Topology and its Applications 265 (2019) 106807

Contents lists available at ScienceDirect

Topology and its Applications

www.elsevier.com/locate/topol

Morse-Bott functions on orthogonal groups

H.I. Bozma [a], W.D. Gillam [b], F. Öztürk [c,*]

[a] Department of Electrical and Electronics Engineering, Boğaziçi University, Turkey
[b] Department of Mathematics and Statistics, Florida International University, United States of America
[c] Department of Mathematics, Boğaziçi University, Turkey

ARTICLE INFO

Article history:
Received 10 December 2018
Received in revised form 14 June 2019
Accepted 1 July 2019
Available online 3 July 2019

Keywords:
Morse-Bott theory
Orthogonal groups
Perfect fillings of tables
Morse-Bott inequalities

ABSTRACT

We make a detailed study of various (quadratic and linear) Morse-Bott trace functions on the orthogonal groups $O(n)$. We describe the critical loci of the quadratic trace function $\mathrm{Tr}(AXBX^T)$ and determine their indices via perfect fillings of tables. We review the basic notions of Morse-Bott cohomology in a simple case where the set of critical points has two connected components. We then use these results to give a new Morse-theoretic computation of the mod 2 Betti numbers of $SO(n)$.

1. Introduction

This paper is devoted to a strictly Morse-theoretic study of various functions on the orthogonal groups $O(n)$. Many of our results surely generalize to the classical groups $U(n)$ and $Sp(n)$ by replacing the base field $\mathbb{R}$ with $\mathbb{C}$ or the quaternion skew-field and making the usual modifications (replace transposes with conjugate-transposes and traces with their real parts). For brevity and simplicity, however, we work strictly with the usual orthogonal groups and leave any generalizations to the interested reader.

We are essentially interested in two classes of functions. The first is the quadratic trace function $f_{A,B} : O(n) \to \mathbb{R}$ given as $f_{A,B}(X) = \mathrm{Tr}(AXBX^T)$, for some fixed, orthogonally diagonalizable $n \times n$ matrices A, B. In the case A and B are symmetric matrices, the question of finding the extrema of $f_{A,B}$ restricted to the signed permutation matrices is nothing but the well-known Quadratic Assignment Problem. The extrema problem of $f_{A,B}$ was worked out by von Neumann in 1937 [13]. In a rather general setting, the fact that $f_{A,B}$ is Morse-Bott was proven in Lie-theoretical terms in [4]. We give an alternative, self-contained proof of this result in Section 2 via a "tour-de-force" of linear algebra. Our analysis there yields a complete

* Corresponding author.
E-mail addresses: bozma@boun.edu.tr (H.I. Bozma), wdgillam@gmail.com (W.D. Gillam), ferit.ozturk@boun.edu.tr (F. Öztürk).

https://doi.org/10.1016/j.topol.2019.07.001
0166-8641/

description of the critical loci of $f_{A,B}$ and their indices. We show that the critical loci are quotients of products of orthogonal groups, the topology of which is explicitly determined by the combinatorial objects called the perfect fillings of tables with margins prescribed by the multiplicities of the eigenvalues of A and B. Moreover the index of each connected component of the critical locus can be computed via the corresponding perfect filling. In the special case when A and B are matrices with distinct eigenvalues, the indices are nothing but the "inversion numbers" of permutations.

Second, we study functions $O(n) \to \mathbb{R}$ obtained by restricting a linear function on the vector space of all $n \times n$ matrices. These functions have been studied in [3], [7], [9], [10] (with an interest in Lusternik–Schnirelmann category), [11] and [12]; in the last two, the authors determine which of those functions are Morse. There are two such functions to which we devote special attention (neither of which is *Morse* for general n). One such is the function $f(X) = \mathrm{Tr}(X)$, originally studied by T. Frankel in [6]. In Section 3 we give a self-contained derivation of Frankel's results which seems simpler to us than Frankel's original approach and which is more in the spirit of the rest of our paper. The gist of these results is that f is Morse-Bott and the critical locus of f is a disjoint union of Grassmannians. In the Appendix we also give a purely combinatorial discussion of some related results of Frankel.

The other linear function of particular interest is the function $f_{nn} : SO(n) \to \mathbb{R}$, $f_{nn}(X) = X_{nn}$ obtained by taking the lower right entry of X. (Any entry would do, but this is a convenient choice.) We show very easily in Section 5 that f_{nn} is Morse-Bott and that the critical locus of f_{nn} is a disjoint union of two copies of $SO(n-1)$. The methods of "Morse-Bott cohomology" (which we treat independently in Section 4 in an original manner catering to our situation) yield a long exact sequence relating the cohomology of $SO(n)$ to that of $SO(n-1)$. The novelty of the Morse-theoretic point of view we take is to interpret the connecting maps in this long exact sequence in Morse-theoretic terms. This allows us to show that these maps are zero with $\mathbb{F}_2$-coefficients. We thus obtain a recursive description of the mod 2 Betti numbers of $SO(n)$ which is easily solved to yield a simple combinatorial formula for these numbers. Although these numbers can be computed in a variety of ways, we believe our Morse-theoretic computation is quite simple and natural. (Together with our combinatorial results in the Appendix, Frankel's study also yields the same Betti number formulae, though that approach is considerably more complicated as it relies on knowing the Betti numbers of Grassmannians, as well as on non-Morse-theoretic results of E. E. Floyd.)

Acknowledgments. The work of the last author is partially supported by Boğaziçi University Research Project BAP-17B06P2.

2. Some Morse-Bott functions on orthogonal groups

Fix two symmetric (equivalently, orthogonally diagonalizable) $n \times n$ matrices A and B and consider the smooth function

$$f = f_{A,B} : O(n) \to \mathbb{R}, \qquad f(X) := \mathrm{Tr}(AXBX^T).$$

If $A' = Q^T AQ$, $B' = R^T BR$ are orthogonal conjugates of A and B, then, using conjugation-invariance of the trace, one sees that $f_{A',B'}(X) = f_{A,B}(QXR^T)$, hence $f_{A',B'}$ is just the composition of $f_{A,B}$ and the automorphism $X \mapsto QXR^T$ of $O(n)$. Therefore, with no real loss of generality, we will assume for the remainder of this section that

$$\begin{aligned} A &= \mathrm{Diag}(a_1 I_{m_1}, \dots, a_s I_{m_s}) \\ B &= \mathrm{Diag}(b_1 I_{n_1}, \dots, b_t I_{n_t}), \end{aligned} \tag{2.1}$$

with $a_i \neq a_k$ for $i \neq k$ and $b_j \neq b_l$ for $j \neq l$. The eigenvalue multiplicities m_i, n_j satisfy

$$\sum_{i=1}^{s} m_i = \sum_{j=1}^{t} n_j = n.$$

Throughout this section, i and k (resp. j and l) will always denote elements of the set $\{1, \dots, s\}$ (resp. $\{1, \dots, t\}$). We write $F := \{X \in O(n) : (Df)(X) = 0\}$ for the critical locus of f.

Lemma 2.1. *For a symmetric $n \times n$ matrix S and a diagonal $n \times n$ matrix A as above, the following are equivalent: (i) AS is symmetric. (ii) $AS = SA$. (iii) $S = \mathrm{Diag}(S_1, \dots, S_s)$ with each S_i a symmetric $m_i \times m_i$ matrix.*

The metric $\langle M, N \rangle := \mathrm{Tr}(M^T N)$ on $T_X O(n) = \{M : MX^T + XM^T = 0\}$ is the unique (up to scaling) bi-invariant Riemannian metric on O(n).

Lemma 2.2. *The gradient of $f = f_{A,B}$ at a point $X \in O(n)$ is given by*

$$(\nabla f)(X) = (AXBX^T - XBX^T A)X.$$

The following are equivalent:

(i) $X \in F$ (i.e. X is a critical point of f).
(ii) $AXBX^T$ is symmetric.
(iii) $AXBX^T = XBX^T A$
(iv) $XBX^T = \mathrm{Diag}(H_1, \dots, H_s)$ with each H_i a symmetric $m_i \times m_i$ matrix.

In fact, the derivative of the extended function at X is given by

$$(Df)(X)(M) = \mathrm{Tr}(AMBX^T + AXBM^T). \tag{2.2}$$

Reading off the gradient with respect to the metric and projecting that onto the tangent space $T_X O(n)$ (see e.g. [2, §4]) the first claim follows. The equivalence of the conditions is evident, after Lemma 2.1 with $S = XBX^T$.

Definition 2.3. A *perfect filling* with margins $(m_1, \dots, m_t; n_1, \dots, n_s)$ is an $s \times t$ matrix ϵ with entries ϵ_{ij} in $\mathbb{N} = \{0, 1, \dots\}$ satisfying

$$\sum_{j=1}^{t} \epsilon_{ij} = m_i \text{ for each } i \in \{1, \dots, s\} \tag{2.3}$$

$$\sum_{i=1}^{s} \epsilon_{ij} = n_j \text{ for each } j \in \{1, \dots, t\} \tag{2.4}$$

Example 2.4. If A and B have distinct eigenvalues, then all m_i and n_j are 1, $s = t = n$, and a perfect filling is an $n \times n$ permutation matrix.

Throughout, we set

$$O(\overline{m}) := \prod_{i=1}^{s} O(m_i), \quad O(\overline{n}) := \prod_{j=1}^{t} O(n_j), \quad O(\epsilon) := \prod_{i,j} O(\epsilon_{ij}).$$

Construction 2.5. Given $Q = (Q[1], \ldots, Q[s]) \in O(\overline{m})$, $R = (R[1], \ldots, R[t]) \in O(\overline{n})$, and a perfect filling ϵ, we construct an $n \times n$ matrix $X = \Phi_\epsilon(Q, R)$ as follows: We write the $m_i \times m_i$ matrix $Q[i]$ in block form

$$Q[i] = \begin{pmatrix} Q[i,1] & \cdots & Q[i,t] \end{pmatrix}, \tag{2.5}$$

with $Q[i,j]$ of size $m_i \times \epsilon_{ij}$. (This makes sense in light of (2.3).) Similarly, we write the $n_j \times n_j$ matrix $R[j]$ in block form

$$R[j] = \begin{pmatrix} R[1,j] & \cdots & R[s,j] \end{pmatrix}, \tag{2.6}$$

with $R[i,j]$ of size $n_j \times \epsilon_{ij}$. (This makes sense in light of (2.4).) We let $X[i,j]$ be the $m_i \times n_j$ matrix defined by

$$X[i,j] := Q[i,j]R[i,j]^T \tag{2.7}$$

and we define X to be the $n \times n$ matrix written in block form as

$$X = \begin{pmatrix} X[1,1] & \cdots & X[1,t] \\ \vdots & & \vdots \\ X[s,1] & \cdots & X[s,t] \end{pmatrix}. \tag{2.8}$$

The Lie group $O(\epsilon)$ acts (smoothly, on the right) on $O(\overline{m}) \times O(\overline{n})$ by setting

$$\begin{aligned} (Q,R) \cdot U &:= (Q \cdot U, R \cdot U) \\ Q \cdot U &:= (Q[1] \cdot U, \ldots, Q[s] \cdot U) \\ R \cdot U &:= (R[1] \cdot U, \ldots, R[t] \cdot U) \\ Q[i] \cdot U &:= \begin{pmatrix} Q[i,1]U[i,1] & \cdots & Q[i,t]U[i,t] \end{pmatrix} \\ R[j] \cdot U &:= \begin{pmatrix} R[1,j]U[1,j] & \cdots & R[s,j]U[s,j] \end{pmatrix} \end{aligned}$$

for $U = (U[i,j]) \in O(\epsilon)$. In other words:

$$\begin{aligned} (Q \cdot U)[i,j] &= Q[i,j]U[i,j] \\ (R \cdot U)[i,j] &= R[i,j]U[i,j]. \end{aligned} \tag{2.9}$$

From (2.9), (2.7), and (2.8), we find $\Phi_\epsilon(Q,R) = \Phi_\epsilon(Q \cdot U, R \cdot U)$.

