

Height functions on compact symmetric spaces

E. Macías-Virgós & M. J. Pereira-Sáez

Monatshefte für Mathematik

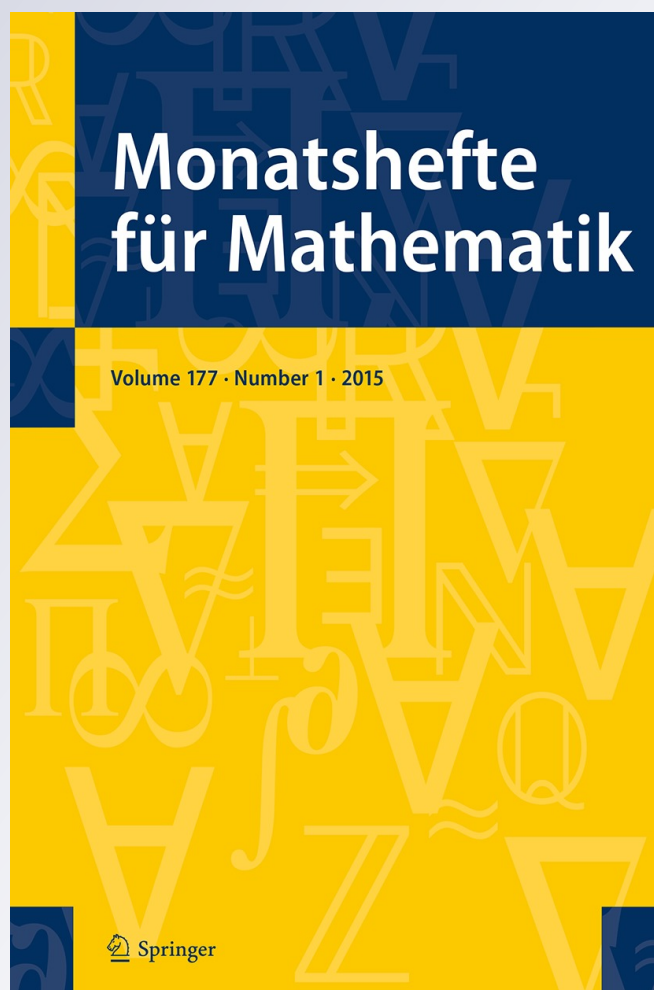
ISSN 0026-9255

Volume 177

Number 1

Monatsh Math (2015) 177:119-140

DOI 10.1007/s00605-014-0696-1



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Wien. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Height functions on compact symmetric spaces

E. Macías-Virgós · M. J. Pereira-Sáez

Received: 7 January 2014 / Accepted: 2 October 2014 / Published online: 24 October 2014
 © Springer-Verlag Wien 2014

Abstract We consider height functions on symmetric spaces $M \cong G/K$ embedded in the associated matrix Lie group G . In particular we study the relationship between the critical sets of the height function on G and its restriction to M . Also we prove that the gradient flow on M can be integrated by means of a generalized Cayley transform; this allows to obtain explicit local charts for the critical submanifolds. Finally, we discuss how to reduce the study of any height function to the case where the ground hyperplane is orthogonal to a real diagonal matrix. This result requires to prove the existence of a polar decomposition adapted to the automorphism defining M . Detailed examples are given.

Keywords Symmetric space · Lie group · Height function · Morse-Bott theory · Critical point · Gradient flow · Cayley transform

Mathematics Subject Classification 58E05 · 53C35

Contents

1 Introduction	120
2 Preliminaries	122

Communicated by A. Cap.

E. Macías-Virgós (✉)
 Facultade de Matemáticas, Universidade de Santiago de Compostela, Santiago de Compostela, Spain
 e-mail: quique.macias@usc.es

M. J. Pereira-Sáez
 Facultade de Economía e Empresa, Universidade da Coruña, A Coruña, Spain
 e-mail: maria.jose.pereira@udc.es

2.1 Lie groups of orthogonal type	122
2.2 Compact symmetric spaces	123
3 Height functions on symmetric spaces	123
3.1 Gradient	124
3.2 Hessian	125
3.3 Height functions on the Lie group	126
4 Cayley transform	127
4.1 Generalized Cayley transform	127
4.2 The Cayley map in symmetric spaces	128
5 Linearization of the gradient	129
5.1 Gradient equation	129
5.2 Local structure of the critical set	131
5.3 Relationship between gradient flows	132
6 Polar decompositions and critical set	132
6.1 Singular value decomposition and polar form	132
6.2 Critical points	133
6.3 Polar decompositions in a symmetric space	134
7 Height functions associated to real diagonal matrices	135
7.1 Reduction to the diagonal case	136
7.2 Description of the critical set	136
7.3 Final example	138
References	139

1 Introduction

A common method in Morse theory is to embed a manifold in a suitable Euclidean space and examine the critical set of some height function. In particular, for symmetric spaces it is possible to construct isometric embeddings in a unified manner [12, 14]. Frankel [4] applied Morse theory to the classical Lie groups and Grassmannians—described in terms of matrices—, and gave a complete description of the critical submanifolds, by using the trace functional as Morse function. Nicolaescu [19] also studied trace functions on Grassmannians. Ramanujam [20–22], Takeuchi [23], Duan [2] and other authors investigated Morse functions on symmetric spaces by choosing a suitable length function. Dynnikov and Veselov [24], considered height functions on the natural embeddings of classical Lie groups and certain symmetric spaces.

The generic formula for such a function is $h_X(A) = \Re \operatorname{Tr}(XA)$, where $\Re \operatorname{Tr}$ is the real part of the trace, and X^* is perpendicular to the hyperplane taken as ground level. A crucial point in order to describe the critical set $\Sigma(h_X^G)$ in the Lie group G is to reduce the matrix X to a diagonal matrix D with non-negative real entries by taking some singular value decomposition UDV^* . In fact, it can be proven that h_X^G is a Morse function (isolated non-degenerate critical points) if and only if the singular values of X are positive and pairwise different. When X^*X has repeated eigenvalues, the function is of Morse-Bott type (that is, there are critical submanifolds instead of isolated points, with non-degenerate Hessian in the transverse directions). An explicit description of the critical set of an arbitrary height function on the Lie group G appears in the authors' paper [6].

For symmetric spaces, however, there is no systematic characterization of Morse functions. In fact, the results for Lie groups led in the past to study only real diagonal

matrices. As a consequence, the behaviour of the height functions on symmetric spaces might seem much more regular than what it actually is. For instance, when $M \cong G/K$ is embedded in the associated Lie group G , Ramanujam [21] states that “the critical submanifolds of G/K are shown to be the intersection of the space G/K and the critical submanifolds of G ” while Dynnikov and Veselov [24] write that “symmetric spaces [...] are invariant by the gradient flow of the height function on the corresponding Lie groups” and that “the restricted flow coincides with the gradient flow of the [restricted] function”. As we shall see, this kind of results only holds in particular—although important—cases, but is no longer true for an arbitrary height function on a symmetric space.

The situation is much more interesting indeed. We shall show that it is possible to compute the critical set of the restricted function h_X^M in terms of the matrix $\widehat{X} := X^* + \sigma(X)$, where σ is the involutive automorphism of G defining the symmetric space. Namely, we obtain the general formula

$$\Sigma \left(h_X^G \right) \cap M \subset \Sigma \left(h_X^M \right) = \Sigma \left(h_{\sigma(\widehat{X})}^G \right) \cap M. \tag{1}$$

Only under additional hypothesis the left inclusion will be an equality. Similarly, the gradient flow in G preserves the symmetric space when $\sigma(X) = X^*$, but not in the general case.

Anyway, we will be able to prove a result of reduction of an arbitrary height function h_X^M to the diagonal case. In order to obtain that result we shall need to show the existence of singular value decompositions which are adapted to the symmetric space. This problem has been widely studied [11, 15]. But in contrast to the already known results, which are of local nature, our decomposition theorem is global and relies in the relationship between the critical points and the polar form of the matrix \widehat{X} . This result is of interest by itself.

Gradient flows related to the minimization of a distance function have been widely studied too. Unlike most Riemannian manifolds, the gradient flows of the height functions on the classical Lie groups can be integrated explicitly. The key idea, found independently by the authors [6] and other authors ([25], cited in [24]) is that this can be achieved by means of the so-called Cayley transform. In this paper we prove that, after a suitable generalization, it is also possible to do so in symmetric spaces. The formulae obtained contain all previous results as particular cases, and may be of interest for applications in optimization and control theory [8, 9, 26].

Finally, we prove that the generalized Cayley transform allows to give explicit local charts for the critical submanifolds of height functions defined on a symmetric space, a result which is completely new. Another consequence is a method to cover the symmetric space by a finite family of contractible open sets, which is suitable for the computation of the Lusternik-Schnirelmann category. In fact, Morse-Bott functions can be effectively used to determine the L.-S. category of some Lie groups and homogeneous spaces [6, 13, 17].

