SEMIDEFINITE PROGRAMMING UPPER BOUNDS FOR PACKING PROBLEMS

FERNANDO MÁRIO DE OLIVEIRA FILHO

20.06.2016

§1. Notation

For x, $y \in \mathbb{R}^n$, let $x \cdot y = x_1y_1 + \cdots + x_ny_n$ denote the Euclidean inner product and write $\|x\| = (x \cdot x)^{1/2}$ for the norm. By $S^{n-1} = \{ x \in \mathbb{R}^n : \|x\| = 1 \}$ we denote the $(n-1)$-dimensional unit sphere and by ω the surface measure on it. A *spherical cap* is the intersection of the unit sphere with a half-space. By $B^n = \{ x \in \mathbb{R}^n : \|x\| \leq 1 \}$ we denote the n-dimensional unit ball.

A *lattice* $\Lambda \subseteq \mathbb{R}^n$ is a discrete subgroup of $\mathbb{R}^n$, i.e., it is the set of all integer linear combinations of some vectors $u_1, \ldots, u_n$ that form a basis of $\mathbb{R}^n$. A nonzero vector $x \in \Lambda$ is a *minimal vector* if it has minimum length among all nonzero vectors in Λ.

The *Hamming cube* is the set $H_n = \mathbb{Z}_2^n$ of all n-bit words. The *Hamming distance* between two words is the number of bits in which they differ. The *weight* of a word x is the number of 1s in it and is denoted by $|x|$, so that the Hamming distance between x, $y \in H_n$ is $|x + y| = |x - y|$.

§2. Packing problems

Below are described the three kinds of packing problems considered in these notes: (i) packings of convex bodies in Euclidean space, (ii) packings of spherical caps on a sphere, and (iii) packings of balls in the Hamming cube. While I have tried to provide some historical background for each problem, it is hardly possible to be complete in such short notes; the interested reader is referred to the excellent book by Conway and Sloane [8] for further references.

Packing convex bodies in Euclidean space. Perhaps the most famous packing problem is the *sphere packing problem*: what is the maximum fraction of $\mathbb{R}^n$ that can be covered with pairwise-nonoverlapping equal balls? A union of nonoverlapping equal balls is a *sphere packing*, and the fraction of space it covers is its *density*. (For the moment this informal definition of density will suffice; a rigorous definition is given in §9.)

For $n = 1$ the solution is trivial: balls are simply intervals, and we can cover the whole real line with them. For $n = 2$ the solution is also familiar; as Coxeter [10] put it:

> The problem of packing, as densely as possible, an unlimited number of equal nonoverlapping circles in a plane was solved millions of years ago by the bees, who found that the best arrangement consists of circles inscribed in the hexagons of the regular tessellation $\{6, 3\}$.

What he means is that for $n = 2$ the optimal packing can be constructed as follows: tile the plane like a honeycomb with regular hexagons and consider for each hexagon its inscribed circle. Alternatively, consider the hexagonal lattice generated by the vectors $(1, 0)$ and $(1/2, \sqrt{3}/2)$ and place centered on each lattice point a circle of radius $1/2$ (Figure 1).

This packing has density $\pi/\sqrt{12} = 0.9069\ldots$. According to Fejes Tóth [13] (Chapter III, §13), its optimality follows from a theorem of Thue presented in 1892; see also the later paper by Thue [38]. Fejes Tóth also gives an elegant proof of this result (ibid., Chapter III, §2).

The only other value of n for which the optimal packing density is known is $n = 3$. The optimal packing was described by Johannes Kepler in his 1611 work *De Nive*

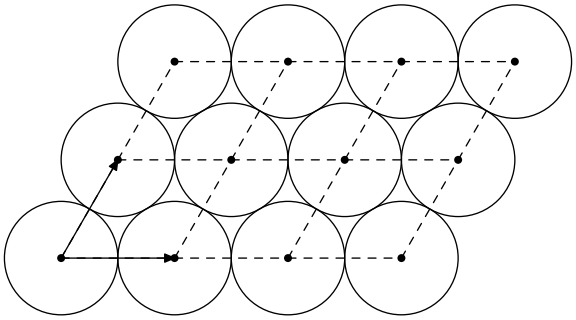

FIGURE 1. The densest packing of circles in the Euclidean plane. The points in the hexagonal lattice are marked, as well as the two vectors that generate the lattice.

Sexangula (On the Six-Cornered Snowflake [20]). Kepler describes (ibid., p. 39) the packing layer by layer. First, pack balls on a plane using a square lattice, so that each ball touches four others. Then there are two ways (Kepler says orders) in which to proceed so as to get a packing for the whole of space:

> [...] the balls from the plane above can be placed on top of the balls from the plane below [that is, shift up the square lattice packing] or instead each ball from the upper plane can lie between four balls of the lower plane. In the first case each ball is touched by four neighbors on the same plane and by one neighbor from the plane above and one from the plane below, so that it lies in the neighborhood of six others [...] This is however not the densest packing. In the second order each ball is not only touched by its four neighbors on the same plane, but also by four neighbors on the plane above and four on the plane below. There are in total then twelve neighbors [...] This is the densest possible packing, and in no other order is it possible to put more balls in a container.

The packing described by Kepler is called the *cubic close-packing*, which is most often used to stack fruits in supermarkets. Another way to describe it is as follows. Consider the lattice consisting of all vectors $x \in \mathbb{Z}^3$ whose coordinates sum to an even number. This is the *face-centered cubic* (or *fcc*) *lattice*. The minimum distance between two distinct vectors in this lattice is $\sqrt{2}$, so if we center on each lattice point a ball of radius $\sqrt{2}/2$, then we obtain a sphere packing in which each ball touches exactly 12 others (Figure 2). The density of this packing is $\pi/\sqrt{18} = 0.7405\ldots$.

Kepler claimed this packing is optimal. A proof, however, was only found in 1998 by Hales [16] with heavy use of computers. (Strictly speaking, Kepler claimed that no other packing *inside a given container* could have higher density, an assertion that fails for many different containers; see Schürmann [32].)

Of course, we may pack any convex body $\mathcal{K} \subseteq \mathbb{R}^n$ (a *convex body* is a compact subset of $\mathbb{R}^n$ with nonempty interior), not only balls. Here two kinds of packings can be considered: *translational packings*, when we want to fill as much of space with pairwise nonoverlapping translated copies of $\mathcal{K}$, and *congruent packings*, when we also allow $\mathcal{K}$ to be rotated. So we consider the two parameters

$$\Delta_{\mathsf{T}}(n, \mathcal{K}) = \text{maximum density of a translational packing of } \mathcal{K} \text{ in } \mathbb{R}^n, \text{ and}$$
$$\Delta_{\mathsf{C}}(n, \mathcal{K}) = \text{maximum density of a congruent packing of } \mathcal{K} \text{ in } \mathbb{R}^n.$$

The sphere packing problem asks for $\Delta_{\mathsf{T}}(n, B^n) = \Delta_{\mathsf{C}}(n, B^n)$.

Problems like the sphere packing problem or the problem of deciding whether space could be tiled by polyhedra were important driving forces in the development of geometry. Hilbert mentioned such problems as part of his 18th problem [17]:

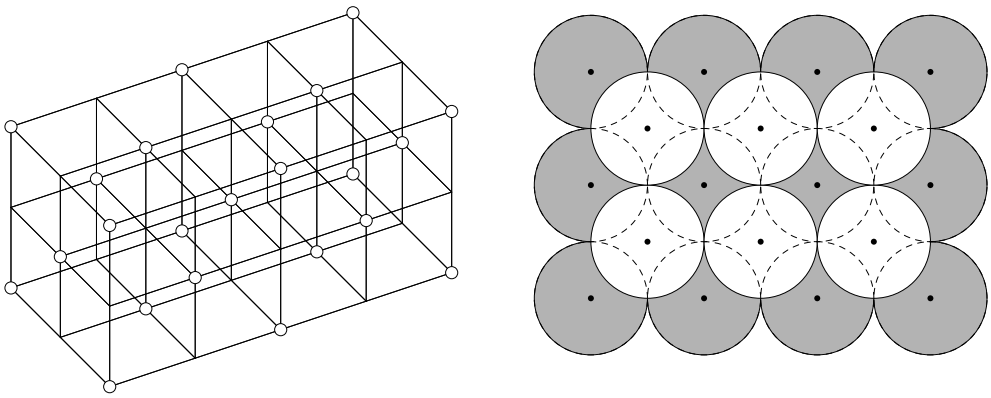

FIGURE 2. On the left, the cubic lattice $\mathbb{Z}^3$ with the points whose coordinates sum to an even number marked; these points form the face-centered cubic lattice. On the right, a top view of two layers of the cubic-close packing. Spheres on the upper layer, in white, are placed over the holes between spheres on the bottom layer, in gray.

> 18. Building up of Space from Congruent Polyhedra.
> (...) I point out the following question, related to the preceding one, and important to number theory and perhaps sometimes useful to physics and chemistry: How can one arrange most densely in space an infinite number of equal solids of given form, e.g., spheres with given radii or regular tetrahedra with given edges (or in prescribed position), that is, how can one so fit them together that the ratio of the filled to the unfilled space may be as great as possible?

The problem of packing congruent copies of a regular tetrahedron in $\mathbb{R}^3$, mentioned by Hilbert above, is perhaps the earliest packing problem that was considered. Its history goes back to Aristotle's refutation of a theory of Plato that assigns different Platonic solids to each of the four elements (so to earth one assigns the cube, to fire the tetrahedron, etc.). Aristotle claimed that tetrahedra can tile space, just as cubes can. The scene for the debate was set for the next several centuries; it is accepted that, only in the fifteenth century, Johannes Müller von Königsberg, known as Regiomontanus, proved that Aristotle's claim was false. His manuscript was lost, but a contemporary manuscript by Francesco Maurolico survives (see Oliveira and Vallentin [27] for a more complete account and references; Struik [36] presents a detailed history of the problem).

Chen, Engel, and Glotzer [6] hold the record for the densest congruent packing of tetrahedra in $\mathbb{R}^3$, with a packing density of ≈ 0.8563 (compare with the packing density for spheres). The best known upper bound for the packing density, of $1 - 2.6\ldots \cdot 10^{-25}$, is due to Gravel, Elser, and Kallus [14].

Packing spherical caps on a sphere. The *kissing number problem* asks for the maximum number of nonoverlapping unit balls in $\mathbb{R}^n$ that can simultaneously touch a central ball. This maximum is denoted by τ_n. Clearly, $\tau_1 = 2$. It is easy to prove that $\tau_2 = 6$. For $n = 3$, however, the problem becomes more interesting.

In the cubic close-packing of Kepler, each ball touches exactly 12 others, so $\tau_3 \geq 12$. Does equality hold? According to Coxeter [9]:

> Among the unpublished papers of David Gregory, H.W. Turnbull found notes of a conversation with Newton in 1694 about the distribution of stars of various magnitudes. The question arose: Can a rigid material sphere be brought into contact with 13 other such spheres of the same size? Gregory said "Yes", and Newton said "No"; but 180 years were to elapse before a conclusive answer was given.

(Coxeter credits Hoppe with proving Newton right in 1874. Conway and Sloane [8] say

"some proofs appeared in the nineteenth century" and mention Hoppe. Schütte and van der Waerden [33] gave a correct and detailed proof; they are usually mentioned in connection with the problem.)

Was Gregory justified in thinking that $\tau_3 = 13$? How can we find an upper bound for τ_3, or for τ_n for that matter?

A *spherical code* with *minimum angular distance* θ is a set $C \subseteq S^{n-1}$ such that the angle between any two distinct points in C is at least θ. The maximum size of a spherical code in S^{n-1} with minimum angular distance θ is denoted by $A(n, \theta)$.