Remark 2.6. Since the columns of $Q[i,j]$ are linearly independent, the action of $U[i,j]$ on $Q[i,j]$ is free. The action of $O(\epsilon)$ on $O(\overline{m}) \times O(\overline{n})$ is therefore a free, smooth action of a compact Lie group on a smooth, compact manifold. The quotient $(O(\overline{m}) \times O(\overline{n}))/O(\epsilon)$ therefore admits a unique smooth manifold structure for which the quotient map is submersive. The quotient is understood to have this smooth structure throughout.

Proposition 2.7. *For Q, R, ϵ, and $X = \Phi_\epsilon(Q,R)$ as in Construction 2.5:*

(i) $X \in O(n)$ and X is a critical point of $f : O(n) \to \mathbb{R}$.
(ii) $XBX^T = \mathrm{Diag}(H_1, \ldots, H_s)$, where H_i is the $m_i \times m_i$ symmetric matrix given by

$$H_i = \sum_{j=1}^{t} b_j Q[i,j]Q[i,j]^T.$$

(iii) The columns of $Q[i,j]$ form an orthonormal basis for the b_j-eigenspace of H_i. In particular, the dimension of this eigenspace is ϵ_{ij}.

Proof. The proof is basically an exercise in multiplying matrices in block form. The matrix $Q[i] \in O(m_i)$ satisfies $Q[i]Q[i]^T = I_{m_i}$ and $Q[i]^T Q[i] = I_{m_i}$. Writing these products in terms of the block form (2.5), we find

$$\sum_{j=1}^{t} Q[i,j]Q[i,j]^T = I_{m_i}, \qquad Q[i,j]^T Q[i,l] = \begin{cases} I_{\epsilon_{ij}}, & j = l \\ 0_{\epsilon_{ij}\times\epsilon_{il}}, & j \neq l \end{cases}. \tag{2.10}$$

Similarly writing $R[j] \in O(n_j)$ in the form (2.6), we find

$$\sum_{i=1}^{s} R[i,j]R[i,j]^T = I_{n_j}, \qquad R[i,j]^T R[k,j] = \begin{cases} I_{\epsilon_{ij}}, & i = k \\ 0_{\epsilon_{ij}\times\epsilon_{kj}}, & i \neq k. \end{cases}. \tag{2.11}$$

To see that $X \in O(n)$, we compute $X^T X = (M[j,l])_{1\le j,l\le t}$ where $M[j,l]$ is the $n_j \times n_l$ matrix given by $M[j,l] = \sum_{i=1}^{s} X[i,j]^T X[i,l]$. Using (2.7), (2.8), (2.10) and (2.11) above, we compute $M[j,l] = I_{n_j}$ when $j = l$ and equals $0_{n_j\times n_l}$ otherwise. This proves that $X \in O(n)$.

For (ii), we compute $XBX^T = (D[i,k])_{1\le i,k\le s}$ where $D[i,k]$ is the $m_i \times m_k$ matrix given by $D[i,k] = \sum_{i=1}^{s}\sum_{j=1}^{t} b_j X[i,j]X[k,j]^T$. Expanding this out using the definition (2.7) of the $X[i,j]$ and (2.10), we find

$$D[i,k] = \begin{cases} \sum_{j=1}^{t} b_j Q[i,j]Q[i,j]^T, & i = k \\ 0_{m_i\times m_k}, & i \neq k. \end{cases}$$

This proves (ii). We have (ii) with $X \in O(n)$ imply (i) by Lemma 2.2.

For (iii), we compute, using (ii) and (2.10):

$$H_i Q[i,j] = \sum_{l=1}^{t} b_l Q[i,l]Q[i,l]^T Q[i,j] = b_j Q[i,j].$$

This shows that the ϵ_{ij} orthonormal columns of $Q[i,j]$ are all in the b_j eigenspace of H_i, so this eigenspace has dimension $\geq \epsilon_{ij}$. Since this is true for each j, H_i is an $m_i \times m_i$ matrix, and (2.3) holds, this last inequality must actually be an equality for every j by basic linear algebra. □

Theorem 2.8. *Construction 2.5 yields a diffeomorphism*

$$\Phi = \coprod_{\epsilon} \Phi_\epsilon : \coprod_{\epsilon} (O(\overline{m}) \times O(\overline{n}))/O(\epsilon) \to F$$

onto the critical locus F of f (the coproduct is over perfect fillings ϵ).

Proof. Fix $X \in F$. Set $H := XBX^T$. Since $X \in F$,

$$H = \mathrm{Diag}(H_1, \ldots, H_s), \tag{2.12}$$

with H_i an $m_i \times m_i$ symmetric matrix (Lemma 2.2). Let $E_{ij} \subseteq \mathbb{R}^{m_i}$ be the b_j-eigenspace of H_i and ϵ_{ij} be its dimension.

Claim 1. *$\epsilon = (\epsilon_{ij})$ is a perfect filling.*

Fix any i. From the block form (2.12) of H we see that any eigenvalue of H_i must also be an eigenvalue of H. Since H is similar to B the eigenvalues of H are the b_j. Therefore the eigenvalues of H_i are *among* the b_j. Since the symmetric $m_i \times m_i$ matrix H_i is diagonalizable, the equality $\sum_j \epsilon_{ij} = m_i$ follows. Now fix any j. From the block form (2.12) of H we see that the b_j-eigenspace E_j of H is the direct sum, over i, of the E_{ij}. Since H is similar to B we have $\dim E_j = n_j$. The equality $\sum_i \epsilon_{ij} = n_j$ follows. This proves the claim.

Now *choose*, for each i, j, an $m_i \times \epsilon_{ij}$ matrix $Q[i,j]$ whose columns form an orthonormal basis for $E_{ij} \subseteq \mathbb{R}^{m_i}$. By definition of H (and the fact that $X \in O(n)$), we have

$$HX = XB. \tag{2.13}$$

Writing H, X, and B in the (respective) block forms (2.12), (2.8), (2.1) and expanding out (2.13), we find

$$H_i X[i,j] = b_j X[i,j]$$

for all i, j. Therefore each column of $X[i,j]$ is in E_{ij}. Since the columns of $Q[i,j]$ form a basis for E_{ij} there is a *unique* $n_j \times \epsilon_{ij}$ matrix $R[i,j]$ such that

$$X[i,j] = Q[i,j]R[i,j]^T. \tag{2.14}$$

Define $Q[i]$ and $R[j]$ from the $Q[i,j]$ and $R[i,j]$ using the usual block forms (2.5), (2.6).

Claim 2. *$Q[i] \in O(m_i)$ for each i and $R[j] \in O(n_j)$ for each j.*

The $Q[i,j]$ have orthonormal columns forming a basis for E_{ij}, so to show that $Q[i] \in O(m_i)$, it is enough to prove that $E_{ij} \perp E_{il}$ for $j \neq l$. This holds by basic linear algebra since E_{ij} and E_{il} are distinct eigenspaces of the *symmetric* matrix H_i. Since $Q[i] \in O(m_i)$ for each i, $X \in O(n)$, and we have the relationship (2.14), we have $R[j] \in O(n_j)$ by the *observation* made in the proof of Proposition 2.7(i). This proves Claim 2.

By Claims 1 and 2, Q, R, and ϵ constructed above are as in Construction 2.5; clearly we have $X = \Phi_\epsilon(Q, R)$ (compare (2.14) and (2.7)). Suppose we also have $X = \Phi_{\epsilon'}(Q', R')$ with Q', R', ϵ' as in Construction 2.5. If we define the $m_i \times m_i$ matrices H_i from X and B as at the beginning of this proof, then by Proposition 2.7(iii), the columns of $Q[i,j]$ form an orthonormal basis for the b_j-eigenspace $E_{ij} \subseteq \mathbb{R}^{m_i}$ of H_i, as do the columns of $Q'[i,j]$. In particular, $Q[i,j]$ and $Q'[i,j]$ must have the same number of columns, so we must have $\epsilon_{ij} = \epsilon'_{ij}$. Furthermore, there is a unique $U[i,j] \in O(\epsilon_{ij})$ with $Q'[i,j] = Q[i,j]U[i,j]$. Combining this with the fact that

$$X[i,j] = Q[i,j]R[i,j]^T = Q'[i,j]R'[i,j]^T$$

(formula (2.7) in Construction 2.5), we deduce

$$Q'[i,j]U[i,j]^T R[i,j]^T = Q'[i,j]R'[i,j]^T. \tag{2.15}$$

By (2.10) (with Q replaced by Q'), we can multiply (2.15) on the left by $Q'[i,j]^T$ to cancel the $Q'[i,j]$'s on the left of each side of (2.15), then transpose to find $R'[i,j] = R[i,j]U[i,j]$. This proves $(Q', R') = (Q, R) \cdot U$, where $U = (U[i,j]) \in O(\epsilon)$.

The above results (and Proposition 2.7(i)) demonstrate that Φ (which is clearly smooth) is bijective. It must therefore be a homeomorphism since its domain is compact and its codomain is Hausdorff. The domain of Φ is a smooth manifold (Remark 2.6), so to conclude that it is a diffeomorphism onto its image F, it remains only to prove that the derivative of each Φ_ϵ is injective when we view Φ_ϵ as a map $(O(\overline{m}) \times$

$O(\overline{n}))/O(\epsilon) \to O(n)$. Equivalently, if we view Φ_ϵ simply as a map $\Phi_\epsilon : O(\overline{m}) \times O(\overline{n}) \to O(n)$, then we must prove that the kernel of each derivative $(D\Phi_\epsilon)(Q, R)$ is precisely the tangent space (at (Q, R)) to the $O(\epsilon)$-orbit of (Q, R). The latter is the image of the Lie derivative (derivative at the identity) of the orbit map $U \mapsto (Q, R) \cdot U$ from $O(\epsilon)$ to $O(\overline{m}) \times O(\overline{n})$.

