On the differential equation $\dot{\Theta} = (\Theta^T - \Theta)\Theta$ with $\Theta \in SO(n)$

Gerd S. Schmidt, Christian Ebenbauer, Frank Allgöwer
Institute for Systems Theory and Automatic Control

May 11, 2014

Abstract

In this note we consider the global convergence properties of the differential equation $\dot{\Theta} = (\Theta^T - \Theta)\Theta$ with $\Theta \in SO(n)$, which is a gradient flow of the function $f : SO(n) \to \mathbb{R}, \Theta \mapsto 2n - 2t(\Theta)$. Many of the presented results are not new, but scattered throughout literature. The motivation of this note is to summarize and extend the convergence results known from literature. Rather than giving an exhaustive list of references, the results are presented in a self-contained fashion.

In this note, we discuss the properties of a function and a differential equation on a smooth manifold. If we speak about a manifold $\mathcal{M}$ of dimension $m$ we always mean a smooth manifold in the sense of [1], i.e. the subset of some $\mathbb{R}^k$ with $k \geq m$ and $\mathcal{M}$ is locally diffeomorphic to $\mathbb{R}^m$. In the context of this note, we need the notions of measure zero and dense. A set $A \subset \mathcal{M}$ of a manifold $\mathcal{M}$ is a set of measure zero if there is a collection of smooth charts $\{U_l, \phi_l\}$ whose domains cover $A$ and such that $\phi_l(A \cap U_l)$ have measure zero in $\mathbb{R}^n$, i.e. the $\phi_l(A \cap U_l)$ can be covered for any $\varepsilon$ by a countable collection of open balls whose volumes sum up to less than $\varepsilon$, for details see e.g. [3, Chapter 10]. To define dense, we need the topological closure $\overline{A}$ of a set $A \subset \mathcal{M}$, i.e. the intersection of all closed sets in $\mathcal{M}$ that contain $A$. A dense subset of a smooth manifold $\mathcal{M}$ is a set $A \subset \mathcal{M}$ such that the topological closure fulfills $\overline{A} = \mathcal{M}$, see e.g. [2, Appendix, Topology].

Here, we consider a function and a differential equation on the set of special orthogonal matrices $SO(n) = \{\Theta \in \mathbb{R}^{n \times n} | \Theta^T = \Theta, \det(\Theta) = 1\}$. $SO(n)$ is a smooth manifold of dimension $\frac{n(n-1)}{2}$ with the subspace topology induced by $\mathbb{R}^{n \times n}$. The tangent space $T_{\Theta} SO(n)$ at $\Theta$ is given by

$$T_{\Theta} SO(n) = \{X \in \mathbb{R}^{n \times n} | X = \Omega \Theta, \Omega \in \mathbb{R}^{n \times n}, \Omega = -\Omega^T\}.$$  

The Riemannian metric $g : T_{\Theta_0} SO(n) \times T_{\Theta_0} SO(n) \to \mathbb{R}$ induced by the standard Euclidean metric on $SO(n)$ is given by

$$g(\Omega_1, \Theta_1 \Theta_2 \Theta) = \text{tr} \left( (\Omega_1 \Theta_2)^T \Omega_2 \Theta \right) = \text{tr} \left( \Omega_1^T \Omega_2 \right).$$

In the following, we define the differential and the Hessian of a function $f : SO(n) \to \mathbb{R}$ at a point $\Theta_0 \in SO(n)$. Let $\Gamma : (-\varepsilon, \varepsilon) \to SO(n)$ be a smooth curve with $\Gamma(t) = \Theta(t)\Gamma(t)$, $\dot{\Gamma}(0) = \Theta_0$ and $\Gamma(0) = \Omega_0 \Theta_0$ with $\Omega_0 \in \mathbb{R}^{n \times n}$ and $\Theta_0 = -\Omega_0^T$. The differential $d f_{\Theta_0} : T_{\Theta_0} SO(n) \to T_f(\Theta_0) \cong \mathbb{R}$ of a function $f : SO(n) \to \mathbb{R}$ at a point $\Theta_0$ evaluated at $\Omega_0 \Theta_0 \in T_{\Theta_0} SO(n)$ is defined by

$$d f_{\Theta_0}(\Omega_0 \Theta_0) = \frac{d}{dt} |_{t=0} (f \circ \Gamma)(t).$$
The critical points of \( f \) are the points \( \Theta_0 \) where \( df_{\Theta_0} \) is not surjective. Because of \( \dim(\text{Ker}(d_f(\Theta_0))) = 1 \), this means that these are the points \( \Theta_0 \) where \( df_{\Theta_0} = 0 \). The gradient of \( f \) is defined as the unique vector field \( \text{grad} \ f \) with

\[
df_{\Theta_0}(\Omega_0 \Theta_0) = g(\text{grad} f(\Theta_0), \Omega_0 \Theta_0),
\]

see e.g. \( \text{[2, Chapter 11]} \). The Hessian \( H_f(\Theta_0) \) of \( f \) at a critical point \( \Theta_0 \) evaluated at \( (\Theta_0, \Theta_0) \) is defined by

\[
H_f(\Theta_0)(\Omega_0 \Theta_0, \Omega_0 \Theta_0) = \frac{df}{dt}igr|_{t=0}(f \circ \Gamma)(t).
\]

