Packing Regular Spherical Polygons

Dissertation

zur

Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Andreas Spomer

2026

Begutachtung:	Prof. Dr. Frank Vallentin
	Prof. Dr. Achill Schürmann
Vorsitz der Prüfungskommission:	Prof. Dr. David Gross
Datum der Disputation:	13.07.2026

Der Leser soll gleichsam in dem großen Garten der Geometrie spazieren geführt werden, und jeder soll sich einen Strauß pflücken können, wie er ihm gefällt.

— David Hilbert, *Anschauliche Geometrie* (1932)

Abstract

This thesis studies packing problems for regular spherical polygons on the two-dimensional unit sphere. It develops techniques for constructing lower and upper bounds for the packing density of spherical polygons with varying sizes and numbers of vertices.

A central contribution of this thesis is the introduction of a necessary and sufficient algebraic criterion for deciding when two congruent regular spherical polygons within a prescribed class have disjoint interiors. This characterization serves as the principal tool underlying both the constructive lower bounds and the analytic upper bounds established in the thesis.

To obtain lower bounds, we formulate the problem of maximizing packing density for a fixed number of regular spherical polygons as an optimization problem on a Riemannian submanifold. We describe numerical methods for approximating stationary points and show how to convert the resulting data into rigorous configurations.

We use three complementary approaches to establish upper bounds. The strongest among them is an extension of the Lovász theta number to Cayley graphs on the special orthogonal group. Its computation is based on the representation theory of the special orthogonal group and leads to optimization problems involving trigonometric sum-of-squares polynomials. By exploiting the symmetries of regular spherical polygons, we significantly reduce the complexity of these problems.

Contents

Part I
Basic notions and techniques

Chapter One

Introduction

1.1 On the history of packing problems

Geometric objects have long played a central role in attempts to describe structures in nature. In his dialogue *Timaeus* [62], Plato associates the classical elements with regular polyhedra, while modern physical models often represent solid matter as an arrangement of densely packed spheres. Although these examples differ substantially in purpose and scientific context, both reflect the broader idea that structure can be studied in geometric terms. In both cases, one is naturally led to the question of how geometric bodies can be arranged in space as efficiently as possible. Problems of this kind are referred to as *packing problems*.

The history of packing problems contains many remarkable developments. The following brief overview is based in part on the expositions given in [29, 30, 50].

The *sphere packing problem* asks for the maximal density of space that can be occupied by nonoverlapping spheres of equal radius. In his work *Strena seu de nive sexangula* from 1611, the German mathematician and astronomer Johannes Kepler described a packing that fills space by a fraction of

$$\frac{\pi}{\sqrt{18}} \approx 0.74048,$$

which he conjectured to be optimal [49]. It took almost 400 years until a proof of this conjecture was published by Thomas Hales in 2005 [45].

While the first proof of the two-dimensional analogue of this problem was published by Thue [77], it is often considered incomplete by modern standards, and the proof is therefore commonly attributed to Fejes Tóth [36].

Apart from the trivial one-dimensional case, the only other dimensions in which the sphere packing problem has been solved are 8 and 24, by Maryna Viazovska [80] and by Cohn, Kumar, Miller, Radchenko, and Viazovska [20], respectively.

Another well-studied packing problem concerns regular tetrahedra. It dates back to Aristotle, who commented on Plato's theory in his treatise *De Caelo*

> "Among surfaces it is agreed that there are three figures which fill the place that contains them — the triangle, the square, and the hexagon: among solids only two, the pyramid and the cube"

[2, translation by Guthrie]. By the term "pyramid" Aristotle refers to the regular tetrahedron, thus claiming that three-dimensional space can be tiled by regular tetrahedra.

However, such a tiling is impossible, as the following argument illustrates. Suppose there exists a tiling of space by regular tetrahedra. We first argue that any such tiling would necessarily be face-to-face, with coinciding edges along shared faces.

If the tiling were not face-to-face, then locally one of two situations would have to occur, namely either an edge intersects the relative interior of a face, or a vertex lies in the relative interior of a face. In the latter case, moving slightly away from the vertex leads again to the first configuration. Hence it suffices to discuss the edge–face case.

Suppose that an edge of one tetrahedron intersects the relative interior of a face of another tetrahedron. At a generic point of intersection, a plane perpendicular to the edge yields a planar subdivision containing the dihedral angle $\theta = \arccos(1/3)$ and a half-plane of angle π. Since the remaining angle $\pi - \theta$ is not an integer multiple of θ, the remaining region cannot be decomposed into sectors of angle θ. Therefore such an intersection configuration is impossible.

In a face-to-face tiling, the dihedral angles around each edge must sum to 2π. Since $5\theta < 2\pi < 6\theta$, at most five tetrahedra can meet along a common edge, leaving a gap that cannot be filled by an additional tetrahedron, as illustrated in Figure 1.1.

The disproof of Aristotle's statement is attributed to Johannes Müller von Königsberg, better known as Regiomontanus. Although only the title of his manuscript has been preserved, it is generally assumed that he possessed the necessary tools to refute the claim.

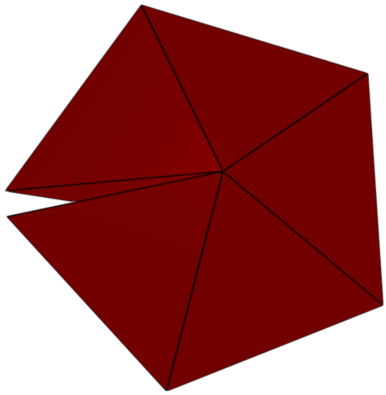


Figure 1.1: Five tetrahedra sharing an edge leave a gap that is too small to accommodate an additional tetrahedron.

The study of dense packings of regular tetrahedra remains an active area of research. At the time of writing, the densest known packing is due to Chen, Engel, and Glotzer [17], which covers space by a fraction of

$$\frac{4000}{4671} \approx 0.85634.$$

Moreover, dense packings have been found for the other four Platonic solids. Torquato and Jiao provide an overview of the currently best-known packings, along with numerical evidence suggesting their optimality [78].

Naturally, the development of increasingly dense packings raises the question of upper bounds for the packing density. For nonspherical bodies, this is substantially more difficult than in the sphere packing case, since the relative orientations of the solids must be taken into account. At present, the only known nontrivial upper bounds for Platonic solids are due to Gravel, Elser, and Kallus, who established an upper bound of $1 - 2.6 \cdot 10^{-25}$ for packings of regular tetrahedra and an upper bound of $1 - 1.4 \cdot 10^{-12}$ for packings of octahedra [43]. As the authors themselves note, these bounds are unlikely to be sharp.

While it is difficult to construct global upper bounds for such packing problems, a local analysis can provide valuable insight. For example, one may consider the question of how many nonoverlapping congruent copies of a Platonic solid can be arranged around a given point in space. This naturally leads to configurations in which the point is a common vertex of all copies. To determine how many solids can meet at this point, we place a sphere of sufficiently small radius centered at the point. The Platonic solids then induce spherical polygons on the surface of this sphere as illustrated in Figure 1.2. For a regular tetrahedron, the induced spherical polygon is an equilateral spherical triangle.

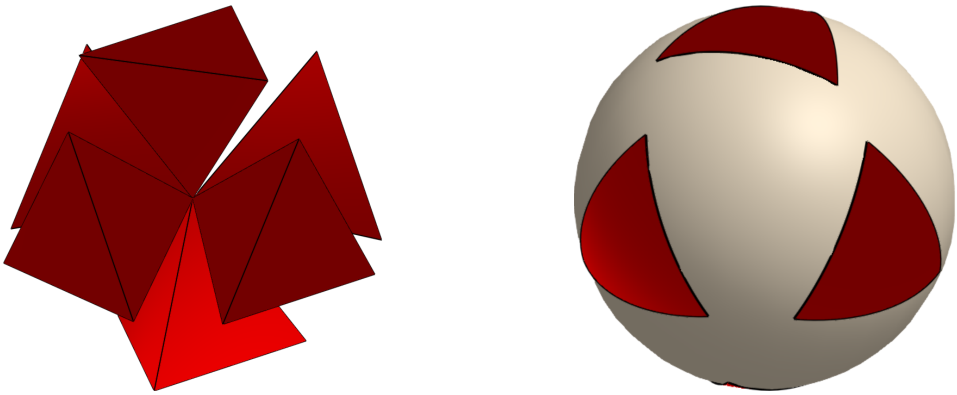


Figure 1.2: Six polygons sharing a vertex induce six spherical triangles on the corresponding sphere.

Thus, intersecting the configuration with a sufficiently small sphere centered at the common vertex transforms the problem into one of packing congruent spherical polygons on the unit sphere. A simple way to obtain an upper bound is to divide the surface area of the sphere by the area of one such polygon. This bound is commonly referred to as the *volume bound*.

In most cases, this trivial upper bound is insufficient to prove optimality. For regular tetrahedra, the volume bound shows that at most 22 nonoverlapping tetrahedra can share a vertex, whereas the largest configuration currently known consists of only 20 tetrahedra.

An explicit construction of such a configuration can be obtained as follows. Consider an icosahedron scaled such that the distance between its center and its vertices is equal to 1. The edge length of the icosahedron then exceeds 1. Consequently, one can place a regular tetrahedron of edge length 1 on each face, with the apex at the center of the icosahedron. This yields a configuration of exactly twenty nonoverlapping tetrahedra sharing a common vertex at the center of the icosahedron, as shown in Figure 1.3.

A different approach to obtaining local bounds is provided by the *kissing number problem*, which asks for the maximum number of nonover-

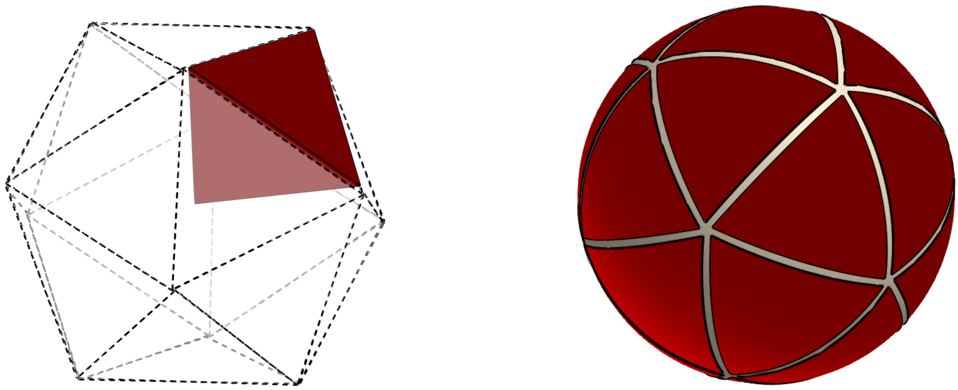


Figure 1.3: The icosahedron defines a configuration of twenty nonoverlapping regular tetrahedra that share a vertex. This configuration corresponds to a packing of twenty spherical triangles.

lapping congruent copies of a convex body that touch a fixed copy.

Famously, the kissing number problem for unit spheres in three dimensions sparked a debate between David Gregory and Isaac Newton. Gregory claimed that thirteen nonoverlapping unit spheres could touch a given unit sphere, whereas Newton argued that the correct number is twelve [71]. The debate was finally settled in 1953 by Schütte and van der Waerden, who proved that the kissing number is 12, thereby confirming Newton's claim [69].

Beyond the trivial one- and two-dimensional cases, the problem has been solved in dimension four by Musin [59], and in dimensions eight and twenty-four by Kabatiansky and Levenshtein [48], and independently by Odlyzko and Sloane [60].

Except for the cube, the kissing numbers of the Platonic solids remain largely unknown. Most known kissing configurations are obtained from corresponding packing constructions. One notable exception is the regular tetrahedron. In her PhD thesis, Chen constructed a kissing configuration of 56 nonoverlapping tetrahedra [16, 50].

To obtain this configuration, arrange four regular tetrahedra around each edge of a fixed tetrahedron. This yields an initial kissing configuration of sixteen tetrahedra. Each vertex of the original tetrahedron is touched by nine other tetrahedra. Additional tetrahedra are then placed in the remaining space around each vertex. Using the icosahedral construction discussed above, one obtains a kissing configuration of exactly 56 tetrahedra.

Whether this construction can be further improved depends on the maximal number of nonoverlapping regular tetrahedra that can share a vertex, and hence on the maximal number of spherical triangles of the corresponding size that can be packed on the sphere. This illustrates that local spherical packing problems arise naturally in the study of packing and kissing configurations of Platonic solids. More generally, it suggests that understanding the arrangement of spherical polygons arising from these configurations is of fundamental importance.

This perspective naturally leads to the central question of this thesis

> *How many nonoverlapping congruent regular spherical polygons of prescribed size can be packed on the two-dimensional sphere?*

While this question is motivated by three-dimensional packing problems, it is also of independent interest. It can be viewed as a packing problem for convex objects in a compact space. Whereas packing problems in non-compact spaces are typically difficult to resolve, as discussed above, the compactness of the unit sphere renders this problem more tractable. At the same time, the curvature of the sphere introduces additional geometric difficulties that do not arise in Euclidean space.

Besides the configuration of twenty spherical triangles mentioned above, the main family of further examples consists of packings of spherical pentagons studied by Tarnai and Gáspár [73, 75, 76]. This problem is of particular interest because of its connection to the structure of virus capsids, see [74], and the references therein. Beyond these examples, no general classification of optimal configurations is known.

Apart from the volume bound, no nontrivial upper bounds are currently known. This reflects the intrinsic difficulty of the problem, which is further compounded by the role of orientations, as is already the case in Euclidean packing problems.

1.2 Outline of the thesis

The present thesis is primarily concerned with the packing problem for regular spherical polygons, and also discusses extensions of the underlying framework to a broader class of spherical polytopes. Its main results build on the paper "Bounding the density of spherical polygon packings" by Fernando Mário de Oliveira Filho, Frank Vallentin, and the author of this thesis [27], while extending several theoretical aspects of that work. For an overview of the mathematical framework underlying this approach, see [29].

The thesis is organized into three main parts. Following this introduction, the first part contains the preliminary material required throughout the thesis, which is presented in Chapter 2.

Chapter 3 develops the foundations of spherical geometry used in the subsequent chapters. Its central contribution is Theorem 3.14, which provides a criterion for determining whether two congruent regular spherical polygons of a prescribed type are interior-disjoint. This result plays a crucial role in the developments presented in the subsequent chapters.

The second part is devoted to the derivation of lower bounds through the construction of explicit configurations. In particular, we introduce a novel approach in which the problem of maximizing the packing density for a fixed number of regular spherical polygons is formulated as an optimization problem on a Riemannian submanifold. In addition, we present a numerical method for approximating stationary points of this problem and explain how these approximations can be used to obtain rigorous results.

Chapter 5 concludes this part with a presentation of the configurations obtained using the proposed method. The configurations are compared with previously known results, including those of Tarnai and Gáspár [73, 75, 76]. To avoid redundancy, we refrain from providing a detailed list of all configurations at this point. A complete list is given in Chapter 9.

The final part of the thesis is devoted to proving the maximality of the configurations obtained in the preceding chapters. In Chapter 6, we derive upper bounds using the framework of topological packing graphs introduced by de Laat and Vallentin [26]. A central result is Theorem 6.1, which gives an upper bound for spherical polygon packings based on the Lovász theta number. Unlike the elementary volume bound, this bound incorporates the dependence of intersections on the relative orientations of the polygons. We also show how upper bounds for spherical polygon packings can be derived from existing upper bounds for the kissing number problem.

The subsequent chapters develop a computational framework tailored to the theta number. Central to this approach are positive type functions, whose characterization requires a complete set of irreducible representations of the three-dimensional special orthogonal group. These representations are constructed in Chapter 7. While the representation-theoretic tools themselves are classical, their application to the present problem appears to be novel.

In Chapter 8, we present a method for computing the theta number through trigonometric polynomial optimization. We give a concise overview of the necessary background, followed by an explanation of how symmetry can be used to reduce the computational complexity of trigonometric polynomial optimization problems. Building on the invariant theory of ordinary polynomials, we introduce a novel framework for the invariant theory of trigonometric polynomials. This approach differs from existing approaches and yields structural results of independent interest. Using this framework, we reformulate the theta number as a trigonometric polynomial optimization problem and show how the symmetries of regular spherical polygons lead to substantial simplifications.

Chapter 9 concludes the thesis by presenting the lower and upper bounds obtained for the considered packing problems. We analyze the implications of these results for the geometric questions posed in the introduction and highlight the cases in which maximality can be established.

Chapter Two

Preliminaries

This chapter collects the fundamental concepts and terminology employed throughout this thesis. The underlying mathematical framework extends significantly beyond the scope of this chapter. Therefore, only the notions required for the later chapters are introduced. Each section is prefaced with references to literature providing a more detailed treatment of the respective topic.

Convex cones [5, 51, 66–68]. A nonempty subset of a real topological vector space that is closed under addition and multiplication by nonnegative scalars is called a *convex cone*. A convex cone $\mathcal{K}$ is *proper* if it is closed, has nonempty interior, and $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$.

Let $\mathcal{K} \subseteq \mathbb{R}^d$ be a convex cone. Its *dual cone* $\mathcal{K}^*$ is defined by

$$\mathcal{K}^* = \left\{ y \in \mathbb{R}^d : \; x^\top y \leq 0 \text{ for all } x \in \mathcal{K} \right\}.$$

While the sign convention is nonstandard in convex optimization, it is commonly used in spherical geometry.

An important class of cones is given by *polyhedral cones*. A polyhedral cone is a finitely generated cone, i.e., a cone of the form

$$\operatorname{cone}\{v_1, \ldots, v_N\} = \left\{ \sum_{k=1}^{N} \lambda_k v_k : \lambda_k \geq 0 \text{ for all } k = 1, \ldots, N \right\},$$

where $v_1, \ldots, v_N \in \mathbb{R}^d$. By the Farkas–Minkowski–Weyl theorem, the dual of a polyhedral cone is again polyhedral [67, Corollary 7.1a].

Each nonzero element y of the dual cone induces a supporting hyperplane of the primal cone

$$\left\{ x \in \mathbb{R}^d : \; y^\top x = 0 \right\}.$$

The intersection of such a hyperplane with a cone $\mathcal{K}$ defines a face of $\mathcal{K}$. One-dimensional faces are called *extreme rays* and $(\dim(\mathcal{K}) - 1)$-dimensional faces are called *facets*.

Let $\mathcal{K} = \text{cone}\{v_1, \ldots, v_N\}$ be a full-dimensional polyhedral cone. A supporting hyperplane of $\mathcal{K}$ defines a facet if and only if the corresponding normal vector lies in an extreme ray of the dual cone. Every extreme ray of a polyhedral cone is spanned by an element of its generating set $\{v_1, \ldots, v_N\}$. Moreover, if $\mathcal{K}$ is full-dimensional, then every facet contains $d-1$ linearly independent generators.

In $\mathbb{R}^3$, the cross product provides a convenient tool for computing facet normals. For $x, y \in \mathbb{R}^3$, it is defined by

$$x \times y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

The cross product is an antisymmetric bilinear map satisfying the *Lagrange identity*

$$(x \times y)^\top (v \times w) = (x^\top v)(y^\top w) - (x^\top w)(y^\top v). \tag{2.1}$$

For a discussion of generalizations of the cross product to higher-dimensional Euclidean spaces, see [56].

Optimization [5, 6, 13, 51, 67]. Let $\mathcal{X}$ be a Hausdorff space. The optimization problems considered in this thesis are of the form

$$\begin{aligned} p = \sup \quad & c(x) \\ \text{s.t.} \quad & x \in \mathcal{S}, \\ & f_j(x) \geq 0 \text{ for all } j = 1, \ldots, m \end{aligned}$$

where $\mathcal{S} \subseteq \mathcal{X}$ and $c, f_1, \ldots, f_m : \mathcal{X} \to \mathbb{R}$ are continuous functions. No convexity assumptions are generally imposed on $\mathcal{S}$ or on the functions $c, f_1, \ldots, f_m$.

A point $x \in \mathcal{S}$ is called a *feasible solution* if $f_j(x) \geq 0$ for all $j = 1, \ldots, m$. If the supremum is attained by a feasible point x, we say that x is an *optimal solution*.

If $\mathcal{X}$ is a real vector space, $\mathcal{S}$ is a convex cone, and $c, f_1, \ldots, f_m$ are linear functionals, we refer to the problem as a *conic optimization problem*.

An important special case arises when $\mathcal{X}$ is the vector space of real symmetric matrices equipped with the trace inner product $\langle A, B \rangle = \text{tr}(AB)$ and $\mathcal{S}$ is the cone of positive semidefinite matrices. The corresponding problem is called a *semidefinite optimization problem*. Semidefinite optimization forms the basis of the numerical framework employed in later chapters. For a detailed background on semidefinite optimization and its applications, see [6, 13, 51].

Matrix Lie groups [12, 31, 37, 38, 46]. Throughout this thesis, the identity element of a general group is denoted by e.

Let V be a d-dimensional vector space equipped with the inner product $\langle \cdot, \cdot \rangle$. A nonempty subset $\mathcal{M} \subseteq V$ is called an n-dimensional *Riemannian submanifold* of V if either

i) $n = d$ and $\mathcal{M}$ is open;

ii) $n = d - k$ for some $k \geq 1$, and for each $x \in \mathcal{M}$ there exist a neighborhood U of x in V and a smooth function $f : U \to \mathbb{R}^k$ such that the differential of f at x has rank k and $\mathcal{M} \cap U = \{y \in U : f(y) = 0\}$.

The inner product on V induces a Riemannian metric on $\mathcal{M}$. A *matrix Lie group* is a closed subgroup of the general linear group over the complex numbers. Every matrix Lie group is a Riemannian submanifold of the ambient matrix space [46, Corollary 3.45].

Throughout this thesis, we are mainly concerned with two families of matrix Lie groups, namely the *special orthogonal group*

$$\mathrm{SO}(d) = \left\{ A \in \mathbb{R}^{d \times d} : \ A^\top A = I, \, \det(A) = 1 \right\}$$

and the *special unitary group*

$$\mathrm{SU}(d) = \left\{ U \in \mathbb{C}^{d \times d} : \ U^* U = I, \, \det(U) = 1 \right\}.$$

Both are compact matrix Lie groups. For a detailed discussion of their topological properties, the reader is referred to [46].

For every compact matrix Lie group Γ, there exists a unique probability measure μ that satisfies

$$\mu(\gamma \, A) = \mu(A \, \gamma) = \mu(A)$$

for all $\gamma \in \Gamma$ and all measurable sets $A \subseteq \Gamma$. This measure is called the *Haar measure*. For a proof of existence and uniqueness, see [37, Chapter 2.2]. The Haar measure induces an inner product on the space $L^2(\Gamma)$.

Let Γ be a matrix Lie group and let V be a Hilbert space. A *group action* of Γ on V is a continuous map $\Gamma \times V \to V$, $(\gamma, v) \mapsto \gamma \cdot v$ with the properties

i) $e \cdot v = v$ for all $v \in V$;

ii) $\gamma_1 \cdot (\gamma_2 \cdot v) = (\gamma_1 \gamma_2) \cdot v$ for all $\gamma_1, \gamma_2 \in \Gamma$ and $v \in V$.

A group action is called linear if the map $v \mapsto \gamma \cdot v$ is linear for each $\gamma \in \Gamma$. All group actions discussed in this thesis are assumed to be linear.

We refer to a vector v as *invariant under the action of* Γ if $\gamma \cdot v = v$. We call a subspace $W \subseteq V$ *invariant* under the action of Γ if $\gamma \cdot W \subseteq W$. An invariant subspace W is *irreducible* if there is no proper closed nonzero invariant subspace of W.

A *representation* is a continuous group homomorphism mapping elements of the group Γ to the general linear group of V. Note that every group action defines a representation and vice versa. A representation is called *irreducible* if its ambient space is irreducible.

Let Γ be a compact matrix Lie group that acts on the vector spaces V_1 and V_2. We call a linear map $A : V_1 \to V_2$ an *equivariant map* if

$$A(\gamma \cdot v) = \gamma \cdot A(v)$$

holds for all $\gamma \in \Gamma$ and $v \in V_1$. If the map A is invertible, the corresponding representations of Γ are called *equivalent*. *Schur's lemma* states that any nonzero equivariant map between irreducible representations is an isomorphism [46, Theorem 4.29].

This notion defines an equivalence relation on the set of irreducible unitary representations. The *dual group* $\hat{\Gamma}$ of Γ is defined as the set of equivalence classes of irreducible unitary representations of Γ. If the group Γ is finite, so is its dual group $\hat{\Gamma}$ [38, Corollary 2.13].

Let Γ be a compact matrix Lie group that acts on a separable complex Hilbert space V by a unitary representation. The Peter–Weyl theorem implies that the action of Γ decomposes V as the direct sum of irreducible subspaces

$$V = \bigoplus_{\rho \in \hat{\Gamma}} \bigoplus_{k=1}^{m_\rho} V_{\rho,k},$$

where the action of Γ on each subspace $V_{\rho,k}$ is induced by the representation ρ and $m_\rho \in \mathbb{N}_0 \cup \{\infty\}$ denotes the corresponding multiplicity. Moreover, each space $V_{\rho,k}$ is finite-dimensional and the set $\hat{\Gamma}$ is countable. For a precise statement of the Peter–Weyl theorem and its implications, see [31, Chapter 7].

Suppose ρ is a finite-dimensional unitary representation of a compact matrix Lie group Γ. The *character* of ρ is the function

$$\chi_\rho : \Gamma \to \mathbb{C}, \qquad \gamma \mapsto \operatorname{tr}(\rho(\gamma)).$$

Using the cyclic property of the trace, one can verify that two equivalent representations define the same character. Hence the character depends only on the equivalence class of the representation.

The character of a representation ρ satisfies

$$\chi_\rho(\gamma) = \chi_\rho(g\gamma g^{-1})$$

for all $\gamma, g \in \Gamma$. Functions with this property are called *class functions.* The characters corresponding to irreducible unitary representations form an orthonormal basis of the closed subspace of $L^2(\Gamma)$ consisting of class functions, with respect to the inner product induced by the Haar measure [46, Theorems 12.15 and 12.18].

Let Γ be a compact matrix Lie group acting linearly and continuously on a Banach space $\mathcal{X}$. We define the *symmetrization operator* by

$$x \mapsto \int_\Gamma \gamma \cdot x \; d\mu(\gamma),$$

where μ is the Haar measure of Γ. The image of $x \in \mathcal{X}$ under the symmetrization operator is denoted by x_{sym}. By construction, $\gamma \cdot x_{\text{sym}} = x_{\text{sym}}$ holds.

Let $\mathcal{S} \subseteq \mathcal{X}$ and let $c, f_1, \ldots, f_m : \mathcal{X} \to \mathbb{R}$ be continuous functions. The optimization problem

$$\begin{aligned} \sup \quad & c(x) \\ \text{s.t.} \quad & x \in \mathcal{S}, \\ & f_j(x) \geq 0 \text{ for all } j = 1, \ldots, m \end{aligned}$$

is said to be Γ-*invariant* if the search space $\mathcal{S}$ is closed under symmetrization, i.e.,

$$\mathcal{S}_{\text{sym}} = \{x_{\text{sym}} : x \in \mathcal{S}\} \subseteq \mathcal{S},$$

and all $x \in \mathcal{S}$ satisfy $f_j(x) = f_j(x_{\text{sym}})$ for all $j = 1, \ldots, m$ as well as $c(x) = c(x_{\text{sym}})$. In particular, the symmetrization of every feasible solution is a feasible solution and attains the same objective value. Thus, restricting the search space to the symmetrized set $\mathcal{S}_{\text{sym}}$ does not affect the optimal value.

Chapter Three

Geometry on the sphere

Spherical geometry provides a natural framework for studying packing problems on the sphere. In contrast to the Euclidean setting, the curvature of the sphere significantly affects the geometry of convex bodies. This necessitates an adaptation of the classical theory of convex bodies to the spherical setting.

We begin by introducing the basic concepts and terminology from spherical geometry used throughout this thesis, following [22, 65, 66].

Rigid motions of convex bodies on the sphere are described by means of an Euler angle parametrization of SO(3), which serves as an important tool throughout this thesis. Our presentation follows [18, 34, 47].

A central issue in packing problems is to determine whether two congruent convex bodies have disjoint interiors. This question is addressed in later sections of this chapter. In particular, Theorem 3.14 provides an algorithmic criterion for deciding whether two congruent regular spherical polygons have disjoint interiors. This theorem is fundamental to the results developed in the subsequent chapters.

3.1 Spherically convex bodies

For an integer $d \geq 2$, let

$$\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : x^\top x = 1\}$$

denote the $(d-1)$–dimensional unit sphere embedded in $\mathbb{R}^d$. We endow $\mathbb{S}^{d-1}$ with the *geodesic metric*

$$\delta(x, y) = \arccos\left(x^\top y\right) \qquad \text{for} \qquad x, y \in \mathbb{S}^{d-1}.$$

The interior and boundary of a set $\mathcal{Q} \subseteq \mathbb{S}^{d-1}$ with respect to the topology induced by the geodesic metric are denoted by $\mathcal{Q}^\circ$ and $\partial\mathcal{Q}$, respectively.

We refer to sets of the form $\mathcal{Q} = \mathcal{K} \cap \mathbb{S}^{d-1}$, where $\mathcal{K} \subseteq \mathbb{R}^d$ is a convex cone, as *spherically convex* sets. If the defining cone is proper, the set forms a *spherically convex body.* The set $\mathcal{Q}^* = \mathcal{K}^* \cap \mathbb{S}^{d-1}$ is called the *dual* of $\mathcal{Q}$. A set $A\mathcal{Q}$ with $A \in \mathrm{SO}(d)$ is a *congruent copy* of $\mathcal{Q}$. A finite collection of congruent copies of $\mathcal{Q}$ is called a *configuration* of $\mathcal{Q}$.

The *geodesic arc* between two points $x, y \in \mathbb{S}^{d-1}$ with $\delta(x,y) \neq \pi$ is given by

$$[x,y] = \{z/\|z\| : \ z = \lambda x + (1-\lambda)y, \ \lambda \in [0,1]\}. \tag{3.1}$$

A set is spherically convex if and only if it is closed under taking geodesic arcs between any pair of non-antipodal points. This characterization is commonly attributed to Robinson [22].

A *spherical cap* of angular radius $\varrho \in (0, \pi/2)$ centered at the standard basis vector $e_d = (0, \dots, 0, 1)$ is defined by

$$\mathcal{C}(\varrho) = \left\{x \in \mathbb{S}^{d-1} : x_d \geq \cos(\varrho)\right\}.$$

It is the closed ball of radius ϱ centered at e_d with respect to the geodesic metric. A spherical cap of the same radius centered at any other point of the sphere is a congruent copy of $\mathcal{C}(\varrho)$. Expressing $\mathcal{C}(\varrho)$ as

$$\mathcal{C}(\varrho) = \mathbb{S}^{d-1} \cap \left\{(x,t) \in \mathbb{R}^{d-1} \times \mathbb{R} : \ \|x\| \leq \tan(\varrho)\, t\right\},$$

shows that it is the intersection of $\mathbb{S}^{d-1}$ with a proper convex cone, and hence a spherically convex body.

Sets of the form $\mathcal{P} = \mathcal{K} \cap \mathbb{S}^{d-1}$, where $\mathcal{K} \subseteq \mathbb{R}^d$ is a proper polyhedral cone, are called *spherical polytopes.* Since the dual of a polyhedral cone is polyhedral, the dual $\mathcal{P}^*$ is again a spherical polytope. The intersections of the extreme rays of $\mathcal{K}$ with $\mathbb{S}^{d-1}$ are called the *vertices* of $\mathcal{P}$.

A subclass of spherical polytopes on $\mathbb{S}^2$ of particular importance is given by the *regular spherical polygons.* Let $\varrho \in (0, \pi/2)$ and $N \geq 3$. For $k \in \mathbb{Z}$, define

$$v_k(\varrho) = \begin{pmatrix} \sin(\varrho)\sin\left((2k-1)\pi/N\right) \\ \sin(\varrho)\cos\left((2k-1)\pi/N\right) \\ \cos(\varrho) \end{pmatrix}. \tag{3.2}$$

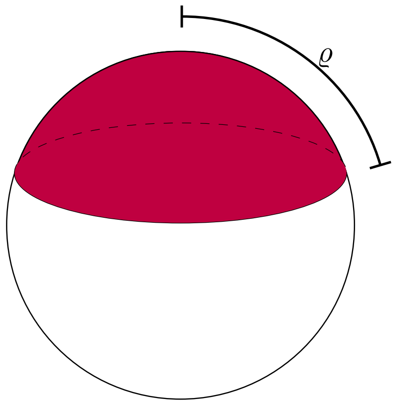


Figure 3.1: The cap $\mathcal{C}(\varrho)$ is the closed geodesic ball of radius ϱ centered at e_d with respect to the geodesic metric. It can be represented as the intersection of $\mathbb{S}^{d-1}$ with a proper convex cone in $\mathbb{R}^d$.

The points $v_k(\varrho)$ lie on the circle of latitude at angular distance ϱ from e_3 and are equally spaced in longitude.

We refer to congruent copies of the set

$$\mathcal{G}_N(\varrho) = \mathbb{S}^2 \cap \operatorname{cone}\{v_1(\varrho), \dots, v_N(\varrho)\} \tag{3.3}$$

as *regular spherical polygons* of angular radius ϱ.

