Found Comput Math (2018) 18:757–788
https://doi.org/10.1007/s10208-017-9353-0

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Interpolation on Symmetric Spaces Via the Generalized Polar Decomposition

Evan S. Gawlik[1] · **Melvin Leok**[1]

Received: 21 May 2016 / Revised: 29 March 2017 / Accepted: 4 May 2017 /
Published online: 26 May 2017
© SFoCM 2017

Abstract We construct interpolation operators for functions taking values in a symmetric space—a smooth manifold with an inversion symmetry about every point. Key to our construction is the observation that every symmetric space can be realized as a homogeneous space whose cosets have canonical representatives by virtue of the generalized polar decomposition—a generalization of the well-known factorization of a real nonsingular matrix into the product of a symmetric positive-definite matrix times an orthogonal matrix. By interpolating these canonical coset representatives, we derive a family of structure-preserving interpolation operators for symmetric space-valued functions. As applications, we construct interpolation operators for the space of Lorentzian metrics, the space of symmetric positive-definite matrices, and the Grassmannian. In the case of Lorentzian metrics, our interpolation operators provide a family of finite elements for numerical relativity that are frame-invariant and have signature which is guaranteed to be Lorentzian pointwise. We illustrate their potential utility by interpolating the Schwarzschild metric numerically.

Keywords Interpolation · Manifold-valued data · Symmetric space · Generalized polar decomposition · Grassmannian · Lorentzian metric · Lie triple system · Geodesic finite element · Karcher mean · Log-Euclidean mean

Communicated by Arieh Iserles.

✉ Evan S. Gawlik
egawlik@ucsd.edu

Melvin Leok
mleok@math.ucsd.edu

1 Department of Mathematics, University of California, San Diego,
9500 Gilman Drive #0112, La Jolla, CA 92093-0112, USA

Mathematics Subject Classification Primary 65D05 · 53C35; Secondary 65N30 · 58J70 · 53B30

1 Introduction

Manifold-valued data and manifold-valued functions play an important role in a wide variety of applications, including mechanics [14,24,42], computer vision and graphics [11,13,15,18–20,30,32,47], medical imaging [7], and numerical relativity [5]. By their very nature, such applications demand that care be taken when performing computations that would otherwise be routine, such as averaging, interpolation, extrapolation, and the numerical solution of differential equations. This paper constructs interpolation and averaging operators for functions taking values in a *symmetric space*—a smooth manifold with an inversion symmetry about every point. Key to our construction is the observation that every symmetric space can be realized as a homogeneous space whose cosets have canonical representatives by virtue of the *generalized polar decomposition*—a generalization of the well-known factorization of a real nonsingular matrix into the product of a symmetric positive-definite matrix times an orthogonal matrix. By interpolating these canonical coset representatives, we derive a family of structure-preserving interpolation operators for symmetric space-valued functions.

Our motivation for constructing such operators is best illustrated by example. Among the most interesting scenarios in which symmetric space-valued functions play a role is numerical relativity. There, the dependent variable in Einstein's equations—the metric tensor—is a function taking values in the space $\mathcal{L}$ of Lorentzian metrics: symmetric, nondegenerate 2-tensors with signature $(3, 1)$. This space is neither a vector space nor a convex set. Rather, it has the structure of a symmetric space. As a consequence, the outputs of basic arithmetic operations on Lorentzian metrics such as averaging, interpolation, and extrapolation need not remain in $\mathcal{L}$. This is undesirable for several reasons. If the metric tensor field is to be discretized with finite elements, then a naive approach in which the components of the metric are discretized with piecewise polynomials may fail to produce a metric field with signature (3, 1) at all points in space-time. Perhaps an even more problematic possibility is that a numerical time integrator used to advance the metric forward in time (e.g., in a $3+1$ formulation of Einstein's equations) might produce metrics with invalid signature. One of the aims of the present paper is to avert these potential dangers altogether by constructing a structure-preserving interpolation operator for Lorentzian metrics. As will be shown, the interpolation operator we derive not only produces interpolants that everywhere belong to $\mathcal{L}$, but it is also frame-invariant: the interpolation operator we derive commutes with the action of the indefinite orthogonal group $O(1, 3)$ on $\mathcal{L}$. Furthermore, our interpolation operator commutes with inversion and interpolates the determinant of the metric tensor in a monotonic manner.

A more subtle example is the space $SPD(n)$ of symmetric positive-definite $n \times n$ matrices. This space forms a convex cone, so arithmetic averaging and linear interpolation trivially produce $SPD(n)$-valued results. Nevertheless, these operations fail to preserve other structures that are important in some applications. For instance, arithmetic averaging does not commute with matrix inversion, and the determinant of the arithmetic average need not be less than or equal to the maximum of the determinants of the data. This may remedied by considering instead the Riemannian mean (also known as the Karcher mean) of symmetric positive-definite matrices with respect to the canonical left-invariant Riemannian metric on $SPD(n)$ [9,33,38]. The Riemannian mean cannot, in general, be expressed in closed form, but it can be computed iteratively and possesses a number of structure-preserving properties; see [9] for details. A less computationally expensive alternative, introduced by Arsigny et al. [6], is to compute the mean of symmetric positive-definite matrices with respect to a log-Euclidean metric on $SPD(n)$. The resulting averaging operator commutes with matrix inversion, prevents overestimation of the determinant, and commutes with similarity transformations that consist of an isometry plus scaling. Both of these constructions turn out to be special cases of the general theory presented in this paper. In our derivation of the log-Euclidean mean, we give a clear geometric explanation of the vector space structure with which Arsigny et al. [6] endow $SPD(n)$ in their derivation, which turns out to be nothing more than a correspondence between a symmetric space ($SPD(n)$) and a Lie triple system [25].

Another symmetric space which we address in this paper is the Grassmannian $Gr(p,n)$, which consists of all p-dimensional linear subspaces of $\mathbb{R}^n$. Interpolation on the Grassmannian is a task of importance in a variety of contexts, including reduced-order modeling [4,48] and computer vision [13,19,30,47]. Not surprisingly, this task has received much attention in the literature; see [2,8] and the references therein. Our constructions in this paper recover some of the well-known interpolation schemes on the Grassmannian, including those that appear in [4,8,13].

There are connections between the present work and geodesic finite elements [21, 22,43,44], a family of conforming finite elements for functions taking values in a Riemannian manifold M. In fact, we recover such elements as a special case; see Sect. 3.3. Since their evaluation amounts to the computation of a weighted Riemannian mean, geodesic finite elements and their derivatives can sometimes be expensive to compute. One of the messages we hope to convey is that when M is a symmetric space, this additional structure enables the construction of alternative interpolants that are less expensive to compute but still possess many of the desirable features of geodesic finite elements.

Our use of the generalized polar decomposition in this paper is inspired by a stream of research [31,40,41] that has, in recent years, cast a spotlight on the generalized polar decomposition's role in numerical analysis. Much of our exposition and notation parallels that which appears in those papers, and we encourage the reader to look there for further insight.

Some of the key contributions of this paper include the following. First, the paper unifies several seemingly disparate interpolation strategies, some of which have been derived in an ad hoc way in the literature. The paper also unveils the geometric underpinnings of these interpolants' structure-preserving properties. These structure-preserving properties are unique to symmetric spaces and lead to important practical consequences, including frame-invariance in the context of Lorentzian metric interpolation. On the practical side, the paper also shows that symmetric spaces admit efficiently computable interpolants. This is significant, since on a general Riemannian manifold M, for instance, it is a simple matter to write down interpolation schemes for M-valued data using the Riemannian exponential map and its inverse, but it is generally not the case that the exponential map can be calculated explicitly, much less inverted. We show that for a symmetric space $\mathcal{S}$, these tasks are tractable. Our use of the generalized polar decomposition plays a key role here, since it reveals not only how to construct a map from a linear space to $\mathcal{S}$, but also how to systematically invert it, a task which would otherwise be nontrivial except in special cases. We also derive formulas for the first and second derivatives of the resulting interpolants. Finally, to our knowledge, the paper introduces the first structure-preserving finite elements for Lorentzian metrics in numerical relativity.

Organization. This paper is organized as follows. We begin in Sect. 2 by reviewing symmetric spaces, Lie triple systems, and the generalized polar decomposition. Then, in Sect. 3, we exploit a correspondence between symmetric spaces and Lie triple systems to construct interpolation operators on symmetric spaces. Finally, in Sect. 4, we specialize these interpolation operators to three examples of symmetric spaces: the space of symmetric positive-definite matrices, the space of Lorentzian metrics, and the Grassmannian. In the case of Lorentzian metrics, we illustrate the potential utility of these interpolation operators by interpolating the Schwarzschild metric numerically.

2 Symmetric Spaces and the Generalized Polar Decomposition

In this section, we review symmetric spaces, Lie triple systems, and the generalized polar decomposition. We describe a well-known correspondence between symmetric spaces and Lie triple systems that will serve in Sect. 3 as a foundation for interpolating functions which take values in a symmetric space. For further background material, we refer the reader to [25,40,41,49].

2.1 Notation and Definitions

Let G be a Lie group and let $\sigma : G \to G$ be an involutive automorphism. That is, $\sigma \neq \text{id.}$ is a bijection satisfying $\sigma(\sigma(g)) = g$ and $\sigma(gh) = \sigma(g)\sigma(h)$ for every $g, h \in G$. Denote by G^σ the subgroup of G consisting of fixed points of σ:

$$G^\sigma = \{g \in G \mid \sigma(g) = g\}.$$

Suppose that G acts transitively on a smooth manifold $\mathcal{S}$ with a distinguished element $\eta \in \mathcal{S}$ whose stabilizer coincides with G^σ. In other words,

$$g \cdot \eta = \eta \iff \sigma(g) = g,$$

where $g \cdot u$ denotes the action of $g \in G$ on an element $u \in \mathcal{S}$. Then there is a bijective correspondence between elements of the homogeneous space G/G^σ and elements of $\mathcal{S}$. On the other hand, the cosets in G/G^σ have canonical representatives by virtue of the *generalized polar decomposition* [40,41]. This decomposition states that any $g \in G$ sufficiently close to the identity $e \in G$ can be written as a product

$$g = pk, \quad p \in G_\sigma,\ k \in G^\sigma, \tag{1}$$

where

$$G_\sigma = \{g \in G \mid \sigma(g) = g^{-1}\}.$$

Moreover, this decomposition is unique in the neighborhood of e on which it exists [41, Theorem 3.1]. As a consequence, there is a bijection between a neighborhood in G_σ of the identity $e \in G_\sigma$ and a neighborhood of the coset $[e] \in G/G^\sigma$. The space G_σ—which, unlike G^σ, is not a subgroup of G—is a symmetric space which is closed under a nonassociative symmetric product $g \cdot h = gh^{-1}g$. Its tangent space at the identity is the space

$$\mathfrak{p} = \{Z \in \mathfrak{g} \mid d\sigma(Z) = -Z\}.$$

Here, $\mathfrak{g}$ denotes the Lie algebra of G, and $d\sigma : \mathfrak{g} \to \mathfrak{g}$ denotes the differential of σ at e, which can be expressed in terms of the Lie group exponential map $\exp : \mathfrak{g} \to G$ via

$$d\sigma(Z) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \sigma(\exp(tZ)).$$

The space $\mathfrak{p}$, which is not a Lie subalgebra of $\mathfrak{g}$, has the structure of a Lie triple system: it is a vector space closed under the double commutator $[\cdot,[\cdot,\cdot]]$. In contrast, the space

$$\mathfrak{k} = \{Z \in \mathfrak{g} \mid d\sigma(Z) = Z\}$$

is a subalgebra of $\mathfrak{g}$, as it is closed under the commutator $[\cdot,\cdot]$. This subalgebra is none other than the Lie algebra of G^σ. The generalized polar decomposition (1) has a manifestation at the Lie algebra level called the *generalized Cartan decomposition*, which decomposes $\mathfrak{g}$ as a direct sum

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}. \tag{2}$$

All of these observations lead to the conclusion that the following diagram commutes:

$$\begin{tikzcd}
\mathfrak{k} \arrow[r, "\exp"] \arrow[d, "\iota"'] & G^\sigma \arrow[dr, "\iota"] & & \\
\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k} \arrow[rr, "\exp"] & & G \arrow[d, "\pi"'] \arrow[dr, "\varphi"] & \\
\mathfrak{p} \arrow[u, "\iota"] \arrow[r, "\exp"'] & G_\sigma \arrow[ur, "\iota"] \arrow[r, "\psi"'] & G/G^\sigma \arrow[r, "\bar{\varphi}"'] & \mathcal{S}
\end{tikzcd}$$

In this diagram, we have used the letter ι to denote the canonical inclusion, $\pi : G \to G/G^\sigma$ the canonical projection, and $\varphi : G \to \mathcal{S}$ the map $\varphi(g) = g \cdot \eta$. The maps ψ and $\bar{\varphi}$ are defined by the condition that the diagram be commutative.

