# Geometric analysis of the condition of the convex feasibility problem

Dissertation

zur Erlangung des Doktorgrades
der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn

vorgelegt von

**Dennis Amelunxen**

Paderborn, den 28. März 2011
(revidierte Version vom 19. August 2011)

Betreuer: Prof. Dr. Peter Bürgisser

Gutachter: Prof. Dr. Peter Bürgisser
Prof. Dr. James Renegar
Prof. Dr. Joachim Hilgert

## Abstract

The focus of this paper is the homogeneous convex feasibility problem, which is the following question: Given an $m$-dimensional subspace of $\mathbb{R}^n$, does this subspace intersect a fixed convex cone solely in the origin or are there further intersection points? This problem includes as special cases the linear, the second order, and the semidefinite feasibility problems, where one simply chooses the positive orthant, a product of Lorentz cones, or the cone of positive semidefinite matrices, respectively. An important role for the running time of algorithms solving the convex feasibility problem is played by Renegar's condition number. The (inverse of the) condition of an input is basically the magnitude of the smallest perturbation, which changes the status of the input, i.e., which makes a feasible input infeasible, or the other way round. We will give an average analysis of this condition for several classes of convex cones, and one that is independent of the underlying convex cone. We will also describe a way of deriving smoothed analyses from our approach. We will achieve these results by adopting a purely geometric viewpoint leading to computations in the Grassmann manifold.

Besides these main results about the random behavior of the condition of the convex feasibility problem, we will obtain a couple of byproduct results in the domain of spherical convex geometry.

## Kurzbeschreibung

Den Mittelpunkt dieser Arbeit bildet das homogene konvexe Lösbarkeitsproblem, welches die folgende Frage ist: Gegeben sei ein $m$-dimensionaler Unterraum des $\mathbb{R}^n$; schneidet dieser Unterraum einen gegebenen konvexen Kegel nur im Ursprung, oder gibt es weitere Schnittpunkte? Dieses Problem umfasst als Spezialfälle das lineare, das quadratische, und das semidefinite Lösbarkeitsproblem, wobei man als konvexen Kegel den positiven Orthanten, ein Produkt von Lorentzkegeln, bzw. den Kegel der positiv semidefiniten Matrizen wählt. Für die Laufzeit von Algorithmen, welche das konvexe Lösbarkeitsproblem lösen, spielt die Renegarsche Konditionszahl eine wichtige Rolle. Die Kondition einer Eingabe, bzw. ihr Inverses, ist gegeben durch die Größe einer kleinsten Störung, welche den Status der Eingabe von 'feasible' zu 'infeasible' bzw. von 'infeasible' zu 'feasible' ändert. Wir werden eine Durchschnittsanalyse dieser Kondition für verschiedene Klassen von konvexen Kegeln angeben, sowie eine, welche unabhängig ist von der Wahl des zugrunde gelegten konvexen Kegels. Wir werden desweiteren einen Weg beschreiben, auf dem auch geglättete Analysen mittels unseres Ansatzes erzielt werden können. Wir erreichen diese Ergebnisse mit Hilfe einer rein geometrischen Sichtweise, welche zu Berechnungen in der Grassmann-Mannigfaltigkeit führt.

Neben diesen Hauptergebnissen über das zufällige Verhalten der Kondition des konvexen Lösbarkeitsproblems werden wir auch einige Nebenergebnisse im Bereich der sphärischen Konvexgeometrie erzielen.

## Acknowledgments

First and foremost, I would like to thank my advisor Peter Bürgisser for his confidence, his encouragement, and for his ongoing support. I am very grateful for so many things, from his advice to apply at the PaSCo research training group, the introduction to the theory of condition numbers, all the way to his efforts in the exhausting process of turning a long and poorly written agglomeration of propositions and remarks into a readable manuscript. And of course, many of the most important ideas and perspectives only became clear during or after one of the extensive afternoon discussions, which I enjoyed very much.

Many thanks to James Renegar and Joachim Hilgert, who kindly agreed to act as referees for this work.

This work was partially supported by DFG grant BU-1371, but the main part of this work has been done during my time at the DFG Research Training Group "Scientific Computation" GRK 693 (PaSCo GK). This institution provided a great framework for free and creative research. I'd like to thank all the people from PaSCo for the comfortable atmosphere and for the universal support.

In Fall 2009 I had the great opportunity to spend a month at the Fields Institute in Toronto during the "Thematic Program on the Foundations of Computational Mathematics". In this inspiring time, besides the great number of most interesting talks I attended, I particularly enjoyed the stimulating conversations and the general creativity-filled air. Thank you to all who made this possible.

Special thanks go to Raphael Hauser and Martin Lotz for the invitation to an interesting and very enjoyable week at the University of Oxford. Hopefully I will get an opportunity to return the favor soon.

Last but not least I would like to thank all former and current members of the working group "Algebraic Complexity and Algorithmic Algebra" for the open and familiar atmosphere. Especially, I'd like to thank Christian Ikenmeyer and Stefan Mengel for useful remarks and for the fruitful and sometimes fundamental discussions during lunch break.

Zu guter Letzt möchte ich noch diese Gelegenheit nutzen und meinen Eltern für die vielen Jahre der Unterstützung danken, und für ihre Hilfe in allen Bereichen außerhalb der Mathematik.

# Contents

# Chapter 1

# Introduction

The focus of this paper is the homogeneous convex feasibility problem, which is the following simple question:

> *Given an m-dimensional subspace of $\mathbb{R}^n$, does this subspace intersect a fixed convex cone solely in the origin or are there further intersection points?*

This problem includes as special cases the linear, the second order, and the semidefinite feasibility problems, where one simply chooses the positive orthant, a product of Lorentz cones, or the cone of positive semidefinite matrices, respectively. We will give an average analysis of Renegar's condition number for several classes of convex cones, and one that is independent of the underlying convex cone. We will also describe a way of deriving smoothed analyses from our approach. We will achieve these results by adopting a purely geometric viewpoint leading to computations in the Grassmann manifold.

## 1.1 Complexity of numerical algorithms

Algorithms in computer science are usually discrete, i.e., they can be described as programs on a Turing machine. The complexity of these algorithms is therefore commonly measured by the amount of time/space the Turing machine needs during the computation. By contrast, numerical algorithms are usually described by operations on real numbers. Taking into account the internal representation of real numbers as floating point numbers, one could translate every (continuous) numerical algorithm into a "Turing machine program" and analyze it just as intrinsically discrete algorithms like for example 3-SAT. But the drawbacks of such a procedure are immediate: First of all, it would make an analysis extremely difficult, and second it would hide most of the essential information. For this reason, it is appropriate to change the model for numerical algorithms and replace the Turing machine by a BSS machine (see [6]), which can process real numbers as real numbers and thus has no need for a painful floating point routine. If we have a decision problem

$$f\colon \mathbb{R}^k \to \{0,1\} \;, \quad A \mapsto f(A) \;,$$

then it may happen that this problem is undecidable, i.e., there exists no BSS machine that computes the function $f$. This happens if the fibers $f^{-1}(0)$ resp. $f^{-1}(1)$ are "too complicated", for example if one of them is the Mandelbrot or a Julia set (cf. [6]). Problems coming from numerics or, as in our case, from convex programming, are not of this type. Broadly speaking, the boundary of the fibers

$\partial f^{-1}(0) = \partial f^{-1}(1)$ are "not too weird" and there are plenty of algorithms that compute $f$.

Assuming that we have a numerical algorithm that computes $f$, how to analyze its running time? It has turned out that for many numerical algorithms the *condition number* plays a decisive role. We may describe the notion of condition in the case of decision problems in the following way. Let us denote the fiber $f^{-1}(1)$ as the set of *feasible inputs*, as opposed to the set of *infeasible inputs*, which shall denote the fiber $f^{-1}(0)$. The interesting part is now the boundary $\partial f^{-1}(1) = \partial f^{-1}(0)$, which we call the set of *ill-posed inputs*. The reason for this convention is that ill-posed inputs should be seen as numerically intractable. More precisely, a slightest perturbation of an ill-posed input will make the output of the algorithm worthless, and not only the input itself, also the intermediate results share this extreme fragility. Of course, as we have mentioned earlier, the BSS machine has infinite precision so this is in theory no problem. But it should be evident that at least every practical algorithm has to handle inputs, which are close to the boundary, i.e., close to the set of ill-posed inputs, with much more care than very well-posed inputs, and usually all hope is lost when the input is ill-posed. The condition number is basically the inverse of the distance to the set of ill-posed inputs. Notice that this quantity only depends on the problem and the input, but not on the specific algorithm being used to solve this problem. This is the most important feature of the condition number and also the reason why the condition number is used for the analysis of the running time, because it is a purely geometric quantity that, as it turns out, captures all the complicated algorithmics. In summary, the higher the condition the closer the input to being ill-posed the worse the running time of the algorithm.

Numerical algorithms may fail on ill-posed inputs. Take for example the inversion procedure of $(n \times n)$-matrices, which is not defined on the set of singular matrices. For this reason, a worst-case analysis usually makes little sense for numerical algorithms because it is simply $\infty$. Instead, one may perform an average analysis, which consists of endowing the input space with a probability distribution, so that the running time becomes a random variable, and then compute the distribution or tail estimates or the expectation of this random variable.

Smale suggested in [51] to use the concepts of condition numbers and average analysis in a 2-part scheme for the analysis of numerical algorithms:

1. Bound the running time $T(A)$ via
$$T(A) \;\leq\; \Big(\text{size}(A) + (\text{log of})\ \text{condition}(A)\Big)^c ,$$
where $\text{size}(A)$ denotes the dimension of the input space, and $c$ is a universal constant; and

2. analyze $\text{condition}(A)$ under random $A$ by giving estimates of the tail
$$\text{Prob}\left[\text{condition}(A) \geq t\right] .$$

In contrast to worst-case analyses in computer science, which are usually assumed to be too pessimistic, as a single bad input may "ruin" the worst-case performance of an algorithm, average case analyses are usually assumed to be too optimistic, or at least not a convincing explanation for an observed good performance of an algorithm, as they strongly depend on the chosen distribution on the inputs. Usually, this distribution is chosen to be a gaussian or a uniform distribution so that the analysis becomes feasible, but these distributions are not likely to represent real-world scenarios.

As a way out of this dilemma, Spielman and Teng developed the new concept of smoothed analysis, which is a blend of worst case and average analysis. Broadly speaking, instead of considering all inputs (worst case), or one random input (average case), one considers all inputs endowed with a certain perturbation. The variation of this perturbation determines whether the smoothed analysis resembles more an average or a worst case analysis, and in some cases (cf. uniform smoothed analysis) it even has the form of an interpolation parameter. Let us give a more precise description of smoothed analysis by considering an input space $\mathcal{M}$, which is endowed with a metric, and which we assume to be compact and endowed with a probability measure (take for example the unit sphere and the usual volume normalized such that the volume of the whole sphere equals 1). Worst case, average, and smoothed analyses of the function $\mathscr{C}\colon \mathcal{M} \to \mathbb{R}$ are then simply the following three quantities

| worst case | average | smoothed |
|---|---|---|
| $\sup_{A\in\mathcal{M}} \mathscr{C}(A)$ | $\mathbb{E}_{A\in U(\mathcal{M})} \mathscr{C}(A)$ | $\sup_{\bar{A}\in\mathcal{M}} \mathbb{E}_{A\in U(B(\bar{A},\sigma))} \mathscr{C}(A)$ , |

where $U(\mathcal{M})$ denotes the uniform probability measure on $\mathcal{M}$, and $U(B(\bar{A},\sigma))$ denotes the uniform probability measure on the ball of radius $\sigma$ around $\bar{A}$. Note that for $\sigma = 0$ smoothed analysis becomes worst case analysis, and for $\sigma = \operatorname{diam}(\mathcal{M})$ smoothed analysis becomes average analysis. For completeness, we should also mention another commonly used perturbation model also known as Gaussian noise. Assuming that the given input space is a euclidean space $\mathbb{R}^N$, this model assumes that the input $A$ is drawn from a normal distribution centered at $\bar{A}$ with variance $\sigma^2$. The role of the specific perturbation is often secondary. In fact, there are robust smoothed analyses for several problems in numerics (see [21] and [13]), where a smoothed analysis for a large class of different perturbation models is provided, all leading to basically equivalent results.

For the matrix condition number there are even smoothed analyses for very general *discrete* perturbations by Tao and Vu (see [55]). Their techniques are very different compared to those used for continuous perturbations, but again the results are similar to the continuous case. It is an open and presumably very difficult question if one can give smoothed analyses under this kind of discrete perturbations for other condition numbers, like the Renegar or the Grassmann condition.

## 1.2 The convex feasibility problem

Let us anticipate some material that we will cover in Chapter 2 so that we can state the main results. In Chapter 2 we will have a closer look at the Renegar and the Grassmann condition.

A linear programming (LP) problem may be given in the following standard form:

$$\text{minimize } c^T x\ , \quad \text{subject to } Ax = b\,,\ x \geq 0\ , \qquad \text{(primal)}$$

$$\text{maximize } b^T y\ , \quad \text{subject to } A^T y \leq c\ , \qquad \text{(dual)}$$

where $A \in \mathbb{R}^{m\times n}$, $c, x \in \mathbb{R}^n$, $b, y \in \mathbb{R}^m$, and the inequalities are meant componentwise. Throughout the paper we will assume that $1 \leq m \leq n-1$, as the case $m \geq n$ is typically trivial.[1] LP-solvers usually work in two steps. First, they transform the

[1] If $m \geq n$ and $b \neq 0$, then for almost all $A \in \mathbb{R}^{m\times n}$ the system of linear equations $Ax = b$ either has a unique solution ($m = n$) or has no solutions ($m > n$).

problem into an equivalent form, which has a trivial or easy-to-find feasible point, i.e., a point $x_0$ satisfying $Ax_0 = b$ and $x_0 \geq 0$ in the primal case, or a point $y_0$ satisfying $A^T y_0 \leq c$ in the dual case. Then in the second step they minimize or maximize the corresponding linear functional via a simplex or an interior point method.

For simplicity, we will assume $b = 0$ and $c = 0$ and concentrate on the feasibility problem. In other words, we are interested in the *homogeneous (linear) feasibility* problem

$$\exists x \neq 0\ , \quad \text{such that } Ax = 0\,,\ x \geq 0\ , \tag{primal}$$

$$\exists y \neq 0\ , \quad \text{such that } A^T y \leq 0\ . \tag{dual}$$

Although it seems that this problem should be much easier than the original linear programming problem, it is in fact basically equivalent due to the duality theorem of linear programming.

The linear programming problem has a vast generalization to what is called *(general) convex programming*. Next, we will describe this generalization. Note that the (componentwise) inequality $v \geq w$, where $v, w \in \mathbb{R}^n$, can be paraphrased by the membership $v - w \in \mathbb{R}^n_+$, where $\mathbb{R}^n_+ = \mathbb{R}_+ \times \ldots \times \mathbb{R}_+$ denotes the positive orthant. In the primal convex programming problem the inequality $x \geq 0$ is thus replaced by the request $x \in C$, where $C \subset \mathbb{R}^n$ denotes a *regular cone*, which means the following:

- $C$ is a convex cone, i.e., for all $x, y \in C$ and positive $\lambda, \mu$ also $\lambda x + \mu y \in C$,
- $C$ is closed,
- $C$ has nonempty interior,
- $C$ does not contain a linear subspace of dimension $\geq 1$.

We call $C$ the *reference cone* of the convex programming problem. The *dual cone* $\breve{C}$ is defined by

$$\breve{C} := \{z \in \mathbb{R}^n \mid z^T x \leq 0\ \forall x \in C\}\ ,$$

and if $C$ is a regular cone then so is its dual $\breve{C}$ (cf. Section 3.1). In the dual convex programming problem the inequality $A^T y \leq c$ is replaced by the request $A^T y - c \in \breve{C}$. A special feature of most reference cones used in convex programming is that they are *self-dual*, which means that $\breve{C} = -C$. See for example the textbook [9] for more on the general convex programming problem.

Our focus lies on the *homogeneous convex feasibility problem*, which is the problem

$$\exists x \neq 0\ , \quad \text{such that } Ax = 0\,,\ x \in C\ , \tag{primal}$$

$$\exists y \neq 0\ , \quad \text{such that } A^T y \in \breve{C}\ , \tag{dual}$$

where $A \in \mathbb{R}^{m \times n}$, and $C \subset \mathbb{R}^n$ is a regular cone. We have already seen that the linear case follows by choosing $C = \mathbb{R}^n_+$ the positive orthant. Further interesting cases are second-order cone programming (SOCP) and semidefinite programming (SDP). They are obtained via the following choices of the reference cone $C$:

$$\begin{aligned}
\text{(LP):} \quad C &= \mathbb{R}^n_+ \\
\text{(SOCP):} \quad C &= \mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k} \\
\text{(SDP):} \quad C &= \mathrm{Sym}^\ell_+ \ := \ \{M \in \mathrm{Sym}^\ell \mid M \text{ is positive semidefinite}\}\ ,
\end{aligned} \tag{1.1}$$

where $\mathcal{L}^k := \{x \in \mathbb{R}^k \mid x_k \geq \|\bar{x}\|, \bar{x} = (x_1, \ldots, x_{k-1})\}$ shall denote the $k$-dimensional Lorentz cone, and $\mathrm{Sym}^\ell := \{M \in \mathbb{R}^{\ell \times \ell} \mid M^T = M\}$ the set of symmetric $(\ell \times \ell)$-matrices. These cones are all self-dual.

In terms of the general framework of decision problems that we described in the last section, we thus have for every regular cone $C \subset \mathbb{R}^n$ and $1 \leq m \leq n-1$ two decision problems

$$f_{\mathrm{P}} \colon \mathbb{R}^{m \times n} \to \{0,1\}\,, \quad A \mapsto \begin{cases} 1 & \text{if } \exists x \neq 0 : Ax = 0\,,\ x \in C \\ 0 & \text{else}\,, \end{cases}$$

$$f_{\mathrm{D}} \colon \mathbb{R}^{m \times n} \to \{0,1\}\,, \quad A \mapsto \begin{cases} 1 & \text{if } \exists y \neq 0 : A^T y \in \breve{C} \\ 0 & \text{else}\,. \end{cases}$$

To ease the notation we define the set of primal/dual feasible/infeasible instances

$$\begin{aligned} \mathcal{F}_{\mathrm{P}} &:= f_{\mathrm{P}}^{-1}(1)\,, \quad \mathcal{I}_{\mathrm{P}} := f_{\mathrm{P}}^{-1}(0)\,, \\ \mathcal{F}_{\mathrm{D}} &:= f_{\mathrm{D}}^{-1}(1)\,, \quad \mathcal{I}_{\mathrm{D}} := f_{\mathrm{D}}^{-1}(0)\,. \end{aligned}$$

A well-known theorem of alternatives says that for almost all $A \in \mathbb{R}^{m \times n}$ either the primal or the dual problem is feasible. More precisely, the boundaries of the fibers of $f_{\mathrm{P}}$ and $f_{\mathrm{D}}$ all coincide with the intersection $\mathcal{F}_{\mathrm{P}} \cap \mathcal{F}_{\mathrm{D}}$, i.e., we have

$$\mathcal{F}_{\mathrm{P}} \cap \mathcal{F}_{\mathrm{D}} \;=\; \partial\mathcal{F}_{\mathrm{P}} \;=\; \partial\mathcal{I}_{\mathrm{P}} \;=\; \partial\mathcal{F}_{\mathrm{D}} \;=\; \partial\mathcal{I}_{\mathrm{D}} \;=:\; \Sigma(C)$$

(see Proposition 2.2.1). This is the set of ill-posed inputs and a central object of our analysis. In terms of the functions $f_{\mathrm{P}}$ and $f_{\mathrm{D}}$ we have for $A \in \mathbb{R}^{m \times n} \setminus \Sigma(C)$

$$f_{\mathrm{P}}(A) = 1 - f_{\mathrm{D}}(A)\,.$$

In summary, we have for every regular cone $C \subset \mathbb{R}^n$ and every $1 \leq m \leq n-1$ a decision problem, which consists of the question whether the input is primal feasible or dual feasible (or ill-posed). A condition number for this feasibility problem is given by *Renegar's condition number*. This condition number is given by the inverse of the relative distance to the set of ill-posed inputs, i.e.,

$$\mathscr{C}_{\mathrm{R}} \colon \mathbb{R}^{m \times n} \setminus \{0\} \to [1, \infty]\,, \quad \mathscr{C}_{\mathrm{R}}(A) := \frac{\|A\|}{d(A, \Sigma(C))}\,,$$

where $\|A\|$ denotes the usual operator norm, and $d(A, \Sigma(C)) = \min\{\|A - A'\| \mid A' \in \Sigma(C)\}$. Equivalently, one can describe the inverse of the Renegar condition as the size of the maximum perturbation of $A$, which does not change the "feasibility property" of $A$,

$$\mathscr{C}_{\mathrm{R}}(A)^{-1} = \max\left\{ r \;\middle|\; \|\Delta A\| \leq r \cdot \|A\| \Rightarrow \begin{pmatrix} A + \Delta A \in \mathcal{F}^{\mathrm{P}} & \text{if } A \in \mathcal{F}^{\mathrm{P}} \\ A + \Delta A \in \mathcal{F}^{\mathrm{D}} & \text{if } A \in \mathcal{F}^{\mathrm{D}} \end{pmatrix} \right\}\,.$$

The Renegar condition is known to control the running time of geometric algorithms like for example the ellipsoid or interior-point methods. For example, in [61] an interior-point algorithm is described that solves the general homogeneous convex feasibility problem (for $C$ a self-scaled cone with a self-scaled barrier function) in

$$O(\sqrt{\nu_C} \cdot \ln(\nu_C \cdot \mathscr{C}_{\mathrm{R}}(A)))$$

interior-point iterations. Here, $\nu_C$ denotes the complexity parameter of a suitable barrier function for the reference cone $C$. The typical barrier functions for the LP-, the SOCP-, and the SDP-cone as defined in (1.1) yield the complexity parameters

$$\begin{aligned} \text{(LP):} \quad \nu_C &= n \\ \text{(SOCP):} \quad \nu_C &= 2 \cdot k \\ \text{(SDP):} \quad \nu_C &= \ell \,. \end{aligned}$$

(See for example [46] for more on this topic.) Additionally, it is shown in [61] that the condition number of the system of equations that is solved in each interior-point step is bounded by a factor of $\mathscr{C}_{\mathrm{R}}(A)^2$. See the references in [61] for further results on the estimate of running times of geometric algorithms for convex programming in terms of the Renegar condition.

The first part of Smale's 2-part scheme for the analysis of the convex feasibility problem is thus a well-studied question. In this work, we will treat the second part of this scheme, namely, we will address the questions about the random behavior of the condition of the convex feasibility problem.

For the analysis of the Renegar condition it turns our that it has the big drawback, that it mixes two causes for bad conditioning (cf. Section 2.2.1). The Grassmann condition is an attempt to overcome this drawback. One way to define it is via

$$\mathscr{C}_{\mathrm{G}}(A) := \begin{cases} \mathscr{C}_{\mathrm{R}}(A^\circ) & \text{if } \mathrm{rk}(A) = m \\ 1 & \text{if } \mathrm{rk}(A) < m \,, \end{cases}$$

where $A^\circ$ denotes the projection of $A$ on the Stiefel manifold $\mathbb{R}_\circ^{m\times n} := \{B \in \mathbb{R}^{m\times n} \mid BB^T = I_m\}$, which we like to call in this context the set of "balanced matrices". One can compute $A^\circ$ by replacing each singular value in $A$ by 1. In Chapter 2 we will discuss this in detail.

The Grassmann condition may be interpreted as a coordinate-free version of the Renegar condition as it solely depends on the kernel of $A$ (which is not immediate from the above definition). Furthermore, the Grassmann condition is the inverse of the sine of the (geodesic) distance to the set of ill-posed inputs in the Grassmann manifold, which is the reason for us to name this quantity the Grassmann condition. See Section 2.3 for more details.

The Grassmann condition is connected to the Renegar condition via the following two inequalities (cf. Theorem 2.3.4 in Section 2.3)

$$\mathscr{C}_{\mathrm{G}}(A) \;\le\; \mathscr{C}_{\mathrm{R}}(A) \;\le\; \kappa(A) \cdot \mathscr{C}_{\mathrm{G}}(A) \,, \tag{1.2}$$

where $\kappa(A)$ denotes the Moore-Penrose condition, i.e., the ratio between the largest and the smallest singular value of $A$. The random behavior, both average and smoothed, of the Moore-Penrose condition is a well-studied object, cf. [24], [20], [14]. So we may content ourselves with results about the random behavior of the Grassmann condition, as this will transfer to results about the Renegar condition through the above inequalities.

## 1.3 Tube formulas

Tube formulas naturally arise in the analysis of condition numbers, as is immediate from the following observation. If the condition number is given by the inverse distance to the set of ill-posed inputs, then the condition of an input $A$ exceeds $t$,

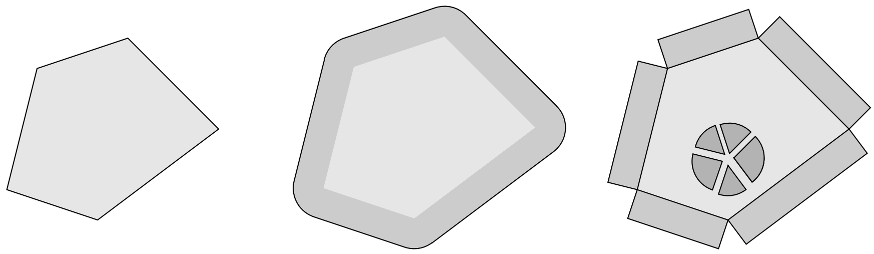


Figure 1.1: The tube around a polytope and its decomposition

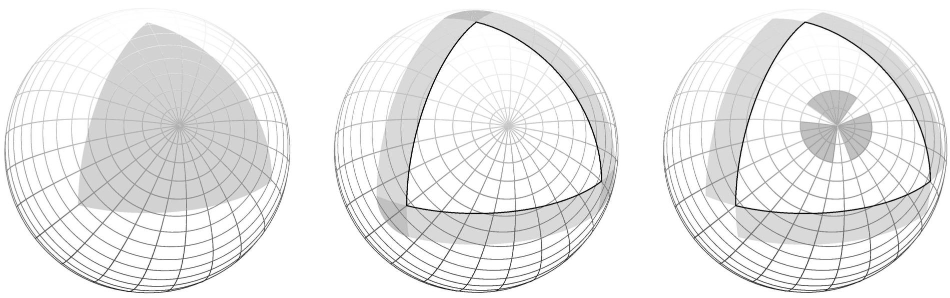


Figure 1.2: The decomposition of the tube around a polyhedral cap

iff the distance of $A$ to the set of ill-posed inputs is less than $1/t$, i.e., iff $A$ lies in the tube of radius $1/t$ around the set of ill-posed inputs. So if the input space is endowed with a probability measure $\mu$, then the probability that the condition number is larger than $t$ equals the $\mu$-volume of the tube of radius $1/t$ around the set of ill-posed inputs.

A prominent theorem by H. Weyl (see [63]) says that the Lebesgue-volume of a tube of radius $r$ around a compact submanifold of euclidean space or of the sphere basically has the form of a polynomial in $r$. In fact, the volume of the tube around a compact convex subset of euclidean space is a polynomial, the so-called Steiner polynomial. See Figure 1.1 for a 2-dimensional example. In the sphere this is not entirely true, as the monomials are replaced by some other functions due to the nonzero curvature of the sphere. So one could say that for small radius the volume of the tube is approximately a polynomial. See Figure 1.2 for an example in the 2-sphere.

We have already mentioned that the Grassmann condition is given by the inverse of the distance to the set of ill-posed inputs in the Grassmann manifold. The main step in our analysis is the derivation of a formula for the volume of the tube around the set of ill-posed inputs. In fact, for the interesting choices of $C$, i.e., those which yield LP, SOCP, or SDP, we will only get upper bounds, but we believe that these tube formulas are fairly sharp.

Estimating these tube formulas to get meaningful tail estimates then still remains a nontrivial task. For this reason we have decided also to include some first-order estimates. By this notion we mean that we only estimate the linear part of the polynomial in the tube formula and forget about the rest. This might seem radical at first sight, but for small radius the linear part is the decisive quantity. The simplification step eliminates numerous technical difficulties, which allows an improvement of the estimates for a large class of convex programs (see the next section). But of course these first-order results have to be taken with a grain of salt, as they are merely an indicator for the true average behavior of the Grassmann condition.

## 1.4 Results

Let $A \in \mathbb{R}^{m\times n}$, $m < n$, be a normal distributed random matrix, i.e., the entries of $A$ are i.i.d. standard normal. We will prove the following tail estimates and estimates of the expected logarithm of the Grassmann condition. If $C$ is *any regular cone* then

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \; < \; 6 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t} \; ,$$

if $t > n^{1.5}$. This yields the following estimate for the expected logarithm of the Grassmann condition

$$\mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] \; < \; 1.5 \cdot \ln(n) + 2 \; ,$$

if $n \geq 4$. This estimate of the expected condition of convex programming is to our knowledge the first known bound that holds in this generality. The first-order estimates show no significant changes except that the tail estimate gets a smaller multiplicative constant, and the assumption $t > n^{1.5}$ may be dropped.

In [18] it was observed that for linear programming the average behavior of the GCC condition number, which is a slight variation of the Renegar condition, may be estimated independently of $n$ and only depending on the smaller quantity $m$. We can achieve this also for the Grassmann condition. In fact, we will show that in the LP-case, i.e., for $C = \mathbb{R}^n_+$, we have for $t > m \geq 8$

$$\begin{aligned} \operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \; &< \; 29 \cdot \sqrt{m} \cdot \frac{1}{t} \; , \\ \mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] \; &< \; \ln(m) + 4 \; . \end{aligned}$$

The first-order estimates show again no significant changes except that the multiplicative constant for the tail estimate is significantly smaller, and the assumption $t > m \geq 8$ may be dropped.

In the definition of the GCC condition number the product structure of the positive orthant $\mathbb{R}^n_+ = \mathbb{R}_+ \times \ldots \times \mathbb{R}_+$ plays a fundamental role. It was therefore not clear to what extent the independence of $n$ in the linear programming case extends to more general classes of convex programming. We will show that for a second-order program with only one inequality, which we call SOCP-1, i.e., for the case $C = \mathcal{L}^n$, it holds that for $t > m \geq 8$

$$\begin{aligned} \operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \; &< \; 20 \cdot \sqrt{m} \cdot \frac{1}{t} \; , \\ \mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] \; &< \; \ln(m) + 3 \; . \end{aligned}$$

The first-order estimates show again only some minor changes in the constants. We may conclude that, although being merely a toy example, the SOCP-1 case is after the LP case the second convex program, where the average behavior of the condition is shown to be independent of the parameter $n$.

These are the nonasymptotic bounds that we will prove. Besides these we have first-order estimates which suggest that the independence of the parameter $n$ holds for *any self-dual cone*, so basically for any cone that is currently used for convex programming. More precisely, we will show that for the SOCP case we have the following first-order estimates

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \;\le\; 4 \cdot m \cdot \frac{1}{t} + \frac{g(n,m)}{t^2} \;,$$

where $g(n,m)$ denotes some function in $n$ and $m$. Note that if we had a nonasymptotic estimate of the form $\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \le c \cdot m \cdot \frac{1}{t}$ where $c$ is some positive constant, then this would imply $\mathbb{E}\,[\ln \mathscr{C}_{\mathrm{G}}(A)]) \le \ln(m) + \ln(c)$. The improvement of the first-order estimate to a nonasymptotic statement remains a (technical) problem that we are confident to solve in the near future.

The same first-order bounds hold for any self-dual cone, if Conjecture 4.4.17 is true. This conjecture states that the sequence of spherical intrinsic volumes form a unimodal sequence for self-dual cones. Log-concavity and unimodality is a highly useful and widely spread phenomenon (cf. the article [54]). The euclidean intrinsic volumes are known to be log-concave and thus unimodal, and we will show that the spherical intrinsic volumes of a product of Lorentz cones form a log-concave and thus unimodal sequence. We believe that Conjecture 4.4.16 and in particular Conjecture 4.4.17 hold, but could not prove it yet. So all this indicates that the independence of the parameter $n$ holds for any self-dual program.

These are the results for the average case. To state also the results of the smoothed analysis appropriately we needed some technical prerequisites that would interrupt the course of this section. Also, these results are only first order estimates, and compared to the average analyses they are not yet good enough to be seen as a true representation of the smoothed behavior of the Grassmann condition. That is why we leave out the discussion of these results at this point and refer to Section 7.3 for the details.

Besides these main results about the random behavior of the condition of the convex feasibility problem, we have obtained a couple of byproduct results that we will state next.

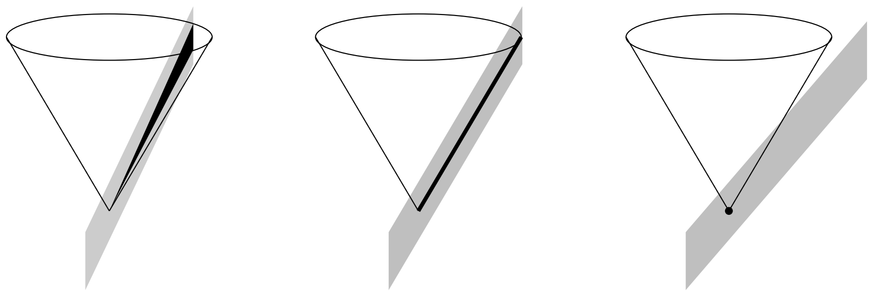
### 1.4.1 Spherical intrinsic volumes

In the context of tube formulas (cf. Section 1.3) one naturally arrives at the (euclidean or spherical) intrinsic volumes. The intrinsic volumes are important invariants of compact convex sets. We will derive a simple formula for the intrinsic volumes of a product of spherical convex sets, which is a direct analogon to the euclidean case. The formula for the euclidean case has been known before, but we will also give a new proof for this formula. We will also prove a simple conversion formula between the euclidean intrinsic volumes of the intersection of a closed convex cone $C$ with the unit ball and the spherical intrinsic volumes of the intersection of $C$ with the unit sphere. See Section 4.4 for more details, and see Chapter B for the proofs.

As a corollary, we will get that the intrinsic volumes of products of spherical balls form a log-concave sequence. The conjecture that this log-concavity property holds for all spherical convex sets is also first formulated in this paper.

### 1.4.2 The semidefinite cone

Finally, we mention one result that might seem secondary at first sight, but that we think has a great potential for further interesting research. This result is the computation of the intrinsic volumes of the cone of positive semidefinite matrices. In



(a) feasible (b) ill-posed (c) infeasible

Figure 1.3: The euclidean setting of the feasibility problem.

Section C.2 we will derive formulas for these intrinsic volumes (cf. also Section 4.4.1) that include a class of integrals that can be seen as an interesting variation of "Mehta's integral". We think that a good understanding of these integrals will lead to interesting new results about semidefinite programming (cf. Section C.3).

## 1.5 Outline

For the different stages of our analysis we will adopt different viewpoints on the homogeneous convex feasibility problem. First, we think of a subspace as the kernel or the image of a matrix; second, we identify a subspace with its intersection with the unit sphere; and third, we think of a subspace as a point in the Grassmann manifold. Each of these viewpoints has its own justification, and it is essential to change the viewpoint in order to get the analysis that we are aiming at.

The interpretation of a subspace as the kernel or the image of a matrix naturally comes from the applications. Manipulating a subspace by an algorithm means manipulating its defining matrix, so this viewpoint is the natural starting point for our analysis. In Chapter 2 we will recall the matrix condition (Moore-Penrose condition), describe the homogeneous convex feasibility problem, and recall the definition of Renegar's condition number. We will then give several equivalent definitions of the Grassmann condition and explain the interrelation between the Renegar and the Grassmann condition.

Replacing the subspaces by subspheres of the unit sphere comes from the specific form of the problem that we analyze. Note that the intersection of a subspace with a cone contains a nonzero point iff it contains a point of norm 1. In other words, nontrivial intersection of the subspace with a cone means nonempty intersection of the corresponding subsphere with the corresponding cap, i.e., with the intersection of the cone with the unit sphere. The advantage of this viewpoint is that the superfluous "cone direction" vanishes and reveals the essential geometry of the feasibility problem, which is in fact *spherical*. See Figure 1.3 and Figure 1.4 for a display of the euclidean and the spherical setting of the feasibility problem. In Chapter 3 we will deal with elementary topics in spherical convex geometry to prepare the ground for the upcoming analysis. Chapter 4 mainly treats Weyl's tube formulas and the consequential notions of (euclidean and spherical) intrinsic volumes.

In the third viewpoint we will consider a subspace simply as a point on the Grassmann manifold. This viewpoint is essential for the analysis as it becomes

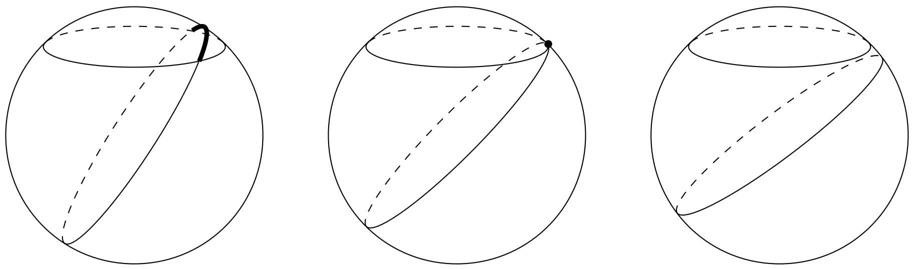



Figure 1.4: The spherical setting of the feasibility problem.

conceptually clear how to perform smoothed and average analyses. In Chapter 5 we will describe how we can perform computations in the Grassmann manifold. We will use this in Chapter 6 where we will prove an extension of Weyl's tube formula to certain hypersurfaces in the Grassmann manifold. This main result then allows us in Chapter 7 to achieve the tail estimates of the Grassmann condition that we already presented in the previous section.

## 1.6 Credits

The overall geometric methods we use for our analysis grew out of the series of papers [16], [15], [17]. In fact, this series was the starting point of our research, and our analysis was an attempt to transfer the results of [17] to the convex feasibility problem.

The concept of condition number in convex programming was introduced by Renegar in the 90's (see [43], [44], [45]). The Grassmann condition, which is the main object of our analysis, was already partly studied by Belloni and Freund [5]; this work led us to the correct relation (1.2) between the Grassmann condition and Renegar's condition number.

As for the spherical geometry and the spherical intrinsic volumes/curvature measures it was a great luck for us having found Glasauer's thesis [30] (see also [31] for a summary of the results). We see the relation between his work and ours as complementary, as his results are in some aspects more general but in some other aspects they are very restricted and even useless for the questions that we try to answer. While his approach mainly uses measure theoretic arguments, we use differential geometry as a basic tool. Our approach is thus more direct and gives more insight in those cases that are interesting for the convex feasibility problem.

Concerning the geometry of the Grassmann manifold the articles [25] and [3] were indispensable for us for the computations that we had to perform and also for the understanding of the great use of fiber bundles.

For surveys on smoothed analysis of condition numbers see [12] and [11] and the references therein. We will mention further sources that we rely on in the corresponding sections.

# Chapter 2

# The Grassmann condition

In this chapter we will recall some facts from numerical linear algebra, discuss the homogeneous convex feasibility problem and Renegar's condition number, and define the main object of this paper, the Grassmann condition.

## 2.1 Preliminaries from numerical linear algebra

Before we discuss the convex feasibility problem, we will review in this section some results from numerical linear algebra that we will need in the course of this paper. Namely, we will recall the matrix condition number, and separately we will discuss two further topics in the following two subsections. The first separate topic is the singular value decomposition, which is of central importance to understand the geometry of a linear operator and that we will also need for the computations in the Grassmann manifold in Chapter 5. The second topic is the projection map onto the Stiefel manifold and the question how a linear operator needs to be perturbed so that its kernel or the image of its adjoint contain a given point. This last subject will be a bit technical, but it will result in a clear connection between the Renegar and the Grassmann condition, that we will present in Section 2.3.

For square matrices $A \in \mathbb{R}^{n\times n}$ the matrix condition number is defined as

$$\kappa(A) := \begin{cases} \|A\| \cdot \|A^{-1}\| & \text{if } \operatorname{rk}(A) = n \\ \infty & \text{if } \operatorname{rk}(A) < n \,. \end{cases}$$

Here, and throughout the paper, $\|A\|$ denotes the usual spectral norm of $A \in \mathbb{R}^{n\times n}$, i.e.,

$$\|A\| = \max\{\|Ax\| \mid x \in \mathbb{R}^n \,,\; \|x\| = 1\} \,,$$

where $\|x\|$ denotes the euclidean norm of $x \in \mathbb{R}^n$.

The definition of the matrix condition number may be generalized to the rectangular case in the following way. Here, and throughout the paper, when we consider rectangular matrices $\mathbb{R}^{m\times n}$, we assume $m \leq n$. Let $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rk}(A) = m$. Interpreting $A$ as a linear operator $A\colon \mathbb{R}^n \to \mathbb{R}^m$, let $\bar{A} := A|_{(\ker A)^\perp}$ denote the restriction of $A$ to the orthogonal complement of the kernel of $A$. The restriction $\bar{A}$ is a linear isomorphism and thus has an inverse $\bar{A}^{-1}$. The *Moore-Penrose inverse* $A^\dagger \in \mathbb{R}^{n\times m}$ of $A$ may now be defined as

$$A^\dagger\colon \mathbb{R}^m \to \mathbb{R}^n \,, \quad A^\dagger(x) := \bar{A}^{-1}(x) \,.$$

The generalization of the matrix condition number to the rectangular case is given by

$$\kappa(A) := \begin{cases} \|A\| \cdot \|A^\dagger\| & \text{if } \operatorname{rk}(A) = m \ , \\ \infty & \text{if } \operatorname{rk}(A) < m \ . \end{cases}$$

The Eckart-Young Theorem (cf. Theorem 2.1.5) characterizes the matrix condition as the relativized inverse distance to the set of rank-deficient matrices. Let us denote the set of rank-deficient $(m \times n)$-matrices and its complement, the set of full-rank $(m \times n)$-matrices, by

$$\begin{aligned} \mathbb{R}^{m\times n}_{\mathrm{rd}} &:= \{A \in \mathbb{R}^{m\times n} \mid \operatorname{rk}(A) < m\} \\ \mathbb{R}^{m\times n}_{*} &:= \mathbb{R}^{m\times n} \setminus \mathbb{R}^{m\times n}_{\mathrm{rd}} = \{A \in \mathbb{R}^{m\times n} \mid \operatorname{rk}(A) = m\} \ . \end{aligned}$$

Besides the spectral norm we also employ the Frobenius norm

$$\|A\|_F = \Big(\sum_{i,j} a_{ij}^2\Big)^{1/2} \ ,$$

where $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. An important property of the spectral norm and of the Frobenius norm is their invariance under multiplication by elements of the orthogonal group $O(n) = \{Q \in \mathbb{R}^{n\times n} \mid Q^T Q = I_n\}$:[1]

$$\|Q_1 \cdot A \cdot Q_2^T\| = \|A\| \ , \quad \|Q_1 \cdot A \cdot Q_2^T\|_F = \|A\|_F$$

for all $A \in \mathbb{R}^{m\times n}$, $Q_1 \in O(m)$, $Q_2 \in O(n)$. If $A \in \mathbb{R}^{m\times n}$ has rank 1, then $A$ can be written in the form $A = u \cdot v^T$ for some $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, and we have

$$\|A\| = \|A\|_F = \|u\| \cdot \|v\| \ , \tag{2.1}$$

where $\|x\|$, $x \in \mathbb{R}^k$, denotes as usual the euclidean norm of $x$.

We denote the metric on $\mathbb{R}^{m\times n}$ induced by the spectral norm by $d(A,B) := \|A - B\|$, and the metric induced by the Frobenius norm by $d_F(A,B) := \|A - B\|_F$. For $A \in \mathbb{R}^{m\times n}$ and $\mathcal{M} \subset \mathbb{R}^{m\times n}$ we define

$$d(A, \mathcal{M}) := \inf\{d(A,B) \mid B \in \mathcal{M}\} \ ,$$

and similarly for $d_F(A, \mathcal{M})$.

**Proposition 2.1.1.** *Let $A \in \mathbb{R}^{m\times n}_*$ be a full-rank $(m\times n)$-matrix. Then the distance of $A$ to the set of rank-deficient matrices is the same for the operator norm as for the Frobenius norm, i.e.,*

$$d(A, \mathbb{R}^{m\times n}_{\mathrm{rd}}) = d_F(A, \mathbb{R}^{m\times n}_{\mathrm{rd}}) \ .$$

*Furthermore, the matrix condition of $A$ is given by the inverse of the relative distance of $A$ to the set of rank-deficient matrices, i.e.,*

$$\kappa(A) = \frac{\|A\|}{d(A, \mathbb{R}^{m\times n}_{\mathrm{rd}})} = \frac{\|A\|}{d_F(A, \mathbb{R}^{m\times n}_{\mathrm{rd}})} \ .$$

*Proof.* This follows from the Eckart-Young Theorem 2.1.5 treated below. □

[1] This is unfortunately the classical notation for the orthogonal group; but the danger of confusing $O(n)$ with the set of at most linearly growing functions is marginal, as the context will make the notation unambiguous.

### 2.1.1 The singular value decomposition

For most of our considerations involving matrices, the singular value decomposition is of fundamental importance, as it reveals the basic geometric properties of the linear operator.

**Theorem 2.1.2** (SVD). *Let $A \in \mathbb{R}^{m\times n}$, $m \le n$. Then there exist orthogonal matrices $Q_1 \in O(m)$, $Q_2 \in O(n)$, and uniquely determined nonnegative constants $\sigma_1 \ge \ldots \ge \sigma_m \ge 0$, such that*

$$A = Q_1 \cdot \begin{pmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_m & 0 & \cdots & 0 \end{pmatrix} \cdot Q_2^T \ . \tag{2.2}$$

*Proof.* See for example [32, Sec. 2.5.3] or [60, Lect. 4]. □

The decomposition of $A$ in (2.2) is called a *singular value decomposition (SVD)* of $A$ and the constants $\sigma_1 \ge \ldots \ge \sigma_m$ are called the singular values of $A$. The singular values have a geometric interpretation as the lengths of the semi-axes of the hyperellipsoid given by the image of the $m$-dimensional unit ball under the map $A$. We state this in the following proposition.

**Proposition 2.1.3.** *Let $A \in \mathbb{R}^{m\times n}$, $m \le n$, and let $B_m$ denote the unit ball in $\mathbb{R}^m$ centered at the origin. Then the singular values $\sigma_1 \ge \ldots \ge \sigma_m$ of $A$ coincide with the lengths of the semi-axes of the hyperellipsoid $A(B_m)$. In particular,*

$$\sigma_1 = \min\{r \mid B_n(r) \supseteq A(B_m)\} \ ,$$
$$\sigma_m = \max\{r \mid B_n(r) \subseteq A(B_m)\} \ ,$$

*where $B_n(r)$ denotes the $n$-dimensional ball of radius $r$ around the origin.*

*Proof.* See for example [60, Lect. 4]. □

With the help of the singular value decomposition we can give new formulas for the aforementioned quantities associated to $A$.

**Corollary 2.1.4.** *Let $A \in \mathbb{R}^{m\times n}$, $m \le n$, have a SVD as in (2.2). Then*

1. *$\|A\| = \sigma_1$ and $\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_m^2}$,*
2. *$A^\dagger = Q_2 \cdot \begin{pmatrix} D \\ 0 \end{pmatrix} \cdot Q_1^T$, with $D := \operatorname{diag}(\sigma_1^{-1}, \ldots, \sigma_m^{-1})$, if $\operatorname{rk}(A) = m$,*
3. *$\|A^\dagger\| = \sigma_m^{-1}$, and $\kappa(A) = \frac{\sigma_1}{\sigma_m}$, if $\operatorname{rk}(A) = m$.*

*Proof.* The first part follows from the invariance of the operator norm and the Frobenius norm under left and right multiplication by orthogonal matrices. For the second part we may assume w.l.o.g. that $Q_1 = I_m$ and $Q_2 = I_n$. For this case the claim is easily verified from the definition of $A^\dagger$. The third part follows from the first two parts. □

Recall that for $A \in \mathbb{R}^{m\times n}$ and $\mathcal{M} \subset \mathbb{R}^{m\times n}$ we have defined the distance $d(A, \mathcal{M}) = \inf\{d(A, B) \mid B \in \mathcal{M}\}$, and similarly for $d_F(A, \mathcal{M})$. We additionally define

$$\operatorname{argmin} d(A, \mathcal{M}) := \{B \in \mathcal{M} \mid d(A, B) = d(A, \mathcal{M})\} \ ,$$

and similarly we define $\operatorname{argmin} d_F(A, \mathcal{M})$. The next result implies Proposition 2.1.1.

**Theorem 2.1.5** (Eckart-Young)**.** *Let $A \in \mathbb{R}_*^{m\times n}$ and let $A$ have a singular value decomposition as in* (2.2)*. Furthermore, let*

$$B := Q_1 \cdot \begin{pmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_{m-1} & 0 & \cdots & 0 \\ & & 0 & 0 & \cdots & 0 \end{pmatrix} \cdot Q_2^T \,. \tag{2.3}$$

*Then $B \in \operatorname{argmin} d(A, \mathbb{R}_{\mathrm{rd}}^{m\times n}) \cap \operatorname{argmin} d_F(A, \mathbb{R}_{\mathrm{rd}}^{m\times n})$. In particular,*

$$d(A, \mathbb{R}_{\mathrm{rd}}^{m\times n}) \;=\; d_F(A, \mathbb{R}_{\mathrm{rd}}^{m\times n}) \;=\; \sigma_m \,.$$

*Proof.* First of all, it is easily seen from the singular value decomposition that $\|Av\| \geq \sigma_m$ for all $v \in S^{n-1}$. Let $B' \in \mathbb{R}_{\mathrm{rd}}^{m\times n}$, and let $v \in \ker B' \cap S^{n-1}$. Then

$$\sigma_m \leq \|Av\| = \|(A - B')v\| \,,$$

which shows that $\|A - B'\| \geq \sigma_m$. So we get $d_F(A, \mathbb{R}_{\mathrm{rd}}^{m\times n}) \geq d(A, \mathbb{R}_{\mathrm{rd}}^{m\times n}) \geq \sigma_m$, and as $B \in \mathbb{R}_{\mathrm{rd}}^{m\times n}$ and $d_F(A, B) = \sigma_m$ we have indeed an equality. □

### 2.1.2 Effects of matrix perturbations

In this section we will strive for a clear picture about how to perturb a given matrix such that the defining subspaces contain some given point. We summarize the result in the next proposition.

For $x \in \mathbb{R}^n \setminus \{0\}$ and $\mathcal{W} \subseteq \mathbb{R}^n$ a linear subspace, we define the angle between $x$ and $\mathcal{W}$ via

$$\sphericalangle(x, \mathcal{W}) := \arccos\left(\tfrac{\|\Pi_{\mathcal{W}}(x)\|}{\|x\|}\right) \in [0, \tfrac{\pi}{2}] \,, \tag{2.4}$$

where $\Pi_{\mathcal{W}}$ denotes the orthogonal projection onto $\mathcal{W}$ (cf. Chapter 3). Note that if $\mathcal{W} = \operatorname{im} A^T$ for some $A \in \mathbb{R}^{m\times n}$, then the orthogonal complement of $\mathcal{W}$ is given by $\mathcal{W}^\perp = \ker A$. Note also that the angle between $x$ and $\mathcal{W}^\perp$ is given by

$$\sphericalangle(x, \mathcal{W}^\perp) = \tfrac{\pi}{2} - \sphericalangle(x, \mathcal{W}) \,.$$

**Theorem 2.1.6.** *Let $A \in \mathbb{R}_*^{m\times n}$ with singular values $\sigma_1 \geq \ldots \geq \sigma_m > 0$, and let $x \in \mathbb{R}^n \setminus \ker(A)$. Furthermore, let $\mathcal{W} := \operatorname{im}(A^T)$, and let $\alpha := \sphericalangle(x, \mathcal{W})$ and $\beta := \sphericalangle(x, \mathcal{W}^\perp) = \frac{\pi}{2} - \alpha$.*

*1a. If $\Delta \in \mathbb{R}^{m\times n}$ is such that $x \in \operatorname{im}(A^T + \Delta^T)$, then*

$$\|\Delta\| \geq \sigma_m \cdot \sin\alpha \,.$$

*1b. There exists $\Delta_0 \in \mathbb{R}^{m\times n}$ such that $x \in \operatorname{im}(A^T + \Delta_0^T)$ and*

$$\|\Delta_0\| = \|\Delta_0\|_F \leq \sigma_1 \cdot \sin\alpha \,.$$

*2a. If $\Delta' \in \mathbb{R}^{m\times n}$ is such that $x \in \ker(A + \Delta')$, then*

$$\|\Delta'\| \geq \sigma_m \cdot \sin\beta \,.$$

*2b. There exists $\Delta_0' \in \mathbb{R}^{m\times n}$ such that $x \in \ker(A + \Delta_0')$ and*

$$\|\Delta_0'\| = \|\Delta_0'\|_F \leq \sigma_1 \cdot \sin\beta \,.$$

For the proof of this proposition and for the analysis of Renegar's condition number in general, it will be crucial to consider matrices whose singular values all coincide, say whose singular values are all equal to 1. Note that this is equivalent to the assumption that $\|A\| = \kappa(A) = 1$ and to the property that the rows of $A$ are orthonormal vectors in $\mathbb{R}^n$. We denote the set of these matrices by

$$\mathbb{R}^{m\times n}_{\circ} := \{A \in \mathbb{R}^{m\times n} \mid \|A\| = \kappa(A) = 1\} , \tag{2.5}$$

and we call these matrices *balanced.*[2] Note that for $m = n$ the set of balanced matrices is the orthogonal group, i.e., $\mathbb{R}^{n\times n}_{\circ} = O(n)$.

While the set of rank-deficient matrices is the boundary of $\mathbb{R}^{m\times n}_{*}$ in $\mathbb{R}^{m\times n}$, the set of balanced matrices may be thought of as the center of $\mathbb{R}^{m\times n}_{*}$. This is specified by the following proposition.

**Proposition 2.1.7.** *The set of balanced matrices is given by*

$$\begin{aligned}\mathbb{R}^{m\times n}_{\circ} &= \operatorname{argmax}\{d(A,\mathbb{R}^{m\times n}_{\mathrm{rd}}) \mid A \in \mathbb{R}^{m\times n} ,\ \|A\| = 1\}\\ &= \operatorname{argmax}\{d(A,\mathbb{R}^{m\times n}_{\mathrm{rd}}) \mid A \in \mathbb{R}^{m\times n} ,\ \|A\|_F = \sqrt{m}\} .\end{aligned}$$

In other words, a matrix $A \in \mathbb{R}^{m\times n}$ with $\|A\| = 1$ or $\|A\|_F = \sqrt{m}$ is balanced iff it maximizes the distance to $\mathbb{R}^{m\times n}_{\mathrm{rd}}$ among all matrices with the corresponding normalization.

*Proof.* From Theorem 2.1.5 we have $d(A,\mathbb{R}^{m\times n}_{\mathrm{rd}}) = \sigma_m$, where $\|A\| = \sigma_1 \geq \ldots \geq \sigma_m$ denote the singular values of $A$. So for $\|A\| = 1$ we have $d(A,\mathbb{R}^{m\times n}_{\mathrm{rd}}) = \|A\| = 1$ iff $A \in \mathbb{R}^{m\times n}_{\circ}$. This implies the first equality.

As for the second equality, note that $\|A\|_F = \sqrt{m}$ implies $\sigma_m \leq 1$. Moreover, for $\|A\|_F = \sqrt{m}$ we have $\sigma_m = 1$ iff $\sigma_m = \sigma_{m-1} = \ldots = \sigma_1 = 1$, i.e., iff $A \in \mathbb{R}^{m\times n}_{\circ}$. □

**Proposition 2.1.8.** *Let $A \in \mathbb{R}^{m\times n}_{*}$ have a SVD as in* (2.2). *Then the set of balanced matrices, which minimize the Frobenius distance to $A$, consists of exactly one element $A^\circ$ where*

$$A^\circ = Q_1 \cdot \begin{pmatrix} 1 & & & 0 & \cdots & 0\\ & \ddots & & \vdots & & \vdots\\ & & 1 & 0 & \cdots & 0\end{pmatrix} \cdot Q_2^T . \tag{2.6}$$

We call $A^\circ$ the *balanced approximation* of $A$.

*Proof.* If $B \in \mathbb{R}^{m\times n}_{\circ}$ then $\|A - B\|_F = \|A^T - B^T\|_F = \|A^T Q_1 - B^T Q_1\|_F$, and the columns of $B^T Q_1$, which we denote by $w_1, \ldots, w_m$, are orthonormal vectors in $\mathbb{R}^n$. Let $v_1, \ldots, v_n$ denote the columns of $Q_2$, so that the columns of $A^T Q_1$ are given by $\sigma_1 v_1, \ldots, \sigma_m v_m$. As

$$\|\sigma_i \cdot v_i - w_i\| \geq |\sigma_i - 1|$$

with equality iff $w_i = v_i$ (use $\sigma_i > 0$), we get

$$\|A - B\|_F^2 \ = \ \|A^T Q_1 - B^T Q_1\|_F^2 \ = \ \sum_{i=1}^{m} \|\sigma_i \cdot v_i - w_i\|^2 \ \geq \ \sum_{i=1}^{m} |\sigma_i - 1|^2$$

with equality iff the columns of $B^T Q_1$ are given by $v_1, \ldots, v_m$. This proves the claim. □

[2] The set of balanced matrices is also called the *Stiefel manifold*. We prefer the term 'balanced matrix' at this point because it emphasizes the property of $A$ defining a linear operator.

The balanced approximation $A^\circ$ also appears in the so-called *polar decomposition* of a linear operator (cf. [32, §4.2.10]). We describe this in the following proposition.

**Proposition 2.1.9.** *Let $A \in \mathbb{R}^{m\times n}_*$ have a SVD as in (2.2). Then there exist $B \in \mathbb{R}^{m\times n}_\circ$ and a symmetric positive definite matrix $S \in \mathbb{R}^{m\times m}$ such that*

$$A = S \cdot B \,. \tag{2.7}$$

*The matrices $S$ and $B$ are uniquely determined and given by*

$$\begin{aligned} B &= A^\circ = (\sqrt{AA^T})^{-1} \cdot A \,, \\ S &= Q_1 \cdot D \cdot Q_1^T = \sqrt{AA^T} \,, \end{aligned}$$

*with $D := \operatorname{diag}(\sigma_1, \ldots, \sigma_m)$.*

*Proof.* If $A$ has a decomposition as in (2.7), then we have

$$A \cdot A^T = S \underbrace{B \cdot B^T}_{=I_m} \underbrace{S^T}_{=S} = S^2 \,.$$

This implies $S = \sqrt{AA^T}$ and $B = (\sqrt{AA^T})^{-1} \cdot A$. In particular, $S$ and $B$ are uniquely determined by the decomposition (2.7).

From the singular value decomposition (2.2) of $A$, we get

$$A = Q_1 \cdot \begin{pmatrix} D & 0 \end{pmatrix} \cdot Q_2^T = \underbrace{Q_1\, D\, Q_1^T}_{\text{symm. pos. def.}} \cdot \underbrace{Q_1 \cdot \begin{pmatrix} I_m & 0 \end{pmatrix} \cdot Q_2^T}_{=A^\circ \in \mathbb{R}^{m\times n}_\circ} \,,$$

which is a decomposition of the form (2.7). From the uniqueness of the decomposition (2.7) it follows that $Q_1\, D\, Q_1^T = \sqrt{AA^T}$ and $A^\circ = (\sqrt{AA^T})^{-1} \cdot A$. □

**Remark 2.1.10.** Let $A \in \mathbb{R}^{m\times n}_*$ and let $A = S \cdot A^\circ$ be the polar decomposition of $A$ as described in Proposition 2.1.9. Then for $x \in \mathbb{R}^n$ and $\Delta, \Delta' \in \mathbb{R}^{m\times n}$ we have

$$\begin{aligned} x \in \operatorname{im}(A^T + \Delta^T) &\iff x \in \operatorname{im}((A^\circ)^T + \Delta^T S^{-1}) \,, \\ x \in \ker(A + \Delta') &\iff x \in \ker(A^\circ + S^{-1}\Delta') \,. \end{aligned}$$

Another useful property of the balanced approximation $A^\circ$ is that it gives an easy formula for the orthogonal projection onto the image of $A$. We formulate this in the following lemma.

**Lemma 2.1.11.** *Let $A \in \mathbb{R}^{m\times n}_*$, and let $\mathcal{W} := \operatorname{im}(A^T)$. Then the orthogonal projection $\Pi_{\mathcal{W}}$ onto $\mathcal{W}$ is given by*

$$\Pi_{\mathcal{W}} = (A^\circ)^T A^\circ \,.$$

*Proof.* Let $x \in \mathbb{R}^n$ be decomposed into $x = y + z$ with $y \in \mathcal{W}$ and $z \in \mathcal{W}^\perp$. As the rows of $A^\circ$ form an orthonormal basis of $\mathcal{W}$, we have $y = (A^\circ)^T \cdot v$ for some $v \in \mathbb{R}^m$, and $A^\circ \cdot z = 0$. Therefore, we have

$$\begin{aligned} (A^\circ)^T A^\circ \cdot x &= (A^\circ)^T A^\circ \cdot (y + z) = (A^\circ)^T A^\circ (A^\circ)^T \cdot v + (A^\circ)^T A^\circ \cdot z \\ &= (A^\circ)^T \cdot v = y = \Pi_{\mathcal{W}}(x) \,. \end{aligned}$$

□

The following Propositions 2.1.12/2.1.14 cover Theorem 2.1.6 for balanced matrices. Here, and throughout the paper, we denote by $S^{n-1}$ the unit sphere in $\mathbb{R}^n$, i.e.,

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\} .$$

**Proposition 2.1.12.** *Let $B \in \mathbb{R}_\circ^{m\times n}$ and let $\mathcal{W} := \mathrm{im}(B^T)$ the image of $B^T$, or equivalently $\ker(B) = \mathcal{W}^\perp$. Furthermore, let $x \in \mathbb{R}^n \setminus \{0\}$, and let $\alpha := \sphericalangle(x, \mathcal{W})$ and $\beta := \sphericalangle(x, \mathcal{W}^\perp) = \frac{\pi}{2} - \alpha$. Then for $\Delta, \Delta' \in \mathbb{R}^{m\times n}$*

$$\begin{aligned} x \in \mathrm{im}(B^T + \Delta^T) \quad &\Rightarrow \quad \|\Delta\| \geq \sin\alpha , \\ x \in \ker(B + \Delta') \quad &\Rightarrow \quad \|\Delta'\| \geq \sin\beta . \end{aligned}$$

*Proof.* If $x \in \mathrm{im}(B^T + \Delta^T)$, then there exist $v \in S^{m-1}$ and $r \in \mathbb{R}$ such that $(B^T + \Delta^T) \cdot v = r \cdot x$. Then we have

$$\|\Delta\| \;\geq\; \|\Delta^T \cdot v\| = \| \underbrace{r \cdot x}_{\in \mathrm{lin}\{x\}} - \underbrace{B^T \cdot v}_{\in S^{n-1}\cap\mathcal{W}} \| \;\geq\; \sin\theta \;\geq\; \sin\alpha ,$$

where $\theta$ denotes the angle between $x$ and $B^T v$.

If $(B + \Delta') \cdot x = 0$, we may compute, using the abbreviation $x^\circ = \|x\|^{-1} \cdot x$,

$$\|\Delta'\| \;\geq\; \|\Delta' \cdot x^\circ\| \;=\; \|B \cdot x^\circ\| \;=\; \|B^T B \cdot x^\circ\| \;=\; \cos\alpha \;=\; \sin\beta ,$$

as $B^T B$ is the orthogonal projection onto $\mathcal{W}$. □

We thus have lower bounds for the norm of perturbations of matrices in $\mathbb{R}_\circ^{m\times n}$ such that the image resp. the kernel contain some given point. We will show next that these lower bounds are sharp by constructing perturbations which have the given Frobenius norms. For the geometric picture it is useful to consider rotations in $\mathbb{R}^n$. A rotation takes place in a 2-dimensional subspace $L$ of $\mathbb{R}^n$, and leaves the orthogonal complement $L^\perp$ fixed. For the definition of such a rotation we additionally need a rotational direction, which is defined by a pair of linearly independent points in $L$.

**Definition 2.1.13.** Let $L \subset \mathbb{R}^n$ a 2-dimensional subspace, and let $p, q \in L$ linearly independent. Then for $\rho \in \mathbb{R}$ we denote by $D_{L,(p,q)}(\rho)$ the matrix of the linear operation, which leaves $L^\perp$ fixed and rotates $L$ by an angle of $\rho$, such that $p$ rotates towards $q$. If $\|p\| = \|q\| = 1$, $p^T q = 0$, and $b_1, \ldots, b_{n-2} \in L^\perp$ are chosen such that the matrix formed by these vectors $Q := \begin{pmatrix} p & q & b_1 & \cdots & b_{n-2} \end{pmatrix} \in O(n)$, then

$$D_{L,(p,q)}(\rho) = Q \cdot \begin{pmatrix} \cos\rho & -\sin\rho & & & \\ \sin\rho & \cos\rho & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \cdot Q^T .$$

**Proposition 2.1.14.** *Let $B \in \mathbb{R}_\circ^{m\times n}$ a balanced matrix with image $\mathcal{W} := \mathrm{im}(B^T)$, let $x \in \mathbb{R}^n \setminus \mathcal{W}^\perp$, $\|x\| = 1$, and let $\alpha := \sphericalangle(x, \mathcal{W})$. Furthermore, let $\Delta, \Delta' \in \mathbb{R}^{m\times n}$ be defined by*

$$\Delta := B \cdot pp^T \cdot (xx^T - I_n) , \quad \Delta' := -B \cdot xx^T ,$$

*where $p \in \mathcal{W}$ denotes the normalized projection of $x$ on $\mathcal{W}$, i.e., $p := (\cos\alpha)^{-1} \cdot B^T Bx$. Then we have $\mathrm{rk}(\Delta), \mathrm{rk}(\Delta') \leq 1$, $\|\Delta\|_F = \|\Delta\| = \sin\alpha$, $\|\Delta'\|_F = \|\Delta'\| = \cos\alpha$, and*

$$x \in \mathrm{im}(B^T + \Delta^T) , \quad x \in \ker(B + \Delta') .$$

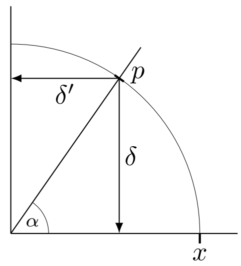



Figure 2.1: The 2-dimensional situation in $L = \mathrm{lin}\{p, x\}$.

*Additionally, if* $x \notin \mathcal{W} \cup \mathcal{W}^\perp$ *and* $\mathcal{W}_\Delta := \mathrm{im}(B^T + \Delta^T)$, $\mathcal{W}_{\Delta'} := \ker(B + \Delta')$, *then*

$$\mathcal{W}_\Delta = D_{L,(p,x)}(\alpha) \cdot \mathcal{W} \quad \text{and} \quad \mathcal{W}_{\Delta'} = D_{L,(x,p)}(\tfrac{\pi}{2} - \alpha) \cdot \mathcal{W}^\perp \tag{2.8}$$

*where* $L := \mathrm{lin}\{p, x\}$.

*Proof.* First of all, let us assume that $Bp = e_1 \in \mathbb{R}^m$ the first canonical basis vector. Afterwards, we will deduce the general statement from this special case.

The first row of $B$ is thus given by $p^T$. As $p^T x = \cos\alpha$ we have $Bx = \cos\alpha \cdot e_1$ and

$$\Delta = e_1 \cdot (\cos\alpha \cdot x - p)^T \ , \quad \Delta' = -\cos\alpha \cdot e_1 \cdot x^T \ .$$

Furthermore, we have $x \in \mathrm{im}(B^T + \Delta^T)$ as

$$\begin{aligned}(B^T + \Delta^T) \cdot e_1 &= B^T Bp + (\cos\alpha \cdot x - p) \cdot e_1^T e_1 = p + \cos\alpha \cdot x - p \\ &= \cos\alpha \cdot x \ ,\end{aligned}$$

and $x \in \ker(B + \Delta')$ as

$$\begin{aligned}(B + \Delta') \cdot x &= Bx - \cos\alpha \cdot e_1 \cdot x^T x = \cos\alpha \cdot e_1 - \cos\alpha \cdot e_1 \\ &= 0 \ .\end{aligned}$$

As for the rank, we have $\mathrm{rk}\,\Delta' = 1$. Moreover, $\mathrm{rk}\,\Delta = 1$ except for the case $x \in \mathcal{W}$, where $\alpha = 0$, $x = p$, and thus $\mathrm{rk}\,\Delta = 0$. Therefore, the norms of $\Delta$ and $\Delta'$ are given by (cf. (2.1))

$$\|\Delta\| = \|\Delta\|_F = \sin\alpha \ , \quad \|\Delta'\| = \|\Delta'\|_F = \cos\alpha \ .$$

As for the claim in (2.8), about $\mathcal{W}_t$ and $\mathcal{W}'_t$ it is readily checked that on the orthogonal complement of $L = \mathrm{lin}\{p, x\}$ the linear operators $B$, $B + \Delta$, and $B + \Delta'$ all coincide, i.e.,

$$B|_{L^\perp} = (B + \Delta)|_{L^\perp} = (B + \Delta')|_{L^\perp} \ .$$

It thus suffices to consider the 2-dimensional situation in $L$. For convenience, let us define

$$\delta := \cos\alpha \cdot x - p \ , \quad \delta' = -\cos\alpha \cdot x \ .$$

Note that $\Delta = e_1 \cdot \delta^T$ and $\Delta' = e_1 \cdot \delta'^T$. See Figure 2.1 for a display of the situation in $L$. From this picture it is readily checked that indeed $\mathcal{W}_\Delta = D_{L,(p,x)}(\alpha) \cdot \mathcal{W}$, and $\mathcal{W}_{\Delta'} = D_{L,(x,p)}(\frac{\pi}{2} - \alpha) \cdot \mathcal{W}^\perp$.

To finish the proof it remains to reduce the general case to the case $Bp = e_1$. Let $Q \in O(m)$ such that $QBp = e_1$, and let

$$\tilde{B} := QB \ , \quad \tilde{\Delta} := Q\Delta \ , \quad \tilde{\Delta}' := Q\Delta' \ .$$

Note that $\tilde{B} + \tilde{\Delta} = Q \cdot (B + \Delta)$ and $\tilde{B} + \tilde{\Delta}' = Q \cdot (B + \Delta')$. The claims hold if $B, \Delta, \Delta'$ are replaced by $\tilde{B}, \tilde{\Delta}, \tilde{\Delta}'$, as $\tilde{B}p = e_1$. But as

$$\operatorname{im}(A^T) = \operatorname{im}((QA)^T) , \quad \ker(A) = \ker(QA) ,$$

for every $A \in \mathbb{R}^{m\times n}$, it is readily checked that the claims also hold for $B, \Delta, \Delta'$. $\square$

**Corollary 2.1.15.** *Let $B \in \mathbb{R}_\circ^{m\times n}$ and $\mathcal{W} := \operatorname{im}(B^T)$. Furthermore, let $x \in \mathbb{R}^n \setminus \mathcal{W}^\perp$, i.e., $x \notin \ker(B)$, and let $\alpha := \sphericalangle(x, \mathcal{W})$ and $\beta := \sphericalangle(x, \mathcal{W}^\perp) = \frac{\pi}{2} - \alpha$. Then*

$$\begin{aligned} \sin\alpha &= \min\{\|\Delta\| \mid x \in \operatorname{im}(B^T + \Delta^T)\} \\ &= \min\{\|\Delta\|_F \mid x \in \operatorname{im}(B^T + \Delta^T)\} , \end{aligned}$$

$$\begin{aligned} \sin\beta &= \min\{\|\Delta'\| \mid x \in \ker(B + \Delta')\} \\ &= \min\{\|\Delta'\|_F \mid x \in \ker(B + \Delta')\} . \end{aligned}$$

*If $x \in \mathcal{W}^\perp \setminus \{0\}$, i.e., $0 \neq x \in \ker(B)$, then the statement still holds if the first two* min *are replaced by* inf*.*

*Proof.* From Proposition 2.1.12 we get

$$\inf\{\|\Delta\|_F \mid x \in \operatorname{im}(B^T + \Delta^T)\} \geq \inf\{\|\Delta\| \mid x \in \operatorname{im}(B^T + \Delta^T)\} \geq \sin\alpha ,$$
$$\inf\{\|\Delta'\|_F \mid x \in \ker(B + \Delta')\} \geq \inf\{\|\Delta'\| \mid x \in \ker(B + \Delta')\} \geq \sin\beta .$$

On the other hand, Proposition 2.1.14 implies that the above inequalities are in fact equalities, and that the minima are attained. In the case $0 \neq x \in \ker(B)$ we have $\alpha = \frac{\pi}{2}$, and $\min\{\|\Delta'\| \mid x \in \ker(B + \Delta')\} = 0 = \sin\beta$. So the second half of the stamement still holds in this case. As for the first half, i.e., the statement involving $\sin\alpha = 1$, one can use a simple perturbation argument to show the claim. The necessity to replace min by inf is easily seen by choosing $n = 2$ and $m = 1$. $\square$

In summary, we have a quite clear picture of how to perturb balanced matrices such that the defining subspaces contain some given point. We can transfer this to the unbalanced situation by using the balancing procedure and thus finish the proof of Theorem 2.1.6.

*Proof of Theorem 2.1.6.* Let $A$ have a singular value decomposition as in (2.2). In particular, $A = S \cdot A^\circ$, where $S = Q_1 \cdot \operatorname{diag}(\sigma_1, \ldots, \sigma_m) \cdot Q_1^T$ (cf. Proposition 2.1.9). As

$$(A^T + \Delta^T) \cdot v \;=\; ((A^\circ)^T S + \Delta^T) \cdot v \;=\; ((A^\circ)^T + \Delta^T S^{-1}) \cdot (Sv) ,$$

we get

$$x \in \operatorname{im}(A^T + \Delta^T) \iff x \in \operatorname{im}((A^\circ)^T + \Delta^T S^{-1}) .$$

If $\Delta$ is such that these conditions are satisfied, then Proposition 2.1.12 implies

$$\sin\alpha \;\leq\; \|\Delta^T S^{-1}\| \;\leq\; \sigma_m^{-1} \cdot \|\Delta\| .$$

In Proposition 2.1.14 we have seen that there exists a perturbation $\Delta_1$ with $x \in \operatorname{im}((A^\circ)^T + \Delta_1^T)$ such that $\|\Delta_1\|_F = \|\Delta_1\| = \sin\alpha$. Therefore, if we define $\Delta_0 := S \cdot \Delta_1$, then we have $x \in \operatorname{im}(A^T + \Delta_0^T)$, and

$$\|\Delta_0\|_F \;=\; \|S \cdot \Delta_1\|_F \;\leq\; \sigma_1 \cdot \|\Delta_1\|_F \;=\; \sigma_1 \cdot \sin\alpha .$$

The claim about the kernel follows analogously with the observation

$$(A + \Delta') \cdot x = (S \cdot A^\circ + \Delta') \cdot x = 0 \iff (A^\circ + S^{-1}\Delta') \cdot x = 0 . \qquad \square$$

## 2.2 The homogeneous convex feasibility problem

In this section we will describe the homogeneous convex feasibility problem and give the definition of Renegar's condition number.

Recall from Section 1.2 that the primal and the dual homogeneous convex feasibility problems are given by

$$\exists x \in \mathbb{R}^n \setminus \{0\}\ , \text{ s.t.} \quad Ax = 0\ ,\ x \in \breve{C}\ , \tag{P}$$

$$\exists y \in \mathbb{R}^m \setminus \{0\}\ , \text{ s.t.} \quad A^T y \in C\ , \tag{D}$$

where $A \in \mathbb{R}^{m\times n}$, $m < n$, and where the reference cone $C \subset \mathbb{R}^n$ is a *regular cone*, i.e., it is a closed convex cone such that both $C$ and its dual $\breve{C}$ have nonempty interior; the dual cone $\breve{C}$ being defined by

$$\breve{C} := \{z \in \mathbb{R}^n \mid z^T x \leq 0\ \forall x \in C\}\ .$$

Note that we have used $\breve{C}$ for the primal problem (P) and $C$ for the dual problem (D). But as $(\breve{C})\breve{} = C$ (cf. [47, Cor. 11.7.2]), this choice only has notational consequences. Recall also that a cone is called *self-dual* if $\breve{C} = -C$. Most of the cones used in convex programming are self-dual, including the cones used in linear programming (LP), second-order programming (SOCP), and semidefinite programming (SDP). See (1.1) for a list of the corresponding reference cones.

We may rephrase (P) by

$$\ker(A) \cap \breve{C} \neq \{0\}\ ,$$

and, using $\operatorname{im}(A^T) = \ker(A)^\perp$, we may rephrase (D) by

$$\operatorname{rk}(A) < m \ \text{ or } \ \operatorname{im}(A^T) \cap C \neq \{0\}\ .$$

This paraphrase of (D) already indicates that the set of rank-deficient $(m \times n)$-matrices $\mathbb{R}^{m\times n}_{\mathrm{rd}}$ will play a role in the geometric understanding of the homogeneous convex feasibility problem.

We define the sets of primal/dual feasible instances by

$$\begin{aligned}
\mathcal{F}^{\mathrm{P}}(C) &:= \{A \in \mathbb{R}^{m\times n} \mid \text{(P) is feasible}\} \\
&= \{A \mid \ker(A) \cap \breve{C} \neq \{0\}\}, \\
\mathcal{F}^{\mathrm{D}}(C) &:= \{A \in \mathbb{R}^{m\times n} \mid \text{(D) is feasible}\} \\
&= \mathbb{R}^{m\times n}_{\mathrm{rd}} \ \dot{\cup}\ \{A \mid \operatorname{rk}(A) = m \text{ and } \operatorname{im}(A^T) \cap C \neq \{0\}\}\ .
\end{aligned}$$

and accordingly the sets of primal/dual infeasible instances by

$$\begin{aligned}
\mathcal{I}^{\mathrm{P}}(C) &:= \mathbb{R}^{m\times n} \setminus \mathcal{F}^{\mathrm{P}}(C) = \{A \in \mathbb{R}^{m\times n} \mid \text{(P) is infeasible}\} \\
&= \{A \mid \ker(A) \cap \breve{C} = \{0\}\}, \\
\mathcal{I}^{\mathrm{D}}(C) &:= \mathbb{R}^{m\times n} \setminus \mathcal{F}^{\mathrm{D}}(C) = \{A \in \mathbb{R}^{m\times n} \mid \text{(D) is infeasible}\} \\
&= \{A \mid \operatorname{rk}(A) = m \text{ and } \operatorname{im}(A^T) \cap C = \{0\}\}\ .
\end{aligned}$$

To ease the notation we will occasionally simply write $\mathcal{F}^{\mathrm{P}}$, $\mathcal{F}^{\mathrm{D}}$, $\mathcal{I}^{\mathrm{P}}$, $\mathcal{I}^{\mathrm{D}}$ instead of $\mathcal{F}^{\mathrm{P}}(C)$, $\mathcal{F}^{\mathrm{D}}(C)$, $\mathcal{I}^{\mathrm{P}}(C)$, $\mathcal{I}^{\mathrm{D}}(C)$.

Note that if $\operatorname{rk}(A) = m$, then being primal/dual feasible/infeasible only depends on $\ker(A)$ respectively $\operatorname{im}(A^T) = \ker(A)^\perp$. In particular, $A$ then satisfies the same feasibility properties as its balanced approximation $A^\circ$, since they define the same subspaces.

Obviously, we have a certain asymmetry in the sets $\mathcal{F}^{\mathrm{P}}$ and $\mathcal{F}^{\mathrm{D}}$, resp. $\mathcal{I}^{\mathrm{P}}$ and $\mathcal{I}^{\mathrm{D}}$, which makes the situation a bit unsightly. Nevertheless, the boundaries of all these sets coincide and form the set of ill-posed inputs, which is the central object in the context of the condition of the convex feasibility problem. Let us formulate this in a proposition.

**Proposition 2.2.1.** *The boundaries of $\mathcal{F}^{\mathrm{P}}$, $\mathcal{F}^{\mathrm{D}}$, $\mathcal{I}^{\mathrm{P}}$, $\mathcal{I}^{\mathrm{D}}$ all coincide and are equal to $\mathcal{F}^{\mathrm{P}} \cap \mathcal{F}^{\mathrm{D}}$, i.e.*

$$\partial\mathcal{F}^{\mathrm{P}} \;=\; \partial\mathcal{F}^{\mathrm{D}} \;=\; \partial\mathcal{I}^{\mathrm{P}} \;=\; \partial\mathcal{I}^{\mathrm{D}} \;=\; \mathcal{F}^{\mathrm{P}} \cap \mathcal{F}^{\mathrm{D}} \,.$$

Before we get to the proof of this proposition, let us give the definition of the set of ill-posed inputs and of the Renegar condition.

**Definition 2.2.2.** The set of ill-posed inputs is defined by

$$\begin{aligned} \Sigma(C) \;&:=\; \mathcal{F}^{\mathrm{P}}(C) \cap \mathcal{F}^{\mathrm{D}}(C) \\ &=\; \partial\mathcal{F}^{\mathrm{P}} \;=\; \partial\mathcal{F}^{\mathrm{D}} \;=\; \partial\mathcal{I}^{\mathrm{P}} \;=\; \partial\mathcal{I}^{\mathrm{D}} \,. \end{aligned}$$

This leads us to the definition of the object we aim to understand. Recall that $d(A,B) = \|A-B\|$ for $A,B \in \mathbb{R}^{m\times n}$, and $d(A,\mathcal{M}) = \inf\{d(A,B) \mid B \in \mathcal{M}\}$, for $\mathcal{M} \subseteq \mathbb{R}^{m\times n}$.

**Definition 2.2.3.** Renegar's condition number is defined by

$$\mathscr{C}_{\mathrm{R}} \colon \mathbb{R}^{m\times n}\setminus\{0\} \to (0,\infty] \,, \quad \mathscr{C}_{\mathrm{R}}(A) := \frac{\|A\|}{d(A,\Sigma(C))} \,.$$

**Remark 2.2.4.**

1. In Section 2.3 (cf. Remark 2.3.5) we will see that $\mathscr{C}_{\mathrm{R}}(A) \geq 1$ for every $A \in \mathbb{R}^{m\times n}\setminus\{0\}$.

2. In Corollary 2.2.6 and Proposition 2.2.7 we will see that $\mathcal{F}^{\mathrm{P}} \cup \mathcal{F}^{\mathrm{D}} = \mathbb{R}^{m\times n}$, where $\mathcal{F}^{\mathrm{P}}$ and $\mathcal{F}^{\mathrm{D}}$ are closed subsets of $\mathbb{R}^{m\times n}$. Knowing this, we could have avoided the definition of $\Sigma = \mathcal{F}^{\mathrm{P}} \cap \mathcal{F}^{\mathrm{D}}$, as

$$d(A,\Sigma) = \begin{cases} d(A,\mathcal{F}^{\mathrm{D}}) & \text{if } A \in \mathcal{F}^{\mathrm{P}} \\ d(A,\mathcal{F}^{\mathrm{P}}) & \text{if } A \in \mathcal{F}^{\mathrm{D}} \end{cases} , \tag{2.9}$$

but the recognition of $\Sigma$ as the central object is of fundamental importance for the understanding of the behavior of the condition. It is for this reason that we prefer the above definition of Renegar's condition number.

3. Another characterization, which is actually the original definition, is given by

$$\mathscr{C}_{\mathrm{R}}(A)^{-1} = \max\left\{ r \,\middle|\, \|\Delta\| \leq r\cdot\|A\| \Rightarrow \begin{pmatrix} A+\Delta \in \mathcal{F}^{\mathrm{P}} & \text{if } A \in \mathcal{F}^{\mathrm{P}} \\ A+\Delta \in \mathcal{F}^{\mathrm{D}} & \text{if } A \in \mathcal{F}^{\mathrm{D}} \end{pmatrix} \right\} \,.$$

The verification of the equivalence is left to the reader.

In the remainder of this section we will give a proof of Proposition 2.2.1. The main step is the following well-known theorem of alternatives.

| | feasible | infeasible |
|---|---|---|
| (P) | $\ker(A) \cap \breve{C} \neq \{0\}$ | $\mathrm{im}(A^T) \cap \mathrm{int}(C) \neq \emptyset$ |
| (D) | $\mathrm{rk}(A) < m$ or $\mathrm{im}(A^T) \cap C \neq \{0\}$ | $\mathrm{rk}(A) = m$ and $\ker(A) \cap \mathrm{int}(\breve{C}) \neq \emptyset$ |

Table 2.1: Overview of the characterizations of $\mathcal{F}^{\mathrm{P}}, \mathcal{I}^{\mathrm{P}}, \mathcal{F}^{\mathrm{D}}, \mathcal{I}^{\mathrm{D}}$.

**Theorem 2.2.5.** *Let $C \subset \mathbb{R}^n$ be a closed convex cone with $\mathrm{int}(C) \neq \emptyset$, and let $\mathcal{W} \subseteq \mathbb{R}^n$ a linear subspace. Then*

$$\mathcal{W} \cap \mathrm{int}(C) = \emptyset \iff \mathcal{W}^\perp \cap \breve{C} \neq \{0\} . \tag{2.10}$$

*In other words,*

$$\textit{either} \quad \mathcal{W} \cap \mathrm{int}(C) \neq \emptyset \quad \textit{or} \quad \mathcal{W}^\perp \cap \breve{C} \neq \{0\} .$$

*Proof.* We first show the '$\Leftarrow$'-direction via contraposition. Let $x \in \mathcal{W} \cap \mathrm{int}(C)$ and $v \in \mathcal{W}^\perp \cap \breve{C}$, so that we need to show $v = 0$. For $\varepsilon > 0$ small enough we have $x + \varepsilon v \in C$, as $x \in \mathrm{int}(C)$. Now, we have $\langle x + \varepsilon v, v\rangle \leq 0$ as $v \in \breve{C}$, and $\langle x + \varepsilon v, v\rangle = \langle x, v\rangle + \varepsilon\langle v, v\rangle = \varepsilon\|v\|^2 \geq 0$ as $x \in \mathcal{W}$ and $v \in \mathcal{W}^\perp$. This implies $\|v\| = 0$ and thus $v = 0$.

To show the '$\Rightarrow$'-direction, let $\mathcal{W} \cap \mathrm{int}(C) = \emptyset$. Let $\Pi\colon \mathbb{R}^n \to \mathcal{W}^\perp$ denote the orthogonal projection onto $\mathcal{W}^\perp$. If we have shown that $\Pi(C) \neq \mathcal{W}^\perp$, then it follows that there exists $v \in \mathcal{W}^\perp \setminus \{0\}$ such that $\langle v, \bar{x}\rangle \leq 0$ for all $\bar{x} \in \Pi(C)$ (cf. for example [47, 11.7.3]). Since $\langle x, v\rangle = \langle \Pi(x), v\rangle \leq 0$ for all $x \in C$, it follows that $v \in \breve{C}$, and thus $\mathcal{W}^\perp \cap \breve{C} \neq \{0\}$.

It remains to show that $\Pi(C) \neq \mathcal{W}^\perp$. To do this indirectly, we assume that $\Pi(C) = \mathcal{W}^\perp$. Let $x \in \mathrm{int}(C)$, and let $y \in C$ such that $\Pi(y) = -\Pi(x)$, which exists by the assumption $\Pi(C) = \mathcal{W}^\perp$. As $x \in \mathrm{int}(C)$ and $y \in C$ it follows that $z := x + y \in \mathrm{int}(C)$. Additionally, $\Pi(z) = \Pi(x + y) = \Pi(x) + \Pi(y) = 0$, i.e., $z \in \mathcal{W}$. So we have $z \in \mathrm{int}(C) \cap \mathcal{W}$, which contradicts the assumption $\mathcal{W} \cap \mathrm{int}(C) = \emptyset$ and thus finishes the proof. □

As a direct corollary, we get a new characterization of the infeasible instances.

**Corollary 2.2.6.** *The primal/dual infeasible instances are given by*

$$\begin{aligned}\mathcal{I}^{\mathrm{P}}(C) &= \left\{A \mid \mathrm{im}(A^T) \cap \mathrm{int}(C) \neq \emptyset\right\}\\ &\subset \mathcal{F}^{\mathrm{D}}(C) ,\\[1ex] \mathcal{I}^{\mathrm{D}}(C) &= \left\{A \mid \mathrm{rk}(A) = m \textit{ and } \ker(A) \cap \mathrm{int}(\breve{C}) \neq \emptyset\right\}\\ &\subset \mathcal{F}^{\mathrm{P}}(C) .\end{aligned}$$

*In particular, we have $\mathcal{F}^{\mathrm{P}}(C) \cup \mathcal{F}^{\mathrm{D}}(C) = \mathbb{R}^{m\times n}$.*

*Proof.* This is an immediate consequence of Theorem 2.2.5. □

See Table 2.1 for an overview of the different characterizations. Another important property of $\mathcal{F}^{\mathrm{P}}(C)$ and $\mathcal{F}^{\mathrm{D}}(C)$ is that they are both closed. Let us formulate this as a proposition.

**Proposition 2.2.7.** *The sets $\mathcal{F}^{\mathrm{P}}(C)$ and $\mathcal{F}^{\mathrm{D}}(C)$ are closed subsets of $\mathbb{R}^{m\times n}$.*

*Proof.* Let $A_k \in \mathcal{F}^{\mathrm{P}}$, $k = 1, 2, \ldots$, be a sequence converging to $\lim_k A_k =: A$. As all the $A_k$ are primal feasible we may choose for every $k$ a point $p_k \in \ker(A_k) \cap \breve{C} \cap S^{n-1}$. By Bolzano-Weierstrass there exists a converging subsequence $p_{k_i}$, and we denote its limit by $\lim_i p_{k_i} =: p$. As $S^{n-1}$ and $\breve{C}$ are closed sets we have $p \in \breve{C} \cap S^{n-1}$. Furthermore, we have

$$A \cdot p = (\lim_i A_{k_i}) \cdot (\lim_i p_{k_i}) = \lim_i \underbrace{A_{k_i} \cdot p_{k_i}}_{=0} = 0 \ ,$$

which shows that $p \in \ker(A)$. So we have $A \in \mathcal{F}^{\mathrm{P}}$, and therefore $\mathcal{F}^{\mathrm{P}}$ is closed.

For the dual case we first note that we can write

$$\mathbb{R}^{m\times n}_{\mathrm{rd}} = \bigcap_{I \in \binom{[n]}{m}} \{A \in \mathbb{R}^{m\times n} \mid \det(A_I) = 0\} \ ,$$

where $[n] = \{1, \ldots, n\}$, $\binom{[n]}{m} = \{I \subseteq [n] \mid |I| = m\}$, and

$$A_I := \begin{pmatrix} a_{1,i_1} & \cdots & a_{1,i_m} \\ \vdots & & \vdots \\ a_{m,i_1} & \cdots & a_{m,i_m} \end{pmatrix} \ ,$$

for $I = \{i_1, \ldots, i_m\}$ with $i_1 < \ldots < i_m$. For every $I \in \binom{[n]}{m}$ the set $\{A \in \mathbb{R}^{m\times n} \mid \det(A_I) = 0\}$ is closed, and thus $\mathbb{R}^{m\times n}_{\mathrm{rd}}$ is closed. Let us denote the other part of $\mathcal{F}^{\mathrm{D}}$ by

$$\mathcal{F}^{\mathrm{D}}_{\circ} := \{A \mid \mathrm{rk}(A) = m \text{ and } \mathrm{im}(A^T) \cap C \neq \{0\}\} \ .$$

Let $B_k \in \mathcal{F}^{\mathrm{D}}$, $k = 1, 2, \ldots$, be a sequence converging to $\lim_k B_k =: B$. We need to show that $B \in \mathcal{F}^{\mathrm{D}}$. As $\mathbb{R}^{m\times n}_{\mathrm{rd}} \subset \mathcal{F}^{\mathrm{D}}$ we may assume that $\mathrm{rk}(B) = m$, and by omitting at most finitely many $B_k$ we may assume w.l.o.g. that $\mathrm{rk}(B_k) = m$, i.e., $B_k \in \mathcal{F}^{\mathrm{D}}_{\circ}$, for all $k$. As in the primal case we find points $q_k \in \mathrm{im}(B_k^T) \cap C \cap S^{n-1}$, and by Bolzano-Weierstrass there exists a converging subsequence $q_{k_i}$ with limit $\lim_i q_{k_i} =: q$. It remains to show that $q \in \mathrm{im}(B^T)$. Here the condition that $B$ has full rank will play a decisive role.

To ease the notation, let us replace $(q_k)$ by the converging subsequence $(q_{k_i})$, and accordingly replace the sequence $(B_k)$ by the corresponding subsequence $(B_{k_i})$. Each $q_k$ has a unique expression as

$$q_k = \sum_{j=1}^{m} r_{k,j} \cdot v_{k,j} \ , \quad r_{kj} \in \mathbb{R} \ ,$$

where $v_{k,1}, \ldots, v_{k,m}$ denote the columns of $B_k^T$. The columns of $B^T$ shall be denoted by $v_1, \ldots, v_m$, so that $\lim_k v_{k,j} = v_j$ for $j = 1, \ldots, m$, and

$$q = \lim_k q_k = \lim_k \Big( \sum_{j=1}^{m} r_{k,j} \cdot v_{k,j} \Big) = \sum_{j=1}^{m} \lim_k (r_{k,j} \cdot v_{k,j}) \ .$$

So all that remains to show is the existence of the limits $\lim_k(r_{k,j})$, $j = 1, \ldots, m$. Because then we have $\lim_k(r_{k,j} \cdot v_{k,j}) = \lim_k(r_{k,j}) \cdot \lim_k(v_{k,j})$ and it follows that $q$ lies in the image of $B^T$. Let $I \in \binom{[n]}{m}$, $I = \{i_1, \ldots, i_m\}$, $i_1 < \ldots < i_m$, such that $\det((B^T)_I) \neq 0$. Again, by omitting at most finitely many $B_k$ we may assume w.l.o.g. that $\det((B_k^T)_I) \neq 0$ for all $k$. With $M_k := (B_k^T)_I$, $\tilde{q}_k := (q_{k,i_1}, \ldots, q_{k,i_m})^T$, and $r_k := (r_{k,1}, \ldots, r_{k,m})^T$ we get

$$r_k = M_k^{-1} \cdot \tilde{q}_k \ .$$

As the limits of both factors on the right-hand side exist, so does the limit of the left-hand side, which finishes the proof. □

**Proposition 2.2.8.** *The closure of $\mathcal{I}^{\mathrm{P}}$ resp. $\mathcal{I}^{\mathrm{D}}$ is given by $\mathcal{F}^{\mathrm{D}}$ resp. $\mathcal{F}^{\mathrm{P}}$, i.e.*

$$\overline{\mathcal{I}^{\mathrm{P}}} = \mathcal{F}^{\mathrm{D}} \ , \quad \overline{\mathcal{I}^{\mathrm{D}}} = \mathcal{F}^{\mathrm{P}} \ .$$

*Proof.* As $\mathcal{F}^{\mathrm{P}}$ and $\mathcal{F}^{\mathrm{D}}$ are both closed, and as $\mathcal{I}^{\mathrm{P}} \subset \mathcal{F}^{\mathrm{D}}$ and $\mathcal{I}^{\mathrm{D}} \subset \mathcal{F}^{\mathrm{P}}$, it suffices to show the following properties.

1. For all $A \in \mathcal{F}^{\mathrm{D}}$ and for all $\varepsilon > 0$ there exists $B \in \mathcal{I}^{\mathrm{P}}$ such that $\|A - B\| \leq \varepsilon$.
2. For all $A \in \mathcal{F}^{\mathrm{P}}$ and for all $\varepsilon > 0$ there exists $B \in \mathcal{I}^{\mathrm{D}}$ such that $\|A - B\| \leq \varepsilon$.

For the first claim, let $A \in \mathcal{F}^{\mathrm{D}}$. We distinguish the cases $\mathrm{im}(A^T) \cap C \neq \{0\}$ and $\mathrm{rk}(A) < m$. If $\mathrm{im}(A^T) \cap C \neq \{0\}$, then it is geometrically clear that there exist arbitrarily small perturbations $B$ of $A$ such that $\mathrm{im}(B^T) \cap \mathrm{int}(C) \neq \emptyset$. If $\mathrm{im}(A^T) \cap C = \{0\}$ and $\mathrm{rk}(A) < m$, then this may be not so clear. So let $\mathrm{rk}(A) < m$, $\varepsilon > 0$, let $v_1, \ldots, v_m$ denote the columns of $A^T$, and w.l.o.g. we may assume $v_m \in \mathrm{lin}\{v_1, \ldots, v_{m-1}\}$. Furthermore, let $w \in \mathrm{int}(C)$, $d := \|w - v_m\| > 0$ (as $\mathrm{im}(A^T) \cap C = \{0\}$) and $\Delta v_m := d^{-1} \cdot (w - v_m)$, and let

$$v_m' := v_m + \varepsilon \cdot \Delta v_m = \Big(1 - \frac{\varepsilon}{d}\Big) \cdot v_m + \frac{\varepsilon}{d} \cdot w \ .$$

We define the perturbation $B \in \mathbb{R}^{m \times n}$ via $B^T := \begin{pmatrix} v_1 & \cdots & v_{m-1} & v_m' \end{pmatrix}$. Then we have

$$\|A - B\| = \|\varepsilon \cdot \Delta v_m\| = \varepsilon \ ,$$

and

$$w = \Big(1 - \frac{d}{\varepsilon}\Big) \cdot v_m + \frac{d}{\varepsilon} \cdot v_m' \ \in \mathrm{im}(B^T) \ ,$$

which shows that $B \in \mathcal{I}^{\mathrm{P}}$, and thus finishes the proof of the first claim.

For the second claim let $A \in \mathcal{F}^{\mathrm{P}}$, i.e., $\ker(A) \cap \breve{C} \neq \{0\}$. Recall our general assumption $m < n$, as we will need it here. If $A$ has full rank, then every small enough perturbation of $A$ has full rank, and the same geometric evidence as noted before shows that there exist arbitrarily small perturbations $B$ of $A$ such that $\mathrm{rk}(B) = m$ and $\ker(B) \cap \mathrm{int}\, \breve{C} \neq \emptyset$. Again, the case $\mathrm{rk}(A) < m$ requires further arguments. Let $v_1, \ldots, v_m$ denote the columns of $A^T$, and w.l.o.g. we may assume $v_m \in \mathrm{lin}\{v_1, \ldots, v_{m-1}\}$. Let $w \in (\ker A \cap \breve{C})$, $w \neq 0$, and let $\Delta v_m \in w^\perp \cap \ker A \cap S^{n-1}$. Denoting $v_m' := v_m + \varepsilon \cdot \Delta v_m$ and $B^T := \begin{pmatrix} v_1 & \cdots & v_{m-1} & v_m' \end{pmatrix}$, we get

$$\|A - B\| = \|\varepsilon \cdot \Delta v_m\| = \varepsilon \ ,$$

$\ker B = \ker A \cap (\Delta v_m)^\perp$. In particular $w \in \ker B$. So we have $B \in \mathcal{F}^{\mathrm{P}}$ and $rk(B) = rk(A) + 1$. Repeating this argument if necessary we arrive at the full-rank case, which we have already discussed. □

*Proof of Proposition 2.2.1.* Follows from Proposition 2.2.7 and Proposition 2.2.8 by the fact that $\partial \mathcal{M} = \overline{\mathcal{M}} \cap \overline{\mathcal{M}^c}$ for every subset $\mathcal{M}$ of a topological space, where $\mathcal{M}^c$ denotes the complement of $\mathcal{M}$. □

### 2.2.1 Intrinsic vs. extrinsic condition

In the last section we have explicitly computed several characterizations of the sets of primal/dual feasible/infeasible instances, and in particular of the set of ill-posed inputs. A quick look at these sets already reveals, that a part of the set of rank-deficient matrices is contained in $\Sigma$. In this section we will elaborate on this interference by the matrix condition, in particular by considering some concrete low-dimensional examples.

Broadly speaking, if a matrix $A$ has a large condition $\mathscr{C}_{\mathrm{R}}(A)$ then this may have two reasons:

1. The subspace defined by $A$ may intersect/miss the reference cone close to the boundary, or

2. the linear operator defined by $A$ may itself be badly conditioned.

Moreover, these two effects interfere and thus make a direct analysis of Renegar's condition hard. To emphasize the different natures of these effects, we call the condition of the operator $A$ the *extrinsic condition* and the condition indicated in 1. the *intrinsic condition*. In the next section we will specify the notion of intrinsic condition by introducing a new condition number. Note that it should be plausible that one can get rid of the extrinsic condition by using a preconditioner that repairs the bad condition of the linear operator. The intrinsic condition on the other hand is truly at the heart of the conditioning problem and thus captures the essential part of the Renegar condition.

The inverse of the matrix condition $\kappa(A)^{-1}$ has a nice geometric interpretation: It is given by the maximal radius of a closed ball around the origin that lies completely in the image of the unit ball in $\mathbb{R}^n$ under the normalized operator $A^{(1)} := \|A\|^{-1} \cdot A$, i.e.,

$$\kappa(A)^{-1} = \max\{r \mid r \cdot B_m \subseteq A^{(1)}(B_n)\} ,$$

where $B_m$ denotes the closed unit ball in $\mathbb{R}^m$ (cf. Proposition 2.1.3). A beautiful characteristic of Renegar's condition number is that in the case $\ker(A) \cap \breve{C} \neq \{0\}$, i.e., in the primal feasible case, it has a similar description. The only difference is that the unit ball $B_n$ is replaced by the intersection of the unit ball with (the dual of) the reference cone $\breve{C}$, i.e., we have

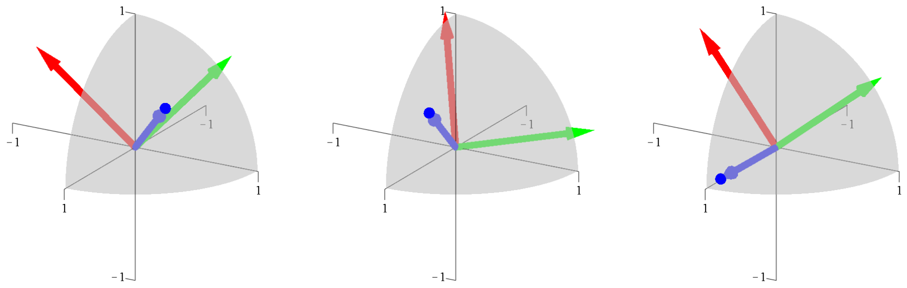
$$\mathscr{C}_{\mathrm{R}}(A)^{-1} = \max\{r \mid r \cdot B_m \subseteq A^{(1)}(B_n \cap \breve{C})\} . \tag{2.11}$$

(See [45] or [42, Cor. 3.6].) We will use this characterization in the following example.

Let us anticipate the definition of the intrinsic condition, that we will give in the next section, and define

$$\mathscr{C}_{\mathrm{G}}(A) \;:=\; \mathscr{C}_{\mathrm{R}}(A^{\circ}) ,$$

where $A^{\circ}$ denotes the balanced approximation of $A$ as defined in (2.6). To get a picture of these quantities let us specialize to the case $n = 3$, $m = 2$, $\breve{C} = \mathbb{R}^3_+$, and $A$ being one of three primal feasible matrices $A_1, A_2, A_3$, that are specified in



(a) $A_1$ (b) $A_2$ (c) $A_3$

Figure 2.2: The singular vectors of $A_1, A_2, A_3$ and the positive orthant.

the following list along with their different conditions.

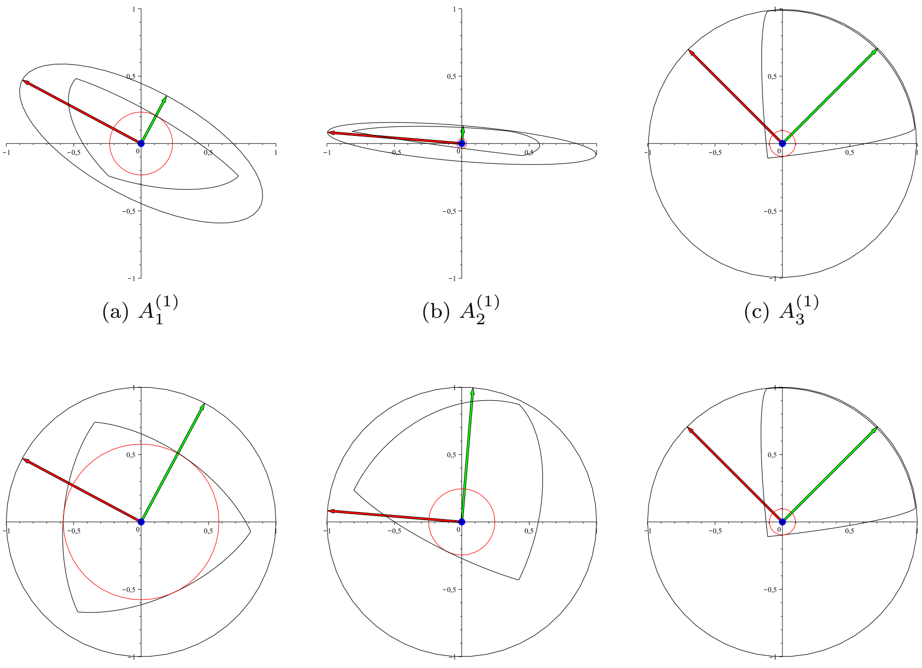
| | $\kappa$ | $\mathscr{C}_\mathrm{G}$ | $\mathscr{C}_\mathrm{R}$ | $\kappa \cdot \mathscr{C}_\mathrm{G}$ |
|---|---|---|---|---|
| $A_1 := \begin{pmatrix} -1 & 3 & -2 \\ -1 & -1 & 2 \end{pmatrix}$ | 2.5 | 1.7 | 4.3 | 4.3 |
| $A_2 := \begin{pmatrix} 5 & 4 & -9 \\ -1 & 1 & 1 \end{pmatrix}$ | 7.5 | 4.1 | 28.0 | 31 |
| $A_3 := \begin{pmatrix} -1 & 9 & -1 \\ -1 & 1 & 9 \end{pmatrix}$ | 1.0 | 10.4 | 10.4 | 10.4 |

The locations of the singular vectors of $A_1, A_2, A_3$ with respect to the positive orthant are shown in Figure 2.2. In these pictures the blue arrow spans the kernel of the respective operator. Figure 2.3 shows the images of the singular vectors of the normalized operators $A_1^{(1)}, A_2^{(1)}, A_3^{(1)}$ as well as the images of the unit sphere and the positive orthant intersected with the unit sphere. It also shows the images of these objects under the balanced operators $A_1^\circ, A_2^\circ, A_3^\circ$.

