UNIVERSITÉ BORDEAUX 1

N^o d'ordre : 5007

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ BORDEAUX 1

Discipline: **Mathématiques Pures**

présentée et soutenue publiquement

par

Alberto PASSUELLO

le 17 décembre 2013

Titre :

Semidefinite programming in combinatorial optimization with applications to coding theory and geometry

JURY

Mme Christine BACHOC	IMB Université Bordeaux 1	Directeur de thèse
M. Dion GIJSWIJT	CWI Amsterdam, TU Delft	Examinateur
M. Didier HENRION	CNRS LAAS Toulouse	Rapporteur
M. Arnaud PÊCHER	LABRI Université Bordeaux 1	Examinateur
M. Jean-Pierre TILLICH	INRIA Rocquencourt	Examinateur
M. Gilles ZÉMOR	IMB Université Bordeaux 1	Examinateur

Institut de Mathématiques de Bordeaux
Université Bordeaux I
351, Cours de la Libération
33405 Talence cedex, France

A ma famille.

Remerciements

Beaucoup de personnes ont contribué à la réalisation de cette thèse. En essayant de n'en oublier aucune, je remonte le long des souvenirs de ces années.

Tout commença le jour où je rencontrai celle qui allait devenir, sous peu, ma directrice de thèse. Je venais de débarquer à l'université de Bordeaux afin d'entamer ma deuxième année de Master dans le cadre du projet ALGANT. Aussi attiré qu'effrayé par le titre du cours "Empilement de sphères, symétries et optimisation" que Christine allait donner cette année là, je lui demandai donc de plus amples informations. Nous nous assîmes face à une feuille blanche et, de façon gentille et directe, Christine m'expliqua en quelques lignes de crayon les liens entre sphères, réseaux et codes, me faisant ainsi entrevoir un monde de mathématiques qui m'était resté inconnu jusque-là. Un an et un mémoire de M2 plus tard, nous commençâmes cette aventure qui se termine aujourd'hui. J'ai eu de la chance d'être dirigé dans mes recherches par une personne aussi disponible et patiente, toujours prête à m'encourager, à trouver de nouvelles idées quand je n'en avais pas, à mettre de la rigueur dans mes intuitions farfelues.

Ensuite, je tiens à exprimer ma reconnaissance à Dion Gijswijt, Didier Henrion, Arnaud Pêcher, Jean-Pierre Tillich et Gilles Zémor pour avoir accepté de former le jury de cette soutenance. Je remercie également Didier Henrion et Eiichi Bannai d'avoir accepté d'être les rapporteurs de cette thèse.

Pendant ces années j'ai eu la chance de voyager et de rencontrer des chercheurs experts qui se trouvent également être des personnes très sympathiques et disponibles, dont je garde beaucoup de bons souvenirs. Il s'agit principalement de Frank Vallentin, Monique Laurent, Cordian Riener, Elisa Gorla et Joachim Rosenthal. Ici à Bordeaux j'ai eu le plaisir de travailler avec Alain Thiery, ce qui fut particulièrement enrichissant.

Comment pourrais-je ne pas penser aussi aux amis avec qui j'ai partagé le bureau 153... Damiano, qui était toujours là le matin quand j'arrivais et qui restait après moi le soir, infatigable, toujours prêt à répondre à mes questions telles que : « Damiano, excuse-moi, quelle est la commande pour copier un fichier avec le terminal ? » Johana, rayonnante de chaleur et pourtant toujours frileuse, qui m'a fait découvrir la Colombie et progresser en espagnol. Rohit et Siddarth, qui m'ont côtoyé ce dernier été, pendant la rédaction de cette thèse.

Parmi tous les doctorants et étudiants de l'IMB, qu'une place soit faite à mes compagnons de route, mes chers camarades italiens. Durant ces années, Nicola, Giovanni et Andrea ont été une présence constante et d'une aide précieuse. Nous avons parcouru le même chemin, duquel je me détournais si souvent et si volontiers, mais sans jamais les perdre de vue.

Mes pensées s'envolent aussi vers mes autres collègues et amis et particulièrement vers Pierre, Nicolas, Soline, Stephanie, Sophie, Nastaran. Qu'ils le sachent ou non, ils ont tous été très importants pour moi.

Il faut aussi remarquer que l'IMB ne serait pas ce qu'il est sans le support de tous les secrétaires, informaticiens et bibliothécaires. Souvent, sans en avoir l'intention, je m'employais à rendre leur travail plus difficile en leur confiant la tâche de remédier à mes manques d'attention. En revanche, cela m'a permis de vérifier qu'ils forment une équipe bien qualifiée et compétente.

Une moindre mais importante partie de mon doctorat m'a vu côtoyer les enseignants et le personnel administratif du département d'informatique de l'IUT Bordeaux 1, en particulier Brigitte Carles, Pierre Moreau et Michel Hickel. Qu'ils soient ici remerciés d'avoir accueilli le néophyte que j'étais en m'accordant leur confiance et cela dans une ambiance respectueuse et amicale.

Je pense également à tous mes amis de Bordeaux et d'Italie, trop nombreux pour être listés ici, dont la plupart seraient par ailleurs gênés de voir leur nom apparaître dans une thèse de maths.

Le remerciement le plus important va sans doute à ma famille, sans laquelle je n'en serais jamais arrivé là.
Mon frère, que chaque été je me réjouissais de rejoindre pour pouvoir mesurer l'ampleur de notre intersection.
Ma mère, qui me guide avec son infinie sagesse silencieuse et dont je porte l'amour où que j'aille.
Mon père, l'exemple le plus important, qui m'a appris et transmis tout ce qu'il y a de bon en moi.
Ils ont su, ces années durant, supporter mon absence sans jamais arrêter de me supporter. Merci.

Résumé

Une nouvelle borne supérieure sur le cardinal des codes de sous-espaces d'un espace vectoriel fini est établie grâce à la méthode de la programmation semidéfinie positive. Ces codes sont d'intérêt dans le cadre du codage de réseau (network coding). Ensuite, par la même méthode, l'on démontre une borne sur le cardinal des ensembles qui évitent une distance donnée dans l'espace de Johnson et qui est obtenue par une variante d'un programme de Schrijver. Les résultats numeriques permettent d'améliorer les bornes existantes sur le nombre chromatique mesurable de l'espace Euclidien. Une hiérarchie de programmes semidéfinis positifs est construite à partir de certaines matrices issues des complexes simpliciaux. Ces programmes permettent d'obtenir une borne supérieure sur le nombre d'indépendance d'un graphe. Aussi, cette hiérarchie partage certaines propriétés importantes avec d'autres hiérarchies classiques. A titre d'exemple, le problème de déterminer le nombre d'indépendance des graphes de Paley est analysé.

MOTS CLÉS : théorie des graphes, nombre d'indépendance, nombre chromatique, SDP, codes projectifs, hiérarchies.

Abstract

We apply the semidefinite programming method to obtain a new upper bound on the cardinality of codes made of subspaces of a linear vector space over a finite field. Such codes are of interest in network coding. Next, with the same method, we prove an upper bound on the cardinality of sets avoiding one distance in the Johnson space, which is essentially Schrijver semidefinite program. This bound is used to improve existing results on the measurable chromatic number of the Euclidean space. We build a new hierarchy of semidefinite programs whose optimal values give upper bounds on the independence number of a graph. This hierarchy is based on matrices arising from simplicial complexes. We show some properties that our hierarchy shares with other classical ones. As an example, we show its application to the problem of determining the independence number of Paley graphs.

Keywords: graph theory, stable number, chromatic number, SDP, projective codes, hierarchies.

Contents

List of Tables

Introduction en français

Depuis son essor, dans les années 1950, la programmation linéaire est devenue un outil fondamental pour modéliser, et souvent résoudre, beaucoup de problèmes dans le domaine de l'optimisation combinatoire. Dès lors, le sujet n'a pas cessé d'attirer un fort intérêt, grâce aussi aux progrès réalisés en algorithmique et au développement de la puissance des moyens de calcul. C'est là que la plupart des problèmes d'optimisation plongent leurs racines. Outre la programmation linéaire, citons aussi la programmation entière et la programmation quadratique. Dans ce travail nous nous intéresserons notamment à la *programmation semidéfinie positive* (SDP).

Rappelons ici une des formulations possibles pour un programme linéaire : il s'agit de trouver un vecteur réel aux entrées non négatives qui maximise une fonction linéaire, le vecteur étant soumis à des contraintes qui sont des inégalités linéaires. Dans un programme semidéfini positif, les vecteurs sont remplacés par des matrices réelles symétriques. Une matrice admissible doit alors être semidéfinie positive, c'est-à-dire que ses valeurs propres doivent être non négatives. Il s'agit d'une généralisation : en effet, si l'on impose la condition que la matrice soit diagonale, l'on se retrouve avec un programme linéaire.

Dans les années 1980, des algorithmes basés sur la méthode des points intérieurs ont vu le jour, d'abord pour les programmes linéaires ([26]) et ensuite pour les programmes semidéfinis positifs ([4], [37], [38]). Ceux-ci permettent de résoudre les programmes semidéfinis positifs d'une façon efficace et avec une très bonne approximation. Entretemps, la programmation semidéfinie positive a été appliquée à une multitude de problèmes, venant d'un vaste éventail de branches des mathématiques. Aujourd'hui encore, la programmation semidéfinie positive semble être loin d'avoir épuisé ses champs d'application, notamment en ce qui concerne l'approximation de problèmes durs à résoudre.

Cette thèse s'inscrit dans une ligne de recherche qui a peut-être son origine dans le travail de Philippe Delsarte sur les schémas d'association ([14]), où l'on trouve un programme linéaire dont la valeur optimale donne une borne supérieure sur la taille maximale d'un code avec distance minimale donnée. Quand ce programme est appliqué aux codes dans l'espace de Hamming, l'on obtient une borne qui a donné pendant longtemps les meilleurs résultats numériques et asymptotiques.

Les inégalités présentes dans le programme linéaire de Delsarte pour l'espace de Hamming H_n peuvent être obtenues en considérant la distribution des distances d'un code. La même distribution résulte en prenant les orbites des paires de points sous l'action du groupe G d'automorphismes de H_n. Les polynômes de Krawtchouk, reliés à la décomposition en composantes irréductibles sous l'action de G de l'espace de fonctions à valeurs complexes de H_n, apparaissent dans les inégalités du programme de Delsarte.

Par ailleurs, la borne de Delsarte peut s'appliquer aux codes d'autres espaces, comme l'espace de Johnson binaire ou la sphère unité de l'espace Euclidien réel. Plus généralement, elle s'applique à tous les espaces 2-points homogènes. La quête de généralisations de cette méthode a conduit à l'utilisation de la programmation semidéfinie positive. En 2005, Schrijver a établi une borne supérieure pour la taille des codes dans l'espace de Hamming en utilisant la programmation semidéfine positive. Dans beaucoup de cas, sa borne donne des résultats meilleurs que celle de Delsarte ([41]). Ensuite, la méthode de Schrijver a été étendue aux codes non-binaires ([22]) et améliorée davantage dans le cas binaire ([21]).

L'idée sous-jacente est de prendre en compte les triplets de points, au lieu des paires. Avec ceci l'on peut obtenir la borne de Schrijver par une démarche similaire à celle rappelée ci-dessus pour le programme de Delsarte. Juste, dans ce cas on considère l'action d'un sous-groupe $G' \leq G$, notamment le sous-groupe stabilisateur d'un point. L'espace H_n muni de cette action n'est pas 2-points homogène. Par conséquent, la décomposition en composantes irréductibles de l'espace de fonctions à valeurs complexes de H_n a des multiplicités plus grandes que 1 et celle-ci est la raison pour laquelle des matrices apparaissent.

Cette idée a déjà été appliquée, avec succès, à d'autres espaces de la théorie des codes, comme la sphère unité de l'espace Euclidien réel. Avec ceci, des nouveaux résultats ont été obtenus pour le problème géométrique du "kissing number" ([10]).

Un point de vue uniforme nous est donné par la théorie des graphes. En effet, un espace métrique peut être vu comme l'ensemble des sommets d'un graphe, où les arêtes correspondent aux distances interdites. Les problèmes cités ci-dessus peuvent alors être abordés en termes d'ensembles indépendants du graphe donné. Ceci nous permet d'utiliser le nombre ϑ de Lovász. En 1979, Lovász définit le nombre ϑ d'un graphe comme la valeur optimale d'un certain programme semidéfini positif associé au graphe ([34]). Ce nombre possède des propriétés remarquables : entre autres, il donne une borne supérieure sur le nombre d'indépendance du graphe. Aussi, quand une petite modification du programme, nommée ϑ', est appliquée à un certain graphe dans l'espace de Hamming, l'on retrouve la borne de Delsarte ([36], [40]).

Par ailleurs, les travaux de recherche plus récents se concentrent sur les *hiérarchies* de programmes semidéfinis positifs ([23], [32], [35]). En ce qui concerne la détermination du nombre d'indépendance α d'un graphe, le premier degré d'une telle hiérarchie donne d'habitude le nombre ϑ de Lovász et les degrés suivants donnent des bornes supérieures de plus en plus strictes pour α. C'est dans ce contexte qu'on peut voir les bornes pour les codes dans l'espace de Hamming basées sur les triplets ([41]) ou les quadruplets ([21]) de points.

Le point de départ de ce travail était l'application de ces méthodes dans le cadre du codage des réseaux (network coding). Il s'agit d'un domaine récent et dynamique de la théorie de l'information, qui capture à la fois l'intérêt des informaticiens et des mathématiciens. Koetter et Kschischang ont formulé le problème d'une communication fiable dans un tel contexte en utilisant des codes de sous-espaces. Cette formulation établit un lien remarquable avec le cadre classique des codes dans l'espace de Hamming. Néanmoins, des différences importantes existent entre ces deux cas, ce qui empêche de réitérer naïvement la même méthode.

Plus précisément, les codes définis par Koetter et Kschischang sont des familles de sous-espaces vectoriels d'un espace vectoriel fixé sur un corps fini. Ces codes sont nommés *codes projectifs* et, dans le cas où tous les sous-espaces ont la même dimension, *codes à dimension constante*. Ils sont les analogues q-aires des codes de Hamming et de Johnson. Ici, q denote le cardinal du corps de base et l'analogie q-aire consiste à remplacer (sous-)ensembles avec (sous-)espaces vectoriels et, par conséquent, le cardinal avec la dimension. Au niveau de l'action de groupe, le groupe symétrique est remplacé par le groupe général linéaire. Il est important, ici, de remarquer une première différence : l'espace de Hamming H_n est muni d'une action transitive, tandis que ce n'est pas le cas dans l'espace projectif. Nous voyons par là que la borne de la programmation linéaire de Delsarte ne peut pas s'appliquer aux codes projectifs. En revanche, nous allons établir une borne de la programmation semidéfinie positive en utilisant le nombre ϑ de Lovász. Afin de calculer cette borne dans des cas non-triviaux, il est nécessaire de symétriser le programme définissant ϑ sous l'action du groupe général linéaire. En ce qui concerne les codes à dimension constante, il y a une meilleure analogie avec les codes à poids constant, ce qui a permis de retrouver la plupart des bornes classiques dans ce cas.

Après cela, nous nous sommes intéressés au nombre chromatique de l'espace Euclidien, dont la détermination est un problème ancien et toujours très ouvert. A ce propos, Frankl et Wilson ont remarqué le rôle joué par les ensembles qui évitent une distance donnée dans l'espace de Johnson. Dans [19] ils établissent une borne supérieure pour le cardinal de tels ensembles, cette borne n'étant valide que pour un choix limité de paramètres. Par là ils arrivent aussi à un résultat asymptotique sur le nombre chromatique Euclidien. Leur borne peut être améliorée et élargie à tout choix de paramètres en utilisant la borne de Delsarte pour les ensembles qui évitent une distance et aussi une borne du type Schrijver sur les triplets de points. Nous allons nous concentrer sur cette dernière en donnant la formulation générale en termes d'un programme semidéfini positif, ainsi que les détails de sa symétrisation. En raison de l'analogie existante entre l'espace de Hamming et l'espace projectif, cette symétrisation a plusieurs points en commun avec celle développée pour les codes projectifs. Plus précisément, les deux peuvent être vues comme des variantes de la diagonalisation par blocs de l'algèbre de Terwilliger de l'espace de Hamming.

Comme l'on peut le voir par ces applications de la méthode SDP, nous avons un bon cadre théorique que nous pouvons appliquer aux problèmes extrémaux de la théorie des graphes mais il est souvent difficile d'obtenir des résultats explicites. En effet, la plupart du travail nécessaire consiste à réduire le programme général selon les symétries de l'espace donné. Des travaux récents sur les hiérarchies de programmes semidéfinis positifs ([23], [32], [35]) peuvent aussi être utilisés pour améliorer certaines bornes, mais ces programmes ont également besoin d'être réduits avant de pouvoir en calculer les solutions. En revanche, chaque programme semidéfini positif peut s'écrire en termes de la plus grande valeur propre de certaines matrices. Pour exemple, dans le cas de ϑ l'on trouve une relaxation classique que, pour les graphes réguliers, donne la borne de Hoffman, borne explicite qui s'écrit par moyen des valeurs propres de la matrice d'adjacence. Notre but est de définir une hiérarchie de programmes semidéfinis positifs qui, tout en améliorant ϑ, se prêtent facilement à l'analyse en termes de valeurs propres, voire qui donnent des généralisations de la borne de Hoffman. A notre avis, le contenu du cinquième chapitre est un premier pas dans cette direction.

Nous y définissons une nouvelle hiérarchie de programmes semidéfinis positifs, dont la valeur optimale donne une borne supérieure sur le nombre d'indépendance d'un graphe. Pour la définir, nous utilisons le langage des complexes simpliciaux, lesquels peuvent s'interpréter comme une généralisation des graphes. Par moyen des opérateurs de bord nous construisons des analogues de deux matrices reliées à ϑ et à la relaxation de Hoffman, notamment la matrice dont toutes les entrées sont 1 et le Laplacian du graphe. Avec ceci, nous définissons la hiérarchie et en analysons les propriétés. Il est naturel qu'une hiérarchie de ce type donne ϑ au premier degré et qu'elle atteigne le nombre d'indépendance à un certain degré fini. Nous montrons cette propriété et d'autres aussi qui généralisent des propriétés du nombre ϑ. Nous ne sommes pas capables de prouver la décroissance de notre hiérarchie, mais nous montrons comment l'obtenir par le moyen d'une légère modification.

A titre d'exemple, nous calculons les valeurs du deuxième degré de notre hiérarchie pour les graphes de Paley. Cette famille de graphes est intéressante en raison de son comportement quasi-aléatoire et de ses applications en théorie des nombres. Il est très difficile d'estimer la taille de la plus grande clique des graphes de Paley. Comme ces graphes sont auto-complémentaires, cela revient à en estimer le nombre d'indépendance. Aussi, en étant fortement réguliers, leur nombre ϑ est donné par la borne de Hoffman. Au moment où cette thèse se termine, nous ne sommes pas en mesure d'améliorer les bornes existantes et le problème reste ouvert.

Structure et résultats de la thèse

Le troisième chapitre[1] est dédié à l'étude du problème de la détermination de bornes supérieures pour le cardinal des codes projectifs et des codes à dimension constante (aussi dits codes de Grassmann). Nous en donnons l'état de l'art et introduisons une borne SDP qui améliore certains résultats existants. La borne est obtenue en symétrisant le programme définissant ϑ sous l'action du groupe général linéaire. En outre, nous montrons que les bornes déjà existantes pour les codes de Grassmann peuvent s'obtenir comme cas particuliers de la borne de Delsarte, laquelle est à la fois un cas particulier du nombre ϑ, tout en montrant ainsi l'intérêt de cette méthode.

Le quatrième chapitre est consacré au nombre chromatique de l'espace Euclidien. Nous y rappelons la définition des graphes de Frankl et Wilson et le rôle qu'ils jouent dans ce contexte. Nous prouvons une borne SDP sur le cardinal des ensembles qui évitent une distance donnée dans l'espace de Johnson. Ce programme semidéfini positif est une variante de celui introduit par Schrijver pour les codes à poids constant ([41]). Le programme est ensuite symétrisé afin de pouvoir en calculer des valeurs explicites. Les résultats ainsi obtenus améliorent les bornes inférieures pour le nombre chromatique Euclidien *mesurable*.

Dans le cinquième chapitre nous définissons une nouvelle séquence, nommée ϑ_k, de programmes semidéfinis positifs qui donnent une borne supérieure pour le nombre d'indépendance d'un graphe. Cette séquence est liée à des matrices venant des complexes simpliciaux. Quelques propriétés de ϑ_k y sont analysées. En particulier, nous prouvons le théorème du sandwich pour tout degré k, ainsi qu'un résultat concernant les homomorphismes de graphes qui généralise une propriété bien connue de ϑ. Ensuite, nous modifions légèrement notre définition de ϑ_k pour assurer la décroissance de la valeur optimale à chaque degré. Nous donnons quelques résultats numériques pour des cycles et des graphes de Paley de petite taille.

Dans le sixième chapitre nous analysons en profondeur le problème de la détermination du nombre d'indépendance des graphes de Paley. Notre programme ϑ_2 est symétrisé, ce qui permet de calculer sa valeur optimale pour un vaste ensemble de paramètres. Ces valeurs sont comparées à celles obtenues par moyen du programme L^2 du [23], aussi rappelé dans le deuxième chapitre. Nous terminons le chapitre par quelques observations sur la structure des graphes de Paley.

[1] Où contenu et résultats de [8] sont reproduits.

Chapter 1

Introduction

Since its development in the 1950's, linear programming has become a fundamental tool in combinatorial optimization, as it gives a general formulation in which a multitude of problems can be expressed and often solved. Research in algorithmic issues and increased computational power have supported the development of the field and this approach gained more and more interest, giving rise to several related optimization problems, like integer programming and quadratic programming. One succesfull generalization is that of *semidefinite programming* (SDP).

A linear programming problem can be formulated as: finding a real vector with non negative entries which maximizes (or minimizes) a linear function, subject to some given linear inequalities. In a semidefinite programming problem we no longer consider vectors but symmetric matrices. Then the required condition is that the matrix is *positive semidefinite*, which means that all its eigenvalues are non negative. This is clearly a more general problem: indeed, requiring the matrix to be diagonal, we go back to a linear program.

In the 1980's, interior point method algorithms were developed for linear programs ([26]) and generalized to semidefinite programs ([4], [37], [38]), thus allowing semidefinite programs to be approximately solved by efficient algorithms. Meanwhile, SDP was applied to model problems arising in a very wide range of areas. Still nowadays, it plays an important role in the development of approximation algorithms for "hard" optimization problems.

Let us move more specifically to the motivation which is behind this work. In coding theory, a linear program was established by Delsarte's work on association schemes ([14]). Its optimal value gives an upper bound on the maximal size of a code with prescribed minimum distance. In particular, it has longtime remained one of the best numerical bounds on the size of codes in the Hamming space, leading also to the best known asymptotics.

The inequalities involved in Delsarte linear program for the Hamming space H_n can be derived by considering the distance distribution of a code. This distance distribution coincides with the orbit decomposition of pairs of words under the group G of automorphisms of H_n. Then the so-called Krawtchouk polynomials are strictly related to the G-irreducible decomposition of the space of complex valued functions on H_n and they are exactly the polynomials occurring in the inequalities of Delsarte linear program.

On the other hand, Delsarte bound equally applies to codes in other spaces, as the binary Johnson space or the real Euclidean unit sphere and, more generally, to all 2-point homogeneous spaces. The quest for generalizations of this method led to use semidefinite programming. In 2005, Schrijver established an SDP upper bound for the size of codes in the Hamming space which strictly improves over Delsarte's one for several parameters ([41]). Next, Schrijver's method was extended to non binary codes ([22]) and it was further improved in the binary case ([21]).

The underlying idea is to consider triples of words instead of pairs. Then it is possible to obtain Schrijver's bound using a similar approach to the one recalled above for Delsarte linear program. The difference is that in this case we consider the action of a subgroup $G' \leq G$, namely the stabilizer of one point. As H_n is not 2-point homogeneous for this action, the irreducible decomposition of the space of complex valued functions on H_n involves multiplicities greater than 1 and, roughly speaking, this is the reason for which one has to deal with matrices.

This idea has already been successfully applied to other spaces occurring in coding theory, like the unit sphere of the Euclidean space, providing new results in an old and still open geometrical problem, namely the kissing number problem ([10]).

One unifying approach is to see all the metric spaces involved as *graphs* where edges coincide with forbidden distances. Then these problems can often be formulated in terms of stable sets in the given graph. This is where Lovász ϑ number comes into play. In 1979, Lovász associated to each graph a semidefinite program, whose optimal value he called the ϑ number of the graph ([34]). Among other remarkable properties, it yields an upper bound on the cardinality of every stable set in the graph. It was early recognized that when a minor modification, called ϑ', is applied to a well chosen graph in the Hamming space, the Delsarte bound is recovered ([36], [40]).

On the other hand, recent research tend to focus on *hierarchies* of semidefinite programs ([23], [32], [35]). When applied to the problem of determining the independence number α of a graph, the first step of a hierarchy usually coincides with the Lovász ϑ number and successive steps give tighter upper bounds on α. Bounds for Hamming codes based on triples ([41]) or quadruples ([21]) of words can be interpreted in this context.

The initial motivation and starting point for this thesis was to understand how this framework could be applied to network coding. Indeed, network coding is a recent and dynamical field in coding theory which is capturing interest from both computer scientists and mathematicians. The coding theoretical formulation of the communication problem in terms of codes of subspaces by Koetter and Kschischang ([29]) establish a parallel with the classical Hamming codes. Nevertheless, important differences exist between the two cases, so a straightforward application of the same method is not possible and one needs to be more careful.

More precisely, codes for network coding are given by collections of linear subspaces of a given vector space over a finite field. Such codes are called *projective codes* or, if all the subspaces have the same dimension, *constant dimension codes*. They are the q-analog of Hamming codes and Johnson codes, respectively. Here, q is the cardinality of the base field and q-analogy is a combinatorial point of view replacing sets with subspaces and cardinality with dimension. Also, if we look at the group action, the symmetric group is replaced with the general linear group. Here we remark a first important difference: the Hamming space H_n has a transitive group action while in the projective space this does not happen. This already tells us that the Delsarte linear programming bound cannot be applied to projective codes. However, a semidefinite programming bound coming from Lovász ϑ number can be established. In order to compute the explicit bound in non trivial cases, the program defining ϑ needs to be symmetrized using the action of the general linear group. About constant dimension codes, the analogy with constant weight codes holds in a better way, so all classical bounds can be generalized and indeed they have already been established.

Next, we remarked that distance avoiding sets in the Johnson space play a role in determining the chromatic number of the Euclidean space, an old and open geometrical problem. Indeed, this was already recognized by Frankl and Wilson in [19], where they give an upper bound on the cardinality of such sets for a limited choice of parameters. With this they obtained an asymptotic result on the chromatic number of the Euclidean space. Their bound can be improved and extended to all parameters by using the Delsarte bound applied to distance avoiding sets and a Schrijver-like bound on triples of words. We mainly focus on this last one, giving its general SDP formulation and the details of the symmetrization of the program. In view of the analogy between Hamming space and projective space, the group theoretical setting that we use in the symmetrization has several common points with the one we developed for projective codes. Actually, both can be seen as variations of the block diagonalization of the Terwilliger algebra of the Hamming space.

As it can be seen from these applications of the SDP method, there is a good theoretical setting that we can apply to extremal problems in graph theory, but it is often hard to obtain explicit results. Indeed, most of the work needed to come up with explicit values of this kind of bounds consists of reducing the general program according to the symmetries of the space in which we are working. Also, recent research on SDP hierarchies ([23], [32], [35]) can be used to improve results in this kind of problems, but such SDPs are not ready to be computed. On the other hand, every semidefinite program has a formulation in terms of the largest eigenvalue of matrices. Looking at ϑ, it is a classical result to get a relaxation which, for regular graphs, yields the more explicit Hoffman bound, involving eigenvalues of the adjacency matrix. Our hope is to derive a hierarchy of semidefinite programs which strengthen ϑ but which remains easy to analyse in terms of eigenvalues and, possibly, which leads to high-order generalizations of the Hoffman bound. We believe that our work in chapter 5 is a first step in this direction.

We build a new hierarchy of semidefinite programs whose optimal value upper bounds the independence number of a graph. To define it, we adopt the framework of simplicial complexes. Indeed, simplicial complexes can be seen as a generalization of graphs. By mean of the boundary maps, we can build high order analogs of two matrices occurring in ϑ and its classical Hoffman relaxation, namely the all one's matrix and the Laplacian of the graph. With this, we define the hierarchy and we discuss its properties. It is natural to ask that a hierarchy of this kind starts from ϑ and reaches the independence number in finitely many steps. We show that this is the case, along with other properties that generalize some properties of ϑ. We are not able to show that our hierarchy is decreasing at any step but we show how this can be insured with a minor modification.

As an application we focus on Paley graphs, computing some values of the second step of our hierarchy. Paley graphs are interesting because of their random like behaviour and their number theoretical meaning. Determine or even estimate their clique number is a very hard open question. As these graphs are self complementary, the clique number coincides with the independence number. Moreover, being strongly regulars, they reach equality in the Hoffman bound so their ϑ number is explicitely known. At this moment, we are not able to give improvements over the existing bounds and the problem remains open.

1.1 Overview of the thesis and results

In the third chapter[1] we deal with the problem of determining upper bounds on the cardinality of codes for network coding, namely the class of projective codes and the subclass of Grassmann codes. We give the state of the art of this problem and we introduce an SDP bound that, for certain parameters, gives better results than the existing ones. This SDP is obtained by symmetrizing the ϑ program under the action of the general linear group. Moreover we show that the existing general bounds for Grassmann codes can be obtained as particular instances of the Delsarte LP bound, which is itself a special case of the SDP defining the ϑ number, thus showing the interest of this approach.

In the fourth chapter we deal with the Euclidean chromatic number. We recall the

[1]This chapter reproduces contents and results of [8].

definition of Frankl-Wilson graphs and the role that they play in this problem. We prove a more general SDP upper bound on the independence number of sets avoiding one distance in the Johnson space. This SDP is a variation of the one introduced by Schrijver for constant weight codes ([41]). The program is symmetrized in order to compute explicit values. The results obtained in this way give improvements on lower bounds for the *measurable* Euclidean chromatic number.

In the fifth chapter we define a new sequence ϑ_k of semidefinite programs which upper bound the independence number of a graph. It is related to matrices arising from abstract simplicial complexes. Some properties of ϑ_k are discussed. In particular, we prove that the sandwich theorem holds for any step of the new sequence, along with a result involving graph homomorphisms which generalizes one well known property of ϑ. Also, we slightly modify the formulation of ϑ_k to insure that the optimal value decreases at each step. We give some numerical results for small cycles and Paley graphs.

In the sixth chapter we analyse in a deeper way the independence number of Paley graphs. Our program ϑ_2 is symmetrized, thus allowing its explicit computation for a wider range of parameters. Then these values are compared to the ones of the program L^2 given in [23] and recalled in chapter 2. We end the chapter by sketching some further considerations on the structure of Paley graphs.

1.2 Notation

- for a natural number n, $[n]$ denotes the set $\{1, 2, \ldots, n\}$.
- for $0-1$ vectors, 0^m or 1^m means that m consecutive coordinates are 0 or 1 respectively. So, for example, $0^2 1^3 01 = (0, 0, 1, 1, 1, 0, 1)$.
- I denotes the identity matrix, whose size will be clear from the context.
- J denotes the square (sometimes rectangular) matrix with all 1's entries.
- $\langle , \rangle$ denotes the scalar product of matrices: given A, B matrices of the same size, $\langle A, B \rangle := \mathrm{Tr}(AB^T)$.
- given a finite set X, we denote $\mathbb{R}^X$ the real vector space of dimension $|X|$ whose basis elements are indexed by the set X; the space $\mathbb{R}^{X \times X}$ is then identified with the vector space of real square matrices with rows and columns indexed by X. Of course, $\mathbb{R}^{n \times n} = \mathbb{R}^{[n] \times [n]}$.
- a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive semidefinite, noted $A \succeq 0$, if one of the following equivalent conditions holds:

  (i) $v^T A v \geq 0$ for all vectors $v \in \mathbb{R}^n$,

  (ii) all eigenvalues of A are non negative,

  (iii) $A = BB^T$ for some $B \in \mathbb{R}^{n \times m}$ with $m \leq n$.

