

$$u(\alpha f + \beta g) = \alpha uf + \beta ug \text{ for all } f, g \in \mathcal{C}(V), \ \alpha, \beta \in \mathbb{C} \text{ and } u \in \Gamma.$$

Linear actions are called *group representations*. More vocabulary: A subspace $S \subseteq \mathcal{C}(V)$ is called Γ-*invariant* if $uS = S$ for all $u \in \Gamma$, i.e. if for every $u \in \Gamma$ and for every $f \in S$ we have $uf \in S$ as well. A nonzero subspace S is called Γ-*irreducible* if $\{0\}$ and S are the only Γ-invariant subspaces of S. Let S and S' be two invariant subspaces. A linear map $T : S \to S'$ is called a Γ-*map* if $T(uf) = uT(f)$ for all $u \in \Gamma$, and $f \in \mathcal{C}(V)$. We say that S and S' are Γ-equivalent if there is a bijective Γ-map between them. If it is clear from the context which group we consider, we frequently omit the prefix "Γ-".

Two little lemmas will make our live easier.

The first one is also called *Maschke's theorem.*

Lemma 3.1. *Let S be a invariant subspace and let $U \subseteq S$ be an invariant subspace of S. Then its orthogonal complement in S, given by*

$$U^{\perp} = \{g \in S : (f, g) = 0 \text{ for all } f \in U\}$$

is invariant as well.

Proof. This comes from the invariant inner product: We have for all $u \in \Gamma$

$$(f, ug) = (u^{-1}f, g) = 0 \text{ whenever } f \in U, \ g \in U^{\perp}. \qquad \square$$

Maschke's theorem implies that a finite-dimensional invariant subspace which is not irreducible splits into an orthogonal sum of irreducible subspaces.

The second lemma is *Schur's lemma.*

Lemma 3.2. *Let S and S' be two irreducible subspaces, and let $T : S \to S'$ be a Γ-map. If S and S' are not equivalent, then $T = 0$. If they are equivalent then either $T = 0$ or T is bijective.*

Proof. The kernel of T is an invariant subspace of S, and the image is an invariant subspace of S'. Since S and S' are irreducible subspaces we are left with four possibilities:

(1) The kernel of T is $\{0\}$ and the image of T is $\{0\}$. This could only happen when $S = \{0\}$.
(2) The kernel of T is $\{0\}$ and the image of T is S'. Then, T is a bijective Γ-map and S and S' are equivalent subspaces.
(3) The kernel of T is S and the image of T is $\{0\}$. Then, T must be the zero map.
(4) The kernel of T is S and the image of T is S'. This could only happen when $S' = \{0\}$. $\square$

We apply Schur's lemma to establish some orthogonality relations which will be essential later.

Lemma 3.3. *Let S and S' be two irreducible subspaces. Let $e_1, \ldots, e_h$ be a complete orthonormal system of S and let $e'_1, \ldots, e'_{h'}$ be one of S'.*

(1) *If S and S' are not equivalent, then they are orthogonal to each other:*
$$(e_i, e'_j) = 0, \text{ for } i = 1, \ldots, h, \ j = 1, \ldots, h'$$

(2) *If S and S' are equivalent and if $T : S \to S'$ is a bijective Γ-map, mapping e_i to e'_i, then there is a constant c so that*
$$(e_i, e'_j) = \begin{cases} c & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Define the linear map $A : S \to S'$ by

$$(9) \qquad\qquad A(e_i) = \sum_{j=1}^{h'} (e'_j, e_i) e'_j.$$

This is a Γ-map. Now the first claim follows immediately from Schur's lemma: We have $A = 0$ and so $(e'_j, e_i) = 0$. For the second claim we assume $A \neq 0$. Then A is bijective by Schur's lemma. Consider the endomorphism $T^{-1}A : S \to S$, $T^{-1}A(e_i) = \sum_{j=1}^{h} (e'_j, e_i) e_j$. Since we work over the complex numbers and $T^{-1}A$ is bijective, it has a nonzero eigenvalue c. The corresponding eigenspace is an invariant subspace of S, so it has to be equal to S. Hence $T^{-1}A$ is c times the identity map, which proves the second claim. $\square$

Exercise 3.4. *Check that the map A defined in (9) is indeed a Γ-map.*

3.2. **The Peter-Weyl theorem.** The theorem of Peter and Weyl [72] and [95] is the starting point of harmonic analysis of compact groups. It connect Fourier analysis with group representations. It shows that the space $\mathcal{C}(V)$ decomposes orthogonally into finite-dimensional irreducible subspaces and that the space $\mathcal{C}(V)$ has a complete orthonormal system which is "in harmony" with the group $\Gamma = \text{Aut}(V)$.

Theorem 3.5. *All irreducible subspaces of $\mathcal{C}(V)$ are of finite dimension. The space $\mathcal{C}(V)$ decomposes orthogonally as*

$$\mathcal{C}(V) = \bigoplus_{k=0,1,\ldots} H_k,$$

and the space H_k decomposes orthogonally as

$$H_k = \bigoplus_{i=1,2,\ldots,m_k} H_{k,i},$$

where $H_{k,i}$ is irreducible, and $H_{k,i}$ is equivalent to $H_{k',i'}$ if and only if $k = k'$. The dimension h_k of $H_{k,i}$ is finite, but the multiplicity m_k can potentially be infinite.

In other words, $\mathcal{C}(V)$ has a complete orthonormal system $e_{k,i,l}$, where $k = 0, 1, \ldots$, $i = 1, 2, \ldots, m_k$, $l = 1, \ldots, h_k$ so that

(1) *the space $H_{k,i}$ spanned by $e_{k,i,1}, \ldots, e_{k,i,h_k}$ is irreducible,*
(2) *the spaces $H_{k,i}$ and $H_{k',i'}$ are equivalent if and only if $k = k'$,*
(3) *there are Γ-maps $\phi_{k,i} : H_{k,1} \to H_{k,i}$ mapping $e_{k,1,l}$ to $e_{k,i,l}$.*

Proof. Consider an element $f \in \mathcal{C}(V)$. Define the positive Γ-invariant kernel $K \in \mathcal{C}(V \times V)$ using invariant integration

$$K(x,y) = \int_\Gamma f(ux)\overline{f(uy)}.$$

The eigenspaces of T_K of nonzero eigenvalues are finite-dimensional. They are invariant because for an eigenfunction $g \in \mathcal{C}(V)$ to an eigenvalue λ we have

$$\begin{aligned}
T_K(ug)(x) &= \int_V K(x,y)g(u^{-1}y)d\mu(y) \\
&= \int_V K(x,uy)g(y)d\mu(y) \\
&= \int_V K(u^{-1}x,y)g(y)d\mu(y) \\
&= \lambda g(u^{-1}x) \\
&= \lambda u g(x).
\end{aligned}$$

Now the statement follows: One breaks the invariant finite-dimensional eigenspaces into irreducible subspaces using Maschke's theorem. The orthonormal system one constructs using Gram-Schmidt orthonormalization. The orthogonality relation follow from Lemma 3.3. Completeness follows form the spectral decomposition of positive kernels. □

3.3. **Bochner's characterization.** The complete orthonormal system $e_{k,i,l}$ of the Peter-Weyl theorem is very useful to characterize Γ-invariant, positive kernels. This is the contents of the following theorem by Bochner [15].

Theorem 3.6. *Let $e_{k,i,l}$ be a complete orthonormal system for $\mathcal{C}(V)$ as in Theorem 3.5. Every Γ-invariant, positive kernel $K \in \mathcal{C}(V \times V)$ can be written as*

(10) $$K(x,y) = \sum_{k=0,1,\ldots} \sum_{i,j=1,2\ldots,m_k} f_{k,ij} \sum_{l=1}^{h_k} e_{k,i,l}(x)\overline{e_{k,j,l}(y)},$$

or more economically as

(11) $$K(x,y) = \sum_{k=0,1,\ldots} \langle F_k, Z_k^{(x,y)} \rangle,$$

with $(F_k)_{ij} = f_{k,ij}$ and $(Z_k^{(x,y)})_{ij} = \sum_{l=1}^{h_k} e_{k,i,l}(x)\overline{e_{k,j,l}(y)}$. Here F_k is Hermitian and positive. The series converges absolutely and uniformly.

