New Bounds
for
Geometric Packing and Coloring
via
Harmonic Analysis and Optimization

Fernando Mario de Oliveira Filho

This research has been carried out at the Centrum Wiskunde & Informatica in Amsterdam.

This research was partially supported by CAPES/Brazil under grant BEX 2421/04-6.

Copyright © 2009 by Fernando Mario de Oliveira Filho.

Printed and bound by Ipskamp Drukkers B.V., The Netherlands.

New Bounds
for
Geometric Packing and Coloring
via
Harmonic Analysis and Optimization

ACADEMISCH PROEFSCHRIFT
ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 1 december 2009, te 10:00 uur

door

Fernando Mario de Oliveira Filho

geboren te São Paulo, Brazilië

Promotiecommissie:

Promotor:
prof. dr. A. Schrijver

Co-promotor:
dr. F. Vallentin

Overige leden:
prof. dr. C. Bachoc
prof. dr. ir. W.H. Haemers
prof. dr. T.H. Koornwinder
prof. dr. M. Laurent
prof. dr. E.M. Opdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Preface

THIS thesis is the end-product of four years of work at the Centrum Wiskunde & Informatica (which for me started as the Centrum voor Wiskunde en Informatica) in Amsterdam. It is the end-product, but by no means could it contain all the experiences I had in the past four years. So I take the opportunity now to acknowledge some who participated, directly and indirectly, in the production of this work.

First, I must thank my supervisor, Lex Schrijver. Needless to say, without him this thesis wouldn't have existed. I am thankful for his guidance in the development of this work, and for his help in all matters concerning my stay. I must here not forget to thank Monique Laurent, who also helped me in many ways during my stay.

While Lex was my supervisor, most of the work in this thesis was done jointly with Frank Vallentin. I shared the same office with Frank for most of the four years of my stay, and working with him was both motivating and fruitful. I must say that I am happy to also count him as a friend.

Most of this work is known to Christine Bachoc, and part of it was also made in collaboration with her. It was always nice to talk to Christine about mathematics, and she always had some good idea to offer. But somehow what keeps coming back to my mind is a certain night at Muiderpoort in which we tested the resistance of a HEMA lock together with Frank, Nebojša, and Hartwig.

My work was certainly made much easier by the staff of CWI. In particular I will remember Susanne van Dam and Bikkie Aldeias. The library staff was also always very helpful, and I thank them for their assistance.

In the past four years I have been blessed with many friends, and they have helped make the last four years of my life so much more enjoyable. So I would like to thank all my friends, in particular some of my friends at CWI: Alexandra Silva, Dion Gijswijt, Erik Jan van Leeuwen, Inken Wohlers, Jana Němcová, Jarek Byrka, José Proença, Nebojša Gvozdenović, and Rodrigo Guimarães. Erik Jan was my roommate for more than three years and I will never forget our trips together.

I also couldn't forget my Brazilian friends, specially those who visited me in Amsterdam: Alex Kusano, Carlos Cardonha, Dairton Bassi, Danilo Sato, Fabio Nishi-

hara, Giuliano Mega, Luciano Soares, Marcos Broinizi, and Peter Kreslins. I will also remember Daniel Martin, specially because of a very interesting conversation we had in Berlin.

Finally, my thanks must go to all my family, who supported me for all this time, specially to my parents Irene and Fernando, my sisters Ana Paula and Ana Carolina, and my grandmother Clélia. Without their love and help I could never have done this.

Fernando Mario de Oliveira Filho
Amsterdam, September 2009

Contents

Chapter One

Introduction and preliminaries

GIVEN a finite graph, to find the maximum number of vertices which are pairwise nonadjacent is a well-known problem in combinatorial optimization. This maximum is usually called the *stability number* of the graph, and several problems can be modeled as the problem of determining the stability number of some finite graph. The stability number is related to the *chromatic number*, which is the minimum number of colors needed to color the vertices of a graph so that no two adjacent vertices are colored with the same color; determining the chromatic number is also a fundamental problem in combinatorial optimization.

Many problems in geometry can be modeled as problems of determining the stability number or chromatic number of some finite graph, but sometimes we need to consider infinite graphs. Moreover, sometimes it might be necessary to work with analogues of the stability number, since infinite graphs might have infinite stable sets. Is it useful to consider such problems in the framework of graphs? As we discuss in Section 1.1, by doing so we gain access to a range of techniques from combinatorial optimization that have been developed to compute bounds for the stability number or chromatic number of finite graphs and we may generalize these techniques to our infinite graphs as well.

In Section 1.2 we present a quick survey of the geometrical problems which motivated many of the results we present in later chapters. In Section 1.3 we give a short outline of the thesis, and finally in Section 1.4 we present some preliminaries and fix some of the notation used in the rest of the text. We usually omit references to the preliminaries during the text; if a term is used but not introduced, then its meaning can often be found in the preliminaries section.

1.1 The stability number of distance graphs

Let X be a metric space with distance function d and take $D \subseteq (0, \infty)$. Consider the following graph, denoted by $G(X, D)$: Its vertices are the points of the space X, and two vertices x, $y \in X$ are adjacent if and only if $d(x, y) \in D$. We say that this graph is a *distance graph* over the metric space (X, d) because its vertex set is X and the adjacency relation is completely characterized by distance.

Many problems of interest can be modeled as problems in distance graphs defined over natural metric spaces. We are particularly interested in the stability number of such graphs. Given a graph $G = (V, E)$, recall that a set $C \subseteq V$ is called *stable* if no two vertices in it are adjacent; the *stability number* of G, denoted by $\alpha(G)$, is the maximum size of any stable set of G.

So we consider our first example of a distance graph. Let $H_n = \{0, 1\}^n$ be the n-dimensional Hamming cube, that is, the set of all binary words of length n. The Hamming cube can be seen as a metric space if one introduces in it the *Hamming distance*: The distance between two words in H_n is just the number of bits in which they differ. Now fix a number $1 < d \leq n$ and consider the distance graph $G(H_n, \{1, \ldots, d-1\})$. A stable set in this graph is a set of words any two of which are at Hamming distance at least d from each other. Such a set is a *binary code* of *length* n and *minimum distance* at least d. The stability number of this graph is thus the maximum size of such a binary code. This is an important parameter in coding theory, and is usually denoted by $A(n, d)$.

Our next example replaces H_n by the $(n-1)$*-dimensional unit sphere*, which is the set

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

where $x \cdot y$ denotes the standard Euclidean inner product in $\mathbb{R}^n$. The unit sphere is also a metric space when one considers the *angular distance*: The distance between points $x, y \in S^{n-1}$ is given by

$$\arccos x \cdot y,$$

so distances are always in the interval $[0, \pi]$, with antipodal points being at distance π.

Now fix $0 < \theta \leq \pi$ and consider the distance graph $G(S^{n-1}, (0, \theta))$. A stable set of this graph is a set of points the distance between any two of which is at least θ. Such a set is called a *spherical code* of *minimum angular distance* at least θ. Notice that the stability number of $G(S^{n-1}, (0, \theta))$ is finite, being then the maximum cardinality of any spherical code of minimum angular distance at least θ. This is also a well-known parameter, which is usually denoted by $A(n, \theta)$. When $\theta = \pi/3$, the stability number of $G(S^{n-1}, (0, \theta))$ is known as the *kissing number* of $\mathbb{R}^n$, since it gives the maximum number of pairwise nonoverlapping unit spheres in $\mathbb{R}^n$ that can simultaneously touch a central unit sphere. For instance, the kissing number of the plane is easily seen to be 6.

As a third example we again consider a distance graph over the sphere. This time, fix a number $0 < \theta \leq \pi$ and consider the graph $G(S^{n-1}, \{\theta\})$, which from now on we denote by $G(S^{n-1}, \theta)$ for short. Stable sets in this graph can be infinite. Indeed, any spherical cap of small enough diameter is a stable set. So the stability number of $G(S^{n-1}, \theta)$ is infinite.

Consider now the surface measure ω on the sphere, which is normalized so that

$$\omega(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

We may now define a measurable version of the stability number of $G(S^{n-1}, \theta)$ by restricting ourselves to measurable stable sets. Namely, the *measurable stability number* of $G(S^{n-1}, \theta)$ is the number

$$\alpha_{\mathrm{m}}(G(S^{n-1}, \theta)) = \sup\{\, \omega(C) : C \subseteq S^{n-1} \text{ is measurable and stable}\,\}.$$

The example of the spherical cap we gave before shows that $\alpha_{\mathrm{m}}(G(S^{n-1}, \theta)) > 0$, and it is immediate that it is finite. We have then a generalization of the concept of stability number.

From the graph on the Hamming cube to the graph $G(S^{n-1}, (0, \theta))$ over the sphere we have jumped from a finite graph to an infinite graph which, however, had finite stability number. Our second jump was from the graph $G(S^{n-1}, (0, \theta))$ to the graph $G(S^{n-1}, \theta)$, which has infinite stability number, but for which we defined a measurable stability number. Now we jump from graphs over the sphere, which is a compact space with a finite measure, to graphs over the n-dimensional Euclidean space, which is not compact.

More specifically, we consider now the distance graph $G(\mathbb{R}^n, \{1\})$, which we abbreviate by $G(\mathbb{R}^n, 1)$. This is a well-known geometric graph, usually called the *unit-distance graph*. A stable set in this graph is a set of points in $\mathbb{R}^n$ no two of which are at distance 1 from each other. Many questions have been asked about this graph. For instance, determining its chromatic number, i.e., the minimum number of colors needed to color the points of $\mathbb{R}^n$ in such a way that no two points at distance 1 get the same color, is a well-known open question in geometry for $n \geq 2$, which we treat in more detail in Section 1.2.

Questions have also been asked about the stable sets of $G(\mathbb{R}^n, 1)$. Notice that they can be infinite, so that the stability number of $G(\mathbb{R}^n, 1)$ is infinite, and they can also have infinite measure. Therefore, instead of considering the measure of a measurable stable set C, we consider its *upper density*, which is given by

$$\overline{\delta}(C) = \limsup_{R \to \infty} \frac{\mathrm{vol}(C \cap [-R, R]^n)}{\mathrm{vol}[-R, R]^n},$$

where $\mathrm{vol}\, X$ simply denotes the Lebesgue measure of set X. Now we may consider the parameter

$$\overline{\alpha}_{\mathrm{m}}(G(\mathbb{R}^n, 1)) = \sup\{\, \overline{\delta}(C) : C \subseteq \mathbb{R}^n \text{ is measurable and stable}\,\},$$

which is a density analogue of the measurable stability number as defined for the graph $G(S^{n-1}, \theta)$. This parameter has also been studied in geometry, being usually denoted by $m_1(\mathbb{R}^n)$; we will discuss this further in Section 1.2.

From our four examples, it is clear that distance graphs and the different concepts of stability number we have considered provide a framework in which one may approach four different problems in the same light. But does this help us to say anything about, for instance, the stability number of $G(H_n, \{1, \ldots, d-1\})$ or about the measurable stability number of $G(S^{n-1}, \theta)$? Well, the problem of determining the stability number of a finite graph is a well-known problem in combinatorial optimization. It is known to be NP-hard, so that there is little hope that an efficient algorithm exists to compute the stability number. But methods have been proposed for the computation of lower and upper bounds for the stability number, so hopefully they can also be used for our distance graphs.

We are particularly interested here in finding upper bounds for the stability number. One upper bound that has been proposed, which provides the best known bounds in many cases, is the Lovász theta number. Introduced by Lovász [29], the *theta number* of a finite graph $G = (V, E)$, denoted by $\vartheta(G)$, is given as the optimal value of a semidefinite programming problem. Namely, $\vartheta(G)$ is equal to the maximum of

$$\sum_{u \in V} \sum_{v \in V} A(u, v)$$

taken over all matrices $A\colon V \times V \to \mathbb{R}$ whose rows and columns are indexed by V and which are such that

$$\begin{aligned} &\textstyle\sum_{v \in V} A(v, v) = 1, \\ &A(u, v) = 0 \quad \text{whenever } u, v \in V \text{ are adjacent, and} \\ &A \text{ is positive semidefinite.} \end{aligned} \tag{1.1}$$

It is easy to see that $\vartheta(G) \geq \alpha(G)$. Indeed, let $C \subseteq V$ be a nonempty stable set of G. Let $\chi^C\colon V \to \{0, 1\}$ denote the characteristic function of C, that is, $\chi^C(v) = 1$ if $v \in C$ and $\chi^C(v) = 0$ otherwise. Then the matrix $A\colon V \times V \to \mathbb{R}$ such that

$$A(u, v) = |C|^{-1} \chi^C(u) \chi^C(v)$$

can be seen to satisfy the properties in (1.1) and moreover

$$\sum_{u \in V} \sum_{v \in V} A(u, v) = |C|,$$

so that $\vartheta(G) \geq |C|$, and therefore, since C is any stable set of G, we have $\vartheta(G) \geq \alpha(G)$. For a finite graph, the theta number can be efficiently computed; that is one of the reasons that makes it an attractive bound.

Can we use the theta number to upper bound the stability number of the distance graph $G(H_n, \{1, \ldots, d-1\})$? Well, this is a finite graph, so the theta number immediately applies. But the matrix we have to consider is now indexed by all binary words of length n, and there are 2^n of them. So the matrix itself is very big, and it is then unlikely that we could use the computer to solve the resulting optimization problem, even for moderate values of n.

But the graph $G(H_n, \{1, \ldots, d-1\})$ is not just any finite graph. It is a distance graph over the Hamming cube, which is a highly symmetrical object. More specifically, the Hamming cube has many isometries, that is, there are many bijections from H_n to itself that preserve distances, and all of these bijections also preserve the adjacency relation of our graph. By exploiting these symmetries, we may reduce the size of the optimization problem, reducing the number of variables from 2^{2n}, which is the number of entries in the matrix A, to only $n + 1$. This allows us to solve the optimization problem and compute the theta number even for large values of n. This is the approach that was proposed by Delsarte [10] in order to bound $A(n, d)$, albeit in a different language.

Now what about our graph $G(S^{n-1}, (0, \theta))$? Can we use the theta number to upper bound its stability number? And what about using the theta number to upper bound the measurable stability number of $G(S^{n-1}, \theta)$ or the maximum density of a stable set of $G(\mathbb{R}^n, 1)$? Maybe the first idea that comes to the mind of someone with a background in combinatorial optimization/computer science when confronted with these questions is to *discretize*, that is, to make our infinite problems finite. For instance, one could try to approximate the infinite graphs by finite ones, and then use the computer. In this thesis, however, we take a different path. Namely, we show how to generalize the Lovász theta number to some kinds of infinite distance graphs — among which are those we discussed above — by considering infinite-dimensional optimization problems, that is, optimization problems with infinitely many variables and/or constraints.

In doing so, we keep our original graphs, instead of approximating them by finite graphs in which many of their original properties might disappear. By exploiting the symmetry of our infinite graphs, we may actually reduce the optimization problems we get to simpler ones, ultimately being able to solve them, sometimes analytically, sometimes with the help of the computer. This is analogous to the approach employed to compute the theta number of $G(H_n, \{1, \ldots, d - 1\})$, in which case it is by considering the symmetries of the Hamming cube that one manages to reduce the size of the original problem, making it possible to solve it. So in many cases we may actually compute the theta number of these infinite graphs, and even when we cannot do it, our optimization problems provide us with a tool that can be used to prove theorems about the stability number of some classes of infinite graphs.

It is then by considering the framework of distance graphs that we manage to use the methods and techniques of combinatorial optimization to help us deal with questions from other areas of mathematics. And it is by keeping ourselves from discretizing the original problems in order to conform them to the traditional framework of combinatorial optimization that we manage to exploit their structure, thus making it possible to successfully apply the optimization techniques.

1.2 The chromatic number of the Euclidean space

The *chromatic number* of $\mathbb{R}^n$, denoted by $\chi(\mathbb{R}^n)$, is the minimum number of colors needed to color the points of $\mathbb{R}^n$ in such a way that no two points at distance 1 get the same color. In other words, $\chi(\mathbb{R}^n)$ is the chromatic number of the infinite graph whose vertex set is $\mathbb{R}^n$ and in which two vertices are adjacent if and only if they lie at distance 1 from each other; this graph is the *unit-distance graph* on $\mathbb{R}^n$.

Obviously, $\chi(\mathbb{R}) = 2$. For $n \geq 2$, the exact value of $\chi(\mathbb{R}^n)$ is unknown. The problem of determining $\chi(\mathbb{R}^n)$ for $n \geq 2$ was proposed by Nelson in 1950 (cf. Soifer [48], Chapter 3) and later considered by mathematicians such as Isbell, Hadwiger, Moser, and Erdős; see the Chapter 3 of Soifer [48] for a detailed historical account.

In 1981, Falconer [13] introduced a restricted version of the parameter $\chi(\mathbb{R}^n)$. Namely, he defined the *measurable chromatic number* of $\mathbb{R}^n$, denoted by $\chi_{\mathrm{m}}(\mathbb{R}^n)$, as the minimum number of colors needed to color the points of $\mathbb{R}^n$ in such a way that no two points at distance 1 get the same color and the sets of points that are colored with a same given color are Lebesgue measurable. In other words, $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is the minimum number of Lebesgue measurable sets needed to partition $\mathbb{R}^n$ in such a way that no two points at distance 1 belong to the same set. Obviously, we have $\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \chi(\mathbb{R}^n)$.

It is readily seen that $\chi_{\mathrm{m}}(\mathbb{R}) = 2$. As in the case of the chromatic number, the exact value of $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is not known for $n \geq 2$. Most of the results in this thesis were motivated by the problem of finding good lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$. So in this section we survey some of the most important results concerning $\chi(\mathbb{R}^n)$ and $\chi_{\mathrm{m}}(\mathbb{R}^n)$. We also survey some of the main results concerning the parameter $m_1(\mathbb{R}^n)$, which is defined as the maximum density a Lebesgue measurable set $C \subseteq \mathbb{R}^n$ can have if it does not contain a pair of points at distance 1 from each other; this parameter is naturally related to $\chi_{\mathrm{m}}(\mathbb{R}^n)$, as we will see later. Our survey does not intend to be exhaustive; for a more complete account, we refer the reader to the survey by Székely [51].

We start with some comments on the parameter $\chi(\mathbb{R}^n)$. First, it is known that

$$4 \leq \chi(\mathbb{R}^2) \leq 7.$$

Here, the lower bound of 4 comes from the *Moser graph*, a graph with 7 vertices which can be realized in the plane with unit segments and which has chromatic number 4; see Figure 1.1 for a picture of the Moser graph. The upper bound of 7 comes from the coloring of the plane also given in Figure 1.1. According to Chapter 3 of the book by Soifer [48], the lower bound is due to Nelson and the upper bound is due to Isbell; both were found in 1950. Both the lower and upper bounds appear in Hadwiger [21], who credits them both to Isbell.

It is also known that

$$6 \leq \chi(\mathbb{R}^3) \leq 15,$$

where the lower bound is due to Nechushtan [35] and the upper bound is due to Coulson [9]. In Table 1.2 we list the best known lower bounds for $\chi(\mathbb{R}^n)$ for dimensions $n = 2, \ldots, 24$.

Frankl and Wilson [16] were the first to prove that $\chi(\mathbb{R}^n)$ grows exponentially with n. They showed that

$$\chi(\mathbb{R}^n) \geq (1 + o(1))1.2^n. \tag{1.2}$$

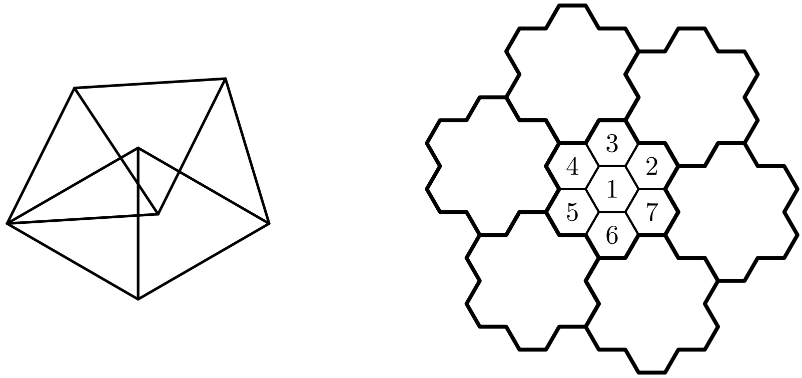

Figure 1.1. On the left we have the Moser graph, which has 7 vertices and chromatic number 4. It is drawn here so that all edges have unit length. On the right we show a coloring of the plane with 7 colors. Each of the regular hexagons has diameter slightly less than one. The points inside each hexagon are colored with one of seven colors; the points in the intersection of two or more hexagons can be colored with any of the colors of the neighboring hexagons. The coloring we show of the seven hexagons at the center is just pasted so as to cover the whole plane.

Later, Raigorodskii [39] gave the slightly better bound

$$\chi(\mathbb{R}^n) \geq (1.239\ldots + o(1))^n. \tag{1.3}$$

As for upper bounds, Larman and Rogers [27] prove that

$$\chi(\mathbb{R}^n) \leq (3 + o(1))^n.$$

Finally, notice that $\chi(\mathbb{R}^n)$ is at least the chromatic number of any finite subgraph of the unit-distance graph on $\mathbb{R}^n$. From a result of de Bruijn and Erdős [7] it follows that $\chi(\mathbb{R}^n)$ is actually equal to the chromatic number of some finite subgraph of the unit-distance graph on $\mathbb{R}^n$. The proof of this result uses the Axiom of Choice. For more on the relation between $\chi(\mathbb{R}^n)$ and the Axiom of Choice, see Chapter 46 of the book by Soifer [48].

As we mentioned before, the study of $\chi_{\mathrm{m}}(\mathbb{R}^n)$ began with Falconer [13], who proved that

$$\chi_{\mathrm{m}}(\mathbb{R}^2) \geq 5 \qquad \text{and} \qquad \chi_{\mathrm{m}}(\mathbb{R}^3) \geq 6.$$

Table 1.2 gives a list of the best lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 2, \ldots, 24$ which are available in the literature. In Chapters 3 and 4 we will improve on the results of this table. In particular, in Chapter 4 we give better lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 3, \ldots, 24$.

The construction of Figure 1.1 and the construction of Coulson [9] both use measurable sets only, hence

$$\chi_{\mathrm{m}}(\mathbb{R}^2) \leq 7 \qquad \text{and} \qquad \chi_{\mathrm{m}}(\mathbb{R}^3) \leq 15.$$

n	Lower bound for $\chi(\mathbb{R}^n)$	Lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$	Upper bound for $m_1(\mathbb{R}^n)$
2	4	5	0.27906976
3	6	6	0.18750000
4	6	8	0.12800000
5	8	11	0.09539473
6	10	15	0.07081295
7	14	19	0.05311365
8	16	30	0.03419769
9	16	35	0.02882153
10	19	45	0.02234835
11	19	56	0.01789325
12	24	70	0.01437590
13	31	84	0.01203324
14	35	102	0.00981770
15	37	119	0.00841374
16	67	148	0.00677838
17	67	174	0.00577854
18	68	194	0.00518111
19	82	263	0.00380311
20	97	315	0.00318213
21	114	374	0.00267706
22	133	526	0.00190205
23	154	754	0.00132755
24	178	933	0.00107286

Table 1.2. These are the best lower bounds for $\chi(\mathbb{R}^n)$ and $\chi_{\mathrm{m}}(\mathbb{R}^n)$ and the best upper bounds for $m_1(\mathbb{R}^n)$ found in the literature. The lower bounds for $\chi(\mathbb{R}^n)$ were all taken from the table compiled by Székely [51] in his survey, except for $\chi(\mathbb{R}^3) \geq 6$, which was given by Nechushtan [35]. The lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 2$ and 3 were given by Falconer [13]; all other lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ were given by Székely and Wormald [54]. Finally, the upper bound for $m_1(\mathbb{R}^2)$ was given by Székely [52]; all other upper bounds for $m_1(\mathbb{R}^n)$ were given by Székely and Wormald [54].

Also the construction of Larman and Rogers [27] uses only measurable sets, so

$$\chi_{\mathrm{m}}(\mathbb{R}^n) \leq (3 + o(1))^n.$$

Since $\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \chi(\mathbb{R}^n)$, the exponential lower bounds for $\chi(\mathbb{R}^n)$ carry over to $\chi_{\mathrm{m}}(\mathbb{R}^n)$. In Chapters 3 and 4 we also give exponential lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$, but they are not better than (1.2).

We know that $\chi(\mathbb{R}) = \chi_{\mathrm{m}}(\mathbb{R}) = 2$, but for $n \geq 2$ it is not known whether $\chi(\mathbb{R}^n)$ coincides with $\chi_{\mathrm{m}}(\mathbb{R}^n)$ or not. Since the Axiom of Choice is needed to construct non-measurable subsets of $\mathbb{R}^n$, if $\chi(\mathbb{R}^n) < \chi_{\mathrm{m}}(\mathbb{R}^n)$, then in order to color the points of $\mathbb{R}^n$ with $\chi(\mathbb{R}^n)$ colors in such a way that no two points at distance 1 get the same color, one would need to use the Axiom of Choice. Moreover, if for some dimension n we have strict inequality, then we would have that all finite subgraphs of the unit-distance graph of $\mathbb{R}^n$ have chromatic number strictly less than $\chi_{\mathrm{m}}(\mathbb{R}^n)$.

We now discuss a parameter which is related to the measurable chromatic number: the maximum density of 1-avoiding sets.

A set $C \subseteq \mathbb{R}^n$ *avoids* distance d if no two points in C are at distance d from each other. A set that avoids distance 1 is also called 1-*avoiding*. Given a Lebesgue measurable set $C \subseteq \mathbb{R}^n$, its *upper density* is defined as

$$\overline{\delta}(C) = \limsup_{R\to\infty} \frac{\mathrm{vol}(C \cap [-R, R]^n)}{\mathrm{vol}[-R, R]^n},$$

where $\mathrm{vol}\, X$ simply denotes the Lebesgue measure of set X. We then consider the parameter

$$m_1(\mathbb{R}^n) = \sup\{\, \overline{\delta}(C) : C \subseteq \mathbb{R}^n \text{ is measurable and 1-avoiding}\,\},$$

i.e., $m_1(\mathbb{R}^n)$ is the maximum upper density a 1-avoiding subset of $\mathbb{R}^n$ can have.

Notice that $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is the minimum number of 1-avoiding measurable sets one needs in order to partition $\mathbb{R}^n$. So we immediately have the inequality

$$m_1(\mathbb{R}^n)\chi_{\mathrm{m}}(\mathbb{R}^n) \geq 1,$$

so that any upper bound for $m_1(\mathbb{R}^n)$ implies a lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$. Finding lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is one of our main motivations in studying $m_1(\mathbb{R}^n)$.

It can be easily shown that $m_1(\mathbb{R}) = 1/2$. The problem of determining $m_1(\mathbb{R}^2)$ was mentioned already in the problem collection of Moser [34]. Later, Székely [52] showed that

$$m_1(\mathbb{R}^2) \leq 12/43 \approx 0.279.$$

Table 1.2 lists the best known upper bounds for $m_1(\mathbb{R}^n)$ for $n = 2, \ldots, 24$ which can be found in the literature. In Chapter 4 we improve on the bounds shown in the table for $n = 2, \ldots, 24$.

We finish by mentioning that the reciprocal of the exponential lower bound for $\chi(\mathbb{R}^n)$ from (1.2), which was given by Frankl and Wilson [16], can be seen to provide an exponential upper bound for $m_1(\mathbb{R}^n)$, showing thus that $m_1(\mathbb{R}^n)$ decays exponentially with n. Namely, we have

$$m_1(\mathbb{R}^n) \leq (1 + o(1))1.2^{-n} \approx (1 + o(1))0.833^n. \tag{1.4}$$

Similarly, the reciprocal of the lower bound given by Raigorodskii [39] (see equation (1.3)) can also be seen to provide an upper bound for $m_1(\mathbb{R}^n)$. In Chapter 4 we also prove that $m_1(\mathbb{R}^n)$ decays exponentially, but our bound is not better than (1.4).

1.3 Outline of the thesis

This thesis is divided into four chapters, including this introductory chapter. The contents of the other three chapters can be summarized as follows:

Chapter 2. The Lovász theta number. We define the Lovász theta number for finite graphs and present some of its basic properties, specially those that are useful to us in later chapters. We also show how to apply the theta number in order to upper bound the maximum cardinality of a binary code with prescribed length and minimum distance, showing that the bound it provides is the same one that was proposed by Delsarte [10], a fact first observed by McEliece, Rodemich, and Rumsey [33] and Schrijver [45]. It is useful to keep this example in mind while reading the other chapters, as the ideas used in this application are analogous to, if simpler than, those used later.

Chapter 3. Distance graphs on the sphere. Here we consider distance graphs on the sphere. We show how to generalize the Lovász theta number to such graphs in order to upper bound the stability number or the measurable stability number. By exploring a relation between the measurable chromatic number of distance graphs on the sphere and the measurable chromatic number of the Euclidean space, we provide improved lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 10, \ldots, 24$. We also consider other applications of the theta number, proving a theorem concerning the behavior of the measurable stability number of a distance graph as progressively more distances are set to define edges. Finally, we use our methods to show that the upper bound of Delsarte, Goethals, and Seidel [11] and Kabatyanskii and Levenshtein [23] for the maximum cardinality of spherical codes with prescribed minimum angular distance can be seen as coming from a generalization of the Lovász theta number, in the same way that the bound of Delsarte [10] for the maximum size of a binary code comes from the theta number for finite graphs, as shown in Chapter 2.

Chapter 4. Distance graphs on the Euclidean space. We next consider the problem of finding upper bounds for $m_1(\mathbb{R}^n)$ (see Section 1.2), the maximum density of a 1-avoiding subset of $\mathbb{R}^n$. We propose a bound based on linear programming and harmonic analysis and we show how to compute it and also how to strengthen it. From this approach we improve on the known upper bounds for $m_1(\mathbb{R}^n)$ for dimensions $n = 2, \ldots, 24$ and on the known lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for dimensions $n = 3, \ldots, 24$.

We also apply our approach to provide upper bounds for the density of distance-avoiding subsets of $\mathbb{R}^n$ as progressively more distances are avoided, proving a stronger version of a result of Furstenberg, Katznelson, and Weiss [17]. Finally, we show how our bounds actually come from a suitable generalization of the Lovász theta number to distance graphs defined over $\mathbb{R}^n$ and we study some of the properties of this generalization.

This thesis is partially based on the following two papers that have been accepted for publication:

- C. Bachoc, G. Nebe, F.M. de Oliveira Filho, and F. Vallentin, Lower bounds for measurable chromatic numbers, `arXiv:0801.1059v2 [math.CO]`, to appear in *Geometric and Functional Analysis*, 2008, 18pp.
- F.M. de Oliveira Filho and F. Vallentin, Fourier analysis, linear programming, and densities of distance avoiding sets in $\mathbb{R}^n$, `arXiv:0808.1822v2 [math.CO]`, to appear in *Journal of the European Mathematical Society*, 2008, 11pp.

1.4 Preliminaries

In this section we make a quick summary of some terms and concepts that appear throughout the thesis and we also fix our notation concerning graph theory, linear algebra, and optimization problems. We strived to keep the necessary preliminaries to a minimum and during the text we often reintroduce some terms that are also defined here, in order to keep the flow.

1.4a. Graph theory. A *graph* is a pair $G = (V, E)$ where V is a nonempty set and $E \subseteq \{ e \subseteq V : |e| = 2 \}$. Set V is the *vertex set* of G and its elements are the *vertices* of G. Set E is the *edge set* of G and its elements are called *edges*. All our graphs are thus *simple*, i.e., they have no loops or parallel edges.

We usually denote edge $\{u, v\}$ simply by uv. We say vertices u, $v \in V$ are *adjacent* if $uv \in E$. When V is finite, we say that G is a *finite* graph.

We say that a graph H is a *subgraph* of graph G if the vertex set of H is a subset of the vertex set of G and the edge set of H is a subset of the edge set of G. The *complementary graph* of a graph $G = (V, E)$, denoted by $\overline{G}$, is the graph with vertex set V in which two distinct vertices are adjacent if and only if they are nonadjacent in G.

Let $G = (V, E)$ be a graph. A set $C \subseteq V$ is *stable* if no two vertices in it are adjacent. The *stability number* of a graph, denoted by $\alpha(G)$, is the maximum size a stable set of G can have.

A *coloring* of G is an assignment of colors to the vertices of G in such a way that no two adjacent vertices are assigned the same color. In a given coloring, the sets of vertices that receive a same color are termed *color classes*. The *chromatic number* of G, denoted by $\chi(G)$, is the minimum number of colors one needs to color G, i.e., it is the minimum number of colors used by any coloring of G.

Notice that in any coloring of G the color classes are stable sets. A coloring is thus a partition of V into stable sets. The following fundamental relation between $\alpha(G)$ and $\chi(G)$ for a finite graph G then follows at once:

$$\alpha(G)\chi(G) \geq |V|. \tag{1.5}$$

So an upper bound for $\alpha(G)$ implies a lower bound for $\chi(G)$.

For a finite graph G, it is an NP-hard problem to compute $\alpha(G)$ or $\chi(G)$. These are some of the first combinatorial problems that were proven to be NP-hard; they figure in the list of problems considered by Karp [24].

An *automorphism* of a graph $G = (V, E)$ is a bijection $\varphi: V \to V$ which preserves adjacency, that is, for any two vertices u, v of G, we have that u is adjacent to v if and only if $\varphi(u)$ is adjacent to $\varphi(v)$. The set of all automorphisms of G, denoted by $\mathrm{Aut}(G)$, is a group under composition, which is called the *automorphism group* of G. A graph G is said to be *vertex-transitive* if for any two vertices u and v of G there is an automorphism φ of G which is such that $\varphi(u) = v$.

1.4b. Vectors and matrices. We denote the transpose of a matrix A by A^T. The conjugate transpose of a complex matrix A is denoted by A^*. For us, a vector $x \in \mathbb{R}^n$ (or $\mathbb{C}^n$) is always a column vector, that is, in matrix operations it is treated as a matrix with n rows and 1 column.

We will often need to use matrices whose rows and columns are indexed by finite sets other than $\{1, \ldots, n\}$. If V is a finite set, then $A: V \times V \to \mathbb{R}$ is a real matrix whose rows and columns are indexed by V and likewise $x: V \to \mathbb{R}$ is a real vector whose entries are indexed by V. Needless to say, all that follows immediately applies to matrices indexed by some finite set V, even though our presentation is in terms of matrices indexed by $\{1, \ldots, n\}$.

Let A be an $n \times n$ matrix, real or complex. We say that A is *symmetric* if $A_{ij} = A_{ji}$ for all i, $j = 1, \ldots, n$. We also write $A(i, j)$ for A_{ij}. A complex $n \times n$ matrix A is *Hermitian* if $A_{ij} = \overline{A_{ji}}$ for all i, $j = 1, \ldots, n$.

A real matrix $A \in \mathbb{R}^{n \times n}$ is *positive semidefinite* if it is symmetric and

$$p^\mathsf{T} A p = \sum_{i,j=1}^{n} A_{ij} p_i p_j \geq 0$$

for every vector $p \in \mathbb{R}^n$. A basic fact about positive semidefinite matrices is the equivalence

$$\begin{aligned} & A \in \mathbb{R}^{n \times n} \text{ is positive semidefinite} \\ & \quad \Longleftrightarrow \text{ there exists a matrix } B \in \mathbb{R}^{n \times n} \text{ such that } A = BB^\mathsf{T} \\ & \quad \Longleftrightarrow A \text{ is symmetric and all of its eigenvalues are nonnegative} \\ & \quad \Longleftrightarrow \text{ there exist } k \text{ and orthogonal vectors } u_1, \ldots, u_k \text{ in } \mathbb{R}^n \\ & \qquad \quad \text{such that } A = u_1 u_1^\mathsf{T} + \cdots + u_k u_k^\mathsf{T}. \end{aligned} \tag{1.6}$$

We say that a complex matrix $A \in \mathbb{C}^{n \times n}$ is *positive semidefinite* if it is Hermitian and

$$p^* A p = \sum_{i,j=1}^{n} A_{ij} \overline{p_i} p_j \geq 0$$

for every vector $p \in \mathbb{C}^n$. The analogue of equivalence (1.6) holds for complex positive semidefinite matrices, with $\mathbb{R}$ replaced by $\mathbb{C}$ and transposes replaced by conjugated transposes.

The *trace* of an $n \times n$ real or complex matrix A is the sum of its diagonal elements and is denoted by $\mathrm{Tr}\, A$. It can be shown that the trace of a matrix equals the sum of its eigenvalues.

Given two matrices A, $B \in \mathbb{R}^{n\times n}$, we let

$$\langle A, B\rangle = \mathrm{Tr}(A^{\mathsf{T}} B).$$

This defines an inner product for the space $\mathbb{R}^{n\times n}$, to which we will often refer to as the *trace inner product*. If A and B are positive semidefinite, then $\langle A, B\rangle \geq 0$, as can be seen from the last equivalence in (1.6). For complex matrices we may likewise define the trace inner product by letting $\langle A, B\rangle = \mathrm{Tr}(A^* B)$.

Finally, we denote by $x \cdot y$ the standard Euclidean inner product between two vectors x, $y \in \mathbb{R}^n$, that is

$$x \cdot y = x^{\mathsf{T}} y = \sum_{i=1}^{n} x_i y_i.$$

We denote the Euclidean norm of a vector $x \in \mathbb{R}^n$ by $\|x\| = (x \cdot x)^{1/2}$.

1.4c. Optimization problems and semidefinite programming. We make heavy use of optimization problems and in relation to them our terminology is quite standard. We assume the reader has some familiarity with linear programming, which we often use throughout the thesis. What we need of linear programming can, in any case, be found in any standard book on the subject, like the book by Schrijver [47]. We also often use semidefinite programming, so we choose to introduce our notation for optimization problems in general while at the same time summarizing some facts about semidefinite programming.

Let $A_1, \ldots, A_m, C \in \mathbb{R}^{n\times n}$ be symmetric matrices and $b_1, \ldots, b_m$ be real numbers. Say we partition the set $\{1, \ldots, m\}$ into sets $I_=$ and $I_\leq$. A *semidefinite programming problem* is the problem of finding the maximum of

$$\langle C, X\rangle \tag{1.7}$$

where X is a matrix such that

$$\begin{aligned} &\langle A_i, X\rangle = b_i \quad \text{for } i \in I_=,\\ &\langle A_i, X\rangle \leq b_i \quad \text{for } i \in I_\leq,\\ &X \in \mathbb{R}^{n\times n} \text{ is positive semidefinite.} \end{aligned} \tag{1.8}$$

We usually express this problem (and also other optimization problems) in the following compact form:

$$\begin{aligned} \max\ &\langle C, X\rangle\\ &\langle A_i, X\rangle = b_i \quad \text{for } i \in I_=,\\ &\langle A_i, X\rangle \leq b_i \quad \text{for } i \in I_\leq,\\ &X \in \mathbb{R}^{n\times n} \text{ is positive semidefinite.} \end{aligned} \tag{1.9}$$

Notice that any linear programming problem can be written as a semidefinite programming problem in the form above; in this case matrices $A_1, \ldots, A_m, C$ are all diagonal.

