Mathematisches
Forschungsinstitut
Oberwolfach

Mitglied der Leibniz-Gemeinschaft

Oberwolfach Preprints

OWP 2015 - 10

ENRIQUE MACÍAS-VIRGÓS, JOHN OPREA, JEFF STROM AND DANIEL TANRÉ

Height Functions on Quaternionic Stiefel Manifolds

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Oberwolfach Preprints (OWP)

Starting in 2007, the MFO publishes a preprint series which mainly contains research results related to a longer stay in Oberwolfach. In particular, this concerns the Research in Pairs-Programme (RiP) and the Oberwolfach-Leibniz-Fellows (OWLF), but this can also include an Oberwolfach Lecture, for example.

A preprint can have a size from 1 - 200 pages, and the MFO will publish it on its website as well as by hard copy. Every RiP group or Oberwolfach-Leibniz-Fellow may receive on request 30 free hard copies (DIN A4, black and white copy) by surface mail.

Of course, the full copy right is left to the authors. The MFO only needs the right to publish it on its website *www.mfo.de* as a documentation of the research work done at the MFO, which you are accepting by sending us your file.

In case of interest, please send a **pdf file** of your preprint by email to *rip@mfo.de* or *owlf@mfo.de*, respectively. The file should be sent to the MFO within 12 months after your stay as RiP or OWLF at the MFO.

There are no requirements for the format of the preprint, except that the introduction should contain a short appreciation and that the paper size (respectively format) should be DIN A4, "letter" or "article".

On the front page of the hard copies, which contains the logo of the MFO, title and authors, we shall add a running number (20XX - XX).

We cordially invite the researchers within the RiP or OWLF programme to make use of this offer and would like to thank you in advance for your cooperation.

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

HEIGHT FUNCTIONS ON QUATERNIONIC STIEFEL MANIFOLDS

ENRIQUE MACÍAS-VIRGÓS, JOHN OPREA, JEFF STROM, AND DANIEL TANRÉ

Abstract. In this note, we study height functions on quaternionic Stiefel manifolds and prove that all these height functions are Morse-Bott. Among them, we characterize the Morse functions and give a lower bound for their number of critical *values.* Relations with the Lusternik-Schnirelmann category are discussed.

1. Introduction

The Lusternik-Schnirelmann category, $\operatorname{cat} X$, of a topological space X is defined as the least integer $m \geq 0$ such that X admits a covering by $m+1$ open sets which are contractible in X [1]. Category has proven useful in areas such as dynamical systems and symplectic geometry as well as homotopy theory, but it has also proven to be notoriously difficult to compute. A longstanding problem has been to compute the LS category of Lie groups. A significant step forward was Singhof's computations $\operatorname{cat} \mathrm{U}(n) = n$ and $\operatorname{cat} \mathrm{SU}(n) = n-1$ [19] of the unitary and special unitary groups, accomplished through a clever use of eigenvalues. Such an approach cannot be carried out for the symplectic groups, $\mathrm{Sp}(n)$, [10]. Some progress has been made, however, for small n and in obtaining lower bounds. Namely, $\operatorname{cat} \mathrm{Sp}(2) = 3$ [17], $\operatorname{cat} \mathrm{Sp}(3) = 5$ [3], and $\operatorname{cat} \mathrm{Sp}(n) \geq n+2$ when $n \geq 3$ [7].

The quaternionic Stiefel manifolds, $X_{n,k} = \mathrm{Sp}(n)/\mathrm{Sp}(n-k)$, have proved somewhat more amenable to category calculations in certain ranges. For instance, $\operatorname{cat} X_{n,k} = k$ when $n \geq 2k$. This result was proved by Nishimoto [15] using the number of eigenvalues of a complex matrix in a fashion similar to Singhof's approach for the unitary group [3]. This has also been proved by Kadzisa and Mimura [9] by a method based on the cone decomposition induced by the Morse-Bott function defined by Frankel in [4]. The analysis of Morse-Bott functions on Lie groups and homogeneous spaces has a long history (see, for instance, [6, 12, 20]). In this paper, we focus our attention on the properties of height functions themselves and relate these to the LS-category when the critical points are isolated [16]. With this method, it was

Date: January 23, 2015.

2010 *Mathematics Subject Classification.* Primary 55M30; Secondary 58C99.

Key words and phrases. Lusternik-Schnirelmann category; quaternionic Stiefel manifold; Morse-Bott theory; height function; critical values.

proved in [11] that the category of the symplectic group $\mathrm{Sp}(n)$ is bounded above by $\binom{n+1}{2}$.

We extend this approach to quaternionic Stiefel manifolds and prove that any height function on $X_{n,k}$ is of Morse-Bott type. In order to do this, we collect basic properties on general height functions on $X_{n,k}$, giving an explicit expression for the gradient and an explicit determination of the critical set. We also prove that any Morse height function on $X_{n,k}$ is perfect and has at least $1+\binom{k+1}{2}$ critical *values.* Moreover, this bound is reached by some height function. This gives an upper bound for the LS-category of $X_{n,k}$, which is not the best possible but shows the limit of the method.

2. Height functions on quaternionic Stiefel manifolds

2.1. **Stiefel manifolds.** Let $\mathbb{H}^n$ be the quaternionic n-space (with the structure of a right $\mathbb{H}$-vector space) endowed with the hermitian product $\langle u, v\rangle = u^*v$. For $0 \leq k \leq n$, let $X_{n,k}$ be the Stiefel manifold of linear maps $\phi\colon \mathbb{H}^k \to \mathbb{H}^n$ which preserve the Hermitian product. The map ϕ can be identified with a matrix x of size $n\times k$ such that $x^*x = I_k$, so the columns of x form an orthonormal k-frame of $\mathbb{H}^n$. We shall represent any element $x \in X_{n,k}$ by two blocks,

$$x = \begin{pmatrix} T \\ P \end{pmatrix},$$

where T, P are quaternionic matrices of order $(n-k)\times k$ and $k \times k$, respectively, which (due to $x^*x = I_k$) verify the relation

$$T^*T + P^*P = I_k. \tag{1}$$

Let $\mathrm{Sp}(n)$ be the Lie group of $n \times n$ matrices A such that $A^*A = I_n$. The linear left action of $\mathrm{Sp}(n)$ on $X_{n,k}$ is transitive and the isotropy group of $x_0 = \begin{pmatrix} 0 \\ I_k \end{pmatrix}$ is isomorphic to $\mathrm{Sp}(n-k)$, so $X_{n,k}$ is diffeomorphic to $\mathrm{Sp}(n)/\mathrm{Sp}(n-k)$.