Proof. It is clear the every kernel of the form (10) or (11) is a positive Γ-invariant kernel. Generally, a kernel $K \in \mathcal{C}(V \times V)$ can be written as a series in the basis $(x, y) \mapsto e_{k,i,l}(x)\overline{e_{k',i',l'}(y)}$. Let $f_{k,i,l;k',i',l'}$ be the corresponding coefficient. We shall show that $f_{k,i,l;k',i',l'} = 0$ if $k \neq k'$ or $l \neq l'$ and that $f_{k,i,l;k,i',l}$ does not depend on l, so that we can set $f_{k,i,l;k,i',l} = f_{k,ii'}$. For this we consider the, possibly degenerate, Γ-invariant inner product on $\mathcal{C}(V)$ defined by K:

$$(f, g)_K = \int_V \int_V f(x) K(x, y) \overline{g(y)} d\mu(x) d\mu(y).$$

For this inner product, the orthogonality relations of Lemma 3.3 apply, which proves the claim about the coefficients $f_{k,i,l;k',i',l'}$. Furthermore, the inner product (3.3) defines an inner product on the space spanned by the vectors

$$\varphi_{k,i} = (e_{k,i,1}, \ldots, e_{k,i,h_k}), \quad i = 1, \ldots, m_k,$$

by

$$(\varphi_{k,i}, \varphi_{k,i'}) = \left(\sum_{l=1}^{h_k} e_{k,i,l}, \sum_{l=1}^{h_k} e_{k,i',l} \right)_K.$$

Hence, $F_k = (f_{k,ii'})$ is positive. $\square$

Bochner's characterization is *the* tool for exploiting symmetry in semidefinite programs. If one wants to optimize over all Γ-invariant kernels, then one only has to optimize over positive F_k's which have size $m_k \times m_k$. The computation of $Z_k^{(x,y)}$ is part of the preprocessing. It can be done by hand, as in the next two sections, or by computer if the group is finite and has moderate size. See the article Babai, Rónyai [3] for algorithmic aspects.

4. Boolean harmonics

In this section we specialize the Peter-Weyl theorem and the characterization of Bochner to the n-dimensional Hamming cube $\{0, 1\}^n$. We apply these results to calculate the theta function for distance graph on $\{0, 1\}^n$. Then we give a complete proof of the specialization using Fourier analysis of Boolean functions.

4.1. **Specializations.** Recall that the space of continuous functions $\mathcal{C}(\{0, 1\}^n)$, also known as complex vectors indexed by $\{0, 1\}^n$, comes with an inner product

$$(f, g) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) \overline{g(x)},$$

which is invariant under the action of the automorphism group. The group has order $2^n n!$ and it is generated by permutations and switches $0 \leftrightarrow 1$. We make extensive use of the functions $\chi_y(x) = (-1)^{y \cdot x}$, which are called *characters*. The character χ_y evaluates to 1 if the Hamming distance between x and y is an even integer and to -1 if it is an odd integer.

The Peter-Weyl theorem specializes as follows.

Theorem 4.1. *The space $\mathcal{C}(\{0, 1\}^n)$ decomposes orthogonally into irreducible subspaces H_k:*

$$\mathcal{C}(\{0, 1\}^n) = H_0 \perp H_1 \perp \ldots \perp H_n.$$

Here H_k denotes the subspace spanned by the functions $\chi_y(x) = (-1)^{y \cdot x}$, with Hamming norm $\|y\| = k$. It is an irreducible subspace of dimension $h_k = \binom{n}{k}$.

In the decomposition all irreducible subspaces have multiplicity 1. We have $H_k = H_{k,1}$ in the notation of Theorem 3.5. Hence, Bochner's characterization (Theorem 3.6) of the $\mathrm{Aut}(\{0,1\}^n)$-invariant, positive kernels only involves positive semidefinite matrices of size 1×1, i.e. nonnegative scalars. The automorphism group acts *distance transitively* on pairs: For every two pairs of points (x,y) and (x',y') having the same Hamming distance $\delta(x,y) = \delta(x',y')$ there is a group element $u \in \mathrm{Aut}(\{0,1\}^n)$ which transforms the pairs into each other $(ux, uy) = (x', y')$. So $\mathrm{Aut}(\{0,1\}^n)$-invariant kernels are actually functions depending only on the distance δ.

Theorem 4.2. *Every* $\mathrm{Aut}(\{0,1\}^n)$-*invariant, positive kernel* $K \in \mathcal{C}(\{0,1\}^n \times \{0,1\}^n)$ *is of the form*

$$K(x,y) = \sum_{k=0}^{n} f_k K_k^n(\delta(x,y)), \quad f_0, \ldots, f_n \geq 0,$$

where K_k^n *denotes a the Krawtchouk polynomial of degree* k.

Krawtchouk polynomials $K_k^n(t)$ *are orthogonal polynomials for the inner product*

$$(P, Q) = \sum_{t=0}^{n} \binom{n}{t} P(t) Q(t),$$

which are normalized by $K_k^n(0) = \binom{n}{k}$. The first few Krawtchouk polynomials are

$$K_0^n(t) = 1,$$
$$K_1^n(t) = -2t + n,$$
$$K_2^n(t) = 2t^2 - 2nt + \binom{n}{2},$$

and in general

$$K_k^n(t) = \sum_{i=0}^{k} \binom{t}{i}\binom{n-t}{k-i}(-1)^i.$$

See the books Szegö [87] or MacWilliams, Sloane [66] for more information on Krawtchouk polynomials.

4.2. **Linear programming bound for binary codes.** Now the calculation of the theta function on the Hamming cube reduces to Delsarte's linear programming bound which was introduced by Delsarte [32]. The connection between Lovász' theta function and Delsarte's linear programming bound was first observed by McEliece, Rodemich, Rumsey Jr. [76] and independently by Schrijver [81].

$$\vartheta'(\{0,1\}^n, \{1, \ldots, d-1\}) = \min \Big\{ \lambda \; : \; K \in \mathcal{C}(\{0,1\}^n \times \{0,1\}^n)$$

$$\text{positive semidefinite,}$$
$$\mathrm{Aut}(\{0,1\}^n)\text{-invariant,}$$
$$K(x,x) = \lambda - 1 \text{ for all } x \in \{0,1\}^n,$$
$$K(x,y) \leq -1 \text{ if } \delta(x,y) \in \{d, \ldots, n\} \Big\}.$$

Now we represent the positive semidefinite matrix $K \in \mathcal{C}(\{0,1\}^n \times \{0,1\}^n)$ which is $\mathrm{Aut}(\{0,1\}^n)$-invariant by $K(x,y) = \sum_{k=0}^{n} f_k K_k^n(\delta(x,y))$ with $f_0, \ldots, f_n \geq 0$,

using the characterization of Bochner, and get the following linear program in the variables $f_0, \ldots, f_n$.

$$\min \Big\{ 1 + \sum_{k=0}^{n} f_k K_k^n(0) \ : \ f_0, \ldots, f_n \geq 0,$$
$$\sum_{k=0}^{n} f_k K_k^n(t) \leq -1 \text{ if } t \in \{d, \ldots, n\} \Big\}.$$

Exercise 4.3. *Calculate $\vartheta'(\{0,1\}^n, \{1, \ldots, d-1\})$ for, say, all $n = 1, \ldots, 128$ and all d.*

4.3. **Fourier analysis on the Hamming cube.** To show Theorem 4.1 and Theorem 4.2 we make use of discrete Fourier analysis of Boolean functions. This theory has recently been successfully developed and applied to problems concerning threshold phenomena and influence in combinatorics, computer science, economics, and political science. E.g. Håstad's result on the hardness of the stability number, we refered earlier to, uses Fourier analysis of Boolean functions. Kalai and Safra [50] survey these developments. One can find further information on Boolean harmonics e.g. in the book Terras [89], in the paper Dunkl [35], or in Delsarte's thesis [32].