If two unit balls touch a central one, then the angle between the contact points is at least $\pi/3$. So a *kissing configuration* in $\mathbb{R}^n$, i.e., an arrangement of nonoverlapping unit balls that simultaneously touch a central one, corresponds to a spherical code in S^{n-1} with minimum angular distance $\pi/3$, and vice versa. In other words, $\tau_n = A(n, \pi/3)$.

Now, given a spherical code with minimum angular distance θ, we may place around each point in the code a spherical cap of angular radius $\theta/2$, and these caps will not overlap. So a spherical code corresponds naturally to a packing of spherical caps on the sphere. Hence, if $K(n, \alpha)$ is a spherical cap in S^{n-1} with angular radius α, then

$$A(n, \theta) \leq \lfloor \omega(S^{n-1})/\omega(K(n, \theta/2)) \rfloor.$$

This simple bound is often called the *volume bound*. For the kissing number problem (i.e., $\theta = \pi/3$) and $n = 2$, the right-hand side is exactly 6, and this bound therefore proves that the arrangement of six circles around a central one is optimal. For $n = 3$ we have $\omega(S^2)/\omega(K(3, \pi/6)) = 14.9282\ldots$, and so $\tau_3 \leq 14$ from the volume bound. It seems then that there is almost enough space to place 15 balls around the central one! Actually, there is so much space left in the cubic close-packing configuration of 12 balls that it is possible to achieve any permutation of the 12 balls by rolling them around in such a way that they never overlap (see Conway and Sloane [8], Chapter 1).

So Gregory, while wrong, was not completely without reason. Aside from the possible proofs that $\tau_3 = 12$ that appeared in the nineteenth century, bounds that show that $\tau_3 \leq 13$ appeared only in the the twentieth century and are not elementary (we will see one such bound in §8).

The kissing number is only known in dimensions 1–4, 8, and 24. Musin [25] showed that $\tau_4 = 24$. Odlyzko and Sloane [26] showed that $\tau_8 = 240$ and $\tau_{24} = 196560$.

Packing balls in the Hamming cube. A *binary code of length n and minimum distance d* is a set $C \subseteq H_n$ such that the Hamming distance between any two distinct words in C is at least d. Let C be a binary code of minimum distance d. When transmitting information, if sender and receiver agree that only words in C are transmitted, then the minimum distance between words implies that words with up to $\lfloor (d-1)/2 \rfloor$ wrong bits can be corrected by the receiver; the correct word is, in this case, the word in C closest to the received word.

Another way to look at this is as follows. If C is a binary code of length n and minimum distance d, then for each word in C we may consider the ball of radius $\lfloor (d-1)/2 \rfloor$ centered on the word (this ball is the set of all words whose Hamming distance from the center is at most $\lfloor (d-1)/2 \rfloor$). These balls are pairwise-disjoint. So, a binary error-correcting code corresponds to a packing of balls in the Hamming cube.

One of the main problems of coding theory is to find large error-correcting codes with given length and minimum distance. In the introduction to his thesis, Delsarte [11] says:

> Research in coding theory may be divided into three main parts. The first way, opened by Shannon [34], consists in a study of the theoretical possibilities offered by the principle of coding for correction of errors in certain communication systems (...). At this level there already arise some algebraic concepts, such as the *minimum*

distance between distinct codewords; among codes having the same *length* n and the same minimum distance d, the best is the one containing the largest number of words.

It is therefore natural that many authors applied themselves to construct "good" codes of fixed parameters n and d.

The maximum size of a binary error-correcting code of length n and minimum distance d is denoted by $A(n, d)$. Determining $A(n, d)$ is an important problem in coding theory, open for most values of n and d (see Chapter 3 of Conway and Sloane [8]).

* * *

Through constructions of sphere packings, spherical codes, or binary codes, one provides lower bounds for the parameters $\Delta_{\mathsf{T}}(n, B^n)$, $A(n, \theta)$, and $A(n, d)$. These notes are concerned however with the nonconstructive task of providing upper bounds for these parameters.

Delsarte [11] proposed a bound for $A(n, d)$ that became known as the *linear programming bound*. Later, Delsarte, Goethals, and Seidel [12] gave an analogous bound for $A(n, \theta)$, and Cohn and Elkies [7] proposed an analogous bound for $\Delta_{\mathsf{T}}(n, \mathcal{K})$. All three linear programming bounds are similar, yet their common root is not immediately clear. This common root is the Lovász theta number (see §5 and §7); identifying it allows us to present a unified treatment and to more easily adapt these results to other situations.

A final remark before we proceed. Binary error-correcting codes come up naturally in the context of communication theory, and their study has been motivated by its development, as noted above. Perhaps less clear is that sphere packings and spherical codes, purely geometrical problems, also have important applications in communication theory. In fact, Shannon considers sphere packings in his foundational work [34], and spherical codes have applications to the design of signals for the Gaussian channel (see Sloane [35]).

§3. Packings and the independence number

Let $G = (V, E)$ be a graph (our graphs never have loops nor parallel edges). A set $C \subseteq V$ is *independent* if x, y are nonadjacent for all x, $y \in C$. The *independence number* of G is the maximum cardinality of any independent set of G and is denoted by $\alpha(G)$.

To each packing problem of the previous section we may associate a *packing graph* in such a way that packings correspond to independent sets and vice versa, and determining the optimal packing density becomes the same as determining the independence number of the packing graph (even if sometimes we have to redefine what the independence number is, as in the case of sphere packings below). Let us now see which packing graph is associated with each problem.

Binary error-correcting codes. Given n, $d > 0$, consider the graph $G(n, d)$ whose vertex set is H_n and in which distinct vertices x, $y \in H_n$ are adjacent if $|x-y| \leq d-1$. Then $C \subseteq H_n$ is independent in $G(n, d)$ if and only if C is a binary code of length n and minimum distance d. So $A(n, d) = \alpha(G(n, d))$.

Spherical codes. Given $n \geq 1$ and $\theta \in (0, \pi]$, let $G(n, \theta)$ be the graph whose vertex set is S^{n-1} and in which distinct vertices x, $y \in S^{n-1}$ are adjacent if $\cos\theta < x \cdot y < 1$. Then $C \subseteq S^{n-1}$ is independent in $G(n, \theta)$ if and only if C is a spherical code with minimum angular distance θ. It follows that $A(n, \theta) = \alpha(G(n, \theta))$.

Note here that, unlike $G(n, d)$, graph $G(n, \theta)$ is infinite. However, it still has a finite independence number.

Translational body packings. Given a convex body $\mathcal{K} \subseteq \mathbb{R}^n$, consider the graph $G_{\mathsf{T}}(n, \mathcal{K})$ whose vertex set is $\mathbb{R}^n$ and in which distinct vertices x, $y \in \mathbb{R}^n$ are ad-

jacent if $(x + \mathcal{K}^\circ) \cap (y + \mathcal{K}^\circ) \neq \emptyset$, where $\mathcal{K}^\circ$ is the interior of $\mathcal{K}$. Then $C \subseteq \mathbb{R}^n$ is independent in $G_\mathsf{T}(n, \mathcal{K})$ if and only if $\bigcup_{x \in C} x + \mathcal{K}$ is a packing of translates of $\mathcal{K}$.

Graph $G_\mathsf{T}(n, \mathcal{K})$ is infinite and so is its independence number. A way around this problem is to define the *size* of an independent set as the density of the corresponding packing and redefine the independence number accordingly. Another option is to use a compactification approach and avoid dealing with infinite independent sets altogether; this is the approach that we shall take in §9. Packing graphs can also be used to model congruent body packings, but the idea is not discussed in these notes.

* * *

The correspondence between packings and independent sets of graphs is more of a rewording than anything else. We are moreover left with huge graphs like $G(n, d)$, or even infinite ones like $G(n, \theta)$ or $G_\mathsf{T}(n, \mathcal{K})$, not to mention that determining the independence number of a graph is a classical NP-hard combinatorial problem. So what do we get by establishing this connection?

When dealing with a graph parameter like the independence number, that is hard to compute, one looks for lower and upper bounds that can be efficiently computed. Lovász [22] proposed an upper bound for the independence number of a finite graph, based on semidefinite programming, that can be efficiently computed. It is a slight variation of this bound that we will adapt and use to compute bounds for the independence number of the packing graphs presented above.

§4. Semidefinite programming basics

This is just a quick summary of the basic facts of semidefinite programming that we will need; for more background see the book by Tunçel [39].

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is *positive semidefinite* if $x^\mathsf{T} A x \geq 0$ for all $x \in \mathbb{R}^n$ or, equivalently, if all eigenvalues of A are nonnegative. (The term positive semidefinite is used exclusively for symmetric matrices.) If A, $B \in \mathbb{R}^{n \times n}$ are matrices, then

$$\langle A, B \rangle = \operatorname{tr} A^\mathsf{T} B = \sum_{i,j=1}^{n} A_{ij} B_{ij}$$

is the *trace inner-product* between A and B. If both A and B are positive semidefinite, then $\langle A, B \rangle \geq 0$.

Let $I_=$ and $I_\leq$ be two finite and disjoint sets of indices, and let $C \in \mathbb{R}^{n \times n}$ and $A_i \in \mathbb{R}^{n \times n}$, for $i \in I_= \cup I_\leq$, be symmetric matrices, and b_i, for $i \in I_= \cup I_\leq$, be real numbers. A *semidefinite programming problem* asks us to maximize the linear function $X \mapsto \langle C, X \rangle$ where $X \in \mathbb{R}^{n \times n}$ ranges over all positive semidefinite matrices satisfying some given linear constraints, namely:

$$\begin{aligned} \langle A_i, X \rangle &= b_i \quad \text{for } i \in I_=, \\ \langle A_i, X \rangle &\leq b_i \quad \text{for } i \in I_\leq. \end{aligned}$$

As usual, we will write this problem in the following compact form:

$$\begin{aligned} \max\ & \langle C, X \rangle \\ & \langle A_i, X \rangle = b_i \quad \text{for } i \in I_=, \\ & \langle A_i, X \rangle \leq b_i \quad \text{for } i \in I_\leq, \\ & X \text{ is positive semidefinite.} \end{aligned} \tag{1}$$

The "max" above should be read as "maximize", and it does not imply that there is an optimal solution: the problem could be infeasible, unbounded, or — and here lies a difference between semidefinite programming and linear programming — the

optimal value could be finite but never attained. If all the matrices C and A_i are diagonal, then the problem above is a linear programming problem.

Just as a linear programming problem has a dual, so does problem (1), namely

$$\begin{array}{l} \min \sum_{i\in I_= \cup I_\leq} y_i b_i \\ \quad \sum_{i\in I_= \cup I_\leq} y_i A_i - C \text{ is positive semidefinite,} \\ \quad y_i \geq 0 \text{ for } i \in I_\leq. \end{array} \tag{2}$$

Here, too, "min" should be read as "minimize". The optimization variables are now the y_i, but (2) can be written in the same form as (1) with a little work. You should try to do so and then compute the dual of the resulting problem to see that the dual of the dual problem is again the primal problem.

Just as in the case of linear programming, primal and dual are related via two important results: weak duality and strong duality. Weak duality asserts that, if X is a feasible solution of (1) and y is a feasible solution of (2), then $\langle C, X\rangle \leq \sum_{i\in I_= \cup I_\leq} y_i b_i$. Indeed, matrices X and $\sum_{i\in I_= \cup I_\leq} y_i A_i - C$ are both positive semidefinite, hence their trace inner-product is nonnegative. But then

$$0 \leq \sum_{i\in I_= \cup I_\leq} y_i \langle A_i, X\rangle - \langle C, X\rangle \leq \sum_{i\in I_= \cup I_\leq} y_i b_i - \langle C, X\rangle,$$

as we wanted.

Strong duality asserts that the optimal values of primal and dual coincide, that is, there is no *duality gap*, and that moreover both optima are attained. Unlike in the case of linear programming, when strong duality always holds, for some pairs of primal-dual semidefinite programs the duality gap maybe be nonzero. A sufficient condition for the duality gap to be zero is to require either primal or dual to be strictly feasible.

