Journal of Geometric Mechanics

https://www.aimspress.com/journal/jgm

Journal of Geometric Mechanics, 15(1): 147–187
DOI: 10.3934/jgm.2023008
Received: 16 November 2022
Revised: 11 January 2023
Accepted: 16 January 2023
Published: 01 February 2023

Research article

A multi-parameter family of metrics on stiefel manifolds and applications

Markus Schlarb*

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

* **Correspondence:** Email: markus.schlarb@mathematik.uni-wuerzburg.de.

Abstract: The real (compact) Stiefel manifold realized as set of orthonormal frames is considered as a pseudo-Riemannian submanifold of an open subset of a vector space equipped with a multi-parameter family of pseudo-Riemannian metrics. This family contains several well-known metrics from the literature. Explicit matrix-type formulas for various differential geometric quantities are derived. The orthogonal projections onto tangent spaces are determined. Moreover, by computing the metric spray, the geodesic equation as an explicit second order matrix valued ODE is obtained. In addition, for a multi-parameter subfamily, explicit matrix-type formulas for pseudo-Riemannian gradients and pseudo-Riemannian Hessians are derived. Furthermore, an explicit expression for the second fundamental form and an explicit formula for the Levi-Civita covariant derivative are obtained. Detailed proofs are included.

Keywords: constrained Lagrangian systems; pseudo-Riemannian gradients; pseudo-Riemannian Hessians; pseudo-Riemannian submanifolds; Riemannian optimization; second fundamental form; sprays; Stiefel manifolds
Mathematics Subject Classification: 53C30, 53B30, 70H03

1. Introduction

Numerical computations on the real (compact) Stiefel manifold viewed as the embedded submanifold $\mathrm{St}_{n,k} = \{X \in \mathbb{R}^{n\times k} \mid X^\top X = I_k\}$ of $\mathbb{R}^{n\times k}$ arise in many branches of applied mathematics like numerical linear algebra and, moreover, in the engineering context, as well. Beside interpolation problems [1], we mention the following examples which are closely linked to optimization. For instance, the symmetric eigenvalue problem can be formulated as an optimization problem on the Stiefel manifold [2]. Moreover, one encounters optimization problems on $\mathrm{St}_{n,k}$ in connection with machine learning [3], multivariate data analysis [4] and computer vision [5, 6].

These problems can be tackled by Riemannian optimization methods, see e.g. [2, 7, 8, 9]. An essential part of their design is the choice of an appropriated Riemannian metric [7, Chap. 1]. The Euclidean metric, see e.g. [2], and the so-called canonical metric, see e.g. [10], are well-known, common choices for the Stiefel manifold. For these two metrics, explicit formulas for Riemannian gradients and Riemannian Hessians of smooth functions are known. Such formulas are desirable for the application of several Riemannian optimization methods. However, there is no reason to restrict to one of these two metrics. In principle, the performance of a Riemannian optimization method could be improved by choosing an alternative metric adapted to the particular function under consideration. For example, the dependence of the speed of convergence of a Riemannian optimization method on the Riemannian metric is investigated in [11] on "Riemannian preconditioning". Moreover, a family of metrics on the generalized Stiefel manifold is introduced in [11] which differs from the family of metrics on $\mathrm{St}_{n,k}$ discussed here.

In this paper, we investigate a $2k$-parameter family of pseudo-Riemannian metrics on $\mathrm{St}_{n,k}$ from an extrinsic point of view. This family does not coincide with the family of metrics considered in [12]. Nevertheless, it contains the Euclidean metric and the so-called canonical metric. In addition, the whole one-parameter family which has been recently introduced in [13] is included. An emphasize is put on deriving explicit formulas for gradients and Hessians suitable for applying them in connection with Riemannian optimization methods. In particular, specific results of the conference paper [14] are reproduced as special cases.

Next we give an overview of this text which is kept as self-contained as possible. We start with endowing $\mathbb{R}^{n\times k}$ with a family of covariant 2-tensors depending on $2k$ parameters, which are invariant under the $O(n)$-left action on $\mathbb{R}^{n\times k}$ by matrix multiplication from the left. For suitable choices of these parameters, the corresponding 2-tensor induces a pseudo-Riemannian metric on an open subset U of $\mathbb{R}^{n\times k}$ such that $\mathrm{St}_{n,k} \subseteq U$ becomes a pseudo-Riemannian submanifold of U. Hence it makes sense to consider the normal bundle of $\mathrm{St}_{n,k}$ and the orthogonal projections onto the tangent spaces of $\mathrm{St}_{n,k}$ which can be described by explicit formulas.

In order to put this extrinsic approach into context to existing works on families of metrics on the Stiefel manifold we also consider $\mathrm{St}_{n,k}$, equipped with our family, as a pseudo-Riemannian reductive homogeneous $SO(n)$-space. This point of view shows that, for the Riemannian case, the family of metrics which is discussed in this text, is partially contained in the family considered in the work [15] on Einstein metrics. Nevertheless, at least to our best knowledge, the family of metrics on $\mathrm{St}_{n,k}$ considered in this paper has never been treated before from an extrinsic point of view.

After this short detour, we come back to the extrinsic approach. We derive an explicit expression for the spray $S\colon T\mathrm{St}_{n,k} \to T(T\mathrm{St}_{n,k})$ associated with the metric. To this end, we exploit a well-known fact, see e.g. [16, Sec. 7.5] for the Riemannian case. The metric spray of a pseudo-Riemannian manifold coincides with the Lagrangian vector field on its tangent bundle associated with the kinetic energy defined by means of the pseudo-Riemannian metric. This allows for computing the metric spray on the tangent bundle TU, where $U \subseteq \mathbb{R}^{n\times k}$ is the open set of which $\mathrm{St}_{n,k}$ is a pseudo-Riemannian submanifold. Eventually, by using a result from [16, Sec. 8.4] on constrained Lagrangian systems, combined with the explicit expression for the orthogonal projections, the metric spray on $T\mathrm{St}_{n,k}$ is computed. As a by-product, the geodesic equation is obtained as an *explicit* second order matrix valued ordinary differential equation (ODE).

Next we derive expressions for pseudo-Riemannian gradients and pseudo-Riemannian Hessians of smooth functions on $\mathrm{St}_{n,k}$ involving only "ordinary" matrix operations. Using the formula for the orthogonal projection onto tangent spaces, we derive an explicit formula for pseudo-Riemannian gradients. Moreover, since we have an expression for the geodesic equation as explicit second order matrix valued ODE, we obtain an explicit formula for pseudo-Riemannian Hessians, too. The expression for the pseudo-Riemannian gradient is valid for all metrics in the $2k$-parameter family, while, for the pseudo-Riemannian Hessian, we restrict ourself to a subfamily depending on $(k + 1)$-parameters in order to obtain formulas which are not too complicated. This $(k+1)$-parameter subfamily still contains the Euclidean metric and the canonical metric as well as the one-parameter family from [13].

Finally, a formula for the second fundamental form of $\mathrm{St}_{n,k}$ considered as pseudo-Riemannian submanifold of an open $U \subseteq \mathbb{R}^{n\times k}$ is derived. We give a concrete expression for the second fundamental form with respect to the metrics in the $(k+1)$-parameter subfamily. By means of the Gauß formula, an explicit matrix-type formula for the Levi-Civita covariant derivative is obtained.

2. Terminology and notations

Throughout this text, except for Section 3.4, we view the real (compact) Stiefel manifold $\mathrm{St}_{n,k}$ as an embedded submanifold of the real $(n \times k)$-matrices $\mathbb{R}^{n\times k}$ which is given by

$$\mathrm{St}_{n,k} = \{X \in \mathbb{R}^{n\times k} \mid X^\top X = I_k\} \subseteq \mathbb{R}^{n\times k}, \quad 1 \leq k \leq n. \tag{2.1}$$

We point out that $\mathrm{St}_{n,k}$ is a proper subset of $\mathbb{R}^{n\times k}$ although the inclusion is denoted by $\mathrm{St}_{n,k} \subseteq \mathbb{R}^{n\times k}$. In the sequel, we often denote proper inclusions by "$\subseteq$". The symbol "$\subset$" is only used if we want to emphasize that an inclusion is not an equality. The tangent bundle of $\mathrm{St}_{n,k}$ is denoted by $T\mathrm{St}_{n,k}$ which is considered as a submanifold of $T\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k}$. More generally, for a manifold M, we denote by TM and T^*M its tangent and cotangent bundle, respectively. In the sequel, if not indicated other-wise, we identify $\mathbb{R}^{n\times k}$ with its dual space $(\mathbb{R}^{n\times k})^*$ via the linear isomorphism

$$\mathbb{R}^{n\times k} \to (\mathbb{R}^{n\times k})^*, \quad V \mapsto \mathrm{tr}\left(V^\top(\cdot)\right) = \left(W \mapsto \mathrm{tr}(V^\top W)\right) \tag{2.2}$$

induced by the Frobenius scalar product. The following characterization of the tangent space of $\mathrm{St}_{n,k}$ at $X \in \mathrm{St}_{n,k}$ considered as subspace of $\mathbb{R}^{n\times k}$ is used frequently

$$T_X\mathrm{St}_{n,k} = \{V \in \mathbb{R}^{n\times k} \mid X^\top V = -V^\top X\} \subseteq \mathbb{R}^{n\times k}. \tag{2.3}$$

We write

$$O(n) = \mathrm{St}_{n,n} = \{R \in \mathbb{R}^{n\times n} \mid R^\top R = RR^\top = I_n\} \tag{2.4}$$

for the orthogonal group and

$$SO(n) = \{R \in \mathbb{R}^{n\times k} \mid R^\top R = RR^\top = I_n \text{ and } \det(R) = 1\} \tag{2.5}$$

for the special orthogonal group. Their Lie algebras coincide and are denoted by

$$\mathfrak{so}(n) = \{\xi \in \mathbb{R}^{n\times n} \mid \xi^\top = -\xi\}. \tag{2.6}$$

Moreover, we write

$$\mathrm{skew}\colon \mathbb{R}^{n\times n} \to \mathfrak{so}(n) \subseteq \mathbb{R}^{n\times n}, \quad A \mapsto \tfrac{1}{2}(A - A^\top) \tag{2.7}$$

for the projection onto $\mathfrak{so}(n)$ whose kernel is given by the set of symmetric matrices $\mathbb{R}^{n\times n}_{\mathrm{sym}}$. The $O(n)$-left action on $\mathbb{R}^{n\times k}$ by matrix multiplication from the left is denoted by

$$\Psi\colon O(n)\times\mathbb{R}^{n\times k}\to\mathbb{R}^{n\times k},\quad (R,X)\mapsto RX. \tag{2.8}$$

By restricting the second argument of Ψ one obtains the $O(n)$-action

$$O(n)\times\mathrm{St}_{n,k}\to\mathrm{St}_{n,k},\quad (R,X)\mapsto RX \tag{2.9}$$

on $\mathrm{St}_{n,k}$ from the left which we denote by Ψ, as well. It is well-known that this $O(n)$-action on $\mathrm{St}_{n,k}$ is transitive. For fixed $R\in O(n)$ we denote the diffeomorphisms induced by the actions from (2.8) and (2.9)

$$\mathbb{R}^{n\times k}\ni X\mapsto RX\in\mathbb{R}^{n\times k}\quad\text{and}\quad\mathrm{St}_{n,k}\ni X\mapsto RX\in\mathrm{St}_{n,k} \tag{2.10}$$

both by Ψ_R.

If $U\subseteq\mathbb{R}^{n\times k}$ is some subset, we write

$$\iota_U\colon U\to\mathbb{R}^{n\times k} \tag{2.11}$$

for the canonical inclusion of U into $\mathbb{R}^{n\times k}$. Moreover, the canonical inclusion of $\mathrm{St}_{n,k}$ into $\mathbb{R}^{n\times k}$ is often denoted by

$$\iota\colon\mathrm{St}_{n,k}\to\mathbb{R}^{n\times k} \tag{2.12}$$

for short.

Next let $\mathrm{pr}\colon F\to M$ be a vector bundle over a manifold M with dual bundle F^*. The smooth sections of F are denoted by $\Gamma^\infty(F)$. Moreover, we denote by $F^{\otimes\ell}$, $\mathrm{S}^\ell(F)$ and $\Lambda^\ell(F)$ the ℓ-th tensor power, the ℓ-th symmetrized tensor power and the ℓ-th antisymmetrized tensor power of F, respectively. In addition, we write $\mathrm{End}(F)\cong F^*\otimes F$ for the endormorphism bundle of F. The vertical bundle of F is denoted by $\mathrm{Ver}(F)\subseteq TF$.

Let $f\colon M\to N$ be a smooth map between manifolds and let $\alpha\in\Gamma^\infty((T^*N)^{\otimes\ell})$ be a covariant tensor field on N. The pullback of α by f is denoted by $f^*\alpha$. If α is a differential form, i.e. $\alpha\in\Gamma^\infty(\Lambda^\ell(T^*M))$, the exterior derivative of α is denoted by $\mathrm{d}\,\alpha$. The tangent map of f is denoted by $Tf\colon TM\to TN$. If f is a map between (open subsets of) finite dimensional $\mathbb{R}$-vector spaces, we write $\mathrm{D}\,f(X)V$ for the derivative of f at X evaluated at V. Sometimes, the tangent map of a smooth map f between arbitrary manifolds at the point X evaluated at a tangent vector V is denoted by $\mathrm{D}\,f(X)V$, as well.

Next let $M\subseteq\mathbb{R}^{n\times k}$ be a submanifold. A vector field $V\colon M\to TM\subseteq\mathbb{R}^{n\times k}\times\mathbb{R}^{n\times k}$ is often implicitly identified with the map $M\to\mathbb{R}^{n\times k}$ defined by its second component which we denote by V, as well, i.e. the "foot point" $X\in M$ is suppressed in our notation. If $S\in\Gamma^\infty(T(TM))$ is a vector field on TM, we view it as a map $S: TM\to T(TM)\subseteq(\mathbb{R}^{n\times k})^4$ usually not suppressing the "foot point" $(X,V)\in TM$.

For a smooth function $F\colon\mathbb{R}^{n\times k}\to\mathbb{R}$ we write $\nabla F(X)$ for the gradient of F at $X\in\mathbb{R}^{n\times k}$ with respect to the Frobenius scalar product, i.e. the unique matrix $\nabla F(X)\in\mathbb{R}^{n\times k}$ with

$$\mathrm{d}\,F\big|_X(V)=\mathrm{tr}\left((\nabla F(X))^\top V\right) \tag{2.13}$$

for all $V\in\mathbb{R}^{n\times k}$. Furthermore $E_{ij}\in\mathbb{R}^{n\times k}$ denotes the matrix whose entries fulfill $(E_{ij})_{f\ell}=\delta_{if}\delta_{j\ell}$ for all $f\in\{1,\dots,n\}$ and $\ell\in\{1,\dots,k\}$ with δ_{if} and $\delta_{j\ell}$ being Kronecker deltas.

Finally, following the convention in [17, Chap. 2], a scalar product is a non-degenerated symmetric bilinear form. Moreover, an inner product is a positive definite symmetric bilinear form.

3. A family of metrics on the stiefel manifold

We start with investigating a $2k$-parameter family of symmetric covariant 2- tensors on $\mathbb{R}^{n\times k}$. For certain choices of these parameters, it defines a pseudo-Riemannian metric on an open subset $U \subseteq \mathbb{R}^{n\times k}$ such that $\mathrm{St}_{n,k} \subseteq U$ becomes a pseudo-Riemannian submanifold of U.

3.1. *A symmetric 2-Tensor on $\mathbb{R}^{n\times k}$ and its pull-back to $\mathrm{St}_{n,k}$*

We introduce a $2k$-parameter family of symmetric covariant 2-tensors on $\mathbb{R}^{n\times k}$.

Lemma 3.1. *Let $D = \mathrm{diag}(D_{11}, \dots, D_{kk}) \in \mathbb{R}^{k\times k}$ and $E = \mathrm{diag}(E_{11}, \dots, E_{kk}) \in \mathbb{R}^{k\times k}$ be both diagonal. Then the point-wise definition*

$$\langle V, W\rangle_X^{D,E} = \mathrm{tr}\,(V^\top W D) + \mathrm{tr}\,(V^\top X X^\top W E) \tag{3.1}$$

with $X \in \mathbb{R}^{n\times k}$ and $V, W \in T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ yields a smooth covariant 2-tensor $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathbb{R}^{n\times k}))$ which is invariant under the $O(n)$-action Ψ defined in (2.8).

Proof. Obviously, (3.1) defines a smooth covariant 2-tensor. Let $R \in O(n)$. Then $\Psi_R^*\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} = \langle\cdot,\cdot\rangle_{(\cdot)}^{D,E}$ holds due to

$$\langle\, \mathrm{D}\,\Psi_R(X)V, \mathrm{D}\,\Psi_R(X)W\rangle_{\Psi_R(X)}^{D,E} = \langle RV, RW\rangle_{RX}^{D,E} = \langle V, W\rangle_X^{D,E}$$

for $X \in \mathbb{R}^{n\times k}$ and $V, W \in T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ showing the Ψ-invariance of $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E}$.

Remark 3.2. Observe that the diagonal entry $E_{ii} \in \mathbb{R}$ of the diagonal matrix $E = \mathrm{diag}(E_{11}, \dots, E_{kk}) \in \mathbb{R}^{k\times k}$ shall not be confused with the matrix $E_{ii} \in \mathbb{R}^{n\times k}$ introduced at the end of Section 2. In the sequel, it should be clear by the context how the symbol E_{ii} has to be understood.

Remark 3.3. Let $E = 0$. Then $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathbb{R}^{n\times k}))$ becomes independent of $X \in \mathbb{R}^{n\times k}$. Hence we may identify $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,0}$ with the symmetric bilinear form

$$\langle\cdot,\cdot\rangle^D \colon \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \to \mathbb{R}, \quad (V, W) \mapsto \langle V, W\rangle^D = \mathrm{tr}(V^\top W D). \tag{3.2}$$

If we want to emphasize that $\langle\cdot,\cdot\rangle^D$ is a symmetric bilinear form on $\mathbb{R}^{n\times k}$, we denote it by $\langle\cdot,\cdot\rangle_{\mathbb{R}^{n\times k}}^D$.

Remark 3.4. The pull-back $\iota^*\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ of $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E}$ with $\iota\colon \mathrm{St}_{n,k} \to \mathbb{R}^{n\times k}$ simplifies for the following values of k:

1. For $k = n$ one has $\mathrm{St}_{n,n} = O(n)$. Thus for $X \in O(n)$ and $V, W \in T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ one obtains

$$\langle V, W\rangle_X^{D,E} = \mathrm{tr}\,(V^\top W(D + E)) = \langle V, W\rangle^{D+E} \tag{3.3}$$

   due to $X^\top X = XX^\top = I_n$, i.e. $\iota^*\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} = \langle\cdot,\cdot\rangle^{D+E}$ holds.
2. For $k = 1$ one has $\mathrm{St}_{n,1} = S^{n-1} \subseteq \mathbb{R}^n$. Using $X^\top V = 0$ for all $X \in S^{n-1}$ and $V \in T_X S^{n-1}$ yields

$$\langle V, W\rangle_X^{D,E} = \langle V, W\rangle^D, \tag{3.4}$$

   i.e. $\iota^*\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E} = \langle\cdot,\cdot\rangle^D$ holds.

Remark 3.5. The pull-back $\iota^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ yields well-known metrics on $\mathrm{St}_{n,k}$ for certain choices of D and E:

1. For $D = I_k$ and $E = 0$ one obtains the Euclidean metric, see e.g. [10], [18, Sec. 23.5] or [2]
2. Setting $D = I_k$ and $E = -\frac{1}{2}I_k$ yields the canonical metric, see e.g. [10] or [18, Sec. 23.5]
3. For $D = 2I_k$ and $E = \nu I_k$ with $\nu = -\frac{2\alpha+1}{\alpha+1}$ and $\alpha \in \mathbb{R} \setminus \{-1\}$ the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ reproduces a one-parameter family which has been introduced in [13], see in particular [13, Eq. (55)].

In order to investigate $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathbb{R}^{n\times k}))$ and its pull-back to $\mathrm{St}_{n,k}$ we first list some properties of $\langle\cdot,\cdot\rangle^D$.

Lemma 3.6. *Let $D = \operatorname{diag}(D_{11},\dots,D_{kk}) \in \mathbb{R}^{k\times k}$ be diagonal. The following assertions are fulfilled:*

1. *The symmetric bilinear form $\langle\cdot,\cdot\rangle^D : \mathbb{R}^{n\times k}\times\mathbb{R}^{n\times k} \to \mathbb{R}$ is a scalar product iff D is invertible.*
2. *The bilinear form $\langle\cdot,\cdot\rangle^D : \mathbb{R}^{n\times k}\times\mathbb{R}^{n\times k} \to \mathbb{R}$ is an inner product iff $D_{ii} > 0$ holds for all $i \in \{1,\dots,k\}$.*
3. *Assume that D is invertible. Then $\langle\cdot,\cdot\rangle^D : \mathbb{R}^{k\times k}\times\mathbb{R}^{k\times k} \to \mathbb{R}$ induces a scalar product on $\mathfrak{so}(k)$ iff*
$$D_{ii} + D_{jj} \neq 0 \tag{3.5}$$
holds for all $i, j \in \{1,\dots,k\}$. This condition is always satisfied for $k = 1$.
4. *Let $k \geq 2$. Then $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} : \mathfrak{so}(k)\times\mathfrak{so}(k) \to \mathbb{R}$ defines an inner product on $\mathfrak{so}(n)$ iff*
$$D_{ii} + D_{jj} > 0 \tag{3.6}$$
holds for all $1 \leq i < j \leq k$. For $k = 1$, this bilinear form defines always an inner product.

Proof. Let $E_{ij} \in \mathbb{R}^{n\times k}$ denote the matrix whose entries fulfill $(E_{ij})_{f\ell} = \delta_{if}\delta_{j\ell}$. Clearly, the set
$$B = \{E_{ij} \mid i \in \{1,\dots,n\} \text{ and } j \in \{1,\dots,k\}\}$$
defines a basis of $\mathbb{R}^{n\times k}$. Thus it suffices to show that for all $E_{ij} \in B$ the associated linear forms
$$\mathbb{R}^{n\times k} \to \mathbb{R}, \quad V \mapsto \langle E_{ij}, V\rangle^D \tag{3.7}$$
are non-zero iff D is invertible. We have
$$\langle E_{ij}, V\rangle^D = \operatorname{tr}(E_{ij}^\top V D) = V_{ij}D_{jj} \tag{3.8}$$
with $V = (V_{ij}) \in \mathbb{R}^{n\times k}$. Equation (3.8) implies that D is invertible iff the linear forms in (3.7) are non-vanishing for all $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,k\}$ showing Claim 1.