A complex-valued function on the n-dimensional Hamming cube $f : \{0,1\}^n \to \mathbb{C}$ has the *Fourier expansion*

$$f(x) = \sum_{y \in \{0,1\}^n} \widehat{f}(y) \chi_y(x),$$

with the characters $\chi_y(x) = (-1)^{y \cdot x}$ and the *Fourier coefficients* $\widehat{f}(y)$. The characters χ_y, $y \in \{0,1\}^n$, which sometimes are called *Walsh functions*, form an orthonormal basis of the space $\mathcal{C}(\{0,1\}^n)$. The one-dimensional subspace spanned by each χ_y is invariant under the abelian group $(\{0,1\}^n, +)$ of addition modulo 2, which corresponds to the switching $0 \leftrightarrow 1$:

$$u\chi_y(x) = \chi_y(x - u) = (-1)^{y \cdot (x-u)} = (-1)^{y \cdot (-u)} \chi_y(x).$$

Under the complete automorphism group of the Hamming cube, i.e. when also permutations are allowed, then the one-dimensional subspaces are grouped together according the Hamming norm of y. The $\binom{n}{k}$-dimensional spaces

$$H_k = \lin\{\chi_y : \|y\| = k\}, \quad k = 0, \ldots, n,$$

are $\mathrm{Aut}(\{0,1\}^n)$-invariant. We shall show that they are irreducible and compute the corresponding basis function Z_k for Bochner's characterization.

For this the notion of zonal spherical function is helpful. A function $f : \{0,1\}^n \to \mathbb{C}$ is called a *zonal spherical function with pole $e \in \{0,1\}^n$* if $f = uf$ for all $u \in \mathrm{Aut}(\{0,1\}^n)$ which stabilize the point e, so that $ue = e$.

Lemma 4.4. *Let $U \subseteq \mathcal{C}(\{0,1\}^n)$ be an invariant subspace and let $e \in \{0,1\}^n$. If the subspace of zonal spherical functions with pole e which lie in U has dimension exactly one, then U is irreducible.*

Proof. Suppose U is not irreducible and U decomposes into nontrivial invariant subspaces $U = V \perp V^\perp$. Let $e_1, \ldots, e_n$ be an orthonormal system of V and let $f_1, \ldots, f_m$ be one of $V^\perp$. Then the functions $z_V(x) = \sum_{i=1}^{n} e_i(e) \overline{e_i(x)}$, and $z_{V^\perp}(x) = \sum_{i=1}^{n} f_i(e) \overline{f_i(x)}$ are two linearly independent zonal spherical functions with pole e. $\qquad\square$

We apply this lemma to H_k. Let $f \in H_k$ be a zonal spherical function with pole $0^n = (0, \ldots, 0)$. The group stabilizing 0^n is the group of all permutations. For every permutation u we have $uf = f$ and in terms of the Fourier expansion:

$$\sum_{y \in \{0,1\}^n, \|y\|=k} \widehat{f}(y) \chi_{uy}(x) = \sum_{y \in \{0,1\}^n, \|y\|=k} \widehat{f}(y) \chi_y(x).$$

So the Fourier coefficients $\widehat{f}(y)$ all have to coincide.

Exercise 4.5. *Show that the H_k's are pairwise not equivalent.*

Let us calculate the basis function $Z_k(x, x')$ for Bochner's characterization (and in this way the zonal spherical functions of H_k because a zonal spherical function with pole e is a multiple of $Z_k(e, \cdot)$ as one sees from the proof of Lemma 4.4). We have

$$Z_k(x, x') = \sum_{\|y\|=k} \chi_u(x) \overline{\chi_u(x')}.$$

Since Z_k is invariant under $\mathrm{Aut}(\{0,1\}^n)$ it only depends on the Hamming distance $t = \delta(x, x')$ and so we can assume that $x = 0$, and $x' = 1^t 0^{n-t}$. A straightforward combinatorial calculation yields

$$Z_k(x, x') = Z_k(0, 1^t 0^{n-t}) = \sum_{i=0}^{k} (-1)^i \binom{t}{i} \binom{n-t}{k-i} = K_k^n(t).$$

Because of Lemma 3.3 we have the orthogonality relation

$$\sum_{t=0}^{n} \binom{n}{t} K_k^n(t) K_{k'}^n(t) = 0, \quad \text{if } k \neq k'.$$

5. Spherical harmonics

In this section we replace the n-dimensional Hamming cube by the unit sphere S^{n-1} in n-dimensional Euclidean space. First we specialize the Peter-Weyl theorem and Bochner's characterization. We apply these results to calculate the theta function for distance graph on the sphere. This gives tight bounds for the kissing number in dimension 8 and 24. Then we give a complete proof of the specialization using the theory of spherical harmonics.

Spherical harmonics have been extensively studied due to their importance in mathematics, physics and engineering. Comprehensive information is available. Here we follow to some extend the book [2, Chapter 9] by Andrews, Askey and Roy, and the paper [28] by Coifman and Weiss which links spherical harmonics with the representation theory of compact topologic groups. A lot of information is contained in the book [92] by Vilenkin and Klimyk. Also the book [68] by Müller is very useful.

Spherical harmonics serve as a full-fledged example of geometric analysis: See the books [14] by Berger and [47] by Helgason. The paper by Stanton [86] emphasizes the similarities between spherical harmonics and Boolean harmonics.

5.1. **Specializations.** Recall that the space of continuous functions $\mathcal{C}(S^{n-1})$ comes with an inner product

$$(f, g) = \int_{S^{n-1}} f(x)\overline{g(x)}d\omega(x),$$

which is invariant under the action of the orthogonal group $\mathrm{O}(\mathbb{R}^n)$, the automorphism group of the unit sphere.

The Peter-Weyl theorem specializes as follows.

Theorem 5.1. *The space of complex-valued continuous functions on the unit sphere decomposes orthogonally into pairwise* $\mathrm{O}(\mathbb{R}^n)$*-irreducible subspaces as follows:*

$$\mathcal{C}(S^{n-1}) = H_0 \perp H_1 \perp H_2 \perp \ldots$$

Here H_k denotes the space of homogeneous polynomial functions of degree k which vanish under the Laplace operator $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$. The space H_k has dimension $h_k = \binom{n+k-1}{n-1} - \binom{n+k-3}{n-1}$.

Since in the decomposition all irreducible subspaces have multiplicity 1, i.e we have $H_k = H_{k,1}$ in the notation of Theorem 3.5, Bochner's characterization of the $\mathrm{O}(\mathbb{R}^n)$-invariant, positive kernels again only involves nonnegative scalars. The orthogonal group acts distance transitively on pairs of points on the sphere. So $\mathrm{O}(\mathbb{R}^n)$-invariant kernels are actually functions depending only on the spherical distance $d(x, y) = \arccos x \cdot y$. It turns out that they are univariate polynomial functions in the inner products $t = x \cdot y$. Schoenberg in [80] was the first who gave this characterization.

Theorem 5.2. *Every positive, $\mathrm{O}(\mathbb{R}^n)$-invariant kernel $K \in \mathcal{C}(S^{n-1} \times S^{n-1})$ can be written as*

$$K(x, y) = \sum_{k=0}^{\infty} f_k P_k^{(\alpha,\alpha)}(x \cdot y), \quad f_0, f_1, \ldots \geq 0.$$

The right hand side converges absolutely and uniformly. Here $P_k^{(\alpha,\alpha)}(t)$ is the normalized Jacobi polynomial of degree k with parameters (α, α), where $\alpha = (n-3)/2$.