Primal problem (1) is *strictly feasible* if it admits a positive *definite* feasible solution X such that $\langle A_i, X\rangle < b_i$ for all $i \in I_\leq$. Dual problem (2) is *strictly feasible* if it admits a feasible solution y such that $\sum_{i\in I_= \cup I_\leq} y_i A_i - C$ is positive *definite* and $y_i > 0$ for all $i \in I_\leq$.

Theorem 1. *If the dual is bounded from below and strictly feasible, then there is an optimal primal solution and no duality gap. Similarly, if the primal is bounded from above and strictly feasible, then there is an optimal dual solution and no duality gap.*

Finally, what makes semidefinite programming attractive as a generalization of linear programming is that there are efficient algorithms for it, both in theory and in practice. In theory, under a few extra assumptions, the ellipsoid method can be used to solve semidefinite programming problems to any desired precision in polynomial time (see the book by Grötschel, Lovász, and Schrijver [15]). In practice, interior point methods can solve moderately-sized problems efficiently, though numerical instability is much more of an issue than with linear programming.

§5. The Lovász theta number

Let $G = (V, E)$ be a finite graph. Lovász [22] introduced a graph parameter $\vartheta(G)$, now called the *Lovász theta number* of G, that can be computed to any desired precision in polynomial time and that provides an upper bound for the independence number of G, that is, $\vartheta(G) \geq \alpha(G)$. Perhaps not surprisingly given all the connections to communication theory mentioned in §2, the theta number was actually introduced in the context of communication theory as an upper bound to the Shannon capacity of a graph.

A variation of the theta number, called the *theta prime number* of G and denoted by $\vartheta'(G)$, was later introduced by McEliece, Rodemich, and Rumsey [24] and Schrijver [31]. There are many equivalent definitions of $\vartheta'(G)$; one of them is as the optimal value of the following semidefinite programming problem:

$$\begin{aligned} \max\ & \langle J, L\rangle \\ & \operatorname{tr} L = 1, \\ & L(x,y) = 0 \quad \text{if } xy \in E, \\ & L\colon V \times V \to \mathbb{R} \text{ is nonnegative and positive semidefinite,} \end{aligned} \tag{3}$$

where J is the all-ones matrix.

This is indeed a semidefinite programming problem in form (1), since all constraints can be directly represented using the trace inner-product. Parameter $\vartheta(G)$ is the optimal value of problem (3) when the constraint that L has to be nonnegative is dropped.

Theorem 2. $\alpha(G) \le \vartheta'(G)$.

Proof. Let C be a nonempty independent set of $G = (V, E)$ and let $\mathbf{1}_C\colon V \to \{0,1\}$ be the characteristic function of C, that is, $\mathbf{1}_C(x) = 1$ if and only if $x \in C$. Then the matrix L such that

$$L(x,y) = \frac{1}{|C|}\mathbf{1}_C(x)\mathbf{1}_C(y)$$

for all x, $y \in V$ is a feasible solution of (3), and hence $\vartheta'(G) \ge \langle J, L\rangle = |C|$. Since C is an arbitrary nonempty independent set, the theorem follows. ∎

The dual of (3) is the problem

$$\begin{aligned} \min\ & \lambda \\ & Z(x,x) \le \lambda \quad \text{for all } x \in V, \\ & Z(x,y) \le 0 \quad \text{if } x \ne y \text{ and } xy \notin E, \\ & Z\colon V \times V \to \mathbb{R} \text{ is symmetric and } Z - J \text{ is positive semidefinite.} \end{aligned} \tag{4}$$

(This problem is not in the form (2), but can be derived from it. This is a simple but good exercise!)

Both primal and dual are strictly feasible and bounded, and hence both optima are attained and there is no duality gap. So $\vartheta'(G)$ can be alternatively defined as the optimal value of (4).

Actually, any feasible solution of (4) provides an upper bound to $\alpha(G)$. A direct proof is as follows. Let C be a nonempty independent set and (Z, λ) be a feasible solution of (4). Then

$$0 \le \sum_{x,y\in C} (Z - J)(x,y) = \sum_{x,y\in C} Z(x,y) - |C|^2 \le \lambda|C| - |C|^2,$$

whence $|C| \le \lambda$.

§6. An upper bound for $A(n,d)$

Graph $G(n,d)$ from §3, which is such that $A(n,d) = \alpha(G(n,d))$, is a finite graph. So $\vartheta'(G(n,d))$ as defined in the previous section provides an upper bound for $A(n,d)$. Since $G(n,d)$ has 2^n vertices, we cannot hope to solve the semidefinite programming problem that defines $\vartheta'(G(n,d))$ even for small values of n. Graph $G(n,d)$ is highly symmetric however, and this can be used to simplify (4) considerably.

An *isometry* of H_n is a bijection $\varphi\colon H_n \to H_n$ that preserves the Hamming distance. The set of all isometries is a group under function composition, denoted

by $\mathrm{Iso}(H_n)$. Isometries can be obtained by flipping and permuting bits. More precisely, given $x_0 \in H_n$ and $\sigma \in \mathcal{S}_n$, we can construct the isometry $x \mapsto x_0 + \sigma x$, where the sum is carried out in $\mathbb{Z}_2^n$ and σx is obtained by permuting the bits of x according to the permutation σ; every isometry is of this form. Every isometry of H_n is an automorphism of $G(n, d)$.

$\mathrm{Iso}(H_n)$ acts on matrices $A\colon H_n \times H_n \to \mathbb{R}$: for an isometry φ we have

$$(\varphi \cdot A)(x, y) = A(\varphi^{-1}x, \varphi^{-1}y) \qquad \text{for all } x,\, y \in H_n.$$

If $\varphi \cdot A = A$ for all $\varphi \in \mathrm{Iso}(H_n)$, then A is *invariant*.

The crucial step in solving (4) is to require Z to be invariant. Then we still get an upper bound (actually, this restriction does not worsen the bound, as we will see below) and problem (4) can be greatly simplified, since invariant and positive semidefinite matrices can be conveniently parametrized.

Theorem 3. *A matrix $A\colon H_n \times H_n \to \mathbb{R}$ is invariant and positive semidefinite if and only if*

$$A(x, y) = \sum_{k=0}^{n} f_k K_k^n(|x - y|) \tag{5}$$

for some nonnegative numbers $f_0, \ldots, f_n$ which are uniquely determined by, and uniquely determine, A.

Here, $K_k^n(t)$ is the *Krawtchouk polynomial* of degree k. Computed on an integer t with $0 \le t \le n$, it is given by

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

Proof. Suppose $A\colon H_n \times H_n \to \mathbb{R}$ is symmetric and invariant. For $u \in H_n$, let $\chi_u\colon H_n \to \mathbb{R}$ be such that $\chi_u(x) = (-1)^{u \cdot x}$. The χ_u are pairwise orthogonal and form therefore a basis of $\mathbb{R}^{H_n}$. Each χ_u is also an eigenvector of A, since

$$\begin{aligned}
(A\chi_u)(x) &= \sum_{y \in H_n} A(x, y)(-1)^{u \cdot y} \\
&= \sum_{y \in H_n} A(0, x + y)(-1)^{u \cdot y} \\
&= \sum_{y \in H_n} A(0, y)(-1)^{u \cdot (x+y)} \\
&= (-1)^{u \cdot x} \sum_{y \in H_n} A(0, y)(-1)^{u \cdot y}.
\end{aligned}$$

The eigenvalue of χ_u depends only on $|u|$. Indeed, if for $v \in H_n$ we have $|v| = |u|$, then there is a permutation $\sigma \in \mathcal{S}_n$ such that $v = \sigma u$. But then

$$\begin{aligned}
\sum_{y \in H_n} A(0, y)(-1)^{\sigma u \cdot y} &= \sum_{y \in H_n} A(0, \sigma y)(-1)^{\sigma u \cdot \sigma y} \\
&= \sum_{y \in H_n} A(0, \sigma y)(-1)^{u \cdot y} \\
&= \sum_{y \in H_n} A(0, y)(-1)^{u \cdot y}.
\end{aligned}$$

(The last identity follows from the invariance of A and from the fact that $\sigma^{-1}0 = 0$.)

For $k = 0, \ldots, n$, denote by λ_k the common eigenvalue of χ_u when $|u| = k$. Write

$$B_k(x,y) = \sum_{\substack{u \in H_n \\ |u|=k}} \chi_u(x)\chi_u(y) = \sum_{\substack{u \in H_n \\ |u|=k}} (-1)^{u\cdot(x+y)}.$$

A simple counting argument shows that $B_k(x,y) = K_k^n(|x-y|)$. Taking $f_k = \lambda_k/2^n$, we have $A = f_0B_0 + \cdots + f_nB_n$. Since A is positive semidefinite, $\lambda_k \geq 0$ and hence $f_k \geq 0$, and so we have the desired expression for $A(x,y)$.

For the converse, each B_k is invariant and positive semidefinite. So if $A = f_0B_0 + \cdots + f_nB_n$ for nonnegative $f_0, \ldots, f_n$, then A is invariant and positive semidefinite.

Finally, the fact that the f_k are uniquely determined by, and uniquely determine, A, follows from the fact that the B_k matrices are linearly independent. ■

Now use this theorem to rewrite (4) for $G(n,d)$, assuming Z to be invariant. Note that the all-ones matrix J is invariant and positive semidefinite; its expansion (5) has only one nonzero coefficient, namely the coefficient of K_0^n, which equals 1. So $Z\colon H_n \times H_n \to \mathbb{R}$ is invariant and such that $Z - J$ is positive semidefinite if and only if

$$Z(x,y) = \sum_{k=0}^{n} f_k K_k^n(|x-y|)$$

with $f_0 \geq 1$ and $f_1, \ldots, f_n$ nonnegative.

Now, if Z is invariant, then all its diagonal entries are the same. Moreover, in $G(n,d)$ we have that x and y are adjacent if $|x-y| \in \{0, \ldots, d-1\}$, so we may rewrite (4) for invariant Z as

$$\begin{array}{ll} \min & \sum_{k=0}^n f_k K_k^n(0) \\ & \sum_{k=0}^n f_k K_k^n(t) \leq 0 \quad \text{for } t = d, \ldots, n, \\ & f_0 \geq 1 \text{ and } f_k \geq 0 \text{ for } k = 1, \ldots, n. \end{array} \tag{6}$$

This is a linear programming problem with $n+1$ variables. Its optimal value provides an upper bound for $A(n,d)$ (actually, any feasible solution provides an upper bound).

The restriction to invariant matrices does not worsen the bound given by (4), i.e., the optimal values of (4) and (6) coincide. Indeed, if (Z,λ) is any feasible solution of (4), then for any $\varphi \in \mathrm{Iso}(H_n)$ we have that $(\varphi \cdot Z, \lambda)$ is also feasible for (4). Hence

$$\overline{Z} = \frac{1}{|\mathrm{Iso}(H_n)|} \sum_{\varphi \in \mathrm{Iso}(H_n)} \varphi \cdot Z$$

is such that $(\overline{Z}, \lambda)$ is feasible for (4), and $\overline{Z}$ is invariant.

Bound (6) was proposed by Delsarte [11]; it is often called *Delsarte's bound* or *linear programming bound*. Its relation to $\vartheta'(G(n,d))$ was observed by McEliece, Rodemich, and Rumsey [24] and Schrijver [31].

We started with a huge semidefinite programming problem and, thanks to symmetry, ended up with a small linear programming problem. Exploiting the symmetry of a problem is a key technique in semidefinite programming; in these notes we will barely scratch the surface.

§7. A generalization of the theta number

Parameter ϑ' was defined for a finite graph and hence does not apply directly to an infinite graph like $G(n,\theta)$. To use ϑ' to give an upper bound for $A(n,\theta)$ we may discretize the infinite graph $G(n,\theta)$ and use the definition of ϑ' from §5, or we may try to extend the definition of §5 to infinite graphs like $G(n,\theta)$; the latter is the approach taken in this section.