Since the Hessian at a critical point is bilinear and symmetric, we have for \( (\Omega_1 + \Omega_2) \Theta_0 \in T_{\Theta_0} \text{SO}(n) \) with \( \Omega_{1,2} = -\Omega^T_{1,2} \) the equality

\[
H_f(\Theta_0)((\Omega_1 + \Omega_2) \Theta_0, (\Omega_1 + \Omega_2) \Theta_0)
= H_f(\Theta_0)(\Omega_1 \Theta_0, \Omega_1 \Theta_0) + 2H_f(\Theta_0)(\Omega_1 \Theta_0, \Omega_2 \Theta_0) + H_f(\Theta_0)(\Omega_2 \Theta_0, \Omega_2 \Theta_0).
\]

As a consequence, the value \( H_f(\Theta_0)(\Omega_2 \Theta_0, \Omega_2 \Theta_0) \) can be computed utilizing the values \( H_f(\Theta_0)(\Omega_2 \Theta_0, \Omega_1 \Theta_0) \) and \( H_f(\Theta_0)(\Omega_1 \Theta_0, \Omega_2 \Theta_0) \). For details on the Hessian at a critical point, see \( \text{[4, Appendix C.5]} \). 

**Lemma 1** Consider the function \( f : \text{SO}(n) \to \mathbb{R}, \Theta \to n - \text{tr} \theta \).

a) The differential \( df_{\Theta_0} \) of \( f \) at \( \Theta_0 \) is given for any \( \Omega_0 \Theta_0 \in T_{\Theta_0} \text{SO}(n) \) by

\[
df_{\Theta_0}(\Omega_0 \Theta_0) = -\frac{1}{2} \text{tr} \left( \Theta_0(\Theta_0 - \Theta_0^T) \Omega_0 \Theta_0 \right)
\]

and the critical points of \( f \) are given by

\[
\mathcal{F} = \{ \Theta_0 \in \text{SO}(n) \mid \Theta_0^T = \Theta_0 \}.
\]

Furthermore, the gradient \( \text{grad} f(\Theta_0) \) at \( \Theta_0 \) is given by

\[
\text{grad} f(\Theta_0) = \frac{1}{2}(\Theta_0 - \Theta_0^T) \Theta_0.
\]

b) The Hessian \( H_f(\Theta_0) \) at a critical point \( \Theta_0 \) is given by

\[
H_f(\Theta_0)(\Omega_1 \Theta_0, \Omega_2 \Theta_0) = \frac{1}{2} \text{tr} \left( \Omega_1^T \Theta_0 \Omega_2 + \Omega_2^T \Theta_0 \Omega_1 \right).
\]

c) The set of critical points \( \mathcal{F} \) has the following properties:

i) \( \mathcal{F} = \bigsqcup_{k=0}^{\frac{2n}{2}} \mathcal{F}_k \) where

\[
\mathcal{F}_k = \{ \Theta_0 \in \text{SO}(n) \mid \Theta_0^T = \Theta_0, \text{ tr } \Theta_0 = n - 4k \}.
\]

ii) Each \( \mathcal{F}_k \) is connected and isolated, i.e. there exists a neighborhood \( U \) of each \( \mathcal{F}_k \) such that \( U \cap \mathcal{F}_l = \emptyset \) for all \( l \neq k \).

iii) \( \mathcal{F}_k \) are compact submanifolds of dimension \( 2k(n-2k) \) and the tangent space at \( \Theta_0 \in \mathcal{F}_k \) is \( T_{\Theta_0} \mathcal{F}_k = \{ \Sigma \in \mathbb{R}^{n \times n} \mid \Sigma = \Sigma^T \} \cap T_{\Theta_0} \text{SO}(n) \).

iv) For every \( k \in \{ 0, 1, \ldots, \frac{n}{2} \} \) and every \( \Theta_0 \in \mathcal{F}_k \) we have

\[
\ker H_f(\Theta_0) = \{ X \in T_{\Theta_0} \text{SO}(n) \mid H_f(\Theta_0)(X,Y) = 0 \text{ for all } Y \in T_{\Theta_0} \text{SO}(n) \} = T_{\Theta_0} \mathcal{F}_k.
\]

d) If \( f \) has a unique minimum at \( \Theta_0 = 1 \), the other critical points are saddle points.
Corollary 2 The differential equation

\[ \dot{\Theta} = (\Theta^T - \Theta)\Theta \]

is the gradient flow of \( f : SO(n) \to \mathbb{R}, \Theta \mapsto 2n - 2\text{tr}(\Theta) \) with respect to the Riemannian metric \((\ref{eq:metric})\).

In the following we prove Lemma 1.

