JOURNAL OF GEOMETRIC MECHANICS
doi:10.3934/jgm.2020031

Volume **13**, Number **1**, March **2021**
pp. **55–72**

A LAGRANGIAN APPROACH TO EXTREMAL CURVES ON STIEFEL MANIFOLDS

Knut Hüper*

Institute of Mathematics, Julius-Maximilians-Universität Würzburg, Germany

Irina Markina

Department of Mathematics, University of Bergen, Norway

Fátima Silva Leite

Institute of Systems and Robotics, and Department of Mathematics
University of Coimbra, Portugal

Abstract. A unified framework for studying extremal curves on real Stiefel manifolds is presented. We start with a smooth one-parameter family of pseudo-Riemannian metrics on a product of orthogonal groups acting transitively on Stiefel manifolds. In the next step Euler-Langrange equations for a whole class of extremal curves on Stiefel manifolds are derived. This includes not only geodesics with respect to different Riemannian metrics, but so-called quasi-geodesics and smooth curves of constant geodesic curvature, as well. It is shown that they all can be written in closed form. Our results are put into perspective to recent related work where a Hamiltonian rather than a Lagrangian approach was used. For some specific values of the parameter we recover certain well-known results.

1. **Introduction.** The main objective of this paper is to present a unified framework for studying extremal curves on the important class of real Stiefel manifolds, using a direct variational approach.

This paper is a tribute to Anthony M. Bloch on the occasion of his 65th anniversary. His work in the area of geometric mechanics and optimal control has been inspirational to us. Without being exhaustive, we name [4, 5, 6, 7, 16], for some of his important contributions in those areas.

In general, the problem of finding extremal curves, e.g., curves that minimize or maximize some functional is an old problem which goes back to the origins of variational calculus. The present paper deals with geometric extremal curves, i.e., those curves that minimize the length functional or in other words find the geodesics on a specific manifold. The manifold we consider here is the set of all orthonormal k-frames on Euclidean n-dimensional space, nowadays known under the name *Stiefel manifold*, cf. [26]. This problem has been considered by several authors, see for instance [7, 9, 11, 18]. The difference in these approaches, is the realization of the Stiefel manifold: it was considered as a homogeneous space under

2020 *Mathematics Subject Classification.* Primary: 58E10, 53C22, 53C30.

Key words and phrases. Extremal curves, geodesics, indefinite metric, Killing form, Lagrange multiplier, smooth distributions, Stiefel manifold, variational calculus.

* Corresponding author.

the action of different groups, or as an embedded submanifold of the space of all $(n \times k)$-matrices. The length functional also varied.

In the present paper we propose to consider a one-parameter family of non-degenerate not necessarily positive definite metrics, that includes most of the previously considered Riemannian metrics. Such an all-embracing approach becomes possible due to the realization of the Stiefel manifold as a homogeneous space under the transitive action of the product group of orthogonal matrices $\mathrm{O}_n \times \mathrm{O}_k$, see Sections 3 and 4. The considered metrics lead to the length functionals. By making use of variational calculus, as a consequence, we obtain a one-parameter family of nonlinear, second order, matrix-valued, ordinary differential equations (ODEs), that surprisingly can be reduced to a one-parameter family of linear, autonomous, second order, matrix ODEs. The solutions of the corresponding one-parameter family of initial value problems, i.e., geodesics, are calculated in closed form for all parameters defining the family of metrics. Besides new formulas and insights, we recover several of the beforementioned previous results. We note that the obtained formulas are written for an arbitrary choice of the isotropy point for the action of $\mathrm{O}_n \times \mathrm{O}_k$, see Section 5. We mention another important feature of our approach. All the closed formulas for extremal curves we derive here share a remarkable property. They all include matrix exponentials exclusively of skew symmetric matrices, predestinated for numerically stable implementations. The skew symmetric matrices to be exponentiated are either of size $k \times k$ or $n \times n$, sometimes of both.

As a further byproduct of our approach we describe a huge family of curves that have constant geodesic curvature with respect to metrics within the one-parameter family. In this part we use a sub-Riemannian approach, that allows to relate the geodesic curves on the Stiefel manifold to special curves on the group $\mathrm{O}_n \times \mathrm{O}_k$ being tangent to distinguished distributions on $\mathrm{O}_n \times \mathrm{O}_k$, associated to a Cartan decomposition of the Lie algebra $\mathfrak{so}_n \times \mathfrak{so}_k$, see Section 6.

2. **Setting.** The orthogonal group O_n in its standard representation is denoted by

$$\mathrm{O}_n := \{Q \in \mathbb{R}^{n \times n} \mid Q^\top Q = I_n\}, \tag{1}$$

where I_n denotes the $(n \times n)$-identity matrix with $n \in \mathbb{N}$. Accordingly, we have the special orthogonal group defined by

$$\mathrm{SO}_n := \{R \in \mathrm{O}_n \mid \det(R) = 1\} \cong \mathrm{O}_n / \mathrm{O}_1 . \tag{2}$$

The Lie algebra of O_n and SO_n, i.e., the set of real skew symmetric $(n \times n)$-matrices, is denoted by

$$\mathfrak{so}_n := \{X \in \mathbb{R}^{n \times n} \mid X = -X^\top\}. \tag{3}$$

The tangent space of O_n, analogously for SO_n, at Q is then

$$T_Q \, \mathrm{O}_n \cong \mathfrak{so}_n \, Q \cong Q \, \mathfrak{so}_n . \tag{4}$$

In the sequel we will mainly use the first isomorphism in (4). The real compact Stiefel manifold (from now on just *Stiefel manifold*) can be defined as the set of orthonormal k-frames in $\mathbb{R}^n$, i.e.,

$$\mathrm{St}_{n,k} := \{X \in \mathbb{R}^{n \times k} \mid X^\top X = I_k\}, \quad 1 \leq k \leq n. \tag{5}$$

The tangent space $T_X \, \mathrm{St}_{n,k}$ is characterized by

$$T_X \, \mathrm{St}_{n,k} = \mathfrak{so}_n \cdot X = \{Z \in \mathbb{R}^{n \times k} \mid X^\top Z \in \mathfrak{so}_k\} = \mathrm{im}(\pi_X^{\tan}), \tag{6}$$

with projection operator

$$\pi_X^{\tan}\colon \mathbb{R}^{n\times k} \to T_X \operatorname{St}_{n,k}, \quad Z \mapsto \Big(I_n - \tfrac{XX^\top}{2}\Big)Z - \tfrac{XZ^\top X}{2}. \tag{7}$$

Besides (5), in a more abstract way, the Stiefel manifold can be considered in several diffeomorphic ways. By applying the regular value theorem to the function $f\colon \mathbb{R}^{n\times k} \to \operatorname{Sym}_k$ defined by $X \mapsto X^\top X - I_k$, where

$$\operatorname{Sym}_k := \{X \in \mathbb{R}^{k\times k} | X = X^\top\} \tag{8}$$

denotes the vector space of symmetric $(k \times k)$-matrices, we see that $\operatorname{St}_{n,k} = f^{-1}(0)$ is a smooth Riemannian submanifold of dimension $\dim \operatorname{St}_{n,k} = nk - k(k+1)/2$. Here $\mathbb{R}^{n\times k}$ is seen as the Riemannian manifold endowed with Euclidean metric, i.e., with the usual Frobenius inner product on each tangent space $T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$

$$\langle \cdot,\cdot\rangle\colon \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \to \mathbb{R}, \quad \langle A, B\rangle := \operatorname{tr}(A^\top B). \tag{9}$$

Two possible further points of view are as follows. Firstly, the group O_n acts transitively on $\operatorname{St}_{n,k}$ via left matrix multiplication

$$\sigma\colon\ \mathrm{O}_n \times \operatorname{St}_{n,k} \to \operatorname{St}_{n,k}, \quad (Q, X) \mapsto QX, \tag{10}$$

with associated map

$$\sigma_X\colon\ \mathrm{O}_n \to \operatorname{St}_{n,k}, \quad Q \mapsto \sigma_X(Q) := \sigma(Q, X) = QX. \tag{11}$$

Secondly, the product group $\mathrm{O}_n \times \mathrm{O}_k$ acts also transitively on $\operatorname{St}_{n,k}$ via

$$\tau\colon\ \mathrm{O}_n \times \mathrm{O}_k \times \operatorname{St}_{n,k} \to \operatorname{St}_{n,k}, \quad \big((U, V), X\big) \mapsto UXV^\top, \tag{12}$$

and associated map

$$\tau_X\colon\ \mathrm{O}_n \times \mathrm{O}_k \to \operatorname{St}_{n,k}, \quad Q \mapsto \tau_X(U, V) := \tau\big((U, V), X\big) = UXV^\top. \tag{13}$$

The smooth maps σ_X and τ_X give rise to two quotient models, namely

$$\operatorname{St}_{n,k} \cong_{\mathrm{O}_n} \mathrm{O}_n \,/\, \mathrm{O}_{n-k}, \qquad \operatorname{St}_{n,k} \cong_{\mathrm{O}_n\times\mathrm{O}_k} (\mathrm{O}_n \times \mathrm{O}_k)/(\mathrm{O}_{n-k} \times \mathrm{O}_k). \tag{14}$$

For the arbitrary chosen $X \in \operatorname{St}_{n,k}$ the associated isotropy subgroup fixing X is isomorphic to O_{n-k}, respectively to $\mathrm{O}_{n-k} \times \mathrm{O}_k$.

Instead of the O_n, respectively $\mathrm{O}_n \times \mathrm{O}_k$, action, we could consider SO_n, respectively $\mathrm{SO}_n \times \mathrm{SO}_k$, actions. However, the situation in this case is a bit more subtle, as e.g., SO_n does not act transitively on $\operatorname{St}_{n,n} \cong_{\mathrm{O}_n} \mathrm{O}_n$. We mention the special cases, e.g., $\operatorname{St}_{n,1} \cong S^{n-1}$, $\operatorname{St}_{n,n-1} \cong \mathrm{SO}_n$ and $\operatorname{St}_{n,k} \cong_{\mathrm{SO}_n} \mathrm{SO}_n \,/\, \mathrm{SO}_{n-k}$ for $1 \le k \le n-1$, cf. [23].