Now we compute these derivatives. For $(M, N) \in T_Q O(\overline{m}) \oplus T_R O(\overline{n})$, we find

$$(D\Phi_\epsilon)(Q,R)(M,N) = \begin{pmatrix} D[1,1] & \cdots & D[1,t] \\ \vdots & & \vdots \\ D[s,1] & \cdots & D[s,t] \end{pmatrix}, \tag{2.16}$$

where

$$D[i,j] = M[i,j]R[i,j]^T + Q[i,j]N[i,j]^T. \tag{2.17}$$

(We have broken the components $M[i]$ of M into blocks $M[i,j]$ of size $m_i \times \epsilon_{ij}$ in the same way we broke up the $Q[i]$ in (2.5). Similarly, we have broken the $N[j]$ into blocks $N[i,j]$ of size $n_j \times \epsilon_{ij}$, just as we broke up the $R[i]$ in (2.6).) The aforementioned Lie derivative takes

$$P = (P[i,j]) \in T_I O(\epsilon) = \bigoplus_{i,j} \mathfrak{so}(\epsilon_{ij})$$

to (M, N) where, when cut into blocks in the usual manner,

$$M[i,j] = Q[i,j]P[i,j] \tag{2.18}$$

$$N[i,j] = R[i,j]P[i,j]. \tag{2.19}$$

Given any $(M, N) \in T_Q O(\overline{m}) \oplus T_R O(\overline{n})$ (not necessarily in the kernel of $(D\Phi_\epsilon)(Q, R)$), we can define $\epsilon_{ij} \times \epsilon_{ij}$ matrices $P[i, j]$ by

$$P[i,j] := Q[i,j]^T M[i,j]. \tag{2.20}$$

Claim 3. *The matrices $P[i, j]$ in (2.20) are skew-symmetric.*

To see this, one can observe that the fact $M[i] \in T_{Q[i]}O(m_i)$ yields immediately to the equality $Q[i, l']^T M[i, l] + M[i, l']^T Q[i, l] = 0$. Setting $l = j$ and $l' = j$ here proves Claim 3. Similarly we note that $R[k', j]^T N[k, j] + N[k', j]^T R[k, j] = 0$.

Now, if (M, N) is actually in the kernel of $(D\Phi_\epsilon)(Q, R)$ (i.e. all the $D[i, j]$ in (2.17) are zero), then of course we claim that (M, N) is the Lie derivative evaluated at the $P[i, j]$ defined in (2.20). To see this, we need to establish (2.18) and (2.19) when the $P[i, j]$ are defined by (2.20). For (2.18), we just multiply (2.20) on the left by $Q[i, j]$ and use (2.10). For (2.19), we multiply $D[i, j] = 0$ (using formula (2.17) for $D[i, j]$) on the left by $Q[i, j]^T$ and use (2.10) and (2.20) to find $P[i, j]R[i, j]^T + N[i, j]^T = 0$. By Claim 3 we can substitute $-P[i, j]^T$ for $P[i, j]$ here to obtain (the transpose of) (2.19). □

Our next task is to describe the Hessian of f at a critical point $X \in F$. We identify $T_X O(n)$ with the space $\mathfrak{so}(n) = T_I O(n)$ of skew-symmetric matrices via the usual isomorphism

$$X : \mathfrak{so}(n) \to T_X O(n), \qquad E \mapsto XE, \tag{2.21}$$

thus we view the Hessian $H(f, X)$ of f at X as a quadratic form on $\mathfrak{so}(n)$.

Lemma 2.9. *Viewing the Hessian $H = H(f, X)$ of f at a critical point $X \in F$ as a quadratic form on $\mathfrak{so}(n)$ as above, it is given by*

$$H(E, N) = \mathrm{Tr}(AX[E, [N, B]]X^T) \tag{2.22}$$

for $E, N \in \mathfrak{so}(n)$, where $[U, V] := UV - VU$.[1]

For the sake of brevity we leave some details in the proof to the reader. The computation is similar to the one in [3, Lemma 3.4] and can also be obtained from general results in [2].

Proof. Trivialize the cotangent bundle of the Lie group $O(n)$ using the isomorphisms (2.21) and view df as a section of this cotangent bundle to obtain a map of smooth manifolds $df : O(n) \to T^*O(n) = O(n) \times \mathfrak{so}(n)^*$ whose first component is the identity (it is a section) and whose second component is given by

$$\begin{aligned} (df)^{ver} &: O(n) \to \mathfrak{so}(n)^* \\ (df)^{ver}(X)(N) &= \mathrm{Tr}(AX[N, B]X^T), \end{aligned} \tag{2.23}$$

as one can see by rewriting (2.2) using the isomorphism (2.21). Using the fact that a matrix $N \in \mathfrak{so}(n)$ is skew-symmetric, one sees that the derivative of $(df)^{ver}$ at $X \in O(n)$ can be written

$$\begin{aligned} (D(df)^{ver})(X) &: T_X O(n) \to \mathfrak{so}(n)^* \\ (D(df)^{ver})(X)(M)(N) &= \mathrm{Tr}(AM[N, B]X^T + AX[N, B]M^T). \end{aligned} \tag{2.24}$$

If we take into account the isomorphism (2.21), then (2.24) becomes the map

$$\begin{aligned} H(X) &: \mathfrak{so}(n) \to \mathfrak{so}(n)^* \\ H(X)(E)(N) &= \mathrm{Tr}(AXE[N, B]X^T + AX[N, B]E^T X^T). \end{aligned} \tag{2.25}$$

By following through any "coordinate-free" construction of the Hessian, one can verify that, when X is a critical point, (2.25) "is" (under the usual correspondence between quadratic forms on a vector space V and linear maps $V \to V^*$) the Hessian of f at X. The proof is completed by using the fact that $E \in \mathfrak{so}(n)$ is skew-symmetric to rewrite (2.25) as in (2.22). $\square$

Definition 2.10. For $1 \le p < q \le n$, let $\mathbb{E}(p, q) \in \mathfrak{so}(n)$ be the skew-symmetric $n \times n$ matrix whose (p, q) entry is -1, whose (q, p) entry is 1, and whose other entries are zero. We refer to the basis for $\mathfrak{so}(n)$ consisting of the $\mathbb{E}(p, q)$ as the *standard basis* for $\mathfrak{so}(n)$. For distinct $p, q \in \{1, \dots, n\}$, let $\mathbb{F}(p, q)$ be the symmetric $n \times n$ matrix whose (p, q) and (q, p) entries are 1 and whose other entries are zero. Note $\mathbb{F}(p, q) = \mathbb{F}(q, p)$. For $p \in \{1, \dots, n\}$, let $\mathbb{D}(p)$ be the $n \times n$ matrix whose (p, p) entry is 1 and whose other entries are zero. For $\sigma \in \mathfrak{S}_n$, let P_σ be the matrix whose ith column is column $\sigma(i)$ of the identity matrix.

For later use, we record the following formulae involving the matrices of Definition 2.10:

$$[\mathbb{E}(p, q), B] = (B_{pp} - B_{qq})\mathbb{F}(p, q) \tag{2.26}$$

[1] It would be incorrect to refer to [,] as the "Lie bracket on $\mathfrak{so}(n)$" because, in (2.22), we apply it to matrices that aren't in $\mathfrak{so}(n)$.

$$[\mathbb{E}(p,q),\mathbb{F}(u,v)] = \begin{cases} 0, & \{p,q\}\cap\{u,v\}=\emptyset \\ \mathbb{F}(q,v), & p=u, q\neq v \\ \mathbb{F}(q,u), & p=v, q\neq u \\ -\mathbb{F}(p,u), & q=v, p\neq u \\ -\mathbb{F}(p,v), & q=u, p\neq v \\ 2\mathbb{D}(q)-2\mathbb{D}(p), & \{p,q\}=\{u,v\} \end{cases} \tag{2.27}$$

$$P_\sigma \mathbb{D}(p) P_\sigma^T = \mathbb{D}(\sigma(p)) \tag{2.28}$$

$$P_\sigma \mathbb{F}(p,q) P_\sigma^T = \mathbb{F}(\sigma(p),\sigma(q)). \tag{2.29}$$

Our first application of these computations is to establish that f is Morse-Bott:

Theorem 2.11. *For $X \in F \subseteq O(n)$ and $M \in T_X O(n)$, the following are equivalent:*

(i) $M \in T_X F \subseteq T_X O(n)$.
(ii) $AMBX^T + AXBM^T$ *is symmetric.*
(iii) $MBX^T + XBM^T = \mathrm{Diag}(W_1,\dots,W_s)$ *with each W_i an $m_i \times m_i$ symmetric matrix.*
(iv) $AMBX^T + AXBM^T = MBX^TA + XBM^TA$.
(v) $BX^TAM + BM^TAX = X^TAMB + M^TAXB$.
(vi) $BX^TAM + BM^TAX$ *is symmetric.*
(vii) $X^TAM + M^TAX = \mathrm{Diag}(U_1,\dots,U_t)$ *with each U_j an $n_j \times n_j$ symmetric matrix.*
(viii) $M \in \operatorname{Ker} H(X)$.

In particular, the equivalence of (i) *and* (viii) *implies that f is Morse-Bott.*

Proof. According to Lemma 2.2, $F = \{X \in O(n) : AXBX^T = XBX^TA\}$. Therefore, a tangent vector $M \in T_X O(n)$ will be in $T_X F$ iff $A(X+\epsilon M)B(X^T+\epsilon M^T)$ (viewed as a matrix with entries in $\mathbb{R}[\epsilon]/\epsilon^2$) is symmetric. Since $AXBX^T$ is symmetric (because $X \in F$), this is equivalent to (ii). This proves the equivalence of (i) and (ii).

The conditions (ii), (iii), and (iv) are equivalent by Lemma 2.1, applied to the symmetric matrix $S = MBX^T + XBM^T$. Similarly, (v), (vi), and (vii) are equivalent by Lemma 2.1, applied to the symmetric matrix $X^TAM + M^TAX$ (and with A there given by B here).

Using the description of $H(X)$ in (2.24), we see that (viii) is equivalent to

$$\mathrm{Tr}(AM[\mathbb{E}(p,q),B]X^T + AX[\mathbb{E}(p,q),B]M^T) = 0$$

for $1 \le p < q \le n$. For any $n \times n$ matrices C and D, we compute

$$\begin{aligned}\mathrm{Tr}(C\mathbb{F}(p,q)D) &= \sum_{r=1}^n (C\mathbb{F}(p,q)D)_{rr} = \sum_{r=1}^n \sum_{u,v=1}^n C_{ru}\mathbb{F}(p,q)_{uv}D_{vr} \\ &= \sum_{r=1}^n C_{rp}D_{qr} + C_{rq}D_{pr} = (DC)_{qp} + (DC)_{pq}.\end{aligned}$$

Using this and (2.26) we compute

$$\begin{aligned}
&\operatorname{Tr}(AM[\mathbb{E}(p,q),B]X^T + AX[\mathbb{E}(p,q),B]M^T) \\
&= (B_{pp} - B_{qq})((X^T AM)_{pq} + (X^T AM)_{qp} + (M^T AX)_{pq} + (M^T AX)_{qp}) \\
&= 2(B_{pp} - B_{qq})(X^T AM + M^T AX)_{pq}.
\end{aligned}$$

This is zero for $1 \le p < q \le n$ iff $X^T AM + M^T AX$ has the block form in (vii) (the blocks must be symmetric since $X^T AM + M^T AX$ is symmetric). This proves the equivalence of (viii) and (vii).