The first matrix $A_1$ has a kernel that hits the reference cone exactly in the center, i.e., the intrinsic condition is best possible. The Renegar condition on the other hand is not as good as it could be due to the non-optimality of the matrix condition. The second matrix has a slightly worse intrinsic condition as the solution set lies closer to the boundary of the reference cone. But the extrinsic/matrix condition is very bad so that the Renegar condition is totally dominated by this effect. The third matrix is nearly balanced but its intrinsic condition is the worst compared to the other two. Finally, note that the product $\kappa(A) \cdot \mathscr{C}_\mathrm{G}(A)$ gives an upper bound for $\mathscr{C}_\mathrm{R}(A)$. This holds in general as we will see in the next section.

## 2.3 Defining the Grassmann condition

In this section we will define the Grassmann condition and give several equivalent characterizations corresponding to the three viewpoints on the homogeneous convex feasibility problem that we described in Section 1.5. We will also establish the relationship between the Renegar and the Grassmann condition that we already announced in Section 1.2.



(a) $A_1^{(1)}$ (b) $A_2^{(1)}$ (c) $A_3^{(1)}$

(d) $A_1^\circ$ (e) $A_2^\circ$ (f) $A_3^\circ$

Figure 2.3: The images under $A_1^{(1)}, A_2^{(1)}, A_3^{(1)}$ and $A_1^\circ, A_2^\circ, A_3^\circ$.

**Definition 2.3.1.** Let $C \subset \mathbb{R}^n$ be a regular cone, and let $1 \leq m \leq n-1$. The *Grassmann condition* is defined by

$$\mathscr{C}_{\mathrm{G}} \colon \mathbb{R}^{m\times n} \setminus \{0\} \to (0,\infty]\,, \; \mathscr{C}_{\mathrm{G}}(A) := \begin{cases} \mathscr{C}_{\mathrm{R}}(A^\circ) & \text{if } \mathrm{rk}(A) = m \\ 1 & \text{if } \mathrm{rk}(A) < m\,, \end{cases}$$

where $A^\circ$ denotes the balanced approximation of $A$ (cf. Proposition 2.1.8/2.1.9).

Before we state the next proposition recall that for $A \in \mathbb{R}_*^{m\times n}$, i.e., $\mathrm{rk}(A) = m$, and for $\mathcal{W} := \mathrm{im}(A^T)$, we have

$$\begin{aligned} A \in \mathcal{F}^{\mathrm{P}} &\iff \mathcal{W}^\perp \cap \breve{C} \neq \{0\} \iff \mathcal{W} \cap \mathrm{int}(C) = \emptyset \\ A \in \mathcal{I}^{\mathrm{P}} &\iff \mathcal{W}^\perp \cap \breve{C} = \{0\} \iff \mathcal{W} \cap \mathrm{int}(C) \neq \emptyset \\ A \in \mathcal{F}^{\mathrm{D}} &\iff \mathcal{W} \cap C \neq \{0\} \iff \mathcal{W}^\perp \cap \mathrm{int}(\breve{C}) = \emptyset \\ A \in \mathcal{I}^{\mathrm{D}} &\iff \mathcal{W} \cap C = \{0\} \iff \mathcal{W}^\perp \cap \mathrm{int}(\breve{C}) \neq \emptyset\,, \end{aligned} \tag{2.12}$$

and $\Sigma = \mathcal{F}^{\mathrm{P}} \cap \mathcal{F}^{\mathrm{D}}$.

**Proposition 2.3.2.** *Let $C \subset \mathbb{R}^n$ be a regular cone, $1 \leq m \leq n-1$, and let $A \in \mathbb{R}_*^{m\times n}$ and $\mathcal{W} := \mathrm{im}(A^T)$.*

1. *Denoting by $\sphericalangle(x,\mathcal{W})$ the angle between $x \in \mathbb{R}^n \setminus \{0\}$ and $\mathcal{W}$ (cf. (2.4)), let*

$$\begin{aligned} \alpha &:= \min\{\sphericalangle(x,\mathcal{W}) \mid x \in C \setminus \{0\}\} \quad , \text{ if } A \in \mathcal{F}^{\mathrm{P}}\,, \\ \beta &:= \min\left\{\sphericalangle(v,\mathcal{W}^\perp) \mid v \in \breve{C} \setminus \{0\}\right\}, \text{ if } A \in \mathcal{F}^{\mathrm{D}}\,. \end{aligned} \tag{2.13}$$

*Then*

$$\mathscr{C}_{\mathrm{G}}(A) = \begin{cases} \frac{1}{\sin\alpha} & \text{if } A \in \mathcal{F}^{\mathrm{P}} , \\ \frac{1}{\sin\beta} & \text{if } A \in \mathcal{F}^{\mathrm{D}} . \end{cases} \tag{2.14}$$

2. *Let* $\Pi_{\mathcal{W}}, \Pi_{\mathcal{W}^\perp}$ *denote the orthogonal projections onto* $\mathcal{W}$ *resp.* $\mathcal{W}^\perp$*, and let*

$$r_{\max} := \begin{cases} \max\{r \mid r\cdot B_n \cap \mathcal{W} \subseteq \Pi_{\mathcal{W}}(\breve{C} \cap B_n)\} & \text{if } A \in \mathcal{F}^{\mathrm{P}} , \\ \max\{r \mid r\cdot B_n \cap \mathcal{W}^\perp \subseteq \Pi_{\mathcal{W}^\perp}(C \cap B_n)\}\} & \text{if } A \in \mathcal{F}^{\mathrm{D}} , \end{cases}$$

*where* $B_n$ *denotes the unit ball in* $\mathbb{R}^n$*. Then*

$$\mathscr{C}_{\mathrm{G}}(A) = \frac{1}{r_{\max}} .$$

Note that from the above proposition it follows that $\mathscr{C}_{\mathrm{G}}(A)$ in fact only depends on the subspace $\mathcal{W} = \operatorname{im}(A^T)$ resp. $\mathcal{W}^\perp = \ker(A)$.

**Remark 2.3.3.** The Grassmann condition as characterized in (2.14) was considered in [5] in the dual feasible case. Also the inequalities in Theorem 2.3.4 relating the Grassmann to the Renegar condition was given there.

*Proof of Proposition 2.3.2.* (1) By definition of Renegar's condition number, we have

$$\mathscr{C}_{\mathrm{G}}(A)^{-1} = \mathscr{C}_{\mathrm{R}}(A^\circ)^{-1} = d(A^\circ, \Sigma(C)) ,$$

as $\|A^\circ\| = 1$. If $A \in \mathcal{F}^{\mathrm{P}}$, then $d(A^\circ, \Sigma) = d(A^\circ, \mathcal{F}^{\mathrm{D}})$ (cf. (2.9) in Remark 2.2.4), and thus

$$\begin{aligned} d(A^\circ, \Sigma) &= \min\left\{\|\Delta\| \mid \operatorname{im}\left((A^\circ)^T + \Delta^T\right) \cap C \neq \{0\} \text{ or } A^\circ + \Delta \in \mathbb{R}^{m\times n}_{\mathrm{rd}}\right\} \\ &= \min\left\{ \min\left\{\|\Delta\| \mid \operatorname{im}\left((A^\circ)^T + \Delta^T\right) \cap C \neq \{0\}\right\} ,\ d(A^\circ, \mathbb{R}^{m\times n}_{\mathrm{rd}}) \right\} . \end{aligned}$$

Note that $d(A^\circ, \mathbb{R}^{m\times n}_{\mathrm{rd}}) = 1$, as $A^\circ \in \mathbb{R}^{m\times n}_\circ$ (cf. Theorem 2.1.5). By Corollary 2.1.15 we have for $x \in \mathbb{R}^n \setminus \mathcal{W}^\perp$ and $\rho := \sphericalangle(x, \mathcal{W})$

$$\min\left\{\|\Delta\| \mid x \in \operatorname{im}\left((A^\circ)^T + \Delta^T\right)\right\} \;=\; \sin\rho ,$$

which also holds for $x \in \mathcal{W}^\perp \setminus \{0\}$, if min is replaced by inf. As $\operatorname{int}(C) \neq \emptyset$ we have $C \not\subseteq \mathcal{W}^\perp$, which implies $\alpha < \frac{\pi}{2}$. So if $x \in C \setminus \{0\}$ is such that the angle $\sphericalangle(x, \mathcal{W})$ is minimal, i.e., $\sphericalangle(x, \mathcal{W}) = \alpha$, then $x \notin \mathcal{W}^\perp$ and there exists $\Delta_0 \in \mathbb{R}^{m\times n}$ such that $x \in \operatorname{im}\left((A^\circ)^T + \Delta_0^T\right)$. On the other hand, any perturbation $\Delta$ such that $\operatorname{im}\left((A^\circ)^T + \Delta^T\right) \cap C \neq \{0\}$ must have a norm of at least $\sin\alpha$, so that we may conclude

$$\min\left\{\|\Delta\| \mid \operatorname{im}\left((A^\circ)^T + \Delta^T\right) \cap C \neq \{0\}\right\} \;=\; \sin\alpha .$$

Altogether, we get

$$d(A^\circ, \Sigma) = \min\{\sin\alpha, 1\} = \sin\alpha .$$

This shows (2.14) in the case $A \in \mathcal{F}^{\mathrm{P}}$.

In the case $A \in \mathcal{F}^{\mathrm{D}}$ we may argue analogously using $d(A^\circ, \Sigma) = d(A^\circ, \mathcal{F}^{\mathrm{P}})$, which implies by Corollary 2.1.15

$$d(A^\circ, \Sigma) = \min\left\{\|\Delta'\| \mid \ker\left(A^\circ + \Delta'\right) \cap \breve{C} \neq \{0\}\right\} = \sin\beta .$$

(2) It remains to show $\sin\alpha = r_{\max}$ resp. $\sin\beta = r_{\max}$. By duality, we may assume w.l.o.g. that $A \in \mathcal{F}^{\mathrm{P}}$. From the characterization of the Renegar condition stated in (2.11) we get

$$\begin{aligned}\sin\alpha = \mathscr{C}_{\mathrm{G}}(A)^{-1} = \mathscr{C}_{\mathrm{R}}(A^\circ)^{-1} & \\ &\overset{(2.11)}{=} \max\{r \mid r\cdot B_m \subseteq A^\circ(B_n \cap \breve{C})\} \\ &= \max\left\{r \mid r\cdot (A^\circ)^T(B_m) \subseteq (A^\circ)^T A^\circ(B_n\cap\breve{C})\right\} ,\end{aligned}$$

where the last equality follows from the fact that the map $(A^\circ)^T\colon \mathbb{R}^m \to \mathbb{R}^n$ is a norm-preserving map. Using Lemma 2.1.11 we may continue

$$\begin{aligned}&= \max\left\{r \mid r\cdot B_n \cap \mathcal{W} \subseteq \Pi_{\mathcal{W}}(B_n\cap\breve{C})\right\} \\ &= r_{\max} .\end{aligned}$$

This finishes the proof. □

**Theorem 2.3.4.** *Let $A \in \mathbb{R}^{m\times n}\setminus\{0\}$. Then*

$$\mathscr{C}_{\mathrm{G}}(A) \;\le\; \mathscr{C}_{\mathrm{R}}(A) \;\le\; \kappa(A)\cdot\mathscr{C}_{\mathrm{G}}(A) . \tag{2.15}$$

**Remark 2.3.5.** Note that as a simple corollary we get $\mathscr{C}_{\mathrm{R}}(A) \ge 1$ for all $A \in \mathbb{R}^{m\times n}\setminus\{0\}$, as $\mathscr{C}_{\mathrm{R}}(A) \ge \mathscr{C}_{\mathrm{G}}(A)$ and $\mathscr{C}_{\mathrm{G}}(A) \ge 1$ by the characterization (2.14).

*Proof of Theorem 2.3.4.* If $\mathrm{rk}(A) < m$, then $\mathscr{C}_{\mathrm{G}}(A) = 1$ and $\kappa(A) = \infty$, and the claim holds trivially. In the following we may thus assume that $\mathrm{rk}(A) = m$.

If $A \in \mathcal{F}^{\mathrm{P}}$, then $d(A,\Sigma) = d(A,\mathcal{F}^{\mathrm{D}})$ (cf. Remark 2.2.4). From Theorem 2.1.6 (1a,b) we get as in the proof of Proposition 2.3.2

$$\sigma_m\cdot\sin\alpha \;\le\; d(A,\Sigma) \;\le\; \sigma_1\cdot\sin\alpha ,$$

where $\sigma_1 \ge \ldots \ge \sigma_m$ denote the singular values of $A$, and $\alpha$ denotes the minimum angle between $C$ and $\mathrm{im}(A^T)$ as defined in (2.13). By Proposition 2.3.2 we have $\mathscr{C}_{\mathrm{G}}(A) = \frac{1}{\sin\alpha}$, and so we get

$$\begin{aligned}\mathscr{C}_{\mathrm{R}}(A) &= \frac{\sigma_1}{d(A,\Sigma)} \le \frac{\sigma_1}{\sigma_m}\cdot\frac{1}{\sin\alpha} = \kappa(A)\cdot\mathscr{C}_{\mathrm{G}}(A) , \\ \mathscr{C}_{\mathrm{R}}(A) &= \frac{\sigma_1}{d(A,\Sigma)} \ge \frac{1}{\sin\alpha} = \mathscr{C}_{\mathrm{G}}(A) .\end{aligned}$$

The case $A \in \mathcal{F}^{\mathrm{D}}$ follows analogously. □

In the remainder of this section we will describe another characterization of the Grassmann condition corresponding to the third viewpoint on the homogeneous convex feasibility problem. This characterization is also the reason for the name 'Grassmann condition'.

For $0 \le m \le n$ the $(n,m)$th *Grassmann manifold* $\mathrm{Gr}_{n,m}$ is defined as the set of $m$-dimensional linear subspaces of $\mathbb{R}^n$

$$\mathrm{Gr}_{n,m} := \{\mathcal{W} \subseteq \mathbb{R}^n \mid \mathcal{W} \text{ lin. subspace}, \, \dim(\mathcal{W}) = m\} .$$

Note that by intersecting each element in $\mathrm{Gr}_{n,m}$ with the unit sphere $S^{n-1}$, we can identify the Grassmann manifold $\mathrm{Gr}_{n,m}$ with the set of $(m-1)$-dimensional subspheres of $S^{n-1}$. In Section 3.2 we will see that this set is endowed with a metric $d_{\mathrm{H}}$, the Hausdorff metric, and forms a complete metric space. We use the same symbol $d_{\mathrm{H}}$ to denote the corresponding metric in $\mathrm{Gr}_{n,m}$.

Another metric on $\mathrm{Gr}_{n,m}$ is given by the geodesic metric, which we denote by $d_{\mathrm{g}}$. This metric comes from the fact that $\mathrm{Gr}_{n,m}$ is a complete Riemannian manifold. In Section 5.3.2 and Section 5.4 we will have a closer look at this. In particular, we will prove the following well-known facts:

- The Hausdorff and the geodesic metric on $\mathrm{Gr}_{n,m}$ are equivalent. In particular, we only have one topology on $\mathrm{Gr}_{n,m}$ that we work with.

- We have $d_{\mathrm{H}}(\mathcal{W}_1,\mathcal{W}_2) \leq \frac{\pi}{2}$ for all $\mathcal{W}_1,\mathcal{W}_2 \in \mathrm{Gr}_{n,m}$.

- The map
$$\mathrm{Gr}_{n,m} \to \mathrm{Gr}_{n,n-m}\ , \quad \mathcal{W} \mapsto \mathcal{W}^{\perp}$$
is an isometry with respect to the Hausdorff metric on $\mathrm{Gr}_{n,m}$ and on $\mathrm{Gr}_{n,n-m}$, and with respect to the geodesic metric on $\mathrm{Gr}_{n,m}$ and $\mathrm{Gr}_{n,n-m}$.

- The surjective maps
$$\begin{aligned} I\colon \mathbb{R}_*^{m\times n} &\to \mathrm{Gr}_{n,m}\ , \quad A \mapsto \operatorname{im} A^T\ , \\ K\colon \mathbb{R}_*^{(n-m)\times n} &\to \mathrm{Gr}_{n,m}\ , \quad A' \mapsto \ker A'\ . \end{aligned}$$
are continuous, open, and closed, with respect to the relative topology on $\mathbb{R}_*^{m\times n}$ respectively $\mathbb{R}_*^{(n-m)\times n}$ (cf. Remark 5.3.6).

  In this section we may use this fact as the definition of the topology on $\mathrm{Gr}_{n,m}$, i.e., the topology on $\mathrm{Gr}_{n,m}$ is given as as the pushforward of the topology on $\mathbb{R}_*^{m\times n}$ resp. $\mathbb{R}_*^{(n-m)\times n}$ via $I$ resp. $K$.

We define the *primal* and *dual Grassmann feasibility problems* by
$$\mathcal{W}^{\perp} \cap \breve{C} \neq \{0\}\ , \tag{GrP}$$
$$\mathcal{W} \cap C \neq \{0\}\ , \tag{GrD}$$
where $\mathcal{W} \in \mathrm{Gr}_{n,m}$, $1 \leq m \leq n-1$, and where the reference cone $C \subset \mathbb{R}^n$ is a regular cone. Notice the full duality of these problems, i.e., there is no structural difference between (GrP) and (GrD). Additionally, the connection between the primal and the dual is given by the isometry $\mathcal{W} \mapsto \mathcal{W}^{\perp}$ (and the local isometry $C \mapsto \breve{C}$; cf. Section 3.2).

We define the sets of primal/dual feasible/infeasible instances, with respect to (GrP) and (GrD), by
$$\begin{aligned} \mathcal{F}_{\mathrm{G}}^{\mathrm{P}} &:= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \text{(GrP) is feasible}\}\ , \\ \mathcal{I}_{\mathrm{G}}^{\mathrm{P}} &:= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \text{(GrP) is infeasible}\}\ , \\ \mathcal{F}_{\mathrm{G}}^{\mathrm{D}} &:= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \text{(GrD) is feasible}\}\ , \\ \mathcal{I}_{\mathrm{G}}^{\mathrm{D}} &:= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \text{(GrD) is infeasible}\}\ . \end{aligned}$$

The relations between these sets are simpler than the relations between $\mathcal{F}^{\mathrm{P}}$, $\mathcal{I}^{\mathrm{P}}$, $\mathcal{F}^{\mathrm{D}}$, and $\mathcal{I}^{\mathrm{D}}$. We describe them in the following proposition.

**Proposition 2.3.6.** *Let $C \subset \mathbb{R}^n$ a regular cone, and let $1 \le m \le n-1$. Then the sets $\mathcal{F}_{\text{G}}^{\text{P}}$ and $\mathcal{F}_{\text{G}}^{\text{D}}$ are closed, $\mathcal{I}_{\text{G}}^{\text{P}}$ and $\mathcal{I}_{\text{G}}^{\text{D}}$ are open, and*

$$\mathcal{F}_{\text{G}}^{\text{P}} = \overline{\mathcal{I}_{\text{G}}^{\text{D}}} \ , \quad \mathcal{F}_{\text{G}}^{\text{D}} = \overline{\mathcal{I}_{\text{G}}^{\text{P}}} \ , \quad \mathcal{I}_{\text{G}}^{\text{P}} = \operatorname{int}(\mathcal{F}_{\text{G}}^{\text{D}}) \ , \quad \mathcal{I}_{\text{G}}^{\text{D}} = \operatorname{int}(\mathcal{F}_{\text{G}}^{\text{P}}) \ . \tag{2.16}$$

*Furthermore, the intersection $\mathcal{F}_{\text{G}}^{\text{P}} \cap \mathcal{F}_{\text{G}}^{\text{D}}$ consists of those $m$-dimensional subspaces, which touch the cone $C$ in the boundary $\partial C$*

$$\mathcal{F}_{\text{G}}^{\text{P}} \cap \mathcal{F}_{\text{G}}^{\text{D}} = \{\mathcal{W} \in \operatorname{Gr}_{n,m} \mid \mathcal{W} \cap \operatorname{int}(C) = \emptyset \text{ and } \mathcal{W} \cap \partial C \neq \{0\}\} \ .$$

*Proof.* If we intersect the sets $\mathcal{F}^{\text{P}}$, $\mathcal{I}^{\text{P}}$, $\mathcal{F}^{\text{D}}$, and $\mathcal{I}^{\text{D}}$ with the set of full-rank $(m \times n)$-matrices, then we get a characterization as stated in (2.12). From this characterization it is seen that the sets $\mathcal{F}_{\text{G}}^{\text{P}}$, $\mathcal{F}_{\text{G}}^{\text{D}}$, $\mathcal{I}_{\text{G}}^{\text{P}}$, $\mathcal{I}_{\text{G}}^{\text{D}}$, are the images of the above intersections under the map $I$, i.e.,

$$\begin{aligned} \mathcal{F}_{\text{G}}^{\text{P}} &= I(\mathcal{F}^{\text{P}} \cap \mathbb{R}_*^{m\times n}) \ , & \mathcal{F}_{\text{G}}^{\text{D}} &= I(\mathcal{F}^{\text{P}} \cap \mathbb{R}_*^{m\times n}) \ , \\ \mathcal{I}_{\text{G}}^{\text{P}} &= I(\mathcal{I}^{\text{P}} \cap \mathbb{R}_*^{m\times n}) \ , & \mathcal{I}_{\text{G}}^{\text{D}} &= I(\mathcal{I}^{\text{P}} \cap \mathbb{R}_*^{m\times n}) \ . \end{aligned}$$

We have seen in Section 2.2 that $\mathcal{F}^{\text{P}}$ and $\mathcal{F}^{\text{D}}$ are closed, and $\mathcal{I}^{\text{P}}$ and $\mathcal{I}^{\text{D}}$ are open. By definition of the relative topology, the same holds for the intersections with $\mathbb{R}_*^{m\times n}$. As the map $I$ is both closed and open, we also have that $\mathcal{F}_{\text{G}}^{\text{P}}$ and $\mathcal{F}_{\text{G}}^{\text{D}}$ are closed, and $\mathcal{I}_{\text{G}}^{\text{P}}$ and $\mathcal{I}_{\text{G}}^{\text{D}}$ are open. Analogously, we have the relations stated in (2.16).

Concerning the statement about the intersection $\mathcal{F}_{\text{G}}^{\text{P}} \cap \mathcal{F}_{\text{G}}^{\text{D}}$, we have for $A \in \mathbb{R}_*^{m\times n}$, $\mathcal{W} = I(A) = \operatorname{im}(A^T)$,

$$\begin{aligned} A \in \Sigma = \mathcal{F}^{\text{P}} \cap \mathcal{F}^{\text{D}} \quad &\overset{(2.16)}{\iff} \quad \mathcal{W} \cap \operatorname{int}(C) = \emptyset \text{ and } \mathcal{W} \cap C \neq \{0\} \\ &\iff \quad \mathcal{W} \cap \operatorname{int}(C) = \emptyset \text{ and } \mathcal{W} \cap \partial C \neq \{0\} \ , \end{aligned}$$

which finishes the proof. □

**Definition 2.3.7.** For $C \subset \mathbb{R}^n$ a regular cone, and $1 \le m \le n-1$ we define

$$\begin{aligned} \Sigma_m(C) &:= \mathcal{F}_{\text{G}}^{\text{P}} \cap \mathcal{F}_{\text{G}}^{\text{D}} \\ &= \{\mathcal{W} \in \operatorname{Gr}_{n,m} \mid \mathcal{W} \cap \operatorname{int}(C) = \emptyset \text{ and } \mathcal{W} \cap \partial C \neq \{0\}\} \ . \end{aligned}$$

The following proposition provides the namesake characterization of the Grassmann condition.

**Proposition 2.3.8.** *Let $C \subset \mathbb{R}^n$ a regular cone, $1 \le m \le n-1$, and let $A \in \mathbb{R}_*^{m\times n}$ and $\mathcal{W} := \operatorname{im}(A^T) \in \operatorname{Gr}_{n,m}$. Then*

$$\mathscr{C}_{\text{G}}(A) = \frac{1}{\sin d_{\text{H}}(\mathcal{W}, \Sigma_m(C))} = \frac{1}{\sin d_{\text{g}}(\mathcal{W}, \Sigma_m(C))} \ ,$$

*where $d_{\text{H}}$ denotes the Hausdorff distance and $d_{\text{g}}$ denotes the geodesic distance in $\operatorname{Gr}_{n,m}$.*

*Proof.* We will show this in Chapter 5 (cf. Corollary 5.5.3). □

# Chapter 3

# Spherical convex geometry

This chapter is devoted to aspects of spherical convex geometry, that we will need for the analysis of the Grassmann condition. The goal of this chapter is to set up the notation for the treatment of convex sets in the sphere, and to describe a broad picture of the characteristics of spherical convex geometry. We include this chapter as we could not find a single source on spherical geometry that covered all the material we need. All what is stated in this chapter is well-known, but for the sake of completeness we provide proofs for most of the statements.

## 3.1 Some basic definitions

From now on we will denote the euclidean distance in $\mathbb{R}^n$ by

$$d^e\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\ , \quad d^e(x,y) = \|x-y\|\ ,$$

and we will denote by

$$d\colon S^{n-1} \times S^{n-1} \to \mathbb{R}\ , \quad d(p,q) = \arccos(\langle p,q\rangle)$$

the spherical distance, i.e., the angle between the points $p$ and $q$. Recall that a set $K^e \subseteq \mathbb{R}^n$ is called convex iff for all $x, y \in K^e$ the line segment between $x$ and $y$ is contained in $K^e$. Analogously, a subset $K \subseteq S^{n-1}$ is called convex iff for all $p, q \in K$ with $q \neq \pm p$ the (unique geodesic) arc between $p$ and $q$, which we denote by $\operatorname{geod}(p,q)$, is contained in $K$. This is equivalent to the condition that the set

$$\operatorname{cone}(K) := \{\lambda \cdot p \mid \lambda \geq 0\, ,\ p \in K\} \subseteq \mathbb{R}^n$$

is convex.

**Definition 3.1.1.** The set of nonempty compact convex sets in euclidean space

$$\mathcal{K}(\mathbb{R}^n) \ :=\ \{K^e \subseteq \mathbb{R}^n \mid K^e \text{ nonempty compact convex}\}$$

is called the set of *convex bodies* in $\mathbb{R}^n$. In the unit sphere we call a set *spherical convex*, if it is closed, convex, and neither the empty set nor the whole sphere

$$\mathcal{K}(S^{n-1}) \ :=\ \{K \subseteq S^{n-1} \mid K \text{ closed and convex, } K \neq \emptyset, \text{ and } K \neq S^{n-1}\}\ .$$

(We exclude the empty set and the whole sphere to avoid pointless case distinctions.)

One of the most important aspects in spherical convex geometry is *duality*. Recall that if $C \subseteq \mathbb{R}^n$ is a convex cone, then its dual is defined via

$$\breve{C} = \{v \in \mathbb{R}^n \mid \langle v, x\rangle \leq 0 \; \forall x \in C\} \ .$$

For $K \in \mathcal{K}(S^{n-1})$, and $C := \operatorname{cone}(K)$, we may therefore define

$$\breve{K} := \breve{C} \cap S^{n-1} \ .$$

We call a cone/spherical convex set *self-dual*, if $\breve{C} = -C$ resp. $\breve{K} = -K$.

**Remark 3.1.2.**
1. The duality map is an involution on the set of spherical convex sets, i.e., $(\breve{K})\breve{} = K$ for all $K \in \mathcal{K}(S^{n-1})$.
2. We may give an intrinsic characterization of the dual via $\breve{K} = \{v \in S^{n-1} \mid d(K, v) \geq \frac{\pi}{2}\}$, where $d(K, v) := \min\{d(p, v) \mid p \in K\}$.
3. The boundary of $\breve{K}$ is given by $\partial\breve{K} = \{v \in S^{n-1} \mid d(K, v) = \frac{\pi}{2}\}$.

In euclidean space every closed convex set admits a (global) projection map, i.e., a map which sends a point $x \in \mathbb{R}^n$ to the uniquely determined point in the convex set, which minimizes the distance to $x$. However, in the sphere a projection map onto a spherical convex set $K \in \mathcal{K}(S^{n-1})$ is in general only defined outside its dual $\breve{K}$. We will see this after we have defined the normal cone and determined the normal cones of a closed convex cone.

**Definition 3.1.3.** Let $K^e \subseteq \mathbb{R}^n$ be a closed convex set (not necessarily compact), and let $\Pi_{K^e}$ denote the projection map onto $K^e$. The *normal cone* of $K^e$ at $x \in K^e$ is defined as

$$N_x(K^e) := \{v \in \mathbb{R}^n \mid \Pi_{K^e}(x + v) = x\} \ .$$

We furthermore define $N_x^S(K^e) := N_x(K^e) \cap S^{n-1}$. For a spherical convex set $K \in \mathcal{K}(S^{n-1})$ and $C := \operatorname{cone}(K)$ we define for $p \in K$

$$N_p(K) := N_p(C) \quad \text{and} \quad N_p^S(K) := N_p(K) \cap S^{n-1} \ .$$

**Remark 3.1.4.**
1. As the term 'normal cone' suggests, $N_x(K^e)$ is indeed a closed convex cone.
2. From the definition of the projection map it is readily seen that $N_x(K^e) = \{0\}$ iff $x \in \operatorname{int}(K^e)$, or equivalently $N_x(K^e) \neq \{0\}$ iff $x \in \partial K^e$. In other words, the normal cone is nontrivial only on the boundary of the convex set.
3. From the fact that $\Pi_{K^e}$ is a (global) map it follows that $\mathbb{R}^n$ has a disjoint decomposition into $\mathbb{R}^n = \dot{\bigcup}_{x \in K^e}(x + N_x(K^e))$.

**Proposition 3.1.5.** *Let $C \subseteq \mathbb{R}^n$ be a closed convex cone, and let $x \in C$. Then*

$$N_x(C) = x^{\perp} \cap \breve{C} \ , \tag{3.1}$$

*where $x^{\perp} = \{y \in \mathbb{R}^n \mid \langle x, y\rangle = 0\}$. In particular, $N_0(C) = \breve{C}$.*

*Proof.* Let $v \in x^\perp \cap \breve{C}$, i.e., $\langle v, x\rangle = 0$ and $\langle v, y\rangle \leq 0$ for all $y \in C$. Therefore, we have for $y \in C \setminus \{x\}$

$$\begin{aligned}\|y - (x+v)\|^2 &= \|y\|^2 - 2\langle y, x+v\rangle + \|x+v\|^2 \\ &= \|y\|^2 - 2\langle y, x\rangle - 2\langle y, v\rangle + \|x\|^2 + \|v\|^2 \\ &= \|x-y\|^2 + \|v\|^2 - 2\langle y, v\rangle \\ &> \|x - (x+v)\|^2 \ .\end{aligned}$$

This shows that $\Pi_C(x+v) = x$, and thus $x^\perp \cap \breve{C} \subseteq N_x(C)$.

For the other inclusion let $v \in N_x(C)$. In order to show $\langle x, v\rangle = 0$ we may assume $x \neq 0$. For $1 \geq \lambda > 0$ or $0 > \lambda$ we have

$$\begin{aligned}\|v\|^2 = \|x - (x+v)\|^2 < \|(1-\lambda)\cdot x - (x+v)\|^2 &= \|\lambda x + v\|^2 \\ &= \lambda^2\|x\|^2 + 2\lambda\langle x, v\rangle + \|v\|^2 \ ,\end{aligned}$$

or equivalently $\langle x, v\rangle > -\frac{\lambda}{2}\cdot\|x\|^2$. This implies $\langle x, v\rangle = 0$, as $\lambda$ may be both positive or negative arbitrarily small.

In order to show $v \in \breve{C}$, let $y \in C \setminus \{x\}$ and $1 \geq \lambda > 0$. We get

$$\begin{aligned}\|v\|^2 = \|x - (x+v)\|^2 \leq \|x + \lambda(y-x) - (x+v)\|^2 &= \|\lambda(y-x) - v\|^2 \\ &= \lambda^2\|y-x\|^2 - 2\lambda\langle y-x, v\rangle + \|v\|^2 \\ &= \lambda^2\|y-x\|^2 - 2\lambda\langle y, v\rangle + \|v\|^2 \ ,\end{aligned}$$

or equivalently $\langle y, v\rangle \leq \frac{\lambda}{2}\cdot\|y-x\|^2$. Since this holds for all $1 \geq \lambda > 0$, we get $\langle y, v\rangle \leq 0$, and therefore $v \in \breve{C}$. This finishes the proof. □

We may now define the projection map for spherical convex sets. For $K \in \mathcal{K}(S^{n-1})$ and $C := \operatorname{cone}(K)$ let

$$\Pi_K\colon S^{n-1}\setminus\breve{K} \to K \ , \quad \Pi_K(p) := \|\Pi_C(p)\|^{-1}\cdot\Pi_C(p) \ , \tag{3.2}$$

where $\Pi_C\colon \mathbb{R}^n \to C$ denotes the projection map onto $C$. This map is well-defined, as $\Pi_C(x) = 0$ iff $x \in \breve{C}$ by Proposition 3.1.5. Moreover, $\Pi_K$ is indeed the projection map onto $K$, which is shown in the following lemma.

**Lemma 3.1.6.** *Let $K \in \mathcal{K}(S^{n-1})$, and let $\Pi_K\colon S^{n-1}\setminus\breve{K} \to K$ be defined as in* (3.2)*. Then for $p \in S^{n-1}\setminus\breve{K}$*

$$\operatorname{argmin} d(p, K) = \{\Pi_K(p)\} \ .$$

*Proof.* Recall that $\breve{K} = \{v \in S^{n-1} \mid d(K, v) \geq \frac{\pi}{2}\}$. So we have $S^{n-1}\setminus\breve{K} = \{p \in S^{n-1} \mid d(K, p) < \frac{\pi}{2}\}$. Furthermore, for $0 \leq \alpha < \frac{\pi}{2}$ we have $d(p, q) = \alpha \iff d_e(p, y) = \sin\alpha$ with $y = \cos(\alpha)\cdot q$ resp. $q = \cos(\alpha)^{-1}\cdot y$. As the sine function is strictly monotone increasing on $[0, \frac{\pi}{2})$ and the cosine function is positive on $[0, \frac{\pi}{2})$, the claim now follows from the uniqueness of the projection on the cone $C = \operatorname{cone}(K)$. □

From 3-dimensional examples it is seen that points in $\breve{K}$ may very well have several closest points in $K$. So the limitation of the domain of $\Pi_K$ is essential.

The notion of the normal cone at a point of the convex set can be extended to faces that we define next.

**Definition 3.1.7.** Let $K^e \subseteq \mathbb{R}^n$ be a closed convex set. An affine hyperplane $H \subset \mathbb{R}^n$ is called a *supporting (affine) hyperplane*, if $H \cap K^e \neq \emptyset$ and $K^e$ lies completely in one of the two closed half-spaces defined by $H$. A subset $F^e \subseteq K^e$ is called an *(exposed) face* of $K^e$, if $F^e = K^e \cap H$, where $H$ is a supporting hyperplane of $K^e$.

For a spherical convex set $K \in \mathcal{K}(S^{n-1})$ we call a subset $F \subseteq K$ a face of $K$, if $\operatorname{cone}(F)$ is a face of $\operatorname{cone}(K)$.

**Remark 3.1.8.** One usually distinguishes between 'faces' and 'exposed faces' (cf. [49, Sec. 2.1]). These terms coincide for polyhedral sets, i.e., for intersections of finitely many half-spaces (see for example [49, Sec. 2.4]). In particular, any polyhedral set can be written as the disjoint union of the relative interiors of its (exposed) faces (see [49, Thm. 2.1.2]). Since we will only talk about faces in the context of polyhedral sets (polyhedral cones), we will drop the additional term 'exposed' for brevity.

It can be shown that for a face $F^e \subseteq K^e$ the normal cone is the same for all points in the relative interior

$$\forall x, y \in \operatorname{relint}(F^e) : N_x(K^e) = N_y(K^e) .$$

See for example [49, Sec. 2.2] for a proof of this fact. We may therefore define for every face $F^e$ of $K^e$

$$N_{F^e}(K^e) := N_x(K^e) \quad \text{and} \quad N^S_{F^e}(K^e) := N^S_x(K^e) ,$$

where $x \in \operatorname{relint}(F^e)$. Analogously, if $K \in \mathcal{K}(S^{n-1})$ and if $F$ is a face of $K$ we define

$$N_F(K) := N_p(K) \quad \text{and} \quad N^S_F(K) := N^S_p(K) ,$$

where $p \in \operatorname{relint}(F^e) \cap S^{n-1}$, $F^e := \operatorname{cone}(F)$.

**Lemma 3.1.9.** *Let $C \subset \mathbb{R}^n$ be a polyhedral cone and let $\mathcal{F}^m(C)$ denote the set of all $m$-dimensional faces of $C$. Then $\breve{C}$ is a polyhedral cone, and for $F \in \mathcal{F}^m(C)$ the normal cone $N_F(C)$ is a $(n-m)$-dimensional face of $\breve{C}$. Furthermore, the map*

$$\mathcal{F}^m(C) \to \mathcal{F}^{n-m}(\breve{C}) , \quad F \mapsto N_F(C) \tag{3.3}$$

*is bijective, and its inverse is given by $\mathcal{F}^{n-m}(\breve{C}) \to \mathcal{F}^m(C)$, $\breve{F} \mapsto N_{\breve{F}}(\breve{C})$.*

*Proof.* See for example [47, § 19] for a proof that $\breve{C}$ is polyhedral. The normal cone $N_F(C)$ is a face of $\breve{C}$ by Proposition 3.1.5. If $F \in \mathcal{F}^m(C)$ and $\mathcal{W} := \operatorname{lin}(F)$ is the $m$-dimensional span of $F$, then it is straightforward to show that $N_F(C) = \mathcal{W}^\perp \cap \breve{C}$ and $N_F(C)$ has dimension $n-m$. This also shows the bijectivity of the map defined in (3.3). Finally, recall that $(\breve{C})\breve{} = C$. Since $F = \operatorname{lin}(F) \cap C$ and $N_F(C) = \operatorname{lin}(F)^\perp \cap \breve{C}$ we get the claimed formula for the inverse. □

To illustrate the concept of normal cones let us consider the positive orthant and compute the normal cones for this important convex cone.

**Example 3.1.10.** Let $C = \mathbb{R}^n_+$ the $n$-dimensional positive orthant. The cone $C$ is a polyhedral cone, i.e., it is the intersection of finitely many (linear) half-spaces ($n$ in this case). Furthermore, the positive orthant is a self-dual cone, i.e., we have $\breve{C} = -C = \mathbb{R}^n_-$. Hence, Proposition (3.1.5) implies that $N_x(C) = -x^\perp \cap C$ for $x \in C$.

A typical $k$-dimensional face $F$ of $C$ with $0 \leq k \leq n$ is given by (in)equalities of the form

$$x_1 \geq 0\,,\; x_2 \geq 0\,,\; \ldots\,,\; x_k \geq 0\,, \quad x_{k+1} = x_{k+2} = \ldots = x_k = 0\,,$$

and its relative interior is defined by the (strict in)equalities

$$x_1 > 0\,,\; x_2 > 0\,,\; \ldots\,,\; x_k > 0\,, \quad x_{k+1} = x_{k+2} = \ldots = x_k = 0\,.$$

Note that the number of $k$-dimensional faces of $C$ is given by $\binom{n}{k}$. We will need this later when we compute the intrinsic volumes of $C$. The normal cone of the above defined face $F$ is given by the (in)equalities

$$x_1 = x_2 = \ldots = x_k = 0\,, \quad x_{k+1} \leq 0\,,\; x_{k+2} \leq 0\,,\; \ldots\,,\; x_n \leq 0\,.$$

### 3.1.1 Minkowski addition and spherical analogs

A central notion of convex geometry in euclidean space is the *Minkowski addition*

$$M_1 + M_2 := \{x + y \mid x \in M_1\,,\; y \in M_2\}\,,$$

where $M_1, M_2 \subseteq \mathbb{R}^n$. If $M_1$ and $M_2$ are both convex, then so is their Minkowski addition $M_1 + M_2$. If additionally both sets are compact, i.e., $M_1$ and $M_2$ are convex bodies in $\mathbb{R}^n$, then so is $M_1 + M_2$. In fact, the Minkowski addition gives the set of convex bodies $\mathcal{K}(\mathbb{R}^n)$ the structure of a commutative semigroup (cf. [49, Sec. 1.7]).

This rich structure of $\mathcal{K}(\mathbb{R}^n)$ due to the Minkowski addition unfortunately does not exist in the set of spherical convex sets $\mathcal{K}(S^{n-1})$. However, some special cases of Minkowski addition do have spherical analogs, which will occupy us in the remainder of this section.

One special case of Minkowski addition is the *direct product*

$$\bar{M}_1 \times \bar{M}_2 = \{(\bar{x}, \bar{y}) \mid \bar{x} \in \bar{M}_1\,,\; \bar{y} \in \bar{M}_2\} \subseteq \mathbb{R}^n\,,$$

where $\bar{M}_1 \subseteq \mathbb{R}^{n_1}$, $\bar{M}_2 \subseteq \mathbb{R}^{n_2}$, and $n := n_1 + n_2$. If we set

$$M_1 := \{(\bar{x}, 0) \mid \bar{x} \in \bar{M}_1\} \subseteq \mathbb{R}^n\,, \quad M_2 := \{(0, \bar{y}) \mid \bar{y} \in \bar{M}_2\} \subseteq \mathbb{R}^n\,,$$

then we get $\bar{M}_1 \times \bar{M}_2 = M_1 + M_2$. This product construction carries over to the spherical setting in the following way.

**Definition 3.1.11.** For $\bar{U}_1 \subseteq S^{n_1-1}$ and $\bar{U}_2 \subseteq S^{n_2-1}$ let

$$\bar{U}_1 \circledast \bar{U}_2 := (\mathrm{cone}(\bar{U}_1) \times \mathrm{cone}(\bar{U}_2)) \cap S^{n-1}\,,$$

where $n := n_1 + n_2$. Recursively, we also define $\bar{U}_1 \circledast \ldots \circledast \bar{U}_k$ for $\bar{U}_i \subseteq S^{n_i-1}$, $i = 1, \ldots, k$.

If $\bar{C}_1 \subseteq \mathbb{R}^{n_1}$ and $\bar{C}_2 \subseteq \mathbb{R}^{n_2}$ are two closed convex cones, then the product of these cones $\bar{C}_1 \times \bar{C}_2$ is also a closed convex cone. Therefore, if $\bar{K}_1 \in \mathcal{K}(S^{n_1-1})$ and $\bar{K}_2 \in \mathcal{K}(S^{n_2-1})$, then $\bar{K}_1 \circledast \bar{K}_2 \in \mathcal{K}(S^{n-1})$, where $n = n_1 + n_2$.

Another special case of Minkowski addition is given by *tubes*.

**Definition 3.1.12.** For $M^e \subseteq \mathbb{R}^n$ and $r \geq 0$ the *tube of radius* $r$ around $M^e$ is defined as

$$\mathcal{T}^e(M^e, r) := \{x \in \mathbb{R}^n \mid \exists y \in M^e : d^e(x, y) \leq r\}\,.$$

For $M \subseteq S^{n-1}$ and $\alpha \geq 0$ the tube of radius $\alpha$ around $M$ is defined as

$$\mathcal{T}(M, \alpha) := \{p \in S^{n-1} \mid \exists q \in M : d(p, q) \leq \alpha\}\,.$$

**Remark 3.1.13.** In the euclidean case the tube of radius $r$ around the set $M^e \subseteq \mathbb{R}^n$ can also be written as the Minkowski addition

$$\mathcal{T}^e(M^e, r) = M^e + B_n(r) \ ,$$

where $B_n(r) := \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$ denotes the closed ball of radius $r$ in $\mathbb{R}^n$ centered at the origin.

In the case where $M^e \subseteq \mathbb{R}^n$ and $M \subseteq S^{n-1}$ are closed sets, we can also write the tubes in the following form

$$\begin{aligned} \mathcal{T}^e(M^e, r) &= \{x \in \mathbb{R}^n \mid d^e(x, M^e) \leq r\} \ , \\ \mathcal{T}(M, \alpha) &= \{p \in S^{n-1} \mid d(p, M) \leq \alpha\} \ , \end{aligned}$$

where $d^e(x, M^e) := \min\{d(x, y) \mid y \in M^e\}$ and similarly for $d(p, M)$.

**Remark 3.1.14.** If $K \in \mathcal{K}(S^{n-1})$ is a spherically convex set, and $C := \operatorname{cone}(K)$, then using the projection function $\Pi_K$ defined in (3.2) it is straightforward to show that for $0 \leq \alpha < \frac{\pi}{2}$

$$\mathcal{T}(K, \alpha) = \mathcal{T}^e(C, \sin(\alpha)) \cap S^{n-1} \ .$$

In euclidean space the tube around a convex body is again a convex body, as it arises as the Minkowki addition of a convex body with a closed ball, which is also a convex body. In contrast to that, if $K \in \mathcal{K}(S^{n-1})$ is a spherical convex set, then it is a nontrivial question if the tube $\mathcal{T}(K, \alpha)$ is again convex. We will treat this question in the following section.

### 3.1.2 On the convexity of spherical tubes

In euclidean space tubes of convex bodies are again convex bodies. We have seen this in the last section. However, in the sphere the situation is very different. In fact, we will see that if $K \in \mathcal{K}(S^{n-1})$ is polyhedral, i.e., $\operatorname{cone}(K)$ is the intersection of finitely many half-spaces, then for all $\alpha > 0$ we have $\mathcal{T}(K, \alpha) \notin \mathcal{K}(S^{n-1})$.

The following lemma collects some trivial cases where a tube around a set fails to be convex.

**Lemma 3.1.15.** *Let $U \subseteq S^{n-1}$, $U \neq \emptyset$, and let $\alpha \geq 0$.*

1. *If $\alpha \geq \pi$ then $\mathcal{T}(U, \alpha) = S^{n-1}$. In particular, $\mathcal{T}(U, \alpha) \notin \mathcal{K}(S^{n-1})$.*
2. *If there is a point $p \in U$ such that also $-p \in U$, then $\mathcal{T}(U, \alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$.*
3. *If $\alpha > \frac{\pi}{2}$ then $\mathcal{T}(U, \alpha) \notin \mathcal{K}(S^{n-1})$.*
4. *If $\alpha = \frac{\pi}{2}$ then $\mathcal{T}(U, \alpha) \in \mathcal{K}(S^{n-1}) \iff U = \{p\}$.*

*Proof.* Let's process the statements one by one:

1. Follows from $U \neq \emptyset$ and $\operatorname{diam}(S^{n-1}) = \pi$.
2. Let us assume that $\mathcal{T}(U, \alpha)$ is convex for some $\alpha > 0$. We will show that in this case $\mathcal{T}(U, \alpha) = S^{n-1}$, which proves the claim as $S^{n-1} \notin \mathcal{K}(S^{n-1})$. Let $q \in S^{n-1} \setminus \{\pm p\}$. The 2-dimensional picture of the subspace $L := \operatorname{lin}\{p, q\}$ shows that $q$ lies on a geodesic segment between two points $q_1, q_2$ in $L$, such that $d(q_1, p) \leq \alpha$ and $d(q_2, -p) \leq \alpha$ (cf. Figure 3.1). The convexity of $\mathcal{T}(U, \alpha)$ thus implies $q \in \mathcal{T}(U, \alpha)$, and hence $T(U, \alpha) = S^{n-1}$.

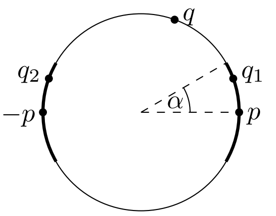



Figure 3.1: Illustration for the proof of Lemma 3.1.15 part (2).

3. It is easily seen that $\mathcal{T}(U, \frac{\pi}{2})$ contains a pair of antipodal points. The claim now follows from part (2) with $U' := \mathcal{T}(U, \frac{\pi}{2})$ and the observation $\mathcal{T}(U, \alpha) = \mathcal{T}(U', \alpha - \frac{\pi}{2})$.
4. The '$\Leftarrow$'-direction is trivial, so let us assume $\mathcal{T}(U, \frac{\pi}{2}) \in \mathcal{K}(S^{n-1})$. If $p, q \in U$ with $q \neq \pm p$, then $\alpha := d(p, q) \in (0, \pi)$, and $L := \text{lin}\{p, q\}$ is a 2-dimensional subspace. Since $L \cap \mathcal{T}(U, \frac{\pi}{2} - \frac{\alpha}{2})$ contains a pair of antipodal points, which is easily verified in $\mathbb{R}^2$, the claim follows from part (3). □

The following proposition shows that in the most important examples of convex cones arising in convex programming, the tube around the corresponding spherical convex set is *never* convex.

**Proposition 3.1.16.** *Let $K \in \mathcal{K}(S^{n-1})$ be a spherical convex set. If the boundary $\partial K$ contains a geodesic segment, then $\mathcal{T}(K, \alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$. If $n \geq 3$ and $K$ is polyhedral and contains more than one point, or if $K = K_1 \circledast K_2$ for some spherical convex sets $K_1$ and $K_2$, then $\mathcal{T}(K, \alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$.*

**Remark 3.1.17.** Proposition 3.1.16 implies that the map

$$\alpha_{\max} \colon \mathcal{K}(S^{n-1}) \to \mathbb{R} , \quad K \mapsto \sup\{\alpha \mid \mathcal{T}(K, \alpha) \in \mathcal{K}(S^{n-1})\}$$

is *not continuous* if $n \geq 3$. Indeed, we will show in Proposition 3.3.4 that the family of polyhedral convex sets lies dense in $\mathcal{K}(S^{n-1})$. So if the map $\alpha_{\max}$ were continuous, then for $n \geq 3$ it had to be the zero map. But this is not true, as is seen by the example of a closed spherical ball. More precisely, if $B_\rho(p)$ denotes the closed spherical ball of radius $\rho$, $0 \leq \rho \leq \frac{\pi}{2}$, around the point $p$, then $\mathcal{T}(B_\rho(p), \alpha) = B_{\rho+\alpha}(p)$. From this it follows that $\alpha_{\max}(B_\rho(p)) = \frac{\pi}{2} - \rho$.

**Remark 3.1.18.** In Corollary 4.1.13, assuming $K \in \mathcal{K}(S^{n-1})$ has smooth boundary, we will give an upper bound for the maximal radius $\alpha_0 \in [0, \pi]$ such that $\mathcal{T}(K, \alpha) \in \mathcal{K}(S^{n-1})$ for all $\alpha \in [0, \alpha_0]$.

*Proof of Proposition 3.1.16.* By Lemma 3.1.15 parts (3) and (4) we may assume that $0 < \alpha < \frac{\pi}{2}$. Denoting $C := \text{cone}(K)$ we have $\mathcal{T}(K, \alpha) = \mathcal{T}^e(C, \sin(\alpha)) \cap S^{n-1}$ (cf. Remark 3.1.14), so we may argue over the cone $C$. Let $p_1, p_2 \in \partial K$, $p_1 \neq \pm p_2$, with $\text{geod}(p_1, p_2) \subset \partial K$, i.e., $(1 - \lambda) \cdot p_1 + \lambda \cdot p_2 \in \partial C$ for all $\lambda \in [0, 1]$. Let $x := \frac{1}{2} \cdot p_1 + \frac{1}{2} \cdot p_2$, and let $v \in N_x^S(C)$. We now consider the points

$$\begin{aligned} q_1 &:= \cos(\alpha) \cdot p_1 + \sin(\alpha) \cdot v , \\ q_2 &:= \cos(\alpha) \cdot p_2 + \sin(\alpha) \cdot v , \\ y &:= \tfrac{1}{2} \cdot q_1 + \tfrac{1}{2} \cdot q_2 \\ &= \cos(\alpha) \cdot x + \sin(\alpha) \cdot v . \end{aligned}$$

Then $\|q_1\| = \|q_2\| = 1$, $\|y\| < 1$ and $d_e(q_1, C) \leq \sin\alpha$, $d_e(q_2, C) \leq \sin\alpha$, and $d_e(y, C) = \sin\alpha$. So $q_1, q_2 \in \mathcal{T}(K,\alpha)$ and $y^\circ := \|y\|^{-1}\cdot y \in \operatorname{geod}(q_1,q_2)$. But $d(y^\circ, K) = \arcsin\left(\frac{\sin\alpha}{\|y\|}\right) > \alpha$, so $y^\circ \notin \mathcal{T}(K,\alpha)$, which shows that $\mathcal{T}(K,\alpha)$ is not convex.

As for the additional statements, if $n \geq 3$ and $K$ is polyhedral and contains more than one point, then $K$ either consists of two antipodal points, in which case $\mathcal{T}(K,\alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$ by Lemma 3.1.15 part (2), or the dimension of the linear hull of $K$ is at least 2. In the second case the boundary $\partial K$ contains a geodesic segment, and hence $\mathcal{T}(K,\alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$.

In the case $K = K_1 \circledast K_2$ for $K_i \in \mathcal{K}(S^{n_i-1})$, $i = 1, 2$, let $p_i \in \partial K_i$, $i = 1, 2$. Then $\{(\cos(\rho)\cdot p_1, \sin(\rho)\cdot p_2) \mid 0 \leq \rho \leq \frac{\pi}{2}\} \subseteq \partial K$, i.e., a geodesic segment is contained in the boundary of $K$, and thus $\mathcal{T}(K,\alpha) \notin \mathcal{K}(S^{n-1})$ for all $\alpha > 0$. □

The assumption that $\mathcal{T}(K,\alpha)$ is convex may not always be satisfiable, but if it is, then we have an important property that we state in the following proposition.

**Proposition 3.1.19.** *Let $K \in \mathcal{K}(S^{n-1})$ and let $\mathcal{T}(K,\alpha_0) \in \mathcal{K}(S^{n-1})$, $\alpha_0 > 0$. Then for all $K' \in \mathcal{K}(S^{n-1})$ such that $K \cap K' = \emptyset$ and $d(K,K') := \min\{d(q,q') \mid q \in K,\ q' \in K'\} < \alpha_0$ there exists a unique pair $(p,p') \in K \times K'$ such that $d(p,p') = d(K,K')$.*

**Lemma 3.1.20.** *Let $K \in \mathcal{K}(S^{n-1})$ and let $\mathcal{T}(K,\alpha_0) \in \mathcal{K}(S^{n-1})$, $\alpha_0 > 0$. Then $\mathcal{T}(K,\alpha) \in \mathcal{K}(S^{n-1})$ for all $0 \leq \alpha \leq \alpha_0$.*

*Proof.* From Lemma 3.1.15 part (3) we get $\alpha_0 \leq \frac{\pi}{2}$, and with part (4) from the same lemma we can easily treat the case $\alpha_0 = \frac{\pi}{2}$. So we may assume $\alpha_0 < \frac{\pi}{2}$. We show the contraposition, i.e., we assume that for some $0 < \alpha < \alpha_0$ we have $\mathcal{T}(K,\alpha) \notin \mathcal{K}(S^{n-1})$ and we need to show that also $\mathcal{T}(K,\alpha_0) \notin \mathcal{K}(S^{n-1})$.

Let $p_1, p_2 \in \mathcal{T}(K,\alpha)$ such that $q \in \operatorname{geod}(p_1,p_2)$ but $q \notin \mathcal{T}(K,\alpha)$, i.e., $\rho := d(q,K) > \alpha$. If $\rho > \alpha_0$, then we are done, as in this case $\mathcal{T}(K,\alpha_0)$ is not convex. So we assume $\rho \leq \alpha_0$ for the rest of the proof. This in particular implies $\rho < \frac{\pi}{2}$ so that the projection of $q$ onto $K$ is well-defined. Let $\tilde{q}$ denote the rotation of $q$ in the 2-dimensional plane $L := \operatorname{lin}\{q, \Pi_K(q)\}$ by an angle of $\alpha_0 - \alpha$ in direction away from $K$ (cf. Figure 3.2(a)). It follows that $d(\tilde{q},K) = d(q,K) + \alpha_0 - \alpha > \alpha_0$. To finish the proof it suffices to show that $\tilde{q}$ lies on a geodesic between points in $\mathcal{T}(K,\alpha_0)$, as this implies that $\mathcal{T}(K,\alpha_0)$ not convex. To show this let

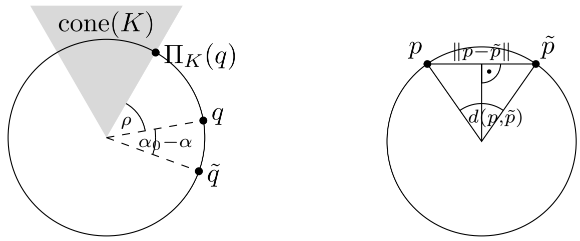
$$p_1 = x_1 + y_1\ ,\ p_2 = x_2 + y_2\ ,\ \text{with}\ \ x_1, x_2 \in L\ ,\ y_1, y_2 \in L^\perp\ .$$

Let $\tilde{x}_1$ and $\tilde{x}_2$ be the results of the same rotation in $L$ that we applied on $q$, and let $\tilde{p}_i := \tilde{x}_i + y_i$, $i = 1, 2$. Then (cf. Figure 3.2(b))

$$\|p_i - \tilde{p}_i\| = \|x_i - \tilde{x}_i\| = \|x_i\|\cdot 2\cdot\sin(\tfrac{\alpha_0-\alpha}{2}) \leq 2\cdot\sin(\tfrac{\alpha_0-\alpha}{2})\ ,$$

which implies $d(p_i,\tilde{p}_i) = 2\cdot\arcsin(\frac{p_i-\tilde{p}_i}{2}) \leq \alpha_0 - \alpha$, $i = 1, 2$. By the triangle inequality for $d(.,.)$ we have $\tilde{p}_1, \tilde{p}_2 \in \mathcal{T}(K,\alpha_0)$. Furthermore, as the above described operation that sends $q$ to $\tilde{q}$ and $p_i$ to $\tilde{p}_i$, $i = 1, 2$, is an element of the orthogonal group, we get $\tilde{q} \in \operatorname{geod}(\tilde{p}_1,\tilde{p}_2)$. This finishes the proof. □

*Proof of Proposition 3.1.19.* We reproduce the proof from [30, Hilfssatz 7.1] in a simplified form. By Lemma 3.1.15 part (3) we have $\alpha_0 \leq \frac{\pi}{2}$ and by Lemma 3.1.20 we have $\mathcal{T}(K,\alpha) \in \mathcal{K}(S^{n-1})$, where $\alpha := d(K,K') < \alpha_0$.




(a) Definition of $\tilde{q}$

(b) $\|p-\tilde{p}\| = 2 \cdot \sin\left(\frac{d(p,\tilde{p})}{2}\right)$

Figure 3.2: Illustrations for the proof of Lemma 3.1.20.

We assume that we have two pairs $(p,p'),(q,q') \in K \times K'$, $(p,p') \neq (q,q')$, which both satisfy $d(p,p') = d(q,q') = d(K,K')$, so that we need to derive a contradiction. The first observation is that $p' \neq \pm q'$, which can be seen in the following way. As $d(p,p') = d(K,K') < \alpha_0 \leq \frac{\pi}{2}$, we have $p = \Pi_K(p')$. Similarly, we have $q = \Pi_K(q')$. So if $p' = q'$ then also $p = q$, contradicting the assumption $(p,p') \neq (q,q')$. If $p' = -q'$ then we have $p', -p' \in \mathcal{T}(K,\alpha)$, and by Lemma 3.1.15 part (2) we get $\mathcal{T}(K,\alpha_0) = \mathcal{T}(\mathcal{T}(K,\alpha),\alpha_0-\alpha) \notin \mathcal{K}(S^{n-1})$, which is a contradiction.

Now that we have $p' \neq \pm q'$ we can argue over the geodesic segment $\operatorname{geod}(p',q')$. As $\operatorname{geod}(p',q') \subset \mathcal{T}(K,\alpha) \cap K'$ and $\mathcal{T}(K,\rho) \cap K' = \emptyset$ for all $0 \leq \rho < \alpha$, we get $\operatorname{geod}(p',q') \subset \{\tilde{p} \mid d(\tilde{p},K) = \alpha\} = \partial\mathcal{T}(K,\alpha)$. But then Proposition 3.1.16 implies that $\mathcal{T}(K,\alpha_0) = \mathcal{T}(\mathcal{T}(K,\alpha),\alpha_0-\alpha) \notin \mathcal{K}(S^{n-1})$, which is a contradiction. This finishes the proof. □

## 3.2 The metric space of spherical convex sets

Besides being important in its own, tubes can be used to define the Hausdorff distances, which turn $\mathcal{K}(\mathbb{R}^n)$ and $\mathcal{K}(S^{n-1})$ into metric spaces. See Figure 3.3 for a small display.

**Definition 3.2.1.** The Hausdorff distance on $\mathcal{K}(\mathbb{R}^n)$ is defined by

$$\begin{aligned} d_{\mathrm{H}}^e(K_1^e,K_2^e) &= \max\left\{ \min\{r \geq 0 \mid K_2^e \subseteq \mathcal{T}^e(K_1^e,r)\},\ \min\{s \geq 0 \mid K_1^e \subseteq \mathcal{T}^e(K_2^e,s)\}\right\} \\ &= \max\left\{ \max\{d^e(K_1^e,y) \mid y \in K_2^e\},\ \max\{d^e(K_2^e,x) \mid x \in K_1^e\}\right\}, \end{aligned}$$

for $K_1^e,K_2^e \in \mathcal{K}(\mathbb{R}^n)$. The Hausdorff distance on $\mathcal{K}(S^{n-1})$ is defined by

$$\begin{aligned} d_{\mathrm{H}}(K_1,K_2) &= \max\left\{ \min\{\alpha \geq 0 \mid K_2 \subseteq \mathcal{T}(K_1,\alpha)\},\ \min\{\beta \geq 0 \mid K_1 \subseteq \mathcal{T}(K_2,\beta)\}\right\} \\ &= \max\left\{ \max\{d(K_1,q) \mid q \in K_2\},\ \max\{d(K_2,p) \mid p \in K_1\}\right\}, \end{aligned}$$

for $K_1,K_2 \in \mathcal{K}(S^{n-1})$.

**Remark 3.2.2.** Recall that for $K_1,K_2 \in \mathcal{K}(S^{n-1})$ we have defined $d(K_1,K_2) = \min\{d(p,q) \mid p \in K_1, q \in K_2\}$. This does not define a metric, as $d(K_1,K_2) = 0$ iff $K_1 \cap K_2 \neq \emptyset$, which is not equivalent to $K_1 = K_2$. In general we have the inequality $d(K_1,K_2) \leq d_{\mathrm{H}}(K_1,K_2)$.

**Proposition 3.2.3.** *The map $d_{\mathrm{H}}^e\colon \mathcal{K}(\mathbb{R}^n) \times \mathcal{K}(\mathbb{R}^n) \to \mathbb{R}$ is a metric on $\mathcal{K}(\mathbb{R}^n)$, and the map $d_{\mathrm{H}}\colon \mathcal{K}(S^{n-1}) \times \mathcal{K}(S^{n-1}) \to \mathbb{R}$ is a metric on $\mathcal{K}(S^{n-1})$. Both $\mathcal{K}(\mathbb{R}^n)$ and $\mathcal{K}(S^{n-1})$ are complete as metric spaces. Additionally, $\mathcal{K}(S^{n-1})$ is compact.*

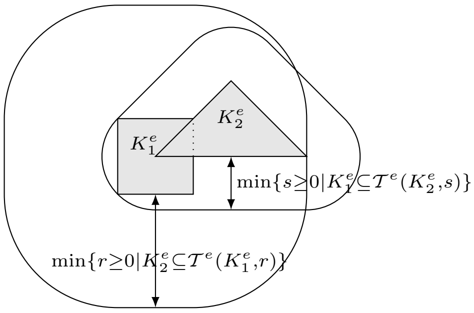



Figure 3.3: The Hausdorff distance between two sets $K_1^e, K_2^e \in \mathcal{K}(\mathbb{R}^2)$.

*Proof.* See for example [49, Sec. 1.8] for the euclidean statements. The statements that the Hausdorff distance is a metric on $\mathcal{K}(S^{n-1})$, and that $\mathcal{K}(S^{n-1})$ is a complete metric space can be shown in the same vein. As for the compactness of $\mathcal{K}(S^{n-1})$, let $\mathcal{C}(S^{n-1})$ denote the set of nonempty closed subsets of $S^{n-1}$. The Hausdorff distance turns $\mathcal{C}(S^{n-1})$ into a compact metric space (cf. [49, p. 56, Note 2]). As $\mathcal{K}(S^{n-1})$ is a closed subset of $\mathcal{C}(S^{n-1})$, it is compact. $\square$

Concerning the spherical Hausdorff metric and the duality map, note that it is not true that the duality map is an isometry. Consider for example a closed spherical ball of radius $\rho$, $0 < \rho < \frac{\pi}{2}$, in $S^{n-1}$ (we will call these convex sets circular caps; cf. Section 3.3), and denote this convex set by $K_\rho$. Then we have $d_{\text{H}}(K_\rho, -K_\rho) = \pi - \rho$. But the dual $\breve{K}_\rho$ is a closed spherical ball of radius $\frac{\pi}{2} - \rho$, so that $d_{\text{H}}(\breve{K}_\rho, -\breve{K}_\rho) = \frac{\pi}{2} + \rho$. So $d_{\text{H}}(K_\rho, -K_\rho) \neq d_{\text{H}}(\breve{K}_\rho, -\breve{K}_\rho)$ for $\rho \neq \frac{\pi}{4}$ and therefore the map $K \mapsto \breve{K}$ is not an isometry. But these are only 'global' effects, i.e., locally the duality map is in fact isometric, which is shown by the following proposition.

**Proposition 3.2.4.** *Let $K_1, K_2 \in \mathcal{K}(S^{n-1})$ with $d_{\text{H}}(K_1, K_2) < \frac{\pi}{2}$. Then*

$$d_{\text{H}}(K_1, K_2) = d_{\text{H}}\big(\breve{K}_1, \breve{K}_2\big) ,$$

*i.e., the duality map is a local isometry.*

Recall that $\breve{K} = \{p \in S^{n-1} \mid d(K, p) \geq \frac{\pi}{2}\}$, cf. Remark 3.1.2. In the proof we will make use of the following lemma.

**Lemma 3.2.5.** *Let $K \in \mathcal{K}(S^{n-1})$. Then for $p \in S^{n-1} \setminus (K \cup \breve{K})$ we have $d(K, p) + d(\breve{K}, p) = \frac{\pi}{2}$. Furthermore, $\mathcal{T}(\breve{K}, \alpha) = \{p \in S^{n-1} \mid d(K, p) \geq \frac{\pi}{2} - \alpha\}$ for $0 \leq \alpha < \frac{\pi}{2}$.*

*Proof.* For the first part of the claim see [30, Hilfssatz 2.1] or [13, Lemma 2.3]. As for the second part, we have $\breve{K} \subseteq \{q \in S^{n-1} \mid d(K, q) \geq \frac{\pi}{2} - \alpha\}$, and for $p \in S^{n-1} \setminus (K \cup \breve{K})$ we have $d(\breve{K}, p) \leq \alpha$ iff $d(K, p) \geq \frac{\pi}{2} - \alpha$. $\square$

*Proof of Proposition 3.2.4.* This was shown in [30, Hilfssatz 2.2]. We reproduce the proof for completeness. Let $\alpha := d_{\text{H}}(K_1, K_2) < \frac{\pi}{2}$. If $p \in S^{n-1}$ is such that $d(K_1, p) < \frac{\pi}{2} - \alpha$, then there exists $q \in K_1 \subseteq \mathcal{T}(K_2, \alpha)$ such that $d(q, p) < \frac{\pi}{2} - \alpha$.

The triangle inequality implies that $d(K_2, p) < \frac{\pi}{2}$. Therefore, we get

$$\begin{aligned}
\breve{K}_2 &= \{p \in S^{n-1} \mid d(K_2, p) \geq \tfrac{\pi}{2}\} \\
&\subseteq \{p \in S^{n-1} \mid d(K_1, p) \geq \tfrac{\pi}{2} - \alpha\} \\
&= \mathcal{T}(\breve{K}_1, \alpha) \ ,
\end{aligned}$$

where the last equality follows from Lemma 3.2.5. By symmetry, we have $\breve{K}_1 \subseteq \mathcal{T}(\breve{K}_2, \alpha)$, and thus $d_{\mathrm{H}}(\breve{K}_1, \breve{K}_2) \leq \alpha$. Interchanging the roles of $K_1, K_2$ and $\breve{K}_1, \breve{K}_2$ yields the claimed result. $\square$

A special class of convex sets in the sphere is the family of subspheres. We distinguish between subspheres and non-subspheres, and use the following conventions: For $k = 0, \ldots, n-2$

$$\begin{aligned}
\mathcal{S}^k(S^{n-1}) &:= \{S \subset S^{n-1} \mid S \text{ is a } k\text{-dim. subsphere}\} \ , \\
\mathcal{S}^*(S^{n-1}) &:= \bigcup_{k=0}^{n-2} \mathcal{S}^k(S^{n-1}) \ , \\
\mathcal{K}^c(S^{n-1}) &:= \mathcal{K}(S^{n-1}) \setminus \mathcal{S}^*(S^{n-1}) \\
&= \{K \in \mathcal{K}(S^{n-1}) \mid K \text{ is not a subsphere}\} \ .
\end{aligned} \tag{3.4}$$

Note that there is a canonical bijection between $\mathcal{S}^k(S^{n-1})$, the set of $k$-dimensional subspheres of $S^{n-1}$, and $\mathrm{Gr}_{n,k+1}$, the set of $(k+1)$-dimensional subspaces of $\mathbb{R}^n$, given by

$$\begin{aligned}
\mathrm{Gr}_{n,k+1} \to \mathcal{S}^k(S^{n-1}) \ , &\qquad \mathcal{W} \mapsto \mathcal{W} \cap S^{n-1} \ , \\
\mathcal{S}^k(S^{n-1}) \to \mathrm{Gr}_{n,k+1} \ , &\qquad S \mapsto \{\lambda p \mid \lambda \in \mathbb{R} \, , \ p \in S\} \ .
\end{aligned}$$

Note that $\breve{S} = S^\perp := \mathcal{W}^\perp \cap S^{n-1}$ for $S \in \mathcal{S}^*(S^{n-1})$ with $\mathcal{W} := \mathrm{lin}(S)$.

As the non-subspheres are central objects for our study, we will call elements in $\mathcal{K}^c(S^{n-1})$ *caps*. Note that this naming is different from other works, where 'cap' may stand for a spherical ball (cf. for example [13]). We will denote spherical balls by the term *circular caps*.

**Remark 3.2.6.** 1. Here is a mnemonic for this specific way of speaking:

*"You can wear caps*  *, but you cannot wear subspheres*  *."*

2. An element $K \in \mathcal{K}(S^{n-1})$ is a cap iff $\exists p \in K : -p \notin K$.

3. An element $K \in \mathcal{K}(S^{n-1})$ is a cap iff its dual $\breve{K}$ is a cap.

The existence of the subspheres leads to a fundamental difference between the metric spaces $\mathcal{K}(\mathbb{R}^n)$ and $\mathcal{K}(S^{n-1})$. Note that $\mathcal{K}(\mathbb{R}^n)$ is path-connected, as for every $K \in \mathcal{K}(\mathbb{R}^n)$ we have the continuous path $[0,1] \to \mathcal{K}(\mathbb{R}^n)$, $t \mapsto t \cdot K$, which connects $K$ with $\{0\} \in \mathcal{K}(\mathbb{R}^n)$. This is not the case in the spherical setting.

**Proposition 3.2.7.** *The decomposition of* $\mathcal{K}(S^{n-1})$ *in its connected components is given by*

$$\mathcal{K}(S^{n-1}) = \mathcal{K}^c(S^{n-1}) \ \dot{\cup} \ \dot{\bigcup_{k=0}^{n-2}} \mathcal{S}^k(S^{n-1}) \ ,$$

*and the components are path-connected (see Figure 3.4).*

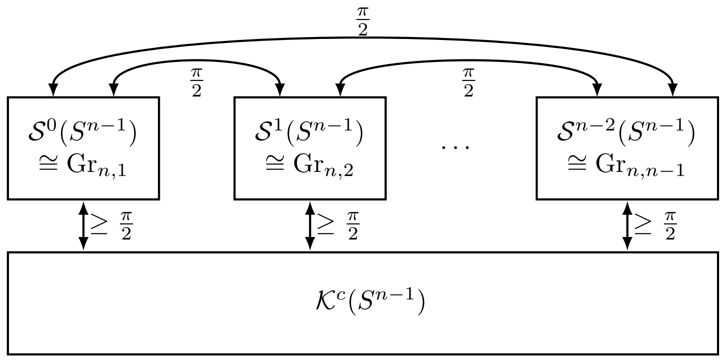



Figure 3.4: The decomposition of $\mathcal{K}(S^{n-1})$ in its connected components and their pairwise Hausdorff distances.

**Lemma 3.2.8.** *If $K \in \mathcal{K}^c(S^{n-1})$ and $S \in \mathcal{S}^*(S^{n-1})$ then $d_{\mathrm{H}}(K, S) \geq \frac{\pi}{2}$. Furthermore, if $S_1 \in \mathcal{S}^{k_1}(S^{n-1})$ and $S_2 \in \mathcal{S}^{k_2}(S^{n-1})$ with $k_1 \neq k_2$, then $d_{\mathrm{H}}(S_1, S_2) = \frac{\pi}{2}$.*

*Proof.* From Theorem 2.2.5 it follows that $K \cap S^\perp \neq \emptyset$ or $\breve{K} \cap S \neq \emptyset$, so in particular $K \not\subseteq \mathcal{T}(S, \alpha)$ for all $0 \leq \alpha < \frac{\pi}{2}$ or $S \not\subseteq \mathcal{T}(K, \alpha)$ for all $0 \leq \alpha < \frac{\pi}{2}$, i.e., $d_{\mathrm{H}}(K, S) \geq \frac{\pi}{2}$. Concerning the subspheres, let $\mathcal{W}_i := \operatorname{lin} S_i$, $i = 1, 2$. If $k_1 \neq k_2$ then either $\dim \mathcal{W}_1 + \dim \mathcal{W}_2^\perp > n$ or $\dim \mathcal{W}_1^\perp + \dim \mathcal{W}_2 > n$, in particular $\dim(\mathcal{W}_1 \cap \mathcal{W}_2^\perp) \geq 1$ or $\dim(\mathcal{W}_1^\perp \cap \mathcal{W}_2) \geq 1$. If $\mathcal{W}_1 \cap \mathcal{W}_2^\perp \neq \{0\}$, then $S_1 \not\subseteq \mathcal{T}(S_2, \alpha)$ for all $0 \leq \alpha < \frac{\pi}{2}$, hence $d_{\mathrm{H}}(S_1, S_2) \geq \frac{\pi}{2}$. Similarly, one gets $d_{\mathrm{H}}(S_1, S_2) \geq \frac{\pi}{2}$ for the case $\mathcal{W}_1^\perp \cap \mathcal{W}_2 \neq \{0\}$. Equality follows from $\mathcal{T}(S, \frac{\pi}{2}) = S^{n-1}$ for any $S \in \mathcal{S}^*(S^{n-1})$. $\square$

**Lemma 3.2.9.** *Let $K \in \mathcal{K}^c(S^{n-1})$. Then there exists $p \in S^{n-1}$ and a continuous path $K_t \in \mathcal{K}^c(S^{n-1})$, $t \in [0, 1]$, such that $K_0 = K$ and $K_1 = \{p\}$.*

*Proof.* Let $C := \operatorname{cone}(K)$, and let $v \in \breve{K}$ such that $-v \notin \breve{K}$ (cf. Remark 3.2.6). Note that $K \not\subseteq v^\perp$. For $q \in S^{n-1}$ and $\rho \geq 0$ let $B(q, \rho) := \mathcal{T}(\{q\}, \rho)$ denote the spherical ball of radius $\rho$ around $q$. Since $v \in \breve{K}$ we have $K \subseteq B(-v, \frac{\pi}{2})$. Furthermore, as $K \not\subseteq v^\perp$, there exists $0 < \alpha \leq \frac{\pi}{2}$ such that $K \cap B(-v, \frac{\pi}{2} - \alpha) \neq \emptyset$. This implies that

$$\alpha_0 := \max\{\alpha \in [0, \tfrac{\pi}{2}] \mid K \cap B(-v, \tfrac{\pi}{2} - \alpha) \neq \emptyset\}$$

is positive, i.e., $\alpha_0 > 0$. If $\alpha_0 = \frac{\pi}{2}$, then $B(-v, \frac{\pi}{2} - \alpha_0) = \{-v\}$, and we set $p := -v$. If $\alpha_0 < \frac{\pi}{2}$, then Proposition 3.1.19 implies that $K \cap B(-v, \frac{\pi}{2} - \alpha_0) = \{p\}$ for some $p \in S^{n-1}$. We may now define the continuous path via

$$K_t := K \cap B(-v, \tfrac{\pi}{2} - t \cdot \alpha_0) . \qquad \square$$

*Proof of Proposition 3.2.7.* The path-connectedness of $\mathcal{S}^k(S^{n-1})$ follows from the path-connectedness of the Grassmann manifold $\mathrm{Gr}_{n,k+1}$ (cf. Chapter 5): Elements $S_1, S_2 \in \mathcal{S}^k(S^{n-1})$ are of the form $S_i = \mathcal{W}_i \cap S^{n-1}$ with $\mathcal{W}_i \in \mathrm{Gr}_{n,k+1}$, $i = 1, 2$. If $\mathcal{W}_t$, $1 \leq t \leq 2$ describes a path in $\mathrm{Gr}_{n,k+1}$ between $\mathcal{W}_1$ and $\mathcal{W}_2$ then so does the path $S_t := \mathcal{W}_t \cap S^{n-1}$ in $\mathcal{S}^k(S^{n-1})$.

For the component $\mathcal{K}^c(S^{n-1})$ we can do the same trick as in the euclidean case, i.e., for $K_1, K_2 \in \mathcal{K}^c(S^{n-1})$ we may 'shrink' $K_i$ to $\{p_i\}$ for some $p_i \in S^{n-1}$, $i = 1, 2$, by Lemma 3.2.9. We can then connect the shrinked sets, thus getting a path from $K_1$ to $K_2$ in $\mathcal{K}^c(S^{n-1})$.

The pairwise disjointness of the sets $\mathcal{S}^c(S^{n-1})$ and $\mathcal{S}^k(S^{n-1})$, $1 \leq k \leq n - 2$, as well as their closedness follows from Lemma 3.2.8. $\square$

## 3.3 Subfamilies of spherical convex sets

In the last section we have seen that the set of spherical convex sets decomposes into subspheres and caps. The set of caps $\mathcal{K}^c(S^{n-1})$ should be seen as the essential part of $\mathcal{K}(S^{n-1})$ containing a variety of sets with diverse properties. We need to specify subfamilies of $\mathcal{K}(S^{n-1})$ with whom we can work in a unified way. In this section we will present the different subfamilies of spherical convex sets, that will appear in our analyses.

We begin with the set of *polyhedral convex sets*.

**Definition 3.3.1.** A spherical convex set $K \in \mathcal{K}(S^{n-1})$ is called *polyhedral* if $\operatorname{cone}(K)$ is the intersection of finitely many $n$-dimensional half-spaces

$$\mathcal{K}^p(S^{n-1}) := \{K \in \mathcal{K}(S^{n-1}) \mid \operatorname{cone}(K) = H_1 \cap \ldots \cap H_k\,,\ H_i = (\text{half-space in } \mathbb{R}^n)\}.$$

**Proposition 3.3.2.** *1. If $K \in \mathcal{K}^p(S^{n-1})$ then also $\breve{K} \in \mathcal{K}^p(S^{n-1})$.*

*2. If $K \in \mathcal{K}^p(S^{n-1})$, then the cone $C := \operatorname{cone}(K)$ can be written in the form $C = \operatorname{cone}(\{p_1, \ldots, p_N\})$ for some $p_1, \ldots, p_N \in S^{n-1}$. Moreover, if $C$ is of the above form, then $K = C \cap S^{n-1}$ is polyhedral.*

*Proof.* See for example [47, Sec. 19]. □

**Example 3.3.3.** Our standard example for a polyhedral convex set is the intersection of the positive orthant with the unit sphere $\mathbb{R}^n_+ \cap S^{n-1}$. We have already computed the normal cones for this spherical convex set in Example 3.1.10, and we will meet this example again on several occasions.

By definition, the set of polyhedral convex sets contains the subspheres of $S^{n-1}$, i.e., $\mathcal{S}^*(S^{n-1}) \subset \mathcal{K}^p(S^{n-1})$. Furthermore, every cap can be approximated by polyhedral caps, which is the content of the following proposition.

**Proposition 3.3.4.** *The family of polyhedral convex sets $\mathcal{K}^p(S^{n-1})$ lies dense in the family of spherical convex sets $\mathcal{K}(S^{n-1})$.*

*Proof.* We reproduce the proof from [30, Hilfssatz 2.5] for completeness. Let $K \in \mathcal{K}(S^{n-1})$ and let $\varepsilon > 0$. We need to find a polyhedral set $P$ with $d_{\mathrm{H}}(K, P) < \varepsilon$. For $p \in S^{n-1}$ and $\rho > 0$ let $B^\circ_\rho(p)$ denote the interior of the circular cap around $p$ of radius $\rho$. The family $\{B^\circ_\varepsilon(p) \mid p \in K\}$ forms an open cover of $K$, and by compactness of $K$ there exists an open subcover $\{B^\circ_\varepsilon(p_1), \ldots, B^\circ_\varepsilon(p_k)\}$. The cone $C := \operatorname{cone}(\{p_1, \ldots, p_k\})$ is a polyhedral cone, and thus $P := C \cap S^{n-1} \in \mathcal{K}^p(S^{n-1})$. Furthermore, we have $P \subseteq K$, as $p_1, \ldots, p_k \in K$. It is verified easily that $K \subset \mathcal{T}(P, \varepsilon)$, which finishes the proof. □

So if we use the notation $\overset{\mathrm{d}}{\subset}$ to denote a dense inclusion, we can summarize briefly (omitting the brackets to ease the notation)

$$\mathcal{S}^* \ \subset\ \mathcal{K}^p \ \overset{\mathrm{d}}{\subset}\ \mathcal{K}\,.$$

Another important subfamily of spherical convex sets is given by the set of *regular caps*.

**Definition 3.3.5.** A spherical convex set $K \in \mathcal{K}(S^{n-1})$ is called *regular* if both $K$ and $\breve{K}$ have nonempty interior

$$\mathcal{K}^r(S^{n-1}) := \{K \in \mathcal{K}(S^{n-1}) \mid \operatorname{int}(K) \neq \emptyset \text{ and } \operatorname{int}(\breve{K}) \neq \emptyset\}\,.$$

Note that subspheres are not regular, i.e., we have $\mathcal{K}^r(S^{n-1}) \subset \mathcal{K}^c(S^{n-1})$.

**Remark 3.3.6.** The property $\mathrm{int}(\breve{K}) \neq \emptyset$ is equivalent to $K \cap -K = \emptyset$. In particular, regular cones (cf. Section 2.2) are exactly those cones $C$, which are of the form $C = \mathrm{cone}(K)$ with $K \in \mathcal{K}^r(S^{n-1})$.

In Section 4.1.2 we will define a subfamily $\mathcal{K}^{\mathrm{sm}}(S^{n-1}) \subset \mathcal{K}^r(S^{n-1})$, which we call the family of *smooth caps*. We will give the precise definition later in Section 4.1.2. Instead, let us have a look at our standard example of a smooth cap.

**Example 3.3.7.** Our standard example of a smooth cap is a circular cap $B_\rho(p) = \{q \in S^{n-1} \mid d(p,q) \leq \rho\}$, where $p \in S^{n-1}$ and $0 < \rho < \frac{\pi}{2}$. A circular cap is self-dual iff $\rho = \frac{\pi}{4}$. In this case, we say that $B_{\pi/4}(p)$ is an $n$-dimensional *Lorentz cap*. As is the case for the positive orthant, we will meet circular caps again on several occasions.

In Section 4.1.2 (cf. Proposition 4.1.10) we will see that smooth caps lie dense in $\mathcal{K}^c(S^{n-1})$. In particular, they lie dense in $\mathcal{K}^r(S^{n-1})$ and $\mathcal{K}^r(S^{n-1})$ lies dense in $\mathcal{K}^c(S^{n-1})$. So we may summarize

$$\mathcal{K}^{\mathrm{sm}} \overset{\mathrm{d}}{\subset} \mathcal{K}^r \overset{\mathrm{d}}{\subset} \mathcal{K}^c .$$

Both polyhedral and smooth convex sets are special cases of stratified convex sets, which we define in the remainder of this section. We will state and prove Weyl's tube formula for (euclidean and for spherical) convex sets in Section 4.3 for stratified convex sets. Besides the pleasing fact that this formula specializes to the polyhedral and the smooth case, the main reason for using stratified sets is that the semidefinite cone is neither smooth nor polyhedral, but it is a stratified cone. We will give the formulas for the intrinsic volumes of the semidefinite cone in Section 4.4.1; the derivation of these formulas is outsurced to Section C.2 in the appendix. Although we do not have a concrete use for these formulas yet, we believe that a good understanding of these quantities will reveal important insights in the complexity of semidefinite programming. In any case, it should be evident that this application of Weyl's tube formula justifies the extra effort it takes to state and prove this formula for stratified convex sets.

Before we give the definition of stratified convex sets, we need to recall some elementary concepts from differential geometry. To keep it as simple as possible we restrict ourselves to submanifolds of euclidean space. See for example the introductory chapters in [53] (although we might use a slightly different notation) for the background of the notions we will treat next.

Let $M \subseteq \mathbb{R}^n$ be a smooth submanifold of euclidean space, where smooth generally means $C^\infty$. We say that $\varphi\colon \mathbb{R}^d \to M$ is a local parametrization of $M$, if $\varphi$ is a (smooth) diffeomorphism between $\mathbb{R}^d$ and an open subset of $M$. The *tangent space* $T_pM$ of $M$ in $p$ is by definition the linear space consisting of all vectors, which arise as velocities in $p$ of curves in $M$ passing through $p$, i.e.,

$$T_pM = \{\dot{c}(0) \mid c\colon \mathbb{R} \to M \text{ smooth curve with } c(0) = p\} .$$

We will use the notation $\dot{c}$ as well as $\frac{dc}{dt}$ to denote the derivative of a curve $c$. The tangent spaces are linear subspaces of $\mathbb{R}^n$, and the *normal space* $T_p^\perp M$ of $M$ in $p$ is defined as the orthogonal complement of $T_pM$ in $\mathbb{R}^n$. So we have an orthogonal decomposition $T_p\mathbb{R}^n = T_pM \oplus T_p^\perp M$, where we use the notation $T_p\mathbb{R}^n = \mathbb{R}^n$ to indicate that we consider the vectors as tangent vectors in $p$. The unit sphere has the tangent spaces $T_pS^{n-1} = p^\perp$, and if $M \subseteq S^{n-1}$ is a submanifold of the unit sphere, we have $T_pM \subseteq T_pS^{n-1} = p^\perp$.

We say that $\hat{M}$ is a *conic manifold*, iff $0 \notin \hat{M}$ and $p \in \hat{M}$ implies that $\lambda p \in \hat{M}$ for all $\lambda > 0$. Note that if we have a submanifold $M$ of $S^{n-1}$, we may define $\hat{M} := \{\lambda p \mid \lambda > 0, p \in M\}$ and treat $\hat{M}$ instead of $M$. Note also that if $\hat{M}$ is a conic manifold and $p \in \hat{M}$, then $p \in T_p\hat{M}$ where we use the canonical identification $T_p\mathbb{R}^n = \mathbb{R}^n$. We call this direction $p \in T_p\hat{M}$ the *cone direction*.

The set of all tangent resp. normal spaces forms the *tangent* resp. *normal bundle* (cf. [53, Ch. 3]). We may consider these bundles as submanifolds of $\mathbb{R}^n \times \mathbb{R}^n$ via

$$TM = \bigcup_{p\in M} \{p\} \times T_pM \ , \quad T^\perp M = \bigcup_{p\in M} \{p\} \times T_p^\perp M \ .$$

Furthermore, we consider the *unit normal bundle*

$$T^{\oplus}M := \bigcup_{p\in M} \{p\} \times (T_p^\perp M \cap S^{n-1}) \ . \tag{3.5}$$

And for $M \subseteq S^{n-1}$ we also consider the *spherical normal bundle*

$$T^S M := \bigcup_{p\in M} \{p\} \times (T_p^\perp M \cap S^{n-1} \cap p^\perp) \ . \tag{3.6}$$

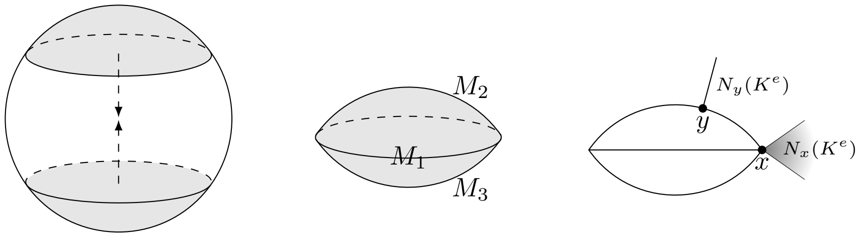
Note that the difference between the unit and the spherical normal bundle lies in the fact that the fiber $\{p\} \times (T_p^\perp M \cap S^{n-1} \cap p^\perp)$, $p \in M$, of the spherical normal bundle lies in the tangent space of the unit sphere $T_pS^{n-1} = p^\perp$.

**Remark 3.3.8.** If $M \subseteq \mathbb{R}^n$ is a smooth $d$-dimensional manifold, then the tangent and the normal bundle are smooth submanifolds of $\mathbb{R}^n \times \mathbb{R}^n$ of dimension $2d$ resp. $d + (n-d) = n$. Furthermore, given a local parametrization $\varphi\colon \mathbb{R}^d \to M$, one can construct local trivializations $\Phi\colon \mathbb{R}^d \times \mathbb{R}^d \to TM$, $\Phi^\perp\colon \mathbb{R}^d \times \mathbb{R}^{n-d} \to T^\perp M$, i.e.,

1. we have
$$P \circ \Phi = \varphi \ , \quad P \circ \Phi^\perp = \varphi \ , \tag{3.7}$$
where $P\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $(x, y) \mapsto x$, denotes the projection onto the first component,
2. for every $u \in \mathbb{R}^d$ the maps
$$\begin{aligned} \Phi_u\colon \mathbb{R}^d \to T_pM \ , &\quad \Phi_u(v) := \Phi(u, v) \ , \\ \Phi_u^\perp\colon \mathbb{R}^{n-d} \to T_p^\perp M \ , &\quad \Phi_u^\perp(v) := \Phi^\perp(u, v) \ , \end{aligned}$$
where $p := \varphi(u)$, are linear.

The unit normal bundle is a hypersurface, i.e., a submanifold of codimension 1, of the normal bundle, and the spherical normal bundle is a hypersurface of the unit normal bundle.

The tangent and the normal bundle are both so-called *vector bundles*, as all fibers of the canonical projection maps (3.7) are vector spaces. Loosely speaking, these bundles are conglomerates of vector spaces, which are conjoined in a smooth way. The unit and the spherical normal bundles are *sphere bundles*, as all fibers are subspheres of the unit sphere. For Weyl's tube formulas in Section 4.3 we need




(a) construction (b) decomposition (c) normal cones

Figure 3.5: A lenticular disc is a nontrivial example for a (euclidean) stratified convex set.

to consider another class of fiber bundles, where each fiber is given by (the relative interior of) a convex cone, respectively its intersection with the unit sphere.

Recall that for a closed convex (not necessarily compact) set $K^e \subseteq \mathbb{R}^n$, and a point $x \in K^e$, we have given the normal cone $N_x(K^e)$ and its intersection with the unit sphere $N_x^S(K^e) = N_x(K^e) \cap S^{n-1}$ (cf. Definition 3.1.3). If we have a manifold $M^e \subseteq \mathbb{R}^n$, which lies in $K^e$, i.e., $M^e \subseteq K^e$, then we define the *duality bundle* $NM^e$ via

$$NM^e \;:=\; \bigcup_{x\in M^e} \{x\} \times \operatorname{relint}(N_x(K^e)) \quad \subseteq \quad T^\perp M^e \;. \tag{3.8}$$

Note that the duality bundle does not depend solely on $M^e$ but on both $M^e$ and the convex set $K^e$. To keep the notation simple, we just write $NM^e$. Note also that without further assumptions, the duality bundle is not necessarily a manifold.

For a spherical convex set $K \in \mathcal{K}(S^{n-1})$ and $M \subseteq K$, we define the *spherical duality bundle* via

$$N^S M \;:=\; \bigcup_{p\in M} \{p\} \times (\operatorname{relint}(N_p(K)) \cap S^{n-1}) \quad \subseteq \quad T^S M \;. \tag{3.9}$$

Note that by definition of the normal cone of $K$ at $p$ (cf. Definition 3.1.3) we have $N_p(K) \subseteq p^\perp$.

We may now define the families of stratified convex sets. As we will state Weyl's tube formulas in Section 4.3 in both the spherical and the euclidean situation, we will define stratified convex sets in both settings as well.

**Definition 3.3.9.** A convex set $K^e \in \mathcal{K}(\mathbb{R}^n)$, resp. $K \in \mathcal{K}(S^{n-1})$, is called *stratified* if it decomposes into a disjoint union $K^e = \dot\bigcup_{i=0}^k M_i^e$, resp. $K = \dot\bigcup_{i=0}^k M_i$, such that:

1. For all $0 \le i \le k$, $M_i^e$ is a smooth connected submanifold of $\mathbb{R}^n$, resp. $M_i$ is a smooth connected submanifold of $S^{n-1}$.

2. For all $0 \le i \le k$ the duality bundle $NM_i^e$, resp. the spherical duality bundle $N^S M_i$ is a smooth manifold.

A stratum $M_i^e$, resp. $M_i$, is called *essential* if $\dim NM_i^e = \dim T^\perp M_i^e = n$, resp. $\dim N^S M_i = \dim T^S M_i = n-2$. Otherwise it is called *negligible*.

**Example 3.3.10.** See Figure 3.5 for a simple nontrivial example of a stratified convex set $K^e \in \mathcal{K}(\mathbb{R}^3)$. The decomposition of $K^e$ is given by

$$K^e = M_0 \,\dot{\cup}\, M_1 \,\dot{\cup}\, M_2 \,\dot{\cup}\, M_3 \;,$$

where $M_0 := \operatorname{int}(K^e)$, and $M_1, M_2, M_3$ are as indicated in the picture. It is easily seen that all $M_i$ are essential.

If we take $x \in M_1$ and replace $M_1$ in the decomposition by $M_1 \setminus \{x\}$ and $\{x\}$, then $M_1 \setminus \{x\}$ is an essential piece, and $\{x\}$ is negligible.

We can also give an example where the condition that the duality bundles are smooth manifolds is violated: Imagine that you take the lenticular disc from Figure 3.5 and let it drop vertically on the plain ground. Imagine further that it hits the floor in a point $x \in M_1$ and slightly bends the edge inwards, so that the resulting normal cone at $x$ only consists of a single ray. Then the bent version of the piece $M_1$ is still a submanifold of the boundary, but the duality bundle fails to be a submanifold, as it does not have a constant dimension.

Usually, we will have $M_0 = \operatorname{int}(K)$, so that $\partial K = \dot{\bigcup}_{i=1}^k M_i$. Both polyhedral and smooth caps are special cases of stratified caps: The natural decomposition of a polyhedral convex set is given by the relative interiors of its faces, and we have already seen that the normal cone is constant on the relative interior of a face. Moreover, we will see that every smooth cap is a stratified cap in Section 4.1.2. So we may summarize

$$\begin{array}{l} \mathcal{K}^p \subset \\ \mathcal{K}^{\mathrm{sm}} \subset \end{array} \mathcal{K}^{\mathrm{str}}.$$

# Chapter 4

# Spherical tube formulas

In this chapter we will state and prove Weyl's euclidean and spherical tube formulas for stratified convex sets. We will also discuss the resulting euclidean and spherical intrinsic volumes. The intrinsic volumes of the semidefinite cone will be given in Section 4.4.1.

## 4.1 Preliminaries

This section is devoted to some differential geometric preliminaries that we need for the spherical tube formulas, as well as for some preliminary computations.

The first topic is the Weingarten map of manifolds which are embedded in euclidean space. We will have to recall some elementary concepts from Riemannian geometry, but as we are working in euclidean space these notions are all accessible without requiring much background in differential geometry. Most of the necessary material can be found for example in the introductory textbook [59]. We will mention further sources in the course of this section.

The second topic is about spherical caps which have a smooth boundary. We will show that the set of these smooth caps lies dense in the set of caps, and we will compute the Weingarten map for tubes around smooth caps.

The third topic is about integration on submanifolds of euclidean space. As a first application we will compute the volume of tubes around subspheres of the unit sphere. From this we will get a set of structural functions of the unit sphere, which will play a prominent role in Weyl's spherical tube formulas.

The fourth and last topic is about several sequences related to the binomial coefficient that will come up in subsequent computations. We will present them in a condensed form, and we will state and prove some properties and identities between them.

### 4.1.1 The Weingarten map for submanifolds of $\mathbb{R}^n$

In this section let $M \subseteq \mathbb{R}^n$ be a smooth manifold. A tangent vector field along a curve $c\colon \mathbb{R} \to M$ is defined to be a map $v\colon \mathbb{R} \to \mathbb{R}^n$ such that $v(t) \in T_{c(t)}M$ for all $t \in \mathbb{R}$. Note that for $p \in M$ we get an orthogonal decomposition $\mathbb{R}^n = T_p\mathbb{R}^n = T_pM \oplus T_p^\perp M$ (cf. Section 3.3).

A tangent vector field $v$ along a curve $c\colon \mathbb{R} \to M$ is said to be *parallel* iff $\dot{v}(t) \in T_{c(t)}^\perp M$ for every $t \in \mathbb{R}$. In this case, the tangent vector $\zeta := v(0) \in T_pM$, $p := c(0)$, is said to be parallel transported along the curve $c$. The following theorem about the existence and the uniqueness of parallel transport, and the angle-preserving property of parallel transport will prove useful for later computations.

**Theorem 4.1.1.** *Let $M \subseteq \mathbb{R}^n$ be a smooth manifold, let $c\colon \mathbb{R} \to M$ be a curve in $M$ with $p := c(0)$, and let $\zeta \in T_pM$. Then there exists a unique parallel vector field $v$ along $c$ such that $v(0) = \zeta$. Furthermore, if $\zeta_1, \zeta_2 \in T_pM$ and $v_1, v_2$ are the corresponding parallel vector fields along $c$, then $\langle v_1(t), v_2(t)\rangle = \langle \zeta_1, \zeta_2\rangle$ for all $t \in \mathbb{R}$. In particular, if $\zeta_1, \ldots, \zeta_d \in T_pM$ form an orthonormal basis of $T_pM$, and if $v_1, \ldots, v_d$ denote the corresponding parallel vector fields along $c$, then $v_1(t), \ldots, v_d(t)$ form an orthonormal basis of $T_{c(t)}M$ for all $t \in \mathbb{R}$.*

*Proof.* See [8, Thm. VII.3.12] (or [59, Ch. 8] for the hypersurface case $d = n-1$). □

**Remark 4.1.2.** In the situation of Theorem 4.1.1 one can also define the parallel transport of a subspace $\mathcal{Y} \subseteq T_pM$ along the curve $c$: Let $\xi_1, \ldots, \xi_k$ be a basis of $\mathcal{Y}$, and let $w_1, \ldots, w_k$ denote the corresponding vector fields along $c$. Then we say that $\mathcal{Y}_t := \mathrm{lin}\{w_1(t), \ldots, w_k(t)\} \subseteq T_{c(t)}M$ is the parallel transport of $\mathcal{Y}$ along $c$ at time $t$. For this definition to make sense it remains to show that $\mathcal{Y}_t$ is independent of the chosen basis $\xi_1, \ldots, \xi_k$. This is verified easily.

For a normal vector $\eta \in T_p^\perp M$ and a curve $c\colon \mathbb{R} \to M$ with $c(0) = p$, we say that $w\colon \mathbb{R} \to \mathbb{R}^n$ is a *normal extension* of $\eta$ along the curve $c$, iff $w(0) = \eta$ and $w(t) \in T_{c(t)}^\perp M$ for all $t \in \mathbb{R}$. In the following lemma we will show that we may always find a normal extension $w$ such that $\dot{w}(0) \in T_pM$, and we may additionally assume that $\|w(t)\| = 1$ for all $t$.

**Lemma 4.1.3.** *Let $M \subseteq \mathbb{R}^n$ be a smooth manifold, let $\eta \in T_p^\perp M$ be a normal vector, and let $c\colon \mathbb{R} \to M$ be a curve with $c(0) = p$. Furthermore, let $\Pi_t$ denote the orthogonal projection onto the normal space $T_{c(t)}^\perp$. Then the curve*

$$w\colon \mathbb{R} \to \mathbb{R}^n\ , \quad w(t) := \Pi_t(\eta)$$

*is a normal extension of $\eta$ along $c$, which satisfies $\dot{w}(0) \in T_pM$. Furthermore, if $\Pi_t(\eta) \neq 0$ for all $t \in \mathbb{R}$, then also the curve $w^\circ$ defined by*

$$w^\circ\colon \mathbb{R} \to \mathbb{R}^n\ , \quad w^\circ(t) := \|w(t)\|^{-1} \cdot w(t)\ , \tag{4.1}$$

*is a normal extension of $\eta$ along $c$ with $\dot{w^\circ}(0) \in T_pM$.*

*Proof.* The curves $w$ and $w^\circ$ are obviously normal extensions of $\eta$, so it remains to show the claim about $\dot{w}(0)$ and $\dot{w^\circ}(0)$. Let $\zeta_1, \ldots, \zeta_d \in T_pM$ form an orthonormal basis of $T_pM$, and let $v_1, \ldots, v_d\colon \mathbb{R} \to TM$ denote their parallel transports along $c$ (cf. Theorem 4.1.1). Furthermore, let $B\colon \mathbb{R} \to \mathbb{R}_\circ^{d\times n}$ be such that the $i$th row of $B(t)$ is given by $v_i(t)^T$. The orthogonal projection $\Pi_t$ is thus given by

$$\Pi_t(\eta) = \eta - B(t)^T B(t) \cdot \eta$$

(cf. Lemma 2.1.11). So we get

$$\dot{w}(0) = -\dot{B}(0)^T B(0) \cdot \eta - B(0)^T \dot{B}(0) \cdot \eta = -B(0)^T \dot{B}(0) \cdot \eta\ ,$$

as $\eta \in T_p^\perp M$, and the projection of $\dot{w}(0)$ on $T_p^\perp M$ is given by

$$\begin{aligned}
\Pi_0(\dot{w}(0)) &= -B(0)^T \dot{B}(0) \cdot \eta - B(0)^T B(0) \cdot (-B(0)^T \dot{B}(0) \cdot \eta) \\
&= -B(0)^T \dot{B}(0) \cdot \eta + B(0)^T \cdot \underbrace{B(0) B(0)^T}_{=I_d} \cdot \dot{B}(0) \cdot \eta \\
&= 0\ .
\end{aligned}$$

This shows that $\dot{w}(0) \in T_pM$.

As for the second claim, note that $\frac{d}{dt}\|w(t)\| = \frac{\langle \dot{w}(t), w(t)\rangle}{\|w(t)\|}$. This implies

$$\dot{w^\circ}(0) = -\frac{\langle \dot{w}(0), \eta\rangle}{\|\eta\|^3} \cdot \eta + \frac{1}{\|p\|} \cdot \dot{w}(0) = \frac{1}{\|p\|} \cdot \dot{w}(0) \in T_pM \ ,$$

as $\dot{w}(0) \in T_pM$ and thus $\langle \dot{w}(0), \eta\rangle = 0$. □

Let $p \in M$, $\zeta \in T_pM$, and $\eta \in T_p^\perp M$. It can be shown that if $c\colon \mathbb{R} \to M$ is a curve with $c(0) = p$ and $\dot{c}(0) = \zeta$, and if $w\colon \mathbb{R} \to \mathbb{R}^n$ is a normal extension of $\eta$ along $c$, then the orthogonal projection of $\dot{w}(0)$ onto $T_pM$ neither depends on the choice of the curve $c$ nor on the choice of the normal extension $w$ of $\eta$ (cf. for example [59, Ch. 14] for the hypersurface case, or [23, Ch. 6] for general Riemannian manifolds). It therefore makes sense to define the map

$$W_{p,\eta}\colon T_pM \to T_pM \ , \quad \zeta \mapsto -\Pi_{T_pM}(\dot{w}(0)) \ ,$$

where $w\colon \mathbb{R} \to \mathbb{R}^n$ is a normal extension of $\eta$ along a curve $c\colon \mathbb{R} \to M$ which satisfies $c(0) = p$ and $\dot{c}(0) = \zeta$, and $\Pi_{T_pM}$ denotes the orthogonal projection onto the tangent space $T_pM$. This map is called the *Weingarten map*.

It can be shown that $W_{p,\eta}$ is a symmetric linear map (cf. [23, Ch. 6]), so that it has $d := \dim M$ real eigenvalues $\kappa_1(p,\eta), \ldots, \kappa_d(p,\eta)$, which are called the *principal curvatures* of $M$ at $p$ in direction $\eta$. The corresponding eigenvectors are called *principal directions*.

When we are working with orientable hypersurfaces, i.e., with submanifolds of codimension 1, which are endowed with a (global) unit normal vector field $\nu\colon M \to T^\perp M$, $\nu(p) \in T_p^\perp M$, $\|\nu(p)\| = 1$, then we abbreviate

$$W_p := W_{p,\nu(p)} \ , \quad \kappa_i(p) := \kappa_i(p, \nu(p)) \ .$$

**Lemma 4.1.4.** *Let $M \subset \mathbb{R}^n$ be an orientable hypersurface, and let $\nu$ be a unit normal vector field of $M$. Then the Weingarten map of $M$ is given by*

$$W_p(\zeta) = -D_p\nu(\zeta)$$

*for all $p \in M$, $\zeta \in T_pM$, where $D_p\nu(\zeta)$ denotes the directional derivative of $\nu$ at $p$ in direction $\zeta$.*

*Proof.* For $p \in M$ and $\zeta \in T_pM$ let $c\colon \mathbb{R} \to M$ with $c(0) = p$ and $\dot{c}(0) = \zeta$. Furthermore, let $w_\nu(t) := \nu(c(t))$. By shrinking the domain of definition of $c$ if necessary, we may assume w.l.o.g. that $\langle w_\nu(t), w_\nu(0)\rangle > 0$ for all $t \in \mathbb{R}$. It follows that the projection of $\nu(p) = w_\nu(0)$ onto the normal space $T_{c(t)}^\perp M = \mathbb{R}\, w_\nu(t)$ is $\neq 0$. Let $w^\circ$ denote the unit normal extension of $\nu(p)$ along $c$ as defined in (4.1). As the normal space $T_{c(t)}^\perp M$ is one-dimensional for all $t$, we get $w_\nu(t) = w^\circ(t)$, and thus $\dot{w_\nu}(0) \in T_pM$ by Lemma 4.1.3. This implies $W_p(\zeta) = \dot{w_\nu}(0) = D_p\nu(\zeta)$. □

If $M \subseteq S^{n-1}$ is a submanifold of the unit sphere then $\hat{M} := \{\lambda p \mid \lambda > 0, p \in M\}$ is a conic manifold, i.e., $0 \notin \hat{M}$ and $x \in \hat{M}$ implies that $\lambda x \in \hat{M}$ for all $\lambda > 0$. The cone direction $x \in T_x\hat{M}$ is a principal direction with corresponding principal curvature 0. Note that we can find principal directions for the remaining principal curvatures in the orthogonal complement of the cone direction, i.e., in $x^\perp$. If $x = p \in S^{n-1}$ then $p^\perp$ is the tangent space of the unit sphere in $p$. We may thus conclude that the Weingarten map of a submanifold $M \subseteq S^{n-1}$ coincides with the Weingarten map of the corresponding conic manifold $\hat{M}$ except for the additional cone direction, which lies in the kernel of the Weingarten map of $\hat{M}$.