3. **Metrics on Stiefel manifolds.** By the transitivity of σ, resp. τ, it follows that σ_X, resp. τ_X, are submersions for all $X \in \operatorname{St}_{n,k}$ and in particular the derivatives (tangent maps)

$$\begin{aligned} \mathrm{D}\,\sigma_X(I_n)\colon\ \mathfrak{so}_n &\to T_X \operatorname{St}_{n,k}, \\ \Omega &\mapsto \Omega X, \end{aligned} \tag{15}$$

$$\begin{aligned} \mathrm{D}\,\tau_X(I_n, I_k)\colon\ \mathfrak{so}_n \times \mathfrak{so}_k &\to T_X \operatorname{St}_{n,k}, \\ (\Omega, \Psi) &\mapsto \Omega X + X\Psi^\top = \Omega X - X\Psi \end{aligned} \tag{16}$$

are both surjective linear maps for all $X \in \operatorname{St}_{n,k}$. These facts can be exploited to define metrics on $\operatorname{St}_{n,k}$. The corresponding constructions have several names in the literature, (i) submersion metrics (for obvious reasons) or (ii) normal metrics (becomes clear below). Note however, that in case of σ_X the resulting metric was called ‘canonical’, cf. [9].

Recall that the Killing form on the Lie algebra $\mathfrak{so}_n$, cf. [10], is the symmetric bilinear form defined as

$$B\colon\ \mathfrak{so}_n \times \mathfrak{so}_n \to \mathbb{R}, \quad B(X,Y) := (n-2)\operatorname{tr}(XY). \tag{17}$$

For $n > 2$ the Killing form is nondegenerate as $\mathfrak{so}_n$ in this case has trivial center. Moreover, it is negative definite as the semisimple group O_n is compact. Consequently, one can take the negative of the Killing form in order to define a left invariant Riemannian metric on O_n as follows

$$\begin{aligned} &\langle\cdot,\cdot\rangle_{\mathrm{O}_n}\colon T_Q\,\mathrm{O}_n \times T_Q\,\mathrm{O}_n \to \mathbb{R},\\ &\qquad (\Omega_1 Q, \Omega_2 Q) \mapsto -\tfrac{B(\Omega_1,\Omega_2)}{n-2} = -\operatorname{tr}(\Omega_1\Omega_2) = \operatorname{tr}\big((\Omega_1 Q)^\top(\Omega_2 Q)\big). \end{aligned} \tag{18}$$

Clearly, this Riemannian metric can equally be considered as the one which is induced by the Euclidean (Frobenius) metric of the embedding space $\mathbb{R}^{n\times n} \supset \mathrm{O}_n$. We will now focus on the group action τ defined by (12). As the linear map $\mathrm{D}\,\tau_X(I_n, I_k)$ is surjective for all $X \in \mathrm{St}_{n,k}$ it induces an isometry between $\ker^\perp \mathrm{D}\,\tau_X(I_n, I_k)$ and $T_X\,\mathrm{St}_{n,k}$ under an additional assumption on the value of α, see below. We proceed as follows.

Definition 3.1. Define the following one-parameter family of indefinite inner products on $\mathfrak{so}_n \times \mathfrak{so}_k$

$$\begin{aligned} &\langle\cdot,\cdot\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k}\colon (\mathfrak{so}_n\times\mathfrak{so}_k)\times(\mathfrak{so}_n\times\mathfrak{so}_k) \to \mathbb{R},\\ &\qquad \big((\Omega_1,\Psi_1),(\Omega_2,\Psi_2)\big) \mapsto -\operatorname{tr}\big(\Omega_1\Omega_2\big) - \tfrac{1}{\alpha}\operatorname{tr}(\Psi_1\Psi_2),\ \alpha \in \mathbb{R}\setminus\{0\}. \end{aligned} \tag{19}$$

Remark 1. The inner product defined in (19) is ad-invariant, that is, for all $A, B, C \in \mathfrak{so}_n \times \mathfrak{so}_k$

$$-\big\langle A, \operatorname{ad}_C B\big\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k} = \big\langle A, [B,C]\big\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k} = \big\langle B, [C,A]\big\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k} = \big\langle B, \operatorname{ad}_C A\big\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k}. \tag{20}$$

By making use of left translations the inner product (19) extends to a bi-invariant (pseudo-)Riemannian metric on $\mathrm{O}_n \times \mathrm{O}_k$

$$\begin{aligned} &\langle\cdot,\cdot\rangle^{(\alpha)}_{\mathrm{O}_n\times\mathrm{O}_k}\colon T_{(U,V)}(\mathrm{O}_n\times\mathrm{O}_k)\times T_{(U,V)}(\mathrm{O}_n\times\mathrm{O}_k) \to \mathbb{R},\\ &\qquad \big((\Omega_1 U, \Psi_1 V), (\Omega_2 U, \Psi_2 V)\big) \mapsto \big\langle(\Omega_1,\Psi_1),(\Omega_2,\Psi_2)\big\rangle^{(\alpha)}_{\mathfrak{so}_n\times\mathfrak{so}_k}\\ &\qquad\qquad = -\operatorname{tr}\big(\Omega_1\Omega_2\big) - \tfrac{\operatorname{tr}(\Psi_1\Psi_2)}{\alpha}. \end{aligned} \tag{21}$$

For $\alpha > 0$ this defines a Riemannian metric on the product group, whereas for $\alpha < 0$ the metric becomes indefinite, i.e. pseudo-Riemannian.

At this stage we recall some terminology and facts from indefinite linear algebra, cf. [12]. Let $\big(V, \langle\cdot,\cdot\rangle_{\mathrm{indef}}\big)$ be an indefinite inner product space. A subspace $S \subset V$ is called nondegenerate with respect to $\langle\cdot,\cdot\rangle_{\mathrm{indef}}$ if $s \in S$ and $\langle s, s'\rangle_{\mathrm{indef}} = 0$ for all $s' \in S$ imply that $s = 0$. The orthogonal *companion* of a *subset* $S \subset V$ is defined by

$$S^\perp := \big\{v \in V \,\big|\, \langle v, s\rangle_{\mathrm{indef}} = 0 \text{ for all } s \in S\big\}. \tag{22}$$

Clearly, $S^\perp$ is a subspace of V. The following proposition is proved e.g. in [12].

Proposition 1. *$S^\perp$ is a direct (orthogonal) complement to S in V iff S is nondegenerate.* □

By means of $\mathrm{D}\,\tau_X(I_n, I_k)$ we will now define a one-parameter family of metrics on $\mathrm{St}_{n,k}$. Denote by K_X the kernel of $\mathrm{D}\,\tau_X(I_n, I_k)$

$$K_X := \ker \mathrm{D}\,\tau_X(I_n, I_k) = \big\{(\Omega,\Psi) \in \mathfrak{so}_n \times \mathfrak{so}_k \;\big|\; \Omega X - X\Psi = 0\big\}. \tag{23}$$

Proposition 2. *For any given $X \in \mathrm{St}_{n,k}$ the subspace K_X is nondegenerate iff $\alpha \neq -1$.*

Proof. Note that $(\Omega, \Psi), (\Gamma, \Xi) \in K_X$ imply $X^\top \Omega X = \Psi$ and $X^\top \Gamma X = \Xi$ as $X \in \mathrm{St}_{n,k}$. By the orthogonal invariance of the trace function we might assume $X = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right]$. We partition $\Omega = \left[\begin{smallmatrix} \Omega_{11} & \Omega_{12} \\ -\Omega_{12}^\top & \Omega_{22} \end{smallmatrix}\right]$ and $\Gamma = \left[\begin{smallmatrix} \Gamma_{11} & \Gamma_{12} \\ -\Gamma_{12}^\top & \Gamma_{22} \end{smallmatrix}\right]$ with $\Omega_{11}, \Gamma_{11} \in \mathfrak{so}_k$. By exploiting properties of the trace function and the fact that $XX^\top = \left[\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right]$ is an orthogonal projection matrix we get

$$
\begin{aligned}
\big\langle (\Omega,\Psi), (\Gamma,\Xi) \big\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} &= -\operatorname{tr}(\Omega\Gamma) - \tfrac{1}{\alpha} \operatorname{tr}(\Psi\Xi) \\
&= -\operatorname{tr}(\Omega\Gamma) - \tfrac{1}{\alpha} \operatorname{tr}\big(\Omega XX^\top \Gamma XX^\top\big) \\
&= -\Big(\operatorname{tr}\big(\tfrac{\alpha+1}{\alpha} \Omega_{11}\Gamma_{11}\big) - 2\operatorname{tr}(\Omega_{12}\Gamma_{12}^\top) + \operatorname{tr}(\Omega_{22}\Gamma_{22}) \Big).
\end{aligned}
\tag{24}
$$

Hence, a given $(\Omega, \Psi) \in K_X$ with $\big\langle (\Omega,\Psi), (\Gamma,\Xi) \big\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} = 0$ for all $(\Gamma, \Xi) \in K_X$ implies $(\Omega, \Psi) = (0, 0)$ iff $\alpha \neq -1$. The result follows. ☐

For $\alpha \neq -1$ we denote by $K_X^\perp$ the orthogonal complement of K_X with respect to the metric (19)

$$
K_X^\perp := \big\{ (\Omega, \Psi) \in \mathfrak{so}_n \times \mathfrak{so}_k \,\big|\, \big\langle (\Omega, \Psi), K_X \big\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} = 0 \big\}. \tag{25}
$$

Consequently, for every $(\Omega, \Psi) \in \mathfrak{so}_n \times \mathfrak{so}_k$ with $\alpha \neq -1$ there is a unique additive decomposition

$$
(\Omega, \Psi) = \Big(\Omega^{K_X} + \Omega^{K_X^\perp}, \Psi^{K_X} + \Psi^{K_X^\perp} \Big), \tag{26}
$$

where $\big(\Omega^{K_X}, \Psi^{K_X}\big) \in K_X$ and $\big(\Omega^{K_X^\perp}, \Psi^{K_X^\perp}\big) \in K_X^\perp$.