It remains only to prove the equivalence of (iv) and (v). The equation in (v) is equivalent to the equation

$$XBX^T AMX^T + XBM^T A = AMBX^T + XM^T AXBX^T \tag{2.30}$$

obtained by multiplying on the left by X and on the right by X^T. Since $X \in F$, we have $AXBX^T = XBX^T A$ (Lemma 2.2), so we can rewrite (2.30) as

$$AXBX^T MX^T + XBM^T A = AMBX^T + XM^T XBX^T A. \tag{2.31}$$

Since $M \in T_X O(n)$ we have $X^T M = -M^T X$, so we can rewrite (2.31) as

$$-AXBM^T + XBM^T A = AMBX^T - MBX^T A$$

or, equivalently, as $XBM^T A + MBX^T A = AXBM^T + AMBX^T$. This equation is equivalent to its transpose, which is nothing but the equation in (iv). The proof is complete. □

Finally, in the remainder of this section, we will use the computations (2.26)-(2.29) to determine the index of the Hessian of f. In order to express our results more simply we will assume (again, with no significant loss in generality) that the diagonal entries of A and B are in non-decreasing order—i.e. $a_1 < \cdots < a_s$ and $b_1 < \cdots < b_t$.

Definition 2.12. A *sign matrix* is a diagonal matrix with diagonal entries in $\{\pm 1\}$. A *signed permutation matrix (SPM)* is a matrix of the form SP_σ with S a sign matrix and $\sigma \in \mathfrak{S}_n$ a permutation.

Lemma 2.13. *Every connected component of F contains a signed permutation matrix.*

Proof. Each connected component of any orthogonal group certainly contains a SPM, so by Theorem 2.8 it suffices to show that if Q, R, ϵ are as in Construction 2.5, and the $Q[i]$ and $R[j]$ are SPMs, then $X = \Phi_\epsilon(Q, R)$ is a SPM. Since $X \in O(n)$ it suffices to show that no row of X contains more than one non-zero entry. To see this, suppose $X_{p,q}, X_{p,r} \neq 0$ for some p, q, r. According to the construction of $X = \Phi_\epsilon(Q, R)$, we would then have

$$X_{p,q} = X[i,j]_{\overline{p},\overline{q}} = \sum_s Q[i,j]_{\overline{p},s} R[i,j]_{\overline{q},s}$$

$$X_{p,r} = X[i,l]_{\overline{p},\overline{r}} = \sum_t Q[i,l]_{\overline{p},t} R[i,l]_{\overline{r},t}$$

for $i, j, l, \overline{p}, \overline{q}, \overline{r}$ determined from p, q, r in the obvious manner. Since $Q[i]$ is a SPM there is a unique $\overline{j}$ such that row $\overline{p}$ of $Q[i, \overline{j}]$ is not identically zero, and, furthermore, there is a unique $\overline{s}$ so that $Q[i, \overline{j}]_{\overline{p},\overline{s}} \neq 0$. Since $X_{p,q}, X_{p,r} \neq 0$, the former fact implies that $j = l = \overline{j}$, and then the latter fact (together with the fact that $R[i, \overline{j}]_{u,\overline{s}} \neq 0$ for at most one u because $R[\overline{j}]$ is a SPM) implies that $\overline{q} = \overline{r}$. Since $j = l$ and $\overline{q} = \overline{r}$ we have $q = r$. □

Theorem 2.14. *Let $X = SP_\sigma$ be a signed permutation matrix. Then:*

(1) X is a critical point of f.
(2) If we let ϵ_{ij} denote the number of non-zero entries in the $m_i \times n_j$ block $X[i,j]$ of X then $\epsilon = (\epsilon_{ij})$ is a perfect filling and there are Q, R as in Construction 2.5 such that $X = \Phi_\epsilon(Q,R)$. (One can even arrange that the $Q[i]$ are SPMs and the $R[j]$ are permutation matrices.)
(3) The Hessian $H = H(f,X)$ of f at X (viewed as a quadratic form on $\mathfrak{so}(n)$ as in Lemma 2.9) is diagonal in the standard basis for $\mathfrak{so}(n)$ and the numbers $H(p,q) := H(\mathbb{E}(p,q), \mathbb{E}(p,q))$ (for $1 \le p < q \le n$) are given by

$$H(p,q) = 2(B_{pp} - B_{qq})(A_{\sigma(q)\sigma(q)} - A_{\sigma(p)\sigma(p)}).$$

(4) The index of $H = H(f,X)$ is equal to the number of $(p,q) \in \{1,\dots,n\}^2$ with $B_{pp} < B_{qq}$ and $A_{\sigma(q)\sigma(q)} > A_{\sigma(p)\sigma(p)}$. In terms of the perfect filling ϵ associated to X in (2), this index can be written

$$\sum_{(i,j)<(k,l)} \epsilon_{ij}\epsilon_{kl},$$

where $(i,j) < (k,l)$ means $i < k$ and $j < l$.

Proof. (1) follows from (2) in light of Proposition 2.7, though it can be seen more directly as follows: Using (2.28) we compute

$$XBX^T = SP_\sigma BP_\sigma^T S = \sum_{p=1}^{n} B_{pp}\mathbb{D}(\sigma(p)).$$

Since this matrix is diagonal, $X \in F$ by Lemma 2.2.

For (2), it is obvious from the construction of the ϵ_{ij} that ϵ is a perfect filling. (For example, $\sum_j \epsilon_{ij} = m_i$ for any fixed i because $\sum_j \epsilon_{ij}$ is just the total number of non-zero entries in some m_i rows of X; this is m_i because X is a SPM.) Fix some i and j. Let $z_1, \dots, z_r \in \{1,\dots,n_j\}$ ($r := n_j - \epsilon_{ij}$) be the identically zero columns of the $m_i \times n_j$ matrix $X[i,j]$. Let $Q[i,j]$ be the $m_i \times \epsilon_{ij}$ matrix obtained from $X[i,j]$ by deleting these columns. Let $R[i,j]^T$ be the $\epsilon_{ij} \times n_j$ matrix such that columns $z_1, \dots, z_r$ of $R[i,j]^T$ are identically zero and such that the $\epsilon_{ij} \times \epsilon_{ij}$ matrix obtained by deleting these zero columns from $R[i,j]^T$ is the identity matrix $I_{\epsilon_{ij}}$. Then we clearly have $X[i,j] = Q[i,j]R[i,j]^T$. For any fixed i, the $m_i \times m_i$ matrix $Q[i] = \begin{pmatrix} Q[i,1] & \cdots & Q[i,t] \end{pmatrix}$ is obtained from the SPM X by taking the m_i rows $\begin{pmatrix} X[i,1] & \cdots & X[i,t] \end{pmatrix}$ of X and then deleting the identically zero columns in the resulting matrix. It is evident from this description of $Q[i]$ that $Q[i]$ is itself a SPM. One can see similarly that, for any fixed j, the $n_j \times n_j$ matrix $R[j] = \begin{pmatrix} R[1,j] & \cdots & R[s,j] \end{pmatrix}$ is a permutation matrix.

The description of the Hessian in (3) follows from (2.22) by using (2.26)-(2.29). (The basic point here is that the diagonal entries of the matrices $\mathbb{F}(p,q)$ are zero, hence the diagonal entries of any product of $\mathbb{F}(p,q)$ and a diagonal matrix will also be zero.)

The first formula for the index in (4) is immediate from (3). (Note that $B_{pp} < B_{qq}$ can occur only if $p < q$ because we are assuming $b_1 < \dots < b_t$.) To see that this is equal to the other formula for the index, first notice that the sum in (4) counts the number of pairs $((p',p),(q',q))$ such that the (p',p) and (q',q) entries of X are non-zero and the (p',p) entry lies in a block $X[i,j]$ further up and to the left in X than the block $X[k,l]$ containing (q',q). Since $X = SP_\sigma$ is a signed permutation matrix we must have $p' = \sigma(p)$ and $q' = \sigma(q)$ for any such pair. Furthermore, the condition that the block $X[k,l]$ containing $(\sigma(q),q)$ is further down and further right than the block $X[i,j]$ containing $(\sigma(p),p)$ is equivalent to the conditions on (p,q) in the first formula for the index. □

Let $F_\epsilon \subseteq F$ be the image of the map Φ_ϵ, so that F is the disjoint union of the F_ϵ by Theorem 2.8. Then:

Corollary 2.15. *The index of $H(f)$ on F_ϵ is constant, given by the sum in Theorem 2.14*(4).

Proof. The index of the Hessian is always locally constant, so it suffices to show that the index is given by the claimed formula on each component of F_ϵ. By Lemma 2.13 each component of F_ϵ contains a signed permutation matrix and by Theorem 2.14 the index at any such permutation matrix is as claimed. □

Example 2.16. Suppose $A = \mathrm{Diag}(a_1, \dots, a_n)$ (for distinct a_p) and $B = \mathrm{Diag}(b_1, \dots, b_n)$ (for distinct b_p). (We continue to assume that the a_p and the b_p are in increasing order.) In this case, Theorems 2.8 and 2.11 show that f is a Morse function whose critical points are *precisely* the signed permutation matrices SP_σ. Fix a signed permutation matrix $X = SP_\sigma$. In this case $B_{pp} < B_{qq}$ (resp. $A_{\sigma(q)\sigma(q)} > A_{\sigma(p)\sigma(p)}$) iff $p < q$ (resp. $\sigma(q) > \sigma(p)$), so by Theorem 2.14 the index of f at X is given by

$$|\{(p, q) : 1 \le p < q \le n \text{ and } \sigma(q) > \sigma(p)\}|$$

which is nothing but the inversion number of σ.

Example 2.17. Another extreme example occurs when $A = aI_n$ and $B = bI_n$. In this case f is constant, so its index at any critical point is zero. This is consistent with Theorem 2.14(4) because there are no such pairs (p, q), nor are there any $(i, j), (k, l)$ with $(i, j) < (k, l)$. (Indeed, there is only one pair (i, j), namely $(1, 1)$.)

3. Linear Morse-Bott functions on $O(n)$

We call a function $f : O(n) \to \mathbb{R}$ *linear* if f is the restriction of a real linear function on the vector space V of $n \times n$ matrices. These functions have been studied in [3], [5], [6], [7], [10], [11], and [12]. Since $\langle A, X\rangle := \mathrm{Tr}(A^T X)$ is a non-degenerate inner product on V, any linear function $f : V \to \mathbb{R}$ is of the form $f(X) := \mathrm{Tr}(A^T X)$ for a unique $A \in V$. We view this (positive definite) inner product as a "constant" Riemannian metric on V. Its restriction to $O(n)$ is the bi-invariant metric on $O(n)$ of §2. The gradient of $f = f_A : O(n) \to \mathbb{R}$ with respect to this metric is easily computed to be

$$(\nabla f)(X) = \frac{1}{2}(A - XA^T X) \in T_X O(n) \tag{3.1}$$

for $X \in O(n)$, when f is regarded as a function $f : O(n) \to \mathbb{R}$. Since the critical points of f are precisely the points of $X \in O(n)$ where $(\nabla f)(X) = 0$, we obtain from (3.1):

Lemma 3.1. *$X \in O(n)$ is a critical point of f_A iff XA^T is symmetric (equivalently $A^T X$ is symmetric).*

The linear functions $f_A : O(n) \to \mathbb{R}$ were studied in [12], where it is shown that f_A is a Morse function on $O(n)$ iff the symmetric matrix AA^T has n distinct eigenvalues. This holds, for example, if $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ with the λ_i distinct *non-negative* real numbers. The resulting functions f_A were studied in [3, §3.3], where their critical points and indices are determined.[2] In the special case where $A = I$, the function $g := f_A$ is given by $g(X) = \mathrm{Tr}(X)$. The function g (or, more precisely, its restriction to $SO(n)$) was studied by Frankel in [6]. From Lemma 3.1 we see that the critical locus of g is exactly the symmetric matrices in $O(n)$.