**Remark 4.1.5.** Let $M \subset \mathbb{R}^n$ be a smooth manifold such that $M \subseteq \partial K$ for some convex set $K$, and let $\eta \in T_p^\perp M$ for some $p \in M$. If additionally $-\eta \in N_p(K)$ (note that the normal vectors in $N_p(K)$ point outwards $K$), then the Weingarten map $W_{p,\eta}$ of $M$ at $p$ in direction $\eta$ is positive semidefinite. See for example [49, Sec. 2.5].

The elementary symmetric functions in the principal curvatures will play a prominent role in the tube formulas. Denoting $d := \dim M$, we define

$$\begin{aligned} \sigma_i(p,\eta) &:= i\text{th elementary symmetric function in } \kappa_1(p,\eta),\ldots,\kappa_d(p,\eta) \\ &= \sum_{1\le j_1<\ldots<j_i\le d} \kappa_{j_1}(p,\eta)\cdots\kappa_{j_i}(p,\eta)\;. \end{aligned} \tag{4.2}$$

For $i = d$ we have $\sigma_d(p,\eta) = \det W_{p,\eta}$, and this quantity is called the *Gaussian curvature* of $M$ at $p$ in direction $\eta$. If we have a hypersurface with a (global) unit normal vector field $\nu$, then we also write $\sigma_i(p)$ instead of $\sigma_i(p,\nu(p))$.

**Example 4.1.6.** As a simple example let us compute the Weingarten map of the boundary of a circular cap. More precisely, let $z \in S^{n-1}$ and $K := B(z,\beta) = \{p \in S^{n-1} \mid d(z,p) \le \beta\}$ the circular cap around $z$ of radius $\beta \in (0,\pi)$. We define $M := \partial K$, which is a hypersurface of $S^{n-1}$. This hypersurface has a global unit normal vector field $\nu$ given by the unit normal vectors pointing inwards $K$. To compute the principal curvatures $\kappa_1(p),\ldots,\kappa_{n-2}(p)$ at $p \in M$, let $\zeta \in T_pM$, $\|\zeta\| = 1$. By making a change of basis in $\mathbb{R}^n$ we may assume without loss of generality that $z = e_1$, $p = \cos(\beta)\,e_1 + \sin(\beta)\,e_2$ and $\zeta = e_3$, where $e_i \in \mathbb{R}^n$ denotes the $i$th canonical basis vector. By this choice of basis, the normal vector $\nu(p)$ is given by $\nu(p) = \sin(\beta)\,e_1 - \cos(\beta)\,e_2$. Consider the rotation

$$Q(\rho) := \begin{pmatrix} 1 & & & & & \\ & \cos(\rho) & -\sin(\rho) & & & \\ & \sin(\rho) & \cos(\rho) & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}.$$

Then $c\colon \mathbb{R} \to M$, $t \mapsto Q\left(\frac{t}{\sin\beta}\right)\cdot p$, describes a curve in $M$ through $p$ with

$$\tfrac{dc}{dt}(0) = \tfrac{1}{\sin\beta}\cdot \dot{Q}(0)\,p = e_3 = \zeta\;.$$

A normal extension of $\nu(p)$ along $c$ is given by

$$\begin{aligned} w(t) &:= Q\left(\tfrac{t}{\sin\beta}\right)\cdot \nu(p) \\ &= (\sin(\beta), -\cos(\beta)\cdot\cos(\tfrac{t}{\sin\beta}), -\cos(\beta)\cdot\sin(\tfrac{t}{\sin\beta}), 0,\ldots,0)^T\;, \end{aligned}$$

and we get $\frac{dw}{dt}(0) = -\cot(\beta)\cdot e_3 = -\cot(\beta)\cdot\zeta$, which implies (cf. Lemma 4.1.4) $W_p(\zeta) = \cot(\beta)\cdot\zeta$. Since this holds for any $\zeta \in T_pM$ with $\|\zeta\| = 1$, we get

$$\kappa_1(p) = \ldots = \kappa_{n-2}(p) = \cot(\beta) \quad \text{for all } p \in M\;,$$

and $\sigma_i(p) = \binom{n-2}{i}\cdot\cot(\beta)^i$.

For $\beta = \frac{\pi}{2}$ we have $M \in \mathcal{S}^{n-2}(S^{n-1})$ and, as $\cos(\frac{\pi}{2}) = 0$, the Weingarten map of $M$ is the zero map. This also holds for subspheres of higher codimension, which is seen in the following way. For $S \in \mathcal{S}^k(S^{n-1})$ and $p \in S$, let $\eta \in T_p^\perp S$ and $L := \mathrm{lin}\{S,\eta\}$. The linear subspace $L$ has dimension $k+2$, and by replacing the euclidean space $\mathbb{R}^n$ by $L$, and $S^{n-1}$ by $L \cap S^{n-1}$, we can reduce the general case to the codimension-1 case. In summary, the Weingarten map of a subsphere of the unit sphere is the zero map, i.e., a subsphere has no curvature relative to the surrounding unit sphere.

In the following lemma we compute the tangent spaces of the normal bundle $T^\perp M \subset \mathbb{R}^n \times \mathbb{R}^n$, where $M \subseteq \mathbb{R}^n$ is a $d$-dimensional manifold, and the spherical normal bundle $T^S M \subset S^{n-1} \times S^{n-1}$, if $M$ additionally lies in $S^{n-1}$ (cf. (3.6)).

**Lemma 4.1.7.** *Let $M \subseteq \mathbb{R}^n$ be a smooth $d$-dimensional manifold. Furthermore, let $(p,\eta) \in T^\perp M$, and let $W_{p,\eta}\colon T_pM \to T_pM$ denote the Weingarten map of $M$ at $p$ in direction $\eta$. Then the tangent space of $T^\perp M$ at $(p,\eta)$ is given by*

$$T_{(p,\eta)}T^\perp M \;=\; \{(\zeta, -W_{p,\eta}(\zeta)) \mid \zeta \in T_pM\} \;\oplus\; \{0\} \times T_p^\perp M \;. \tag{4.3}$$

*If additionally $M \subseteq S^{n-1}$, and if $(p,\eta) \in T^S M$, i.e., $\|\eta\| = 1$ and $\langle p,\eta\rangle = 0$, then the tangent space of $T^S M$ at $(p,\eta)$ is given by*

$$T_{(p,\eta)}T^S M \;=\; \{(\zeta, -W_{p,\eta}(\zeta)) \mid \zeta \in T_pM\} \;\oplus\; \{0\} \times \left(T_p^\perp M \cap p^\perp \cap \eta^\perp\right) \;. \tag{4.4}$$

Note that the decompositions in (4.3) and (4.4) are orthogonal decompositions.

*Proof.* The right-hand side of (4.3) is a $n$-dimensional subspace of $\mathbb{R}^n \times \mathbb{R}^n$. In order to show the equality in (4.3) it thus suffices to show that the right-hand lies in the tangent space $T_{(p,\eta)}T^\perp M$.

Let $\zeta \in T_pM$ and let $c\colon \mathbb{R} \to M$ be a curve such that $c(0) = p$ and $\dot{c}(0) = \zeta$. Furthermore, let $w_1\colon \mathbb{R} \to \mathbb{R}^n$ be a normal extension of $\eta$ along $c$, such that $\dot{w}_1(0) \in T_pM$ (cf. Lemma 4.1.3). The composite curve $t \mapsto (c(t), w_1(t))$ then describes a curve in the normal bundle $T^\perp M$ with $(c, w_1)(0) = (p,\eta)$. Furthermore, we have

$$\tfrac{d}{dt}(c(t), w_1(t))(0) = (\dot{c}(0), \dot{w}_1(0)) = (\zeta, -W_{p,\eta}(\zeta)) \;.$$

This shows that $(\zeta, -W_{p,\eta}(\zeta) \in T_{(p,\eta)}T^\perp M$.

As for the second summand of the right-hand side in (4.3), note that $T_p^\perp M$ is a linear space so that $T_p^\perp M$ coincides with the tangent space of $T_p^\perp M$ at $\eta$. If $w_2\colon \mathbb{R} \to T_p^\perp M$ is a curve with $w_2(0) = \eta$, then we have $(p, w_2(t)) \in T^\perp M$ for all $t$ and $(p, w_2(0)) = (p,\eta)$. Therefore, we have

$$\tfrac{d}{dt}(p, w_2(t))(0) = (0, \dot{w}_2(0)) \in T_{(p,\eta)}T^\perp M \;.$$

This shows that $\{0\} \times T_p^\perp M \subset T_{(p,\eta)}T^\perp M$, and thus finishes the proof of the equality in (4.3).

As for the claim about the spherical normal bundle we now additionally assume $\|\eta\| = 1$ and $\langle p,\eta\rangle = 0$. Furthermore, let $\hat{M} := \{\lambda p \mid \lambda > 0, p \in M\}$ denote the conic manifold corresponding to $M$. Note that

$$T_q^\perp \hat{M} = T_q^\perp M \cap q^\perp \quad \text{for all } q \in M. \tag{4.5}$$

As above, it suffices to show that the right-hand side of (4.4) lies in the tangent space of $T^S M$ in $(p,\eta)$.

For $\zeta \in T_pM$ we consider a curve $c\colon \mathbb{R} \to M$ such that $c(0) = p$ and $\dot{c}(0) = \zeta$. The normal direction $\eta \in T_p^\perp M$ also lies in $T_p^\perp \hat{M}$, as $\langle p,\eta\rangle = 0$. Let $w_1\colon \mathbb{R} \to \mathbb{R}^n$ be a normal extension of $\eta$ along $c$ w.r.t. $\hat{M}$, i.e., $w_1(0) = \eta$ and $w_1(t) \in T_{c(t)}^\perp \hat{M}$. By Lemma 4.1.3 we may choose $w_1$ such that $\dot{w}_1(0) \in T_p\hat{M}$. Furthermore, if the domain of the curve $c$ is shrinked to a sufficiently small interval around 0, then we may choose $w_1$ such that additionally $\|w_1(t)\| = 1$ for all $t$ (cf. Lemma 4.1.3). The resulting composite curve $t \mapsto (c(t), w_1(t))$ then describes a curve in the spherical normal bundle $T^S M$ by (4.5). Denoting by $\hat{W}_{p,\eta}$ the Weingarten map of $\hat{M}$ at $p$ in direction $\eta$, we get

$$\tfrac{d}{dt}(c(t),w_1(t))(0) = (\zeta, \dot{w}_1(0)) = (\zeta, -\hat{W}_{p,\eta}(\zeta)) = (\zeta, -W_{p,\eta}(\zeta)) \ .$$

This shows that $(\zeta, -W_{p,\eta}(\zeta)) \in T_{(p,\eta)}T^S M$.

As for the second summand of the right-hand side in (4.4), note that $T_p^\perp M \cap p^\perp \cap \eta^\perp$ is the tangent space of $T_p^\perp M \cap p^\perp \cap S^{n-1}$ at $\eta$. So the only change of the proof of (4.3) consists in restricting the co-domain of the curve $w_2$ and to consider $w_2 \colon \mathbb{R} \to T_p^\perp M \cap p^\perp \cap S^{n-1}$. The same reasoning as above shows the equality in (4.4). □

We finish this section with a lemma that will be an essential part of the computations in Section 6.3.

**Lemma 4.1.8.** *Let $M \subset S^{n-1}$ be a hypersurface of the unit sphere, let $\hat{M} := \{\lambda\, p \mid \lambda > 0, p \in M\}$ be the corresponding conic hypersurface of $\mathbb{R}^n$, and let $\nu$ be a unit normal field of $\hat{M}$. Furthermore, let $c\colon \mathbb{R} \to M$ be a smooth curve, let $\zeta_1, \ldots, \zeta_{n-2}$ be an orthonormal basis of $T_pM$, where $p := c(0)$, and let $v_i \colon \mathbb{R} \to \mathbb{R}^n$ be the parallel transport w.r.t. $M$ of $\zeta_i$ along $c$, $i = 1, \ldots, n-2$. Then $Q\colon \mathbb{R} \to \mathbb{R}^{n\times n}$, given by*

$$Q(t) := \begin{pmatrix} c(t) & v_1(t) & \cdots & v_{n-2}(t) & \nu(c(t)) \end{pmatrix} \ ,$$

*satisfies $Q(t) \in O(n)$ for all $t \in \mathbb{R}$. Furthermore,*

$$\dot{Q}(0) = Q(0) \cdot \left( \begin{array}{c|ccc|c} 0 & -a_1 & \cdots & -a_{n-2} & 0 \\ \hline a_1 & 0 & \cdots & 0 & -b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-2} & 0 & \cdots & 0 & -b_{n-2} \\ \hline 0 & b_1 & \cdots & b_{n-2} & 0 \end{array} \right) ,$$

*with*

$$\dot{c}(0) = \sum_{i=1}^{n-2} a_i \cdot \zeta_i \quad \text{and} \quad W_p(\dot{c}(0)) = \sum_{i=1}^{n-2} b_i \cdot \zeta_i \ ,$$

*where $W_p$ denotes the Weingarten map of $M$ at $p$.*

*Proof.* By Theorem 4.1.1 we have that $v_1(t), \ldots, v_{n-2}(t)$ is an orthonormal basis of $T_{c(t)}M$. Furthermore, as $T_qS^{n-1} = q^\perp$, we have $\langle c(t), v_i(t)\rangle = 0$, and as $\nu(c(t)) \in T^\perp_{c(t)}\hat{M}$, we have $\langle v_i(t), \nu(c(t))\rangle = 0$ for all $t \in \mathbb{R}$. Finally, we have $\langle c(t), \nu(c(t))\rangle = 0$, as $\hat{M}$ is a conic manifold, which implies $c(t) \in T_{c(t)}\hat{M}$. This shows that $Q(t) \in O(n)$ for all $t \in \mathbb{R}$.

As $Q$ describes a curve in $O(n)$, it follows that $\frac{dQ}{d\rho}(0) = Q(0) \cdot U$ with skew-symmetric $U \in \mathbb{R}^{n\times n}$, i.e., $U^T = -U$ (cf. Section 5.2). As $\dot{c}(0) = \sum_{i=1}^{n-2} a_i \cdot \zeta_i$, the first column of $U$ is given by $(0, a_1, \ldots, a_{n-2}, 0)^T$, and by skew-symmetry this also gives us the first row. The zero matrix in the middle follows from the fact that the $\zeta_i$ are parallel transported along $c$. Finally, the last column of $U$ follows from Lemma 4.1.4, and the last row follows again from skew-symmetry. □

### 4.1.2 Smooth caps

In this section we will consider convex sets with smooth boundaries. We will treat the question of approximation by these sets and we will compute the curvature of tubes around them.

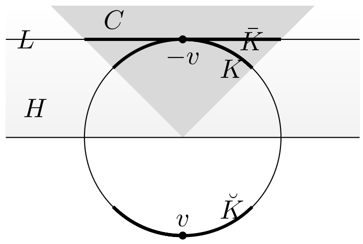



Figure 4.1: Illustration for the proof of Proposition 4.1.10.

**Definition 4.1.9.** For $n \geq 2$ the set of *smooth convex bodies* is defined as

$$\mathcal{K}^{\mathrm{sm}}(\mathbb{R}^n) := \left\{ K \in \mathcal{K}(\mathbb{R}^n) \,\middle|\, \begin{array}{c} \mathrm{int}(K) \neq \emptyset \text{ and } \partial K \text{ is a smooth hypersurface} \\ \text{in } \mathbb{R}^n \text{ with nowhere vanishing Gaussian curvature} \end{array} \right\} .$$

Analogously, for $n \geq 3$ the set of *smooth caps* is defined as

$$\mathcal{K}^{\mathrm{sm}}(S^{n-1}) := \left\{ K \in \mathcal{K}(S^{n-1}) \,\middle|\, \begin{array}{c} K \in \mathcal{K}^r(S^{n-1}) \text{ and } \partial K \text{ is a smooth} \\ \text{hypersurface in } S^{n-1} \text{ with nowhere} \\ \text{vanishing Gaussian curvature} \end{array} \right\}$$

(cf. Definition 3.3.5 for the definition of $\mathcal{K}^r(S^{n-1})$).

For $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, $M := \partial K$, we denote by $\nu\colon M \to \mathbb{R}^n$ the unit normal field pointing inwards the cap $K$. This is well-defined, by the following arguments. The normal cone is contained in the 1-dimensional normal space, i.e., $N_p(K) \subseteq T_p^\perp M$, for $p \in M$. For $\eta \in N_p(K) \setminus \{0\}$ we have $\langle \eta, q\rangle \leq 0$ for all $q \in K$. Moreover, as the interior of $K$ is by definition non-empty, there exists $q_0 \in K$ such that $\langle \eta, q\rangle < 0$. This implies that $N_p(K)$ is a half-line, and there exists a unique element $\eta \in N_p(K)$ of length 1. The direction $\nu(p)$ is defined to be the vector $-\eta$, so that $\langle \nu(p), q\rangle \geq 0$ for all $q \in K$.

**Proposition 4.1.10.** *The set of smooth convex bodies lies dense in the set of convex bodies and the set of smooth convex caps lies dense in the set of convex caps, i.e.,*

$$\mathcal{K}^{\mathrm{sm}}(\mathbb{R}^n) \overset{d}{\subset} \mathcal{K}(\mathbb{R}^n) , \quad \mathcal{K}^{\mathrm{sm}}(S^{n-1}) \overset{d}{\subset} \mathcal{K}^c(S^{n-1}) .$$

*Proof.* The euclidean statement is originally due to Minkowski (cf. [7, §6]). See [48] for a more recent proof. We will only deduce the spherical from the euclidean statement.

First of all, every $K \in \mathcal{K}^c(S^{n-1})$ can be approximated by a sequence $(K_i)_i$ in $\mathcal{K}^c(S^{n-1})$ such that all $K_i$ lie in a fixed open half-space. This is seen in the following way. Let $v \in \breve{K}$ such that $-v \notin \breve{K}$ (cf. Remark 3.2.6), and let $H_i := B(-v, \rho_i)$ the circular cap of radius $\rho_i$ around $-v$, with $\rho_i := \frac{\pi}{2} - \frac{1}{i}$. Then we define $K_i := H_i \cap K$. The fact that $-v \notin \breve{K}$ implies that $K \not\subset v^\perp$, and thus $K_i \neq \emptyset$ for large enough $i$. Furthermore, being the intersection of two convex sets, $K_i$ is again convex. So we get $K_i \in \mathcal{K}^c(S^{n-1})$ for all large enough $i$, and every $K_i$ lies in the open half-space $\{x \mid \langle x, v\rangle < 0\}$. As $K_i \subseteq K$ and $K \subseteq \mathcal{T}(K_i, \frac{1}{i})$, we also have $K_i \to K$ in the Hausdorff metric.

So we may assume w.l.o.g. that $K \in \mathcal{K}^c(S^{n-1})$ lies in an open half-space $H = \{x \mid \langle x, v\rangle < 0\}$. Let $L := \{x \mid \langle x, v\rangle = -1\}$ denote the affine hyperplane through $-v$

that is orthogonal to $v$, and let $\bar{K} := C \cap L$, where $C := \operatorname{cone}(K)$. See Figure 4.1 for a small display. If $\varphi\colon L \to \mathbb{R}^{n-1}$ denotes a linear isometry, then we get that $K^e := \varphi(\bar{K})$ is a convex body in $\mathbb{R}^{n-1}$, i.e., $K^e \in \mathcal{K}(\mathbb{R}^{n-1})$. By the euclidean statement of the proposition we can find smooth convex bodies $K_i^e \in \mathcal{K}^{\mathrm{sm}}(\mathbb{R}^{n-1})$ such that $K_i^e \to K^e$. We define

$$\bar{K}_i := \varphi^{-1}(K_i^e)\ , \quad C_i := \operatorname{cone}(\bar{K}_i)\ , \quad K_i := C_i \cap S^{n-1}\ .$$

Then we have $K_i \in \mathcal{K}^c(S^{n-1})$, and the boundary of $K_i$ is smooth. From the approximation property $K_i^e \to K^e$ it follows that $K_i \to K$, which is verified easily. It remains to show that the Gaussian curvature of $\partial K_i$ does not vanish.

If the Gaussian curvature of $\partial K_i$ vanishes in $p \in \partial K_i$, then the Weingarten map of $\partial K_i$ in $p$ has a nontrivial kernel. This implies that the Weingarten map of $\partial C_i \setminus \{0\}$ in $p$ has a kernel of dimension at least 2, as the cone direction adds a dimension to the kernel of the Weingarten map. It follows that the Weingarten map of $\partial \bar{K}_i$ at $x$, where $\{x\} = L \cap \mathbb{R} \cdot p$, has a nontrivial kernel, as the dimension can drop at most by 1. But this means that the Gaussian curvature of $\bar{K}_i$ vanishes at $x$, which contradicts the assumption on $K_i^e$ and thus finishes the proof. □

The following proposition summarizes the most important properties of smooth caps.

**Proposition 4.1.11.** 1. *If $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ then also $\breve{K} \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$.*

2. *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, and let $p \in \partial K$. Then we have $\nu(p)^\perp \cap K = \{p\}$ and the map $p \mapsto -\nu(p)$ describes a diffeomorphism between the boundary of $K$ and the boundary of its dual $\breve{K}$.*

3. *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, $p \in \partial K$, and let $\kappa_1, \ldots, \kappa_{n-2}$ denote the principal curvatures of $\partial K$ in $p$. Then the principal curvatures of $\breve{K}$ at $-\nu(p)$ are given by $\kappa_1^{-1}, \ldots, \kappa_{n-2}^{-1}$.*

We will deduce this proposition from the following lemma.

**Lemma 4.1.12.** *Let $M \subset S^{n-1}$ be a smooth hypersurface of $S^{n-1}$ with unit normal vector field $\nu\colon M \to T^\perp M$ and with principal curvatures $\kappa_1(p), \ldots, \kappa_{n-2}(p)$ at $p \in M$. Furthermore, for $\alpha \in \mathbb{R}$ let $f_\alpha\colon M \to S^{n-1}$ be given by*

$$f_\alpha(p) := \cos(\alpha) \cdot p - \sin(\alpha) \cdot \nu(p)\ ,$$

*and let the image of $f_\alpha$ be denoted by $M_\alpha$. If $f_\alpha$ is injective, and if for all $p \in M$*

$$\prod_{i=1}^{n-2} (\cos(\alpha) + \sin(\alpha) \cdot \kappa_i(p)) \neq 0\ ,$$

*then $M_\alpha$ is a smooth hypersurface of $S^{n-1}$. In this case, the tangent space of $M_\alpha$ at $f_\alpha(p)$ coincides with the tangent space of $M$ at $p$, and the map $\nu_\alpha\colon M_\alpha \to S^{n-1}$, $\nu_\alpha(f_\alpha(p)) := \cos(\alpha) \cdot \nu(p) + \sin(\alpha) \cdot p$, is a unit normal field of $M_\alpha$. Furthermore, if $\zeta \in T_pM$ is a principal direction of $M$ at $p$ with principal curvature $\kappa$, then $\zeta$ is a principal direction of $M_\alpha$ at $f_\alpha(p)$ with principal curvature (w.r.t. the normal field $\nu_\alpha$)*

$$\frac{\cos(\alpha) \cdot \kappa - \sin(\alpha)}{\cos(\alpha) + \sin(\alpha) \cdot \kappa}\ .$$

*Proof.* In order to show that $M_\alpha$ is a smooth manifold we use the inverse function theorem. As the map $f_\alpha$ is injective, it suffices to show that the derivative $D_p f_\alpha$ has full rank for all $p \in M$. Let $p \in M$ and let $\zeta \in T_pM$ be a principal direction w.r.t. the principal curvature $\kappa$. Recall that by Lemma 4.1.4 we have $D_p\nu(\zeta) = -W_p(\zeta)$, where $W_p$ denotes the Weingarten map of $M$ in $p$. Thus, using the linearity of the directional derivative, we get

$$\begin{aligned} D_p f_\alpha(\zeta) &= \cos(\alpha) \cdot D_p \operatorname{id}(\zeta) - \sin(\alpha) \cdot D_p\nu(\zeta) \\ &= \cos(\alpha) \cdot \zeta - \sin(\alpha) \cdot (-W_p(\zeta)) \\ &= (\cos(\alpha) + \sin(\alpha) \cdot \kappa) \cdot \zeta \ . \end{aligned} \tag{4.6}$$

Since $\cos(\alpha) + \sin(\alpha) \cdot \kappa \neq 0$, and as this holds for all principal curvatures at $p$, we get that $W_p$ has full rank, and thus, by the inverse function theorem, that $M_\alpha$ is a smooth hypersurface of $S^{n-1}$.

The fact that the tangent space of $M_\alpha$ at $f_\alpha(p)$ coincides with $T_pM$ follows from (4.6) and the fact that one can chose a basis of $T_pM$ consisting of principal directions. The fact that the normal space of $M_\alpha$ at $f_\alpha(p)$ is spanned by $\nu_\alpha(p)$ follows from the properties $\langle \nu_\alpha(p), f_\alpha(p)\rangle = 0$ and $\langle \nu_\alpha(p), \zeta\rangle = 0$ for all $\zeta \in T_pM$. Hence, $\nu_\alpha$ defines a unit normal field on $M_\alpha$.

As for the claim about the principal curvatures of $M_\alpha$, let again $\zeta \in T_pM$ be a principal direction w.r.t. the principal curvature $\kappa$, and let $c_\alpha(t) := f_\alpha(c(t))$, where $c\colon \mathbb{R} \to M$ is a curve with $c(0) = p$ and $\dot{c}(0) = \zeta$. If we denote $w(t) := \nu(c(t))$, we have $\dot{w}(0) = D_p\nu(\zeta) = -W_p(\zeta) = -\kappa\,\zeta$. Furthermore, (4.6) yields $\dot{c_\alpha}(0) = (\cos(\alpha) + \sin(\alpha) \cdot \kappa) \cdot \zeta$. If we denote

$$w_\alpha(t) := \nu_\alpha(c_\alpha(t)) = \cos(\alpha) \cdot w(t) + \sin(\alpha) \cdot c(t) \ ,$$

then we get

$$\begin{aligned} -\dot{w_\alpha}(0) &= -\cos(\alpha) \cdot \dot{w}(0) - \sin(\alpha) \cdot \dot{c}(0) \\ &= \cos(\alpha) \cdot \kappa \cdot \zeta - \sin(\alpha) \cdot \zeta \\ &= \frac{\cos(\alpha) \cdot \kappa - \sin(\alpha)}{\cos(\alpha) + \sin(\alpha) \cdot \kappa} \cdot \dot{c_\alpha}(0) \ . \end{aligned}$$

This shows the claim about the principal curvatures of $M_\alpha$ and thus finishes the proof. □

*Proof of Proposition 4.1.11.* To ease the notation, let $M := \partial K$. The map

$$f\colon M \to S^{n-1} \ , \quad p \mapsto -\nu(p) \ ,$$

is injective by the following arguments. Assume that $p_1, p_2 \in M$, $p_1 \neq p_2$, with $\nu(p_1) = \nu(p_2) =: \eta$. As $K$ lies in an open half-space we also have $p_1 \neq -p_2$, and there is a unique geodesic arc $\operatorname{geod}(p_1, p_2)$ between $p_1$ and $p_2$. This geodesic arc is contained in $K$ by the convexity of $K$. Furthermore, $K$ lies in the half-space $\{x \in \mathbb{R}^n \mid \langle x, \eta\rangle \geq 0\}$. The geodesic arc $\operatorname{geod}(p_1, p_2) \subset \operatorname{lin}\{p_1, p_2\}$ lies in the hyperplane $\eta^\perp$. This implies $\operatorname{geod}(p_1, p_2) \subset \partial K = M$. Moreover, $\nu(p) = \eta$ for all $p \in \operatorname{geod}(p_1, p_2)$. This implies that the Gaussian curvature is zero along this arc, which contradicts the assumption $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$.

The image of $f$ lies in $\breve{K}$, as $-\nu(p) \in N_p(K) \subseteq \breve{C}$, where $C := \operatorname{cone}(K)$ (cf. Proposition 3.1.5). Moreover, we have $N_p^S(K) = p^\perp \cap \breve{K} = \{-\nu(p)\}$, so that $f(p)$ lies in the boundary $\partial\breve{K} =: \breve{M}$. Therefore, reducing the co-domain of $f$ and identifying the resulting map with $f$, we have an injective map $f\colon M \to \breve{M}$. This map is also surjective by the following arguments. The dual of $\breve{C}$ is again the primal cone $C$, i.e., $(\breve{C})^{\breve{}} = C$ (cf. [47, Cor. 11.7.2]). Therefore, for $\eta \in \breve{M} = \partial\breve{K}$ there exists $p \in N^S_\eta(\breve{K}) \subseteq K$. As $\langle \eta, p\rangle = 0$, we have $\eta = -\nu(p) = f(p)$. This shows that the map $f$ is surjective.

Applying Lemma 4.1.12 with $\alpha = \frac{\pi}{2}$ (note that $f = f_{\pi/2}$), we get that $\breve{M}$ is a smooth hypersurface of $S^{n-1}$. This proves part (2) of the claim.

Furthermore, Lemma 4.1.12 implies that the principal curvatures of $\breve{M}$ at $f(p)$ are given by $-\kappa_1^{-1}(p), \dots, -\kappa_{n-2}^{-1}(p)$, where the corresponding unit normal field is given by $f(p) \mapsto p$. This normal field points outwards the cap $\breve{K}$, so that the principal curvatures at $f(p)$ w.r.t. the unit normal field pointing inwards $\breve{K}$ are given by $\kappa_1^{-1}(p), \dots, \kappa_{n-2}^{-1}(p)$. This proves part (3) of the claim.

The first part of the claim follows from $K \in \mathcal{K}^r(S^{n-1}) \iff \breve{K} \in \mathcal{K}^r(S^{n-1})$ (cf. Section 3.3), and part (3). □

We finish this section with another corollary from Lemma 4.1.12 about the maximum radius $\alpha_0$ for which the tube $\mathcal{T}(K, \alpha_0)$ around a smooth cap $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ is still convex. The idea is to use the fact that the boundary of a convex cap, if it is a smooth submanifold of $S^{n-1}$, always has a positive semidefinite Weingarten map (cf. Remark 4.1.5). From Lemma 4.1.12 we get a formula for the Weingarten map of the boundary of a tube. So as soon as this fails to be positive semidefinite, the tube cannot be convex.

**Corollary 4.1.13.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and let $\alpha_0 := \sup\{\alpha \mid \mathcal{T}(K,\alpha) \in \mathcal{K}(S^{n-1})\}$. Then*

$$\alpha_0 \le \arctan(\min\{\kappa_{\min}(p) \mid p \in M\}) , \tag{4.7}$$

*where $M := \partial K$, and $\kappa_{\min}(p)$ denotes the minimum principal curvature of $M$ at $p$.*

*Proof.* We first show that for $0 \le \alpha < \frac{\pi}{2}$ the map $f_\alpha\colon M \to S^{n-1}$ given by $f_\alpha(p) = \cos(\alpha)\cdot p - \sin(\alpha)\cdot\nu(p)$ (cf. Lemma 4.1.12) is a bijection between $M$ and the boundary of $\mathcal{T}(K,\alpha)$. If $C := \operatorname{cone}(K)$ denotes the cone defined by $K$, then the (euclidean) projection map $\Pi_C\colon \mathbb{R}^n \to C$ satisfies

$$\Pi_C^{-1}(p) = p - \mathbb{R}\,\nu(p)$$

for all $p \in M \subset \partial C$ (cf. Section 3.1). This implies that for $0 \le \alpha < \frac{\pi}{2}$ we have $\Pi_C(f_\alpha(p)) = \cos(\alpha)\cdot p$, and thus $d(f_\alpha(p), K) = \alpha$, i.e., $f_\alpha(p) \in \partial\mathcal{T}(K,\alpha)$. Moreover, any point $p_\alpha \in \partial\mathcal{T}(K,\alpha)$ is of the form $p_\alpha = f_\alpha(p)$, and $f_\alpha(p) = f_\alpha(q)$ implies $p = q$, since $p = \cos(\alpha)^{-1} \cdot \Pi_C(f_\alpha(p))$.

Lemma 4.1.12 thus implies that $M_\alpha := \mathcal{T}(K,\alpha)$ is a smooth hypersurface if

$$\cos(\alpha) + \sin(\alpha)\cdot\kappa_i(p) \neq 0 \quad \text{for all } 1 \le i \le n-2,\ p \in M,$$

where $\kappa_1(p), \dots, \kappa_{n-2}(p)$ denote the principal curvatures of $M$ at $p$. In this case, the map $f_\alpha\colon M \to M_\alpha$ is a diffeomorphism and the principal curvatures of $M_\alpha$ at $f_\alpha(p)$ are given by $(\cos(\alpha)\cdot\kappa_i(p) - \sin(\alpha))/(\cos(\alpha) + \sin(\alpha)\cdot\kappa_i(p))$, $i = 1, \dots, n-2$.

As $M$ is the boundary of the smooth cap $K$, the Weingarten map of $M$ is positive definite at each point (cf. Remark 4.1.5 and Definition 4.1.9). This implies that $\cos(\alpha) + \sin(\alpha)\cdot\kappa_i(p) > 0$ for all $i = 1, \dots, n-2$, $p \in M$, $0 \le \alpha < \frac{\pi}{2}$. In particular, $M_\alpha$ is a smooth hypersurface for all $0 \le \alpha < \frac{\pi}{2}$.

If $\mathcal{T}(K,\alpha)$ is convex, then the Weingarten map of $M_\alpha$ is positive semidefinite (cf. Remark 4.1.5), i.e.,

$$\frac{\cos(\alpha)\cdot\kappa_i(p) - \sin(\alpha)}{\cos(\alpha) + \sin(\alpha)\cdot\kappa_i(p)} \ge 0 \quad \text{for all } p \in M,\ i = 1, \dots, n-2.$$

This is equivalent to $\alpha \leq \arctan(\min\{\kappa_{\min}(p) \mid p \in M\}) < \frac{\pi}{2}$. □

**Remark 4.1.14.** We believe that the inequality for $\alpha_0$ in (4.7) is an equality. The missing argument for a proof of this equality is that the positive semidefiniteness of the Weingarten map implies convexity. Theorems of this kind are known for the euclidean case (cf. [49, Sec. 5]). The transition to the spherical setting should be doable as in the proof of Proposition 4.1.10.

### 4.1.3 Integration on submanifolds of $\mathbb{R}^n$

In this section we will describe integration on submanifolds of euclidean space. Later, in Section 5.1, we will describe integration on general Riemannian manifolds, but as we need to know about integration on submanifolds of $\mathbb{R}^n$ for Weyl's tube formulas, we will treat this special case already at this point. The basic facts that we state in the following paragraphs can be found for example in [52, Ch. 3]. As a first application we will compute the volume of tubes around subspheres of the unit sphere.

Let $M \subseteq \mathbb{R}^n$ be a smooth submanifold of dimension $d$, and let $\varphi\colon \mathbb{R}^d \to M$ be a smooth parametrization of an open subset $U := \mathrm{im}(\varphi) \subseteq M$, i.e., $\varphi$ is a smooth diffeomorphism. Denoting by $D_x\varphi\colon \mathbb{R}^d \to T_{\varphi(x)}M$ the derivative of $\varphi$ in $x \in \mathbb{R}^d$, the integral of an integrable function $f\colon M \to \mathbb{R}$ over $U$ is defined via

$$\int_{y\in U} f(y)\, dM := \int_{x\in\mathbb{R}^d} f(\varphi(x)) \cdot |\det(D_x\varphi)|\, dx\ . \tag{4.8}$$

It is essential that the definition in (4.8) is independent of the chosen parametrization $\varphi$ of $M$, i.e.,

$$\int_{x\in\mathbb{R}^d} f(\varphi(x)) \cdot |\det(D_x\varphi)|\, dx = \int_{x\in\mathbb{R}^d} f(\psi(x)) \cdot |\det(D_x\psi)|\, dx\ ,$$

if $\psi\colon \mathbb{R}^d \to M$ is another smooth parametrization with $\mathrm{im}(\psi) = U$. This follows from the transformation theorem (cf. for example [52, Thm. 3-13]).

More generally, let $(U_i)_i$ be a sequence of open subsets of $M$ which cover $M$, and let $(\varphi_i)_i$ be a partition of unity subordinate to the open cover $(U_i)_i$. This means that the $\varphi_i\colon M \to [0,1]$ are smooth functions with $\varphi_i(p) = 0$ for $p \notin U_i$, such that for every $p \in M$ there exists a neighborhood $U$ of $p$, such that all but a finite number of $\varphi_i$ are zero on $U$, and $\sum_{i=1}^{\infty} \varphi_i(p) = 1$. If $f\colon M \to \mathbb{R}$ is an integrable function, then the integral of $f$ over $M$ is defined via

$$\int_M f\, dM := \sum_{i=1}^{\infty} \int_{U_i} f \cdot \varphi_i\, dU_i\ .$$

See [52, Ch. 3] for the details of this definition. The $d$-dimensional volume of $M$ is defined via

$$\mathrm{vol}_d\, M := \int_M 1\, dM\ .$$

This notion of volume coincides with the usual (Lebesgue) volume on $\mathbb{R}^d$ if $M$ is an open subset of $\mathbb{R}^d$.

An important tool in our computations will be the *smooth coarea formula*. Before we can state this we need to define the *Normal Jacobian* of a (surjective) linear operator.

If $A\colon V \to W$ is a surjective linear operator between euclidean vector spaces $V$ and $W$ of dimensions $n := \dim V \geq \dim W =: d$, then the Normal Jacobian of $A$ is defined as

$$\operatorname{ndet}(A) := |\det(A|_{\ker(A)^\perp})| \;, \tag{4.9}$$

where $A|_{\ker(A)^\perp}$ denotes the restriction of $A$ to the orthogonal complement of the kernel of $A$. Obviously, if $m = n$ then $\operatorname{ndet}(A) = |\det(A)|$, so the Normal Jacobian provides a natural generalization of the absolute value of the determinant.

**Lemma 4.1.15** (Coarea Formula)**.** *Let $M_1, M_2 \subseteq \mathbb{R}^n$ be smooth submanifolds of $\mathbb{R}^n$, and let $\varphi\colon M_1 \to M_2$ be a smooth map such that $D_x\varphi\colon T_xM_1 \to T_{\varphi(x)}M_2$ is surjective for almost all $x \in M_1$. Then for all integrable functions $f\colon M_1 \to \mathbb{R}$*

$$\int_{x\in M_1} f(x)\,dM_1 = \int_{y\in M_2}\int_{x\in\varphi^{-1}(y)} \frac{f(x)}{\operatorname{ndet}(D_x\varphi)}\,d\varphi^{-1}(y)\,dM_2 \;, \tag{4.10}$$

$$\int_{x\in M_1} f(x)\cdot\operatorname{ndet}(D_x\varphi)\,dM_1 = \int_{y\in M_2}\int_{x\in\varphi^{-1}(y)} f(x)\,d\varphi^{-1}(y)\,dM_2 \;. \tag{4.11}$$

*In particular, if $\varphi\colon M_1 \to M_2$ is a diffeomorphism and $g\colon M_2 \to \mathbb{R}$ integrable, then*

$$\int_{x\in M_1} g(\varphi(x))\cdot|\det(D_x\varphi)|\,dM_1 = \int_{y\in M_2} g(y)\,dM_2 \;. \tag{4.12}$$

*Proof.* See [37, 3.8] and [26, 3.2.11]. □

**Remark 4.1.16.** The inner integrals in (4.11) over the fiber $\varphi^{-1}(y)$ are well-defined for almost all $y \in M_2$. This follows from Sard's lemma (cf. for example [52, Thm. 3-14]), which implies that almost all $y \in M_2$ are regular values, i.e., the differential $D_x\varphi$ has full rank for all $x \in \varphi^{-1}(y)$. The fibers $\varphi^{-1}(y)$ of regular values $y$ are smooth submanifolds of $M_1$ and therefore the integral over $\varphi^{-1}(y)$ is well-defined.

The following lemma is an important step for the proof of Weyl's tube formula. Recall from Section 3.3 that the normal bundle $T^\perp M \subset \mathbb{R}^n \times \mathbb{R}^n$ of a smooth manifold $M \subseteq \mathbb{R}^n$ is a manifold of dimension $n$. Furthermore, we have a canonical projection $P\colon T^\perp M \to M$, $P\colon (p,\eta) \mapsto p$ (cf. Remark 3.3.8). In the following lemma we compute the Normal Jacobian of the derivative of this projection. Recall that we have computed the tangent spaces of the (unit) normal bundle in Lemma 4.1.7.

**Lemma 4.1.17.** *Let $M \subseteq \mathbb{R}^n$ be a smooth manifold of dimension $d := \dim M$, and let $P\colon T^\perp M \to M$ denote the canonical projection $P\colon (p,\eta) \mapsto p$. Then for $(p,\eta) \in T^\perp M$ the Normal Jacobian of $D_{(p,\eta)}P$ is given by*

$$\operatorname{ndet}(D_{(p,\eta)}P) = \prod_{i=1}^{d} \left(1 + \kappa_i(p,\eta)^2\right)^{-\frac{1}{2}} \;,$$

*where $\kappa_1(p,\eta), \ldots, \kappa_d(p,\eta)$ denote the principal curvatures of $M$ at $p$ in direction $\eta$. If additionally $M \subseteq S^{n-1}$, and if $P'\colon T^S M \to M$ denotes the canonical projection of the spherical normal bundle, then*

$$\operatorname{ndet}\left(D_{(p,\eta)}P'\right) = \operatorname{ndet}(D_{(p,\eta)}P)$$

*for all $(p,\eta) \in T^S M$.*

*Proof.* In Lemma 4.1.7 it was shown that the tangent space of $T^\perp M$ at $(p,\eta)$ is given by $T_{(p,\eta)}T^\perp M = L_1 \oplus L_2$ with

$$L_1 := \{(\zeta, -W_{p,\eta}(\zeta)) \mid \zeta \in T_pM\}\ , \quad L_2 := \{0\} \times T_p^\perp M\ ,$$

where $W_{p,\eta}\colon T_pM \to T_pM$ denotes the Weingarten map of $M$ at $p$ in direction $\eta$. Note that $L_1 \perp L_2$. Furthermore, we clearly have

$$D_{(p,\eta)}P(\xi_1,\xi_2) = \xi_1\ , \quad \text{for } (\xi_1,\xi_2) \in T_{(p,\eta)}T^\perp M \subset \mathbb{R}^n \times \mathbb{R}^n\ .$$

This implies that

$$\ker D_{(p,\eta)}P = L_2\ , \quad (\ker D_{(p,\eta)}P)^\perp = L_1\ .$$

To define a basis in $L_1$, let $\zeta_1,\ldots,\zeta_d \in T_pM$ be an orthonormal basis consisting of principal directions, i.e., $W_{p,\eta}(\zeta_i) = \kappa_i(p,\eta)\cdot\zeta_i$, $i=1,\ldots,d$. The corresponding vectors $(\zeta_i, -\kappa_i\cdot\zeta_i)$, $i=1,\ldots,d$, provide an orthogonal basis of $L_1$ of lengths $\|(\zeta_i,-\kappa_i\cdot\zeta_i)\| = \sqrt{1+\kappa_i(p,\eta)^2}$. Furthermore, this orthogonal basis is mapped onto the orthonormal basis $\zeta_1,\ldots,\zeta_d$ of $T_pM$. It follows that the Normal Jacobian of $D_{(p,\eta)}P$ is given by

$$\operatorname{ndet}(D_{(p,\eta)}P) = \prod_{i=1}^d \left(1+\kappa_i(p,\eta)^2\right)^{-\frac{1}{2}}\ .$$

The claim about the spherical normal bundle follows analogously. □

As a first example for the usefulness of the coarea formula, we will compute the volume of tubes around subspheres of $S^{n-1}$. Throughout this paper we use the notation

$$\mathcal{O}_k \ := \ \operatorname{vol}_k S^k \ = \ \frac{2\pi^{\frac{k+1}{2}}}{\Gamma(\frac{k+1}{2})} \tag{4.13}$$

$$\mathcal{O}_{n-1,k}(\alpha) \ := \ \operatorname{vol}_{n-1}\mathcal{T}(S,\alpha)\ , \tag{4.14}$$

where $S \in \mathcal{S}^k(S^{n-1})$, and $0 \le \alpha \le \frac{\pi}{2}$. Note that we have

$$\frac{(k-1)\cdot\mathcal{O}_k}{\mathcal{O}_{k-2}} = (k-1)\cdot\frac{2\pi^{\frac{k+1}{2}}}{\frac{k-1}{2}\cdot\Gamma(\frac{k-1}{2})}\cdot\frac{\Gamma(\frac{k-1}{2})}{2\pi^{\frac{k-1}{2}}} = 2\pi\ . \tag{4.15}$$

**Proposition 4.1.18.** *The volume of the $\alpha$-tube, $0 \le \alpha \le \frac{\pi}{2}$, around a subsphere $S \in \mathcal{S}^k(S^{n-1})$, $0 \le k \le n-2$, is given by*

$$\mathcal{O}_{n-1,k}(\alpha) = \mathcal{O}_k\cdot\mathcal{O}_{n-2-k}\cdot\int_0^\alpha \cos(\rho)^k\cdot\sin(\rho)^{n-2-k}\,d\rho\ .$$

*Furthermore, the volume of a circular cap $B(z,\beta)$ of radius $\beta\in[0,\pi]$ is given by*

$$\operatorname{vol} B(z,\beta) = \mathcal{O}_{n-2}\cdot\int_0^\beta \sin(\rho)^{n-2}\,d\rho\ .$$

*Proof.* Using the continuity of both sides we may assume w.l.o.g. $0<\alpha<\frac{\pi}{2}$. Let $S\in\mathcal{S}^k(S^{n-1})$, and consider the open subset

$$\mathcal{T}_\alpha := \mathcal{T}(S,\alpha)\setminus(S\cup\partial\mathcal{T}(S,\alpha))\ .$$

of $S^{n-1}$. Note that $\mathcal{T}(S,\alpha)\setminus\mathcal{T}_\alpha$ is a nullset, so that we have $\operatorname{vol}\mathcal{T}(S,\alpha)=\operatorname{vol}\mathcal{T}_\alpha$. Furthermore, we have a differentiable bijection

$$\begin{array}{rcl}\varphi\colon S\times S^\perp\times(0,\alpha) & \to & \mathcal{T}_\alpha\\ (p,q,\rho) & \mapsto & \cos(\rho)\cdot p+\sin(\rho)\cdot q\ .\end{array}$$

In order to compute the determinant of the derivative, let $\zeta\in T_pS$ and $\eta\in T_qS^\perp$. Note that $\zeta$ and $\eta$ are orthogonal to $p$ and $q$. We have

$$\begin{aligned}D_{(p,q,\rho)}\varphi(\zeta,0,0)&=\cos(\rho)\cdot\zeta\\ D_{(p,q,\rho)}\varphi(0,\eta,0)&=\sin(\rho)\cdot\eta\\ D_{(p,q,\rho)}\varphi(0,0,1)&=-\sin(\rho)\cdot p+\cos(\rho)\cdot q\ .\end{aligned}$$

Note that the vector $-\sin(\rho)\cdot p+\cos(\rho)\cdot q$ has unit length and it is orthogonal to the vectors $\zeta$ and $\eta$. Therefore, we get

$$\begin{aligned}|\det(D_{(p,q,\rho)}\varphi)|&=\left|\det\begin{pmatrix}\cos(\rho)\cdot I_k & & \\ & \sin(\rho)\cdot I_{n-2-k} & \\ & & 1\end{pmatrix}\right|\\ &=\cos(\rho)^k\cdot\sin(\rho)^{n-2-k}\ .\end{aligned}$$

From (4.12) applied to $\varphi$ we get

$$\begin{aligned}\mathcal{O}_{n-1,k}(\alpha)&=\operatorname{vol}\mathcal{T}_\alpha=\int_{\mathcal{T}_\alpha}1\,dS^{n-1}\\ &=\int_{S\times S^\perp\times(0,\alpha)}\cos(\rho)^k\cdot\sin(\rho)^{n-2-k}\,d(S\times S^\perp\times(0,\alpha))\\ &=\operatorname{vol}S\cdot\operatorname{vol}S^\perp\cdot\int_0^\alpha\cos(\rho)^k\cdot\sin(\rho)^{n-2-k}\,d\rho\ .\end{aligned}$$

The claim about the circular caps follows analogously by taking $k=0$ and replacing the 0-subsphere, which consists of a pair of antipodal points, by a single point. □

**Corollary 4.1.19.** *For $k,\ell\in\mathbb{N}$ we have*

$$\int_0^{\frac{\pi}{2}}\cos(\rho)^k\cdot\sin(\rho)^\ell\,d\rho=\frac{\mathcal{O}_{k+\ell+1}}{\mathcal{O}_k\cdot\mathcal{O}_\ell}\ .$$

*Proof.* Let $S\in\mathcal{S}^j(S^{n-1})$ be a subsphere of $S^{n-1}$. The complement of the open tube of radius $\frac{\pi}{2}$ around $S$ is given by $S^\perp$, which is a nullset. Therefore, we have

$$\mathcal{O}_{n-1}=\mathcal{O}_{n-1,j}(\tfrac{\pi}{2})=\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}\cdot\int_0^{\frac{\pi}{2}}\cos(\rho)^j\cdot\sin(\rho)^{n-2-j}\,d\rho$$

by Proposition 4.1.18. The claim follows by choosing $j:=k$ and $n:=k+\ell+2$. □

### 4.1.4 The binomial coefficient and related quantities

In the previous section we already encountered the quantity $\mathcal{O}_k=\operatorname{vol}_k S^k$ in (4.13). In this section we will present related quantities that will come up in the computations, and we will state and prove some properties and identities between them.

This collection will prove useful in the subsequent sections, but it may safely be skipped on a first reading.

Central to all the quantities, that we will deal with, is the $\Gamma$-function, which we consider as a function $\Gamma\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$. In the following we summarize some (well-known) properties of the $\Gamma$-function that we will make use of (cf. [1, § 6.1]):

1. For $x \in \mathbb{R}_{>0}$ we have $\Gamma(x+1) = x \cdot \Gamma(x)$.
2. We have $\Gamma(1) = 1$ and $\Gamma(n) = (n-1)!$ for $n \in \mathbb{Z}_{>0}$. Furthermore, we have $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
3. The function $\mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto \ln(\Gamma(x))$, is convex. This is equivalent to the inequality
$$\Gamma(x)^2 \leq \Gamma(x-c) \cdot \Gamma(x+c)\ ,$$
for $x, c \in \mathbb{R}$, $x > |c|$.
4. For $x \in \mathbb{R}_{>0}$ we have the duplication formula
$$\Gamma(2x) = \tfrac{1}{\sqrt{\pi}} \cdot 2^{2x-1} \cdot \Gamma(x) \cdot \Gamma(x + \tfrac{1}{2})\ . \tag{4.16}$$
5. For $x \in \mathbb{R}_{>0}$ we have the estimate
$$\Gamma(x + \tfrac{1}{2}) < \sqrt{x} \cdot \Gamma(x)\ . \tag{4.17}$$
This estimate is asymptotically sharp, i.e., $\Gamma(x+\frac{1}{2}) \sim \sqrt{x} \cdot \Gamma(x)$ for $x \to \infty$, where $f(x) \sim g(x)$ means $\frac{f(x)}{g(x)} \to 1$ for $x \to \infty$.

Besides the volume of the $k$th unit sphere, we will come across the volume of the $k$th unit ball $B_k \subset \mathbb{R}^k$. We denote this by

$$\omega_k \ := \ \mathrm{vol}_k\, B_k \ = \ \frac{\mathcal{O}_{k-1}}{k} \ = \ \frac{\pi^{\frac{k}{2}}}{\Gamma(\frac{k+2}{2})} \quad \text{for } k > 0, \text{ and put } \omega_0 := 1\ .$$

It is convenient for us to extend the binomial coefficient with the help of the $\Gamma$-function to also include half-integers.[1] For $n, m \in \mathbb{Z}$, $-1 \leq m \leq n+1$, we define

$$\binom{n/2}{m/2} := \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+2}{2}) \cdot \Gamma(\frac{n-m+2}{2})}\ . \tag{4.18}$$

Furthermore, besides the binomial coefficient we define the *flag coefficients* $\left[ {n \atop m} \right]$ for $n, m \in \mathbb{N}$, $n \geq m$, via

$$\begin{bmatrix} n \\ m \end{bmatrix} := \frac{\sqrt{\pi} \cdot \Gamma(\frac{n+1}{2})}{\Gamma(\frac{m+1}{2}) \cdot \Gamma(\frac{n-m+1}{2})}\ . \tag{4.19}$$

These coefficients were defined in [36, Ch. 6], and they can be interpreted as continuous analogues of the binomial coefficients. The following proposition provides some identities between the binomial coefficient and the flag coefficients.

[1] Note that we might as well extend the binomial coefficient to a function $\mathbb{R} \times \mathbb{R} \to [-\infty, \infty]$ via $(x, y) \mapsto \binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(y+1) \cdot \Gamma(x-y+1)}$. The restriction to half-integers is only for convenience, as we will only need these values.

**Proposition 4.1.20.** *1. For $n,m\in\mathbb{N}$, $n\geq m$,*

$$\binom{n}{m} = \frac{\sqrt{\pi}\cdot\Gamma(\frac{n+1}{2})\cdot\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+1}{2})\cdot\Gamma(\frac{m+2}{2})\cdot\Gamma(\frac{n-m+1}{2})\cdot\Gamma(\frac{n-m+2}{2})}\ . \tag{4.20}$$

*In particular,*

$$\binom{n}{m} = \begin{bmatrix} n \\ m \end{bmatrix}\cdot\binom{n/2}{m/2}\ . \tag{4.21}$$

*2. For $n,m\in\mathbb{N}$, $n\geq m$,*

$$\binom{n/2}{m/2} = \frac{\omega_m\cdot\omega_{n-m}}{\omega_n}\quad ,\qquad \begin{bmatrix} n \\ m \end{bmatrix} = \frac{\mathcal{O}_m\cdot\mathcal{O}_{n-m}}{2\cdot\mathcal{O}_n}\ . \tag{4.22}$$

*In particular,*

$$\binom{n}{m}\cdot\frac{\omega_n}{\omega_m\cdot\omega_{n-m}} = \begin{bmatrix} n \\ m \end{bmatrix}\ , \tag{4.23}$$

$$\binom{n}{m}\cdot\frac{2\cdot\mathcal{O}_n}{\mathcal{O}_m\cdot\mathcal{O}_{n-m}} = \binom{n/2}{m/2}\ . \tag{4.24}$$

*3. For $n,m\to\infty$ such that also $(n-m)\to\infty$*

$$\begin{bmatrix} n \\ m \end{bmatrix} \sim \sqrt{\frac{\pi}{2}}\cdot\sqrt{\frac{m(n-m)}{n}}\cdot\binom{n/2}{m/2}\ . \tag{4.25}$$

*Proof.* Equation (4.20) follows from the duplication formula of the Γ-function via

$$\begin{aligned}\binom{n}{m} &= \frac{\Gamma(n+1)}{\Gamma(m+1)\cdot\Gamma(n-m+1)} \\ &= \frac{\frac{1}{\sqrt{\pi}}\cdot 2^n\cdot\Gamma(\frac{n+1}{2})\cdot\Gamma(\frac{n+2}{2})}{\frac{1}{\sqrt{\pi}}\cdot 2^m\cdot\Gamma(\frac{m+1}{2})\cdot\Gamma(\frac{m+2}{2})\cdot\frac{1}{\sqrt{\pi}}\cdot 2^{n-m}\cdot\Gamma(\frac{n-m+1}{2})\cdot\Gamma(\frac{n-m+2}{2})} \\ &= \frac{\sqrt{\pi}\cdot\Gamma(\frac{n+1}{2})\cdot\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+1}{2})\cdot\Gamma(\frac{m+2}{2})\cdot\Gamma(\frac{n-m+1}{2})\cdot\Gamma(\frac{n-m+2}{2})}\ .\end{aligned}$$

Equations (4.21)–(4.24) follow from (4.20) by plugging in the definitions of the corresponding quantities.

As for the asymptotics stated in (4.25), we compute

$$\begin{aligned}\sqrt{\frac{\pi}{2}}\cdot\sqrt{\frac{m(n-m)}{n}}\cdot\binom{n/2}{m/2} &= \sqrt{\frac{\pi}{2}}\cdot\sqrt{\frac{m(n-m)}{n}}\cdot\frac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{m}{2}+1)\cdot\Gamma(\frac{n-m}{2}+1)} \\ &= \frac{\sqrt{\pi}\cdot\sqrt{\frac{n}{2}}\cdot\Gamma(\frac{n}{2})}{\sqrt{\frac{m}{2}}\cdot\Gamma(\frac{m}{2})\cdot\sqrt{\frac{n-m}{2}}\cdot\Gamma(\frac{n-m}{2})} \sim \frac{\sqrt{\pi}\cdot\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m+1}{2})\cdot\Gamma(\frac{n-m+1}{2})} = \begin{bmatrix} n \\ m \end{bmatrix}\ ,\end{aligned}$$

where we have used the asymptotics $\sqrt{x}\cdot\Gamma(x)\sim\Gamma(x+\frac{1}{2})$ for $x\to\infty$. □

We finish this section with a discussion about a particularly important property, which is shared by many sequences that we will come across. This property is *log-concavity*. Additionally, we also mention *unimodality* of sequences.

**Definition 4.1.21.** Let $(a_i)_{i=0,\dots,N}$ be a sequence of real numbers, where $N \in \mathbb{N}$ or $N = \infty$.

1. The sequence $(a_i)_i$ is called *log-concave* iff $a_i^2 \geq a_{i-1} \cdot a_{i+1}$ for all $0 < i < N$.
2. The sequence $(a_i)_i$ is called *unimodal* iff there exists $0 \leq M \leq N$ such that $a_0 \leq a_1 \leq \ldots \leq a_M$ and $a_M \geq a_{M+1} \geq \ldots \geq a_N$.
3. A zero element of the sequence $a_i = 0$ for some $0 < i < N$ is called an *internal zero* iff there exist $0 \leq j < i$ and $i < k \leq N$ such that $a_j \neq 0$ and $a_k \neq 0$.

Note that a sequence of positive real numbers $(a_i)_i$ is log-concave iff the sequence $(\ln a_i)_i$ is concave. The following proposition collects some properties of log-concave and unimodal sequences.

**Proposition 4.1.22.** *1. Let $(a_i)_{i=0,\dots,N}$ be a log-concave sequence of real numbers, where $N \in \mathbb{N}$ or $N = \infty$. If $a_i > 0$ for all $i$ or if $a_i \geq 0$ and $(a_i)_i$ has no internal zeros, then $(a_i)_i$ is a unimodal sequence.*

*2. If $(a_i)_i$ is a log-concave sequence, then so is the sequence $(\sqrt{a_i})_i$.*

*3. If $(a_i)_i, (b_i)_i$ are log-concave sequences, then so is the product sequence $(a_i \cdot b_i)_i$.*

*4. Let $a_0, \dots, a_m, b_0, \dots, b_n \in \mathbb{R}_{\geq 0}$, and let $(c_i)_i$ denote their convolution, i.e.,*

$$c_i := \sum_{k=\max\{0,i-n\}}^{\min\{i,m\}} a_k \cdot b_{i-k} \ .$$

*(a) If $(a_i)_i$ and $(b_i)_i$ are log-concave sequences with no internal zeros, then also the sequence $(c_i)_i$ is log-concave and with no internal zero.*

*(b) If $(a_i)_i$ and $(b_i)_i$ are symmetric, i.e., $a_i = a_{m-i}$ and $b_j = b_{n-j}$ for all $i$ and $j$, and if both $(a_i)_i$ and $(b_i)_i$ are unimodal, then also the sequence $(c_i)_i$ is symmetric and unimodal.*

*Proof.* Parts (1)–(3) follow directly from the definition. For a proof of part (4) see for example [54, Prop. 1 & 2]. □

**Proposition 4.1.23.** *The following sequences of positive numbers are log-concave:*

*1. For $n \in \mathbb{N}$:*

*(a)* $\left[{n \atop 0}\right], \left[{n \atop 1}\right], \dots, \left[{n \atop n}\right]$

*(b)* $\binom{n/2}{-1/2}, \binom{n/2}{0/2}, \binom{n/2}{1/2}, \dots, \binom{n/2}{(n+1)/2}$

*(c)* $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$

*2. For $m \in \mathbb{N}$:*

*(a)* $\left[{m \atop m}\right], \left[{m+1 \atop m}\right], \left[{m+2 \atop m}\right], \dots$

*(b)* $\binom{m/2}{m/2}, \binom{(m+1)/2}{m/2}, \binom{(m+2)/2}{m/2}, \dots$

*(c)* $\binom{m}{m}, \binom{m+1}{m}, \binom{m+2}{m}, \dots$

*3.* $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots$

*4.* $\omega_0, \omega_1, \omega_2, \ldots$.

*Proof.* All these claims follow from the log-convexity of the Γ-function. We exemplarily show part (1.a) of the claim:

$$\begin{aligned}\frac{\left[\begin{smallmatrix}n\\i\end{smallmatrix}\right]^2}{\left[\begin{smallmatrix}n\\i-1\end{smallmatrix}\right]\cdot\left[\begin{smallmatrix}n\\i+1\end{smallmatrix}\right]} &= \frac{\pi\cdot\Gamma(\frac{n+1}{2})^2}{\Gamma(\frac{i+1}{2})^2\cdot\Gamma(\frac{n-i+1}{2})^2}\cdot\frac{\Gamma(\frac{i}{2})\cdot\Gamma(\frac{n-i+2}{2})}{\sqrt{\pi}\cdot\Gamma(\frac{n+1}{2})}\cdot\frac{\Gamma(\frac{i+2}{2})\cdot\Gamma(\frac{n-i}{2})}{\sqrt{\pi}\cdot\Gamma(\frac{n+1}{2})}\\ &= \frac{\Gamma(\frac{i}{2})\cdot\Gamma(\frac{i+2}{2})}{\Gamma(\frac{i+1}{2})^2}\cdot\frac{\Gamma(\frac{n-i+2}{2})\cdot\Gamma(\frac{n-i}{2})}{\Gamma(\frac{n-i+1}{2})^2}\\ &\geq 1\ ,\end{aligned}$$

where the inequality follows from the log-convexity of the Γ-function. □

See [54] for more on log-concavity in diverse areas of mathematics.

## 4.2 The (euclidean) Steiner polynomial

Before we make the computations in the sphere let us have a look at the euclidean situation. As early as 1840, J. Steiner found that the volume of the tube of radius $r$ around a convex body $K \subset \mathbb{R}^n$ has the form of a polynomial in $r$. See Figure 1.1 in Section 1.3 for a 2-dimensional example, which shows that at least in dimension $n = 2$ Steiner's formula seems obvious.

The *intrinsic volumes* $V_i^e(K)$ of $K \in \mathcal{K}(\mathbb{R}^n)$ are defined as (scaled versions of) the coefficients of the Steiner polynomial. More precisely,

$$\mathrm{vol}_n\, \mathcal{T}^e(K, r) = \sum_{i=0}^{n} \omega_i \cdot V_{n-i}^e(K)\cdot r^i \tag{4.26}$$

(cf. [49, Sec. 4.2]). The quantities $V_i^e(K)$ are called the intrinsic volumes of $K$ because they do not depend on the embedding of $K$ in $\mathbb{R}^n$, i.e., considering the convex body $K \subset \mathbb{R}^{\tilde{n}}$ with $\tilde{n} \geq n$ will yield the same intrinsic volumes.

For $K \in \mathcal{K}(\mathbb{R}^n)$ we have

$$V_n^e(K) = \mathrm{vol}_n(K)\ ,\quad V_{n-1}^e(K) = \frac{\mathrm{vol}_{n-1}(\partial K)}{2}\ ,\quad V_0^e(K) = 1\ .$$

Furthermore, in the special case where $K$ is a polytope, the intrinsic volumes are given by

$$V_i^e(K) = \sum_F \mathrm{vol}_i(F)\cdot\frac{\mathrm{vol}_{n-i-1}(N_F^S)}{\mathcal{O}_{n-i-1}} \tag{4.27}$$

(cf. [49, Sec. 4.2]), where the summation is over all $i$-dimensional faces $F$ of $K$, and $N_F^S$ denotes the intersection of the outer cone to $K$ in $F$ with the unit sphere (cf. Section 3.1).

**Example 4.2.1.** As a first example, we will compute the intrinsic volumes of the $n$-dimensional unit ball $B_n$. If $B_n(r) \subset \mathbb{R}^n$ denotes the $n$-dimensional ball of radius $r$, then we have $\mathcal{T}^e(B_n, r) = B_n(1 + r)$. Therefore, we get

$$\mathrm{vol}\,\mathcal{T}^e(B_n, r) = \mathrm{vol}\, B_n(1 + r) = (1 + r)^n \cdot \omega_n = \sum_{i=0}^{n}\binom{n}{i}\cdot\omega_n\cdot r^i\ ,$$

which implies

$$V_i^e(B_n) = \binom{n}{i} \cdot \frac{\omega_n}{\omega_{n-i}} \; .$$

Note that if we define for $K \in \mathcal{K}(\mathbb{R}^n)$ a modified version $\tilde{V}_i^e$ of intrinsic volumes via

$$\tilde{V}_i^e(K) := \frac{V_i^e(K)}{\omega_i} \; ,$$

then these modified intrinsic volumes are also independent of the embedding of $K$ in euclidean space. Furthermore, we get by Proposition 4.1.20

$$\tilde{V}_i^e(B_n) = \binom{n}{i} \cdot \frac{\omega_n}{\omega_i \cdot \omega_{n-i}} = \begin{bmatrix} n \\ i \end{bmatrix} \; .$$

Recall that in Section 3.1.1 we have discussed the Minkowski addition in euclidean space. This Minkowski addition leads to a vast generalization of intrinsic volumes, the notion of *mixed volumes*. We will not need this notion as it has no direct spherical analog. But we mention the fact that the Alexandrov-Fenchel inequality for mixed volumes (cf. [49, Ch. 5]) implies that the sequence of (euclidean) intrinsic volumes is always log-concave, which we formulate in the following proposition.

**Proposition 4.2.2.** *For $K \in \mathcal{K}(\mathbb{R}^n)$ the sequence $\tilde{V}_0^e(K), \ldots, \tilde{V}_n^e(K)$ is log-concave. In particular, the sequence $V_0^e(K), \ldots, V_n^e(K)$ is log-concave, i.e.*

$$V_i^e(K)^2 \geq V_{i-1}^e(K) \cdot V_{i+1}^e(K) \; , \quad i = 1, \ldots, n-1 \; .$$

*Proof.* For the sake of completeness we include the derivation from the Alexandrov-Fenchel inequality by adopting the notation of [49, Ch. 5] and by referring to this for the necessary definitions. From [49, (5.1.26)] we have for $K_1, K_2 \in \mathcal{K}(\mathbb{R}^n)$ and $\lambda_1, \lambda_2 \in \mathbb{R}_+$

$$\mathrm{vol}_n(\lambda_1 \cdot K_1 + \lambda_2 \cdot K_2) = \sum_{i=0}^{n} \binom{n}{i} \cdot \lambda_1^i \cdot \lambda_2^{n-i} \cdot V(K_1[i], K_2[n-i]) \; ,$$

where $V(K_1[i], K_2[n-i])$ denotes a particular mixed volume. Choosing $K_1 := K$ and $K_2 := B_n$, and using $\mathcal{T}^e(K, r) = K + r\, B_n$, we get

$$\mathrm{vol}_n\, \mathcal{T}^e(K, r) = \sum_{k=0}^{n} \binom{n}{k} \cdot V(K[n-k], B_n[k]) \cdot r^k \; .$$

Comparing this with the definition of the intrinsic volumes, we get

$$\begin{aligned} \tilde{V}_i^e(K) &= \binom{n}{i} \cdot \frac{1}{\omega_{n-i} \cdot \omega_i} \cdot V(K[i], B_n[n-i]) \\ &= \begin{bmatrix} n \\ i \end{bmatrix} \cdot \frac{1}{\omega_n} \cdot V(K[i], B_n[n-i]) \; , \end{aligned}$$

where the second equality follows from Proposition 4.1.20. The Alexandrov-Fenchel inequality (cf. [49, Sec. 6.3]) implies that for $1 \leq i \leq n-1$

$$V(K[i], B_n[n-i])^2 \geq V(K[i+1], B_n[n-i-1]) \cdot V(K[i-1], B_n[n-i+1]) \; .$$

Using this and the log-concavity of the sequence $([{n \atop i}])_i$ (cf. Proposition 4.1.23), we compute for $1 \leq i \leq n-1$

$$\begin{aligned}
&\frac{\tilde{V}_i^e(K)^2}{\tilde{V}_{i-1}^e(K) \cdot \tilde{V}_{i+1}^e(K)} \\
&= \frac{[{n \atop i}]^2}{[{n \atop i-1}] \cdot [{n \atop i+1}]} \cdot \frac{V(K[i], B_n[n-i])^2}{V(K[i+1], B_n[n-i-1]) \cdot V(K[i-1], B_n[n-i+1])} \\
&\geq 1\ .
\end{aligned}$$

This shows that the sequence $\tilde{V}_0^e(K), \ldots, \tilde{V}_n^e(K)$ is log-concave. The fact that the sequence $V_0^e(K), \ldots, V_n^e(K)$ is log-concave follows from the log-concavity of $\omega_0, \omega_1, \omega_2, \ldots$ and the identity $V_i^e(K) = \tilde{V}_i^e(K) \cdot \omega_i$. □

A spherical analog of Proposition 4.2.2 is unknown. We will formulate one in Conjecture 4.4.16.

## 4.3 Weyl's tube formulas

In this section we will derive formulas for the volume of the tube around a stratified convex set in $\mathbb{R}^n$ or $S^{n-1}$, respectively. The coefficients occuring, i.e., the intrinsic volumes and its spherical generalizations, turn out to be certain *integrals of curvature* over the boundary of the convex set. These formulas are well-known and originally due to H. Weyl, cf. [63]. The reason for us to include these computations is that they will serve as a model for the computation in the Grassmann manifold. Furthermore, the form in which we state these tube formulas is particularly useful for the computations in Chapter B in the appendix, where we will prove some simple calculation rules for the intrinsic volumes (cf. Section 4.4), which appear to be new. And in Chapter C in the appendix we will use the spherical tube formula to compute the intrinsic volumes of the semidefinite cone (cf. Section 4.4.1), which also seems to have never been done before.

While this section is solely devoted to the derivation of the tube formulas, we will treat in the forthcoming section the intrinsic volumes, which evolve from these formulas. We rely in this section basically on Weyl's original paper [63] and on the treatment in [17] and in [12].

Before making the first definitions let us have a look at another example, similar to the euclidean polytope in Figure 1.1. The intersection of the positive orthant $\mathbb{R}^n_+$ with the unit sphere $S^{n-1}$ is a spherical polytope. Figure 1.2 in Section 1.3 shows this spherical polytope as well as the decomposition of the tube around it for the special case $n = 3$.

This example shows that the volume of the tube will not be a polynomial function, but may nevertheless have a similar structure. It turns out that the only change one has to do in the formula (4.26), is to replace the monomials $r^k$ by functions which arise in the volume of the tube around subspheres of $S^{n-1}$.

Recall from (4.13) and (4.14) that we denote the volume of the $k$-dimensional unit sphere by $\mathcal{O}_k = \mathrm{vol}_k\, S^k$, and the volume of the $\alpha$-tube around a $k$-sphere in $S^{n-1}$ by $\mathcal{O}_{n-1,k}(\alpha) = \mathrm{vol}_{n-1}\, \mathcal{T}(S, \alpha)$, $S \in \mathcal{S}^k(S^{n-1})$. Note that $\mathcal{O}_{n-1,k}(\frac{\pi}{2}) = \mathcal{O}_{n-1}$. We further use the notation

$$I_{n,j}(\alpha) := \frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_j \cdot \mathcal{O}_{n-2-j}} = \int_0^\alpha \cos(\rho)^j \cdot \sin(\rho)^{n-2-j}\, d\rho\ . \tag{4.28}$$

Note that for small values of $\alpha$ we have approximately $\sin(\alpha) \sim \alpha$ and $\cos(\alpha) \sim 1$, and thus for $\alpha \to 0$

$$I_{n,j}(\alpha) \quad \sim \quad \int_0^\alpha \rho^{n-2-j}\, d\rho \;=\; \tfrac{1}{n-1-j} \cdot \alpha^{n-1-j} \;. \tag{4.29}$$

So the $I$-functions should be thought of as spherical substitutes for the monomials, which appear in the euclidean tube formula.

**Proposition 4.3.1.** *Let $n \geq 2$ and let $\emptyset \neq U \subseteq \mathbb{R}$ open. Then the functions $I_{n,j}\colon U \to \mathbb{R}$, $0 \leq j \leq n-2$ and the constant function $\mathbf{1}\colon U \to \mathbb{R}$, $\mathbf{1}(\alpha) := 1$, are linearly independent.*

*Proof.* It is easily seen that the functions $\cos(\alpha)^j \cdot \sin(\alpha)^{n-2-j} = I'_{n,j}(\alpha)$, are linearly independent on an open interval. If we have $a_1 \cdot \mathbf{1} + a_2 \cdot I_{n,0} + \ldots + a_n \cdot I_{n,n-2} = 0$, then $a_2 \cdot I'_{n,0} + \ldots + a_n \cdot I'_{n,n-2} = 0$ and the linear independence of $I'_{n,j}$, $0 \leq j \leq n-2$ implies that $a_2 = \ldots = a_n = 0$, and therefore also $a_1 = 0$. □

We state Weyl's tube formulas for (euclidean or spherical) stratified convex sets. Recall from Definition 3.3.9 that we call a (euclidean or spherical) convex set $K$ stratified if it decomposes into a disjoint union of smooth connected submanifolds of $\mathbb{R}^n$ resp. $S^{n-1}$ such that the duality bundles (cf. (3.8)/(3.9)) also form smooth manifolds. The classification into essential and negligible pieces (cf. Definition 3.3.9) is justified by the following theorem.

**Theorem 4.3.2.** *1. Let $K \in \mathcal{K}(\mathbb{R}^n)$ be a stratified convex body with decomposition $K = \dot\bigcup_{i=0}^{\tilde{k}} M_i$, such that $M_0 = \operatorname{int} K$, and such that $M_1, \ldots, M_k$ are the essential and $M_{k+1}, \ldots, M_{\tilde{k}}$, $k \leq \tilde{k}$, are the negligible pieces. Furthermore, let $d_i$ be the dimension of the stratum $M_i$. Then for $r \geq 0$*

$$\operatorname{vol}_n \mathcal{T}^e(K,r) = \operatorname{vol}_n K + \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \sum_{i=1}^{k} \int_{x\in M_i} \int_{\eta \in N_x^S(K)} \sigma_{d_i-j}^{(i)}(x,-\eta)\, d\eta\, dx \;,$$

*where $\sigma_\ell^{(i)}(x,-\eta)$ denotes the $\ell$th elementary symmetric function in the principal curvatures of $M_i$ at $x$ in direction $-\eta$ (cf. (4.2)), and $\sigma_\ell(x,-\eta) := 0$ if $\ell < 0$.*

*2. Let $K \in \mathcal{K}(S^{n-1})$ be a stratified spherical convex set with decomposition $K = \dot\bigcup_{i=0}^{\tilde{k}} M_i$, such that $M_0 = \operatorname{int}(K)$, and such that $M_1, \ldots, M_k$ are the essential and $M_{k+1}, \ldots, M_{\tilde{k}}$, $k \leq \tilde{k}$, are the negligible pieces. Furthermore, let $d_i$ be the dimension of the stratum $M_i$. Then for $0 \leq \alpha \leq \frac{\pi}{2}$*

$$\operatorname{vol}_{n-1} \mathcal{T}(K,\alpha) = \operatorname{vol}_{n-1} K + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \sum_{i=1}^{k} \int_{p\in M_i} \int_{\eta \in N_p^S(K)} \sigma_{d_i-j}^{(i)}(p,-\eta)\, d\eta\, dp \;,$$

*where $\sigma_\ell^{(i)}(p,-\eta)$ denotes the $\ell$th elementary symmetric function in the principal curvatures of $M_i$ at $p$ in direction $-\eta$ (cf. (4.2)), and $\sigma_\ell(p,-\eta) := 0$ if $\ell < 0$.*

Note that the normal directions $\eta \in N_p^S$ point outwards $K$. Therefore, the principal curvatures $\kappa_i(p,-\eta)$ are nonnegative (cf. Remark 4.1.5).

Before we give the proof, let us consider the special cases of smooth and polyhedral caps.

**Corollary 4.3.3.** *Let $K \in \mathcal{K}(S^{n-1})$ have a smooth boundary $M := \partial K$. Then for $0 \leq \alpha \leq \frac{\pi}{2}$*

$$\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) = \mathrm{vol}_{n-1}\,K + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \int_{p \in M} \sigma_{n-2-j}(p)\,dM\,.$$

*Proof.* Take the decomposition $M_0 := \mathrm{int}(K)$, $M_1 := \partial K = M$. $\square$

The following result for the sphere is analogous to (4.27) for euclidean space.

**Corollary 4.3.4.** *Let $K \in \mathcal{K}^p(S^{n-1})$ be a polyhedral cap and let $\mathcal{F}_j$ denote the set of all $j$-dimensional faces of $K$. Then for $0 \leq \alpha \leq \frac{\pi}{2}$*

$$\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) = \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \sum_{F \in \mathcal{F}_j} \mathrm{vol}_j(F) \cdot \mathrm{vol}_{n-2-j}(N_F^S)\,.$$

*Proof.* The boundary of the polyhedral cap $K$ has the natural stratification

$$\partial K = \dot{\bigcup_{j=0}^{n-2}} \dot{\bigcup_{F \in \mathcal{F}_j}} \mathrm{relint}(F)\,,$$

where each piece is easily seen to be essential. Furthermore, as the normal cone is constant on each of the pieces $\mathrm{relint}(F)$, it follows that the corresponding duality bundles are smooth manifolds. Since the relative interior of every face of $K$ is an open subset of a subsphere of $S^{n-1}$, and as the curvatures of subspheres of $S^{n-1}$ are zero, we get

$$\sigma_j^F(p,\eta) = \begin{cases} 0 & \text{if } j > 0 \\ 1 & \text{if } j = 0 \end{cases}\,,$$

where the superscript $.^F$ shall indicate the dependence on the face $F$. From Theorem 4.3.2 part (2) we get

$$\begin{aligned}\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) &= \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} \sum_{F \in \mathcal{F}_j} I_{n,j}(\alpha) \cdot \int\limits_{p \in F} \int\limits_{\eta \in N_p^S} 1\,dN_p^S\,dF \\ &= \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \sum_{F \in \mathcal{F}_j} \mathrm{vol}_j(F) \cdot \mathrm{vol}_{n-2-j}(N_F^S)\,. \quad \square\end{aligned}$$

Before we give the proof of Theorem 4.3.2 let us also compute the volume of the tubes around circular caps and around the cap defined by the positive orthant.

**Example 4.3.5.** Let $K = B(z,\beta) \subset S^{n-1}$, $0 < \beta \leq \pi/2$, the circular cap of radius $\beta$ around $z$. In Example 4.1.6 we have seen that the principal curvatures in $p \in M = \partial K$ are given by $\kappa_1(p) = \ldots = \kappa_{n-2}(p) = \cot(\beta)$, if we choose the unit normal field pointing inwards $K$. So from Corollary 4.3.3 we get

$$\begin{aligned}\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) &= \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \int_{p \in M} \binom{n-2}{n-2-j} \cdot (\cot\beta)^{n-2-j}\,dM \\ &= \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \binom{n-2}{j} \cdot (\cot\beta)^{n-2-j} \cdot \mathrm{vol}_{n-2}(M) \\ &= \mathrm{vol}_{n-1}(K) + \mathcal{O}_{n-2} \cdot \sum_{j=0}^{n-2} \binom{n-2}{j} \cdot (\cos\beta)^{n-2-j} \cdot (\sin\beta)^j \cdot I_{n,j}(\alpha)\,,\end{aligned}$$

where the last equality follows from $\operatorname{vol} M = (\sin\beta)^{n-2} \cdot \mathcal{O}_{n-2}$.

Note that in the case of circular caps we may compute the volume of the $\alpha$-tube in a different way by taking the circular cap of radius $\beta+\alpha$. Using Proposition 4.1.18 we get for $0 \le \alpha \le \pi - \beta$

$$\begin{aligned}
\operatorname{vol}_{n-1} \mathcal{T}(K,\alpha) &= \operatorname{vol}_{n-1}\left(B(z,\beta+\alpha)\right)\\
&= \mathcal{O}_{n-2} \cdot \int_0^{\alpha+\beta} \sin(\rho)^{n-2}\, d\rho\\
&= \mathcal{O}_{n-2} \cdot \int_0^{\beta} \sin(\rho)^{n-2}\, d\rho + \mathcal{O}_{n-2} \cdot \int_0^{\alpha} \sin(\beta+\rho)^{n-2}\, d\rho\\
&= \operatorname{vol}_{n-1}(K) + \mathcal{O}_{n-2} \cdot \int_0^{\alpha} (\sin\rho\cos\beta + \cos\rho\sin\beta)^{n-2}\, d\rho\\
&\overset{(4.28)}{=} \operatorname{vol}_{n-1}(K) + \mathcal{O}_{n-2} \cdot \sum_{j=0}^{n-2} \binom{n-2}{j} \cdot (\cos\beta)^{n-2-j} \cdot (\sin\beta)^j \cdot I_{n,j}(\alpha)\ ,
\end{aligned}$$

which coincides with the result of the first computation.

**Example 4.3.6.** Let $K = \mathbb{R}^n_+ \cap S^{n-1}$ be the intersection of the positive orthant with the unit sphere. The face structure of $K$ is simple. A typical $k$-dimensional face $F$ of $K$ is given by equations of the form

$$x_1 = x_2 = \ldots = x_{n-1-k} = 0\ , \quad x_{n-k} > 0\,,\ \ldots\ ,\ x_n > 0\ , \quad \sum_{i=1}^n x_i^2 = 1\ . \tag{4.30}$$

The number of $k$-dimensional faces of $K$ is given by $\binom{n}{n-1-k}$ and the $k$-dimensional volume of such a face is $\frac{1}{2^{k+1}} \cdot \mathcal{O}_k$. The dual $N_F^S$ of the face $F$ defined in (4.30) is given by the equations (cf. Example 3.1.10)

$$x_1 < 0\,,\ x_2 < 0\,,\ \ldots\ ,\ x_{n-1-k} < 0\ , \quad x_{n-k} = \ldots = x_n = 0\ , \quad \sum_{i=1}^n x_i^2 = 1\ ,$$

and we have $\operatorname{vol}_{n-2-k} N_F^S = \frac{1}{2^{n-1-k}} \cdot \mathcal{O}_{n-2-k}$. So from Corollary 4.3.4 we get

$$\begin{aligned}
\operatorname{vol} \mathcal{T}(K,\alpha) &= \operatorname{vol}_{n-1}(K) + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \sum_{F \in \mathcal{F}_j} \operatorname{vol}_j(F) \cdot \operatorname{vol}_{n-2-j}(N_F^S)\\
&= \frac{\mathcal{O}_{n-1}}{2^n} + \sum_{j=0}^{n-2} I_{n,j}(\alpha) \cdot \binom{n}{n-1-j} \cdot \frac{1}{2^{j+1}} \cdot \mathcal{O}_j \cdot \frac{1}{2^{n-1-j}} \cdot \mathcal{O}_{n-2-j}\\
&= \frac{\mathcal{O}_{n-1}}{2^n} + \sum_{j=0}^{n-2} \mathcal{O}_{n-1,j}(\alpha) \cdot \frac{\binom{n}{j+1}}{2^n}\ .
\end{aligned}$$

In the remainder of this section we will give the proof of Weyl's tube formula for spherical convex sets as stated in Theorem 4.3.2 part (2).

*Proof of Theorem 4.3.2.* We will only prove part (2) of the theorem. The euclidean statement in part (1) follows analogously, and the proof even simplifies as some subtleties of the spherical setting do not appear in the euclidean setting.

First of all, due to continuity we may assume w.l.o.g. that $0 < \alpha < \frac{\pi}{2}$, so that $\mathcal{T}(K,\alpha) \subseteq S^{n-1} \setminus \breve{K}$. The projection map $\Pi_K\colon S^{n-1} \setminus \breve{K} \to K$ (cf. (3.2) in Section 3.1) provides the following decomposition of $S^{n-1}$

$$S^{n-1} \setminus \breve{K} = \dot{\bigcup_{i=0}^{\tilde{k}}} \Pi_K^{-1}(M_i) \ .$$

Denoting $\overline{\mathcal{T}_i}(\alpha) := \mathcal{T}(K,\alpha) \cap \Pi_K^{-1}(M_i)$ we get

$$\mathrm{vol}_{n-1}\, \mathcal{T}(K,\alpha) = \sum_{i=0}^{\tilde{k}} \mathrm{vol}_{n-1}\, \overline{\mathcal{T}_i}(\alpha) \ . \tag{4.31}$$

For $i = 0$ we have $M_0 = \mathrm{int}(K)$, and thus $\overline{\mathcal{T}_0}(\alpha) = \mathrm{int}(K)$. This implies

$$\mathrm{vol}_{n-1}\, \overline{\mathcal{T}_0}(\alpha) = \mathrm{vol}_{n-1}(K) \ . \tag{4.32}$$

Now, let us fix an index $1 \le i \le \tilde{k}$. The set $\overline{\mathcal{T}_i}(\alpha)$ is given by

$$\overline{\mathcal{T}_i}(\alpha) = \left\{ \cos(\rho)\cdot p + \sin(\rho)\cdot \eta \mid p \in M_i \ , \ \eta \in N_p^S(K) \ , \ 0 \le \rho \le \alpha \right\} \ .$$

Besides that, we define the set $\mathcal{T}_i(\alpha) \subseteq \overline{\mathcal{T}_i}(\alpha)$ via

$$\mathcal{T}_i(\alpha) := \left\{ \cos(\rho)\cdot p + \sin(\rho)\cdot \eta \mid p \in M_i \ , \ \eta \in N_p^\circ(K) \ , \ 0 < \rho < \alpha \right\} \ ,$$

where $N_p^\circ(K) := \mathrm{relint}(N_p(K)) \cap S^{n-1}$. It is easily seen that the complement of $\mathcal{T}_i(\alpha)$ in $\overline{\mathcal{T}_i}(\alpha)$ is a nullset, so that we have

$$\mathrm{vol}_{n-1}\, \overline{\mathcal{T}_i}(\alpha) = \mathrm{vol}_{n-1}\, \mathcal{T}_i(\alpha) \ . \tag{4.33}$$

Recall that $N^S M_i$ denotes the spherical duality bundle of $M_i$ (with respect to $K$) (cf. Section 3.3). This bundle is by the definition of stratified caps a smooth manifold. We now consider the map

$$\varphi_i\colon N^S M_i \times (0,\alpha) \to \mathcal{T}_i(\alpha) \ , \quad (p,\eta,\rho) \mapsto \cos(\rho)\cdot p + \sin(\rho)\cdot \eta \ . \tag{4.34}$$

This map is bijective by definition of $N^S M_i$ (cf. (3.9)); in the special case where $M_i$ has codimension 1 in $S^{n-1}$ we have seen this in the proof of Corollary 4.1.13. Furthermore, $\varphi_i$ is smooth by assumption on the smoothness of $M_i$ and $N^S M_i$. By the assumption that $M_1,\ldots,M_k$ are the essential pieces and $M_k,\ldots,M_{\tilde{k}}$ are negligible, we have

$$\dim N^S M_i \begin{cases} = n-2 & \text{if } i \le k \\ < n-2 & \text{if } k < i \le \tilde{k} \ . \end{cases}$$

This implies that the image of $\varphi_i$, i.e., the set $\mathcal{T}_i(\alpha)$, has volume 0 if $M_i$ is a negligible piece:

$$\mathrm{vol}_{n-1}\, \mathcal{T}_i(\alpha) = 0 \ , \quad \text{for } k < i \le \tilde{k} \ . \tag{4.35}$$

For the rest of the proof let us drop the index $i$ and we assume that $M = M_i$, $1 \le i \le k$, is an essential stratum of the given decomposition. To finish the proof we need to show that

$$\mathrm{vol}_{n-1}\, \mathcal{T}(\alpha) = \sum_{j=0}^{d} I_{n,j}(\alpha) \cdot \int\limits_{p\in M} \int\limits_{\eta \in N_p^S(K)} \sigma_{d-j}(p,-\eta)\, d\eta\, dp \ . \tag{4.36}$$

We will achieve this by computing the derivative of the function $\varphi$ from (4.34), so that we can apply the transformation formula from Lemma 4.1.15 to compute the volume of $\mathcal{T}(\alpha)$.

Note that $N^S M \subseteq T^S M$, and $\dim N^S M = \dim T^S M$ as $M$ is an essential piece (cf. Definition 3.3.9). Therefore, the spherical duality bundle $N^S M$ is an open subset of the spherical normal bundle $T^S M$. In particular, the tangent space of $N^S M$ at $(p,\eta)$ coincides with the tangent space of $T^S M$ at $(p,\eta)$, which we have computed in Lemma 4.1.7. Therefore, the tangent space of $N^S M\times(0,\alpha)$ at $(p,\eta,\rho)$ is given by

$$T_{(p,\eta,\rho)}(N^S M\times(0,\alpha)) = T_{(p,\eta)}T^S M\times\mathbb{R} = (L_1\oplus L_2)\times\mathbb{R}\ . \tag{4.37}$$

Here, $L_1$ and $L_2$ are defined via

$$L_1 := \{(\zeta,-W_{p,\eta}(\zeta))\mid \zeta\in T_pM\}\ ,\quad L_2 := \{0\}\times(T_p^\perp M\cap p^\perp\cap\eta^\perp)\ ,$$

where $W_{p,\eta}\colon T_pM\to T_pM$ denotes the Weingarten map of $M$ at $p$ in direction $\eta$. If $\zeta_1,\dots,\zeta_d$ is an orthonormal basis of $T_pM$ consisting of principal directions, i.e., $W_{p,\eta}(\zeta_i)=\kappa_i(p,\eta)\,\zeta_i$, then the vectors $\xi_1,\dots,\xi_d$, where

$$\xi_i := (\zeta_i,-\kappa_i(p,\eta)\,\zeta_i)\in L_1\ ,$$

describe an orthogonal basis of $L_1$. Note that $\|\xi_i\|=\sqrt{1+\kappa_i(p,\eta)^2}$. Furthermore, if $\eta_1,\dots,\eta_{n-d-2}$ denotes an orthonormal basis of $T_p^\perp M\cap p^\perp\cap\eta^\perp$, then

$$(0,\eta_1)\ ,\ \dots\ ,\ (0,\eta_{n-d-2})\in L_2$$

is an orthonormal basis of $L_2$.

As for the derivative of the map $\varphi\colon(p,\eta,\rho)\mapsto\cos(\rho)\cdot p+\sin(\rho)\cdot\eta$, note that for fixed $\rho\in(0,\alpha)$, the map

$$\tilde\varphi\colon N^S M\to\mathcal{T}(\alpha)\ ,\quad (p,\eta)\mapsto\cos(\rho)\cdot p+\sin(\rho)\cdot\eta$$

is linear. Therefore, we get

$$\begin{aligned}
D_{(p,\eta,\rho)}\varphi(\xi_i,0) &= D_{(p,\eta)}\tilde\varphi(\zeta_i,-\kappa_i(p,\eta)\,\zeta_i) &&= (\cos(\rho)-\sin(\rho)\cdot\kappa_i(p,\eta))\cdot\zeta_i\ ,\\
D_{(p,\eta,\rho)}\varphi(0,\eta_j,0) &= D_{(p,\eta)}\tilde\varphi(0,\eta_j) &&= \sin(\rho)\cdot\eta_j\ ,\\
D_{(p,\eta,\rho)}\varphi(0,0,1) &= \tfrac{d}{d\rho}(\cos(\rho)\cdot p+\sin(\rho)\cdot\eta) = -\sin(\rho)\cdot p+\cos(\rho)\cdot\eta\ ,
\end{aligned}$$

where $1\le i\le d$ and $1\le j\le n-d-2$. Note that the $n-1$ vectors

$$\zeta_1,\dots,\zeta_d\ ,\ \eta_1,\dots,\eta_{n-d-2}\ ,\ -\sin(\rho)\cdot p+\cos(\rho)\cdot\eta$$

are orthonormal. Note also that we have $-\kappa_i(p,\eta)=\kappa_i(p,-\eta)$. Therefore, the Normal Jacobian of $D_{(p,\eta,\rho)}\varphi$ is the absolute value of the determinant of the matrix

$$\left(\begin{array}{ccc|ccc|c}
\frac{\cos(\rho)+\sin(\rho)\cdot\kappa_1(p,-\eta)}{\sqrt{1+\kappa_1(p,\eta)^2}} & & & & & & \\
 & \ddots & & & 0 & & 0\\
 & & \frac{\cos(\rho)+\sin(\rho)\cdot\kappa_d(p,-\eta)}{\sqrt{1+\kappa_d(p,\eta)^2}} & & & & \\
\hline
 & & & \sin(\rho) & & & \\
 & 0 & & & \ddots & & 0\\
 & & & & & \sin(\rho) & \\
\hline
 & 0 & & & 0 & & 1
\end{array}\right).$$

As $\kappa_i(p,-\eta) \geq 0$ (cf. Remark 4.1.5), we have

$$\operatorname{ndet}(D_{(p,\eta,\rho)}\varphi) = \sin(\rho)^{n-d-2} \cdot \prod_{i=1}^{d} \frac{\cos(\rho) + \sin(\rho) \cdot \kappa_i(p,-\eta)}{\sqrt{1+\kappa_i(p,\eta)^2}} \,.$$

By Lemma 4.1.15 we get

$$\begin{aligned}
\operatorname{vol}_{n-1} \mathcal{T}(\alpha) &= \int\limits_{(p,\eta,\rho)\in N^S M\times(0,\alpha)} \operatorname{ndet}(D_{(p,\eta,\rho)}\varphi)\; d(p,\eta,\rho) \\
&= \int\limits_{(p,\eta)\in N^S M} \int_0^\alpha \sin(\rho)^{n-d-2} \cdot \prod_{i=1}^{d} \frac{\cos(\rho) + \sin(\rho)\cdot \kappa_i(p,-\eta)}{\sqrt{1+\kappa_i(p,\eta)^2}}\; d\rho\; d(p,\eta) \,. \qquad (4.38)
\end{aligned}$$

By Lemma 4.1.17 and another application of the coarea formula in Lemma 4.1.15 to the projection $N^S M \to M$, we may continue as

$$\begin{aligned}
(4.38) &= \int\limits_{p\in M} \int\limits_{\eta\in N_p^S(K)} \int_0^\alpha \sin(\rho)^{n-d-2} \cdot \prod_{i=1}^{d} (\cos(\rho) + \sin(\rho)\cdot\kappa_i(p,-\eta))\; d\rho\; d\eta\; dp \\
&= \int\limits_{p\in M} \int\limits_{\eta\in N_p^S(K)} \int_0^\alpha \sin(\rho)^{n-d-2} \cdot \sum_{j=0}^{d} \cos(\rho)^j \cdot \sin(\rho)^{d-j} \cdot \sigma_{d-j}(p,-\eta)\, d\rho\, d\eta\, dp \\
&\overset{(4.28)}{=} \sum_{j=0}^{d} I_{n,j}(\alpha) \cdot \int\limits_{p\in M} \int\limits_{\eta\in N_p^S(K)} \sigma_{d-j}(p,-\eta)\; d\rho\; d\eta\; dp \,.
\end{aligned}$$

This shows (4.36) and thus finishes the proof. □

## 4.4 Spherical intrinsic volumes

In this section we will describe the notion of spherical intrinsic volume. We rely in our presentation on [30] and [29].

Central for the definition of the spherical intrinsic volumes is the following statement about the decomposition of the volume of the tube around a convex set.

**Proposition 4.4.1.** *For $K \in \mathcal{K}(S^{n-1})$ and $0 \leq \alpha \leq \frac{\pi}{2}$ the volume of the $\alpha$-tube around $K$ is given by*

$$\operatorname{vol}_{n-1} \mathcal{T}(K,\alpha) = \operatorname{vol}_{n-1}(K) + \sum_{j=0}^{n-2} V_j(K) \cdot \mathcal{O}_{n-1,j}(\alpha) \,, \qquad (4.39)$$

*for some continuous functions $V_j \colon \mathcal{K}(S^{n-1}) \to \mathbb{R}$, $0 \leq j \leq n-2$.*

Note that if the volume of the tube, $\operatorname{vol}_{n-1} \mathcal{T}(K,\alpha)$, is of the form as stated in (4.39), then the values $V_j(K)$, $0 \leq j \leq n-2$, are uniquely determined. This follows from the linear independence of the functions $\mathcal{O}_{n-1,j}$ and the constant function (cf. Proposition 4.3.1). For the proof of Proposition 4.4.1 we need the following simple fact about the continuity of the volume of the tube.

**Lemma 4.4.2.** *For $\alpha \geq 0$ the function*

$$\mathcal{K}(S^{n-1}) \to \mathbb{R}\,, \quad K \mapsto \mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) \tag{4.40}$$

*is uniformly continuous.*

*Proof.* Let $(K_i)_i$ be a sequence of spherical convex sets in $\mathcal{K}(S^{n-1})$, which converges to $K \in \mathcal{K}(S^{n-1})$. It suffices to show that $\mathrm{vol}_{n-1}\,\mathcal{T}(K_i,\alpha)$ converges to $\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha)$. This implies that the function in (4.40) is continuous, and thus uniformly continuous, as $\mathcal{K}(S^{n-1})$ is a compact metric space (cf. Proposition 3.2.3).

For all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d_H(K, K_i) < \varepsilon$, i.e., $K \subseteq \mathcal{T}(K_i,\varepsilon)$ and $K_i \subseteq \mathcal{T}(K,\varepsilon)$, for all $i \geq N$. This implies that $\mathcal{T}(K,\alpha) \subseteq \mathcal{T}(K_i,\alpha+\varepsilon)$ and $\mathcal{T}(K_i,\alpha) \subseteq \mathcal{T}(K,\alpha+\varepsilon)$ for all $i \geq N$. In particular, we have $\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) \leq \mathrm{vol}_{n-1}\,\mathcal{T}(K_i,\alpha+\varepsilon)$ and $\mathrm{vol}_{n-1}\,\mathcal{T}(K_i,\alpha) \leq \mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha+\varepsilon)$ for all $i \geq N$. We thus get

$$\begin{aligned} \mathrm{vol}\,\mathcal{T}(K,\alpha) \;&\leq\; \liminf_{i\to\infty} \mathrm{vol}\,\mathcal{T}(K_i,\alpha+\varepsilon)\,, \\ \limsup_{i\to\infty} \mathrm{vol}\,\mathcal{T}(K_i,\alpha) \;&\leq\; \mathrm{vol}\,\mathcal{T}(K,\alpha+\varepsilon)\,. \end{aligned}$$

Letting $\varepsilon \to 0$, we get $\lim_{i\to\infty} \lim \mathrm{vol}\,\mathcal{T}(K_i,\alpha) = \mathrm{vol}\,\mathcal{T}(K,\alpha)$. □

*Proof of Proposition 4.4.1.* From Weyl's tube formula in Theorem 4.3.2 we know that (4.39) holds for $K \in \mathcal{K}^{\mathrm{str}}(S^{n-1})$, which is a dense subset of $\mathcal{K}(S^{n-1})$. The idea is to use the uniform continuity of the map (4.40) to extend the validity of (4.39) to all of $\mathcal{K}(S^{n-1})$. Let

$$f_1 := \mathbf{1}\,,\; f_2 := \mathcal{O}_{n-1,0}\,,\; f_3 := \mathcal{O}_{n-1,1}\,,\; \ldots\,,\; f_n := \mathcal{O}_{n-1,n-2}\; \in \mathbb{R}^{[0,\pi/2]}\,,$$

where $\mathbb{R}^{[0,\pi/2]} := \{f\colon [0,\frac{\pi}{2}] \to \mathbb{R}\}$, and $\mathbf{1}$ denotes the constant function $\mathbf{1}(\alpha) = 1$. The fact that $f_1,\ldots,f_n$ are linear independent (cf. Proposition 4.3.1) implies that there are $\alpha_1,\ldots,\alpha_n \in [0,\frac{\pi}{2}]$ such that the matrix $A \in \mathbb{R}^{n\times n}$, where the $i$th row is given by $(f_1(\alpha_i), f_2(\alpha_i),\ldots,f_n(\alpha_i))$, is nonsingular. (This can be proved by induction on $n$, similar to the finite-dimensional case.) Let $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ denote the inverse of the matrix $A$. For $K \in \mathcal{K}(S^{n-1})$ let $g_K$ denote the function

$$g_K\colon [0,\tfrac{\pi}{2}] \to \mathbb{R}\,, \quad g_K(\alpha) := \mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha)\,.$$

For a stratified cap $K \in \mathcal{K}^{\mathrm{str}}(S^{n-1})$ Theorem 4.3.2 implies that $g_K$ lies in the span of the functions $f_1,\ldots,f_n$, so that we have

$$g_K(\alpha) = \mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) = \sum_{i=1}^{n} c_i(K)\cdot f_i(\alpha)$$

for some constants $c_1(K),\ldots,c_n(K)$ depending on $K$. Using the matrix $B$, the constants $c_i(K)$ can be expressed in the values of $g_K$ at $\alpha_1,\ldots,\alpha_n$ via

$$c_i(K) = \sum_{j=1}^{n} b_{ij}\cdot g_K(\alpha_j)\,.$$

By Lemma 4.4.2 it follows that the functions $c_i\colon \mathcal{K}^{\mathrm{str}}(S^{n-1}) \to \mathbb{R}$ are uniformly continuous and thus have a unique continuous extension $c_i\colon \mathcal{K}(S^{n-1}) \to \mathbb{R}$. Furthermore, for $K \in \mathcal{K}(S^{n-1})$ and $K_\ell \in \mathcal{K}^{\mathrm{str}}(S^{n-1})$ such that $K_\ell \to K$ for $\ell \to \infty$ we have

$$\mathrm{vol}_{n-1}\,\mathcal{T}(K,\alpha) = \lim_{\ell\to\infty} g_{K_\ell}(\alpha) = \sum_{i=1}^{n} \lim_{\ell\to\infty} (c_i(K_\ell))\cdot f_i(\alpha) = \sum_{i=1}^{n} c_i(K)\cdot f_i(\alpha)\,,$$

for all $\alpha \in [0, \frac{\pi}{2}]$. Defining $V_j(K) := c_{j+2}(K)$ for $0 \leq j \leq n-2$ thus finishes the proof. □

**Definition 4.4.3.** For $-1 \leq j \leq n-1$ the $j$th *(spherical) intrinsic volume* is a function

$$V_j \colon \mathcal{K}(S^{n-1}) \cup \{\emptyset, S^{n-1}\} \to \mathbb{R} \,,$$

defined as follows. For $0 \leq j \leq n-2$ and $K \in \mathcal{K}(S^{n-1})$ the value of $V_j$ in $K$ is the quantity $V_j(K)$ defined in (4.39). Furthermore, for $j \in \{-1, n-1\}$ and $K \in \mathcal{K}(S^{n-1})$ the $j$th intrinsic volume of $K$ is defined as

$$V_{n-1}(K) := \frac{\mathrm{vol}_{n-1}(K)}{\mathcal{O}_{n-1}} \quad , \qquad V_{-1}(K) := \frac{\mathrm{vol}_{n-1}(\breve{K})}{\mathcal{O}_{n-1}} \,.$$

Lastly, for $K \in \{\emptyset, S^{n-1}\}$

$$V_j(\emptyset) := \begin{cases} 1 & \text{if } j = -1 \\ 0 & \text{else} \end{cases} \quad , \qquad V_j(S^{n-1}) := \begin{cases} 1 & \text{if } j = n-1 \\ 0 & \text{else} \end{cases} \,.$$

In the following proposition we collect the formulas for the intrinsic volumes that arise from Weyl's tube formula.

**Proposition 4.4.4.** *Let $K \in \mathcal{K}(S^{n-1})$ and $0 \leq j \leq n-2$.*

1. *If $K$ is a stratified spherical convex set with decomposition $K = \dot{\bigcup}_{i=0}^{\tilde{k}} M_i$, such that $M_0 = \mathrm{int}(K)$, and such that $M_1, \ldots, M_k$ are the essential and $M_{k+1}, \ldots, M_{\tilde{k}}$, $k \leq \tilde{k}$, are the negligible pieces, then*

$$V_j(K) \;=\; \frac{1}{\mathcal{O}_j \cdot \mathcal{O}_{n-2-j}} \cdot \sum_{i=1}^{k} \int_{p\in M_i} \int_{\eta \in N_p^S(K)} \sigma_{d_i-j}^{(i)}(p, -\eta)\, d\eta\, dp \,,$$

*where $d_i$ denotes the dimension of the stratum $M_i$, $\sigma_\ell^{(i)}(p, -\eta)$ denotes the $\ell$th elementary symmetric function in the principal curvatures of $M_i$ at $p$ in direction $-\eta$ (cf. (4.2)), and $\sigma_\ell(p, -\eta) := 0$ if $\ell < 0$.*

2. *If $K$ is a smooth cap with boundary $M := \partial K$, then*

$$V_j(K) \;=\; \frac{1}{\mathcal{O}_j \cdot \mathcal{O}_{n-2-j}} \cdot \int_{p\in M} \sigma_{n-2-j}(p)\, dM \,.$$

3. *If $K$ is a polyhedral cap with $j$-dimensional faces $\mathcal{F}_j$, then*

$$V_j(K) \;=\; \sum_{F \in \mathcal{F}_j} \frac{\mathrm{vol}_j(F)}{\mathcal{O}_j} \cdot \frac{\mathrm{vol}_{n-2-j}(N_F^S)}{\mathcal{O}_{n-2-j}} \,.$$

*Proof.* Follows from Theorem 4.3.2, Corollary 4.3.3, and Corollary 4.3.4. □

Before we continue with properties of the intrinsic volumes, let us consider the important special cases for $K$ being the intersection of the positive orthant with the unit sphere.

**Example 4.4.5.** Let $K = \mathbb{R}^n_+ \cap S^{n-1}$. Then from Example 4.3.6 we get for $0 \leq j \leq n-2$
$$V_j(K) = \frac{\binom{n}{j+1}}{2^n} \, .$$
This formula also holds for $j \in \{-1, n-1\}$ as is seen easily.