We now introduce some terminology that also applies to other optimization problems. Function (1.7) is called the *objective function*. Conditions (1.8) are the *constraints* of the problem. Matrix X is the *variable* matrix.

A matrix $X \in \mathbb{R}^{n\times n}$ that satisfies (1.8) is said to be a *feasible solution* of problem (1.9). If (1.9) has a feasible solution, then it is said to be *feasible*, otherwise we say it is *infeasible*. If $X \in \mathbb{R}^{n\times n}$, then $\langle C, X\rangle$ is its *objective value*. The maximum itself is the *optimal value* of problem (1.9) and a feasible solution whose objective value is the optimal value is an *optimal solution*.

Needless to say, we could just as well have started with a minimization problem instead of a maximization problem. Sometimes when we deal with optimization problems, it might not be clear that the maximum/minimum exists, so we use supremum/infimum instead.

Like a linear programming problem, a semidefinite programming problem such as (1.9) has a dual problem. The *dual problem* of (1.9) is

$$\begin{aligned} \min\ & y_1b_1 + \cdots + y_mb_m \\ & y_1A_1 + \cdots + y_mA_m - C \text{ is positive semidefinite}, \\ & y_i \geq 0 \quad \text{for } i \in I_{\leq}. \end{aligned} \tag{1.10}$$

Here, the optimization variables are $y_1, \ldots\ y_m$. We also note that this problem can be put in form (1.9), as one can show after a bit of work.

It is easy to prove that *weak duality* holds, i.e., that the minimum is at least the maximum. Indeed, if X is a feasible solution of (1.9) and $(y_1, \ldots, y_m)$ is a feasible solution of (1.10), then since for positive semidefinite matrices A and B we have $\langle A, B\rangle \geq 0$, it follows that

$$\begin{aligned} 0 &\leq \langle y_1A_1 + \cdots + y_mA_m - C, X\rangle \\ &= y_1\langle A_1, X\rangle + \cdots + y_m\langle A_m, X\rangle - \langle C, X\rangle \\ &\leq y_1b_1 + \cdots + y_mb_m - \langle C, X\rangle, \end{aligned}$$

as we wanted.

Strong duality, i.e., that the minimum and the maximum coincide, does not always hold for primal/dual pairs of semidefinite programs. A simple sufficient condition for strong duality to hold is known, however. We say that (1.9) is *strictly feasible* if it has a feasible solution X which is positive definite and such that $\langle A_i, X\rangle < b_i$ for all $i \in I_{\leq}$. Here, observe that we require X to be *positive definite*, that is, all its eigenvalues must be positive.

If (1.9) is strictly feasible and has finite optimal value, then also the dual problem has a finite optimal value and the optima coincide. For more on duality theory of semidefinite programming, see e.g. the book by Nesterov and Nemirovskii [36].

We finish our discussion on semidefinite programming with a note on complexity. By the use of the ellipsoid method (cf. Grötschel, Lovász, and Schrijver [19]), semidefinite programs can be solved in polynomial time to any fixed precision. In practice, however, interior-point algorithms, which are also polynomial-time, are the method of choice in dealing with semidefinite programs.

Chapter Two

The Lovász theta number

ONE of the parameters that can be shown to provide an upper bound to the stability number of a finite graph is the Lovász theta number, introduced by Lovász [29] in 1979. It can be described as the optimal solution of a semidefinite programming problem, and thus can be efficiently computed. In many cases, the bounds provided by the theta number and derived methods are the best known (see, e.g., the thesis of Gvozdenović [20]).

In this chapter, we define the theta number and quickly review some of its most important properties. Our presentation is very selective, as we keep in mind the developments of Chapters 3 and 4, where we generalize the theta number to some kinds of infinite graphs.

Finally, we remind the reader that all the graph-theoretic notions and the semidefinite programming terminology we use are presented in Section 1.4. We also stress that all graphs considered in this chapter are finite.

2.1 Definition and basic properties

Let $G = (V, E)$ be a graph. In a 1979 paper [29], Lovász introduced a parameter $\vartheta(G)$ that provides an upper bound for the stability number $\alpha(G)$ and that moreover can be efficiently computed. The parameter $\vartheta(G)$ has many more interesting properties, some of which we will study later in this chapter. We begin by defining $\vartheta(G)$ as the optimal value of the following semidefinite programming problem:

$$\begin{aligned} \max\ & \textstyle\sum_{u,v\in V} A(u,v) \\ & \textstyle\sum_{v\in V} A(v,v) = 1, \\ & A(u,v) = 0 \quad \text{if } u \text{ is adjacent to } v, \\ & A\colon V\times V \to \mathbb{R} \text{ is positive semidefinite.} \end{aligned} \tag{2.1}$$

We start by showing that

$$\vartheta(G) \geq \alpha(G). \tag{2.2}$$

The proof is simple. Let $C \subseteq V$ be a nonempty stable set and $\chi^C: V \to \{0,1\}$ be its characteristic function, which is equal to 1 only for the elements of C. The matrix $A: V \times V \to \mathbb{R}$ such that

$$A(u,v) = |C|^{-1}\chi^C(u)\chi^C(v)$$

is feasible for problem (2.1). Indeed, the first constraint is trivially satisfied and the second set of constraints is also satisfied since C is stable. Finally, A is positive semidefinite since $A = |C|^{-1}\chi^C(\chi^C)^{\mathsf{T}}$. Moreover, we have that the sum of all the entries of A is equal to $|C|$, so $\vartheta(G) \geq |C|$, and it follows that $\vartheta(G) \geq \alpha(G)$.

Notice that problem (2.1) is strictly feasible (cf. Section 1.4c) and has a finite optimal value. So strong duality holds between it and its dual, which then provides an equivalent formulation for $\vartheta(G)$, namely

$$\begin{array}{ll} \min & \lambda \\ & Z(v,v) = \lambda - 1 \quad \text{for all } v \in V, \\ & Z(u,v) = -1 \qquad \text{if } u, v \in V,\ u \neq v, \text{ are nonadjacent}, \\ & Z: V \times V \to \mathbb{R} \text{ is positive semidefinite.} \end{array} \tag{2.3}$$

It is instructive to show directly from (2.3) that $\vartheta(G) \geq \alpha(G)$. To do so, let Z be any feasible solution of (2.3) and $C \subseteq V$ be any nonempty stable set of G. Let λ be such that $Z(v,v) = \lambda - 1$ for all $v \in V$. Then, since Z is positive semidefinite and C is stable,

$$0 \leq \sum_{u,v \in C} Z(u,v) = |C|(\lambda - 1) - (|C|^2 - |C|),$$

and it follows that $\lambda \geq |C|$ and hence $\vartheta(G) \geq |C|$. Notice that in fact any feasible solution of (2.3) provides an upper bound to $\alpha(G)$; this is a reflection of the weak duality relation between (2.3) and (2.1).

By using (2.3) we can also prove that

$$\vartheta(G) \leq \chi(\overline{G}), \tag{2.4}$$

where $\overline{G}$ is the complementary graph of G.

To prove (2.4), suppose we have stable sets $C_1, \ldots, C_k$ of $\overline{G}$ which partition V, where $k = \chi(\overline{G})$. Consider the matrix

$$Z = k \sum_{i=1}^{k} \chi^{C_i}(\chi^{C_i})^{\mathsf{T}} - J,$$

where J is the all one matrix. Clearly, Z satisfies the first two sets of constraints of (2.3) and its diagonal elements are all equal to $\chi(\overline{G}) - 1$. So, if we prove that Z is positive semidefinite, we will have proven that Z is feasible for (2.3), and hence (2.4) will follow.

To prove Z is positive semidefinite, let $p: V \to \mathbb{R}$ be any vector. For $i = 1, \ldots, k$, let $\pi_i = (\chi^{C_i})^{\mathsf{T}} p$. Since

$$J = (\chi^{C_1} + \cdots + \chi^{C_k})(\chi^{C_1} + \cdots + \chi^{C_k})^{\mathsf{T}},$$

it follows that

$$\begin{aligned} p^\mathsf{T} Z p &= k \sum_{i=1}^{k} \pi_i^2 - \sum_{i,j=1}^{k} \pi_i \pi_j \\ &= \sum_{i=1}^{k} \sum_{j=i+1}^{k} (\pi_i - \pi_j)^2 \\ &\geq 0, \end{aligned}$$

and hence Z is positive semidefinite.

Combining (2.2) with (2.4), we get the following theorem of Lovász [29].

Theorem 2.1. *We have* $\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G})$ *for every graph* G. ■

2.2 Other properties of the theta number

We will now quickly review some of the most important properties of the theta number. The first property we mention is the monotonicity of $\vartheta(G)$, that is, if H is a subgraph of G and has the same vertex set as G, then

$$\vartheta(H) \geq \vartheta(G).$$

Indeed, let V be the common vertex set of G and H. Suppose $Z\colon V \times V \to \mathbb{R}$ is an optimal solution of formulation (2.3) for $\vartheta(H)$. Then Z is also feasible for (2.3) when we consider the edge set of G, hence $\vartheta(H) \geq \vartheta(G)$, as we wanted.

The next property is analogous to the inequality

$$\alpha(G)\chi(G) \geq |V|,$$

which holds for a graph $G = (V, E)$. It states that, for any graph $G = (V, E)$,

$$\vartheta(G)\vartheta(\overline{G}) \geq |V|. \tag{2.5}$$

To prove (2.5), let Z be an optimal solution of formulation (2.3) of $\vartheta(G)$. We then have $Z(v, v) = \vartheta(G) - 1$ for all $v \in V$. Consider the matrix

$$A = (\vartheta(G)|V|)^{-1}(Z + J).$$

Notice A is feasible for formulation (2.1) of $\vartheta(\overline{G})$, so

$$\begin{aligned} \vartheta(\overline{G}) &\geq \sum_{u,v \in V} A(u, v) \\ &= \langle J, A \rangle \\ &= (\vartheta(G)|V|)^{-1}(\langle J, Z \rangle + \langle J, J \rangle) \\ &\geq (\vartheta(G)|V|)^{-1}\langle J, J \rangle \\ &= \vartheta(G)^{-1}|V|, \end{aligned}$$

proving (2.5). Here we used that $\langle J, Z\rangle \geq 0$ since both J and Z are positive semidefinite.

For vertex-transitive graphs we get equality in (2.5). To see this, let $G = (V, E)$ be a vertex-transitive graph. Let A be an optimal solution of formulation (2.1) for $\vartheta(G)$. We say A is *invariant* under $\mathrm{Aut}(G)$, the automorphism group of G, if

$$A(\varphi(u), \varphi(v)) = A(u, v) \qquad \text{for all } \varphi \in \mathrm{Aut}(G) \text{ and } u, v \in V.$$

We may assume that A is invariant. Indeed, if φ is an automorphism of G, then

$$(\varphi \cdot A)(u, v) = A(\varphi(u), \varphi(v))$$

is also an optimal solution of (2.1). So the matrix

$$\sum_{\varphi \in \mathrm{Aut}(G)} |\,\mathrm{Aut}(G)|^{-1}(\varphi \cdot A)$$

is also an optimal solution which is moreover invariant.

So we may assume A is an invariant optimal solution. We observe in particular that, since G is vertex-transitive, all diagonal entries of A coincide, and hence are equal to $1/|V|$.

Consider then the matrix

$$Z = (|V|^2/\vartheta(G))A - J.$$

We claim that Z is a feasible solution of formulation (2.3) for $\vartheta(\overline{G})$. To show this, it suffices to prove that Z is positive semidefinite, as the other constraints are trivially satisfied.

To prove Z is positive semidefinite, we first observe that $\mathbf{1}$, the all one vector, is an eigenvector of A. To see this, consider two vertices u and v. Then there is an automorphism φ of G such that $\varphi(u) = v$. Then

$$\sum_{w \in V} A(v, w) = \sum_{w \in V} A(\varphi(u), w) = \sum_{w \in V} A(u, \varphi^{-1}(w)) = \sum_{w \in V} A(u, w),$$

and it follows that $\mathbf{1}$ is indeed an eigenvector of A. Moreover, since the sum of all entries of A is $\vartheta(G)$, the eigenvalue associated with $\mathbf{1}$ is $\vartheta(G)/|V|$.

The vector $\mathbf{1}$ is also an eigenvector of J, with associated eigenvalue $|V|$; moreover $\mathbf{1}$ is the only (up to scalar multiples) eigenvector of J with nonzero associated eigenvalue. With this it becomes at once clear that Z is positive semidefinite.

We have thus proved that $\vartheta(\overline{G}) \leq |V|/\vartheta(G)$, since the diagonal elements of A are all equal to $1/|V|$. So we have $\vartheta(G)\vartheta(\overline{G}) \leq |V|$ and, together with (2.5), we obtain the identity

$$\vartheta(G)\vartheta(\overline{G}) = |V|.$$

The following theorem summarizes the results of this section. These results are also contained in Lovász [29].

Theorem 2.2. *Let G and H be graphs over the same vertex set V. We have:*

(i) *if H is a subgraph of G, then $\vartheta(H) \geq \vartheta(G)$;*
(ii) *$\vartheta(G)\vartheta(\overline{G}) \geq |V|$;*
(iii) *if G is vertex-transitive, then $\vartheta(G)\vartheta(\overline{G}) = |V|$.* ∎

2.3 An application to binary codes

It is natural to ask if there are further constraints that can be added to the formulation of the theta number of a graph in order to strengthen it. McEliece, Rodemich, and Rumsey [33], and independently Schrijver [45], proposed adding the constraint $A \geq 0$ to (2.1), i.e., requiring the matrix to be nonnegative. They thus introduced the parameter ϑ', which obviously also upper bounds $\alpha(G)$. In using the notation ϑ' for this parameter, we in fact follow the choice of Schrijver; McEliece, Rodemich, and Rumsey use α_L instead of ϑ'.

One application of the parameter ϑ' is in giving an upper bound for the sizes of binary codes with prescribed minimum distance. A *binary code* is a subset C of the Hamming cube $H_n = \{0,1\}^n$. The code C is said to have *length* n. If x and y are two codewords, the *Hamming distance* between them is

$$\mathrm{dist_H}(x, y) = |\{i : x_i \neq y_i\}|.$$

The *minimum distance* of the code C is the minimum distance between any pair of distinct codewords in C. We denote by $A(n, d)$ the maximum size a binary code of length n and minimum distance at least d can have. The parameter $A(n, d)$ is of particular interest in coding theory; see, e.g., the book by MacWilliams and Sloane [31].

Say we want to find $A(n, d)$. Consider the graph G whose vertex set is $\{0, 1\}^n$ and in which two vertices are adjacent if and only if the Hamming distance between them is strictly less than d. Then $A(n, d)$ is the stability number of G, and $\vartheta'(G)$ provides an upper bound for $A(n, d)$.

In other words, the following semidefinite programming problem gives an upper bound for $A(n, d)$:

$$\begin{aligned}\max\ & \textstyle\sum_{x,y\in H_n} A(x, y)\\ & \textstyle\sum_{x\in H_n} A(x, x) = 1,\\ & A(x, y) = 0 \quad \text{if } 0 < \mathrm{dist_H}(x, y) < d,\\ & A\colon H_n \times H_n \to \mathbb{R} \text{ is nonnegative and positive semidefinite.}\end{aligned} \tag{2.6}$$

In its present form, problem (2.6) is not of much use, since the size of the matrix A is exponential in n. A simple observation can help us overcome this problem however.

Let $\mathrm{Iso}(H_n)$ be the *isometry group* of the Hamming cube, that is, $\mathrm{Iso}(H_n)$ is the group of all bijections $\sigma\colon H_n \to H_n$ which preserve Hamming distances. (Notice that any element of $\mathrm{Iso}(H_n)$ is also an automorphism of the graph G we defined above.)

A matrix $A\colon H_n \times H_n \to \mathbb{R}$ is *invariant* under $\mathrm{Iso}(H_n)$ if

$$A(\sigma(x), \sigma(y)) = A(x, y) \qquad \text{for all } \sigma \in \mathrm{Iso}(H_n) \text{ and } x,\, y \in H_n.$$

Let A be an optimal solution of (2.6). We may assume that A is invariant, as otherwise the matrix

$$\sum_{\sigma \in \mathrm{Iso}(H_n)} |\,\mathrm{Iso}(H_n)|^{-1}(\sigma \cdot A)$$

is an optimal solution of (2.6) which is invariant, and we can take it instead of A. Here, $\sigma \cdot A$ is the matrix such that $(\sigma \cdot A)(x, y) = A(\sigma(x), \sigma(y))$.

So, since we may restrict ourselves to invariant matrices, we may use the following result to our advantage.

Theorem 2.3. *A matrix $A\colon H_n \times H_n \to \mathbb{R}$ is invariant and positive semidefinite if and only if*

$$A(x, y) = \sum_{k=0}^{n} f_k K_k^n(\mathrm{dist}_{\mathrm{H}}(x, y))$$

for nonnegative numbers f_0, …, f_n, where $K_k^n(t)$ is the Krawtchouk polynomial of degree k, which for integer t, $0 \le t \le n$, is given by

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

Using this theorem, we may rewrite (2.6) to obtain the equivalent problem

$$\begin{array}{ll}\max & 2^{2n} f_0 \\ & \sum_{k=0}^{n} f_k \binom{n}{k} = 2^{-n}, \\ & \sum_{k=0}^{n} f_k K_k^n(t) = 0 \qquad \text{for } t = 1, \ldots, d-1, \\ & \sum_{k=0}^{n} f_k K_k^n(t) \ge 0 \qquad \text{for } t = 0, \ldots, n, \\ & f_k \ge 0 \qquad\qquad\qquad\quad \text{for } k = 0, \ldots, n.\end{array}$$

Notice that the above problem is a linear programming problem with $n+1$ variables, which can be efficiently solved on a computer. This upper bound to $A(n, d)$ was introduced by Delsarte [10] in his doctoral thesis in the more general framework of (symmetric) association schemes and is generally known as the *linear programming bound*. The observation that it can be obtained from ϑ' is due to McEliece, Rodemich, and Rumsey [33] and Schrijver [45].

We finish this section with a proof of Theorem 2.3.

Proof of Theorem 2.3. If $A\colon H_n \times H_n \to \mathbb{R}$ is an invariant matrix, then $A(x, y)$ only depends on the distance between x and y, because if $\mathrm{dist}_{\mathrm{H}}(x, y) = \mathrm{dist}_{\mathrm{H}}(x', y')$, then there is $\sigma \in \mathrm{Iso}(H_n)$ such that $\sigma(x) = x'$ and $\sigma(y) = y'$. So such an invariant matrix A can be seen as a one variable function $A\colon \{0, \ldots, n\} \to \mathbb{R}$, and vice versa. This means that $\mathcal{W}$, the vector space of all invariant matrices $A\colon H_n \times H_n \to \mathbb{R}$, has dimension $n + 1$.

Let us find a basis for $\mathcal{W}$. For $u \in H_n$, consider the function $\chi_u\colon H_n \to \mathbb{R}$ such that $\chi_u(x) = (-1)^{u \cdot x}$, where $x \cdot y = \sum_{i=1}^n x_i y_i$. We observe that the functions χ_u, $u \in H_n$, are orthogonal with respect to the standard inner product in $\mathbb{R}^{H_n}$.

For $k = 0, \ldots, n$, consider the matrix $B_k\colon H_n \times H_n \to \mathbb{R}$ given by

$$B_k(x, y) = \sum_{\substack{u \in H_n \\ |u| = k}} \chi_u(x)\chi_u(y), \tag{2.7}$$

where $|u| = |\{i : u_i \neq 0\}|$ is the *weight* of the codeword u. We claim that each B_k is invariant.

Indeed, if σ is any isometry of the Hamming cube and x, y are any two codewords, then by expanding (2.7) using the definition of the χ_u we get

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

If we consider addition of codewords as an operation over $\mathrm{GF}(2)$, then since σ is an isometry, codewords $x + y$ and $\sigma(x) + \sigma(y)$ must both have the same weight, which is equal to $\mathrm{dist}_{\mathrm{H}}(x, y)$. Since the sum is taken over all codewords u of a given weight, it follows that $B_k(\sigma(x), \sigma(y)) = B_k(x, y)$, and B_k is invariant as claimed.

Now, since the χ_u are orthogonal, the matrices B_k are also orthogonal with respect to the trace inner product on $\mathbb{R}^{H_n \times H_n}$, i.e., $\langle B_k, B_l \rangle = \mathrm{Tr}(B_k^{\mathsf{T}} B_l) = 0$ if $k \neq l$. From this, and since $\mathcal{W}$ has dimension $n+1$, we conclude that $B_0, \ldots, B_n$ is an orthogonal basis of $\mathcal{W}$.

So, if $A\colon H_n \times H_n \to \mathbb{R}$ is a positive semidefinite invariant matrix, we must have

$$A = \sum_{k=0}^n f_k B_k \tag{2.8}$$

for some numbers f_k, which must be nonnegative since $\langle A, B_k \rangle \geq 0$ for all k, as both A and B_k are positive semidefinite. Conversely, if A is written in the form (2.8) with nonnegative numbers f_k, then A is invariant and positive semidefinite.

To finish, we need only note that $B_k(x, y) = K_k^n(\mathrm{dist}_{\mathrm{H}}(x, y))$, what follows from the definitions of B_k and K_k^n. ■

2.4 Notes

As we said in the beginning of the chapter, our presentation is very selective. We mention, however, two good sources of further information on the theta number: Chapter 67 of the book by Schrijver [46] and the nice survey paper by Knuth [26].

Though now the theta number is usually introduced as a polynomially computable bound for the stability number or the chromatic number, at the time of its introduction this was not known to be the case (in fact, it was not yet known that linear programming problems could be solved in polynomial time). Lovász [29] remarks, however, after proving strong duality between (2.1) and (2.3), that such a result provides a "good characterization of the value $\vartheta(G)$", so the computational complexity aspect was already present.

Chapter Three

Distance graphs on the sphere

THE Lovász theta number, which we defined in Chapter 2, provides a way to upper bound the stability number of a finite graph. Also for an infinite graph can we consider the notion of stability number and related notions and then the question arises of how to provide bounds for such parameters of an infinite graph. In this chapter we consider infinite graphs whose vertex set is the $(n-1)$-dimensional unit sphere and we show how to generalize the theta number for these graphs. As a consequence of our study, we improve the lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$, the measurable chromatic number of $\mathbb{R}^n$ (see Section 1.2) for $n = 10, \ldots, 24$. This chapter is partially based on the paper by Bachoc, Nebe, Oliveira, and Vallentin [3].

3.1 An infinite graph

Recall we denote the standard Euclidean inner product between vectors $x, y \in \mathbb{R}^n$ by $x \cdot y$. The $(n-1)$-*dimensional unit sphere* is the set

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

From now on we assume $n \geq 2$, unless otherwise noted. Let $D \subseteq [-1, 1)$ be a closed subset of the real line. Consider the graph $G(S^{n-1}, D)$ whose vertex set is S^{n-1} and in which two vertices x and y are adjacent if and only if $x \cdot y \in D$. Notice $G(S^{n-1}, D)$ is an infinite graph. We say it is a *distance graph* because its vertex set is a metric space and because the adjacency of any two vertices depends only on the distance between them.

Stable sets in $G(S^{n-1}, D)$ can be infinite, so its stability number is infinite. If we restrict ourselves to Lebesgue measurable stable sets, however, we may give an alternative definition of the stability number of $G(S^{n-1}, D)$.

For $n \geq 1$, let ω be the surface measure on S^{n-1}, normalized so that

$$\omega(S^{n-1}) = \omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

Then the *measurable stability number* of $G(S^{n-1}, D)$ is

$$\alpha_{\mathrm{m}}(G(S^{n-1}, D)) = \sup\{\, \omega(C) : C \subseteq S^{n-1} \text{ is stable and measurable}\,\}.$$

It is easy to see that $\alpha_{\mathrm{m}}(G(S^{n-1}, D)) > 0$ for every closed subset $D \subseteq [-1, 1)$ of the real line.

We may define the *chromatic number* of $G(S^{n-1}, D)$, which we here denote by $\chi(G(S^{n-1}, D))$, simply as the minimum number of colors needed to color the points of S^{n-1} in such a way that if two points are adjacent, then their colors are different. Equivalently, this is the minimum number of stable sets (not necessarily measurable) one needs in order to partition S^{n-1}. Lovász [30] studied the parameter $\chi(G(S^{n-1}, D))$ for the case in which D is a singleton.

It is easy to see that $\chi(G(S^{n-1}, D)) < \infty$ for any closed $D \subseteq [-1, 1)$: It suffices to partition the sphere into parts of very small diameter, so that each part is a stable set. Notice also that $\chi(G(S^{n-1}, D))$ is at least the chromatic number of any finite subgraph of $G(S^{n-1}, D)$. From a result of de Bruijn and Erdős [7], we know in fact that $\chi(G(S^{n-1}, D))$ is equal to the chromatic number of some finite subgraph.

In Section 1.2, we have defined the chromatic number $\chi(\mathbb{R}^n)$ of the Euclidean space $\mathbb{R}^n$, which bears resemblance to the chromatic number of $G(S^{n-1}, D)$. There we also defined the measurable chromatic number $\chi_{\mathrm{m}}(\mathbb{R}^n)$ of $\mathbb{R}^n$, which was introduced by Falconer [13]. Following this idea of Falconer, we may define the *measurable chromatic number* of $G(S^{n-1}, D)$, which we denote by $\chi_{\mathrm{m}}(G(S^{n-1}, D))$, as the minimum number of measurable stable sets needed to partition S^{n-1}. Notice this is the same as the chromatic number of $G(S^{n-1}, D)$, except that now we require the color classes to be measurable sets. In analogy to inequality (1.5), which holds for finite graphs, we have

$$\alpha_{\mathrm{m}}(G(S^{n-1}, D))\, \chi_{\mathrm{m}}(G(S^{n-1}, D)) \geq \omega_n. \tag{3.1}$$

In this chapter we will study $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$ and $\chi_{\mathrm{m}}(G(S^{n-1}, D))$, which are both more amenable to the application of analytic techniques than $\chi(G(S^{n-1}, D))$. We are specially interested in the case $D = \{t\}$ for some $-1 \leq t < 1$. This case is of particular interest since

$$\chi_{\mathrm{m}}(G(S^{n-1}, t)) \leq \chi_{\mathrm{m}}(\mathbb{R}^n) \qquad \text{for every } -1 \leq t < 1, \tag{3.2}$$

where we write $G(S^{n-1}, t) = G(S^{n-1}, \{t\})$. To see this, notice that if we have a coloring of $\mathbb{R}^n$ with measurable color classes in which no color class contains a pair of points at distance 1, we may scale this coloring until no color class contains a pair of points at distance $\sqrt{2 - 2t}$. Then the intersection of the scaled coloring with S^{n-1} will give a coloring of $G(S^{n-1}, t)$ with measurable color classes and at most $\chi_{\mathrm{m}}(\mathbb{R}^n)$ colors.

3.2 Preliminaries on functional analysis

We wish to generalize the Lovász theta number to the infinite graphs on the sphere that we consider. To this end, we need a generalized notion of matrix that will allow us to speak of matrices indexed by S^{n-1}. The notion we need here is that of a *kernel*, an object studied in functional analysis.

So our goal in this section is to give a summary of all the concepts of functional analysis that we will need. We will develop the theory for functions over S^{n-1}, but everything remains unchanged if we replace the sphere by other measure spaces, provided they satisfy some requirements. In Chapter 4, for instance, we will need to deal with functions defined over the cube $[-R, R]^n$. All theorems presented here also hold for these functions. We also develop the theory for the case of real-valued functions, what is sufficient for our purposes. Finally, everything we present here can be found in classical books on functional analysis, like the books by Riesz and Sz.-Nagy [42] or Reed and Simon [40].

By $L^2(S^{n-1})$ we denote the Hilbert space of square-integrable real-valued functions defined over S^{n-1}, with respect to the measure ω and with inner product

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

for f, $g \in L^2(S^{n-1})$. As usual, we write $\|f\| = (f, f)^{1/2}$ for the L^2-norm of f.

Likewise we consider the Hilbert space $L^2(S^{n-1} \times S^{n-1})$, equipped with the inner product

$$\langle A, B\rangle = \int_{S^{n-1}} \int_{S^{n-1}} A(x, y)B(x, y)\, d\omega(x)d\omega(y)$$

for A, $B \in L^2(S^{n-1} \times S^{n-1})$. We also write $\|A\| = \langle A, A\rangle^{1/2}$ for the L^2-norm of A.

The elements of $L^2(S^{n-1} \times S^{n-1})$ are called *Hilbert-Schmidt kernels*, or *kernels* for short. They are much like matrices in many senses. Like a matrix, a kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ also defines a linear operator $A\colon L^2(S^{n-1}) \to L^2(S^{n-1})$. Indeed, if $f \in L^2(S^{n-1})$, we let

$$(Af)(x) = \int_{S^{n-1}} A(x, y)f(y)\, d\omega(y).$$

The integral above exists for almost all $x \in S^{n-1}$ and the function Af so defined is square-integrable. Moreover, if both A and f are continuous, then Af is also continuous.

A kernel A is *symmetric* if $A(x, y) = A(y, x)$ for all x, $y \in S^{n-1}$. In the same way that matrices are said to be positive semidefinite, kernels are said to be positive. More specifically, a kernel A is *positive* if it is symmetric and if for every $p \in L^2(S^{n-1})$ we have

$$(Ap, p) = \int_{S^{n-1}} \int_{S^{n-1}} A(x, y)p(x)p(y)\, d\omega(x)d\omega(y) \geq 0.$$

Bochner [5] made the following observation:

$$\text{if } A\colon S^{n-1}\times S^{n-1}\to\mathbb{R} \text{ is continuous, then } A \text{ is positive if and only if for any choice } x_1,\ldots,x_N \text{ of finitely many points in } S^{n-1}, \text{ the matrix } \big(A(x_i,x_j)\big)_{i,j=1}^{N} \text{ is positive semidefinite.} \tag{3.3}$$

This observation also holds for kernels defined over spaces other than $S^{n-1}\times S^{n-1}$ (like $[-R,R]^n\times[-R,R]^n$), provided they satisfy some simple requirements.

The notion of positivity for a kernel is thus directly analogous to the notion of positive semidefiniteness for a matrix. Our choice of using, in relation to kernels, the term “positive” instead of “positive semidefinite”, is just a matter of following tradition.

Let $A\in L^2(S^{n-1}\times S^{n-1})$ be a kernel. A nonzero function $f\in L^2(S^{n-1})$ is an *eigenfunction* of A if $Af=\lambda f$ for some real number λ. The scalar λ is the *eigenvalue* associated with the eigenfunction f. Every nonzero symmetric Hilbert-Schmidt kernel has a nonzero eigenvalue. Even more strongly, the following spectral decomposition theorem holds:

Theorem 3.1 (Hilbert-Schmidt theorem)**.** *Let $A\in L^2(S^{n-1}\times S^{n-1})$ be a symmetric kernel. Then there is a complete orthonormal system $\varphi_1,\varphi_2,\ldots$ of $L^2(S^{n-1})$ consisting of eigenfunctions of A with associated eigenvalues $\lambda_1,\lambda_2,\ldots$ such that*

$$A(x,y)=\sum_{i=1}^{\infty}\lambda_i\varphi_i(x)\varphi_i(y) \tag{3.4}$$

with convergence according to the L^2-norm. In particular, $\|A\|^2=\sum_{i=1}^{\infty}\lambda_i^2$.

A consequence is that, if $f\in L^2(S^{n-1})$ is any function, then

$$Af=\sum_{i=1}^{\infty}\lambda_i(f,\varphi_i)\varphi_i$$

in the sense of L^2 convergence. This means that there is no eigenfunction of A with nonzero associated eigenvalue that is orthogonal to all the functions φ_i for which $\lambda_i\neq 0$. Observe also that it becomes obvious that A is positive if and only if $\lambda_i\geq 0$ for $i=1,2,\ldots$, i.e., if and only if all its eigenvalues are nonnegative.

The Hilbert-Schmidt theorem also implies that if A and B are two positive kernels, then $\langle A,B\rangle\geq 0$. There is a simpler proof of this fact if both A and B are continuous. In this case we need only notice that the kernel $C(x,y)=A(x,y)B(x,y)$ is also positive. This follows from the analogous statement for positive semidefinite matrices (which can be easily proved with the help of the spectral theorem for matrices) coupled with (3.3). Then $\langle A,B\rangle=(C\mathbf{1},\mathbf{1})\geq 0$, where $\mathbf{1}$ is the constant one function.

If we consider only continuous and positive kernels, we may guarantee absolute and uniform convergence in the development (3.4). This is known as Mercer’s theorem:

Theorem 3.2 (Mercer's theorem)**.** *Let* $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ *be a continuous and positive kernel. Then there is a complete orthonormal system* $\varphi_1, \varphi_2, \ldots$ *of* $L^2(S^{n-1})$ *consisting of continuous eigenfunctions of* A *with nonnegative associated eigenvalues* λ_1, $\lambda_2, \ldots$ *such that*

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

with absolute and uniform convergence over $S^{n-1} \times S^{n-1}$.

Let $A \in L^2(S^{n-1} \times S^{n-1})$ and consider its development (3.4). If the series $\sum_{i=1}^{\infty} \lambda_i$ converges absolutely, we say that A is *trace-class* and that its *trace* is equal to the value of the series. Not all Hilbert-Schmidt kernels are trace class, but Mercer's theorem implies that all continuous and positive kernels are trace-class. Indeed, if A is a continuous and positive kernel, then since we have absolute and uniform convergence for (3.4) and since all eigenvalues are nonnegative we also have

$$\int_{S^{n-1}} A(x, x)\, d\omega(x) = \sum_{i=1}^{\infty} \lambda_i \int_{S^{n-1}} \varphi_i(x)^2\, d\omega(x) = \sum_{i=1}^{\infty} \lambda_i = \sum_{i=1}^{\infty} |\lambda_i|.$$

Notice how in this case the analogy between kernels and matrices is strong, as the trace of a matrix is both the sum of its diagonal elements and the sum of its eigenvalues.

3.3 A generalization of the theta number

Now we have all the notions we need in order to generalize the Lovász theta number to our graph $G(S^{n-1}, D)$, where $D \subseteq [-1, 1)$ is a closed subset of the real line. We define $\vartheta(G(S^{n-1}, D))$ as the optimal value of the following optimization problem:

$$\begin{array}{ll} \sup & \int_{S^{n-1}} \int_{S^{n-1}} A(x, y)\, d\omega(x) d\omega(y) \\ & \int_{S^{n-1}} A(x, x)\, d\omega(x) = 1, \\ & A(x, y) = 0 \quad \text{if } x \cdot y \in D, \\ & A\colon S^{n-1} \times S^{n-1} \to \mathbb{R} \text{ is a continuous and positive kernel.} \end{array} \tag{3.5}$$

This is an infinite-dimensional semidefinite programming problem which is very similar to formulation (2.1) for the theta number of finite graphs. The main property of the theta number, i.e., that it gives an upper bound to the stability number, is also true for our generalization:

Theorem 3.3. *We have* $\vartheta(G(S^{n-1}, D)) \geq \alpha_{\mathrm{m}}(G(S^{n-1}, D))$.

In the proof, it is crucial to use the fact that ω is a regular measure. In a regular measure space, measurable sets can be approximated arbitrarily well by open or closed sets: If C is measurable, then for every $\varepsilon > 0$ there exist sets $X \subseteq C \subseteq Y$, where X is closed and Y is open, such that $\omega(C \setminus X) < \varepsilon$ and $\omega(Y \setminus C) < \varepsilon$.

Proof of Theorem 3.3. The proof is similar to the proof of the analogous statement for finite graphs (cf. Section 2.1), only slightly more complicated because of analysis.

To start we fix $\varepsilon > 0$. Let $C \subseteq S^{n-1}$ be a measurable stable set of $G(S^{n-1}, D)$ such that $\omega(C) \geq \alpha_{\mathrm{m}}(G(S^{n-1}, D)) - \varepsilon$. Since ω is regular, we may assume that C is closed, otherwise we just find a stable set with measure closer to $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$ and consider a suitable inner approximation of it by a closed set.

So C is closed and therefore also compact. Then, since also D is compact, there must be a number $\delta > 0$ such that $|x \cdot y - t| > \delta$ for all $x, y \in C$ and $t \in D$. This means that, for small enough $\xi > 0$, the set

$$B(C, \xi) = \{\, x \in S^{n-1} : \exists\, y \in C \text{ such that } \|x - y\| < \xi \,\}$$

is stable.

Now the function $f\colon S^{n-1} \to [0, 1]$ such that

$$f(x) = \xi^{-1} \cdot \max\{\xi - d(x, C), 0\},$$

where $d(x, C) = \min\{\, \|x - y\| : y \in C \,\}$, is continuous and we have $f(x) = 1$ for $x \in C$ and $f(x) = 0$ for $x \in S^{n-1} \setminus B(C, \xi)$. So the kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ given by

$$A(x, y) = (f, f)^{-1} f(x) f(y)$$

is feasible for (3.5).

Let us now estimate the objective value of A. Recall that $B(C, \xi)$ is stable. Then we have

$$(f, f) \leq \omega(B(C, \xi)) \leq \alpha_{\mathrm{m}}(G(S^{n-1}, D))$$

and

$$\int_{S^{n-1}} \int_{S^{n-1}} f(x) f(y)\, d\omega(x) d\omega(y) \geq \omega(C)^2 \geq (\alpha_{\mathrm{m}}(G(S^{n-1}, D)) - \varepsilon)^2,$$

implying

$$\int_{S^{n-1}} \int_{S^{n-1}} A(x, y)\, d\omega(x) d\omega(y) \geq \frac{(\alpha_{\mathrm{m}}(G(S^{n-1}, D)) - \varepsilon)^2}{\alpha_{\mathrm{m}}(G(S^{n-1}, D))}$$

and, since ε is arbitrary, the theorem follows. ■

The proof above is indeed similar to the proof for finite graphs. In particular, note that the function f we used to define A is nothing more than a suitable continuous approximation of the characteristic function $\chi^C\colon S^{n-1} \to \{0, 1\}$ of the stable set C we chose. This we had to do, since the kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that

$$A(x, y) = \omega(C)^{-1} \chi^C(x) \chi^C(y)$$

is not continuous, even though it satisfies all other constraints of (3.5). Dropping the continuity condition is, however, impossible, as for instance if $D = \{t\}$ we would have constraints $A(x, y) = 0$ for x, y such that $x \cdot y = t$, and these constraints would become trivial without the continuity condition.