2.2. **Height functions.** Let $\mathbb{H}^{n\times k}$ be the vector space of quaternionic matrices of size $n\times k$. As a real vector space, it is isomorphic to $\mathbb{R}^{4(n\times k)}$ with *euclidean* norm given by $|x|^2 = \mathrm{Tr}(x^*x)$. It follows that the euclidean inner product is given by $[y, x] = \Re\mathrm{Tr}(y^*x)$, where $\Re\mathrm{Tr}$ is the real part of the trace, and the height of x with respect to the hyperplane orthogonal to a given matrix ω^* is given by $a\,\Re\mathrm{Tr}(\omega x) + b$, where a and b are real constants. Let $h_\omega\colon \mathbb{H}^{n\times k} \to \mathbb{R}$ be the function

$$h_\omega(x) = \Re\mathrm{Tr}(\omega x).$$

Since h_ω is $\mathbb{R}$-linear, its gradient at any point is ω^*. If we denote ω by a block matrix $\left(\delta\middle|D\right)$, with blocks δ, D of size $k\times(n-k)$ and $k\times k$ respectively, then

$$h_\omega\left(\begin{pmatrix} T \\ P \end{pmatrix}\right) = \Re\mathrm{Tr}(\omega x) = \Re\mathrm{Tr}(\delta T + DP). \tag{2}$$

When $n = k$, the notation above means $\omega = D \in \mathbb{H}^{n\times n}$, $x = P$ and the Stiefel manifold is the group $\mathrm{Sp}(n)$. The corresponding height function is that considered in [11].

2.3. **Height functions on $X_{n,k}$.** Let us now consider the restriction $f_\omega\colon X_{n,k}\to\mathbb{R}$ of the height function h_ω to the Stiefel manifold. We want to choose the matrix ω in such a way that f_ω has isolated critical points and few critical *levels.* In order to determine the critical points, we first need expressions for the gradient and the Hessian of f_ω.

Proposition 2.1. *The gradient of $f_\omega\colon X_{n,k}\to\mathbb{R}$ at the point $x\in X_{n,k}$ is given by*

$$(\mathrm{grad} f_\omega)_x = \omega^* - (1/2)x(\omega x + x^*\omega^*).$$

Proof. Begin with $x_0 = \begin{pmatrix} 0 \\ I_k \end{pmatrix}$. The tangent space, $T_{x_0}(X_{n,k})$, is obtained as the image of the tangent space at $I_n\in\mathrm{Sp}(n)$ by the canonical projection $\mathrm{Sp}(n)\to X_{n,k}$. It is therefore formed by the matrices $\begin{pmatrix} X \\ Y \end{pmatrix}$ such that $Y+Y^*=0$. (This may also be seen by differentiating a curve of matrices satisfying (1).) Moreover, the orthogonal subspace $(T_{x_0}(X_{n,k}))^\perp$ is formed by the matrices $\begin{pmatrix} 0 \\ Z \end{pmatrix}$ such that $Z - Z^* = 0$. So we see that the projection of an arbitrary matrix $y = \begin{pmatrix} X \\ Y \end{pmatrix}\in\mathbb{H}^{n\times k}$ onto $T_{x_0}(X_{n,k})$ is given by

$$\pi_0(y) = \begin{pmatrix} X \\ (1/2)(Y - Y^*) \end{pmatrix}. \tag{3}$$

Now, let $x = \begin{pmatrix} T \\ P \end{pmatrix}\in X_{n,k}$ and $\pi\colon\mathbb{H}^{n\times k}\to T_x(X_{n,k})$ be the projection. The gradient of $f_\omega\colon X_{n,k}\to\mathbb{R}$ at x is the projection $\pi(\omega^*)$ of the gradient ω^* of h_ω. Since the action of $\mathrm{Sp}(n)$ on $X_{n,k}$ is transitive, there is some $A\in\mathrm{Sp}(n)$ such that $x = Ax_0$, so $T_x(X_{n,k}) = A\cdot T_{x_0}(X_{n,k})$. Thus, we may consider ω^* as a vector in the tangent space of $\mathbb{H}^{n\times k}$ at x, translate it via $A^* = A^{-1}$ to the tangent space of $\mathbb{H}^{n\times k}$ at x_0, where the projection is already determined by (3), and return to the tangent space at x via A. As the action is an isometry, we obtain

$$(\mathrm{grad} f_\omega)_x = A\pi_0(A^*\omega^*). \tag{4}$$

The matrix A can be written as $A = \begin{pmatrix} \alpha & T \\ \beta & P \end{pmatrix} \in \mathrm{Sp}(n)$, with blocks $\alpha \in \mathbb{H}^{(n-k)\times(n-k)}$ and $\beta \in \mathbb{H}^{k\times(n-k)}$, the condition $AA^* = I_n$ being equivalent to

$$\alpha\alpha^* + TT^* = I_{n-k},\ \beta\alpha^* + PT^* = 0,\ \beta\beta^* + PP^* = I_k. \tag{5}$$

We replace A, x, ω in (4), by their decomposition in blocks. Using (3) and (5) we obtain,

$$(\mathrm{grad} f)_{\begin{pmatrix} T \\ P \end{pmatrix}} = \begin{pmatrix} \delta^* - \frac{1}{2}T\big((\delta T + DP) + (\delta T + DP)^*\big) \\ \\ D^* - \frac{1}{2}P\big((\delta T + DP) + (\delta T + DP)^*\big) \end{pmatrix}, \tag{6}$$

which is the expression of the statement. □

The study of height functions can be simplified by using the singular value decompositions, see [21, Theorem 7.2]. Let $\omega^* = U \begin{pmatrix} 0 \\ S \end{pmatrix} V^*$ be a singular value decomposition of $\omega^* = \big(\delta \big| D\big)^*$. In this decomposition, we have $U \in \mathrm{Sp}(n)$, $V \in \mathrm{Sp}(k)$ and $S \in \mathbb{H}^{k\times k}$ is a block diagonal matrix

$$S = \begin{pmatrix} 0_{n_0} & & & \\ & s_1 I_{n_1} & & \\ & & \ddots & \\ & & & s_r I_{n_r} \end{pmatrix}, \quad n_0 + n_1 + \cdots + n_r = k, \tag{7}$$

with real numbers $0 < s_1 < \cdots < s_r$. The next result reduces the determination of the gradient and the Hessian to the case $\omega = \omega_0$.