It is not hard to come up with an extension of (4) to infinite graphs. For instance, we could extend the concept of positive semidefiniteness to functions $A\colon V \times V \to \mathbb{R}$ with infinite V by calling such a function *positive semidefinite* if $\big(A(x_i, x_j)\big)_{i,j=1}^{N}$ is positive semidefinite for every choice $x_1, \ldots, x_N$ of finitely many points in V. With this definition, (4) can be extended directly and the proof that any feasible solution provides an upper bound to the independence number works as given in §5.

For this extension we imposed no conditions on V. On the one hand, therefore, we get a very general bound, but on the other hand positive semidefinite functions $A\colon V \times V \to \mathbb{R}$ can be very general objects and we have little hope of dealing with them in all generality successfully.

So for our extension we impose some extra constraints on the vertex set, narrowing the space of functions we consider. More precisely, we require V to be a compact, Hausdorff, and separable measure space with a Radon measure μ that is nonzero on open sets, and then restrict our functions to be continuous kernels on V.

In what follows we need some concepts from functional analysis, which can be found in any standard book such as the one by Riesz and Sz.-Nagy [29].

Let V and μ be as above. For f, $g \in L^2(V)$, write

$$(f, g) = \int_V f(x)g(x)\, d\mu(x).$$

(We usually deal with real-valued functions. In §10, complex-valued functions will be needed, but it will always be clear when a function is complex-valued.) A *kernel* is a function in $L^2(V \times V)$. As a matrix $A \in \mathbb{R}^{n\times n}$ defines a linear transformation $A\colon \mathbb{R}^n \to \mathbb{R}^n$, so does a kernel $K \in L^2(V \times V)$ define an operator $K\colon L^2(V) \to L^2(V)$: for $f \in L^2(V)$ and $x \in V$,

$$(Kf)(x) = \int_V K(x, y)f(y)\, d\mu(y).$$

(You have to prove that the integral converges for almost all $x \in V$ and that $Kf \in L^2(V)$.)

A kernel K is *symmetric* if $K(x, y) = K(y, x)$ for all x, $y \in V$ or, alternatively, if K is self-adjoint as an operator. A symmetric kernel K is *positive* if for all $\rho \in L^2(V)$ we have

$$(K\rho, \rho) = \int_V \int_V K(x, y)\rho(x)\rho(y)\, d\mu(x)d\mu(y) \geq 0.$$

So a positive kernel is the analogue of a positive semidefinite matrix; we say "positive" instead of "positive semidefinite" just to follow tradition.

The Spectral Theorem can be generalized to kernels. We say that a number λ is an *eigenvalue* of a kernel $K \in L^2(V\times V)$ if there is a nonzero function $f \in L^2(V)$ such that $Kf = \lambda f$, in which case f is an *eigenfunction associated* with λ. The Hilbert-Schmidt theorem states that if K is a symmetric kernel, then there is a complete orthonormal system f_1, f_2, ... of $L^2(V)$ and real numbers λ_1, λ_2, ..., where λ_i is an eigenvalue of K and f_i is an associated eigenfunction of λ_i, such that

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i f_i(x) f_i(y)$$

with convergence in the L^2 norm. As a consequence we may prove that a symmetric kernel K is positive if and only if all its eigenvalues are nonnegative, as is the case for a matrix.

Finally, we come to the extension of (4). Let $G = (V, E)$ be a graph where V is a compact, Hausdorff, and separable measure space with a Radon measure that is nonzero on open sets. Let $\vartheta'(G)$ be the optimal value of

$$\begin{aligned} \min\ & \lambda \\ & Z(x,x) \le \lambda \quad \text{for all } x \in V, \\ & Z(x,y) \le 0 \quad \text{if } x \ne y \text{ and } xy \notin E, \\ & Z\colon V \times V \to \mathbb{R} \text{ is a continuous and symmetric kernel} \\ & \qquad \text{and } Z - J \text{ is positive}, \end{aligned} \tag{7}$$

where $J\colon V \times V \to \mathbb{R}$ is the constant 1 kernel. Note that, if V is a finite set, then positive kernels are just matrices and the problem above becomes (4).

Theorem 4. *If (Z, λ) is a feasible solution of (7), then $\alpha(G) \le \lambda$. In particular, $\alpha(G) \le \vartheta'(G)$.*

Proof. The assumptions made on V allow us to use the following observation of Bochner [4]: a continuous and symmetric kernel $K\colon V \times V \to \mathbb{R}$ is positive if and only if $\big(K(x_i, x_j)\big)_{i,j=1}^{N}$ is positive semidefinite for every choice $x_1, \ldots, x_N$ of finitely many points in V.

Then the proof in §5 follows through. Indeed, if (Z, λ) is feasible for (7) and $C \subseteq V$ is a nonempty independent set, then

$$0 \le \sum_{x,y \in C} (Z - J)(x,y) = \sum_{x,y \in C} Z(x,y) - |C|^2 \le \lambda |C| - |C|^2,$$

whence $|C| \le \lambda$ as needed. ∎

Note in particular that if (7) is feasible then $\alpha(G)$ is finite.

This theorem applies directly to graph $G(n, \theta)$, since $V = S^{n-1}$ together with the surface measure satisfies all requirements above; this application is the topic of the next section. Since $\mathbb{R}^n$ is noncompact, we cannot apply the theorem directly to $G_{\mathsf{T}}(n, \mathcal{K})$; to do so we will need a compactification step (see §9).

For our extension of ϑ' we used the dual formulation (4). One may define a primal problem of which (7) is the dual using a general theory of convex optimization (see e.g. the book by Barvinok [3]), by optimizing over measures on $V \times V$ instead of kernels.

§8. An upper bound for $A(n, \theta)$

The sphere S^{n-1} is a compact Hausdorff and separable space and the surface measure ω over S^{n-1} is a Radon measure that is nonzero on open sets. So Theorem 4 holds for $V = S^{n-1}$ and $\vartheta'(G(n, \theta))$ as defined in (7) is an upper bound for $\alpha(G(n, \theta)) = A(n, \theta)$. The question is however: how can $\vartheta'(G(n, \theta))$ be computed? Or, simpler still, how can we find a feasible solution of (7)?

As in §6, the answer is to exploit the symmetry of $G(n, \theta)$. The orthogonal group $\mathrm{O}(n) = \{\, A \in \mathbb{R}^{n \times n} : A^{\mathsf{T}} A = I \,\}$ acts on S^{n-1}, the action of $A \in \mathrm{O}(n)$ taking x to Ax, and this action preserves the inner product: for all $x, y \in S^{n-1}$ we have $Ax \cdot Ay = x \cdot y$. Then every $A \in \mathrm{O}(n)$ gives an automorphism of $G(n, \theta)$.

$\mathrm{O}(n)$ acts on kernels $K\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$: for $A \in \mathrm{O}(n)$ we have

$$(A \cdot K)(x, y) = K(A^{-1}x, A^{-1}y) \qquad \text{for all } x, y \in S^{n-1}.$$

If $A \cdot K = K$ for all $A \in \mathrm{O}(n)$, then K is *invariant*. The key idea is then to restrict ourselves in (7) to invariant kernels. We still get an upper bound to $A(n, \theta)$ (actually, as in the case of binary codes in §6, this restriction does not worsen the bound given by (7); see below), but now we are able to use the following characterization of positive and invariant kernels due to Schoenberg [30] (cf. Theorem 3 above):

Theorem 5. *A kernel $K\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ is continuous, positive, and invariant if and only if*

$$K(x,y) = \sum_{k=0}^{\infty} f_k P_k^n(x \cdot y) \tag{8}$$

for some nonnegative numbers f_0, f_1, ... such that $\sum_{k=0}^{\infty} f_k$ converges, in which case the series in (8) converges absolutely and uniformly over $S^{n-1} \times S^{n-1}$. The numbers f_k are uniquely determined by, and uniquely determine, K.

Here P_k^n is the Jacobi polynomial of degree k and parameters $\alpha = \beta = (n-3)/2$ (see, e.g., the book by Szegö [37]). These univariate polynomials can be obtained from the sequence of polynomials $1, t, t^2, \ldots$ by applying the Gram-Schmidt orthogonalization process with respect to the inner product

$$(\varphi, \psi)_n = \int_{-1}^{1} \varphi(t)\psi(t)(1-t^2)^{(n-3)/2}\, dt$$

for $\varphi, \psi\colon [-1,1] \to \mathbb{R}$. We shall normalize the polynomials P_k^n so that $P_k^n(1) = 1$. The Jacobi polynomials can be easily computed with a recurrence; see (4.5.1) in the book by Szegö [37].

Let us now use the theorem above to rewrite (7). Since P_0^n is the constant 1 polynomial, a symmetric kernel $Z\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ is continuous and such that $Z - J$ is positive if and only if

$$Z(x,y) = \sum_{k=0}^{\infty} f_k P_k^n(x \cdot y) \tag{9}$$

with $f_0 \geq 1$ and $f_1, f_2, \ldots$ nonnegative such that $\sum_{k=0}^{\infty} f_k$ converges. Recall also that $x, y \in S^{n-1}$ are adjacent in $G(n,\theta)$ if $\cos\theta < x \cdot y < 1$. Then (7) can be rewritten, giving us the problem

$$\begin{aligned} \min\ & \textstyle\sum_{k=0}^{\infty} f_k \\ & \textstyle\sum_{k=0}^{\infty} f_k P_k^n(t) \leq 0 \quad \text{if } -1 \leq t \leq \cos\theta, \\ & f_0 \geq 1 \text{ and } f_1, f_2, \ldots \text{ nonnegative with } \textstyle\sum_{k=0}^{\infty} f_k < \infty. \end{aligned} \tag{10}$$

This is a linear programming problem with infinitely many variables (one for each $k \geq 0$) and infinitely many constraints (one for each $t \in [-1, \cos\theta]$). In practice, to find a feasible solution of (10), and therefore an upper bound for $A(n,\theta)$, we have to turn this problem into a finite linear program.

The first step is easy: fix some $d > 0$ and set $f_k = 0$ if $k > d$, that is, truncate expansion (9). We get a linear program with finitely many variables and the bound thus obtained is still valid.

Dealing with the infinitely many constraints is trickier. One approach is to use *sampling*: choose a finite set $S \subseteq [-1, \cos\theta]$ and consider only the constraints in (10) for $t \in S$. If S is a fine enough sample, one would hope that the optimal solution of the resulting linear program would also be feasible for the original, but that has to be checked. In any case, if the finite sample S is not enough to produce a feasible solution, but produces an almost feasible one, by playing with the variable f_0 and rescaling the solution we may obtain a feasible solution to the original problem, at the cost of loosing a bit of quality. It is also possible to use semidefinite programming to deal with the constraints of (10) in such a way as to avoid sampling; we will see how in §12.

It is an excellent exercise to compute a table of bounds given by (10) for the kissing number $\tau_n = A(n, \pi/3)$. For $n = 3$, taking $d = 13$, we get the bound $13.15833\ldots$, proving that $\tau_3 \leq 13$. Compare this with the volume bound of §2, which only shows that $\tau_3 \leq 14$.

To obtain (10) from (7) we restricted ourselves to invariant kernels, and so the optimal value of (10) could be strictly worse than that of (7), which is $\vartheta'(G(n,\theta))$. This is however not the case. The argument runs similar to that in §6. Let μ be the Haar measure over $\mathrm{O}(n)$, normalized so that $\mu(\mathrm{O}(n)) = 1$ (this is the Radon measure invariant under $\mathrm{O}(n)$; for more information see e.g. the book by Mattila [23]). If (Z, λ) is feasible for (7), then for every $A \in \mathrm{O}(n)$ we have that $(A \cdot Z, \lambda)$ is also feasible for (7). But then the kernel

$$\overline{Z}(x,y) = \int_{\mathrm{O}(n)} Z(A^{-1}x, A^{-1}y)\, d\mu(A)$$

is invariant and such that $(\overline{Z}, \lambda)$ is feasible for (7).