2.2 Correspondence Between Symmetric Spaces and Lie Triple Systems

An important feature of the diagram above is that the maps along its bottom row—when restricted to suitable neighborhoods of the neutral elements $0 \in \mathfrak{p}$, $e \in G_\sigma$, $[e] \in G/G^\sigma$, and the distinguished element $\eta \in \mathcal{S}$—are diffeomorphisms [25, p. 104, p. 124, p. 253]. In particular, the composition

$$F = \bar{\varphi} \circ \psi \circ \exp \tag{3}$$

(or, equivalently, $F = \varphi \circ \iota \circ \exp$) provides a diffeomorphism from a neighborhood of $0 \in \mathfrak{p}$ to a neighborhood of $\eta \in \mathcal{S}$, given by

$$F(P) = \exp(P) \cdot \eta$$

for $P \in \mathfrak{p}$. The space $\mathfrak{p}$, being a vector space, offers a convenient space to perform computations (such as averaging, interpolation, extrapolation, and the numerical solution of differential equations) that might otherwise be unwieldy on the space $\mathcal{S}$. This is analogous to the situation that arises when working with the Lie group G. Often, computations on G are more easily performed by mapping elements of G to the Lie algebra $\mathfrak{g}$ via the inverse of the exponential map (or an approximation thereof), performing computations in $\mathfrak{g}$, and mapping the result back to G via the exponential map (or an approximation thereof).

We remark that the analogy just drawn between computing on Lie groups and computing on symmetric spaces is in fact more than a mere resemblance; the latter situation directly generalizes the former. Indeed, any Lie group G can be realized as a symmetric space by considering the action of $G \times G$ on G given by $(g, h) \cdot k = gkh^{-1}$. The stabilizer of $e \in G$ is the diagonal of $G \times G$, which is precisely the subgroup fixed by the involution $\sigma(g, h) = (h, g)$. In this setting, one finds that the map (3) takes $(X, -X) \in \mathfrak{g} \times \mathfrak{g}$ to $\exp(2X) \in G$. This shows that, up to a trivial modification, the map (3) reduces to the Lie group exponential map if $\mathcal{S}$ happens to be a Lie group.

An additional feature of the map (3) is its equivariance with respect to the action of G^σ on $\mathcal{S}$ and $\mathfrak{p}$. Specifically, for $g \in G$, let $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ denote the adjoint action of G on $\mathfrak{g}$:

$$\mathrm{Ad}_g Z = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g \exp(tZ) g^{-1}.$$

In a slight abuse of notation, we will write

$$\mathrm{Ad}_g Z = gZg^{-1}$$

in this paper, bearing in mind that the above equality holds in the sense of matrix multiplication for any matrix group. The following lemma shows that $F \circ \mathrm{Ad}_g\big|_{\mathfrak{p}} = g \cdot F$ for every $g \in G^\sigma$. Note that this statement makes implicit use of the (easily verifiable) fact that Ad_g leaves $\mathfrak{p}$ invariant when $g \in G^\sigma$; that is, $gPg^{-1} \in \mathfrak{p}$ for every $g \in G^\sigma$ and every $P \in \mathfrak{p}$.

Lemma 2.1 *For every $P \in \mathfrak{p}$ and every $g \in G^\sigma$,*

$$g \cdot F(P) = F(gPg^{-1}).$$

Proof Note that $g \in G^\sigma$ implies $g^{-1} \in G^\sigma$, so $g^{-1} \cdot \eta = \eta$. Hence, since $\exp(gPg^{-1}) = g \exp(P) g^{-1}$,

$$\begin{aligned} F(gPg^{-1}) &= \exp(gPg^{-1}) \cdot \eta \\ &= g \exp(P) g^{-1} \cdot \eta \\ &= g \exp(P) \cdot \eta \\ &= g \cdot F(P) \end{aligned}$$

□

We finish this section by remarking that σ induces a family of symmetries on $\mathcal{S}$ as follows. Define $s_\eta : \mathcal{S} \to \mathcal{S}$ by setting

$$s_\eta(g \cdot \eta) = \sigma(g) \cdot \eta$$

for each $g \in G$. Note that s_η is well defined, fixes η, and has differential equal to minus the identity. Furthermore, by definition, the following diagram commutes:

$$\begin{tikzcd} \mathfrak{p} \arrow[r, "\exp"] \arrow[d, "d\sigma"'] & G \arrow[r, "\varphi"] \arrow[d, "\sigma"'] & \mathcal{S} \arrow[d, "s_\eta"] \\ \mathfrak{p} \arrow[r, "\exp"'] & G \arrow[r, "\varphi"'] & \mathcal{S} \end{tikzcd}$$

Written another way,

$$s_\eta(F(P)) = F(-P) \tag{4}$$

for every $P \in \mathfrak{p}$. In a similar manner, a symmetry at each point $h \cdot \eta \in \mathcal{S}$ can be defined via

$$s_{h\cdot\eta}(g \cdot \eta) = h \cdot s_\eta(h^{-1} g \cdot \eta) = h\sigma(h^{-1} g) \cdot \eta.$$

If $\mathcal{S}$ admits a G-invariant Riemannian metric, then the maps F and $s_{h\cdot\eta}$ have particularly notable interpretations. Any such metric induces a canonical connection on $\mathcal{S}$ [35, Thoerem 3.3]. With respect to this connection, F may be identified with the

Riemannian exponential map $\mathrm{Exp}_\eta : T_\eta\mathcal{S} \to \mathcal{S}$ upon identifying $\mathfrak{p}$ with $T_\eta\mathcal{S}$ via $\mathfrak{p} \cong \mathfrak{g}/\mathfrak{k} = T_{[e]}(G/G^\sigma) \cong T_\eta\mathcal{S}$ [35, Theorem 3.2(3)]. In addition, the map $s_{h\cdot\eta}$ is an isometry that sends $\mathrm{Exp}_{h\cdot\eta}(X)$ to $\mathrm{Exp}_{h\cdot\eta}(-X)$ for every $X \in T_{h\cdot\eta}\mathcal{S}$ [35, p. 231]. As an important special case, note that $\mathcal{S}$ admits a G-invariant Riemannian metric whenever G^σ is compact [35, p. 245]. Examples of symmetric spaces that do not admit G-invariant Riemannian metrics include the space of symmetric 4×4 matrices with signature (3, 1) (see Sect. 4.1.2) and the affine Grassmannian manifold consisting of p-dimensional affine subspaces of $\mathbb{R}^n$ [49, Section 7.5].

2.3 Generalizations

The construction above can be generalized by replacing the exponential map in (3) with a different local diffeomorphism. One example is given by fixing an element $\bar{g} \in G$ and replacing $\exp : \mathfrak{p} \to G_\sigma$ in (3) with the map

$$P \mapsto \psi^{-1}([\bar{g}\exp(P)]) = \psi^{-1}(\pi(\bar{g}\exp(P))). \tag{5}$$

The output of this map is nothing more than the factor p in the generalized polar decomposition $\bar{g}\exp(P) = pk$, $p \in G_\sigma$, $k \in G^\sigma$. The map (3) then becomes

$$F_{\bar{g}}(P) = \bar{g}\exp(P) \cdot \eta. \tag{6}$$

This generalization of (3) has the property that it provides a diffeomorphism between a neighborhood of $0 \in \mathfrak{p}$ and a neighborhood of $\bar{g} \cdot \eta \in \mathcal{S}$ rather than η. Note that when $\bar{g} = e$ (the identity element), this map coincides with (3). A calculation similar to the proof of Lemma 2.1 shows that the map $(\bar{g}, P) \mapsto F_{\bar{g}}(P)$ is G^σ-equivariant, in the sense that

$$F_{h\bar{g}h^{-1}}(hPh^{-1}) = h \cdot F_{\bar{g}}(P) \tag{7}$$

for every $h \in G^\sigma$ and every $P \in \mathfrak{p}$. Furthermore,

$$s_{\bar{g}\cdot\eta}(F_{\bar{g}}(P)) = F_{\bar{g}}(-P) \tag{8}$$

for every $P \in \mathfrak{p}$. These identities are summarized in the following pair of diagrams, the first of which commutes for every $h \in G^\sigma$, and the second of which commutes for every $\bar{g} \in G$.

$$\begin{tikzcd} G \times \mathfrak{p} \arrow[r, "f"] \arrow[d, "\Psi_h"'] & \mathcal{S} \arrow[d, "\Phi_h"] \\ G \times \mathfrak{p} \arrow[r, "f"'] & \mathcal{S} \end{tikzcd} \qquad \begin{tikzcd} \mathfrak{p} \arrow[r, "F_{\bar{g}}"] \arrow[d, "d\sigma"'] & \mathcal{S} \arrow[d, "s_{\bar{g}\cdot\eta}"] \\ \mathfrak{p} \arrow[r, "F_{\bar{g}}"'] & \mathcal{S} \end{tikzcd}$$

Here, we have denoted $f(\bar{g}, P) = F_{\bar{g}}(P)$, $\Psi_h(\bar{g}, P) = (h\bar{g}h^{-1}, hPh^{-1})$, and $\Phi_h(u) = h \cdot u$.

More generally, one may consider replacing the exponential map in (5) with any retraction $R : \mathfrak{g} \to G$ [1, p. 55]. For instance, if G is a quadratic matrix group, one may choose R equal to the Cayley transform, or more generally, any diagonal Padé approximant of the matrix exponential [12].

3 Interpolation on Symmetric Spaces

In this section, we exploit the correspondence between symmetric spaces and Lie triple systems discussed in Sects. 2.2–2.3 in order to interpolate functions which take values in a symmetric space.

3.1 A Structure-Preserving Interpolant

Consider the task of interpolating m elements $u_1, u_2, \ldots, u_m \in \mathcal{S}$. To facilitate the exposition, we will think of these elements as the values of a smooth function $u : \Omega \to \mathcal{S}$ defined on a domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, at locations $x^{(1)}, x^{(2)}, \ldots, x^{(m)} \in \Omega$, although this point of view is not essential in what follows. Our goal is thus to construct a function $\mathcal{I}u : \Omega \to \mathcal{S}$ that satisfies $\mathcal{I}u(x^{(i)}) = u_i$, $i = 1, 2, \ldots, m$, and has a desired level of regularity (e.g., continuity). We assume that for each $x \in \Omega$, $u(x)$ belongs to the range of map (3). We may then interpolate $u_1, u_2, \ldots, u_m$ by interpolating $F^{-1}(u_1), F^{-1}(u_2), \ldots, F^{-1}(u_m) \in \mathfrak{p}$ and mapping the result back to $\mathcal{S}$ via F. More precisely, set

$$\mathcal{I}u(x) = F(\hat{\mathcal{I}}P(x)), \tag{9}$$

where $P(x) = F^{-1}(u(x))$ and $\hat{\mathcal{I}}P : \Omega \to \mathfrak{p}$ is an interpolant of $F^{-1}(u_1)$, $F^{-1}(u_2)$, $\ldots$, $F^{-1}(u_m)$. Then $\mathcal{I}u$ interpolates the data while fulfilling the following important properties.