Next we prove Claim 2. Let $0 \neq V = (V_{ij}) \in \mathbb{R}^{n\times k}$. Then $\langle V, V\rangle^D > 0$ holds iff $D_{ii} > 0$ for $i \in \{1,\dots,k\}$ due to
$$\langle V, V\rangle^D = \operatorname{tr}(V^\top V D) = \sum_{i=1}^{k}\sum_{j=1}^{n} V_{ji}^2 D_{ii}.$$

We now prove Claim 3. For $k = 1$ the assertion is trivial due to $\dim(\mathfrak{so}(1)) = 0$. For $k \geq 2$ the set $\{E_{ij} - E_{ji} \mid 1 \leq i < j \leq k\}$ is a basis of $\mathfrak{so}(k)$. Thus $\langle\cdot,\cdot\rangle^D$ induces a scalar product on $\mathfrak{so}(k)$ iff the linear forms
$$\mathfrak{so}(k) \to \mathbb{R}, \quad A \mapsto \langle E_{ij} - E_{ji}, A\rangle^D$$

are non-vanishing for all $1 \le i < j \le k$. Writing $A = (A_{ij}) = (-A_{ji}) \in \mathfrak{so}(k)$ we compute

$$\langle E_{ij} - E_{ji}, A\rangle^D = \langle E_{ij}, A\rangle^D - \langle E_{ji}, A\rangle^D = A_{ij}D_{jj} - A_{ji}D_{ii} = A_{ij}(D_{jj} + D_{ii})$$

showing that $\langle \cdot, \cdot\rangle^D$ defines a scalar product on $\mathfrak{so}(k)$ iff

$$D_{ii} + D_{jj} \neq 0, \quad i, j \in \{1, \dots, k\}$$

holds. Here we exploited that $D_{ii} + D_{ii} \neq 0$ is automatically fulfilled because D is invertible.

It remains to prove Claim 4. The case $k = 1$ is trivial due to $\mathfrak{so}(1) = \{0\}$. Thus assume $k \ge 2$. Let $A = (A_{ij}) \in \mathfrak{so}(k)$. Exploiting $A_{ij} = -A_{ji}$ we calculate

$$\langle A, A\rangle^D = \tfrac{1}{2}\operatorname{tr}(A^\top AD) + \tfrac{1}{2}\operatorname{tr}(A^\top AD) = \tfrac{1}{2}\sum_{i,j=1}^{k} A_{ij}^2(D_{ii} + D_{jj}). \tag{3.9}$$

Using $A_{ii} = 0$ we conclude that $\langle A, A\rangle^D > 0$ holds for all $0 \neq A \in \mathfrak{so}(k)$ iff $D_{ii} + D_{jj} > 0$ is fulfilled for all $1 \le i < j \le k$.

The next lemma shows that $\langle \cdot, \cdot\rangle^{D,E}_{(\cdot)}$ induces a pseudo-Riemannian metric on the Stiefel manifold for certain choices of D and E.

Lemma 3.7. *Let $D = \operatorname{diag}(D_{11}, \dots, D_{kk}) \in \mathbb{R}^{k\times k}$ and $E = \operatorname{diag}(E_{11}, \dots, E_{kk}) \in \mathbb{R}^{k\times k}$ be both diagonal and let $X \in \mathrm{St}_{n,k}$. Then the following assertions are fulfilled:*

1. *Let $1 \le k < n$. The bilinear form*
$$\langle \cdot, \cdot\rangle^{D,E}_X \colon T_X\mathbb{R}^{n\times k} \times T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \to \mathbb{R} \tag{3.10}$$
is a scalar product iff D and $D + E$ are both invertible. For $k = n$ the bilinear form in (3.10) defines a scalar product iff $D + E$ is invertible.
2. *Assume that (3.10) defines a scalar product. Then the pull-back $\iota^*\langle \cdot, \cdot\rangle^{D,E}_{(\cdot)}$ to $\mathrm{St}_{n,k}$ defines a pseudo-Riemannian metric on $\mathrm{St}_{n,k}$, i.e.*
$$\langle \cdot, \cdot\rangle^{D,E}_X \colon T_X\mathrm{St}_{n,k} \times T_X\mathrm{St}_{n,k} \to \mathbb{R} \tag{3.11}$$
is a scalar product on $T_X\mathrm{St}_{n,k}$, iff the condition
$$D_{ii} + E_{ii} + D_{jj} + E_{jj} \neq 0, \quad i, j \in \{1, \dots, k\} \tag{3.12}$$
holds.
3. *Assume that (3.10) defines a scalar product. For $2 \le k \le n-1$ the symmetric covariant 2-tensor $\iota^*\langle \cdot, \cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ is a Riemannian metric on $\mathrm{St}_{n,k}$, i.e.*
$$\langle \cdot, \cdot\rangle^{D,E}_X \colon T_X\mathrm{St}_{n,k} \times T_X\mathrm{St}_{n,k} \to \mathbb{R} \tag{3.13}$$
is an inner product on $T_X\mathrm{St}_{n,k}$, iff the conditions $D_{ii} > 0$ for all $i \in \{1, \dots, k\}$ and
$$D_{ii} + E_{ii} + D_{jj} + E_{jj} > 0, \quad 1 \le i < j \le k \tag{3.14}$$
are fulfilled. For $k = 1$ one obtains a Riemannian metric iff $D_{11} > 0$ holds. For $k = n$ the tensor $\iota^*\langle \cdot, \cdot\rangle^{D,E}_{(\cdot)}$ defines a Riemannian metric iff $D_{ii} + E_{ii} + D_{jj} + E_{jj} > 0$ holds for all $1 \le i < j \le n$.

Proof. Since the $O(n)$-left action Ψ on $\mathbb{R}^{n\times k}$ defined in (2.8) is isometric with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ by Lemma 3.1 and, moreover, Ψ restricts to a transitive action on $\mathrm{St}_{n,k}$ it suffices to prove the claims for a single point $X_0 \in \mathrm{St}_{n,k}$.

We first consider the case $k = n$. Then $\langle\cdot,\cdot\rangle^{D,E}_X = \langle\cdot,\cdot\rangle^{D+E}$ holds for all $X \in \mathrm{St}_{n,n} = O(n)$ by Remark 3.4, Claim 1. Hence $\langle\cdot,\cdot\rangle^{D,E}_X$ is non-degenerated iff $D+E$ is invertible according to Lemma 3.6, Claim 1. Next we consider the case $1 \le k < n$. We choose $X_0 = I_{n,k}$, where

$$I_{n,k} = \begin{bmatrix} I_k \\ 0 \end{bmatrix} \in \mathrm{St}_{n,k},$$

and write

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \in \mathbb{R}^{n\times k} \quad \text{and} \quad W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \in \mathbb{R}^{n\times k}$$

with $V_1, W_1 \in \mathbb{R}^{k\times k}$ and $V_2, W_2 \in \mathbb{R}^{(n-k)\times k}$. By this notation and identifying $T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ we calculate

$$\begin{aligned} \langle V, W\rangle^{D,E}_{I_{n,k}} &= \mathrm{tr}\Big(\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^\top \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} D\Big) + \mathrm{tr}\Big(\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^\top \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} E\Big) \\ &= \mathrm{tr}\,(V_1^\top W_1(D+E)) + \mathrm{tr}\,(V_2^\top W_2 D). \end{aligned} \tag{3.15}$$

By (3.15) and Lemma 3.6, Claim 1, the bilinear form $\langle\cdot,\cdot\rangle^{D,E}_{I_{n,k}}$ defines a scalar product on $T_X\mathbb{R}^{n\times k}$ iff D and $D+E$ are both invertible.

Next we assume that D and $D+E$ are choosen such that $\langle\cdot,\cdot\rangle^{D,E}_X$ defines a scalar product on $T_X\mathbb{R}^{n\times k}$ for each $X \in \mathrm{St}_{n,k}$. We now prove Claim 2 for $1 \le k \le n-1$. To this end, it is sufficient to show that

$$\langle\cdot,\cdot\rangle^{D,E}_{I_{n,k}} \colon T_{I_{n,k}}\mathrm{St}_{n,k} \times T_{I_{n,k}}\mathrm{St}_{n,k} \to \mathbb{R} \tag{3.16}$$

is a scalar product iff (3.12) holds. The tangent space $T_{I_{n,k}}\mathrm{St}_{n,k}$ is given by

$$T_{I_{n,k}}\mathrm{St}_{n,k} = \Big\{\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \Big| V_1 \in \mathfrak{so}(k) \text{ and } V_2 \in \mathbb{R}^{(n-k)\times k}\Big\} \subseteq T_{I_{n,k}}\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}, \tag{3.17}$$

see e.g. [10, Sec. 2.2.1]. Thus we may write $V, W \in T_X\mathrm{St}_{n,k}$ as

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \in \mathbb{R}^{n\times k} \quad \text{and} \quad W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \in \mathbb{R}^{n\times k}$$

with $V_1, W_1 \in \mathfrak{so}(k)$ and $V_2, W_2 \in \mathbb{R}^{(n-k)\times k}$. We now obtain

$$\iota^*\langle V, W\rangle^{D,E}_{I_{n,k}} = \mathrm{tr}\,(V_1^\top W_1(D+E)) + \mathrm{tr}\,(V_2^\top W_2 D) \tag{3.18}$$

analogously to (3.15). Clearly, Equation (3.18) defines a scalar product on $T_{I_{n,k}}\mathrm{St}_{n,k}$ iff

$$\mathfrak{so}(k) \times \mathfrak{so}(k) \to \mathbb{R}, \quad (V_1, W_1) \mapsto \mathrm{tr}\,(V_1^\top W_1(D+E))$$

yields a scalar product on $\mathfrak{so}(k)$ and

$$\mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{(n-k)\times k} \to \mathbb{R}, \quad (V_2, W_2) \mapsto \mathrm{tr}(V_2^\top W_2 D)$$

defines a scalar product on $\mathbb{R}^{(n-k)\times k}$. By applying Lemma 3.6, Claim 3 we obtain the desired result. Next we consider the case $k = n$. By exploiting the $O(n)$-invariance of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ and $T_{I_n}\mathrm{St}_{n,n} = \mathfrak{so}(n)$ as well as $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} = \langle\cdot,\cdot\rangle^{D+E}$ for $k = n$, Claim 2 follows by Lemma 3.6, Claim 3.

It remains to prove Claim 3. We first consider the case $2 \leq k \leq n-1$. Since the bilinear form on $T_{I_{n,k}}\mathrm{St}_{n,k}$ induced by $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is given by (3.18), the desired result is a consequence of Lemma 3.6, Claim 2 and Lemma 3.6, Claim 4. For $k = 1$, we observe that $\iota^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is independent of E due to $X^\top V = 0$ for all $X \in \mathrm{St}_{n,1}$ and $V \in T_X\mathrm{St}_{n,1}$, see also Remark 3.4, Claim 2. Hence (3.18) implies that $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is positive definite iff

$$\mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{(n-k)\times k} \to \mathbb{R}, \quad (V_2, W_2) \mapsto \mathrm{tr}(V_2^\top W_2 D)$$

is positive definite. The desired result follows by Lemma 3.6, Claim 2. For $k = n$, the assertion holds due to $\langle\cdot,\cdot\rangle^{D,E}_X = \langle\cdot,\cdot\rangle^{D+E}$ for all $X \in \mathrm{St}_{n,n} = O(n)$ by Lemma 3.6, Claim 3.

The next lemma generalizing [14, Lem. 2] shows that there is an open neigbourhood $U \subseteq \mathbb{R}^{n\times k}$ of $\mathrm{St}_{n,k}$ such that $\mathrm{St}_{n,k} \subseteq (U, \iota_U^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ is a pseudo-Riemannian submanifold. This fact is crucial for the following discussion.

Lemma 3.8. *Let $D, E \in \mathbb{R}^{k\times k}$ be both diagonal such that for each $X \in \mathrm{St}_{n,k}$*

$$\langle\cdot,\cdot\rangle^{D,E}_X \colon T_X\mathbb{R}^{n\times k} \times T_X\mathbb{R}^{n\times k} \to \mathbb{R} \tag{3.19}$$

defines a scalar product on $T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ which induces a scalar product on $T_X\mathrm{St}_{n,k} \subseteq T_X\mathbb{R}^{n\times k}$. Then there exists an open neighbourhood $U \subseteq \mathbb{R}^{n\times k}$ of $\mathrm{St}_{n,k}$ such that $\iota_U^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*U))$ is a pseudo-Riemannian metric on U and $(\mathrm{St}_{n,k}, \iota^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ is a pseudo-Riemannian submanifold of $(U, \iota_U^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$.

Proof. We identify $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathbb{R}^{n\times k}))$ with the continuous map

$$\varphi \colon \mathbb{R}^{n\times k} \to \mathrm{S}^2((\mathbb{R}^{n\times k})^*), \quad X \mapsto \langle\cdot,\cdot\rangle^{D,E}_X = ((V, W) \mapsto \langle V, W\rangle^{D,E}_X).$$

The bilinear form $\varphi(X) = \langle\cdot,\cdot\rangle^{D,E}_X \in \mathrm{S}^2((\mathbb{R}^{n\times k})^*)$ is a scalar product for all $X \in \mathrm{St}_{n,k}$ by assumption. Hence, by the continuity of φ, there is an on open neighbourhood U_X of X in $\mathbb{R}^{n\times k}$ such that $\varphi(\widetilde{X}) \in \mathrm{S}^2((\mathbb{R}^{n\times k})^*)$ is non-degnerated for all $\widetilde{X} \in U_X$. We set

$$U = \bigcup_{X\in\mathrm{St}_{n,k}} U_X.$$

Then $U \subseteq \mathbb{R}^{n\times k}$ is open as a union of open sets and fulfills $\mathrm{St}_{n,k} \subseteq U$ by definition. Moreover, $\varphi(\widetilde{X})$ is non-dengenerated for all $\widetilde{X} \in U$ by construction. Hence $\iota_U^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ defines a pseudo-Riemannian metric on U such that $\mathrm{St}_{n,k} \subseteq (U, \iota_U^*\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ is a pseudo-Riemannian submanifold.

Obviously, the inclusion $\mathrm{St}_{n,k} \subseteq U$ from Lemma 3.8 is always proper since $\mathrm{St}_{n,k}$ is closed in $\mathbb{R}^{n\times k}$ while U is open in $\mathbb{R}^{n\times k}$.

Notation 3.9. From now on, unless indicated otherwise, pull-backs of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ to submanifolds of $\mathbb{R}^{n\times k}$ are suppressed in the notation.

3.2. Number of parameters

In the case $k = n$, the $2k$-parameter family of covariant 2-tensors $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is actually a k-parameter family by Remark 3.4, Claim 1. Indeed, $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ depends only on $D + E$. Hence one may ask if there exits always such an over-parameterization.

Lemma 3.10. *Let $D = \operatorname{diag}(D_{11}, \ldots, D_{kk}) \in \mathbb{R}^{k\times k}$ be some diagonal matrix. Then the following assertions are fulfilled:*

1. *The bilinear form $\langle\cdot,\cdot\rangle^D : \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \to \mathbb{R}$ vanishes identically iff $D = 0$ holds.*
2. *The restriction $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} : \mathfrak{so}(k) \times \mathfrak{so}(k) \to \mathbb{R}$ of $\langle\cdot,\cdot\rangle^D : \mathbb{R}^{k\times k} \times \mathbb{R}^{k\times k} \to \mathbb{R}$ fulfills the following assertions:*
   (a) *For $k = 1$ one has $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} = 0$ for all $D \in \mathbb{R}^{1\times1} \cong \mathbb{R}$.*
   (b) *For $k = 2$ one has $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} = 0$ iff $D_{11} + D_{22} = 0$ holds.*
   (c) *For $k \geq 3$ one has $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} = 0$ iff $D = 0$ holds.*

Proof. Let $E_{ij} \in \mathbb{R}^{n\times k}$ the matrix whose entries fulfill $(E_{ij})_{f\ell} = \delta_{if}\delta_{j\ell}$. Then

$$\langle E_{ij}, V\rangle^D = V_{ij}D_{jj}, \quad i \in \{1, \ldots, n\},\ j \in \{1, \ldots, k\}, \tag{3.20}$$

where $V = (V_{ij}) \in \mathbb{R}^{n\times k}$. Since $\langle\cdot,\cdot\rangle^D = 0$ holds iff the linear forms $\langle E_{ij}, \cdot\rangle^D : \mathbb{R}^{n\times k} \to \mathbb{R}$ vanishes for all $1 \leq i \leq n$ and $1 \leq j \leq k$, the first claim follows by (3.20).

Next, we consider $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} : \mathfrak{so}(k) \times \mathfrak{so}(k) \to \mathbb{R}$. Clearly, it vanishes for $k = 1$ for all $D \in \mathbb{R}^{1\times1}$ due to $\mathfrak{so}(1) = \{0\}$. We now assume $k \geq 2$. Then $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} : \mathfrak{so}(k) \times \mathfrak{so}(k) \to \mathbb{R}$ vanishes iff the linear forms

$$\langle E_{ij} - E_{ji}, \cdot\rangle^D : \mathfrak{so}(k) \to \mathbb{R} \tag{3.21}$$

vanish for all $1 \leq i < j \leq k$. Writing $A = (A_{ij}) = (-A_{ji}) \in \mathfrak{so}(k)$ we obtain

$$\langle E_{ij} - E_{ji}, A\rangle^D = A_{ij}D_{ii} - A_{ji}D_{jj} = A_{ij}(D_{ii} + D_{jj}).$$

Thus the linear forms (3.21) are zero iff $D_{ii} + D_{jj} = 0$ holds for all $1 \leq i < j \leq k$. For $k = 2$ this is equivalent to $D_{11} + D_{22} = 0$. It remains to consider the case $k \geq 3$. The conditions $D_{ii} + D_{jj} = 0$ for all $1 \leq i < j \leq k$ include the conditions

$$D_{11} + D_{ii} = 0 \iff D_{11} = -D_{ii} \quad \text{for all} \quad 2 \leq i \leq k \tag{3.22}$$

and

$$D_{(k-1)(k-1)} + D_{kk} = 0. \tag{3.23}$$

In particular $D_{11} = -D_{k-1}$ and $D_{11} = -D_{kk}$ holds. Plugging these identities into (3.23) yields

$$-D_{11} - D_{11} = -2D_{11} = 0 \iff D_{11} = 0.$$

Hence (3.22) implies $D_{ii} = 0$ for all $2 \leq i \leq k$. Therefore $\langle\cdot,\cdot\rangle^D\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} = 0$ iff $D = 0$ as desired. ∎

The next lemma justifies calling $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ a $2k$-parameter family provided that $3 \leq k \leq n - 1$ holds.

Lemma 3.11. *Let*

$$\mathbb{R}^{k\times k}_{\mathrm{diag}} = \{\operatorname{diag}(D_{11},\ldots,D_{kk}) \mid D_{11},\ldots,D_{kk} \in \mathbb{R}\} \subseteq \mathbb{R}^{k\times k}$$

denote the k-dimensional real vector space of $(k\times k)$*-diagonal matrices. Moreover, define*

$$\psi\colon \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \to \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k})), \quad (D,E) \mapsto \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}. \tag{3.24}$$

Then ψ *is a linear map which fulfills the following assertions depending on k and n:*

1. *For* $k = 1 = n$*, one has* $\dim(\operatorname{im}(\psi)) = 0$ *and* $\ker(\psi) = \mathbb{R}\times\mathbb{R}$.
2. *For* $k = 1$ *and* $n > 1$ *one has* $\dim(\operatorname{im}(\psi)) = 1$ *and* $\ker(\psi) = \{(0,E) \mid E \in \mathbb{R}\} \subseteq \mathbb{R}\times\mathbb{R}$.
3. *For* $k = 2 = n$ *one has* $\dim(\operatorname{im}(\psi)) = 1$ *and*
$$\ker(\psi) = \{((D_{11},D_{22}),(E_{11},-D_{11}-D_{22}-E_{11})) \mid D_{11},D_{22},E_{11} \in \mathbb{R}\} \subseteq \mathbb{R}^{2\times 2}_{\mathrm{diag}} \times \mathbb{R}^{2\times 2}_{\mathrm{diag}}.$$
4. *For* $2 < k < n$ *one has* $\dim(\operatorname{im}(\psi)) = 2k$ *and* $\ker(\psi) = \{0\} \subseteq \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}}$.
5. *For* $k = n > 2$ *one has* $\dim(\operatorname{im}(\psi)) = k$ *and* $\ker(\psi) = \{(D,-D) \mid D \in \mathbb{R}^{k\times k}_{\mathrm{diag}}\} \subseteq \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}}$.

Proof. Clearly, the map ψ is linear. Next we define the linear map

$$\widetilde{\psi}\colon \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \to \mathrm{S}^2(T^*_{I_{n,k}}\mathrm{St}_{n,k}), \quad (D,E) \mapsto \langle\cdot,\cdot\rangle^{D,E}_{I_{n,k}}.$$

Obviously, for each $(D,E) \in \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}}$ one has $(\psi(D,E))(I_{n,k}) = \langle\cdot,\cdot\rangle^{D,E}_{I_{n,k}} = \widetilde{\psi}(D,E)$. Since $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is invariant under the transitive $O(n)$-action Ψ on $\mathrm{St}_{n,k}$ according to Lemma 3.1, this yields

$$(D,E) \in \ker(\psi) \iff (D,E) \in \ker(\widetilde{\psi}). \tag{3.25}$$

Moreover, the equivalence

$$(D,E) \in \ker(\widetilde{\psi}) \quad \iff \quad (\langle V,W\rangle^{D,E}_{I_{n,k}} = 0 \ \text{ for all } \ V,W \in T_{I_{n,k}}\mathrm{St}_{n,k}) \tag{3.26}$$

is clearly fulfilled. We again write

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \in T_{I_{n,k}}\mathrm{St}_{n,k} \quad \text{and} \quad W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \in T_{I_{n,k}}\mathrm{St}_{n,k}$$

with $V_1, W_1 \in \mathfrak{so}(k)$ and $V_2, W_2 \in \mathbb{R}^{(n-k)\times k}$. By this notation and the description of $\ker(\widetilde{\psi})$ from (3.26), we study each case separately:

1. Obviously, for $k = 1 = n$ the claim $\ker(\widetilde{\psi}) = \mathbb{R}\times\mathbb{R}$ is correct due to $T_{I_1}\mathrm{St}_{1,1} = \{0\}$ implying $\dim(S^2(T^*_{I_1}\mathrm{St}_{1,1})) = 0$.
2. For $k = 1$ and $n > 1$ we have
$$(\widetilde{\psi}(D,E))(V,W) = \operatorname{tr}(V_1^\top W_1(D+E)) + \operatorname{tr}(V_2^\top W_2 D) = \langle V_1,W_1\rangle^{D+E}\big|_{\mathfrak{so}(1)\times\mathfrak{so}(1)} + \langle V_2,W_2\rangle^{D}_{\mathbb{R}^{(n-1)\times 1}}. \tag{3.27}$$
Clearly, Equation (3.27) vanishes iff $D = 0$ holds independent of the value of $D+E$ by Lemma 3.10. Hence the kernel of ψ is given by $\ker(\widetilde{\psi}) = \{(0,E) \mid E \in \mathbb{R}\}$
3. For $k = 2 = n$ we have
$$(\widetilde{\psi}(D,E))(V,W) = \langle V,W\rangle^{D+E}\big|_{\mathfrak{so}(2)\times\mathfrak{so}(2)}.$$
Lemma 3.10 yields $\widetilde{\psi}(D,E) = 0$ iff $(D+E)_{11} + (D+E)_{22} = 0$ is fulfilled. Therefore we obtain
$$\ker(\widetilde{\psi}) = \{((D_{11},D_{22}),(E_{11},-D_{11}-D_{22}-E_{11})) \mid D_{11},D_{22},E_{11} \in \mathbb{R}\}.$$

4. We now consider the case $3 \leq k \leq n-1$. Then one has

$$(\widetilde{\psi}(D,E))(V,W) = \operatorname{tr}(V_1^\top W_1(D+E)) + \operatorname{tr}(V_2^\top W_2 D) = \langle V_1, W_1\rangle^{D+E}\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)} + \langle V_2, W_2\rangle^D_{\mathbb{R}^{(n-k)\times k}}.$$

   By Lemma 3.10, we have $\widetilde{\psi}(D,E) = 0$ iff $D = 0$ and $D + E = 0$ holds. Therefore the kernel of ψ is given by $\ker(\widetilde{\psi}) = \{(D,E) \in \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \mid D = 0 = E\} = \{0\}$.
5. It remains to consider the case $k = n \geq 3$. We obtain

$$(\widetilde{\psi}(D,E))(V,W) = \operatorname{tr}(V^\top W(D+E)) = \langle V, W\rangle^{D+E}\big|_{\mathfrak{so}(k)\times\mathfrak{so}(k)}.$$

   for all $V, W \in T_{I_n}\mathrm{St}_{n,n} = \mathfrak{so}(n)$. Thus $\widetilde{\psi}(D,E) = 0$ holds iff $D + E = 0$ is fulfilled by Lemma 3.10. Hence the kernel of $\widetilde{\psi}$ is given by $\ker(\widetilde{\psi}) = \{(D,-D) \mid D \in \mathbb{R}^{k\times k}_{\mathrm{diag}}\}$.

