Polynomial Optimization and Discrete Geometry

Advisor : Christine Bachoc
Rujuta .S. Joshi

01/07/2015

Acknowledgements

I would like to thank Professor Christine Bachoc, my Advisor for providing guidence and all the support. I am also grateful to her for encouraging me to pursue this thesis topic .I would also like to thank my family for their constant support.

Contents

Chapter 1

Introduction

Problem Statement : How many non overlapping regular tetrahedra, having a common vertex can be arranged in $\mathbb{R}^3$?

Solution to this problem (say T(3)) is known to satisfy $20 \leq T(3) \leq 22$). Thesis aims at trying to solve the above problem with the help of polynomial optimization.

Thesis is divided into 4 parts. The first part explains basic concepts required further. Second part deals with the theory of polynomial optimization. In part 3 we study spherical harmonics and apply polynomial optimization to the kissing number problem.Eventually in part 4 we try to bound T(3) with tools developed so far.

Chapter 2

Preliminaries

Some concepts have been introduced ,which will be used later in the thesis .

Definition 2.0.0.1. A polynomial optimization problem is to optimize value of f $\in$ $\mathbb{R}[x_1, ..., x_n]$ over a set K described by some $g_1, ..., g_m \in$ $\mathbb{R}[x_1, , , , x_n]$ .Let us consider computing infimum of a polynomial f $\in$ $\mathbb{R}[x_1, ..., x_n]$ over $K = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, ..., g_m(x) \geq 0\}$.

$$f_{min} := inf\{f(x) \mid g_1(x) \geq 0, ..., g_m(x) \geq 0\} \tag{2.1}$$

Notations : Let N = $\mathbb{N} \cup \{0\}, t \geq 0$; $N_t^n := \{(\alpha_1, .., \alpha_n) \in N^n \mid \sum_{i=1}^n \alpha_i \leq t\}$; $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, ..., x_n]$. Let $\alpha \in N^n$ (say $\alpha = (\alpha_1, .., \alpha_n)$) ; $|\alpha| := \sum \alpha_i$. Then $\mathbf{x}^\alpha = x_1^{\alpha_1}...x_n^{\alpha_n}$; $\mathbb{R}[\mathbf{x}]_t := \{f \in \mathbb{R}[\mathbf{x}] \mid degf \leq t\}$; $[\mathbf{x}]_t := (x^\alpha)_{|\alpha| \leq t}$ (In some fixed order) Now if f $\in \mathbb{R}[\mathbf{x}]_t$. Then deg f $\leq t$. Therefore coefficient of f can be expressed in a vector form $[f] = [f_\alpha]_{\alpha \in N_t^n}$: f_α coefficient of $\mathbf{x}^\alpha$ (in same order as above) So $f = [f]^t[x]_t$. Let I be an ideal in $\mathbb{R}[\mathbf{x}]$,then $I_t = \{f \in I \mid degf \leq t\}$

2.0.1 Measure theory

Let X be a set. If 2^X is collection of subsets of X.

Definition 2.0.1.1. σ - algebra Σ : Let $\Sigma \subseteq 2^X$ then Σ is a $\sigma - algebra$ if

1. $X \in \Sigma$
2. if $A \in \Sigma$, then $X\setminus$ A in Σ
3. Σ is closed under countable union. i.e If $A_1, A_2, ... \in \Sigma$ then $\cup_{i \in \mathbb{N}} A_i \in \Sigma$

Definition 2.0.1.2. Measure μ : Measure μ on a set X (with $\sigma - algebra$ Σ) is a function from Σ to $\mathbb{R} \cup \{\infty\}$ satisfying

1. $\mu(A) \geq 0 \; \forall \; A \in \Sigma$
2. $\mu(\phi) = 0$
3. Let $A_1, A_2, ...$ be countable pairwise disjoint subsets of X in Σ, then $\mu(\cup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mu(A_i)$

Definition 2.0.1.3. Borel measure on a set X : X is locally compact and Hausdorff space.Let Σ be the smallest $\sigma - algebra$ containing all open sets in X.And a measure μ defined on this $\sigma - algebra$ is a borel measure.

Definition 2.0.1.4. Dirac measure :(let X be set with given $\sigma - algebra\ \Sigma$).Then dirac measure w.r.t. a fixed point $x \in X$ is

$$\begin{aligned} \mu(A) &= 0\ if\ x \notin A \\ &= 1\ if\ x \in A \end{aligned}$$

where A $\in$ Σ.

Definition 2.0.1.5. Probability measure : μ measure on X with $\sigma - algebra\ \Sigma$ is a probability measure if μ takes values in $[0, 1]$ and $\mu(X) = 1$.

2.0.2 Algebraic Geometry

Let I be an ideal in $\mathbb{R}[\mathbf{x}]$.

Definition 2.0.2.1. Radical of I : $\sqrt{I}$:=$\{f \in \mathbb{R}[\mathbf{x}] \mid f^m \in I$ for some $m \geq 1\}$. I is said to a radical ideal if $I = \sqrt{I}$.

Definition 2.0.2.2. Real radical of I : $\sqrt[\mathbb{R}]{I} := \{f \in \mathbb{R}[\mathbf{x}] \mid f^{2m} + p_1^2 + ... + p_k^2 \in I$ for some $m \geq 1$ and $p_1, .., p_k \in \mathbb{R}[\mathbf{x}]\}$. I is said to be a real radical ideal if $I = \sqrt[\mathbb{R}]{I}$.

Definition 2.0.2.3. $V_{\mathbb{C}}(I)$ **:** $\{(a_1, ..., a_n) \in \mathbb{C}^n \mid f(a_1, ..., a_n) = 0\ \forall\ f \in I\}$. It is called a complex variety.

Definition 2.0.2.4. $V_{\mathbb{R}}(I)$ **:** $\{(a_1, ..., a_n) \in \mathbb{R}^n \mid f(a_1, ..., a_n) = 0\ \forall\ f \in I\} = V_{\mathbb{C}}(I) \cap \mathbb{R}^n$ .It is a called a real variety.

Definition 2.0.2.5. $I(V_{\mathbb{C}}(I))$:= $\{f \in \mathbb{R}[\mathbf{x}] \mid f(a_1, ..., a_n) = 0\ \forall\ (a_1, ..., a_n) \in V_{\mathbb{C}}(I)\}$.

Definition 2.0.2.6. $I(V_{\mathbb{R}}(I))$:= $\{f \in \mathbb{R}[\mathbf{x}] \mid f(a_1, ..., a_n) = 0\ \forall\ (a_1, ..., a_n) \in V_{\mathbb{R}}(I)\}$.

Lemma 2.0.2.7. $I \subseteq \sqrt{I} \subseteq I(V_{\mathbb{C}}(I))$

Proof. $f \in I \Rightarrow f^1 \in I$,Therefore $f \in I \Rightarrow f \in \sqrt{I}$. $f \in \sqrt{I} \Rightarrow f^m \in I$ for some $m \geq 1$. Therefore $f^m(a_1, ..., a_n) = 0\ \forall\ (a_1, ..., a_n) \in V_{\mathbb{C}}(I) \Rightarrow f(a_1, ..., a_n) = 0\ \forall\ (a_1, ..., a_n) \in V_{\mathbb{C}}(I) \Rightarrow f \in I(V_{\mathbb{C}}(I))$ ☐

Lemma 2.0.2.8. $I \subseteq \sqrt[\mathbb{R}]{I} \subseteq I(V_{\mathbb{R}}(I))$

Proof. $f \in I \Rightarrow f^2 \in I$, Therefore $f \in I \Rightarrow f \in \sqrt[\mathbb{R}]{I}$. $f \in \sqrt[\mathbb{R}]{I} \Rightarrow f^{2m} + p_1^2 + ... + p_k^2 \in I$. Let $(a_1, ..., a_n) \in V_{\mathbb{R}}(I)$ then $(f^{2m}+p_1^2+...+p_k^2)(a_1, ..., a_n) = 0$ but $(a_1, ..., a_n) \in \mathbb{R}^n \Rightarrow p_i(a_1, ..., a_n) \in \mathbb{R}\ \forall\ i$ and $f^m(a_1, ..., a_n) \in \mathbb{R} \Rightarrow f^{2m}(a_1, ..., a_n) = 0 \Rightarrow f(a_1, ..., a_n) = 0 \Rightarrow f \in I(V_{\mathbb{R}}(I))$

☐

Theorem 2.0.2.9. ***Hilbert's Nulltellensatz and Real Nullstellensatz thm:*** *I ideal in* $\mathbb{R}[\boldsymbol{x}]$. *Then* $\sqrt{I} = I(V_{\mathbb{C}}(I))$ *and* $\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I))$.*(For Hilbert's Nulltellensatz refer to Serre Lang's and for real Nulltellensatz refer [3])*

Lemma 2.0.2.10. $I \subseteq I(V_{\mathbb{C}}(I)) \subseteq I(V_{\mathbb{R}}(I))$.

Proof. $V_{\mathbb{R}}(I)$: $V_{\mathbb{C}}(I) \bigcap \mathbb{R}^n$ Therefore f $\in I(V_{\mathbb{C}}(I)) \Rightarrow f(a_1, ..., a_n) = 0\ \forall\ (a_1, ..., a_n) \in V_{\mathbb{C}}(I)$ which implies $f(a_1, ..., a_n) = 0 \forall (a_1, ..., a_n) \in V_{\mathbb{R}}(I) \Rightarrow f \in I(V_{\mathbb{R}}(I))$. Therefore, $I \subseteq I(V_{\mathbb{C}}(I)) \subseteq I(V_{\mathbb{R}}(I))$. ☐

Theorem 2.0.2.11. *If I is a real radical ideal and* $\mid V_{\mathbb{R}}(I) \mid < \infty$*, then* $V_{\mathbb{C}}(I) = V_{\mathbb{R}}(I)$.

Proof. If I is a real radical ideal then $I \subseteq I(V_{\mathbb{C}}(I)) \subseteq \sqrt[\mathbb{R}]{I} = I$. It implies $I(V_{\mathbb{C}}(I)) = \sqrt{I} = I = I(V_{\mathbb{R}}(I)) \implies$ I is radical and $I(V_{\mathbb{C}}(I)) = I(V_{\mathbb{R}}(I))$. Now if I is a real radical and $|V_{\mathbb{R}}(I)| < \infty$,then $V_{\mathbb{R}}(I) = V_{\mathbb{C}}(J)$ for some ideal J.($\because$ if $|V_{\mathbb{R}}(I)| = 1$ say $a = (a_1, ..., a_n) = V_{\mathbb{R}}(I)$,then $x_i - a_i \in \mathbb{R}[x_1, ..., x_n] \; \forall \; 1 \leq i \leq n$ and $V_{\mathbb{C}}((x_1 - a_1, ..., x_n - a_n)) = V_{\mathbb{R}}(I)$. Now if $|V_{\mathbb{R}}(I)| = m < \infty$ then we get ideals $J_1, ..., J_m$ such that each point in $V_{\mathbb{R}}(I) = V_{\mathbb{C}}(J_i)$ for some $1 \leq j \leq m$ and therefore $V_{\mathbb{R}}(I) = V_{\mathbb{C}}(J_1 ... J_m))$ and so $I(V_{\mathbb{C}}(I)) = I(V_{\mathbb{R}}(I)) = I(V_{\mathbb{C}}(J))$. $V_{\mathbb{C}}(I(V_{\mathbb{C}}(J))) = V_{\mathbb{C}}(J)$ ($\because V_{\mathbb{C}}(J) \subseteq V_{\mathbb{C}}(I(V_{\mathbb{C}}(J)))$ and if a $\in V_{\mathbb{C}}(I(V_{\mathbb{C}}(J)))$ then $\forall \; f \in I(V_{\mathbb{C}}(J)); f(a) = 0$. But by Hilbert Nullstellensatz $I(V_{\mathbb{C}}(J)) = \sqrt{J}.J \subseteq \sqrt{J}$,Therefore $\forall \; f \in J; f(a) = 0$. So $a \in V_{\mathbb{C}}(J))$. Similarly $V_{\mathbb{C}}(I(V_{\mathbb{C}}(I))) = V_{\mathbb{C}}(I)$. So $V_{\mathbb{C}}(I) = V_{\mathbb{C}}(J) = V_{\mathbb{R}}(I)$.

□

Proposition 2.0.2.12. *Let I be an ideal in* $\mathbb{R}[x_1, ..., x_n]$*.Then* $\mid V_{\mathbb{C}}(I) \mid < \infty$ *iff* $\mathbb{R}[x_1, ..., x_n]/I$ *is finite dimensional as a vector space.(For proof refer to [1])*

Interpolation Polynomials :

Theorem 2.0.2.13. *Let* $V \subseteq \mathbb{R}^n$ *be finite set.Then there exist polynomials* $p_v \in \mathbb{R}[x_1, ..., x_n] \; \forall \; v \in V$ *satisfying* $p_v(u) = \delta_{u,v} \; \forall \; u, v \in V$*.Then we have that for any polynomial* $f \in \mathbb{R}[x_1, .., x_n]$

$$f - \sum_{v \in V_{\mathbb{C}}(I)} f(v) p_v \in I(V_{\mathbb{C}}(I)) \tag{2.2}$$

Proof. Fix v.$\forall u \neq v \; \exists$ component i_u such that $v(i_u) \neq u(i_u)$. Define

$$p_v := \prod_{u \in V_{\mathbb{C}}(I) \backslash v} (x(i_u) - u(i_u))/(p(i_u) - u(i_u)) \tag{2.3}$$

According to this definition $p_v(v) = 1$ and $p_v(u) = 0 \; \forall \; u \neq v \in V_{\mathbb{C}}(I)$.Let $f \in \mathbb{R}[x_1, .., x_n]$.For any u in $V_{\mathbb{C}}(I)$ we have $(f - \sum_{v \in V_{\mathbb{C}}(I)} f(v) p_v)(u) = f(u) - \sum_{v \in V_{\mathbb{C}}(I)} f(v) p_v(u) = f(u) - f(u) = 0$.So by definition of $I(V_{\mathbb{C}}(I)), f - \sum_{v \in V_{\mathbb{C}}(I)} f(v) p_v \in I(V_{\mathbb{C}}(I))$. □

2.0.3 Linear Functionals

Definition 2.0.3.1. Linear Functional : Let $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ be a sequence of real numbers.Corresponding linear functional L on $\mathbb{R}[x_1, ..., x_n]$ is given by

$$\begin{gathered} \text{L} : \mathbb{R}[x_1, .., x_n] \longrightarrow \mathbb{R} \\ x^\alpha \longmapsto L(x^\alpha) = y_\alpha \\ f = \textstyle\sum_\alpha f_\alpha x^\alpha \longmapsto L(f) = \sum_\alpha f_\alpha y_\alpha \end{gathered}$$

Definition 2.0.3.2. Moment Matrix : Let $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ be a sequence of real numbers then

$$M(y) := (y_{\alpha+\beta})_{\alpha,\beta \in N^n} ......(It\ is\ an\ infinite\ matrix) \tag{2.4}$$

If g is a polynomial define

$$(g * y)_\alpha = \sum_{\delta \leq deg(g)} g_\delta y_{\alpha+\delta} \tag{2.5}$$

Lemma 2.0.3.3. *Let p be any polynomial in* $\mathbb{R}[x_1, .., x_n]$*.Let* $y = (y_\alpha)_{\alpha \in N^n}$ *be a sequence of real numbers.Let L be the corresponding linear functional.Then* $M(y) \succeq 0$ *iff* $L(p^2) \geq 0 \; \forall \; p \in \mathbb{R}[x_1, .., x_n]$ *and* $M(g * y) \succeq 0$ *iff* $L(gp^2) \geq 0 \; \forall \; p \in \mathbb{R}[x_1, .., x_n]$.