Proposition 3.1 *Suppose that $\hat{\mathcal{I}}$ commutes with Ad_g for every $g \in G^\sigma$. That is,*

$$\hat{\mathcal{I}}(gPg^{-1})(x) = g\hat{\mathcal{I}}P(x)g^{-1}$$

for every $x \in \Omega$ and every $g \in G^\sigma$. Then $\mathcal{I}$ is G^σ-equivariant. That is,

$$\mathcal{I}(g \cdot u)(x) = g \cdot \mathcal{I}u(x) \tag{10}$$

for every $x \in \Omega$ and every $g \in G^\sigma$ sufficiently close to the identity.

Proof The claim is a straightforward consequence of Lemma 2.1. □

Note that g must be sufficiently close to the identity in (10) to ensure that $g \cdot u_i$ belongs to the range of the map (3) for each $i = 1, 2, \ldots, m$.

Proposition 3.2 *Suppose that $\hat{\mathcal{I}}$ commutes with $d\sigma\big|_{\mathfrak{p}}$. That is,*

$$\hat{\mathcal{I}}(-P)(x) = -\hat{\mathcal{I}}P(x)$$

for every $x \in \Omega$. Then $\mathcal{I}$ commutes with s_η. That is,

$$\mathcal{I}(s_\eta(u))(x) = s_\eta(\mathcal{I}u(x))$$

for every $x \in \Omega$.

Proof The claim is a straightforward consequence of (4). □

The preceding propositions apply, in particular, to any interpolant $\hat{\mathcal{I}}P : \Omega \to \mathfrak{p}$ of the form

$$\hat{\mathcal{I}}P(x) = \sum_{i=1}^{m} \phi_i(x) P(x^{(i)})$$

with scalar-valued shape functions $\phi_i : \Omega \to \mathbb{R}$, $i = 1, 2, \dots, m$, satisfying $\phi_i(x^{(j)}) = \delta_{ij}$, where δ_{ij} denotes the Kronecker delta. By the propositions above, such an interpolant gives rise to a G^σ-equivariant interpolant $\mathcal{I}u : \Omega \to \mathcal{S}$ that commutes with s_η, given by

$$\mathcal{I}u(x) = F\left(\sum_{i=1}^{m} \phi_i(x) F^{-1}(u_i)\right). \tag{11}$$

Written more explicitly,

$$\mathcal{I}u(x) = \exp(\hat{\mathcal{I}}P(x)) \cdot \eta, \tag{12}$$

where

$$\hat{\mathcal{I}}P(x) = \sum_{i=1}^{m} \phi_i(x) F^{-1}(u_i). \tag{13}$$

Note that the interpolation strategy above resembles the ones used in, for instance, [6,16,23,50].

3.2 Derivatives of the Interpolant

The relations (12–13) lead to an explicit formula for the derivatives of $\mathcal{I}u(x)$ with respect to each of the coordinate directions x_j, $j = 1, 2, \dots, d$. Namely,

$$\frac{\partial \mathcal{I}u}{\partial x_j}(x) = \mathrm{dexp}_{\hat{\mathcal{I}}P(x)} \frac{\partial \hat{\mathcal{I}}P}{\partial x_j}(x) \cdot \eta, \tag{14}$$

where

$$\frac{\partial \hat{\mathcal{I}} P}{\partial x_j}(x) = \sum_{i=1}^{m} \frac{\partial \phi_i}{\partial x_j}(x) F^{-1}(u_i)$$

and $\mathrm{dexp}_X Y$ denotes the differential of exp at $X \in \mathfrak{g}$ in the direction $Y \in \mathfrak{g}$.

An explicit formula for $\mathrm{dexp}_X Y$ is the series

$$\mathrm{dexp}_X Y = \exp(X) \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} \mathrm{ad}_X^k Y,$$

where $\mathrm{ad}_X Y = [X, Y]$ denotes the adjoint action of $\mathfrak{g}$ on itself [29, p. 55]. In practice, one may truncate this series to numerically approximate $\mathrm{dexp}_X Y$. Note that while the exact value of $\mathrm{dexp}_X Y$ belongs to $\mathfrak{p}$ whenever $X, Y \in \mathfrak{p}$, this need not be true of its truncated approximation. However, this is of little import since any spurious $\mathfrak{k}$-components in such a truncation act trivially on η in (14).

While the series expansion of $\mathrm{dexp}_X Y$ is valid on any finite-dimensional Lie group, more efficient methods are available for the computation of $\mathrm{dexp}_X Y$ when G is a matrix group. Arguably, the simplest is to make use of the identity [26,37, p. 253]

$$\exp \begin{pmatrix} X & Y \\ 0 & X \end{pmatrix} = \begin{pmatrix} \exp(X) & \mathrm{dexp}_X Y \\ 0 & \exp(X) \end{pmatrix}. \tag{15}$$

More sophisticated approaches with better numerical properties can be found in [3,26, pp. 253–259].

The identity (15) can be leveraged to derive formulas for higher-order derivatives of $\mathcal{I}u(x)$, provided of course that G is a matrix group. As shown in Appendix A, we have

$$\frac{\partial^2 \mathcal{I} u}{\partial x_j \partial x_k}(x) = A \cdot \eta \tag{16}$$

for each $j, k = 1, 2, \ldots, d$, where A denotes the $(1, 4)$ block of the matrix

$$\exp \begin{pmatrix} X & Y & Z & W \\ 0 & X & 0 & Z \\ 0 & 0 & X & Y \\ 0 & 0 & 0 & X \end{pmatrix},$$

and $X = \hat{\mathcal{I}} P(x)$, $Y = \frac{\partial \hat{\mathcal{I}} P}{\partial x_j}(x)$, $Z = \frac{\partial \hat{\mathcal{I}} P}{\partial x_k}(x)$, and $W = \frac{\partial^2 \hat{\mathcal{I}} P}{\partial x_j \partial x_k}(x)$.

3.3 Generalizations

More generally, by fixing an element $\bar{g} \in G$ and adopting map (6) instead of F, we obtain interpolation schemes of the form

$$\mathcal{I}_{\bar{g}} u(x) = F_{\bar{g}} \left(\sum_{i=1}^{m} \phi_i(x) F_{\bar{g}}^{-1}(u_i) \right) = \bar{g} \exp \left(\sum_{i=1}^{m} \phi_i(x) F_{\bar{g}}^{-1}(u_i) \right) \cdot \eta. \tag{17}$$

Here, we must of course assume that u_i belongs to the range of $F_{\bar{g}}$ for each $i = 1, 2, \dots, m$. This interpolant is therefore suitable for interpolating elements of $\mathcal{S}$ in a neighborhood of $\bar{g} \cdot \eta$. Using the fact that $F_{hg}(P) = h \cdot F_g(P)$ for every $h, g \in G$ and every $P \in \mathfrak{p}$, one finds that this interpolant is equivariant under the action of the full group G, in the sense that

$$\mathcal{I}_{h\bar{g}}(h \cdot u)(x) = h \cdot \mathcal{I}_{\bar{g}} u(x) \tag{18}$$

for every $x \in \Omega$ and every $h \in G$ sufficiently close to the identity. On the other hand, the equivariance of $F_{\bar{g}}$ under the action of the subgroup G^σ [recall (7)] implies that

$$\mathcal{I}_{h\bar{g}h^{-1}}(h \cdot u)(x) = h \cdot \mathcal{I}_{\bar{g}} u(x) \tag{19}$$

for every $x \in \Omega$ and every $h \in G^\sigma$ sufficiently close to the identity. Comparing (18) with (19) leads to the conclusion that this interpolant is invariant under post-multiplication of $\bar{g}$ by elements of G^σ; that is,

$$\mathcal{I}_{\bar{g}h} u(x) = \mathcal{I}_{\bar{g}} u(x) \tag{20}$$

for every $x \in \Omega$ and every $h \in G^\sigma$ sufficiently close to the identity. Finally, as a consequence of (8),

$$\mathcal{I}_{\bar{g}}(s_{\bar{g}\cdot\eta}(u))(x) = s_{\bar{g}\cdot\eta}(\mathcal{I}_{\bar{g}} u(x))$$

for every $x \in \Omega$.

A natural choice for $\bar{g}$ is not immediately evident, but one heuristic is to select $j \in \{1, 2, \dots, m\}$ and set $\bar{g}$ equal to a representative of the coset $\bar{\varphi}^{-1}(u_j)$. A more interesting option is to allow $\bar{g}$ to vary with x and to define $\bar{g}(x)$ implicitly via

$$\bar{g}(x) \cdot \eta = \mathcal{I}_{\bar{g}(x)} u(x). \tag{21}$$

Equivalently,

$$\sum_{i=1}^{m} \phi_i(x) F_{\bar{g}(x)}^{-1}(u_i) = 0. \tag{22}$$

A method for computing the interpolant $\mathcal{I}_{\bar{g}(x)} u(x)$ numerically is self-evident. Namely, one performs the fixed-point iteration suggested by (21), as we explain in greater detail in Sect. 4.

In what follows, we show that if G^σ is compact, so that $\mathcal{S}$ admits a G-invariant Riemannian metric, then (22) characterizes $\bar{g}(x) \cdot \eta \in \mathcal{S}$ as the weighted Riemannian mean of $u_1, u_2, \dots, u_m$. Recall that in this setting, the map $F : \mathfrak{p} \to \mathcal{S}$ sending P to

$\exp(P) \cdot \eta$ may be identified with the Riemannian exponential map $\mathrm{Exp}_\eta : T_\eta\mathcal{S} \to \mathcal{S}$ upon identifying $\mathfrak{p}$ with $T_\eta\mathcal{S}$.

Lemma 3.3 *Suppose that G^σ is compact, so that $\mathcal{S}$ admits a G-invariant Riemannian metric. If $\bar{g}(x) \in G$ is a solution of (21) [or, equivalently, (22)], then $\bar{g}(x) \cdot \eta \in \mathcal{S}$ locally minimizes*

$$\sum_{i=1}^{m} \phi_i(x)\, \mathrm{dist}\, (\bar{g}(x) \cdot \eta, u_i)^2 \tag{23}$$

among elements of $\mathcal{S}$, where $\mathrm{dist} : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ *denotes the geodesic distance on $\mathcal{S}$.*

Proof For each $i = 1, 2, \ldots, m$, let $P_i = F^{-1}_{\bar{g}(x)}(u_i)$, so that $\bar{g}(x)\exp(P_i) \cdot \eta = u_i$. Since the metric on $\mathcal{S}$ is G-invariant, the identity $\exp(P_i) \cdot \eta = \mathrm{Exp}_\eta(P_i)$ implies that $u_i = \bar{g}(x) \cdot \mathrm{Exp}_\eta(P_i) = \mathrm{Exp}_{\bar{g}(x)\cdot\eta}(\bar{g}(x)P_i)$. Equivalently, $P_i = \bar{g}(x)^{-1}\mathrm{Exp}^{-1}_{\bar{g}(x)\cdot\eta} u_i$. This shows that (22) is equivalent to

$$\sum_{i=1}^{m} \phi_i(x) \mathrm{Exp}^{-1}_{\bar{g}(x)\cdot\eta} u_i = 0.$$

The latter equation is precisely the equation which characterizes minimizers of (23); see [33, Theorem 1.2]. □

Notice that minimizers of (23) are precisely geodesic finite elements on G/G^σ, as described in [21,22,43,44]. We refer the reader to those articles for further information about the approximation properties of these interpolants, as well as the convergence properties of iterative algorithms used to compute them.