Corollary 2.2. *Let f_ω and f_{ω_0} be the height functions on $X_{n,k}$ with respect to $\omega = \big(\delta\big|D\big)$ and $\omega_0 = \big(0\big|S\big)$. Then the gradient and Hessian of f_ω are related to those of f_{ω_0} as follows.*

a) *For any $x \in X_{n,k}$, we have $(\mathrm{grad} f_\omega)_x = U(\mathrm{grad} f_{\omega_0})_{U^*xV}V^*$.*

b) *For any $x \in X_{n,k}$ and $W \in T_x(X_{n,k})$, we have*

$$(\mathcal{H}f_\omega)_x(W) = U\big((\mathcal{H}f_{\omega_0})_{U^*xV}(U^*WV)\big)V^*.$$

Proof. The equality in a) follows from Proposition 2.1. Using this equality, the Hessian, $(\mathcal{H}f_\omega)_x \colon T_x(X_{n,k}) \to T_x(X_{n,k})$ can be computed as

$$\begin{aligned} (\mathcal{H}f_\omega)_x(W) &= \left(\frac{d}{dt}(\mathrm{grad} f_\omega)_{x+tW}\right)_{|t=0} \\ &= U\left(\frac{d}{dt}(\mathrm{grad} f_{\omega_0})_{U^*(x+tW)V}\right)_{|t=0} V^* \\ &= U\big((\mathcal{H}f_{\omega_0})_{U^*xV}(U^*WV)\big)V^*. \end{aligned}$$

□

2.4. **Critical points.** We now characterize the critical points of the height function on $X_{n,k}$. Let $x = \begin{pmatrix} T \\ P \end{pmatrix} \in X_{n,k}$ and $\omega = \begin{pmatrix} \delta | D \end{pmatrix}$, with $\delta \in \mathbb{H}^{k\times(n-k)}$ and $D \in \mathbb{H}^{k\times k}$ as above.

Proposition 2.3. *The point x is critical for the height function, $f_\omega \colon X_{n,k} \to \mathbb{R}$ if, and only if, $\omega^* = x\omega x$. In this case, the matrix $\omega x = \delta T + DP$ is Hermitian.*

Proof. From Proposition 2.1 we deduce that x is a critical point of f_ω if, and only if, $2\omega^* = x(\omega x + x^*\omega^*)$. Since $x^*x = I_k$, we have $2x^*\omega^* = \omega x + x^*\omega^*$, then $(\omega x)^* = \omega x$. Moreover $2\omega^* = x(2\omega x)$, hence $\omega^* = x\omega x$.

Conversely, if $\omega^* = x\omega x$, it is $x^*\omega^* = \omega x$, so ωx is Hermitian. Then, $x(\omega x + x^*\omega^*) = 2x\omega x = 2\omega^*$, that is, $(\mathrm{grad} f)_x = 0$. □

Corollary 2.4. *If the point $x = \begin{pmatrix} T \\ P \end{pmatrix}$ is critical for the height function, $f_\omega \colon X_{n,k} \to \mathbb{R}$, then the following formulae hold:*

$$\delta^* = T(\delta T + DP), \tag{8}$$

$$D^* = P(\delta T + DP), \tag{9}$$

$$\delta\delta^* + DD^* = (\delta T + DP)^2. \tag{10}$$

Proof. The first two formulae follow from Equation (6) and the fact that $\omega x = \delta T + DP$ is Hermitian, as proved in Proposition 2.3. The third one is $\omega\omega^* = (\omega x)^2$, which is also an immediate consequence of Proposition 2.3. □

In the case $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$, Proposition 2.3 can be simplified as follows.

Corollary 2.5. *Let $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$ be as in (7) and $\omega = U\omega_0 V^*$, with $U \in \mathrm{Sp}(n)$, $V \in \mathrm{Sp}(k)$. Then, the following properties hold.*

a) *The point $x = \begin{pmatrix} T \\ P \end{pmatrix}$ is critical for the height function, f_{ω_0}, if, and only if, $TSP = 0$ and $PSP = S$. In this case, the matrix SP is Hermitian.*

b) *The point x is a critical point of f_ω if, and only if, U^*xV is a critical point of f_{ω_0}.*

Proof. The statement a) is a rewriting of Proposition 2.3 in this particular case. Property b) is a direct consequence of Corollary 2.2.a. □

Recall that $\langle , \rangle$ denotes the Hermitian product in $\mathbb{H}^n$. Even if the matrix ω is not a square matrix, we have $\langle \omega\omega^* v, v\rangle = \langle \omega^* v, \omega^* v\rangle = |\omega v|^2 \geq 0$, for all $v \in \mathbb{H}^k$. Therefore, the matrix

$$\Delta = \omega\omega^* = \delta\delta^* + DD^*$$

is *positive semi-definite.*

Proposition 2.6. *The point x is critical for the height function, $f_\omega \colon X_{n,k} \to \mathbb{R}$ if, and only if, there exists a Hermitian square root Y of Δ such that $\omega^* = xY$. In this case $Y = \omega x$.*

Proof. First, let $\omega^* = xY$ with $Y = Y^*$ and $Y^2 = \Delta$. Then $x\omega x = x(Y^*x^*)x = xY(x^*x) = xY = \omega^*$, so x is a critical point by Proposition 2.3. Moreover from $x^*x = I_k$ it follows $x^*\omega^* = Y$, so $Y = Y^* = \omega x$.