Lemma 3.2. *For $\alpha \neq -1$ and arbitrary $X \in \mathrm{St}_{n,k}$, the unique orthogonal projection operator*

$$
\pi_X \colon \mathfrak{so}_n \times \mathfrak{so}_k \to K_X, \quad (\Omega, \Psi) \mapsto \big(\Omega^{K_X}, \Psi^{K_X}\big) \tag{27}
$$

is given by

$$
\begin{aligned}
\Omega^{K_X} &= \big(I_n - XX^\top\big)\Omega\big(I_n - XX^\top\big) + \tfrac{\alpha}{\alpha+1} XX^\top \Omega XX^\top + \tfrac{1}{\alpha+1} X\Psi X^\top \\
&= \Omega - \Big(\Omega XX^\top + XX^\top \Omega - \tfrac{2\alpha+1}{\alpha+1} XX^\top \Omega XX^\top - \tfrac{1}{\alpha+1} X\Psi X^\top \Big), \\
\Psi^{K_X} &= \tfrac{\alpha}{\alpha+1} X^\top \Omega X + \tfrac{1}{\alpha+1} \Psi \\
&= \Psi - \tfrac{\alpha}{\alpha+1} \big(\Psi - X^\top \Omega X\big).
\end{aligned}
\tag{28}
$$

The complementary orthogonal projection operator is as

$$
\begin{aligned}
\pi_X^\perp := \mathrm{id} - \pi_X \colon\ & \mathfrak{so}_n \times \mathfrak{so}_k \to K_X^\perp, \\
& (\Omega, \Psi) \mapsto \big(\Omega^{K_X^\perp}, \Psi^{K_X^\perp}\big) = \big(\Omega - \Omega^{K_X}, \Psi - \Psi^{K_X}\big),
\end{aligned}
\tag{29}
$$

given by

$$
\begin{aligned}
\Omega^{K_X^\perp} &= XX^\top \Omega + \Omega XX^\top - \tfrac{2\alpha+1}{\alpha+1} XX^\top \Omega XX^\top - \tfrac{1}{\alpha+1} X\Psi X^\top, \\
\Psi^{K_X^\perp} &= \tfrac{\alpha}{\alpha+1} \big(\Psi - X^\top \Omega X\big).
\end{aligned}
\tag{30}
$$

Proof. In fact, a direct computation shows that $\pi_X^2 = \pi_X$ exploiting the fact that $I_n - XX^\top$ and $XX^\top$ denote complementary orthogonal projection operators $\mathbb{R}^{n\times k} \to \mathbb{R}^{n\times k}$. Moreover, by using (28) and several times $X^\top X = I_k$, we compute

$$\begin{aligned}\Omega^{K_X} X - X\Psi^{K_X} &= \tfrac{\alpha}{\alpha+1} XX^\top \Omega X + \tfrac{1}{\alpha+1} X\Psi - X\big(\tfrac{\alpha}{\alpha+1} X^\top \Omega X + \tfrac{1}{\alpha+1}\Psi\big)\\ &= 0,\end{aligned} \tag{31}$$

as required. □

Remark 2. Note that K_X does not depend on the choice of the α-metric as it is defined by the α-independent linear map $\mathrm{D}\,\tau_X(I_n, I_k)$, only. The projection operator π_X, however, depends on the α-metric, therefore we have the presence of the parameter α in the components of $\big(\Omega^{K_X}, \Psi^{K_X}\big) \in K_X$.

Corollary 1. *The special case $X = \left[\begin{smallmatrix} I_k \\ 0\end{smallmatrix}\right]$ might be of interest. Partition an arbitrary*

$$(\Omega, \Psi) = \left(\begin{bmatrix} A & -B^\top \\ B & C\end{bmatrix}, \Psi\right) \in \mathfrak{so}_n \times \mathfrak{so}_k,\ A \in \mathfrak{so}_k,\ B \in \mathbb{R}^{k\times(n-k)},\ C \in \mathfrak{so}_{n-k}, \tag{32}$$

accordingly. Then for this case, the explicit form of the decomposition (26) is as

$$(\Omega, \Psi) = \underbrace{\left(\begin{bmatrix} \frac{\Psi+\alpha A}{\alpha+1} & 0 \\ 0 & C\end{bmatrix}, \frac{\Psi+\alpha A}{\alpha+1}\right)}_{=:(\Omega^{K_X}, \Psi^{K_X})\in K_X} + \underbrace{\left(\begin{bmatrix} \frac{A-\Psi}{\alpha+1} & -B^\top \\ B & 0\end{bmatrix}, \frac{\alpha(\Psi-A)}{\alpha+1}\right)}_{=:(\Omega^{K_X^\perp}, \Psi^{K_X^\perp})\in K_X^\perp}. \tag{33}$$

Proof. Indeed,

$$\begin{aligned}\Big\langle\Big(\Omega_1^{K_X^\perp}, \Psi_1^{K_X^\perp}\Big), \Big(\Omega_2^{K_X}, \Psi_2^{K_X}\Big)\Big\rangle &= \left\langle\left(\begin{bmatrix} \frac{A_1-\Psi_1}{\alpha+1} & -B_1^\top \\ B_1 & 0\end{bmatrix}, \frac{\alpha(\Psi_1-A_1)}{\alpha+1}\right), \left(\begin{bmatrix} \frac{\Psi_2+\alpha A_2}{\alpha+1} & 0 \\ 0 & C_2\end{bmatrix}, \frac{\Psi_2+\alpha A_2}{\alpha+1}\right)\right\rangle\\ &= -\operatorname{tr}\frac{(A_1-\Psi_1)(\Psi_2+\alpha A_2)}{(\alpha+1)^2} - \frac{1}{\alpha}\operatorname{tr}\frac{\alpha(\Psi_1-A_1)(\Psi_2+\alpha A_2)}{(\alpha+1)^2}\\ &= 0.\end{aligned}$$

The claim follows by counting parameters in the matrices $A_1, A_2, B_1, C_2, \Psi_1, \Psi_2$. □

Fix an arbitrary $X \in \mathrm{St}_{n,k}$ and denote

$$\mathfrak{g} := \mathfrak{so}_n \times \mathfrak{so}_k, \quad \mathfrak{k} := K_X, \quad \mathfrak{p} := K_X^\perp. \tag{34}$$

Then one has the following relations

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \supset \mathfrak{k}. \tag{35}$$

The direct sum in the first relation is orthogonal. The second relation reflects the fact that $\mathfrak{k} = K_X$ is the Lie algebra of the isotropy group isomorphic to $\mathrm{O}_{n-k} \times \mathrm{O}_k$ that fixes the point $X \in \mathrm{St}_{n,k}$. The third relation follows by

$$\big\langle [\mathfrak{p}, \mathfrak{k}], \mathfrak{k}\big\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} = \big\langle \mathfrak{p}, [\mathfrak{k}, \mathfrak{k}]\big\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} = 0, \tag{36}$$

due to ad-invariance of the inner product and the first two properties in (35). For the last property it is sufficient to show it at the point $X = \left[\begin{smallmatrix} I_k \\ 0\end{smallmatrix}\right]$. In this case we denote $D := \frac{\Psi+\alpha A}{\alpha+1}$ in (33). Then

$$\left(\left[\begin{smallmatrix} D & 0 \\ 0 & C\end{smallmatrix}\right], D\right) \in K_X = \mathfrak{k}, \qquad \left(\begin{bmatrix} A-D & B \\ -B^\top & 0\end{bmatrix}, \alpha(D-A)\right) \in K_X^\perp = \mathfrak{p}. \tag{37}$$

In the sequel, $E_{i,j}$ denotes the $(n\times n)$-matrix with entry (i,j) equal to 1, entry (j,i) equal to -1 and all other entries equal to 0. The matrices $E_{i,j}$ satisfy the following commutator properties:

$$[E_{i,j}, E_{f,l}] = -\delta_{il} E_{j,f} - \delta_{jf} E_{i,l} + \delta_{if} E_{j,l} + \delta_{jl} E_{i,f},$$

where δ_{ij} denotes the Kronecker delta. Taking into consideration the structure of the matrices in $\mathfrak{k}$ and $\mathfrak{p}$, we see that

$$\{E_{i,j},\ 1 \le i < j \le k\} \cup \{E_{k+i,k+j},\ 1 \le i < j \le n-k\} \tag{38}$$

forms a basis for $\mathfrak{k}$, while

$$\{E_{i,j},\ 1 \le i < j \le k\} \cup \{E_{i,k+j},\ 1 \le i \le k,\, 1 \le j \le n-k\} \tag{39}$$

forms a basis for $\mathfrak{p}$. Moreover, for any $E_{k+i,k+j} \in \mathfrak{k}$ with $1 \le i < j \le n-k$ there are $E_{l,k+i}, E_{l,k+j} \in \mathfrak{p}$ with $1 \le l \le k$ such that

$$E_{k+i,k+j} = [E_{l,k+i}, E_{l,k+j}]\,. \tag{40}$$

Analogously, for any $E_{i,j} \in \mathfrak{k}$, $1 \le i < j \le k$, there are $E_{i,k+l}, E_{j,k+l} \in \mathfrak{p}$ with $1 \le l \le n-k$, such that

$$E_{i,j} = [E_{i,k+l}, E_{j,k+l}]\,.$$

Thus for any chosen $X \in \mathrm{St}_{n,k}$ we have decomposed the Lie algebra $\mathfrak{so}_n \times \mathfrak{so}_k$ into a direct sum, orthogonal with respect to the α-inner product, satisfying (36).

Definition 3.3. Let $X \in \mathrm{St}_{n,k}$ be arbitrary. Consider $\xi_1, \xi_2 \in T_X \mathrm{St}_{n,k}$ with $\xi_i := \Omega_i X - X\Psi_i$, $\Omega_i \in \mathfrak{so}_n$, $\Psi_i \in \mathfrak{so}_k$ and $i \in \{1,2\}$. We define a smooth family of normal or submersion (pseudo-)Riemannian metrics on the Stiefel manifold $\mathrm{St}_{n,k}$ via

$$\begin{aligned}\langle\cdot,\cdot\rangle^{(\alpha)}_{\mathrm{St}} \colon T_X \mathrm{St}_{n,k} \times & T_X \mathrm{St}_{n,k} \to \mathbb{R},\\ &(\xi_1,\xi_2) \mapsto -\operatorname{tr}\big(\Omega_1^{K_X^\perp}\Omega_2^{K_X^\perp}\big) - \tfrac{1}{\alpha}\operatorname{tr}\big(\Psi_1^{K_X^\perp}\Psi_2^{K_X^\perp}\big).\end{aligned} \tag{41}$$

For computational purposes it is certainly desirable to have an explicit formula for $\Omega^{K_X^\perp}$ and $\Psi^{K_X^\perp}$ purely in terms of $\xi \in T_X \mathrm{St}_{n,k}$ and $X \in \mathrm{St}_{n,k}$.