The equality $\ker(\psi) = \ker(\widetilde{\psi})$ is satisfied according to (3.25). Moreover, we have

$$\dim(\operatorname{im}(\psi)) = \dim(\mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}}) - \dim(\ker(\psi)) = 2k - \dim(\ker(\widetilde{\psi}))$$

as desired.

Remark 3.12. Lemma 3.7, Claim 3 shows that the set of all parameters

$$\{(D,E) \in \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \mid \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \text{ defines a pseudo-Riemannian metric on } \mathrm{St}_{n,k}\}$$

contains the non-empty subset $\{(D,E) \in \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \mid D_{ii} > 0 \text{ and } E_{ii} > 0 \text{ for all } i \in \{1,\ldots,k\}\}$ which is open in $\mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}}$. Moreover, the linear map $\psi\colon \mathbb{R}^{k\times k}_{\mathrm{diag}} \times \mathbb{R}^{k\times k}_{\mathrm{diag}} \to \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ is injective for $2 < k < n$ according to Lemma 3.11. This point of view justifies calling $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ a $2k$-parameter family at least for $2 < k < n$. For other choices of k and n one has rather a $(\dim(\operatorname{im}(\psi)))$-parameter family of metrics. However, ignoring this over parameterization, we call them $2k$-parameter family, nevertheless.

3.3. Orthogonal projections onto tangent spaces

The Stiefel manifold $\mathrm{St}_{n,k}$ endowed with $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ can be viewed as a pseudo-Riemannian submanifold of $(U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ with some suitable open $U \subseteq \mathbb{R}^{n\times k}$ by Lemma 3.8. Consequently, for any given point $X \in \mathrm{St}_{n,k}$, we may consider the orthogonal projection

$$P_X\colon T_X\mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k} \subseteq \mathbb{R}^{n\times k},$$

where $T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ is endowed with the scalar product $\langle\cdot,\cdot\rangle^{D,E}_X$. Moreover, it makes sense to consider the normal space $N_X\mathrm{St}_{n,k} = (T_X\mathrm{St}_{n,k})^\perp \subseteq \mathbb{R}^{n\times k}$ with respect to $\langle\cdot,\cdot\rangle^{D,E}_X\colon T_X\mathbb{R}^{n\times k} \times T_X\mathbb{R}^{n\times k} \to \mathbb{R}$.

Notation 3.13. From now on, unless indicated otherwise, we always assume that $D, E \in \mathbb{R}^{k\times k}$ are both diagonal matrices such that $\langle\cdot,\cdot\rangle^{D,E}_X$ defines a scalar product on $\mathbb{R}^{n\times k}$ for each $X \in \mathrm{St}_{n,k}$ and $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ induces a pseudo-Riemannian metric on $\mathrm{St}_{n,k}$. In particular, we may assume that D and $D+E$ are both invertible. In view of Lemma 3.7, Claim 1 this assumption is of no restriction. For the case $k = n$, we replace D by $D + E$ and E by 0, if necessary.

Lemma 3.14. *Let $D = \operatorname{diag}(D_{11},\ldots,D_{kk}) \in \mathbb{R}^{k\times k}$ be invertible such that $D_{ii} + D_{jj} \neq 0$ holds for all $i, j \in \{1,\ldots,k\}$. Then the following assertions are fulfilled:*

1. *The orthogonal complement of* $\mathfrak{so}(k)$ *in* $\mathbb{R}^{k\times k}$ *with respect to the scalar product* $\langle\cdot,\cdot\rangle^D$ *is given by*

$$\mathfrak{so}(k)^{\perp_D} = \{A \in \mathbb{R}^{k\times k} \mid AD = (AD)^\top\} = \{\Lambda D^{-1} \mid \Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\} \subseteq \mathbb{R}^{k\times k}. \tag{3.28}$$

Moreover, $\mathfrak{so}(k) \oplus \mathfrak{so}(k)^{\perp_D} = \mathbb{R}^{k\times k}$ *holds.*

2. *The orthogonal projection*

$$\pi^D \colon \mathbb{R}^{k\times k} \to \mathfrak{so}(k) \subseteq \mathbb{R}^{k\times k}, \quad A \mapsto \pi^D(A) \tag{3.29}$$

onto $\mathfrak{so}(k)$ *with respect* $\langle\cdot,\cdot\rangle^D$ *is entry-wise given by*

$$\pi^D(A)_{ij} = \tfrac{1}{D_{ii}+D_{jj}}(AD - DA^\top)_{ij} = \tfrac{1}{D_{ii}+D_{jj}}(A_{ij}D_{jj} - A_{ji}D_{ii}), \quad i,j \in \{1,\dots,k\}. \tag{3.30}$$

Proof. We first determine $\mathfrak{so}(k)^{\perp_D}$. To this end, we calculate

$$\begin{aligned}\mathfrak{so}(k)^{\perp_D} &= \{A \in \mathbb{R}^{k\times k} \mid \langle A, B\rangle^D = 0 \text{ for all } B \in \mathfrak{so}(n)\}\\ &= \{A \in \mathbb{R}^{k\times k} \mid \operatorname{tr}((AD)^\top B) = 0 \text{ for all } B \in \mathfrak{so}(n)\}\\ &= \{A \in \mathbb{R}^{k\times k} \mid AD = (AD)^\top \in \mathbb{R}^{k\times k}_{\mathrm{sym}} \text{ is symmetric }\}.\end{aligned}$$

Let $\Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}$. Then $(\Lambda D^{-1})D = \Lambda = \Lambda^\top = D(\Lambda D^{-1})^\top$ showing $\{\Lambda D^{-1} \mid \Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\} \subseteq \mathfrak{so}(k)^{\perp_D}$. The equality $\mathfrak{so}(k)^{\perp_D} = \{\Lambda D^{-1} \mid \Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\}$ follows by counting dimensions. By Lemma 3.6, Claim 3 the assumptions on D ensure that $\langle\cdot,\cdot\rangle^D$ induces a scalar product on $\mathfrak{so}(k)$. Hence $\mathfrak{so}(k) \oplus \mathfrak{so}(k)^{\perp_D} = \mathbb{R}^{k\times k}$ holds, see e.g. [17, Chap. 2, Lem. 23].

It remains to prove Claim 2. To this end, we show $\operatorname{im}(\pi^D) = \mathfrak{so}(k)$ and $\ker(\pi^D) = \mathfrak{so}(k)^{\perp_D}$ as well as $\pi^D\big|_{\mathfrak{so}(k)} = \mathrm{id}_{\mathfrak{so}(k)}$. We first prove $\operatorname{im}(\pi^D) \subseteq \mathfrak{so}(n)$. Let $A = (A_{ij}) \in \mathbb{R}^{k\times k}$. We compute

$$((\pi^D(A))^\top)_{ij} = \pi^D(A)_{ji} = \tfrac{1}{D_{jj}+D_{ii}}(A_{ji}D_{ii} - A_{ij}D_{jj}) = -\tfrac{1}{D_{ii}+D_{jj}}(A_{ij}D_{jj} - A_{ji}D_{ii}) = -\pi^D(A)_{ij}.$$

for $i,j \in \{1,\dots,k\}$ showing $\operatorname{im}(\pi^D) \subseteq \mathfrak{so}(k)$. Moreover, for $A \in \mathfrak{so}(k)$, i.e. $A_{ij} = -A_{ji}$, we have

$$\pi^D(A)_{ij} = \tfrac{1}{D_{ii}+D_{jj}}(A_{ij}D_{jj} - (-A_{ij})D_{ii}) = \tfrac{1}{D_{ii}+D_{jj}}A_{ij}(D_{jj} + D_{ii}) = A_{ij}.$$

This yields $\pi^D(A) = A$ for all $A \in \mathfrak{so}(k)$, i.e. $\pi^D\big|_{\mathfrak{so}(k)} = \mathrm{id}_{\mathfrak{so}(k)}$. Moreover, the inclusion $\operatorname{im}(\pi^D) \subseteq \mathfrak{so}(k)$ is in fact an equality. Next let $A \in \mathfrak{so}(k)^{\perp_D}$. Then $AD = DA^\top$ holds according to Claim 1 implying

$$\pi^D(A)_{ij} = \tfrac{1}{D_{ii}+D_{jj}}(AD - DA^\top)_{ij} = 0.$$

Thus $\pi^D\big|_{\mathfrak{so}^{\perp_D}} = 0$ follows.

The formula for π^D can be rewritten in terms of the so-called Hadamard or Schur product. For matrices $A, B \in \mathbb{R}^{k\times k}$, it is entry-wise defined by

$$(A \circledcirc B)_{ij} = A_{ij}B_{ij}, \quad i,j \in \{1,\dots,k\}. \tag{3.31}$$

Remark 3.15. Let $\mu \in \mathbb{R}^{k\times k}$ be defined entry-wise by

$$\mu_{ij} = \tfrac{1}{D_{ii}+D_{jj}}, \quad i,j \in \{1,\dots,k\}. \tag{3.32}$$

Then the projection $\pi^D \colon \mathbb{R}^{n\times k} \to \mathfrak{so}(k)$ from Lemma 3.14 can be rewritten as

$$\pi^D(A) = \mu \circledcirc (AD - DA^\top), \quad A \in \mathbb{R}^{k\times k}. \tag{3.33}$$

Corollary 3.16. *Let $0 \neq \beta \in \mathbb{R}$ and define $D = \beta I_k$. Then, for each $A \in \mathbb{R}^{k\times k}$ the map π^D from Lemma 3.14 simplifies to*

$$\pi^{\beta I_k}(A) = \tfrac{1}{2}(A - A^\top) = \operatorname{skew}(A). \tag{3.34}$$

Proof. The desired result follows by a straightforward calculation exploiting $D_{ii} = \beta \neq 0$ for all $i \in \{1, \dots, k\}$.

We determine the normal spaces of $\mathrm{St}_{n,k}$ with respect $\langle \cdot, \cdot \rangle_{(\cdot)}^{D,E}$ generalizing [19, Chap. 1, Lem. 3.15] and [14, Lem. 3].

Lemma 3.17. *The normal space $N_X\mathrm{St}_{n,k} = (T_X\mathrm{St}_{n,k})^\perp \subseteq T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ at $X \in \mathrm{St}_{n,k}$ with respect to $\langle \cdot, \cdot \rangle_X^{D,E}$ is given by*

$$N_X\mathrm{St}_{n,k} = \{X\Lambda(D+E)^{-1} \in \mathbb{R}^{n\times k} \mid \Lambda = \Lambda^\top \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\}. \tag{3.35}$$

Proof. Clearly, the set $\{X\Lambda(D+E)^{-1} \in \mathbb{R}^{n\times k} \mid \Lambda = \Lambda^\top \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\}$ is a linear subspace of $\mathbb{R}^{n\times k}$ of dimension $(k^2 + k)/2$ being the image of the injective linear map

$$\mathbb{R}^{k\times k}_{\mathrm{sym}} \to \mathbb{R}^{n\times k}, \quad \Lambda \mapsto X\Lambda(D+E)^{-1}.$$

Moreover, every matrix $V = X\Lambda(D+E)^{-1}$ with $\Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}$ is orthogonal to the tangent space $T_X\mathrm{St}_{n,k}$. Indeed, we have for $W \in T_X\mathrm{St}_{n,k}$

$$\begin{aligned}\langle V, W\rangle_X^{D,E} &= \operatorname{tr}\left((X\Lambda(D+E)^{-1})^\top WD\right) + \operatorname{tr}\left((X\Lambda(D+E)^{-1})^\top XX^\top WE\right)\\ &= \operatorname{tr}\left(\Lambda^\top (X^\top W)\right)\\ &= 0\end{aligned}$$

due to $\Lambda = \Lambda^\top$ and $X^\top W = -W^\top X$. Therefore $\{X\Lambda(D+E)^{-1} \in \mathbb{R}^{n\times k} \mid \Lambda = \Lambda^\top \in \mathbb{R}^{k\times k}_{\mathrm{sym}}\} \subseteq N_X\mathrm{St}_{n,k}$ follows. By counting dimensions, this inclusion is in fact an equality.

Theorem 3.18. *Let $X \in \mathrm{St}_{n,k}$. The orthogonal projection of $T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$ onto $T_X\mathrm{St}_{n,k} \subseteq \mathbb{R}^{n\times k}$ with respect to $\langle \cdot, \cdot \rangle_X^{D,E}$ is given by*

$$P_X\colon \mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k} \subseteq \mathbb{R}^{n\times k}, \quad V \mapsto P_X(V) = V - XX^\top V + X\pi^{D+E}(X^\top V). \tag{3.36}$$

Proof. We first show $\operatorname{im}(P_X) = T_X\mathrm{St}_{n,k}$. Let $X \in \mathrm{St}_{n,k}$ and $V \in \mathbb{R}^{n\times k}$. One calculates

$$X^\top(P_X(V)) = X^\top(V - XX^\top V + X\pi^{D+E}(X^\top V)) = X^\top V - X^\top V + \pi^{D+E}(X^\top V) = \pi^{D+E}(X^\top V).$$

Moreover, using $\operatorname{im}(\pi^{D+E}) = \mathfrak{so}(n)$, we obtain

$$(P_X(V))^\top X = (V - XX^\top V + X\pi^{D+E}(X^\top V))^\top X = V^\top X - V^\top X + (\pi^{D+E}(X^\top V))^\top = -\pi^{D+E}(X^\top V).$$

Hence $X^\top(P_X(V)) = \pi^{D+E}(X^\top V) = -(P_X(V))^\top X$ follows, i.e. $\operatorname{im}(P_X) \subseteq T_X\mathrm{St}_{n,k}$ as desired.

We now assume $V \in T_X\mathrm{St}_{n,k}$. By using $X^\top V = -V^\top X$ and $\pi^D\big|_{\mathfrak{so}(n)} = \mathrm{id}_{\mathfrak{so}(n)}$, we calculate

$$P_X(V) = V - XX^\top V + X\pi^D(X^\top V) = V - XX^\top V + X(X^\top V) = V$$

proving $P_X\big|_{T_X\mathrm{St}_{n,k}} = \mathrm{id}_{T_X\mathrm{St}_{n,k}}$ and implying that $\operatorname{im}(P_X) \subseteq T_X\mathrm{St}_{n,k}$ is indeed an equality.

It remains to show $\ker(P_X) = (T_X\mathrm{St}_{n,k})^\perp$. Let $V \in N_X\mathrm{St}_{n,k}$. We may write $V = X\Lambda(D+E)^{-1}$ with some suitable symmetric matrix $\Lambda \in \mathbb{R}^{k\times k}_{\mathrm{sym}}$ by exploiting Lemma 3.17. Consequently, we have

$$\begin{aligned} P_X(V) &= P_X(X\Lambda(D+E)^{-1}) \\ &= X\Lambda(D+E)^{-1} - XX^\top(X\Lambda(D+E)^{-1}) + X\pi^{D+E}(X^\top X\Lambda(D+E)^{-1}) \\ &= X\pi^{D+E}(\Lambda(D+E)^{-1}) \\ &= 0, \end{aligned}$$

by using Lemma 3.14, Claim 1 which shows $\pi^{D+E}(\Lambda(D+E)^{-1}) = 0$.

Theorem 3.18 reproduces several results known in the literature.

Remark 3.19. Let $X \in \mathrm{St}_{n,k}$. We obtain the following special cases for $P_X\colon \mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k}$ by using Corollary 3.16:

1. For $D = I_k$ and $E = 0$ we get the formula
$$P_X(V) = V - XX^\top V + X\,\mathrm{skew}(X^\top V) = (I_n - \tfrac{1}{2}XX^\top)V - \tfrac{1}{2}XV^\top X \tag{3.37}$$
that can be found for example in [2, Ex. 3.6.2] or [10, Eq. (2.4)]
2. More generally, for $D = 2I_k$ and $E = \nu I_n$ with $\nu \in \mathbb{R}\setminus\{-2\}$ one obtains
$$P_X(V) = V - XX^\top V + X\,\mathrm{skew}(X^\top V) = V - \tfrac{1}{2}XX^\top V - \tfrac{1}{2}XV^\top X \tag{3.38}$$
reproducing the orthogonal projection from [14, Prop. 2].

Next we determine an orthonormal basis of $(T_{I_{n,k}}\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ which allows for computing the signature of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$, as well.

Remark 3.20. We define the subsets $B_1, B_2 \subseteq \mathbb{R}^{n\times k}$ such that $B = B_1 \cup B_2$ is an orthonormal basis of $(T_{I_{n,k}}\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{I_{n,k}})$. To this end, let $E_{ij} \in \mathbb{R}^{n\times k}$ denote the matrix whose entries fulfill $(E_{ij})_{f\ell} = \delta_{if}\delta_{j\ell}$ as usual. We set $B_1 = \emptyset$ for $k = 1$ and define

$$B_1 = \Big\{\tfrac{1}{\sqrt{|s_{ij}|}}(E_{ij} - E_{ji}) \,\Big|\, s_{ij} = D_{ii} + E_{ii} + D_{jj} + E_{jj},\ 1 \le i < j \le k\Big\}, \quad 2 \le k \le n. \tag{3.39}$$

Moreover, we set

$$B_2 = \Big\{\tfrac{1}{\sqrt{|D_{jj}|}}E_{ij} \,\Big|\, k+1 \le i \le n, \quad 1 \le j \le k\Big\}, \quad 1 \le k < n. \tag{3.40}$$

and $B_2 = \emptyset$ for $k = n$. A straightforward calculation shows $\langle V, W\rangle^{D,E}_{I_{n,k}} = 0$ for all $V, W \in B$ with $V \neq W$. Moreover, for $V = W \in B$ one obtains

$$\Big\langle \tfrac{1}{\sqrt{|s_{ij}|}}(E_{ij} - E_{ji}), \tfrac{1}{\sqrt{|s_{ij}|}}(E_{ij} - E_{ji})\Big\rangle^{D,E}_{I_{n,k}} = \tfrac{s_{ij}}{|s_{ij}|} = \pm 1, \quad 1 \le i < j \le k \tag{3.41}$$

and

$$\Big\langle \tfrac{1}{\sqrt{|D_{jj}|}}E_{ij}, \tfrac{1}{\sqrt{|D_{jj}|}}E_{ij}\Big\rangle^{D,E}_{I_{n,k}} = \tfrac{D_{jj}}{|D_{jj}|} = \pm 1, \quad k+1 \le i \le n,\ 1 \le j \le k. \tag{3.42}$$

Hence B is in fact an orthonormal basis. Thus we may compute the signature of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$. The number of negative signs associated with $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$, named index in [17, Chap. 2, Def. 18], is given by

$$s = \sharp\{(i,j) \mid 1 \le i < j \le k \text{ and } s_{ij} < 0\} + (n-k)\cdot\sharp\{j \mid 1 \le j \le k \text{ and } D_{jj} < 0\}, \tag{3.43}$$

where $\sharp S$ denotes the number of elements in the finite set S.

3.4. Stiefel manifolds as reductive homogeneous spaces

Before we continue with the extrinsic approach, we briefly discuss the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ on $\mathrm{St}_{n,k}$ viewed as a pseudo-Riemannian reductive homogeneous $SO(n)$-space. This point of view allows for relating $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ to the metrics investigated in [15]. For general properties of reductive homogeneous space we refer to [17, Chapter 11] as well as [18, Section 23.4].

Throughout this subsection, we assume $1 \leq k \leq n-1$ and $n \geq 3$. Then the Killing form on $SO(n)$ given by

$$\langle \xi, \eta \rangle = (n-2)\operatorname{tr}(\xi\eta), \quad \xi, \eta \in \mathfrak{so}(n)$$

is negative definite, see e.g. [18, Sec. 21.6]. In addition, $\mathrm{St}_{n,k}$ is diffeomorphic to the reductive homogeneous space $SO(n)/SO(n-k)$, where $SO(n-k)$ is realized as a closed subgroup of $SO(n)$ via

$$SO(n-k) \cong \left\{ \begin{bmatrix} I_k & 0 \\ 0 & R \end{bmatrix} \middle| R \in SO(n-k) \right\} \subseteq SO(n)$$

and a reductive split is given by $\mathfrak{so}(n) = \mathfrak{h} \oplus \mathfrak{m}$, where

$$\mathfrak{h} = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & \xi_{22} \end{bmatrix} \middle| \xi_{22} \in \mathfrak{so}(n-k) \right\} \quad \text{and} \quad \mathfrak{m} = \left\{ \begin{bmatrix} \xi_{11} & -\xi_{21}^{\top} \\ \xi_{21} & 0 \end{bmatrix} \middle| \xi_{11} \in \mathfrak{so}(k),\ \xi_{21} \in \mathbb{R}^{(n-k)\times k} \right\},$$

see e.g. [18, Sec. 23.5]. In particular, since $SO(n) \times \mathrm{St}_{n,k} \ni (R, X) \mapsto RX \in \mathrm{St}_{n,k}$ is a transitive $SO(n)$-left action whose stabilizer subgroup of $I_{n,k}$ coincides with $SO(n-k) \subseteq SO(n)$, the map

$$\mathrm{pr}\colon SO(n) \to \mathrm{St}_{n,k} \cong SO(n)/SO(n-k), \quad R \mapsto RI_{n,k} \tag{3.44}$$

is a surjective submersion which induces a $SO(n)$-equivariant diffeomorphism

$$\check{\mathrm{pr}}\colon SO(n)/SO(n-k) \to \mathrm{St}_{n,k}, \quad R \cdot SO(n-k) \mapsto RI_{n,k}. \tag{3.45}$$

Here $R \cdot SO(n-k) \in SO(n)/SO(n-k)$ denotes the coset defined by $R \in SO(n)$. We refer to [20, Thm. 6.4] and [21, Thm. 21.18] for more details on diffeomorphisms associated with transitive actions.