Conversely, if x is a critical point, then $\omega^* = x(\omega x)$, and the matrix $Y = \omega x$ is Hermitian. Moreover $(\omega x)^2 = (\omega x)(\omega x)^* = \omega(x^*x)\omega^* = \omega\omega^* = \Delta$. □

Proposition 2.7. *Let $\omega_0 = \left(0 \middle| S\right)$, $x = \begin{pmatrix} T \\ P \end{pmatrix} \in X_{n,k}$, $x_0 = \begin{pmatrix} 0 \\ I_k \end{pmatrix}$ and $W \in T_x X_{n,k}$. Let $A \in \mathrm{Sp}(n)$ such that $x = Ax_0$ and $W = AW_0$, with $W_0 = \begin{pmatrix} X \\ Y \end{pmatrix} \in T_{x_0} X_{n,k}$. Then, if x is a critical point of f_{ω_0}, the Hessian of f_{ω_0} is given by,*

$$(\mathcal{H} f_{\omega_0})_x(W) = -(1/2)A \begin{pmatrix} 2XSP \\ YSP + SPY \end{pmatrix}.$$

Proof. We first observe that the existence of A comes from the transitivity of the action of $\mathrm{Sp}(n)$. For the determination of the Hessian, we follow a classical procedure (see [5]) which consists of three steps.

(i) First, we extend the gradient computed in Proposition 2.1 to a vector field on $\mathbb{H}^{n\times k}$, by

$$(\widetilde{\mathrm{grad}} f_{\omega_0})_M = \omega_0^* - (1/2)M(\omega_0 M + M^*\omega_0^*).$$

(ii) Secondly, we determine the covariant derivative $\nabla_W(\widetilde{\mathrm{grad}} f_{\omega_0})$. Since the covariant derivative in $\mathbb{H}^{n\times k}$ is the usual derivative, we find after computation,

$$\left(\frac{d}{dt}(\widetilde{\mathrm{grad}} f_{\omega_0})_{M+tW}\right)_{|t=0} = (-1/2)\left[M(\omega_0 W + W^*\omega_0^*) + W(\omega_0 M + M^*\omega_0^*)\right]. \quad (11)$$

(iii) Finally, the Hessian consists of the projection of the expression (11) onto the tangent space, $T_x X_{n,k}$. We denote $A = \begin{pmatrix} \alpha & T \\ \beta & P \end{pmatrix} \in \mathrm{Sp}(n)$ and recall that in $W_0 = \begin{pmatrix} X \\ Y \end{pmatrix} \in T_{x_0} X_{n,k}$, the matrix X is arbitrary and Y is anti-hermitian, i.e., $Y + Y^* = 0$. Set

$$(\Gamma_{\omega_0})_x(W) = \left(\frac{d}{dt}(\widetilde{\mathrm{grad}} f_{\omega_0})_{M+tW}\right)_{|t=0}.$$

Replacing M by x and ω_0, W by their values in (11) gives,

$$(\Gamma_{\omega_0})_x \begin{pmatrix} U \\ V \end{pmatrix} = -\frac{1}{2}\left\{ \begin{pmatrix} T \\ P \end{pmatrix} (SV + V^*S) + \begin{pmatrix} U \\ V \end{pmatrix} (SP + P^*S) \right\}.$$

So

$$\begin{aligned} A^* (\Gamma_{\omega_0})_x \begin{pmatrix} U \\ V \end{pmatrix} &= -\frac{1}{2}\left\{ \begin{pmatrix} 0 \\ I \end{pmatrix} (SV + V^*S) + \begin{pmatrix} X \\ Y \end{pmatrix} (SP + P^*S) \right\} \\ &= -\frac{1}{2}\left\{ \begin{pmatrix} 0 \\ SV + V^*S \end{pmatrix} + \begin{pmatrix} XSP + XP^*S \\ YSP + YP^*S \end{pmatrix} \right\}. \end{aligned}$$

As x is a critical point, we have $SP = P^*S$ and we obtain

$$A^* (\Gamma_{\omega_0})_x \begin{pmatrix} U \\ V \end{pmatrix} = -\frac{1}{2}\left\{ \begin{pmatrix} 0 \\ SV + V^*S \end{pmatrix} + \begin{pmatrix} 2XSP \\ 2YSP \end{pmatrix} \right\}.$$

So, the image by the projection onto $T_{x_0}X_{n,k}$, recalled in (3), equals,

$$-\frac{1}{2}\begin{pmatrix} 2XSP \\ YSP + SPY \end{pmatrix}.$$

Finally, a left translation by A gives the projection onto $T_xX_{n,k}$ and we get the formula of the statement. □

2.5. **Critical sets.** In this paragraph, we characterize the critical sets of any height function on a quaternionic Stiefel manifold.

Theorem 2.8. *The critical set of the height function $f_\omega \colon X_{n,k} \to \mathbb{R}$ is diffeomorphic to the product*

$$X_{n-k+n_0,n_0} \times \Sigma(n_1) \times \cdots \times \Sigma(n_r),$$

where $X_{n-k+n_0,n_0} = \mathrm{Sp}(n-k+n_0)/\mathrm{Sp}(n-k)$ is a Stiefel manifold and $\Sigma(m)$ denotes a disjoint union $G_{0,m} \sqcup G_{1,m} \sqcup \cdots \sqcup G_{m,m}$ of Grassmannians, $G_{i,m} = \mathrm{Sp}(m)/(\mathrm{Sp}(i) \times \mathrm{Sp}(m-i))$.

Proof of Theorem 2.8. According to Corollary 2.5, it is sufficient to consider the case f_{ω_0} with $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$ and $S = \mathrm{diag}(0_{n_0}, s_1 I_{n_1}, \ldots, s_r I_{n_r})$ a non-negative real diagonal block matrix. From Corollary 2.5 we know that $x = \begin{pmatrix} T \\ P \end{pmatrix}$ is a critical point of f_{ω_0} if, and only if, $TSP = 0$ and $PSP = S$. Moreover, we have also that SP is Hermitian and $S^2 = (SP)^2$.