Proposition 3. *Let $X \in \mathrm{St}_{n,k}$ and $\xi \in T_X \mathrm{St}_{n,k}$, with $\xi := \Omega X - X\Psi$. Then,*

$$\begin{aligned}\Omega^{K_X^\perp} &= \xi X^\top - X\xi^\top + \tfrac{2\alpha+1}{\alpha+1} X\xi^\top X X^\top,\\ \Psi^{K_X^\perp} &= -\tfrac{\alpha}{\alpha+1} X^\top \xi.\end{aligned} \tag{42}$$

Proof. Plug $\xi = \Omega X - X\Psi$ into (42) and compare with (30), giving the result. $\square$

Corollary 2. *The normal or submersion (pseudo-)Riemannian metric on the Stiefel manifold $\mathrm{St}_{n,k}$ induced by the group action of $\mathrm{O}_n \times \mathrm{O}_k$ defined by (13) with $\mathrm{O}_n \times \mathrm{O}_k$ endowed with the pseudo-Riemannian metric defined by (21) expressed exclusively by $X \in \mathrm{St}_{n,k}$ and $\xi \in T_X \mathrm{St}_{n,k}$ now boils down to*

$$\big\langle \xi_1, \xi_2 \big\rangle^{(\alpha)}_{\mathrm{St}} = \operatorname{tr}\big(2\xi_1\xi_2^\top\big) - \tfrac{2\alpha+1}{\alpha+1}\operatorname{tr}\big(\xi_1\xi_2^\top X X^\top\big). \tag{43}$$

$\square$

Hereafter we refer to the members of the α-parameter family of metrics as α-metrics.

4. **Special cases of α-metrics.** For certain values of α we now put our results into perspective with more or less recently published works.

4.1. **Case 1: The limit $\alpha \to 0$ and the normal or submersion metric induced by σ.** By taking the limit $\alpha \to 0$ in (43), the right hand side is remarkably still a Riemannian one. Actually this metric coincides with the normal or submersion metric induced by the transitive group action of O_n on $\mathrm{St}_{n,k}$ by left multiplication. Namely in this case we have

$$\lim_{\alpha\to 0} \langle \xi_1, \xi_2\rangle^{(\alpha)}_{\mathrm{St}} = \operatorname{tr}\left(2\xi_1\xi_2^\top\right) - \operatorname{tr}\left(\xi_1\xi_2^\top X X^\top\right). \tag{44}$$

For convenience one might compare the above formula choosing $X = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right]$, $\xi_i = \left[\begin{smallmatrix} A_i \\ B_i \end{smallmatrix}\right] \in T_X \mathrm{St}_{n,k}$ where $A_i \in \mathfrak{so}_k$ and $B_i \in \mathbb{R}^{n-k\times k}$ with formula (2.22) in [9] (up to a factor 1/2), i.e. in the notation of Corollary 1

$$\langle \xi_1, \xi_2\rangle^{(0)}_{\mathrm{St}} = \operatorname{tr}(2A_1A_2^\top + 2B_1B_2^\top) - \operatorname{tr}(A_1A_2^\top) = \operatorname{tr}(A_1A_2^\top + 2B_1B_2^\top). \tag{45}$$

Let us look on the situation a little bit more carefully. Formula (30) implies that $\frac{1}{\alpha}\Psi^{K_\perp} = \frac{1}{1+\alpha}(\Psi - X^T\Omega X) \to \Psi - X^T\Omega X$ as $\alpha \to 0$. Therefore

$$\tfrac{1}{\alpha} \operatorname{tr}\left(\Psi_1^{K_\perp}\Psi_2^{K_\perp}\right) = \alpha \operatorname{tr}\left(\tfrac{1}{\alpha}\Psi_1^{K_\perp} \tfrac{1}{\alpha}\Psi_2^{K_\perp}\right) \to 0 \quad \text{as} \quad \alpha \to 0. \tag{46}$$

It implies that the limit in (41) exists as $\alpha \to 0$ and it is equal to $-\operatorname{tr}\left(\Omega_1^{K_\perp}\Omega_2^{K_\perp}\right) = \operatorname{tr}(A_1A_2^\top + 2B_1B_2^\top)$ by (45). Now consider the case $X = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right]$. Then according to (33) we have

$$\mathfrak{so}_n \times \mathfrak{so}_k \ni (\Omega, \Psi) = \underbrace{\left(\left[\begin{smallmatrix} \Psi & 0 \\ 0 & C \end{smallmatrix}\right], \Psi\right)}_{\in K_X} + \underbrace{\left(\left[\begin{smallmatrix} A-\Psi & -B^\top \\ B & 0 \end{smallmatrix}\right], 0\right)}_{\in K_X^\perp}, \tag{47}$$

where K_X needs to be orthogonal to $K_X^\perp$ with respect to the inner product (19) as $\alpha \to 0$, i.e., if we denote

$$(\Omega_1, \Psi_1) = \left(\left[\begin{smallmatrix} \Psi & 0 \\ 0 & C \end{smallmatrix}\right], \Psi\right), \qquad (\Omega_2, \Psi_2) = \left(\left[\begin{smallmatrix} A-\Psi & -B^\top \\ B & 0 \end{smallmatrix}\right], 0\right), \tag{48}$$

then

$$\left\langle (\Omega_1, \Psi_1), (\Omega_2, \Psi_2)\right\rangle^{(\alpha)}_{\mathfrak{so}_n \times \mathfrak{so}_k} = \operatorname{tr}\left((A-\Psi)\Psi\right) - \tfrac{1}{\alpha}\operatorname{tr}(\Psi \cdot 0) = 0. \tag{49}$$

Since we know that the limit of the metric exists, it will correspond to the metric space where $\Psi = 0$. In this case we obtain

$$\mathfrak{so}_n \times \mathfrak{so}_k \ni (\Omega, 0) = \underbrace{\left(\left[\begin{smallmatrix} 0 & 0 \\ 0 & C \end{smallmatrix}\right], 0\right)}_{\in K_X = \mathfrak{k}} + \underbrace{\left(\left[\begin{smallmatrix} A & -B^\top \\ B & 0 \end{smallmatrix}\right], 0\right)}_{\in K_X^\perp = \mathfrak{p}} \in \mathfrak{so}_n \times \{0\} \cong \mathfrak{so}_n\,. \tag{50}$$

It indeed corresponds to the action of $\mathrm{O}(n)$ on the Stiefel manifold with $\mathfrak{so}_n = \mathfrak{p} \oplus \mathfrak{k}$, see [18, Formulas (17) and (18)].

4.2. **Case 2: $\alpha = -\frac{1}{2}$ and the Euclidean metric induced by $\mathrm{St}_{n,k} \subset \mathbb{R}^{n\times k}$.** For $\alpha = -\frac{1}{2}$ we start with a pseudo-Riemannian metric on $\mathrm{O}_n \times \mathrm{O}_k$ as α is negative. It turns out, its restriction to the horizonal space $K_X^\perp$ becomes Riemannian. Actually, this metric is exactly the one you get by considering $\mathrm{St}_{n,k} \subset \mathbb{R}^{n\times k}$ as a Riemannian submanifold of the Euclidean space $\mathbb{R}^{n\times k}$ endowed with Frobenius inner product. Explicitly, we get

$$\langle \xi_1, \xi_2\rangle^{(-1/2)}_{\mathrm{St}_{n,k}} = \operatorname{tr}(2\xi_1\xi_2^\top), \tag{51}$$

i.e., twice the usual Euclidean inner product on $\mathbb{R}^{n\times k}$. This result is not too surprising. Indeed, consider the group action

$$\beta\colon\ \mathrm{O}_n \times \mathrm{O}_k \times \mathbb{R}^{n\times k} \to \mathbb{R}^{n\times k}, \quad \big((U,V), X\big) \mapsto UXV^\top, \tag{52}$$

with

$$\beta_X\colon\ \mathrm{O}_n \times \mathrm{O}_k \to \mathbb{R}^{n\times k}, \quad (U,V)\mapsto UXV^\top. \tag{53}$$

The orbits of β are precisely those matrix sets which have fixed singular values. Certainly, on each of these orbits the product group acts transitively. Therefore the restriction $\beta_X(I,I)|_{X-\mathrm{orbit}}$ is a submersion for any X. The Stiefel manifold is equal to one of these orbits, the result follows by Proposition 3.

4.3. **Case 3: The limit $\alpha \to \infty$.** This approach gives rise to quasi-geodesic curves on $\mathrm{St}_{n,k}$. For convenience, we only look at $X = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right]$. Indeed, if we take a limit as $\alpha \to \infty$ in (33) we come to the decomposition

$$\mathfrak{so}_n \times \mathfrak{so}_k \ni (\Omega,\Psi) = \underbrace{\left(\left[\begin{smallmatrix} A & 0 \\ 0 & C \end{smallmatrix}\right], A\right)}_{\in K_X = \mathfrak{k}} + \underbrace{\left(\left[\begin{smallmatrix} 0 & -B^\top \\ B & 0 \end{smallmatrix}\right], \Psi - A\right)}_{\in K_X^\perp = \mathfrak{p}}. \tag{54}$$

This corresponds to the quasi-geodesic horizontal distribution in [18, formulas (20) and (21)] generated by the $\mathfrak{p}$-subspace within the Cartan decomposition (35). The inner product that makes the direct sum in (35) orthogonal was not presented in [18] and this case was treated differently. See also [19] for a slightly different perspective.

Remark 3. Note that in the literature *quasi-geodesics* appear with different meanings, not necessarily related in an obvious way.

In numerics, e.g., in [22], those curves are denoted by quasi-geodesics, which are smooth curves approximating geodesics in some sense. Furthermore, in [1] and [20] the connection of this type of quasi-geodesics to a retraction approximating the Riemannian exponential map is made. See, e.g., [2] for the concept of retractions applied to numerics.

The notion of quasi-geodesics in the sense of M. Gromov, however, is derived from the general theory of metric spaces and the notion of quasi-isometries between those spaces. This gives a precise mathematical meaning to *coarse* spaces and *coarse* structures as part of geometry, i.e., studying metric spaces from a large scale point of view, see e.g., [24], [25].

In our paper, the term quasi-geodesic refers to the former concept, i.e., quasi-geodesics are considered to be smooth curves close to geodesics in some sense, although nowhere we comment explicitly on corresponding retractions. More important, however, the curves we consider here, all have constant geodesic curvature.