Proof.

$$\begin{aligned} L(p^2) &= \sum_{\alpha} (p^2)_\alpha y_\alpha \\ &= \sum_{\alpha} (\sum_{\beta+\gamma=\alpha} (p)_\beta (p)_\gamma) y_\alpha \\ &= \sum_{\alpha} \sum_{\beta+\gamma=\alpha} (p)_\beta (p)_\gamma y_{\beta+\gamma} \\ &= \sum_{\beta,\gamma} (p)_\beta (p)_\gamma y_{\beta+\gamma} \\ &= pM(y)p^t \end{aligned}$$

So $L(p^2) = pM(y)p^t \ \forall \ p \ \in \ \mathbb{R}[x_1, ..., x_n]$.So $L(p^2) \geq 0 \ \forall \ p \ \in \ \mathbb{R}[x_1, ..x_n]$ iff $M(y) \succeq 0$.

$$\begin{aligned} L(gp^2) &= \sum_{\alpha} (gp^2)_\alpha y_\alpha \\ &= \sum_{\alpha} (\sum_{\delta} g_\delta (p^2)_{\alpha-\delta}) y_\alpha \\ &= \sum_{\alpha} \sum_{\delta} g_\delta \sum_{\beta+\gamma=\alpha-\delta} (p)_\beta (p)_\gamma y_\alpha \\ &= \sum_{\delta,\beta,\gamma} g_\delta (p)_\beta (p)_\gamma y_{\delta+\beta+\gamma} \\ &= pM(g * y)p^t \end{aligned}$$

So $L(gp^2) = pM(g * y)p^t \ \forall \ p \ \in \ \mathbb{R}[x_1, .., x_n]$.So $L(gp^2) \geq 0 \ \forall \ p \ \in \ \mathbb{R}[x_1, ..., x_n]$ iff $M(g * y) \succeq 0$. □

Chapter 3

Polynomial Optimization

3.1 Semidefinite Optimization

In this section we see what is a semidefinite program and its dual . Its application to the max cut problem is summarized.

3.1.1 Semidefinite Program

Definition 3.1.1.1. Convex Cone : Let K $\neq \varnothing$ be a subset of $\mathbb{R}^n$. K is a convex cone if $\forall \alpha, \beta \in \mathbb{R}_{\geq 0}$ and $\forall x, y \in$ K , $\alpha x + \beta y \in$ K.

Definition 3.1.1.2. Dual Cone of K : K* := { y $\in \mathbb{R}^n$ | $\langle$ y , x $\rangle \geq 0$ $\forall$ x $\in$ K } where $\langle , \rangle$ is inner product defined on $\mathbb{R}^n$.

Consider $S^n = \{$ A $\in M_n(\mathbb{R}) \mid A = A^T\}$. It is a set of $n \times n$ symmetric matrices so dim S^n = n(n+1)/2.

Definition 3.1.1.3. Inner Product on S^n :

$$\langle A, B \rangle = Tr(A^T B) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij} \tag{3.1}$$

where $A = (a_{ij})$ $B = (b_{ij})$. A $\in$ S^n is positive semidefinite if $\forall$ v $\in$ $\mathbb{R}^n$ $v^t A v \geq 0$.

Proposition 3.1.1.4. *Equivalent conditions for being positive semidefinite :*
(i) $\forall$ v $\in$ $\mathbb{R}^n$ $v^t A v \geq 0$
(ii) All eigenvalues of A are non negative. $A = \lambda_1 e_1.e_1^t + ...\lambda_n e_n.e_n^t$*,where* λ_i*are eigenvalues.*
(iii) $A = LL^T$

Proof. $(i) \Rightarrow (ii)$ let λ be an eigenvalue of A and v corresponding eigenvector. $Av = \lambda$ v.So $v^t A v = \lambda v^t v$. $v^t v > 0$ and $v^t A v \geq 0 \Rightarrow \lambda \geq 0$.By Spectral theorem $\exists$ an orthogonal matrix U s.t. $A = UDU^T$ where D is diagonal matrix $(\lambda_1, ..., \lambda_n)$ with λ_i being eigenvalues of A.Let U= $(e_1, ..., e_n)$ where e_i are column vectors.So $A = \lambda_1 e_1.e_1^t + ...\lambda_n e_n.e_n^t$.

$(ii) \Rightarrow (iii)$ $A = \sum_{i=1}^{n}$ $\lambda_i e_i e_i^t$ where e_i are orthogonal vectors forming U. All λ_i non negative $\Rightarrow$ $\delta_i = \sqrt{\lambda_i}$ $\in$ $\mathbb{R}$.Let C =diagonal matrix $(\delta_1, ..., \delta_n)$ and let $L = UC$ then $A = LL^T$.

$(iii) \Rightarrow (i)$ Let v $\in$ $\mathbb{R}$. $v^t A v = v^t L L^T v = (v^t L)(v^t L)^t \geq 0$.

□

Definition 3.1.1.5. Cone of positive semidefinite matrices :
$S^n_+ := \{A \in S^n \mid$ A is positive semidefinite$\}$.It is a convex cone (can easily be proved from definition).

Proposition 3.1.1.6. *S^n_+ is self dual .i.e $S^{n*}_+ = S^n_+$*

Proof. Let B $\in S^n$ s.t. $\langle A, B\rangle \geq 0 \; \forall \; A \in S^n_+$.Let v $\in \mathbb{R}^n$. $v^t Bv = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j b_{ij}$.Let $V = vv^t$.Then $(i,j)^{th}$ entry of V is $v_i v_j$.By Proposition 3.1.1.4 (iii) above $V = vv^t \Rightarrow V \in S^n_+$.Therefore, $\langle V, B\rangle \geq 0$.So $\langle V, B\rangle = Tr(V^T B) = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j b_{ij} = v^t Bv \geq 0$.v was arbitray vector in $\mathbb{R}$ .So $B \in S^n_+$.Therefore, by definition of dual cone $S^{n*}_+ \subseteq S^n_+$.Let A , A' $\in S^n_+$.Then from proof of $(ii) \Rightarrow (iii)$ of Proposition 3.1.1.4 we know that $A = \sum_{i=1}^{n} \lambda_i u_i u_i^t$ with λ_i non negative.So $\langle A, A'\rangle = \langle A', A\rangle = Tr(A'^T A) = Tr(A'^T(\sum_{i=1}^{n} \lambda_i u_i u_i^t)) = \sum_{i=1}^{n} \lambda_i Tr(A'^T u_i u_i^t)$.But from the first part of this proof we know that $Tr(A'^T u_i u_i^t) = u_i^t A' u_i$.So $\langle A, A'\rangle = \sum_{i=1}^{n} \lambda_i u_i^t A' u_i$ with λ_i non negative and A' positive semidefinite.So $\langle A, A'\rangle \geq 0$.Therefore, by definition of dual cone if $A \in S^n_+$ then A $\in S^{n*}_+$.So $S^n_+ \subseteq S^{n*}_+$.So $S^{n*}_+ = S^n_+$. □

Definition 3.1.1.7. Semidefinite program : Let C, $A_1, ..., A_r \in S^n$ and $b = (b_1, ..., b_r) \in \mathbb{R}^n$

Standard Primal form :

$$p^* = sup\{\langle C, X\rangle \mid \langle A_i, X\rangle = b_i \forall i = 1, ..., r \; and \; X \succeq 0\} \tag{3.2}$$

Standard Dual form :

$$d^* = inf\{b^t y \mid \sum_{i=1}^{r} y_i A_i - C \succeq 0\} \tag{3.3}$$

Proposition 3.1.1.8. *$p^* \leq d^*$.i.e Weak duality always holds*

Proof. Let X be feasible for (3.2) and y be feasible for (3.3).So X $\succeq 0$ and $\sum_{i=1}^{r} y_i A_i - C \succeq 0$.So by Proposition 3.1.1.6 and definition of dual cone $\langle X, \sum_{i=1}^{r} y_i A_i - C\rangle \geq 0$.So $\langle X, \sum_{i=1}^{r} y_i A_i - C\rangle = \sum_{i=1}^{r} y_i \langle X, A_i\rangle - \langle X, C\rangle \geq 0$.Therefore, $\sum_{i=1}^{r} y_i b_i \geq \langle C, X\rangle$.So $b^t y \geq \langle C, X\rangle$.So $p^* \leq d^*$. □

If p^* is bounded above and $\exists$ X $\in S^n$ which is strictly feasible for (3.2) (i.e X feasible s.t X $\succ 0$) then $p^* = d^*$.Similarly if d^* is bounded from below and $\exists$ y $\in \mathbb{R}^r$ which is strictly feasible for (3.3)(i.e $\sum_{i=1}^{r} y_i A_i - C \succ 0$) then $p^* = d^*$.With the help of convexity theory it can be shown that in the latter case $\exists X_0$ feasible for (3.2) such that $d^* \leq \langle C, X\rangle$ (see [1] for proof).So then weak duality implies $p^* = d^*$.With Ellipsoid method we can solve SDP in polynomial time.Interioir point methods provide efficient algorithms to solve semidefinite programs (upto any precision).

3.1.2 Applications of Semidefinite Programs to Combinatorial Problems

Max Cut Problem: Given a graph $G = (V, E)$ where $V = v_i; i \in I$ with I index set of vertices. E is set of edges of G. Let denote weight assigned to edge between v_i, v_j(if exists) by $w_{\{i,j\}}$. Max cut problem asks us to partition the set of vertices V in 2 sets such that the total weight of edges crossing the partition is maximum.

We can reformulate it in terms of a polynomial optimization problem. To each vertex $v_i \in \mathbb{R}^I$ we assign a variable x_i such that it satisfies $x_i^2 = 1$. Any solution $\underline{x} = (x_i)_{i\in I} \in \mathbb{R}^I$satisfying $x_i^2 = 1\ \forall i \in I$ gives us a partition of vertices (say $I_1 = \{i \in I \mid x_i = 1\}$ and $I_2 = \{i \in I \mid x_i = -1\}$) . And for any partition we have a solution $\underline{x} = (x_i)_{i\in I}$satisfying $x_i^2 = 1\ \forall i \in I$. Now say we have a partition corresponding to $\underline{x}$. Then the total weight of edges crossing the partition is $\sum_{\{i,j\}\in E,}(1 - x_ix_j)w_{\{i,j\}}/2$. $\because$ If $v_i, v_j \in$ same side of partition then $x_i = x_j$. $\therefore x_ix_j = 1$. So weight of edge between v_i, v_j(if exists) is not counted in the sum. And if v_i, v_j belong to different sides then $x_i = -x_j$. $\therefore 1 - x_ix_j = 2$ and so $1/2$ appears in the sum. So Max cut problem can be reformulated as

$$mxcut(G, w) = max_{x\in\mathbb{R}^I}\{\sum_{\{i,j\}\in E}(1 - x_ix_j)w_{\{i,j\}}/2 \mid x_i^2 = 1\forall i \in I\} \tag{3.4}$$

we have an SDP relaxation to this problem.Consider the following SDP problem

$$mxcutsdp(G, w) = max_{X\in S^n}\{\sum_{\{i,j\}\in E}(1 - X_{ij})w_{\{i,j\}}/2 \mid X_{ii} = 1\forall i \in I \ and\ X \succeq 0\} \tag{3.5}$$

Now observe the feasible region for (2.4).Let x $\in \mathbb{R}^I$ such that $x_i^2 = 1\forall i \in I$.Let X = $x.x^t$(positive semidefinite by Proposition 3.1.1.4).$X_{ii} = x_i^2 = 1$.So X is feasible for (2.5).And in this case ,$\sum_{\{i,j\}\in E}(1 - X_{ij})w_{\{i,j\}}/2 = \sum_{\{i,j\}\in E}(1 - x_ix_j)w_{\{i,j\}}/2$.Therefore set on which max is calculated for (2.4) $\subseteq$ set on which max is calculated for (2.5).
Therefore,mxcut(G,w) $\leq$ mxcutsdp(G,w).So using semidefinite programming we get a bound on max cut.

3.1.3 Lovász sandwich inequalities

Let $G = (V, E)$ be a graph.

Definition 3.1.3.1. Stable Set :$S \subseteq V$ is a stable set with respect to G if $\forall\ v_i, v_j$ $\in$S,$\{i, j\} \notin$E.(i.e there are no edges in the subgraph induced by S in G)

Definition 3.1.3.2. Stability number of G $\alpha(G)$:Stability number is the maximum cardinality of a stable set in G.

Definition 3.1.3.3. Characteristic vector of $S \subseteq V$: χ^S :

$$\begin{aligned}\chi^S(i) &= 1 \ \ if\ v_i \in S\\ &= 0 \ \ if\ v_i \notin S\end{aligned}$$

Now we reformulate $\alpha(G)$ in terms of a polynomial optimization problem.To each vertex v_i we assign a variable x_i.Now consider $\underline{x} = (x_i) \in \mathbb{R}^I$ satisfying $x_i^2 = x_i\forall i$ $\in$I and $x_ix_j = 0\ \forall\{i, j\} \in E$.(So x_i =1 or 0 $\forall i \in I$).Consider the set $S = \{v_i \in V|x_i = 1\}$.If $v_i, v_j \in S\ i \neq j$,then $\{i, j\} \notin$ E,because if it did it would imply $x_ix_j = 0 \Rightarrow 1 = 0$ giving contradiction.So S is a stable set.$\sum_{i\in I} x_i = \sum_{i\in S} x_i + \sum_{i\in I-S} x_i = \sum_{i\in S} x_i = |S|$.So $\forall\underline{x}\ \in \mathbb{R}^I$ satisfying $x_i^2 = x_i\forall i\ \in I$ and $x_ix_j = 0\ \forall\{i, j\}$ $\in$E,$\sum_{i\in I} x_i$ gives cardinality of a stable set in G.And given any stable set S in G,χ^S satisfies the above conditions with $\sum_{i\in I} x_i = |S|$.So,

$$\alpha(G) = max_{\underline{x}\in\mathbb{R}^I}\{\sum_{i\in I} x_i \mid x_ix_j = 0\ \forall\ \{i, j\} \in E\ and\ x_i^2 = x_i\ \forall\ i \in I\} \tag{3.6}$$

Definition 3.1.3.4. Theta number of G ϑ(G) : is defined by

$$\vartheta(G) = max_{X \in S^n}\{\sum X_{ij} \mid X_{ij} = 0 \;\forall\; \{i,j\} \;\in E, Tr(X) = 1 \; and \; X \succeq 0\} \tag{3.7}$$

Definition 3.1.3.5. Chromatic number of G $\chi(G)$: Minimum number of colours required to colour vertices of G such that no 2 adjacent (vertices with an edge between them) vertices have same colour.

$\bar{G}$ is complement graph of G.

