Concepts and techniques of optimization on the sphere

O. P. Ferreira* A. N. Iusem, † S. Z. Németh ‡

November 13, 2013

Abstract

In this paper some concepts and techniques of Mathematical Programming are extended in an intrinsic way from the Euclidean space to the sphere. In particular, the notion of convex functions, variational problem and monotone vector fields are extended to the sphere and several characterizations of these notions are shown. As an application of the convexity concept, necessary and sufficient optimality conditions for constrained convex optimization problems on the sphere are derived.

Keywords: Sphere, convex function in the sphere, spheric constrained optimization, variational problem, monotone vector fields.

1 Introduction

It is natural to extend the concepts and techniques of Optimization from the Euclidean space to the Euclidean sphere. This has been done frequently before. The motivation of this extension is either of purely theoretical nature or aims at obtaining efficient algorithms; see [2, 6, 23, 24, 26, 27, 28, 29]. Indeed, many optimization problems are naturally posed on the sphere, which has a specific underlining algebraic structure that could be exploited to greatly reduce the cost of obtaining the solutions; see [23, 24, 28, 29]. Besides the theoretical interest, constrained optimization problems on the sphere also have a wide range of applications in many different areas of study such as numerical multilinear algebra (see, e.g., [18]), solid mechanics (see, e.g., [9]), signal processing (see,

*IME/UFG, Campus II- Caixa Postal 131, Goiânia, GO, 74001-970, Brazil (e-mail:orizon@ufg.br). The author was supported in part by FUNAPE/UFG, CNPq Grants 201112/2009-4, 475647/2006-8 and PRONEX–Optimization(FAPERJ/CNPq).

†Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Rio de Janeiro, RJ, 22460-320, Brazil (email:iusp@impa.br). This author was supported in part by CNPq grant no. 301280/86 and PRONEX-Otimização(FAPERJ/CNPq).

‡School of Mathematics, The University of Birmingham, The Watson Building, Edgbaston, Birmingham B15 2TT, United Kingdom (e-mail:nemeths@for.mat.bham.ac.uk). The author was supported in part by the Hungarian Research Grant OTKA 60480.

e.g., [19, 25]) and quantum mechanics (see, e.g., [1]). For instance, consider the generic constrained optimization problem on the sphere $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \ : \ \|x\| = 1\}$:

$$\min\{f(x) : x \in C\}, \qquad C \subseteq \mathbb{S}^n. \tag{1}$$

For $C = \mathbb{S}^n$ and a quadratic form $f(x) = x^T Q x$, the problem in (1) becomes a minimal eigenvalue problem, that is, finding the spectral norm of the matrix $-Q$ (see, e.g., [23]). Problem (1) includes as particular cases the problem of deciding the non-negativity of a homogeneous multivariate polynomial over the sphere (see, e.g., [14, 20, 21]) as well as the Bi-Quadratic Optimization problem over unit spheres (see, e.g., [17]). For quadratic functions it also contains the trust region problem that appears in many nonlinear programming algorithms as a sub-problem, see [3]. Let us state the facility location problem which is also a particular instance of the problem in (1): Let $p_1, p_2, \cdots, p_m \in C \subseteq \mathbb{S}^n$ and let $c_1, c_2, \cdots, c_m$ positive numbers and $f(p) = \sum_{i=1}^m c_i d(p_i, p)$, where d is the distance on the surface of the sphere. So, the spherical facility location problem is $\min\{\sum_{i=1}^m c_i d(p_i, p) : x \in C\}$ (see, e.g.,[6, 10, 13, 28]) .

The aim of this paper is to extend some concepts and techniques of Mathematical Programming from the Euclidean space to the Euclidean sphere in an intrinsic way. For extending these concepts we first study some important properties of the intrinsic distance from a fixed point; for instance, we present the spectral decomposition of its Hessian. Then we extend to the sphere the concept of convex functions, variational problem and monotone vector fields. In particular, we present the first and second order characterizations of convex functions and, as an application, we obtain the necessary an sufficient optimality conditions for convex constrained optimization problems on the sphere. We also present some basics properties related to the variational problem.

The structure of this paper is as follows. In Section 2, we recall some notations, definitions and basic properties about the geometry of the sphere used throughout the paper. In Section 2.1 we present some important properties of the intrinsic distance from a fixed point. In Section 3 we consider some properties of convex set in the sphere. In Section 4 we study the basic properties of convex functions on the sphere. In Section 5 we obtain sufficient optimality conditions for constrained optimization problems on the sphere. In Section 6 we study the basics properties of variational problem in the sphere. In Section 7 we define the monotonicity of a vector field on the sphere and show that the gradient vector field of a differentiable convex function on the sphere is intrinsically monotone. We conclude this paper by making some final remarks in Section 8.

2 Basics results about the sphere

In this section we recall some notations, definitions and basic properties about the geometry of the sphere used throughout the paper. They can be found in many introductory books on Riemannian and differential Geometry, for example in [4], [5] and [22].

Let $\langle , \rangle$ be the *Euclidean inner product*, with corresponding *norm* denoted by $\| \ \|$. Throughout the paper the *n-dimensional Euclidean sphere* and its *tangent hyperplane at a point p* are denoted by

$$\mathbb{S}^n := \{p = (p_1, \ldots, p_{n+1}) \in \mathbb{R}^{n+1} : \|p\| = 1\}, \qquad T_p\mathbb{S}^n := \{v \in \mathbb{R}^n : \langle p, v\rangle = 0\},$$

respectively. Let I be the $(n+1) \times (n+1)$ identity matrix. The *projection onto the tangent hyperplane* $T_p\mathbb{S}^n$ is the linear mapping defined by

$$I - pp^T : \mathbb{R}^{n+1} \to T_p\mathbb{S}^n, \tag{2}$$

where p^T denotes the transpose of the vector p.

The *intrinsic distance on the sphere* between two arbitrary points $p, q \in \mathbb{S}^n$ is defined by

$$d(p, q) := \arccos\langle p, q\rangle. \tag{3}$$

It can be shown that the intrinsic distance $d(p, q)$ between two arbitrary points $p, q \in \mathbb{S}^n$ is obtained by minimizing the *arc length functional* ℓ,

$$\ell(c) := \int_a^b \|c'(t)\| dt,$$

over the set of all piecewise continuously differentiable curves $c : [a, b] \to \mathbb{S}^n$ joining p to q, i.e., such that $c(a) = p$ and $c(b) = q$. Moreover, d is a distance in $\mathbb{S}^n$ and $(\mathbb{S}^n, d)$ is a complete metric space, so that $d(p, q) \geq 0$ for all $p, q \in \mathbb{S}^n$, and $d(p, q) = 0$ if and only if $p = q$. It is easy to check also that $d(p, q) \leq \pi$ for all $p, q \in \mathbb{S}^n$, and $d(p, q) = \pi$ if and only if $p = -q$.

The intersection curve of a plane though the origin of $\mathbb{R}^{n+1}$ with the sphere $\mathbb{S}^n$ is called a *geodesic*. A geodesic segment $\gamma : [a, b] \to \mathbb{S}^n$ is said to be *minimal* if its arc length is equal the intrinsic distance between its end points, i.e., if $\ell(\gamma) := \arccos\langle\gamma(a), \gamma(b)\rangle$. We say that γ is a *normalized geodesic* if $\|\gamma'\| = 1$. If $p, q \in \mathbb{S}^n$ are such that $q \neq p$ and $q \neq -p$, then the unique *segment of minimal normalized geodesic from to p to q* is

$$\gamma_{pq}(t) = \left(\cos t - \frac{\langle p, q\rangle \sin t}{\sqrt{1 - \langle p, q\rangle^2}}\right) p + \frac{\sin t}{\sqrt{1 - \langle p, q\rangle^2}}\, q, \qquad t \in [0,\ d(p, q)]. \tag{4}$$

Let $p \in \mathbb{S}^n$ and $v \in T_p\mathbb{S}^n$ such that $\|v\| = 1$. The minimal segment of geodesic connecting p to $-p$, starting at p with velocity v at p is given by

$$\gamma_{p\{-p\}}(t) := \cos(t)\, p + \sin(t)\, v, \qquad t \in [0,\ \pi]. \tag{5}$$

The *exponential mapping* $\exp_p : T_p\mathbb{S}^n \to \mathbb{S}^n$ is defined by $\exp_p v := \gamma_v(1)$, where γ_v is the geodesic defined by its initial position p, with velocity v at p. Hence,

$$\exp_p v := \begin{cases} \cos(\|v\|)\, p + \sin(\|v\|)\, \dfrac{v}{\|v\|}, & v \in T_p\mathbb{S}^n/\{0\}, \\ p, & v = 0. \end{cases} \tag{6}$$