4.4. **Case 4: $\alpha = 1$.** This case is classical and corresponds to the submersion, where the standard trace inner product on $\mathfrak{so}_n \times \mathfrak{so}_k$ is considered, see, for instance, [11] and [18].

4.5. **Case 5: $\alpha < -1$.** In this case we always get a pseudo-Riemannian metric on $\mathrm{St}_{n,k}$ which is not Riemannian. We are currently not aware of any applications.

Remark 4. We now briefly discuss two extreme cases, namely $\mathrm{St}_{n,1} = S^{n-1}$ and $\mathrm{St}_{n,n} = \mathrm{O}_n$.

For the sphere S^{n-1} the α-metric becomes independent of α. Indeed, (43) reduces to the scaled Euclidean inner product on $T_X S^{n-1} \subset \mathbb{R}^n$ as for $i \in \{1,2\}$ we have $T_X S^{n-1} \ni \xi_i \perp X$. Explicitly,

$$\langle \xi_1, \xi_2 \rangle^{(\alpha)}_{\mathrm{St}_{n,1}} = \operatorname{tr}\left(2\xi_1\xi_2^\top\right) - \tfrac{2\alpha+1}{\alpha+1}\operatorname{tr}\big(\xi_1 \underbrace{\xi_2^\top X}_{=0} X^\top\big) = 2\xi_2^\top \xi_1.$$

For the orthogonal group O_n the α-metric (43) simplifies to

$$\langle \xi_1, \xi_2 \rangle^{(\alpha)}_{\mathrm{St}_{n,n}} = \operatorname{tr}\left(2\xi_1\xi_2^\top\right) - \tfrac{2\alpha+1}{\alpha+1} \operatorname{tr}\Big(\xi_1\xi_2^\top \underbrace{XX^\top}_{=I_n}\Big) = \tfrac{1}{\alpha+1} \operatorname{tr}\left(\xi_1\xi_2^\top\right).$$

In particular, it becomes independent of X. Moreover, for any $\alpha \in \mathbb{R} \setminus \{-1\}$ the metric is either positive definite or negative definite. In other words, it is always a nonzero multiple of the Killing form. In the limit $\alpha \to \infty$ the metric becomes identical to zero.

5. **Geodesics on $\mathrm{St}_{n,k}$ with respect to different values of α.** The purpose of this section is to derive a one-parameter family of Euler-Lagrange equations describing the critical points of the energy functional on $\mathrm{St}_{n,k}$ with respect to the metric (43). It is a remarkable fact that one is able to show, independent of the value of α, that this family of nonlinear, second order matrix ODEs is equivalent to a family of *linear*, time-invariant, second order matrix ODEs. To write down a closed form solution for the corresponding initial value problem is then straightforward.

Rather than solving a variational problem directly on $\mathrm{St}_{n,k}$ where we would need an α-dependent formula of the covariant derivative we proceed with a Lagrange multiplier approach in the space of smooth curves in $\mathbb{R}^{n\times k}$.

Extend the metric (43) to a function $T_X\mathbb{R}^{n\times k} \times T_X\mathbb{R}^{n\times k} \to \mathbb{R}$ in a straightforward way, saying

$$\langle \xi_1, \xi_2 \rangle^{(\alpha)}_{\mathbb{R}^{n\times k}} := \operatorname{tr}\left(2\xi_1\xi_2^\top\right) - \tfrac{2\alpha+1}{\alpha+1} \operatorname{tr}\left(\xi_1\xi_2^\top XX^\top\right) \tag{55}$$

for any $X \in \mathbb{R}^{n\times k}$ and any $\xi_1, \xi_2 \in T_X\mathbb{R}^{n\times k}$. We now study the following problem from variational calculus. Find the critical points of

$$\begin{aligned} F\colon C^\infty(\mathbb{R}, \mathbb{R}^{n\times k} \times \mathrm{Sym}_k) &\to \mathbb{R}, \\ (X,S) &\mapsto \tfrac{1}{2}\int_0^1 \Big(\langle \dot X, \dot X\rangle^{(\alpha)}_{\mathbb{R}^{n\times k}} + \operatorname{tr}\left(S(X^\top X - I_k)\right)\Big)\mathrm{d}t. \end{aligned} \tag{56}$$

Here $S\colon \mathbb{R} \to \mathrm{Sym}_k$ serves as a matrix-valued Lagrange multiplier. Following the usual approach, cf. [28], consider an admissible variational family of the curves X and S fulfilling boundary conditions

$$\begin{aligned} X_\varepsilon(t) &:= X(t) + \varepsilon Y(t) \in \mathbb{R}^{n\times k}, \quad Y(0) = Y(1) = 0, \\ S_\varepsilon(t) &:= S(t) + \varepsilon T(t) \in \mathrm{Sym}_k, \quad T(0) = T(1) = 0. \end{aligned} \tag{57}$$

A critical point of (56) has to satisfy $f'(\varepsilon)|_{\varepsilon=0} = 0$ where

$$f\colon (-\delta,\delta) \to \mathbb{R}, \qquad f(\varepsilon) := F(X_\varepsilon, S_\varepsilon). \tag{58}$$

We can therefore state $f'(0) = 0$, iff for all admissible variations Y, T we have

$$\int_0^1 \Big(\langle \dot X, \dot Y\rangle^{(\alpha)}_{\mathbb{R}^{n\times k}} + \tfrac{1}{2}\operatorname{tr}\left(T(X^\top X - I_k)\right) + \operatorname{tr}(SX^\top Y)\Big)\mathrm{d}t = 0. \tag{59}$$

By partial integration, respecting the boundary conditions, equation (59) is equivalent to the system

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\left((2I_n - \tfrac{2\alpha+1}{\alpha+1}XX^\top)\dot X\right) + \tfrac{2\alpha+1}{\alpha+1}\dot X \dot X^\top X - XS = 0, \\ X^\top X = I_k. \end{aligned} \tag{60}$$

Exploiting the symmetry of S we get rid of S in (60). A byproduct of these calculations is the relation

$$X^\top \ddot X = \ddot X^\top X. \tag{61}$$

Moreover, using also

$$X^\top X = I_k \implies \dot X^\top X + X^\top \dot X = 0 \implies \ddot X^\top X + X^\top \ddot X + 2\dot X^\top \dot X = 0, \tag{62}$$

we find an explicit form for the Lagrange multiplier

$$S = -2\big(\dot X^\top \dot X + \tfrac{2\alpha+1}{\alpha+1}(X^\top \dot X)^2\big). \tag{63}$$

We can now write down an α-family of Euler-Lagrange equations to the variational problem (56), it is the set

$$\begin{aligned} \ddot X + \tfrac{2\alpha+1}{\alpha+1}(\dot X \dot X^\top X) + X\big(\tfrac{2\alpha+1}{\alpha+1}(X^\top \dot X)^2 + \dot X^\top \dot X\big) &= 0,\\ X^\top X &= I_k. \end{aligned} \tag{64}$$

Equations (64) discribe an explicit, time-variant, second order, highly nonlinear matrix ODE on $\mathbb{R}^{n\times k}$ with solutions lying on the Stiefel manifold. We want to look for solutions of the associated α-family of initial value problems (IVP).

$$\begin{aligned} \ddot X &= -\tfrac{2\alpha+1}{\alpha+1}(\dot X \dot X^\top X) - X\Big(\tfrac{2\alpha+1}{\alpha+1}(X^\top \dot X)^2 + \dot X^\top \dot X\Big), & \dot X(0) &=: \dot X_0,\\ X^\top X &= I_k, & X(0) &=: X_0. \end{aligned} \tag{65}$$

Theorem 5.1. *For all $\alpha \neq -1$ the IVP (65) is equivalent to a second order* linear *time-invariant IVP.*

Proof. We streamline notation a bit and introduce for convenience

$$v := \tfrac{2\alpha+1}{\alpha+1}. \tag{66}$$

We rewrite the right hand side of the first equation in (65) and identify certain invariants. Indeed,

$$\begin{aligned} &-v\Big(\dot X \cdot \underbrace{\dot X^\top X}_{\in \mathfrak{so}_k \text{ by } (62)}\Big) - X\Big(v(X^\top \dot X)^2 + \dot X^\top \dot X\Big)\\ &= -v(\dot X \dot X^\top X) - X\Big(v(X^\top \dot X)^2 + \dot X^\top \dot X\Big) + \underbrace{(\dot X X^\top \dot X - \dot X X^\top \dot X)}_{=0}\\ &= -\dot X\,\underbrace{(1-v)(X^\top \dot X)}_{=\Psi^{K_X^\perp}(t) \text{ by } (42)} \underbrace{-\big(vXX^\top \dot X X^\top + X\dot X^\top - \dot X X^\top\big)}_{=\Omega^{K_X^\perp}(t) \text{ by } (42)}\dot X\\ &= \Omega^{K_X^\perp}(t)\cdot \dot X - \dot X \cdot \Psi^{K_X^\perp}(t), \end{aligned} \tag{67}$$

setting $\xi = \dot X(t)$ for $X(t)$ being the solution of (64).