  We mention that semidefinite matrices form a self-dual cone. We write $A \succeq B$ to signify $A - B \succeq 0$.
- for X a finite set, we will often identify functions $F : X \times X \to \mathbb{R}$ with matrices $F \in \mathbb{R}^{X \times X}$, by $F(x, y) = F_{x,y}$. Positive semidefiniteness remains valid in both notations.
- when considering complex matrices, all works the same replacing symmetric with *hermitian* and transpose with *conjugate transpose.*

Chapter 2

Background

In this chapter we give the necessary background material. The first part is nothing else that basic representation theory. The results that we recall there will be used in chapter 3, 4 and 6 to decompose certain group representations. The second part sets up some basic vocabulary of graph theory. The third part is about semidefinite programming and it is at the core of what is done in this thesis. The classic Lovász ϑ number is introduced. Although all definitions and results are commonly known in the literature, they are reproduced here, along with some proofs, as they will serve as a model for the work that we present further on in chapter 5. In particular, we briefly sketch the Lasserre hierarchy for the independence number ([32]), along with a variant given by Gvozdenović, Laurent and Vallentin ([23]). We then outline, following [6], a general strategy to symmetrize the ϑ program under a group action, using tools of representation theory. A section on coding theory will serve to introduce some basic vocabulary and to give the first example of symmetrization of an SDP, showing that in the Hamming space the ϑ' program coincides with Delsarte program. Finally, the block diagonalization of the Terwilliger algebra of the Hamming scheme ([41], [45]) is explained, as it will serve to establish some results in chapter 4.

2.1 Representation theory of finite groups

We will consider here the representation theory of finite groups in characteristic zero, in particular over the complex field $\mathbb{C}$. A classic reference is given by the book of Jean-Pierre Serre, [42], which we will essentially follow for this introductory part. The goal is to understand how a given representation decomposes into irreducible subrepresentations and how to find such decomposition in an explicit way.

2.1.1 First definitions

Definition 2.1.1. *Let k be a field and G a finite group. A representation of G over k is a homomorphism $\rho : G \to GL(V)$, where V is a vector space over k.*

If V has finite dimension n over k, we say that the representation has *degree* n. The choice of a basis of V allows us to express $\rho(g)$ as an invertible matrix and to define a *matrix representation* $\rho' : G \to GL_n(k)$. Different choices of basis result in representations which are isomorphic in the sense of the following definition.

Definition 2.1.2. *Two representations of the same group G, $\rho : G \to GL(V)$ and $\rho' : G \to GL(V')$, are said to be* isomorphic *if there exists an isomorphism $\tau : V \to V'$, such that, for every $g \in G$ the following diagram is commutative:*

$$\begin{array}{ccc} V & \xrightarrow{\rho(g)} & V \\ \downarrow{\scriptstyle \tau} & & \downarrow{\scriptstyle \tau} \\ V' & \xrightarrow{\rho'(g)} & V' \end{array}$$

Clearly, isomorphic representations have the same degree. From the matrix point of view, $\rho(g)$ and $\rho'(g)$ are conjugated by an invertible matrix.

Examples :

- The trivial representation of G of degree n is defined by $\rho(g) := Id(V)$ for every $g \in G$, where $\dim(V) = n$. Of course, this definition is given modulo isomorphism, as it depends on the choice of V.

- Let V be the vector space over k whose basis vectors are indexed by elements of G: $V := \langle e_h : h \in G \rangle$. The regular representation of G is defined by $\rho(g)(e_h) := e_{gh}$ extended by linearity. Its degree is $|G|$.

Definition 2.1.3. *Given two representations of the same group G, $\rho : G \to GL(V)$ and $\rho' : G \to GL(V')$, we define their sum $(\rho \oplus \rho') : G \to GL(V \oplus V')$ by $(\rho \oplus \rho')(g) = (\rho(g), \rho'(g))$.*

The degree of the sum is the sum of the degrees. If ρ and ρ' are given in matrix representation, then a corresponding matrix for $(\rho \oplus \rho')(g)$ is

$$\begin{pmatrix} \rho(g) & 0 \\ 0 & \rho'(g) \end{pmatrix}$$

Definition 2.1.4. *Let $\rho : G \to GL(V)$ be a representation of G and W a linear subspace of V. If W is stable with respect to the action of G, that is, if for every $g \in G$, $\rho(g)(W) \subset W$ then the restriction $\rho_{|W} : G \to GL(W)$ is a* subrepresentation *of ρ, of degree* $\dim(W)$.

Every representation has two trivial subrepresentations: the ones defined by the trivial subspaces $\{0\}$ and V.

Definition 2.1.5. *A representation is* irreducible *if it doesn't have non trivial subrepresentations.*

Examples :

- Every representation of degree 1 is irreducible.

The following lemma is a basic result which plays a central role in the development of representation theory.

Lemma 2.1.6 (Schur)**.** *Let $\rho : G \to GL(V)$ and $\rho' : G \to GL(V')$ be two irreducible representations and let $f : V \to V'$ be a linear homomorphism which commutes with the action of G, that is, $f \circ \rho(g) = \rho'(g) \circ f$ for every $g \in G$. Then $f = 0$ unless $\rho \simeq \rho'$, in which case f is an isomorphism. If moreover the ground field is algebraically closed and $V = V'$, then f is a scalar multiple of the identity.*

For the rest of the chapter we will consider only finite dimensional representations over the complex field $\mathbb{C}$, which is enough for the purpose of this thesis.

2.1.2 Decomposition of a representation

Theorem 2.1.7. *Let $\rho : G \to GL(V)$ be a reducible representation and $\rho_{|W} : G \to GL(W)$ one of its non trivial subrepresentations. There exists a linear subspace W' which is a complement of W in V and which is stable with respect to the action of G. In other words, $V = W \oplus W'$ and $\rho = (\rho_{|W} \oplus \rho_{|W'})$.*

Proof. Let $\langle , \rangle$ be a scalar product on V. We define a G–invariant scalar product $\langle , \rangle_G$ by setting

$$\langle v, v' \rangle_G := \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)(v), \rho(g)(v') \rangle$$

Then the orthogonal complement of W with respect to $\langle , \rangle_G$ has the desired properties. □

Iterating the theorem above yields the following classical result.

Corollary 2.1.8 (Maschke theorem)**.** *Every representation decomposes as a direct sum of irreducible subrepresentations.*

Remark 2.1.9. This theorem remains valid for any base field k, as long as its characteristic doesn't divide $|G|$.

Remark 2.1.10. From another equivalent point of view we can think about representations of G over k as modules of the group algebra $k[G]$. In this case, Maschke theorem states that $k[G]$ is a semisimple algebra.

Remark 2.1.11. The decomposition announced in the previous corollary is in general not unique. Uniqueness is provided by the decomposition into isotypic components, as we explain next.

Let $V = V_1 \oplus \cdots \oplus V_m$ be a decomposition of V in irreducible subrepresentations and let $\{W_1, W_2, \ldots W_r\}$ be the collection of irreducible subrepresentations of V, modulo isomorphism. The *isotypic components* of V are defined by $\mathcal{I}_j := \oplus_{i:V_i \simeq W_j} V_i$. So $\mathcal{I}_j \simeq m_j W_j$ where $m_j := \mathrm{card}\{i : V_i \simeq W_j\}$ is the multiplicity of W_j in V. We obtain that

$$V = \oplus_j \mathcal{I}_j \simeq \oplus_j m_j W_j \tag{2.1.1}$$

This decomposition is unique in the sense that each $\mathcal{I}_j$ is uniquely determined and does not depend on the original decomposition of V into irreducibles. In general, every isotypic component $\mathcal{I}_j$ has several irreducible decompositions but clearly all of them are isomorphic. In particular, the multiplicity and the dimension of every W_j are uniquely determined.

2.1.3 Characters

As we will see, characters encode a lot of properties of representations and they are a fundamental tool for performing calculations. We list here the main results.

Definition 2.1.12. *The character of a representation $\rho : G \to GL(V)$ is the function $\chi_\rho : G \to \mathbb{C}$ defined by $\chi_\rho(g) := \mathrm{Tr}(\rho(g))$.*

From a well known property of the trace, it follows that characters are *class functions*, that is, they are invariant on conjugacy classes of G. Moreover,

Theorem 2.1.13. *The characters of irreducible pairwise non isomorphic representations of G form a basis of the space of class functions on G.*

Corollary 2.1.14. *The number of irreducible pairwise non isomorphic representations of G equals the number of conjugacy classes of G.*

Given two characters χ_ρ, $\chi_{\rho'}$ we consider their scalar product, given by

$$\langle \chi_\rho, \chi_{\rho'} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \overline{\chi_{\rho'}(g)}$$

where $\overline{\cdot}$ means the complex conjugate.

Then the next result says that the basis of theorem 2.1.13 is an orthonormal one.

Theorem 2.1.15. *If ρ is an irreducible representation, then $\langle \chi_\rho, \chi_\rho \rangle = 1$. If $\rho \not\simeq \rho'$, then $\langle \chi_\rho, \chi_{\rho'} \rangle = 0$.*

Corollary 2.1.16. *Consider the decomposition (2.1.1). For every j, we have that $\langle \chi_V, \chi_{W_j} \rangle = m_j$.*

Corollary 2.1.17. *A representation ρ is irreducible if and only if $\langle \chi_\rho, \chi_\rho \rangle = 1$.*

By computing the character of the regular representation of a group G, we see that every irreducible representation W_j of G is contained in the regular representation with multiplicity equal to its degree n_j. This yields the *sum of squares* formula:

Proposition 2.1.18. *Let $\{W_1, W_2, \dots W_r\}$ be the collection of irreducible representations of a group G, modulo isomorphism. Set $n_j := \deg(W_j)$. Then*

$$|G| = \sum_{j=1}^{r} n_j^2$$

The next theorem reveals the importance of characters in the study of representations.

Theorem 2.1.19. *Two representations are isomorphic if and only if they have the same character.*

2.1.4 Restriction and induction

Here we see how representations of a group can be constructed from those of its subgroups or from those of larger groups. Let $H \leq G$ be two finite groups.

Definition 2.1.20. *If $\rho : G \to GL(V)$ is a representation of G, then the restriction $\rho_{|H} : H \to GL(V)$ is a representation of H, denoted $Res_H^G(\rho)$ or $\rho \downarrow_H^G$.*

Definition 2.1.21. *If $\mu : H \to GL(V)$ is a representation of H, then the induced representation is built as follows. Consider the space $V' := \oplus_{i=1}^m x_i V$, where $\{x_1, \dots, x_m\}$ is a system of representatives of G/H. For $g \in G$ and for every i, there exists an index j and an element $h \in H$ such that $gx_i = x_j h$. Then we obtain a representation of G, $\mu' : G \to GL(V')$ by setting $\mu'(g)(x_i v) := x_j(\mu(h)v)$. The induced representation is denoted $Ind_H^G(\mu)$ or $\mu \uparrow_H^G$.*

Using an analog notation for characters, we can mention the following result which shows the duality between restriction and induction.

Theorem 2.1.22 (Frobenius reciprocity)**.** *Let χ be a character of G and ψ be a character of H. We have*

$$\langle \chi, Ind_H^G(\psi) \rangle = \langle Res_H^G(\chi), \psi \rangle$$

2.2 Notations and basic definitions of graph theory

A graph consists of a set V of *vertices* together with a set $E \subset \binom{V}{2}$ of *edges*. The notation is $\mathcal{G} = (V, E)$. If not explicitely stated, all graphs which appear in this thesis will be *finite* graphs, i.e. with a finite number of vertices. Moreover, the definition implies that all graphs appearing here will be undirected, without loops and multiple edges. In chapter 3 we will define a network as a *directed graph*, where E is taken to be a collection of ordered pairs of vertices.

We will usually simplify the notation, writing edges as vv' instead of $\{v, v'\}$. Given two vertices $v, v' \in V$, we say that v is *adjacent* to v' if $vv' \in E$. This is a symmetric relation and it can be encoded in the *adjacency matrix* $A(\mathcal{G})$ of $\mathcal{G}$, defined as the matrix whose rows and columns are indexed by V and whose entries are

$$A(\mathcal{G})_{v,v'} := \begin{cases} 1 & \text{if } vv' \in E \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $A(\mathcal{G})$ is a real symmetric matrix and we mention that its spectrum encodes some important properties of the graph.

An adjacent vertex is called a *neighbour*. The number of neighbours of a vertex v is called the *degree* of v, noted $\deg(v)$. A graph is called *regular* if all of its vertices have the same degree d, which in this case is also the largest eigenvalue of the adjacency matrix. Some easy examples of regular graphs are complete graphs $K_n = ([n], \binom{[n]}{2})$, null (or empty) graphs $([n], \emptyset)$ and cycles $C_n = ([n], \{ij : j - i = 1\})$.

A graph is called *strongly regular* if it is regular and moreover the number of common neighbours of a pair of vertices v, v' depends only on whether vv' is an edge or not, hence taking only two values λ, μ, respectively. In that case, the graph is said to be (n, d, λ, μ)-strongly regular, where n is the number of vertices and d the degree.

A *subgraph* of $\mathcal{G}$ is a graph $\mathcal{H}$ whose vertex and edge sets are subsets of those of $\mathcal{G}$. The subgraph $\mathcal{H} = (V', E')$ is called an *induced* subgraph of $\mathcal{G} = (V, E)$ if $E' = E \cap \binom{V'}{2}$. If $V' = V$, then $\mathcal{H}$ is called a *spanning* subgraph of $\mathcal{G}$.

Two graphs $\mathcal{G}, \mathcal{G}'$ are *isomorphic* if there exists a bijection between their vertex sets which preserves the adjacency relation. Such a map is called an *isomorphism* and we write $\mathcal{G} \simeq \mathcal{G}'$. When $\mathcal{G} = \mathcal{G}'$, we talk about *automorphisms* of $\mathcal{G}$.

Given a graph $\mathcal{G} = (V, E)$, the *complement graph* is defined as $\overline{\mathcal{G}} = (V, \overline{E})$ where $\overline{E}$ is the complement of E in $\binom{V}{2}$.

A set of pairwise non adjacent vertices of $\mathcal{G}$ is called a *stable set* (or *independent set*) of $\mathcal{G}$. The cardinality of the largest independent set is called the *independence number* of $\mathcal{G}$, noted $\alpha(\mathcal{G})$. The dual concept is that of a clique. A set of pairwise adjacent vertices of $\mathcal{G}$ is called a *clique* of $\mathcal{G}$. So a clique is an induced subgraph which

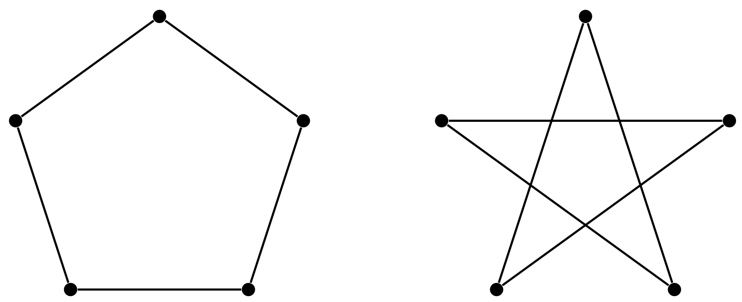

Figure 2.a: The cycle of length 5 is isomorphic to its complement

is isomorphic to a complete graph. The cardinality of the largest clique is called the *clique number* of $\mathcal{G}$, noted $\omega(\mathcal{G})$. We note that an independent set of $\mathcal{G}$ is a clique of $\overline{\mathcal{G}}$, so that $\alpha(\mathcal{G}) = \omega(\overline{\mathcal{G}})$.

A k-*colouring* of $\mathcal{G}$ is a map $c : V \to [k]$ such that $c(v) \neq c(v')$ for all $vv' \in E$. In other words, it is a partition of the vertex set into k independent sets. The least integer k such that a k-colouring exists is called the *chromatic number* of $\mathcal{G}$ and denoted $\chi(\mathcal{G})$.

Determining α and χ is a fundamental problem in theoretical computer science, which is known to be NP-hard ([27]).

2.3 Semidefinite programming and Lovász ϑ number

In this third part we introduce some of the main tools that we will use in this thesis, namely linear programs (LPs) and semidefinite programs (SDPs). We mention that they are particular convex programs and thus can be analyzed in the general setting of *convex optimization problems* ([46]), although this point of view is more general and not needed for our work. A link between graphs and SDPs is provided by the ϑ number, introduced by Lovász in 1979. This number have several interesting properties. What will be more interesting here is that, for any graph, the optimal value of the program defining ϑ gives an upper bound on its independence number. Stronger upper bounds are provided by some hierarchies of SDPs and we recall two of them.

2.3.1 LPs and SDPs

A *linear program* (LP) is an optimization problem in which we ask for maximizing (or minimizing) a linear function under some constraints given by linear inequalities.

A linear program can be given in its classical *primal* form:

$$\sup\left\{c^T x \quad \text{subject to :} \quad \begin{array}{l} x \geq 0 \\ Ax \leq b \end{array}\right\} \tag{2.3.1}$$

where $c, x, b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$, $x_1, \dots, x_n$ being the variables. The *dual* of the above program is the following:

$$\inf\left\{y^T b \quad \text{subject to :} \quad \begin{array}{l} y \geq 0 \\ y^T A \geq c^T \end{array}\right\}$$

A vector x satisfying all constraints in 2.3.1 is called a *feasible solution*. The program is called feasible or unfeasible according to the existence of such a vector. The scalar product $c^T x$ is called the *objective function* and, for x feasible, $c^T x$ is called the *objective value* of x. A vector x is *optimal* if its objective value is maximal amongst all objective values of feasible solutions. In such case, $c^T x$ is called the *optimal value* of 2.3.1. A feasible problem is said to be *unbounded* if the objective function can assume arbitrarily large positive objective values. Otherwise, it is said to be *bounded*. The same vocabulary (with the obvious variations) is used for the dual program.

A pair of primal and dual linear programs are moreover related by two duality theorems called weak duality and strong duality. The *weak duality* theorem says that any objective value of the dual is larger than any objective value of the primal. The *strong duality* theorem states that if one of the programs is bounded feasible, then the other is also bounded feasible and their optimal values coincide.

A real *semidefinite program* (SDP) is an optimization problem of the following form:

$$\sup\left\{ \langle C, X\rangle \quad \text{subject to:} \quad \begin{array}{l} X \succeq 0, \\ \langle A_i, X\rangle = b_i \text{ for } i = 1, \dots, m \end{array} \right\} \tag{2.3.2}$$

where

- $C, A_1, \dots, A_m$ are given symmetric matrices,
- $b_1, \dots, b_m$ are given real values,
- X is a variable matrix,
- $\langle A, B\rangle = \text{trace}(AB)$ is the usual scalar product between matrices,
- $X \succeq 0$ means that X is symmetric positive semidefinite.

This formulation includes linear programming as the special case when all matrices involved are diagonal matrices. A program like (2.3.2) is called in *primal* form. A semidefinite program can also be expressed in its *dual* form:

$$\inf\left\{\ b^T y \quad \text{subject to:} \quad \textstyle\sum_{i=1}^m y_i A_i \succeq C\ \right\} \tag{2.3.3}$$

So, given $(C, A_1, \ldots, A_m, b_1, \ldots, b_m)$ we will talk about the two programs above as the pair of primal and dual SDP's associated to such data. The definitions of feasibility and optimality for semidefinite programs are the same as for linear programs. We refer to the vaste literature on SDP (e.g. [44], [47]) for further details on this topic. Let us just mention that semidefinite programs can be approximated in polynomial time within any specified accuracy by the ellipsoid algorithm or by practically efficient interior point methods.

The duality theory holds as well in this case.

Proposition 2.3.1 (Weak duality)**.** *For every X primal feasible and for every y dual feasible, we have*

$$\langle C, X\rangle \leq b^T y \tag{2.3.4}$$

In particular, the optimal value of the primal program is less or equal than the optimal value of the dual program. Moreover, whenever we have feasible solutions X and y such that equality holds in 2.3.4, then they are optimal solutions for their respective programs.

The quantity $b^T y - \langle C, X\rangle$ is called *duality gap* between the primal feasible solution X and the dual feasible solution y. To announce the strong duality theorem for SDP, we need one last definition: we say that (a solution of) the program 2.3.2 or 2.3.3 is *strictly feasible* if the matrix is positive definite, i.e. positive semidefinite and non singular.

Proposition 2.3.2 (Strong duality)**.** *If one of the primal or of the dual program is bounded and strictly feasible, then the other one is bounded feasible, it has an optimal solution and the two optimal values coincide.*

2.3.2 Lovász ϑ number

2.3.2.1 Shannon capacity

Given two graphs $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$, their *strong product* $\mathcal{G} \boxtimes \mathcal{H}$ is defined as the graph with vertex set the cartesian product $V \times V'$ and with edges $\{(a, b)(a', b') :\ a = a' \text{ or } aa' \in E \ \text{ and } \ b = b' \text{ or } bb' \in E\}$. Let $\mathcal{G}^k$ denote the strong product of $\mathcal{G}$ with itself k times.

Definition 2.3.3. *The Shannon capacity of a graph* $\mathcal{G}$ *is defined by*

$$\Theta(\mathcal{G}) = \sup_k \sqrt[k]{\alpha(\mathcal{G}^k)} = \lim_{k\to+\infty} \sqrt[k]{\alpha(\mathcal{G}^k)}$$

The number $\Theta(\mathcal{G})$ was introduced by Shannon in 1956 ([43]) with the following interpretation. Suppose that vertices of the graph $\mathcal{G}$ represent letters of an alphabet and that the adjacency relation means that one letter can be confused with the other. Then $\alpha(\mathcal{G}^k)$ is the maximum number of messages of length k such that no two of them can be confused and $\Theta(\mathcal{G})$ can be seen as an information theoretical parameter representing the effective size of an alphabet in a communication model represented by $\mathcal{G}$.

We have that $\alpha(\mathcal{G}) \leq \Theta(\mathcal{G})$ and equality does not hold in general. For instance, consider C_5, the cycle of length 5. Clearly $\alpha(C_5) = 2$ and Lovász showed that $\Theta(C_5) = \sqrt{5}$ ([34]). He obtained this result by introducing a general upper bound on Θ, for any graph. This upper bound is called the *Lovász* ϑ *function*.

2.3.2.2 Definition and properties of ϑ

There are several equivalent definitions for Lovász ϑ function. Here we list only the two that are more suitable to handle and which we will consider all along this thesis.

Definition 2.3.4 ([34], theorem 4)**.** *Let* $\mathcal{G} = (V, E)$ *be a graph, then*

$$\vartheta(\mathcal{G}) := \max\Big\{ \sum_{x,y\in V} F(x,y) \ : \ \begin{array}{l} F \in \mathbb{R}^{V\times V},\ F \succeq 0 \\ \sum_{x\in V} F(x,x) = 1 \\ F(x,y) = 0 \text{ if } xy \in E \end{array} \Big\} \tag{2.3.5}$$

Equivalently, by duality,

Lemma 2.3.5 ([34], theorem 3)**.** *Let* $\mathcal{G} = (V, E)$ *be a graph, then*

$$\vartheta(\mathcal{G}) = \min\Big\{ \lambda_{max}(Z) \ : \ \begin{array}{l} Z \in \mathbb{R}^{V\times V},\ Z \text{ symmetric} \\ Z_{x,y} = 1 \text{ if } x = y \text{ or } xy \in \overline{E} \end{array} \Big\} \tag{2.3.6}$$

where $\lambda_{max}(Z)$ *denotes the largest eigenvalue of* Z*.*

Remark 2.3.6. Taking $F = (1/|V|)I$, we see that a strictly feasible solution exists for the primal program defining $\vartheta(\mathcal{G})$, hence strong duality holds and the definition of $\vartheta(\mathcal{G})$ carries no ambiguity.

This function is interesting as it can be used to approximate numbers which are hard to compute, like the independence number or the chromatic number of a graph, as we will see next. Moreover, being formulated as the optimal value of a semidefinite program, an approximation is computable in polynomial time in the size of the graph. Finally, we notice that every feasible solution of the primal or of the dual gives respectively a lower or an upper bound of the optimal value.

Proposition 2.3.7 ([34], lemma 3)**.** *For any graph $\mathcal{G}$, $\alpha(\mathcal{G}) \leq \vartheta(\mathcal{G})$.*

Proof. Let $S \subset V$ be an independent set. Let $\mathbf{1}_S \in \mathbb{R}^V$ be its characteristic vector (with 1 in entries indexed by elements of S and 0 elsewhere). Define the matrix

$$X_S := \frac{1}{|S|}\mathbf{1}_S\mathbf{1}_S^T \in \mathbb{R}^{V\times V} .$$

By construction, this matrix is positive semidefinite of trace 1 and its entries are 0 when at least one of the indices lies outside S. So it fulfils the condition of the program (2.3.5) and it has objective value $|S|$. The optimal value being the maximum among all feasible values, we have $|S| \leq \vartheta(\mathcal{G})$ for any independent set S, from which we derive the announced inequality. □

Remark 2.3.8. The program given in 2.3.5 is one of the equivalent formulations of Lovász original $\vartheta(\mathcal{G})$. If the additional constraint that F takes non negative values is added, the optimal value gives a sharper bound for $\alpha(\mathcal{G})$ denoted $\vartheta'(\mathcal{G})$.

Proposition 2.3.9 ([34], theorem 1)**.** *For any graph $\mathcal{G}$, $\Theta(\mathcal{G}) \leq \vartheta(\mathcal{G})$.*

Proof. By previous proposition and the inequality $\vartheta(\mathcal{G} \boxtimes \mathcal{H}) \leq \vartheta(\mathcal{G})\vartheta(\mathcal{H})$ proved in [34] (lemma 2), we have $\alpha(\mathcal{G}^k) \leq \vartheta(\mathcal{G}^k) \leq \vartheta(\mathcal{G})^k$. □

Now we recall and prove another famous result, known as the *sandwich theorem*.

Proposition 2.3.10 ([34], corollary 3)**.** *For any graph $\mathcal{G}$, $\alpha(\mathcal{G}) \leq \vartheta(\mathcal{G}) \leq \chi(\overline{\mathcal{G}})$.*

Proof. The first inequality has already been established. We prove the second one. A coloring of $\overline{\mathcal{G}}$ of size $c := \chi(\overline{\mathcal{G}})$ is a partition $(C_1, \ldots, C_c)$ of the vertex set in cliques of $\mathcal{G}$. We can reorder the vertex set according to such partition. Then the block matrix

$$Z := \begin{pmatrix} I & J & J & \ldots \\ J & I & J & \ldots \\ J & J & I & \ddots \end{pmatrix}$$

where blocks correspond to sets of the partition, is feasible for the program (2.3.6). Here we abuse a little of the notation of J, as the non diagonal blocks need not be square blocks. Consider the matrix $cI - Z$ which has diagonal blocks of the form $(c-1)I$ and non diagonal blocks of the form $-J$. As all of its minors are positive, $cI - Z$ is positive semidefinite and thus the largest eigenvalue of Z is less than or equal to c. As Z is feasible for (2.3.6), we have $\vartheta(\mathcal{G}) \leq \lambda_{max}(Z) \leq c = \chi(\overline{\mathcal{G}})$. □

Now we show how ϑ behaves when two graphs are related by homomorphism in the sense of the following definition.

Definition 2.3.11. *Let $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$ be two graphs. A graph homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$, is an application φ from V to V', which preserves edges, that is $(i, j) \in E \Rightarrow (\varphi(i), \varphi(j)) \in E'$.*

Proposition 2.3.12. *If there exists a homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$, then $\vartheta(\overline{\mathcal{G}}) \leq \vartheta(\overline{\mathcal{H}})$.*

Proof. Take any feasible solution F of the program defining $\vartheta(\overline{\mathcal{G}})$. In particular $F_{i,j} = 0$ whenever $i \neq j$, $ij \notin E$. Define $F' \in \mathbb{R}^{V' \times V'}$ by

$$(F')_{u,v} := \sum_{\substack{i \,:\, \varphi(i)=u \\ j \,:\, \varphi(j)=v}} F_{i,j}$$

Then, using the fact that $(u, v) \notin E' \Rightarrow (i, j) \notin E$ for all i, j such that $\varphi(i) = u, \varphi(j) = v$ we see that F' is a feasible solution of the program defining $\vartheta(\overline{\mathcal{H}})$ and that its objective value is the same as the one of F. The announced inequality follows. ☐

We conclude this overview of classical facts about the ϑ function recalling two more results from the original paper of Lovász.

Proposition 2.3.13 ([34, corollary 2 and theorem 8])**.** *For all graphs $\mathcal{G} = (V, E)$, $\vartheta(\mathcal{G})\vartheta(\overline{\mathcal{G}}) \geq |V|$. If moreover the graph is vertex transitive, equality holds.*

There are few families of graphs for which the ϑ number is established. One of these is the family of cycles C_n. If $n = 2m$ is even, then C_{2m} is a bipartite graph and it is easy to see that $\alpha(C_{2m}) = \vartheta(C_{2m}) = m$.

Proposition 2.3.14 ([34], corollary 5)**.** *For n odd,*

$$\vartheta(C_n) = \frac{n \cos(\pi/n)}{1 + \cos(\pi/n)}$$

2.3.3 Hierarchies of SDPs for the independence number

In many combinatorial problems it is helpful to obtain an efficient description of an approximation $P^{(1)}$ of a certain *polytope* P. For example, P can be given as the convex hull of the solution vectors of a linear $0-1$ program, and $P^{(1)} \supseteq P$ can be obtained by mean of linear (or semidefinite) constraints. In this case, $P^{(1)}$ is called a linear (or semidefinite) *relaxation* of P. More generally, given an optimization problem written in the form of a (linear, semidefinite, ...) program $\mathcal{P}$, a *relaxation* of $\mathcal{P}$ is another program $\mathcal{P}^{(1)}$ with the properties that the feasible region of $\mathcal{P}^{(1)}$ contains the feasible region of $\mathcal{P}$. A sequence of the kind

$$P^{(1)} \supseteq P^{(2)} \supseteq \cdots \supseteq P^{(\alpha)} = P$$

is called a *hierarchy of relaxations* of the polytope P. Analogously, if dealing with programs, we will talk about hierarchies of relaxations of a program $\mathcal{P}$. This is a large domain in optimization theory, but for the purpose of this thesis, we restrict ourselves to hierarchies of semidefinite programs.

Lovász and Schrijver ([35]) and Lasserre ([32]) have defined hierarchies of semidefinite programs, as relaxations of (linear or non linear) $0-1$ programs. The hierarchy of Lasserre refines the one of Lovász and Schrijver, as it is shown in [33]. In [23], the authors introduced a new hierarchy. They then applied the second and third steps of their hierarchy to the problem of calculating the independence number of Paley graphs. Here we sketch the details of the Lasserre hierarchy for the independence number ([32]), along with the variant of Gvozdenović, Laurent and Vallentin ([23]).

2.3.3.1 The Lasserre hierarchy for the independence number

Given a graph $\mathcal{G} = (V, E)$, the *stable set problem* is: find a stable subset $S \subset V$ of maximal cardinality $|S| = \alpha(\mathcal{G})$. The *stable set polytope* is defined as the convex hull of the characteristic vectors of all stable subsets of V:

$$\mathrm{STAB}(\mathcal{G}) := \mathrm{conv}\{\mathbf{1}_S \in \{0,1\}^V : (S \times S) \cap E = \emptyset\}$$

Then we have

$$\alpha(\mathcal{G}) = \max\{\sum_{i \in V} x_i : x \in \mathrm{STAB}(\mathcal{G})\} \tag{2.3.7}$$

Lasserre hierarchy applies to the stable set problem, giving approximations of $\mathrm{STAB}(\mathcal{G})$. As a consequence of 2.3.7, it applies to bound the independence number. We are going to describe this last formulation.

Given a finite set V, let $\binom{V}{t}$ denote the collection of subsets of V with cardinality t and $\binom{V}{\leq t}$ the collection of subsets of V with cardinality less or equal than t. The Lasserre hierarchy is based on moment matrices:

Definition 2.3.15. *The moment matrix of a vector $y \in \mathbb{R}^{\binom{V}{\leq 2t}}$ is the matrix $M_t(y) \in \mathbb{R}^{\binom{V}{\leq t} \times \binom{V}{\leq t}}$ defined by:*

$$M_t(y)_{I,J} := y_{I \cup J}$$

Definition 2.3.16. *Given a graph $\mathcal{G} = (V, E)$, for any $t \in \mathbb{N}$ we define*

$$\ell^{(t)}(\mathcal{G}) := \sup\Big\{ \sum_{i \in V} y_i \ : \ \begin{array}{l} y \in \mathbb{R}^{\binom{V}{\leq 2t}} \\ M_t(y) \succeq 0 \\ y_\emptyset = 1 \\ y_{\{i,j\}} = 0 \text{ if } ij \in E \end{array} \Big\} \tag{2.3.8}$$

The program 2.3.8 is called the t-th iterate (or step) of the Lasserre hierarchy for the stable set problem. It can be verified that $\ell^{(1)}(\mathcal{G}) = \vartheta(\mathcal{G})$, so we can see $\ell^{(t)}$ as a generalization of Lovász ϑ number. It is worth to notice that, for a feasible solution y, the second and fourth constraints imply that $y_I = 0$ whenever I contains an edge. The next results show the interest of this hierarchy.