Theorem 3.1.3.6. ***Lovsz Inequality :***

$$\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G}) \tag{3.8}$$
$$\alpha(\bar{G}) \leq \vartheta(\bar{G}) \leq \chi(G) \tag{3.9}$$

Proof. (3.9) follows from (3.8) .So its enough to prove (3.8).

$\alpha(G) \leq \vartheta(G)$:
Let $\underline{x}$ $\in$ feasible region for $\alpha(G)$. Therefore from above we know it corresponds to a stable set S with $\Sigma_{i \in I} x_i = \;\; |S|$ Consider X $= \underline{x}.\underline{x}^t \;/\; |S|$. So $X \succeq 0$ by Proposition 3.1.1.4

$$\begin{aligned} Tr(X) &= \Sigma_{i \in I} {x_i}^2 / \,|S| \\ &= \Sigma_{i \in I} x_i / \,|S| \\ &= 1 \end{aligned}$$

because ${x_i}^2 = x_i \;\forall\, i \in I$. $X_{ij} = x_i x_j = 0 \quad$ if $\;\{i,j\} \in$ E. Therefore X $\;\in$ feasible region for $\vartheta(G)$. And

$$\begin{aligned} \sum X_{ij} &= \sum_{i \in I} X_{ii} + \sum_{i \neq j} X_{ij} \\ &= \sum_{i \in I} x_i^2/|S| + \sum_{i \neq j : \{i,j\} \in E} x_i x_j/|S| + \sum_{i \neq j : \{i,j\} \notin E} x_i x_j/|S| \\ &= \sum_{i \in S} x_i^2/|S| + \sum_{i \neq j : x_i, x_j \in S} x_i x_j/|S| \qquad (\because x_i = 0 \; if \; x_i \notin S) \\ &= (\sum_{i \in S} x_i)^2/|S| \\ &= |S|^2/|S| = |S| \end{aligned}$$

Therefore $\;\alpha(G) \leq \vartheta(G)$ .....(because set on which max is $\leq$ set on which max calculated for $\alpha(G)$ is calculated for $\vartheta(G)$)

$\vartheta(G) \leq \chi(\bar{G})$:
Lets say vertices of $\bar{G}$ can be colored with r colors s.t no 2 adjacent vertices in $\bar{G}$ have same color.Let C_i be the set of vertices colored with i^{th} color. Let y_i be the characteristic vector of C_i. $\;C_i \cap C_j = \phi$ for $i \neq j$ and $\cup_{i=1}^{r} C_i$ is all vertices in G. Therefore $e = \sum_{i=1}^{r} y_i$ where e $=$

$[1,...,1] \in \mathbb{R}^{|I|}$. Now consider a matrix X feasible for (3.7).

$$\begin{aligned}
\langle X, \sum_{i=1}^{r} y_i y_i^t \rangle &= \sum_{i=1}^{r} \langle X, y_i y_i^t \rangle \\
&= \sum_{i=1}^{r} \sum_{i_1,i_2} X_{i_1,i_2} y_{i_{i1}} y_{i_{i2}} \\
&= \sum_{i=1}^{r} \sum_{v_{i_1},v_{i_2} \in C_i} X_{i_1 i_2} \qquad (\because if v_j \notin C_i \; y_{i_j} = 0) \\
&= \sum_{i=1}^{n} \sum_{v_j \in C_i} X_{jj}
\end{aligned}$$

Because $if i_1 \neq i_2$ and $v_{i_1}, v_{i_2} \in C_i \Rightarrow$ there is no edge between v_{i_1}, v_{i_2} $in\ \bar{G} \Rightarrow$ there is an edge between v_{i_1}, v_{i_2} $inG \therefore X$ is feasible for $\Rightarrow X_{i_1 i_2} = 0$. Therefore $\langle X, \sum_{i=1}^{r} y_i y_i^t \rangle = T_r(X) = 1$ .Now consider,

$$\begin{aligned}
Y &= \sum_{i=1}^{r} (ry_i - e)(ry_i - e)^t \\
&= \sum_{i=1}^{r} r^2 y_i y_i^t - (\sum_{i=1}^{r} ry_i)e^t - e\sum_{i=1}^{r} ry_i^t + \sum_{i=1}^{r} ee^t \\
&= \sum_{i=1}^{r} r^2 y_i y_i^t - ree^t
\end{aligned}$$

So $Y \succeq 0$ (by prop (1.1)) and we have X $\succeq 0$.
Therefore $\langle X, Y \rangle \geq 0$. Therefore $r^2 \langle X, \sum_{i=1}^{r} y_i y_i^t \rangle$- r $\langle X, ee^t \rangle \geq 0$. So $r^2 \geq r(\sum X_{ij})$. So $r \geq \sum X_{ij}$. Therefore $\vartheta(G) \leq \chi(\bar{G})$.

□

Definition 3.1.3.7. Clique number:w(G): Let G=(V,E) be a graph. A clique is a graph in which every two distict vertices are joint by an edge.w(G) is the maximum cardinality of a clique contained in G.

Definition 3.1.3.8. Perfect Graph : G =(V,E) is a perfect graph if ∀ induced subgraph H of G w(H)= $\chi(H)$

Theorem 3.1.3.9. ***The Strong Perfect Graph Theorem****: A Graph G=(V,E)is perfect iff G neither contains an odd cycle of length atleast 5 nor complement of such a cycle as an induced subgraph.*

Berge conjectured this theorem in 1961 and in 2004 this theorem was proved by Maria Chudnovsky,Neil Robertson,Paul Seymour and Robin Thomas. Ref [2]

Theorem 3.1.3.10. ***Weak Perfect Graph Theorem****: If G = (V,E) is a perfect graph then $\bar{G}$is also a perfect graph.*

Remark 3.1. Strong implies weak perfect theorem (just from def of $\bar{G}$). Therefore for a perfect graph G .$\bar{G}$ is perfect.So $\chi(\bar{G})$ = w($\bar{G}$). A clique in $\bar{G}$ corresponds to a stable set in G. So $\alpha(G) \geq w(\bar{G})$= $\chi(\bar{G})$. We already know $\alpha(G) \leq \chi(\bar{G}) \therefore \alpha(G) = \chi(\bar{G})$ for perfect graphs.∴ in case of perfect graphs sandwich inequality => $\alpha(G) = \vartheta(G) = \chi(\bar{G})$. So by using semidefinite program we can calculate $\vartheta(G)$ and so $\alpha(G)$.

3.2 Sum of squares

3.2.1 Relation between sum of squares and being positive

We see how to reformulate "a polynomial being a sum of squares" in terms of a semidefinite program .In this section we study the relation between positivity of a polynomial and it being a sum of squares with the help of Putinar's result. Lasserre's Hierarchy is summarized as well .

For any $f \in \mathbb{R}[x], f \in \mathbb{R}[x]_{2t}$ for some $t \geq 0$. So $f = [f]^t[x]_{2t}$ where $[f] = (f_\alpha)_{|\alpha| \leq 2t}$ coefficient vector .

Theorem 3.2.1.1. *Let $f \in \mathbb{R}[x]_{2t}$. Then f is a sum of squares iff $S \neq \phi$,where*

$$S = \{ X \in S^{N_t^n} \mid \sum_{\beta,\gamma \in N_t^n \text{ such that } \beta+\gamma=\alpha} X_{\beta,\gamma} = f_\alpha \ \forall \ \alpha \in N_{2t}^n \text{ and } X \succeq 0\} \tag{3.10}$$

Proof. $\Rightarrow$: Let f be a sum of squares. So $f = p_1^2 + ... + p_m^2$ for some $m \geqslant 1$ and $p_i \in \mathbb{R}[X]_t \ \forall \ 1 \leqslant i \leqslant m$. $P_i^2 = [x]_t^t[p_i][p_i]^t[x_t] \quad \forall \ 1 \leqslant i \leqslant m$. So

$$[f]^T[u]_{2t} = [x]_t^T(\sum_{i=1}^m [p_i].[p_i]^t)[u]_t$$

We have $[p_i].[p_i]^t \succeq 0 \ \ \forall \ 1 \leqslant i \leqslant m$.Let $[y]_t$ be any vector (compatible with p_i),then

$$[y]^T(\sum_{i=1}^m [p_i][p_i]^t)[y] = \sum_{i=1}^m [y]^T([p_i][p_i]^t)[y] \geq 0$$

So $\sum_{i=1}^m [p_i][p_i]^t \succeq 0$. Let P $=\sum_{i=1}^m [p_i][p_i]^t$. Then,

$$[f]^t[u]_{2t} = [x]_t^T P[x]_t = \sum_{|\beta| \leqslant t \, , |\gamma| \leqslant t} P_{\beta\gamma} x^\beta x^\gamma \tag{3.11}$$

Equating coefficient of α where $|\alpha| \leq 2t$ we get

$$f_\alpha = \sum_{s.t \ \beta+\gamma=\alpha \text{ and } |\beta|,|\gamma| \leq t} P_{\beta,\gamma} \tag{3.12}$$

$\forall \ \alpha$ with $|\alpha| \leq 2t$. Therefore P $\in$ S.So f is a sum of squares $\implies \ S \neq \phi$.

$\Leftarrow$:

If $S \neq \phi$, then $\exists \ X \succeq 0 \in S^{|N_t^n|}$ such that

$$\sum_{\beta,\gamma \in N_t^n \ \ s.t. \ \beta+\gamma=\alpha} X_{\beta,\gamma} = f_\alpha \ \forall \ \alpha \ \in N_{2t}^n \tag{3.13}$$

From this we get, $[f]^T[x]_{2t} = [x]_t^T X[x]_t$. From proof of proposition 3.1.3,we get that X$=\lambda_1 e_1 e_1^t + ... + \lambda_m e_m e_m^t$ where m$=|N_t^n|$ with $\lambda_i \geq 0$. Therefore $\sqrt{\lambda_i} \ \in \mathbb{R}$. So

$$[f]^T[x]_{2t} = \sum_{i=1}^m [x]_t^T[\sqrt{\lambda_i}e_i][\sqrt{\lambda_i}e_i]^t[x]_t.$$

Therefore f $= \sum {p_i}^2$ where ${p_i}^2 = [x]_t^T[\sqrt{\lambda_i}e_i][\sqrt{\lambda_i}e_i]^t[x]_t$. So $S \neq \phi$ implies f is a sum of squares. □

Theorem 3.2.1.2. *Let* $g_1, ..., g_m \in \mathbb{R}[x_1, ..., x_n]$. *Let deg(*$g_i$*)=* d_i $\forall 1 \leq i \leq m$. *Then f* $= (p_{01}^2 + ...p_{0k_0}^2) + g_1(p_{11}^2 + ...p_{1k_1}^2) + ... + g_m(p_{m1}^2 + .. + p_{mk_m}^2)$ *for some* $p_{ij} \in \mathbb{R}[x_1, .., x_n]$ *iff* $S_(g_1, ...g_m) \neq \phi$ *where*

$$S_(g_1, ...g_m) := \{(X_0, X_1, ..X_m) \,|X_0 \in S^{N_t^n}, X_i \in S^{N^n_{\lfloor (2t-d_i)/2 \rfloor}} \; \forall 1 \leq i \leq m, X_i \succeq 0 \; \forall \; 0 \leq i \leq m$$
$$\text{and } f_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma} \; \forall \; \alpha \in N_{2t}^n\}.$$

Proof. $\Rightarrow$:
Let f $= (p_{01}^2 + ...p_{0k_0}^2) + g_1(p_{11}^2 + ...p_{1k_1}^2) + ... + g_m(p_{m1}^2 + .. + p_{mk_m}^2)$. Similarly as before we can write

$$[f]^T[x]_{2t} = [x]_t^T P_0 [x]_t + [g_1]^T[x][x]_{t-\lceil d_1/2 \rceil}^T P_1 [x]_{t-\lceil d_1/2 \rceil} + ... + [g_m]^T[x][x]_{t-\lceil d_m/2 \rceil}^T P_m [x]_{t-\lceil d_m/2 \rceil}$$

where $P_i \succeq 0 \; \forall \; 0 \leq i \leq m$. Equating the coefficients we get,

$$f_\alpha = \sum_{\beta+\gamma=\alpha} P_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} P_{i\beta,\gamma} \tag{3.14}$$

$\forall \; \alpha$. Therefore $S_(g_1, ..., g_m) \neq \phi$.

$\Leftarrow$:
If $S_(g_1, ..., g_m) \neq \phi$,we get $(X_0, .., X_m)$ such that

$$f_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma} \; \forall \; \alpha \tag{3.15}$$

So we have

$$[f]^T[x] = [x]_t^T X_0 [x]_t + [g_1]^T[x][x]^T X_1 [x] + ... + [g_m]^T[x][x]^T X_m [x]$$

As we saw before if $X_i \succeq 0$, we can write $[x]^T X_i [x] = p_{i1}^2 + ... + p_{ik_i}^2$ for some $k_i \geq 0$. Therefore,

$$f = (p_{01}^2 + ... + p_{0k_0}^2) + g_1(p_{11}^2 + ... + p_{1k_1}^2) + ... + g_m(p_{m1}^2 + ... + p_{mk_m}^2) \tag{3.16}$$

□

Remark 3.2. $g(x_1, .., x_n)$ is a sum of squares implies $g(x_1, .., x_n) \geq 0$.But $g(x_1, .., x_n) \geq 0$ need not imply that $g(x_1, .., x_n)$is a sum of squares.

If K is of the form described in preliminaries and if K is compact then we can use results of Schmdgen and Putinar to characterize positivity of f over K. Let $g = (g_1, ..., g_m)$ be used to describe K.

Definition 3.2.1.3. Q(g) :

$$Q(g) := \{\sigma_0 + \sigma_1 g_1 + ... + \sigma_m g_m | \sigma_i \textit{ is sum of squares } \forall \; 0 \leq i \leq m\} \tag{3.17}$$

Definition 3.2.1.4. $Q_t(g)$ **:**

$$Q_t(g) := \{\sigma_0 + \sigma_1 g_1 + ... + \sigma_m g_m \, |\sigma_i \textit{ is sum of squares } \forall \; 0 \leq i \leq m$$
$$\textit{and } deg(\sigma_i g_i) \leq 2t \textit{ and } deg(\sigma_0) \leq 2t\}$$

Definition 3.2.1.5. $\Gamma(g)$**:** $\Gamma(g)$ is defined as a quadratic module generated by $g^e := g_1^{e_1}...g_m^{e_m}$ where $e \in \{0,1\}^m$

Archimedean Condition: $\exists$ R>0 such that R-$x_1^2 - .... - x_n^2 \ \in Q(g)$

Lemma 3.2.1.6. *Archimedean condition holds implies K is compact.*

Proof. K= $\{x \in \mathbb{R}^n | g_1(x) \geq 0, ..., g_m(x) \geq 0\}$. So K is closed. And if Archimedean condition holds, then K is bounded (because ($let\ a = (a_1, ..., a_n) \ \in K$ Archimedean condition holds $\implies \ \exists\ R > 0 s.t R - x_1^2 - .... - x_n^2 = \sigma_0 + \sigma_1 g_1 + ... + \sigma_m g_m$ for some σ_i sum of squares. Therefore$R - a_1^2 - ... - a_n^2 = \sigma_0(a) + \sigma_1(a) g_1(a) + ... + \sigma_m(a) g_m(a)$. So on K, $R - (a_1^2 + ... + a_n^2) \geq 0$. Therefore $|a| \leq R$)). So K is closed and bounded. Therefore K is compact. □

Theorem 3.2.1.7. ***(Schmüdgen)*** *:Let K be compact.f(x)>0* $\forall\ x \ \in K \implies f \ \in \Gamma(g)$.