It is easy to prove that $\gamma_{tv}(1) = \gamma_v(t)$ for all t. Therefore, for all $t \in \mathbb{R}$ we have

$$\exp_p tv := \begin{cases} \cos(t\|v\|)\, p + \sin(t\|v\|)\, \dfrac{v}{\|v\|}, & v \in T_p\mathbb{S}^n/\{0\}, \\ p, & v = 0. \end{cases} \tag{7}$$

We will also use the expression above for denoting the geodesic starting at $p \in \mathbb{S}^n$ with velocity $v \in T_p\mathbb{S}^n$ at p. The *inverse of the exponential mapping* is given by

$$\exp_p^{-1} q := \begin{cases} \dfrac{\arccos\langle p, q\rangle}{\sqrt{1 - \langle p, q\rangle^2}} \left(I - pp^T\right) q, & q \notin \{p, -p\}, \\ 0, & q = p. \end{cases} \tag{8}$$

It follows from (3) and (8) that

$$d(p, q) = \|\exp_q^{-1} p\|, \qquad p, q \in \mathbb{S}^n. \tag{9}$$

Let $\Omega \subset \mathbb{S}^n$ be an open set. The *gradient on the sphere* of a differentiable function $f : \Omega \to \mathbb{R}$ at a point $p \in \Omega$ is the vector defined by

$$\operatorname{grad} f(p) := \left[I - pp^T\right] Df(p) = Df(p) - \langle Df(p), p\rangle\, p, \tag{10}$$

where $Df(p) \in \mathbb{R}^{n+1}$ is the usual gradient of f at $p \in \Omega$. A *vector field* on $\Omega \subset \mathbb{S}^n$ is a smooth mapping $X : \Omega \to \mathbb{R}^{n+1}$ such that $X(p) \in T_p\mathbb{S}^n$. The *covariant derivative* of X at $p \in \Omega$ is map $\nabla X(p) : T_p\mathbb{S}^n \to T_p\mathbb{S}^n$ given by

$$\nabla X(p) := [I - pp^T] DX(p),$$

where $DX(p)$ is the usual derivative of X at p. The *Hessian on the sphere* of a twice differentiable function $f : \Omega \to \mathbb{R}$ at a point $p \in \Omega$ is the map $\nabla \operatorname{grad} f(p) := \operatorname{Hess} f(p) : T_p\mathbb{S}^n \to T_p\mathbb{S}^n$ given by

$$\operatorname{Hess} f(p) := \left[I - pp^T\right] \left[D^2 f(p) - \langle Df(p), p\rangle I\right], \tag{11}$$

where $D^2 f(p)$ is the usual Hessian (Euclidean Hessian) of the function f at a point p.

Let $I \subset \mathbb{R}$ be an open interval, $\Omega \subset \mathbb{S}^n$ an open set and $\gamma : I \to \Omega$ a geodesic segment. If $f : C \to \mathbb{R}$ is a differentiable function then, since $\gamma'(t) \in T_{\gamma(t)}\mathbb{S}^n$ for all $t \in I$, the equality (10) implies

$$\frac{d}{dt} f(\gamma(t)) = \left\langle \operatorname{grad} f(\gamma(t)), \gamma'(t)\right\rangle = \left\langle Df(\gamma(t)), \gamma'(t)\right\rangle, \qquad \forall\, t \in I. \tag{12}$$

and if the function f is twice differentiable then it holds that

$$\begin{aligned} \frac{d^2}{dt^2} f(\gamma(t)) &= \left\langle \operatorname{Hess} f(\gamma(t))\gamma'(t), \gamma'(t)\right\rangle \\ &= \left\langle D^2 f(\gamma(t))\gamma'(t), \gamma'(t)\right\rangle + \langle Df(\gamma(t)), \gamma(t)\rangle \left\langle \gamma'(t), \gamma'(t)\right\rangle, \qquad \forall\, t \in I. \end{aligned} \tag{13}$$

We end this section by stating two standard notations. We denote the *open* and the *closed ball* with radius $\delta > 0$ and center in $p \in \mathbb{S}^n$ by $B_\delta(p) := \{q \in \mathbb{S}^n : d(p, q) < \delta\}$ and $\bar{B}_\delta(p) := \{q \in \mathbb{S}^n : d(p, q) \leq \delta\}$ respectively.

2.1 Properties of the intrinsic distance

In this section, we present some important properties of the intrinsic distance from a fixed point. In particular, we present the spectral decomposition of the Hessian of the intrinsic distance.

The *intrinsic distance function on the sphere from the fixed point* $q \in \mathbb{S}^n$ is the mapping $d_q : \mathbb{S}^n \to \mathbb{R}$ defined by

$$d_q(p) := \arccos\langle p, q\rangle. \tag{14}$$

The intrinsic distance function on the sphere $\mathbb{S}^n$ satisfies the following important properties, which are an immediate consequence of its definition:

i) $d_p(q) = d_q(p)$, for all $p, q \in \mathbb{S}^n$;

ii) $0 \leq d_p(q) \leq \pi$, for all $p, q \in \mathbb{S}^n$;

iii) $d_q(p) = 0$ if and only if $p = q$;

iv) $d_q(p) = \pi$ if and only if $p = -q$.

Equation (9) can be rewritten as

$$d_q(p) = \|\exp_q^{-1} p\|, \qquad p, q \in \mathbb{S}^n. \tag{15}$$

The intrinsic distance d_q from q is twice differentiable at $p \in \mathbb{S}^n \backslash \{q, -q\}$. By combining (10) and (14), we can easily see that the *gradient of the distance from* q at $p \in \mathbb{S}^n \backslash \{q, -q\}$ is given by

$$\operatorname{grad} d_q(p) := -\frac{1}{\sqrt{1 - \langle p, q\rangle^2}} \left[I - pp^T\right] q. \tag{16}$$

Moreover, using (11) and (14), we obtain after some algebra that the *Hessian of the distance from* q at $p \in \mathbb{S}^n \backslash \{q, -q\}$ is given by

$$\operatorname{Hess} d_q(p) := \frac{\langle p, q\rangle}{\sqrt{1 - \langle p, q\rangle^2}} \left[I - pp^T\right] \left[I - \frac{1}{1 - \langle p, q\rangle^2} qq^T\right]. \tag{17}$$

Before presenting the spectral decomposition of the intrinsic distance from a fixed point on the sphere, we need a technical result.

Lemma 1. Let $p, q \in \mathbb{S}^n$. If $|\langle p, q\rangle| \neq 1$, then the following statements hold:

i) $\dim\left(T_p\mathbb{S}^n \cap T_q\mathbb{S}^n\right) = n - 1$;

ii) $\langle q - \langle p, q\rangle p, v\rangle = 0, \qquad \forall\, v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$;

as a consequence, taking an orthonormal base of the subspace $T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$, say $\{v_1, \ldots, v_{n-1}\}$ and defining $v_n = (q - \langle p, q\rangle p)/\|q - \langle p, q\rangle p\|$, the set $\{v_1, \ldots, v_{n-1}, v_n\}$ is an orthonormal base of $T_p\mathbb{S}^n$.

Proof. Elementary. $\square$

In the next lemma we present a spectral decomposition of the intrinsic distance from a fixed point on the sphere. The results in this lemma and the next one are closely related to Theorems IV.1 and Corollary IV.2 in [8].

Lemma 2. Take $q \in \mathbb{S}^n$ and let $\operatorname{Hess} d_q(p) : T_p\mathbb{S}^n \to T_p\mathbb{S}^n$ be the Hessian of the intrinsic distance from q at the point $p \in \mathbb{S}^n \backslash \{q, -q\}$. Then,

$$\operatorname{Hess} d_q(p)\ (q - \langle p, q\rangle p) = 0, \qquad \operatorname{Hess} d_q(p)\, v = \frac{\langle p, q\rangle}{\sqrt{1 - \langle p, q\rangle^2}}\, v, \qquad \forall\, v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n. \tag{18}$$

As a consequence, $\lambda_1 = 0$ and $\lambda_2 = \langle p, q\rangle/\sqrt{1 - \langle p, q\rangle^2}$ are the unique eigenvalues of $\operatorname{Hess} d_q(p)$, with algebraic multiplicity 1 and $n-1$, respectively. Moreover, if $\langle p, q\rangle \geq 0$, then the Hessian $\operatorname{Hess} d_q(p)$ is positive semidefinite, and if $\langle p, q\rangle \leq\ 0$, then the Hessian $\operatorname{Hess} d_q(p)$ is negative semidefinite.