In the sequel, we construct a scalar product

$$\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}\colon \mathfrak{so}(n) \times \mathfrak{so}(n) \to \mathbb{R}$$

on $\mathfrak{so}(n)$ which induces a left-invariant metric on $SO(n)$ such that (3.44) becomes a pseudo-Riemannian submersion. In addition, equipping $SO(n)/SO(n-k)$ with this submersion metric turns (3.45) into a $SO(n)$-equivariant isometry to $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$.

Throughout this section we denote by $D, E \in \mathbb{R}^{k\times k}$ diagonal matrices such that D and $D+E$ are both invertible, see also Notation 3.13.

Lemma 3.21. *Let $E_{ij} \in \mathbb{R}^{n\times n}$ be the matrix whose entries fulfill $(E_{ij})_{f\ell} = \delta_{if}\delta_{j\ell}$ and let $F = D+E \in \mathbb{R}^{k\times k}$. Then*

$$A\colon \mathfrak{so}(n) \to \mathfrak{so}(n), \quad \xi = \begin{bmatrix} \xi_{11} & -\xi_{21}^{\top} \\ \xi_{21} & \xi_{22} \end{bmatrix} \mapsto A(\xi) = \begin{bmatrix} \operatorname{skew}(\xi_{11}(D+E)) & -\frac{1}{2}D\xi_{21}^{\top} \\ \frac{1}{2}\xi_{21}D & \xi_{22} \end{bmatrix} \tag{3.46}$$

is linear, where $\xi_{11} \in \mathfrak{so}(k)$, $\xi_{22} \in \mathfrak{so}(n-k)$ and $\xi_{21} \in \mathbb{R}^{(n-k)\times k}$. Moreover, evaluating A at the basis $\{(E_{ij} - E_{ji}) \mid 1 \leq i < j \leq k\}$ of $\mathfrak{so}(n)$ yields

$$A(E_{ij} - E_{ji}) = \begin{cases} \frac{F_{ii}+F_{jj}}{2}(E_{ij} - E_{ji}) & \text{if } 1 \leq i < j \leq k, \\ \frac{1}{2}D_{jj}(E_{ij} - E_{ji}) & \text{if } k+1 \leq i \leq n,\ 1 \leq j \leq k, \\ E_{ij} - E_{ji} & \text{if } k+1 \leq i < j \leq n. \end{cases} \tag{3.47}$$

In particular, $A\colon \mathfrak{so}(n) \to \mathfrak{so}(n)$ as well as its restriction $A\big|_{\mathfrak{m}}\colon \mathfrak{m} \to \mathfrak{m}$ are linear isomorphisms.

Proof. Clearly, A is linear. We show (3.47) by using the definition of A in (3.46). First we consider the case $1 \leq i < j \leq k$. Then $E_{ij} - E_{ji}$ is mapped by A to

$$A(E_{ij} - E_{ji}) = \begin{bmatrix} \operatorname{skew}\big((\widehat{E}_{ij} - \widehat{E}_{ji})F\big) & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \operatorname{skew}\big(\widehat{E}_{ij}F_{jj} - \widehat{E}_{ji}F_{ii}\big) & 0 \\ 0 & 0 \end{bmatrix} = \tfrac{F_{ii}+F_{jj}}{2}(E_{ij} - E_{ji}),$$

with $\widehat{E}_{ij} \in \mathbb{R}^{k\times k}$ defined by $(\widehat{E}_{ij})_{f\ell} = \delta_{if}\delta_{j\ell}$. Next assume $k+1 \leq i \leq n$ and $1 \leq j \leq k$. One obtains

$$A(E_{ij} - E_{ji}) = \tfrac{1}{2}D_{jj}(E_{ij} - E_{ji}).$$

The equality $A(E_{ij} - E_{ji}) = E_{ij} - E_{ji}$ for $k+1 \leq i < j \leq n$ is obvious.

Lemma 3.22. *Define*

$$\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)} \colon \mathfrak{so}(n) \times \mathfrak{so}(n) \to \mathbb{R}, \quad (\xi,\eta) \mapsto \langle\xi,\eta\rangle^{\mathrm{red}(D,E)} = \operatorname{tr}(\xi^\top A(\eta)), \tag{3.48}$$

where $A\colon \mathfrak{so}(n) \to \mathfrak{so}(n)$ is the linear map from Lemma 3.21. Then the following assertions are fulfilled:

1. *$\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ is a scalar product on $\mathfrak{so}(n)$.*
2. *The restriction of $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ to $\mathfrak{m}$ defines a scalar product $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}\colon \mathfrak{m}\times\mathfrak{m} \to \mathbb{R}$ on $\mathfrak{m}$.*
3. *Writing*
$$\xi = \begin{bmatrix} \xi_{11} & -\xi_{21}^\top \\ \xi_{21} & \xi_{22} \end{bmatrix} \in \mathfrak{so}(n) \quad \textit{and} \quad \eta = \begin{bmatrix} \eta_{11} & -\eta_{21}^\top \\ \eta_{21} & \eta_{22} \end{bmatrix} \in \mathfrak{so}(n) \tag{3.49}$$
with $\xi_{11}, \eta_{11} \in \mathfrak{so}(k)$ and $\xi_{21}, \eta_{21} \in \mathbb{R}^{(n-k)\times k}$ yields
$$\langle\xi,\eta\rangle^{\mathrm{red}(D,E)} = \operatorname{tr}(\xi_{11}^\top \operatorname{skew}(\eta_{11}(D+E))) + \operatorname{tr}(\xi_{21}^\top\eta_{21}D) + \operatorname{tr}(\xi_{22}^\top\eta_{22}). \tag{3.50}$$
4. *$\langle\cdot,\cdot\rangle^{red(D,E)}$ is $\operatorname{Ad}(SO(k))$-invariant.*
5. *Declaring $T_e\operatorname{pr}\colon T_eSO(n) \to T_{\operatorname{pr}(e)}(SO(n)/SO(n-k))$ as an isometry defines a $SO(n)$-invariant pseudo-Riemannian metric on $SO(n)/SO(n-k)$ such that $\operatorname{pr}\colon SO(n) \to SO(n)/SO(n-k)$ is a pseudo-Riemannian submersion, where $SO(n)$ is equipped with the left-invariant metric defined by $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$.*

Proof. Obviously, $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ is a bilinear form. Using the notation introduced in (3.49) one calculates

$$\begin{aligned}\langle\xi,\eta\rangle^{\mathrm{red}(D,E)} &= \operatorname{tr}\left(\begin{bmatrix} \xi_{11} & -\xi_{21}^\top \\ \xi_{21} & \xi_{22} \end{bmatrix}^\top \begin{bmatrix} \operatorname{skew}(\eta_{11}(D+E)) & -\frac{1}{2}D\eta_{21}^\top \\ \frac{1}{2}\eta_{21}D & \eta_{22} \end{bmatrix}\right) \\ &= \operatorname{tr}(\xi_{11}^\top(\operatorname{skew}(\eta_{11}(D+E))) + \operatorname{tr}(\xi_{21}^\top\eta_{21}D) + \operatorname{tr}(\xi_{22}^\top\eta_{22}) \\ &= \langle\eta,\xi\rangle^{\mathrm{red}(D,E)}.\end{aligned} \tag{3.51}$$

Hence $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ is symmetric. Claim 3 follows by (3.51), as well. Moreover, $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ is a scalar product since $A\colon \mathfrak{so}(n) \to \mathfrak{so}(n)$ is a linear isomorphism by Lemma 3.21 showing Claim 1. Claim 2 follows since $A\big|_m\colon \mathfrak{m} \to \mathfrak{m}$ is an isomorphism, too.

In order to show the $\operatorname{Ad}(SO(n-k))$-invariance we calculate

$$\operatorname{Ad}_g(\xi) = \begin{bmatrix} I_k & 0 \\ 0 & R \end{bmatrix}\begin{bmatrix} \xi_{11} & -\xi_{21} \\ \xi_{21} & \xi_{22} \end{bmatrix}\begin{bmatrix} I_k & 0 \\ 0 & R \end{bmatrix}^\top = \begin{bmatrix} \xi_{11} & -(R\xi_{21})^\top \\ R\xi_{21} & R\xi_{22}R^\top \end{bmatrix}$$

for $\xi = \begin{bmatrix} x_{11} & -\xi_{21}^\top \\ \xi_{21} & \xi_{22} \end{bmatrix} \in \mathfrak{so}(n)$ and $g = \begin{bmatrix} I_k & 0 \\ 0 & R \end{bmatrix} \in SO(n-k) \subseteq SO(n)$ implying

$$\begin{aligned} \langle \mathrm{Ad}_g(\xi), \mathrm{Ad}_g(\eta) \rangle^{\mathrm{red}(D,E)} &= \mathrm{tr}\,(\xi_{11}^\top(\,\mathrm{skew}(\eta_{11}(D+E)))) + \mathrm{tr}\,(((R\xi_{21})^\top R\eta_{21} D) + \mathrm{tr}\,((R\xi_{22}R^\top)^\top R\eta_{22}R^\top) \\ &= \langle \xi, \eta \rangle^{\mathrm{red}(D,E)} \end{aligned}$$

as desired.

It remains to prove Claim 5. By (3.50) the vector spaces $\mathfrak{m} \subseteq \mathfrak{so}(n)$ and $\mathfrak{h} \subseteq \mathfrak{so}(n)$ are orthogonal complements with respect to $\langle \cdot, \cdot \rangle^{\mathrm{red(D,E)}}$. Moreover, by exploiting the $\mathrm{Ad}\,(SO(n-k))$-invariance of $\langle \cdot, \cdot \rangle^{\mathrm{red}(D,E)}$, this claim follows by [18, Prop. 23.23] which extends to the pseudo-Riemannian setting because its proof only relies on the non-degeneracy of the metric.

After this preparation, we are in the position to show that $\langle \cdot, \cdot \rangle^{\mathrm{red(D,E)}}$ has indeed the desired property. To this end, the tangent map of (3.44) at $I_n \in SO(n)$ is determined as

$$T_{I_n} \mathrm{pr} \colon \mathfrak{so}(n) \to T_{I_{n,k}} \mathrm{St}_{n,k}, \quad \xi \mapsto \xi I_{n,k}. \tag{3.52}$$

Proposition 3.23. *Let $SO(n)/SO(n-k)$ be equipped with the pseudo-Riemannian metric constructed by means of the scalar product $\langle \cdot, \cdot \rangle^{\mathrm{red}(D,E)} \colon \mathfrak{m} \times \mathfrak{m} \to \mathbb{R}$ and let $\mathrm{St}_{n,k}$ be endowed with the metric $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$.*

1. *The restriction of* (3.52) *to* $\mathfrak{m}$*, i.e. the linear map*

$$T_{I_n} \mathrm{pr}\big|_{\mathfrak{m}} \colon \mathfrak{m} \to T_{I_{n,k}} \mathrm{St}_{n,k}, \quad \xi \mapsto \xi I_{n,k} \tag{3.53}$$

is an isometry, where $T_{I_{n,k}} \mathrm{St}_{n,k}$ is equipped with the scalar product $\langle \cdot, \cdot \rangle^{D,E}_{I_{n,k}}$.

2. *The $SO(n)$-equivariant diffeomorphism* (3.45) *is an isometry.*

Proof. We write $\xi, \eta \in \mathfrak{m}$ as

$$\xi = \begin{bmatrix} \xi_{11} & -\xi_{21}^\top \\ \xi_{21} & 0 \end{bmatrix} \quad \text{and} \quad \eta = \begin{bmatrix} \eta_{11} & -\eta_{21}^\top \\ \eta_{21} & 0 \end{bmatrix}$$

with $\xi_{11}, \eta_{11} \in \mathfrak{so}(k)$ as well as $\xi_{21}, \eta_{21} \in \mathbb{R}^{(n-k)\times k}$ and compute

$$\begin{aligned} \Big\langle T_{I_n} \mathrm{pr}\, \xi, T_{I_n} \mathrm{pr}\, \eta \Big\rangle^{D,E}_{\mathrm{pr}(I_n)} &= \Big\langle \begin{bmatrix} \xi_{11} \\ \xi_{21} \end{bmatrix}, \begin{bmatrix} \eta_{11} \\ \eta_{21} \end{bmatrix} \Big\rangle^{D,E}_{I_{n,k}} \\ &= \mathrm{tr}\,(\xi_{11}^\top(\,\mathrm{skew}(\eta_{11}(D+E)))) + \mathrm{tr}\,(\xi_{21}^\top \eta_{21} D) \\ &= \langle \xi, \eta \rangle^{\mathrm{red}(D,E)}, \end{aligned}$$

where the last equality holds by Lemma 3.22, Claim 3. It remains to show Claim 2. Since the metric on $SO(n)/SO(n-k)$ induced by $\langle \cdot, \cdot \rangle^{\mathrm{red}(D,E)}$ and the metric $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$ on $\mathrm{St}_{n,k}$ are both $SO(n)$-invariant, the map $\check{\mathrm{pr}} \colon SO(n)/SO(n-k) \to \mathrm{St}_{n,k}$ is an isometry by Claim 1 due to its $SO(n)$-equivariance.

Proposition 3.23 allows for relating the metric $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$ to the metrics on $\mathrm{St}_{n,k}$ defined in [15, Eq. (3.2)]. In order to compare these metrics we introduce some notation following [15]. We choose $k_1, \ldots, k_s \in \mathbb{N}$ with

$$k_1 + \cdots + k_s = k \quad \text{and} \quad k_i \geq 2 \text{ for all } i \in \{1, \ldots, s\}$$

and write

$$D = \mathrm{diag}(\widetilde{D}_{11} I_{k_1}, \ldots, \widetilde{D}_{ss} I_{k_s}) \quad \text{and} \quad E = \mathrm{diag}(\widetilde{E}_{11} I_{k_1}, \ldots, \widetilde{E}_{ss} I_{k_s}), \tag{3.54}$$

where $\widetilde{D}_{ii}, \widetilde{E}_{ii} \in \mathbb{R}$. Using the notation from [15] with $\mathfrak{p} = \mathfrak{m}$ and $t = 1$ we rewrite $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ as

$$\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)} = \sum_{i=1}^{s} \tfrac{\widetilde{D}_{ii}+\widetilde{E}_{ii}+\widetilde{D}_{ii}+\widetilde{E}_{ii}}{2(n-2)}\langle\cdot,\cdot\rangle_{\mathfrak{p}_i} + \sum_{1\le i<j\le s} \tfrac{\widetilde{D}_{ii}+\widetilde{E}_{ii}+\widetilde{D}_{jj}+\widetilde{E}_{jj}}{2(n-2)}\langle\cdot,\cdot\rangle_{\mathfrak{p}_{ij}} + \sum_{i=1}^{s} \tfrac{\widetilde{D}_{ii}}{2(n-2)}\langle\cdot,\cdot\rangle_{\mathfrak{p}_{i,s+1}}.$$

Here $\langle\cdot,\cdot\rangle_{\mathfrak{p}_{ij}} = (n-2)\,\mathrm{tr}\left((\cdot)^\top(\cdot)\right)\big|_{\mathfrak{p}_{ij}\times\mathfrak{p}_{ij}}$ denotes the Killing form on $\mathfrak{so}(n)$ scaled by -1 restricted to $\mathfrak{p}_{ij}$. Hence $\langle\cdot,\cdot\rangle^{\mathrm{red}(D,E)}$ coincides with the inner product defined in [15, Eq. (3.2)], where $x_i = x_{ii}$ and

$$x_{ij} = \begin{cases} \frac{\widetilde{D}_{ii}+\widetilde{E}_{ii}+\widetilde{D}_{jj}+\widetilde{E}_{jj}}{2(n-2)} & \text{if } 1 \le i \le j \le s \\ \frac{\widetilde{D}_{ii}}{2(n-2)} & \text{if } j = s+1 \text{ and } 1 \le i \le s, \end{cases}$$

provided that D and E are defined as in (3.54) as well as

$$\widetilde{D}_{ii} > 0, \quad i \in \{1,\dots,s\} \quad \text{and} \quad \widetilde{D}_{ii} + \widetilde{E}_{ii} + \widetilde{D}_{jj} + \widetilde{E}_{jj} > 0, \quad i,j \in \{1,\dots,s\}$$

holds. This can be seen by observing that for $\xi \in \mathfrak{m} = \mathfrak{p}$ the unique decomposition of ξ into sums of $\xi_{ij} \in \mathfrak{p}_{ij}$ can be rewritten in terms of block matrices as

$$\xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \cdots\cdots\cdots & \xi_{1,s} & \xi_{1,s+1} \\ \xi_{21} & \xi_{22} & \cdots\cdots\cdots & \xi_{2,s} & \xi_{2,s+1} \\ \vdots & & \ddots & \vdots & \vdots \\ \xi_{s,1} & \cdots\cdots\cdots & & \xi_{s,s} & \xi_{s,s+1} \\ \xi_{s+1,1} & \cdots & \xi_{s+1,s-1} & \xi_{s+1,s} & 0 \end{bmatrix}.$$

Finally, we point out that the Einstein metrics discussed in [15, Sec. 6] yield the following equations for D and E

$$x = \tfrac{\widetilde{D}_{ii}+\widetilde{E}_{ii}}{n-2} \text{ for } 1 \le i \le s \quad y = \tfrac{\widetilde{D}_{ii}+\widetilde{E}_{ii}+\widetilde{D}_{jj}+\widetilde{E}_{jj}}{2(n-2)} \text{ for } 1 \le i < j \le s, \quad z = \tfrac{\widetilde{D}_{ii}}{2(n-2)} \text{ for } 1 \le i \le s,$$

where x, y, z denote the parameters of the metric from [15, Eq. (6.2)]. Thus

$$D = 2(n-2)zI_k \implies D + E = 2z(n-2)I_k + E = (n-2)xI_k \implies E = (n-2)(x-2z)I_k$$

and therefore $y = \frac{(n-2)(2z+(x-2z)+2z+(x-2z))}{2(n-2)} = x$ holds for $x, z \in \mathbb{R}$. In particular, the metrics on $\mathrm{St}_{n,k}$ defined by $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ contain only the two $SO(n) \times SO(k)$-invariant Einstein metrics from [15], the so-called Jensen metrics. However, they do not contain the "new" Einstein metrics from that paper.

Remark 3.24. Although the "new" Einstein metrics form [15] are not contained in the family of metrics on $\mathrm{St}_{n,k}$ defined by $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$, we are not able to rule out that the family $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ includes Einstein metrics different from the Jensen metrics. However, searching for Einstein metrics in $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is out of the scope of this text.

4. Sprays and geodesic equations

The goal of this section is to derive an explicit expression for the spray $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$ associated with the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$. An expression for S yields an expression for the geodesic equation with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ as an explicit second order ODE, as well.

First we recall the definition of a metric spray, also known as spray associated with a metric, from [22, Chap. 8, §4] whose existence and uniqueness is proven in [22, Chap. 8, Thm. 4.2]. For general properties of sprays we refer to [22, Chap. 4, §3-4]. Moreover, a discussion of the relation of sprays to torsion-free covariant derivatives can be found in [22, Chap. 8 §2].

Definition 4.1. Let $(M, \langle\cdot,\cdot\rangle)$ be a pseudo-Riemannian manifold. The metric spray $S \in \Gamma^\infty(T(TM))$ is the unique spray which is associated with the Levi-Civita covariant derivative defined by the pseudo-Riemannian metric $\langle\cdot,\cdot\rangle$.

An expression of a metric spray in local coordinates is given in (4.2) below. Next we discuss the relation of metric sprays to Lagrangian mechanics.

Let $(M, \langle\cdot,\cdot\rangle)$ be pseudo-Riemannian and let $\omega_0 \in \Gamma^\infty(\Lambda^2 T^*(T^*M))$ denote the canonical symplectic form on T^*M. It is given by

$$\omega_0 = -\,\mathrm{d}\,\theta_0$$

with $\theta_0 \in \Gamma^\infty(T^*(T^*M))$ being the canonical 1-form on T^*M. We refer to [16, Sec. 6.2] for the definition of ω_0 and θ_0. Consider the Lagrange function

$$L\colon TM \to \mathbb{R}, \quad v_x \mapsto L(v_x) = \tfrac{1}{2}\langle v_x, v_x\rangle_x.$$

Let $\mathrm{F}L\colon TM \to T^*M$ denote the fiber derivative of L defined by

$$((\mathrm{F}L)(v_x))(w_x) = \tfrac{\mathrm{d}}{\mathrm{d}t}L(v_x + tw_x)\Big|_{t=0}, \quad x \in M, \quad v_x, w_x \in T_xM,$$

see e.g. [16, Eq. (7.2.1)]. The pullback

$$\omega_L = (\mathrm{F}L)^*\omega_0$$

is a closed 2-from on TM, the so-called Lagrangian 2-form, see [16, Sec. 7.2]. In addition, ω_L is non-degenerated, i.e. symplectic, since $\mathrm{F}L\colon TM \to T^*M$ is a diffeomorphism due to

$$\mathrm{F}L\colon TM \to T^*M, \quad v_x \mapsto \mathrm{F}L(v_x) = \langle v_x, \cdot\rangle \tag{4.1}$$

by [16, Eq. (7.5.3)]. Moreover, the energy

$$E_L\colon TM \to \mathbb{R}, \quad v_x \mapsto ((\mathrm{F}L)(v_x))(v_x) - L(v_x)$$

associated with L fulfills $E_L = L$, see e.g. [16, Sec. 7.3]. Let $X_{E_L} \in \Gamma^\infty(T(TM))$ denote the Lagrangian vector field and write $i_{X_{E_L}}\omega_L$ for the insertion of X_{E_L} into the first argument of ω_L as usual. Then X_{E_L} is uniquely determined by

$$i_{X_{E_L}}\omega_L = \mathrm{d}\,E_L \quad \iff \quad \omega_L(X_{E_L}, V) = \mathrm{d}\,E_L(V) \text{ for all } V \in \Gamma^\infty(T(TM)).$$

according to [16, Sec. 7.3]. Moreover, the Lagrangian vector field X_{E_L} coincides with the spray associated with the metric $\langle\cdot,\cdot\rangle$, see e.g. [16, Sec. 7.5]. It is exactly the so-called canonical spray from [22, Chap. 7, §7] which coincides with the metric spray, see [22, Chap. 8, Thm. 4.2]. Finally, we mention a local expression for sprays, see e.g. [22, Chap. 8, §4]. A metric spray $S : TM \to T(TM)$ can be represented in a chart $(TU, (x, v))$ of TM induced by a chart (U, x) of M by

$$S(x, v) = (x, v, v, -\Gamma_x(v, v)). \tag{4.2}$$

Here Γ_x denotes the quadratic map defined by $(\Gamma_x(v,v))^k = \Gamma_{ij}^k(x)v^i v^j$ using Einstein summation convention, where Γ_{ij}^k are the Christoffel symbols of the Levi-Civita covariant derivative with respect to the chart (U, x). In order to apply these general results to our particular situation, we introduce some notation.