Note that the intrinsic volumes of the positive orthant coincide with the binomial distribution with $n$ trials and probability of success equal to $\frac{1}{2}$. The deeper reason for this curious fact is an alternative characterization for the intrinsic volumes of spherical polytopes that we give in Proposition 4.4.6. Before we state this, recall that the boundary of a polyhedral cone $C$ decomposes into the relative interiors of its faces (cf. Remark 3.1.8)
$$\partial C = \dot{\bigcup_{F \text{ face of } C}} \operatorname{relint}(F) \, .$$
So for $x \in C$ we can define
$$\operatorname{face}(x) := \begin{cases} C & \text{if } x \in \operatorname{int}(C) \\ F & \text{if } x \in \operatorname{relint}(F) \text{ and } F \text{ a face of } C. \end{cases} \tag{4.41}$$

**Proposition 4.4.6.** *Let $K \in \mathcal{K}^p(S^{n-1})$ be a spherical polytope and let $C := \operatorname{cone}(K)$ denote the corresponding polyhedral cone. Furthermore, let $\Pi_C \colon \mathbb{R}^n \to C$ denote the projection map onto $C$, and let $d_C$ denote the function*
$$d_C \colon \mathbb{R}^n \to \{0, 1, 2, \ldots, n\} \, , \quad x \mapsto \dim(\operatorname{face}(\Pi_C(x))) \, , \tag{4.42}$$
*where* $\operatorname{face}(x)$ *is defined as in* (4.41)*. Then the* $j$*th intrinsic volume of* $K$ *is given by*
$$V_j(K) = \operatorname*{Prob}_{p \in S^{n-1}} [d_C(p) = j+1] \, , \tag{4.43}$$
*where $p \in S^{n-1}$ is drawn uniformly at random.*

*Proof.* The claim is true for $j \in \{-1, n-1\}$, as
$$d_C(p) = n \iff (\dim C = n \text{ and } p \in C) \, ,$$
$$d_C(p) = 0 \iff (\dim \breve{C} = n \text{ and } p \in \breve{C}) \, .$$
In the remainder of the proof we therefore assume $0 \leq j \leq n-2$.

Let $F^e \subseteq C$ be a face of $C$ of dimension $\dim F^e = j+1$, and let $N = N_{F^e}(C)$ be the normal cone of $C$ in $F^e$. W.l.o.g. we may also assume that $\operatorname{lin}(F^e) = \mathbb{R}^{j+1} \times \{0\}$, so that
$$F^e = \tilde{F}^e \times \{0\} \quad \text{and} \quad N = \{0\} \times \tilde{N}$$
with $\tilde{F}^e \subseteq \mathbb{R}^{j+1}$ and $\tilde{N} \subseteq \mathbb{R}^{n-j-1}$. The inverse image of $F^e$ under the projection map $\Pi_C$ is given by $\Pi_C^{-1}(F^e) = \tilde{F}^e \times \tilde{N}$.

Now, if $x = (x_1, x_2) \in \mathbb{R}^{j+1} \times \mathbb{R}^{n-j-1}$ is drawn at random with respect to the $n$-dimensional normal distribution, i.e., $x \in \mathcal{N}(0, I_n)$, then we get
$$\begin{aligned} \operatorname*{Prob}_{x \in \mathcal{N}(0,I_n)} \left[x \in \Pi_C^{-1}(F^e)\right] &= \operatorname*{Prob}_{x_1 \in \mathcal{N}(0,I_{j+1})} \left[x_1 \in \tilde{F}^e\right] \cdot \operatorname*{Prob}_{x_2 \in \mathcal{N}(0,I_{n-j-1})} \left[x_2 \in \tilde{N}\right] \\ &= \frac{\operatorname{vol}(\tilde{F}^e \cap S^j)}{\mathcal{O}_j} \cdot \frac{\operatorname{vol}(\tilde{N} \cap S^{n-2-j})}{\mathcal{O}_{n-2-j}} \, . \end{aligned}$$

Altogether, we get

$$\begin{aligned}\operatorname*{Prob}_{p\in S^{n-1}}[d_C(p)=j+1] &= \operatorname*{Prob}_{x\in\mathcal{N}(0,I_n)}[d_C(x)=j+1]\\ &= \sum_{F\in\mathcal{F}_j(K)} \frac{\mathrm{vol}_j(F)}{\mathcal{O}_j}\cdot\frac{\mathrm{vol}_{n-2-j}(N_F^S)}{\mathcal{O}_{n-2-j}}\\ &\overset{\text{Prop. 4.4.4}}{=} V_j(K)\ . \end{aligned}$$

□

Proposition 4.4.6 explains why the intrinsic volumes of the positive orthant are given by a binomial distribution:

**Example 4.4.7.** Let $C = \mathbb{R}^n_+$ and $K = C\cap S^{n-1}$. For $x \in \mathbb{R}^n$ the projection $\Pi_C(x)$ is given by the element $\bar{x} = (\bar{x}_1,\dots,\bar{x}_n)$, where $\bar{x}_i = \begin{cases} x_i & \text{if } x_i \geq 0\\ 0 & \text{else}\end{cases}$. From this it is easily seen that the function $d_C$ from (4.42) is given by

$$d_C(x) = |\{i \mid x_i > 0\}|\ .$$

If $p \in S^{n-1}$ is drawn uniformly at random, then the probability that the $i$th component of $p$ is positive is given by $\frac{1}{2}$. Therefore, by (4.43) the $j$th intrinsic volume is given by

$$\begin{aligned}V_j(K) &= (\text{probability of } j+1 \text{ ‘heads’ when tossing } n \text{ fair coins})\\ &= \frac{\binom{n}{j+1}}{2^n}\ .\end{aligned}$$

Another important example is where $K \subset S^{n-1}$ is a circular cap. We will consider this in the next example.

**Example 4.4.8.** Let $K = B(z,\beta)$, $0 < \beta \leq \pi/2$, be a circular cap. Then from Example 4.3.5 we get for $0\leq j\leq n-2$

$$\begin{aligned}V_j(K) &= \frac{\mathcal{O}_{n-2}}{\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}}\cdot\binom{n-2}{j}\cdot\sin(\beta)^j\cdot\cos(\beta)^{n-2-j}\\ &\overset{(4.24)}{=} \binom{(n-2)/2}{j/2}\cdot\frac{\sin(\beta)^j\cdot\cos(\beta)^{n-2-j}}{2}\ . \end{aligned}\tag{4.44}$$

For Lorentz caps, i.e., $\beta = \frac{\pi}{4}$, we have $\sin(\frac{\pi}{4}) = \cos(\frac{\pi}{4}) = \frac{1}{\sqrt{2}}$, and (4.44) simplifies further to

$$V_j(K) = \frac{\binom{(n-2)/2}{j/2}}{2^{n/2}}\ .$$

Furthermore, we get from Proposition 4.1.18 that

$$\begin{aligned}V_{n-1}(K) &= \frac{\mathcal{O}_{n-2}}{\mathcal{O}_{n-1}}\cdot\int_0^\beta \sin(\rho)^{n-2}\,d\rho\\ &\overset{(4.18)}{=} \binom{(n-2)/2}{(n-1)/2}\cdot\frac{n-1}{2}\cdot\int_0^\beta \sin(\rho)^{n-2}\,d\rho\ ,\end{aligned}$$

and similarly

$$V_{-1}(K) = \binom{(n-2)/2}{-1/2}\cdot\frac{n-1}{2}\cdot\int_0^{\frac{\pi}{2}-\beta}\sin(\rho)^{n-2}\,d\rho\ .$$

For $\beta = \frac{\pi}{4}$ and using the hypergeometric function (cf. Remark D.3.3 in Section D.3 in the appendix) this simplifies to

$$\begin{aligned} V_{n-1}(K) &= V_{-1}(K) \\ &= \frac{\binom{(n-2)/2}{-1/2}}{2^{n/2}} \cdot \frac{{}_2F_1(\frac{n-1}{2}, \frac{1}{2}; \frac{n+1}{2}; \frac{1}{2})}{\sqrt{2}} \qquad \left[\sim \frac{\binom{(n-2)/2}{-1/2}}{2^{n/2}} \ , \ \text{for } n \to \infty\right] . \end{aligned}$$

An important property of the intrinsic volumes of circular caps is their log-concavity. We show this in the following proposition.[2]

**Proposition 4.4.9.** *Let $K = B(z, \beta)$, $0 \le \beta \le \pi/2$, be a circular cap. Then the sequence $V_{-1}(K), V_0(K), \ldots, V_{n-1}(K)$ is log-concave and has no internal zeros.*

*Proof.* From (4.44) we get for $1 \le j \le n-3$

$$\frac{V_j(K)^2}{V_{j-1}(K) \cdot V_{j+1}(K)} = \frac{\binom{(n-2)/2}{j/2}^2}{\binom{(n-2)/2}{(j-1)/2} \cdot \binom{(n-2)/2}{(j+1)/2}} \ge 1 \ ,$$

where the inequality follows from Proposition 4.1.23. For $j = n-2$ we compute

$$\begin{aligned} \frac{V_{n-2}(K)^2}{V_{n-3}(K) \cdot V_{n-1}(K)} &= \underbrace{\frac{\binom{(n-2)/2}{(n-2)/2}^2}{\binom{(n-2)/2}{(n-3)/2} \cdot \binom{(n-2)/2}{(n-1)/2}}}_{\ge 1} \cdot \frac{\sin(\beta)^{n-1}}{\cos(\beta) \cdot (n-1) \cdot \int_0^\beta \sin(\rho)^{n-2} \, d\rho} \\ &\ge \frac{\sin(\beta)^{n-1}}{(n-1) \cdot \int_0^\beta \cos(\rho) \cdot \sin(\rho)^{n-2} \, d\rho} = 1 \ , \end{aligned}$$

and similarly for $j = 0$. Finally, as $V_j(K) > 0$ for all $-1 \le j \le n-1$, the sequence in particular has no internal zeros. □

The following proposition lists some elementary properties of the intrinsic volumes.

**Proposition 4.4.10.** *1. The intrinsic volumes are nonnegative, i.e., $V_j(K) \ge 0$ for all $-1 \le j \le n-1$ and for all $K \in \mathcal{K}(S^{n-1}) \cup \{\emptyset, S^{n-1}\}$.*

*2. If $S \in \mathcal{S}^i(S^{n-1})$ then $V_j(S) = \left\{ \begin{matrix} 1 & \text{if } j = i \\ 0 & \text{else} \end{matrix} \right\}$.*

*3. For $K \in \mathcal{K}(S^{n-1})$ the intrinsic volumes of the dual $\breve{K}$ are given by*

$$V_j(K) = V_{n-2-j}(\breve{K}) \ .$$

*4. For all $K \in \mathcal{K}(S^{n-1}) \cup \{\emptyset, S^{n-1}\}$*

$$\sum_{j=-1}^{n-1} V_j(K) = 1 \ .$$

*In particular, $V_j(K) \le 1$.*

[2] Also the intrinsic volumes of the positive orthant form a log-concave sequence. But as the positive orthant is only a special second-order cone, this observation also follows from Corollary 4.4.14.

5. *For all convex caps $K \in \mathcal{K}^c(S^{n-1})$*

$$\sum_{\substack{j=-1\\ j\equiv 0 \bmod 2}}^{n-1} V_j(K) = \sum_{\substack{j=-1\\ j\equiv 1 \bmod 2}}^{n-1} V_j(K) = \tfrac{1}{2}\ .$$

*In particular, if $K$ is a cap then $V_j(K) \leq \frac{1}{2}$.*

*Proof.* (1) For polytopes $K \in \mathcal{K}^p(S^{n-1})$ the intrinsic volumes are nonnegative, which is immediate from Proposition 4.4.4. For $K \in \mathcal{K}(S^{n-1})$ the nonnegativity of the intrinsic volumes follows from their continuity and from the fact that the polytopes $\mathcal{K}^p(S^{n-1})$ form a dense subset of $\mathcal{K}(S^{n-1})$.

(2) The claim about the intrinsic volumes of subspheres follows directly from the definition of the intrinsic volumes.

(3) Recall that for $K \in \mathcal{K}^p(S^{n-1})$ we have a bijection between the $j$-dimensional faces of $K$ and the $(n-2-j)$-dimensional faces of $\breve{K}$ (cf. Lemma 3.1.9). Using this it is straightforward to deduce $V_j(K) = V_{n-2-j}(\breve{K})$ for $K \in \mathcal{K}^p(S^{n-1})$ from Proposition 4.4.4. The general validity of this identity follows from the continuity of the intrinsic volumes and an approximation argument as above (cf. Proposition 3.2.4). Alternatively, one may use Proposition 4.1.11 to first show $V_j(K) = V_{n-2-j}(\breve{K})$ for $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and then apply the approximation argument.

(4) The fact that the intrinsic volumes add up to 1 either follows from Proposition 4.4.6 or from the general fact $\mathcal{T}(K, \frac{\pi}{2}) = S^{n-1} \setminus \mathrm{int}(\breve{K})$, which implies

$$\mathcal{O}_{n-1} - \mathrm{vol}_{n-1}\breve{K} = \mathrm{vol}_{n-1}\mathcal{T}(K, \tfrac{\pi}{2}) = \mathrm{vol}_{n-1}(K) + \sum_{j=0}^{n-2} \underbrace{\mathcal{O}_{n-1,j}(\tfrac{\pi}{2})}_{=\mathcal{O}_{n-1}} \cdot V_j(K)\ .$$

(5) Finally, the last statement follows from the Gauss-Bonnet formula in spherical space. See [30, Sec. 4.3] or the references given in [29, p. 5] for proofs of this fact. □

Next, we will state the principal kinematic formula in a similar form as given in [12, Sec. 1.7.3]. To achieve this we define the *intrinsic volume polynomial* and the *reverse volume polynomial* via

$$V(K;X) := V_{-1}(K) + V_0(K)\cdot X + \ldots + V_{n-1}(K)\cdot X^n\ , \tag{4.45}$$

$$V^{\mathrm{rev}}(K;X) := V_{n-1}(K) + V_{n-2}(K)\cdot X + \ldots + V_{-1}(K)\cdot X^n\ , \tag{4.46}$$

where $X$ denotes a formal variable. Note that while $V(K;X)$ is independent of whether $K$ is considered as an element in $\mathcal{K}(S^{n-1})$ or in $\mathcal{K}(S^{N-1})$ for some $N \geq n$, this is not the case for the reverse volume polynomial $V^{\mathrm{rev}}(K;X)$. For example, if $S \in \mathcal{S}^k(S^{n-1})$ (cf. (3.4)), then $V(S;X) = X^{k+1}$, while $V^{\mathrm{rev}}(S;X) = X^{n-k-1}$.

Note also that from $V_j(K) = V_{n-2-j}(\breve{K})$ (cf. Proposition 4.4.10) we get

$$V(\breve{K};X) = V^{\mathrm{rev}}(K;X)\ ,\quad V^{\mathrm{rev}}(\breve{K};X) = V(K;X)\ .$$

**Theorem 4.4.11** (Principal kinematic formula)**.** *Let $K, K' \in \mathcal{K}(S^{n-1})$, and let $Q \in O(n)$ be chosen uniformly at random. Then the expected reverse volume polynomial of the intersection $K \cap (Q\cdot K')$ is given* $\mathrm{mod}\, X^n$ *by*

$$\mathbb{E}\,[V^{\mathrm{rev}}(K \cap (Q\cdot K');X)] \;\equiv\; V^{\mathrm{rev}}(K;X)\cdot V^{\mathrm{rev}}(K';X) \mod X^n\ .$$

*For the intrinsic volumes this means that for* $0 \leq j \leq n-1$

$$\begin{aligned} \mathbb{E}\left[V_j(K \cap (Q \cdot K'))\right] &= \sum_{i=j}^{n-1} V_i(K) \cdot V_{n-1+j-i}(K') \quad \textit{and} \\ \mathbb{E}\left[V_{-1}(K \cap (Q \cdot K'))\right] &= 1 - \sum_{j=0}^{n-1} \sum_{i=j}^{n-1} V_i(K) \cdot V_{n-1+j-i}(K') \;. \end{aligned}$$

*In particular, if* $S \in \mathcal{S}^k(S^{n-1})$ *is chosen uniformly at random, then for* $0 \leq j \leq n-1$

$$\begin{aligned} \mathbb{E}\left[V_j(K \cap S)\right] &= \begin{cases} V_{j+n-1-k}(K) & \textit{if } j \leq k \\ 0 & \textit{if } j > k \;, \end{cases} \\ \mathbb{E}\left[V_{-1}(K \cap S)\right] &= V_{-1}(K) + \ldots + V_{n-2-k}(K) \;. \end{aligned}$$

*Proof.* See for example [30, Ch. 5] or the references given in [12, Sec. 1.7.3]. □

As a corollary we get the following result about the probability that a randomly chosen subspace intersects a given cap. In terms of the convex feasibility problem, the following result provides a formula for the probability that a random instance (defined by a matrix with i.i.d. normal distributed entries) is feasible.

**Corollary 4.4.12.** *Let* $K \in \mathcal{K}^c(S^{n-1})$ *and let* $S \in \mathcal{S}^k(S^{n-1})$, $0 \leq k \leq n-2$, *be chosen uniformly at random. Then*

$$\mathrm{Prob}[K \cap S \neq \emptyset] = 2 \cdot \sum_{\substack{j=n-1-k \\ j \equiv n-1-k \bmod 2}}^{n-1} V_j(K) \;.$$

*For the special case* $K = \mathbb{R}^n_+ \cap S^{n-1}$ *we have*

$$\mathrm{Prob}[K \cap S \neq \emptyset] = \frac{1}{2^{n-1}} \cdot \sum_{j=0}^{k} \binom{n-1}{j} \;.$$

The second statement is usually attributed to Wendel [62], who proved it in a direct way.

*Proof.* First, note that the intersection $K \cap S$ is either empty or again a convex cap. Since $V_j(\emptyset) = 0$ if $j \neq -1$ and $V_{-1}(\emptyset) = 1$, we get from Proposition 4.4.10 part (5) for all $S \in \mathcal{S}^k(S^{n-1})$

$$2 \cdot \sum_{\substack{j=-1 \\ j \equiv 0 \bmod 2}}^{n-1} V_j(K \cap S) = \begin{cases} 1 & \text{if } K \cap S \neq \emptyset \\ 0 & \text{if } K \cap S = \emptyset \end{cases} \;. \tag{4.47}$$

From Theorem 4.4.11 we get for $S \in \mathcal{S}^k(S^{n-1})$ chosen uniformly at random

$$\mathbb{E}[V_j(K \cap S)] = \begin{cases} V_{j+n-1-k}(K) & \text{if } j \leq k \\ 0 & \text{if } j > k \end{cases} \;.$$

Therefore, taking the expectation of both sides in (4.47) yields

$$\begin{aligned}\mathrm{Prob}[K\cap S\neq\emptyset] &= 2\cdot\sum_{\substack{j=-1\\ j\equiv 0 \bmod 2}}^{n-1}\mathbb{E}[V_j(K\cap S)] = 2\cdot\sum_{\substack{j=0\\ j\equiv 0 \bmod 2}}^{k} V_{j+n-1-k}(K)\\ &= 2\cdot\sum_{\substack{\ell=n-1-k\\ \ell\equiv n-1-k \bmod 2}}^{n-1} V_\ell(K)\ ,\end{aligned}$$

where we changed the summation by using $\ell := n-1+j-k$.

For the special case $K=\mathbb{R}^n_+\cap S^{n-1}$ recall from Example 4.4.7 that $V_\ell(K)=\binom{n}{\ell+1}/2^n$. Changing the summation again by setting $i:=n-1-\ell$, we get

$$\mathrm{Prob}[K\cap S\neq\emptyset] = 2\cdot\sum_{\substack{i=0\\ i\equiv k \bmod 2}}^{k}\frac{\binom{n}{n-1-i+1}}{2^n} = \sum_{\substack{i=0\\ i\equiv k \bmod 2}}^{k}\frac{\binom{n}{i}}{2^{n-1}}\ .$$

Using the identity $\binom{n}{i}=\binom{n-1}{i-1}+\binom{n-1}{i}$ we get

$$\sum_{\substack{i=0\\ i\equiv k \bmod 2}}^{k}\binom{n}{i} = \sum_{i=0}^{k}\binom{n-1}{i}\ ,$$

which finishes the proof. □

In the realm of convex optimization the direct product construction $C_1\times\ldots\times C_k$, where each $C_i$ is a closed convex cone, is an important tool. It is therefore of particular interest to have a simple formula for the intrinsic volumes of products in terms of the intrinsic volumes of the components.

Fortunately, such a general rule exists. It says that the intrinsic volumes of a product of caps arise from the intrinsic volumes of the components via convolution. In fact, this holds for the spherical case as well as for the euclidean case.

The euclidean case has been treated in [36, Sec. 9.7]. The formula for the spherical case seems to be new. We will give proofs for both calculation rules in the appendix (although the statement is in the euclidean case not new, the proof we give in the appendix is new).

Analogous to the definition of the spherical intrinsic volume polynomial we define for $K^e\in\mathcal{K}(\mathbb{R}^n)$ the euclidean intrinsic volume polynomial $V^e(K^e;X)$ via

$$V^e(K^e;X) := \sum_{j=0}^{n} V_j^e(K^e)\cdot X^j\ , \tag{4.48}$$

where $X$ denotes a formal variable.

**Proposition 4.4.13.** *1. Let $K_1^e, K_2^e$ be convex bodies. Then*

$$V^e(K_1^e\times K_2^e;X) = V^e(K_1^e;X)\cdot V^e(K_2^e;X)\ . \tag{4.49}$$

*2. Let $K_1, K_2$ be spherical convex sets. Then*

$$V(K_1\circledast K_2;X) = V(K_1;X)\cdot V(K_2;X) \tag{4.50}$$

*(cf. Section 3.1.1 for the definition of $K_1\circledast K_2$).*

*Proof.* See Section B.2 and Section B.1 (a proof for the euclidean statement can also be found in [36, Sec. 9.7]). □

**Corollary 4.4.14.** *1. Let $K_1, K_2$ be spherical convex sets such that the sequences of their intrinsic volumes are log-concave and with no internal zeros each. Then also the sequence of intrinsic volumes of $K := K_1 \circledast K_2$ is log-concave and has no internal zeros.*

*2. If $C = \mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k}$ is a second-order cone, then the intrinsic volumes of $K := C \cap S^{n-1}$ form a log-concave sequence with no internal zeros.*

*3. Let $K_1, K_2$ be self-dual caps such that the sequences of their intrinsic volumes are unimodal each. Then also the sequence of intrinsic volumes of $K := K_1 \circledast K_2$ is unimodal.*

*Proof.* Parts (1) and (3) follow from Proposition 4.4.13 and Proposition 4.1.22. Part (2) additionally follows from Proposition 4.4.9, where it was shown that the intrinsic volumes of a circular cap form a log-concave sequence with no internal zeros. □

**Remark 4.4.15.** Note that the calculation rule from Proposition 4.4.13 provides yet another way to compute the intrinsic volumes of the positive orthant: For $C_1 := \mathbb{R}_+$, $K_1 := C_1 \cap S^0$, we have $\mathrm{vol}_0(K_1) = \mathrm{vol}_0(\breve{K}_1) = 1$ and therefore $V(K_1; X) = \frac{1}{2} \cdot X + \frac{1}{2}$. As $\mathbb{R}_+^n = \mathbb{R}_+ \times \ldots \times \mathbb{R}_+$ we may apply (4.50) to get

$$V(K; X) = V(K_1; X)^n = \frac{1}{2^n} \cdot (X+1)^n = \frac{1}{2^n} \cdot \sum_{j=0}^{n} \binom{n}{j} \cdot X^j \ ,$$

where $K := \mathbb{R}_+^n \cap S^{n-1}$, yielding $V_{j-1}(K) = \binom{n}{j}/2^n$.

In Section 4.2 we have seen that the sequence of euclidean intrinsic volumes is log-concave. We formulate the spherical analog as a conjecture that (if true) would prove very useful for the average analysis of the Grassmann condition in Section 7.1.

**Conjecture 4.4.16.** For $K \in \mathcal{K}(S^{n-1})$ the sequence $V_{-1}(K), V_0(K), \ldots, V_{n-1}(K)$ is log-concave, i.e.

$$V_i(K)^2 \geq V_{i-1}(K) \cdot V_{i+1}(K) \ , \quad \text{for } i = 0, \ldots, n-2 \ .$$

In fact, for the average analysis in Section 7.1 it would be enough to know that the following weaker conjecture is true.

**Conjecture 4.4.17.** For $K \subseteq S^{n-1}$ a self-dual cap the sequence of intrinsic volumes $V_0(K), \ldots, V_{n-2}(K)$ is unimodal, i.e., we have

$$V_0(K) \leq \ldots \leq V_{\lfloor \frac{n}{2} \rfloor}(K) = V_{\lceil \frac{n}{2} \rceil}(K) \geq \ldots \geq V_{n-2}(K) \ .$$

The last topic of this section is about the relation between euclidean and spherical intrinsic volumes. Note that if we have a closed convex cone $C$, then instead of intersecting $C$ with the unit sphere $S^{n-1}$ we may as well intersect $C$ with the unit ball $B_n$. The result is not a spherical convex set but a (euclidean) convex body. Thus, let

$$K := C \cap S^{n-1} \ , \quad K^e := C \cap B_n \ ,$$

where $B_n$ denotes the $n$-dimensional unit ball. The following proposition provides a link between the euclidean intrinsic volumes of $K^e$ and the spherical intrinsic volumes of $K$. This result also appears to be new. We will give the proof in the appendix.

**Proposition 4.4.18.** *Let $C \subseteq \mathbb{R}^n$ be a closed convex cone, and let $K := C \cap S^{n-1}$ and $K^e := C \cap B_n$. Then*

$$V_i^e(K^e) \;=\; \sum_{j=i}^{n} V_i^e(B_j) \cdot V_{j-1}(K) \;=\; \sum_{j=i}^{n} \binom{j}{i} \cdot \frac{\omega_j}{\omega_{j-i}} \cdot V_{j-1}(K) \,.$$

*Proof.* The second equality follows from Example 4.2.1. We will give the rest of the proof in Section B.3. □

The following remarks shall comment on the natural approach of using the link between the euclidean and spherical intrinsic volumes provided by Proposition 4.4.18 to attack Conjecture 4.4.16. In a nutshell, we have come to the conclusion that Proposition 4.4.18 does not help to get a proof for Conjecture 4.4.16.

**Remark 4.4.19.** It is not possible to deduce Conjecture 4.4.16 solely from the log-concavity of the euclidean intrinsic volumes and the transformation formula from Proposition 4.4.18. More precisely, if $a_0, \ldots, a_n$ is a log-concave sequence of positive numbers, and if $b_0, \ldots, b_n$ and $c_0, \ldots, c_n$ are defined via

$$a_i = \sum_{j=i}^{n} \binom{j}{i} \cdot b_j \;, \quad a_i = \sum_{j=i}^{n} \binom{j}{i} \cdot \frac{\omega_j}{\omega_i} \cdot c_j \;, \quad i = 0, \ldots, n \;,$$

then neither the sequence $b_0, \ldots, b_n$ nor the sequence $c_0, \ldots, c_n$ need to be log-concave. Counter-example: Take $a = (\exp(3), \exp(3), \exp(2.5), \exp(1))$. Then $b \approx (9.46, 3.88, 4.03, 2.72)$ and $c \approx (6.57, 4.05, 6.75, 2.72)$, and neither $b$ nor $c$ is log-concave.

However, there probably is a connection between the log-concavity of the euclidean and the spherical intrinsic volumes. Unfortunately, even if this connection exists, it goes in the 'wrong' direction.

**Remark 4.4.20.** In [10, Thm. 2.5.3] it is shown that if $a_0, \ldots, a_n$ is a log-concave sequence of positive numbers, then the sequence $b_0, \ldots, b_n$ defined by

$$b_i := \sum_{j=i}^{n} \binom{j}{i} \cdot a_j$$

is also log-concave. It is tempting to believe that also the sequence $c_0, \ldots, c_n$ defined by

$$c_i := \sum_{j=i}^{n} \binom{j}{i} \cdot \frac{\omega_j}{\omega_{j-i}} \cdot a_j$$

is log-concave. If this were true, then a positive answer to Conjecture 4.4.16 would imply that the sequence $V_i^e(K^e)$, $i = 0, \ldots, n$, where $K^e = C \cap B_n$ and $C \subseteq \mathbb{R}^n$ a closed convex cone, is log-concave. But this is true by the Alexandrov-Fenchel inequality (cf. Proposition 4.2.2).

### 4.4.1 Intrinsic volumes of the semidefinite cone

In this section we will state the formulas for the intrinsic volumes of the semidefinite cone. We will provide a more detailed discussion of the occurring formulas, as well as the proofs for all statements in this section in Chapter C in the appendix.

The semidefinite cone is defined in the euclidean space

$$\mathrm{Sym}^n = \{A \in \mathbb{R}^{n\times n} \mid A^T = A\} \;,$$

which is a linear subspace of $\mathbb{R}^{n\times n}$ of dimension $\frac{n(n+1)}{2} =: t(n)$. Let $S(\mathrm{Sym}^n)$ denote the unit sphere in this subspace.

In order to get nice formulas for the intrinsic volumes we introduce some more notation. For $z = (z_1, \ldots, z_n)$ we denote the Vandermonde determinant by

$$\Delta(z) := \prod_{1\le i<j\le n} (z_i - z_j) \;.$$

For $0 \le r \le n$ we can decompose the Vandermonde determinant into (cf. (C.5) in Section C.1)

$$\Delta(z) = \sum_{\ell=0}^{r\cdot(n-r)} (-1)^\ell \cdot \Delta_{r,\ell}(z) \;,$$

where

$$\Delta_{r,\ell}(z) \;:=\; \Delta(x)\cdot\Delta(y)\cdot\sigma_\ell(x^{-1}\otimes y)\cdot\prod_{i=1}^r x_i^{n-r} \;,$$

with $x := (z_1, \ldots, z_r)$, $y := (z_{r+1}, \ldots, z_n)$, $\sigma_\ell$ denoting the $\ell$th elementary symmetric function, and

$$x^{-1}\otimes y := \left(\frac{y_1}{x_1},\ldots,\frac{y_1}{x_r},\frac{y_2}{x_1},\ldots,\frac{y_2}{x_r},\ldots,\frac{y_{n-r}}{x_1},\ldots,\frac{y_{n-r}}{x_r}\right) \in \mathbb{R}^{r\cdot(n-r)} \;.$$

Lastly, we define for $0 \le r \le n$ and $0 \le \ell \le r(n-r)$

$$J(n,r,\ell) \;:=\; \int\limits_{z\in\mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}}\cdot|\Delta_{r,\ell}(z)|\,dz \;. \tag{4.51}$$

**Proposition 4.4.21.** *Let* $\mathrm{Sym}^n_+$ *denote the* $n$*th semidefinite cone*

$$\mathrm{Sym}^n_+ := \{A \in \mathrm{Sym}^n \mid A \textit{ is pos. semidef.}\} \;,$$

*and let* $K_n := \mathrm{Sym}^n_+ \cap S(\mathrm{Sym}^n)$ *denote the corresponding cap. The* $k$*th intrinsic volume of* $K_n$*,* $-1 \le k \le t(n)-1 = \frac{n(n+1)}{2} - 1$*, is given by*

$$V_k(K_n) = \frac{1}{n!\cdot 2^{\frac{n}{2}}\cdot\prod_{d=1}^n \Gamma(\frac{d}{2})}\cdot\sum_{r=0}^n \binom{n}{r}\cdot J(n,r,k+1-t(n-r)) \;,$$

*where* $J(n,r,\ell)$ *is defined as in* (4.51) *for* $0 \le \ell \le r(n-r)$*, resp.* $J(n,r,\ell) := 0$ *for the remaining cases. In particular,*

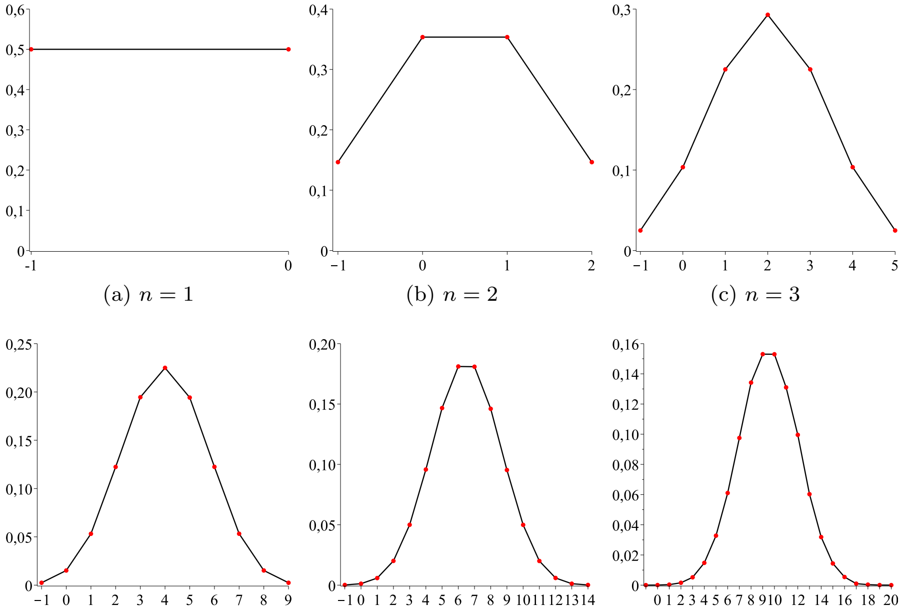
$$V_{-1}(K_n) = V_{t(n)-1}(K_n) = \frac{1}{n!\cdot 2^{\frac{n}{2}}\cdot\prod_{d=1}^n\Gamma(\frac{d}{2})}\cdot\int\limits_{z\in\mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}}\cdot|\Delta(z)|\,dz \;. \tag{4.52}$$

*Proof.* See Section C.2 in the appendix. □

**Remark 4.4.22.** Up to our knowledge, only the values $V_{-1}(K_n) = V_{t(n)-1}(K_n)$ have been known before, as these coincide with the probability that a random matrix from the $n$th Gaussian orthogonal ensemble (cf. Section C.1) is positive definite. It is known that this probability equals the above given term in (4.52) (cf. for example [22]). The computation of the remaining intrinsic volumes seems to not have been done before.

See Figure 4.2 for a graphical display of the intrinsic volumes of $K_n$ for $n = 1,\ldots,6$. In Figure 4.3 we have displayed the logarithms of the intrinsic volumes, to illustrate the conjectured log-concavity (cf. Conjecture 4.4.16). The observable inaccuracy for $n = 6$ results from the fact that the values of the intrinsic volumes of $K_n$ for $n = 4,5,6$ are obtained via numerical approximation. See Chapter C for the details.

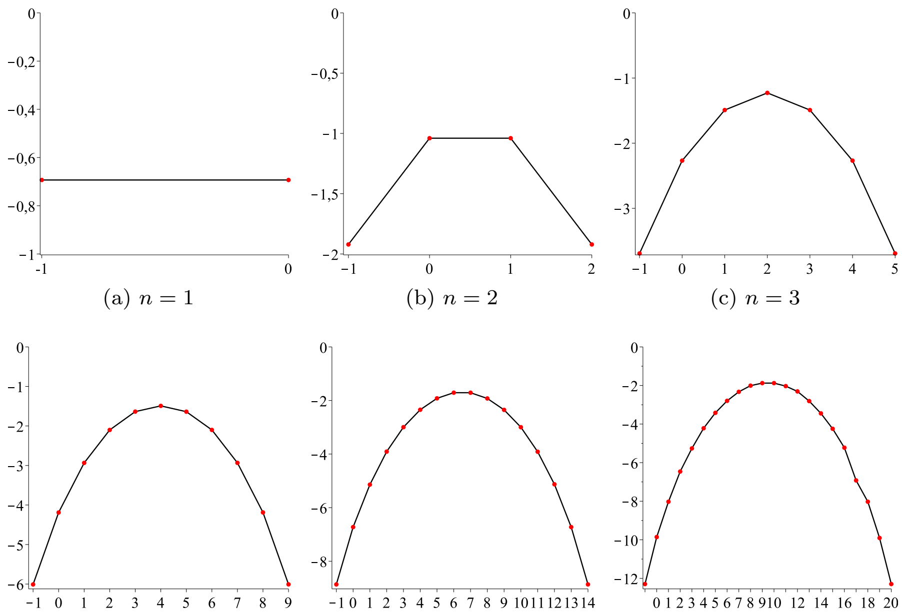
In Section C.3 we will formulate a couple of open questions related to the intrinsic volumes of the semidefinite cone.



(a) $n = 1$ (b) $n = 2$ (c) $n = 3$

(d) $n = 4$ (e) $n = 5$ (f) $n = 6$

Figure 4.2: The intrinsic volumes of the semidefinite cone.



(a) $n = 1$ (b) $n = 2$ (c) $n = 3$

(d) $n = 4$ (e) $n = 5$ (f) $n = 6$

Figure 4.3: The logarithms of the intrinsic volumes of the semidefinite cone.

# Chapter 5

# Computations in the Grassmann manifold

The objective of this chapter is to prepare the ground for the computations in Chapter 6 by describing a differential geometric view on $\mathrm{Gr}_{n,m}$. This is greatly inspired by the articles [25] and [3], although we will only loosely refer to these sources within this chapter.

## 5.1 Preliminary: Riemannian manifolds

We adopt the notation from [23]. It is not the aim of this section to reproduce all the well-known material that may be necessary for a full understanding, but only to pick out the concepts that we need, to recall the corresponding notions, and to set up the notation. For more thorough introductions including proofs we refer to [23], [8], and [19, Ch. 1], or any of the numerous introductory books on Riemannian geometry.

Let $M$ be a smooth manifold, where smooth always means $C^\infty$. For $p \in M$ we denote the tangent space of $M$ in $p$ by $T_pM$. The disjoint union of all these tangent spaces gives the tangent bundle of $M$, denoted by $TM$. The tangent bundle carries the structure of a smooth manifold and has a projection map $\Pi\colon TM \to M$, $T_pM \ni v \mapsto p$. A (local) vector field $X$ on $M$ is a section of the tangent bundle, i.e., a smooth function $X\colon M \rightarrowtail TM$, which satisfies $\Pi(X(p)) = p$ for all $p \in M$ for which $X(p)$ is defined. Here and throughout, the notation $\rightarrowtail$ shall indicate that the map may only be defined on an open subset of the domain.

It is convenient, and we will make extensive use of this, to define vector fields along curves. For example, if we have a curve $c\colon J \to M$, $J \subseteq \mathbb{R}$ an open interval, then a vector field $V$ along $c$ is a map $V\colon J \to TM$, such that $\Pi(V(t)) = c(t)$. Note that if $c\colon J \to M$ is a curve on $M$, then $c$ naturally defines a vector field along $c$ which is given by $t \mapsto \dot{c}(t) = \frac{dc}{dt}(t)$. More generally, if we have a map $\varphi\colon M_1 \to M_2$ between manifolds $M_1, M_2$, then the derivative of $\varphi$ defines a map between the tangent bundles $TM_1$ and $TM_2$. We denote this derivative by $D\varphi\colon TM_1 \to TM_2$, and we denote by $D_p\varphi\colon T_pM_1 \to T_{\varphi(p)}M_2$, $p \in M_1$, the restriction of $D\varphi$ to $T_pM_1$, which is a linear map.

A *Riemannian metric* on $M$ is a family of bilinear forms $\langle .,. \rangle_p$, $p \in M$, which varies smoothly in $p$. This means that if $X, Y$ are local vector fields on the same open subset $U \subseteq M$, the map $U \to \mathbb{R}$, $p \mapsto \langle X(p), Y(p) \rangle_p$, is smooth. A smooth map $\varphi\colon M_1 \to M_2$ between Riemannian manifolds $M_1, M_2$ of dimensions $n_1 \geq n_2$ is called a *Riemannian submersion* iff for every $p \in M$ there exists an orthonormal basis $v_1, \ldots, v_{n_1}$ of $T_pM_1$ and an orthonormal basis $w_1, \ldots, w_{n_2}$ of $T_pM_2$ such that $D_p\varphi(v_i) = w_i$ for all $i = 1, \ldots, n_2$ and $D_p\varphi(v_j) = 0$ for all $j = n_2 + 1, \ldots, n_1$.

A first application of the Riemannian metric is that it defines a *connection* on $M$, i.e. a way to differentiate vector fields.[1] So, if we are given a vector field $V\colon J \to TM$ along a curve $c\colon J \to M$, then there is a well-defined vector field $\nabla V\colon J \to TM$, the *covariant derivative* of $V$ along $c$. If $M$ is embedded in euclidean space and if the Riemannian metric is the usual scalar product in euclidean space, then $\nabla V(t)$ is given by the orthogonal projection of $\frac{dV}{dt}(t)$ onto the tangent space of $M$ in $c(t)$. Thus, the derivative $\nabla V$ may be thought of as the change of $V$ relative to $M$.

Another application of the Riemannian metric is that it defines a way to measure the lengths of curves on $M$. More precisely, for every $p \in M$ the bilinear form $\langle ., . \rangle_p$ on $T_pM$ defines a norm $\|.\|_p$ on $T_pM$ via $\|v\|_p := \sqrt{\langle v, v\rangle_p}$, $v \in T_pM$. With this norm one can define the length of a curve $c\colon J \to M$ via

$$\text{length}(c) \; := \; \int_J \|\dot{c}(t)\|_{c(t)}\, dt \; .$$

Furthermore, this also defines a notion of distance on $M$ via

$$d(p, q) := \inf \left\{ \text{length}(c) \;\middle|\; \begin{matrix} c\colon [0,1] \to M \text{ a piecewise smooth curve} \\ \text{with } c(0) = p \text{ and } c(1) = q \end{matrix} \right\} \; . \tag{5.1}$$

This distance is in fact a metric, which is called the *geodesic metric* on $M$. We will explain this nomenclature in the following paragraphs.

A *geodesic* on $M$ is a curve $c\colon J \to M$, for which the derivative of the field $\dot{c}$ is the zero vector field, i.e., for which

$$\nabla \dot{c}(t) = 0 \; , \quad \text{for all } t \in J \; .$$

Informally, a geodesic on $M$ is a curve which does not bend relative to $M$. It can be shown that for every $p \in M$ there exists a radius $r > 0$ such that for every $v \in T_pM \setminus \{0\}$ with $s := \|v\|_p < r$ there exists a unique geodesic $c\colon (-r, r) \to M$ with $c(0) = p$ and $\dot{c}(0) = \frac{1}{s} \cdot v$. In fact, the assignment $v \mapsto c(s)$ yields a map

$$\exp_p\colon T_pM \rightarrowtail M \; ,$$

which is known as the *exponential map*.

For $r > 0$ small enough the exponential map is a diffeomorphism between the open ball of radius $r$ around the origin in $T_pM$ and a corresponding open neighborhood of $p$ in $M$. Such a neighborhood of $p$ in $M$ is called a *normal neighborhood* of $p$. It can be shown that if $q$ lies in a normal neighborhood around $p$, then there exists a unique path of shortest length between $p$ and $q$, and this path can be described by a geodesic. Therefore, the distance $d(p, q)$ as defined in (5.1) is in this case the length of a geodesic, which justifies the name geodesic distance (cf. [19, § I.6]).

In Section 4.1.3 we have described integration on submanifolds of euclidean space. This naturally transfers to general Riemannian manifolds in the following way. If $\varphi\colon M_1 \to M_2$ is a smooth map between Riemannian manifolds, then for $p \in M_1$ the map $D_p\varphi\colon T_pM_1 \to T_{\varphi(p)}M_2$ is a linear map between inner product spaces. If the differential $D_p\varphi$ is surjective, then the Normal Jacobian of $D_p\varphi$ is defined as in (4.9). We can use this to define the integral of an integrable function

[1] This is known as the Theorem of Levi-Civita (see e.g. [23, Thm. 3.6]).

$f\colon M \to \mathbb{R}$ over an open subset $U \subseteq M$, which has a parametrization, i.e., a smooth diffeomorphism, $\varphi\colon \mathbb{R}^m \to U$, via

$$\int\limits_{p\in U} f(p)\, dM := \int\limits_{x\in\mathbb{R}^m} f(\varphi(x)) \cdot \operatorname{ndet}(D_x\varphi)\, dx \; . \tag{5.2}$$

It can be shown that this definition does not depend on the specific parametrization $\varphi$ of $U$. Furthermore, using a partition of unity as indicated in Section 4.1.3, one can define the integral of $f$ over $M$, which we denote by $\int_M f\, dM$. In particular, the volume of a subset $U \subseteq M$ is defined via

$$\operatorname{vol} U := \int\limits_{p\in M} 1_U(p)\, dM \; ,$$

where $1_U$ denotes the characteristic function of $U$. From the Coarea Formula that we stated in Lemma 4.1.15 one can deduce the following corollaries which we will use for several computations of volumes of Riemannian manifolds.

**Corollary 5.1.1.** *Let $\varphi\colon M_1 \to M_2$ be a smooth surjective map between Riemannian manifolds $M_1, M_2$, such that the derivative $D_p\varphi$ is surjective for almost all $p \in M_1$. Then*

$$\operatorname{vol} M_1 = \int\limits_{q\in M_2} \int\limits_{p\in\varphi^{-1}(q)} \frac{1}{\operatorname{ndet}(D_p\varphi)}\, d\varphi^{-1}(q)\, dM_2 \; . \tag{5.3}$$

*If additionally* $\dim M_1 = \dim M_2$*, then*

$$\operatorname{vol} M_2 \le \int\limits_{q\in M_2} \#\varphi^{-1}(q)\, dM_2 = \int\limits_{p\in M_1} \operatorname{ndet}(D_p\varphi)\, dM_1 \; , \tag{5.4}$$

*where $\#\varphi^{-1}(q)$ denotes the number of elements in the fiber $\varphi^{-1}(q)$. Furthermore, if $\varphi$ is a diffeomorphism then*

$$\operatorname{vol} M_2 = \int\limits_{p\in M_1} \operatorname{ndet}(D_p\varphi)\, dM_1 \; . \tag{5.5}$$

*Proof.* Equations (5.3) and (5.4) follow from (4.10) and (4.11) in Lemma 4.1.15, respectively, using suitable partitions of unity. Equation (5.5) follows from (5.4). □

## 5.2 Orthogonal group

In this section we will recall some elementary facts about the orthogonal group $O(n)$, that will be crucial for the computations in the Grassmann manifold $\mathrm{Gr}_{n,m}$. At some points it is beneficial to know that the orthogonal group is a compact Lie group and thus has some particularly nice properties. For this reason we will also state some elementary definitions and well-known facts about Lie groups with the single purpose of simplifying some arguments about the orthogonal group, which is our only interest.

A *Lie group $G$* is a smooth manifold which is at the same time a group, and for which the group operations $(x, y) \mapsto x \cdot y$ and $x \mapsto x^{-1}$ are smooth maps. The most important example of a Lie group is the general linear group $\mathrm{Gl}_n(\mathbb{R})$. It can be shown that a closed subgroup of $\mathrm{Gl}_n(\mathbb{R})$ is again a Lie group, a so-called linear Lie group. Thus, the orthogonal group $O(n) = \{Q \in \mathbb{R}^{n\times n} \mid Q^TQ = I_n\}$, which is obviously a closed subgroup of $\mathrm{Gl}_n(\mathbb{R})$, is a linear Lie group.

If $G$ is a Lie group then for every $g \in G$ one has a diffeomorphism given by the (left) multiplication map $L_g\colon G \to G$, $L_g(g') := g \cdot g'$. In particular, the differential $D_eL_g\colon T_eG \to T_gG$, where $e \in G$ denotes the identity element, maps the tangent space at $e$ to the tangent space at $g$. The tangent space at $e$ is also called the *Lie algebra* of $G$.

For the orthogonal group, being a linear Lie group and thus embedded in euclidean space $\mathbb{R}^{n\times n}$, one can identify the Lie algebra with a linear subspace of $\mathbb{R}^{n\times n}$. It turns out that this Lie algebra is given by the set of skew-symmetric $(n \times n)$-matrices $\mathrm{Skew}_n = \{U \in \mathbb{R}^{n\times n} \mid U^T = -U\}$. In particular, the dimension of $O(n)$ is given by $\dim \mathrm{Skew}_n = \frac{n(n-1)}{2}$.

Moreover, the Lie algebra of $O(n)$, $\mathrm{Skew}_n$, has a natural inner product inherited from the euclidean space $\mathbb{R}^{n\times n}$. The following two observations are immediate. First, not only the tangent space at the identity matrix, but any tangent space of $O(n)$ can be identified with a linear subspace of $\mathbb{R}^{n\times n}$. More precisely, the tangent space $T_QO(n)$ can be identified with $Q\cdot\mathrm{Skew}_n = \{Q\cdot U \mid U \in \mathrm{Skew}_n\}$. In particular, $T_QO(n)$ has a natural inner product inherited from $\mathbb{R}^{n\times n}$. The second observation is that the inner product in $\mathbb{R}^{n\times n}$ is invariant under (left and right) multiplication with orthogonal matrices.

Combining these observations we may conclude that we have an inner product on each tangent space of $O(n)$, which is invariant under both left and right multiplication with elements from $O(n)$, i.e. we have a bi-invariant Riemannian metric on $O(n)$. It turns out that it is beneficial to scale this inner product by a factor of $\frac{1}{2}$, so that we define

$$\langle QU_1, QU_2\rangle_Q := \tfrac{1}{2} \cdot \mathrm{tr}(U_1^T \cdot U_2) \tag{5.6}$$

for $Q \in O(n)$ and $U_1, U_2 \in \mathrm{Skew}_n$. Observe that we have a canonical basis for $\mathrm{Skew}_n$ given by $\{E_{ij} - E_{ji} \mid 1 \le j < i \le n\}$, where $E_{ij}$ denotes the $(i, j)$th elementary matrix, i.e., the matrix which is zero everywhere except for the $(i, j)$th entry which is 1. This basis is orthogonal and by the choice of the scaling factor it is also orthonormal. In fact, this is the reason to use the scaling factor in the first place. In the euclidean metric on $\mathbb{R}^{n\times n}$ these basis vectors have length $\sqrt{2}$.

Note that the fact that the Riemannian metric is (up to the scaling by $\frac{1}{2}$) inherited from $\mathbb{R}^{n\times n}$ makes the induced connection $\nabla$ on $O(n)$ particularly simple. It is just the orthogonal projection of the derivative in $\mathbb{R}^{n\times n}$ onto the corresponding tangent space of $O(n)$.

Being a smooth manifold, $O(n)$ also has an exponential map. For compact Lie groups, the exponential map is always globally defined. Moreover, it can be shown (see for example [8, Sec. IV.6/VII.8]) that for $O(n)$ with the Riemannian metric chosen as above the exponential map coincides with the exponential function on matrices. More precisely, for $Q \cdot U \in T_QO(n)$, $U \in \mathrm{Skew}_n$, we have

$$\exp_Q(QU) \;=\; Q \cdot \exp_{I_n}(U) \;=\; Q \cdot \sum_{k=0}^{\infty} \frac{U^k}{k!} \;. \tag{5.7}$$

Note also that we have

$$Q^{-1} \cdot \exp_{I_n}(U) \cdot Q = \exp_{I_n}(Q^{-1} \cdot U \cdot Q) \;. \tag{5.8}$$

For $U = E_{ij} - E_{ji}$ and $I_{ij} := E_{ii} + E_{jj}$, $i > j$, we have

$$U^0 = I_n \ , \quad U^1 = U \ , \quad U^2 = -I_{ij} \ , \quad U^3 = -U \ ,$$
$$U^4 = I_{ij} \ , \quad U^5 = U \ , \qquad \ldots \ .$$

From this computation it is straightforward to compute the infinite series $\sum_{k=0}^{\infty} \frac{(\rho \cdot U)^k}{k!}$ resulting in

$$\sum_{k=0}^{\infty} \frac{(\rho \cdot U)^k}{k!} = \cos(\rho) \cdot (E_{ii} + E_{jj}) + \sin(\rho) \cdot (E_{ij} - E_{ji}) + \sum_{k \neq i,j} E_{kk} \ .$$

We thus have an explicit formula for the geodesics on $O(n)$ in direction of the canonical basis vectors $E_{ij} - E_{ji}$, $i > j$. Analogously, a simple computation shows that for $m \leq \frac{n}{2}$ and

$$U = \begin{pmatrix} 0 & -A & 0 \\ A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times n} \ , \quad A := \operatorname{diag}(\alpha_1, \ldots, \alpha_m)$$

we get

$$\exp_{I_n}(\rho \cdot U) = \begin{pmatrix} C_{A,\rho} & -S_{A,\rho} & 0 \\ S_{A,\rho} & C_{A,\rho} & 0 \\ 0 & 0 & I_{n-2m} \end{pmatrix} \ , \tag{5.9}$$

where

$$C_{A,\rho} := \operatorname{diag}(\cos(\rho \cdot \alpha_1), \ldots, \cos(\rho \cdot \alpha_m)) \ ,$$
$$S_{A,\rho} := \operatorname{diag}(\sin(\rho \cdot \alpha_1), \ldots, \sin(\rho \cdot \alpha_m)) \ .$$

Geometrically, the following happens by the transformation (5.9): For $1 \leq i \leq m$ the vectors $e_i$ and $e_{i+m}$, are rotated in the plane $\operatorname{lin}\{e_i, e_{i+m}\}$ with velocity $\alpha_i$, and this happens for all $1 \leq i \leq m$ simultaneously.

For later use we compute the volume of the orthogonal group.

**Proposition 5.2.1.** *The volume of the orthogonal group with respect to the Riemannian metric as defined in* (5.6) *is given by*

$$\operatorname{vol} O(n) = \prod_{i=0}^{n-1} \mathcal{O}_i = \frac{2^n \cdot \pi^{\frac{n^2+n}{4}}}{\prod_{d=1}^{n} \Gamma(\frac{d}{2})} \ . \tag{5.10}$$

**Remark 5.2.2.** If we do not scale the Riemannian metric as in (5.6) but simply take the inner product as inherited from $\mathbb{R}^{n \times n}$, then this causes a scaling of the volume of $O(n)$ by $\sqrt{2}^{\dim O(n)} = 2^{\frac{n(n-1)}{4}}$. This explains the discrepancy between the formula for $\operatorname{vol} O(n)$ in (5.10) and the computation in [26, 3.2.28(5)].

*Proof of Proposition 5.2.1.* Let $e_1 \in \mathbb{R}^n$ denote the first canonical basis vector in $\mathbb{R}^n$, and let us consider the smooth surjection

$$\varphi \colon O(n) \to S^{n-1} \ , \quad Q \mapsto Q \cdot e_1 \ .$$

The fiber of $e_1 = (1, 0, \ldots, 0)^T$ is given by

$$\varphi^{-1}(e_1) = \left\{ \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \bar{Q} \\ 0 & \end{pmatrix} \middle| \, \bar{Q} \in O(n-1) \right\} \ ,$$

which is isometric to $O(n-1)$ and in particular has the same volume. This also holds for the fiber of any other $p \in S^{n-1}$.

Next, we will show that $\varphi$ is a Riemannian submersion so that in particular the Normal Jacobian of $\varphi$ is given by $\mathrm{ndet}_Q\,\varphi = 1$. If we have shown this then we are done as we may then apply the coarea formula (5.3) from Corollary 5.1.1, which yields

$$\operatorname{vol} O(n) = \int\limits_{p\in S^{n-1}} \int\limits_{Q\in\varphi^{-1}(p)} 1\; d\varphi^{-1}(p)\; dS^{n-1} = \mathcal{O}_{n-1}\cdot \operatorname{vol} O(n-1)\ .$$

By induction, using $O(1)=\{1,-1\}$ and $\operatorname{vol} O(1) = 2 = \mathcal{O}_0$, we get

$$\operatorname{vol} O(n) = \prod_{i=0}^{n-1} \mathcal{O}_i = \prod_{i=0}^{n-1} \frac{2\pi^{\frac{i+1}{2}}}{\Gamma(\frac{i+1}{2})} = \frac{2^n\cdot \pi^{\frac{n^2+n}{4}}}{\prod_{d=1}^{n}\Gamma(\frac{d}{2})}\ .$$

So it remains to show that $\varphi$ is a Riemannian submersion. W.l.o.g. we may restrict ourselves to analyze the derivative in $Q = I_n$. To compute the derivative $D_{I_n}\varphi$ we consider the geodesic in direction $E_{ij}-E_{ji}$, $1\le j<i\le n$,

$$\gamma_{ij}(\rho) := \cos(\rho)\cdot(E_{ii}+E_{jj}) + \sin(\rho)\cdot(E_{ij}-E_{ji}) + \sum_{k\neq i,j} E_{kk}\ ,$$

so that $\frac{d\gamma_{ij}}{d\rho}(0) = E_{ij}-E_{ji}$. The composition of $\gamma_{ij}$ with $\varphi$ is given by $\varphi(\gamma_{i1}(\rho)) = \cos(\rho)\cdot e_1 + \sin(\rho)\cdot e_i$ and $\varphi(\gamma_{ij}(\rho)) = e_1$ if $j\neq 1$. Therefore, we get

$$D_{I_n}\varphi(E_{ij}-E_{ji}) = \begin{cases} e_i & \text{if } j=1 \\ 0 & \text{if } j\neq 1\ . \end{cases}$$

In other words, the orthonormal basis vectors $E_{i1}-E_{1i}$, $i=2,\ldots,n$, are sent to the orthonormal basis $e_2,\ldots,e_n$ of $T_{e_1}S^{n-1}$, and the other basis vectors $E_{ij}-E_{ji}$, $j\neq 1$, are sent to the zero vector. Thus, $\varphi$ is a Riemannian submersion, which finishes the proof. □

## 5.3 Quotients of the orthogonal group

This section is devoted to the description of the Stiefel manifold $\mathrm{St}_{n,m}$ and the Grassmann manifold $\mathrm{Gr}_{n,m}$, where $1\le m\le n-1$, as quotients of the orthogonal group $O(n)$. The Stiefel manifold consists of all $m$-tuples of orthonormal vectors in $\mathbb{R}^n$ and the Grassmann manifold consists of all $m$-dimensional linear subspaces of $\mathbb{R}^n$, i.e.,

$$\begin{aligned} \mathrm{St}_{n,m} &= \{B\in\mathbb{R}^{n\times m} \mid B^TB = I_m\} \\ \mathrm{Gr}_{n,m} &= \{\mathcal{W}\subseteq\mathbb{R}^n \mid \mathcal{W} \text{ an } m\text{-dimensional subspace}\}\ . \end{aligned}$$

Note that we have already encountered both of them in slightly different forms (cf. (2.5) in Section 2.1.2, and (3.4) in Section 3.2)

$$\mathrm{St}_{n,m} \cong \mathbb{R}^{m\times n}_{\circ}\ ,\quad \mathrm{Gr}_{n,m}\cong \mathcal{S}^{m-1}(S^{n-1})\ .$$

In fact, there are a number of different ways to identify these sets (cf. the discussion in [25]), but it turns out that considering both of them as quotients of the orthogonal group is particularly beneficial for the computations. We will explain this next by first discussing the general concept of a homogeneous space.

In the following let $G$ be a Lie group and let $H$ be a closed Lie subgroup of $G$, i.e., $H$ is a closed subset of $G$ and furthermore a subgroup $G$. It can be shown that if $H$ is a closed subgroup of $G$, then it is also a submanifold of $G$. The space of left cosets of $H$ in $G$ is denoted by

$$G/H := \{gH \mid g \in G\} .$$

Note that since the left multiplication map $L_g\colon G \to G$, $L_g(g') = g \cdot g'$, is a diffeomorphism, each coset $gH$ is a submanifold of $G$. We use the notation $[g] := gH$ for $g \in G$.

By definition of $G/H$ there is a natural projection map

$$\Pi\colon G \to G/H , \quad g \mapsto [g] .$$

Furthermore, the set of left cosets $G/H$ carries a transitive $G$-action via

$$G \times G/H \to G/H , \quad (g_1, [g_2]) \mapsto [g_1 g_2] .$$

It turns out that $G/H$ even carries a natural manifold structure, which is shown by the following theorem.

**Theorem 5.3.1.** *Let $G$ be a Lie group and $H$ a closed Lie subgroup. Then there exists a unique $C^\infty$-manifold structure on $G/H$ such that the projection $\Pi\colon G \to G/H$ is smooth, and such that for every $g \in G$ there exists an open subset $U \subseteq G/H$ and a smooth map $\psi\colon U \to G$ with $\Pi(\psi(u)) = u$ for all $u \in U$ and $\psi(u_g) = g$ for some $u_g \in U$.*

$$\begin{tikzcd} g \in G \arrow[rr, "\Pi"] & & G/H \\ & u_g \in U \arrow[ul, mapsto] \arrow[ul, "\psi"'] \arrow[ur, hook] & \end{tikzcd}$$

*The $G$-action $G \times G/H \to G/H$, $(g_1, [g_2]) \mapsto [g_1 g_2]$, is a smooth map, and the dimension of $G/H$ is given by* $\dim G/H = \dim G - \dim H$.

*Proof.* See for example [8, Thm. IV.9.2]. □

For our computations this theorem is not quite sufficient. The following lemma will allow us to describe an explicit model for the tangent spaces in $G/H$ and with it also a Riemannian metric etc. Note that an element $g \in G$ lies in the submanifold $[g] = gH \subseteq G$, which implies that the tangent space of $[g]$ in $g$ is a linear subspace of the tangent space of $G$ in $g$. The tangent space $T_g[g]$ is also called the *vertical space* of $G/H$ in $g$.

**Lemma 5.3.2.** *Let $G$ be a Lie group and $H$ a closed Lie subgroup. Furthermore, let $g \in G$ and let $L \subset T_gG$ be a linear subspace such that $T_gG = T_g[g] \oplus L$. Then there exists an open ball $B$ around the origin in $L$ such that the restriction of the composition $\Pi \circ \exp_g$ parametrizes an open neighborhood $U$ of $[g] \in G/H$, i.e.,*

$$\begin{tikzcd} T_gG \arrow[r, "\exp_g"] & G \arrow[r, "\Pi"] & G/H \\ L \supseteq B \arrow[r, "\sim"] \arrow[u, phantom, "\cup|"] & \exp_g(B) \arrow[r, "\sim"] \arrow[u, phantom, "\cup|"] & U \arrow[u, phantom, "\cup|"] \\ 0 \arrow[r, mapsto] \arrow[u, phantom, "\Cup"] & g \arrow[r, mapsto] \arrow[u, phantom, "\Cup"] & {[g]} \arrow[u, phantom, "\Cup"] \end{tikzcd} \quad . \tag{5.11}$$

*In particular, if $D_0$ denotes the differential of $\Pi \circ \exp_g$ in the origin of $T_gG$, then the map $L \to T_{[g]}G/H$, $v \mapsto D_0(v)$, is a linear isomorphism.*

*Proof.* See [34, Lemma II.4.1] for the case $g = e$. The case for general $g \in G$ follows easily by using the left multiplication map. $\square$

The direct summand $L$ of the vertical space $T_g[g]$ thus provides a model of the tangent space of $G/H$ at $[g]$. The question is of course, how the direct summand $L$ should be chosen in order to provide a "good" model of the tangent space of the quotient $G/H$. If the Lie group $G$ is endowed with a Riemannian metric, then one can choose $L$ as the orthogonal complement of the vertical space $T_g[g]$. The orthogonal complement $(T_g[g])^\perp$ of the vertical space is called the *horizontal space* of $G/H$ in $g$. From (5.11) we get a linear isomorphism between the horizontal space and the tangent space of the quotient

$$T_gG \supseteq (T_g[g])^\perp \xrightarrow{\sim} T_{[g]}G/H \ , \quad v \mapsto D_0(v) \ , \tag{5.12}$$

where $D_0$ denotes the differential of $\Pi \circ \exp_g$ in the origin of $T_gG$. Note that for a different representative $gh$, $h \in H$, of the coset $[g]$, we get a linear isomorphism between the horizontal spaces of $G/H$ in $g$ and in $gh$, i.e.,

$$\begin{tikzcd}
(T_g[g])^\perp \arrow[dr, "\sim"] \arrow[d, dashed, "\tau_{g,h}"'] & \\
(T_{gh}[g])^\perp \arrow[r, "\sim"] & T_{[g]}G/H
\end{tikzcd} \quad . \tag{5.13}$$

In order to transfer the Riemannian metric on $G$ to a Riemannian metric on the quotient $G/H$ one could think of declaring the linear isomorphism in (5.12) an isometry. This is possible if the Riemannian metric on $G$ is bi-invariant, i.e., invariant under left and right multiplication, as this implies that the induced isomorphisms $\tau_{g,h}$ between the horizontal spaces are isometries. We will verify this in the following paragraphs for the special case $G = O(n)$.

Recall that (cf. Section 5.2) in the case $G = O(n)$ the tangent space at the identity element is given by $T_{I_n}O(n) = \mathrm{Skew}_n$, and the tangent space at $Q \in O(n)$ is given by $T_QO(n) = Q \cdot \mathrm{Skew}_n$. For a subgroup $H \subseteq O(n)$ let us denote the horizontal space of $O(n)/H$ at the identity $I_n$ by

$$\overline{\mathrm{Skew}_n} := (T_{I_n}H)^\perp \ \subseteq \ T_{I_n}O(n) = \mathrm{Skew}_n \ .$$

The horizontal space of $O(n)/H$ at $Q$ is then given by

$$(T_Q[Q])^\perp = Q \cdot \overline{\mathrm{Skew}_n} \ \subseteq \ T_QO(n) = Q \cdot \mathrm{Skew}_n \ .$$

Let us now fix an element $Q \in O(n)$, and an element $h \in H$, and let us compute the isomorphism $\tau_{Q,h}$ between the horizontal spaces of $O(n)/H$ at $Q$ and at $Qh$ (cf. (5.13)). For $U \in \overline{\mathrm{Skew}_n}$ we have

$$Q\,U = \tfrac{d}{dt} \exp_Q(t \cdot QU)(0) \in Q \cdot \overline{\mathrm{Skew}_n} \ ,$$

and the corresponding tangent vector in $T_{[Q]}O(n)/H$ is given by

$$\tfrac{d}{dt} \left[\exp_Q(t \cdot QU)\right](0) \ .$$

To see what the image of $QU$ under the map $\tau_{Q,h}$ is, we compute

$$\begin{aligned}\big[\exp_Q(t\cdot QU)\big] &= \big[\exp_Q(t\cdot QU)\cdot h\big] \overset{(5.7)}{=} \big[Q\cdot \exp_{I_n}(t\cdot U)\cdot h\big]\\ &= \big[Qh\cdot h^{-1}\cdot \exp_{I_n}(t\cdot U)\cdot h\big] \overset{(5.8)}{=} \big[Qh\cdot \exp_{I_n}(t\cdot h^{-1}U\,h)\big]\\ &\overset{(5.7)}{=} \big[\exp_{Qh}(t\cdot Qh\cdot h^{-1}U\,h)\big] \ .\end{aligned}$$

This shows that we have

$$\tau_{Q,h}(Q\,U) = Qh\cdot h^{-1}U\,h\ .$$

In particular, the transition map $\tau_{Q,h}$ is an isometry (cf. (5.6)).

Now that we have computed the transition functions $\tau_{Q,h}$ between the horizontal spaces, we see that we may identify the tangent space of the quotient $O(n)/H$ with the following set

$$T_{[Q]}O(n)/H \ \cong\ \big([Q]\times \overline{\mathrm{Skew}_n}\big)\,/\sim\ , \tag{5.14}$$

where we define the equivalence relation $\sim$ via

$$\begin{aligned}(Q_1,U_1)\sim(Q_2,U_2)\ :\Longleftrightarrow\ &\tau_{Q_1,h}(Q_1\,U_1) = Q_2\,U_2\ \text{, i.e.,}\\ &h^{-1}U_1h = U_2\ \text{, with } h := Q_1^{-1}\cdot Q_2\in H\ .\end{aligned} \tag{5.15}$$

We denote the equivalence class defined by $(Q,U)$ by $[Q,U]$, so that from now we write a tangent vector in $T_{[Q]}O(n)/H$ in the form

$$[Q,U]\in T_{[Q]}O(n)/H\ ,$$

where $U\in\overline{\mathrm{Skew}_n}$.

As the transition maps $\tau_{Q,h}$ are isometries, we may define an inner product on $T_{[Q]}O(n)/H$ by declaring the linear isomorphism in (5.12) between the horizontal space $Q\cdot\overline{\mathrm{Skew}_n}$ and the tangent space $T_{[Q]}O(n)/H$ an isometry. This amounts to declaring an inner product on the right-hand side of (5.14) via

$$\big\langle [Q,U_1]\,,\ [Q,U_2]\big\rangle := \langle U_1,U_2\rangle_{I_n} \overset{(5.6)}{=} \tfrac{1}{2}\cdot \mathrm{tr}(U_1^T\cdot U_2)\ , \tag{5.16}$$

for $U_1,U_2\in\overline{\mathrm{Skew}_n}$.

We have thus defined an $O(n)$-invariant Riemannian metric on the quotient $O(n)/H$. It can be shown that the following properties hold for this Riemannian metric:

1. The projection $\Pi\colon O(n)\to O(n)/H$ is a Riemannian submersion.
2. The differential of the projection $\Pi\colon O(n)\to O(n)/H$ is given by the orthogonal projection onto the horizontal space.
3. The exponential map on $O(n)/H$ is given in the following way. For $[Q]\in O(n)/H$ and $U\in\overline{\mathrm{Skew}_n}$ we have

$$\overline{\exp}_{[Q]}([Q,U]) = [\exp_Q(Q\,U)]\ , \tag{5.17}$$

where $\overline{\exp}$ and exp shall denote the exponential maps of $O(n)/H$ and $O(n)$, respectively. In other words, the projection $\Pi$ maps geodesics in $O(n)$ on geodesics in $O(n)/H$, and every geodesic in $O(n)/H$ may be obtained this way.

Before specializing the subgroup $H$ so that $O(n)/H$ yields a model for the Stiefel or the Grassmann manifold, let us compute the volume of the homogeneous space $O(n)/H$.

**Proposition 5.3.3.** *Let $H$ be a closed Lie subgroup of the orthogonal group $O(n)$. Then the volume of the homogeneous space $O(n)/H$ is given by*

$$\operatorname{vol} O(n)/H = \frac{\operatorname{vol} O(n)}{\operatorname{vol} H} .$$

*Proof.* The proof goes analogous to the proof of Proposition 5.2.1.[2] The projection map $\Pi\colon O(n) \to O(n)/H$ is a Riemannian submersion, and for each element $\tilde{H} \in O(n)/H$ the fiber $\Pi^{-1}(\tilde{H}) = \tilde{H} \subset O(n)$ is isometric to $H$. Applying the co-area formula (5.3) from Corollary 5.1.1 yields

$$\operatorname{vol} O(n) = \int\limits_{\tilde{H}\in O(n)/H} \int\limits_{Q\in\tilde{H}} 1 \, d\tilde{H} \, dO(n)/H = \operatorname{vol} O(n)/H \cdot \operatorname{vol} H . \qquad \square$$

### 5.3.1 Stiefel manifold

The $(n,m)$th Stiefel manifold $\mathrm{St}_{n,m}$, $m \leq n$, consists of all $m$-tuples of orthonormal vectors in $\mathbb{R}^n$. Instead of considering only $m$-tuples we may as well consider all $n$-tuples of orthonormal vectors in $\mathbb{R}^n$, i.e., all (ordered) orthonormal bases, and identify those for which the first $m$ components coincide. This last description of $\mathrm{St}_{n,m}$ can be written as a quotient of the orthogonal group $O(n)$. To make this precise, let

$$H := \left\{ \begin{pmatrix} I_m & 0 \\ 0 & \bar{Q} \end{pmatrix} \middle| \bar{Q} \in O(n-m) \right\} \subseteq O(n) .$$

Note that $H \cong O(n-m)$. In particular, $H$ is a closed Lie subgroup of $O(n)$. Considering the homogeneous space $O(n)/H$ we have that the left cosets $Q_1 H = Q_2 H$ iff the first $m$ columns of $Q_1$ and $Q_2$ coincide. Therefore, we get a model for the Stiefel manifold via

$$\mathrm{St}_{n,m} \cong O(n)/H .$$

As $H \cong O(n-m)$, we also write $\mathrm{St}_{n,m} \cong O(n)/O(n-m)$ implicitly assuming the embedding of $O(n-m)$ in $O(n)$ as given by $H$.

The vertical and the horizontal space of $O(n)/H$ in $I_n$ are given by

$$T_{I_n} H = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & V \end{pmatrix} \middle| V \in \mathrm{Skew}_{n-m} \right\} ,$$

$$\overline{\mathrm{Skew}_n} = (T_{I_n} H)^\perp = \left\{ \begin{pmatrix} U & -R^T \\ R & 0 \end{pmatrix} \middle| U \in \mathrm{Skew}_m \, , \; R \in \mathbb{R}^{(n-m)\times m} \right\} .$$

Furthermore, the induced Riemannian metric on $\mathrm{St}_{n,m}$ is given by (cf. (5.16))

$$\left\langle \left[Q, \begin{pmatrix} U_1 & -R_1^T \\ R_1 & 0 \end{pmatrix}\right], \left[Q, \begin{pmatrix} U_2 & -R_2^T \\ R_2 & 0 \end{pmatrix}\right] \right\rangle = \tfrac{1}{2} \cdot \operatorname{tr}(U_1^T \cdot U_2) + \operatorname{tr}(R_1^T \cdot R_2) . \qquad (5.18)$$

From Proposition 5.3.3 we get that the volume of the Stiefel manifold with respect to this Riemannian metric is given by

$$\operatorname{vol} \mathrm{St}_{n,m} = \frac{\operatorname{vol} O(n)}{\operatorname{vol} O(n-m)} \overset{\text{Prop. 5.2.1}}{=} \prod_{i=n-m}^{n-1} \mathcal{O}_i = \frac{2^m \cdot \pi^{\frac{2nm-m^2+m}{4}}}{\prod_{d=n-m+1}^{n} \Gamma(\frac{d}{2})} . \qquad (5.19)$$

[2]In fact, Proposition 5.2.1 can be deduced from Proposition 5.3.3, as $S^{n-1} \cong \mathrm{St}_{n,1}$ and thus $\mathcal{O}_{n-1} = \operatorname{vol} \mathrm{St}_{n,1} = \frac{\operatorname{vol} O(n)}{\operatorname{vol} O(n-1)}$ (cf. Section 5.3.1).

**Remark 5.3.4.** As in Remark 5.2.2 we point out the difference between the volume of $\mathrm{St}_{n,m}$ with respect to the Riemannian metric on $\mathrm{St}_{n,m}$ as defined in (5.18) and the volume of $\mathrm{St}_{n,m}$ if we consider it as a submanifold of $\mathbb{R}^{n\times m}$ and take the volume with respect to the volume form induced by this inclusion. To get the latter volume we need to scale the result in (5.19) by $\sqrt{2}^{\dim \mathrm{St}_{n,m}} = 2^{\frac{m(m-1)}{4}+\frac{m(n-m)}{2}}$, which yields the volume $2^{\frac{m}{2}} \cdot (2\pi)^{\frac{2nm-m^2+m}{4}} \cdot \prod_{d=n-m+1}^{n} \Gamma(\frac{d}{2})^{-1}$.

### 5.3.2 Grassmann manifold

The $(n,m)$th Grassmann manifold $\mathrm{Gr}_{n,m}$, $m \le n$, consists of all $m$-dimensional subspaces of $\mathbb{R}^n$. We may identify an $m$-dimensional subspace $\mathcal{W}$ with the set of all orthonormal bases of this subspace. Moreover, we may also identify $\mathcal{W}$ with the set of all orthonormal bases of $\mathbb{R}^n$ whose first $m$ components span $\mathcal{W}$ and whose last $n-m$ components span its orthogonal complement.[3] This, in turn, leads to a description of $\mathrm{Gr}_{n,m}$ as a homogeneous space. We define

$$H := \left\{ \begin{pmatrix} \bar{Q} & 0 \\ 0 & \bar{\bar{Q}} \end{pmatrix} \middle| \, \bar{Q} \in O(m) \, , \; \bar{\bar{Q}} \in O(n-m) \right\} \subseteq O(n) \, ,$$

so that $H \cong O(m) \times O(n-m)$ is in particular a closed Lie subgroup of $O(n)$. As for the homogeneous space $O(n)/H$ it is straightforward that two left cosets $Q_1H, Q_2H \in O(n)/H$ coincide iff the first $m$ columns of $Q_1$ and the first $m$ columns of $Q_2$ span the same subspace $\mathcal{W}$, or equivalently the last $n-m$ columns of $Q_1$ and the last $n-m$ columns of $Q_2$ span the same subspace $\mathcal{W}^\perp$. We thus get

$$\mathrm{Gr}_{n,m} \cong O(n)/H \, ,$$

which we also write in the form $\mathrm{Gr}_{n,m} \cong O(n)/(O(m)\times O(n-m))$, implicitly assuming the embedding of $O(m)\times O(n-m)$ in $O(n)$ as given by $H$.

The vertical and the horizontal space of $O(n)/H$ in $I_n$ are given by

$$T_{I_n}H = \left\{ \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \middle| \, U \in \mathrm{Skew}_m \, , \; V \in \mathrm{Skew}_{n-m} \right\} \, ,$$

$$\overline{\mathrm{Skew}_n} = (T_{I_n}H)^\perp = \left\{ \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \middle| \, R \in \mathbb{R}^{(n-m)\times m} \right\} \, .$$

Furthermore, the induced Riemannian metric on $\mathrm{Gr}_{n,m}$ is given by (cf. (5.16))

$$\left\langle \left[Q, \begin{pmatrix} 0 & -R_1^T \\ R_1 & 0 \end{pmatrix}\right], \left[Q, \begin{pmatrix} 0 & -R_2^T \\ R_2 & 0 \end{pmatrix}\right] \right\rangle = \mathrm{tr}(R_1^T \cdot R_2) \, . \tag{5.20}$$

Note that for $h = \begin{pmatrix} \bar{Q} & 0 \\ 0 & \bar{\bar{Q}} \end{pmatrix} \in H$, i.e., $\bar{Q} \in O(m)$ and $\bar{\bar{Q}} \in O(n-m)$, we have

$$\left[Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix}\right] \overset{(5.15)}{=} \left[Qh, h^{-1} \cdot \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \cdot h\right] = \left[Qh, \begin{pmatrix} 0 & -(\bar{\bar{Q}}^T \cdot R \cdot \bar{Q})^T \\ \bar{\bar{Q}}^T \cdot R \cdot \bar{Q} & 0 \end{pmatrix}\right] \, .$$

Note that the rank of the matrix $R$ coincides with the rank of $\bar{\bar{Q}}^T \cdot R \cdot \bar{Q}$. So for $v := \left[Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix}\right] \in T_{[Q]}\,\mathrm{Gr}_{n,m}$ we may define the *rank of* $v$ via

$$\mathrm{rk}(v) := \mathrm{rk}(R) \, . \tag{5.21}$$

[3] Note that this might be interpreted as a primal-dual view on (GrP) and (GrD) (cf. Section 2.3).

Moreover, if $R$ has the singular value decomposition (assuming $m \leq \frac{n}{2}$ for notational reasons)

$$R = \bar{\bar{Q}} \cdot \begin{pmatrix} A \\ 0 \end{pmatrix} \cdot \bar{Q}^T ,$$

with $A = \operatorname{diag}(\alpha_1, \dots, \alpha_m)$, then we have

$$\left[Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix}\right] = \left[Qh, \begin{pmatrix} 0 & -A & 0 \\ A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right] . \tag{5.22}$$

Note that the rank of the tangent vector $v$ is given by

$$\operatorname{rk}(v) = |\{i \mid \alpha_i > 0\}| .$$

The Grassmann manifold is a compact and connected manifold, and from Proposition 5.3.3 we get that the volume of $\mathrm{Gr}_{n,m}$ with respect to the above Riemannian metric is given by

$$\begin{aligned} \operatorname{vol} \mathrm{Gr}_{n,m} &= \frac{\operatorname{vol} O(n)}{\operatorname{vol} O(m) \cdot \operatorname{vol} O(n-m)} \overset{\text{Prop. 5.2.1}}{=} \frac{\prod_{i=0}^{n-1} \mathcal{O}_i}{\prod_{i=0}^{m-1} \mathcal{O}_i \cdot \prod_{i=0}^{n-m-1} \mathcal{O}_i} \\ &= \frac{\prod_{i=n-m}^{n-1} \mathcal{O}_i}{\prod_{i=0}^{m-1} \mathcal{O}_i} = \pi^{\frac{m(n-m)}{2}} \cdot \prod_{d=1}^{m} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{n-m+d}{2})} . \end{aligned} \tag{5.23}$$

The scaling of the Riemannian metric is in some sense arbitrary and only motivated by the resulting computations. That is why it is sometimes helpful not to consider the volume but the relative volume

$$\operatorname{rvol} \mathcal{M} := \frac{\operatorname{vol} \mathcal{M}}{\operatorname{vol} \mathrm{Gr}_{n,m}}$$

for measurable $\mathcal{M} \subseteq \mathrm{Gr}_{n,m}$. In fact, the relative volume is a probability measure on $\mathrm{Gr}_{n,m}$ and it can be described without considering $\mathrm{Gr}_{n,m}$ as a homogeneous space.

**Proposition 5.3.5.** *Let $\mathcal{M} \subseteq \mathrm{Gr}_{n,m}$, and let $B \in \mathbb{R}^{n \times m}$ and $B' \in \mathbb{R}^{(n-m) \times n}$ be random matrices, where each entry of $B$ and $B'$ is chosen i.i.d. normal distributed. Then*

$$\operatorname{rvol} \mathcal{M} = \operatorname{Prob}[\operatorname{im} B \in \mathcal{M}] = \operatorname{Prob}[\ker B' \in \mathcal{M}] . \tag{5.24}$$

*Proof.* The measure rvol on $\mathrm{Gr}_{n,m}$ is a Borel measure, which is invariant under the action of the orthogonal group $O(n)$, and which is normalized via $\operatorname{rvol}(\mathrm{Gr}_{n,m}) = 1$. It is known that these properties uniquely determine the measure rvol (cf. for example [39, Ch. 3]). This fact easily implies (5.24), as we will see next.

Recall that $\mathbb{R}^{n \times m}_*$ denotes the set of full rank $(n \times m)$-matrices. Let $\mu$ and $\mu'$ denote the measures on $\mathbb{R}^{n \times m}_*$ respectively $\mathbb{R}^{(n-m) \times n}_*$, which result from taking the entries in $B \in \mathbb{R}^{n \times m}_*$ respectively $B \in \mathbb{R}^{(n-m) \times n}_*$ i.i.d. standard normal distributed (note that $B \in \mathbb{R}^{n \times m}_*$ and $B \in \mathbb{R}^{(n-m) \times n}_*$ with probability 1). Furthermore, let $I$ and $K$ denote the maps

$$\begin{aligned} I \colon \mathbb{R}^{n \times m}_* &\to \mathrm{Gr}_{n,m} , \quad B \mapsto \operatorname{im} B , \\ K \colon \mathbb{R}^{(n-m) \times n}_* &\to \mathrm{Gr}_{n,m} , \quad B' \mapsto \ker B' . \end{aligned} \tag{5.25}$$

Then we have

$$\mathrm{Prob}\,[\mathrm{im}\, B \in \mathcal{M}] = \int\limits_{I^{-1}(\mathcal{W})} d\mu = \int\limits_{\mathcal{W}} d\mu_* \;,$$

$$\mathrm{Prob}\,[\ker B' \in \mathcal{M}] = \int\limits_{K^{-1}(\mathcal{W})} d\mu' = \int\limits_{\mathcal{W}} d\mu'_* \;,$$

where $\mu_*$ and $\mu'_*$ denote the pushforwards of the measures $\mu$ and $\mu'$ onto $\mathrm{Gr}_{n,m}$ via the maps $I$ and $K$, respectively. These pushforwards are also orthogonal invariant probability measures on $\mathrm{Gr}_{n,m}$. From the above mentioned uniqueness we thus get that the integrals coincide with rvol. □

**Remark 5.3.6.** We obtain further representations of the Grassmann manifold by identifying a subspace $\mathcal{W} \subseteq \mathbb{R}^n$ with all its bases or with all bases of its complement. Formally, this means that we can identify $\mathrm{Gr}_{n,m}$ with $\mathbb{R}_*^{m\times n}/\,\mathrm{Gl}_m$ or $\mathbb{R}_*^{(n-m)\times n}/\,\mathrm{Gl}_{n-m}$. Note that $\mathbb{R}_*^{m\times n}$ and $\mathbb{R}_*^{(n-m)\times n}$ are homogeneous spaces, for example $\mathbb{R}_*^{m\times n} \cong \mathrm{Gl}_n\,/\,\mathrm{Gl}_{n-m}$, where $\mathrm{Gl}_{n-m}$ is identified with a corresponding subgroup of $\mathrm{Gl}_n$. Moreover, the maps $I$ and $K$ as defined in (5.25) are Riemannian submersions and thus continuous, open, and closed maps.

## 5.4 Geodesics in $\mathrm{Gr}_{n,m}$

In this section we will have a closer look at geodesics in $\mathrm{Gr}_{n,m}$. Central to the understanding of the geometry of the Grassmann manifold is the notion of principal angles, which was originally defined by Jordan [35].

**Definition 5.4.1.** Let $1 \le m \le M \le n-1$, and let $\mathcal{W}_1 \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2 \in \mathrm{Gr}_{n,M}$. Furthermore, let $X_1 \in \mathbb{R}^{n\times m}$ and $X_2 \in \mathbb{R}^{n\times M}$ be such that the columns of $X_i$ form an orthonormal basis of $\mathcal{W}_i$, $i = 1, 2$. The principal angles $\alpha_1 \le \ldots \le \alpha_m \in [0, \frac{\pi}{2}]$ between $\mathcal{W}_1$ and $\mathcal{W}_2$ are defined as the arccosines of the singular values of the matrix $X_1^T X_2 \in \mathbb{R}^{m\times M}$, i.e.,

$$X_1^T X_2 = Q_1 \begin{pmatrix} \cos(\alpha_1) & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \cos(\alpha_m) & 0 & \cdots & 0 \end{pmatrix} \cdot Q_2^T \;,$$

where $Q_1 \in O(m)$ and $Q_2 \in O(M)$.

Note that we have $\|X_1\| = \|X_2\| = 1$, as the columns of $X_1$ and $X_2$ are orthonormal vectors in $\mathbb{R}^n$. This implies $\sigma_1 = \|X_2^T X_1\| \le \|X_2^T\| \cdot \|X_1\| = 1$, so that the arccosines of the singular values are well-defined. Furthermore, it is easily seen that the principal angles are independent of the above chosen orthonormal bases $X_1, X_2$ (cf. Proposition D.2.2).

**Remark 5.4.2.** The number of principal angles, which are 0, gives the dimension of the intersection $\mathcal{W}_1 \cap \mathcal{W}_2$:

$$\alpha_1 = \ldots = \alpha_k = 0\,,\ \alpha_{k+1} > 0 \iff k = \dim(\mathcal{W}_1 \cap \mathcal{W}_2)\;.$$

The '$\Leftarrow$'-direction follows from the fact that we can choose $X_1$ and $X_2$ such that the first $k$ columns coincide and describe an orthonormal basis of $\mathcal{W}_1 \cap \mathcal{W}_2$. The '$\Rightarrow$'-direction is also verified easily.

We will analyze principal angles in more details in Chapter D in the appendix. One important property of the principal angles is that they parametrize pairs of subspaces up to orthogonal transformation (cf. Proposition D.2.6). In particular, if we have any function $f$ defined on $\mathrm{Gr}_{n,m} \times \mathrm{Gr}_{n,m}$ satisfying $f(Q\mathcal{W}_1, Q\mathcal{W}_2) = f(\mathcal{W}_1, \mathcal{W}_2)$ for all $\mathcal{W}_1, \mathcal{W}_2 \in \mathrm{Gr}_{n,m}$, $Q \in O(n)$, then this function $f$ only depends on the principal angles. This means that there exists a (symmetric) function $g$ defined on $[0, \frac{\pi}{2}]^m$ such that $f(\mathcal{W}_1, \mathcal{W}_2) = g(\alpha_1, \ldots, \alpha_m)$, where $\alpha_1, \ldots, \alpha_m$ denote the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$. This emphasizes the central role of the principal angles for the Grassmann manifold.

Another important property of the principal angles is that they are basically invariant under the duality map. The precise statement is given in the following proposition.

**Proposition 5.4.3.** *Let $1 \leq m \leq M \leq n-1$, and let $\mathcal{W}_1 \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2 \in \mathrm{Gr}_{n,M}$. Then the nonzero principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$ coincide with the nonzero principal angles between $\mathcal{W}_1^\perp$ and $\mathcal{W}_2^\perp$.*

*Proof.* See Corollary D.2.5. □

Note that the bijection $\mathrm{Gr}_{n,m} \to \mathcal{S}^{m-1}(S^{n-1})$, $\mathcal{W} \mapsto \mathcal{W} \cap S^{n-1}$, implies that the Hausdorff distance on $\mathcal{S}^{m-1}(S^{n-1})$ transfers to a metric on $\mathrm{Gr}_{n,m}$, which we also denote by $d_{\mathrm{H}}$,

$$d_{\mathrm{H}}(\mathcal{W}_1, \mathcal{W}_2) := d_{\mathrm{H}}(\mathcal{W}_1 \cap S^{n-1}, \mathcal{W}_2 \cap S^{n-1}) \; .$$

The following proposition shows how this metric is related to the principal angles.

**Proposition 5.4.4.** *Let $\mathcal{W}_1, \mathcal{W}_2 \in \mathrm{Gr}_{n,m}$, and let $\alpha_1 \leq \ldots \leq \alpha_m$ denote the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$. Then the Hausdorff distance between $\mathcal{W}_1$ and $\mathcal{W}_2$ is given by*

$$d_{\mathrm{H}}(\mathcal{W}_1, \mathcal{W}_2) = \|(\alpha_1, \ldots, \alpha_m)\|_\infty = \alpha_m \; . \tag{5.26}$$

*Furthermore, the map $\mathrm{Gr}_{n,m} \to \mathrm{Gr}_{n,n-m}$, $\mathcal{W} \mapsto \mathcal{W}^\perp$, is an isometry if $\mathrm{Gr}_{n,m}$ and $\mathrm{Gr}_{n,n-m}$ are both endowed with the Hausdorff metric.*

*Proof.* Let $S_i := \mathcal{W}_i \cap S^{n-1}$, $i = 1, 2$, denote the subspheres defined by $\mathcal{W}_1$ and $\mathcal{W}_2$, and let $X_i \in \mathbb{R}^{n \times m}$ be such that the columns of $X_i$ form an orthonormal basis of $\mathcal{W}_i$, $i = 1, 2$. From the definition of the principal angles in Definition 5.4.1 and from the geometric interpretation of the minimum singular value in Proposition 2.1.3 we get

$$\cos \alpha_m = \max\{r \mid X_1^T X_2(B_m) \supseteq B_m(r)\} \; ,$$

where $B_m(r) \subset \mathbb{R}^m$ denotes the ball of radius $r$, and $B_m := B_m(1)$. As $X_1 \colon \mathbb{R}^m \to \mathbb{R}^n$ is a linear isometry, we have

$$X_1^T X_2(B_m) \supseteq B_m(r) \iff X_1 X_1^T X_2(B_m) \supseteq X_1(B_m(r)) \; .$$

Furthermore, $X_i(B_m(r))$ equals the intersection of $\mathcal{W}_i$ with $B_n(r)$, and the linear map $X_i X_i^T$ equals the orthogonal projection onto $\mathcal{W}_i$ (cf. Lemma 2.1.11), $i = 1, 2$. This yields

$$\begin{aligned} \cos \alpha_m &= \max\{r \mid \Pi_{\mathcal{W}_1}(\mathcal{W}_2 \cap B_n) \supseteq \mathcal{W}_1 \cap B_n(r)\} \\ &= \max\{\|x\| \mid x \in \Pi_{\mathcal{W}_1}(S_2)\} \\ &= \min\{\cos(d(q, S_1)) \mid q \in S_2\} \\ &= \cos(\max\{d(q, S_1) \mid q \in S_2\}) \; . \end{aligned}$$

By symmetry this also holds with interchanged roles of $\mathcal{W}_1$ and $\mathcal{W}_2$, so that we may conclude

$$\begin{aligned}\alpha_m &= \max\{d(q,S_1) \mid q \in S_2\} = \max\{d(p,S_2) \mid p \in S_1\} \\ &= d_{\mathrm{H}}(S_1,S_2) \ .\end{aligned}$$

The claim that the map $\mathcal{W} \mapsto \mathcal{W}^\perp$ is an isometry follows from (5.26) and Proposition 5.4.3. □

The following fundamental lemma yields concrete formulas for the geodesics in the Grassmann manifold $\mathrm{Gr}_{n,m}$. We will assume that $m \le \frac{n}{2}$ to simplify the notation.

**Lemma 5.4.5.** *Let $\mathcal{W} \in \mathrm{Gr}_{n,m}$, with $m \le \frac{n}{2}$, and let $v \in T_{\mathcal{W}}\,\mathrm{Gr}_{n,m}$. Then there exists $Q \in O(n)$ such that $[Q] = \mathcal{W}$ and*

$$v = \left[ Q, \begin{pmatrix} 0 & -A & 0 \\ A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \ , \tag{5.27}$$

*where $A = \mathrm{diag}(\alpha_1,\dots,\alpha_m)$ with $\alpha_1,\dots,\alpha_m \ge 0$. Furthermore, the exponential map $\overline{\exp}$ on $\mathrm{Gr}_{n,m}$ in $\mathcal{W}$ in direction $v$ is given by*

$$\overline{\exp}_{\mathcal{W}}(\rho \cdot v) = \left[ Q \cdot \begin{pmatrix} C_{A,\rho} & -S_{A,\rho} & 0 \\ S_{A,\rho} & C_{A,\rho} & 0 \\ 0 & 0 & I_{n-2m} \end{pmatrix} \right] \ , \tag{5.28}$$

*where*

$$\begin{aligned} C_{A,\rho} &= \mathrm{diag}(\cos(\rho\cdot\alpha_1),\dots,\cos(\rho\cdot\alpha_m)) \ , \\ S_{A,\rho} &= \mathrm{diag}(\sin(\rho\cdot\alpha_1),\dots,\sin(\rho\cdot\alpha_m)) \ . \end{aligned}$$

*Proof.* In (5.22) we have seen that for every $v \in T_{\mathcal{W}}\,\mathrm{Gr}_{n,m}$ we can find a representative $Q \in O(n)$ with $[Q] = \mathcal{W}$, such that $v$ is of the form (5.27).

The exponential map on $\mathrm{Gr}_{n,m}$ in $\mathcal{W} = [Q]$ in direction $v = [Q,U]$ is given by (cf. (5.17))

$$\overline{\exp}_{\mathcal{W}}(\rho \cdot v) = \left[\exp_Q(\rho \cdot Q\,U)\right] = \left[Q \cdot \exp_{I_n}(\rho \cdot U)\right] \ ,$$

where exp shall denote the exponential map on $O(n)$. From (5.9) in Section 5.2 we thus get the claimed formula for $\overline{\exp}_{\mathcal{W}}(\rho \cdot v)$. □

**Example 5.4.6.** Let us consider the special case where the rank of the tangent vector $v \in T_{\mathcal{W}}\,\mathrm{Gr}_{n,m}$ (cf. (5.21)) is $\mathrm{rk}(v) = 1$. If the tangent vector has unit length, i.e., $\|v\| = 1$, then we can find $Q \in O(n)$ such that $[Q] = \mathcal{W}$, and

$$v = \left[ Q, \left( \begin{array}{c|ccc|c} 0 & 0 & \cdots & 0 & -1 \\ \hline 0 & & & & 0 \\ \vdots & & 0 & & \vdots \\ 0 & & & & 0 \\ \hline 1 & 0 & \cdots & 0 & 0 \end{array} \right) \right] \ .$$

The exponential map on $\mathrm{Gr}_{n,m}$ at $\mathcal{W}$ in direction $v$ is thus given by

$$\overline{\exp}_{\mathcal{W}}(\rho \cdot v) = [Q \cdot Q_\rho] \ , \quad Q_\rho := \begin{pmatrix} \cos(\rho) & & -\sin(\rho) \\ & I_{n-2} & \\ \sin(\rho) & & \cos(\rho) \end{pmatrix} \ .$$

Denoting by $c\colon \mathbb{R} \to \mathrm{Gr}_{n,m}$ the curve $c(\rho) := \overline{\exp}_{\mathcal{W}}(\rho \cdot v)$, let us also compute the tangent vector $\dot{c}(\rho)$, to get accustomed to the notation. First of all, we have

$$\tfrac{d}{d\rho}(Q \cdot Q_\rho)(\rho) = \tfrac{d}{dt}(Q \cdot Q_{\rho+t})(0) = Q \cdot Q_\rho \cdot \tfrac{d}{dt} Q_t(0) = Q \cdot Q_\rho \cdot \begin{pmatrix} 0 & & & & -1 \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 1 & & & & 0 \end{pmatrix} .$$

This implies that

$$\dot{c}(\rho) = \left[ Q \cdot Q_\rho, \begin{pmatrix} 0 & & & & -1 \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 1 & & & & 0 \end{pmatrix} \right] . \tag{5.29}$$

Note that this is the correct result, as the exponential map is known to "parallel transport its velocity".

In the following proposition we summarize the most important global properties of the exponential map on $\mathrm{Gr}_{n,m}$.

**Proposition 5.4.7.** *Let $\mathcal{W} = [Q] \in \mathrm{Gr}_{n,m}$, and consider the open set*

$$\mathcal{U} := \left\{ \left[ Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \right] \,\middle|\, R \in \mathbb{R}^{(n-m)\times m} ,\ \|R\| < \frac{\pi}{2} \right\} \subset T_{\mathcal{W}} \mathrm{Gr}_{n,m} , \tag{5.30}$$

*where $\|R\|$ denotes the operator norm of $R$. Furthermore, let $\overline{\mathcal{U}}$ denote the closure of $\mathcal{U}$, and let $\partial\mathcal{U}$ denote the boundary of $\mathcal{U}$. Then the following holds.*

1. *The exponential map $\overline{\exp}_{\mathcal{W}}$ is injective on $\mathcal{U}$.*
2. *The exponential map $\overline{\exp}_{\mathcal{W}}$ is surjective on $\overline{\mathcal{U}}$.*
3. *If $v = \left[ Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \right] \in \overline{\mathcal{U}}$ and $\mathcal{W}' = \overline{\exp}_{\mathcal{W}}(v)$, then the curve*

$$[0,1] \to \mathrm{Gr}_{n,m} , \quad \rho \mapsto \overline{\exp}_{\mathcal{W}}(\rho \cdot v) \tag{5.31}$$

*is a shortest length geodesic between $\mathcal{W}$ and $\mathcal{W}'$. In particular, we have $d_{\mathrm{g}}(\mathcal{W}, \mathcal{W}') = \|v\| = \|R\|_F$.*

4. *For $v \in \partial\mathcal{U}$ we have $\overline{\exp}_{\mathcal{W}}(v) = \overline{\exp}_{\mathcal{W}}(-v)$, so that the injectivity radius of $\mathrm{Gr}_{n,m}$ is $\frac{\pi}{2}$.*

*Proof.* See Section D.2 in the appendix. □

With the help of Proposition 5.4.7 we can now show how the geodesic metric on $\mathrm{Gr}_{n,m}$ is related to the principal angles. In particular, we will see that the Hausdorff metric and the geodesic metric are equivalent.

**Corollary 5.4.8.** *For $\mathcal{W}_1, \mathcal{W}_2 \in \mathrm{Gr}_{n,m}$ let $d_{\mathrm{g}}(\mathcal{W}_1, \mathcal{W}_2)$ denote the geodesic distance between $\mathcal{W}_1$ and $\mathcal{W}_2$. Then*

$$d_{\mathrm{g}}(\mathcal{W}_1, \mathcal{W}_2) = \|(\alpha_1, \ldots, \alpha_m)\|_2 = \sqrt{\alpha_1^2 + \ldots + \alpha_m^2} , \tag{5.32}$$

*where $\alpha_1 \leq \ldots \leq \alpha_m$ denote the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$. In particular, the Hausdorff metric and the geodesic metric are equivalent.*

*Furthermore, the map $\mathrm{Gr}_{n,m} \to \mathrm{Gr}_{n,n-m}$, $\mathcal{W} \mapsto \mathcal{W}^\perp$, is an isometry if $\mathrm{Gr}_{n,m}$ and $\mathrm{Gr}_{n,n-m}$ are both endowed with the geodesic metric.*

*Proof.* Let us fix $\mathcal{W} \in \mathrm{Gr}_{n,m}$, and let $\mathcal{U} \subset T_{\mathcal{W}}\,\mathrm{Gr}_{n,m}$ be defined as in (5.30). By Proposition 5.4.7, we get that every $\mathcal{W}' \in \mathrm{Gr}_{n,m}$ is of the form $\mathcal{W}' = \overline{\exp}_{\mathcal{W}}(v)$ for some $v = \left[Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix}\right] \in \overline{\mathcal{U}}$. Moreover, the geodesic distance between $\mathcal{W}$ and $\mathcal{W}'$ is given by $d_{\mathrm{g}}(\mathcal{W}, \mathcal{W}') = \|R\|_F$. It remains to show that the principal angles between $\mathcal{W}$ and $\mathcal{W}'$ are given by the singular values $\alpha_1 \geq \ldots \geq \alpha_m$ of $R$.

Note that as $v \in \overline{\mathcal{U}}$, we have $\alpha_1 = \|R\| \leq \frac{\pi}{2}$. By (5.22) we may assume w.l.o.g. that $R = \begin{pmatrix} A \\ 0 \end{pmatrix}$, with $A = \mathrm{diag}(\alpha_1, \ldots, \alpha_m)$. To simplify the notation we may furthermore assume that $\mathcal{W} = [I_n]$ and $v = \left[I_n, \begin{pmatrix} 0 & -A & 0 \\ A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]$.

A formula for the unit-speed geodesic (5.31) is given in (5.28). It follows that an orthonormal basis of $\overline{\exp}_{\mathcal{W}}(v)$ is given by

$$\{\cos(\alpha_i) \cdot e_i + \sin(\alpha_i) \cdot e_{m+i} \mid i = 1, \ldots, m\} ,$$

where $e_1, \ldots, e_n \in \mathbb{R}^n$ denote the canonical basis vectors. Note that an orthonormal basis of $\mathcal{W} = [I_n]$ is given by $\{e_i \mid i = 1, \ldots, m\}$. As $\alpha_i \leq \frac{\pi}{2}$, we have $\cos(\alpha_i) \geq 0$, $i = 1, \ldots, m$. By Definition 5.4.1 we get that the principal angles between $\mathcal{W}$ and $\overline{\exp}_{\mathcal{W}}(v)$ are given by $\alpha_1, \ldots, \alpha_m$. This shows (5.32).

The equivalence of the Hausdorff metric and the geodesic metric follows from (5.32) and (5.26) and the equivalence of the 2-norm and the $\infty$-norm.

The claim that the map $\mathcal{W} \mapsto \mathcal{W}^\perp$ is an isometry, if $\mathrm{Gr}_{n,m}$ and $\mathrm{Gr}_{n,n-m}$ are both endowed with the geodesic metric, follows from (5.32) and Proposition 5.4.3. □

**Remark 5.4.9.** We content ourselves in this paper with the Hausdorff and the geodesic metric on $\mathrm{Gr}_{n,m}$. See [25, §4.3] for a longer list of common metrics on $\mathrm{Gr}_{n,m}$ with the corresponding expressions in the principal angles.

## 5.5 Closest elements in the Sigma set

In this section we will provide a transition to the main result of this paper which is the Grassmannian tube formula that we will prove in Chapter 6. Similar to as we did for linear operators in Section 2.1.2 we will analyze how to perturb linear subspaces so that the result contains a given point. This will give us a clear picture about the geodesics towards the set $\Sigma_m(C)$, $C$ being a regular cone (cf. Section 2.3).

Recall that for $C \subset \mathbb{R}^n$ a regular cone the set $\Sigma_m(C)$ was given by

$$\Sigma_m(C) = \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \mathcal{W} \cap \mathrm{int}(C) = \emptyset \text{ and } \mathcal{W} \cap \partial C \neq \{0\}\}$$

(cf. Definition 2.3.7). For a single point $p \in S^{n-1}$ we define

$$\Sigma_m(p) := \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid p \in \mathcal{W}\} .$$

Note that we trivially have

$$\Sigma_m(C) \subset \bigcup_{p \in \partial K} \Sigma_m(p) , \tag{5.33}$$

where $K = C \cap S^{n-1}$.

Next, we will treat the question 'Given some $\mathcal{W}_0 \notin \Sigma_m(p)$, what is the/a closest element in $\Sigma_m(p)$?'. The following construction is natural: Let $q$ be the projection of $p$ on the subsphere $S_0 := \mathcal{W}_0 \cap S^{n-1}$, i.e., $q = \|\Pi_{\mathcal{W}_0}(p)\|^{-1} \cdot \Pi_{\mathcal{W}_0}(p)$, where $\Pi_{\mathcal{W}_0}$ denotes the orthogonal projection onto $\mathcal{W}_0$. The subspace $\mathcal{W}_1 := \bar{\mathcal{W}} + \mathbb{R}\,p$, where $\bar{\mathcal{W}} := q^\perp \cap \mathcal{W}_0$, lies in $\Sigma_m(p)$. It is obtained by rotating the point $q \in \mathcal{W}_0$ onto $p$ while leaving the orthogonal complement $q^\perp \cap \mathcal{W}_0$ fixed. It is geometrically plausible that this yields a closest point in $\Sigma_m(p)$. Proposition 5.5.2 below shows that this is true both for the geodesic or the Hausdorff distance. Moreover, the above described point in $\Sigma_m(p)$ is the unique element, which minimizes the geodesic distance to $\mathcal{W}_0$. Before we give this proposition we introduce a notation, which describes the above construction of $\mathcal{W}_1$.

**Definition 5.5.1.** For $\mathcal{W} \in \mathrm{Gr}_{n,m}$ and $x \in \mathbb{R}^n \setminus \mathcal{W}^\perp$ we define $\mathcal{W}(\to x) \in \mathrm{Gr}_{n,m}$ via

$$\mathcal{W}(\to x) := (y^\perp \cap \mathcal{W}) + \mathbb{R}\,x\ ,$$

where $y := \Pi_{\mathcal{W}}(x)$, and $\Pi_{\mathcal{W}}$ denotes the orthogonal projection onto $\mathcal{W}$.

Recall that $d(.,.)$ denotes the spherical distance in $S^{n-1}$ (cf. Section 3.1).