The *Jacobi polynomials* with parameters (α, β) are orthogonal polynomials for the measure $(1 - t)^{\alpha}(1 + t)^{\beta}dt$ on the interval $[-1, 1]$. We denote by $P_k^{(\alpha,\beta)}$ the *normalized Jacobi polynomial* of degree k with normalization $P_k^{(\alpha,\beta)}(1) = 1$. The first few Jacobi polynomials are

$$P_0^{(\alpha,\alpha)}(t) = 1,$$
$$P_1^{(\alpha,\alpha)}(t) = t,$$
$$P_2^{(\alpha,\alpha)}(t) = \frac{n}{n-1}t^2 - \frac{1}{n-1}.$$

For more information on Jacobi polynomials we refer to the books [1] by Abramowitz, Stegun, [87] by Szegö, [2] by Andrews, Askey and Roy, and [62] by Lebedev.

5.2. Linear programming bound for the unit sphere.

Before we give a proof of these two theorems, we demonstrate how to apply them to calculate the theta function for distance graphs on the unit sphere. Again the original semidefinite program degenerates to a linear program because $\mathcal{C}(S^{n-1})$ decomposes multiplicity-free into $\mathrm{Aut}(S^{n-1})$-irreducible subspaces. This linear program was studied by Delsarte, Goethals and Seidel in [33] and by Kabatiansky and Levenshtein in [49], but see also Sloane [29, Chapter 9].

A case for which the calculation of the theta function is particularly simple is the graph $G(S^{n-1}, (0, \pi/2))$. Every vertex is adjacent to a pointed half-sphere. Despite its simplicity this case already carries all the features of the more complicated kissing number cases $G(S^7, (0, \pi/3))$ and $G(S^{23}, (0, \pi/3))$.

We show that the vertices of a regular cross polytope (the $2n$ points $\pm e_i$, with $i = 1, \ldots, n$) form an optimal spherical code. We say that N points on the unit sphere S^{n-1} form an *optimal spherical code* if they maximize the minimal distance among all N-point configuration on S^{n-1}.

Using Theorem 5.2 the theta function $\vartheta'(G(S^{n-1}, (0, \pi/2)))$ simplifies to

$$\inf \left\{ 1 + \sum_{k=0}^{\infty} f_k P_k^{(\alpha,\alpha)}(0) \ : \ f_0, f_1, \ldots \geq 0, \right.$$
$$\left. \sum_{k=0}^{\infty} f_k P_k^{(\alpha,\alpha)}(\cos\theta) \leq -1 \text{ if } \theta \in [\pi/2, \pi] \right\}.$$

Consider a polynomial $F(t) = \sum_{k=0}^{d} f_k P_k^{(\alpha,\alpha)}(t)$ with nonnegative coefficients f_k. If $F(t) \leq -1$ for all $t \in [-1, 0]$, then

$$F(0) + 1 \geq \vartheta'(G(S^{n-1}, (0, \pi/2))) \geq \alpha(G(S^{n-1}, (0, \pi/2))).$$

We give an explicit F which proves the sharp bound of $2n$. Generally, if a polynomial proves a sharp bound, then all inequalities in Theorem 2.7 have to be equalities. So we have $F(t) = -1$ for all t which occur as inner product.

We make the Ansatz

$$F(t) = f_0 + f_1 t + f_2 \left(\frac{n}{n-1} t^2 - \frac{1}{n-1} \right)$$

and in order that the bound is sharp we have to have

$$
\begin{aligned}
1 + F(1) &= 2n, \\
F(0) &= -1, \\
F(-1) &= -1.
\end{aligned}
$$

So we consider the system of linear equations

$$
\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -\frac{1}{n-1} \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 2n-1 \\ -1 \\ -1 \end{pmatrix},
$$

which happens to have the nonnegative solution $f_0 = 0$, $f_1 = n$, $f_2 = n - 1$.

Exercise 5.3. *Why does this show that the vertices of a regular cross polytope form an optimal spherical code?*

Exercise 5.4. *Show that $\tau_7 = \vartheta'(G(S^7, (0, \pi/3))) = 240$.*

Exercise 5.5. *Show that $\vartheta'(G(S^{24}, (0, \pi/3))) = 196560$.*

Exercise 5.6. *Show that the vertices of the icosahedron form an optimal spherical code.*

For more on optimal spherical codes see [13], [63], [31], [27], [11], [7]. Recently, Wang [94] gave a list of conjectural optimal spherical codes up to dimension 8 and up to 27 points. It is a challenge to find proofs of their optimality.

5.3. **Harmonic polynomials.** To establish the specialization of the Peter-Weyl theorem for the case of the unit sphere we consider harmonic polynomials.

For the vector space of complex polynomial functions $f : \mathbb{C}^n \to \mathbb{C}$ which are homogeneous and are of degree d we write $\mathrm{Pol}_{=d}(\mathbb{C}^n)$. Each $f \in \mathrm{Pol}_{=d}(\mathbb{C}^n)$ satisfies the equation $f(\alpha x) = \alpha^d f(x)$ for all $\alpha \in \mathbb{C}$ and $x \in \mathbb{C}^n$. The dimension of $\mathrm{Pol}_{=d}(\mathbb{C}^n)$ is $\binom{n+d-1}{n-1}$. We define the polynomial $\omega \in \mathrm{Pol}_{=2}(\mathbb{C}^n)$ by $\omega(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$.

Furthermore, we introduce two differential operators: The *Nabla operator*

$$\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)^t$$

and the *Laplace operator*

$$\Delta = \nabla^t \nabla = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2} = \omega(\nabla),$$

which maps $\mathrm{Pol}_{=d}(\mathbb{C}^n)$ to $\mathrm{Pol}_{=d-2}(\mathbb{C}^n)$. It will be convenient to have the following inner product on $\mathrm{Pol}_{=d}(\mathbb{C}^n)$:

$$\langle f, g \rangle = \frac{1}{d!} f(\nabla) \overline{g},$$

for which the basis of monomials is an orthogonal basis with

$$\langle x_1^{m_1} \cdots x_n^{m_n}, x_1^{m_1} \cdots x_n^{m_n} \rangle = \frac{m_1! \cdots m_n!}{d!}.$$

This inner product might look mysterious, but it turns out to be very natural if one considers polynomials as symmetric tensors as we point out in the appendix.

It is a simple and helpful fact that multiplication with ω is the adjoint to the Laplace operator:

Lemma 5.7. *For all $f \in \mathrm{Pol}_{=d-2}(\mathbb{C}^n)$ and for all $g \in \mathrm{Pol}_{=d}(\mathbb{C}^n)$ we have*

$$\langle \omega f, g \rangle = \langle f, \Delta g \rangle.$$

Proof. This follows immediately from the obvious fact that the evaluation map $f \mapsto f(\nabla)$ is an algebra isomorphism between the algebra of polynomials $\mathbb{C}[x_1, \ldots, x_n]$ and the algebra of differential operators $\mathbb{C}[\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}]$. In particular, this implies $(\omega f)(\nabla) = f(\nabla) \omega(\nabla)$, and hence $\langle \omega f, g \rangle = \frac{1}{d!} f(\nabla) \overline{(\Delta g)} = \langle f, \Delta g \rangle$. □

The kernel of the Laplace operator is the space of *harmonic polynomials* which we denote by

$$\mathrm{Harm}_d = \{f \in \mathrm{Pol}_{=d}(\mathbb{C}^n) : \Delta f = 0\}.$$

Theorem 5.8. *The space* $\mathrm{Pol}_{=d}(\mathbb{C}^n)$ *decomposes orthogonally into harmonic polynomials as follows:*

$$\mathrm{Pol}_{=d}(\mathbb{C}^n) = \mathrm{Harm}_d \perp \omega\,\mathrm{Harm}_{d-2} \perp \omega^2\,\mathrm{Harm}_{d-4} \perp \dots .$$

The dimension of Harm_k, *which we denote by* h_k, *is equal to* $\binom{n+k-1}{n-1} - \binom{n+k-3}{n-1}$.