Theorem 3.2.1.8. ***(Putinar)*** *:Lets assume archimedean condition holds.Then f(x)>0* $\forall\ x \ \in K \implies f \ \in Q(g)$.

3.2.2 Lasserre Hierarchy

Let t$\geq \lceil deg(f)/2 \rceil$.Lasserre introduced relaxations to the polynomial optimization problem based on Putinar's result.

Consider

$$f_t^{sos} = sup_{\lambda\ \in \mathbb{R}} \{\lambda : f - \lambda \ \in Q_t(g)\} \tag{3.18}$$

Lemma 3.2.2.1. $f_t^{sos} \leq f_{min}$

Proof. Let λ be such that $f - \lambda \ \in \ Q_t(g) \subseteq Q(g)$. So $f - \lambda > 0$ on K. $f_{min} = inf_K f(x)$. So $f_{min} \geq f_t^{sos}$. □

Theorem 3.2.2.2. ***Lasserre Hierarchy :***

$$f_t^{sos} \leq f_{t+1}^{sos} \leq ..... \leq f_{min} \tag{3.19}$$

Proof. Now $Q_{t+i}(g) \subseteq Q_{t+i+1}(g)\ \forall\ i \ \in \mathbb{N}$. Therefore

$$\{\lambda \ \in \mathbb{R} | f - \lambda \ \in Q_{t+i}(g)\} \subseteq \{\lambda \ \in \mathbb{R} | f - \lambda \ \in Q_{t+i+1}(g)\}\ \forall\ i \ \in \mathbb{N}.$$

So $f_{t+i}^{sos} \leq f_{t+i+1}^{sos}\ \forall\ i \ \in \mathbb{N}$. Therefore we get $f_t^{sos} \leq f_{t+1}^{sos} \leq ..... \leq f_{min} = inf_K f(x)$ by using lemma 3.2.2.1. □

We know that $f - \lambda \ \in Q_t(g) \iff S_(g_1, ..., g_m)^{f-\lambda} \neq \phi$,where

$$S_(g_1, ..., g_m)^{f-\lambda} = \{(X_0, X_1, ..X_m)\ | X_0 \ \in S^{N_t^n}, X_i \ \in S^{N^n_{\lfloor (2t-d_i)/2 \rfloor}}\ \forall\ 1 \leq i \leq m, X_i \succeq 0\ \forall\ 0 \leq i \leq m$$

$$and\ (f-\lambda)_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma}$$

$$\forall\ \alpha \ \in N_{2t}^n\}$$

So ,

$$S_(g_1,...,g_m)^{f-\lambda} = \{(X_0, X_1, ..X_m)\ |X_0\ \in S^{N_t^n}, X_i\ \in S^{N^n_{\lfloor(2t-d_i)/2\rfloor}}\ \forall\ 1 \leq i \leq m, X_i \succeq 0$$
$$\forall\ 0 \leq i \leq m\ and\ f_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma}$$
$$\forall\ \alpha\ \in N^n_{2t} - (0, ..., 0)\ and\ f_{(0,....,0)} - \lambda = X_{0_{00}} + \sum_{i=1}^{m} g_{i0} X_{i_{00}}\}$$

Therefore ,

$$\{\lambda : f - \lambda\ \in Q_t(g)\} = \{f_{(0,...,0)} - X_{0_{00}} - \sum_{i=1}^{m} g_{i0} X_{i_{00}}\ |X_0\ \in S^{N_t^n}, X_i\ \in S^{N^n_{\lfloor(2t-d_i)/2\rfloor}}\ \forall\ 1 \leq i \leq m, X_i \succeq 0$$
$$\forall\ 0 \leq i \leq m$$
$$and\ f_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma}$$
$$\forall\ \alpha\ \in N^n_{2t} - (0, ..., 0)\}$$

Therefore ,

$$f_t^{sos} = f_{(0,...,0)} + sup\{-X_{0_{00}} - \sum_{i=1}^{m} g_{i0} X_{i_{00}}\ |X_0\ \in S^{N_t^n}, X_i\ \in S^{N^n_{\lfloor(2t-d_i)/2\rfloor}}$$
$$\forall\ 1 \leq i \leq m, X_i \succeq 0\ \forall\ 0 \leq i \leq m$$
$$and\ f_\alpha = \sum_{\beta+\gamma=\alpha} X_{0\beta,\gamma} + \sum_{i=1}^{m} \sum_{\delta \leq d_i} g_{i\delta} \sum_{\beta+\gamma=\alpha-\delta} X_{i\beta,\gamma}$$
$$\forall\ \alpha\ \in N^n_{2t} - (0, ..., 0)\}$$

Let

$$\text{X}= \begin{bmatrix} X_0 & 0 & 0 & \dots \\ 0 & X_1 & 0 & \dots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & X_m \end{bmatrix}.$$

Define $C_0 = (C_{0_{\beta,\gamma}})_{\beta,\gamma\ \in N_t^n}$ where $C_{0_{00}} = -1$ and everywhere else 0. Define $C_i = (C_{i_{\beta,\gamma}})_{\beta,\gamma\ \in N^n_{t-\lceil d_i/2\rceil}}$ where $C_{i_{00}} = -g_{i0}$ and everywhere else 0 $\forall\ 1 \leq i \leq m$. Define

$$\text{C}= \begin{bmatrix} C_0 & 0 & 0 & \dots \\ 0 & C_1 & 0 & \dots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & C_m \end{bmatrix}$$

$\forall\ \alpha\ \in N_t^n \backslash (0, ..., 0)$ define $A_{\alpha 0} = (A_{\alpha 0_{\beta,\gamma}})_{\beta,\gamma\ \in N_t^n}$ where $A_{\alpha 0_{\beta,\gamma}} = 1$ if $\beta + \gamma = \alpha$ and otherwise 0. Define $A_{\alpha i} = (A_{\alpha i_{\beta,\gamma}})_{\beta,\gamma\ \in N^n_{t-\lceil (d_i)/2\rceil}}$ where $A_{\alpha i_{\beta,\gamma}} = g_{i\delta}$ where $\beta + \gamma = \alpha - \delta$ and otherwise 0. Define

$$A_\alpha = \begin{bmatrix} A_{\alpha 0} & 0 & 0 & \dots \\ 0 & A_{\alpha 1} & 0 & \dots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & A_{\alpha m} \end{bmatrix}$$

Now one can check that

$$f_t^{sos} = f_{(0,....,0)} + sup\{\langle C, X\rangle | X \succeq 0 \; and \; \langle A_\alpha, X\rangle = f_\alpha \; \forall \; \alpha \; \in N_t^n \backslash (0,...,0)\} \tag{3.20}$$

The dual program can be expressed as

$$f_t^{mom} = f_{(0,..,0)} + inf\{\sum_{\alpha \; \in N_t^n \backslash (0,...,0)} f_\alpha y_\alpha | \sum_{\alpha \; \in N_t^n \backslash (0,...,0)} y_\alpha A_\alpha - C \succeq 0\} \tag{3.21}$$

Therefore ,$f_t^{mom} = f_{(0,..,0)} + inf\{\sum_{\alpha \; \in N_t^n} f_\alpha y_\alpha | \sum_{\alpha \; \in N_t^n \backslash (0,...,0)} y_\alpha A_\alpha - C \succeq 0 \; and \; y_{(0,...,0)} = 1\}$

$$\sum_{\alpha \; \in N_t^n \backslash (0,...,0)} y_\alpha A_\alpha - C = \begin{bmatrix} M(y) & 0 & 0 & \dots \\ 0 & M(g_1 * y) & 0 & \dots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & M(g_m * y) \end{bmatrix}$$

when $y_{(0,...,0)} = 1$,M(y) and $M(g_i * y)$ are moment matrices described in preliminaries.

Therefore ,$\sum_{\alpha \; \in N_t^n \backslash (0,...,0)} y_\alpha A_\alpha - C \succeq 0 \iff M(y) \succeq 0 \; and \; M(g_i * y) \succeq 0 \; \forall \; 1 \leq i \leq m \; and \; y_{(0,...,0)} = 1$.By result from preliminaries, $M(y) \succeq 0 \iff L(p) \geq 0 \; \forall \; p$:a sum of squares and $M(g_i * y) \succeq 0 \iff L(p) \geq 0 \; \forall \; p \; \in \; g_i \times$ (a sum of squares) where L corresponds to Linear functional associated to y_α.And $y_{(0,...,0)} = 1 \iff L(1) = 1$.Therefore,

$$f_t^{mom} = inf_{L \; \in \mathbb{R}[x]_{2t}^*}\{L(f) | L(1) = 1 \; and \; L(p) \geq 0 \; \forall \; Q_t(g)\}..........................(MOMt) \tag{3.22}$$

where $\mathbb{R}[x]_{2t}^*$ is a set of linear functionals on $\mathbb{R}[x]_{2t}$.$f_{min} = inf_K f(x) := f_{min}$.

Theorem 3.2.2.3. ***Lasserre:*** *Assume that the Archimedean condition holds.then* $f_{min} = \lim_{t \to \infty} f_t^{sos}$

Proof. So we have to prove that given any $\epsilon > 0 \; \exists \; t_0$ such that $f_{t_0}^{sos} \geq f_{min} - \epsilon$.(Because f_t^{sos} is a non decreasing sequence such that $f_t^{sos} \leq f_{min} \; \forall \; t)$.$f_{min} = inf_K f(x)$. So $f - f_{min} \geq 0$ on K. Therefore for any $\epsilon > 0, f - f_{min} + \epsilon > 0$ on K.So by Putinar's result $f - f_{min} + \epsilon \; \in \; Q(g)$. Therefore $f - f_{min} + \epsilon \; \in \; Q_{t_0}(g)$ for some t_0. So $f_t^{sos} \geq f_{min} - \epsilon$.(By definition of f_t^{sos}.) So $f_{min} = \lim_{t \to \infty} f_t^{sos}$ □

3.3 Moments

Let μ be a measure on K.Define linear functional L_μ by

$$L_\mu(f) = \int_K f(x) d\mu = \sum_\alpha f_\alpha \int_K x^\alpha d\mu \tag{3.23}$$

From calculations in the previous section we see that $f_t^{sos} \leq f_t^{mom}$ (by weak duality).

Lemma 3.3.0.4. $f_{min} = inf\{L_\mu(f) | \mu$ *is a probability measure* $\}$.

Proof. Consider $inf\{L_\mu(f)|\mu\ is\ a\ probability\ measure\}$. We have that $f_{min} = inf\{f(x)|g_1(x) \geq 0, ..., g_m(x) \geq 0\}$.If we fix any x in K and consider Dirac measure associated to x ,then

$$L_\mu(f) = \int_{y\ \in K} f(y)d\mu = \int_{y\ \in K} f(y)d\delta_x(y) = f(x). \tag{3.24}$$

And Dirac measure is a probability measure. So $inf\{L_\mu(f)|\mu\ is\ a\ probability\ measure\} \leq f_{min}$ and

$$L_\mu(f) = \int_{y\ \in K} f(y)d\mu . L_\mu(f) \geq f_{min} \int_{y\ \in K} d\mu = f_{min} \tag{3.25}$$

whenever μ is a probability measure. □

Definition 3.3.0.5. f_t^{mom} **:**

$$f_t^{mom} = inf\{L(f)|L(1) = 1\ and\ L(p) \geq 0\ \forall\ p\ \in\ Q_t(g)\} \tag{3.26}$$

Theorem 3.3.0.6. ***Haviland :*** *$L = L_\mu$ for some measure μ on K iff L is nonnegative on P(K) where P(K)= $\{p\ \in\ \mathbb{R}[x_1, ..., x_n]|p \geq 0\ on\ K\}$*

$Q_t(g) \subseteq P(K)$ and μ is a probability measure implies $L_\mu(1) = 1$.Therefore

$$\{L_\mu(f)|\mu\ is\ a\ probability\ measure\} \subseteq \{L(f)|L(1) = 1\ and\ L(p) \geq 0\ \forall\ p\ \in\ Q_t(g)\}$$

So $f_t^{mom} \leq f_{min}$. So we have

$$f_t^{sos} \leq f_t^{mom} \leq f_{min} \tag{3.27}$$

So [Lasserre thm] implies that if archimedean condition holds then

$$\lim_{t\to\infty} f_t^{sos} = \lim_{t\to\infty} f_t^{mom} = f_{min} \tag{3.28}$$

If for some t optimal value of f_t^{mom} is $L_\mu(f)$ (where μ is a probability measure),then $f_{min} = inf_\mu\{L_\mu(f)|\mu\ is\ a\ probability\ measure\ on\ K\} \geq f_t^{mom}$.Therefore $f_{min} = f_t^{mom}$ for that t.

Let L be a linear functional.Then M(L):= $(L(x^\alpha x^\beta))_{\alpha,\beta}$. Ker M(L) = $\{p\ \in\ \mathbb{R}[x_1, ...x_n]|p^T M(L) = 0\} = \{p\ \in\ \mathbb{R}[x_1, ..x_n]|L(pq) = p^T M(L)q = 0\ \forall\ q\ \in\ \mathbb{R}[x_1, ..x_n]\}$.Ker M(L) is an ideal in $\mathbb{R}[x_1, .., x_n]$.If M(L) $\succeq 0$,then $L(p^2) = 0 \implies p\ \in$Ker M(L)..($\because L(p^2) = 0 \implies p^T M(L)p = 0 \implies p^T N N^T p = 0 (\because M(L) \succeq 0) \implies (N^T p)^T N^T p = 0 \implies N^T p = 0 \implies p^T N N^T = 0 \implies p\ \in$ Ker M(L)).