**Proposition 5.5.2.** *Let $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$, $S_0 := \mathcal{W}_0 \cap S^{n-1}$, and $p \in S^{n-1} \setminus \mathcal{W}_0^\perp$. Furthermore, let*

$$\mathcal{W}_1 := \mathcal{W}_0(\to p)\ . \tag{5.34}$$

*Then the following holds:*

1. $\mathcal{W}_1 \in \Sigma_m(p)$,
2. $d_{\mathrm{H}}(\mathcal{W}_0, \mathcal{W}_1) = d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}_1) = d(p, S_0)$,
3. $d(p, S_0) = d_{\mathrm{H}}(\mathcal{W}_0, \Sigma_m(p)) = d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(p))$,
4. *$\mathcal{W}_1$ is the unique element in $\Sigma_m(p)$, which minimizes the geodesic distance to $\mathcal{W}_0$, i.e.,*

$$\{\mathcal{W}_1\} = \{\mathcal{W}_0(\to p)\} = \mathrm{argmin}\{d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}) \mid \mathcal{W} \in \Sigma_m(p)\}\ . \tag{5.35}$$

As a corollary we get a complete picture about closest elements in $\Sigma_m(C)$.

**Corollary 5.5.3.** *Let $C \subset \mathbb{R}^n$ be a regular cone and let $K := C \cap S^{n-1}$.*

1. *Let $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$ and $S_0 := \mathcal{W}_0 \cap S^{n-1}$ such that $\mathcal{W}_0 \cap C = \{0\}$. Then the distance of $\mathcal{W}_0$ to $\Sigma_m(C)$ is given by*

$$d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(C)) = d_{\mathrm{H}}(\mathcal{W}_0, \Sigma_m(C)) = d(S_0, K)\ . \tag{5.36}$$

*Furthermore, the elements in $\Sigma_m(C)$, which minimize the geodesic distance to $\mathcal{W}_0$ are given by*

$$\begin{aligned}&\mathrm{argmin}\{d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}) \mid \mathcal{W} \in \Sigma_m(C)\} \\ &\qquad = \big\{\mathcal{W}_0(\to p) \mid p \in \mathrm{argmin}\{d(p', S_0) \mid p' \in \partial K\}\big\}\ .\end{aligned} \tag{5.37}$$

2. *If $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$ is such that $\mathcal{W}_0 \cap C \neq \{0\}$ then*

$$d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(C)) = d_{\mathrm{H}}(\mathcal{W}_0, \Sigma_m(C)) = d(S_0^\perp, \breve{K})\ , \tag{5.38}$$

*where $S_0^\perp = \mathcal{W}_0^\perp \cap S^{n-1}$.*

*Proof.* (1) If we have $\mathcal{W}_0 \cap C = \{0\}$, then the distance to $\Sigma_m(C)$ is given by the distance to the set $\{\mathcal{W} \mid \mathcal{W} \cap C \neq \{0\}\}$ (cf. Definition 2.3.7), i.e.,

$$d_*(\mathcal{W}_0, \Sigma_m(C)) = \min\{d_*(\mathcal{W}_0, \mathcal{W}) \mid \mathcal{W} \cap C \neq \{0\}\} ,$$

where $d_* = d_{\mathrm{g}}$ or $d_* = d_{\mathrm{H}}$. Since $\{\mathcal{W} \mid \mathcal{W} \cap C \neq \{0\}\} = \bigcup_{p \in K} \Sigma_m(p)$, we get

$$\begin{aligned} \min\{d_*(\mathcal{W}_0, \mathcal{W}) \mid \mathcal{W} \cap C \neq \{0\}\} &= \min\{d_*(\mathcal{W}_0, \Sigma_m(p)) \mid p \in K\} \\ &\overset{\text{Prop. 5.5.2}}{=} \min\{d(S_0, p) \mid p \in K\} \\ &= d(S_0, K) , \end{aligned}$$

which shows (5.36). As for the equality in (5.37) note that we have

$$d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(C)) = d(S_0, K) \overset{\text{Prop. 5.5.2}}{=} \min\{d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(p)) \mid p \in \partial K\} . \tag{5.39}$$

From the inclusion in (5.33) and from (5.35) we thus get

$$\begin{aligned} &\operatorname{argmin}\{d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}) \mid \mathcal{W} \in \Sigma_m(C)\} \\ &\qquad \subseteq \big\{\mathcal{W}_0(\to p) \mid p \in \operatorname{argmin}\{d(p', S_0) \mid p' \in \partial K\}\big\} . \end{aligned}$$

In fact, this inclusion also holds if the geodesic distance is replaced by the Hausdorff distance. For the other inclusion we need to argue about the specific form of the geodesic metric.

Let $p \in \operatorname{argmin}\{d(p', S_0) \mid p' \in K\}$, and let $\mathcal{W}_1 := \mathcal{W}_0(\to p)$. By Proposition 5.5.2 the element $\mathcal{W}_1$ is the unique element in $\Sigma_m(p)$, which minimizes the distance to $\mathcal{W}_0$. Note that $\dim(\mathcal{W}_0 \cap \mathcal{W}_1) = m - 1$, so that the first $m-1$ principal angles between $\mathcal{W}_0$ and $\mathcal{W}_1$ are 0. The unique shortest-length geodesic from $\mathcal{W}_0$ to $\mathcal{W}_1$ rotates the point $q := \|\Pi_{\mathcal{W}_0}(p)\|^{-1} \cdot \Pi_{\mathcal{W}_0}(p)$ onto the point $p$ and leaves the subspace $\bar{\mathcal{W}} := q^\perp \cap \mathcal{W}_0$ invariant. Since $d(p, S_0) = d(K, S_0)$, it follows that $\mathcal{W} \cap C = \{0\}$ for all $\mathcal{W}$ on the geodesic between $\mathcal{W}_0$ and $\mathcal{W}_1$. From this it follows that $\mathcal{W}_1 \in \Sigma_m(C)$, and as we have (5.39), the inclusion '$\supseteq$' in (5.37) follows. This finishes the proof of part (1) of the claim.

As for part (2) of the claim, recall from Proposition 5.4.4 and Corollary 5.4.8 that the map $\mathrm{Gr}_{n,m} \to \mathrm{Gr}_{n,n-m}$, $\mathcal{W} \to \mathcal{W}^\perp$, is an isometry if we consider both $\mathrm{Gr}_{n,m}$ and $\mathrm{Gr}_{n,n-m}$ being endowed with the geodesic metric or both being endowed with the Hausdorff metric. Furthermore, under this map we have a bijection $\Sigma_m(C) \to \Sigma_{n-m}(\breve{C})$, so that $d_*(\mathcal{W}, \Sigma_m(C)) = d_*(\mathcal{W}^\perp, \Sigma_{n-m}(\breve{C}))$, where $d_* = d_{\mathrm{g}}$ or $d_* = d_{\mathrm{H}}$. So (5.38) follows from (5.36) via this duality. □

*Proof of Proposition 5.5.2.* First of all, as $p \in \mathcal{W}_1$ we have $\mathcal{W}_1 \in \Sigma_m(p)$, so part (1) of the claim is trivial.

Let $\bar{X} \in \mathbb{R}^{n \times (m-1)}$ be such that the columns of $\bar{X}$ form an orthonormal basis of $\bar{\mathcal{W}} \subseteq \mathcal{W}_1 \cap \mathcal{W}_2$. If we set

$$X_0 := \begin{pmatrix} q & \bar{X} \end{pmatrix} , \; X_1 := \begin{pmatrix} p & \bar{X} \end{pmatrix} \in \mathbb{R}^{n \times m} , \tag{5.40}$$

then the columns of $X_i$ form an orthonormal basis of $\mathcal{W}_i$, $i = 0, 1$. We get

$$X_0^T \cdot X_1 = \begin{pmatrix} q^T p & 0 \\ 0 & I_{m-1} \end{pmatrix} ,$$

and $\arccos(q^T p) = d(p, q) = d(p, S_0)$. This implies

$$d_{\mathrm{H}}(\mathcal{W}_0, \mathcal{W}_1) = d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}_1) = d(p, S_0) .$$

In particular, we have $d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(p)) \leq d(p, S_0)$.

As for part (3) of the claim note that if $\mathcal{W}_1' \in \Sigma_m(p)$, i.e., $p \in \mathcal{W}_1'$, then

$$d_{\mathrm{g}}(\mathcal{W}_0, \mathcal{W}_1') \geq d_{\mathrm{H}}(\mathcal{W}_0, \mathcal{W}_1') \geq \min\{\alpha \mid S_1' \subseteq \mathcal{T}(S_0, \alpha)\} \geq d(p, S_0) \ .$$

This finishes the proof of the equalities $d_{\mathrm{g}}(\mathcal{W}_0, \Sigma_m(p)) = d_{\mathrm{H}}(\mathcal{W}_0, \Sigma_m(p)) = d(p, S_0)$.

To show part (4) of the claim let again $\mathcal{W}_1' \in \Sigma_m(p)$. Furthermore, let $\bar{X}' \in \mathbb{R}^{n\times(m-1)}$ be such that the columns of $\bar{X}'$ form an orthonormal basis of $p^\perp \cap \mathcal{W}_1'$, and let $X_1' := \begin{pmatrix} p & \bar{X}' \end{pmatrix} \in \mathbb{R}^{n\times m}$. Then the columns of $X_1'$ form an orthonormal basis of $\mathcal{W}_1'$, and with $X_0$ defined as in (5.40) we get

$$X_0^T X_1' = \begin{pmatrix} q^T p & 0 \\ 0 & \bar{X}^T \bar{X}' \end{pmatrix} \ .$$

In particular, the singular values of $X_0^T X_1'$ are given by $q^T p$ and the singular values of $\bar{X}^T \bar{X}'$. The singular values of $\bar{X}^T \bar{X}'$ are all $\leq 1$, as the largest of them is given by the operator norm, and we have $\|\bar{X}^T \bar{X}'\| \leq \|\bar{X}\| \cdot \|\bar{X}'\| = 1$. The singular values are $= 1$, i.e., the arccosines of them are $= 0$, iff the columns of $\bar{X}$ and the columns of $\bar{X}'$ span the same subspace. This equivalence is verified easily. As $\mathcal{W}_1 = \bar{\mathcal{W}} + \mathbb{R}\, p$, we get part (4) of the claim. $\square$

# Chapter 6

# A tube formula for the Grassmann bundle

In this chapter we will prove the main result of this paper, a formula for the volume of the tube of the Sigma set in the Grassmann manifold. This formula is a generalization of Weyl's classical tube formula for the sphere and can be interpreted as an average case analysis of the Grassmann condition of the convex feasibility problem.

## 6.1 Main results

Throughout this chapter let $C \subset \mathbb{R}^n$ be a regular cone, and let $K = C \cap S^{n-1}$. Recall that in Section 2.3 we have defined the sets

$$\begin{aligned} \mathcal{F}^{\mathrm{D}}_{\mathrm{G}}(C) &= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \mathcal{W} \cap C \neq \{0\}\} \;, \\ \mathcal{F}^{\mathrm{P}}_{\mathrm{G}}(C) &= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \mathcal{W}^{\perp} \cap \breve{C} \neq \{0\}\} \;. \end{aligned}$$

Furthermore, the interior of the primal feasible set $\mathcal{F}^{\mathrm{P}}_{\mathrm{G}}(C)$ is given by the dual infeasible set $\mathcal{I}^{\mathrm{D}}_{\mathrm{G}}(C) = \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \mathcal{W} \cap C = \{0\}\}$. The set of ill-posed inputs (cf. Definition 2.3.7) is given by

$$\begin{aligned} \Sigma_m(C) &= \mathcal{F}^{\mathrm{D}}_{\mathrm{G}}(C) \cap \mathcal{F}^{\mathrm{P}}_{\mathrm{G}}(C) \\ &= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \mathcal{W} \cap \mathrm{int}(C) = \emptyset \text{ and } \mathcal{W} \cap \partial C \neq \{0\}\} \;. \end{aligned} \tag{6.1}$$

To ease the notation we will occasionally write $\Sigma_m(K)$ instead of $\Sigma_m(C)$, or we will drop the brackets and simply write $\Sigma_m$.

Recall from Corollary 5.5.3 that for $\mathcal{W} \in \mathrm{Gr}_{n,m}$ the geodesic distance of $\mathcal{W}$ to $\Sigma_m$ coincides with the Hausdorff distance of $\mathcal{W}$ to $\Sigma_m$. To take this into account, let $d_*(.,.)$ denote in the following paragraph the geodesic or the Hausdorff distance.

As the Grassmann condition of a point $\mathcal{W} \in \mathrm{Gr}_{n,m}$ is given by

$$\mathscr{C}_{\mathrm{G}}(\mathcal{W}) = \frac{1}{\sin d_*(\mathcal{W}, \Sigma_m)}$$

(cf. Proposition 2.3.8), we define the (primal/dual) tube of radius $\alpha$ around $\Sigma_m$ via

$$\begin{aligned} \mathcal{T}(\Sigma_m, \alpha) &:= \{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid d_*(\mathcal{W}, \Sigma_m) \leq \alpha\} \;, \\ \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) &:= \mathcal{T}(\Sigma_m, \alpha) \cap \mathcal{F}^{\mathrm{P}}_{\mathrm{G}} \;, \\ \mathcal{T}^{\mathrm{D}}(\Sigma_m, \alpha) &:= \mathcal{T}(\Sigma_m, \alpha) \cap \mathcal{F}^{\mathrm{D}}_{\mathrm{G}} \;, \end{aligned}$$

where $0 \leq \alpha \leq \frac{\pi}{2}$. Note that

$$\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \cup \mathcal{T}^{\mathrm{D}}(\Sigma_m, \alpha) = \mathcal{T}(\Sigma_m, \alpha) \,,$$
$$\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \cap \mathcal{T}^{\mathrm{D}}(\Sigma_m, \alpha) = \Sigma_m \,.$$

Note also that the isometry $\mathrm{Gr}_{n,m} \to \mathrm{Gr}_{n,n-m}$, $\mathcal{W} \mapsto \mathcal{W}^{\perp}$, induces a bijection between $\mathcal{T}^{\mathrm{D}}(\Sigma_m(C), \alpha)$ and $\mathcal{T}^{\mathrm{P}}(\Sigma_{n-m}(\breve{C}), \alpha)$. So it suffices to know formulas for the (relative) volume of the primal tube $\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha)$. The primal tube has the following simple characterization

$$\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) = \{\mathrm{lin}(S) \mid S \in \mathcal{S}^{m-1}(S^{n-1}) \,,\ S \cap \mathrm{int}(K) = \emptyset \,,\ d(S, K) \leq \alpha\} \,, \quad (6.2)$$

where $\mathcal{S}^{m-1}(S^{n-1})$ denotes the set of $(m-1)$-dimensional subspheres of $S^{n-1}$ (cf. Section 3.2). This characterization follows from the first part of Corollary 5.5.3.

**Theorem 6.1.1.** *Let $C \subset \mathbb{R}^n$ be a regular cone and let $K = C \cap S^{n-1}$. Then for $1 \leq m \leq n-1$ and $0 \leq \alpha \leq \frac{\pi}{2}$*

$$\mathrm{rvol}\ \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \leq \frac{2m(n-m)}{n} \binom{n/2}{m/2} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} |d_{ij}^{nm}| \cdot I_{n,i}(\alpha) \,, \quad (6.3)$$

*where*

- $\binom{n/2}{m/2} = \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+2}{2}) \cdot \Gamma(\frac{n-m+2}{2})}$ *and* $\left[\begin{smallmatrix} n-2 \\ j \end{smallmatrix}\right] = \frac{\sqrt{\pi} \cdot \Gamma(\frac{n-1}{2})}{\Gamma(\frac{j+1}{2}) \cdot \Gamma(\frac{n-j-1}{2})}$ *(cf. Section 4.1.4),*
- *$V_j(K)$ denotes the $j$th spherical intrinsic volume of $K$ (cf. Section 4.4),*
- *$I_{n,i}(\alpha) = \int_0^{\alpha} \cos(\rho)^i \cdot \sin(\rho)^{n-2-i}\, d\rho$ (cf. Section 4.3), and*
- *the constants $d_{ij}^{nm}$ are defined for $i + j + m \equiv 1 \pmod 2$ and*

$$0 \leq \tfrac{i-j}{2} + \tfrac{m-1}{2} \leq m-1 \,, \quad 0 \leq \tfrac{i+j}{2} - \tfrac{m-1}{2} \leq n-m-1 \quad (6.4)$$

*via*

$$d_{ij}^{nm} \;:=\; (-1)^{\frac{i-j}{2} - \frac{m-1}{2}} \cdot \frac{\binom{m-1}{\frac{i-j}{2} + \frac{m-1}{2}} \cdot \binom{n-m-1}{\frac{i+j}{2} - \frac{m-1}{2}}}{\binom{n-2}{j}} \,, \quad (6.5)$$

*and $d_{ij}^{nm} := 0$ otherwise.*

*Furthermore, if $\alpha_0 := \sup\{\alpha \mid \mathcal{T}(K, \alpha) \in \mathcal{K}^c(S^{n-1})\}$, then for $0 \leq \alpha \leq \alpha_0$*

$$\mathrm{rvol}\ \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) = \frac{2m(n-m)}{n} \binom{n/2}{m/2} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot I_{n,i}(\alpha) \,. \quad (6.6)$$

**Remark 6.1.2.** The most important part of the above theorem is the estimate in (6.3), which holds for all $0 \leq \alpha \leq \frac{\pi}{2}$. For the interesting choices of the cone $C$ we always have $\alpha_0 = 0$ (cf. Proposition 3.1.16), so the second part (6.6) of the theorem yields no statement for these cases. Moreover, we will prove the equality in (6.6) in Section 6.5 only for the case $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$. The reason for this restriction lies in the fact that the map

$$\alpha_{\max} \colon \mathcal{K}(S^{n-1}) \to \mathbb{R} \,, \quad K \mapsto \sup\{\alpha \mid \mathcal{T}(K, \alpha) \in \mathcal{K}(S^{n-1})\}$$

$$D_{7,1} = \begin{pmatrix} 1 &&&&& \\ & 1 &&&& \\ && 1 &&& \\ &&& 1 && \\ &&&& 1 & \\ &&&&& 1 \end{pmatrix}, \qquad D_{7,6} = \begin{pmatrix} &&&&& -1 \\ &&&& 1 & \\ &&& -1 && \\ && 1 &&& \\ & -1 &&&& \\ 1 &&&&& \end{pmatrix}$$

$$D_{7,2} = \begin{pmatrix} & -\frac{1}{5} &&&& \\ 1 & 0 & -\frac{2}{5} &&& \\ & \frac{4}{5} & 0 & -\frac{3}{5} && \\ && \frac{3}{5} & 0 & -\frac{4}{5} & \\ &&& \frac{2}{5} & 0 & -1 \\ &&&& \frac{1}{5} & \end{pmatrix}, \qquad D_{7,5} = \begin{pmatrix} &&&& \frac{1}{5} & \\ &&& -\frac{2}{5} & 0 & 1 \\ && \frac{3}{5} & 0 & -\frac{4}{5} & \\ & -\frac{4}{5} & 0 & \frac{3}{5} && \\ 1 & 0 & -\frac{2}{5} &&& \\ & \frac{1}{5} &&&& \end{pmatrix}$$

$$D_{7,3} = \begin{pmatrix} && \frac{1}{10} &&& \\ & -\frac{2}{5} & 0 & \frac{3}{10} && \\ 1 & 0 & -\frac{3}{5} & 0 & \frac{3}{5} & \\ & \frac{3}{5} & 0 & -\frac{3}{5} & 0 & 1 \\ && \frac{3}{10} & 0 & -\frac{2}{5} & \\ &&& \frac{1}{10} && \end{pmatrix}, \qquad D_{7,4} = \begin{pmatrix} &&& -\frac{1}{10} && \\ && \frac{3}{10} & 0 & -\frac{2}{5} & \\ & -\frac{3}{5} & 0 & \frac{3}{5} & 0 & -1 \\ 1 & 0 & -\frac{3}{5} & 0 & \frac{3}{5} & \\ & \frac{2}{5} & 0 & -\frac{3}{10} && \\ && \frac{1}{10} &&& \end{pmatrix}$$

Table 6.1: The coefficient matrix $D_{n,m} = (d_{ij}^{nm})_{i,j=0,\ldots,n-2}$.

is *not continuous* if $n \geq 3$ (cf. Remark 3.1.17). In Section A.3 in the appendix we will use the kinematic formula for an alternative computation of $\operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha)$ for $0 \leq \alpha \leq \alpha_0$. As a corollary, this computation will allow us to transfer (6.6) from $\mathcal{K}^{\mathrm{sm}}(S^{n-1})$ to $\mathcal{K}(S^{n-1})$.

We may summarize that the assumption $0 \leq \alpha \leq \alpha_0$ is a very strong assumption, which is not satisfied for most cases of interest. The estimate in (6.3) is a stable result, which can be used for an average analysis of the Grassmann condition (cf. Chapter 7).

**Remark 6.1.3.** We may (partly) recover from (6.6) Weyl's tube formula by considering the case $m = 1$. In this case we have $d_{ij}^{n1} = \delta_{ij}$. Using the identities from Proposition 4.1.20, and using $\mathcal{O}_j \cdot \mathcal{O}_{n-2-j} \cdot I_{n,j}(\alpha) = \mathcal{O}_{n-1,j}(\alpha)$ and $\mathcal{O}_{k-1} = k \cdot \omega_k$, we get

$$\begin{aligned} \operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_1, \alpha) &\overset{(6.6)}{=} 2 \cdot \frac{n-1}{n} \cdot \frac{2 \cdot \omega_{n-1}}{\omega_n} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \frac{\mathcal{O}_j \cdot \mathcal{O}_{n-2-j}}{2 \cdot \mathcal{O}_{n-2}} \cdot I_{n,j}(\alpha) \\ &= \frac{2}{\mathcal{O}_{n-1}} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \mathcal{O}_{n-1,j}(\alpha) \\ &\overset{(4.39)}{=} 2 \cdot \frac{\operatorname{vol}_{n-1} \mathcal{T}(K, \alpha) - \operatorname{vol}_{n-1} K}{\mathcal{O}_{n-1}} . \end{aligned}$$

This result is correct as $\mathcal{T}^{\mathrm{P}}(\Sigma_1, \alpha)$ is the projective image of $\mathcal{T}(K, \alpha) \setminus \operatorname{int}(K)$, and $\operatorname{vol} \mathbb{P}^{n-1} = \frac{\mathcal{O}_{n-1}}{2}$.

Nevertheless, this only yields part of Weyl's spherical tube formula, which holds for all $0 \leq \alpha \leq \frac{\pi}{2}$, whereas (6.6) only holds for $0 \leq \alpha \leq \alpha_0$. The determination of the largest $\alpha$, for which (6.6) holds, remains an open problem for $1 < m < n - 1$.

The twisting coefficients $d_{ij}^{nm}$ appear in the following formal polynomial identity

$$(X-Y)^{m-1}\cdot(1+XY)^{n-m-1}=\sum_{i,j=0}^{n-2}\binom{n-2}{j}\cdot d_{ij}^{nm}\cdot X^{n-2-j}\cdot Y^{n-2-i}$$

(cf. proof of Proposition 6.4.6 in Section 6.4 below). In order to get a better feeling for these coefficients, let us have a look at the coefficient matrices $D_{n,m}\in\mathbb{R}^{(n-1)\times(n-1)}$ where $D_{n,m}=(d_{ij}^{nm})_{i,j=0,\ldots,n-2}$. See Table 6.1 for a display of $D_{n,m}$ for some concrete values for $n$ and $m$. This table shows that the nonzero coefficients $d_{ij}^{nm}$ lie in a "perforated" rectangle in $D_{n,m}$, which is determined by the inequalities (6.4) and the parity condition $i+j+m\equiv 1 \pmod 2$. The table also shows the symmetries

$$d_{i,n-2-j}^{n,n-m}=\pm d_{ij}^{nm}\;,\quad d_{n-2-i,n-2-j}^{nm}=\pm d_{ij}^{nm}\;.$$

Observe that a comparison between (6.6) and Weyl's tube formula (4.39) reveals that basically the $I$-functions in the classical tube formula are replaced by the "twisted $I$-functions" $\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha)$. In Section A.3 we will give some alternative formulas for the twisted $I$-functions by making use of the principal kinematic formula.

In the following corollary we derive from Theorem 6.1.1 results about the relative volume of the whole tube $\mathcal{T}(\Sigma_m,\alpha)$ from those results about the "half-tube" $\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha)$.

**Corollary 6.1.4.** *Let the notation be as in Theorem 6.1.1. Then for* $0\leq\alpha\leq\alpha_0$

$$\text{rvol}\;\mathcal{T}(\Sigma_m,\alpha)=\frac{4m(n-m)}{n}\binom{n/2}{m/2}\cdot\sum_{\substack{j=0\\ j\equiv n-m-1\\ \text{mod } 2}}^{n-2} V_j(K)\cdot\begin{bmatrix}n-2\\ j\end{bmatrix}\cdot\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha)\;,\quad(6.7)$$

*and for all* $0\leq\alpha\leq\frac{\pi}{2}$

$$\text{rvol}\;\mathcal{T}(\Sigma_m,\alpha)\leq\frac{4m(n-m)}{n}\binom{n/2}{m/2}\cdot\sum_{j=0}^{n-2} V_j(K)\cdot\begin{bmatrix}n-2\\ j\end{bmatrix}\cdot\sum_{i=0}^{n-2}|d_{ij}^{nm}|\cdot I_{n,i}(\alpha)\,.\quad(6.8)$$

*Proof.* The isometry $\mathrm{Gr}_{n,m}\to\mathrm{Gr}_{n,n-m}$, $\mathcal{W}\mapsto\mathcal{W}^\perp$, induces a bijection between $\mathcal{T}^{\mathrm{D}}(\Sigma_m(C),\alpha)$ and $\mathcal{T}^{\mathrm{P}}(\Sigma_{n-m}(\breve{C}),\alpha)$. So we get

$$\begin{aligned}\text{rvol}\;\mathcal{T}(\Sigma_m(C),\alpha)&=\text{rvol}\;\mathcal{T}^{\mathrm{P}}(\Sigma_m(C),\alpha)+\text{rvol}\;\mathcal{T}^{\mathrm{D}}(\Sigma_m(C),\alpha)\\&=\text{rvol}\;\mathcal{T}^{\mathrm{P}}(\Sigma_m(C),\alpha)+\text{rvol}\;\mathcal{T}^{\mathrm{P}}(\Sigma_{n-m}(\breve{C}),\alpha)\;,\end{aligned}$$

and Theorem 6.1.1 yields formulas resp. estimates for these quantities. Since $V_j(\breve{K})=V_{n-2-j}(K)$ (cf. Proposition 4.4.10) and

$$d_{i,n-2-j}^{n,n-m}\;=\;(-1)^{n-m-1-j}\cdot d_{ij}^{nm}\;,$$

we get

$$\sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot I_{n,i}(\alpha) + \sum_{j=0}^{n-2} V_j(\breve{K}) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot I_{n,i}(\alpha)$$
$$= \sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} (d_{ij}^{nm} + d_{i,n-2-j}^{nm}) \cdot I_{n,i}(\alpha)$$
$$= 2 \cdot \sum_{\substack{j=0 \\ j \equiv n-m-1 \\ \bmod 2}}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot I_{n,i}(\alpha) .$$

This implies the exact formula (6.7). The estimate (6.8) follows analogously with the observation $|d_{i,n-2-j}^{n,n-m}| = |d_{ij}^{nm}|$, so that we have no cancellation in this case. □

Finally, we consider the asymptotics of the tube formulas for $\alpha \to 0$. Note that for the $I$-functions we have the asymptotics (cf. Section 4.3)

$$I_{n,i}(\alpha) \sim \tfrac{1}{n-1-i} \cdot \alpha^{n-1-i} \quad \text{for } \alpha \to 0 . \tag{6.9}$$

From this we can derive an asymptotic estimate of the relative volume of the tube around $\Sigma_m$, that we state in the following corollary.

**Corollary 6.1.5.** *Let the notation be as in Theorem 6.1.1. If $\alpha_0 > 0$, then for $\alpha \to 0$ we have*

$$\begin{aligned} \operatorname{rvol} \mathcal{T}(\Sigma_m, \alpha) &= 8 \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n-m}{2})} \cdot V_{n-m-1}(K) \cdot \alpha + O(\alpha^2) \qquad (6.10) \\ &< 4 \cdot \sqrt{m(n-m)} \cdot V_{n-m-1}(K) \cdot \alpha + O(\alpha^2) , \end{aligned}$$

*where the constant in the $O$-notation may depend on $n$, $m$, and $K$. In the case $\alpha_0 = 0$ the equality* (6.10) *still holds as an inequality.*

*Proof.* The asymptotics of the $I$-functions in (6.9) show that the linear term of the tube formulas (6.7) and (6.8) arises from the summands which involve $I_{n,n-2}(\alpha)$. The constant $d_{ij}^{nm}$ for $i = n-2$ is zero except for $j = n-m-1$, where it has the value $\binom{n-2}{m-1}^{-1}$. Using the identities from Proposition 4.1.20 we compute

$$\begin{aligned} \frac{4m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2} \cdot \left[\begin{smallmatrix} n-2 \\ m-1 \end{smallmatrix}\right]}{\binom{n-2}{m-1}} &\overset{(4.21)}{=} \frac{4m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2}}{\binom{(n-2)/2}{(m-1)/2}} \\ &\overset{(4.18)}{=} 8 \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n-m}{2})} . \end{aligned}$$

From the estimate $\Gamma(z + \frac{1}{2}) < \sqrt{z} \cdot \Gamma(z)$ (cf. Section 4.1.4) we finally get

$$\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n-m}{2})} < \frac{\sqrt{m(n-m)}}{2} .$$ □

**Remark 6.1.6.** The asymptotic question as treated in Corollary 6.1.5 falls in the domain of so-called contact measures. This was initiated by Firey in [27] (see [50, §4] for an overview of the development of this topic). While originally only considered in euclidean space, contact measures were also considered in more general spaces like the sphere and hyperbolic space, and even in general homogeneous spaces (see [56], [57], [58]). So the formula (6.10) has been known before.

### 6.1.1 Proof strategy

The main idea of the proof is the following. First of all, we will show that for the estimate (6.3) it suffices to assume that $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ is a smooth cap. In this case, if a subspace $\mathcal{W} \in \mathrm{Gr}_{n,m}$ lies in $\Sigma_m$, i.e., if it touches the cap $K$, then it intersects $K$ in exactly one point. This intersection point $p$ lies in the boundary $\partial K =: M$, which is a hypersurface of the unit sphere $S^{n-1}$. So $\mathcal{W}$ is a subspace of $T_pM + \mathbb{R}\,p$. Conversely, if we choose a point $p \in M$, and if we choose a $(m-1)$-dimensional subspace $\mathcal{Y}$ in the tangent space $T_pM$, then the $m$-dimensional subspace $\mathcal{W} := \mathcal{Y} + \mathbb{R}\,p$ touches the cap $K$ in the point $p$ and thus lies in $\Sigma_m$.

This construction indicates that for $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ the set $\Sigma_m$ is a smooth manifold. More precisely, denoting by $\mathrm{Gr}(M, m-1)$ the $(m-1)$th Grassmann bundle, i.e.,

$$\mathrm{Gr}(M, m-1) = \{(p, \mathcal{Y}) \mid p \in M\ ,\ \mathcal{Y} \subseteq T_pM\ (m-1)\text{-dimensional subspace}\}\ ,$$

we will show in Section 6.2 the map

$$E_m \colon\ \mathrm{Gr}(M, m-1) \to \Sigma_m\ ,\quad (p, \mathcal{Y}) \mapsto \mathcal{Y} + \mathbb{R}\,p$$

is a diffeomorphism. The main steps in the proof of Theorem 6.1.1 are now the following.

1. For the inequality (6.3) reduce the general case to the case $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$; the equality (6.6) we will only prove in the case $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ (see Section A.3 in the appendix for the transfer to general $K \in \mathcal{K}(S^{n-1})$).

2. Describe a parametrization of $\Sigma_m$, which shows that $\Sigma_m$ is an orientable hypersurface of $\mathrm{Gr}_{n,m}$, and find formulas for its tangent and its normal spaces.

3. If, starting at a point $\mathcal{W} \in \Sigma_m$, we track the geodesic on $\mathrm{Gr}_{n,m}$ in normal direction up to a distance of $\alpha$, and if we do this for every $\mathcal{W} \in \Sigma_m$, then we will get a surjection of the tube $\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha)$ resp. $\mathcal{T}^{\mathrm{D}}(\Sigma_m, \alpha)$ depending on the normal direction. More precisely, let us denote

$$\Upsilon \colon \Sigma_m \times \mathbb{R} \to \mathrm{Gr}_{n,m}\ ,\quad (\mathcal{W}, \rho) \mapsto \overline{\exp}_{\mathcal{W}}(\rho \cdot \nu_\Sigma(\mathcal{W}))\ ,$$

   where $\nu_\Sigma$ denotes a unit normal field on $\Sigma$. Restricting the second component to an interval $[0, \alpha]$ resp. $[-\alpha, 0]$ implies that the image of $\Upsilon$ is the primal tube of radius $\alpha$ resp. the dual tube of radius $\alpha$, depending on the direction of $\nu_\Sigma$. With the aid of the coarea formula, we will get an upper bound for the volume of the corresponding $\alpha$-tube around $\Sigma_m$.

4. It will turn out that in order to get the formulas in Theorem 6.1.1 we will have to compute the expectation of a certain (twisted) characteristic polynomial. Computing this expectation we will finish the proof of Theorem 6.1.1 (for $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ in the equality case).

We finish this section with a lemma that covers the first step. The following three sections will treat the remaining steps, and Section 6.5 will combine the different steps to complete the proof of Theorem 6.1.1.

**Lemma 6.1.7.** *For $\alpha \geq 0$ the function*

$$\mathcal{K}(S^{n-1}) \to \mathbb{R}\ ,\quad K \mapsto \mathrm{rvol}\,\mathcal{T}^{\mathrm{P}}(\Sigma_m(K), \alpha)$$

*is uniformly continuous.*

*Proof.* The proof goes analogously to the proof of Lemma 4.4.2. It suffices to show that the map is continuous, as $\mathcal{K}(S^{n-1})$ is compact (cf. Proposition 3.2.3). From the characterization of the primal tube in (6.2) we get for $K_1, K_2 \in \mathcal{K}(S^{n-1})$

$$K_1 \subseteq \mathcal{T}(K_2,\varepsilon) \;\Rightarrow\; \mathcal{T}^{\mathrm{P}}(\Sigma_m(K_1),\alpha) \subseteq \mathcal{T}^{\mathrm{P}}(\Sigma_m(K_2),\alpha+\varepsilon)\ . \tag{6.11}$$

Let $(K_i)_i$ be a sequence in $\mathcal{K}(S^{n-1})$, which converges to $K \in \mathcal{K}(S^{n-1})$. We need to show that $\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha)$ converges to $\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha)$.

For all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d_H(K,K_i) < \varepsilon$, i.e., $K \subseteq \mathcal{T}(K_i,\varepsilon)$ and $K_i \subseteq \mathcal{T}(K,\varepsilon)$, for all $i \geq N$. From (6.11) we thus get for $i \geq N$

$$\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha) \subseteq \mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha+\varepsilon)\ , \quad \mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha) \subseteq \mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha+\varepsilon)\ .$$

In particular, we have

$$\begin{aligned}\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha) &\leq \operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha+\varepsilon)\ ,\\ \operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha) &\leq \operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha+\varepsilon)\ ,\end{aligned}$$

for all $i \geq N$. This implies

$$\begin{aligned}\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha) &\leq \liminf_{i\to\infty}\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha+\varepsilon)\ ,\\ \limsup_{i\to\infty}\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha) &\leq \operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha+\varepsilon)\ .\end{aligned}$$

Letting $\varepsilon \to 0$, we get $\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K),\alpha) = \lim_{i\to\infty}\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m(K_i),\alpha)$. □

## 6.2 Parametrizing the Sigma set

The goal of this section is to show that for $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ the set $\Sigma_m(K)$ is a smooth orientable hypersurface of $\mathrm{Gr}_{n,m}$, and to describe its tangent and normal spaces.

Before we state the main theorem let us recall the notion of the Grassmann bundle. Let $M$ be a $d$-dimensional Riemannian manifold. The $k$th Grassmann bundle over $M$, $0 \leq k \leq d$, is given by

$$\begin{aligned}\mathrm{Gr}(M,k) &= \{(p,\mathcal{Y}) \mid p \in M\ ,\ \mathcal{Y} \subseteq T_pM\ k\text{-dimensional subspace}\}\\ &= \bigcup_{p\in M}\{p\}\times \mathrm{Gr}(T_pM,k)\ ,\end{aligned}$$

where $\mathrm{Gr}(T_pM,k) := \{\mathcal{Y} \subseteq T_pM\ k\text{-dimensional subspace}\}$.

To see that the Grassmann bundle is indeed a manifold, let $\varphi\colon \mathbb{R}^d \to M$ be a parametrization of an open subset $U$ of $M$. This defines a local trivialization of the Grassmann bundle

$$\Phi\colon \mathbb{R}^d \times \mathrm{Gr}_{d,k} \xrightarrow{\sim} \mathrm{Gr}(U,k)\ , \quad (x,\mathcal{X}) \mapsto \big(\varphi(x), D_x\varphi(\mathcal{X})\big)\ . \tag{6.12}$$

If $\psi\colon \mathbb{R}^{k(d-k)} \to \mathrm{Gr}_{d,k}$ denotes a parametrization of an open subset of the Grassmann manifold $\mathrm{Gr}_{d,k}$, then we can combine this with the local trivialization $\Phi$ via

$$\Psi\colon \mathbb{R}^d \times \mathbb{R}^{k(d-k)} \to \mathrm{Gr}(M,k)\ , \quad (x,X) \mapsto \Phi(x,\psi(X))\ . \tag{6.13}$$

The map $\Psi$ is thus a parametrization of an open subset of the Grassmann bundle $\mathrm{Gr}(M,k)$. For the proof that $\mathrm{Gr}(M,k)$ is indeed a smooth manifold, it remains to show that different parametrizations $\tilde{\varphi}$ and $\tilde{\psi}$ give rise to a smooth transition map

$$\Psi^{-1} \circ \tilde{\Psi}\colon \mathbb{R}^d \times \mathbb{R}^{k(d-k)} \rightarrowtail \mathbb{R}^d \times \mathbb{R}^{k(d-k)} .$$

This is a straightforward exercise.

Note that the dimension of $\mathrm{Gr}(M,k)$ is given by

$$\dim \mathrm{Gr}(M,k) = d + k(d-k) ,$$

where $d = \dim M$. In particular, for $d = n-2$ and $k = m-1$ we get

$$\begin{aligned}\dim \mathrm{Gr}(M,m-1) &= n-2+(m-1)(n-m-1) = m(n-m)-1\\ &= \dim \mathrm{Gr}_{n,m} -1 .\end{aligned}$$

**Theorem 6.2.1.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and $M := \partial K$, and let $E_m$ denote the map*

$$E_m\colon \mathrm{Gr}(M,m-1) \to \mathrm{Gr}_{n,m} , \quad (p,\mathcal{Y}) \mapsto \mathcal{Y} + \mathbb{R}\, p . \tag{6.14}$$

*Then $E_m$ is an injective smooth map whose image is given by $\Sigma_m(K)$. In particular, $\Sigma_m(K)$ is a hypersurface in $\mathrm{Gr}_{n,m}$ and isomorphic to the Grassmann bundle $\mathrm{Gr}(M,m-1)$. Additionally, the hypersurface $\Sigma_m(K)$ is orientable, i.e., there exists a global unit normal vector field $\nu_\Sigma$ on $\Sigma_m(K)$.*

The fact that $E_m$ is a bijection between $\mathrm{Gr}(M,m-1)$ and $\Sigma_m$ follows from simple arguments from spherical convex geometry. Therefore, we will treat this claim in the following lemma.

**Lemma 6.2.2.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and $M := \partial K$, and let $E_m$ be defined as in* (6.14). *Then $E_m$ is a bijection between $\mathrm{Gr}(M,m-1)$ and $\Sigma_m$, and its inverse is given by*

$$E_m^{-1}(\mathcal{W}) = (p, \mathcal{W} \cap p^\perp) , \quad \text{where } \mathcal{W} \cap K = \{p\} . \tag{6.15}$$

*Proof.* As in Section 4.1.2 let $\nu\colon M \to \mathbb{R}^n$ denote the unit normal field such that $\nu(p)$ points inside the cap $K$ for $p \in M = \partial K$. If $(p,\mathcal{Y}) \in \mathrm{Gr}(M,m-1)$, then $\mathcal{W} := \mathcal{Y} + \mathbb{R}\, p$ does not intersect the interior of the cone $C := \mathrm{cone}(K)$, as $\mathcal{W} \subseteq \nu(p)^\perp$. Therefore, by (6.1) we have $\mathcal{W} \in \Sigma_m$. On the other hand, if $\mathcal{W} \in \Sigma_m$, then by Proposition 4.1.11 the intersection $\mathcal{W} \cap K$ consists of a single point $p$. Moreover, for $\mathcal{Y} := \mathcal{W} \cap p^\perp \in \mathrm{Gr}(T_pM, m-1)$ we have $E_m(p,\mathcal{Y}) = \mathcal{W}$. We may conclude, that $E_m$ is a bijection between $\mathrm{Gr}(M,m-1)$ and $\Sigma_m$, and its inverse is given by (6.15). □

To clarify the role of convexity for the proof of Theorem 6.2.1 we will drop the convexity assumption in the following lemmas.

**Lemma 6.2.3.** *Let $M \subset S^{n-1}$ be a submanifold and let $0 \leq m-1 \leq \dim M =: d$. Then the map $E_m$ as defined in* (6.14) *is smooth.*

*Proof.* This claim follows from choosing a parametrization $\varphi\colon \mathbb{R}^d \to M$ and checking that the map

$$\tilde{\Phi}\colon \mathbb{R}^d \times \mathrm{Gr}_{d,m-1} \to \mathrm{Gr}_{n,m} , \quad (x,\mathcal{Y}) \mapsto \mathbb{R}\, \varphi(x) + D_x\varphi(\mathcal{Y})$$

is smooth. This is again a straightforward exercise. The restriction of $E_m$ to $\mathrm{Gr}(U,m-1)$ is then given by $\tilde{\Phi} \circ \Phi^{-1}$, where $\Phi$ is defined as in (6.12), and thus a smooth map. □

For the computation of the derivative of the map $E_m$, we do not need an explicit model of the tangent spaces of the Grassmann bundle $\mathrm{Gr}(M, m-1)$. Nevertheless, it might be helpful to know about the natural decomposition of the tangent space in the direct sum of the vertical space and the horizontal space, that we describe in the following remark.

**Remark 6.2.4.** Let $M \subset S^{n-1}$ be a hypersurface and $1 \leq m \leq n-1$. For $(p, \mathcal{Y}) \in \mathrm{Gr}(M, m-1)$ the tangent space of $\mathrm{Gr}(M, m-1)$ in $(p, \mathcal{Y})$ has a decomposition in the direct sum

$$T_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1) \ = \ T^v_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1) \oplus T^h_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1) \ .$$

The components $T^v_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1)$ and $T^h_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1)$ are called the *vertical space* and the *horizontal space*, respectively. The vertical space is given by the tangent space of the fiber $\{p\} \times \mathrm{Gr}(T_pM, m-1)$, which is a submanifold of the fiber bundle $\mathrm{Gr}(M, m-1)$, i.e.,

$$T^v_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1) = T_{(p,\mathcal{Y})}(\{p\} \times \mathrm{Gr}(T_pM, m-1)) \ .$$

As for the horizontal space, let $c\colon \mathbb{R} \to M$, $c(0) = p$, be a curve through $p$, and let $\mathcal{Y}_t$ denote the parallel transport of $\mathcal{Y}$ along $c$ at time $t$ (cf. Remark 4.1.2). Then the map $c_h\colon \mathbb{R} \to \mathrm{Gr}(M, m-1)$, $c_h(t) = (c(t), \mathcal{Y}_t)$, is a curve through $c_h(0) = (p, \mathcal{Y})$ and thus defines a tangent vector $\dot{c_h}(0) \in T_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1)$. It can be shown that this tangent vector only depends on the tangent vector $\dot{c}(0) \in T_pM$, and that the induced map

$$T_pM \to T_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1)$$

is a linear injection. The horizontal space is defined as the image of this linear injection, so that

$$T^h_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1) \simeq T_pM \ .$$

Additionally, it can be shown that the intersection of the horizontal and the vertical space only consists of the zero vector, which is geometrically obvious.

For the computation of the derivative of the map $E_m$ we need to use the specific model of $\mathrm{Gr}_{n,m}$ that we described in Section 5.3.2. Recall that we have identified the Grassmann manifold with the homogeneous space $\mathrm{Gr}_{n,m} \cong O(n)/(O(m) \times O(n-m))$. See Section 5.3 and Section 5.3.2 for a description of the tangent spaces of $\mathrm{Gr}_{n,m}$ and the Riemannian metric on them. The following lemma separately describes the images of vertical vectors and of horizontal vectors in $T_{(p,\mathcal{Y})} \mathrm{Gr}(M, m-1)$ (cf. Remark 6.2.4) under the derivative of $E_m$.

**Lemma 6.2.5.** *Let $M \subset S^{n-1}$ be a hypersurface with unit normal field $\nu$, and let $1 \leq m \leq n-1$. Furthermore, let $(p, \mathcal{Y}) \in \mathrm{Gr}(M, m-1)$, and let $\zeta_1, \ldots, \zeta_{n-2}$ be an orthonormal basis of $T_pM$ such that $\mathcal{Y} = \mathrm{lin}\{\zeta_1, \ldots, \zeta_{m-1}\}$. Then*

$$Q := \begin{pmatrix} p & \zeta_1 & \cdots & \zeta_{n-2} & \nu(p) \end{pmatrix} \in O(n) \ , \tag{6.16}$$

*and $\mathcal{W} := E_m(p, \mathcal{Y}) = [Q]$, where $E_m$ is defined as in (6.14). The derivative of the map $E_m$ at $(p, \mathcal{Y})$ is given in the following way:*

1. *(vertical) Let $w\colon \mathbb{R} \to \mathrm{Gr}_{n-2,m-1}$ be a curve through $w(0) = \mathbb{R}^{m-1} \times \{0\}$. Furthermore, let $c_v\colon \mathbb{R} \to \mathrm{Gr}(M, m-1)$ be defined by*

$$c_v(t) := \big(p\,,\ \{x_1\,\zeta_1 + \ldots + x_{n-2}\,\zeta_{n-2} \mid (x_1, \ldots, x_{n-2}) \in w(t)\}\big) \ , \tag{6.17}$$

*so that* $c_v(0) = (p, \mathcal{Y})$. *Then*

$$D_{(p,\mathcal{Y})}E_m(\dot{c_v}(0)) = \left[Q\,,\,\begin{pmatrix} 0 & -R_v^T \\ R_v & 0 \end{pmatrix}\right]\,, \tag{6.18}$$

*where*

$$R_v = \left(\begin{array}{c|c} \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & \bar{X} \\ \hline 0 & 0 \cdots 0 \end{array}\right) \in \mathbb{R}^{(n-m)\times m}\,,$$

*with* $\bar{X} \in \mathbb{R}^{(n-m-1)\times(m-1)}$ *given by* $\dot{w}(0) = \left[I_{n-2}, \begin{pmatrix} 0 & -\bar{X}^T \\ \bar{X} & 0 \end{pmatrix}\right]$.

2. *(horizontal) Let* $c\colon \mathbb{R} \to M$ *with* $c(0) = p$ *and* $\zeta := \dot{c}(0) \in T_pM$, *and let* $\mathcal{Y}_t$ *denote the parallel transport of* $\mathcal{Y}$ *along* $c$ *at time* $t$ *(cf. Theorem 4.1.1, resp. Remark 4.1.2). Then*

$$c_h\colon \mathbb{R} \to \mathrm{Gr}(M, m-1)\,, \quad c_h(t) := (c(t), \mathcal{Y}_t) \tag{6.19}$$

*satisfies* $c_h(0) = (p, \mathcal{Y})$. *The image of the tangent vector* $\dot{c_h}(0)$ *under the derivative of* $E_m$ *is given by*

$$D_{(p,\mathcal{Y})}E_m(\dot{c_h}(0)) = \left[Q\,,\,\begin{pmatrix} 0 & -R_h^T \\ R_h & 0 \end{pmatrix}\right]\,, \tag{6.20}$$

*where*

$$R_h = \left(\begin{array}{c|c} \begin{matrix} a_m \\ \vdots \\ a_{n-2} \end{matrix} & 0 \\ \hline 0 & b_1 \cdots b_{m-1} \end{array}\right) \in \mathbb{R}^{(n-m)\times m}\,,$$

*and the coefficients* $a_1, \ldots, a_{n-2}$, $b_1, \ldots, b_{n-2}$ *being given by*

$$\zeta = \sum_{i=1}^{n-2} a_i \cdot \zeta_i \quad \text{and} \quad W_p(\zeta) = \sum_{i=1}^{n-2} b_i \cdot \zeta_i\,,$$

*with* $W_p$ *denoting the Weingarten map of* $M$ *at* $p$ *(cf. Section 4.1.1).*

*Proof.* As for the first part, let the curve $w$ in $\mathrm{Gr}_{n-2,m-1}$ be represented by the curve $\bar{Q}\colon \mathbb{R} \to O(n-2)$ through $\bar{Q}(0) = I_{n-2}$, i.e., $w(t) = [\bar{Q}(t)]$. It is easily seen that the image of the curve $c_v$ under the map $E_m$ is given by

$$E_m \circ c_v(t) = [Q_v(t)]\,, \quad Q_v(t) := Q \cdot \begin{pmatrix} 1 & & \\ & \bar{Q}(t) & \\ & & 1 \end{pmatrix}\,. \tag{6.21}$$

Note that $Q_v(0) = Q$, as $\bar{Q}(0) = I_{n-2}$. Therefore, the derivative of $E_m \circ c_v(t)$ in 0 is given as stated in (6.18).

For the second part, let $v_i\colon \mathbb{R} \to \mathbb{R}^n$ be the parallel transport of $\zeta_i$ along $c$, $i = 1, \ldots, n-2$, and let

$$Q_h(t) := \begin{pmatrix} c(t) & v_1(t) & \cdots & v_{n-2}(t) & \nu(c(t)) \end{pmatrix} \in O(n)\,. \tag{6.22}$$

The fact that $Q_h(t)$ is an orthogonal matrix follows from the fact that the vectors $v_1(t), \ldots, v_{n-2}(t)$ form an orthonormal basis of $T_{c(t)}M$ (cf. Theorem 4.1.1), and $T_qM \subset T_qS^{n-1} = q^\perp$ for $q \in M$. It follows that $Q_h(0) = Q$, and we have

$$E_m \circ c_h(t) = [Q_h(t)]\,.$$

From Lemma 4.1.8 we know that the derivative of $Q_h$ at 0 is given by

$$\dot{Q}_h(0) = Q \cdot \begin{pmatrix} 0 & -a_1 & \cdots & -a_{n-2} & 0 \\ a_1 & 0 & \cdots & 0 & -b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-2} & 0 & \cdots & 0 & -b_{n-2} \\ 0 & b_1 & \cdots & b_{n-2} & 0 \end{pmatrix} . \tag{6.23}$$

From this we get that the derivative of $[Q_h(t)]$ at 0 is given by (6.20). $\square$

In the setting of Lemma 6.2.5 we collect all tangent vectors in $T_{\mathcal{W}} \operatorname{Gr}_{n,m}$ of the form (6.18) in the set $T^v_{\mathcal{W}}$, i.e.,

$$T^v_{\mathcal{W}} := \{D_{(p,\mathcal{Y})}E_m(\dot{c_v}(0)) \mid c_v \colon \mathbb{R} \to \operatorname{Gr}(M, m-1) \text{ given as in (6.17)}\} . \tag{6.24}$$

Furthermore, we collect all tangent vectors of the form (6.20) in the set $T^h_{\mathcal{W}}$, i.e.,

$$T^h_{\mathcal{W}} := \{D_{(p,\mathcal{Y})}E_m(\dot{c_h}(0)) \mid c_h \colon \mathbb{R} \to \operatorname{Gr}(M, m-1) \text{ given as in (6.19)}\} . \tag{6.25}$$

In other words, the sets $T^v_{\mathcal{W}}$ and $T^h_{\mathcal{W}}$ are the images of the vertical and the horizontal space (cf. Remark 6.2.4) of $\operatorname{Gr}(M, m-1)$ under the derivative of $E_m$, respectively. The following lemma gives a more detailed description of these sets.

**Corollary 6.2.6.** *Let the setting be as in Lemma 6.2.5. The sets $T^v_{\mathcal{W}}$ and $T^h_{\mathcal{W}}$ as defined in* (6.24) *and* (6.25) *are linear subspaces of $T_{\mathcal{W}} \operatorname{Gr}_{n,m}$ of dimensions*

$$\dim(T^v_{\mathcal{W}}) = (m-1)(n-m-1) , \quad \dim(T^h_{\mathcal{W}}) = n-m-1+\operatorname{rk}(W_{p,\mathcal{Y}}) ,$$

*where $W_{p,\mathcal{Y}}$ is defined by*

$$W_{p,\mathcal{Y}} \colon \mathcal{Y} \to \mathcal{Y} , \quad W_{p,\mathcal{Y}}(\zeta) := \Pi_{\mathcal{Y}} \circ W_p(\zeta) , \tag{6.26}$$

*$\Pi_{\mathcal{Y}}$ denoting the orthogonal projection onto $\mathcal{Y}$. A basis for $T^v_{\mathcal{W}}$ is given by*

$$\xi^v_{ij} := \left[ Q , \left( \begin{array}{cc|c|c} \multicolumn{2}{c|}{\multirow{2}{*}{0}} & 0 & 0 \\ \cline{3-4} & & -\bar{E}^T_{ij} & 0 \\ \hline 0 & \multicolumn{1}{|c|}{\bar{E}_{ij}} & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ \cline{1-2} 0 & \multicolumn{1}{|c|}{0} & & \end{array} \right) \right] , \quad 1 \leq i \leq n-m-1 , \ 1 \leq j \leq m-1 ,$$

*where $\bar{E}_{ij} \in \mathbb{R}^{(n-m-1)\times(m-1)}$ denotes the $(i,j)$th elementary matrix. The space $T^h_{\mathcal{W}}$ is spanned by the vectors*

$$\xi^h_k := \left[ Q , \left( \begin{array}{cc|c|c} \multicolumn{2}{c|}{\multirow{2}{*}{0}} & -\hat{a}^T_k & 0 \\ \cline{3-4} & & 0 & -\bar{b}_k \\ \hline \hat{a}_k & \multicolumn{1}{|c|}{0} & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ \cline{1-2} 0 & \multicolumn{1}{|c|}{\bar{b}^T_k} & & \end{array} \right) \right] , \quad 1 \leq k \leq n-2 ,$$

*where*

$$\hat{a}_k = \begin{cases} (0 \,, \ \ldots \ , \ 0)^T & \textit{if } 1 \leq k \leq m-1 \\ (0, \ldots, 0, \underset{\substack{\uparrow \\ k-m+1}}{1}, 0, \ldots, 0)^T & \textit{if } m \leq k \leq n-2 , \end{cases}$$

$$\bar{b}_k = \left( \ \langle W_p(\zeta_k), \zeta_1 \rangle \ , \ \ldots \ , \ \langle W_p(\zeta_k), \zeta_{m-1} \rangle \ \right)^T .$$

*If $\operatorname{rk}(W_{p,\mathcal{Y}}) = m-1$, then $\xi^h_1, \ldots, \xi^h_{n-2}$ form a basis of $T^h_{\mathcal{W}}$.*

*Proof.* From (6.18) in Lemma 6.2.5 it follows that $T^v_{\mathcal{W}}$ is a linear space of dimension $(m-1)(n-m-1)$, and an orthonormal basis of $T^v_{\mathcal{W}}$ is given by the vectors $\xi^v_{ij}$. As for the set $T^h_{\mathcal{W}}$, let $\xi \in T^h_{\mathcal{W}}$ be given as in (6.20). The coefficient vector $\hat{a}$ is then the representation of the projection of $\zeta$ onto $\mathcal{Y}^\perp$ w.r.t. the basis $\zeta_m, \ldots, \zeta_{n-2}$ of $\mathcal{Y}^\perp$. Furthermore, the coefficient vector $\bar{b}$ is the representation of $W_{p,\mathcal{Y}}(\zeta)$ w.r.t. the basis $\zeta_1, \ldots, \zeta_{m-1}$ of $\mathcal{Y}$. It follows that $T^h_{\mathcal{W}}$ is a linear subspace of the tangent space $T_{\mathcal{W}} \operatorname{Gr}_{n,m}$. As for the dimension, note that writing the vectors $\hat{a}_1, \ldots, \hat{a}_{n-2}$ and $\bar{b}_1, \ldots, \bar{b}_{n-2}$ in a matrix yields

$$\begin{pmatrix} \bar{b}_1 & \cdots & \bar{b}_{n-2} \\ \hat{a}_1 & \cdots & \hat{a}_{n-2} \end{pmatrix} = \begin{pmatrix} \bar{B} & * \\ 0 & I_{n-m-1} \end{pmatrix} , \tag{6.27}$$

where $\bar{B} \in \mathbb{R}^{(m-1)\times(m-1)}$ denotes the transformation matrix of $W_{p,\mathcal{Y}}$ with respect to the basis $\zeta_1, \ldots, \zeta_{m-1}$ of $\mathcal{Y}$. Therefore, the dimension of $T^h_{\mathcal{W}}$ is given by the rank of the matrix in (6.27), and we have

$$\dim(T^h_{\mathcal{W}}) = n - m - 1 + \operatorname{rk}(W_{p,\mathcal{Y}}) . \qquad \square$$

We may now give the proof of Theorem 6.2.1.

*Proof of Theorem 6.2.1.* In Lemma 6.2.2 we have seen that $E_m$ is bijective, and in Lemma 6.2.3 we have seen that $E_m$ is smooth. For the proof that $E_m$ is an embedding, it remains to show that the derivative of $E_m$ has everywhere full rank.

The sets $T^v_{\mathcal{W}}$ and $T^h_{\mathcal{W}}$ defined in (6.24) and (6.25) lie by definition in the image of the derivative of $E_m$ at $(p, \mathcal{Y})$. In Corollary 6.2.6 we have seen that $T^v_{\mathcal{W}}$ and $T^h_{\mathcal{W}}$ are linear subspaces of $T_{\mathcal{W}} \operatorname{Gr}_{n,m}$, which lie orthogonal to each other. Furthermore, we have seen that the dimension of $T^h_{\mathcal{W}}$ depends on the rank of the restriction of the Weingarten map to $\mathcal{Y}$, as defined in (6.26). Recall that the Weingarten map $W_p$ is positive definite, as $M = \partial K$ and $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ (cf. Remark 4.1.5 and Definition 4.1.9). This implies that also $W_{p,\mathcal{Y}}$ is positive definite, and therefore $\operatorname{rk}(W_{p,\mathcal{Y}}) = m - 1$. It follows that

$$\begin{aligned} \operatorname{rk}(D_{(p,\mathcal{Y})} E_m) &\geq \dim(T^v_{\mathcal{W}}) + \dim(T^h_{\mathcal{W}}) = m(n-m) - 1 \\ &= \dim \operatorname{Gr}(M, m-1) . \end{aligned}$$

So $E_m$ is an embedding of $\operatorname{Gr}(M, m-1)$ in $\operatorname{Gr}_{n,m}$, and its image is given by $\Sigma_m$. In particular, $\Sigma_m$ is a hypersurface of $\operatorname{Gr}_{n,m}$, as the dimension of $\operatorname{Gr}(M, m-1)$ is given by $\dim \operatorname{Gr}(M, m-1) = \dim \operatorname{Gr}_{n,m} - 1$. Moreover, the tangent space of $\Sigma_m$ at $\mathcal{W}$ decomposes into $T_{\mathcal{W}} \Sigma_m = T^v_{\mathcal{W}} \oplus T^h_{\mathcal{W}}$, and we can define a unit normal field $\nu_\Sigma$ on $\Sigma_m$ by setting

$$\nu_\Sigma(\mathcal{W}) := \left[ Q \,, \left( \begin{array}{c|c} 0 & \begin{array}{c|c} 0 & -1 \\ \hline 0 & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & 0 \\ \hline 1 & 0 \end{array} & 0 \end{array} \right) \right] \in T_{\mathcal{W}} \operatorname{Gr}_{n,m} , \tag{6.28}$$

where $Q \in O(n)$ is defined as in (6.16). This finishes the proof. $\square$

**Remark 6.2.7.** In the above proof we have shown that the rank of the derivative of $E_m$ is at least as big as $\dim(T^v_{\mathcal{W}}) + \dim(T^h_{\mathcal{W}})$. In fact, as the tangent space of $\operatorname{Gr}(M, m-1)$ decomposes into the vertical and the horizontal space (cf. Remark 6.2.4), and as $T^v_{\mathcal{W}}$ and $T^h_{\mathcal{W}}$ are the images of these spaces under the derivative of $E_m$, we have

$$\operatorname{rk}(D_{(p,\mathcal{Y})} E_m) = \dim(T^v_{\mathcal{W}}) + \dim(T^h_{\mathcal{W}}) .$$

Therefore, $(p, \mathcal{Y}) \in \mathrm{Gr}(M, m-1)$ is a critical point of the map $E_m$, iff the restricted Weingarten map $W_{p,\mathcal{Y}}$ is rank-deficient.

**Remark 6.2.8.** Recall that $\Sigma_m = \mathcal{F}_{\mathrm{G}}^{\mathrm{D}} \cap \mathcal{F}_{\mathrm{G}}^{\mathrm{P}}$, so $\mathrm{Gr}_{n,m} \setminus \Sigma_m$ decomposes into two disjoint components. The normal field $\nu_\Sigma$ as defined (6.28) points into the component $\mathcal{F}_{\mathrm{G}}^{\mathrm{D}}$. This is seen in the following way. Let the notation be as in Theorem 6.2.1. Defining $w\colon \mathbb{R} \to \mathrm{Gr}_{n,m}$, $w(\rho) := [Q \cdot Q_\rho]$, with

$$Q_\rho := \begin{pmatrix} \cos(\rho) & & -\sin(\rho) \\ & I_{n-2} & \\ \sin(\rho) & & \cos(\rho) \end{pmatrix} ,$$

we get $w(0) = \mathcal{W}$, and $\dot{w}(0) = \nu_\Sigma(\mathcal{W})$. The point $p_\rho := \cos(\rho)\, p + \sin(\rho)\, \nu(p)$ lies in $w(\rho)$, and for small enough $\rho > 0$ we have $p_\rho \in \mathrm{int}(K)$, as the unit normal field $\nu$ of $M$ is chosen such that $\nu(p)$ points inside the cap $K$. Therefore, for small enough $\rho > 0$ we have $w(\rho) \in \mathcal{F}_{\mathrm{G}}^{\mathrm{D}}$.

From now on we assume that $M = \partial K$ with $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$. Furthermore, we denote the sets $T_{\mathcal{W}}^v$ and $T_{\mathcal{W}}^h$ from (6.24) and (6.25) by $T_{\mathcal{W}}^v \Sigma_m$ and $T_{\mathcal{W}}^h \Sigma_m$, respectively. These subspaces are called the *vertical* and the *horizontal space* of $\Sigma_m$ at $\mathcal{W}$. Note that the tangent space of $\Sigma_m$ at $\mathcal{W}$ decomposes orthogonally into the vertical and the horizontal space, i.e.,

$$T_{\mathcal{W}} \Sigma_m \;=\; T_{\mathcal{W}}^v \Sigma_m \;\oplus\; T_{\mathcal{W}}^h \Sigma_m \;.$$

Note also that a basis of $T_{\mathcal{W}} \Sigma_m$ is thus given by

$$\xi_{ij}^v \;,\; i = 1, \ldots, n-m-1, j = 1, \ldots, m-1 \;; \quad \xi_k^h \;,\; k = 1, \ldots, n-2 \;,$$

defined as in Corollary 6.2.6.

We finish this section with a view on the canonical projection map

$$\Pi_M\colon \Sigma_m \to M \;, \quad \mathcal{W} \mapsto p \;, \text{ where } \mathcal{W} \cap K = \{p\} \;. \tag{6.29}$$

The following lemma is about the fiber $\Pi_M^{-1}(p)$, $p \in M$. Note that if we have chosen an orthonormal basis of $T_pM$, we may identify the set $\mathrm{Gr}(T_pM, m-1)$ with $\mathrm{Gr}_{n-2,m-1}$. In particular, $\mathrm{Gr}(T_pM, m-1)$ is endowed with a canonical Riemannian metric. This Riemannian metric is independent of the chosen orthonormal basis of $T_pM$. In fact, this also follows from the following lemma.

**Lemma 6.2.9.** *Let $\Pi_M\colon \Sigma_m \to M$ denote the canonical projection map as given in* (6.29)*, and let $p \in M$. Then the fiber $\Pi_M^{-1}(p)$ is a submanifold of $\Sigma_m$, which is isometric to* $\mathrm{Gr}(T_pM, m-1)$ *via the mutually inverse maps*

$$\begin{aligned} \Pi_M^{-1}(p) &\to \mathrm{Gr}(T_pM, m-1) \;, & \mathcal{W} &\mapsto \mathcal{W} \cap p^\perp \;, \\ \mathrm{Gr}(T_pM, m-1) &\to \Pi_M^{-1}(p) \;, & \mathcal{Y} &\mapsto \mathcal{Y} + \mathbb{R}\, p \;. \end{aligned}$$

*Additionally, the Normal Jacobian of the derivative of $\Pi_M$ at $\mathcal{W} \in \Pi_M^{-1}(p)$ is given by*

$$\mathrm{ndet}(D_{\mathcal{W}} \Pi_M) = \det(W_{p,\mathcal{Y}})^{-1} \;,$$

*where $W_{p,\mathcal{Y}}$ denotes the restriction of the Weingarten map of $M$ at $p$ to the subspace $\mathcal{Y} := \mathcal{W} \cap p^\perp$, as defined in* (6.26).

*Proof.* The projection $\Pi_M$ can be written in the form

$$\Pi_M = \Pi_1 \circ E_m^{-1} \;, \tag{6.30}$$

where $E_m$ is defined as in (6.14), and $\Pi_1\colon \operatorname{Gr}(M, m-1) \to M$ denotes the projection on the first component. Therefore, we have

$$\Pi_M^{-1}(p) = E_m(\{p\} \times \operatorname{Gr}(T_pM, m-1)) \;.$$

In particular, $\Pi_M^{-1}(p)$ is a submanifold of $\Sigma_m$, which is diffeomorphic to $\{p\} \times \operatorname{Gr}(T_pM, m-1) \simeq \operatorname{Gr}(T_pM, m-1)$ via the above given maps. As for the claim that this is an isometry, note that (6.30) implies that the kernel of the derivative of $\Pi_M$ is given by the vertical space, i.e.,

$$T_{\mathcal{W}}\Pi_M^{-1}(p) = \ker(D_{\mathcal{W}}\Pi_M) = T_{\mathcal{W}}^v\Sigma_m \;.$$

An orthonormal basis of the vertical space is given by $\{\xi_{ij}^v \mid 1 \le i \le n-m-1, 1 \le j \le m-1\}$. It is easily seen that this orthonormal basis maps to an orthonormal basis in $\operatorname{Gr}(T_pM, m-1) \simeq \operatorname{Gr}_{n-2,m-1}$.

As for the claim about the Normal Jacobian, note that the orthogonal complement of the kernel of the derivative of $\Pi_M$ is given by the horizontal space. A basis of $T_{\mathcal{W}}^h\Sigma_m$ is given by $\xi_1^h, \ldots, \xi_{n-2}^h$, defined as in Corollary 6.2.6. It is easily seen that we have

$$D_{\mathcal{W}}\Pi_M(\xi_k^h) = \zeta_k \in T_pM \;, \quad k = 1, \ldots, n-2 \;.$$

As $\zeta_1, \ldots, \zeta_{n-2}$ describe an orthonormal basis of $T_pM$, the Normal Jacobian of $D_{\mathcal{W}}\Pi_M$ is given by the inverse of the volume of the parallelepiped spanned by the vectors $\xi_1^h, \ldots, \xi_{n-2}^h$ in $T_{\mathcal{W}}^h\Sigma_m$. This volume is given by (cf. (6.27)) $|\det(W_{p,\mathcal{Y}})| = \det(W_{p,\mathcal{Y}})$, as $W_p$ and thus also $W_{p,\mathcal{Y}}$ is positive definite. □

## 6.3 Computing the tube

In this section we will compute the Normal Jacobian of the canonical surjection $\Upsilon$ of the tube around $\Sigma_m$, given by

$$\Upsilon\colon \Sigma_m \times \mathbb{R} \to \operatorname{Gr}_{n,m} \;, \quad (\mathcal{W}, \rho) \mapsto \overline{\exp}_{\mathcal{W}}(\rho \cdot \nu(\mathcal{W})) \;, \tag{6.31}$$

where $\overline{\exp}$ denotes the exponential map in $\operatorname{Gr}_{n,m}$, and $\nu_\Sigma\colon \Sigma_m \to T^\perp\Sigma_m$ denotes the unit normal field as defined in (6.28). Geometrically, the image $\Upsilon(\mathcal{W}, \rho)$ is obtained by rotating the intersection point $p \in M \cap \mathcal{W}$, i.e., $\mathcal{W} \cap K = \{p\}$, for an angle $\rho$ in normal direction away from the cap $K$, and keeping the subspace $p^\perp \cap \mathcal{W}$ fixed (cf. Section 5.5).

**Proposition 6.3.1.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and $M := \partial K$, and let $\Sigma_m := \Sigma_m(K)$. For $\mathcal{W} \in \Sigma_m$ and $\mathcal{W} \cap K = \{p\}$, let $\zeta_1, \ldots, \zeta_{n-2}$ be an orthonormal basis of $T_pM$ such that the intersection $\mathcal{Y} := \mathcal{W} \cap p^\perp$ is given by $\mathcal{Y} = \operatorname{lin}\{\zeta_1, \ldots, \zeta_{m-1}\}$. Then the Normal Jacobian of $D_{(\mathcal{W},\rho)}\Upsilon$, where $\Upsilon$ is defined as in* (6.31)*, is given by*

$$\operatorname{ndet}(D_{(\mathcal{W},\rho)}\Upsilon) = \frac{\left|\det\left(\begin{pmatrix}\cos(\rho)\cdot\bar{I} & 0\\ 0 & -\sin(\rho)\cdot\hat{I}\end{pmatrix}\cdot\Lambda + \begin{pmatrix}\sin(\rho)\cdot\bar{I} & 0\\ 0 & \cos(\rho)\cdot\hat{I}\end{pmatrix}\right)\right|}{\det(W_{p,\mathcal{Y}})} \;,$$

*where $\bar{I}, \hat{I}$ denote the $(m-1)$th and $(n-m-1)$th identity matrix, respectively, $W_{p,\mathcal{Y}}$ denotes the restricted Weingarten map (cf. (6.26)), and where $\Lambda$ denotes the representation matrix of the Weingarten map $W_p$ of $M$ at $p$, i.e.,*

$$\Lambda = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{n-2,1} \\ \vdots & & \vdots \\ \lambda_{1,n-2} & \cdots & \lambda_{n-2,n-2} \end{pmatrix} , \quad \lambda_{k\ell} := \langle W_p(\zeta_k), \zeta_\ell \rangle \; .$$

*Proof.* Let $Q \in O(n)$ be given as in (6.16), so that $\mathcal{W} = [Q]$. Then the map $\Upsilon$ is given by (cf. Example 5.4.6)

$$\Upsilon(\mathcal{W}, \rho) = [Q \cdot Q_\rho] \; , \quad Q_\rho := \begin{pmatrix} \cos(\rho) & & -\sin(\rho) \\ & I_{n-2} & \\ \sin(\rho) & & \cos(\rho) \end{pmatrix} \; .$$

Furthermore, we have seen in (5.29) in Example 5.4.6 (in a slightly different notation) that the derivative of $\Upsilon$ is in the second component given by

$$D_{(\mathcal{W},\rho)}\Upsilon(0,1) = \left[ Q \cdot Q_\rho \, , \left( \begin{array}{cc|c|c} \multicolumn{2}{c|}{\multirow{2}{*}{0}} & 0 & -1 \\ \cline{3-4} \multicolumn{2}{c|}{} & 0 & 0 \\ \hline 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ \cline{1-2} 1 & 0 & \multicolumn{2}{c}{} \end{array} \right) \right] \; . \tag{6.32}$$

Here, we use the same block decomposition as in Lemma 6.2.6.

The computation of the derivative of $\Upsilon$ on the first component is more complicated. In the following paragraph we describe an outline of the general approach.

Let $\xi \in T_{\mathcal{W}} \operatorname{Gr}_{n,m}$ be given by $\xi = \left[ Q, \left( \begin{smallmatrix} 0 & -R^T \\ R & 0 \end{smallmatrix} \right) \right]$, with $R \in \mathbb{R}^{(n-m)\times m}$. Furthermore, let $\tilde{Q}\colon \mathbb{R} \to O(n)$ be such that $\tilde{Q}(0) = Q$ and such that the induced curve in $\operatorname{Gr}_{n,m}$ satisfies $[\tilde{Q}(t)] \in \Sigma_m$ and $\frac{d}{dt}[\tilde{Q}(t)](0) = \xi$. Recall from Section 5.3.2 that in this case the matrix $\left( \begin{smallmatrix} 0 & -R^T \\ R & 0 \end{smallmatrix} \right) \in \overline{\operatorname{Skew}_n}$ is the orthogonal projection of the matrix $U \in \operatorname{Skew}_n$, which is defined by $\frac{d}{dt}\tilde{Q}(0) = QU \in T_Q O(n)$, onto the horizontal space $\overline{\operatorname{Skew}_n}$. We will choose the curve $\tilde{Q}$ such that the map $\Upsilon$ is given by

$$\Upsilon([\tilde{Q}(t)], \rho) = [\tilde{Q}(t) \cdot Q_\rho] \; . \tag{6.33}$$

In this case we have (cf. Example 5.4.6)

$$D_{(\mathcal{W},\rho)}\Upsilon(\xi, 0) = \left[ Q \cdot Q_\rho \, , \begin{pmatrix} 0 & -R_\xi^T \\ R_\xi & 0 \end{pmatrix} \right] \; ,$$

where $Q \cdot Q_\rho \cdot \left( \begin{smallmatrix} 0 & -R_\xi^T \\ R_\xi & 0 \end{smallmatrix} \right) \in T_{Q\cdot Q_\rho} O(n)$ is the horizontal component of

$$\tfrac{d}{dt}(\tilde{Q}(t) \cdot Q_\rho)(0) = QU \cdot Q_\rho = (Q \cdot Q_\rho) \cdot (Q_\rho^T \cdot U \cdot Q_\rho) \in T_{Q\cdot Q_\rho} O(n) \; .$$

So in order to compute the derivative of $\Upsilon$ we need to

- choose a basis of $T_{\mathcal{W}} \operatorname{Gr}_{n,m}$,
- find corresponding curves $\tilde{Q}$ in $O(n)$ whose images in $\operatorname{Gr}_{n,m}$ lie in $\Sigma_m$, and which satisfy (6.33),
- compute the (horizontal component of) the conjugation $Q_\rho^T \cdot U \cdot Q_\rho$ of the corresponding skew-symmetric matrix $U$.

Of course, the basis we choose will be the basis that we described in Corollary 6.2.6.

For $\xi^v \in T^v_{\mathcal{W}}\Sigma_m$ given by

$$\xi^v = \left[ Q\,, \left(\begin{array}{c|c} 0 & \begin{array}{c|c} 0 & 0 \\ \hline -\bar{X}^T & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & \bar{X} \\ \hline 0 & 0 \end{array} & 0 \end{array}\right) \right]$$

with $\bar{X} \in \mathbb{R}^{(n-m-1)\times(m-1)}$ as in (6.18) in Lemma 6.2.5, a defining curve in $O(n)$ is given by (cf. (6.21))

$$Q_v(t) = Q \cdot \begin{pmatrix} 1 & & \\ & \bar{Q}(t) & \\ & & 1 \end{pmatrix} ,$$

where $\bar{Q} \colon \mathbb{R} \to O(n-2)$ with $\bar{Q}(0) = I_{n-2}$ is such that $\frac{d}{dt}\bar{Q}(0) = \left(\begin{smallmatrix} 0 & -\bar{X}^T \\ \bar{X} & 0 \end{smallmatrix}\right)$. In this case we have

$$\tfrac{d}{dt}Q_v(0) = Q \cdot \left(\begin{array}{c|c} 0 & \begin{array}{c|c} 0 & 0 \\ \hline -\bar{X}^T & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & \bar{X} \\ \hline 0 & 0 \end{array} & 0 \end{array}\right) .$$

It is easily seen that for $\mathcal{W}^v_t := [Q_v(t)]$ we have $\mathcal{W}^v_t \cap K = \{p\}$ for all $t$. Moreover, the first column of $Q_v(t)$ is given by $p$, and the last column of $Q_v(t)$ is given by $\nu(p)$. This implies that we indeed have

$$\Upsilon([Q_v(t)], \rho) = [Q_v(t) \cdot Q_\rho] .$$

Therefore, as

$$Q_\rho^T \cdot \left(\begin{array}{c|c} 0 & \begin{array}{c|c} 0 & 0 \\ \hline -\bar{X}^T & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & \bar{X} \\ \hline 0 & 0 \end{array} & 0 \end{array}\right) \cdot Q_\rho = \left(\begin{array}{c|c} 0 & \begin{array}{c|c} 0 & 0 \\ \hline -\bar{X}^T & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & \bar{X} \\ \hline 0 & 0 \end{array} & 0 \end{array}\right) ,$$

we have

$$D_{(\mathcal{W},\rho)}\Upsilon(\xi^v, 0) = \left[ Q \cdot Q_\rho\,, \left(\begin{array}{c|c} 0 & \begin{array}{c|c} 0 & 0 \\ \hline -\bar{X}^T & 0 \end{array} \\ \hline \begin{array}{c|c} 0 & \bar{X} \\ \hline 0 & 0 \end{array} & 0 \end{array}\right) \right] . \tag{6.34}$$

For $\xi^h \in T^h_{\mathcal{W}}\Sigma_m$ given by

$$\xi^h = \left[ Q\,, \left(\begin{array}{c|c} 0 & \begin{array}{c|c} -\hat{a}^T & 0 \\ \hline 0 & -\bar{b} \end{array} \\ \hline \begin{array}{c|c} \hat{a} & 0 \\ \hline 0 & \bar{b}^T \end{array} & 0 \end{array}\right) \right]$$

with $\hat{a} \in \mathbb{R}^{n-m-1}$ and $\bar{b} \in \mathbb{R}^{m-1}$ as in (6.20) in Lemma 6.2.5, a defining curve in $O(n)$ is given by $t \mapsto Q_h(t)$, defined as in (6.22). Note that (cf. (6.23))

$$\tfrac{d}{dt}Q_h(0) = Q \cdot \left(\begin{array}{c|c|c} 0 & -a^T & 0 \\ \hline a & 0 & -b \\ \hline 0 & b^T & 0 \end{array}\right) ,$$

where $a = (a_1, \dots, a_{n-2})^T$ and $b = (b_1, \dots, b_{n-2})^T$. If we define $\mathcal{W}_t^h := [Q_h(t)]$, then from the definition of $Q_h(t)$ in (6.22), the intersection with the cap $K$ is given by $\mathcal{W}_t^h \cap K = \{(\text{first column of } Q_h(t))\}$, and the corresponding normal direction is given by the last column of $Q_h(t)$. This implies that again, we have

$$\Upsilon([\tilde{Q}_h(t)], \rho) = [\tilde{Q}_h(t) \cdot Q_\rho] \ .$$

Furthermore, we have

$$Q_\rho^T \cdot \left(\begin{array}{c|c|c} 0 & -a^T & 0 \\ \hline a & 0 & -b \\ \hline 0 & b^T & 0 \end{array}\right) \cdot Q_\rho = \left(\begin{array}{c|c|c} 0 & -ca^T + sb^T & 0 \\ \hline ca - sb & 0 & -sa - cb \\ \hline 0 & sa^T + cb^T & 0 \end{array}\right) ,$$

where we use the abbreviations

$$sa := \sin(\rho) \cdot a \ , \quad sb := \sin(\rho) \cdot b \ , \quad ca := \cos(\rho) \cdot a \ , \quad cb := \cos(\rho) \cdot b \ .$$

This finally yields

$$D_{(\mathcal{W},\rho)}\Upsilon(\xi^v, 0) = \left[ Q \cdot Q_\rho, \left( \begin{array}{cc|cc} \multicolumn{2}{c|}{\multirow{2}{*}{0}} & -c\hat{a}^T + s\hat{b}^T & 0 \\ \cline{3-4} & & 0 & -s\bar{a} - c\bar{b} \\ \hline c\hat{a} - s\hat{b} & 0 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ \cline{1-2} 0 & s\bar{a}^T + c\bar{b}^T & & \end{array} \right) \right] , \tag{6.35}$$

where we use the notation

$$\bar{a} := (a_1, \dots, a_{m-1}) \ , \quad \hat{a} := (a_m, \dots, a_{n-2}) \ , \quad s\bar{a} := \sin(\rho) \cdot \bar{a} \ , \quad c\hat{a} := \cos(\rho) \cdot \hat{a} \ ,$$
$$\bar{b} := (b_1, \dots, b_{m-1}) \ , \quad \hat{b} := (b_m, \dots, b_{n-2}) \ , \quad c\bar{b} := \cos(\rho) \cdot \bar{b} \ , \quad s\hat{b} := \sin(\rho) \cdot \hat{b} \ .$$

Recall that in Corollary 6.2.6 we have identified a basis of $T_\mathcal{W}\Sigma_m$ given by $\xi_{ij}^v$ and $\xi_k^h$, where $1 \le i \le n-m-1$, $1 \le j \le m-1$, and $1 \le k \le n-2$. Note that additionally to the vectors $\hat{a}_1, \dots, \hat{a}_{n-2} \in \mathbb{R}^{n-m-1}$ and $\bar{b}_1, \dots, \bar{b}_{n-2} \in \mathbb{R}^{m-1}$, which appear in the definition of the $\xi_k^h$, we may evidently define the vectors $\bar{a}_1, \dots, \bar{a}_{n-2} \in \mathbb{R}^{m-1}$ and $\hat{b}_1, \dots, \hat{b}_{n-2} \in \mathbb{R}^{n-m-1}$ so that

$$\begin{pmatrix} \bar{a}_1 & \cdots & \bar{a}_{n-2} \\ \hat{a}_1 & \cdots & \hat{a}_{n-2} \end{pmatrix} = I_{n-2} \ , \quad \begin{pmatrix} \bar{b}_1 & \cdots & \bar{b}_{n-2} \\ \hat{b}_1 & \cdots & \hat{b}_{n-2} \end{pmatrix} = \Lambda \ . \tag{6.36}$$

We thus have a basis of $T_\mathcal{W}\Sigma_m \times T_\rho\mathbb{R}$ given by

$$\begin{aligned} \{(0,1)\} \ &\cup \ \{(\xi_{ij}^v, 0) \mid 1 \le i \le n-m-1 \, , \ 1 \le j \le m-1\} \\ &\cup \ \{(\xi_k^h, 0) \mid 1 \le k \le n-2\} \ . \end{aligned}$$

Let us define

$$P(\mathcal{W}, \rho) := \left( \begin{array}{l} \text{parallelepiped in } T_\mathcal{W}\Sigma_m \times T_\rho\mathbb{R} \text{ spanned by the vectors} \\ (0,1) \, , \ (\xi_{ij}^v, 0) \, , \ (\xi_k^h, 0) \end{array} \right) ,$$

$$DP(\mathcal{W}, \rho) := \left( \begin{array}{l} \text{parallelepiped in } T_{\mathcal{W}_\rho} \mathrm{Gr}_{n,m} \text{ spanned by the vectors} \\ D_{(\mathcal{W},\rho)}\Upsilon(0,1) \, , \ D_{(\mathcal{W},\rho)}\Upsilon(\xi_{ij}^v, 0) \, , \ D_{(\mathcal{W},\rho)}\Upsilon(\xi_k^h, 0) \end{array} \right) .$$

Then we have

$$\operatorname{ndet}(D_{(\mathcal{W},\rho)}\Upsilon) = \frac{\operatorname{vol} DP(\mathcal{W},\rho)}{\operatorname{vol} P(\mathcal{W},\rho)} \ . \tag{6.37}$$

As for the denominator, the tangent vectors $(0,1)$ and $(\xi_{ij}^v, 0)$ are orthonormal, and they are orthogonal to the vectors $(\xi_k^h, 0)$. The vectors $(\xi_k^h, 0)$ span a parallelepiped of volume $\det(W_{p,\mathcal{Y}})$ (cf. proof of Lemma 6.2.9). This implies that

$$\operatorname{vol} P(\mathcal{W},\rho) = \det(W_{p,\mathcal{Y}}) \ .$$

As for the numerator in (6.37), we get from (6.32), (6.34), and (6.35)

$$\begin{aligned}
\operatorname{vol} DP(\mathcal{W},\rho) &= \left|\det\begin{pmatrix} s\bar{a}_1 + c\bar{b}_1 & \cdots & s\bar{a}_{n-2} + c\bar{b}_{n-2} \\ c\hat{a}_1 - s\hat{b}_1 & \cdots & c\hat{a}_{n-2} - s\hat{b}_{n-2}\end{pmatrix}\right| \\
&= \left|\det\left(\begin{pmatrix} c\bar{b}_1 & \cdots & c\bar{b}_{n-2} \\ -s\hat{b}_1 & \cdots & -s\hat{b}_{n-2}\end{pmatrix} + \begin{pmatrix} s\bar{a}_1 & \cdots & s\bar{a}_{n-2} \\ c\hat{a}_1 & \cdots & c\hat{a}_{n-2}\end{pmatrix}\right)\right| \\
&\overset{(6.36)}{=} \left|\det\left(\begin{pmatrix} \cos(\rho)\cdot\bar{I} & 0 \\ 0 & -\sin(\rho)\cdot\hat{I}\end{pmatrix}\cdot\Lambda + \begin{pmatrix} \sin(\rho)\cdot\bar{I} & 0 \\ 0 & \cos(\rho)\cdot\hat{I}\end{pmatrix}\right)\right| \ .
\end{aligned}$$

This finishes the proof of Proposition 6.3.1. □

Having computed the Normal Jacobian of the map $\Upsilon$ we have finished “half” of the proof of Theorem 6.1.1. The following computation shows what remains to be done. This computation only serves motivational purposes so that we may be somewhat generous in omitting details.

For small enough $\alpha > 0$ the volume of the primal tube around $\Sigma_m$ is given by (cf. Corollary 5.1.1)

$$\begin{aligned}
\operatorname{vol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) &= \int_{\Sigma_m}\int_{-\alpha}^{0} \operatorname{ndet}(D_{(\mathcal{W},\rho)}\Upsilon)\, d\rho\, d\mathcal{W} \\
&= \int_{\Sigma_m}\int_{-\alpha}^{0} \frac{\left|\det\left(\begin{pmatrix} \cos(\rho)\cdot\bar{I} & 0 \\ 0 & -\sin(\rho)\cdot\hat{I}\end{pmatrix}\cdot\Lambda + \begin{pmatrix} \sin(\rho)\cdot\bar{I} & 0 \\ 0 & \cos(\rho)\cdot\hat{I}\end{pmatrix}\right)\right|}{\det(W_{p,\mathcal{Y}})}\, d\rho\, d\mathcal{W},
\end{aligned}$$

where $\Lambda = \Lambda(\mathcal{W})$ as in Proposition 6.3.1. Using the projection map $\Pi_M\colon \Sigma_m \to M$, and using Lemma 6.2.9, we may continue as

$$= \int_{p\in M}\ \int_{\mathcal{Y}\in\mathrm{Gr}(T_pM,m-1)}\ \int_{-\alpha}^{0} \left|\det\left(\begin{pmatrix} \cos(\rho)\cdot\bar{I} & 0 \\ 0 & -\sin(\rho)\cdot\hat{I}\end{pmatrix}\cdot\Lambda + \begin{pmatrix} \sin(\rho)\cdot\bar{I} & 0 \\ 0 & \cos(\rho)\cdot\hat{I}\end{pmatrix}\right)\right| d\rho\, d\mathcal{Y}\, dp.$$

Substituting $t := -\tan\rho$ and $\tau := \tan\alpha$ yields (using $\sin(\arctan(t)) = t/\sqrt{1+t^2}$, $\cos(\arctan(t)) = 1/\sqrt{1+t^2}$, and $\frac{d}{dt}\arctan(t) = 1/(1+t^2)$)

$$= \int_0^{\tau}\int_{p\in M}\ \int_{\mathcal{Y}\in\mathrm{Gr}(T_pM,m-1)} \frac{\left|\det\left(\begin{pmatrix} \bar{I} & 0 \\ 0 & t\cdot\hat{I}\end{pmatrix}\cdot\Lambda + \begin{pmatrix} -t\cdot\bar{I} & 0 \\ 0 & \hat{I}\end{pmatrix}\right)\right|}{(1+t^2)^{n/2}}\, d\mathcal{Y}\, dp\, dt \ .$$

Using Fubini's Theorem, and rescaling the volume element on $\mathrm{Gr}(T_pM, m-1)$ so that we get a probability measure, we see that we need to compute an expectation of a certain generalized characteristic polynomial. We will call this the twisted characteristic polynomial (depending on $\mathcal{Y} \in \mathrm{Gr}(T_pM, m-1)$), and we will compute the expectation w.r.t. $\mathcal{Y} \in \mathrm{Gr}(T_pM, m-1)$ in the following section.

## 6.4 The expected twisted characteristic polynomial

In this section we will compute the expectation of the twisted characteristic polynomial (cf. Definition 6.4.2 below). For convenience we use in this section the subsitutions $k := n-2$ and $\ell := m-1$.

For $J \in \binom{[k]}{i} = \{i\text{-element subsets of } [k]\}$, $[k] = \{1, \ldots, k\}$, let us denote by $\mathrm{pm}_J(\Lambda)$ the $J$th principal minor of $\Lambda \in \mathbb{R}^{k\times k}$, i.e.

$$\mathrm{pm}_J(\Lambda) = \det(\Lambda_J) \;, \quad \Lambda_J = \begin{pmatrix} \lambda_{j_1,j_1} & \lambda_{j_1,j_2} & \cdots & \lambda_{j_1,j_i} \\ \lambda_{j_2,j_1} & \lambda_{j_2,j_2} & \cdots & \lambda_{j_2,j_i} \\ \vdots & \vdots & & \vdots \\ \lambda_{j_i,j_1} & \lambda_{j_i,j_2} & \cdots & \lambda_{j_i,j_i} \end{pmatrix} ,$$

if $J = \{j_1, \ldots, j_i\}$, $j_1 < j_2 < \ldots < j_i$. It is well-known that the usual characteristic polynomial can be written in terms of the principal minors. We will describe this in detail in the following lemma, as we will need a corresponding statement for the twisted characteristic polynomial.

**Lemma 6.4.1.** *The characteristic polynomial of $\Lambda \in \mathbb{R}^{k\times k}$ is given by*

$$\det(\Lambda - t\cdot I_k) = \sum_{i=0}^{k} (-1)^{k-i} \cdot \sigma_i(\Lambda) \cdot t^{k-i} ,$$

*where $\sigma_i(\Lambda)$ denotes the sum of all principal minors of $\Lambda$ of size $i$, i.e.,*

$$\sigma_i(\Lambda) = \sum_{J \in \binom{[k]}{i}} \mathrm{pm}_J(\Lambda) .$$

*Proof.* If $v_1, \ldots, v_k$ denote the columns of $\Lambda$, we can write the determinant as a function in the columns via

$$\det(\Lambda - t\cdot I_k) = \det(v_1 - t\,e_1, v_2 - t\,e_2, \ldots, v_k - t\,e_k) .$$

Using the multilinearity of det we get

$$\det(\Lambda - t\cdot I_k) = \sum_{J\subseteq [k]} \det(w_{J,1}, w_{J,2}, \ldots, w_{J,k}) , \tag{6.38}$$

where $w_{J,i} = v_i$, if $i \in J$, and $w_{J,i} = -t\,e_i$, if $i \notin J$. Expanding the determinant in the columns $w_{J,i} = -t\,e_i$, i.e., for $i \notin J$, yields

$$\det(w_{J,1}, w_{J,2}, \ldots, w_{J,k}) = (-t)^{k-|J|} \cdot \det(\Lambda_J) = (-t)^{k-|J|} \cdot \mathrm{pm}_J(\Lambda) .$$

Arranging the summands in (6.38) according to the size of $J$ finally yields

$$\det(\Lambda - t\cdot I_k) = \sum_{i=0}^{k} (-t)^{k-i} \cdot \sum_{J\in\binom{[k]}{i}} \mathrm{pm}_J(\Lambda) = \sum_{i=0}^{k} (-1)^{k-i} \cdot \sigma_i(\Lambda) \cdot t^{k-i} . \qquad \square$$

For diagonalizable $\Lambda$ the quantity $\sigma_i(\Lambda)$ is the evaluation of the $i$th elementary symmetric function in the eigenvalues of $\Lambda$. But for clarity we do not generally assume that $\Lambda$ is diagonalizable. Note that $\sigma_i(\Lambda) = \sigma_i(B^{-1} \cdot \Lambda \cdot B)$ for invertible $B$, as the characteristic polynomial is invariant under conjugation.

**Definition 6.4.2.** Let $\Lambda \in \mathbb{R}^{k\times k}$ and let $0 \leq \ell \leq k$. We define the $\ell$th *twisted characteristic polynomial* of $\Lambda$ as

$$\mathrm{ch}_\ell(\Lambda, t) := \det\left( \begin{pmatrix} I_\ell & 0 \\ 0 & t\, I_{k-\ell} \end{pmatrix} \cdot \Lambda + \begin{pmatrix} -t\, I_\ell & 0 \\ 0 & I_{k-\ell} \end{pmatrix} \right) .$$

Furthermore, we define the $\ell$th *positive twisted characteristic polynomial* of $\Lambda$ as

$$\mathrm{ch}_\ell^+(\Lambda, t) := \det\left( \begin{pmatrix} I_\ell & 0 \\ 0 & t\, I_{k-\ell} \end{pmatrix} \cdot \Lambda + \begin{pmatrix} t\, I_\ell & 0 \\ 0 & I_{k-\ell} \end{pmatrix} \right) .$$

Note that for $\ell = k$ we get that $\mathrm{ch}_k(\Lambda, t)$ is the usual characteristic polynomial, whereas for $\ell = 0$ we get $\mathrm{ch}_0(\Lambda, t) = \det(I_k + t\Lambda) = \sum_{i=0}^k \sigma_i(\Lambda) \cdot t^i$.

**Remark 6.4.3.** It should be noted that from a coordinate-free viewpoint there is an important difference between the usual characteristic polynomial and the twisted characteristic polynomial of a matrix $\Lambda$. The usual characteristic polynomial only depends on the linear map defined by $\Lambda$, whereas the twisted characteristic polynomial depends on the linear map and the subspace $\mathbb{R}^\ell \times \{0\}$. In fact, if one changes the bases of $\mathbb{R}^\ell \times \{0\}$ and of $\{0\} \times \mathbb{R}^{k-\ell}$, then the twisted characteristic polynomial stays invariant, i.e.,

$$\mathrm{ch}_\ell\left( \left( \begin{smallmatrix} B_1 & 0 \\ 0 & B_2 \end{smallmatrix} \right)^{-1} \cdot \Lambda \cdot \left( \begin{smallmatrix} B_1 & 0 \\ 0 & B_2 \end{smallmatrix} \right), t \right) = \mathrm{ch}_\ell\left(\Lambda, t\right) ,$$

for $B_1 \in \mathrm{Gl}_\ell$, $B_2 \in \mathrm{Gl}_{k-\ell}$. The same observation also applies for $\mathrm{ch}_\ell^+(\Lambda, t)$.

The following definition takes account of the above remarked invariance property of $\mathrm{ch}_\ell(\Lambda, t)$ and $\mathrm{ch}_\ell^+(\Lambda, t)$.

**Definition 6.4.4.** Let $\varphi\colon \mathbb{R}^k \to \mathbb{R}^k$ be a linear map, and let $\mathcal{Y} \in \mathrm{Gr}_{k,\ell}$. We define

$$\mathrm{ch}_\mathcal{Y}(\varphi, t) := \mathrm{ch}_\ell(\Lambda, t) \; , \quad \mathrm{ch}_\mathcal{Y}^+(\varphi, t) := \mathrm{ch}_\ell^+(\Lambda, t) \; ,$$

where $\Lambda \in \mathbb{R}^{k\times k}$ denotes the representation matrix of $\varphi$ with respect to a basis $b_1, \ldots, b_k$ of $\mathbb{R}^k$, which satisfies $\mathcal{Y} = \mathrm{lin}\{b_1, \ldots, b_\ell\}$ and $\mathcal{Y}^\perp = \mathrm{lin}\{b_{\ell+1}, \ldots, b_k\}$.

For the proof of Theorem 6.1.1 we will be interested in the expected value of $\mathrm{ch}_\mathcal{Y}(\varphi, t)$ and $\mathrm{ch}_\mathcal{Y}^+(\varphi, t)$ if $\mathcal{Y} \in \mathrm{Gr}_{k,\ell}$ is chosen uniformly at random. The following lemma shows that we may as well argue over the coordinate dependent twisted characteristic polynomial.

**Lemma 6.4.5.** *Let $\varphi\colon \mathbb{R}^k \to \mathbb{R}^k$ be a linear map and let $\Lambda$ be the representation matrix of $\varphi$ w.r.t. the canonical basis of $\mathbb{R}^k$. Then we have*

$$\begin{aligned} \mathop{\mathbb{E}}_{\mathcal{Y}}\left[\mathrm{ch}_\mathcal{Y}(\varphi, t)\right] &= \mathop{\mathbb{E}}_{Q}\left[\mathrm{ch}_\ell(Q^T \Lambda\, Q, t)\right] \; , \\ \mathop{\mathbb{E}}_{\mathcal{Y}}\left[\mathrm{ch}_\mathcal{Y}^+(\varphi, t)\right] &= \mathop{\mathbb{E}}_{Q}\left[\mathrm{ch}_\ell^+(Q^T \Lambda\, Q, t)\right] \; , \end{aligned}$$

*where $\mathcal{Y} \in \mathrm{Gr}_{k,\ell}$ and $Q \in O(k)$ are chosen uniformly at random.*

*Proof.* Let $Q \in O(k)$, and let $\mathcal{Y} := [Q] \in \mathrm{Gr}_{k,\ell}$. Note that $\mathcal{Y}$ is the linear subspace spanned by the first $\ell$ columns $b_1, \ldots, b_\ell$ of $Q$. The representation matrix of $\varphi$ w.r.t. the basis $b_1, \ldots, b_k$ of $\mathbb{R}^k$ is given by $Q^T \Lambda Q$. Therefore, we have

$$\mathrm{ch}_{\mathcal{Y}}(\varphi, t) = \mathrm{ch}_\ell(Q^T \Lambda Q, t) , \quad \mathrm{ch}_{\mathcal{Y}}^+(\varphi, t) = \mathrm{ch}_\ell^+(Q^T \Lambda Q, t) .$$

If $Q \in O(k)$ is chosen uniformly at random, then also $\mathcal{Y} = [Q] \in \mathrm{Gr}_{k,\ell}$ is chosen uniformly at random (cf. Section 5.3.2). This finishes the proof. □

**Proposition 6.4.6.** *Let $\Lambda \in \mathbb{R}^{k\times k}$ and $\ell \le k$. Then we have for $Q \in O(k)$ chosen uniformly at random*

$$\mathop{\mathbb{E}}_{Q}\left[\mathrm{ch}_\ell(Q^T \cdot \Lambda \cdot Q, t)\right] \;=\; \sum_{i,j=0}^{k} d_{ij} \cdot \sigma_{k-j}(\Lambda) \cdot t^{k-i} , \tag{6.39}$$

$$\mathop{\mathbb{E}}_{Q}\left[\mathrm{ch}_\ell^+(Q^T \cdot \Lambda \cdot Q, t)\right] \;=\; \sum_{i,j=0}^{k} |d_{ij}| \cdot \sigma_{k-j}(\Lambda) \cdot t^{k-i} , \tag{6.40}$$

*where the coefficients $d_{ij}$ are given for $i + j + \ell \equiv 0 \bmod 2$ and*

$$0 \le \tfrac{i-j}{2} + \tfrac{\ell}{2} \le \ell , \quad 0 \le \tfrac{i+j}{2} - \tfrac{\ell}{2} \le k - \ell ,$$

*by*

$$d_{ij} \;=\; (-1)^{\frac{i-j}{2}-\frac{\ell}{2}} \cdot \frac{\binom{\ell}{\frac{i-j}{2}+\frac{\ell}{2}} \cdot \binom{k-\ell}{\frac{i+j}{2}-\frac{\ell}{2}}}{\binom{k}{j}} , \tag{6.41}$$

*and $d_{ij} = 0$ else. Additionally, if $\Lambda$ is positive semidefinite, then for $t \ge 0$*

$$|\mathrm{ch}_\ell(\Lambda, t)| \le \mathrm{ch}_\ell^+(\Lambda, t) ,$$

*so that in this case*

$$\mathop{\mathbb{E}}_{Q}\Big[\big|\,\mathrm{ch}_\ell(Q^T \cdot \Lambda \cdot Q, t)\big|\Big] \;\le\; \sum_{i,j=0}^{k} |d_{ij}| \cdot \sigma_{k-j}(\Lambda) \cdot t^{k-i} . \tag{6.42}$$

**Remark 6.4.7.** The coefficients $d_{ij}$ defined in (6.41) coincide with the coefficients $d_{ij}^{nm}$ defined in (6.5), where $n := k + 2$, $m := \ell + 1$. See Table 6.1 for the values of $d_{ij}$ resp. $d_{ij}^{nm}$ for some concrete examples.

The proof of Proposition 6.4.6 is quite simple. The basic idea is to dissect the polynomials and to consider the expectations of the principal minors. For the expectation of the principal minors we exploit the invariance of the elementary symmetric functions. The only remaining difficulty is then the computation of the coefficients.

For $0 \le r \le k$ the $r$th leading principal minor is the principal minor for $J = [r]$. Let us denote this by

$$\mathrm{lpm}_r(\Lambda) := \mathrm{pm}_{[r]}(\Lambda) = \det(\Lambda_{[r]}) , \quad \Lambda_{[r]} = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix} .$$

By permuting the rows and columns of $\Lambda$ we can relate the principal minors with each other. More precisely, let $J = \{j_1, \ldots, j_r\}$, and let $\pi$ be any permutation of $[k]$

such that $\pi(i) = j_i$ for all $i = 1,\dots,r$. If $M_\pi$ denotes the permutation matrix according to $\pi$, i.e., $M_\pi \cdot e_i = e_{\pi(i)}$, then $\Lambda_J = (M_\pi^T \cdot \Lambda \cdot M_\pi)_{[r]}$, and therefore

$$\mathrm{pm}_J(\Lambda) = \mathrm{lpm}_r(M_\pi^T \cdot \Lambda \cdot M_\pi) .$$

The single principal minors may not be so interesting, as they depend strongly on the matrix $\Lambda$, in contrast to the $\sigma_r$, which only depend on the transformation, i.e., the conjugacy class of $\Lambda$. But we can change this by considering the averaged principal minors.

**Lemma 6.4.8.** *Let $\Lambda \in \mathbb{R}^{k\times k}$, and let $Q \in O(k)$ be chosen uniformly at random. Then for $J \in \binom{[k]}{r}$ we have*

$$\mathop{\mathbb{E}}_{Q}\big[\mathrm{pm}_J(Q^T \cdot \Lambda \cdot Q)\big] \;=\; \mathop{\mathbb{E}}_{Q}\big[\mathrm{lpm}_r(Q^T \cdot \Lambda \cdot Q)\big] \;=\; \binom{k}{r}^{-1} \sigma_r(\Lambda) . \tag{6.43}$$

*Proof.* For the first equality let $J = \{j_1,\dots,j_r\}$, let $\pi$ be any permutation of $[k]$ such that $\pi(i) = j_i$ for all $i = 1,\dots,r$, and let $M_\pi$ denote the permutation matrix according to $\pi$. We have seen that $\mathrm{pm}_J(\Lambda) = \mathrm{lpm}_r(M_\pi^T \cdot \Lambda \cdot M_\pi)$. This implies

$$\begin{aligned}\mathop{\mathbb{E}}_{Q}\big[\mathrm{pm}_J(Q^T \cdot \Lambda \cdot Q)\big] &= \mathop{\mathbb{E}}_{Q}\big[\mathrm{lpm}_r(M_\pi^T \cdot Q^T \cdot \Lambda \cdot Q \cdot M_\pi)\big] \\ &\overset{\tilde{Q}:=QM_\pi}{=} \mathop{\mathbb{E}}_{\tilde{Q}}\Big[\mathrm{lpm}_r(\tilde{Q}^T \cdot \Lambda \cdot \tilde{Q})\Big] ,\end{aligned}$$

where we have used the fact that right multiplication by the fixed element $M_\pi$ leaves the uniform distribution on $O(k)$ invariant. This also implies

$$\begin{aligned}\mathop{\mathbb{E}}_{Q}\big[\mathrm{lpm}_r(Q^T \cdot \Lambda \cdot Q)\big] &= \binom{k}{r}^{-1} \sum_{J\in\binom{[k]}{r}} \mathop{\mathbb{E}}_{Q}\big[\mathrm{pm}_J(Q^T \cdot \Lambda \cdot Q)\big] \\ &= \binom{k}{r}^{-1} \mathop{\mathbb{E}}_{Q}\Big[\sum_{J\in\binom{[k]}{r}} \mathrm{pm}_J(Q^T \cdot \Lambda \cdot Q)\Big] \\ &\overset{\text{Lem. 6.4.1}}{=} \binom{k}{r}^{-1} \mathop{\mathbb{E}}_{Q}\big[\sigma_r(Q^T \cdot \Lambda \cdot Q)\big] = \binom{k}{r}^{-1} \sigma_r(\Lambda) .\end{aligned}$$

□

Before we give the proof of Proposition 6.4.6 we may give a useful reformulation of Lemma 6.4.8 in the following corollary.

**Corollary 6.4.9.** *Let $\varphi\colon \mathbb{R}^k \to \mathbb{R}^k$ be a linear map. Furthermore, for $\mathcal{Y} \in \mathrm{Gr}_{k,\ell}$, let $\varphi_{\mathcal{Y}}$ denote the linear map*

$$\varphi_{\mathcal{Y}}\colon \mathcal{Y} \to \mathcal{Y} , \quad \varphi_{\mathcal{Y}}(x) := \Pi_{\mathcal{Y}} \circ \varphi(x) ,$$

*where $\Pi_{\mathcal{Y}}\colon \mathbb{R}^k \to \mathcal{Y}$ denotes the orthogonal projection onto $\mathcal{Y}$. Then, for uniformly random $\mathcal{Y} \in \mathrm{Gr}_{k,\ell}$, we have*

$$\mathop{\mathbb{E}}_{\mathcal{Y}}[\det(\varphi_{\mathcal{Y}})] = \binom{k}{\ell}^{-1} \sigma_\ell(\varphi) .$$

*Proof.* The determinant of $\varphi_{\mathcal{Y}}$ is given by the $\ell$th leading principal minor of the representation matrix of $\varphi$, if the basis $b_1,\dots,b_k$ of $\mathbb{R}^k$ is chosen such that $\mathcal{Y} = \mathrm{lin}\{b_1,\dots,b_\ell\}$ and $\mathcal{Y}^\perp = \mathrm{lin}\{b_{\ell+1},\dots,b_k\}$. The claim now immediately follows from Lemma 6.4.5 and Lemma 6.4.8. □

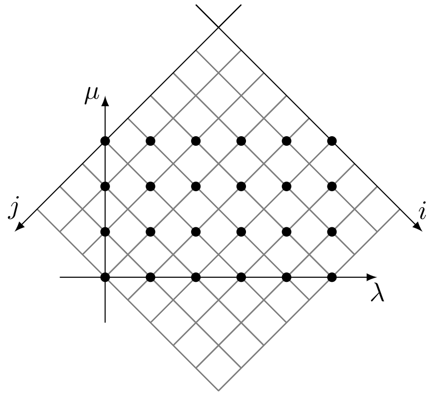



Figure 6.1: Illustration of the change of summation in (6.46) ($k = 8$, $\ell = 5$).