The vertices of the dual polytope $\mathcal{G}_N(\varrho)^*$ are denoted by $h_j(\varrho)$, where $j = 1, \dots, N$. Their indices are chosen so that $h_j(\varrho)^\top v_{j-1}(\varrho) = h_j(\varrho)^\top v_j(\varrho) = 0$ for $j = 1, \dots, N$. An explicit formula is obtained from the cross product

$$h_k(\varrho) = \frac{v_{k-1}(\varrho) \times v_k(\varrho)}{\|v_{k-1}(\varrho) \times v_k(\varrho)\|}. \tag{3.4}$$

This formula extends the definition of h_k to all $k \in \mathbb{Z}$. From

$$v_{k-1}(\varrho) \times v_k(\varrho) = 2\sin(\varrho)\sin(\pi/N) \begin{pmatrix} \cos(\varrho)\sin(2(k-1)\pi/N) \\ \cos(\varrho)\cos(2(k-1)\pi/N) \\ -\sin(\varrho)\cos(\pi/N) \end{pmatrix}$$

we deduce that the dual of a regular spherical polygon is again a regular spherical polygon, although generally with a different angular radius.

The angular distance between two vertices v_k and v_ℓ of $\mathcal{G}_N(\varrho)$ is given by

$$\delta(v_k(\varrho), v_\ell(\varrho)) = \arccos\left(\sin(\varrho)^2 \cos(2(k-\ell)\pi/N) + \cos(\varrho)^2\right). \tag{3.5}$$

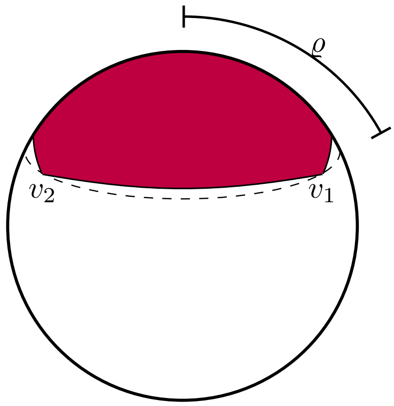


Figure 3.2: A regular spherical polygon $\mathcal{G}_N(\varrho)$ on $\mathbb{S}^2$. Its vertices lie on the circle of latitude at angular distance ϱ from e_3. The polygon is obtained by intersecting $\mathbb{S}^2$ with the polyhedral cone spanned by these vertices.

Throughout this thesis, $\varrho \in (0, \pi/2)$ and $N \geq 3$ are assumed unless stated otherwise. Whenever no ambiguity arises, we omit the dependence on ϱ in the notation.

The packing problem for a spherically convex body $\mathcal{Q} \subseteq \mathbb{S}^{d-1}$ is defined as the optimization problem

$$\begin{aligned} \tau(\mathcal{Q}) = \max \quad & |\mathcal{R}| \\ \text{s.t.} \quad & \mathcal{R} \subseteq \mathrm{SO}(d) \\ & (A\mathcal{Q} \cap B\mathcal{Q})^\circ = \emptyset \quad \text{for all } A, B \in \mathcal{R},\ A \neq B. \end{aligned}$$

The quantity $\tau(\mathcal{Q})$ is called the *packing number* of $\mathcal{Q}$. Any feasible solution is called a *packing* of $\mathcal{Q}$.

A basic symmetry of the packing problem is that pairwise feasibility depends only on relative rotations. Indeed, for all $A, B \in \mathrm{SO}(d)$,

$$(A\mathcal{Q} \cap B\mathcal{Q})^\circ = \emptyset \qquad \text{if and only if} \qquad (\mathcal{Q} \cap A^\top B\mathcal{Q})^\circ = \emptyset. \tag{3.6}$$

A packing $\mathcal{R} \subseteq \mathrm{SO}(d)$ is called *maximal* if $|\mathcal{R}| = \tau(\mathcal{Q})$. We refer to a maximal packing of spherical caps or regular spherical polygons as *optimal* if any increase in the corresponding angular radius makes a packing of the same cardinality infeasible.

The main focus of this thesis is on packing problems for regular spherical polygons, namely the problems $\tau(\mathcal{G}_N(\varrho))$, where $N \geq 3$ is the number of vertices and $\varrho \in (0, \pi/2)$ denotes the angular radius of the polygon.

Other notable examples of packing problems arise from spherical caps. For a given angular radius $\varrho \in (0, \pi/2)$, the packing problem $\tau(\mathcal{C}(\varrho))$ with $\mathcal{C}(\varrho) \subseteq \mathbb{S}^{d-1}$ is equivalent to determining the largest spherical code whose pairwise inner products are at most $\cos(2\varrho)$. In particular, the classical kissing number problem corresponds to $\tau(\mathcal{C}(\pi/6))$.

3.2 Euler angles

A parametrization of $\mathrm{SO}(d)$ provides a convenient framework for describing rigid motions on the sphere. In this section, we restrict to the three-dimensional case, where rotations admit a classical parametrization in terms of Euler angles.

Let $\theta \in \mathbb{R}$. The two basic rotation matrices are given by

$$R_X(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

and

$$R_Z(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

These describe the rotations by the angle θ about the x-axis and z-axis, respectively. The matrices satisfy

$$\begin{aligned} R_Z(\theta)^\top &= R_Z(-\theta), & R_X(\theta)^\top &= R_X(-\theta), \\ R_Z(\theta_1+\theta_2) &= R_Z(\theta_1)R_Z(\theta_2), & R_X(\theta_1+\theta_2) &= R_X(\theta_1)R_X(\theta_2). \end{aligned}$$

We define the parametrization $A : \mathbb{R}^3 \to \mathrm{SO}(3)$ by

$$A(\alpha, \beta, \gamma) = R_Z(\gamma)R_X(\beta)R_Z(\alpha). \tag{3.7}$$

Since each basic rotation lies in $\mathrm{SO}(3)$, we have $A(\alpha, \beta, \gamma) \in \mathrm{SO}(3)$ for every triple $(\alpha, \beta, \gamma) \in \mathbb{R}^3$.

Proposition 3.1. *For every $R \in \mathrm{SO}(3)$, there exists a triple $(\alpha, \beta, \gamma) \in [-\pi, \pi)^3$ such that*

$$R = R_Z(\gamma)R_X(\beta)R_Z(\alpha).$$

In particular, the map A is surjective.

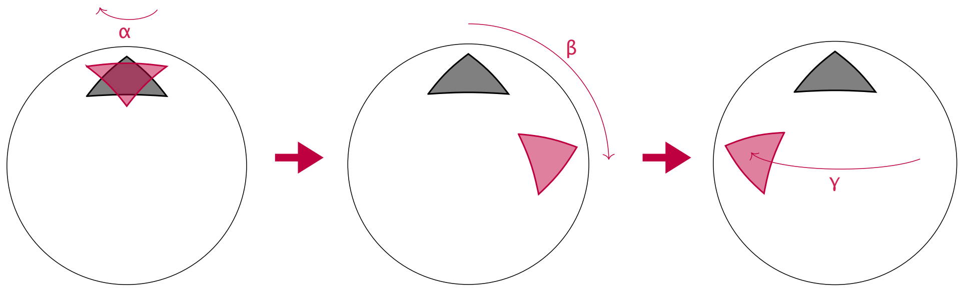


Figure 3.3: A decomposition of a rotation in SO(3) into three successive rotations, first about the z-axis by α, then about the x-axis by β, and finally about the z-axis by γ.

Proof. Let $R \in \mathrm{SO}(3)$ and denote the vector Re_3 by v. Since $v \in \mathbb{S}^2$, there exist $\beta, \gamma \in [-\pi, \pi)$ such that

$$v = \begin{pmatrix} \sin(\beta)\sin(\gamma) \\ \sin(\beta)\cos(\gamma) \\ \cos(\beta) \end{pmatrix} = R_Z(\gamma)R_X(\beta)e_3.$$

From

$$R_X(-\beta)R_Z(-\gamma)Re_3 = R_X(-\beta)R_Z(-\gamma)v = e_3$$

we deduce that the matrix $R_X(-\beta)R_Z(-\gamma)R$ leaves e_3 invariant. Thus, it acts as a rotation in the plane orthogonal to e_3. Hence, there exists an $\alpha \in [-\pi, \pi)$ such that $R_X(-\beta)R_Z(-\gamma)R = R_Z(\alpha)$. Rearranging yields $R = R_Z(\gamma)R_X(\beta)R_Z(\alpha)$. $\square$

The underlying idea extends to $\mathrm{SO}(d)$ for arbitrary $d \geq 3$ by successively eliminating coordinates with *Givens rotations*. For a background on Givens rotations, we refer to [42, Chapter 5.1].

Let $(\alpha, \beta, \gamma) \in [-\pi, \pi)^3$. Up to periodicity, there exist two distinct parameter triples representing the inverse of $A(\alpha, \beta, \gamma)$

$$A(\alpha, \beta, \gamma)^\top = A(-\gamma, -\beta, -\alpha) = A(\pi - \gamma, \beta, \pi - \alpha), \tag{3.8}$$

which can be verified by direct computation using standard trigonometric identities. Combining these identities yields

$$A(\alpha, \beta, \gamma) = A(\alpha - \pi, -\beta, \gamma - \pi). \tag{3.9}$$

Consequently, the map A can be restricted to the domain $[-\pi, \pi) \times [0, \pi] \times [-\pi, \pi)$ without loss of its surjectivity. When restricted to this domain, the parametrization becomes injective, except at the singular values $\beta = 0$ and $\beta = \pi$. The injectivity for $0 < \beta < \pi$ can be shown by comparing the entries of the matrices and exploiting the uniqueness of the trigonometric representation.

We refer to a triple $(\alpha, \beta, \gamma) \in [-\pi, \pi) \times [0, \pi] \times [-\pi, \pi)$ that satisfies $R = A(\alpha, \beta, \gamma)$ as a set of *Euler angles* associated with R.

3.3 Characterizing intersection

A central step in solving the packing problem defined by a spherical polytope $\mathcal{P} \subseteq \mathbb{S}^{d-1}$ is to characterize the set

$$\{A \in \mathrm{SO}(d) : (\mathcal{P} \cap A\mathcal{P})^\circ \neq \emptyset\}\,.$$

For general dimensions and general spherical polytopes, this problem appears to be difficult to settle.

In classical convex geometry, two convex bodies $\mathcal{Q}_1, \mathcal{Q}_2 \subseteq \mathbb{R}^d$ have disjoint interiors if and only if there exists a separating hyperplane, i.e., there exists $h \in \mathbb{R}^d \setminus \{0\}$ and $b \in \mathbb{R}$ such that

$$h^\top x \leq b \text{ for all } x \in \mathcal{Q}_1 \quad \text{and} \quad h^\top x \geq b \text{ for all } x \in \mathcal{Q}_2. \tag{3.10}$$

In the case of interior-disjoint polytopes in $\mathbb{R}^d$, a natural first guess is that a separating hyperplane is induced by the normal vector that defines a facet of one of the two polytopes. This would reduce the separation problem to a finite set of candidate hyperplanes determined by the facets. Example 3.2 shows that this reduction fails already in dimension three.

By contrast, in dimensions one and two, the same idea yields a finite intersection test based on the facet-defining hyperplanes, thereby giving a convenient algorithm for detecting intersections by checking all facet-defining hyperplanes. While we omit a formal proof of these cases, it can be obtained using the techniques discussed in this section.

Our aim is to transfer this approach to spherical polytopes on the two-dimensional sphere. Since a spherical polytope is represented as the intersection of $\mathbb{S}^2$ with a polyhedral cone, it is natural to formulate separation in terms of the underlying cones. Any hyperplane separating two convex cones must pass through the origin, so the analogue of (3.10) is obtained with $b = 0$. This suggests considering hyperplanes induced by facets of the associated polyhedral cones or, equivalently, by the facets of the corresponding spherical polytope. The following examples illustrate fundamental obstructions that arise when attempting to transfer this method to the spherical setting.

Example 3.2. Consider the two regular spherical triangles $\mathcal{G}_3(\varrho)$, with $\varrho = \arccos\left(\sqrt{2/3}\right)$ and its congruent copy $A(0,\pi,\pi/6)\,\mathcal{G}_3(\varrho)$. Since the second Euler angle equals π, the spherical polygons lie in opposite hemispheres and are therefore interior-disjoint. The inner products listed in Table 3.1 show that for every facet-defining hyperplane of $\mathcal{G}_3(\varrho)$ there exist two vertices of $A(0,\pi,\pi/6)\,\mathcal{G}_3(\varrho)$ that lie on opposite sides of that hyperplane and vice versa.

The convex hull of the three vertices of either spherical triangle together with the origin is a regular Euclidean simplex. The two simplices obtained by this construction are also interior-disjoint. Since the facet hyperplanes of these simplices coincide with the hyperplanes induced by the facets of the spherical polygons, the same obstruction also appears in the Euclidean setting.

This phenomenon arises whenever the two spherical polytopes are contained in opposite hemispheres. In these cases, the two copies are separated by the hyperplane corresponding to the equator, whose normal vector is given by the standard basis vector e_3. Consequently, when searching for separating hyperplanes, one must also consider this hyperplane in addition to those induced by the facets.

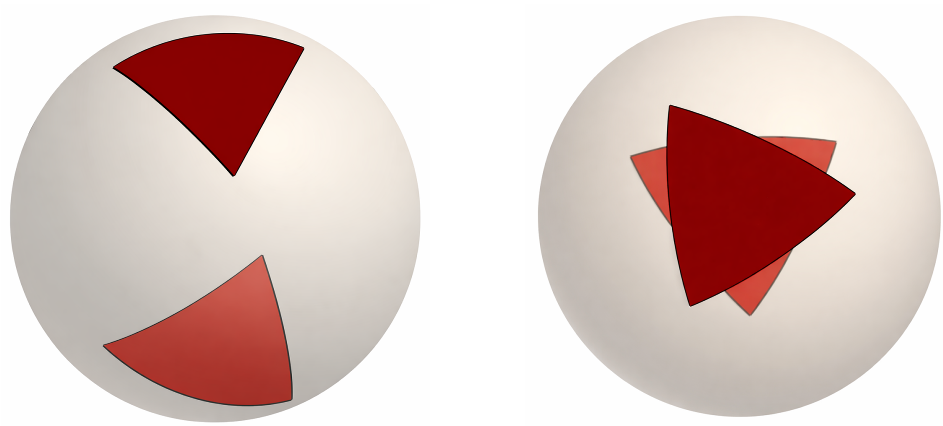


Figure 3.4: Two interior-disjoint regular spherical triangles for which none of the facet-defining hyperplanes separates the two polygons.

	v_1	v_2	v_3
h_1	-0.199239	0.743570	0.272166
h_2	0.743570	0.272166	-0.199239
h_3	0.272166	-0.199239	0.743570

Table 3.1: Inner products $h_k^\top A(0,\pi,\pi/6)\, v_\ell$, where v_1, v_2, v_3 denote the vertices of $\mathcal{G}_3(\varrho)$ and h_1, h_2, h_3 are the corresponding facet-defining normal vectors. By symmetry of the construction, these values coincide with $h_k^\top A(0,\pi,\pi/6)^\top v_\ell$.

Example 3.3. As the angular radius increases, regular spherical polygons occupy a larger region of $\mathbb{S}^2$. Consider the regular spherical triangles $\mathcal{G}_3(\varrho)$ and $A(\alpha,\beta,\gamma)\mathcal{G}_3(\varrho)$ with

$$\varrho = \arccos\left(\left(4-\sqrt{5}\right)/11\right) \quad \text{and} \quad (\alpha,\beta,\gamma) = (7/30\,\pi, 9/8\,\pi, 1/10\,\pi)\,.$$

The two copies are interior-disjoint, as illustrated in Figure 3.5, although none of their facets defines a separating hyperplane. Moreover, even the hyperplane corresponding to the equator fails to separate the two spherical polygons, see Tables 3.2 and 3.3 for numerical approximations of the corresponding inner products.

Among regular spherical polygons, this phenomenon occurs only in the triangular case. It is closely related to the magnitude of their inner angles, as becomes apparent in the subsequent analysis.

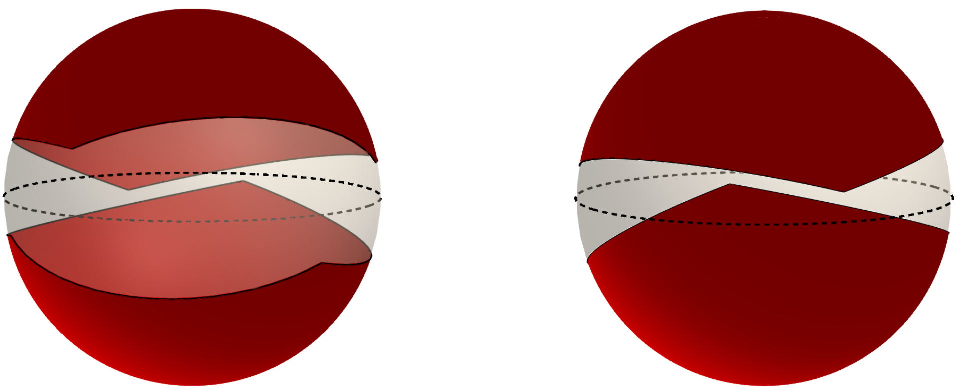


Figure 3.5: The two spherical polygons are interior-disjoint, although neither a facet-defining hyperplane nor the equator-induced hyperplane is separating. The dashed line indicates the equator.

A	v_1	v_2	v_3
h_1	0.161082	0.652073	−0.444560
h_2	0.499185	0.168635	−0.233293
h_3	−0.013494	0.402902	0.075569
e_3	−0.226686	−0.428860	0.211093

Table 3.2: Inner products $h_k^\top A(\alpha,\beta,\gamma)\, v_\ell$ and $e_3^\top A(\alpha,\beta,\gamma)\, v_\ell$, where v_1, v_2, v_3 are the vertices and h_1, h_2, h_3 are the facet-defining normals of $\mathcal{G}_3(\varrho)$.

$A^\top$	v_1	v_2	v_3
h_1	0.402902	0.075569	−0.013494
h_2	0.652073	−0.444560	0.161082
h_3	0.168635	−0.233293	0.499185
e_3	−0.428860	0.211093	−0.226686

Table 3.3: Inner products $h_k^\top A(\alpha,\beta,\gamma)^\top v_\ell$ and $e_3^\top A(\alpha,\beta,\gamma)^\top v_\ell$, where v_1, v_2, v_3 are the vertices and h_1, h_2, h_3 are the facet-defining normals of $\mathcal{G}_3(\varrho)$.

Let $\mathcal{K}$ be a polyhedral cone, and let $\mathcal{P} = \mathcal{K} \cap \mathbb{S}^2$ be the corresponding spherical polytope. Consider two congruent copies of $\mathcal{P}$ defined by rotations $A, B \in \mathrm{SO}(3)$. By the symmetry described in (3.6), we may assume $B = I$. The spherical polytopes $\mathcal{P}$ and $A\mathcal{P}$ intersect in their interiors if and only if the polyhedral cones $\mathcal{K}$ and $A\mathcal{K}$ intersect in their interiors. To determine whether $(\mathcal{K} \cap A\mathcal{K})^\circ$ is empty, we can use the dual cone $\mathcal{K}^*$. Every nonzero vector in $\mathcal{K}^* \cap -A\mathcal{K}^*$ defines the normal of a separating hyperplane passing through the origin. Normalizing such a vector to lie on $\mathbb{S}^2$ yields the equivalent condition

$$(\mathcal{P} \cap A\mathcal{P})^\circ = \emptyset \qquad \text{if and only if} \qquad \mathcal{P}^* \cap -A\mathcal{P}^* \neq \emptyset.$$

From this point onward, we restrict attention to spherical polytopes $\mathcal{P}$ that are *rotationally antipodal-symmetric*, i.e., there exists a rotation $R \in \mathrm{SO}(3)$ such that $-\mathcal{P} = R\mathcal{P}$. This property is preserved under duality. Assuming rotational antipodal-symmetry allows the separation condition to be reformulated as an intersection condition between $\mathcal{P}^*$ and a rotated copy of $\mathcal{P}^*$.

The facet-defining normals of $\mathcal{P}$ are precisely the vertices of the dual polytope $\mathcal{P}^*$. Consequently, checking whether a facet-defining hyperplane separates $\mathcal{P}$ and $A\mathcal{P}$ is equivalent to verifying whether the corresponding vertex belongs to $\mathcal{P}^* \cap AR\mathcal{P}^*$. The hyperplane corresponding to the equator with respect to a regular spherical polygon is represented by the barycenter of its dual.

As demonstrated in Example 3.3, it is not sufficient to examine only the vertices and the barycenter in order to determine whether two congruent spherical polygons intersect in their interior. However, there exists a class of spherical polygons for which this approach is sufficient.

The two-dimensional Euclidean analogue of this problem was studied by Mascarenhas and Birgin [55]. They introduced a finite set of so-called *sentinels*, which suffice to detect intersections of congruent copies obtained by Euclidean motions of a given polytope. As discussed in the introductory example of [55], not every polytope admits a finite set of sentinels. Two triangles can be made to intersect arbitrarily close to a vertex of each triangle, thereby avoiding any prescribed finite sentinel set.

A key obstruction is the presence of inner angles smaller than $\pi/2$. The techniques of Mascarenhas and Birgin require all inner angles to be at least $\pi/2$. In the Euclidean plane, no triangle satisfies this condition. In contrast, for regular spherical polygons $\mathcal{G}_N(\varrho)$, the inner angles tend to π as the angular radius ϱ increases.

In the following section, we adapt the results of Mascarenhas and Birgin to the case of spherical polygons. We will prove the following theorem, which will play a central role in the subsequent development.

Theorem 3.4. *Let either $N = 3$ and $\varrho \in (0, \arccos(\sqrt{5}/3)]$, or $N \geq 4$ and $\varrho \in (0, \pi/2)$. For any $A \in \mathrm{SO}(3)$, the regular spherical polygon $\mathcal{G}_N(\varrho)$ and its congruent copy $A\mathcal{G}_N(\varrho)$ have disjoint interiors if and only if at least one of the following holds*

i) a facet-defining hyperplane of the polygon $\mathcal{G}_N(\varrho)$ separates $\mathcal{G}_N(\varrho)$ from $A\mathcal{G}_N(\varrho)$, i.e., there exists a vertex h of $\mathcal{G}_N(\varrho)^*$ such that $h^\top Av \geq 0$ for all vertices v of $\mathcal{G}_N(\varrho)$;

ii) a facet-defining hyperplane of the polygon $A\mathcal{G}_N(\varrho)$ separates $\mathcal{G}_N(\varrho)$ from $A\mathcal{G}_N(\varrho)$, i.e., there exists a vertex h of $\mathcal{G}_N(\varrho)^*$ such that $(Ah)^\top v \geq 0$ for all vertices v of $\mathcal{G}_N(\varrho)$;

iii) the equator-induced hyperplane separates $\mathcal{G}_N(\varrho)$ from $A\mathcal{G}_N(\varrho)$, i.e., $e_3^\top Av \leq 0$ holds for all vertices v of $\mathcal{G}_N(\varrho)$.

A generalization of Theorem 3.4 can be established for a broader class of spherical polytopes, which we call *admissible*. In the next section, we provide a characterization of admissible spherical polytopes and establish a corresponding set of sentinels for the respective dual polygons.

3.3.1 Using sentinels to detect intersection

Throughout this section, we adopt the following conventions. The vertices $v_1, \dots, v_N$ of a given spherical polytope are indexed cyclically in such a way that consecutive vertices form a facet. Indices are understood modulo N, i.e., $v_k = v_\ell$ whenever $k \equiv \ell \pmod N$. The shortest side length of $\mathcal{P}$ is denoted by $\Delta(\mathcal{P}) = \min_{j=1,\dots,N} \delta(v_j, v_{j+1})$.

Definition 3.5. Let $\mathcal{P}$ be a spherical polytope. A finite set $S \subseteq \mathcal{P}$ is called a *sentinel arrangement* for $\mathcal{P}$ if

$$\mathcal{P} \cap A\mathcal{P} \neq \emptyset \quad \text{if and only if} \quad S \cap A\mathcal{P} \neq \emptyset \text{ or } AS \cap \mathcal{P} \neq \emptyset$$

for all $A \in \mathrm{SO}(3)$. The elements of such a set S are called *sentinels*.

As discussed in Section 3.3, the vertices of the dual polytope are precisely the normalized facet-defining normals of the original polytope. This allows the separation problem for two congruent copies of $\mathcal{P}$ to be reformulated as an intersection problem of the dual polytopes. In this section, we construct, under suitable geometric assumptions, a finite sentinel arrangement for $\mathcal{P}^*$ which detects these intersections.

The construction is motivated by the planar result of Mascarenhas and Birgin [55]. In their setting, the sentinel set consists of all vertices of the polygon together with boundary points that regulate the distance between consecutive sentinels and one additional interior point. This interior point ensures that any sufficiently large overlap which avoids the sentinels on the boundary is nevertheless detected by the interior sentinel.

We define the *forbidden region* of the spherical polytope $\mathcal{P}$ around its j-th vertex v_j by

$$\mathcal{F}_j(\mathcal{P}) = \{u \in \mathcal{P} : \text{there exist } x \in [v_{j-1}, v_j] \text{ and } y \in [v_j, v_{j+1}] \\ \text{with } \delta(x,y) < \Delta(\mathcal{P}) \text{ and } u \in [x,y]\}.$$

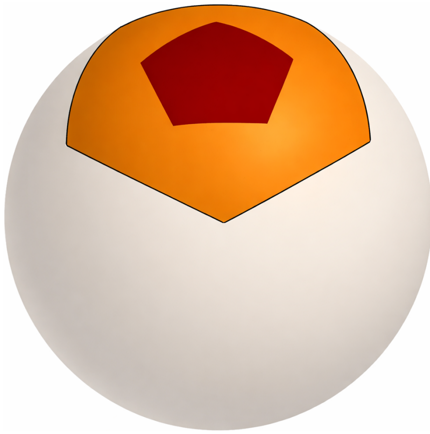


Figure 3.6: A regular spherical polygon together with its forbidden region $\mathcal{F}(\mathcal{P})$, shown in orange.

The *forbidden region* of $\mathcal{P}$ is defined as the union of the forbidden regions around all of its vertices

$$\mathcal{F}(\mathcal{P}) = \bigcup_{j=1}^{N} \mathcal{F}_j(\mathcal{P}).$$

The forbidden region consists of all points that can be reached by a geodesic segment connecting two edges incident to a common vertex, where the endpoints of the segment are at distance less than the minimum edge length of $\mathcal{P}$.

Lemma 3.6. *Let $\mathcal{P}$ be a spherical polytope satisfying either*

i) $N \geq 4$, all inner angles of $\mathcal{P}$ are at least $\pi/2$, and the longest side length is at most $\pi/2$;

ii) $\mathcal{P}$ is a regular spherical triangle.

Let $x \in [v_1, v_2]$ and $y \in [v_2, v_3]$ be points such that $x, y \neq v_2$ and $\delta(x, y) < \Delta(\mathcal{P})$. Then the spherical triangle

$$\mathcal{T} = \operatorname{cone}\{x, y, v_2\} \cap \mathbb{S}^2$$

is contained in the forbidden region around v_2.

Moreover, if $\mathcal{P}$ is a regular spherical triangle, then $\mathcal{P} \setminus \mathcal{F}(\mathcal{P})$ is spherically convex.

Proof. For $z = x, y$ we parametrize the geodesic arc $[z, v_2]$ by

$$z_t = \varphi_z(t)\, z + \varphi_z(1-t)\, v_2$$

where

$$\varphi_z(t) = \frac{\sin(\delta_z t)}{\sin(\delta_z)} \qquad \text{and} \qquad \delta_z = \delta(z, v_2).$$

Then $z_0 = v_2$, $z_1 = z$, $\|z_t\| = 1$, and $\delta(z_t, v_2) = \delta_z t$, so this parametrization is of constant speed. It is related to the parametrization in (3.1) by

$$\lambda = \frac{\sin(t\delta_z)}{\sin(t\delta_z) + \sin((1-t)\delta_z)}.$$

We claim

$$\mathcal{T} = \bigcup_{t\in[0,1]} [x_t, y_t]\,.$$

Indeed, the union of the geodesic arcs is included in the spherical triangle due to its spherical convexity.

Conversely, let $u \in \mathcal{T}$. Then there exist $\lambda_x, \lambda_y, \lambda_v \geq 0$ such that

$$u = \lambda_x x + \lambda_y y + \lambda_v v_2.$$

Since $\lambda_v = 0$ implies $u \in [x, y]$, we proceed under the assumption $\lambda_v \neq 0$.

If $u = v_2$, then $u = x_0 = y_0$, so the claim is immediate. Hence assume $u \neq v_2$. In this case we show that there exists $t \in (0, 1]$ such that x_t, y_t, u are coplanar. Equivalently,

$$\begin{aligned} 0 &= \det(x_t|y_t|u) \\ &= (\lambda_v\varphi_x(t)\varphi_y(t) - \lambda_x\varphi_x(1-t)\varphi_y(t) - \lambda_y\varphi_x(t)\varphi_y(1-t))\det(x|y|v_2). \end{aligned}$$

Since x, y, v_2 are linearly independent, we have

$$0 = \lambda_v\, \varphi_x(t)\varphi_y(t) - \lambda_x\, \varphi_x(1-t)\varphi_y(t) - \lambda_y\, \varphi_x(t)\varphi_y(1-t).$$

Rearranging terms yields

$$1 = \frac{\lambda_x\, \varphi_x(1-t)}{\lambda_v\, \varphi_x(t)} + \frac{\lambda_y\, \varphi_y(1-t)}{\lambda_v\, \varphi_y(t)}.$$

The function

$$t \mapsto \frac{\lambda_x\, \varphi_x(1-t)}{\lambda_v\, \varphi_x(t)} + \frac{\lambda_y\, \varphi_y(1-t)}{\lambda_v\, \varphi_y(t)}$$

is continuous on $(0,1]$, equals 0 at $t = 1$, and tends to infinity as $t \to 0^+$. By the intermediate value theorem, it attains the value 1 for some $t' \in (0,1)$. Now $u \in [x_{t'}, y_{t'}]$ follows from construction.

To prove that $\mathcal{T} \subseteq \mathcal{F}_2(\mathcal{P})$, it remains to show that

$$\delta(x_t, y_t) \leq \delta(x,y) < \Delta(\mathcal{P}) \qquad \text{for all } t \in [0,1].$$

Let α denote the inner angle of $\mathcal{P}$ at v_2. By the Lagrange identity (2.1)

$$\cos(\alpha) = \frac{(v_2 \times x_t)^\top (v_2 \times y_t)}{\|v_2 \times x_t\|\, \|v_2 \times y_t\|} = \frac{\cos(\delta(x_t,y_t)) - \cos(\delta_x t)\cos(\delta_y t)}{\sin(\delta_x t)\sin(\delta_y t)}.$$

Define the function f by

$$f(t) = \cos(\delta(x_t, y_t)) = \cos(\delta_x t)\cos(\delta_y t) + \sin(\delta_x t)\sin(\delta_y t)\cos(\alpha).$$

We first consider the case $N \geq 4$ and $\alpha \geq \pi/2$. By assumption $\delta_x, \delta_y \leq \pi/2$, thus f is nonincreasing on $[0,1]$. Therefore $\delta(x_t, y_t) \leq \delta(x,y) < \Delta(\mathcal{P})$ for all $t \in [0,1]$, and consequently $\mathcal{T} \subseteq \mathcal{F}_2(\mathcal{P})$.

It remains to treat the case where $\mathcal{P}$ is a regular spherical triangle. In this case, the dependence of f on the inner angle α can be made explicit using the Lagrange identity

$$\cos(\alpha) = \frac{(v_2 \times v_1)^\top (v_2 \times v_3)}{\|v_2 \times v_1\| \|v_2 \times v_3\|} = \frac{c}{1+c}$$

where $c = \cos(\Delta(\mathcal{P}))$. By applying elementary trigonometric identities, we obtain

$$f(t) = \frac{1}{2(1+c)}\cos((\delta_x + \delta_y)t) + \frac{1+2c}{2(1+c)}\cos((\delta_x - \delta_y)t).$$

Since the side length of a regular spherical triangle is at most $2\pi/3$, we have $c \geq -1/2$. Thus both coefficients in the last expression are nonnegative.

If $\delta_x + \delta_y \leq \pi$, the function f is monotonically decreasing on $[0,1]$. Suppose, for contradiction, that $\delta_x + \delta_y > \pi$. Now $\Delta(\mathcal{P}) \leq 2\pi/3$

implies $\delta_x + \delta_y \in (\pi, 4\pi/3)$. Since the cosine function is increasing on $(\pi, 4\pi/3)$, we get

$$\begin{aligned} c &< \cos(\delta(x,y)) \\ &= \frac{1}{2(1+c)}\cos(\delta_x+\delta_y) + \frac{1+2c}{2(1+c)}\cos(\delta_x-\delta_y) \\ &\le \frac{\cos(2\Delta(\mathcal{P}))}{2(1+c)} + \frac{1+2c}{2(1+c)} \\ &= \frac{(2c^2-1)+(2c+1)}{2(1+c)} \\ &= c. \end{aligned}$$

This contradiction shows that $\delta_x + \delta_y \le \pi$. Hence f is nonincreasing on $[0,1]$, and the arguments used in the previous case give $\mathcal{T} \subseteq \mathcal{F}_2(\mathcal{P})$.

It remains to prove that $\mathcal{P} \setminus \mathcal{F}(\mathcal{P})$ is spherically convex if $\mathcal{P}$ is a regular spherical triangle. By definition of the forbidden region, every point of $\mathcal{F}(\mathcal{P})$ lies in a spherical triangle of the form considered above.