Theorem 3.3.0.7. ***(Curto and Fialkow) :****Let L be a linear functional.If $M(L) \succeq 0$ and rank M(L) $= r < \infty$,then L has a unique representing measure μ.*

Proof. Let $J = KerM(L)$. p $\in \sqrt[\mathbb{R}]{J} \implies \exists k, p_1, ..., p_s$ such that $p^{2k} + \sum_{i=1}^{s} p_i^2\ \in J$ .Therefore $L(p^{2k} + \sum_{i=1}^{s} p_i^2) = 0$.As M(L) $\succeq 0$, $L(p_i^2) \geq 0\ \forall\ 1 \leq i \leq s$ and $L(p^{2k}) \geq 0$.Therefore,$L(p^{2k}) = 0$.So M(L) $\succeq 0$ implies $p^k\ \in J$.If k is even we can again derive $p^{k_1}\ \in J$ for $k_1 = k/2$. if k is odd $p^{k+1}\ \in J$ then again we get $p^{(k+1)/2}\ \in J$.Continuing in this way we get $p\ \in J$.Therefore $\sqrt[\mathbb{R}]{J} = J$.So J is real radical ideal.M(L) has finite rank r.Let columns indexed by $x^{\alpha_1}, ..., x^{\alpha_\gamma}$ be maximal linearly independent set of columns.Then $\lambda_1 x^{\bar{\alpha}_1} + ... + \lambda_r x^{\bar{\alpha}_r} = 0\ \in\ \mathbb{R}[x_1, ...x_n]/J \implies \lambda_1 x^{\alpha_1} + ... + \lambda_r x^{\alpha_r}\ \in\ J$.Therefore $L((\lambda_1 x^{\alpha_1} + ... + \lambda_r x^{\alpha_r}).x^\gamma) = 0\ \forall\ \gamma\ \in\ N^n$.Therefore $\lambda_1 L(x^{\alpha_1+\gamma}) + ... + \lambda_r L(x^{\alpha_r+\gamma}) = 0\ \forall\ \gamma\ \in\ N^n$.This implies $\lambda_i = 0\ \forall 1 \leq i \leq r$..($\because$ columns corresponding to $x^{\alpha_i}\ \in\ M(L)$ are linearly independent.Therefore dim$\mathbb{R}[x_1, ...x_n]/J \geq r$.And if $\exists$ any β such that $x^{\alpha_1}, ..., x^{\alpha_r}, x^{\beta}$ are linearly independent ,then columns in M(L) corresponding to $(x^{\alpha_1}, ..., x^{\alpha_r}, x^{\beta})$ are also linearly independent.($\because$ if not $\exists \lambda_1, ..., \lambda_{r+1}$such that λ_1(column corres tox^{α_1})+........+λ_{r+1}(column corres tox^{β}) $= 0 \implies L(\lambda_1 x^{\alpha_1} + ... + \lambda_r x^r + \lambda_{r+1} x^{\beta}) = 0 \implies \lambda_1 x^{\alpha_1} + ... + \lambda_r x^r + \lambda_{r+1} x^{\beta} \in J$ contradicting $x^{\alpha_1}, ..., x^{\alpha_r}, x^{\beta}$ linearly independent $\in \mathbb{R}[x_1, ...x_n]/J$ over $\mathbb{R}$).Therefore dim$\mathbb{R}[x_1, ...x_n]/J \geq r$.Therefore dim$\mathbb{R}[x_1, ...x_n]/J = r$. J is a real radical ideal.Therefore a radical ideal.So by proposition in preliminaries $V_{\mathbb{C}}(J) = V_{\mathbb{R}}(J)$ and $|V_{\mathbb{C}}(J)| = |V_{\mathbb{R}}(J)| = r$.Let $V_{\mathbb{C}}(J) = \{v_1, ..., v_r\} \subseteq \mathbb{R}^n$.Let $p_{v_i} \in \mathbb{R}[x_1, ..., x_n]$be interpolation polynomials at points of $V_{\mathbb{R}}(J)$.(Interpolation polynomials are described explicitly in preliminaries)$p_{v_i}(v_j) = \delta_{i,j}.(p_{v_i}^2 - p_{v_i})(v_j) = p_{v_i}^2(v_j) - p_{v_i}(v_j) = 0$.Therefore $p_{v_i}^2 - p_{v_i}(v_j) = 0 \; \forall \; 1 \leq j \leq r, 1 \leq i \leq r$..So $p_{v_i}^2 - p_{v_i} \in I(V_{\mathbb{C}}(J)) \; \forall \; 1 \leq i \leq r$ .But $I(V_{\mathbb{C}}(J)) = J(\because \; J \; is \; radical)$.Therefore $p_{v_i}^2 - p_{v_i} \in J \; \forall \; 1 \leq i \leq r$.Therefore $L(p_{v_i}^2) - L(p_{v_i}) = 0 \; \forall \; 1 \leq i (\leq r.L(p_{v_i}^2) > 0 \; \forall \; 1 \leq i \leq r.(\because \; M(L) \succeq 0 \implies L(p_{v_i}^2) \geq 0 \; \forall \; 1 \leq i \leq r$ and $L(p_{v_i}^2) = 0 \implies p_{v_i} \in J$(as $M(L) \succeq 0$)giving contradiction as $p_{v_i}(v_i) \neq 0$). Consider $\mu = \sum_{i=1}^r L(p_{v_i}) \delta_{v_i}$($\delta v_i$ dirac measure with respect to v_i).Then for any f$\in \mathbb{R}[x_1, ...x_n]$.$L_\mu(f) = \int_K f(x) d\mu = \sum_{i=1}^r L(p_{v_i}) \int_K f(x) d\delta_{v_i} = \sum_{i=1}^r L(p_{v_i}) f(v_i)$.By lemma $f - \sum_{v_i \in V_{\mathbb{C}}(J)} f(v_i) p_{v_i} \in I(V_{\mathbb{C}}(J)) = J$.Therefore $L(f) = \sum_{v_i \in V_{\mathbb{C}}(J)} f(v_i) L(p_{v_i})$.So $L_\mu(f) = L(f) \; \forall \; f \in \mathbb{R}[x_1, ..., x_n]$..So $L = L_\mu$. ☐

Definition 3.3.0.8. Truncated moment matrix of L : $L \in \mathbb{R}[x_1, ..., x_n]_{2t}^*$

$$M_t(L) := (L(x^\alpha x^\beta))_{\alpha, \beta \in N_t^n} \tag{3.29}$$

Definition 3.3.0.9. Flat Extenstion :$M_t(L)$ is a flat extension of$M_{t-1}(L)$ if rank $M_t(L)$ =rank $M_{t-1}(L)$.

Theorem 3.3.0.10. *Let $L \in \mathbb{R}[x_1, ..., x_n]_{2t}^*$..If $M_t(L)$ is a flat extension of $M_{t-1}(L)$ then $\exists \tilde{L} \in \mathbb{R}[x_1, ..., x_n]^*$ such that $\tilde{L} = L$ on $\mathbb{R}[x_1, ..., x_n]_{2t}$ and such that rank $M(\tilde{L}) = rank M_t(L)$.*

Proof. Consider I ideal generated by $Ker M_t(L) \in \mathbb{R}[x_1, ..., x_n], M_t(L)$ is a flat extension of $M_{t-1}(L)$.Therefore rank $M_t(L)$ =rank $M_{t-1}(L)$.So columns corresponding to x^α where $|\alpha| = t$ can be expressed in terms of columns corresponding to x^β with $|\beta| \leq t - 1$.Therefore $\bar{x^\alpha}$ with $|\alpha| = t$ can be expressed in terms of $\bar{x^\beta}$ with $|\beta| \leq t-1 \in \mathbb{R}[x_1, ..., x_n]/I$.So any f$\in \mathbb{R}[x_1, ..., x_n]$ can be expressed in terms of x^α with$|\alpha| \leq t$(modulo I).Now define $\tilde{L}(f)$ such that if $\bar{f} = \bar{g}$ for some g in $\mathbb{R}[x_1, ..., x_n]_{2t}$,then $\tilde{L}(f) = L(g)$.(well defined because if $\bar{g} = \bar{h}, g, hin\mathbb{R}[x_1, ..., x_n]_{2t}$ then $g - h \in Ker \; M_t(L)$.Therefore $L(g) = L(h)$).And rank $M(\tilde{L})$ = rank $M_t(L)$.(because $\bar{f} \in \mathbb{R}[x_1, .., x_n]/I$ can be expressed in terms of $\bar{x^\alpha}; |\alpha| \leq t$) ☐

Theorem 3.3.0.11. *Let $L \in \mathbb{R}[x_1, ..., x_n]_{2t}^*$ such that $M_t(L) \succeq 0, M_{t-\lceil d_j/2 \rceil}(g_j L) \succeq 0 \; \forall \; 1 \leq j \leq m$ and rank$M_t(L)$ =rank $M_{t-d_K}(L)$ where $d_K = max\{\lceil d_j/2 : 1 \leq j \leq mrceil\}$.Then L has a representing measure μ such that supp(μ) $\subseteq K$*

Proof. Similarly as in previous theorem,rank $M_t(L)$ = rank $M_{t-d_k}(L) \implies \exists \; \tilde{L} \in \mathbb{R}[x_1, ..., x_n]^*$ such that $\tilde{L} = L$ on $\mathbb{R}[x_1, .., x_n]_{2t}$ and such that rank $M(\tilde{L})$ = rank $M_t(L)$.By proposition in preliminaries $M(\tilde{L}) \succeq 0$ iff $\tilde{L}(\sigma) \geq 0 \; \forall \; \sigma$ sum of squares.$\tilde{L}(\sigma) = L(\sigma')$ for some σ' sum of squares and $M_t(L) \succeq 0$.Therefore $L(\sigma') \geq 0$.So $M(\tilde{L}) \succeq 0$.So by theorem of Curto and Fialkow $\tilde{L}$ and so L has a representing measure μ such that $L = \sum_{i=1}^r L(p_{v_i}) L_{v_i}$ where supp(μ) = $\{v_1, ..., v_r\} \subseteq \mathbb{R}^n$ with r= rank $M(\tilde{L})$ = rank $M_t(L)$.Now to show that $v_i \in K \; \forall \; 1 \leq i \leq r$. Rank $M_t(L)$ = rank $M_{t-d_K}(L)$.Therefore every p_{v_i} can be written in terms of polynomials of deg atmost $t - d_K$(modulo I= ideal generated by Ker $M_t(L)$.Say $p_{v_i} = h_i$(modulo I)$\forall \; 1 \leq i \leq r$.Therefore $p_{v_i} - h_i \in I \; \forall \; 1 \leq i \leq r$.For any $1 \leq j \leq m, M_{t-\lceil d_i/2 \rceil}(g_j L) \succeq 0$. and deg $h_i \leq t - d_K \leq t - \lceil d_j/2 \rceil$.Therefore $(g_j L)(h_i^2) \geq 0$.So $L(g_j h_i^2) \geq 0$.$p_{v_i}^2 - h_i^2 \in I \; \forall \; 1 \leq i \leq r$.Therefore $L(g_j h_i^2) = L(g_j p_{v_i}^2) \geq 0 \; \forall \; 1 \leq i \leq r$. and $\forall \; 1 \leq j \leq m$.Therefore $L(g_j p_{v_i}^2) = \sum_{i=1}^r L(p_{v_i}) L_{v_i}(g_j p_{v_i}^2) = g_j(v_i) \geq 0 \; \forall \; 1 \leq i \leq r$ and $\forall \; 1 \leq j \leq m$.So $v_i \in K \; \forall \; 1 \leq i \leq r$.Therefore supp($\mu$) $\subseteq$ K. ☐

Theorem 3.3.0.12. *Let $L \in \mathbb{R}[x_1, ..., x_n]^*_{2t}$ be an optimal solution of (MOMt).Assume L satisfies rank $M_t(L) =$ rank M_{t-d_K}. Then $f_t^{mom} = f_{min}$*

Proof. L is optimal solution of (MOMt).Therefore $L(p) \geq 0 \; \forall \; p \in Q_t(g)$.So $M_t(L) \succeq 0$ and $M_{t-\lceil d_j/2 \rceil} \succeq 0 \; \forall \; 1 \leq j \leq m$.rank $M_t(L) =$ rank M_{t-d_K}.So by previous theorem L has a representing measure μ with supp$(\mu) \subseteq K$.Let supp$(\mu) = \{v_1, ..., v_r\} \subseteq \mathbb{R}^n$.Then L $= \sum_{i=1}^r L(p_{v_i}) L_{v_i}$.Therefore $f_t^{mom} = L(f) = \sum_{i=1}^r L(p_{v_i}) f(v_i) (v_i \in$ supp$(\mu) \subseteq K \; \forall \; 1 \leq i \leq r$ and $L(p_{v_i}) = L(p_{v_i}^2) \geq 0 \implies f(v_i) \geq f_{min} \; \forall \; 1 \leq i \leq r)$.So $f_t^{mom} \geq (\sum_{i=1}^r L(p_{v_i})) f_{min}$..$L(1) = 1$ .So $\sum_{i=1}^r L(p_{v_i}) = 1$.$\therefore$ $f_t^{mom} \geq f_{min}$.Therefore $f_t^{mom} = f_{min}$. □

Chapter 4

Kissing Number

4.1 Spherical Harmonics

Here we go through concepts of spherical harmonics, needed to derive the addition theorem. This chapter contains description of gegenbauer polynomials . It discusses the ' kissing number problem ' and how to give a bound for it.

Laplace equation in n variables.

$$\sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = 0 \tag{4.1}$$

Definition 4.1.0.13. Harmonic polynomial: Homogeneous polynomials which satisfy Laplace equation are called harmonic.

Definition 4.1.0.14. Spherical harmonic: Spherical harmonic in n variables is restriction of homogeneous polynomial (say u) in n variables satisfying the Laplace equation to unit sphere S^{n-1} in $\mathbb{R}^n$

Definition 4.1.0.15. $V_{k,n}$: Homogeneous polynomials of deg k in n variables form a vector place over $\mathbb{R}$.We denote this space by $V_{k,n}$.

Now to calculate the dimension of $V_{k,n}$.

Lemma 4.1.0.16. *dim* $V_{k,n} = \binom{n-1+k}{n-1} = \binom{n-1+k}{k}$

Proof. Let $\mathbf{x} = (x_1, .., x_n)$. $(\mathbf{x}^{\alpha})_{|\alpha|=k} span V_{k,n}$. And they are linearly independent. $dim V_{k,n} = |S|$ where $S = \{\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{Z}^n_{\geq 0} \mid |\alpha| = \alpha_1 + ... + \alpha_n = k\}$. Now consider any arrangement of n-1 lines and k dots .Each arrangement gives rise to a distinct α. And each α gives us a unique arrangement. Therefore $|S|$ is the number of arrangements of n-1 lines and k dots i.e $\binom{n-1+k}{n-1} = \binom{n-1+k}{k}$ □

Denote dim of $V_{k,n}$ by $d_{k,n}$. So $d_{k,n} = \binom{n-1+k}{n-1}$. All homogeneous polynomials are not harmonic.Let $\Delta = \sum \partial^2/\partial x_i^2$ be Laplace operator. Δ is a linear operator. Harmonic polynomials of deg k in n- variables also form a vector space over $\mathbb{R}$ Let us denote it by $W_{k,n}$. Let p(x) be a homogeneous polynomial of deg k in n variables. We can write p(x) as

$$p(x) = \sum_{j=0}^{k} A_{k-j}(x_1, ..., x_{n-1}) x_n^j \tag{4.2}$$

where $A_m(x_1, .., x_{n-1})$ is homogeneous polynomial of degree m in n-1 variables.p(x) harmonic implies $\Delta p(x) = 0$. So ,

$$\begin{aligned}
\Delta p(x) &= \sum_{j=0}^{k} \Delta(A_{k-j}(x_1, ..., x_{n-1})x_n^j) \\
&= \sum_{j=0}^{k}(\sum_{i=1}^{n-1}(\partial^2(A_{k-j}(x_1, ..., x_{n-1})x_n^j)/\partial x_i^2) + (\partial^2(A_{k-j}(x_1, ..., x_{n-1})x_n^j)/\partial x_n^2)) \\
&= \sum_{j=0}^{k-2} \Delta(A_{k-j}(x_1, ..., x_{n-1}))x_n^j + \sum_{j=2}^{k} j(j-1)(A_{k-j}(x_1, ..., x_{n-1}))x_n^{j-2} \\
\therefore 0 &= \sum_{j=0}^{k-2}(A_{k-j}(x_1, ..., x_{n-1}) + (j+2)(j+1)A_{k-j-2}(x_1, ..., x_{n-1}))x_n^j
\end{aligned}$$

So

$$\Delta A_{k-j}(x_1, .., x_n) = -(j+2)(j+1)A_{k-j-2}(x_1, .., x_n) \ \forall \ 0 \leq j \leq k-2 \tag{4.3}$$

If we have A_k and A_{k-1} then we can compute $A_j \ \forall \ 0 \leq j \leq k-2$.So we have p.We can define $\phi : V_{k,n-1} \times V_{k-1,n-1} \to W_{k,n}$ by $\phi(A_k, A_k - 1) =$ the corresponding harmonic polynomial computed using (3.3).ϕ is linear and bijective.So by null rank theorem dim$W_{k,n} =$ dim $(V_{k,n-1} \times V_{k-1,n-1}) = d_{k,n-1} + d_{k-1,n-1}$.Therefore dim $W_{k,n} = \binom{n-2+k}{k} + \binom{n+k-3}{k-1}$.