Proof. Since $p \in \mathbb{S}^n \backslash \{q, -q\}$, we have $|\langle p, q\rangle| \neq 1$, which implies from (17) that the Hessian is well defined. As $q^T q = 1$, simple calculations give

$$\left[I - \frac{1}{1 - \langle p, q\rangle^2}\, qq^T\right] (q - \langle p, q\rangle p) = -\langle p, q\rangle p.$$

On the other hand, $[I - pp^T](-\langle p, q\rangle p) = 0$, which combined with the latter equality and (17), implies the first equality in (18), and we also have that λ_1 is an eigenvalue of the Hessian. For proving the second equality in (18), note that definitions of $T_p\mathbb{S}^n$ and $T_q\mathbb{S}^n$ imply that

$$p^T v = 0, \qquad q^T v = 0, \qquad \forall\, v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n.$$

So, the second inequality in (18) follows from (17) and the last two equalities. In particular, the Hessian is a multiple of the identity in the subspace $T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$ and, since $\dim T_p\mathbb{S}^n = n$, we conclude, using Lemma 1, that the eigenvalues λ_1 and λ_2 have algebraic multiplicity 1 and $n-1$, respectively, proving the first statement.

For proving the second statement, let $\{v_1, \ldots, v_{n-1}\}$ be an orthonormal base of the subspace $T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$. Since $|\langle p, q\rangle| \neq 1$, we can define $v_n = (q - \langle p, q\rangle p)/\|q - \langle p, q\rangle p\|$. So, Lemma 1 implies that $\{v_1, \ldots, v_{n-1}, v_n\}$ is an orthonormal base of $T_p\mathbb{S}^n$. Therefore, given $u \in T_p\mathbb{S}^n$, there exist $a_1, \ldots, a_{n-1}, a_n \in \mathbb{R}$ such that $u = a_1 v_1 + \cdots + a_{n-1} v_{n-1} + a_n v_n$, which, using the first statement, entails

$$\langle \operatorname{Hess} d_q(p) u, u\rangle = \lambda_2 (a_1^2 + \cdots + a_{n-1}^2),$$

completing the proof of the second statement. $\square$

Take $q \in \mathbb{S}^n$ and define $\rho_q : \mathbb{S}^n \to \mathbb{R}$ as

$$\rho_q(p) := \frac{1}{2} d_q^2(p). \tag{19}$$

Lemma 3. Take $q \in \mathbb{S}^n$ and define $\operatorname{Hess}\rho_q(p) : T_p\mathbb{S}^n \to T_p\mathbb{S}^n$ as the Hessian of ρ_q at the point $p \in \mathbb{S}^n \backslash \{q, -q\}$. Then, the following equalities hold:

$$\operatorname{Hess}\rho_q(p)\ (q - \langle p, q\rangle p) = q - \langle p, q\rangle p, \qquad \operatorname{Hess}\rho_q(p)\, v = \frac{\langle p, q\rangle \arccos\langle p, q\rangle}{\sqrt{1 - \langle p, q\rangle^2}}\, v, \tag{20}$$

for all $v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$, As a consequence, $\mu_1 = 1$ and $\mu_2 = \langle p, q\rangle \arccos\langle p, q\rangle / \sqrt{1 - \langle p, q\rangle^2}$ are the unique eigenvalues of $\operatorname{Hess}\rho_q(p)$, with algebraic multiplicity 1 and $n-1$, respectively. Moreover, if $\langle p, q\rangle > 0$, then the Hessian $\operatorname{Hess}\rho_q(p)$ is positive definite.

Proof. Using the definition of ρ_q in (19) and (11), it is easy to conclude, after some algebra, that

$$\operatorname{Hess}\rho_q(p) = d_q(p) \operatorname{Hess} d_q(p) + [I - pp^T] Dd_q(p) Dd_q(p)^T, \tag{21}$$

where $Dd_q(p)$ is the usual derivative of d_q at the point p. Since $Dd_q(p) = -q/\sqrt{1 - \langle p, q\rangle^2}$, we have

$$Dd_q(p) Dd_q(p)^T = \frac{1}{1 - \langle p, q\rangle^2}\, qq^T. \tag{22}$$

As $q^T q = 1$, it follows from the last equality that $Dd_q(p) Dd_q(p)^T (q - \langle p, q\rangle p) = q$. On the other hand, $[I - pp^T]q = q - \langle p, q\rangle p$. Hence, we obtain that

$$[I - pp^T] Dd_q(p) Dd_q(p)^T (q - \langle p, q\rangle p) = q - \langle p, q\rangle p.$$

Therefore, combining the last equality, equation (21) and the first equality in (18), we get that

$$\operatorname{Hess}\rho_q(p)(q - \langle p, q\rangle p) = q - \langle p, q\rangle p,$$

which is the first equality in (20). For proving the second one, note first that the definition of $T_q\mathbb{S}^n$ implies that $q^T v = 0$ for all $v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$. So, using (22), we have

$$[I - pp^T] Dd_q(p) Dd_q(p)^T v = 0, \qquad \forall\, v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n.$$

Hence, equation (21) implies that $\operatorname{Hess}\rho_q(p)v = d_q(p) \operatorname{Hess} d_q(p)v$ for all $v \in T_p\mathbb{S}^n \cap T_q\mathbb{S}^n$. Thus, using the second equality in (18) and the definition of $d_q(p)$ in (14), the second equality in (20) follows. The remainder of our proof requires arguments similar to those in the proof of Lemma 2 (note that in the final part of the proof we must invoke the fact that $\arccos\langle p, q\rangle > 0$, which holds because $p \neq q$). $\square$

The *distance* to a set $C \in \mathbb{S}^n$ is the function $d_C : \mathbb{S}^n \to \mathbb{R}$ defined by

$$d_C(p) := \inf \{ d_p(q) \ : \ q \in C \}. \tag{23}$$

Since the sphere endowed with the Riemannian distance is a metric space we have the following results.

Proposition 1. Let $C \in \mathbb{S}^n$ be a nonempty subset. Then

$$|d_C(p) - d_C(q)| \leq d(p, q), \qquad \forall\ p, q \in \mathbb{S}^n.$$

In particular, the function d_C is continuous.

3 Convex sets on the sphere

In this section we present some properties of the convex sets of the sphere. It is worth to remark that the convex sets on the sphere $\mathbb{S}^n$ are closely related to the pointed convex cones in the Euclidean space $\mathbb{R}^{n+1}$.

Definition 1. The set $C \subseteq \mathbb{S}^n$ is said to be *spherically convex* if for any p, $q \in C$ all the minimal geodesic segments joining p to q are contained in C.

We assume for convenience that *from now on all spherically convex sets are nonempty proper subsets of the sphere*.

For each set $A \subset \mathbb{S}^n$, let $K_A \subset \mathbb{R}^{n+1}$ be the *cone spanned by* A, namely,

$$K_A := \{tp \,:\, p \in A,\ t \in [0, +\infty)\}\,. \tag{24}$$

Clearly, K_A is the smallest cone which contains A. In the next result we relate a spherically convex set with the cone spanned by it, but first we need another definition. A convex cone $K \subset \mathbb{R}^{n+1}$ is said to be *pointed* if $K \cap (-K) \subseteq \{0\}$, or equivalently, if K does not contain straight lines through the origin. The following result is proved in [7].

Proposition 2. The set C is spherically convex if and only if the cone K_C is convex and pointed.

Let $C \subset \mathbb{S}^n$ be a spherically convex set. The *spherical polar set* of the set C is intrinsically defined by

$$C^{\ominus} := \left\{q \in \mathbb{S}^n \,:\, d(p,q) \geq \frac{\pi}{2}, \forall\, p \in C\right\}. \tag{25}$$

Since the function $[-1, 1] \ni t \mapsto \arccos(t)$ is decreasing, it is easy to conclude that

$$C^{\ominus} = \{q \in \mathbb{S}^n \,:\, \langle p, q\rangle \leq 0, \forall\, p \in C\}. \tag{26}$$

Let $K^{-} := \{y \in R^{n+1} : \langle x, y\rangle \leq 0,\ \forall\, x \in K\}$ be the *polar cone* of the cone K, K_C^{-} be the polar cone of the cone K_C and $K_{C^{\ominus}}$ be the cone spanned by $C^{\ominus}$, as defined in (24). The next proposition is an immediate consequence of (24), together with the definition and properties of the polar cone.

Proposition 3. Let $C \subset \mathbb{S}^n$ be a spherically convex set with nonempty (intrinsic) interior. The polar set $C^{\ominus}$ of C satisfies the following properties:

(i) $K_{C^{\ominus}} = K_C^{-}$;

(ii) K_C^{-} is pointed. As a consequence, $C^{\ominus}$ is spherically convex;

(iii) $C^{\ominus}$ is always closed. Furthermore, $C^{\ominus\ominus}$ is equal to the closure of C.