The contents of the paper are as follows. In Sect. 2 we fix some notations and basic definitions. We shall consider the symmetric space $M \cong G/K$ associated to the involutive automorphism σ of the compact Lie group G , where K is the fixed

point subgroup. Then M can be embedded in a natural manner in G by means of the so-called Cartan embedding.

In Sect. 3 we consider an arbitrary height function on M , which is defined by the formula $h_X^M(A) = \Re \operatorname{Tr}(XA)$, where the matrix X^* is orthogonal to the ground hyperplane. We then compute the gradient and the Hessian. In particular the point $A \in M$ is a critical point of h_X^M if and only if $\widehat{X} = A\sigma(\widehat{X})A$, where $\widehat{X} := X^* + \sigma(X)$. As a particular case we compute the critical points of the function h_X^G on the Lie group G ; this time the critical equation is $X^* = AXA$. Thus we obtain formula (1) and we show with an example that the first inclusion may be strict.

In Sect. 4 we introduce a generalization of the Cayley transform between the Lie group G and its Lie algebra \mathfrak{g} . More precisely, each point $A \in G$ has a neighbourhood $\Omega_G(A)$ which is mapped diffeomorphically onto the tangent space $T_{A^*}G$. We then verify that this diffeomorphism passes naturally to the symmetric space $G/K \cong M \subset G$, thus providing a covering of M by contractible open sets $\Omega_M(A) = \Omega_G(A) \cap M$.

In Sect. 5 we prove that the Cayley transform allows to linearize the gradient flow and to solve explicitly the gradient equation for symmetric spaces. This has important consequences for the knowledge of the critical set of h_X^M , since the Cayley transform will define a local chart $\Omega_M(A) \cap \Sigma(h_X^M)$ which is modeled on the kernel of the Hessian. Finally we compare the gradient flows of the function h_X^G and its restriction h_X^M to the symmetric space M .

In Sect. 6 we prove—under some hypothesis—that the matrix \widehat{X} admits a singular value decomposition and a polar form which are adapted to the automorphism σ defining the symmetric space. More explicitly, since $\sigma(\widehat{X}) = \widehat{X}^*$ we prove that there is a polar decomposition $\widehat{X} = S\Omega$ such that the orthogonal part verifies $\sigma(\Omega) = \Omega^*$ and the hermitian part verifies $\sigma(S) = \Omega^*S\Omega$. Analogously there is a SVD decomposition $\widehat{X} = UDV^*$ such that the matrix $\Theta = U^*\sigma(V)$ verifies $\sigma(\Theta) = \Theta^*$. Then the critical set of h_X^M is diffeomorphic to the critical set of $h_D^{M'}$, where D is a real diagonal matrix and M' is a new symmetric space associated to the automorphism $\sigma'(X) = \Theta\sigma(X)\Theta^*$. This simplifies greatly the study of height functions, as we show by a detailed final example.

2 Preliminaries

We begin by recalling several basic definitions and results.

2.1 Lie groups of orthogonal type

Let \mathbb{K} be one of the algebras \mathbb{R} (reals), \mathbb{C} (complex) or \mathbb{H} (quaternions). For the matrix $A \in \mathbb{K}^{n \times n}$ we denote by A^* its conjugate transpose. Let $O(n, \mathbb{K}) = \{A \in \mathbb{K}^{n \times n} : AA^* = I\}$ be the compact Lie group of orthogonal (resp. unitary, symplectic) matrices and let G be the connected component of the identity. Thus, G will be either $SO(n)$, $U(n)$ or $Sp(n)$. In the real case we shall consider sometimes the whole Lie group $O(n)$. The Lie algebra of G is formed by the skew-symmetric (resp. skew-Hermitian) matrices, $\mathfrak{g} = \{X \in \mathbb{K}^{n \times n} : X + X^* = 0\}$.

The Riemannian metric induced on $G \subset \mathbb{K}^{n \times n}$ by the usual inner product $\langle X, Y \rangle = \Re \operatorname{Tr}(X^*Y)$ is bi-invariant ($\Re \operatorname{Tr}$ denotes the real part of the trace).

2.2 Compact symmetric spaces

Let $\sigma : G \rightarrow G$ be an involutive automorphism and let $K = \{B \in G : \sigma(B) = B\}$ be the closed Lie subgroup of fixed points.

Assumption In this paper we shall assume that the automorphism σ is the restriction of an involutive automorphism $\sigma : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n}$ of unital algebras. We shall also assume that $\sigma(X^*) = \sigma(X)^*$ for all $X \in \mathbb{K}^{n \times n}$. These conditions are not too restrictive; for instance, all the compact irreducible Riemannian symmetric spaces in Cartan's classification fulfill them (see [7, p. 518]).

The Lie algebra of K is $\mathfrak{k} = \{X \in \mathfrak{g} : \sigma(X) = X\}$ and the tangent space $T_{[I]}G/K$ is isomorphic to $\mathfrak{m} = \{X \in \mathfrak{g} : \sigma(X) = -X\}$. Since $\operatorname{Ad}(k)(\mathfrak{m}) = \mathfrak{m}$ for all $k \in K$, there is an invariant Riemannian metric on G/K which has null torsion and parallel curvature [3, 7, 27]. The homogeneous space G/K is called a *globally symmetric compact space*.

Let $\gamma : G/K \hookrightarrow G$ be the embedding given by $\gamma([B]) = B\sigma(B)^{-1}$. It is an isometry (up to the constant 2).

Proposition 1 ([3, p. 185]) *Assume that G/K is connected. Then the image $M = \gamma(G/K)$ of γ is the connected component N_I of the identity of the submanifold*

$$N = \left\{ B \in G : \sigma(B) = B^{-1} \right\}.$$

See [12] for several examples. The manifold M will be called the *Cartan model* of the symmetric space G/K . The isometric action of G induced by γ on M is given by $l_B^M(A) = BA\sigma(B)^{-1}$, for $B \in G, A \in M$.

Example 1 The Lie group G itself is a symmetric space defined by the automorphism $\sigma : G \times G \rightarrow G \times G$ with $\sigma(B_1, B_2) = (B_2, B_1)$. The fixed point set is the diagonal Δ . The diffeomorphism $G \cong (G \times G)/\Delta$ is given by $B \cong [(B, I)]$. The submanifold $N \subset G \times G$ is formed by all pairs (B, B^{-1}) .

Proposition 2 *For any point $A \in M$, the tangent space is*

$$T_A M = \{Y \in \mathbb{K}^{n \times n} : YA^* + AY^* = 0, \sigma(Y) = -Y^*\}.$$

Proof Since $A = \gamma([B])$ for some $B \in G$, it is $T_A M = (l_B^M)_{*I}(T_I M)$, so $Y \in T_A M$ if and only if $(l_B^M)^{-1}(Y) = B^*Y\sigma(B) \in T_I M$ because l_B^M is a linear action. The result follows from the definition of \mathfrak{m} .

3 Height functions on symmetric spaces

First of all we shall consider an arbitrary height function on the vector space $\mathbb{K}^{n \times n}$. Then we shall study how it behaves when restricted to the Cartan model M of a

symmetric space. Finally we shall obtain as a particular case the corresponding results for the Lie group G .

The height function $h_X : \mathbb{K}^{n \times n} \rightarrow \mathbb{R}$ with respect to an hyperplane perpendicular to $X^* \in \mathbb{K}^{n \times n}$, where $X \neq 0$, is given, up to a constant, by

$$h_X(Y) = \langle X^*, Y \rangle = \Re \operatorname{Tr}(XY).$$

3.1 Gradient

Since h_X is \mathbb{R} -linear, its gradient at $Y \in \mathbb{K}^{n \times n}$ is $(\operatorname{grad} h_X)_Y = X^*$. Let $h_X^M : M \rightarrow \mathbb{R}$ be the restriction of h_X to the Cartan model $M \subset G \subset \mathbb{K}^{n \times n}$ of the symmetric space G/K .

We denote $\widehat{X} := X^* + \sigma(X)$. Notice that $\sigma(\widehat{X}) = \widehat{X}^*$.