Lemma 2.3.17. *For any t, $\alpha(\mathcal{G}) \le \ell^{(t)}(\mathcal{G})$, with equality for $t \ge \alpha(\mathcal{G})$.*

Proof. For the first statement, it suffices to show that $|S| \le \ell^{(t)}(\mathcal{G})$ whenever $S \subset V$ is a stable set. For this, we fix S stable and we define a vector $y_S \in \mathbb{R}^{\binom{V}{\le 2t}}$ by

$$(y_S)_A := \begin{cases} 1 & \text{if } A \subset S \\ 0 & \text{if not} \end{cases}$$

Then $M_t(y_S)$ is positive semidefinite as it is a submatrix of $y_S^T y_S$. The two last constraints are obviously verified. So y_S is a feasible solution and $\sum_{i \in V} (y_S)_i = |S|$. The second statement follows from the fact that the Lasserre hierarchy refines the Sherali-Adams hierarchy (see [33] for details). $\square$

The following proposition is easy to establish.

Proposition 2.3.18. *For any t, $\ell^{(t+1)}(\mathcal{G}) \le \ell^{(t)}(\mathcal{G})$.*

Remark 2.3.19. If in 2.3.8 we add the constraint that the entries of $M_t(y)$ are non negatives, then the new program $\ell^{(t)}_{\ge 0}$ gives a stronger upper bound on the independence number of graphs. Moreover, $\ell^{(1)}_{\ge 0}$ coincides with ϑ' as defined in remark 2.3.8.

2.3.3.2 The hierarchy L^t for the independence number

Note that the computation of 2.3.8 involves a matrix of size $O(|V|^t)$ and $O(|V|^{2t})$ variables. In [23], the authors considered a less costly variation of the Lasserre hierarchy, denoted L^t. Let us sketch their construction. As before, we restrict to the case of upper bounds on the independence number. We need to consider the following principal submatrices of moment matrices:

Definition 2.3.20. *Given a vector $y \in \mathbb{R}^{\binom{V}{\le 2t}}$ and a subset $T \subset V$ of cardinality $|T| = t - 1$ we define the matrix $M(T, y)$ as the principal submatrix of $M_t(y)$ whose rows and columns are indexed by the following set:*

$$\bigcup_{S \subset T} \{S, S \cup \{i\}\}_{i \in V}$$

Note that to define $M(T, y)$ for a set T of cardinality $t-1$, we use only the entries of y indexed by $\binom{V}{\le t+1}$. Then we can define the new hierarchy:

Definition 2.3.21. *Given a graph $\mathcal{G} = (V, E)$, for any $t \in \mathbb{N}$ we define*

$$L^t(\mathcal{G}) \;:=\; \sup\left\{ \sum_{i\in V} y_i \;:\; \begin{array}{l} y \in \mathbb{R}^{\binom{V}{\leq t+1}} \\ M(T,y) \succeq 0 \;\forall\, T \in \binom{V}{t-1} \\ y_\emptyset = 1 \\ y_{\{i,j\}} = 0 \text{ if } ij \in E \end{array} \right\} \tag{2.3.9}$$

The following properties are easily verified:

Proposition 2.3.22. *For any graph $\mathcal{G}$, for any $t \in \mathbb{N}$, the number $L^t(\mathcal{G})$ has the following properties:*

- *for $t = 1$, $L^1(\mathcal{G}) = \ell^{(1)}(\mathcal{G}) = \vartheta(\mathcal{G})$;*
- *$\ell^{(t)}(\mathcal{G}) \leq L^t(\mathcal{G})$;*
- *$\alpha(\mathcal{G}) \leq L^t(\mathcal{G})$, with equality for $t \geq \alpha(\mathcal{G})$;*
- *$L^{t+1}(\mathcal{G}) \leq L^t(\mathcal{G})$.*

So the bound on the independence number provided by this hierarchy is not as good as the one of Lasserre $\ell^{(t)}$ but we note that the complexity has decreased. Indeed, the program 2.3.9 involves $\frac{|V|^{t+1}}{(t+1)!} + O(|V|^t)$ variables and (with some further manipulation) $\frac{2^{t-1}|V|^{t-1}}{(t-1)!} + O(|V|^{t-2})$ matrices of size $n+1$ (see [23] for details).

2.4 Symmetrization and examples

To symmetrize a semidefinite program means to exploit its symmetries (coming from some group action) in order to obtain another program which is easier to compute and *equivalent* to the original one, in the sense that the symmetrization transforms feasible solutions in feasible solutions and leaves the objective value invariant. As a consequence, the symmetrized program will have the same optimal value as the original one. Consider the ϑ program (2.3.5). In order to be calculated explicitely for "big" graphs, ϑ needs to be defined by a less costly program. This can be done thanks to automorphisms of the graph and using tools from representation theory. We explain how in this final part of chapter 2, giving the link to one of the fundamental problems in coding theory. We follow [6], to which we refer for a more complete treatment of the subject. At the end, we sketch the block diagonalization of the Terwilliger algebra of the Hamming scheme, because we will need it in chapter 4. It was first obtained by Schrijver ([41]) in order to derive a bound for codes in the Hamming space which improves over Delsarte's one. Here we follow the equivalent reformulation in terms of group representations given in [45].

2.4.1 Symmetrization of the SDP defining ϑ

Fix a graph $\mathcal{G} = (V, E)$ and a subgroup G of its automorphism group. The group G acts on $\mathbb{R}^{V\times V}$ by $(g.F)(x, y) = F(g^{-1}.x, g^{-1}.y)$. The program 2.3.5 is *invariant* under this action: this means that for all $g \in G$ and for all feasible solution F, $g.F$ is still feasible and it has the same objective value as F. It follows that 2.3.5 has an optimal solution which is G-invariant. In order to prove this, remark that starting from a feasible F, we can build its average over G,

$$\frac{1}{|G|}\sum_{g\in G} g.F \ ,$$

which is feasible, G-invariant and with the same objective value as F. It follows that the original program is equivalent to the one where the constraint that F is G-invariant is added. A G-invariant function satisfies $F(x, y) = F(g.x, g.y)$ for every $g \in G$ and so it takes constant values $f_1, \dots, f_m$ on the orbits $V_1, \dots, V_m$ of G on pairs of elements of V. Assume that $V_1 \cup \dots \cup V_\ell = \{(x, x) : x \in V\}$. Then the program 2.3.5 is equivalent to the following one:

$$\begin{aligned} \max\{ \ \textstyle\sum_{i=1}^m |V_i| f_i \ \ : \ \ & F \in (\mathbb{R}^{V\times V})^G, \ F \succeq 0 \\ & \textstyle\sum_{i=1}^{\ell} |V_i| f_i = 1 \\ & f_i = 0 \text{ if } V_i \subset E \qquad \} \end{aligned} \tag{2.4.1}$$

where $(\mathbb{R}^{V\times V})^G$ denotes the G-invariant real matrices. The space of G-invariant complex matrices $(\mathbb{C}^{V\times V})^G$ is a *matrix* $*$*-algebra*, i.e. a vector space which is closed under taking conjugate transpose and under multiplication. Note that $I \in (\mathbb{C}^{V\times V})^G$. Then the Artin-Wedderburn theorem says that $(\mathbb{C}^{V\times V})^G$ is isomorphic to a direct sum of matrix algebras

$$\begin{aligned} \varphi : (\mathbb{C}^{V\times V})^G \ &\longrightarrow \ \bigoplus_{k=1}^{d} \mathbb{C}^{m_k\times m_k} \\ F \ &\longrightarrow \ (F_k)_{k=1}^d \end{aligned}$$

The isomorphism φ preserves eigenvalues and hence positive semidefiniteness. The matrix $\varphi(F)$ is in block diagonal form, so

$$F \succeq 0 \ \Leftrightarrow \ \varphi(F) \succeq 0 \ \Leftrightarrow \ F_k \succeq 0 \ \forall \ k = 1, \dots, d \ .$$

The advantage of considering the matrices F_k is that it is computationally easier to solve a SDP program involving several smaller blocks than a SDP program involving one large matrix.

Now we sketch a general method to construct such an isomorphism. Consider the space $\mathbb{C}^V$ with the action of G given by $(g.f)(x) = f(g^{-1}.x)$ and the G-invariant inner product $\langle f, f' \rangle := \frac{1}{|V|} \sum_{x \in V} f(x)\overline{f'(x)}$. This gives a representation of G which we can decompose in irreducibles

$$\mathbb{C}^V = \left(H_{1,1} \perp \cdots \perp H_{1,m_1}\right) \perp \cdots \perp \left(H_{d,1} \perp \cdots \perp H_{d,m_d}\right)$$

where parentheses correspond to isotypic components, i.e. we have $H_{k,i} \simeq H_{k',i'}$ if and only if $k = k'$.

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 the complex numbers $\langle g.e_{k,i,s}, e_{k,i,t} \rangle$ do not depend on i (such a basis exists precisely because the G-isomorphism class of $H_{k,i}$ does not depend on i). From this data, we introduce the $m_k \times m_k$ matrices $E_k(x, y)$ with coefficients $E_{k,i,j}(x, y)$:

$$E_{k,i,j}(x, y) := \sum_{s=1}^{d_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)} \tag{2.4.2}$$

These matrices are sometimes called *zonal matrices* to say that they are G-invariant. Moreover, it can be proven that the matrices $E_{k,i,j}$ form a basis of the algebra of G-invariant matrices, hence

$$F \in (\mathbb{C}^{V \times V})^G \;\Leftrightarrow\; F(x, y) = \sum_k \langle F_k, \overline{E_k(x, y)} \rangle$$

and sending F to $(F_k)_{k=1}^d$ gives the announced isomorphism.

The theorem which characterizes positive definite G-invariant functions is known as *Bochner theorem*:

Theorem 2.4.1 ([11]). *A function $F \in \mathbb{C}^{V \times V}$ is G-invariant and positive semidefinite if and only if it can be written as*

$$F(x, y) = \sum_{k=1}^{d} \langle F_k, \overline{E_k(x, y)} \rangle \tag{2.4.3}$$

where $F_k \in \mathbb{C}^{m_k \times m_k}$ is hermitian positive semidefinite.

Note that the entries of each F_k can be written as the scalar product of F and $E_{k,i,j}$, so that they can be expressed as a function of the variables $y_i := |V_i| f_i$. Once this description is obtained, program 2.3.5 becomes

$$\max\Big\{ \sum_{i=1}^m y_i \;:\; \begin{array}{l} F_k \succeq 0 \;\forall\; k = 1, \ldots, d \\ \sum_{i=1}^{\ell} y_i = 1 \\ y_i = 0 \text{ if } V_i \subset E \end{array} \Big\} \tag{2.4.4}$$

The size of the program has been reduced from $|V|^2$ variables to $m = \sum_{k=1}^d (m_k)^2$.

Remark 2.4.2. Of course, in 2.4.4 we consider only the F_k such that the corresponding F has real entries. Whenever the group G has only real representations, we can start from decomposing $\mathbb{R}^V$ and work only in the real case.

2.4.2 Application to coding theory

A *code* $\mathcal{C}$ is a subset of a metric space (X, d). The *minimum distance* of $\mathcal{C}$, denoted $d(\mathcal{C})$ is the least value of pairwise distances of elements of $\mathcal{C}$. The classic framework is that of the (binary) *Hamming space* $H_n = \{0, 1\}^n$, i.e. the n-tuples of $0-1$ words, together with the *Hamming distance*

$$d(x, y) := |\{1 \leq i \leq n : x_i \neq y_i\}|$$

The Hamming weight of a word x is defined as $|x| := d(x, 0^n)$. A scalar product is defined on H_n by $x.y := \sum_{i=1}^n x_i y_i \mod 2$. The cardinality of the intersection of the support of x and the support of y is denoted $|x \cap y|$.

Intuitively, as the cardinality of a code grows, its minimum distance decreases. So it is interesting to study the number $\mathcal{A}(X, d) := \max\{|\mathcal{C}| : \mathcal{C} \subset X, d(\mathcal{C}) = d\}$, the maximal cardinality of a code of X with prescribed minimum distance d. The determination of this number is hard, so the goal is usually to find lower and upper bounds on it. While lower bounds mainly come from explicit constructions, upper bounds need to be set up by theoretical arguments. One of the best upper bounds for binary Hamming codes was established by Delsarte and it is obtained as the optimal value of a linear program.

Theorem 2.4.3. *(Delsarte linear programming upper bound, [14])*

$$\begin{aligned} \mathcal{A}(H_n, d) \quad \leq \sup \Big\{ 1 + \textstyle\sum_{i=1}^n x_i \quad : \quad & x_i \geq 0 \\ & x_1 = \cdots = x_{d-1} = 0 \\ & 1 + \textstyle\sum_{i=1}^n x_i K_k(i) \geq 0 \ \forall \ k = 0, \ldots, n \ \Big\} \end{aligned} \tag{2.4.5}$$

where K_k are the Krawtchouk polynomials defined by

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

Lovász ϑ' program applies to bound the maximal cardinality $\mathcal{A}(X, d)$ of codes in a metric space X with prescribed minimal distance d. Indeed $\mathcal{A}(X, d) = \alpha(\mathcal{G})$ where $\mathcal{G}$ is the graph with vertex set X and edges set $\{xy : 0 < d(x, y) < d\}$. So, we obtain

Theorem 2.4.4. *(The semidefinite programming bound for codes)*

$$\begin{aligned} \mathcal{A}(X, d) \quad \leq \sup \Big\{ \textstyle\sum_{X^2} F(x, y) \quad : \quad & F \in \mathbb{R}^{X \times X}, \ F \succeq 0, \ F \geq 0, \\ & \textstyle\sum_X F(x, x) = 1 \\ & F(x, y) = 0 \quad \text{if} \quad 0 < d(x, y) < d \ \Big\} \end{aligned} \tag{2.4.6}$$

$$\mathcal{A}(X,d) \le \inf\Big\{\ t/\lambda \ :\ F \in \mathbb{R}^{X\times X},\ F - \lambda \succeq 0 \qquad F(x,x) \le t \ \ \textit{for all } x \in X \qquad F(x,y) \le 0 \ \ \textit{if}\ \ d(x,y) \ge d \Big\} \tag{2.4.7}$$

The second inequality (2.4.7) is obtained with the dual formulation of the semidefinite program (2.4.6). It should be noted that any feasible solution of the semidefinite program in (2.4.7) leads to an upper bound for $\mathcal{A}(X,d)$.

Most of the spaces occurring in coding theory are of exponential size but they are endowed with group actions which allow to reduce the size of the SDP above. The Hamming space is endowed with the $\mathrm{Aut}(H_n) = T \rtimes S_n$ action, where $T \simeq (H_n,+)$ acts by translations and S_n by permuting the coordinates. It was first recognized in [36] and [40] that in the Hamming space, the program 2.4.6 symmetrized by the $T \rtimes S_n$ action coincides with the Delsarte linear program 2.4.5. As an application of the method introduced in section 2.4.1, we sketch the symmetrization process.

Let X denote H_n. The program 2.4.6 is invariant for the action of $G := T \rtimes S_n$ and thus we can restrict to G-invariant functions F. As the G-orbits of pairs are characterized by the pairwise distance, such F satisfies

$$F(x,y) = \tilde{F}(d(x,y))$$

and takes $n+1$ values $f_0, f_1, \dots, f_n$ on the respective orbits. Then we can immediately translate the program 2.4.6, except for the semidefiniteness constraint which we analyse separately. Indeed, the values of the function verify:

$$f_0 = \frac{1}{|X|} \qquad f_1 = \dots = f_{d-1} = 0 \qquad f_d, \dots, f_n \ge 0$$

and the objective function writes as

$$|X| f_0 + \sum_{i=1}^{n} |V_i| f_i = 1 + \sum_{i=1}^{n} |V_i| f_i$$

where $V_i := \{(x,y) \in X \times X : d(x,y) = i\}$.

The set of irreducible characters $\{\chi_z : z \in H_n\}$ is an orthonormal basis of $\mathbb{C}^X$. Remember that these characters are given by $\chi_z(x) := (-1)^{z.x}$. Putting together characters where z has the same weight, we obtain

$$\mathbb{C}^X = P_0 \oplus P_1 \oplus \dots \oplus P_n \tag{2.4.8}$$

where $P_k := \bigoplus_{|z|=k} \mathbb{C}\chi_z$ is G-irreducible of dimension $\binom{n}{k}$.

Next, we build the zonal functions. As the decomposition 2.4.8 is multiplicity free, the zonal functions are 1×1-matrices. They are given by:

$$E_k(x, y) := \sum_{|z|=k} \chi_z(x)\chi_z(y)$$

Denote $K_k(d(x, y)) := E_k(x, y)$. The function K_k is well defined and it can be proved that it is a polynomial of degree k. Explicitely:

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

Then we know, by theorem 2.4.1, that F is positive semidefinite if and only if it writes as a linear combination of the functions K_k with positive coefficients. These coefficients are given by

$$\frac{1}{|X|^2} \sum_{(x,y)\in X} F(x, y)K_k(d(x, y)) = \frac{1}{|X|^2} \sum_{i=0}^{n} |V_i| f_i K_k(i)$$

Now, calling $x_i = |V_i|f_i$ we find exactly program 2.4.5.

Indeed, it is a general fact that for *2-point homogeneous* spaces, the program ϑ' after symmetrization, reduces to a linear program where different families of orthogonal polynomials come into play. Here, a 2-point homogeneous space is a metric space (X, d) on which a group G acts transitively and such that the orbits on pairs are parametrized by the distance, i.e. $(x, y) \sim_G (x', y') \Leftrightarrow d(x, y) = d(x', y')$.

Apart from the Hamming space, another example of 2-point homogeneous space is given by the binary *Johnson* space J_n^w, defined as the set of $0 - 1$ words with Hamming weight w, together with the action of the symmetric group S_n.

2.4.3 Block diagonalization of the Terwilliger algebra of H_n

The reduction of the ϑ' program for the Hamming space amounts to block diagonalize the algebra $(\mathbb{C}^{H_n \times H_n})^{\mathrm{Aut}(H_n)}$. In this section we replace $\mathrm{Aut}(H_n)$ with one of its stabilizer subgroups, namely $S_n = \mathrm{Stab}_{\mathrm{Aut}(H_n)}(0^n)$. The algebra $(\mathbb{R}^{H_n \times H_n})^{S_n}$ coincides with the so called Terwilliger algebra of H_n. Its block diagonalization is originally contained in the paper of Schrijver, [41], where it is applied in order to obtain improvements on Delsarte LP bound. Here we adopt an equivalent point of view, but different from Schrijver's original one, which allows us to use the framework introduced in section 2.4.1. For notations and main results we follow [45] and the references therein. The decomposition 2.4.9 and the link with Hahn polynomials were proven by Delsarte in [15].

Take the notation $X := H_n$. The space X splits into the S_n-orbits $X_0, \dots, X_n$, where $X_w = J_n^w$ is the Johnson space. Hence

$$\mathbb{R}^X = \mathbb{R}^{X_0} \oplus \mathbb{R}^{X_1} \oplus \cdots \oplus \mathbb{R}^{X_n}$$

The irreducible modules for the action of the symmetric group on n elements are the *Specht modules* $\mathfrak{S}^\lambda$. They are indexed by partitions λ of n. The following decomposition is a classical fact:

$$\mathbb{R}^{X_w} = \begin{cases} H_{0,w} \oplus \cdots \oplus H_{w,w} & \text{if } 0 \leq w \leq \lfloor n/2 \rfloor \\ H_{0,w} \oplus \cdots \oplus H_{n-w,w} & \text{else} \end{cases} \tag{2.4.9}$$

where $H_{k,w} := \mathfrak{S}^{(n-k,k)}$ is of dimension $d_k := \binom{n}{k} - \binom{n}{k-1}$.

The isotypic components of $\mathbb{R}^X$ are

$$\mathcal{I}_k := H_{k,k} \oplus \cdots \oplus H_{k,n-k} \simeq H_{k,k}^{m_k}$$

for $k = 0, \dots, \lfloor n/2 \rfloor$. Here the multiplicity is given by $m_k := n - 2k + 1$. Now, let $\{e_{kk1}, \dots, e_{kkd_k}\}$ be an orthonormal basis of $H_{k,k}$. Applying the *valuation* operator

$$\begin{array}{llll} \psi_{k,i}: & \mathbb{R}^{X_k} & \longrightarrow & \mathbb{R}^{X_i} \\ & f & \longrightarrow & [\ x \to \sum\{f(y) : |y| = k, y \subset x\}\ ] \end{array}$$

we obtain an orthogonal basis $\{e_{ki1}, \dots, e_{kid_k}\}$ of $H_{k,i}$, whenever $i > k$ (see [15]).

To each isotypic component we associate the zonal matrix $E_k \in \mathbb{R}^{m_k \times m_k}$ as explained in 2.4.2:

$$E_{k,i,j}(x,y) := \frac{1}{|X|} \sum_{\ell=1}^{d_k} e_{ki\ell}(x) e_{kj\ell}(y)$$

The zonal matrices have an explicit description in terms of Hahn polynomials.

Definition 2.4.5. *The Hahn polynomials associated to the parameters n, i, j with $0 \leq i \leq j \leq n$ are the polynomials $Q_k(n,i,j;z)$ with $0 \leq k \leq \min(i, n-j)$ uniquely determined by the properties:*

- *Q_k has degree k in the variable z*
- *They are orthogonal polynomials for the weights*

$$0 \leq h \leq i \quad w(n,i,j;h) = \binom{i}{h}\binom{n-i}{j-i+h}$$

- *$Q_k(0) = 1$.*

Theorem 2.4.6. *If $k \leq i \leq j \leq n-k$, $|x| = i$, $|y| = j$, we have*

$$E_{k,i,j}(x,y) = |X| \frac{\binom{j-k}{i-k}\binom{n-2k}{j-k}}{\binom{n}{j}\binom{j}{i}} Q_k(n,i,j;i-|x\cap y|)$$

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

This explicit description along with Bochner theorem 2.4.1 gives a block diagonalization of the algebra of S_n-invariant matrices, together with a description of those who are positive semidefinite.

Chapter 3

Bounds for projective codes from semidefinite programming

3.1 Introduction

In network coding theory, as introduced in [3], some information is transmitted through a network (i.e. a directed graph), possibly having several sources and several receivers, and a certain number of intermediate nodes. Such information is modeled as vectors of fixed length over a finite field $\mathbb{F}_q$, called *packets*. To improve the performance of the communication, intermediate nodes should forward random linear $\mathbb{F}_q$-combinations of the packets they receive. This is the approach taken in the non-coherent communication case, that is, when the structure of the network is not known a priori [24]. Hence, if no errors occur, what is globally preserved all along the network is the *vector space* spanned by the packets injected at the sources. This observation led Koetter and Kschischang [29] to model network codes as subsets of the projective space $\mathcal{P}(\mathbb{F}_q^n)$, the set of linear subspaces of $\mathbb{F}_q^n$, or of the Grassmann space $\mathcal{G}_q(n,k)$, the subset of those subspaces of $\mathbb{F}_q^n$ having dimension k. Subsets of $\mathcal{P}(\mathbb{F}_q^n)$ are called *projective codes* while subsets of the Grassmann space will be referred to as *constant-dimension codes* or *Grassmann codes*.

As usual in coding theory, in order to protect the system from errors, it is desirable to select the elements of the code so that they are pairwise as far as possible with respect to a suitable distance. The *subspace distance* defined by

$$d_S(U,V) = \dim(U+V) - \dim(U\cap V) = \dim U + \dim V - 2\dim(U\cap V)$$

was introduced in [29] for this purpose. It is then natural to ask how large a code with a given minimal distance can be. Formally, we define

$$\begin{cases} A_q(n,d) := \max\{|\mathcal{C}| \ : \ \mathcal{C}\subset\mathcal{P}(\mathbb{F}_q^n),\ d_S(\mathcal{C})\geq d\} \\ A_q(n,k,2\delta) := \max\{|\mathcal{C}| \ : \ \mathcal{C}\subset\mathcal{G}_q(n,k),\ d_S(\mathcal{C})\geq 2\delta\} \end{cases}$$

where $d_S(\mathcal{C})$ denotes the minimal subspace distance among distinct elements of a code $\mathcal{C}$. In this chapter we will discuss upper estimates for these numbers.

The Grassmann space $\mathcal{G}_q(n,k)$ is a homogeneous space under the action of the linear group $GL_n(\mathbb{F}_q)$. Moreover, this group acts *distance transitively* when this space is endowed with the subspace distance. We mean here that the orbits of $GL_n(\mathbb{F}_q)$ acting on pairs of elements of $\mathcal{G}_q(n,k)$ are characterized by the value that this distance takes on them; in other words the Grassmann space is 2-point homogeneous under this action. Due to this property, codes and designs in this space can be analysed in the framework of Delsarte theory, in the same way as for other classical spaces in coding theory such as the Hamming space and the binary Johnson space. In fact, $\mathcal{G}_q(n,k)$ turns out to be a *q-analog* in the sense of combinatorics of the binary Johnson space as shown in [15]. The linear group plays the role of the symmetric group for the Johnson space, while the dimension replaces the weight function.

Unsurprisingly, the classical bounds (anticode, Hamming, Johnson, Singleton) have been derived for the Grassmann codes [29, 48, 49]. The more sophisticated Delsarte linear programming bound was already obtained in [15]; however numerical computations indicate that it is not better than the anticode bound. Moreover, the Singleton and anticode bounds have the same asymptotic which is attained by a family of Reed-Solomon like codes built in [29] and closely connected to the rank-metric Gabidulin codes.

In contrast, the projective space has a much nastier behaviour, essentially because it is not 2-point homogeneous, in fact not even homogeneous under the action of a group. For example, the balls in this space have a size that depends not only on their radius, but also on the dimension of their center. Consequently, bounds for projective codes are much harder to settle. Etzion and Vardy in [17] provide a bound in the form of the optimal value of a linear program, which is derived by elementary reasoning involving packing issues. The Etzion-Vardy bound is the only successful generalization of the classical bounds to the projective space.

In this chapter we settle semidefinite programming bounds for projective codes and we compare them with the above mentioned bounds. For the projective space, the symmetrization was announced in [9] (see also [5]). The program remains a semidefinite program but fortunately with polynomial size.

This chapter is organized as follows. In Section 2 we review the classical bounds for constant dimension codes and the Etzion-Vardy bound for projective codes. In Section 3 we show that most of the bounds for Grassmann codes can be derived from the semidefinite programming method. In Section 4 we reduce the ϑ program by the action of the group $GL_n(\mathbb{F}_q)$. In Section 5 we present some numerical results obtained with this method and we compare them with the Etzion-Vardy method for $q = 2$ and $n \leq 16$. Another distance of interest on the projective space, the so-called *injection distance* was introduced in [30]. We show how to adapt the Etzion-Vardy bound as well as the semidefinite programming bound to this case.

3.2 Elementary bounds for Grassmann and projective codes

3.2.1 Bounds for Grassmann codes

In this section we review the classical bounds for $A_q(n, k, 2\delta)$. We note that the subspace distance takes only even values on the Grassmann space and that one can restrict to $k \leq n/2$ by the relation $A_q(n, k, 2\delta) = A_q(n, n-k, 2\delta)$, which follows by considering orthogonal subspaces.

We recall the definition of the q-analog of the binomial coefficient that counts the number of k-dimensional subspaces of a fixed n-dimensional space over $\mathbb{F}_q$, i.e. the number of elements of $\mathcal{G}_q(n, k)$.

Definition 3.2.1. *The q-ary binomial coefficient is defined by*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1) \dots (q - 1)}$$

The sphere-packing bound [29]

$$A_q(n, k, 2\delta) \leq \frac{|\mathcal{G}_q(n, k)|}{|B_k(\delta - 1)|} = \frac{\begin{bmatrix} n \\ k \end{bmatrix}_q}{\sum_{m=0}^{\lfloor(\delta-1)/2\rfloor} \begin{bmatrix} k \\ m \end{bmatrix}_q \begin{bmatrix} n-k \\ m \end{bmatrix}_q q^{m^2}}. \tag{3.2.1}$$

It follows as usual from the observation that the balls of radius $\delta - 1$ centered at the elements of a code $\mathcal{C} \subset \mathcal{G}_q(n, k)$ with minimal distance 2δ, are pairwise disjoint, and have the same cardinality $\sum_{m=0}^{\lfloor(\delta-1)/2\rfloor} \begin{bmatrix} k \\ m \end{bmatrix}_q \begin{bmatrix} n-k \\ m \end{bmatrix}_q q^{m^2}$.

The Singleton bound [29]

$$A_q(n, k, 2\delta) \leq \begin{bmatrix} n - \delta + 1 \\ k - \delta + 1 \end{bmatrix}_q. \tag{3.2.2}$$

It is obtained by the introduction of a "puncturing" operation on the code.

The anticode bound [48]

An *anticode* of diameter e is a subset of a metric space whose pairwise distinct elements are at distance less or equal than e. The general anticode bound ([14]) states that, given a metric space X which is homogeneous under the action of a group G, for every code $\mathcal{C} \subset X$ with minimal distance d and for every anticode $\mathcal{A} \in X$ of diameter $d-1$, we have

$$|\mathcal{C}| \leq \frac{|X|}{|\mathcal{A}|} .$$

Spheres of given radius r are one example of anticodes of diameter $2r$: indeed if we take $\mathcal{A}$ to be a sphere of radius $\delta-1$ in $\mathcal{G}_q(n,k)$, we recover the sphere-packing bound. Obviously, to obtain the strongest bound, we have to choose the largest anticodes of given diameter, which in our case are not the spheres. The largest anticode $\mathcal{A}$ of diameter $2\delta-2$ in the finite Grassmannian was described in [20]: it consists of all elements of $\mathcal{G}_q(n,k)$ which contain a fixed $(\delta-1)$-dimensional subspace. Taking such $\mathcal{A}$ in the general anticode bound, we recover *the* (best) anticode bound for $\mathcal{G}_q(n,k)$:

$$A_q(n,k,2\delta) \leq \frac{\begin{bmatrix} n \\ k-\delta+1 \end{bmatrix}_q}{\begin{bmatrix} k \\ k-\delta+1 \end{bmatrix}_q} = \frac{(q^n-1)(q^{n-1}-1)\dots(q^{n-k+\delta}-1)}{(q^k-1)(q^{k-1}-1)\dots(q^{\delta}-1)} \tag{3.2.3}$$

It follows from the previous discussion that the anticode bound improves the sphere-packing bound. Moreover, the anticode bound is usually stronger than the Singleton bound, with equality only in the cases $n=k$ or $\delta=1$ ([49]).

The first and second Johnson-type bound [49]

$$A_q(n,k,2\delta) \leq \left\lfloor \frac{(q^n-1)(q^k-q^{k-\delta})}{(q^k-1)^2-(q^n-1)(q^{k-\delta}-1)} \right\rfloor \tag{3.2.4}$$

as long as $(q^k-1)^2-(q^n-1)(q^{k-\delta}-1) > 0$, and

$$A_q(n,k,2\delta) \leq \left\lfloor \frac{q^n-1}{q^k-1} A_q(n-1,k-1,2\delta) \right\rfloor . \tag{3.2.5}$$

These bounds were obtained in [49] through the construction of a binary constant-weight code associated to every constant-dimension code. Iterating the latter, one obtains

$$A_q(n,k,2\delta) \leq \left\lfloor \frac{q^n-1}{q^k-1} \left\lfloor \frac{q^{n-1}-1}{q^{k-1}-1} \dots \left\lfloor \frac{q^{n-k+\delta}-1}{q^{\delta}-1} \right\rfloor \dots \right\rfloor \right\rfloor . \tag{3.2.6}$$

If the floors are removed from the right hand side of (3.2.6), the anticode bound is recovered, so (3.2.6) is stronger. In the particular case of $\delta=k$ and if $n \not\equiv 0 \bmod k$, (3.2.6) was sharpened in [17] to:

$$A_q(n,k,2k) \leq \left\lfloor \frac{q^n-1}{q^k-1} \right\rfloor - 1 \tag{3.2.7}$$

In contrast, for $\delta = k$ and if k divides n, we have equality in (3.2.6), because of the existence of *spreads* (see [17]):

$$A_q(n,k,2k) = \frac{q^n - 1}{q^k - 1}$$

Summing up, the strongest upper bound for constant dimension codes reviewed so far comes by putting together (3.2.6) and (3.2.7):

Theorem 3.2.2. *If $n - k \not\equiv 0 \mod \delta$, then*

$$A_q(n,k,2\delta) \leq \left\lfloor \frac{q^n - 1}{q^k - 1} \left\lfloor \dots \left\lfloor \frac{q^{n-k+\delta+1} - 1}{q^{\delta+1} - 1} \left(\left\lfloor \frac{q^{n-k+\delta} - 1}{q^\delta - 1} \right\rfloor - 1 \right) \right\rfloor \dots \right\rfloor \right\rfloor$$

otherwise

$$A_q(n,k,2\delta) \leq \left\lfloor \frac{q^n - 1}{q^k - 1} \left\lfloor \dots \left\lfloor \frac{q^{n-k+\delta+1} - 1}{q^{\delta+1} - 1} \left\lfloor \frac{q^{n-k+\delta} - 1}{q^\delta - 1} \right\rfloor \right\rfloor \dots \right\rfloor \right\rfloor.$$

3.2.2 A bound for projective codes

Here we turn our attention to codes whose codewords have not necessarily the same dimension, and we review the bound obtained by Etzion and Vardy in [17] for these codes. The idea is to split a code $\mathcal{C}$ into subcodes $\mathcal{C}_k = \mathcal{C} \cap \mathcal{G}_q(n,k)$ of constant dimension, and then to design linear constraints on the cardinality D_k of each $\mathcal{C}_k$, coming from packing constraints.