Problem (3.5) might provide a bound for $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$, but in the way it is presented it is not clear how to actually solve it. This is similar to the situation we were in after introducing problem (2.6) in order to bound the parameter $A(n, d)$. Similarly to that case, the way out of this problem starts with the following observation: In solving (3.5), one may restrict oneself to kernels that are invariant under the orthogonal group, which is the isometry group of S^{n-1}. We show in the next section how this observation can be used to reduce the problem to a more tractable form.

Finally, notice that our generalization does not use any property specific to the sphere. So it can also be used for other compact metric spaces with regular measures. In Section 4.6b, for instance, we use it for distance graphs on the space $[-R, R]^n$.

3.4 Exploiting symmetry with a theorem of Schoenberg

The *orthogonal group of the* n*-dimensional Euclidean space* is defined as

$$\mathrm{O}(\mathbb{R}^n) = \{\, T \in \mathbb{R}^{n\times n} : T^{\mathsf{T}} T = I \,\},$$

where I is the identity matrix. The surface measure on the sphere is *invariant* under $\mathrm{O}(\mathbb{R}^n)$, i.e., if $X \subseteq S^{n-1}$ is measurable and $T \in \mathrm{O}(\mathbb{R}^n)$, then

$$\omega(T \cdot X) = \omega(\{\, Tx : x \in X \,\}) = \omega(X).$$

This implies that, if $f\colon S^{n-1} \to \mathbb{R}$ is a measurable function and $T \in \mathrm{O}(\mathbb{R}^n)$, then

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

So also the inner product $(\cdot, \cdot)$ is *invariant* under $\mathrm{O}(\mathbb{R}^n)$, that is, for $T \in \mathrm{O}(\mathbb{R}^n)$ and $f, g \in L^2(S^{n-1})$, we have $(T \cdot f, T \cdot g) = (f, g)$, where $(T \cdot f)(x) = f(Tx)$ for all $x \in S^{n-1}$.

A kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ is *invariant* under $\mathrm{O}(\mathbb{R}^n)$ if

$$A(Tx, Ty) = A(x, y) \qquad \text{for all } T \in \mathrm{O}(\mathbb{R}^n) \text{ and } x,\, y \in S^{n-1}.$$

So, if A is invariant, the value of $A(x, y)$ depends only on $x \cdot y$. We claim that, in solving (3.5), we may restrict ourselves to invariant kernels.

To prove this claim we need to use the Haar measure over $\mathrm{O}(\mathbb{R}^n)$. The orthogonal group $\mathrm{O}(\mathbb{R}^n)$ is a topological group with the topology inherited from $\mathbb{R}^{n\times n}$. This means that with this topology it is a topological Hausdorff space and that group operations (in this case, matrix multiplication and inversion) are continuous. Moreover, $\mathrm{O}(\mathbb{R}^n)$ is compact, hence there is a finite Radon measure μ over it which is invariant under the action of the group, i.e., for every $T \in \mathrm{O}(\mathbb{R}^n)$ and measurable $\mathcal{X} \subseteq \mathrm{O}(\mathbb{R}^n)$,

$$\mu(\mathcal{X}) = \mu(\{\, TX : X \in \mathcal{X} \,\}) = \mu(\{\, XT : X \in \mathcal{X} \,\}).$$

This invariant measure is unique up to multiplication by scalars and is called the *Haar measure*. For a proof of the existence of the Haar measure, see the book by Halmos [22]. The book by Mattila [32] contains a discussion of the properties of the Haar measure for the orthogonal group.

The main use of the Haar measure is in invariant integration. If $f\colon \mathrm{O}(\mathbb{R}^n) \to \mathbb{R}$ is any measurable function and X and Y are any two orthogonal matrices, then

$$\int_{\mathrm{O}(\mathbb{R}^n)} f(XTY)\, d\mu(T) = \int_{\mathrm{O}(\mathbb{R}^n)} f(T)\, d\mu(T).$$

Lovász [28] gives a simple, combinatorial way to construct an invariant integration for functions defined over a compact topological group. From such an integration, the existence of the Haar measure can also be derived.

So we can prove our claim that we may restrict ourselves to invariant kernels. Let A be any feasible solution of (3.5). Consider the kernel $\overline{A}\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that

$$\overline{A}(x, y) = \int_{\mathrm{O}(\mathbb{R}^n)} A(Tx, Ty)\, d\mu(T),$$

where μ is the Haar measure over $\mathrm{O}(\mathbb{R}^n)$ normalized such that $\mu(\mathrm{O}(\mathbb{R}^n)) = 1$.

From our discussion above, $\overline{A}$ is invariant under $\mathrm{O}(\mathbb{R}^n)$. With the help of Fubini's theorem, one may also verify that $\overline{A}$ is feasible for (3.5); moreover

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

and from this our claim follows at once.

We may now use a theorem of Schoenberg [44], which characterizes the invariant, positive, and continuous kernels over the sphere in terms of Jacobi polynomials. To state the theorem, we must therefore define the Jacobi polynomials.

Our discussion on the Jacobi polynomials basically follows Szegö [50]. The *Jacobi polynomials* with *parameters* (α, β), α, $\beta > -1$, are the orthogonal polynomials with respect to the weight function $(1-u)^\alpha (1+u)^\beta$ on the interval $[-1, 1]$.

This means that we consider the Hilbert space $L^2([-1, 1])$ with inner product

$$(f, g)_{\alpha,\beta} = \int_{-1}^{1} f(u) g(u)\, (1-u)^\alpha (1+u)^\beta\, du. \tag{3.6}$$

Then we consider the polynomial functions

$$1,\, u,\, u^2,\, u^3,\, \ldots \tag{3.7}$$

and apply the Gram-Schmidt orthogonalization process to them, with respect to the inner product (3.6), in the order of the sequence above.

Since each function in (3.7) is linearly independent of those which come before it in the sequence, from this process we will obtain functions

$$P_0^{(\alpha,\beta)},\, P_1^{(\alpha,\beta)},\, P_2^{(\alpha,\beta)},\, P_3^{(\alpha,\beta)},\, \ldots \tag{3.8}$$

which are by construction polynomials and pairwise orthogonal and such that the linear span of $P_0^{(\alpha,\beta)}, \ldots, P_k^{(\alpha,\beta)}$ is the same as that of $1, \ldots, u^k$. So it must be that $P_k^{(\alpha,\beta)}$ is a polynomial exactly of degree k. Notice, moreover, that during the process we have little choice, so each $P_k^{(\alpha,\beta)}$ is uniquely determined up to scalar multiples.

Finally, since we may write any polynomial as a linear combination of the polynomials $P_k^{(\alpha,\beta)}$, Weierstrass' approximation theorem tells us that (3.8) is actually a complete orthogonal system for $L^2([-1,1])$, i.e., for every $f \in L^2([-1,1])$, we may write

$$f = \sum_{k=0}^{\infty} \lambda_k P_k^{(\alpha,\beta)}$$

with convergence in the norm $\|f\|_{\alpha,\beta} = (f,f)_{\alpha,\beta}^{1/2}$.

The polynomials $P_k^{(\alpha,\beta)}$ are the ones we call Jacobi polynomials with parameters (α,β). They are usually normalized so that

$$P_k^{(\alpha,\beta)}(1) = \binom{k+\alpha}{k} = \frac{\Gamma(k+\alpha+1)}{\Gamma(k+1)\Gamma(\alpha+1)},$$

and we keep this normalization to prevent confusion when alluding to facts concerning Jacobi polynomials which are to be found in external references. Sometimes it is more convenient to use a different normalization, however. Namely we also consider the polynomials $\overline{P}_k^{(\alpha,\beta)}$ which are such that

$$\overline{P}_k^{(\alpha,\beta)}(u) = \frac{P_k^{(\alpha,\beta)}(u)}{P_k^{(\alpha,\beta)}(1)},$$

and hence $\overline{P}_k^{(\alpha,\beta)}(1) = 1$.

Now we are ready to state the following theorem due to Schoenberg [44]:

Theorem 3.4. *Let $\alpha = (n-3)/2$. A kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ is continuous, positive, and invariant under $\mathrm{O}(\mathbb{R}^n)$ if and only if*

$$A(x,y) = \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(x \cdot y) \tag{3.9}$$

for some nonnegative numbers $f_0, f_1, \ldots$ such that $\sum_{k=0}^{\infty} f_k$ converges, in which case the series in (3.9) *converges absolutely and uniformly over $S^{n-1} \times S^{n-1}$.*

This theorem is very similar to Mercer's theorem (Theorem 3.2), the eigenfunctions being hidden in the Jacobi polynomials. This will become clear once we present a proof of Theorem 3.4 in Section 3.8. For now, we stress that the guarantee of absolute and uniform convergence is linked to the continuity and positiveness of the kernel A, as was the case in Mercer's theorem.

Let us now try to better understand the statement of Theorem 3.4 before we use it to deal with problem (3.5). A function $f\colon S^{n-1} \to \mathbb{R}$ is a *zonal spherical function* with *pole* $e \in S^{n-1}$ if $f(x)$ depends only on the inner product between e and x. In other words, $f(x) = f(y)$ if $e \cdot x = e \cdot y$. So a zonal spherical function f can also be seen as a function of one variable defined over the domain $[-1, 1]$. For $u \in [-1, 1]$, we will write $f(u)$ for the common value of f for all $x \in S^{n-1}$ with $e \cdot x = u$.

The weight function $(1-u^2)^{(n-3)/2}$ is obtained from the projection of the surface measure ω of the sphere S^{n-1} onto the interval $[-1, 1]$. More specifically, if we fix a point $e \in S^{n-1}$, then for any measurable subset U of $[-1, 1]$ we have

$$\omega(\{\, x \in S^{n-1} : e \cdot x \in U \,\}) = \int_U \omega_{n-1}(1-u^2)^{(n-3)/2}\,du,$$

where $\omega_{n-1} = \omega(S^{n-2})$. This means that, if $f, g\colon S^{n-1} \to \mathbb{R}$ are square-integrable zonal spherical functions with the same pole e, then

$$(f, g) = \int_{S^{n-1}} f(x)g(x)\,d\omega(x) = \int_{-1}^{1} f(u)g(u)\,\omega_{n-1}(1-u^2)^{(n-3)/2}\,du.$$

For $k = 0, 1, \ldots$, consider the kernel $E_k^n\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that

$$E_k^n(x, y) = \overline{P}_k^{(\alpha,\alpha)}(x \cdot y),$$

where $\alpha = (n-3)/2$. This is an invariant kernel. So, if we fix a point e on the sphere, the function $y \mapsto E_k^n(e, y)$ is a zonal spherical function of pole e. Then, using again the invariance of E_k^n and since the inner product $(\cdot,\cdot)$ is also invariant, we have for $k \neq l$ that

$$\begin{aligned}
\langle E_k^n, E_l^n\rangle &= \int_{S^{n-1}}\int_{S^{n-1}} E_k^n(x, y)E_l^n(x, y)\,d\omega(x)d\omega(y)\\
&= \int_{S^{n-1}}\int_{S^{n-1}} E_k^n(e, y)E_l^n(e, y)\,d\omega(x)d\omega(y)\\
&= \omega_n \int_{S^{n-1}} E_k^n(e, y)E_l^n(e, y)\,d\omega(y)\\
&= \omega_n\omega_{n-1}\int_{-1}^{1} \overline{P}_k^{(\alpha,\alpha)}(u)\overline{P}_l^{(\alpha,\alpha)}(u)\,(1-u^2)^\alpha\,du\\
&= 0,
\end{aligned}$$

since $\overline{P}_k^{(\alpha,\alpha)}$ is orthogonal to $\overline{P}_l^{(\alpha,\alpha)}$ with respect to the weight function $(1-u^2)^\alpha$.

So we see that the kernels E_k^n are pairwise orthogonal, and at once it is clear that the numbers f_k in the right-hand side of (3.9) are given by $f_k = \|E_k^n\|^{-1}\langle A, E_k^n\rangle$.

We may now use Theorem 3.4 to deal with problem (3.5). Suppose we restrict ourselves in (3.5) to invariant kernels. If A is invariant, the constraint

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

is the same as $A(x, x) = \omega_n^{-1}$ for all $x \in S^{n-1}$, since all diagonal entries coincide. The objective function is just $\langle E_0^n, A\rangle$. So, if we write $A = \sum_{k=0}^{\infty} f_k E_k^n$, then since the E_k^n are pairwise orthogonal we have $\langle E_0^n, A\rangle = \omega_n^2 f_0$ for the objective function. So we may rewrite (3.5) equivalently as

$$\begin{array}{lll} \sup & \omega_n^2 f_0 & \\ & \sum_{k=0}^{\infty} f_k = \omega_n^{-1}, & \\ & \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(t) = 0 & \text{for all } t \in D, \\ & f_k \geq 0 & \text{for } k = 0, 1, \ldots. \end{array} \tag{3.10}$$

Above, in the first constraint, we made use of the fact that

$$A(x, x) = \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(x \cdot x) = \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(1) = \sum_{k=0}^{\infty} f_k,$$

since we use the normalization $\overline{P}_k^{(\alpha,\alpha)}(1) = 1$.

Now (3.10) is a linear programming problem, albeit an infinite one, with infinitely many variables and constraints. Solving this problem is relatively simple when $D = \{t\}$ for some $-1 \leq t < 1$, as we discuss in the next section. Also when D is finite can problem (3.10) be used to obtain information about $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$, as we show in Section 3.6.

3.5 Solving the problem for one inner product

We now discuss how to solve problem (3.10) when $D = \{t\}$ for $-1 \leq t < 1$. This is the main case of interest to us because of relation (3.2), which makes a connection between $\chi_{\mathrm{m}}(G(S^{n-1}, t))$ and $\chi_{\mathrm{m}}(\mathbb{R}^n)$.

When $D = \{t\}$, problem (3.10) becomes

$$\begin{array}{lll} \sup & \omega_n^2 f_0 & \\ & \sum_{k=0}^{\infty} f_k = \omega_n^{-1}, & \\ & \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(t) = 0, & \\ & f_k \geq 0 & \text{for } k = 0, 1, \ldots. \end{array} \tag{3.11}$$

For $n \geq 2$ and $-1 \leq t < 1$ we define

$$l_n(t) = \inf\{\, \overline{P}_k^{(\alpha,\alpha)}(t) : k = 0, 1, \ldots \,\}, \tag{3.12}$$

where $\alpha = (n-3)/2$. The infimum above is finite for every $-1 \leq t < 1$, as we show in the next two sections. For $n \geq 3$ we will show later that the infimum is always attained; for $n = 2$ this is not the case for all t.

The first property of $l_n(t)$ we observe is that it is negative for $-1 \leq t < 1$. For suppose not. Then we have $\overline{P}_k^{(\alpha,\alpha)}(t) \geq 0$ for $k = 0, 1, \ldots$. This implies that problem (3.11) is either infeasible, what from Theorem 3.3 we know is not true, or that its optimal value is zero. This case is also impossible, or from Theorem 3.3 we would have that $0 = \vartheta(G(S^{n-1}, t)) \geq \alpha_{\mathrm{m}}(G(S^{n-1}, t)) > 0$, a contradiction.

In the case $D = \{t\}$, one can express $\vartheta(G(S^{n-1}, t))$ in terms of $l_n(t)$ alone, as we show in the next theorem.

Theorem 3.5. *For $n \geq 2$ and $-1 \leq t < 1$ we have*

$$\vartheta(G(S^{n-1}, t) = \frac{\omega_n l_n(t)}{l_n(t) - 1}.$$

If, moreover, the infimum in the definition of $l_n(t)$ is attained, then problem (3.11) *admits an optimal solution.*

Proof. Let $\alpha = (n-3)/2$. The following linear programming problem on two variables and infinitely many constraints is a dual of (3.11):

$$\begin{aligned} \inf\ & \omega_n^{-1} z_0 \\ & z_0 + z_1 \geq \omega_n^2, \\ & z_0 + z_1 \overline{P}_k^{(\alpha,\alpha)}(t) \geq 0 \quad \text{for } k = 1, 2, \ldots. \end{aligned} \tag{3.13}$$

It is easy to show that weak duality holds between (3.11) and (3.13). Indeed, if f_0, $f_1, \ldots$ is a feasible solution of (3.11) and (z_0, z_1) is a feasible solution of (3.13), then

$$\omega_n^2 f_0 \leq \sum_{k=0}^{\infty} f_k(z_0 + z_1 \overline{P}_k^{(\alpha,\alpha)}(t)) = \omega_n^{-1} z_0.$$

One may easily check that

$$z_0 = \frac{\omega_n^2 l_n(t)}{l_n(t) - 1} \qquad \text{and} \qquad z_1 = \frac{\omega_n^2}{1 - l_n(t)}$$

is a feasible solution of (3.13). From weak duality we then get that

$$\vartheta(G(S^{n-1}, t)) \leq \frac{\omega_n l_n(t)}{l_n(t) - 1}. \tag{3.14}$$

We now show how to construct a feasible solution for (3.11). To this end, let k be such that $\overline{P}_k^{(\alpha,\alpha)}(t) < 0$. Then

$$f_0 = \frac{\overline{P}_k^{(\alpha,\alpha)}(t)}{\omega_n(\overline{P}_k^{(\alpha,\alpha)}(t) - 1)} \qquad \text{and} \qquad f_k = \frac{1}{\omega_n(1 - \overline{P}_k^{(\alpha,\alpha)}(t))} \tag{3.15}$$

is a feasible solution of (3.11), and so

$$\vartheta(G(S^{n-1}, t)) \geq \frac{\omega_n \overline{P}_k^{(\alpha,\alpha)}(t)}{\overline{P}_k^{(\alpha,\alpha)}(t) - 1}.$$

Now, by changing k, we may make $\overline{P}_k^{(\alpha,\alpha)}(t)$ as close to $l_n(t)$ as we like. So, using (3.14), we see that

$$\vartheta(G(S^{n-1},t)) = \frac{\omega_n l_n(t)}{l_n(t)-1}.$$

Moreover, if the infimum in the definition of $l_n(t)$ is attained, so that $l_n(t) = \overline{P}_k^{(\alpha,\alpha)}(t)$ for some k, then we see that, for this k in particular, (3.15) is in fact an optimal solution of (3.11), as we wanted. ■

So, in order to compute $\vartheta(G(S^{n-1},t))$ for $-1 \leq t < 1$, one only has to compute $l_n(t)$. In the rest of this section we will study the behavior of $l_n(t)$ and see how it can be computed.

3.5a. The behavior of $l_2(t)$. We start with the case $n = 2$, which is rather pathological. There are basically three possibilities for $l_2(t)$, depending on the value of t. To establish the behavior of $l_2(t)$, we use the fact that

$$\overline{P}_k^{(-\frac{1}{2},-\frac{1}{2})}(\cos\theta) = \cos k\theta \tag{3.16}$$

(cf. (5.1.1) in Andrews, Askey, and Roy [1]).

Now fix $-1 \leq t < 1$ and let $0 < \theta \leq \pi$ be such that $t = \cos\theta$. We distinguish three cases. Suppose first that $\theta = (p/q)\pi$, where p and q are relatively prime natural numbers and p is odd. Then we have $\overline{P}_q^{(-\frac{1}{2},-\frac{1}{2})}(t) = -1$, and so $l_2(t) = -1$ and the infimum is in this case attained.

Suppose now $\theta = (p/q)\pi$ where p and q are relatively prime natural numbers and p is even. In this case we have $\cos k\theta > -1$ for $k = 0, 1, \ldots$. Also, if we have $k \equiv l \pmod q$, then

$$\overline{P}_k^{(-\frac{1}{2},-\frac{1}{2})}(t) = \overline{P}_l^{(-\frac{1}{2},-\frac{1}{2})}(t),$$

and so we see that for some $k = 1, \ldots, q-1$ we have $l_2(t) = \overline{P}_k^{(-\frac{1}{2},-\frac{1}{2})}(t) > -1$.

Finally, suppose θ is not a positive rational multiple of π. In particular, then, the angles $k\theta$, for $k \geq 0$ integer, are all different modulo 2π. One may further prove that in this case the set $\{\theta, 2\theta, 3\theta, \ldots\}$ is dense in the circle, and so $l_2(t) = -1$. The infimum is not attained in this case, though.

The bound given by $\vartheta(G(S^{n-1},t))$ is tight for some values of t, in particular when $t = \cos\pi/p$ for some positive integer p. Then we have $\vartheta(G(S^{n-1},t)) = \pi$, and the union of the open arcs of the unit circle that are of the form $(2j\pi/p, (2j+1)\pi/p)$ for $j = 0, \ldots, p-1$ is a stable set of $G(S^{n-1},t)$ of measure π, as illustrated in Figure 3.1.

For $t = \cos 2\pi/3$, one has $l_2(t) = \cos 2\pi/3$, and so $\vartheta(G(S^{n-1},t)) = 2\pi/3$. Then the bound is also tight, since one may choose an open arc of length $2\pi/3$ as a stable set in $G(S^{n-1},t)$.

For all other values of t, it is not clear whether the bound given by the theta number is tight.

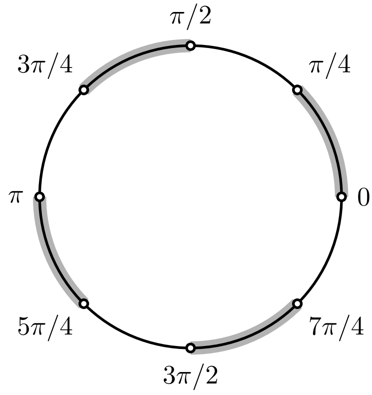

Figure 3.1. The shaded set above, consisting of the union of the open arcs of the form $(2j\pi/4, (2j+1)\pi/4)$ for $j = 0, \ldots, 3$, is a stable set of the graph $G(S^1, \cos \pi/4)$ which has measure π.

3.5b. The basic behavior of $l_n(t)$ for $n \geq 3$. From now on we assume $n \geq 3$. Recall that we use two normalizations for the Jacobi polynomials. The Jacobi polynomials $P_k^{(\alpha,\beta)}$ are normalized so that

$$P_k^{(\alpha,\beta)}(1) = \binom{k+\alpha}{k} = \frac{\Gamma(k+\alpha+1)}{\Gamma(k+1)\Gamma(\alpha+1)}. \tag{3.17}$$

This is the normalization commonly found in the literature about Jacobi polynomials. We also consider a different normalization for the Jacobi polynomials, namely the polynomials $\overline{P}_k^{(\alpha,\beta)}$ which are such that

$$\overline{P}_k^{(\alpha,\beta)}(u) = \frac{P_k^{(\alpha,\beta)}(u)}{P_k^{(\alpha,\beta)}(1)}.$$

We first establish that the infimum in the definition (3.12) of $l_n(t)$ is attained for all $-1 \leq t < 1$. Indeed, let $\alpha = (n-3)/2$. Since $\overline{P}_k^{(\alpha,\alpha)}(1) = 1$ for all k, using the formula

$$P_k^{(\alpha,\beta)}(-u) = (-1)^k P_k^{(\beta,\alpha)}(u) \tag{3.18}$$

(cf. (6.4.23) in Andrews, Askey, and Roy [1]) we see that $l_n(-1) = -1$.

For $-1 < t < 1$, the existence of $l_n(t)$ for $n \geq 3$ follows immediately from the fact that, for $\alpha \geq 0$ and $-1 < t < 1$,

$$\lim_{k\to\infty} \overline{P}_k^{(\alpha,\beta)}(t) = 0. \tag{3.19}$$

This follows from Theorem 8.21.8 in Szegö [50] together with (3.17). The theorem provides an asymptotic formula for the polynomials $P_k^{(\alpha,\beta)}$ from which it follows

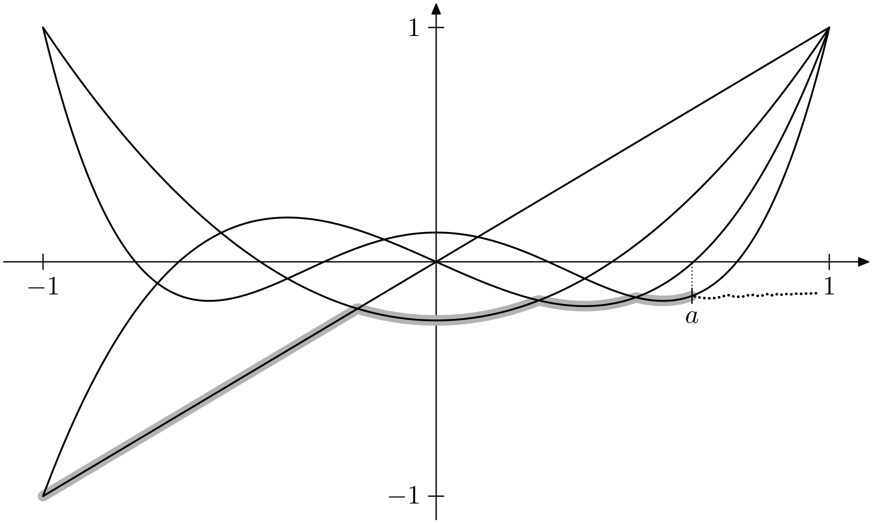

Figure 3.2. In the subinterval $[-1, a]$ of $[-1, 1]$, function l_5 is the minimum of the Jacobi polynomials $\overline{P}_k^{(1,1)}$ for $k = 1, \dots, 4$. These polynomials are plotted above, and the graph of l_5 in $[-1, a]$ is highlighted in gray. The plot of the graph of l_5 continues as a dotted line in the interval $[a, 1)$.

that $P_k^{(\alpha,\beta)}(t) \to 0$ as $k \to \infty$ for any choice of α, $\beta > -1$ and any fixed number $-1 < t < 1$. From (3.17) we see that, for $\alpha \geq 0$, we have $P_k^{(\alpha,\beta)}(1) \geq 1$, so together with the asymptotic formula we have (3.19).

The asymptotic formula for $P_k^{(\alpha,\beta)}$ given in Theorem 8.21.8 of Szegö [50] is uniform in any interval $[a, b] \subseteq (-1, 1)$. So for any $[a, b] \subseteq (-1, 1)$, (3.19) holds uniformly for all $t \in [a, b]$. This implies that, for every interval $[a, b] \subseteq (-1, 1)$, there is a number k_0 such that, for every $t \in [a, b]$,

$$l_n(t) = \min\{\, \overline{P}_k^{(\alpha,\alpha)}(t) : k = 0, \dots, k_0 \,\}, \tag{3.20}$$

where $\alpha = (n-3)/2$. So we see that $l_n(t)$ as a function of t is continuous in $[-1, 1)$. This is so because, in any interval $[a, b] \subseteq (-1, 1)$, function $l_n(t)$ is the minimum of finitely many continuous functions, what implies that $l_n(t)$ is continuous in $(-1, 1)$. The continuity of $l_n(t)$ at -1 follows from the fact that $l_n(-1) = -1$ and $\overline{P}_1^{(\alpha,\alpha)}(u) = u$, as we see from (3.18). Figure 3.2 shows the graphs of the first few Jacobi polynomials $\overline{P}_k^{(\alpha,\alpha)}$ for $\alpha = (n-3)/2$ with $n = 5$, together with the graph of $l_5(t)$.

We now come to the problem of computing $l_n(t)$ for a given $-1 < t < 1$. The numbers $\overline{P}_k^{(\alpha,\alpha)}(t)$, for $k = 0, 1, \dots$, can be computed with the help of the recurrence formula for Jacobi polynomials (formula (4.5.1) in Szegö [50]). This formula, adapted to the normalization of the polynomials $\overline{P}_k^{(\alpha,\alpha)}$, is

$$\overline{P}_k^{(\alpha,\alpha)}(u) = \frac{2k+2\alpha-1}{k+2\alpha} u \overline{P}_{k-1}^{(\alpha,\alpha)}(u) - \frac{k-1}{k+2\alpha} \overline{P}_{k-2}^{(\alpha,\alpha)}(u), \quad k \geq 2, \\ \overline{P}_1^{(\alpha,\alpha)}(u) = u, \qquad \overline{P}_0^{(\alpha,\alpha)}(u) = 1. \tag{3.21}$$

To find $l_n(t)$ we know from (3.20) that we only need to compute $\overline{P}_k^{(\alpha,\alpha)}(t)$, where again $\alpha = (n-3)/2$, for $k = 0, \ldots, k_0$ for some k_0 and take the minimum. The asymptotic formula of Theorem 8.21.8 of Szegö [50], however, does not provide precise error estimates, so it does not help us in finding k_0.

To estimate k_0, and therefore know when we may stop computing $\overline{P}_k^{(\alpha,\alpha)}(t)$ in order to find $l_n(t)$, we may use a modified version of a formula by Feldheim and Vilenkin (cf. Corollary 6.7.3 in Andrews, Askey, and Roy [1]). This formula will allow us to estimate $|\overline{P}_k^{(\alpha,\alpha)}(t)|$ for $\alpha \geq 0$ and $k \geq k_0$ for any given k_0. In Corollary 6.7.3 of Andrews, Askey, and Roy [1] the formula is stated in terms of the *ultraspherical polynomials* C_k^λ, which are such that

$$C_k^\lambda(u) = \frac{\Gamma(\lambda+1/2)\Gamma(k+2\lambda)}{\Gamma(2\lambda)\Gamma(k+\lambda+1/2)} P_k^{(\lambda-\frac{1}{2},\lambda-\frac{1}{2})}(u). \tag{3.22}$$

If in the formula we take $\nu = \alpha + 1/2$ and $\lambda = 0$ we then get, for $\alpha \geq 0$,

$$\overline{P}_k^{(\alpha,\alpha)}(\cos\theta) = \frac{2\Gamma(\alpha+1)}{\Gamma(\frac{1}{2})\Gamma(\alpha+\frac{1}{2})} \int_0^{\pi/2} \cos^{2\alpha}\phi\,(1-\sin^2\theta\cos^2\phi)^{k/2} \\ \cdot p_k(\theta,\phi)\,d\phi, \tag{3.23}$$

where $p_k(\theta,\phi) = \overline{P}_k^{(-\frac{1}{2},-\frac{1}{2})}(\cos\theta(1-\sin^2\theta\cos^2\phi)^{-1/2})$.

So let us see how to use this formula in order to estimate $|\overline{P}_k^{(\alpha,\alpha)}(\cos\theta)|$ for all $k \geq k_0$ and fixed $0 < \theta < \pi$. Fix θ and let I denote the integral in (3.23). From the Cauchy-Schwarz inequality we know that

$$|I|^2 \leq \int_0^{\pi/2} \cos^{4\alpha}\phi\,(1-\sin^2\theta\cos^2\phi)^k\,d\phi \cdot \int_0^{\pi/2} (p_k(\theta,\phi))^2\,d\phi. \tag{3.24}$$

From (3.16) we see that the second integral above is at most $\pi/2$. So let us look at the integral

$$\int_0^{\pi/2} \cos^{4\alpha}\phi\,(1-\sin^2\theta\cos^2\phi)^k\,d\phi. \tag{3.25}$$

As $0 < \theta < \pi$, we have $\sin^2\theta\cos^2\phi > 0$ for $0 \leq \phi < \pi/2$. So, as $k \to \infty$, this integral tends to zero. This also holds uniformly for all θ contained in any interval $[a,b] \subseteq (0,\pi)$. Notice this gives another proof of (3.19) for the case $\alpha = \beta$ and also an alternative proof of the existence of a number k_0 such that (3.20) holds. So we also have another proof of the continuity of $l_n(t)$ for $n \geq 3$.

More important though is the fact that we may upper bound (3.25) for any given k by using, for instance, some numerical integration method. So, since (3.25), as a function of k and for fixed $0 < \theta < \pi$, is decreasing, by upper bounding (3.25) for some k_0 we get from (3.23) and (3.24) an upper bound for $|\overline{P}_k^{(\alpha,\alpha)}(\cos\theta)|$ for all $k \geq k_0$.

As an example, say we wish to compute $l_4(0.999)$. We begin by computing $\overline{P}_k^{(\frac{1}{2},\frac{1}{2})}(0.999)$ for $k = 0, \ldots, 3500$ and then taking the minimum of all these numbers, which is attained at $k = 99$. The minimum is a rational number, as is clear from (3.21), and its first few digits are -0.217258. Now we use (3.23) and (3.24) to upper bound $|\overline{P}_k^{(\frac{1}{2},\frac{1}{2})}(0.999)|$ for $k \geq k_0$. If we take $k_0 = 3500$, we see that

$$|\overline{P}_k^{(\frac{1}{2},\frac{1}{2})}(0.999)| \leq 0.2090\ldots$$

for all $k \geq k_0$. So it becomes clear that $l_4(0.999) = \overline{P}_{99}^{(\frac{1}{2},\frac{1}{2})}(0.999)$.

We have observed in practice that, for any given $-1 < t < 1$ and $\alpha \geq 0$, the sequence

$$\overline{P}_0^{(\alpha,\alpha)}(t), \overline{P}_1^{(\alpha,\alpha)}(t), \overline{P}_2^{(\alpha,\alpha)}(t), \ldots$$

is decreasing until $\min\{\, \overline{P}_k^{(\alpha,\alpha)}(t) \;:\; k = 0, 1, \ldots \,\}$ is found. We can only prove this is the case for some values of t, however. Notice that, if this were true for all $-1 < t < 1$, then computing $l_n(t)$ would be much simpler. We present this and other questions about Jacobi polynomials in Section 3.5d.

3.5c. A formula for the limit of $l_n(t)$ as $t \to 1$. The aim of this section is to prove the following theorem:

Theorem 3.6. *For $n \geq 3$ we have*

$$\lim_{t\to 1} l_n(t) = \Gamma(\alpha+1)\Big(\frac{2}{j_{\alpha+1,1}}\Big)^{\alpha} J_\alpha(j_{\alpha+1,1}),$$

where $\alpha = (n-3)/2$, J_α is the Bessel function of the first kind of order α, and $j_{\alpha+1,1}$ is the first positive zero of $J_{\alpha+1}$, the Bessel function of the first kind of order $\alpha+1$.

Bessel functions will play a much more prominent role in Chapter 4, here we only need a few facts relating Bessel functions of the first kind to the Jacobi polynomials. Background material on Bessel functions is provided in Section 4.3. This material is, however, not necessary for understanding the rest of this section.

From Theorem 3.6 one may derive the following corollary:

Corollary 3.7. *For $n \geq 3$, setting $\alpha = (n-3)/2$, we have*

$$\chi_{\mathrm{m}}(\mathbb{R}^n) \geq 1 - \frac{1}{\Gamma(\alpha+1)}\Big(\frac{j_{\alpha+1,1}}{2}\Big)^{\alpha}\frac{1}{J_\alpha(j_{\alpha+1,1})}.$$

Proof. We use Theorem 3.6 together with Theorem 3.5 to see that

$$\lim_{t\to 1}\frac{\omega_n}{\vartheta(G(S^{n-1},t))} = 1 - \frac{1}{\Gamma(\alpha+1)}\Big(\frac{j_{\alpha+1,1}}{2}\Big)^{\alpha}\frac{1}{J_\alpha(j_{\alpha+1,1})}.$$

Now we apply Theorem 3.3, (3.1), and (3.2), finishing the proof. ■

n	Previous lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$	Lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ from Corollary 3.7
3	6	4
4	8	6
5	11	9
6	15	13
7	19	19
8	30	26
9	35	35
10	45	48
11	56	64
12	70	85
13	84	113
14	102	147
15	119	191
16	148	248
17	174	319
18	194	408
19	263	521
20	315	663
21	374	840
22	526	1,061
23	754	1,336
24	933	1,679

Table 3.3. The table shows the best lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ previously known together with the lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ from Corollary 3.7. The lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^3)$ was given by Falconer [13]; all other lower bounds were given by Székely and Wormald [54]. Notice that we have improvements for $n = 10, \ldots, 24$.

Bessel functions may be numerically evaluated to any desired accuracy with the help of a computer; the zeros of Bessel functions can also be computed to any desired accuracy (cf. Section 4.3). So we may use Corollary 3.7 to compute lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$. By doing so we were able to improve on the best previously known lower bounds for $n = 10, \ldots, 24$. Table 3.3 shows the numbers we obtained compared with the best previously known lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$; notice that, since $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is integer, the table actually shows the ceiling of the lower bound provided by Corollary 3.7.

From Corollary 3.7 one may also derive an exponential lower bound for the growth of $\chi_{\mathrm{m}}(\mathbb{R}^n)$ as a function of n. Namely, it is possible to prove that

$$\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \Big(\frac{2}{e} + o(1)\Big)^{-n/2} = (1.165\ldots + o(1))^n.$$

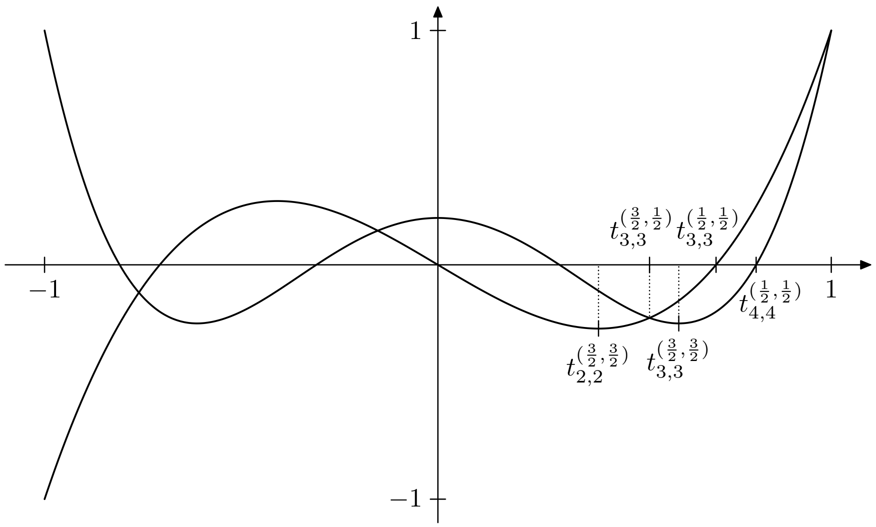

Figure 3.4. Here we have the graphs of $\overline{P}_3^{(\frac{1}{2},\frac{1}{2})}$ and $\overline{P}_4^{(\frac{1}{2},\frac{1}{2})}$. This plot illustrates some properties of the Jacobi polynomials, in particular the interlacing property and identities (3.26) and (3.27).

In Section 4.4 we obtain the same lower bound from fundamentally the same formula given in Corollary 3.7. To obtain this lower bound, more properties of Bessel functions and their zeros are necessary, so we defer further discussion on this issue to Section 4.4. We observe however that our lower bound, though exponential, is not better than the best known asymptotic lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$, which were discussed in Section 1.2.

We now set out to prove Theorem 3.6. Recall that we assume $n \geq 3$. We begin with a characterization of $l_n(t)$ for some specific values of t. To present this characterization we first state some properties of the Jacobi polynomials which we will need. Some of these properties hold for more general families of orthogonal polynomials, but we state them for the case of Jacobi polynomials. In reading this section, it is useful to keep Figure 3.4 in mind; it illustrates some of the properties of the Jacobi polynomials we discuss below.