3.4 Interpolation Error Estimates

Error estimates for interpolants of the form (9) can be derived by appealing to the smoothness of the map $F : \mathfrak{p} \to \mathcal{S}$ and the approximation properties of $\hat{\mathcal{I}}P$. Roughly speaking, the interpolant (9) inherits the approximation properties enjoyed by $\hat{\mathcal{I}}P$ under mild assumptions. To see this, consider the setting in which $\mathcal{S}$ is embedded in $\mathbb{R}^n$ for some $n \geq 1$. Denote by $Du \in \mathbb{R}^{n\times d}$ and $DF \in \mathbb{R}^{n\times n}$ the matrices of partial derivatives of u and F, viewed as maps from $\Omega \subset \mathbb{R}^d$ and $\mathfrak{p} \simeq \mathbb{R}^n$, respectively, to $\mathbb{R}^n$. Define $DP \in \mathbb{R}^{n\times d}$, $D\hat{\mathcal{I}}P \in \mathbb{R}^{n\times d}$, and $D\mathcal{I}u \in \mathbb{R}^{n\times d}$ similarly. Our goal in what follows is to bound the norms of $\mathcal{I}u(x) - u(x)$ and $D\mathcal{I}u(x) - Du(x)$ at a point $x \in \Omega$ by the norms of $\hat{\mathcal{I}}P(x) - P(x)$ and $D\hat{\mathcal{I}}P(x) - DP(x)$. We use $\|\cdot\|$ to denote any vector norm (when the argument is a vector) and the corresponding induced matrix norm (when the argument is a matrix).

Proposition 3.4 *Assume that DP is bounded on Ω, and assume that F and DF are Lipschitz on a set $\mathcal{U} \subset \mathfrak{p}$ whose interior contains the closure of $P(\Omega) = \{P(x) \mid x \in \Omega\}$. Define*

$$C_0 = \sup_{x\in\Omega} \|DP(x)\|,$$
$$C_1 = \sup_{\substack{A,B\in\mathcal{U}\\A\neq B}} \frac{\|F(A)-F(B)\|}{\|A-B\|},$$
$$C_2 = \sup_{\substack{A,B\in\mathcal{U}\\A\neq B}} \frac{\|DF(A)-DF(B)\|}{\|A-B\|}.$$

If $\sup_{x\in\Omega}\|\hat{\mathcal{I}}P(x)-P(x)\|$ *is sufficiently small, then for every* $x\in\Omega$*,*

$$\|\mathcal{I}u(x)-u(x)\| \leq C_1\|\hat{\mathcal{I}}P(x)-P(x)\| \tag{24}$$

and

$$\begin{aligned}\|D\mathcal{I}u(x)-Du(x)\| &\leq C_1\|D\hat{\mathcal{I}}P(x)-DP(x)\| \\ &\quad + C_2\|\hat{\mathcal{I}}P(x)-P(x)\|\left(C_0+\|D\hat{\mathcal{I}}P(x)-DP(x)\|\right).\end{aligned} \tag{25}$$

Proof If $\sup_{x\in\Omega}\|\hat{\mathcal{I}}P(x)-P(x)\|$ is sufficiently small, then $\hat{\mathcal{I}}P(\Omega)=\{\hat{\mathcal{I}}P(x)\mid x\in\Omega\}\subseteq\mathcal{U}$. Inequality (24) then follows immediately from the definition of C_1, since $\mathcal{I}u=F\circ\hat{\mathcal{I}}P$ and $u=F\circ P$. Moreover, the chain rule implies

$$\begin{aligned}D\mathcal{I}u(x)-Du(x) &= DF(\hat{\mathcal{I}}P(x))D\hat{\mathcal{I}}P(x)-DF(P(x))DP(x)\\ &= \left[DF(\hat{\mathcal{I}}P(x))-DF(P(x))\right]D\hat{\mathcal{I}}P(x)\\ &\quad + DF(P(x))\left[D\hat{\mathcal{I}}P(x)-DP(x)\right].\end{aligned}$$

Hence, noting that $\|DF(P(x))\|\leq C_1$, we have

$$\|D\mathcal{I}u(x)-Du(x)\| \leq C_2\|\hat{\mathcal{I}}P(x)-P(x)\|\|D\hat{\mathcal{I}}P(x)\| + C_1\|D\hat{\mathcal{I}}P(x)-DP(x)\|.$$

This proves inequality (25), since

$$\begin{aligned}\|D\hat{\mathcal{I}}P(x)\| &\leq \|D\hat{\mathcal{I}}P(x)-DP(x)\|+\|DP(x)\|\\ &\leq \|D\hat{\mathcal{I}}P(x)-DP(x)\|+C_0.\end{aligned}$$

□

The preceding proposition implies that the error in $\mathcal{I}u$ is controlled pointwise by the error in $\hat{\mathcal{I}}P$, and the error in $D\mathcal{I}u$ is controlled pointwise by the error in $D\hat{\mathcal{I}}P$, up to the addition of terms that are, in typical applications, small in comparison with $D\hat{\mathcal{I}}P-DP$. Needless to say, analogous estimates with obvious modifications hold for the interpolant (17).

It should be noted that these estimates depend on the choice of embedding of $\mathcal{S}$ in $\mathbb{R}^n$. Inequality (24) can be easily expressed more intrinsically by replacing the left-hand side with the geodesic distance between $\mathcal{I}u(x)$ and $u(x)$, and replacing C_1 with the appropriately modified Lipschitz constant. Intrinsic variants of inequality (25) are not as easy to derive, and it would be interesting to do so following the lead of [22] and [21].

4 Applications

In this section, we apply the general theory above to several symmetric spaces, including the space of symmetric positive-definite matrices, the space of Lorentzian metrics, and the Grassmannian.

4.1 Symmetric Matrices with Fixed Signature

Let n be a positive integer and let p and q be nonnegative integers satisfying $p+q=n$. Consider the set

$$\mathcal{L} = \{L \in GL_n(\mathbb{R}) \mid \text{signature}(L) = (q, p)\},$$

where signature(L) denotes the signature of a nonsingular symmetric matrix L—an ordered pair indicating the number of positive and negative eigenvalues of L. The general linear group $GL_n(\mathbb{R})$ acts transitively on $\mathcal{L}$ via the group action

$$A \cdot L = ALA^T,$$

where $A \in GL_n(\mathbb{R})$ and $L \in \mathcal{L}$. Let $J = \text{diag}(-1, \ldots, -1, 1, \ldots, 1)$ denote the diagonal $n \times n$ matrix with p entries equal to -1 and q entries equal to 1. The stabilizer of J in $GL_n(\mathbb{R})$ is the indefinite orthogonal group [34, pp. 70-71]

$$O(p, q) = \{Q \in GL_n(\mathbb{R}) \mid QJQ^T = J\}.$$

Its elements are precisely those matrices that are fixed points of the involutive automorphism

$$\begin{aligned} \sigma : GL_n(\mathbb{R}) &\to GL_n(\mathbb{R}) \\ A &\mapsto JA^{-T}J, \end{aligned}$$

where A^{-T} denotes the inverse transpose of a matrix $A \in GL_n(\mathbb{R})$. In contrast, the set of matrices which are mapped by σ to their inverses is

$$Sym_J(n) = \{P \in GL_n(\mathbb{R}) \mid PJ = JP^T\}.$$

The setting we have just described is an instance of the general theory presented in Sect. 2.1, with $G = GL_n(\mathbb{R})$, $G^\sigma = O(p, q)$, $G_\sigma = Sym_J(n)$, $\mathcal{S} = \mathcal{L}$, and $\eta = J$.

It follows that the generalized polar decomposition (1) of a matrix $A \in GL_n(\mathbb{R})$ (sufficiently close to the identity matrix I) with respect to σ reads [27, Theorem 5.1]

$$A = PQ, \quad P \in Sym_J(n),\ Q \in O(p,q). \tag{26}$$

The generalized Cartan decomposition (2) decomposes an element Z of the Lie algebra $\mathfrak{gl}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$ of the general linear group as a sum

$$Z = X + Y, \quad X \in \mathfrak{sym}_J(n),\ Y \in \mathfrak{o}(p,q),$$

where

$$\mathfrak{sym}_J(n) = \{X \in \mathfrak{gl}_n(\mathbb{R}) \mid XJ = JX^T\}$$

and

$$\mathfrak{o}(p,q) = \{Y \in \mathfrak{gl}_n(\mathbb{R}) \mid YJ + JY^T = 0\}$$

denotes the Lie algebra of $O(p,q)$.

We can now write down the map $F : \mathfrak{sym}_J(n) \to \mathcal{L}$ defined abstractly in (3), which provides a diffeomorphism between a neighborhood of the zero matrix and a neighborhood of J. By definition,

$$\begin{aligned} F(X) &= \exp(X)J\exp(X)^T \\ &= \exp(X)\exp(X)J \\ &= \exp(2X)J, \end{aligned} \tag{27}$$

where the second line follows from the fact that $\exp(X) \in Sym_J(n)$ whenever $X \in \mathfrak{sym}_J(n)$. Notice that F maps straight lines in $\mathfrak{sym}_J(n)$ passing through the zero matrix to curves in $Sym_J(n)$ passing through J.

The inverse of F can likewise be expressed in closed form. This can be obtained directly by solving (27) for X, but it is instructive to see how to derive the same result by inverting each of the maps appearing in the composition (3). To start, note that explicit formulas for the matrices P and Q in the decomposition (26) of a matrix $A \in GL_n(\mathbb{R})$ are known [28, Theorem 2.3]. Provided that AJA^TJ has no nonpositive real eigenvalues, we have

$$\begin{aligned} P &= (AJA^TJ)^{1/2}, \\ Q &= (AJA^TJ)^{-1/2}A, \end{aligned}$$

where $B^{1/2}$ denotes the principal square root of a matrix B, and $B^{-1/2}$ denotes the inverse of $B^{1/2}$. Thus, if $A \cdot J = AJA^T = L \in \mathcal{L}$ and if LJ has no nonpositive real eigenvalues, then the factor P in the polar decomposition (26) of A is given by

$$P = (LJ)^{1/2}.$$

It follows that for such a matrix L,

$$F^{-1}(L) = \log\left((LJ)^{1/2}\right),$$

where $\log(B)$ denotes the principal logarithm of a matrix B. We henceforth denote by $\mathcal{L}_*$ the set of matrices $L \in \mathcal{L}$ for which LJ has no nonpositive real eigenvalues, so that $F^{-1}(L)$ is well defined for $L \in \mathcal{L}_*$.

The right-hand side of (29) can be simplified using the following property of the matrix logarithm, whose proof can be found in [26, Theorem 11.2]: If a square matrix B has no nonpositive real eigenvalues, then

$$\log(B^{1/2}) = \frac{1}{2}\log(B).$$

From this it follows that

$$F^{-1}(L) = \frac{1}{2}\log(LJ) \tag{28}$$

for $L \in \mathcal{L}_*$. This formula, of course, could have been obtained directly from (27), but we have chosen a more circuitous derivation to give a concrete illustration of the theory presented in Sect. 2.

Substituting (27) and (28) into (11) gives the following heuristic for interpolating a set of matrices $L_1, L_2, \ldots, L_m \in \mathcal{L}_*$—thought of as the values of a smooth function $L : \Omega \to \mathcal{L}_*$ at points $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$ in a domain Ω—at a point $x \in \Omega$:

$$\mathcal{I}L(x) = \exp\left(\sum_{i=1}^{m}\phi_i(x)\log(L_i J)\right)J. \tag{29}$$

Here, as before, the functions $\phi_i : \Omega \to \mathbb{R}$, $i = 1, 2, \ldots, m$, denote scalar-valued shape functions with the property that $\phi_i(x^{(j)}) = \delta_{ij}$. The right-hand side of (29) can be rewritten in an equivalent way if one uses the fact that the matrix exponential commutes with conjugation and the matrix logarithm commutes with conjugation when its argument has no nonpositive real eigenvalues. Since $L_i J$ has no nonpositive real eigenvalues for each i, and since $J^{-1} = J$, a short calculation shows that

$$\mathcal{I}L(x) = J\exp\left(\sum_{i=1}^{m}\phi_i(x)\log(JL_i)\right). \tag{30}$$

In addition to satisfying $\mathcal{I}L(x) \in \mathcal{L}$ for every $x \in \Omega$, the interpolant so defined enjoys the following properties, which generalize the observations made in Theorems 3.13 and 4.2 of [6].