Lemma 1 *The projection of $Z \in \mathbb{K}^{n \times n}$ onto T_1G is its skew-symmetric part, $\frac{1}{2}(Z - Z^*)$. The projection of the latter onto $\mathfrak{m} = T_1M$ is its skew-invariant part, that is, $\frac{1}{4}[(Z - \sigma(Z)) - (Z^* - \sigma(Z)^*)]$.*

Proposition 3 *The gradient of h_X^M at any point $A \in M$ is the projection of $\operatorname{grad} h_X$ onto $T_A M$, that is,*

$$\left(\operatorname{grad} h_X^M\right)_A = \frac{1}{4}(\widehat{X} - A\sigma(\widehat{X})A). \tag{2}$$

Proof Using Lemma 1,

$$\left(\operatorname{grad} h_X^M\right)_I = \frac{1}{4}[(X^* - \sigma(X)^*) - (X - \sigma(X))].$$

Let $A = B\sigma(B)^{-1} = l_B^M(I)$. Since the translations l_B^M are isometries we have

$$\left(\operatorname{grad} h_X^M\right)_A = l_B^M \left(\operatorname{grad}(h_X^M \circ l_B^M)_I\right) = l_B^M \left(\left(\operatorname{grad} h_{\sigma(B)^*XB}\right)_I\right).$$

Then we apply the action to

$$\left((B^*X^*\sigma(B) - \sigma(B)^*\sigma(X^*)B) - (\sigma(B^*)XB - B^*\sigma(X)\sigma(B))\right)$$

and we obtain

$$(X^* - B\sigma(B)^*\sigma(X)^*B\sigma(B)^* - B\sigma(B)^*XB\sigma(B)^* + \sigma(X)) = (\widehat{X} - A\widehat{X}^*A).$$

Corollary 1 *$A \in M$ is a critical point of h_X^M if and only if $\widehat{X} = A\sigma(\widehat{X})A$, where $\widehat{X} := X^* + \sigma(X)$.*

Remark 1 Instead of height, one can consider the distance to X^* . Since $|A - X^*|^2 = a h_X^M(A) + b$, with $a, b \in \mathbb{R}$, both functions have the same critical points in M .

Geometrically, these are the points where the line $\overrightarrow{AX^*}$ is perpendicular to $T_A M$ [18, p. 35].

3.2 Hessian

Let $A \in M$. Since M and G are submanifolds of $\mathbb{K}^{n \times n}$, in order to compute the Hessian we have to extend the vector field $\text{grad } h_X^M$ to all matrices [5, p. 72], taking for instance

$$(\widetilde{\text{grad } h_X^M})_B = \frac{1}{4} (\widehat{X} - B\sigma(\widehat{X})B).$$

The covariant derivative ∇ in $\mathbb{K}^{n \times n}$ is the usual derivative. Then for $W \in T_A M$ the Hessian $(\nabla_W^M \text{grad})_A$ is the projection of

$$\begin{aligned} (\nabla_W \widetilde{\text{grad}})_A &= \lim_{t \rightarrow 0} \frac{1}{t} [(\widetilde{\text{grad}})_{A+tW} - (\widetilde{\text{grad}})_A] \\ &= \lim_{t \rightarrow 0} \frac{1}{4t} [\widehat{X} - (A + tW)\sigma(\widehat{X})(A + tW) - (\widehat{X} - A\sigma(\widehat{X})A)] \\ &= -\frac{1}{4} (A\sigma(\widehat{X})W + W\sigma(\widehat{X})A), \end{aligned}$$

which belongs to $T_A M$. Then we have proved:

Proposition 4 *The Hessian $H(h_X^M)_A : T_A M \rightarrow T_A M$ of the height function $h_X^M : M \rightarrow \mathbb{R}$ is given by*

$$H(h_X^M)_A(W) = -\frac{1}{4} (A\sigma(\widehat{X})W + W\sigma(\widehat{X})A).$$

An easy computation shows that: (i) A is a critical point of h_X^M if and only if the matrix \widehat{X}^*A is Hermitian; (ii) the matrix W belongs to the tangent space $T_A M$ if and only if WA^* is skew-Hermitian and $\sigma(W) = W^*$; (iii) W is in the kernel of the Hessian if in addition the matrix \widehat{X}^*W is Hermitian.

Example 2 Let us consider the complex Grassmannian $U(2)/(U(1) \times U(1))$ defined by the automorphism $\sigma(A) = I_{1,1}AI_{1,1}$, where $I_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The Cartan model is the sphere $S^2 \subset U(2) \cong S^3 \times S^1$ formed by the matrices $\begin{pmatrix} s & -\bar{z} \\ z & s \end{pmatrix}$ where $(s, z) \in \mathbb{R} \times \mathbb{C}$ verifies $s^2 + |z|^2 = 1$.

Let us take on M the function h_X^M with $X = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Then $\widehat{X} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ and the critical points are the two poles $\pm I$. The tangent space $T_{\varepsilon I} M$ is formed by the matrices $W = \begin{pmatrix} 0 & z \\ -\bar{z} & 0 \end{pmatrix}$ with $z \in \mathbb{C}$. The Hessian is $(Hh_X^M)_{\varepsilon I}(W) = (-\varepsilon/2)W$, so h_X^M is a Morse function on M . On the other hand, as we will see (cf. Sect. 2.3, also [6]), the critical set of h_X^G on the Lie group $U(2)$ is formed by the two circles $U(1) \times \{\pm 1\}$ of matrices $\begin{pmatrix} \alpha & 0 \\ 0 & \pm 1 \end{pmatrix}$ with $\alpha \in \mathbb{C}, |\alpha| = 1$.

3.3 Height functions on the Lie group

In [6] the authors studied the height functions on the Lie group G of orthogonal type. Let us sketch how to recover those results by considering G as the symmetric space of Example 1.

Proposition 5 *The gradient of $h_X^G: G \rightarrow \mathbb{R}$ at $A \in G$ is $(\text{grad } h_X^G)_A = \frac{1}{2}(X^* - AXA)$.*

Proof First, the Cartan embedding $\gamma: G \rightarrow G \times G$ can be extended to a map $\gamma: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times n} = \mathbb{K}^{2n \times 2n}$ by putting $\gamma(X) = \begin{pmatrix} X & 0 \\ 0 & X^* \end{pmatrix}$. This \mathbb{R} -linear map is an isometry (up to the constant $1/2$). Let $M = \gamma(G)$. According to Proposition 3, the gradient of $h_{\gamma(X)}^M$ at $\gamma(A) \in M$ is

$$\frac{1}{2}(\widehat{\gamma(X)} - \gamma(A)\sigma(\widehat{\gamma(X)})\gamma(A)) = \frac{1}{2}\gamma(X^* - AXA).$$

On the other hand,

$$\left(\text{grad } h_X^G\right)_A = 2\gamma_{*A}^{-1} \left(\text{grad}(h_X^G \circ \gamma^{-1})\right)_{\gamma(A)} = \gamma^{-1} \left(\text{grad } h_{\gamma(X)}^M\right)_{\gamma(A)}$$

and the result follows. □

Remark 2 Compare formulae in (2) and (5). A similar computation is valid *mutatis mutandi* for the Hessian [6, p. 329], the gradient flow [6, p. 330] and the local structure of the critical set [6, p. 331] in the group G . In all formulae it is enough to substitute $2X^*$ for \widehat{X} .

From Propositions 3 and 5 we obtain:

Corollary 2 *Let $M \subset G$ be the Cartan model of the symmetric space G/K . Then the critical set in M of the height function h_X^M is $\Sigma(h_X^M) = \Sigma(h_{\sigma(\widehat{X})}^G) \cap M$.*

Corollary 3 $\Sigma(h_X^G) \cap M \subset \Sigma(h_X^M)$.

Proof This is an obvious consequence of Remark 1. Alternatively, notice that $X^* = AXA$ implies $\widehat{X} = A\sigma(\widehat{X})A$ when $\sigma(A) = A^*$. □

According to Ramanujam [21, p. 219], when $X = I$ the critical points of the height function h_X^M are just the points of $\Sigma(h_X^G)$ that belong to M . The same result is true when X is a real diagonal matrix for the symmetric spaces studied by Duan [2] and Dynnikov-Veselov [24].

Corollary 4 *If $\sigma(X) = X^*$, then the critical points of the height function h_X^M on the symmetric space M verify that $\Sigma(h_X^M) = \Sigma(h_X^G) \cap M$.*

However, in the generic case the preceding result no longer holds, as the following example shows.

Example 3 Let us consider the symmetric space $Sp(1)/U(1)$. It is defined by the automorphism $\sigma(X) = -\mathbf{i}X\mathbf{i}$. The Cartan model M is the sphere $S^2 \subset Sp(1) = S^3$ formed by the unit quaternions $q = s + \mathbf{j}z$ such that $s \in \mathbb{R}, z \in \mathbb{C}$, with $s^2 + |z|^2 = 1$. Notice that q has a null \mathbf{i} -coordinate.

Now we consider the height function h_X with $X = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

(i) On the group $G = Sp(1)$, by Proposition 5 we have that the critical points of h_X^G are the elements $q = t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \in Sp(1)$ such that $-(\mathbf{i} + \mathbf{j} + \mathbf{k})\bar{q} = q(\mathbf{i} + \mathbf{j} + \mathbf{k})$. From this equation we deduce that $t = 0$ and $x = y = z$. But since q is unitary, we have that the only critical points in $G = Sp(1)$ are $\Sigma(h_X^G) = \{\pm \frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})\}$. These two points are not in M because they have a non-null \mathbf{i} -coordinate.