First, all k zeros of $P_k^{(\alpha,\beta)}$ are real and distinct and are contained in the interval $(-1, 1)$ (cf. Theorem 3.3.1 in Szegö [50]). The zeros of the Jacobi polynomials also satisfy the *interlacing property*. Let $t_{k,1}^{(\alpha,\beta)} < \cdots < t_{k,k}^{(\alpha,\beta)}$ denote the zeros of $P_k^{(\alpha,\beta)}$ in ascending order and set $t_{k,0}^{(\alpha,\beta)} = -1$ and $t_{k,k+1}^{(\alpha,\beta)} = 1$. Then the interlacing property says that in each interval $(t_{k,i}^{(\alpha,\beta)}, t_{k,i+1}^{(\alpha,\beta)})$, for $i = 0, \ldots, k$, there is exactly one zero of $P_{k+1}^{(\alpha,\beta)}$ (cf. Theorem 3.3.2 in Szegö [50]).

Since all the zeros of $P_k^{(\alpha,\beta)}$ lie in the interval $(-1, 1)$ and are real and distinct and since we know from (3.17) that $P_k^{(\alpha,\beta)}(1) > 0$, it must be that the rightmost extremum z of $P_k^{(\alpha,\beta)}$ is a minimum and that $P_k^{(\alpha,\beta)}(z) < 0$. It is actually the case that the rightmost extremum of $P_k^{(\alpha,\beta)}$ occurs at $t_{k-1,k-1}^{(\alpha+1,\beta+1)}$, the rightmost zero of $P_{k-1}^{(\alpha+1,\beta+1)}$. This is so since

$$\frac{dP_k^{(\alpha,\beta)}(u)}{du} = \frac{(k+\alpha+\beta+1)}{2} P_{k-1}^{(\alpha+1,\beta+1)}(u) \tag{3.26}$$

(cf. (6.3.8) in Andrews, Askey, and Roy [1]), so in fact we see that the extrema of $P_k^{(\alpha,\beta)}$ coincide with the zeros of $P_{k-1}^{(\alpha+1,\beta+1)}$.

Fix $\alpha \geq 0$. For some values of t, we know exactly which polynomial achieves the minimum $\min\{\,\overline{P}_j^{(\alpha,\alpha)}(t) : j = 0, 1, \ldots\,\}$. Namely, if t is the rightmost extremum of $\overline{P}_k^{(\alpha,\alpha)}$, which is a minimum, then we know that the minimum above is equal to $\overline{P}_k^{(\alpha,\alpha)}(t)$, as we state in the next theorem.

Theorem 3.8. *For $\alpha \geq 0$ and $k \geq 2$ we have*

$$\min\{\,\overline{P}_j^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}) : j = 0, 1, \ldots\,\} = \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}).$$

In particular, for $n \geq 3$, $\alpha = (n-3)/2$, and $k \geq 2$ we have

$$l_n(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}) = \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}).$$

To prove the theorem we give some further facts about Jacobi polynomials that we will need and which will also be of use later on. We begin by observing that the points where $\overline{P}_k^{(\alpha,\alpha)}$ and $\overline{P}_{k+1}^{(\alpha,\alpha)}$ coincide in the interval $(-1,1)$ correspond to the zeros of $P_k^{(\alpha+1,\alpha)}$, as is clear from the formula

$$(k+\alpha+1)(1-u)\overline{P}_k^{(\alpha+1,\alpha)}(u) = (\alpha+1)(\overline{P}_k^{(\alpha,\alpha)}(u) - \overline{P}_{k+1}^{(\alpha,\alpha)}(u)), \tag{3.27}$$

which is adapted to the normalization of $\overline{P}_k^{(\alpha,\alpha)}$ from (6.4.20) in Andrews, Askey, and Roy [1].

We now show that

the rightmost zero of $P_k^{(\alpha+1,\alpha)}$ in the interval $[-1,1]$ is the only zero of $P_k^{(\alpha+1,\alpha)}$ that lies in the interval $(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}, t_{k,k}^{(\alpha+1,\alpha+1)})$. (3.28)

To show this we need the identity

$$P_k^{(\alpha+1,\alpha)} = \frac{k+2\alpha+2}{2k+2\alpha+2} P_k^{(\alpha+1,\alpha+1)} + \frac{k+\alpha+1}{2k+2\alpha+2} P_{k-1}^{(\alpha+1,\alpha+1)}, \tag{3.29}$$

which follows from (6.4.21) in Andrews, Askey, and Roy [1] by taking $\beta = \alpha + 1$ and applying (3.18).

Now, from (3.29) and the interlacing property, one sees that $P_k^{(\alpha+1,\alpha)}$ has different signs at the endpoints of the intervals $(t_{k-1,i-1}^{(\alpha+1,\alpha+1)}, t_{k,i}^{(\alpha+1,\alpha+1)})$ for $i = 1, \ldots, k$, where we take $t_{k-1,0}^{(\alpha+1,\alpha+1)} = -1$. So there is a zero of $P_k^{(\alpha+1,\alpha)}$ in each such interval. But then, since $P_k^{(\alpha+1,\alpha)}$ has exactly k zeros, it must be that there is exactly one zero of $P_k^{(\alpha+1,\alpha)}$ in each of the k intervals, and (3.28) follows.

Finally, we need two properties of the rightmost extremum of $\overline{P}_k^{(\alpha,\alpha)}$. The first one is given in (6.4.19) of Andrews, Askey, and Roy [1] and implies, in particular, that for $\alpha > -1/2$,

$$|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)})| > |\overline{P}_{k+1}^{(\alpha,\alpha)}(t_{k,k}^{(\alpha+1,\alpha+1)})| \qquad \text{for } k = 2, 3, \ldots, \tag{3.30}$$

i.e., the rightmost extremum of $\overline{P}_k^{(\alpha,\alpha)}$ decreases in absolute value with k.

The second property is stated in (6.4.24) of Andrews, Askey, and Roy [1] but only for the case $\alpha = 0$. It says that

$$\min\{\, \overline{P}_k^{(\alpha,\alpha)}(u) : u \in [0,1]\,\} = \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}) \qquad \text{for } k = 2, 3, \ldots, \tag{3.31}$$

i.e., the global minimum of $\overline{P}_k^{(\alpha,\alpha)}$ in $[0,1]$ is attained at $t_{k-1,k-1}^{(\alpha+1,\alpha+1)}$, its rightmost extremum in $[0,1]$.

To prove (3.31) we first consider the function

$$g(u) = (\overline{P}_k^{(\alpha,\alpha)}(u))^2 + \frac{1-u^2}{k(k+2\alpha+1)}\left(\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du}\right)^2.$$

We then compute g' and use the identity

$$(1-u^2)\frac{d^2\overline{P}_k^{(\alpha,\alpha)}(u)}{du^2} - (2\alpha+2)u\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du} + k(k+2\alpha+1)\overline{P}_k^{(\alpha,\alpha)}(u) = 0,$$

which is adapted from (6.3.9) in Andrews, Askey, and Roy [1] to the normalization of $\overline{P}_k^{(\alpha,\alpha)}$, to obtain

$$g'(u) = \frac{(4\alpha+2)u}{k(k+2\alpha+1)}\left(\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du}\right)^2.$$

So we see that the polynomial g' takes positive values on $(0,1]$ and hence g is increasing on $[0,1]$. So we have

$$g(t_{k-1,i}^{(\alpha+1,\alpha+1)}) < g(t_{k-1,i+1}^{(\alpha+1,\alpha+1)})$$

whenever $i < k-1$ and $t_{k-1,i}^{(\alpha+1,\alpha+1)} \geq 0$. By using (3.26) we then get

$$(\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,i}^{(\alpha+1,\alpha+1)}))^2 < (\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,i+1}^{(\alpha+1,\alpha+1)}))^2$$

whenever $i < k-1$ and $t_{k-1,i}^{(\alpha+1,\alpha+1)} \geq 0$.

Now, we see from (3.18) and (3.26) that, when k is odd, then $\overline{P}_k^{(\alpha,\alpha)}(0) = 0$, and when k is even, then 0 is an extremum of $\overline{P}_k^{(\alpha,\alpha)}$. So, since $t_{k-1,i}^{(\alpha+1,\alpha+1)}$ are the extrema of $\overline{P}_k^{(\alpha,\alpha)}$, it is clear that (3.31) follows.

Proof of Theorem 3.8. Let $t = t_{k-1,k-1}^{(\alpha+1,\alpha+1)}$. We first show that the sequence

$$\overline{P}_0^{(\alpha,\alpha)}(t), \ldots, \overline{P}_k^{(\alpha,\alpha)}(t)$$

is decreasing. Indeed, for $j < k$, we have

$$t_{j,j}^{(\alpha+1,\alpha)} \leq t_{k-1,k-1}^{(\alpha+1,\alpha)} < t_{k-1,k-1}^{(\alpha+1,\alpha+1)} = t.$$

The first inequality above comes from the interlacing property and the second inequality follows from (3.28). So t lies to the right of the rightmost zero of $\overline{P}_j^{(\alpha+1,\alpha)}$, hence $\overline{P}_j^{(\alpha+1,\alpha)}(t) > 0$. Then it is clear from (3.27) that $\overline{P}_j^{(\alpha,\alpha)}(t) > \overline{P}_{j+1}^{(\alpha,\alpha)}(t)$, as we wanted.

Now, for $j > k$, we also have $\overline{P}_j^{(\alpha,\alpha)}(t) > \overline{P}_k^{(\alpha,\alpha)}(t)$. Indeed, from (3.30) and (3.31) we know that

$$\overline{P}_k^{(\alpha,\alpha)}(t) < \overline{P}_j^{(\alpha,\alpha)}(t_{j-1,j-1}^{(\alpha+1,\alpha+1)}) = \min\{\, \overline{P}_j^{(\alpha,\alpha)}(u) : u \in [0,1]\,\}$$

for all $j > k$. This finishes the proof. ■

An interesting corollary of Theorem 3.8 is the following:

Corollary 3.9. *For $n \geq 3$ and $k = 2, 3, \ldots,$ we have $l_n(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}) < l_n(t)$ for all $t > t_{k-1,k-1}^{(\alpha+1,\alpha+1)}$, where $\alpha = (n-3)/2$.*

Proof. Follows from Theorem 3.8 together with (3.30) and (3.31). ■

There are two things left for a proof of Theorem 3.6. The first is the following simple observation: We have

$$l_n(t) \leq \overline{P}_{k+1}^{(\alpha,\alpha)}(t_{k,k}^{(\alpha+1,\alpha)}) \;\text{ for }\; t \in [t_{k-1,k-1}^{(\alpha+1,\alpha+1)}, t_{k,k}^{(\alpha+1,\alpha+1)}] \;\text{ and }\; k \geq 2, \tag{3.32}$$

where $\alpha = (n-3)/2$. Indeed, we know from (3.28) that $t_{k,k}^{(\alpha+1,\alpha)}$ is the only zero of $\overline{P}_k^{(\alpha+1,\alpha)}$ that lies in the interval $(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}, t_{k,k}^{(\alpha+1,\alpha+1)})$. Hence by (3.27) it is clear that $t_{k,k}^{(\alpha+1,\alpha)}$ is the only point in this interval where $\overline{P}_k^{(\alpha,\alpha)}$ equals $\overline{P}_{k+1}^{(\alpha,\alpha)}$. From the interlacing property applied to $\overline{P}_{k-1}^{(\alpha+1,\alpha+1)}$ and $\overline{P}_k^{(\alpha+1,\alpha+1)}$ together with (3.26), we know that $\overline{P}_k^{(\alpha,\alpha)}$ is increasing in $[t_{k-1,k-1}^{(\alpha+1,\alpha+1)}, t_{k,k}^{(\alpha+1,\alpha+1)}]$, whereas $\overline{P}_{k+1}^{(\alpha,\alpha)}$ is decreasing in this interval. So we have that

$$\min\{\overline{P}_k^{(\alpha,\alpha)}(t), \overline{P}_{k+1}^{(\alpha,\alpha)}(t)\} \leq \overline{P}_k^{(\alpha,\alpha)}(t_{k,k}^{(\alpha+1,\alpha)}) = \overline{P}_{k+1}^{(\alpha,\alpha)}(t_{k,k}^{(\alpha+1,\alpha)})$$

for all $t \in [t_{k-1,k-1}^{(\alpha+1,\alpha+1)}, t_{k,k}^{(\alpha+1,\alpha+1)}]$, and so with the definition of $l_n(t)$ we have (3.32).

The last ingredient we need for the proof of Theorem 3.6 is the following result:

Theorem 3.10. *For $\alpha > -1/2$ and $\beta > -1$ we have that, with the notation of Theorem 3.6,*

$$\lim_{k\to\infty} \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\beta)}) = \Gamma(\alpha+1)\Big(\frac{2}{j_{\alpha+1,1}}\Big)^{\alpha} J_\alpha(j_{\alpha+1,1}).$$

A very similar result is proven by Wong and Zhang [57] (see (1.6) in their paper), but since they are interested in providing asymptotic formulas, their proof is considerably more complicated than the one we present. Before presenting the proof, however, we first use this theorem to prove Theorem 3.6.

Proof of Theorem 3.6. From Theorem 3.10 we know that both limits

$$\lim_{k\to\infty} \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha+1)}) \qquad \text{and} \qquad \lim_{k\to\infty} \overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha)})$$

are equal to

$$\Gamma(\alpha+1)\Big(\frac{2}{j_{\alpha+1,1}}\Big)^{\alpha} J_\alpha(j_{\alpha+1,1}),$$

where $\alpha = (n-3)/2$. Theorem 6.1.1 in Szegö [50] implies that the zeros of $\overline{P}_k^{(\alpha,\beta)}$ are dense in $[-1,1]$. From this and the interlacing property we see that both sequences $(t_{k,k}^{(\alpha+1,\alpha+1)})_{k\geq 0}$ and $(t_{k,k}^{(\alpha+1,\alpha)})_{k\geq 0}$ tend to 1 as $k \to \infty$. So, by using Corollary 3.9, (3.32), and Theorem 3.8, we are done. ■

We finish this section with a proof of Theorem 3.10.

Proof of Theorem 3.10. In our proof we need two facts which relate Bessel functions of the first kind with Jacobi polynomials. The first fact we state relates the rightmost zero of the polynomial $P_k^{(\alpha,\beta)}$ with the first positive zero of the Bessel function of the first kind J_α of order α. For $k = 1, 2, \ldots$, let $0 < \theta_k < \pi$ be such that $\cos\theta_k = t_{k,k}^{(\alpha,\beta)}$. Then (cf. Theorem 4.14.1 in Andrews, Askey, and Roy [1])

$$\lim_{k\to\infty} k\theta_k = j_{\alpha,1}, \tag{3.33}$$

where $j_{\alpha,1}$ is the first positive zero of J_α.

The second fact we need is the following formula, adapted from Theorem 4.11.6 in Andrews, Askey, and Roy [1] to take into account the normalization of $\overline{P}_k^{(\alpha,\alpha)}$:

$$\lim_{k\to\infty} \overline{P}_k^{(\alpha,\alpha)}\Big(\cos\frac{u}{k}\Big) = \Gamma(\alpha+1)\Big(\frac{2}{u}\Big)^{\alpha} J_\alpha(u). \tag{3.34}$$

Now, to prove the theorem, we estimate the difference

$$\Big|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\beta)}) - \Gamma(\alpha+1)\Big(\frac{2}{j_{\alpha+1,1}}\Big)^{\alpha} J_\alpha(j_{\alpha+1,1})\Big|,$$

which we upper bound by

$$\begin{aligned}&\Big|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\beta)}) - \overline{P}_k^{(\alpha,\alpha)}\Big(\cos\frac{j_{\alpha+1,1}}{k}\Big)\Big| \\ &\qquad + \Big|\overline{P}_k^{(\alpha,\alpha)}\Big(\cos\frac{j_{\alpha+1,1}}{k}\Big) - \Gamma(\alpha+1)\Big(\frac{2}{j_{\alpha+1,1}}\Big)^{\alpha} J_\alpha(j_{\alpha+1,1})\Big|.\end{aligned}$$

From (3.34) we see that the second term tends to 0 as $k \to \infty$. So we estimate the first term. Let θ_{k-1} be such that $\cos\theta_{k-1} = t_{k-1,k-1}^{(\alpha+1,\beta)}$ and $0 < \theta_{k-1} < \pi$. Applying the mean value theorem twice we get

$$\begin{aligned}
&\Big|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\beta)}) - \overline{P}_k^{(\alpha,\alpha)}\Big(\cos\frac{j_{\alpha+1,1}}{k}\Big)\Big| \\
&\qquad \le \Big(\max_{u\in[-1,1]}\Big|\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du}\Big|\Big)\Big|\cos\theta_{k-1} - \cos\frac{j_{\alpha+1,1}}{k}\Big| \\
&\qquad \le \Big(\max_{u\in[-1,1]}\Big|\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du}\Big|\Big)\Big(\max_{\theta\in I_k}|\sin\theta|\Big)\Big|\theta_{k-1} - \frac{j_{\alpha+1,1}}{k}\Big|,
\end{aligned} \tag{3.35}$$

where I_k is the interval with extremes θ_{k-1} and $j_{\alpha+1,1}/k$.

Then, by using (3.33) we obtain

$$\begin{aligned}
\Big|\theta_{k-1} - \frac{j_{\alpha+1,1}}{k}\Big| &= \Big|\theta_{k-1} - \frac{j_{\alpha+1,1}}{k-1} + \frac{j_{\alpha+1,1}}{k(k-1)}\Big| \\
&\le \Big|\frac{(k-1)\theta_{k-1} - j_{\alpha+1,1}}{k-1}\Big| + \Big|\frac{j_{\alpha+1,1}}{k(k-1)}\Big| \\
&= o(1/k).
\end{aligned} \tag{3.36}$$

So for $\theta \in I_k$ we have

$$|\sin\theta| \le |\theta| \le \frac{j_{\alpha+1,1}}{k} + \Big|\theta_{k-1} - \frac{j_{\alpha+1,1}}{k}\Big| = O(1/k). \tag{3.37}$$

Now, if we adapt (3.26) to the normalization of $\overline{P}_k^{(\alpha,\alpha)}$ we get the formula

$$\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du} = \frac{k(k+2\alpha+1)}{2\alpha+2}\overline{P}_{k-1}^{(\alpha+1,\alpha+1)}(u). \tag{3.38}$$

Also, for $\alpha > -1/2$, we have $|\overline{P}_k^{(\alpha,\alpha)}(u)| \le 1$ for $-1 \le u \le 1$; this is stated in p. 302 of Andrews, Askey, and Roy [1] in terms of the ultraspherical polynomials, which we defined in (3.22). Then it is clear from (3.38) that in the interval $[-1,1]$ we have

$$\Big|\frac{d\overline{P}_k^{(\alpha,\alpha)}(u)}{du}\Big| = O(k^2),$$

and this together with (3.35), (3.36), and (3.37) implies that

$$\lim_{k\to\infty}\Big|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\beta)}) - \overline{P}_k^{(\alpha,\alpha)}\Big(\cos\frac{j_{\alpha+1,1}}{k}\Big)\Big| = 0,$$

as we wanted. ■

3.5d. Open questions. We now give some open questions regarding the Jacobi polynomials and their relation to the function l_n and to $\vartheta(G(S^{n-1},t))$.

In Section 3.5b we discussed how to compute $l_n(t)$ for $n \ge 3$ and $-1 \le t < 1$. The following question is related to this issue.

Question 1. *Let* $\alpha \geq 0$ *and* $-1 < t < 1$. *Let* k *be such that*

$$\min\{\, \overline{P}_j^{(\alpha,\alpha)}(t) : j = 0, 1, \ldots \,\} = \overline{P}_k^{(\alpha,\alpha)}(t).$$

Is it true that the sequence

$$\overline{P}_0^{(\alpha,\alpha)}(t), \overline{P}_1^{(\alpha,\alpha)}(t), \ldots, \overline{P}_k^{(\alpha,\alpha)}(t)$$

is decreasing?

If the answer to this question is "yes", then computing $l_n(t)$ for $n \geq 3$ and any $-1 < t < 1$ becomes a simple matter of computing $\overline{P}_k^{(\alpha,\alpha)}(t)$ for $k = 0, 1, \ldots,$ until the sequence ceases to be decreasing, at which point we know we have found the minimum. Notice that in the proof of Theorem 3.8 we have actually shown that Question 1 admits an affirmative answer when $t = t_{k-1,k-1}^{(\alpha+1,\alpha+1)}$ for any given $k \geq 2$.

Our second question is also related to the computation of $l_n(t)$, indeed it concerns a complete characterization of l_n. We showed in Theorem 3.8 that, if $\alpha \geq 0$ and $t = t_{k-1,k-1}^{(\alpha+1,\alpha+1)}$ for any given $k \geq 2$, then

$$\min\{\, \overline{P}_j^{(\alpha,\alpha)}(t) : j = 0, 1, \ldots \,\} = \overline{P}_k^{(\alpha,\alpha)}(t).$$

So we know in this case exactly at which degree the minimum is achieved.

From Figure 3.2 it seems, in fact, that the minimum above is achieved by $\overline{P}_k^{(\alpha,\alpha)}$ for every $t \in [a, b]$, where a is the last point in $(0, 1)$ where $\overline{P}_{k-1}^{(\alpha,\alpha)}$ and $\overline{P}_k^{(\alpha,\alpha)}$ coincide, and likewise b is the last point in $(0, 1)$ where $\overline{P}_k^{(\alpha,\alpha)}$ and $\overline{P}_{k+1}^{(\alpha,\alpha)}$ coincide. If we recall identity (3.27), we are then lead to the following question:

Question 2. *Let* $\alpha \geq 0$ *and set* $t_{0,0}^{(\alpha+1,\alpha)} = -1$. *Is it true that, for* $k \geq 0$ *and all* $t \in [t_{k,k}^{(\alpha+1,\alpha)}, t_{k+1,k+1}^{(\alpha+1,\alpha)}]$, *we have*

$$\min\{\, \overline{P}_j^{(\alpha,\alpha)}(t) : j = 0, 1, \ldots \,\} = \overline{P}_{k+1}^{(\alpha,\alpha)}(t)?$$

An affirmative answer would lead to a complete characterization of the behavior of l_n for $n \geq 3$.

The last question we present concerns the bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ given in Corollary 3.7. The bound from the corollary comes from the inequality

$$\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \lim_{t \to 1} \frac{\omega_n}{\vartheta(G(S^{n-1}, t))}, \tag{3.39}$$

which follows from the fact that $\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \chi_{\mathrm{m}}(G(S^{n-1}, t))$ for all $-1 \leq t < 1$. It is then reasonable to ask whether the limit above provides the best lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$. In other words, we may ask the following question:

Question 3. *For* $n \geq 3$, *is it so that* $\lim_{t\to 1} \vartheta(G(S^{n-1}, t)) \leq \vartheta(G(S^{n-1}, u))$ *for every* $-1 \leq u < 1$?

An affirmative answer would show that taking the limit in (3.39) gives a lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ at least as good as $\omega_n/\vartheta(G(S^{n-1},t))$ for any given $-1 \leq t < 1$.

Question 3 is strongly related to a conjecture concerning Jacobi polynomials which is open at least since 1975. We now briefly discuss this connection.

From Theorem 3.5 we know that the answer to Question 3 is "yes" if and only if

$$\lim_{t \to 1} l_n(t) \geq l_n(u)$$

for all $-1 \leq u < 1$. Then, by combining (3.32) with Theorems 3.6 and 3.10, we see that, if

$$|\overline{P}_k^{(\alpha,\alpha)}(t_{k-1,k-1}^{(\alpha+1,\alpha)})| > |\overline{P}_{k+1}^{(\alpha,\alpha)}(t_{k,k}^{(\alpha+1,\alpha)})| \qquad \text{for } k = 2, 3, \ldots \tag{3.40}$$

with $\alpha = (n-3)/2$, then Question 3 admits an affirmative answer.

By using a result of Gerard and Roberts [18] we may reformulate (3.40). They show that, for α, $\beta > -1$,

$$P_k^{(\alpha+1,\beta)}(u) = P_k^{(\alpha,\beta)}(u) \qquad \text{if } u \text{ is an extremum of } P_k^{(\alpha,\beta)}.$$

Using this together with (3.18) and (3.26) we may then show that, for $\alpha > 0$,

$$\overline{P}_k^{(\alpha,\alpha)}(u) = \overline{P}_k^{(\alpha,\alpha-1)}(u) \qquad \text{if } u \text{ is an extremum of } \overline{P}_k^{(\alpha,\alpha-1)}. \tag{3.41}$$

This can also be seen to hold for $\alpha = 0$ if one defines

$$\overline{P}_k^{(0,-1)}(u) = \frac{\overline{P}_k^{(0,0)}(u) + \overline{P}_{k-1}^{(0,0)}(u)}{2}.$$

Now, using (3.26) together with (3.41), we may rewrite (3.40) as

$$|\overline{P}_k^{(\alpha,\alpha-1)}(t_{k-1,k-1}^{(\alpha+1,\alpha)})| > |\overline{P}_{k+1}^{(\alpha,\alpha-1)}(t_{k,k}^{(\alpha+1,\alpha)})| \qquad \text{for } k = 2, 3, \ldots. \tag{3.42}$$

Notice the resemblance between this and (3.30).

Both (3.30) and (3.42) are special cases of a conjecture presented on page 190 of Szegö [50]. The conjecture states that, for $\alpha > \beta > -1/2$,

$$|\overline{P}_k^{(\alpha,\beta)}(t_{k-1,k-1}^{(\alpha+1,\beta+1)})| > |\overline{P}_{k+1}^{(\alpha,\beta)}(t_{k,k}^{(\alpha+1,\beta+1)})| \qquad \text{for } k = 2, 3, \ldots. \tag{3.43}$$

Notice that a proof of this conjecture would provide an affirmative answer to Question 3 for $n \geq 4$. Wong and Zhang [57] show that the conjecture is asymptotically true, i.e., that (3.43) holds for all large enough k.

For $\alpha = 0$ it was proven by Wong and Zhang [58] that (3.42) holds with the inequalities reversed (this was a conjecture of Askey, cf. Wong and Zhang [58]). It is no surprise then that the answer to Question 3 is "no" when $n = 3$. Indeed, we have

$$l_3(t_{1,1}^{(1,0)}) = -1/3,$$

whereas from Theorem 3.6 we know that

$$\lim_{t \to 1} l_3(t) = -0.4027\ldots.$$

The difference between the limit and $l_3(t_{1,1}^{(1,0)})$ is however not enough to provide a better lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^3)$, since $\omega_3/\vartheta(G(S^2, t_{1,1}^{(1,0)})) = 4$, but also the ceiling of the lower bound of Corollary 3.7 is 4.

3.6 A theorem concerning many forbidden inner products

Let $t_1, t_2, \ldots$ be a sequence of numbers in $[0, 1)$ that converges to 1. Then

$$\alpha_{\mathrm{m}}(G(S^{n-1}, \{t_1, t_2, \ldots\})) = 0.$$

This is a simple consequence of a generalization — due to Weil (cf. Weil [56], page 50; see also Stromberg [49] for a simple proof) — of Steinhaus' theorem, which implies that if $C \subseteq S^{n-1}$ has positive measure, then there is a $-1 < t_0 < 1$ such that for all t with $t_0 < t < 1$, set C contains a pair of points with inner product t. In this section we show the following stronger result for $n \geq 3$.

Theorem 3.11. *Let* $t_1, t_2, \ldots$ *be a sequence of numbers in* $[0, 1)$ *that converges to* 1. *Then, for* $n \geq 3$, *we have that* $\alpha_{\mathrm{m}}(G(S^{n-1}, \{t_1, \ldots, t_N\})) \to 0$ *as* $N \to \infty$.

This theorem does not hold for $n = 2$. Indeed, for $p \geq 1$ integer, let $C_p \subseteq S^1$ be the union of the open arcs of the form $(2j\pi/p, (2j+1)\pi/p)$ for $j = 0, \ldots, p-1$ (the same construction was considered in Section 3.5a; see in particular Figure 3.1). Notice that no two points in C_p form an angle of $(2j+1)\pi/p$ with the origin for any integer $j \geq 0$. In particular, the angle between any two points in C_{3^k} is not of the form $\pi/3^j$ for $j = 1, \ldots, k$, so C_{3^k} is a stable set of

$$G(S^1, \{\cos \pi/3^j : j = 1, \ldots, k\}).$$

But then, since $\omega(C_{3^k}) = \pi$ for all $k \geq 1$, we immediately see that Theorem 3.11 fails for the sequence

$$\cos \pi/3, \cos \pi/3^2, \cos \pi/3^3, \ldots,$$

which converges to 1.

Theorem 3.11 is an analogue for the sphere of a theorem of Falconer [14] for the plane. In particular, the construction presented above to show that Theorem 3.11 fails for $n = 2$ is similar to a construction of Falconer [14] for the real line. We consider this theorem of Falconer and other related results in Section 4.5 and the treatment we present there is analogous to the one presented here.

We prove Theorem 3.11 by showing that it is a consequence of a stronger result, namely of the following theorem.

Theorem 3.12. *Let* $n \geq 3$. *There is a function* $r\colon (0, 1] \times [0, 1) \to (0, 1)$ *such that if* $t_1, \ldots, t_N \in [0, 1)$ *with* $t_1 < \cdots < t_N$ *satisfy*

$$t_{i+1} \geq r(\varepsilon, t_i) \quad \textit{for } i = 1, \ldots, N-1 \tag{3.44}$$

for a given $0 < \varepsilon \leq 1$, *then*

$$\alpha_{\rm m}(G(S^{n-1}, \{t_1, \ldots, t_N\})) \leq \omega_n \cdot \frac{\lambda_n^N + \varepsilon(N-1)}{1 + \lambda_n + \cdots + \lambda_n^N + \varepsilon(N-1)},$$

where $\lambda_n = \lambda_n(t_1) = |\min\{\, l_n(u) : t_1 \leq u < 1 \,\}|$.

This theorem is an analogue to the sphere of Theorem 4.5 of Section 4.5. It seems a bit artificial, so before providing a proof we first discuss some of its consequences.

First, from Section 3.5b we know that the number $\lambda_n(t)$ exists for all $-1 \leq t < 1$. Moreover, from Corollary 3.9 we know that $\lambda_n(t) \leq |l_n(0)|$ for all $0 \leq t < 1$. Then, from Theorem 3.8 and (3.21) we have that, for $0 \leq t < 1$,

$$\lambda_n(t) \leq |l_n(0)| = \frac{1}{n-1} \leq \frac{1}{2}. \tag{3.45}$$

The first consequence of Theorem 3.12 that we present is the following corollary.

Corollary 3.13. *For* $n \geq 3$, *let* $L_n = |\lim_{t\to 1} l_n(t)|$. *Then*

$$\inf_{t_1,\ldots,t_N \in [0,1)} \alpha_{\rm m}(G(S^{n-1}, \{t_1, \ldots, t_N\})) \leq \omega_n \cdot \frac{L_n^N(1 - L_n)}{1 - L_n^{N+1}}.$$

Proof. We know from Section 3.5c that L_n exists. Moreover $\lim_{t\to 1} \lambda_n(t) = L_n$. So, by taking $\varepsilon \to 0$ and $t_1 \to 1$ in Theorem 3.12, and since $1 + a + \cdots + a^N = (1 - a^{N+1})/(1 - a)$, we are done. ■

This corollary together with (3.45) shows that the measure of stable sets decrease exponentially with the number of inner products which define edges. Another useful corollary of Theorem 3.12 is the following weakening of it.

Corollary 3.14. *Let* $n \geq 3$ *and* $N \geq 2$. *If numbers* $t_1, \ldots, t_N \in [0, 1)$ *are given such that* $t_1 < \cdots < t_N$ *and*

$$t_{i+1} \geq r(\varepsilon, t_i) \qquad \textit{for } i = 1, \ldots, N-1,$$

where $\varepsilon = \lambda_n^{N+1}/((1 - \lambda_n)(N-1))$ *with* r *and* $\lambda_n = \lambda_n(t_1)$ *as in the statement of Theorem 3.12, then* $\alpha_{\rm m}(G(S^{n-1}, \{t_1, \ldots, t_N\})) \leq \omega_n 2^{-N}$.

Proof. We apply Theorem 3.12 and (3.45) to obtain

$$\begin{aligned} \alpha_{\rm m}(G(S^{n-1}, \{t_1, \ldots, t_N\})) &\leq \omega_n \cdot \frac{\lambda_n^N + \varepsilon(N-1)}{1 + \lambda_n + \cdots + \lambda_n^N + \varepsilon(N-1)} \\ &= \omega_n \cdot \frac{\lambda_n^N + \lambda_n^{N+1}/(1-\lambda_n)}{(1 - \lambda_n^{N+1})/(1 - \lambda_n) + \lambda_n^{N+1}/(1 - \lambda_n)} \\ &= \omega_n(\lambda_n^N(1 - \lambda_n) + \lambda_n^{N+1}) \\ &\leq \omega_n 2^{-N}, \end{aligned}$$

where we use that $\lambda_n(1 - \lambda_n) \leq 1/4$. ■

From this weakened version of Theorem 3.12 we may prove Theorem 3.11.

Proof of Theorem 3.11. If $t_1, t_2, \ldots$ is a sequence of numbers in $[0, 1)$ that converges to 1, then for every $N \geq 2$ we may find N numbers among $t_1, t_2, \ldots$ which satisfy the hypothesis of Corollary 3.14, so $\alpha_{\mathrm{m}}(G(S^{n-1}, \{t_1, t_2, \ldots\})) \leq \omega_n 2^{-N}$. Since this is true for every $N \geq 2$, we are done. ■

We end this section with a proof of Theorem 3.12. To this end it will be useful to introduce a dual of problem (3.10). When $D = \{t_1, \ldots, t_N\}$, problem (3.10) becomes

$$\begin{array}{lll} \sup\ \omega_n^2 f_0 & & \\ & \sum_{k=0}^{\infty} f_k = \omega_n^{-1}, & \\ & \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(t_i) = 0 & \text{for } i = 1, \ldots, N, \\ & f_k \geq 0 & \text{for } k = 0, 1, \ldots. \end{array} \tag{3.46}$$

We consider the following dual for it:

$$\begin{array}{ll} \inf\ \omega_n^{-1} z_0 & \\ \quad z_0 + z_1 + \cdots + z_N \geq \omega_n^2, & \\ \quad z_0 + z_1 \overline{P}_k^{(\alpha,\alpha)}(t_1) + \cdots + z_N \overline{P}_k^{(\alpha,\alpha)}(t_N) \geq 0 & \text{for } k = 1, 2, \ldots. \end{array} \tag{3.47}$$

As in the case of finite-dimensional linear programming, weak duality holds between (3.46) and (3.47). That is to say that, if $f_0, f_1, \ldots$ is a feasible solution of (3.46) and $(z_0, z_1, \ldots, z_N)$ is a feasible solution of (3.47), then

$$\omega_n^2 f_0 \leq \sum_{k=0}^{\infty} f_k (z_0 + z_1 \overline{P}_k^{(\alpha,\alpha)}(t_1) + \cdots + z_N \overline{P}_k^{(\alpha,\alpha)}(t_N)) = \omega_n^{-1} z_0.$$

So any feasible solution of (3.47) upper bounds $\vartheta(G(S^{n-1}, \{t_1, \ldots, t_N\}))$.

Proof of Theorem 3.12. Write $\alpha = (n-3)/2$ and let $0 < \varepsilon \leq 1$ and $0 < t < 1$ be given. We claim that we may pick a number $0 < u_0 < 1$ with the following property: For all $u \geq u_0$ and all $k \geq 0$, if $\overline{P}_k^{(\alpha,\alpha)}(u) \leq 1 - \varepsilon$, then $|\overline{P}_k^{(\alpha,\alpha)}(s)| < \varepsilon$ for all $0 \leq s \leq t$.

To prove the claim, recall from Section 3.5b that $\overline{P}_k^{(\alpha,\alpha)}(s) \to 0$ as $k \to \infty$ uniformly in the interval $[0, t]$. So we can take a k_0 such that $|\overline{P}_k^{(\alpha,\alpha)}(s)| < \varepsilon$ for all $0 \leq s \leq t$ and $k \geq k_0$. Now, since each $\overline{P}_k^{(\alpha,\alpha)}$ is continuous and such that $\overline{P}_k^{(\alpha,\alpha)}(1) = 1$, we may pick a number $0 < u_0 < 1$ such that $\overline{P}_k^{(\alpha,\alpha)}(u) > 1 - \varepsilon$ for all $u \geq u_0$ and $k \leq k_0$, and u_0 will have the property we claimed.

Set $r(\varepsilon, t) = u_0$. Suppose numbers $t_1, \ldots, t_N \in [0, 1)$ with $t_1 < \cdots < t_N$ are given which satisfy (3.44). We claim that, for $1 \leq j \leq N$,

$$\sum_{i=1}^{j} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i) \geq -\lambda_n^j - \varepsilon(j-1) \qquad \text{for all } k \geq 0. \tag{3.48}$$

Before proving the claim, let us show how to apply it in order to prove the theorem. Taking $j = N$ we see that

$$\sum_{i=1}^{N} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i) \geq -\lambda_n^N - \varepsilon(N-1) \qquad \text{for all } k \geq 0.$$

So, letting $S = 1 + \lambda_n + \cdots + \lambda_n^N + \varepsilon(N-1)$, we may set

$$z_0 = \omega_n^2 \cdot \frac{\lambda_n^N + \varepsilon(N-1)}{S} \qquad \text{and} \qquad z_i = \omega_n^2 \cdot \frac{\lambda_n^{i-1}}{S} \quad \text{for } i = 1, \ldots, N$$

and check that this is a feasible solution of (3.47). But then, from the weak duality relation between (3.47) and (3.46) together with Theorem 3.3, the theorem follows.