[2] Strictly-speaking, the functions studied in [3, §3.3] are not linear as written there, but are obtained from the aforementioned linear functions by composing with *affine*-linear automorphisms of $\mathbb{R}$, as is evident from the expression in [3, §3.3].

In [6] a diffeomorphism

$$\coprod_k \mathbb{G}(2k,n) \to F \subseteq SO(n), \qquad \Lambda \mapsto X(\Lambda), \tag{3.2}$$

is established, where $\mathbb{G}(2k,n)$ is the Grassmannian of $2k$-dimensional linear subspaces $\Lambda \subseteq \mathbb{R}^n$ and F denotes the critical locus of $g: SO(n) \to \mathbb{R}$. This diffeomorphism can be obtained easily as follows: Given $\Lambda \in \mathbb{G}(2k,n)$, let $X = X(\Lambda)$ be the unique linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ such that Λ (resp. $\Lambda^\perp$) is the (-1)-eigenspace (resp. 1-eigenspace) of X. Obviously X is self-adjoint (so $X = X^T$), $X^2 = I$, and $X \in SO(n)$ (because the dimension $2k$ of Λ is even), so $X \in F$. This yields a map as in (3.2) which is clearly smooth. A smooth inverse for this map can be constructed as follows: Fix $X \in F$. Since X is symmetric, it is (orthogonally) diagonalizable (over $\mathbb{R}$, so its eigenvalues are real) and distinct eigenspaces of X are orthogonal. Furthermore $X \in O(n)$, so its eigenvalues have magnitude 1. Therefore, if we let $\Lambda = \Lambda(X)$ be the (-1)-eigenspace of X, then we have an orthogonal direct sum decomposition $\mathbb{R}^n = \Lambda \oplus \Lambda^\perp$ with $\Lambda^\perp$ equal to the 1-eigenspace of X. Since $X \in SO(n)$, the dimension of Λ must be even ($2k$, say). Evidently $X \mapsto \Lambda(X)$ is the inverse of $\Lambda \mapsto X(\Lambda)$.

The Hessian of g (and, more generally, of any linear function $f_A : O(n) \to \mathbb{R}$) is easily described. As in §2, we view the Hessian of f_A as a quadratic form on $\mathfrak{so}(n)$ via the isomorphism (2.21). The proof of the following lemma is essentially the same as the proof of Lemma 2.9.

Lemma 3.2. *The Hessian $H = H(f_A, X)$ of the linear function f_A at a critical point X, viewed as a quadratic form on $\mathfrak{so}(n)$ as above, is given for $E, N \in \mathfrak{so}(n)$ by*

$$H(E,N) = \mathrm{Tr}(A^T XEN).$$

Remark 3.3. The bilinear form H of Lemma 3.2 is not generally symmetric when $X \notin F$. When $X \in F$ the symmetricity can be verified directly.

We can see from Lemma 3.2 that $g : O(n) \to \mathbb{R}$ is Morse-Bott, as follows: Fix any $X \in F \subseteq O(n)$, $M \in T_X O(n)$. Since F consists of the symmetric matrices in $O(n)$ we have $M \in T_X F$ iff $X + \epsilon M \in SO(n, \mathbb{R}[\epsilon]/\epsilon^2)$ is symmetric, which, since X is symmetric, is equivalent to saying M is symmetric. It is obvious from basic properties of the trace that $\mathrm{Tr}(MN) = 0$ whenever M is symmetric and N is skew-symmetric. On dimension grounds we therefore have an orthogonal direct sum decomposition

$$\begin{aligned} V &= \{M \in V : M = M^T\} \oplus \{N \in V : N = -N^T\} \\ &= \{M \in V : M = M^T\} \oplus \mathfrak{so}(n). \end{aligned}$$

Therefore M is in $T_X F$ iff $\mathrm{Tr}(MN) = 0$ for all $N \in \mathfrak{so}(n)$. It is immediate to see that this latter condition is equivalent to M being in the kernel of $H(g,X)$. This proves that the kernel of $H(g,X)$ is precisely $T_X F$ and therefore g is Morse-Bott.

Since the Grassmannians $\mathbb{G}(k,n)$ are connected, the index of the Hessian of g is determined by its values at the critical points $-I_k \oplus I_{n-k}$. These indices are easily computed (see e.g. [5, Proposition 1.2] or [6, Lemma 3]). We present the proof in our setup:

Lemma 3.4. *When $X = -I_k \oplus I_{n-k}$ and $A = I$, the quadratic form $H(E,N) = \mathrm{Tr}(XEN)$ on $\mathfrak{so}(n)$ of Lemma 3.2 is diagonal in the standard basis (Definition 2.10) for $\mathfrak{so}(n)$. The numbers $H(p,q) := H(\mathbb{E}(p,q), \mathbb{E}(p,q))$ are given by*

$$H(p,q) = \begin{cases} -2, & k < p < q \leq n \\ 0, & 1 \leq p \leq k < q \leq n \\ 2, & 1 \leq p < q \leq k. \end{cases}$$

The index of H is $\iota(k) := \binom{n-k}{2}$.

Proof. For standard basis vectors $\mathbb{E}(p,q), \mathbb{E}(u,v) \in \mathfrak{so}(n)$, we see that $\mathbb{E}(p,q)\mathbb{E}(u,v)$ has no non-zero diagonal entries when $(p,q) \neq (u,v)$. Since the effect of multiplying on the left by X is simply to multiply the first k rows by -1, the matrices $X\mathbb{E}(p,q)\mathbb{E}(u,v)$ still have no non-zero diagonal entries when $(p,q) \neq (u,v)$. These matrices therefore have trace zero, which shows that H is diagonal in the standard basis. The non-zero diagonal entries of $\mathbb{E}(p,q)^2$ are precisely the (p,p)-entry and the (q,q)-entry, both of which are -1. Taking into account the effect of multiplying on the left by X, we arrive at the formula for the $H(p,q) = \mathrm{Tr}(X\mathbb{E}(p,q)^2)$. The formula for the index amounts to counting the number of pairs (p,q) with $k < p < q \leq n$. □

The "Morse-Bott inequalities" (i.e. the existence of a spectral sequence going from the cohomology of the critical locus F shifted by the index of the Hessian to the cohomology of $SO(n)$) for g imply that

$$\dim \mathrm{H}^i(SO(n), \mathbb{F}) \leq \sum_k \dim \mathrm{H}^{i-\iota(2k)}(\mathbb{G}(2k,n), \mathbb{F}) \tag{3.3}$$

for all i, n and any field $\mathbb{F}$ (with $\iota(2k)$ as in Lemma 3.4). When $\mathbb{F} = \mathbb{F}_2$, Frankel shows in [6] that the inequalities (3.3) are in fact *equalities*. This is done by appeal to a result of E. E. Floyd [6, Theorem A] asserting that, since F is the fixed locus of the involution $X \mapsto X^{-1} = X^T$ of the smooth, compact manifold $SO(n)$, the sum of the mod 2 Betti numbers of F is bounded above by the sum of the mod 2 Betti numbers of $SO(n)$. We believe that it is possible to give a purely Morse theoretic proof of this fact (cf. §5), but we have not attempted this. In the Appendix we give a purely combinatorial proof of Frankel's mod 2 Betti number relationship.

Example 3.5. When $n = 3$ the critical locus F of $g : SO(3) \to \mathbb{R}$ is the disjoint union of $\{I\} = \mathbb{G}(0,3)$ (index 3) and $\mathbb{G}(2,3) \cong \mathbb{RP}^2$ (index 0). The equalities of mod 2 Betti numbers above amount to the equality of (mod 2) Poincaré polynomials

$$p(SO(3)) = p(\mathbb{RP}^3) = 1 + t + t^2 + t^3 = p(\mathbb{RP}^2) + t^3 p(\{I\}).$$

4. Simple Morse-Bott cohomology

Let X be a smooth compact manifold of dimension d. For simplicity, we assume X is connected. A *simple Morse-Bott function* is a non-constant Morse-Bott function $f : X \to \mathbb{R}$ whose critical locus F consists only of points where f obtains its maximum or minimum value. Throughout this section, f will be a simple Morse-Bott function on X.

We have $F = F_0 \coprod F_k$, where F_0 (resp. F_k) is the index 0 (resp. k) critical locus consisting of minima (resp. maxima) for f. The image of f must be a closed interval $[a,b] \subseteq \mathbb{R}$ with $F_0 = f^{-1}(a)$, $F_k = f^{-1}(b)$. Note that k is also the codimension of F_k in X, since the Hessian of f must be negative definite on the normal bundle of F_k in X. Let m be the codimension of F_0 in X.

Fix a Riemannian metric on f. The gradient ∇f is the vector field on X dual to the 1-form df under the metric. Integration of ∇f yields a smooth action of the group $G = (\mathbb{R}, +)$ on X, denoted $\tau \cdot x$. The action fixes F and is free on $X \setminus F$. For any $x \in X \setminus F$, the function $t \mapsto f(\tau \cdot x)$ is a strictly increasing function of $\tau \in \mathbb{R}$ approaching b (resp. a) at $t \to \infty$ (resp. $\tau \to -\infty$). The quotient $(X \setminus F)/G =: M$ can be (and will be) identified with any regular fiber $f^{-1}(c)$ $(c \in (a,b))$ of f, thus we view M as a closed subspace of X contained in $X \setminus F$. There is a continuous source map

$$\begin{aligned} s : X \setminus F_k &\to F_0 \\ s(x) &:= \lim_{\tau \to -\infty} \tau \cdot x \end{aligned}$$

which is a locally trivial $\mathbb{R}^m$ bundle and whose restriction to M is a locally trivial S^{m-1} bundle (any regular level set of f intersects any fiber of τ in a sphere). (See [1, Theorem A.9]. It seems that $X \setminus F_k$ should be diffeomorphic to the normal bundle $N = N_{F_0/X}$ by a diffeomorphism exchanging s and the projection $N \to F_0$ but in [1] this is only shown to hold locally. One can easily show that, in the situation we shall consider in §5, one does have such a global diffeomorphism.) Similarly, there is a continuous target map

$$\begin{aligned} t : X \setminus F_0 &\to F_k \\ t(x) &:= \lim_{\tau \to +\infty} \tau \cdot x \end{aligned}$$

which is a locally trivial $\mathbb{R}^k$ bundle and whose restriction to M is a locally trivial S^{k-1} bundle.

Fix a "coefficient" ring A. The properties of t mentioned above imply that $\mathrm{R}\, t_! \underline{A}_{X \setminus F_0}$ is locally isomorphic to $\underline{A}_{F_k}[-k]$. We assume that t is *oriented* (with respect to A) in the sense that there is an isomorphism

$$\eta : \mathrm{R}\, t_! \underline{A}_{X \setminus F_0} \to \underline{A}_{F_k}[-k]$$

(i.e. an isomorphism of $\underline{A}_{F_k}$ modules $\mathrm{R}^k\, t_! \underline{A}_{X \setminus F_0} \cong \underline{A}_{F_k}$). Set $Y = f^{-1}[a,c]$, $Z = f^{-1}[c,b]$.