The upper bound for $A(n,\theta)$ given by (10) is called the *linear programming bound*. It was introduced by Delsarte, Goethals, and Seidel [12] and quickly became a fundamental tool in the study of spherical codes. For instance, it was used by Odlyzko and Sloane [26] to prove that $\tau_8 = 240$ and $\tau_{24} = 196560$; Musin [25] developed a stronger version of the bound to show that $\tau_4 = 24$, and Bachoc and Vallentin [2] strengthened the bound further, via semidefinite programming, to provide some of the best upper bounds for $A(n,\theta)$ currently known. Finally, the relation between (10) and (7), which mimics the relation between Delsarte's linear programming bound for $A(n,d)$ and $\vartheta'(G(n,d))$, was observed by Bachoc, Nebe, Oliveira, and Vallentin [1].

§9. Periodic packings and compactification

The bound given by problem (7) applies directly to $G(n,\theta)$, but not to $G_{\mathsf{T}}(n,\mathcal{K})$ since $\mathbb{R}^n$ is not compact. A simple compactification trick will suffice to put us back on track, however. In order to describe the trick, we have to give a rigorous definition for the notion of density, which was informally introduced in §2.

Let $A \subseteq \mathbb{R}^n$ be a measurable set. We say that A has *density* $\delta(A)$ if for every $p \in \mathbb{R}^n$ we have that

$$\delta(A) = \lim_{T\to\infty} \frac{\operatorname{vol} A \cap (p + [-T,T]^n)}{\operatorname{vol}[-T,T]^n},$$

where $\operatorname{vol} X$ is the Lebesgue measure of $X \subseteq \mathbb{R}^n$. (In particular, the limit above must exist for every $p \in \mathbb{R}^n$.)

Not every set has a density, but every measurable set $A \subseteq \mathbb{R}^n$ has an *upper density*

$$\overline{\delta}(A) = \sup_{p\in\mathbb{R}^n} \limsup_{T\to\infty} \frac{\operatorname{vol} A \cap (p + [-T,T]^n)}{\operatorname{vol}[-T,T]^n}.$$

So for a convex body $\mathcal{K} \subseteq \mathbb{R}^n$ we may define

$$\Delta_{\mathsf{T}}(n,\mathcal{K}) = \sup\{\, \overline{\delta}(P) : P \text{ a translational packing of } \mathcal{K} \,\} \quad \text{and}$$
$$\Delta_{\mathsf{C}}(n,\mathcal{K}) = \sup\{\, \overline{\delta}(P) : P \text{ a congruent packing of } \mathcal{K} \,\}.$$

Both these parameters remain unchanged if we restrict ourselves to packings with density. Even more: they remain unchanged if we restrict ourselves to periodic packings.

A set $A \subseteq \mathbb{R}^n$ is *periodic* if there is a lattice $\Lambda \subseteq \mathbb{R}^n$ that leaves A invariant, i.e., $x + A = A$ for all $x \in \Lambda$. Such a lattice Λ is a *periodicity lattice* of A. Such a periodic set A repeats itself in translated copies of a *fundamental region*

$$F = \{\, \alpha_1 u_1 + \cdots + \alpha_n u_n : 0 \le \alpha_i < 1 \text{ for } i = 1, \ldots, n \,\}$$

of Λ, where $u_1, \ldots, u_n$ is a basis of Λ (see Figure 3). If A is measurable, it has a density and

$$\delta(A) = \frac{\operatorname{vol} A \cap F}{\operatorname{vol} F}.$$

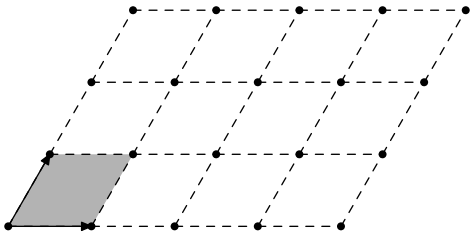

FIGURE 3. The hexagonal lattice in $\mathbb{R}^2$, generated by the two vectors shown, and a fundamental region in gray.

Suppose now P is a packing, either translational or congruent, of $\mathcal{K}$. Given $\varepsilon > 0$, there are $p \in \mathbb{R}^n$ and T (which we may pick arbitrarily large) such that

$$\left| \overline{\delta}(P) - \frac{\operatorname{vol} P \cap (p + [-T,T]^n)}{\operatorname{vol}[-T,T]^n} \right| < \frac{\varepsilon}{2}. \tag{11}$$

Let Q be the union of all copies of $\mathcal{K}$ in P that intersect $(p+[-T+d, T-d]^n)$, where $d = \operatorname{diam} \mathcal{K}$. By picking T large enough, the ratio $\operatorname{vol}[-T+d, T-d]^n / \operatorname{vol}[-T,T]^n$ can be made as close to 1 as wanted, and so we may assume that

$$\left| \frac{\operatorname{vol} P \cap (p + [-T,T]^n)}{\operatorname{vol}[-T,T]^n} - \frac{\operatorname{vol} Q}{\operatorname{vol}[-T,T]^n} \right| < \frac{\varepsilon}{2}. \tag{12).}$$

Now, let P' be the packing obtained by pasting copies of Q in a periodic fashion, that is, set

$$P' = \bigcup_{x \in 2T\mathbb{Z}^n} x + Q.$$

This is indeed a packing, since we erased a border of width d when making Q. Moreover, this is a periodic packing with periodicity lattice $2T\mathbb{Z}^n$, and from (11) and (12) we get $|\,\overline{\delta}(P) - \delta(P')| < \varepsilon$.

So to compute $\Delta_{\mathsf{T}}(n, \mathcal{K})$ or $\Delta_{\mathsf{C}}(n, \mathcal{K})$ we may restrict ourselves to periodic packings. This allows us to work with a packing graph having a compact vertex set. Indeed, let $\Lambda \subseteq \mathbb{R}^n$ be a lattice and consider the graph $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$ whose vertex set is $\mathbb{R}^n/\Lambda$ and in which distinct vertices $x, y \in \mathbb{R}^n/\Lambda$ are adjacent if there is $v \in \Lambda$ such that $(x + v + \mathcal{K}^\circ) \cap (y + \mathcal{K}^\circ) \neq \emptyset$. If $C \subseteq \mathbb{R}^n/\Lambda$ is independent in $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$, then

$$\bigcup_{v \in \Lambda} \bigcup_{x \in C} x + v + \mathcal{K}$$

is a packing of translates of $\mathcal{K}$ that is periodic with periodicity lattice Λ. Conversely, periodic translational packings of $\mathcal{K}$ with periodicity lattice Λ correspond to independent sets in $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$, and hence the maximum density of a periodic translational packing of $\mathcal{K}$ with periodicity lattice Λ is

$$\frac{\alpha(G_{\mathsf{T}}(n, \Lambda, \mathcal{K})) \operatorname{vol} \mathcal{K}}{\operatorname{vol}(\mathbb{R}^n/\Lambda).}$$

Graph $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$ has a nice compact vertex set and therefore the machinery of §7 applies: $\vartheta'(G_{\mathsf{T}}(n, \Lambda, \mathcal{K}))$ as defined by (7) provides an upper bound to $\alpha(G_{\mathsf{T}}(n, \Lambda, \mathcal{K}))$. The only problem is that we do not know the periodicity lattice Λ in advance — usually, we are interested in considering lattices with larger and larger minimal vectors. A way around this issue is the topic of the next section.

§10. An upper bound for $\Delta_{\mathsf{T}}(n, \mathcal{K})$

Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a convex body. To get an upper bound for $\Delta_{\mathsf{T}}(n, \mathcal{K})$, the idea is to define an optimization problem from any feasible solution of which we will be able to derive a feasible solution of (7) for $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$ for any lattice $\Lambda \subseteq \mathbb{R}^n$ with large enough minimal vectors.

To describe this optimization problem we need some notions from the theory of harmonic analysis on $\mathbb{R}^n$. What we need is little, however; more background can be found in the book by Katznelson [19].

In this section we deal with complex-valued functions. A function $f\colon \mathbb{R}^n \to \mathbb{C}$ is *rapidly decreasing* if it is infinitely differentiable and any of its derivatives, multiplied by any polynomial on the variables $x_1, \ldots, x_n$, vanishes at infinity.

Let $f \in L^1(\mathbb{R}^n)$. The *Fourier transform* of f computed at $u \in \mathbb{R}^n$ is

$$\widehat{f}(u) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i u \cdot x}\, dx.$$

If f is rapidly decreasing, then its Fourier transform is also rapidly decreasing and we have the *inversion formula*

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(u) e^{2\pi i u \cdot x}\, du.$$

A function $f \in L^\infty(\mathbb{R}^n)$ is of *positive type* if $f(x) = \overline{f(-x)}$ for all $x \in \mathbb{R}^n$ and for all $\rho \in L^1(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)\rho(x)\overline{\rho(y)}\, dxdy \geq 0. \tag{13}$$

A function $f \in L^1(\mathbb{R}^n)$ with $f(x) = \overline{f(-x)}$ for all $x \in \mathbb{R}^n$ is of positive type if and only if its Fourier transform is nonnegative. It is easy to give a proof of this for rapidly decreasing functions by using the inversion formula. Indeed, let $f\colon \mathbb{R}^n \to \mathbb{C}$ be rapidly decreasing and such that $f(x) = \overline{f(-x)}$ for all $x \in \mathbb{R}^n$. Then its Fourier transform is real-valued; given $\rho \in L^1(\mathbb{R}^n)$, use the inversion formula to substitute $f(x-y)$ in (13) and exchange integrals to get

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)\rho(x)\overline{\rho(y)}\, dxdy = \int_{\mathbb{R}^n} \widehat{f}(u)|\widehat{\rho}(u)|^2\, du. \tag{14}$$

So, if $\widehat{f}(u) \geq 0$ for all $u \in \mathbb{R}^n$, then (14) is nonnegative for all $\rho \in L^1(\mathbb{R}^n)$. On the other hand, if $\widehat{f}(u) < 0$ for some $x \in \mathbb{R}^n$, then we may choose a $\rho \in L^1(\mathbb{R}^n)$ such that (14) is negative.

The optimization problem promised at the beginning of the section is:

$$\begin{aligned} \min\ & f(0) \\ & f(x) \leq 0 \qquad \text{if } \mathcal{K}^\circ \cap (x + \mathcal{K}^\circ) = \emptyset, \\ & \widehat{f}(0) \geq \operatorname{vol} \mathcal{K}, \\ & f\colon \mathbb{R}^n \to \mathbb{R} \text{ is rapidly decreasing and of positive type.} \end{aligned} \tag{15}$$

Theorem 6. *If f is any feasible solution of (15), then $\Delta_{\mathsf{T}}(n, \mathcal{K}) \leq f(0)$.*

Proof. Let f be a feasible solution of (15) and take a lattice $\Lambda \subseteq \mathbb{R}^n$ whose minimal vectors have length at least as large as the diameter of $\mathcal{K}$. Then for $x, y \in \mathbb{R}^n/\Lambda$ set

$$Z(x, y) = \frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)}{\operatorname{vol} \mathcal{K}} \sum_{v \in \Lambda} f(x - y + v).$$

We show now that (Z, λ) with $\lambda = \operatorname{vol}(\mathbb{R}^n/\Lambda) f(0)/\operatorname{vol}\mathcal{K}$ is a feasible solution of (7) for the graph $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$.

First, since f is rapidly decreasing, the infinite sum above converges for all x, y and the resulting kernel is continuous.

Next, since the minimal vectors of Λ have length at least as large as the diameter of $\mathcal{K}$, if $v \in \Lambda$ is nonzero then $\mathcal{K}^\circ \cap (v + \mathcal{K}^\circ) = \emptyset$, and hence $f(v) \le 0$. But then $Z(x, x) \le \lambda$ for all $x \in \mathbb{R}^n/\Lambda$.