Proof. As stated above, \( \Gamma : (-\varepsilon, \varepsilon) \to SO(n) \) is a differentiable curve with \( \dot{\Gamma}(t) = \Omega(t)\Gamma(t), \dot{\Gamma}(0) = \Theta_0 \) and \( \dot{\Gamma}(0) = \Omega_0\Theta_0 \) with \( \Omega_0 \in \mathbb{R}^{n \times n} \) and \( \Theta_0 = -\Omega_0^T \). Then

\[
df_{\Theta_0}(\Omega_0\Theta_0) = d\left|_{t=0} \right. (n - \text{tr}(\Gamma(t))) = -\text{tr}(\Omega_0\Theta_0)
= -\frac{1}{2} \text{tr}\left( \Theta_0^T\Omega_0^T + \Omega_0\Theta_0 \right)
= -\frac{1}{2} \text{tr}\left( (\Theta_0 - \Theta_0^T)\Omega_0 \right)
= -\frac{1}{2} \text{tr}\left( \Theta_0^T(\Theta_0 - \Theta_0^T)\Omega_0 \right).
\]

Therefore, the critical points of \( f \) are given by

\[
\{ \Theta_0 \in SO(n) | \Theta_0 = \Theta_0^T \}.
\]

With the definition of the Riemannian metric by \((\ref{eq:metric})\), the gradient \( \text{grad } f \) at \( \Theta_0 \) is given by

\[
\text{grad } f(\Theta_0) = \frac{1}{2}(\Theta_0 - \Theta_0^T)\Theta_0^T.
\]

(7) Let \( \Theta_0 \) denote a critical point of \( f \). As stated above, \( \Gamma : (-\varepsilon, \varepsilon) \to SO(n) \) is a differentiable curve with \( \dot{\Gamma}(t) = \Omega(t)\Gamma(t), \dot{\Gamma}(0) = \Theta_0 \) and \( \dot{\Gamma}(0) = \Omega_0\Theta_0 \) with \( \Omega_0 \in \mathbb{R}^{n \times n} \) and \( \Theta_0 = -\Omega_0^T \). Then

\[
H_f(\Theta_0)(\Omega_0\Theta_0, \Omega_0\Theta_0) = \left. \frac{d^2}{dt^2} \right|_{t=0} (f(\Gamma(t))) = -\text{tr}(\dot{\Gamma}(t))|_{t=0}
= -\text{tr}(\dot{\Omega}(t)\Gamma(t) + \Omega(t)\dot{\Gamma}(t))|_{t=0}
= -\text{tr}(\Omega_0^T\Theta_0 = \text{tr}\left( \Omega_0^T\Theta_0 \Omega_0 = \text{tr}\left( \Theta_0^T\Omega_0^T\Theta_0\Omega_0 \right)\right),
\]

where we utilized that \( \text{tr}(\Omega(t)\Theta_0) = 0 \) since \( \Omega(t) \) is skew symmetric for all \( t \) and \( \Theta_0 = \Theta_0^T \) since \( \Theta_0 \) is a critical point. Utilizing \((\ref{eq:metric})\) we get

\[
H_f(\Theta_0)(\Omega_1\Theta_0, \Omega_2\Theta_0) = \frac{1}{2}\left(H_f(\Theta_0)((\Omega_1 + \Omega_2)\Theta_0, (\Omega_1 + \Omega_2)\Theta_0) - H_f(\Theta_0)(\Omega_2\Theta_0, \Omega_2\Theta_0)\right)
= \frac{1}{2} \text{tr}\left( (\Omega_1 + \Omega_2)^T\Theta_0(\Omega_1 + \Omega_2) - \Omega_1^T\Theta_0\Omega_1 - \Omega_2^T\Theta_0\Omega_2 \right)
= \frac{1}{2} \text{tr}\left( \Omega_1^T\Theta_0\Omega_2 + \Omega_2^T\Theta_0\Omega_1 \right).
\]

(8) Since \( \Theta_0 \) is symmetric, \( \Theta_0 \) is orthogonally diagonalizable, i.e., \( \Theta_0 = \Pi^T\Lambda\Pi \) for some diagonal \( D \) and orthonormal \( \Pi \) where the columns of \( \Pi \) are eigenvectors of \( \Theta_0 \). Since \( \Pi^T\Theta_0\Pi = I \), we get \( D^2 = I \) and consequently the eigenvalues are \( \pm 1 \). Since \( \Theta_0 \neq I \) and \( \det(\Theta_0) \neq 1 \), we always have an even number of negative eigenvalues. A similarity transformation leaves the trace invariant, hence a critical point \( \Theta_0 \) fulfills

\[
\text{tr}(\Theta_0) = n - 4k,
\]
where $k \in \{0, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \}$ is the number of eigenvalue pairs which are −1. [ii] We start by showing that each $\mathcal{F}_k$ is path connected and thus connected. Let $k \in \{0, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \}$ be arbitrary but fixed and let $\Theta_1, \Theta_2 \in \mathcal{F}_k$. Then there are orthogonal $\Pi_1, \Pi_2$ such that $\Theta_1 = \Pi_1^T D \Pi_1$ and $\Theta_2 = \Pi_2^T D \Pi_2$. Furthermore, there are real skew-symmetric matrices $\Omega_1, \Omega_2$ such that $\Pi_1 = \exp(\Omega_1)$ and $\Pi_2 = \exp(\Omega_2)$ with exp denoting the matrix exponential. Then $\alpha : [0, 1] \to SO(n)$ defined by