Lemma 4.1 *Let $Q \in O(p, q)$. If $\tilde{L}_i = QL_iQ^T$, $i = 1, 2, \ldots, m$, and if Q is sufficiently close to the identity matrix, then*

$$\mathcal{I}\tilde{L}(x) = Q\,\mathcal{I}L(x)\,Q^T.$$

for every $x \in \Omega$.

Proof Apply Proposition 3.1. □

Lemma 4.2 *If* $\tilde{L}_i = JL_i^{-1}J$, $i = 1, 2, \ldots, m$, *then*

$$\mathcal{I}\tilde{L}(x) = J\,(\mathcal{I}L(x))^{-1}\,J.$$

for every $x \in \Omega$.

Proof Apply Proposition 3.2, noting that if $L \in \mathcal{L}$ and $L = A \cdot J = AJA^T$, $A \in GL_n(\mathbb{R})$, then $s_\eta(L) = \sigma(A) \cdot J = \sigma(A)J\sigma(A)^T = (JA^{-T}J)J(JA^{-T}J)^T = JA^{-T}JA^{-1}J = JL^{-1}J$. □

Note that the preceding two lemmas can be combined to conclude that if $\tilde{L}_i = L_i^{-1}$, $i = 1, 2, \ldots, m$, then

$$\mathcal{I}\tilde{L}(x) = (\mathcal{I}L(x))^{-1}.$$

To see this, observe that $L_i^{-1} = J(JL_i^{-1}J)J^T$ and $J \in O(p,q)$.

Lemma 4.3 *If* $\sum_{i=1}^m \phi_i(x) = 1$ *for every* $x \in \Omega$, *then*

$$\det \mathcal{I}L(x) = \prod_{i=1}^m (\det L_i)^{\phi_i(x)}$$

for every $x \in \Omega$.

Proof Using the identities $\det \exp(A) = \exp(\mathrm{tr}(A))$ and $\mathrm{tr}(\log(A)) = \log(\det A)$, we have

$$\begin{aligned}
\det \mathcal{I}L(x) &= \det\left(\exp\left(\sum_{i=1}^m \phi_i(x)\log(L_iJ)\right)\right)\det J \\
&= \exp\left(\mathrm{tr}\left(\sum_{i=1}^m \phi_i(x)\log(L_iJ)\right)\right)\det J \\
&= \exp\left(\sum_{i=1}^m \phi_i(x)\mathrm{tr}\left(\log(L_iJ)\right)\right)\det J \\
&= \exp\left(\sum_{i=1}^m \phi_i(x)\log\left(\det(L_iJ)\right)\right)\det J
\end{aligned}$$

$$= \left(\prod_{i=1}^{m} \det(L_i J)^{\phi_i(x)} \right) \det J$$
$$= \left(\prod_{i=1}^{m} \det(L_i)^{\phi_i(x)} \det(J)^{\phi_i(x)} \right) \det J$$

The conclusion then follows from the fact that $\sum_{i=1}^{m} \phi_i(x) = 1$ and $\det J = \pm 1$. $\square$

Generalizations. As explained abstractly in Sect. 3.3, the interpolation formula (30) can be generalized by fixing an element $\bar{A} \in GL_n(\mathbb{R})$ and replacing (27) with the map

$$F_{\bar{A}}(X) = \bar{A}\exp(X) J \left(\bar{A}\exp(X)\right)^T = \bar{A} F(X) \bar{A}^T = \bar{A}\exp(2X) J \bar{A}^T.$$

The inverse of this map reads

$$F_{\bar{A}}^{-1}(L) = \frac{1}{2}\log(\bar{A}^{-1} L \bar{A}^{-T} J).$$

Substituting into (17) gives

$$\begin{aligned}
\mathcal{I}_{\bar{A}} L(x) &= F_{\bar{A}} \left(\sum_{i=1}^{m} \phi_i(x) F_{\bar{A}}^{-1}(L_i) \right) \\
&= \bar{A}\exp\left(2 \sum_{i=1}^{m} \phi_i(x) \frac{1}{2} \log(\bar{A}^{-1} L_i \bar{A}^{-T} J) \right) J \bar{A}^T \\
&= \bar{L} \left(J \bar{A}^T \right)^{-1} \exp\left(\sum_{i=1}^{m} \phi_i(x) \log(\bar{A}^{-1} L_i \bar{A}^{-T} J) \right) J \bar{A}^T,
\end{aligned}$$

where $\bar{L} = \bar{A} J \bar{A}^T$. Using the fact that the matrix exponential commutes with conjugation and the matrix logarithm commutes with conjugation when its argument has no nonpositive real eigenvalues, we conclude that

$$\begin{aligned}
\mathcal{I}_{\bar{A}} L(x) &= \bar{L} \exp\left(\sum_{i=1}^{m} \phi_i(x) \left(J \bar{A}^T \right)^{-1} \log(\bar{A}^{-1} L_i \bar{A}^{-T} J) J \bar{A}^T \right) \\
&= \bar{L} \exp\left(\sum_{i=1}^{m} \phi_i(x) \log(\bar{L}^{-1} L_i) \right),
\end{aligned} \tag{31}$$

provided that $\bar{L}^{-1} L_i$ has no nonpositive real eigenvalues for each i.

Rather than fixing $\bar{A}$, one may choose to define $\bar{A}$ implicitly via (21); that is,

$$\bar{A}(x) J \bar{A}(x)^T = \mathcal{I}_{\bar{A}(x)} L(x).$$

The output of the resulting interpolation scheme is the solution $\bar{L}$ to the equation

$$\sum_{i=1}^{m} \phi_i(x) \log\left(\bar{L}^{-1} L_i\right) = 0, \tag{32}$$

which can be computed with a fixed-point iteration.

Algorithms. In summary, we have derived the following pair of algorithms for interpolating matrices in the space $\mathcal{L}$ of nonsingular symmetric matrices with signature (q, p). The first of these algorithms implements (31), which reduces to (30) when $\bar{L}$ is taken equal to J. The algorithm implicitly requires its inputs to have the property that for each $i = 1, 2, \ldots, m$, the matrix $\bar{L}^{-1} L_i$ has no nonpositive real eigenvalues.

Algorithm 1 Interpolation of symmetric matrices with fixed signature

Require: Matrices $\{L_i \in \mathcal{L}\}_{i=1}^m$, shape functions $\{\phi_i : \Omega \to \mathbb{R}\}_{i=1}^m$, point $x \in \Omega$, matrix $\bar{L} \in \mathcal{L}$

1: **return** $\bar{L} \exp\left(\sum_{i=1}^m \phi_i(x) \log\left(\bar{L}^{-1} L_i\right)\right)$

The second algorithm solves (32) and requires the same constraint on its inputs as Algorithm 1. Observe that Algorithm 1 is equivalent to Algorithm 2 if one terminates the fixed-point iteration after the first iteration.

Algorithm 2 Iterative interpolation of symmetric matrices with fixed signature

Require: Matrices $\{L_i \in \mathcal{L}\}_{i=1}^m$, shape functions $\{\phi_i : \Omega \to \mathbb{R}\}_{i=1}^m$, point $x \in \Omega$, initial guess $\bar{L} \in \mathcal{L}$, tolerance $\varepsilon > 0$

1: **while** $\left\| \sum_{i=1}^m \phi_i(x) \log\left(\bar{L}^{-1} L_i\right) \right\| > \varepsilon$ **do**

2: $\quad \bar{L} = \bar{L} \exp\left(\sum_{i=1}^m \phi_i(x) \log\left(\bar{L}^{-1} L_i\right)\right)$

3: **end while**

4: **return** $\bar{L}$

4.1.1 Symmetric Positive-Definite Matrices

When $J = I$, the preceding theory provides structure-preserving interpolation schemes for the space $SPD(n)$ of symmetric positive-definite matrices. Formula (30) is the weighted log-Euclidean mean introduced by [6], and equation (32) gives the weighted Riemannian mean (or Karcher mean) of symmetric positive-definite matrices [9,33,38]. The latter observation can be viewed as a consequence of Lemma 3.3, which applies in this setting since $O(n)$ is compact.

We remark that the interpolation formula (30) on $SPD(n)$ was devised in [6] by endowing $SPD(n)$ with what the authors term a "novel vector space structure." This vector space structure is nothing more than that obtained by identifying $SPD(n)$ with the Lie triple system $\mathfrak{sym}_I(n)$ via map (28), as we have done here.

4.1.2 Lorentzian Metrics

When $n = 4$ and $J = \mathrm{diag}(-1, 1, 1, 1)$, the preceding theory provides structure-preserving interpolation schemes for the space of Lorentzian metrics—the space of symmetric, nonsingular matrices having signature $(3, 1)$. Lemma 4.1 states that the interpolation operator (30) in this setting commutes with Lorentz transformations. By choosing, for instance, Ω equal to a four-dimensional simplex (or a four-dimensional hypercube) and $\{\phi_i\}_i$ equal to scalar-valued Lagrange polynomials (or tensor products of Lagrange polynomials) on Ω, one obtains a family of Lorentzian metric-valued finite elements. These elements are capable of interpolating Lorentzian metric-valued functions whose components are continuous on the closure of Ω.

In view of their potential application to numerical relativity, we have numerically computed the interpolation error committed by such elements when approximating the Schwarzschild metric, which is an explicit solution to Einstein's equations outside of a spherical mass [10, p. 193]. In Cartesian coordinates, this metric reads

$$L(t,x,y,z) = \begin{pmatrix} -\left(1-\frac{R}{r}\right) & 0 & 0 & 0 \\ 0 & 1+\left(\frac{R}{r-R}\right)\frac{x^2}{r^2} & \left(\frac{R}{r-R}\right)\frac{xy}{r^2} & \left(\frac{R}{r-R}\right)\frac{xz}{r^2} \\ 0 & \left(\frac{R}{r-R}\right)\frac{xy}{r^2} & 1+\left(\frac{R}{r-R}\right)\frac{y^2}{r^2} & \left(\frac{R}{r-R}\right)\frac{yz}{r^2} \\ 0 & \left(\frac{R}{r-R}\right)\frac{xz}{r^2} & \left(\frac{R}{r-R}\right)\frac{yz}{r^2} & 1+\left(\frac{R}{r-R}\right)\frac{z^2}{r^2} \end{pmatrix}, \tag{33}$$

where R (the Schwarzschild radius) is a positive constant (which we take equal to 1 in what follows) and $r = \sqrt{x^2+y^2+z^2} > R$. We interpolated this metric over the region $U = \{0\} \times [2,3] \times [2,3] \times [2,3]$ on a uniform $N \times N \times N$ grid of cubes using formula (30) elementwise, with shape functions $\{\phi_i\}_i$ given by tensor products of Lagrange polynomials of degree k. The results in Table 1 indicate that the L^2-error

$$\|\mathcal{I}L - L\|_{L^2(U)} = \left(\int_U \|\mathcal{I}L(t,x,y,z) - L(t,x,y,z)\|_F^2 \, dx\, dy\, dz\right)^{1/2} \tag{34}$$

(which we approximated with numerical quadrature) converges to zero with order 2 and 3, respectively, when using polynomials of degree $k = 1$ and $k = 2$. Here, $\|\cdot\|_F$ denotes the Frobenius norm. In addition, Table 1 indicates that the error in the H^1-seminorm (referred to abusively as the H^1-error in Table 1)

$$|\mathcal{I}L - L|_{H^1(U)} = \left(\int_U \sum_{j=1}^{4} \left\| \frac{\partial \mathcal{I}L}{\partial \xi_j}(t,x,y,z) - \frac{\partial L}{\partial \xi_j}(t,x,y,z) \right\|_F^2 dx\, dy\, dz\right)^{1/2} \tag{35}$$

converges to zero with order 1 and 2, respectively, when using polynomials of degree $k = 1$ and $k = 2$. Here, we have denoted $\xi = (t, x, y, z)$.