Claim 5.2. *For a fixed solution $X(t)$ to (65), both linear operators $\Psi^{K_X^\perp}(t)\colon \mathbb{R} \to \mathfrak{so}_k$ and $\Omega^{K_X^\perp}(t)\colon \mathbb{R} \to \mathfrak{so}_n$ are constant.*

Proof. (Of Claim 5.2) Firstly,

$$\dot\Psi^{K_X^\perp}(t) = (1-v)\tfrac{\mathrm{d}}{\mathrm{d}t}(X^\top \dot X) = (1-v)(\dot X^\top \dot X + X^\top \ddot X) = 0, \tag{68}$$

by combining (61) and (62). Secondly,

$$\begin{aligned} \dot\Omega^{K_X^\perp}(t) &= -v\big(\dot X \underbrace{X^\top \dot X}_{\in \mathfrak{so}_k} X^\top + X \underbrace{\tfrac{\mathrm{d}}{\mathrm{d}t}(X^\top \dot X)}_{=\dot\Psi^{K_X^\perp}(t)=0} X^\top + XX^\top \dot X \dot X^\top\big)\\ &\quad - (\dot X \dot X^\top + X\ddot X^\top - \ddot X X^\top - \dot X \dot X^\top)\\ &= -v\big[XX^\top, \dot X \dot X^\top\big] - X\ddot X^\top + \ddot X X^\top = 0. \end{aligned} \tag{69}$$

The last equality in (69) follows by inserting twice the Euler-Lagrange equation (64). ☐

Consequently, $\Psi^{K_X^\perp}(t) = \Psi^{K_X^\perp}(0) =: \Psi_0^{K_X^\perp}$ and $\Omega^{K_X^\perp}(t) = \Omega^{K_X^\perp}(0) =: \Omega_0^{K_X^\perp}$. Now we combine both constant operators to one single constant linear operator

$$\Delta_0 \colon \mathbb{R}^{n\times k} \to \mathbb{R}^{n\times k}, \qquad X \mapsto \Delta_0 \circ X := \Omega_0^{K_X^\perp} X - X\Psi_0^{K_X^\perp}. \tag{70}$$

The Euler-Lagrange equation can now be rewritten as

$$\ddot{X} = \Delta_0 \circ \dot{X}. \tag{71}$$

Furthermore,

$$\begin{aligned} \Delta_0 \circ X &= \Omega_0^{K_X^\perp} X - X\Psi_0^{K_X^\perp} \\ &= -\big(vXX^\top \dot{X} X^\top + X\dot{X}^\top - \dot{X}X^\top\big)\cdot X - X\cdot(1-v)X^\top \dot{X} \\ &= \dot{X}. \end{aligned} \tag{72}$$

Consequently, the Euler-Lagrange equations for our variational calculus problem get the simple form

$$\ddot{X} = \Delta_0 \circ \Delta_0 \circ X, \quad X^\top X = I_k, \tag{73}$$

a second order, explicit, time-independent, linear ODE on the Stiefel manifold. ☐

Remark 5. Notice, that for convenience, throughout our paper the notation $\mathrm{e}^X \equiv \exp(X)$ means exclusively the matrix exponential for the square matrix X.

Corollary 3. *The unique solution of the initial value problem*

$$\ddot{X} = \Delta_0 \circ \Delta_0 \circ X, \quad X(0) = X_0 \in \mathrm{St}_{n,k}, \quad \dot{X}(0) = \dot{X}_0 \in T_{X_0}\mathrm{St}_{n,k} \tag{74}$$

for finding extremal curves with respect to the α-metric defined by (43) on the Stiefel manifold $\mathrm{St}_{n,k}$ is

$$\begin{aligned} X(t) &= \mathrm{e}^{t\Omega_0^{K_X^\perp}}\, X_0\, \mathrm{e}^{-t\Psi_0^{K_X^\perp}} \\ &= \mathrm{e}^{t(-vX_0X_0^\top \dot{X}_0 X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0 X_0^\top)} \cdot X_0 \cdot \mathrm{e}^{-t(1-v)X_0^\top \dot{X}_0} \\ &= \mathrm{e}^{t(-vX_0X_0^\top \dot{X}_0 X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0 X_0^\top)} \cdot \mathrm{e}^{-t(1-v)X_0X_0^\top \dot{X}_0 X_0^\top} \cdot X_0 \\ &= \exp\Big(t\big(-\tfrac{2\alpha+1}{\alpha+1}X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top\big)\Big)\cdot X_0 \cdot \exp\big(t\tfrac{\alpha}{\alpha+1}X_0^\top \dot{X}_0\big) \\ &= \exp\Big(t\big(-\tfrac{2\alpha+1}{\alpha+1}X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top\big)\Big)\exp\big(t\tfrac{\alpha}{\alpha+1}X_0X_0^\top \dot{X}_0X_0^\top\big)X_0. \end{aligned} \tag{75}$$

Proof. The proof is by direct verification, alternatively see subsection 11 of chapter 1 in [8] for handling this type of linear time-invariant matrix-valued ODE. ☐

We are now in the position to study the special cases for choosing certain values for α from Section 4 again, this time in terms of closed formulas for extremal curves.

5.1. Case 1: The limit $\alpha \to 0$ and the normal or submersion metric induced by σ.

The corresponding Euler-Lagrange equation appeared already in [9], but it was presented without proof. However, the closed form solution of the IVP

$$\begin{aligned} \ddot{X} &= -\dot{X}\dot{X}^\top X - X\Big((X^\top \dot{X})^2 + \dot{X}^\top \dot{X}\Big), \quad & \dot{X}(0) &=: \dot{X}_0, \\ X^\top X &= I_k, & X(0) &=: X_0. \end{aligned} \tag{76}$$

in terms of *arbitrary* initial values seems to be new,

$$X(t) = \mathrm{e}^{t(-X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot X_0. \tag{77}$$

Consider again the special isotropy point

$$X_0 = \begin{bmatrix} I_k \\ 0 \end{bmatrix},\ (\Omega, \Psi) = \left(\begin{bmatrix} A & -B^\top \\ B & C \end{bmatrix}, \Psi \right) \in \mathfrak{so}_n \times \mathfrak{so}_k,\ \dot{X}_0 = \Omega X_0 - X_0\Psi = \begin{bmatrix} A \\ B \end{bmatrix}. \tag{78}$$

With this choice of initial values, the solution of IVP (76), i.e. (77), becomes

$$X(t) = \mathrm{e}^{t\begin{bmatrix} A & -B^\top \\ B & 0 \end{bmatrix}} \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \tag{79}$$

in accordance with [11] and [18].

5.2. **Case 2:** $\alpha = -\frac{1}{2}$ **and the Euclidean metric induced by** $\mathrm{St}_{n,k} \subset \mathbb{R}^{n\times k}$**.** Here the corresponding (surprisingly simple) Euler-Lagrange equation is well-known. The closed form solution of the corresponding IVP appeared for the first time in [15], [27], see also [13] for a low dimensional example.

$$\begin{aligned} \ddot{X} &= -X\dot{X}^\top \dot{X}, & \dot{X}(0) &=: \dot{X}_0, \\ X^\top X &= I_k, & X(0) &=: X_0. \end{aligned} \tag{80}$$

The solution for (80) in terms of *arbitrary* initial values is as

$$\begin{aligned} X(t) &= \mathrm{e}^{t(-X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot X_0 \cdot \mathrm{e}^{-tX_0^\top \dot{X}_0} \\ &= \mathrm{e}^{t(-X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot \mathrm{e}^{-tX_0X_0^\top \dot{X}_0X_0^\top} \cdot X_0. \end{aligned} \tag{81}$$

Choose again (78) as initial values. Now (81) becomes

$$X(t) = \mathrm{e}^{t\begin{bmatrix} 2A & -B^\top \\ B & 0 \end{bmatrix}} \begin{bmatrix} I_k \\ 0 \end{bmatrix} \mathrm{e}^{-tA} = \mathrm{e}^{t\begin{bmatrix} 2A & -B^\top \\ B & 0 \end{bmatrix}} \mathrm{e}^{-t\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}} \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \tag{82}$$

again in accordance with [11] and [18].

One might compare the two alternatives in (81) with the cumbersome formula appearing in subsection 2.2.2 of [9].

Remark 6. The IVP (80) appears also in a more general context in [7], where the authors study variational problems on Stiefel manifolds with a very different family of left invariant Riemannian metrics compared to ours'. Their family of Riemannian metrics depends on an $(n \times n)$-diagonal positive definite matrix Λ. However, for $\Lambda = I_n$, (80) is equivalent to (21) and (22) in [7]. But note, our n corresponds to an N in [7].

5.3. **Case 3: The limit** $\alpha \to \infty$**.** The associated Euler-Lagrange equation was to the best knowledge of the authors never published before:

$$\begin{aligned} \ddot{X} &= -2(\dot{X}\dot{X}^\top X) - X\Big(2(X^\top \dot{X})^2 + \dot{X}^\top \dot{X}\Big), & \dot{X}(0) &=: \dot{X}_0, \\ X^\top X &= I_k, & X(0) &=: X_0. \end{aligned} \tag{83}$$

The solution for IVP (83) in terms of *arbitrary* initial values is as

$$\begin{aligned} X(t) &= \mathrm{e}^{t(-2X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot X_0 \cdot \mathrm{e}^{tX_0^\top \dot{X}_0} \\ &= \mathrm{e}^{t(-2X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot \mathrm{e}^{tX_0X_0^\top \dot{X}_0X_0^\top} \cdot X_0. \end{aligned} \tag{84}$$

Choose once again (78) as initial values. Then (84) becomes

$$X(t) = \mathrm{e}^{t\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}} \begin{bmatrix} I_k \\ 0 \end{bmatrix} \mathrm{e}^{tA} = \mathrm{e}^{t\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}} \mathrm{e}^{t\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}} \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \tag{85}$$

again in accordance with [18] and [19].

Remark 7. The curves (84) and (85) showed up recently in a different context, cf. [18], [19] and [20]. In [20] the authors used these curves to solve a boundary value problem, whereas in [18, 19] they were analyzed from a purely geometric point of view.

5.4. **Case 4:** $\alpha = 1$. The associated Euler-Lagrange equation is as

$$\begin{aligned} \ddot{X} &= -\tfrac{3}{2}(\dot{X}\dot{X}^\top X) - X\Big(\tfrac{3}{2}(X^\top \dot{X})^2 + \dot{X}^\top \dot{X}\Big), & \dot{X}(0) &=: \dot{X}_0, \\ X^\top X &= I_k, & X(0) &=: X_0. \end{aligned} \tag{86}$$

The solution for IVP (86) in terms of *arbitrary* initial values is as

$$\begin{aligned} X(t) &= \mathrm{e}^{t(-\frac{3}{2}X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot X_0 \cdot \mathrm{e}^{\frac{t}{2}X_0^\top \dot{X}_0} \\ &= \mathrm{e}^{t(-\frac{3}{2}X_0X_0^\top \dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top)} \cdot \mathrm{e}^{\frac{t}{2}X_0X_0^\top \dot{X}_0X_0^\top} \cdot X_0. \end{aligned} \tag{87}$$

Choose (78) as initial values. Then (87) becomes

$$\begin{aligned} X(t) &= \exp\left(t\begin{bmatrix} A/2 & -B^\top \\ B & 0 \end{bmatrix}\right)\begin{bmatrix} I_k \\ 0 \end{bmatrix}\mathrm{e}^{tA/2} \\ &= \exp\left(t\begin{bmatrix} A/2 & -B^\top \\ B & 0 \end{bmatrix}\right)\exp\left(t\begin{bmatrix} A/2 & 0 \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} I_k \\ 0 \end{bmatrix}, \end{aligned} \tag{88}$$

again in accordance with [11] and [18]. However, (86) seems to be new.