Notation 4.2. Throughout this section $U \subseteq \mathbb{R}^{n\times k}$ denotes an open subset of $\mathbb{R}^{n\times k}$ with the property from Lemma 3.8. Moreover, we denote by $\widetilde{L}$ the Lagrange function

$$\widetilde{L}\colon TU \to \mathbb{R}, \quad (X, V) \mapsto \widetilde{L}(X, V) = \tfrac{1}{2}\langle V, V\rangle_X^{D,E}, \tag{4.3}$$

where we identify $TU \cong U \times \mathbb{R}^{n\times k}$ as usual.

We use the formula for $\omega_0 \in \Gamma^\infty(\Lambda^2 T^*(T^*U))$ on T^*U given in the next remark.

Remark 4.3. The canonical symplectic form $\omega_0 \in \Gamma^\infty(\Lambda^2 T^*(T^*U))$ on T^*U is given by

$$\omega_0\big|_{(X,V)}((X, V, Y, Z), (X, V, \widetilde{Y}, \widetilde{Z})) = \operatorname{tr}(Y^\top \widetilde{Z}) - \operatorname{tr}(\widetilde{Y}^\top Z), \tag{4.4}$$

for $(X, V, Y, Z), (X, V, \widetilde{Y}, \widetilde{Z}) \in T(T^*U)$ identifying $T(T^*U) \cong U \times (\mathbb{R}^{n\times k})^* \times \mathbb{R}^{n\times k} \times (\mathbb{R}^{n\times k})^*$ as well as $\mathbb{R}^{n\times k} \cong (\mathbb{R}^{n\times k})^*$ via $V \mapsto \operatorname{tr}(V^\top(\cdot))$. Indeed, Equation (4.4) follows by the local formula for the canonical symplectic form ω_0 on T^*U, see e.g. [16, Sec. 6.2], applied to the gobal chart $(U, \mathrm{id}_U) = (U, X_{ij})$.

4.1. Lagrangian 2-Form

We now calculate the Lagrangian 2-from $\omega_{\widetilde{L}} = (\mathrm{F}\widetilde{L})^*\omega_0$. To this end, we first determine the fiber derivative $\mathrm{F}\widetilde{L}\colon TU \to T^*U$ and its tangent map.

Lemma 4.4. *For $(X, V) \in TU$ the fiber derivative $\mathrm{F}\widetilde{L}\colon TU \to T^*U$ of $\widetilde{L}$ is given by*

$$\mathrm{F}\widetilde{L}(X, V) = (X, \operatorname{tr}((VD + XX^\top VE)^\top(\cdot))). \tag{4.5}$$

Proof. Let $(X, V), (X, W) \in TU$. We have $(\mathrm{F}\widetilde{L}(X, V))(X, W) = \langle V, W\rangle_X^{D,E}$ by the Definition of $\widetilde{L}$ and (4.1). Using the definition of $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E}$ and exploiting properties of the trace we obtain

$$(\mathrm{F}\widetilde{L}(X, V))(X, W) = \operatorname{tr}(V^\top WD) + \operatorname{tr}(V^\top XX^\top WE) = \operatorname{tr}((VD + XX^\top VE)^\top W)$$

as desired.

Lemma 4.5. *The tangent map $T(\mathrm{F}\widetilde{L})\colon T(TU) \to T(T^*U)$ is given by*

$$(T(\mathrm{F}\widetilde{L}))(X, V, Y, Z) = (\mathrm{F}\widetilde{L}(X, V), Y, \operatorname{tr}((ZD + YX^\top VE + XY^\top VE + XX^\top ZE)^\top(\cdot)))$$

for $(X, V, Y, Z) \in T(TU) \cong U \times (\mathbb{R}^{n\times k})^3$, where we identify $T(T^*U) \cong U \times (\mathbb{R}^{n\times k})^* \times \mathbb{R}^{n\times k} \times (\mathbb{R}^{n\times k})^*$.

Proof. Let $(X, V, Y, Z) \in T(TU)$. The smooth curve $\gamma\colon (-\epsilon, \epsilon) \ni t \mapsto (X + tY, V + tZ) \in TU$, for $\epsilon > 0$ sufficiently small, fulfills $\gamma(0) = (X, V)$ with $\dot{\gamma}(0) = (Y, Z)$. Then

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{F}\widetilde{L}(\gamma(t))\big|_{t=0} = (Y, \operatorname{tr}((ZD + YX^\top VE + XY^\top VE + XX^\top ZE)^\top(\cdot))).$$

This yields the desired result.

Lemma 4.6. *The Lagrangian* 2*-form* $\omega_{\widetilde{L}} = (\mathrm{F}\widetilde{L})^*\omega_0 \in \Gamma^\infty(\Lambda^2 T^*(TU))$ *is given by*

$$\begin{aligned}\omega_{\widetilde{L}}\big|_{(X,V)}((X,V,Y,Z),(X,V,\widetilde{Y},\widetilde{Z})) &= \operatorname{tr}\big(Y^\top(\widetilde{Z}D + \widetilde{Y}X^\top VE + X\widetilde{Y}^\top VE + XX^\top\widetilde{Z}E)\big)\\ &\quad - \operatorname{tr}\big(\widetilde{Y}^\top(ZD + YX^\top VE + XY^\top VE + XX^\top ZE)\big)\end{aligned} \tag{4.6}$$

with $(X,V) \in TU \cong U \times \mathbb{R}^{n\times k}$ *and* $(X,V,Y,Z),(X,V,\widetilde{Y},\widetilde{Z}) \in T_{(X,V)}TU$.

Proof. Using the formula for $\omega_0 \in \Gamma^\infty(\Lambda^2 T^*(T^*U))$ from Remark 4.3, a straightforward calculation shows that $\omega_{\widetilde{L}} = (\mathrm{F}\widetilde{L})^*\omega_0$ is given by (4.6). To this end, the formulas from Lemma 4.4 and Lemma 4.5 are plugged into the definition of the pull-back $(\mathrm{F}\widetilde{L})^*\omega_0$.

4.2. Sprays on TU

Next the spray $\widetilde{S} \in \Gamma^\infty(T(TU))$ associated with $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is calculated exploiting $\widetilde{S} = X_{E_{\widetilde{L}}}$, where $X_{E_{\widetilde{L}}}$ is the Lagrangian vector field. A closed form expression for $\widetilde{S}(X,V)$ is obtained for all $(X,V) \in \mathrm{St}_{n,k} \times \mathbb{R}^{n\times k} \subseteq TU$.

Lemma 4.7. *For* $(X,V) \in TU$ *and* $(X,V,Y,Z) \in T_{(X,V)}TU$ *one has*

$$\mathrm{d}\,E_{\widetilde{L}}\big|_{(X,V)}(X,V,Y,Z) = \operatorname{tr}(V^\top ZD) + \operatorname{tr}(Z^\top XX^\top VE) + \operatorname{tr}(V^\top YX^\top VE). \tag{4.7}$$

Proof. Let $(X,V),(Y,Z) \in TU$. We calculate

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}E_{\widetilde{L}}(X+tY,V+tZ)\big|_{t=0} &= \tfrac{1}{2}\Big(\operatorname{tr}(Z^\top VD + V^\top ZD)\\ &\quad + \operatorname{tr}(Z^\top XX^\top VE + V^\top YX^\top VE + V^\top XY^\top VE + V^\top XX^\top ZE)\Big).\end{aligned}$$

Using properties of the trace yields the desired result.

Next we consider a linear matrix equation of a certain form. We need to solve this equation for computing the metric spray on TU, see Proposition 4.9. Moreover, one encounters this equation in the proof of Proposition 5.2 on pseudo-Riemannian gradients below.

Lemma 4.8. *Let* $D,E \in \mathbb{R}^{k\times k}_{\mathrm{diag}}$ *such that* D *and* $D+E$ *are both invertible and let* $W \in \mathbb{R}^{n\times k}$. *Moreover, let* $U \subseteq \mathbb{R}^{n\times k}$ *be open with the property from Lemma 3.8. Then for* $X \in U$ *the linear equation*

$$\widetilde{\Gamma}D + XX^\top\widetilde{\Gamma}E = W \tag{4.8}$$

has a unique solution in terms of $\widetilde{\Gamma}$. *Moreover, for* $X \in \mathrm{St}_{n,k}$, *it is explicitly given by*

$$\widetilde{\Gamma} = (W - XX^\top W(D+E)^{-1}E)D^{-1}. \tag{4.9}$$

Proof. For each $X \in U$ the linear map $\phi\colon \mathbb{R}^{n\times k} \ni \widetilde{\Gamma} \mapsto \widetilde{\Gamma}D + XX^\top\widetilde{\Gamma}E \in \mathbb{R}^{n\times k}$ is an isomorphism since the bilinear form

$$\mathbb{R}^{n\times k}\times\mathbb{R}^{n\times k} \to \mathbb{R}, \quad (Y,Z) \mapsto \operatorname{tr}(V^\top\phi(W)) = \langle V,W\rangle^{D,E}_X$$

is non-degenerated by assumption. Hence (4.8) admits a unique solution. Now assume $X \in \mathrm{St}_{n,k}$. We briefly explain how (4.9) can be derived. By exploiting $X^\top X = I_k$, Equation (4.8) implies

$$X^\top W = X^\top\widetilde{\Gamma}D + X^\top\widetilde{\Gamma}E = X^\top\widetilde{\Gamma}(D+E) \quad\iff\quad X^\top\widetilde{\Gamma} = X^\top W(D+E)^{-1}.$$

Plugging $X^\top\widetilde{\Gamma} = X^\top W(D+E)^{-1}$ into (4.8) yields

$$\widetilde{\Gamma}D + X(X^\top W(D+E)^{-1})E = W \quad\iff\quad \widetilde{\Gamma} = (W - X(X^\top W(D+E)^{-1})E)D^{-1}.$$

A straightforward calculation shows that $\widetilde{\Gamma}$ is indeed a solution of (4.8).

Proposition 4.9. *The spray $\widetilde{S} \in \Gamma^\infty(T(TU))$ associated with the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is given by*

$$\widetilde{S}(X,V) = (X,V,V,-\widetilde{\Gamma}) = (X,V,V,-\widetilde{\Gamma}_X(V,V)) \tag{4.10}$$

for all $(X,V) \in TU \cong U \times \mathbb{R}^{n\times k}$. Here $\widetilde{\Gamma} = \widetilde{\Gamma}_X(V,V) \in \mathbb{R}^{n\times k}$ depending on $(X,V) \in TU$ is the unique solution of the linear equation

$$\widetilde{\Gamma}D + XX^\top\widetilde{\Gamma}E = VX^\top VE + XV^\top VE - VEV^\top X \tag{4.11}$$

in terms of $\widetilde{\Gamma}$ with fixed $(X,V) \in TU$. Moreover, for $(X,V) \in \mathrm{St}_{n,k} \times \mathbb{R}^{n\times k}$ one has

$$\begin{aligned}\widetilde{\Gamma}_X(V,V) &= (VX^\top VE + XV^\top VE - VEV^\top X)D^{-1}\\ &\quad + (XX^\top VEV^\top X - X(X^\top V)^2E - XV^\top VE)(D+E)^{-1}ED^{-1}.\end{aligned} \tag{4.12}$$

Proof. Using $\widetilde{S} = X_{E_{\widetilde{L}}}$ we compute $\widetilde{S}$ via solving $i_{X_{E_{\widetilde{L}}}}\omega_{\widetilde{L}} = \mathrm{d}\,E_{\widetilde{L}}$ for $X_{E_{\widetilde{L}}}$, i.e. $\widetilde{S} = X_{E_{\widetilde{L}}}$ fulfills

$$\omega_{\widetilde{L}}(X_{E_{\widetilde{L}}}(X,V),(X,V,Y,Z)) = \mathrm{d}\,E_{\widetilde{L}}\big|_{(X,V)}(X,V,Y,Z). \tag{4.13}$$

for all $(X,V,Y,Z) \in T(TU)$. Since $\omega_{\widetilde{L}}$ is non-degenerated, $X_{E_{\widetilde{L}}}$ is uniquely determined by (4.13). The local form of a metric spray, see (4.2), motivates the Ansatz

$$X_{E_{\widetilde{L}}}(X,V) = (X,V,V,-\widetilde{\Gamma}_X(V,V)) = (X,V,V,-\widetilde{\Gamma})$$

with $\widetilde{\Gamma} = \widetilde{\Gamma}_X(V,V) \in \mathbb{R}^{n\times k}$ depending on $(X,V) \in TU$. Inserting $X_{E_{\widetilde{L}}}$ into $\omega_{\widetilde{L}}$ from Lemma 4.6 yields the 1-form

$$\begin{aligned}(i_{X_{E_{\widetilde{L}}}}\omega_{\widetilde{L}})\big|_{(X,V)}(X,V,Y,Z) &= \omega_{\widetilde{L}}\big|_{(X,V)}(X_{E_{\widetilde{L}}}(X,V),(X,V,Y,Z))\\ &= \mathrm{tr}\Big(V^\top\Big(ZD + YX^\top VE + XY^\top VE + XX^\top ZE\Big)\\ &\quad - \mathrm{tr}\Big(Y^\top\Big(-\widetilde{\Gamma}D + VX^\top VE + XV^\top VE - XX^\top\widetilde{\Gamma}E\Big)\Big)\end{aligned} \tag{4.14}$$

with $(X,V) \in TU$ and $(X,V,Y,Z) \in T(TU)$. Using (4.14) and the formula for $\mathrm{d}\,E_{\widetilde{L}}$ from Lemma 4.7, the equation $i_{X_{E_{\widetilde{L}}}}\omega_{\widetilde{L}} = \mathrm{d}\,E_{\widetilde{L}}$ becomes

$$\begin{aligned}&\mathrm{tr}(V^\top ZD) + \mathrm{tr}(Z^\top XX^\top VE) + \mathrm{tr}(V^\top YX^\top VE)\\ &\quad = \mathrm{tr}\Big(V^\top\Big(ZD + YX^\top VE + XY^\top VE + XX^\top ZE\Big)\\ &\qquad - \mathrm{tr}\Big(Y^\top\Big(-\widetilde{\Gamma}D + VX^\top VE + XV^\top VE - XX^\top\widetilde{\Gamma}E\Big)\Big)\end{aligned} \tag{4.15}$$

for all $(X,V,Y,Z) \in TU$. Clearly, Equation (4.15) is equivalent to

$$\mathrm{tr}(Y^\top(VEV^\top X)) = \mathrm{tr}\left(Y^\top(-\widetilde{\Gamma}D + VX^\top VE + XV^\top VE - XX^\top\widetilde{\Gamma}E)\right)$$

for all $Y \in \mathbb{R}^{n\times k}$. This can be equivalently rewritten as

$$\widetilde{\Gamma}D + XX^\top\widetilde{\Gamma}E = VX^\top VE + XV^\top VE - VEV^\top X \tag{4.16}$$

showing the first claim.

We now assume $X \in \mathrm{St}_{n,k}$. Writing $W = VX^\top VE + XV^\top VE - VEV^\top X$ and invoking Lemma 4.8 in order to solve (4.16) for $\widetilde{\Gamma}$ yields

$$\begin{aligned}\widetilde{\Gamma} &= WD^{-1} - XX^\top W(D+E)^{-1}ED^{-1}\\ &= (VX^\top VE + XV^\top VE - VEV^\top X)D^{-1}\\ &\quad + (XX^\top VEV^\top X - X(X^\top V)^2E - XV^\top VE)(D+E)^{-1}ED^{-1}\end{aligned}$$

as desired.

Remark 4.10. Obviously, for $E = 0$, Proposition 4.9 implies $\widetilde{\Gamma}_X(V,V) = 0$ for all $(X,V) \in TU$.

Proposition 4.9 admits a relatively simple expression for $\widetilde{S} \in \Gamma^\infty(T(TU))$ evaluated at $(X,V) \in T\mathrm{St}_{n,k}$ for a subfamily of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$. Since this subfamily will be discussed several times below, it deserves its own notation.

Notation 4.11. We write $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ for the covariant 2-tensor $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ which is obtained by specifying $E = \nu I_k$ with $\nu \in \mathbb{R}$, i.e.

$$\langle V,W\rangle^{D,\nu}_X = \mathrm{tr}\,(V^\top WD) + \nu\,\mathrm{tr}\,(V^\top XX^\top W), \quad X \in \mathbb{R}^{n\times k} \text{ and } V,W \in T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}.$$

Unless indicated otherwise, pull-backs of $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathbb{R}^{n\times k}))$ to submanifolds of $\mathbb{R}^{n\times k}$ are omitted in the notation. Moreover, we assume that D and ν are chosen such that $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)})$ is a pseudo-Riemannian submanifold. In particular, we assume that D and $D + \nu I_k$ are both invertible.

Corollary 4.12. *The spray $\widetilde{S} \in \Gamma^\infty(T(TU))$ on TU associated with $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*U))$ evaluated at $(X,V) \in T\mathrm{St}_{n,k}$ is given by*

$$\widetilde{S}(X,V) = (X,V,V,-\widetilde{\Gamma}_X(V,V)), \tag{4.17}$$

where

$$\widetilde{\Gamma}_X(V,V) = \Big(2\nu VX^\top V + \nu XV^\top V(D(D+\nu I_k)^{-1}) - 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1}\Big)D^{-1}. \tag{4.18}$$

Proof. Let $(X,V) \in T\mathrm{St}_{n,k}$ and write $\widetilde{\Gamma} = \widetilde{\Gamma}_X(V,V)$ for short. Plugging $E = \nu I_k$ into Formula (4.12) from Proposition 4.9 and using $X^\top V = -V^\top X$ we obtain

$$\begin{aligned}\widetilde{\Gamma} &= \Big(VX^\top VE + XV^\top VE - VEV^\top X\Big)D^{-1}\\ &\quad + \Big(XX^\top VEV^\top X - X(X^\top V)^2E - XV^\top VE\Big)(D+E)^{-1}ED^{-1}\\ &= \nu\Big(VX^\top V + XV^\top V - VV^\top X\Big)D^{-1} + \nu^2\Big(XX^\top VV^\top X - X(X^\top V)^2 - XV^\top V\Big)(D+\nu I_k)^{-1}D^{-1}\\ &= \nu\Big(VX^\top V + XV^\top V + VX^\top V\Big)D^{-1} + \nu^2\Big(-X(X^\top V)^2 - X(X^\top V)^2 - XV^\top V\Big)(D+\nu I_k)^{-1}D^{-1}\\ &= \Big(2\nu VX^\top V + XV^\top V(\nu I_k) - 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1} - XV^\top V(\nu^2(D+\nu I_k)^{-1})\Big)D^{-1}\\ &= \Big(2\nu VX^\top V + XV^\top V(\nu I_k - \nu^2(D+\nu I_k)^{-1}) - 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1}\Big)D^{-1}\\ &= \Big(2\nu VX^\top V + \nu XV^\top V(D(D+\nu I_k)^{-1}) - 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1}\Big)D^{-1},\end{aligned}$$

where the last equality holds due to

$$(\nu I_k - \nu^2(D + \nu I_k)^{-1})_{ii} = \nu - \tfrac{\nu^2}{D_{ii}+\nu} = \tfrac{\nu(D_{ii}+\nu)-\nu^2}{D_{ii}+\nu} = \nu(D(D + \nu I_k)^{-1})_{ii}$$

for $i \in \{1, \dots, k\}$.

4.3. Sprays on Stiefel Manifolds

We now determine the spray $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$ associated with the metric $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$. To this end, a result from [16, Prop. 8.4.1] is exploited which is stated for Riemannian manifolds. The proof works for pseudo-Riemannian manifolds, as well, since it only exploits the non-degeneracy of the metric. We reformulate it in the following proposition.

Proposition 4.13. *Let $M \subseteq \widetilde{M}$ be a pseudo-Riemannian submanifold of a pseudo-Riemannian manifold $(\widetilde{M}, \langle \cdot, \cdot \rangle)$ and let $\widetilde{S} \in \Gamma^\infty(T(T\widetilde{M}))$ denote the metric spray on $T\widetilde{M}$. Then the spray $S \in \Gamma^\infty(T(TM))$ on TM associated with the induced pseudo-Riemannian metric is given by*

$$S = TP \circ \widetilde{S}\big|_{TM} \colon TM \to T(TM), \tag{4.19}$$

where $P\colon T\widetilde{M}\big|_M \to TM$ denotes the vector bundle morphism that is defined fiber-wise by the orthogonal projections $P_x\colon T_x\widetilde{M} \to T_xM \subseteq T_x\widetilde{M}$ with respect to $\langle \cdot, \cdot \rangle$, where $x \in M$.

Lemma 4.14. *The tangent map $TP\colon T(\mathrm{St}_{n,k} \times \mathbb{R}^{n\times k}) \to T(T\mathrm{St}_{n,k})$ of*

$$P\colon \mathrm{St}_{n,k} \times \mathbb{R}^{n\times k} \to T\mathrm{St}_{n,k}, \quad (X, V) \mapsto (X, P_X(V)), \tag{4.20}$$

where $P_X(V) = V - XX^\top V + X\pi^{D+E}(X^\top V)$ is the orthogonal projection from Theorem 3.18, is given by

$$TP(X, V, Y, Z) = (X, V, Y, Z - XY^\top V - XX^\top Z + X\pi^{D+E}(Y^\top V + X^\top Z))) \tag{4.21}$$

for all $(X, V, Y, Z) \in T(\mathrm{St}_{n,k} \times \mathbb{R}^{n\times k}) \cong T\mathrm{St}_{n,k} \times (\mathbb{R}^{n\times k})^2$.

Proof. By exploiting $\pi^{D+E}(X^\top V) = X^\top V$ due to $X^\top V = -V^\top X \in \mathfrak{so}(k)$ for $(X, V) \in T\mathrm{St}_{n,k}$ one calculates

$$\begin{aligned} T_{(X,V)}P(Y, Z) &= (Y, Z - YX^\top V - XY^\top V - XX^\top Z + Y\pi^{D+E}(X^\top V) + X\pi^{D+E}(Y^\top V + X^\top Z)) \\ &= (Y, Z - XY^\top V - XX^\top Z + X\pi^{D+E}(Y^\top V + X^\top Z)), \end{aligned}$$

where $(X, V, Y, Z) \in T(\mathrm{St}_{n,k} \times \mathbb{R}^{n\times k})$.

Theorem 4.15. *The spray $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$ associated with $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$ is given by*

$$S(X, V) = (X, V, V, -\widetilde{\Gamma}_X(V, V) - XV^\top V + XX^\top \widetilde{\Gamma}_X(V, V) + X\pi^{D+E}(V^\top V - X^\top \widetilde{\Gamma}_X(V, V))) \tag{4.22}$$

for all $(X, V) \in T\mathrm{St}_{n,k}$. Here $\widetilde{\Gamma}_X(V, V) \in \mathbb{R}^{n\times k}$ depending on $(X, V) \in T\mathrm{St}_{n,k}$ is given by

$$\begin{aligned} \widetilde{\Gamma}_X(V, V) &= (VX^\top VE + XV^\top VE - VEV^\top X)D^{-1} \\ &\quad + (XX^\top VEV^\top X - X(X^\top V)^2E - XV^\top VE)(D + E)^{-1}ED^{-1}. \end{aligned} \tag{4.23}$$

Proof. One can view $\mathrm{St}_{n,k}$ equipped with $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ as a pseudo-Riemannian submanifold of $(U,\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ according to Lemma 3.8. Let $\widetilde{S}\in\Gamma^\infty(T(TU))$ be the metric spray on TU determined in Proposition 4.9. Then $S = TP\circ\widetilde{S}\big|_{T\mathrm{St}_{n,k}}$ holds by Proposition 4.13. Using Lemma 4.14 yields

$$\begin{aligned}S(X,V) &= TP\circ\widetilde{S}\big|_{T\mathrm{St}_{n,k}}(X,V) = TP(X,V,V,-\widetilde{\Gamma}_X(V,V))\\ &= (X,V,V,-\widetilde{\Gamma}_X(V,V) - XV^\top V + XX^\top\widetilde{\Gamma}_X(V,V) + X\pi^{D+E}(V^\top V - X^\top\widetilde{\Gamma}_X(V,V)))\end{aligned}$$

for all $(X,V)\in T\mathrm{St}_{n,k}$ as desired.