*Proof of Proposition 6.4.6.* Using the multilinearity of the determinant and arguing as we did for the usual characteristic polynomial in Lemma 6.4.1, we see that

$$\mathrm{ch}_\ell(\Lambda, t) = \sum_{J \subset [k]} (-1)^{c_1(J)} \cdot \mathrm{pm}_J(\Lambda) \cdot t^{c_2(J)} , \tag{6.44}$$

for some integers $c_1(J), c_2(J)$. This consideration also shows that for $\mathrm{ch}_\ell^+$ we get the same expansion except for the sign $(-1)^{c_1(J)}$, i.e.,

$$\mathrm{ch}_\ell^+(\Lambda, t) = \sum_{J \subset [k]} \mathrm{pm}_J(\Lambda) \cdot t^{c_2(J)} . \tag{6.45}$$

Averaging the twisted characteristic polynomial yields

$$\begin{aligned}
\mathop{\mathbb{E}}_Q \left[\mathrm{ch}_\ell(Q^T \cdot \Lambda \cdot Q, t)\right] &= \sum_{J \subset [k]} (-1)^{c_1(J)} \cdot \mathop{\mathbb{E}}_Q \left[\mathrm{pm}_J(Q^T \cdot \Lambda \cdot Q)\right] \cdot t^{c_2(J)} \\
&\overset{(6.43)}{=} \sum_{J \subset [k]} \frac{(-1)^{c_1(J)}}{\binom{k}{|J|}} \cdot \sigma_{|J|}(\Lambda) \cdot t^{c_2(J)} \\
&= \sum_{i,j=0}^{k} \tilde{d}_{ij} \cdot \sigma_{k-j}(\Lambda) \cdot t^{k-i} ,
\end{aligned}$$

for some constants $\tilde{d}_{ij}$. To compute these constants let us consider the matrices $\Lambda = s \cdot I_k$. For this choice of $\Lambda$ we have $\mathrm{ch}_\ell(s \cdot I_k, t) = (s - t)^\ell \cdot (1 + s \cdot t)^{k-\ell}$. As $\sigma_{k-j}(s \cdot I_k) = \binom{k}{j} \cdot s^{k-j}$ and $Q^T \cdot s \cdot I_k \cdot Q = s \cdot I_k$, we get

$$\begin{aligned}
(s - t)^\ell \cdot (1 + s \cdot t)^{k-\ell} = \mathrm{ch}_\ell(s \cdot I_k, t) &= \mathop{\mathbb{E}}_Q \left[\mathrm{ch}_\ell(Q^T \cdot s \cdot I_k \cdot Q, t)\right] \\
&= \sum_{i,j=0}^{k} \tilde{d}_{ij} \cdot \binom{k}{j} \cdot s^{k-j} \cdot t^{k-i} .
\end{aligned}$$

Let us expand the first term so that we can make a comparison of the coefficients to get the $\tilde{d}_{ij}$. We have

$$
\begin{aligned}
(s-t)^\ell \cdot (1+s{\cdot}t)^{k-\ell} &= \left(\sum_{\lambda=0}^{\ell}\binom{\ell}{\lambda}(-1)^{\ell-\lambda}\cdot s^\lambda\cdot t^{\ell-\lambda}\right)\cdot\left(\sum_{\mu=0}^{k-\ell}\binom{k-\ell}{\mu}s^\mu\cdot t^\mu\right)\\
&= \sum_{\lambda=0}^{\ell}\sum_{\mu=0}^{k-\ell}(-1)^{\ell-\lambda}\binom{\ell}{\lambda}\binom{k-\ell}{\mu}\cdot s^{\lambda+\mu}\cdot t^{\ell-\lambda+\mu} \qquad (6.46)\\
&\overset{\substack{i=k-\ell+\lambda-\mu\\ j=k-\lambda-\mu}}{=} \sum_{\substack{i,j=0\\ i+j+\ell\equiv 0\\ (\mathrm{mod}\ 2)}}^{k}(-1)^{\frac{i-j}{2}-\frac{\ell}{2}}\binom{\ell}{\frac{i-j}{2}+\frac{\ell}{2}}\binom{k-\ell}{k-\frac{i+j+\ell}{2}}\cdot s^{k-j}\cdot t^{k-i},
\end{aligned}
$$

where we interpret $\binom{n}{m}=0$ if $m<0$ or $m>n$, i.e., the above summation over $i,j$ in fact only runs over the rectangle determined by the inequalities $0\le\frac{i-j}{2}+\frac{\ell}{2}\le\ell$ and $0\le k-\frac{i+j+\ell}{2}\le k-\ell$. See Figure 6.1 for an illustration of the change of summation. Note that the reverse substitution is given by $\lambda=\frac{i-j+\ell}{2}$ and $\mu=k-\frac{i+j+\ell}{2}$.

Comparing the coefficients of the two expressions of $(s-t)^\ell\cdot(1+s\cdot t)^{k-\ell}$ reveals that indeed $\tilde{d}_{ij}=d_{ij}$ as defined in (6.41). This shows the equality in (6.39).

The equality in (6.40) is shown analogously with the observation

$$\mathrm{ch}_\ell^+(s\cdot I_k,t)=(s+t)^\ell\cdot(1+s\cdot t)^{k-\ell}\ .$$

As for the additional claim, note that for positive semidefinite $\Lambda$ every principal minor is nonnegative, i.e., $\mathrm{pm}_J(\Lambda)\ge 0$ for all $J\subset[k]$. Therefore, if $t\ge 0$, we get from (6.44) and (6.45)

$$
\begin{aligned}
|\mathrm{ch}_\ell(\Lambda,t)| &\overset{(6.44)}{=} \Big|\sum_{J\subset[k]}(-1)^{c_1(J)}\cdot\mathrm{pm}_J(\Lambda)\cdot t^{c_2(J)}\Big|\\
&\le \sum_{J\subset[k]}\mathrm{pm}_J(\Lambda)\cdot t^{c_2(J)} \overset{(6.45)}{=} \mathrm{ch}_\ell^+(\Lambda,t)\ .
\end{aligned}
$$

This finishes the proof of Proposition 6.4.6. □

## 6.5 Proof of Theorem 6.1.1

Before we give the proof of Theorem 6.1.1 we compute the volume of the set $\Sigma_m(K)$ for $K\in\mathcal{K}^{\mathrm{sm}}(S^{n-1})$. Recall that in Section 6.2 we have shown that $\Sigma_m(K)$ is a hypersurface of $\mathrm{Gr}_{n,m}$. In particular, it has a well-defined volume given by the integral of the constant 1-function over $\Sigma_m(K)$.

**Lemma 6.5.1.** *For $K\in\mathcal{K}^{\mathrm{sm}}(S^{n-1})$ we have*

$$\mathrm{vol}\,\Sigma_m(K)=\frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}\cdot\mathcal{O}_{m-1}\cdot\mathcal{O}_{n-m-1}}{\binom{n-2}{m-1}}\cdot V_{n-m-1}(K)\ .$$

*Proof.* Consider the projection map $\Pi_M\colon\Sigma_m\to M$. From Corollary 5.1.1 and from Lemma 6.2.9, we get

$$
\begin{aligned}
\mathrm{vol}\,\Sigma_m &= \int_{\Sigma_m}1\,d\mathcal{W} \overset{(5.5)}{=} \int_{p\in M}\int_{\mathcal{W}\in\Pi_M^{-1}(p)}\mathrm{ndet}(D_{\mathcal{W}}\Pi_M)^{-1}\,d\mathcal{W}\,dp\\
&\overset{\text{Lem. 6.2.9}}{=} \int_{p\in M}\int_{\mathcal{Y}\in\mathrm{Gr}(T_pM,m-1)}\det(W_{p,\mathcal{Y}})\,d\mathcal{W}\,dp\ , \qquad (6.47)
\end{aligned}
$$

where $W_{p,\mathcal{Y}}$ denotes the linear map

$$W_{p,\mathcal{Y}}\colon \mathcal{Y} \to \mathcal{Y}\ ,\quad W_{p,\mathcal{Y}}(\zeta) = \Pi_{\mathcal{Y}} \circ W_p(\zeta)\ ,$$

$\Pi_{\mathcal{Y}}$ denoting the orthogonal projection onto $\mathcal{Y}$. Normalizing the volume element on $\mathrm{Gr}(T_pM, m-1)$ and using Corollary 6.4.9, we get

$$\begin{aligned}(6.47) &= \int_{p\in M} \mathrm{vol}\,\mathrm{Gr}_{n-2,m-1} \cdot \mathop{\mathbb{E}}_{\mathcal{Y}}\left[\det(W_{p,\mathcal{Y}})\right]\, dp \\ &\overset{\text{Cor. 6.4.9}}{=} \frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}}{\binom{n-2}{m-1}} \cdot \int_{p\in M} \sigma_{m-1}(W_p)\, dp\ . \end{aligned} \tag{6.48}$$

Note that $\sigma_\ell(W_p)$ coincides with the values $\sigma_\ell(p)$, the $\ell$th elementary symmetric function in the principal curvatures of $M$ at $p$. So with Proposition 4.4.4 we may conclude

$$(6.48) = \frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1} \cdot \mathcal{O}_{m-1} \cdot \mathcal{O}_{n-m-1}}{\binom{n-2}{m-1}} \cdot V_{n-m-1}(K)\ . \qquad \square$$

We will now give the proof of Theorem 6.1.1. Recall that the intrinsic volumes $V_j(K)$ are continuous in $K$ (cf. Proposition 4.4.1), and the set of smooth caps $\mathcal{K}^{\mathrm{sm}}(S^{n-1})$ lies dense in $\mathcal{K}^c(S^{n-1})$ (cf. Proposition 4.1.10); in particular, it lies dense in $\mathcal{K}^r(S^{n-1})$. Furthermore, in Lemma 6.1.7 we have seen that the map

$$\mathcal{K}(S^{n-1}) \to \mathbb{R}\ ,\quad K \mapsto \mathrm{rvol}\,\mathcal{T}^{\mathrm{P}}(\Sigma_m(K), \alpha)$$

is (uniformly) continuous. Therefore, in order to show the inequality (6.3) we may assume w.l.o.g. that $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$. We will assume this for the rest of this section.

Recall from Section 6.2 that for $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, the set $\Sigma_m$ is an orientable hypersurface of $\mathrm{Gr}_{n,m}$, where the chosen unit normal field of $\Sigma_m$, denoted by $\nu_\Sigma$ (cf. (6.28)) points into the component $\mathcal{F}^{\mathrm{D}}_{\mathrm{G}}$ (cf. Remark 6.2.8). This implies that the image $\Upsilon(\Sigma_m \times [-\alpha, 0])$ (cf. (6.31)) covers the primal tube $\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha)$. Applying the coarea formula in Corollary 5.1.1, inequality (5.4), to the map $\Upsilon$ yields

$$\mathrm{vol}\,\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \le \int_{\Sigma_m} \int_{-\alpha}^{0} \mathrm{ndet}(D_{(\mathcal{W},\rho)}\Upsilon)\, d\rho\, d\mathcal{W}\ .$$

Using the notation of Proposition 6.3.1 (note that the matrix $\Lambda$ depends on the subspace $\mathcal{Y} = \mathcal{W} \cap p^\perp \in \mathrm{Gr}(T_pM, m-1)$) we may continue as

$$= \int_{\Sigma_m} \int_{-\alpha}^{0} \frac{\left|\det\left(\begin{pmatrix}\cos(\rho)\cdot \bar{I} & 0 \\ 0 & -\sin(\rho)\cdot \hat{I}\end{pmatrix} \cdot \Lambda + \begin{pmatrix}\sin(\rho)\cdot \bar{I} & 0 \\ 0 & \cos(\rho)\cdot \hat{I}\end{pmatrix}\right)\right|}{|\det(\bar{\Lambda})|}\, d\rho\, d\mathcal{W},$$

and with the help of the projection map $\Pi_M\colon \Sigma_m \to M$ and Lemma 6.2.9 we may continue

$$= \int_{p\in M} \int_{\mathrm{Gr}(T_pM, m-1)} \int_{-\alpha}^{0} \left|\det\left(\begin{pmatrix}\cos(\rho)\cdot \bar{I} & 0 \\ 0 & -\sin(\rho)\cdot \hat{I}\end{pmatrix} \cdot \Lambda + \begin{pmatrix}\sin(\rho)\cdot \bar{I} & 0 \\ 0 & \cos(\rho)\cdot \hat{I}\end{pmatrix}\right)\right|\, d\rho\, d\mathcal{Y}\, dp.$$

Subsituting $t := -\tan\rho$ and $\tau := \tan\alpha$ yields (using $\sin(\arctan(t)) = t/\sqrt{1+t^2}$, $\cos(\arctan(t)) = 1/\sqrt{1+t^2}$, and $\frac{d}{dt}\arctan(t) = 1/(1+t^2)$)

$$= \int\limits_{p\in M} \int\limits_{\mathrm{Gr}(T_pM,m-1)} \int_0^\tau (1+t^2)^{-n/2} \cdot |\mathrm{ch}_{\mathcal{Y}}(W_p,t)| \, dt \, d\mathcal{Y} \, dp \, .$$

Rescaling the volume element on $\mathrm{Gr}(T_pM, m-1)$ finally yields

$$= \int\limits_{p\in M} \int_0^\tau (1+t^2)^{-n/2} \cdot \mathrm{vol}\,\mathrm{Gr}_{n-2,m-1} \cdot \mathop{\mathbb{E}}_{\mathcal{Y}}\Big[\, |\mathrm{ch}_{\mathcal{Y}}(W_p,t)| \,\Big] \, dt \, dp \, ,$$

where the expectation is w.r.t. $\mathcal{Y} \in \mathrm{Gr}(T_pM, m-1)$ chosen uniformly at random.

From Lemma 6.4.5 and Proposition 6.4.6 we get

$$\mathop{\mathbb{E}}_{\mathcal{Y}}\Big[\, |\mathrm{ch}_{\mathcal{Y}}(W_p,t)| \,\Big] \leq \sum_{i,j=0}^{n-2} |d_{ij}^{nm}| \cdot \sigma_{n-2-j}(p) \cdot t^{n-2-i} \, . \tag{6.49}$$

So we may conclude

$$\begin{aligned} \mathrm{vol}\,\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) &\leq \int\limits_{p\in M} \int_0^\tau \frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}}{(1+t^2)^{n/2}} \cdot \sum_{i,j=0}^{n-2} \left|d_{ij}^{nm}\right| \cdot \sigma_{n-2-j}(p) \cdot t^{n-2-i} \, dt \, dp \\ &= \mathrm{vol}\,\mathrm{Gr}_{n-2,m-1} \cdot \sum_{i,j=0}^{n-2} \left|d_{ij}^{nm}\right| \cdot \int_0^\tau \frac{t^{n-2-i}}{(1+t^2)^{n/2}} \, dt \cdot \int\limits_{p\in M} \sigma_{n-2-j}(p) \, dp \, . \end{aligned}$$

Reversing the substitution $t = \tan\rho$ and $\tau = \tan\alpha$ yields

$$\int_0^\tau \frac{t^{n-2-i}}{(1+t^2)^{n/2}} \, dt = \int_0^\alpha \cos(\rho)^i \cdot \sin(\rho)^{n-2-i} \, d\rho = I_{n,i}(\alpha) \, .$$

In Proposition 4.4.4 we have seen that

$$\int\limits_{p\in M} \sigma_{n-2-j}(p) \, dp = \mathcal{O}_j \cdot \mathcal{O}_{n-2-j} \cdot V_j(K) \, .$$

From (5.23) in Section 5.3.2 we get

$$\frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}}{\mathrm{vol}\,\mathrm{Gr}_{n,m}} = \frac{\prod_{i=n-m-1}^{n-3} \mathcal{O}_i}{\prod_{i=0}^{m-2} \mathcal{O}_i} \cdot \frac{\prod_{i=0}^{m-1} \mathcal{O}_i}{\prod_{i=n-m}^{n-1} \mathcal{O}_i} = \frac{\mathcal{O}_{m-1} \cdot \mathcal{O}_{n-m-1}}{\mathcal{O}_{n-2} \cdot \mathcal{O}_{n-1}} \, . \tag{6.50}$$

So finally, using the identities in Proposition 4.1.20, we compute

$$\begin{aligned} \mathrm{rvol}\,\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) &\leq \frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}}{\mathrm{vol}\,\mathrm{Gr}_{n,m}} \cdot \sum_{i,j=0}^{n-2} \left|d_{ij}^{nm}\right| \cdot \int_0^\tau \frac{t^{n-2-i}}{(1+t^2)^{n/2}} \, dt \cdot \int\limits_{p\in M} \sigma_{n-2-j}(p) \, dp \\ &= \frac{\mathcal{O}_{m-1} \cdot \mathcal{O}_{n-m-1}}{\mathcal{O}_{n-2} \cdot \mathcal{O}_{n-1}} \cdot \sum_{i,j=0}^{n-2} \left|d_{ij}^{nm}\right| \cdot I_{n,i}(\alpha) \cdot \mathcal{O}_j \cdot \mathcal{O}_{n-2-j} \cdot V_j(K) \\ &= \frac{2m(n-m)}{n} \cdot \frac{\omega_m \cdot \omega_{n-m}}{\omega_n} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \frac{\mathcal{O}_j \cdot \mathcal{O}_{n-2-j}}{2 \cdot \mathcal{O}_{n-2}} \cdot \sum_{i=0}^{n-2} \left|d_{ij}^{nm}\right| \cdot I_{n,i}(\alpha) \\ &\overset{(4.22)}{=} \frac{2m(n-m)}{n} \cdot \binom{n/2}{m/2} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} |d_{ij}^{nm}| \cdot I_{n,i}(\alpha) \, . \end{aligned}$$

This finishes the proof of the inequality (6.3) in Theorem 6.1.1.

As for the exact formula in the case $0 \leq \alpha \leq \alpha_0$, let us first describe the most natural approach, although we will go an alternative way, which seems more comfortable in our situation. Recapitulating the above estimate of rvol $\mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha)$, we notice that there are two steps where we possibly lose exactness. The first step lies in the fact that the map $\Upsilon$ might not be injective on $\Sigma_m \times [-\alpha, 0]$. This is harmless, as we will show that $\Upsilon$ is indeed injective on $\Sigma_m \times (-\alpha_0, 0]$. The second critical step is the estimate in (6.49). If we knew that we may drop the absolute value, i.e., if we had

$$\mathrm{ch}_{\mathcal{Y}}(W_p, t) \geq 0 \ , \quad \text{for } t \leq \tan(\alpha_0) \ , \tag{6.51}$$

then we would get

$$\mathop{\mathbb{E}}_{\mathcal{Y}}\Big[\, |\mathrm{ch}_{\mathcal{Y}}(W_p, t)| \,\Big] = \mathop{\mathbb{E}}_{\mathcal{Y}}\big[\, \mathrm{ch}_{\mathcal{Y}}(W_p, t)\big] = \sum_{i,j=0}^{n-2} d_{ij}^{nm} \cdot \sigma_{n-2-j}(p) \cdot t^{n-2-i} \ .$$

So this would give the equality statement in Theorem 6.1.1. The problem in this approach is to show (6.51). More precisely, it is easily seen that $\mathrm{ch}_{\mathcal{Y}}(W_p, t) \geq 0$ for all $t \leq \varepsilon$, where $\varepsilon > 0$ is some constant only depending on $K$. It is not so easy to show that one may take $\varepsilon \geq \tan(\alpha_0)$.

Therefore, we show the equality statement by some alternative method. First of all, we will show that $\Upsilon$ is injective on $\Sigma_m \times (-\alpha_0, 0]$. Let $\mathcal{W}_\rho := \Upsilon(\mathcal{W}, \rho)$ with $\rho \in (-\alpha_0, 0)$. If $q \in S^{n-1}$ denotes the point, which is obtained by rotating $p$ by an angle of $\rho$ in normal direction away fram $K$, then we have $\mathcal{W}_\rho = \mathcal{W}(\to q)$ (cf. the comments after (6.31); and cf. Definition 5.5.1). We get that $\mathcal{W}_\rho \cap K = \emptyset$, and the (spherical) distance between $S_\rho := \mathcal{W}_\rho \cap S^{n-1}$ and $K$ is given by $d(\mathcal{W}_\rho, \Sigma_m) = \rho$ (cf. Proposition 5.5.2). By Proposition 3.1.19 the pair $(p, q)$ is the unique pair in $K \times S_\rho$, which has spherical distance $\rho$. Therefore, by (5.37) in Corollary 5.5.3, $\mathcal{W}$ is the unique element in $\Sigma_m$, which minimizes the distance to $\mathcal{W}_\rho$. This shows that $\Upsilon$ is injective on $\Sigma_m \times (-\alpha_0, 0]$. It also shows that

$$\Big\{ \tilde{\mathcal{W}} \in \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \mid d(\tilde{\mathcal{W}}, \Sigma_m(K)) = \rho \Big\} \;=\; \Sigma_m(K_\rho) \ , \tag{6.52}$$

where $K_\rho := \mathcal{T}(K, \rho) \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ (by definition of $\alpha_0$).

Considering the distance function

$$\mathrm{dist}\colon \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \to \mathbb{R} \ , \quad \mathrm{dist}(\mathcal{W}) := d(\mathcal{W}, \Sigma_m) \ ,$$

we have the following commutative diagram

$$\begin{tikzcd} \Sigma_m \times [-\alpha, 0] \arrow[r, "\Upsilon"] \arrow[dr, "|\Pi_2|"'] & \mathcal{T}^{\mathrm{P}}(\Sigma_m, \alpha) \arrow[d, "\mathrm{dist}"] \\ & \mathbb{R} \end{tikzcd} \ ,$$

where $|\Pi_2|(\mathcal{W}, \rho) := |\rho|$. The Normal Jacobian of the map dist is easily seen to be 1 (cf. Example 5.4.6). Moreover, the fibers of dist are given by

$$\mathrm{dist}^{-1}(\rho) \overset{(6.52)}{=} \Sigma_m(K_\rho) \ ,$$

where $K_\rho = \mathcal{T}(K,\rho) \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$. Using the coarea formula in Corollary 5.1.1, we get

$$\begin{aligned}\operatorname{vol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) &= \int_0^\alpha \operatorname{vol}(\Sigma_m(K_\rho))\, d\rho \\ &\overset{\text{Lem. 6.5.1}}{=} \int_0^\alpha \frac{\operatorname{vol}\mathrm{Gr}_{n-2,m-1}\cdot \mathcal{O}_{m-1}\cdot \mathcal{O}_{n-m-1}}{\binom{n-2}{m-1}} \cdot V_{n-m-1}(K_\rho)\, d\rho\ .\end{aligned}$$

Note that

$$\frac{\operatorname{vol}\mathrm{Gr}_{n-2,m-1}\cdot \mathcal{O}_{m-1}\cdot \mathcal{O}_{n-m-1}}{\operatorname{vol}\mathrm{Gr}_{n,m}\cdot\binom{n-2}{m-1}} \overset{(6.50)}{=} \frac{\mathcal{O}_{m-1}^2\cdot \mathcal{O}_{n-m-1}^2}{\binom{n-2}{m-1}\cdot \mathcal{O}_{n-2}\cdot \mathcal{O}_{n-1}}\ .$$

In Proposition A.2.1 in Section A.2 in the appendix we will show that

$$V_{n-m-1}(K_\rho) = \binom{(n-2)/2}{(m-1)/2}\cdot \sum_{i,j=0}^{n-2} d_{ij}^{nm}\cdot \begin{bmatrix} n-2\\ j\end{bmatrix}\cdot \cos(\rho)^i\cdot \sin(\rho)^{n-2-i}\cdot V_j(K)\ .$$

Using this, we get

$$\begin{aligned}\operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) &= \frac{\operatorname{vol}\mathrm{Gr}_{n-2,m-1}\cdot \mathcal{O}_{m-1}\cdot \mathcal{O}_{n-m-1}}{\operatorname{vol}\mathrm{Gr}_{n,m}\cdot\binom{n-2}{m-1}}\cdot \int_0^\alpha V_{n-m-1}(K_\rho)\, d\rho \\ &= \frac{\binom{(n-2)/2}{(m-1)/2}\cdot \mathcal{O}_{m-1}^2\cdot \mathcal{O}_{n-m-1}^2}{\binom{n-2}{m-1}\cdot \mathcal{O}_{n-2}\cdot \mathcal{O}_{n-1}}\cdot \sum_{i,j=0}^{n-2} d_{ij}^{nm}\cdot \begin{bmatrix} n-2\\ j\end{bmatrix} \\ &\qquad \cdot \int_0^\alpha \cos(\rho)^i\cdot \sin(\rho)^{n-2-i}\, d\rho\cdot V_j(K)\ .\end{aligned}$$

With Proposition 4.1.20 we finally compute

$$\begin{aligned}\frac{\binom{(n-2)/2}{(m-1)/2}\cdot \mathcal{O}_{m-1}^2\cdot \mathcal{O}_{n-m-1}^2}{\binom{n-2}{m-1}\cdot \mathcal{O}_{n-2}\cdot \mathcal{O}_{n-1}} &\overset{(4.21)}{=} \frac{\mathcal{O}_{m-1}^2\cdot \mathcal{O}_{n-m-1}^2}{\begin{bmatrix} n-2\\ m-1\end{bmatrix}\cdot \mathcal{O}_{n-2}\cdot \mathcal{O}_{n-1}} \overset{(4.22)}{=} \frac{2\cdot \mathcal{O}_{m-1}\cdot \mathcal{O}_{n-m-1}}{\mathcal{O}_{n-1}} \\ &= \frac{2m(n-m)}{n}\cdot \frac{\omega_m\cdot\omega_{n-m}}{\omega_n} \overset{(4.22)}{=} \frac{2m(n-m)}{n}\cdot \binom{n/2}{m/2}\ ,\end{aligned}$$

which finishes the proof. □

# Chapter 7

# Estimations

In this chapter we will perform several analyses of the Grassmann condition number. These will be average analyses, but we will also give a first attempt of a smoothed analysis.

We will start with a first-order average analysis, which means that instead of using the whole tube formula for tail estimates we will only use the leading term. This implies of course that the results have to be taken with a grain of salt due to this inaccuracy. But these first-order results are still interesting as they might reveal the actual behavior of the condition, which is harder to be demonstrated in the more complicated full tube formula. In particular, we will not only be able to give estimates which are independent of the cone involved, but also improvements for a large class of cone families, which suggest a general independence of the expected condition from the dimension of the cone if some weak conditions are satisfied.

Second, we will give estimates of the full tube formula, carrying the most important 1st order estimates over to the full setting. The results will be slightly worse than suggested by the results of the first-order approach.

Third, we will explain how one can obtain smoothed analyses of the Grassmann condition. We will perform a first-order smoothed analysis to illustrate this, but this will merely be a proof of concept as the results are not yet satisfactory.

Table 7.1 summarizes the results of the average analyses, i.e., the tail estimates of the Grassmann condition of a matrix $A \in \mathbb{R}^{n\times m}$, $m < n$, assuming that $A$ be a normal distributed random matrix, i.e., the entries of $A$ are i.i.d. standard normal. Here, the abbreviations LP, SOCP-1, SOCP stand for the following choices of the reference cone $C$:

$$\begin{aligned}
\text{(LP):}\quad C &= \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0\ \forall i = 1,\ldots,n\} \\
\text{(SOCP-1):}\quad C &= \mathcal{L}^n = \{x \in \mathbb{R}^n \mid x_n^2 \geq x_1^2 + \ldots + x_{n-1}^2\} \\
\text{(SOCP):}\quad C &= \mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k}\ , \quad n_1 + \ldots + n_k = n\ .
\end{aligned}$$

Table 7.2 shows the corresponding estimates of the expectation of the logarithm of the Grassmann condition.

The following proposition provides the link between the average analysis of the Grassmann condition and the geometric volume computation of Chapter 6. Note that $\mathscr{C}_{\mathrm{G}}(A) \geq 1$ (cf. Remark 2.3.5).

**Proposition 7.0.2.** *Let $C \subset \mathbb{R}^n$ be a regular cone. If $A \in \mathbb{R}^{n\times m}$, $m < n$, is a normal distributed random matrix, then for $t \geq 1$ and $\alpha := \arcsin(1/t)$*

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] = \operatorname{rvol} \mathcal{T}(\Sigma_m(C), \alpha)\ .$$

| | 1st order | full |
|---|---|---|
| –any cone– | $2 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t}$ | $6 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t}$, if $t > n^{\frac{3}{2}}$ |
| LP | $2.3 \cdot \sqrt{m} \cdot \frac{1}{t}$ | $29 \cdot \sqrt{m} \cdot \frac{1}{t}$, if $t > m \geq 8$ |
| SOCP-1 | $1.6 \cdot \sqrt{m} \cdot \frac{1}{t}$ | $20 \cdot \sqrt{m} \cdot \frac{1}{t}$, if $t > m \geq 8$ |
| SOCP | $4 \cdot m \cdot \frac{1}{t}$ | — |
| –any self-dual cone– (assuming Conjecture 4.4.17) | $4 \cdot m \cdot \frac{1}{t}$ | — |

Table 7.1: Estimates of the tail $\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t]$.

| | $\mathbb{E}[\ln \mathscr{C}_{\mathrm{G}}(A)] < \ldots$ |
|---|---|
| –any cone– | $1.5 \cdot \ln(n) + 1.8$, if $n \geq 4$ |
| LP | $\ln(m) + 3.4$, if $m \geq 8$ |
| SOCP-1 | $\ln(m) + 3$, if $m \geq 8$ |

Table 7.2: Estimates of the expectation of $\ln \mathscr{C}_{\mathrm{G}}(A)$.

*Proof.* If $\mathcal{W} := \mathrm{im}(A^T)$, then with probability 1, we have $\mathcal{W} \in \mathrm{Gr}_{n,m}$. Moreover, the induced distribution on $\mathrm{Gr}_{n,m}$ is the uniform probability distribution (cf. Proposition 5.3.5). Furthermore, by Proposition 2.3.8 we have $\mathscr{C}_{\mathrm{G}}(A) = 1/\sin d_*(\mathcal{W}, \Sigma_m(C))$, where $d_*$ may denote either the Hausdorff- or the geodesic distance on $\mathrm{Gr}_{n,m}$. Therefore, we have

$$\mathscr{C}_{\mathrm{G}}(A) > t \iff \mathcal{W} \in \mathcal{T}(\Sigma_m(C), \alpha) ,$$

with $\alpha := \arcsin(1/t)$. □

The following lemma provides an easy transfer of tail estimates to estimates of the expectation.

**Lemma 7.0.3.** *Let $X$ be a random variable taking values $\geq 1$, and for $t \geq 1$ let*

$$\mathrm{Prob}[X > t] < c \cdot m^{d_1} \cdot (n-m)^{d_2} \cdot \frac{1}{t} .$$

*Then the expectation of the logarithm of $X$ is bounded from above by*

$$\mathbb{E}[\ln X] < d_1 \cdot \ln(m) + d_2 \cdot \ln(n-m) + \ln(c) + 1 ,$$

*for some constants $c, d_1, d_2 > 0$.*

*Proof.* This is shown by the following computation (put $r := \ln c$)

$$\begin{aligned}
\mathbb{E}\,[\ln X] &= \int_0^\infty \mathrm{Prob}[\ln(X) > s]\,ds \\
&< \ln(m^{d_1}\cdot(n-m)^{d_2}) + r + \int_{\ln(m^{d_1}\cdot(n-m)^{d_2})+r}^\infty c\cdot m^{d_1}\cdot(n-m)^{d_2}\cdot\exp(-s)\,ds \\
&= d_1\cdot\ln(m) + d_2\cdot\ln(n-m) + r + c\cdot\exp(-r) \\
&= d_1\cdot\ln(m) + d_2\cdot\ln(n-m) + \ln(c) + 1\ .
\end{aligned}$$

□

## 7.1 Average analysis – 1st order

In this section we will give first-order average analyses of the Grassmann condition, i.e., we will approximate the volume of the tube around $\Sigma_m$ by the volume of the (lower-dimensional) Sigma set. We use the following asymptotic notation for $t \to \infty$

$$\begin{aligned}
&\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \lesssim\ f(n,m)\cdot\frac{1}{t} \\
&\qquad :\Longleftrightarrow\ \mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \le\ f(n,m)\cdot\frac{1}{t} + \frac{g(n,m)}{t^2} \quad \text{for all } t>0\ ,
\end{aligned}$$

for some function $g(n,m)$.

Note that by Proposition 7.0.2 and by Corollary 6.1.5 we have, using $K := C \cap S^{n-1}$,

$$\begin{aligned}
\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &\lesssim 8\cdot\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})}\cdot\frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n-m}{2})}\cdot V_{n-m-1}(K)\cdot\frac{1}{t} \qquad (7.1)\\
&< 4\cdot\sqrt{m(n-m)}\cdot V_{n-m-1}(K)\cdot\frac{1}{t}\ ,
\end{aligned}$$

as $\arcsin(1/t) \sim 1/t$.

**Theorem 7.1.1.** *Let $A \in \mathbb{R}^{n\times m}$, $m<n$, be a normal distributed random matrix, i.e., the entries of $A$ are i.i.d. standard normal.*

1. *Let $C \subset \mathbb{R}^n$ be* any regular cone*. Then*

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \lesssim\ 2\cdot\sqrt{m(n-m)}\cdot\frac{1}{t}\ . \qquad (7.2)$$

2. *(LP) Let $C = \mathbb{R}^n_+$ be the positive orthant. Then*

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \lesssim\ 2.3\cdot\sqrt{m}\cdot\frac{1}{t}\ . \qquad (7.3)$$

3. *(SOCP-1) Let $C = \mathcal{L}^n$ be the nth Lorentz cone. Then*

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \lesssim\ 1.6\cdot\sqrt{m}\cdot\frac{1}{t}\ . \qquad (7.4)$$

4. *(SOCP) Let $C = \mathcal{L}^{n_1}\times\ldots\times\mathcal{L}^{n_k}$ be any second-order cone. Then*

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \ \lesssim\ 4\cdot m\cdot\frac{1}{t}\ . \qquad (7.5)$$

*5. Let $C \subset \mathbb{R}^n$ be* any self-dual cone. *If Conjecture 4.4.17 is true, then*

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \;\lesssim\; 4 \cdot m \cdot \frac{1}{t} \;. \tag{7.6}$$

Occasionally, we will use estimates or computations that are easily verified by hand or with the help of a computer algebra system. If we do this during a longer computation, then we will mark this with the symbol at the corresponding step.

In the remainder of this section we will give the proof of Theorem 7.1.1. We will show the tail estimates (7.2)–(7.6) one at a time.

*Proof of Theorem 7.1.1(1).* If $C \subset \mathbb{R}^n$ is a regular cone, then $K := C \cap S^{n-1}$ is a cap, i.e., $K \in \mathcal{K}^c(S^{n-1})$. In particular, $V_j(K) \le \frac{1}{2}$ (cf. Proposition 4.4.10). From (7.1) we thus get

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim 4 \cdot \sqrt{m(n-m)} \cdot V_{n-m-1}(K) \cdot \frac{1}{t} \le 2 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t} \;.$$

This proves the tail estimate (7.2). □(1)

*Proof of Theorem 7.1.1(2).* Let $C = \mathbb{R}^n_+$ be the positive orthant. The intrinsic volumes of $K = C \cap S^{n-1}$ are given by (cf. Remark 4.4.15)

$$V_j(K) = \frac{\binom{n}{j+1}}{2^n} \;, \tag{7.7}$$

for $-1 \le j \le n-1$. From (7.1) we thus get

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim \underbrace{4 \cdot \sqrt{m(n-m)} \cdot \frac{\binom{n}{m}}{2^n}}_{=:f_m(n)} \cdot \frac{1}{t} \;.$$

Using Proposition 4.1.22 and Proposition 4.1.23 it is easily seen that for fixed $m$ the sequence $f_m(m+1), f_m(m+2), f_m(m+3), \ldots$ is log-concave. In particular, it has at most one maximum. This maximum lies at $n = 2m+1$, as

$$\frac{f_m(2m)}{f_m(2m+1)} = 2 \cdot \frac{\sqrt{m}}{\sqrt{m+1}} \cdot \frac{m+1}{2m+1} = \frac{\sqrt{m \cdot (m+1)}}{\frac{m+(m+1)}{2}} < 1 \;,$$

and

$$\frac{f_m(2m+2)}{f_m(2m+1)} = \frac{1}{2} \cdot \frac{\sqrt{m+2}}{\sqrt{m+1}} \cdot \frac{2m+2}{m+2} = \frac{m+1}{\sqrt{m+1} \cdot \sqrt{m+2}} < 1 \;.$$

Therefore, using

$$\sqrt{m(m+1)} \cdot \frac{\binom{2m+1}{m}}{2^{2m+1}} \;<\; \frac{\sqrt{m}}{\sqrt{\pi}} \;,$$

we get

$$\begin{aligned}\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &\lesssim 4 \cdot \sqrt{m(m+1)} \cdot \frac{\binom{2m+1}{m}}{2^{2m+1}} \cdot \frac{1}{t} < 4 \cdot \frac{\sqrt{m}}{\sqrt{\pi}} \cdot \frac{1}{t} \\ &< 2.3 \cdot \sqrt{m} \cdot \frac{1}{t} \;.\end{aligned}$$

This proves the tail estimate (7.3). □(2)

**Remark 7.1.2.** Note that if the ratio $\frac{m}{n} =: c$ is fixed, then by (7.1) and (7.7)

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim 4 \cdot \sqrt{c(1-c)} \cdot n \cdot \frac{\binom{n}{c \cdot n}}{2^n} \cdot \frac{1}{t} ,$$

which goes exponentially to 0 if $c \neq \frac{1}{2}$ (cf. Section A.1).

*Proof of Theorem 7.1.1(3).* Let $C = \mathcal{L}^n$ be the $n$th Lorentz cone. Recall that the intrinsic volumes of $K = C \cap S^{n-1}$ are given for $0 \leq j \leq n-2$ by (cf. Example 4.4.8)

$$V_j(K) = \frac{\binom{(n-2)/2}{j/2}}{2^{n/2}} . \tag{7.8}$$

From (7.1) we thus get

$$\begin{aligned} \operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \;&\lesssim\; 8 \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n-m}{2})} \cdot \frac{\binom{(n-2)/2}{(m-1)/2}}{2^{n/2}} \cdot \frac{1}{t} \\ &\overset{(4.18)}{=} \underbrace{4 \cdot \frac{m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2}}{2^{n/2}}}_{=: g_m(n)} \cdot \frac{1}{t} . \end{aligned} \tag{7.9}$$

Is is easily seen, that for fixed $m$ the sequence $(\frac{n-m}{n} \mid n = m+1, m+2, \ldots)$ is log-concave. Moreover, using Proposition 4.1.22 and Proposition 4.1.23 it is easily seen that for fixed $m$ the sequence $g_m(m+1), g_m(m+2), g_m(m+3), \ldots$ is log-concave. In particular, it has at most one maximum. This maximum lies at $n = 2m+1$, as

$$\frac{g_m(2m)}{g_m(2m+1)} = \frac{4 \cdot \frac{m^2}{2m} \cdot \frac{\binom{m}{m/2}}{2^m}}{4 \cdot \frac{m(m+1)}{2m+1} \cdot \frac{\binom{m+1/2}{m/2}}{2^{m+1/2}}} = \frac{1}{\sqrt{2}} \cdot \frac{2m+1}{m+1} \cdot \frac{\binom{m}{m/2}}{\binom{m+1/2}{m/2}} < 1 ,$$

and

$$\frac{g_m(2m+2)}{g_m(2m+1)} = \frac{4 \cdot \frac{m(m+2)}{2m+2} \cdot \frac{\binom{m+1}{m/2}}{2^{m+1}}}{4 \cdot \frac{m(m+1)}{2m+1} \cdot \frac{\binom{m+1/2}{m/2}}{2^{m+1/2}}} = \frac{1}{\sqrt{2}} \cdot \frac{m+2}{m+1} \cdot \frac{2m+1}{2m+2} \cdot \frac{\binom{m+1}{m/2}}{\binom{m+1/2}{m/2}} < 1 .$$

Therefore, using

$$\frac{m+1}{2m+1} \cdot \frac{\binom{m+1/2}{m/2}}{2^m} < \frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{m}} ,$$

we get

$$\begin{aligned} \operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim 4 \cdot \frac{m(m+1)}{2m+1} \cdot \frac{\binom{m+1/2}{m/2}}{2^{m+1/2}} \cdot \frac{1}{t} &< \frac{4}{\sqrt{2\pi}} \cdot \sqrt{m} \cdot \frac{1}{t} \\ &< 1.6 \cdot \sqrt{m} \cdot \frac{1}{t} . \end{aligned}$$

This proves the tail estimate (7.4). □(3)

**Remark 7.1.3.** Note as in Remark 7.1.2 that if the ratio $\frac{m}{n} =: c$ is fixed then by (7.1) and (7.8)

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim 4 \cdot c(1-c) \cdot \frac{\binom{n/2}{c \cdot n/2}}{2^{c \cdot n/2}} \cdot \frac{1}{t} .$$

which goes exponentially to 0 if $c \neq \frac{1}{2}$ (cf. Section A.1).

For the tail estimates (7.5) and (7.6) we need the following lemma, which provides upper bounds for the intrinsic volumes of self-dual caps, assuming that the sequence of intrinsic volumes is unimodal.

**Lemma 7.1.4.** *If $K \in \mathcal{K}^c(S^{n-1})$ is a self-dual cap, and if the sequence of intrinsic volumes $V_0(K), V_1(K), \ldots, V_{n-2}(K)$ is unimodal, then*

$$V_i(K) \leq \begin{cases} \frac{1}{n-1-2i} & \textit{for } 0 \leq i \leq \frac{n-2}{2} \\ \frac{1}{2i-n+3} & \textit{for } \frac{n-2}{2} \leq i \leq n-2\,. \end{cases} \tag{7.10}$$

*Proof.* Recall that by self-duality we have $V_i(K) = V_{n-2-i}(K)$, so that the unimodality of the sequence of intrinsic volumes implies that $V_{i-1}(K) \leq V_i(K)$, if $i \leq \frac{n-2}{2}$, and $V_i(K) \geq V_{i+1}(K)$, if $i \geq \frac{n-2}{2}$. Also note that as $\sum_{i=-1}^{n-1} V_i(K) = 1$, we have $V_{i_1}(K) + \ldots + V_{i_k}(K) \leq 1$ for all $-1 \leq i_1 < i_2 < \ldots < i_k \leq n-1$. Let us first treat the case where $n$ is even. We have

$$\begin{aligned} 1 &\geq V_{\frac{n-2}{2}}(K)\,, \\ 1 &\geq V_{\frac{n-2}{2}-1}(K) + V_{\frac{n-2}{2}}(K) + V_{\frac{n-2}{2}+1}(K) \geq 3 \cdot V_{\frac{n-2}{2}-1}(K) \\ &\vdots \\ 1 &\geq (1+2j) \cdot V_{\frac{n-2}{2}-j}(K)\,, \end{aligned}$$

or equivalently $V_{\frac{n-2}{2}-j}(K) = V_{\frac{n-2}{2}+j}(K) \leq \frac{1}{1+2j}$. Substituting $i = \frac{n-2}{2} - j$ or $i = \frac{n-2}{2} + j$, respectively, yields (7.10) for $n$ even.

For $n$ odd we have

$$\begin{aligned} 1 &\geq V_{\frac{n-2}{2}-\frac{1}{2}}(K) + V_{\frac{n-2}{2}+\frac{1}{2}}(K) = 2 \cdot V_{\frac{n-3}{2}}(K)\,, \\ 1 &\geq V_{\frac{n-2}{2}-\frac{3}{2}}(K) + V_{\frac{n-2}{2}-\frac{1}{2}}(K) + V_{\frac{n-2}{2}+\frac{1}{2}}(K) + V_{\frac{n-2}{2}+\frac{3}{2}}(K) \geq 4 \cdot V_{\frac{n-3}{2}-1}(K) \\ &\vdots \\ 1 &\geq (2+2j) \cdot V_{\frac{n-3}{2}-j}(K)\,, \end{aligned}$$

or equivalently $V_{\frac{n-3}{2}-j}(K) = V_{\frac{n-1}{2}+j}(K) \leq \frac{1}{2+2j}$. Substituting $i = \frac{n-3}{2} - j$ or $i = \frac{n-1}{2} + j$, respectively, yields (7.10) for $n$ odd. This finishes the proof. □

*Proof of Theorem 7.1.1(4),(5).* Let $C \subset \mathbb{R}^n$ be a self-dual cone such that the sequence of intrinsic volumes $V_0(K), V_1(K), \ldots, V_{n-2}(K)$, where $K = C \cap S^{n-1}$, is unimodal. This is the case for $C = \mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k}$ a second-order cone (cf. Corollary 4.4.14), or for any self-dual cone, if Conjecture 4.4.17 is true. From (7.1) and Lemma 7.1.4 we thus get

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \quad \lesssim \quad \underbrace{4 \cdot \sqrt{m(n-m)} \cdot \begin{Bmatrix} \frac{1}{2m-n+1} & \text{if } m \geq \frac{n}{2} \\ \frac{1}{n-2m+1} & \text{if } m \leq \frac{n}{2} \end{Bmatrix}}_{=:h_m(n)} \cdot \frac{1}{t}$$

For fixed $m$ the sequence $h_m(m+1), h_m(m+2), h_m(m+3), \ldots$ has its maximum in $n = 2m$, as for $n \leq 2m$

$$\frac{h_m(n-1)}{h_m(n)} \overset{[n \leq 2m]}{=} \sqrt{\frac{n-m-1}{n-m}} \cdot \frac{2m-n+1}{2m-n+2} < 1\,,$$

and for $n \geq 2m$

$$\frac{h_m(n+1)}{h_m(n)} \overset{[n\geq 2m]}{=} \sqrt{\frac{n+1-m}{n-m}} \cdot \frac{n-2m+1}{n-2m+2}$$
$$= \sqrt{\frac{(n-m)\cdot(n-2m+1)^2+(n-2m+1)^2}{(n-m)\cdot(n-2m+1)^2+(n-m)\cdot 2(n-2m+1)+(n-m)}}$$
$$< \sqrt{\frac{(n-m)\cdot(n-2m+1)^2+(n-m)(n-2m+1)}{(n-m)\cdot(n-2m+1)^2+(n-m)\cdot 2(n-2m+1)}}$$
$$\leq 1 \ .$$

Therefore, we may conclude

$$\mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] \lesssim 4 \cdot m \cdot \frac{1}{t} \ .$$

This proves the tail estimates (7.5) and (7.6) and therefore finishes the proof. □(4),(5)

## 7.2 Average analysis – full

In this section we will estimate the complete tube formula thus getting a full average analysis of the Grassmann condition.

**Theorem 7.2.1.** *Let $A \in \mathbb{R}^{n\times m}$, $m < n$, be a normal distributed random matrix, i.e., the entries of $A$ are i.i.d. standard normal.*

1. *Let $C \subset \mathbb{R}^n$ be* any regular cone. *Then*

$$\begin{aligned} \mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &< 6 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t} \ , && \text{if } t > n^{\frac{3}{2}} \ , \\ \mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] &< 1.5 \cdot \ln(n) + 1.8 \ , && \text{if } n \geq 4 \ . \end{aligned} \tag{7.11}$$

2. *(LP) Let $C = \mathbb{R}^n_+$ be the positive orthant, and let $t > m \geq 8$. Then*

$$\begin{aligned} \mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &< 29 \cdot \sqrt{m} \cdot \frac{1}{t} \ , \\ \mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] &< \ln(m) + 3.4 \ . \end{aligned} \tag{7.12}$$

3. *(SOCP-1) Let $C = \mathcal{L}^n$ be the nth Lorentz cone, and let $t > m \geq 8$. Then*

$$\begin{aligned} \mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &< 20 \cdot \sqrt{m} \cdot \frac{1}{t} \ , \\ \mathbb{E}\left[\ln \mathscr{C}_{\mathrm{G}}(A)\right] &< \ln(m) + 3 \ . \end{aligned} \tag{7.13}$$

The following lemma collects some estimates that we will use in the proof of Theorem 7.2.1.

**Lemma 7.2.2.** *Let $i, k, \ell, m, n \in \mathbb{N}$ with $n \geq 2$ and $1 \leq m \leq n-1$.*

1. *We have*

$$\frac{\Gamma(\frac{m+\ell+1}{2})}{\Gamma(\frac{m}{2})} \leq \sqrt{\frac{m}{2}} \cdot \left(\frac{m+\ell}{2}\right)^{\frac{\ell}{2}} \ . \tag{7.14}$$

*2. For* $0 \le k \le m-1$ *and* $0 \le i-k \le n-m-1$ *we have*

$$\left(\frac{m+i-2k}{n-m-i+2k}\right)^{\frac{i-2k}{2}} < n^{\frac{i}{2}} \;. \tag{7.15}$$

*3. For* $0 \le i \le n-2$ *we have*

$$\sum_{k=0}^{n-2} \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} = \binom{n-2}{i} \;. \tag{7.16}$$

*4. For* $0 \le \alpha \le \frac{\pi}{2}$, $t := \sin(\alpha)^{-1}$, *and* $n \ge 3$, *we have*

$$\sum_{i=0}^{n-2} \binom{n-2}{i} \cdot n^{\frac{i}{2}} \cdot I_{n,n-2-i}(\alpha) \;<\; \frac{3}{t} \,, \qquad \textit{if } t > n^{\frac{3}{2}} \,, \tag{7.17}$$

$$\sum_{i=0}^{n-2} \binom{n-2}{i} \cdot I_{n,n-2-i}(\alpha) \;<\; \exp\left(\frac{n}{m}\right) \cdot \frac{1}{t} \,, \qquad \textit{if } t > m \,. \tag{7.18}$$

*5. For* $x \ge 0$ *and* $0 \le y \le x$ *let the binomial coefficient be extended to* $\binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(y+1)\cdot\Gamma(x-y+1)}$. *Then*

$$\binom{x}{y} \le \binom{x}{x/2} < \sqrt{\frac{2}{\pi \cdot x}} \cdot 2^x \;. \tag{7.19}$$

*Proof.* (1) For the first estimate we distinguish the cases $\ell$ odd and $\ell$ even. Using $\Gamma(x+1) = x \cdot \Gamma(x)$, we get for $\ell$ odd

$$\frac{\Gamma(\frac{m+\ell+1}{2})}{\Gamma(\frac{m}{2})} = \prod_{a=0}^{\frac{\ell-1}{2}} \left(\frac{m}{2} + a\right) \le \frac{m}{2} \cdot \left(\frac{m+\ell-1}{2}\right)^{\frac{\ell-1}{2}} \le \sqrt{\frac{m}{2}} \cdot \left(\frac{m+\ell}{2}\right)^{\frac{\ell}{2}} \;.$$

Using additionally $\Gamma(x+\frac{1}{2}) < \sqrt{x} \cdot \Gamma(x)$ (cf. Section 4.1.4), we get for $\ell$ even

$$\frac{\Gamma(\frac{m+\ell+1}{2})}{\Gamma(\frac{m}{2})} = \prod_{a=0}^{\frac{\ell}{2}-1} \left(\frac{m+1}{2} + a\right) \cdot \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} < \left(\frac{m+\ell}{2}\right)^{\frac{\ell}{2}} \cdot \sqrt{\frac{m}{2}} \;.$$

(2) As for the second estimate, we distinguish the cases $i \ge 2k$ and $i \le 2k$. From $0 \le k \le m-1$ and $0 \le i-k \le n-m-1$ we get

$$\begin{aligned} 1 \;\le\; & \quad m-k+i-k & \le\; n-1 \,, \\ 1 \;\le\; & n-(m-k+i-k) & \le\; n-1 \,. \end{aligned}$$

For $i \ge 2k$ we thus get

$$\left(\frac{m+i-2k}{n-m-i+2k}\right)^{\frac{i-2k}{2}} \le (n-1)^{\frac{i-2k}{2}} < n^{\frac{i}{2}},$$

and for $i \le 2k$

$$\left(\frac{m+i-2k}{n-m-i+2k}\right)^{\frac{i-2k}{2}} = \left(\frac{n-m+2k-i}{m-2k+i}\right)^{\frac{2k-i}{2}} \le (n-1)^{\frac{2k-i}{2}} < n^{\frac{i}{2}} \;.$$

(3) The third point is the so-called Vandermonde's identity, which follows from the following polynomial identity

$$\begin{aligned}\sum_{i=0}^{n-2}\binom{n-2}{i}\cdot X^i = (1+X)^{n-2} &= (1+X)^{m-1}\cdot(1+X)^{n-m-1}\\ &= \left(\sum_{k=0}^{n-2}\binom{m-1}{k}\cdot X^k\right)\cdot\left(\sum_{\ell=0}^{n-2}\binom{n-m-1}{\ell}\cdot X^\ell\right)\\ &= \sum_{i=0}^{n-2}\left(\sum_{k=0}^{n-2}\binom{m-1}{k}\cdot\binom{n-m-1}{i-k}\right)\cdot X^i\,.\end{aligned}$$

(4) The $I$-functions have been estimated in [17, Lemma 2.2] in the following way. Let $\varepsilon := \sin(\alpha) = \frac{1}{t}$. For $i < n-2$

$$I_{n,n-2-i}(\alpha) = \int_0^\alpha \cos(\rho)^{n-2-i}\cdot\sin(\rho)^i\,d\rho \le \frac{\varepsilon^{i+1}}{i+1}\,, \tag{7.20}$$

and for $i = n-2$

$$\begin{aligned}I_{n,0}(\alpha) = \int_0^\alpha \sin(\rho)^{n-2}\,d\rho &\le \frac{\mathcal{O}_{n-1}\cdot\varepsilon^{n-1}}{2\mathcal{O}_{n-2}} = \frac{\sqrt{\pi}}{2}\cdot\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}\cdot\varepsilon^{n-1}\\ &< \sqrt{\frac{\pi}{2(n-2)}}\cdot\varepsilon^{n-1}\,.\end{aligned}$$

With these estimates we get

$$\begin{aligned}&\sum_{i=0}^{n-2}\binom{n-2}{i}\cdot n^{\frac{i}{2}}\cdot I_{n,n-2-i}(\alpha)\\ &\quad\le \sum_{i=0}^{n-2}\binom{n-2}{i}\cdot n^{\frac{i}{2}}\cdot\frac{\varepsilon^{i+1}}{i+1} + \left(\sqrt{\frac{\pi}{2(n-2)}} - \frac{1}{n-1}\right)\cdot\varepsilon^{n-1}\cdot n^{\frac{n-2}{2}}\\ &\quad< \varepsilon\cdot\left(\sum_{i=0}^{n-2}\binom{n-2}{i}\cdot n^{\frac{i}{2}}\cdot\varepsilon^i + \frac{1.4}{\sqrt{n}}\cdot\varepsilon^{n-2}\cdot n^{\frac{n-2}{2}}\right)\\ &\quad= \varepsilon\cdot\left(\left(1+\sqrt{n}\cdot\varepsilon\right)^{n-2} + 1.4\cdot\varepsilon^{n-2}\cdot n^{\frac{n-3}{2}}\right)\,,\end{aligned}$$

and for $\varepsilon < n^{-\frac{3}{2}}$ we may continue

$$\begin{aligned}&< \varepsilon\cdot\left(\underbrace{\left(1+\frac{1}{n}\right)^{n-2}}_{<\exp(1)} + 1.4\cdot n^{\frac{3}{2}-n}\right)\\ &< 3\cdot\varepsilon\,.\end{aligned}$$

Similarly we compute the last estimate

$$\sum_{i=0}^{n-2}\binom{n-2}{i}\cdot I_{n,n-2-i}(\alpha) \overset{\text{(as above)}}{<} \varepsilon\cdot\left((1+\varepsilon)^{n-2} + \frac{1.4}{\sqrt{n}}\cdot\varepsilon^{n-2}\right)\,, \tag{7.21}$$

and for $\varepsilon < \frac{1}{m}$ we finally get

$$
\begin{aligned}
(7.21) < \varepsilon \cdot \left( \left( 1 + \frac{1}{m} \right)^{n-2} + \frac{1.4}{\sqrt{n} \cdot m^{n-2}} \right) \\
< \; \varepsilon \cdot \exp\left( \frac{n}{m} \right) \; .
\end{aligned}
$$

(5) The extended binomial coefficient is obviously symmetric around $x/2$, i.e.,

$$\frac{x}{x/2 - c} = \frac{x}{x/2 + c} \; ,$$

for $0 \leq c \leq \frac{x}{2}$. Furthermore, for fixed $x \geq 0$ the functions $\Gamma(y+1), \Gamma(x-y+1)$ are log-convex (cf. Section 4.1.4). It follows that their product is also log-convex, and the inverse of their product is log-concave. Therefore, the function

$$[0, x] \to \mathbb{R} \; , \quad y \mapsto \frac{\Gamma(x+1)}{\Gamma(y+1) \cdot \Gamma(x-y+1)} = \binom{x}{y}$$

is log-concave. The symmetry around $x/2$ thus implies that also its maximum has to lie in $x/2$. So we have

$$\binom{x}{y} \leq \binom{x}{x/2} = \frac{\Gamma(x+1)}{\Gamma(\frac{x+2}{2})^2} \; .$$

By the duplication formula for the $\Gamma$-function (cf. (4.16)), we have

$$\frac{\Gamma(x+1)}{\Gamma(\frac{x+2}{2})^2} = \frac{2^x \cdot \Gamma(\frac{x+1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{x+2}{2})} < \sqrt{\frac{2}{\pi \cdot x}} \cdot 2^x \; . \qquad \square$$

*Proof of Theorem 7.2.1.* As in the proof of Theorem 7.1.1 we begin with the derivation of the estimates of the expectation of $\ln(\mathscr{C}_{\mathrm{G}}(A))$ from the tail estimates. For the general case we have

$$
\begin{aligned}
\mathbb{E}\left[\ln(\mathscr{C}_{\mathrm{G}}(A))\right] &= \int_0^\infty \mathrm{Prob}[\ln(\mathscr{C}_{\mathrm{G}}(A)) > s] \, ds \\
&< 1.5 \cdot \ln(n) + r + \int_{\ln(n^{3/2})+r}^\infty 6 \cdot \sqrt{m(n-m)} \cdot \exp(-s) \, ds \\
&= 1.5 \cdot \ln(n) + r + 6 \cdot \underbrace{\frac{\sqrt{m(n-m)}}{n^{3/2}}}_{\leq 2^{-3/2}} \cdot \exp(-r) \\
&< 1.5 \cdot \ln(n) + 1.8 \; ,
\end{aligned}
$$

if we choose $r := \ln\left(\frac{3}{\sqrt{2}}\right)$. Analogously, we get for the remaining cases LP and SOCP-1 with $c = 29$ or $c = 20$, respectively,

$$
\begin{aligned}
\mathbb{E}\left[\ln(\mathscr{C}_{\mathrm{G}}(A))\right] < \ln(m) + r + \int_{\ln(m)+r}^\infty c \cdot \sqrt{m} \cdot \exp(-s) \, ds = \ln(m) + \ln\left(\tfrac{c}{\sqrt{m}}\right) + 1 \\
< \ln(m) + \ln\left(\tfrac{c}{\sqrt{8}}\right) + 1 \; ,
\end{aligned}
$$

where $r := \ln\left(\frac{c}{\sqrt{m}}\right)$, and where the last inequality holds for $m \geq 8$.

It remains to show the tail estimates in (7.11)–(7.13). Recall from Proposition 7.0.2 and Corollary 6.1.4 that

$$\begin{aligned} \mathrm{Prob}[\mathscr{C}_{\mathrm{G}}(A) > t] &= \mathrm{rvol}\, \mathcal{T}(\Sigma_m(C), \alpha) \\ &\leq \frac{4m(n-m)}{n} \binom{n/2}{m/2} \cdot \sum_{j=0}^{n-2} V_j(K) \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \sum_{i=0}^{n-2} |d_{ij}^{nm}| \cdot I_{n,i}(\alpha)\,, \end{aligned} \tag{7.22}$$

where $\alpha = \arcsin\left(\frac{1}{t}\right)$, and

$$|d_{ij}^{nm}| \;=\; \frac{\binom{m-1}{\frac{i-j}{2}+\frac{m-1}{2}} \cdot \binom{n-m-1}{\frac{i+j}{2}-\frac{m-1}{2}}}{\binom{n-2}{j}}\;, \quad \text{if } i+j+m \equiv 1 \pmod 2\;,$$

and $|d_{ij}^{nm}| = 0$ else. Using the decomposition $\binom{n-2}{j} = \left[\begin{smallmatrix} n-2 \\ j \end{smallmatrix}\right] \cdot \binom{(n-2)/2}{j/2}$ (cf. Proposition 4.1.20), we get

$$\begin{aligned} \frac{4m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2} \cdot \left[\begin{smallmatrix} n-2 \\ j \end{smallmatrix}\right]}{\binom{n-2}{j}} &= \frac{4m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2}}{\binom{(n-2)/2}{j/2}} \\ &\overset{(4.18)}{=} \frac{4m(n-m)}{n} \cdot \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+2}{2}) \cdot \Gamma(\frac{n-m+2}{2})} \cdot \frac{\Gamma(\frac{j+2}{2}) \cdot \Gamma(\frac{n-j}{2})}{\Gamma(\frac{n}{2})} \\ &= 8 \cdot \frac{\Gamma(\frac{j+2}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\cdot\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-m}{2})}\;. \end{aligned}$$

From (7.22) and changing the summation via $i \leftarrow n-2-i$ and $j \leftarrow n-2-j$, we thus get

$$\begin{aligned} \mathrm{rvol}\, \mathcal{T}(\Sigma_m(C), \alpha) \leq 8 \cdot \sum_{\substack{i,j=0 \\ i+j+m\equiv 1 \\ (\mathrm{mod}\ 2)}}^{n-2} & V_{n-2-j}(K) \cdot \frac{\Gamma(\frac{j+2}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-m}{2})} \cdot \binom{m-1}{\frac{i-j}{2}+\frac{m-1}{2}} \\ & \cdot \binom{n-m-1}{\frac{i+j}{2}-\frac{m-1}{2}} \cdot I_{n,n-2-i}(\alpha)\;, \end{aligned} \tag{7.23}$$

where we interpret $\binom{k}{\ell} = 0$ if $\ell < 0$ or $\ell > k$, i.e., the above summation over $i,j$ in fact only runs over the rectangle determined by the inequalities $0 \leq \frac{i-j}{2} + \frac{m-1}{2} \leq m-1$ and $0 \leq \frac{i+j}{2} - \frac{m-1}{2} \leq n-m-1$. As the summation runs only over those $i,j$, for which $i+j+m \equiv 1 \bmod 2$, we may replace the summation over $j$ by a summation over $k = \frac{i-j}{2} + \frac{m-1}{2}$. The inequalities then transform into $0 \leq k \leq m-1$ and $0 \leq i-k \leq n-m-1$. So we have

$$\begin{aligned} (7.23) =\; 8 \cdot \sum_{i,k=0}^{n-2} V_{n-m-1-i+2k}(K) \cdot \frac{\Gamma(\frac{m+i-2k+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m-i+2k+1}{2})}{\Gamma(\frac{n-m}{2})} \\ \cdot \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha)\;. \end{aligned} \tag{7.24}$$

In order to continue, we need to estimate the intrinsic volumes. We therefore have to split up the argumentation according to the three claims in Theorem 7.2.1.

(1) For any cap $K \in \mathcal{K}^c(S^{n-1})$, we have $V_j(K) \leq \frac{1}{2}$ (cf. Proposition 4.4.10). Using Lemma 7.2.2 we have

$$\begin{aligned}
(7.24) \overset{(7.14)}{\leq} \; & 2 \cdot \sqrt{m(n-m)} \cdot \sum_{i,k=0}^{n-2} \left(\frac{m+i-2k}{2}\right)^{\frac{i-2k}{2}} \cdot \left(\frac{n-m-i+2k}{2}\right)^{\frac{-i+2k}{2}} \\
& \qquad \cdot \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha) \\
\overset{(7.15)}{\leq} \; & 2 \cdot \sqrt{m(n-m)} \cdot \sum_{i=0}^{n-2} I_{n,n-2-i}(\alpha) \cdot n^{\frac{i}{2}} \cdot \sum_{k=0}^{n-2} \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \\
\overset{(7.16)}{=} \; & 2 \cdot \sqrt{m(n-m)} \cdot \sum_{i=0}^{n-2} I_{n,n-2-i}(\alpha) \cdot n^{\frac{i}{2}} \cdot \binom{n-2}{i} \\
\overset{(7.17)}{<} \; & 6 \cdot \sqrt{m(n-m)} \cdot \frac{1}{t}\,, \quad \text{if } t > n^{\frac{3}{2}}\,.
\end{aligned}$$

This shows the tail estimate in (7.14).

(2) Let $C = \mathbb{R}^n_+$ be the positive orthant. Recall that for this case we have (cf. Remark 4.4.15)

$$V_j(K) = \frac{\binom{n}{j+1}}{2^n} \overset{(4.20)}{=} \frac{\sqrt{\pi} \cdot \Gamma(\frac{n+1}{2}) \cdot \Gamma(\frac{n+2}{2})}{\Gamma(\frac{j+2}{2}) \cdot \Gamma(\frac{j+3}{2}) \cdot \Gamma(\frac{n-j}{2}) \cdot \Gamma(\frac{n-j+1}{2}) \cdot 2^n}\,.$$

Continuing again from (7.24) we get

$$\begin{aligned}
\mathrm{rvol}\mathcal{T}(\Sigma_m(\mathbb{R}^n_+), \alpha) \leq \; & 8 \cdot \sum_{i,k=0}^{n-2} \frac{\binom{n}{m+i-2k}}{2^n} \cdot \frac{\Gamma(\frac{m+i-2k+1}{2})}{\Gamma(\frac{m}{2})} \cdot \frac{\Gamma(\frac{n-m-i+2k+1}{2})}{\Gamma(\frac{n-m}{2})} \\
& \cdot \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha) \\
= \; & 8\sqrt{\pi} \cdot \sum_{i,k=0}^{n-2} \frac{\Gamma(\frac{n+1}{2}) \cdot \Gamma(\frac{n+2}{2})}{2^n \cdot \Gamma(\frac{m+i-2k+2}{2}) \cdot \Gamma(\frac{n-m-i+2k+2}{2}) \cdot \Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \\
& \cdot \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha) \\
\overset{(4.18)}{=} \; & \frac{8\sqrt{\pi}}{2^n} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \cdot \sum_{i,k=0}^{n-2} \binom{n/2}{(m+i-2k)/2} \cdot \binom{m-1}{k} \\
& \qquad \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha) \\
\overset{(7.19)}{<} \; & \frac{16}{\sqrt{n} \cdot 2^{n/2}} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \cdot \sum_{i,k=0}^{n-2} \binom{m-1}{k} \cdot \binom{n-m-1}{i-k} \cdot I_{n,n-2-i}(\alpha) \\
\overset{(7.16)}{=} \; & \frac{16}{\sqrt{n} \cdot 2^{n/2}} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \cdot \sum_{i=0}^{n-2} \binom{n-2}{i} \cdot I_{n,n-2-i}(\alpha) \\
\overset{(7.18)}{<} \; & \frac{16}{\sqrt{n} \cdot 2^{n/2}} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \cdot \exp\left(\frac{n}{m}\right) \cdot \frac{1}{t}\,, \quad \text{if } t > m\,.
\end{aligned}$$

Estimating $\Gamma(\frac{n+1}{2}) < \sqrt{\frac{n}{2}} \cdot \Gamma(\frac{n}{2})$ (cf. (4.17)), and rewriting

$$\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2})} \overset{(4.18)}{=} \frac{m(n-m)}{2n} \cdot \binom{n/2}{m/2} ,$$

we finally get

$$\begin{aligned} \operatorname{rvol} \mathcal{T}(\Sigma_m(\mathbb{R}^n_+), \alpha) &< \sqrt{2} \cdot 8 \cdot \frac{\Gamma(\frac{n}{2}) \cdot \exp(\frac{n}{m})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m}{2}) \cdot 2^{\frac{n}{2}}} \cdot \frac{1}{t} \\ &= \sqrt{2} \cdot 4 \cdot \frac{m(n-m)}{n} \cdot \frac{\binom{n/2}{m/2}}{2^{\frac{n}{2}}} \cdot \exp(\tfrac{n}{m}) \cdot \frac{1}{t} \\ &\overset{(7.9)}{=} \sqrt{2} \cdot g_m(n) \cdot \exp(\tfrac{n}{m}) \cdot \frac{1}{t} . \end{aligned}$$

In the proof of Theorem 7.1.1 we have seen that for fixed $m$ the sequence $g_m(m+1)$, $g_m(m+2), g_m(m+3), \ldots$ is log-concave. Additionally, the sequence $\exp(\frac{m+1}{m})$, $\exp(\frac{m+2}{m}), \exp(\frac{m+3}{m}), \ldots$ is log-concave, which implies that also the sequence

$$(g_m(n) \cdot \exp(\tfrac{n}{m}) \mid n = m+1, m+2, \ldots)$$

is log-concave. Moreover, the maximum of this last sequence lies between $2m+5$ and $2m+7$ for $m \geq 8$, as

$$\frac{g_m(2m+4) \cdot \exp(\frac{2m+4}{m})}{g_m(2m+5) \cdot \exp(\frac{2m+5}{m})} = \frac{\Gamma(\frac{2m+4}{2})}{\Gamma(\frac{2m+5}{2})} \cdot \frac{\Gamma(\frac{m+5}{2})}{\Gamma(\frac{m+4}{2})} \cdot \frac{\sqrt{2}}{\exp\left(\frac{1}{m}\right)} < 1 ,$$

$$\frac{g_m(2m+8) \cdot \exp(\frac{2m+8}{m})}{g_m(2m+7) \cdot \exp(\frac{2m+7}{m})} = \frac{\Gamma(\frac{2m+8}{2})}{\Gamma(\frac{2m+7}{2})} \cdot \frac{\Gamma(\frac{m+7}{2})}{\Gamma(\frac{m+8}{2})} \cdot \frac{\exp\left(\frac{1}{m}\right)}{\sqrt{2}} < 1 , \quad \text{for } m \geq 8 .$$

So in order to get an estimate of $g_m(n) \cdot \exp(\frac{n}{m})$ for $m \geq 8$, we only have to check the cases $n \in \{2m+5, 2m+6, 2m+7\}$. The following asymptotics is verified easily (for example with a computer algebra system)

$$\frac{g_m(2m+k) \cdot \exp(\frac{2m+k}{m})}{\sqrt{m}} \xrightarrow{m \to \infty} \frac{4 \cdot \exp(2)}{\sqrt{2\pi}} < 12 .$$

With this asymptotics in mind, it is straightforward to verify that for $m \geq 7$

$$g_m(2m+5) \cdot \exp(\tfrac{2m+5}{m}) < 20 \cdot \sqrt{m}$$

$$g_m(2m+6) \cdot \exp(\tfrac{2m+6}{m}) < 20 \cdot \sqrt{m}$$

$$g_m(2m+7) \cdot \exp(\tfrac{2m+7}{m}) < 20 \cdot \sqrt{m} .$$

As $\sqrt{2} \cdot 20 < 29$, we finally get

$$\operatorname{rvol} \mathcal{T}(\Sigma_m(\mathbb{R}^n_+), \alpha) < 29 \cdot \sqrt{m} \cdot \frac{1}{t} ,$$

for $t > m \geq 8$.

(3) Lastly, we will treat the case $C=\mathcal{L}^n$. Recall that in this case we have for $0\le j\le n-2$ (cf. Example 4.4.8)

$$V_j(K)=\frac{\binom{(n-2)/2}{j/2}}{2^{\frac{n}{2}}}=\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{j+2}{2})\cdot\Gamma(\frac{n-j}{2})}\cdot\frac{1}{2^{\frac{n}{2}}}\;.$$

Continuing from (7.24) we get

$$\begin{aligned}
&\operatorname{rvol}\mathcal{T}(\Sigma_m(\mathcal{L}^n),\alpha)\\
&\quad\le 8\cdot\sum_{i,k=0}^{n-2}\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{m}{2})\cdot\Gamma(\frac{n-m}{2})}\cdot\frac{1}{2^{\frac{n}{2}}}\cdot\binom{m-1}{k}\cdot\binom{n-m-1}{i-k}\cdot I_{n,n-2-i}(\alpha)\\
&\quad\overset{(7.16)}{=}8\cdot\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{m}{2})\cdot\Gamma(\frac{n-m}{2})}\cdot\frac{1}{2^{\frac{n}{2}}}\cdot\sum_{i=0}^{n-2}\binom{n-2}{i}\cdot I_{n,n-2-i}(\alpha)\\
&\quad\overset{(7.18)}{=}8\cdot\frac{\Gamma(\frac{n}{2})\cdot\exp\left(\frac{n}{m}\right)}{\Gamma(\frac{m}{2})\cdot\Gamma(\frac{n-m}{2})\cdot 2^{\frac{n}{2}}}\cdot\frac{1}{t}\\
&\quad=g_m(n)\cdot\exp(\tfrac{n}{m})\cdot\frac{1}{t}\\
&\quad\overset{\text{(see above)}}{<}20\cdot\sqrt{m}\cdot\frac{1}{t}\;,
\end{aligned}$$

for $t>m\ge 8$. This finishes the proof. $\square$

## 7.3 Smoothed analysis – 1st order

In this section we will perform a smoothed analysis of first-order. By this we mean the following: Recall that for smooth $K\in\mathcal{K}^{\mathrm{sm}}(S^{n-1})$ the set $\Sigma_m(K)$ is a hypersurface of $\mathrm{Gr}_{n,m}$. In particular, it has an induced (lower-dimensional) volume functional. A first-order smoothed analysis means that we will give estimates of the (lower-dimensional) volume of the intersection of $\Sigma_m(K)$ with a ball $B(\mathcal{W}_0,\beta)$ in $\mathrm{Gr}_{n,m}$ divided by the (full-dimensional) volume of the ball $B(\mathcal{W}_0,\beta)$.

We already face a problem at this high-level approach, as it is not clear what metric on $\mathrm{Gr}_{n,m}$ one should choose, and what role this metric plays for the results. We will choose the Hausdorff metric, as this will allow a nice argumentation. It should be evident though, particularly in view of the transition from matrices to subspaces $\mathbb{R}^{m\times n}_*\to\mathrm{Gr}_{n,m}$, $A\mapsto\mathrm{im}(A^T)$, that the geodesic metric on $\mathrm{Gr}_{n,m}$ would be the more significant choice. But this would add another difficulty to the analysis, so that we restrict to the Hausdorff metric at this point. We define for $\mathcal{W}_0\in\mathrm{Gr}_{n,m}$, $0<\beta<\frac{\pi}{2}$,

$$B_{\mathrm{H}}(\mathcal{W}_0,\beta):=\{\mathcal{W}\in\mathrm{Gr}_{n,m}\mid d_{\mathrm{H}}(\mathcal{W}_0,\mathcal{W})<\beta\}\;,$$

the ball of radius $\beta$ around $\mathcal{W}_0$. A smoothed analysis of $\mathscr{C}_{\mathrm{G}}$ corresponds to estimating the volume of the intersection of $B_{\mathrm{H}}(\mathcal{W}_0,\beta)$ with the tube around $\Sigma_m$ relative to the volume of $B_{\mathrm{H}}(\mathcal{W}_0,\beta)$.

In Section D.3 (cf. Proposition D.3.4) we will show that the relative volume of $B_{\mathrm{H}}(\mathcal{W}_0,\beta)$ can be estimated from below via

$$\operatorname{rvol}B_{\mathrm{H}}(\mathcal{W}_0,\beta)\ge\sin(\beta)^{m(n-m)}\cdot\begin{bmatrix}n\\m\end{bmatrix}^{-1}\;.\tag{7.25}$$

So it remains to estimate the volume of the intersection of the tube around $\Sigma_m$ with the ball around $\mathcal{W}_0$.

As already mentioned, we will describe a first attempt to estimate the (lower-dimensional) volume of the intersection of $\Sigma_m$ with the ball around $\mathcal{W}_0$. Here, we will see that the elegant argumentation in Section 6.4 will inevitably fail. We will try a natural but coarse approach to overcome this obstacle. The result is stated in the following proposition. It is not yet satisfactory, as we think[1] that the Grassmann condition should allow a tail estimate of the form $n^{O(1)} \cdot \frac{1}{\sin(\beta)} \cdot \frac{1}{t}$. But at least we have a first proof of concept that the overall approach will work for smoothed analyses. A fully developed smoothed analysis remains open for future research.

**Theorem 7.3.1.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ be a smooth cap. Furthermore, let $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$, and let $0 < \beta < \frac{\pi}{2}$. Then we have*

$$\frac{\operatorname{vol}\left(\Sigma_m(K) \cap B_{\mathrm{H}}(\mathcal{W}_0, \beta)\right)}{\operatorname{vol} B_{\mathrm{H}}(\mathcal{W}_0, \beta)} < \sin(\beta)^{2-3m} \cdot m \cdot \begin{bmatrix} n \\ m-1 \end{bmatrix} \cdot \binom{n-m}{m} \cdot \binom{n-2}{m-1} .$$

*This gives rise to a first order smoothed analysis of*

$$\operatorname{Prob}[\mathscr{C}_{\mathrm{G}}(\mathcal{W}) > t] \lesssim 2 \cdot \sin(\beta)^{2-3m} \cdot m \cdot \begin{bmatrix} n \\ m-1 \end{bmatrix} \cdot \binom{n-m}{m} \cdot \binom{n-2}{m-1} \cdot \frac{1}{t} \tag{7.26}$$

$$< 2 \cdot \sin(\beta)^{2-3m} \cdot m \cdot \sqrt{n} \cdot \binom{n}{m}^{2.5} \cdot \frac{1}{t} , \tag{7.27}$$

*where $\mathcal{W}$ is chosen uniformly at random in a ball of radius $\beta$ w.r.t. the Hausdorff metric.*

Note that the asymptotic of the right-hand side in (7.27) is roughly of the form $n^{O(m)} \cdot \frac{1}{\sin(\beta)^{3m-2}} \cdot \frac{1}{t}$. This is certainly not satisfactory, but for fixed $m$ it is at least polynomial in $n$, which shows that it is not completely worthless.

The proof of Theorem 7.3.1 is basically a refinement of the proof of Lemma 6.5.1. It is convenient to use the concept of curvature measures, which can be interpreted as local versions of the intrinsic volumes. We will only define them for smooth caps $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, and also only for $0 \leq i \leq n-2$. See [30] and the references therein for more about the spherical curvature measures.

**Definition 7.3.2.** Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ and let $M := \partial K$. Then for $0 \leq i \leq n-2$ the *ith curvature measure* $\Phi_i(K, .)$ is defined via

$$\Phi_i(K, B) := \frac{1}{\mathcal{O}_i \cdot \mathcal{O}_{n-2-i}} \cdot \int_{B \cap M} \sigma_{n-2-i}(p)\, dp , \tag{7.28}$$

where $B \subseteq S^{n-1}$ is a Borel set.

**Remark 7.3.3.** Note that by Proposition 4.4.4 we have $\Phi_i(K, S^{n-1}) = V_i(K)$. Furthermore, we have $\Phi_i(K, \emptyset) = 0$, and

$$\Phi_i(K, B) \leq \Phi_i(K, S^{n-1}) = V_i(K) \leq \tfrac{1}{2} ,$$

where the first inequality follows from the nonnegativity of the function $\sigma_i(p)$, and the second inequality follows from Proposition 4.4.10.

[1]We derive this conjecture from the average analysis in Section 7.2 and from a comparison to other smoothed analysis results (cf. the results listed in [12]).

We will need the following local version of the kinematic formula.

**Theorem 7.3.4.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, and let $B \subseteq S^{n-1}$ be a Borel set. Then for $0 \le i \le n-2$ we have*

$$\Phi_{n-2-i}(K,B) \;=\; \mathop{\mathbb{E}}_{S\in\mathcal{S}^{i+1}(S^{n-1})}\left[\Phi_0(K\cap S, B\cap S)\right] \;, \tag{7.29}$$

*where $S \in \mathcal{S}^{i+1}(S^{n-1})$ is chosen uniformly at random.*

Note that if $M \subset S^{n-1}$ is a manifold and $S \subset S^{n-1}$ is a uniformly random subsphere of some fixed dimension, then with probability 1 the intersection $M \cap S$ is again a smooth manifold or empty. So the quantity $\Phi_0(K\cap S, B\cap S)$ in (7.29) is well-defined with probability 1.

*Proof.* This is a special case of Korollar 5.2.2 in [30] and a special case of Theorem 2.7 in [17]. More precisely, for $B \subseteq S^{n-1}$ open and $U := B\cap M$ the formula given in [17, Thm 2.7] is

$$\int_U \sigma_i(p)\,dp = C(p,i)\cdot \mathop{\mathbb{E}}_{S\in\mathcal{S}^{i+1}(S^{n-1})}\left[\int_{U\cap S}\sigma_i(p)\,dp\right] \;, \tag{7.30}$$

where $p := n+1$, and the constant $C(p,i)$ is given by[2] $C(p,i) := (p-i-1)\cdot\binom{p-1}{i}\cdot\frac{\mathcal{O}_{p-1}\cdot\mathcal{O}_p}{\mathcal{O}_i\cdot\mathcal{O}_{i+1}\cdot\mathcal{O}_{p-i-2}}$. Using Proposition 4.1.20 we can simplify this constant to

$$\begin{aligned}
C(p,i) &\overset{(4.21)}{=} (p-i-1)\cdot\binom{(p-1)/2}{i/2}\cdot\begin{bmatrix}p-1\\ i\end{bmatrix}\cdot\frac{\mathcal{O}_{p-1}\cdot\mathcal{O}_p}{\mathcal{O}_i\cdot\mathcal{O}_{i+1}\cdot\mathcal{O}_{p-i-2}}\\
&\overset{(4.22)}{=} (p-i-1)\cdot\frac{\omega_i\cdot\omega_{p-1-i}}{\omega_{p-1}}\cdot\frac{\mathcal{O}_i\cdot\mathcal{O}_{p-1-i}}{2\cdot\mathcal{O}_{p-1}}\cdot\frac{\mathcal{O}_{p-1}\cdot\mathcal{O}_p}{\mathcal{O}_i\cdot\mathcal{O}_{i+1}\cdot\mathcal{O}_{p-i-2}}\\
&= \frac{\mathcal{O}_{i-1}}{i\cdot\mathcal{O}_{i+1}}\cdot\frac{(p-1)\cdot\mathcal{O}_p}{\mathcal{O}_{p-2}}\cdot\frac{\mathcal{O}_{p-1-i}}{2}\\
&\overset{(4.15)}{=} \frac{\mathcal{O}_{p-1-i}}{2}\;.
\end{aligned}$$

Having simplified the constant $C(p,i)$, we see that (7.30) is indeed equivalent to (7.29) (note that $p = n-1$). □

With the local kinematic formula we can give an estimate of the curvature measure if the Borel set is a tube around a subsphere. (This curvature measure will appear in the proof of Theorem 7.3.1.) The following lemma is a generalization of [13, Lem. 4.6]. The proof is a straightforward extension of the proof given in [13].

**Lemma 7.3.5.** *Let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, and let $S_0 \in \mathcal{S}^{m-1}(S^{n-1})$ be an $(m-1)$-subsphere. Then for $0 \le \beta \le \frac{\pi}{2}$, we have*

$$\Phi_{n-m-1}(K,\mathcal{T}(S_0,\beta)) \;\le\; \sin(\beta)^{n-2m}\cdot\frac{m}{2(n-m)}\cdot\binom{n-m}{m}\;.$$

[2] In [17, Thm 2.7] there is a typo in the definition of the constant $C(p,i)$. Namely, the binomial coefficient $\binom{p-1}{i}$ is falsely replaced by $\frac{p-1}{i}$ (cp. [13, Thm. 4.2]).

*Proof.* The proof goes analogous to the proof of [13, Lemma 4.6] (cf. also [17, Prop. 3.2]). Using the kinematic formula in Theorem 7.3.4 and the fact that the curvature measures may be estimated by $\frac{1}{2}$ (cf. Remark 7.3.3), we get

$$\begin{aligned} \Phi_{n-m-1}(K,B) &= \underset{S\in\mathcal{S}^m(S^{n-1})}{\mathbb{E}} [\Phi_0(K\cap S, B\cap S)] \\ &\leq \underset{S\in\mathcal{S}^m(S^{n-1})}{\mathbb{E}} \left[\begin{pmatrix} 0 & \text{if } B\cap S=\emptyset \\ \frac{1}{2} & \text{if } B\cap S\neq\emptyset \end{pmatrix}\right] \\ &= \tfrac{1}{2}\cdot \underset{S\in\mathcal{S}^m(S^{n-1})}{\operatorname{Prob}} [B\cap S\neq\emptyset] \ . \end{aligned}$$

Setting $B := \mathcal{T}(S_0,\beta)$, we get

$$\begin{aligned} \operatorname*{Prob}_S [\mathcal{T}(S_0,\beta)\cap S\neq\emptyset] &= \operatorname*{Prob}_S [\exists p\in S : d(p,S_0)\leq\beta] \\ &= \operatorname*{Prob}_S [\sphericalangle_{\min}(S,S_0)\leq\beta] \ , \end{aligned}$$

where $\sphericalangle_{\min}(S,S_0)$ shall denote the minimum principal angle between $S_0$ and $S$. In Corollary D.3.5 in Section D.3 in the appendix we will show that

$$\operatorname*{Prob}_S [\sphericalangle_{\min}(S,S_0)\leq\beta] \leq \frac{m(n-2m)}{n-m}\cdot\binom{n-m}{m}\cdot I_{n-2m+2,1}(\beta) \ .$$

Using the estimate of the $I$-function, that we stated in (7.20), we get

$$I_{n-2m+2,1}(\beta) = \int_0^\beta \cos(\rho)\cdot\sin(\rho)^{n-2m-1}\,d\rho \ \leq \ \frac{\sin(\beta)^{n-2m}}{n-2m} \ .$$

Putting everything together finally yields

$$\Phi_{n-m-1}(K,B) \ \leq \ \frac{1}{2}\cdot\frac{m(n-2m)}{n-m}\cdot\binom{n-m}{m}\cdot\frac{\sin(\beta)^{n-2m}}{n-2m} \ . \qquad \square$$

For convenience, we formulate another simple lemma that we will use in the proof of Theorem 7.3.1.

**Lemma 7.3.6.** *Let $\mathcal{W}\in\mathrm{Gr}_{n,m}$, and let $H\in\mathrm{Gr}_{n,n-1}$ be a hyperplane that contains $\mathcal{W}$. Furthermore, let $p\in S^{n-1}\setminus H^\perp$, and let $\bar p\in H\cap S^{n-1}$ denote the normalization of the projection of $p$ on $H$, i.e., $\bar p = \|\Pi_H(p)\|^{-1}\cdot\Pi_H(p)$, where $\Pi_H$ denotes the orthogonal projection on $H$. Then $d^e(\bar p,\mathcal{W})\leq d^e(p,\mathcal{W})$.*

*Proof.* W.l.o.g. we may assume that $\mathcal{W}=\mathbb{R}^m\times\{0\}$ and $H=\mathbb{R}^{n-1}\times\{0\}$. If $p=(x_1,\ldots,x_n)^T$, then $\bar p = (1-x_n^2)^{-\frac{1}{2}}\cdot(x_1,\ldots,x_{n-1},0)^T$, and we have

$$\begin{aligned} d^e(p,\mathcal{W}) &= \sqrt{x_{m+1}^2+\ldots+x_n^2}\ , \\ d^e(\bar p,\mathcal{W}) &= (1-x_n^2)^{-\frac{1}{2}}\cdot\sqrt{x_{m+1}^2+\ldots+x_{n-1}^2}\ . \end{aligned}$$

As $(1-x_n^2)\cdot(x_{m+1}^2+\ldots+x_n^2)-(x_{m+1}^2+\ldots+x_{n-1}^2) = x_n^2\cdot(1-(x_{m+1}^2+\ldots+x_n^2))\geq 0$, we have

$$x_{m+1}^2+\ldots+x_n^2 \geq \frac{x_{m+1}^2+\ldots+x_{n-1}^2}{1-x_n^2}\ ,$$

which implies $d^e(p,\mathcal{W})\geq d^e(\bar p,\mathcal{W})$. $\square$

*Proof of Theorem 7.3.1.* To ease the notation, let us abbreviate

$$\tilde{\Sigma} := \Sigma_m \cap B_{\mathrm{H}}(\mathcal{W}_0, \beta) .$$

Recall that $\Pi_M \colon \Sigma_m \to M$ denotes the canonical projection from $\Sigma_m$ to $M$, where $M = \partial K$ as usual. Let us denote the restriction of $\Pi_M$ to $\tilde{\Sigma}$ by

$$\tilde{\Pi} \colon \tilde{\Sigma} \to M , \quad \tilde{\Pi}(\mathcal{W}) = \Pi_M(\mathcal{W}) .$$

Furthermore, we denote the image of $\tilde{\Pi}$ by $\tilde{M}$. Note that $\mathcal{W} \in B_{\mathrm{H}}(\mathcal{W}_0, \beta)$ implies that for the corresponding subspheres we have $S \subset \mathcal{T}(S_0, \beta)$, where $S := \mathcal{W} \cap S^{n-1}$ and $S_0 := \mathcal{W}_0 \cap S^{n-1}$. Therefore, we have

$$\tilde{M} \subseteq \mathcal{T}(S_0, \beta) \cap M . \tag{7.31}$$

From the coarea formula (cf. Corollary 5.1.1) we get the following formula for the volume of $\tilde{\Sigma}$

$$\operatorname{vol} \tilde{\Sigma} = \int_{\tilde{\Sigma}} 1 \, d\mathcal{W} \overset{(5.5)}{=} \int_{p \in \tilde{M}} \int_{\mathcal{W} \in \tilde{\Pi}^{-1}(p)} \operatorname{ndet}(D_{\mathcal{W}} \tilde{\Pi})^{-1} \, d\mathcal{W} \, dp .$$

Note that as $\tilde{\Pi}$ is the restriction of $\Pi_M$ to an open subset of the domain, their derivatives coincide. Denoting by $\tilde{\mathrm{Gr}}$ the set of elements $\mathcal{Y} \in \mathrm{Gr}(T_pM, m-1)$ such that $\mathcal{W} := \mathbb{R}\, p + \mathcal{Y}$ lies in the ball around $\mathcal{W}_0$, i.e.,

$$\tilde{\mathrm{Gr}} := \{\mathcal{Y} \in \mathrm{Gr}(T_pM, m-1) \mid \mathbb{R}\, p + \mathcal{Y} \in B_{\mathrm{H}}(\mathcal{W}_0, \beta)\} ,$$

we get

$$\operatorname{vol} \tilde{\Sigma} = \int_{p \in \tilde{M}} \int_{\mathcal{W} \in \tilde{\Pi}^{-1}(p)} \operatorname{ndet}(D_{\mathcal{W}} \Pi_M)^{-1} \, d\mathcal{W} \, dp \overset{\text{Lem. 6.2.9}}{=} \int_{p \in \tilde{M}} \int_{\mathcal{Y} \in \tilde{\mathrm{Gr}}} \det(W_{p,\mathcal{Y}}) \, d\mathcal{W} \, dp ,$$

where $W_{p,\mathcal{Y}}$ denotes as usual the restriction of the Weingarten map $W_p$ of $M$ at $p$ to $\mathcal{Y}$. This is the point where we face the problem that we can not argue as in Section 6.4, as the integral is not over the whole fiber $\mathrm{Gr}(T_pM, m-1)$ but only over $\tilde{\mathrm{Gr}}$. We now have two possibilities for a coarse estimation:

1. We extend the set $\tilde{\mathrm{Gr}}$ to all of $\mathrm{Gr}(T_pM, m-1)$, and use the arguments from Section 6.4.
2. We estimate $\det(W_{p,\mathcal{Y}})$ by $\sigma_{m-1}(p)$ (losing a factor of $\binom{n-2}{m-1}$, compared to Corollary 6.4.9) and estimate the volume of $\tilde{\mathrm{Gr}}$.

It turns out that the first approach is too coarse, so that we will only pursue the second approach.

So we have

$$\int_{\mathcal{Y} \in \tilde{\mathrm{Gr}}} \det(W_{p,\mathcal{Y}}) \, d\mathcal{W} \;\leq\; \int_{\mathcal{Y} \in \tilde{\mathrm{Gr}}} \sigma_{m-1}(p) \, d\mathcal{W} \;=\; \sigma_{m-1}(p) \cdot \operatorname{vol}(\tilde{\mathrm{Gr}}) .$$

To estimate the volume of $\tilde{\mathrm{Gr}}$, we define for $p \in M$

$$\overline{\mathcal{W}_0}(p) := (\text{orthogonal projection of } \mathcal{W}_0 \text{ on } \nu(p)^{\perp}) ,$$

where $\nu(p)$ denotes the unit normal field pointing inwards the cap $K$ (cf. Section 4.1.2). The space $\overline{\mathcal{W}_0}(p)$ is a linear subspace of the hyperplane $\nu(p)^\perp$ of dimension at most $m$. In fact, for almost all $p \in M$ the dimension of $\overline{\mathcal{W}_0}(p)$ is exactly $m$, as is easily checked. In the following we assume that $\dim \overline{\mathcal{W}_0}(p) = m$, i.e., $\overline{\mathcal{W}_0}(p) \in \mathrm{Gr}_{n,m}$. Note that $\overline{\mathcal{W}_0}(p)$ does not necessarily lie in $\Sigma_m$. Nevertheless, we have by Lemma 7.3.6

$$d_{\mathrm{H}}(\mathcal{W}_0, \mathcal{W}) < \beta \;\Rightarrow\; d_{\mathrm{H}}(\overline{\mathcal{W}_0}(p), \mathcal{W}) < \beta \;, \tag{7.32}$$

if $\mathcal{W} \in \Sigma_m$ and $\mathcal{W} \cap K = \{p\}$. So we get

$$\mathrm{vol}(\tilde{\mathrm{Gr}}) \leq \mathrm{vol}\,\mathrm{Gr}_{n-2,m-1} \cdot \operatorname*{Prob}_{\bar{\mathcal{W}} \in \mathrm{Gr}_{n-2,m-1}} \left[\sphericalangle_{\max}(\mathbb{R}^m, \bar{\mathcal{W}}) \leq \beta\right] \;.$$

In Corollary D.3.5 in Section D.3 in the appendix we will see that

$$\begin{aligned} \operatorname*{Prob}_{\bar{\mathcal{W}} \in \mathrm{Gr}_{n-2,m-1}} \left[\sphericalangle_{\max}(\mathbb{R}^m, \bar{\mathcal{W}}) \leq \beta\right] &= \sin(\beta)^{(m-1)(n-m-1)-(m-1)} \\ &= \sin(\beta)^{2-m-n+m(n-m)} \;. \end{aligned}$$

Using (7.25), and using

$$\begin{aligned} \int_{\tilde{M}} \sigma_{m-1}(p)\, dp &\overset{(7.31)}{\leq} \int_{\mathcal{T}(S_0,\beta)\cap M} \sigma_{m-1}(p)\, dp \\ &\overset{(7.28)}{=} \mathcal{O}_{m-1} \cdot \mathcal{O}_{n-m-1} \cdot \Phi_{n-m-1}(K, \mathcal{T}(S_0,\beta)) \;, \end{aligned}$$

we may therefore conclude

$$\begin{aligned} \frac{\mathrm{vol}\,\tilde{\Sigma}}{\mathrm{vol}\,B_{\mathrm{H}}(\mathcal{W}_0,\beta)} &= \frac{\mathrm{vol}\,\tilde{\Sigma}}{\mathrm{vol}\,\mathrm{Gr}_{n,m} \cdot \mathrm{rvol}\,B_{\mathrm{H}}(\mathcal{W}_0,\beta)} \\ &\leq \frac{\mathrm{vol}\,\mathrm{Gr}_{n-2,m-1}}{\mathrm{vol}\,\mathrm{Gr}_{n,m}} \cdot \begin{bmatrix} n \\ m \end{bmatrix} \cdot \sin(\beta)^{2-m-n} \cdot \int_{p\in\tilde{M}} \sigma_{m-1}(p)\, dp \\ &\overset{(6.50)}{=} \frac{\mathcal{O}_{m-1}^2 \cdot \mathcal{O}_{n-m-1}^2}{\mathcal{O}_{n-2} \cdot \mathcal{O}_{n-1}} \cdot \begin{bmatrix} n \\ m \end{bmatrix} \cdot \sin(\beta)^{2-m-n} \cdot \Phi_{n-m-1}(K, \mathcal{T}(S_0,\beta)) \\ &\overset{\text{Lem. 7.3.5}}{\leq} \frac{\mathcal{O}_{m-1}^2 \cdot \mathcal{O}_{n-m-1}^2}{\mathcal{O}_{n-2} \cdot \mathcal{O}_{n-1}} \cdot \begin{bmatrix} n \\ m \end{bmatrix} \cdot \sin(\beta)^{2-3m} \cdot \frac{m}{2(n-m)} \cdot \binom{n-m}{m} \;. \end{aligned}$$

It remains to polish the constant. More precisely, in order to get (7.26) we will show that

$$\frac{\mathcal{O}_{m-1}^2 \cdot \mathcal{O}_{n-m-1}^2}{\mathcal{O}_{n-2} \cdot \mathcal{O}_{n-1}} \cdot \begin{bmatrix} n \\ m \end{bmatrix} \cdot \frac{1}{2(n-m)} = \begin{bmatrix} n \\ m-1 \end{bmatrix} \cdot \binom{n-2}{m-1} \;. \tag{7.33}$$

For this we use again Proposition 4.1.20:

$$\begin{aligned}
&\frac{\mathcal{O}_{m-1}^2\cdot\mathcal{O}_{n-m-1}^2}{\mathcal{O}_{n-2}\cdot\mathcal{O}_{n-1}}\cdot\begin{bmatrix}n\\m\end{bmatrix}\cdot\frac{1}{2(n-m)}\cdot\begin{bmatrix}n\\m-1\end{bmatrix}^{-1}\cdot\binom{n-2}{m-1}^{-1}\\
&\overset{(4.21)}{\underset{(4.22)}{=}}\frac{\mathcal{O}_{m-1}^2\cdot\mathcal{O}_{n-m-1}^2}{\mathcal{O}_{n-2}\cdot\mathcal{O}_{n-1}}\cdot\frac{\mathcal{O}_m\cdot\mathcal{O}_{n-m}}{2\cdot\mathcal{O}_n}\cdot\frac{1}{2(n-m)}\\
&\qquad\qquad\cdot\frac{2\cdot\mathcal{O}_n}{\mathcal{O}_{m-1}\cdot\mathcal{O}_{n-m+1}}\cdot\frac{2\cdot\mathcal{O}_{n-2}}{\mathcal{O}_{m-1}\cdot\mathcal{O}_{n-m-1}}\cdot\frac{\omega_{n-2}}{\omega_{m-1}\cdot\omega_{n-m-1}}\\
&=\frac{\mathcal{O}_{n-m-1}}{\mathcal{O}_{n-m+1}}\cdot\frac{\mathcal{O}_m\cdot\mathcal{O}_{n-m}}{\mathcal{O}_{n-1}}\cdot\frac{1}{n-m}\cdot\frac{\omega_{n-2}}{\omega_{m-1}\cdot\omega_{n-m-1}}\\
&=\frac{\mathcal{O}_{n-m-1}}{\mathcal{O}_{n-m+1}}\cdot\frac{\mathcal{O}_m\cdot\mathcal{O}_{n-m}}{\mathcal{O}_{n-1}}\cdot\frac{1}{n-m}\cdot\frac{(m-1)\cdot(n-m-1)}{n-2}\cdot\frac{\mathcal{O}_{n-3}}{\mathcal{O}_{m-2}\cdot\mathcal{O}_{n-m-2}}\\
&=\frac{\mathcal{O}_{n-m-1}}{(n-m)\cdot\mathcal{O}_{n-m+1}}\cdot\frac{(m-1)\cdot\mathcal{O}_m}{\mathcal{O}_{m-2}}\cdot\frac{(n-m-1)\cdot\mathcal{O}_{n-m}}{\mathcal{O}_{n-m-2}}\cdot\frac{\mathcal{O}_{n-3}}{(n-2)\cdot\mathcal{O}_{n-1}}\\
&\overset{(4.15)}{=}1\,.
\end{aligned}$$

This shows (7.33) and thus (7.26).

For the inequality (7.27) we need to show that

$$\begin{bmatrix}n\\m-1\end{bmatrix}\cdot\binom{n-2}{m-1}\cdot\binom{n-m}{m}<\sqrt{n}\cdot\binom{n}{m}^{2.5}\,.$$

First of all, we have

$$\frac{\binom{n-m}{m}}{\binom{n}{m}}=\frac{(n-m)(n-m-1)\cdots(n-2m)}{n(n-1)\cdots(n-m)}\le 1\,,$$

$$\frac{\binom{n-2}{m-1}}{\binom{n}{m}}=\frac{m(n-m)}{n(n-1)}\le\frac{\frac{n^2}{4}}{n(n-1)}=\frac{n}{4(n-1)}\le\frac{1}{2}\,,$$

so that it remains to show

$$\begin{bmatrix}n\\m-1\end{bmatrix}<2\cdot\sqrt{n}\cdot\binom{n}{m}^{0.5}\,. \tag{7.34}$$

Recall from Proposition 4.1.20 that $\binom{n}{m}=\left[\begin{smallmatrix}n\\m\end{smallmatrix}\right]\cdot\binom{n/2}{m/2}$. Using the estimate

$$\Gamma(x+\tfrac{1}{2})<\sqrt{x}\cdot\Gamma(x)$$

(cf. (4.17)), we get

$$\begin{aligned}
\binom{n/2}{m/2}&\overset{(4.18)}{=}\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{m+2}{2})\cdot\Gamma(\frac{n-m+2}{2})}=\frac{n}{2}\cdot\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{m+2}{2})\cdot\Gamma(\frac{n-m+2}{2})}\\
&>\frac{n}{2}\cdot\frac{\sqrt{\frac{2}{n}}\cdot\Gamma(\frac{n+1}{2})}{\sqrt{\frac{m+1}{2}}\cdot\Gamma(\frac{m+1}{2})\cdot\sqrt{\frac{n-m+1}{2}}\cdot\Gamma(\frac{n-m+1}{2})}\\
&\overset{(4.19)}{=}\sqrt{\frac{2n}{(m+1)(n-m+1)\cdot\pi}}\cdot\begin{bmatrix}n\\m\end{bmatrix}\,.
\end{aligned}$$

This implies

$$\binom{n}{m} = \begin{bmatrix} n \\ m \end{bmatrix} \cdot \binom{n/2}{m/2} > \sqrt{\frac{2n}{(m+1)(n-m+1)\cdot \pi}} \cdot \begin{bmatrix} n \\ m \end{bmatrix}^2 . \tag{7.35}$$

Furthermore, we estimate

$$\begin{aligned} \frac{\left[\begin{smallmatrix} n \\ m-1 \end{smallmatrix}\right]}{\left[\begin{smallmatrix} n \\ m \end{smallmatrix}\right]} &= \frac{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-m+2}{2})}{\Gamma(\frac{m+1}{2}) \cdot \Gamma(\frac{n-m+1}{2})} = \frac{2}{m} \cdot \frac{\Gamma(\frac{m+2}{2})}{\Gamma(\frac{m+1}{2})} \cdot \frac{\Gamma(\frac{n-m+2}{2})}{\Gamma(\frac{n-m+1}{2})} \\ &< \frac{2}{m} \cdot \sqrt{\frac{m+1}{2}} \cdot \sqrt{\frac{n-m+1}{2}} = \frac{\sqrt{(m+1)(n-m+1)}}{m} . \end{aligned} \tag{7.36}$$

So finally we get

$$\begin{aligned} \frac{\left[\begin{smallmatrix} n \\ m-1 \end{smallmatrix}\right]}{\sqrt{\binom{n}{m}}} &\overset{(7.35)}{<} \left(\frac{(m+1)(n-m+1)\cdot \pi}{2n}\right)^{\frac{1}{4}} \cdot \frac{\left[\begin{smallmatrix} n \\ m-1 \end{smallmatrix}\right]}{\left[\begin{smallmatrix} n \\ m \end{smallmatrix}\right]} \\ &\overset{(7.36)}{<} \left(\frac{(m+1)(n-m+1)\cdot \pi}{2n}\right)^{\frac{1}{4}} \cdot \frac{\sqrt{(m+1)(n-m+1)}}{m} \\ &= \left(\frac{\pi}{2}\right)^{\frac{1}{4}} \cdot \underbrace{\left(\frac{m+1}{m}\right)^{\frac{3}{4}}}_{\leq 2^{3/4}} \cdot \underbrace{\left(\frac{n-m+1}{n}\right)^{\frac{1}{4}}}_{\leq 1} \cdot \underbrace{\frac{\sqrt{n-m+1}}{m^{\frac{1}{4}}}}_{\leq \sqrt{n}} \\ &< 2 \cdot \sqrt{n} , \end{aligned}$$

which shows (7.34) and thus finishes the proof. □

# Appendix A

# Miscellaneous

This chapter is mainly devoted to outsourced computations that are too long for an inclusion in the text, but that we give for the sake of completeness.

## A.1 On a threshold phenomenon in the sphere

In the average analysis of the Grassmann condition we have seen that the ratio $\frac{m}{n}$ plays a significant role for the asymptotic behavior of the expectation of the Grassmann condition (cf. Remark 7.1.2 and Remark 7.1.3). In this section we will compute some elementary asymptotics in the sphere that may give a reason for the appearance of this ratio.

To start with, it is well known that the volume of $S^{n-1}$ concentrates around any hypersphere, i.e., any $(n-2)$-dimensional subsphere of $S^{n-1}$. In formulas, this means the following. For every $\alpha > 0$ we have

$$\lim_{n\to\infty} \frac{\operatorname{vol} \mathcal{T}(S^{n-2}, \alpha)}{\operatorname{vol} S^{n-1}} = 1 \,,$$

where we consider $S^{n-2}$ as an embedded hypersphere in $S^{n-1}$. What if we don't want to consider hyperspheres, but general $m$-dimensional subspheres? It turns out that the ratio $\frac{m}{n}$ is the decisive quantity. This is shown by the following proposition. In broad terms, its statement is that for large $n$ the relative volume of the $\alpha$-tube around an $m$-dimensional subsphere is almost zero if $\alpha < \arccos(\sqrt{\frac{m}{n}})$ and almost 1 if $\alpha > \arccos(\sqrt{\frac{m}{n}})$. So there is an interesting threshold phenomenon that generalizes the above mentioned observation about the concentration of the volume around hyperspheres. Moreover, this threshold phenomenon could be the reason for the appearance of the ratio $\frac{m}{n}$ in the average analysis of the Grassmann condition.