For the sake of comparison, Table 2 shows the interpolation errors committed when applying componentwise polynomial interpolation to the same problem. Within each element, the value of this interpolant at a point $\xi = (t, x, y, z)$ lying in the element is given by

$$\mathcal{I}L(\xi) = \sum_{i=1}^{m} \phi_i(\xi) L_i, \tag{36}$$

where $\{\phi_i\}_i$ are tensor products of Lagrange polynomials of degree k and $\{L_i\}_i$ are the values of L at the corresponding degrees of freedom. The errors committed by this interpolation scheme are very close to those observed in Table 1 for the structure-preserving scheme (30).

For this particular numerical example, the componentwise polynomial interpolant (36) has correct signature (3, 1) for every $(t, x, y, z) \in U$. This need not hold in general. For example, consider the metric tensor

$$L(t, x, y, z) = \begin{pmatrix} -6\sin^2(2\pi x) + 3\sin^2(\pi x) & 3\cos(2\pi x) & 0 & 0 \\ 3\cos(2\pi x) & 2\sin^2(2\pi x) + 2\sin^2(\pi x) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Though not a solution to Einstein's equations, this metric tensor nonetheless has signature (3, 1) everywhere. Indeed, a numerical calculation verifies that at all points (t, x, y, z), the matrix $L(t, x, y, z)$ has eigenvalues $\lambda_-, 1, 1, \lambda_+$ satisfying $\lambda_- \le \alpha$ and $\lambda_+ \ge \beta$ with $\alpha \approx -0.54138$ and $\beta \approx 2.23064$. Interpolating this metric componentwise with linear polynomials (over the region same region U as above) produces a metric with signature (4, 0) at 32 quadrature points (out of 64 total) on the coarsest grid ($N = 2$). The essence of the problem is that for any integer k, any t, any y, and any z, the average of $L(t, k/2, y, z)$ and $L(t, (k+1)/2, y, z)$ is

$$\frac{1}{2}\left(L(t, k/2, y, z) + L(t, (k+1)/2, y, z)\right) = \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

which shows (by continuity of the interpolant) that the componentwise linear interpolant (36) on the coarsest grid ($N = 2$) is positive definite on an open subset of U. In contrast, the structure-preserving scheme (30) automatically generates an interpolant with correct signature (3, 1) at all points (t, x, y, z).

4.2 The Grassmannian

Let p and n be positive integers satisfying $p < n$. Consider the Grassmannian $Gr(p, n)$, which consists of all p-dimensional linear subspaces of $\mathbb{R}^n$. Any element $\mathcal{V} \in Gr(p, n)$ can be written as the span of p vectors $v_1, v_2, \ldots, v_p \in \mathbb{R}^n$. The orthogonal group $O(n)$ acts transitively on $Gr(p, n)$ via the action

Table 1 Error incurred when interpolating the Schwarzschild metric (33) over the region $U = \{0\} \times [2,3] \times [2,3] \times [2,3]$ using the formula (30). The interpolant was computed elementwise on a uniform $N \times N \times N$ grid of cubes, with shape functions $\{\phi_i\}_i$ on each cube given by tensor products of Lagrange polynomials of degree k

	$k=1$				$k=2$			
N	L^2-error	Order	H^1-error	Order	L^2-error	Order	H^1-error	Order
2	3.296×10^{-3}		2.840×10^{-2}		1.740×10^{-4}		2.469×10^{-3}	
4	8.383×10^{-4}	1.975	1.421×10^{-2}	0.998	2.173×10^{-5}	3.001	6.204×10^{-4}	1.993
8	2.105×10^{-4}	1.994	7.108×10^{-3}	0.999	2.715×10^{-6}	3.000	1.553×10^{-4}	1.998
16	5.268×10^{-5}	1.998	3.554×10^{-3}	1.000	3.393×10^{-7}	3.000	3.883×10^{-5}	1.999

Table 2 Error incurred when interpolating the Schwarzschild metric (33) over the region $U = \{0\} \times [2,3] \times [2,3] \times [2,3]$ using the componentwise interpolation formula (36). The interpolant was computed elementwise on a uniform $N \times N \times N$ grid of cubes, with shape functions $\{\phi_i\}_i$ on each cube given by tensor products of Lagrange polynomials of degree k

	$k=1$				$k=2$			
N	L^2-error	Order	H^1-error	Order	L^2-error	Order	H^1-error	Order
2	2.899×10^{-3}		2.952×10^{-2}		1.757×10^{-4}		2.508×10^{-3}	
4	7.372×10^{-4}	1.975	1.477×10^{-2}	0.999	2.201×10^{-5}	2.996	6.296×10^{-4}	1.994
8	1.851×10^{-4}	1.994	7.383×10^{-3}	1.000	2.753×10^{-6}	2.999	1.576×10^{-4}	1.998
16	4.632×10^{-5}	1.998	3.692×10^{-3}	1.000	3.442×10^{-7}	3.000	3.940×10^{-5}	1.999

$$A \cdot \mathrm{span}(v_1, v_2, \dots, v_p) = \mathrm{span}(Av_1, Av_2, \dots, Av_p),$$

where $A \in O(n)$. For convenience, we will sometimes write $A\mathcal{V}$ as shorthand for $\mathrm{span}(Av_1, Av_2, \dots, Av_p)$. Let $e_1, e_2, \dots, e_n$ be the canonical basis for $\mathbb{R}^n$. The stabilizer of $\mathrm{span}(e_1, e_2, \dots, e_p)$ in $O(n)$ is the subgroup

$$O(p) \times O(n-p) = \left\{ \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \mid A_1 \in O(p),\ A_2 \in O(n-p) \right\}.$$

The elements of $O(p) \times O(n-p)$ are precisely those matrices in $O(n)$ that are fixed points of the involutive automorphism

$$\begin{aligned} \sigma : O(n) &\to O(n) \\ A &\mapsto JAJ, \end{aligned}$$

where

$$J = \begin{pmatrix} -I_p & 0 \\ 0 & I_{n-p} \end{pmatrix},$$

and I_p and I_{n-p} denote the $p \times p$ and $(n-p) \times (n-p)$ identity matrices, respectively. The matrices in $O(n)$ that are mapped to their inverses by σ constitute the space

$$Sym_J(n) \cap O(n) = \{P \in O(n) \mid PJ = JP^T\}.$$

The generalized polar decomposition of a matrix $A \in O(n)$ in this setting thus reads

$$A = PQ, \quad P \in Sym_J(n) \cap O(n),\ Q \in O(p) \times O(n-p). \tag{37}$$

The corresponding generalized Cartan decomposition reads

$$Z = X + Y, \quad X \in \mathfrak{sym}_J(n) \cap \mathfrak{o}(n),\ Y \in \mathfrak{o}(p) \times \mathfrak{o}(n-p),$$

where, for each m, $\mathfrak{o}(m)$ denotes the space of antisymmetric $m \times m$ matrices,

$$\mathfrak{o}(p) \times \mathfrak{o}(n-p) = \left\{ \begin{pmatrix} Y_1 & 0 \\ 0 & Y_2 \end{pmatrix} \mid Y_1 \in \mathfrak{o}(p),\ Y_2 \in \mathfrak{o}(n-p) \right\},$$

and

$$\begin{aligned} \mathfrak{sym}_J(n) \cap \mathfrak{o}(n) &= \{X \in \mathfrak{o}(n) \mid XJ = JX^T\} \\ &= \left\{ \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \mid B \in \mathbb{R}^{(n-p)\times p} \right\}. \end{aligned}$$

The map $F : \mathfrak{sym}_J(n) \cap \mathfrak{o}(n) \to Gr(p,n)$ is given by

$$F(X) = \operatorname{span}(\exp(X)e_1, \exp(X)e_2, \dots, \exp(X)e_p).$$

The inverse of F can be computed (naively) as follows. Given an element $\mathcal{V} \in Gr(p,n)$, let $a_1, a_2, \dots, a_p$ be an orthonormal basis for $\mathcal{V}$. Extend this basis to an orthonormal basis $a_1, a_2, \dots, a_n$ of $\mathbb{R}^n$. Then

$$F^{-1}(\mathcal{V}) = \log(P),$$

where $P \in Sym_J(n) \cap O(n)$ is the first factor in the generalized polar decomposition (37) of $A = (a_1\, a_2\, \cdots\, a_n)$. Note that this map is independent of the chosen bases for $\mathcal{V}$ and its orthogonal complement in $\mathbb{R}^n$. Indeed, if $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_p$ is any other orthonormal basis for $\mathcal{V}$ and $\tilde{a}_{p+1}, \tilde{a}_{p+2}, \dots, \tilde{a}_n$ is any other basis for the orthogonal complement of $\mathcal{V}$, then there is a matrix $R \in O(p) \times O(n-p)$ such that $\tilde{A} = AR$, where $\tilde{A} = (\tilde{a}_1\, \tilde{a}_2\, \cdots\, \tilde{a}_n)$. The generalized polar decomposition of $\tilde{A}$ is thus $\tilde{A} = P\tilde{Q}$, where $\tilde{Q} = QR$.

More generally, we may opt to fix an element $\bar{A} \in O(n)$ and consider interpolants of the form (17) using the map

$$F_{\bar{A}}(X) = \operatorname{span}(\bar{A}\exp(X)e_1, \bar{A}\exp(X)e_2, \dots, \bar{A}\exp(X)e_p), \tag{38}$$

The inverse of this map, in analogy with the preceding paragraph, is

$$F_{\bar{A}}^{-1}(\mathcal{V}) = \log(P), \tag{39}$$

where now $P \in Sym_J(n) \cap O(n)$ is the first factor in the generalized polar decomposition (37) of $\bar{A}^T A$, where $A \in O(n)$ is a matrix whose first p and last $n-p$ columns, respectively, form orthonormal bases for $\mathcal{V}$ and its orthogonal complement.

Algorithms. We now turn our attention to the computation of the interpolant (17) in this setting. A naive implementation using the steps detailed above for computing $F_{\bar{A}}$ and its inverse would lead to an algorithm for computing the interpolant having complexity $O(n^3)$. Remarkably, the computation of (17) can be performed in $O(np^2)$ operations, as we now show. The resulting algorithm turns out to be identical to that proposed in [4]. The fact that this algorithm scales linearly with n is noteworthy, as it renders this interpolation scheme practical for applications in which $n \gg p$.

The derivation of the algorithm hinges upon the following two lemmas, which, when combined, allow for a computation of the interpolant while operating solely on matrices of size $n \times p$ or smaller. The first lemma gives a useful formula for $F_{\bar{A}}(X)$. In it, we make use of the thin singular value decomposition [46, p. 27], and we adopt the following notation. If Θ is a diagonal matrix, we write $\cos(\Theta)$ and $\sin(\Theta)$ to denote the diagonal matrices with diagonal entries $(\cos(\Theta))_{ii} = \cos(\Theta_{ii})$ and $(\sin(\Theta))_{ii} = \sin(\Theta_{ii})$, respectively.

Lemma 4.4 *Let*

$$\bar{A} = \begin{pmatrix} \bar{A}_1 & \bar{A}_2 \end{pmatrix} \in O(n)$$

with $\bar{A}_1 \in \mathbb{R}^{n\times p}$ *and* $\bar{A}_2 \in \mathbb{R}^{n\times(n-p)}$*, and let* $X = \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \in \mathfrak{sym}_J(n) \cap \mathfrak{o}(n)$ *with* $B \in \mathbb{R}^{(n-p)\times p}$*. Then*

$$\bar{A}\exp(X)\begin{pmatrix} I_p \\ 0 \end{pmatrix} = \bar{A}_1 V \cos(\Theta) V^T + U \sin(\Theta) V^T,$$

where $U \in \mathbb{R}^{n\times p}$, $\Theta \in \mathbb{R}^{p\times p}$*, and* $V \in \mathbb{R}^{p\times p}$ *denote the factors in the thin singular value decomposition*

$$\bar{A}_2 B = U \Theta V^T. \tag{40}$$

In particular, $F_{\bar{A}}(X)$ *is the space spanned by the columns of* $\bar{A}_1 V \cos(\Theta) V^T + U \sin(\Theta) V^T$*. Equivalently, since* V *is orthogonal,* $F_{\bar{A}}(X)$ *is the space spanned by the columns of* $\bar{A}_1 V \cos(\Theta) + U \sin(\Theta)$.