Proof. We first show that Δ maps $\mathrm{Pol}_{=d}(\mathbb{C}^n)$ onto $\mathrm{Pol}_{=d-2}(\mathbb{C}^n)$: By Lemma 5.7 we have for $f \in (\Delta(\mathrm{Pol}_{=d-2}(\mathbb{C}^n)))^{\perp}$ and for $g \in \mathrm{Pol}_{=d}(\mathbb{C}^n)$

$$\langle \omega f, g \rangle = \langle f, \Delta g \rangle = 0,$$

hence ωf, and so f, must be 0. The dimension formula for linear maps implies the desired formula for h_k.

With similar arguments one shows that

$$\mathrm{Pol}_{=d}(\mathbb{C}^n) = \mathrm{Harm}_d \perp \omega\,\mathrm{Pol}_{=d-2}(\mathbb{C}^n).$$

Before we can inductively derive the orthogonal decomposition as stated in the theorem we only have to verify that $\omega\,\mathrm{Harm}_{d-2}$ is orthogonal to $\omega^2\,\mathrm{Pol}_{=d-4}(\mathbb{C}^n)$. The following identity will be helpful: for all $p \in \mathrm{Pol}_{=d}(\mathbb{C}^n)$, we have $\Delta \omega p = 2(n+2d)p + \omega \Delta p$ (in the calculation, notice that $\sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i} p = dp$). Then,

$$\begin{aligned}
\langle \omega f, \omega^2 g \rangle &= \langle f, \Delta \omega^2 g \rangle \\
&= \langle f, 2(n+2(d-2))\omega g + \omega \Delta \omega g \rangle \\
&= \langle f, \omega(2(n+2(d-2))g + \Delta \omega g) \rangle \\
&= 0,
\end{aligned}$$

for all $f \in \mathrm{Harm}_{d-2}$ and $g \in \mathrm{Pol}_{=d-4}(\mathbb{C}^n)$. $\square$

By $\mathrm{Pol}_{\leq d}(S^{n-1})$ we denote the vector space of complex-valued polynomial functions on S^{n-1} which are of degree at most d. This is a subspace of $\mathcal{C}(S^{n-1})$ so it is equipped with the inner product

$$(f, g) = \int_{S^{n-1}} f(x)\overline{g(x)}d\omega_n(x).$$

Recall that the measures on S^{n-1} and S^{n-2} are related by

$$d\omega_n(x) = (1 - x_1^2)^{(n-3)/2}dx_1 d\omega_{n-1}((x_2^2 + \dots + x_n^2)^{-1/2}(x_2, \dots, x_n)).$$

The orthogonal group acts on the space of polynomial functions by $uf(x) = f(u^{-1}x)$. The space of harmonic polynomials Harm_d is invariant under this action. Hence, we have for all $f \in \mathrm{Harm}_d$ and all $u \in \mathrm{O}(\mathbb{R}^n)$ that $uf \in \mathrm{Harm}_d$. This follows from the fact that $\Delta(uf) = u(\Delta f)$. Another more conceptual way to see that Harm_d is an $\mathrm{O}(\mathbb{R}^n)$-invariant space uses the fact that the inner product $\langle \cdot, \cdot \rangle$ is invariant under the action of $\mathrm{O}(\mathbb{R}^n)$. We will give a proof of this in the appendix where the connection between symmetric tensors and polynomials will also make the inner product $\langle \cdot, \cdot \rangle$ more transparent. It is obvious that the summand $\omega\,\mathrm{Pol}_{=d-2}(\mathbb{C}^n)$ in the decomposition (5.3) is an $\mathrm{O}(\mathbb{R}^n)$-invariant space. Then, its orthogonal complement, Harm_d, is an $\mathrm{O}(\mathbb{R}^n)$-invariant space, too.

The space of *spherical harmonics*

$$H_k = \{f_{|S^{n-1}} : f \in \mathrm{Harm}_k\}$$

is also an $\mathrm{O}(\mathbb{R}^n)$-invariant space. The spaces Harm_k and H_k are equivalent as $\mathrm{O}(\mathbb{R}^n)$-invariant spaces: the restriction map $\varphi(f) = f_{|S^{n-1}}$ is a bijective linear map

$\varphi : \mathrm{Harm}_k \to H_k$ with $\varphi(uf) = u\varphi(f)$ for all $f \in \mathrm{Harm}_k$ and all $u \in \mathrm{O}(\mathbb{R}^n)$. Note that the inverse map of φ is given by $\varphi^{-1}(f)(x) = \|x\|^k f\left(\frac{x}{\|x\|}\right)$.

The following construction will turn out to be important: For an orthonormal system $e_1, \ldots, e_{h_k}$ of H_k we define the function

$$z_k : S^{n-1} \times S^{n-1} \to \mathbb{C} \quad \text{by} \quad z_k(x,y) = \sum_{i=1}^{h_k} e_i(x)\overline{e_i}(y).$$

Proposition 5.9. *The following properties hold for the function z_k:*
  (a) z_k does not dependent on the choice of the orthonormal system of H_k.
  (b) For all $u \in \mathrm{O}(\mathbb{R}^n)$ and for all $x, y \in S^{n-1}$ we have $z_k(ux, uy) = z_k(x, y)$.

Proof. (a) Let $e'_1, \ldots, e'_{h_k}$ be a second orthonormal system of H_k. Let u be the orthogonal matrix $u = (u_{ij}) \in \mathrm{O}(\mathbb{R}^{h_k})$ with $ue_i = e'_i$, $i = 1, \ldots, h_k$. Then,

$$
\begin{aligned}
\sum_{i=1}^{h_k} e'_i(x)\overline{e'_i(y)} &= \sum_{i=1}^{h_k} (ue_i)(x)\overline{(ue_i)(y)} \\
&= \sum_{i=1}^{h_k}\sum_{j=1}^{h_k} u_{ji}e_j(x)u_{ji}\overline{e_j(y)} \\
&= \sum_{j=1}^{h_k}\left(\sum_{i=1}^{h_k} u_{ji}u_{ji}\right)e_j(x)\overline{e_j(y)} \\
&= \sum_{j=1}^{h_k} e_j(x)\overline{e_j(y)}.
\end{aligned}
$$

(b) Since the inner product on $\mathrm{Pol}_{\leq d}(S^{n-1})$ is invariant under $\mathrm{O}(\mathbb{R}^n)$, it follows that $u^{-1}e_i$, $i = 1, \ldots, h_k$ is an orthonormal system of H_k. Now (b) is a consequence of (a). $\qquad\square$

As in the case of Boolean harmonics we make use of the concept of zonal spherical functions to prove the irreducibility of H_k. A function $f : S^{n-1} \to \mathbb{C}$ is called a *zonal spherical function with pole* $e \in S^{n-1}$ if $f = uf$ for all $u \in \mathrm{Stab}(\mathrm{O}(\mathbb{R}^n), e) = \{u \in \mathrm{O}(\mathbb{R}^n) : ue = e\}$. Property (b) of Proposition 5.9 implies that $x \mapsto z_k(e, x)$ is a zonal spherical function with pole e. In direct analogy to Lemma 4.4 we have:

Lemma 5.10. *Let $U \subseteq \mathrm{Pol}_{\leq d}(S^{n-1})$ be an $\mathrm{O}(\mathbb{R}^n)$-invariant space and let $e \in S^{n-1}$ be a point on the sphere. If the dimension of the space of zonal spherical functions with pole e which lie in U is exactly one, then U is an $\mathrm{O}(\mathbb{R}^n)$-irreducible space.*

Now the specialization of the Peter-Weyl theorem given in Theorem 5.1 follows form the following theorem.

Theorem 5.11. *The space of complex-valued polynomials restricted to the sphere which are of degree at most d decomposes orthogonally into the spaces of spherical harmonics:*

$$\mathrm{Pol}_{\leq d}(S^{n-1}) = H_0 \perp H_1 \perp \ldots \perp H_d.$$

The H_k are $\mathrm{O}(\mathbb{R}^n)$-irreducible spaces which are pairwise inequivalent.