We define a hemisphere of the sphere as a certain sub-level of the intrinsic distance from a fixed point. More precisely, the *open hemisphere* and the *closed hemisphere* with *pole* $p \in \mathbb{S}^n$ are defined by

$$S_p^n := \{q \in \mathbb{S}^n : d(p,q) < \pi/2\} = \{q \in \mathbb{S}^n : \langle p, q\rangle > 0\}$$

and

$$\bar{S}_p^n := \{q \in \mathbb{S}^n : d(p,q) \le \pi/2\} = \{q \in \mathbb{S}^n : \langle p, q\rangle \ge 0\},$$

respectively. The following result is proved in [7].

Corollary 1. If $C \subset \mathbb{S}^n$ is a closed spherically convex set, then there exist $p \in \mathbb{S}^n$ such that $C \subset S_p^n$.

Let $C \subset \mathbb{S}^n$ be a closed spherically convex set and $\mathbb{P}(C)$ the set of all subsets of C. The *projection mapping* $P_C(\,.\,) : \mathbb{S}^n \to \mathbb{P}(C)$ *onto the set* C is defined by

$$\begin{aligned} P_C(p) &:= \{\bar{p} \in C : d(p,\bar{p}) \le d(p,q), \forall\, q \in C\} \\ &= \{\bar{p} \in C : \langle p, q\rangle \le \langle p, \bar{p}\rangle, \forall\, q \in C\}\,, \end{aligned} \tag{27}$$

that is, it is the set of minimizers of the function $C \ni q \mapsto d(p,q)$. The minimal value of the function $C \ni q \mapsto d(p,q)$ is called the *distance of* p *from* C and it is denoted by $d_C(p)$. Hence, using this new notation, and equations (25), we can rewrite the spherical polar of C as

$$C^{\ominus} = \left\{p \in \mathbb{S}^n \,:\, d_C(p) \ge \frac{\pi}{2}\right\}.$$

An immediate consequence of the second equality in (27) is the montonicity of the projection mapping (see [7]), stated as follows:

Corollary 2. Let $C \subset \mathbb{S}^n$ be a nonempty closed spherically convex set. Then the projection mapping $P_C(\,.\,) : \mathbb{S}^n \to \mathbb{P}(C)$ onto the set C satisfies

$$\langle \bar{p} - \bar{q}, p - q\rangle \ge 0, \qquad \forall\, \bar{p} \in P_C(p),\ \forall\, \bar{q} \in P_C(q).$$

The next two results are important properties of the projection onto the set C; the proofs can be found in [7].

Proposition 4. Let $C \subset \mathbb{S}^n$ be a closed spherically convex set. Consider $p \in \mathbb{S}^n$ and $\bar{p} \in C$. If $\bar{p} \in P_C(p)$, then

$$\left\langle \left(I - \bar{p}\bar{p}^T\right) p, \left(I - \bar{p}\bar{p}^T\right) q\right\rangle \le 0, \qquad \forall\, q \in C,$$

or equivalently, $\left(I - \bar{p}\bar{p}^T\right) p = p - \langle p, \bar{p}\rangle \bar{p} \in K_{C^{\ominus}}$.

Proposition 5. Let $C \subset \mathbb{S}^n$ be a closed spherically convex set. Consider $p \in \mathbb{S}^n$ and $\bar{p} \in C$ and assume that $\langle p, \bar{p}\rangle > 0$. The following statements are equivalent:

i) $\bar{p} \in P_C(p)$

ii) $\left\langle \left(I - \bar{p}\bar{p}^T\right) p, \left(I - \bar{p}\bar{p}^T\right) q \right\rangle \le 0$, for all $\quad q \in C$.

iii) $\left(I - \bar{p}\bar{p}^T\right) p = p - \langle p, \bar{p}\rangle \bar{p} \in K_{C^\ominus}$.

Moreover, $P_C(p)$ is a singleton.

Proposition 6. Let $C \subset \mathbb{S}^n$ be a nonempty closed convex set. Let $\bar{p} \in \mathbb{S}^n$ and assume that $C \subset \mathbb{S}^n_{\bar{p}}$. Then $P_C : \mathbb{S}^n_{\bar{p}} \to C$ is continuous.

Proof. Let $\{p^k\} \subset \mathbb{S}^n_{\bar{p}}$ be such that $\lim_{k\to+\infty} p^k = p$. Since $\{P_C(p^k)\} \subset \mathbb{S}^n_{\bar{p}}$, the sequence $\{P_C(p^k)\}$ is bounded. Let q be a cluster point of $\{P_C(p^k)\}$ and let $\{p^{k_j}\}$ be such that $\lim_{k\to+\infty} P_C(p^{k_j}) = q$. As $C \subset \mathbb{S}^n_{\bar{p}}$, it follows from Proposition 5 that $\{P_C(p^k)\}$ is a singleton. Hence,

$$d_C(p^{k_j}) = d(p^{k_j}, P_C(p^{k_j})), \qquad \forall\, k_j.$$

Using Proposition 1 and letting $j \to \infty$, we have $d_C(p) = d(p,q)$. Since C is closed, it follows that $q \in C$, which together with (23), (27) and $d_C(p) = d(p,q)$ imply that $q = P_C(p)$, because $P_C(p)$ is a singleton. Therefore, the sequence $\{P_C(p^k)\}$ has only one cluster point, namely, $P_C(p)$. Thus, $\lim_{k\to+\infty} P_C(p^k) = P_C(p)$ and the proof is concluded. □

4 Convex functions on the sphere

In this section we study the basic properties of convex functions on the sphere. In particular, we present the first and second order characterizations of differentiable convex functions on the sphere.

Definition 2. Let $C \subset \mathbb{S}^n$ be a spherically convex set and $I \subset \mathbb{R}$ an interval. A function $f : C \to \mathbb{R}$ is said to be spherically convex (respectively, strictly spherically convex) if for any minimal geodesic segment $\gamma : I \to C$, the composition $f \circ \gamma : I \to \mathbb{R}$ is convex (respectively, strictly convex) in the usual sense.

It follows from the above definition that $f : C \to \mathbb{R}$ is a spherically convex function if and only if the *epigraph*

$$\mathrm{epi} f := \{(p, \mu) : p \in C,\ \mu \in \mathbb{R},\ \mu \ge f(p)\},$$

is convex in $\mathbb{S}^n \times \mathbb{R}$. Moreover, if $f : C \to \mathbb{R}$ is a spherically convex function, then the sub-level sets $\{p \in C : \ f(p) \le k\}$ are spherically convex sets for all $k \in \mathbb{R}$.

Remark 1. If $C = \mathbb{S}^n$ and $f : C \to \mathbb{R}$ is spherically convex, then f is constant; that is, there is no non-contant spherically convex function defined on the whole sphere. Of course, there exist non-constant spherically convex functions defined on proper spherically convex subsets $C \subset \mathbb{S}^n$. We will present several examples at the end of this section.

Proposition 7. Let $C \subset \mathbb{S}^n$ be an open spherically convex set and $f : C \to \mathbb{R}$ a differentiable function. The function f is spherically convex if and only if

$$f(q) \geq f(p) + \langle \operatorname{grad} f(p), \exp_p^{-1} q \rangle, \qquad \forall\, p, q \in C,\ q \neq p,$$

or equivalently,

$$f(q) \geq f(p) + \frac{\arccos\langle p, q\rangle}{\sqrt{1 - \langle p, q\rangle^2}} \left\langle Df(p)\,,\, [I - pp^T]q \right\rangle, \qquad \forall\, p, q \in C,\ q \neq p,$$

Proof. Using (12), the usual characterization of scalar convex functions implies that, for all minimal geodesic segment $\gamma : I \to C$, the composition $f \circ \gamma : I \to \mathbb{R}$ is convex if and only if

$$f(\gamma(t_2)) \geq f(\gamma(t_1)) + \left\langle Df(\gamma(t_1)), \gamma'(t_1) \right\rangle (t_2 - t_1), \qquad \forall\ t_2, t_1 \in I.$$

Note that if $\gamma : [0, 1] \to C$ is the minimal geodesic segment from $p = \gamma(0)$ to $q = \gamma(1)$, then it may be represented as $\gamma(t) = \exp_p t \exp_p^{-1} q$ and $\gamma'(0) = \exp_p^{-1} q$. Therefore, the first inequality of the proposition is an immediate consequence of the inequality above, Definition 2 and equation (8). For concluding the proof, note that equations (10) and (8) imply the equivalence between the two inequalities of the lemma. □

Proposition 8. Let $C \subset \mathbb{S}^n$ be an open spherically convex set and $f : C \to \mathbb{R}$ a differentiable function. The function f is spherically convex if and only if the gradient vector field $\operatorname{grad} f$ on the sphere satisfies the following inequality

$$\langle \operatorname{grad} f(p), \exp_p^{-1} q \rangle + \langle \operatorname{grad} f(q), \exp_q^{-1} p \rangle \leq 0, \qquad \forall\, p, q \in C.$$

or equivalently,

$$\langle Df(p) - Df(q), p - q \rangle + (\langle p, q\rangle - 1) \left[\langle Df(p), p \rangle + \langle Df(q), q \rangle \right] \geq 0, \qquad \forall\, p, q \in C.$$