(ii) Nevertheless, there are points of M that are critical points for the height function restricted to the Cartan model $M \subset Sp(1)$ of $Sp(1)/U(1)$. This time, by Proposition 3, the condition for a point $q \in M$ to be critical for h_X^M is $\widehat{X} = q\sigma(\widehat{X})q$, where $\widehat{X} = X^* + \sigma(X) = -2(\mathbf{j} + \mathbf{k})$. So, from the condition $-2\bar{q}(\mathbf{j} + \mathbf{k}) = 2(\mathbf{j} + \mathbf{k})q$ we obtain that $\Sigma(h_X^M) = \{\pm \frac{1}{\sqrt{2}}(\mathbf{j} + \mathbf{k})\}$.

4 Cayley transform

In this Section we integrate explicitly the gradient flow and give local charts for the critical submanifolds.

4.1 Generalized Cayley transform

The following *generalized Cayley map* was defined by the authors in [6]. Let $A \in G$, that is, $AA^* = I$. We consider the open set of matrices

$$\Omega(A) = \{X \in \mathbb{K}^{n \times n} : A + X \text{ is invertible}\}.$$

Definition 1 The *Cayley transform centered at A* is the map $c_A : \Omega(A) \rightarrow \Omega(A^*)$ defined by

$$c_A(X) = (I - A^*X)(A + X)^{-1} = (A + X)^{-1}(I - XA^*).$$

It is a diffeomorphism, with $c_A^{-1} = c_{A^*}$.

Remark 3 When $A = I$ one obtains the classical Cayley map $c(X) = (I - X)(I + X)^{-1}$ which is defined for the matrices X that do not have -1 as an eigenvalue. Notice that $c_A(X) = c(A^*X)A^*$.

Theorem 1 ([6, p. 328]) *Let $\Omega_G(A) := \Omega(A) \cap G$. The map c_A induces a diffeomorphism $\Omega_G(A) \cong T_{A^*}G$. In particular, the open subspace $\Omega_G(A) \subset G$ is contractible.*

Proof We sketch the proof for the sake of completeness. We shall use the properties of c_A that appear in [6, p. 327].

We first prove that $c_A(\Omega_G(A)) \subset T_{A^*}G$. Let $B \in \Omega_G(A)$, then $B^* = B^{-1}$ and

$$c_A(B)^* = c_{A^*}(B^*) = c_{A^*}(B^{-1}) = -Ac_A(B)A.$$

It follows that the matrix $c_A(B)A$ is skew-Hermitian, that is, $c_A(B) \in T_{A^*}G$.

Now we prove that $T_{A^*}G \subset \Omega(A^*)$. If there exists $Y \in T_{A^*}G$ such that $Y + A^*$ is not invertible then $\exists v \neq 0$ such that $(Y + A^*)v = 0$, hence $A^*v = -Yv$ and $-v = AYv$. That means that -1 is an eigenvalue of AY , which is impossible because the eigenvalues of a skew-Hermitian matrix must have null real part.

Finally, we prove that $c_{A^*}(T_{A^*}G) \subset \Omega_G(A)$. Let $Y \in T_{A^*}G \subset \Omega(A^*)$. Then $c_{A^*}(Y) \in G$. Indeed, $AY \in T_I G$ is skew-Hermitian, hence $I - Y^*A^*$ is invertible, so $c_{A^*}(Y) = c_A(Y^*)^* = [(A + Y^*)^{-1}(I - Y^*A^*)]^*$ is invertible. Moreover, as $-A^*Y^*A^* = Y$, it happens that

$$Y^* = -AYA = -Ac_A(c_{A^*}(Y))A = c_{A^*}(c_{A^*}(Y)^{-1})$$

while

$$Y^* = c_A(c_{A^*}(Y))^* = c_{A^*}(c_{A^*}(Y)^*).$$

Since c_{A^*} is injective it follows that $c_{A^*}(Y)^* = (c_{A^*}(Y))^{-1}$. □

4.2 The Cayley map in symmetric spaces

We shall verify in the next paragraphs that the preceding properties of the Cayley transform in the Lie group G are naturally inherited by the Cartan model $M \subset G$ of the symmetric space G/K .

Lemma 2 *If $A \in G$ then $c_{\sigma(A)} \circ \sigma = \sigma \circ c_A$ on $\Omega(A)$, or, equivalently, $\sigma \circ c_{\sigma(A)} = c_A \circ \sigma$ on $\Omega(\sigma(A))$.*

Proof As the matrix A is orthogonal, so is $\sigma(A)$. Also we know that, $\sigma(A^*) = \sigma(A)^*$. Now, let $X \in \Omega(A)$, that is, the matrix $A + X$ is invertible. Then $\sigma(A + X) = \sigma(A) + \sigma(X)$ is also invertible, hence $\sigma(X) \in \Omega(\sigma(A))$.

So we can compose σ and $c_{\sigma(A)}$ and we have

$$\begin{aligned} c_{\sigma(A)}(\sigma(X)) &= (I - \sigma(A)^*\sigma(X)) \cdot (\sigma(A) + \sigma(X))^{-1} \\ &= (I - \sigma(A^*X)) \cdot \sigma(A + X)^{-1} \\ &= \sigma(I - A^*X) \cdot \sigma(A + X)^{-1} \\ &= \sigma(c_A(X)). \end{aligned} \quad \square$$

Theorem 2 *Let $M \subset G$ be the Cartan model of the symmetric space G/K . Let $A \in M$. Then $\Omega_M(A) := \Omega(A) \cap M$ is a contractible open subspace of M .*

Proof From Theorem 1 we know that $\Omega_G(A) \cong T_{A^*}G$ can be contracted to A by the contraction $\nu : \Omega_G(A) \times [0, 1] \rightarrow \Omega_G(A)$ given by

$$\nu(X, t) = c_{A^*}(tc_A(X)).$$

When $A \in M$ we shall use the same contraction for $\Omega_M(A)$. Then it is enough to prove that when $X \in M$ the contraction $\nu(X, t)$ remains in M for all $t \in [0, 1]$. Since $M = N_I$ (Proposition 1), it suffices to prove that $\nu(X, t) \in N$ for all t , that is, $\sigma(\nu(X, t)) = \nu(X, t)^{-1}$. This can be shown as follows.

First, since $A, X \in M \subset N \subset G$, we have $\sigma(A) = A^{-1} = A^*$ and $\sigma(X) = X^{-1}$. Moreover, let $\nu(t) = \nu(X, t)$; then, since $X \in \Omega_G(A)$, it is $\nu(t) \in \Omega_G(A)$ (Theorem 1), hence $\nu(t)^{-1} = \nu(t)^* \in \Omega_G(A^*)$. On the other hand, $\nu(t) \in \Omega_G(A)$ implies that $\sigma(\nu(t)) \in \Omega_G(\sigma(A)) = \Omega_G(A^*)$. Finally, from Lemma 2 and the properties of the generalized Cayley transform [6, p. 327], we obtain

$$c_{A^*}(\nu(t)^{-1}) = -Ac_A(\nu(t))A = -Atc_A(X)A = tc_{A^*}(X^{-1})$$

while

$$\begin{aligned} c_{A^*}(\sigma(\nu(t))) &= \sigma(c_{\sigma(A^*)}(\nu(t))) = \sigma(c_A(\nu(t))) = \sigma(tc_A(X)) = t\sigma(c_A(X)) \\ &= tc_{\sigma(A)}(\sigma(X)) = tc_{A^*}(\sigma(X)) = tc_{A^*}(X^{-1}), \end{aligned}$$

and the result follows because c_{A^*} is injective. □

Remark 4 The existence of a covering by contractible open sets is of interest when computing the *Lusternik-Schnirelmann category* of M [6, 13].

5 Linearization of the gradient

The generalized Cayley transform studied above allows to linearize the gradient flow of any height function on a symmetric space. The results in this Section generalize Volchenko and Kozachko's result ([25], cited in [24]) for the classical Cayley transform in a Lie group.

5.1 Gradient equation

Theorem 3 *Let h_X^M be an arbitrary height function on the symmetric space M . Let A be a critical point. Then the solution of the gradient equation $4\alpha' = \widehat{X} - \alpha\sigma(\widehat{X})\alpha$, with initial condition $\alpha_0 \in \Omega_M(A)$, is the image by the Cayley transform c_{A^*} of the curve*

$$\beta(t) = \exp\left(\frac{-t}{4}A^*\widehat{X}\right)\beta_0\exp\left(\frac{-t}{4}\widehat{X}A^*\right), \tag{3}$$

where $\widehat{X} = X^* + \sigma(X)$ and $\beta_0 = c_A(\alpha_0) \in T_{A^*}M$.