Remark 4.16. We often denote the spray on $T\mathrm{St}_{n,k}$ associated with $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ from Theorem 4.15 by

$$S(X,V) = (X,V,V,-\Gamma) = (X,V,V,-\Gamma_X(V,V)),$$

i.e. we write $-\Gamma$ or $-\Gamma_X(V,V)$ for the fourth component of S. For $(X,V)\in T\mathrm{St}_{n,k}$ it is given by

$$-\Gamma_X(V,V) = -\widetilde{\Gamma}_X(V,V) - XV^\top V + XX^\top\widetilde{\Gamma}_X(V,V) + X\pi^{D+E}(V^\top V - X^\top\widetilde{\Gamma}_X(V,V)) \tag{4.24}$$

according to Theorem 4.15, where $\widetilde{\Gamma}_X(V,V)$ is determined by (4.23). Obviously, Equation (4.24) yields a well-defined expression for all $X\in\mathbb{R}^{n\times k}$ and $V\in\mathbb{R}^{n\times k}$ which is quadratic in V. Hence, by polarization, (4.24) can be viewed as the definition of the smooth map

$$\Gamma\colon U\to \mathrm{S}^2((\mathbb{R}^{n\times k})^*)\otimes\mathbb{R}^{n\times k},\quad X\mapsto ((V,W)\mapsto\Gamma_X(V,W)). \tag{4.25}$$

Clearly, Equation (4.25) yields a smooth extension of the fourth component of the metric spray $S\in\Gamma^\infty(T(T\mathrm{St}_{n,k}))$. This extension is used in Proposition 6.5 and Proposition 6.8 below.

Corollary 4.17. *The spray $S\in\Gamma^\infty(T(T\mathrm{St}_{n,k}))$ associated with the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ from Theorem 4.15 has the following properties:*

1. *The spray $S\in\Gamma^\infty(T(T\mathrm{St}_{n,k}))$ is complete.*
2. *The maximal integral curve $\mathbb{R}\ni t\mapsto\Phi^S_t((X_0,V_0)) = (X(t),V(t))\in T\mathrm{St}_{n,k}$ of S through the point $(X_0,V_0)\in T\mathrm{St}_{n,k}$ at $t=0$ fulfills the explicit non-linear first order ODE*

$$\begin{aligned}\dot{X} &= V\\ \dot{V} &= -\widetilde{\Gamma}_X(V,V) - XV^\top V + XX^\top\widetilde{\Gamma}_X(V,V) + X\pi^{D+E}(V^\top V - X^\top\widetilde{\Gamma}_X(V,V)),\end{aligned} \tag{4.26}$$

   with initial condition $(X(0),V(0)) = (X_0,V_0)\in T\mathrm{St}_{n,k}$ writing $X = X(t)$ and $V = V(t)$ for short.
3. *Let $\mathrm{pr}\colon T\mathrm{St}_{n,k}\to\mathrm{St}_{n,k}$ be the canonical projection. The curve $\mathbb{R}\ni t\mapsto \mathrm{pr}\circ\Phi^S_t(X_0,V_0) = X(t)\in\mathrm{St}_{n,k}$ is a geodesic with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ through the point $X(0) = X_0\in\mathrm{St}_{n,k}$ with initial velocity $\dot{X}(0) = V_0\in T_{X_0}\mathrm{St}_{n,k}$.*
4. *The geodesic equation on $\mathrm{St}_{n,k}$ with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is given by the non-linear explicit second order ODE*

$$\ddot{X} = -\widetilde{\Gamma}_X(\dot{X},\dot{X}) - X\dot{X}^\top\dot{X} + XX^\top\widetilde{\Gamma}_X(\dot{X},\dot{X}) + X\pi^{D+E}(\dot{X}^\top\dot{X} - X^\top\widetilde{\Gamma}_X(\dot{X},\dot{X})) \tag{4.27}$$

   with initial conditions $X(0) = X_0\in\mathrm{St}_{n,k}$ and $\dot{X}(0) = \dot{X}_0\in T_{X_0}\mathrm{St}_{n,k}$.

Proof. We first show that S is complete. The transitive $O(n)$-action Ψ acts on $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ by isometries according to Lemma 3.1, i.e. $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ is a compact pseudo-Riemannian homogeneous manifold. Hence completeness follows by [23]. The other statements are well-known consequences of general properties of sprays associated with a metric, see e.g. [16, Sec. 7.5], combined with the explicit formula for $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$ from Theorem 4.15.

The formula for the metric spray S from Theorem 4.15 admits a simplification for $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$.

Corollary 4.18. *For* $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)})$ *the metric spray is given by* $S(X,V) = (X, V, V, -\Gamma_X(V,V))$ *with*

$$-\Gamma_X(V,V) = 2\nu VV^\top XD^{-1} + 2\nu X(X^\top V)^2 D^{-1} - XV^\top V + X\pi^{D+\nu I_k}(V^\top V) \tag{4.28}$$

for $(X,V) \in T\mathrm{St}_{n,k}$. *Moreover, the geodesic equation reads*

$$\ddot{X} = 2\nu \dot{X}\dot{X}^\top XD^{-1} + 2\nu X(X^\top \dot{X})^2 D^{-1} - X\dot{X}^\top \dot{X} + X\pi^{D+\nu I_k}(\dot{X}^\top \dot{X}). \tag{4.29}$$

Proof. Let $(X,V) \in T\mathrm{St}_{n,k}$. Using the formula for $\widetilde{\Gamma}_X(V,V)$ from Corollary 4.12 we calculate

$$\begin{aligned}
X^\top\widetilde{\Gamma}_X(V,V) &= X^\top\Big(2\nu VX^\top V + \nu XV^\top V(D(D+\nu I_k)^{-1}) - 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1}\Big)D^{-1} \\
&= 2\nu X^\top VX^\top VD^{-1} + \nu V^\top V((D+\nu I_k)^{-1}D)D^{-1} - 2\nu^2(X^\top V)^2(D+\nu I_k)^{-1}D^{-1} \\
&= 2\nu(X^\top V)^2 D^{-1} - 2\nu^2(X^\top V)^2(D+\nu I_k)^{-1}D^{-1} + \nu V^\top V((D+\nu I_k)^{-1}) \\
&= (X^\top V)^2\Big(2\nu D^{-1} - 2\nu^2(D+\nu I_k)^{-1}D^{-1}\Big) + \nu V^\top V((D+\nu I_k)^{-1}) \\
&= 2\nu(X^\top V)^2(D+\nu I_k)^{-1} + \nu V^\top V(D+\nu I_k)^{-1} \\
&= \nu\Big(V^\top V + 2(X^\top V)^2\Big)(D+\nu I_k)^{-1},
\end{aligned}$$

where the identity

$$(2\nu D^{-1} - 2\nu^2(D+\nu I_k)^{-1}D^{-1})_{ii} = \tfrac{2(\nu D_{ii}+\nu^2)-2\nu^2}{(D_{ii}+\nu)D_{ii}} = \tfrac{2\nu}{D_{ii}+\nu} = 2\nu((D+\nu I_k)^{-1})_{ii}$$

is used. This yields

$$XX^\top\widetilde{\Gamma}_X(V,V) = \nu X\Big(V^\top V + 2(X^\top V)^2\Big)(D+\nu I_k)^{-1}.$$

Moreover, using the symmetry of $\nu(V^\top V + 2(X^\top V)^2) \in \mathbb{R}^{k\times k}_{\mathrm{sym}}$ we obtain by Lemma 3.14, Claim 1

$$\pi^{D+\nu I_k}(X^\top\widetilde{\Gamma}_X(V,V)) = \pi^{D+\nu I_k}(\nu(V^\top V + 2(X^\top V)^2)(D+\nu I_k)^{-1}) = 0.$$

Therefore $\Gamma_X(V,V)$ can be obtained by Theorem 4.15 via calculating

$$\begin{aligned}
-\Gamma_X(V,V) &= -\widetilde{\Gamma}_X(V,V) - XV^\top V + XX^\top\widetilde{\Gamma}_X(V,V) + X\pi^{D+\nu I_k}(V^\top V - X^\top\widetilde{\Gamma}_X(V,v)) \\
&= \Big(-2\nu VX^\top VD^{-1} - \nu XV^\top V(D+\nu I_k)^{-1} + 2\nu^2 X(X^\top V)^2(D+\nu I_k)^{-1}D^{-1}\Big) - XV^\top V \\
&\quad + \Big(\nu XV^\top V(D+\nu I_k)^{-1} + 2\nu X(X^\top V)^2(D+\nu I_k)^{-1}\Big) + X\pi^{D+\nu I_k}(V^\top V) \\
&= 2\nu VV^\top XD^{-1} + 2X(X^\top V)^2(D+\nu I_k)^{-1}(\nu^2 D^{-1} + \nu I_k) - XV^\top V + X\pi^{D+\nu I_k}(V^\top V) \\
&= 2\nu VV^\top XD^{-1} + 2\nu X(X^\top V)^2 D^{-1} - XV^\top V + X\pi^{D+\nu I_k}(V^\top V),
\end{aligned}$$

where the last equality follows due to

$$((D+\nu I_k)^{-1}(\nu^2 D^{-1} + \nu I_k))_{ii} = \tfrac{(\nu^2/D_{ii})+\nu}{D_{ii}+\nu} = \tfrac{\nu(\nu+D_{ii})}{D_{ii}(\nu+D_{ii})} = \nu(D^{-1})_{ii}.$$

This yields the desired result.

Remark 4.19. Corollary 4.18 generalizes the geodesic equation from [13]. Indeed, setting $D = 2I_k$ and $\nu = -\frac{2\alpha+1}{\alpha+1}$ with $\alpha \in \mathbb{R} \setminus \{-1\}$ yields

$$-\Gamma_X(V, V) = \nu V V^\top X + \nu X(X^\top V)^2 - X V^\top V \tag{4.30}$$

due to $\pi^{(2+\nu)I_k}(V^\top V) = \operatorname{skew}(V^\top V) = 0$ in accordance with [13, Eq. (65)].

Remark 4.20. We are not aware of an explicit solution of the geodesic equation for general diagonal matrices D and E. To our best knowledge, an explicit solution is only known for the special case $D = 2I_k$ and $E = \nu I_k$, see [13]. Nevertheless, one could exploit that $(T\mathrm{St}_{n,k}, \omega_{(T\iota)^*\widetilde{L}}, (T\iota)^*\widetilde{L})$ defines a Hamiltonian system whose Hamiltonian vector field is given by the metric spray $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$. This point of view would allow to study the geodesic equation using the theory of integrable systems. However, investigating these aspects in detail is out of the scope of this paper. In this context, we only refer to [24], where geodesic flows on the cotangent bundle $T^*\mathrm{St}_{n,k}$ and their integrability are studied.

5. Pseudo-Riemannian gradients and pseudo-Riemannian Hessians

We now determine pseudo-Riemannian gradients and pseudo-Riemannian Hessians of smooth functions on $\mathrm{St}_{n,k}$. Specific results from [14] are generalized, where similar ideas were used to obtain the gradients and Hessians of smooth function on $\mathrm{St}_{n,k}$ with respect to the one-parameter family of metrics from [13]. Moreover, similar formulas for gradients and Hessians on $\mathrm{St}_{n,k}$ with respect to a family of metrics corresponding to $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$, where $D = \alpha_0 I_k$ and $E = (\alpha_1 - \alpha_0) I_k$ with $\alpha_0, \alpha_1 \in \mathbb{R}$, i.e. a scaled version of the metrics introduced in [13], are independently obtained in [25].

Notation 5.1. From now on, unless indicated otherwise, we denote by $U \subseteq \mathbb{R}^{n\times k}$ an open subset with the property from Lemma 3.8.

5.1. Pseudo-Riemannian Gradients

We first determine the gradient of a smooth function on $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ with respect to the metric $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)} \in \Gamma^\infty(\mathrm{S}^2(T^*\mathrm{St}_{n,k}))$. Let $\sharp_{D,E}\colon T^*_X\mathrm{St}_{n,k} \to T_X\mathrm{St}_{n,k}$ denote the sharp map associated with $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$, i.e. the inverse of the flat map $\flat\colon T_X\mathrm{St}_{n,k} \ni V \mapsto \langle V, \cdot \rangle^{D,E}_X \in T^*_X\mathrm{St}_{n,k}$. Then $\operatorname{grad} f \in \Gamma^\infty(T\mathrm{St}_{n,k})$ is the unique vector field that fulfills

$$\langle \operatorname{grad} f(X), V \rangle^{D,E}_X = \mathrm{d} f\big|_X(V) \quad \Longleftrightarrow \quad \operatorname{grad} f(X) = \big(\mathrm{d} f\big|_X(\cdot)\big)^{\sharp_{D,E}} \tag{5.1}$$

for all $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$, see e.g. [26, Sec. 8.1] for the Riemannian case, which clearly extends to the pseudo-Riemannian case.

Proposition 5.2. *Let $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ be smooth with some smooth extension $F\colon U \to \mathbb{R}$. Then the gradient of f at $X \in \mathrm{St}_{n,k}$ with respect to $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$ is given by*

$$\operatorname{grad} f(X) = \nabla F(X) D^{-1} - X X^\top \nabla F(X) D^{-1} + X \pi^{D+E}\big(X^\top \nabla F(X)(D^{-1} - (D+E)^{-1} E D^{-1})\big). \tag{5.2}$$

Proof. We first compute the gradient of $F\colon U \to \mathbb{R}$ with respect to $\langle \cdot, \cdot \rangle^{D,E}_{(\cdot)}$. Let $X \in U$. Then $\operatorname{grad} F(X) \in \mathbb{R}^{n\times k}$ fulfills

$$\langle \operatorname{grad} F(X), V \rangle^{D,E}_X = \mathrm{d} F\big|_X(V) = \operatorname{tr}\big((\nabla F(X))^\top V\big) \tag{5.3}$$

for all $V \in T_X\mathbb{R}^{n\times k} \cong \mathbb{R}^{n\times k}$. Using the definition of $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$, Equation (5.3) can be rewritten as

$$\operatorname{tr}\left(V^\top(\operatorname{grad}F(X)D + XX^\top \operatorname{grad}F(X)E)\right) = \operatorname{tr}\left(V^\top \nabla F(X)\right).$$

Since $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ is non-degenerated, this is equivalent to the linear equation

$$\operatorname{grad}F(X)D + XX^\top \operatorname{grad}F(X)E = \nabla F(X) \tag{5.4}$$

in terms of $\operatorname{grad}F(X)$. Now assume $X \in \mathrm{St}_{n,k}$. Then the unique solution of (5.4) is given by

$$\operatorname{grad}F(X) = \nabla F(X)D^{-1} - XX^\top \nabla F(X)(D+E)^{-1}ED^{-1}$$

according to Lemma 4.8. Next, we use the well-known formula $\operatorname{grad}f(X) = P_X(\operatorname{grad}F(X))$, where $P_X\colon \mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k}$ is determined in Theorem 3.18. One calculates

$$\begin{aligned}
\operatorname{grad}f(X) &= P_X(\nabla F(X)D^{-1} - XX^\top \nabla F(X)(D+E)^{-1}ED^{-1}) \\
&= \left(\nabla F(X)D^{-1} - XX^\top \nabla F(X)D^{-1} + X\pi^{D+E}(X^\top \nabla F(X)D^{-1})\right) \\
&\quad - \left(XX^\top \nabla F(X)(D+E)^{-1}ED^{-1} - XX^\top(XX^\top \nabla F(X)(D+E)^{-1}ED^{-1}) \right. \\
&\quad \left. + X\pi^{D+E}(X^\top XX^\top \nabla F(X)(D+E)^{-1}ED^{-1})\right) \\
&= \nabla F(X)D^{-1} - XX^\top \nabla F(X)D^{-1} + X\pi^{D+E}(X^\top \nabla F(X)(D^{-1} - (D+E)^{-1}ED^{-1}))
\end{aligned}$$

for $X \in \mathrm{St}_{n,k}$ as desired.

Next we specialize the formula for the gradient to the subfamily $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$.

Proposition 5.3. *Let $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ be smooth with some smooth extension $F\colon U \to \mathbb{R}$. Then the gradient of f with respect to $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ is given by*

$$\operatorname{grad}f(X) = \nabla F(X)D^{-1} - XX^\top \nabla F(X)D^{-1} + X\pi^{D+\nu I_k}(X^\top \nabla F(X)(D+\nu I_k)^{-1}) \tag{5.5}$$

for all $X \in \mathrm{St}_{n,k}$.

Proof. Using Proposition 5.2 we obtain for $X \in T_X\mathrm{St}_{n,k}$

$$\begin{aligned}
\operatorname{grad}f(X) &= \nabla F(X)D^{-1} - XX^\top \nabla F(X)D^{-1} + X\pi^{D+\nu I_k}(X^\top \nabla F(X)(D^{-1} - \nu(D+\nu I_k)^{-1}D^{-1})) \\
&= \nabla F(X)D^{-1} - XX^\top \nabla F(X)D^{-1} + X\pi^{D+\nu I_k}(X^\top \nabla F(X)(D+\nu I_k)^{-1}),
\end{aligned}$$

where the identity

$$(D^{-1} - \nu(D+\nu I_k)^{-1}D^{-1})_{ii} = \tfrac{1}{D_{ii}} - \tfrac{\nu}{(D_{ii}+\nu)D_{ii}} = \tfrac{D_{ii}+\nu-\nu}{(D_{ii}+\nu)D_{ii}} = ((D+\nu I_k)^{-1})_{ii}$$

is used to obtain the last equality.

Corollary 5.4. *Let $\alpha \in T^*_X\mathrm{St}_{n,k}$ be given by*

$$\alpha = \operatorname{tr}\left(V^\top(\cdot)\right)\colon T_X\mathrm{St}_{n,k} \to \mathbb{R}, \quad W \mapsto \operatorname{tr}(V^\top W) \in \mathbb{R}, \tag{5.6}$$

where $V \in \mathbb{R}^{n\times k}$ is some matrix. The sharp map $\sharp_{D,E}\colon T_X^*\mathrm{St}_{n,k} \to T_X\mathrm{St}_{n,k}$ associated with $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,E}$ applied to α is given by

$$\alpha^{\sharp_{D,E}} = \big(\operatorname{tr}(V^\top(\cdot))\big)^{\sharp_{D,\nu I_k}} = VD^{-1} - XX^\top VD^{-1} + X\pi^{D+E}(X^\top V(D^{-1} - (D+E)^{-1}ED^{-1})). \tag{5.7}$$

Specializing $E = \nu I_k$ yields the sharp map with respect to $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,\nu}$ applied to α, namely

$$\alpha^{\sharp_{D,\nu I_k}} = \big(\operatorname{tr}(V^\top(\cdot))\big)^{\sharp_{D,\nu I_k}} = VD^{-1} - XX^\top VD^{-1} + X\pi^{D+\nu I_k}(X^\top V(D+\nu I_k)^{-1}). \tag{5.8}$$

Proof. Consider the smooth function $F\colon \mathbb{R}^{n\times k} \ni X \mapsto \operatorname{tr}(V^\top X) \in \mathbb{R}$ and set $f = F\big|_{\mathrm{St}_{n,k}}\colon \mathrm{St}_{n,k} \to \mathbb{R}$. Then $\mathrm{d}F\big|_X(W) = \operatorname{tr}(V^\top W)$ and thus $\nabla F(X) = V$ follows. Applying Proposition 5.2 and Proposition 5.3, respectively, yields the desired result because of (5.1).

Corollary 5.5. *Proposition 5.3 reproduces some results known from the literature as special cases:*

1. *For $D = I_k$ and $\nu = 0$ one has*

$$\operatorname{grad} f(X) = \nabla F(X) - \tfrac{1}{2}XX^\top\nabla F(X) - \tfrac{1}{2}X(\nabla F(X))^\top X. \tag{5.9}$$

   This coincides with the gradient with respect to the Euclidean metric, see e.g. [2].

2. *For $D = I_k$ and $\nu = -\frac{1}{2}$, one has*

$$\operatorname{grad} f(X) = \nabla f(X) - X(\nabla F(X))^\top X \tag{5.10}$$

   reproducing the formula for the gradient from [10, Eq. (2.53)].

3. *For $D = 2I_k$ and $-2 \neq \nu \in \mathbb{R}$ the gradient of f simplifies to*

$$\operatorname{grad} f(X) = \tfrac{1}{2}\big(\nabla f(X) - \tfrac{\nu+1}{2+\nu}XX^\top\nabla F(X) - \tfrac{1}{2+\nu}X(\nabla F(X))^\top X\big) \tag{5.11}$$

   reproducing the expression for the gradient from [14, Thm. 1].

Proof. These formulas follow by straightforward calculations by plugging the particular choices for D and ν into the expression for $\operatorname{grad} f$ from Proposition 5.3.

5.2. Pseudo-Riemannian Hessians

Next we determine the pseudo-Riemannian Hessian of a smooth function $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$. Here we only consider the subfamily $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,\nu}$ in order to obtain formulas which are not too complicated.

Lemma 5.6. *Let $X \in \mathrm{St}_{n,k}$, $V \in T_X\mathrm{St}_{n,k}$ and let $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ be smooth with some smooth extension $F\colon U \to \mathbb{R}$. The Hessian of f with respect $\langle\cdot,\cdot\rangle_{(\cdot)}^{D,\nu}$ considered as quadratic form is given by*

$$\begin{aligned}\operatorname{Hess}(f)\big|_X(V,V) = {}& \mathrm{D}^2 F(X)(V,V) \\ &+ \mathrm{D}F(X)(2\nu VV^\top XD^{-1} + 2\nu X(X^\top V)^2D^{-1} - XV^\top V + X\pi^{D+\nu I_k}(V^\top V)),\end{aligned} \tag{5.12}$$

where $X \in \mathrm{St}_{n,k}$ and $V, W \in T_X\mathrm{St}_{n,k}$.