We have a commutative diagram of $\underline{A}_X$ modules with exact rows

$$\begin{tikzcd} 0 \arrow[r] & \underline{A}_{Z \setminus M} \arrow[r] & \underline{A}_Z \arrow[r] & \underline{A}_M \arrow[r] & 0 \\ 0 \arrow[r] & \underline{A}_{X \setminus Y} \arrow[r] \arrow[u, equal] \arrow[d] & \underline{A}_X \arrow[r] \arrow[u] \arrow[d, equal] & \underline{A}_Y \arrow[r] \arrow[u] \arrow[d] & 0 \\ 0 \arrow[r] & \underline{A}_{X \setminus F_0} \arrow[r] & \underline{A}_X \arrow[r] & \underline{A}_{F_0} \arrow[r] & 0 \end{tikzcd} \tag{4.1}$$

where we have suppressed notation for proper pushforwards to X (this is the usual pushforward for the closed subspaces F_0, M, Y, and Z and the "extension by zero" for the open subspaces $Z \setminus M = X \setminus Y$ and $X \setminus F_0$).

The map $\underline{A}_{X \setminus Y} \to \underline{A}_{X \setminus F_0}$ in (4.1) becomes an isomorphism when $\mathrm{R}\, t_!$ is applied: Indeed, by the base change theorem for proper direct images, it suffices to show that, for any $x \in F_k$, the map $(t|X \setminus Y)^{-1}(x) \hookrightarrow t^{-1}(x)$ induces an isomorphism on compactly supported cohomology. This map is homeomorphic to the inclusion of the open unit ball into $\mathbb{R}^k$, so this is indeed the case. The "orientation" isomorphism η from above therefore also yields an isomorphism $\mathrm{R}\, t_! \underline{A}_{X \setminus Y} \cong \underline{A}_{F_k}[-k]$ which we also call η.

Applying $\mathrm{R}\, t_!$ to (the triangle associated to) the top row of (4.1) and using this isomorphism, we obtain a triangle

$$\underline{A}_{F_k}[-k] \longrightarrow \mathrm{R}\, t_! \underline{A}_Z \longrightarrow \mathrm{R}\, t_* \underline{A}_M \xrightarrow{t_*} \underline{A}_{F_k}[1-k] \tag{4.2}$$

in the derived category $\mathbf{D}(\underline{A}_{F_k})$ (note that $\mathrm{R}\, t_! \underline{A}_M = \mathrm{R}\, t_* \underline{A}_M$ because $t|M : M \to F_k$ is a sphere bundle, so it is proper). Applying $\mathrm{R}\,\Gamma = \mathrm{R}\,\Gamma(X, \_\_)$ to (4.1) and using the isomorphism(s) η yields a commutative diagram

$$\begin{tikzcd}
\mathrm{R}\,\Gamma(F_k, A)[-k] \arrow[r] & \mathrm{R}\,\Gamma(Z, A) \arrow[r] & \mathrm{R}\,\Gamma(M, A) \arrow[r, "t_*"] & \mathrm{R}\,\Gamma(F_k, A)[1-k] \\
\mathrm{R}\,\Gamma(F_k, A)[-k] \arrow[r] \arrow[u, equal] \arrow[d, "\cong"'] & \mathrm{R}\,\Gamma(X, A) \arrow[r] \arrow[u] \arrow[d, equal] & \mathrm{R}\,\Gamma(Y, A) \arrow[r] \arrow[u] \arrow[d, "\cong"'] & \mathrm{R}\,\Gamma(F_k, A)[1-k] \arrow[u, equal] \arrow[d, "\cong"'] \\
\mathrm{R}\,\Gamma_!(X \setminus F_0, A) \arrow[r] & \mathrm{R}\,\Gamma(X, A) \arrow[r] & \mathrm{R}\,\Gamma(F_0, A) \arrow[r] & \mathrm{R}\,\Gamma_!(X \setminus F_0, A)[1]
\end{tikzcd}$$

in $\mathbf{D}(A)$ where the rows are triangles, $\Gamma_!$ is the compactly supported global sections functor, and the vertical arrows define maps of triangles. The map $\mathrm{R}\,\Gamma(Y, A) \to \mathrm{R}\,\Gamma(F_0, A)$ is the map induced by the inclusion $F_0 \hookrightarrow Y$. This inclusion induces isomorphisms on cohomology because it is retracted by the map $s|Y : Y \to F_0$, which is a closed disc bundle and hence induces isomorphisms on cohomology. We thus see that the map $\mathrm{R}\,\Gamma(F_0, A) \to \mathrm{R}\,\Gamma(M, A)$ given by composing the inverse of $\mathrm{R}\,\Gamma(Y, A) \to \mathrm{R}\,\Gamma(F_0, A)$ and $\mathrm{R}\,\Gamma(Y, A) \to \mathrm{R}\,\Gamma(M, A)$ (this map is the one induced by the inclusion $M \hookrightarrow Y$) is nothing but the map induced by $s|M : M \to F_0$. We thus obtain a triangle

$$\mathrm{R}\,\Gamma(F_k, A)[-k] \longrightarrow \mathrm{R}\,\Gamma(X, A) \longrightarrow \mathrm{R}\,\Gamma(F_0, A) \xrightarrow{t_* s^*} \mathrm{R}\,\Gamma(F_k, A)[1-k]$$

in $\mathbf{D}(A)$ and hence also a long exact sequence

$$\cdots \xrightarrow{t_* s^*} \mathrm{H}^{i-k}(F_k) \longrightarrow \mathrm{H}^i(X) \longrightarrow \mathrm{H}^i(F_0) \xrightarrow{t_* s^*} \mathrm{H}^{i+1-k}(F_k) \longrightarrow \cdots \tag{4.3}$$

in cohomology (coefficients in A understood).

The map $t_* : \mathrm{H}^*(M) \to \mathrm{H}^{*+1-k}(F_k)$ appearing in (4.3) satisfies the *Projection Formula*

$$t_*(t^*(\alpha) \cdot \beta) = \alpha \cdot t_*(\beta)$$

for $\alpha \in \mathrm{H}^*(F_k)$, $\beta \in \mathrm{H}^*(M)$. This holds as a matter of "general nonsense" owing to the construction of the latter map from the map $t_* : \mathrm{R}\, t_* \underline{A}_M \to \underline{A}_{F_k}[1-k]$ in (4.2). We will recall the details for the reader's convenience. A cohomology class $\alpha \in \mathrm{H}^r(F_k)$ is the same thing as a $\mathbf{D}(\underline{A}_{F_k})$-morphism $\alpha : \underline{A}_{F_k} \to \underline{A}_{F_k}[r]$. Similarly, $\beta \in \mathrm{H}^q(M)$ is a map $\beta : \underline{A}_M \to \underline{A}_M[q]$. The cup product $t^*(\alpha) \cdot \beta$ corresponds to the composition

$$\underline{A}_M \xrightarrow{t^{-1}\alpha} \underline{A}_M[r] \xrightarrow{\beta[r]} \underline{A}_M[r+q].$$

The adjunction map $\mathrm{Id} \to t_* t^{-1}$ yields a natural transformation $\mathrm{Id} \to Rt_* t^{-1}$. Evaluating this on $\underline{A}_{F_k}$ yields a natural map $\theta : \underline{A}_{F_k} \to \mathrm{R}\, t_* \underline{A}_M$. The cohomology class $t_*\beta \in \mathrm{H}^{q+1-k}(F_k)$ corresponds to the composition

$$\underline{A}_{F_k} \xrightarrow{\theta} \mathrm{R}\, t_* \underline{A}_M \xrightarrow{\mathrm{R}\, t_* \beta} \mathrm{R}\, t_* \underline{A}_M[q] \xrightarrow{t_*[q]} \underline{A}_{F_k}[q+1-k].$$

We have a commutative diagram

$$\begin{tikzcd}
\underline{A}_{F_k} \arrow[r, "\alpha"] \arrow[d, "\theta"'] & \underline{A}_{F_k}[r] \arrow[d, "\theta[r]"'] & \underline{A}_{F_k}[q+r+1-k] \\
\mathrm{R}\, t_* \underline{A}_M \arrow[r, "\mathrm{R}\, t_* t^{-1}\alpha"] & \mathrm{R}\, t_* \underline{A}_M[r] \arrow[r, "\mathrm{R}\, t_* \beta[r]"] & \mathrm{R}\, t_* \underline{A}_M[q+r] \arrow[u, "t_*[q+r]"]
\end{tikzcd}$$

in $\mathbf{D}(\underline{A}_{F_k})$. The two ways around this diagram are the two sides of the Projection Formula (right first is the LHS, down first is the RHS).

The long exact sequence (4.3) gives rise to short exact sequences

$$0 \to \mathrm{Cok}(\delta_{i-1}) \to \mathrm{H}^i(X) \to \mathrm{Ker}(\delta_i) \to 0,$$

where

$$\delta_i = t_* s^* : \mathrm{H}^i(F_0) \to \mathrm{H}^{i+1-k}(F_k).$$

5. A particularly nice Morse-Bott function on $SO(n)$

We now specialize the discussion of §3 to the case where $A = \mathrm{Diag}(0, \ldots, 0, 1)$. In this case $f = f_A : SO(n) \to \mathbb{R}$ is given by $f(X) = X_{nn}$ (the lower right entry of X). This function f arises as the composition of the map

$$p : SO(n) \to S^{n-1}$$
$$p(X) := (X_{n1}, \ldots, X_{nn})$$

and the usual height function

$$h : S^{n-1} \to \mathbb{R}$$
$$h(x_1, \ldots, x_n) := x_n.$$

The map p is submersive — in fact it is an $SO(n-1)$ principal bundle (see below). The height function is Morse, hence f is Morse-Bott. The critical points of the height function are just the points $(0, \ldots, 0, \pm 1)$ of S^{n-1} where the height is minimum or maximum, hence the critical loci of f are just the points of $SO(n)$ mapped to these two points by p.

Explicitly, the critical locus F of f is $F_0 \coprod F_{n-1}$ where

$$F_0 = \{X \in SO(n) : X_{nn} = -1\}$$
$$F_{n-1} = \{X \in SO(n) : X_{nn} = 1\}.$$

We identify F_{n-1} with $SO(n-1)$ by taking $Q \in SO(n-1)$ to the block-diagonal matrix $Q \oplus 1 \in F_{n-1}$. Throughout, we let $J := \mathrm{Diag}(-1, 1, \ldots, 1) \in O(n)$. Geometrically, J is a reflection across the hyperplane $e_1^\perp$. We have $\det J = -1$ and $J^2 = I$. We identify F_0 with $SO(n-1)$ by taking $Q \in SO(n-1)$ to the block-diagonal matrix $JQ \oplus -1 \in F_0$.

Since f takes its minimum (resp. maximum) value on F_0 (resp. F_{n-1}), the Hessian of f must be positive (resp. negative) definite on the normal bundle of F_0 (resp. F_{n-1}). Therefore, F_0 is of index 0 and F_{n-1} is of index $n - 1 = \mathrm{codim}(F_{n-1} \subseteq SO(n))$.

The moduli space of flow lines $M = f^{-1}(0)$ is just the set of $X \in SO(n)$ with $X_{nn} = 0$. We shall make use of the map $\pi : M \to S^{n-2}$ defined by $\pi(X) := (X_{1n}, \dots, X_{n-1,n})$. I.e., $\pi(X)$ is the *right column* of X, excepting the lower right entry, which is zero. The map π is not to be confused with the restriction of $p : SO(n) \to S^{n-1}$ to M. The latter can also be viewed as a map $M \to S^{n-2}$, defined using the *bottom row* rather than the *right column* (always excepting the lower right entry).