The latter equality implies that S is the modulus of SP and [2, Theorem 5.5] implies the existence of $U \in \mathrm{Sp}(k)$ such that $SP = US$. As $SP = US$ is Hermitian, we get $SU = U^*S$, hence

$$S = USU.$$

A direct computation (see [12, Lemma 6]) implies that the unitary matrix U is of the form $U = \mathrm{diag}(U_0, U_1, \ldots, U_r)$ with $U_i U_i^* = I_{n_i}$ for $0 \le i \le k$, and $U_i = U_i^*$ for $1 \le i \le k$.

Now, we decompose T in $T = (T_0, \ldots, T_r)$, where T_j is a block of size $(n-k) \times n_j$. The equality $TSP = 0$ implies $s_j T_j U_j = 0$ and thus $T_j = 0$ for $j \ge 1$.

Secondly, we decompose P in $P = \begin{pmatrix} P_{00} & P_{01} & \ldots & P_{0r} \\ P_{10} & P_{11} & \ldots & P_{1r} \\ \vdots & & & \\ P_{r0} & P_{r1} & \ldots & P_{rr} \end{pmatrix}$, where P_{ij} is a block of size $n_i \times n_j$. The equality $S = PSP = PUS$ gives $P_{ij} = 0$ if $i \neq j$ and $P_{ii} = U_i$ for $1 \le i \le r$.

In conclusion, $\begin{pmatrix} T \\ P \end{pmatrix}$ has for its first columns $\begin{pmatrix} T_0 \\ P_{00} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and the other columns are of the form

$$\begin{pmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ U_1 & & \\ & \ddots & \\ & & U_r \end{pmatrix}.$$

The equality $T^*T + P^*P = I_k$ implies $T_0^* T_0 + P_{00}^* P_{00} = I_{n_0}$, so these first columns represent an element of the Stiefel manifold X_{n-k+n_0, n_0}. Finally each Hermitian matrix $U_i \in \mathrm{Sp}(n_i)$ verifies $U_i^2 = I_{n_i}$, so it can be written as $U_i = V D_i V^*$, where D_i is a diagonal matrix with ± 1 on the diagonal. The orbit for the action of $\mathrm{Sp}(n_i)$ by conjugation of such D_i is diffeomorphic to some Grassmannian, $\mathrm{Sp}(n_i)/(\mathrm{Sp}(l) \times \mathrm{Sp}(n_i - l))$. □

2.6. **Height functions as Morse functions: Indices and critical values.** By definition, a function on a compact manifold is Morse-Bott if the Hessian is non-degenerate in the directions transverse to the critical set. Moreover, such a function is Morse if it has a finite number of critical points. The next result characterizes the height functions that are Morse on a quaternionic Stiefel manifold. (Note that the case of real Stiefel manifolds is developed in [18, Theorem 1.2].)

Theorem 2.9. *Height functions, $f_\omega \colon X_{n,k} \to \mathbb{R}$, on quaternionic Stiefel manifolds satisfy the following properties.*

a) *Any height function is Morse-Bott.*

b) *A height function, f_ω, is a Morse function if, and only if, the singular values of $\omega\omega^*$ are positive and pairwise different. In this case, the critical set is diffeomorphic to the set of points* $\begin{pmatrix} 0 \\ E \end{pmatrix}$ *where* $E = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_k) \in \operatorname{Sp}(k)$*, with* $\varepsilon_i = \pm 1$.

Proof. Property b) is a direct consequence of a) and Theorem 2.8. We begin the proof of Property a) for the height function f_{ω_0}, with $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$, see (7). Observe that the definition of Morse-Bott means that the kernel of the Hessian equals the tangent space to the critical set. Let $A \in \operatorname{Sp}(n)$ such that $x = Ax_0$ with $x_0 = \begin{pmatrix} 0 \\ I \end{pmatrix}$ and x a critical point.

The *tangent space to the critical set* Σ can be computed as usual (here, the coordinates correspond to the lines of the matrix) from the equations established in Corollary 2.5.a, i.e., we have

$$W = \begin{pmatrix} U \\ V \end{pmatrix} \in T_x\Sigma \text{ if, and only if, } \begin{cases} USP &= 0 \\ VSP + PSV &= 0. \end{cases} \tag{12}$$

Let $W = \begin{pmatrix} U \\ V \end{pmatrix} = AW_0$ with $W_0 = \begin{pmatrix} X \\ Y \end{pmatrix}$. According to Proposition 2.7, the *kernel of the Hessian of* f_{ω_0} is characterized by,

$$W = \begin{pmatrix} U \\ V \end{pmatrix} \in \ker(\mathcal{H} f_{\omega_0})_x \text{ if, and only if, } \begin{cases} XSP &= 0 \\ YSP + SPY &= 0. \end{cases} \tag{13}$$

We observe

$$\begin{aligned} \begin{cases} XSP &= 0 \\ YSP + SPY &= 0 \end{cases} &\iff \begin{pmatrix} X \\ Y \end{pmatrix} SP + \begin{pmatrix} 0 \\ SPY \end{pmatrix} = 0 \\ &\iff A\begin{pmatrix} X \\ Y \end{pmatrix} SP + A\begin{pmatrix} 0 \\ SPY \end{pmatrix} = 0. \\ &\iff \begin{pmatrix} U \\ V \end{pmatrix} SP + A\begin{pmatrix} 0 \\ SPY \end{pmatrix} = 0. \end{aligned}$$

So the equivalence of the systems (12) and (13) is a consequence of

$$\textit{Claim: } A\begin{pmatrix} 0 \\ SPY \end{pmatrix} = \begin{pmatrix} 0 \\ PSV \end{pmatrix}. \tag{14}$$