Proof First of all, let us prove that $\beta(t) \in T_{A^*}M$ (cf. Proposition 2) for all $t \in \mathbb{R}$. Notice that being A a critical point implies that $A\widehat{X}^* = \widehat{X}^*A^*$, so this matrix is Hermitian. Now, using that $A \exp(XA) = \exp(AX)A$ for arbitrary matrices A, X , and that β_0 is in $T_{A^*}M$, we have

$$\begin{aligned} A^*\beta(t)^* &= A^* \exp\left(\frac{-t}{4}A\widehat{X}^*\right) \beta_0^* \exp\left(\frac{-t}{4}\widehat{X}^*A\right) \\ &= A^* \exp\left(\frac{-t}{4}\widehat{X}A^*\right) \beta_0^* \exp\left(\frac{-t}{4}\widehat{X}^*A\right) \\ &= \exp\left(\frac{-t}{4}A^*\widehat{X}\right) A^* \beta_0^* \exp\left(\frac{-t}{4}\widehat{X}^*A\right) \\ &= \exp\left(\frac{-t}{4}A^*\widehat{X}\right) (-\beta_0A) \exp\left(\frac{-t}{4}\widehat{X}^*A\right) \\ &= -\exp\left(\frac{-t}{4}A^*\widehat{X}\right) \beta_0 \exp\left(\frac{-t}{4}A\widehat{X}^*\right) A \\ &= -\exp\left(\frac{-t}{4}A^*\widehat{X}\right) \beta_0 \exp\left(\frac{-t}{4}\widehat{X}A^*\right) A \\ &= -\beta(t)A. \end{aligned}$$

Besides, we know that σ preserves the product, so

$$\sigma(\beta(t)) = \sigma\left(\exp\left(\frac{-t}{4}A^*\widehat{X}\right)\right) \sigma(\beta_0) \sigma\left(\exp\left(\frac{-t}{4}\widehat{X}A^*\right)\right). \tag{4}$$

Remember that $\sigma(\widehat{X}) = (\widehat{X})^*$ and that $\sigma(A) = A^{-1} = A^*$ because $A \in M \subset N$. Making use again of the algebraic properties of σ and knowing that $\beta_0 \in T_{A^*}M$ implies $\sigma(\beta_0) = \beta_0^*$, we have that (4) equals

$$\exp\left(\frac{-t}{4}\sigma(A^*\widehat{X})\right) \beta_0^* \exp\left(\frac{-t}{4}\sigma(\widehat{X}A^*)\right) = \exp\left(\frac{-t}{4}A\widehat{X}^*\right) \beta_0^* \exp\left(\frac{-t}{4}\widehat{X}^*A\right) = \beta(t)^*.$$

This proves that $\beta(t) \in T_{A^*}M$.

Now we shall verify that $c_{A^*}(\beta)$ is a solution of the gradient equation in Theorem 3. By differentiating β we have that

$$\beta' = (-1/4) (\beta\widehat{X}A^* + A^*\widehat{X}\beta). \tag{5}$$

Let $\alpha = c_{A^*} \circ \beta = (I - A\beta)(A^* + \beta)^{-1}$, then $\alpha \cdot (A^* + \beta) = I - A\beta$. Taking derivatives, $\alpha' \cdot (A^* + \beta) + \alpha\beta' = -A\beta'$, that is, $\alpha' \cdot (A^* + \beta) = -(A + \alpha) \cdot \beta'$.

It is not hard to check that the inverse of $A^* + \beta$ is $\frac{1}{2}(A + c_{A^*}(\beta)) = \frac{1}{2}(A + \alpha)$. Then

$$\alpha' = -\frac{1}{2}(A + \alpha)\beta'(A + \alpha) = \frac{1}{8}(A + \alpha)(\beta\widehat{X}A^* + A^*\widehat{X}\beta)(A + \alpha). \tag{6}$$

Moreover,

$$\beta = c_A(\alpha) = (I - A^*\alpha)(A + \alpha)^{-1} = (A + \alpha)^{-1}(I - \alpha A^*)$$

implies that $\beta(A + \alpha) = I - A^*\alpha$ and $(A + \alpha)\beta = I - \alpha A^*$. Therefore, using Eq. (6), we have that

$$\begin{aligned} 8\alpha' &= (A + \alpha)\beta\widehat{X}A^*(A + \alpha) + (A + \alpha)A^*\widehat{X}\beta(A + \alpha) \\ &= (I - \alpha A^*)\widehat{X}A^*(A + \alpha) + (A + \alpha)A^*\widehat{X}(I - A^*\alpha) \\ &= 2(\widehat{X} - \alpha A^*\widehat{X}A^*\alpha) \\ &= 2(\widehat{X} - \alpha\sigma(\widehat{X})\alpha). \end{aligned} \quad \square$$

By using the definitions $\beta_0 = (I - A^*\alpha_0)(A + \alpha_0)^{-1}$ and $c_{A^*}(\beta(t)) = A(A^* - \beta(t))(A^* + \beta(t))^{-1}$, and the property $A^* \exp(\widehat{X}A^*)A = \exp(A^*\widehat{X})$, one finally obtains an explicit formula for the solution, namely

$$\begin{aligned} \alpha(t) &= A \left(\sinh\left(\frac{t}{4}A^*\widehat{X}\right) + \cosh\left(\frac{t}{4}A^*\widehat{X}\right)A^*\alpha_0 \right) \\ &\quad \times \left(\cosh\left(\frac{t}{4}A^*\widehat{X}\right) + \sinh\left(\frac{t}{4}A^*\widehat{X}\right)A^*\alpha_0 \right)^{-1}. \end{aligned} \quad (7)$$

This formula, for the particular case of a Lie group G , the classical Cayley transform c_I and the particular height function h_D^G where D is a real diagonal matrix is due to Dynnikov and Veselov [24].

5.2 Local structure of the critical set

Giving a local chart for the critical set of a Morse-Bott function is another new application of the generalized Cayley transform.

Let $h_X^M(A) = \Re \operatorname{Tr}(XA)$ be a height function on the symmetric space M . Let us denote by $\Sigma(h_X^M)$ the critical set. Given a critical point $A \in \Sigma(h_X^M)$, let $S^M(A)$ be the vector space

$$S^M(A) = \{\beta_0 \in T_{A^*M} : A^*\widehat{X}\beta_0 + \beta_0\widehat{X}A^* = 0\}. \quad (8)$$

Observe that $S^M(A)$ is isomorphic to the kernel of the Hessian $H(h_X^M)_A$.

Theorem 4 *The generalized Cayley transform induces a diffeomorphism*

$$c_{A^*} : S^M(A) \rightarrow \Sigma(h_X^M) \cap \Omega_M(A).$$

Proof From Theorem 3, we know that $c_{A^*}(\beta_0)$ is a critical point if and only if $\beta'(t) = 0$ for all t , that is, the curve $\beta(t)$ is constant. Let us see that this is equivalent to the condition $\beta_0 \in S^M(A)$.

If $\beta(t) = \beta_0$ for all t then $\beta'(0) = 0$, which implies from Eq. (5) that $A^* \widehat{X} \beta_0 + \beta_0 \widehat{X} A^* = 0$.

Conversely, from $A^* \widehat{X} \beta_0 = -\beta_0 \widehat{X} A^*$ it follows that

$$\exp\left(-\frac{t}{4} A^* \widehat{X}\right) \beta_0 = \beta_0 \exp\left(\frac{t}{4} \widehat{X} A^*\right),$$

that is, $\beta(t) = \beta_0$ (Eq. 3). □

5.3 Relationship between gradient flows

As in Corollary 4, if one restricts to the case $X = I$ or more generally to $X = D$ a real diagonal matrix such that $\sigma(D) = D$, then the gradient flow of h_X^G will be tangent to the symmetric space M , embedded into G . That means that the restricted flow coincides with the gradient flow of the height function restricted to M , with respect to the induced metric, as in [24, Cor. 2.1]. But in general the symmetric space will not be invariant under the gradient flow. In fact, the gradient flow of h_X^G could even be transverse to M .

Example 4 Let us consider again Example 3. Take the critical point $A = (1/\sqrt{2})(\mathbf{j} + \mathbf{k}) \in M$. According to formula (7) the gradient flow line of h_X^M passing through $\alpha_0 = 1$ is given by $\alpha^M(t) = \operatorname{sech} t\sqrt{2} - \mathbf{j}(\tanh t\sqrt{2}) \frac{1-\mathbf{i}}{\sqrt{2}} \in M$. On the other hand, the flow line of h_X^G in the group G passing through the same point $\alpha_0 = 1$ is $\alpha^G(t) = \operatorname{sech}(t\sqrt{3}) - \tanh(t\sqrt{3}) \frac{\mathbf{i}+\mathbf{j}+\mathbf{k}}{\sqrt{3}}$, where we have chosen the Cayley transform corresponding to the critical point $A = (1/\sqrt{3})(\mathbf{i} + \mathbf{j} + \mathbf{k})$. Notice that $\alpha^G(t) \notin M$ for $t \neq 0$.

6 Polar decompositions and critical set

The study of Morse-Bott functions can be considerably simplified by means of the so-called *singular value decomposition*. Also there is an interesting relationship between polar forms and critical points. The main result in this Section will be the existence of decompositions which are adapted to the automorphism σ that defines the symmetric space.

6.1 Singular value decomposition and polar form

The following constructions are well known [10]. For the quaternionic case see [16, 28].