The group $G := SO(n-1)$ acts (on the left, smoothly) on $SO(n)$ by letting $g \in G$ act on $SO(n)$ via left multiplication by $g \oplus 1 \in SO(n)$. This is an action through isometries of $SO(n)$ making p a principal G-bundle. In particular, p is G-invariant, and hence so is $f = hp$. Hence G also acts naturally on F_0, F_{n-1}, and M, making the source map $s : M \to F_0$ and target map $t : M \to F_{n-1}$ equivariant. Under the identifications $F_0 = SO(n-1)$, $F_{n-1} = SO(n-1)$ from above, the G action on F_{n-1} is identified with the action of $SO(n-1)$ on itself by left multiplication ($g \cdot Q = gQ$), while the G action on F_0 is identified with the action of $SO(n-1)$ on itself defined by $g \cdot Q = JgJQ$. Note that the map $\pi : M \to S^{n-2}$ defined above is G-*equivariant* (not G-*invariant*) when $G = SO(n-1)$ acts on S^{n-2} by left multiplication, as usual. The latter action is transitive, hence:

(*) For any $X \in M$, there is a $g \in G$ such that the right column of $g \cdot X$ is e_1.

The general formula (3.1), specialized to our case ($A = \mathrm{Diag}(0, \dots, 0, 1)$) yields the following explicit formula for the gradient of our Morse-Bott function f:

$$(\nabla f)(X) = \frac{1}{2} \begin{pmatrix} -X_{1n}X_{n1} & -X_{1n}X_{n2} & \cdots & -X_{1n}X_{nn} \\ -X_{2n}X_{n1} & -X_{2n}X_{n2} & \cdots & -X_{2n}X_{nn} \\ \vdots & \vdots & & \vdots \\ -X_{n-1,n}X_{n1} & -X_{n-1,n}X_{n2} & \cdots & -X_{n-1,n}X_{nn} \\ -X_{nn}X_{n1} & -X_{nn}X_{n2} & \cdots & 1 - X_{nn}^2 \end{pmatrix}.$$

The *key observation* about this formula is that rows $2, \dots, n-1$ of $(\nabla f)(X)$ will be zero whenever X has $X_{2n} = X_{3n} = \cdots = X_{n-1,n} = 0$. This implies that, for such an X, rows $2, \dots, n-1$ of X must remain constant along the gradient flow of X (and hence also at the two limit points of the flow of X, when X is not a critical point). This observation and **(*)** above allow us to describe the source and target maps s, t explicitly, as follows: Consider some $X \in M$ whose right column is e_1. Write X in the block form

$$X = \begin{pmatrix} 0 & 1 \\ V & 0 \\ v & 0 \end{pmatrix} \tag{5.1}$$

where V is $(n-2) \times (n-1)$ and v is $1 \times (n-1)$. Note that $\begin{pmatrix} V \\ v \end{pmatrix}$ is in $O(n-1)$ and has determinant $(-1)^{n+1}$, hence $\det \begin{pmatrix} v \\ V \end{pmatrix} = -1$. Since rows $2, \dots, n-1$ of $s(X) \in F_0 \subseteq SO(n)$ must be the same as those of X, we must have

$$s(X) = \begin{pmatrix} v & 0 \\ V & 0 \\ 0 & -1 \end{pmatrix} \in F_0 \subseteq SO(n)$$

Under our identification $F_0 = SO(n-1)$, we therefore have

$$s(X) = \begin{pmatrix} -v \\ V \end{pmatrix} \in SO(n-1) \tag{5.2}$$

(recall that this identification involves a left multiplication by J, which changes the sign of the first row). Similar considerations show that, for X as above, $t(X) \in F_{n-1} \subseteq SO(n)$ is given by

$$t(X) = \begin{pmatrix} -v & 0 \\ V & 0 \\ 0 & 1 \end{pmatrix} \in F_{n-1} \subseteq SO(n).$$

Under our identification $F_{n-1} = SO(n-1)$, we therefore have

$$t(X) = \begin{pmatrix} -v \\ V \end{pmatrix} \in SO(n-1). \tag{5.3}$$

Since the maps $s : M \to SO(n-1)$ and $t : M \to SO(n-1)$ are G-equivariant, they are determined by (5.2) and (5.3) in light of **(*)**.

The maps s and t, while not the same, are still "close to being equal" in a sense we will now make precise. Since $SO(n-1)$ is a Lie group, the set $\mathrm{Hom}(M, SO(n-1))$ has a natural group structure, hence we can consider the map $st^{-1} : M \to SO(n-1)$. (If s and t are "kind of the same," then st^{-1} ought to be "close to the identity".) To describe this map, first notice that, for any $L \in \mathbb{RP}^{n-2}$ (i.e. any 1-dimensional linear subspace of $\mathbb{R}^{n-1}$), the reflection $\rho(L^\perp)$ across the hyperplane $L^\perp$ is an orthogonal linear transformation of $\mathbb{R}^{n-1}$ of determinant -1. Following this reflection by our reflection J yields an element $r(L) := J\rho(L^\perp) \in SO(n-1)$. This defines a smooth map $r : \mathbb{RP}^{n-2} \to SO(n-1)$.

Proposition 5.1. *The map $st^{-1} : M \to SO(n-1)$ is equal to the composition of the map $\pi : M \to S^{n-2}$, the quotient projection $q : S^{n-2} \to \mathbb{RP}^{n-2}$, and the map $r : \mathbb{RP}^{n-2} \to SO(n-1)$ discussed above.*

Proof. Given any $Y \in M$, by **(*)** we can find some $g \in G = SO(n-1)$ and some $X \in M$ with right column e_1 so that $Y = g \cdot X$. If we write this X in the block form (5.1), then, as we saw above,

$$s(Y) = s(g \cdot X) = g \cdot s(X) = JgJ \begin{pmatrix} -v \\ V \end{pmatrix}$$

$$t(Y) = t(g \cdot X) = g \cdot t(X) = g \begin{pmatrix} -v \\ V \end{pmatrix},$$

hence $(st^{-1})(Y) = JgJg^{-1} = JgJg^T$. Now write g in the block form

$$g = \begin{pmatrix} w & W \end{pmatrix}$$

where w is $(n-1) \times 1$ and W is $(n-1) \times (n-2)$. The fact that $gg^T = I$ is equivalent to $WW^T = I - ww^T$, and, using this, we compute

$$\begin{aligned} gJg^T &= \begin{pmatrix} w & W \end{pmatrix} J \begin{pmatrix} w^T \\ W^T \end{pmatrix} \\ &= \begin{pmatrix} w & W \end{pmatrix} \begin{pmatrix} -w^T \\ W^T \end{pmatrix} \\ &= -ww^T + WW^T \\ &= I - 2ww^T. \end{aligned}$$

Since the matrix ww^T is the orthogonal projection onto (the span of) w, we see that $I - 2ww^T = \rho(w^\perp)$ is the reflection across $w^\perp$. We have

$$Y = g \cdot X = \begin{pmatrix} w & W & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ V & 0 \\ v & 0 \end{pmatrix} = \begin{pmatrix} WV & w \\ v & 0 \end{pmatrix},$$

so $\pi(Y) = w$, $(q\pi)(y)$ is the span of w, and $(rq\pi)(y) = J\rho(w^\perp)$ is reflection across $w^\perp$ followed by J, which is indeed equal to $(st^{-1})(Y) = JgJg^{-1}$ since we just saw that $gJg^{-1} = \rho(w^\perp)$. □

We can use our study of the Morse-Bott function f to calculate the Betti numbers $b_i(n) := \dim \mathrm{H}^i(SO(n), \mathbb{F}_2)$ of $SO(n)$ with coefficients in the two element field $\mathbb{F}_2$. We will show that these are given by

$$b_i(n) = |\{S \subseteq \{1, \dots, n-1\} : \sum_{s\in S} s = i\}|.$$

Using the fact that a subset $S \subseteq \{1, \dots, n-1\}$ either contains $n-1$ or does not, we obtain

$$b_i(n) = b_i(n-1) + b_{i+1-n}(n-1) \tag{5.4}$$

for all $n > 1$. The formula (5.4), together with the fact that $b_0(1) = 1$ and $b_i(1) = 0$ for $i \neq 0$, uniquely determines all the $b_i(n)$. In terms of the Poincaré polynomials $p_n(t) := \sum_i b_i(n)t^i$, the description of the $b_i(1)$ above is equivalent to $p_1(t) = 1$ and formula (5.4) is equivalent to

$$p_n(t) = p_{n-1}(t) + t^{n-1}p_{n-1}(t) \tag{5.5}$$

for all $n > 1$. Clearly (5.5), together with the fact that $p_1(t) = 1$, uniquely determines the $p_n(t)$. Since $SO(1)$ is a point, its Poincaré polynomial is indeed $p_1(t) = 1$, so to show that our formula for the $b_i(n)$ is correct, it is therefore enough to show that the Poincaré polynomials of $SO(n)$ satisfy (5.5) for all $n > 1$. We assume $n > 1$ from now on.

The long exact sequence (4.3) for our Morse function f relates the cohomology of $SO(n)$ to the cohomology of F_0 and F_{n-1} (with a degree shift of $n-1$ for the latter), both of which are identified with $SO(n-1)$. From this sequence, we see that (5.5) is equivalent to saying that the maps

$$t_*s^* : \mathrm{H}^i(SO(n-1), \mathbb{F}_2) \to \mathrm{H}^{i+1-n}(SO(n-1), \mathbb{F}_2)$$

are all zero.