$$t \mapsto \exp \left( \Omega_1^T (1-t) \right) \exp \left( \Omega_2^T t \right) D \exp(\Omega_2 t) \exp(\Omega_1 (1-t))$$

is a smooth curve in $\mathcal{F}_k$ which connects $\Theta_1$ and $\Theta_2$. Since $\Theta_1, \Theta_2 \in \mathcal{F}_k$ were arbitrary, this implies the path-connectedness of $\mathcal{F}_k$. To show that $\mathcal{F}_k$ is isolated, we utilize that $n - 4 l = f |_{X \neq \emptyset} f |_{X} = n - 4 k$ for $l \neq k$. Then there is a $\epsilon(l)$ with $(n - 4 l - \epsilon(l), n - 4 l + \epsilon(l)) \cap (n - 4 k - \epsilon(l), n - 4 k + \epsilon(l)) = \emptyset$ for $k \neq l$. As a consequence, the intersection of the preimage of these sets under $f$ is empty. Since $f$ is continuous and both, $(n - 4 l - \epsilon(l), n - 4 l + \epsilon(l))$ and $(n - 4 k - \epsilon(l), n - 4 k + \epsilon(l))$ are open, their preimages are open and contain $\mathcal{F}_l$ and $\mathcal{F}_k$ respectively. With $U_k(l) = f^{-1}((n - 4 l - \epsilon(l), n - 4 l + \epsilon(l)))$ we thus have $U_k(l) \cap \mathcal{F}_l = \emptyset$. Since this is possible for every $l \in \{0, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \}$ and since a finite intersection of open sets is an open set, we find an open neighborhood $U$ of $\mathcal{F}_k$ such that $U \cap \mathcal{F}_l = \emptyset$ for all $l \in \{0, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \}$ and since the property that the $\mathcal{F}_k$ are submanifolds is given in [5]. The tangient space follows from [7].

Let $k \in \{0, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \}$ be arbitrary but fixed and $\Theta_0 \in \mathcal{F}_k$. Since $\ker H_f(\Theta_0) \cap T_{\Theta_0} \mathcal{F}_k$ is always true, we have to check $\ker H_f(\Theta_0) \subset T_{\Theta_0} \mathcal{F}_k$. Since every critical point is symmetric, there is an orthogonal $\Pi$ such that $\Pi^T \Theta_0 \Pi = D$ where $D$ is a diagonal matrix with non-zero diagonal elements. Furthermore, we know that

$$H_f(\Theta_0)(\Omega_2 \Theta_0, \Omega_2 \Theta_0) = -\frac{1}{2} \text{tr} \left( \Omega_1^T \Theta_0 \Omega_2 + \Omega_1^T \Theta_0 \Omega_2 \right)$$

which $\tilde{\Omega}_1 = \Pi^T \Omega_1$ and $\tilde{\Omega}_2 = \Pi^T \Omega_2$. As a consequence

$$\ker H_f(\Theta_0) = \left\{ X \in T_{\Theta_0} SO(n) \mid H_f(\Theta_0)(X, Y) = 0 \text{ for all } Y \in T_{\Theta_0} SO(n) \right\} = \left\{ X \in T_{\Theta_0} SO(n) \mid \text{tr} \left( X^T D \tilde{\Omega}_2^2 D - D \tilde{\Omega}_2^2 DX \right) = 0 \text{ for all } \tilde{\Omega}_2 D \in T_{\Theta_0} SO(n) \right\} = \left\{ X \in T_{\Theta_0} SO(n) \mid \text{tr} \left( \tilde{\Omega}_2 D(X^T - X)D \right) = 0 \text{ for all } \tilde{\Omega}_2 D \in T_{\Theta_0} SO(n) \right\}.$$
Since $\Theta_0 = \Theta_0^T$, $\Theta_0$ is orthogonally diagonalizable, i.e. $\Theta_0 = \Pi^T D \Pi$ for some diagonal $D$ and orthogonal $\Pi$. Therefore

$$H_f((\Omega_0 \Theta_0), (\Omega_0 \Theta_0)) = \text{tr} \left( \Omega_0^T \Theta_0 \Omega_0 \right) = - \text{tr} \left( \Omega_0^T D \right),$$

where $\Omega_0 = \Pi \Omega_0 \Pi^T$ is skew symmetric. Consequently, $\text{tr} \left( \Omega_0^T \Theta_0 \Omega_0 \right)$ is definite for all skew symmetric $\Omega_0$ at a critical point $\Theta_0$ if and only if $\text{tr} \left( \Omega_0^T D \right)$ is definite for all skew symmetric $\Omega_0$ where $D = \Pi \Theta_0 \Pi^T$ is diagonal. Thus, we have to consider $H_f$ only for diagonal $D$, i.e.