For each such triangle $\mathcal{T}$, its complement in $\mathcal{P}$ can be written as

$$\mathcal{P} \setminus \mathcal{T} = \left\{u \in \mathcal{P} :\ h^\top u > 0\right\},$$

where h defines the facet of $\mathcal{T}$ containing x, y, with $h^\top v_2 < 0$. Using this half-space representation, we conclude that $\mathcal{P} \setminus \mathcal{T}$ is closed under taking geodesic arcs between pairs of points. Since $\mathcal{P} \setminus \mathcal{F}(\mathcal{P})$ is the intersection of these spherically convex complements, it is spherically convex as well. □

If only the first case of the lemma is considered, the proof becomes substantially simpler, see [27].

Lemma 3.7. *Let $\mathcal{P}$ be a spherical polytope with four or more vertices, and suppose that all inner angles of $\mathcal{P}$ are at least $\pi/2$. Then the geodesic distance between any two points lying on nonadjacent facets of $\mathcal{P}$ is at least $\Delta(\mathcal{P})$.*

Proof. It suffices to show that the $\Delta(\mathcal{P})$-neighborhood of any facet of $\mathcal{P}$, restricted to $\partial\mathcal{P}$, intersects only that facet and its two adjacent facets.

Let $[v_1, v_2]$ be a facet of $\mathcal{P}$ and let $h \in \mathcal{P}^*$ such that $h^\top v_1 = h^\top v_2 = 0$. Moreover, let $[v_1, u_1]$ and $[v_2, u_2]$ be the inward-pointing geodesic arcs through v_1 and v_2, respectively, which are orthogonal to the facet $[v_1, v_2]$ and satisfy $\delta(v_1, u_1) = \delta(v_2, u_2) = \Delta(\mathcal{P})$. Since $N \geq 4$ implies $\Delta(\mathcal{P}) \leq \pi/2$, and since the inner angles of $\mathcal{P}$ at v_1 and v_2 are at least $\pi/2$, the points u_1 and u_2 are contained in $\mathcal{P}$.

By spherical convexity, $\mathcal{P}$ contains the spherical quadrilateral

$$\mathcal{Q} := \operatorname{cone}\{v_1, v_2, u_1, u_2\} \cap \mathbb{S}^2.$$

Choose $n \in \mathbb{S}^2$ such that $n^\top u_1 = n^\top u_2 = 0$. Geometrically, the hyperplane defined by n is obtained by rotating the hyperplane defined by h towards the interior of $\mathcal{P}$ by an angle strictly greater than $\Delta(\mathcal{P})$. In particular, choosing a coordinate system such that $[v_1, v_2]$ is on the equator and the inward orthogonal geodesics are the meridians, a direct coordinate calculation shows that every point of $[u_1, u_2]$ has distance at least $\Delta(\mathcal{P})$ from $[v_1, v_2]$. Equality is attained at the endpoints u_1 and u_2.

Consequently, any point of $\mathcal{P}$ whose distance from $[v_1, v_2]$ is strictly smaller than $\Delta(\mathcal{P})$ cannot lie beyond the side $[u_1, u_2]$ of the quadrilateral $\mathcal{Q}$. Hence, inside $\mathcal{P}$, such a point either lies in $\mathcal{Q}$ or in the $\Delta(\mathcal{P})$-neighborhood of either v_1 or v_2. These neighborhoods can be covered by constructing analogous quadrilaterals for the adjacent edges.

Thus no point on a facet nonadjacent to $[v_1, v_2]$ can have distance smaller than $\Delta(\mathcal{P})$ from $[v_1, v_2]$. Since the facet $[v_1, v_2]$ was arbitrary, the claim follows. $\square$

Theorem 3.8. *Let $\mathcal{P}$ be a spherical polytope with at least four vertices such that all inner angles are at least $\pi/2$, the longest side length is at most $\pi/2$, and the set $\mathcal{P} \setminus \mathcal{F}(\mathcal{P})$ is nonempty. Let $S \subseteq \mathcal{P}$ be a finite set such that*

i) all vertices of $\mathcal{P}$ lie in S;

ii) S contains at least one element that is not in $\mathcal{F}(\mathcal{P})$;

iii) every geodesic subarc of a facet of $\mathcal{P}$ that contains no point of S has length less than $\Delta(\mathcal{P})$.

The set S is a sentinel arrangement for $\mathcal{P}$.

Proof. Let $\mathcal{P}$ be a spherical polytope, $S \subseteq \mathcal{P}$ as described above, and let $A \in \mathrm{SO}(3)$. Denote $A\mathcal{P}$ by $\mathcal{P}'$ and AS by S'. The vertices of $\mathcal{P}$ and

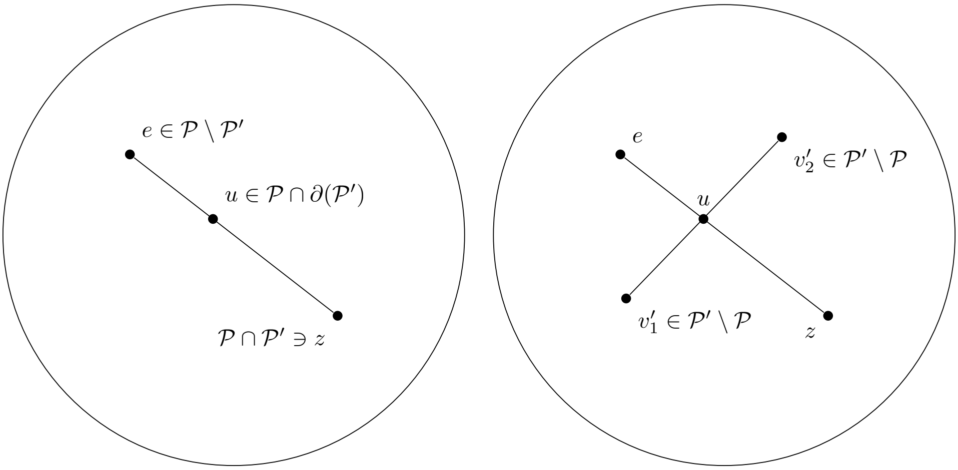


Figure 3.7: A conceptual image of the elements mentioned in the first two paragraphs of the proof of Theorem 3.8.

$\mathcal{P}'$ are denoted by $v_1, \ldots, v_N$ and $v_1', \ldots, v_N'$, respectively. Arguing by contradiction, assume $\mathcal{P} \cap \mathcal{P}' \neq \emptyset$ and

$$S \cap \mathcal{P}' = \emptyset \qquad \text{and} \qquad S' \cap \mathcal{P} = \emptyset.$$

Choose $z \in \mathcal{P} \cap \mathcal{P}'$ and $e \in S$ that does not lie in the forbidden region $\mathcal{F}(\mathcal{P})$. Since $e \notin \mathcal{P}'$, the geodesic arc $[e, z]$ crosses a face, say $[v_1', v_2']$, of $\mathcal{P}'$ at a point u different from v_1' and v_2'.

By assumption, $v_1', v_2' \notin \mathcal{P}$. Thus, there exist two facets of $\mathcal{P}$ each separating one of the points v_1' and v_2' from $\mathcal{P}$. These facets intersect the facet $[v_1', v_2']$ at the points x and y. The geodesic arc $[x, y]$ is contained in $\mathcal{P} \cap \mathcal{P}'$. Therefore it contains no point of S', which implies $\delta(x, y) < \Delta(\mathcal{P})$. Since the distance between the facets is less than $\Delta(\mathcal{P})$, Lemma 3.7 implies that they must share a vertex v_j, which by assumption is not contained in $\mathcal{P}'$. Assume $x \in [v_{j-1}, v_j]$ and $y \in [v_j, v_{j+1}]$.

The vertex v_j and the points x, y define a spherical triangle $\mathcal{T}$. Depending on which side of the facet $[v_1', v_2']$ the vertex v_j is located either $e \in \mathcal{T}$ or $z \in \mathcal{T}$ has to hold.

1. Case $e \in \mathcal{T}$: The distance between x and y is less than $\Delta(\mathcal{P})$. Moreover, $N \geq 4$ implies $\Delta(\mathcal{P}) \leq \pi/2$. Hence, the first part of Lemma 3.6 implies $e \in \mathcal{F}(\mathcal{P})$, which contradicts the assumption.

2. Case $z \in \mathcal{T}$: Recall that the vertex v_j is not contained in $\mathcal{P}'$. Therefore, there exists a facet $[v_k', v_{k+1}']$ that intersects the geodesic

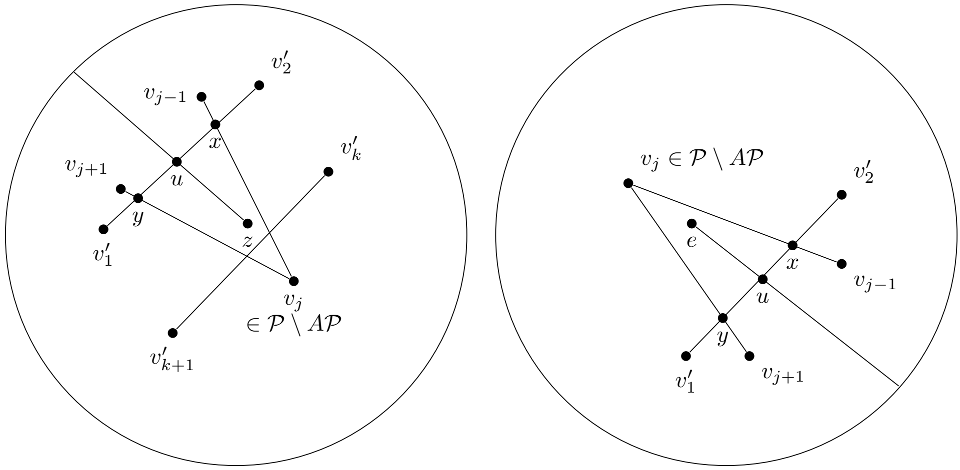


Figure 3.8: The two cases that need to be distinguished depending on which side of the facet $[v'_1, v'_2]$ the vertex v_j is located. Note that the inner angles are not accurately depicted.

arc $[z, v_j]$. Moreover, the facet $[v'_k, v'_{k+1}]$ intersects both $[v_{j-1}, v_j]$ and $[v_j, v_{j+1}]$. The geodesic arc on the facet $[v_j, v_{j+1}]$ between its intersections with $[v'_1, v'_2]$ and $[v'_k, v'_{k+1}]$ is contained in both $\mathcal{P}$ and $\mathcal{P}'$, so it does not contain an element of S. Consequently, the distance between $[v'_1, v'_2]$ and $[v'_k, v'_{k+1}]$ is less than $\Delta(\mathcal{P})$. By Lemma 3.7, the facets $[v'_1, v'_2]$ and $[v'_k, v'_{k+1}]$ are adjacent. Without loss of generality, we assume $k = 2$.

We choose a coordinate system such that $x = (0, 0, 1)$ and the facet-defining normal vector of $[v'_1, v'_2]$ equals $h = (0, -1, 0)$. The facets $[v'_1, v'_2]$ and $[v'_2, v'_3]$ form an angle of at least $\pi/2$. For $[v'_2, v'_3]$ to intersect $[v_{j-1}, v_j]$, the angle between $[x, v'_1]$ and $[x, v_j]$ has to be greater than $\pi/2$. Thus, there exists $\beta > 0$ such that one can move from x in the direction perpendicular to $[v'_1, v'_2]$ while staying in the triangle $\mathcal{T}$, until one reaches a point x' on the boundary of $\mathcal{T}$. This movement is induced by rotating x along the axis $(1, 0, 0)$ i.e., $x' = R_X(\beta)x$.

From there, we continue moving parallel to $[x, y]$ until we reach that geodesic arc at the point y'. Since the inner angle of $\mathcal{T}$ at v_j is at least $\pi/2$, we stay inside $\mathcal{T}$.

By construction, y' is one of the two points where the hyperplanes defined by h and $R_X(\beta)h$ intersect. An explicit computation shows $y' = (\pm 1, 0, 0)$. Hence, the angular distance between x and y' equals

$\pi/2$. From $y' \in [x, y]$ we obtain

$$\Delta(\mathcal{P}) > \delta(x, y) \geq \delta(x, y') = \pi/2,$$

which contradicts $\Delta(\mathcal{P}) \leq \pi/2$. $\square$

Observe that the proof only uses the existence of a single sentinel in $\mathcal{P} \setminus \mathcal{F}(\mathcal{P})$ for the unrotated copy. It does not require a corresponding point in $A\mathcal{P} \setminus \mathcal{F}(A\mathcal{P})$. Thus, for detecting intersections between two congruent copies, it suffices to place such an additional sentinel on one of the two copies, provided the vertices and facet sentinels are included symmetrically as above.

Consider a regular spherical polygon $\mathcal{G}_N(\varrho)$ with $N \geq 4$ and $\varrho \in (0, \pi/2)$. By (3.5), its side length is at most $\pi/2$. Its inner angle can be obtained from its dual by $\pi - \Delta(\mathcal{G}_N(\varrho)^*)$. Thus, it is at least $\pi/2$.

In the next section, we show that the complement of the forbidden region inside $\mathcal{G}_N(\varrho)$ contains the barycenter. Combined with Theorem 3.8, this will imply that for every regular spherical polygon with at least four vertices, the set consisting of the vertices together with the barycenter forms a sentinel arrangement.

The family of regular spherical triangles $\mathcal{G}_3(\varrho)$ with $\varrho \in (0, \pi/2)$ is not covered by Theorem 3.8. Depending on the angular radius ϱ, the side length of $\mathcal{G}_3(\varrho)$ may attain values up to $2\pi/3$. The argument used in Theorem 3.8 relies on the bound $\Delta(\mathcal{P}) \leq \pi/2$, which fails in general for $\mathcal{G}_3(\varrho)$. Therefore, the proof has to be modified in the triangular case.

Theorem 3.9. *Let $\mathcal{G}_3(\varrho)$ be a regular spherical triangle of angular radius $\varrho \in (0, \pi/2)$ whose inner angle is at least $\pi/2$, and suppose that the complement of the forbidden region inside $\mathcal{G}_3(\varrho)$ is nonempty. Every finite set $S \subseteq \mathcal{G}_3(\varrho)$ containing all vertices and at least one point of $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ is a sentinel arrangement.*

Proof. Arguing as in the proof of Theorem 3.8, assume for contradiction that $\mathcal{P} \cap A\mathcal{P} \neq \emptyset$, while $S \cap A\mathcal{P} = \emptyset$ and $AS \cap \mathcal{P} = \emptyset$. Choose $e \in S \setminus \mathcal{F}(\mathcal{P})$ and $z \in \mathcal{P} \cap A\mathcal{P}$. The same construction yields a spherical triangle

$$\mathcal{T} = \operatorname{cone}\{x, y, v_j\} \cap \mathbb{S}^2$$

with x, y on the two facets incident to v_j, and the proof splits into the two cases $e \in \mathcal{T}$ and $z \in \mathcal{T}$.

Since $\mathcal{G}_3(\varrho)$ satisfies the second condition of Lemma 3.6, the case $e \in \mathcal{T}$ is immediate.

Assume that $z \in \mathcal{T}$. To derive a contradiction analogous to the one obtained in the proof of Theorem 3.8, we make explicit the dependence of the side length of $\mathcal{G}_3(\varrho)$ and the inner angle α on the angular radius ϱ.

Denote $\cos(\varrho)$ by t and $\Delta(\mathcal{G}_3(\varrho))$ by Δ. Equation (3.5) yields

$$\cos(\Delta) = (3t^2 - 1)/2.$$

Using the Lagrange identity (2.1) to compute the inner angle gives

$$\cos(\alpha) = \frac{\cos(\Delta)}{1 + \cos(\Delta)} = \frac{3t^2 - 1}{3t^2 + 1}.$$

Choose the vertices v_1', v_2', v_3' as well as x, y and h as in the proof of Theorem 3.8.

As in the proof of Theorem 3.8, we start at x and follow the geodesic orthogonal to the facet $[v_1', v_2']$ for a distance of β until we reach the boundary of $\mathcal{T}$. In contrast to the preceding proof, where the direction is changed by an angle of $\pi/2$, we change direction by the inner angle α and proceed along the corresponding arc until it meets the facet $[x, y]$ at y'. The choice of rotation angle ensures that this arc cannot leave $\mathcal{T}$ through the facet $[v_j, y]$. We obtain the desired contradiction by verifying $\delta(x, y') \geq \Delta$.

By construction, y' is one of the two antipodal intersection points of the hyperplanes associated with h and n, where

$$n = R_X(\beta)\, R_Z(\alpha - \pi/2)\, h = \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha)\cos(\beta) \\ \sin(\alpha)\sin(\beta) \end{pmatrix}.$$

Up to sign, y' is given by

$$y' = \left(n_1^2 + n_3^2\right)^{-1/2} \begin{pmatrix} n_3 \\ 0 \\ -n_1 \end{pmatrix},$$

which allows us to compute the absolute value of the inner product of y' with x

$$\left|x^\top y'\right| = \sqrt{\frac{n_1^2}{(n_1^2 + n_3^2)}}.$$

To determine the sign of the inner product, recall from the previous proof that if $\alpha = \pi/2$, the inner product of x and y' must be equal to zero. As α increases, the inner product decreases, so it has to be nonpositive.

An explicit computation of $x^\top y'$ yields

$$x^\top y' = -\sqrt{\frac{(3t^2-1)^2}{(3t^2-1)^2 + 12t^2s^2}},$$

where we denote $\sin(\beta)$ by s.

Now, $\cos(\Delta) \geq x^\top y'$ if and only if

$$\frac{3t^2-1}{2} \geq -\sqrt{\frac{(3t^2-1)^2}{(3t^2-1)^2 + 12t^2s^2}}.$$

For $t \in [1/\sqrt{3}, 1]$ the inequality is straightforward, so we continue assuming $t \in [0, 1/\sqrt{3})$. Elementary manipulations yield

$$0 \leq -4t^2s^2 - 3t^4 + 2t^2 + 1.$$

Using $s^2 \leq 1$ gives

$$-4t^2s^2 - 3t^4 + 2t^2 + 1 \geq -3t^4 - 2t^2 + 1 = (1-3t^2)(t^2+1)$$

which is nonnegative for all $t \in [0, 1/\sqrt{3})$. Therefore, $\delta(x, y') \geq \Delta$, yielding the desired contradiction. □

In the next section, we characterize the regular spherical triangles that satisfy the assumptions of Theorem 3.9.

Theorems 3.8 and 3.9 identify a class of spherical polytopes for which intersections of congruent copies can be detected using only finitely many test points. Consequently, only finitely many candidates need to be considered when looking for separating hyperplanes for congruent copies of their respective dual polytopes.

Definition 3.10. Let $\mathcal{P}$ be a spherical polytope and let $\mathcal{P}^*$ be its dual polytope. Denote the vertices of $\mathcal{P}$ and $\mathcal{P}^*$ by $v_1, \ldots, v_N$ and $h_1, \ldots, h_N$, respectively. We refer to $\mathcal{P}$ as admissible if

i) $\mathcal{P}$ is rotationally antipodal-symmetric, i.e., there exists $A \in \mathrm{SO}(3)$ such that $-\mathcal{P} = A\mathcal{P}$;

ii) all inner angles of $\mathcal{P}^*$ are at least $\pi/2$, i.e., $v_j^\top v_{j+1} \geq 0$ for all $j = 1, \ldots, N$;

iii) the set $\mathcal{P}^* \setminus \mathcal{F}(\mathcal{P}^*)$ is nonempty;

iv) either $N \geq 4$ and the longest side length of $\mathcal{P}^*$ is at most $\pi/2$, i.e., $h_j^\top h_{j+1} \geq 0$, or $\mathcal{P}^*$ is a regular spherical triangle.

Although our techniques apply to all admissible spherical polytopes, in the remainder of this thesis we focus primarily on the subclass of admissible regular spherical polygons. Their symmetry leads to optimization problems that remain computationally tractable.

3.3.2 Proof of the intersection theorem

The proof of Theorem 3.4 requires a characterization of admissible regular spherical polygons. While every regular spherical polygon with at least four vertices is admissible, regular spherical triangles are admissible precisely up to a threshold on their angular radius.

A natural candidate for proving that the complement of the forbidden region is nonempty is the barycenter of the polygon under consideration. For the dual of a regular spherical polygon, this point is given by $-e_3$ due to the sign convention used in the definition of dual cones.

Lemma 3.11. *Let $\varrho \in (0, \pi/2)$ and $N \geq 4$. The regular spherical polygon $\mathcal{G}_N(\varrho)$ is admissible. Moreover, the barycenter $-e_3$ lies in $\mathcal{G}_N(\varrho)^* \setminus \mathcal{F}(\mathcal{G}_N(\varrho)^*)$.*

Proof. The symmetries of a regular spherical polygon allow for an immediate verification of its rotational antipodal-symmetry.

By (3.5), the distances between adjacent vertices of both $\mathcal{G}_N(\varrho)$ and $\mathcal{G}_N(\varrho)^*$ are bounded above by $\pi/2$. Hence, the corresponding inner products of adjacent vertices are nonnegative.

Thus, showing that $-e_3$ does not lie in the forbidden region of $\mathcal{G}_N(\varrho)^*$ verifies admissibility.

Denote the vertices of $\mathcal{G}_N(\varrho)^*$ by $h_1, \ldots, h_N$. The forbidden region of the regular spherical polygon around the j-th vertex is contained in the spherical triangle defined by the vertices h_{j-1}, h_j, h_{j+1}. If $N \geq 5$, the barycenter is not contained in any of these triangles, so it does not lie in the forbidden region.

In the $N = 4$ case, the barycenter lies on the facet of the corresponding spherical triangle that is given by the geodesic arc joining the two opposite vertices h_{j-1}, h_{j+1}. Equation (3.5) shows that their angular distance is at least $\Delta(\mathcal{G}_4(\varrho)^*)$. Thus, the point $-e_3$ does not lie in the forbidden region. $\square$

Deciding whether the complement of the forbidden region inside a regular spherical triangle is nonempty requires a more delicate argument. As in the previous case, the barycenter is an apparent choice. However, it is a priori unclear how to determine the shortest geodesic arc joining two adjacent facets and passing through the barycenter.

Lemma 3.12. *Let $\mathcal{G}_3(\varrho)$ be a spherical triangle of angular radius $\varrho \in (0, \pi/2)$. The complement $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ is nonempty if and only if the angular side length is at least* $\arccos(-1/4)$*. Moreover, the latter condition implies that the barycenter e_3 lies in $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$.*

Proof. We first show that if $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ is nonempty, then $e_3 \in \mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$. Recall from Lemma 3.6 that the set $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ is spherically convex. Since $\mathcal{G}_3(\varrho)$ is regular, $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ is invariant under the cyclic group generated by $R_Z(2\pi/3)$. Hence, if $\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$ contains a point, it also contains its orbit under the corresponding cyclic group action. Averaging and normalizing yields a fixed point of this action in $\mathcal{G}_3(\varrho)$, namely e_3. By spherical convexity, $e_3 \in \mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$. Since the converse implication is immediate, we obtain

$$\mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho)) \neq \emptyset \qquad \text{if and only if} \qquad e_3 \in \mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho)).$$

To determine whether $e_3 \in \mathcal{G}_3(\varrho) \setminus \mathcal{F}(\mathcal{G}_3(\varrho))$, consider two adjacent facets $[v_1, v_2]$ and $[v_2, v_3]$. We seek points on these facets such that the geodesic arc joining them passes through e_3 and has minimal length. To this end, we parametrize points on the Euclidean line segments between the vertices by

$$\begin{aligned} x(t) &= t\, v_1 + (1-t)\, v_2, \\ y(s) &= s\, v_2 + (1-s)\, v_3 \end{aligned}$$

for $s, t \in [0, 1]$.

The desired points are obtained by maximizing the objective function

$$\frac{x(t)^\top y(s)}{\|x(t)\|\,\|y(s)\|},$$

subject to the constraint that e_3, $x(t)$, and $y(s)$ are coplanar. Since e_3 is a scalar multiple of $v_1 + v_2 + v_3$, the coplanarity constraint can be stated as

$$\begin{aligned} 0 &= \det(x(t)|y(s)|v_1 + v_2 + v_3) \\ &= (3st - s - 2t + 1)\det(v_1|v_2|v_3). \end{aligned}$$

From the linear independence of v_1, v_2, and v_3 we conclude $3st - 2t - s + 1 = 0$, which is solved by $s = \frac{2t-1}{3t-1}$. Observe that $s \in [0,1]$ implies $t \in [1/2, 1]$ or $t = 0$. The latter gives $x(0) = y(1) = v_2$ and therefore can be disregarded. We continue by restricting t to the interval $[1/2, 1]$. To analyze the objective function, we define

$$\begin{aligned} f(t) &= x(t)^\top y\left(\tfrac{2t-1}{3t-1}\right) = \frac{2t^2(c-1)}{3t-1} + 1, \\ g(t) &= x(t)^\top x(t) = 2t(1-c)(t-1) + 1 \end{aligned}$$

where $c \in (-1/2, 1)$ denotes the cosine of the side length of $\mathcal{G}_3(\varrho)$. By symmetry, $y(s)^\top y(s) = g(s)$. Hence the objective function becomes

$$\frac{f(t)}{\sqrt{g(t)\, g\left(\frac{2t-1}{3t-1}\right)}}. \tag{3.11}$$

After differentiation with respect to t and clearing the strictly positive denominator, the numerator is

$$2t(c-1)(2c+1)(3t-2)(2t^2c + 4t^2 - 3t + 1).$$

It follows from $2t^2c + 4t^2 - 3t + 1 \geq 3t^2 - 3t + 1 > 0$ that the derivative vanishes only at $t = 2/3$. Examining the sign of the derivative shows that (3.11) attains its maximum on $[1/2, 1]$ at $t = 2/3$ and the corresponding value is

$$\frac{8c+1}{4c+5}.$$

Therefore every geodesic segment connecting the two adjacent facets and passing through e_3 has length at least Δ if and only if

$$c \geq \frac{8c+1}{4c+5},$$

or equivalently $\cos(\Delta) = c \leq -1/4$. $\square$

The following lemma identifies exactly those angular radii for which the admissibility conditions are satisfied.

Lemma 3.13. *A regular spherical triangle $\mathcal{G}_3(\varrho)$ is admissible if and only if* $0 < \varrho \leq \arccos\left(\sqrt{5}/3\right)$.

Proof. As in the case $N \geq 4$, the rotational antipodal-symmetry of $\mathcal{G}_3(\varrho)$ follows from its symmetry. Thus, to show that the regular spherical triangle $\mathcal{G}_3(\varrho)$ is admissible, we verify that $\mathcal{G}_3(\varrho)^* \setminus \mathcal{F}(\mathcal{G}_3(\varrho)^*) \neq \emptyset$ and that the inner angle of the dual polygon $\mathcal{G}_3(\varrho)^*$ is at least $\pi/2$.

Let Δ denote the angular side length of $\mathcal{G}_3(\varrho)$. Using (3.5), we obtain

$$\cos(\Delta) = (3\cos(\varrho)^2 - 1)/2.$$

The cosine of the side length of the dual spherical polygon can be computed by

$$\frac{(v_1 \times v_2)^\top (v_2 \times v_3)}{\|v_1 \times v_2\| \|v_2 \times v_3\|} = -\frac{\cos(\Delta)}{1 + \cos(\Delta)} = \frac{1 - 3\cos(\varrho)^2}{1 + 3\cos(\varrho)^2},$$

where the second equality follows from the Lagrange identity (2.1).

By Lemma 3.12, the set $\mathcal{G}_3(\varrho)^* \setminus \mathcal{F}(\mathcal{G}_3(\varrho)^*)$ is nonempty if and only if the side length is at least $\arccos(-1/4)$. Expressing this condition in terms of the angular radius yields $\varrho \leq \arccos\left(\sqrt{5}/3\right)$.

It remains to verify that this also implies the condition on the inner angles required in the definition of admissibility. The inner angle of the dual polygon $\mathcal{G}_3(\varrho)^*$ is at least $\pi/2$ if and only if the inner products of consecutive facet-defining normal vectors are nonnegative. These vectors coincide with the vertices of $\mathcal{G}_3(\varrho)$. By (3.5), this condition is equivalent to

$$\varrho \leq \arccos\left(1/\sqrt{3}\right).$$

Consequently, a regular spherical triangle $\mathcal{G}_3(\varrho)$ is admissible if and only if $\varrho \leq \arccos\left(\sqrt{5}/3\right)$. □

We now prove Theorem 3.4 by combining the main ideas developed in this chapter.

Proof of Theorem 3.4. Let $A \in \mathrm{SO}(3)$. Assume that either $N \geq 4$ and $\varrho \in (0, \pi/2)$, or $N = 3$ and $\varrho \in \left(0, \arccos\left(\sqrt{5}/3\right)\right]$. The regular spherical polygons $\mathcal{G}_N(\varrho)$ and $A\mathcal{G}_N(\varrho)$ have disjoint interiors if and only if their dual polygons satisfy $\mathcal{G}_N(\varrho)^* \cap -A\mathcal{G}_N(\varrho)^* \neq \emptyset$. Since regular spherical polygons are rotationally antipodal-symmetric, there exists $B \in \mathrm{SO}(3)$ such that $-A\mathcal{G}_N(\varrho)^* = B\mathcal{G}_N(\varrho)^*$.

By Lemmas 3.11 and 3.13, the spherical polygon $\mathcal{G}_N(\varrho)$ is admissible under the present assumptions. Hence, the set consisting of the vertices of $\mathcal{G}_N(\varrho)^*$ together with the barycenter $-e_3$ satisfies the conditions of Theorem 3.8 and Theorem 3.9, respectively. Consequently, the set forms a sentinel arrangement for $\mathcal{G}_N(\varrho)^*$.

Therefore, $\mathcal{G}_N(\varrho)^* \cap B\mathcal{G}_N(\varrho)^* \neq \emptyset$ if and only if the intersection contains either the barycenter or a vertex of one of the two polygons. As remarked at the end of the proof of Theorem 3.8, it is not necessary to test whether the barycenter of $B\mathcal{G}_N(\varrho)^*$ lies in $\mathcal{G}_N(\varrho)^*$.

If the intersection contains a vertex h of $\mathcal{G}_N(\varrho)^*$, then $h^\top Av \geq 0$ holds for all vertices Av of $A\mathcal{G}_N(\varrho)$. If it contains a vertex $-Ah$ of $-A\mathcal{G}_N(\varrho)^*$, then $h^\top A^\top v \geq 0$ must hold for all vertices v of $\mathcal{G}_N(\varrho)$. Finally, if the barycenter $-e_3$ of $\mathcal{G}_N(\varrho)^*$ lies in the intersection, then $e_3^\top Av \leq 0$ holds for all vertices Av of $A\mathcal{G}_N(\varrho)$. □

We conclude the chapter with an alternative version of Theorem 3.4, in which the third separation condition is replaced by a more directly verifiable one. This reformulation will allow us to relate the packing problem for regular spherical polygons of angular radius ϱ to the corresponding packing problem for spherical caps of the same radius.

Theorem 3.14 (Intersection theorem for regular spherical polygons)**.** *Let either $N = 3$ and $\varrho \in \left(0, \arccos\left(\sqrt{5}/3\right)\right]$, or $N \geq 4$ and $\varrho \in (0, \pi/2)$. For any $A \in \mathrm{SO}(3)$ with associated Euler angles (α, β, γ), the regular spherical polygon $\mathcal{G}_N(\varrho)$ and its congruent copy $A\mathcal{G}_N(\varrho)$ have disjoint interiors if and only if at least one of the following conditions is satisfied:*

- *i) a facet-defining hyperplane of the polygon $\mathcal{G}_N(\varrho)$ separates $\mathcal{G}_N(\varrho)$ from $A\mathcal{G}_N(\varrho)$, i.e., there exists a vertex h of $\mathcal{G}_N(\varrho)^*$ such that $h^\top Av \geq 0$ for all vertices v of $\mathcal{G}_N(\varrho)$;*
- *ii) a facet-defining hyperplane of the polygon $A\mathcal{G}_N(\varrho)$ separates $\mathcal{G}_N(\varrho)$ from $A\mathcal{G}_N(\varrho)$, i.e., there exists a vertex h of $\mathcal{G}_N(\varrho)^*$ such that $(Ah)^\top v \geq 0$ for all vertices v of $\mathcal{G}_N(\varrho)$;*
- *iii) the spherical caps of angular radius ϱ centered at e_3 and Ae_3 have disjoint interiors, i.e., $\cos(\beta) \leq \cos(2\varrho)$.*

Proof. Let N and ϱ satisfy the assumptions of the theorem and let $A \in \mathrm{SO}(3)$ with associated Euler angles (α, β, γ). If the first or the second condition holds, the corresponding hyperplane separates the two polygons, and hence their interiors are disjoint. If the circumscribed spherical caps have disjoint interiors, the same must hold for $\mathcal{G}_N(\varrho)$ and $A\mathcal{G}_N(\varrho)$.

For the converse implication, it remains to show that the third condition of Theorem 3.4 implies that the circumscribed spherical caps have disjoint interiors.

Let φ denote the angular radius of the inscribed spherical cap of $\mathcal{G}_N(\varrho)$, i.e., the spherical distance between the barycenter e_3 and the midpoint of a facet. Using (3.5), we obtain

$$\cos(\varphi) = \frac{e_3^\top(v_1 + v_2)}{\|v_1 + v_2\|} = \sqrt{\frac{2t^2}{(1-c)t^2 + c + 1}},$$

where we set $t = \cos(\varrho)$ and $c = \cos(2\pi/N)$.