Proof. Using (12), the usual first order characterization of convex functions implies that, for all minimal geodesic segments $\gamma : I \to C$, the composition $f \circ \gamma : I \to \mathbb{R}$ is convex if and only if

$$\left[\left\langle Df(\gamma(t_2)), \gamma'(t_2) \right\rangle - \left\langle Df(\gamma(t_1)), \gamma'(t_1) \right\rangle \right] (t_2 - t_1) \geq 0, \qquad \forall\ t_2, t_1 \in I.$$

Note that if $\gamma : [0, 1] \to C$ is the segment of minimal geodesic from $p = \gamma(0)$ to $q = \gamma(1)$, then it may be represented as $\gamma(t) = \exp_p t \exp_p^{-1} q$ and $\gamma'(0) = \exp_p^{-1} q$. Therefore, the first inequality of the proposition follows by combining the previous inequality with Definition 2 and (8). For concluding the proof, note that equations (10) and (8) imply the equivalence between the two inequalities of the lemma. □

Proposition 9. Let $C \subset \mathbb{S}^n$ be a spherically convex set and $f : C \to \mathbb{R}$ be a twice differentiable function. The function f is spherically convex if and only if the Hessian $\operatorname{Hess} f$ on the sphere satisfies the following inequality

$$\langle \operatorname{Hess} f(p)v, v\rangle \geq 0, \qquad \forall\, p \in \mathbb{S}^n, \forall\, v \in T_p\mathbb{S}^n,$$

or equivalently,

$$\left\langle D^2 f(p)v, v\right\rangle + \langle Df(p), p\rangle \langle v, v\rangle \geq 0, \qquad \forall\, p \in \mathbb{S}^n, \forall\, v \in T_p\mathbb{S}^n,$$

where $D^2 f(p)$ is the usual Hessian and $Df(p)$ is the usual gradient of f at a point $p \in \Omega$. If the above inequalities are strict then f is strictly spherically convex.

Proof. Using (13), the usual second order characterization of spherically convex functions implies that, for all minimal geodesic segment $\gamma : I \to C$, the composition $f \circ \gamma : I \to \mathbb{R}$ is convex if and only if

$$\left\langle D^2 f(\gamma(t))\gamma'(t), \gamma'(t)\right\rangle + \langle Df(\gamma(t)), \gamma(t)\rangle \left\langle \gamma'(t), \gamma'(t)\right\rangle \geq 0, \qquad \forall\, t \in I.$$

If the last inequality is strict then $f\circ\gamma$ is strictly convex. Therefore, the result follows by combining the above inequality with Definition 2. For concluding the proof, note that equation (11) implies the equivalence between the two inequalities of the lemma. □

Example 1. Fix $q \in \mathbb{S}^n$. The function $d_q(\cdot) : B_{\pi/2}(q) \to \mathbb{R}$ is spherically convex.

In general, taking a spherically convex set $C \subset B_{\pi/2}(p)$, the function $d_q(\cdot) : C \to \mathbb{R}$ is spherically convex. Indeed, since $-q \notin B_{\pi/2}(q)$, the spherical convexity of $d_q(\cdot)$ follows by combining Lemma 2 with Proposition 9.

Example 2. Fix $q \in \mathbb{S}^n$. The function define $\rho_q : S^n_q :\to \mathbb{R}$ as

$$\rho_q(p) := \frac{1}{2} d^2_q(p).$$

is strictly spherically convex. In general, taking a spherically convex set $C \subset S^n_q$, the function $\rho_q : C \to \mathbb{R}$ is strictly spherically convex. Indeed, since $-q \notin S^n_q$, the spherical convexity of ρ_q follows by combining Lemma 3 with Proposition 9.

Example 3. Take $\tilde{p} = (0, \cdots, 0, 1) \in R^{n+1}$ and $C = \{p = (p_1 \cdots, p_{n+1}) \in \mathbb{S}^n \,:\, p_{n+1} > 0\}$. The function $\psi : C \to \mathbb{R}$ defined by $\psi(p) = -\ln(\pi/2 - d(\tilde{p}, p))$ is spherically convex. Indeed, since $-\tilde{p} \notin C$ and $\pi/2 - d(\tilde{p}, p) > 0$, the spherical convexity of ψ follows by combining equation (11), Lemma 2 and Proposition 9.

Example 4. Let $p = (p_1 \cdots, p_{n+1})$ and $S_{++} = \{p \in \mathbb{S}^n \,:\, p_1 > 0, \cdots, p_{1+n} > 0\}$. The function $\varphi : S_{++} \to \mathbb{R}$ defined by $\varphi(p) = -\sum_{i=1}^{n+1} \ln(p_i)$ is spherically convex. The spherical convexity of φ follows from equation (11) and Proposition 9.

5 Optimization problems on the sphere

In this section we will present sufficient optimality conditions for constrained optimization problems on the sphere. Let $\Omega \subset \mathbb{S}^n$ be an open set and $f : \Omega \to \mathbb{R}$ be a differentiable function. Consider the following nonlinear programming problem

$$\min\{f(p)\ :\ p \in C\}. \tag{28}$$

Proposition 10. Let $C \subset \Omega$ be a spherically convex set . If $\bar{p} \in C$ is a solution of the problem (28) then

$$\left\langle Df(\bar{p}), [I - \bar{p}\bar{p}^T]p \right\rangle = \left\langle \operatorname{grad} f(\bar{p}), [I - \bar{p}\bar{p}^T]p \right\rangle \geq 0, \quad \forall\ p \in C.$$

Proof. The above equality follows easily from (10). Take $p \in C$ and let $\bar{p} \in C$ be a solution to problem (28). Let

$$[0, 1] \ni t \mapsto \gamma(t) = \exp_{\bar{p}}(t \exp_{\bar{p}}^{-1} p),$$

be the geodesic from $\bar{p}$ to p. Since C is spherically convex and $p, \bar{p} \in C$, we conclude that $\gamma(t) \in C$ for all $t \in [0, 1]$. Hence, as $\bar{p} \in C$ is a solution to the problem in (28), we have

$$\frac{f(\gamma(t)) - f(\bar{p})}{t} \geq 0, \qquad \forall\, t \in [0, 1].$$

Taking the limit in the above inequality when t tends to zero, we obtain, using (12), that

$$\langle \operatorname{grad} f(\bar{p}), \gamma'(0)\rangle \geq 0.$$

As $\gamma'(0) = \exp_{\bar{p}}^{-1} p$, the result follows from the previous inequality by using (8) and taking in account that $\arccos\langle \bar{p}, p\rangle \geq 0$. $\square$

Proposition 11. Let $C \subset \Omega$ be a spherically convex set and f be a spherically convex function in C. The point $\bar{p} \in C$ is a solution of the problem in (28) if and only if

$$\left\langle Df(\bar{p}), [I - \bar{p}\bar{p}^T]p \right\rangle = \left\langle \operatorname{grad} f(\bar{p}), [I - \bar{p}\bar{p}^T]p \right\rangle \geq 0, \qquad \forall\, p \in C.$$

Proof. If the point $\bar{p} \in C$ is a solution of (40) then the inequality follows from Proposition 10. Conversely, take p, $\bar{p} \in C$, $p \neq \bar{p}$ and assume that $\left\langle Df(\bar{p}), [I - \bar{p}\bar{p}^T]p \right\rangle \geq 0$. As f is spherically convex in C, we conclude from Proposition 7 that

$$f(p) \geq f(\bar{p}) + \frac{\arccos\langle \bar{p}, p\rangle}{\sqrt{1 - \langle \bar{p}, p\rangle^2}} \left\langle Df(\bar{p})\,,\, [I - \bar{p}\bar{p}^T]p \right\rangle, \qquad \forall\, p \in C,\ p \neq \bar{p}.$$

Since $\left\langle Df(\bar{p})\,,\, [I - \bar{p}\bar{p}^T]p \right\rangle \geq 0$ and $\arccos\langle \bar{p}, p\rangle \geq 0$, the latter inequality implies that $f(p) \geq f(\bar{p})$, for all $p \in C$. Hence, $\bar{p}$ is a solution of the problem in (40). $\square$

In view of the well known optimality conditions for spherically convex optimization problems, the proof of the next result is an immediate consequence of the definitions of the intrinsic distance and the projection.