Let $B(V,e) := \{U \in \mathcal{P}(\mathbb{F}_q^n) : d_S(U,V) \leq e\}$ denote the ball of center V and radius e. If $\dim V = i$, we have that

$$|B(V,e)| = \sum_{k=0}^{e} \sum_{j=0}^{k} \begin{bmatrix} i \\ j \end{bmatrix}_q \begin{bmatrix} n-i \\ k-j \end{bmatrix}_q q^{j(k-j)}$$

We define $c(i,k,e) := |B(V,e) \cap \mathcal{G}_q(n,k)|$ for V of dimension i. It is not difficult to prove that

$$c(i,k,e) = \sum_{j=\lceil \frac{i+k-e}{2} \rceil}^{\min\{k,i\}} \begin{bmatrix} i \\ j \end{bmatrix}_q \begin{bmatrix} n-i \\ k-j \end{bmatrix}_q q^{(i-j)(k-j)}. \qquad (3.2.8)$$

Theorem 3.2.3 (Linear programming upper bound for codes in $\mathcal{P}(\mathbb{F}_q^n)$, [17])**.**

$$A_q(n, 2e+1) \leq \sup \left\{ \sum_{k=0}^{n} x_k \ : \ x_k \leq A_q(n,k,2e+2) \ \forall \ k = 0,\dots,n \right.$$
$$\left. \sum_{i=0}^{n} c(i,k,e) x_i \leq \begin{bmatrix} n \\ k \end{bmatrix}_q \ \forall \ k = 0,\dots,n \right\}$$

Proof. For $\mathcal{C} \subset \mathcal{P}(\mathbb{F}_q^n)$ of minimal distance $2e+1$, define $D_0, \dots, D_n$ by $D_k = |\mathcal{C} \cap \mathcal{G}_q(n,k)|$. Then $\sum_{k=0}^n D_k = |\mathcal{C}|$ and each D_k represents the cardinality of a subcode of $\mathcal{C}$ of constant dimension k, so it is upper bounded by $A_q(n,k,2e+2)$. Moreover, balls of radius e centered at the codewords are pairwise disjoint, so the sets $B(V,e) \cap \mathcal{G}_q(n,k)$ for $V \in \mathcal{C}$ are pairwise disjoint subsets of $\mathcal{G}_q(n,k)$. So

$$\sum_{V \in \mathcal{C}} |B(V,e) \cap \mathcal{G}_q(n,k)| \le |\mathcal{G}_q(n,k)|.$$

Because the number of elements of $B(V,e) \cap \mathcal{G}_q(n,k)$ only depends on $i = \dim(V)$, k and e and equals $c(i,k,e)$ by definition, we obtain the second constraint

$$\sum_{i=0}^n c(i,k,e) D_i \le \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

So $|\mathcal{C}|$ is upper bounded by the optimal value of the announced linear program in real variables x_i. □

Remark 3.2.4. Of course, in view of explicit computations, if the exact value of $A_q(n,k,2e+2)$ is not available, it can be replaced in the linear program of Theorem 3.2.3 by an upper bound.

3.3 The semidefinite programming method on the Grassmannian

The Grassmann space $\mathcal{G}_q(n,k)$ is 2-point homogeneous for the action of $G = GL_n(q)$ and its associated zonal polynomials are computed in [15]. They belong to the family of q-Hahn polynomials, which are q-analogs of the Hahn polynomials related to the binary Johnson space and introduced in definition 2.4.5.

Definition 3.3.1. *The q-Hahn polynomials associated to the parameters n, s, t with $0 \le s \le t \le n$ are the polynomials $Q_k(n,s,t;u)$ with $0 \le k \le \min(s, n-t)$ uniquely determined by the properties:*

- *Q_k has degree k in the variable $[u] = q^{1-u} \begin{bmatrix} u \\ 1 \end{bmatrix}_q$*
- *They are orthogonal polynomials for the weights*

$$0 \le i \le s \quad w(n,s,t;i) = \begin{bmatrix} s \\ i \end{bmatrix}_q \begin{bmatrix} n-s \\ t-s+i \end{bmatrix}_q q^{i(t-s+i)}$$

- *$Q_k(0) = 1$.*

To be more precise, in the Grassmann space $\mathcal{G}_q(n,k)$, the zonal polynomials are associated to the parameters $s = t = k$. The other parameters will come into play when we will analyse the full projective space in Section 4. The resulting linear programming bound is explicitly stated in [15]:

Theorem 3.3.2. *(Delsarte's linear programming bound [15]).*

$$A_q(n,k,2\delta) \leq \inf\Big\{1 + f_1 + \cdots + f_k \ : \ f_i \geq 0 \ \forall\ i = 1,\dots,k \\ F(u) \leq 0 \ \forall\ u = \delta,\dots,k\Big\}$$

where $F(u) = 1 + \sum_{i=1}^{k} f_i Q_i(u)$ *and* $Q_i(u) = Q_i(n,k,k;u)$ *as in Definition 3.3.1.*

3.3.1 Bounds for Grassmann codes

In order to show the power of the semidefinite programming bound, we will verify that most of the bounds recalled in section 3.2 for Grassmann codes can be obtained from Corollary 2.4.4 or Theorem 3.3.2. To that end, in each case we construct a suitable feasible solution of (2.4.7).

The Singleton bound.

The function

$$F(x,y) = \sum_{\dim(z)=n-\delta+1} \ \sum_{\substack{w \subset z, \\ \dim(w)=k-\delta+1}} \mathbf{1}(z \cap x = w)\mathbf{1}(z \cap y = w)$$

where $x \to \mathbf{1}(z \cap x = w)$ denotes the characteristic function of the set $\{x \in \mathcal{G}_q(n,k) \ : \ z \cap x = w\}$, is obviously positive semidefinite. One can verify that F is a feasible solution of (2.4.7) and leads to the Singleton bound (3.2.2).

The sphere-packing and anticode bounds.

The sphere-packing bound and the anticode bound in $\mathcal{G}_q(n,k)$ can also be obtained directly, with the functions

$$F(x,y) = \sum_{\dim(z)=k} \mathbf{1}_{B(z,\delta-1)}(x)\mathbf{1}_{B(z,\delta-1)}(y)$$

and

$$F(x,y) = \sum_{\dim(z)=k-\delta+1} \mathbf{1}(z \subset x)\mathbf{1}(z \subset y)\ .$$

In general, the anticode bound $|\mathcal{C}| \leq |X|/|\mathcal{A}|$ can be derived from (2.4.7), using the function $F(x,y) = \sum_{g \in G} \mathbf{1}_{\mathcal{A}}(g.x)\mathbf{1}_{\mathcal{A}}(g.y)$.

Remark 3.3.3. An interesting open point is to prove whether or not the anticode bound coincides with the Delsarte bound for constant dimension codes. We already proved one inequality and in our numerical computations the two bounds give the same value for several different parameters.

The first Johnson-type bound.

We want to apply Delsarte's linear programming bound of Theorem 3.3.2 with a function F of degree 1, i.e. $F(u) = f_0 Q_0(u) + f_1 Q_1(u)$. According to [15] the first q-Hahn polynomials are

$$Q_0(u) = 1 \quad , \quad Q_1(u) = \left(1 - \frac{(q^n-1)(1-q^{-u})}{(q^k-1)(q^{n-k}-1)}\right).$$

In order to construct a feasible solution of the linear program, we need a pair (f_0, f_1) of positive numbers for which $F(u) = f_0 + f_1 Q_1(u)$ is non-positive for $u = \delta, \ldots, k$. Then $1 + f_1/f_0$ will be an upper bound for $A_q(n, k, 2\delta)$. As $Q_1(u)$ is decreasing, the optimal choice of (f_0, f_1) satisfies $F(\delta) = 0$. So $f_1/f_0 = -1/Q_1(\delta)$ and we need $Q_1(\delta) < 0$. We obtain (3.2.4):

$$A_q(n, k, 2\delta) \leq 1 + \frac{f_1}{f_0} = 1 - \frac{1}{Q_1(\delta)} = \frac{(q^n-1)(q^k - q^{k-\delta})}{(q^k-1)^2 - (q^n-1)(q^{k-\delta}-1)}.$$

The second Johnson-type bound.

Here we recover an inequality on the optimal value $B_q(n, k, 2\delta)$ of (2.4.7) in the case $X = \mathcal{G}_q(n, k)$ (with the subspace distance) which is similar to (3.2.5):

$$B_q(n, k, 2\delta) \leq \frac{q^n-1}{q^k-1} B_q(n-1, k-1, 2\delta).$$

Let (F, t, λ) be an optimal solution for the program (2.4.7) in $\mathcal{G}_q(n-1, k-1)$ relative to the minimal distance 2δ, i.e. F satisfies the conditions: $F \succeq \lambda$, $F(x, x) \leq t$, $F(x, y) \leq 0$ if $d(x, y) \geq 2\delta$, and moreover $t/\lambda = B_q(n-1, k-1, 2\delta)$.

We consider the function G on $\mathcal{G}_q(n, k) \times \mathcal{G}_q(n, k)$ given by

$$G(x, y) = \sum_{|D|=1} \mathbf{1}(D \subset x)\mathbf{1}(D \subset y) F(x \cap H_D, y \cap H_D)$$

where, for every one-dimensional space D, H_D is an arbitrary hyperplane such that $D \oplus H_D = \mathbb{F}_q^n$. It can be verified that the function G is a feasible solution of the program (2.4.7) in $\mathcal{G}_q(n, k)$ for the minimal distance 2δ, and we obtain that

$$B_q(n, k, 2\delta) \leq \frac{t}{\lambda} \frac{q^n-1}{q^k-1} = \frac{q^n-1}{q^k-1} B_q(n-1, k-1, 2\delta) \ .$$

Remark 3.3.4. in [17], another Johnson-type bound is given:

$$A_q(n,k,2\delta) \leq \frac{q^n-1}{q^{n-k}-1} A_q(n-1,k,2\delta) \ .$$

which follows easily from the second Johnson-type bound combined with the equality $A_q(n,k,2\delta) = A_q(n,n-k,2\delta)$. Similarly as above, one can show that an analogous inequality holds for the semidefinite programming bounds.

3.4 Semidefinite programming bounds for projective codes

In this section we reduce the semidefinite programs (2.4.6), (2.4.7) for the entire projective space, using the action of the group $G = GL_n(\mathbb{F}_q)$. We follow the general method described in section 2.4.1. As we saw, the semidefinite program can be restricted to G-invariant functions F. It remains to obtain an explicit description of the G-invariant positive semidefinite functions on the projective space. The projective space endowed with the G action is the q-analog of the Hamming space where the action is restricted to S_n. So the symmetrization process is analog to the one given in section 2.4.3. Indeed, in [15], Delsarte treats the two cases in a common framework.

3.4.1 The G-invariant positive semidefinite functions on the projective space

In order to obtain a suitable expression for these functions, we exploit the decomposition of the space of real-valued functions under the action of G. We take the following notations: let $X = \mathcal{P}(\mathbb{F}_q^n)$, $X_k = \mathcal{G}_q(n,k)$ and let $\mathbb{R}^X = \{f : X \to \mathbb{R}\}$. In the paper [15], Delsarte showed that the irreducible decomposition of the $\mathbb{R}^{X_k}$ under the action of G is given by the *harmonic subspaces* $H_{k,i}$:

$$\mathbb{R}^{X_k} = H_{0,k} \oplus H_{1,k} \oplus \cdots \oplus H_{\min\{k,n-k\},k} \tag{3.4.1}$$

Here, $H_{k,k}$ is the kernel of the *differentiation operator*

$$\begin{array}{llll} \delta_k : & \mathbb{R}^{X_k} & \longrightarrow & \mathbb{R}^{X_{k-1}} \\ & f & \longrightarrow & [\ x \to \sum\{f(y) : |y| = k, x \subset y\}\ ] \end{array}$$

and $H_{k,i}$ is the image of $H_{k,k}$ under the *valuation operator*

$$\begin{array}{llll} \psi_{ki} : & \mathbb{R}^{X_k} & \longrightarrow & \mathbb{R}^{X_i} \\ & f & \longrightarrow & [\ x \to \sum\{f(y) : |y| = k, y \subset x\}\ ] \end{array}$$

for $k \leq i \leq n-k$. Because δ_k is surjective, we have $h_k := \dim(H_{k,k}) = \begin{bmatrix} n \\ k \end{bmatrix}_q - \begin{bmatrix} n \\ k-1 \end{bmatrix}_q$. Moreover, ψ_{ki} commutes with the action of G so $H_{k,i}$ is isomorphic to $H_{k,k}$. Putting together the spaces $\mathbb{R}^{X_k}$ one gets the global picture:

$$\begin{array}{lcccccccccccccc}
\mathbb{R}^X & = & \mathbb{R}^{X_0} & \oplus & \mathbb{R}^{X_1} & \oplus & \dots & \oplus & \mathbb{R}^{X_{\lfloor \frac{n}{2} \rfloor}} & \oplus & \dots & \oplus & \mathbb{R}^{X_{n-1}} & \oplus & \mathbb{R}^{X_n} \\
 & & & & & & & & & & & & & & \\
\mathcal{I}_0 & = & H_{0,0} & \oplus & H_{0,1} & \oplus & \dots & & \dots & & \dots & \oplus & H_{0,(n-1)} & \oplus & H_{0,n} \\
\mathcal{I}_1 & = & & & H_{1,1} & \oplus & \dots & & \dots & & \dots & \oplus & H_{1,(n-1)} & & \\
\mathcal{I}_2 & = & & & & & \dots & & \dots & & \dots & & & & \\
\vdots & & & & & & \ddots & & & & & & & & \\
\vdots & & & & & & & \ddots & & & & & & & \\
\mathcal{I}_{\lfloor \frac{n}{2} \rfloor} & = & & & & & & & H_{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor} & & & & & &
\end{array}$$

Here, the columns give the irreducible decomposition (3.4.1) of the spaces $\mathbb{R}^{X_k}$. The irreducible components which lie in the same row are all isomorphic, and together they form the isotypic components

$$\mathcal{I}_k := H_{k,k} \oplus H_{k,k+1} \oplus \cdots \oplus H_{k,n-k} \simeq H_{k,k}^{n-2k+1} \ .$$

Starting from this decomposition, we build the zonal matrices $E_k(x,y)$ [6, section 3.3] in the following way. We take an isotypic component $\mathcal{I}_k$ and we fix an orthonormal basis $(e_{kk1}, \dots, e_{kkh_k})$ of $H_{k,k}$. Applying the valuation operator ψ_{ki}, we get an orthogonal basis $(e_{ki1}, \dots, e_{kih_k})$ of $H_{k,i}$. Then we define

$$E_{kst}(x,y) = \frac{1}{h_k} \sum_{i=1}^{h_k} e_{ksi}(x) e_{kti}(y) \ .$$

Theorem 2.4.1 gives the expression of the G-invariant positive semidefinite functions in terms of the matrices $E_k(x,y)$:

Theorem 3.4.1. *The function $F \in \mathbb{R}^{X \times X}$ is positive semidefinite and G-invariant if and only if it can be written as*

$$F(x,y) = \sum_{k=0}^{\lfloor n/2 \rfloor} \langle F_k, E_k(x,y) \rangle \qquad (3.4.2)$$

where $F_k \in \mathbb{R}^{(n-2k+1)\times(n-2k+1)}$ and F_k is positive semidefinite.

Now we need to compute the E_k's. They are zonal matrices: in other words, for all $k \leq s,t \leq n-k$, for all $g \in G, E_{kst}(x,y) = E_{kst}(gx,gy)$. This means that

E_{kst} is actually a function of the variables which parametrize the orbits of G on $X \times X$. It is easy to see that the orbit of the pair (x, y) is characterized by the triple $(\dim(x), \dim(y), \dim(x \cap y))$.

The next theorem gives an explicit expression of E_{kst}, in terms of the polynomials Q_k of Definition 3.3.1.

Theorem 3.4.2. *If $k \le s \le t \le n-k$, $\dim(x) = s$, $\dim(y) = t$,*

$$E_{kst}(x, y) = |X| \frac{\begin{bmatrix} t-k \\ s-k \end{bmatrix}_q \begin{bmatrix} n-2k \\ t-k \end{bmatrix}_q}{\begin{bmatrix} n \\ t \end{bmatrix}_q \begin{bmatrix} t \\ s \end{bmatrix}_q} q^{k(t-k)} Q_k(n, s, t; s - \dim(x \cap y))$$

If $\dim(x) \neq s$ or $\dim(y) \neq t$, $E_{kst}(x, y) = 0$.

We note that the weights involved in the orthogonality relations of the polynomials Q_k have a combinatorial meaning:

Lemma 3.4.3. *[16] Given $x \in X_s$, the number of elements $y \in X_t$ such that $\dim(x \cap y) = s - i$ is equal to $w(n, s, t; i)$.*

Proof. (of Theorem 3.4.2) By construction, $E_{kst}(x, y) \neq 0$ only if $\dim(x) = s$ and $\dim(y) = t$, so in this case E_{kst} is a function of $(s - \dim(x \cap y))$. Accordingly, for $k \le s \le t \le n-k$, we introduce $P_{k,s,t}$ such that $E_{k,s,t}(x, y) = P_{k,s,t}(s - \dim(x \cap y))$. Now we want to relate $P_{k,s,t}$ to the q-Hahn polynomials. We start with two lemmas: one obtains the orthogonality relations satisfied by $P_{k,s,t}$ and the other computes $P_{k,s,t}(0)$.

Lemma 3.4.4. *With the above notations,*

$$P_{k,s,t}(0) = \dim(x) \frac{\begin{bmatrix} t-k \\ s-k \end{bmatrix}_q \begin{bmatrix} n-2k \\ t-k \end{bmatrix}_q}{\begin{bmatrix} n \\ t \end{bmatrix}_q \begin{bmatrix} t \\ s \end{bmatrix}_q} q^{k(t-k)}. \tag{3.4.3}$$

Proof. We have $P_{k,s,t}(0) = E_{k,s,t}(x, y)$ for all x, y with $\dim(x) = s$, $\dim(y) = t$, $x \subset y$.

Hence

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

With the relation $\psi_{s,t} \circ \psi_{k,s} = \begin{bmatrix} t-k \\ s-k \end{bmatrix}_q \psi_{k,t}$,

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

and we obtain

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

where the last equality follows from [15, Theorem 3]. □

Lemma 3.4.5. *With the above notations,*

$$
\sum_{i=0}^{s} w(n,s,t;i) P_{k,s,t}(i) P_{l,s,t}(i) = \delta_{k,l} |X|^2 \frac{\begin{bmatrix} n-2k \\ s-k \end{bmatrix}_q \begin{bmatrix} n-2k \\ t-k \end{bmatrix}_q q^{k(s+t-2k)}}{\begin{bmatrix} n \\ s \end{bmatrix}_q h_k}. \tag{3.4.4}
$$

Proof. We compute $\Sigma := \sum_{y\in X} E_{k,s,t}(x,y) E_{l,s',t'}(y,z)$.

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

We obtain, with $t=s'$, $t'=s$, $x=z\in X_s$, taking account of $E_{l,t,s}(y,x)=E_{l,s,t}(x,y)$,

$$
\sum_{y\in X_t} E_{k,s,t}(x,y)E_{l,s,t}(x,y)=\delta_{k,l}|X|\frac{\left[\begin{smallmatrix} n-2k\\ t-k\end{smallmatrix}\right]_q q^{k(t-k)}}{h_k}E_{k,s,s}(x,x).
$$

The above identity becomes in terms of $P_{k,s,t}$

$$
\sum_{y\in X_t} P_{k,s,t}(s-\dim(x\cap y))P_{l,s,t}(s-\dim(x\cap y))=\delta_{k,l}|X|\frac{\left[\begin{smallmatrix} n-2k\\ t-k\end{smallmatrix}\right]_q q^{k(t-k)}}{h_k}P_{k,s,s}(0).
$$

Taking account of (3.4.3) and Lemma 3.4.3, we obtain (3.4.4). □

We have proved that the functions $P_{k,s,t}$ satisfy the same orthogonality relations as the q-Hahn polynomials. So we are done if $P_{k,s,t}$ is a polynomial of degree at most k in the variable $[u]=[\dim(x\cap y)]$. This property is proved in the case $s=t$ in [15, Theorem 5] and extends to $s\le t$ with a similar line of reasoning. The multiplicative factor between $P_{k,s,t}(u)$ and $Q_k(n,s,t;u)$ is then given by $P_{k,s,t}(0)$ and the proof of Theorem 3.4.2 is completed. □

3.4.2 The reduction of the program (2.4.6) for projective codes

A function $F\in\mathbb{R}^{X\times X}$ is G-invariant if it can be written as

$$
F(x,y)=\tilde{F}(\dim(x),\dim(y),\dim(x\cap y))
$$

So we introduce a function $\tilde{F}$ with the property that $\tilde{F}(s,t,i) = F(x,y)$ for $x, y \in X$ with $\dim(x) = s, \dim(y) = t, \dim(x \cap y) = i$. Let

$$N_{sti} := |\{(x,y) \in X \times X : \dim(x) = s, \dim(y) = t, \dim(x \cap y) = i\}|$$

and

$$\Omega(d) := \{(s,t,i) : 0 \leq s,t \leq n,\ i \leq \min(s,t),\ s+t \leq n+i, \\ \text{either } s = t = i \text{ or } s+t-2i \geq d\}. \tag{3.4.5}$$

Then, (2.4.6) becomes:

$$A_q(n,d) \leq \sup\Big\{ \sum_{s,t,i} N_{sti}\tilde{F}(s,t,i) \quad : \quad \tilde{F} \in \mathbb{R}^{[n]^3},\ \tilde{F} \succeq 0,\ \tilde{F} \geq 0, \\ \sum_{s=0}^{n} N_{sss}\tilde{F}(s,s,s) = 1, \\ \tilde{F}(s,t,i) = 0 \text{ if } (s,t,i) \notin \Omega(d) \Big\}$$

where, of course, $\tilde{F} \succeq 0$ means that the corresponding F is positive semidefinite.

Then, we introduce the variables $x_{sti} := N_{sti}\tilde{F}(s,t,i)$. It is straightforward to rewrite the program in terms of these variables, except for the condition $\tilde{F} \succeq 0$. From Theorem 4.1, this is equivalent to the semidefinite conditions $F_k \succeq 0$, where the matrices F_k are given by the scalar product of F and E_k:

$$(F_k)_{lj} = \frac{1}{|X|^2} \sum_{X^2} F(x,y)E_k(x,y)_{lj} = \frac{1}{|X|^2} \sum_{s,t,i} x_{sti}\tilde{E}_k(s,t,i)_{lj}.$$

We can substitute the value of $\tilde{E}_k(s,t,i)_{lj}$ using Theorem 4.2; in particular it is 0 when $(l,j) \neq (s,t)$, and, when $(l,j) = (s,t)$:

$$(F_k)_{st} = \frac{1}{|X|} \sum_{i} x_{sti} \frac{\begin{bmatrix} t-k \\ s-k \end{bmatrix}_q \begin{bmatrix} n-2k \\ t-k \end{bmatrix}_q}{\begin{bmatrix} n \\ t \end{bmatrix}_q \begin{bmatrix} t \\ s \end{bmatrix}_q} q^{k(t-k)} Q_k(n,s,t;s-i). \tag{3.4.6}$$

Finally we obtain:

Theorem 3.4.6.

$$A_q(n,d) \leq \sup\Big\{ \sum_{(s,t,i)\in\Omega(d)} x_{sti} \quad : \quad (x_{sti})_{(s,t,i)\in\Omega(d)},\ x_{sti} \geq 0, \\ \sum_{s=0}^{n} x_{sss} = 1, \\ F_k \succeq 0 \ \ \textit{for all } k = 0,\ldots,\lfloor n/2 \rfloor \Big\}$$

where $\Omega(d)$ is defined in (3.4.5) *and the matrices F_k are given in (3.4.6).*

Remark 3.4.7. A projective code $\mathcal{C}$ with minimal distance d provides a feasible solution of the above program, given by:

$$x_{sti} = \frac{1}{|\mathcal{C}|}|\{(x, y) \in \mathcal{C} : \dim(x) = s, \dim(y) = t, \dim(x \cap y) = i\}.$$

In particular, we have

$$\sum_{t,i} x_{sti} = |\mathcal{C} \cap \mathcal{G}_q(n, s)|,$$

so, if $A_q(n, s, d)$ (or an upper bound) is known, the additional inequality

$$\sum_{t,i} x_{sti} \leq A_q(n, s, d)$$

can be added to the semidefinite program of Theorem 3.4.6 in order to tighten the upper bound for $A_q(n, d)$.

3.5 Numerical results

Here we give the explicit results obtained for the upper bounds on $A_2(n, d)$, hence for $q = 2$. In Table 1, we consider the subspace distance d_S while Table 2 is related to the injection distance d_i recently introduced in [30].

3.5.1 The subspace distance

The first column of Table 3.1 displays the upper bound obtained from Etzion-Vardy linear program recalled in Theorem 3.2.3. The second column contains the upper bound from the semidefinite program of Theorem 3.4.6, strengthened by the inequalities (see Remark 3.4.7):

$$\sum_{t,i} x_{sti} \leq A_2(n, s, 2\lceil d/2 \rceil) \quad \text{for all } s = 0, \ldots, n.$$

In both programs, $A_2(n, k, 2\delta)$ is replaced by its upper bound in Theorem 3.2.2.

3.5.2 Some additional inequalities

In [17], Etzion and Vardy have found additional inequalities on the unknowns of their linear program in the specific case of $n = 5$ and $d = 3$. With this, they could improve their bound to the exact value $A_2(5, 3) = 18$. In this section we establish analogous inequalities for other values of the parameters (n, d).

	E-V LP	SDP
$A_2(4,3)$	6	6
$A_2(5,3)$	20	20
$A_2(6,3)$	124	124
$A_2(7,3)$	832	776
$A_2(7,5)$	36	35
$A_2(8,3)$	9365	9268
$A_2(8,5)$	361	360
$A_2(9,3)$	114387	107419
$A_2(9,5)$	2531	2485
$A_2(10,3)$	2543747	2532929
$A_2(10,5)$	49451	49394
$A_2(10,7)$	1224	1223
$A_2(11,5)$	693240	660285
$A_2(11,7)$	9120	8990
$A_2(12,7)$	323475	323374
$A_2(12,9)$	4488	4487
$A_2(13,7)$	4781932	4691980
$A_2(13,9)$	34591	34306
$A_2(14,9)$	2334298	2334086
$A_2(14,11)$	17160	17159
$A_2(15,11)$	134687	134095
$A_2(16,13)$	67080	67079

Table 3.1: Bounds for the subspace distance

Theorem 3.5.1. *Let $\mathcal{C} \subset \mathcal{P}(\mathbb{F}_q^n)$, of minimal subspace distance d, and let $D_k := |\mathcal{C} \cap \mathcal{G}_q(n,k)|$. Then, if*

$$d + 2\lceil d/2 \rceil + 2 < 2n < 2d + 2\lceil d/2 \rceil + 2,$$

we have:

- $D_{2n-d-\lceil d/2 \rceil -1} \leq 1$*;*
- *if* $D_{2n-d-\lceil d/2 \rceil -1} = 1$ *then*

$$D_{\lceil d/2 \rceil} \leq \frac{q^n - q^{2n-d-\lceil d/2 \rceil -1}}{q^{\lceil d/2 \rceil} - q^{n-d-1}}.$$

Proof. It is clear that $D_i \leq 1$ for $0 \leq i < \lceil d/2 \rceil$. Moreover, for all $x, y \in \mathcal{C} \cap \mathcal{G}(n, \lceil d/2 \rceil)$, $\dim(x \cap y) = 0$. We want to show that $D_{2n-d-\lceil d/2 \rceil -1} \leq 1$. Indeed assume by contradiction $x \neq y \in \mathcal{C} \cap \mathcal{G}(n, 2n-d-\lceil d/2 \rceil -1)$, we have

$$\begin{cases} 4n-2d-2\lceil d/2 \rceil -2 \leq n+\dim(x\cap y) \\ d \leq 4n-2d-2\lceil d/2 \rceil -2-2\dim(x\cap y) \end{cases}$$

leading to

$$\begin{cases} 2\dim(x\cap y) \geq 6n-4d-4\lceil d/2 \rceil -4 \quad (*) \\ 2\dim(x\cap y) \leq 4n-3d-2\lceil d/2 \rceil -2 \quad (**) \end{cases}$$

To obtain a contradiction, we must have $(*) > (**)$ which is equivalent to the hypothesis $2n > d+2\lceil d/2 \rceil +2$.

With a similar reasoning, we prove that, for all $x \in \mathcal{C} \cap \mathcal{G}(n, \lceil d/2 \rceil)$ and $w \in \mathcal{C} \cap \mathcal{G}(n, 2n-d-\lceil d/2 \rceil -1)$, $\dim(x \cap w) = n-d-1$. Indeed,

$$\begin{cases} 2n-d-1 \leq n+\dim(x\cap w) \\ d \leq 2n-d-1-2\dim(x\cap w) \end{cases}$$

so

$$\begin{cases} \dim(x\cap w) \geq n-d-1 \\ \dim(x\cap w) \leq n-d-1/2 \end{cases}$$

which yields the result.

Now we assume $D_{2n-d-\lceil d/2 \rceil -1} = 1$. Let $w \in \mathcal{C} \cap \mathcal{G}(n, 2n-d-\lceil d/2 \rceil -1)$. Let $\mathcal{U}$ denote the union of the subspaces x belonging to $\mathcal{C} \cap \mathcal{G}(n, \lceil d/2 \rceil)$. We have $|\mathcal{U}| = 1 + D_{\lceil d/2 \rceil}(q^{\lceil d/2 \rceil}-1)$ and $|\mathcal{U} \cap w| = 1 + D_{\lceil d/2 \rceil}(q^{n-d-1}-1)$. On the other hand, $|\mathcal{U} \backslash (\mathcal{U} \cap w))| \leq |\mathbb{F}_q^n \backslash w|$, leading to

$$D_{\lceil d/2 \rceil}(q^{\lceil d/2 \rceil} - q^{n-d-1}) \leq q^n - q^{2n-d-\lceil d/2 \rceil -1} \ .$$

□

In several cases, adding the inequalities proved in the above theorem to the programs lead to a lower optimal value, however we found that only in one case other than $(n, d) = (5, 3)$, the final result, after rounding to the previous integer, is improved. It is the case $(n, d) = (7, 5)$, where $D_3 \leq 17$ and, by Theorem 3.5.1, if $D_5 = 1$ then $D_3 \leq 16$. So we can add $D_3 + D_5 \leq 17$ and $D_2 + D_4 \leq 17$, leading to:

Theorem 3.5.2.

$$A_2(7, 5) \leq 34.$$

This bound can be obtained with both the linear program of Theorem 3.2.3 and the semidefinite program of Theorem 3.4.6.

3.5.3 The injection distance

Recently, a new metric has been considered in the framework of projective codes, the so-called *injection* metric, introduced in [30]. The *injection distance* between two subspaces $U, V \in \mathcal{P}(\mathbb{F}_q^n)$ is defined by

$$d_i(U,V) = \max\{\dim(U), \dim(V)\} - \dim(U \cap V)$$

When restricted to the Grassmann space, i.e. when U, V have the same dimension, the new distance coincides with the subspace distance (up to multiplication by 2) so nothing new comes from considering one distance instead of the other. But in general we have the relation ([30])

$$d_i(U,V) = \frac{1}{2}d_S(U,V) + \frac{1}{2}|\dim(U) - \dim(V)|$$

where d_S denotes the subspace distance.
It is straightforward to modify the programs in order to produce bounds for codes on the new metric space $(\mathcal{P}(\mathbb{F}_q^n), d_i)$. Let

$$A_q^{inj}(n,d) = \max\{|\mathcal{C}| \ : \ \mathcal{C} \subset \mathcal{P}(\mathbb{F}_q^n),\ d_i(\mathcal{C}) \geq d\}.$$

For constant dimension codes, we have $A_q^{inj}(n,k,d) = A_q(n,k,2d)$.