Now, if distinct x, $y \in \mathbb{R}^n/\Lambda$ are nonadjacent in $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$, then for all $v \in \Lambda$ we have that $(x + v + \mathcal{K}^\circ) \cap (y + \mathcal{K}^\circ) = \emptyset$, and this happens if and only if $\mathcal{K}^\circ \cap (x - y + v + \mathcal{K}^\circ) = \emptyset$. So $f(x - y + v) \le 0$ for all $v \in \Lambda$, and $Z(x, y) \le 0$.

It remains to show that $Z - J$ is positive. To this end we will show that all eigenvalues of $Z - J$ are nonnegative. The *dual lattice* of Λ is

$$\Lambda^* = \{\, u \in \mathbb{R}^n : u \cdot v \in \mathbb{Z} \text{ for all } v \in \Lambda \,\}.$$

The functions $x \mapsto e^{2\pi i u \cdot x}$, for $u \in \Lambda^*$, are periodic with periodicity lattice Λ: for $v \in \Lambda$ and all $x \in \mathbb{R}^n$ we have $e^{2\pi i u \cdot (x+v)} = e^{2\pi i u \cdot x}$. So these are functions in $L^2(\mathbb{R}^n/\Lambda)$; they actually form a complete orthonormal system of $L^2(\mathbb{R}^n/\Lambda)$ with respect to the inner product

$$(\varphi, \psi) = \frac{1}{\operatorname{vol}(\mathbb{R}^n/\Lambda)} \int_{\mathbb{R}^n/\Lambda} \varphi(x)\overline{\psi(x)}\, dx$$

for φ, $\psi \in L^2(\mathbb{R}^n/\Lambda)$.

Function $x \mapsto e^{2\pi i u \cdot x}$ is an eigenfunction of Z with eigenvalue

$$\frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)\widehat{f}(u)}{\operatorname{vol}\mathcal{K}}.$$

Indeed,

$$\int_{\mathbb{R}^n/\Lambda} Z(x, y) e^{2\pi i u \cdot y}\, dy = \frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)}{\operatorname{vol}\mathcal{K}} \int_{\mathbb{R}^n/\Lambda} \sum_{v \in \Lambda} f(x - y + v) e^{2\pi i u \cdot y}\, dy.$$

Exchange the integral with the sum above and note that, since the exponential is periodic with periodicity lattice Λ, the result is the same as integrating f over $\mathbb{R}^n$:

$$\begin{aligned}
\int_{\mathbb{R}^n/\Lambda} Z(x, y) e^{2\pi i u \cdot y}\, dy &= \frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)}{\operatorname{vol}\mathcal{K}} \int_{\mathbb{R}^n} f(x - y) e^{2\pi i u \cdot y}\, dy \\
&= \frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)}{\operatorname{vol}\mathcal{K}} \int_{\mathbb{R}^n} f(y) e^{2\pi i u \cdot (x-y)}\, dy \\
&= \frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)\widehat{f}(u)}{\operatorname{vol}\mathcal{K}} e^{2\pi i u \cdot x},
\end{aligned}$$

as we wanted.

Now, the constant 1 kernel J has only one eigenfunction with nonzero eigenvalue, namely the constant 1 function, whose eigenvalue is $\operatorname{vol}(\mathbb{R}^n/\Lambda)$. The constant 1 function is also an eigenfunction of Z (take $u = 0$ in $x \mapsto e^{2\pi i u \cdot x}$) with eigenvalue

$$\frac{\operatorname{vol}(\mathbb{R}^n/\Lambda)\widehat{f}(0)}{\operatorname{vol}\mathcal{K}} \ge \operatorname{vol}(\mathbb{R}^n/\Lambda),$$

since $\widehat{f}(0) \ge \operatorname{vol}\mathcal{K}$.

So we see that all eigenvalues of $Z - J$ are nonnegative: the eigenvalue of the constant 1 function is nonnegative, and for all other $u \in \Lambda^*$ the eigenvalue of $x \mapsto e^{2\pi i u \cdot x}$ is nonnegative since $\widehat{f}(u) \ge 0$ as f is of positive type.

So (Z, λ) is a feasible solution of (7) for the graph $G_{\mathsf{T}}(n, \Lambda, \mathcal{K})$. Hence from Theorem 4 it follows that $\alpha(G_{\mathsf{T}}(n, \Lambda, \mathcal{K})) \leq \lambda$. Recall from the previous section that the maximum density of a periodic packing of translates of $\mathcal{K}$ is

$$\frac{\alpha(G_{\mathsf{T}}(n, \Lambda, \mathcal{K})) \operatorname{vol} \mathcal{K}}{\operatorname{vol}(\mathbb{R}^n/\Lambda)} \leq \frac{\lambda \operatorname{vol} \mathcal{K}}{\operatorname{vol}(\mathbb{R}^n/\Lambda)} = f(0).$$

Since Λ is any lattice with large enough minimal vectors, the theorem follows. ∎

If we rewrite f using the inversion formula, then (15) becomes a linear programming problem with infinitely many variables (one variable $\widehat{f}(u)$ for each $u \in \mathbb{R}^n$) and infinitely many constraints. The bound given by (15) is commonly called the *linear programming bound* (compare with the linear programming bound for binary codes of §6 or spherical codes of §8). We now require f to be nonpositive in a noncompact set, and this makes it much harder to apply sampling and linear programming as could be done for the linear programming bound for spherical codes in §8. One way to find feasible solutions of (15) is to use polynomial optimization methods with sums of squares. The essential theory is presented in §11, then applied to the bound for spherical codes in §12 and to (15) in §13.

Theorem 6 was originally proven by Cohn and Elkies [7] and then used to find upper bounds for the maximum density of sphere packings. Their approach to finding feasible solutions of (15) is however different from the one in §13.

§11. Sum-of-squares polynomials and semidefinite programming

If a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ can be written as a sum of squares, that is, if there are polynomials $q_1, \ldots, q_m$ such that $p = q_1^2 + \cdots + q_m^2$, then p is nonnegative everywhere. Having a sum-of-squares decomposition is a sufficient condition for nonnegativity, but it is not in general necessary: Hilbert [18] showed that nonnegativity is equivalent to having a sum-of-squares decomposition only for (i) univariate polynomials, (ii) polynomials of degree two, and (iii) two-variable polynomials of degree 4. (Exercise: show that a univariate polynomial is nonnegative if and only if it is the sum of at most two squares.)

So requiring a polynomial to be a sum of squares is usually more restrictive than requiring it to be nonnegative. It is however an NP-complete problem to decide whether a given polynomial is nonnegative (prove it), while deciding whether a given polynomial is a sum of squares is a more tractable computational problem in practice, since it can be reduced to the problem of deciding whether a certain semidefinite programming problem is feasible. Though the complexity of this latter problem is still open, it can be solved in practice by most solvers in many cases of practical interest.

For $d \geq 0$, let $\mathbb{R}[x_1, \ldots, x_n]_{\leq d}$ denote the space of polynomials of degree at most d. If B is any finite set of polynomials, let $v_B \colon B \to B$ be such that $v_B(p) = p$ for all $p \in B$. The connection between sums of squares and semidefinite programming is established by the following theorem.

Theorem 7. *Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial of degree $2d$ and let B be a basis of $\mathbb{R}[x_1, \ldots, x_n]_{\leq d}$. Then p is a sum of squares if and only if there is a positive semidefinite matrix $Q \colon B \times B \to \mathbb{R}$ such that $p = v_B^{\mathsf{T}} Q v_B$.*

Proof. Say there is a positive semidefinite matrix Q such that $p = v_B^{\mathsf{T}} Q v_B$. Then we have $Q = u_1 u_1^{\mathsf{T}} + \cdots + u_m u_m^{\mathsf{T}}$ for some vectors $u_i \colon B \to \mathbb{R}$ and writing $q_i = u_i^{\mathsf{T}} v_B$ we have that each q_i is a polynomial and

$$p = v_B^{\mathsf{T}} Q v_B = v_B^{\mathsf{T}} u_1 u_1^{\mathsf{T}} v_B + \cdots + v_B^{\mathsf{T}} u_m u_m^{\mathsf{T}} v_B = q_1^2 + \cdots + q_m^2,$$

hence p is a sum of squares.

Conversely, say $p = q_1^2 + \cdots + q_m^2$ for some $q_1, \ldots, q_m$. Each q_i has degree at most d and thus can be expressed as a linear combination of the polynomials in B. So for $i = 1, \ldots, m$ let $u_i : B \to \mathbb{R}$ be such that $q_i = u_i^\mathsf{T} v_B$. Then $p = v_B^\mathsf{T} Q v_B$, with $Q = u_1 u_1^\mathsf{T} + \cdots + u_m u_m^\mathsf{T}$ positive semidefinite. ∎

Given a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $2d$, how can we set up a semidefinite programming problem that is feasible if and only if p is a sum of squares? We start by picking a basis B of $\mathbb{R}[x_1, \ldots, x_n]_{\leq d}$. Then from Theorem 7 we know that p is a sum of squares if and only if there is a positive semidefinite matrix $Q : B \times B \to \mathbb{R}$ such that

$$p = v_B^\mathsf{T} Q v_B = \langle v_B v_B^\mathsf{T}, Q \rangle.$$

This identity is not a linear constraint on the entries of Q, but an identity between polynomials. Both sides of the identity are polynomials of degree $2d$, so to check the identity we have to expand both sides on a basis $B_=$ of $\mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ and compare coefficients.

More precisely, for $q \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ and $r \in B_=$, let $\operatorname{coeff}(r, q)$ be the coefficient of r in the expansion of q on the basis $B_=$. For $r \in B_=$, write $\operatorname{coeff}(r, v_B v_B^\mathsf{T})$ for the real matrix obtained from $v_B v_B^\mathsf{T}$ by applying $\operatorname{coeff}(r, \cdot)$ entrywise. Then $p = \langle v_B v_B^\mathsf{T}, Q \rangle$ if and only if

$$\langle \operatorname{coeff}(r, v_B v_B^\mathsf{T}), Q \rangle = \operatorname{coeff}(r, p) \qquad \text{for each } r \in B_=. \tag{16}$$

So p is a sum of squares if and only if there is a positive semidefinite matrix satisfying the $|B_=|$ linear constraints above.

Let us work out an example in detail. Say we are given the polynomial $p(x) = x^8 + 2x^6 + 3x^4 + 4x^3 + 4$ and we are asked to find its global minimum

$$\begin{aligned} \min\{\, p(x) : x \in \mathbb{R} \,\} &= \max\{\, \lambda : p(x) - \lambda \geq 0 \text{ for all } x \in \mathbb{R} \,\} \\ &= \max\{\, \lambda : p(x) - \lambda \text{ is a sum of squares} \,\}. \end{aligned}$$

(The last identity follows from the fact that nonnegativity is equivalent to being a sum of squares for univariate polynomials.)