We now prove (3.48) by induction. For $j = 1$, the statement is obviously true. Now suppose the statement is true for some $1 \leq j < N$; we show it is also true for $j + 1$. To this end, let $k \geq 0$ be an integer. If $\overline{P}_k^{(\alpha,\alpha)}(t_{j+1}) > 1 - \varepsilon$, then by using the induction hypothesis and since $\lambda_n \leq 1$ we get

$$\begin{aligned}\sum_{i=1}^{j+1} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i) &= \lambda_n^j \overline{P}_k^{(\alpha,\alpha)}(t_{j+1}) + \sum_{i=1}^{j} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i)\\ &\geq \lambda_n^j(1-\varepsilon) - \lambda_n^j - \varepsilon(j-1)\\ &\geq -\varepsilon j.\end{aligned}$$

If, on the other hand, $\overline{P}_k^{(\alpha,\alpha)}(t_{j+1}) \leq 1 - \varepsilon$, we know from the choice of the t_i that $|\overline{P}_k^{(\alpha,\alpha)}(t_i)| < \varepsilon$ for $i = 1, \ldots, j$. But then we have

$$\begin{aligned}\sum_{i=1}^{j+1} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i) &= \lambda_n^j \overline{P}_k^{(\alpha,\alpha)}(t_{j+1}) + \sum_{i=1}^{j} \lambda_n^{i-1} \overline{P}_k^{(\alpha,\alpha)}(t_i)\\ &\geq -\lambda_n^{j+1} - \varepsilon j,\end{aligned}$$

proving (3.48). ■

3.7 The theta number of the complementary graph

In Chapter 2, when we studied the theta number for finite graphs, we derived relations between the theta number of a graph G and that of its complement $\overline{G}$. These relations were given in Theorem 2.2. We try now to decide what the theta number of the complementary graph of the infinite graph $G(S^{n-1}, D)$ should be in order to see which of the properties of the theta number for finite graphs extend to this infinite case.

Let $D \subseteq [0, 1)$ be a closed subset of the real line. Consider the graph $\overline{G}(S^{n-1}, D)$, the complement of $G(S^{n-1}, D)$. Its vertices are the points of the $(n-1)$-dimensional unit sphere S^{n-1}. Two distinct vertices x, $y \in S^{n-1}$ are adjacent in $\overline{G}(S^{n-1}, D)$ if $x \cdot y \notin D$.

How could we define $\vartheta(\overline{G}(S^{n-1}, D))$? The definition which worked for the graph $G(S^{n-1}, D)$ does not work here anymore: The optimization problem (3.5), adapted to the graph $\overline{G}(S^{n-1}, D)$, is infeasible. Indeed, the only continuous kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ satisfying $A(x, y) = 0$ for all $x \neq y$ such that $x \cdot y \notin D$ is the zero kernel, and this kernel is not feasible for (3.5).

This happens because, while the complement $\overline{G}(S^{n-1}, D)$ of $G(S^{n-1}, D)$ is also a distance graph, in $G(S^{n-1}, D)$ there is a whole interval of inner products close to 1 that do not define edges, whereas in $\overline{G}(S^{n-1}, D)$ there are inner products arbitrarily close to 1 which define edges. It is this difference that accounts for the fact that, while formulation (3.5) works for $G(S^{n-1}, D)$, it fails for $\overline{G}(S^{n-1}, D)$.

Moreover, this difference between $G(S^{n-1}, D)$ and $\overline{G}(S^{n-1}, D)$ also accounts for the different nature of the stable sets of both graphs. While in $G(S^{n-1}, D)$ stable sets can have positive measure, in $\overline{G}(S^{n-1}, D)$ they are always finite and of bounded cardinality. So $\overline{G}(S^{n-1}, D)$ has finite stability number.

For finite graphs we presented two equivalent formulations for the theta number, problems (2.1) and (2.3), which were obtained from the duality theory of semidefinite programming. Our generalization of the theta number for $G(S^{n-1}, D)$, problem (3.5), was based on (2.1). If we base ourselves on (2.3), however, we may also define a version of the theta number for $\overline{G}(S^{n-1}, D)$.

The *theta number* of $\overline{G}(S^{n-1}, D)$, denoted by $\vartheta(\overline{G}(S^{n-1}, D))$, is the optimal value of the following optimization problem:

$$\begin{aligned} \inf\ & \lambda \\ & Z(x, x) = \lambda - 1 \quad \text{for all } x \in S^{n-1}, \\ & Z(x, y) = -1 \qquad \text{if } x \cdot y \in D, \\ & Z\colon S^{n-1} \times S^{n-1} \to \mathbb{R} \text{ is continuous and positive.} \end{aligned} \tag{3.49}$$

It is *a priori* not clear that this optimization problem is feasible — we will prove feasibility in a moment. For now, we observe that the following inequality holds, as expected:

$$\vartheta(\overline{G}(S^{n-1}, D)) \geq \alpha(\overline{G}(S^{n-1}, D)). \tag{3.50}$$

Assume (3.49) is feasible and let C be a nonempty stable set of $\overline{G}(S^{n-1}, D)$ and Z be a feasible solution of (3.49). Recall C is finite. Let λ be such that $Z(x, x) = \lambda - 1$ for all $x \in S^{n-1}$. Then, since Z is continuous and positive and since C is stable, by using (3.3) we get

$$0 \leq \sum_{x,y \in C} Z(x, y) = (\lambda - 1)|C| - (|C|^2 - |C|),$$

and this implies that $\lambda \geq |C|$, whence (3.50) follows. Note also how this is a dual approach: While any feasible solution of (3.49) provides an upper bound for $\alpha(\overline{G}(S^{n-1}, D))$, our proof that (3.5) gives an upper bound for the parameter $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$ shows only that an optimal solution of that optimization problem gives an upper bound.

We now prove that

$$\vartheta(G(S^{n-1}, D))\,\vartheta(\overline{G}(S^{n-1}, D)) = \omega_n, \tag{3.51}$$

along the way proving that (3.49) is feasible. Identity (3.51) is in direct analogy with the similar identity that is valid for vertex-transitive finite graphs (see Theorem 2.2).

We begin by showing inequality "$\leq$" in (3.51), proving at the same time that problem (3.49) is feasible. So let $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ be a feasible solution of (3.5) and assume that A is invariant under $\mathrm{O}(\mathbb{R}^n)$ (invariance under the orthogonal group was defined in Section 3.4). Then, in particular, all diagonal entries of A are equal to ω_n^{-1}. Let $J\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ be the constant one kernel and consider the kernel

$$Z = \frac{\omega_n^2}{\langle J, A\rangle} A - J.$$

We claim that Z is feasible for (3.49).

To prove this claim it suffices to check that Z is positive, since it clearly satisfies all other constraints of (3.49). To prove that Z is positive, notice that $\mathbf{1}$, the constant one function, is an eigenfunction of A, with associated eigenvalue $\langle J, A\rangle/\omega_n$. Indeed, since A is invariant under the orthogonal group, for every $x \in S^{n-1}$ we have

$$\int_{S^{n-1}} A(x, y)\, d\omega(y) = \frac{\langle J, A\rangle}{\omega_n}.$$

Now $\mathbf{1}$ is the only eigenfunction of J, with associated eigenvalue ω_n. Then it follows from the Hilbert-Schmidt theorem (Theorem 3.1) that Z is positive, as we wanted.

So Z is a feasible solution of (3.49), and hence $\vartheta(\overline{G}(S^{n-1}, D)) \leq 1 + Z(x, x)$ for any $x \in S^{n-1}$. So we have

$$\langle J, A\rangle\, \vartheta(\overline{G}(S^{n-1}, D)) \leq \langle J, A\rangle(1 + Z(x, x)) = \langle J, A\rangle \frac{\omega_n}{\langle J, A\rangle} = \omega_n.$$

Now, since this holds for any invariant feasible solution A of (3.5) and since we may restrict ourselves in (3.5), without loss of generality, to invariant kernels (cf. Section 3.4), we have

$$\vartheta(G(S^{n-1}, D))\,\vartheta(\overline{G}(S^{n-1}, D)) \leq \omega_n.$$

We now show inequality "$\geq$" in (3.51). To this end, let Z be a feasible solution of (3.49) and let λ be such that $Z(x, x) = \lambda - 1$ for all $x \in S^{n-1}$. Then

$$A = (\lambda\omega_n)^{-1}(Z + J)$$

is feasible for (3.5), as can be easily checked. So we know that $\vartheta(G(S^{n-1}, D)) \geq \langle J, A\rangle$ and hence

$$\vartheta(G(S^{n-1}, D))\lambda \geq \langle J, A\rangle\lambda = (\lambda\omega_n)^{-1}(\langle J, Z\rangle + \langle J, J\rangle)\lambda \geq \omega_n,$$

since $\langle J, Z\rangle \geq 0$ as both J and Z are positive. Now, since Z is any feasible solution of (3.49), we must have

$$\vartheta(G(S^{n-1}, D))\,\vartheta(\overline{G}(S^{n-1}, D)) \geq \omega_n,$$

and so our proof of (3.51) is finished.

To solve (3.49) we may apply the same idea we applied before in relation to (3.5). First, we restrict ourselves to kernels that are invariant under $\mathrm{O}(\mathbb{R}^n)$, and in doing so no generality is lost (cf. Section 3.4). Then we apply Theorem 3.4 to decompose the kernel Z. We arrive at the following optimization problem, equivalent to (3.49):

$$\begin{aligned} \inf\ & 1 + \textstyle\sum_{k=0}^{\infty} g_k \\ & \textstyle\sum_{k=0}^{\infty} g_k P_k^{(\alpha,\alpha)}(t) = -1 && \text{for } t \in D, \\ & g_k \geq 0 && \text{for } k = 0, 1, \ldots, \end{aligned} \tag{3.52}$$

where $\alpha = (n-3)/2$.

3.7a. Bounds from finite subgraphs. We now use formulation (3.49) to derive a result concerning finite subgraphs of $G(S^{n-1}, D)$, which is analogous to the monotonicity of the theta number for finite graphs (cf. item (i) in Theorem 2.2).

Let $H = (V, E)$ be a finite subgraph of $G(S^{n-1}, D)$. Then

$$\vartheta(\overline{H}) \leq \chi(H) \leq \chi_{\mathrm{m}}(G(S^{n-1}, D)).$$

This means that a strategy to find lower bounds for $\chi_{\mathrm{m}}(G(S^{n-1}, D))$ is to compute the theta number of the complements of finite subgraphs of $G(S^{n-1}, D)$. This, however, cannot yield better bounds than $\omega_n/\vartheta(G(S^{n-1}, D))$, as we proceed to show.

Indeed, let Z be a feasible solution of (3.49). Then the matrix $\big(Z(x,y)\big)_{x,y \in V}$ is a feasible solution for formulation (2.3) of $\vartheta(\overline{H})$. With (3.51), this implies that

$$\vartheta(\overline{H}) \leq \vartheta(\overline{G}(S^{n-1}, D)) = \frac{\omega_n}{\vartheta(G(S^{n-1}, D))},$$

as we wanted.

Upper bounds for the stability number of a finite subgraph $H = (V, E)$ of the graph $G(S^{n-1}, D)$ also provide upper bounds for the measurable stability number of $G(S^{n-1}, D)$. Indeed we have

$$\frac{\alpha_{\mathrm{m}}(G(S^{n-1}, D))}{\omega_n} \leq \frac{\alpha(H)}{|V|}. \tag{3.53}$$

To see this, let C be a measurable stable set of $G(S^{n-1}, D)$. Then $|C \cap T \cdot V| \leq \alpha(H)$ for all $T \in \mathrm{O}(\mathbb{R}^n)$, where

$$T \cdot V = \{\, Tx : x \in V \,\}.$$

Let μ be the Haar measure over $\mathrm{O}(\mathbb{R}^n)$ normalized so that $\mu(\mathrm{O}(\mathbb{R}^n)) = 1$. Then for any $x \in S^{n-1}$ and measurable $X \subseteq S^{n-1}$,

$$\mu(\{\, T \in \mathrm{O}(\mathbb{R}^n) : Tx \in X \,\}) = \frac{\omega(X)}{\omega_n}$$

(cf. Theorem 3.7 in Mattila [32]). With this we have that

$$\begin{aligned}\frac{\omega(C)}{\omega_n}|V| &= \sum_{x\in V}\int_{\mathrm{O}(\mathbb{R}^n)} |C\cap\{Tx\}|\,d\mu(T)\\ &= \int_{\mathrm{O}(\mathbb{R}^n)} |C\cap T\cdot V|\,d\mu(T)\\ &\le \alpha(H)\end{aligned}$$

and, since this holds for every measurable stable set C of $G(S^{n-1}, D)$, (3.53) follows.

Now, since $\vartheta(H)\ge\alpha(H)$, we have

$$\frac{\vartheta(H)}{|V|}\ge\frac{\alpha(H)}{|V|}\ge\frac{\alpha_{\mathrm{m}}(G(S^{n-1},D))}{\omega_n},$$

and then by computing the theta number of finite subgraphs of $G(S^{n-1}, D)$ we may find upper bounds for $\alpha_{\mathrm{m}}(G(S^{n-1}, D))$. But these bounds cannot be better than the bound provided by $\vartheta(G(S^{n-1}, D))$. Indeed, since H is finite, from Theorem 2.2 we know that

$$\vartheta(H)\,\vartheta(\overline{H})\ge|V|.$$

But then, since $\vartheta(\overline{H})\le\vartheta(\overline{G}(S^{n-1}, D))$, as we observed above, and from (3.51), we have

$$\frac{\vartheta(H)}{|V|}\ge\frac{1}{\vartheta(\overline{H})}\ge\frac{\vartheta(G(S^{n-1},D))}{\omega_n},$$

as we wanted.

3.7b. An application to spherical codes. Finally, we would like to discuss an application of (3.49) to the problem of spherical codes. A *spherical code* in S^{n-1} of *minimum angular distance* θ, $0<\theta\le\pi$, is a subset C of S^{n-1} such that the angle between any two distinct points in C is at least θ. We define $A(n,\theta)$ to be the maximum size that a spherical code in S^{n-1} of minimum angular distance θ can have.

Delsarte, Goethals, and Seidel [11] and Kabatyanskii and Levenshtein [23] proposed an upper bound for $A(n,\theta)$ based on linear programming, which is analogous to Delsarte's linear programming bound for binary codes (cf. Section 2.3). We show now how to obtain this bound from a suitable generalization of the theta number, in analogy with the similar observation that McEliece, Rodemich, and Rumsey [33] and Schrijver [45] made in relation to Delsarte's linear programming bound for binary codes (cf. Section 2.3).

Fix a number $-1\le t<1$ and consider the graph $G(S^{n-1}, [-1,t])$. The stable sets in the complement $\overline{G}(S^{n-1}, [-1,t])$ of $G(S^{n-1}, [-1,t])$ are precisely the spherical codes in S^{n-1} of minimum angular distance $\theta=\arccos t$. So the stability number of $\overline{G}(S^{n-1}, [-1,t])$ is equal to the parameter $A(n,\theta)$.

We may consider problem (3.49) for $\overline{G}(S^{n-1}, [-1, t])$ and at the same time strengthen it in the direction of $\alpha(\overline{G}(S^{n-1}, [-1, t])$. So we obtain the following optimization problem, the optimal value of which we denote by $\vartheta'(\overline{G}(S^{n-1}, [-1, t]))$:

$$\begin{aligned} \inf\ & \lambda \\ & Z(x,x) = \lambda - 1 \quad \text{for all } x \in S^{n-1}, \\ & Z(x,y) \leq -1 \quad\quad \text{if } x \cdot y \in [-1, t], \\ & Z\colon S^{n-1} \times S^{n-1} \to \mathbb{R} \text{ is continuous and positive.} \end{aligned} \tag{3.54}$$

Let now $C \subseteq S^{n-1}$ be a nonempty stable set of $\overline{G}(S^{n-1}, [-1, t])$. Let Z be a feasible solution of (3.54) and λ be such that $Z(x, x) = \lambda - 1$ for all $x \in S^{n-1}$. Then, using (3.3), we have

$$0 \leq \sum_{x,y \in C} Z(x, y) \leq (\lambda - 1)|C| - (|C|^2 - |C|).$$

This shows that $\lambda \geq |C|$ and so

$$\vartheta'(\overline{G}(S^{n-1}, [-1, t])) \geq \alpha(\overline{G}(S^{n-1}, [-1, t])) = A(n, \theta).$$

Note that in (3.54) we require that $A(x, y) \leq -1$ when $x \cdot y \in [-1, t]$ instead of requiring equality to hold. As we see above this still provides an upper bound for $A(n, \theta)$, and the upper bound is possibly tighter than if we had required equality to hold. We also remark that $\vartheta'(\overline{G}(S^{n-1}, [-1, t]))$ is analogous to the parameter ϑ' defined for finite graphs in Section 2.3.

In order to solve (3.54), we apply Theorem 3.4, so that $\vartheta'(\overline{G}(S^{n-1}, [-1, t]))$ is the optimal value of the following infinite linear program:

$$\begin{aligned} \inf\ & 1 + \textstyle\sum_{k=0}^{\infty} g_k \\ & \textstyle\sum_{k=0}^{\infty} g_k P_k^{(\alpha,\alpha)}(u) \leq -1 \quad \text{for all } u \in [-1, t], \\ & g_k \geq 0 \quad\quad\quad\quad\quad\quad\quad\ \ \text{for } k = 0, 1, \ldots, \end{aligned}$$

where $\alpha = (n - 3)/2$.

This is exactly the linear programming bound of Delsarte, Goethals, and Seidel [11] and Kabatyanskii and Levenshtein [23]. Solving this optimization problem is not as straight-forward as it was to solve (3.5). Since any feasible solution of it provides an upper bound for $A(n, \theta)$, however, we may for instance simply consider a finite number of variables and use some optimization approach to find good feasible solutions.

3.8 A proof of Schoenberg's theorem

Our goal in this section is to give a proof of Theorem 3.4, which was originally stated and proved by Schoenberg [44]. We aim at presenting a short and simple proof. The reader who is familiar with representation theory will likely notice many concepts from representation theory hidden in the proof. We choose not to make our use of such concepts explicit as they are not essential to the proof we present. Finally, it is important to emphasize the similarity between this proof and the proof of Theorem 2.3.

We start with a lemma from which Theorem 3.4 will follow after a bit of analytical work.

Lemma 3.15. *The kernel* $E_k^n\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ *such that* $E_k^n(x, y) = \overline{P}_k^{(\alpha,\alpha)}(x \cdot y)$, *with* $\alpha = (n-3)/2$ *for* $n \geq 2$, *is positive for every* k.

We will prove the lemma in a moment — proving it is actually most of the work. First, however, we use it to prove the theorem.

Proof of Theorem 3.4. We first show that, if f_0, f_1, ... are nonnegative numbers such that $\sum_{k=0}^{\infty} f_k$ converges, then the series

$$\sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(x \cdot y) \tag{3.55}$$

converges absolutely and uniformly for all x, $y \in S^{n-1}$.

Indeed, since we know from Lemma 3.15 that the kernel E_k^n is positive, we must have that $|\overline{P}_k^{(\alpha,\alpha)}(u)| \leq \overline{P}_k^{(\alpha,\alpha)}(1) = 1$ for all $u \in [-1, 1]$. This immediately implies that the series

$$\sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(u)$$

converges absolutely for every $u \in [-1, 1]$, and hence (3.55) converges absolutely for all x, $y \in S^{n-1}$.

Now, for every $u \in [-1, 1]$ and every integer $m \geq 0$ we also have that

$$\Bigl|\sum_{k=m}^{\infty} f_k P_k^{(\alpha,\alpha)}(u)\Bigr| \leq \sum_{k=m}^{\infty} f_k,$$

and so (3.55) also converges uniformly for all x, $y \in S^{n-1}$, as we wanted.

With the above observation, if we are given nonnegative numbers f_0, f_1, ... such that $\sum_{k=0}^{\infty} f_k$ converges, then the kernel $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that

$$A(x, y) = \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}(x \cdot y)$$

is continuous. From Lemma 3.15 it is also positive, and sufficiency follows.

Now we prove necessity. Let $A\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ be a continuous, positive, and invariant kernel. Kernel A is invariant, so let $h\colon [-1, 1] \to \mathbb{R}$ be the function such that $A(x, y) = h(x \cdot y)$ for all x, $y \in S^{n-1}$.

We mention in Section 3.4 that the polynomials $\overline{P}_0^{(\alpha,\alpha)}$, $\overline{P}_1^{(\alpha,\alpha)}$, ... form a complete orthogonal system of $L^2([-1, 1])$, which we equip with the inner product $(\cdot,\cdot)_{\alpha,\alpha}$. So there are numbers $f_0, f_1, \ldots$ such that

$$h = \sum_{k=0}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)} \tag{3.56}$$

with convergence in the L^2-norm.

We first claim that the f_k are all nonnegative. To see this, recall from Section 3.4 that, since the $\overline{P}_k^{(\alpha,\alpha)}$ are pairwise orthogonal with respect to the inner product $(\cdot,\cdot)_{\alpha,\alpha}$, the kernels E_k^n are pairwise orthogonal with respect to the inner product $\langle\cdot,\cdot\rangle$. But then, since from (3.56) we have

$$A = \sum_{k=0}^{\infty} f_k E_k^n$$

in the sense of L^2 convergence and since, by Lemma 3.15, each E_k^n is positive, we must have that the f_k are all nonnegative. (This follows from the fact that, if A and B are positive kernels, then $\langle A, B\rangle \geq 0$, cf. Section 3.2.)

To finish, we show that the series $\sum_{k=0}^{\infty} f_k$ converges. To this end, consider for $m = 0, 1, \ldots$ the function

$$h_m(u) = h(u) - \sum_{k=0}^{m} f_k \overline{P}_k^{(\alpha,\alpha)}(u) \qquad \text{for all } u \in [-1,1].$$

These are continuous functions. Moreover, since we have

$$h_m = \sum_{k=m+1}^{\infty} f_k \overline{P}_k^{(\alpha,\alpha)}$$

in the sense of L^2 convergence, from Lemma 3.15 it follows that for each m the kernel $A_m(x,y) = h_m(x \cdot y)$ is positive.

This implies in particular that $h_m(1) \geq 0$ for all m. But then we have

$$h(1) - \sum_{k=0}^{m} f_k = h(1) - \sum_{k=0}^{m} f_k \overline{P}_k^{(\alpha,\alpha)}(1) = h_m(1) \geq 0,$$

and we conclude that the series of nonnegative terms $\sum_{k=0}^{\infty} f_k$ converges to a number less than or equal to $h(1)$, as we wanted. ∎

The part of the proof in which we deal with convergence is basically a simplified version of the proof of Mercer's theorem (cf. Riesz and Sz.-Nagy [42]). All that is left to do now is to prove the lemma.

Proof of Lemma 3.15. An n-variable polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is *harmonic* if it is homogeneous and vanishes under the Laplace operator, that is, if $\Delta p = 0$, where

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$$

is the *Laplace operator*.

Recall that for $L^2(S^{n-1})$ we use the inner product

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

We may also treat polynomials p, $q \in \mathbb{R}[x_1, \ldots, x_n]$ as functions in $L^2(S^{n-1})$ by considering their restrictions to the sphere. It is in this sense that we also write (p, q), i.e., when we write such an expression we consider p and q restricted to the sphere.

Claim A. If p, $q \in \mathbb{R}[x_1, \ldots, x_n]$ *are harmonic and of distinct degrees, then* $(p, q) = 0$.

Proof. This follows easily from Green's identity (equation (1.1) in Axler, Bourdon, and Ramey [2]; see also Section 9.4 of Andrews, Askey, and Roy [1]). Let p_k, p_l be two harmonic polynomials of degrees k and l, respectively. Green's identity states that

$$\begin{aligned}\int_{\|x\|\leq 1} & p_k(x)(\Delta p_l)(x) - p_l(x)(\Delta p_k)(x)\,dx \\ &= \int_{S^{n-1}} p_k(x)(\nabla p_l)(x) \cdot x - p_l(x)(\nabla p_k)(x) \cdot x\,d\omega(x).\end{aligned} \tag{3.57}$$

Note that, since p_k, p_l are harmonic, both sides above are equal to 0.

Now, since p_k is a homogeneous polynomial of degree k,

$$(\nabla p_k) \cdot x = x_1 \frac{\partial p_k}{\partial x_1} + \cdots + x_n \frac{\partial p_k}{\partial x_n} = k p_k,$$

and similarly $(\nabla p_l) \cdot x = l p_l$. Substituting back into the right-hand side of (3.57) we get

$$(l - k) \int_{S^{n-1}} p_k(x) p_l(x)\,d\omega(x) = 0,$$

as we wanted. ◀

The restriction of a harmonic polynomial in $\mathbb{R}[x_1, \ldots, x_n]$ to S^{n-1} is called a *spherical harmonic*. Because of Claim A, we may define the *degree* of a nonzero spherical harmonic to be the degree of a harmonic polynomial from which it is obtained. Actually, it is true that two harmonic polynomials which coincide in S^{n-1} coincide everywhere, so that each spherical harmonic corresponds to exactly one harmonic polynomial. We do not need this more general fact for the rest of the proof, however.

For $k = 0, 1, \ldots$, let H_k be the subspace of $L^2(S^{n-1})$ generated by the spherical harmonics of degree k. We equip H_k with the inner product $(\cdot, \cdot)$. We claim that $H_k \neq \{\mathbf{0}\}$ for all k. To prove this, we need to exhibit for every k a harmonic polynomial of degree k that does not vanish on the sphere. Consider then the two-variable polynomial $(x_1 + i x_2)^k$. This is a homogeneous polynomial of degree k with complex coefficients. It also vanishes under the Laplace operator, as can be seen after a simple application of the chain rule.

This polynomial has complex coefficients, but from it we may obtain a polynomial with real coefficients that is harmonic. To do so, expand $(x_1 + ix_2)^k$ to obtain $k+1$ monomials. Now, simply delete the imaginary unit i from each monomial in which it appears. For instance, in the case of $k = 4$, we have the polynomial

$$(x_1 + ix_2)^4 = x_1^4 + 4ix_1^3x_2 - 6x_1^2x_2^2 - 4ix_1x_2^3 + x_2^4,$$

and from it we obtain the polynomial

$$x_1^4 + 4x_1^3x_2 - 6x_1^2x_2^2 - 4x_1x_2^3 + x_2^4.$$

The polynomial obtained after the deletion of the imaginary unit has real coefficients and one can easily check that it is harmonic and of degree k. Moreover, if we consider this polynomial as a polynomial in $\mathbb{R}[x_1, \ldots, x_n]$, we immediately see that when it is restricted to the sphere it is nonzero, proving the claim.

Let $e_{k,1}, \ldots, e_{k,h_k}$ be an orthonormal basis of H_k, where h_k is the dimension of H_k. Consider the kernel $B_k \colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that

$$B_k(x, y) = \sum_{i=1}^{h_k} e_{k,i}(x)e_{k,i}(y). \tag{3.58}$$

Each B_k is nonzero, continuous, and positive. From Claim A it also follows that the B_k, $k = 0, 1, \ldots$, are pairwise orthogonal with respect to $\langle \cdot, \cdot \rangle$.

Claim B. Each kernel B_k is invariant under $\mathrm{O}(\mathbb{R}^n)$.

Proof. Let $f \colon S^{n-1} \to \mathbb{R}$ and $T \in \mathrm{O}(\mathbb{R}^n)$. We define $T \cdot f$ as the function such that $(T \cdot f)(x) = f(Tx)$ for all $x \in S^{n-1}$. Recall from Section 3.4 that the inner product $(\cdot, \cdot)$ is invariant under the orthogonal group, that is, if $f, g \in L^2(S^{n-1})$, then $(T \cdot f, T \cdot g) = (f, g)$.

For each $x \in S^{n-1}$, the function $\mathrm{ev}_x \colon H_k \to \mathbb{R}$ such that $\mathrm{ev}_x(f) = f(x)$ is a linear functional. Since H_k is finite dimensional, this means that for each $x \in S^{n-1}$ there is an $r_x \in H_k$ such that $(f, r_x) = f(x)$ for all $f \in H_k$. Moreover, if $T \in \mathrm{O}(\mathbb{R}^n)$, then $r_{Tx} = T^{-1} \cdot r_x$. Indeed, we have that, for $f \in H_k$,

$$\begin{aligned} (f, T^{-1} \cdot r_x) &= (T \cdot f, r_x) \\ &= (T \cdot f)(x) \\ &= f(Tx), \end{aligned}$$

as we wanted.

Now note that, since $e_{k,1}, \ldots, e_{k,h_k}$ is an orthonormal basis of H_k,

$$B_k(x, y) = \sum_{i=1}^{h_k} (e_{k,i}, r_x)(e_{k,i}, r_y) = (r_x, r_y).$$

But we also have

$$(r_{Tx}, r_{Ty}) = (T^{-1} \cdot r_x, T^{-1} \cdot r_y) = (r_x, r_y),$$

and so $B_k(Tx, Ty) = B_k(x, y)$, proving that B_k is invariant. ◀

We finally come to the last claim we need to complete the proof.

Claim C. For each $k = 0, 1, \ldots$ *there is a real one-variable polynomial* R_k *of degree* k *such that* $B_k(x, y) = R_k(x \cdot y)$ *for all* $x, y \in S^{n-1}$.

Proof. Let $e = (1, 0, \ldots, 0) \in S^{n-1}$. From (3.58), we may see $p(x) = B_k(e, x)$ as an n-variable homogeneous polynomial of degree k. For $u \in [-1, 1]$, let $w(u) = (u, (1 - u^2)^{1/2}, 0, \ldots, 0) \in S^{n-1}$. Using Claim B, we conclude that $B_k(x, y) = p(w(x \cdot y))$ for all $x, y \in S^{n-1}$.

For $u \in [-1, 1]$, let us then compute $p(w(u))$. In computing $p(w(u))$, we need only look at monomials of p which contain only variables x_1 and x_2. There must be at least one such monomial, as B_k is nonzero.

Now let p_0 be the polynomial consisting of all monomials of p that are of the form $cx_1^{k_1} x_2^{k_2}$ with k_2 even; likewise, let p_1 be the polynomial consisting of all monomials $cx_1^{k_1} x_2^{k_2}$ of p that are such that k_2 is odd. So, for $u \in [-1, 1]$, we have

$$p(w(u)) = p_0(w(u)) + p_1(w(u)). \tag{3.59}$$

We claim that, for $u \in [-1, 1]$, we actually have

$$p(w(u)) = p_0(w(u)).$$

To see this, for $u \in [-1, 1]$ let $w'(u) = (u, -(1 - u^2)^{1/2}, 0, \ldots, 0) \in S^{n-1}$. Since B_k is invariant and $e \cdot w(u) = e \cdot w'(u)$, we must have $p(w(u)) = p(w'(u))$. But notice that $p_0(w(u)) = p_0(w'(u))$, so that from (3.59) we see that $p_1(w(u)) = p_1(w'(u))$ and, since $p_1(w'(u)) = -p_1(w(u))$, we must have that $p_1(w(u)) = 0$ for all $u \in [-1, 1]$, and our claim follows.

So it is at once clear that $R_k(u) = p_0(w(u))$ can be seen as a one-variable polynomial of degree k. Moreover, $B_k(x, y) = R_k(x \cdot y)$ for all $x, y \in S^{n-1}$. ◀

Since the kernels B_k, $k = 0, 1, \ldots$, are pairwise orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$, the polynomials R_k, $k = 0, 1, \ldots$, are pairwise orthogonal with respect to the inner product $(\cdot, \cdot)_{\alpha,\alpha}$, i.e., they are orthogonal with respect to the weight function $(1 - u^2)^{(n-3)/2}$ (cf. Section 3.4). But then, since $R_k(1) > 0$, R_k must be a positive multiple of $\overline{P}_k^{(\alpha,\alpha)}$. So also B_k is a positive multiple of E_k^n, and hence E_k^n is positive for $k = 0, 1, \ldots$. ■

Chapter Four

Distance graphs on the Euclidean space

In Chapter 3 we showed how to generalize the Lovász theta function to distance graphs over the $(n-1)$-dimensional unit sphere in order to upper bound the measurable stability number of such graphs. Our main motivation in doing so was to find lower bounds for the measurable chromatic number of $\mathbb{R}^n$, and we could improve on previous lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 10, \ldots, 24$.

In this chapter we use a more direct method to find lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$, namely, we work directly with $\mathbb{R}^n$ instead of S^{n-1}. As a consequence we are able to improve on previously known bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 3, \ldots, 24$. We also provide bounds for $m_1(\mathbb{R}^n)$, the maximum density of 1-avoiding sets, which improve on previously known bounds for $n = 2, \ldots, 24$. Our method can be seen as a generalization of the Lovász theta number for distance graphs on $\mathbb{R}^n$; this view is explored in detail in Section 4.6. This chapter is partially based on the paper by Oliveira and Vallentin [37].

4.1 Distance-avoiding sets in Euclidean space

A set $C \subseteq \mathbb{R}^n$ *avoids* distance d if for all $x, y \in C$ we have $\|x - y\| \neq d$. Sometimes we say that a set that avoids distance d is a *d-avoiding set*.

Given a Lebesgue measurable set $C \subseteq \mathbb{R}^n$, its *upper density* is defined as

$$\overline{\delta}(C) = \limsup_{R \to \infty} \frac{\operatorname{vol}(C \cap [-R, R]^n)}{\operatorname{vol}[-R, R]^n},$$

where $[-R, R]^n$ is the n-dimensional cube of side $2R$ with center at the origin and $\operatorname{vol} X$ simply denotes the Lebesgue measure of set X. In this chapter, we will investigate the maximum density that measurable distance-avoiding sets can achieve, that is, we investigate the parameter

$$m_{d_1,\ldots,d_N}(\mathbb{R}^n) = \sup\{\, \overline{\delta}(C) : C \subseteq \mathbb{R}^n \text{ is measurable and avoids distances } d_1, \ldots, d_N \,\}.$$

We are chiefly interested in finding upper bounds for $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$, in particular for the case of 1-avoiding sets, i.e., for $m_1(\mathbb{R}^n)$. We do so because of the connection between $m_1(\mathbb{R}^n)$ and $\chi_{\mathrm{m}}(\mathbb{R}^n)$, the measurable chromatic number of $\mathbb{R}^n$. Recall from Section 1.2 that $\chi_{\mathrm{m}}(\mathbb{R}^n)$ is the minimum number of measurable 1-avoiding sets needed to partition $\mathbb{R}^n$, hence

$$m_1(\mathbb{R}^n)\chi_{\mathrm{m}}(\mathbb{R}^n) \geq 1, \tag{4.1}$$

so that any upper bound for $m_1(\mathbb{R}^n)$ implies a lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$.

In the next section, we prove the main theorem of this chapter, which provides a tool to compute upper bounds for $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$. To state the theorem, we first introduce the real-valued function Ω_n, which is defined on the nonnegative reals and is such that

$$\Omega_n(\|u\|) = \frac{1}{\omega_n}\int_{S^{n-1}} e^{iu\cdot\xi}\, d\omega(\xi), \tag{4.2}$$

for $u \in \mathbb{R}^n$, where ω is the surface measure on the $(n-1)$-dimensional unit sphere

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

Here $x \cdot y$ denotes the standard Euclidean inner product between vectors x and y. Measure ω is normalized so that

$$\omega(S^{n-1}) = \omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

From (4.2) we immediately see that $\Omega_n(0) = 1$. Moreover, for $n = 1$ we simply have $\Omega_n(t) = \cos t$ for all $t \geq 0$. For general n, the function Ω_n also admits an expression in terms of Bessel functions, a fact which allows us to compute good numerical approximations of Ω_n and also to apply analytical tools. We will use this expression, which we present in Section 4.3, in our applications of the main theorem. Figure 4.1 shows a graph of function Ω_4 which illustrates the typical behavior of the functions Ω_n for $n \geq 2$. In particular, we see that the first extremum of Ω_4 is a global minimum and that $\Omega_4(t) \to 0$ as $t \to \infty$. We will discuss these and other properties of Ω_n in Section 4.3.

Having defined the function Ω_n, we may finally present our main theorem.

Theorem 4.1. *Let $d_1, \ldots, d_N$ be any given positive real numbers. Suppose there are numbers $z_0, z_1, \ldots, z_N$ such that*

$$z_0 + z_1 + \cdots + z_n \geq 1, \tag{4.3}$$
$$z_0 + z_1\Omega_n(td_1) + \cdots + z_N\Omega_n(td_N) \geq 0 \quad \textit{for all } t > 0. \tag{4.4}$$

Then $m_{d_1,\ldots,d_N}(\mathbb{R}^n) \leq z_0$.

Notice that Theorem 4.1 says that any feasible solution of the optimization problem

$$\inf\{\, z_0 : (z_0, z_1, \ldots, z_N) \text{ satisfies (4.3)–(4.4)} \,\}$$

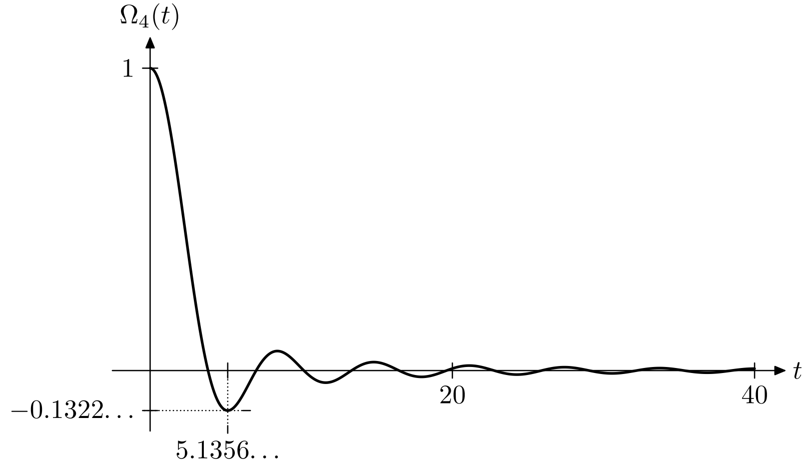

Figure 4.1. The graph of Ω_4.

provides an upper bound for $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$. This is a linear programming problem with $N+1$ variables and infinitely many constraints. In fact, we use concepts from linear programming in our proof of Theorem 4.1.

In Section 4.4 we use Theorem 4.1 in order to provide upper bounds for the densities of 1-avoiding sets, and hence also lower bounds for the measurable chromatic number of $\mathbb{R}^n$; in Section 4.5 we use the theorem to upper bound the densities of sets avoiding many distances. Finally, in Section 4.6 we generalize our results to other, slightly more general, distance graphs over $\mathbb{R}^n$ and show how the bound of Theorem 4.1 can be obtained from a suitable generalization of the theta number.

Indeed, distance-avoiding sets in $\mathbb{R}^n$ can be seen as stable sets in a distance graph defined over $\mathbb{R}^n$. Let D be a nonempty compact subset of the positive real line — in particular, $0 \notin D$. Consider the graph $G(\mathbb{R}^n, D)$ whose vertices are the points of $\mathbb{R}^n$ and in which two vertices x, $y \in \mathbb{R}^n$ are adjacent if and only if $\|x - y\| \in D$. This graph is a distance graph in $\mathbb{R}^n$, i.e., adjacency in this graph is completely characterized in terms of distance, and the stable sets of this graph correspond to sets which avoid all distances in D. It is this more general setting that we consider in Section 4.6, where we discuss how the theta number can be generalized to the graphs $G(\mathbb{R}^n, D)$.

4.2 Harmonic analysis and a proof of the main theorem

We need some basic facts from harmonic analysis in order to prove Theorem 4.1. What we present here can be found, e.g., in the book by Katznelson [25].