5.5. **Case 5:** $\alpha < -1$. In this case we always get a true pseudo-Riemannian metric on $\mathrm{St}_{n,k}$, i.e., one which is not Riemannian. We are currently not aware of any application.

Remark 8. Again we comment on the two extreme cases $\mathrm{St}_{n,1} = S^{n-1}$ and $\mathrm{St}_{n,n} = \mathrm{O}_n$.

For the sphere S^{n-1} the geodesic curves (75) simplify to curves describing great circles, independent of α. Indeed,

$$\begin{aligned} X(t) &= \exp\Big(t\big(-\tfrac{2\alpha+1}{\alpha+1}X_0\underbrace{X_0^\top\dot{X}_0}_{=0}X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top\big)\Big)\exp\big(t\tfrac{\alpha}{\alpha+1}X_0\underbrace{X_0^\top\dot{X}_0}_{=0}X_0^\top\big)\cdot X_0 \\ &= \exp\big(t(\dot{X}_0X_0^\top - X_0\dot{X}_0^\top)\big)\cdot X_0 = X_0\cos\big(t\|\dot{X}_0\|\big) + \dot{X}_0\frac{\sin(t\|\dot{X}_0\|)}{\|\dot{X}_0\|}. \end{aligned}$$

Note that the last equality even makes sense for $\|\dot{X}_0\| \to 0$ as the limit of the quotient in the last summand nevertheless exists. Moreover, to see that these circles are actually great circles, one might insert $t = \frac{\pi}{\|\dot{X}_0\|}$ to realize that X_0 and $-X_0$ as well, i.e. antipodes, lie on the curve $X(t)$.

For the orthogonal group O_n the geodesic curves (75) reduce to curves specified by one-parameter subgroups acting on X_0, independent of α. In fact, exploiting $\dot{X}_0X_0^\top = -X_0^\top\dot{X}_0 \in \mathfrak{so}_n$ we get

$$\begin{aligned} X(t) &= \exp\Big(t\big(-\tfrac{2\alpha+1}{\alpha+1}\underbrace{X_0X_0^\top}_{=I_n}\dot{X}_0X_0^\top - X_0\dot{X}_0^\top + \dot{X}_0X_0^\top\big)\Big)\exp\big(t\tfrac{\alpha}{\alpha+1}\underbrace{X_0X_0^\top}_{=I_n}\dot{X}_0X_0^\top\big)\cdot X_0 \\ &= \exp\big(t(\dot{X}_0X_0^\top)\big)\cdot X_0. \end{aligned}$$

These results are certainly fully in accordance with well-known facts.

6. **Curves of constant geodesic curvature with respect to α-metrics.** The relationship to a sub-Riemannian problem will be explained in this section. First we present and adapt a result from [18] that was obtained earlier, cf. [17], [21]. Consider $\mathrm{O}_n \times \mathrm{O}_k$ endowed with the α-metric defined in (21). Let $\mathfrak{p} \oplus \mathfrak{k}$ be the corresponding Cartan decomposition of $\mathfrak{so}_n \times \mathfrak{so}_k$ as it is written in (34) and (35). By making use of the translation we define an α-family of left invariant horizontal distributions $\mathcal{H}$ on $\mathrm{O}_n \times \mathrm{O}_k$ that are orthogonal to the vertical distribution $\mathcal{V}$ with respect to the bi-invariant α-metric (21). The *vertical distribution* $\mathcal{V}$ is equal to $\mathfrak{k} = K_X$ at the identity whereas the *horizontal distribution* $\mathcal{H}$ is equal to $\mathfrak{p} = K_X^{\perp}$ at the identity. Thus $T(\mathrm{O}_n \times \mathrm{O}_k) = \mathcal{H} \oplus \mathcal{V}$, where $\mathcal{H}$ is the bracket generating distribution according to the fourth condition in (35). We say that a smooth curve $c\colon [0,1] \to \mathrm{O}_n \times \mathrm{O}_k$ is *horizontal* if $\dot{c}(t) \in \mathcal{H}_{c(t)}$. From now on we assume that $\langle \cdot\,, \cdot \rangle^{(\alpha)}$ is positive definite when it is restricted to $\mathfrak{p}$. The *sub-Riemannian distance* function $d_{\mathrm{subR}}(g_1, g_2)$ on the group $\mathrm{O}_n \times \mathrm{O}_k$ is defined as the infimum of the lengths among all horizontal curves connecting the points g_1 and g_2 on $\mathrm{O}_n \times \mathrm{O}_k$. The reader can find more about sub-Riemannian geometry in [3, 21].

Definition 6.1. A sub-Riemannian geodesic $g\colon [0,1] \to \mathrm{O}_n \times \mathrm{O}_k$ is a smooth horizontal curve that locally realizes the sub-Riemannian distance function d_{subR}.

Proposition 4. *A sub-Riemannian geodesic on $\mathrm{O}_n \times \mathrm{O}_k$ tangent to $\mathcal{H}$ is given by*

$$g(t) = g_0 \operatorname{Exp}\big(t(P_{\mathfrak{p}} + P_{\mathfrak{k}})\big) \operatorname{Exp}(-tP_{\mathfrak{k}}), \quad g_0 \in \mathrm{O}_n \times \mathrm{O}_k, \quad P_{\mathfrak{p}} \in \mathfrak{p},\ P_{\mathfrak{k}} \in \mathfrak{k}, \tag{89}$$

with initial velocity $P_{\mathfrak{p}} + P_{\mathfrak{k}}$ and initial point $g_0 \in \mathrm{O}_n \times \mathrm{O}_k$, cf. [17, 21]. *The Riemannian geodesics on $\mathrm{St}_{n,k} = (\mathrm{O}_n \times \mathrm{O}_k)/(\mathrm{O}_k \times \mathrm{O}_{n-k})$ are exactly projections of sub-Riemannian geodesics* (89) *for which $P_{\mathfrak{k}} = 0$, cf.* [17].

Remark 9. Here, Exp denotes the exponential map for the product group $\mathrm{O}_n \times \mathrm{O}_k$, at this stage not to be confused with the ordinary matrix exponential $\mathrm{e}^X \equiv \exp(X)$.

Proof. Translated into our terminology, the sub-Riemannian geodesics (89) on the product group $\mathrm{O}_n \times \mathrm{O}_k$ (only at the identity for convenience) are of the form

$$\begin{aligned}
&\mathbb{R} \to \mathrm{O}_n \times \mathrm{O}_k,\\
&\quad t \mapsto \Big(\exp\big(t(\Omega_0^{K_X^{\perp}} + \Omega_0^{K_X})\big), \exp\big(-t(\Psi_0^{K_X^{\perp}} + \Psi_0^{K_X})\big)\Big) \circ \Big(\exp\big(-t\Omega_0^{K_X}\big), \exp\big(t\Psi_0^{K_X}\big)\Big)\\
&\qquad = \Big(\exp\big(t(\Omega_0^{K_X^{\perp}} + \Omega_0^{K_X})\big)\exp\big(-t\Omega_0^{K_X}\big), \exp\big(-t(\Psi_0^{K_X^{\perp}} + \Psi_0^{K_X})\big)\exp\big(t\Psi_0^{K_X}\big)\Big),
\end{aligned}$$

which, for $P_{\mathfrak{k}} = 0$, i.e., $\big(\Omega_0^{K_X}, \Psi_0^{K_X}\big) = (0,0)$, project to Riemannian geodesics on $\mathrm{St}_{n,k}$, which start at X and emanate in direction $\Omega_0^{K_X^{\perp}} X - X\Psi_0^{K_X^{\perp}}$, as

$$\mathbb{R} \to \mathrm{St}_{n,k}, \quad t \mapsto \mathrm{e}^{t\Omega_0^{K_X^{\perp}}} X\, \mathrm{e}^{-t\Psi_0^{K_X^{\perp}}}. \tag{90}$$

This is in accordance with Corollary 3, namely the first line of (75), as required. □

We rephrase here a result from proposition 5 in [18], adapted to our notations but there under more general assumptions.

Proposition 5. *The projection of the sub-Riemannian geodesic* (89) *onto the Stiefel manifold $\mathrm{St}_{n,k} = (\mathrm{O}_n \times \mathrm{O}_k)/(\mathrm{O}_k \times \mathrm{O}_{n-k})$ is a curve of constant geodesic curvature relative to the Riemannian α-metric defined in* (41). *The geodesic curvature of the projection is equal to $\|[P_{\mathfrak{p}}, P_{\mathfrak{k}}]\|$, where the norm is understood in terms of the α-metric.* □

As a consequence of Proposition 5 we immediately obtain

Corollary 4. *For any $X \in \mathrm{St}_{n,k}$ and any $(\Omega, \Psi) \in \mathfrak{so}_n \times \mathfrak{so}_k$ all smooth curves of type*

$$\mathbb{R} \to \mathrm{St}_{n,k}, \quad t \mapsto \mathrm{e}^{t\Omega} X \,\mathrm{e}^{-t\Psi} \tag{91}$$

are curves of constant geodesic curvature relative to the Riemannian α-metric defined in (41). □

6.1. **Alternative approach to obtain formula (89).** Equation (89) for a sub-Riemannian geodesic tangent to the left invariant distribution obtained from $\mathfrak{p}$ can be also obtained by variational methods, incorporating Lagrange multipliers for both, holonomic and nonholonomic constraints, [14]. To give an outline of that approach, we first embed $\mathrm{O}_n \times \mathrm{O}_k$ into the space $\mathbb{R}^{(n+k)\times(n+k)}$ of square (n, k)-block diagonal matrices with square $(n \times n)$ and $(k \times k)$-blocks on the main diagonal. By making use of left translations we identify the tangent space $T_X(\mathbb{R}^{n\times n} \oplus \mathbb{R}^{k\times k}) \cong \mathbb{R}^{n\times n} \oplus \mathbb{R}^{k\times k}$ with the Lie algebra $\mathfrak{g} := \mathfrak{gl}_n \oplus \mathfrak{gl}_k$ of (n, k)-block diagonal matrices. We decompose

$$\mathfrak{g} = \mathrm{Sym}_n \oplus \mathrm{Sym}_k \oplus \mathfrak{so}_n \oplus \mathfrak{so}_k = \mathrm{Sym}_n \oplus \mathrm{Sym}_k \oplus \mathfrak{p} \oplus \mathfrak{k}, \tag{92}$$

where $\mathfrak{p} \oplus \mathfrak{k}$ is the Cartan decomposition of $\mathfrak{so}_n \times \mathfrak{so}_k$. The direct sums (92) are orthogonal with respect to the inner product

$$\langle A, B\rangle^{(\alpha)} = \mathrm{tr}(A_1^\top B_1) + \tfrac{1}{\alpha}\,\mathrm{tr}(A_2^\top B_2). \tag{93}$$