The third condition of Theorem 3.4 implies that $\mathcal{G}_N(\varrho)$ and the inscribed spherical cap of $A\mathcal{G}_N(\varrho)$ are separated by the hyperplane induced by the equator, which implies $\pi/2 + \varphi \leq \delta(e_3, Ae_3) = \beta$. Therefore, it suffices to verify that $2\varrho \leq \pi/2 + \varphi$, or equivalently

$$\cos(2\varrho) \geq \cos(\pi/2 + \varphi).$$

Now, $\cos(2\varrho) = 2t^2 - 1$, whereas

$$\cos\left(\frac{\pi}{2} + \varphi\right) = -\sin(\varphi) = -\sqrt{1 - \cos(\varphi)^2} = -\sqrt{\frac{(1-t^2)(1+c)}{(1-c)t^2 + c + 1}}.$$

If $t \in \left[1/\sqrt{2}, 1\right]$, then $2t^2 - 1 \geq 0$, while $\cos(\pi/2 + \varphi) \leq 0$, so the desired inequality is immediate.

It remains to consider the case $t \in \left[0, 1/\sqrt{2}\right)$. In this range, the inequality $\cos(2\varrho) \geq \cos(\pi/2 + \varphi)$ can be expressed as

$$1 - 2t^2 \leq \sqrt{\frac{(1-t^2)(1+c)}{(1-c)t^2 + c + 1}}.$$

Squaring both sides and simplifying yields

$$4ct^2(t^2 - 1)^2 + 2t^2(1 - 2t^4) \geq 0.$$

Since $c \in [-1/2, 1)$, the left-hand side is bounded from below by

$$4ct^2(t^2-1)^2 + 2t^2(1-2t^4) \geq 2t^4(2-3t^2),$$

which is nonnegative for $t \in \left[0, 1/\sqrt{2}\right]$. □

It remains open whether the bound on the angular radius in Theorem 3.14 is optimal for regular spherical triangles. The obstruction discussed in Example 3.3 suggests that the inner angle of the dual polygon should be at least $\pi/2$. For regular spherical triangles, this condition is equivalent to

$$\varrho \leq \arccos\left(1/\sqrt{3}\right).$$

This is a strictly larger range than the one covered by the admissibility argument above. Numerical experiments suggest that the theorem may remain valid throughout this full range.

Conjecture 3.15. *Theorem 3.14 remains valid for every regular spherical triangle of angular radius at most* $\arccos\left(1/\sqrt{3}\right)$.

Part II

Optimization-based constructions of polygon packings

CHAPTER FOUR

Optimization over the set of packings

Optimal solutions of the spherical cap packing problem tend to be highly symmetric. The high degree of symmetry inherent to spherical caps often gives rise to configurations with well-defined geometric structure.

In contrast, optimal packings of objects with less symmetry are significantly more irregular. While certain spherical polytopes admit highly regular arrangements that cover a substantial portion of the sphere, the construction of rigorous and provably dense packings for general spherical polytopes remains a challenging problem.

So far, only a limited number of explicit constructions for dense packings of regular spherical polygons are known. In a series of works, Tarnai and Gáspár described packings of regular spherical pentagons with cardinalities ranging from 3 to 12, as well as 24, 36, and 72 [73, 75, 76].

In this chapter, we introduce a novel approach for computing packings consisting of n spherical polygons by solving a constrained optimization problem defined on the Riemannian submanifold SO(3). The objective is to determine the largest angular radius for which a valid packing of regular spherical polygons exists. We will explain how to model this problem and how to compute stationary points numerically. To transform the output of these computations into rigorously defined packings, we develop a rounding procedure that projects the numerical output onto the field of rational numbers.

4.1 The Riemannian structure of packing configurations

The problem of finding an optimal packing of exactly $n \geq 3$ regular spherical polygons with $N \geq 3$ vertices amounts to maximizing the angular radius ϱ under the constraint that there exist n rotation matrices that define a packing. By exploiting the rotational invariance of the problem, we may assume without loss of generality that there exists an optimal configuration that contains the identity, which leads to the following optimization problem

$$\begin{array}{ll} \max & \varrho \\ \text{s.t.} & \varrho \in (0, \pi/2), \\ & A_1, \ldots, A_{n-1} \in \mathrm{SO}(3), \\ & (\mathcal{G}_N(\varrho) \cap A_i \, \mathcal{G}_N(\varrho))^\circ = \emptyset \quad \text{for } 1 \leq i \leq n-1, \\ & (A_i \, \mathcal{G}_N(\varrho) \cap A_j \, \mathcal{G}_N(\varrho))^\circ = \emptyset \quad \text{for } 1 \leq i < j \leq n-1. \end{array} \tag{4.1}$$

By Theorem 3.14, the interior-disjointness condition for two congruent copies of an admissible regular spherical polygon $\mathcal{G}_N(\varrho)$ can be expressed in terms of finitely many inequalities involving the functions

$$\begin{aligned} h_c &\colon (0, \pi/2) \times \mathrm{SO}(3) \to \mathbb{R}, \quad (\varrho, A) \mapsto \cos(2\varrho) - A_{3,3}, \\ h_{k,\ell} &\colon (0, \pi/2) \times \mathrm{SO}(3) \to \mathbb{R}, \quad (\varrho, A) \mapsto h_k(\varrho)^\top A \, v_\ell(\varrho), \end{aligned}$$

where v_ℓ and h_k are defined as in (3.2) and (3.4), respectively.

For $A, B \in \mathrm{SO}(3)$ and $\varrho \in (0, \pi/2)$, the interiors of $A \, \mathcal{G}_N(\varrho)$ and $B \, \mathcal{G}_N(\varrho)$ are disjoint if and only if the ordered pair (A, B) belongs to the union of the following three sets

$$\begin{aligned} \Omega_\varrho^1 &= \left\{ (A, B) \in \mathrm{SO}(3)^2 \colon \; h_c(\varrho, A^\top B) \geq 0 \right\}, \\ \Omega_\varrho^2 &= \bigcup_{k=1}^N \bigcap_{\ell=1}^N \left\{ (A, B) \in \mathrm{SO}(3)^2 \colon \; h_{k,\ell}(\varrho, A^\top B) \geq 0 \right\}, \\ \Omega_\varrho^3 &= \bigcup_{k=1}^N \bigcap_{\ell=1}^N \left\{ (A, B) \in \mathrm{SO}(3)^2 \colon \; h_{k,\ell}(\varrho, B^\top A) \geq 0 \right\}. \end{aligned}$$

Hence, for $N \geq 4$, Problem (4.1) can equivalently be formulated as

$$\begin{aligned} \max \quad & \varrho \\ \text{s.t.} \quad & \varrho \in (0, \pi/2), \\ & A_1, \dots, A_{n-1} \in \mathrm{SO}(3), \\ & (I, A_i) \in \Omega_\varrho^1 \cup \Omega_\varrho^2 \cup \Omega_\varrho^3 \quad \text{for } 1 \leq i \leq n-1, \\ & (A_i, A_j) \in \Omega_\varrho^1 \cup \Omega_\varrho^2 \cup \Omega_\varrho^3 \quad \text{for } 1 \leq i < j \leq n-1. \end{aligned} \tag{4.2}$$

If $N = 3$, the formulation above is valid for $\varrho \in \big(0, \arccos\big(\sqrt{5}/3\big)\big]$, since regular spherical triangles are admissible only in this range.

The interior-disjointness constraint in packing problems for spherical caps reduces to Ω_ϱ^1, which leads to the simplified problem

$$\begin{aligned} \max \quad & \varrho \\ \text{s.t.} \quad & \varrho \in (0, \pi/2), \\ & A_1, \dots, A_{n-1} \in \mathrm{SO}(3), \\ & (I, A_i) \in \Omega_\varrho^1 \qquad \text{for } 1 \leq i \leq n-1, \\ & (A_i, A_j) \in \Omega_\varrho^1 \qquad \text{for } 1 \leq i < j \leq n-1. \end{aligned} \tag{4.3}$$

Numerical approximations to stationary points of these optimization problems can be obtained using the *Riemannian augmented Lagrangian method* developed by Liu and Boumal [53]. In the following section, we provide a brief overview of this method and explain how it can be applied to the present setting.

4.2 Numerical methods

The *Riemannian augmented Lagrangian method* extends the classical augmented Lagrangian framework, for which we refer to [9], to optimization problems posed on Riemannian submanifolds. More precisely, it is used to compute numerical approximations of first-order stationary points of problems of the form

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in \mathcal{M}, \\ & h_k(x) \leq 0 \quad \text{for } 1 \leq k \leq m, \end{aligned} \tag{4.4}$$

where $\mathcal{M}$ is a Riemannian submanifold and $f, h_1, \dots, h_m$ are twice continuously differentiable functions $\mathcal{M} \to \mathbb{R}$.

For a given penalty parameter $\rho > 0$, the augmented Lagrangian function is defined as

$$\Lambda_\rho(x, \lambda) = f(x) + \frac{\rho}{2} \sum_{k=1}^{m} \max\left\{0, \frac{\lambda_k}{\rho} + h_k(x)\right\}^2 .$$

Under the stated assumptions, the function Λ_ρ is continuously differentiable on $\mathcal{M} \times \mathbb{R}^m$.

The main steps of the Riemannian augmented Lagrangian method can be summarized as follows

1. For a fixed penalty parameter ρ and fixed nonnegative Lagrange multipliers $\lambda \in \mathbb{R}^m_{\geq 0}$, compute a first-order stationary point of the function $\Lambda_\rho(\cdot, \lambda)$ over $\mathcal{M}$. This can be achieved using standard techniques of smooth Riemannian optimization, such as gradient descent or Newton-type methods. An overview of such techniques can be found in [12].

2. Update the Lagrange multipliers and the penalty parameter based on the current constraint violations.

Under suitable assumptions, iterating these two steps yields an approximation of a first-order stationary point on $\mathcal{M}$ that satisfies the imposed constraints. For a detailed discussion of the convergence of this method, see [53].

So far, we have considered constraint sets of the form

$$x \in \{x \in \mathcal{M} : \; h(x) \leq 0\}$$

for a sufficiently smooth function h. However, the optimization problem (4.2) involves finite unions and intersections of such sets. A constraint of the form

$$x \in \bigcup_{k=1}^{m} \{x \in \mathcal{M} : \; h_k(x) \leq 0\}$$

can be modeled by requiring non-positivity of the pointwise minimum over all the defining constraint functions, whereas intersections can be enforced using the pointwise maximum.

This observation allows for a reformulation of (4.2) as an instance of (4.4), albeit with nonsmooth constraint functions. By Rademacher's theorem [70, Chapter 2.1, Theorem 1.4], the set of nonsmooth points for these functions has measure zero. Consequently, the lack of smoothness does not pose significant difficulties in numerical practice. To avoid this ambiguity altogether, one may either choose a subgradient or use a suitable smooth approximation of the minimum and maximum functions.

To obtain rigorous packings, we project the floating-point output of the Riemannian augmented Lagrangian method onto the field of rational numbers using the *Cayley transform*. Let A be a matrix such that none of its eigenvalues equals -1. Its Cayley transform is defined as $(I-A)\,(I+A)^{-1}$. Outside a singular set of measure zero, the Cayley transform defines a self-inverse map between $\mathrm{SO}(3)$ and the space of skew-symmetric matrices.

Consider a floating-point matrix obtained by solving Problem (4.2) numerically. To project the matrix onto $\mathrm{SO}(3) \cap \mathbb{Q}^{3\times 3}$, we start by applying the Cayley transform. By rounding each entry to a nearby rational number, we obtain a rational matrix $A_{\mathbb{Q}} \in \mathbb{Q}^{3\times 3}$. Since $A_{\mathbb{Q}}$ is not necessarily skew-symmetric, we restore that property by replacing $A_{\mathbb{Q}}$ with $(A_{\mathbb{Q}} - A_{\mathbb{Q}}^{\top})/2$. Applying the Cayley transform to this matrix yields a rational element of $\mathrm{SO}(3)$.

After projecting the numerical output onto the field of rational numbers, we need to ensure that the resulting configuration defines a valid packing. The computed angular radius ϱ could be too large or unnecessarily small. To approximate the maximal feasible angular radius, we perform a binary search in ϱ.

Chapter Five

Configurations

Using the proposed manifold optimization approach, we compute packings for regular spherical triangles, squares, and pentagons with cardinalities ranging from 3 to 24. The optimization is implemented in the `Julia` programming language [8] using the `Manopt` and `Manifolds` packages [3, 7]. All scripts, together with a comprehensive database of the computed configurations, are available at [28].

This chapter provides a summary and analysis of the resulting configurations. A complete list of all constructed packings, together with their corresponding angular radii, is provided in Chapter 9. Figures 5.1, 5.2, and 5.3 illustrate the dependence of the angular radius on the cardinality.

To assess the reliability of the proposed approach, we also compute optimized packings of spherical caps with cardinalities ranging from 5 to 24 by numerically solving (4.3). Table 5.1 compares the obtained angular radii with the best-known values reported in the database of spherical codes maintained by Henry Cohn [19]. Up to negligible numerical deviations, the computed radii agree with those in the database, which provides strong evidence for the reliability of the optimization procedure.

For regular spherical triangles, the admissibility criterion imposes the bound $\varrho \leq \arccos(\sqrt{5}/3)$. Among the computed triangle configurations, this condition is satisfied only from cardinality $n = 14$ onward, as illustrated in Figure 5.1. For smaller cardinalities, the method still produces configurations that, assuming Conjecture 3.15, define valid packings. The facets of an octahedron induce a packing of 8 regular spherical triangles whose dual polygons have inner angles equal to $\pi/2$. This configuration serves as the starting point for our computations, since the inner angle of the corresponding polygon increases in the subsequent packings.

n	Best radius	Manifold opt.	Difference
$5^\star$	0.785398163397	0.785398163346	5.07×10^{-11}
$6^\star$	0.785398163397	0.785398163347	5.02×10^{-11}
$7^\star$	0.679539948816	0.679539948739	7.63×10^{-11}
$8^\star$	0.653263580858	0.653263580805	5.37×10^{-11}
$9^\star$	0.615479708670	0.615479708605	6.51×10^{-11}
$10^\star$	0.577239916709	0.577239916649	5.96×10^{-11}
$11^\star$	0.553574358897	0.553574358834	6.26×10^{-11}
$12^\star$	0.553574358897	0.553574358833	6.32×10^{-11}
$13^\star$	0.498611796219	0.498611796104	1.14×10^{-10}
$14^\star$	0.485817371443	0.485817371369	7.37×10^{-11}
15	0.468253077154	0.468253077068	8.61×10^{-11}
16	0.455918360799	0.455918360677	1.23×10^{-10}
17	0.445847224575	0.445847224435	1.40×10^{-10}
18	0.432463395991	0.432463395488	5.03×10^{-10}
19	0.416190463888	0.416190457510	6.38×10^{-9}
20	0.413913874894	0.413913873373	1.52×10^{-9}
21	0.398050462820	0.398047141625	3.32×10^{-6}
22	0.390431560119	0.390431558872	1.25×10^{-9}
23	0.381441395654	0.381441386802	8.85×10^{-9}
$24^\star$	0.381273869375	0.381273868311	1.06×10^{-9}

Table 5.1: Comparison of best-known angular radii for spherical cap packings and values obtained via the manifold optimization approach. The $\star$ indicates that the corresponding configuration is proven to be optimal.

To identify configurations with distinguished structure, we examine the discrete differences of the angular radius as a function of the cardinality. Deviations from the otherwise monotonic decay indicate configurations with increased geometric regularity or symmetry, which are also apparent in the corresponding plots.

The most notable cases identified in this way are the packings of 8 and 20 regular spherical triangles, of 6 regular spherical squares, and of 12 regular spherical pentagons. These configurations are induced by the facets of the Platonic solids, namely the octahedron, icosahedron, cube, and dodecahedron, and therefore form exact tilings of the sphere.

The configurations with cardinalities immediately preceding these tiling configurations are of comparable angular radius. They arise from

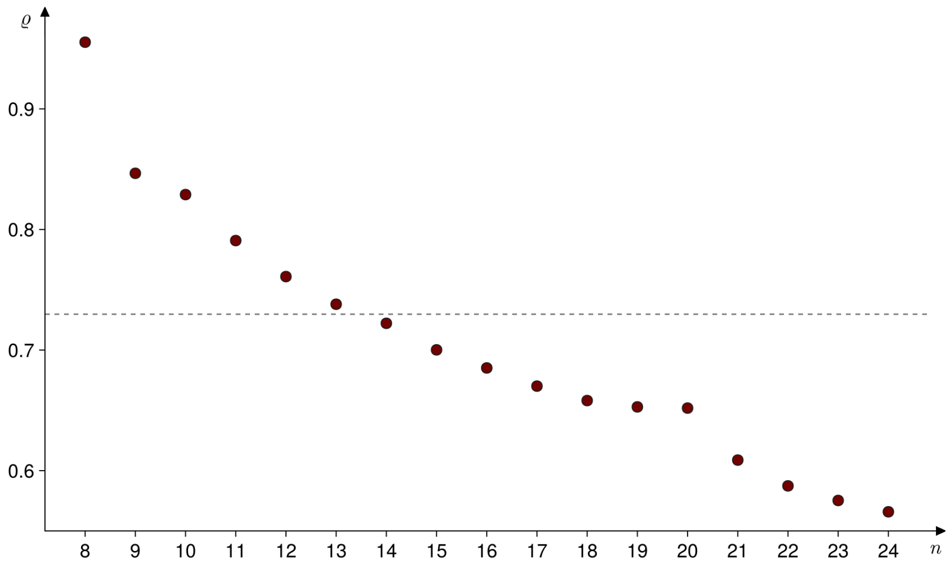


Figure 5.1: Constructed packings of n regular spherical triangles and their corresponding angular radii. All configurations with inner angles of at least $\pi/2$ are included. The dashed line at $\varrho = \arccos\left(\sqrt{5}/3\right)$ indicates the threshold for admissibility. Configurations above this threshold are treated as conjectural.

removing a single polygon from a tiling configuration and re-optimizing the resulting arrangement. The resulting packings retain much of the original geometric structure, indicating that these tilings act as stable templates for locally optimal configurations. An example of this phenomenon is illustrated in Figure 5.4.

Additional notably dense configurations occur for 10 and 24 spherical triangles, 12, 20, 22, and 24 spherical squares, as well as 4, 6, 20, 22, and 24 spherical pentagons.

The angular radii obtained for packings of regular spherical pentagons agree, up to numerical precision, with those reported by Tarnai and Gáspár. For $n = 7, 9, 10$, and 24, however, the constructed configurations exhibit strictly larger angular radii, thereby improving upon the previously reported values, see Table 5.2.

The constructed packings yield lower bounds on the problem of determining the maximal number of congruent copies of a Platonic solid that can share a common vertex. As discussed in the introductory chapter, this problem can be reformulated as a packing problem for

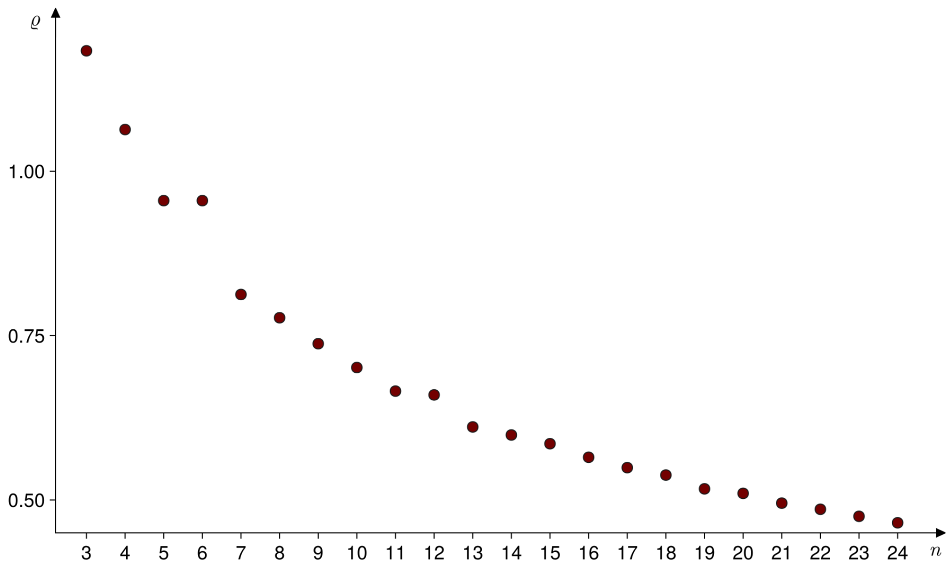


Figure 5.2: Constructed packings of n regular spherical squares and their corresponding angular radii.

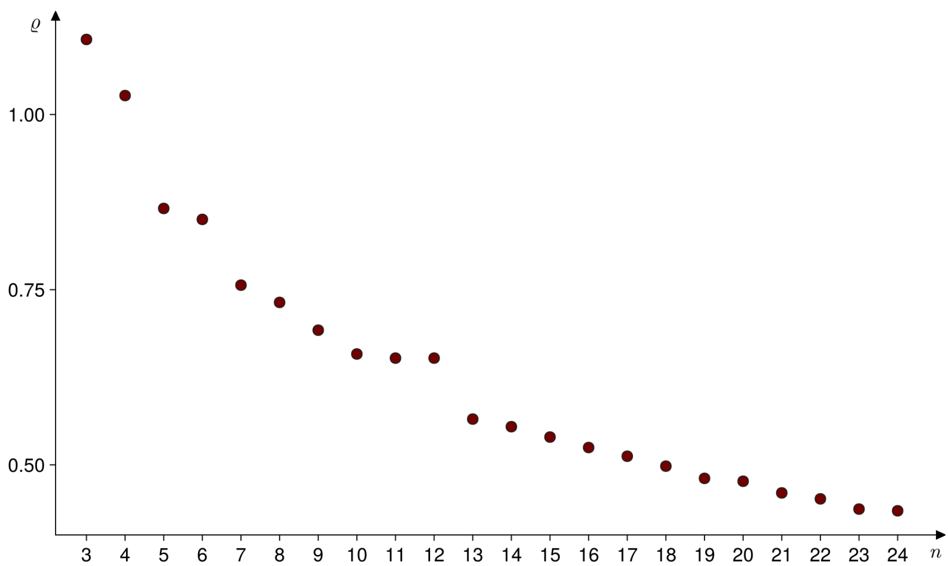


Figure 5.3: Constructed packings of n regular spherical pentagons and their corresponding angular radii.

n	Tarnai and Gáspár	Manifold opt.	Difference
3	1.107148717794	1.107148717286	5.07×10^{-10}
4	1.027058947449	1.027058933447	1.40×10^{-8}
5	0.865924548398	0.865924529734	1.86×10^{-8}
6	0.850308808328	0.850308804251	4.07×10^{-9}
7	0.755559002214	0.756498153903	9.39×10^{-4}
8	0.731849437364	0.731849436898	4.65×10^{-10}
9	0.673172689642	0.692196532848	1.90×10^{-2}
10	0.657631771837	0.658224309839	5.93×10^{-4}
24	0.433685276841	0.435221769445	1.53×10^{-3}

Table 5.2: Comparison of angular radii for packings of regular spherical pentagons reported by Tarnai and Gáspár with those obtained via the manifold optimization approach.

the corresponding regular spherical polygon. Any packing of such a spherical polygon that has the same number of vertices and a larger angular radius provides a corresponding lower bound.

From the computed packings, we recover the known lower bound of 20 for the tetrahedron and obtain lower bounds of 7 for the octahedron and 4 for the icosahedron. In the case of the dodecahedron, the associated spherical triangle packing is not admissible and therefore the method is not applicable. The case of the cube is immediate.

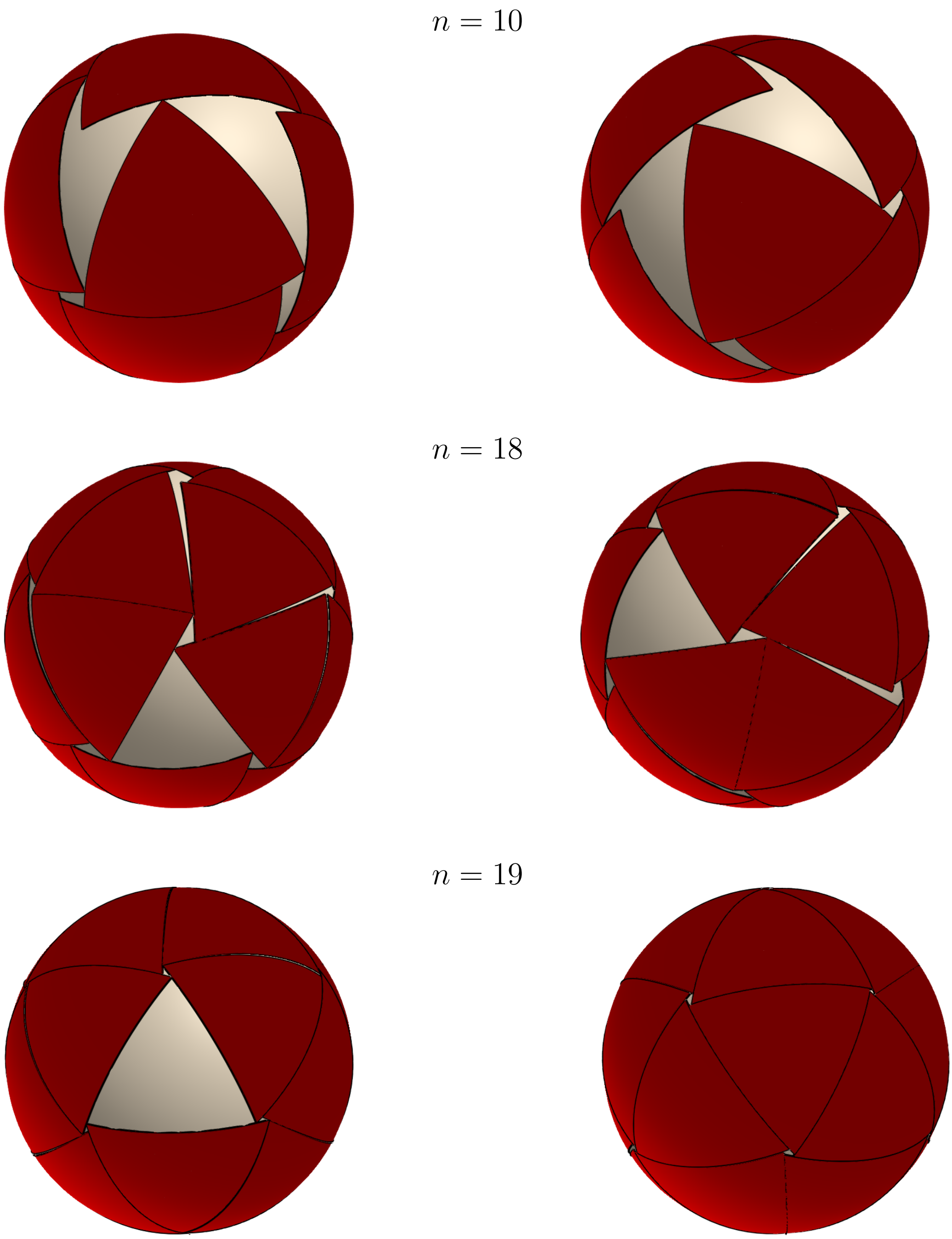


Figure 5.4: Examples of packings of regular spherical triangles with cardinalities 10, 18, and 19. The configuration of cardinality 10 is conjectured to form a packing, whereas the other two configurations satisfy the admissibility criterion. For each configuration, two complementary views of the sphere are shown. Polygons intersecting the boundary of one view continue on the opposite side in the corresponding complementary view.

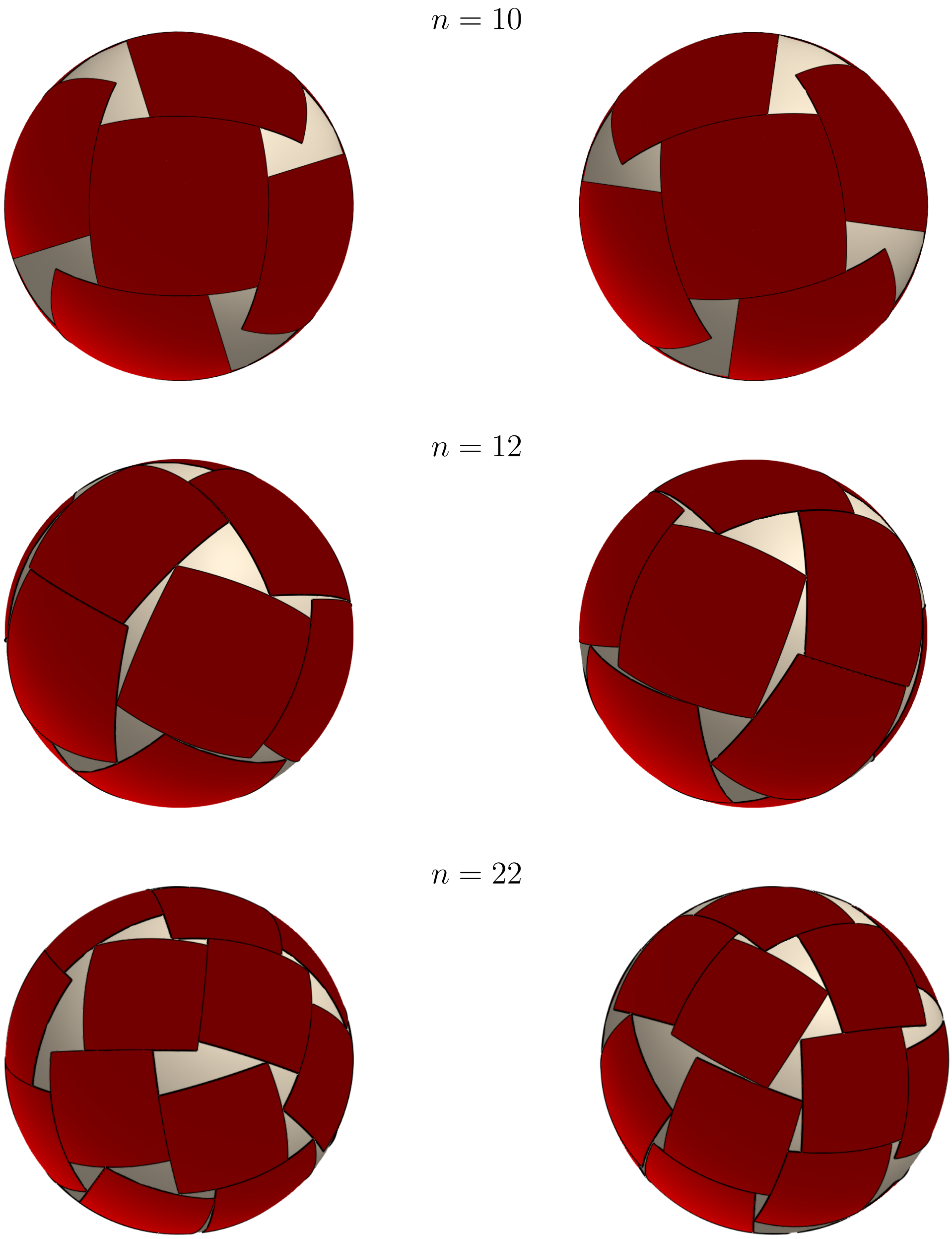


Figure 5.5: Examples of packings of regular spherical squares with cardinalities 10, 12, and 22. For each configuration, two complementary views of the sphere are shown. Polygons intersecting the boundary of one view continue on the opposite side in the corresponding complementary view.