To modify the linear program of Etzion and Vardy for this new distance, we need to write down the "sphere-packing" like constraint. The cardinality of balls in $\mathcal{P}(\mathbb{F}_q^n)$ for the injection distance can be found in [28]. Let $B^{inj}(V,e)$ denotes the ball of center V and radius e. If $\dim(V) = i$, we have

$$\begin{aligned} |B^{inj}(V,e)| = & \sum_{r=0}^{e} q^{r^2} \begin{bmatrix} i \\ r \end{bmatrix}_q \begin{bmatrix} n-i \\ r \end{bmatrix}_q \\ & + \sum_{r=0}^{e}\sum_{\alpha=1}^{r} q^{r(r-\alpha)} \left(\begin{bmatrix} i \\ r \end{bmatrix}_q \begin{bmatrix} n-i \\ r-\alpha \end{bmatrix}_q + \begin{bmatrix} i \\ r-\alpha \end{bmatrix}_q \begin{bmatrix} n-i \\ r \end{bmatrix}_q \right). \end{aligned}$$

We define $c^{inj}(i,k,e) := |B^{inj}(V,e) \cap \mathcal{G}_q(n,k)|$ where $\dim(V) = i$. We set $\alpha := |i-k|$.

$$c^{inj}(i,k,e) = \begin{cases} \sum_{r=0}^{e} q^{r(r-\alpha)} \begin{bmatrix} i \\ r \end{bmatrix}_q \begin{bmatrix} n-i \\ r-\alpha \end{bmatrix}_q & \text{if } i \geq k \\ \sum_{r=0}^{e} q^{r(r-\alpha)} \begin{bmatrix} i \\ r-\alpha \end{bmatrix}_q \begin{bmatrix} n-i \\ r \end{bmatrix}_q & \text{if } i \leq k \end{cases}$$

We obtain:

Theorem 3.5.3 (Linear programming upper bound for codes in $\mathcal{P}(\mathbb{F}_q^n)$ for the injection distance)**.**

$$\begin{aligned} A_q^{inj}(n,d) \leq \sup\Big\{ \textstyle\sum_{k=0}^{n} x_k \ \ : \ \ & x_k \leq A_q^{inj}(n,k,d) \ \forall\ k = 0,\dots,n \\ & \textstyle\sum_{i=0}^{n} c^{inj}(i,k,e) x_i \leq \begin{bmatrix} n \\ k \end{bmatrix}_q \ \forall\ k = 0,\dots,n \Big\} \end{aligned}$$

For the semidefinite programming bound, we only need to change the definition of $\Omega(d)$; let

$$\Omega^{inj}(d) := \{(s,t,i) : 0 \le s,t \le n,\ i \le \min(s,t),\ s+t \le n+i, \\ \text{either } s=t=i \text{ or } \max(s,t)-i \ge d\}. \tag{3.5.1}$$

Then, we have:

Theorem 3.5.4.

$$A_q^{inj}(n,d) \le \sup\Big\{ \sum_{(s,t,i)\in\Omega^{inj}(d)} x_{sti} \ :\ (x_{sti})_{(s,t,i)\in\Omega^{inj}(d)},\ x_{sti} \ge 0, \\ \sum_{s=0}^{n} x_{sss} = 1, \\ F_k \succeq 0 \ \textit{for all } k = 0,\ldots,\lfloor n/2 \rfloor \Big\}$$

where $\Omega^{inj}(d)$ *is defined in* (3.5.1) *and the matrices* F_k *are given in (3.4.6).*

Table 3.2 displays the numerical computations obtained from the two programs, in the same manner as for Table 1.

	E-V LP	SDP
$A_2^{inj}(7,3)$	37	37
$A_2^{inj}(8,3)$	362	364
$A_2^{inj}(9,3)$	2533	2536
$A_2^{inj}(10,3)$	49586	49588
$A_2^{inj}(10,4)$	1229	1228
$A_2^{inj}(11,4)$	9124	9126
$A_2^{inj}(12,4)$	323778	323780
$A_2^{inj}(12,5)$	4492	4492
$A_2^{inj}(13,5)$	34596	34600
$A_2^{inj}(14,6)$	17167	17164
$A_2^{inj}(15,6)$	134694	134698
$A_2^{inj}(16,7)$	67087	67084

Table 3.2: Bounds for the injection distance

Remark 3.5.5. We observe that the bound obtained for $A_2(n,4e+1)$ is most of the time slightly larger than the one obtained for $A_2^{inj}(n,2e+1)$. In [28], the authors notice that their constructions lead to codes that are slightly better for the injection distance that for the subspace distance. So both experimental observations tend to indicate that $A_2(n,4e+1)$ should be larger than $A_2^{inj}(n,2e+1)$.

Note: All values of the linear programs have been obtained with the free solver LRS 4.2, while all values of the semidefinite programs have been obtained with the solvers SDPA or SDPT3, available on the NEOS website (http://www.neos-server.org/neos/).

Chapter 4

The chromatic number of the Euclidean space

The motivation for the work presented in this chapter is the problem of determining the chromatic number of the Euclidean space $\mathbb{R}^n$, that is, the least number of subsets which partition $\mathbb{R}^n$ in such a way that no one of them contains points at distance 1. When trying to solve this problem, the idea of using well chosen embeddings in $\mathbb{R}^n$ of some finite abstract graphs arises in a natural way. Indeed, in 1951, De Bruijn and Erdös ([12]) proved that the chromatic number of $\mathbb{R}^n$ can be determined, at least theoretically, by using this approach.

Using a so called "intersection theorem", Frankl and Wilson ([19]) established a bound on the cardinality of sets avoiding one distance in the Johnson space and, as a consequence, they derived an aymptotic result on the chromatic number of $\mathbb{R}^n$ which is still the best known. The drawback of their approach is that their bound is not good in low dimensions and moreover it works only for a limited choice of parameters.

On the other hand, bounds in the Johnson space coming from the semidefinite programming method apply to all possible parameters, as we show in the third section. The Delsarte linear programming bound is recalled. We prove a new semidefinite programming bound on sets avoiding one distance in the Johnson space by considering triples of words. This bound is a variant of Schrijver's bound on constant weight codes ([41]) and it is at least as good as Delsarte's one. We symmetrize it using a group theoretical setting in the spirit of this thesis. In most of the cases we computed, it improves the bound of Frankl and Wilson.

In [13], the authors established a linear program whose optimal value gives a lower bound for the *measurable* chromatic number of $\mathbb{R}^n$. Also, they showed how to strengthen it by including constraints on regular simplices. In the fifth section we show how our bounds on sets avoiding one distance in the Johnson space can be used in their linear program. Better results than the existing ones are obtained for $n = 9, \ldots, 23$.

4.1 Introduction

Definition 4.1.1. *The chromatic number of the real Euclidean space* $\mathbb{R}^n$*, denoted* $\chi(\mathbb{R}^n)$*, is the minimum number of colors needed to color* $\mathbb{R}^n$ *in such a way that points at distance* 1 *receive different colors.*

In other words, consider *the unit distance graph of* $\mathbb{R}^n$: it is the (infinite) graph with vertex set $\mathbb{R}^n$ and edges between points at Euclidean distance 1. Then $\chi(\mathbb{R}^n)$ is the chromatic number of this graph according to the definition given in chapter 2. It is an old and open problem to determine $\chi(\mathbb{R}^n)$, even for small dimensions.

The easiest case is that of $n = 1$, where the chromatic number of the real line is 2. The classes of coloring are $\bigcup_{i \in \mathbb{N}} [2i, 2i+1[$ and its complement (up to translation). So $\chi(\mathbb{R}) = 2$. Surprisingly, no other value is known. Consider for instance the plane: determining the chromatic number of $\mathbb{R}^2$ is known as the Hadwiger–Nelson problem. Looking at an equilateral triangle with side length 1, it becomes clear that in a coloring of the plane, the three vertices will receive different colors. This allows us to establish the lower bound $\chi(\mathbb{R}^2) \geq 3$. The role played by the triangle can be generalized as follows.

Definition 4.1.2. *A unit distance graph in* $\mathbb{R}^n$ *is a finite graph which can be embedded in* $\mathbb{R}^n$ *in such a way that its adjacent vertices are exactly the vertices at Euclidean distance* 1*. In other words, it is a finite induced subgraph of the unit distance graph of* $\mathbb{R}^n$*.*

Clearly, every cycle graph is a unit distance graph in the plane when seen as a regular polygon of side length 1. But, as the chromatic number of a graph is 2 or 3, according to the parity of n, we cannot hope to use cycles to improve the previous bound.

The Moser graph is a unit distance graph in $\mathbb{R}^2$ whose chromatic number is 4, thus leading to the lower bound $4 \leq \chi(\mathbb{R}^2)$.

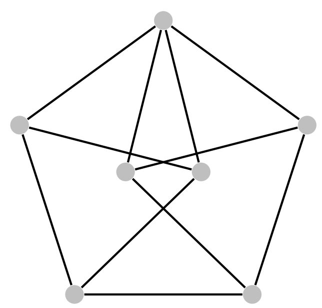

Figure 4.a: Moser unit distance graph

On the other hand, the classic hexagonal tiling of the plane realizes a 7-coloring leading to the upper bound $\chi(\mathbb{R}^2) \leq 7$. No better bounds are known in dimension 2.

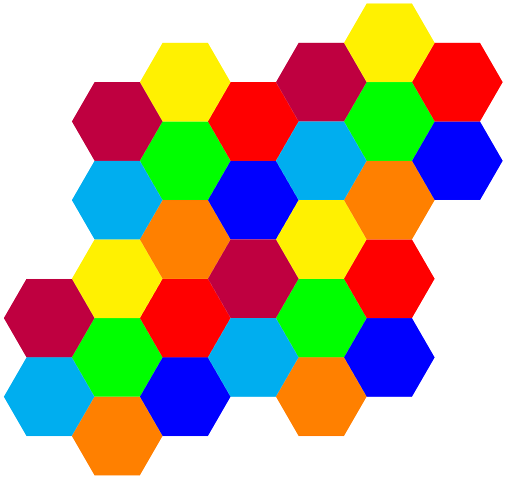

Figure 4.b: Hexagonal tiling

Let us mention also that, up to now, the best estimate in dimension 3 is given by $6 \leq \chi(\mathbb{R}^3) \leq 15$ (see the overview in [31]).

In general, it is known by the De Bruijn-Erdos theorem ([12]) that $\chi(\mathbb{R}^n)$ is attained by a finite unit distance graph $\mathcal{G} \subset \mathbb{R}^n$.

4.2 Frankl-Wilson graphs

Of course, for any graph $\mathcal{G} = (V, E)$ we have

$$\chi(\mathcal{G})\alpha(\mathcal{G}) \geq |V|$$

In particular, if $\mathcal{G}$ is a unit distance graph embedded in $\mathbb{R}^n$, we find

$$\chi(\mathbb{R}^n) \geq \chi(\mathcal{G}) \geq \frac{|V|}{\alpha(\mathcal{G})} \tag{4.2.1}$$

so any upper bound on $\alpha(\mathcal{G})$ will result in a lower bound for $\chi(\mathbb{R}^n)$.

Definition 4.2.1 ([19]). *Given n, let $p < (n+1)/2$ be a power of a prime number. The* Frankl-Wilson graph *associated to the pair (n,p) is the graph $J(n,p)$ with vertex set $V = \frac{1}{\sqrt{2p}} J_n^{2p-1}$ and edges set $E = \{xy : x.y = (p-1)/2p\}$.*

So we are considering vectors of length n with weight $2p-1$ where non zero entries take the constant value $1/\sqrt{2p}$. Note that Frankl-Wilson graphs can be embedded in $\mathbb{R}^n$ as unit distance graphs. In fact, denoting $\mathbf{x}$ the image of x in $\mathbb{R}^n$, we see that

$$d(\mathbf{x}, \mathbf{y}) = 1 \Leftrightarrow \sqrt{\sum_i (\mathbf{x}_i - \mathbf{y}_i)^2} = \sqrt{2p} \Leftrightarrow |x \cap y| = p - 1 \Leftrightarrow xy \in E$$

Moreover, as all words have constant weight, they can actually be embedded in $\mathbb{R}^{n-1}$. Following the previous discussion, we find that

$$\chi(\mathbb{R}^{n-1}) \geq \frac{\binom{n}{2p-1}}{\alpha(J(n,p))}$$

Let us now focus on $\alpha(J(n,p))$. Applying their intersection theorem ([19]), Frankl and Wilson found

$$\alpha(J(n,p)) \leq \binom{n}{p-1} \tag{4.2.2}$$

which, optimizing over p, leads to $\chi(\mathbb{R}^n) \geq (1.207+o(1))^n$, provided n is large enough. In 2000, Raigorodskii improved the lower bound to $(1.239 + o(1))^n$ ([39]).

In what follows, we will compare the Frankl-Wilson bound 4.2.2 to other bounds coming from SDP. Notably, we have the ϑ' number which, after symmetrization, reduces to a linear program, and the Schrijver-like bound involving triples of words.

The independent sets of the Frankl-Wilson graph $J(n,p)$ are the constant weight codes in the Johnson space J_n^w, for $w = 2p - 1$, with no words at Hamming distance $d = 2p$ (i.e. Johnson distance $\delta := d/2 = p$). Thus the independence number of $J(n,p)$ is the maximal size of such a code. While the Frankl-Wilson bound is restricted to the choice of parameters of definition 4.2.1, the bounds that we give in the next section apply to all choices of parameters $w \leq n/2$ and δ.

4.3 Bounds for sets avoiding one distance in the Johnson space

Let $w \leq n/2$ and define $J(n,w,i)$ as the graph with vertex set J_n^w with an edge xy whenever $|x \cap y| = i$. Frankl-Wilson graphs correspond, up to a normalization factor, to $J(n, 2p-1, p-1)$ and this is precisely the case in which the bound 4.2.2 holds. With the corresponding normalization factor, these graphs $J(n,w,i)$ are also unit distance graphs in $\mathbb{R}^n$. Moreover, as they have constant weight, they can actually be embedded in $\mathbb{R}^{n-1}$.

4.3.1 The LP bound

As the Johnson space J_n^w is 2-point homogeneous for the action of the symmetric group S_n, the ϑ' number for graphs in this space reduces to Delsarte linear program. The reduction, contained in [15], is essentially the same as for the symmetrization of the Grassmann space, which indeed is the q-analog of J_n^w (see section 3.3.3). It involves a simplified version of the Hahn polynomials, namely the family associated to parameters n, w, w in definition 2.4.5.

Definition 4.3.1. *The Hahn polynomials associated to the parameters n, w with $0 \leq w \leq n$ are the polynomials $Q_k(t)$ with $0 \leq k \leq w$ uniquely determined by the properties:*

- *Q_k has degree k,*
- *for all $0 \leq k < l \leq w$,*

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

- *$Q_k(0) = 1$.*

Theorem 4.3.2. *(Delsarte's linear programming bound [15]).*
The maximum number of words in J_n^w avoiding Johnson distance δ is upper bounded by the optimal value of the following linear program:

$$\inf \Big\{ 1 + f_1 + \cdots + f_w \ : \ \begin{array}{l} f_k \geq 0 \ \forall\ k = 1, \ldots, w \\ F(t) \leq 0 \ \forall\ t = 1, \ldots, \delta - 1, \delta + 1, \ldots, w \Big\} \end{array}$$

where $F(t) = 1 + \sum_{k=1}^{w} f_k Q_k(t)$ and $Q_k(t)$ is as in the definition above.

Remark 4.3.3. From the previous discussion, applying theorem 4.3.2 with $\delta = w - i$ yields an upper bound on $\alpha(J(n, w, i))$ which in the rest of the chapter will be referred to as *LP bound.*

4.3.2 The SDP bound on triples of words

The bound that we establish in this section is a strengthening of the ϑ' number, obtained by constraints on triples of words. It was first obtained in [41], where it is applied to constant weight codes. We will see that the symmetrization needed is closely related to the block diagonalization of the Terwilliger algebra (section 2.4.3), from which we take some notations and results. We restrict here to the case $w < n/2$. Set $X := J_n^w$ and fix $z_0 := 1^w 0^{n-w}$. Then S_n acts on X and $\mathrm{Stab}_{S_n}(z_0) = S_w \times S_{n-w}$. We start by giving the general SDP formulation of our bound.

Theorem 4.3.4. *(The SDP bound on triples).*
The maximum number of points in a metric space X avoiding distance δ is upper bounded by the optimal value of the following semidefinite program:

$$\sup \left\{ \begin{array}{lll} \sum_{X^2} F(x,y,x) & : & F(x,y,z) = F(\{x,y,z\}) \\ & & \forall\ z \in X, \left[(x,y) \to F(x,y,z)\right] \succeq 0 \\ & & F(x,y,z) \geq 0\ \forall\ x,y,z \\ & & \sum_X F(x,x,x) = 1 \\ & & F(x,y,z) = 0 \quad \textit{if} \quad \delta \in \{d(x,y), d(x,z), d(y,z)\} \end{array} \right\} \tag{4.3.1}$$

where the first constraint means that F is invariant under permutation of the variables, that is, the value of $F(x,y,z)$ depends only on the set $\{x,y,z\}$.

Proof. For every $\mathcal{C} \subset X$ with no words at pairwise distance δ, the function

$$F(x,y,z) := \begin{cases} 1/|\mathcal{C}| & \text{if all } x,y,z \in \mathcal{C} \\ 0 & \text{else} \end{cases}$$

is feasible for the program 4.3.1 and its objective value is

$$\sum_{X^2} F(x,y,x) = \frac{1}{|\mathcal{C}|}|\mathcal{C}|^2 = |\mathcal{C}| \ .$$

□

Remark 4.3.5. It is easily seen that this bound improves the bound of ϑ' (remark 2.3.8) applied to the corresponding graph. Indeed, for a feasible F, set $G(x,y) := F(x,y,x)$. Then G is feasible for ϑ' and its objective value is the same as the one of F. In the Johnson space, this means that the SDP bound on triples is always at least as good as Delsarte LP bound.

In the case of the Johnson space, the program 4.3.1 need to be symmetrized to reduce its size in order to compute explicit optimal values. This can be done similarly to what we explained for ϑ in section 2.4.1.

Description of the orbits and first reduction

Consider the action of S_n on triples of words. For $(x,y,z) \in X^3$ we define $a := |x \cap z|$, $b := |y \cap z|$, $c := |x \cap y|$, $d := |x \cap y \cap z|$ and we have:

Theorem 4.3.6. *Two elements of X^3 are in the same orbit under the action of S_n if and only if they give rise to the same 4-tuple (a, b, c, d). Moreover the collection of orbits is parametrized by the following set:*

$$\Omega(n,w) := \left\{ \begin{array}{lcl} (a,b,c,d) & : & 0 \le d \le a,b,c \le w \\ & & \max\{a+b, a+c, b+c\} \le w+d \\ & & 2w \le n + \min\{a,b,c\} \\ & & 3w + d \le n + a + b + c \end{array} \right\} \tag{4.3.2}$$

Proof. The action of the symmetric group preserves cardinality of intersection of words, so the "only if" part obviously holds. Conversely, let us define the standard representative (x_0, y_0, z_0) of the S_n-orbit parametrized by (a, b, c, d) as:

$$x_0 := 1^d 1^{a-d} 0^{b-d} 0^{w-a-b+d} 1^{c-d} 1^{w-a-c+d} 0^{w-b-c+d} 0 \dots 0$$

$$y_0 := 1^d 0^{a-d} 1^{b-d} 0^{w-a-b+d} 1^{c-d} 0^{w-a-c+d} 1^{w-b-c+d} 0 \dots 0$$

$$z_0 := 1^d 1^{a-d} 1^{b-d} 1^{w-a-b+d} 0^{c-d} 0^{w-a-c+d} 0^{w-b-c+d} 0 \dots 0$$

Then it is clear that for any other triple with the same parameters (a, b, c, d) we can act with S_n by permuting each 0-1 block to transform such triple into the standard one (x_0, y_0, z_0).

As for the second statement of the proof, given (a, b, c, d) satisfying the constraints in 4.3.2 we can build (x_0, y_0, z_0) as above. Conversely, given any triple (x, y, z), the first constraint is clear, the second one comes by putting together relations of the kind $(x \cap z) \cup (y \cap z) \subset z$, the third simply by $|x \cup z| \le n$ and the last one by $|x \cup y \cup z| \le n$. □

The program 4.3.1 is invariant for the action of S_n, so we can restrict to consider S_n-invariant functions F. From the preceding theorem, these functions satisfy $F(x, y, z) = \hat{F}(a, b, c, d)$. Let x_{abcd} denote the cardinality of the respective orbit for the action of S_n on X. Then, program 4.3.1 is equivalent to:

$$\sup \left\{ \sum_{b=0}^{w} x_{wbbb} \hat{F}(w,b,b,b) \;:\; \begin{array}{l} \hat{F}(a,b,c,d) = \hat{F}(\{a,b,c\},d) \\ \forall\; z \in X, \Big[(x,y) \to F(x,y,z)\Big] \succeq 0 \\ \hat{F}(a,b,c,d) \ge 0 \;\forall\; (a,b,c,d) \in \Omega(n,w) \\ x_{wwww} \hat{F}(w,w,w,w) = 1 \\ \hat{F}(a,b,c,d) = 0 \text{ if } \delta \in \{w-a, w-b, w-c\} \end{array} \right\}$$

Now we focus on the second constraint. The following proposition is easy to prove.

Proposition 4.3.7. *If the function $(x, y, z) \to F(x, y, z)$ is G-invariant, then for any $z_0 \in X$, the function $(x, y) \to F(x, y, z_0)$ is $Stab_G(z_0)$-invariant.*

So we need a description of the positive semidefinite functions which are invariant under the stabilizer of one point. As S_n acts transitively on X, all the stabilizer subgroups are conjugated. In what follows we fix the point $z_0 := 1^w 0^{n-w}$ and $\mathrm{Stab}_{S_n}(z_0) = S_w \times S_{n-w}$. We are going to block diagonalize the algebra $(\mathbb{R}^{X\times X})^{S_w \times S_{n-w}}$ with an explicit isomorphism which preserves positive semidefiniteness. To do this, we can follow the method given in sections 2.4.1 and 2.4.3. It will turn out that we are in fact block diagonalizing a tensor product of Terwilliger algebras, namely the ones associated to H_w and H_{n-w}. As a consequence, the zonal functions that we build will be a product of zonal functions of the respective spaces.

Decomposition of $\mathbb{R}^X$ under $S_w \times S_{n-w}$

Clearly, X has a partition into orbits $X_i := \{x \in X : |x \cap z_0| = i\}$ which yields

$$\mathbb{R}^X = \mathbb{R}^{X_0} \oplus \mathbb{R}^{X_1} \oplus \cdots \oplus \mathbb{R}^{X_w}$$

Yet, this decomposition is not irreducible. To decompose it further, note that

$$\mathbb{R}^{X_i} \simeq \mathbb{R}^{J_w^i} \otimes \mathbb{R}^{J_{n-w}^{w-i}}$$

The irreducible decomposition of the Johnson space into Specht modules was given in 2.4.9:

$$\mathbb{R}^{J_w^i} = \begin{cases} H_{0,i}^w \oplus \cdots \oplus H_{i,i}^w & \text{if } i \leq \lfloor w/2 \rfloor \\ H_{0,i}^w \oplus \cdots \oplus H_{w-i,i}^w & \text{else} \end{cases}$$

$$\mathbb{R}^{J_{n-w}^{w-i}} = \begin{cases} H_{0,w-i}^{n-w} \oplus \cdots \oplus H_{w-i,w-i}^{n-w} & \text{if } w-i \leq \lfloor (n-w)/2 \rfloor \\ H_{0,w-i}^{n-w} \oplus \cdots \oplus H_{n-2w+i,w-i}^{n-w} & \text{else} \end{cases}$$

where $H_{k,i}^w \simeq \mathfrak{S}^{(w-k,k)}$ and $H_{h,w-i}^{n-w} \simeq \mathfrak{S}^{(n-w-h,h)}$. So we get

$$\begin{aligned} \mathbb{R}^{X_i} &= \left(H_{0,i}^w \oplus \cdots \oplus H_{\min\{i,w-i\},i}^w\right) \otimes \left(H_{0,w-i}^{n-w} \oplus \cdots \oplus H_{\min\{w-i,n-2w+i\},w-i}^{n-w}\right) \\ &\simeq \left(H_{0,i}^w \otimes H_{0,w-i}^{n-w}\right) \oplus \cdots \oplus \left(H_{\min\{i,w-i\},i}^w \otimes H_{\min\{w-i,n-2w+i\},w-i}^{n-w}\right) \end{aligned}$$

We know that the irreducible modules for the action of the direct product of groups are the tensor products of irreducible modules for the action of each of the group factor. We know that Specht modules are the irreducible modules for the action of the symmetric group. So, in the case of the product of symmetric groups $S_w \times S_{n-w}$ we have decomposed each $\mathbb{R}^{X_i}$ in its irreducible components.

Theorem 4.3.8. *The irreducible decomposition of the space* $\mathbb{R}^X$ *under the action of* $S_w \times S_{n-w}$ *is given by:*

$$\begin{aligned}\mathbb{R}^X &= \bigoplus_{i=0}^{w} \bigoplus_{k=0}^{\min\{i,w-i\}} \bigoplus_{h=0}^{\min\{w-i,n-2w+i\}} H^w_{k,i} \otimes H^{n-w}_{h,w-i} \\ &\simeq \bigoplus_{k=0}^{\lfloor w/2 \rfloor} \bigoplus_{h=0}^{\lfloor (n-w)/2 \rfloor} H_{k,h}^{m_{k,h}}\end{aligned}$$

where

$$H_{k,h} := H^w_{k,k} \otimes H^{n-w}_{h,h} \simeq \mathfrak{S}^{(w-k,k)} \otimes \mathfrak{S}^{(n-w-h,h)}$$

is of dimension

$$d_{k,h} := \left[\binom{w}{k} - \binom{w}{k-1}\right] \cdot \left[\binom{n-w}{h} - \binom{n-w}{h-1}\right]$$

and the multiplicity is given by

$$\begin{aligned}m_{k,h} &:= |\{k, \dots, w-k\} \cap \{h, \dots, n-w-h\}| \\ &= \max\Big\{ \min\{w-k, n-w-h\} - \max\{k,h\}, 0 \Big\}\end{aligned}$$

Proof. From the previous discussion. To prove the multiplicity formula, it is enough to remark that (an isomorphic copy of) $H_{k,h}$ is contained in $\mathbb{R}^{X_{w-j}}$ if and only if $j \in \{k, \dots, w-k\} \cap \{h, \dots, n-w-h\}$. Note that, depending on n and w, some of the multiplicities can be 0. □

Construction of the zonal matrices

Recall from section 2.4.3 the following facts:

- For $0 \leq k \leq \lfloor w/2 \rfloor$, we can take an orthonormal basis $\{e^w_{k,k,1}, \dots, e^w_{k,k,d_k}\}$ for $H^w_{k,k}$ inside $\mathbb{R}^{J^k_w}$.
- Analogously, for $0 \leq h \leq \lfloor (n-w)/2 \rfloor$, we can take an orthonormal basis $\{e^{n-w}_{h,h,1}, \dots, e^{n-w}_{h,h,d_h}\}$ for $H^{n-w}_{h,h}$ inside $\mathbb{R}^{J^h_{n-w}}$.
- Here $d_k = \binom{w}{k} - \binom{w}{k-1}$ and $d_h = \binom{n-w}{h} - \binom{n-w}{h-1}$.
- Whenever $i > k$ and $j > h$, by mean of the corresponding valuation operators, we obtain basis $\{e^w_{k,i,1}, \dots, e^w_{k,i,d_k}\}$ of $H^w_{k,i}$ and $\{e^{n-w}_{h,j,1}, \dots, e^{n-w}_{h,j,d_h}\}$ of $H^{n-w}_{h,j}$.

With this we can build a basis for the copy $H^w_{k,i} \otimes H^{n-w}_{h,w-i}$ of $H_{k,h}$ inside $\mathbb{R}^{X_i}$ as

$$\{e^w_{k,i,1} e^{n-w}_{h,w-i,1}, \dots, e^w_{k,i,d_k} e^{n-w}_{h,w-i,d_h}\}$$

Now, the zonal matrices satisfy:

$$E_{k,h} \in \mathbb{R}^{m_{k,h} \times m_{k,h}} \quad \text{for} \quad \begin{array}{l} k = 0, \dots, \lfloor w/2 \rfloor \\ h = 0, \dots, \lfloor (n-w)/2 \rfloor \end{array}$$

$$E_{k,h,i,j}(x,y) \neq 0 \text{ only if } |x \cap z_0| = i, |y \cap z_0| = j$$

and they are given by:

$$\begin{aligned} E_{k,h,i,j}(x,y) &= \frac{1}{|X|} \sum_{\ell=0}^{d_k} \sum_{\ell'=0}^{d_h} e^w_{k,i,\ell}(x) e^{n-w}_{h,w-i,\ell'}(x) e^w_{k,j,\ell}(y) e^{n-w}_{h,w-j,\ell'}(y) \\ &= \frac{1}{|X|} \Big(\sum_{\ell} e^w_{k,i,\ell}(x) e^w_{k,j,\ell}(y) \Big) \Big(\sum_{\ell'} e^{n-w}_{h,w-i,\ell'}(x) e^{n-w}_{h,w-j,\ell'}(y) \Big) \\ &= \frac{1}{|X|} E^w_{k,i,j}(x,y) E^{n-w}_{h,w-i,w-j}(x,y) \end{aligned}$$

where, by theorem 2.4.6,

$$E^w_{k,i,j}(x,y) = |H_w| \frac{\binom{j-k}{i-k}\binom{w-2k}{j-k}}{\binom{w}{j}\binom{j}{i}} Q_k(w,i,j;\ i - |x \cap y \cap z_0|)$$

and

$$E^{n-w}_{h,w-i,w-j}(x,y) = |H_{n-w}| \frac{\binom{w-j-h}{w-i-h}\binom{n-w-2h}{w-j-h}}{\binom{n-w}{w-j}\binom{w-j}{w-i}} Q_h(n-w, w-i, w-j;\ w-i-|x \cap y \cap \overline{z_0}|)$$

The final reduction

We go back to program 4.3.1. By Bochner theorem 2.4.1 we know that the function $F_z(x,y) := F(x,y,z)$ is positive semidefinite and $(S_w \times S_{n-w})$-invariant if and only if there exist positive semidefinite matrices $F_{k,h}$ such that

$$F_z(x,y) = \sum \langle F_{k,h}, E_{k,h}(x,y) \rangle \ .$$

We can recover the matrices $F_{k,h}$ by mean of the scalar product

$$F_{k,h} = \frac{1}{|X|^2} \sum_{x,y \in X} F_z(x,y) E_{k,h}(x,y)$$

thus obtaining an explicit formula in terms of the variables $y_{abcd} := x_{abcd} \hat{F}(a,b,c,d)$. This yields the final reduction of program 4.3.1 in J^w_n:

Theorem 4.3.9. *The maximum number of words in J_n^w avoiding Johnson distance δ is upper bounded by the optimal value of the following semidefinite program:*

$$\sup\left\{ \sum_{b=0}^{w} y_{wbbb} \ : \ \begin{array}{l} y_{abcd} = y_{\sigma(a)\sigma(b)\sigma(c)d} \ \forall \ \sigma \in S_3 \\ F_{k,h} \succeq 0 \\ y_{abcd} \geq 0 \ \forall \ (a,b,c,d) \in \Omega(n,w) \\ y_{wwww} = 1 \\ y_{abcd} = 0 \ \textit{if} \ \delta \in \{w-a, w-b, w-c\} \end{array} \right\}$$

where, for $0 \leq k \leq \lfloor w/2 \rfloor$, $0 \leq h \leq \lfloor (n-w)/2 \rfloor$, $F_{k,h} \in \mathbb{R}^{m_{k,h} \times m_{k,h}}$ is defined by:

$$(F_{k,h})_{a,b} = \sum_{c,d} y_{abcd} \frac{\binom{b-k}{a-k}\binom{w-2k}{b-k}\binom{w-b-h}{w-a-h}\binom{n-w-2h}{w-b-h}}{\binom{w}{b}\binom{b}{a}\binom{n-w}{w-b}\binom{w-b}{w-a}}$$
$$Q_k(w,a,b;\ a-d) Q_h(n-w, w-a, w-b;\ w-a-c+d)$$

Remark 4.3.10. From the previous discussion, applying theorem 4.3.9 with $\delta = w - i$ yields an upper bound on $\alpha(J(n,w,i))$ which in the rest of the chapter will be referred to as *SDP bound.*

4.4 Numerical results

In the following table we compare

- the Frankl-Wilson bound 4.2.2,
- the linear programming bound of theorem 4.3.2,
- the SDP bound of theorem 4.3.9

on the independence number of graphs $J(n,w,i)$ for some choice of parameters. We recall that the F-W bound applies only in the case $(w,i) = (2p-1, p-1)$ for some p power of a prime. Apart from these cases, we have chosen the graphs $J(n,w,i)$ that we will use later on in tables 4.2 and 4.3.

Remark 4.4.1. Here we see that for small parameters the SDP bound is stronger than the bound of Frankl and Wilson. What is not known is the asymptotic behaviour of the SDP bound. In particular, can it beat the $(1.207)^n$ obtained with Frankl-Wilson graphs?