**Proposition A.1.1.** *Let $c \in [0,1]$, and let $m\colon \mathbb{N} \to \mathbb{N}$ be such that $m(n) \leq n$ for all $n \in \mathbb{N}$, and $\lim \frac{m(n)}{n} = c$. Then*

$$\lim_{n\to\infty} \frac{\operatorname{vol} \mathcal{T}(S^{m(n)-1}, \alpha)}{\operatorname{vol} S^{n-1}} = \begin{cases} 0 & \text{if } \alpha < \arccos(\sqrt{c}) \\ 1 & \text{if } \alpha > \arccos(\sqrt{c}) \,. \end{cases}$$

We will prove this proposition by making use of the kinematic formula and the following well-known fact about the binomial distribution.

**Proposition A.1.2.** *For $n \in \mathbb{N}$ and $c \in [0,1]$ let*

$$f_{n,c}(k) := \binom{n}{k} \cdot c^k \cdot (1-c)^{n-k} \ ,$$
$$F_{n,c}(x) := \sum_{k=0}^{\lfloor xn \rfloor} f_{n,c}(k) \ .$$

*Then for $n \to \infty$ the functions $F_{n,c}$ converge pointwise to*

$$F_{n,c} \to F_c \ , \quad F_c(x) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x > c \ . \end{cases}$$

*Proof.* This is readily deduced from Chebyshev's inequality

$$\mathrm{Prob}\big[|\mu - X| \geq a\big] \leq \frac{\sigma^2}{a^2} \ ,$$

where $\mu := \mathbb{E}[X]$ denotes the expectationo of $X$, and $\sigma^2$ denotes the variance of $X$. More precisely, the expectation of the binomial distribution is given by $\mu = n \cdot c$, and the variance is given by $\sigma^2 = n \cdot c(1-c)$. So for $x < c$ we have

$$F_{n,c}(x) = \sum_{k=0}^{\lfloor xn \rfloor} f_{n,c}(k) \leq \mathrm{Prob}\big[|\mu - X| \geq n(c-x)\big] \leq \frac{n \cdot c(1-c)}{n^2 \cdot (c-x)^2} \stackrel{n\to\infty}{\longrightarrow} 0 \ .$$

Similarly, we get that $F_{n,c}(x) \to 1$ for $n \to \infty$, if $x > c$. □

*Proof of Proposition A.1.1.* Observe that for $p \in S^{n-1}$ and $S \in \mathcal{S}^k(S^{n-1})$ we have

$$p \in \mathcal{T}(S, \alpha) \iff S \cap B(p, \alpha) \neq \emptyset \ .$$

It is irrelevant if we fix $S_0 \in \mathcal{S}^k(S^{n-1})$ and choose $p \in S^{n-1}$ uniformly at random, or if we fix $p_0 \in S^{n-1}$ and choose $S \in \mathcal{S}^k(S^{n-1})$ uniformly at random, which is seen by the following small computation

$$\begin{aligned}
\frac{\mathrm{vol}\,\mathcal{T}(S_0,\alpha)}{\mathrm{vol}\,S^{n-1}} &= \mathop{\mathrm{Prob}}_{p \in S^{n-1}}[p \in \mathcal{T}(S_0,\alpha)] = \mathop{\mathrm{Prob}}_{Q \in O(n)}[Qp_0 \in \mathcal{T}(S_0,\alpha)] \\
&= \mathop{\mathrm{Prob}}_{Q \in O(n)}[p_0 \in \mathcal{T}(Q^{-1}S_0,\alpha)] = \mathop{\mathrm{Prob}}_{S \in \mathcal{S}^k(S^{n-1})}[p_0 \in \mathcal{T}(S,\alpha)] \\
&= \mathop{\mathrm{Prob}}_{S \in \mathcal{S}^k(S^{n-1})}[B(p_0,\alpha) \cap S \neq \emptyset] \ .
\end{aligned}$$

Applying the kinematic formula as stated in Corollary 4.4.12 and using the formula for the intrinsic volumes of circular caps as derived in Example 4.4.8 yields

$$\begin{aligned}
\mathrm{Prob}[B(p_0,\alpha) \cap S \neq \emptyset] &= 2 \cdot \sum_{\substack{j=n-1-k \\ j \equiv n-1-k \\ (\mathrm{mod}\ 2)}}^{n-1} V_j(B(p_0,\alpha)) \\
&= \sum_{\substack{j=n-1-k \\ j \equiv n-1-k \\ (\mathrm{mod}\ 2)}}^{n-2} \frac{\Gamma(\frac{n-2}{2}+1)}{\Gamma(\frac{j}{2}+1) \cdot \Gamma(\frac{n-2-j}{2}+1)} \cdot \sin(\alpha)^j \cdot \cos(\alpha)^{n-2-j} \\
&\quad + \delta_{\mathrm{even}}(k) \cdot V_{n-1}(B(p_0,\alpha)) \ ,
\end{aligned}$$

where $\delta_{\text{even}}(k) = 1$ if $k$ is even, and $\delta_{\text{even}}(k) = 0$ if $k$ is odd. The single term $V_{n-1}(B(p_0, \alpha))$, which we have computed in Proposition 4.1.18, is irrelevant for the asymptotics that we are interested in, as

$$V_{n-1}(B(p_0,\alpha)) = \frac{\mathcal{O}_{n-2}}{\mathcal{O}_{n-1}} \cdot \int_0^\alpha \sin(\rho)^{n-2}\, d\rho \to 0 \quad \text{for } n \to \infty\,.$$

With $h := \frac{n-2}{2}$, $\ell := \frac{n-2-j}{2}$, and $c := \cos^2\alpha$ we may rewrite

$$\frac{\Gamma(\frac{n-2}{2}+1)}{\Gamma(\frac{j}{2}+1)\cdot\Gamma(\frac{n-2-j}{2}+1)} \cdot \sin(\alpha)^j \cdot \cos(\alpha)^{n-2-j} = \binom{h}{\ell} \cdot (1-c)^{h-\ell} \cdot c^\ell\,.$$

For $n$ even and $k$ odd we thus get

$$\mathrm{Prob}[B(p_0,\alpha)\cap S \neq \emptyset] = \sum_{i=0}^{\frac{k-1}{2}} \binom{\frac{n-2}{2}}{i} \cdot (\cos^2\alpha)^i \cdot (1-\cos^2\alpha)^{\frac{n-2}{2}-i}\,,$$

i.e., the value of the distribution function of the binomial distribution with parameters $\frac{n-2}{2}$ and $\cos^2\alpha$ at $\frac{k-1}{2}$. For the case $n$ and $k$ even and for the case $n$ odd this does not hold as nice as in the above case. But for the asymptotics this is not so important so that the claimed statement follows from the well-known asymptotics stated in Proposition A.1.2. □

## A.2 Intrinsic volumes of tubes

In this section we will compute the intrinsic volumes of tubes of convex sets, assuming that these tubes are again convex.

**Proposition A.2.1.** *Let $K \in \mathcal{K}^c(S^{n-1})$ and $\alpha > 0$ such that the tube $K_\alpha := \mathcal{T}(K,\alpha) \in \mathcal{K}^c(S^{n-1})$ is still convex. Then the intrinsic volumes of the tube $K_\alpha$ are given by*

$$V_{n-m-1}(K_\alpha) = \binom{(n-2)/2}{(m-1)/2} \sum_{i,j=0}^{n-2} d_{ij}^{nm} \cdot \begin{bmatrix} n-2 \\ j \end{bmatrix} \cdot \cos(\alpha)^i \cdot \sin(\alpha)^{n-2-i} \cdot V_j(K)\,,$$

*where $1 \le m \le n-1$, $d_{ij}^{nm}$ as defined in (6.5) in Theorem 6.1.1.*

The coefficients $d_{ij}^{nm}$ arose from the polynomial identity (cf. proof of Proposition 6.4.6)

$$(X-Y)^{m-1}\cdot(1+XY)^{n-m-1} = \sum_{i,j=0}^{n-2} \binom{n-2}{j} \cdot d_{ij}^{nm} \cdot X^{n-2-j} \cdot Y^{n-2-i}\,.$$

In the following lemma we will show that they also appear in a slightly different polynomial identity.

**Lemma A.2.2.** *For $n \ge 2$, $0 \le j \le n-2$, and $1 \le m \le n-1$ we have the following identity of formal polynomials*

$$(1-XY)^j \cdot (X+Y)^{n-2-j} = \sum_{m=1}^{n-1}\sum_{i=0}^{n-2} \binom{n-2}{m-1} \cdot d_{ij}^{nm} \cdot X^{n-2-i} \cdot Y^{m-1}\,. \qquad \text{(A.1)}$$

*Proof.* Expanding the terms yields

$$
\begin{aligned}
&(1-XY)^j \cdot (X+Y)^{n-2-j} \\
&= \left( \sum_{k=0}^{j} \binom{j}{k} \cdot (-1)^{j-k} \cdot X^{j-k} \cdot Y^{j-k} \right) \cdot \left( \sum_{\ell=0}^{n-2-j} \binom{n-2-j}{\ell} \cdot X^{n-2-j-\ell} \cdot Y^{\ell} \right) \\
&\overset{(*)}{=} \sum_{i,m} \binom{j}{\frac{i+j}{2} - \frac{m-1}{2}} \cdot \binom{n-2-j}{\frac{i-j}{2} + \frac{m-1}{2}} \cdot (-1)^{\frac{i-j}{2} - \frac{m-1}{2}} \cdot X^{n-2-i} \cdot Y^{m-1} \,,
\end{aligned}
$$

where in $(*)$ we have used the substitutions $k = \frac{i+j}{2} - \frac{m-1}{2}$ and $\ell = \frac{i-j}{2} + \frac{m-1}{2}$, and where the last summation runs over all $0 \le i \le n-2$, $1 \le m \le n-1$, which satisfy $i + j + m \equiv 1 \pmod 2$. We finally compute

$$
\begin{aligned}
\frac{\binom{n-2}{m-1} \cdot \binom{m-1}{\frac{i-j}{2} + \frac{m-1}{2}} \cdot \binom{n-m-1}{\frac{i+j}{2} - \frac{m-1}{2}}}{\binom{n-2}{j}} &= \frac{(n-2)!}{(m-1)! \cdot (n-m-1)!} \cdot \frac{j! \cdot (n-2-j)!}{(n-2)!} \\
&\quad \cdot \frac{(m-1)!}{(\frac{i-j}{2} + \frac{m-1}{2})! \cdot (\frac{m-1}{2} - \frac{i-j}{2})!} \cdot \frac{(n-m-1)!}{(\frac{i+j}{2} - \frac{m-1}{2})! \cdot (n-2-\frac{m-1}{2} - \frac{i+j}{2})!} \\
&= \frac{j!}{(\frac{i+j}{2} - \frac{m-1}{2})! \cdot (\frac{m-1}{2} - \frac{i-j}{2})!} \cdot \frac{(n-2-j)!}{(\frac{i-j}{2} + \frac{m-1}{2})! \cdot (n-2-\frac{m-1}{2} - \frac{i+j}{2})!} \\
&= \binom{j}{\frac{i+j}{2} - \frac{m-1}{2}} \cdot \binom{n-2-j}{\frac{i-j}{2} + \frac{m-1}{2}} \,.
\end{aligned}
$$

This finishes the proof. □

For the following lemma recall that $I_{n,j}(\alpha) = \int_0^\alpha \cos(\rho)^j \cdot \sin(\rho)^{n-2-j}\, d\rho$.

**Lemma A.2.3.** *For $n \ge 2$, $0 \le j \le n-2$, and $\alpha, \beta \in \mathbb{R}$ we have*

$$
I_{n,j}(\alpha+\beta) = I_{n,j}(\alpha) + \sum_{m=1}^{n-1} \binom{n-2}{m-1} \cdot I_{n,n-m-1}(\beta) \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot \cos(\alpha)^i \cdot \sin(\alpha)^{n-2-i} \,. \tag{A.2}
$$

*Proof.* Using the addition theorem for sin and cos, we compute

$$
\begin{aligned}
&\cos(\alpha+\rho)^j \cdot \sin(\alpha+\rho)^{n-2-j} \\
&= (\cos(\alpha)\cdot\cos(\rho) - \sin(\alpha)\cdot\sin(\rho))^j \cdot (\sin(\alpha)\cdot\cos(\rho) + \cos(\alpha)\cdot\sin(\rho))^{n-2-j} \\
&= \cos(\alpha)^{n-2} \cdot \cos(\rho)^{n-2} \cdot (1 - \tan(\alpha)\tan(\rho))^j \cdot (\tan(\alpha) + \tan(\rho))^{n-2-j} \\
&\overset{(A.1)}{=} \cos(\alpha)^{n-2} \cdot \cos(\rho)^{n-2} \cdot \sum_{m=1}^{n-1} \sum_{i=0}^{n-2} \binom{n-2}{m-1} \cdot d_{ij}^{nm} \cdot \tan(\alpha)^{n-2-i} \cdot \tan(\rho)^{m-1} \\
&= \sum_{m=1}^{n-1} \binom{n-2}{m-1} \cdot \sin(\rho)^{m-1} \cdot \cos(\rho)^{n-m-1} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot \cos(\alpha)^i \cdot \sin(\alpha)^{n-2-i} \,.
\end{aligned}
$$

From this we get

$$
\begin{aligned}
I_{n,j}(\alpha+\beta) &= I_{n,j}(\alpha) + \int_0^\beta \cos(\alpha+\rho)^j \cdot \sin(\alpha+\rho)^{n-2-j}\, d\rho \\
&= I_{n,j}(\alpha) + \sum_{m=1}^{n-1} \binom{n-2}{m-1} \cdot I_{n,n-m-1}(\beta) \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot \cos(\alpha)^i \cdot \sin(\alpha)^{n-2-i} \,.
\end{aligned}
$$

□

*Proof of Proposition A.2.1.* The volume of the tube around $K_\alpha$ is given by

$$\begin{aligned}
\operatorname{vol}\mathcal{T}(K_\alpha,\beta) &= \operatorname{vol}\mathcal{T}(K,\alpha+\beta)\\
&\overset{(4.39)}{=} \operatorname{vol}K + \sum_{j=0}^{n-2} I_{n,j}(\alpha+\beta)\cdot\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}\cdot V_j(K)\\
&\overset{(A.2)}{=} \operatorname{vol}K + \sum_{j=0}^{n-2}\Big(I_{n,j}(\alpha) + \sum_{m=1}^{n-1}\binom{n-2}{m-1}\cdot I_{n,n-m-1}(\beta)\\
&\qquad\qquad \cdot\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\Big)\cdot\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}\cdot V_j(K)\\
&\overset{(4.39)}{=} \operatorname{vol}K_\alpha + \sum_{m=1}^{n-1} I_{n,n-m-1}(\beta)\cdot\sum_{i,j=0}^{n-2} d_{ij}^{nm}\cdot\binom{n-2}{m-1}\cdot\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}\\
&\qquad\qquad \cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\cdot V_j(K)\,.
\end{aligned}$$

Therefore, again by (4.39), we have

$$\begin{aligned}
V_{n-m-1}(K_\alpha) &= \sum_{i,j=0}^{n-2} d_{ij}^{nm}\cdot\frac{\binom{n-2}{m-1}\cdot\mathcal{O}_j\cdot\mathcal{O}_{n-2-j}}{\mathcal{O}_{m-1}\cdot\mathcal{O}_{n-m-1}}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\cdot V_j(K)\\
&\overset{(4.24),(4.22)}{=} \sum_{i,j=0}^{n-2} d_{ij}^{nm}\cdot\binom{(n-2)/2}{(m-1)/2}\cdot\begin{bmatrix}n-2\\ j\end{bmatrix}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\cdot V_j(K)\,.
\end{aligned}$$

This finishes the proof. □

## A.3 On the twisted $I$-functions

In this section we will fill the gap in the proof of the equality case (6.6) in Theorem 6.1.1, that we only proved for $K\in\mathcal{K}^{\mathrm{sm}}(S^{n-1})$ in Section 6.5. So far we used a differential geometric approach for the proof of Theorem 6.1.1, but now we will use a more general convex geometric approach. The main idea is to use the kinematic formula, resp. Corollary 4.4.12, to derive an alternative expression for $\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha)$. More precisely, let $K\in\mathcal{K}^{\mathrm{r}}(S^{n-1})$, and $\Sigma_m:=\Sigma_m(K)$. Then for $\mathcal{W}\in\operatorname{Gr}_{n,m}$ we have

$$\mathcal{W}\in\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha)\iff\mathcal{W}\cap\mathcal{T}(K,\alpha)\neq\emptyset \text{ and } \mathcal{W}\cap\operatorname{int}(K)=\emptyset\,.$$

If the tube $K_\alpha:=\mathcal{T}(K,\alpha)$ is convex, then we can use Corollary 4.4.12 to compute the probability of the right-hand side, assuming that $\mathcal{W}\in\operatorname{Gr}_{n,m}$ is chosen uniformly at random. Note that this coincides with the relative volume of $\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha)$. The result will depend on the intrinsic volumes of $K_\alpha$, which we have computed in Proposition A.2.1. Unfortunately, we will not get the formula (6.6) directly, but only via a short detour over the twisted $I$-functions, by which we mean the functions

$$\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha)\,.$$

These functions appear in the tube formula (6.6) and can be interpreted as a substitute of the $I$-functions from the spherical tube formula (4.39).

The overall argumentation is as follows. The kinematic formula yields an alternative expression of $\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha)$ that holds for all regular cones. Comparing this with the formula we gave in (6.6) yields an alternative expression for the twisted $I$-functions. Replacing the corresponding expression in the formula that we derived from the kinematic formula then shows that (6.6) indeed holds for all $K \in \mathcal{K}^r(S^{n-1})$.

This argumentation is a little ugly, but it yields the desired result. It also shows that

1. the discontinuity of the map $\alpha_{\max}$ (cf. Remark 3.1.17) has a lot of negative consequences,

2. the case where the tube $K_\alpha = \mathcal{T}(K,\alpha)$ is convex can be handled by quite different methods.

**Proposition A.3.1.** *Let the setting be as in Theorem 6.1.1. Then for* $0 \le \alpha \le \alpha_0$

$$\operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) = 2\cdot \sum_{j=0}^{n-2} V_j(K)\Bigg( -\delta(n,m,j) + \delta_{\mathrm{odd}}(m)\cdot \frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}}$$
$$+ \sum_{\substack{k=n-m\\ k\equiv n-m\\ \pmod 2}}^{n-2} \binom{(n-2)/2}{k/2} \cdot \sum_{i=0}^{n-2} d_{ij}^{n,n-k-1}\cdot \begin{bmatrix} n-2\\ j\end{bmatrix}\cdot \cos(\alpha)^i\cdot \sin(\alpha)^{n-2-i}\Bigg) . \quad (A.3)$$

*where* $\delta_{\mathrm{odd}}(m) = 1$ *if* $m$ *is odd and* $\delta_{\mathrm{odd}}(m) = 0$ *if* $m$ *is even, and* $\delta(n,m,j) = 1$ *if* $j \in \{n-m, n-m+2, \ldots\}$ *and* $\delta(n,m,j) = 0$ *else.*

*Proof.* Recall from Corollary 4.4.12 that for $\mathcal{W} \in \mathrm{Gr}_{n,m}$ chosen uniformly at random we have

$$\operatorname{Prob}[K\cap\mathcal{W}\neq\emptyset] = 2\cdot \sum_{\substack{k=n-m\\ k\equiv n-m\\ \pmod 2}}^{n-1} V_k(K) .$$

If $K \in \mathcal{K}(S^{n-1})$ is such that $K_\alpha := \mathcal{T}(K,\alpha) \in \mathcal{K}(S^{n-1})$ we thus get

$$\begin{aligned}
&\operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) = \operatorname{Prob}[\mathcal{W}\cap K_\alpha\neq\emptyset] - \operatorname{Prob}[\mathcal{W}\cap K\neq\emptyset]\\
&= 2\cdot \sum_{\substack{k=n-m\\ k\equiv n-m\\ \pmod 2}}^{n-2} (V_k(K_\alpha)-V_k(K)) + 2\cdot\delta_{\mathrm{odd}}(m)\cdot \tfrac{1}{\mathcal{O}_{n-1}}\cdot(\operatorname{vol} K_\alpha - \operatorname{vol}(K))\\
&\overset{(4.39)}{=} 2\cdot \sum_{\substack{k=n-m\\ k\equiv n-m\\ \pmod 2}}^{n-2} (V_k(K_\alpha)-V_k(K)) + 2\cdot\delta_{\mathrm{odd}}(m)\cdot \sum_{j=0}^{n-2} \frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}}\cdot V_j(K) . \quad (A.4)
\end{aligned}$$

By Proposition A.2.1 we have for $0 \le k \le n-2$

$$V_k(K_\alpha) = \binom{(n-2)/2}{k/2}\cdot \sum_{i,j=0}^{n-2} d_{ij}^{n,n-k-1}\cdot \begin{bmatrix} n-2\\ j\end{bmatrix}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\cdot V_j(K) .$$

From (A.4) we thus get (after a small rearrangement of the summation terms)

$$\operatorname{rvol} \mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) = 2\cdot\sum_{j=0}^{n-2} V_j(K)\Bigg( -\delta(n,m,j)+\delta_{\mathrm{odd}}(m)\cdot\frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}} \\ + \sum_{\substack{k=n-m\\ k\equiv n-m\\ (\mathrm{mod}\ 2)}}^{n-2} \binom{(n-2)/2}{k/2}\cdot\sum_{i=0}^{n-2} d_{ij}^{n,n-k-1}\cdot\begin{bmatrix} n-2\\ j\end{bmatrix}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i}\Bigg)\,.$$

This finishes the proof. □

**Corollary A.3.2.** *For $1\le m\le n-1$, $0\le j\le n-2$, and $\alpha\in[0,\frac{\pi}{2}]$ we have*

$$\frac{m(n-m)}{n}\cdot\binom{n/2}{m/2}\cdot\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha) = \sum_{i=0}^{n-2}\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i} \\ \cdot\sum_{\substack{\ell=1\\ \ell\equiv m-1\\ (\mathrm{mod}\ 2)}}^{m-1}\binom{(n-2)/2}{(\ell-1)/2}\cdot d_{ij}^{n\ell} - \frac{\delta(n,m,j)}{\left[\begin{smallmatrix} n-2\\ j\end{smallmatrix}\right]} + \frac{\delta_{\mathrm{odd}}(m)}{\left[\begin{smallmatrix} n-2\\ j\end{smallmatrix}\right]}\cdot\frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}}\,,$$

*where $d_{ij}^{nm}$ defined as in Theorem 6.1.1, and $\delta_{\mathrm{odd}}$ and $\delta$ as in Proposition A.3.1.*

*Proof.* Let $K=B(z,\beta)$, with $0<\beta\le\frac{\pi}{2}$, be a circular cap. For this choice of $K$ we have $\alpha_0=\frac{\pi}{2}-\beta$ and $K_\alpha=B(z,\alpha+\beta)$). For $0\le j\le n-2$ we have

$$V_j(K) = \binom{(n-2)/2}{j/2}\cdot\frac{\cos(\beta)^{n-2-j}\cdot\sin(\beta)^j}{2}$$

(cf. Example 4.4.8). By (A.3) we thus have for $0\le\alpha\le\frac{\pi}{2}-\beta$

$$\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) = \sum_{j=0}^{n-2}\binom{(n-2)/2}{j/2}\cdot\cos(\beta)^{n-2-j}\cdot\sin(\beta)^j \\ \cdot\Bigg(\sum_{\substack{k=n-m\\ k\equiv n-m\\ (\mathrm{mod}\ 2)}}^{n-2}\sum_{i=0}^{n-2} d_{ij}^{n,n-k-1}\cdot\binom{(n-2)/2}{k/2}\cdot\begin{bmatrix} n-2\\ j\end{bmatrix}\cdot\cos(\alpha)^i\cdot\sin(\alpha)^{n-2-i} \\ -\delta(n,m,j)+\delta_{\mathrm{odd}}(m)\cdot\frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}}\Bigg)\,. \tag{A.5}$$

On the other hand, by (6.6) in Theorem 6.1.1 we have

$$\operatorname{rvol}\mathcal{T}^{\mathrm{P}}(\Sigma_m,\alpha) = \frac{2m(n-m)}{n}\binom{n/2}{m/2}\cdot\sum_{j=0}^{n-2} V_j(K)\cdot\begin{bmatrix} n-2\\ j\end{bmatrix}\cdot\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha) \\ = \sum_{j=0}^{n-2}\binom{(n-2)/2}{j/2}\cdot\cos(\beta)^{n-2-j}\cdot\sin(\beta)^j\cdot\frac{m(n-m)}{n}\cdot\binom{n/2}{m/2}\cdot\begin{bmatrix} n-2\\ j\end{bmatrix} \\ \cdot\sum_{i=0}^{n-2} d_{ij}^{nm}\cdot I_{n,i}(\alpha)\,. \tag{A.6}$$

Using that the functions $\cos(\beta)^{n-2-j} \cdot \sin(\beta)^j$, $0 \le j \le n-2$, are linearly independent, a comparison of the coefficients between (A.5) and (A.6) yields

$$\begin{aligned}
&\frac{m(n-m)}{n} \cdot \binom{n/2}{m/2} \cdot \sum_{i=0}^{n-2} d_{ij}^{nm} \cdot I_{n,i}(\alpha) \\
&= \sum_{i=0}^{n-2} \cos(\alpha)^i \cdot \sin(\alpha)^{n-2-i} \cdot \sum_{\substack{k=n-m \\ k\equiv n-m \\ (\mathrm{mod}\ 2)}}^{n-2} \binom{(n-2)/2}{k/2} \cdot d_{ij}^{n,n-k-1} \\
&\qquad - \frac{\delta(n,m,j)}{\left[\begin{smallmatrix} n-2 \\ j \end{smallmatrix}\right]} + \frac{\delta_{\mathrm{odd}}(m)}{\left[\begin{smallmatrix} n-2 \\ j \end{smallmatrix}\right]} \cdot \frac{\mathcal{O}_{n-1,j}(\alpha)}{\mathcal{O}_{n-1}} \,.
\end{aligned}$$

This all holds for $0 < \beta \le \frac{\pi}{2}$ and $0 \le \alpha \le \frac{\pi}{2} - \beta$, and therefore for all $0 \le \alpha \le \frac{\pi}{2}$. □

**Corollary A.3.3.** *The formula* (6.6) *in Theorem 6.1.1 holds for all* $K \in \mathcal{K}^r(S^{n-1})$.

*Proof.* The formula for rvol $\mathcal{T}(\Sigma_m(K), \alpha)$ given in (A.3) in Proposition A.3.1 holds for all $K \in \mathcal{K}^r(S^{n-1})$. In Corollary A.3.2 we have seen that the difference between (A.3) and the formula (6.6) given in Theorem 6.1.1 consists of an alternative representation of the twisted $I$-functions. Interchanging these different representations transfers (A.3) into (6.6). The validity of (A.3) for all $K \in \mathcal{K}^r(S^{n-1})$ thus also transfers to (6.6). □

# Appendix B

# Some computation rules for intrinsic volumes

In this chapter we will provide the proofs for Proposition 4.4.13 and Proposition 4.4.18.

## B.1 Spherical products

We begin with the spherical statement of Proposition 4.4.13. Let $K_1 \in \mathcal{K}(S^{n_1-1})$, $K_2 \in \mathcal{K}(S^{n_2-1})$, and let (cf. Section 3.1.1)

$$K := K_1 \circledast K_2 = (C_1 \times C_2) \cap S^{n-1} ,$$

where $C_i := \mathrm{cone}(K_i) \subset \mathbb{R}^{n_i}$, $i = 1, 2$, and $n := n_1 + n_2$. Recall that the intrinsic volume polynomial of $K$ is defined via (cf. (4.45))

$$V(K; X) = V_{-1}(K) + V_0(K) \cdot X + \ldots + V_{n-1}(K) \cdot X^n .$$

The spherical statement of Proposition 4.4.13 says that the intrinsic volume polynomial of $K$ is given by the product of the intrinsic volume polynomials of $K_1$ and $K_2$, i.e.,

$$V(K; X) = V(K_1; X) \cdot V(K_2; X) . \tag{B.1}$$

We thus have to show that for $-1 \leq j \leq n-1$

$$V_j(K) = \sum_{k=0}^{j+1} V_{k-1}(K_1) \cdot V_{j-k}(K_2) . \tag{B.2}$$

The first step in the proof is to show that it suffices to prove (B.1) for polyhedral $K_1$ and $K_2$. Indeed, recall that the set of polyhedral convex sets $\mathcal{K}^p(S^{n-1})$ lies dense in $\mathcal{K}(S^{n-1})$ (cf. Section 3.3). If $K_i \in \mathcal{K}(S^{n_i-1})$ is approximated by the sequence $(K_{i,j})_j$ in $\mathcal{K}^p(S^{n_i-1})$, $i = 1, 2$, then

$$\lim_{j\to\infty} (\underbrace{K_{1,j} \circledast K_{2,j}}_{=:K_j}) \overset{(*)}{=} \lim_{j\to\infty} K_{1,j} \circledast \lim_{j\to\infty} K_{2,j} = K_1 \circledast K_2 = K ,$$

where the step $(*)$ follows from the continuity of the map

$$\mathcal{K}(S^{n_1-1}) \times \mathcal{K}(S^{n_2-1}) \to \mathcal{K}(S^{n-1}) , \quad (K_1, K_2) \mapsto K_1 \circledast K_2 ,$$

which is verified easily. As the intrinsic volumes are continuous, we get

$$\begin{aligned} V(K;X) &= \lim_{j\to\infty} V(K_j;X) \overset{(**)}{=} \lim_{j\to\infty} \left(V(K_{1,j};X)\cdot V(K_{2,j};X)\right) \\ &= \lim_{j\to\infty} V(K_{1,j};X)\cdot \lim_{j\to\infty} V(K_{2,j};X) = V(K_1;X)\cdot V(K_2;X)\ , \end{aligned}$$

where $(**)$ follows, if we have shown (B.1), resp. (B.2), for polyhedral $K_1$, $K_2$.

We will deduce (B.2) for polyhedral $K_1$, $K_2$ from the characterization of the intrinsic volumes in Proposition 4.4.6. Before we can do this we have to recall some well-known facts about polyhedral cones.

Let $C_1 \subseteq \mathbb{R}^{n_1}$ and $C_2 \subseteq \mathbb{R}^{n_1}$, and thus also $C = C_1 \times C_2 \subseteq \mathbb{R}^n$ be polyhedral cones. If $\Pi_{C_1}, \Pi_{C_2}, \Pi_C$ denote the projection maps onto $C_1, C_2, C$, respectively, we get for $x = (x_1, x_2) \in \mathbb{R}^n$

$$\Pi_C(x) = (\Pi_{C_1}(x_1), \Pi_{C_2}(x_2))\ .$$

Moreover, if $\mathcal{F}(C)$ denotes the set of faces of $C$, then we have

$$\mathcal{F}(C) = \{F_1 \times F_2 \mid F_1 \in \mathcal{F}(C_1)\ ,\ F_2 \in \mathcal{F}(C_2)\}\ .$$

Recall that for $x \in C$ we have defined (cf. (4.41))

$$\operatorname{face}(x) := \begin{cases} C & \text{if } x \in \operatorname{int}(C) \\ F & \text{if } x \in \operatorname{relint}(F) \text{ and } F \text{ a face of } C. \end{cases}$$

Let $\operatorname{face}_1$ and $\operatorname{face}_2$ denote the corresponding functions with $C$ being replaced by $C_1$ and $C_2$, respectively. Then for $x = (x_1, x_2) \in C = C_1 \times C_2$, we have

$$\operatorname{face}(x) = \operatorname{face}_1(x_1) \times \operatorname{face}_2(x_2)\ . \tag{B.3}$$

Recall also that we have defined the function

$$d_C\colon \mathbb{R}^n \to \{0,1,2,\ldots,n\}\ ,\quad x \mapsto \dim(\operatorname{face}(\Pi_C(x)))\ .$$

As $C$ is a cone, we have $d_C(\lambda\, x) = d_C(x)$ for all $x \in \mathbb{R}^n$, $\lambda > 0$. From (B.3) we get for $x = (x_1, x_2) \in \mathbb{R}^n$

$$d_C(x) = d_{C_1}(x_1) + d_{C_2}(x_2)\ . \tag{B.4}$$

In Proposition 4.4.6 we have shown that the $j$th intrinsic volume of $K = C \cap S^{n-1}$ is given by

$$V_j(K) = \operatorname*{Prob}_{p \in S^{n-1}} [d_C(p) = j+1]\ ,$$

where $p \in S^{n-1}$ is drawn uniformly at random. If $x \in \mathbb{R}^n$ is a normal distributed variable $x \sim \mathcal{N}(0, I_n)$, then with probability 1 we have $x \neq 0$. Moreover, the induced probability distribution on $S^{n-1}$ via $x \mapsto \|x\|^{-1} \cdot x$ is the uniform distribution. As $d_C(\lambda\, x) = d_C(x)$ for all $x \in \mathbb{R}^n$, $\lambda > 0$, we get

$$V_j(K) = \operatorname*{Prob}_{x \sim \mathcal{N}(0, I_n)} [d_C(x) = j+1]\ . \tag{B.5}$$

The normal distribution has the pleasing fact that

$$x = (x_1, x_2) \sim \mathcal{N}(0, I_n) \iff x_1 \sim \mathcal{N}(0, I_{n_1})\ ,\ x_2 \sim \mathcal{N}(0, I_{n_2})\ .$$

Therefore, we get

$$V_j(K) \overset{(B.5)}{=} \underset{\substack{x_1\sim\mathcal{N}(0,I_{n_1})\\ x_2\sim\mathcal{N}(0,I_{n_2})}}{\operatorname{Prob}} [d_C(x_1,x_2)=j+1] \overset{(B.4)}{=} \underset{x_1,x_2}{\operatorname{Prob}} [d_{C_1}(x_1)+d_{C_2}(x_2)=j+1]$$

$$= \sum_{k=0}^{j+1} \underset{x_1}{\operatorname{Prob}} [d_{C_1}(x_1)=k] \cdot \underset{x_2}{\operatorname{Prob}} [d_{C_2}(x_2)=j+1-k]$$

$$\overset{(B.5)}{=} \sum_{k=0}^{j+1} V_{k-1}(K_1)\cdot V_{j-k}(K_2) \ .$$

This shows (B.2) for polyhedral $K_1$, $K_2$, and thus finishes the proof. □

## B.2 Euclidean products

In this section we will give the proof of the euclidean statement of Proposition 4.4.13. We need to show that for euclidean convex bodies $K_1 \in \mathcal{K}(\mathbb{R}^{n_1})$, $K_2 \in \mathcal{K}(\mathbb{R}^{n_2})$, and $K := K_1 \times K_2 \in \mathcal{K}(\mathbb{R}^n)$, $n = n_1+n_2$, we have

$$V^e(K;X) = V^e(K_1;X)\cdot V^e(K_2;X) \ ,$$

where $V^e(K;X)$ denotes the euclidean intrinsic volume polynomial (cf. (4.48))

$$V^e(K;X) = \sum_{j=0}^{n} V_j^e(K)\cdot X^j \ .$$

So we need to show that for $0 \le j \le n$

$$V_j^e(K) = \sum_{k=0}^{j} V_k^e(K_1)\cdot V_{j-k}^e(K_1) \ . \tag{B.6}$$

In contrast to the spherical case, the euclidean intrinsic volumes do not form a probability distribution. Moreover, we do not have a characterization of the euclidean intrinsic volumes similar to the characterization of the spherical intrinsic volumes from Proposition 4.4.6. The proof of (B.6) will therefore be completely different to the proof of (B.2). Instead of arguing over polyhedral convex sets, we will argue over smooth convex bodies.

Let $K \in \mathcal{K}(\mathbb{R}^n)$ such that $\operatorname{int}(K) \neq \emptyset$, and such that the boundary $M := \partial K$ is smooth and oriented such that the normal direction points inwards the convex body. Then from (the euclidean version of) Weyl's tube formula in Theorem 4.3.2 we get

$$\operatorname{vol}_n \mathcal{T}^e(K,r) = \operatorname{vol}_n K + \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \int_{x\in M} \sigma_{n-1-j}(x)\,dx \ .$$

Comparing this formula with the Steiner polynomial

$$\operatorname{vol}_n \mathcal{T}^e(K,r) = \sum_{i=0}^{n} \omega_i \cdot V_{n-i}^e(K)\cdot r^i \ . \tag{B.7}$$

yields $V_n^e(K) = \operatorname{vol}_n K$, and for $0 \le j \le n-1$

$$V_j^e(K) = \frac{1}{(n-j)\cdot \omega_{n-j}} \cdot \int_{x\in M} \sigma_{n-1-j}(x)\,dx \ . \tag{B.8}$$

Using an approximation argument similar to the spherical case, we may assume for the proof of (B.6) w.l.o.g. that $K_1$ and $K_2$ have nonempty interiors, and the boundaries of $K_1$ and $K_2$ are smooth. We denote these by

$$M_1 := \partial K_1 \ , \quad M_2 := \partial K_2 \ .$$

Furthermore, let $\nu_i\colon M_i \to \mathbb{R}^{n_i}$ denote the unit normal field of $M_i$ such that for $x_i \in M_i$ the vector $\nu_i(x_i)$ points inwards the convex body $K_i$, $i = 1, 2$.

The boundary of the product $K = K_1 \times K_2$ decomposes via

$$\partial K = (\mathrm{int}(K_1) \times M_2) \,\dot\cup\, (M_1 \times \mathrm{int}(K_2)) \,\dot\cup\, (M_1 \times M_2) \ .$$

For $x = (x_1, x_2) \in \partial K$ the normal cone of $K$ at $x$ is thus given by

$$N_x(K) = \begin{cases} \{0\} \times \mathbb{R}_- \cdot \nu_2(x_2) & \text{if } x \in \mathrm{int}(K_1) \times M_2 \\ \mathbb{R}_- \cdot \nu_1(x_1) \times \{0\} & \text{if } x \in M_1 \times \mathrm{int}(K_2) \\ \mathbb{R}_- \cdot \nu_1(x_1) \times \mathbb{R}_- \cdot \nu_2(x_2) & \text{if } x \in M_1 \times M_2 \ , \end{cases} \tag{B.9}$$

where $\mathbb{R}_- = \{r \in \mathbb{R} \mid r \leq 0\}$. Let us denote

$$\begin{aligned} M_0^\times &:= \mathrm{int}(K) = \mathrm{int}(K_1) \times \mathrm{int}(K_2) \\ M_1^\times &:= \mathrm{int}(K_1) \times M_2 \\ M_2^\times &:= M_1 \times \mathrm{int}(K_2) \\ M_3^\times &:= M_1 \times M_2 \ , \end{aligned}$$

so that $K$ decomposes into

$$K = M_0^\times \,\dot\cup\, M_1^\times \,\dot\cup\, M_2^\times \,\dot\cup\, M_3^\times \ .$$

Moreover, the pieces $M_i^\times$, $i = 0, 1, 2, 3$, are smooth manifold, and from (B.9) it is easily seen that the duality bundles $NM_i$ (cf. (3.8)) are also smooth manifolds. It follows that $K$ is a stratified convex body. Furthermore, all pieces $M_i^\times$, $i = 0, 1, 2, 3$, are essential.

From the euclidean part of Theorem 4.3.2 we thus get (note that $M_1^\times$ and $M_2^\times$ have codimension 1 in $\mathbb{R}^n$)

$$\begin{aligned} \mathrm{vol}_n \mathcal{T}^e(K, r) = \mathrm{vol}_n K &+ \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \int_{x \in M_1^\times} \sigma_{n-1-j}^{(\times,1)}(x)\, dx \\ &+ \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \int_{x \in M_2^\times} \sigma_{n-1-j}^{(\times,2)}(x)\, dx \\ &+ \sum_{j=0}^{n-2} \frac{r^{n-j}}{n-j} \cdot \int_{x \in M_3^\times} \int_{\eta \in N_x^S} \sigma_{n-2-j}^{(\times,3)}(x, -\eta)\, d\eta\, dx \ , \end{aligned} \tag{B.10}$$

where the notation $\sigma^{(\times,i)}$ shall indicate the dependence on $M_i^\times$.

The volume of $K$ is obviously given by

$$\mathrm{vol}_n K = \mathrm{vol}_{n_1} K_1 \cdot \mathrm{vol}_{n_2} K_2 = V_{n_1}^e(K_1) \cdot V_{n_2}^e(K_2) \ . \tag{B.11}$$

As for the principal curvatures of the hypersurfaces $M_1^\times$ and $M_2^\times$ in $\mathbb{R}^n$, it is easily seen that these are given by

$$\begin{aligned} x=(x_1,x_2)\in M_1^\times \ &: \ \underbrace{0,\ldots,0}_{n_1\text{-times}},\ \kappa_1^{(2)}(x_2),\ldots,\kappa_{n_2-1}^{(2)}(x_2)\ ,\\ x=(x_1,x_2)\in M_2^\times \ &: \ \kappa_1^{(1)}(x_1),\ldots,\kappa_{n_1-1}^{(1)}(x_1)\ ,\ \underbrace{0,\ldots,0}_{n_2\text{-times}}\ , \end{aligned}$$

where $\kappa^{(1)}$ and $\kappa^{(2)}$ shall denote the corresponding principal curvatures of $M_1$ and $M_2$, respectively. For the elementary symmetric polynomials in the principal curvatures we thus get for $x=(x_1,x_2)\in M_1^\times$ and $n-1-j>n_2-1$

$$\sigma_{n-1-j}^{(\times,1)}(x)=0\ ,$$

and therefore

$$\int\limits_{x\in M_1^\times}\sigma_{n-1-j}^{(\times,1)}(x)\,dx=0\ . \tag{B.12}$$

For $n-1-j\le n_2-1$ we get

$$\begin{aligned} \sigma_{n-1-j}^{(\times,1)}(x)&=\sigma_{n-1-j}(\underbrace{0,\ldots,0}_{n_1\text{-times}},\ \kappa_1^{(2)}(x_2),\ldots,\kappa_{n_2-1}^{(2)}(x_2))\\ &=\sigma_{n-1-j}(\kappa_1^{(2)}(x_2),\ldots,\kappa_{n_2-1}^{(2)}(x_2))\\ &=\sigma_{n-1-j}^{(2)}(x_2)\ , \end{aligned}$$

where the notation $\sigma^{(1)}$ shall indicate the dependence on $M_1$. From this we get

$$\begin{aligned} \int\limits_{x\in M_1^\times}\sigma_{n-1-j}^{(\times,1)}(x)\,dx&=\int\limits_{x_1\in\mathrm{int}(K_1)}\ \int\limits_{x_2\in M_2}\sigma_{n-1-j}^{(2)}(x_2)\,dx\\ &=\mathrm{vol}_{n_1}K_1\cdot\int\limits_{x_2\in M_2}\sigma_{n-1-j}^{(2)}(x_2)\,dx\\ &\overset{(B.8)}{=}V_{n_1}^e(K_1)\cdot(n-j)\cdot\omega_{n-j}\cdot V_{j-n_1}^e(K_2)\ . \end{aligned} \tag{B.13}$$

Similarly, we get

$$\int\limits_{x\in M_2^\times}\sigma_{n-1-j}^{(\times,2)}(x)\,dx=\begin{cases}0 & \text{if } n-1-j>n_1\\ (n-j)\cdot\omega_{n-j}\cdot V_{j-n_2}^e(K_1)\cdot V_{n_2}^e(K_2) & \text{if } n-1-j\le n_1\ .\end{cases} \tag{B.14}$$

It remains to treat the contribution of $M_3^\times$. Note that $M_3^\times$ is a submanifold of codimension 2 in $\mathbb{R}^n$. The normal cone of $K$ at $x=(x_1,x_2)\in M_3^\times$ is the 2-dimensional product of the half-lines $\mathbb{R}_-\cdot\nu_1(x_1)$ and $\mathbb{R}_-\cdot\nu_2(x_2)$ (cf. (B.9)). As for the intersection $N_x^S(K)=N_x(K)\cap S^{n-1}$ we get

$$N_x^S(K)=\{(-\cos(\rho)\cdot\nu_1(x_1),-\sin(\rho)\cdot\nu_2(x_2))\mid 0\le\rho\le\tfrac{\pi}{2}\}\ . \tag{B.15}$$

It is straightforward to verify that the principal curvatures of $M_3^\times$ at $x=(x_1,x_2)$ in direction $(c\,\nu_1(x_1),s\,\nu_2(x_2))$ are given by

$$c\cdot\kappa_1^{(1)}(x_1),\ldots,c\cdot\kappa_{n_1-1}^{(1)}(x_1)\ ,\ s\cdot\kappa_1^{(2)}(x_2),\ldots,s\cdot\kappa_{n_2-1}^{(2)}(x_2)\ . \tag{B.16}$$

Note that the symmetric functions in these curvatures can be written in the form

$$\begin{aligned}
&\sigma_\ell(c\cdot\kappa_1^{(1)}(x_1),\dots,c\cdot\kappa_{n_1-1}^{(1)}(x_1)\,,\ s\cdot\kappa_1^{(2)}(x_2),\dots,s\cdot\kappa_{n_2-1}^{(2)}(x_2)) \qquad \text{(B.17)}\\
&\quad=\sum_{k=0}^{\ell} c^k\cdot\sigma_k\left(\kappa_1^{(1)}(x_1),\dots,\kappa_{n_1-1}^{(1)}(x_1)\right)\cdot s^{\ell-k}\cdot\sigma_{\ell-k}\left(\kappa_1^{(2)}(x_2),\dots,\kappa_{n_2-1}^{(2)}(x_2)\right)\ ,
\end{aligned}$$

with the usual convention $\sigma_N(a_1,\dots,a_n)=0$, if $N>n$. From (B.15), (B.16), and (B.17), we thus get

$$\begin{aligned}
&\int_{x\in M_3^\times}\int_{\eta\in N_x^S}\sigma_{n-2-j}^{(\times,3)}(x,-\eta)\,d\eta\,dx\\
&=\int_{x\in M_3^\times}\int_0^{\frac{\pi}{2}}\sigma_{n-2-j}^{(\times,3)}(x,(-\cos(\rho)\,\nu_1(x_1),-\sin(\rho)\,\nu_2(x_2)))\,d\rho\,dx\\
&=\sum_{k=0}^{n-2-j}\int_0^{\frac{\pi}{2}}\cos(\rho)^k\cdot\sin(\rho)^{n-2-j-k}\,d\rho\cdot\int_{x_1\in M_1}\sigma_k^{(1)}(x_1)\,dx_1\cdot\int_{x_2\in M_2}\sigma_{n-2-j-k}^{(2)}(x_2)\,dx_2\ .
\end{aligned}$$

Recall that in Corollary 4.1.19 we have seen that $\int_0^{\frac{\pi}{2}}\cos(\rho)^k\cdot\sin(\rho)^\ell\,d\rho=\frac{\mathcal{O}_{k+\ell+1}}{\mathcal{O}_k\cdot\mathcal{O}_\ell}$. We may thus continue

$$\begin{aligned}
&=\sum_{k=0}^{n-2-j}\frac{\mathcal{O}_{n-1-j}}{\mathcal{O}_k\cdot\mathcal{O}_{n-2-j-k}}\cdot\int_{x_1\in M_1}\sigma_k^{(1)}(x_1)\,dx_1\cdot\int_{x_2\in M_2}\sigma_{n-2-j-k}^{(2)}(x_2)\,dx_2\\
&\overset{\text{(B.8)}}{=}\sum_{k=0}^{n-2-j}(n-j)\cdot\omega_{n-j}\cdot V_{n_1-1-k}^e(K_1)\cdot V_{j+k+1-n_1}^e(K_2)\ . \qquad \text{(B.18)}
\end{aligned}$$

Combining (B.10), (B.11), (B.12), (B.13), (B.14), and (B.18) yields

$$\begin{aligned}
\mathrm{vol}_n\,\mathcal{T}^e(K,r)&=V_{n_1}^e(K_1)\cdot V_{n_2}^e(K_2)+\sum_{j=0}^{n-1}r^{n-j}\cdot\omega_{n-j}\cdot V_{n_1}^e(K_1)\cdot V_{j-n_1}^e(K_2)\\
&\quad+\sum_{j=0}^{n-1}r^{n-j}\cdot\omega_{n-j}\cdot V_{j-n_2}^e(K_1)\cdot V_{n_2}^e(K_2)\\
&\quad+\sum_{j=0}^{n-2}\sum_{k=0}^{n-2-j}r^{n-j}\cdot\omega_{n-j}\cdot V_{n_1-1-k}^e(K_1)\cdot V_{j+k+1-n_1}^e(K_2)\\
&=\sum_{j=0}^{n}\sum_{k=-1}^{n-1-j}r^{n-j}\cdot\omega_{n-j}\cdot V_{n_1-1-k}^e(K_1)\cdot V_{j+k+1-n_1}^e(K_2)\\
&\overset{[\ell:=n_1-1-k]}{=}\sum_{j=0}^{n}r^{n-j}\cdot\omega_{n-j}\cdot\sum_{\ell=j-n_2}^{n_1}V_\ell^e(K_1)\cdot V_{j-\ell}^e(K_2)\ .
\end{aligned}$$

Comparing this with the Steiner polynomial in (B.7) shows that indeed

$$V_j^e(K)=\sum_{\ell=j-n_2}^{n_1}V_\ell^e(K_1)\cdot V_{j-\ell}^e(K_2)\ ,$$

which was to be shown. □

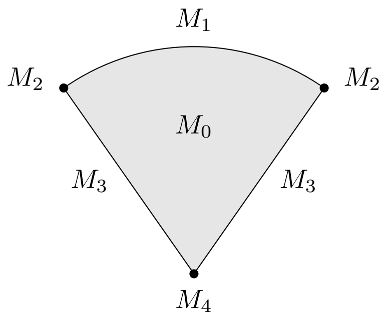



Figure B.1: The decomposition of the cone stub $K^e$.

## B.3 Euclidean vs. spherical intrinsic volumes

In this section we will give the proof of Proposition 4.4.18. Let $C \subseteq \mathbb{R}^n$ be a closed convex cone, and let

$$K^e := C \cap B_n \ , \quad K := C \cap S^{n-1} \ ,$$

where $B_n$ denotes the $n$-dimensional unit ball. The statement of Proposition 4.4.18 is that the euclidean resp. the spherical intrinsic volumes of $K^e$ and $K$ are related by the formula

$$V_j^e(K^e) = \sum_{\ell=j}^{n} \binom{\ell}{j} \cdot \frac{\omega_\ell}{\omega_{\ell-j}} \cdot V_{\ell-1}(K) \ . \tag{B.19}$$

To prove this equality, we first consider the case where $C \in \mathrm{Gr}_{n,m}$ is an $m$-dimensional subspace. In this case, the set $K$ is an $(m-1)$-dimensional subsphere of $S^{n-1}$, and $V_j(K) = \delta_{j,m-1}$, i.e., $V_{m-1}(K) = 1$ and $V_j(K) = 0$ if $j \neq m-1$. Furthermore, we have (cf. Example 4.2.1) $V_i^e(K^e) = \binom{m}{i} \cdot \frac{\omega_m}{\omega_{m-i}}$, $0 \leq i \leq m$, and $V_i^e(K^e) = 0$ if $i > m$. This shows (B.19) in the case where $C$ is a linear subspace.

It remains to prove (B.19) where $C$ is convex cone, which is not a linear subspace. Equivalently, it remains to prove the case where $K = C \cap S^{n-1}$ is a cap. By an approximation argument as in Section B.1 we may assume w.l.o.g. that $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$ is a smooth cap.

For the rest of this section let $K \in \mathcal{K}^{\mathrm{sm}}(S^{n-1})$, and let $M := \partial K$ denote the smooth boundary of $K$. Furthermore, let $K^\circ$ denote the interior of $K$ (w.r.t. the topology on $S^{n-1}$). The boundary of the cone stub $K^e$ decomposes into

$$\partial K^e = M_0 \,\dot\cup\, M_1 \,\dot\cup\, M_2 \,\dot\cup\, M_3 \,\dot\cup\, M_4 \ ,$$

where

$$M_0 := \mathrm{int}(K^e) \ , \qquad M_1 := K^\circ \ , \qquad M_2 := M \ ,$$
$$M_3 := \{\lambda \cdot p \mid \lambda \in (0,1) \ , \ p \in M\} \ , \qquad M_4 := \{0\} \ .$$

See Figure B.1 for a picture of this decomposition. The dimensions of these pieces are given by

$$\dim M_0 = n \ , \qquad \dim M_1 = n-1 \ , \qquad \dim M_2 = n-2 \ ,$$
$$\dim M_3 = n-1 \ , \qquad \dim M_4 = 0 \ .$$

The euclidean version of Weyl's tube formula yields

$$\begin{aligned}\operatorname{vol}_n \mathcal{T}^e(K^e, r) &= \operatorname{vol}_n(K^e) + \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \int\limits_{x\in M_1} \sigma^{(1)}_{n-1-j}(x)\, dx \\ &+ \sum_{j=0}^{n-2} \frac{r^{n-j}}{n-j} \cdot \int\limits_{x\in M_2} \int\limits_{\eta\in N_x^S} \sigma^{(2)}_{n-2-j}(x,-\eta)\, d\eta\, dx \\ &+ \sum_{j=0}^{n-1} \frac{r^{n-j}}{n-j} \cdot \int\limits_{x\in M_3} \sigma^{(3)}_{n-1-j}(x)\, dx + \frac{r^n}{n} \cdot \int\limits_{\eta\in N_0^S} 1\, d\eta \end{aligned} \tag{B.20}$$

The principal curvatures in $M_1$ and $M_3$ are given by

$$\begin{aligned} x \in M_1 &: 1, \ldots, 1 \\ x = \lambda\, p \in M_3 &: \frac{\kappa_1(p)}{\lambda}, \ldots, \frac{\kappa_{n-2}(p)}{\lambda}, 0, \end{aligned}$$

where $p \in M$, and $\kappa_i(p)$ denotes the $i$th principal curvature of $M$ at $p$. This follows from the fact that $M_1$ is an open subset of $S^{n-1}$, and $M_3$ is an open subset of $\partial C \setminus \{0\}$.

As for the piece $M_2$, the normal cone $N_p(K^e)$ for $p \in M_2$ is given by

$$N_p(K^e) = \mathbb{R}_+\, p + \mathbb{R}_-\, \nu(p) ,$$

where $\nu$ denotes the unit normal field on $M$, which points inwards the cap $K$. The intersection of the normal cone with the unit sphere is thus given by

$$N_p^S(K^e) = \{\sin(\rho)\, p - \cos(\rho)\, \nu(p) \mid 0 \le \rho \le \tfrac{\pi}{2}\} .$$

Note that $p \in T_p\mathbb{R}^n$ is the normal vector of the unit sphere. So the principal curvatures of $M_2$ at $p$ in direction $p$ are $1, \ldots, 1$ ($(n-2)$-times). More generally, we get that the principal curvatures of $M_2$ at $p$ in direction $s\, p - c\, \nu(p)$ are given by

$$s + c\, \kappa_1(p), \ldots, s + c\, \kappa_{n-2}(p) .$$

The value of the $\ell$th elementary symmetric function in these principal curvatures is given by

$$\begin{aligned} \sigma_\ell(s + c\,\kappa_1(p), \ldots, s + c\,\kappa_{n-2}(p)) &= \sum_{i_1<\ldots<i_\ell} (s + c\,\kappa_{i_1}(p)) \cdots (s + c\,\kappa_{i_\ell}(p)) \\ &= \sum_{i_1<\ldots<i_\ell} \sum_{k=0}^{\ell} s^{\ell-k} \cdot c^k \cdot \sigma_k(\kappa_{i_1}(p), \ldots, \kappa_{i_k}(p)) \\ &= \sum_{k=0}^{\ell} s^{\ell-k} \cdot c^k \cdot \sum_{i_1<\ldots<i_\ell} \sigma_k(\kappa_{i_1}(p), \ldots, \kappa_{i_k}(p)) \\ &\overset{(*)}{=} \sum_{k=0}^{\ell} s^{\ell-k} \cdot c^k \cdot \binom{n-2-k}{\ell-k} \cdot \sigma_k(\kappa_1(p), \ldots, \kappa_{n-2}(p)) , \end{aligned}$$

where $(*)$ follows from the following combinatorial argument. Consider a single summand of $\sigma_k(\kappa_1(p), \ldots, \kappa_{n-2}(p))$, which is the product of $k$ principal curvatures with indices, say, $j_1,\dots,j_k$. This summand appears in $\sum_{i_1<\ldots<i_\ell}\sigma_k(\kappa_{i_1}(p),\dots,\kappa_{i_k}(p))$ iff $\{j_1,\dots,j_k\}\subseteq\{i_1,\dots,i_\ell\}$. The number of choices of $i_1<\ldots<i_\ell$ for this to happen is given by the binomial coefficient $\binom{n-2-k}{\ell-k}$.

From (B.20) we thus get

$$\begin{aligned}
&\mathrm{vol}_n\,\mathcal{T}^e(K^e,r)=\mathrm{vol}_n(K^e)+\sum_{j=0}^{n-1}\frac{r^{n-j}}{n-j}\cdot\binom{n-1}{n-1-j}\cdot\mathrm{vol}_{n-1}(K)\\
&+\sum_{j=0}^{n-2}\frac{r^{n-j}}{n-j}\cdot\int_{p\in M}\int_0^{\frac{\pi}{2}}\sum_{k=0}^{n-2-j}\sin(\rho)^{n-2-j-k}\cdot\cos(\rho)^k\cdot\binom{n-2-k}{n-2-j-k}\sigma_k(p)\,d\rho\,dp\\
&+\sum_{j=1}^{n-1}\frac{r^{n-j}}{n-j}\cdot\int_0^1\int_{p\in M}\frac{\lambda^{n-2}}{\lambda^{n-1-j}}\cdot\sigma_{n-1-j}(p)\,dp\,d\lambda+\frac{r^n}{n}\cdot\mathrm{vol}_{n-1}(\breve{K})
\end{aligned}$$

(note that changing the integration over $M_3$ to an integration over $M$ yields a factor of $\lambda^{n-2}$). The volume of $K^e$ is given by

$$\mathrm{vol}_n(K^e)=\omega_n\cdot\frac{\mathrm{vol}_{n-1}(K)}{\mathcal{O}_{n-1}}=\omega_n\cdot V_{n-1}(K)\ .$$

Furthermore, we have

$$\begin{aligned}
\frac{1}{n-j}\cdot\binom{n-1}{n-1-j}\cdot\mathrm{vol}_{n-1}(K)&=\frac{(n-1)!\cdot n}{(n-j)\cdot(n-1-j)!\cdot j!}\cdot\frac{\mathcal{O}_{n-1}}{n}\cdot V_{n-1}(K)\\
&=\binom{n}{j}\cdot\omega_n\cdot V_{n-1}(K)\ ,
\end{aligned}$$

and

$$\int_0^{\frac{\pi}{2}}\sin(\rho)^{n-2-j-k}\cdot\cos(\rho)^k\,d\rho\overset{\text{Cor. 4.1.19}}{=}\frac{\mathcal{O}_{n-1-j}}{\mathcal{O}_k\cdot\mathcal{O}_{n-2-j-k}}\ ,$$

$$\int_{p\in M}\sigma_k(p)\,dp\overset{\text{Prop. 4.4.4}}{=}\mathcal{O}_k\cdot\mathcal{O}_{n-2-k}\cdot V_{n-2-k}(K)\ .$$

This implies

$$\begin{aligned}
&\frac{1}{n-j}\cdot\int_{p\in M}\int_0^{\frac{\pi}{2}}\sum_{k=0}^{n-2-j}\sin(\rho)^{n-2-j-k}\cdot\cos(\rho)^k\cdot\binom{n-2-k}{n-2-j-k}\sigma_k(p)\,d\rho\,dp\\
&\quad=\frac{1}{n-j}\cdot\sum_{k=0}^{n-2-j}\frac{\mathcal{O}_{n-1-j}}{\mathcal{O}_k\cdot\mathcal{O}_{n-2-j-k}}\cdot\binom{n-2-k}{n-2-j-k}\cdot\mathcal{O}_k\cdot\mathcal{O}_{n-2-k}\cdot V_{n-2-k}(K)\\
&\quad=\frac{\mathcal{O}_{n-1-j}}{n-j}\cdot\sum_{k=0}^{n-2-j}\frac{n-1-j-k}{\mathcal{O}_{n-2-j-k}}\cdot\frac{\mathcal{O}_{n-2-k}}{n-1-k}\\
&\qquad\qquad\cdot\frac{(n-1-k)\cdot(n-2-k)!}{(n-1-j-k)\cdot(n-2-j-k)!\cdot j!}\cdot V_{n-2-k}(K)\\
&\quad=\omega_{n-j}\cdot\sum_{k=0}^{n-2-j}\frac{\omega_{n-1-k}}{\omega_{n-1-j-k}}\cdot\binom{n-1-k}{j}\cdot V_{n-2-k}(K)\\
&\quad\overset{[\ell:=n-1-k]}{=}\omega_{n-j}\cdot\sum_{\ell=j+1}^{n-1}\frac{\omega_\ell}{\omega_{\ell-j}}\cdot\binom{\ell}{j}\cdot V_{\ell-1}(K)\ .
\end{aligned}$$

Additionally, we have

$$\int\limits_{p\in M} \sigma_{n-1-j}(p)\,dp \overset{\text{Prop. 4.4.4}}{=} \mathcal{O}_{j-1}\cdot\mathcal{O}_{n-1-j}\cdot V_{j-1}(K)\ ,$$

which implies

$$\begin{aligned}\frac{1}{n-j}\cdot\int_0^1 \int\limits_{p\in M} \lambda^{j-1}\cdot\sigma_{n-1-j}(p)\,dp\,d\lambda &= \frac{\mathcal{O}_{n-1-j}}{n-j}\cdot\frac{\mathcal{O}_{j-1}}{j}\cdot\cdot V_{j-1}(K)\\ &= \omega_{n-j}\cdot\omega_j\cdot\cdot V_{j-1}(K)\ .\end{aligned}$$

Lastly, we have

$$\frac{1}{n}\cdot\mathrm{vol}_{n-1}(\breve{K}) = \frac{\mathcal{O}_{n-1}}{n}\cdot V_{-1}(K) = \omega_n\cdot V_{-1}(K)\ .$$

Putting everything together, we get

$$\begin{aligned}\mathrm{vol}_n\,\mathcal{T}^e(K^e,r) &= \omega_n\cdot V_{n-1}(K) + \sum_{j=0}^{n-1} r^{n-j}\cdot\binom{n}{j}\cdot\omega_n\cdot V_{n-1}(K)\\ &\qquad + \sum_{j=0}^{n-2} r^{n-j}\cdot\omega_{n-j}\cdot\sum_{\ell=j+1}^{n-1}\frac{\omega_\ell}{\omega_{\ell-j}}\cdot\binom{\ell}{j}\cdot V_{\ell-1}(K)\\ &\qquad + \sum_{j=1}^{n-1} r^{n-j}\omega_{n-j}\cdot\omega_j\cdot\cdot V_{j-1}(K) + r^n\cdot\omega_n\cdot V_{-1}(K)\\ &= \sum_{j=0}^{n} r^{n-j}\cdot\binom{n}{j}\cdot\omega_n\cdot V_{n-1}(K)\\ &\qquad + \sum_{j=0}^{n-2} r^{n-j}\cdot\omega_{n-j}\cdot\sum_{\ell=j+1}^{n-1}\frac{\omega_\ell}{\omega_{\ell-j}}\cdot\binom{\ell}{j}\cdot V_{\ell-1}(K)\\ &\qquad + \sum_{j=0}^{n-1} r^{n-j}\omega_{n-j}\cdot\omega_j\cdot V_{j-1}(K)\ .\\ &= \sum_{j=0}^{n} r^{n-j}\cdot\omega_{n-j}\cdot\sum_{\ell=j}^{n}\frac{\omega_\ell}{\omega_{\ell-j}}\cdot\binom{\ell}{j}\cdot V_{\ell-1}(K)\ .\end{aligned}$$

Comparing this with the Steiner polynomial

$$\mathrm{vol}_n\,\mathcal{T}^e(K,r) = \sum_{j=0}^{n} r^{n-j}\cdot\omega_{n-j}\cdot V_j^e(K)$$

shows the equality in (B.19) and thus finishes the proof. □

# Appendix C

# The semidefinite cone

In this chapter we will compute the intrinsic volumes of the cone of positive semidefinite matrices. We regard this computation as a first step towards a full understanding of the role of the intrinsic volumes in the domain of semidefinite programming. In Section C.3 we will formulate some observations, open questions, and conjectures concerning the intrinsic volumes of the semidefinite cone.

## C.1 Preliminary: Some integrals appearing

In this section we will try to relate the integrals that will come up in the formulas of the intrinsic volumes of the semidefinite cone to the so-called Selberg integral and the Mehta integral (cf. for example [28]). We will also introduce the notation that we will use in the formulas of the intrinsic volumes of the semidefinite cone. Selberg's and Mehta's integrals can be solved exactly, which we think is also possible for the integrals from the formulas of the intrinsic volumes of the semidefinite cone. We rely in this short account mainly on the article [28]. See this and the references therein for more details about the Selberg and the Mehta integral.

Let the Vandermonde determinant be denoted by

$$\Delta(t) := \prod_{1\le i<j\le n} (t_i - t_j) \ ,$$

where $t = (t_1, \ldots, t_n)$. The Selberg integral is the following identity

$$\begin{aligned} S_n(\alpha,\beta,\gamma) &:= \int\limits_{[0,1]^n} |\Delta(t)|^{2\gamma} \cdot \prod_{i=1}^n t_i^{\alpha-1} \cdot (1-t_i)^{\beta-1} \, dt \\ &= \prod_{j=0}^{n-1} \frac{\Gamma(\alpha+j\gamma)\Gamma(\beta+j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(\alpha+\beta+(n+j-1)\gamma)\Gamma(1+\gamma)} \ , \end{aligned} \tag{C.1}$$

which holds for complex $\alpha, \beta, \gamma$ satisfying

$$\begin{aligned} &\Re(\alpha) > 0 \\ &\Re(\beta) > 0 \\ &\Re(\gamma) > -\min\left\{\frac{1}{n}, \frac{\Re(\alpha)}{n-1}, \frac{\Re(\beta)}{n-1}\right\} \ . \end{aligned}$$

The Mehta integral is given by the identity

$$M_n(\gamma) := \int_{\mathbb{R}^n} e^{-\frac{\|z\|^2}{2}} \cdot |\Delta(z)|^{2\gamma}\, dz = (2\pi)^{\frac{n}{2}} \cdot \prod_{j=1}^{n} \frac{\Gamma(1+j\gamma)}{\Gamma(1+\gamma)} ,$$

where $\Re(\gamma) > -\frac{1}{n}$, and it can be derived from the Selberg integral via (cf. [28])

$$\lim_{L\to\infty} (2L)^{n+n(n-1)\gamma} \cdot 2^{n\cdot L^2} \cdot S_n\big(\tfrac{L^2}{2}+1, \tfrac{L^2}{2}+1, \gamma\big) = (2\pi)^{\frac{n}{2}} \cdot M_n(\gamma) .$$

Setting the parameter $\gamma := \frac{1}{2}$, and restricting the integration to the positive orthant, we define

$$M_n^+ := \int_{\mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}} \cdot |\Delta(z)|\, dz . \tag{C.2}$$

A random matrix $A \in \mathrm{Sym}^n = \{B \in \mathbb{R}^{n\times n} \mid B^T = B\}$ is said to be from the *Gaussian orthogonal ensemble*, $A \in \mathrm{GOE}(n)$, iff the entries $a_{ij}$ are chosen in the following way:

- For $1 \le i \le n$ the entries $a_{ii}$ are i.i.d. standard normal distributed,
- for $1 \le i < j \le n$ the entries $a_{ij}$ are i.i.d. normal distributed with mean 0 and variance $\frac{1}{2}$,
- for $1 \le j < i \le n$ the entry $a_{ij}$ is set to $a_{ij} := a_{ji}$.

The probability that a random matrix from the $n$th Gaussian orthogonal ensemble is positive definite is given by (cf. for example [22])

$$\operatorname*{Prob}_{A\in\mathrm{GOE}(n)}[A \text{ is positive definite}] = \frac{1}{n! \cdot 2^{n/2} \cdot \prod_{d=1}^{n} \Gamma(\frac{d}{2})} \cdot M_n^+ . \tag{C.3}$$

For $n = 1, 2, 3$ this probability and also its asymptotics for $n \to \infty$ are known (cf. [22]). We have

$$\begin{aligned}
M_1^+ &= \sqrt{\frac{\pi}{2}} \\
M_2^+ &= \sqrt{\pi} \cdot (2 - \sqrt{2}) \\
M_3^+ &= \frac{3\pi}{\sqrt{2}} - 6 \\
M_n^+ &\sim n! \cdot 2^{n/2} \cdot \prod_{d=1}^{n} \Gamma(\tfrac{d}{2}) \cdot \exp\big(-n^2 \cdot \tfrac{\ln 3}{4}\big) , \quad \text{for } n \to \infty .
\end{aligned}$$

Knowing an exact expression for $M_n^+$, though maybe not too exciting on its own, might serve as a building block for understanding the type of integrals that we will describe next.

For $0 \le r \le n$ the Vandermonde determinant may be decomposed in the following way

$$\begin{aligned}\Delta(z) &= \prod_{1\le i<j\le n}(z_i - z_j)\\ &= \left(\prod_{1\le i<j\le r}(z_i - z_j)\right)\cdot\left(\prod_{r+1\le i<j\le n}(z_i - z_j)\right)\cdot\left(\prod_{i=1}^{r}\prod_{j=r+1}^{n}(z_i - z_j)\right)\\ &\overset{(*)}{=} \Delta(x)\cdot\Delta(y)\cdot\sum_{\ell=0}^{r\cdot(n-r)}(-1)^{\ell}\cdot\sigma_\ell(x^{-1}\otimes y)\cdot\prod_{i=1}^{r}x_i^{n-r}\end{aligned}$$

where $x := (z_1,\dots,z_r)$, $y := (z_{r+1},\dots,z_n)$, $\sigma_\ell$ denotes the $\ell$th elementary symmetric function, and

$$x^{-1}\otimes y := \left(\frac{y_1}{x_1},\dots,\frac{y_1}{x_r},\frac{y_2}{x_1},\dots,\frac{y_2}{x_r},\dots,\frac{y_{n-r}}{x_1},\dots,\frac{y_{n-r}}{x_r}\right)\in\mathbb{R}^{r\cdot(n-r)}\ .$$

The last equality $(*)$ is seen in the following way:

$$\begin{aligned}\prod_{i=1}^{r}\prod_{j=r+1}^{n}(z_i - z_j) &= \sum_{S=(s_{ij})\in\{0,1\}^{r\times(n-r)}}\prod_{i=1}^{r}\prod_{j=1}^{n-r}\begin{pmatrix}x_i & \text{if } s_{ij}=0\\ -y_j & \text{if } s_{ij}=1\end{pmatrix}\\ &= \sum_{\ell=0}^{r(n-r)}(-1)^{\ell}\cdot\sum_{\substack{S\in\{0,1\}^{r\times(n-r)}\\ \sum s_{ij}=\ell}}\prod_{i=1}^{r}\prod_{j=1}^{n-r}\begin{pmatrix}x_i & \text{if } s_{ij}=0\\ y_j & \text{if } s_{ij}=1\end{pmatrix}\\ &= \sum_{\ell=0}^{r(n-r)}(-1)^{\ell}\cdot\sigma_\ell(x^{-1}\otimes y)\cdot\prod_{i=1}^{r}x_i^{n-r}\ .\end{aligned}$$

So if we define

$$\Delta_{r,\ell}(z) \ := \ \Delta(x)\cdot\Delta(y)\cdot\sigma_\ell(x^{-1}\otimes y)\cdot\prod_{i=1}^{r}x_i^{n-r}\ , \tag{C.4}$$

then we have

$$\Delta(z) = \sum_{\ell=0}^{r\cdot(n-r)}(-1)^{\ell}\cdot\Delta_{r,\ell}(z)\ . \tag{C.5}$$

Analogous to (C.2) we define for $0\le r\le n$ and $0\le\ell\le r(n-r)$

$$J(n,r,\ell) \ := \int\limits_{z\in\mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}}\cdot|\Delta_{r,\ell}(z)|\,dz\ . \tag{C.6}$$

Note that we get $J(n,0,0) = J(n,n,0) = M_n^+$.

Abbreviating $\ell^* := r(n-r)-\ell$ we have

$$\begin{aligned}\Delta_{n-r,\ell^*}(y,x) &= \Delta_{n-r}(y)\cdot\Delta_r(x)\cdot\sigma_{\ell^*}(y^{-1}\otimes x)\cdot\prod_{j=1}^{n-r}y_j^{r}\\ &= \Delta_r(x)\cdot\Delta_{n-r}(y)\cdot\sigma_\ell(x^{-1}\otimes y)\cdot\prod_{i=1}^{r}x_i^{n-r} = \Delta_{r,\ell}(x,y)\ .\end{aligned}$$

| | $r=0$ | $r=1$ |
|---|---|---|
| $\ell=0$ | $\sqrt{\frac{\pi}{2}}$ | $\sqrt{\frac{\pi}{2}}$ |

(a) $n=1$

| | $r=0$ | $r=1$ | $r=2$ |
|---|---|---|---|
| $\ell=0$ | $\sqrt{\pi}\cdot(2-\sqrt{2})$ | $\sqrt{\frac{\pi}{2}}$ | $\sqrt{\pi}\cdot(2-\sqrt{2})$ |
| $\ell=1$ | – | $\sqrt{\frac{\pi}{2}}$ | – |

(b) $n=2$

| | $r=0$ | $r=1$ | $r=2$ | $r=3$ |
|---|---|---|---|---|
| $\ell=0$ | $\frac{3\pi}{\sqrt{2}}-6$ | $\pi(\sqrt{2}-1)$ | $\pi(1-\frac{1}{\sqrt{2}})$ | $\frac{3\pi}{\sqrt{2}}-6$ |
| $\ell=1$ | – | 2 | 2 | – |
| $\ell=2$ | – | $\pi(1-\frac{1}{\sqrt{2}})$ | $\pi(\sqrt{2}-1)$ | – |

(c) $n=3$

Table C.1: The values of $J(n,r,\ell)$ for $n=1,2,3$.

This implies the identity

$$J(n,r,\ell)=J(n,n-r,r(n-r)-\ell)\ . \tag{C.7}$$

Furthermore, we will show in Section C.2 (cf. Remark C.2.1) the relation

$$\sum_{r=0}^{n}\binom{n}{r}\cdot\sum_{\ell=0}^{r(n-r)}J(n,r,\ell)\ =\ n!\cdot 2^{\frac{n}{2}}\cdot\prod_{d=1}^{n}\Gamma\left(\tfrac{d}{2}\right)\ . \tag{C.8}$$

Table C.1 shows the values of $J(n,r,\ell)$ for $n=1,2,3$. These are easily computed by hand or with a computer algebra system. Although for higher dimensions the standard algorithms in a computer algebra system will not yield a good result, one can still approximate the value of the integrals by a Monte-Carlo method. The results of a Monte-Carlo approximation of $J(n,r,\ell)$, $n=4,5,6$, with $10^7$ samples are shown in Table C.2.

## C.2 The intrinsic volumes of the semidefinite cone

In this section we will derive the formulas for the intrinsic volumes of the semidefinite cone that we already stated in Proposition 4.4.21. Recall that the semidefinite cone is defined in $\mathrm{Sym}^n=\{A\in\mathbb{R}^{n\times n}\mid A^T=A\}$, which is a linear subspace of $\mathbb{R}^{n\times n}$ of dimension $\frac{n(n+1)}{2}=:t(n)$. An orthogonal basis for $\mathrm{Sym}^n$ is given by $B_{ij}$, $1\le i\le j\le n$,

$$B_{ij}:=\begin{cases}E_{ii} & \text{if } i=j\\ E_{ij}+E_{ji} & \text{if } i\neq j,\end{cases} \tag{C.9}$$

where $E_{ij}$ denotes as usual the $(i,j)$th elementary matrix. Note that $\|B_{ij}\|_F=\sqrt{2}$ for $i\neq j$. We denote the positive semidefinite cone and its intersection with the unit sphere $S(\mathrm{Sym}^n)$ by

$$\mathrm{Sym}^n_+=\{A\in\mathrm{Sym}^n\mid A \text{ is pos. semidef.}\}\ \text{ resp. }\quad K_n=\mathrm{Sym}^n_+\cap S(\mathrm{Sym}^n)\ .$$

| | $r=0$ | $r=1$ | $r=2$ | $r=3$ | $r=4$ |
|---|---|---|---|---|---|
| $\ell=0$ | 0.370 | 1.328 | 1.237 | 0.572 | 0.371 |
| $\ell=1$ | – | 2.758 | 3.995 | 2.006 | – |
| $\ell=2$ | – | 2.004 | 5.654 | 2.761 | – |
| $\ell=3$ | – | 0.571 | 4.004 | 1.328 | – |
| $\ell=4$ | – | – | 1.237 | – | – |

(a) $n=4$

| | $r=0$ | $r=1$ | $r=2$ | $r=3$ | $r=4$ | $r=5$ |
|---|---|---|---|---|---|---|
| $\ell=0$ | 0.20 | 1.39 | 2.02 | 1.03 | 0.34 | 0.20 |
| $\ell=1$ | – | 3.65 | 8.83 | 5.24 | 1.68 | – |
| $\ell=2$ | – | 3.63 | 16.87 | 12.82 | 3.61 | – |
| $\ell=3$ | – | 1.66 | 18.75 | 18.68 | 3.67 | – |
| $\ell=4$ | – | 0.34 | 12.88 | 16.81 | 1.39 | – |
| $\ell=5$ | – | – | 5.26 | 8.80 | – | – |
| $\ell=6$ | – | – | 1.03 | 2.04 | – | – |

(b) $n=5$

| | $r=0$ | $r=1$ | $r=2$ | $r=3$ | $r=4$ | $r=5$ | $r=6$ |
|---|---|---|---|---|---|---|---|
| $\ell=0$ | 0.11 | 1.55 | 4.15 | 2.85 | 0.83 | 0.20 | 0.11 |
| $\ell=1$ | – | 4.96 | 21.35 | 18.33 | 6.07 | 1.31 | – |
| $\ell=2$ | – | 6.31 | 52.95 | 60.42 | 21.00 | 3.95 | – |
| $\ell=3$ | – | 4.12 | 77.03 | 118.06 | 46.62 | 6.40 | – |
| $\ell=4$ | – | 1.31 | 73.62 | 164.72 | 72.22 | 5.03 | – |
| $\ell=5$ | – | 0.21 | 48.28 | 165.04 | 79.12 | 1.64 | – |
| $\ell=6$ | – | – | 21.68 | 121.68 | 52.71 | – | – |
| $\ell=7$ | – | – | 5.79 | 59.52 | 21.55 | – | – |
| $\ell=8$ | – | – | 0.86 | 18.22 | 4.03 | – | – |
| $\ell=9$ | – | – | – | 2.74 | – | – | – |

(c) $n=6$

Table C.2: The values of $J(n,r,\ell)$ for $n=4,5,6$, approximated by a Monte-Carlo method with $10^7$ samples.

We will show that the $k$th intrinsic volume of $K_n$, $-1 \leq k \leq t(n)-1 = \frac{n(n+1)}{2}-1$, is given by

$$V_k(K_n) = \frac{1}{n! \cdot 2^{\frac{n}{2}} \cdot \prod_{d=1}^{n} \Gamma(\frac{d}{2})} \cdot \sum_{r=0}^{n} \binom{n}{r} \cdot J(n, r, k+1-t(n-r)) \;,$$

where $J(n,r,\ell)$ is defined as in (C.6) for $0 \leq \ell \leq r(n-r) = t(n)-t(r)-t(n-r)$, resp. $J(n,r,\ell) := 0$ for the remaining cases. In particular, we have (cf. (C.3); cf. also Remark 4.4.22)

$$V_{-1}(K_n) = V_{t(n)-1}(K_n) = \frac{1}{n! \cdot 2^{\frac{n}{2}} \cdot \prod_{d=1}^{n} \Gamma(\frac{d}{2})} \cdot \int\limits_{z \in \mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}} \cdot |\Delta(z)| \, dz \;.$$

Note that the condition $J(n,r,k+1-t(n-r)) > 0$ can be rewritten as two inequalities

$$\begin{aligned} & J(n,r,k+1-t(n-r)) > 0 \\ & \qquad \iff \begin{pmatrix} k+1 \;\geq\; t(n-r) \;, \\ k+1 \;\leq\; r(n-r)+t(n-r) = t(n)-t(r) \end{pmatrix} . \end{aligned} \tag{C.10}$$

**Remark C.2.1.** 1. Recall that the intrinsic volumes form a discrete probability distribution (cf. Proposition 4.4.10), so that in particular $\sum_{k=-1}^{t(n)-1} V_k(K_n) = 1$. It is easily seen that in this sum every integral $J(n,r,\ell)$, $0 \leq \ell \leq r(n-r)$, appears exactly once, so that we get

$$1 = \sum_{k=-1}^{t(n)-1} V_k(K_n) = \frac{1}{n! \cdot 2^{\frac{n}{2}} \cdot \prod_{d=1}^{n} \Gamma(\frac{d}{2})} \cdot \sum_{r=0}^{n} \binom{n}{r} \cdot \sum_{\ell=0}^{r(n-r)} J(n,r,\ell) \;.$$

This shows the relation (C.8) of the integrals $J(n,r,\ell)$.

2. Note that the self-duality of the semidefinite cone implies

$$V_k(K_n) = V_{t(n)-2-k}(K) \;.$$

(cf. Proposition 4.4.10). This equality is verified easily using the relation (C.7), i.e., using $J(n,r,\ell) = J(n,n-r,r(n-r)-\ell)$.

Using the (approximated) values of $J(n,r,\ell)$ in Table C.1 and Table C.2 we may compute the intrinsic volumes of $K_n$ for $n = 1,\ldots,6$. The result is shown in Figure 4.2 in Section 4.4.1.

The following proposition gives a full description of the boundary structure of the semidefinite cone. Recall from Definition 3.1.7 that a face of a closed convex set is the intersection of the convex set with a supporting hyperplane.

**Proposition C.2.2.** *The faces of* $\mathrm{Sym}^n_+$ *are parametrized by the subspaces of* $\mathbb{R}^n$*. More precisely, for* $L \in \mathrm{Gr}_{n,r}$ *the set* $\{A \in \mathrm{Sym}^n_+ \mid \mathrm{im}(A) \subseteq L\}$ *is a face of* $\mathrm{Sym}^n_+$ *of dimension* $t(r) = \frac{r(r+1)}{2}$*. On the other hand, every face of* $\mathrm{Sym}^n_+$ *has dimension* $t(r)$ *for some* $0 \leq r \leq n$ *and is of the above form. Moreover, every face of* $\mathrm{Sym}^n_+$ *is of the form*

$$\left\{ Q \begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix} Q^T \,\middle|\, A' \in \mathrm{Sym}^r_+ \right\} \;, \tag{C.11}$$

*where $Q \in O(n)$ and $0 \leq r \leq n$. The normal cone at the face defined in* (C.11) *is given by*

$$\left\{ Q \begin{pmatrix} 0 & 0 \\ 0 & -A'' \end{pmatrix} Q^T \,\middle|\, A'' \in \operatorname{Sym}_+^{n-r} \right\} . \tag{C.12}$$

*Proof.* See for example [4, II.12] or [41]. □

Note that (C.11) and (C.12) show that when analyzing a face of $\operatorname{Sym}_+^n$, by choosing an appropriate basis of $\mathbb{R}^n$ we may assume without loss of generality that this face is of the form $\operatorname{Sym}_+^r \times \{0\}$ with corresponding normal cone $\{0\} \times (-\operatorname{Sym}_+^{n-r})$.

Next, we will describe the stratified structure of the semidefinite cone. Before we can do this, we need to prepare some notation. Let us define the eigenvalue map Eig, that maps each nonzero positive semidefinite matrix onto the ordered vector of its nonzero eigenvalues, i.e.,

$$\operatorname{Eig}\colon \operatorname{Sym}_+^n \setminus \{0\} \ni A \mapsto (\lambda_1, \ldots, \lambda_r) ,$$

if $\lambda_1 \geq \ldots \geq \lambda_r > 0$ are the positive eigenvalues of $A$. We will need to distinguish the matrices according to their eigenvalue patterns. Therefore, we introduce the notation

$$(\rho_1, \ldots, \rho_m) \vDash r \;:\Longleftrightarrow\; \rho_1, \ldots, \rho_m \in \mathbb{Z}_{>0} ,\; \rho_1 + \ldots + \rho_m = r .$$

Furthermore, we define the eigenvalue pattern of a (nonzero) positive semidefinite matrix $A$ via

$$\operatorname{patt}(A) := (\rho_1, \ldots, \rho_m) , \quad \text{iff } \lambda_1 = \ldots = \lambda_{\rho_1} > \lambda_{\rho_1+1} = \ldots = \lambda_{\rho_1+\rho_2} > \ldots ,$$

where $(\lambda_1, \ldots, \lambda_r) = \operatorname{Eig}(A)$. Note that in the above notation we have

$$\operatorname{patt}(A) \vDash \operatorname{rk}(A) = \rho_1 + \ldots + \rho_m .$$

Using this notation, $K_n$ decomposes into

$$K_n = \dot{\bigcup}_{r=1}^{n} \dot{\bigcup}_{\rho \vDash r} M_{n,\rho} , \tag{C.13}$$

with

$$M_{n,\rho} := \{ A \in K_n \mid \operatorname{patt}(A) = \rho \} . \tag{C.14}$$

Note that $\operatorname{int}(K_n) = \dot{\bigcup}_{\rho \vDash n} M_{n,\rho}$ and $\partial K_n = \dot{\bigcup}_{r=1}^{n-1} \dot{\bigcup}_{\rho \vDash r} M_{n,\rho}$.

In the following proposition we will show that the semidefinite cap $K_n$ satisfies all conditions of a stratified convex cap (cf. Definition 3.3.9). Moreover, we will show that the decomposition in (C.13) is a valid stratification of $K_n$, and we will determine the essential and the negligible pieces in this decomposition.

**Proposition C.2.3.** *The set $M_{n,\rho}$, $\rho \vDash r \leq n$, defined in* (C.14) *is a smooth submanifold of the unit sphere $S(\operatorname{Sym}^n)$. Furthermore, the spherical duality bundle $N^S M_{n,\rho}$ (cf.* (3.9)*) is a smooth manifold for all $\rho \vDash n$. In particular, $K_n$ is a stratified convex set.*

*The pieces $\{ M_{n,1^{(r)}} \mid 1 \leq r \leq n \}$, where $1^{(r)} := (1, 1, \ldots, 1) \vDash r$, are essential and all the other pieces $M_{n,\rho}$ are negligible.*

*Proof.* Let $\rho = (\rho_1, \ldots, \rho_m) \vDash r \leq n$. We define the set $P_{n,\rho} \subset S^{n-1}$ via

$$P_{n,\rho} := \left\{ \lambda \in S^{n-1} \;\middle|\; \begin{array}{l} \lambda_1 = \ldots = \lambda_{\rho_1} > \lambda_{\rho_1+1} = \ldots = \lambda_{\rho_1+\rho_2} > \ldots \\ \ldots \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0 \end{array} \right\} . \tag{C.15}$$

It is easily seen that the set $P_{n,\rho}$ is a $(m-1)$-dimensional submanifold of $S^{n-1}$ (there are $m$ degrees of freedom for the values of the blocks of $\lambda$; minus 1 degree of freedom because of the restriction $\lambda \in S^{n-1}$).

Furthermore, we define the subgroup $O(n,\rho)$ of $O(n)$ via

$$\begin{aligned} O(n,\rho) &:= \left\{ \begin{pmatrix} Q_1 & & & \\ & \ddots & & \\ & & Q_m & \\ & & & Q' \end{pmatrix} \;\middle|\; Q_i \in O(\rho_i), Q' \in O(n-r) \right\} \\ &\cong O(\rho_1) \times \ldots \times O(\rho_m) \times O(n-r) . \end{aligned}$$

From Section 5.3 we know that the homogeneous space $O(n)/O(n,\rho)$ is a smooth manifold of dimension

$$\begin{aligned} \dim O(n)/O(n,\rho) &= \dim O(n) - \sum_{i=1}^{m} \dim O(\rho_i) - \dim O(n-r) \\ &= t(n-1) - \sum_{i=1}^{m} t(\rho_i - 1) - t(n-r-1) . \end{aligned}$$

We now consider the map

$$\psi_{n,\rho} \colon P_{n,\rho} \times O(n) \to M_{n,\rho} , \quad (\lambda, Q) \mapsto Q \cdot \operatorname{diag}(\lambda) \cdot Q^T . \tag{C.16}$$

This map is clearly smooth and surjective. Concerning the fiber of $A \in M_{n,\rho}$, we may assume w.l.o.g. that $A = \operatorname{diag}(\mu)$. Note that we have

$$Q \cdot \operatorname{diag}(\lambda) \cdot Q^T = \operatorname{diag}(\mu) \iff \lambda = \mu \text{ and } Q \cdot \operatorname{diag}(\lambda) = \operatorname{diag}(\lambda) \cdot Q .$$

Furthermore, it is easily checked that for $\lambda \in P_{n,\rho}$ we have

$$Q \cdot \operatorname{diag}(\lambda) = \operatorname{diag}(\lambda) \cdot Q \iff Q \in O(n,\rho) .$$

We may thus conclude that

$$\psi_{n,\rho}(\lambda, Q) = A \iff \lambda = \mu \text{ and } Q \in O(n,\rho) . \tag{C.17}$$

This implies that the map $\psi_{n,\rho}$ factorizes over the product $P_{n,\rho} \times O(n)/O(n,\rho)$, i.e., we have a commutative diagram

$$\begin{tikzcd} P_{n,\rho} \times O(n) \arrow[d, "\Pi"'] \arrow[dr, "\psi_{n,\rho}"] & \\ P_{n,\rho} \times O(n)/O(n,\rho) \arrow[r, dashed, "\overline{\psi_{n,\rho}}"'] & M_{n,\rho} \end{tikzcd} , \tag{C.18}$$

where $\Pi \colon P_{n,\rho} \times O(n) \to P_{n,\rho} \times O(n)/O(n,\rho)$ denotes the canonical projection map. Moreover, (C.17) implies that the map $\overline{\psi_{n,\rho}}$ is a bijection. Using Lemma 5.3.2 it is straightforward to show that the differential of $\overline{\psi_{n,\rho}}$ also has full rank, i.e., $\overline{\psi_{n,\rho}}$ is an immersion, and thus $\overline{\psi_{n,\rho}}$ is a diffeomorphism. The compactness of the domain $P_{n,\rho} \times O(n)/O(n,\rho)$ ensures that $M_{n,\rho}$ is indeed a submanifold of $S(\mathrm{Sym}^n)$.

As for the claim about the duality bundle, note that for $A = Q \cdot \operatorname{diag}(\lambda) \cdot Q^T$, $\lambda \in P_{n,\rho}$, the normal cone $N_A(K_n)$ is given by (cf. Proposition C.2.2)

$$N_A(K_n) = \left\{ Q \begin{pmatrix} 0 & 0 \\ 0 & -A'' \end{pmatrix} Q^T \,\middle|\, A'' \in \operatorname{Sym}_+^{n-r} \right\} .$$

Using the notation

$$K_{n-r}^\circ := \operatorname{int}(K_{n-r}) ,$$

where the interior is of course taken with respect to the topology on $S(\operatorname{Sym}_{n-r})$, we can define the map

$$\begin{aligned} \Psi_{n,\rho} \colon P_{n,\rho} \times O(n) \times K_{n-r}^\circ &\to N^S M_{n,\rho} , \\ (\lambda, Q, A'') &\mapsto \left( \psi(\lambda, Q) \, , \; Q \begin{pmatrix} 0 & 0 \\ 0 & -A'' \end{pmatrix} Q^T \right) . \end{aligned}$$

This map is smooth and surjective. As in (C.18), the map $\Psi_{n,\rho}$ factorizes over $P_{n,\rho} \times O(n)/O(n,\rho) \times K_{n-r}^\circ$, and thus yields a diffeomorphism

$$\overline{\Psi_{n,\rho}} \colon P_{n,\rho} \times O(n)/O(n,\rho) \times K_{n-r}^\circ \to N^S M_{n,\rho} .$$

In particular, $N^S M_{n,\rho}$ is a smooth manifold of dimension

$$\begin{aligned} \dim N^S M_{n,\rho} &= \dim(P_{n,\rho} \times O(n)/O(n,\rho) \times K_{n-r}^\circ) \\ &= (m-1) + \left( t(n-1) - \sum_{i=1}^m t(\rho_i - 1) - t(n-r-1) \right) + (t(n-r) - 1) . \end{aligned}$$

Note that we have

$$t(n-1) - t(n-r-1) + t(n-r) - 1 = t(n) - 1 - r ,$$

and

$$m - 1 - \sum_{i=1}^m t(\rho_i - 1) \begin{cases} = r - 1 & \text{if } \rho = 1^{(r)} \\ < r - 1 & \text{else} . \end{cases}$$

Therefore, we have

$$\dim N^S M_{n,\rho} \begin{cases} = t(n) - 2 & \text{if } \rho = 1^{(r)} \\ < t(n) - 2 & \text{else} , \end{cases}$$

i.e., the pieces $M_{n,1^{(r)}}$, $1 \le r \le n$, are essential, and all other pieces are negligible. □

From now on we may restrict the computations to the essential pieces $M_{n,1^{(r)}}$. We furthermore use the notation

$$P_r := P_{r,1^{(r)}} = \left\{ \lambda \in S^{r-1} \mid \lambda_1 > \lambda_2 > \ldots > \lambda_r > 0 \right\} . \tag{C.19}$$

Note that we have a natural bijection between $P_r$ and $P_{n,1^{(r)}}$ given by

$$E_{n,r} \colon P_r \to P_{n,1^{(r)}} , \quad \lambda \mapsto (\lambda, \underbrace{0, \ldots, 0}_{n-r}) , \tag{C.20}$$

which is easily seen to be an isometry.

The subgroup $O(n,\rho)$ for $\rho = 1^{(r)}$ is given by

$$O(n,1^{(r)}) = \underbrace{\{\pm 1\} \times \ldots \times \{\pm 1\}}_{r \text{ times}} \times O(n-r)\ . \tag{C.21}$$

In the following proposition we will replace the map $\overline{\psi_{n,1^{(r)}}}$, which parametrizes $M_{n,1^{(r)}}$ via $O(n)/O(n,1^{(r)})$, by a map $\varphi_{n,r}$, which parametrizes $M_{n,1^{(r)}}$ via $\mathrm{St}_{n,r} = O(n)/O(n-r)$. The advantage is that we have already discussed the homogeneous space $\mathrm{St}_{n,r}$ in Section 5.3.1.

**Proposition C.2.4.** *Let $1 \le r \le n$, and let $\varphi_{n,r}$ be defined via*

$$\varphi_{n,r} \colon P_r \times \mathrm{St}_{n,r} \to M_{n,1^{(r)}}\ , \quad (\lambda,[Q]) \mapsto Q \cdot \mathrm{diag}(E_{n,r}(\lambda)) \cdot Q^T\ ,$$

*where $P_r$ is defined as in* (C.19)*, and $E_{n,r}$ is defined as in* (C.20)*. The map $\varphi_{n,r}$ is a $2^r$-fold smooth covering, i.e., $\varphi_{n,r}$ is a local diffeomorphism, and for every $A \in M_{n,1^{(r)}}$ we have $|\varphi_{n,r}^{-1}(A)| = 2^r$. The Normal Jacobian of $\varphi_{n,r}$ at $(\lambda,[Q]) \in P_r \times \mathrm{St}_{n,r}$ is given by*

$$\mathrm{ndet}(D_{(\lambda,[Q])}\varphi_{n,r}) = 2^{r(n-r)/2+r(r-1)/4} \cdot \prod_{i=1}^{r} \lambda_i^{n-r} \cdot \Delta(\lambda)\ ,$$

*where $\Delta(\lambda) = \prod_{1\le i<j\le r}(\lambda_i - \lambda_j)$ denotes the Vandermonde determinant.*

*Proof.* From (C.21) it is easily seen that we have a canonical projection map

$$\Pi_r \colon\ \mathrm{St}_{n,r} \to O(n)/O(n,1^{(r)})\ ,$$

which, for $Q \in O(n)$, maps the coset of $Q$ in $\mathrm{St}_{n,r} = O(n)/O(n-r)$ to the coset of $Q$ in $O(n)/O(n,1^{(r)})$. Roughly speaking, a tuple of orthonormal vectors maps to the tuple of the corresponding directions. So the projection map $\Pi_r$ is in fact a $2^r$-fold covering. Now, we can write the map $\varphi_{n,r}$ in the form

$$\varphi_{n,r} = \overline{\psi_{n,1^{(r)}}} \circ (E_{n,r}, \Pi_r)\ ,$$

where $\overline{\psi_{n,1^{(r)}}}$ is defined as in (C.18). As $\overline{\psi_{n,1^{(r)}}}$ is a diffeomorphism, it follows that $\varphi_{n,r}$ is a $2^r$-fold covering.

Concerning the differential $D_{(\lambda,[Q])}\varphi_{n,r}$ note that w.l.o.g. we may assume $Q = I_n$. For $\zeta \in T_\lambda P_r$ we have

$$D_{(\lambda,[I_n])}\varphi_{n,r}(\zeta,0) = \mathrm{diag}(\zeta)\ . \tag{C.22}$$

As for the second component, recall from Section 5.3.1 that for $\overline{\mathrm{Skew}_n}$ we have an orthonormal basis given by

$$\{E_{ij} - E_{ji} \mid 1 \le j < i \le r\} \ \cup\ \{E_{ij} - E_{ji} \mid r+1 \le i \le n\,,\ 1 \le j \le r\}\ ,$$

where $E_{ij}$ denotes the $(i,j)$th elementary matrix. Let $\eta_{ij}$ denote the corresponding tangent vector in $T_{[I_n]}\,\mathrm{St}_{n,r}$, i.e.,

$$\eta_{ij} := [I_n\ ,\ E_{ij} - E_{ji}] \in T_{[I_n]}\,\mathrm{St}_{n,r}\ , \tag{C.23}$$

where either $1 \le j < i \le r$, or $r+1 \le i \le n$ and $1 \le j \le r$. Furthermore, let $U_{ij} \colon \mathbb{R} \to O(n)$ be a curve such that the induced curve in $\mathrm{St}_{n,r}$, $t \mapsto [U_{ij}(t)]$, defines the direction $\eta_{ij}$, i.e., $\frac{d}{dt}[U_{ij}(t)](0) = \eta_{ij}$. Then we may compute the derivative of $\varphi_{n,r}$ in the second component via

$$\begin{aligned} D_{(\lambda,[I_n])}(0,\eta_{ij}) &= \tfrac{d}{dt}\big(U_{ij}(t)\cdot \operatorname{diag}(\lambda,0,\dots,0)\cdot U_{ij}(t)^T\big)(0) \\ &= \big(\underbrace{\tfrac{d}{dt}U_{ij}(0)}_{E_{ij}-E_{ji}}\big)\cdot \operatorname{diag}(\lambda,0,\dots,0) + \operatorname{diag}(\lambda,0,\dots,0)\cdot\big(\underbrace{\tfrac{d}{dt}U_{ij}(t)^T(0)}_{E_{ji}-E_{ij}}\big) \\ &= \begin{cases} (\lambda_j-\lambda_i)\cdot(E_{ij}+E_{ji}) & \text{if } 1\le j<i\le r \\ \lambda_j\cdot(E_{ij}+E_{ji}) & \text{if } r+1\le i\le n,\ 1\le j\le r\,. \end{cases} \end{aligned} \tag{C.24}$$

Note that the direction $\eta_{ij}$ has length 1, while the direction $E_{ij}+E_{ji}$ has length $\sqrt{2}$. Taking this into account, we get from (C.22) and (C.24) that

$$\begin{aligned} \operatorname{ndet}(D_{(\lambda,[Q])}\varphi_{n,r}) &= \prod_{1\le j<i\le r}(\sqrt{2}\cdot(\lambda_j-\lambda_i))\cdot \prod_{\substack{r+1\le i\le n\\ 1\le j\le r}}(\sqrt{2}\cdot\lambda_j) \\ &= \sqrt{2}^{r(r-1)/2}\cdot \prod_{1\le j<i\le r}(\lambda_j-\lambda_i)\cdot \sqrt{2}^{(n-r)r}\cdot \prod_{1\le j\le r}\lambda_j^{n-r}\,. \qquad \square \end{aligned}$$

**Remark C.2.5.** The map $\varphi_{n,r}$ has the important property of $O(n)$-*equivariance*, by which we mean the following. The domain $P_r\times \mathrm{St}_{n,r}$ can be endowed with an $O(n)$-action via $\tilde{Q}\bullet(\lambda,[Q]) := (\lambda,[\tilde{Q}Q])$, and the codomain $M_{n,1^{(r)}}\subset S(\mathrm{Sym}^n)$ is endowed with an $O(n)$-action via conjugation, i.e., $\tilde{Q}\bullet A := \tilde{Q}\cdot A\cdot \tilde{Q}^T$. Then $\varphi_{n,r}$ satisfies for all $(\lambda,[Q])\in P_r\times \mathrm{St}_{n,r}$ and all $\tilde{Q}\in O(n)$ the equivariance property $\varphi_{n,r}(\tilde{Q}\bullet(\lambda,[Q])) = \tilde{Q}\bullet\varphi_{n,r}(\lambda,[Q])$.

It remains to compute the principal curvatures of the essential pieces $M_{n,1^{(r)}}$ before we can use Theorem 4.3.2 to compute the intrinsic volumes of $K_n$. We will do this in the following lemma. Recall from the proof of Proposition C.2.3 that the dimension of $M_{n,1^{(r)}}$ is given by

$$\begin{aligned} \dim M_{n,1^{(r)}} &= r-1+t(n-1)-t(n-r-1) \\ &= r(n-r)+t(r)-1\,. \end{aligned}$$

**Lemma C.2.6.** *Let $A = Q\cdot \operatorname{diag}(\lambda,0,\dots,0)\cdot Q^T\in M_{n,1^{(r)}}$, where $\lambda\in P_r$. Furthermore, let $A''\in \mathrm{Sym}_+^{n-r}$, so that $B := Q\cdot\begin{pmatrix}0 & 0\\ 0 & -A''\end{pmatrix}\cdot Q^T\in N_A(K_n)$ is a vector in the normal cone of $K_n$ in $A$. If $\mu_1\ge\dots\ge\mu_{n-r}\ge 0$ denote the eigenvalues of $A''$, then the principal curvatures of $M_{n,1^{(r)}}$ at $A$ in direction $-B$ are given by*

$$\frac{\mu_1}{\lambda_1},\dots,\frac{\mu_{n-r}}{\lambda_1},\ \frac{\mu_1}{\lambda_2},\dots,\frac{\mu_{n-r}}{\lambda_2},\ \dots,\ \frac{\mu_1}{\lambda_r},\dots,\frac{\mu_{n-r}}{\lambda_r},\ \underbrace{0,\dots,0}_{t(r)-1}\,.$$

*Proof.* Using the $O(n)$-equivariance property of $\varphi_{n,r}$ (cf. Remark C.2.5) we may assume w.l.o.g. that $A = \operatorname{diag}(\lambda)$, and $A'' = \operatorname{diag}(\mu_1,\dots,\mu_{n-r})$. From Proposition C.2.4 we get that the tangent space of $M_{n,1^{(r)}}$ at $A$ is given by

$$T_A M_{n,1^{(r)}} = D\varphi_{n,r}(T_\lambda P_r\times T_{[I_n]}\,\mathrm{St}_{n,r})\,.$$

It is easily seen that all the vectors in $D\varphi_{n,r}(T_\lambda P_r\times\{0\})$ are principal directions with principal curvature 0, thus giving $r-1$ of the claimed $t(r)-1$ zero curvatures.

As for the second component, let as in the proof of Proposition C.2.4 $U_{ij}(t) \in O(n)$ be a curve such that the corresponding curve $[U_{ij}(t)]$ in $\mathrm{St}_{n,r}$ defines the direction $\eta_{ij} \in T_{[I_n]}\,\mathrm{St}_{n,r}$ (cf. (C.23)). The corresponding curve

$$t \mapsto \varphi_{n,r}(\lambda, [U_{ij}(t)]) = U_{ij}(t) \cdot \mathrm{diag}(\lambda, 0, \ldots, 0) \cdot U_{ij}(t)^T$$

defines the image $D_{(\lambda,[I_n])}(0, \eta_{ij})$ (cf. (C.24)). Moreover, we may define a normal extension of $-B = \mathrm{diag}(0, \ldots, 0, \mu_1, \ldots, \mu_{n-r})$ along this curve via

$$v(t) := U_{ij}(t) \cdot \mathrm{diag}(0, \ldots, 0, \mu_1, \ldots, \mu_{n-r}) \cdot U_{ij}(t)^T \ .$$

Differentiating this normal extension yields

$$\begin{aligned}\tfrac{d}{dt}v(0) &= (E_{ij} - E_{ji}) \cdot \mathrm{diag}(0, \mu) + \mathrm{diag}(0, \mu) \cdot (E_{ji} - E_{ij}) \\ &= \begin{cases} 0 & \text{if } 1 \leq j < i \leq r \\ -\mu_{i-r} \cdot (E_{ij} + E_{ji}) & \text{if } r+1 \leq i \leq n,\ 1 \leq j \leq r \ . \end{cases}\end{aligned}$$

Comparing this with the values of $D_{(\lambda,[I_n])}(0, \eta_{ij})$ in (C.24) yields

$$\tfrac{d}{dt}v(0) = \begin{cases} 0 \cdot D_{(\lambda,[I_n])}(0, \eta_{ij}) & \text{if } 1 \leq j < i \leq r \\ -\frac{\mu_{i-r}}{\lambda_j} \cdot D_{(\lambda,[I_n])}(0, \eta_{ij}) & \text{if } r+1 \leq i \leq n,\ 1 \leq j \leq r \ . \end{cases}$$

We may conclude that the direction $D_{(\lambda,[I_n])}(0, \eta_{ij})$ is a principal direction with curvature 0 resp. $-\left(-\frac{\mu_{i-r}}{\lambda_j}\right) = \frac{\mu_{i-r}}{\lambda_j}$ (cf. Section 4.1.1). The final computation $t(r-1) + r - 1 = t(r) - 1$ shows that indeed $t(r) - 1$ principal curvatures of $M_{n,1^{(r)}}$ are 0, which finishes the proof. □

Before we give the proof of Proposition 4.4.21, we state another small lemma, that will come in handy for an integral conversion, that we will have to make.