Since $n > 1$, the relative dimension $n-1$ of t is > 0, hence it follows from the Projection Formula that $t_*t^* = 0$, so it will be enough to show that $s^* = t^*$ as maps

$$\mathrm{H}^i(SO(n-1), \mathbb{F}_2) \to \mathrm{H}^i(M, \mathbb{F}_2).$$

To see this, we write s as the composition of

$$st^{-1} \times t : M \to SO(n-1) \times SO(n-1)$$

and the multiplication map

$$\mu : SO(n-1) \times SO(n-1) \to SO(n-1).$$

We are working over a field, so we have a Künneth Formula isomorphism

$$\begin{aligned} \mathrm{H}^*(SO(n-1), \mathbb{F}_2) \otimes \mathrm{H}^*(SO(n-1), \mathbb{F}_2) &= \mathrm{H}^*(SO(n-1) \times SO(n-1), \mathbb{F}_2) \\ \alpha \otimes \beta &\mapsto (\pi_1^*\alpha)(\pi_2^*\beta). \end{aligned}$$

In terms of this isomorphism, μ^* can be written

$$\mu^*\beta = \beta \otimes 1 + 1 \otimes \beta + \sum_j \beta'_j \otimes \beta''_j,$$

where the β'_j and β''_j have positive degree.[3]

Now we get to the (next) *key observation*: According to Proposition 5.1, the map st^{-1} factors through the quotient map $q : S^{n-2} \to \mathbb{RP}^{n-2}$. But the maps

$$q^* : \mathrm{H}^i(\mathbb{RP}^{n-2}, \mathbb{F}_2) \to \mathrm{H}^i(S^{n-2}, \mathbb{F}_2)$$

are zero for $i \neq 0$ (this is obvious for $i \neq n-2$ and holds when $i = n-2$ because q has degree 2), hence $(st^{-1})^*\alpha = 0$ whenever α has positive degree. The equality that we want to establish, $s^* = t^*$, is obvious in degree 0, and, for β of positive degree, we compute

$$\begin{aligned} s^*\beta &= ((st^{-1}) \times t)^*\mu^*\beta \\ &= ((st^{-1})^* \otimes t^*)(\beta \otimes 1 + 1 \otimes \beta + \sum_j \beta'_j \otimes \beta''_j) \\ &= (st^{-1})^*\beta + t^*\beta + \sum_j (st^{-1})^*(\beta'_j))(t^*(\beta''_j)) \\ &= t^*\beta. \end{aligned}$$

Appendix A

Recall from §3 that Frankel, in [6], obtained—via Morse Theory—a relationship between the mod 2 Betti numbers of $SO(n)$ and those of the Grassmannians $\mathbb{G}(2k, n)$. Here we will give two alternative derivations of this relationship: the first uses the usual combinatorial description of the mod 2 Betti numbers of Grassmannians and the combinatorial formula for the mod 2 Betti numbers of $SO(n)$ given in §5 (which is also the "usual textbook" description appearing, for example, in [8, §3.D]), while the other uses alternative combinatorial descriptions of these Betti numbers obtained via Morse Theory in [3]. These results can be interpreted in various ways. For example, our first result can be viewed as a derivation of the "usual" combinatorial formula for the mod 2 Betti numbers of $SO(n)$, starting from Frankel's relationship. Similarly, our two results taken together can be viewed as a combinatorial proof that the formula for the numbers $\dim \mathrm{H}^i(O(n), \mathbb{F}_2)$ in [3, §3] coincides with the "usual one".

It will be convenient to work with $O(n)$ instead of $SO(n)$ to avoid various parity considerations. For our purposes, Frankel's relationship is most naturally written

$$\dim \mathrm{H}^i(O(n), \mathbb{F}_2) = \sum_k \dim \mathrm{H}^{i-\iota(k)}(\mathbb{G}(k, n), \mathbb{F}_2), \tag{A.1}$$

where $\iota(k) := \binom{k}{2}$. Since $\mathbb{G}(k, n) \cong \mathbb{G}(n-k, n)$, the RHS of (A.1) is unchanged if we instead take $\iota(k) := \binom{n-k}{2}$. It can be seen similarly that (A.1) is equivalent to

$$\dim \mathrm{H}^i(SO(n), \mathbb{F}_2) = \sum_k \dim \mathrm{H}^{i-\iota(2k)}(\mathbb{G}(2k, n), \mathbb{F}_2),$$

[3] It is easy to prove that this holds for any H-space. See [8, Page 283]. In fact, the Hopf algebra structure on $\mathrm{H}^*(SO(n-1), \mathbb{F}_2)$ can be shown to be primitive [8, Page 298], meaning the terms in the sum over j above are not actually present, but we do not need to make use of this fact.

using the latter $\iota(k)$. (This is the form of Frankel's relationship in §3 and [6].)

Definition A.1. For $n \in \mathbb{Z}^+$, set $[n] := \{1, \ldots, n\}$. For $S \subseteq [n]$, set $\sum S := \sum_{s \in S} s$,

$$\deg S := |\{(p,q) \in ([n] \setminus S) \times S : p < q\}| = \left(\sum S\right) - \binom{|S|+1}{2},$$

$$\operatorname{sdeg} S := \binom{|S|}{2} + \deg S = |\{(p,q) \in [n] \times S : p < q\}|.$$

If n is not clear from context we will write $\deg_n S$ (resp. $\operatorname{sdeg}_n S$) for $\deg S$ (resp. $\operatorname{sdeg} S$). Define

$$\begin{aligned} C_i(k,n) &:= \{S \subseteq [n] : |S| = k,\ \deg S = i\} \\ C_i(n) &:= \{S \subseteq [n] : \operatorname{sdeg} S = i\} \\ c_i(k,n) &:= |C_i(k,n)| \\ c_i(n) &:= |C_i(n)| \\ b_i(n) &:= |\{S \subseteq [n-1] : \sum S = i\}|. \end{aligned}$$

The numbers $b_i(n)$ defined above were also defined in §5, where it was shown that $\dim \mathrm{H}^i(SO(n), \mathbb{F}_2) = b_i(n)$.

It is well-known that $\dim \mathrm{H}^i(\mathbb{G}(k,n), \mathbb{F}_2) = c_i(k,n)$, hence the RHS of (A.1) is nothing but $c_i(n)$. Frankel's relationship is therefore equivalent to the purely combinatorial formula

$$2b_i(n) = c_i(n). \tag{A.2}$$

To establish (A.2), first note that, when $n = 1$, both sides of (A.2) are 2 (resp. 0) when $i = 0$ (resp. when $i \neq 0$). (We have $C_0(1) = \{\emptyset, \{1\}\}$.) As in §5, we see easily that

$$b_i(n) = b_i(n-1) + b_{i+1-n}(n-1) \tag{A.3}$$

for all i and all $n > 1$. Together with the known values of the $b_i(1)$, formula (A.3) determines all the $b_i(n)$. To establish (A.2) it remains only to show that

$$c_i(n) = c_i(n-1) + c_{i+1-n}(n-1) \tag{A.4}$$

for all i and all $n > 1$. To do this, first observe that, for any $S \subseteq [n-1]$ (which we can also regard as a subset of $[n]$), we have

$$\deg_{n-1} S = \deg_n S \tag{A.5}$$

$$\operatorname{sdeg}_{n-1} S = \operatorname{sdeg}_n S. \tag{A.6}$$

Next observe that

$$\begin{aligned} &\{(p,q) \in ([n] \setminus (S \textstyle\coprod \{n\})) \times (S \coprod \{n\}) : p < q\} \\ &= \{(p,q) \in ([n-1] \setminus S) \times S : p < q\} \textstyle\coprod \{(p,n) : p \in ([n-1] \setminus S)\}. \end{aligned}$$

Taking cardinalities yields

$$\deg_n(S \coprod \{n\}) = n - 1 - |S| + \deg_{n-1} S. \tag{A.7}$$

Adding (A.7) and the formula

$$\binom{|S|+1}{2} = \binom{|S|}{2} + |S|$$

yields

$$\operatorname{sdeg}_n(S \coprod \{n\}) = n - 1 + \operatorname{sdeg}_{n-1} S. \tag{A.8}$$

By (A.6) the rule $S \mapsto S$ defines a map $C_i(n-1) \to C_i(n)$ and by (A.8) the rule $S \mapsto S \coprod \{n\}$ defines a map $C_{i+1-n}(n-1) \to C_i(n)$. The coproduct of these two maps yields a map

$$C_i(n-1) \coprod C_{i+1-n}(n-1) \to C_i(n)$$

which is clearly bijective because any $S \in C_i(n)$ either contains n or doesn't. Taking cardinalities yields the recursion (A.4).

For $S \subseteq [n]$, let $\sigma_S := \operatorname{Diag}(\epsilon_1, \ldots, \epsilon_n)$, where ϵ_i is -1 if $i \in S$ and 1 otherwise. Fix real numbers $\lambda_1, \ldots, \lambda_n$ with $0 < \lambda_1 < \cdots < \lambda_n$ and set $A := \operatorname{Diag}(\lambda_1, \ldots, \lambda_n)$. In [3, §3], Duan shows that the function $f : O(n) \to \mathbb{R}$ defined by $f(X) = -\operatorname{Tr}(AX)$ is Morse, with critical points given by the σ_S, and with the index of σ_S given by $(\sum S) - |S|$.[4] In [3, Theorem 6] it is also shown that f is "nice enough" that, with $\mathbb{F}_2$ coefficients, the boundary maps in the Morse complex are zero, hence $\dim \mathrm{H}^i(O(n), \mathbb{F}_2)$ is just the number of critical points of f of index i. Frankel's relationship (A.1) is easily obtained from Duan's results via the computation

$$\begin{aligned}
\sum_k \dim \mathrm{H}^{i-\iota(k)}(\mathbb{G}(k,n), \mathbb{F}_2) &= \sum_k c_{i-\iota(k)}(k,n) \\
&= \sum_k |\{S \subseteq [n] : |S| = k,\ \deg S = i - \binom{k}{2}\}| \\
&= \sum_k |\{S \subseteq [n] : |S| = k,\ (\textstyle\sum S) - |S| = i\}| \\
&= |\{S \subseteq [n] : (\textstyle\sum S) - |S| = i\}| \\
&= \dim \mathrm{H}^i(O(n), \mathbb{F}_2),
\end{aligned}$$

where the only non-obvious inequality is the last one, which follows from the aforementioned results of Duan.

References

[1] D.M. Austin, P.J. Braam, Morse-Bott theory and equivariant cohomology, in: The Floer Memorial Volume, vol. 133, 1995, pp. 123–183.

[2] P.A. Absil, R. Mahony, J. Trumpf, An extrinsic look at the Riemannian Hessian, in: Geometric Science of Information, in: Lecture Notes in Comput. Sci., vol. 8085, Springer, Heidelberg, 2013, pp. 361–368.

[3] H. Duan, Morse functions and cohomology of homogeneous spaces, in: Topics in Cohomological Studies of Algebraic Varieties: Impanga Lecture Notes, Birkhauser, 2005, pp. 105–133.

[4] J.J. Duistermaat, J.A.C. Kolk, V.S. Varadarajan, Functions, flows and oscillatory integrals on flag manifolds and conjugacy classes in real semisimple Lie groups, Compos. Math. 49 (3) (1983) 309–398.

[4] As mentioned earlier, Duan's function is actually twice this one plus a constant, but that is of no consequence.

[5] A.P. Veselov, I.A. Dynnikov, Integrable gradient flows and Morse theory, St. Petersburg Math. J. 8 (3) (1997) 429–446.
[6] T. Frankel, Critical submanifolds of the classical groups and Stiefel manifolds, in: 1965 Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 37–53.
[7] A. Gómez-Tato, E. Macías-Virgós, M.J. Pereira-Sáez, Trace map, Cayley transform and LS category of Lie groups, Ann. Glob. Anal. Geom. 39 (3) (2011) 325–335.
[8] A. Hatcher, Algebraic Topology, Cambridge University Press, Cambridge, 2002.
[9] U. Helmke, J.B. Moore, Optimization and Dynamical Systems, Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1994.
[10] E. Macías-Virgós, M.J. Pereira-Sáez, Height functions on compact symmetric spaces, Monatshefte Math. 177 (1) (2015) 119–140.
[11] E. Macías-Virgós, M.J. Pereira-Sáez, D. Tanré, Morse theory and the Lusternik-Schnirelmann category of quaternionic Grassmannians, Proc. Edinb. Math. Soc. 60 (2) (2017) 441–449.
[12] Anant R. Shastri, B. Subhash, Linear Morse functions on orbit spaces, J. Ramanujan Math. Soc. 25 (2) (2010) 147–167.
[13] J. von Neumann, Some matrix-inequalities and metrization of matric-space, Mitt. Forsch.-Inst. Math. Mech. Univ. Tomsk 1 (1937) 286–299.