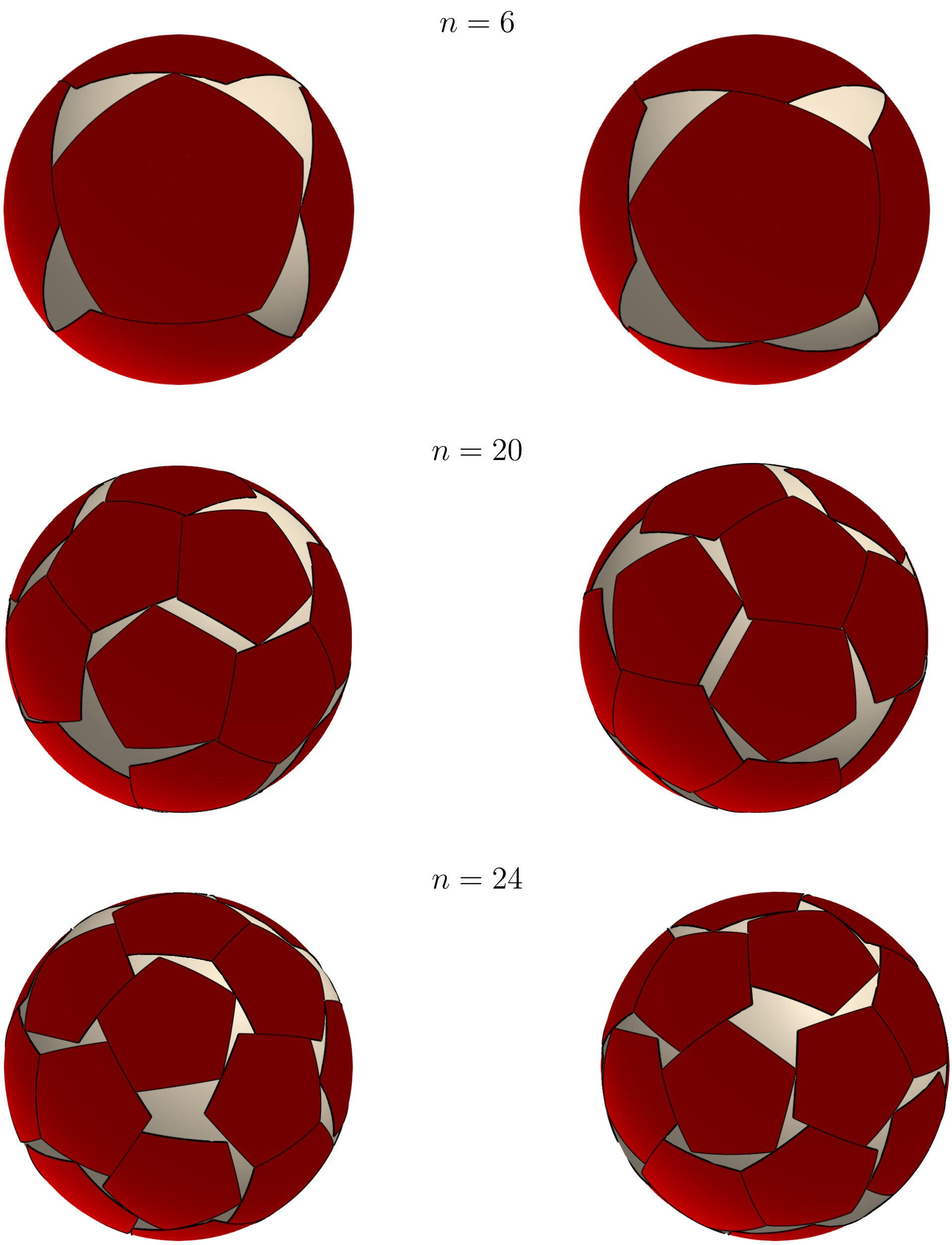


Figure 5.6: Examples of packings of regular spherical pentagons with cardinalities 6, 20, and 24. For each configuration, two complementary views of the sphere are shown. Polygons intersecting the boundary of one view continue on the opposite side in the corresponding complementary view.

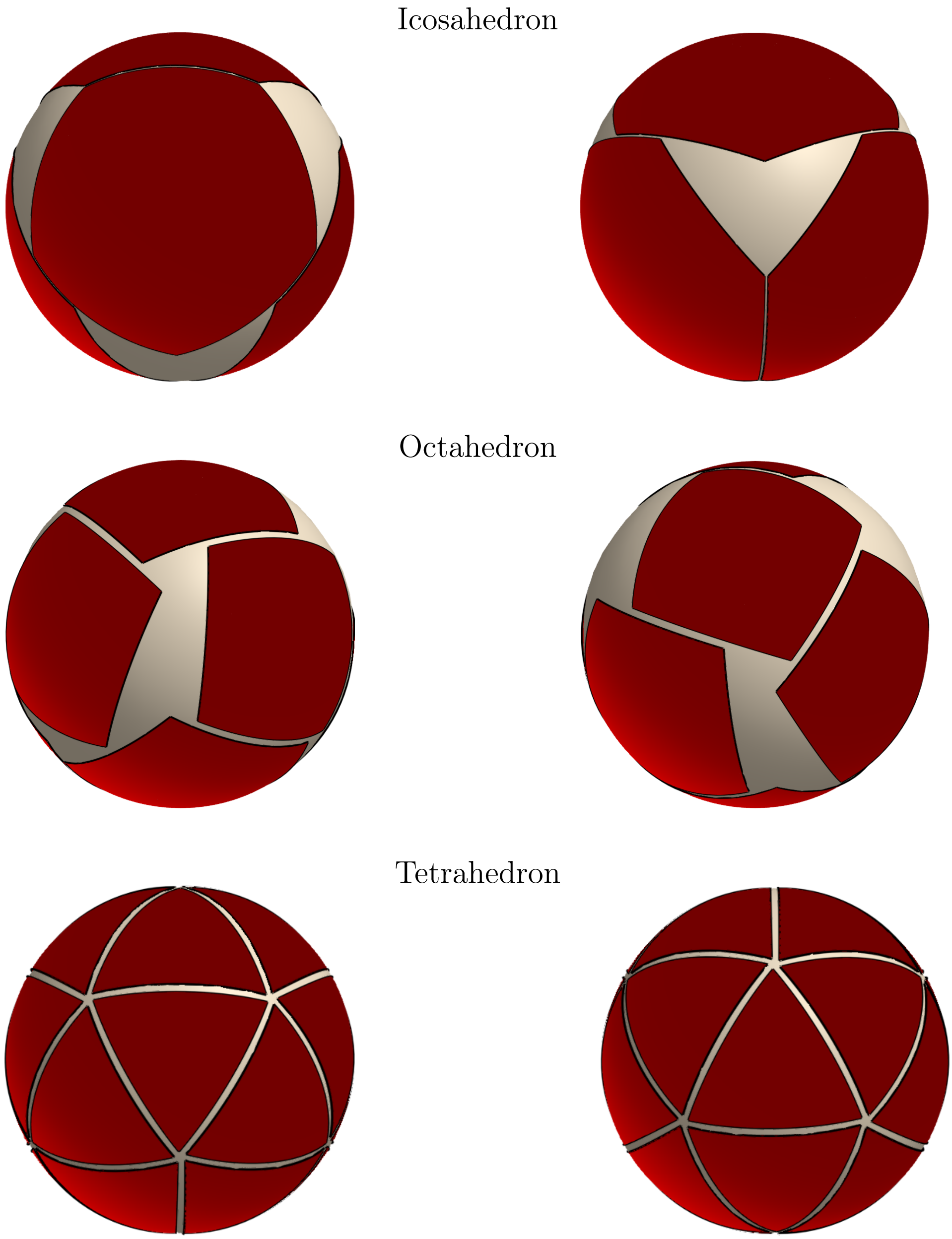


Figure 5.7: Packings of spherical polygons corresponding to lower bounds on the vertex-sharing problem of Platonic solids. Shown are the configurations associated with the icosahedron, octahedron, and tetrahedron. For each case, two complementary views of the sphere are displayed. Polygons intersecting the boundary of one view continue on the opposite side in the corresponding complementary view.

Part III
Proving maximality

Chapter Six

Upper bounds for packing numbers

After establishing dense configurations that serve as lower bounds for the corresponding packing numbers, we address the question of their maximality. This requires upper bounds on the packing number. In order to prove maximality, it is typically not necessary to obtain sharp upper bounds. Indeed, any upper bound whose floor agrees with the cardinality of the configuration under consideration suffices to establish maximality.

Let $\mathcal{P} \subseteq \mathbb{S}^{d-1}$ be a spherical polytope, and let ω denote the surface measure on $\mathbb{S}^{d-1}$. We define the volume bound for the packing problem for $\mathcal{P}$ by

$$\frac{\omega(\mathbb{S}^{d-1})}{\omega(\mathcal{P})},$$

which is an immediate upper bound for the packing number $\tau(\mathcal{P})$.

For regular spherical polygons on $\mathbb{S}^2$, Girard's theorem [21] yields an explicit formula for the volume of a regular spherical polygon, and hence for the corresponding volume bound

$$\frac{\omega(\mathbb{S}^2)}{\omega(\mathcal{G}_N(\varrho))} = \frac{4\pi}{2N\phi + (2-N)\pi}, \tag{6.1}$$

where

$$\phi = \arccos\left(\frac{\cos(\varrho)\sin(\pi/N)}{\sqrt{1-\sin(\pi/N)^2\,\sin(\varrho)^2}}\right).$$

Since the volume bound does not incorporate the relative positions of congruent copies, it is often too weak to establish maximality of a given packing.

For this reason, we introduce two additional upper bounds for packing numbers in this chapter, both of which exploit geometric information. We discuss the relationships among these bounds and examine the role of symmetries.

A standard way to model packing problems is to formulate them as independent set problems in suitable graphs. This approach is widely used in discrete geometry, see [79] for several examples. Since the underlying vertex set may be infinite, one needs a suitable framework to define the usual graph-theoretical objects on such graphs. De Laat and Vallentin provide a framework for independent set problems on graphs equipped with a particular topology [26], which we employ throughout this chapter.

A graph $G = (V, E)$ is a *topological packing graph* if the vertex set is endowed with a Hausdorff topology and every finite clique is contained in an open clique. For the remainder of this chapter, we assume that V is a compact space. This assumption ensures, in particular, that the independence number is finite.

To model the packing problem for a given spherical polytope $\mathcal{P} \subseteq \mathbb{S}^{d-1}$, we choose $\mathrm{SO}(d)$ to be the vertex set. Two rotations $A, B \in \mathrm{SO}(d)$ are adjacent if $A\mathcal{P} \cap B\mathcal{P}$ has nonempty interior. Thus, packings correspond to independent sets and vice versa.

The resulting graph has additional structure. The vertex set $\mathrm{SO}(d)$ forms a matrix Lie group, and by the symmetry relation (3.6), two rotations A and B are adjacent if and only if I and $A^{-1}B$ are adjacent. Such graphs are known as *Cayley graphs*.

6.1 The Lovász theta number

In 1979, Lovász introduced an upper bound on the independence number of a finite graph, known as the theta number [54]. In its original formulation, the theta number is defined as an optimization problem over the cone of positive semidefinite matrices indexed by the vertex set of the graph. In order to extend this notion to infinite graphs, it is necessary to identify an appropriate analogue of positive semidefinite matrices in an infinite-dimensional setting.

Let V be a compact Hausdorff space. A natural replacement is given by positive semidefinite kernels. A continuous Hermitian kernel $K\colon V \times V \to \mathbb{C}$ is called *positive semidefinite* if, for any finite collection of points $v_1, \dots, v_m \in V$, the matrix $\big[K(v_k, v_\ell)\big]_{k,\ell=1}^m$ is positive semidefinite. The set of such kernels forms a convex cone.

Given a topological packing graph $G = (V, E)$, we define the *theta number* by

$$\begin{aligned} \inf \quad & \lambda \\ \text{s.t.} \quad & \lambda \in \mathbb{R}_{\geq 0}, \\ & K\colon V \times V \to \mathbb{R} && \text{continuous and positive semidefinite,} \\ & K(v,v) \leq \lambda - 1 && \text{for all } v \in V, \\ & K(u,v) \leq -1 && \text{for all } \{u,v\} \notin E \text{ and } u \neq v. \end{aligned} \tag{6.2}$$

Let $I \subseteq V$ be an independent set and let (λ, K) be a feasible solution of (6.2). Then

$$0 \leq \sum_{u,v \in I} K(u,v) = \sum_{u \in I} K(u,u) + \sum_{\substack{u,v \in I \\ u \neq v}} K(u,v) \leq |I|\,(\lambda - |I|),$$

which implies $|I| \leq \lambda$. Consequently, the theta number provides an upper bound on the independence number of G.

Let G be a Cayley graph defined on a compact matrix Lie group Γ equipped with its Haar measure μ. Any continuous kernel $K\colon \Gamma \times \Gamma \to \mathbb{C}$ can be symmetrized by

$$K_{\text{sym}}(\gamma_1, \gamma_2) = \int_\Gamma K(g^{-1}\gamma_1, g^{-1}\gamma_2)\, d\mu(g).$$

The resulting kernel is invariant under the action of Γ, defined by

$$(g \cdot K)(\gamma_1, \gamma_2) = K(g^{-1}\gamma_1, g^{-1}\gamma_2).$$

In particular, $K_{\text{sym}}(\gamma_1, \gamma_2) = K_{\text{sym}}(e, \gamma_1^{-1}\gamma_2)$ holds, which implies that the value of $K_{\text{sym}}(\gamma_1, \gamma_2)$ depends only on $\gamma_1^{-1}\gamma_2$. Thus, invariant kernels may be identified with functions on Γ. Conversely, every continuous function $f\colon \Gamma \to \mathbb{C}$ defines a continuous kernel by $K(\gamma_1, \gamma_2) = f(\gamma_1^{-1}\gamma_2)$.

We say that a function $f\colon \Gamma \to \mathbb{C}$ is *of positive type* if the associated kernel $(\gamma_1, \gamma_2) \mapsto f\left(\gamma_1^{-1}\gamma_2\right)$ is Hermitian and positive semidefinite. The symmetrization operator defines a surjective map from the cone of positive semidefinite kernels onto the cone of positive type functions.

Moreover, if (λ, K) is feasible for (6.2), so is $(\lambda, K_{\mathrm{sym}})$. Consequently, restricting attention to kernels arising from positive type functions does not weaken the bound. The theta number therefore simplifies to

$$\begin{array}{lll} \inf & f(e)+1 & \\ \text{s.t.} & f\colon \Gamma \to \mathbb{R} & \text{is continuous and of positive type,} \\ & f(\gamma) \leq -1 & \text{for all } \gamma \in \Gamma \setminus \{e\} \text{ such that } \{e,\gamma\} \notin E. \end{array} \tag{6.3}$$

Applying the preceding framework to packings of spherical polytopes yields the following theorem.

Theorem 6.1. *Let $\mathcal{P} \subseteq \mathbb{S}^{d-1}$ be a spherical polytope, and let $f\colon\ \mathrm{SO}(d) \to \mathbb{R}$ be a continuous positive type function such that*

$$f(A) \leq -1 \qquad \textit{whenever} \qquad (\mathcal{P} \cap A\mathcal{P})^{\circ} = \emptyset.$$

Then the packing number $\tau(\mathcal{P})$ is bounded above by $f(I)+1$. Moreover, there exists such a function f for which

$$f(I)+1 = \frac{\omega(\mathbb{S}^{d-1})}{\omega(\mathcal{P})},$$

where ω denotes the surface measure on $\mathbb{S}^{d-1}$.

Proof. The first part of the theorem follows directly from the preceding discussion. It remains to prove the second assertion. Let $\mathcal{P} \subseteq \mathbb{S}^{d-1}$ be a spherical polytope, let ω be the surface measure on $\mathbb{S}^{d-1}$, and let μ denote the Haar measure on $\mathrm{SO}(d)$. We define a function on $\mathrm{SO}(d)$ by

$$h(A) = \mu(\mathcal{R} \cap A\mathcal{R}) = \int_{\mathrm{SO}(d)} \mathbb{1}_{\mathcal{R}}(R)\, \mathbb{1}_{\mathcal{R}}(A^{\top} R)\, d\mu(R),$$

where $\mathcal{R} = \{R \in \mathrm{SO}(d) : Re_d \in \mathcal{P}\}$, and $\mathbb{1}_{\mathcal{R}}$ denotes the indicator function of the set $\mathcal{R}$. Since h is a convolution of two functions in $L^{\infty}(\mathrm{SO}(d))$, it is continuous by [37, Proposition 2.39].

Using the invariance of the Haar measure, we obtain

$$h(A^{\top} B) = \mu(A\mathcal{R} \cap B\mathcal{R}) = \int_{\mathrm{SO}(d)} \mathbb{1}_{A\mathcal{R}}(R)\, \mathbb{1}_{B\mathcal{R}}(R)\, d\mu(R).$$

Thus, for every finite set $\{A_1, \ldots, A_m\} \subseteq \mathrm{SO}(d)$, the corresponding matrix $[h(A_k^{\top} A_{\ell})]_{k,\ell=1}^{m}$ is a Gram matrix and therefore positive semidefinite. Consequently, h is of positive type.

From

$$\begin{aligned}\int_{\mathrm{SO}(d)} h(A^\top B)\,d\mu(B) &= \int_{\mathrm{SO}(d)}\int_{\mathrm{SO}(d)} \mathbb{1}_{\mathcal{R}}(R)\,\mathbb{1}_{\mathcal{R}}(B^\top AR)\,d\mu(R)\,d\mu(B)\\ &= \int_{\mathrm{SO}(d)} \mathbb{1}_{\mathcal{R}}(R) \int_{\mathrm{SO}(d)} \mathbb{1}_{\mathcal{R}}(B^\top AR)\,d\mu(B)\,d\mu(R)\\ &= \mu(\mathcal{R})^2,\end{aligned}$$

we conclude that the constant function is an eigenfunction of the integral operator induced by the kernel $(A,B)\mapsto h(A^\top B)$ with corresponding eigenvalue $\mu(\mathcal{R})^2$.

Using the fact that the normalized surface measure $\omega/\omega(\mathbb{S}^{d-1})$ is the pushforward of the Haar measure μ through the map $A\mapsto Ae_d$, we obtain

$$h(A) = \frac{\omega(\mathcal{P}\cap A\mathcal{P})}{\omega(\mathbb{S}^{d-1})}.$$

The function

$$f(A) = \frac{h(A)}{\mu(\mathcal{R})^2} - 1$$

satisfies the conditions of the theorem. In fact, f is continuous and $f(A) = -1$ if $\mathcal{P}$ and $A\mathcal{P}$ are interior-disjoint. The constant kernel has rank one, and its only nonzero eigenspace is the space of constant functions with corresponding eigenvalue 1. Hence, all eigenvalues of the kernel induced by f are nonnegative. Mercer's theorem [41, Theorem 4.14] implies that f is of positive type. The claim follows from $f(I)+1 = \omega(\mathbb{S}^{d-1})/\omega(\mathcal{P})$. □

The second part of Theorem 6.1 shows that the theta number is at least as strong as the volume bound. To construct efficient bounds, we aim to optimize over the cone of positive type functions while ensuring that all constraints are satisfied. The subsequent chapters develop the tools required for this purpose.

This chapter concludes by considering the theta number for spherical cap packings and explaining how it can be applied to the packing problem for spherical polytopes.

6.2 Upper bounds via subgraphs

Let $G=(V,E)$ be a topological packing graph and let $G'=(V,E')$ be a subgraph with $E'\subseteq E$. Then every independent set in G is also

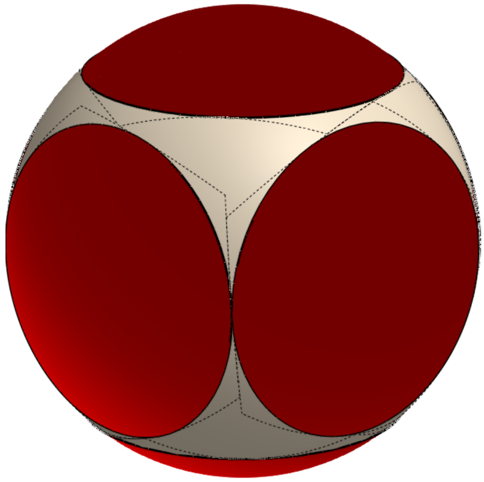


Figure 6.1: The inscribed spherical caps of a packing of regular spherical polygons form a spherical cap packing.

independent in G', and hence the independence number of G' provides an upper bound for the independence number of G. In particular, removing edges from a graph can only increase its independence number. At the same time, edge removal may enlarge the automorphism group, resulting in a more symmetric graph. This additional symmetry often simplifies the computation of the independence number.

We apply this principle to the graph G corresponding to the packing problem for a spherical polytope $\mathcal{P} \subseteq \mathbb{S}^{d-1}$. Let $\mathcal{C} \subseteq \mathcal{P}$ be the inscribed spherical cap of $\mathcal{P}$. The graph associated with the packing problem for $\mathcal{C}$ is a subgraph of G, which implies $\tau(\mathcal{P}) \leq \tau(\mathcal{C})$.

The theta number for the spherical cap packing problem coincides with the linear programming bound of Delsarte, Goethals, and Seidel [33]. They showed that the packing number of a spherical cap of angular radius $\varrho \in (0, \pi/2)$ on $\mathbb{S}^{d-1}$ is bounded above by the optimal value of the following optimization problem

$$\begin{aligned} \min \quad & f(1) + 1 \\ \text{s.t.} \quad & f(t) = \sum_{k=0}^{m} \lambda_k \, C_k^{\alpha}(t) \quad \text{with } \lambda_0, \ldots, \lambda_m \geq 0, \\ & f(t) \leq -1 \qquad\qquad\quad \text{for all } t \in [-1, \cos(2\varrho)], \end{aligned} \tag{6.4}$$

where C_k^{α} denotes the Gegenbauer polynomial of degree k with parameter $\alpha = d/2 - 1$. For background on their properties, see [1].

We call a packing *Delsarte-tight* if its cardinality equals the Delsarte bound.

More recent developments have led to stronger bounds, most notably those due to Bachoc and Vallentin [4] and to de Laat, Leijenhorst and de Muinck Keizer [25]. For the computations in Chapter 9, we employ the bound by Bachoc and Vallentin. Although weaker in principle compared to the bound by de Laat, Leijenhorst and de Muinck Keizer, the Bachoc and Vallentin bound is considerably faster to compute, and no substantial improvements relevant to this work are expected.

The main advantage of the Delsarte bound lies in its simplicity. While the stronger bounds rely on numerical optimization techniques to construct feasible solutions, the Delsarte bound often allows for explicit, analytic constructions of feasible solutions.

For certain packings, it is possible to slightly reduce the radius of the spherical caps while still being able to certify maximality via (6.4). Consider a maximal packing of n spherical caps of angular radius ϱ. We define the *stability region* of such a packing as the largest interval $I \subseteq [-1, 1]$ containing $\cos(2\varrho)$ such that for every $t \in I$, the optimal value of (6.4) lies in $[n, n+1)$ when the interval in the second constraint is replaced by $[-1, t]$. If no such interval exists, the stability region is defined to be empty. We call the stability region of a packing *trivial* if it is empty or consists of the single point $\{\cos(2\varrho)\}$.

A nontrivial stability region for a packing in $\mathbb{S}^2$ provides a useful tool for describing the asymptotic behavior of the packing problem for regular spherical polygons $\mathcal{G}_N(\varrho)$ as $N \to \infty$.

Theorem 6.2. *Let $A_1, \ldots, A_n \in \mathrm{SO}(3)$ define a maximal packing of spherical caps of radius $\varrho \in (0, \pi/2)$ and suppose that this packing has a nontrivial stability region. There exists $N_0 \in \mathbb{N}$ such that the configuration $A_1, \ldots, A_n$ is a maximal packing for all regular spherical polygons $\mathcal{G}_N(\varrho)$ with $N \geq N_0$.*

Proof. By construction $\mathcal{G}_N(\varrho) \subseteq \mathcal{C}(\varrho)$, and hence the configuration $A_1, \ldots, A_n$ defines a packing for $\mathcal{G}_N(\varrho)$ for all $N \geq 3$.

The inscribed radius of the regular spherical polygon $\mathcal{G}_N(\varrho)$ is given by

$$\varphi_N = \arccos\left(\frac{\cos(\varrho)}{\sqrt{1 - \sin(\pi/N)^2\,\sin(\varrho)^2}}\right),$$

and converges to ϱ as $N \to \infty$. Thus, there exists $N_0 \in \mathbb{N}$ such that $\cos(2\varphi_N)$ lies in the stability region for all $N \geq N_0$. The Delsarte

Name	ϱ	ε	N_0
Equilateral triangle	$\pi/3$	$1/6$	6
Regular tetrahedron	$\arccos(-1/3)/2$	$1/12$	8
Regular octahedron	$\pi/4$	$1/27$	12
Regular icosahedron	$\arccos\left(1/\sqrt{5}\right)/2$	$(17-7\sqrt{5})/198$	25

Table 6.1: Delsarte-tight packings on $\mathbb{S}^2$, their packing radii ϱ, the stability parameters ε from [15], and the smallest integers N_0 such that $\cos(2\varphi_N) \in [\cos(2\varrho), \cos(2\varrho)+\varepsilon)$ for all $N \geq N_0$.

bound then certifies the maximality of the corresponding packing of the inscribed spherical caps, which in turn implies the maximality of the associated regular spherical polygon packing. □

In general, no systematic method is known for computing the stability region of a given packing. Böröczky and Glazyrin proved that every Delsarte-tight packing has a nontrivial stability region, see [15, Theorem 5]. More precisely, for any Delsarte-tight packing of spherical caps of radius ϱ, they provided an explicit $\varepsilon > 0$ such that the interval $[\cos(2\varrho), \cos(2\varrho)+\varepsilon)$ is contained in the stability region.

Table 6.1 lists well-known Delsarte-tight packings of spherical caps, together with their packing radii ϱ, the corresponding parameter ε, and the smallest integer N_0 such that $\cos(2\varphi_N) \in [\cos(2\varrho), \cos(2\varrho)+\varepsilon)$ for all $N \geq N_0$.

A similar asymptotic analysis for packings not listed in the table would require an investigation of stability regions associated with stronger bounds. Although no systematic approach is currently known, numerical computations may still provide useful evidence or partial information.

Chapter Seven

Positive type functions

In order to construct feasible functions for Theorem 6.1, we study the cone of positive type functions on SO(3). Recall that a function $f\colon \Gamma \to \mathbb{C}$ is of positive type if, for every finite set $\{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$, the matrix $\big[f(\gamma_k^{-1}\gamma_\ell)\big]_{k,\ell=1}^m$ is Hermitian positive semidefinite.

Throughout the remainder of this thesis, the trace inner product of the complex matrices $A, B \in \mathbb{C}^{d\times d}$ is denoted by $\langle A, B\rangle = \mathrm{tr}(A^*B)$.

A complete characterization of continuous functions of positive type on compact separable groups was given by Bochner [11].

Theorem 7.1 (Bochner's Theorem)**.** *Let Γ be a compact and separable group, and let $\hat{\Gamma}$ denote its dual group. A function $f\colon \Gamma \to \mathbb{C}$ is continuous and of positive type if and only if it admits an expansion of the form*

$$f(\gamma) = \sum_{\rho\in\hat{\Gamma}} \langle F_\rho, \rho(\gamma)\rangle,$$

where the coefficient matrices F_ρ are Hermitian positive semidefinite for all $\rho \in \hat{\Gamma}$, $\sum_{\rho\in\hat{\Gamma}} \mathrm{tr}(F_\rho) < \infty$, and the series converges absolutely.

We briefly indicate why every function of this form is of positive type. For a given finite set $\{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$, a vector $v \in \mathbb{C}^m$, and an irreducible representation $\rho \in \hat{\Gamma}$, the matrix

$$\sum_{k,\ell=1}^{m} \rho(\gamma_k^{-1}\gamma_\ell)\, \overline{v_k}\, v_\ell = \left(\sum_{k=1}^{m} \rho(\gamma_k)\, v_k\right)^* \left(\sum_{\ell=1}^{m} \rho(\gamma_\ell)\, v_\ell\right)$$

is positive semidefinite. Consequently, each term

$$\sum_{k,\ell=1}^{m} \langle F_\rho, \rho(\gamma_k^{-1}\gamma_\ell)\rangle\, \overline{v_k}\, v_\ell$$

is a nonnegative real number.

To apply Bochner's theorem to SO(3), we require a complete set of its irreducible representations. This chapter provides a survey of the classical construction of these representations, following [1, 14, 18, 72]. In particular, we exploit the well-known relationship between SO(3) and SU(2) to describe their dual groups and give an explicit construction of a complete set of irreducible representations.

7.1 The relation between SU(2) and SO(3)

For $d \in \mathbb{N}$, the *special unitary group* $\mathrm{SU}(d)$ is defined as

$$\mathrm{SU}(d) := \{U \in \mathbb{C}^{d\times d} : \; U^* U = I, \; \det(U) = 1\}.$$

Every element of SU(2) can be written in the form

$$\begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \in \mathbb{C}^{2\times 2} \qquad \text{with} \qquad |a|^2 + |b|^2 = 1. \tag{7.1}$$

Such a matrix is uniquely determined by $|a|$ and by the arguments of a and b. Assuming $ab \neq 0$, we define the parameters $(\alpha, \beta, \gamma) \in [-2\pi, 2\pi) \times [0, \pi] \times [-\pi, \pi)$ as the unique solution to

$$a = e^{i(\alpha+\gamma)/2} \cos(\beta/2), \qquad b = i\, e^{i(\alpha-\gamma)/2} \sin(\beta/2).$$

For every $U \in \mathrm{SU}(2)$ there exists a triple of parameters $(\alpha, \beta, \gamma) \in [-2\pi, 2\pi) \times [0, \pi] \times [-\pi, \pi)$ such that

$$U = U(\alpha, \beta, \gamma) = U_Z(\gamma) U_X(\beta) U_Z(\alpha),$$

where

$$U_Z(\theta) = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}, \qquad U_X(\theta) = \begin{pmatrix} \cos(\theta/2) & i\sin(\theta/2) \\ i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}.$$

The parameters (α, β, γ) are referred to as *Euler angles* of SU(2).

We relate SU(2) to SO(3) by constructing a group homomorphism using the map $\xi : \mathbb{R}^3 \to \{M \in \mathbb{C}^{2\times 2} : \; M = M^*, \, \mathrm{tr}(M) = 0\}$ given by

$$\xi(x) = \begin{pmatrix} x_3 & x_1 - i x_2 \\ x_1 + i x_2 & -x_3 \end{pmatrix}.$$

Note that ξ is linear and invertible.

Theorem 7.2. *The map* $\Phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ *defined by*

$$\Phi(U)(x) = \xi^{-1}(U\xi(x)U^*)$$

is a surjective group homomorphism with kernel $\{\pm I\}$.

Proof. We first verify that Φ is a well-defined group homomorphism. To this end, let $U \in \mathrm{SU}(2)$ be given as in (7.1). Evaluating $\Phi(U)$ on the standard basis vectors e_1, e_2, e_3 yields

$$\begin{aligned}
\Phi(U)(e_1) &= \big(\mathrm{Re}\,(\overline{a}^2 - b^2)\,,\quad \mathrm{Im}\,(\overline{a}^2 - b^2)\,,\quad -2\,\mathrm{Re}\,(ab)\big)\,,\\
\Phi(U)(e_2) &= \big(-\mathrm{Im}\,(\overline{a}^2 + b^2)\,,\quad \mathrm{Re}\,(\overline{a}^2 + b^2)\,,\quad -2\,\mathrm{Im}\,(ab)\big)\,,\\
\Phi(U)(e_3) &= \big(2\,\mathrm{Re}\,(\overline{a}b)\,,\quad 2\,\mathrm{Im}\,(\overline{a}b)\,,\quad |a|^2 - |b|^2\big)\,,
\end{aligned}$$

where we represented U as in (7.1). Thus, $\Phi(U)$ is given by

$$\Phi\,(U) = \begin{pmatrix} \mathrm{Re}\,(\overline{a}^2 - b^2) & -\mathrm{Im}\,(\overline{a}^2 + b^2) & 2\,\mathrm{Re}\,(\overline{a}b) \\ \mathrm{Im}\,(\overline{a}^2 - b^2) & \mathrm{Re}\,(\overline{a}^2 + b^2) & 2\,\mathrm{Im}\,(\overline{a}b) \\ -2\,\mathrm{Re}\,(ab) & -2\,\mathrm{Im}\,(ab) & |a|^2 - |b|^2 \end{pmatrix}.$$

Substituting the Euler parametrization of SU(2) into the above expression shows

$$\Phi(U_Z(\theta)) = R_Z(\theta) \qquad \text{and} \qquad \Phi(U_X(\theta)) = R_X(\theta), \tag{7.2}$$

where $R_Z(\theta)$ and $R_X(\theta)$ denote the rotations defined in Section 3.2.

By evaluating the composition of the two linear maps $\Phi(U_1)$ and $\Phi(U_2)$ at $x \in \mathbb{R}^3$, we get

$$\begin{aligned}
(\Phi(U_1) \circ \Phi(U_2))(x) &= \Phi(U_1)\left(\xi^{-1}(U_2\,\xi(x)\,U_2^*)\right) \\
&= \xi^{-1}(U_1\,U_2\,\xi(x)\;U_2^*\,U_1^*) \\
&= \Phi(U_1\,U_2)(x).
\end{aligned} \tag{7.3}$$

Equations (7.2) and (7.3) show that the map Φ is a well-defined group homomorphism. Together with the Euler-angle decomposition of SO(3), Equation (7.2) implies that Φ is surjective.

We conclude by computing the kernel of Φ. Let $U \in \ker(\Phi)$. Then for all $x \in \mathbb{R}^3$,

$$x = \Phi(U)(x) = \xi^{-1}\,(U\,\xi(x)\,U^*)$$

implies $\xi(x) = U\,\xi(x)\,U^*$ and therefore $U\xi(x) = \xi(x)\,U$. Thus, the matrix U must commute with every Hermitian traceless matrix, from which we deduce $U = \pm I$. $\square$

Theorem 7.2 relates the representations of SO(3) to the representations of SU(2). In particular, let ρ be an irreducible representation of SO(3). The composition $\rho \circ \Phi$ defines an irreducible representation of SU(2) that is trivial on $\{\pm I\}$.

On the other hand, consider an irreducible representation ρ of SU(2) with $\rho(-I) = \rho(I)$. Since $\ker(\Phi) = \{\pm I\}$, the map Φ factors through the quotient and defines an isomorphism $\Phi_r : \mathrm{SU}(2)/\{\pm I\} \to \mathrm{SO}(3)$. Thus, $\rho \circ \Phi_r^{-1}$ is well-defined on SO(3) and induces an irreducible representation of SO(3). This establishes a bijection between the irreducible representations of SO(3) and those of SU(2) that are trivial on $\{\pm I\}$.