(n, w, i)	$\alpha(J(n, w, i))$	F-W bound	LP bound	SDP bound
$(6, 3, 1)$	4	6	4	4
$(7, 3, 1)$	5	7	5	5
$(8, 3, 1)$	8	8	8	8
$(9, 3, 1)$	8	9	11	8
$(10, 5, 2)$	27	45	30	27
$(11, 5, 2)$	37	55	42	37
$(12, 5, 2)$	57	66	72	57
$(12, 6, 2)$			130	112
$(13, 5, 2)$		78	109	72
$(13, 6, 2)$			191	148
$(14, 7, 3)$		364	290	184
$(15, 7, 3)$		455	429	261
$(16, 7, 3)$		560	762	464
$(16, 8, 3)$			1315	850
$(17, 7, 3)$		680	1215	570
$(17, 8, 3)$			2002	1090
$(18, 9, 4)$		3060	3146	1460
$(19, 9, 4)$		3876	4862	2127
$(20, 9, 3)$			13765	6708
$(20, 9, 4)$		4845	8840	3625
$(21, 9, 4)$		5985	14578	4875
$(21, 10, 4)$			22794	8639
$(22, 9, 4)$		7315	22333	6480
$(22, 11, 5)$			36791	11360
$(23, 9, 4)$		8855	32112	8465
$(23, 11, 5)$			58786	17055
$(24, 9, 4)$		10626	38561	10796
$(24, 12, 5)$			172159	53945
$(25, 9, 4)$		12650	46099	13720
$(26, 13, 6)$		230230	453169	101494
$(27, 13, 6)$		296010	742900	163216

Table 4.1: Upper bounds for $\alpha(J(n, w, i))$

4.5 Density of sets avoiding distance 1

The lower bounds on the Euclidean chromatic number obtained by the results of table 4.1 together with the inequality 4.2.1 don't improve over the existing ones. In this section we will explain how those results can be used to improve bounds on the so-called *measurable chromatic number* of the Euclidean space, that is in the case when the color classes are required to be measurable with respect to the Lebesgue measure of $\mathbb{R}^n$.

Definition 4.5.1. *The measurable chromatic number of the real Euclidean space $\mathbb{R}^n$, denoted $\chi_m(\mathbb{R}^n)$, is the minimum number of colors needed to color $\mathbb{R}^n$ in such a way that points at distance 1 receive different colors and so that points receiving the same color form Lebesgue measurable sets.*

Clearly, we have the inequality $\chi_m(\mathbb{R}^n) \geq \chi(\mathbb{R}^n)$. For the plane, it is only known that $5 \leq \chi_m(\mathbb{R}^2) \leq 7$ ([18]). We introduce the density of a measurable set A:

$$\delta(A) := \limsup_{r \to +\infty} \frac{\operatorname{vol}(A \cap B_n(r))}{\operatorname{vol}(B_n(r))}$$

and the *extreme density* of a set avoiding distance 1:

$$m_1(\mathbb{R}^n) := \sup \left\{ \delta(A) : A \subset \mathbb{R}^n \text{ is measurable}, d(x,y) \neq 1 \ \forall\, x, y \in A \right\}$$

As a coloring gives a partition of $\mathbb{R}^n$ and $\delta(\mathbb{R}^n) = 1$, we have the relation

$$\chi_m(\mathbb{R}^n) \cdot m_1(\mathbb{R}^n) \geq 1$$

so we can focus on upper bounds for the extreme density of sets avoiding distance 1 in order to find lower bounds for the measurable chromatic number.

In [13], an upper bound on m_1 is given in the form of the optimal value of a linear program (see [13] for the definition of the function Ω_n):

Theorem 4.5.2 ([13], theorem 1.1)**.**

$$m_1(\mathbb{R}^n) \leq \inf \left\{ z_0 \quad : \quad \begin{array}{l} z_0 + z_1 \geq 1 \\ z_0 + \Omega_n(t) z_1 \geq 0 \ \forall\, t \geq 0 \end{array} \right\} \tag{4.5.1}$$

Moreover, an explicit optimal solution of the program above can be described in terms of the absolute minimum of Ω_n.

The linear program above can be strengthened by adding linear inequalities coming by well chosen finite unit distance graphs in $\mathbb{R}^n$. Indeed, in [13] it is shown how the bound 4.5.1 can be modified by using the regular simplex. With this new program, improvements are obtained on $m_1(\mathbb{R}^n)$ (hence also on $\chi_m(\mathbb{R}^n)$) for $n = 4, \ldots, 24$.

The same reasoning can be applied with a unit distance graph other than the regular simplex and theorem 4.5.1 becomes:

Theorem 4.5.3. *If $\mathcal{G} = (V, E)$ is a unit distance graph in $\mathbb{R}^n$, then*

$$m_1(\mathbb{R}^n) \leq \inf \Big\{ z_0 + z_2 \tfrac{\alpha(\mathcal{G})}{|V|} \ : \ \begin{array}{l} z_2 \geq 0 \\ z_0 + z_1 + z_2 \geq 1 \\ z_0 + z_1 \Omega_n(t) + z_2 \frac{1}{|V|} \sum_{i=1}^{|V|} \Omega_n(r_i t) \geq 0 \\ \textit{for all} \ \ t > 0 \end{array} \Big\} \tag{4.5.2}$$

where, for $x_i \in V$, $r_i := \|x_i\|$.

Remark 4.5.4. The program 4.5.2 involves an uncountable set of linear constraints. The method to solve a linear program of this kind is to discretize an interval $]0, M]$, then solve the linear program for t restricted to such discretized values and finally (if needed) adjust the optimal solution by adding a constant factor.

At this point, it is clear how to use the upper bounds on $\alpha(J(n, w, i))$ of table 4.1 in theorem 4.5.2 to upper bound $m_1(\mathbb{R}^n)$.

4.6 Numerical results

In the following two tables we give improved upper bounds on $m_1(\mathbb{R}^n)$ and lower bounds on $\chi_m(\mathbb{R}^n)$. For $n = 9, \dots, 23$ we give:

- the best previously known bound (from the table in [13]),
- the bound of theorem 4.5.2 applied with regular simplices, established in [13],
- the bound of theorem 4.5.2 together with our choice of the unit distance graph $\mathcal{G} = J(n, w, i)$. The "*" indicates that the upper bound on $\alpha(\mathcal{G})$ from our theorem 4.3.9 is used.

n	previous	(4.5.2) with $\mathcal{G}$ simplex	(4.5.2) with $\mathcal{G} = J(n,w,i)$	$J(n,w,i)$
9	0.0288215	0.0187324	0.01678	J(10,5,2)
10	0.0223483	0.0138079	0.01269	J(11,5,2)
11	0.0178932	0.0103166	0.0088775	J(12,6,2)*
12	0.0143759	0.00780322	0.006111	J(13,6,2)*
13	0.0120332	0.00596811	0.00394332	J(14,7,3)*
14	0.00981770	0.00461051	0.00300286	J(15,7,3)*
15	0.00841374	0.00359372	0.00242256	J(16,8,3)*
16	0.00677838	0.00282332	0.00161645	J(17,8,3)*
17	0.00577854	0.00223324	0.00110487	J(18,9,4)*
18	0.00518111	0.00177663	0.00084949	J(19,9,4)*
19	0.00380311	0.00141992	0.00074601	J(20,9,3)*
20	0.00318213	0.00113876	0.00046909	J(21,10,4)*
21	0.00267706	0.00091531	0.00031431	J(22,11,5)*
22	0.00190205	0.00073636	0.00024621	J(23,11,5)*
23	0.00132755	0.00059204	0.0002122678	J(24,12,5)

Table 4.2: Upper bounds for $m_1(\mathbb{R}^n)$

n	previous	(4.5.2) with $\mathcal{G}$ simplex	(4.5.2) with $\mathcal{G} = J(n,w,i)$	$J(n,w,i)$
9	35	54	60	J(10,5,2)
10	48	73	79	J(11,5,2)
11	64	97	113	J(12,6,2)*
12	85	129	164	J(13,6,2)*
13	113	168	254	J(14,7,3)*
14	147	217	334	J(15,7,3)*
15	191	279	413	J(16,8,3)*
16	248	355	619	J(17,8,3)*
17	319	448	906	J(18,9,4)*
18	408	563	1178	J(19,9,4)*
19	521	705	1341	J(20,9,3)*
20	662	879	2132	J(21,10,4)*
21	839	1093	3182	J(22,11,5)*
22	1060	1359	4062	J(23,11,5)*
23	1336	1690	4712	J(24,12,5)*

Table 4.3: Lower bounds for $\chi_m(\mathbb{R}^n)$

Chapter 5

Hierarchy of semidefinite programs from simplicial complexes

In this chapter we introduce a new sequence ϑ_k of semidefinite programs whose optimal value upper bound the independence number of a graph. To define it, we adopt the framework of simplicial complexes, which we introduce in the second section. In the third section we give the definition of ϑ_k. We show that, as for the hierarchies recalled in chapter 2, the first step coincides with the ϑ number and, for a graph $\mathcal{G}$, the $\alpha(\mathcal{G})$-th step yields the exact value $\alpha(\mathcal{G})$. We prove that the sandwich theorem holds for any step of the new sequence ϑ_k, along with a result involving graph homomorphisms which generalizes one well known property of ϑ. A drawback is that we don't know whether this "hierarchy" is decreasing at any step or not. To insure the decreasing property, in the fourth section we slightly modify the SDP formulation of ϑ_k and we show that this modification is compatible with all the properties of ϑ_k proved before. In the fifth section, some values of the second step of our hierarchy are computed for cycle graphs and Paley graphs. They show significant improvements over the first step.

5.1 Introduction

Let A be the adjacency matrix of a graph $\mathcal{G} = (V, E)$ and D be the diagonal matrix with $D_{i,i} = \deg(i)$. The matrix $L_0 := D - A$ is called the *Laplacian* of the graph $\mathcal{G}$. Recall the dual formulation for Lovász ϑ number of $\mathcal{G}$:

$$\vartheta(\mathcal{G}) = \min\{\lambda_{\max}(Z) \ : \ Z_{i,j} = 1 \ \text{ if } i = j \text{ or } ij \in \overline{E}\} \tag{5.1.1}$$

We note that the constraint on Z is that $Z = J$ outside of the support of A. So it is natural to optimize over the set

$$Z = J + \gamma A, \qquad \gamma \in \mathbb{R}.$$

Also, the eigenvalues of Z are much easier to analyse if J and A commute i.e. if the graph is regular. To make the connection with the Laplacian introduced above, it is the case $L_0 = dI - A$ if d is the degree of $\mathcal{G}$. We can rewrite

$$Z = J + \gamma dI - \gamma L_0, \qquad \gamma \in \mathbb{R}.$$

If $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ denote the eigenvalues of L_0, because $L_0 J = 0$ we have $\lambda_n = 0$ and the eigenvalues of Z are: $n + \gamma d$, $\gamma(d - \lambda_i)$ $(1 \leq i \leq n-1)$. The optimal choice for γ is obtained for

$$\min_{\gamma} \Big\{ \max\{n + \gamma d,\ \gamma(d - \lambda_i)\ (1 \leq i \leq n-1)\} \Big\}$$

which is equal to

$$\frac{n(\lambda_{\max}(L_0) - d)}{\lambda_{\max}(L_0)}$$

in which we recognize the expression of the well-known *Hoffman bound*:

Theorem 5.1.1. *(Hoffman bound) Let $\mathcal{G}$ be a regular graph of degree d, let A be its adjacency matrix. We have*

$$\alpha(\mathcal{G}) \leq \vartheta(\mathcal{G}) \leq \frac{n\lambda_{\min}(A)}{\lambda_{\min}(A) - d}.$$

Moreover, the last inequality is an equality for edge transitive graphs and for strongly regular graphs.

In general, L_0 and J commute between them, so restricting to matrices

$$Z = J + \sum_{i \geq 0} \gamma_i L_0^i$$

in 5.1.1, we obtain a linear program.

The original motivation for the work in this chapter is an attempt to build a SDP hierarchy based on generalizations of matrices J and L_0, with the hope to get high order Hoffman-like eigenvalues bounds. We give the details of the construction of this new SDP hierarchy, then we discuss its properties and some possible modifications. High order generalizations of matrices J and L_0 arise in a natural way when considering simplicial complexes.

5.2 Simplicial complexes

An *abstract simplicial complex* is a family Δ of finite subsets of a given set V such that whenever $A \in \Delta$ and $B \subset A$, then $B \in \Delta$. Elements of a simplicial complex are called *faces* and the *dimension* of a face $A \in \Delta$ is defined as $\dim(A) := |A| - 1$. The set of faces of dimension k is denoted by Δ_k. We remark that abstract simplicial complexes can be thought of as a generalization of graphs. Indeed, a graph is nothing else than a complex with faces of dimension 0 (vertices) and 1 (edges), apart from the empty set. For more details on simplicial complexes, see [25].

The goal of this section is to define a pair of matrices associated to a simplicial complex, one of which is a generalization of the Laplacian of a graph. For the rest of this section, take $V = \{v_1, \dots, v_n\}$ to be a finite set of cardinality n.

Let $W := \mathbb{R}^V$, let $\bigwedge^k W$ denote the k-th exterior power of W and let the boundary operator δ_k be defined as:

$$\begin{aligned} \delta_k : \textstyle\bigwedge^{k+1} W &\longrightarrow \textstyle\bigwedge^k W \\ s_0 \wedge \cdots \wedge s_k &\longrightarrow \textstyle\sum_{j=0}^k (-1)^j (s_0 \wedge \cdots \wedge s_k)^j \end{aligned} \tag{5.2.1}$$

where $\{s_0, \dots, s_k\} \subset \{v_1, \dots, v_n\}$ and $(s_0 \wedge \cdots \wedge s_k)^j$ means that s_j has been removed from $s_0 \wedge \cdots \wedge s_k$, leaving the others in place. The above definition makes sense because $\delta_k(s_{\sigma(0)} \wedge \cdots \wedge s_{\sigma(k)}) = \varepsilon(\sigma)\delta_k(s_0 \wedge \cdots \wedge s_k)$ for all permutation σ of $\{0, \dots, k\}$. We set $\delta_0(s_0) = 1 \in \mathbb{R} = \bigwedge^0 W$.

We have for all $k \geq 1$,

$$\delta_{k-1} \circ \delta_k = 0.$$

The real vector spaces $\bigwedge^k W$ are endowed with the standard inner product. The adjoint operator δ_k^* is given by:

$$\begin{aligned} \delta_k^* : \textstyle\bigwedge^k W &\longrightarrow \textstyle\bigwedge^{k+1} W \\ s_0 \wedge \cdots \wedge s_{k-1} &\longrightarrow \textstyle\sum_{v \in V} v \wedge s_0 \wedge \cdots \wedge s_{k-1}. \end{aligned} \tag{5.2.2}$$

Let Δ be an abstract simplicial complex on V, let $\mathcal{C}_k$ be the subspace of $\bigwedge^{k+1} W$ generated by $s_0 \wedge \cdots \wedge s_k$ where $\{s_0, \dots, s_k\} \in \Delta_k$ (in particular $\mathcal{C}_{-1} = \mathbb{R} = \bigwedge^0 W$). The dimension d_k of $\mathcal{C}_k$ is obviously equal to the number of faces of Δ of dimension k. We remark that $\delta_k(\mathcal{C}_k) \subset \mathcal{C}_{k-1}$. We keep the same notation δ_k for the restricted application $\delta_k : \mathcal{C}_k \to \mathcal{C}_{k-1}$ and δ_k^* for its adjoint operator, which is given by

$$\delta_k^*(s_0 \wedge \cdots \wedge s_{k-1}) = \sum_{v \in V \;:\; \{v, s_0, \dots, s_{k-1}\} \in \Delta} v \wedge s_0 \wedge \cdots \wedge s_{k-1} \ .$$

We introduce

$$\mathcal{L}_{k-1} := \delta_k \delta_k^* : \mathcal{C}_{k-1} \longrightarrow \mathcal{C}_{k-1}$$

and

$$\mathcal{L}'_{k-1} := \delta_{k-1}^* \delta_{k-1} : \mathcal{C}_{k-1} \longrightarrow \mathcal{C}_{k-1}.$$

In order to write $\mathcal{L}_{k-1}$ and $\mathcal{L}'_{k-1}$ in matrix form, we need to choose basis of the spaces $\mathcal{C}_k$. For that we put an arbitrary order relation among elements of V and we will express all matrices of linear transformations of the spaces $\mathcal{C}_k$ in the standard basis, given by the faces $s_0 \wedge \cdots \wedge s_k$ where $s_0 < \cdots < s_k$. Let $N_k \in \mathbb{R}^{d_{k-1} \times d_k}$ be the matrix of δ_k. We have $\delta_{k-1} \circ \delta_k = 0$ so $N_{k-1}N_k = 0$. The matrices L_{k-1} of $\mathcal{L}_{k-1}$ and L'_{k-1} of $\mathcal{L}'_{k-1}$ are then:

$$L_{k-1} := N_k N_k^T \qquad L'_{k-1} := N_{k-1}^T N_{k-1} \tag{5.2.3}$$

We have

$$L_{k-1},\ L'_{k-1} \in \mathbb{R}^{d_{k-1} \times d_{k-1}} \qquad L_{k-1}L'_{k-1} = L'_{k-1}L_{k-1} = 0. \tag{5.2.4}$$

For all $s \in \Delta_{k-1}$, let $\deg(s)$ denote the number of k-dimensional faces that contain s. We note that the number of $(k-2)$-dimensional faces that are contained in s is necessarily equal to k as all subsets of s belong to Δ. Moreover, for $s, s' \in \binom{V}{k}$, such that $|s \cap s'| = k-1$, we define a sign rule $\varepsilon(s, s') \in \{-1, +1\}$ as follows: let $s \cup s' = \{s_0 < \cdots < s_k\} \in \binom{V}{k+1}$ and $(s \cup s') \backslash (s \cap s') = \{s_i, s_{i'}\}$. Then $\varepsilon(s, s') := (-1)^{i+i'-1}$. We compute, for $s, s' \in \Delta_{k-1}$:

$$\begin{aligned} (L'_{k-1})_{s,s'} &= \sum_{r \in \Delta_{k-2}} (N_{k-1})_{r,s}(N_{k-1})_{r,s'} \\ &= \begin{cases} k & \text{if } s = s' \\ \varepsilon(s,s') & \text{if } |s \cup s'| = k+1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{5.2.5}$$

and

$$\begin{aligned} (L_{k-1})_{s,s'} &= \sum_{t \in \Delta_k} (N_k)_{s,t}(N_k)_{s',t} \\ &= \begin{cases} \deg(s) & \text{if } s = s' \\ -\varepsilon(s,s') & \text{if } s \cup s' \in \Delta_k \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{5.2.6}$$

We remark that the entries corresponding to (s, s') with $|s \cup s'| \geq k+2$ are equal to zero in both matrices, and that $(L_{k-1})_{s,s'} = -(L'_{k-1})_{s,s'}$ if $s \cup s' \in \Delta_k$.

5.2.1 Examples

Example 1: Let $\mathcal{G} = (V, E)$ be a graph and take $\Delta = \{E, V, \emptyset\}$. We have $\delta_1(s_0 \wedge s_1) = s_1 - s_0$ and $\delta_0(s_k) = 1$.

$$\mathcal{C}_1 \xrightarrow[N_1]{\delta_1} \mathcal{C}_0 \xrightarrow[N_0]{\delta_0} \mathbb{R} \tag{5.2.7}$$

We have

$$\mathcal{L}'_0(v) = \delta_0^* \delta_0(v) = \delta_0^*(1) = \sum_{u \in V} u$$

while

$$\begin{aligned}\mathcal{L}_0(v) = \delta_1\delta_1^*(v) &= \delta_1(\sum_{u\in V\ :\ uv\in E} u\wedge v)\\ &= \sum_{u\in V\ :\ uv\in E}(v-u)\\ &= \deg(v)v - \sum_{u\in V\ :\ uv\in E} u\end{aligned}$$

so, in terms of matrices, we see that

$$L_0' = J \quad \text{and} \quad L_0 = D - A \tag{5.2.8}$$

where D is the diagonal matrix with $\deg(u)$ on the diagonal and A is the adjacency matrix of $\mathcal{G}$. We recognize in L_0 the Laplacian of $\mathcal{G}$.

Example 2: Let $\mathcal{G} = (V, E)$ be a graph with no isolated vertex. Let Δ be the set T of 3-sets of vertices $t = \{v_1, v_2, v_3\}$ such that t contains at least an edge, together with all subsets of elements of T.

$$\mathcal{C}_2 \xrightarrow[N_2]{\delta_2} \mathcal{C}_1 \xrightarrow[N_1]{\delta_1} \mathcal{C}_0 \xrightarrow[N_0]{\delta_0} \mathbb{R} \tag{5.2.9}$$

with $\delta_2(s_0\wedge s_1\wedge s_2) = s_1\wedge s_2 - s_0\wedge s_2 + s_0\wedge s_1$. We have

$$\begin{aligned}\mathcal{L}_1'(u\wedge v) &= \delta_1^*\delta_1(u\wedge v) = \delta_1^*(v-u)\\ &= \sum_{w\in V}(w\wedge v - w\wedge u)\\ &= 2(u\wedge v) + \sum_{w\in V,\ w\neq u,v}(w\wedge v - w\wedge u)\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_1(u\wedge v) &= \delta_2\delta_2^*(u\wedge v)\\ &= \delta_1(\sum_{w\in V\ :\ uvw\in T} w\wedge u\wedge v)\\ &= \sum_{w\in V\ :\ uvw\in T}(u\wedge v - w\wedge v + w\wedge u)\\ &= \deg(uv)(u\wedge v) - \sum_{w\in V\ :\ uvw\in T}(w\wedge v - w\wedge u)\end{aligned}$$

where $\deg(uv) = |\{t\in T\ :\ uv\subset t\}|$. We note that

$$(\mathcal{L}_1 + \mathcal{L}_1')(u\wedge v) = (\deg(u\wedge v) + 2)(u\wedge v) + \sum_{w\in V\ :\ uvw\notin T}(w\wedge v - w\wedge u).$$

For the matrices,

$$(L'_1)_{x\wedge y,u\wedge v} = \begin{cases} 2 & \text{if } xy = uv \\ 0 & \text{if } |xyuv| = 4 \\ \pm 1 & \text{if } |xyuv| = 3 \end{cases} \tag{5.2.10}$$

where $|xyuv|$ stands for the the cardinal of $\{x, y, u, v\}$, and the $-$ sign occurs in the last case iff $x < y = u < v$ or $u < v = x < y$.

$$(L_1)_{x\wedge y,u\wedge v} = \begin{cases} \deg(xy) & \text{if } xy = uv \\ 0 & \text{if } |xyuv| = 4 \\ 0 & \text{if } |xyuv| = 3 \text{ and } xyuv \notin T \\ \pm 1 & \text{if } |xyuv| = 3 \text{ and } xyuv \in T \end{cases} \tag{5.2.11}$$

and the $+$ sign occurs in the last case iff $x < y = u < v$ or $u < v = x < y$.

We analyse the spectrum of L'_1: from the above expressions it is easy to check that $L'_1 N_1^T = n N_1^T$, hence $(L'_1)^2 = n L'_1$, from which we derive that L'_1 has eigenvalues n and 0. The multiplicity of n is $\mathrm{Tr}(L'_1)/n = n-1$ and the eigenspace of n is $\mathrm{Im}(N_1^T)$.

Example 3: The above example extends to $(k+1)$-subsets. We define Δ as the set T of $(k+1)$-subsets of V that contain at least an edge, together with all subsets of V with at most k elements. Then $L_{k-1},\ L'_{k-1} \in \mathbb{R}^{\binom{V}{k}\times\binom{V}{k}}$ and

$$\begin{aligned} (L'_{k-1})_{s,s'} &= \sum_{r\in\binom{V}{k-1}} (N_{k-1})_{r,s}(N_{k-1})_{r,s'} \\ &= \begin{cases} k & \text{if } |s\cup s'| = k \text{ i.e. } s = s' \\ 0 & \text{if } |s\cup s'| \geq k+2 \\ \varepsilon(s,s') & \text{if } |s\cup s'| = k+1 \end{cases} \end{aligned} \tag{5.2.12}$$

and

$$\begin{aligned} (L_{k-1})_{s,s'} &= \sum_{t\in T} (N_k)_{s,t}(N_k)_{s',t} \\ &= \begin{cases} \deg(s) & \text{if } |s\cup s'| = k \text{ i.e. } s = s' \\ 0 & \text{if } |s\cup s'| \geq k+2 \\ 0 & \text{if } |s\cup s'| = k+1 \text{ and } s\cup s' \notin T \\ -\varepsilon(s,s') & \text{if } |s\cup s'| = k+1 \text{ and } s\cup s' \in T \end{cases} \end{aligned} \tag{5.2.13}$$

We note that $\{\Delta_i,\ i \leq k-1\}$ is the complete complex which has trivial homology. We analyse the spectrum of L'_{k-1}: from (5.2.5) it is easy to check that $L'_{k-1} N_{k-1}^T = n N_{k-1}^T$ hence $(L'_{k-1})^2 = n L'_{k-1}$ from which we derive that L'_{k-1} has eigenvalues n and 0. The multiplicity of n is $\mathrm{Tr}(L'_{k-1})/n = k\binom{n}{k}/n$ and the eigenspace of

n is $\mathrm{Im}(N_{k-1}^T)$.

Example 4: Another complex of great interest is the cliques complex, where for $i \leq k$, Δ_i is the set Δ_i^{cl} of cliques on $(i+1)$ vertices in the graph. We note that the union of two k-cliques that intersect on $k-1$ vertices is a clique if and only if the two vertices that do not belong to the intersection are connected. The formulas for the coefficients of L_{k-1} and L'_{k-1} are the same as above.

Example 5: More generally, if $\mathcal{G}$ is a $(k+1)$-uniform hypergraph, i.e. a set of vertices V together with a set S of $(k+1)$-subsets of V called hyperedges, we can take for Δ the set S of hyperedges together with all their subsets. For example S can be the set of induced subgraphs of $\mathcal{G}$ that belong to a given isomorphism class. The same formulas as in Example 3 give the coefficients of L_{k-1} and L'_{k-1}. The matrix L_{k-1} (sometimes $L_{k-1} + L'_{k-1}$) is called the Laplacian of $\mathcal{G}$.

5.3 Generalized ϑ numbers

Recall from chapter 2 the following formulations for Lovász ϑ number of a graph $\mathcal{G} = (V, E)$:

$$\begin{array}{rcllcl}
\vartheta(\mathcal{G}) & = & \max\{ & \langle J, X\rangle & : & X \succeq 0 \\
 & & & & & \langle I, X\rangle = 1 \\
 & & & & & X_{i,j} = 0 \text{ if } ij \in E \ \ \} \\
 & = & \min\{ & t & : & tI - J + \sum_{ij\in E} x_{i,j}E_{i,j} \succeq 0 \ \ \} \\
 & = & \min\{ & \lambda_{\max}(Z) & : & Z_{i,j} = 1 \text{ for } i = j \text{ and } ij \in \overline{E} \ \ \} \\
 & = & \min\{ & \lambda_{\max}(Z) & : & Z = J + T,\ T_{ij} = 0 \text{ for } ij \notin E \ \ \}
\end{array}$$

By analogy, we introduce, with the notations of Example 3, a new function ϑ_k:

Definition 5.3.1. *For any graph $\mathcal{G}$, we define*

$$\begin{array}{rlll}
\vartheta_k(\mathcal{G}) := \max\Big\{\langle L'_{k-1}, X\rangle & : & X \succeq 0, & \\
 & & \langle I, X\rangle = 1, & \\
 & & X_{s,s'} = 0 & \textit{if } |s \cup s'| \geq k+2 \textit{ or} \\
 & & & s \cup s' \textit{ contains an edge} \\
 & & \varepsilon(s,s')X_{s,s'} = \varepsilon(t,t')X_{t,t'} & \textit{if } s \cup s' = t \cup t' \\
 & & & \textit{and } |s \cup s'| = k+1 \quad \Big\}
\end{array}$$

Since this program is not strictly feasible, we don't know if strong duality holds. However, its dual formulation is given in the following lemma:

Lemma 5.3.2. *For any graph $\mathcal{G}$, we have*

$$\vartheta_k(\mathcal{G}) \leq \min\left\{ t \ : \ tI - L'_{k-1} + M + \sum x_{trr'}(E_{tr} - E_{tr'}) \succeq 0 \right\}$$

where M is supported on the pairs (s, s') such that $|s \cup s'| \geq k+2$ or $s \cup s'$ contains an edge, the sum runs over the set

$$\{(t, r, r') : |t| = k+1, t \text{ without edges }, |r| = |r'| = k-1, r \neq r'\}$$

and E_{tr} denotes the matrix with a single upper non zero coefficient corresponding to indices s, s' such that $s \cap s' = r$, $s \cup s' = t$ and equal to $\varepsilon(s, s')$. Equivalently:

$$\vartheta_k(\mathcal{G}) \leq \min\left\{ \lambda_{\max}(Z) \ : \ \begin{array}{l} Z = L'_{k-1} + T, \\ T_{s,s} = 0 \text{ if } s \text{ does not contain an edge} \end{array} \right\}$$

where, for $k \geq 2$, T verifies $\varepsilon(s, s')T_{s,s'} + \varepsilon(v, v')T_{v,v'} = 0$ for all $\{s, s'\} \neq \{v, v'\}$ such that $s \cup s' = v \cup v'$ has cardinality $k+1$ and contains no edges.

At this point, several remarks are needed.

Remark 5.3.3. Clearly, $\vartheta_1 = \vartheta$. For $k = 2$ we are considering a matrix X whose rows and columns are indexed on pair of vertices; the only entries of X that are not set to zero are the diagonal ones corresponding to an index which does not contain an edge, and the ones whose union of row index and column index is a triangle without edges. In this last case, entries corresponding to the same triangle have the same absolute value, the sign is given according to the rule we introduced in section 5.2.

Remark 5.3.4. If we partition the set $\binom{V}{k}$ into two subsets, the subset of cocliques and the subset of k-tuples containing at least one edge, then a feasible matrix X for (5.3.1) satisfies $X_{s,s'} = 0$ if one of s, s' is in the second part. So X has the form:

$$X = \left(\begin{array}{c|c} X' & 0 \\ \hline 0 & 0 \end{array}\right)$$

and we can replace in $\vartheta_k(\mathcal{G})$ all matrices by matrices indexed by cocliques. In this way we obtain a program which is equivalent to ϑ_k. Moreover in this case, strong duality holds with the same argument as for ϑ. In particular, for $k = 2$, the restricted matrix X' is indexed by the set $\overline{E}$ of non-adjacent pairs.

Remark 5.3.5. Note that $\vartheta_k(\mathcal{G}) = -\infty$ whenever $k > \alpha(\mathcal{G})$, so from now on we restrict ourselves to the cases $k \in \{1, 2, \ldots, \alpha(\mathcal{G})\}$.

We want to construct some feasible matrices Z for the dual formulation of ϑ_k. More precisely, because L'_{k-1} and L_{k-1} commute, we want Z to be a linear combination of L'_{k-1}, and powers of L_{k-1}. So we introduce:

$$\vartheta_k^{LP}(\mathcal{G}) := \min\Big\{ \lambda_{\max}(Z) \ : \ \begin{array}{l} Z = \gamma_{-1}L'_{k-1} + \sum_{i\geq 0}\gamma_i L^i_{k-1} \\ (\gamma_{-1}-1)L'_{k-1} + \sum_{i\geq 0}\gamma_i L^i_{k-1})_{s,s'} = 0 \text{ for} \\ (|s\cup s'| \leq k+1 \text{ containing no edges} \end{array} \Big\}$$

Then $\vartheta_k^{LP}(\mathcal{G})$ is the optimal value of a linear program.

5.3.1 Properties of ϑ_k

We are going to see that the number ϑ_k conserve some well-known properties of ϑ, recalled in section 2.2.