Corollary 3. Let $C \subset \mathbb{S}^n$ be a nonempty closed spherically convex set. Take $q \in \mathbb{S}^n$ and $\bar{q} \in C$. Assume that $\langle q, \bar{q} \rangle > 0$. The following statements are equivalent:

i) $\bar{q} \in P_C(q)$.

ii) $\left\langle [I - \bar{q}\bar{q}^T]q, [I - \bar{q}\bar{q}^T]p \right\rangle \leq 0, \ \forall\, p \in C$.

iii) $[I - \bar{q}\bar{q}^T]q = q - \langle q, \bar{q} \rangle \bar{q} \in K_{C^{\ominus}}$.

Moreover, $P_C(q)$ is a singleton.

Proof. First note that $C \cap S^n_q$ is spherically convex. Since $\langle q, \bar{q} \rangle > 0$, we have $\bar{q} \in C \cap S^n_q$. Hence, from the definition of the projection in (27) we conclude that $\bar{q} \in P_C(q)$ if and only if $\bar{q} = \operatorname{argmin}\{\rho_q(p) : p \in C \cap S^n_q\}$, where $\rho_q(p) := d^2_q(p)/2$. Therefore, since ρ_q is strictly spherically convex in $C \cap S^n_q$ and

$$\operatorname{grad} \rho_q(\bar{q}) = -\frac{\arccos\langle q, \bar{q} \rangle}{\sqrt{1 - \langle q, \bar{q} \rangle^2}} \left[I - \bar{q}\bar{q}^T \right] q, \tag{29}$$

the equivalence of items (i) and (ii) follows by applying Proposition 11. Moreover, the strict spherical convexity of ρ_q implies that $P_C(q)$ is a singleton. The equivalence of items (ii) and (iii) follows trivially from the equality

$$\left\langle [I - \bar{q}\bar{q}^T]q, [I - \bar{q}\bar{q}^T]p \right\rangle = \left\langle [I - \bar{q}\bar{q}^T]q, p \right\rangle,$$

equation (24) and the definition of the spherical polar. □

The *intrinsic diameter* of a nonempty closed spherically convex set $C \subset \mathbb{S}^n$ is defined as the maximum of the intrinsic distance between two points of the set C; that is,

$$\operatorname{diam}(C) := \sup\{d(p, q) : p, q \in C\}, \tag{30}$$

where d is the intrinsic distance on the sphere as defined in (3). The next definition is equivalent to the definition of antipodal pair of a convex cone given by Iusem and Seeger in [11, 12].

Definition 3. Let $C \subset \mathbb{S}^n$ be a nonempty closed spherically convex set. The pair $(u, v) \in \mathbb{S}^n \times \mathbb{S}^n$ is called an antipodal pair of C if $u, v \in C$ and $d(u, v) = \operatorname{diam}(C)$.

Let $C \subset \mathbb{S}^n$ be a closed spherically convex set. The *spherical dual set* of C is defined by

$$C^{\oplus} := \left\{ q \in \mathbb{S}^n : d_p(q) \leq \frac{\pi}{2}, \ \forall\, p \in C \right\}. \tag{31}$$

Since the function $[-1,1] \ni t \mapsto \arccos(t)$ is decreasing, it is easy to conclude from (14) that

$$C^{\oplus} := \{q \in \mathbb{S}^n : \langle p, q\rangle \geq 0, \ \forall\, p \in C\}\,. \tag{32}$$

Equations (26) and (32) imply that $C^{\oplus} = -C^{\ominus}$. In view of (32), the next theorem is an immediate consequence of Theorem 4.1 of [11]. For the sake of completeness we present here an intrinsic proof.

Theorem 1. Let $C \subset \mathbb{S}^n$ be a nonempty closed spherically convex set with nonempty interior. If (u, v) is an antipodal pair of C and $v \neq u$ then it holds that:

$$\frac{u - \langle u, v\rangle v}{\sqrt{1-\langle u, v\rangle^2}} \in C^{\oplus}, \qquad\qquad \frac{v - \langle u, v\rangle u}{\sqrt{1-\langle u, v\rangle^2}} \in C^{\oplus}. \tag{33}$$

Proof. The definition of antipodal pairs of C in (30) implies that $v \in C$ is a maximizer of the distance function $d_u : \mathbb{S}^n \backslash \{u, -u\} \to \mathbb{R}$,

$$d_u(p) = \arccos\langle u, p\rangle,$$

on the spherically convex set C, that is, $v = \operatorname{argmin}\{-d_u(p) : p \in C\}$. Hence, using Proposition 10, we conclude that $v \in C$ satisfies:

$$\left\langle Dd_u(v), [I - vv^T]p \right\rangle \leq 0, \qquad p \in C.$$

Since $Dd_u(v) = -u/\sqrt{1-\langle u, v\rangle^2}$, we obtain from the previous inequality that $\langle u, [I - vv^T]p\rangle \geq 0$, for all $p \in C$, or equivalently, $\langle [I - vv^T]u, p\rangle \geq 0$, for all $p \in C$, which, taking into account (32), is equivalent to the first inclusion in the statement of the theorem. A similar argument can be used to prove the second inclusion. □

6 Variational problem on the sphere

In this section we define variational problems in the sphere and study some of their basic properties; for instance, we give a characterization of their solution sets.

In this section we assume that all spherically convex sets $C \subset \mathbb{S}^n$ have nonempty (intrinsic) interior. The *normal cone mapping* associated to the set C on the sphere $C \ni p \mapsto N_C(p) \in T_p\mathbb{S}^n$ is defined by

$$N_C(p) := \begin{cases} \{v \in T_p\mathbb{S}^n : \langle v, q\rangle \leq 0, \ \forall\, q \in C\}\,, & \text{for} \quad p \in C, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{34}$$

From (8) it is easy to see that $\langle v, q\rangle \leq 0$ if only if $\langle v, \exp_p^{-1} q\rangle \leq 0$. Therefore, since in the Euclidean space $\mathbb{R}^{n+1}$ we have $\exp_p^{-1} q = q - p$, we conclude that the above definition extends the usual definition of normal cone mapping from the Euclidean space to the sphere. The next proposition is an immediate consequence of (24), (26) and (34) .

Proposition 12. Let $C \subset \mathbb{S}^n$ be a spherically convex set. For each $p \in C$ there holds

$$N_C(p) = T_p\mathbb{S}^n \cap K_{C^\ominus} = \left\{ v \in \mathbb{R}^{n+1} \ : \ \langle v, p\rangle = 0,\ \langle v, q\rangle \leq 0,\ \ \forall\, q \in C \right\}.$$

Corollary 4. Let $C \subset \mathbb{S}^n$ be a spherically convex and $p, q \in \mathbb{S}^n$. The following items are equivalent:

i) $q \in N_C(p)$;

ii) $p \in N_{C^\ominus}(q)$;

ii) $p \in C,\ q \in C^\ominus,\ \langle p, q\rangle = 0$.

Proof. The result is immediate by combining (26) and Proposition 12. $\square$

Let X be a vector field on the sphere and $C \subset \mathbb{S}^n$ be a closed spherically convex set. The *spheric variational problem* associated to X and C is defined as the inclusions

$$p \in C \subset \mathbb{S}^n, \qquad X(p) + N_C(p) \ni 0. \tag{35}$$

From the definition in (34) and the definition of the tangent plane $T_p\mathbb{S}^n$, (35) is equivalent to

$$p \in C, \qquad \langle X(p), p\rangle = 0, \quad \langle X(p), q\rangle \geq 0, \quad \forall\, q \in C. \tag{36}$$

Remark 2. When p is in the intrinsic interior of C is easy to see that (36) is equivalent to the equation $X(p) = 0$.

Using the definition of the dual spheric set $C^\oplus$ and (24), the conditions in (36) are equivalent to

$$p \in C, \qquad \langle X(p), p\rangle = 0, \qquad X(p) \in K_{C^\oplus}. \tag{37}$$

Remark 3. If $C = \mathbb{R}^{n+1}_+ \cap \mathbb{S}^n$ then the latter conditions become:

$$p \geq 0, \qquad X(p) \geq 0, \qquad \langle X(p), p\rangle = 0, \qquad p \in \mathbb{S}^n, \tag{38}$$

which define the *spheric complementarity problem.*

Proposition 13. Let $C \subset \mathbb{S}^n$ be a closed spherically convex set, X be a vector field on $\mathbb{S}^n$ and $p \in C$. Then $X(p) + N_C(p) \ni 0$ if and only if $P_C(\exp_p(-rX(p))) = p$ for all $r > 0$ such that $\|rX(p)\| < \pi$.