Let $Y \in \mathbb{K}^{n \times n}$ be an arbitrary matrix. Since the matrix YY^* is Hermitian positive-semidefinite (hereafter abbreviated to H.p.-s.), its eigenvalues are real and non-negative, say $0, t_1^2, \dots, t_k^2$, with multiplicities n_0, n_1, \dots, n_k , respectively. Let us consider the matrix with diagonal blocks

$$D = \begin{pmatrix} \overline{0_{n_0}} & & & \\ & \overline{t_1 I_{n_1}} & & \\ & & \ddots & \\ & & & \overline{t_k I_{n_k}} \end{pmatrix} \in \mathbb{K}^{n \times n}, \tag{9}$$

where $0 < t_1 < \dots < t_k$ and $n_0 + \dots + n_k = n$. Then there exist orthogonal (resp. unitary, symplectic) matrices U, V such that $Y = UDV^*$ (singular value decomposition). As a consequence we have the left polar decomposition $Y = S\Omega$, where the matrix $\Omega = UV^*$ is orthogonal and the matrix $S = UDU^*$ is H.p.-s. In fact, S is uniquely determined as the only H.p.-s. square root of $YY^* = UD^2U^*$.

Analogously we have the right polar decomposition $Y = \Omega S'$, where $S' = VDV^* = \Omega^* S \Omega$ is the only H.p.-s. square root of $Y^*Y = VD^2V^*$. We shall denote $S = (YY^*)^{1/2}$ and $S' = (Y^*Y)^{1/2}$ the H.p.-s. matrices above. Notice that the orthogonal part Ω in the polar decompositions is not unique, excepting when Y is invertible.

Conversely, from any polar decomposition $Y = S\Omega$ it is possible to deduce a singular value decomposition because $S = (YY^*)^{1/2}$ is diagonalizable, say $S = UDU^*$. Then taking $V = \Omega^*U$ gives $Y = UDV^*$.

Lemma 3 *Let Σ be an Hermitian square root of YY^* . Then there exists some orthogonal matrix W such that $S = WDW^*$ and $\Sigma = W\Delta W^*$, where Δ is a real diagonal matrix verifying $\Delta^2 = D^2$ (that is, Δ only differs from D by the signs of the entries).*

Proof Since Σ commutes with $S^2 = \Sigma^2$, they can be simultaneously diagonalized. Moreover, the ordering of the diagonal of D can be recovered by conjugation with a real orthogonal matrix. □

6.2 Critical points

Let us consider the function $h_Y^G(A) = \Re \operatorname{Tr}(YA)$. According to Proposition 5, each critical point $A \in G$ determines a decomposition $Y = (YA)A^*$ of Y , where $\Sigma = YA$ is Hermitian but no necessarily positive-semidefinite. Let us call *almost polar* such a decomposition. Since $\Sigma^2 = YY^*$ we know from Lemma 3 that $\Sigma = W\Delta W^*$, where

$$\Delta = D \cdot \begin{pmatrix} \overline{0_{n_0}} & & & \\ & \overline{E_1} & & \\ & & \ddots & \\ & & & \overline{E_k} \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad E_i = \begin{pmatrix} \varepsilon_i^1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \varepsilon_i^{n_i} \end{pmatrix}, \quad \varepsilon_i^j = \pm 1.$$

Then the value of the function at the critical point A is

$$h_Y^G(A) = \Re \operatorname{Tr}(\Sigma) = t_1 \operatorname{Tr} E_1 + \dots + t_k \operatorname{Tr} E_k. \tag{10}$$

Remark 5 When $\mathbb{K} = \mathbb{R}$, the orthogonal part of the polar decomposition may not be in the connected component $G = SO(n)$ of the identity. In order to avoid cumbersome technicalities in the next Proposition, when dealing with the real case we should consider that the group is $O(n)$.

Proposition 6 *The critical point $A \in G$ is a global maximum of h_Y^G if and only if the decomposition $(YA)A^*$ is a true polar decomposition (i.e. the Hermitian matrix $\Sigma = YA$ is positive-semidefinite).*

Proof If A is a global maximum for h_Y^G , then it is a critical point where the function value is that given in (10). But since Y actually has a polar decomposition ΩS it happens that the maximum value is

$$\Re \operatorname{Tr}(S) = \Re \operatorname{Tr}(D) = n_1 t_1 + \cdots + n_k t_k. \tag{11}$$

Comparing both values it follows that all $\varepsilon_{ij} = +1$, that is, $\Delta = D$ and $\Sigma = S$. Notice that the global maximum may not be unique due to the non-uniqueness of Ω when Y is singular.

Conversely, if YA is positive semi-definite then the critical value $h_X(A)$ is that given in (11). This value is then bigger than all the other critical values, by formula (10). But since G is compact, the function actually has a global maximum, which is a critical point. Then it must be in the same level that the point A . \square

Remark 6 When $Y = SA^*$ is a polar decomposition it is well known that A maximizes the distance of Y to the orthogonal matrices. In the same way, it maximizes the function $\Re \operatorname{Tr}(YB)$, $B \in G$.

6.3 Polar decompositions in a symmetric space

According to Corollary 2, the critical points of h_X^M are the critical points of $h_{\sigma(\widehat{X})}^G$ that lie in M . This is why we are interested in the matrix $\widehat{X} = X^* + \sigma(X)$. It verifies $\sigma(\widehat{X}) = \widehat{X}^*$.

Assumption Let $\widehat{X} = UDV^*$ be a singular value decomposition. We shall assume that $\sigma(D)$ is positive-semidefinite. This hypothesis is not too restrictive. For the symmetric spaces in Cartan's classification it happens that either $\sigma(D) = D$ or $\sigma(D) = -JDJ$, with $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, which has the effect of permuting the entries of D in two blocks. The latter case corresponds to the symmetric spaces $U(2n)/Sp(n)$ and $SO(2n)/U(n)$.

Theorem 5 (adapted polar decomposition) *Let $Y \in \mathbb{K}^{n \times n}$ be a matrix such that $\sigma(Y) = Y^*$. Assume that $\sigma(D)$ is positive semi-definite for the matrix D of singular values of Y . Then there exists a polar decomposition $Y = S\Omega$ such that $\sigma(\Omega) = \Omega^*$ and $\sigma(S) = \Omega^*S\Omega$.*

Proof First, let us assume that Y is invertible. Then the polar decompositions $Y = S\Omega$ (left) and $Y = \Omega(\Omega^*S\Omega)$ (right) are unique. Clearly the matrix $\sigma(\Omega)$ is orthogonal, while $\sigma(S)$ is Hermitian. Moreover $\sigma(S) = \sigma(U)\sigma(D)\sigma(U)^*$, hence it is H.p.-s. by the assumption. Then $Y = \sigma(Y)^* = \sigma(\Omega)^*\sigma(S)$ is a (right) polar decomposition, so by uniqueness $\sigma(\Omega) = \Omega^*$ and $\sigma(S) = \Omega^*S\Omega$.

We now deal with the non-invertible case. Let \mathfrak{Y} be the real vector space

$$\mathfrak{Y} = \{Y \in \mathbb{K}^{n \times n} : \sigma(Y) = Y^*\}.$$

Then $\mathfrak{Y} \cap G$ is the manifold N defined in Proposition 1. Let $\delta : \mathfrak{Y} \times G \rightarrow \mathbb{R}$ be the continuous function

$$\delta(Y, A) = \Re \operatorname{Tr}(YA).$$

According to the first part of the proof and to Proposition 6, we know that for each invertible matrix Y the function $\delta(Y, -) : G \rightarrow \mathbb{R}$ has a unique global maximum Ω which belongs to N . Then from Lemma 4 below it follows that all functions $\delta(Y, -)$, $Y \in \mathfrak{Y}$, have at least one maximum in N . But maxima correspond to polar decompositions by Proposition 6, so the result follows. \square

Lemma 4 *Let $\mathfrak{I} \subset \mathfrak{Y}$ be the dense subspace of invertible matrices. Assume that each map $\delta(Y, -)$, with $Y \in \mathfrak{I}$, has a unique global maximum which belongs to N . Then all maps $\delta(Y, -)$, with $Y \in \mathfrak{Y}$, have at least one maximum which belongs to N .*

Proof The invertible matrices form a dense subset in \mathfrak{Y} , because for any $Y \in \mathfrak{Y}$ it is

$$\sigma(Y + \varepsilon I) = \sigma(Y) + \varepsilon \sigma(I) = (Y + \varepsilon I)^*.$$

Let Y_0 be a non invertible matrix in \mathfrak{Y} . Let $A_0 \in G$ be any global maximum of the function $\delta(Y_0, -) : G \rightarrow \mathbb{R}$ and let $t_0 = \delta(Y_0, A_0)$. Then there exists a sequence (Y_n, A_n) converging to (Y_0, A_0) , where $Y_n \in \mathfrak{Y}$ is invertible and $A_n \in G$, such that $t_n = \delta(Y_n, A_n)$ converges to t_0 . Let $N_n \in G$ be the sole global maximum of $\delta(Y_n, -)$. Since N is compact, there exists a subsequence (N_n) which converges to some $N_0 \in N$. Then taking limits in $\delta(Y_n, A_n) \leq \delta(Y_n, N_n)$ we obtain that $t_0 = \delta(Y_0, A_0) \leq \delta(Y_0, N_0)$, hence $N_0 \in N$ corresponds to the maximum t_0 . \square

Corollary 5 *Under the assumption, there exists a right polar decomposition $Y = \Omega S'$ such that $\sigma(\Omega) = \Omega^*$ and $\sigma(S') = \Omega S' \Omega^*$.*

Corollary 6 (adapted SVD) *There exists a singular value decomposition $Y = U D V^*$ such that the matrix $\Theta = U^* \sigma(V)$ verifies $\sigma(\Theta) = \Theta^*$.*

Proof Let us take the polar decomposition $Y = S \Omega$ given by Theorem 5 and the corresponding singular value decomposition $Y = U D V^*$. Then $\sigma(UV^*) = \sigma(\Omega) = \Omega^* = V U^*$. \square

7 Height functions associated to real diagonal matrices

In this section we show that, under mild hypothesis, the study of any height function h_X^M on the symmetric space M can be reduced to the particular case $h_D^{M'}$ where D is a real non-negative diagonal matrix and $M' \subset G$ is diffeomorphic to M .