Since p has degree 8, we take $B = \{1, x, x^2, x^3, x^4\}$. Then $v_B^\mathsf{T} = (1, x, x^2, x^3, x^4)$ and

$$v_B v_B^\mathsf{T} = \begin{pmatrix} 1 & x & x^2 & x^3 & x^4 \\ x & x^2 & x^3 & x^4 & x^5 \\ x^2 & x^3 & x^4 & x^5 & x^6 \\ x^3 & x^4 & x^5 & x^6 & x^7 \\ x^4 & x^5 & x^6 & x^7 & x^8 \end{pmatrix}.$$

Take $B_= = \{1, x, \ldots, x^8\}$ and write $F_k = \operatorname{coeff}(x^k, v_B v_B^\mathsf{T})$ for $k = 0, \ldots, 8$. Here are, for instance, F_0 and F_4:

$$F_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad F_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then for $\lambda \in \mathbb{R}$ we have that $p(x) - \lambda$ is a sum of squares if and only if there is a positive semidefinite matrix $Q : B \times B \to \mathbb{R}$ satisfying

$$\begin{array}{lll} \langle F_0, Q \rangle = 4 - \lambda, & \langle F_3, Q \rangle = 4, & \langle F_6, Q \rangle = 2, \\ \langle F_1, Q \rangle = 0, & \langle F_4, Q \rangle = 3, & \langle F_7, Q \rangle = 0, \\ \langle F_2, Q \rangle = 0, & \langle F_5, Q \rangle = 0, & \langle F_8, Q \rangle = 1. \end{array} \tag{17}$$

So the global minimum of p is the optimal value of the semidefinite programming problem

$$\begin{aligned} &\max \lambda \\ &\quad Q \text{ and } \lambda \text{ satisfy (17)}, \\ &\quad Q\colon B \times B \to \mathbb{R} \text{ is positive semidefinite and } \lambda \in \mathbb{R}. \end{aligned}$$

(Note λ is also a variable in this problem.)

We will sometimes write semidefinite programming problems with "polynomial constraints", that is, constraints such as $p = \langle v_B v_B^{\mathsf{T}}, Q \rangle$ that are identities between two polynomials. It is then implied that to transform the problem into a true semidefinite programming problem one has to pick a basis to express the polynomial identity, as described above.

So far we have discussed the relation between sums of squares and nonnegativity everywhere, but what if we want a polynomial to be nonnegative inside a given domain?

If the domain is a *basic closed semialgebraic set*, that is, a set

$$D = \{\, x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0 \,\}$$

where $g_i \in \mathbb{R}[x_1, \ldots, x_n]$, then the theory can be easily extended. Indeed, if there are sum-of-squares polynomials q_0, q_1, …, q_m such that

$$p = q_0 + q_1 g_1 + \cdots + q_m g_m,$$

then p is nonnegative in D. The converse holds under certain conditions; see e.g. a theorem of Putinar [28].

For univariate polynomials the situation is much simpler. A classical theorem of Lukács (see Theorem 1.21.1 in the book by Szegö [37]) asserts that a polynomial $p \in \mathbb{R}[x]$ is nonnegative on the interval $[a, b]$ if and only if there are polynomials q, q' such that

$$p(x) = q(x)^2 + (x - a)(b - x)q'(x)^2$$

if p has even degree, or

$$p(x) = (x - a)q(x)^2 + (b - x)q'(x)^2$$

if p has odd degree.

It is usual in the literature to always use the standard monomial basis (i.e., the basis of the space of polynomials consisting of all monomials) for both B and $B_=$. This restriction does not make the theory any simpler and, though it is true that the standard monomial basis is used in most practical applications, sometimes picking another basis is essential for the numerical stability of the resulting problem. This will be the case in §13.

§12. Application to the linear programming bound for spherical codes

In §8 we used sampling to compute the linear programming bound (10) for spherical codes. Let us now use sums of squares to describe a semidefinite programming problem any feasible solution of which gives an upper bound to $A(n, \theta)$.

As in §8, we start by fixing $d > 0$ and truncating the sums in (10) after $k = 2d$. So we wish to solve the problem

$$\begin{aligned} &\min \textstyle\sum_{k=0}^{2d} f_k \\ &\quad \textstyle\sum_{k=0}^{2d} f_k P_k^n(t) \leq 0 \quad \text{if } -1 \leq t \leq \cos\theta, \\ &\quad f_0 \geq 1 \text{ and } f_1, \ldots, f_{2d} \geq 0. \end{aligned}$$

n	Lower bound	Upper bound	n	Lower bound	Upper bound
3	12	13.158330	14	1606	3530.606349
4	24	25.558429	15	2564	5451.677955
5	40	46.337573	16	4320	8364.000000
6	72	82.631215	17	5346	12373.384615
7	126	140.162445	18	7398	18199.285714
8	240	240.000000	19	10668	26771.000000
9	306	380.099072	20	17400	39655.000000
10	500	595.828789	21	27720	59693.117647
11	582	915.389530	22	49896	88391.875000
12	840	1416.090277	23	93150	130340.050063
13	1154	2234.378143	24	196560	196560.000000

TABLE 4. Table of lower and upper bounds for the kissing number. The lower bounds come from Table 1.5 in Conway and Sloane [8], except for $n = 13$ and 14, in which case they were provided by Zinoviev and Ericson [40]. The upper bounds were computed by solving problem (18) numerically with $d = 11$; the standard monomial basis was used throughout. My gratitude goes to Fabrício Caluza Machado for computing the table above.

Notice that $\sum_{k=0}^{2d} f_k P_k^n(t)$ is a univariate polynomial on t. Therefore, it is nonpositive on $[-1, \cos\theta]$ if and only if there are sum-of-squares polynomials q, q' such that

$$\sum_{k=0}^{2d} f_k P_k^n(t) = -q(t) - (t+1)(\cos\theta - t)q'(t).$$

Here, q has degree up to $2d$ and q' has degree up to $2(d-1)$.

Let B and B' be bases of $\mathbb{R}[t]_{\le d}$ and $\mathbb{R}[t]_{\le d-1}$ respectively. Then q is a sum of squares if and only if there is a positive semidefinite matrix $R\colon B \times B \to \mathbb{R}$ such that $q = \langle v_B v_B^\mathsf{T}, R\rangle$. Similarly, q' is a sum of squares if and only if there is a positive semidefinite matrix $R'\colon B' \times B' \to \mathbb{R}$ such that $q' = \langle v_{B'} v_{B'}^\mathsf{T}, R'\rangle$. So any feasible solution of the semidefinite programming problem with polynomial constraints

$$\begin{aligned} \min\ & \textstyle\sum_{k=0}^{2d} f_k \\ & \textstyle\sum_{k=0}^{2d} f_k P_k^n(t) + \langle v_B v_B^\mathsf{T}, R\rangle + \langle (t+1)(\cos\theta - t) v_{B'} v_{B'}^\mathsf{T}, R'\rangle = 0, \\ & f_0 \ge 1 \text{ and } f_1, \ldots, f_{2d} \ge 0, \\ & R\colon B \times B \to \mathbb{R} \text{ and } R'\colon B' \times B' \to \mathbb{R} \text{ positive semidefinite} \end{aligned} \tag{18}$$

gives an upper bound to $A(n, \theta)$.

Table 4 shows bounds for the kissing number obtained by solving the problem above for different values of n. It is a good exercise to try to duplicate these results.

§13. Application to the linear programming bound for sphere packings

Let us now use sums of squares to describe a semidefinite programming problem any feasible solution of which provides an upper bound for $\Delta_\mathsf{T}(n, \mathcal{K})$, where $\mathcal{K} \subseteq \mathbb{R}^n$ is a ball of radius $1/2$. This process is similar to that described in the previous section, but slightly more involved.

Radial functions. If $\mathcal{K}$ is a ball of radius $1/2$, then $\mathcal{K}^\circ \cap (x + \mathcal{K}^\circ) = \emptyset$ if and only if $\|x\| \ge 1$. Hence the nonpositivity constraint on f in (15) becomes

$$f(x) \le 0 \quad \text{if } \|x\| \ge 1.$$

To simplify (15), we restrict the choice of f to radial functions. A function $f\colon \mathbb{R}^n \to \mathbb{C}$ is *radial* if $f(x)$ depends only on $\|x\|$, that is, $f(x) = f(y)$ if $\|x\| = \|y\|$. Since $\mathcal{K}$

is a ball, this restriction does not worsen the bound given by (15). The argument is already familiar. If f is a feasible solution of (15) and $A \in \mathrm{O}(n)$ is an orthogonal matrix, then function $A \cdot f$ such that $(A \cdot f)(x) = f(A^{-1}x)$ is also a feasible solution of (15) with the same objective value. Hence if μ is the Haar measure over $\mathrm{O}(n)$ normalized so that $\mu(\mathrm{O}(n)) = 1$, then the function $\overline{f}$ such that

$$\overline{f}(x) = \int_{\mathrm{O}(n)} f(A^{-1}x)\, d\mu(A)$$

is a radial function that is a feasible solution of (15) with $\overline{f}(0) = f(0)$. (The only technical point here is to argue that $\overline{f}$ is rapidly decreasing.)

A parametrization of $\widehat{f}$, and positive typeness. Function f is radial if and only if its Fourier transform $\widehat{f}$ is radial. Moreover, since f is of positive type and real-valued, we have $f(x) = f(-x)$ for all $x \in \mathbb{R}^n$, and this holds if and only if $\widehat{f}$ is real-valued and such that $\widehat{f}(u) = \widehat{f}(-u)$ for all $u \in \mathbb{R}^n$ (prove it).

We specify f via its Fourier transform $\widehat{f}$. Fix an integer $d > 0$ and set $\widehat{f}(u) = p(\|u\|)e^{-\pi\|u\|^2}$, where p is an even univariate polynomial of degree at most $2d$. Then f is of positive type if and only if p is nonnegative, and since p is univariate this happens if and only if p is a sum of squares.

Computing f from $\widehat{f}$, and constraints on f. To compute f from $\widehat{f}$ given as above we use the following theorem.

Theorem 8. *If $k \geq 0$ and $x \in \mathbb{R}^n$, then*

$$\int_{\mathbb{R}^n} \|u\|^{2k} e^{-\pi\|u\|^2} e^{2\pi u \cdot x}\, du = k!\, \pi^{-k} L_k^{n/2-1}(\pi\|x\|^2) e^{-\pi\|x\|^2}.$$

Here, L_k^α is the Laguerre polynomial of degree k and parameter α. For a proof of this theorem, see the paper by de Laat, Oliveira, and Vallentin [21].

The Laguerre polynomials L_k^α of parameter $\alpha > -1$ are univariate polynomials that can be obtained, aside from normalization, by applying the Gram-Schmidt orthogonalization process to the polynomials 1, t, t^2, $\ldots$ with respect to the inner product

$$(\varphi, \psi)_\alpha = \int_0^\infty \varphi(t)\psi(t)t^\alpha e^{-t}\, dt$$

for φ, $\psi \colon \mathbb{R} \to \mathbb{R}$. There are simple recurrence relations that can be used to compute L_k^α; see e.g. (5.1.10) in the book by Szegö [37].

If $p(t) = \sum_{k=0}^d a_k t^{2k}$, then using Theorem 8 we see that

$$f(x) = \sum_{k=0}^d a_k k!\, \pi^{-k} L_k^{n/2-1}(\pi\|x\|^2) e^{-\pi\|x\|^2}.$$

In other words, from Theorem 8 we see how to define a linear transformation $\mathcal{F}$ on the space of univariate even polynomials so that $f(x) = \mathcal{F}(p)(\|x\|)e^{-\pi\|x\|^2}$. Note moreover that $\mathcal{F}$ preserves the degree of a polynomial to which it is applied.

Since p is univariate, we have $f(x) \leq 0$ if $\|x\| \geq 1$ if and only if there are sum-of-squares polynomials q, q' such that

$$\mathcal{F}(p)(s) = -q(s) - (s^2 - 1)q'(s).$$

Note moreover that the degree of q is at most $2d$, and similarly the degree of q' is at most $2(d-1)$. Since p is even and $s^2 - 1$ is also even, we may assume that q and q' are even.

Full formulation. In our formulation we have to ensure that p, q, and q' are even. Instead of being a burden in the form of extra constraints, this allows us to break down the matrices representing these polynomials into two blocks each.

To simplify the discussion let us assume $d > 0$ is odd. Let P_0, P_1, ... be a sequence of univariate polynomials such that P_k has degree k. Then $P_k(t^2)$ is an even polynomial of degree $2k$ and $tP_k(t^2)$ is an odd polynomial of degree $2k+1$.