Proposition 5.3.6. *For all $k \leq \alpha(\mathcal{G})$, $\alpha(\mathcal{G}) \leq \vartheta_k(\mathcal{G})$, with equality when $k = \alpha(\mathcal{G})$.*

Proof. Let S be an independent set of $\mathcal{G}$. Let $V_k^{\mathcal{G}}$ denote the set of all k-subsets of vertices of $\mathcal{G}$. Define the matrix $X_k^S \in \mathbb{R}^{V_k^{\mathcal{G}}\times V_k^{\mathcal{G}}}$ by:

$$(X_k^S)_{s,s'} = \begin{cases} (L'_{k-1})_{s,s'} & \text{if } s\cup s' \subset S \\ 0 & \text{else} \end{cases}$$

Because X_k^S is a submatrix of L'_{k-1}, it is positive semidefinite. We compute

$$\langle X_k^S, I\rangle = k\binom{|S|}{k} \tag{5.3.1}$$

and, with (5.2.12),

$$\begin{aligned} \langle X_k^S, L'_{k-1}\rangle &= k^2\binom{|S|}{k} + \sum_{\substack{|s\cup s'|=k+1 \\ s\cup s'\subseteq S}} 1 \\ &= k^2\binom{|S|}{k} + (k+1)k\binom{|S|}{k+1} \\ &= k\binom{|S|}{k}|S|. \end{aligned} \tag{5.3.2}$$

Moreover, from the definition of X_k^S and the fact that S is an independent set, it is clear that $(X_k^S)_{s,s'} = 0$ if $s\cup s'$ contains an edge, or if $|s\cup s'| \geq k+2$. So the matrix $X_k^S/(k\binom{|S|}{k})$ is feasible for $\vartheta_k(\mathcal{G})$, its objective value is $|S|$ and we can conclude that

$\alpha(\mathcal{G}) \leq \vartheta_k(\mathcal{G})$.
Now, if $k = \alpha(\mathcal{G})$, a subset of $V(\mathcal{G})$ of size greater or equal to $k+1$ must contain an edge. So, if $s \neq s' \in V_k^{\mathcal{G}}$, $s \cup s'$ contains an edge, and the non diagonal coefficients of a primal feasible matrix X of (5.3.1) are equal to zero. So, taking account of the fact that the diagonal coefficients of L'_{k-1} are equal to k, we have $\langle L'_{k-1}, X\rangle = k\langle I, X\rangle = k = \alpha(\mathcal{G})$. □

For the rest of the chapter we take the following notation. Let $V_k^{\mathcal{G}}$ denote the set of all k-subsets of vertices of $\mathcal{G}$. For any k, a graph homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$ induces a linear application $\tilde{\varphi}_k$ of $\mathbb{R}^{V_k^{\mathcal{G}} \times V_k^{\mathcal{G}}}$ into $\mathbb{R}^{V_k^{\mathcal{H}} \times V_k^{\mathcal{H}}}$, which goes as follows:

$$\tilde{\varphi}_k(X)_{u,u'} := \sum_{\substack{s \,:\, \varphi(s)=u \\ s' \,:\, \varphi(s')=u'}} X_{s,s'}$$

where $u, u' \in V_k^{\mathcal{H}}$ and $s, s' \in V_k^{\mathcal{G}}$.

Then the next theorem generalizes proposition 2.3.12.

Theorem 5.3.7. *If there exists a graph homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$, then, for any k, $\vartheta_k(\overline{\mathcal{G}}) \leq \vartheta_k(\overline{\mathcal{H}})$.*

Proof. We begin the proof by two remarks. The first is that, given an order on the vertices of $\mathcal{H}$, we can order the vertices of $\mathcal{G}$ in such a way that φ becomes an increasing function; it follows that $\varepsilon(\varphi(s), \varphi(s')) = \varepsilon(s, s')$ for all $s, s' \in V_k^{\mathcal{G}}$ such that $|s \cup s'| = k+1$, $\varphi(s), \varphi(s') \in V_k^{\mathcal{H}}$, $|\varphi(s) \cup \varphi(s')| = k+1$.
Now, let X be a feasible solution of (5.3.1) with respect to the graph $\overline{\mathcal{G}}$. The second remark is that $X_{s,s'} = 0$ if $\varphi(s) \notin V_k^{\mathcal{H}}$ or $\varphi(s') \notin V_k^{\mathcal{H}}$, because in that case s or s' contains an edge of $\overline{\mathcal{G}}$. For the same reason, $X_{s,s'} = 0$ if $s \neq s'$ are such that $\varphi(s) = \varphi(s')$. We will use these two facts all along the proof. Our goal now is to show that the matrix $\tilde{\varphi}_k(X)$ is feasible for the program (5.3.1) with respect to the graph $\overline{\mathcal{H}}$. Such matrix is clearly symmetric, let us see that it is positive semidefinite: for all vector $\alpha \in \mathbb{R}^{V_k^{\mathcal{H}}}$,

$$\begin{aligned}
\sum_{u,u'} \alpha_u (\tilde{\varphi}_k(X))_{u,u'}\, \alpha_{u'} &= \sum_{u,u'} \alpha_u \Big(\sum_{\substack{s \,:\, \varphi(s)=u \\ s' \,:\, \varphi(s')=u'}} X_{s,s'} \Big)\, \alpha_{u'} \\
&= \sum_{\substack{s,s' \in V_k^{\mathcal{G}} \\ \varphi(s),\varphi(s') \in V_k^{\mathcal{H}}}} \alpha_{\varphi(s)} X_{s,s'}\, \alpha_{\varphi(s')} \\
&= \sum_{s,s' \in V_k^{\mathcal{G}}} \alpha_{\varphi(s)} X_{s,s'}\, \alpha_{\varphi(s')} \geq 0
\end{aligned}$$

Let us see also that $\tilde{\varphi}_k(X)$ has trace 1:

$$\begin{aligned}\mathrm{Tr}(\tilde{\varphi}_k(X)) &= \sum_u (\tilde{\varphi}_k(X))_{u,u} = \sum_u \sum_{\substack{s\ :\ \varphi(s)=u \\ s'\ :\ \varphi(s')=u}} X_{s,s'} \\ &= \sum_u \sum_{s\ :\ \varphi(s)=u} X_{s,s} = \sum_{s\ :\ \varphi(s)\in V_k^{\mathcal{H}}} X_{s,s} = \mathrm{Tr}(X) = 1\ .\end{aligned}$$

Next, if $|u \cup u'| \geq k+2$, then $|s \cup s'| \geq k+2$ for all s, s' such that $\varphi(s) = u$ and $\varphi(s') = u'$, hence $(\tilde{\varphi}_k(X))_{u,u'} = 0$. Also, $s \cup s'$ contains an edge of $\overline{\mathcal{G}}$ if $u \cup u'$ contains an edge of $\overline{\mathcal{H}}$, so $(\tilde{\varphi}_k(X))_{u,u'} = 0$ also in this case.
Finally, take $u \cup u' = v \cup v'$ of cardinality $k+1$. Then the collection of $s \cup s'$ where (s, s') is an inverse image of (u, u') and the collection of $t \cup t'$ where (t, t') is an inverse image of (v, v'), coincide. So we verify that

$$\begin{aligned}\varepsilon(u,u')(\tilde{\varphi}_k(X))_{u,u'} &= \sum_{\substack{\varphi(s)=u \\ \varphi(s')=u'}} \varepsilon(u,u') X_{s,s'} \\ &= \sum_{\substack{\varphi(s)=u \\ \varphi(s')=u'}} \varepsilon(s,s') X_{s,s'} \\ &= \sum_{\substack{\varphi(t)=v \\ \varphi(t')=v'}} \varepsilon(t,t') X_{t,t'} \\ &= \sum_{\substack{\varphi(t)=v \\ \varphi(t')=v'}} \varepsilon(v,v') X_{t,t'} = \varepsilon(v,v')(\tilde{\varphi}_k(X))_{v,v'}\ .\end{aligned}$$

Up to now, we have shown that the matrix $\tilde{\varphi}_k(X)$ is a feasible solution of the program (5.3.1) with respect to the graph $\overline{\mathcal{H}}$. Now we calculate its objective value:

$$\begin{aligned}\langle L'_{k-1}, \tilde{\varphi}_k(X)\rangle &= 2 + \sum_{u\neq u'} \varepsilon(u,u')(\tilde{\varphi}_k(X))_{u,u'} \\ &= 2 + \sum_{u\neq u'} \sum_{\substack{\varphi(s)=u \\ \varphi(s')=u'}} \varepsilon(u,u') X_{s,s'} \\ &= 2 + \sum_{u\neq u'} \sum_{\substack{\varphi(s)=u \\ \varphi(s')=u'}} \varepsilon(s,s') X_{s,s'} \\ &= 2 + \sum_{s\neq s'} \varepsilon(s,s') X_{s,s'} = \langle L'_{k-1}, X\rangle\ .\end{aligned}$$

Every feasible matrix for $\vartheta_k(\overline{\mathcal{G}})$ gives a feasible matrix for $\vartheta_k(\overline{\mathcal{H}})$ with the same objective value. This concludes the proof. □

Now we want to establish an analog of the sandwich theorem 2.3.10 for every ϑ_k. For this, we need the following equivalent definition of the chromatic number of a graph.

Proposition 5.3.8. *The chromatic number $\chi(\mathcal{G})$ of a graph $\mathcal{G}$ is the least natural number ℓ such that there exists a homomorphism $\mathcal{G} \to \mathcal{K}_\ell$ where $\mathcal{K}_\ell$ denotes the complete graph on ℓ vertices.*

Lemma 5.3.9. *If $\ell = \chi(\overline{\mathcal{G}})$, then $\vartheta_k(\mathcal{G}) \leq \vartheta_k(\overline{\mathcal{K}_\ell})$.*

Proof. By previous definition and theorem 5.3.7. □

Proposition 5.3.10. *If n is the number of vertices of $\mathcal{G}$, then $\vartheta_k(\mathcal{G}) \leq n$ for all k.*

Proof. A feasible solution of the dual program is given by the matrix L'_{k-1} itself, which has largest eigenvalue n. □

Corollary 5.3.11. *For all k, $\vartheta_k(\overline{\mathcal{K}_\ell}) = \ell$.*

Proof. By proposition 5.3.6 together with proposition 5.3.10. □

From all previous results, we obtain the *sandwich theorem* for ϑ_k.

Theorem 5.3.12. *For all graph $\mathcal{G}$ and all k, $\alpha(\mathcal{G}) \leq \vartheta_k(\mathcal{G}) \leq \chi(\overline{\mathcal{G}})$.*

Remark 5.3.13. To prove corollary 5.3.11 we only need the special case of proposition 5.3.6 when $\mathcal{G} = \overline{\mathcal{K}_\ell}$. Then the general inequality $\alpha(\mathcal{G}) \leq \vartheta_k(\mathcal{G})$ can also be derived from theorem 5.3.7 and corollary 5.3.11 via the clique number ω. By definition, $\omega(\mathcal{G})$ is the largest cardinality of a clique of $\mathcal{G}$ and it can also be defined as the largest natural number m such that there exists a homomorphism $\mathcal{K}_m \to \mathcal{G}$. As independent sets correspond to cliques in the complement graph, $\omega(\overline{\mathcal{G}}) = \alpha(\mathcal{G})$.

5.4 A decreasing hierarchy

We have seen that the sequence of programs ϑ_k defined in the previous section has some remarkable properties, namely the sandwich theorem and the fact that for any graph $\mathcal{G}$ we reach the exact value $\alpha(\mathcal{G})$ at the $\alpha(\mathcal{G})$-th step. Nevertheless another desirable property for a sequence of this kind would be to be decreasing as a function of k. At the moment we are not able neither to prove nor to disprove this in the general case, although practical computations show $\vartheta_2 < \vartheta_1$. So in this section we consider a slightly different program obtained from ϑ_k by adding the requirement that certain matrices are positive semidefinite. We show that this new sequence of programs conserve all properties of the original one and, moreover, it is decreasing with k.

Recall that $V_k^{\mathcal{G}}$ denotes the set of all k-subsets of vertices of $\mathcal{G}$. For each $k \geq 2$ we define the following linear application

$$\tau_{k,k-1} : \mathbb{R}^{V_k^{\mathcal{G}} \times V_k^{\mathcal{G}}} \to \mathbb{R}^{V_{k-1}^{\mathcal{G}} \times V_{k-1}^{\mathcal{G}}}$$

by

- if $u \in V_{k-1}^{\mathcal{G}}$, then $\tau_{k,k-1}(X)_{u,u} := \frac{1}{k} \sum_{u \subset s} X_{s,s}$
- if $u, u' \in V_{k-1}^{\mathcal{G}}$ are such that $|u \cup u'| = k$, then

$$\varepsilon(u,u')\tau_{k,k-1}(X)_{u,u'} := \frac{1}{k(k-1)} X_{u\cup u',u\cup u'} + \frac{1}{(k+1)k(k-1)} \sum_{\substack{|s\cup s'|=k+1 \\ u\cup u' \subset s\cup s'}} \varepsilon(s,s') X_{s,s'}$$

- $\tau_{k,k-1}(X)_{u,u'} := 0$ elsewhere.

Remark that $\varepsilon(u,u')\tau_{k,k-1}(X)_{u,u'}$ is a function of $u \cup u'$. For each $k > h \geq 2$, we denote by $\tau_{k,h}$ the composition $\tau_{h+1,h} \circ \cdots \circ \tau_{k-1,k-2} \circ \tau_{k,k-1}$.

Definition 5.4.1. *For each graph $\mathcal{G}$ we define $\hat{\vartheta}_k(\mathcal{G})$ as the optimal value of the program 5.3.1 to which we add the constraints $\tau_{k,k-1}(X) \succeq 0$, $\tau_{k,k-2}(X) \succeq 0$, …, $\tau_{k,1}(X) \succeq 0$.*

Proposition 5.4.2. *For all $k \leq \alpha(\mathcal{G})$, $\alpha(\mathcal{G}) \leq \hat{\vartheta}_k(\mathcal{G})$, with equality when $k = \alpha(\mathcal{G})$.*

Proof. Almost all has already been proved in Proposition 5.3.6. Let S be an independent set of $\mathcal{G}$ of maximal size and recall the definition of X_k^S:

$$(X_k^S)_{s,s'} = \begin{cases} (L'_{k-1})_{s,s'} & \text{if } s \cup s' \subset S \\ 0 & \text{else} \end{cases}$$

It remains to show that all matrices $\tau_{k,h}(X_k^S)$ are positive semidefinite. It is not difficult to see that $\tau_{k,k-1}(X_k^S)_{u,u} = |S| - k + 1$ whenever $u \subset S$ and that $\tau_{k,k-1}(X_k^S)_{u,u'} = \frac{|S|-k+1}{k-1}\varepsilon(u,u')$ whenever $u \cup u' \subset S$ and that all other entries are zero. So we recognize that $\tau_{k,k-1}(X_k^S) = \frac{|S|-k+1}{k-1} X_{k-1}^S$ which is positive semidefinite. Iterating this computation, we find that all others $\tau_{k,h}(X_k^S)$ are positive scalar multiples of X_h^S respectively, hence positive semidefinite. This concludes the proof. $\square$

Theorem 5.4.3. *For all k, $\hat{\vartheta}_k(\mathcal{G}) \leq \hat{\vartheta}_{k-1}(\mathcal{G})$.*

Proof. Let X be a feasible solution of the program $\hat{\vartheta}_k(\mathcal{G})$ and note Y the matrix $\tau_{k,k-1}(X)$. We will see that this matrix is a feasible solution of $\hat{\vartheta}_{k-1}(\mathcal{G})$. Let us verify that Y has trace 1, as the other constraints are easily seen to be satisfied.

$$\begin{aligned}
\mathrm{Tr}(Y) &= \sum_u Y_{u,u} = \sum_u \frac{1}{k} \sum_{s\,:\,u\subset s} X_{s,s} \\
&= \sum_s X_{s,s}\Big(\frac{1}{k}\sum_{u\,:\,u\subset s} 1\Big) = \sum_s X_{s,s} = \mathrm{Tr}(X) = 1
\end{aligned}$$

The objective value of this solution is

$$\begin{aligned}
\langle L'_{k-2}, Y\rangle &= (k-1)\mathrm{Tr}(Y) + \sum_{u,u'\,:\,|u\cup u'|=k} \varepsilon(u,u')Y_{u,u'} \\
&= (k-1) + \underbrace{\sum_{s\,:\,|s|=k} X_{s,s}\Big(\frac{1}{k(k-1)}\sum_{u,u'\,:\,u\cup u'=s} 1\Big)}_{\mathrm{Tr}(X)=1} \\
&\quad + \sum_{u,u'\,:\,|u\cup u'|=k} \frac{1}{(k+1)k(k-1)} \sum_{\substack{|s\cup s'|=k+1 \\ u\cup u'\subset s\cup s'}} \varepsilon(s,s')X_{s,s'} \\
&= k + \sum_{s,s'\,:\,|s\cup s'|=k+1} \varepsilon(s,s')X_{s,s'}\Big(\frac{1}{(k+1)k(k-1)} \sum_{\substack{|u\cup u'|=k \\ u\cup u'\subset s\cup s'}} 1\Big) \\
&= k + \sum_{s,s'\,:\,|s\cup s'|=k+1} \varepsilon(s,s')X_{s,s'} \\
&= \langle L'_{k-1}, X\rangle
\end{aligned}$$

and this yields the announced inequality. □

Remark 5.4.4. Obviously, $\hat{\vartheta}_k(\mathcal{G}) \leq \vartheta_k(\mathcal{G})$. Although the optimal solutions of both programs coincide in the cases we computed, it seems that to insure the decreasing of our programs, $\hat{\vartheta}_k$ need to be considered, as the condition $\tau_{k,k-1}(X) \succeq 0$ is not always satisfied when X is a feasible solution of ϑ_k. Of course, for practical computations one can consider simply ϑ_k.

Recall that a graph homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$ induces a linear application $\tilde{\varphi}_k$ of $\mathbb{R}^{V_k^{\mathcal{G}} \times V_k^{\mathcal{G}}}$ into $\mathbb{R}^{V_k^{\mathcal{H}} \times V_k^{\mathcal{H}}}$.

Proposition 5.4.5. *The following diagram is commutative:*

$$
\begin{array}{ccc}
\mathbb{R}^{V_k^{\mathcal{G}}\times V_k^{\mathcal{G}}} & \xrightarrow{\tau_{k,k-1}} & \mathbb{R}^{V_{k-1}^{\mathcal{G}}\times V^{\mathcal{G}}k-1} \\
\Big\downarrow \tilde{\varphi}_k & & \Big\downarrow \tilde{\varphi}_{k-1} \\
\mathbb{R}^{V_k^{\mathcal{H}}\times V_k^{\mathcal{H}}} & \xrightarrow{\tau_{k,k-1}} & \mathbb{R}^{V_{k-1}^{\mathcal{H}}\times V_{k-1}^{\mathcal{H}}}
\end{array}
$$

Proof. Take $X \in \mathbb{R}^{V_k^{\mathcal{G}}\times V_k^{\mathcal{G}}}$. Let us call $Y = \tau_{k,k-1}(\tilde{\varphi}_k(X))$ and $Z = \tilde{\varphi}_{k-1}(\tau_{k,k-1}(X))$. Along this proof we will make use of the remarks which appear in the proof of theorem (5.3.7). We calculate

$$Y_{r,r} = \frac{1}{k}\sum_{v\,:\,r\subset v}(\tilde{\varphi}_k(X))_{v,v} = \frac{1}{k}\sum_{v\,:\,r\subset v}\ \sum_{s\,:\,\varphi(s)=v} X_{s,s} = \frac{1}{k}\sum_{s\,:\,r\subset\varphi(s)} X_{s,s}$$

and, for $|r\cup r'| = k$,

$$
\begin{aligned}
\varepsilon(r,r')Y_{r,r'} &= \frac{1}{k(k-1)}\tilde{\varphi}_k(X)_{r\cup r',r\cup r'} + \frac{1}{(k+1)k(k-1)}\sum_{r\cup r'\subset v\cup v'}\varepsilon(v,v')\tilde{\varphi}_k(X)_{v,v'} \\
&= \frac{1}{k(k-1)}\sum_{\varphi(s)=r\cup r'} X_{s,s} + \frac{1}{(k+1)k(k-1)}\sum_{r\cup r'\subset v\cup v'}\ \sum_{\substack{\varphi(s)=v\\ \varphi(s')=v'}}\varepsilon(s,s')X_{s,s'}
\end{aligned}
$$

On the other hand,

$$Z_{r,r} = \sum_{u\,:\,\varphi(u)=r}\frac{1}{k}\sum_{u\,:\,u\subset s} X_{s,s} = \frac{1}{k}\sum_{s\,:\,\varphi^{-1}(r)\subset s} X_{s,s}$$

and, for $|r\cup r'| = k$,

$$
\begin{aligned}
Z_{r,r'} &= \sum_{\substack{\varphi(u)=r\\ \varphi(u')=r'}}(\tau_{k,k-1}(X))_{u,u'} \\
&= \sum_{\substack{\varphi(u)=r\\ \varphi(u')=r'}}\frac{1}{k(k-1)}X_{u\cup u',u\cup u'} + \sum_{\substack{\varphi(u)=r\\ \varphi(u')=r'}}\frac{1}{(k+1)k(k-1)}\sum_{u\cup u'\subset s\cup s'}\varepsilon(s,s')X_{s,s'}
\end{aligned}
$$

Now it is easy to see that $Y = Z$. □

Corollary 5.4.6. *If there exists a graph homomorphism* $\varphi : \mathcal{G} \to \mathcal{H}$*, then, for any* k*,* $\hat{\vartheta}_k(\overline{\mathcal{G}}) \leq \hat{\vartheta}_k(\overline{\mathcal{H}})$*.*

Proof. It follows from theorem 5.3.7 and proposition 5.4.5, taking account of the fact that $\tilde{\varphi}_k$ preserves positivity of matrices, as it is shown in the proof of theorem 5.3.7. $\square$

The *sandwich theorem* for $\hat{\vartheta}_k$ holds as well with the same proof as for ϑ_k.

Theorem 5.4.7. *For all graph* $\mathcal{G}$ *and all* k*,* $\alpha(\mathcal{G}) \leq \hat{\vartheta}_k(\mathcal{G}) \leq \chi(\overline{\mathcal{G}})$*.*

5.5 Numerical results

In this section we first introduce another program strictly related to ϑ_k and then give some numerical results on two classical families of graphs.

$$\vartheta_k^+(\mathcal{G}) := \max\Big\{\langle L_{k-1}^+, X\rangle \;:\; \begin{array}{lll} X \succeq 0, & & \\ \langle I, X\rangle = 1, & & \\ X_{s,s'} = 0 & \text{if } |s \cup s'| \geq k+2 \text{ or} & \\ & s \cup s' \text{ contains an edge,} & \\ X_{s,s'} = X_{t,t'} & \text{if } s \cup s' = t \cup t' & \\ & \text{and } |s \cup s'| = k+1 & \Big\} \end{array} \tag{5.5.1}$$

where all matrices are indexed on $V_k^{\mathcal{G}}$ and $L_{k-1}^+ := (N_{k-1}^+)^T (N_{k-1}^+)$ with $N_{k-1}^+ \in \mathbb{R}^{V_{k-1}^{\mathcal{G}} \times V_k^{\mathcal{G}}}$ defined by $(N_{k-1}^+)_{r,s} := 1$ if and only if $r \subset s$, and $(N_{k-1}^+)_{r,s} := 0$ elsewhere. So

$$(L_{k-1}^+)_{s,s'} = \begin{cases} k & \text{if } |s \cup s'| = k \text{ i.e. } s = s' \\ 0 & \text{if } |s \cup s'| \geq k+2 \\ 1 & \text{if } |s \cup s'| = k+1 \end{cases} \tag{5.5.2}$$

In other words, ϑ_k^+ is the program ϑ_k on which we have forgotten all sign relations. Proposition 5.3.6 and theorem 5.3.7 hold also in this case. Although this program does not have a clear algebraic description, it is interesting for computations, as it seems to be competitive with the original one.

In table 5.1 we compare the optimal values of $\vartheta = \vartheta_1$, ϑ_2 and ϑ_2^+ for some cycle graphs C_n and Paley graphs $P(n)$. See next chapter for more details on this last family.

Note: All values of the semidefinite programs have been obtained with the solvers SDPA or SDPT3, available on the NEOS website (http://www.neos-server.org/neos/).

	ϑ	ϑ_2	ϑ_2^+
C_5	2.236	2	2
C_7	3.317	3	4
C_9	4.36	4.019	4.119
C_{11}	5.386	5	5
C_{15}	7.417	7.007	7.021
C_{23}	11.446	11.003	11.016
$P(13)$	3.605	3	4
$P(17)$	4.123	3.242	4.044
$P(29)$	5.385	4.443	4.228
$P(37)$	6.082	4.611	4.354
$P(41)$	6.403	5.177	5
$P(53)$	7.280	5.438	5.756

Table 5.1: Values of ϑ, ϑ_2, ϑ_2^+ on some cycles and Paley graphs

5.6 Bounds for generalized independence numbers

We define the generalized independence number $\alpha_k(\mathcal{G})$ to be the maximal number of elements of a subset S of $V(\mathcal{G})$ such that S does not contain any $(k+1)$-clique. So $\alpha_1(G)$ is the usual independence number $\alpha(\mathcal{G})$. Here we use a degenerated version of the cliques complex considered in Example 4. More precisely, we take for Δ_k the set of $(k+1)$-cliques and for Δ_{k-1} the set of all k-subsets. We note that then L'_{k-1} is the same as before, while

$$L_{k-1} = \left(\begin{array}{c|c} L^{cl}_{k-1} & 0 \\ \hline 0 & 0 \end{array}\right)$$

where we have partitioned the index set Δ_{k-1} in cliques and non-cliques, and L^{cl}_{k-1} is the Laplacian matrix for the cliques complex. Then:

$$\vartheta_k^{cl}(\mathcal{G}) := \max\left\{ \langle L'_{k-1}, X\rangle \ : \ \begin{array}{ll} X \succeq 0, & \\ \langle I, X\rangle = 1, & \\ X_{s,s'} = 0 & \text{if } |s\cup s'| \geq k+2 \text{ or} \\ & s\cup s' \in \Delta_k \\ \varepsilon(s,s')X_{s,s'} = \varepsilon(t,t')X_{t,t'} & \text{if } s\cup s' = t\cup t' \\ & \text{and } |s\cup s'| = k+1 \end{array} \right\} \tag{5.6.1}$$

Proposition 5.6.1. *For all $k \leq \alpha_k(\mathcal{G})$, $\alpha_k(\mathcal{G}) \leq \vartheta_k^{cl}(\mathcal{G})$*

Proof. The same as for Proposition 5.3.6. □

Open questions

In our opinion, this chapter has some points that deserve further discussion.

- First of all, can one prove or disprove the decreasing property for ϑ_k? Alternatively, is there an easier practical way to insure it?
- Secondly, it would be nice to have an explicit formula of $\vartheta_2(\mathcal{G})$ when $\mathcal{G}$ runs over some families of graphs, e.g. cycle graphs.
- Next, can one find an analog of Hoffman bound for ϑ_2? If yes, which property should the graph have in order to reach equality in the given bound?
- Finally, the relation with existing hierarchies has still to be investigated.

Chapter 6

Some results on Paley graphs

In this final chapter we focus on Paley graphs. It is an open and interesting problem to determine their clique number, also in view of number theoretical applications. As these graphs are self complementary, the clique number coincides with the independence number. In the second section we symmetrize our program ϑ_2 for Paley graphs and we calculate some explicit values. In the third section we explore the circulant structure of a Paley graph and its subgraphs. We show that we can reduce the problem of determining the independence number from the original graph to the induced subgraph of non squares. In the fourth and last section, we show that the ϑ number of a circulant graph is in fact a linear program.

6.1 Introduction

Definition 6.1.1. *Let $q \equiv 1 \mod 4$ be a power of a prime. The Paley graph of order q is $P(q) := (V, E)$ where $V = \mathbb{F}_q$ and $E = \{xy : x - y \text{ is a square } \mod q, x \neq y\}$.*

In other words, $P(q)$ is the Cayley graph on the additive group $\mathbb{F}_q$ with generating set Q, the subgroup of squares in $\mathbb{F}_q^{\times}$. The condition $q \equiv 1 \mod 4$ is equivalent to say that -1 is a square modulo q and this is needed in order to get undirected edges. The Paley graph $P(q)$ is isomorphic to its complement, via multiplication by a non square. The automorphism group of $P(q)$ contains $\Gamma(q) := (\mathbb{F}_q, +) \rtimes Q$ and acts transitively on the edges as well as on the non edges. So $P(q)$ is a strongly regular graph. If $q = 4t + 1$, its parameters are $v = 4t + 1$, $k = 2t$, $\lambda = t - 1$, $\mu = t$. Being strongly regular, its ϑ number reaches the Hoffman bound

$$\vartheta(P(q)) = \sqrt{q}$$

Noting that $\vartheta(P(q))\vartheta(\overline{P(q)}) = |V| = q$ because $P(q)$ is vertex transitive and that $P(q) \simeq \overline{P(q)}$ leads to the same conclusion.

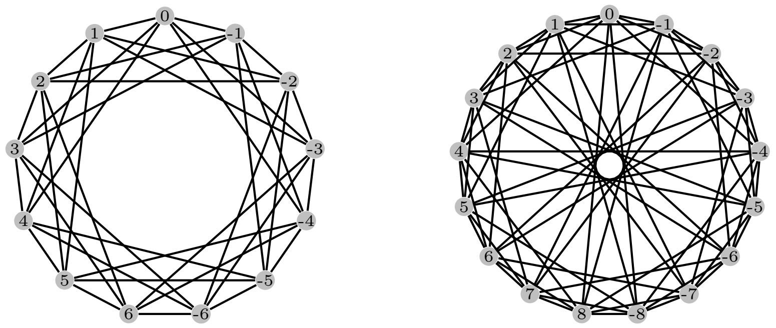

Figure 6.a: Paley graphs of order 13 and 17

Paley graphs are interesting because of their quasi-random behaviour. The independence number of Paley graphs $P(p)$ has been calculated for primes p up to 7000 by James B. Shearer ([1]); Geoffrey Exoo extended it to $p < 10000$ ([2]). If $q = q_0^2$ is a square, then the subfield $\mathbb{F}_{q_0}$ is a clique of size $\sqrt{q}$ so in this case $\alpha(P(q)) = \sqrt{q}$. If q is not a square, $\alpha(P(q))$ is not known.

In the remaining we restrict to the case when $q = p$ is a prime number. We already calculated some values of ϑ and ϑ_2 for Paley graphs in table 5.1. In order to compute more values of $\vartheta_2(P(p))$ we need to symmetrize it under $\mathrm{Aut}(P(p))$. We analyse this action in the next section.

6.2 Symmetry reduction of ϑ_2 for Paley graphs

Let us take the following notations: $\mathbb{Z}_p := \mathbb{Z}/p\mathbb{Z}$, $\Gamma_0 := \mathbb{Z}_p \rtimes Q$ and $\Gamma := \mathbb{Z}_p \rtimes \mathbb{Z}_p^\times$. We let α denote a primitive element in $\mathbb{Z}_p$. An element of Γ is denoted (t, s) with $t \in \mathbb{Z}_p$ and $s \in \mathbb{Z}_p^\times$. The group multiplication is $(t, s)(t', s') = (t + st', ss')$ and the inverse is $(t, s)^{-1} = (-ts^{-1}, s^{-1})$. The action of Γ on $\mathbb{Z}_p$ is given by $(t, s).x = t + sx$. We sometimes simplify $(t, 1)$ to t and $(0, s)$ to s.

6.2.1 Orbits

The group Γ_0 acts transitively on $V := \mathbb{Z}_p$. The 2-subsets split into two orbits: the edges and the non-edges of $P(p)$. Let us now consider the orbits on 3-subsets (also called triangles). It is clear that the number of edges in each triangle is an invariant. For $i = 0, 1, 2, 3$, let $\mathcal{T}_i$ denote the set of triangles containing i edges. Multiplication by α sends bijectively $\mathcal{T}_0$ to $\mathcal{T}_3$ and $\mathcal{T}_1$ to $\mathcal{T}_2$. Using the property that $P(p)$ is strongly regular with parameters $(p, (p-1)/2, (p-5)/4, (p-1)/4)$, we count $|\mathcal{T}_3| = |\mathcal{T}_0| = p(p-1)(p-5)/48$ and $|\mathcal{T}_1| = |\mathcal{T}_2| = p(p-1)^2/16$. The following lemma describes the orbits of Γ_0 on $\mathcal{T}_0$ and on $\mathcal{T}_1$:

Lemma 6.2.1. *For the action of* Γ_0 *on* $\mathcal{T}_1$*, we have:*

(1) Every orbit contains an element of the type $\{0, 1, \beta\}$ *where* $\beta \in \overline{Q}$ *and* $\beta - 1 \in \overline{Q}$.

(2) The triangle $\{0, 1, \beta'\}$ *is in the same orbit as* $\{0, 1, \beta\}$ *if and only if* β' *belongs to* $\{\beta, 1 - \beta\}$.

(3) The stabiliser of $\{0, 1, \beta\}$ *is trivial, unless* $2 \in \overline{Q}$ *and* $\beta = 2^{-1}$*, in which case it is of order* 2 *and is equal to* $\{(0, 1), (1, -1)\}$.

For the action of Γ_0 *on* $\mathcal{T}_0$*, we have:*

(1) Every orbit contains an element of the type $\{0, \alpha, \beta\}$ *where* $\beta \in \overline{Q}$ *and* $\beta - \alpha \in \overline{Q}$.

(2) The triangle $\{0, \alpha, \beta'\}$ *is in the same orbit as* $\{0, \alpha, \beta\}$ *if and only if* β' *belongs to* $\{\beta, \alpha^2\beta^{-1}, \alpha - \beta, \alpha^2(\alpha - \beta)^{-1}, \alpha\beta^{-1}(\beta - \alpha), \alpha\beta(\beta - \alpha)^{-1}\}$.