Proof The formula is proved in [17, Theorem 2.3].

The next lemma gives a useful formula for $F_{\bar{A}}^{-1}(\mathcal{V})$. Closely related formulas appear without proof in [4,8,13] and elsewhere, so we give a proof here for completeness.

Lemma 4.5 *Let $\bar{A} = \begin{pmatrix} \bar{A}_1 & \bar{A}_2 \end{pmatrix} \in O(n)$ be as in Lemma (4.4), and let $\mathcal{V} \in Gr(p,n)$. Let*

$$A = \begin{pmatrix} A_1 & A_2 \end{pmatrix} \in O(n)$$

be such that the columns of A_1 and A_2, respectively, form orthonormal bases for $\mathcal{V}$ and its orthogonal complement. Assume that $\bar{A}_1^T A_1$ is invertible. Then

$$F_{\bar{A}}^{-1}(\mathcal{V}) = \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix},$$

where

$$B = \bar{A}_2^T U \arctan(\Sigma) V^T, \tag{41}$$

and $U \in \mathbb{R}^{n \times p}$, $\Sigma \in \mathbb{R}^{p \times p}$, and $V \in \mathbb{R}^{p \times p}$ denote the factors in the thin singular value decomposition

$$(I - \bar{A}_1 \bar{A}_1^T) A_1 (\bar{A}_1^T A_1)^{-1} = U \Sigma V^T. \tag{42}$$

Proof It is enough to check that if B is given by (41), then the image of $\begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}$ under $F_{\bar{A}}$ is $\mathcal{V}$. In other words, we must check that the columns of

$$\bar{A} \exp \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} I_p \\ 0 \end{pmatrix} \tag{43}$$

span $\mathcal{V}$. To this end, observe that by the orthogonality of $\bar{A}$,

$$\bar{A}_2 \bar{A}_2^T U = (I - \bar{A}_1 \bar{A}_1^T) U = U, \tag{44}$$

where the last equality follows from (42) upon noting that $(I - \bar{A}_1 \bar{A}_1^T)$ is a projection. Thus, by inspection of (41), the thin singular value decomposition of $\bar{A}_2 B$ is

$$\bar{A}_2 B = U \Theta V^T,$$

where $\Theta = \arctan \Sigma$. Now by Lemma 4.4,

$$\bar{A} \exp \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} I_p \\ 0 \end{pmatrix} = \bar{A}_1 V \cos(\Theta) V^T + U \sin(\Theta) V^T. \tag{45}$$

Using (42), this simplifies to

$$\begin{aligned} \bar{A} \exp \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} I_p \\ 0 \end{pmatrix} &= \bar{A}_1 V \cos(\Theta) V^T + (I - \bar{A}_1 \bar{A}_1^T) A_1 (\bar{A}_1^T A_1)^{-1} V \Sigma^{-1} \sin(\Theta) V^T \\ &= \bar{A}_1 V \cos(\Theta) V^T + (I - \bar{A}_1 \bar{A}_1^T) A_1 (\bar{A}_1^T A_1)^{-1} V \cos(\Theta) V^T \\ &= A_1 (\bar{A}_1^T A_1)^{-1} V \cos(\Theta) V^T. \end{aligned}$$

Observe that since $\Sigma = \tan(\Theta)$ is finite, the diagonal entries of $\cos(\Theta)$ are nonzero. Thus, $(\bar{A}_1^T A_1)^{-1} V \cos(\Theta) V^T$ is invertible, so we conclude that the columns of (43) span the same space that is spanned by the columns of A_1, namely $\mathcal{V}$. □

The preceding two lemmas lead to the following algorithm, which coincides with that introduced in [4], for computing the interpolant

$$\mathcal{I}_{\bar{A}} \mathcal{V}(x) = F_{\bar{A}} \left(\sum_{i=1}^{m} \phi_i(x) F_{\bar{A}}^{-1}(\mathcal{V}^{(i)}) \right) \tag{46}$$

of elements $\mathcal{V}^{(1)}, \mathcal{V}^{(2)}, \ldots, \mathcal{V}^{(m)}$ of $Gr(p,n)$. Note that the computational complexity of this algorithm is $O(np^2)$. In particular, owing to the identity (44), the $(n-p) \times n$ matrix $\bar{A}_2$ plays no role in the algorithm, despite its worrisome appearance in (40) and (41).

Algorithm 3 Interpolation on the Grassmannian $Gr(p,n)$

Require: Subspaces $\{\mathcal{V}^{(i)} \in Gr(p,n)\}_{i=1}^m$, shape functions $\{\phi_i : \Omega \to \mathbb{R}\}_{i=1}^m$, point $x \in \Omega$, matrix $\bar{A}_1 \in \mathbb{R}^{n \times p}$ with orthonormal columns

1: $Z \leftarrow 0_{n \times p}$
2: **for** $i = 1, 2, \ldots, m$ **do**
3: Let $A_1 \in \mathbb{R}^{n \times p}$ be a matrix whose columns form an orthonormal basis for $\mathcal{V}^{(i)}$.
4: Compute the thin singular value decomposition

$$(I - \bar{A}_1 \bar{A}_1^T) A_1 (\bar{A}_1^T A_1)^{-1} = U \Sigma V^T,$$

with $U \in \mathbb{R}^{n \times p}$, $\Sigma \in \mathbb{R}^{p \times p}$, and $V \in \mathbb{R}^{p \times p}$.
5: $Z \mathrel{+}= \phi_i(x) U \arctan(\Sigma) V^T$
6: **end for**
7: Compute the thin singular value decomposition

$$Z = U \Theta V^T,$$

with $U \in \mathbb{R}^{n \times p}$, $\Theta \in \mathbb{R}^{p \times p}$, and $V \in \mathbb{R}^{p \times p}$.
8: $A \leftarrow \bar{A}_1 V \cos(\Theta) + U \sin(\Theta)$
9: **return** $\mathrm{span}(a_1, a_2, \ldots, a_p)$, where a_j denotes the j^{th} column of A.

Note that the output of Algorithm 3 is independent of the choice of orthonormal basis made for each $\mathcal{V}^{(i)}$ in Line 3 of Algorithm 3. This can be checked directly by observing that a change of basis corresponds to post-multiplication of A_1 by a matrix $R \in O(p)$, leaving $(I - \bar{A}_1 \bar{A}_1^T) A_1 (\bar{A}_1^T A_1)^{-1}$ invariant. Similarly, the output of the algorithm is invariant under post-multiplication of $\bar{A}_1$ by any matrix $R \in O(p)$, since it can be checked that such a transformation leaves the output of Line 8 invariant. This last statement leads to the conclusion that

$$\mathcal{I}_{\bar{A}Q} \tilde{\mathcal{V}}(x) = \mathcal{I}_{\bar{A}} \tilde{\mathcal{V}}(x) \tag{47}$$

for any $Q \in O(p) \times O(n-p)$, which reaffirms (20).

The interpolant so constructed enjoys the following additional property.

Lemma 4.6 *The interpolant (46) commutes with the action of $O(n)$ on $Gr(p,n)$. That is, if $Q \in O(n)$ and $\tilde{\mathcal{V}}^{(i)} = Q\mathcal{V}^{(i)}$, $i = 1, 2, \dots, m$, then*

$$\mathcal{I}_{Q\bar{A}}\tilde{\mathcal{V}}(x) = Q\mathcal{I}_{\bar{A}}\mathcal{V}(x)$$

for every $x \in \Omega$.

Proof Apply (18). □

Another $O(n)$-equivariant interpolant on $Gr(p,n)$ is given abstractly by (22). In this setting, this interpolant is obtained by solving

$$\sum_{i=1}^{m} \phi_i(x) F_{\bar{A}}^{-1}(\mathcal{V}^{(i)}) = 0$$

for $\bar{A}$ and outputting the space spanned by the first p columns of $\bar{A}$. Algorithmically, this amounts to wrapping a fixed-point iteration around Algorithm 3, as detailed below.

Algorithm 4 Iterative interpolation on the Grassmannian $Gr(p,n)$

Require: Subspaces $\{\mathcal{V}^{(i)} \in Gr(p,n)\}_{i=1}^m$, shape functions $\{\phi_i : \Omega \to \mathbb{R}\}_{i=1}^m$, point $x \in \Omega$, matrix $\bar{A}_1 \in \mathbb{R}^{n\times p}$ with orthonormal columns

1: **repeat**
2: Use Algorithm 3 to compute the interpolant of $\{\mathcal{V}^{(i)}\}_{i=1}^m$ at x, storing the result as a matrix $A \in \mathbb{R}^{n\times p}$ (i.e., the matrix A appearing in line 8 of Algorithm 3).
3: $\bar{A}_1 \leftarrow A$
4: **until** converged
5: **return** $\mathrm{span}(a_1, a_2, \dots, a_p)$, where a_j denotes the j^{th} column of $\bar{A}_1$.

Since $O(p) \times O(n-p)$ is compact, Lemma 3.3 shows that Algorithm 4 produces the weighted Riemannian mean on $Gr(p,n)$. This interpolant has been considered previously by several authors, including [8,13,21].

4.3 Lie Groups

It was remarked in Sect. 2.2 that any Lie group G can be realized as a symmetric space $(G\times G)/\mathrm{diag}(G\times G)$, since $\mathrm{diag}(G\times G) = \{(g,g) \mid g \in G\}$ fulfills two roles simultaneously: it is the stabilizer of $e \in G$ under the action of $G \times G$ on G given by $(g,h)\cdot k = gkh^{-1}$, and it is the subgroup $(G\times G)^\sigma$ of fixed points of the involutive automorphism $\sigma(g,h) = (h,g)$. In the notation of Sect. 2, one checks that $(G\times G)_\sigma = \{(g,h) \mid g = h^{-1}\}$, $\mathfrak{p} = \{(X,-X) \mid X \in \mathfrak{g}\}$, $\mathfrak{k} = \{(X,X) \mid X \in \mathfrak{g}\}$, and $F(X,-X) = \exp(2X)$. Thus, the interpolant (11) of a collection of elements

$g_1, g_2, \dots, g_m \in G$ reads

$$\mathcal{I}g(x) = \exp\left(\sum_{i=1}^{m} \phi_i(x) \log(g_i)\right).$$

This is of course a standard strategy for interpolation on Lie groups that enjoys widespread use [39], and it belongs to a broad class of methods that perform interpolation on Lie groups by mapping elements of G to $\mathfrak{g}$ and back [36,45]. This interpolant, being $(G \times G)^\sigma$-equivariant, commutes with the adjoint action of G on itself. That is, $\mathcal{I}(hgh)^{-1}(x) = h\mathcal{I}g(x)h^{-1}$ for every $h \in G$ sufficiently close to the identity $e \in G$. However, it is not G-equivariant ($\mathcal{I}(hg)(x) \neq h\mathcal{I}g(x)$ in general), and it has the disadvantage of requiring that each g_i be close to e. The variant (21) overcomes these limitations by seeking a solution $\bar{g} \in G$ to the equation

$$\bar{g} = \exp\left(\sum_{i=1}^{m} \phi_i(x) \log(\bar{g}^{-1} g_i)\right),$$

which defines a geodesic finite element [22,43,44] if G is equipped with a bi-invariant metric (so that the Lie group exponential and Riemannian exponential maps coincide). The latter interpolant exists whenever $g_1, g_2, \dots, g_m$ are sufficiently close to one another [22, Theorem 3.1], and it is manifestly G-equivariant.

5 Conclusion

This paper has presented a family of structure-preserving interpolation operators for functions taking values in a symmetric space $\mathcal{S}$. We accomplished this by identifying $\mathcal{S}$ with a homogeneous space G/G^σ and interpolating coset representatives obtained from the generalized polar decomposition. The resulting interpolation operators enjoy equivariance with respect to the action of G^σ on $\mathcal{S}$, equivariance with respect to the action of certain geodesic symmetries on $\mathcal{S}$, and optimal approximation properties. The application of these interpolation schemes seems intriguing, particularly in the context of numerical relativity, where they provide structure-preserving finite elements for the metric tensor.