Proof. The result is trivial if $X(p) = 0$. Now, we assume that $X(p) \neq 0$. Since $X(p) \in T_p\mathbb{S}^n$, i.e., $\langle X(p), p\rangle = 0$, we conclude from Proposition 12 that $-X(p) \in N_C(p)$ if and only if $-X(p) \in K_{C^\ominus}$. As $X(p) \in T_p\mathbb{S}^n$ and $\|rX(p)\| < \pi$, we have $\cos(\|rX(p)\|) = \left\langle \exp_p(-rX(p)), p \right\rangle$. Thus using (6) we have

$$\exp_p(-rX(p)) - \left\langle \exp_p(-rX(p)), p \right\rangle p = -\sin(\|rX(p)\|)\, \frac{X(p)}{\|X(p)\|}.$$

Hence, as $\sin(\|rX(p)\|) \geq 0$ and $-X(p) \in N_C(p)$ if only if $-X(p) \in K_{C^\ominus}$, we conclude that $-X(p) \in N_C(p)$ if only if

$$\exp_p(-rX(p)) - \left\langle \exp_p(-rX(p)), p \right\rangle p \in K_{C^\ominus},$$

for all $r > 0$ such that $\|rX(p)\| < \pi$. Therefore, the result follows from the equivalence between items i) and iii) of Proposition 5 (see Proposition 6 of [7]). □

Let X be a vector field on the sphere $\mathbb{S}^n$ and $r > 0$. We define the map $\Phi : \mathbb{S}^n \to \mathbb{S}^n$ as $\Phi(p) = \exp_p X(p)$, that is

$$\Phi(p) := \begin{cases} \cos(\|rX(p)\|)\, p + \sin(\|rX(p)\|)\, \dfrac{X(p)}{\|X(p)\|}, & X(p) \neq 0, \\ p, & X(p) = 0. \end{cases} \tag{39}$$

Note that if X is continuous then the map Φ is also continuous. The next proposition was first proved in a more general setting in [16] (see also [15]). Here we will give a proof which uses Proposition 13.

Proposition 14. Let $C \subset \mathbb{S}^n$ be a closed spherically convex set and X be a vector field on the sphere $\mathbb{S}^n$. If X is continuous then the spheric variational problem (35) associated to C has a closed and nonempty solution set.

Proof. Take $r > 0$ and let $\Phi : \mathbb{S}^n \to \mathbb{S}^n$ be as defined in (39). Then, since C is spherically convex and the map Φ is continuous, we conclude from Proposition 6 that the function $\Psi : C \to C$ defined by

$$\Psi(p) = P_C \circ \Phi(p)) = P_C(\exp_p(-rX(p))),$$

is continuous. From Proposition 13 the solution set of the the variational problem (35) associated to C is

$$F = \{p \in C \;:\; \Psi(p) = p\},$$

which, due to the continuity of the map Φ, is closed. From Proposition 2 the cone K_C is convex and pointed. By pointedness of C we have $\operatorname{int}(K_C^+) \neq \emptyset$, where $K_C^+ := \{p \in R^{n+1} : \langle p, q\rangle \geq 0, \ \forall\, q \in K_C\}$ is the dual of the cone K_C. Take $\tilde{p} \in \operatorname{int}(K_C^+)$, $\alpha > 0$, and consider the sets

$$H_{\tilde{p},\alpha} := \left\{q \in R^{n+1} : \langle \tilde{p}, q\rangle = \alpha,\ \|q\| \leq 1\right\}, \qquad S^n_{\tilde{p},\alpha} := \{q \in \mathbb{S}^n : \langle \tilde{p}, q\rangle \geq \alpha\}.$$

Let $D = K_C \cap H_{\tilde{p},\alpha}$. Define now $f : H_{\tilde{p},\alpha} \to S^n_{\tilde{p},\alpha}$ as $f(q) = ||q||^{-1}q$. It is easy to see that f is a homeomorphism between $H_{\tilde{p},\alpha}$ and $S^n_{\tilde{p},\alpha}$, and hence its restriction to D is a homeomorphism between D and its image $f(D)$. We claim that $f(D) = C$. Clearly, $f(D) \subset C$, because for $q \in D \subset K$, we have $f(q) \in K$, since $f(q)$ is a positive multiple of q, and so $f(q) \in K \cap S^n = C$. Take now $p \in C$. For checking that $p = f(q)$ for some q in D, it suffices to prove that there exists $\beta > 0$ such that

βp belongs to D, i.e. such that $\alpha = \langle \tilde{p}, \beta p\rangle = \beta\langle \tilde{p}, p\rangle$, which occurs for $\beta = \alpha/\langle \tilde{p}, p\rangle$, which is well defined because $\langle \tilde{p}, p\rangle > 0$, using the facts that $\tilde{p} \in \operatorname{int}(K_C^+)$ and $p \in C \subset K$. Hence the claim is established, so that $f(D) = C$ and hence C is homeomorphic to D. Now, compactness of C implies compactness of D, and also D is convex because it is the intersection of two convex sets: the cone K_C and the hyperplane $H_{\tilde{p},\alpha}$. Therefore, C is homeomorphic to a compact and convex set, namely, D. Since, the fixed point property is a topological property, we can apply Brouwer's fixed point theorem to conclude the existence of a least one such fixed point for the function Ψ, that is, $F \neq \emptyset$, concluding the proof. □

Proposition 15. Let $\Omega \subset \mathbb{S}^n$ be an open set, $C \subset \Omega$ be a closed spherically convex set and $f : \Omega \to \mathbb{R}$ be a differentiable function. If the point $\bar{p} \in C$ is a local solution of the optimization problem

$$\min\{f(p) \,:\, p \in C\}, \tag{40}$$

then $\bar{p} \in C$ is solution of the variational inequality

$$p \in C, \qquad \operatorname{grad} f(p) + N_C(p) \ni 0. \tag{41}$$

Moreover, if f is a spherically convex function in C then the point $\bar{p} \in C$ is a global solution of (40) if and only if it is a solution of (41).

Proof. Let $\bar{p} \in C$ be a local solution of (40). The spheric convexity of C implies that, for any $p \in C$ and $t \in [0, 1]$, we have

$$\gamma(t) = \exp_{\bar{p}}\left(t\exp_{\bar{p}}^{-1} p\right) \in C.$$

Since $\bar{p} \in C$ is a local solution of (40), the latter equality implies that 0 is a local minimum of $f \circ \gamma : [0, 1] \to \mathbb{R}$. So,

$$\langle \operatorname{grad} f(\bar{p}), \exp_{\bar{p}}^{-1} p\rangle = (f \circ \gamma)'(0) \geq 0, \qquad \forall\, p \in C.$$

As $\operatorname{grad} f(\bar{p}) \in T_{\bar{p}}\mathbb{S}^n$, it is easy to conclude from (8) and the last inequality that

$$\langle \operatorname{grad} f(\bar{p}), p\rangle \geq 0, \qquad \forall\, p \in C.$$

Thus, using the definition in (34) we obtain $-\operatorname{grad} f(\bar{p}) \in N_C(\bar{p})$, which implies that $\bar{p}$ is solution of the variational inequality in (41) and the first statement is proved.

For proving the second statement, it is sufficient to prove that if $\bar{p} \in C$ is a solution of (41) then $\bar{p} \in C$ is also solution of (40). Assume that $\bar{p} \in C$ is solution of (41), that is, $-\operatorname{grad} f(\bar{p}) \in N_C(\bar{p})$. Moreover, assume that f is spherically convex. Since $-\operatorname{grad} f(\bar{p}) \in N_C(\bar{p})$, we conclude from (34) and (10) that

$$\left\langle \left[I - \bar{p}\bar{p}^T\right] Df(\bar{p}), p\right\rangle = \langle \operatorname{grad} f(\bar{p}), p\rangle \geq 0, \qquad \forall\, p \in C.$$

Using the last inequality we obtain, after some simple algebraic manipulation, that

$$\left\langle Df(\bar{p}), \left[I - \bar{p}\bar{p}^T\right] p\right\rangle = \left\langle \left[I - \bar{p}\bar{p}^T\right] Df(\bar{p}), p\right\rangle \geq 0, \qquad \forall\, p \in C.$$

On the other hand, since f is a spherically convex function, we have, using Proposition 7,

$$f(p) \geq f(\bar{p}) + \frac{\arccos\langle \bar{p}, p\rangle}{\sqrt{1-\langle \bar{p}, p\rangle^2}} \left\langle Df(\bar{p})\,,\, [I - \bar{p}\bar{p}^T]p \right\rangle, \qquad \forall\, p \in C.$$

Therefore, combining the last two inequalities, we obtain that $f(p) \geq f(\bar{p})$ for all $p \in C$, concluding the proof. □

Corollary 5. Let $C \subset \mathbb{S}^n$ be a closed spherically convex set, $p \in \mathbb{S}^n$ and $\bar{p} \in C$ such that $\langle p, \bar{p}\rangle > 0$. Then $P_C(p) = \bar{p}$ if and only if $\exp^{-1}_{\bar{p}} p \in N_C(\bar{p})$.