Write $B_0 = \{P_0(t^2), \ldots, P_{\lfloor d/2 \rfloor}(t^2)\}$ and $B_1 = \{tP_0(t^2), \ldots, tP_{\lfloor d/2 \rfloor}(t^2)\}$ and set $B = B_0 \cup B_1$. We know that p is a sum of squares if and only if there is a positive semidefinite matrix $Q\colon B \times B \to \mathbb{R}$ such that $p = \langle v_B v_B^\mathsf{T}, Q \rangle$. Let $Q^-\colon B \times B \to \mathbb{R}$ be the matrix such that $Q^-(r,s) = Q(r,s)$ if r, $s \in B_i$ for $i = 0$, 1 and $Q^-(r,s) = -Q(r,s)$ otherwise. This matrix is positive semidefinite. Moreover

$$p(-t) = \langle (v_B v_B^\mathsf{T})(-t), Q \rangle = \langle (v_B v_B^\mathsf{T})(t), Q^- \rangle.$$

But then, since $p(t) = p(-t)$, we have

$$p = (1/2)(\langle v_B v_B^\mathsf{T}, Q \rangle + \langle v_B v_B^\mathsf{T}, Q^- \rangle) = \langle v_B v_B^\mathsf{T}, (1/2)(Q + Q^-) \rangle.$$

Now $(1/2)(Q + Q^-)$ is a positive semidefinite matrix with a zero in position (r,s) for all $(r,s) \in B_0 \times B_1 \cup B_1 \times B_0$. So this matrix gives a representation of p and is composed of two blocks corresponding to B_0 and B_1.

We may represent polynomial q similarly. As for polynomial q', it has degree at most $2(d-1)$, so we take $B'_0 = \{P_0(t^2), \ldots, P_{\lfloor d/2 \rfloor - 1}(t^2)\}$ and proceed similarly to define B'_1.

Extend the linear transformation $\mathcal{F}$ to matrices of even polynomials by applying it entrywise. For short, write $V_i = v_{B_i} v_{B_i}^\mathsf{T}$ and $V'_i = v_{B'_i} v_{B'_i}^\mathsf{T}$. Then each feasible solution of the following semidefinite programming problem with polynomial identity constraints provides an upper bound for $\Delta_\mathsf{T}(n, \mathcal{K})$ (cf. (15)):

$$\begin{aligned}
\min\ & \langle \mathcal{F}(V_0)(0), Q_0 \rangle + \langle \mathcal{F}(V_1)(0), Q_1 \rangle \\
& \langle \mathcal{F}(V_0)(s), Q_0 \rangle + \langle \mathcal{F}(V_1)(s), Q_1 \rangle + \langle V_0(s), R_0 \rangle + \langle V_1(s), R_1 \rangle \\
& \qquad + \langle (s^2-1)V'_0(s), R'_0 \rangle + \langle (s^2-1)V'_1(s), R'_1 \rangle = 0, \\
& \langle V_0(0), Q_0 \rangle + \langle V_1(0), Q_1 \rangle \geq \operatorname{vol} \mathcal{K}, \\
& Q_i,\ R_i\colon B_i \times B_i \to \mathbb{R} \text{ and } R'_i\colon B'_i \times B'_i \to \mathbb{R} \text{ are positive semidefinite.}
\end{aligned} \tag{19}$$

Choosing bases and computational results. To solve problem (19) we need to choose polynomials P_0, P_1, ... We also need to choose a basis $B_=$ of the space of even polynomials of degree up to $2d$ so as to transform (19) into a semidefinite programming problem.

We could pick the standard monomial basis both times. This is a very poor choice in practice: the resulting problems are ill-conditioned and numerical instability prevents solvers from finding solutions even for small values of d.

A better choice, arrived at by extensive guessing and experimentation (see the paper by de Laat, Oliveira, and Vallentin [21], §5.3), is to take

$$P_k(t) = \mu_k^{-1} L_k^{n/2-1}(2\pi t),$$

where μ_k is the largest absolute value of any coefficient of $L_k^{n/2-1}(2\pi t)$, and set $B_= = \{P_0(t^2), \ldots, P_d(t^2)\}$.

Table 5 shows bounds for $\Delta_\mathsf{T}(n, \mathcal{K})$ computed by solving (19) for different values of d. Again, it is a good exercise to reproduce this table. The lower bounds on the

n	Lower bound	$d = 21$	$d = 25$
2	0.28868	0.28867518	0.28867505
3	0.17678	0.18615810	0.18615285
4	0.12500	0.13126244	0.13125761
5	0.08839	0.09976847	0.09972443
6	0.07217	0.08089432	0.08084123
7	0.06250	0.06939099	0.06931809
8	0.06250	0.06273191	0.06253436
9	0.04419	0.05951452	0.05910841
10	0.03906	0.05868910	0.05033605
11	0.03516	0.06035374	0.00385678
12	0.03704	0.06549422	0.05775411
13	0.03516	0.07455412	0.00164939
14	0.03608	0.08571855	0.07182631
15	0.04419	0.10084887	0.00003525

TABLE 5. Table with lower bounds for the sphere packing density, together with upper bounds computed by solving problem (19) numerically for $d = 21$ and $d = 25$. Densities are actually the *center density*, the number of centers of unit balls per unit volume; to get the density, multiply by the volume of a unit ball. These are numerical results, not all trustworthy; see below.

table are taken from Cohn and Elkies [7]; compare this table to the table at the end of their paper.

It should be noted that the numbers on the table were computed by solving problem (19) with the CSDP [5] solver, which works with double-precision floating-point arithmetic. No rigorous verification was done on the solutions obtained, so the numbers obtained cannot be said to provide bounds. As the dimension and degree increase, the problems become more unstable. Notice for instance the result for $n = 15$ and $d = 25$: it is clearly wrong! When using numerical solvers to get bounds, the work is not over when the problem is solved. Rigorous verification of the solutions — along with the many tricks involved in tweaking an infeasible solution to turn it into a feasible one — is an essential part of the process. See for instance the paper by de Laat, Oliveira, and Vallentin [21] for one approach to rigorous verification of solutions.

§14. References

[1] C. Bachoc, G. Nebe, F.M. de Oliveira Filho, and F. Vallentin, Lower bounds for measurable chromatic numbers, *Geometric and Functional Analysis* 19 (2009) 645–661.

[2] C. Bachoc and F. Vallentin, New upper bounds for kissing numbers from semidefinite programming, *Journal of the American Mathematical Society* 21 (2008) 909–924.

[3] A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics 54, American Mathematical Society, Providence, Rhode Island, 2002.

[4] S. Bochner, Hilbert distances and positive definite functions, *Annals of Mathematics* 42 (1941) 647–656.

[5] B. Borchers, CSDP, a C library for semidefinite programming, *Optimization Methods and Software* 11/12 (1999) 613–623.

[6] E.R. Chen, M. Engel, and S.C. Glotzer, Dense crystalline dimer packings of regular tetrahedra, *Discrete & Computational Geometry* 44 (2010) 253–280.

[7] H. Cohn and N. Elkies, New upper bounds on sphere packings I, *Annals of Mathematics* 157 (2003) 689–714.

[8] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices, and Groups*, Grundlehren der mathematischen Wissenschaften 290, Springer-Verlag, New York, 1988.

[9] H.S.M. Coxeter, An upper bound for the number of equal nonoverlapping spheres that can touch another of the same size, in: *Proceedings of Symposia in Pure Mathematics, Vol. VII*, American Mathematical Society, Providence, 1963, pp. 53–71.

[10] H.S.M. Coxeter, The problem of packing a number of equal nonoverlapping circles on a sphere, *Transactions of the New York Academy of Sciences* 24 (1962) 320–331.

[11] P. Delsarte, *An Algebraic Approach to the Association Schemes of Coding Theory*, Philips Research Reports Supplements 1973 No. 10, Philips Research Laboratories, Eindoven, 1973.

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

[13] L. Fejes Tóth, *Lagerungen in der Ebene auf der Kugel und im Raum*, Grundlehren der mathematischen Wissenschaften 65, Springer-Verlag, Berlin, 1972.

[14] S. Gravel, V. Elser, and Y. Kallus, Upper bound on the packing density of regular tetrahedra and octahedra, *Discrete & Computational Geometry* 46 (2011) 799–818.

[15] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Algorithms and Combinatorics 2, Springer-Verlag, Berlin, 1988.

[16] T.C. Hales, A proof of the Kepler conjecture, *Annals of Mathematics* 162 (2005) 1065–1185.

[17] D. Hilbert, Mathematical problems, *Bullentin of the American Mathematical Society* 8 (1902) 437–479.

[18] D. Hilbert, Über die Darstellung definiter Formen als Summe von Formenquadraten, *Mathematische Annalen* 32 (1888) 342–350.

[19] Y. Katznelson, *An Introduction to Harmonic Analysis*, John Wiley & Sons, Inc., New York, 1968.

[20] J. Kepler, *Vom sechseckigen Schnee* (translated from the Latin with commentary and afterword by L. Dunsch), Hellerau-Verlag, Dresden, 2005.

[21] D. de Laat, F.M. de Oliveira Filho, and F. Vallentin, Upper bounds for packings of spheres of several radii, to appear in *Forum of Mathematics, Sigma*, 2012, 31pp., arXiv:1206.2608.

[22] L. Lovász, On the Shannon capacity of a graph, *IEEE Transactions on Information Theory* IT-25 (1979) 1–7.

[23] P. Mattila, *Geometry of Sets and Measures in Euclidean Space: Fractals and Rectifiability*, Cambridge Studies in Advanced Mathematics 44, Cambridge University Press, Cambridge, 1995.

[24] R.J. McEliece, E.R. Rodemich, and H.C. Rumsey, The Lovász bound and some generalizations, *Journal of Combinatorics, Information & System Sciences* 3 (1978) 134–152.

[25] O.R. Musin, The kissing number in four dimensions, *Annals of Mathematics* 168 (2008) 1–32.

[26] A.M. Odlyzko and N.J.A. Sloane, New bounds on the number of unit spheres that can touch a unit sphere in n dimensions, *Journal of Combinatorial Theory, Series A* 26 (1979) 210–214.

[27] F.M. de Oliveira Filho and F. Vallentin, Mathematical optimization for packing problems, *SIAG/OPT Views and News* 23 (2015) 5–14.

[28] M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana University Mathematics Journal* 42 (1993) 969–984.

[29] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original), Dover Books on Advanced Mathematics, Dover Publications, Inc., New York, 1990.

[30] I.J. Schoenberg, Metric spaces and completely monotone functions, *Annals of Mathematics* 39 (1938) 811–841.

[31] A. Schrijver, A comparison of the Delsarte and Lovász bounds, *IEEE Transactions on Information Theory* IT-25 (1979) 425–429.

[32] A. Schürmann, On packing spheres into containers, *Documenta Mathematica* 11 (2006) 393–406.

[33] K. Schütte and B.L. van der Waerden, Das Problem der dreizehn Kugeln, *Mathematische Annalen* 125 (1953) 325–334.

[34] C.E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal* 27 (1948) 379–423, 623–656.

[35] N.J.A. Sloane, Tables of sphere packings and spherical codes, *IEEE Transactions on Information Theory* IT-27 (1981) 327–338.

[36] D.J. Struik, Het probleem "de impletione loci", *Niew Archief voor Wiskunde II* 15 (1926) 121–137.

[37] G. Szegö, *Orthogonal Polynomials* (Fourth Edition), American Mathematical Society Colloquium Publications Volume XXIII, American Mathematical Society, Providence, 1975.

[38] A. Thue, Über die dichste Zusammenstellung von kongruenten Kreisen in einer Ebene, *Skripter udgivne af Videnskabs-selskabet i Christiania*, 1910, 7pp.

[39] L. Tunçel, *Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization*, Fields Institute Monographs, American Mathematical Society, Providence, 2010.

[40] V.A. Zinoviev and T. Ericson, New lower bounds for contact numbers in small dimensions, *Problems of Information Transmission* 35 (1999) 287–294.

F.M. de Oliveira Filho, Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, 05508-090 São Paulo/SP, Brazil.