To establish this equality, we first deduce from Theorem 2.8 that

$$\begin{pmatrix} T \\ P \end{pmatrix} = \begin{pmatrix} T_0 & 0 \\ P_{00} & 0 \\ 0 & U_{1\ldots r} \end{pmatrix},$$

where $U_{1\dots r} = \mathrm{diag}(U_1, \dots, U_r)$ with $U_i \in \mathrm{Sp}(n_i)$. We decompose the matrix A as

$$A = \begin{pmatrix} \alpha & T_0 & 0 \\ \beta_0 & P_{00} & 0 \\ \beta_{1\dots r} & 0 & U_{1\dots r} \end{pmatrix}$$

with $\beta_0 \in \mathbb{H}^{n_0 \times (n_0-k)}$, $\beta_{1,\dots,r} \in \mathbb{H}^{(k-n_0)\times(n-k)}$. The equality $AA^* = I_n$ implies $\beta_{1\dots r}U^*_{1\dots r} = 0$, so $\beta_{1\dots r} = 0$. Finally we have

$$S\beta = \begin{pmatrix} 0 & 0 & \dots & 0 \\ & s_1 I & & \\ & & \ddots & \\ & & & s_r I \end{pmatrix} \begin{pmatrix} \beta_{00} & \beta_{01} & \dots & \beta_{0r} \\ 0 & 0 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & 0 \end{pmatrix} = 0.$$

Now, we compute, by using $TSP = 0$ and $PSP = S$, see Corollary 2.5.a,

$$A \begin{pmatrix} 0 \\ SPY \end{pmatrix} = \begin{pmatrix} \alpha & T \\ \beta & P \end{pmatrix} \begin{pmatrix} 0 \\ SPY \end{pmatrix} = \begin{pmatrix} TSPY \\ PSPY \end{pmatrix} = \begin{pmatrix} 0 \\ SY \end{pmatrix}.$$

From $\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \alpha & T \\ \beta & P \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$, we get $V = \beta X + PY$ and

$$PSV = PS\beta X + PSPY = PSPY = SY.$$

The claim (14) is proved and the critical point x is non-degenerate.

So, Property a) is proved for f_{ω_0}. The general case of f_ω follows from the transformation rule of the Hessian (Corollary 2.2.b) and the determination of critical points (see Corollary 2.5.b). □

If we denote by α_i the number of critical points of index i and by b_i the i-th Betti number of $X_{n,k}$, we know, from classical Morse theory, that $\alpha_i \geq b_i$, for all i. A Morse function is called *perfect* if we have $\alpha_i = b_i$, for all i.

Corollary 2.10. *Each Morse height function on $X_{n,k}$ is a perfect Morse function.*

Proof. From their description in Theorem 2.9, it is clear that a Morse height function has exactly 2^k critical points. On the other side, we know [14, Theorem 3.10] that, for $k > 0$, the cohomology of $X_{n,k}$, with $\mathbb{Z}$ coefficients, is an exterior algebra,

$$H^*(X_{n,k}) = \wedge(y_{n-k+1}, \dots, y_n), \tag{15}$$

with y_i of degree $4i-1$. Its Poincaré series is $P(t) = (1+t^{4(n-k)+3}) \cdots (1+t^{4n-1})$. The sum of coefficients is 2^k and we get $\sum_{i\geq 0} \alpha_i = \sum_{i \geq 0} b_i$. With the Morse inequalities, $\alpha_i \geq b_i$, this implies $b_i = \alpha_i$, for all i. □

We have proved the previous Corollary without the determination of the index of a critical point. The next result specifies this index.

Proposition 2.11. *Let* $x = \begin{pmatrix} 0 \\ E \end{pmatrix}$ *be a critical point of a Morse height function, with* $E = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_k)$, $\varepsilon_i = \pm 1$. *Suppose that* $\varepsilon_{i_1}, \ldots, \varepsilon_{i_p}$ *are the positive entries in* E, *with* $i_1 < \cdots < i_p$. *Then the index of* x *is given by*

$$\operatorname{Ind}(x) = p(4(n-k) - 1) + 4(i_1 + \cdots + i_p).$$

Proof. According to Theorem 2.9, the matrix defining the height function can be assumed to have the form $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$, with $S = \operatorname{diag}(s_1, \ldots, s_k)$ and $0 < s_1 < \cdots < s_k$. A critical point of the associated height function, f_{ω_0}, is of the form $x = \begin{pmatrix} 0 \\ \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_k) \end{pmatrix}$, with $\varepsilon_i = \pm 1$.

We refer to Proposition 2.7. The point $x = \begin{pmatrix} 0 \\ E \end{pmatrix}$ is the image of $x_0 = \begin{pmatrix} 0 \\ I \end{pmatrix}$ by the left action of $\operatorname{Sp}(n)$, so there is some A such that $Ax = x_0$. Since the isotropy of x_0 is $\operatorname{Sp}(n-k)$, the matrix A is not unique but we can choose $A = \begin{pmatrix} I & 0 \\ 0 & E \end{pmatrix} \in \operatorname{Sp}(n)$ and we have

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & E \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X \\ EY \end{pmatrix}.$$

Therefore, the result of Proposition 2.7 can be written as

$$\begin{aligned} (\mathcal{H} f_{\omega_0})_x \begin{pmatrix} U \\ V \end{pmatrix} &= -\frac{1}{2} \begin{pmatrix} I & 0 \\ 0 & E \end{pmatrix} \begin{pmatrix} 2XSE \\ (YSE + SEY) \end{pmatrix} \\ &= -\frac{1}{2} \begin{pmatrix} 2XSE \\ EYSE + ESEY \end{pmatrix} \\ &= -\frac{1}{2} \begin{pmatrix} 2USE \\ VSE + ESV \end{pmatrix} \end{aligned}$$

The Hessian is a self-adjoint linear map, whose eigenvalues are of two types:

(i) $-\varepsilon_j s_j$, for $1 \leq j \leq k$. The corresponding eigenvectors are the matrices all of whose terms are zero, except the j^{th}-column of U and the element v_{jj} of V. As EV is skew-hermitian, the associated eigenspace is of (real) dimension $4(n-k) + 3$.

(ii) $(-1/2)(\varepsilon_i s_i + \varepsilon_j s_j)$, with $1 \leq i < j \leq k$. The corresponding eigenvectors are the matrices all of whose terms are zero, except v_{ij} and $v_{ji} = -v_{ij}^*$. The associated eigenspace is of (real) dimension 4.