$$H_f((\Omega_0 \Theta_0), (\Omega_0 \Theta_0)) = - \text{tr} \left( \Omega_0^T D \right) = - \sum_{1 \leq k \leq n} (\Omega_0^T)_k \, D_{kk} = - \sum_{1 \leq k \leq n} (\Omega_0)_k (\Omega_0)_k D_{kk}$$

$$= \sum_{1 \leq k \leq n} ((\Omega_0)_k)^2 \, D_{kk} = \sum_{1 \leq k \leq n} ((\Omega_0)_k)^2 \, D_{kk} \quad \text{where} \quad D_{kk} = D_{kk} + D_{ll} ((\Omega_0)_k)^2.$$  

The critical points $\Theta_0$ are such that $\Theta_0 \in \text{SO}(n)$ are symmetric, therefore all eigenvalues of $\Theta_0$ are real. Observe now that $\Theta_0 = \Pi^T D \Pi$ with orthogonal $\Pi$ implies $D^2 = I$. Hence, the eigenvalues are $D_{kk} \in \{-1, 1\}$ and since $\Theta_0 \in \text{SO}(n)$ the number of $-1$-eigenvalues is even. Consequently we have to determine the definiteness of $H_f$ by considering (19) for all diagonal matrices $D$ with $\pm 1$ on the diagonals where the number of $-1$ entries is zero or even.

Suppose first, that all $D_{kk}$ are equal to 1, i.e. $D = I$. The associated $\Theta_0$ is $\Theta_0 = \Pi^T D \Pi = \Pi^T \Pi = I$. Then, then implies $H_f((\Omega_0 \Theta_0), (\Omega_0 \Theta_0)) = \sum_{1 \leq k \leq n} 2 ((\Omega_0)_k)^2$, i.e. $H_f > 0$ for all skew-symmetric $\Omega_0$. Thus $H_f$ is positive definite if $\Theta_0 = I$. Suppose now there is an even number of eigenvalues $D_{kk}$ equal to $-1$. Then, there are indices $l, k$ such that $D_{kk} = -1$ and $D_{ll} = 1$, and therefore there are skew symmetric $\Omega_0$ such that $H_f((\Omega_0 \Theta_0), (\Omega_0 \Theta_0)) < 0$. As consequence, $H_f$ is indefinite at a critical point $\Theta_0$ where $\Theta_0$ has an even number of negative eigenvalues $D_{kk} = -1$. Therefore, $\Theta_0 = I$ is the only local (global) minimum of $f$. All other critical points are saddle points.}

**Definition 3 (on p.21)** Let $\mathcal{M}$ be a smooth Riemannian manifold and $f : \mathcal{M} \rightarrow \mathbb{R}$ be a smooth function. Denote the set of critical points of $f$ by $C(f)$. $f$ is called Morse-Bott function provided the following conditions are satisfied:

a) $f$ has compact sublevel sets.

b) $C(f) = \bigcup_{j=1}^k \mathcal{N}_j$ where $\mathcal{N}_j$ are disjoint, closed and connected submanifolds of $\mathcal{M}$ and $f$ is constant on $\mathcal{N}_j$ for $j = 1, \ldots, k$.

c) $\ker H_f(x) = T_x \mathcal{N}_j$ for all $x \in \mathcal{N}_j$ and all $j = 1, \ldots, k$.

**Lemma 4** $f : \text{SO}(n) \rightarrow \mathbb{R}, \Theta \mapsto n - \text{tr}(\Theta)$ is a Morse-Bott function.

**Proof.** We show only Definition 3, since 2 and 3 were shown in Lemma 1 $\text{SO}(n)$ is compact, hence $f$ attains its minimal and its maximal value on $\text{SO}(n)$. The minimal value of $f$ is zero, the maximal value is $2n - 2$ for $n$ odd and $2n$ for $n$ even. If $n$ is odd we thus have

$$L_\varepsilon = \{ \Theta \in \text{SO}(n) | f(\Theta) \leq c \} = \begin{cases} f^{-1}([0, c]) & \text{for } c \leq 2n - 2 \\ f^{-1}([0, 2n - 2]) & \text{for } c > 2n - 2. \end{cases}$$

If $n$ is even we have

$$L_\varepsilon = \{ \Theta \in \text{SO}(n) | f(\Theta) \leq c \} = \begin{cases} f^{-1}([0, c]) & \text{for } c \leq 2n \\ f^{-1}([0, 2n]) & \text{for } c > 2n. \end{cases}$$
Since \( f \) is continuous, the preimage of a closed set is a closed set and since \( SO(n) \) is bounded, its subsets are bounded as well. Since \( SO(n) \subset \mathbb{R}^{n \times n} \), the boundedness and closedness of the sublevel sets implies their compactness. The convergence properties of the gradient flow associated with \( \Theta \mapsto n - \text{tr}(\Theta) \) are thus determined by the following proposition.

**Proposition 5** [4, Proposition 3.9] Let \( f : \mathcal{M} \to \mathbb{R} \) be a Morse-Bott function on a Riemannian manifold \( \mathcal{M} \). The \( \omega \)–limit set \( \omega(x) \) of \( x \in \mathcal{M} \) with respect to the gradient flow of \( f \) is a single critical point of \( f \). Every solution of the gradient flow converges to an equilibrium point.

To give a more detailed specification the convergence behavior of the gradient flow of \( \Theta \mapsto n - \text{tr}(\Theta) \) we need the following result.