Proof. Using Lemma 5.10 we show that the H_k's are irreducible. Let $f \in H_k$ be a zonal spherical function with pole $e = (1, 0, \ldots, 0)^t$. We consider f as a polynomial in Harm_k. Then we can write

$$f(x_1, \ldots, x_n) = \sum_{i=0}^{k} x_1^{k-i} p_i(x_2, \ldots, x_n),$$

where p_i is a homogeneous polynomial of degree i which is invariant under the action of the stabilizer group $\mathrm{Stab}(\mathrm{O}(\mathbb{R}^n), e)$. That means that p_i is a radial function, a function which only depends on the norm of $(x_2, \ldots, x_n)$. Because p_i is homogeneous of degree i we have $p_i(x_2, \ldots, x_n) = c_i(x_2^2 + \cdots + x_n^2)^{i/2}$ for some constant c_i. Hence, $p_i = 0$ whenever i is an odd integer. So we can write

$$f(x_1, \ldots, x_n) = \sum_{i=0}^{k/2} c_i x_1^{k-2i} (x_2^2 + \cdots + x_n^2)^i.$$

Since f is a harmonic polynomial we have (check this)

$$0 = \Delta f = \sum_{i=1}^{k/2} (\alpha_i c_i + \beta_i c_{i-1}) x_1^{k-2i} (x_2^2 + \cdots + x_n^2)^{i-1},$$

where $\alpha_i = 2i(n + 2i - 3)$, $\beta_i = (k - 2i + 1)(k - 2i + 2)$, and so $c_i = (-1)^i c_0 \frac{\beta_1 \cdots \beta_i}{\alpha_1 \cdots \alpha_i}$, for $i = 1, \ldots, k/2$, which shows that f is determined by c_0 only. Thus, the space of zonal spherical functions is one-dimensional.

If $n > 2$, then, by looking at the dimension h_k, it follows immediately that H_k are pairwise inequivalent and also the orthogonality relation between them. The case $n = 2$ we leave as an exercise (see below). $\qquad\square$

Exercise 5.12. *Make the link between the spherical harmonics for the unit circle S^1 and the classical Fourier series of 2π-periodic functions. Show how the the functions $\cos kx$, $\sin kx$, with $k = 0, 1, 2, \ldots$, appear as orthonormal system for the spaces H_k. Show that H_k and $H_{k'}$ with $k \neq k'$ are not equivalent.*

5.4. **Addition formula.** To establish Bochner's characterization for the sphere we still have to calculate $Z_k(x, y)$. For this we do not need to write down an orthonormal system of H_k explicitly. Studying the orthogonality relations and their interplay with zonal spherical function suffices here.

Theorem 5.13. *Let $e_1, \ldots, e_{h_k}$ be an orthonormal system of H_k. For $x, y \in S^{n-1}$ we have the addition formula*

$$P_k^{((n-3)/2, (n-3)/2)}(x \cdot y) = \alpha_k \sum_{l=1}^{h_k} e_l(x) \overline{e_l(y)},$$

for a positive constant α_k.

Proof. Let z_k and z_k' be zonal spherical functions for H_k and $H_{k'}$, with $k \neq k'$, both with pole $e = (1, 0, \ldots, 0)$. By the orthogonality relation between nonequivalent $\mathrm{O}(\mathbb{R}^n)$-irreducible subspaces it follows

$$(z_k, z_{k'}) = \int_{S^{n-1}} z_k(x) \overline{z_{k'}(x)} d\omega(x) = 0$$

The functions z_k and $z_{k'}$ are invariant under the stabilizer group $\mathrm{Stab}(\mathrm{O}(\mathbb{R}^n), e)$. So they only depend on the first coordinate. One can use (5.3) to express the integral above as one involving only x_1:

$$\int_{-1}^{1} z_k(x_1, \ldots, x_n)\overline{z_{k'}(x_1, \ldots, x_n)}(1 - x_1^2)^{(n-3)/2}dx_1 = 0.$$

Hence, the zonal spherical functions z_k are multiples of Jacobi polynomials:

$$z_k(x_1, \ldots, x_n) = \alpha_k P_k^{((n-3/2),(n-3)/2)}(x_1),$$

and hence

$$\sum_{l=1}^{h_k} e_l(x)\overline{e_l(y)} = \alpha_k P_k^{((n-3/2),(n-3)/2)}(x \cdot y).$$

The constant α_k is positive because $P_k^{((n-3/2),(n-3)/2)}(1) = 1$. $\square$

Appendix A. Symmetric tensors and polynomials

In this section we write V for $\mathbb{C}^n$. The symmetric group S_d of permutations $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$ acts on $V^{\otimes d}$ by permuting positions:

$$\sigma(v_1 \otimes \cdots \otimes v_d) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(d)},$$

with $\sigma \in \mathrm{S}_d$. The subspace of *symmetric d-tensors over V* is

$$S_d(V^{\otimes d}) = \{u \in V^{\otimes d} : \sigma u = u \text{ for all } \sigma \in \mathrm{S}_d\}.$$

The inner product on $V^{\otimes d}$ is given by

$$\langle v_1 \otimes \cdots \otimes v_d, w_1 \otimes \cdots \otimes w_d \rangle = \prod_{i=1}^{d} v_i \cdot \overline{w_i}.$$

Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{C}^n$. Define the linear map

$$\pi : V^{\otimes d} \to \mathrm{Pol}_{=d}(\mathbb{C}^n), \quad \text{by } e_{i_1} \otimes \cdots \otimes e_{i_d} \mapsto x_{i_1} \cdots x_{i_d}.$$

For an n-tupel $(m_1, \ldots, m_n)$ of nonnegative integers with $m_1 + \cdots + m_n = d$ define the *symmetric monomial*

$$s_{(m_1, \ldots, m_n)} = \sum_{(i_1, \ldots i_d)} e_{i_1} \otimes \cdots \otimes e_{i_d},$$

where we sum over all $(i_1, \ldots, i_d)$ so that the number $j \in \{1, \ldots, n\}$ occurs in $(i_1, \ldots, i_d)$ with multiplicity m_j, e.g.

$$s_{(1,2)} = e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1.$$

The symmetric monomials form a basis of $S_d(V)$. We have

$$\pi\big(s_{(m_1, \ldots, m_n)}\big) = \frac{d!}{m_1! \cdots m_n!} x_1^{m_1} \cdots x_n^{m_n}.$$

Hence, the map π restricted to symmetric d-tensors over V is an isometry due to the specific choice of the inner product $\langle \cdot, \cdot \rangle$ in $\mathrm{Pol}_{=d}(\mathbb{C}^n)$.

The orthogonal group $\mathrm{O}(\mathbb{R}^n)$ acts on $V^{\otimes d}$ by $u(v_1 \otimes \cdots \otimes v_d) = uv_1 \otimes \cdots \otimes uv_d$, with $u \in \mathrm{O}(\mathbb{R}^n)$. Let us verify that π commutes with the action of $\mathrm{O}(\mathbb{R}^n)$, i.e.