Let $f\colon \mathbb{R}^n \to \mathbb{C}$ be a function. We say that f is *periodic* if there exists a basis $b_1, \ldots, b_n$ of $\mathbb{R}^n$ such that

$$f(x + \alpha_1 b_1 + \cdots + \alpha_n b_n) = f(x)$$

for all $x \in \mathbb{R}^n$ and $\alpha_1, \ldots, \alpha_n$ integer. The lattice

$$\Lambda = \{\, \alpha_1 b_1 + \cdots + \alpha_n b_n : \alpha_i \in \mathbb{Z} \,\}$$

is a *periodicity lattice* of f. Any version of Λ scaled by an integer factor is also a periodicity lattice of f. For our purposes, when we work with periodic functions, the particular lattice will not be important.

A periodic function with periodicity lattice Λ repeats itself in translated copies of the *fundamental domain* $D = \{\, \lambda_1 b_1 + \cdots + \lambda_n b_n : 0 \leq \lambda_i < 1 \,\}$ of Λ centered at points of Λ. It can also be seen as a function defined over the domain $\mathbb{R}^n/\Lambda$.

Let $f\colon \mathbb{R}^n \to \mathbb{C}$ be a measurable periodic function with periodicity lattice Λ. Let D be the fundamental domain of Λ. We say that f is *square-integrable* if

$$\lim_{R\to\infty} \frac{1}{\operatorname{vol}[-R,R]^n} \int_{[-R,R]^n} |f(x)|^2\, dx = \frac{1}{\operatorname{vol} D} \int_D |f(x)|^2\, dx < \infty.$$

Notice that the condition expressed in terms of the limit above has the advantage of being independent of the periodicity lattice of f.

If f, $g\colon \mathbb{R}^n \to \mathbb{C}$ are two square-integrable periodic functions, we write

$$\langle f, g\rangle = \lim_{R\to\infty} \frac{1}{\operatorname{vol}[-R,R]^n} \int_{[-R,R]^n} f(x)\overline{g(x)}\, dx.$$

This is clearly well-defined when f and g have both the same periodicity lattice Λ. In this case, if D is the fundamental domain of Λ, we simply have

$$\langle f, g\rangle = \frac{1}{\operatorname{vol} D} \int_D f(x)\overline{g(x)}\, dx.$$

Also when f and g have different periodicity lattices is $\langle f, g\rangle$ well-defined, as can be seen after a bit of work.

The space of all square-integrable periodic functions with a common periodicity lattice Λ is a Hilbert space when endowed with the inner product $\langle \cdot, \cdot \rangle$ — let us denote it by $L^2(\mathbb{R}^n/\Lambda)$.

Let Λ^* be the *dual lattice* of Λ, that is,

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

The functions

$$e^{iu\cdot x} \qquad \text{for } u \in 2\pi\Lambda^*,$$

where $2\pi\Lambda^*$ is the lattice Λ^* scaled by 2π, form a complete orthonormal system of $L^2(\mathbb{R}^n/\Lambda)$. So, for every function $f \in L^2(\mathbb{R}^n/\Lambda)$, we have

$$f(x) = \sum_{u\in 2\pi\Lambda^*} \widehat{f}(u) e^{iu\cdot x} \tag{4.5}$$

with convergence in the L^2-norm, where $\widehat{f}(u) = \langle f, e^{iu\cdot x}\rangle$ is the *Fourier coefficient* of f *associated with* u. We also have *Parseval's formula*: If f, $g \in L^2(\mathbb{R}^n/\Lambda)$, then

$$\langle f, g\rangle = \sum_{u\in 2\pi\Lambda^*} \widehat{f}(u)\overline{\widehat{g}(u)}. \tag{4.6}$$

Actually, if $f \in L^2(\mathbb{R}^n/\Lambda)$ and $u \notin 2\pi\Lambda^*$, then $\widehat{f}(u) = \langle f, e^{iu\cdot x}\rangle = 0$. So we may omit Λ^* in (4.5) and (4.6) if we like. For instance, (4.5) becomes

$$f(x) = \sum_{u\in\mathbb{R}^n} \widehat{f}(u)e^{iu\cdot x}.$$

We just observe that only countably many of the Fourier coefficients are nonzero and that the order of summation is irrelevant. We may similarly rewrite (4.6) in order to hide Λ^*.

The observation of the last paragraph is related to the fact that we may define a space of *almost periodic functions* (cf. Katznelson [25], Chapter VI) that includes all continuous periodic functions and in which $\langle\cdot,\cdot\rangle$ plays the role of the inner product. In this space, which is however not a Hilbert space, as it is not complete, Parseval's formula still holds and functions might have Fourier coefficients supported at any countable subset of $\mathbb{R}^n$. This more complicated framework is however not needed for our purposes.

Proof of Theorem 4.1. Let $C \subseteq \mathbb{R}^n$ be a nonempty measurable set that avoids distances $d_1, \ldots, d_N$. Denote by $\chi^C\colon \mathbb{R}^n \to \{0, 1\}$ its characteristic function, whose support is exactly C. We may assume without loss of generality that χ^C is periodic.

Indeed, let R be such that $\operatorname{vol}(C \cap [-R, R]^n)/\operatorname{vol}[-R, R]^n$ is close to the upper density of C and also such that $\operatorname{vol}[-R + d, R - d]^n/\operatorname{vol}[-R, R]^n$, with $d = \max\{d_1, \ldots, d_N\}$, differs from 1 only negligibly. Then we construct a set that avoids distances $d_1, \ldots, d_N$ by tiling $\mathbb{R}^n$ with copies of $C \cap [-R + d, R - d]^n$ centered at the points of the lattice $2R\mathbb{Z}^n$. This new set has upper density close to that of C and its characteristic function is periodic. By taking R large enough, we can make the density of the new distance-avoiding set as close to $\overline{\delta}(C)$ as we like.

So χ^C is periodic and, since C is measurable, χ^C is square-integrable. For a vector $y \in \mathbb{R}^n$, we denote by $C + y$ the translation of C by y and we write $C - y = C + (-y)$. The following two properties are crucial for our argument:

$$\begin{aligned}\langle\chi^C, \mathbf{1}\rangle &= \overline{\delta}(C),\\ \langle\chi^{C-y}, \chi^C\rangle &= \overline{\delta}(C \cap (C - y)) \quad \text{for all } y \in \mathbb{R}^n,\end{aligned}$$

where $\mathbf{1}$ is the constant one function. In particular, we have $\langle\chi^C, \chi^C\rangle = \overline{\delta}(C)$ and $\langle\chi^{C-y}, \chi^C\rangle = 0$ if y has norm equal to $d_1, \ldots, d_N$. Also, $\langle\chi^C, \mathbf{1}\rangle$ is just $\widehat{\chi}^C(0)$, the Fourier coefficient of χ^C computed at 0.

Consider now the function $\varphi\colon \mathbb{R}^n \to \mathbb{C}$ such that $\varphi(y) = \langle\chi^{C-y}, \chi^C\rangle$. This is the so-called *autocorrelation function* of χ^C. Function χ^{C-y} is periodic and square-integrable and its periodicity lattice is the same as that of function χ^C. Moreover, $\chi^{C-y}(x) = \chi^C(x + y)$, so we may express the Fourier coefficients of χ^{C-y}

in terms of those of χ^C: We simply have $\widehat{\chi}^{C-y}(u) = \widehat{\chi}^C(u)e^{iu\cdot y}$. So, from Parseval's formula we have

$$\varphi(y) = \overline{\delta}(C \cap (C - y)) = \sum_{u \in \mathbb{R}^n} |\widehat{\chi}^C(u)|^2 e^{iu\cdot y} \tag{4.7}$$

for all $y \in \mathbb{R}^n$. Notice that this series converges absolutely and uniformly everywhere in $\mathbb{R}^n$.

By taking spherical averages of φ we construct a radial function $\overline{\varphi}\colon \mathbb{R}^n \to \mathbb{R}$, that is, the values of $\overline{\varphi}$ depend only on the norm of the argument. In other words, we set

$$\overline{\varphi}(y) = \frac{1}{\omega_n} \int_{S^{n-1}} \varphi(\|y\|\xi)\, d\omega(\xi).$$

Combining the expansion (4.7) for φ with formula (4.2), we obtain the following expression for $\overline{\varphi}$:

$$\overline{\varphi}(y) = \sum_{t \geq 0} \alpha(t)\Omega_n(t\|y\|), \tag{4.8}$$

where $\alpha(t)$ is the sum of $|\widehat{\chi}^C(u)|^2$ for vectors u having norm t. So the $\alpha(t)$ are real and nonnegative, also only countably many of them are nonzero. Furthermore, $\alpha(0) = |\widehat{\chi}^C(0)|^2 = \overline{\delta}(C)^2$ and $\sum_{t\geq 0} \alpha(t) = \overline{\varphi}(0) = \overline{\delta}(C)$.

So, the optimal value of the following optimization problem in variables $\alpha(t)$ gives an upper bound for the density of any measurable set that avoids the distances $d_1, \ldots, d_N$:

$$\begin{array}{ll} \sup & \alpha(0) \\ & \sum_{t\geq 0} \alpha(t) = 1, \\ & \sum_{t \geq 0} \alpha(t)\Omega_n(td_i) = 0 \quad \text{for } i = 1, \ldots, N, \\ & \alpha(t) \geq 0 \qquad\qquad\qquad\quad \text{for all } t \geq 0. \end{array} \tag{4.9}$$

Recall that above all but countably many of the variables are zero. Notice also the normalization $\sum_{t\geq 0} \alpha(t) = 1$ which we employed.

Problem (4.9) is a linear programming problem with infinitely many variables and $N + 1$ constraints. A dual of it is the problem

$$\begin{array}{ll} \inf & z_0 \\ & z_0 + z_1 + \cdots + z_N \geq 1, \\ & z_0 + z_1\Omega_n(td_1) + \cdots + z_N\Omega_n(td_N) \geq 0 \quad \text{for all } t > 0, \end{array} \tag{4.10}$$

which is a linear programming problem with $N + 1$ variables and infinitely many constraints.

As usual, weak duality holds between (4.9) and (4.10): If $\alpha(t)$ is feasible for (4.9) and $(z_0, z_1, \ldots, z_N)$ is feasible for (4.10), then

$$\alpha(0) \leq \sum_{t \geq 0} \alpha(t)(z_0 + z_1\Omega_n(td_1) + \cdots + z_N\Omega_N(td_N)) = z_0,$$

what finishes the proof of the theorem. ■

4.3 Some facts on Bessel functions

Theorem 4.1 is given in terms of the function Ω_n defined in (4.2). To apply the theorem, we need to know more about Ω_n. We know from (4.2) that $\Omega_1(t) = \cos t$, but what about larger values of n? For $n \geq 2$, expression (4.2) no longer provides a lot of information about Ω_n, for instance it is not clear how to easily compute approximations to the function evaluated at specific points. Fortunately, Ω_n admits an expression in terms of the Bessel function of the first kind, a well-known function whose properties are well-understood. In this section we shall develop this connection and as a result derive those properties of Ω_n which are of interest to us.

Our discussion on Bessel functions is oriented towards our study of Ω_n. At first it was tempting to prove most of the facts about Bessel functions that we use, but the choice of what to prove and what not to prove seemed artificial, so proofs of facts concerning Bessel functions are mostly omitted. We refer the reader to the book by Watson [55] for a thorough account on the theory of Bessel functions — we will follow this book in the rest of this section. Chapter 4 of the book by Andrews, Askey, and Roy [1] also covers most of the material we need.

Let α be a real number. The *Bessel function of the first kind* of *order* α is the function

$$J_\alpha(t) = \sum_{k=0}^{\infty} \frac{(-1)^k (t/2)^{2k+\alpha}}{k!\,\Gamma(k+\alpha+1)} \tag{4.11}$$

of one complex variable t. The series on the right-hand side — a hypergeometric series — converges absolutely for every t (except for $t = 0$ when $\alpha < 0$) and uniformly in any compact subset of $\mathbb{C}$ (not containing the origin, when $\alpha < 0$). This follows from an application of the ratio test (cf. Watson [55], Section 3.13). Hence J_α is an analytic function of t for all t (except for $t = 0$, when $\alpha < 0$), and differentiation and integration may be carried out term-by-term. By using term-by-term differentiation, one may easily check that J_α is a solution of the differential equation

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - \alpha^2)y = 0, \tag{4.12}$$

which is called *Bessel's equation for functions of order* α.

Lommel (cf. Watson [55], Section 3.1) provided an alternative definition for J_α when $\alpha > -1/2$, which is

$$J_\alpha(t) = \frac{1}{\pi^{1/2}\Gamma(\alpha+1/2)}\left(\frac{t}{2}\right)^\alpha \int_0^\pi \cos(t\cos\theta)\sin^{2\alpha}\theta\, d\theta. \tag{4.13}$$

Notice $\alpha > -1/2$ is necessary for the integral to converge. For a proof of the equivalence between (4.11) and (4.13), see Section 3.3 of Watson [55].

From Lommel's definition we may derive a relation between Ω_n and $J_{(n-2)/2}$ when $n \geq 2$. For suppose $n \geq 2$. Notice that the function $e^{iu\cdot\xi}$, for $u \in \mathbb{R}^n$ and $\xi \in S^{n-1}$, is (as a function of ξ) a zonal spherical function with pole $u/\|u\|$. (For basic facts on zonal spherical functions, see Section 3.4.) So we may write, equivalently to (4.2),

$$\Omega_n(t) = \frac{\omega_{n-1}}{\omega_n} \int_{-1}^{1} e^{ist}(1-s^2)^{(n-3)/2}\,ds = \frac{\Gamma(n/2)}{\pi^{1/2}\Gamma((n-1)/2)} \int_{-1}^{1} \cos(st)(1-s^2)^{(n-3)/2}\,ds$$

for all $t \geq 0$.

Writing $\alpha = (n-2)/2$ and making the substitution $s = \cos\theta$ we get

$$\Omega_n(t) = \frac{\Gamma(n/2)}{\pi^{1/2}\Gamma(\alpha+1/2)} \int_0^{\pi} \cos(t\cos\theta)\sin^{2\alpha}\theta\,d\theta,$$

and comparing with (4.13) we obtain

$$\Omega_n(t) = \Gamma\Big(\frac{n}{2}\Big)\Big(\frac{2}{t}\Big)^{(n-2)/2} J_{(n-2)/2}(t) \tag{4.14}$$

for $t > 0$, and $\Omega_n(0) = 1$, as can be seen from (4.2). This expression was given by Schoenberg [43].

Actually, (4.14) also applies when $n = 1$, as then we have $\Omega_1(t) = \cos t$ and

$$J_{-1/2}(t) = \Big(\frac{2}{\pi t}\Big)^{1/2} \cos t$$

(cf. Watson [55], equation (6) in Section 3.4).

From (4.11) one can approximate J_α to any desired degree of accuracy. According to Watson [55], Section 20.1, Bessel [4] was the first to provide a table for J_0 and J_1, which are called I^0 and I^1 in his paper, respectively. Today, most mathematical software allows for the efficient computation of J_α to very high accuracy. Watson's book also contains tables for J_α for some values of α.

So expression (4.14) at least helps us evaluate Ω_n. It also provides us with the means to investigate other properties of Ω_n — that is what we do now as we investigate the behavior of the extrema of Ω_n.

The following formula will be of great help in our investigation:

$$\frac{dt^{-\alpha}J_\alpha(t)}{dt} = -t^{-\alpha}J_{\alpha+1}(t).$$

This follows immediately from (4.11). By using this formula we conclude that

$$\Omega_n'(t) = -\Gamma\Big(\frac{n}{2}\Big)\Big(\frac{2}{t}\Big)^{(n-2)/2} J_{n/2}(t) \tag{4.15}$$

for $t > 0$. So the positive extrema of Ω_n correspond to the positive zeros of $J_{n/2}$. Hence, in order to understand the extrema of Ω_n, we first state some properties of the zeros of J_α.

It can be shown (cf. Watson [55], Section 15.2) that J_α has infinitely many positive zeros. Hence also J_α' has infinitely many positive zeros. Then, since J_α is analytic everywhere, except maybe at the origin, the set $\{ t > 0 : J_\alpha(t) = 0 \}$ has no cluster point in the positive real line, and hence we may consider the sequence

$$0 < j_{\alpha,1} < j_{\alpha,2} < \cdots$$

of all positive zeros of J_α given in order of magnitude. Similarly, we consider the sequence

$$0 < j'_{\alpha,1} < j'_{\alpha,2} < \cdots$$

of all positive zeros of J'_α. We also observe that J_α has no double zero, except maybe for the origin. That is to say that no complex number $t \neq 0$ is a zero of both J_α and J'_α, as otherwise repeated differentiation of (4.12) would yield that all derivatives of J_α vanish at t, and from the analyticity of J_α we would then have that J_α is identically zero, a contradiction.

Another important property of the positive zeros of J_α is the *interlacing property* (cf. Watson [55], Section 15.22). It states that the zeros of J_α are interlaced with those of $J_{\alpha+1}$, i.e., between any two zeros of J_α there is a zero of $J_{\alpha+1}$ and vice versa. When $\alpha > -1$, the case which is of interest to us, we even know that $j_{\alpha,1}$ is to the left of $j_{\alpha+1,1}$, so that the interlacing property implies that

$$0 < j_{\alpha,1} < j_{\alpha+1,1} < j_{\alpha,2} < j_{\alpha+1,2} < j_{\alpha,3} < \cdots . \tag{4.16}$$

From what we know so far about the zeros of J_α, we already have a quite clear picture of the general behavior of the functions Ω_n, and we can see that this picture agrees with the behavior illustrated in Figure 4.1. Indeed, from (4.13) one may easily check that for all small enough positive t, $J_\alpha(t) > 0$ when $\alpha > -1/2$. Then, from (4.15), $\Omega'_n(t) < 0$ for all small enough positive t. Since Ω'_n changes signs at the points $j_{n/2,1}$, $j_{n/2,2}$, ..., if we put $j_{n/2,0} = 0$, we conclude that Ω_n is decreasing in the intervals $[j_{n/2,k}, j_{n/2,k+1})$ when $k \geq 0$ is even and is increasing in any such interval when $k \geq 0$ is odd. Also, $j_{n/2,1}$, $j_{n/2,2}$, ... are the extrema of Ω_n, and from the above discussion it is clear that $j_{n/2,k}$ is a minimum of Ω_n if k is odd, and a maximum if k is even. From (4.16) it then follows that every minimum of Ω_n is negative and that every maximum is positive. All this can be trivially seen to apply for $n = 1$ as well, as in this case $\Omega_1(t) = \cos t$ and $\Omega'_1(t) = -\sin t$.

We show now a result about the extrema of Ω_n for $n \geq 2$ which will play an important role in Section 4.4. Namely, we will show that the extrema of Ω_n decrease in absolute value as t increases, or in other words

$$|\Omega_n(j_{n/2,1})| > |\Omega_n(j_{n/2,2})| > |\Omega_n(j_{n/2,3})| > \cdots . \tag{4.17}$$

In particular, this implies that $j_{n/2,1}$ is the global minimum of Ω_n. The assumption $n \geq 2$ is here crucial, as all extrema of Ω_1 have the same magnitude.

To prove (4.17), we first prove a similar result for J_α when $\alpha \geq 0$. Namely, we show that

$$|J_\alpha(j'_{\alpha,1})| > |J_\alpha(j'_{\alpha,2})| > |J_\alpha(j'_{\alpha,3})| > \cdots . \tag{4.18}$$

This is proven in Watson [55], Section 15.31, for the more general *cylinder functions*, which include the Bessel function of the first kind as a special case. Anyway, we present a proof below since this is the main step in the proof of (4.17).

To prove (4.18) we use the inequalities

$$j_{\alpha,1} > \alpha, \quad j'_{\alpha,1} > \alpha, \tag{4.19}$$

valid for $\alpha \geq 0$ (cf. Watson [55], Section 15.3). Actually, now we only need the latter, but the former will also be important elsewhere.

Consider the function

$$L(t) = (J_\alpha(t))^2 + \frac{t^2(J'_\alpha(t))^2}{t^2 - \alpha^2}.$$

Since J_α is a solution of (4.12), we have

$$L'(t) = \frac{-2t^3(J'_\alpha(t))^2}{(t^2 - \alpha^2)^2},$$

and so L is decreasing for $t > \alpha$. Together with (4.19), this implies that

$$L(j'_{\alpha,1}) > L(j'_{\alpha,2}) > L(j'_{\alpha,3}) > \cdots,$$

whence (4.18) follows.

From (4.18) we have that, for $k = 1, 2, \ldots,$

$$|J_\alpha(j'_{\alpha,k})| > |J_\alpha(t)| \qquad \text{for all } t > j'_{\alpha,k}, \tag{4.20}$$

and with this we may prove (4.17). Indeed, let $\alpha = (n-2)/2$ and fix some integer $k > 0$. Let l be such that $j'_{\alpha,l}$ is the first zero of J'_α to the right of $j_{\alpha,k}$. Clearly, $j'_{\alpha,l} < j_{\alpha,k+1}$. So, since from (4.15) and (4.16), $j_{n/2,k} = j_{\alpha+1,k}$ is the only extremum of Ω_n between $j_{\alpha,k}$ and $j_{\alpha,k+1}$, together with (4.14) and (4.20) we obtain

$$|\Omega_n(j_{n/2,k})| \geq |\Omega_n(j'_{\alpha,l})| > |\Omega_n(t)| \qquad \text{for all } t \geq j_{n/2,k+1},$$

and (4.17) follows.

Finally, there are two more facts we wish to state about Bessel functions and Ω_n. The first one provides a bound for the values of $J_\alpha(t)$ for $t \geq 0$. From (4.20) we see that $J_\alpha(t)$ is bounded at infinity for $\alpha \geq 0$. But more is known: We indeed have that

$$|J_\alpha(t)| \leq 1 \quad \text{and} \quad |J_{\alpha+1}(t)| \leq 1/\sqrt{2} \tag{4.21}$$

whenever $\alpha \geq 0$ (cf. Watson [55], equation (10) in Section 13.42).

The last fact concerns the asymptotic behavior of Ω_n. For $n \geq 2$, we have

$$\lim_{t \to \infty} \Omega_n(t) = 0. \tag{4.22}$$

For $n \geq 3$, this follows from (4.14) together with the fact that J_α is bounded at infinity for $\alpha \geq 0$. However, the statement follows for all $n \geq 2$ from the asymptotic formula for J_α, for $\alpha \geq 0$ (cf. Watson [55], equation (1) in Section 7.21), which already shows that $J_\alpha(t) \to 0$ as $t \to \infty$.

4.4 Sets that avoid one distance

We show now how to use Theorem 4.1 to find upper bounds to the densities of sets that avoid one distance, which without loss of generality we assume to be 1, as $m_1(\mathbb{R}^n) = m_d(\mathbb{R}^n)$ for any $d > 0$. So we present upper bounds for $m_1(\mathbb{R}^n)$.

Recall that we have shown in Section 4.3 that the global minimum of Ω_n occurs at $j_{n/2,1}$, the first positive zero of $J_{n/2}$, and that it is a negative number. The following result then follows after a simple application of Theorem 4.1.

Theorem 4.2. *For $n \geq 1$ we have*

$$m_1(\mathbb{R}^n) \leq \frac{\Omega_n(j_{n/2,1})}{\Omega_n(j_{n/2,1}) - 1} \qquad \textit{and} \qquad \chi_{\mathrm{m}}(\mathbb{R}^n) \geq \frac{\Omega_n(j_{n/2,1}) - 1}{\Omega_n(j_{n/2}, 1)}. \tag{4.23}$$

Proof. Take

$$z_0 = \frac{-\Omega_n(j_{n/2,1})}{1 - \Omega_n(j_{n/2,1})} \qquad \text{and} \qquad z_1 = \frac{1}{1 - \Omega_n(j_{n/2,1})} \tag{4.24}$$

and apply Theorem 4.1 and (4.1). ∎

The bound given by Theorem 4.2 is actually the best possible bound that can be obtained from Theorem 4.1. In other words, it can be shown that z_0 and z_1 given as in (4.24) are an optimal solution of the optimization problem (4.9). This is easy to see since, in the case of sets avoiding one distance, problem (4.9) has only two variables, and solving it amounts to solving a two-by-two system.

The bounds of Theorem 4.2 can be computed, with the help of a computer, to any desired accuracy for any fixed n. In Table 4.2 we show the numbers so obtained, compared to previous best known upper bounds for $m_1(\mathbb{R}^n)$ and lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 2, \ldots, 24$. For $n = 1$, the bounds obtained from Theorem 4.2 are tight, yielding the trivial results $m_1(\mathbb{R}) = 1/2$ and $\chi_{\mathrm{m}}(\mathbb{R}) = 2$. The upper bounds for $m_1(\mathbb{R}^n)$ from Theorem 4.2 are better than previously known bounds for $n = 3, \ldots, 24$. The lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ coming from the theorem improve on the previously known lower bounds for $n = 4, \ldots, 24$.

We also observe that the bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ in (4.23) is actually the bound obtained in Section 3.5c with our approach of using bounds for the stability number of distance graphs on the sphere to bound $\chi_{\mathrm{m}}(\mathbb{R}^n)$. In that section, however, we could only conclude that the bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ in (4.23) was a lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^{n+1})$, thus obtaining a weaker result.

From Theorem 4.2 we may prove that $m_1(\mathbb{R}^n)$ decreases exponentially with n, and so that $\chi_{\mathrm{m}}(\mathbb{R}^n)$ increases exponentially with n. Indeed, from (4.19) we know that $j_{n/2,1} > n/2 > (n-2)/2$ and this, together with (4.21) and (4.14), implies that, for $n \geq 2$,

$$|\Omega_n(j_{n/2,1})| \leq \Gamma\Big(\frac{n}{2}\Big)\Big(\frac{2}{(n-2)/2}\Big)^{(n-2)/2}.$$

n	Previous upper bound for $m_1(\mathbb{R}^n)$	Bound for $m_1(\mathbb{R}^n)$ from Theorem 4.2	Previous lower bound for $\chi_m(\mathbb{R}^n)$	Bound for $\chi_m(\mathbb{R}^n)$ from Theorem 4.2
2	0.27906976	0.28711938	5	4
3	0.18750000	0.17846503	6	6
4	0.12800000	0.11682584	8	9
5	0.09539473	0.07933457	11	13
6	0.07081295	0.05537340	15	19
7	0.05311365	0.03948200	19	26
8	0.03419769	0.02863560	30	35
9	0.02882153	0.02106105	35	48
10	0.02234835	0.01567163	45	64
11	0.01789325	0.01177701	56	85
12	0.01437590	0.00892554	70	113
13	0.01203324	0.00681436	84	147
14	0.00981770	0.00523614	102	191
15	0.00841374	0.00404638	119	248
16	0.00677838	0.00314283	148	319
17	0.00577854	0.00245212	174	408
18	0.00518111	0.00192105	194	521
19	0.00380311	0.00151057	263	663
20	0.00318213	0.00119181	315	840
21	0.00267706	0.00094321	374	1,061
22	0.00190205	0.00074859	526	1,336
23	0.00132755	0.00059567	754	1,679
24	0.00107286	0.00047513	933	2,105

Table 4.2. The table shows the best upper bounds for $m_1(\mathbb{R}^n)$ and the best lower bounds for $\chi_m(\mathbb{R}^n)$ previously known, together with the upper bounds for $m_1(\mathbb{R}^n)$ and the lower bounds for $\chi_m(\mathbb{R}^n)$ coming from Theorem 4.2. The upper bound for $m_1(\mathbb{R}^2)$ was given by Székely [52] and all other upper bounds for $m_1(\mathbb{R}^n)$ were given by Székely and Wormald [54]. The lower bounds for $\chi_m(\mathbb{R}^n)$ for $n = 2$ and 3 were given by Falconer [13]. All other lower bounds for $\chi_m(\mathbb{R}^n)$ were given by Székely and Wormald [54].

Using Stirling's formula $\Gamma(\alpha+1) \sim \sqrt{2\pi\alpha}(\alpha/e)^\alpha$, we then have that, asymptotically,

$$|\Omega_n(j_{n/2,1})| \leq \sqrt{\pi(n-2)}\Big(\frac{2}{e}\Big)^{(n-2)/2} \leq \Big(\frac{2}{e} + o(1)\Big)^{n/2}. \tag{4.25}$$

So, from Theorem 4.2, we derive the following asymptotic bounds:

$$m_1(\mathbb{R}^n) \leq \Big(\frac{2}{e} + o(1)\Big)^{n/2} = (0.857\ldots + o(1))^n$$

and

$$\chi_{\mathrm{m}}(\mathbb{R}^n) \geq \Big(\frac{2}{e} + o(1)\Big)^{-n/2} = (1.165\ldots + o(1))^n.$$

Our bounds are exponential, but not better than the best known asymptotic bounds for both $m_1(\mathbb{R}^n)$ and $\chi_{\mathrm{m}}(\mathbb{R}^n)$, which were discussed in Section 1.2.

4.4a. Simplices and improved bounds. By strengthening problem (4.9), which is used in the proof of Theorem 4.1, it is possible to get better bounds for $m_1(\mathbb{R}^n)$ than those provided by Theorem 4.2. The idea here is to add to (4.9) constraints coming from unit simplices in $\mathbb{R}^n$. In what follows we show how to modify the proof of Theorem 4.1 in order to make use of such new constraints. Understanding the proof of that theorem is thus essential in order to read this section.

Let $C \subseteq \mathbb{R}^n$ be a measurable 1-avoiding set and suppose $v_1, \ldots, v_{n+1}$ are the vertices of a unit simplex in $\mathbb{R}^n$, so that $\|v_i - v_j\| = 1$ if $i \neq j$. Then, if $i \neq j$, we have $(C - v_i) \cap (C - v_j) = \emptyset$. So

$$\overline{\delta}(C) \geq \overline{\delta}\left(\bigcup_{i=1}^{n+1} C \cap (C - v_i)\right) = \sum_{i=1}^{n+1} \overline{\delta}(C \cap (C - v_i)). \tag{4.26}$$

If we assume, as we do in the proof of Theorem 4.1, that C is such that its characteristic function $\chi^C\colon \mathbb{R}^n \to \{0,1\}$ is periodic, then we may consider the function $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\varphi(y) = \langle \chi^{C-y}, \chi^C \rangle,$$

the autocorrelation function of χ^C, which we also used in the proof of Theorem 4.1. Then (4.26) implies that φ satisfies

$$\varphi(v_1) + \cdots + \varphi(v_{n+1}) \leq \varphi(0). \tag{4.27}$$

In the proof of Theorem 4.1, to obtain problem (4.9), we first take spherical averages of φ, i.e., we consider the function $\overline{\varphi}\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\overline{\varphi}(y) = \frac{1}{\omega_n} \int_{S^{n-1}} \varphi(\|y\|\xi)\, d\omega(\xi).$$

Inequality (4.27) then implies that also $\overline{\varphi}$ satisfies

$$\overline{\varphi}(v_1) + \cdots + \overline{\varphi}(v_{n+1}) \leq \overline{\varphi}(0). \tag{4.28}$$

To prove this, we need a simple fact about the relation between the Haar measure over the orthogonal group $\mathrm{O}(\mathbb{R}^n)$ and the measure ω over S^{n-1}. (For basic facts on $\mathrm{O}(\mathbb{R}^n)$ and the Haar measure defined over it, see Section 3.4.) Let μ be the Haar measure over $\mathrm{O}(\mathbb{R}^n)$, normalized so that $\mu(\mathrm{O}(\mathbb{R}^n)) = 1$. Then for any $x \in S^{n-1}$ and measurable $X \subseteq S^{n-1}$,

$$\mu(\{\, T \in \mathrm{O}(\mathbb{R}^n) : Tx \in X \,\}) = \frac{\omega(X)}{\omega_n}$$

(cf. Theorem 3.7 in Mattila [32]). So averaging on the sphere with respect to ω is the same as symmetrizing with respect to $\mathrm{O}(\mathbb{R}^n)$, i.e.,

$$\overline{\varphi}(y) = \frac{1}{\omega_n} \int_{S^{n-1}} \varphi(\|y\|\xi)\, d\omega(\xi) = \int_{\mathrm{O}(\mathbb{R}^n)} \varphi(Ty)\, d\mu(T). \tag{4.29}$$

Then, since if $v_1, \ldots, v_{n+1}$ are the vertices of a unit simplex, also $Tv_1, \ldots, Tv_{n+1}$ are the vertices of a unit simplex for every $T \in \mathrm{O}(\mathbb{R}^n)$, by symmetrizing both sides of (4.27) with respect to the orthogonal group we get (4.28).

Let now $\mathcal{S}$ be a finite set of $(n+1)$-tuples of numbers which correspond to the squared norms of the vertices of unit simplices in $\mathbb{R}^n$. In other words, for each $s = (s_1, \ldots, s_{n+1}) \in \mathcal{S}$ there is a unit simplex in $\mathbb{R}^n$ with vertices $v_1, \ldots, v_{n+1}$ such that $\|v_i\|^2 = s_i$. Since $\overline{\varphi}$ is radial, i.e., since its values depend only on the norm of its argument, inequality (4.28) does not depend on the actual vectors $v_1, \ldots, v_{n+1}$, but rather on their norms. So for each $s \in \mathcal{S}$ we have

$$\overline{\varphi}(\sqrt{s_1}) + \cdots + \overline{\varphi}(\sqrt{s_{n+1}}) \leq \overline{\varphi}(0), \tag{4.30}$$

where $\overline{\varphi}(t)$, for a given number $t \geq 0$, is the common value of $\overline{\varphi}$ for vectors of norm t.

Recall that we have expression (4.8) for $\overline{\varphi}$, so the left-hand side of (4.30) can be equivalently written as

$$\sum_{i=1}^{n+1} \sum_{t \geq 0} \alpha(t) \Omega_n(t\sqrt{s_i}) = \sum_{t \geq 0} \alpha(t) \sum_{i=1}^{n+1} \Omega_n(t\sqrt{s_i}).$$

Taking into account our normalization $\sum_{t \geq 0} \alpha(t) = 1$, one then obtains the optimization problem

$$\begin{array}{lll} \sup & \alpha(0) & \\ & \sum_{t\geq 0} \alpha(t) = 1, & \\ & \sum_{t\geq 0} \alpha(t)\Omega_n(t) = 0, & \\ & \sum_{t\geq 0} \alpha(t) \sum_{i=1}^{n+1} \Omega_n(t\sqrt{s_i}) \leq 1 & \text{for } s \in \mathcal{S}, \\ & \alpha(t) \geq 0 & \text{for all } t \geq 0, \end{array} \tag{4.31}$$

the optimal value of which provides an upper bound for $m_1(\mathbb{R}^n)$.

A dual of problem (4.31) is the optimization problem

$$\begin{array}{lll} \inf & z_0 + \sum_{s\in\mathcal{S}} w_s & \\ & z_0 + z_1 + \sum_{s\in\mathcal{S}} (n+1) w_s \geq 1, & \\ & z_0 + z_1\Omega_n(t) + \sum_{s\in\mathcal{S}} w_s \sum_{i=1}^{n+1} \Omega_n(t\sqrt{s_i}) \geq 0 & \text{for all } t > 0, \\ & w_s \geq 0 & \text{for all } s \in \mathcal{S}. \end{array} \tag{4.32}$$

Weak duality holds between (4.31) and (4.32), and so any feasible solution of (4.32) provides an upper bound for $m_1(\mathbb{R}^n)$.

By solving problem (4.32) with an appropriate set $\mathcal{S}$ of simplices for each dimension, we could improve on the existing bounds for $m_1(\mathbb{R}^n)$ for $n = 2, \ldots, 24$,

n	Previous upper bound for $m_1(\mathbb{R}^n)$	Bound for $m_1(\mathbb{R}^n)$ from problem (4.32)	Previous lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$	Bound for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ from problem (4.32)
2	0.27906976	0.26841200	5	4
3	0.18750000	0.16560900	6	7
4	0.12800000	0.11293700	8	9
5	0.09539473	0.07528450	11	14
6	0.07081295	0.05157090	15	20
7	0.05311365	0.03612710	19	28
8	0.03419769	0.02579710	30	39
9	0.02882153	0.01873240	35	54
10	0.02234835	0.01380790	45	73
11	0.01789325	0.01031660	56	97
12	0.01437590	0.00780322	70	129
13	0.01203324	0.00596811	84	168
14	0.00981770	0.00461051	102	217
15	0.00841374	0.00359372	119	279
16	0.00677838	0.00282332	148	355
17	0.00577854	0.00223324	174	448
18	0.00518111	0.00177663	194	563
19	0.00380311	0.00141992	263	705
20	0.00318213	0.00113876	315	879
21	0.00267706	0.00091531	374	1,093
22	0.00190205	0.00073636	526	1,359
23	0.00132755	0.00059204	754	1,690
24	0.00107286	0.00047489	933	2,106

Table 4.3. The table shows the best upper bounds for $m_1(\mathbb{R}^n)$ and the best lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ previously known, together with the upper bounds for $m_1(\mathbb{R}^n)$ and lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ coming from problem (4.32) with the use of specific lists of simplices, as described below. The upper bound for $m_1(\mathbb{R}^2)$ was given by Székely [52] and all other upper bounds for $m_1(\mathbb{R}^n)$ were given by Székely and Wormald [54]. The lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 2$ and 3 were given by Falconer [13]. All other lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ were given by Székely and Wormald [54].

and these improved bounds also imply improved bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for dimensions $n = 3, \ldots, 24$. These new bounds are summarized in Table 4.3. We now explain how to solve problem (4.32) and also the method we used to choose the sets of simplices.

To find a (good) feasible solution for (4.32) for a given set $\mathcal{S}$ of simplices we use the computer: A simple analytical solution like the one we provided in Theorem 4.2 seems to be no longer possible. To be able to use the computer, we need to discretize problem (4.32), which is infinite, and work with numerical approximations of Ω_n.

An explanation of how to proceed with this discretization and get rigorous bounds at the end is thus in order.

To discretize (4.32), we first consider only the constraints

$$S(t) = z_0 + z_1\Omega_n(t) + \sum_{s\in\mathcal{S}} w_s \sum_{i=1}^{n+1} \Omega_n(t\sqrt{s_i}) \geq 0 \tag{4.33}$$

for $t \in (0, L]$ for some $L > 0$. Then we discretize the interval $(0, L]$ by considering only the points in it that are of the form $k\varepsilon$, where $k > 0$ is an integer and $\varepsilon > 0$ is a suitably small number. We compute the necessary values of Ω_n with d digits of precision and solve the resulting linear program with the help of a computer.

We then obtain a solution (z_0, z_1, w) of this linear program. This will probably not be a feasible solution of (4.32), but for some $\eta \geq 0$, $(z_0 + \eta, z_1, w)$ will be feasible. All we have to do is estimate η and hope that, if L is large enough and ε small enough, η will also be small.