7.2 Irreducible representations of SU(2)

Consider the canonical grading of the bivariate polynomial ring $\mathbb{C}[z_1, z_2]$

$$\mathbb{C}[z_1, z_2] = \bigoplus_{k=0}^{\infty} V_{k/2},$$

where $V_{k/2}$ is the subspace of $\mathbb{C}[z_1, z_2]$ consisting of homogeneous polynomials of degree k. We adopt the half-integer convention $\ell = k/2$, commonly used in the literature. Throughout this chapter, sums over half-integers increase in steps of one, i.e. a sum ranging from $-3/2$ to $3/2$ uses the indices $-3/2, -1/2, 1/2, 3/2$. The same convention is used for index ranges of the form $a, \ldots, b$ with half-integer endpoints.

We equip each space V_ℓ with the inner product

$$\langle p, q \rangle = \overline{p}(\nabla)\, q(z),$$

where $p(\nabla)$ is the polynomial differential operator obtained by replacing each z_i by $\partial/\partial z_i$. An orthonormal basis of V_ℓ is given by

$$e_n(z) = ((\ell - n)!(\ell + n)!)^{-1/2}\, z_1^{\ell - n} z_2^{\ell + n}$$

for $n = -\ell, \ldots, \ell$.

We define an action of SU(2) on a polynomial $p \in V_\ell$ by

$$(U \cdot p)(z) = p\left(U^* z\right). \tag{7.4}$$

This action induces a representation on V_ℓ, which we denote by ρ_ℓ.

Lemma 7.3. *For any nonnegative half-integer ℓ, the representation ρ_ℓ defined by* (7.4) *is unitary.*

Proof. Consider the kernel $K\colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ given by

$$K(z,w) = \sum_{n=-\ell}^{\ell} \overline{e_n(z)}\, e_n(w).$$

Evaluating K at $z, w \in \mathbb{C}^2$ yields

$$K(z,w) = \sum_{n=-\ell}^{\ell} \frac{(\overline{z_1} w_1)^{\ell-n} (\overline{z_2} w_2)^{\ell+n}}{(\ell+n)!(\ell-n)!} = \frac{(z^* w)^{2\ell}}{(2\ell)!},$$

which implies that the kernel is invariant under the action of $U \in \mathrm{SU}(2)$ defined by $U \cdot K(z,w) = K(U^* z, U^* w)$. For a fixed $U \in \mathrm{SU}(2)$, we expand $U \cdot K(z,w)$ as

$$\begin{aligned}
U \cdot K(z,w) &= \sum_{n=-\ell}^{\ell} \overline{e_n(U^* z)}\, e_n(U^* w) \\
&= \sum_{n=-\ell}^{\ell} \left(\sum_{m=-\ell}^{\ell} [\rho_\ell(U)^*]_{m,n}\, \overline{e_m(z)} \right) \left(\sum_{m'=-\ell}^{\ell} [\rho_\ell(U)]_{n,m'}\, e_{m'}(w) \right) \\
&= \sum_{m,\, m'=-\ell}^{\ell} [\rho_\ell(U)^* \rho_\ell(U)]_{m,m'}\, \overline{e_m(z)}\, e_{m'}(w).
\end{aligned}$$

Comparing coefficients of $K(z,w)$ and $U \cdot K(z,w)$ with respect to the linearly independent kernels $\overline{e_m(z)}\, e_{m'}(w)$, we get $\rho_\ell(U)^* \rho_\ell(U) = I$. □

7.2.1 Explicit formulas

To proceed, we derive an explicit expression for the matrix elements of the representation ρ_ℓ, i.e., we compute

$$[\rho_\ell(U)]_{m,n} = \langle e_m, U \cdot e_n \rangle.$$

To this end, we consider the univariate polynomial obtained by restriction $(U \cdot e_n)(z, 1)$. The coefficient of the monomial $z^{\ell-m}$ in this expression agrees, up to normalization, with $\langle e_m, U \cdot e_n \rangle$. Hence, the matrix elements can be deduced from the Taylor expansion at $z = 0$.

Using the definition of the action (7.4), together with the explicit form of U from (7.1), we obtain

$$[\rho_\ell(U)]_{m,n} = C_{m,n} \frac{d^{\ell-m}}{dz^{\ell-m}} \left(\overline{a}z + \overline{b}\right)^{\ell-n} (-bz+a)^{\ell+n} \Big|_{z=0} \tag{7.5}$$

where

$$C_{m,n} = \left(\frac{(\ell+m)!}{(\ell-n)!(\ell+n)!(\ell-m)!} \right)^{1/2}.$$

Decomposing U in terms of Euler angles leads to the following factorization of the representation

$$\rho_\ell\left(U(\alpha,\beta,\gamma)\right) = \rho_\ell\left(U_Z(\gamma)\right) \rho_\ell\left(U_X(\beta)\right) \rho_\ell\left(U_Z(\alpha)\right). \tag{7.6}$$

Consequently, it suffices to determine the matrices corresponding to $U_Z(\theta)$ and $U_X(\theta)$.

Applying (7.5) to $U_Z(\theta)$ yields a diagonal matrix. More precisely,

$$[\rho_\ell\left(U_Z(\theta)\right)]_{m,m} = e^{im\theta}. \tag{7.7}$$

As a result, (7.6) simplifies to

$$[\rho_\ell\left(U(\alpha,\beta,\gamma)\right)]_{m,n} = e^{i(n\alpha+m\gamma)} \left[\rho_\ell\left(U_X(\beta)\right)\right]_{m,n}.$$

To compute $\rho_\ell\left(U_X(\beta)\right)$, we substitute $t = \cos(\beta)$, which allows us to express $U_X(\beta)$ as $W(\cos(\beta))$ where

$$W(t) = \begin{pmatrix} ((1+t)/2)^{1/2} & i\,((1-t)/2)^{1/2} \\ i\,((1-t)/2)^{1/2} & ((1+t)/2)^{1/2} \end{pmatrix}.$$

Using the expression (7.5), one verifies that the matrix elements satisfy the symmetries

$$[\rho_\ell(W(t))]_{m,n} = \overline{[\rho_\ell(W(t)^*)]_{m,n}} = \overline{[\rho_\ell(W(t))^*]_{m,n}} = [\rho_\ell(W(t))]_{n,m},$$

as well as

$$[\rho_\ell(W(t))]_{m,n} = (-1)^\ell\, \overline{[\rho_\ell(W(-t))]_{m,-n}}.$$

These symmetries allow us to restrict attention to indices satisfying $m \geq |n|$, since all other entries can be recovered from these.

The matrix $W(t)$ is of the form (7.1) where

$$a = ((1+t)/2)^{1/2} \qquad \text{and} \qquad b = i\,((1-t)/2)^{1/2}.$$

First, we assume that $t \in (-1, 1)$, so that $a, b \neq 0$. We substitute $x = |a|^2 - |b|^2 - 2\overline{a}bz$. A direct computation using $|a|^2 + |b|^2 = 1$ gives

$$\overline{a}z + \overline{b} = (1 - x)/(2b) \qquad \text{and} \qquad -bz + a = (1 + x)/(2\overline{a}).$$

Substituting into (7.5), we obtain

$$(-1)^{\ell-m}\, 2^{-\ell-m}\, C_{m,n}\, (\overline{a})^{-m-n}\, b^{n-m}\, \frac{d^{\ell-m}}{dx^{\ell-m}}\, (1-x)^{\ell-n}\, (1+x)^{\ell+n}\bigg|_{x=t}.$$

The *Rodrigues formula* [1, Section 2.5] relates the expression above to the *Jacobi polynomials.* We denote the Jacobi polynomial of degree k with respect to the parameters α and β in the standard normalization by $P_k^{(\alpha,\beta)}$. For an introduction to Jacobi polynomials, we refer to [1, Chapters 2 and 6].

Applying the Rodrigues formula and evaluating at $t = \cos(\beta)$ yields

$$i^{n-m}\, C'_{m,n} \sin(\beta/2)^{m-n} \cos(\beta/2)^{m+n}\, P_{\ell-m}^{(m-n,m+n)}(\cos(\beta)), \tag{7.8}$$

with

$$C'_{m,n} = (\ell - m)!\, C_{m,n} = \left(\frac{(\ell-m)!(\ell+m)!}{(\ell-n)!(\ell+n)!}\right)^{1/2}.$$

The derivation above assumes $m \geq |n|$ as well as $t \in (-1, 1)$. However, a direct verification shows that the expression (7.8) extends to the boundary values $t = \pm 1$, while the remaining matrix elements are determined by the symmetries established earlier.

In classical notation, we define

$$P^{\ell}_{m,n}(\cos(\beta)) = C'_{m,n} \sin(\beta/2)^{m-n} \cos(\beta/2)^{m+n}\, P_{\ell-m}^{(m-n,m+n)}(\cos(\beta)).$$

Combining this with the Euler-angle decomposition (7.6), we arrive at the explicit formula

$$[\rho_\ell(U(\alpha, \beta, \gamma))]_{m,n} = i^{n-m}\, e^{i(n\alpha+m\gamma)}\, P^{\ell}_{m,n}(\cos(\beta)).$$

We conclude by giving an explicit description of continuous functions of positive type on SU(2) and SO(3). Irreducibility and completeness of the constructed representations are established in the subsequent section. For the present discussion, we proceed under this assumption.

As shown in Section 7.1, the irreducible representations of SO(3) are precisely those representations of SU(2) that are trivial on $\{\pm I\}$. From (7.5) we deduce that $\rho_\ell(-I) = \rho_\ell(I)$ holds if and only if ℓ is a nonnegative integer. Bochner's theorem yields the following characterization.

Theorem 7.4. *A function $f : \mathrm{SU}(2) \to \mathbb{C}$ parameterized by Euler angles is continuous and of positive type if and only if for every nonnegative half-integer ℓ there exists a Hermitian positive semidefinite matrix F^ℓ such that*

$$f(\alpha,\beta,\gamma) = \sum_{\ell\in\mathbb{N}_0/2} \sum_{m,n=-\ell}^{\ell} \left[F^\ell\right]_{n,m} i^{n-m}\, e^{i(n\alpha+m\gamma)}\, P^\ell_{m,n}(\cos(\beta)),$$

as well as $\sum_{\ell\in\mathbb{N}_0/2} \operatorname{tr}(F^\ell) < \infty$. The latter guarantees absolute convergence of the series.

A function $f : \mathrm{SO}(3) \to \mathbb{C}$ parameterized by Euler angles is continuous and of positive type if and only if for every nonnegative integer ℓ there exists a Hermitian positive semidefinite matrix F^ℓ such that

$$f(\alpha,\beta,\gamma) = \sum_{\ell\in\mathbb{N}_0} \sum_{m,n=-\ell}^{\ell} \left[F^\ell\right]_{n,m} i^{n-m}\, e^{i(n\alpha+m\gamma)}\, P^\ell_{m,n}(\cos(\beta)),$$

with $\sum_{\ell\in\mathbb{N}_0} \operatorname{tr}(F^\ell) < \infty$. The latter guarantees absolute convergence of the series.

7.2.2 Completeness and irreducibility

To justify Theorem 7.4, we show that the family of representations ρ_ℓ defined by the action (7.4) forms a complete set of irreducible unitary representations. This section follows the exposition of Bröcker and tom Dieck in [14, Section 2.5].

Theorem 7.5. *For a given nonnegative half-integer ℓ, the representation ρ_ℓ associated with the action* (7.4) *is irreducible.*

Proof. Let $A : V_\ell \to V_\ell$ be a linear ρ_ℓ-equivariant map for a given nonnegative half-integer ℓ. We show that A is a scalar multiple of the identity. Recall from (7.7) that $\rho_\ell(U_Z(\theta))$ is a diagonal matrix whose entries are given by $[\rho_\ell(U_Z(\theta))]_{m,m} = e^{im\theta}$. Consequently, every matrix that commutes with $\rho_\ell(U_Z(\theta))$ must be diagonal.

On the other hand, all matrix coefficients of $\rho_\ell(U_X(\theta))$ are nontrivial functions of θ. Hence, any diagonal matrix commuting with $\rho_\ell(U_X(\theta))$ must be a scalar multiple of the identity. $\square$

For a half-integer ℓ, denote the character of ρ_ℓ by χ_ℓ. By diagonalizing $U \in \mathrm{SU}(2)$ one obtains a unitary matrix Q and $\theta \in [-\pi, \pi)$

such that $U = Q\,U_Z(2\theta)\,Q^*$. Since the trace is cyclic, we have $\chi_\ell(U) = \chi_\ell(U_Z(2\theta))$. Equation (7.7) allows us to compute the characters of ρ_ℓ explicitly

$$\chi_\ell(U_Z(2\theta)) = \sum_{m=-\ell}^{\ell} e^{2im\theta} = \frac{\sin\left((2\ell+1)\,\theta\right)}{\sin\left(\theta\right)}.$$

From $e^{2im\theta} + e^{-2im\theta} = 2\cos(2m\theta)$ we conclude that the functions $\chi_{j/2}(U_Z(2\theta))$ with $j = 0, \ldots, 2\ell$ span the same vector space as the trigonometric functions $1, \cos(\theta), \ldots, \cos(2\ell\theta)$. Hence, the linear span of the associated even 2π-periodic functions arising from the characters is uniformly dense in the space of even continuous 2π-periodic functions. Note that the family of functions $\{\sin((2\ell+1)\theta)/\sin\theta \,:\, \ell \in \mathbb{N}_0/2\}$ is orthogonal with respect to the inner product

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\,g(\theta)\,\sin\left(\theta\right)^2\,d\theta.$$

Let f be a class function. Expanding $f(U)$ and $f(U_Z(\theta))$ in terms of the character basis and in terms of the associated 2π-periodic functions, respectively, yields

$$\int_{\mathrm{SU}(2)} f(U)\,d\mu(U) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(U_Z(2\theta))\,\sin(\theta)^2\,d\theta,$$

where μ is the Haar measure on SU(2).

Theorem 7.6. *Every irreducible unitary representation is isomorphic to one of the representations ρ_ℓ defined by the action* (7.4).

Proof. Assume that there exists an irreducible representation that is not isomorphic to one of the representations ρ_ℓ. The character of that representation is orthogonal to every character χ_ℓ. Hence, there exists a nonzero even continuous 2π-periodic function that is orthogonal to every $\chi_\ell(U_Z(2\theta))$. This contradicts the density of the cosine functions in the space of even continuous 2π-periodic functions. $\square$

Chapter Eight

Computing the theta number

The set of rotations that satisfy the interior-disjointness constraint in Theorem 6.1 gives rise to an infinite family of linear constraints. Polynomial optimization provides a natural framework for treating such infinite constraint families and has become a central tool in modern geometric optimization. A survey of geometric optimization problems and their formulation via polynomial techniques is given in [79].

In the present setting, trigonometric polynomials arise naturally. Dumitrescu developed an adaptation of polynomial optimization to the trigonometric case [35]. In Section 8.1, we summarize the relevant aspects, following [35, Chapters 3–4].

Section 8.2 investigates the impact of symmetries on trigonometric polynomial optimization problems. Our approach is based on the theory of symmetry-adapted optimization of ordinary polynomials, developed by Gatermann and Parrilo [40]. Although symmetries in trigonometric polynomial optimization have previously been considered, for example, by Metzlaff [57, 58], our treatment employs a different group action, leading to a distinct formulation of invariance.

We conclude the chapter by formulating the theta number as a trigonometric polynomial optimization problem and subsequently exploiting the symmetries of regular spherical polygons to simplify the resulting problem.

8.1 Trigonometric polynomial optimization

Let $\theta = (\theta_1, \dots, \theta_n)$ be a vector of indeterminates. The Laurent ring of *trigonometric polynomials* is given by $\mathbb{C}\left[e^{\pm i\theta_1}, \dots, e^{\pm i\theta_n}\right]$. A polynomial $p \in \mathbb{C}\left[e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_n}\right]$ can be expressed as

$$p(\theta) = \sum_{k \in \mathbb{Z}^n} \lambda_k \, e^{ik^\top \theta}, \tag{8.1}$$

where $\lambda_k \in \mathbb{C}$ and $\lambda_k = 0$ for all but finitely many k. A trigonometric polynomial is called *Hermitian* if it is real-valued. Note that a trigonometric polynomial is Hermitian if and only if its coefficients in (8.1) satisfy $\lambda_{-k} = \overline{\lambda_k}$.

The *partial degree* in θ_j of a trigonometric polynomial is defined by

$$\deg_j(p) = \max\left\{|k_j| : \; k \in \mathbb{Z}^n, \, \lambda_k \neq 0\right\}.$$

The *degree* of a trigonometric polynomial p is the vector $[\deg_j(p)]_{j=1}^n$.

A Hermitian trigonometric polynomial σ is a *sum-of-squares* if there exists a vector $b(\theta)$ of trigonometric polynomials and a Hermitian positive semidefinite matrix Q such that

$$\sigma(\theta) = b(\theta)^* Q \, b(\theta). \tag{8.2}$$

Such vectors b are referred to as sum-of-squares bases for σ. After a linear change of basis, the sum-of-squares basis may be chosen to consist of the standard trigonometric monomials $e^{ik^\top \theta}$, i.e., for every sum-of-squares polynomial σ there exists a Hermitian positive semidefinite matrix Q such that

$$\sigma(\theta) = \sum_{k,\ell \in \mathbb{Z}^n} Q_{k,\ell} \, e^{i(k-\ell)^\top \theta}. \tag{8.3}$$

Given Hermitian trigonometric polynomials $q_1, \ldots, q_m$, a set of the form

$$\Sigma = \{\theta \in [-\pi, \pi)^n : \; q_1(\theta) \geq 0, \ldots, q_m(\theta) \geq 0\} \tag{8.4}$$

is called a *semialgebraic set*. We refer to the polynomials $q_1(\theta), \ldots, q_m(\theta)$ as the *generators* of Σ.

For semialgebraic sets $\Sigma_1, \ldots, \Sigma_m$, the problem

$$\begin{array}{lll} \min & p(0) & \\ \text{s.t.} & p & \text{Hermitian trigonometric polynomial} \\ & p(\theta) \geq 0 & \text{for all } \theta \in \Sigma_1 \cup \ldots \cup \Sigma_m \end{array} \tag{8.5}$$

is called a *trigonometric polynomial optimization problem.*

To construct feasible solutions for trigonometric polynomial optimization problems, we will make use of sum-of-squares polynomials. For a semialgebraic set as in (8.4), consider the Hermitian trigonometric polynomial

$$p(\theta) = \sigma_0(\theta) + \sum_{j=1}^{m} \sigma_j(\theta)\, q_j(\theta), \tag{8.6}$$

where $\sigma_0, \ldots, \sigma_m$ are sum-of-squares polynomials. Since sum-of-squares polynomials are nonnegative by construction, p is nonnegative on Σ. In fact, every trigonometric polynomial that is strictly positive on a semialgebraic set admits a sum-of-squares representation, which constitutes a key distinction from classical polynomial optimization.

Theorem 8.1. *Let $q_1, \ldots, q_m$ be Hermitian trigonometric polynomials that generate the semialgebraic set*

$$\Sigma = \{\theta \in [-\pi, \pi)^n : q_1(\theta) \geq 0, \ldots, q_m(\theta) \geq 0\}.$$

If p is a Hermitian trigonometric polynomial that is strictly positive on Σ, then there exist trigonometric sum-of-squares polynomials $\sigma_0, \ldots, \sigma_m$ such that

$$p(\theta) = \sigma_0(\theta) + \sum_{j=1}^{m} \sigma_j(\theta) q_j(\theta).$$

The proof of Theorem 8.1 relies on classical results from real algebraic geometry. Let $\mathbb{R}[x]$ denote the polynomial ring in the variables $x = (x_1, \ldots, x_n)$. In analogy to the trigonometric case, a polynomial $\sigma \in \mathbb{R}[x]$ is a sum-of-squares if there exists a finite vector of polynomials $b(x)$ and a symmetric positive semidefinite matrix Q such that

$$\sigma(x) = b(x)^\top Q\, b(x).$$

Given $q_1, \ldots, q_m \in \mathbb{R}[x]$, the *quadratic module* generated by these polynomials is defined as

$$\mathcal{QM}(q_1, \ldots, q_m) = \left\{ \sigma_0 + \sum_{j=1}^{m} q_j \sigma_j : \ \sigma_0, \ldots, \sigma_m \text{ are sum-of-squares} \right\}.$$

A quadratic module is *Archimedean* if there exists $N \in \mathbb{N}$ such that

$$N - \sum_{j=1}^{n} x_j^2 \in \mathcal{QM}(q_1, \ldots, q_m).$$

Theorem 8.2 (Putinar's Positivstellensatz). *Let $q_1, \ldots, q_m \in \mathbb{R}[x]$ such that the quadratic module $\mathcal{QM}(q_1, \ldots, q_m)$ is Archimedean. A polynomial $p \in \mathbb{R}[x]$ that is strictly positive on*

$$\{x \in \mathbb{R}^n : \; q_1(x) \geq 0, \ldots, q_m(x) \geq 0\}$$

belongs to $\mathcal{QM}(q_1, \ldots, q_m)$.

For a proof of Theorem 8.2, the reader is referred to the original work of Putinar [63].

Proof of Theorem 8.1. To apply Putinar's Positivstellensatz to Hermitian trigonometric polynomials, we perform a change of variables. A Hermitian trigonometric polynomial can be expressed in terms of sine and cosine using

$$\begin{aligned}
p(\theta) &= \sum_{k \in \mathbb{Z}^n} \lambda_k \, e^{ik^\top \theta} \\
&= \frac{1}{2} \left(\sum_{k \in \mathbb{Z}^n} \lambda_k \, e^{ik^\top \theta} + \overline{\lambda_k \, e^{ik^\top \theta}} \right) \\
&= \sum_{k \in \mathbb{Z}^n} \operatorname{Re}(\lambda_k) \, \cos\left(k^\top \theta\right) - \operatorname{Im}(\lambda_k) \, \sin\left(k^\top \theta\right).
\end{aligned}$$

Substituting $x_j = \cos(\theta_j)$, $y_j = \sin(\theta_j)$ and using elementary trigonometric identities, we obtain a real polynomial P in the variables $(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n)$ such that

$$p(\theta) = P(\cos(\theta_1), \ldots, \cos(\theta_n), \sin(\theta_1), \ldots, \sin(\theta_n)).$$

Consider a semialgebraic set Σ given as in (8.4) and a trigonometric polynomial p which is positive on Σ. The positivity of p on Σ is equivalent to the positivity of P on

$$\begin{aligned}
\big\{(x, y) \in \mathbb{R}^{2n} : \; Q_1(x, y) \geq 0, \ldots, Q_m(x, y) \geq 0, \\
x_1^2 + y_1^2 = 1, \ldots, x_n^2 + y_n^2 = 1\big\},
\end{aligned}$$

where $Q_1, \ldots, Q_m$ are obtained from $q_1, \ldots, q_m$ by the same substitution. The relation

$$x_j^2 + y_j^2 = 1 \quad \text{if and only if} \quad 1 - x_j^2 - y_j^2 \geq 0 \text{ and } x_j^2 + y_j^2 - 1 \geq 0$$

shows that the polynomial

$$n - \left(\sum_{j=1}^{n} x_j^2 + y_j^2 \right)$$

lies in the associated quadratic module. Thus the quadratic module is Archimedean and the claim follows from Putinar's Positivstellensatz. □

After fixing a degree bound, solving the corresponding sum-of-squares relaxation of a trigonometric polynomial optimization problem reduces to semidefinite programming over Hermitian positive semidefinite matrices. By separating real and imaginary parts, this formulation can equivalently be expressed as a semidefinite program over real symmetric matrices of twice the dimension, see [81]. Under mild assumptions, semidefinite programs can be solved efficiently [23, 44].

On the other hand, the size of the underlying semidefinite program scales with the polynomial degree required for an exact or sufficiently accurate sum-of-squares certificate. In fact, classical polynomial optimization is NP-hard in general, and many of these hardness results can be adapted to the trigonometric setting. In particular, Blum et al. showed that the feasibility problem for polynomial optimization is NP-hard [10, Chapter 5.4], and O'Donnell exhibited examples requiring exponential bit complexity [61].

The trigonometric setting presents an additional structural difficulty. Cancellations among monomial terms imply that the degree is not multiplicative. As a result, controlling the degree required for sum-of-squares representations is subtle, and even problems of apparently low degree may give rise to large-scale semidefinite programs. Concrete examples illustrating this phenomenon can be found in [35, Chapter 3.5.2]. This behavior is consistent with the absence of uniform degree bounds for sum-of-squares representations on compact algebraic varieties, as shown by Scheiderer [64].

8.2 Invariants in trigonometric polynomial optimization

Given a vector of variables $\theta = (\theta_1, \ldots, \theta_n)$, consider the trigonometric polynomial ring $\mathbb{C}\left[e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_n}\right]$. Define a ring isomorphism by

$$\begin{aligned}\Phi : \mathbb{C}\left[e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_n}\right] &\to \mathbb{C}[z, \overline{z}]/\mathfrak{I}_n \\ e^{i\theta_j} &\mapsto z_j && \text{for } j = 1, \ldots, n, \\ e^{-i\theta_j} &\mapsto \overline{z}_j && \text{for } j = 1, \ldots, n,\end{aligned}$$

where $\mathbb{C}[z, \overline{z}]$ is the polynomial ring in $2n$ independent variables and the ideal $\mathfrak{I}_n$ is given by

$$\mathfrak{I}_n = \langle z_j \, \overline{z}_j - 1, \; j = 1, \ldots, n \rangle.$$

For notational convenience, we write $\xi = (z, \overline{z})$.

Consider a finite group Γ acting on $\mathbb{C}[\xi]$ via a representation ρ of appropriate dimension, i.e.,

$$(\gamma \cdot P)(\xi) = P(\rho(\gamma^{-1}) \, \xi).$$

This induces an action on trigonometric polynomials via

$$\gamma \cdot p = \Phi^{-1}(\gamma \cdot \Phi(p)).$$

For such an action to be well-defined, we require that the ideal $\mathfrak{I}_n$ is invariant under the action of Γ, i.e.,

$$\langle \left[\rho(\gamma^{-1}) \, \xi\right]_j \; \left[\rho(\gamma^{-1}) \, \xi\right]_{n+j} - 1, \, j = 1, \ldots, n \rangle \subseteq \mathfrak{I}_n$$

for all $\gamma \in \Gamma$. A trigonometric polynomial p with $\gamma \cdot p = p$ is called *invariant.*

We restrict our attention to group actions that leave the space of Hermitian trigonometric polynomials invariant. In terms of the matrix representation, this means that, for every $\gamma \in \Gamma$, the matrix $\rho(\gamma)$ admits a block decomposition

$$\rho(\gamma) = \begin{pmatrix} \rho(\gamma)_{11} & \rho(\gamma)_{12} \\ \overline{\rho(\gamma)_{12}} & \overline{\rho(\gamma)_{11}} \end{pmatrix},$$

where $\rho(\gamma)_{11}, \, \rho(\gamma)_{12} \in \mathbb{C}^{n \times n}$.

Theorem 8.3. *Let Γ be a finite group and let ρ be a representation of Γ inducing a well-defined action on trigonometric polynomials under which the Hermitian polynomials form an invariant space. For every $\gamma \in \Gamma$, each row of the matrix $\rho(\gamma)$ contains exactly one nonzero entry, which is of modulus 1. Moreover, $\rho(\gamma)$ is unitary.*

Proof. Let ρ be a representation of Γ satisfying the assumptions of the theorem and let $\gamma \in \Gamma$. Denote the j-th row of $\rho(\gamma^{-1})_{11}$ and of $\rho(\gamma^{-1})_{12}$ by v, w, respectively. Since $\mathfrak{I}_n$ is invariant under the action of Γ, we deduce

$$\begin{aligned} 1 &\equiv (z^\top v + \overline{z}^\top w)(w^* z + v^* \overline{z}) \\ &= z^\top (v\, w^*) z + z^\top (v v^*)\, \overline{z} + \overline{z}^\top (w w^*)\, z + \overline{z}^\top (w v^*) \overline{z} \end{aligned}$$

modulo $\mathfrak{I}_n$. Reducing the right-hand side modulo $\mathfrak{I}_n$ and comparing coefficients yields

$$\begin{aligned} v_k\, \overline{w_\ell} + v_\ell\, \overline{w_k} &= 0 \qquad \text{for } k, \ell = 1, \ldots, n, \\ v_k\, \overline{v_\ell} + w_\ell\, \overline{w_k} &= 0 \qquad \text{for } k, \ell = 1, \ldots, n \text{ with } k \neq \ell, \\ \sum_{k=1}^{n} \left(|v_k|^2 + |w_k|^2 \right) &= 1. \end{aligned}$$

These equations imply that the row (v, w) contains exactly one nonzero entry that is of modulus 1.

Thus, each row contains exactly one entry of absolute value one, with these entries occurring in distinct columns, and the remaining entries vanish. This implies that the rows form an orthonormal system. Hence, $\rho(\gamma^{-1})$ is unitary. □

Theorem 8.3 implies that the only group actions under consideration are of the form

$$(\gamma \cdot p)(\theta_1, \ldots, \theta_n) = p(\pm\theta_{\tau(1)} + \phi_1, \ldots, \pm\theta_{\tau(n)} + \phi_n) \qquad (8.7)$$

where τ is an element of the symmetric group $\mathrm{Sym}(n)$ and ϕ lies in the n-dimensional torus $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$. Consequently, Γ is isomorphic to a subgroup of the semidirect product

$$(\mathrm{Sym}(n) \ltimes (\mathbb{Z}/2\mathbb{Z})^n) \ltimes (\mathbb{R}/(2\pi\mathbb{Z}))^n.$$

This yields an induced action of the group Γ on the n-dimensional torus $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$. We call a subset of $[-\pi, \pi)^n$ *invariant* if its image under the canonical projection onto the torus is invariant under the induced group action on the torus.

When characterizing invariant sum-of-squares polynomials, it seems natural to assume that there exists a basis consisting of invariant functions. This assumption is false in general, as the following example demonstrates.

Example 8.4. Consider the ring of univariate Hermitian trigonometric polynomials invariant under the action $\theta \mapsto -\theta$. This ring is generated by $\cos(\theta)$. The sum-of-squares polynomial $\sigma(\theta) = \sin(\theta)^2$ is invariant under the action, but cannot be represented as a sum-of-squares polynomial in a cosine basis.

Hence, we require a more sophisticated way to construct a symmetry-adapted sum-of-squares basis. Let Γ be a group acting as in Theorem 8.3 and let σ be an invariant sum-of-squares polynomial of the form

$$\sigma(\theta) = b(\theta)^* Q\, b(\theta),$$

where $b(\theta)$ denotes the standard trigonometric monomial basis of length d and $Q \in \mathbb{C}^{d \times d}$ is Hermitian positive semidefinite. Theorem 8.3 guarantees the existence of a unitary representation $U(\gamma)$ such that

$$\gamma \cdot \sigma(\theta) = b(\theta)^* U(\gamma)^* Q\, U(\gamma)\, b(\theta),$$

for all $\gamma \in \Gamma$.

This induces a group action of the opposite group of Γ on the space of Hermitian matrices $\gamma \cdot Q = U(\gamma)^* Q\, U(\gamma)$. The opposite group is defined on the same underlying set as Γ, but with the group operation applied in the reverse order. Since the action preserves semidefiniteness, an averaging argument allows us to assume that the matrix Q is invariant. Moreover, the opposite group of Γ acts on $\mathbb{C}^d$ by $\gamma \cdot v = U(\gamma)^* v$.

If ρ is a unitary representation of Γ, its Hermitian conjugate ρ^* defines a representation of its opposite group. This construction yields a bijection between representations of Γ and representations of its opposite group. To avoid switching between Γ and its opposite group, we exclusively consider representations of Γ in the following.

Recall from Chapter 2 that the action of the opposite group decomposes $\mathbb{C}^d$ into invariant subspaces

$$\bigoplus_{\rho \in \hat{\Gamma}} \bigoplus_{k=1}^{m_\rho} V_{\rho,k},$$

where $V_{\rho,k}$ are irreducible subspaces such that $V_{\rho,k}$ and $V_{\rho',k'}$ are equivalent if and only if $\rho = \rho'$ and m_ρ denotes the multiplicity of the corresponding representation. Hence, the space of complex-valued matrices can be decomposed as follows

$$\begin{aligned}\mathbb{C}^d \otimes \left(\mathbb{C}^d\right)^* &= \left(\bigoplus_{\rho\in\hat{\Gamma}} \bigoplus_{k=1}^{m_\rho} V_{\rho,k}\right) \otimes \left(\bigoplus_{\rho'\in\hat{\Gamma}} \bigoplus_{k'=1}^{m_{\rho'}} V_{\rho',k'}\right)^* \\ &= \bigoplus_{\rho,\rho'\in\hat{\Gamma}} \bigoplus_{k=1}^{m_\rho} \bigoplus_{k'=1}^{m_{\rho'}} V_{\rho,k} \otimes (V_{\rho',k'})^*,\end{aligned}$$

where the dual space of a vector space V is denoted by V^*. By Schur's lemma, the space of Γ-equivariant linear maps between $V_{\rho,k}$ and $V_{\rho',k'}$ is zero if $\rho \neq \rho'$, and consists of scalar multiples of the identity if $\rho = \rho'$. Thus, the space of invariant matrices decomposes as

$$\bigoplus_{\rho\in\hat{\Gamma}} I_{d_\rho} \otimes \mathbb{C}^{m_\rho \times m_\rho}$$

where d_ρ is the dimension of the representation ρ. Therefore, we obtain a block-diagonalization of the matrix Q

$$\bigoplus_{\rho\in\hat{\Gamma}} Q_\rho,$$

where every block Q_ρ can be further decomposed into d_ρ identical blocks.