$\pi(us) = a\pi(s)$ for all $s \in S_d(V)$ and $u = (u_{i,j}) \in O(\mathbb{R}^n)$. It suffices to consider the basis $s_{(m_1,\ldots,m_n)}$. We have

$$\pi\big(us_{(m_1,\ldots,m_n)}\big) = \pi\Big(\sum_{(i_1,\ldots,i_d)} ue_{i_1} \otimes \cdots \otimes ue_{i_d} \Big)$$

$$= \pi\Big(\sum_{(i_1,\ldots,i_d)} \sum_{j_1,\ldots,j_d=1}^{n} u_{j_1,i_1} \cdots u_{j_d,i_d} e_{j_1} \otimes \cdots \otimes e_{j_d} \Big)$$

$$= \sum_{(i_1,\ldots,i_d)} \Big(\sum_{j=1}^{n} u_{j,i_1} x_j\Big) \cdots \Big(\sum_{j=1}^{n} u_{j,i_d} x_j\Big)$$

$$= \frac{d!}{m_1! \cdots m_n!}\Big(\sum_{j=1}^{n} u_{j,1} x_j\Big)^{m_1} \cdots \Big(\sum_{j=1}^{n} u_{j,n} x_j\Big)^{m_n}$$

$$= u\pi(s_{(m_1,\ldots,m_n)}).$$

We finish by showing that the inner product on $\mathrm{Pol}_{=d}(\mathbb{C}^n)$ is invariant under the action of the orthogonal group. For this note that the inner product on $V^{\otimes d}$ is invariant under this group action. From the previous considerations we have for all $u \in O(\mathbb{R}^n)$ the chain of equalities:

$$\langle uf, ug \rangle = \langle \pi^{-1}(uf), \pi^{-1}(ug) \rangle$$

$$= \langle u\pi^{-1}(f), u\pi^{-1}(g) \rangle$$

$$= \langle \pi^{-1}(f), \pi^{-1}(g) \rangle$$

$$= \langle f, g \rangle.$$

Appendix B. Further reading

In the last years many results were obtained for semidefinite programs which are symmetric. This was done for a variety of problems and applications:

— interior point algorithms (Kanno, Ohsaki, Murota, Katoh [51], de Klerk, Pasechnik [53], Murota, Kanno, Kojima, Kojima [69]),

— polynomial optimization (Gatermann and Parrilo [37] and Jansson, Lasserre, Riener, Theobald [48], Bosse [17], Laurent [61]),

— truss topology optimization (Bai, de Klerk, Pasechnik, Sotirov [10]),

— quadratic assignment problem (de Klerk, Sotirov [57]),

— fast mixing Markov chains on graphs (Boyd, Diaconis, Xiao [19], Boyd, Diaconis, Parrilo, Xiao [18]),

— graph coloring (Gvozdenović, Laurent [43], [44], Gvozdenović [42]), — crossing numbers for complete binary graphs (de Klerk, Pasechnik, Schrijver [55]),

— travelling salesman problem (de Klerk, Pasechnik, Sotirov [56]),

— coding theory (Schrijver [82], Gijswijt, Schrijver, Tanaka [39], de Klerk, Pasechnik [54], de Klerk, Newman, Pasechnik, Sotirov [52], Laurent [60], Bachoc [4], Vallentin [91], Bachoc, Vallentin [8]),

— low distortion embeddings (Vallentin [90])

— geometry (Bachoc, Vallentin [5], [6], [7], Bachoc, Nebe, de Oliveira Filho, Vallentin [9]).

References

[1] M. Abramowitz, I.A. Stegun, *Handbook of mathematical functions*, National Bureau of Standards Appl. Series **55**, U.S. Dept. Commerce, 1972.

[2] G.E. Andrews, R. Askey, R. Roy, *Special functions*, Cambridge University Press, 1999.

[3] L. Babai, L. Rónyai, *Computing irreducible representations of finite groups*, Math. Comp. **55** (1990), 705–722.

[4] C. Bachoc, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52** (2006), 2111–2125.

[5] C. Bachoc, F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), 909–924.

[6] C. Bachoc, F. Vallentin, *Semidefinite programming, multivariate orthogonal polynomials, and codes in spherical caps*, preprint, arXiv:math/0610856v2 [math.MG], October 2007, to appear in Europ. J. Combin.

[7] C. Bachoc, F. Vallentin, *Optimality and uniqueness of the (4,10,1/6) spherical code*, preprint, arXiv:math/0708.3947v2 [math.MG], May 2008, to appear in J. Combin. Theory Ser. A.

[8] C. Bachoc, F. Vallentin, *More semidefinite programming bounds (extended abstract)*, p. 129–132 in Proceedings "DMHF 2007: COE Conference on the Development of Dynamic Mathematics with High Functionality", October 2007, Fukuoka, Japan.

[9] C. Bachoc, G. Nebe, F.M. de Oliveira Filho, F. Vallentin, *Lower bounds for measurable chromatic numbers*, preprint, arXiv:math/0801.1059v1 [math.CO], January 2008, to appear in Geom. Funct. Anal.

[10] Y. Bai, E. de Klerk, D.V. Pasechnik, R. Sotirov, *Exploiting group symmetry in truss topology optimization*, preprint, Optimization Online, April 2007.

[11] B. Ballinger, G. Blekherman, H. Cohn, N. Giansiracusa, E. Kelly, A. Schürmann, *Experimental study of energy-minimizing point configurations on spheres*, preprint, arXiv:math/0611451v2 [math.MG], November 2006, to appear in Experiment. Math.

[12] E. Bannai, T. Ito, *Algebraic combinatorics I: Association schemes*, Benjamin/Cummings, 1984.

[13] E. Bannai, N.J.A. Sloane, *Uniqueness of certain spherical codes*, Canad J. Math. **33** (1981), 437–449.

[14] M. Berger, *A Panoramic View of Riemannian Geometry*, Springer Verlag, 2003.

[15] S. Bochner, *Hilbert distances and positive definite functions*, Ann. of Math. **42** (1941), 647–656.

[16] A. Borel, *Hermann Weyl and Lie Groups*, pp. 53–82 in Hermann Weyl, 1885–1985, Centenary lectures at the ETH Zürich (K. Chandrasekharan ed.), Springer Verlag, 1986.

[17] H. Bosse, *Symmetric, positive polynomials, which are not sums of squares*, preprint, Series: CWI. Probability, Networks and Algorithms [PNA], Nr. E0706, 2007.

[18] S. Boyd, P. Diaconis, P.A. Parrilo, L. Xiao, *Symmetry analysis of reversible Markov chains*, Internet Mathematics **2** (2005).

[19] S. Boyd, P. Diaconis, L. Xiao, *Fastest mixing Markov chain on a graph*, SIAM Rev. **46** (2004), 667–689.

[20] T. Bröcker, T. tom Dieck, *Representations of compact Lie groups*, Springer Verlag, 1985.

[21] A.E. Brouwer, *Table of general binary codes*,
http://www.win.tue.nl/~aeb/codes/binary-1.html

[22] A.E. Brouwer, A.M. Cohen, A. Neumaier, *Distance-regular graphs*, Springer Verlag, 1989.

[23] R. Carter, G. Siegel, I. Macdonald, *Lectures on Lie groups and Lie algebras*, Cambridge University Press, 1995.

[24] B. Casselman, *The difficulties of kissing in three dimensions*, Notices Amer. Math. Soc. **51** (2004), 884–885.

[25] H. Cohn, N.D. Elkies, *New upper bounds on sphere packings I*, Ann. of Math. **157** (2003), 689–714.

[26] H. Cohn, A. Kumar, *Universally optimal distributions of points on spheres*, J. Amer. Math. Soc. **20** (2007), 99–148.

[27] H. Cohn, A. Kumar, *Uniqueness of the (22, 891, 1/4) spherical code*, New York J. Math. **13** (2007), 147–157.

[28] R.R. Coifman, G. Weiss, *Representations of compact groups and spherical harmonics*, L'Einseignement math. **14** (1968), 121–173.

[29] J.H. Conway, N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer Verlag, 1988.

[30] R. Courant, D. Hilbert, *Methods of mathematical physics*, Interscience Publishers, 1953.

[31] H. Cuypers, *A note on the tight spherical 7-designs in $\mathbb{R}^{23}$ and 5-designs in $\mathbb{R}^7$*, Des. Codes Cryptogr. **34** (2005), 333–337.

[32] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.

[33] P. Delsarte, J.M. Goethals, J.J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.

[34] P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series **11**, Institute of Mathematical Statistics, Hayward, CA, 1988.

[35] C.F. Dunkl, *A Krawtchouk polynomial addition theorem and wreath product of symmetric groups*, Indiana Univ. Math. J. **25** (1976), 335–358.

[36] U. Feige, *Randomized graph products, chromatic numbers, and the Lovász ϑ-function*, Combinatorica **17** (1997), 79–90.

[37] K. Gatermann, P.A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Alg. **192** (2004), 95–128.

[38] D.C. Gijswijt, *Matrix algebras and semidefinite programming techniques for codes*, PhD Thesis, University of Amsterdam, 2005.