To estimate η, we observe first that if we choose $L' \geq L$ large enough, since we know $\Omega_n(t) \to 0$ as $t \to \infty$, $S(t)$ will be small enough for all $t \geq L'$. To estimate how small it will be, we need only know how to bound the value of $|\Omega_n(t)|$ for $t \geq L'$ for any given $L' > 0$. One approach is to use (4.14), but an approach that provides better bounds is as follows: Estimate the rightmost zero z of $\Omega_n'(t)$ that lies in the interval $[0, L']$ and then observe that $|\Omega_n(t)| \leq |\Omega_n(z)|$ for all $t \geq z$, as follows from (4.17).

Next, we need to estimate the minimum of $S(t)$ for $t \in (0, L']$. One approach to do so that works well with small numbers of simplices is to compute numerical approximations to the zeros of the derivative of $S(t)$ which lie in the interval $(0, L']$. We may then evaluate $S(t)$ at these points, and in doing so compute its minimum in the interval $(0, L']$.

Finally, in all our computations we work with some finite precision d. This means that, for instance, if $s \in \mathcal{S}$, instead of computing $\Omega_n(t\sqrt{s_i})$ when needed, we compute an approximation to Ω_n evaluated at an approximation of $t\sqrt{s_i}$. If we choose our precision d to be large enough, however, the errors we make will not add too much to η. One way to estimate how good our precision has to be is to use the fact that

$$|\Omega_n(t) - \Omega_n(u)| \leq |t - u|$$

for all t, $u \geq 0$.

This will follow from the mean value theorem once we establish that

$$|\Omega_n'(t)| \leq 1 \tag{4.34}$$

for all $t \geq 0$. For $n = 2$ this is clear from (4.15) and (4.21). For $n \geq 4$ and $t \geq n$, (4.34) also follows from (4.15), since we know from (4.21) that $|J_{n/2}(t)| \leq 1$ for all $t \geq 0$ and since $\Gamma(\alpha+1) \leq \alpha^\alpha$ for all $\alpha \geq 1$. For $n = 3$ we may check numerically that $\Gamma(3/2)(2/3)^{1/2} \leq 1$, and so (4.34) also follows for all $t \geq n$.

Now, for $0 \leq t \leq n$, we rewrite (4.15) by comparing it with formula (4.14) for Ω_{n+2}, obtaining

$$\Omega_n'(t) = -\frac{t}{n}\Omega_{n+2}(t).$$

Then, since $|\Omega_{n+2}(t)| \leq 1$ for all $t \geq 0$, (4.34) follows.

We now briefly explain how to find good simplices for use in (4.32). For $n \geq 4$, we use only one simplex: The unit simplex centered at the origin. All vertices of this simplex have the same norm $\sqrt{1/2 - 1/(2n+2)}$, so (4.33) amounts to

$$z_0 + z_1\Omega_n(t) + w(n+1)\Omega_n(t\sqrt{1/2 - 1/(2n+2)}) \geq 0 \quad \text{for all } t > 0.$$

For $n = 2$, we take as $\mathcal{S}$ the set consisting of the following triples (a, b, c) of squared norms:

$$\begin{aligned}&(2.4, 2.4, 0.3603\ldots),\\ &(3.1, 3.1, 6.5240\ldots), \text{ and}\\ &(3.7, 3.7, 7.4171\ldots).\end{aligned} \tag{4.35}$$

The last number in the above triples is a root of

$$3(a^2 + b^2 + c^2 + 1) - (a + b + c + 1)^2. \tag{4.36}$$

This ensures that the determinant of the matrix

$$\begin{pmatrix} a & \frac{1}{2}(a+b-1) & \frac{1}{2}(a+c-1) \\ \frac{1}{2}(a+b-1) & b & \frac{1}{2}(b+c-1) \\ \frac{1}{2}(a+c-1) & \frac{1}{2}(b+c-1) & c \end{pmatrix} \tag{4.37}$$

is zero for each triple in (4.35). The matrix above is also positive semidefinite for each such triple, so if one sees it as the Gram matrix of vectors v_1, v_2, and $v_3 \in \mathbb{R}^2$ with squared norms a, b, and c, one has that v_1, v_2, and v_3 are the vertices of a unit simplex in $\mathbb{R}^2$.

To find the triples in (4.35) we generated all triples (a, b, c) with a, $b = 0.1j$ for $j = 1, \ldots, 40$ and c a root of (4.36) such that matrix (4.37) is positive semidefinite. We then observed that only the triples in (4.35) were used in the solution of (4.32), in the sense that only for these triples the associated variables were nonzero.

For $n = 3$, we used a similar approach, which lead us to the set $\mathcal{S}$ of quadruples (a, b, c, d) of squared norms which consists of the quadruples

$$\begin{aligned}&(0.3, 0.4, 0.4, 0.4171\ldots),\\ &(1.9, 1.9, 1.9, 0.1893\ldots), \text{ and}\\ &(2.0, 2.0, 2.0, 0.2251\ldots),\end{aligned}$$

where the last number in each quadruple above is a root of

$$4(a^2 + b^2 + c^2 + d^2 + 1) - (a + b + c + d + 1)^2.$$

A similar approach to the one described above to generate simplices for $n = 2$ and 3 could in principle be used to generate simplices for higher dimensions. The computational cost of such an approach however becomes prohibitively large as n increases, and we observed that the gains are not substantial.

4.5 Sets that avoid many distances

Two important results are known about the behavior of $m_{d_1,\dots,d_N}(\mathbb{R}^n)$ in relation to the distances $d_1, \dots, d_N$. The first result we consider is the following theorem of Furstenberg, Katznelson, and Weiss [17].

Theorem 4.3. *If $C \subseteq \mathbb{R}^2$ has positive upper density, then there is a number d_0 such that for all $d > d_0$ there are points $x, y \in C$ with $\|x - y\| = d$.*

In other words, the theorem says that a planar set that avoids an unbounded sequence of distances has upper density zero. This was originally a conjecture made by Székely [53].

Bourgain [6] and Falconer and Marstrand [15] also presented proofs of this theorem. Both their papers predate the paper of Furstenberg, Katznelson, and Weiss [17], which only appeared in 1990. Their result was known as far back as 1983 though, as Erdős [12] states in his paper included in the proceedings of the June 26–July 2, 1983 *Oberwolfach Meeting on Measure Theory* that "the problem remained unsolved until a few weeks ago when B. Weiss proved Székely's conjecture". It is interesting to notice that Erdős attributes the proof to Weiss alone; Falconer and Marstrand [15] follow Erdős, citing his Oberwolfach paper, but Bourgain [6] already cites a preprint by Katznelson and Weiss as the origin of the result, and refers to the theorem as the "Katznelson-Weiss theorem". Note also that the paper by Furstenberg, Katznelson, and Weiss [17] contains also a result about sets avoiding triangles, which is related to Theorem 4.3. More recently, Bukh [8] proved a stronger result which implies Theorem 4.3; we will consider it in more detail below. Also Quas [38] presented a proof of Theorem 4.3, which is based on the second moment method in probability.

The second theorem is a theorem of Falconer [14] very similar to Theorem 4.3. This theorem also settled a conjecture of Székely [53], and was proven using a variant of the method of Falconer and Marstrand [15].

Theorem 4.4. *Let $d_1, d_2, \dots$ be a sequence of positive numbers which converges to zero. Then $m_{d_1,\dots,d_N}(\mathbb{R}^2) \to 0$ as $N \to \infty$.*

Recall that in Section 3.6 we presented an analogue of this theorem for distance graphs on the sphere. The proof of the above theorem we present later on is closely related to the proof of this analogue for the sphere.

Note that, although both theorems are stated and have been proved for $\mathbb{R}^2$, they are valid for all $n \geq 2$ since we have

$$m_{d_1,\dots,d_N}(\mathbb{R}^{n+1}) \leq m_{d_1,\dots,d_N}(\mathbb{R}^n) \tag{4.38}$$

for $n \geq 1$. This follows from the fact that, if a measurable set $C \subseteq \mathbb{R}^{n+1}$ avoids distances $d_1, \ldots, d_N$, then the intersection of C with any translation of the hyperplane $\{ x \in \mathbb{R}^{n+1} : x_{n+1} = 0 \}$ gives rise to a subset of $\mathbb{R}^n$ which also avoids distances $d_1, \ldots, d_N$, and which therefore has upper density at most $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$. This implies that also C has upper density at most $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$.

Both theorems fail, however, for $n = 1$. Indeed, the set

$$C = \bigcup_{k \in \mathbb{Z}} (2k, 2k+1)$$

avoids every odd integer distance, but has upper density $1/2$, and so Theorem 4.3 is false for $\mathbb{R}$. To see that also Theorem 4.4 is false for $\mathbb{R}$, we use the following argument by Falconer [14]: Consider the sequence $1, 3^{-1}, 3^{-2}, \ldots$, which converges to zero. For each integer $p \geq 1$, the set $\{ 3^{-p} x : x \in C \}$ avoids distances $1, 3^{-1}, \ldots, 3^{-p}$, and has upper density $1/2$, hence it is clear that Theorem 4.4 is also false for $\mathbb{R}$.

Finally, notice that in the statement of Theorem 4.4 we consider ever-growing, but always finite, sets of distances to be forbidden. This is so because if a planar set avoids any given zero-convergent sequence of distances, then it cannot have positive upper density. This follows from Steinhaus' theorem, a theorem in real analysis which states that, if a subset C of $\mathbb{R}^2$ has positive measure, then the difference set $\{ x - y : x, y \in C \}$ contains a neighborhood of the origin.

Recently, Bukh [8] showed that, as the ratios $d_2/d_1, \ldots, d_N/d_{N-1}$ between the distances $d_1, \ldots, d_N$ go to infinity, so does $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$ go to $(m_1(\mathbb{R}^n))^N$. He also showed that $m_{d_1,\ldots,d_N}(\mathbb{R}^n) \geq (m_1(\mathbb{R}^n))^N$, whatever the distances $d_1, \ldots, d_N$ might be, hence we have

$$\inf_{d_1,\ldots,d_N > 0} m_{d_1,\ldots,d_N}(\mathbb{R}^n) = (m_1(\mathbb{R}^n))^N.$$

Bukh's result provides alternative proofs of both Theorem 4.3 and Theorem 4.4. It also shows that the density of distance-avoiding sets drops exponentially as more distances are avoided. In this section, we shall deduce from Theorem 4.1 a result of the same sort of Bukh's result which, though weaker, already implies both theorems from the beginning of the section; it is as follows:

Theorem 4.5. *Let $n \geq 2$ and denote by λ_n the absolute value of the global minimum of Ω_n. For every $0 < \varepsilon \leq 1$ there is a number $r = r(\varepsilon)$ such that if $d_1, \ldots, d_N$ are any distances satisfying*

$$d_2/d_1 > r, \ldots, d_N/d_{N-1} > r, \tag{4.39}$$

then

$$m_{d_1,\ldots,d_N}(\mathbb{R}^n) \leq \frac{\lambda_n^N + \varepsilon(N-1)}{1 + \lambda_n + \cdots + \lambda_n^N + \varepsilon(N-1)}.$$

Proof. Only two properties of Ω_n are used in the proof: (i) the fact that $\Omega_n(0) = 1$ and that it is continuous at zero (cf. (4.2)), and (ii) the fact that $\Omega_n(t) \to 0$ as $t \to \infty$ (cf. (4.22)). The latter property is crucial: It shows exactly where the argument would fail — as it should! — for $n = 1$, as $\Omega_1(t) = \cos t$ does not converge to zero as $t \to \infty$.

Let $0 < \varepsilon \leq 1$ be given. Since $\Omega_n(0) = 1$ and since it is continuous, there must be a number $t_0 > 0$ such that $\Omega_n(t) > 1 - \varepsilon$ for all $0 \leq t \leq t_0$. Likewise, since $\Omega_n(t) \to 0$ as $t \to \infty$, there must be a number $t_1 > t_0$ such that $|\Omega_n(t)| < \varepsilon$ for all $t \geq t_1$. We set $r(\varepsilon) = t_1/t_0$.

Now let distances $d_1, \ldots, d_N$ be given such that (4.39) is satisfied. We claim that, for $1 \leq j \leq N$,

$$\sum_{i=j}^{N} \lambda_n^{N-i} \Omega_n(td_i) \geq -\lambda_n^{N-j+1} - \varepsilon(N-j) \qquad \text{for all } t \geq 0. \tag{4.40}$$

Before proving the claim, we show how to apply it. By taking $j = 1$, we have

$$\sum_{i=1}^{N} \lambda_n^{N-i} \Omega_n(td_i) \geq -\lambda_n^{N} - \varepsilon(N-1) \qquad \text{for all } t \geq 0.$$

So, letting $S = 1 + \lambda_n + \cdots + \lambda_n^N + \varepsilon(N-1)$, we may set

$$z_0 = \frac{\lambda_n^N + \varepsilon(N-1)}{S} \qquad \text{and} \qquad z_i = \frac{\lambda_n^{N-i}}{S} \quad \text{for } i = 1, \ldots, N$$

and apply Theorem 4.1, finishing the proof.

We now prove (4.40) by induction. For $j = N$, the statement is obviously true. Now suppose the statement is true for some $1 < j \leq N$; we show it is also true for $j - 1$. To this end, consider any number $t \geq 0$. Suppose first $t \leq t_0/d_{j-1}$. Then, from the choice of t_0, $\Omega_n(td_{j-1}) > 1 - \varepsilon$. Using the induction hypothesis we then get

$$\begin{aligned}
\sum_{i=j-1}^{N} \lambda_n^{N-i} \Omega_n(td_i) &= \lambda_n^{N-j+1} \Omega_n(td_{j-1}) + \sum_{i=j}^{N} \lambda_n^{N-i} \Omega_n(td_i) \\
&\geq \lambda_n^{N-j+1}(1-\varepsilon) - \lambda_n^{N-j+1} - \varepsilon(N-j) \\
&\geq -\varepsilon(N-j+1),
\end{aligned}$$

where we use the fact that $\lambda_n \leq 1$, as follows from (4.2).

Now suppose $t \geq t_0/d_{j-1}$. Then, for $j \leq i \leq N$ we have $td_i \geq t_0 d_i/d_{j-1} \geq t_0 r = t_1$, hence $|\Omega_n(td_i)| < \varepsilon$. Since $\lambda_n \leq 1$, we then must have

$$\begin{aligned}
\sum_{i=j-1}^{N} \lambda_n^{N-i} \Omega_n(td_i) &= \lambda_n^{N-j+1} \Omega_n(td_{j-1}) + \sum_{i=j}^{N} \lambda_n^{N-i} \Omega_n(td_i) \\
&\geq -\lambda_n^{N-j+2} - \varepsilon(N-j+1),
\end{aligned}$$

and (4.40) follows. ∎

Recall from Section 4.3 that $\lambda_n = |\Omega_n(j_{n/2,1})|$, where $j_{n/2,1}$ is the first positive zero of $J_{n/2}$. From this, we may check numerically for $n = 2$ and 3 that $\lambda_n < 1$.

For $n \geq 4$, one may check that $\lambda_n < 1$ by combining (4.19) and (4.21) with (4.14). So the bound provided by Theorem 4.5 is actually exponential in N.

Taking $\varepsilon \to 0$, we have the following corollary:

Corollary 4.6. *For $n \geq 2$ we have*

$$\inf_{d_1,\ldots,d_N>0} m_{d_1,\ldots,d_N}(\mathbb{R}^n) \leq \frac{\lambda_n^N(1-\lambda_n)}{1-\lambda_n^{N+1}}.$$

Proof. We have $1+\lambda_n+\cdots+\lambda_n^N = (1-\lambda_n^{N+1})/(1-\lambda_n)$. Now apply Theorem 4.5 with $\varepsilon \to 0$. ■

The following corollary is a weakening of Theorem 4.5, which nonetheless already shows the exponential decay of $\inf\{\, m_{d_1,\ldots,d_N}(\mathbb{R}^n) \,:\, d_1,\ldots,d_N > 0\,\}$ in terms of N.

Corollary 4.7. *Let $n \geq 2$. For every $N \geq 1$ there is a number $r = r(N)$ such that if $d_1, \ldots, d_N$ are any distances satisfying*

$$d_2/d_1 > r, \ldots, d_N/d_{N-1} > r, \tag{4.41}$$

then $m_{d_1,\ldots,d_N}(\mathbb{R}^n) \leq 2^{-N}$.

Proof. We prove the result for $n = 2$, and hence by (4.38) it will follow for all $n > 2$ as well. We know (cf. Section 4.3) that the global minimum of Ω_2 occurs at $j_{1,1}$, the first positive zero of J_1, so we may check that $\lambda_2 \leq 1/2$, and hence the result holds for $N = 1$ in $\mathbb{R}^2$, as follows from Theorem 4.2.

So we may assume $N > 1$. Let then $\varepsilon = \lambda_2^{N+1}/((1-\lambda_2)(N-1))$ and apply Theorem 4.5 with $n = 2$. We then obtain a number r depending only on ε, and hence only on N, such that if $d_1, \ldots, d_N$ satisfy (4.41), then

$$\begin{aligned} m_{d_1,\ldots,d_N}(\mathbb{R}^n) &\leq \frac{\lambda_2^N + \varepsilon(N-1)}{1+\lambda_2+\cdots+\lambda_2^N+\varepsilon(N-1)} \\ &= \frac{\lambda_2^N + \lambda_2^{N+1}/(1-\lambda_2)}{(1-\lambda_2^{N+1})/(1-\lambda_2)+\lambda_2^{N+1}/(1-\lambda_2)} \\ &= \lambda_2^N(1-\lambda_2)+\lambda_2^{N+1} \\ &\leq 2^{-N}, \end{aligned}$$

where we use that $\lambda_2(1-\lambda_2) \leq 1/4$. ■

Though a weakening of Theorem 4.5, Corollary 4.7 already implies both Theorem 4.3 and Theorem 4.4. Indeed, to see that it implies Theorem 4.3, notice that if a set $C \subseteq \mathbb{R}^2$ avoids an unbounded sequence of distances, then for every $N \geq 1$ we may find in this sequence distances $d_1, \ldots, d_N$ satisfying (4.41), and hence the density of C is at most 2^{-N}. Since this is true for every N, set C cannot have positive density. The proof of Theorem 4.4 from Corollary 4.7 is similar.

Finally, we remark that Theorem 4.5 actually provides a bound exponential in both the dimension n and the number of distances N. This is so since also λ_n decreases exponentially with the dimension, as we showed in Section 4.4.

4.6 Relation to the theta number

Theorem 4.1 provides a bound for $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$ by means of an optimization problem. In this section, we aim at making the connection between this problem and the Lovász theta number clear. Along the way we discuss alternative methods to bound $m_{d_1,\ldots,d_N}(\mathbb{R}^n)$ and prove that they are all equivalent to the approach of Theorem 4.1.

4.6a. Two infinite graphs. We will consider here a slightly more general setting than before. Let D be a nonempty compact subset of the positive real line — in particular, $0 \notin D$. We let $m_D(\mathbb{R}^n)$ be the maximum upper density that a subset of $\mathbb{R}^n$ which avoids all distances in D can have.

In what follows we will consider two infinite graphs:

(i) the graph $G(\mathbb{R}^n, D)$ whose vertices are the points of $\mathbb{R}^n$ and in which two vertices $x, y \in \mathbb{R}^n$ are adjacent if and only if $\|x - y\| \in D$; and
(ii) the graph $G([-R, R]^n, D)$, $R > 0$, whose vertices are the points in the cube $[-R, R]^n$, with $x, y \in [-R, R]^n$ adjacent if and only if $\|x - y\| \in D$.

The stable sets of both graphs are sets which avoid all distances in D. Note that $m_D(\mathbb{R}^n)$ is a density analogue of the stability number for the graph $G(\mathbb{R}^n, D)$. The measurable stability number of $G([-R, R]^n, D)$ we define as

$$\alpha_{\mathrm{m}}(G([-R, R]^n, D)) = \sup\{\, \mathrm{vol}\, C : C \subseteq [-R, R]^n \text{ is measurable and stable}\,\}.$$

Notice that this is analogous to our approach for the sphere (cf. Section 3.1). Observe moreover that, since D is compact and $0 \notin D$, both parameters $m_D(\mathbb{R}^n)$ and $\alpha_{\mathrm{m}}(G([-R, R]^n, D))$ are positive.

4.6b. Three theta numbers. We now make use of kernels defined over the compact set $[-R, R]^n \times [-R, R]^n$. In Section 3.2 we presented material on kernels defined over the sphere. The concepts and results of that section all continue to hold when the sphere is replaced by $[-R, R]^n$.

Following our approach for the sphere (cf. Section 3.3), we define the theta number of $G([-R, R]^n, D)$, denoted by $\vartheta(G([-R, R]^n, D))$, as the optimal value of the following optimization problem:

$$\begin{aligned}
\sup\ & \textstyle\int_{[-R,R]^n} \int_{[-R,R]^n} A(x, y)\, dx dy \\
& \textstyle\int_{[-R,R]^n} A(x, x)\, dx = 1, \\
& A(x, y) = 0 \quad \text{if } \|x - y\| \in D, \\
& A\colon [-R, R]^n \times [-R, R]^n \to \mathbb{R} \text{ is a continuous and positive kernel.}
\end{aligned} \tag{4.42}$$

By repeating the arguments in the proof of Theorem 3.3 we then obtain the inequality

$$\vartheta(G([-R, R]^n, D)) \geq \alpha_{\mathrm{m}}(G([-R, R]^n, D)).$$

One may then observe that

$$\lim_{R\to\infty} \frac{\vartheta(G([-R, R]^n, D))}{\mathrm{vol}[-R, R]^n} \geq m_D(\mathbb{R}^n), \tag{4.43}$$

provided the limit exists — we will see later that this is the case; for now we just assume this to be true. This is our first approach to bound $m_D(\mathbb{R}^n)$.

The second approach we present is closer to that of Theorem 4.1 and works directly with $\mathbb{R}^n$. Let $f\colon \mathbb{R}^n \to \mathbb{C}$ be a continuous function. We say that f is of *positive type* if for every choice $x_1, \ldots, x_N$ of finitely many points in $\mathbb{R}^n$, the matrix $\big(f(x_i - x_j)\big)_{i,j=1}^N$ is positive semidefinite. If f is of positive type, then $f(0) \geq 0$ and $f(x) = \overline{f(-x)}$ for all $x \in \mathbb{R}^n$. Moreover, since for every $x \in \mathbb{R}^n$ the matrix

$$\begin{pmatrix} f(0) & f(x) \\ f(-x) & f(0) \end{pmatrix}$$

has to be positive semidefinite, we have $|f(x)| \leq f(0)$ for every $x \in \mathbb{R}^n$.

Recall the inner product $\langle \cdot, \cdot \rangle$ introduced in Section 4.2. Let $\vartheta_R(G(\mathbb{R}^n, D))$ be the optimal value of the following optimization problem:

$$\begin{aligned} \sup\ & \langle \varphi, \mathbf{1} \rangle \\ & \varphi(0) = 1, \\ & \varphi(x) = 0 \quad \text{if } \|x\| \in D, \\ & \varphi\colon \mathbb{R}^n \to \mathbb{R} \text{ is continous, periodic with periodicity lattice } R\mathbb{Z}^n, \\ & \qquad \text{and of positive type.} \end{aligned} \tag{4.44}$$

Here, $\mathbf{1}$ is the constant one function. Notice we require φ to be real-valued. One can actually prove, however, that the optimal value remains unchanged if φ is allowed to take complex values as well. We introduce the constraint because, as we will see, when we construct feasible solutions for this problem from distance-avoiding sets, we naturally get real-valued functions.

Let $C \subseteq \mathbb{R}^n$ be a measurable set which avoids all distances in D and which is such that χ^C, its characteristic function, has periodicity lattice $R\mathbb{Z}^n$ for some $R > 0$. Consider the function

$$\varphi(x) = \langle \chi^{C-x}, \chi^C \rangle, \tag{4.45}$$

the autocorrelation function of χ^C, also used in the proof of Theorem 4.1 (cf. Section 4.2). Notice φ is real-valued and periodic with periodicity lattice $R\mathbb{Z}^n$. It is also of positive type, since

$$\varphi(x-y) = \langle \chi^{C-(x-y)}, \chi^C \rangle = \langle \chi^{C-x}, \chi^{C-y} \rangle$$

for all $x, y \in \mathbb{R}^n$. We also have $\varphi(0) = \overline{\delta}(C)$ and $\varphi(x) = 0$ if $\|x\| \in D$.

Let us now compute the objective value of φ for (4.44). We have $\langle \varphi, \mathbf{1} \rangle = \widehat{\varphi}(0)$, the Fourier coefficient of φ computed at 0. Now recall that, by applying Parseval's formula, one may express the Fourier coefficients of φ in terms of those of χ^C, obtaining the expansion

$$\varphi(x) = \sum_{u \in \mathbb{R}^n} |\widehat{\chi}^C(u)|^2 e^{iu \cdot x} \tag{4.46}$$

for φ (cf. (4.7)), whence we see immediately that $\widehat{\varphi}(0) = |\widehat{\chi}^C(0)|^2 = \overline{\delta}(C)^2$.

So, if we show that φ is a continuous function, then $\overline{\delta}(C)^{-1}\varphi$ will be feasible for (4.44) and will have objective value $\overline{\delta}(C)$. Then it will follow that

$$\vartheta_R(G(\mathbb{R}^n, D)) \geq \overline{\delta}(C)$$

for every set $C \subseteq \mathbb{R}^n$ avoiding all the distances in D for which χ^C is periodic with periodicity lattice $R\mathbb{Z}^n$. Since we argued in the proof of Theorem 4.1 in Section 4.2 that the densities of such periodic distance-avoiding sets come arbitrarily close to $m_D(\mathbb{R}^n)$ as $R \to \infty$, we then could conclude that

$$\lim_{R\to\infty} \vartheta_R(G(\mathbb{R}^n, D)) \geq m_D(\mathbb{R}^n), \tag{4.47}$$

provided that the limit exists. In Section 4.6c we prove that the limit indeed exists; for now we simply assume this fact. So we obtain our second approach to bound $m_D(\mathbb{R}^n)$.

To show that φ is continuous we need a simple trick that will be useful again later, so we present it as a lemma for future reference.

Lemma 4.8. *For $f\colon \mathbb{R}^n \to \mathbb{C}$ and $x \in \mathbb{R}^n$, let $f_x\colon \mathbb{R}^n \to \mathbb{C}$ be the function such that $f_x(z) = f(z+x)$. If $f\colon \mathbb{R}^n \to \mathbb{C}$ is a square-integrable periodic function, then the function $F(x,y) = \langle f_x, f_y\rangle$ is continuous.*

Proof. If f is continuous, then we are done. If not, let Λ be a periodicity lattice of f and recall that continuous functions are dense in $L^2(\mathbb{R}^n/\Lambda)$ in the topology induced by the norm $\|g\| = \langle g, g\rangle^{1/2}$. So fix $\varepsilon > 0$ and let $g \in L^2(\mathbb{R}^n/\Lambda)$ be a continuous function such that $\|f - g\| < \varepsilon$. Then also $\|f_x - g_x\| < \varepsilon$. We assume without loss of generality that $\|f\| = 1$. Since g is continuous, if x and x' are close enough, we will have $\|g_x - g_{x'}\| < \varepsilon$, and then, using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |F(x,y) - F(x',y)| &= |\langle f_x - f_{x'}, f_y\rangle| \\ &\leq |\langle f_x - g_x, f_y\rangle| + |\langle g_x - g_{x'}, f_y\rangle| + |\langle g_{x'} - f_{x'}, f_y\rangle| \\ &\leq \|f_x - g_x\|\|f_y\| + \|g_x - g_{x'}\|\|f_y\| + \|g_{x'} - f_{x'}\|\|f_y\| \\ &< 3\varepsilon. \end{aligned}$$

Therefore, if (x,y) is close enough to (x',y'), then

$$|F(x,y) - F(x',y')| \leq |F(x,y) - F(x',y)| + |F(x',y) - F(x',y')| < 6\varepsilon,$$

proving that F is continuous. ■

From the lemma, it is obvious that the autocorrelation function φ is continuous, as we wanted.

The third and final approach we present for bounding $m_D(\mathbb{R}^n)$ is closely related to (4.44). In it, we just drop the constraint that φ be periodic. In other words, we let $\vartheta(G(\mathbb{R}^n, D))$ be the optimal value of the following optimization problem:

$$\begin{aligned} \sup\ & \langle \varphi, \mathbf{1}\rangle \\ & \varphi(0) = 1, \\ & \varphi(x) = 0 \quad \text{if } \|x\| \in D, \\ & \varphi\colon \mathbb{R}^n \to \mathbb{R} \text{ is continous and of positive type.} \end{aligned} \tag{4.48}$$

Once we see that the objective function is defined for every function $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ which is continuous and of positive type, it will be immediate that

$$\vartheta(G(\mathbb{R}^n, D)) \geq m_D(\mathbb{R}^n),$$

since any feasible solution of (4.44) for any $R > 0$ is also feasible for (4.48).

To see that the objective function of (4.48) is well-defined, we use the following theorem of Bochner (cf. Reed and Simon [41], Theorem IX.9):

Theorem 4.9 (Bochner's theorem). *A function $f\colon \mathbb{R}^n \to \mathbb{C}$ is continuous and of positive type if and only if*

$$f(x) = \int_{\mathbb{R}^n} e^{iu\cdot x}\, d\nu(u) \tag{4.49}$$

for some finite nonnegative Borel measure ν.

Notice that Bochner's theorem is valid for any continuous function of positive type, in particular for periodic functions. If $f\colon \mathbb{R}^n \to \mathbb{C}$ is a periodic, continuous function of positive type with periodicity lattice Λ, then the measure ν given by Bochner's theorem is supported at the points of the scaled dual lattice $2\pi\Lambda^*$, thus corresponding to the Fourier coefficients of f. In this case, the integral is just a (possibly infinite) sum. Compare this situation to the situation of the autocorrelation function φ of χ^C given in (4.45). It is periodic, continuous, and of positive type, and we had the expansion (4.46), from which we see that $\widehat{\varphi}(u) = |\chi^C(u)|^2$ for all $u \in 2\pi\Lambda^*$, so that all Fourier coefficients of φ are nonnegative. Notice also that the Fourier expansion of φ is just (4.49) with $\nu(u) = \widehat{\varphi}(u)$, where we write $\nu(u)$ instead of $\nu(\{u\})$ for short.

Now, the fact that the objective function of (4.48) is well-defined follows immediately from the following theorem.

Theorem 4.10. *Let $f\colon \mathbb{R}^n \to \mathbb{C}$ be a continuous function of positive type and let ν be the measure given by Bochner's theorem applied to f. Then $\langle f, \mathbf{1}\rangle = \nu(0)$.*

Proof. For $\delta > 0$, write

$$B_\delta = \bigcup_{i=1}^{n} \{\, u \in \mathbb{R}^n : 0 \neq |u_i| < \delta \,\}.$$

Now fix $\varepsilon > 0$. Then for every $\delta > 0$ there is an $L \geq 0$ such that, for all $R \geq L$,

$$\Big|\frac{1}{\operatorname{vol}[-R, R]^n} \int_{[-R,R]^n} e^{iu\cdot x}\, dx\Big| < \varepsilon \tag{4.50}$$

for all $u \notin B_\delta$, $u \neq 0$.

To see this, let $b_1, \ldots, b_n$ be the canonical orthonormal basis of $\mathbb{R}^n$ and take some $u \notin B_\delta$ such that $u \neq 0$. One possible periodicity lattice Λ of the function $e^{iu\cdot x}$ is generated by the vectors $b_1', \ldots, b_n'$ such that

$$b_i' = \begin{cases} b_i, & \text{if } u_i = 0;\\ (2\pi/u\cdot b_i)b_i, & \text{otherwise.}\end{cases}$$

Since $u \notin B_\delta$, the fundamental domain $D = \{\, \lambda_1 b'_1 + \cdots + \lambda_n b'_n : 0 \leq \lambda_i < 1 \,\}$ of this lattice is a parallelepiped whose sides have length at most $\max\{1, 2\pi\delta^{-1}\}$.

Let $d = \max\{1, 2\pi\delta^{-1}\}$. Take L to be big enough so that, for all $R \geq L$, $\operatorname{vol}([-R,R]^n \setminus [-R+d, R-d]^n)/\operatorname{vol}[-R,R]^n < \varepsilon$. Any cube $[-R,R]^n$ for any $R \geq L$ can be almost completely partitioned into translated copies of the fundamental domain of Λ, except for a region at the border, which will be completely contained in $[-R,R]^n \setminus [-R+d, R-d]^n$. Inside any copy of the fundamental domain D of Λ, the integral of $e^{iu\cdot x}$ equals zero, as for $u \neq 0$ we have

$$\int_D e^{iu\cdot x}\, dx = \langle e^{iu\cdot x}, \mathbf{1} \rangle = 0.$$

Then it follows from the choice of L that (4.50) holds for every $R \geq L$ and $u \notin B_\delta$, $u \neq 0$, as we wanted.

Now, since $\bigcap_{\delta > 0} B_\delta = \emptyset$ and since ν is a finite nonnegative Borel measure, there must be a $\delta > 0$ such that $\nu(B_\delta) < \varepsilon$. If we then let L be such that (4.50) holds for all $R \geq L$ and $u \notin B_\delta$ such that $u \neq 0$, then for all $R \geq L$ we have

$$\begin{aligned}
&\Big| \frac{1}{\operatorname{vol}[-R,R]^n} \int_{[-R,R]^n} f(x)\, dx - \nu(0) \Big| \\
&\quad = \Big| \frac{1}{\operatorname{vol}[-R,R]^n} \int_{[-R,R]^n} \int_{\mathbb{R}^n} e^{iu\cdot x}\, d\nu(u) dx - \nu(0) \Big| \\
&\quad = \Big| \int_{\mathbb{R}^n} \frac{1}{\operatorname{vol}[-R,R]^n} \int_{[-R,R]^n} e^{iu\cdot x}\, dx d\nu(u) - \nu(0) \Big| \\
&\quad < (\nu(\mathbb{R}^n) + 1)\varepsilon.
\end{aligned}$$

The last inequality follows from splitting $\mathbb{R}^n$, the integration domain, into B_δ, $\{0\}$, and $\mathbb{R}^n \setminus (B_\delta \cup \{0\})$. It is now obvious that $\langle f, \mathbf{1} \rangle = \nu(0)$, as we wanted. ∎

Before proceeding, we observe that the resemblance between (4.44) or (4.48) and formulation (2.1) for the theta number of a finite graph is particularly clear when one sees the connection between functions of positive type and positive translation invariant kernels over $\mathbb{R}^n \times \mathbb{R}^n$.

Such a kernel is just a function $A\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$. It is *translation invariant* if $A(x+z, y+z) = A(x,y)$ for all x, y, $z \in \mathbb{R}^n$. If A is continuous, we say it is *positive* if for every choice $x_1, \ldots, x_N$ of finitely many points in $\mathbb{R}^n$, the matrix $\big(A(x_i, x_j)\big)_{i,j=1}^N$ is positive semidefinite. Notice then that if $f\colon \mathbb{R}^n \to \mathbb{C}$ is a continuous function of positive type, then $A\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ such that $A(x,y) = f(x-y)$ is a continuous, positive, translation invariant kernel. Conversely, to a continuous, positive, and translation invariant kernel corresponds a continuous function of positive type.

Notice also that, if $f\colon \mathbb{R}^n \to \mathbb{C}$ is any continuous function of positive type and $A\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is such that $A(x,y) = f(x-y)$, then

$$\lim_{R\to\infty} \frac{1}{(\operatorname{vol}[-R,R]^n)^2} \int_{[-R,R]^n} \int_{[-R,R]^n} A(x,y)\, dx dy = \langle f, \mathbf{1} \rangle. \tag{4.51}$$

One easy way to see this is to use Bochner's theorem and an argument similar to the one used above to show that $\langle f, \mathbf{1} \rangle = \nu(0)$.

So one can state (4.44) or (4.48) in terms of kernels. It is clear how to do so: The constraints can be directly translated and, because of (4.51), the objective function becomes, for a kernel $A\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

$$\lim_{R\to\infty} \frac{1}{(\operatorname{vol}[-R,R]^n)^2} \int_{[-R,R]^n} \int_{[-R,R]^n} A(x,y)\,dxdy.$$

4.6c. The existence of the limits and equivalence of the bounds. We now prove the following theorem:

Theorem 4.11. *The limits in* (4.43) *and* (4.47) *exist and*

$$\lim_{R\to\infty} \frac{\vartheta(G([-R,R]^n, D))}{\operatorname{vol}[-R,R]^n} = \lim_{R\to\infty} \vartheta_R(G(\mathbb{R}^n, D)) = \vartheta(G(\mathbb{R}^n, D)). \tag{4.52}$$

Proof. We first show that

$$\lim_{R\to\infty} \vartheta_R(G(\mathbb{R}^n, D)) = \vartheta(G(\mathbb{R}^n, D)). \tag{4.53}$$

Our strategy is to, given a feasible solution φ of (4.48), construct for each $R > 4d$ a feasible solution φ_R of (4.44) with periodicity lattice $R\mathbb{Z}^n$, in such a way that

$$\lim_{R\to\infty} \langle \varphi_R, \mathbf{1} \rangle = \langle \varphi, \mathbf{1} \rangle.$$

If we can do that, then it will follow that the limit in (4.53) exists and also that it is equal to $\vartheta(G(\mathbb{R}^n, D))$. This is so since, for all $R > 0$, we have $\vartheta_R(G(\mathbb{R}^n, D)) \leq \vartheta(G(\mathbb{R}^n, D))$, as any feasible solution of (4.44) for any $R > 0$ is also a feasible solution of (4.48), and then it is clear that by taking R large enough we can make $\vartheta_R(G(\mathbb{R}^n, D))$ as close to $\vartheta(G(\mathbb{R}^n, D))$ as we like.

For our analysis we fix some $L > d$; later we will take $L = R/2 - d$ for any given $R > 4d$. Let $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ be any feasible solution of (4.48) and consider the kernel $A_L\colon [-L,L]^n \times [-L,L]^n \to \mathbb{R}$ such that

$$A_L(x,y) = \varphi(x-y).$$

This is a continuous and positive kernel, so from Mercer's theorem (Theorem 3.2, but for kernels over $[-L,L]^n \times [-L,L]^n$ instead of $S^{n-1} \times S^{n-1}$) we know that there are continuous functions $f_i\colon [-L,L]^n \to \mathbb{R}$ such that

$$A_L(x,y) = \sum_{i=1}^{\infty} f_i(x) f_i(y) \tag{4.54}$$

with absolute and uniform convergence over $[-L,L]^n \times [-L,L]^n$.