**Lemma 6** Let \( \mathcal{M} \) be a smooth and compact Riemannian manifold of dimension \( m \), \( f : \mathcal{M} \to \mathbb{R} \) a Morse-Bott function and denote the set of critical points of \( f \) by \( \mathcal{C} \). Let \( \mathcal{N} \) be a fixed connected component of \( C(f) \) of dimension \( n \). If at least one of the \( m - n \) eigenvalues with nonzero real part of the linearization of \( \nabla f \) at some \( x \in \mathcal{N} \) has a real part greater than zero, then the set \( A \) of initial conditions \( x_0 \in \mathcal{M} \) for which the solutions \( t \mapsto \Phi(t,x_0) \) of the gradient flow \( \dot{x} = - \nabla f(x) \) converge towards \( \mathcal{N} \), i.e.

\[
A = \{ x_0 \in \mathcal{M} | \lim_{t \to \infty} \Phi(t,x_0) \in \mathcal{N} \},
\]

has measure zero. Furthermore \( \mathcal{M} \setminus A \) is dense in \( \mathcal{M} \), i.e. \( \overline{\mathcal{M} \setminus A} = \mathcal{M} \).

**Proof.** The goal of the proof is to show that \( A \) has measure zero and that \( \mathcal{M} \setminus A \) is dense. We show this in the following way. First, we consider the set of points lying in a suitable neighborhood of \( \mathcal{N} \) and which contains the orbits of the solutions of the gradient flow \( \dot{x} = - \nabla f(x) \) which eventually converge towards \( \mathcal{N} \). We utilize a result from [4] to conclude that this set has measure zero and \( \mathcal{M} \setminus A \) without this set is dense. Then, we utilize this set to derive the same result for \( A \) utilizing the properties of the flow of the gradient vector field on \( \mathcal{M} \).

In the following, we apply [4, Proposition 4.1]. This proposition concerns the case of \( n \) times continuously differentiable vector field \( v : \mathbb{R}^l \to \mathbb{R}^l \) together with submanifold of equilibria \( \mathcal{N} \) in \( \mathbb{R}^l \) under the assumption that \( \mathcal{N} \) is normally hyperbolic with respect to \( v \). Normal hyperbolicity of \( \mathcal{N} \) means that the linearization of the vector field \( v \) at \( x \in \mathcal{N} \) has \( n - \text{dim} \mathcal{N} \) eigenvalues with real parts different from zero. Under these assumptions, there exists a neighborhood \( \mathcal{U} \) of \( \mathcal{N} \) such that any solution \( t \mapsto \Phi(t,x_0) \) of \( \dot{x} = v(x) \) with initial condition \( x_0 \) and with a forward orbit \( \Phi([0,\infty);x_0) \) in \( \mathcal{U} \) lies on the stable manifold \( W^s_{\text{loc}}(p) \) of a point \( p \in \mathcal{U} \). \( W^s_{\text{loc}}(p) \) is defined by

\[
W^s_{\text{loc}}(p) = \{ x \in \mathcal{U} | \lim_{t \to +\infty} \Phi(t,x) = p \}.
\]

We can always embed \( \mathcal{M} \) into \( \mathbb{R}^l \) for \( l \) large enough, see e.g. [4, Chapter 10], therefore we can utilize [4] also for our case of a vector field on a manifold \( \mathcal{M} \).

Since \( \mathcal{M} \) is compact and \( f \) is smooth, we have a global flow \( \Phi : \mathbb{R} \times \mathcal{M} \to \mathcal{M} \), which means that \( t \mapsto \Phi(t,x_0) \) is a solution of the gradient flow \( \dot{x} = - \nabla f(x) \) defined for all \( t \in \mathbb{R} \) and with \( \Phi(0,x_0) = x_0 \), see e.g. [4, Chapter 17]. Furthermore \( \Phi(t,\cdot) : \mathcal{M} \to \mathcal{M} \) is a diffeomorphism for every \( t \in \mathbb{R} \). Since \( f \) is a Morse-Bott function, \( \mathcal{N} \) is normally hyperbolic, i.e. the linearization of the gradient flow at any \( x \in \mathcal{N} \) has exactly \( m - n \) eigenvalues with real parts different from zero, see [4, p. 183]. According to [4, Proposition 4.1], we have a neighborhood \( \mathcal{U} \) of \( \mathcal{N} \) such that for every solution \( \Phi(t,x) \) with \( x \in \mathcal{N} \) and a forward orbit \( \Phi([0,\infty);x_0) \) in \( \mathcal{U} \), the solution has to lie in one \( W^s_{\text{loc}}(p) \) with \( p \in \mathcal{N} \). We know from [4, Proposition 3.2], that if we choose \( \mathcal{U} \) small enough, then the local stable manifold \( W^s_{\text{loc}}(\mathcal{N}) \) of \( \mathcal{N} \) given by

\[
W^s_{\text{loc}}(\mathcal{N}) = \bigcup_{p \in \mathcal{N}} W^s_{\text{loc}}(p)
\]