Recall that the index of the critical point $x = \begin{pmatrix} 0 \\ E \end{pmatrix}$, with $E = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_k)$, is the total dimension of the eigenspaces associated to negative eigenvalues. Let $\varepsilon_{i_1}, \ldots, \varepsilon_{i_p}$ be the positive entries, with $i_1 < \cdots < i_p$. In case (i), negative eigenvalues are exactly the ε_{i_k}'s, with $k = 1, \ldots, p$. In case (ii), they are the pairs (i, j), with $i < j$ and $\varepsilon_j = +1$. In conclusion we get the formula

$$\operatorname{Ind}(x) = p(4(n-k)+3) + 4((i_1-1)+\cdots+(i_p-1)),$$

which is equivalent to that of the statement. □

Finally, we give a lower bound for the number of critical *values* of a Morse height function which is of interest in the determination of the LS-category.

Proposition 2.12. *The number of critical values of a Morse height function on $X_{n,k}$ is at least $1+\binom{k+1}{2}$. Moreover, there exists a height function for which the bound is sharp.*

Proof. Let $\omega_0 = \begin{pmatrix} 0 | S \end{pmatrix}$, with $S = \operatorname{diag}(s_1, \ldots, s_k)$, $0 < s_1 < \cdots < s_k$. The value of f_{ω_0} at the critical point $x = \begin{pmatrix} 0 \\ E \end{pmatrix}$, with $E = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_k)$, $\varepsilon_i = \pm 1$, is $f_{\omega_0}(x) = \sum_{i=0}^k \varepsilon_i s_i$. Therefore, the number of critical values is equal to the cardinality of the set $\left\{\sum_{i=0}^k \varepsilon_i s_i \mid \varepsilon_i = \pm 1\right\}$. We observe that

$$\begin{aligned}
\#\left\{\sum_{i=0}^k \varepsilon_i s_i \mid \varepsilon_i = \pm 1\right\} &= \#\left\{\sum_{i=0}^k \varepsilon_i s_i + s_1 + \cdots + s_k \mid \varepsilon_i = \pm 1\right\} \\
&= \#\left\{\sum_{i=0}^k f_i s_i \mid f_i \in \{0, 2\}\right\} \\
&= \#\left\{\sum_{i=0}^k g_i t_i \mid g_i \in \{0, 1\}\right\},
\end{aligned}$$

with $t_i = 2s_i$. The latter set has a cardinality greater than, or equal to, $1+\binom{k+1}{2}$ because it contains the following distinct elements:

$0, t_1, \ldots, t_k, t_k + t_1, \ldots, t_k + t_{k-1}, t_k + t_{k-1} + t_1, \ldots, t_k + t_{k-1} + t_{k-2}, \ldots, t_k + \cdots + t_1.$

This lower bound is reached in the case $s_i = i$, as a direct computation shows (see [11]). □

3. LS category of Stiefel manifolds

3.1. **Properties of LS category.** The definition of the LS-category of a topological space, X, has been recalled in the introduction. We list here some basic properties of it, referring to [1] for more details.

(1) The LS-category is a homotopy type invariant.

(2) By definition, the *cup length* of a space X is the largest integer ℓ such that there exists a product $x_1 \cdots x_\ell \neq 0$, with $x_i \in \tilde{H}^*(X;A)$. Here the coefficient ring A may vary and the cup length may be considered for any coefficients. Then

$$\operatorname{cup}(X) \leq \operatorname{cat} X.$$

For instance, from Equation (15), it appears that the cup length of $H(X_{n,k};\mathbb{Z})$ equals k.

(3) Let X be an $(n-1)$-connected CW-complex, then

$$\operatorname{cat} X \leq (\dim X)/n.$$

(4) If M is a smooth compact manifold and $\operatorname{crit}(M)$ denotes the minimum number of critical points for any smooth function $f\colon M \to \mathbb{R}$, then

$$\operatorname{cat} M + 1 \leq \operatorname{crit}(M).$$

In fact, we shall use a refined version of property (4) which was observed in [16].

Theorem 3.1. *Let M be a connected compact manifold and let $f\colon M \to \mathbb{R}$ be a smooth function with isolated critical points. Then*

$$\operatorname{cat} M + 1 \leq \#\{\text{critical values of } f\}.$$

This result is the reason we are interested in finding Morse functions with as few critical *values* as possible. In the opposite way, we know (see [13, Lemma 2.8, page 17]) that any Morse function with critical points, $x_1, \dots, x_\ell$, can be approached by a Morse function, g, with the same critical points and such that $g(x_i) \neq g(x_j)$ if $i \neq j$.

3.2. **LS category of** $\operatorname{Sp}(n)$**.** The case of $\operatorname{Sp}(n)$, which corresponds to $k = n$, was already studied in [11]. It was shown there that, when $\omega = D = \operatorname{diag}(1, 2, \dots, n)$ the height function f_ω has $\binom{n+1}{2} + 1$ different critical values. Then, by Theorem 3.1,

$$\operatorname{cat} \operatorname{Sp}(n) \leq \binom{n+1}{2}.$$

3.3. **The Stiefel manifolds** $X_{n,k}$**.** In Proposition 2.12, we proved that, for any Morse height function on $X_{n,k}$, the cardinality of the set of different critical values is always greater than, or equal to, $\binom{k+1}{2} + 1$. Theorem 3.1 implies

$$\operatorname{cat} X_{n,k} \leq \binom{k+1}{2}. \tag{16}$$

Recall from [8, Proposition 2.1, Page 15], that $X_{n,k}$ is of dimension $2k(2n-k+1) - k = 4nk - 2k^2 + k$ and, of connectivity plus 1 equal to $4n - 4k + 3$. We want to compare the bound given by the dimension and connectivity to (16). For that, we study the sign of the difference

$$\begin{aligned} A(n,k) &= \frac{4nk-2k^2+k}{4n-4k+3} - \frac{k(k+1)}{2} \\ &= \frac{4nk-4nk^2+4k^3-3k^2-k}{8n-8k+6}. \end{aligned}$$

As $0 \leq k \leq n$, the sign of $A(n,k)$ is the same as the sign of

$$\begin{aligned} B(n,k) &= 4k^2 - k(4n+3) + 4n - 1 \\ &= 4(k-1)(k-(n-(1/4))). \end{aligned}$$

This quadratic polynomial in k has two zeros, $k_1 = 1$ and $k_2 = n - (1/4)$. Thus $B(n,k)$ is only positive when $k = n$ and this is the case of $\mathrm{Sp}(n)$ already studied in [11].