Polar coordinates in n-dimensions$(r, \theta_1, ..., \theta_{n-2}, \phi)$

$$\begin{aligned}
x_1 &= rcos(\theta_1) \quad (4.4)\\
x_2 &= rsin(\theta_1)cos(\theta_2) \quad (4.5)\\
&\vdots \quad (4.6)\\
x_{n-1} &= rsin(\theta_1)sin(\theta_2)\ldots sin(\theta_{n-2})cos(\phi) \quad (4.7)\\
x_n &= rsin(\theta_1)sin(\theta_2)\ldots sin(\theta_{n-2})sin(\phi) \quad (4.8)
\end{aligned}$$

with $0 \leq \theta_i \leq \pi$ and $0 \leq \phi \leq 2\pi$.
We can define inner product on the space of real continuous functions on S^{n-1} by

Definition 4.1.0.17. $\langle f, g\rangle$ **:**

$$\langle f, g\rangle = \int_{S^{n-1}} f(\xi)g(\xi)dw(\xi) \tag{4.9}$$

Let $H_k(x_1, ..., x_n)$ be harmonic homogeneous polynomial of degree k and $H_j(x_1, ..., x_n)$ be harmonic homogeneous polynomial of degree j.

Theorem 4.1.0.18. *Harmonic homogeneous polynomials of different degrees are orthogonal.*

Proof. H_k, H_j harmonic implies $\Delta H_k = 0 = \Delta H_j$.So

$$\int_{x_1^2+....+x_n^2\leq 1}(H_j(x_1, ..., x_n)\Delta H_k(x_1, ..., x_n) - H_k(x_1, ..., x_n)\Delta H_j(x_1, ..., x_n))dx_1dx_2....dx_n = 0$$

By Gauss-Green's theorem, we get

$$LHS = \int_{|\xi|=1} (H_j(\xi)\partial/\partial r H_k(r\xi)|_{r=1} - H_k(\xi)\partial/\partial r H_j(r\xi)|_{r=1}) dw(\xi)$$

$H_k, H_j$are homogeneous of deg k ,deg j respectively.Therefore

$$\partial/\partial r H_k(r\xi)|_{r=1} = \partial/\partial r(r^k H_k(\xi))|_{r=1} = kr^{k-1}H_k(\xi)|_{r=1} = kH_k(\xi)$$
$$\partial/\partial r H_j(r\xi)|_{r=1} = \partial/\partial r(r^j H_j(\xi))|_{r=1} = jr^{j-1}H_j(\xi)|_{r=1} = jH_j(\xi)$$

So LHS= $\int_{S^{n-1}}(k-j)H_j(\xi)H_k(\xi)dw(\xi) = 0$. $dw(\xi)$ is invariant measure on surface of S^{n-1}. So if $k \neq j$ then $\langle H_j(\xi), H_k(\xi)\rangle = 0$.Therefore homogeneous harmonic polynomials of different degrees are orthogonal. □

Denote dim $W_{k,n}$ by $c_{k,n} = \binom{n+k-2}{k} + \binom{n+k-3}{k-1}$.

With respect to the above inner product we can use Grahm-Schmidt orthogonalization to obtain an orthonormal basis of $W_{k,n}$.Let $S_{k,j}$ for j= $1, ..., c_{k,n}$ be the orthonormal basis thus obtained.

Let O be an orthogonal $n \times n$ matrix i.e

$$O^T O = Id = OO^T.$$

Then we have a map from $\mathbb{R}^n$ to $\mathbb{R}^n$ given by $x \mapsto Ox$. Then scalar product

$$\begin{aligned}(Ox, Oy) &= (Ox)^T(Ox)\\ &= x^T O^T Oy\\ &= x^T y\\ &= (x, y)\end{aligned}$$

Lemma 4.1.0.19. $S_{k,j}(Ox) \in W_{k,n}$

Proof. Consider $\Delta(S_{k,j}(Ox)))$,

$$\Delta(S_{k,j}(Ox))) = (\partial/\partial x_1, ..., \partial/\partial x_n) \begin{bmatrix}\partial/\partial x_1\\ \vdots\\ \partial/\partial x_n\end{bmatrix} S_{k,j}(Ox)$$

Let $y = Ox$ then $(\partial/\partial x_1, ..., \partial/\partial x_n) = (\partial/\partial y_1, ..., \partial/\partial y_n)O$. Therefore $\Delta_{with\ respect\ to\ x_1,...,x_n}(S_{k,j}(Ox)) = \Delta_{with\ respect\ to\ O(x_1,...,x_n)}(S_{k,j}(Ox)) = 0$ □

So each $S_{k,j}(Ox)$ can be written in terms of $S_{k,j}(x)$; $j = 1, ..., c_{k,n}$ uniquely.Let $S_{k,j}(Ox) = \sum_{l=1}^{c_{k,n}} A^k_{jl} S_{k,l}(x)$ for $j = 1, ..., c_{k,n}$.Let $A^k = (A^k_{j,l})_{jl}$ be a $c_{k,n} \times c_{k,n}$ matrix.Then

$$\begin{aligned}((A^k)^T A^k)_{jl} &= \sum_{i=1}^{c_{k,n}} A^k_{ij} A^k_{il}\\ &= \int_{S^{n-1}} S_{k,j}(Ox)S_{k,l}(Ox)dw(x)\\ &= \int_{S^{n-1}} S_{k,j}(x)S_{k,l}(x)dw(x) = \delta_{jl}\ (asdwisinvariant.).\end{aligned}$$

Therefore A^k is orthogonal.

Let $\eta \in S^{n-1}$.Then η can be expressed in such form: $\eta = t(1,0,...,0) + \sqrt{1-t^2}\eta'$ where $|\eta'| = 1$ and η' is of the form $(0,*,*,...,*).(\eta,(1,0,..,0)) = t = cos\theta_1$ where $\eta = (1,\theta_1,...,\theta_{n-2},\phi)$ in polar coordinates.By relations given by (*) we get
$dx_1...dx_n = r^{n-1}sin^{n-2}\theta_1 \ldots \sin^2\theta_{n-3}sin\theta_{n-2}drd\theta_1...d\theta_{n-2}d\phi$
So $dw_n = rsin^{n-2}\theta_1 d\theta_1 dw_{n-1}$.Therefore on S^{n-1} i.e when $r=1$ we have

$$dw_n = sin^{n-2}\theta_1 d\theta_1 dw_{n-1} = (\sqrt{1-t^2}^{n-2}/\sqrt{1-t^2})dtdw_{n-1} = \sqrt{1-t^2}^{n-3}dtdw_{n-1} \tag{4.10}$$

Consider

$$F_k(x,\tilde{\eta}) = \left[S_{k,1}(x),\ldots,S_{k,c_{k,n}}(x)\right]\begin{bmatrix}S_{k,1}(\tilde{\eta})\\ \vdots \\ S_{k,c_{k,n}}(\tilde{\eta})\end{bmatrix}. \tag{4.11}$$

If we consider F_k as a function of x then it is a homogeneous polynomial of degree k. And it will be harmonic as each $S_{k,j}(x)$ is harmonic for $j = 1,..,c_{k,n}$. Let O be an orthogonal matrix which fixes η. Then

$$F_k(Ox,O\eta) = \left[S_{k,1}(Ox),\ldots,S_{k,c_{k,n}}(Ox)\right]\begin{bmatrix}S_{k,1}(O\tilde{\eta})\\ \vdots \\ S_{k,c_{k,n}}()\tilde{\eta})\end{bmatrix}$$

$$So\ F_k(Ox,\eta) = \left[S_{k,1}(Ox),\ldots,S_{k,c_{k,n}}(Ox)\right](A^K)^T A^K\begin{bmatrix}S_{k,1}(O\tilde{\eta})\\ \vdots \\ S_{k,c_{k,n}}()\tilde{\eta})\end{bmatrix}\ldots\ldots\ldots\ldots(Because\ O\eta = \eta)$$

So $F_k(Ox,\eta) = F_k(x,\eta)$ ....(Because A^K is orthogonal)

Therefore F_k as a function of x is invariant under all orthogonal transformations that fix η. F_k only depends on scalar product $(x,\tilde{\eta})$. So we can write $F_k(x,\tilde{\eta}) = b_k P_k((x,\tilde{\eta}))$.....(***) for some constant b_k. We can normalize it by taking $P_k((\eta,\eta)) = P_k(1) = 1$.
So $\sum_{j=1}^{c_{k,n}}(S_{k,j}(\eta))^2 = F_k(\eta,\eta) = b_k$. Therefore $\int_{S^{n-1}} b_k dw(\eta) = \sum_{j=1}\int_{S^{n-1}}(S_{k,j}(\eta))^2 dw(\eta)$.$\therefore$ $b_k w_n = \int_{S^{n-1}} b_k dw(\eta) = c_{k,n}$......($\because$ $S_{k,j} : j = 1,...,c_{k,n}$ is an orthonormal basis for $W_{k,n}$ with respect to(3.4) ......(II)where w_n is surface area of S^{n-1}. So $b_k = c_{k,n}/w_n$.

$F_k(x,\eta) = \sum_{j=1}^{c_{k,n}} S_{k,j}(x)S_{k,j}(\eta) = c_{k,n}/w_n P_k((x,\eta))$.

$$c_{k,n}^2/w_n^2 P_k^2((x,\eta)) = \sum_{j=1}^{c_{k,n}}\sum_{i=1}^{c_{k,n}} S_{k,j}(x)S_{k,j}(\eta)S_{k,i}(x)S_{k,i}(\eta).$$

$$\begin{aligned}
c_{k,n}^2/w_n^2\int_{S^{n-1}} P_k^2((x,\eta))dw(\eta) &= \sum_{j=1}^{c_{k,n}}\sum_{i=1}^{c_{k,n}} S_{k,j}(x)S_{k,i}(x)\int_{S^{n-1}} S_{k,j}(\eta)S_{k,i}(\eta)dw(\eta)\\
&= \sum_{j=1}^{c_{k,n}}(S_{k,j}(x))^2.....(by(II))\\
&= F_k(x,x) = b_k P_k((x,x))\\
&= c_{k,n}/w_n P_k(1) = c_{k,n}/w_n
\end{aligned}$$

So $\int_{S^{n-1}} P_k^2((x,\eta))dw(\eta) = w_n/c_{k,n}$

$F_k(x,\eta) = b_k P_k((x,\eta)) = \sum_{j=1}^{c_{k,n}} S_{k,j}(x)S_{k,j}(\eta)$

$F_k'(x,\eta) = b_k' P_k'((x,\eta)) = \sum_{j=1}^{c_{k',n}} S_{k',j}(x)S_{k',j}(\eta)$

Therefore $b_k b_k' \int_{S^{n-1}} P_k((x,\eta))P_k'((x,\eta))dw(\eta) = \sum_{j=1}^{c_{k,n}} \sum_{i=1}^{c_{k',n}} S_{k,j}(x)S_{k',i}(x) \int_{S^{n-1}} S_{k,j}(\eta)S_{k',i}(\eta)dw(\eta)$ 0 (if $k \neq k'$) ......(by (3.5))

$$\int_{S^{n-1}} P_k((x,\eta))P_k'((x,\eta))dw(\eta) = (w_n/c_{k,n})\delta_{k,k'} \tag{4.12}$$

Now let x $= (1,0,..,0)$..Let $t = (x,\eta)$.Therefore using (3.6) we get $\int_{S^{n-1}} P_k((x,\eta))P_k'((x,\eta)) = \int_{-1}^{1} P_k(t)P_k'(t)(\sqrt{1-t^2})^{n-3}dt \int dw_{n-1}$....(because as η varies on S^{n-1},t varies from -1 to 1)
$= w_{n-1}\int_{-1}^{1} P_k(t)P_k'(t)(\sqrt{1-t^2})^{n-3}dt$

$$(w_n/(w_{n-1}c_{k,n}))\delta_{k,k'} = \int_{-1}^{1} P_k(t)P_k'(t)(\sqrt{1-t^2})^{n-3}dt \tag{4.13}$$

4.2 Gegenbauer Polynomials

Ultraspherical polynomials are defined in terms of their generating function.For a given α, C_n^α are coefficients of t^n in $1/(1-2xt+t^2)^\alpha$. i.e

$$\sum_{n=0}^{\infty} C_n^\alpha(x)t^n = 1/(1-2xt+t^2)^\alpha$$

The above equation (3.8) implies $P_k(t) = C.C_k^{n-2/2}(t)$ for some constant C. (For this implication refer to [4] or [5]). But we have normalized so that $P_k(1) = 1$. Therefore $C = 1/C_k^{n-2/2}$.. Therefore $P_k(t) = C_k^{n-2/2}(t)/C_k^{n-2/2}(1)$

Definition 4.2.0.20. Gegenbauer Polynomial (of deg k in n variables) :
$G_k^n(t) := C_k^{n-2/2}(t)/C_k^{n-2/2}(1)$

So $P_k(t) = G_k^n(t).G_0^n(t) = P_0(t)$. But $P_0(t)$ is constant as it is homogeneous polynomial of deg 0.But $P_0(1) = 1$. So $G_0^n(t) = 1\ \forall\ n$.

By (***)we see that $F_k(x,\tilde{\eta}) = b_k P_k((x,\tilde{\eta}))$
We know $b_k = C_{k,n}/w_n$

Therefore $F_k(x,\tilde{\eta}) = \sum_{j=1}^{C_{k,n}} S_{k,j}(x)S_{k,j}(\tilde{\eta}) = C_{k,n}/w_n G_k^n((x,\tilde{\eta}))$.
So we have proved the addition theorem.