**Lemma C.2.7.** *Let $f\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be a homogeneous function of degree $d$, i.e., $f(x) = \|x\|^d \cdot f(\|x\|^{-1} \cdot x)$. Then for $U \subset S^{n-1}$ a Borel set*

$$\int\limits_{p \in U} f(p)\, dp = \frac{1}{2^{\frac{n+d}{2}-1} \cdot \Gamma(\frac{n+d}{2})} \cdot \int\limits_{x \in \hat{U}} e^{-\frac{\|x\|^2}{2}} \cdot f(x)\, dx \ ,$$

*where $\hat{U} = \{s \cdot p \mid s > 0\,,\ p \in U\}$.*

*Proof.* The normal Jacobian of the projection $\Pi\colon \mathbb{R}^n \setminus \{0\} \to S^{n-1}$, $\Pi(x) = \|x\|^{-1} \cdot x$, is given by $\mathrm{ndet}(D_x\Pi) = \|x\|^{-(n-1)}$. From the coarea formula in Lemma 4.1.15 we thus get

$$\begin{aligned}\int\limits_{x \in \hat{U}} e^{-\frac{\|x\|^2}{2}} \cdot f(x)\, dx &= \int\limits_{p \in U} \int_0^\infty s^{n-1} \cdot e^{-s^2/2} \cdot f(s \cdot p)\, ds\, dp \\ &= \int\limits_{p \in U} f(p)\, dp \cdot \int_0^\infty s^{n-1+d} \cdot e^{-s^2/2}\, ds \ .\end{aligned}$$

Substituting $t := s^2/2$, and using the well-known formula $\int_0^\infty t^{z-1} \cdot e^{-t}\, dt = \Gamma(z)$, we finally get

$$\begin{aligned}\int_0^\infty s^{n-1+d} \cdot e^{-s^2/2}\, ds &= \int_0^\infty \sqrt{2t}^{\,n-1+d} \cdot e^{-t} \cdot \frac{1}{\sqrt{2t}}\, dt \\ &= 2^{\frac{n+d}{2}-1} \cdot \int_0^\infty t^{\frac{n+d}{2}-1} \cdot e^{-t}\, dt = 2^{\frac{n+d}{2}-1} \cdot \Gamma(\tfrac{n+d}{2}) \ . \quad \square\end{aligned}$$

*Proof of Proposition 4.4.21.* We use the notation

$$\begin{aligned} d(n,r) &:= \dim M_{n,1^{(r)}} = r(n-r) + t(r) - 1 \\ &= t(n) - t(n-r) - 1 \ . \end{aligned}$$

From Proposition 4.4.4 and Proposition C.2.6 we get for $0 \le k \le t(n) - 2$

$$V_k(K_n) \;=\; \frac{1}{\mathcal{O}_k \cdot \mathcal{O}_{t(n)-2-k}} \cdot \sum_{r=1}^{n-1} \int\limits_{A \in M_{n,1^{(r)}}} \int\limits_{B \in N_A^S} \sigma^{(r)}_{d(n,r)-k}(A,B)\, dN_A^S\, dM_{n,1^{(r)}} \ ,$$

where $\sigma^{(r)}_{d(n,r)-k}$ shall denote the dependence on $M_{n,1^{(r)}}$. From Lemma C.2.6 we know the principal curvatures of $M_{n,1^{(r)}}$. Note that the $\ell$th elementary symmetric function in these principal curvatures is given by $\sigma_\ell(\mu \otimes \lambda^{-1})$. Using the coarea formula (cf. Lemma 4.1.15) and Proposition C.2.4 we can transform the integral over $M_{n,1^{(r)}}$ to an integral over $P_r \times \mathrm{St}_{n,r}$. Similarly, we can transform the integral over $N_A^S$ to an integral over $P_{n-r} \times O(n-r)$, as $\mathrm{St}_{n-r,n-r} = O(n-r)$. This yields

$$\begin{aligned} &\int\limits_{A \in M_{n,1^{(r)}}} \int\limits_{B \in N_A^S} \sigma^{(r)}_{d(n,r)-k}(A,B)\, dN_A^S\, dM_{n,1^{(r)}} \\ &\quad = \frac{1}{2^r} \cdot \int\limits_{P_r \times \mathrm{St}_{n,r}} \frac{1}{2^{n-r}} \cdot \int\limits_{P_{n-r} \times O(n-r)} 2^{r(n-r)/2 + r(r-1)/4} \cdot \prod_{i=1}^r \lambda_i^{n-r} \cdot \Delta(\lambda) \\ &\qquad\qquad \cdot 2^{\frac{(n-r)(n-r-1)}{4}} \cdot \Delta(\mu) \cdot \sigma_{d(n,r)-k}(\mu \otimes \lambda^{-1})\, d(\lambda, [Q_1])\, d(\mu, [Q_2]) \\ &\quad = 2^{\frac{n(n-1)}{4} - n} \cdot \mathrm{vol}\,\mathrm{St}_{n,r} \cdot \mathrm{vol}\, O(n-r) \cdot \int\limits_{P_r} \int\limits_{P_{n-r}} \Delta(\lambda) \cdot \Delta(\mu) \cdot \prod_{i=1}^r \lambda_i^{n-r} \\ &\qquad\qquad \cdot \sigma_{d(n,r)-k}(\mu \otimes \lambda^{-1})\, d\lambda\, d\mu \\ &\quad \overset{(*)}{=} \frac{2^{\frac{n(n-1)}{4}} \cdot \pi^{\frac{t(n)}{2}}}{\prod_{d=1}^n \Gamma(\frac{d}{2})} \cdot \int\limits_{P_r} \int\limits_{P_{n-r}} \underbrace{\Delta(\lambda) \cdot \Delta(\mu) \cdot \prod_{i=1}^r \lambda_i^{n-r} \cdot \sigma_{d(n,r)-k}(\mu \otimes \lambda^{-1})}_{=: f(\lambda,\mu)}\, d\lambda\, d\mu \ , \end{aligned}$$

where $(*)$ follows from (cf. Proposition 5.2.1 and (5.19))

$$\mathrm{vol}\,\mathrm{St}_{n,r} \cdot \mathrm{vol}\, O(n-r) = \mathrm{vol}\, O(n) = \frac{2^n \cdot \pi^{\frac{n^2+n}{4}}}{\prod_{d=1}^n \Gamma(\frac{d}{2})} \ .$$

Note that the integrand $f(\lambda,\mu)$ coincides with the function (cf. (C.4))

$$f(\lambda,\mu) = \Delta_{r,d(n,r)-k}(\lambda,\mu) \ .$$

Note also that the integrand $f(\lambda,\mu)$ is homogeneous in $\lambda$ of degree

$$r(n-r) + \frac{r(r-1)}{2} - d(n,r) + k = k + 1 - r \ .$$

Furthermore, it is homogeneous in $\mu$ of degree

$$\frac{(n-r)(n-r-1)}{2} + d(n,r) - k = \frac{n(n-1)}{2} + r - (k+1) \ .$$

Using Lemma C.2.7 twice we get

$$\int_{P_r}\int_{P_{n-r}} f(\lambda,\mu)\,d\lambda\,d\mu = \frac{1}{2^{\frac{k-1}{2}}\cdot\Gamma(\frac{k+1}{2})}\cdot\frac{1}{2^{\frac{n(n+1)}{4}-\frac{k+3}{2}}\cdot\Gamma(\frac{n(n+1)}{4}-\frac{k+1}{2})}\cdot\int_{\hat P_r}\int_{\hat P_{n-r}} e^{-\frac{\|x\|^2+\|y\|^2}{2}}\cdot f(x,y)\,dy\,dx\;.$$

We may simplify the constant to

$$\frac{1}{2^{\frac{k-1}{2}}\cdot\Gamma(\frac{k+1}{2})}\cdot\frac{1}{2^{\frac{n(n+1)}{4}-\frac{k+3}{2}}\cdot\Gamma(\frac{n(n+1)}{4}-\frac{k+1}{2})} = \frac{1}{2^{\frac{n(n+1)}{4}}}\cdot\frac{2}{\Gamma(\frac{k+1}{2})}\cdot\frac{2}{\Gamma(\frac{n(n+1)}{4}-\frac{k+1}{2})}\overset{(4.13)}{=}\frac{1}{2^{\frac{n(n+1)}{4}}}\cdot\frac{\mathcal{O}_k\cdot\mathcal{O}_{t(n)-2-k}}{\pi^{\frac{n(n+1)}{4}}}\;.$$

All in all, we get for $0\le k\le t(n)-2$

$$V_k(K_n) = \frac{1}{2^{\frac{n}{2}}\cdot\prod_{d=1}^{n-1}\Gamma(\frac{d}{2})}\cdot\sum_{r=1}^{n-1}\int_{\hat P_r}\int_{\hat P_{n-r}} e^{-\frac{\|(x,y)\|^2}{2}}\cdot f(x,y)\,dx\,dy\;. \tag{C.25}$$

It is easily seen that if $\tilde x\in\mathbb{R}^r$ is obtained from $x$ by permuting the entries of $x$, and if the same holds for $\tilde y\in\mathbb{R}^{n-r}$ and $y$, then we have

$$|f(\tilde x,\tilde y)| = |f(x,y)|\;.$$

Note that $f(x,y)>0$ for $x\in P_r$ and $y\in P_{n-r}$. Using this, we can rewrite the integral (C.25) via

$$\begin{aligned}\int_{\hat P_r}\int_{\hat P_{n-r}} e^{-\frac{\|(x,y)\|^2}{2}}\cdot f(x,y)\,dx\,dy &= \frac{1}{r!}\cdot\int_{\mathbb{R}^r_+}\frac{1}{(n-r)!}\cdot\int_{\mathbb{R}^{n-r}_+} e^{-\frac{\|(x,y)\|^2}{2}}\cdot|f(x,y)|\,dx\,dy\\ &= \frac{1}{r!\cdot(n-r)!}\cdot\int_{z\in\mathbb{R}^n_+} e^{-\frac{\|z\|^2}{2}}\cdot|\Delta_{r,d(n,r)-k}(z)|\,dz\\ &\overset{(C.6)}{=}\frac{1}{r!\cdot(n-r)!}\cdot J(n,r,t(n)-t(n-r)-1-k)\;.\end{aligned}$$

So from (C.25) we finally get for $0\le k\le t(n)-2$

$$\begin{aligned}V_k(K_n) &= \frac{1}{n!\cdot2^{\frac{n}{2}}\cdot\prod_{d=1}^{n}\Gamma(\frac{d}{2})}\cdot\sum_{r=1}^{n-1}\binom{n}{r}\cdot J(n,r,t(n)-t(n-r)-1-k)\\ &\overset{[s:=n-r]}{=}\frac{1}{n!\cdot2^{\frac{n}{2}}\cdot\prod_{d=1}^{n}\Gamma(\frac{d}{2})}\cdot\sum_{s=1}^{n-1}\binom{n}{s}\cdot J(n,n-s,t(n)-t(s)-(k+1))\\ &\overset{(C.7)}{=}\frac{1}{n!\cdot2^{\frac{n}{2}}\cdot\prod_{d=1}^{n}\Gamma(\frac{d}{2})}\cdot\sum_{s=1}^{n-1}\binom{n}{s}\cdot J(n,s,k+1-t(n-s))\\ &\overset{(*)}{=}\frac{1}{n!\cdot2^{\frac{n}{2}}\cdot\prod_{d=1}^{n}\Gamma(\frac{d}{2})}\cdot\sum_{s=0}^{n}\binom{n}{s}\cdot J(n,s,k+1-t(n-s))\;,\end{aligned}$$

where $(*)$ follows from $J(n,0,k+1-t(n)) \neq 0 \;\Rightarrow\; k+1-t(n)=0$, and $J(n,n,k+1) \neq 0 \;\Rightarrow\; k+1=0$. The intrinsic volumes $V_{t(n)-1}(K_n) = V_{-1}(K_n)$ are given by the relative volume of $K_n$, which equals the relative volume of $M_{n,1^{(n)}}$. This is computed as above in an analogous and simpler manner (setting $r=n$). We skip the details of this last computation. □

## C.3 Observations, open questions, conjectures

In this section we will formulate some open questions about the intrinsic volumes of the semidefinite cone.

We begin with a specialization of Conjecture 4.4.16, which seems the more plausible after a look at Figure 4.3 in Section 4.4.1.

**Conjecture C.3.1.** The intrinsic volumes of the semidefinite cone form a log-concave sequence, i.e., for $n \geq 1$ we have

$$V_k(K_n)^2 \;\geq\; V_{k-1}(K_n)\cdot V_{k+1}(K_n)\;, \quad \text{for all } 0 \leq k \leq t(n)-2\;.$$

The following question is motivated from the admittedly few exact values of $J(n,r,\ell)$ that we know (cf. Table C.1).

**Question C.3.2.** Is there a closed formula (for example in terms of the $\Gamma$-function) for the integrals $J(n,r,\ell)$?

Closely related, but probably easier to answer is the following question.

**Question C.3.3.** What are the orders of magnitude of the intrinsic volumes of the semidefinite cone?

Inspired by the motivational experiment in [40, Sec. 3] we formulate the next conjecture for the traditional (SDP) problem, which consists of minimizing/maximizing a linear functional over the intersection of the semidefinite cone with an affine $m$-dimensional subspace. Let a random (SDP)-instance be given in the following way:

- The linear functional, which is to be minimized/maximized is defined by a uniformly random point in $S(\mathrm{Sym}_n)$;
- the affine subspace, which is to be intersected with the semidefinite cone, is given by a uniformly random linear subspace of dimension $m$, and a uniformly random direction from $S(\mathrm{Sym}_n)$.

It is known that the rank $r$ of the optimal solution of an (SDP) problem satisfies with probability 1 the inequalities

$$t(n-r) \leq m \leq t(n)-t(r) \tag{C.26}$$

(cf. [41, Cor. 3.3.4]; cf. also [40, Sec. 3]). It is remarkable, that these inequalities coincide with the inequalities (C.10) for the summands of the intrinsic volumes, if we substitute $m=k+1$. Supported by this observation, but mainly motivated by the geometric picture resulting from the computation in Section C.2 we formulate the following conjecture.

**Conjecture C.3.4.** Let $0 \leq r \leq n$ such that the inequalities (C.26) are satisfied. The probability, that the rank of the optimal solution of a random (SDP) instance (in the sense described above) is $r$, can be given in terms of the integrals $J(n,r,\ell)$.

In the paper [40] the authors analyzed the algebraic degree of semidefinite programming. Avoiding the definition of this notion of degree, we make the following observation. The algebraic degree attributes to each $r$, which satisfies the inequalities (C.26) a positive integer. Using the table of the algebraic degree in [40, Table 2] we may compute the ratios of the algebraic degrees, where $r$ runs over the interval determined by (C.26). The result is shown in Table C.3.

Also the summands of the intrinsic volume $V_k(K_n)$ (assuming $m = k + 1$ for comparison) attribute to each $r$, which satisfies the inequalities (C.26) a positive real number. Table C.4 shows the ratios of these summands.

Comparing these two tables, we see some minor differences, but on the whole, we can observe a notable coherence. This leads us to the final question we would like to formulate.

**Question C.3.5.** What is the relation between the summands of the intrinsic volumes and the algebraic degree of semidefinite programming?

| | $n = 2$ | | $n = 3$ | | $n = 4$ | | $n = 5$ | | $n = 6$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $m$ | $r$ | percent | $r$ | percent | $r$ | percent | $r$ | percent | $r$ | percent |
| 1 | 1 | 100 | 2 | 100 | 3 | 100 | 4 | 100 | 5 | 100 |
| 2 | 1 | 100 | 2 | 100 | 3 | 100 | 4 | 100 | 5 | 100 |
| 3 | | | 2 | 50 | 3 | 61.54 | 4 | 66.67 | 5 | 69.57 |
| | | | 1 | 50 | 2 | 38.46 | 3 | 33.33 | 4 | 30.43 |
| 4 | | | | | 3 | 21.05 | 4 | 30.77 | 5 | 36.36 |
| | | | 1 | 100 | 2 | 78.95 | 3 | 69.23 | 4 | 63.64 |
| 5 | | | | | | | 4 | 7.17 | 5 | 12.5 |
| | | | 1 | 100 | 2 | 100 | 3 | 92.83 | 4 | 87.5 |
| 6 | | | | | | | | | 5 | 2.07 |
| | | | | | 2 | 78.95 | 3 | 89.23 | 4 | 90.67 |
| | | | | | 1 | 21.05 | 2 | 10.77 | 3 | 7.25 |
| 7 | | | | | 2 | 38.46 | 3 | 65 | 4 | 75.22 |
| | | | | | 1 | 61.54 | 2 | 35 | 3 | 24.78 |
| 8 | | | | | | | 3 | 35 | 4 | 51.32 |
| | | | | | 1 | 100 | 2 | 65 | 3 | 48.68 |
| 9 | | | | | | | 3 | 10.77 | 4 | 27.83 |
| | | | | | 1 | 100 | 2 | 89.23 | 3 | 72.17 |

Table C.3: Ratios of the algebraic degrees (based on [40, Table 2]).

| | $n = 2$ | | $n = 3$ | | $n = 4$ | | $n = 5$ | | $n = 6$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $k+1$ | $r$ | percent | $r$ | percent | $r$ | percent | $r$ | percent | $r$ | percent |
| 1 | 1 | 100 | 2 | 100 | 3 | 100 | 4 | 100 | 5 | 100 |
| 2 | 1 | 100 | 2 | 100 | 3 | 100 | 4 | 100 | 5 | 100 |
| 3 | | | 2 | 50 | 3 | 59.78 | 4 | 64.80 | 5 | 65.71 |
| | | | 1 | 50 | 2 | 40.22 | 3 | 36.20 | 4 | 34.29 |
| 4 | | | | | 3 | 18.11 | 4 | 25.76 | 5 | 30.36 |
| | | | 1 | 100 | 2 | 81.89 | 3 | 74.24 | 4 | 69.64 |
| 5 | | | | | | | 4 | 5.12 | 5 | 8.38 |
| | | | 1 | 100 | 2 | 100 | 3 | 95.88 | 4 | 91.62 |
| 6 | | | | | | | | | 5 | 1.18 |
| | | | | | 2 | 81.86 | 3 | 90.19 | 4 | 91.87 |
| | | | | | 1 | 18.14 | 2 | 9.81 | 3 | 6.95 |
| 7 | | | | | 2 | 40.19 | 3 | 65.72 | 4 | 75.19 |
| | | | | | 1 | 59.81 | 2 | 34.28 | 3 | 24.81 |
| 8 | | | | | | | 3 | 34.44 | 4 | 49.26 |
| | | | | | 1 | 100 | 2 | 65.56 | 3 | 50.74 |
| 9 | | | | | | | 3 | 9.76 | 4 | 24.61 |
| | | | | | 1 | 100 | 2 | 90.24 | 3 | 75.39 |

Table C.4: Ratios of the summands in $V_{k+1}(K_n)$.

# Appendix D

# On the distribution of the principal angles

The line of research we followed in Chapter 6 may broadly be summarized by saying that we tried to understand the relative positioning of a cap and a subsphere (cf. Section 3.2; in particular Figure 3.4). In this chapter we will discuss the relative positioning of two subspheres. This is a classic field tracing back to Jordan [35], who first defined the principal angles between two subspaces. The goal of this chapter is to further explain the relation between singular values/vectors and principal angles/directions, and to derive the distribution of the principal angles. The results are not new, but as we avoid results from multivariate statistics (except for a formula for the hypergeometric function of scalar matrix argument) we will derive a homogeneous picture of the principal angles. We also believe that a good understanding of these concepts in the differential geometric setting is essential for an aspired (good) smoothed analysis of the Grassmann condition.

## D.1 Singular vectors

Recall that the singular value decomposition of a matrix $A \in \mathbb{R}^{m\times n}$, $m \leq n$, asserts that there exist orthogonal matrices $Q_1 \in O(m)$, $Q_2 \in O(n)$, and uniquely determined nonnegative constants $\sigma_1 \geq \ldots \geq \sigma_m \geq 0$, such that

$$A = Q_1 \cdot \begin{pmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_m & 0 & \cdots & 0 \end{pmatrix} \cdot Q_2^T \tag{D.1}$$

(cf. Theorem 2.1.2). While the singular values of a matrix are unique, this is not the case for the tuple $(Q_1, Q_2)$ such that (D.1) holds. To characterize all singular value decompositions of a single matrix $A$ we make the following definition.

**Definition D.1.1.** For $A \in \mathbb{R}^{m\times n}$ and $\sigma > 0$ we define

$$\begin{aligned} \mathrm{SV}_A(\sigma) &:= \{(v,u) \in \mathbb{R}^n \times \mathbb{R}^m \mid Av = \sigma u\,,\ A(v^\perp) \subseteq u^\perp\,,\ \|v\| = \|u\|\}\,, \\ \mathrm{SV}_A(0) &:= \ker A \times (\operatorname{im} A)^\perp\,. \end{aligned}$$

We call $\mathrm{SV}_A(\sigma)$ the set of *singular vectors* corresponding to $\sigma$. We furthermore denote the projections of $\mathrm{SV}_A(\sigma)$ onto the first and onto the second component by

$$\mathrm{SV}_A^1(\sigma) := \Pi_1(\mathrm{SV}_A(\sigma))\,, \quad \mathrm{SV}_A^2(\sigma) := \Pi_2(\mathrm{SV}_A(\sigma))\,,$$

where $\Pi_1\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\Pi_2\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ denote the projections.

**Proposition D.1.2.** *For $A \in \mathbb{R}^{m\times n}$ and $\sigma \geq 0$ the set of singular vectors is a linear subspace of $\mathbb{R}^n \times \mathbb{R}^m$. Furthermore, if $\sigma_1 \neq \sigma_2$ then $\mathrm{SV}^i_A(\sigma_1) \cap \mathrm{SV}^i_A(\sigma_2) = \{0\}$ for $i = 1, 2$.*

*Proof.* Clearly, the set $\mathrm{SV}_A(0) = \ker A \times (\operatorname{im} A)^\perp$ is a vector space. We show the case $\sigma > 0$ in several steps. More precisely, we will show the following properties about $\mathrm{SV}_A(\sigma)$:

1. $\mathrm{SV}_A(\sigma)$ is closed,
2. if $(v, u) \in \mathrm{SV}_A(\sigma)$ and $c \in \mathbb{R}$ then $(cv, cu) \in \mathrm{SV}_A(\sigma)$,
3. if $(v_1, u_1), (v_2, u_2) \in \mathrm{SV}_A(\sigma)$ with $\|v_1\| = \|v_2\| = 1$, then $(v_1 + v_2, u_1 + u_2) \in \mathrm{SV}_A(\sigma)$.

If we have shown these properties then we are done, as for $(v_1, u_1), (v_2, u_2) \in \mathrm{SV}_A(\sigma)$ the vector $(v_1 + v_2, u_1 + u_2)$ can be approximated by elements in $\mathrm{SV}_A(\sigma)$ using properties 2 and 3, and thus lies in $\mathrm{SV}_A(\sigma)$ by property 1.

Properties 1 and 2 are verified easily. In order to show property 3, we assume that $(v_1, u_1), (v_2, u_2) \in \mathrm{SV}_A(\sigma)$ with $\|v_1\| = \|v_2\| = 1$. Trivially, we have $A(v_1 + v_2) = \sigma(u_1 + u_2)$. It remains to show that $\langle x, v_1 + v_2\rangle = 0 \Rightarrow \langle Ax, u_1 + u_2\rangle = 0$, and $\|v_1 + v_2\| = \|u_1 + u_2\|$.

Note that the map $x \mapsto \langle x, v_i\rangle v_i$ is the orthogonal projection onto $\mathbb{R}v_i$. From $A(v_i^\perp) \subseteq u_i^\perp$ we get $\langle Ax, u_i\rangle = \langle x, v_i\rangle \cdot \langle Av_i, u_i\rangle = \sigma\langle x, v_i\rangle$, $i = 1, 2$. This implies $\langle Ax, u_1 + u_2\rangle = \langle \sigma x, v_1 + v_2\rangle$, and thus

$$\begin{aligned}\|u_1 + u_2\|^2 &= \langle u_1 + u_2, u_1 + u_2\rangle = \langle \sigma^{-1} \cdot A(v_1 + v_2), u_1 + u_2\rangle \\ &= \langle v_1 + v_2, v_1 + v_2\rangle = \|v_1 + v_2\|^2 .\end{aligned}$$

Furthermore, if $\langle x, v_1 + v_2\rangle = 0$ then $\langle Ax, u_1 + u_2\rangle = \langle \sigma x, v_1 + v_2\rangle = 0$, which finishes the proof of property 3.

For the additional claim let $\sigma_1 \neq \sigma_2$, and let $(v_i, u_i) \in \mathrm{SV}_A(\sigma_i)$, $i = 1, 2$. We need to show that $v_1 \neq v_2$ and $u_1 \neq u_2$. For this we first treat the case $\sigma_2 = 0$. If $u_1 = 0$ then $v_1 = 0$, so let $u_1 \neq 0$. As $v_2 \in \ker A$ and $Av_1 = \sigma_1 u_1 \neq 0$, we have $v_1 \neq v_2$. Moreover, since $u_1 \in \operatorname{im} A$ and $u_2 \in (\operatorname{im} A)^\perp$, we have $u_1 \neq u_2$.

Finally, we assume $\sigma_1, \sigma_2 > 0$, $\sigma_1 \neq \sigma_2$. W.l.o.g. we may also assume $\|v_1\| = \|v_2\| = 1$. As $\|Av_i\| = \sigma_i\|u_i\| = \sigma_i$, $i = 1, 2$, we get $v_1 \neq v_2$. In particular, we have $\langle v_1, v_2\rangle < 1$. Assuming $\sigma_1 > \sigma_2$, we get $\langle u_1, u_2\rangle = \sigma_1^{-1}\langle Av_1, u_2\rangle = \frac{\sigma_2}{\sigma_1}\langle v_1, v_2\rangle < 1$, and thus $u_1 \neq u_2$. □

In the following proposition we will see that the spaces of left and right singular vectors characterize all singular value decompositions of $A$.

**Proposition D.1.3.** *For $A \in \mathbb{R}^{m\times n}$, $m \leq n$, the linear spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ decompose into orthogonal sums of the spaces of singular vectors of $A$, i.e.,*

$$\mathbb{R}^n = \bigoplus_{\sigma\geq 0} \mathrm{SV}^1_A(\sigma) , \quad \mathbb{R}^m = \bigoplus_{\sigma\geq 0} \mathrm{SV}^2_A(\sigma) , \tag{D.2}$$

*and $\mathrm{SV}^1_A(\sigma) \perp \mathrm{SV}^1_A(\sigma')$ and $\mathrm{SV}^2_A(\sigma) \perp \mathrm{SV}^2_A(\sigma')$ for $\sigma \neq \sigma'$.*

*Furthermore, if* $\sigma_1 \geq \ldots \geq \sigma_m$ *denote the singular values of* $A$*, then* $Q_1 \in O(m)$ *and* $Q_2 \in O(n)$ *define a SVD of* $A$ *as in* (2.2) *iff*

$$\begin{aligned} (v_i, u_i) &\in \mathrm{SV}_A(\sigma_i)\ , \quad \textit{for } i = 1, \ldots, m\ , \\ (v_j, 0) &\in \mathrm{SV}_A(0) \quad , \quad \textit{for } j = m+1, \ldots, n\ , \end{aligned} \tag{D.3}$$

*where* $v_1, \ldots, v_n$ *denote the columns of* $Q_2$ *and* $u_1, \ldots, u_m$ *denote the columns of* $Q_1$.

*Proof.* We first show that $Q_1 \in O(m)$ and $Q_2 \in O(n)$ determine a SVD of $A$ as in (2.2) iff the relations in (D.3) hold. Let $k := \mathrm{rk}(A)$ so that $\sigma_1 \geq \ldots \geq \sigma_k(A) > 0$ and $\sigma_{k+1} = \ldots = \sigma_m = 0$. If $Q_1 \in O(m)$ and $Q_2 \in O(n)$ satisfy (2.2), then denoting the columns of $Q_1$ by $u_1, \ldots, u_m$ and the columns of $Q_2$ by $v_1, \ldots, v_n$, we have

$$\begin{aligned} Av_i &= \sigma_i u_i\ , \quad \text{for } i = 1, \ldots, k\ , \\ Av_i &= 0 \quad\ \ , \quad \text{for } i = k+1, \ldots, m\ , \\ Av_j &= 0 \quad\ \ , \quad \text{for } j = m+1, \ldots, n\ , \end{aligned} \tag{D.4}$$

These relations along with the property $Q_1 \in O(m)$ and $Q_2 \in O(n)$ are equivalent to (D.3).

On the other hand, if (D.3) holds for some $Q_1 \in O(m)$ and $Q_2 \in O(n)$, then (D.4) holds, which may be reformulated as

$$AQ_2 = Q_1 \cdot \begin{pmatrix} S & & 0 \\ & 0 & 0 \end{pmatrix}\ ,$$

where $S := \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$. In particular, $Q_1$ and $Q_2$ determine a SVD of $A$ as in (2.2).

As for the claim about the orthogonal decompositions, note that from Theorem 2.1.2 we have the existence of a SVD. The columns $u_1, \ldots, u_m$ of $Q_1$ form an orthonormal basis of $\mathbb{R}^m$, and the columns $v_1, \ldots, v_n$ of $Q_2$ form an orthonormal basis of $\mathbb{R}^n$. The memberships $v_j \in \mathrm{SV}_A^1(\sigma_j)$ for $j = 1, \ldots, n$ and $u_i \in \mathrm{SV}_A^2(\sigma_i)$ for $i = 1, \ldots, m$, along with the properties of the spaces of singular vectors shown in Proposition D.1.2, imply the decompositions of $\mathbb{R}^m$ and $\mathbb{R}^n$ as stated in (D.2). □

**Corollary D.1.4.**

1. *For* $\sigma > 0$ *the scaled restriction of the linear map defined by* $A$*, given by*
$$\mathrm{SV}_A^1(\sigma) \to \mathrm{SV}_A^2(\sigma)\ , \quad v \mapsto \sigma^{-1} \cdot Av\ ,$$
*is bijective and preserves the scalar product. In particular, if* $\sigma > 0$ *then* $\|v\| = \|u\|$ *for all* $(v, u) \in \mathrm{SV}_A(\sigma)$.
2. *The spaces of singular vectors corresponding to the maximum singular vector have the simpler characterization* $\mathrm{SV}_A^1(\|A\|) = \{v \mid \|Av\| = \|A\| \cdot \|v\|\}$ *and* $\mathrm{SV}_A(\|A\|) = \{(v, u) \mid Av = \|A\| \cdot u\}$.

*Proof.* Let the notation be as in Proposition D.1.3.

1. We define the index set $I \subseteq \{1, \ldots, m\}$ via $I = \{i \mid \sigma_i = \sigma\}$. Then the columns $\{v_i \mid i \in I\}$ of $Q_2$ form an orthonormal basis of $\mathrm{SV}_A^1(\sigma)$, and the columns $\{u_i \mid i \in I\}$ of $Q_1$ form an orthonormal basis of $\mathrm{SV}_A^2(\sigma)$. The claim follows from the fact that these orthonormal bases are mapped onto each other by the above defined scaled restriction $v \mapsto \sigma^{-1} \cdot Av$.
2. For $A = 0$ the claim is trivial, so let us assume $\|A\| > 0$. It suffices to show the first claim, i.e., that $\|Av\| = \|A\| \cdot \|v\|$ implies $v \in \mathrm{SV}_A^1(\|A\|)$. From the orthogonal decomposition of $\mathbb{R}^n$ in (D.2), and since $\|Av\| = \sigma \cdot \|v\| < \|A\| \cdot \|v\|$ for $v \in \mathrm{SV}_A^1(\sigma)$, $\sigma < \|A\|$, we get $v \in \mathrm{SV}_A^1(\|A\|)$ iff $\|Av\| = \|A\| \cdot \|v\|$, which proves the claim. □

## D.2 Principal directions

In this section we will further explain the relation between principal angles/directions and singular values/vectors. Among other things, we will show that the principal angles determine the relative positions of subspaces. In other words, we will show that the principal angles are the invariants of the product $\mathrm{Gr}_{n,m} \times \mathrm{Gr}_{n,M}$ under the canonical action of the orthogonal group $O(n)$. We will also fill the gap in the description of the global properties of $\mathrm{Gr}_{n,m}$ that we left open in Section 5.4 (cf. Proposition 5.4.7).

Recall from Definition 5.4.1 that for $\mathcal{W}_1 \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2 \in \mathrm{Gr}_{n,M}$, where $1 \leq m \leq M \leq n-1$, the principal angles $\alpha_1 \leq \ldots \leq \alpha_m \in [0, \frac{\pi}{2}]$ between $\mathcal{W}_1$ and $\mathcal{W}_2$ are defined in the following way: Let $X_1 \in \mathbb{R}^{n\times m}$ and $X_2 \in \mathbb{R}^{n\times M}$ be such that the columns of $X_i$ form an orthonormal basis of $\mathcal{W}_i$, $i = 1, 2$. Then the principal angles are the arccosines of the singular values of the matrix $X_1^T X_2 \in \mathbb{R}^{m\times M}$, i.e.,

$$X_1^T X_2 = Q_1 \begin{pmatrix} \cos(\alpha_1) & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \cos(\alpha_m) & 0 & \cdots & 0 \end{pmatrix} \cdot Q_2^T \ ,$$

where $Q_1 \in O(m)$ and $Q_2 \in O(M)$. In the following definition we will define the concept of principal directions.

**Definition D.2.1.** Let $\mathcal{W}_1 \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2 \in \mathrm{Gr}_{n,M}$, where $1 \leq m \leq M \leq n-1$, and let $X_1 \in \mathbb{R}^{n\times m}$ and $X_2 \in \mathbb{R}^{n\times M}$ be such that the columns of $X_i$ form an orthonormal basis of $\mathcal{W}_i$, $i = 1, 2$. The vector space of *principal directions* of $\mathcal{W}_1$ and $\mathcal{W}_2$ to the angle $\alpha \in [0, \frac{\pi}{2}]$ is defined by

$$\mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha) := \{(X_1 u, X_2 v) \mid (u, v) \in \mathrm{SV}_{X_2^T X_1}(\cos\alpha)\} \subseteq \mathbb{R}^n \times \mathbb{R}^n \ .$$

Furthermore, we denote the projections of $\mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha)$ onto the first and onto the second component by

$$\mathrm{PD}^1_{\mathcal{W}_1,\mathcal{W}_2}(\alpha) := \Pi_1(\mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha)) \ , \quad \mathrm{PD}^2_{\mathcal{W}_1,\mathcal{W}_2}(\alpha) := \Pi_2(\mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha)) \ ,$$

where $\Pi_1 \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\Pi_2 \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ denote the projections.

**Proposition D.2.2.** *Let the notation be as in Definition D.2.1. Then neither the principal angles nor the vector spaces of principal directions depend on the specific choice of* $X_1$ *and* $X_2$.

*Proof.* Let $X_1'$ and $X_2'$ be a different choice of orthonormal bases of $\mathcal{W}_1$ and $\mathcal{W}_2$. Then there exist $Q_1 \in O(m)$ and $Q_2 \in O(M)$ such that $X_1' = X_1 \cdot Q_1$ and $X_2' = X_2 \cdot Q_2$. In particular, we have $(X_2')^T X_1' = Q_2^T X_2^T X_1 Q_1$, which has the same singular values as $X_2^T X_1$.

Furthermore, we have that $(u, v)$ is a pair of singular vectors for $X_2^T X_1$ iff $(u', v') := (Q_1^T u, Q_2^T v)$ is a pair of singular vectors for $(X_2')^T X_1'$. But then

$$X_1' u' = X_1 Q_1 Q_1^T u = X_1 u \quad \text{and} \quad X_2' v' = X_2 Q_2 Q_2^T v = X_2 v \ ,$$

which finishes the proof. □

The following proposition summarizes the most important geometric properties of the spaces of principal directions.

**Proposition D.2.3.** *Let the notation be as in Definition D.2.1 except that we write* $\mathrm{PD}(\alpha)$ *instead of* $\mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha)$ *to ease the notation.*

1. *We have orthogonal decompositions*

$$\mathcal{W}_1 = \bigoplus_{\alpha\in[0,\frac{\pi}{2}]} \mathrm{PD}^1(\alpha)\ , \quad \mathcal{W}_2 = \bigoplus_{\alpha\in[0,\frac{\pi}{2}]} \mathrm{PD}^2(\alpha)\ .$$

2. *Let $\alpha < \frac{\pi}{2}$. If $x \in \mathrm{PD}^1(\alpha)$ then there exists a unique $y \in \mathrm{PD}^2(\alpha)$ such that $(x, y) \in \mathrm{PD}(\alpha)$. Furthermore, if $(x, y) \in \mathrm{PD}(\alpha)$ then $\|x\| = \|y\|$.*

3. *For $\alpha = 0$:*

$$\mathrm{PD}(0) = \{(x, x) \mid x \in \mathcal{W}_1 \cap \mathcal{W}_2\}\ .$$

   *In particular, we have $\mathrm{PD}^1(\alpha) = \mathrm{PD}^2(\alpha) = \mathcal{W}_1 \cap \mathcal{W}_2$.*

4. *For $\alpha = \frac{\pi}{2}$:*

$$\mathrm{PD}(\tfrac{\pi}{2}) = \mathcal{W}_1 \cap \mathcal{W}_2^\perp \ \times\ \mathcal{W}_1^\perp \cap \mathcal{W}_2\ .$$

   *In particular, we have $\mathrm{PD}^1(\alpha) = \mathcal{W}_1 \cap \mathcal{W}_2^\perp$ and $\mathrm{PD}^2(\alpha) = \mathcal{W}_1^\perp \cap \mathcal{W}_2$.*

5. *For $0 < \alpha < \frac{\pi}{2}$: Let $(x_1, y_1), \ldots, (x_k, y_k) \in \mathrm{PD}(\alpha)$ such that $x_1, \ldots, x_k$ form an orthonormal basis of $\mathrm{PD}^1(\alpha)$, and let $L_i := \mathrm{lin}\{x_i, y_i\}$, $i = 1, \ldots, k$. Then we get:*
   (a) *$y_1, \ldots, y_k$ is an orthonormal basis of $\mathrm{PD}^2(\alpha)$,*
   (b) *$\langle x_i, y_i\rangle = \cos(\alpha)$, in particular $\dim L_i = 2$ for $i = 1, \ldots, k$,*
   (c) *$L_i \perp L_j$ for $i \neq j$.*

*Proof.* We process the statements one by one:

1. This follows directly from the orthogonal decompositions (D.2) in Proposition D.1.3.

2. Let $\sigma := \cos(\alpha) > 0$. For $x \in \mathrm{PD}^1(\alpha)$ there exists a unique $u \in \mathrm{SV}^1_{X_2^T X_1}(\sigma)$ such that $x = X_1 u$. For $v := \sigma^{-1} \cdot X_2^T X_1 u$, and only for this choice of $v$, we have $(u, v) \in \mathrm{SV}_{X_2^T X_1}(\sigma)$. Therefore $y = X_2 v$ is uniquely determined by the property $(x, y) \in \mathrm{PD}(\alpha)$. Furthermore, we have $\|X_1 u\| = \|u\|$ and $\|X_2 v\| = \|v\|$, and by Corollary D.1.4 part (1) we also have $\|u\| = \|v\|$.

3. Let $(x, y) \in \mathrm{PD}(0)$, with $x = X_1 u$ and $y = X_2 v$, $(u, v) \in \mathrm{SV}_{X_2^T X_1}(1)$. Then we have $\Pi_{\mathcal{W}_2}(x) = X_2 X_2^T X_1 u = X_2 v = y$, and since $\|x\| = \|y\|$ by part (2), we get $x = y$.

   On the other hand, if $x \in \mathcal{W}_1 \cap \mathcal{W}_2 \setminus \{0\}$, then there exist $u \in \mathbb{R}^m \setminus \{0\}$ and $v \in \mathbb{R}^M \setminus \{0\}$ such that $x = X_1 u$ and $x = X_2 v$. Therefore $X_2^T X_1 u = v$, and since $\|u\| = \|v\| > 0$, we get $\|X_2^T X_1\| \geq 1$. So we have $\|X_2^T X_1\| = 1$ and by Corollary D.1.4 part (2) we get $\mathrm{SV}_{X_2^T X_1}(1) = \{(u', v') \mid X_2^T X_1 u' = v'\}$. In particular, $(u, v) \in \mathrm{SV}_{X_2^T X_1}(1)$ and $(x, x) \in \mathrm{PD}(0)$.

4. Recall that $\mathrm{SV}_{X_2^T X_1}(0) = \ker X_2^T X_1 \times (\mathrm{im}\, X_2^T X_1)^\perp$. For the first component we compute

$$\begin{aligned} X_2^T X_1 u = 0 &\iff X_2 X_2^T X_1 u = 0 \\ &\iff X_1 u \in \mathcal{W}_2^\perp\ . \end{aligned}$$

For the second component we compute

$$\begin{aligned} v \in (\operatorname{im} X_2^T X_1)^\perp &\iff \langle v, X_2^T X_1 u\rangle = 0 \quad \forall u \in \mathbb{R}^m \\ &\iff \langle X_2 v, X_2 X_2^T X_1 u\rangle = 0 \quad \forall u \in \mathbb{R}^m \\ &\iff X_2 v \in (\Pi_{\mathcal{W}_2}(\mathcal{W}_1))^\perp \\ &\iff X_2 v \in \mathcal{W}_1^\perp \ . \end{aligned}$$

5. The vectors $y_1, \ldots, y_k$ form an orthonormal basis of $\mathrm{PD}^2(\alpha)$ by Corollary D.1.4 part (1). Furthermore, denoting $\sigma := \cos(\alpha)$, we compute

$$\langle x_i, y_j\rangle = \langle \Pi_{\mathcal{W}_2}(x_i), y_j\rangle = \langle X_2 X_2^T X_1 u_i, y_j\rangle = \langle X_2 \cdot \sigma v_i, X_2 v_j\rangle = \sigma \cdot \delta_{ij} \ ,$$

where $\delta_{ij}$ denotes the Kronecker-delta. This shows the claims in part (b) and (c). □

**Corollary D.2.4.** *Let $\mathcal{W}_1 \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2 \in \mathrm{Gr}_{n,M}$, where $1 \le m \le M \le n-1$. Furthermore, let $\alpha_1 \le \ldots \le \alpha_m$ denote the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$, and let $d := \mathcal{W}_1 \cap \mathcal{W}_2$ and $d^\perp := \mathcal{W}_1 \cap \mathcal{W}_2^\perp$, so that*

$$0 = \alpha_1 = \ldots = \alpha_d < \alpha_{d+1} \le \ldots \le \alpha_{m-d^\perp} < \alpha_{m-d^\perp+1} = \ldots = \alpha_m = \tfrac{\pi}{2} \ .$$

*Then we have an orthogonal decomposition*

$$\begin{aligned} \mathbb{R}^n = (\mathcal{W}_1 \cap \mathcal{W}_2) \;\oplus\; L_{d+1} \;\oplus\; \ldots \;\oplus\; L_{m-d^\perp} \\ \oplus\; (\mathcal{W}_1 \cap \mathcal{W}_2^\perp) \;\oplus\; (\mathcal{W}_1^\perp \cap \mathcal{W}_2) \;\oplus\; (\mathcal{W}_1 + \mathcal{W}_2)^\perp \ , \end{aligned} \tag{D.5}$$

*where $L_i = \mathrm{lin}\{x_i, y_i\}$ with $(x_i, y_i) \in \mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha_i)$ for $i = d+1, \ldots, m-d^\perp$.*

*Proof.* First of all, we have the orthogonal decomposition $\mathbb{R}^n = (\mathcal{W}_1 + \mathcal{W}_2) \oplus (\mathcal{W}_1 + \mathcal{W}_2)^\perp$. Using the decompositions of $\mathcal{W}_1$ and $\mathcal{W}_2$ in Proposition D.2.3 we get the stated decomposition of $\mathcal{W}_1 + \mathcal{W}_2$. □

**Corollary D.2.5.** *Let the notation be as in Definition D.2.1. The nonzero principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$ coincide with the nonzero principal angles between $\mathcal{W}_1^\perp$ and $\mathcal{W}_2^\perp$. More precisely,*

$$\forall\, \alpha \in (0, \tfrac{\pi}{2}] : \dim \mathrm{PD}_{\mathcal{W}_1,\mathcal{W}_2}(\alpha) = \dim \mathrm{PD}_{\mathcal{W}_1^\perp,\mathcal{W}_2^\perp}(\alpha) \ .$$

*Proof.* Let $\mathbb{R}^n$ be decomposed as in (D.5). Denoting $L_i^1 := \mathrm{lin}\{x_i\}$ and $L_i^2 := \mathrm{lin}\{y_i\}$ for $d+1 \le i \le m - d^\perp$, we have

$$\begin{aligned} \mathcal{W}_1 &= (\mathcal{W}_1 \cap \mathcal{W}_2) \;\oplus\; L_{d+1}^1 \;\oplus\; \ldots \;\oplus\; L_{m-d}^1 \;\oplus\; (\mathcal{W}_1 \cap \mathcal{W}_2^\perp) \ , \\ \mathcal{W}_2 &= (\mathcal{W}_1 \cap \mathcal{W}_2) \;\oplus\; L_{d+1}^2 \;\oplus\; \ldots \;\oplus\; L_{m-d}^2 \;\oplus\; (\mathcal{W}_1^\perp \cap \mathcal{W}_2) \ . \end{aligned} \tag{D.6}$$

Let $\hat{L}_i^1$ denote the orthogonal complement of $L_i^1$ in $L_i$, and let $\hat{L}_i^2$ denote the orthogonal complement of $L_i^2$ in $L_i$. The decompositions of $\mathcal{W}_1$ and $\mathcal{W}_2$ in (D.6) imply that its orthogonal complements $\mathcal{W}_1^\perp$ and $\mathcal{W}_2^\perp$ decompose in the following way

$$\begin{aligned} \mathcal{W}_1^\perp &= (\mathcal{W}_1^\perp \cap \mathcal{W}_2) \;\oplus\; \hat{L}_{d+1}^1 \;\oplus\; \ldots \;\oplus\; \hat{L}_{m-d}^1 \;\oplus\; (\mathcal{W}_1 + \mathcal{W}_2)^\perp \ , \\ \mathcal{W}_2^\perp &= (\mathcal{W}_1 \cap \mathcal{W}_2^\perp) \;\oplus\; \hat{L}_{d+1}^2 \;\oplus\; \ldots \;\oplus\; \hat{L}_{m-d}^2 \;\oplus\; (\mathcal{W}_1 + \mathcal{W}_2)^\perp \ . \end{aligned}$$

The angle between $\hat{L}_i^1$ and $\hat{L}_i^2$ is the same as the angle between $L_i^1$ and $L_i^2$. Choosing appropriate bases shows that these are indeed the principal angles and thus finishes the proof. □

The following proposition shows that the principal angles determine the relative positions of subspaces.

**Proposition D.2.6.** *Let $\mathcal{W}_1, \mathcal{W}_1' \in \mathrm{Gr}_{n,m}$ and $\mathcal{W}_2, \mathcal{W}_2' \in \mathrm{Gr}_{n,M}$, where $1 \le m \le M \le n-1$. Furthermore, let $\alpha_1 \le \ldots \le \alpha_m$ denote the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$, and let $\alpha_1' \le \ldots \le \alpha_m'$ denote the principal angles between $\mathcal{W}_1'$ and $\mathcal{W}_2'$. Then*

$$\exists Q \in O(n) : Q(\mathcal{W}_1) = \mathcal{W}_1' \, , \; Q(\mathcal{W}_2) = \mathcal{W}_2' \iff \forall i = 1, \ldots, m : \alpha_i = \alpha_i' \, .$$

*Proof.* If $\mathcal{W}_1' = Q(\mathcal{W}_1)$ and $\mathcal{W}_2' = Q(\mathcal{W}_2)$ for some $Q \in O(n)$, and if $X_1$ and $X_2$ are orthonormal bases for $\mathcal{W}_1$ and $\mathcal{W}_2$ respectively, then $QX_i$ is an orthonormal basis for $\mathcal{W}_i'$, $i = 1, 2$. Since $(QX_1)^T QX_2 = X_1^T Q^T Q X_2 = X_1^T X_2$, the principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$ coincide with the principal angles between $\mathcal{W}_1'$ and $\mathcal{W}_2'$.

On the other hand, if all the principal angles coincide, then there is an orthogonal transformation such that the orthogonal decomposition of $\mathbb{R}^n$ as stated in Corollary D.2.4 transforms into a corresponding decomposition of $\mathbb{R}^n$ with $\mathcal{W}_1$ and $\mathcal{W}_2$ being replaced by $\mathcal{W}_1'$ and $\mathcal{W}_2'$. The decomposition of $\mathcal{W}_1$ and $\mathcal{W}_2$ as given in (D.6) shows that the orthogonal transformation can be chosen in such a way that additionally $\mathcal{W}_1$ goes into $\mathcal{W}_1'$ and $\mathcal{W}_2$ goes into $\mathcal{W}_2'$. □

We finish this section with a supplement to Section 5.4. More precisely, in Proposition 5.4.7 we have listed a couple of global properties of $\mathrm{Gr}_{n,m}$ that we did not prove. We fill this gap with the help of the principal directions.

The statement of Proposition 5.4.7 was the following: Let $\mathcal{W} = [Q] \in \mathrm{Gr}_{n,m}$, and let

$$\mathcal{U} := \left\{ \left[ Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \right] \middle| \, R \in \mathbb{R}^{(n-m)\times m} \, , \; \|R\| < \frac{\pi}{2} \right\} \subset T_{\mathcal{W}} \, \mathrm{Gr}_{n,m} \, ,$$

where $\|R\|$ denotes the operator norm of $R$. Furthermore, let $\overline{\mathcal{U}}$ denote the closure of $\mathcal{U}$, and let $\partial\mathcal{U}$ denote the boundary of $\mathcal{U}$. Then the following holds.

1. The exponential map $\overline{\exp}_{\mathcal{W}}$ is injective on $\mathcal{U}$.
2. The exponential map $\overline{\exp}_{\mathcal{W}}$ is surjective on $\overline{\mathcal{U}}$.
3. If $v = \left[ Q, \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \right] \in \overline{\mathcal{U}}$ and $\mathcal{W}' = \overline{\exp}_{\mathcal{W}}(v)$, then the curve
$$[0,1] \to \mathrm{Gr}_{n,m} \, , \quad \rho \mapsto \overline{\exp}_{\mathcal{W}}(\rho \cdot v)$$
is a shortest length geodesic between $\mathcal{W}$ and $\mathcal{W}'$. In particular, we have $d_{\mathrm{g}}(\mathcal{W}, \mathcal{W}') = \|v\| = \|R\|_F$.
4. For $v \in \partial\mathcal{U}$ we have $\overline{\exp}_{\mathcal{W}}(v) = \overline{\exp}_{\mathcal{W}}(-v)$, so that the injectivity radius of $\mathrm{Gr}_{n,m}$ is $\frac{\pi}{2}$.

*Proof of Proposition 5.4.7.* A general argument, the so-called *Hopf-Rinow Theorem* (cf. for example [19, Thm. I.7.1]), implies that each pair of elements in $\mathrm{Gr}_{n,m}$ can be joined by a geodesic. We will argue this via principal angles/directions, as this will give us the extensive statement of Proposition 5.4.7.

For notational simplicity, let us assume $m \le \frac{n}{2}$; the case $m \ge \frac{n}{2}$ follows analogously. Let $\mathcal{W}, \mathcal{W}' \in \mathrm{Gr}_{n,m}$, and let $\alpha_1, \ldots, \alpha_m$ denote the principal angles between $\mathcal{W}$ and $\mathcal{W}'$. By Corollary D.2.4 and Proposition D.2.3 we can find an orthogonal basis $x_1, \ldots, x_n$ of $\mathbb{R}^n$ such that $\mathcal{W} = \operatorname{lin}\{x_1, \ldots, x_m\}$ and $\mathcal{W}' = \operatorname{lin}\{y_1, \ldots, y_m\}$, where $y_i = \cos(\alpha_i) \cdot x_i + \sin(\alpha_i) \cdot x_{m+i}$ for $i = 1, \ldots, m$. Let $Q \in O(n)$ be such that the $i$th column of $Q$ is $x_i$, and let $C$ and $S$ denote the diagonal matrices $C = \operatorname{diag}(\cos(\alpha_1), \ldots, \cos(\alpha_m))$, $S = \operatorname{diag}(\sin(\alpha_1), \ldots, \sin(\alpha_m))$. Setting

$$v := \left[ Q, \begin{pmatrix} 0 & -A & 0 \\ A & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \in T_{\mathcal{W}} \operatorname{Gr}_{n,m} , \quad \text{with } A = \operatorname{diag}(\alpha_1, \ldots, \alpha_m) , \tag{D.7}$$

it follows that $\mathcal{W}'$ is given by

$$\mathcal{W}' = \overline{\exp}_{\mathcal{W}}(v) = \left[ Q \cdot \begin{pmatrix} C & -S & 0 \\ S & C & 0 \\ 0 & 0 & I_{n-2m} \end{pmatrix} \right] . \tag{D.8}$$

In particular, we have shown that the image of $\overline{\mathcal{U}}$ under the exponential map covers $\operatorname{Gr}_{n,m}$.

On the other hand, we have seen in Lemma 5.4.5 that any geodesic through $\mathcal{W}$ can be brought to the form (D.8). From such a representation one can easily deduce the principal angles and the spaces of principal directions. Furthermore, using the decomposition of $\mathbb{R}^n$ as shown in Corollary D.2.4, the remaining statements of Proposition 5.4.7 are a small but notation-consuming exercise. □

## D.3 Computing the distribution of the principal angles

In this section we will compute the volume of metric balls of the Grassmann manifold. More precisely, we will compute the Normal Jacobian of the parametrization of $\operatorname{Gr}_{n,m}$, which involves the principal angles to a fixed $M$-dimensional subspace of $\mathbb{R}^n$, where $M$ may be any integer within $m \leq M \leq n - m$ (this generalization is needed for the smoothed analysis in Section 7.3). The parametrization will be defined in (D.9). It exploits the geometry of the principal directions as described in Proposition D.2.3.

We derive the Normal Jacobian of the parametrization via differential geometric methods. See [2] for a different approach via multivariate statistics. In this paper (cf. [2, Thm. 1]) the volume of metric balls in $\operatorname{Gr}_{n,m}$ w.r.t. the Hausdorff metric was computed. We will obtain this in Proposition D.3.4.

The following proposition is the main result of this section. It includes the formula of the above mentioned Normal Jacobian. The proof of this proposition is deferred to the end of this section.

**Proposition D.3.1.** *Let $1 \leq m \leq M \leq n-1$ with $m + M \leq n$. Using the identification $\operatorname{St}_{n,m} = O(n)/O(n-m)$ and $\operatorname{Gr}_{n,m} = O(n)/(O(m) \times O(n-m))$, let*

$$\varphi\colon \operatorname{St}_{M,m} \times \operatorname{St}_{n-M,m} \times \mathbb{R}^m \to \operatorname{Gr}_{n,m}$$

$$([Q_1],[Q_2],v) \mapsto \left[ \left( \begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array} \right) \cdot \left( \begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array} \right) \right] , \tag{D.9}$$

*where* $C = \operatorname{diag}(\cos(v_1), \dots, \cos(v_m))$, $S = \operatorname{diag}(\sin(v_1), \dots, \sin(v_m))$. *The Normal Jacobian of* $\varphi$ *is given by*

$$\operatorname{ndet}(D_{([Q_1],[Q_2],v)}\varphi) = \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^{m} \left|\sin(v_i)^{n-M-m} \cdot \cos(v_i)^{M-m}\right| \ .$$

*In particular, for* $M = m$ *we have*

$$\operatorname{ndet}(D_{(Q_1,[Q_2],v)}\varphi) = \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^{m} \left|\sin(v_i)^{n-2m}\right| \ .$$

Note that the map $\varphi$ as defined in (D.9) is well-defined, which is seen by the following small computation. For $Q_1' \in O(M)$ such that $[Q_1'] = [Q_1]$ in $\mathrm{St}_{M,m}$ and for $Q_2' \in O(n-M)$ such that $[Q_2'] = [Q_2]$ in $\mathrm{St}_{n-M,m}$, we can find $\bar{Q}_1 \in O(M-m)$ and $\bar{Q}_2 \in O(n-M-m)$ such that

$$Q_1' = Q_1 \cdot \begin{pmatrix} I_m & 0 \\ 0 & \bar{Q}_1 \end{pmatrix} \ , \quad Q_2' = Q_2 \cdot \begin{pmatrix} I_m & 0 \\ 0 & \bar{Q}_2 \end{pmatrix} \ .$$

We thus get

$$\begin{aligned}
&\left(\begin{array}{c|c} Q_1' & 0 \\ \hline 0 & Q_2' \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \\
&= \left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \cdot \left(\begin{array}{cc|cc} I_m & & 0 & \\ & \bar{Q}_1 & & 0 \\ \hline 0 & & I_m & \\ & 0 & & \bar{Q}_2 \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \ .
\end{aligned}$$

The second product on the right-hand side of the above equation is commutative, so that we get in $\mathrm{Gr}_{n,m} = O(n)/(O(m) \times O(n-m))$

$$\begin{aligned}
&\left[\left(\begin{array}{c|c} Q_1' & 0 \\ \hline 0 & Q_2' \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right)\right] \\
&= \left[\left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \cdot \left(\begin{array}{cc|cc} I_m & & 0 & \\ & \bar{Q}_1 & & 0 \\ \hline 0 & & I_m & \\ & 0 & & \bar{Q}_2 \end{array}\right)\right] \\
&= \left[\left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right)\right] \ .
\end{aligned}$$

This shows the well-definedness of $\varphi$ as defined in (D.9).

Let us first use the result to compute the volume of Grassmannian balls. We say that a metric $d_* \colon \mathrm{Gr}_{n,m} \times \mathrm{Gr}_{n,m} \to \mathbb{R}$ is induced by the symmetric function

$f_*\colon [0,\frac{\pi}{2}]^m \to \mathbb{R}$, i.e., $f_*(\alpha_1,\ldots,\alpha_m) = f_*(\alpha_{P(1)},\ldots,\alpha_{P(m)})$ for any permutation $P$, if

$$d_*(\mathcal{W}_1,\mathcal{W}_2) = f_*(\alpha) \;,$$

for $\mathcal{W}_1,\mathcal{W}_2 \in \mathrm{Gr}_{n,m}$, where $\alpha \in \mathbb{R}^m$ is the vector of principal angles between $\mathcal{W}_1$ and $\mathcal{W}_2$. By Proposition D.2.6 any orthogonally invariant metric on $\mathrm{Gr}_{n,m}$ is induced by such a symmetric function.

For the computation of the volume of Grassmannian balls we define the following transformation

$$T\colon [0,\tfrac{\pi}{2}]^m \to [0,1]^m \;, \quad (v_1,\ldots,v_m) \mapsto (\sin(v_1)^2,\ldots,\sin(v_m)^2) \;.$$

**Proposition D.3.2.** *Let $d_*$ be a metric on $\mathrm{Gr}_{n,m}$ induced by $f_*\colon [0,\frac{\pi}{2}]^m \to \mathbb{R}$. For $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$ and $\beta \geq 0$ let $B_*(\mathcal{W}_0,\beta)$ denote the ball of radius $\beta$ around $\mathcal{W}_0$ w.r.t. the metric $d_*$. The relative volume of $B_*(\mathcal{W}_0,\beta)$ is given by*

$$\begin{aligned}\mathrm{rvol}\, B_*(\mathcal{W}_0,\beta) = \frac{\pi^{\frac{m}{2}}}{m!} &\cdot \prod_{i=0}^{m-1} \frac{\Gamma(\frac{n-i}{2})}{\Gamma(\frac{i+1}{2})^2 \cdot \Gamma(\frac{n-m-i}{2})} \\ &\cdot \int\limits_{T(K_*(\beta))} \prod_{i<j} |s_i - s_j| \cdot \prod_{i=1}^{m} s_i^{\frac{n-2m-1}{2}} \cdot (1-s_i)^{-\frac{1}{2}}\, ds \;,\end{aligned}$$

*where $K_*(\beta) := \big\{v \in [0,\frac{\pi}{2}]^m \mid f_*(v) < \beta\big\}$.*

*Proof.* For generic $\mathcal{W} \in \mathrm{Gr}_{n,m}$ the principal angles $\alpha_1,\ldots,\alpha_m$ between $\mathcal{W}$ and $\mathcal{W}_0$ are mutually distinct and lie in the open interval $(0,\frac{\pi}{2})$, i.e., we have

$$0 < \alpha_1 < \ldots < \alpha_m < \frac{\pi}{2} \;.$$

To see that this holds generically, note that $\alpha_1 = 0$ iff $\mathcal{W} \cap \mathcal{W}_0 \neq \{0\}$, and $\alpha_m = \frac{\pi}{2}$ iff $\mathcal{W} \cap \mathcal{W}_0^\perp \neq \{0\}$. Furthermore, $\alpha_i = \alpha_j$ for some $i \neq j$ iff the singular values of the matrix $X_0^T X$, where $X_0, X \in \mathbb{R}^{n\times m}$ are such that the columns of $X_0$ form an orthonormal basis of $\mathcal{W}_0$ and the columns of $X$ form an orthonormal basis of $\mathcal{W}$, are not all mutually distinct. These are finitely many events with probability 0 and thus altogether have probability 0.

When the principal angles are strictly increasing then the principal directions are uniquely determined up to multiplication by $\pm 1$ (cf. Section D.2). So if we consider the function

$$\begin{aligned}\varphi\colon O(m) \times \mathrm{St}_{n-m,m} \times K_*^<(\beta) &\to \mathrm{Gr}_{n,m} \\ (Q_1,[Q_2],v) &\mapsto \left[\left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \cdot \left(\begin{array}{c|cc} C & -S & \\ \hline S & C & \\ & & I_{n-2m} \end{array}\right)\right] \;,\end{aligned}$$

where $K_*^<(\beta) := \big\{v \in (0,\frac{\pi}{2})^m \mid f_*(v) < \beta \,,\; v_1 < \ldots < v_m\big\}$, then we will get a $2^m$-fold covering of the image of $\varphi$, which lies dense in $B_*(\mathcal{W}_0,\beta)$, the ball w.r.t. $d_*$ of radius $\beta$ around $\mathcal{W}_0$. Applying the smooth coarea formula, Proposition D.3.1 (for

$M = m$) implies that

$$\begin{aligned}
\operatorname{vol} B_*(\mathcal{W}_0,\beta) &= \frac{1}{2^m}\cdot \int\limits_{O(m)}\ \int\limits_{\mathrm{St}_{n-m,m}}\ \int\limits_{K_*^<(\beta)} \operatorname{ndet}_{(Q_1,[Q_2],v)}(\varphi)\, dv\, d[Q_2]\, dQ_1 \\
&= \frac{\operatorname{vol} O(m)\cdot \operatorname{vol}\mathrm{St}_{n-m,m}}{2^m} \\
&\quad\cdot \int\limits_{K_*^<(\beta)} \prod_{i<j} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^m \sin(v_i)^{n-2m}\, dv\ .
\end{aligned}$$

We may change the integration over $K_*^<(\beta)$ to an integration over $K_*(\beta)$ if we divide the result by $m!$. Dividing by $\operatorname{vol}\mathrm{Gr}_{n,m}$ to get the relative volume, we get

$$\begin{aligned}
\operatorname{rvol} B_*(\mathcal{W}_0,\beta) &= \frac{\operatorname{vol} O(m)\cdot \operatorname{vol}\mathrm{St}_{n-m,m}}{\operatorname{vol}\mathrm{Gr}_{n,m}\cdot m!\cdot 2^m} \\
&\quad\cdot \int\limits_{K_*(\beta)} \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^m \sin(v_i)^{n-2m}\, dv\ .
\end{aligned}$$

Using Proposition 5.2.1, (5.19), and (5.23), the constant computes as

$$\begin{aligned}
&\frac{\operatorname{vol} O(m)\cdot \operatorname{vol}\mathrm{St}_{n-m,m}}{\operatorname{vol}\mathrm{Gr}_{n,m}\cdot m!\cdot 2^m} \\
&= \frac{2^m\cdot \pi^{\frac{m^2+m}{4}}}{\prod_{d=1}^m \Gamma(\frac{d}{2})}\cdot \frac{2^m\cdot \pi^{\frac{2nm-3m^2+m}{4}}}{\prod_{d=n-2m+1}^{n-m}\Gamma(\frac{d}{2})}\cdot \frac{\prod_{d=1}^m \Gamma(\frac{n-m+d}{2})}{\pi^{\frac{m(n-m)}{2}}\cdot \prod_{d=1}^m \Gamma(\frac{d}{2})}\cdot \frac{1}{m!\cdot 2^m} \\
&= \frac{2^m\cdot \pi^{\frac{m}{2}}}{m!}\cdot \prod_{i=0}^{m-1} \frac{\Gamma(\frac{n-i}{2})}{\Gamma(\frac{i+1}{2})^2\cdot \Gamma(\frac{n-m-i}{2})}\ .
\end{aligned}$$

It remains to change the integration over $K_*(\beta)$ to an integration over $T(K_*(\beta))$. Substituting $v$ by $s := T(v) = (\sin(v_1)^2,\ldots,\sin(v_m)^2)$ yields

$$\begin{aligned}
&\int\limits_{K_*(\beta)} \prod_{i<j} \left|\sin^2 v_i\cdot \cos^2 v_j - \cos^2 v_i\cdot \sin^2 v_j\right| \cdot \prod_{i=1}^m \sin(v_i)^{n-2m}\, dv \\
&= \int\limits_{T(K_*(\beta))} \prod_{i=1}^m \frac{1}{2\sqrt{s_i\cdot(1-s_i)}}\cdot \prod_{i<j}|s_i-s_j|\cdot \prod_{i=1}^m s_i^{\frac{n-2m}{2}}\, ds \\
&= 2^{-m}\cdot \int\limits_{T(K_*(\beta))} \prod_{i<j}|s_i-s_j|\cdot \prod_{i=1}^m s_i^{\frac{n-2m-1}{2}}\cdot(1-s_i)^{-\frac{1}{2}}\, ds. \qquad \square
\end{aligned}$$

For the Hausdorff metric the volume of the metric ball can be written in terms of the so-called *hypergeometric function*. We will describe the few basic facts about this function that we will make use of in the following remark.

**Remark D.3.3.** Let $X \in \mathbb{C}^{m\times m}$ be a complex symmetric $(m\times m)$-matrix, with $\|X\| < 1$, and let $a,b,c\in\mathbb{C}$, with $c \notin \mathbb{Z}\cup \frac{1}{2}\cdot\mathbb{Z}$. The *Gaussian hypergeometric function of matrix argument* is defined as the convergent series

$${}_2F_1(a,b;c;X) = \sum_{k=0}^{\infty}\sum_{\kappa\vdash k} \frac{(a)_\kappa\cdot(b)_\kappa}{(c)_\kappa}\cdot\frac{C_\kappa(X)}{k!}\ ,$$

where the second summation runs over all partitions $\kappa$ of $k$, $C_\kappa(X)$ denotes the *zonal polynomial* of $X$ corresponding to $\kappa$ (cf. [38, Def. 7.2.1]), and $(a)_\kappa$ denotes the *generalized Pochhammer symbol*

$$(a)_\kappa = \prod_{i=1}^{\ell} \prod_{j=0}^{k_i-1} \left( a - \frac{i-1}{2} + j \right) ,$$

if $\kappa = (\kappa_1, \dots, \kappa_\ell)$. If $X = r \cdot I_m$ is a scalar multiple of the identity matrix, then the hypergeometric function has the integral representation

$$\begin{aligned} {}_2F_1(a, b; c; r \cdot I_m) = {} & \frac{\pi^{\frac{m}{2}}}{m!} \cdot \prod_{d=0}^{m-1} \frac{\Gamma(c - \frac{d}{2})}{\Gamma(\frac{m-d}{2}) \cdot \Gamma(a - \frac{d}{2}) \cdot \Gamma(c - a - \frac{d}{2})} \\ & \cdot \int_{[0,1]^m} \prod_{i<j} |s_i - s_j| \cdot \prod_{i=1}^{m} \frac{s_i^{a - \frac{m+1}{2}} \cdot (1 - s_i)^{c-a-\frac{m+1}{2}}}{(1 - r \cdot s_i)^b} \, ds_1 \cdots ds_m , \end{aligned} \tag{D.10}$$

holding for the same restrictions on $a$ and $c$ as before, and $0 < r < 1$ (see for example [33, (3.1)] or [28, (3.16)]).

Note that if we set $b = 0$ then using the Selberg integral $S_n(\alpha, \beta, \gamma)$ (cf. (C.1) in Section C.1), we get

$$\begin{aligned} {}_2F_1(a, 0; c; r \cdot I_m) & = \frac{\pi^{\frac{m}{2}}}{m!} \cdot \prod_{d=0}^{m-1} \frac{\Gamma(c - \frac{d}{2})}{\Gamma(\frac{m-d}{2}) \cdot \Gamma(a - \frac{d}{2}) \cdot \Gamma(c - a - \frac{d}{2})} \\ & \qquad \cdot S_m \left( a - \tfrac{m-1}{2}, c - a - \tfrac{m-1}{2}, \tfrac{1}{2} \right) \\ & = 1 , \end{aligned}$$

which may also be easily deduced from the definition of ${}_2F_1$ as an infinite series. Also, for $a, b, c \in \mathbb{R}_+$ we have

$${}_2F_1(a, b; c; r \cdot I_m) \geq 1 , \tag{D.11}$$

as all summands are real and nonnegative and the first summand equals 1.

**Proposition D.3.4.** *For $\mathcal{W}_0 \in \mathrm{Gr}_{n,m}$ and $\beta \in [0, \frac{\pi}{2}]$ let $B_{\mathrm{H}}(\mathcal{W}_0, \beta)$ denote the ball of radius $\beta$ around $\mathcal{W}_0$ w.r.t. the Hausdorff metric $d_{\mathrm{H}}$. The relative volume of $B_{\mathrm{H}}(\mathcal{W}_0, \beta)$ is given by*

$$\begin{aligned} \operatorname{rvol} B_{\mathrm{H}}(\mathcal{W}_0, \beta) & = \sin(\beta)^{m(n-m)} \cdot \begin{bmatrix} n \\ m \end{bmatrix}^{-1} \cdot {}_2F_1(\tfrac{n-m}{2}, \tfrac{1}{2}; \tfrac{n+1}{2}; \sin^2\beta \cdot I_m) \\ & \geq \sin(\beta)^{m(n-m)} \cdot \begin{bmatrix} n \\ m \end{bmatrix}^{-1} . \end{aligned}$$

*Proof.* From Proposition D.3.2 we get

$$\begin{aligned} \operatorname{rvol} B_{\mathrm{H}}(\mathcal{W}_0, \beta) = {} & \frac{\pi^{\frac{m}{2}}}{m!} \cdot \prod_{i=0}^{m-1} \frac{\Gamma(\frac{n-i}{2})}{\Gamma(\frac{i+1}{2})^2 \cdot \Gamma(\frac{n-m-i}{2})} \\ & \cdot \int_{T(K_{\mathrm{H}}(\beta))} \prod_{i<j} |s_i - s_j| \cdot \prod_{i=1}^{m} s_i^{\frac{n-2m-1}{2}} \cdot (1 - s_i)^{-\frac{1}{2}} \, ds , \end{aligned}$$

where $K_{\mathrm{H}}(\beta) = [0,\beta]^m$. Note that $T(K_{\mathrm{H}}(\beta)) = [0,\sin^2\beta]^m$. Scaling the domain by $\frac{1}{\sin^2\beta}$ to get an integration over $[0,1]^m$ and using the integral representation (D.10) of the hypergeometric function of scalar matrix argument yields

$$\begin{aligned}
&\int_{[0,\sin^2\beta]^m} \prod_{i<j} |s_i - s_j| \cdot \prod_{i=1}^m s_i^{\frac{n-2m-1}{2}} \cdot (1-s_i)^{-\frac{1}{2}}\, ds \\
&\qquad = (\sin^2\beta)^{\frac{m(m-1)}{2} + m\cdot\frac{n-2m-1}{2} + m} \\
&\qquad\qquad \int_{[0,1]^m} \prod_{i<j} |s_i - s_j| \cdot \prod_{i=1}^m s_i^{\frac{n-2m-1}{2}} \cdot (1-\sin^2\beta\cdot s_i)^{-\frac{1}{2}}\, ds \\
&\qquad = \sin(\beta)^{m(n-m)} \cdot \frac{m!}{\pi^{\frac{m}{2}}} \cdot \prod_{d=0}^{m-1} \frac{\Gamma(\frac{m-d}{2})\cdot\Gamma(a-\frac{d}{2})\cdot\Gamma(c-a-\frac{d}{2})}{\Gamma(c-\frac{d}{2})} \\
&\qquad\qquad \cdot {}_2F_1(\tfrac{n-m}{2}, \tfrac{1}{2}; \tfrac{n+1}{2}; \sin^2\beta\cdot I_m) .
\end{aligned}$$

So the constant mostly cancels leaving $\frac{\Gamma(\frac{m+1}{2})\cdot\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{1}{2})\cdot\Gamma(\frac{n+1}{2})} = \begin{bmatrix} n \\ m \end{bmatrix}^{-1}$. The estimate follows from (D.11), which finishes the proof. □

In the following corollary we will deduce from Proposition D.3.1 some formulas for probabilities involving the relative positions of an $m$-dimensional to an $M$-dimensional subspace of $\mathbb{R}^n$.

**Corollary D.3.5.** *Let $1 \le m \le M \le n-1$ with $m + M \le n$, and let $\mathcal{W}_0 \in \mathrm{Gr}_{n,M}$ fixed. If $\mathcal{W} \in \mathrm{Gr}_{n,m}$ is chosen uniformly at random then*

$$\begin{aligned}
\mathrm{Prob}[\sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W}) \le \beta] &= \sin(\beta)^{m(n-M)} \cdot \prod_{d=1}^m \frac{\Gamma(\frac{d+1}{2})\cdot\Gamma(\frac{n-m+d}{2})}{\Gamma(\frac{M-m+d}{2})\cdot\Gamma(\frac{n-M+1+d}{2})} \\
&\qquad \cdot {}_2F_1(\tfrac{n-M}{2}, \tfrac{m+1-M}{2}; \tfrac{n-M+m+1}{2}; \sin^2\beta\cdot I_m) ,
\end{aligned}$$

*where $\sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W})$ denotes the largest principal angle between $\mathcal{W}_0$ and $\mathcal{W}$. In particular, if $M = m+1$ then*

$$\mathrm{Prob}[\sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W}) \le \beta] = \sin(\beta)^{m(n-m)-m} .$$

*Furthermore, for the smallest principal angle $\sphericalangle_{\min}(\mathcal{W}_0,\mathcal{W})$ between $\mathcal{W}_0$ and $\mathcal{W}$ we have the estimate*

$$\begin{aligned}
&\mathrm{Prob}[\sphericalangle_{\min}(\mathcal{W}_0,\mathcal{W}) \le \beta] \\
&\qquad \le I_{n-2m+2,M-m}(\beta) \cdot \frac{m(n-M-m+1)}{n-M+1} \cdot \begin{bmatrix} n-m \\ M-1 \end{bmatrix} \cdot \binom{(n-M+1)/2}{m/2} ,
\end{aligned}$$

*where again $I_{n,j}(\beta) = \int_0^\beta \cos(\rho)^j \cdot \sin(\rho)^{n-2-j}\, d\rho$. For $M = m+1$ this simplifies to*

$$\mathrm{Prob}[\sphericalangle_{\min}(\mathcal{W}_0,\mathcal{W}) \le \beta] \le I_{n-2m+2,1}(\beta) \cdot \frac{m(n-2m)}{n-m} \cdot \binom{n-m}{m} .$$

*Proof.* The arguments are similar to those in the proof of Proposition D.3.2, so we may skip some technicalities. Considering the function $\varphi\colon \mathrm{St}_{m,M} \times \mathrm{St}_{n-M,m} \times U \to$

$\mathrm{Gr}_{n,m}$ defined in (D.9) we see that for $U = U_1 := [0,\beta]^m$ we get a $(2^m \cdot m!)$-fold covering (of a dense subset) of the set

$$\{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W}) \le \beta]$$

and for $U = U_2 := \{v \in \mathbb{R}^m \mid 0 < v_1 < \beta\,,\ v_1 < v_i < \frac{\pi}{2}\ \forall i = 2,\ldots,m\}$ we get a $(2^m \cdot (m-1)!)$-fold covering (of a dense subset) of the set

$$\{\mathcal{W} \in \mathrm{Gr}_{n,m} \mid \sphericalangle_{\min}(\mathcal{W}_0,\mathcal{W}) \le \beta\}\ .$$

The transformation formula and Proposition D.3.1 thus imply

$$\begin{aligned}\mathrm{Prob}[\sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W}) \le \beta] &= \frac{\mathrm{vol}\,\mathrm{St}_{M,m} \cdot \mathrm{vol}\,\mathrm{St}_{n-M,m}}{\mathrm{vol}\,\mathrm{Gr}_{n,m} \cdot m! \cdot 2^m} \\ &\quad \cdot \int_{U_1} \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^m \sin(v_i)^{n-m-M} \cdot \cos(v_i)^{M-m}\, dv\end{aligned}$$

and

$$\begin{aligned}\mathrm{Prob}[\sphericalangle_{\min}(\mathcal{W}_0,\mathcal{W}) \le \beta] &= \frac{\mathrm{vol}\,\mathrm{St}_{M,m} \cdot \mathrm{vol}\,\mathrm{St}_{n-M,m}}{\mathrm{vol}\,\mathrm{Gr}_{n,m} \cdot (m-1)! \cdot 2^m} \\ &\quad \cdot \int_{U_2} \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^m \sin(v_i)^{n-m-M} \cdot \cos(v_i)^{M-m}\, dv\ .\end{aligned}$$

The constant computes as

$$\frac{\mathrm{vol}\,\mathrm{St}_{M,m} \cdot \mathrm{vol}\,\mathrm{St}_{n-M,m}}{\mathrm{vol}\,\mathrm{Gr}_{n,m} \cdot m! \cdot 2^m} = \frac{2^m \cdot \pi^{\frac{m}{2}}}{m!} \cdot \prod_{d=1}^m \frac{\Gamma(\frac{n-m+d}{2})}{\Gamma(\frac{d}{2}) \cdot \Gamma(\frac{M-m+d}{2}) \cdot \Gamma(\frac{n-m-M+d}{2})}\ .$$

The same arguments as in the proof of Proposition D.3.4 yield the claim about $\mathrm{Prob}[\sphericalangle_{\max}(\mathcal{W}_0,\mathcal{W}) \le \beta]$, where the simplification in the special case $M = m+1$ follows from the fact ${}_2F_1(a,0;c;r \cdot I_m) = 1$ (cf. Remark D.3.3).

As for the statement about the smallest principal angle, we compute

$$\begin{aligned}&\int_0^\beta \int_{v_1}^{\frac{\pi}{2}} \cdots \int_{v_1}^{\frac{\pi}{2}} \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \\ &\qquad\qquad \cdot \prod_{i=1}^m \sin(v_i)^{n-m-M} \cdot \cos(v_i)^{M-m}\, dv_m \cdots dv_2\, dv_1 \\ &\overset{s_i := \sin^2 v_i}{=} \int_0^{\sin^2\beta} \int_{s_1}^1 \cdots \int_{s_1}^1 \prod_{i=1}^m \frac{1}{2\sqrt{s_i(1-s_i)}} \cdot \prod_{j<i} |s_i - s_j| \\ &\qquad\qquad \cdot \prod_{i=1}^m s_i^{\frac{n-m-M}{2}} \cdot (1-s_i)^{\frac{M-m}{2}}\, ds_m \cdots ds_2\, ds_1 \\ &= \frac{1}{2^m} \cdot \int_0^{\sin^2\beta} s_1^{\frac{n-m-M-1}{2}} \cdot (1-s_1)^{\frac{M-m-1}{2}} \int_{s_1}^1 \cdots \int_{s_1}^1 \prod_{i=2}^m (s_i - s_1) \cdot \prod_{2\le j<i} |s_i - s_j| \\ &\qquad\qquad \cdot \prod_{i=2}^m s_i^{\frac{n-m-M-1}{2}} \cdot (1-s_i)^{\frac{M-m-1}{2}}\, ds_m \cdots ds_2\, ds_1\ .\end{aligned}$$

By estimating $s_i - s_1 \leq s_i$ and by extending the integrals from the interval $[s_1, 1]$ to $[0, 1]$ we may continue

$$
\begin{aligned}
&\leq \frac{1}{2^m} \cdot \int_0^{\sin^2\beta} s_1^{\frac{n-m-M-1}{2}} \cdot (1-s_1)^{\frac{M-m-1}{2}}\, ds_1 \cdot \int_0^1 \cdots \int_0^1 \prod_{2\leq j<i} |s_i - s_j| \\
&\qquad\qquad \cdot \prod_{i=2}^{m} s_i^{\frac{n-m-M+1}{2}} \cdot (1-s_i)^{\frac{M-m-1}{2}}\, ds_m \cdots ds_2 \\
&\leq \frac{1}{2^m} \cdot \int_0^{\sin^2\beta} s_1^{\frac{n-m-M-1}{2}} \cdot (1-s_1)^{\frac{M-m-1}{2}}\, ds_1 \cdot S_{m-1}\left(\tfrac{n-m-M+3}{2}, \tfrac{M-m+1}{2}, \tfrac{1}{2}\right) ,
\end{aligned}
$$

where $S_n(x, y, z)$ shall denote the Selberg integral (cf. Section C.1). Evaluating the Selberg integral, we may continue

$$
\begin{aligned}
&= \frac{1}{2^m} \cdot \int_0^{\beta} 2 \cdot \sin v_1 \cdot \cos v_1 \cdot \sin(v_1)^{n-m-M-1} \cdot \cos(v_1)^{M-m-1}\, ds_1 \\
&\qquad\qquad \cdot \frac{(m-1)!}{\pi^{\frac{m-1}{2}}} \cdot \prod_{d=0}^{m-2} \frac{\Gamma(\frac{m-1-d}{2}) \cdot \Gamma(\frac{n-M-d+1}{2}) \cdot \Gamma(\frac{M-1-d}{2})}{\Gamma(\frac{n-d}{2})} \\
&= \frac{(m-1)!}{2^{m-1} \cdot \pi^{\frac{m-1}{2}}} \cdot I_{n-2m+2,M-m}(\beta) \cdot \prod_{d=0}^{m-2} \frac{\Gamma(\frac{m-1-d}{2}) \cdot \Gamma(\frac{n-M-d+1}{2}) \cdot \Gamma(\frac{M-1-d}{2})}{\Gamma(\frac{n-d}{2})} \ .
\end{aligned}
$$

Combining this with the above given formula for $\mathrm{Prob}[\sphericalangle_{\min}(\mathcal{W}_0, \mathcal{W}) \leq \beta]$ we get

$$
\begin{aligned}
\mathrm{Prob}[\sphericalangle_{\min}(\mathcal{W}_0, \mathcal{W}) \leq \beta] &\leq 2\sqrt{\pi} \cdot I_{n-2m+2,M-m}(\beta) \\
&\qquad \cdot \frac{\Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{M}{2}) \cdot \Gamma(\frac{n-M-m+1}{2})} \cdot \frac{\Gamma(\frac{n-M+1}{2})}{\Gamma(\frac{m}{2}) \cdot \Gamma(\frac{n-M-m+2}{2})} \ .
\end{aligned}
$$

A straightforward computation, using the identities in Proposition 4.1.20, finishes the proof. □

We finish this section with the proof of Proposition D.3.1.

*Proof of Proposition D.3.1.* For $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, let

$$C_x := \mathrm{diag}(\cos(x_1), \ldots, \cos(x_m)) \ , \quad S_x := \mathrm{diag}(\sin(x_1), \ldots, \sin(x_m)) \ .$$

We need to compute the Normal Jacobian of the function $\varphi$ defined in (D.9); so first, we will have to compute the differential $D\varphi$ in $([Q_1], [Q_2], v)$. As the domain of $\varphi$ is the direct product $\mathrm{St}_{M,m} \times \mathrm{St}_{n-M,m} \times U$, $U \subseteq \mathbb{R}^m$, the tangent space also decomposes into a direct product. We will consider these components separately. Recall that we have given an extensive description of the Stiefel and the Grassmann manifold in Section 5.3.1 and Section 5.3.2, which we will make use of in the following argumentation.

By symmetry, we may assume w.l.o.g. that $Q_1 = I_M$ and $Q_2 = I_{n-M}$. Let $\mathcal{W}_v := \varphi([I_M], [I_{n-M}], v)$. For the first component we consider a curve $\bar{c}_1$ in $\mathrm{St}_{M,m} = O(M)/O(M-m)$ defined by a curve $c_1$ in $O(M)$:

$$
\begin{aligned}
&c_1 \colon \mathbb{R} \to O(M) \, , \; c_1(0) = I_M \, , \; \frac{dc_1}{dt}(0) = \begin{pmatrix} U_1 & -R_1^T \\ R_1 & 0 \end{pmatrix} \ , \\
&\bar{c}_1 \colon \mathbb{R} \to \mathrm{St}_{M,m} \, , \; \bar{c}_1(t) = [c_1(t)] \, , \; \frac{d\bar{c}_1}{dt}(0) = \left[ I_M, \begin{pmatrix} U_1 & -R_1^T \\ R_1 & 0 \end{pmatrix} \right] \ ,
\end{aligned}
$$

where $U_1 \in \mathrm{Skew}_m$ and $R_1 \in \mathbb{R}^{(M-m)\times m}$. Let us abbreviate $C := C_v$ and $S := S_v$. Then we get

$$\varphi(\bar{c}_1(t), [I_{n-M}], v) = [Q_1(t)] \ ,$$

where $Q_1(t) \in O(n)$ is given by

$$Q_1(t) = \left(\begin{array}{c|c} c_1(t) & 0 \\ \hline 0 & I_{n-M} \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) .$$

Note that

$$Q_1(0) = \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) =: Q_v \ , \tag{D.12}$$

and $[Q_v] = \mathcal{W}_v \in \mathrm{Gr}_{n,m}$. We compute

$$\begin{aligned} \frac{dQ_1}{dt}(0) &= \left(\begin{array}{c|c} \frac{dc_1}{dt}(0) & 0 \\ \hline 0 & 0 \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \\ &= Q_v \cdot Q_v^{-1} \cdot \left(\begin{array}{cc|c} U_1 & -R_1^T & 0 \\ R_1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right) \cdot Q_v \\ &\overset{(*)}{=} Q_v \cdot \left(\begin{array}{cc|cc} CU_1C & -CR_1^T & -CU_1S & 0 \\ R_1C & 0 & -R_1S & 0 \\ \hline -SU_1C & SR_1^T & SU_1S & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \ , \end{aligned}$$

where $(*)$ is verified easily (note that $Q_v^{-1} = Q_v^T$). We get

$$\begin{aligned} D\varphi\left(\tfrac{d\bar{c}_1}{dt}(0), 0, 0\right) &= \frac{d\varphi(\bar{c}_1(t), [I_{n-M}], v)}{dt}(0) = \frac{d[Q_1(t)]}{dt}(0) \\ &= \left[ Q_v, \left(\begin{array}{cc|cc} \multicolumn{2}{c|}{0} & -CU_1S & 0 \\ & & 0 & 0 \\ \hline -SU_1C & 0 & \multicolumn{2}{c}{0} \\ 0 & 0 & & \end{array}\right) \right] \in T_{[Q_v]} \mathrm{Gr}_{n,m} \ . \end{aligned}$$

This settles the first component. As for the second component, let us consider the curves

$$\begin{aligned} &c_2 \colon \mathbb{R} \to O(n-M)\,,\ c_2(0) = I_{n-M}\,,\ \frac{dc_2}{dt}(0) = \begin{pmatrix} U_2 & -R_2^T \\ R_2 & 0 \end{pmatrix} \ , \\ &\bar{c}_2 \colon \mathbb{R} \to \mathrm{St}_{n-M,m}\,,\ \bar{c}_2(t) = [c_2(t)]\,,\ \frac{d\bar{c}_2}{dt}(0) = \left[ I_{n-M}, \begin{pmatrix} U_2 & -R_2^T \\ R_2 & 0 \end{pmatrix} \right] \ , \end{aligned}$$

where $U_2 \in \mathrm{Skew}_m$ and $R_2 \in \mathbb{R}^{(n-M-m)\times m}$. Then we get

$$\varphi([I_M], \bar{c}_2(t), v) = [Q_2(t)] \ ,$$

where $Q_2(t) \in O(n)$ is given by

$$Q_2(t) = \left(\begin{array}{c|c} I_M & 0 \\ \hline 0 & c_2(t) \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) .$$

Note again, that $Q_2(0) = Q_v$. As above, we compute

$$\begin{aligned} \frac{dQ_2}{dt}(0) &= \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \frac{dc_2}{dt}(0) \end{array}\right) \cdot \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \\ &= Q_v \cdot Q_v^{-1} \cdot \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & U_2 & -R_2^T \\ 0 & R_2 & 0 \end{array}\right) \cdot Q_v \\ &\overset{(**)}{=} Q_v \cdot \left(\begin{array}{cc|cc} SU_2S & 0 & SU_2C & -SR_2^T \\ 0 & 0 & 0 & 0 \\ \hline CU_2S & 0 & CU_2C & -CR_2^T \\ R_2S & 0 & R_2C & 0 \end{array}\right) , \end{aligned}$$

where $(**)$ is again verified easily. We get

$$\begin{aligned} D\varphi\left(0, \tfrac{d\bar{c}_2}{dt}(0), 0\right) &= \frac{d\varphi([I_M], \bar{c}_1(t), v)}{dt}(0) = \frac{d[Q_2(t)]}{dt}(0) \\ &= \left[ Q_v, \left(\begin{array}{cc|cc} \multicolumn{2}{c|}{0} & SU_2C & -SR_2^T \\ & & 0 & 0 \\ \hline CU_2S & 0 & \multicolumn{2}{c}{0} \\ R_2S & 0 & & \end{array}\right) \right] \in T_{[Q_v]} \operatorname{Gr}_{n,m} . \end{aligned}$$

This settles the second component. As for the third, we consider the curve

$$c_3 \colon \mathbb{R} \to \mathbb{R}^m , \; c_3(0) = v , \; \frac{dc_3}{dt}(0) = \zeta ,$$

where $\zeta \in T_v\mathbb{R}^m = \mathbb{R}^m$. Then we get

$$\varphi([I_M], [I_{n-M}], c_3(t)) = [Q_3(t)] ,$$

where

$$Q_3(t) = \left(\begin{array}{cc|cc} C_{c_3(t)} & & -S_{c_3(t)} & \\ & I_{M-m} & & 0 \\ \hline S_{c_3(t)} & & C_{c_3(t)} & \\ & 0 & & I_{n-m-M} \end{array}\right) .$$

We compute

$$
\begin{aligned}
\frac{dQ_3}{dt}(0) &= \left(\begin{array}{cc|cc} -S\cdot\operatorname{diag}(\zeta) & & -C\cdot\operatorname{diag}(\zeta) & \\ & 0 & & 0 \\ \hline C\cdot\operatorname{diag}(\zeta) & & -S\cdot\operatorname{diag}(\zeta) & \\ & 0 & & 0 \end{array}\right) \\
&= \left(\begin{array}{cc|cc} C & & -S & \\ & I_{M-m} & & 0 \\ \hline S & & C & \\ & 0 & & I_{n-m-M} \end{array}\right) \cdot \left(\begin{array}{cc|cc} 0 & & -\operatorname{diag}(\zeta) & \\ & 0 & & 0 \\ \hline \operatorname{diag}(\zeta) & & 0 & \\ & 0 & & 0 \end{array}\right) \\
&= Q_v \cdot \left(\begin{array}{cc|cc} 0 & & -\operatorname{diag}(\zeta) & \\ & 0 & & 0 \\ \hline \operatorname{diag}(\zeta) & & 0 & \\ & 0 & & 0 \end{array}\right) .
\end{aligned}
$$

We get

$$
\begin{aligned}
D\varphi\left(0,0,\tfrac{dc_3}{dt}(0)\right) &= \frac{d\varphi([I_M],[I_{n-M}],c_3(t))}{dt}(0) = \frac{d[Q_3(t)]}{dt}(0) \\
&= \left[ Q_v, \left(\begin{array}{cc|cc} 0 & & -\operatorname{diag}(\zeta) & \\ & & & 0 \\ \hline \operatorname{diag}(\zeta) & & 0 & \\ & 0 & & \end{array}\right)\right] \in T_{[Q_v]}\operatorname{Gr}_{n,m} .
\end{aligned}
$$

This settles the third component.

We have thus a full description of the differential $D\varphi$. To compute the normal Jacobian it remains to specify orthonormal bases for the tangent spaces $T_v\mathbb{R}$, $T_{[I_M]}\operatorname{St}_{M,m}$, $T_{[I_{n-M}]}\operatorname{St}_{n-M,m}$, set up the corresponding Jacobi matrix, and compute its determinant.

Let us write $E^k_{ij}$ for the $(i,j)$th elementary matrix of format $k\times k$. Let us furthermore define

$$
\begin{aligned}
\xi_{ij} &:= \left[I_M, E^M_{ij} - E^M_{ji}\right] \in T_{[I_M]}\operatorname{St}_{M,m} , \\
\eta_{ij} &:= \left[I_{n-M}, E^{n-M}_{ij} - E^{n-M}_{ji}\right] \in T_{[I_{n-M}]}\operatorname{St}_{n-M,m} .
\end{aligned}
$$

Then we have the following orthonormal bases of $T_{[I_M]}\operatorname{St}_{M,m}$ and $T_{[I_{n-M}]}\operatorname{St}_{n-M,m}$:

$$
\begin{aligned}
T_{[I_M]}\operatorname{St}_{M,m} : \xi_{ij} , &\quad 1\le j<i\le m \ \text{ or } (m+1\le i\le M\ ,\ 1\le j\le m) , \\
T_{[I_{n-M}]}\operatorname{St}_{n-M,m} : \eta_{ij} , &\quad 1\le j<i\le m \ \text{ or } (m+1\le i\le n-M\ ,\ 1\le j\le m) ,
\end{aligned}
$$

To get a nice form of the Jacobi matrix let us choose the following order of the basis vectors of

$$
T_{[I_M]}\operatorname{St}_{M,m} \times T_{[I_{n-M}]}\operatorname{St}_{n-M,m} \times T_v\mathbb{R}^m \ :
$$

1. the canonical basis of $T_v\mathbb{R}^m = \mathbb{R}^m$,
2. the first half of $T_{[I_M]}\operatorname{St}_{M,m}$ consisting of
$$
\xi_{ij}\ ,\quad 1\le j<i\le m\ ,
$$
3. the first half of $T_{[I_{n-M}]}\operatorname{St}_{n-M,m}$ consisting of
$$
\eta_{ij}\ ,\quad 1\le j<i\le m\ ,
$$

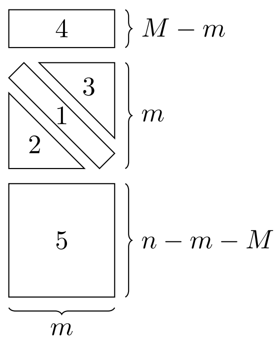
4. the second half of $T_{[I_M]}\,\mathrm{St}_{M,m}$ consisting of
$$\xi_{ij}\;, \quad m+1 \le i \le M\;,\; 1 \le j \le m\;,$$
5. the second half of $T_{[I_{n-M}]}\,\mathrm{St}_{n-M,m}$ consisting of
$$\eta_{ij}\;, \quad m+1 \le i \le n-M\;,\; 1 \le j \le m\;.$$

To set up the Jacobi matrix of $\varphi$, it remains to specify an order of the basis vectors of $T_{[Q_v]}\,\mathrm{Gr}_{n,m}$. Recall that this space is given by

$$T_{[Q_v]}\,\mathrm{Gr}_{n,m} = \left\{ \begin{pmatrix} 0 & -R^T \\ R & 0 \end{pmatrix} \,\middle|\, R \in \mathbb{R}^{(n-m)\times m} \right\}\;.$$

So we may identify each tangent vector with a $((n-m)\times m)$-matrix. Specifying an orthonormal basis of $T_{[Q_v]}\,\mathrm{Gr}_{n,m}$ thus means to identify an order in which to read the entries of this matrix. It turns out that the following order yields a particularly nice form of the Jaobi matrix:

1. the diagonal elements of the middle $m\times m$ submatrix
2. the strictly lower diagonal elements in the middle $m \times m$ submatrix
3. the strictly upper diagonal elements in the middle $m \times m$ submatrix
4. the upper $(M-m)\times m$ submatrix (row by row)
5. the lower $(n-m-M)\times m$ submatrix (row by row).




Now that we have made the necessary specifications we can compute the Jacobi matrix which turns out to be the following

$$\begin{pmatrix}
I_m & & & & & & & & \\
& -SC_v & CS & & & & & & \\
& CS & -SC & & & & & & \\
& & & C & & & & & \\
& & & & \ddots & & & & \\
& & & & & C & & & \\
& & & & & & S & & \\
& & & & & & & \ddots & \\
& & & & & & & & S
\end{pmatrix}
\begin{array}{l}
\}\, m \\
\}\, \frac{m(m-1)}{2} \\
\}\, \frac{m(m-1)}{2} \\
\left.\begin{array}{l} \}\, m \\ \\ \}\, m \end{array}\right\} (M-m)\text{-times} \\
\left.\begin{array}{l} \}\, m \\ \\ \}\, m \end{array}\right\} (n-m-M)\text{-times}\;,
\end{array}$$

where

$$SC := \mathrm{diag}(\sin v_2 \cdot \cos v_1, \sin v_3 \cdot \cos v_1, \sin v_3 \cdot \cos v_2, \ldots, \sin v_m \cdot \cos v_{m-1})$$
$$CS := \mathrm{diag}(\cos v_2 \cdot \sin v_1, \cos v_3 \cdot \sin v_1, \cos v_3 \cdot \sin v_2, \ldots, \cos v_m \cdot \sin v_{m-1})\;.$$

Using the fact that

$$\det \left(\begin{array}{ccc|ccc}
a_1 & & & b_1 & & \\
& \ddots & & & \ddots & \\
& & a_k & & & b_k \\
\hline
c_1 & & & d_1 & & \\
& \ddots & & & \ddots & \\
& & c_k & & & d_k
\end{array}\right) = \det \begin{pmatrix}
a_1 & b_1 & & & \\
c_1 & d_1 & & & \\
& & \ddots & & \\
& & & a_k & b_k \\
& & & c_k & d_k
\end{pmatrix} = \prod_{i=1}^{k} a_i d_i - b_i c_i\;,$$

we finally get that the normal Jacobian of $\varphi$ is given by

$$\operatorname{ndet}(D_{([I_M],[I_{n-M}],v)}\varphi) = \prod_{j<i} \left|\sin^2 v_i \cos^2 v_j - \cos^2 v_i \sin^2 v_j\right| \cdot \prod_{i=1}^{m} \left|\sin(v_i)^{n-m-M} \cdot \cos(v_i)^{M-m}\right| . \qquad \square$$

# Bibliography


[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] P.-A. Absil, A. Edelman, and P. Koev. On the largest principal angle between random subspaces. *Linear Algebra Appl.*, 414(1):288–294, 2006.

[3] P.-A. Absil, R. Mahony, and R. Sepulchre. Riemannian geometry of Grassmann manifolds with a view on algorithmic computation. *Acta Appl. Math.*, 80(2):199–220, 2004.

[4] A. Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[5] A. Belloni and R. M. Freund. A geometric analysis of Renegar's condition number, and its interplay with conic curvature. *Math. Program.*, 119(1, Ser. A):95–107, 2009.

[6] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation.* Springer, 1998.

[7] T. Bonnesen and W. Fenchel. *Theorie der konvexen Körper.* Springer-Verlag, Berlin, 1974. Berichtigter Reprint.

[8] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry*, volume 120 of *Pure and Applied Mathematics.* Academic Press Inc., Orlando, FL, second edition, 1986.

[9] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge University Press, Cambridge, 2004.

[10] F. Brenti. Unimodal, log-concave and Pólya frequency sequences in combinatorics. *Mem. Amer. Math. Soc.*, 81(413):viii+106, 1989.

[11] P. Bürgisser. Smoothed analysis of condition numbers. Invited Paper in Section 15: Mathematical Aspects of Computer Science, International Congress of Mathematicians, Hyderabad, India, 2010.

[12] P. Bürgisser. Smoothed analysis of condition numbers. In *Foundations of computational mathematics, Hong Kong 2008*, volume 363 of *London Math. Soc. Lecture Note Ser.*, pages 1–41. Cambridge Univ. Press, Cambridge, 2009.

[13] P. Bürgisser and D. Amelunxen. Robust smoothed analysis of a condition number for linear programming. *Math. Program.*, pages 1–31, 2010.

[14] P. Bürgisser and F. Cucker. Smoothed analysis of Moore–Penrose Inversion. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2769–2783, 2010.

[15] P. Bürgisser, F. Cucker, and M. Lotz. General formulas for the smoothed analysis of condition numbers. *C. R. Math. Acad. Sci. Paris*, 343(2):145–150, 2006.

[16] P. Bürgisser, F. Cucker, and M. Lotz. Smoothed analysis of complex conic condition numbers. *J. Math. Pures et Appl.*, 86:293–309, 2006.

[17] P. Bürgisser, F. Cucker, and M. Lotz. The probability that a slightly perturbed numerical analysis problem is difficult. *Math. Comp.*, 77(263):1559–1583, 2008.

[18] P. Bürgisser, F. Cucker, and M. Lotz. Coverage processes on spheres and condition numbers for linear programming. *Ann. Probab.*, 38(2):570–604, 2010.

[19] I. Chavel. *Riemannian geometry*, volume 98 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2006. A modern introduction.

[20] Z. Chen and J. J. Dongarra. Condition numbers of Gaussian random matrices. *SIAM J. Matrix Anal. Appl.*, 27(3):603–620 (electronic), 2005.

[21] F. Cucker, R. Hauser, and M. Lotz. Adversarial smoothed analysis. *J. Complexity*, 26(3):255–262, 2010.

[22] J.-P. Dedieu and G. Malajovich. On the number of minima of a random polynomial. *J. Complexity*, 24(2):89–108, 2008.

[23] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[24] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. of Matrix Anal. and Applic.*, 9:543–556, 1988.

[25] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353 (electronic), 1999.

[26] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[27] W. J. Firey. An integral-geometric meaning for lower order area functions of convex bodies. *Mathematika*, 19:205–212, 1972.

[28] P. J. Forrester and S. O. Warnaar. The importance of the Selberg integral. *Bull. Amer. Math. Soc. (N.S.)*, 45(4):489–534, 2008.

[29] F. Gao, D. Hug, and R. Schneider. Intrinsic volumes and polar sets in spherical space. *Math. Notae*, 41:159–176 (2003), 2001/02. Homage to Luis Santaló. Vol. 1 (Spanish).

[30] S. Glasauer. Integralgeometrie konvexer Körper im sphärischen Raum. Thesis, Univ. Freiburg i. Br., 1995.

[31] S. Glasauer. Integral geometry of spherically convex bodies. *Diss. Summ. Math.*, 1(1-2):219–226, 1996.

[32] G. Golub and C. Van Loan. *Matrix computations.* Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[33] R. D. Gupta and D. St. P. Richards. Hypergeometric functions of scalar matrix argument are expressible in terms of classical hypergeometric functions. *SIAM J. Math. Anal.*, 16(4):852–858, 1985.

[34] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80 of *Pure and Applied Mathematics.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978.

[35] C. Jordan. Essai sur la géométrie à $n$ dimensions. *Bull. Soc. Math. France*, 3:103–174, 1875.

[36] D. A. Klain and G.-C. Rota. *Introduction to geometric probability.* Lezioni Lincee. [Lincei Lectures]. Cambridge University Press, Cambridge, 1997.

[37] F. Morgan. *Geometric measure theory.* Academic Press Inc., San Diego, CA, second edition, 1995. A beginner's guide.

[38] R. J. Muirhead. *Aspects of multivariate statistical theory.* John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.

[39] L. Nachbin. *The Haar integral.* D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965.

[40] J. Nie, K. Ranestad, and B. Sturmfels. The algebraic degree of semidefinite programming. *Math. Program.*, 122(2):379–405, 2009.

[41] G. Pataki. The geometry of semidefinite programming. In *Handbook of semidefinite programming*, volume 27 of *Internat. Ser. Oper. Res. Management Sci.*, pages 29–65. Kluwer Acad. Publ., Boston, MA, 2000.

[42] J. Peña. Understanding the geometry of infeasible perturbations of a conic linear system. *SIAM J. Optim.*, 10(2):534–550 (electronic), 2000.

[43] J. Renegar. Some perturbation theory for linear programming. *Math. Programming*, 65(1, Ser. A):73–91, 1994.

[44] J. Renegar. Incorporating condition measures into the complexity theory of linear programming. *SIAM J. Optim.*, 5(3):506–524, 1995.

[45] J. Renegar. Linear programming, complexity theory and elementary functional analysis. *Math. Programming*, 70(3, Ser. A):279–351, 1995.

[46] J. Renegar. *A mathematical view of interior-point methods in convex optimization.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[47] R. T. Rockafellar. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[48] M. Schmuckenschlaeger. A simple proof of an approximation theorem of H. Minkowski. *Geom. Dedicata*, 48(3):319–324, 1993.

[49] R. Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[50] R. Schneider and J. A. Wieacker. Integral geometry. In *Handbook of convex geometry, Vol. A, B*, pages 1349–1390. North-Holland, Amsterdam, 1993.

[51] S. Smale. Complexity theory and numerical analysis. In *Acta numerica, 1997*, volume 6 of *Acta Numer.*, pages 523–551. Cambridge Univ. Press, Cambridge, 1997.

[52] M. Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus.* W. A. Benjamin, Inc., New York-Amsterdam, 1965.

[53] M. Spivak. *A comprehensive introduction to differential geometry. Vol. I.* Publish or Perish Inc., Wilmington, Del., third edition, 2005.

[54] R. P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 500–535. New York Acad. Sci., New York, 1989.

[55] T. Tao and V. Vu. The condition number of a randomly perturbed matrix. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 248–255. ACM, New York, 2007.

[56] E. Teufel. Kinematische Berührformeln in Räumen konstanter Krümmung. *Abh. Math. Sem. Univ. Hamburg*, 58:255–266, 1988.

[57] E. Teufel. Kinematische Berührformeln in Räumen konstanter Krümmung. II. *Abh. Math. Sem. Univ. Hamburg*, 59:107–122, 1989.

[58] E. Teufel. Kinematische Berührformeln in Riemannschen homogenen Räumen. *Results Math.*, 15(3-4):361–381, 1989.

[59] J. A. Thorpe. *Elementary topics in differential geometry.* Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1994.

[60] L. N. Trefethen and D. Bau, III. *Numerical linear algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[61] J. C. Vera, J. C. Rivera, J. Peña, and Yao Hui. A primal-dual symmetric relaxation for homogeneous conic systems. *J. Complexity*, 23(2):245–261, 2007.

[62] J. G. Wendel. A problem in geometric probability. *Math. Scand.*, 11:109–111, 1962.

[63] H. Weyl. On the Volume of Tubes. *Amer. J. Math.*, 61(2):461–472, 1939.