7.1 Reduction to the diagonal case

We begin with the following Lemma, which is a straightforward exercise.

Lemma 5 *Let Θ be an orthogonal matrix such that $\sigma(\Theta) = \Theta^*$. Let $\sigma': G \rightarrow G$ be the involutive automorphism $\sigma'(X) = \Theta\sigma(X)\Theta^*$ and let K' be its fixed point subgroup. Then the manifolds N, N' associated to σ, σ' respectively by Proposition 1 verify that $N' = \Theta N$.*

Let $h_X^M : M \rightarrow \mathbb{R}$ be a height function on the symmetric space $M = G/K$ and let $\widehat{X} = X^* + \sigma(X)$. Let $\widehat{X} = UDV^*$ be an adapted singular value decomposition as in the preceding Section (recall that we are assuming that $\sigma(D)$ is positive semi-definite). Then, from

$$\sigma(V)\sigma(D)\sigma(V)^* = \sigma(S) = \Omega S \Omega^* = UDU^*$$

it follows that $\sigma(D) = \Theta^* D \Theta$. Now, from Lemma 5 we have that $\sigma'(D) = D$, which implies that the matrix $\widehat{D}' := D^* + \sigma'(D)$ equals $2D$, while $\sigma'(\widehat{D}') = 2D$.

Let M' be the symmetric space associated to σ' . Hence (Proposition 3) the gradient of the height function $h_D^{M'} : M' \rightarrow \mathbb{R}$ at the point $B \in M'$ is

$$\left(\text{grad } h_D^{M'}\right)_B = \frac{1}{2}(D - BDB). \tag{12}$$

Proposition 7 *Assume $M = N$. Then the point $A \in M$ is a critical point of h_X^M if and only if $U^*AV \in M'$ is a critical point of $h_D^{M'}$, where $\widehat{X} = UDV^*$ is an adapted SVD.*

Proof In fact, by Proposition 3 and formula (12) we have that

$$\left(\text{grad } h_X^M\right)_A = 2U \left(\left(\text{grad } h_D^{M'}\right)_{U^*AV} \right) V^*$$

because $V^*\sigma(U) = \Theta = U^*\sigma(V)$. Now, $U^*AV \in N'$ because

$$\sigma'(U^*AV) = V^*\sigma(U)\sigma(U^*AV)\sigma(V)^*U = V^*A^*U = (U^*AV)^*.$$

Finally, $M = N$ implies $\Theta \in M$ and $M' = N'$.

Analogously we have for the Hessian that

$$\left(Hh_X^M\right)_A(W) = 2U \left(\left(Hh_D^{M'}\right)_{U^*AV}(U^*WV) \right) V^*.$$

7.2 Description of the critical set

According to the previous section we can assume that $X = D$ is a diagonal matrix like in (9), and that $\sigma(D) = D$, hence $\widehat{D} = 2D$.

Lemma 6 *Let A be a critical point of h_D^G , that is, $D = ADA$. Then A can be decomposed into boxes, of size n_0, n_1, \dots, n_k respectively,*

$$A = \begin{pmatrix} \underline{A_0} & & & 0 \\ & \underline{A_1} & & \\ & & \ddots & \\ 0 & & & \underline{A_k} \end{pmatrix}$$

such that $A_0A_0^* = I$ and $A_i^2 = I, A_i = A_i^*$, for $1 \leq i \leq k$.

Proof Write boxes

$$A = \begin{pmatrix} A_0 & A_{01} & A_{02} & \dots \\ A_{10} & A_1 & A_{12} & \dots \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

From the condition $DA^* = AD$, it follows that $A_{0j} = 0$ and $A_i = A_i^*, i, j = 1, \dots, k$. Moreover $t_i A_{ji}^* = t_j A_{ij}$ for $i < j$. From the conditions $AA^* = I = A^*A$ it follows that $A_0A_0^* = I$ and $A_{i0} = 0, i = 1, \dots, k$. Also

$$A_{11}^2 + A_{12}A_{12}^* + \dots + A_{1k}A_{1k}^* = I = A_{11}^2 + A_{21}^*A_{21} + \dots + A_{k1}^*A_{k1}$$

that implies

$$(t_2^2 - t_1^2)A_{12}A_{12}^* + \dots + (t_k^2 - t_1^2)A_{1k}A_{1k}^* = 0$$

from which it follows easily that $A_{12} = \dots = A_{1k} = 0$. For the other rows the computation is similar. □

Remark 7 Recall that $\Sigma(h_D^M) = \Sigma(h_D^G) \cap M$. A precise description of the critical set of h_D^G can be deduced from the previous Lemma, see [6, p. 329] and [17]. Then h_D^G is a Morse function if and only if $\dim S^M(A) = 0$ (cf. Sect. 5.2), which is equivalent to $n_0 = 0$ and $n_1 = \dots = n_k = 1$.

From Proposition 5 we obtain the following Corollary,—this result appears in Brockett’s paper [1, p. 771], where it is attributed to Shayman, see also [8, p. 99].

Corollary 7 *The height function h_X^G in the Lie group G is a Morse function if and only if the singular values of the matrix X are positive and pairwise different.*

The following Example illustrates all the results in this Section.

7.3 Final example

Let $G = Sp(2)$ and $\sigma(A) = -\mathbf{i}A\mathbf{i}$, then the fixed point subgroup is $K = U(2)$. The symmetric space $G/K = Sp(2)/U(2)$ can be identified with the manifold of complex structures on \mathbb{H}^2 which are compatible with the hermitian product. These are the matrices $\mathcal{J} \in Sp(2)$ such that $\mathcal{J}^2 = -I$. Let $\mathcal{J}_0 = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{pmatrix}$. Then the Cartan embedding $\gamma: G/K \rightarrow Sp(2)$ is $\gamma(\mathcal{J}) = -\mathcal{J}\mathcal{J}_0$. As a consequence, it can be proven that the Cartan model M equals the manifold N of matrices such that $\sigma(A) = A^*$ (Proposition 1). Explicitly, it is formed by the diagonal matrices

$$\begin{pmatrix} \alpha & 0 \\ 0 & \delta \end{pmatrix}, \quad |\alpha| = |\delta| = 1, \Re(\alpha\mathbf{i}) = \Re(\delta\mathbf{i}) = 0,$$

jointly with the matrices

$$\begin{pmatrix} \alpha & -\mathbf{i}\bar{\beta}\mathbf{i} \\ \beta & \beta\bar{\alpha}\mathbf{i}\beta^{-1}\mathbf{i} \end{pmatrix}, \quad \beta \neq 0, |\alpha|^2 + |\beta|^2 = 1, \Re(\alpha\mathbf{i}) = 0.$$

Let us take $X = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \in \mathbb{H}^{2 \times 2}$ with $x = 1 + \mathbf{j}$ and $y = \mathbf{i} + \mathbf{j}$.

First, we study the function h_X^G on the Lie group G . Since $XX^* = \begin{pmatrix} |x|^2 & 0 \\ 0 & |y|^2 \end{pmatrix} = 2I$, we have the s.v.d. $X = UDV^*$ with $U = \frac{1}{\sqrt{2}}X$, $D = \text{diag}(\sqrt{2}, \sqrt{2})$ and $V = I$. The critical set $\Sigma(h_X^G)$ is then diffeomorphic to $\Sigma(h_D^G)$, which is a disjoint union $\Sigma(2) = G_0^2 \sqcup G_1^2 \sqcup G_2^2$ of Grassmannians [6], namely, two points and $Sp(2)/(Sp(1) \times Sp(1)) \cong S^4$. More precisely, from the gradient equation $D = ADA$ it follows that the three components are $\{\pm I\}$ and the sphere

$$\left\{ \begin{pmatrix} s & \bar{\beta} \\ \beta & -s \end{pmatrix} : s \in \mathbb{R}, s^2 + |\beta|^2 = 1 \right\},$$

which are the orbits by the adjoint action of I , $-I$ and $\pm P = \pm \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ respectively [4]. From the index of each orbit one easily obtains the critical values of the height function.