Theorem 4.2.0.21. ***Addition Theorem :***

$$G_k^n((x,\tilde{\eta})) = w_n/C_{k,n}\sum_{j=1}^{C_{k,n}} S_{k,j}(x)S_{k,j}(\tilde{\eta}) \tag{4.14}$$

Theorem 4.2.0.22. ***(Schoenberg):****If $X = (X_{ij})_{i,j}$ is a $N \times N$ matrix on $\mathbb{R}$ s.t $X \succeq 0$ with rank atmost n, then $G_k^n(X) \succeq 0$.*

Proof. $X \succeq 0$ implies $X = LL^t$. $rank(X) = rank(L) \leq n$.Therefore we can choose L s.t L is an N× n matrics. Therefore each row is a vector in $\mathbb{R}^n$. Let i^{th} row be r_i
Then

$$X = [\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{pmatrix}](r_1 \quad r_2 \quad ... \quad r_N)$$

Therefore $X_{ij} = (r_i, r_j) \forall 1 \leq i, j \leq N$. So $G_k^n(X_{ij}) = G_k^n((r_i, r_j))....(As\, r_i, r_j \in \mathbb{R}^n\ well\ defined)$

Addition theorem implies

$$G_k^n(X_{ij}) = w_n/C_{k,n} \sum_{l=1}^{C_{k,n}} S_{k,l}(r_i) S_{k,l}(r_j) \quad \forall 1 \leq i, j \leq N$$

Therefore matrix $G_k^n(X) = w_n/C_{k,n} MM^T\ where\ M = (S_{k,j}(r_i))_{i,j}$ So $G_k^n(X) \succeq 0$. □

4.3 Kissing Number

Given a sphere A in dim n. Kissing number k(n) is the maximum number of spheres of same size as A that can touch A simultaneously without overlapping each other.
For the case n =2,it is easy to show that k(2)=6 using the diagram below.

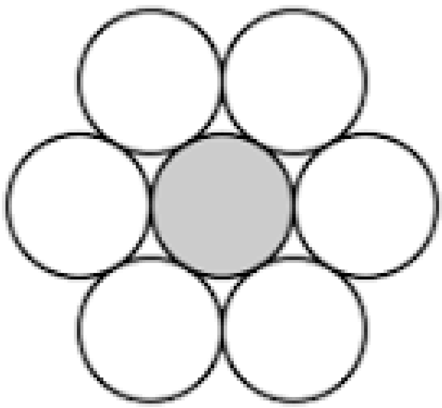

Figure 4.1: Kissing Number for $n = 2$

There is no space for a 7^{th} circle can be shown using contradiction. Kissing numbers are known for n=1,2,3,4,8 and 24.For n=3 regular icosahedron gives us a configuration where 12 spheres touch given sphere without overlapping.(Icosahedron has 20 faces and 12 vertices such that five faces meet at each vertex.If we consider spheres touching the sphere at these 12 vertices we get the desired configuration).But a lot of space is left even after placing 12 spheres touching the center one.Therefore its hard to know if the above is the unique configuration.

For n=4 24-cell provides a configuration.Therefore k(4)≥ 24. Musin proved in 2003 that k(4) =24 . For n=8 root latice E_8 provides a configuration.It is known that k(8)=240.Leech latice gives a configuration for n=24. k(24) is known to be 196560.
Delsarte, Goethals and Seidel method can be used to find good bounds on the kissing numbers.

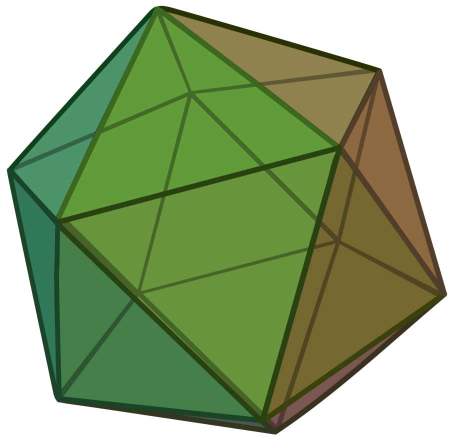

Figure 4.2: Icosahedron

Theorem 4.3.0.23. ***Delsarte, Goethals and Seidel***: *If* $f(t)=\sum_{k=0}^{d} c_k G_k^n(t)$ *where* $G_k^n(t)$ *are GegenBauer Polynomial with* $c_0 > 0$ *and* $c_k \geq 0 \ \forall \ k = 1...d$ *and* $f(t) \leq 0 \ \forall \ t \in [-1, 1/2]$ *then* $k(n) \leqslant f(1)/c_0$

Consider the following : Let A be unit sphere centered at origin in $\mathbb{R}^n$. Lets say its possible that N unit spheres touch A without overlapping .Let $x_1, ..., x_N$ denote the points where they touch A. No two spheres overlap. Therefore $\langle x_{i,i} \rangle = 1 \ \forall \ 1 \leq i \leq N$. So $\langle x_i, x_j \rangle \leq 1/2 \ \ \forall \ i \neq j$ (Because θ between lines joining centers should be $\geq$ 60 if they don't overlap. Therefore $cos\theta \leq cos60 = 1/2$. So $\langle x_i, x_j \rangle \leq 1/2$)

Consider the matrix X such that

$$X_{ij} = \langle x_i, x_j \rangle \therefore X = [\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}](x_1 \quad x_2 \quad ... \quad x_N)$$

Therefore X is positive semidefinite. And as $x_i \in \mathbb{R}^n \ \ \forall \ \ 1 \leq i \leq n \ \ rank \ \ x \leq n.$
Consider the set

$$S = \{X \in S^N | X \succeq 0, x_{ii} = 1 \ \ \forall \ \ 1 \leq i \leq N, x_{ij} \leq 1/2 \ \ \forall \ i \neq j \ \ rank \ x \quad \leq n\} \tag{4.15}$$

If $S \neq \phi$ then $\exists \ \ x \in S^N s.t. X \succeq 0, x_{ii} = 1 \ \& \ x_{ij} \leq 1/2$ and rank $x \subseteq n$.
So $X = LL^T$ for some L a $N \times n \ matrix$.
So we can consider rows of L as points on A and so we get a configuration for N points.

Returning to the proof

Proof. According to the explanation above k(n) corresponds to a matrix $X \succeq 0$ with $x_{ii} = 1 \forall 1 \leqslant i \leqslant k(n)$ and $x_{ij} \leqslant 1/2 \ \forall i \neq j$ and rank x $\leq n$.So applying Schoenberg's theorem we get $G_k^n(X) \succeq 0$.Therefore sum of all entries of $G_k^n(X) \geq 0$.

$$\sum_{j=1}^{k(n)} \sum_{i=1}^{k(n)} G_k^n(X_{ij}) \geq 0 \tag{4.16}$$

$$\sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)} f(X_{ij}) = \sum_{i=1}^{k(n)} f(X_{ii}) + \sum_{i\neq j} f(X_{ij})$$
$$= k(n)f(1) + \sum_{i\neq j} f(X_{ij})$$

$X_{ij} \leq 1/2$ if $i \neq j$ and $f \leq 0$ on [-1,1/2]. Therefore $\sum_{i\neq j} f(X_{ij}) \leq 0$.So,

$$\sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)} f(X_{ij}) \leq k(n)f(1) \tag{4.17}$$

Now calculating again we get

$$\sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)} f(X_{ij}) = \sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)}\sum_{k=0}^{d} c_k G_k^n(X_{ij})$$
$$= \sum_{k=0}^{d} c_k \sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)} G_k^n(X_{ij})$$
$$\geq c_0 \sum_{j=1}^{k(n)}\sum_{i=1}^{k(n)} G_0^n(X_{ij}) ...(because\ we\ have (1.2)\ and\ c_k \geq 0)$$

We have $G_0^n(X_{ij}) = 1\ \forall\ i, j$. Therefore

$$\sum_{i-1}^{k(n)}\sum_{j=1}^{k(n)} f(X_{ij}) \geq c_0(k(n))^2 \tag{4.18}$$

So $k(n)f(1) \geq c_0(k(n))^2$.So

$$k(n) \leq f(1)/c_0 \tag{4.19}$$

□

Now lets try to find bound on k(n) using Delsarte's method and semidefinite optimization.Consider the following program for some fixed D.

$$min_n = min_F\{F(1)|F = \sum_{k=0}^{D} \lambda_k G_k^n, \lambda_k \geq 0, \lambda_0 = 1\ and\ F(t) \leq 0\ \forall\ t\ \in\ [-1, 1/2]\} \tag{4.20}$$

Then by Delsarte's theorem $F(1)/\lambda_0 = F(1) \geq k(n)\ \forall\ F$ satisfying the condition.Therefore $min_n \geq k(n)$

As we increase D we will get better bounds.Now we have to convert this problem to a semidefinite optimization problem.

$F(t) \leq 0\ \forall\ t\ \in\ [-1, 1/2] \iff -F(t) \geq 0$ on [-1,1/2].[-1,1/2] can be reformulated as $K = \{t\ \in\ \mathbb{R}|g_1(t) := (1/2 - t)(t + 1) \geq 0\}$.If we can show that archimedean condition holds for $Q(g)$ where $g = (g_1)$.Then by Putinar's theorem we get that $-F(t) > 0$ on [-1,1/2] $\iff -F\ \in\ Q(g)$.Then we can replace the program by

$$min_n = min_F\{F(1)|F = \sum_{k=0}^{D} \lambda_k G_k^n, \lambda_k \geq 0, \lambda_0 = 1\ and\ -F\ \in\ Q(g)\} \tag{4.21}$$

In chapter 2 section 2 we saw that the condition $-F\ \in\ Q_r(g)$ for some r can be replaced by a semidefinite program.By varying r and D and using semidefinite optimization we can get bounds on k(n).Refer appendix for the program.

Chapter 5

Triangle Packing

Problem Statement : What is the maximum number of regular tetrahedron that we can pack in unit sphere S^2 having a common vertex origin so that none of them overlap?

This problem corresponds to finding the maximum number of equilateral spherical triangles with edge length $\pi/3$ that cover the sphere without overlapping.It is known that this number T(3) satisfies $20 \leq T(3) \leq 22$. The upperbound can be found by dividing the surface area of sphere by area of a spherical equilateral triangle of edge $\pi/3$.Surface area of sphere is 4π and area of spherical triangle can be calculated using Girard's theorem.Icosahedron gives us a configuration for packing 20 tetrahedrons in a sphere.($\therefore T(3) \geq$20)As we reformulated 'kissing number problem' in terms of points on the sphere,we try to reformulate this problem. We can denote vertices of spherical triangle as (x_1, x_2, x_3) with certain conditions so that they form equilateral spherical triangle.So we have to find maximum number of triples (x_1, x_2, x_3) on sphere such that they form equilateral triangle and no two overlap.So we need to find condition in terms of (x_1, x_2, x_3) and (y_1, y_2, y_3) that will imply that the 2 triangles donot overlap. So we need to find conditions depending on scalar products (x_i, y_j) such that the 2 triangles donot overlap.

Definition 5.0.0.24. Let Ω be the set of nine tuples $(a_{11}, a_{12}, a_{13}, ..., a_{33})$ such that if

$$a_{ij} = (x_i, y_j) \ \forall \ i, j \ for \ 2 \ triangles \ (x_1, x_2, x_3) \ and \ (y_1, y_2, y_3)$$

then (x_1, x_2, x_3) and (y_1, y_2, y_3) donot overlap.

Let us say we can arrange N non overlapping tetrahedrons with common vertex in $\mathbb{R}^3$. Let origin be the common vertex. Let $T_1, ..., T_N$ be the corresponding spherical equilateral triangles on S^2, where $T_i = \{x^i_1, x^i_2, x^i_3\} \ \forall \ 1 \leq i \leq N$. So

$$(x^i_j, x^i_j) = 1 \ \forall \ 1 \leq i \leq N \ and \ \forall \ 1 \leq j \leq 3 \quad (5.1)$$
$$(x^i_j, x^i_{j'}) = 1/2 \ \forall \ 1 \leq i \leq N \ and \ j \neq j'. \quad (5.2)$$

Definition 5.0.0.25. Denote

$$X^{k,l}_{i,j} = (x^k_i, x^l_j)$$
$$X = (A_{kl})_{k,l}$$
$$where \ A_{kl} \ is \ itself \ a \ matrix \ given \ by$$
$$A_{kl} = (X^{k,l}_{i,j})_{i,j}.$$

One can check that $\forall\ 1 \leq k \leq N$

$$A_{kk} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

By definition of matrix X it is evident that $X \succeq 0$. As each $x_j^i \in \mathbb{R}^3$,rank X ≤ 3. For any A_{kl} , let a_{kl} be point $(X_{11}^{kl}, X_{12}^{kl}, ..., X_{33}^{kl})$ in $\mathbb{R}^9$.

Definition 5.0.0.26.

$$P_k^n(z) = P_k^n(z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9) := \sum_{i=1}^{9} G_k^n(z_i) \tag{5.3}$$

where $z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9)$

Theorem 5.0.0.27. *If* $P = \sum_{s=0}^{D} f_s P_s^3$ *with* $f_0 = 1$ *and* $f_s \geq 0$ $\forall 1 \leq s \leq D$ *and if P(z)* $\leq 0\ \forall\ z \in \Omega$*, then*

$$N \leq \frac{P(a_{ll})}{9} \tag{5.4}$$

(N and A_{ll} *are as mentioned above)*

Proof. X $\succeq 0$ implies $G_s^3(X) \succeq 0$ by Schoenberg's theorem for all $s = 0, ..., D$. So sum of entries is greater than or equal to 0. So

$$\sum_{i,j,k,l} G_s^3(X_{i,j}^{k,l}) \geq 0 \tag{5.5}$$

Consider

$$\begin{aligned} \sum_{k,l} P(a_{kl}) &= \sum_{k,l} \sum_{s=0}^{D} f_s P_s^3(a_{kl}) \\ &= \sum_{k,l} \sum_{s=0}^{D} f_s \sum_{i,j} G_s^3(X_{i,j}^{k,l}) \\ &= \sum_{s=0}^{D} f_s \sum_{i,j,k,l} G_s^3(X_{i,j}^{k,l}) \\ &\geq \sum_{i,j,k,l} G_0^3(X_{i,j}^{k,l}) \end{aligned}$$

Because $f_k \geq 0$ and (5.12). Therefore

$$\sum_{k,l} P(a_{kl}) \geq 9N^2 \tag{5.6}$$

Again calculating $\sum_{k,l} P(a_{kl})$ we get,

$$\begin{aligned} \sum_{k,l} P(a_{kl}) &= NP(a_{ll}) + \sum_{k \neq l} P(a_{kl}) \\ &\leq NP(a_{ll}) \end{aligned}$$

Because P is less than or equal to 0 on Ω. So we get

$$\sum_{k,l} P(a_{kl}) \leq NP(a_{ll}) \tag{5.7}$$

So (5.13) and (5.14) imply

$$N \leq \frac{P(a_{ll})}{9} \tag{5.8}$$

□

Now to express Ω in terms of algebraic inequalities.