[39] D.C. Gijswijt, A. Schrijver, H. Tanaka, *New upper bounds for nonbinary codes*, J. Combin. Theory Ser. A **13** (2006), 1717–1731.

[40] R. Goodman, N.R. Wallach, *Representations and invariants of the classical groups*, Cambridge University Press, 1998.

[41] K.I. Gross, *On the evolution of noncommutative harmonic analysis*, Amer. Math. Monthly, **85** (1978), 525–548.

[42] N. Gvozdenović, *Approximating the stability number and the chromatic number of a graph via semidefinite programming*, PhD Thesis, University of Amsterdam, 2008.

[43] N. Gvozdenović, M. Laurent, *The operator Ψ for the chromatic number of a graph*, SIAM J. Optim. **19** (2008), 572–591.

[44] N. Gvozdenović, M. Laurent, *Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization*, SIAM J. Optim. **19** (2008), 592-615.

[45] A. Haar, *Der Massbegriff in der Theorie der kontinuierlichen Gruppen*, Ann. Math. **34** (1933), 147–169.

[46] J. Håstad, *Clique is hard to approximate within $n^{1-\epsilon}$*, Acta Mathematica, **182**, (1999) 105–142.

[47] S. Helgason, *Groups and geometric analysis*, Academic Publishers, 1984.

[48] L. Jansson, J.B. Lasserre, C. Riener, T. Theobald, *Exploiting symmetries in SDP-relaxations for polynomial optimization*, preprint, Optimization Online, September 2006.

[49] G.A. Kabatiansky, V.I. Levenshtein, *Bounds for packings on a sphere and in space*, Problems of Information Transmission **14** (1978), 1–17.

[50] G. Kalai, M. Safra, *Threshold phenomena and influence* p. 25–60 in Computational Complexity and Statistical Physics, (A.G. Percus, G. Istrate, C. Moore, eds.) Oxford University Press, 2006.

[51] Y. Kanno, M. Ohsaki, K. Murota, N. Katoh, *Group symmetry in interior-point methods for semidefinite program*, Optimization and Engineering **2** (2001) 293–320.

[52] E. de Klerk, M.W. Newman, D.V. Pasechnik, R. Sotirov, *On the Lovasz θ-number of almost regular graphs with application to Erdős-Rényi graphs*, Discussion paper from Tilburg University, Center for economic research, 2006.

[53] E. de Klerk, D.V. Pasechnik, *Solving SDP's in non-commutative algebras part I: the dual-scaling algorithm*, Discussion paper from Tilburg University, Center for economic research, 2005.

[54] E. de Klerk, D.V. Pasechnik, *A note on the stability number of an orthogonality graph*, European J. Combin. **28** (2007) 1971–1979.

[55] E. de Klerk, D.V. Pasechnik, A. Schrijver, *Reduction of symmetric semidefinite programs using the regular $*$-representation*, Math. Prog. **109** (2007) 613–624.

[56] E. de Klerk, D.V. Pasechnik, R. Sotirov, *On Semidefinite Programming Relaxations of the Travelling Salesman Problem*, Discussion paper from Tilburg University, Center for economic research, 2007.

[57] E. de Klerk, R. Sotirov, *Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem*, preprint, Optimization Online, June 2007.

[58] D.E. Knuth, *The sandwich theorem*, Electron. J. Combin. **1** (1994), 48 pp.

[59] J. Köbler, U. Schöning, J. Torán, *The graph isomorphism problem: its structural complexity*, Birkhäuser, 1993.

[60] M. Laurent, *Strengthened semidefinite programming bounds for codes*, Math. Prog. **109** (2007), 239–261.

[61] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, preprint, February 2008, to appear in IMA volume Emerging Applications of Algebraic Geometry (M. Putinar, S. Sullivant, eds.)

[62] N.N. Lebedev, *Special functions and their applications*, Dover Publication Inc., 1972.

[63] V.I. Levenshtein, *On bounds for packing in n-dimensional Euclidean space*, Soviet Math. Dokl. **20** (1979), 417–421.

[64] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inf. Th. **25** (1979), 1–7.

[65] L. Lovász, *Discrete and continuous: Two sides of the same?*, Geom. Funct. Anal., Special volume — GAFA2000, 359–382.

[66] F.J. MacWilliams, N.J.A. Sloane, *The theory of error-correcting codes*, North Holland, 1988.

[67] B.D. McKay, `nauty` *User's Guide (Version 2.4)*, 2006, http://cs.anu.edu.au/people/bdm/nauty/nug-2.4b3.pdf .

[68] C. Müller, *Spherical Harmonics*, Springer Verlag, Heidelberg, 1966.

[69] K. Murota, Y. Kanno, M. Kojima and S. Kojima, *A Numerical Algorithm for Block-Diagonal Decomposition of Matrix *-Algebras*, METR 2007-52, Department of Mathematical Informatics, University of Tokyo, Japan, September 2007.

[70] O.R. Musin, *The kissing number in four dimensions*, Ann. of Math. **168** (2008), 1–32.

[71] O.R. Musin, *Multivariate positive definite functions on spheres*, preprint, arXiv:math/07010.83805v1 [math.MG], October 2007.

[72] F. Peter, H. Weyl, *Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe*, Math. Ann. **97** (1927), 737–755.

[73] F. Pfender, G.M. Ziegler, *Kissing numbers, sphere packings and some unexpected proofs*, Notices Amer. Math. Soc. **51** (2004), 873–883.

[74] M. Reed, B. Simon, *Methods of modern mathematical physics I: Functional analysis*, Academic Press, 1972.

[75] F. Riesz, B. Sz.-Nagy, *Functional analysis*, Dover Publication Inc., 1990.

[76] R.J. McEliece, E.R. Rodemich, H. C. Rumsey Jr., *The Lovász bound and some generalizations*, J. Combin. Inf. System Sci. **3** (1978) 134–152.

[77] 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.

[78] F.M. de Oliveira Filho, F. Vallentin, *Fourier analysis, linear programming, and distance avoiding sets in $\mathbb{R}^n$*, preprint, arXiv:0808.1822v1 [math.CO], August 2008.

[79] B.E. Sagan, *The symmetric group, representations, combinatorial algorithms, and symmetric functions*, Springer Verlag, 2001.

[80] I.J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108.

[81] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inf. Th. **25** (1979), 425–429.

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

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

[84] J.P. Serre, *Linear representations of finite groups*, Springer Verlag, 1977.

[85] P. Slodowy, *The early development of the representation theory of simple Lie groups: A. Hurwitz, I. Schur, H. Weyl*, Jber. d. Dt. Math.-Verein. **101** (1999), 97–115.

[86] D. Stanton, *An introduction to group representations and orthogonal polynomials*, Orthogonal Polynomials: Theory and Practice, ed. P. Nevai, Proceedings of NATO-ASI, Columbus, 1990, 419–433.

[87] G. Szegö, *Orthogonal polynomials*, Volume XXIII, Colloqium Publications, American Mathematical Society, 1967.

[88] S.J. Taylor, *Introduction to measure and integration*, Cambridge University Press, 1973.

[89] A. Terras, *Fourier analysis on finite groups and applications*, Cambridge University Press, 1999.

[90] F. Vallentin, *Optimal embeddings of distance transitive graphs into Euclidean spaces*, J. Combin. Theory Ser. B **98** (2008), 95–104.

[91] F. Vallentin, *Symmetry in semidefinite programs*, preprint, arXiv:math/0706.4233v1 [math.OC], June 2007, to appear in Linear Algebra and Appl.

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

[93] E.B. Vinberg, *Linear representations of groups*, Birkhäuser, 1989.

[94] J. Wang, *Finding and investigating exact spherical codes*, preprint, arXiv:math/0805.0776v1 [math.MG], May 2008.

[95] H. Weyl, *Harmonics on manifolds*, Ann. of Math. **35** (1934), 486–499.

F. Vallentin, Centrum voor Wiskunde en Informatica (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

E-mail address: f.vallentin@cwi.nl