Finally we have an explicit description of the critical set as $\Sigma(h_X^G) = V\Sigma(h_D^G)U^*$, that is the disjoint union of $\{\pm U^*\}$ and $G_1^2U^*$ (Proposition 7, see also [17, p. 3]). It could be computed directly from the gradient equation $X^* = AXA$. Notice that $\Sigma(h_X^G) \cap M = \emptyset$, that is, none of the critical points of h_X^G is contained in the symmetric space $M \cong G/K$.

Now we can compute the critical points of the restriction h_X^M of the height function to $M \subset G$. We shall use the formula $\Sigma(h_X^M) = \Sigma(h_{\sigma(\widehat{X})}^G) \cap M$. (Corollary 2), where $\widehat{X} = X^* + \sigma(X)$.

We have $\sigma(\widehat{X}) = \widehat{X}^* = \begin{pmatrix} x_0 & 0 \\ 0 & y_0 \end{pmatrix}$, where $x_0 = \bar{x} - \mathbf{i}x\mathbf{i} = 2\mathbf{j}$ and $y_0 = \bar{y} - \mathbf{i}y\mathbf{i} = 2 + 2\mathbf{j}$. Notice that $|x| = |y|$ but $|x_0| \neq |y_0|$. This time we take the SVD $\widehat{X}^* = UDV^*$ where $U = \begin{pmatrix} \mathbf{j} & 0 \\ 0 & \frac{1}{\sqrt{2}}(1 + \mathbf{j}) \end{pmatrix}$, $D = \text{diag}(2, 2\sqrt{2})$ and $V = I$. This is an adapted SVD (Corollary 6), that is, the matrix $\Theta = U^*\sigma(V) = \begin{pmatrix} -\mathbf{j} & 0 \\ 0 & \frac{1}{\sqrt{2}}(1 - \mathbf{j}) \end{pmatrix}$ verifies $\sigma(\Theta) = \Theta^*$.

The critical set $\Sigma(h_D^G)$ is diffeomorphic to $\Sigma(1) \times \Sigma(1)$, that is, four points, because $\Sigma(1) = G_0^1 \sqcup G_1^1$ [6]. Explicitly,

$$\Sigma(h_D^G) = \{A \in Sp(2) : DA^* = AD\} = \{\pm I, \pm P\}.$$

Now, $\Sigma(h_{\sigma(\widehat{X})}^G) = V\Sigma(h_D^G)U^* = \{\pm U^*, \pm PU^*\}$, and these four points are in M , then it follows that h_X^M is a Morse function with four critical points.

Finally, we shall verify that $\Sigma(h_X^M) = V\Sigma(h_D^{M'})U^*$, where $M' = \Theta M$ (Lemma 5); remember that $M = N$ in the present example.

According to Lemma 5 the new automorphism $\sigma' : G \rightarrow G$ is given by

$$\sigma'(A) = \Theta\sigma(A)\Theta^* = U^*(-\mathbf{i}A\mathbf{i})U = -\mathbf{i}UAU^*\mathbf{i}$$

and we check that $\sigma'(D) = -\mathbf{i}UDU^*\mathbf{i} = -\mathbf{i}D\mathbf{i} = D$, and $\widehat{D}' = D^* + \sigma'((D)) = 2D$. Then

$$\Sigma(h_D^{M'}) = \Sigma(h_{\sigma'(\widehat{D}')}^G \cap M') = \Sigma(h_{2D}^G) \cap M'.$$

We already know that $\Sigma(h_{2D}^{M'}) = \Sigma(h_D^{M'}) = \{\pm I, \pm P\}$. On the other hand $M' = \Theta M = U^*M$, and $\Sigma(h_D^G) \cap M' = \{\pm I, \pm P\}$. Then $V\Sigma(h_D^{M'})U^* = \{\pm IU^*, \pm PU^*\}$, which equals $\Sigma(h_X^M)$ as promised.

Acknowledgments Partially supported by Research Project MTM2013-41768-P, MINECO, Spain.

References

1. Brockett, R.W.: Least squares matching problems. *Linear Algebra Appl.* **761**, 123–124 (1989)
2. Duan, H.: Morse functions and cohomology of homogeneous spaces. In: Pragacz, P. (ed.) *Topics in Cohomological Studies of Algebraic Varieties*. Impanga Lecture Notes Mathematics, pp. 105–133. Birkhäuser, Boston (2005)
3. Fomenko, A.T.: *Differential Geometry and Topology*, Contemporary Soviet Mathematics. Consultants Bureau, New York-London (1987)
4. Frankel, T.: Critical submanifolds of the classical groups and Stiefel manifolds. *Differ. and Combinat. Topology, Sympos.* Marston Morse, Princeton, pp. 37–53 (1963)
5. Gallot, S., Hulin, D., Lafontaine, J.: *Riemannian Geometry*. Universitext. Springer, Berlin (2004)

6. Gómez-Tato, A., Macías-Virgós, E., Pereira-Sáez, M.J.: Trace map, Cayley transform and LS category of Lie groups. *Ann. Global Anal. Geom.* **39**(3), 325–335 (2011)
7. Helgason, S.: *Differential Geometry, Lie Groups and Symmetric Spaces*. Academic Press, Massachusetts (1978)
8. Helmke, U., Moore, J.B.: *Optimization and Dynamical Systems*. Springer, London (1996)
9. Helmke, U., Shayman, M.A.: Critical points of matrix least squares distance functions. *Linear Algebra Appl.* **215**, 1–19 (1995)
10. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, Cambridge (1985)
11. Munthe-Kaas, H.Z., Quispel, G.R.W., Zanna, A.: Generalized polar decompositions on lie groups with involutive automorphisms. *Found. Comput. Math.* **1**, 297–324 (2001)
12. Kadzisa, H., Mimura, M.: Cartan models and cellular decompositions of symmetric Riemannian spaces. *Topol. Appl.* **156**, 348–364 (2008)
13. Kadzisa, H., Mimura, M.: Morse-Bott functions and the Lusternik-Schnirelmann category. *J. Fixed Point Theory Appl.* **10**, 63–85 (2011)
14. Kobayashi, S.: Isometric imbeddings of compact symmetric spaces. *Tôhoku Math. J.* **20**, 21–25 (1968)
15. Lawson, J.D.: Polar and Ol'shanskii decompositions. *J. Reine Angew. Math.* **448**, 191–219 (1994)
16. Loring, T.A.: Factorization of matrices of quaternions. *Expo. Math.* **30**(3), 250–267 (2012)
17. Macías-Virgós, E., Pereira-Sáez, M.J.: An upper bound for the Lusternik–Schnirelmann category of the symplectic group. *Math. Proc. Camb. Philos. Soc.* **155**(2), 271–276 (2013)
18. Milnor, J.W.: Morse theory. Based on lecture notes. In: Spivak, M., Wells, R. (eds.) *Annals of Mathematics Studies*, No.51. Princeton University Press VI, Princeton, N.J. (1963)
19. Nicolaescu, L.I.: Morse theory on Grassmannians. *An. Ştiinţ. Univ. Al. I. Cuza Iaşi, Ser. Nouă, Mat.* **40**(1), 25–46 (1994)
20. Ramanujam, S.: An application of Morse theory to some homogeneous spaces. *Tôhoku Math. J.* **21**(3), 344–353 (1969)
21. Ramanujam, S.: Morse theory of certain symmetric spaces. *J. Differ. Geom.* **3**, 213–229 (1969)
22. Ramanujam, S.: Topology of classical groups. *Osaka J. Math.* **6**, 243–249 (1969)
23. Takeuchi, M.: Nice functions on symmetric spaces. *Osaka J. Math.* **6**(2), 283–289 (1969)
24. Veselov, A.P., Dynnikov, I.A.: Integrable gradient flows and Morse theory. *St. Petersburg. Math. J.* **8**(3), 429–446 (1997); translation of *Algebra Anal.* **8**(3), 78–103 (1996)
25. Volchenko, K.Y., Kozachko, A.N.: Integrable gradient flows on classic manifolds and Morse theory. *Moscow Univ. Math. Bull.* **52**(3), 9–15 (1997); translation of *Vestnik. Moskow. Univ. Ser. Mat. I*, **3**, 9–15 (1997)
26. Wu, R.B., Chakrabarti, R., Rabitz, H.: Critical landscape topology for optimization on the symplectic group. *J. Optim. Theory Appl.* **145**, 387–406 (2010)
27. Ziller, W.: Lie groups, representation theory and symmetric spaces. Chapter 6. 2010. <http://www.math.upenn.edu/wziller/math650/LieGroupsReps.pdf>
28. Zhang, F.: Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251**, 21–57 (1997)