Acknowledgements

This research was supported through the program "Research in Pairs" at the Mathematisches Forschungsinstitut Oberwolfach in 2013. The authors thank the MFO for its generosity and hospitality. The first author was also supported by a European Science Foundation ACAT travel grant. The second author was partially supported by a grant from the Simons Foundation (#244393 to John Oprea). The fourth author is partially supported by the MICINN grant MTM2013-41768-P, ANR-11-BS01-002-01 "HOGT" and ANR-11-LABX-0007-01 "CEMPI".

References

[1] Octav Cornea, Gregory Lupton, John Oprea, and Daniel Tanré, *Lusternik-Schnirelmann category*, Mathematical Surveys and Monographs, vol. 103, American Mathematical Society, Providence, RI, 2003. MR 1990857 (2004e:55001)

[2] Douglas R. Farenick and Barbara A. F. Pidkowich, *The spectral theorem in quaternions*, Linear Algebra Appl. **371** (2003), 75–102. MR 1997364 (2004m:15026)

[3] Lucía Fernández-Suárez, Antonio Gómez-Tato, Jeffrey Strom, and Daniel Tanré, *The Lusternik-Schnirelmann category of* Sp(3), Proc. Amer. Math. Soc. **132** (2004), no. 2, 587–595. MR 2022385 (2004m:55005)

[4] Theodore Frankel, *Critical submanifolds of the classical groups and Stiefel manifolds*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 37–53. MR 0196766 (33 #4952)

[5] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine, *Riemannian geometry*, third ed., Universitext, Springer-Verlag, Berlin, 2004. MR 2088027 (2005e:53001)

[6] A. Gómez-Tato, E. Macías-Virgós, and M. J. Pereira-Sáez, *Trace map, Cayley transform and LS category of Lie groups*, Ann. Global Anal. Geom. **39** (2011), no. 3, 325–335. MR 2769302 (2012d:22011)

[7] Norio Iwase and Mamoru Mimura, *L-S categories of simply-connected compact simple Lie groups of low rank*, Categorical decomposition techniques in algebraic topology (Isle of Skye, 2001), Progr. Math., vol. 215, Birkhäuser, Basel, 2004, pp. 199–212. MR 2039767 (2005a:55001)

[8] I. M. James, *The topology of Stiefel manifolds*, Cambridge University Press, Cambridge-New York-Melbourne, 1976, London Mathematical Society Lecture Note Series, No. 24. MR 0431239 (55 #4240)

[9] Hiroyuki Kadzisa and Mamoru Mimura, *Morse-Bott functions and the Lusternik-Schnirelmann category*, J. Fixed Point Theory Appl. **10** (2011), no. 1, 63–85. MR 2825740 (2012k:58016)

[10] E. Macías-Virgós and M. J. Pereira-Sáez, *Symplectic matrices with predetermined left eigenvalues*, Linear Algebra Appl. **432** (2010), no. 1, 347–350. MR 2566482 (2010m:15020)

[11] ______, *An upper bound for the Lusternik-Schnirelmann category of the symplectic group*, Math. Proc. Cambridge Philos. Soc. **155** (2013), no. 2, 271–276. MR 3091519

[12] ______, *Height functions on compact symmetric spaces*, Monatshefte für Mathematik (2014), 1–22.

[13] John Milnor, *Lectures on the h-cobordism theorem*, Notes by L. Siebenmann and J. Sondow, Princeton University Press, Princeton, N.J., 1965. MR 0190942 (32 #8352)

[14] Mamoru Mimura and Hirosi Toda, *Topology of Lie groups. I, II*, Translations of Mathematical Monographs, vol. 91, American Mathematical Society, Providence, RI, 1991, Translated from the 1978 Japanese edition by the authors. MR 1122592 (92h:55001)

[15] Tetsu Nishimoto, *On the Lusternik-Schnirelmann category of Stiefel manifolds*, Topology Appl. **154** (2007), no. 9, 1956–1960. MR 2319267 (2008a:55003)

[16] Yuli B. Rudyak and Felix Schlenk, *Lusternik-Schnirelmann theory for fixed points of maps*, Topol. Methods Nonlinear Anal. **21** (2003), no. 1, 171–194. MR 1980143 (2004d:55002)

[17] Paul A. Schweitzer, *Secondary cohomology operations induced by the diagonal mapping*, Topology **3** (1965), 337–355. MR 0182969 (32 #451)

[18] Anant R. Shastri and B. Subhash, *Linear Morse functions on orbit spaces*, J. Ramanujan Math. Soc. **25** (2010), no. 2, 147–167. MR 2642452 (2011j:58018)

[19] Wilhelm Singhof, *On the Lusternik-Schnirelmann category of Lie groups*, Math. Z. **145** (1975), no. 2, 111–116. MR 0391075 (52 #11897)

[20] A. P. Veselov and I. A. Dynnikov, *Integrable gradient flows and Morse theory*, Algebra i Analiz **8** (1996), no. 3, 78–103. MR 1402289 (97h:58036)

[21] Fuzhen Zhang, *Quaternions and matrices of quaternions*, Linear Algebra Appl. **251** (1997), 21–57. MR 1421264 (97h:15020)

Institute of Mathematics, Department of Geometry and Topology, University of Santiago de Compostela, 15782 Spain

E-mail address: quique.macias@usc.es

Department of Mathematics, Cleveland State University, Cleveland OH, 44115 USA

E-mail address: j.oprea@csuohio.edu

Department of Mathematics, Western Michigan University, Kalamazoo, MI, 49008-5200 USA

E-mail address: Jeff.Strom@wmich.edu

Département de Mathématiques, Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France

E-mail address: Daniel.Tanre@univ-lille1.fr