Proof. The geodesic $\gamma\colon \mathbb{R} \to \mathrm{St}_{n,k}$ through $\gamma(0) = X \in \mathrm{St}_{n,k}$ with initial velocity $\dot{\gamma}(0) = V \in T_X\mathrm{St}_{n,k}$ fulfills the explicit second order ODE

$$\begin{aligned}\ddot{\gamma}(t) &= 2\nu\dot{\gamma}(t)\dot{\gamma}(t)^\top\gamma(t)D^{-1} + 2\nu\gamma(t)\big(\gamma(t)^\top\dot{\gamma}(t)\big)^2D^{-1} \\ &\quad - \gamma(t)\dot{\gamma}(t)^\top\dot{\gamma}(t) + \gamma(t)\pi^{D+\nu I_k}\big(\dot{\gamma}(t)^\top\dot{\gamma}(t)\big)\end{aligned} \tag{5.13}$$

according to Corollary 4.18. Evaluating (5.13) at $t = 0$ yields

$$\ddot{\gamma}(0) = 2\nu VV^\top XD^{-1} + 2\nu X(X^\top V)^2D^{-1} - XV^\top V + X\pi^{D+\nu I_k}(V^\top V) \tag{5.14}$$

due to the initial conditions $\gamma(0) = X$ and $\dot{\gamma}(0) = V$. The Hessian of f considered as quadratic form can be determined as

$$\mathrm{Hess}(f)\big|_X(V,V) = \tfrac{\mathrm{d}^2}{\mathrm{d}t^2}(f\circ\gamma)(t)\big|_{t=0}, \tag{5.15}$$

see e.g. [26, Prop. 8.3] for the Riemannian case, which clearly extends to pseudo-Riemannian manifolds. Using $f = F\big|_{\mathrm{St}_{n,k}}$, Formula (5.15) yields

$$\mathrm{Hess}(f)\big|_X(V,V) = \mathrm{D}^2 F(X)(\dot{\gamma}(0),\dot{\gamma}(0)) + \mathrm{D}\,F(X)\ddot{\gamma}(0) \tag{5.16}$$

by the chain rule. Plugging (5.14) into (5.16) yields the desired result.

Theorem 5.7. *Let $X \in \mathrm{St}_{n,k}$ and $V, W \in T_X\mathrm{St}_{n,k}$. Moreover, define $\widetilde{D} = D + \nu I_k$. The Hessian of a smooth function $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ with smooth extension $F\colon U \to \mathbb{R}$ with respect $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ is given by*

$$\begin{aligned}\mathrm{Hess}(f)\big|_X(V,W) &= \mathrm{tr}\big((\mathrm{D}(\nabla F)(X)V)^\top W\big) \\ &\quad + \nu\,\mathrm{tr}\big((XD^{-1}(\nabla F(X))^\top V + \nabla F(X)D^{-1}X^\top V)^\top W\big) \\ &\quad + \nu\,\mathrm{tr}\big((XV^\top XX^\top\nabla F(X)D^{-1} + XX^\top\nabla F(X)D^{-1}V^\top X)^\top W\big) \\ &\quad - \tfrac{1}{2}\,\mathrm{tr}\big((VX^\top\nabla F(X) + V(\nabla F(X))^\top X)^\top W\big) \\ &\quad + \tfrac{1}{2}\,\mathrm{tr}\big((V\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1})\widetilde{D} - V\widetilde{D}\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1}))^\top W\big).\end{aligned} \tag{5.17}$$

Proof. Let $(X,V),(X,W) \in T\mathrm{St}_{n,k}$. We obtain for the Hessian of f as symmetric 2-tensor

$$\begin{aligned}\mathrm{Hess}(f)\big|_X(V,W) &= \mathrm{tr}\big((\mathrm{D}(\nabla F)(X)V)^\top W\big) \\ &\quad + \nu\,\mathrm{tr}\big((\nabla F(X))^\top(VW^\top + WV^\top)XD^{-1}\big) \\ &\quad + \nu\,\mathrm{tr}\big((\nabla F(X))^\top X(X^\top VX^\top W + X^\top WX^\top V)D^{-1}\big) \\ &\quad - \tfrac{1}{2}\,\mathrm{tr}\big((\nabla F(X))^\top X(V^\top W + W^\top V)\big) \\ &\quad + \tfrac{1}{2}\,\mathrm{tr}\big((\nabla F(X))^\top X\pi^{D+\nu I_k}(V^\top W + W^\top V)\big)\end{aligned} \tag{5.18}$$

by applying polarization to the quadratic form obtained in Lemma 5.6 and using the identities

$$\mathrm{D}\,F(X)V = \mathrm{tr}\big((\nabla F(X))^\top V\big) \quad \text{and} \quad \mathrm{D}^2 F(X)(V,W) = \mathrm{tr}\big((\mathrm{D}(\nabla F)(X)V)^\top W\big).$$

Next, we set $\widetilde{D} = D+\nu I_k$ which is invertible according to Notation 3.13. Since the orthogonal projection $\pi^{\widetilde{D}}\colon \mathbb{R}^{k\times k} \to \mathfrak{so}(k) \subseteq \mathbb{R}^{k\times k}$ is self-adjoint with respect to the scalar product

$$\langle\cdot,\cdot\rangle^{\widetilde{D}}\colon \mathbb{R}^{k\times k}\times\mathbb{R}^{k\times k} \to \mathbb{R}, \quad (V,W) \mapsto \mathrm{tr}\,(V^\top W\widetilde{D})$$

on $\mathbb{R}^{k\times k}$, we calculate

$$\begin{aligned}\operatorname{tr}\left((\nabla F(X))^\top X\pi^{\widetilde{D}}(V^\top W + W^\top V)\right) &= \operatorname{tr}\left((X^\top\nabla F(X)\widetilde{D}^{-1})^\top \pi^{\widetilde{D}}(V^\top W + W^\top V)\widetilde{D}\right)\\ &= \langle X^\top\nabla F(X)\widetilde{D}^{-1}, \pi^{\widetilde{D}}(V^\top W + W^\top V)\rangle^{\widetilde{D}}\\ &= \langle \pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1}), V^\top W + W^\top V\rangle^{\widetilde{D}}\\ &= \operatorname{tr}\left((V\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1})\widetilde{D} - V\widetilde{D}\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1}))^\top W\right)\end{aligned} \tag{5.19}$$

by exploiting $\operatorname{im}(\pi^{\widetilde{D}}) = \mathfrak{so}(k)$. The desired result follows by rewriting (5.18) using well-known properties of the trace and applying (5.19) to the last summand of (5.18).

Corollary 5.8. *5.7: Let $D = 2I_k$ and $-2 \neq \nu \in \mathbb{R}$. Then the Hessian of the smooth function $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ with respect to $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ reads*

$$\begin{aligned}\operatorname{Hess}(f)\big|_X(V,W) = &\operatorname{tr}\left((\mathrm{D}(\nabla F)(X)V)^\top W\right)\\ &+ \tfrac{\nu}{2}\operatorname{tr}\left((X(\nabla F(X))^\top V + \nabla F(X)X^\top V)^\top W\right)\\ &+ \tfrac{\nu}{2}\operatorname{tr}\left((XV^\top XX^\top \nabla F(X) + XX^\top\nabla F(X)V^\top X)^\top W\right)\\ &- \tfrac{1}{2}\operatorname{tr}\left((VX^\top\nabla F(X) + V(\nabla F(X))^\top X)^\top W\right)\end{aligned} \tag{5.20}$$

with $X \in \mathrm{St}_{n,k}$ and $V, W \in T_X\mathrm{St}_{n,k}$ reproducing the formula from [14, Thm. 2].

Proof. We set $D = 2I_k$ in Theorem 5.7. Obviously, $\widetilde{D} = (2+\nu)I_k$ holds. Hence

$$\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1})\widetilde{D} = \pi^{\widetilde{D}}(X^\top\nabla F(X)) = \widetilde{D}\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1})$$

is fulfilled by the linearity of $\pi^{\widetilde{D}}\colon \mathbb{R}^{n\times k} \to \mathfrak{so}(k) \subseteq \mathbb{R}^{n\times k}$. Thus the last summand of (5.17) vanishes.

Theorem 5.7 yields an expression for the Hessian of $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ as covariant 2-tensor. However, for applications, see e.g. [2, Chap. 6], an expression for the Hessian of f viewed as section of $\mathrm{End}(T\mathrm{St}_{n,k})$ is desirable. Thus, following [14, Re. 6], we state the next remark and the next corollary.

Remark 5.9. In [2, Eq. (6.3)] the Hessian of a smooth function $f\colon M \to \mathbb{R}$ on a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ endowed with a covariant derivative ∇ is defined as

$$\widetilde{\operatorname{Hess}}(f)\big|_x(v_x) = \nabla_{v_x}\operatorname{grad} f\big|_x$$

for $x \in M$ and $v_x \in T_xM$. In particular, $\widetilde{\operatorname{Hess}}(f) \in \Gamma^\infty(\mathrm{End}(TM))$ holds. If ∇ is chosen as the Levi-Civita covariant derivative ∇^{LC}, then $\widetilde{\operatorname{Hess}}(f)$ is related to $\operatorname{Hess}(f) \in \Gamma^\infty(\mathrm{S}^2(T^*M))$ via

$$\langle\widetilde{\operatorname{Hess}}(f)\big|_x(v), w\rangle = \langle\nabla^{\mathrm{LC}}_{v_x}\operatorname{grad} f\big|_x, w_x\rangle = \operatorname{Hess}(f)\big|_x(v_x, w_x), \tag{5.21}$$

where $x \in M$ and $v_x, w_x \in T_xM$, see e.g. [26, Prop. 8.1] for a proof for the Riemannian case. Clearly, Equation (5.21) holds in the pseudo-Riemannian case, too. We rewrite (5.21) equivalently as

$$\langle\widetilde{\operatorname{Hess}}(f)\big|_x(v_x), \cdot\rangle = \operatorname{Hess}(f)\big|_x(v_x, \cdot). \tag{5.22}$$

Applying the sharp map $\sharp\colon T_x^*M \to T_xM$ associated with $\langle\cdot,\cdot\rangle$ on both sides of (5.22) yields

$$\widetilde{\operatorname{Hess}}(f)\big|_x(v_x) = \left(\operatorname{Hess}(f)\big|_x(v_x,\cdot)\right)^\sharp. \tag{5.23}$$

Corollary 5.10. *Let $f\colon \mathrm{St}_{n,k} \to \mathbb{R}$ be smooth with some smooth extension $F\colon U \to \mathbb{R}$. The Hessian of f with respect to $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ considered as a section of $\mathrm{End}(T\mathrm{St}_{n,k})$ is given by*

$$\begin{aligned}\widetilde{\mathrm{Hess}}(f)\big|_X(V) = L^{D,\nu}_X\Big(&\mathrm{D}(\nabla F)(X)V + \nu(XD^{-1}(\nabla F(X))^\top V + \nabla F(X)D^{-1}X^\top V)\\ &+ \nu(XV^\top XX^\top \nabla F(X)D^{-1} + XX^\top \nabla F(X)D^{-1}V^\top X)\\ &- \tfrac{1}{2}(VX^\top \nabla F(X) + V(\nabla F(X))^\top X)\\ &+ \tfrac{1}{2}(V\pi^{\widetilde{D}}(X^\top \nabla F(X)\widetilde{D}^{-1})\widetilde{D} - V\widetilde{D}\pi^{\widetilde{D}}(X^\top\nabla F(X)\widetilde{D}^{-1}))\Big)\end{aligned}$$

for $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$, where $\widetilde{D} = D + \nu I_k$ and $L^{D,\nu}_X\colon \mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k} \subseteq \mathbb{R}^{n\times k}$ is the linear map given by

$$L^{D,\nu}_X(V) = VD^{-1} - XX^\top VD^{-1} + X\pi^{D+\nu I_k}(X^\top V(D+\nu I_k)^{-1}).$$

Proof. We have already obtained $\mathrm{Hess}(f)$ in Theorem 5.7 in such a form that the formula for the sharp map from Corollary 5.4 can be applied to $\mathrm{Hess}(f)\big|_X(V,\cdot) \in T^*_X\mathrm{St}_{n,k}$. Now Remark 5.9 yields the desired result.

6. Second fundamental form and Levi-Civita covariant derivative

In this section, we compute the second fundamental form of $\mathrm{St}_{n,k}$ considered as pseudo-Riemannian submanifold of $(U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$. Moreover, an expression for the Levi-Civita covariant derivative on $\mathrm{St}_{n,k}$ is derived. We first recall Notation 5.1. Unless indicated otherwise, we denote by $U \subseteq \mathbb{R}^{n\times k}$ an open neighbourhood of $\mathrm{St}_{n,k}$ with the property from Lemma 3.8.

6.1. Levi-Civita covariant derivative on U

We consider the Levi-Civita covariant derivative $\widetilde{\nabla^{\mathrm{LC}}}$ on U with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$. Recall Proposition 4.9. For $(X,V) \in TU$ the spray $\widetilde{S} \in \Gamma^\infty(T(TU))$ associated with the metric $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ on U is given by

$$\widetilde{S}(X,V) = (X, V, V, -\widetilde{\Gamma}_X(V,V)) \tag{6.1}$$

where $\widetilde{\Gamma}_X(V,V)$ is the unique solution of the linear equation (4.11). We now discuss how $\widetilde{\Gamma}_X(V,V) \in \mathbb{R}^{n\times k}$ is related to the Christoffel symbols of the Levi-Civita covariant derivative $\widetilde{\nabla^{\mathrm{LC}}}$ on $(U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$. To this end, we view $\mathrm{id}_U\colon U \ni X \mapsto X \in U$ as the global chart $(U, \mathrm{id}_U) = (U, X_{ij})$ of U and identify the coordinate vector fields $\frac{\partial}{\partial X_{ij}}$ with the constant functions $U \ni X \mapsto E_{ij} \in \mathbb{R}^{n\times k}$. Then (6.1) is a coordinate expression for the metric spray $\widetilde{S}$ with respect to the global chart $(TU, (X_{ij}, V_{ij}))$ induced by the chart (U, X_{ij}), see also Proposition 4.9. Thus the local form of metric sprays, see (4.2), implies that the entry $(\widetilde{\Gamma}_X(V,V))_{ij}$ fulfills

$$(\widetilde{\Gamma}_X(V,V))_{ij} = \sum_{a,c=1}^{n}\sum_{b,d=1}^{k} \widetilde{\Gamma}^{(i,j)}_{(a,b),(c,d)}\Big|_X V_{ab}V_{cd}, \tag{6.2}$$

where $V = (V_{ij}) \in \mathbb{R}^{n\times k}$ and the functions $\widetilde{\Gamma}^{(i,j)}_{(a,b),(c,d)}\colon U \ni X \mapsto \widetilde{\Gamma}^{(i,j)}_{(a,b),(c,d)}\big|_X \in \mathbb{R}$ denote the Christoffel symbols of $\widetilde{\nabla^{\mathrm{LC}}}$ with respect to $(U, (X_{ij}))$. Hence $\widetilde{\nabla^{\mathrm{LC}}}$ can be expressed with respect to the global chart (U, X_{ij}) as

$$\widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W}\big|_X = \mathrm{D}\,\widetilde{W}(X)\widetilde{V}\big|_X + \widetilde{\Gamma}_X(\widetilde{V}\big|_X, \widetilde{W}\big|_X), \tag{6.3}$$

for vector fields $\widetilde{V}, \widetilde{W} \in \Gamma^\infty(TU)$ and $X \in U$, see e.g. [27, Chap. 4]. A similar "matrix notation" for Christoffel symbols has already appeared in [10, Sec. 2.2.3], where, in addition, it is mentioned that (for fixed $X \in U$) the symmetric bilinear map $\mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \ni (V, W) \mapsto \widetilde{\Gamma}_X(V, W) \in \mathbb{R}^{n\times k}$ can be obtained from the quadratic map $\mathbb{R}^{n\times k} \ni V \mapsto \widetilde{\Gamma}_X(V, V) \in \mathbb{R}^{n\times k}$ by polarization. Hence the Christoffel symbols on U can be identified with the smooth map

$$\widetilde{\Gamma} \colon U \to \mathrm{S}^2((\mathbb{R}^{n\times k})^*) \otimes \mathbb{R}^{n\times k}, \quad X \mapsto ((V, W) \mapsto \widetilde{\Gamma}_X(V, W)). \tag{6.4}$$

The "Christoffel symbols" from [10, Sec. 2.2.3] will be discussed in Remark 6.11 below.

Next we give an expression for the Levi-Civita covariant derivative ∇^{LC} on $\mathrm{St}_{n,k}$ with respect $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$. We refer to Proposition 6.8 as well as Corollary 6.9 below for an alternative formula for ∇^{LC}.

Proposition 6.1. *Let $V, W \in \Gamma^\infty(T\mathrm{St}_{n,k})$. The Levi-Civita covariant derivative on $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ is given by*

$$\nabla^{\mathrm{LC}}_V W\big|_X = P_X\Big(\mathrm{D}\,\widetilde{W}(X)V\big|_X + \widetilde{\Gamma}_X(V\big|_X, W\big|_X)\Big) \tag{6.5}$$

for all $X \in \mathrm{St}_{n,k}$, where $\widetilde{V} \in \Gamma^\infty(TU)$ is a smooth extensions of V. Here $\widetilde{\Gamma}$ is defined by (6.4). *Moreover, $P_X \colon \mathbb{R}^{n\times k} \to T_X\mathrm{St}_{n,k}$ is the orthogonal projection with respect to $\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)}$ from Theorem 3.18.*

Proof. Since $\mathrm{St}_{n,k}$ is a pseudo-Riemannian submanifold of $(U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$, the result follows by (6.3) due to

$$\nabla^{\mathrm{LC}}_V W\big|_X = P_X\Big(\widetilde{\nabla^{\mathrm{LC}}_V}\widetilde{W}\big|_X\Big),$$

see e.g. [17, Chap. 4, Lem. 3].

6.2. Second fundamental form

We now consider the second fundamental form, also called shape operator, of $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$. We refer to [17, Chap. 4] for general properties of pseudo-Riemanian submanifolds and the second fundamental form, see also [27, Chap. 8] for the Riemannian case. Using these references, we briefly introduce the notation which is used in the sequel subsections.

Let M be a pseudo-Riemannian submanifold of a pseudo-Riemannian manifold $(\widetilde{M}, \langle\cdot,\cdot\rangle)$. The corresponding Levi-Civita covariant derivatives on M and $\widetilde{M}$ are denoted by ∇^{LC} and $\widetilde{\nabla^{\mathrm{LC}}}$, respectively. Moreover, let $NM \to M$ be the normal bundle of M and let $\mathrm{II} \in \Gamma^\infty((\mathrm{S}^2(T^*M)) \otimes NM)$ be the second fundamental form of M, see e.g. [17, Chap. 4, Lem. 4], defined by

$$\mathrm{II}(V, W)\big|_x = P^\perp_x\Big(\widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W}\big|_x\Big), \quad x \in M, \quad V, W \in \Gamma^\infty(TM), \tag{6.6}$$

where $\widetilde{V}, \widetilde{W} \in \Gamma^\infty(T\widetilde{M})$ are smooth extensions of $V, W \in \Gamma^\infty(TM)$, respectively, and $P^\perp_x \colon T_x\widetilde{M} \to N_xM$ denotes the orthogonal projection onto the normal space $N_xM = (T_xM)^\perp$. The Levi-Civita covariant derivative on M fulfills

$$\nabla^{\mathrm{LC}}_V W = \widetilde{\nabla^{\mathrm{LC}}_V}\widetilde{W} - \mathrm{II}(V, W) \tag{6.7}$$

for all $V, W \in \Gamma^\infty(TM)$ by [17, Chap. 4]. Here $\widetilde{W}$ is again some smooth extension of W to $\widetilde{M}$. The identity (6.7) is named Gauß formula in [27, Thm. 8.2], which includes a proof for the Riemannian case, as well.

Lemma 6.2. *Define* $\nabla\colon \Gamma^\infty(T\widetilde{M}) \times \Gamma^\infty(T\widetilde{M}) \to \Gamma^\infty(T\widetilde{M})$ *by*

$$\nabla_{\widetilde{V}}\widetilde{W} = \widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W} - \widetilde{\mathrm{II}}(\widetilde{V}, \widetilde{W}), \quad \widetilde{V}, \widetilde{W} \in \Gamma^\infty(T\widetilde{M}), \tag{6.8}$$

where $\widetilde{\mathrm{II}} \in \Gamma^\infty((\mathrm{S}^2T^*\widetilde{M}) \otimes T\widetilde{M})$ *denotes a smooth extension of the second fundamental form* II *on* M *to* $\widetilde{M}$. *Then* ∇ *is a covariant derivative on* $\widetilde{M}$ *whose restriction to* M *coincides with* ∇^{LC}, *i.e.*

$$\nabla^{\mathrm{LC}}_V W\big|_x = \nabla_{\widetilde{V}}\widetilde{W}\big|_x \tag{6.9}$$

holds for all $x \in M$ *and* $V, W \in \Gamma^\infty(TM)$ *with smooth extensions* $\widetilde{V}, \widetilde{W} \in \Gamma^\infty(T\widetilde{M})$. *Moreover, the Christoffel symbols of* ∇ *with respect to the local chart* (U, x) *of* $\widetilde{M}$ *are given by*

$$\Gamma^k_{ij} = \widetilde{\Gamma}^k_{ij} - \widetilde{\mathrm{II}}^k_{ij}. \tag{6.10}$$

Here $\widetilde{\mathrm{II}}^k_{ij}$ *is defined by* $\widetilde{\mathrm{II}}(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}) = \widetilde{\mathrm{II}}^k_{ij}\frac{\partial}{\partial x^k}$ *using Einstein summation convention and* $\widetilde{\Gamma}^k_{ij}$ *denote the Christoffel symbols of* $\widetilde{\nabla^{\mathrm{LC}}}$ *with respect to the chart* (U, x).

Proof. Obviously, the definition of ∇ yields a covariant derivative on M. Moreover, the Gauß formula (6.7) implies

$$\nabla^{\mathrm{LC}}_{\widetilde{V}}\widetilde{W}\big|_x = \widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W}\big|_x - \mathrm{II}\big|_x(\widetilde{V}\big|_x, \widetilde{W}\big|_x) = \nabla_{\widetilde{V}}\widetilde{W}\big|_x$$

for all $x \in M$ and all $V, W \in \Gamma^\infty(TM)$ with smooth extensions $\widetilde{V}, \widetilde{W} \in \Gamma^\infty(T\widetilde{M})$, respectively.

It remains to show the formula for the Christoffel symbols. Let (U, x) be a local chart of $\widetilde{M}$. Using Einstein summation convention one obtains

$$\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \widetilde{\nabla^{\mathrm{LC}}}_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} - \widetilde{\mathrm{II}}(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}) = \widetilde{\Gamma}^k_{ij}\frac{\partial}{\partial x^k} - \widetilde{\mathrm{II}}^k_{ij}\frac{\partial}{\partial x^k} = (\widetilde{\Gamma}^k_{ij} - \widetilde{\mathrm{II}}^k_{ij})\frac{\partial}{\partial x^k}$$

showing the desired result.

Remark 6.3. The definition of the covariant derivative ∇ on $\widetilde{M}$ in Lemma 6.2 depends on the choice of the smooth extension $\widetilde{\mathrm{II}}$ of II. Nevertheless, Equation (6.9) is independent of the extension $\widetilde{\mathrm{II}}$ of II.

Reformulating [16, Cor. 8.4.2] yields the next lemma which allows for computing the second fundamental form of $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$.