(3) The stabiliser of $\{0, \alpha, \beta\}$ *is trivial, unless we are in one of the following cases:*

- **(i)** $2 \in Q$ *and* $\beta = 2\alpha, -\alpha, \alpha/2$*, all three triangles are in the same orbit, and the stabiliser is* $\{\pm 1\}$.
- **(ii)** $3 \in Q$ *and* $\alpha^2 + \beta^2 - \alpha\beta = 0$*, the two corresponding triangles are in the same orbit, and the stabiliser has order* 3 *generated by* $\gamma := (\alpha, -\alpha\beta^{-1})$.

To summarize, every orbit except maybe one or two has length $p(p-1)/2$ so the number of orbits of Γ_0 acting on $\mathcal{T}_0$ is essentially $(p-5)/24$.

6.2.2 The irreducible representations of Γ_0

The irreducible representations of Γ_0 are the following:

- We have the characters of degree one lifted from Q, through the quotient map $\Gamma_0 \to \Gamma_0/\mathbb{Z}_p \simeq Q$. There are $(p-1)/2$ such characters.
- We have the induced representations from characters of $(\mathbb{Z}_p, +)$. Let χ be a non trivial character of $(\mathbb{Z}_p, +)$. To fix ideas, let $\zeta_p \in \mathbb{C}$ be a p-th root of 1 and $\chi(t) := \zeta_p^t$. The representation $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi$ is irreducible and has degree $(p-1)/2$. It can be described as $R = \oplus_{a \in Q}\mathbb{C}e_a$ with the action $(t, 1).e_a = \chi(ta^{-1})e_a$ and $(0, s).e_a = e_{sa}$ (and thus $(t, s).e_a = \chi(ts^{-1}a^{-1})e_{sa}$). We see from its characters that among $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi^r, r \in \mathbb{Z}_p^\times$, there are two classes of isomorphism corresponding to $r \in Q$ and $r \in \overline{Q}$.

To summarize, we have $(p-1)/2$ representations of degree 1 and 2 representations of degree $(p-1)/2$. It matches with $|\Gamma_0| = p(p-1)/2 = 1^2 \cdot (p-1)/2 + ((p-1)/2)^2 \cdot 2$.

6.2.3 The decomposition of $\bigwedge^2 \mathbb{C}^{\mathbb{Z}_p}$ under Γ_0

Let $W = \mathbb{C}^{\mathbb{Z}_p} = \oplus_{i=0}^{p-1} \mathbb{C} e_i$. The action of $\gamma = (t,s) \in \Gamma_0$ on e_i is thus $\gamma e_i = e_{t+si}$ and on $e_i \wedge e_j$ it is $\gamma(e_i \wedge e_j) = e_{t+si} \wedge e_{t+sj}$. In view of the symmetrization of ϑ_2 we need a precise description of the irreducible decomposition of $\bigwedge^2 W$, namely we need an explicit orthonormal basis for each irreducible subspace, and when there are isomorphic irreducible subspaces, we need basis in which the action of Γ_0 is expressed by the same matrices.

We consider the natural ordering $0 < 1 < \cdots < (p-1)$ for the elements of $\mathbb{Z}_p$ and thus the standard basis of $\Lambda^2 W$ is $\{e_i \wedge e_j,\ 0 \leq i < j \leq (p-1)\}$.

Theorem 6.2.2. *The irreducible characters appearing in the decomposition of $\Lambda^2 W$ under the action of Γ_0 are:*

- *the $(p-1)/4$ characters lifted from characters ψ of Q such that $\psi(-1) = -1$, each with multiplicity 2,*
- *the characters $Ind_{\mathbb{Z}_p}^{\Gamma_0} \chi$ and $Ind_{\mathbb{Z}_p}^{\Gamma_0} \chi^\alpha$, where χ is a character of $(\mathbb{Z}_p, +)$, each with multiplicity $(p-1)/2$.*

Proof. Let ψ be a multiplicative character of Q. We again denote ψ its lift to Γ_0. We apply the idempotent $\frac{1}{|\Gamma_0|} \sum_{\gamma \in \Gamma_0} \psi(\gamma^{-1})\gamma$ to $e_0 \wedge e_1$ leading to

$$v_1^\psi := \frac{1}{|\Gamma_0|} \sum_{s \in Q} \psi(s^{-1}) \sum_{t \in \mathbb{Z}_p} e_t \wedge e_{t+s}$$

We see that $e_t \wedge e_{t+s} = \pm e_0 \wedge e_1$ if and only if $(t,s) = (0,1)$ or $(t,s) = (1,-1)$. In the last case $e_t \wedge e_{t+s} = e_1 \wedge e_0 = -e_0 \wedge e_1$ so $v_1^\psi \neq 0$ if and only if $\psi(-1) = -1$. Moreover, if $\psi(-1) = -1$, the expression of v_1^ψ on the standard basis is:

$$v_1^\psi = \frac{1}{|\Gamma_0|} \sum_{\substack{i<j \\ j-i \in Q}} 2\psi((j-i)^{-1})(e_i \wedge e_j).$$

So $v_1^\psi \neq 0$ and $\gamma v_1^\psi = \psi(\gamma) v_1^\psi$. That is, the space generated by v_1^ψ is G-stable and G acts on it with character ψ. In a similar way, we construct an element v_2^ψ such that $\gamma v_2^\psi = \psi(\gamma) v_2^\psi$:

$$\begin{aligned} v_2^\psi &:= \frac{1}{|\Gamma_0|} \sum_{s \in Q} \psi(s^{-1}) \sum_{t \in \mathbb{Z}_p} e_t \wedge e_{t+s\alpha} \\ &= \frac{1}{|\Gamma_0|} \sum_{\substack{i<j \\ j-i \in \overline{Q}}} 2\psi((j-i)^{-1}\alpha)(e_i \wedge e_j). \end{aligned}$$

We note that it is orthogonal to v_1^ψ.

Now we look for the subspaces isomorphic to $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi$, where χ is a character of $(\mathbb{Z}_p, +)$. We consider, for $\ell \neq 0$,

$$u_\ell^\chi := \frac{1}{p}\sum_{t\in\mathbb{Z}_p} \chi(-t)(e_t \wedge e_{t+\ell}).$$

We have that $t.u_\ell^\chi = \chi(t)u_\ell^\chi$, $u_\ell^\chi = -\chi(\ell)u_{-\ell}^\chi$ and $\langle u_\ell^\chi, u_{\ell'}^\chi\rangle = 0$ if $\ell \neq \pm\ell' \mod p$. So the character χ of $\mathbb{Z}_p$ occurs with multiplicity $(p-1)/2$ in $\langle u_\ell^\chi \,:\, \ell = 1, \ldots, (p-1)/2\rangle$. The character $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi$ of Γ_0 occurs with multiplicity at least $(p-1)/2$. Indeed, the subspaces $R_\ell := \oplus_{s\in Q} s.u_\ell^\chi$ for $\ell = 1, \ldots, (p-1)/2$, realize this representation, and the action of Γ_0 in the basis $\{s.u_\ell^\chi, s \in Q\}$ is as described in the subsection 6.3.2. Moreover, the vectors $\{s.u_\ell^\chi, s \in Q, \ell = 1, \ldots (p-1)/2\}$ are pairwise orthogonal: $\langle s.u_\ell^\chi, s.u_{\ell'}^\chi\rangle = \langle u_\ell^\chi, u_{\ell'}^\chi\rangle = 0$ if $\ell \neq \ell'$ and $\langle s.u_\ell^\chi, t.u_{\ell'}^\chi\rangle = 0$ if $s \neq t$ because $\mathbb{Z}_p$ acts with different characters on this vectors. It remains to express $s.u_\ell^\chi$ in the standard basis. We have

$$\begin{aligned}
s.u_\ell^\chi &= \frac{1}{p}\sum_{t\in\mathbb{Z}_p} \chi(-t)(e_{st} \wedge e_{st+s\ell}) \\
&= \frac{1}{p}\sum_{t\in\mathbb{Z}_p} \chi(-ts^{-1})(e_t \wedge e_{t+s\ell}) \\
&= \frac{1}{p}\Big(\sum_{\substack{i<j \\ j-i=s\ell}} \chi(-is^{-1})(e_i \wedge e_j) - \sum_{\substack{i<j \\ j-i=-s\ell}} \chi(-js^{-1})(e_i \wedge e_j)\Big)
\end{aligned}$$

To summarize, we have found in $\bigwedge^2 W$: the $(p-1)/4$ characters ψ of Q such that $\psi(-1) = -1$, each of dimension 1 and with multiplicity 2 and the characters $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi$ and $\mathrm{Ind}_{\mathbb{Z}_p}^{\Gamma_0}\chi^\alpha$, each of dimension $(p-1)/2$ and with multiplicity $(p-1)/2$. Indeed,

$$\dim(\bigwedge^2 W) = \binom{p}{2} = 2\cdot(p-1)/4 + 2\cdot((p-1)/2)\cdot((p-1)/2)$$

so the lower bounds found for the multiplicities are equalities. □

6.2.4 Expression of the Γ_0-invariant matrices

Let $X \in \mathbb{R}^{\binom{p}{2}\times\binom{p}{2}}$, then by theorem 2.4.1 we know that X is Γ_0-invariant and positive semidefinite if and only if

$$X(ij, k\ell) = \sum_{\psi\,:\,\psi(-1)=-1} \langle A_\psi, \overline{E_\psi(ij,k\ell)}\rangle + \langle A_\chi, \overline{E_\chi(ij,k\ell)}\rangle + \langle A_{\chi^\alpha}, \overline{E_{\chi^\alpha}(ij,k\ell)}\rangle$$

where A_ψ, A_χ, A_{χ^α} are hermitian positive semidefinite matrices of size respectively 2, $(p-1)/2$, $(p-1)/2$.

For $1 \le a, b \le 2$,

$$\begin{aligned}\left(E_\psi\right)_{a,b}(ij,k\ell) &= \langle v_a^\psi/\|v_a^\psi\|, e_i \wedge e_j\rangle \overline{\langle v_b^\psi/\|v_b^\psi\|, e_k \wedge e_\ell\rangle} \\ &= \begin{cases} \frac{4\psi((j-i)^{-1}(l-k)\alpha^{a-b})}{p(p-1)} & \text{if } \begin{cases} \left(\frac{j-i}{p}\right) = (-1)^{a-1} \\ \left(\frac{l-k}{p}\right) = (-1)^{b-1} \end{cases} \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

For $1 \le a, b \le (p-1)/2$,

$$\left(E_\chi\right)_{a,b}(ij,k\ell) = \sum_{s\in Q} \langle s.u_a^\chi/\|s.u_a^\chi\|, e_i \wedge e_j\rangle \overline{\langle s.u_b^\chi/\|s.u_b^\chi\|, e_k \wedge e_\ell\rangle}$$

$$= \begin{cases} \frac{1}{p}\Big(\chi((k-i)s^{-1}) + \chi((j-l)s^{-1})\Big) & \text{if } a^{-1}(j-i) = b^{-1}(l-k) =: s \text{ and } s \in Q \\ -\frac{1}{p}\Big(\chi((l-i)s^{-1}) + \chi((j-k)s^{-1})\Big) & \text{if } a^{-1}(j-i) = -b^{-1}(l-k) =: s \text{ and } s \in Q \\ 0 & \text{otherwise} \end{cases}$$

The same formula holds for E_{χ^α} replacing χ with χ^α.

6.2.5 The final reduction

We are now ready to symmetrize the program 5.3.1 for Paley graphs for the case $k = 2$. Recall the original formulation:

$$\vartheta_2(P(p)) = \max\Big\{\langle L_1', X\rangle \ : \ \begin{array}{l} X \in \mathbb{R}^{\binom{p}{2}\times\binom{p}{2}} \succeq 0 \\ \langle I, X\rangle = 1 \\ X_{s,s'} \neq 0 \text{ only if } s = s' \in \overline{E} \text{ or } s \cup s' \text{ is} \\ \qquad\qquad \text{a triangle without edges} \\ X_{ab,ac} = -X_{ab,bc} = X_{ac,bc} \ \forall\ a < b < c \end{array}\Big\}$$

This program is clearly invariant for the Γ_0 action. Then, using the definition of L_1' and lemma 6.3.1, the objective function writes as

$$\begin{aligned}\langle L_1', X\rangle &= 2\sum_{a<b} X_{ab,ab} + 6\sum_{a<b<c} X_{ab,ac} \\ &= 2\mathrm{Tr}(X) + 6\sum_{a<b<c\ :\ abc\in\mathcal{T}_0} X_{ab,ac} \\ &= 2 + 6\sum_{\beta\in\Omega} C_\beta X_{0\alpha,0\beta}\end{aligned}$$

where Ω is such that $\{(0,\alpha,\beta) : \beta \in \Omega\}$ is the collection of the canonical representatives of the orbits of Γ_0 on $\mathcal{T}_0$, and C_β is the cardinality of the corresponding orbit.

The trace of X is given by

$$\mathrm{Tr}(X) = \sum_{a<b\ :\ ab\in\overline{E}} X_{ab,ab} = \frac{p(p-1)}{4} X_{0\alpha,0\alpha}$$

Then, with the zonal matrices for the action of Γ_0 derived before, we can announce the final reduction of the program:

Theorem 6.2.3. *With the previous notation, for every p prime,*

$$\alpha(P(p)) \leq 2 + 6\max\left\{ \begin{array}{ll} \sum_{\beta\in\Omega} C_\beta X_{0\alpha,0\beta} & : \quad A_\psi, A_\chi, A_{\chi^\alpha} \succeq 0 \\ & \quad \frac{p(p-1)}{4} X_{0\alpha,0\alpha} = 1 \end{array} \right\}$$

where $A_\psi, A_\chi, A_{\chi^\alpha}$ are obtained as the scalar product of X with $E_\psi, E_\chi, E_{\chi^\alpha}$ respectively:

$$\left(A_\rho\right)_{a,b} = \frac{1}{\dim(\rho)} \sum_{i<j,\ k<\ell} X(ij,kl)\overline{\left(E_\rho\right)_{a,b}(ij,k\ell)}.$$

In the following table we compare the bound of theorem 6.3.3 with the bound $L^2(P(p))$ in [23] (see section 2.2.3) for several values of p. The reason is that both bounds involve triples of words.

p	$\sqrt{p}$	$\vartheta_2(P(p))$	$L^2(P(p))$	$\alpha(P(p))$
61	7.810	5.874	5.465	5
73	8.544	6.236	5.973	5
89	9.433	6.304	6.304	5
97	9.848	7.517	7.398	6
101	10.049	6.898	6.611	5
109	10.440	7.366	7.366	6
113	10.630	7.676	7.599	7
137	11.704	8.623	8.200	7
149	12.206	8.750	8.231	7
157	12.529	8.851	8.707	7
173	13.152	10.020	9.426	8
181	13.453	9.466	9.112	7
193	13.892	9.301	9.210	7
197	14.035	9.902	9.226	8
229	15.132	10.707	10.290	9
233	15.264	10.537	10.182	7
241	15.524	10.073	9.891	7
257	16.031	10.469	10.247	7
269	16.401	11.450	10.624	8
277	16.643	11.282	10.340	8
281	16.763	11.034	10.605	7
293	17.117	11.706	10.937	8
313	17.691	12.182	11.551	8
317	17.804	12.535	12.337	9
337	18.357	12.129	11.658	9
401	20.024	13.329	12.753	9
509	22.561	14.974	14.307	9
601	24.515	16.694	16.077	11
701	26.476	17.581	16.857	10
809	28.442	18.439	17.371	11

Table 6.1: Values of $\vartheta_2(P(p))$

Note: All values of the semidefinite programs have been obtained with the solvers SDPA or SDPT3, available on the NEOS website (http://www.neos-server.org/neos/).

6.3 Circulant subgraphs of Paley graphs

Definition 6.3.1. *Given $n \geq 2$ and $J \subset \{1, \ldots, n-1\}$ such that $J = -J$, the circulant graph $C_n(J)$ is the graph with vertices $\mathbb{Z}/n\mathbb{Z}$ and edges $\{ij : i - j \in J\}$.*

Equivalently, a circulant graph is a graph whose adjacency matrix is circulant, i.e. all of its rows are obtained by a cyclic shift of the first one. In particular, a circulant graph is a Cayley graph on an additive group with generating set J. Clearly, $C_n(J)$ is regular of degree $|J|$.

Let $\overline{J} = \{1, \ldots, n-1\} \setminus J$; then $C_n(\overline{J})$ is the complementary graph of $C_n(J)$. We introduce $J_0 = J \cap \{1, \ldots, \lfloor n/2 \rfloor\}$ and note that $J = J_0 \cup -J_0$. We have $|J_0| = |J|/2$ if $n/2 \notin J$ and $|J_0| = (|J| + 1)/2$ otherwise. Obviously $C_n(J) \simeq C_n(aJ)$ for $a \in \mathbb{Z}_n^\times$.

Because translations by $\mathbb{Z}_n$ act transitively on vertices of $C_n(J)$, we have

$$\vartheta(C_n(J))\vartheta(C_n(\overline{J})) = n. \tag{6.3.1}$$

Let Q be the subgroup of squares of $\mathbb{Z}_p^\times$ and $\overline{Q} = \mathbb{Z}_p^\times \setminus Q$, so that we have the disjoint union $\mathbb{Z}_p = Q \cup \overline{Q} \cup \{0\}$.

Proposition 6.3.2. *The induced subgraphs Q and $\overline{Q}$ of the Paley graph $P(p)$ are circulant graphs of order $n = (p-1)/2$.*

Proof. If α is a primitive element of $\mathbb{Z}_p$ and $\beta = \alpha^2$, then Q is the graph with vertices $\{\beta^i, 0 \leq i \leq (p-3)/2\}$ and edges between β^i and β^j if and only if $\beta^i - \beta^j \in Q$, or equivalently $\beta^{i-j} - 1 \in Q$. So Q is isomorphic to the circulant graph of order $(p-1)/2$ with associated set

$$J(Q) = \{j \in [(p-3)/2] : \beta^j \in 1 + Q\}$$

and similarly $\overline{Q} \simeq C_{(p-1)/2}(J(\overline{Q}))$ for

$$J(\overline{Q}) = \{j \in [(p-3)/2] : \beta^j \in 1 + \overline{Q}\}.$$

We note that $|J(Q)| = (p-5)/4$ and $|J(\overline{Q})| = (p-1)/4$, due to strong regularity of Paley graphs. □

On the other hand we have

Theorem 6.3.3.

$$\alpha(P(p)) = \alpha(\overline{Q}) + 1$$

Proof. An independent set for $P(p)$ can be assumed to contain 0 after translation. Then the non zero elements are contained in $\overline{Q}$. Conversely, the union of an independent set of $\overline{Q}$ and $\{0\}$ makes an independent set of $P(p)$. □

So, $\alpha(P(p)) \leq \vartheta(\overline{Q}) + 1$ and, as $\overline{Q}$ is a subgraph of $P(p)$, $\vartheta(\overline{Q}) \leq \vartheta(P(p))$.

6.4 Formulations for the ϑ number of a circulant graph

Remember that $\overline{Q}$ is a circulant graph of even order $n = (p-1)/2$. So for the rest of the section we assume that n is even and we look at $\vartheta(C_n(J))$ more in details. Due to transitivity of $\mathbb{Z}_n$, the ϑ number is a linear program.

Theorem 6.4.1.

$$\vartheta(C_n(J)) = \max\{\ nf_0 \ : \ \begin{array}{l} f \in \mathbb{R}^{1+n/2} \\ f_0, f_1, \ldots, f_{n/2} \geq 0 \\ \sum_{\ell=0}^{n/2} f_\ell \cos\left(\frac{2\pi\ell j}{n}\right) = \delta_{j,0} \quad j \in \{0\} \cup J_0\ \} \end{array} \tag{6.4.1}$$

$$\vartheta(C_n(J)) = \min\{\ ng_0 \ : \ \begin{array}{l} g \in \mathbb{R}^{1+|J_0|} \\ \sum_{j \in \{0\} \cup J_0} g_j \cos\left(\frac{2\pi\ell j}{n}\right) \geq \delta_{\ell,0} \quad \ell \in [0, \ldots, n/2]\ \} \end{array} \tag{6.4.2}$$

Proof. By definition, we have

$$\vartheta(C_n(J)) = \max\{\ \sum F(\gamma,\delta) \ : \ \begin{array}{l} F \in \mathbb{R}^{n \times n}, F \succeq 0 \\ \sum F(\gamma,\gamma) = 1 \\ F(\gamma,\delta) = 0 \text{ if } \gamma - \delta \in J\ \} \end{array}$$

Such program is invariant for the additive action of $\mathbb{Z}/n\mathbb{Z}$ on itself, thus we can restrict to consider functions F such that $F(\gamma + x, \delta + x) = F(\gamma, \delta)$ for any $x \in \mathbb{Z}/n\mathbb{Z}$. Such functions can be written as $F(\gamma,\delta) = \tilde{F}(\gamma - \delta)$ for some $\tilde{F} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{R}$ such that $\tilde{F}(-j) = \tilde{F}(j)$.

Recall that the characters of $\mathbb{Z}/n\mathbb{Z}$ are given by:

$$\begin{aligned} \chi_\ell \ : \ \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{C}^\times \\ j &\longrightarrow e^{2i\pi j\ell/n} \end{aligned}$$

for $\ell = 0, \ldots, n-1$. Then $\tilde{F}$ expands as

$$\tilde{F}(j) = \sum_{\ell=0}^{n-1} f_\ell \chi_\ell(j) = \sum_{\ell=0}^{n-1} f_\ell e^{2i\pi j\ell/n}$$

with the following properties:

- $F \succeq 0 \Leftrightarrow f_\ell \geq 0\ \forall\ \ell$

- $\tilde{F}$ real-valued implies that $f_\ell = f_{n-\ell}\ \forall\ \ell$
- it follows that $\tilde{F}(j) = f_0 + \sum_{\ell=1}^{n/2} 2f_\ell \cos(2\pi j\ell/n)$
- we have

$$\sum_{\gamma,\delta\in\mathbb{Z}/n\mathbb{Z}} F(\gamma,\delta) = n \sum_{j\in\mathbb{Z}/n\mathbb{Z}} \tilde{F}(j) = n\Big(nf_0 + \sum_\ell 2f_\ell \sum_{j=0}^{n-1} \cos(2\pi j\ell/n)\Big) = n^2 f_0$$

$$\sum_{\gamma\in\mathbb{Z}/n\mathbb{Z}} F(\gamma,\gamma) = n\tilde{F}(1) = n(f_0 + 2f_1 + \cdots + 2f_{n/2})$$

$$\tilde{F}(j) = 0 \text{ for } j \in J_0$$

So, renaming $f_0 = nf_0$ and $f_\ell = 2nf_\ell$ for $\ell = 1, \dots, n-1$, we find 6.5.1. By duality, 6.5.2 follows. □

We note $J_0^c = \{1, \dots, n/2\} \setminus J_0$. For the Paley graph, $J_0^c(\overline{Q}) = J_0(Q)$. For $p \equiv 1 \bmod 8$ they have the same number of elements. For $p \equiv 5 \bmod 8$ they differ by 1. Then the idea is that the two optimal values $\vartheta(\overline{Q})$ and $\vartheta(Q)$ are rather close because the sets $J_0(\overline{Q})$ and $J_0(Q)$ are random enough, with the consequence that the cosinus are "interlaced" enough.

We have computed some values of the linear program $\vartheta(\overline{Q})$ up to $p < 20000$. The optimal value is around $\sqrt{(p-1)/2}$, but it doesn't seem to converge.

Remark 6.4.2. The optimal value of (6.5.1) for $\overline{Q}$ is obtained for f^* such that some of the coordinates $f_\ell^* = 0$. Let us denote $\{0\} \cup L$ the support of the optimal solution f^*. Then $|L| = |J_0|$ so that f^* is uniquely obtained by solving a linear system (this is because the vertices of the polytope on which we are optimizing are all obtained this way: by setting an appropriate number of coordinates to zero). Then the optimal value for Q has support $\{0\} \cup \overline{L}$. However we couldn't so far find a relationship between J_0 and L.

Here the discussion remains open. From the relation $\vartheta(Q)\vartheta(\overline{Q}) = (p-1)/2$ and the considerations above, one can hope that, roughly speaking,

$$\vartheta(Q) \approx \vartheta(\overline{Q}) \approx \sqrt{(p-1)/2}$$

This would lead to a minor improvement on the estimate of $\alpha(P(p))$ with respect to the known bound $\alpha(P(p)) \leq \sqrt{p}$. Nevertheless it seems that it is really hard to improve this bound, especially with a multiplicative factor. In [7] the authors prove $\alpha(P(p)) \leq \sqrt{p} - 1$ under some conditions on p.

Bibliography

[1] http://www.research.ibm.com/people/s/shearer/indpal.html.

[2] http://ginger.indstate.edu/ge/PALEY/index.html.

[3] R. Ahlswede, N. Cai, S.-Y.R. Li, and R.W. Yeung. Network information flow. *Information Theory, IEEE Transactions On*, 46(4):1204–1216, July 2000.

[4] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, pages 13–51, 1995.

[5] C. Bachoc. Applications of semidefinite programming to coding theory. Information Theory Workshop (ITW), 2010.

[6] C. Bachoc, D. Gijswijt, A. Schrijver, and F. Vallentin. Invariant semidefinite programs. In M. Anjos and J.-B. Lasserre, editors, *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *International series in operations research and management science*. Springer 2012.

[7] C. Bachoc, M. Matolcsi, and I. Z. Ruzsa. Squares and difference sets in finite fields. arXiv:1305.0577.

[8] C. Bachoc, A. Passuello, and F. Vallentin. Bounds for projective codes from semidefinite programming. *Adv. Math. Commun.*, 7:127–145, 2013.

[9] C. Bachoc and F. Vallentin. More semidefinite programming bounds. Proceeding of DMHF, 2007.

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

[11] S. Bochner. Hilbert distances and positive definite functions. *Annals Of Mathematics*, 41:647–656, 1941.

[12] N. G. de Bruijn and P. Erdos. A colour problem for infinite graphs and a problem in the theory of relations. *Nederl. Akad. Wetensch. Proc. Ser. A*, 54:371–373, 1951.

[13] F.M. de Oliveira Filho and F. Vallentin. Fourier analysis, linear programming and densities of distance avoiding sets in $\mathbb{R}^n$. *J. Eur. Math. Soc.*, 12:1417–1428, 2010.

[14] P. Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.*, 1973.

[15] P. Delsarte. Hahn polynomials, discrete harmonics and t-designs. *SIAM J. Appl. Math.*, 34, 1978.

[16] C. F. Dunkl. An addition theorem for some q-Hahn polynomials. *Monat. Math*, 85:5–37, 1977.

[17] T. Etzion and A. Vardy. Error-correcting codes in projective space. *IEEE Trans. Inform. Theory*, 57(2):1165–1173, February 2011.

[18] K. J. Falconer. The realization of distances in measurable subsets covering $\mathbb{R}^n$. *J. Combin. Theory Ser. A*, page 187–189, 1981.

[19] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1:357–368, 1981.

[20] P. Frankl and R. M. Wilson. The Erdos-Ko-Rado theorem for vector spaces. *J. Combin. Theory, Series A*, 43:228–236, 1986.

[21] D. Gijswijt, H. Mittelmann, and A. Schrijver. Semidefinite code bounds based on quadruple distances. *IEEE Transactions on Information Theory*, 58:2697–2705, 2012.

[22] D. Gijswijt, A. Schrijver, and H. Tanaka. New upper bounds for nonbinary codes. *J. Combin. Th. Ser. A*, 13:1717–1731, 2006.

[23] N. Gvozdenovic, M. Laurent, and F. Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. *Operations Research Letters*, 37:27–31, 2009.

[24] T. Ho, R. Koetter, M. Médard, D. R. Karger, and M. Effros. The benefits of coding over routing in a randomized setting. Proc. IEEE Int. Symp. Information Theory, Yokohama, Japan, June 2003.

[25] J. Jonsson. *Simplicial complexes of graphs*. Lecture notes in Mathematics. Springer, 2008.

[26] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, page 373–395, 1984.

[27] R. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. R.E. Miller and J. W. Thatcher, eds., Plenum, New York, 1972.

[28] A. Khaleghi and F. R. Kschischang. Projective space codes for the injection metric. pages 9–12. Proceedings of the Canadian workshop on information theory, Ottawa, Canada, May 2009.

[29] R. Koetter and F. R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Transactions on Information Theory*, 54(8):3579 – 3591, August 2008.

[30] F. R. Kschischang and D. Silva. On metrics for error correction in network coding. *IEEE Transactions on Information Theory*, 55(12):5479–5490, 2009.

[31] A. Kupavskii and A. Raigorodskii. On the chromatic numbers of small-dimensional Euclidean spaces. volume 34, pages 435–439. European Conference on Combinatorics, Graph Theory and Applications EuroComb, 2009.

[32] J.B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear $0-1$ programs. *SIAM J. Optim.*, 12:756–769, 2002.

[33] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for $0-1$ programming. *Math. Oper. Res.*, 28:470–496, 2003.

[34] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25:1–5, 1979.

[35] L. Lovász and A. Schrijver. Cones of matrices and set-functions and $0-1$ optimization. *SIAM J. Optim.*, 1:166–190, 1991.

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

[37] Yu. E. Nesterov and A. S. Nemirovski. Conic formulation of a convex programming problem and duality. *Optim. Methods Softw.*, pages 95–115, 1992.

[38] Yu. E. Nesterov and A. S. Nemirovski. *Interior point polynomial algorithms in convex programming.* SIAM Publications. SIAM, 1994.

[39] A. M. Raigorodskii. On the chromatic number of a space. *Russian Math. Surveys*, 55:351–352, 2000.

[40] A. Schrijver. A comparison of the Delsarte and Lovász bound. *IEEE Trans. Inform. Theory*, 25:425–429, 1979.

[41] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Trans. Inf. Theory*, 51:2859–2866, 2005.

[42] J.-P. Serre. *Linear representations of finite groups*, volume 42 of *Graduate texts in mathematics*. Springer 2001.

[43] C. E. Shannon. The zero-error capacity of a noisy channel. *IRE Transactions on Information Theory*, 2(3):8–19, 1956.

[44] M.J. Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001.

[45] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra and Appl.*, 430:360–369, 2009.

[46] L. Vandenberghe and S. Boyd. *Convex optimization*. Cambridge University Press 2004.

[47] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38:49–95, 1996.

[48] H. Wang, C. Xing, and R. Safavi-Naini. Linear authentication codes: bounds and constructions. *IEEE Trans. Inform. Theory*, 49, 866-872.

[49] S.T. Xia and F.W. Fu. Johnson type bounds on constant dimension codes. *Designs, Codes, Crypto.*, 50:163–172, February 2009.

Programmation semidéfinie positive dans l'optimisation combinatoire avec applications à la théorie des codes correcteurs et à la géométrie.

Résumé : Une nouvelle borne supérieure sur le cardinal des codes de sous-espaces d'un espace vectoriel fini est établie grâce à la méthode de la programmation semidéfinie positive. Ces codes sont d'intérêt dans le cadre du codage de réseau (network coding). Ensuite, par la même méthode, l'on démontre une borne sur le cardinal des ensembles qui évitent une distance donnée dans l'espace de Johnson et qui est obtenue par une variante d'un programme de Schrijver. Les résultats numeriques permettent d'améliorer les bornes existantes sur le nombre chromatique mesurable de l'espace Euclidien. Une hiérarchie de programmes semidéfinis positifs est construite à partir de certaines matrices issues des complexes simpliciaux. Ces programmes permettent d'obtenir une borne supérieure sur le nombre d'indépendance d'un graphe. Aussi, cette hiérarchie partage certaines propriétés importantes avec d'autres hiérarchies classiques. A titre d'exemple, le problème de déterminer le nombre d'indépendance des graphes de Paley est analysé.

Mots clés : théorie des graphes, nombre d'indépendance, nombre chromatique, SDP, codes projectifs, hiérarchies.

Semidefinite programming in combinatorial optimization with applications to coding theory and geometry.

Abstract: We apply the semidefinite programming method to obtain a new upper bound on the cardinality of codes made of subspaces of a linear vector space over a finite field. Such codes are of interest in network coding. Next, with the same method, we prove an upper bound on the cardinality of sets avoiding one distance in the Johnson space, which is essentially Schrijver semidefinite program. This bound is used to improve existing results on the measurable chromatic number of the Euclidean space. We build a new hierarchy of semidefinite programs whose optimal values give upper bounds on the independence number of a graph. This hierarchy is based on matrices arising from simplicial complexes. We show some properties that our hierarchy shares with other classical ones. As an example, we show its application to the problem of determining the independence number of Paley graphs.

Keywords: graph theory, stable number, chromatic number, SDP, projective codes, hierarchies.

Discipline : Mathématiques pures

Institut de Mathématiques de Bordeaux
Université Bordeaux 1,
351, cours de la libération,
33405 Talence