Proof. Let $p \in \mathbb{S}^n$ and $\bar{p} \in C$ such that $\langle p, \bar{p}\rangle > 0$. It follows from Example 2 that one half of the square of the intrinsic distance from $p \in \mathbb{S}^n$, i.e., the function $\rho_p : \mathbb{S}^n \backslash \{p, -p\} \to \mathbb{R}$ defined as

$$\rho_p(q) := \frac{1}{2} d_p^2(q) = \frac{1}{2} \arccos^2\langle q, p\rangle,$$

is differentiable and strictly spherically convex in the hemisphere

$$S_p^n := \{q \in \mathbb{S}^n : d_p(q) < \pi/2\} = \{q \in \mathbb{S}^n : \langle q, p\rangle > 0\},$$

which has the point p as a its pole. As $\bar{p} \in C \cap S_p^n$, the definition of the projection in (27) implies that $P_C(p) = \bar{p}$ if and only if $\bar{p} = \operatorname{argmin}\{\rho_p(q) : q \in C \cap S_p^n\} = \operatorname{argmin}\{\rho_p(q) : q \in C\}$. Hence, using Proposition 15, and equations (8) and (29), we conclude that $P_C(p) = \bar{p}$ if and only if

$$\exp^{-1}_{\bar{p}} p = -\operatorname{grad} \rho_p(\bar{p}) \in N_C(\bar{p}),$$

which is the desired result. □

7 Monotone vector fields on the sphere

In this section we define the monotonicity of a vector field on the sphere. In particular, we show that the gradient vector field of a differentiable spherically convex function is monotone.

Definition 4. Let $C \subset \mathbb{S}^n$ be a spherically convex set. The vector field $C \ni p \mapsto X(p) \in T_p\mathbb{S}^n$ on the sphere is said to be *spherically monotone* if the following inequality holds:

$$\langle X(p), q\rangle + \langle X(q), p\rangle \leq 0, \qquad \forall\ p,\ q \in C.$$

Remark 4. Since $X(p) \in T_p\mathbb{S}^n$ and $X(q) \in T_q\mathbb{S}^n$, the inequality in the above definition becomes

$$\langle X(p), [I - pp^T]q\rangle + \langle X(q), [I - qq^T]p\rangle \leq 0, \qquad \forall\ p,\ q \in C.$$

Since $\arccos\langle p,q\rangle/\sqrt{1-\langle p,q\rangle^2}\geq 0$, it follows from (8) that the previous inequality is equivalent to the following one:

$$\langle X(p), \exp_p^{-1}q\rangle + \langle X(q), \exp_q^{-1}p\rangle \leq 0, \qquad \forall\ p,\ q\in C.$$

Now, the exponential mapping in the Euclidean space is $\exp_p v = p+v$. Hence, its inverse is $\exp_p^{-1}q = q-p$. So, the above inequality in the Euclidean space is equivalent to

$$\langle X(p)-X(q), p-q\rangle \geq 0, \qquad \forall\ p,\ q\in C,$$

which is the usual expression defining a monotone operator. Moreover, it is easy to see that the inequality in Definition 4 is equivalent to

$$\langle X(p)-X(q), p-q\rangle \geq 0, \qquad \forall\ p,\ q\in C,$$

where C now is a spherically convex set in $\mathbb{S}^n$.

Proposition 16. Let $C\subset\mathbb{S}^n$ be a spherically convex set and $f : C\to\mathbb{R}$ be a differentiable function. Then, f is spherically convex if and only if its gradient vector field $C\ni p\mapsto \operatorname{grad} f(p)\in T_p\mathbb{S}^n$ is spherically monotone, that is,

$$\langle \operatorname{grad} f(p), q\rangle + \langle \operatorname{grad} f(q), p\rangle \leq 0, \qquad \forall\ p,\ q\in C.$$

Proof. The result follows from the equivalence of the inequalities of Propositions 8 and 16 (similarly to the ideas of Remark 4). □

8 Final remarks

This paper is a continuation of [7], where we studied some basic intrinsic properties of the spherically convex functions and we only slightly touched the optimization theory in this new context. We expect that the results of this paper become a first step towards a more general theory, including algorithms for solving spherically convex optimization problems. We forsee further progress in this topic in the nearby future.

References

[1] G. Dahl, J. M. Leinaas, J. Myrheim, and E. Ovrum. A tensor product matrix approximation problem in quantum physics. *Linear Algebra Appl.*, 420(2-3):711–725, 2007.

[2] P. Das, N. R. Chakraborti, and P. K. Chaudhuri. Spherical minimax location problem. *Comput. Optim. Appl.*, 18(3):311–326, 2001.

[3] J. E. Dennis, Jr. and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[4] M. P. do Carmo. *Differential geometry of curves and surfaces*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1976. Translated from the Portuguese.

[5] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[6] Z. Drezner and G. O. Wesolowsky. Minimax and maximin facility location problems on a sphere. *Naval Res. Logist. Quart.*, 30(2):305–312, 1983.

[7] O. P. Ferreira, A. N. Iusem, and S. Z. Nemeth. Projections onto convex sets on the sphere. *Journal of Global Optimization*, 57:663–676 (2013).

[8] R. Ferreira, J. Xavier, J. Costeira, and V. Barroso. Newton algorithms for riemannian distance related problems on connected locally symmetric manifolds. *Thechnical Repor: Institute for Systems and Robotics (ISR), Signal and Image Processing Group (SPIG), Instituto Superior Tecnico (IST)*, 2008.

[9] D. Han, H. H. Dai, and L. Qi. Conditions for strong ellipticity of anisotropic elastic materials. *J. Elasticity*, 97(1):1–13, 2009.

[10] S. He, Z. Li, and S. Zhang. Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Math. Program.*, 125(2, Ser. B):353–383, 2010.

[11] A. Iusem and A. Seeger. On pairs of vectors achieving the maximal angle of a convex cone. *Math. Program.*, 104(2-3, Ser. B):501–523, 2005.

[12] A. Iusem and A. Seeger. Searching for critical angles in a convex cone. *Math. Program.*, 120(1, Ser. B):3–25, 2009.

[13] I. N. Katz and L. Cooper. Optimal location on a sphere. *Comput. Math. Appl.*, 6(2):175–196, 1980.

[14] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, New York, 2009.

[15] C. Li and J.-C. Yao. Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the solution set and the proximal point algorithm. *SIAM J. Control Optim.*, 50(4):2486–2514, 2012.

[16] S.-L. Li, C. Li, Y.-C. Liou, and J.-C. Yao. Existence of solutions for variational inequalities on Riemannian manifolds. *Nonlinear Anal.*, 71(11):5695–5706, 2009.

[17] C. Ling, J. Nie, L. Qi, and Y. Ye. Biquadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM J. Optim.*, 20(3):1286–1310, 2009.

[18] L. Qi. Eigenvalues of a real supersymmetric tensor. *J. Symbolic Comput.*, 40(6):1302–1324, 2005.

[19] L. Qi and K. L. Teo. Multivariate polynomial minimization and its application in signal processing. *J. Global Optim.*, 26(4):419–433, 2003.

[20] L. Qi, F. Wang, and Y. Wang. Z-eigenvalue methods for a global polynomial optimization problem. *Math. Program.*, 118(2, Ser. A):301–316, 2009.

[21] B. Reznick. Some concrete aspects of Hilbert's 17th Problem. In *Real algebraic geometry and ordered structures (Baton Rouge, LA, 1996)*, volume 253 of *Contemp. Math.*, pages 251–272. Amer. Math. Soc., Providence, RI, 2000.

[22] T. Sakai. *Riemannian geometry*, volume 149 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI, 1996. Translated from the 1992 Japanese original by the author.

[23] S. T. Smith. Optimization techniques on Riemannian manifolds. In *Hamiltonian and gradient flows, algorithms and control*, volume 3 of *Fields Inst. Commun.*, pages 113–136. Amer. Math. Soc., Providence, RI, 1994.

[24] A. M.-C. So. Deterministic approximation algorithms for sphere constrained homogeneous polynomial optimization problems. *Math. Program.*, 129(2, Ser. B):357–382, 2011.

[25] S. Weiland and F. van Belzen. Singular value decompositions and low rank approximations of tensors. *IEEE Trans. Signal Process.*, 58(3, part 1):1171–1182, 2010.

[26] G.-L. Xue. A globally convergent algorithm for facility location on a sphere. *Comput. Math. Appl.*, 27(6):37–50, 1994.

[27] G. L. Xue. On an open problem in spherical facility location. *Numer. Algorithms*, 9(1-2):1–12, 1995.

[28] L. Zhang. On the convergence of a modified algorithm for the spherical facility location problem. *Oper. Res. Lett.*, 31(2):161–166, 2003.

[29] X. Zhang, C. Ling, and L. Qi. The best rank-1 approximation of a symmetric tensor and related spherical optimization problems. *SIAM J. Matrix Anal. Appl.*, 33(3):806–821, 2012.