is a smooth submanifold of dimension \( m + k \) where \( k < m - n \) is the number of eigenvalues with real part smaller than zero. Since the stable manifold \( W^s_{\text{loc}}(\mathcal{N}) \) is a submanifold of \( \mathcal{M} \) with smaller dimension than \( \mathcal{M} \), the stable manifold has measure zero and \( \mathcal{M} \setminus W^s_{\text{loc}}(\mathcal{N}) \) is dense in \( \mathcal{M} \), see [4, Theorem 10.5].
Let $A$ be defined by (20). Define $A_1$ by

$$A_1 = \{ x \in A | \forall t: \theta(t,x) \in \mathcal{W} \}$$

and let $A_k$ for $k \geq 2$ be defined by

$$A_k = \{ x \in A \setminus (A_1 \cup \ldots \cup A_{k-1}) | \forall t: \theta(t,x) \in \mathcal{W} \}.$$  

If $x \in A$, then there is an integer $k \in \mathbb{N}$ such that $x \in A_k$. As a consequence

$$A = \cup_{k \in \mathbb{N}} A_k.$$  

Because of (24), $\theta(k, A_k) \subset \mathcal{W}$ for every $A_k$. Moreover, [2, Proposition 4.1] implies that

$$\phi(k, A_k) \subset W^s_{\text{loc}}(\mathcal{N}).$$  

As subset of a set of measure zero, $\theta(k, A_k)$ has measure zero, see e.g. [2, Lemma A.60(b)]. Since $\theta(k, \cdot): \mathcal{M} \rightarrow \mathcal{M}$ is a diffeomorphism, this means that $A_k$ also has measure zero, see e.g. [2, Lemma 10.1]. According to (25), $A$ is a countable union of $A_k$, i.e. $A$ is a countable union of sets of measure zero. Therefore, $A$ has measure zero and as a consequence $\mathcal{M} \setminus A$ is dense.

To finally derive the global stability properties of the identity matrix $I \in \mathbb{R}^{n \times n}$ for the gradient flow of $\Theta \mapsto n - \text{tr}(\Theta)$ and thus also for the differential equation (10), we linearize the gradient flow around the equilibria.

**Lemma 7** The convergence properties of the gradient flow of $f: SO(n) \rightarrow \mathbb{R}, \Theta \mapsto n - \text{tr}(\Theta)$ are the following:

a) The $\omega$-limit set of any solution is contained in the set of equilibria given by (1), i.e.

$$\mathcal{F} = \{ \Theta_0 \in SO(n) | \Theta_0^T = \Theta_0 \}.$$  

b) The equilibrium $I$ is locally exponentially stable and all other equilibria are unstable.

c) The set of initial conditions for which the solutions of the gradient flow $\dot{x} = -\text{grad } f(x)$ of $f$ converge towards $I$ is dense in $SO(n)$ and the set of initial conditions for which the solutions of the gradient flow converge to the other equilibria has measure zero.

**Corollary 8** The identity matrix $I$ is an almost globally asymptotically stable equilibrium for the differential equation

$$\dot{\Theta} = (\Theta^T - \Theta) \Theta.$$  

In the following we prove Lemma 7.

**Proof.** [2] is a consequence of Lemma 4 and Proposition 5. To prove the property [2] we linearize the gradient flow with the vector field $\text{grad } f$ defined by $\Theta \mapsto \frac{1}{2}(\Theta^T - \Theta) \Theta$ around the equilibria. To do this directly we compute

$$d\text{grad } f_{\Theta_0}(X) = \frac{1}{2} \frac{d}{dt} |_{t=0} (\text{grad } f) \circ \Gamma(t)$$  

where $\Theta_0$ is an equilibrium and $\Gamma: (-\varepsilon, \varepsilon) \rightarrow SO(n)$ is smooth with $\Gamma(0) = \Theta_0, \Gamma'(0) = X$ and $X \in T_{\Theta_0}SO(n)$. This yields

$$d\text{grad } f_{\Theta_0}(X) = \frac{1}{2} \left( \Gamma^T(t) \Gamma'(t) + \Gamma'(t) \Gamma(t) - \Gamma'(t) \Gamma(t) \right) |_{t=0}$$  

$$= \frac{1}{2} (\Theta^T \Theta_0 + \Theta_0^T X - X \Theta_0 - \Theta_0 X)$$  

$$= \frac{1}{2} (\Theta_0 X + X \Theta_0).$$

7
since $\Theta_0 = \Theta_0^T$ and $X = \Omega_0 \Theta_0$ for a $\Omega_0 \in \mathbb{R}^{n \times n}$ with $\Omega_0 = -\Omega_0^T$. Consequently the linearization of the gradient flow at an equilibrium $\Theta_0$ is given by

$$
\dot{X} = -\frac{1}{2} \left( \Theta_0^T X + X \Theta_0 \right),
$$

(29)

where $X \in T_{\Theta_0} SO(n)$. Note that due to the simple nature of the Riemannian metric (2) and the connection of the linearization of a gradient flow to the Hessian, we could have obtained (29) directly from (14). More precisely, utilize

$$
H_f(\Theta_0)(X,X) = \text{tr}(X^T \dot{\Theta}_0 X) = \text{vec}(\Theta_0 X^T) \text{vec}(X)
$$