Acknowledgements EG has been supported in part by NSF under Grants DMS-1411792, DMS-1345013. ML has been supported in part by NSF under Grants DMS-1010687, CMMI-1029445, DMS-1065972, CMMI-1334759, DMS-1411792, DMS-1345013.

Appendix A: Second-Order Derivatives of the Matrix Exponential

In this section, we prove (16) by showing that if $\hat{\mathcal{I}}P : \Omega \to \mathbb{R}^{n\times n}$ is a smooth matrix-valued function defined on a domain $\Omega \subset \mathbb{R}^d$, then, for each $j, k = 1, 2, \dots, d$, the matrix $\frac{\partial^2}{\partial x_j \partial x_k} \exp(\hat{\mathcal{I}}P(x))$ is given by reading off the (1, 4) block of

$$\exp\begin{pmatrix} X & Y & Z & W \\ 0 & X & 0 & Z \\ 0 & 0 & X & Y \\ 0 & 0 & 0 & X \end{pmatrix}, \tag{48}$$

where $X = \hat{\mathcal{I}}P(x)$, $Y = \frac{\partial \hat{\mathcal{I}}P}{\partial x_j}(x)$, $Z = \frac{\partial \hat{\mathcal{I}}P}{\partial x_k}(x)$, and $W = \frac{\partial^2 \hat{\mathcal{I}}P}{\partial x_j \partial x_k}(x)$. To prove this, recall first the identity (15), which can be written as

$$\left.\frac{d}{dt}\right|_{t=0} \exp(U + tV) = \mathcal{R}\left[\exp\begin{pmatrix} U & V \\ 0 & U \end{pmatrix}\right] \tag{49}$$

for any square matrices U and V of equal size, where $\mathcal{R}$ denotes the map which sends a $2l \times 2l$ matrix B to the $l \times l$ submatrix of B consisting of the intersection of the first l rows and last l columns of B. Now observe that with X, Y, Z, and W defined as above,

$$\begin{aligned} \frac{\partial^2}{\partial x_j \partial x_k} \exp(\hat{\mathcal{I}}P(x)) &= \left.\frac{\partial^2}{\partial s\, \partial t}\right|_{s=t=0} \exp(X + tY + sZ + stW) \\ &= \left.\frac{\partial}{\partial s}\right|_{s=0} \left.\frac{\partial}{\partial t}\right|_{t=0} \exp(X + sZ + t(Y + sW)) \\ &= \left.\frac{\partial}{\partial s}\right|_{s=0} \mathcal{R}\left[\exp\begin{pmatrix} X + sZ & Y + sW \\ 0 & X + sZ \end{pmatrix}\right] \\ &= \mathcal{R}\left[\left.\frac{\partial}{\partial s}\right|_{s=0} \exp\begin{pmatrix} X + sZ & Y + sW \\ 0 & X + sZ \end{pmatrix}\right] \end{aligned}$$

Using (49) again, we have

$$\left.\frac{\partial}{\partial s}\right|_{s=0} \exp\begin{pmatrix} X + sZ & Y + sW \\ 0 & X + sZ \end{pmatrix} = \mathcal{R}\left[\exp\begin{pmatrix} X & Y & Z & W \\ 0 & X & 0 & Z \\ 0 & 0 & X & Y \\ 0 & 0 & 0 & X \end{pmatrix}\right].$$

This shows that

$$\frac{\partial^2}{\partial x_j \partial x_k} \exp(\hat{\mathcal{I}}P(x)) = \mathcal{R}\left[\mathcal{R}\left[\exp\begin{pmatrix} X & Y & Z & W \\ 0 & X & 0 & Z \\ 0 & 0 & X & Y \\ 0 & 0 & 0 & X \end{pmatrix}\right]\right],$$

which is precisely the $(1, 4)$ block of matrix (48).

References

1. P.-A. Absil, R. Mahony, and R. Sepulchre. Optimization Algorithms on Matrix Mani- folds. Princeton University Press, 2009.
2. P.-A. Absil, R. Mahony, and R. Sepulchre. Riemannian geometry of Grassmann manifolds with a view on algorithmic computation. Acta Applicandae Mathematica 80.2 (2004), pp. 199–220.
3. A. H. Al-Mohy and N. J. Higham. Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation. SIAM Journal on Matrix Analysis and Applications 30.4 (2009), pp. 1639–1657
4. D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. AIAA Journal 46.7 (2008), pp. 1803–1813.
5. D. N. Arnold. Numerical problems in general relativity. Numerical Mathematics and Advanced Applications (P. Neittaanmki, T. Tiihonen, and P. Tarvainen, eds.), World Scientific (2000), pp. 3–15.
6. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. SIAM Journal on Matrix Analysis and Applications 29.1 (2007), pp. 328–347.
7. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magnetic Resonance in Medicine 56.2 (2006), pp. 411–421.
8. E. Begelfor and M. Werman. Affine invariance revisited. Conference on Computer Vision and Pattern Recognition. IEEE. 2006, pp. 2087–2094.
9. R. Bhatia. The Riemannian mean of positive matrices. Matrix Information Geometry. Springer, 2013, pp. 35–51.
10. S. M. Carroll. Spacetime and Geometry: An Introduction to General Relativity. San Francisco, CA, USA: Addison Wesley, 2004.
11. E. Celledoni, M. Eslitzbichler, and A. Schmeding. Shape Analysis on Lie Groups with Applications in Computer Animation. arXiv preprint arXiv:1506.00783 (2015).
12. E. Celledoni and A. Iserles. Approximating the exponential from a Lie algebra to a Lie group. Math. Comp. 69.232 (2000), pp. 1457–1480.
13. J.-M. Chang, C. Peterson, M. Kirby, et al. Feature patch illumination spaces and Karcher compression for face recognition via Grassmannians. Advances in Pure Mathematics 2.04 (2012), p. 226.
14. F Demoures et al. Discrete variational Lie group formulation of geometrically exact beam dynamics. Numerische Mathematik 130.1 (2015), pp. 73-123.
15. I. L. Dryden and K. V. Mardia. Statistical Shape Analysis: With Applications in R. Wiley, 2016.
16. T. Duchamp, G. Xie, and T. P.-Y. Yu. Single basepoint subdivision schemes for manifold-valued data: time-symmetry without space-symmetry. Foundations of Com- putational Mathematics 13.5 (2013), pp. 693–728. doi:10.1007/s10208-013-9144-1.
17. A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. SIAM Journal on Matrix Analysis and Applications 20.2 (1998), pp. 303–353.
18. P. T. Fletcher, C. Lu, and S. Joshi. Statistics of shape via principal geodesic analysis on Lie groups. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 1. IEEE. 2003, pp. 1–7.
19. K. A. Gallivan, A. Srivastava, X. Liu, and P. Van Dooren. Efficient algorithms for inferences on grassmann manifolds. 2003 IEEE Workshop on Statistical Signal Processing. IEEE. 2003, pp. 315–318.
20. F. de Goes, B. Liu, M. Budninskiy, Y. Tong, and M. Desbrun. Discrete 2-Tensor Fields on Triangulations. Computer Graphics Forum. Vol. 33. 5.Wiley Online Library. 2014, pp. 13–24.
21. P. Grohs. Quasi-interpolation in Riemannian manifolds. IMA Journal of Numerical Analysis 33.3 (2013), pp. 849–874.
22. P. Grohs, H. Hardering, and O. Sander. Optimal a priori discretization error bounds for geodesic finite elements. Foundations of Computational Mathematics 15.6 (2015), pp. 1357–1411.
23. P. Grohs, M. Sprecher, and T. Yu. Scattered manifold-valued data approximation. Numerische Mathematik 135.4 (2017), pp. 987–1010. doi:10.1007/s00211-016-0823-0.
24. J. Hall and M. Leok. Lie group spectral variational integrators. Foundations of Computational Mathematics, pp. 1–59.
25. S. Helgason. Differential Geometry, Lie Groups, and Symmetric Spaces. Vol. 80. Academic press, 1979.
26. N. J. Higham. Functions of Matrices: Theory and Computation. SIAM, 2008.

27. N. J. Higham. J-orthogonal matrices: Properties and generation. SIAM review 45.3 (2003), pp. 504–519.
28. N. J. Higham, C. Mehl, and F. Tisseur. The canonical generalized polar decomposition. SIAM Journal on Matrix Analysis and Applications 31.4 (2010), pp. 2163–2180.
29. J. Hilgert and K.-H. Neeb. Structure and Geometry of Lie Groups. Springer, 2011.
30. Y. Hong et al. Geodesic regression on the Grassmannian. Computer Vision-ECCV 2014. Springer, 2014, pp. 632–646.
31. A. Iserles and A. Zanna. Efficient computation of the matrix exponential by generalized polar decompositions. SIAM Journal on Numerical Analysis 42.5 (2005), pp. 2218–2256.
32. T. Jiang et al. Frame field generation through metric customization. ACM Transactions on Graphics (TOG) 34.4 (2015), p. 40.
33. H. Karcher. Riemannian center of mass and mollifier smoothing. Communications on Pure and Applied Mathematics 30.5 (1977), pp. 509–541.
34. A. W. Knapp. Lie Groups: Beyond an Introduction. Vol. 140. Springer, 2013.
35. S. Kobayashi and K. Nomizu. Foundations of Differential Geometry. Vol. 2.Wiley New York, 1969.
36. A. Marthinsen. Interpolation in Lie groups. SIAM Journal on Numerical Analysis 37.1 (1999), pp. 269–285.
37. R. Mathias. A chain rule for matrix functions and applications. SIAM Journal on Matrix Analysis and Applications 17.3 (1996), pp. 610–620.
38. M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM Journal on Matrix Analysis and Applications 26.3 (2005), pp. 735–747.
39. A. Mota, W. Sun, J. T. Ostien, J. W. Foulk, and K. N. Long. Lie-group interpolation and variational recovery for internal variables. Computational Mechanics 52.6 (2013), pp. 1281–1299.
40. H. Z. Munthe-Kaas, G. R.W. Quispel, and A. Zanna. Symmetric spaces and Lie triple systems in numerical analysis of differential equations. BIT Numerical Mathematics 54.1 (2014), pp. 257–282.
41. H. Z. Munthe-Kaas, G. Quispel, and A. Zanna. Generalized polar decompositions on Lie groups with involutive automorphisms. Foundations of Computational Mathematics 1.3 (2001), pp. 297–324.
42. O. Sander. Geodesic finite elements for Cosserat rods. International Journal for Numerical Methods in Engineering 82.13 (2010), pp. 1645–1670.
43. O. Sander. Geodesic finite elements of higher order. IMA J. Numer. Anal. 36.1 (2016), pp. 238–266.
44. O. Sander. Geodesic finite elements on simplicial grids. International Journal for Numerical Methods in Engineering 92.12 (2012), pp. 999–1025.
45. T. Shingel. Interpolation in special orthogonal groups. IMA Journal of Numerical Analysis (2008).
46. L. N. Trefethen and D. Bau III. Numerical Linear Algebra. Vol. 50. SIAM, 1997.
47. P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa. Statistical computations on Grassmann and Stiefel manifolds for image and video-based recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 33.11 (2011), pp. 2273–2286.
48. F. Vetrano, C. Le Garrec, G. D. Mortchelewicz, and R. Ohayon. Assessment of strategies for interpolating POD based reduced order models and application to aeroelasticity. Journal of Aeroelasticity and Structural Dynamics 2.2 (2012).
49. J. Wallner, E. N. Yazdani, and A. Weinmann. Convergence and smoothness analysis of subdivision rules in Riemannian and symmetric spaces. Advances in Computational Mathematics 34.2 (2011), pp. 201–218.
50. E. N. Yazdani and T. P.-Y. Yu. On Donoho's Log-Exp subdivision scheme: choice of retraction and time-symmetry. Multiscale Modeling and Simulation 9.4 (2011), pp. 1801–1828.