Here $A = (A_1, A_2), B = (B_1, B_2) \in \mathfrak{g}$, meaning that $A_1, B_1 \in \mathfrak{gl}_n$ and $A_2, B_2 \in \mathfrak{gl}_k$. The product $\langle \cdot\,, \cdot\rangle^{(\alpha)}$ is ad-invariant, indefinite, and non-degenerate for $\alpha \neq -1$. We define the bi-invariant indefinite metric on $T(\mathbb{R}^{n\times n} \oplus \mathbb{R}^{k\times k})$ by taking left translations of the product $\langle \cdot\,, \cdot\rangle^{(\alpha)}$. Define the functional

$$X(t) \mapsto \tfrac{1}{2}\int_0^1 \Big(\big\langle \dot X, \dot X\big\rangle^{(\alpha)} + \big\langle S, X^T X - I_{n+k}\big\rangle^{(\alpha)} + \big\langle P_{\mathfrak{k}}, X^T \dot X\big\rangle^{(\alpha)}\Big)\mathrm{d}t. \tag{94}$$

Here $\frac{1}{2}\int_0^1 \langle \dot X, \dot X\rangle^{(\alpha)}\mathrm{d}t$ is the energy functional that is equal to the square of the α-norm of a curve $X(t) \in \mathbb{R}^{n\times n} \oplus \mathbb{R}^{k\times k}$. The summand $\frac{1}{2}\int_0^1 \langle S, X^T X - I_{n+k}\rangle^{(\alpha)}\mathrm{d}t$ is a requirement that $X(t) \in \mathrm{O}_n \times \mathrm{O}_k$ and S is an (n, k)-block diagonal symmetric matrix in $\mathbb{R}^{n\times n}\oplus\mathbb{R}^{k\times k}$, that serves as a Lagrange multiplier to enforce the holonomic constraint. It implies that $X^T \dot X \in \mathfrak{so}_n \times \mathfrak{so}_k$. The last term $\frac{1}{2}\int_0^1 \langle P_{\mathfrak{k}}, X^T \dot X\rangle^{(\alpha)}\mathrm{d}t$ incorporates the Lagrange multiplier $P_{\mathfrak{k}} \in \mathfrak{k}$ to enforce the nonholonomic constraint. It will ensure that $X^T \dot X \in \mathfrak{p}$.

By making calculations similar to those in section 4 of [14] and looking for the critical point of (94), we obtain the geodesic equation

$$\ddot X(t) = X(t)\,\mathrm{Exp}(tP_{\mathfrak{k}})\Big(P_{\mathfrak{p}}^2 + [P_{\mathfrak{k}}, P_{\mathfrak{p}}]\Big)\,\mathrm{Exp}(-tP_{\mathfrak{k}}), \tag{95}$$

on $\mathrm{O}_n \times \mathrm{O}_k$ for a sub-Riemannian geodesic. Here $P_{\mathfrak{k}} \in \mathfrak{k}$, and $P_{\mathfrak{p}} \in \mathfrak{p}$, can be considered as data for specifying an initial velocity vector. It is a result of the corresponding calculations showing that the Lagrange multiplier $P_{\mathfrak{k}}$ is actually constant. By the substitution

$$Y(t) := X(t)\,\mathrm{Exp}(tP_{\mathfrak{k}}) \tag{96}$$

the equation (95) is reduced to the linear second order differential equation with constant coefficients

$$\ddot Y(t) - 2\dot Y(t)P_{\mathfrak{k}} + Y(t)\Big(P_{\mathfrak{k}}^2 - P_{\mathfrak{p}}^2 - [P_{\mathfrak{k}}, P_{\mathfrak{p}}]\Big) = 0, \tag{97}$$

having the unique solution

$$Y(t) = Y(0)\,\mathrm{Exp}\big(t(P_{\mathfrak{p}} + P_{\mathfrak{k}})\big), \tag{98}$$

with $Y(0) = X(0)$. Combining (96) and (98) we obtain (89). Note that, eventually, in this subsection as $X \in \mathrm{O}_n \times \mathrm{O}_k$ was realized as a matrix $X = \left[\begin{smallmatrix} X_1 & 0 \\ 0 & X_2 \end{smallmatrix}\right]$ with $X_1 \in \mathrm{O}_n$ and $X_2 \in \mathrm{O}_k$, we got $\mathrm{Exp}(X) = \exp\big(\left[\begin{smallmatrix} X_1 & 0 \\ 0 & X_2 \end{smallmatrix}\right]\big) = \left[\begin{smallmatrix} \mathrm{e}^{X_1} & 0 \\ 0 & \mathrm{e}^{X_2} \end{smallmatrix}\right]$.

Acknowledgments. The first author has been supported by German BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA. The second author was partially supported by the project Pure Mathematics in Norway, funded by the Trond Mohn Foundation. The third author acknowledges Fundação para a Ciênca e a Tecnologia (FCT) and COMPETE 2020 program for the financial support to the project UIDB/00048/2020.

REFERENCES

[1] P.-A. Absil, R. Mahony and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.

[2] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens and M. Shub, Newton's method on Riemannian manifolds and a geometric model for the human spine, *IMA J. Numer. Anal.*, **22** (2002), 359–390.

[3] A. Agrachev, D. Barilari and U. Boscain, *A Comprehensive Introduction to Sub-Riemannian Geometry*, Cambridge University Press, Cambridge, 2020.

[4] A. Bloch, L. Colombo, R. Gupta and D. Martín de Diego, A geometric approach to the optimal control of nonholonomic mechanical systems, in *Analysis and Geometry in Control Theory and Its Applications*, Springer, Cham, 2015, 35–64.

[5] A. M. Bloch, *Nonholonomic Mechanics and Control*, 2nd edition, Springer, New York, 2015.

[6] A. M. Bloch, J. E. Marsden and D. V. Zenkov, Nonholonomic dynamics, *Notices Amer. Math. Soc.*, **52** (2005), 324–333.

[7] A. M. Bloch, P. E. Crouch and A. K. Sanyal, A variational problem on Stiefel manifolds, *Nonlinearity*, **19** (2006), 2247–2276.

[8] R. W. Brockett, *Finite dimensional linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.

[9] A. Edelman, T. A. Arias and S. T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.*, **20** (1998), 303–353.

[10] J. Faraut, *Analysis on Lie Groups. An Introduction*, Cambridge University Press, Cambridge, 2008.

[11] Y. N. Fedorov and B. Jovanović, Geodesic flows and Neumann systems on Stiefel varieties: Geometry and integrability, *Math. Z.*, **270** (2012), 659–698.

[12] I. Gohberg, P. Lancaster and L. Rodman, *Indefinite Linear Algebra and Applications*, Birkhäuser Verlag, Basel, 2005.

[13] K. Hüper, M. Kleinsteuber and F. Silva Leite, Rolling Stiefel manifolds, *Internat. J. Systems Sci.*, **39** (2008), 881–887.

[14] K. Hüper, I. Markina and F. Silva Leite, An extrinsic approach to sub-Riemannian geodesics on the orthogonal group, in *CONTROLO 2020, Proceedings of the 14th APCA International Conference on Automatic Control and Soft Computing*, vol. 695, LNEE, Bragança, Portugal, Springer 2020, 274–283.

[15] K. Hüper and F Ullrich, Real Stiefel manifolds: An extrinsic point of view, In *13th APCA International Conference on Automatic Control and Soft Computing (Controlo 2018)*, Ponta Delgada, Azores, Portugal, 2018, 13–18.

[16] I. I. Hussein and A. M. Bloch, Optimal control of underactuated nonholonomic mechanical systems, *IEEE Trans. Automat. Control*, **53** (2008), 668–682.

[17] V. Jurdjevic, *Optimal Control and Geometry: Integrable Systems*, Cambridge University Press, Cambridge, 2016.

[18] V. Jurdjevic, I. Markina and F. Silva Leite, Extremal curves on Stiefel and Grassmann manifolds, *Journal of Geometric Analysis*, **30** (2020), 3948–3978.

[19] V. Jurdjevic, F. Silva Leite and K. Krakowski, The geometry of quasi-geodesics on Stiefel manifolds, in *13th APCA International Conference on Automatic Control and Soft Computing (Controlo 2018)*, Ponta Delgada, Azores, Portugal, 2018, 213–218.

[20] K. A. Krakowski, L. Machado, F. Silva Leite and J. Batista, A modified Casteljau algorithm to solve interpolation problems on Stiefel manifolds, *J. Comput. Appl. Math.*, **311** (2017), 84–99.

[21] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, American Mathematical Society, Providence, RI, 2002.

[22] Y. Nishimori and S. Akaho, Learning algorithms utilizing quasi-geodesic flows on the Stiefel manifold, *Neurocomputing*, **67** (2005), 106–135.

[23] A. L. Onishchik, *Topology of Transitive Transformation Groups*, Johann Ambrosius Barth Verlag GmbH, Leipzig, 1994.

[24] J. Roe, *Lectures on Coarse Geometry*, American Mathematical Society, Providence, RI, 2003.

[25] J. Roe, What is . . . a coarse space?, *Notices Amer. Math. Soc.*, **53** (2006), 668–669.

[26] E. Stiefel, Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten, *Comment. Math. Helv.*, **8** (1935), 305–353.

[27] F. Ullrich, Rolling maps for real Stiefel manifolds, Master's thesis, Institute of Mathematics, Julius-Maximilians-Universität Würzburg, Germany, 2017.

[28] E. Zeidler, *Nonlinear Functional Analysis and Its Applications*, vol. 3, Springer-Verlag, New York, 1985.

Received for publishing March 2020.

E-mail address: hueper@mathematik.uni-wuerzburg.de
E-mail address: irina.markina@uib.no
E-mail address: fleite@mat.uc.pt