It is vital to understand how the change of basis affects the action of Γ on the sum-of-squares polynomial σ. Let ρ be an irreducible unitary representation of degree d_ρ, and let $e_{\rho,k,1},\ldots,e_{\rho,k,d_\rho}$ be an orthonormal basis of $V_{\rho,k}$. Consider the vector-valued trigonometric polynomial

$$h_{\rho,k}(\theta) = \left[(e_{\rho,k,\ell})^* b(\theta)\right]_{\ell=1}^{d_\rho}.$$

By construction, the componentwise action of Γ on $h_{\rho,k}$ is given by

$$\gamma \cdot h_{\rho,k} = \rho(\gamma)\, h_{\rho,k}.$$

In analogy to linear functions, we call $h_{\rho,k}$ a ρ-equivariant trigonometric polynomial with respect to the group action of Γ.

Using this change of basis, σ can be expressed as

$$\sigma(\theta) = \sum_{\rho\in\hat{\Gamma}} h_\rho(\theta)^* Q_\rho\, h_\rho(\theta),$$

where $h_\rho(\theta)$ are $(\rho \otimes I)$-equivariant polynomials. By computing, for each irreducible representation of Γ, a basis of appropriate degree for the corresponding equivariant vector space, we can reduce the size of matrices in a sum-of-squares decomposition. This significantly reduces the computational time required to solve trigonometric polynomial optimization problems numerically.

To compute such a basis, consider a finite group Γ with a well-defined action on trigonometric polynomials such that the space of Hermitian polynomials is invariant, and let ρ be a d_ρ-dimensional unitary representation of Γ. We define a projection from the space of d_ρ-dimensional vector-valued trigonometric polynomials onto the space of ρ^*-equivariant vector-valued trigonometric polynomials

$$p \mapsto \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \rho(\gamma)\,(\gamma \cdot p).$$

This is an adaptation of the projection operator for ordinary polynomials, as described in [39]. A projection onto the space of ρ-equivariant vector-valued trigonometric polynomials is obtained via the opposite group.

An equivariant basis up to a given degree can be computed by applying the projection operator corresponding to each irreducible representation of Γ to the standard monomial basis.

At first sight, the appearance of the opposite group may seem artificial. However, it is a consequence of the chosen convention for the action on polynomials. In many situations, it is more convenient to let a group element act by $(\gamma \cdot P)(\xi) = P(\rho(\gamma)\xi)$ rather than by $(\gamma \cdot P)(\xi) = P(\rho(\gamma^{-1})\xi)$. The former defines an action of the opposite group and therefore leads to equivariants transforming according to the irreducible representations of the original group Γ.

Since the groups relevant in this chapter are abelian, they are canonically isomorphic to their opposite groups. We therefore suppress this distinction in what follows.

We now discuss two examples of symmetry reduction that are particularly relevant for our application.

Example 8.5. We revisit Example 8.4 in a multivariate setting. Consider the action of $\mathbb{Z}/2\mathbb{Z}$ given by $\theta \mapsto -\theta$. Since the group is abelian, the irreducible representations are given by the characters $\chi_0(k) = 1$ and $\chi_1(k) = (-1)^k$ for $k \in \mathbb{Z}/2\mathbb{Z}$. By applying the projection operators corresponding to both representations to the monomial basis, one obtains the sum-of-squares bases

$$b_0(\theta) = [\cos(k^\top \theta)]_{k\in\mathbb{Z}^n} \quad \text{and} \quad b_1(\theta) = [\sin(k^\top \theta)]_{k\in\mathbb{Z}^n}.$$

Here, we scaled the sine terms to ensure that their coefficients are real.

A Γ-invariant sum-of-squares polynomial can be decomposed as

$$\sigma(\theta) = b_0(\theta)^\top Q_0\, b_0(\theta) + b_1(\theta)^\top Q_1\, b_1(\theta).$$

In particular, the sum-of-squares basis is real-valued. Consequently,

$$\begin{aligned}\sigma(\theta) &= (1/2)\left(\sigma(\theta) + \overline{\sigma(\theta)}\right)\\ &= (1/2)\left(b_0(\theta)^\top \left(Q_0 + \overline{Q_0}\right) b_0(\theta) + b_1(\theta)^\top \left(Q_1 + \overline{Q_1}\right) b_1(\theta)\right)\end{aligned}$$

yields a sum-of-squares decomposition consisting of real symmetric matrices.

The transition from complex to real matrices additionally simplifies numerical computations since Hermitian positive semidefinite matrices are represented by real positive semidefinite matrices of twice their size. A real-valued trigonometric basis can be obtained whenever the symmetry group contains an action of the form $\theta \mapsto -\theta + \pi\, b$ with $b \in \{0, 1\}^n$.

Example 8.6. For a given integer $N \in \mathbb{N}$, the group $\mathbb{Z}/N\mathbb{Z}$ acts on trigonometric polynomials in n variables by

$$k \cdot p(\theta_1, \ldots, \theta_n) = p(\theta_1 + 2\pi k/N, \theta_2, \ldots, \theta_n).$$

The group action affects only one variable, so we omit the dependence on $(\theta_2, \ldots, \theta_n)$ for the remainder of this example and denote θ_1 by θ. Since the group is abelian, the irreducible representations are the characters $\chi_\ell(k) = e^{2\pi i k\ell/N}$ for $k, \ell \in \mathbb{Z}/N\mathbb{Z}$. Applying the projection operator corresponding to the ℓ-th character to the standard monomial basis vector yields

$$b_\ell(\theta) = e^{i\ell\theta} \left[e^{iNk\theta}\right]_{k\in\mathbb{Z}}.$$

In particular, we have

$$b_\ell(\theta) = e^{i\ell\theta}\, b_0(\theta),$$

so that for any Hermitian matrix Q,

$$b_\ell(\theta)^* Q\, b_\ell(\theta) = b_0(\theta)^* Q\, b_0(\theta).$$

This implies that, unlike in the introductory example, every invariant sum-of-squares polynomial can be expressed using an invariant basis. Thus, the entire computation can be carried out in the subring of invariant trigonometric polynomials $\mathbb{C}[e^{\pm iN\theta}]$.

Consider a trigonometric polynomial that is positive on a semialgebraic set invariant under the action of a finite group Γ. Such a polynomial admits a sum-of-squares decomposition of the form (8.6). If the generators of the semialgebraic set are all invariant, we can symmetrize each sum-of-squares polynomial individually. While not every set of generators is necessarily invariant, one can always find an invariant set of generators.

Proposition 8.7. *Let Γ be a finite group that induces a well-defined action on trigonometric polynomials such that the space of Hermitian trigonometric polynomials is invariant. For every Γ-invariant semialgebraic set*

$$\Sigma = \{\theta \in [-\pi, \pi)^n : \ p_1(\theta) \geq 0, \ldots, \ p_m(\theta) \geq 0\}$$

there exist Γ-invariant Hermitian trigonometric polynomials $q_1, \ldots, q_{m'}$ such that

$$\Sigma = \{\theta \in [-\pi, \pi)^n : \ q_1(\theta) \geq 0, \ldots, \ q_{m'}(\theta) \geq 0\}.$$

The corresponding statement for ordinary polynomials was proved by de Laat and Leijenhorst [52, Appendix A]. We adapt their arguments to the trigonometric case.

Proof. Since the semialgebraic set is invariant, we have

$$\Sigma = \bigcap_{j=1}^{m} \bigcap_{\gamma \in \Gamma} \{\theta \in [-\pi, \pi)^n : \ \gamma \cdot p_j(\theta) \geq 0\}.$$

It therefore suffices to verify that a set of the form

$$\Sigma = \bigcap_{\gamma \in \Gamma} \{\theta \in [-\pi, \pi)^n : \gamma \cdot p(\theta) \geq 0\}$$

is a semialgebraic set generated by Γ-invariant polynomials.

To this end, define the invariant polynomials

$$q_j = \sum_{\substack{S \subseteq \Gamma \\ |S| = j}} \prod_{\gamma \in S} \gamma \cdot p$$

for $j = 1, \ldots, |\Gamma|$ as well as the semialgebraic set

$$\Sigma' = \left\{\theta \in [-\pi, \pi)^n : \; q_1(\theta) \geq 0, \ldots, \; q_{|\Gamma|}(\theta) \geq 0\right\}.$$

We claim $\Sigma = \Sigma'$. The definition of q_j implies $\Sigma \subseteq \Sigma'$.

The reverse inclusion is shown by considering the complements. Assume that there exist $\theta \in [-\pi, \pi)^n$ and $\gamma \in \Gamma$ such that $\gamma \cdot p(\theta) < 0$ and $q_1(\theta) \geq 0, \ldots, q_{|\Gamma|}(\theta) \geq 0$. Since the functions q_j are invariant, applying the group action preserves the inequalities $q_j(\theta) \geq 0$. Using (8.7), we may replace θ with a suitable point in its orbit, which gives $\theta' \in [-\pi, \pi)^n$ such that $p(\theta') < 0$, while $q_j(\theta') \geq 0$ for all j.

Consider the polynomials

$$r_j = \sum_{\substack{S \subseteq \Gamma \setminus \{e\} \\ |S| = j}} \prod_{\gamma \in S} \gamma \cdot p.$$

Note that $q_j(\theta') = p(\theta')\, r_{j-1}(\theta') + r_j(\theta')$. Using $r_{|\Gamma|}(\theta') = 0$ in combination with the assumption gives $r_{|\Gamma|-1}(\theta') \leq 0$. In the same way, one can show $r_{|\Gamma|-2}(\theta') \leq 0$. In fact, the argument can be used inductively to show $r_1(\theta') \leq 0$. Hence,

$$q_1(\theta') = p(\theta') + r_1(\theta') \leq p(\theta') < 0,$$

which contradicts the assumption. □

8.3 A sum-of-squares formulation for the theta number

Consider a real-valued function of positive type on SO(3), as described in Theorem 7.4, with only finitely many nonzero coefficient matrices. Then f admits a finite expansion of the form

$$f(\alpha, \beta, \gamma) = \sum_{\ell=0}^{d} \langle F^\ell, Y^\ell(\alpha, \beta, \gamma)\rangle + \overline{\langle F^\ell, Y^\ell(\alpha, \beta, \gamma)\rangle} \qquad (8.8)$$

where

$$\left[Y^\ell(\alpha, \beta, \gamma)\right]_{m,n} = i^{n-m}\, e^{i(n\alpha + m\gamma)}\, P^\ell_{m,n}(\cos(\beta)).$$

In particular, f is a Hermitian trigonometric polynomial in the Euler angles (α, β, γ).

Let $N \geq 3$ and $\varrho \in (0, \pi/2)$ such that $\mathcal{G}_N(\varrho)$ is admissible. By Theorem 3.14, the interior-disjointness condition

$$(\mathcal{G}_N(\varrho) \cap A(\alpha, \beta, \gamma)\, \mathcal{G}_N(\varrho))^\circ = \emptyset$$

holds if and only if (α, β, γ) belongs to the union of

$$\Sigma_1 = \left\{(\alpha, \beta, \gamma) \in [-\pi, \pi)^3 : \ \cos(2\varrho) - \cos(\beta) \geq 0\right\},$$
$$\Sigma_2 = \bigcup_{k=1}^{N} \bigcap_{\ell=1}^{N} \left\{(\alpha, \beta, \gamma) \in [-\pi, \pi)^3 : \ h_k^\top A(\alpha, \beta, \gamma)\, v_\ell \geq 0\right\},$$
$$\Sigma_3 = \bigcup_{k=1}^{N} \bigcap_{\ell=1}^{N} \left\{(\alpha, \beta, \gamma) \in [-\pi, \pi)^3 : \ h_k^\top A(\alpha, \beta, \gamma)^\top v_\ell \geq 0\right\}.$$

Here, $v_1, \ldots, v_N$ and $h_1, \ldots, h_N$ denote the vertices of $\mathcal{G}_N(\varrho)$ and $\mathcal{G}_N(\varrho)^*$, respectively. Both $\cos(2\varrho) - \cos(\beta)$ and $h_k^\top A(\alpha, \beta, \gamma)\, v_\ell$ are Hermitian trigonometric polynomials in (α, β, γ).

By Theorem 6.1, an upper bound for $\tau(\mathcal{G}_N(\varrho))$ is obtained by solving the optimization problem

$$\begin{array}{lll} \inf & f(0,0,0) & \\ \text{s.t.} & F^0, \ldots, F^d & \text{Hermitian positive semidefinite,} \\ & f(\alpha, \beta, \gamma) & \text{given as in (8.8),} \\ & f(\alpha, \beta, \gamma) \leq -1 & \text{for all } (\alpha, \beta, \gamma) \in \Sigma_1 \cup \Sigma_2 \cup \Sigma_3. \end{array} \tag{8.9}$$

In principle, problem (8.9) can be approached using sum-of-squares methods. However, as the truncation parameter d increases, the degree of the underlying trigonometric polynomials and the size of the matrix variables grow rapidly. The resulting semidefinite programs quickly become computationally intractable, even for moderate values of d. Consequently, symmetry reduction techniques are essential for obtaining meaningful numerical bounds.

From the symmetry present in any packing problem (3.6), one obtains

$$(\mathcal{G}_N(\varrho) \cap A\mathcal{G}_N(\varrho))^\circ = \emptyset \quad \text{if and only if} \quad \left(\mathcal{G}_N(\varrho) \cap A^\top \mathcal{G}_N(\varrho)\right)^\circ = \emptyset.$$

Moreover, every real-valued function $f : \mathrm{SO}(3) \to \mathbb{R}$ of positive type is invariant under inversion of the argument $f(A) = f(A^\top)$. Hence,

$$f(\alpha, \beta, \gamma) \leq -1 \text{ on } \Sigma_2 \quad \text{if and only if} \quad f(\alpha, \beta, \gamma) \leq -1 \text{ on } \Sigma_3.$$

Thus, it suffices to impose the constraint on Σ_2.

Recall from Section 3.2 that the inverse of $A(\alpha, \beta, \gamma)$ can be represented in terms of Euler angles as

$$A(\alpha, \beta, \gamma)^\top = A(-\gamma, -\beta, -\alpha) = A(\pi - \gamma, \beta, \pi - \alpha).$$

These relations define a group action of $(\mathbb{Z}/2\mathbb{Z})^2$ on the space of trigonometric polynomials in the variables (α, β, γ) by

$$\begin{aligned} (1,0) \cdot p(\alpha, \beta, \gamma) &= p(-\gamma, -\beta, -\alpha), \\ (0,1) \cdot p(\alpha, \beta, \gamma) &= p(\pi - \gamma, \beta, \pi - \alpha), \end{aligned} \tag{8.10}$$

which leaves the optimization problem (8.9) invariant.

Additional symmetries arise from the symmetry group of $\mathcal{G}_N(\varrho)$, namely the *dihedral group* D_N. The dihedral group can be realized as a subgroup of the orthogonal group generated by

$$R = \begin{pmatrix} \cos(2\pi/N) & \sin(2\pi/N) & 0 \\ -\sin(2\pi/N) & \cos(2\pi/N) & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In particular, D_N admits the semidirect product decomposition $D_N = \mathbb{Z}/2\mathbb{Z} \ltimes \mathbb{Z}/N\mathbb{Z}$.

Using the D_N-symmetry, we observe that for $A \in \mathrm{SO}(3)$, the interior-disjointness condition

$$(\mathcal{G}_N(\varrho) \cap A\mathcal{G}_N(\varrho))^\circ = \emptyset$$

is equivalent to

$$\left(\mathcal{G}_N(\varrho) \cap \left(HAG^\top\right) \mathcal{G}_N(\varrho)\right)^\circ = \emptyset \qquad \text{for all } G, H \in D_N.$$

This yields an action of $D_N \times D_N$ on the space of matrices defined by

$$(G, H) \cdot A = HAG^\top.$$

However, this action does not preserve $\mathrm{SO}(3)$, since some elements of D_N reverse orientation. To obtain a well-defined action on $\mathrm{SO}(3)$, we restrict the action to the subgroup $\mathbb{Z}/2\mathbb{Z} \ltimes (\mathbb{Z}/N\mathbb{Z})^2$ generated by (R, I), (I, R), and (S, S). For every $(G, H) \in \mathbb{Z}/2\mathbb{Z} \ltimes (\mathbb{Z}/N\mathbb{Z})^2$ and $A \in \mathrm{SO}(3)$, one has $(G, H) \cdot A \in \mathrm{SO}(3)$.

Using the relations $SR_Z(\theta)S^\top = R_Z(-\theta)$ and $SR_X(\theta)S^\top = R_X(\theta)$, one obtains the induced action on Euler angles

$$\begin{aligned}
(R, I) \cdot A(\alpha, \beta, \gamma) &= A(\alpha - 2\pi/N, \beta, \gamma),\\
(I, R) \cdot A(\alpha, \beta, \gamma) &= A(\alpha, \beta, \gamma + 2\pi/N),\\
(S, S) \cdot A(\alpha, \beta, \gamma) &= A(-\alpha, \beta, -\gamma).
\end{aligned}$$

This gives rise to an action of $\mathbb{Z}/2\mathbb{Z} \ltimes (\mathbb{Z}/N\mathbb{Z})^2$ on the space of trigonometric polynomials in the variables (α, β, γ)

$$\begin{aligned}
(R, I) \cdot p(\alpha, \beta, \gamma) &= p(\alpha - 2\pi/N, \beta, \gamma),\\
(I, R) \cdot p(\alpha, \beta, \gamma) &= p(\alpha, \beta, \gamma + 2\pi/N),\\
(S, S) \cdot p(\alpha, \beta, \gamma) &= p(-\alpha, \beta, -\gamma),
\end{aligned}$$

which leaves the constraints in (8.9) invariant.

Collecting the symmetries discussed above leads to a group action of the semidirect product $\Gamma = (\mathbb{Z}/2\mathbb{Z})^3 \ltimes (\mathbb{Z}/N\mathbb{Z})^2$

$$\begin{aligned}
(1,0,0,0,0) \cdot p(\alpha, \beta, \gamma) &= p(-\gamma, -\beta, -\alpha),\\
(0,1,0,0,0) \cdot p(\alpha, \beta, \gamma) &= p(\pi - \gamma, \beta, \pi - \alpha),\\
(0,0,1,0,0) \cdot p(\alpha, \beta, \gamma) &= p(-\alpha, \beta, -\gamma),\\
(0,0,0,1,0) \cdot p(\alpha, \beta, \gamma) &= p(\alpha - 2\pi/N, \beta, \gamma),\\
(0,0,0,0,1) \cdot p(\alpha, \beta, \gamma) &= p(\alpha, \beta, \gamma + 2\pi/N),
\end{aligned} \tag{8.11}$$

where $(\mathbb{Z}/2\mathbb{Z})^3$ acts on $(\mathbb{Z}/N\mathbb{Z})^2$ via the induced transformations on the Euler angles.

Averaging any feasible solution of (8.9) over Γ yields an invariant feasible solution. This is achieved by averaging the matrix-valued functions Y^ℓ over the group Γ

$$\left[Y^\ell_{\mathrm{sym}}(\alpha, \beta, \gamma)\right]_{m,n} = \frac{1}{|\Gamma|} \sum_{g \in \Gamma} g \cdot \left[Y^\ell(\alpha, \beta, \gamma)\right]_{m,n}.$$

For indices $(m, n) \equiv (0, 0) \pmod N$, this yields

$$\left(i^{n-m}/2\right) \left(\cos(n\alpha + m\gamma) + (-1)^{n-m} \cos(m\alpha + n\gamma)\right) P^\ell_{m,n}(\cos(\beta)),$$

and it vanishes otherwise. Hence, only matrix entries with indices divisible by N need to be considered, which reduces the dimension of the coefficient matrices.

The invariance

$$Y^{\ell}_{\mathrm{sym}}(\alpha,\beta,\gamma) = Y^{\ell}_{\mathrm{sym}}(-\gamma,-\beta,-\alpha) = Y^{\ell}_{\mathrm{sym}}(\alpha,\beta,\gamma)^*$$

implies that Y^{ℓ}_{sym} is Hermitian and, consequently, $\langle F^{\ell}, Y^{\ell}_{\mathrm{sym}}(\alpha,\beta,\gamma)\rangle$ is real-valued for a Hermitian matrix F^{ℓ}.

In general, the actions of $(0,0,0,1,0)$ and $(0,0,0,0,1)$ do not preserve the objective value. The symmetry-adapted formulation should therefore be understood as a restriction of the original theta problem to a smaller, invariant search space, rather than as an equivalent reformulation.

The generator $\cos(2\varrho)-\cos(\beta)$ of the semialgebraic set Σ_1 is invariant under the group action of Γ. This allows for a symmetry-adapted sum-of-squares decomposition of polynomials strictly positive on Σ_1. Invariance under

$$(1,1,1,0,0)\cdot p(\alpha,\beta,\gamma) = p(\pi-\alpha,-\beta,\pi-\gamma)$$

further implies that such decompositions can be realized using real symmetric matrices as discussed in Example 8.5.

Although Σ_2 is not invariant under the action of Γ, the identity

$$A(\alpha,\beta,\gamma) = A(\alpha-\pi,-\beta,\gamma-\pi)$$

shows that each generator is invariant under the action of $(1,1,0,0,0)$, which induces a $\mathbb{Z}/2\mathbb{Z}$-symmetry. Invariance of f under the action of $(0,0,0,0,1)$ allows the set Σ_2 to be reduced to

$$\Sigma_2' = \bigcap_{\ell=1}^{N} \left\{(\alpha,\beta,\gamma)\in[-\pi,\pi)^3 : \, h_1^{\top} A(\alpha,\beta,\gamma) v_{\ell} \geq 0\right\}.$$

Replacing h_1 by any other vertex h_j yields a semialgebraic set in the same orbit, and hence an equivalent constraint.

Our choice of coordinates in (3.2) implies $Sh_1 = h_1$, which makes the set invariant under the action of (S,S). The semialgebraic set Σ_2' is also invariant under the action of (R,I), but generally not under (I,R). Consequently, the symmetries present in Σ_2' are induced by $(\mathbb{Z}/2\mathbb{Z})^2 \ltimes \mathbb{Z}/N\mathbb{Z}$ with the corresponding action

$$\begin{aligned}
(1,0,0)\cdot p(\alpha,\beta,\gamma) &= p(-\alpha,\beta,-\gamma),\\
(0,1,0)\cdot p(\alpha,\beta,\gamma) &= p(\alpha-\pi,-\beta,\gamma-\pi),\\
(0,0,1)\cdot p(\alpha,\beta,\gamma) &= p(\alpha-2\pi/N,\beta,\gamma).
\end{aligned} \tag{8.12}$$

Proposition 8.7 implies that there exists an invariant generating set for Σ_2'. In addition, the symmetry

$$(1,1,0)\cdot p(\alpha,\beta,\gamma) = p(-\alpha+\pi,-\beta,-\gamma+\pi)$$

ensures that the positive semidefinite matrices in sum-of-squares representations of polynomials strictly positive on Σ_2' can be assumed to be real.

The preceding discussion leads to the symmetry-reduced relaxation of the theta number

$$\begin{array}{lll} \inf & \displaystyle\sum_{\ell=0}^{d}\langle F^\ell, Y^\ell_{\mathrm{sym}}(0,0,0)\rangle & \\ \text{s.t.} & F^0,\dots,F^d & \text{Hermitian positive semidefinite,} \\ & \displaystyle\sum_{\ell=0}^{d}\langle F^\ell, Y^\ell_{\mathrm{sym}}(\alpha,\beta,\gamma)\rangle \le -1 & \text{for all } (\alpha,\beta,\gamma)\in\Sigma_1\cup\Sigma_2'. \end{array}$$

Recall from Example 8.6 that a symmetry-adapted sum-of-squares basis incorporating the $(\mathbb{Z}/N\mathbb{Z})^2$ symmetry in (8.11) lies in the invariant ring $\mathbb{C}\left[e^{\pm iN\alpha}, e^{\pm i\beta}, e^{\pm iN\gamma}\right]$. Thus the cyclic shift symmetries can be incorporated at the level of the monomial basis by restricting to the invariant ring.

The remaining symmetries then decompose this space into equivariant subspaces for the $(\mathbb{Z}/2\mathbb{Z})^3$-action. Hence, computing the irreducible representations of the full semidirect product $(\mathbb{Z}/2\mathbb{Z})^3 \ltimes (\mathbb{Z}/N\mathbb{Z})^2$ is not required. The same reasoning applies to the group action in (8.12).

Chapter Nine

Implementation and Results

In this chapter, we present lower and upper bounds for the packing problems associated with the configurations constructed in Part II of this thesis. For each configuration, we computed the volume bound (6.1), the symmetry-reduced theta number described in Section 8.3, and the Bachoc-Vallentin bound [4]. The Bachoc-Vallentin bound is not directly applied to the polygon packing problem, but to the spherical cap packing problem associated with the caps inscribed in the corresponding polygons, as explained in Section 6.2.

The computations of the theta number and the Bachoc-Vallentin bound were carried out in the `Julia` programming language [8], using the `ClusteredLowRankSolver` package [52], which supports arbitrary-precision floating-point arithmetic. A database containing the rational matrices defining each configuration as well as the scripts used for the computations is available in [28].

Tables 9.2, 9.3, and 9.4 present the resulting lower and upper bounds. For comparison, we also computed the theta number for the spherical cap packing problem using the approach of Delsarte, Goethals, and Seidel [33], discussed in Section 6.2. To this end, polynomials of degree 30 were used. The corresponding values are listed in Table 9.1. Figures 9.1 and 9.2 illustrate how the various bounds depend on the angular radii. Since the packing number is an integer, any gap of less than 1 between a lower and an upper bound certifies maximality.

As in Chapter 5, we consider all configurations of regular spherical triangles with cardinality at least 8. Our framework guarantees that all configurations of cardinality at least 14 are packings. Conjecture 3.15 suggests that the remaining configurations also define packings. These are marked as "conjectured" in Table 9.2.

Although the solutions were obtained numerically, the usage of arbitrary-precision arithmetic allows us to control feasibility errors. Consequently, the computed solutions can be perturbed into feasible ones with only negligible effect on the objective value. For a detailed discussion of this aspect in the context of a related problem, we refer to [24].

Computing the theta number for a packing problem associated with regular spherical triangles at truncation degree 14 requires several hours of computation time, whereas a truncation degree of 20 requires more than six days and approximately `128 GB` of memory. The sum-of-squares formulation requires both the truncation degree d and the quotient $\lfloor d/N \rfloor$ to be even, so not all truncation degrees are admissible. As a result, the degree of a constructed solution in the variables $e^{iN\alpha}$ and $e^{iN\gamma}$ is at most 6 for regular spherical triangles and at most 4 for regular spherical squares and pentagons. More powerful hardware or more efficient solvers may allow for higher degrees and thereby lead to improved bounds.

To compute the Bachoc-Vallentin bound for the packing problem associated with the inscribed spherical caps, we used polynomials of total degree at most 28. In some cases, there exist packings of spherical caps whose cardinality exceeds that of the cap packing induced by a polygon packing. In those cases, we report the corresponding cardinality in parentheses next to the Bachoc-Vallentin bound. The existence of such packings implies that upper bounds for the packing number of the inscribed spherical caps cannot be used to certify the maximality of the corresponding polygon packing configuration.

For many configurations, the theta number alone suffices to certify maximality. These are marked by a $\star$ symbol next to the angular radius. Configurations for which maximality can be certified by the theta number and, independently, by at least one of the other two bounds, are indicated by a † symbol. The symbol $-$ indicates that the theta number improves upon the floors of both the volume bound and the Bachoc-Vallentin bound, without certifying maximality.

Across all considered configurations, the theta number provides the strongest upper bounds and is most effective for certifying maximality. In particular, it frequently reduces the gap to below 1, thereby proving maximality. The Bachoc-Vallentin bound becomes increasingly competitive as the number of vertices increases. While it is generally weaker for triangles, its performance improves noticeably for squares and is strongest for pentagons among the considered cases. By contrast, the volume bound is computationally inexpensive but typically less tight, and only certifies maximality in a limited number of cases.

These results also yield improved bounds for the related problem of determining how many copies of a given Platonic solid can share a common vertex. For the icosahedron, all three bounds certify the maximality of our configuration. In the case of the octahedron, we construct a configuration of 7 interior-disjoint octahedra sharing a vertex. While the volume bound and the Bachoc-Vallentin bound yield an upper bound of 9, the theta number improves this to an upper bound of 8. Although the theta number gives a better bound than the volume bound for the packing problem corresponding to the tetrahedron, we were unable to improve the known lower bound of 20 or the known upper bound of 22.

Lower and upper bounds for packing problems for spherical caps

Radius	*Lower bound*	*Volume bound*	Theta number
0.955316⋆	4	4.73205080	4.00000000
0.785398	5	6.82842981	6.00000000
0.785398⋆	6	6.82842981	6.00000000
0.679539⋆	7	9.00343295	7.68883886
0.653263⋆	8	9.71368150	8.27952346
0.615479⋆	9	10.8990037	9.51918464
0.577239⋆	10	12.3435877	10.7914035
0.553574	11	13.3914519	12.0000000
0.553574⋆	12	13.3914519	12.0000000
0.498611	13	16.4267859	14.4969321
0.485817	14	17.2851527	15.2280842
0.468253	15	18.5801287	16.44861331
0.455918	16	19.5804350	17.36885997
0.445847	17	20.4594672	18.13548770
0.432463	18	21.7240749	19.39281488
0.416190⋆	19	23.4290463	19.39281488
0.413913	20	23.6837874	21.17864874
0.398050	21	25.5815348	22.78903891
0.390431	22	26.5763470	23.69705747
0.381441	23	27.8277034	24.81252556
0.381273⋆	24	27.8519341	24.83514047

Table 9.1: Lower and upper bounds for packing problems for spherical caps as a function of the angular radius. Lower bounds are given by explicit configurations that can be found in the references given in [19], while upper bounds are obtained via the volume bound and the theta number. The ⋆ symbol indicates configurations whose maximality is certified by the theta number. All numerical values are truncated to six decimal places for the radii and to eight decimal places for the bounds. Packings of cardinality 5 to 14 and 24 are known to be maximal [19].

Lower and upper bounds for packing problems for regular spherical triangles

Radius	*Lower bound*	*Volume bound*	*Bachoc-Vallentin*	*d*	Theta number	*Remarks*
0.955316†	8	8.00000000	(9) 9.00000000	14	8.00148720	conjectured
0.846584	9	10.84079445	(12) 12.72238983	14	10.43859589	conjectured
0.828930⋆	10	11.41373633	(12) 13.49278286	14	10.97882053	conjectured
0.790798	11	12.78772029	(15) 15.65588685	14	12.31168623	conjectured
0.760924^-	12	14.01415753	(17) 17.34929797	14	13.55069849	conjectured
				20	13.53801952	
0.738005^-	13	15.05923710	(18) 18.94361211	14	14.62806719	conjectured
0.719838	14	15.96041040	(20) 20.34188158	14	15.57438551	
0.699345^-	15	17.06344004	(20) 21.78736915	14	16.77117697	
0.685159	16	17.88615258	(22) 23.08083263	14	17.67500325	
				20	17.62851173	
0.669803	17	18.83686063	(24) 24.57551675	14	18.71223464	
				20	18.67127997	
0.658112	18	19.60600155	(24) 25.66413630	14	19.57509688	
				20	19.55107105	
0.652634†	19	19.98080594	(24) 26.21183173	14	19.99058786	
				20	19.98296276	
0.651805†	20	20.03839317	(24) 26.29552502	14	20.04764259	
				20	20.03822089	
0.615479	20	22.79466514	(28) 30.43073692	14	22.59030673	tetrahedra
				20	22.51176943	sharing a vertex
0.604417	21	23.73485658	(30) 31.77968284	14	23.48417323	
				20	23.40372464	
0.587346^-	22	25.29181773	(32) 34.02763309	14	24.99197449	
				20	24.90075396	
0.570870	23	26.92948343	(34) 36.66255105	14	26.59150848	
				20	26.48770382	
0.558125^-	24	28.29737925	(36) 38.71347302	14	27.93923717	
				20	27.82052189	

Table 9.2: Lower and upper bounds for packing problems for regular spherical triangles. The lower bounds are given by explicit configurations, while upper bounds are obtained via the volume bound, the Bachoc-Vallentin bound, and the theta number. All numerical values are truncated to six decimal places for the radii and to eight decimal places for the bounds.