For each $i = 1, 2, \ldots$, consider the function $g_i\colon \mathbb{R}^n \to \mathbb{R}$ which, in translated copies of the cube $[-L,L]^n$ centered on points of the lattice $2(L+d)\mathbb{Z}^n$, is a copy

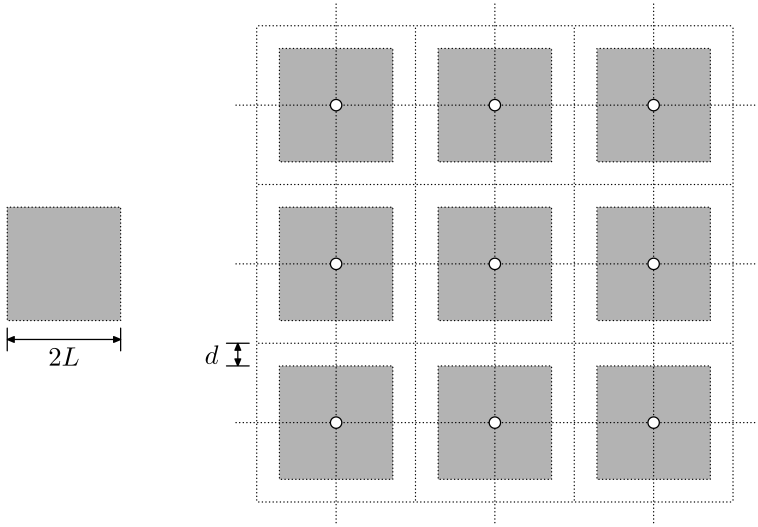

Figure 4.4. The figure illustrates the process of obtaining a periodic function $g\colon \mathbb{R}^n \to \mathbb{R}$ from a function $f\colon [-L, L]^n \to \mathbb{R}$ in the case $n = 2$. Function $f\colon [-L, L]^2 \to \mathbb{R}$ is defined on the gray square of side $2L$ shown on the left. Function g is defined over $\mathbb{R}^2$; we show a part of it on the right. The white dots $\circ$ mark the points of the lattice $2(L + d)\mathbb{Z}^2$. Notice that, inside each gray square, which is a translated copy of the domain of f, function g is just a copy of f.

of f_i, being zero everywhere else — Figure 4.4 illustrates this construction for $n = 2$. Notice that each function g_i is periodic with periodicity lattice $2(L + d)\mathbb{Z}^n$.

Consider then the kernel $B_L\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$B_L(x, y) = \sum_{i=1}^{\infty} g_i(x) g_i(y). \tag{4.55}$$

Since we have absolute and uniform convergence in (4.54) and since each g_i is periodic, from B_L we may construct a translation invariant kernel $\overline{B}_L\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by setting

$$\overline{B}_L(x, y) = \lim_{T\to\infty} \frac{1}{\operatorname{vol}[-T, T]^n} \int_{[-T,T]^n} B_L(x + z, y + z)\, dz. \tag{4.56}$$

Since $\overline{B}_L$ is translation invariant, there is a function $\psi_L\colon \mathbb{R}^n \to \mathbb{R}$ that is such that $\overline{B}_L(x, y) = \psi_L(x - y)$ for all $x, y \in \mathbb{R}^n$. We claim that the function

$$\varrho_L = \frac{\operatorname{vol}[-L - d, L + d]^n}{\operatorname{vol}[-L, L]^n} \psi_L$$

is a feasible solution of (4.44) with periodicity lattice $2(L + d)\mathbb{Z}^n$.

We now proceed to prove this claim. Since each g_i is periodic with periodicity lattice $2(L + d)\mathbb{Z}^n$, it is easy to check that also ϱ_L is periodic with periodicity lattice $2(L + d)\mathbb{Z}^n$. Moreover, from the construction of the g_i and (4.56), we see that $\overline{B}_L(x, y) = 0$ whenever $\|x - y\| \in D$, and so $\varrho_L(x) = 0$ whenever $\|x\| \in D$.

To prove our claim that ϱ_L is feasible for (4.44), we still have to check that it is continuous and of positive type and that $\varrho_L(0) = 1$.

We first show that ϱ_L is continuous. To this end, notice that, since we have absolute and uniform convergence in (4.54), from the construction of the g_i we see that we also have absolute and uniform convergence in (4.55), and so we may rewrite (4.56) as

$$\overline{B}_L(x, y) = \sum_{i=1}^{\infty} \langle (g_i)_x, (g_i)_y \rangle, \tag{4.57}$$

where $(g_i)_x(z) = g_i(x + z)$ for all x, $z \in \mathbb{R}^n$.

To prove that ϱ_L is continuous, we first argue that the series in (4.57) converges absolutely and uniformly over $\mathbb{R}^n \times \mathbb{R}^n$. Indeed, since $\varphi(0) = 1$, we have from (4.54) that

$$\operatorname{vol}[-L, L]^n = \int_{[-L,L]^n} A_L(x, x)\, dx = \sum_{i=1}^{\infty} (f_i, f_i), \tag{4.58}$$

where for two square-integrable functions f, g: $[-L, L]^n \to \mathbb{R}$ we write

$$(f, g) = \int_{[-L,L]^n} f(x) g(x)\, dx.$$

So the series of nonnegative terms on the right-hand side of (4.58) converges. Now notice that, for $i = 1, 2, \ldots,$

$$\langle g_i, g_i \rangle = \frac{(f_i, f_i)}{\operatorname{vol}[-L - d, L + d]^n}, \tag{4.59}$$

so also the series of nonnegative terms

$$\sum_{i=1}^{\infty} \langle g_i, g_i \rangle$$

converges. Next, observe that, if x, $y \in \mathbb{R}^n$, then $|\langle (g_i)_x, (g_i)_y \rangle| \le \langle g_i, g_i \rangle$. From this we see that the series in (4.57) converges absolutely for every x, $y \in \mathbb{R}^n$. It also converges uniformly for every x, $y \in \mathbb{R}^n$ since, for $m \ge 1$,

$$\sum_{i=m}^{\infty} |\langle (g_i)_x, (g_i)_y \rangle| \le \sum_{i=m}^{\infty} \langle g_i, g_i \rangle.$$

Now, we know from Lemma 4.8 that, for each $i = 1, 2, \ldots,$ the function $(x, y) \mapsto \langle (g_i)_x, (g_i)_y \rangle$ is continuous. Then, since convergence in (4.57) is uniform, $\overline{B}_L$ is continuous, and hence also ϱ_L is continuous, as we wanted. Now that we know that ϱ_L is continuous, the fact that it is of positive type follows immediately from expression (4.57) for $\overline{B}_L$.

Finally, we show that $\varrho_L(0) = 1$. From (4.57), using (4.58) and (4.59), we see that

$$\varrho_L(0) = \frac{\operatorname{vol}[-L-d, L+d]^n}{\operatorname{vol}[-L, L]^n} \sum_{i=1}^{\infty} \langle g_i, g_i \rangle = \frac{1}{\operatorname{vol}[-L, L]^n} \sum_{i=1}^{\infty} (f_i, f_i) = 1,$$

as we wanted.

So we just proved that ϱ_L is feasible for (4.44) with periodicity lattice $2(L+d)\mathbb{Z}^n$. Now we estimate its objective value, to show that it approaches the objective value of φ as $L \to \infty$.

From (4.51) we see that

$$\lim_{L\to\infty} \frac{1}{(\operatorname{vol}[-L, L]^n)^2} \int_{[-L,L]^n} \int_{[-L,L]^n} A_L(x, y)\, dxdy = \langle \varphi, \mathbf{1} \rangle. \tag{4.60}$$

Now note that

$$\int_{[-L,L]^n} \int_{[-L,L]^n} A_L(x, y)\, dxdy = \sum_{i=1}^{\infty} (f_i, \mathbf{1})^2. \tag{4.61}$$

By the construction of the g_i we have

$$\langle g_i, \mathbf{1} \rangle = \frac{(f_i, \mathbf{1})}{\operatorname{vol}[-L-d, L+d]^n},$$

and so using (4.51) again we obtain for ψ_L that

$$\begin{aligned} \langle \psi_L, \mathbf{1} \rangle &= \lim_{T\to\infty} \frac{1}{(\operatorname{vol}[-T, T]^n)^2} \int_{[-T,T]^n} \int_{[-T,T]^n} \overline{B}_L(x, y)\, dxdy \\ &= \sum_{i=1}^{\infty} \langle g_i, \mathbf{1} \rangle^2 \\ &= \frac{1}{(\operatorname{vol}[-L-d, L+d]^n)^2} \sum_{i=1}^{\infty} (f_i, \mathbf{1})^2. \end{aligned}$$

Combining this with (4.60), (4.61) and the definition of ϱ_L, we see at once that

$$\lim_{L\to\infty} \langle \varrho_L, \mathbf{1} \rangle = \langle \varphi, \mathbf{1} \rangle. \tag{4.62}$$

For each $L > d$ we have built a feasible solution ϱ_L of (4.44) with periodicity lattice $2(L+d)\mathbb{Z}^n$ such that (4.62) holds. Given $R > 4d$, we set $L = R/2 - d > d$ and $\varphi_R = \varrho_L$. Then we have for each $R > 4d$ a feasible solution of (4.44) with periodicity lattice $R\mathbb{Z}^n$ and the limit of their objective values is the objective value of φ, as we wanted in the first place.

So we just established (4.53). In an analogous way, one may prove that

$$\lim_{R\to\infty} \frac{\vartheta(G([-R, R]^n, D))}{\operatorname{vol}[-R, R]^n} = \vartheta(G(\mathbb{R}^n, D)),$$

establishing the existence of the limit above and finally also (4.52). ∎

4.6d. Spherical averages and a reformulation. It is not immediately clear how to solve problem (4.48). To simplify it, the trick is again to exploit symmetry, in this case by noticing that in (4.48) we may restrict ourselves to radial functions φ.

Recall that a function $f\colon \mathbb{R}^n \to \mathbb{C}$ is radial if its values depend only on the norm of its argument, that is, if $f(x) = f(y)$ whenever $\|x\| = \|y\|$. To see that we may restrict ourselves to radial functions, let φ be a feasible solution of (4.48). Consider then the function $\overline{\varphi}\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\overline{\varphi}(x) = \frac{1}{\omega_n} \int_{S^{n-1}} \varphi(\|x\|\xi)\, d\omega(\xi),$$

which is radial. We claim that it is feasible for (4.48) and also that $\langle \overline{\varphi}, \mathbf{1} \rangle = \langle \varphi, \mathbf{1} \rangle$.

It is clear that $\overline{\varphi}(0) = 1$ and that $\overline{\varphi}(x) = 0$ whenever $\|x\| \in D$. It is also clear that $\overline{\varphi}$ is continuous. So we show that it is also of positive type. To see this, recall from (4.29) that, if $\mathrm{O}(\mathbb{R}^n)$ is the orthogonal group and μ is the Haar measure over it normalized so that $\mu(\mathrm{O}(\mathbb{R}^n)) = 1$, then

$$\frac{1}{\omega_n} \int_{S^{n-1}} \varphi(\|x\|\xi)\, d\omega(\xi) = \int_{\mathrm{O}(\mathbb{R}^n)} \varphi(Tx)\, d\mu(T). \tag{4.63}$$

If $x_1, \ldots, x_N$ is any choice of finitely many vectors in $\mathbb{R}^n$ and if $p_1, \ldots, p_N$ are any real numbers, then since φ is of positive type,

$$\sum_{i,j=1}^{N} \overline{\varphi}(x_i - x_j) p_i p_j = \int_{\mathrm{O}(\mathbb{R}^n)} \sum_{i,j=1}^{N} \varphi(Tx_i - Tx_j) p_i p_j \, d\mu(T) \geq 0,$$

and we see that $\overline{\varphi}$ is also of positive type.

Finally, using (4.63) again, we get

$$\langle \overline{\varphi}, \mathbf{1} \rangle = \int_{\mathrm{O}(\mathbb{R}^n)} \langle T \cdot \varphi, \mathbf{1} \rangle \, d\mu(T),$$

where $(T \cdot \varphi)(x) = \varphi(Tx)$ for all $x \in \mathbb{R}^n$. So we see at once that $\langle \overline{\varphi}, \mathbf{1} \rangle = \langle \varphi, \mathbf{1} \rangle$.

So we have proven our claim that we may restrict ourselves in (4.48) to radial functions. Then we may use the following theorem of Schoenberg [43]:

Theorem 4.12. *A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is continuous, radial, and of positive type, if and only if*

$$f(x) = \int_0^\infty \Omega_n(t\|x\|)\, d\alpha(t) \tag{4.64}$$

for every $x \in \mathbb{R}^n$, where α is a finite, nonnegative Borel measure, and Ω_n is given by (4.2).

Proof. The proof is almost immediate from Bochner's theorem and (4.2). From Bochner's theorem, f is continuous and of positive type if and only if there is a measure ν so that (4.49) holds. But then, f is radial, continuous, and of positive type if and only if, for every $x \in \mathbb{R}^n$,

$$\begin{aligned} f(x) &= \frac{1}{\omega_n} \int_{S^{n-1}} f(\|x\|\xi)\, d\omega(\xi) \\ &= \frac{1}{\omega_n} \int_{S^{n-1}} \int_{\mathbb{R}^n} e^{iu\cdot(\|x\|\xi)}\, d\nu(u) d\omega(\xi) \\ &= \int_{\mathbb{R}^n} \frac{1}{\omega_n} \int_{S^{n-1}} e^{i(u\|x\|)\cdot\xi}\, d\omega(\xi) d\nu(u) \\ &= \int_{\mathbb{R}^n} \Omega_n(\|u\|\|x\|)\, d\nu(u), \end{aligned}$$

where in the last identity we used (4.2).

Now, consider the measure α over $[0, \infty)$ which for a Borel set $X \subseteq [0, \infty)$ is such that

$$\alpha(X) = \nu(\{\, u \in \mathbb{R}^n : \|u\| \in X \,\}).$$

Then we have

$$f(x) = \int_0^\infty \Omega_n(t\|x\|)\, d\alpha(t),$$

as we wanted. ■

Notice that we already used Theorem 4.12 implicitly in the proof of Theorem 4.1. Compare, for instance, expression (4.64) with (4.8).

Recall that, if $f \colon \mathbb{R}^n \to \mathbb{C}$ is continuous and of positive type and is given as (4.49), then $\langle f, \mathbf{1} \rangle = \nu(0)$, as follows from Bochner's theorem. Then, by using Theorem 4.12, we may rewrite (4.48) equivalently as

$$\begin{aligned} \sup\ & \alpha(0) \\ & \textstyle\int_0^\infty 1\, d\alpha(t) = 1, \\ & \textstyle\int_0^\infty \Omega_n(td)\, d\alpha(t) = 0 \quad \text{if } d \in D, \\ & \alpha \text{ is a finite, nonnegative Borel measure.} \end{aligned} \tag{4.65}$$

Here we use the fact that $\Omega_n(0) = 1$, as follows from (4.2).

Notice the similarity between the above problem and (4.9). The only change is that the sums are now integrals and α is now a measure. We can also write a dual for (4.65) such that weak duality holds. If we do so for the case in which we have $D = \{d_1, \ldots, d_N\}$, then we get the dual problem

$$\begin{aligned} \inf\ & z_0 \\ & z_0 + z_1 + \cdots + z_N \geq 1, \\ & z_0 + z_1\Omega_n(td_1) + \cdots + z_N\Omega_n(td_N) \geq 0 \quad \text{for all } t > 0, \end{aligned} \tag{4.66}$$

which is exactly the same problem we considered in the proof of Theorem 4.1. We may also prove in this case that strong duality holds between (4.65) and (4.66). From this, we immediately see that $\vartheta(G(\mathbb{R}^n, D))$ gives exactly the same bound as Theorem 4.1.

4.6e. The complementary graph. In Section 3.7, we showed how to define a theta number for the complements of distance graphs on the sphere, in such a way that the relations that one verifies between the theta number of a finite graph and that of its complement also carry over to the infinite graphs on the sphere. In this section we aim at doing the same for the graphs $G(\mathbb{R}^n, D)$, where D is again a compact subset of the positive real line.

The complement of $G(\mathbb{R}^n, D)$, denoted by $\overline{G}(\mathbb{R}^n, D)$, is the graph whose vertex set is $\mathbb{R}^n$ and in which $x, y \in \mathbb{R}^n$, $x \neq y$, are adjacent if and only if $\|x - y\| \notin D$. Since D is compact and $0 \notin D$, the graph $\overline{G}(\mathbb{R}^n, D)$ has finite stability number.

How should we define the theta number for $\overline{G}(\mathbb{R}^n, D)$? Certainly, formulation (4.48) no longer works: A continuous function φ such that $\varphi(0) = 1$ must also be positive in a neighborhood of the origin, but to be feasible in (4.48) we would require $\varphi(x) = 0$ whenever $x \neq 0$ and $\|x\| \notin D$, and since D is compact and does not contain the origin, we would get a contradiction. So (4.48) is infeasible for $\overline{G}(\mathbb{R}^n, D)$.

A formulation based on formulation (2.3) for the theta number of a finite graph provides however a working alternative. Let $\vartheta(\overline{G}(\mathbb{R}^n, D))$ be the optimal value of the optimization problem

$$\begin{aligned} \inf\ & 1 + \psi(0) \\ & \psi(x) = -1 \quad \text{whenever } \|x\| \in D, \\ & \psi\colon \mathbb{R}^n \to \mathbb{R} \text{ is a continuous function of positive type.} \end{aligned} \tag{4.67}$$

It is not immediately clear that (4.67) is even feasible; this we will prove in a moment. Before that, we observe that any feasible solution of (4.67) provides an upper bound to the stability number of $\overline{G}(\mathbb{R}^n, D)$. Indeed, if $C \subseteq \mathbb{R}^n$ is a nonempty stable set of $\overline{G}(\mathbb{R}^n, D)$ — which is, as observed above, finite — and ψ is a feasible solution of (4.67), then

$$0 \leq \sum_{x,y \in C} \psi(x - y) = |C|\psi(0) - |C|(|C| - 1),$$

and hence $|C| \leq 1 + \psi(0)$, as we wanted.

Recall from Theorem 2.2 that, if $G = (V, E)$ is a finite vertex-transitive graph, then

$$\vartheta(G)\vartheta(\overline{G}) = |V|.$$

We now prove the analogous identity for graphs of the form $G(\mathbb{R}^n, D)$; note these graphs are also vertex-transitive. More specifically, we prove

$$\vartheta(G(\mathbb{R}^n, D))\vartheta(\overline{G}(\mathbb{R}^n, D)) = 1. \tag{4.68}$$

We begin by showing the inequality "$\leq$" in (4.68), and along the way we also show that (4.67) is feasible. Let φ be a feasible solution of (4.48). We claim that the function

$$\psi(x) = \langle \varphi, \mathbf{1} \rangle^{-1} \varphi(x) - 1$$

is a feasible solution of (4.67).

It is clear that ψ is continuous and that $\psi(x) = -1$ whenever $\|x\| \in D$. Now, to see that ψ is of positive type, let ν be the measure given by Bochner's theorem applied to φ. From Theorem 4.10 we know that

$$\langle \varphi, \mathbf{1} \rangle = \nu(0).$$

Let δ be the pure point measure over $\mathbb{R}^n$ which is 1 at the origin. Then the measure

$$\tilde{\nu} := \langle \varphi, \mathbf{1} \rangle^{-1} \nu - \delta$$

is a nonnegative Borel measure. Moreover, since ν was given by Bochner's theorem applied to φ, we have

$$\psi(x) = \int_{\mathbb{R}^n} e^{iu \cdot x} \, d\tilde{\nu}(u),$$

and so we see that ψ is also of positive type. With this, notice that we just proved that (4.67) is feasible, since (4.48) is feasible.

Finally, observe that we have

$$\langle \varphi, \mathbf{1} \rangle \vartheta(\overline{G}(\mathbb{R}^n, D)) \leq \langle \varphi, \mathbf{1} \rangle (1 + \psi(0)) = 1,$$

and since φ is any feasible solution of (4.48), it follows that

$$\vartheta(G(\mathbb{R}^n, D)) \vartheta(\overline{G}(\mathbb{R}^n, D)) \leq 1.$$

To see the reverse inequality, let ψ be a feasible solution of (4.67). Then the function

$$\varphi(x) = \frac{1 + \psi(x)}{1 + \psi(0)}$$

is feasible for (4.48). Indeed, it is a continuous function of positive type that satisfies $\varphi(0) = 1$ and $\varphi(x) = 0$ if $\|x\| \in D$. Now, we know that

$$\langle \varphi, \mathbf{1} \rangle = \frac{1 + \langle \psi, \mathbf{1} \rangle}{1 + \psi(0)} \geq \frac{1}{1 + \psi(0)},$$

whence

$$\vartheta(G(\mathbb{R}^n, D))(1 + \psi(0)) \geq \langle \varphi, \mathbf{1} \rangle (1 + \psi(0)) \geq 1$$

and, since this holds for any feasible solution ψ of (4.67), we have the inequality

$$\vartheta(G(\mathbb{R}^n, D)) \vartheta(\overline{G}(\mathbb{R}^n, D)) \geq 1,$$

finishing the proof of (4.68).

4.6f. Bounds from finite subgraphs. To finish, we investigate the idea of using the theta number of finite subgraphs of $G(\mathbb{R}^n, D)$ to bound $m_D(\mathbb{R}^n)$ or $\chi_{\mathrm{m}}(G(\mathbb{R}^n, D))$. Analogous results were presented for distance graphs on the sphere in Section 3.7a.

Let $H = (V, E)$ be a finite subgraph of $G(\mathbb{R}^n, D)$. Clearly, $\chi_{\mathrm{m}}(G(\mathbb{R}^n, D)) \geq \chi(H)$. Since, from Theorem 2.1, $\vartheta(\overline{H}) \leq \chi(H)$, we have

$$\vartheta(\overline{H}) \leq \chi_{\mathrm{m}}(G(\mathbb{R}^n, D)).$$

So computing the theta number of complements of finite subgraphs of $G(\mathbb{R}^n, D)$ is a strategy to find lower bounds for $\chi_{\mathrm{m}}(G(\mathbb{R}^n, D))$.

The bounds computed with such an approach are, however, not better than the bound obtained from computing $\vartheta(G(\mathbb{R}^n, D))$. Indeed, let ψ be a feasible solution of (4.67). Then the matrix $Z\colon V \times V \to \mathbb{R}$ such that

$$Z(x, y) = \psi(x - y)$$

is a feasible solution of formulation (2.3) for $\vartheta(\overline{H})$, and has objective value equal to that of ψ. Together with (4.68), this implies that

$$\vartheta(\overline{H}) \leq \vartheta(\overline{G}(\mathbb{R}^n, D)) = \frac{1}{\vartheta(G(\mathbb{R}^n, D))}, \qquad (4.69)$$

so by computing the theta number of finite subgraphs of $G(\mathbb{R}^n, D)$ one cannot do better than by computing $\vartheta(G(\mathbb{R}^n, D))$ in order to bound $\chi_{\mathrm{m}}(G(\mathbb{R}^n, D))$.

Bounds on the stability numbers of finite subgraphs of $G(\mathbb{R}^n, D)$ can also be used to bound $m_D(\mathbb{R}^n)$. Indeed, if $H = (V, E)$ is a finite subgraph of $G(\mathbb{R}^n, D)$, one has

$$m_D(\mathbb{R}^n) \leq \frac{\alpha(H)}{|V|}.$$

To see this, let C be a measurable stable set of $G(\mathbb{R}^n, D)$. Then $|C \cap (V + y)| \leq \alpha(H)$ for every $y \in \mathbb{R}^n$. But then

$$\begin{aligned} \overline{\delta}(C)|V| &= \sum_{x \in V} \limsup_{R \to \infty} \frac{1}{\mathrm{vol}[-R, R]^n} \int_{[-R,R]^n} |C \cap \{x + y\}|\, dy \\ &= \limsup_{R \to \infty} \frac{1}{\mathrm{vol}[-R, R]^n} \int_{[-R,R]^n} |C \cap (V + y)|\, dy \\ &\leq \alpha(H), \end{aligned}$$

as we wanted (this was observed by Larman and Rogers [27]). Since $\vartheta(H) \geq \alpha(H)$, we then have

$$\frac{\vartheta(H)}{|V|} \geq \frac{\alpha(H)}{|V|} \geq m_D(\mathbb{R}^n).$$

Recall from Theorem 2.2 that for the finite graph H we have that

$$\vartheta(H)\vartheta(\overline{H}) \geq |V|.$$

So, using (4.69) we get

$$\frac{\vartheta(H)}{|V|} \geq \frac{1}{\vartheta(\overline{H})} \geq \vartheta(G(\mathbb{R}^n, D)),$$

and we see that one cannot find better bounds for $m_D(\mathbb{R}^n)$ by computing the theta numbers of finite subgraphs of $G(\mathbb{R}^n, D)$ than by computing $\vartheta(G(\mathbb{R}^n, D))$.

Bibliography

[1] G.E. Andrews, R. Askey, and R. Roy, *Special Functions*, Encyclopedia of Mathematics and its Applications 71, Cambridge University Press, Cambridge, 1999.

[2] S. Axler, P. Bourdon, and W. Ramey, *Harmonic Function Theory*, Second Edition, Graduate Texts in Mathematics 137, Springer-Verlag, New York, 2001.

[3] C. Bachoc, G. Nebe, F.M. de Oliveira Filho, and F. Vallentin, Lower bounds for measurable chromatic numbers, `arXiv:0801.1059v2 [math.CO]`, to appear in *Geometric and Functional Analysis*, 2008, 18pp.

[4] F.W. Bessel, Untersuchung des Theils der planetarischen Störungen, welcher aus der Bewegung von der Sonne entsteht, *Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften*, *Mathemat. Klasse* (1824), 52pp.

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

[6] J. Bourgain, A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$, *Israel Journal of Mathematics* 54 (1986) 307–316.

[7] N.G. de Bruijn and P. Erdős, A colour problem for infinite graphs and a problem in the theory of relations, *Indagationes Mathematicae* 13 (1951) 371–373.

[8] B. Bukh, Measurable sets with excluded distances, *Geometric and Functional Analysis* 18 (2008) 668–697.

[9] D. Coulson, A 15-colouring of 3-space omitting distance one, *Discrete Mathematics* 256 (2002) 83–90.

[10] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Research Reports Supplements 1973 No. 10, Philips Research Laboratories, Eindhoven, 1973.

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

[12] P. Erdős, Some combinatorial, geometric and set theoretic problems in measure theory, in: *Measure Theory Oberwolfach 1983* (Proceedings of the Conference held at Oberwolfach, June 26–July 2, 1983; D. Kölzow and D. Maharam-Stone, eds.), Lecture Notes in Mathematics 1089, Springer-Verlag, Berlin, 1984, pp. 321–327.

[13] K.J. Falconer, The realization of distances in measurable subsets covering $\mathbb{R}^n$, *Journal of Combinatorial Theory, Series A* 31 (1981) 184–189.

[14] K.J. Falconer, The realization of small distances in plane sets of positive measure, *Bulletin of the London Mathematical Society* 18 (1986) 475–477.

[15] K.J. Falconer and J.M. Marstrand, Plane sets with positive density at infinity contain all large distances, *Bulletin of the London Mathematical Society* 18 (1986) 471–474.

[16] P. Frankl and R.M. Wilson, Intersection theorems with geometric consequences, *Combinatorica* 1 (1981) 357–368.

[17] H. Furstenberg, Y. Katznelson, and B. Weiss, Ergodic theory and configurations in sets of positive density, in: *Mathematics of Ramsey Theory* (J. Nešetřil and V. Rödl, eds.), Springer-Verlag, Berlin, 1990, pp. 184–198.

[18] B. Gerard and L. Roberts, Graphical discovery of a new identity for Jacobi polynomials, *American Mathematical Monthly* 105 (1998) 163–166.

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

[20] N. Gvozdenović, *Approximating the Stability Number and the Chromatic number of a Graph via Semidefinite Programming*, PhD thesis, University of Amsterdam, 2008.

[21] H. Hadwiger, Ungelöste probleme Nr. 40, *Elemente der Mathematik* 16 (1961) 103–104.

[22] P.R. Halmos, *Measure Theory*, D. Van Nostrand Company, Inc., New York, 1950.

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

[24] R.M. Karp, Reducibility among combinatorial problems, in: *Complexity of Computer Computations* (Proceedings of a symposium on the Complexity of Computer Computations, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1972; R.E. Miller, J.W. Thatcher, eds.), Plenum Press, New York, 1972, pp. 85–103.

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

[26] D.E. Knuth, The sandwich theorem, *Electronic Journal of Combinatorics* 1 (1994) 48pp.

[27] D.G. Larman and C.A. Rogers, The realization of distances within sets in Euclidean space, *Mathematika* 19 (1972) 1–24.

[28] L. Lovász, Discrete and continuous: two sides of the same?, *Geometric and Functional Analysis*, Special Volume, Part I (2000) 359–382.

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

[30] L. Lovász, Self-dual polytopes and the chromatic number of distance graphs on the sphere, *Acta Scientiarum Mathematicarum* 45 (1983) 317–323.

[31] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Mathematical Library 16, Twelfth impression, North-Holland, Amsterdam, 2006.

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

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

[34] W. Moser, *Research Problems in Discrete Geometry*, McGill University, Montreal, 1981.

[35] O. Nechushtan, On the space chromatic number, *Discrete Mathematics* 256 (2002) 499–507.

[36] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics 13, Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[37] F.M. de Oliveira Filho and F. Vallentin, Fourier analysis, linear programming, and densities of distance avoiding sets in $\mathbb{R}^n$, `arXiv:0808.1822v2 [math.CO]`, to appear in *Journal of the European Mathematical Society*, 2008, 11pp.

[38] A. Quas, Distances in positive density sets in $\mathbb{R}^d$, *Journal of Combinatorial Theory, Series A* 116 (2009) 979–987.

[39] A.M. Raigorodskii, On the chromatic number of a space, *Russian Mathematical Surveys* 55 (2000) 351–352.

[40] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, New York, 1972.

[41] M. Reed and B. Simon, *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-adjointness*, Academic Press, New York, 1975.

[42] 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.

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

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

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

[46] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency, Volume B*, Algorithms and Combinatorics 24, Springer-Verlag, Berlin, 2003.

[47] A. Schrijver, *Theory of Linear and Integer Programming*, John Wiley & Sons, Chichester, 1986.

[48] A. Soifer, *The Mathematical Coloring Book*, Springer, New York, 2009.

[49] K. Stromberg, An elementary proof of Steinhaus's theorem, *Proceedings of the American Mathematical Society* 36 (1972) 308.

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

[51] L.A. Székely, Erdős on unit distances and the Szemerédi-Trotter theorems, in: *Paul Erdős and His Mathematics II* (Papers from the conference held in Budapest, July 4–11, 1999; G. Halász, L. Lovász, M. Simonovits, and V.T. Sós, eds.), Bolyai Society Mathematical Studies 11, János Bolyai Mathematical Society, Budapest; Springer-Verlag, Berlin, 2002, pp. 649–666.

[52] L.A. Székely, Measurable chromatic number of geometric graphs and sets without some distances in Euclidean space, *Combinatorica* 4 (1984) 213–218.

[53] L.A. Székely, Remarks on the chromatic number of geometric graphs, in: *Graphs and other combinatorial topics* (Proceedings of the Third Czechoslovak Symposium on Graph Theory, held in Prague, August 24th to 27th, 1982; M. Fiedler, ed.), Teubner-Texte zur Mathematik 59, Teubner, Leipzig, 1983, pp. 312–315.

[54] L.A. Székely and N.C. Wormald, Bounds on the measurable chromatic number of $\mathbb{R}^n$, *Discrete Mathematics* 75 (1989) 343–372.

[55] G.N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1922.

[56] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Publications de l'Institut de Mathématique de l'Université de Strasbourg IV, deuxième édition, Hermann, Paris, 1965.

[57] R. Wong and J.-M. Zhang, Asymptotic monotonicity of the relative extrema of Jacobi polynomials, *Canadian Journal of Mathematics* 46 (1994) 1318–1337.

[58] R. Wong and J.-M. Zhang, On the relative extrema of the Jacobi polynomials $P_n^{(0,-1)}(x)$, *SIAM Journal on Mathematical Analysis* 25 (1994) 776–811.

Index

Samenvatting

De meeste ontwikkelingen in dit proefschrift werden gemotiveerd door het volgende probleem: *Wat is het kleinste aantal kleuren dat nodig is om de punten in* $\mathbb{R}^n$ *te kleuren zodanig dat geen twee punten op afstand* 1 *van elkaar dezelfde kleur krijgen en dat de verzamelingen punten met dezelfde kleur Lebesgue-meetbaar zijn?* Dit kleinste aantal is $\chi_{\mathrm{m}}(\mathbb{R}^n)$, het *meetbare kleuringsgetal (measurable chromatic number)* van $\mathbb{R}^n$. Het is eenvoudig om in te zien dat $\chi_{\mathrm{m}}(\mathbb{R}) = 2$, maar al voor $n = 2$ is het alleen bekend dat $5 \leq \chi_{\mathrm{m}}(\mathbb{R}^2) \leq 7$. Een van de hoofddoelen bij de start van dit onderzoek was om betere ondergrenzen te vinden voor $\chi_{\mathrm{m}}(\mathbb{R}^n)$. We bestuderen ook de parameter $m_1(\mathbb{R}^n)$, de grootste dichtheid die een Lebesgue-meetbare deelverzameling van $\mathbb{R}^n$ kan hebben als deze geen twee punten met onderlinge afstand 1 mag bevatten. Deze parameter is gerelateerd aan $\chi_{\mathrm{m}}(\mathbb{R}^n)$, aangezien

$$\chi_{\mathrm{m}}(\mathbb{R}^n) m_1(\mathbb{R}^n) \geq 1.$$

In dit proefschrift modelleren we het probleem van het vaststellen van $\chi_{\mathrm{m}}(\mathbb{R}^n)$ en $m_1(\mathbb{R}^n)$ als een probleem over oneindige grafen. We breiden ideeën van de combinatorische optimalisering uit die voor eindige grafen gelden — in het bijzonder het Lovász theta getal — tot de oneindige grafen die wij bekijken. Door de symmetrie van deze grafen uit te buiten met hulp van Fourieranalyse kunnen we optimaliseringsproblemen definiëren en oplossen waarvan de optimale waardes grenzen geven voor $\chi_{\mathrm{m}}(\mathbb{R}^n)$ en $m_1(\mathbb{R}^n)$. Op deze manier verbeteren we de bestaande ondergrenzen voor $\chi_{\mathrm{m}}(\mathbb{R}^n)$ voor $n = 3, \ldots, 24$ en de bestaande bovengrenzen voor $m_1(\mathbb{R}^n)$ voor $n = 2, \ldots, 24$.

Behalve bovenstaande resultaten kunnen we nog meer resultaten afleiden met onze technieken. In het bijzonder kunnen we bewijzen dat $\chi_{\mathrm{m}}(\mathbb{R}^n)$ exponentieel groeit in n, ondanks dat onze schatting van de groeisnelheid niet beter is dan bestaande schattingen. Daarnaast kunnen we een numerieke versie krijgen van een stelling van Furstenberg, Katznelson en Weiss, die stelt dat een Lebesgue-meetbare verzameling $C \subseteq \mathbb{R}^n$ die geen paren punten op onderlinge afstand d_1, d_2, ... bevat dichtheid nul moet hebben. Onze resultaten tonen aan dat de dichtheid van een verzameling $C \subseteq \mathbb{R}^n$ die geen paren punten op onderlinge afstand $d_1, \ldots, d_N$ bevat exponentieel afneemt in N, mits de afstanden $d_1, \ldots, d_N$ uitdijen.

We ontwikkelen ook technieken voor grafen die over de sfeer gedefinieerd zijn. De resultaten die we voor zulke grafen verkrijgen geven ook ondergrenzen voor $\chi_{\mathrm{m}}(\mathbb{R}^n)$, die beter zijn dan de bestaande ondergrenzen voor $n = 10, \ldots, 24$. Voor de sfeer bewijzen we ook een versie van een stelling van Falconer, die lijkt op de zojuist beschreven stelling van Furstenberg, Katznelson en Weiss.

Summary

MOST of the developments in this thesis were motivated by the following problem: *What is the minimum number of colors one needs in order to color the points of* $\mathbb{R}^n$ *so that no two points at distance* 1 *get the same color and so that the sets of points that are colored with the same color are Lebesgue measurable?* This minimum is denoted by $\chi_{\mathrm{m}}(\mathbb{R}^n)$ and is called the *measurable chromatic number* of $\mathbb{R}^n$. It is easy to see that $\chi_{\mathrm{m}}(\mathbb{R}) = 2$, but already for $n = 2$ it is only known that $5 \leq \chi_{\mathrm{m}}(\mathbb{R}^2) \leq 7$. One of our main goals when we started this research was to find improved lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$. We also study the parameter $m_1(\mathbb{R}^n)$, the maximum density that a Lebesgue measurable subset of $\mathbb{R}^n$ can have if it does not contain two points at distance 1 from each other. This parameter is related to $\chi_{\mathrm{m}}(\mathbb{R}^n)$, since

$$\chi_{\mathrm{m}}(\mathbb{R}^n) m_1(\mathbb{R}^n) \geq 1.$$

In this thesis, we model the problems of determining $\chi_{\mathrm{m}}(\mathbb{R}^n)$ or $m_1(\mathbb{R}^n)$ as problems over infinite graphs. We then extend ideas from combinatorial optimization — most importantly, the Lovász theta number — which apply to finite graphs to the infinite graphs we consider. By exploiting the symmetry of these graphs, with the help of harmonic analysis, we are able to define and solve optimization problems whose optimal values provide bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ or $m_1(\mathbb{R}^n)$. In doing so we improve on existing lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$ for $n = 3, \ldots, 24$ and on existing upper bounds for $m_1(\mathbb{R}^n)$ for $n = 2, \ldots, 24$.

We may also derive further results from our methods. In particular, it is possible to prove that $\chi_{\mathrm{m}}(\mathbb{R}^n)$ grows exponentially with n, even though our estimate for the growth ratio is not better than previously known estimates. Also, we obtain a numerical version of a theorem of Furstenberg, Katznelson, and Weiss, which states that a Lebesgue measurable set $C \subseteq \mathbb{R}^n$ that does not contain pairs of points at distances $d_1, d_2, \ldots$ from each other, for some unbounded sequence $d_1, d_2, \ldots$, must have density zero. Our result shows that the density of a set $C \subseteq \mathbb{R}^n$ that does not contain pairs of points at distances $d_1, \ldots, d_N$ decreases exponentially with N, as long as the distances $d_1, \ldots, d_N$ space out.

We also develop methods for graphs defined over the sphere. Results we obtain

from such graphs also provide lower bounds for $\chi_{\mathrm{m}}(\mathbb{R}^n)$, which are better than previously known lower bounds for $n = 10, \ldots, 24$. For the sphere we also prove a version of a theorem of Falconer, which is close to the theorem by Furstenberg, Katznelson, and Weiss discussed above.