Lemma 6.4. *Let* $M \subseteq \widetilde{M}$ *be a pseudo-Riemannian submanifold of* $(\widetilde{M}, \langle\cdot,\cdot\rangle)$. *Moreover, we denote by* $\widetilde{S} \in \Gamma^\infty(T(T\widetilde{M}))$ *and* $S \in \Gamma^\infty(T(TM))$ *the metric sprays on* TM *and* $T\widetilde{M}$, *respectively. Then*

$$(S - \widetilde{S})(v_x) = (\mathrm{II}\big|_x(v_x, v_x))^{\mathrm{ver}}\big|_{v_x}, \tag{6.11}$$

holds for all $x \in M$ *and* $v_x \in T_xM$, *where*

$$(\cdot)^{\mathrm{ver}}\big|_{v_x}\colon T_xM \to \mathrm{Ver}_{v_x}(TM) \subseteq T_{v_x}(TM)$$

is the vertical lift and $\mathrm{II} \in \Gamma^\infty((\mathrm{S}^2(T^*M)) \otimes NM)$ *is the second fundamental form of* $M \subseteq \widetilde{M}$.

Proof. This is a direct consequence of [16, Cor. 8.4.2] as well as the definition II recalled in (6.6).

Lemma 6.4 applied to $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ yields an expression for the second fundamental form.

Proposition 6.5. *Consider* $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ *as pseudo-Riemannian submanifold. Then the following assertions are fulfilled:*

1. *The second fundamental form of* $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ *is given by*
$$\mathrm{II}\big|_X(V, W) = \widetilde{\Gamma}_X(V, W) - \Gamma_X(V, W) \tag{6.12}$$
for all $X \in \mathrm{St}_{n,k}$ *and* $V, W \in T_X\mathrm{St}_{n,k}$*, where* Γ_X *and* $\widetilde{\Gamma}_X$ *denote the symmetric bilinear maps associated with the quadratic maps defined by the sprays* $S \in \Gamma^\infty(T(T\mathrm{St}_{n,k}))$ *and* $\widetilde{S} \in \Gamma^\infty(T(TU))$*, respectively.*
2. *A smooth extension* $\widetilde{\mathrm{II}} \in \Gamma^\infty((\mathrm{S}^2(T^*U)) \otimes TU)$ *of* II *is given*
$$\widetilde{\mathrm{II}}\big|_X(V, W) = \tilde{\Gamma}_X(V, W) - \Gamma_X(V, W), \tag{6.13}$$
for all $X \in U$ *and* $V, W \in T_XU \cong \mathbb{R}^{n\times k}$*, Here we view* $\Gamma_X(V, W)$ *as in Remark 4.16, i.e. as the smooth map* $\Gamma\colon U \to \mathrm{S}^2((\mathbb{R}^{n\times k})^*) \otimes \mathbb{R}^{n\times k}$ *defined in* (4.25).

Proof. Lemma 6.4 applied to $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$ implies that
$$S(X, V) - \widetilde{S}(X, V) = (\mathrm{II}\big|_X(V, V))^{\mathrm{ver}}\big|_{(X,V)} \tag{6.14}$$
holds for all $(X, V) \in T\mathrm{St}_{n,k}$. The vertical lift for fixed $(X, V) \in TU$ is the linear isomorphism
$$(\cdot)^{\mathrm{ver}}\big|_{(X,V)}\colon TU \to \mathrm{Ver}(TU)_{(X,V)}, \quad (X, W) \mapsto (X, W)^{\mathrm{ver}}\big|_{(X,V)} = (X, V, 0, W),$$
according to its local expression, see e.g. [20, Sec. 8.12]. Thus
$$\mathrm{II}\big|_X(V, V) = -\Gamma_X(V, V) - (-\widetilde{\Gamma}_X(V, V)) = \widetilde{\Gamma}_X(V, V) - \Gamma_X(V, V)$$
follows by (6.14). Since the quadratic map $T_X\mathrm{St}_{n,k} \ni V \mapsto \widetilde{\Gamma}_X(V, V) - \Gamma_X(V, V) \in \mathbb{R}^{n\times k}$ determines uniquely the associated symmetric billinear map, Claim 1 is shown. Now Claim 2 is obvious.

The second fundamental from can be simplified for all metrics in the subfamily $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$.

Corollary 6.6. *The second fundamental form of* $\mathrm{St}_{n,k} \subseteq (U, \langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)})$ *is given by*
$$\begin{aligned}\mathrm{II}\big|_X(V, W) = &-\tfrac{1}{2}X(V^\top W + W^\top V)D(D + \nu I_k)^{-1} + \nu X(X^\top VX^\top W + X^\top WX^\top V)(D + \nu I_k)^{-1}\\ &+ \tfrac{1}{2}X\pi^{D+\nu I_k}(V^\top W + W^\top V)\end{aligned} \tag{6.15}$$
for all $X \in \mathrm{St}_{n,k}$ *and* $V, W \in T_X\mathrm{St}_{n,k}$.

Proof. Let $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$. We first compute the quadratic map associated with II. Using Corollary 4.12 and Corollary 4.18, Proposition 6.5 implies
$$\begin{aligned}\mathrm{II}\big|_X(V, V) &= \widetilde{\Gamma}_X(V, V) - \Gamma_X(V, V)\\ &= \Big(2\nu VX^\top VD^{-1} + \nu XV^\top V(D + \nu I_k)^{-1} - 2\nu^2X(X^\top V)^2(D + \nu I_k)^{-1}D^{-1}\Big)\\ &\quad + \Big(2\nu VV^\top XD^{-1} + 2\nu X(X^\top V)^2D^{-1} - XV^\top V + X\pi^{D+\nu I_k}(V^\top V)\Big)\\ &= XV^\top V\Big(\nu(D + \nu I_k)^{-1} - I_k\Big) + 2\nu X(X^\top V)^2\Big(D^{-1} - \nu(D + \nu I_k)^{-1}D^{-1}\Big) + X\pi^{D+\nu I_k}(V^\top V)\\ &= -XV^\top VD(D + \nu I_k)^{-1} + 2\nu X(X^\top V)^2(D + \nu I_k)^{-1} + X\pi^{D+\nu I_k}(V^\top V).\end{aligned} \tag{6.16}$$

Here we exploited

$$(\nu(D+\nu I_k)^{-1}-I_k)_{ii}=\tfrac{\nu}{D_{ii}+\nu}-1=\tfrac{\nu-(D_{ii}+\nu)}{D_{ii}+\nu}=-\tfrac{D_{ii}}{D_{ii}+\nu}=-(D(D+\nu)^{-1})_{ii}$$

as well as

$$(D^{-1}-\nu(D+\nu I_k)^{-1}D^{-1})_{ii}=\tfrac{1}{D_{ii}}(1-\tfrac{\nu}{D_{ii}+\nu})=\tfrac{1}{D_{ii}}\tfrac{D_{ii}+\nu-\nu}{D_{ii}+\nu}=((D+\nu I_k)^{-1})_{ii}.$$

The desired result follows by polarization.

Corollary 6.7. *For $D=2I_k$ and $E=\nu I_k$ the second fundamental form is given by*

$$\mathrm{II}\big|_X(V,W)=-\tfrac{1}{2+\nu}X(V^\top W+W^\top V)+\tfrac{\nu}{2+\nu}X(X^\top VX^\top W+X^\top WX^\top V) \tag{6.17}$$

for $X\in\mathrm{St}_{n,k}$ and $V,W\in T_X\mathrm{St}_{n,k}$.

Proof. Plugging $D=2I_k$ into the formula from Corollary 6.6 the claim follows by a straightforward calculation by exploiting $\pi^{2I_k+\nu I_k}=\mathrm{skew}\colon\mathbb{R}^{k\times k}\to\mathfrak{so}(k)$.

6.3. *Levi-Civita Covariant Derivative on* $\mathrm{St}_{n,k}$

Next we derive an alternative expression for the Levi-Civita covariant derivative on $(\mathrm{St}_{n,k},\langle\cdot,\cdot\rangle^{D,E}_{(\cdot)})$.

Proposition 6.8. *The covariant derivative ∇ on U from Lemma 6.2 fulfills for $\widetilde{V},\widetilde{W}\in\Gamma^\infty(TU)$ and $X\in U$*

$$\nabla_{\widetilde{V}}\widetilde{W}\big|_X=\widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W}\big|_X-\widetilde{\mathrm{II}}\big|_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)=\mathrm{D}\,\widetilde{W}(X)\widetilde{V}\big|_X+\Gamma_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X), \tag{6.18}$$

where Γ denotes the smooth map $U\to\mathrm{S}^2((\mathbb{R}^{n\times k})^*)\otimes\mathbb{R}^{n\times k}$ defined in (4.25). *If $\widetilde{V},\widetilde{W}\in\Gamma^\infty(TU)$ are smooth extensions of $V,W\in\Gamma^\infty(T\mathrm{St}_{n,k})$, respectively, then*

$$\nabla^{\mathrm{LC}}_V W\big|_X=\mathrm{D}\,\widetilde{W}(X)V\big|_X+\Gamma_X(V\big|_X,W\big|_X) \tag{6.19}$$

is satisfied for all $X\in\mathrm{St}_{n,k}$.

Proof. Using Lemma 6.2 and Proposition 6.5 we compute

$$\begin{aligned}\nabla_{\widetilde{V}}\widetilde{W}\big|_X&=\widetilde{\nabla^{\mathrm{LC}}_{\widetilde{V}}}\widetilde{W}\big|_X-\widetilde{\mathrm{II}}\big|_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)\\&=\Big(\mathrm{D}\,\widetilde{W}(X)\widetilde{V}\big|_X+\widetilde{\Gamma}_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)\Big)-\Big(\widetilde{\Gamma}_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)-\Gamma_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)\Big)\\&=\mathrm{D}\,\widetilde{W}(X)\widetilde{V}\big|_X+\Gamma_X(\widetilde{V}\big|_X,\widetilde{W}\big|_X)\end{aligned}$$

for $V,W\in\Gamma^\infty(TU)$ and $X\in U$ showing (6.18). If $\widetilde{V},\widetilde{W}$ are smooth extensions of $V,W\in\Gamma^\infty(T\mathrm{St}_{n,k})$, respectively, we obtain

$$\nabla_{\widetilde{V}}\widetilde{W}\big|_X=\nabla^{\mathrm{LC}}_V W\big|_X=\mathrm{D}\,\widetilde{W}(X)V\big|_X+\Gamma_X(V\big|_X,W\big|_X)$$

for all $X\in\mathrm{St}_{n,k}$ by Lemma 6.2 proving (6.19).

Proposition 6.8 yields a more explicit formula for the subfamily $\langle\cdot,\cdot\rangle^{\nu}_{(\cdot)}$.

Corollary 6.9. *Let $V, W \in \Gamma^\infty(T\mathrm{St}_{n,k})$ and let $\widetilde{W} \in \Gamma^\infty(TU)$ be smooth extension of W. The Levi-Civita covariant derivative on $\mathrm{St}_{n,k}$ with respect to the metric $\langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)}$ is given by*

$$\nabla^{\mathrm{LC}}_V W\big|_X = \mathrm{D}\,\widetilde{W}(X)V\big|_X + \Gamma_X(V\big|_X, W\big|_X) \tag{6.20}$$

for $X \in \mathrm{St}_{n,k}$, where

$$\begin{aligned}\Gamma_X(V, W) = &-\nu(VW^\top + WV^\top)XD^{-1} - \nu X(X^\top VX^\top W + X^\top WX^\top V)D^{-1}\\ &+ \tfrac{1}{2}X(V^\top W + W^\top V) - \tfrac{1}{2}X\pi^{D+\nu I_k}(V^\top W + W^\top V)\end{aligned} \tag{6.21}$$

writing $V = V\big|_X$ and $W = W\big|_X$ for short.

Proof. The quadratic map $\Gamma_X\colon T_X\mathrm{St}_{n,k} \ni V \mapsto \Gamma_X(V, V) \in \mathbb{R}^{n\times k}$ is determined in Corollary 4.18. The associated symmetric bilinear map $T_X\mathrm{St}_{n,k} \times T_X\mathrm{St}_{n,k} \ni (V, W) \mapsto \Gamma_X(V, W) \in \mathbb{R}^{n\times k}$ can be obtained by polarization. Now Proposition 6.8 yields the desired result.

Corollary 6.9 yields an expression for the covariant derivative with respect to the family of metrics introduced in [13].

Corollary 6.10. *Using the notation from Corollary 6.9 one obtains for ∇^{LC} on $(\mathrm{St}_{n,k}, \langle\cdot,\cdot\rangle^{D,\nu}_{(\cdot)})$ with $D = 2I_k$ and $-2 \neq \nu \in \mathbb{R}$*

$$\nabla^{\mathrm{LC}}_V W\big|_X = \mathrm{D}\,\widetilde{W}(X)V - \tfrac{\nu}{2}(VW^\top + WV^\top)X - \tfrac{\nu}{2}X(X^\top VX^\top W + X^\top WX^\top V) + \tfrac{1}{2}X(V^\top W + W^\top V). \tag{6.22}$$

Proof. Plugging $D = 2I_k$ into the formula from Corollary 6.9 yields the desired result by exploiting $\pi^{2I_k+\nu I_k}(V^\top W + W^\top V) = \mathrm{skew}(V^\top W + W^\top V) = 0$ for all $V, W \in T_X\mathrm{St}_{n,k}$.

By setting $D = \alpha_0 I_k$ and $E = (\alpha_1 - \alpha_0)I_k$ for $\alpha_0, \alpha_1 \in \mathbb{R}$, Corollary 6.9 reproduces [25, Eq. (5.4)], where this expression has been obtained independently. Formulas for ∇^{LC} of a similar form as in Proposition 6.8 or Corollary 6.9 have already appeared in the literature in [10, 25], see also [28, Sec. 4]. In the next remark we relate the summand $\Gamma_X(V, W)$ in these formulas to the Christoffel symbols of the covariant derivative ∇ on the open $U \subseteq \mathbb{R}^{n\times k}$.

Remark 6.11. Consider the smooth map $\Gamma\colon U \ni X \mapsto ((V, W) \mapsto \Gamma_X(V, W)) \in \mathrm{S}^2((\mathbb{R}^{n\times k})^*) \otimes \mathbb{R}^{n\times k}$ from (4.25) in Remark 4.16. The Christoffel symbols of the covariant derivative $\nabla_V W = \widetilde{\nabla^{\mathrm{LC}}_V} W - \widetilde{\mathrm{II}}(V, W)$ on U with respect to $(U, \mathrm{id}_U) = (U, X_{ij})$ corresponds to the entries of the matrix $\Gamma_X(V, W)$ by Proposition 6.8. More precisely, we again identify the coordinate vector field $\frac{\partial}{\partial X_{ij}}$ with the map $U \ni X \mapsto E_{ij} \in \mathbb{R}^{n\times k}$. Then the (i, j)-entry of $\Gamma_X(V, W) \in \mathbb{R}^{n\times k}$ is given by a formula similar to (6.2), namely

$$(\Gamma_X(V, W))_{ij} = \sum_{a,c=1}^{n}\sum_{b,d=1}^{k} \Gamma^{(i,j)}_{(a,b),(c,d)} V_{ab} W_{cd}, \tag{6.23}$$

where $\Gamma^{(i,j)}_{(a,b),(c,d)}\colon U \to \mathbb{R}$ are the Christoffel symbols of ∇ with respect to the global chart (U, X_{ij}), see Lemma 6.2 and Proposition 6.8. We point out that the map Γ in (4.25) corresponds to the Christoffel symbols of the covariant derivative ∇ on U but it *cannot* correspond to the Christoffel symbols of ∇^{LC} on $\mathrm{St}_{n,k}$ due to $\dim(\mathrm{St}_{n,k}) < nk = \dim(U)$. Nevertheless, if ∇ is applied to vector fields which are tangent to $\mathrm{St}_{n,k}$ evaluated at points $X \in \mathrm{St}_{n,k}$, it yields the same result as ∇^{LC} by Proposition 6.8.

A similar expression for "Christoffel symbols" has already appeared in [10] for the so-called canonical metric as well as for the Euclidean metric, however, without relating them to the Christoffel symbols of the covariant derivative ∇ on U. Indeed, by exploiting Corollary 6.9, for $D = I_k$ and $\nu = 0$ we obtain

$$\Gamma_X(V, W) = \tfrac{1}{2} X(V^\top W + W^\top V)$$

reproducing Γ_e in [10, Sec. 2.2.3]. Analogously, setting $D = I_k$ and $\nu = -\frac{1}{2}$ in Corollary 6.9 yields

$$\Gamma_X(V, W) = \tfrac{1}{2}(VW^\top + WV^\top)X + \tfrac{1}{2}XV^\top(I_n - XX^\top)W + \tfrac{1}{2}XW^\top(I_n - XX^\top)V$$

for $X \in \mathrm{St}_{n,k}$ and $V, W \in T_X\mathrm{St}_{n,k}$. This expression coincides with [10, Eq. (2.49)].

7. Conclusions

We investigated a multi-parameter family of metrics on an open $U \subseteq \mathbb{R}^{n\times k}$ such that $\mathrm{St}_{n,k} \subseteq U$ becomes a pseudo-Riemannian submanifold. The corresponding geodesic equation for $\mathrm{St}_{n,k}$ as explicit matrix-valued second order ODE was derived by computing the metric spray on $T\mathrm{St}_{n,k}$. In principle, this approach to determine the geodesic equation is not limited to $\mathrm{St}_{n,k}$. It seems to be applicable to a pseudo-Riemannian submanifold of an open subset of a vector space as soon as the metric spray on the open subset and the tangent map of the orthogonal projection are known. Beside the geodesic equation, several other quantities related to the geometry of the pseudo-Riemannian submanifold $\mathrm{St}_{n,k} \subseteq U$ were determined in terms of explicit matrix-type formulas. In particular, the expressions for pseudo-Riemannian gradients and pseudo-Riemannian Hessians could pave the way for designing new optimization methods on $\mathrm{St}_{n,k}$. Moreover, we reproduced several well-known results from the literature putting them into a new perspective.

Acknowledgments

This work has been supported by the German Federal Ministry of Education and Research (BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA).

Conflict of interest

The author declares there is no conflict of interest.

References

1. K. A. Krakowski, L. Machado, F. Silva Leite, J. Batista, A modified Casteljau algorithm to solve interpolation problems on Stiefel manifolds, *J. Comput. Appl. Math.*, **311** (2017), 84–99. https://doi.org/10.1016/j.cam.2016.07.018
2. P. A. Absil, R. Mahony, R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton,Princeton University Press, 2008. https://doi.org/10.1515/9781400830244
3. T. Kanamori, A. Takeda, Non-convex Optimization on Stiefel Manifold and Applications to Machine Learning, in *Neural Information Processing*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, 109–116.

4. N. Trendafilov, M. Gallo, *Multivariate data analysis on matrix manifolds—(with Manopt)*, Springer Series in the Data Sciences, Springer, Cham, 2021. https://doi.org/10.1007/978-3-030-76974-1
5. Y. Ma, J. Kosecka, S. Sastry, Motion Estimation in Computer Vision: Optimization on Stiefel Manifolds, in *Proceedings of the 37th IEEE Conference on Decision and Control*, 1998, 3751–3756.
6. U. Helmke, K. Hüper, P. Lee, d J. Moore, Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold, *Int. J. Comput. Vision*, **74** (2007), 117–136. https://doi.org/10.1007/s11263-006-0005-0
7. N. Boumal, An introduction to optimization on smooth manifolds, To appear with Cambridge University Press, 2022. https://www.nicolasboumal.net/book
8. C. Udrişte, *Convex functions and optimization methods on Riemannian manifolds*, vol. 297 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1994.
9. S. T. Smith, *Geometric Optimization Methods for Adaptive Filtering*, PhD thesis, Harvard University, Cambridge, 1993.
10. A. Edelman, T. A. Arias, S. T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.*, **20** (1999), 303–353. https://doi.org/10.1137/S0895479895290954
11. B. Mishra, R. Sepulchre, Riemannian Preconditioning, *SIAM Journal on Optimization*, **26** (2016), 635–660. https://doi.org/10.1137/140970860
12. A. M. Bloch, P. E. Crouch, A. K. Sanyal, A variational problem on Stiefel manifolds, *Nonlinearity*, **19** (2006), 2247–2276. http://dx.doi.org/10.1088/0951-7715/19/10/002
13. K. Hüper, I. Markina, F. Silva Leite, A Lagrangian approach to extremal curves on Stiefel manifolds, *J. Geom. Mech.*, **13** (2021), 55–72. https://doi.org/10.3934/jgm.2020031
14. M. Schlarb, K. Hüper, Optimization on Stiefel Manifolds, in *CONTROLO 2022* (eds. L. Brito Palma, R. Neves-Silva and L. Gomes), Springer International Publishing, Cham, 2022, 363–374.
15. A. Arvanitoyeorgos, V. V. Dzhepko, Y. G. Nikonorov, Invariant Einstein metrics on some homogeneous spaces of classical Lie groups, *Canad. J. Math.*, **61** (2009), 1201–1213. https://doi.org/10.4153/CJM-2009-056-2
16. J. E. Marsden, T. S. Ratiu, *Introduction to mechanics and symmetry*, vol. 17 of Texts in Applied Mathematics, 2nd edition, Springer-Verlag, New York, 1999. https://doi.org/10.1007/978-0-387-21792-5
17. B. O'Neill, *Semi-Riemannian geometry*, vol. 103 of Pure and Applied Mathematics, Academic Press, Inc., New York, 1983.
18. J. Gallier, J. Quaintance, *Differential geometry and Lie groups—a computational perspective*, vol. 12 of Geometry and Computing, Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-46040-2
19. U. Helmke, J. B. Moore, *Optimization and dynamical systems*, Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1994. https://doi.org/10.1007/978-1-4471-3467-1

20. P. W. Michor, *Topics in differential geometry*, vol. 93 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2008. https://doi.org/10.1090/gsm/093

21. J. M. Lee, *Introduction to smooth manifolds*, vol. 218 of Graduate Texts in Mathematics, 2nd edition, Springer, New York, 2013. https://doi.org/10.1007/978-1-4419-9982-5

22. S. Lang, *Differential and Riemannian manifolds*, vol. 160 of Graduate Texts in Mathematics, 3rd edition, Springer-Verlag, New York, 1995. https://doi.org/10.1007/978-1-4612-4182-9

23. J. Marsden, On Completeness of Homogeneous Pseudo-Riemannian Manifolds, *Indiana. Univ. Math. J.*, **22** (1972), 1065–1066. https://doi.org/10.1512/iumj.1973.22.22089

24. Y. N. Fedorov, B. Jovanović, Geodesic flows and Neumann systems on Stiefel varieties: geometry and integrability, *Math. Z.*, **270** (2012), 659–698. https://doi.org/10.1007/s00209-010-0818-y

25. D. Nguyen, Operator-valued formulas for Riemannian Gradient and Hessian and families of tractable metrics, 2021. https://arxiv.org/abs/2009.10159

26. J. Gallier, J. Quaintance, *Differential geometry and Lie groups—a second course*, vol. 13 of Geometry and Computing, Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-46047-1

27. J. M. Lee, *Introduction to Riemannian manifolds*, vol. 176 of Graduate Texts in Mathematics, 2nd edition, Springer, Cham, 2018. https://doi.org/10.1007/978-3-319-91755-9

28. D. Nguyen, Riemannian geometry with differentiable ambient space and metric operator, 2021. https://arxiv.org/abs/2105.01583