Lower and upper bounds for packing problems for regular spherical squares

Radius	*Lower bound*	*Volume bound*	*Bachoc-Vallentin*	*d*	Theta number	*Remarks*
1.183199†	3	3.70443713	3.00000000	10	3.00000001	
1.063440†	4	4.71541086	4.07790842	10	4.10116100	
0.955316	5	6.00000117	(6) 6.00000000	10	6.00000044	
				18	6.00000026	
0.955316†	6	6.00000000	6.00000000	8	6.00000463	
0.812554*	7	8.58304028	(8) 8.13119410	10	7.84845101	
0.785398^{-}	7	9.24441273	(9) 9.00000000	10	8.46487808	octahedra
				18	8.43452493	sharing a vertex
0.777160*	8	9.45911352	(9) 9.10977817	10	8.67678741	
0.737706*	9	10.59062399	(10) 10.27965041	10	9.72908492	
0.701465*	10	11.80498479	(12) 12.08022888	10	10.87300195	
0.665589^{-}	11	13.20988402	(12) 13.08564377	10	12.36460904	
				18	12.24240918	
0.659825*	12	13.45739970	13.35693928	10	12.61827032	
0.611102^{-}	13	15.83957294	(15) 16.21135247	10	14.96350191	
				18	14.70525606	
0.598860^{-}	14	16.53186711	(16) 16.95237174	10	15.70271462	
				18	15.36791722	
0.585592^{-}	15	17.33205714	(17) 17.82667450	10	16.60940922	
				18	16.16157960	
0.564943	16	18.69188151	(18) 19.54161047	10	18.18025701	
0.549177^{-}	17	19.83538800	(20) 20.76935334	10	19.48304253	
				18	18.59577345	
0.537891^{-}	18	20.71669278	(20) 21.67448098	10	20.53747812	
				18	19.43932673	
0.516950^{-}	19	22.50799105	(22) 23.89372360	10	22.82038760	
				18	21.21727058	
0.510062^{-}	20	23.14608098	(24) 24.61734992	10	23.66899224	
				18	21.84775700	
0.495188^{-}	21	24.61617987	(24) 26.24964661	10	25.64625742	
				18	23.29803247	
0.485876^{-}	22	25.60629269	(26) 27.32759700	10	27.00764135	
				18	24.26292605	
0.475163^{-}	23	26.81816363	(27) 28.70062328	10	28.74508182	
				18	25.44302605	
0.465337^{-}	24	28.00434523	(28) 30.17773550	10	30.52343961	
				18	26.62321441	

Table 9.3: Lower and upper bounds for packing problems for regular spherical squares. The lower bounds are given by explicit configurations, while upper bounds are obtained via the volume bound, the Bachoc-Vallentin bound, and the theta number. All numerical values are truncated to six decimal places for the radii and to eight decimal places for the bounds.

Lower and upper bounds for packing problems for regular spherical pentagons

Radius	*Lower bound*	*Volume bound*	*Bachoc-Vallentin*	*d*	Theta number	*Remarks*
$1.107148^\dagger$	3	4.00000000	3.23606797	14	3.23606992	
$1.027058^\dagger$	4	4.67521434	4.02586260	14	4.00425871	
$1.017221^\dagger$	4	4.76984107	4.04611156	10	4.06501130	icosahedra sharing a vertex
$0.865924^\star$	5	6.67214088	(6) 6.01513743	14	5.94289949	
$0.850308^\dagger$	6	6.92983044	6.05170751	14	6.22491869	
$0.756498^\star$	7	8.83521938	(8) 8.00493858	14	7.91739923	
$0.731849^\dagger$	8	9.46296954	8.44605070	14	8.63164917	
$0.692196^\dagger$	9	10.61854972	9.56183171	14	9.99379488	
0.658224	10	11.78061879	11.03162960	14	11.57505633	
				20	11.53768330	
0.652358	11	12.00000004	(12) 12.00000000	14	12.00000393	
$0.652358^\dagger$	12	12.00000510	12.00000001	10	12.00002878	
0.565329	13	16.10441278	(14) 15.26026568	14	15.32016393	
0.554476	14	16.75654644	(15) 15.89359509	14	15.95264213	
0.539673	15	17.71049491	(16) 16.83442133	14	16.89037273	
0.524729	16	18.75684983	(17) 17.87168328	14	17.89517658	
0.512308	17	19.69727207	(18) 18.91986202	14	18.82818266	
0.498146^-	18	20.85668953	(19) 20.21849698	14	19.96762912	
				20	19.62178909	
0.480722	19	22.42647447	(20) 21.66650360	14	21.49920555	
				20	21.05564494	
0.476634^-	20	22.81997390	(21) 22.05813478	14	21.88236952	
				20	21.39579388	
0.459979^-	21	24.53327575	(24) 24.03727825	14	23.60037910	
				20	22.91769941	
0.451462^-	22	25.48362122	(24) 24.98748702	14	24.58375506	
				20	23.81047251	
0.436889^-	23	27.24080150	(25) 26.77138644	14	26.51402878	
				20	25.41736863	
0.435221^-	24	27.45326251	(25) 26.98022302	14	26.75588236	
				20	25.61687007	

Table 9.4: Lower and upper bounds for packing problems for regular spherical pentagons. The lower bounds are given by explicit configurations, while upper bounds are obtained via the volume bound, the Bachoc-Vallentin bound, and the theta number. All numerical values are truncated to six decimal places for the radii and to eight decimal places for the bounds.

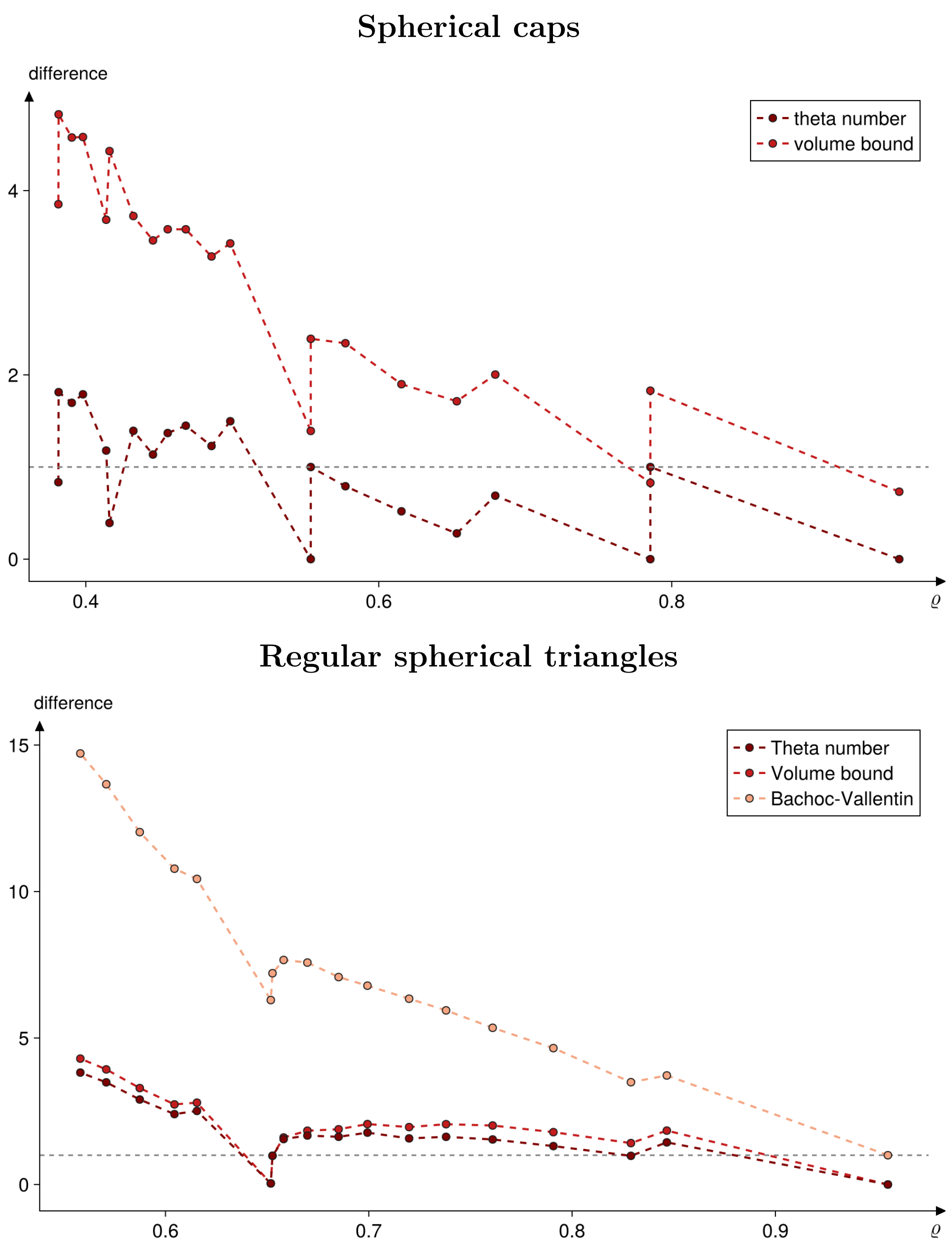
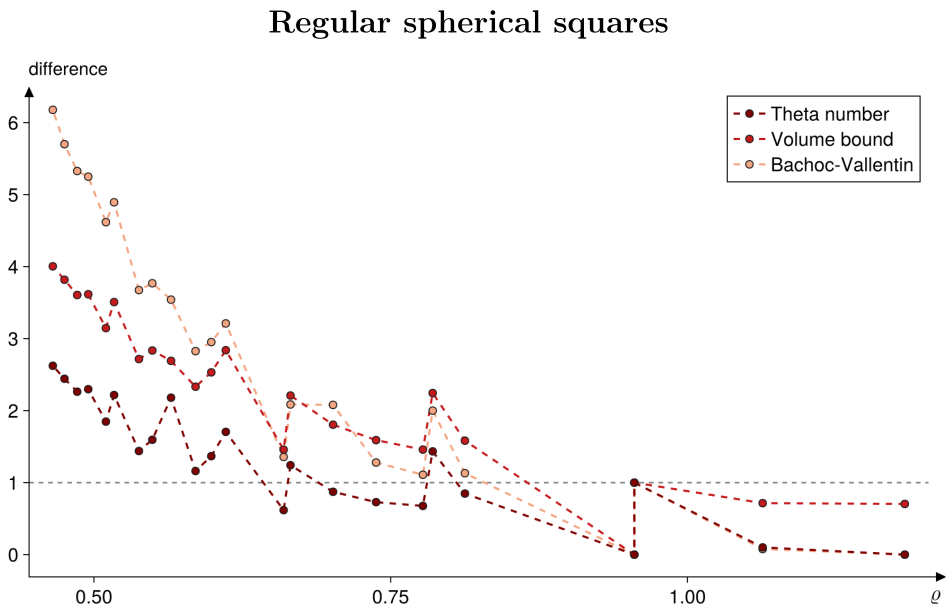


Figure 9.1: Differences between the lower bound and the corresponding upper bounds for packing problems for spherical caps and regular spherical triangles. The discrete data points obtained from Tables 9.1 and 9.2 are connected by dashed lines for visual guidance. The horizontal dashed line at 1 marks the threshold below which an upper bound certifies maximality.



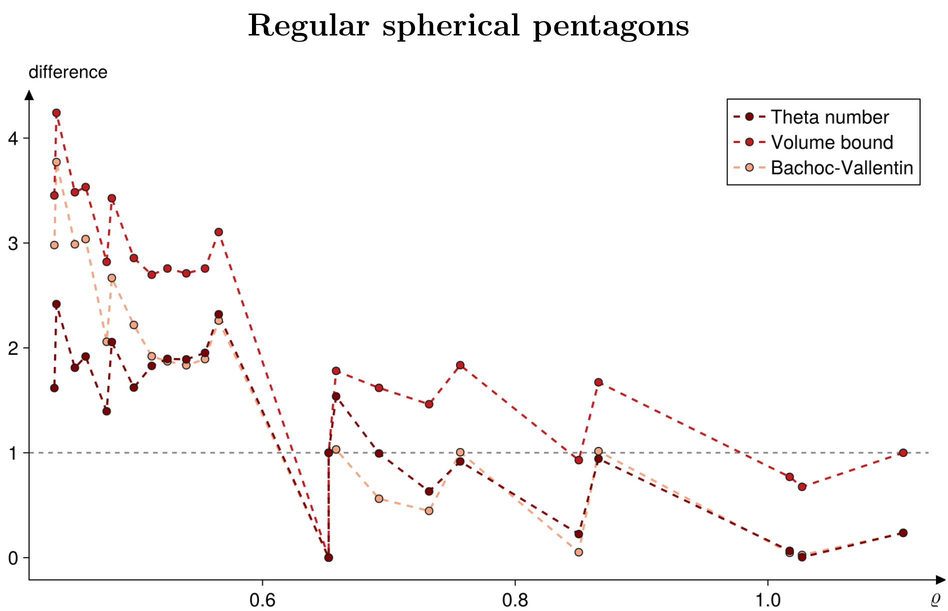


Figure 9.2: Differences between the lower bound and the corresponding upper bounds for packing problems for regular spherical squares and regular spherical pentagons. The discrete data points obtained from Tables 9.3 and 9.4 are connected by dashed lines for visual guidance. The horizontal dashed line at 1 marks the threshold below which an upper bound certifies maximality.

Bibliography

[1] G. Andrews, R. Askey, and R. Roy, *Special Functions*, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999.

[2] Aristotle, *On the Heavens: With an English Translation by W. K. C. Guthrie*, ser. Loeb classical library. Harvard University Press, 1986.

[3] S. D. Axen, M. Baran, R. Bergmann, and K. Rzecki, "Manifolds.jl: An extensible Julia framework for data analysis on manifolds," *ACM Transactions on Mathematical Software*, vol. 49, no. 4, p. 1–23, 2023. [Online]. Available: http://doi.org/10.1145/3618296

[4] C. Bachoc and F. Vallentin, "New upper bounds for kissing numbers from semidefinite programming," *Journal of the American Mathematical Society*, vol. 21, no. 3, p. 909–924, 2008. [Online]. Available: http://doi.org/10.1090/S0894-0347-07-00589-9

[5] A. Barvinok, *A Course in Convexity*, ser. Graduate studies in mathematics. American Mathematical Soc., 2002.

[6] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, ser. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2001.

[7] R. Bergmann, "Manopt.jl: Optimization on manifolds in Julia," *Journal of Open Source Software*, vol. 7, no. 70, p. 3866, 2022.

[8] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Review*, vol. 59, no. 1, pp. 65–98, 2017. [Online]. Available: http://doi.org/10.1137/141000671

[9] E. G. Birgin and J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, ser. Fundamentals of Algorithms. Philadelphia, PA: SIAM, 2014. [Online]. Available: https://epubs.siam.org/doi/book/10.1137/1.9781611973365

[10] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*. Berlin, Heidelberg: Springer-Verlag, 1997.

[11] S. Bochner, "Hilbert distances and positive definite functions," *Annals of Mathematics*, vol. 42, no. 3, pp. 647–656, 1941. [Online]. Available: http://www.jstor.org/stable/1969252

[12] N. Boumal, *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023. [Online]. Available: https://www.nicolasboumal.net/book

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[14] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, ser. Graduate Texts in Mathematics. Springer Berlin Heidelberg, 2003.

[15] K. J. Böröczky and A. Glazyrin, "Stability of optimal spherical codes," 2023. [Online]. Available: https://arxiv.org/abs/1711.06012

[16] E. R. Chen, "A picturebook of tetrahedral packings," Ph.D. Thesis, University of Michigan, 2008.

[17] E. R. Chen, M. Engel, and S. C. Glotzer, "Dense crystalline dimer packings of regular tetrahedra," *Discrete & Computational Geometry*, vol. 44, no. 2, p. 253–280, 2010. [Online]. Available: http://doi.org/10.1007/s00454-010-9273-0

[18] G. Chirikjian and A. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotation and Motion Groups*. CRC Press, 2000.

[19] H. Cohn, "Small spherical and projective codes," https://hdl.handle.net/1721.1/15354, 2004, dataset, MIT DSpace repository.

[20] H. Cohn, A. Kumar, S. Miller, D. Radchenko, and M. Viazovska, "The sphere packing problem in dimension 24," *Annals of Mathematics*, vol. 185, no. 3, 2017. [Online]. Available: http://doi.org/10.4007/annals.2017.185.3.8

[21] H. Coxeter, *Introduction to Geometry*, ser. Wiley Classics Library. Wiley, 1989.

[22] L. Danzer, B. Grünbaum, and V. Klee, "Helly's theorem and its relatives," in *Convexity*. American Mathematical Society, 1963, pp. 101–180.

[23] E. de Klerk and F. Vallentin, "On the Turing model complexity of interior point methods for semidefinite programming," *SIAM Journal on Optimization*, vol. 26, no. 3, p. 1944–1961, 2016. [Online]. Available: http://doi.org/10.1137/15M103114X

[24] D. de Laat, F. M. de Oliveira Filho, and F. Vallentin, "Upper bounds for packings of spheres of several radii," *Forum of Mathematics, Sigma*, vol. 2, p. e23, 2014.

[25] D. de Laat, N. M. Leijenhorst, and W. H. H. de Muinck Keizer, "Optimality and uniqueness of the D_4 root system," 2024. [Online]. Available: https://arxiv.org/abs/2404.18794

[26] D. de Laat and F. Vallentin, "A semidefinite programming hierarchy for packing problems in discrete geometry," *Mathematical Programming*, vol. 151, no. 2, p. 529–553, 2014. [Online]. Available: http://doi.org/10.1007/s10107-014-0843-4

[27] F. M. de Oliveira Filho, A. Spomer, and F. Vallentin, "Bounding the density of spherical polygon packings," 2026. [Online]. Available: https://arxiv.org/abs/2604.21451

[28] ——, "Data and code for: Bounding the density of spherical polygon packings," 2026. [Online]. Available: http://doi.org/10.7910/DVN/PNH45K

[29] F. M. de Oliveira Filho and F. Vallentin, "Mathematical optimization for packing problems," 2015. [Online]. Available: https://arxiv.org/abs/1403.1166

[30] ——, “Computing upper bounds for the packing density of congruent copies of a convex body,” *Bolyai Society Mathematical Studies*, pp. 155–188, 2018.

[31] A. Deitmar and S. Echterhoff, *Principles of Harmonic Analysis*, ser. Universitext. Springer International Publishing, 2014.

[32] Delft High Performance Computing Centre (DHPC), “DelftBlue Supercomputer (Phase 2),” https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2, 2024.

[33] P. Delsarte, J. Goethals, and J. Seidel, “Spherical codes and designs,” *Geometriae Dedicata*, vol. 6, pp. 363–388, 1977.

[34] J. Diebel, “Representing attitude: Euler angles, unit quaternions, and rotation vectors,” *Matrix*, vol. 58, 2006.

[35] B. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications.* Springer, 2007.

[36] L. Fejes Tóth, G. Fejes Tóth, and W. Kuperberg, *Lagerungen: Arrangements in the Plane, on the Sphere, and in Space*, ser. Grundlehren der mathematischen Wissenschaften. Springer International Publishing, 2023, vol. 360.

[37] G. Folland, *A Course in Abstract Harmonic Analysis*, ser. Studies in Advanced Mathematics. Taylor & Francis, 1994.

[38] W. Fulton and J. Harris, *Representation Theory: A First Course*, ser. Graduate Texts in Mathematics. Springer New York, 1991.

[39] K. Gatermann, “Semi-invariants, equivariants and algorithms,” *Applicable Algebra in Engineering, Communication and Computing*, vol. 7, 1996. [Online]. Available: http://doi.org/10.1007/BF01191379

[40] K. Gatermann and P. A. Parrilo, “Symmetry groups, semidefinite programs, and sums of squares,” *Journal of Pure and Applied Algebra*, vol. 192, p. 95–128, 2004. [Online]. Available: http://doi.org/10.1016/j.jpaa.2003.12.011

[41] A. Gheondea, “An operator theoretical approach to Mercer’s theorem,” 2025. [Online]. Available: https://arxiv.org/abs/2512.06475

[42] G. H. Golub and C. F. van Loan, *Matrix Computations*, 4th ed. JHU Press, 2013. [Online]. Available: http://www.cs.cornell.edu/cv/GVL4/golubandvanloan.htm

[43] S. Gravel, V. Elser, and Y. Kallus, “Upper bound on the packing density of regular tetrahedra and octahedra,” *Discrete & Computational Geometry*, vol. 46, no. 4, p. 799–818, 2010. [Online]. Available: http://doi.org/10.1007/s00454-010-9304-x

[44] M. Grötschel, L. Lovász, and A. Schrijver, “The ellipsoid method and its consequences in combinatorial optimization,” *Combinatorica*, vol. 1, no. 2, pp. 169–197, 1981.

[45] T. C. Hales, “A proof of the Kepler conjecture,” *Annals of Mathematics*, vol. 162, no. 3, pp. 1065–1185, 2005.

[46] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, ser. Graduate Texts in Mathematics. Cham: Springer International Publishing, 2015.

[47] H. A. Hashim, “Special orthogonal group SO(3), Euler angles, angle-axis, Rodriguez vector and unit-quaternion: Overview, mapping and challenges,” 2024. [Online]. Available: https://arxiv.org/abs/1909.06669

[48] G. A. Kabatiansky and V. I. Levenshtein, “On bounds for packings on a sphere and in space,” *Problemy Peredachi Informatsii*, vol. 14, no. 1, pp. 3–25, 1978.

[49] J. Kepler, *Ioannis Kepleri S. C. Maiest. Mathematici Strena Seu De Niue Sexangula*. Tampach, 1611.

[50] J. Lagarias and C. Zong, “Mysteries in packing regular tetrahedra,” *Notices of the American Mathematical Society*, vol. 59, 2012.

[51] M. Laurent and F. Vallentin, “Semidefinite optimization,” https://www.mi.uni-koeln.de/opt/wp-content/uploads/2015/10/laurent_vallentin_sdo_2012_05.pdf, 2012.

[52] N. Leijenhorst and D. de Laat, “Solving clustered low-rank semidefinite programs arising from polynomial optimization,” 2023. [Online]. Available: https://arxiv.org/abs/2202.12077

[53] C. Liu and N. Boumal, “Simple algorithms for optimization on Riemannian manifolds with constraints,” 2019. [Online]. Available: https://arxiv.org/abs/1901.10000

[54] L. Lovász, “On the Shannon capacity of a graph,” *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.

[55] W. F. Mascarenhas and E. G. Birgin, “Using sentinels to detect intersections of convex and nonconvex polygons,” *Computational & Applied Mathematics*, vol. 29, 2010.

[56] W. S. Massey, “Cross products of vectors in higher dimensional Euclidean spaces,” *The American Mathematical Monthly*, vol. 90, pp. 697–701, 1983. [Online]. Available: http://www.jstor.org/stable/2323537

[57] T. Metzlaff, “Symmetry adapted bases for trigonometric optimization,” *ACM Commun. Comput. Algebra*, vol. 57, p. 137–140, 2023. [Online]. Available: http://doi.org/10.1145/3637529.3637535

[58] ——, “On symmetry adapted bases in trigonometric optimization,” *Journal of Symbolic Computation*, vol. 127, p. 102369, 2025. [Online]. Available: http://doi.org/10.1016/j.jsc.2024.102369

[59] O. R. Musin, “The kissing number in four dimensions,” 2006. [Online]. Available: https://arxiv.org/abs/math/0309430

[60] A. Odlyzko and N. Sloane, “New bounds on the number of unit spheres that can touch a unit sphere in n dimensions,” *Journal of Combinatorial Theory, Series A*, vol. 26, no. 2, pp. 210–214, 1979. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0097316579900748

[61] R. O’Donnell, “SOS Is Not Obviously Automatizable, Even Approximately,” in *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 67. Schloss Dagstuhl

– Leibniz-Zentrum für Informatik, 2017, pp. 59:1–59:10. [Online]. Available: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ITCS.2017.59

[62] Plato, "Timaeus," in *Timaeus. Critias. Cleitophon. Menexenus. Epistles*, ser. Loeb Classical Library. Harvard University Press, 1929, no. 234. [Online]. Available: https://www.loebclassics.com/view/plato-philosopher_timaeus/1929/pb_LCL234.235.xml

[63] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana University Mathematics Journal*, vol. 42, pp. 969–984, 1993. [Online]. Available: http://www.jstor.org/stable/24897130

[64] C. Scheiderer, "Non-existence of degree bounds for weighted sums of squares representations," *Journal of Complexity*, vol. 21, pp. 823–844, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0885064X05000300

[65] R. Schneider and W. Weil, *Stochastic and Integral Geometry*, ser. Probability and Its Applications. Springer Berlin Heidelberg, 2008.

[66] R. Schneider, *Convex Cones: Geometry and Probability*, 1st ed., ser. Lecture Notes in Mathematics, 2319. Cham: Springer International Publishing, 2022.

[67] A. Schrijver, *Theory of Linear and Integer Programming.* Chichester: John Wiley & Sons, 1986, reprinted 1998.

[68] ——, "A course in combinatorial optimization," 2003. [Online]. Available: https://homepages.cwi.nl/~lex/files/dict.pdf

[69] K. Schütte and B. L. van der Waerden, "Das problem der dreizehn kugeln," *Mathematische Annalen*, vol. 125, pp. 325–334, 1953. [Online]. Available: https://eudml.org/doc/160314

[70] L. Simon, *Lectures on Geometric Measure Theory*, ser. Proceedings of the Centre for Mathematical Analysis. Centre for Mathematical Analysis, Australian National University, 1984.

[71] G. G. Szpiro, *Kepler's Conjecture: How Some of the Greatest Minds in History Helped Solve One of the Oldest Math Problems in the World.* Hoboken, NJ: Wiley, 2003.

[72] J. D. Talman, *Special Functions: A Group Theoretic Approach.* New York: W. A. Benjamin, Inc., 1968, based on lectures by Eugene P. Wigner.

[73] T. Tarnai and Z. Gáspár, "Packing of equal regular pentagons on a sphere," *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 457, pp. 1043–1058, 2001.

[74] T. Tarnai, Z. Gáspár, and L. Szalai, "Pentagon packing models for "all-pentamer" virus structures," *Biophysical Journal*, vol. 69, no. 2, pp. 612–618, 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0006349595799386

[75] T. Tarnai and Z. Gáspár, "Packing of 5 regular pentagons on a sphere," *Acta Technica Academiae Scientiarum Hungaricae*, vol. 106, no. 3–4, pp. 275–282, 1994.

[76] ——, "Packing of regular pentagons on a sphere," in *Intuitive Geometry*, ser. Colloquia Mathematica Societatis János Bolyai, K. Böröczky and G. Fejes Tóth, Eds. Amsterdam: North-Holland, 1994, vol. 63, pp. 475–480.

[77] A. Thue, *Über die dichteste Zusammenstellung von kongruenten Kreisen in einer Ebene.* J. Dybwad, 1910.

[78] S. Torquato and Y. Jiao, "Dense packings of polyhedra: Platonic and Archimedean solids," *Phys. Rev. E*, vol. 80, p. 041104, 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.80.041104

[79] F. Vallentin, "Conic optimization for extremal geometry," 2025. [Online]. Available: https://arxiv.org/abs/2510.06960

[80] M. Viazovska, "The sphere packing problem in dimension 8," *Annals of Mathematics*, vol. 185, no. 3, 2017. [Online]. Available: http://doi.org/10.4007/annals.2017.185.3.7

[81] J. Wang, "A more efficient reformulation of complex SDP as real SDP," 2024. [Online]. Available: https://arxiv.org/abs/2307.11599

List of Symbols

Sets

$\emptyset$	empty set
$\mathbb{N}$	natural numbers
$\mathbb{N}_0$	nonnegative integers
$\mathbb{Z}$	integers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$L^p(X)$	space of p-integrable functions on a measure space X
$\langle p_1, \ldots, p_n \rangle$	ideal generated by polynomials $p_1, \ldots, p_n$
$\lvert R \rvert$	cardinality of a finite set R
$x \equiv y$	x and y belong to the same equivalence class with respect to a given equivalence relation

Linear Algebra

$\overline{z}$	complex conjugate of z, applied entrywise to vectors and matrices
$\mathrm{Re}(z)$	real part of z
$\mathrm{Im}(z)$	imaginary part of z
$\lVert x \rVert$	Euclidean norm of a vector x
$x \times y$	cross product of vectors x and y (p. 12)
I	identity matrix of appropriate dimension
A^{-1}	inverse of A
$A^{\top}$	transpose of A
U^*	conjugate transpose of U
$\det(A)$	determinant of A
$\mathrm{tr}(A)$	trace of A
$V \otimes W$	tensor product of vector spaces V and W
$\dim(V)$	dimension of V
$\mathrm{cone}\{v_1, \ldots, v_N\}$	polyhedral cone generated by $v_1, \ldots, v_N$ (p. 11)
$\mathcal{K}^*$	dual cone of $\mathcal{K}$ (p. 11)

Spherical Geometry

$\mathbb{S}^{d-1}$	$(d-1)$-dimensional unit sphere (p. 17)
δ	geodesic metric (p. 17)
$\mathcal{Q}^\circ$	interior of $\mathcal{Q}$ in $\mathbb{S}^{d-1}$ (p. 17)
$\partial\mathcal{Q}$	boundary of $\mathcal{Q}$ in $\mathbb{S}^{d-1}$ (p. 17)
$\mathcal{Q}^*$	dual of a spherically convex set $\mathcal{Q}$ (p. 18)
$[x, y]$	geodesic arc between non-antipodal points x and y (p. 18)
$\mathcal{C}(\varrho)$	spherical cap of angular radius ϱ (p. 18)
$\mathcal{G}_N(\varrho)$	regular spherical polygon of angular radius ϱ (p. 19)
$v_k(\varrho)$	k-th vertex of $\mathcal{G}_N(\varrho)$ (p. 18)
$h_k(\varrho)$	k-th vertex of $\mathcal{G}_N(\varrho)^*$ (p. 19)
$\tau(\mathcal{Q})$	packing number of a spherically convex body $\mathcal{Q}$ (p. 20)
$\Delta(\mathcal{P})$	shortest side length of a spherical polytope $\mathcal{P}$ (p. 28)
$\mathcal{F}_j(\mathcal{P})$	forbidden region of a spherical polytope $\mathcal{P}$ around the j-th vertex (p. 28)
$\mathcal{F}(\mathcal{P})$	forbidden region of a spherical polytope $\mathcal{P} \subseteq \mathbb{S}^2$ (p. 29)

Groups

e	identity element of a group
γ^{-1}	inverse of a group element γ
$\gamma \cdot v$	action of a group element γ on a vector v (p. 13)
χ_ρ	character of a representation ρ (p. 15)
$G \ltimes H$	semidirect product of groups G and H
$\widehat{\Gamma}$	dual group of Γ (p. 14)
$\mathrm{SO}(d)$	special orthogonal group in dimension d (p. 13)
$\mathrm{SU}(d)$	special unitary group in dimension d (p. 13)
$\mathbb{Z}/N\mathbb{Z}$	cyclic group with N elements
$\mathrm{Sym}(n)$	symmetric group on n symbols
$R_X(\theta)$	rotation matrix by angle θ about the x-axis (p. 21)
$R_Z(\theta)$	rotation matrix by angle θ about the z-axis (p. 21)
$A(\alpha, \beta, \gamma)$	Euler angle parametrization of SO(3) (p. 21)
(α, β, γ)	Euler angles parametrizing rotations in SO(3) and SU(2)
$U_Z(\theta)$	basic rotation in SU(2) by angle θ (p. 76)
$U_X(\theta)$	basic rotation in SU(2) by angle θ (p. 76)
x_{sym}	image of x under the symmetrization operator (p. 15)

Acknowledgements

On some days, working on mathematics can be extraordinarily fulfilling, while on others, it can be endlessly frustrating. It is good to have people around with whom one can enjoy the fascinating aspects and get through the difficult moments. I was lucky to have such people around me, and I would like to express my gratitude to them.

First, I would like to mention my former PhD colleagues *Arne Heimendahl* and *Karla Leipold*. Over the last couple of years, we have had quite a few adventures together. If either of you ever wants to play a game of mathematical-theorem charades or needs a helping hand with carrying an old couch to your office, let me know. *Marc Zimmermann* has earned my highest appreciation for mixing the best cocktails ever served by a mathematician. Moreover, I would like to thank *Annette Koenen*, the good fairy of our group, who solves problems none of us even knew existed.

Special thanks go to my supervisor, *Frank Vallentin*. Attending your lectures as an undergraduate student sparked an interest that I never knew I had and inspired me to pursue a career in academia. By accepting me as a PhD student, you gave me the opportunity to go on this wonderful journey, and I could not be more grateful for that. Thank you for everything you have done for me, both before and after I became a PhD student under your supervision.

I am very glad to have been a member of our little group, and I look forward to sharing a Lavazza coffee with any of you in the near future.

I would also like to thank *Fernando Mário de Oliveira Filho*. Your open and generous manner makes collaborating with you very enjoyable, and your feedback helped me improve as a mathematician.

During my time as a PhD student, I had the pleasure of sharing both inspiring and light-hearted conversations with many other researchers, including but not limited to *Bram Bekker*, *Nando Leijenhorst*, *Willem de Muinck Keizer*, *Felix Kirschner*, *Daniel Brosch*, and

Nicolas Nagel.

Finally, I would like to thank my friends and family for their support over the last couple of years. In particular, I thank *Jana Michael*, who supported me throughout the creation of this thesis.

I am grateful to *Karla Leipold*, *Bram Bekker*, and *Marc Zimmermann* for proofreading parts of this thesis.

Technical contributions

The computations underlying the results presented in this thesis were carried out on the TU Delft supercomputer DelftBlue, provided by the Delft High Performance Computing Centre [32], and on a system with Samsung ECC DDR4-2667 memory, access to which was kindly provided by the group of Christian Sohler. Special thanks go to Martin Olschewski, who set up access to the latter.

Digital tools, including ChatGPT, were used in a supportive manner for language editing of parts of the text. All mathematical content, proofs, computations, interpretations, and conclusions were developed, checked, and are the sole responsibility of the author.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten – noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss des Prüfungsverfahrens der Promotion nicht vornehmen werde. Die Bestimmungen der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der der Dissertation zugrundeliegenden Arbeiten und der schriftlich oder elektronisch verfassten Dissertation beachtet habe, und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen

Teilpublikationen

- F.M. de Oliveira, A. Spomer, F. Vallentin, *Bounding the density of spherical polygon packings* [27]

Köln, den 25.05.2026

Andreas Spomer