(30)

$$
= \text{vec}(X)^T (I \otimes \Theta_0 + (I \otimes \Theta_0)^T) \text{vec}(X)
$$

$$
= -\frac{1}{2} \text{vec}(X)^T \text{vec}(\Theta_0 X + X \Theta_0).
$$

If $\Theta_0 = I$, then the linearization is

$$
\dot{X} = -X,
$$

(31)

which shows that the equilibrium $I$ is locally exponentially stable. Now consider the linearization at the other equilibrium points, i.e. $\Theta_0 \neq I$ and $\Theta_0 = \Theta_0^T$. Since $\Theta_0$ is symmetric, $\Theta_0$ is orthonormally diagonalizable, i.e. $\Theta_0 = \Pi^T D \Pi$ for some diagonal $D$ and orthonormal $\Pi$ where the columns of $\Pi$ are eigenvectors of $\Theta_0$. Since $\Theta_0^T \Theta_0 = I$, we get $D^T = I$ and consequently the eigenvalues are $\pm 1$. Since $\Theta_0 \neq I$ and $\det(\Theta_0) = 1$, we always have an even number of negative eigenvalues with associated eigenvectors $v_1, \ldots, v_k$. Set $\overline{U} = v_1 v_1^T - v_2 v_2^T$ and $X = \overline{U} \Theta_0 \in T_{\Theta_0} SO(n)$. Therefore

$$
-\frac{1}{2} (\Theta_0 X + X \Theta_0) = -\frac{1}{2} \left( \Theta_0^T \overline{U} \Theta_0 + \overline{U} \Theta_0 \Theta_0 \right)
$$

(32)

$$
= -\frac{1}{2} \left( \Theta_0 (v_1 v_1^T - v_2 v_2^T) \Theta_0 + (v_1 v_1^T - v_2 v_2^T) \Theta_0 \Theta_0 \right)
$$

$$
= -\frac{1}{2} \left( (-v_1 v_1^T + v_2 v_2^T) \Theta_0 + (-v_1 v_1^T + v_2 v_2^T) \Theta_0 \right) = \overline{U} \Theta_0.
$$

Therefore, $X = \overline{U} \Theta_0$ is an eigenvector of the operator defined by the right hand side of (29). Since the associated eigenvalue is positive (one), the linearization (29) is unstable. Consequently, the linearization of the gradient flow at the equilibria $\Theta_0$ with $\Theta_0 \neq I$ and $\Theta_0 = \Theta_0^T$ is unstable, which proves (b). Denote the flow of $\dot{\Theta} = -\text{grad} f(\Theta)$ by $\phi: \mathbb{R} \times SO(n) \to SO(n)$ and by $B_k$ the set

$$
B_k = \{ x_0 \in SO(n) | \lim_{t \to \infty} \phi(t,x_0) \in F_k \},
$$

(33)

i.e. the set of initial conditions that converges to the connected component $F_k$ of the set of critical points $F$ given in Lemma [3]. Because of Proposition [5] we are certain that any solution of the gradient flow converges to the critical set of $f$ and as a consequence,$SO(n) = \bigcup_{k=0}^{\lfloor n^2/2 \rfloor} B_k$. Then $B_0$ is the set of initial conditions for which the flow converges to $I$ and $B = \bigcup_{k=1}^{n^2/2} B_k$ is the set of initial conditions for which the flow converges to any of the other critical points. In Lemma [5] we showed that $B_k$ has measure zero and that $SO(n) \setminus B_k$ is dense in $SO(n)$. Since $B$ is the union of a finite number of sets of measure zero, it has measure zeros, see e.g. [2, Lemma 10.1]. In particular $B_0 = SO(n) \setminus B$ is dense.

**Bibliography**

[1] V. Guillemin, Differential Topology, Prentice-Hall, Inc., 1974.
[2] J. M. Lee, Introduction to Smooth Manifolds, Vol. 218 of Graduate Texts in Mathematics, Springer, 2006.

[3] J. Milnor, Topology from the differentiable viewpoint, Princeton Landmarks in Mathematics, Princeton University Press, 1997, originally published by University Press of Virginia, 1965.

[4] U. Helmke, J. B. Moore, Optimization and Dynamical Systems, Springer, 1994.

[5] T. Frankel, Critical submanifolds of the classical groups and Stiefel manifolds, in: S. S. Cairns (Ed.), Differential and Combinatorial Topology — A Symposium in Honor of Marston Morse, Princeton University Press, 1965, pp. 37–54.

[6] B. Aulbach, Continuous and Discrete Dynamics near Manifolds of Equilibria, Vol. 1058 of Lecture Notes in Mathematics, Springer, 1984.

[7] D. McDuff, D. Salamon, Introduction to Symplectic Topology, Clarendon Press, 1998.

[8] D. M. Austin, P. J. Braam, Morse-Bott theory and equivariant cohomology, in: H. Hofer, C. H. Taubes, A. Weinstein, Z. Eduard (Eds.), The Floer Memorial Volume, Vol. 133 of Progress in Mathematics, Birkhäuser, 1995, pp. 123–184.