Let $\{x, y, z\}$ be a spherical equilateral triangle of edge length $\pi/3$ on S^2. These 3 points lie on a plane. Centre of this planar triangle is given by (x+y+z)/3 . Centre of the spherical triangle $\{x, y, z\}$ lies on the line through origin and (x+y+z)/3 and on S^2. So spherical centre is given by

$$C_{x,y,z} = k(x + y + z)/3$$

For some scalar k. But $C_{x,y,z}$ lies on S^2 so

$$k^2((x + y + z),(x + y + z))/9 = 1$$
$$k^2(\frac{2}{3}) = 1$$

Because $(x, x) = 1 = (y, y) = (z, z)$ and $(x, y) = (y, z) = (x, z)$. So $k = \sqrt{\frac{3}{2}}$. So

$$C_{x,y,z} = \frac{(x + y + z)}{\sqrt{6}} \tag{5.9}$$

Let X $= \{x_1, x_2, x_3\}, Y = \{y_1, y_2, y_3\}$ be 2 spherical equilateral triangles of edge length $\pi/3$ on S^2. Let r be the angular distance between C_{x_1,x_2,x_3}, x_1. Denote it by $\theta(C_{x_1,x_2,x_3}, x_1)$

Proposition 5.0.0.28. *If $\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) \geq 2r$, then spherical triangles X and Y donot overlap.(Note that $\theta(C_{y_1,y_2,y_3}, y_1)$= $\theta(C_{x_1,x_2,x_3}, x_1) = \theta(C_{x_1,x_2,x_3}, x_2) = \theta(C_{x_1,x_2,x_3}, x_1)$. Same for Y.)*

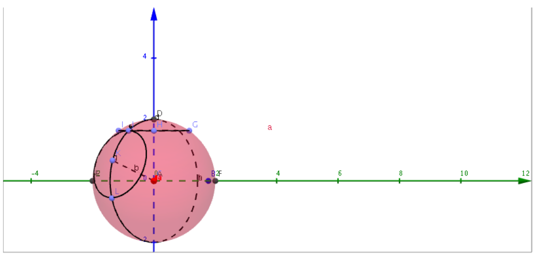

Figure 5.1: spherical caps

Lemma 5.0.0.29. *The condition that $\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) \geq 2r$ can be rewritten in terms of scalar product , which is given by*

$$\sum_{i,j\ \in\{1,2,3\}} (x_i, y_j) \leq 2 \tag{5.10}$$

Proof. $\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) \geq 2r$ implies

$$\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) \geq 2\theta(C_{x_1,x_2,x_3}, x_1) \tag{5.11}$$

So

$$cos(\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3})) \leq cos(2\theta(C_{x_1,x_2,x_3}, x_1))$$
$$cos(\theta(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3})) \leq 2cos(\theta(C_{x_1,x_2,x_3}, x_1))^2 - 1$$

As $C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}, C_{x_1,x_2,x_3}, x_1$ lie on S^2, we have

$$(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) \leq 2(C_{x_1,x_2,x_3}, x_1)^2 - 1$$

But by (5.9)

$$(C_{x_1,x_2,x_3}, C_{y_1,y_2,y_3}) = \frac{\sum_{i,j\ \in\{1,2,3\}}(x_i, y_j)}{6}$$
$$(C_{x_1,x_2,x_3}, x_1) = \frac{2}{\sqrt{6}}$$

So,

$$\sum_{i,j\ \in\{1,2,3\}} (x_i, y_j) \leq 2$$

□

So we know that X and Y donot overlap if (5.6) is satisfied.

Now we have to consider the case when $\sum_{i,j\ \in\{1,2,3\}}(x_i, y_j) > 2$.

For any u $\neq v$ points on S^2 forming an edge of a spherical equilateral triangle with edge $\pi/3$, $\{u, v, u \wedge v\}$ form a basis for $\mathbb{R}^3$. ($(u \wedge v)$ is point on S^2 such that ,that vector is normal to plane spanned by u and v) . So for any point h in $\mathbb{R}^3$ we can write

$$h = au + bv + c(u \wedge v) \tag{5.12}$$

So as $(u, u) = 1 = (v, v)$ and $((u \wedge v), u) = 0 = ((u \wedge v), v)$

$$(h, u) = a(u, u) + b(v, u) + c((u \wedge v), u) = a + \frac{b}{2}$$
$$(h, v) = a(u, v) + b(v, v) + c((u \wedge v), v) = \frac{a}{2} + b$$

Solving these equations we get

$$a = \frac{4}{3}(h, u) - \frac{2}{3}(h, v) \tag{5.13}$$
$$b = \frac{4}{3}(h, v) - \frac{2}{3}(h, u) \tag{5.14}$$

Now let h_1, h_2 be 2 points in $\mathbb{R}^3$. So again we can write

$$h_1 = a_1u + b_1v + c_1(u \wedge v)$$
$$h_2 = a_2u + b_2v + c_2(u \wedge v)$$

So

$$\begin{aligned}(h_1,h_2) &= ((h_1 = a_1u + b_1v + c_1(u\wedge v)),(h_2 = a_2u + b_2v + c_2(u\wedge v)))\\ &= ((a_1u+b_1v),(a_2u+b_2v)) + c_1c_2\\ &= (a_1a_2 + b_1b_2 + \frac{a_1b_2 + a_2b_1}{2}) + c_1c_2\end{aligned}$$

$(h_1, u\wedge v) = c_1$ and $(h_2, u\wedge v) = c_2$. So h_1, h_2 belong to different halfspaces created by plane spanned by u,v iff $c_1c_2 \leq 0$. i.e iff

$$(h_1,h_2) - (a_1a_2 + b_1b_2 + \frac{a_1b_2 + a_2b_1}{2}) \leq 0 \tag{5.15}$$

Definition 5.0.0.30. $\bar{H}(x,y;z)$**:** Vector x and vector y span a plane. This plane creates 2 halfspaces. $\bar{H}(x,y;z)$ is the halfspace different from halfspace containing z.

$$X \subset \bar{H}(y_1,y_2;y_3) \tag{5.16}$$
$$\iff \{x_i, y_3\}\ satisfy\ (5.15)\ \forall\ i\ \in\ \{1,2,3\} \tag{5.17}$$

Let $h_1 = x_i$ and $h_2 = y_3$. So by (5.13) and (5.14) we have

$$\begin{aligned}a_1 &= \frac{4}{3}(h_1,y_1) - \frac{2}{3}(h_1,y_2) = \frac{4}{3}(x_i,y_1) - \frac{2}{3}(x_i,y_2)\\ b_1 &= \frac{4}{3}(h_1,y_2) - \frac{2}{3}(h_1,y_1) = \frac{4}{3}(x_i,y_2) - \frac{2}{3}(x_i,y_1)\\ a_2 &= \frac{4}{3}(h_2,y_1) - \frac{2}{3}(h_2,y_2) = \frac{4}{3}(y_3,y_1) - \frac{2}{3}(y_3,y_2) = \frac{1}{3}\\ b_2 &= \frac{4}{3}(h_2,y_2) - \frac{2}{3}(h_2,y_1) = \frac{4}{3}(y_3,y_2) - \frac{2}{3}(y_3,y_1) = \frac{1}{3}\end{aligned}$$

So by (5.15)

$$\mathrm{X} \subset \bar{H}(y_1,y_2;y_3) \iff$$

$$(x_i,y_3) - (\frac{4}{9}(x_i,y_1) - \frac{2}{9}(x_i,y_2) + \frac{4}{9}(x_i,y_2) - \frac{2}{9}(x_i,y_1) + \frac{\frac{4}{9}(x_i,y_1) - \frac{2}{9}(x_i,y_2)}{2} + \frac{\frac{4}{9}(x_i,y_2) - \frac{2}{9}(x_i,y_1)}{2}) \leq 0$$

So

$$\mathrm{X} \subset \bar{H}(y_1,y_2;y_3) \iff$$

$$(x_i,y_3) - \frac{(x_i,y_1) + (x_i,y_2)}{3} \leq 0$$

$\forall\ i\ \in\ \{1,2,3\}$. Similarly

$$\mathrm{Y} \subset \bar{H}(x_1,x_2;x_3) \iff$$

$$(y_i, x_3) - \frac{(y_i, x_1) + (y_i, x_2)}{3} \leq 0$$

Claim 5.0.0.31. *If $\sum_{i,j\ \in \{1,2,3\}} (x_i, y_j) > 2$ then X and Y donot overlap iff atleast one of the following is true.*

$$X \subset \bar{H}(y_1, y_2; y_3)\ i.e\ (x_i, y_3) - \frac{(x_i, y_1) + (x_i, y_2)}{3} \leq 0\ \forall\ i \in\ \{1, 2, 3\}$$
$$X \subset \bar{H}(y_2, y_3; y_1)\ i.e\ (x_i, y_1) - \frac{(x_i, y_2) + (x_i, y_3)}{3} \leq 0\ \forall\ i \in\ \{1, 2, 3\}$$
$$X \subset \bar{H}(y_1, y_3; y_2)\ i.e\ (x_i, y_2) - \frac{(x_i, y_1) + (x_i, y_3)}{3} \leq 0\ \forall\ i \in\ \{1, 2, 3\}$$
$$Y \subset \bar{H}(x_1, x_2; x_3)\ i.e\ (y_i, x_3) - \frac{(y_i, x_1) + (y_i, x_2)}{3} \leq 0\ \forall\ i\ \in\ \{1, 2, 3\}$$
$$Y \subset \bar{H}(x_2, x_3; x_1)\ i.e\ (y_i, x_1) - \frac{(y_i, x_2) + (y_i, x_3)}{3} \leq 0\ \forall\ i\ \in\ \{1, 2, 3\}$$
$$Y \subset \bar{H}(x_1, x_3; x_2)\ i.e\ (y_i, x_2) - \frac{(y_i, x_1) + (y_i, x_3)}{3} \leq 0\ \forall\ i\ \in\ \{1, 2, 3\}$$

Proof. $\Leftarrow$ is clear.
$\Rightarrow$:
Consider this diagram. Let the spherical triangle in the diagram be X .

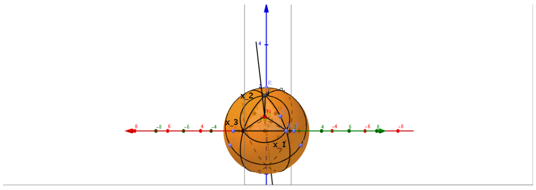

Figure 5.2: spherical caps

Let Y be any other spherical equilateral triangle.If $Y \subset \bar{H}(x_1, x_3; x_2)$ then we are done. If not there are 3 cases.If we consider each case we get atleast one of the above conditions.

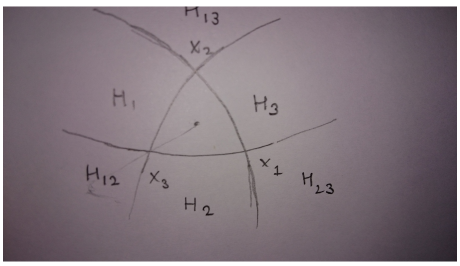

Figure 5.3: Kissing Number for $n = 2$

1. 2 points in $\bar{H}(x_1, x_3; x_2)$ and 1 in other halfspace. wlg call those 2 points y_1 and y_2
   (a) y_1, y_2 in H_2 : Atleast one edge y_3, y_1 or y_3, y_2 of spherical equilateral triangle passes through interior of X so X and Y overlap.
   (b) y_1 in H_{12} and y_2 in H_{23} not possible because we want equilateral triangle.
   (c) y_1 in H_{12} and y_2 in H_2.Then we can have
   from figure we can see that X subset of $\bar{H}(y_2, y_3; y_2)$.

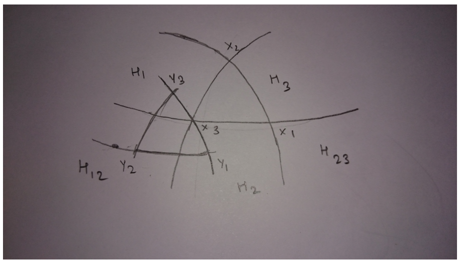

Figure 5.4: Kissing Number for $n = 2$

2. 1 point in $\bar{H}(x_1, x_3; x_2)$ and 2 in other.
3. All 3 in halfspace different from $\bar{H}(x_1, x_3; x_2)$.

□

Now as we did in kissing number problem we can express Ω in a semidefinite program with the help of polynomial optimization.

5.1 Appendix

Here I am adding the code for kissing number problem.
load("/Applications/sage/build/pkgs/SDP/SDP.py")
load("/Users/satishjoshi/Desktop/documents/Gegenbauerpoly.sage")
load("/Applications/sage/gcoeff.sage")
load("/Applications/sage/sos.sage")
load("/Applications/sage/soskiss.sage")

$$\begin{aligned}
c &= matrix(RR, 2*D+1, 1, lambdai, j : G_coeff(n, j, i)) \\
G &= matrix(RR, 2*D+1, 2*D, lambdai, j : G_coeff(n, j+1, i)) \\
G_0 &= transpose(G) \\
G_1 &= transpose(matrix(RR, 2*D+1, (D+1)^2, lambdai, j : sos(D, i, j))) \\
G_2 &= transpose(matrix(RR, 2*D+1, D^2, lambdai, j : soskiss(D, i, j))) \\
h_0 &= transpose(matrix([1 for\ i\ in\ [0..(2*D-1)]])) \\
h_1 &= matrix(RR, D+1, D+1, lambdai, j : 0) \\
h_2 &= matrix(RR, D, D, lambdai, j : 0) \\
c &= matrix_converter(c,' cvxopt') \\
G_0 &= matrix_converter(G_0,' cvxopt') \\
G_1 &= matrix_converter(G_1,' cvxopt') \\
G_2 &= matrix_converter(G_2,' cvxopt') \\
h_0 &= matrix_converter(h_0,' cvxopt') \\
h_1 &= matrix_converter(h_1,' cvxopt') \\
h_2 &= matrix_converter(h_2,' cvxopt')
\end{aligned}$$

$h_s = [h_1, h_2]$

$G_s = [G_1, G_2]$

import cvxopt from cvxopt import matrix, solvers

sol = solvers.sdp$(c, G_0, h_0, Gs = G_s, hs = h_s)$
print sol['x']

Gegenbauerpoly.sage

def G(n,k) :
R.<x>=CC[ ]
a=(n-2)/2
c=$sage.functions.orthogonal\_polys.gegenbauer(k, a, 1)$
b=$sage.functions.orthogonal\_polys.gegenbauer(k, a, x)$
if k>0:
return b/c
else:

```
return 0*x + b/c
```

gcoeff.sage

```
defG_coeff(n, k, d) :
if d<k+1 :
R.<x>=CC[]
load("/Users/satishjoshi/Desktop/documents/Gegenbauerpoly.sage")
p=G(n,k)
v=p.coefficients(sparse=False)
return v[d]
else :
return 0
```

sos.sage

```
def sos(D,i,j):
a=j%(D + 1)
b=int(j/(D+1))
if a+b==i:
return 1
else:
return 0
```

soskiss.sage

```
def soskiss(D,i,j):
v=[1/2, -1/2, -1]
a=j%D
b=int(j/D)
c=i-(a+b)
if 0 ≤ c ≤ 2 :
return v[c]
else:
return 0
```

Bibliography

[1] *Monique Laurent : Semidefinite Optimization*

[2] *Chudnovsky,Robertson,Seymour,Thomas ; The strong perfect graph theorem; Annals of mathematics,164(2006)*

[3] *Henri Lombardi ; Effective real Nullstellensatz and variants*

[4] *Andrews -Askey-Roy - Special - functions - 1999*

[5] *Feng Dai and Yuan Xu : Spherical Harmonics*

Bibliography

[1] **Monique Laurent** ;*Semidefinite Optimization*; May 22, 2012

[2] **Maria Chudnovsky, Neil Robertson,Paul Seymour, Robin Thomas** ;*The strong perfect graph theorem* ; Annals of mathematics ;**164**(2006),**51-229**

[3] **Henri Lombardi** ; *Effective real Nullstellensatz and variants*

[4] **Andrews ,Askey,Roy** ;*Special functions* Cambridge University Press;(1999)

[5] **Feng Dai ,Yuan Xu** ; *Spherical Harmonics* ; arXiv:1304.2585